



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
MESTRADO ACADÊMICO EM MATEMÁTICA

ERNANDES FERREIRA DO NASCIMENTO NETO

DESIGUALDADE DE RUELLE

FORTALEZA

2023

ERNANDES FERREIRA DO NASCIMENTO NETO

DESIGUALDADE DE RUELLE

Dissertação apresentada ao Curso de Mestrado Acadêmico em Matemática do Programa de Pós-Graduação em Matemática do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Matemática. Área de Concentração: Sistemas Dinâmicos.

Orientador: Prof. Dr. Yuri Gomes Lima.

Coorientador: Prof. Dr. Mauricio José Poletti Merlo.

FORTALEZA

2023

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

N194d Nascimento Neto, Ernandes Ferreira do.

Desigualdade de Ruelle / Ernandes Ferreira do Nascimento Neto. – 2023.
79 f. : il.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Matemática, Fortaleza, 2023.

Orientação: Prof. Dr. Yuri Lima Gomes.

Coorientação: Prof. Dr. Mauricio José Poletti Merlo.

1. Entropia. 2. Expoentes de Lyapunov. 3. Desigualdade de Ruelle. I. Título.

CDD 510

ERNANDES FERREIRA DO NASCIMENTO NETO

DESIGUALDADE DE RUELLE

Dissertação apresentada ao Curso de Mestrado Acadêmico em Matemática do Programa de Pós-Graduação em Matemática do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Matemática. Área de Concentração: Sistemas Dinâmicos.

Aprovada em: 06/07/2023.

BANCA EXAMINADORA

Prof. Dr. Yuri Gomes Lima. (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Mauricio José Poletti
Merlo. (Coorientador)
Universidade Federal do Ceará (UFC)

Prof^ª. Dr^ª. Aline de Melo Machado.
Universidade Federal do Ceará (UFC)

Para minha mãe, Maria da Conceição, pois
sem ela, jamais teria sonhado em estar aqui.

AGRADECIMENTOS

Agradeço, acima de tudo, a Deus por ter me mantido firme e paciente ao longo de toda a minha carreira acadêmica, o que possibilitou o sucesso na realização deste trabalho.

Expresso minha gratidão à minha família, especialmente à minha mãe, Maria da Conceição, por ter sido um apoio emocional constante em minha vida.

Gostaria de agradecer aos meus orientadores, Prof. Dr. Yuri Gomes Lima e Prof. Dr. Maurício José Poletti Merlo, por estarem sempre disponíveis para me auxiliar neste trabalho, fornecendo materiais de estudo, conselhos, amizade, sugestões e ensinamentos que foram fundamentais para o êxito desta pesquisa.

Também desejo expressar minha gratidão ao meu colega, Paulo Henryque, que me convenceu a ingressar na pós-graduação em Matemática após dois anos sem estudar a disciplina.

Gostaria de agradecer aos amigos e colegas da Pós-Graduação em Matemática da UFC, Alisson Conceição, Allen Ibiapina, Breno Cavalcante, Caio Adler, Johnatan Costa, Jônatas Aquino, Jônatas Marinho, Junior Bessa, Karen Silva, Idalina Maria, Matheus Andrade, Matheus Mendes, Michael Silva, Samuel Martins, Samuel Netto e Vinicius Prado, que proporcionaram muitos momentos de reflexão e alegria ao longo de todo o curso.

Não posso deixar de mencionar todos os professores do Departamento de Matemática da UFC no campus do PICI, que contribuíram para o meu crescimento acadêmico e pessoal.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

"Não importa. Tente novamente. Erre novamente. Erre melhor."

(Samuel Beckett)

RESUMO

Neste trabalho, apresentaremos uma introdução à entropia métrica de Kolmogorov-Sinai, que é um invariante de transformações que preservam uma probabilidade. Também discutiremos os Expoentes de Lyapunov como uma ferramenta para descrever o comportamento assintótico de uma transformação. O principal resultado desse trabalho é a Desigualdade de Ruelle, que estabelece uma relação entre os dois conceitos acima, em uma variedade Riemanniana compacta, através de uma desigualdade. Por fim, destacamos a importância da compacidade nesse contexto, mostrando um exemplo devido a Riquelme em que a desigualdade não é satisfeita.

Palavras-chave: entropia; expoentes de Lyapunov; desigualdade de Ruelle.

ABSTRACT

In this work, we will present an introduction to the Kolmogorov-Sinai metric entropy, which is an invariant of transformations that preserve a probability. We will also discuss the Lyapunov Exponents as a tool to describe the asymptotic behavior of a transformation. The main result of this work is the Ruelle Inequality, which establishes a relationship between the two concepts above, on a compact Riemannian manifold, through an inequality. Finally, we highlight the importance of compactness in this context, showing an example due to Riquelme where the inequality is not satisfied.

Keywords: entropy; Lyapunov exponents; Ruelle inequality.

SUMÁRIO

1	INTRODUÇÃO	9
2	INTRODUÇÃO À TEORIA DA ENTROPIA MÉTRICA . .	11
2.1	Entropia de uma partição	11
2.2	Entropia de uma Transformação que Preserva Medida	15
2.3	Teorema de Kolmogorov-Sinai	23
2.4	Entropia Local	29
3	TEOREMA DE OSELEDETS	32
3.1	Cociclos Lineares	32
3.2	Expoentes de Lyapunov Extremais	35
3.3	Teorema de Oseledets	39
4	DESIGUALDADE DE RUELLE	52
4.1	Desigualdade de Ruelle	52
5	UM CONTRA-EXEMPLO PARA A DESIGUALDADE DE RUELLE NO CASO NÃO-COMPACTO	61
5.1	Construção da variedade Riemanniana não-compacta	61
5.2	O difeomorfismo suave ϕ	69
5.3	Contra-exemplo para a Desigualdade de Ruelle	72
	REFERÊNCIAS	79

1 INTRODUÇÃO

A desigualdade de Ruelle, ou desigualdade de Margulis-Ruelle, é um resultado que relaciona dois importantes conceitos da Teoria Ergódica para sistemas diferenciáveis. Esses conceitos são a entropia métrica, definida por Andrei Kolmogorov e Yakov Sinai, e os expoentes de Lyapunov, introduzidos por Aleksandr M. Lyapunov. Ambos são amplamente estudados na área de sistemas dinâmicos.

No segundo capítulo, iremos introduzir a definição de entropia métrica de uma partição finita ou enumerável. As ideias aqui são motivadas pela teoria da informação. Ao longo do capítulo, iremos definir a entropia de um sistema dinâmico mensurável com base na definição de uma partição. Também serão estudadas algumas propriedades dessas partições. Apresentaremos, ainda, maneiras de calcular a entropia de alguns sistemas dinâmicos, bem como provar o Teorema de Kolmogorov-Sinai, o que motiva a definição de partição geradora. Por fim, enunciaremos o Teorema de Brin-Katok, uma outra alternativa para o cálculo da entropia.

No terceiro capítulo, iremos definir o que seria um cociclo linear. A partir daí, faremos alguns exemplos, sendo o mais relevante para este trabalho o cociclo derivada. Apresentaremos também o Teorema Subaditivo de Kingman, com o qual será possível provar o Teorema Ergódico de Birkhoff. Vamos enunciar o Teorema de Oseledets, o que nos permitirá dar a definição de expoente de Lyapunov. Por fim, faremos uma demonstração detalhada do Teorema de Oseledets para o caso de dimensão dois. A prova segue de perto a de [11].

No quarto capítulo, faremos uma prova da desigualdade de Ruelle para difeomorfismos de classe C^1 em uma variedade Riemanniana compacta, que será realizada através de uma série de lemas. Um desses lemas envolve a construção de uma partição que é essencial ao longo da prova. A compacidade da variedade é de extrema importância para a finitude dos elementos da partição. Faremos primeiro o caso em que o sistema é ergódico e, depois, usaremos a decomposição ergódica para tratar o caso de medidas de probabilidade invariantes em geral.

No quinto capítulo, veremos que a compacidade da variedade na desigualdade de Ruelle não pode ser removida. Neste capítulo, será construída uma variedade não compacta, a partir de uma transformação de intercâmbio de intervalos com entropia positiva, usando uma construção do tipo fluxo de suspensão. Após isso, será definido o

fluxo de suspensão, o que permitirá a definição do difeomorfismo procurado e da medida de probabilidade invariante, para a qual haverá entropia métrica positiva e expoentes de Lyapunov zero, caracterizando um contraexemplo da desigualdade de Ruelle em variedades não compactas.

2 INTRODUÇÃO À TEORIA DA ENTROPIA MÉTRICA

A questão se dois sistemas de Bernoulli são ergodicamente equivalentes era um grande problema nas décadas iniciais do desenvolvimento formal da teoria ergódica. Esse problema foi respondido negativamente por Kolmogorov e Sinai, que introduziu para este propósito um novo invariante: a *entropia métrica*. Mais tarde, Ornstein provou que este invariante é completo dentro da classe dos sistemas de Bernoulli: dois sistemas de Bernoulli são ergodicamente equivalentes se e somente se suas entropias métricas forem iguais.

Para motivar a definição de entropia métrica dada por Kolmogorov-Sinai, vamos considerar a situação básica da Teoria da Informação. Consideremos um canal de comunicação que transmite, sucessivamente, certos símbolos. Esse canal pode ser um télégrafo transmitindo pontos e traços, segundo o código Morse, uma fibra ótica, transmitindo zeros e uns, segundo o código binário ASCII, ou qualquer outro sistema de transmissão sequencial de informação. O objetivo é medir a *entropia* do canal, ou seja, a quantidade de informação transmitida, em média, a cada unidade de tempo. A seguir daremos algumas definições e resultados sobre a entropia métrica, que podem ser encontrados nas referências [7, 10, 12].

2.1 Entropia de uma partição

Seja $(M, \mathcal{F}, \mu)$ um espaço de probabilidade. Entendemos por *partição* de M uma família enumerável (finita ou infinita) $\mathcal{A} = \{A \in \mathcal{F}\}$ tal que $\mu(A \cap B) = 0$ para todos $A, B \in \mathcal{A}$ distintos e $\mu(\cup_{A \in \mathcal{A}} A) = 1$. Dadas duas partições $\mathcal{A}$ e $\mathcal{B}$ dizemos que $\mathcal{A} = \mathcal{B} \bmod 0$ se para cada $A \in \mathcal{A}$ com medida positiva podemos encontrar um elemento $B \in \mathcal{B}$ com medida positiva tal que $\mu(A \Delta B) = 0$.

Denotamos por $\mathcal{A}(x)$ o elemento da partição $\mathcal{A}$ que contém um dado ponto x . Dadas duas partições $\mathcal{A}$ e $\mathcal{B}$ de M , a soma das partições $\mathcal{A}$ e $\mathcal{B}$ é a partição

$$\mathcal{A} \vee \mathcal{B} = \{A \cap B : A \in \mathcal{A} \text{ e } B \in \mathcal{B}\}. \quad (2.1)$$

A cada partição $\mathcal{A}$ de M associamos a respectiva *função informação*

$$I_{\mathcal{A}} : M \rightarrow \mathbb{R}, I_{\mathcal{A}}(x) = -\log \mu(\mathcal{A}(x)). \quad (2.2)$$

Temos a seguinte definição:

Definição 2.1.1 (Entropia da partição) A entropia de uma partição $\mathcal{A}$ é dada por

$$H_\mu(\mathcal{A}) = \int I_{\mathcal{A}} d\mu = \sum_{A \in \mathcal{A}} -\mu(A) \log \mu(A). \quad (2.3)$$

Como é usual na teoria, façamos a seguinte convenção: $0 \log 0 = 0$. Note que se $\varphi: (0, \infty) \rightarrow \mathbb{R}$ é dada por $\varphi(x) = -x \log x$, então derivando duas vezes tem-se $\varphi'' < 0$. Portanto, φ é côncava, e daí pela Desigualdade de Jensen temos

$$t_1 \varphi(x_1) + \cdots + t_n \varphi(x_n) \leq \varphi(t_1 x_1 + \cdots + t_n x_n) \quad (2.4)$$

para todos $x_1, \dots, x_n \geq 0$ e $t_1, \dots, t_n \geq 0$ com $t_1 + \cdots + t_n = 1$. A igualdade ocorre se e somente se $x_1 = \cdots = x_n$.

Dizemos também que duas partições $\mathcal{A}$ e $\mathcal{B}$ são *independentes* se $\mu(A \cap B) = \mu(A)\mu(B)$ para todos $A \in \mathcal{A}, B \in \mathcal{B}$. Nesse caso, $I_{\mathcal{A} \vee \mathcal{B}} = I_{\mathcal{A}} + I_{\mathcal{B}}$, donde $H_\mu(\mathcal{A} \vee \mathcal{B}) = H_\mu(\mathcal{A}) + H_\mu(\mathcal{B})$. Em geral, para partições quaisquer, essa igualdade é uma desigualdade, como será visto mais abaixo.

Exemplo 2.1.2 Considere $M = [0, 1]$ munido com a σ -de Borel e considere a medida de Lebesgue. Para cada natural $n \geq 1$ considere a partição $\mathcal{A}^n$ em subintervalos $\left(\frac{i-1}{2^n}, \frac{i}{2^n}\right]$, onde $i = 1, \dots, 2^n$. Então

$$H_\mu(\mathcal{A}^n) = \sum_{i=1}^{2^n} -\frac{1}{2^n} \log 2^{-n} = n \log 2.$$

Exemplo 2.1.3 Seja $M = \{0, \dots, d-1\}^{\mathbb{N}}$ munido com a medida produto $\mu = \nu^{\mathbb{N}}$, onde ν é um vetor de probabilidade em $\{0, \dots, d-1\}$. Denotemos $p_i = \nu(\{i\})$ para cada $i \in \{0, \dots, d-1\}$. Para cada $n \geq 1$, seja $\mathcal{A}^n$ a partição de M em cilindros $[0; a_1, \dots, a_n]$ de comprimento n que começam na posição 0. Mas antes, note que

$$H_\mu(\mathcal{A}) = \sum_{A \in \mathcal{A}} \mu(A) \log \mu(A) = \sum_{i=0}^{d-1} -p_i \log p_i.$$

Agora, a entropia de $\mathcal{A}^n$ é

$$\begin{aligned} H_\mu(\mathcal{A}^n) &= \sum_{a_1, \dots, a_n} -p_{a_1} \cdots p_{a_n} \log(p_{a_1} \cdots p_{a_n}) \\ &= \sum_{j=1}^n \sum_{a_1, \dots, a_n} -p_{a_1} \cdots p_{a_j} \cdots p_{a_n} \log p_{a_j} \\ &= \sum_j \sum_{a_j} -p_{a_j} \log p_{a_j} \sum_{a_i, i \neq j} p_{a_1} \cdots p_{a_{j-1}} p_{a_{j+1}} \cdots p_{a_n} \end{aligned}$$

Uma vez que $\sum_{a_i, i \neq j} p_{a_1} \cdots p_{a_{j-1}} p_{a_{j+1}} \cdots p_{a_n} = 1$ para cada i , segue que

$$H_\mu(\mathcal{A}^n) = \sum_{j=1}^n \sum_{a_j=0}^{d-1} -p_{a_j} \log p_{a_j} = \sum_{j=1}^n \sum_{i=0}^{d-1} -p_i \log p_i = -n \sum_{i=0}^{d-1} p_i \log p_i = n H_\mu(\mathcal{A}).$$

Proposição 2.1.4 *Toda partição finita tem entropia finita. De fato, $H_\mu(\mathcal{A}) \leq \log \#\mathcal{A}$ e vale a igualdade se e somente se $\mu(A) = \frac{1}{\#\mathcal{A}}$ para todo $A \in \mathcal{A}$.*

Demonstração: Seja $\mathcal{A} = \{A_1, A_2, \dots, A_n\}$ e considere os números $t_i = \frac{1}{n}$ e $x_i = \mu(A_i)$.

Usando a desigualdade (2.4):

$$\frac{1}{n} H_\mu(\mathcal{A}) = \sum_{i=1}^n t_i \varphi(x_i) \leq \varphi\left(\sum_{i=1}^n t_i x_i\right) = \varphi\left(\frac{1}{n}\right) = \frac{\log n}{n}.$$

Portanto, $H_\mu(\mathcal{A}) \leq \log n$. Ademais, vale a igualdade se, e somente se, $\mu(A_i) = \frac{1}{n}$ para todo $i = 1, \dots, n$.

□

Ao longo desse trabalho, considera-se apenas partições com **entropia finita**.

Definição 2.1.5 (Entropia condicional) *Sejam $\mathcal{A}$ e $\mathcal{B}$ partições. A entropia condicional de $\mathcal{A}$ com relação a $\mathcal{B}$ é definida por*

$$H_\mu(\mathcal{A}|\mathcal{B}) = \sum_{A \in \mathcal{A}} \sum_{B \in \mathcal{B}} -\mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(B)}. \quad (2.5)$$

Intuitivamente, a equação (2.5) mede a informação dada pela partição $\mathcal{A}$ uma vez que se conhece a informação da partição $\mathcal{B}$. Veja que $H_\mu(\mathcal{A}|\mathcal{M}) = H_\mu(\mathcal{A})$ para toda partição $\mathcal{A}$, onde $\mathcal{M} = \{M\}$. Além disso, quando as partições $\mathcal{A}$ e $\mathcal{B}$ são independentes, tem-se $H_\mu(\mathcal{A}|\mathcal{B}) = H_\mu(\mathcal{A})$. De fato,

$$\begin{aligned} H_\mu(\mathcal{A}|\mathcal{B}) &= \sum_{A \in \mathcal{A}} \sum_{B \in \mathcal{B}} -\mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(B)} = \sum_{A \in \mathcal{A}} \sum_{B \in \mathcal{B}} -\mu(A) \mu(B) \log \frac{\mu(A) \mu(B)}{\mu(B)} \\ &= \sum_{A \in \mathcal{A}} -\mu(A) \log \mu(A) \sum_{B \in \mathcal{B}} \mu(B) = \sum_{A \in \mathcal{A}} -\mu(A) \log \mu(A) = H_\mu(\mathcal{A}). \end{aligned}$$

Definição 2.1.6 (Refinamento de Partições) *Sejam $\mathcal{A}$ e $\mathcal{B}$ partições, dizemos que a partição $\mathcal{A}$ é menos fina que a partição $\mathcal{B}$, e escreve-se $\mathcal{A} \prec \mathcal{B}$, se para todo $B \in \mathcal{B}$ tem-se $B \subset A$, para algum $A \in \mathcal{A}$, a menos de medida nula.*

Note que, a soma $\mathcal{A} \vee \mathcal{B}$ é menos fina que toda partição $\mathcal{C}$ tal que $\mathcal{A} \prec \mathcal{C}$ e $\mathcal{B} \prec \mathcal{C}$. De fato, dada uma partição $\mathcal{C}$ tal que $\mathcal{A} \prec \mathcal{C}$ e $\mathcal{B} \prec \mathcal{C}$, para todo $C \in \mathcal{C}$ existe $A \in \mathcal{A}$ e existe $B \in \mathcal{B}$ tal que $C \subset A$ e $C \subset B$ daí $C \subset A \cap B$, a menos de medida nula. Assim, $\mathcal{A} \vee \mathcal{B} \prec \mathcal{C}$.

Lema 2.1.7 *Sejam $\mathcal{A}, \mathcal{B}$ e $\mathcal{C}$ partições com entropia finita. Então:*

- a) $H_\mu(\mathcal{A} \vee \mathcal{B} | \mathcal{C}) = H_\mu(\mathcal{A} | \mathcal{C}) + H_\mu(\mathcal{B} | \mathcal{A} \vee \mathcal{C})$.
- b) Se $\mathcal{A} \prec \mathcal{B}$, então $H_\mu(\mathcal{A} | \mathcal{C}) \leq H_\mu(\mathcal{B} | \mathcal{C})$ e $H_\mu(\mathcal{C} | \mathcal{B}) \leq H_\mu(\mathcal{C} | \mathcal{A})$. Em particular, se $\mathcal{C} = \{M\}$, tem-se $H_\mu(\mathcal{A}) \leq H_\mu(\mathcal{B})$.
- c) $H_\mu(\mathcal{C} | \mathcal{B}) \leq H_\mu(\mathcal{C})$.
- d) $H_\mu(\mathcal{A} \vee \mathcal{B}) \leq H_\mu(\mathcal{A}) + H_\mu(\mathcal{B})$.
- e) $\mathcal{A} \prec \mathcal{B}$ se, e somente se, $H_\mu(\mathcal{A} | \mathcal{B}) = 0$.

Demonstração: Usando a definição 2.1.5 no item a), tem-se

$$\begin{aligned}
 H_\mu(\mathcal{A} \vee \mathcal{B} | \mathcal{C}) &= \sum_{A, B, C} -\mu(A \cap B \cap C) \log \frac{\mu(A \cap B \cap C)}{\mu(C)} \\
 &= \sum_{A, B, C} -\mu(A \cap B \cap C) \log \frac{\mu(A \cap B \cap C)}{\mu(A \cap C)} + \sum_{A, B, C} -\mu(A \cap B \cap C) \log \frac{\mu(A \cap C)}{\mu(C)} \\
 &= \sum_{A, B, C} -\mu((A \cap C) \cap B) \log \frac{\mu((A \cap C) \cap B)}{\mu(A \cap C)} + \sum_{A, C} -\mu(A \cap C) \log \frac{\mu(A \cap C)}{\mu(C)} \\
 &= H_\mu(\mathcal{B} | \mathcal{A} \vee \mathcal{C}) + H_\mu(\mathcal{A} | \mathcal{C}),
 \end{aligned}$$

como queríamos. Façamos a primeira parte do item b). Uma vez que $\mathcal{A} \prec \mathcal{B}$, usando o item a), temos

$$H_\mu(\mathcal{B} | \mathcal{C}) = H_\mu(\mathcal{A} \vee \mathcal{B} | \mathcal{C}) = H_\mu(\mathcal{A} | \mathcal{C}) + H_\mu(\mathcal{B} | \mathcal{A} \vee \mathcal{C}) \geq H_\mu(\mathcal{A} | \mathcal{C}).$$

Agora façamos segunda parte do item b). Veja que para qualquer $A \in \mathcal{A}$ e $C \in \mathcal{C}$, tem-se

$$\frac{\mu(C \cap A)}{\mu(A)} = \sum_{B \subset A} \frac{\mu(B)}{\mu(A)} \cdot \frac{\mu(B \cap C)}{\mu(B)}.$$

Uma vez que $\sum_{B \subset A} \frac{\mu(B)}{\mu(A)} = 1$, usando (2.4), tem-se

$$\varphi\left(\frac{\mu(C \cap A)}{\mu(A)}\right) = \varphi\left(\sum_{B \subset A} \frac{\mu(B)}{\mu(A)} \cdot \frac{\mu(B \cap C)}{\mu(B)}\right) \geq \sum_{B \subset A} \frac{\mu(B)}{\mu(A)} \varphi\left(\frac{\mu(B \cap C)}{\mu(B)}\right),$$

para todo $A \in \mathcal{A}$ e para todo $C \in \mathcal{C}$. Por conseguinte,

$$\begin{aligned}
 H_\mu(\mathcal{C} | \mathcal{A}) &= \sum_{A, C} \mu(A) \varphi\left(\frac{\mu(C \cap A)}{\mu(A)}\right) \geq \sum_{A, C} \mu(A) \sum_{B \subset A} \frac{\mu(B)}{\mu(A)} \varphi\left(\frac{\mu(B \cap C)}{\mu(B)}\right) \\
 &= \sum_{A, C} \sum_{B \subset A} -\mu(B \cap C) \log \frac{\mu(B \cap C)}{\mu(B)} = H_\mu(\mathcal{C} | \mathcal{B}).
 \end{aligned}$$

Para o item c), tome $\mathcal{A} = \{M\}$ na segunda parte do item b), daí $H_\mu(\mathcal{C}|\mathcal{B}) \leq H_\mu(\mathcal{C}|\mathcal{A}) = H_\mu(\mathcal{C})$.

Para o item d), fazendo $\mathcal{C} = \{M\}$ no item a) e usando o item c), tem-se

$$H_\mu(\mathcal{A} \vee \mathcal{B}) = H_\mu(\mathcal{A}) + H_\mu(\mathcal{B}|\mathcal{A}) \leq H_\mu(\mathcal{A}) + H_\mu(\mathcal{B}).$$

Para o item e), usando a definição 2.1.5, $H_\mu(\mathcal{A}|\mathcal{B}) = 0$ se, e somente se, para todos $A \in \mathcal{A}$ e $B \in \mathcal{B}$, tem-se

$$\mu(A \cap B) = 0 \text{ ou } \frac{\mu(A \cap B)}{\mu(B)} = 1.$$

Ou seja, ou B é disjunto de A (a menos de medida nula) ou B está contido em A (a menos de medida nula). Em outras palavras, $H_\mu(\mathcal{A}|\mathcal{B}) = 0$ se, e somente se, $\mathcal{A} \prec \mathcal{B}$.

□

Dados M, N espaços mensuráveis, seja $f : M \rightarrow N$ uma transformação mensurável, e considere μ uma probabilidade em M . Assim, $f_*\mu$ é uma probabilidade em N tal que para todo mensurável $B \in N$ tem-se $f_*\mu(B) = \mu(f^{-1}(B))$. Ademais, se $\mathcal{A}$ é uma partição em N , então $f^{-1}(\mathcal{A}) = \{f^{-1}(A) : A \in \mathcal{A}\}$ é uma partição em M . Por definição,

$$\begin{aligned} H_\mu(f^{-1}(\mathcal{A})) &= \sum_{A \in \mathcal{A}} -\mu(f^{-1}(A)) \log \mu(f^{-1}(A)) \\ &= \sum_{A \in \mathcal{A}} f_*\mu(A) \log f_*\mu(A) = H_{f_*\mu}(\mathcal{A}). \end{aligned} \quad (2.6)$$

Em particular, quando $M = N$ e μ for uma medida invariante por f , tem-se

$$H_\mu(f^{-1}(\mathcal{A})) = H_\mu(\mathcal{A}), \text{ para toda partição } \mathcal{A} \text{ em } M. \quad (2.7)$$

2.2 Entropia de uma Transformação que Preserva Medida

Seja $f : M \rightarrow M$ uma transformação mensurável que preserva a medida de probabilidade μ . Dada uma partição $\mathcal{A}$ de M com entropia finita, para cada $n \geq 1$, denotamos

$$\mathcal{A}^n = \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) = \{A_0 \cap f^{-1}(A_1) \cap \cdots \cap f^{-n+1}(A_{n-1}) : A_0, \dots, A_{n-1} \in \mathcal{A}\} \text{ para cada } n \geq 1.$$

Veja que o elemento $\mathcal{A}^n(x)$ que contém $x \in M$ é:

$$\mathcal{A}^n(x) = \mathcal{A}(x) \cap f^{-1}(\mathcal{A}(f(x))) \cap \cdots \cap f^{-n+1}(\mathcal{A}(f^{n-1}(x))).$$

A sequência $\{\mathcal{A}^n\}_{n \geq 1}$ é não-decrescente, isto é, $\mathcal{A}^n \prec \mathcal{A}^{n+1}$. Segue do item b) do Lema 2.1.7 que a sequência das entropias $\{H_\mu(\mathcal{A}^n)\}_{n \geq 1}$ é não-decrescente.

Definição 2.2.1 (Entropia de um sistema dada uma partição) *Seja $(M, \mathcal{F}, \mu, f)$ um sistema que preserva medida. Dada uma partição $\mathcal{A}$ de M , definimos a entropia de f com relação à partição $\mathcal{A}$ como sendo*

$$h_\mu(f, \mathcal{A}) = \lim_n \frac{1}{n} H_\mu(\mathcal{A}^n) = \lim_n \frac{1}{n} H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right). \quad (2.8)$$

Vamos mostrar que (2.8) está bem definido, isto é, que o limite existe. Para isto, precisaremos da seguinte proposição:

Proposição 2.2.2 (Fekete) *Se $\{a_n\}_{n \geq 1}$ é uma sequência de números reais tal que $a_{m+n} \leq a_m + a_n$ para todos $m, n \geq 1$, então $\lim_n \frac{a_n}{n}$ existe e é igual a $\inf_n \frac{a_n}{n}$.*

Demonstração: Fixe $p \geq 1$. Para cada $n \geq 1$ podemos escrever $n = kp + i$, onde $0 \leq i < p$, então

$$\frac{a_n}{n} = \frac{a_{kp+i}}{kp+i} \leq \frac{a_{kp}}{kp} + \frac{a_i}{kp} \leq \frac{ka_p}{kp} + \frac{a_i}{kp} = \frac{a_p}{p} + \frac{a_i}{kp}$$

Fazendo $n \rightarrow \infty$ temos $k \rightarrow \infty$, daí

$$\limsup_n \frac{a_n}{n} \leq \frac{a_p}{p} \implies \limsup_n \frac{a_n}{n} \leq \inf_{p \geq 1} \frac{a_p}{p}.$$

Por outro lado, sempre temos

$$\inf_{p \geq 1} \frac{a_p}{p} \leq \liminf_n \frac{a_n}{n}.$$

Assim, $\lim_n \frac{a_n}{n}$ existe e é igual a $\inf_n \frac{a_n}{n}$.

□

Lema 2.2.3 *Seja $(M, \mathcal{F}, \mu, f)$ um sistema que preserva medida. Então $H_\mu(\mathcal{A}^{m+n}) \leq H_\mu(\mathcal{A}^m) + H_\mu(\mathcal{A}^n)$ para todos $m, n \geq 1$.*

Demonstração: Por Definição, tem-se $\mathcal{A}^{m+n} = \bigvee_{i=0}^{m+n-1} f^{-i}(\mathcal{A}) = \mathcal{A}^m \vee f^{-m}(\mathcal{A}^n)$. Usando o item d) do Lema 2.1.7, tem-se

$$H_\mu(\mathcal{A}^{m+n}) = H_\mu(\mathcal{A}^m \vee f^{-m}(\mathcal{A}^n)) \leq H_\mu(\mathcal{A}^m) + H_\mu(f^{-m}(\mathcal{A}^n)). \quad (2.9)$$

Por outro lado, como μ é uma medida invariante por f , usando (2.7), $H_\mu(f^{-m}(\mathcal{A}^n)) = H_\mu(\mathcal{A}^n)$. Usando este fato em (2.9), obtemos o resultado do lema.

□

Desse modo, combinando a Proposição 2.2.2 com o Lema 2.2.3, obtemos que o limite em (2.8) existe e vale

$$\lim_n \frac{1}{n} H_\mu(\mathcal{A}^n) = \lim_n \frac{1}{n} H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right) = \inf_n \frac{1}{n} H_\mu(\mathcal{A}^n). \quad (2.10)$$

Observe que esta entropia é tanto maior quanto mais fina for a partição. Com efeito, se $\mathcal{A} \prec \mathcal{B}$, então $\mathcal{A}^n \prec \mathcal{B}^n$, para todo n . Pelo item b) do Lema 2.1.7, tem-se $H_\mu(\mathcal{A}^n) \leq H_\mu(\mathcal{B}^n)$ para todo n . Desse modo,

$$\mathcal{A} \prec \mathcal{B} \implies h_\mu(\mathcal{A}) \leq h_\mu(\mathcal{B}). \quad (2.11)$$

Definição 2.2.4 (Entropia de uma transformação) *Seja $(M, \mathcal{F}, \mu, f)$ um sistema que preserva medida num espaço de probabilidade. A entropia do sistema (f, μ) é dada por*

$$h_\mu(f) = \sup \{ h_\mu(f, \mathcal{A}) : \mathcal{A} \text{ é uma partição com entropia finita} \}. \quad (2.12)$$

Exemplo 2.2.5 *Considere a transformação binária $f : [0, 1] \rightarrow [0, 1]$, dado por $f(x) = 2x \bmod 1$. Sabe-se que f preserva μ , a medida de Lebesgue no intervalo. Considere a partição $\mathcal{A} = \left(0, \frac{1}{2}\right) \cup \left(\frac{1}{2}, 1\right)$. Assim para todo $n \geq 1$,*

$$\begin{aligned} f^{-1}(\mathcal{A}) &= \left(0, \frac{1}{4}\right) \cup \left(\frac{1}{4}, \frac{1}{2}\right) \cup \left(\frac{1}{2}, \frac{3}{4}\right) \cup \left(\frac{3}{4}, 1\right) \\ &\vdots \\ f^{-n+1}(\mathcal{A}) &= \left(0, \frac{1}{2^n}\right) \cup \dots \cup \left(\frac{2^n-1}{2^n}, 1\right). \end{aligned}$$

Assim, $\mathcal{A}^n = \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) = \left\{ \left(\frac{i-1}{2^n}, \frac{i}{2^n}\right) : i = 1 \dots, 2^n \right\}$. Pelo Exemplo 2.1.2, tem-se

$$h_\mu(f, \mathcal{A}) = \lim_n \frac{1}{n} H_\mu(\mathcal{A}^n) = \lim_n \frac{1}{n} \cdot n \log 2 = \log 2.$$

Será visto na próxima seção que $h_\mu(f) = \log 2$, ou seja, a partição $\mathcal{A}$ realiza o supremo em (2.12).

Exemplo 2.2.6 *Cosidere o shift $\sigma : \Sigma \rightarrow \Sigma$, onde $\Sigma = \{0, \dots, d-1\}^{\mathbb{N}}$ munido com a medida de Bernoulli $\mu = \nu^{\mathbb{N}}$. Considere a partição $\mathcal{A}$ em Σ em cilindros do tipo $A_i = [0; i]$, com $i = 0, \dots, d-1$. Assim, um elemento típico de $\mathcal{A}^n$ é da forma*

$$A_{i_0} \cap \sigma^{-1}(A_{i_1}) \cap \dots \cap \sigma^{-(n-1)}(A_{i_{n-1}}) = \{ \{x_n\} \in \Sigma : x_0 = i_0, \dots, x_{n-1} = i_{n-1} \},$$

que possui medida $p_{i_0} \cdots p_{i_{n-1}}$. Usando o Exemplo 2.1.3, segue que

$$h_\mu(\sigma, \mathcal{A}) = \lim_n \frac{1}{n} H_\mu(\mathcal{A}^n) = \lim_n \frac{1}{n} \left(n \sum_{i=0}^{d-1} -p_i \log p_i \right) = \sum_{i=0}^{d-1} -p_i \log p_i.$$

Também veremos na próxima seção que $h_\mu(\sigma) = \sum_{i=0}^{d-1} -p_i \log p_i$, e portanto, a partição $\mathcal{A}$ atinge o supremo em (2.12).

Proposição 2.2.7 *Seja $f : M \rightarrow M$ uma transformação que preserva medida no espaço de probabilidade $(M, \mathcal{F}, \mu)$, e considere $\mathcal{A}$ uma partição de M . Então $\left\{ \frac{1}{n} H_\mu(\mathcal{A}^n) \right\}_{n \geq 1}$ decresce para $h_\mu(f, \mathcal{A})$.*

Demonstração: Vamos mostrar, por indução, que

$$H_\mu(\mathcal{A}^n) = H_\mu(\mathcal{A}) + \sum_{j=1}^{n-1} H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^j f^{-i}(\mathcal{A}) \right. \right). \quad (2.13)$$

Para $n = 1$ o resultado é claro, agora, vamos assumir que seja verdade para $n = p$, e provar que também vale para $n = p + 1$. De fato, usando o item a) do Lema 2.1.7 quando $\mathcal{C} = \{M\}$ e o fato de μ ser invariante por f , tem-se

$$\begin{aligned} H_\mu(\mathcal{A}^{p+1}) &= H_\mu \left(\bigvee_{i=1}^p f^{-i}(\mathcal{A}) \vee \mathcal{A} \right) = H_\mu \left(\bigvee_{i=1}^p f^{-i}(\mathcal{A}) \right) + H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^p f^{-i}(\mathcal{A}) \right. \right) \\ &= H_\mu \left(\bigvee_{i=0}^{p-1} f^{-i}(\mathcal{A}) \right) + H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^p f^{-i}(\mathcal{A}) \right. \right) \quad (\text{usando a hipótese de indução.}) \\ &= H_\mu(\mathcal{A}) + \sum_{j=1}^{p-1} H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^j f^{-i}(\mathcal{A}) \right. \right) + H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^p f^{-i}(\mathcal{A}) \right. \right) \\ &= H_\mu(\mathcal{A}) + \sum_{j=1}^p H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^j f^{-i}(\mathcal{A}) \right. \right) \end{aligned}$$

o que prova a igualdade em (2.13). Desse modo, usando o item b) do Lema 2.1.7 e (2.13) vamos ter $H_\mu(\mathcal{A}^n) \geq n H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^n f^{-i}(\mathcal{A}) \right. \right)$, daí

$$\begin{aligned} n H_\mu \left(\bigvee_{i=0}^n f^{-i}(\mathcal{A}) \right) &= n \left[H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right) + H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^n f^{-i}(\mathcal{A}) \right. \right) \right] \\ &\leq (n+1) H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right), \end{aligned}$$

e portanto

$$\frac{1}{n+1} H_\mu(\mathcal{A}^{n+1}) \leq \frac{1}{n} H_\mu(\mathcal{A}^n).$$

□

Proposição 2.2.8 *Seja $f : M \rightarrow M$ uma transformação que preserva medida no espaço de probabilidade $(M, \mathcal{F}, \mu)$. Para toda partição $\mathcal{A}$ com entropia finita, tem-se*

$$h_\mu(f, \mathcal{A}) = \lim_n H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^n f^{-i}(\mathcal{A}) \right. \right).$$

Demonstração: Usando a Definição em (2.8) e o resultado em (2.13), temos

$$h_\mu(f, \mathcal{A}) = \lim_n \frac{1}{n} H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right) = \lim_n \frac{1}{n} \sum_{j=1}^{n-1} H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^j f^{-i}(\mathcal{A}) \right. \right).$$

Por outro lado, o item *b*) do Lema 2.1.7 garante que a sequência $\left\{ H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^n f^{-i}(\mathcal{A}) \right. \right) \right\}_{n \geq 1}$ é não-crescente. Em particular, $\lim_n H_\mu \left(\mathcal{A} \left| \bigvee_{i=1}^n f^{-i}(\mathcal{A}) \right. \right)$ existe e, consequentemente, coincide com limite Cesàro da igualdade anterior.

□

Teorema 2.2.9 *Seja $f : M \rightarrow M$ uma transformação que preserva medida no espaço de probabilidade $(M, \mathcal{F}, \mu)$. Dadas $\mathcal{A}$ e $\mathcal{B}$ partições com entropia finita, tem-se:*

- a)* $h_\mu(f, \mathcal{A}) \leq H_\mu(\mathcal{A})$.
- b)* $h_\mu(f, \mathcal{A} \vee \mathcal{B}) \leq h_\mu(f, \mathcal{A}) + h_\mu(f, \mathcal{B})$.
- c)* $\mathcal{A} \prec \mathcal{B} \Rightarrow h_\mu(f, \mathcal{A}) \leq h_\mu(f, \mathcal{B})$.
- d)* $h_\mu(f, \mathcal{A}) \leq h_\mu(f, \mathcal{B}) + H_\mu(\mathcal{A}|\mathcal{B})$.
- e)* $h_\mu(f, f^{-1}(\mathcal{A})) = h_\mu(f, \mathcal{A})$.
- f)* Se $k \geq 1$, $h_\mu(f, \mathcal{A}) = h_\mu(f, \mathcal{A}^k)$.
- g)* Se f é invertível e $k \geq 1$, então

$$h_\mu(f, \mathcal{A}) = h_\mu \left(f, \bigvee_{i=-k}^{k-1} f^{-i}(\mathcal{A}) \right).$$

Demonstração:

- a)* Usando o item *d*) do Lema 2.1.7 e (2.7), temos para todo $n \geq 1$:

$$\begin{aligned} \frac{1}{n} H_\mu(\mathcal{A}^n) &\leq \frac{1}{n} \sum_{i=0}^{n-1} H_\mu(f^{-i}(\mathcal{A})) \\ &= \frac{1}{n} \sum_{i=0}^{n-1} H_\mu(\mathcal{A}) \\ &= H_\mu(\mathcal{A}). \end{aligned}$$

Passando o limite sobre n , obtemos o resultado.

b) Usando novamente o item $d)$ do Lema 2.1.7, obtem-se para todo $n \geq 1$ que

$$\begin{aligned} H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A} \vee \mathcal{B}) \right) &= H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \vee \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{B}) \right) \\ &\leq H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right) \vee H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{B}) \right). \end{aligned}$$

Agora basta dividir por n e passar o limite.

c) Se $\mathcal{A} \prec \mathcal{B}$, então para todo $B \in \mathcal{B}$ existe algum $A \in \mathcal{A}$ tal que $B \subset A$, donde $f^{-i}(B) \subset f^{-i}(A)$ para cada $i \geq 0$. Assim:

$$\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \prec \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{B}).$$

Usando o item $b)$ do Lema 2.1.7, obtemos

$$H_\mu(\mathcal{A}^n) = H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right) \leq H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{B}) \right) = H_\mu(\mathcal{B}^n).$$

Agora basta dividir por n e passar o limite.

d) Pelos itens $a)$, $b)$ e $d)$ do Lema 2.1.7, tem-se para todo $n \geq 1$ que

$$\begin{aligned} H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right) &\leq H_\mu \left(\left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right) \vee \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{B}) \right) \right) \\ &= H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{B}) \right) + H_\mu \left(\left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right) \middle| \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{B}) \right) \right) \\ &\leq H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{B}) \right) + \sum_{i=0}^{n-1} H_\mu \left(f^{-i}(\mathcal{A}) \middle| \left(\bigvee_{j=0}^{n-1} f^{-j}(\mathcal{B}) \right) \right) \\ &\leq H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{B}) \right) + \sum_{i=0}^{n-1} H_\mu(f^{-i}(\mathcal{A}) | f^{-i}(\mathcal{B})) \\ &= H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{B}) \right) + n H_\mu(\mathcal{A} | \mathcal{B}), \end{aligned}$$

onde na última igualdade usamos (2.7). Agora basta dividir por n e passar o limite.

e) Usando (2.7) tem-se

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

assim,

$$h_\mu(f, f^{-1}(\mathcal{A})) = \lim_n \frac{1}{n} H \left(\bigvee_{i=1}^n f^{-i}(\mathcal{A}) \right) = \lim_n \frac{1}{n} H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right) = h_\mu(f, \mathcal{A}).$$

f) Usando a definição dada em (2.8) temos

$$\begin{aligned}
 h_\mu \left(f, \bigvee_{i=0}^{k-1} f^{-i}(\mathcal{A}) \right) &= \lim_n \frac{1}{n} H_\mu \left(\bigvee_{j=0}^{n-1} f^{-j} \left(\bigvee_{i=0}^{k-1} f^{-i}(\mathcal{A}) \right) \right) \\
 &= \lim_n \frac{1}{n} H_\mu \left(\bigvee_{i=0}^{n+k-2} f^{-i}(\mathcal{A}) \right) \\
 &= \lim_n \left(\frac{n+k-1}{n} \right) \frac{1}{n+k-1} H_\mu \left(\bigvee_{i=0}^{n+k-2} f^{-i}(\mathcal{A}) \right) \\
 &= h_\mu(f, \mathcal{A}).
 \end{aligned}$$

g) Usando (2.7) recursivamente e o item f) acima, temos

$$\begin{aligned}
 h_\mu \left(f, \bigvee_{i=-k}^{k-1} f^{-i}(\mathcal{A}) \right) &= h_\mu \left(f, \bigvee_{i=0}^{2k-1} f^{-i}(\mathcal{A}) \right) \\
 &= h_\mu(f, \mathcal{A}).
 \end{aligned}$$

□

Teorema 2.2.10 *Seja $(M, \mathcal{F}, \mu, f)$ um sistema que preserva medida em um espaço de probabilidade. Valem as seguintes afirmações.*

- a) Para todo $k > 0$, $h_\mu(f^k) = kh_\mu(f)$.
- b) Se f é invertível, então $h_\mu(f^k) = |k|h_\mu(f)$ para todo $k \in \mathbb{Z}$.

Demonstração:

a) Vamos mostrar primeiro que

$$h_\mu \left(f^k, \bigvee_{i=0}^{k-1} f^{-i}(\mathcal{A}) \right) = kh_\mu(f, \mathcal{A}) \text{ se } k > 0.$$

De fato,

$$\begin{aligned}
 h_\mu \left(f^k, \bigvee_{i=0}^{k-1} f^{-i}(\mathcal{A}) \right) &= \lim_n \frac{1}{n} H_\mu \left(\bigvee_{j=0}^{n-1} f^{-kj} \left(\bigvee_{i=0}^{k-1} f^{-i}(\mathcal{A}) \right) \right) \\
 &= \lim_n \frac{k}{kn} H_\mu \left(\bigvee_{i=0}^{nk-1} f^{-i}(\mathcal{A}) \right) \\
 &= kh_\mu(f, \mathcal{A}).
 \end{aligned}$$

Então,

$$\begin{aligned}
 kh_\mu(f) &= k \cdot \sup_{\mathcal{A}} h_\mu(f, \mathcal{A}) = \sup_{\mathcal{A}} h_\mu \left(f^k, \bigvee_{i=0}^{k-1} f^{-i}(\mathcal{A}) \right) \\
 &\leq \sup_{\mathcal{B}} h_\mu(f^k, \mathcal{B}) = h_\mu(f^k).
 \end{aligned}$$

Por outro lado, pelo item *c*) do Teorema 2.2.9, tem-se

$$h_\mu(f^k, \mathcal{A}) \leq h_\mu \left(f^k, \bigvee_{i=0}^{k-1} f^{-i}(\mathcal{A}) \right) = kh_\mu(f, \mathcal{A}),$$

daí $h_\mu(f^k) \leq kh_\mu(f)$, obtendo a segunda desigualdade.

b) Basta mostrar que $h_\mu(f^{-1}) = h_\mu(f)$ e para tal, basta mostrar que $h_\mu(f^{-1}, \mathcal{A}) = h_\mu(f, \mathcal{A})$ para toda partição com entropia finita. Por (2.7), tem-se

$$\begin{aligned} H_\mu \left(\bigvee_{i=0}^{n-1} f^i(\mathcal{A}) \right) &= H_\mu \left(f^{-(n-1)} \left(\bigvee_{i=0}^{n-1} f^i(\mathcal{A}) \right) \right) \\ &= H_\mu \left(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A}) \right), \end{aligned}$$

para toda partição com entropia finita.

□

Definição 2.2.11 (Equivalência Ergódica) Dizemos que os sistemas $(M, \mathcal{F}_1, f, \mu)$ e $(N, \mathcal{F}_2, g, \nu)$, que preservam medidas de probabilidade μ e ν respectivamente, são **ergodicamente equivalentes** se existem conjuntos mensuráveis $X \subset M$ e $Y \subset N$ com $\mu(M \setminus X) = 0$ e $\nu(N \setminus Y) = 0$ e uma bijeção mensurável $\varphi : X \rightarrow Y$ com inversa mensurável tal que

$$\varphi_*\mu = \nu \quad \text{e} \quad \varphi \circ f = g \circ \varphi.$$

Teorema 2.2.12 Se $(M, \mathcal{F}_1, f, \mu)$ e $(N, \mathcal{F}_2, g, \nu)$ são dois sistemas ergodicamente equivalentes, então $h_\mu(f) = h_\nu(g)$.

Demonstração: Seja $\varphi : X \rightarrow Y$ como na Definição 2.2.11. Suponha, sem perda de generalidade, que $X = M$ e $Y = N$. Considere uma partição $\mathcal{B}$ de N com entropia finita, e seja $\mathcal{A} = \{\varphi^{-1}(B) : B \in \mathcal{B}\}$. Afirmamos que $\mathcal{A}$ é uma partição em M . Como φ é mensurável, temos $\mathcal{A} \subset \mathcal{F}_1$. Ademais, dados $B_1, B_2 \in \mathcal{B}$ temos

$$\varphi^{-1}(B_1) \cap \varphi^{-1}(B_2) = \varphi^{-1}(B_1 \cap B_2) = \varphi^{-1}(\emptyset) = \emptyset,$$

e vale que $M = \bigcup_{B \in \mathcal{B}} \varphi^{-1}(B)$, o que prova afirmação. Agora, para todo $n \in \mathbb{N}$ e para todo $B_i \in \mathcal{B}$, temos que

$$\begin{aligned} D = \bigcap_{j=0}^{n-1} f^{-j}(\varphi^{-1}(B_i)) &= \bigcap_{j=0}^{n-1} (f^{-j} \circ \varphi^{-1})(B_i) \\ &= \bigcap_{j=0}^{n-1} (\varphi^{-1} \circ g^{-j})(B_i) \\ &= \varphi^{-1} \left(\bigcap_{j=0}^{n-1} g^{-j}(B_i) \right), \end{aligned}$$

assim $D \in \varphi^{-1}(\mathcal{B}^n)$. Isso prova que $\mathcal{A}^n = \varphi^{-1}(\mathcal{B}^n)$, logo para cada $n \in \mathbb{N}$ temos

$$H_\mu(\mathcal{A}^n) = \sum_{B \in \mathcal{B}^n} -\mu(\varphi^{-1}(B)) \log \mu(\varphi^{-1}(B)) = \sum_{B \in \mathcal{B}^n} -\nu(B) \log \nu(B) = H_\nu(\mathcal{B}^n).$$

Assim,

$$h_\mu(f, \mathcal{A}) = \lim_n \frac{1}{n} H_\mu(\mathcal{A}^n) = \lim_n \frac{1}{n} H_\nu(\mathcal{B}^n) = h_\nu(g, \mathcal{B}).$$

Dessa forma, para qualquer partição $\mathcal{B}$ de N temos $h_\mu(f) \geq h_\nu(g, \mathcal{B})$, logo $h_\mu(f) \geq h_\nu(g)$.

Por simetria, temos $h_\mu(f) \leq h_\nu(g)$, o que conclui a prova. □

2.3 Teorema de Kolmogorov-Sinai

Em geral, a principal dificuldade no cálculo da entropia está no cálculo do supremo na Definição (2.12). Nessa seção, veremos métodos que vão auxiliar na tarefa do cálculo da entropia de muitos casos, onde poderemos encontrar partições $\mathcal{A}$ que realizam o supremo, ou seja, $h_\mu(f) = h_\mu(f, \mathcal{A})$.

Lema 2.3.1 *Dado $k \geq 1$ e $\varepsilon > 0$, existe $\delta > 0$ tal que, para quaisquer partições finitas $\mathcal{A} = \{A_1, \dots, A_k\}$ e $\mathcal{B} = \{B_1, \dots, B_k\}$,*

$$\mu(A_i \triangle B_i) < \delta \text{ para todo } i = 1, \dots, k \implies H_\mu(\mathcal{B}|\mathcal{A}) < \varepsilon.$$

Demonstração: Fixe $\varepsilon > 0$ e $k \geq 1$. Pela continuidade da função $\varphi : [0, 1] \rightarrow \mathbb{R}$, $\varphi(x) = -x \log x$, existe $\rho > 0$ tal que $\varphi(x) < \frac{\varepsilon}{k^2}$ para todo $x \in [0, \rho] \cup (1 - \rho, 1]$. Tome $\delta = \frac{\rho}{k}$. Dadas as partições $\mathcal{A}$ e $\mathcal{B}$ como no enunciado, denote por $\mathcal{C}$ a partição cujos elementos são

interseções $A_i \cap B_j$ com $i \neq j$ e também o conjunto $\bigcup_{i=1}^k A_i \cap B_i$. Note que $\mu(A_i \cap B_j) \leq \mu(A_i \triangle B_i) < \delta$ para todo $i \neq j$ e

$$\mu\left(\bigcup_{i=1}^k A_i \cap B_i\right) \geq \sum_{i=1}^k (\mu(A_i) - \mu(A_i \triangle B_i)) > \sum_{i=1}^k (\mu(A_i) - \delta) = 1 - \rho.$$

Portanto,

$$H_\mu(\mathcal{C}) = \sum_{C \in \mathcal{C}} \varphi(\mu(C)) < \#\mathcal{C} \frac{\varepsilon}{k^2} \leq \varepsilon.$$

Segue imediatamente de (2.1) que $\mathcal{A} \vee \mathcal{B} = \mathcal{A} \vee \mathcal{C}$. Então pelos itens *a*), *c*) do Lema 2.1.7 segue que

$$\begin{aligned} H_\mu(\mathcal{B}|\mathcal{A}) &= H_\mu(\mathcal{A} \vee \mathcal{B}) - H_\mu(\mathcal{A}) = H_\mu(\mathcal{A} \vee \mathcal{C}) - H_\mu(\mathcal{A}) \\ &= H_\mu(\mathcal{C}|\mathcal{A}) \leq H_\mu(\mathcal{C}) < \varepsilon. \end{aligned}$$

□

Convém observar que, para quaisquer partições $\mathcal{A}$, $\mathcal{B}$ e $\mathcal{C}$ com entropias finitas tem-se

$$H_\mu(\mathcal{A}|\mathcal{C}) \leq H_\mu(\mathcal{A}|\mathcal{B}) + H_\mu(\mathcal{B}|\mathcal{C}). \quad (2.14)$$

De fato, uma vez que $\mathcal{A} \prec \mathcal{A} \vee \mathcal{B}$ e $\mathcal{B} \prec \mathcal{B} \vee \mathcal{C}$, à luz dos itens *a*), *b*) do Lema 2.1.7 segue que

$$\begin{aligned} H_\mu(\mathcal{A}|\mathcal{C}) &\leq H_\mu(\mathcal{A} \vee \mathcal{B}|\mathcal{C}) = H_\mu(\mathcal{B}|\mathcal{C}) + H_\mu(\mathcal{A}|\mathcal{B} \vee \mathcal{C}) \\ &\leq H_\mu(\mathcal{B}|\mathcal{C}) + H_\mu(\mathcal{A}|\mathcal{B}), \end{aligned}$$

o que prova (2.14).

Proposição 2.3.2 *Seja $\mathcal{D}$ uma álgebra que gera a σ -álgebra dos conjuntos mensuráveis de M , a menos de medida nula. Para toda partição $\mathcal{B}$ de M com entropia finita e todo $\varepsilon > 0$, existe alguma partição finita $\mathcal{A} \subset \mathcal{D}$ tal que $H_\mu(\mathcal{B}|\mathcal{A}) < \varepsilon$.*

Demonstração: Vamos reduzir o enunciado para caso em que $\mathcal{B}$ é uma partição finita. Assuma que $\mathcal{B}$ é enumerável com elementos B_j , $j = 1, 2, \dots$. Para cada $k \geq 1$, considere a partição finita

$$\mathcal{B}_k = \left\{ B_1, \dots, B_k, M \setminus \bigcup_{j=1}^k B_j \right\}.$$

Afirmção 2.3.1 *Para toda partição com entropia finita:*

$$\lim_k H_\mu(\mathcal{B}|\mathcal{B}_k) = 0.$$

Prova da Afirmção: Seja $B_0 = M \setminus \bigcup_{j=1}^k B_j$. Por definição,

$$H_\mu(\mathcal{B}|\mathcal{B}_k) = \sum_{i=0}^k \sum_{j \geq 1} -\mu(B_i \cap B_j) \log \frac{\mu(B_i \cap B_j)}{\mu(B_j)}.$$

Todas as parcelas com $i \geq 1$ são nulas, uma vez que $\mu(B_i \cap B_j) = 0$ quando $i \neq j$ e $\mu(B_i \cap B_j) = \mu(B_i)$ quando $i = j$. temos Para $i = 0$, temos nesse caso que $\mu(B_i \cap B_j)$ é igual a zero se $j \leq k$ e igual a $\mu(B_j)$ se $j > k$. Portanto,

$$H_\mu(\mathcal{B}|\mathcal{B}_k) = \sum_{j>k} -\mu(B_j) \log \frac{\mu(B_j)}{\mu(B_0)} \leq \sum_{j>k} -\mu(B_j) \log \mu(B_j).$$

Como $\mathcal{B}$ tem entropia finita, o lado direito da expressão converge a zero quando $k \rightarrow \infty$, o que prova a afirmação.

Agora, dado $\varepsilon > 0$, pela Afirmção 2.3.1, fixe $k \geq 1$ de tal modo que $H_\mu(\mathcal{B}|\mathcal{B}_k) < \frac{\varepsilon}{2}$. Seja $\delta > 0$ qualquer. Para cada $i = 1, \dots, k$ existe $D_i \in \mathcal{D}$ tal que

$$\mu(B_i \triangle D_i) < \frac{\delta}{2k^2}. \quad (2.15)$$

Considere $A_1 = D_1$ e $A_i = D_i \setminus \bigcup_{j=1}^{i-1} D_j$ para $i = 2, \dots, k$ e também $A_0 = M \setminus \bigcup_{j=1}^k D_j$. Veja que $\mathcal{A} = \{A_1, \dots, A_k, A_0\}$ é partição de M e que $A_i \in \mathcal{D}$ para todo i . Para $i = 1, \dots, k$, temos que $A_i \triangle D_i = D_i \setminus A_i = D_i \cap \left(\bigcup_{j=1}^{i-1} D_j \right)$. Dado qualquer x neste conjunto, existe $j < i$ tal que $x \in D_i \cap D_j$. Como $B_i \cap B_j = \emptyset$, segue $x \in (D_i \setminus B_i) \cup (D_j \setminus B_j)$. Assim,

$$A_i \triangle D_i \subset \bigcup_{j=1}^{i-1} (D_j \setminus B_j) \subset \bigcup_{j=1}^{i-1} (D_j \triangle B_j)$$

e, portanto, $\mu(A_i \triangle D_i) < \frac{i\delta}{2k^2} \leq \frac{\delta}{2k}$. Usando (2.15), segue que

$$\mu(A_i \triangle B_i) < \frac{\delta}{2k^2} + \frac{\delta}{2k} \leq \frac{\delta}{k} \text{ para } i = 1, \dots, k. \quad (2.16)$$

Ademais, $A_0 \triangle B_0 \subset \bigcup_{i=0}^k A_i \triangle B_i$ uma vez que $\mathcal{A}$ e $\mathcal{B}_k$ são partições de M . Logo, (2.16) implica

$$\mu(A_0 \triangle B_0) < \delta. \quad (2.17)$$

Note que as partições $\mathcal{B}_k$ e $\mathcal{A}$ satisfazem as hipóteses do Lema 2.3.1. Supondo que $\delta > 0$ é suficientemente pequeno, as relações (2.16) e (2.17) implicam que $H_\mu(\mathcal{B}_k|\mathcal{A}) < \frac{\varepsilon}{2}$. Portanto, pela desigualdade em (2.14), temos

$$H_\mu(\mathcal{B}|\mathcal{A}) \leq H_\mu(\mathcal{B}|\mathcal{B}_k) + H_\mu(\mathcal{B}_k|\mathcal{A}) < \varepsilon,$$

como queríamos. □

Corolário 2.3.3 *Seja $\mathcal{A}_1 \prec \dots \prec \mathcal{A}_n \prec \dots$ uma sequência não-decrescente de partições com entropia finita tais que $\bigcup_{n=1}^{\infty} \mathcal{A}_n$ gera a σ -álgebra dos conjuntos mensuráveis M , a menos de medida nula. Então vale $\lim_n H_\mu(\mathcal{B}|\mathcal{A}_n) = 0$ para toda partição $\mathcal{B}$ com entropia finita.*

Demonstração: Para cada n , seja $\mathcal{D}_n$ a álgebra gerada por $\mathcal{A}_n$. Note que os elementos de $\mathcal{D}_n$ são as uniões finitas de elementos de $\mathcal{A}_n$ juntamente com os complementos dessas uniões. Ademais, $\mathcal{D} = \bigcup_n \mathcal{D}_n$ é a álgebra gerada por $\bigcup_{n=1}^{\infty} \mathcal{A}_n$. Considere qualquer $\varepsilon > 0$. Usando a Proposição 2.3.2, existe uma partição finita $\mathcal{A} \subset \mathcal{D}$ tal que $H_\mu(\mathcal{B}|\mathcal{A}) < \varepsilon$. Logo, uma vez que $\mathcal{A}$ é finita, existe $m \geq 1$ tal que $\mathcal{A} \subset \mathcal{A}_m$ e, portanto, $\mathcal{A}$ é menos fina que $\mathcal{A}_m$. Usando o item b) do Lema 2.1.7,

$$H_\mu(\mathcal{B}|\mathcal{A}_n) \leq H_\mu(\mathcal{B}|\mathcal{A}_m) \leq H_\mu(\mathcal{B}|\mathcal{A}) < \varepsilon \text{ para todo } n \geq m,$$

o que prova o corolário. □

Teorema 2.3.4 (Kolmogorov-Sinai) *Seja $\mathcal{A}_1 \prec \dots \prec \mathcal{A}_n \prec \dots$ uma sequência não-decrescente de partições com entropia finita tais que $\bigcup_{n=1}^{\infty} \mathcal{A}_n$ gera a σ -álgebra dos conjuntos mensuráveis M , a menos de medida nula. Então*

$$h_\mu(f) = \lim_n h_\mu(f, \mathcal{A}_n).$$

Demonstração: O limite sempre existe, pois a propriedade em (2.11) implica que $\{h_\mu(f|\mathcal{A}_n)\}_{n \geq 1}$ é não-decrescente. A definição de entropia em (2.12) nos diz que $h_\mu(f) \geq h_\mu(f, \mathcal{A}_n)$, para todo $n \geq 1$. Assim, basta mostrar que $h_\mu(f, \mathcal{B}) \leq \lim_n h_\mu(f, \mathcal{A}_n)$ para toda partição $\mathcal{B}$ com entropia finita. Pelo item d) do Teorema 2.2.9, temos

$$h_\mu(f, \mathcal{B}) \leq h_\mu(f, \mathcal{A}_n) + H_\mu(\mathcal{B}|\mathcal{A}_n), \text{ para todo } n.$$

Passando o limite e usando o Corolário 2.3.3, obtem-se $h_\mu(f, \mathcal{B}) \leq \lim_n h_\mu(f, \mathcal{A}_n)$ para toda partição $\mathcal{B}$ com entropia finita.

□

Corolário 2.3.5 *Seja $\mathcal{A}$ uma partição com entropia finita tal que a união dos iterados $\mathcal{A}^n = \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{A})$, $n \geq 1$, gera a σ -álgebra dos conjuntos mensuráveis de M , a menos de medida nula. Então $h_\mu(f) = h_\mu(f, \mathcal{A})$,*

Demonstração: Basta aplicar o Teorema 2.3.4 à sequência $\mathcal{A}^n$ e usar o item f) do Teorema 2.2.9 para obter

$$h_\mu(f) = \lim_n h_\mu(f, \mathcal{A}^n) = \lim_n h_\mu(f, \mathcal{A}) = h_\mu(f, \mathcal{A}).$$

□

Veja que, no Exemplo 2.2.6, temos que $\mathcal{A} = \{[0; a_1]; a_1 = 0, \dots, d-1\}$ satisfaz o Corolário 2.3.5.

Corolário 2.3.6 *Suponha que o sistema (f, μ) é invertível. Seja $\mathcal{A}$ uma partição com entropia finita tal que a união dos iterados $\mathcal{A}^{\pm n} = \bigvee_{i=-n}^{n-1} f^{-i}(\mathcal{A})$, $n \geq 1$, gera a σ -álgebra dos conjuntos mensuráveis de M , a menos de medida nula. Então $h_\mu(f) = h_\mu(f, \mathcal{A})$.*

Demonstração: Basta aplicar o Teorema 2.3.4 à sequência $\mathcal{A}^{\pm n}$ e usar o item g) do Teorema 2.2.9 para obter que

$$h_\mu(f) = \lim_n h_\mu(f, \mathcal{A}^{\pm n}) = \lim_n h_\mu(f, \mathcal{A}) = h_\mu(f, \mathcal{A}).$$

□

Definição 2.3.7 (Partições Geradoras) *Seja $f : M \rightarrow M$ uma transformação que preserva a medida μ no espaço de probabilidade $(M, \mathcal{F}, \mu)$. Dada uma partição $\mathcal{A}$ que satisfaz o Corolário 2.3.5 ou o Corolário 2.3.6, dizemos que $\mathcal{A}$ é uma partição geradora unilateral, no caso do Corolário 2.3.5, e bilateral, no caso do Corolário 2.3.6.*

Corolário 2.3.8 *Suponha que o sistema (f, μ) é invertível e existe alguma partição $\mathcal{A}$ com entropia finita tal que $\bigcup_{n=1}^{\infty} \mathcal{A}^n$ gera a σ -álgebra dos conjuntos mensuráveis de M , a menos de medida nula. Então $h_\mu(f) = 0$.*

Demonstração: Usando a Proposição 2.2.8 e o Corolário 2.3.5:

$$h_\mu(f) = h_\mu(f, \mathcal{A}) = \lim_n H_\mu \left(\mathcal{A} \left| f^{-1}(\mathcal{A}^n) \right. \right).$$

Como $\bigcup_n \mathcal{A}^n$ gera a σ -álgebra dos conjuntos mensuráveis de M , $\bigcup_n f^{-1}(\mathcal{A}^n)$ gera a σ -álgebra $f^{-1}(\mathcal{F})$. Mas $f^{-1}(\mathcal{F}) = \mathcal{F}$, pois f é invertível. Logo, pelo Corolário 2.3.3, temos que $H_\mu(\mathcal{A} | f^{-1}(\mathcal{A}^n))$ converge a zero quando $N \rightarrow \infty$. Segue que $h_\mu(f) = h_\mu(f, \mathcal{A}) = 0$. $\square$

Vamos agora supor que M é um espaço métrico, munido com a σ -álgebra de Borel. Dada $\mathcal{A}$ uma partição de M , defini-se $\text{diam } \mathcal{A} = \sup\{\text{diam } A : A \in \mathcal{A}\}$.

Corolário 2.3.9 *Seja $\mathcal{A}_1 \prec \dots \mathcal{A}_n \prec \dots$ uma sequência não-decrescente de partições com entropia finita tais que $\text{diam } \mathcal{A}_n(x) \rightarrow 0$ para μ -quase todo ponto $x \in M$. Então*

$$h_\mu(f) = \lim_n h_\mu(f, \mathcal{A}_n).$$

Demonstração: Pelo Teorema de Kolmogorov-Sinai, é suficiente mostrar que a σ -álgebra gerada por $\bigcup_{j=1}^{\infty} \mathcal{A}_j$ contém os borelianos. Para tal, considere U um aberto, não vazio, qualquer de M . Como $\text{diam } \mathcal{A}_n \rightarrow 0$ quando $n \rightarrow \infty$, para cada $x \in U$ existe $k \geq 1$ tal que $A_x = \mathcal{A}_k(x) \subset U$. Veja que A_x pertence à σ -álgebra gerada por $\bigcup_n \mathcal{A}_n$. Doravante, tal álgebra é enumerável, já que ela é formada pelas uniões finitas de elementos das partições $\mathcal{A}_n$. Em particular, o conjunto dos valores tomado por A_x é enumerável. Segue que $U = \bigcup_{x \in U} A_x$ também está na σ -álgebra gerada por $\bigcup_{j=1}^{\infty} \mathcal{A}_j$. Isto prova que a σ -álgebra gerada por $\bigcup_{j=1}^{\infty} \mathcal{A}_j$ contém todos os abertos e, portanto, contém os conjuntos borelianos. $\square$

Exemplo 2.3.10 *Seja $f : \mathbb{S}^1 \rightarrow \mathbb{S}^1$ um homeomorfismo e seja μ uma medida de probabilidade invariante qualquer. Dada uma partição finita $\mathcal{A}$ de $\mathbb{S}^1$ em subintervalos, denotemos por $x_1, \dots, x_m$ os seus pontos extremos. Para todo $j \geq 1$, a partição $f^{-j}(\mathcal{A})$ está formada pelos subintervalos de $\mathbb{S}^1$ determinados pelos pontos $f^{-j}(x_i)$. Isso implica que, para cada $n \geq 1$, os elementos de $\mathcal{A}^n$ têm seus extremos no conjunto*

$$\{f^{-j}(x_i); j = 0, \dots, n-1 \text{ e } i = 1, \dots, m\}.$$

Em particular, $\#\mathcal{A}^n \leq mn$. Então, pela Proposição 2.1.4,

$$h_\mu(f|\mathcal{A}) = \lim_n \frac{1}{n} H_\mu(\mathcal{A}^n) \leq \lim_n \frac{1}{n} \log \#\mathcal{A}^n \leq \lim_n \frac{1}{n} \log mn = 0.$$

Para ver que $h_\mu(f) = 0$, basta tomar uma sequência qualquer de partições finitas em intervalos com diâmetro indo para zero e aplicar o Corolário 2.3.9.

Corolário 2.3.11 *Seja $\mathcal{A}$ uma partição com entropia finita tal que, para μ -quase todo ponto $x \in M$, tem-se $\text{diam } \mathcal{A}^n(x) \rightarrow 0$. Então $h_\mu(f) = h_\mu(f, \mathcal{A})$.*

Demonstração: Basta aplicar o Corolário 2.3.9 à sequência $\mathcal{A}^n$ e usar o item f) do Teorema 2.2.9. □

Note que, no Exemplo 2.2.5, a partição $\mathcal{A} = \left[0, \frac{1}{2}\right) \cup \left(\frac{1}{2}, 1\right]$ satisfaz as hipóteses do Corolário 2.3.11, logo $h_\mu(f) = \log 2$.

2.4 Entropia Local

Nesta seção, iremos discutir uma visão complementar da entropia, de caráter mais local, esse é o conteúdo do Teorema de Brin-Katok. Para tal, precisaremos da definição de bolas dinâmicas:

Definição 2.4.1 *Considere uma transformação contínua $f : M \rightarrow M$ num espaço métrico compacto. Dados $x \in M$, $n \geq 1$ e $\varepsilon > 0$, a bola dinâmica de tamanho n e raio ε em torno de x é dada*

$$B(x, n, \varepsilon) = \{y \in M : d(f^i(x), f^i(y)) < \varepsilon \text{ para todo } i = 0, \dots, n-1\}.$$

Ou seja,

$$B(x, n, \varepsilon) = B_\varepsilon(x) \cap f^{-1}(B_\varepsilon(f(x))) \dots f^{-n+1}(B_\varepsilon(f^{n-1}(x)))$$

. Agora, dada uma medida de probabilidade μ invariante por f , defina:

$$\begin{aligned} h_\mu^+(f, \varepsilon, x) &= \limsup_n -\frac{1}{n} \log \mu(B(x, n, \varepsilon)) \text{ e} \\ h_\mu^-(f, \varepsilon, x) &= \liminf_n -\frac{1}{n} \log \mu(B(x, n, \varepsilon)). \end{aligned}$$

Temos assim o seguinte resultado:

Teorema 2.4.2 (Brin-Katok) *Seja μ uma medida de probabilidade invariante por f . Os limites*

$$\lim_{\varepsilon \rightarrow 0} h_{\mu}^{+}(f, \varepsilon, x) \text{ e } \lim_{\varepsilon \rightarrow 0} h_{\mu}^{-}(f, \varepsilon, x)$$

existem e são iguais para μ -quase todo $x \in M$. Ademais, denotando por $h_{\mu}(f, x)$ o valor em comum entre eles, a função $h_{\mu}(f, \cdot)$ é μ -integrável e

$$h_{\mu}(f) = \int h_{\mu}(f, x) d\mu.$$

Demonstração: Vide [4].

□

Exemplo 2.4.3 *Seja $f_A : \mathbb{T}^d \rightarrow \mathbb{T}^d$ um endomorfismo linear induzido no toro $\mathbb{T}^2$ por uma matriz invertível A com coeficientes inteiros. Seja μ a medida de Haar no $\mathbb{T}^2$. Então*

$$h_{\mu}(f_A) = \sum_{i=1}^d \log |\lambda_i|,$$

onde $\lambda_1, \dots, \lambda_d$ são os autovalores de A , contando a multiplicidade.

Demonstração: Vamos supor que A é diagonalizável. Seja $v_1, \dots, v_d$ a base normalizada de $\mathbb{R}^d$ de autovetores. Seja u o número de autovalores de A com valor absoluto maior que 1. Dado $x \in \mathbb{T}^d$, todo ponto y em uma vizinhança de x pode ser escrito da forma

$$y = x + \sum_{i=1}^d t_i v_i$$

com $t_1, \dots, t_d$ próximo de 0. Dado $\varepsilon > 0$, denote por $D(x, \varepsilon)$ o conjunto dos pontos y dessa forma como $|t_i| < \varepsilon$ para todo $i = 1, \dots, d$. Ademais, para cada $n \geq 1$, pode-se considerar

$$D(x, n, \varepsilon) = \{y \in \mathbb{T}^d : f_A^j(y) \in D(f^j(x), \varepsilon) \text{ para todo } j = 0, \dots, n-1\}.$$

Notemos que $f_A^j(y) = f_A^j(x) + \sum_{i=1}^d t_i \lambda_i^j v_i$ para todo $n \geq 1$. Desse modo

$$D(x, n, \varepsilon) = \left\{ x + \sum_{i=1}^d t_i \lambda_i v_i : |\lambda_i^n t_i| < \varepsilon \text{ para } i \leq u \text{ e } |t_i| < \varepsilon \text{ para } i > u \right\}.$$

Dessa maneira, existe uma constante $K_1 > 0$ que depende apenas de A tal que

$$K_1^{-1} \varepsilon^d \prod_{i=1}^u |\lambda_i|^{-n} \leq \mu(D(x, n, \varepsilon)) \leq K_1 \varepsilon^d \prod_{i=1}^u |\lambda_i|^{-n}$$

para todos $x \in \mathbb{T}^d$, $n \geq 1$ e $\varepsilon > 0$. Existe uma outra constante $K_2 > 0$ que depende apenas de A de modo que

$$B_{K_2^{-1}\varepsilon}(x) \subset D(x, \varepsilon) \subset B_{K_2\varepsilon}(x)$$

para todo $x \in \mathbb{T}^d$ e $\varepsilon > 0$ pequeno. Então, $B(x, n, K_2^{-1}\varepsilon) \subset D(x, n, \varepsilon) \subset B(x, n, K_2\varepsilon)$ para todo $n \geq 1$. Tomando $K = K_1 K_2$ e combinando os resultados, no obtemos

$$K^{-1}\varepsilon^d \prod_{i=1}^u |\lambda_i|^{-n} \leq \mu(D(x, n, \varepsilon)) \leq K\varepsilon^d \prod_{i=1}^u |\lambda_i|^{-n}$$

para todo $x \in \mathbb{T}^d$, $n \geq 1$ e $\varepsilon > 0$. Assim,

$$h_\mu^+(f_A, \varepsilon, x) = h_\mu^-(f_A, \varepsilon, x) = \lim_n \frac{1}{n} \log \mu(B(x, n, \varepsilon)) = \sum_{i=1}^u \log |\lambda_i|$$

para todo $x \in \mathbb{T}^d$ e $\varepsilon > 0$ pequeno. Portanto, usando o Teorema de Brin-Katok

$$h_\mu(f_A) = h_\mu(f_A, x) = \sum_{i=1}^u \log |\lambda_i|$$

para μ -quase todo ponto x .

O caso geral é tratado de maneira análogo usando a forma canônica de Jordan.

□

3 TEOREMA DE OSELEDTS

Neste capítulo vamos introduzir o segundo personagem da desigualdade de Ruelle, que são os expoentes de Lyapunov. Essa ferramenta é definida a partir do Teorema de Oseledets, os expoentes de Lyapunov ajudar a compreender melhor o comportamento assintótico de certas transformações. As definições e resultados encontrados neste capítulo podem ser encontrados nas referências [7] e [11].

3.1 Cociclos Lineares

Vamos introduzir o objeto básico de estudo, que são os cociclos lineares, bem como algumas definições que nos acompanharão ao longo deste trabalho.

Definição 3.1.1 (Norma e co-norma) *Dada $A \in GL(d, \mathbb{R})$, define-se a norma de A por*

$$\|A\| = \max \left\{ \frac{\|Av\|}{\|v\|} : v \neq 0 \right\} = \max \{ \|Av\| : \|v\| \leq 1 \}, \quad (3.1)$$

e a co-norma de A por

$$m(A) = \min \left\{ \frac{\|Av\|}{\|v\|} : v \neq 0 \right\}. \quad (3.2)$$

Dada duas matrizes $A, B \in GL(d, \mathbb{R})$, se $v \in \mathbb{R}^d$ então

$$\|(AB)(v)\| = \|A(Bv)\| \leq \|A\| \cdot \|Bv\| \leq \|A\| \cdot \|B\| \cdot \|v\|,$$

daí, tomando o supremo,

$$\|AB\| \leq \|A\| \cdot \|B\|.$$

Afirmamos que $m(A) = \|A^{-1}\|^{-1}$. Para ver isso, a definição (3.1) dá que

$$\|A^{-1}\| \geq \frac{\|A^{-1}v\|}{\|v\|}$$

para $v \neq 0$. Tomando $w \in \mathbb{R}^d$ tal que $Aw = v$ (que existe pois A é invertível), temos

$$\|A^{-1}\| \geq \frac{\|w\|}{\|Aw\|} \implies \|A^{-1}\|^{-1} \leq \frac{\|Aw\|}{\|w\|},$$

e daí pela definição em (3.2) segue que $\|A^{-1}\|^{-1} \leq m(A)$. Por outro lado, temos para todo $v \neq 0$ que

$$\|Av\| \geq m(A) \cdot \|v\| \implies 1 \geq m(A) \cdot \frac{\|v\|}{\|Av\|} \implies 1 \geq m(A) \cdot \frac{\|A^{-1}w\|}{\|w\|}.$$

Essa última desigualdade vale para todo $w \neq 0$, logo tomando o máximo em $w \neq 0$ obtemos que

$$1 \geq m(A) \cdot \|A^{-1}\| \implies m(A) \leq \|A^{-1}\|^{-1}.$$

Obtemos assim que $m(A) = \|A^{-1}\|^{-1}$, conforme afirmado.

Definição 3.1.2 (Skew product) *Dado um sistema dinâmico $(M, \mathcal{F}, \mu, f)$, um skew product sobre f é um sistema dinâmico F agindo no espaço produto $M \times N$ de modo que o seguinte diagrama comuta*

$$\begin{array}{ccc} M \times N & \xrightarrow{F} & M \times N \\ \pi \downarrow & & \downarrow \pi \\ M & \xrightarrow{f} & M \end{array}$$

isto é, $\pi \circ F = f \circ \pi$, onde $\pi : M \times N \rightarrow M$ é a projeção na primeira coordenada.

Estamos interessados em skew products que agem linearmente na segunda coordenada, logo N precisa ser um espaço vetorial. Essas aplicações serão chamadas de cociclos lineares.

Definição 3.1.3 (Cociclo linear) *Dado um sistema dinâmico $(M, \mathcal{F}, \mu, f)$ e uma aplicação mensurável $A : M \rightarrow GL(d, \mathbb{R})$, definimos o $GL(d, \mathbb{R})$ -cociclo linear sobre $(M, \mathcal{F}, \mu, f)$ como sendo a transformação*

$$\begin{aligned} F : M \times \mathbb{R}^d &\longrightarrow M \times \mathbb{R}^d \\ (x, v) &\longmapsto (f(x), A(x) \cdot v) \end{aligned}$$

Como o par (f, A) define $F = F_{f,A}$, também o chamaremos de cociclo. Às vezes, nós também chamaremos a aplicação A de cociclo. Note que para todo $n \geq 1$ tem-se $F^n(x, v) = (f^n(x), A^n(x) \cdot v)$, onde

$$A^n(x) = A(f^{n-1}(x)) \cdots A(f(x))A(x). \quad (3.3)$$

Quando f for invertível, F também será. Além disso, para todo $n \geq 1$ tem-se $F^{-n}(x, v) = (f^{-n}(x), A^{-n}(x) \cdot v)$, onde

$$A^{-n}(x) = A(f^{-n}(x)) \cdots A(f^{-1}(x)). \quad (3.4)$$

Veja que de (3.3) e (3.4) tem-se para todo $m, n \in \mathbb{Z}$ e $x \in M$ que

$$A^{m+n}(x) = A^n(f^m(x))A^m(x). \quad (3.5)$$

Vejam os alguns exemplos.

Exemplo 3.1.4 *Um problema muito estudado em probabilidade é o produto aleatório de matrizes. Para tal, considere um conjunto de matrizes com uma lei probabilidade. Escolhendo ao acaso matrizes desse conjunto, fazemos o produto delas. Em sistemas dinâmicos, tem-se o interesse em saber o que ocorrerá com esses produtos matriciais. Isso pode ser modelado do seguinte modo. Seja $M = GL(d, \mathbb{R})^{\mathbb{Z}}$, e considere o shift bilateral*

$$\begin{aligned} f: M &\longrightarrow M \\ (\alpha_k)_k &\longmapsto (\alpha_{k+1})_k. \end{aligned}$$

Considere a aplicação

$$\begin{aligned} A: M &\longrightarrow GL(d, \mathbb{R}) \\ (\alpha_k)_k &\longmapsto \alpha_0, \end{aligned}$$

e seja $F: M \times \mathbb{R}^d \rightarrow M \times \mathbb{R}^d$ o cociclo linear definido por A sobre f . Veja que a k -ésima iteração de F é dada por:

$$F^k(x, v) = ((\alpha_{n+k})_n, (\alpha_{n-1} \cdots \alpha_1 \alpha_0) \cdot v).$$

Dada uma probabilidade ν no espaço $GL(d, \mathbb{R})$, podemos considerar a medida produto $\mu = \nu^{\mathbb{Z}}$, que é caracterizada por

$$\mu(\{(\alpha_k)_k; \alpha_i \in E_i, \dots, \alpha_j \in E_j\}) = \nu(E_i) \cdots \nu(E_j)$$

para cada $i \leq j$ e quaisquer mensuráveis $E_i, \dots, E_j \subset GL(d, \mathbb{R})$. É claro que a medida μ é invariante pelo shift.

Exemplo 3.1.5 *Vamos chamar de cociclo linear localmente constante à seguinte construção mais geral. Seja $(Y, \mathcal{Y}, \eta)$ um espaço de probabilidade, e considere $N = Y^{\mathbb{Z}}$ dotada da σ -álgebra produto $\mathcal{G} = \mathcal{Y}^{\mathbb{Z}}$ com a medida produto $\vartheta = \eta^{\mathbb{Z}}$. Considere o shift $g: N \rightarrow N$, e seja $B: N \rightarrow GL(d, \mathbb{R})$ uma função mensurável qualquer, dependendo apenas da coordenada zero, ou seja, $B(y) = \beta(y_0)$, onde $\beta: Y \rightarrow GL(d, \mathbb{R})$ é mensurável. Defina o cociclo linear*

$G : N \times \mathbb{R}^d \rightarrow N \times \mathbb{R}^d$ definido por B sobre g . Note que G é semi-conjugado ao cociclo F do exemplo anterior, onde $\nu = \beta_*\eta$, como no diagrama abaixo

$$\begin{array}{ccc} N \times \mathbb{R}^d & \xrightarrow{G} & N \times \mathbb{R}^d \\ \psi \downarrow & & \downarrow \psi \\ M \times \mathbb{R}^d & \xrightarrow{F} & M \times \mathbb{R}^d \end{array}$$

onde $\psi((x_k)_k, v) = ((\beta(x_k))_k, v)$.

Exemplo 3.1.6 (Cociclo derivada) Considere um difeomorfismo $f : M \rightarrow M$ no toro $M = \mathbb{T}^d$ de dimensão $d \geq 1$, que preserva a orientação. Sabe-se que existem campos vetoriais $X_1, \dots, X_d$ em $\mathbb{T}^d$ de forma que $\{X_1(p), \dots, X_d(p)\}$ é uma base para o espaço tangente $T_p M$, para cada $p \in M$, tais campos existem uma vez que o toro é uma variedade paralelizável. Seja μ uma medida invariante por f . O cociclo derivada de f é

$$\begin{aligned} F : M \times \mathbb{R}^d &\longrightarrow M \times \mathbb{R}^d \\ (p, v) &\longmapsto (f(p), A(p) \cdot v), \end{aligned}$$

onde $A(p) \in GL(d, \mathbb{R})$ é matriz, com respeito às bases dadas pelos campos vetoriais, da derivada $df(p) : T_p M \rightarrow T_{f(p)} M$. Veja que estudar a dinâmica do cociclo F é estudar a dinâmica da derivada $df(p)$. A regra da cadeia nos permite obter (3.5).

3.2 Expoentes de Lyapunov Extremais

Vamos agora definir o maior expoente de Lyapunov, que fornece uma taxa exponencial de crescimento da norma de cociclos.

Definição 3.2.1 (Família subaditiva) Considere $(M, \mathcal{F}, \mu)$ um espaço de probabilidade e $f : M \rightarrow M$ uma aplicação que preserva a medida μ . Dizemos que uma família de funções mensuráveis $\varphi_n : M \rightarrow [-\infty, +\infty)$, $n \geq 1$, é subaditiva com respeito a f se

$$\varphi_{m+n} \leq \varphi_m + \varphi_n \circ f^m, \text{ para todos } m, n \geq 1. \quad (3.6)$$

Exemplo 3.2.2 Para toda função mensurável $\psi : M \rightarrow \mathbb{R}$, considere a soma de Birkhoff

$$\psi_n = \sum_{j=0}^{n-1} \psi \circ f^j. \quad (3.7)$$

Note que

$$\psi_{m+n} = \sum_{j=0}^{m+n-1} \psi \circ f^j = \sum_{j=0}^{m-1} \psi \circ f^j + \sum_{j=0}^{n-1} \psi \circ f^{j+m} = \psi_m + \psi_n \circ f^m.$$

Em particular, a sequência $(\psi_n)_n$ é subaditiva.

Exemplo 3.2.3 Para toda função mensurável $A: M \rightarrow GL(d, \mathbb{R})$, considere a sequência $\varphi_n(x) = \log \|A^n(x)\|$, onde $A^n(x)$ é dado em (3.3) e (3.4). Para todo $x \in M$, temos:

$$\begin{aligned} \varphi_{m+n}(x) &= \log \|A^{m+n}(x)\| = \log \|A^n(f^m(x))A^m(x)\| \\ &\leq \log(\|A^n(f^m(x))\| \|A^m(x)\|) = \log \|A^n(f^m(x))\| + \log \|A^m(x)\| \\ &= \varphi_n \circ f^m(x) + \varphi_m(x). \end{aligned}$$

Logo, $(\varphi_n)_n$ é subaditiva.

Teorema 3.2.4 (Subaditivo de Kingman) Seja μ uma probabilidade invariante por $f: M \rightarrow M$ e seja $\varphi_n: M \rightarrow \mathbb{R}$, $n \geq 1$, uma sequência subaditiva de funções mensuráveis tal que $\varphi_1^+ = \max\{\varphi_1, 0\} \in L^1(\mu)$. Então a sequência $\left(\frac{\varphi_n}{n}\right)_n$ converge em μ -quase todo ponto para uma função f -invariante $\varphi: M \rightarrow \mathbb{R}$. Além disso, $\varphi^+ \in L^1(\mu)$ e

$$\int \varphi \, d\mu = \lim_n \frac{1}{n} \int \varphi_n \, d\mu = \inf_n \frac{1}{n} \int \varphi_n \, d\mu \in [-\infty, +\infty).$$

Demonstração: Vide [7, Teorema 3.3.3] ou [11, Teorema 3.3].

□

Corolário 3.2.5 (Teorema Ergódico de Birkhoff) Seja μ uma probabilidade invariante por $f: M \rightarrow M$ e seja $\varphi \in L^1(\mu)$. Então

$$\bar{\varphi}(x) = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \varphi \circ f^j(x) \tag{3.8}$$

existe em μ -quase todo ponto. Ademais, a função $\bar{\varphi}$ é f -invariante e μ -integrável com

$$\int \bar{\varphi} \, d\mu = \int \varphi \, d\mu.$$

Demonstração: Segue do Exemplo (3.2.2) que $(\varphi_n)_n = \left(\sum_{j=0}^{n-1} \varphi \circ f^j\right)_n$ é subaditiva, e portanto o Teorema 3.2.4 diz que (3.8) converge para μ -quase todo ponto, e o limite é

μ -integrável e f -invariante. Além disso, como μ é invariante por f , tem-se

$$\begin{aligned}\int \bar{\varphi} \, d\mu &= \lim_n \frac{1}{n} \int \sum_{j=0}^{n-1} \varphi \circ f^j \, d\mu = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \int \varphi \circ f^j \, d\mu \\ &= \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \int \varphi \, d\mu = \int \varphi \, d\mu.\end{aligned}$$

□

Corolário 3.2.6 *Seja μ uma probabilidade invariante por $f : M \rightarrow M$ e seja $\varphi : M \rightarrow \mathbb{R}$, uma função tal que $\psi = \varphi \circ f - \varphi$ é μ -integrável. Então*

$$\lim_n \frac{1}{n} \varphi(f^n(x)) = 0$$

para μ -quase todo $x \in M$. Em particular, a afirmação vale quando $\varphi \in L^1(\mu)$.

Demonstração: Aplicando o Corolário 3.2.5 à função integrável $\psi = \varphi \circ f - \varphi$, temos que a sequência $\left(\frac{1}{n} \varphi \circ f^n\right)$ converge para alguma função mensurável $\bar{\varphi}$. De fato,

$$\begin{aligned}\bar{\varphi}(x) &= \lim_n \frac{1}{n} \sum_{j=0}^{n-1} (\varphi \circ f - \varphi) \circ f^j(x) = \lim_n \frac{1}{n} \left(\sum_{j=0}^{n-1} (\varphi \circ f^{j+1})(x) - \sum_{j=0}^{n-1} (\varphi \circ f^j)(x) \right) \\ &= \lim_n \frac{1}{n} (\varphi \circ f^n(x) - \varphi(x)) \\ &= \lim_n \frac{1}{n} \varphi \circ f^n(x).\end{aligned}$$

Para $k \geq 1$, defina $A_k = \{x \in M : |\varphi(x)| \leq k\}$. Fixe k . Pelo Teorema de Recorrência de Poincaré, para μ -quase todo ponto $x \in A_k$ existem infinitos $n_j \in \mathbb{N}$ tais que $f^{n_j}(x) \in A_k$, ou seja, $|\varphi \circ f^{n_j}(x)| \leq k$. Logo

$$|\bar{\varphi}(x)| = \lim_{n_j} \left| \frac{1}{n_j} \varphi \circ f^{n_j}(x) \right| \leq \lim_{n_j} \frac{k}{n_j} = 0,$$

o que implica que $\bar{\varphi}(x) = 0$ para μ -quase todo ponto em A_k . Como $M = \bigcup_{k \geq 1} A_k$, segue o resultado.

□

Vamos agora enunciar e demonstrar o teorema que vai nos permitir definir os expoentes extremais de Lyapunov.

Teorema 3.2.7 (Furstenberg-Kesten) *Sejam $(M, \mathcal{F}, \mu)$ um espaço de probabilidade, $f : M \rightarrow M$ uma transformação mensurável que preserva a medida μ , e $A : M \rightarrow GL(d, \mathbb{R})$ uma aplicação mensurável tal que $\log^{\pm 1} \|A\| = \max\{\log \|A^{\pm 1}\|, 0\} \in L^1(\mu)$. Então*

$$\lambda_{\max}(x) = \lim_n \frac{1}{n} \log \|A^n(x)\| \quad \text{e} \quad \lambda_{\min}(x) = \lim_n \frac{1}{n} \log \|(A^n(x))^{-1}\|^{-1} \quad (3.9)$$

existem para μ -quase todo ponto $x \in M$. Além disso, as funções $\lambda_{\max}$ e $\lambda_{\min}$ são f -invariantes e μ -integráveis com

$$\int \lambda_{\max}(x) \, d\mu(x) = \lim_n \frac{1}{n} \int \log \|A^n(x)\| \, d\mu(x)$$

e

$$\int \lambda_{\min}(x) \, d\mu(x) = \lim_n \frac{1}{n} \int \log \|(A^n(x))^{-1}\|^{-1} \, d\mu(x).$$

Demonstração: Defina as seguintes funções

$$\varphi_n(x) = \log \|A^n(x)\| \quad \text{e} \quad \psi_n(x) = \log \|(A^n(x))^{-1}\|.$$

Desse modo,

$$\varphi_1(x) = \log \|A(x)\| \quad \text{e} \quad \psi_1(x) = \log \|(A(x))^{-1}\|.$$

Segue da hipótese que φ_1^+ e ψ_1^+ são integráveis em relação à medida μ , consequentemente $\varphi_1^+(x), \psi_1^+(x) \in [-\infty, +\infty)$ para μ -quase todo ponto $x \in M$. Finalmente, tendo em vista o Exemplo 3.2.3, as sequências φ_n e ψ_n são subaditivas, daí, usando o Teorema 3.2.4, vale que

$$\lambda_{\max}(x) = \lim_n \frac{1}{n} \log \|A^n(x)\| \quad \text{e} \quad \lambda_{\min}(x) = \lim_n \frac{1}{n} \log \|(A^n(x))^{-1}\|^{-1}$$

existem para μ -quase todo ponto $x \in M$ e satisfazem as demais condições elencadas no enunciado do teorema.

□

Definição 3.2.8 *As funções $\lambda_{\max}$ e $\lambda_{\min}$ são chamadas de **expoentes de Lyapunov extremais**.*

3.3 Teorema de Oseledets

Agora, iremos enunciar o Teorema Multiplicativo de Oseledets que generaliza o conceito de expoentes de Lyapunov. Daremos uma prova desse teorema em um caso particular, em dimensão baixa. No que segue, vamos considerar $(M, \mathcal{F}, \mu)$ um espaço de probabilidade, $f : M \rightarrow M$ uma transformação mensurável que preserva a medida μ , $A : M \rightarrow \text{GL}(d, \mathbb{R})$ uma aplicação mensurável e $F : M \times \mathbb{R}^d \rightarrow M \times \mathbb{R}^d$ o cociclo $F(x, v) = (f(x), A(x) \cdot v)$.

Definição 3.3.1 (Filtração) *Uma filtração em $\mathbb{R}^d$ é uma família não-decrescente $W^1 \supsetneq \dots \supsetneq W^k \supsetneq \{0\}$ de subespaços vetoriais no espaço euclidiano $\mathbb{R}^d$, com $k \leq d$. A filtração é completa quando $k = d$ e $\dim W^j = d + 1 - j$ para todo $j = 1, \dots, d$.*

Teorema 3.3.2 (Oseledets - caso não-invertível) *Se $\log^+ \|A^{\pm 1}\| \in L^1(\mu)$, então para μ -quase todo $x \in M$, existe um número inteiro $k = k(x)$, números reais $\lambda_1(x) > \dots > \lambda_k(x)$ e uma filtração $\mathbb{R}^d = V_x^1 \supsetneq \dots \supsetneq V_x^k \supsetneq \{0\}$ tal que para todo $i \in \{1, \dots, k\}$:*

- a) $k(f(x)) = k(x)$, $\lambda_i(f(x)) = \lambda_i(x)$ e $A(x) \cdot V_x^i = V_{f(x)}^i$;
- b) as aplicações $x \mapsto k(x)$, $x \mapsto \lambda_i(x)$ e $x \mapsto V_x^i$ são mensuráveis;
- c) $\lim_n \frac{1}{n} \log \|A^n(x)v\| = \lambda_i(x)$ para todo $v \in V_x^i \setminus V_x^{i+1}$.

Demonstração: Consulte [11, Teorema 4.1].

□

Definição 3.3.3 (Expoentes de Lyapunov) *Os números λ_i que satisfazem as condições do Teorema 3.3.2 são chamados expoentes de Lyapunov. Eles satisfazem $\lambda_{\max} = \lambda_1$ e $\lambda_{\min} = \lambda_k$.*

Quando μ é ergódica, segue do item b) do Teorema 3.3.2 que os valores de $k(x)$ e cada um dos expoentes de Lyapunov $\lambda_i(x)$ são constantes em um conjunto de medida total. Vamos chamar de $d_i(x) = \dim V_x^i - \dim V_x^{i+1}$ a multiplicidade do expoente de Lyapunov correspondente $\lambda_i(x)$. O espectro de Lyapunov de F é o conjunto de todos os expoentes de Lyapunov, cada um contado com multiplicidade. O espectro de Lyapunov é simples se todos os expoentes de Lyapunov tiverem multiplicidade 1 ou, de forma equivalente, se a filtração for completa.

Quando f for invertível, o Teorema de Oseledets admite um enunciado mais preciso.

Teorema 3.3.4 (Oseledets - caso invertível) *Se $f : M \rightarrow M$ for invertível e $\log^+ \|A^{\pm 1}\| \in L^1(\mu)$, então para μ -quase todo $x \in M$ existe uma decomposição $\mathbb{R}^d = E_x^1 \oplus \cdots \oplus E_x^k$ tal que para todo $i = 1, \dots, k$:*

- a) $A(x) \cdot E_x^i = E_{f(x)}^i$ e $V_x^i = \bigoplus_{j=1}^i E_x^j$;
- b) $\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \|A^n(x)v\| = \lambda_i(x)$ para todo $v \in E_x^i$ diferente de zero;
- c) $\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \left| \sin \angle \left(\bigoplus_{i \in I} E_{f^n(x)}^i, \bigoplus_{j \in J} E_{f^n(x)}^j \right) \right| = 0$ sempre que $I \cap J = \emptyset$.

Acima, o ângulo $\angle(V, W)$ entre os subespaços $V, W \subset \mathbb{R}^d$ é o menor ângulo entre dois vetores não-nulos $v \in V$ e $w \in W$.

Demonstração: Vide [11, Teorema 4.2].

□

No restante desse capítulo, vamos enunciar e provar o Teorema de Oseledets quando $d = 2$.

Teorema 3.3.5 (Oseledets em dimensão dois - caso não-invertível) *Se $\log^+ \|A^{\pm 1}\| \in L^1(\mu)$, então para μ -quase todo ponto $x \in M$ uma das condições abaixo vale:*

- 1) $\lambda_{\max}(x) = \lambda_{\min}(x)$ e $\lim_n \frac{1}{n} \log \|A^n(x)v\| = \lambda_{\max}(x)$, para todo $v \in \mathbb{R}^2 \setminus \{0\}$.
- 2) $\lambda_{\min}(x) < \lambda_{\max}(x)$ e existe uma reta $E_x^s \subset \mathbb{R}^2$ tal que

$$\lim_n \frac{1}{n} \log \|A^n(x)v\| = \begin{cases} \lambda_{\min}(x), & \text{se } v \in E_x^s \setminus \{0\} \\ \lambda_{\max}(x), & \text{se } v \in \mathbb{R}^2 \setminus E_x^s. \end{cases} \quad (3.10)$$

Ademais, o espaço E_x^s depende mensuravelmente do ponto x e é invariante pelo cociclo, ou seja, $A(x) \cdot E_x^s = E_{f(x)}^s$ para μ -quase todo ponto $x \in M$.

Demonstração: Assumimos inicialmente que $A : M \rightarrow \text{SL}(2, \mathbb{R})$ (o caso geral será deduzido a partir deste). Considere $x \in M$ como na conclusão do Teorema 3.2.7. Como $d = 2$, tem-se $\|A^n(x)\| = \|(A^n(x))^{-1}\|$. Daí,

$$\lambda_{\max}(x) + \lambda_{\min}(x) = \lim_n \frac{1}{n} \left(\log \|A^n(x)\| + \log \|A^n(x)\|^{-1} \right) = 0.$$

Assim podemos considerar $\lambda(x) = \lambda_{\max}(x) = -\lambda_{\min}(x)$. Primeiro, seja $x \in M$ de modo que $\lambda(x) = 0$. Para todo $n \in \mathbb{N}$ e todo $v \in \mathbb{R}^2 \setminus \{0\}$ temos

$$\|A^n(x)\|^{-1} \|v\| \leq \|A^n(x)v\| \leq \|A^n(x)\| \|v\|.$$

Tomando logaritmo na desigualdade e dividindo por n obtemos que

$$\frac{1}{n} \log(\|A^n(x)\|^{-1} \|v\|) \leq \frac{1}{n} \log \|A^n(x)v\| \leq \frac{1}{n} \log(\|A^n(x)\| \|v\|)$$

e daí, passando o limite, segue que

$$\lim_n \frac{1}{n} \log \|A^n(x)\|^{-1} \leq \lim_n \frac{1}{n} \log \|A^n(x)v\| \leq \lim_n \frac{1}{n} \log \|A^n(x)\|,$$

ou seja:

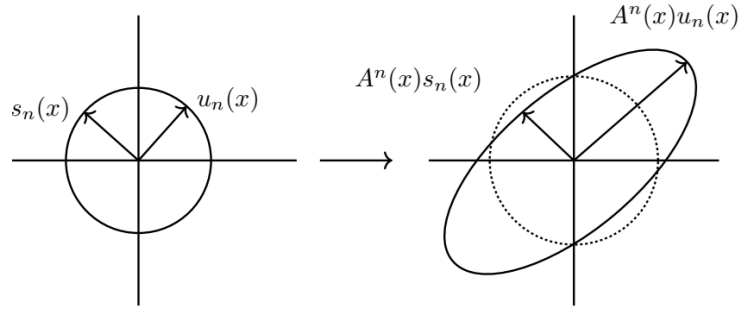
$$-\lambda(x) \leq \lim_n \frac{1}{n} \log \|A^n(x)v\| \leq \lambda(x).$$

Assim, quando $\lambda(x) = 0$, obtemos o caso 1).

Agora, vamos supor que $\lambda(x) > 0$. Seja $n_0 \in \mathbb{N}$ tal que para $n \geq n_0$ tem-se $\|A^n(x)\| \approx e^{n\lambda(x)} > 1$. Tome vetores ortonormais $s_n(x)$ e $u_n(x)$, respectivamente, tais que

$$\|A^n(x)s_n(x)\| = \|A^n(x)\|^{-1} \text{ e } \|A^n(x)u_n(x)\| = \|A^n(x)\|.$$

Figura 1 – Expansão e contração de s_n e u_n .



Fonte: Elaborado pelo autor.

Agora, iremos provar que $(s_n(x))_n$ converge, isto é, que o subespaço gerado por $s_n(x)$ convergirá para o espaço procurado.

Lema 3.3.6 *O ângulo $\alpha_n = \angle(s_n(x), s_{n+1}(x))$ decresce exponencialmente, com*

$$\limsup_n \frac{1}{n} \log |\sin \alpha_n| \leq -2\lambda(x).$$

Demonstração: Seja $\alpha_n = \angle(s_n(x), s_{n+1}(x))$. Com respeito à base ortonormal $\{s_{n+1}(x), u_{n+1}(x)\}$, o vetor $s_n(x)$ se escreve como

$$s_n(x) = \cos \alpha_n \cdot s_{n+1}(x) + \sin \alpha_n \cdot u_{n+1}(x).$$

A imagem por $A^{n+1}(x)$ desta base é, por definição, ortonormal, e portanto:

$$\begin{aligned} \|A^{n+1}(x)s_n(x)\| &= \|\cos \alpha_n(A^{n+1}(x)s_{n+1}(x)) + \sin \alpha_n(A^{n+1}(x)u_{n+1}(x))\| \\ &\geq \|\sin \alpha_n(A^{n+1}(x)u_{n+1}(x))\| \\ &= |\sin \alpha_n| \|A^{n+1}(x)\|. \end{aligned} \quad (3.11)$$

Por outro lado, por (3.5), tem-se

$$\begin{aligned} \|A^{n+1}(x)s_n(x)\| &= \|A(f^n(x))A^n(x)s_n(x)\| \\ &\leq \|A(f^n(x))\| \cdot \|A^n(x)s_n(x)\| \\ &= \|A(f^n(x))\| \cdot \|A^n(x)\|^{-1}. \end{aligned} \quad (3.12)$$

Combinando as desigualdades (3.11) e (3.12), temos

$$|\sin \alpha_n| \leq \frac{\|A(f^n(x))\|}{\|A^n(x)\| \|A^{n+1}(x)\|}.$$

Tomando logaritmo e dividindo por n , temos

$$\limsup_n \frac{\log |\sin \alpha_n|}{n} \leq \limsup_n \frac{\log \|A(f^n(x))\|}{n} - \limsup_n \frac{\log \|A^n(x)\|}{n} - \limsup_n \frac{\log \|A^{n+1}(x)\|}{n}. \quad (3.13)$$

Para $\varphi = \log \|A\|$, o Corolário 3.2.6 garante que, para μ -quase todo ponto $x \in M$, temos

$$\lim_n \frac{1}{n} \log \|A(f^n(x))\| = 0.$$

Desse modo, de (3.13) e pelo Furstenberg-Kersten tem-se que

$$\limsup_n \frac{1}{n} \log |\sin \alpha_n| \leq -2\lambda(x).$$

□

Lema 3.3.7 *A sequência $(s_n(x))_n$ é de Cauchy no espaço projetivo.*

Demonstração: Seja $\varepsilon > 0$ tal que $-2\lambda(x) + \varepsilon < 0$. Pelo Lema 3.3.6 e da definição de limite superior, existe $n_0 \in \mathbb{N}$ tal que para todo $n \geq n_0$ tem-se

$$|\sin \alpha_n| \leq e^{n(-2\lambda(x) + \varepsilon)}. \quad (3.14)$$

Considere o intervalo $K = [0, 2\pi]$. Se o ângulo entre $s_n(x)$ e $s_{n+1}(x)$ está em $\left[0, \frac{2\pi}{3}\right]$, sabemos que $\sin\left(\frac{\alpha_n}{2}\right) \leq \sin \alpha_n$, daí,

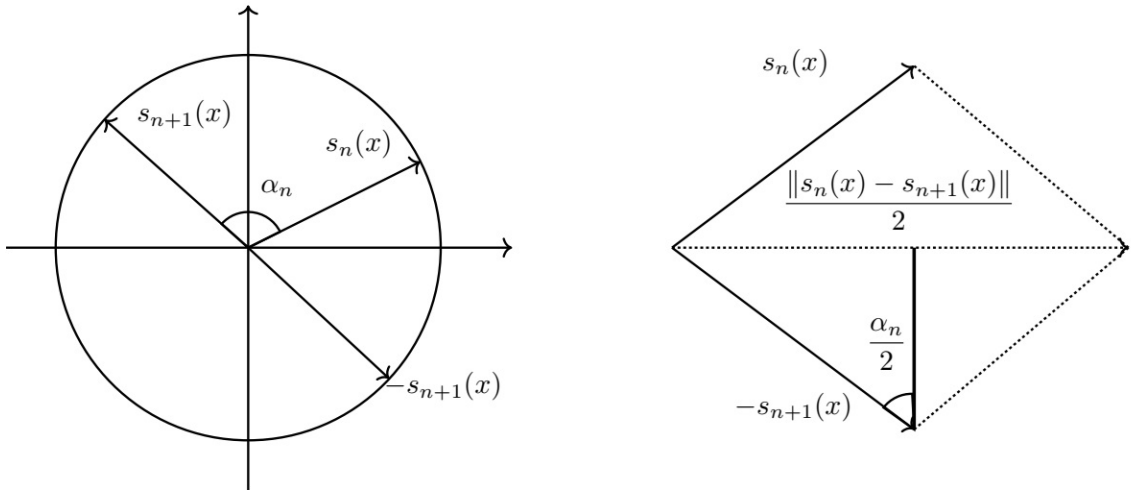
$$\frac{1}{2} \|s_n(x) - s_{n+1}(x)\| = \sin\left(\frac{\alpha_n}{2}\right) \leq \sin \alpha_n \quad (3.15)$$

Pela desigualdade (3.14) vamos ter

$$\|s_n(x) - s_{n+1}(x)\| \leq 2e^{n(-2\lambda(x) + \varepsilon)}. \quad (3.16)$$

Se o ângulo entre $s_n(x)$ e $s_{n+1}(x)$ está em $K \setminus \left[0, \frac{2\pi}{3}\right]$, trocando alguns $s_j(x)$ por $-s_j(x)$ se necessário para completar o losango, como as diagonais são bissetrizes e a altura de um triângulo isósceles, como na figura, pela definição do seno de um ângulo, tem-se o mesmo resultado (3.16).

Figura 2 – Losango gerador po s_n e s_{n-1} .



Fonte: Elaborado pelo autor.

Consequentemente, existe $C > 0$ tal que para n suficientemente grande vale

$$\|s_n(x) - s_{n+1}(x)\| \leq Ce^{n(-2\lambda(x) + \varepsilon)}. \quad (3.17)$$

Em particular, a sequência $(s_n(x))_n$ é de Cauchy.

□

Vamos definir $s(x) = \lim_n s_n(x)$ sempre que o limite existir. Seja $\beta_n(x)$ o ângulo entre $s(x)$ e $s_n(x)$. Como α_n converge para 0 exponencialmente, o mesmo ocorre para $\beta_n(x)$, com a mesma taxa exponencial. Logo

$$\limsup_n \frac{1}{n} \log |\sin \beta_n(x)| \leq -2\lambda(x). \quad (3.18)$$

Lema 3.3.8 *O vetor $s(x)$ é contraído à taxa de $\lambda(x)$, isto é:*

$$\lim_n \frac{1}{n} \log \|A^n(x)s(x)\| = -\lambda(x).$$

Demonstração: Seja $\beta_n(x) = \angle(s(x), s_n(x))$. Logo, $s(x)$ se escreve como

$$s(x) = \cos \beta_n(x) s_n(x) + \sin \beta_n(x) u_n(x),$$

na base ortonormal $\{s_n(x), u_n(x)\}$. Aplicando $A^n(x)$ em $s(x)$ temos

$$\begin{aligned} \|A^n(x)s(x)\| &= \|\cos \beta_n(x) A^n(x)s_n(x) + \sin \beta_n(x) A^n(x)u_n(x)\| \\ &\geq \|\cos \beta_n(x) A^n(x)s_n(x)\| \\ &= |\cos \beta_n(x)| \|A^n(x)\|^{-1}. \end{aligned}$$

Então,

$$\begin{aligned} \liminf_n \frac{1}{n} \log \|A^n(x)s(x)\| &\geq \liminf_n \frac{1}{n} (|\cos \beta_n(x)| \|A^n(x)\|^{-1}) \\ &= \liminf_n \frac{1}{n} \log |\cos \beta_n(x)| - \lim_n \frac{1}{n} \log \|A^n(x)\| \\ &= 0 - \lambda(x) = -\lambda(x) \end{aligned}$$

e portanto

$$\liminf_n \frac{1}{n} \log \|A^n(x)s(x)\| \geq -\lambda(x). \quad (3.19)$$

Por outro lado, como $\|A^n(x)s(x)\| \leq |\cos \beta_n(x)| \|A^n(x)s_n(x)\| + |\sin \beta_n(x)| \|A^n(x)u_n(x)\|$,

tem-se

$$\begin{aligned}
& \limsup_n \frac{1}{n} \log \|A^n(x)s(x)\| \\
& \leq \limsup_n \frac{1}{n} \log (|\cos \beta_n(x)| \|A^n(x)s_n(x)\| + |\sin \beta_n(x)| \|A^n(x)u_n(x)\|) \\
& = \max \left\{ \limsup_n \frac{1}{n} \log |\cos \beta_n(x)| \|A^n(x)s_n(x)\|, \limsup_n \frac{1}{n} \log |\sin \beta_n(x)| \|A^n(x)u_n(x)\| \right\} \\
& \leq \max \left\{ \limsup_n \frac{1}{n} \log \|A^n(x)\|^{-1}, \limsup_n \frac{1}{n} \log |\sin \beta_n(x)| + \limsup_n \frac{1}{n} \log \|A^n(x)\| \right\} \\
& \leq \max \{-\lambda(x), -2\lambda(x) + \lambda(x)\} \\
& = -\lambda(x).
\end{aligned}$$

Desse modo,

$$\lim_n \frac{1}{n} \log \|A^n(x)s(x)\| = -\lambda(x).$$

□

Note que o Lema 3.3.8 nos dá uma idéia de quem poderia ser E_x^s no teorema, mas para tal, precisaremos de mais uma lema. Seja $v \in \mathbb{R}^2$ que não seja colinear com $s(x)$ e seja $\gamma_n(x)$ o ângulo entre os vetores $s_n(x)$ e v .

Lema 3.3.9 *Se $v \in \mathbb{R}^2 \setminus \{0\}$ não é colinear com $s(x)$, então*

$$\lim_n \frac{1}{n} \log \|A^n(x)v\| = \lambda(x). \quad (3.20)$$

Demonstração: Podemos escrever $v = \cos \gamma_n(x)s_n(x) + \sin \gamma_n(x)u_n(x)$, logo

$$\begin{aligned}
\|A^n(x)v\| &= \|\cos \gamma_n(x)(A^n(x)s_n(x)) + \sin \gamma_n(x)(A^n(x)u_n(x))\| \\
&\geq |\sin \gamma_n(x)| \|A^n(x)u_n(x)\| \\
&= |\sin \gamma_n(x)| \|A^n(x)\|.
\end{aligned}$$

Como v não é colinear com $s(x)$, dado $\varepsilon > 0$ existe $n_0 \in \mathbb{N}$ tal que $\varepsilon < |\sin \gamma_n(x)| \leq 1$ para todo $n \geq n_0$, logo

$$\begin{aligned}
\liminf_n \frac{1}{n} \log \|A^n(x)v\| &\geq \liminf_n \frac{1}{n} \log |\sin \gamma_n(x)| \|A^n(x)\| \\
&= \liminf_n \frac{1}{n} \log \varepsilon + \liminf_n \frac{1}{n} \log \|A^n(x)\| \\
&= 0 + \lambda(x)
\end{aligned}$$

Assim,

$$\liminf_n \frac{1}{n} \log \|A^n(x)v\| \geq \lambda(x).$$

Por outro lado, uma vez que $\|A^n(x)v\| \leq \|A^n(x)\| \|v\|$, temos

$$\limsup_n \frac{1}{n} \log \|A^n(x)v\| \leq \lim_n \frac{1}{n} \log \|A^n(x)\| = \lambda(x).$$

Assim,

$$\lim_n \frac{1}{n} \log \|A^n(x)v\| = \lambda(x). \quad (3.21)$$

□

Lema 3.3.10 $A(x)s(x)$ é colinear a $s(f(x))$.

Demonstração: Com efeito, à luz do Lema 3.3.8, tem-se

$$\begin{aligned} -\lambda(x) &= \lim_n \frac{1}{n+1} \log \|A^{n+1}s(x)\| \\ &= \lim_n \frac{n}{n+1} \cdot \frac{1}{n} \log \|A^n(f(x))A(x)s(x)\| \\ &= \lim_n \frac{1}{n} \log \|A^n(f(x))A(x)s(x)\|. \end{aligned}$$

Usando o Lema 3.3.9 para $v \in \mathbb{R}^2 \setminus \{0\}$ não colinear a $s(f(x))$, temos

$$\begin{aligned} \lim_n \frac{1}{n} \log \|A^n(f(x))v\| &= \lambda(f(x)) \\ &= \lambda(x). \end{aligned}$$

Como $\lambda(x) \neq 0$, segue que $A(x)s(x)$ é necessariamente colinear a $s(f(x))$.

□

Parte final da prova do Teorema 3.3.5. Tome $E_x^s \subset \mathbb{R}^2$ como sendo a reta $\mathbb{R}s(x) = \{as(x); a \in \mathbb{R}\}$ gerada por $s(x)$. O Lema 3.3.8 prova a primeira afirmação em (3.10), o Lema 3.3.9 prova a segunda afirmação em (3.10). Ademais, temos $A(x)E_x^s = E_{f(x)}^s$, pelo Lema 3.3.10.

Agora vamos estender o resultado para $\text{GL}(2, \mathbb{R})$. Defina $c : M \rightarrow \mathbb{R}$ por $c(x) = |\det A(x)|^{\frac{1}{2}}$. Seja $B : M \rightarrow \text{SL}(2, \mathbb{R})$ tal que $A(x) = c(x)B(x)$ para $x \in M$. Sabemos que para cada $x \in M$ tem-se $|\det A(x)^+| \leq \|A(x)^+\|^2$. Logo, se $\log^+ \|A^{\pm 1}\| \in L^1(\mu)$ então $\log^+ c \in L^1(\mu)$. Como $\|A(x)\| = |c(x)| \|B(x)\|$, temos

$$\int \log^+ \|A(x)\| d\mu(x) = \int \log^+ c(x) d\mu(x) + \int \log^+ \|B(x)\| d\mu(x)$$

e portanto $\log^+ \|B\| \in L^1(\mu)$. Analogamente, $\log^+ \|B^{-1}\| \in L^1(\mu)$. Assim, para todo $v \in \mathbb{R}^2 \setminus \{0\}$, usando o Teorema de Binet para o produto finito de matrizes, temos que para μ -quase todo ponto $x \in M$:

$$\begin{aligned} \lim_n \frac{1}{n} \log \|A^n(x)v\| &= \lim_n \frac{1}{n} \log \left\| \prod_{j=0}^{n-1} c(f^j(x)) \cdot B^n(x)v \right\| \\ &= \lim_n \left[\frac{1}{n} \sum_{j=0}^{n-1} \log c(f^j(x)) + \frac{1}{n} \log \|B^n(x)v\| \right] \end{aligned}$$

Pelo Teorema 3.8, existe $t(x) = \lim_n \frac{1}{n} \sum_{j=0}^{n-1} \log c(f^j(x))$. Também existe, pela parte já provada desse teorema, o limite $\lim_n \frac{1}{n} \log \|B^n(x)v\|$, assim:

a) ou, $\lambda_{\max}(x) = \lambda_{\min}(x) = t(x)$ e

$$\lim_n \frac{1}{n} \log \|A^n(x)v\| = t(x),$$

b) ou, $\lambda_{\max}(x) > \lambda_{\min}(x)$ e existe uma reta $E_x^s \subset \mathbb{R}^2$ tal que

$$\lim_n \frac{1}{n} \log \|A^n(x)v\| = \begin{cases} \lambda_{\min}(x), & \text{se } v \in E_x^s \setminus \{0\} \\ \lambda_{\max}(x), & \text{se } v \in \mathbb{R}^2 \setminus E_x^s. \end{cases}$$

Note que $E_x^s = t(x) + \bar{E}_x^s$, onde $\bar{E}_x^s$ é o subespaço vetorial unidimensional dado pela parte já provada para o cociclo B . Ademais, $A(x) \cdot E_x^s = E_{f(x)}^s$.

□

Procedemos agora para a versão do teorema, assumindo que a transformação é invertível.

Teorema 3.3.11 (Caso Invertível) *Se $f : M \rightarrow M$ for invertível e $\log^+ \|A^\pm\| \in L^1(\mu)$, então para μ -quase todo ponto $x \in M$, uma das condições abaixo vale:*

1. $\lambda_{\max}(x) = \lambda_{\min}(x)$ e $\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log \|A^n(x)v\| = \lambda_{\max}(x)$, para todo $v \in \mathbb{R}^2 \setminus \{0\}$,
2. ou, $\lambda_{\min}(x) < \lambda_{\max}(x)$ e existe uma decomposição $\mathbb{R}^2 = E_x^u \oplus E_x^s$ tal que

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \log \|A^n(x)v\| = \begin{cases} \lambda_{\min}(x), & \text{se } v \in E_x^s \setminus \{0\} \\ \lambda_{\max}(x), & \text{se } v \in \mathbb{R}^2 \setminus E_x^s. \end{cases} \quad (3.22)$$

$$\lim_{n \rightarrow -\infty} \frac{1}{n} \log \|A^n(x)v\| = \begin{cases} \lambda_{\max}(x), & \text{se } v \in E_x^u \setminus \{0\} \\ \lambda_{\min}(x), & \text{se } v \in \mathbb{R}^2 \setminus E_x^u. \end{cases} \quad (3.23)$$

Além disso, como no Teorema 3.3.5, $A(x) \cdot E_x^u = E_{f(x)}^u$ e $A(x) \cdot E_x^s = E_{f(x)}^s$ e o ângulo entre essas duas retas decresce subexponencialmente ao longo de órbitas, ou seja:

$$\lim_n \frac{1}{n} \log |\sin \angle(E_{f^n(x)}^s, E_{f^n(x)}^u)| = 0. \quad (3.24)$$

Antes de fazermos a demonstração do Teorema 3.3.11, precisaremos da seguinte proposição.

Proposição 3.3.12 *Seja $A \in SL(2, \mathbb{R})$, então para todos os vetores $x, y \in \mathbb{R}^2 \setminus \{0\}$ tem-se*

$$\frac{1}{\|A\| \|A^{-1}\|} \leq \frac{|\sin \angle(A(x), A(y))|}{|\sin \angle(x, y)|} \leq \|A\| \|A^{-1}\|. \quad (3.25)$$

Demonstração: Seja $\varphi_A : \mathbb{RP}^2 \rightarrow \mathbb{RP}^2$, onde $\mathbb{RP}^2$ é o espaço projetivo real de dimensão 2, dada por $\varphi_A(x) = \frac{A(x)}{\|A(x)\|}$. Para $x \in \mathbb{RP}^2$ e $v \in T_x \mathbb{RP}^2$, temos

$$\begin{aligned} d(\varphi_A)_x \cdot v &= \lim_{t \rightarrow 0} \frac{1}{t} (\varphi_A(x + tv) - \varphi_A(x)) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \left(\frac{A(x + tv)}{\|A(x + tv)\|} - \frac{A(x)}{\|A(x)\|} \right). \end{aligned}$$

Caso φ_A seja Lipschitz, então o resultado segue. De fato, sabe-se que dados $p, q \in \mathbb{R} \setminus \{0\}$, temos

$$\left\| \frac{p}{\|p\|} - \frac{q}{\|q\|} \right\| \leq \max \left\{ \frac{1}{\|p\|}, \frac{1}{\|q\|} \right\} \|p - q\|. \quad (3.26)$$

Assim, para vermos que φ_A é Lipschitz, basta usar(3.26), donde

$$\begin{aligned} \|d(\varphi_A(x))v\| &= \left\| \lim_{t \rightarrow 0} \frac{1}{t} (\varphi_A(x + tv) - \varphi_A(x)) \right\| \\ &\leq \lim_{t \rightarrow 0} \frac{1}{|t|} \left\| \frac{A(x + tv)}{\|A(x + tv)\|} - \frac{A(x)}{\|A(x)\|} \right\| \\ &\leq \lim_{t \rightarrow 0} \frac{1}{|t|} \max \left\{ \frac{1}{\|A(x + tv)\|}, \frac{1}{\|A(x)\|} \right\} |t| \|A(v)\| \\ &= \lim_{t \rightarrow 0} \max \left\{ \frac{1}{\|A(x + tv)\|}, \frac{1}{\|A(x)\|} \right\} \|A(v)\| \\ &= \frac{\|A(v)\|}{\|A(x)\|} \\ &\leq \frac{\|A\| \|v\|}{m(A)}. \end{aligned}$$

Daí,

$$\frac{m(A)}{\|A\|} \leq \|d\varphi_A(x)\| \leq \frac{\|A\|}{m(A)}.$$

Logo, uma vez que $d(\varphi_A(x), \varphi_A(y)) = |\sin \angle(A(x), A(y))|$, temos

$$\frac{m(A)}{\|A\|} \leq \frac{|\sin \angle(A(x), A(y))|}{|\sin \angle(x, y)|} \leq \frac{\|A\|}{m(A)}.$$

□

Demonstração: *Prova do Teorema 3.3.11:* Vamos tratar o caso em que A toma valores em $\mathrm{SL}(2, \mathbb{R})$ (a extensão para $\mathrm{GL}(d, \mathbb{R})$ é feita de modo similar ao que foi feito no Teorema 3.3.5). Assim, considere $x \in M$ como na conclusão do Teorema 3.2.7, e $\lambda(x) = \lambda_{\max}(x) = -\lambda_{\min}(x)$. O caso em que $\lambda(x) = 0$ segue diretamente do Teorema 3.3.5 aplicado a F e F^{-1} . Agora, vamos assumir que $\lambda(x) > 0$. Sejam $E_x^s = \mathbb{R}s(x)$ e $E_x^u = \mathbb{R}u(x)$ os subespaços dados pelo Teorema 3.3.5 para F e F^{-1} , respectivamente. Precisamos mostrar que essas duas retas são transversais.

Lema 3.3.13 *Os vetores $s(x)$ e $u(x)$ são não colineares para μ -quase todo ponto de $\{x : \lambda(x) > 0\}$.*

Demonstração: Basta mostrar que

$$\lim_{n \rightarrow -\infty} \frac{1}{n} \log \|A^n(x)|E_x^s\| = -\lambda(x)$$

para μ -quase todo ponto em $\{x : \lambda(x) > 0\}$. Sejam

$$\psi_n(x) = \frac{1}{-n} \log \|A^{-n}(x)|E_x^s\| \quad \text{e} \quad \varphi_n(x) = \frac{1}{-n} \log \|(A^n(x)|E_x^s)^{-1}\|.$$

Usando o Teorema 3.3.5 para F^{-1} , para μ -quase todo ponto $x \in M$ a sequência ψ_n converge, digamos para $\psi(x)$. Usando (3.5), temos

$$(A^n(f^{-n}(x))|E_{f^{-n}(x)}^s)^{-1} = A^{-n}(x)|E_x^s.$$

Desse modo,

$$\begin{aligned} \varphi_n \circ f^{-n}(x) &= \frac{1}{-n} \log \|(A^n(f^{-n}(x))|E_{f^{-n}(x)}^s)^{-1}\| \\ &= \frac{1}{-n} \log \|A^{-n}(x)|E_x^s\| \\ &= \psi_n(x). \end{aligned}$$

Assim, $\varphi_n \circ f^{-n} = \psi_n$. Por (3.10), para μ -quase todo ponto $x \in M$, o espaço gerado por $s(x)$ é contraído à taxa exponencial de $-\lambda(x)$, ou seja, φ_n converge para $-\lambda$. Em particular, a sequência φ_n converge para $-\lambda$ em medida, isto é, para todo $\varepsilon > 0$ tem-se

$$\lim_n \mu(\{x \in M : |\varphi_n(x) + \lambda(x)| > \varepsilon\}) = 0. \quad (3.27)$$

Como λ é f -invariante, para todo $k \in \mathbb{Z}$ temos

$$\begin{aligned} \{x \in M : |\varphi_n(x) + \lambda(x)| > \varepsilon\} &= \{x \in M : |\varphi_n(f^k(x)) + \lambda(f^k(x))| > \varepsilon\} \\ &= \{x \in M : |\varphi_n(f^k(x)) + \lambda(x)| > \varepsilon\}. \end{aligned}$$

Em particular,

$$\begin{aligned} \lim_n \mu(\{x \in M : |\varphi_n(x) + \lambda(x)| > \varepsilon\}) &= \lim_n \mu(\{x \in M : |\varphi_n(f^{-n}(x)) + \lambda(x)| > \varepsilon\}) \\ &= \lim_n \mu(\{x \in M : |\psi_n(x) + \lambda(x)| > \varepsilon\}). \end{aligned}$$

De (3.27), temos

$$\lim_n \mu(\{x \in M : |\psi_n(x) + \lambda(x)| > \varepsilon\}) = 0. \quad (3.28)$$

Tendo em vista (3.28), concluímos que ψ_n converge em medida para $-\lambda$, e como ψ_n converge para ψ , então $\psi = -\lambda$ para μ -quase todo ponto. Assim, temos

$$\lim_n \frac{1}{n} \log \|A^{-n}(x)E_x^s\| = -\lambda(x).$$

Como

$$\lim_n \frac{1}{n} \log \|A^{-n}(x)E_x^u\| = \lambda(x),$$

segue que $E_x^s \neq E_x^u$.

□

Lema 3.3.14 *Seja $\theta(y) = \angle(E_y^s, E_y^u)$. Para μ -quase todo ponto x com $\lambda(x) > 0$, temos*

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

Demonstração: Usando a Proposição 3.3.12, temos

$$\frac{m(A(x))}{\|A(x)\|} \leq \frac{|\sin \angle(A(x)E_x^s, A(x)E_x^u)|}{|\sin \angle(E_x^s, E_x^u)|} \leq \frac{\|A(x)\|}{m(A(x))}.$$

Relembre que $A(x) \in \text{SL}(2, \mathbb{R})$, logo $m(A(x)) = \|A(x)\|^{-1}$. Usando que os espaços E_x^s e E_x^u são invariantes pelo cociclo, obtemos

$$m(A(x))^2 \leq \frac{|\sin \theta(f(x))|}{|\sin \theta(x)|} \leq \|A(x)\|^2.$$

Tomando logaritmo nesta desigualdade, obtemos

$$|\log |\sin \theta(f(x))| - \log |\sin \theta(x)|| \leq 2 \log \|A(x)\|.$$

Integrando ambos os lados dessa desigualdade, temos

$$\int |\log |\sin \theta(f(x))| - \log |\sin \theta(x)|| d\mu(x) \leq 2 \int \log \|A(x)\| d\mu(x) < \infty.$$

Portanto, $\log |\sin \theta(f(x))| - \log |\sin \theta(x)| \in L^1(\mu)$. À luz do Corolário 3.2.6 aplicado à função $\rho(x) = \log |\sin(x)|$ temos

$$\lim_{n \rightarrow \pm\infty} \frac{1}{n} \log |\sin \angle(E_{f^n(x)}^s, E_{f^n(x)}^u)| = 0.$$

□

Isto finaliza a prova do Teorema 3.3.11.

□

4 DESIGUALDADE DE RUELLE

Neste capítulo, provaremos um resultado básico da teoria ergódica para difeomorfismos, que relaciona entropia e expoentes de Lyapunov. O resultado é uma desigualdade, conhecida como **desigualdade de Ruelle**, e foi provado por David Ruelle na década de 1970. Mencionamos que Grigory Margulis provou uma versão particular da desigualdade anteriormente a Ruelle, para aplicações que preservam volume. Por essa razão, alguns autores chamam-na de desigualdade de Margulis-Ruelle. Os resultados deste capítulo podem ser encontrados em [5].

4.1 Desigualdade de Ruelle

Teorema 4.1.1 (Desigualdade de Ruelle) *Seja $f \in \text{Diff}^1(M)$, onde M é uma variedade compacta sem bordo. Para toda $\mu \in \mathcal{M}_f(M)$, tem-se*

$$h_\mu(f) \leq \int \left(\sum_{\lambda_i(x) > 0} m_i(x) \lambda_i(x) \right) d\mu(x). \quad (4.1)$$

Por simplicidade da demonstração, restringimos a prova ao caso em que $M \subset \mathbb{R}^d$ com a métrica euclidiana. Isso evita o uso de mapas exponenciais que, embora causem pouca distorção local da métrica, carregam a notação.

Primeiro vamos supor que μ é ergódica. Dessa maneira, teremos que $\lambda_i(x) = \lambda_i$ é constante para μ -quase todo ponto. Logo, basta provar que

$$h_\mu(f) \leq \sum_{\lambda_i > 0} m_i \lambda_i. \quad (4.2)$$

Podemos supor que $\sup_{x \in M} \|df_x\| \geq 1$, pois quando $\sup_{x \in M} \|df_x\| < 1$, existe $\sigma > 0$ tal que $\sup_{x \in M} \|df_x\| \leq \sigma < 1$. Dessa maneira, pelo Teorema de Mudança de Variável temos

$$\begin{aligned} \text{Vol}(M) &= \text{Vol}(f(M)) = \int_{f(M)} dx = \int_M |\det df_y| dy \\ &\leq \int_M \sqrt[d]{\|df_y\|} dy \leq \int_M \sqrt[d]{\sigma} dy < \int_M dy = \text{Vol}(M), \end{aligned}$$

o que é um absurdo.

A prova utilizará uma série de lemas. O primeiro deles envolve apenas a diferenciabilidade de f .

Lema 4.1.2 Para todo m , existe $\varepsilon_0 = \varepsilon_0(f^m) > 0$ tal que para todo $\varepsilon \in (0, \varepsilon_0]$ e todo $x \in M$ tem-se

$$f^m(B_\varepsilon(x)) \subset 2df_x^m(B_\varepsilon(0)) + f^m(x). \quad (4.3)$$

Demonstração: Como $M \subset \mathbb{R}^d$, o mapa exponencial $\exp_x : T_x M \rightarrow M$ é dado por $\exp_x(v) = v + x$. Fixado $x \in M$, escreva $A = df_x^m$ e considere $F = \exp_{f^m(x)}^{-1} \circ f^m \circ \exp_x$. Temos que $F(0) = 0$ e F é diferenciável, com $dF_0 = A$. Como $F(v) = f^m(v + x) - f^m(x)$, queremos mostrar que existe $\varepsilon_0 = \varepsilon_0(f^m) > 0$ tal que para todo $\varepsilon \in (0, \varepsilon_0]$ vale

$$F(B_\varepsilon(0)) \subset 2A(B_\varepsilon(0)),$$

ou seja, para cada $z \in B_\varepsilon(0)$ devemos encontrar $z' \in B_\varepsilon(0)$ tal que $F(z) = 2A(z')$. Seja $\beta := \inf\{m(df_x^m) : x \in M\}$. Pela diferenciabilidade de F , existe $\delta_0 > 0$ tal que

$$\|z\| < \delta_0 \implies F = A + E \text{ com } \|E(z)\| \leq \beta\|z\|. \quad (4.4)$$

Dado z com $\|z\| < \delta_0$, temos

$$F(z) = 2A(z') \iff 2A(z') = A(z) + E(z) \iff 2z' = z + A^{-1}E(z) \iff z' = \frac{z + A^{-1}E(z)}{2}.$$

Por (4.4), temos $\|E(z)\| \leq \beta\|z\| \leq \|A^{-1}\|^{-1}\|z\|$ e, portanto,

$$\|A^{-1}E(z)\| \leq \|A^{-1}\| \cdot \|E(z)\| \leq \|z\|.$$

Assim $\|z'\| \leq \|z\|$. Tomando $\varepsilon_0 := \min\left\{\delta_0, \frac{\rho}{\sup_{x \in M} \|df_x^m\| + 1}\right\}$ onde ρ é o raio de injetividade de M , o resultado segue. □

O próximo lema constrói partições que se relacionam bem com bolas.

Lema 4.1.3 Seja $(M, \mathcal{F}, \mu, g)$ um sistema que preserva medida. Para todo $r > 0$, existe uma partição $\mathcal{A}$ com as seguintes propriedades:

- a) $\text{diam}(\mathcal{A}) \leq 2r$.
- b) Para todo $A \in \mathcal{A}$, existe $y \in A$ tal que

$$B_{\frac{r}{2}}(y) \subset A \subset \overline{B_r(y)}.$$

Demonstração: Usando a compacidade de M , podemos considerar $\Gamma = \{z_1, \dots, z_\ell\} \subset M$ um conjunto de pontos r -separado maximal. Temos

$$d(z_i, z_j) > r, \text{ para todos } z_i, z_j \in \Gamma, i \neq j.$$

Como Γ é maximal, para todo $x \in M$ existe $z \in \Gamma$ tal que $d(x, z) \leq r$, caso contrário o conjunto $\Gamma \cup \{x\}$ é r -separado, o que contradiz a maximalidade de Γ . Assim

$$M = \bigcup_{k=1}^{\ell} \overline{B_r(z_k)}.$$

Para cada $z_i \in \Gamma$, defina

$$A'_i = \{y \in M : d(y, z_i) \leq d(y, z_j) \text{ para todo } j \neq i\}.$$

Veja que para cada $j \in \{1, \dots, \ell\}$, A'_j é mensurável e, ainda, $\{A'_1, \dots, A'_\ell\}$ cobre M , pois M é compacta (mas não constitui uma partição pois os A'_i podem se intersectar). Dado $x \in B_{\frac{r}{2}}(z_i)$, temos para todo $j \neq i$

$$r < d(z_i, z_j) \leq d(z_i, x) + d(x, z_j) < \frac{r}{2} + d(x, z_j) \implies \frac{r}{2} < d(x, z_j), \quad (4.5)$$

onde $x \in A'_i$ e $x \notin A'_j$ para todo $j \neq i$. Por outro lado, se existe $y \in A'_i \setminus \overline{B_r(z_i)}$, então para todo $j \neq i$ tem-se

$$r < d(y, z_i) \leq d(y, z_j) \implies r < d(y, z_j),$$

ou seja, $\Gamma \cup \{y\}$ é r -separado, o que contradiz a maximalidade de Γ . Portanto, para todo $j \in \{1, \dots, \ell\}$ temos

$$B_{\frac{r}{2}}(z_j) \subset A'_j \subset \overline{B_r(z_j)}. \quad (4.6)$$

Vamos agora definir a partição $\{A_1, \dots, A_\ell\}$ por

$$A_1 = A'_1 \text{ e } A_j = A'_j \setminus \left(\bigcup_{i=1}^{j-1} A'_i \right) \text{ para } j = 2, \dots, \ell.$$

Note que para todo $j \in \{1, \dots, \ell\}$ tem-se que A_j é mensurável e, por construção, $A_i \cap A_j = \emptyset$ para todos $i \neq j$. Assim, $\mathcal{A} = \{A_1, \dots, A_\ell\}$ define uma partição de M . Por (4.5), para todo j vale que $B_{\frac{r}{2}}(z_j) \subset A_j$. Ademais, segue de (4.6) e da construção da partição que $A_j \subset \overline{B_r(z_j)}$. Em particular, $\text{diam}(\mathcal{A}) \leq 2r$, o que prova os itens a) e b).

□

Assim, se m for suficientemente grande e $\eta > 0$, pelos Lemas 4.1.3 e 4.1.2 temos que para todo $\varepsilon \in (0, \min\{\varepsilon_0(f^m), \varepsilon_1(f^m, \eta)\})$ existe uma partição $\mathcal{A}$ de M com as seguintes propriedades:

a) $\text{diam}(\mathcal{A}) < \varepsilon$.

b)

$$h_\mu(f^m, \mathcal{A}) \geq h_\mu(f^m) - \eta. \quad (4.7)$$

c) existe $r \in (0, \varepsilon/2)$ tal que para todo $A \in \mathcal{A}$ existe $y \in A$ tal que

$$B_{\frac{r}{2}}(y) \subset A \subset \overline{B_r(y)}; \quad (4.8)$$

d) Para todo $x \in M$ vale

$$f^m(B_{2r}(x)) \subset 2df_x^m(B_{2r}(0)) + f^m(x). \quad (4.9)$$

Fixe uma tal partição. Como f é invertível, podemos aplicar a Proposição 2.2.8 e obter que

$$\begin{aligned} h_\mu(f^m, \mathcal{A}) &= h_\mu(f^{-m}, \mathcal{A}) \\ &= \lim_k H_\mu(\mathcal{A} | f^m(\mathcal{A}) \vee f^{2m}(\mathcal{A}) \vee \dots \vee f^{km}(\mathcal{A})) \end{aligned} \quad (4.10)$$

$$\begin{aligned} &= \inf_{k \geq 1} H_\mu(\mathcal{A} | f^m(\mathcal{A}) \vee f^{2m}(\mathcal{A}) \vee \dots \vee f^{km}(\mathcal{A})) \\ &\leq H_\mu(\mathcal{A} | f^m(\mathcal{A})) \\ &= \sum_{D \in f^m(\mathcal{A})} \mu(D) \sum_{A \in \mathcal{A}} -\frac{\mu(A \cap D)}{\mu(D)} \log \left(\frac{\mu(A \cap D)}{\mu(A)} \right) \end{aligned} \quad (4.11)$$

$$\begin{aligned} &= \sum_{D \in f^m(\mathcal{A})} \mu(D) H_\mu(\mathcal{A} | D) \\ &\leq \sum_{D \in f^m(\mathcal{A})} \mu(D) \log \#\{A \in \mathcal{A} : A \cap D \neq \emptyset\}, \end{aligned} \quad (4.12)$$

onde na última desigualdade utilizamos a Proposição 2.1.4. Agora, temos o interesse em limitar $\#\{A \in \mathcal{A} : A \cap D \neq \emptyset\}$. Para tal, façamos os seguintes lemas. O primeiro deles estabelece uma estimativa uniforme, independente das propriedades dos expoentes de Lyapunov.

Lema 4.1.4 *Para todo $D \in f^m(\mathcal{A})$ tem-se*

$$\#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} \leq 12^d \sup_{x \in M} \|df_x\|^{md}. \quad (4.13)$$

Demonstração: Seja $r' = r/2$. Sabemos que para todo $A \in \mathcal{A}$ existe $y \in A$ tal que $B_{r'}(y) \subset A \subset \overline{B_r(y)} \subset B_{2r}(y)$. Em particular,

$$\text{diam}(A) \leq 2r. \quad (4.14)$$

Seja $D = f^m(A) \in f^m(\mathcal{A})$. Pela desigualdade do Valor Médio, temos

$$\text{diam}(D) \leq \sup_{x \in A} \|df_x^m\| \cdot \text{diam}(A') \leq \sup_{x \in M} \|df\|^m \cdot 2r. \quad (4.15)$$

Dado $\delta > 0$, uma δ -vizinhança de D é definida por

$$B_\delta(D) := \bigcup_{x \in D} B_\delta(x).$$

Caso $A \cap D \neq \emptyset$ então, usando (4.8), temos $A \subset B_{2r}(D)$. Note que $\text{diam}(B_{2r}(D)) < \text{diam}(D) + 4r$, pois dados $x, y \in B_{2r}(D)$, existem $x' \in D$ e $y' \in D$ tais que $\|x - x'\| < 2r$ e $\|y - y'\| < 2r$ e portanto pela desigualdade triangular temos

$$\|x - y\| \leq \|x - x'\| + \|x' - y'\| + \|y - y'\| < \|x' - y'\| + 4r.$$

Dessa maneira $B_{2r}(D) \subset B_{\text{diam}(D)+4r}(z)$ para qualquer $z \in D$. Usando (4.15), temos

$$\begin{aligned} \text{Vol}(B_{2r}(D)) &\leq \text{Vol}(B_{\text{diam}(D)+4r}(z)) \leq \omega_d \cdot (\text{diam}(D) + 4r)^d \\ &\leq \omega_d \cdot \left(\sup_{x \in M} \|df_x\|^m \cdot 2r + 4r \right)^d \\ &= \omega_d \cdot \left(\sup_{x \in M} \|df_x\|^m + 2 \right)^d (2r)^d, \end{aligned} \quad (4.16)$$

onde $\omega_d = \text{Vol}(B_1(0))$. Como $B_{\frac{r}{2}}(y) \subset A$, temos $\omega_d \cdot \left(\frac{r}{2}\right)^d \leq \text{Vol}(A)$. Assim

$$\#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} \cdot \omega_d \cdot \left(\frac{r}{2}\right)^d \leq \sum_{A \cap D \neq \emptyset} \text{Vol}(A).$$

Como $\text{Vol } D \leq \text{Vol } B_{2r}(D)$, usando (4.16), tem-se

$$\#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} \cdot \omega_d \cdot \left(\frac{r}{2}\right)^d \leq \omega_d \cdot \left(\sup_{x \in M} \|df_x\|^m + 2 \right)^d (2r)^d.$$

Por conseguinte,

$$\begin{aligned} \#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} &\leq \left(\sup_{x \in M} \|df_x\|^m + 2 \right)^d 4^d \\ &\leq \sup_{x \in M} \|df_x\|^{md} \left(\frac{1}{2} + \frac{1}{\sup_{x \in M} \|df_x\|^m} \right)^d 2^d 4^d \\ &\leq 12^d \cdot \sup_{x \in M} \|df_x\|^{md}, \end{aligned}$$

onde na última desigualdade usamos que $\sup_{x \in M} \|df_x\|^m \geq 1$ para m grande.

□

Observe que o lema acima não envolve a dinâmica de f , tampouco as propriedades de expansão e contração determinadas pelos expoentes de Lyapunov. Nossa meta agora é levar em conta a dinâmica de f e refinar a desigualdade (4.13). Para tal, vamos introduzir o conjunto dos pontos “quase regulares”. Seja

$$R_{m,\eta} = \left\{ x \in M : \left| \frac{1}{k} \log \sigma_i(df_x^k) - \lambda_i \right| < \eta, \text{ para todo } k \geq m \text{ e } i \right\},$$

onde λ_i é o i -ésimo expoente de Lyapunov da medida μ . Observe que $R_{m,\eta}$ é mensurável e para todo $\eta > 0$ vale que

$$\mu(M \setminus R_m(f, \eta)) \rightarrow 0 \text{ quando } m \rightarrow \infty. \quad (4.17)$$

Lema 4.1.5 *Seja $D \in f^m(\mathcal{A})$ tal que $f^{-m}(D) \cap R_{m,\eta} \neq \emptyset$. Então*

$$\#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} \leq 48^d \cdot e^{m\eta d} \prod_{\lambda_i > 0} e^{m_i(\lambda_i + \eta)m}. \quad (4.18)$$

Demonstração: Sejam $A_0 = f^{-m}(D)$ e $x \in A_0 \cap R_{m,\eta}$. Seja $B := B_{2r}(x)$. Por (4.8) e (4.9), temos que $D = f^m(A_0) \subset f^m(B) \subset E_0 := 2df_x^m(B_{2r}(0)) + f^m(x)$. Assim, para todo $A \in \mathcal{A}$ tal que $A \cap D \neq \emptyset$ tem-se $A \cap E_0 \neq \emptyset$. Ainda por (4.8), sabe-se que existe $y \in A$ com $B_{\frac{r}{2}}(y) \subset A$, donde

$$\omega_d \cdot \left(\frac{r}{2}\right)^d \leq \text{Vol}(A)$$

e portanto

$$\#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} \cdot \omega_d \cdot \left(\frac{r}{2}\right)^d \leq \sum_{A \cap D \neq \emptyset} \text{Vol}(A). \quad (4.19)$$

Seja $E_1 = B_{2r}(E_0)$. Se $A \cap E_0 \neq \emptyset$ então $A \subset E_1$, logo por (4.19) temos

$$\#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} \cdot \omega_d \cdot \left(\frac{r}{2}\right)^d \leq \text{Vol}(E_1). \quad (4.20)$$

Vamos estimar o volume de E_1 colocando-o dentro de um elipsóide que se relaciona com E_0 . Sejam $\lambda_1, \dots, \lambda_\ell$ os expoentes de Lyapunov positivos, e considere seus respectivos valores singulares $\sigma_1(df_x^m), \dots, \sigma_\ell(df_x^m)$. Seja $E := E(k \cdot 8r \cdot \sigma_1(df_x^m), \dots, k \cdot 8r \cdot \sigma_\ell(df_x^m), k \cdot 8r \cdot e^{m\eta}, \dots, k \cdot 8r \cdot e^{m\eta})$ um elipsóide, onde $k > 0$ será determinado de forma que $E_1 \subset E$. Veja que $\sigma_i(df_x^m) \leq e^{m\eta}$ para todo $\lambda_i \leq 0$. Seja $\bar{x} \in E_1$. Escreva $\bar{x} = x + x'$ com $x \in E_0$ e

$\|x'\| < 2r$. Usando que $(a+b)^2 \leq 2(a^2+b^2)$ para $a, b \in \mathbb{R}$, temos

$$\begin{aligned}
\sum_{i=1}^{\ell} \frac{(x_i + x'_i)^2}{(k \cdot 8r \cdot \sigma_i(df_x^m))^2} + \sum_{i=\ell+1}^d \frac{(x_i + x'_i)^2}{(k \cdot 8r \cdot e^{m\eta})^2} &\leq \frac{2}{k^2} \left(\sum_{i=1}^{\ell} \frac{x_i^2}{(8r \cdot \sigma_i(df_x^m))^2} + \sum_{i=\ell+1}^d \frac{x_i^2}{(8r \cdot e^{m\eta})^2} \right) \\
&+ \frac{2}{k^2} \left(\sum_{i=1}^{\ell} \frac{(x'_i)^2}{(8r \cdot \sigma_i(df_x^m))^2} + \sum_{i=\ell+1}^d \frac{(x'_i)^2}{(8r \cdot e^{m\eta})^2} \right) \\
&\leq \frac{2}{k^2} \left(\sum_{i=1}^d \frac{x_i^2}{(8r \cdot \sigma_i(df_x^m))^2} \right) + \frac{2}{k^2} \left(\sum_{i=1}^d \frac{(x'_i)^2}{(2r)^2} \right) \\
&\leq \frac{4}{k^2}.
\end{aligned}$$

Tomando $k = 3$, segue que $E_1 \subset E$. Logo

$$\begin{aligned}
\text{Vol}(E_1) &\leq \text{Vol}(E) = \omega_d \cdot 3^d \cdot (8r)^d \prod_{i=1}^{\ell} \sigma_i(df_x^m) \prod_{i=\ell+1}^d e^{m\eta} \\
&\leq \omega_d \cdot 3^d \cdot (8r)^d \cdot e^{m\eta d} \prod_{\lambda_i > 0} e^{m(\lambda_i + \eta)m_i}.
\end{aligned} \tag{4.21}$$

Combinando (4.20) e (4.21), obtemos

$$\#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} \cdot \omega_d \cdot \left(\frac{r}{2}\right)^d \leq \omega_d \cdot 3^d \cdot (8r)^d \cdot e^{m\eta d} \prod_{\lambda_i > 0} e^{m(\lambda_i + \eta)m_i},$$

o que implica

$$\#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} \leq 48^d \cdot e^{m\eta d} \prod_{\lambda_i > 0} e^{m(\lambda_i + \eta)m_i}.$$

□

Agora temos todos os ingredientes para estabelecer a desigualdade de Ruelle para medidas ergódicas.

Demonstração: (Desigualdade de Ruelle - caso ergódico) Combinando (4.7) e (4.10), temos

$$h_\mu(f^m) - \eta \leq \sum_{D \in f^m(\mathcal{A})} \mu(D) \log \#\{A \in \mathcal{A} : A \cap D \neq \emptyset\}. \tag{4.22}$$

Usando os Lemas 4.1.4 e 4.1.5 e o fato de que μ é f -invariante, segue que

$$\begin{aligned}
h_\mu(f^m) - \eta &\leq \sum_{D \in f^m(\mathcal{A})} \mu(D) \log \#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} \\
&= \sum_{f^{-m}(D) \cap R_{m,\eta} = \emptyset} \mu(f^{-m}(D)) \log \#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} + \\
&\quad \sum_{f^{-m}(D) \cap R_{m,\eta} \neq \emptyset} \mu(f^{-m}(D)) \log \#\{A \in \mathcal{A} : A \cap D \neq \emptyset\} \\
&\leq \sum_{f^{-m}(D) \cap R_{m,\eta} = \emptyset} \mu(D) \log \left(12^d \sup_{x \in M} \|df_x\|^{md} \right) + \\
&\quad \sum_{f^{-m}(D) \cap R_{m,\eta} \neq \emptyset} \mu(D) \log \left(48^d \cdot e^{m\eta d} \prod_{\lambda_i > 0} e^{m(\lambda_i + \eta)m_i} \right) \\
&\leq \left[d \log 12 + md \log \left(\sup_{x \in M} \|df_x\| \right) \right] \mu(M \setminus R_{m,\eta}) + d \log 48 + \\
&\quad m\eta d + \sum_{\lambda_i > 0} m(\lambda_i + \eta)m_i. \tag{4.23}
\end{aligned}$$

Dividindo por m , obtemos que

$$h_\mu(f) \leq \frac{\eta}{m} + \left[\frac{d \log 12}{m} + d \log \left(\sup_{x \in M} \|df_x\| \right) \right] \mu(M \setminus R_{m,\eta}) + \frac{d \log 48}{m} + \eta d + \sum_{\lambda_i > 0} (\lambda_i + \eta)m_i.$$

Fazendo $m \rightarrow \infty$, segue de (4.17) que

$$h_\mu(f) \leq \eta d + \sum_{\lambda_i > 0} (\lambda_i + \eta)m_i.$$

Como η é arbitrário, fazendo $\eta \rightarrow 0$ concluímos que

$$h_\mu(f) \leq \sum_{\lambda_i > 0} m_i,$$

o que estabelece (4.2).

□

Agora, vamos supor que μ não é ergódica. Seja $\{\mu_P : P \in \mathcal{P}\}$ a decomposição ergódica de μ . Devemos escrever a entropia $h_\mu(f)$ em termos de sua decomposição ergódica, e para tal precisaremos do seguinte teorema.

Teorema 4.1.6 (Jacobs) *Suponha que M é um espaço métrico completo separável. Seja $(M, \mathcal{F}, \mu, f)$ um sistema que preserva medida, onde μ é uma de probabilidade, e seja $\{\mu_P : P \in \mathcal{P}\}$ a sua decomposição ergódica. Então*

$$h_\mu(f) = \int h_{\mu_P}(f) d\hat{\mu}(P). \tag{4.24}$$

Demonstração: Vide [7, Teorema 9.6.2].

□

Demonstração: (Desigualdade de Ruelle - caso não-ergódico) Como cada μ_P é ergódica, podemos utilizar o resultado mostrado para medidas ergódicas. Dessa maneira, usando o Teorema 4.1, temos

$$\begin{aligned}
 h_\mu(f) &= \int h_{\mu_P}(f) d\hat{\mu}(P) \\
 &\leq \int \left(\sum_{i:\lambda_i} m_i(P) \lambda_i(P) \right) d\hat{\mu}(P) \\
 &= \int \left(\sum_{\lambda_i(P(x))>0} m_i(P(x)) \lambda_i(P(x)) \right) d\mu(x) \\
 &= \int \left(\sum_{\lambda_i(x)>0} m_i(x) \lambda_i(x) \right) d\mu(x).
 \end{aligned}$$

□

Corolário 4.1.7 *Seja $f \in \text{Diff}^1(M)$, onde M é uma variedade compacta sem bordo. Se $\mu \in \mathcal{M}_f(M)$ ergódica é tal que $h_\mu(f) > 0$, então μ possui pelo menos um expoente de Lyapunov positivo e pelo menos um negativo. Em particular, quando M é uma superfície, então seus dois expoentes de Lyapunov são tais que $\lambda_{\max} > 0$ e $\lambda_{\min} < 0$.*

Demonstração: Usando a Desigualdade Ruelle, temos

$$0 < h_\mu(f) \leq \sum_{\lambda_i > 0} m_i \lambda_i,$$

o que implica a existência de algum i tal que $\lambda_i > 0$. Por outro lado, como $h_\mu(f) = h_\mu(f^{-1})$, segue que f^{-1} tem um expoente positivo. Uma vez que o espectro de f^{-1} é simétrico ao de f , segue que existe j tal que $\lambda_j < 0$.

□

5 UM CONTRA-EXEMPLO PARA A DESIGUALDADE DE RUELLE NO CASO NÃO-COMPACTO

Neste capítulo, vamos mostrar que a compacidade da variedade no Teorema 4.1.1 não pode ser retirada. Será feita a construção de um sistema dinâmico similar a um fluxo de suspensão sobre uma transformação de intercâmbio de intervalos, cujo comportamento local é o de uma translação. Os resultados desse capítulo foram obtidos por Riquelme [9].

5.1 Construção da variedade Riemanniana não-compacta

No que segue, iremos introduzir o conceito de intercâmbio de intervalos, que será utilizado na construção da variedade.

Definição 5.1.1 (Transformação de Intercâmbio de Intervalos) *Uma transformação de intercâmbio de intervalos (TII) é uma transformação invertível $T : [0, 1) \rightarrow [0, 1)$ satisfazendo as seguintes condições:*

- a) *Existe uma sequência estritamente crescente $\{x_i\}_{i \geq 0} \subset [0, 1)$ e uma sequência $\{a_i\}_{i \geq 0} \subset \mathbb{R}$ tal que $x_0 = 0$, $\lim_{i \rightarrow \infty} x_i = 1$ e $T(x) = x + a_i$ para todo $x \in [x_i, x_{i+1})$;*
- b) *O único ponto de acumulação do conjunto $\{x_i + a_i\} \cup \{x_{i+1} + a_{i+1}\}_{i \geq 0}$ é 1.*

Denote por m a medida de Lebesgue em $[0, 1)$ e seja T uma TII. Como T é definido por partes por translações, T preserva a medida m . Vamos denotar por $\mathcal{I}$ a partição de $[0, 1)$ definida por $\{[x_i, x_{i+1})\}_{i \geq 0}$. Essa partição satisfaz a seguintes propriedade de entropia.

Proposição 5.1.2 (Blume) *Seja T uma TII. Se $h_m(T) > 0$, então $H_m(\mathcal{I}) = \infty$.*

Demonstração: Vide [3, Corolário 3.3].

□

Definição 5.1.3 (Sistema aperiódico) *Seja (X, m, T) um sistema dinâmico ergódico. Dizemos que T é aperiódico se é invertível e o conjunto dos pontos periódicos tem m -medida nula.*

O próximo teorema nos diz que estudar sistemas aperiódicos é equivalente a estudar TII's.

Teorema 5.1.4 (Arnoux-Ornstein-Weiss) *Todo sistema dinâmico aperiódico é ergodicamente equivalente a uma TII munida da medida de Lebesgue.*

Demonstração: Vide [2, Teorema 2].

□

Para qualquer $h \in (0, \infty]$, existe um shift de Bernoulli tal que sua entropia métrica é igual a h . Como o conjunto dos pontos periódicos do shift é enumerável e a medida de Bernoulli não tem átomos, eles têm medida nula, segue do Teorema 5.1.4 que existe uma TII que é ergodicamente equivalente a um shift de Bernoulli com entropia h . Uma vez que a entropia é invariante por conjugação, segue que para todo $h \in (0, \infty]$ existe uma TII T em $[0, 1)$ com entropia $h_m(T) = h$.

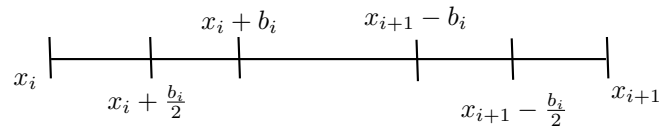
Escreva $I = [0, 1)$, e seja $T : I \rightarrow I$ uma TII. Escreva $I_0 = [x_0, x_1)$ e $I_i = [x_i, x_{i+1})$ para todo $i \geq 1$. Por simplicidade, sejam $l_i = m(I_i)$ e $S = \{x_i\}_{i \geq 0}$. Como T é uma translação por partes, ela é suave em $I \setminus S$.

Agora, vamos construir uma função $r : I \rightarrow \mathbb{R}^+ \cup \{+\infty\}$ tal que $r|_{I \setminus S}$ é suave e $\lim_{x \rightarrow x_i} r(x) = +\infty$ para todo $i \geq 0$. Ademais, essa função será constante e igual 1 em um conjunto de medida de Lebesgue grande. Para todo $i \geq 0$, tomamos $0 < b_i < l_i/2$, a ser escolhido *a posteriori*. Divida I_i em cinco subintervalos:

$$\begin{aligned} I_{i,1} &= \left[x_i, x_i + \frac{b_i}{2} \right), \quad I_{i,2} = \left[x_i + \frac{b_i}{2}, x_i + b_i \right), \quad I_{i,3} = [x_i + b_i, x_{i+1} - b_i), \\ I_{i,4} &= \left[x_{i+1} - b_i, x_{i+1} - \frac{b_i}{2} \right) \text{ e } I_{i,5} = \left[x_{i+1} - \frac{b_i}{2}, x_{i+1} \right). \end{aligned}$$

Veja a figura abaixo:

Figura 3 – Partição do Intervalo I_i .



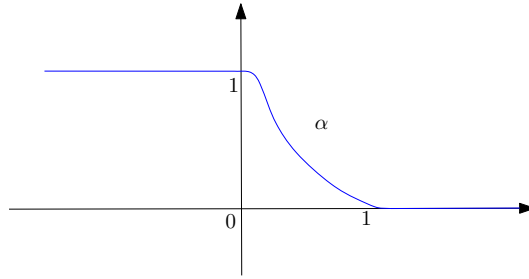
Fonte: Elaborado pelo autor.

Seja $\alpha : \mathbb{R} \rightarrow [0, 1]$ uma função suave tal que $\alpha|_{(-\infty, 0]} \equiv 1$, a restrição $\alpha|_I$ é estritamente decrescente e $\alpha|_{[1, +\infty)} \equiv 0$.

Considere também $\gamma_{i,2} : I_{i,2} \rightarrow [0, 1]$ e $\gamma_{i,4} : I_{i,4} \rightarrow [0, 1]$ definidas por

$$\gamma_{i,2}(x) = \frac{x - (x_i + \frac{b_i}{2})}{\frac{b_i}{2}} \quad \text{e} \quad \gamma_{i,4}(x) = \frac{x - (x_{i+1} - b_i)}{\frac{b_i}{2}}.$$

Figura 4 – Gráfico da função α .



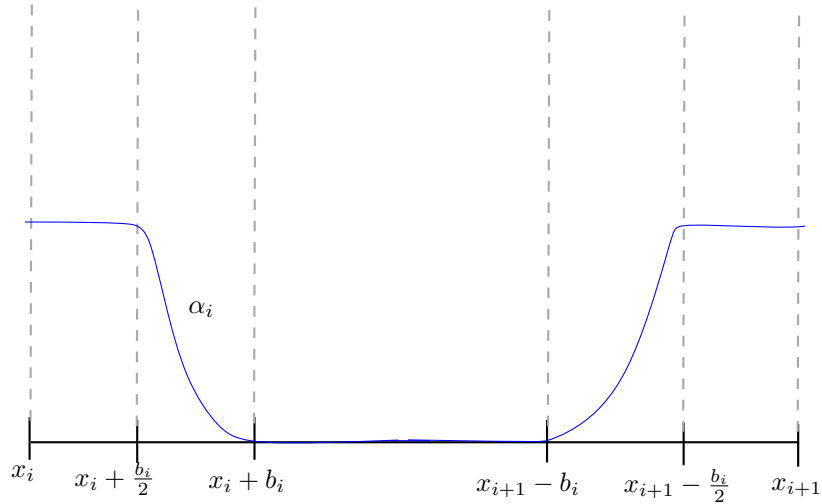
Fonte: Elaborado pelo autor.

Finalmente, considere $\alpha_i : I_i \rightarrow \mathbb{R}$ definida por

$$\alpha_i(x) = \begin{cases} 1, & \text{se } x \in I_{i,1}, \\ (\alpha \circ \gamma_{i,2})(x), & \text{se } x \in I_{i,2}, \\ 0, & \text{se } x \in I_{i,3}, \\ 1 - (\alpha \circ \gamma_{i,4})(x), & \text{se } x \in I_{i,4}, \\ 1, & \text{se } x \in I_{i,5}. \end{cases}$$

Veja que para todo $i \geq 0$, a função α_i é suave por construção, veja seu gráfico abaixo.

Figura 5 – Gráfico da função α_i .



Fonte: Elaborado pelo autor

Defina a função $f_i : I_i \rightarrow \mathbb{R}^+$ por

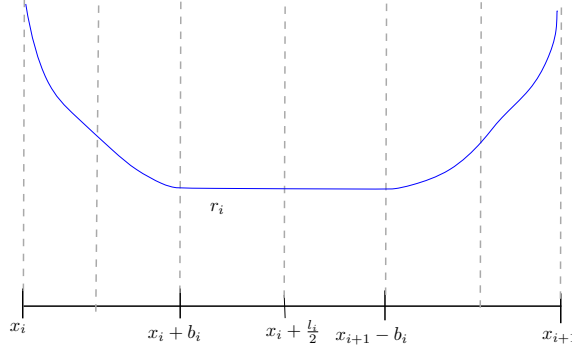
$$f_i(x) = \begin{cases} 1 - \log\left(\frac{x - x_i}{b_i}\right), & \text{se } x \in (x_i, x_i + \frac{b_i}{2}], \\ 1 - \log\left(\frac{x_{i+1} - x}{b_i}\right), & \text{se } x \in [x_{i+1} - \frac{b_i}{2}, x_{i+1}). \end{cases} \quad (5.1)$$

A função f_i é suave em $(x_i, x_i + \frac{l_i}{2})$ e $(x_{i+1} - \frac{l_i}{2}, x_{i+1})$. Defina agora a função $r_i : I_i \rightarrow \mathbb{R}$ por

$$r_i(x) = \alpha_i(x)f_i(x) + (1 - \alpha_i(x))$$

Tal função é suave em I_i . De fato, o único ponto em que ela poderia não ser suave é

Figura 6 – Gráfico da Função r_i .



Fonte: Elaborado pelo autor.

$x = x_i + \frac{l_i}{2} \in I_{i,3}$. Note também que r_i é igual a 1 em $I_{i,3}$, concluímos que r_i é suave. Veja também que $r_i = f_i$ em $I_{i,1} \cup I_{i,5}$.

Finalmente, definimos $r : [0, 1) \rightarrow R^+$ por

$$r(x) = \begin{cases} r_i(x), & \text{se } x \in I_i, \\ +\infty, & \text{caso contrário.} \end{cases} \quad (5.2)$$

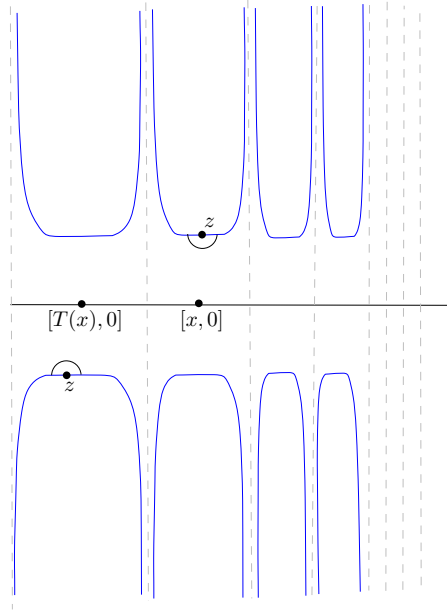
Desse modo, vamos considerar T e r como na Definição 5.1.1 e (5.2), respectivamente. Definimos o espaço topológico $M = M(T, r)$ igual ao quociente $I \times \mathbb{R} / \sim$ com a topologia induzida de $I \times \mathbb{R}$, onde a relação de equivalência $\sim$ é definida por $(x, r(x)) \sim (T(x), -r(x))$.

Chame $\pi : I \times \mathbb{R} \rightarrow M$ a projeção canônica definida por essa relação. Por simplicidade, vamos denotar por $[x, y]$ a imagem do ponto $(x, y) \in I \times \mathbb{R}$ por π . No que segue, vamos mostrar primeiro que M admite uma estrutura suave não-compacta, e em seguida que M possui uma métrica Riemanniana. Seja M^* o subconjunto de M definido por

$$M^* = \{[x, y] \in M : x \in (0, 1), -r(T^{-1}(x)) < y < r(x)\}$$

e seja F o subconjunto de M definido por

$$F = \{[x, r(x)] \in M : x \in I \setminus S\}.$$

Figura 7 – Espaço topológico M .

Fonte: Elaborado pelo autor.

Teorema 5.1.5 *O espaço topológico $M = M(T, r)$ possui uma estrutura suave de variedade.*

Demonstração: Vamos considerar duas famílias de cartas locais de M . Para $z = [x, y] \in M^*$, seja $\varepsilon > 0$ tal que a bola euclidiana centrada em (x, y) e raio ε , definida por $B_\varepsilon((x, y))$, está contida em $\pi^{-1}(M^*)$. A carta ao redor de z é definida pela inversa $\psi_z^e := \pi^{-1} : \pi(B_\varepsilon((x, y))) \rightarrow B_\varepsilon((x, y))$, a qual chamaremos de **carta de Tipo 1**.

Agora, se $z = [x, y] \in F$ a definição da carta ao redor de z é um pouco mais delicada. Considere $\tilde{x} \in I_j$, para algum $j \geq 0$. Escolha números reais $0 < \varepsilon < \frac{1}{2} \min\{|\tilde{x} - x_j|, |x_{j+1} - \tilde{x}|\}$ e $0 < \eta < \frac{1}{2}$, e defina os seguintes conjuntos:

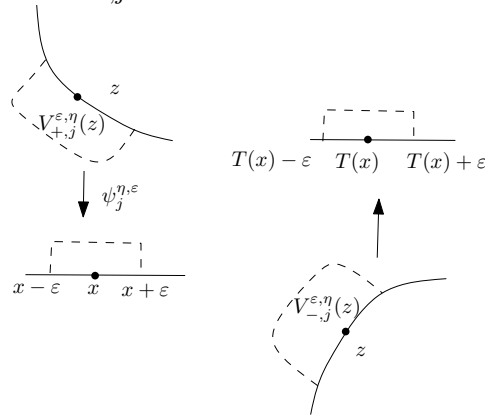
$$\begin{aligned} V_{+,j}^{\varepsilon,\eta}(z) &= \{(x, y) : |x - \tilde{x}| < \varepsilon, \quad r(x) - \eta < y \leq r(x)\} \quad \text{e} \\ V_{-,j}^{\varepsilon,\eta}(z) &= \{(x, y) : |x - T(\tilde{x})| < \varepsilon, \quad -r(x) \leq y < -r(x) + \eta\}. \end{aligned}$$

Veja que o conjunto $V^{\varepsilon,\eta}(z) = \pi(V_{+,j}^{\varepsilon,\eta}(z) \cup V_{-,j}^{\varepsilon,\eta}(z))$ é uma vizinhança aberta de z . Ademais, a aplicação $\psi_z^{\varepsilon,\eta} : V^{\varepsilon,\eta}(z) \rightarrow \mathbb{R}^2$ definida por

$$\psi_z^{\varepsilon,\eta}([x, y]) = \begin{cases} (x, y - r(x)), & \text{se } (x, y) \in V_{+,j}^{\varepsilon,\eta}(z), \\ (x, y + r(x)), & \text{se } (x, y) \in V_{-,j}^{\varepsilon,\eta}(z) \end{cases}$$

é um homeomorfismo, com inversa dada por $(\psi_z^{\varepsilon,\eta})^{-1} : V_{+,j}^{\varepsilon,\eta}(z) \cup V_{-,j}^{\varepsilon,\eta}(z) \rightarrow V^{\varepsilon,\eta}(z)$ dada

Figura 8 – Vizinhas $V_{+,j}^{\varepsilon,\eta}(z)$ e $V_{-,j}^{\varepsilon,\eta}(z)$.



Fonte: Elaborado pelo autor.

por

$$(\psi_z^{\varepsilon,\eta})^{-1}(x,y) = \begin{cases} [x, y + r(x)], & \text{se } (x,y) \in V_{+,j}^{\varepsilon,\eta}(z) \\ [x, y - r(x)], & \text{se } (x,y) \in V_{-,j}^{\varepsilon,\eta}(z). \end{cases}$$

Chamaremos $\psi_z^{\varepsilon,\eta}$ de **carta de Tipo 2**. Agora vamos mostrar que as transições entre cartas são suaves. Quando temos uma transição entre cartas do Tipo 1 ou cartas do Tipo 2, a transição é simplesmente a identidade em $\mathbb{R}^2$, e, portanto, uma aplicação suave. Assim, considere ψ_1 uma carta de Tipo 1 e ψ_2 uma carta do Tipo 2.

A transição $\psi_2 \circ \psi_1^{-1}$ será da forma $(x,y) \mapsto (x, y - r(x))$ ou $(x,y) \mapsto (T(x), y - r(x))$, dependendo apenas do domínio $V_{+,j}^{\varepsilon,\eta}(z)$ ou $V_{-,j}^{\varepsilon,\eta}(z)$ que contém (x,y) . Se considerarmos a transição $\psi_1 \circ \psi_2^{-1}$, teremos as formas $(x,y) \mapsto (x, y + r(x))$ ou $(x,y) \mapsto (T^{-1}(x), y - r(T^{-1}(x)))$, dependendo apenas dos domínios como anteriormente. Em ambos os casos, as transições são suaves, pois T e r são suaves em seus respectivos domínios.

□

Observe que M^* é a imagem de $N = \{(x,y) \in \mathbb{R}^2 : x \in (0,1), -r(T^{-1}(x)) < y < r(x)\}$ pela projeção canônica π . Como $\pi|_N$ é um difeomorfismo, a métrica euclidiana $\tilde{g}^e$ em $\mathbb{R}^2$ induz de modo canônico uma métrica Riemanniana $g^e := (\pi|_N^{-1})^* \tilde{g}^e$ em M^* . A métrica g^e não pode ser estendida para toda a variedade M , pois nas cartas ao redor de pontos de F , onde r não é localmente constante, a métrica euclidiana sofre distorção.

Definição 5.1.6 *Seja M uma variedade suave. Sejam g^1 e g^2 duas métricas Riemannianas em M . Dizemos que g^1 e g^2 são pontualmente equivalentes se para todo $p \in M$ existe uma*

constante $C(p) \geq 1$ tal que para todo $v \in T_p M$ tem-se

$$C(p)^{-1} \leq \frac{g_p^1(v, v)}{g_p^2(v, v)} \leq C(p).$$

Proposição 5.1.7 *A variedade suave $M = M(T, r)$ admite uma métrica Riemanniana g , que é pontualmente equivalente à métrica g^e quando restrita ao conjunto M^* .*

Demonstração: Seja $z \in F$, escolha $\varepsilon > 0$ e escolha $0 < \delta < \frac{1}{2}$ tal que a carta de Tipo 2 ao redor de z esteja bem definida. Defina a métrica h^δ em $V^{\varepsilon, \delta}$ dada por $(\psi_z^{\varepsilon, \delta})^* \tilde{g}^e$. Como a transição entre duas cartas de Tipo 2 é uma translação, segue que h^δ está bem definido. Considere o seguinte conjunto:

$$R^\delta = \{[x, y] \in M : x \in I \setminus S, -r(T^{-1}(x)) < y < -r(T^{-1}(x)) + \delta \text{ ou } r(x) - \delta < y \leq r(x)\} \quad (5.3)$$

Note que $R^\delta \subset V^{\varepsilon, \delta}(w)$ para algum $w \in F$. Considere a função suave $\rho_\delta : M \rightarrow [0, 1]$ tal que $\rho_\delta|_{M \setminus R^\delta} \equiv 1$, $\rho_\delta|_F \equiv 0$ e $0 < \rho_\delta < 1$ em qualquer outro caso. Desse modo, a métrica Riemanniana definida por $g^\delta = \rho_\delta g^e + (1 - \rho_\delta) h^\delta$ coincide com g^e em $M \setminus R^\delta$.

Afirmção: As métricas g^e e g^δ são pontualmente equivalentes em M^* .

De fato, seja $z = [x, y] \in M^*$. Vamos denotar por $\|\cdot\|_z^\delta$ respectivamente $\|\cdot\|_z^e$ a norma induzida por g^δ (respectivamente g^e) em $T_z M$. Sejam $\tilde{z} \in F$ e $\varepsilon, \delta > 0$ tais que $\psi_{\tilde{z}}^{\varepsilon, \delta}$ está bem definido. Seja $\|d_z \psi_{\tilde{z}}^{\varepsilon, \delta}\|$ a norma do operador linear

$$d_z \psi_{\tilde{z}}^{\varepsilon, \delta} : (T_z M, g_z^e) \rightarrow (\mathbb{R}^2, g^e).$$

Então, se $z \in V_{+, j}^{\varepsilon, \delta}(\tilde{z})$ em coordenadas locais temos

$$d_z \psi_{\tilde{z}}^{\varepsilon, \delta} = \begin{pmatrix} 1 & 0 \\ -r'(x) & 1 \end{pmatrix} \quad \text{e} \quad (d_z \psi_{\tilde{z}}^{\varepsilon, \delta})^{-1} = \begin{pmatrix} 1 & 0 \\ r'(x) & 1 \end{pmatrix}.$$

Por outro lado, se $z \in V_{-, j}^{\varepsilon, \delta}(\tilde{z})$, teremos em coordenadas locais

$$d_z \psi_{\tilde{z}}^{\varepsilon, \delta} = \begin{pmatrix} 1 & 0 \\ r'(T^{-1}(x)) & 1 \end{pmatrix} \quad \text{e} \quad (d_z \psi_{\tilde{z}}^{\varepsilon, \delta})^{-1} = \begin{pmatrix} 1 & 0 \\ -r'(T^{-1}(x)) & 1 \end{pmatrix}.$$

Vamos limitar $\|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}\|$. Para tal, podemos supor, sem perda de generalidade, que $z \in V_{+,j}^{\varepsilon,\delta}(\tilde{z})$. Temos $(d_z\psi_{\tilde{z}}^{\varepsilon,\delta})(d_z\psi_{\tilde{z}}^{\varepsilon,\delta})^*$, logo:

$$\begin{aligned} (d_z\psi_{\tilde{z}}^{\varepsilon,\delta})(d_z\psi_{\tilde{z}}^{\varepsilon,\delta})^* &= \begin{pmatrix} 1 & 0 \\ -r'(x) & 1 \end{pmatrix} \begin{pmatrix} 1 & -r'(x) \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 1 & -r'(x) \\ -r'(x) & 1 + (r'(x))^2 \end{pmatrix}, \end{aligned}$$

cuja equação característica é

$$\det \begin{pmatrix} 1 - \lambda & -r'(x) \\ -r'(x) & 1 + (r'(x))^2 - \lambda \end{pmatrix} = 0 \longrightarrow 1 - (2 + (r'(x))^2)\lambda + \lambda^2 = 0,$$

donde encontramos o maior autovalor $\lambda = \frac{2 + (r'(x))^2 + \sqrt{(r'(x))^4 + 4(r'(x))^2}}{2}$. Por um lado, $\lambda \geq 1$ o que implica $\|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}\| = \sqrt{\lambda} \geq 1$. Por outro,

$$\begin{aligned} \lambda &= \frac{2 + (r'(x))^2 + \sqrt{(r'(x))^4 + 4(r'(x))^2}}{2} \leq \frac{2(1 + (r'(x))^2 + |r'(x)|)}{2} \\ &= 1 + (r'(x))^2 + |r'(x)| \\ &\leq (1 + |r'(x)|)^2 \\ &\leq 4(1 + |r'(x)|)^2, \end{aligned}$$

logo $\|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}\| \leq 2(1 + |r'(x)|)$. Usando o mesmo argumento para $\|(d_z\psi_{\tilde{z}}^{\varepsilon,\delta})^{-1}\|$, concluimos que $1 \leq \|(d_z\psi_{\tilde{z}}^{\varepsilon,\delta})^{-1}\| \leq 2(1 + |r'(x)|)$. Assim, para todo $v \in T_z M$ temos

$$\begin{aligned} (\|v\|_z^\delta)^2 &= g_z^\delta(v, v) = \rho_\delta(z)g_z^e(v, v) + (1 - \rho_\delta(z))h_z^\delta(v, v) \\ &= \rho_\delta(z)g_z^e(v, v) + (1 - \rho_\delta(z))g_z^e(d_z\psi_{\tilde{z}}^{\varepsilon,\delta}(v), d_z\psi_{\tilde{z}}^{\varepsilon,\delta}(v)) \\ &\leq \rho_\delta(z)g_z^e(v, v) + (1 - \rho_\delta(z))\|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}\|^2 g_z^e(v, v) \\ &\leq \rho_\delta(z)\|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}\|^2 g_z^e(v, v) + (1 - \rho_\delta(z))\|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}\|^2 g_z^e(v, v) \\ &= \|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}\|^2 g_z^e(v, v) = \|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}\|^2 (\|v\|_z^e)^2. \end{aligned}$$

Por outro lado,

$$\begin{aligned} (\|v\|_z^e)^2 &= (\|(d_z\psi_{\tilde{z}}^{\varepsilon,\delta})^{-1} \cdot d_z\psi_{\tilde{z}}^{\varepsilon,\delta}(v)\|_z^e)^2 \\ &\leq \|(d_z\psi_{\tilde{z}}^{\varepsilon,\delta})^{-1}\|^2 (\|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}(v)\|_z^e)^2 \\ &= \|(d_z\psi_{\tilde{z}}^{\varepsilon,\delta})^{-1}\|^2 ((1 - \rho_\delta(z))\|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}(v)\|_z^e + \rho_\delta(z)\|d_z\psi_{\tilde{z}}^{\varepsilon,\delta}(v)\|_z^e)^2 \\ &\leq (2 + 2|r'(x)|)^2 ((1 - \rho_\delta(z))(2 + 2|r'(x)|)\|v\|_z^e + \rho_\delta(z)(2 + 2|r'(x)|)\|v\|_z^e)^2 \\ &= (2 + 2|r'(x)|)^4 (\|v\|_z^\delta)^2. \end{aligned}$$

Para $z = [x, y] \in M^*$, defina $C(z) = (2 + 2r'(x))^2$. Logo, para todo $v \in T_z M$ temos

$$C(z)^{-1} \cdot \|v\|_z^e \leq \|v\|_z^\delta \leq C(z) \cdot \|v\|_z^e,$$

o que prova a afirmação. □

5.2 O difeomorfismo suave ϕ

Agora, vamos considerar a variedade Riemanniana (M, g^δ) construída na seção anterior.

Definição 5.2.1 (Fluxo de suspensão sobre M^*) *Um fluxo de suspensão (ϕ^t) em M^* é um fluxo tal que para todo $t \in \mathbb{R}$ a aplicação $\phi^t : M^* \rightarrow M^*$ satisfaz $\phi^t([x, y]) = [x, y + t]$.*

Vamos denotar por $\phi := \phi^1$ a transformação igual ao tempo $t = 1$ do fluxo, veja a Figura 9. Em outras palavras,

$$\phi([x, y]) = \begin{cases} [x, y + 1], & \text{se } y + 1 < r(x) \\ [T(x), y + 1 - 2r(x)], & \text{se } y + 1 \geq r(x). \end{cases}$$

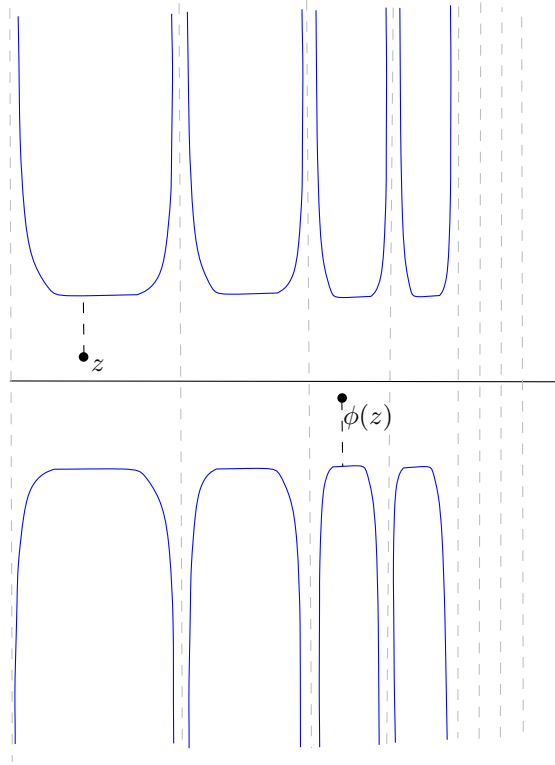
Veja a figura abaixo:

Proposição 5.2.2 *A aplicação $\phi : M \rightarrow M$ é suave.*

Demonstração: Sejam $z = [x, y] \in M$ e ψ_α, ψ_β cartas locais. Vamos mostrar que o mapa $\phi_{\alpha, \beta} := \psi_\alpha \circ \phi \circ \psi_\beta^{-1}$ é suave em z . Se ψ_α e ψ_β são cartas do Tipo 1, então $\phi_{\alpha, \beta}$ é da forma $(x, y) \mapsto (x, y + 1)$ ou $(x, y) \mapsto (T(x), y + 1 - r(x))$, logo $\phi_{\alpha, \beta}$ é suave em z nesse caso. Se ψ_α é uma carta do Tipo 1 e ψ_β do Tipo 2, então $\phi_{\alpha, \beta}$ é igual a $(x, y) \mapsto (T(x), y + 1 - r(x))$, que é suave pois suas coordenadas o são. Quando ψ_α é uma carta de Tipo 2 e ψ_β do Tipo 1, então $\phi_{\alpha, \beta}$ é igual a $(x, y) \mapsto (T(x), y + 1 - r(x))$, que também é suave. Por fim, se ψ_α e ψ_β são cartas de Tipo 2, então $\phi_{\alpha, \beta}$ é igual a $(x, y) \mapsto (x, y + 1)$ ou $(x, y) \mapsto (x, y + 1 - r(x))$. Assim, tem-se que $\phi_{\alpha, \beta}$ é suave, e portanto, ϕ é suave em M . □

Seja $\|d\phi\|^\delta$ a norma do operador em relação a g^δ . Quando $z \in M^* \cap \phi^{-1}(M^*)$, podemos considerar também a norma $\|d_z \phi\|^e$ de $d\phi$ com respeito à norma euclidiana g^e . A próxima proposição nos diz como comparar essas duas normas.

Figura 9 – Tempo 1 do fluxo de suspensão.



Fonte: Elaborado pelo autor.

Proposição 5.2.3 *Existe uma função mensurável $\beta : M^* \cap \phi^{-1}(M^*) \rightarrow \mathbb{R}^+$ tal que*

$$\|d_z \phi\|^\delta \leq \beta(z) \|d_z \phi\|^e \quad (5.4)$$

para todo $z \in M^* \cap \phi^{-1}(M^*)$.

Demonstração: Usando a Proposição 5.1.7, para todos $z \in M^* \cap \phi^{-1}(M^*)$ e $v \in T_z M$, temos

$$\begin{aligned} \|d_z \phi(v)\|_{\phi(z)}^\delta &\leq C(\phi(z)) \|d_z \phi(v)\|_{\phi(z)}^e \\ &\leq C(\phi(z)) \cdot \|d_z \phi\|^e \cdot \|v\|_z^e \\ &\leq C(\phi(z)) \cdot C(z) \cdot \|d_z \phi\|^e \cdot \|v\|_z^\delta \\ &\leq (2 + 2|r'(x)|)^4 \cdot \max\{(2 + 2|r'(x)|)^2, (2 + 2|r'(T(x))|)^2\} \|d_z \phi\|^e \cdot \|v\|_z^\delta. \end{aligned}$$

Lembramos que g^e e g^δ coincide em $M \setminus R^\delta$. Seja K^δ definido por $K = \{[x, y] \in M : -r(T^{-1}(x)) + \delta < y < r(x) - (1 + \delta)\}$. Então podemos tomar β por

$$\beta(z) = (2 + 2|r'(x)|)^4 \cdot \max\{(2 + 2|r'(x)|)^2, (2 + 2|r'(T(x))|)^2\} \mathbb{1}_{M \setminus K^\delta} + \mathbb{1}_{K^\delta}, \quad (5.5)$$

o que termina a proposição.

□

Lembre que para $z = [x, y] \in M^* \cap \phi^{-1}(M^*)$, temos as seguintes representações para $d_z\phi$ em coordenadas:

- Se $y + 1 < r(x)$, temos

$$d_z\phi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

- Se $y + 1 \geq r(x)$ com $x \in I \setminus S$, temos

$$d_z\phi = \begin{pmatrix} 1 & 0 \\ -2r'(x) & 1 \end{pmatrix}.$$

Vamos agora definir uma medida de probabilidade invariante pelo fluxo (ϕ^t) . Para todo conjunto boreliano $A \subset M$, considere a medida $\tilde{\mu}$ dada por

$$\tilde{\mu}(A) = \int_0^1 \int_{-r(T^{-1}(x))}^{r(x)} \mathbb{1}_{\pi^{-1}(A)}(x, y) dy dx.$$

Como a medida de Lebesgue m é invariante por translações e ϕ^t age por translações em linhas verticais, a medida $\tilde{\mu}$ é ϕ^t -invariante.

Proposição 5.2.4 *A medida $\tilde{\mu}$ é finita.*

Demonstração: É suficiente mostrar que $\int_0^1 r(x) dm$ é finita. Considere r restrito a I_i com $i \geq 0$, assim $r_i(x) = \alpha_i(x)f_i(x) + (1 - \alpha_i(x)) \leq f_i(x)$, onde $f_i(x)$ é dada por (5.1).

Temos as seguintes estimativas:

- Quando $x \in I_{i,1} \cup I_{i,2}$, temos $r_i(x) \leq 2 - \log\left(\frac{x-x_i}{b_i}\right)$.
- Quando $x \in I_{i,4} \cup I_{i,5}$, temos $r_i(x) \leq 2 - \log\left(\frac{x_{i+1}-x}{b_i}\right)$.
- Quando $x \in I_{i,3}$, temos $r_i(x) \equiv 1$.

Desse modo,

$$\begin{aligned} \int_0^1 \int_0^{r(x)} 1 dy dx &= \sum_{i \geq 0} \int_{I_i} r(x) dx \\ &\leq \sum_{i \geq 0} \int_{x_i}^{x_i+b_i} \left(2 - \log \frac{x-x_i}{b_i}\right) dx + \int_{x_{i+1}-b_i}^{x_{i+1}} \left(2 - \log \frac{x_{i+1}-x}{b_i}\right) dx \\ &\quad + \int_{x_i+b_i}^{x_{i+1}-b_i} dx \\ &= \sum_{i \geq 0} 2 \left(\int_0^{b_i} \left(2 - \log \frac{x}{b_i}\right) dx \right) + l_i - 2b_i \\ &= \sum_{i \geq 0} 4b_i + l_i \\ &\leq 5, \end{aligned}$$

o que conclui a prova.

□

Definição 5.2.5 (Medida invariante para o fluxo (ϕ^t)) Definimos a medida de probabilidade μ em M como sendo a normalização da medida $\tilde{\mu}$. Essa medida é ϕ -invariante.

5.3 Contra-exemplo para a Desigualdade de Ruelle

Nesta seção, vamos dar um exemplo de um difeomorfismo em uma variedade **não-compacta** que não satisfaz a desigualdade (4.1). Dessa forma, concluímos que a condição de compacidade é necessária para a Desigualdade de Ruelle ser verificada.

Teorema 5.3.1 (Riquelme) Para todo $h \in (0, \infty)$, existem uma variedade Riemanniana não compacta M , um difeomorfismo $f : M \rightarrow M$ de classe C^∞ e uma medida de probabilidade μ em M que é f -invariante tal que $h_\mu(f) = h$ e para μ -quase todo ponto de M os expoentes de Lyapunov são iguais a zero. Em outras palavras:

$$0 = \int \sum_{\lambda_i(x) > 0} m_i(x) \lambda_i(x) d\mu < h_\mu(f) = h < \infty. \quad (5.6)$$

Para provar o Teorema 5.3.1 vamos precisar de alguns lemas. Para tal, considere uma TII (I, m, T) com entropia $h \in (0, \infty]$. Considere também a variedade Riemanniana (M, g^δ) construída na Seção 5.1. Sejam também a aplicação C^∞ dada por $\phi = \phi^1$ e a medida de probabilidade μ construída na seção anterior, construída na Seção 5.2, que é ϕ -invariante.

Lema 5.3.2 Defina a função

$$h(x) = \begin{cases} 2 + 2|r'(x)|, & \text{se } x \in I \setminus S \\ 0, & \text{se caso contrário.} \end{cases}$$

Se $-\sum_{i \geq 0} b_i \log b_i < \infty$, então $\log^+(h)$ é m -integrável.

Demonstração: É suficiente provar que $x \mapsto \log(1 + |r'(x)|)$ é m -integrável em I , pois

$$\log^+(h(x)) \leq \log 2 + \log(1 + |r'(x)|).$$

No intervalo I_i , temos $r(x) = \alpha_i(x)f_i(x) + (1 - \alpha_i(x))$. Assim,

$$|r'(x)| = |\alpha'_i(x)f_i(x) + \alpha_i(x)f'_i(x) - \alpha'_i(x)| \leq |\alpha'_i(x)||f_i(x) - 1| + |f'_i(x)|,$$

e daí

$$|r'(x)| \leq \begin{cases} \frac{b_i}{x-x_i}, & \text{se } x \in I_{i,1} \\ \frac{C}{b_i}, & \text{se } x \in I_{i,2} \\ 0, & \text{se } x \in I_{i,3} \\ \frac{C}{b_i}, & \text{se } x \in I_{i,4} \\ \frac{b_i}{x_{i+1}-x} & \text{se } x \in I_{i,5}, \end{cases}$$

onde $C \geq 1$ é uma constante que depende apenas de $\sup_{x \in \mathbb{R}} |\alpha'(x)| < \infty$. Em particular,

$$\begin{aligned} \int_0^1 \log^+(1 + |r'(x)|) dx &= \sum_{i \geq 0} \sum_{j=1}^5 \int_{I_{i,j}} \log(1 + |r'(x)|) dx \\ &\leq \sum_{i \geq 0} \int_{I_{i,1}} \log\left(1 + \frac{b_i}{x-x_i}\right) dx \\ &\quad + \int_{I_{i,2}} \log\left(1 + \frac{C}{b_i}\right) dx + \int_{I_{i,3}} \log(1) dx \\ &\quad + \int_{I_{i,4}} \log\left(1 + \frac{C}{b_i}\right) dx + \int_{I_{i,5}} \log\left(1 + \frac{b_i}{x_{i+1}-x}\right) dx \\ &\leq \sum_{i \geq 0} (3 + \log(2C)) l_i - b_i \log(b_i) \\ &\leq 3 + \log(2C) - \sum_{i \geq 0} b_i \log(b_i) \\ &< +\infty, \end{aligned}$$

como queríamos. □

Note que na prova do Lema 5.3.2 é crucial assumir que $-\sum_{i \geq 0} b_i \log b_i < \infty$. Em particular, se a TII possui entropia $h > 0$, segue da Proposição 5.1.2 que $-\sum_{i \geq 0} l_i \log l_i = \infty$, em cujo caso não podemos escolher b_i uniformemente proporcional a l_i de modo que $-\sum_{i \geq 0} b_i \log b_i < \infty$.

Lema 5.3.3 *Se $-\sum_{i \geq 0} b_i \log b_i < \infty$, então*

$$\int \log^+ \|d\phi\|^e d\mu < \infty \text{ e } \int \log^+ \|d\phi^{-1}\|^e d\mu < \infty.$$

Demonstração: Note que $M^* \cap \phi^{-1}(M^*)$ é uma conjunto de medida total com relação a μ , uma vez que $(M^* \cap \phi^{-1}(M^*))^c = (M^*)^c \cup (\phi^{-1}(M^*))^c$ tem medida nula, pois $(M^*)^c =$

$\{[x, y] : y = r(x) \text{ ou } y = r(T^{-1}(x))\}$ possui medida nula e

$$(\phi^{-1}(M^*))^c = \{[x, y] : y + 1 = -r(T^{-1}(x)) \text{ ou } y + 1 = r(x)\}$$

também possui medida nula. Sabemos que a diferencial $d_z\phi$ é a matriz identidade se $z \in \{[x, y] \in M : -r(T^{-1}(x)) < y < r(x) - 1\}$ e a matriz

$$\begin{pmatrix} 1 & 0 \\ -2r'(x) & 1 \end{pmatrix}$$

se $z \in \{[x, y] \in M : r(x) - 1 < y < r(x)\}$. Em particular, $\|d_z\phi\|^e \leq 2 + 2|r'(x)|$, logo

$$\int \log^+ \|d_z\phi\|^e d\tilde{\mu} \leq \int_0^1 \int_{r(x)-1}^{r(x)} \log^+ \|d_z\phi\|^e dx dy \leq \int \log^+(h) dm < +\infty.$$

O Lema 5.3.2 nos diz que a última integral é finita. De modo similar, prova-se a finitude da segunda integral.

□

Como vamos trabalhar com $\|d_z\phi\|^\delta$, precisamos também do seguinte lema.

Lema 5.3.4 *Se $-\sum_{i \geq 0} b_i \log b_i < \infty$, então*

$$\int \log^+ \|d\phi\|^\delta d\mu < \infty \text{ e } \int \log^+ \|d\phi^{-1}\|^\delta d\mu < \infty.$$

Demonstração: Pela Proposição 5.2.3, temos

$$\begin{aligned} \int \log^+ \|d_z\phi\|^\delta d\mu(z) &\leq \int \log^+(\beta(z) \|d_z\phi\|^e) d\mu(z) \\ &\leq \int \log^+ \|d_z\phi\|^e d\mu(z) + \int \log^+ \beta(z) d\mu(z). \end{aligned}$$

Pelo Lema 5.3.3, a primeira integral é finita. Para a segunda integral, temos

$$\begin{aligned} \int \log^+ \beta(z) d\tilde{\mu} &= \int_0^1 \int_{-r(T^{-1}(x))}^{-r(T^{-1}(x))+\delta} \log^+ \beta[x, y] dx dy \\ &\quad + \int_0^1 \int_{r(x)-(1+\delta)}^{r(x)} \log^+ \beta([x, y]) dx dy \\ &\leq (1+2\delta) \int_0^1 4 \log(2 + 2|r'(x)|) + 2 \log(2 + 2|r'(T(x))|) dx \\ &= 6(1+2\delta) \int \log^+(h) dm. \end{aligned}$$

Na última igualdade foi usado o fato de que T preserva a medida de Lebesgue dx em $(0, 1)$.

De maneira similar, prova-se que a segunda integral é finita.

□

Antes de prosseguir com a prova do Teorema 5.3.1, precisamos de uma maneira sucinta de calcular a entropia de ϕ em função da TII e de r . Para isso, usaremos a fórmula de Abramov:

$$h_\mu(\phi) = \frac{h_m(T)}{2 \int r dm}. \quad (5.7)$$

Uma prova desta fórmula pode ser encontrada em [8, p.90]. Ao longo da prova, também faremos uso do seguinte teorema.

Teorema 5.3.5 (Aaronson) *Seja $(M, \mathcal{B}, m)$ um espaço de Lebesgue, onde m é uma medida de probabilidade. Suponha que $T : M \rightarrow M$ é uma transformação ergódica que preserva a medida m . Se $h : M \rightarrow \mathbb{R}$ é uma função mensurável tal que $\log^+(|h|)dm < \infty$, então*

$$\lim_n \frac{1}{n} \log^+ \left| \sum_{i=0}^{n-1} h(T^i(x)) \right| = 0$$

para m -quase todo ponto $x \in M$.

Demonstração: Vide [1, Teorema 2.3.1].

□

Observação 5.3.1 *Se $h_m(T) = \infty$, como a função r é m -integrável, segue da fórmula de Abramov que a entropia do fluxo de suspensão $h_\mu(\phi)$ é infinita. Escolhendo os b_i 's de modo que $-\sum_{i \geq 0} b_i \log b_i < \infty$, o Lema 5.3.4 nos permite aplicar o Teorema de Oseledets, e em particular*

$$\int \sum_{i: \lambda_i(x) > 0} m_i(x) \lambda_i(x) d\mu < \infty.$$

Veja que a entropia de ϕ com respeito a μ é maior que a soma dos expoentes de Lyapunov positivos, o que contradiz a Desigualdade de Ruelle.

Demonstração: (Teorema 5.3.1)

Suponha que $0 < h_m(T) < \infty$. Vamos mostrar que os expoentes de Lyapunov em relação à métrica g^δ são nulos em μ -quase todo ponto. Para tal, primeiro vamos calcular os expoentes de Lyapunov em relação à métrica g^e sobre M^* .

Como M^* não é um conjunto ϕ -invariante, vamos trabalhar no conjunto ϕ -invariante $\bigcap_{k \in \mathbb{Z}} \phi^k(M^*)$, que possui medida total em relação a μ . De fato, note que

$\left(\bigcap_{k \in \mathbb{Z}} \phi^k(M^*)\right)^c = \bigcup_{k \in \mathbb{Z}} (\phi^k(M^*))^c$ tem medida nula, pois é a união enumerável de conjuntos de medida nula. Para calcular os expoentes de Lyapunov em relação g^e , faremos uso do Teorema 5.3.5. Para utilizar esse teorema, precisamos que o sistema $(M, \mathcal{F}, \mu, \phi)$ seja ergódico, o que não é necessariamente verdade. Por outro lado, o fluxo (ϕ^t) é ergódico, caso contrário existe um conjunto invariante por (ϕ^t) com medida em $(0, 1)$; tal conjunto necessariamente é da forma

$$\{[x, y] : x \in A \text{ e } -r(T^{-1}(x)) \leq y \leq r(x)\}$$

e portanto $A \subset [0, 1)$ é invariante pela TII e possui medida $0 < m(A) < 1$, o que gera uma contradição, pois a TII é ergódica. De todo modo, a ergodicidade de (ϕ^t) não implica na ergodicidade de ϕ^t para um t particular. Por outro lado, essa condição implica que existem infinitos $t \in \mathbb{R}$ tais que ϕ^t é ergódico, veja [6, p.220].

Considere τ tal que ϕ^τ é ergódico e sejam $(\lambda_i)_{i \leq k}$ os expoentes de Lyapunov de ϕ . Os expoentes de Lyapunov de ϕ^τ são $(\tau \lambda_i)_{i \leq k}$. De fato, dado $v \in T_z M \setminus \{0\}$ temos

$$\min_{\substack{0 \leq t \leq 1 \\ y \in M}} \{m(d_y \phi^t)\} \cdot \|d_z \phi^{\lfloor \tau n \rfloor}(v)\| \leq \|d_z \phi^{\tau n}(v)\| \leq \max_{\substack{0 \leq t \leq 1 \\ y \in M}} \{\|d_y \phi^t\|\} \cdot \|d_z \phi^{\lfloor \tau n \rfloor}(v)\|.$$

Passando logaritmo e dividindo por n , por um lado temos

$$\begin{aligned} \frac{1}{n} \log \|d_z \phi^{\tau n}\| &\leq \frac{1}{n} \log \left[\max_{\substack{0 \leq t \leq 1 \\ y \in M}} \{\|d_y \phi^t\|\} \cdot \|d_z \phi^{\lfloor \tau n \rfloor}(v)\| \right] \\ &= \frac{\lfloor \tau n \rfloor}{\lfloor \tau n \rfloor \cdot n} \log \left[\max_{\substack{0 \leq t \leq 1 \\ y \in M}} \{\|d_y \phi^t\|\} \cdot \|d_z \phi^{\lfloor \tau n \rfloor}(v)\| \right] \\ &\leq \tau \frac{1}{\lfloor \tau n \rfloor} \log \left[\max_{\substack{0 \leq t \leq 1 \\ y \in M}} \{\|d_y \phi^t\|\} \cdot \|d_z \phi^{\lfloor \tau n \rfloor}(v)\| \right] \end{aligned} \quad (5.8)$$

e pelo outro temos

$$\tau \frac{(n-1)}{n} \cdot \frac{1}{\lfloor \tau n \rfloor} \log \left[\min_{\substack{0 \leq t \leq 1 \\ y \in M}} \{m(d_y \phi^t)\} \cdot \|d_z \phi^{\lfloor \tau n \rfloor}(v)\| \right] \leq \frac{1}{n} \log \|d_z \phi^{\tau n}(v)\|. \quad (5.9)$$

Fazendo $n \rightarrow \infty$ em (5.8) e (5.9), obtemos o afirmado.

Em particular, se a soma dos expoentes de Lyapunov de ϕ^τ é zero, então o mesmo acontece com a soma dos expoentes de Lyapunov de ϕ . Desse modo, podemos supor, sem perda de generalidade, que ϕ é uma mapa ergódico.

Para todo $i \geq 0$ escolha as constantes $0 < b_i < l_i/2$ de forma que $-\sum_{i \geq 0} b_i \log b_i < \infty$. Seja $z = [x, y] \in \bigcap_{k \in \mathbb{Z}} \phi^k(M^*)$. Para todo $n \geq 0$, tome $k(n)$ de forma que $\phi^n(z) =$

$[T^{k(n)}(x), y']$ para algum $y' \in (-r(T^{k(n)-1}(x)), r(T^{k(n)}(x)))$. Então

$$d_z \phi^n = \begin{pmatrix} 1 & 0 \\ -2 \sum_{j=0}^{k(n)} r'(T^{k(n)}(x)) & 1 \end{pmatrix},$$

daí

$$\|d_z \phi^n\|^e \leq 2 + 2 \sum_{j=0}^{k(n)} |r'(T^{k(n)}(x))| \leq \sum_{j=0}^{k(n)} (2 + 2|r'(T^{k(n)}(x))|) \leq \sum_{j=0}^{n-1} h(T^j(x)),$$

onde h é definida no Lema 5.3.2. Usando o Teorema 5.3.5, temos

$$\lim_n \frac{1}{n} \log \|d_z \phi^n\|^e \leq \lim_n \frac{1}{n} \log \left(\sum_{j=0}^{n-1} h(T^j(x)) \right) = 0.$$

Em particular, os expoentes de Lyapunov positivos de ϕ com respeito à métrica g^e são nulos para μ -quase todo ponto de M , pois para todo $v \in T_z M$ temos

$$\lim_n \frac{1}{n} \log \|d_z \phi^n(v)\|^e \leq \lim_n \frac{1}{n} \log (\|d_z \phi^n\|^e \|v\|^e) = \lim_n \frac{1}{n} \log \|d_z \phi^n\|^e = 0.$$

O mesmo argumento, aplicado a ϕ^{-1} , implica que os expoentes de Lyapunov negativos de ϕ são nulos em μ -quase todo ponto de M com respeito a g^e .

Dados $z \in M$ e $v \in T_z M$, seja $\lambda^\delta(v)$ o expoente de Lyapunov de v com respeito à métrica g^δ . Se $z \in \bigcap_{k \in \mathbb{Z}} \phi^k(M^*)$, pela Proposição 5.2.3 temos

$$\begin{aligned} \lambda^\delta(v) &= \lim_n \frac{1}{n} \log \|d_z \phi^n(v)\|_{\phi^n(z)}^\delta \\ &\leq \lim_n \frac{1}{n} \log \|d_z \phi^n\|^\delta \|v\|_z^\delta \\ &= \lim_n \frac{1}{n} \log \|d_z \phi^n\|^\delta \\ &\leq \lim_n \frac{1}{n} \log (2 + 2|r'(x)|)^4 \\ &\quad + \lim_n \frac{1}{n} \log [\max\{(2 + 2|r'(x)|)^2, (2 + 2|r'(T^{k(n)}(x))|)^2\}] \\ &\quad + \lim_n \frac{1}{n} \log \|d_z \phi^n\|^e \\ &\leq \lim_n \frac{1}{n} \log h(T^{k(n)}(x)). \end{aligned}$$

Como $\log^+ h$ é m -integrável, o Corolário 3.2.6 implica que

$$0 \leq \lim_n \frac{1}{n} \log h(T^{k(n)}(x)) = \lim_n \frac{k(n)}{n} \frac{1}{k(n)} \log h(T^{k(n)}(x)) \leq \lim_n \frac{1}{k(n)} \log h(T^{k(n)}(x)) = 0$$

para m -quase todo ponto $x \in I$. Portanto, os expoentes de Lyapunov positivos de ϕ com respeito a métrica g^δ são nulos em μ -quase todo ponto. Utilizando o mesmo argumento para ϕ^{-1} , tem-se a mesma conclusão para os expoentes de Lyapunov negativos.

Usando a fórmula de Abramov em 5.7, segue que

$$h_\mu(\phi) = \frac{h_m(T)}{2 \int r dm}. \quad (5.10)$$

Usando o fato $h_\mu(\phi^s) = |s|h_\mu(\phi)$ para todo $s \in \mathbb{R}$ (isso também é devido a Abramov), temos

$$h_\mu(f) = h, \quad (5.11)$$

onde $f = \phi^2 \int r dm$. Como os expoentes de Lyapunov de f são todos nulos para μ -quase todo ponto de M , o Teorema 5.3.1 está provado.

□

REFERÊNCIAS

- [1] AARONSON, J. **An introduction to infinite ergodic theory**. Mathematical Surveys and Monographs, Providence, RI, v. 50, 1997.
- [2] ARNOUX, P; ORNSTEIN, D. S.; WEISS, B. Cutting and stacking, interval exchanges and geometric models. **Israel J. Math.**, Israel, v. 50, n. 1-2, p. 160–168, 1985.
- [3] BLUME, F. An entropy estimate for infinite interval exchange transformations. **Math. Z.**, Germany, v. 272, n. 1-2, p. 17–29, 2012.
- [4] BRIN, M.; KATOK, A. On local entropy. *In*: PALLIS, J. (ed.). **Geometric dynamics**: Rio de Janeiro, Brazil, 1981. Berlin: Springer-Verlag, 1983. (Lecture Notes in Mathematics, v. 1007). p. 30–38.
- [5] HASSELBLATT, B.; KATOK, A. **Introduction to the modern theory of dynamical systems**. [S. n.]: Cambridge, 1995.
- [6] LOSERT, V.; SCHMIDT, K. A class of probability measures on groups arising from some problems in ergodic theory, Probability measures on groups (Proc. Fifth Conf., Oberwolfach, 1978). **Lecture Notes in Math.**, Berlin, v. 706, p. 220–238, 1979.
- [7] OLIVEIRA, K.; VIANA, M. **Fundamentos da teoria ergódica**. 2 ed. Rio de Janeiro: IMPA, 2019.
- [8] PARRY, P.; POLLICOTT, M. **Zeta functions and the periodic orbit structure of hyperbolic Dynamics**. Montrouge: Société mathématique de France, 1990. (Astérisque, v. 187-188).
- [9] RIQUELME, F. Counterexamples to Ruelle’s inequality in the noncompact case. **Annales de l’Institut Fourier**, v. 67, n. 1, p.23-41, 2017.
- [10] SARIG, O. **Lecture notes on ergodic theory**, [S.l.] Apr.3, 2023. Disponível em: <https://www.weizmann.ac.il/math/sarigo/sites/math.sarigo/files/uploads/ergodicnotes.pdf> Acesso em: 20 abr. 2023..
- [11] VIANA, M. **Lectures on lyapunov exponents**, 1st ed. Reino Unido: Cambridge University Press, 2014.
- [12] WALTERS, P. **An introduction to ergodic theory** . New York: Springer, 2000.