UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE MATEMÁTICA

WAGNER RANTER GOUVEIA DA SILVA

Existência e Unicidade de Medidas Maximizantes para Classes Robustas de Difeomorfismos Locais

Maceió
2012

WAGNER RANTER GOUVEIA DA SILVA

Existência e Unicidade de Medidas Maximizantes para Classes Robustas de Difeomorfismos Locais

Dissertação de Mestrado na área de concentração de Sistemas Dinâmicos submetida em 25 de Março de 2013 à Banca Examinadora, designada pelo Colegiado do Programa de Pós-Graduação em Matemática da Universidade Federal de Alagoas, como parte dos requisitos necessários à obtenção do grau de mestre em Matemática.

Orientador: Prof. Dr. Krerley Irraciel Martins de Oliveira

Maceió

2012

Minha família.

AGRADECIMENTOS

Ao meu Deus pela força de nunca desistir e pelas oportunidades a oferecida:

À minha família, em especial aos minha mãe Mª de Fátima de Gouveia por tudo, ao meu irmão Waldemar Gouveia (Gaspar) por todos os momentos juntos

Ao meu orientador Prof. Dr. Krerley Oliveira pelo apoio, conselhos e muitas conversas sobre matemática.

Aos Professor Walter Huaraca pela conversas sobre matemática, trabalho, festas, namoradas, enfim, sobre minha vida. Agradeço, também, pelos conselhos, pela orientação na graduação a qual ajudou-me a trilhar o mestrado com menos dificuldade e, por fim, agradeço a amizade.

Aos meus amigos de turma por todos os momentos juntos - sejam estudando ou jogando boliche. Em especial ao Rafael Alvarez, Rafael Nobrega, Davi Lima, pelas conversas e discussões que me ajudaram neste trabalho, ao Abrão Mendes pela amizade e solidariedade, aos Felipe Leandro, Allan George, Marcos Ranieri pelas "resenha".

Aos funcionários do IM, em especial, Sinvaldo, Dona Marieleide (Minha Patroa), Dona Maria e Seu Manuel pelas brincadeiras e conversas descontraídas.

À minha namorada Brigida Villar, pelo companheirismo.

À CAPES REUNE pelo apoio financeiro durante estes dois anos.

RESUMO

Neste trabalho estudamos a prova da existência de medidas que maximizam a entropia métrica para um conjunto aberto de difeomorfismos locais de classe C^1 de uma variedade Riemanniana compacta, conexa.
Nesse contexto, também, foi visto que a entropia topológica coincide com logaritmo do grau. Além disso, temos que as medidas maximizantes são automedidas do operador de Ruelle-Perron-Frobenius.

Quando a aplicação é topologicamente mixing, a medida maximizante é única e positiva em todos os conjuntos abertos.

Palavras Chaves: Medida maximizante, Grau, O Operador de Ruelle-Perron-Frobenius.

ABSTRACT

At this work, we study the proof of the existence of measures that maximaizing the metric entropy for a open set of local diffeomorphisms of classes C^1 of a compact, connected Riemannian manifold.
In this context, we also see that the topological entropy coincides with the logarithm of the degree. Furthermore, we have that the maximizing measures are eigenmeasures of the Ruelle-Perron-Frobenius operator.

When the map is topologically mixing, the maximizing measure is unique and positive on every open set.

Keywords: Maximizing measure, Degree, The Ruelle-Perron-Frobenius Operator.

LISTA DE FIGURAS

SUMÁRIO

INTRODUÇÃO

O objetivo desse trabalho é estudar a prova do seguinte teorema, dada por Krerley Oliveira e Marcelo Viana em (2).

Teorema 0.0.1. *Seja $f : M \to M$ um difeomorfismo local de classe C^1 e grau $p \geq 1$. Se f satisfaz:*

$$\max_{1 \leq k \leq d-1} \max_{x \in M} \limsup_{n \to \infty} \frac{1}{n} \log ||(Df_x^n)^{\wedge_k}|| < \log p. \tag{1}$$

Então $h_{top}(f) = \log p$, e qualquer automedida maximal μ do dual do operador de transferência $\mathcal{L}$ é uma medida maximizante. Em particular, existe alguma medida maximizante para f. Se f é topologicamente mixing então a medida maximizante é única e positiva nos abertos.

Então para tal, precisaremos de algumas noções básicas que dividimos em três capítulos.

No primeiro capítulo, daremos algumas noções básicas e exemplos sobre entropia métrica e entropia topológica que são partes fundamentais do Teorema 3.0.4.

No segundo capítulo, desenvolveremos técnicas que serão úteis na demonstração do Teorema 3.0.4. Na Seção 2.1 mostraremos a existência de automedida maximal para o operador de Ruelle-Perron-Frobenius, depois, mostraremos que tais medidas tem propriedades interessantes. Na Seção 2.2, daremos algumas noções sobre produto exterior, expoentes de Lyapunov e mostraremos um lema que será de fundamental importância para o decorrer do trabalho. Na Seção 2.3, daremos algumas noções de tempos hiperbólicos que serão úteis para mostrarmos a existência de partições geradoras e usaremos o teorema de Rokhlin (Teorema 2.5.3) para termos uma estimativa da entropia topológica.

No terceiro capítulo, reproduziremos a prova dada por Oliveira e Viana do teorema enunciado acima. Primeiro, provaremos a existência de uma medida maximizante e, além disso, que uma automedida maximal também é uma medida maximizante. Depois, se f como no teorema é topologicamente mixing então mostraremos que tal medida maximizante é única.

1 NOÇÕES BÁSICAS

Neste capítulo apresentaremos os conceitos e resultados necessários para o desenvolvimento deste trabalho.

1.1 Medidas invariantes e ergódicas

Seja $(M,\mathcal{B},\nu)$ um espaço de probabilidade. Considere $f: M \to M$ uma transformação mensurável, dizemos que ν é f-invariante se $\nu(f^{-1}(B)) = \nu(B)$ para todo $B \in \mathcal{B}$.

Daremos uma condição suficiente para que exista uma medida invariante.

Teorema 1.1.1 (Existência de medidas invariantes)**.** *Seja $f: M \to M$ uma transformação contínua num espaço métrico compacto. Então existe pelo menos uma medida de probabilidade em M invariante por f.*

O ponto principal na demonstração é considerar uma certa topologia no conjunto $\mathcal{M}_1(M)$ das medidas de probabilidade em M, que chamamos de topologia fraca*. A ideia é que duas medidas são consideradas próximas se as integrais com relação a cada uma dessas duas medidas estejam próximas para (muitas) funções contínuas. Temos que com essa topologia $\mathcal{M}_1(M)$ é compacto, daí construímos uma sequência de medida de modo que seja f-invariante então possui uma subsequência convergente para uma medida f-invariante.

Exemplo 1.1.1. *Seja $f:[0,1] \to [0,1]$, definida por $f(x) = 10x \pmod 1$. A medida de Lebesgue é f-invariante.*

De fato, seja $(a,b) \subset [0,1]$. Temos que $f^{-1}(a,b) = \cup_{k=0}^{9}(a+k/10, b+k/10)$, daí segue que $m(f^{-1}(a,b)) = \sum_{k=0}^{9} m(a+k/10, b+k/10) = m(a,b)$. Logo pela aditividade de m segue que vale para uniões finitas de intervalos. Agora, a família das uniões finitas de intervalos é uma álgebra que gera a σ-álgebra de Borel de $[0,1]$. Logo, pelo Teorema 4.4.8 (Apêndice), temos que m é f-invariante.

Daremos agora a definição de topologicamente mixing.

Definição 1.1.2. Dizemos que $f : M \to M$ é topologicamente mixing se dado qualquer conjunto aberto $U \subset M$ existe $N \in \mathbb{N}$ tal que $f^N(U) = M$.

Dizemos que (f,ν) é um sistema ergódico se dado $B \in \mathcal{B}$ tal que $f^{-1}(B) = B$ então ou $\nu(B) = 0$ ou $\nu(B) = 1$.

1.2 Entropia

Para motivar a definição de entropia de Kolmogorov-Sinai, vamos considerar a seguinte situação básica da Teoria da Informação. Consideremos um canal de comunicação que transmite, sucessivamente, certos símbolos. Esse canal pode ser um telégrafo transmitindo pontos e traços, segundo o antigo código Morse, uma fibra ótima, 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.

Em Teoria da Informação é usual considerar logaritmos na base 2, porque essencialmente todos os canais de informação que encontramos na prática são binários. No entanto, em Teoria Ergódica é mais comum considerar logaritmos naturais (base e), e nós faremos o mesmo. Por definição, a *quantidade de informação* associada a um caracter $a \in \mathcal{A}$ está dada por

$$I(a) = -\log p_a$$

onde p_a é a probabilidade (frequência) do caracter a. A *informação média* associada ao alfabeto $\mathcal{A}$ é dada por

$$I(\mathcal{A}) = \sum_a -p_a I(a) = \sum_a -p_a \log p_a.$$

Mais geralmente, a informação associada a uma palavra $a_1...a_n$ é

$$I(a_1...a_n) = -\log p_{a_1...a_n}$$

onde a probabilidade $p_{a_1...a_n}$ da palavra é, usualmente, maior que o produto $p_{a_1}...p_{a_n}$ das probabilidades das suas letras (vale a igualdade no caso independente). Denotando por $\mathcal{A}^n$ o conjunto de todas as palavras de comprimento n, definimos

$$I(\mathcal{A}^n) = \sum_{a_1,\dots,a_n} -p_{a_1...a_n} \log p_{a_1...a_n}.$$

Finalmente, a *entropia* do canal de comunicação é definida por:

$$I = \lim \frac{1}{n} I(\mathcal{A}^n). \tag{1.1}$$

1.2.1 Entropia de uma partição

Queremos adaptar estas ideias ao nosso contexto em Teoria Ergódica. A principal diferença é que, enquanto em Teoria da Informação o alfabeto $\mathcal{A}$ é discreto(finito), em geral, esse não é necessariamente o caso para o espaço de estados da maioria dos sistemas dinâmicos interessantes. Esse ponto é resolvido fazendo uso de partições, finitas ou enumeráveis, do espaço de estados.

Seja $(M, \mathcal{B}, \nu)$ um espaço de probabilidade. Neste capítulo, uma *partição* é uma família finita ou enumerável $\mathcal{P}$ de subconjuntos de M disjuntos dois-a-dois e cuja a união tem medida total. Denotamos por $\mathcal{P}(x)$ o elemento da partição que contém x. A *soma* $\mathcal{P} \vee \mathcal{Q}$ de duas partições $\mathcal{P}$ e $\mathcal{Q}$ é uma partição cujos elementos são as interseções $P \cap Q$ com $P \in \mathcal{P}$ e $Q \in \mathcal{Q}$. Mais geralmente, dada qualquer família de partições $\mathcal{P}_n$, definimos

$$\bigvee_n \mathcal{P}_n = \{\cap_n P_n : P_n \in \mathcal{P}_n, \forall n \in \mathbb{N}\}.$$

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

$$I_{\mathcal{P}} : M \to \mathbb{R}, \quad I_{\mathcal{P}}(x) = -\log \nu(\mathcal{P}(x)).$$

É claro que a função $I_{\mathcal{P}}$ é mensurável. Então chamamos *entropia, ou informação média*, da partição $\mathcal{P}$ ao número

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

Como é usual na teoria da integral de Lebesgue, fazemos a convenção de que $0 \log 0 = \lim_{x \to 0} x \log x = 0$.

Relacionamos com isto, a seguinte observação será útil em diversas ocasiões. Considere a função $\phi : (0, \infty) \to \mathbb{R}$ da da por $\phi(x) = -x \log x$. Derivando duas vezes vemos que $\phi'' < 0$. Portanto ϕ é côncava:

$$t_1 \phi(x_1) + \cdots + t_k \phi(x_k) \leq \phi(t_1 x_1 + \cdots + t_k x_k) \tag{1.2}$$

para todo $x_1, ..., x_k$ e $t_1, ..., t_k \geq 0$ com $t_1 + \cdots + t_k = 1$. Além disso, a concavidade é estrita: vale a igualdade em (1.2) se, e somente se, $x_1 = \cdots = x_k$.

Dizemos que duas partições $\mathcal{P}$ e $\mathcal{Q}$ são *independentes* se $\nu(P \cap Q) = \nu(P)\nu(Q)$ para todo $P \in \mathcal{P}$ e todo $Q \in \mathcal{Q}$. Nesse caso, $I_{\mathcal{P}\vee\mathcal{Q}} = I_{\mathcal{P}} + I_{\mathcal{Q}}$ e, portanto,

$$H_\nu(\mathcal{P} \vee \mathcal{Q}) = H_\nu(\mathcal{P}) + H_\nu(\mathcal{Q}).$$

Exemplo 1.2.1. *Seja $M = \{1, ..., d\}^{\mathbb{N}}$ munido de uma medida produto $\nu = \eta^{\mathbb{N}}$. Denotamos $p_i = \eta(\{i\})$ para cada $i \in \{1, ..., d\}$. Para cada $n \geq 1$, seja $\mathcal{P}^n$ a partição de M em cilindros $[0; a_1, ..., a_n]$ de comprimento n. A entropia de $\mathcal{P}^n$ é*

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

A última soma é igual a 1, uma vez que $\sum_i p_i = 1$. Portanto,

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

Chamaremos de *entropia condicional* de uma partição $\mathcal{P}$ com relação a uma partição $\mathcal{Q}$ ao número

$$H_\nu(\mathcal{P}/\mathcal{Q}) = \sum_{P\in\mathcal{P}}\sum_{Q\in\mathcal{Q}} -\nu(P\cap Q)\log\frac{\nu(P\cap Q)}{\nu(Q)}. \tag{1.3}$$

Intuitivamente, ele mede a informação adicional fornecida pela partição $\mathcal{P}$ uma vez conhecida a informação da partição $\mathcal{Q}$. É claro que $H_\nu(\mathcal{P}/\mathcal{M}) = H_\nu(\mathcal{P})$ para todo $\mathcal{P}$, onde $\mathcal{M}$ denota a partição trivial $\mathcal{M} = \{M\}$. Além disso, se $\mathcal{P}$ e $\mathcal{Q}$ são independentes então $H_\nu(\mathcal{P}/\mathcal{Q}) = H_\nu(\mathcal{P})$.

Dadas duas partições, $\mathcal{P}$ e $\mathcal{Q}$ dizemos que $\mathcal{P}$ é *menos fina* que $\mathcal{Q}$, e escrevemos $\mathcal{P} \prec \mathcal{Q}$, se todo elemento de $\mathcal{Q}$ está contido em algum elemento de $\mathcal{P}$, a menos de medida nula. A soma $\mathcal{P} \vee \mathcal{Q}$ é, precisamente, a menos fina de todas as partições $\mathcal{R}$ tais que $\mathcal{P} \prec \mathcal{R}$ e $\mathcal{Q} \prec \mathcal{R}$.

Agora daremos algumas propriedades da entropia condicional de uma partição.

Lema 1.2.1. *Sejam $\mathcal{P}, \mathcal{Q}$ e $\mathcal{R}$ partições com entropia finita. Então,*

(a) $H_\nu(\mathcal{P} \vee \mathcal{Q}/\mathcal{R}) = H_\nu(\mathcal{P}/\mathcal{R}) + H_\nu(\mathcal{Q}/\mathcal{P} \vee \mathcal{Q})$;

(b) *se $\mathcal{P} \prec \mathcal{Q}$ então $H_\nu(\mathcal{P}/\mathcal{R}) \leq H_\nu(\mathcal{Q}/\mathcal{R})$ e $H_\nu(\mathcal{R}/\mathcal{P}) \geq H_\nu(\mathcal{R}/\mathcal{Q})$.*

(c) $\mathcal{P} \prec \mathcal{Q}$ *se, e somente se,* $H_\nu(\mathcal{P}/\mathcal{Q}) = 0$.

Observe, em particular, que se tomarmos $\mathcal{P} = \mathcal{M}$ no item (b) do Lema 1.2.1 obtemos que

$$H_\nu(\mathcal{R}/\mathcal{Q}) \leq H_\nu(\mathcal{R}) \quad \text{para quaisquer partições } \mathcal{Q} \text{ e } \mathcal{R}. \tag{1.4}$$

Além disso, tomando $\mathcal{R} = \mathcal{M}$ no item (a), segue que

$$H_\nu(\mathcal{P} \vee \mathcal{Q}) = H_\nu(\mathcal{P}) + H\nu(\mathcal{Q}/\mathcal{P}) \leq H_\nu(\mathcal{P}) + H_\nu(\mathcal{Q}). \tag{1.5}$$

Seja $f : M \to N$ uma transformação mensurável e seja ν uma probabilidade em M. Então $f_*\nu$ é uma probabilidade em N. Além disso, se $\mathcal{P}$ é uma partição de N então $f^{-1}(\mathcal{P}) = \{f^{-1}(P) : P \in \mathcal{P}\}$ é uma partição de M. Por definição:

$$\begin{aligned} H_\nu(f^{-1}(\mathcal{P})) &= \sum_{P \in \mathcal{P}} -\nu(f^{-1}(P)) \log \nu(f^{-1}(p)) \\ &= \sum_{P \in \mathcal{P}} -f_*\nu(P) \log f_*\nu(P) \\ &= H_{f_*\nu}(\mathcal{P}). \end{aligned}$$

Em particular, se $M = N$ e a medida ν é invariante por f então

$$H_\nu(f^{-1}(\mathcal{P})) = H_\nu(\mathcal{P}) \ \text{ para toda partição } \ \mathcal{P}. \tag{1.6}$$

Também precisaremos da seguinte propriedade de continuidade.

Lema 1.2.2. *Dado* $k \geq 1$ *e* $\varepsilon > 0$ *existe* $\delta > 0$ *tal que, para quaisquer partições finitas* $\mathcal{P} = \{P_1, ..., P_k\}$ *e* $\mathcal{Q} = \{Q_1, ..., Q_k\}$,

$$\nu(P_i \Delta Q_i) < \delta \quad \textit{para todo}\, i = 1, ..., k \Rightarrow H_\nu(\mathcal{Q}/\mathcal{P}) < \varepsilon,$$

onde $\nu(P_i \Delta Q_i) = \nu(P_i \backslash Q_i) + \nu(Q_i \backslash P_i)$.

Demonstração. Fixe $\varepsilon > 0$ e $k \geq 1$. Pela continuidade da função $\phi : [0,1] \to \mathbb{R}, \phi(x) = -x \log x$, existe $\rho > 0$ tal que $\phi(x) < \varepsilon / k^2$ para todo $x \in [0, \rho) \cup (1-\rho, 1]$. Tome $\delta = \rho / k$. Dadas duas partições $\mathcal{P}$ e $\mathcal{Q}$ como no enunciado, denote por $\mathcal{R}$ a partição cujos elementos são as interseções $P_i \cap Q_j$ com $i \neq j$ e também o conjunto $\cup_{i=1}^k P_i \cap Q_i$. Note que $\nu(P_i \cap Q_j) \leq \nu(P_i \Delta Q_i) < \delta$ para todo $i \neq j$ e

$$\nu(\bigcup_{i=1}^k P_i \cap Q_i) \geq \sum_{i=1}^k (\nu(P_i) - \nu(P_i \Delta Q_i)) > \sum_{i=1}^k (\nu(P_i) - \delta) = 1 - \rho.$$

Portanto,

$$H_\nu(\mathcal{R}) = \sum_{R\in\mathcal{R}} \phi(\nu(R)) < \#\mathcal{R}\frac{\varepsilon}{k^2}.$$

É claro da definição que $\mathcal{P}\vee\mathcal{Q} = \mathcal{P}\vee\mathcal{R}$. Então, usando (1.4) e (1.5),

$$\begin{aligned} H_\nu(\mathcal{Q}/\mathcal{P}) = H_\nu(\mathcal{P}\vee\mathcal{Q}) - H_\nu(\mathcal{P}) &= H_\nu(\mathcal{P}\vee\mathcal{R}) - H_\nu(\mathcal{P}) \\ = H_\nu(\mathcal{R}/\mathcal{P}) \le H_\nu(\mathcal{R}) < \varepsilon. \end{aligned}$$

□

1.2.2 Entropia de um sistema dinâmico

Seja $f: M \to M$ uma transformação mensurável preservando uma medida de probabilidade ν. A noção de entropia do sistema (f,ν), apresentada a seguir, é inspirado pela ideia de entropia de um canal de comunicação definida por (1.1)

Dada uma partição $\mathcal{P}$ de M com entropia finita, denotamos

$$\mathcal{P}^n = \bigvee_{i=0}^{n-1} f^{-i}(\mathcal{P}) \text{ para cada } n \ge 1.$$

Observe que o elemento $\mathcal{P}^n(x)$ que contém $x\in M$ está dado por:

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

É claro que a sequência $\mathcal{P}^n$ é não-decrescente, ou seja, $\mathcal{P}^n \prec \mathcal{P}^{n+1}$ para todo n. Portanto, a sequência das entropias $H_\nu(\mathcal{P}^n)$ também é não-decrescente.

Daremos uma definição e consequentemente um lema que será útil no desenvolver da teoria.

Definição 1.2.3. A sequência $(a_n)_{n\in\mathbb{N}}$ é dita subaditiva quando $a_{m+n} \le a_m + a_n$.

Lema 1.2.4. *Seja a_n uma sequência subaditiva. Então:* $\dfrac{a_n}{n} \to L = \inf\{\dfrac{a_n}{n}\}$.

Demonstração. Seja $L = \inf\{\frac{a_n}{n}\}$. Dado $\varepsilon > 0, \exists N$ tal que $\frac{a_N}{N} \le L + \varepsilon$. Se $n > N$, escrevemos $n = qN + r$ com $0 \le r \le N-1$.

$$\frac{a_n}{n} = \frac{a_{qN+r}}{qN+r} \le \frac{a_{qN}}{n} + \frac{a_r}{n} \le \frac{qa_N}{n} + \frac{a_r}{n} \le \frac{qa_N}{qN} + \frac{a_r}{n} \le L + \varepsilon + \frac{1}{n}\sup\{a_1,...,a_{N-1}\}.$$

Se n for suficientemente grande, $L \le \dfrac{a_n}{n} \le L + 2\varepsilon$. □

Mostraremos agora que sequência $(H_\nu(\mathcal{P}^n))_{n\geq 1}$ é subaditiva:

Lema 1.2.5. *$H_\nu(\mathcal{P}^{n+m}) \leq H_\nu(\mathcal{P}^n) + H_\nu(\mathcal{P}^m)$ para todo $m,n \geq 1$.*

Demonstração. Por definição, $\mathcal{P}^{m+n} = \vee_{i=0}^{m+n-1} f^{-i}(\mathcal{P}) = \mathcal{P}^m \vee f^{-m}(\mathcal{P}^n)$. Portanto, usando (1.5),

$$H_\nu(\mathcal{P}^{m+n}) \leq H_\nu(\mathcal{P}^m) + H_\nu(f^{-m}(\mathcal{P}^n)). \tag{1.7}$$

Por outro lado, como ν é invariante por f, a propriedade (1.6) implica que

$$H_\nu(f^{-m}(\mathcal{P}^n)) = H_\nu(\mathcal{P}^n)$$

para todo m,n. Substituindo esse fato em (1.7) obtemos a conclusão do lema. □

Chamaremos *entropia de f com relação à medida ν e à partição $\mathcal{P}$* o limite

$$h_\nu(f,\mathcal{P}) = \lim_n \frac{1}{n} H_\nu(\mathcal{P}^n) = \inf_n \frac{1}{n} H_\nu(\mathcal{P}^n). \tag{1.8}$$

Observe que esta entropia é tanto maior quanto mais fina for a partição. De fato, se $\mathcal{P} \prec \mathcal{Q}$ então $\mathcal{P}^n \prec \mathcal{Q}^n$ para todo o n. Usando o Lema 1.2.1, segue que $H_\nu(\mathcal{P}^n) \leq H_\nu(\mathcal{Q}^n)$ para todo n. Consequentemente,

$$\mathcal{P} \prec \mathcal{Q} \Rightarrow h_\nu(f,\mathcal{P}) \leq h_\nu(f,\mathcal{Q}) \tag{1.9}$$

Finalmente, a *entropia* do sistema (f,ν) é definida por

$$h_\nu(f) = \sup_{\mathcal{P}} h_\nu(f,\mathcal{P}), \tag{1.10}$$

onde o supremo é tomado sobre todas as partições com entropia finita. Uma observação útil é que a definição não é afetada se consideramos o supremo apenas sobre as partições finitas.

Exemplo 1.2.2. *Considere a transformação decimal $f : [0,1] \to [o,1]$, dada por $f(x) = 10x(\mod 1)$. Como observamos anteriormente, f preserva a medida de Lebesgue no intervalo m. Seja $\mathcal{P}$ a partição de $[0,1]$ por intervalos da forma $((k-1)/10, k/10]$ com $k = 1,...,10$. Então $\mathcal{P}^n$ é a partição do $[0,1]$ por intervalos da forma $((k-1)/10^n, k/10^n]$ com $k = 1,...,10^n$. Assim, como*

$$H_m(\mathcal{P}^n) = \sum_{k=1}^{10^n} -10^{-n} \log 10^{-n} = n \log 10.$$

Segue-se

$$h_m(f,\mathcal{P}) = \lim_n H_m(\mathcal{P}^n) = \log 10.$$

Usando a teoria que será desenvolvida na seção seguinte, veremos que $h_m(f)$ *também é igual a* $\log 10$*, ou seja,* $\mathcal{P}$ *realiza o supremo na definição (1.10).*

Este lema será útil na prova do teorema de Kolmogorov-Sinai.

Lema 1.2.6. $h_\nu(f,\mathcal{Q}) \leq h_\nu(f,\mathcal{P}) + H_\nu(\mathcal{Q}/\mathcal{P})$ *para quaisquer partições* $\mathcal{P}$ *e* $\mathcal{Q}$ *com entropia finita.*

Demonstração. Pelo Lema 1.2.1, para todo $n \geq 1$ vale que

$$\begin{aligned} H_\nu(\mathcal{Q}^{n+1}/\mathcal{P}^{n+1}) &= H_\nu(\mathcal{Q}^n \vee f^{-n}(\mathcal{Q})/\mathcal{P}^n \vee f^{-n}(\mathcal{P})) \\ &\leq H_\nu(\mathcal{Q}^n/\mathcal{P}^n) + H_\nu(f^{-n}(\mathcal{Q})/f^{-n}(\mathcal{P})). \end{aligned}$$

O último termo é igual a $H_\nu(\mathcal{Q}/\mathcal{P})$, poque a medida ν é f-invariante. Portanto, a relação anterior prova que

$$H_\nu(\mathcal{Q}^n/\mathcal{P}^n) \leq n H_\nu(\mathcal{Q}/\mathcal{P}) \quad \text{para todo } n \geq 1. \tag{1.11}$$

Usando o Lema 1.2.1, uma vez mais, segue que

$$H_\nu(\mathcal{Q}^n) \leq H_\nu(\mathcal{P}^n \vee \mathcal{Q}^n) = H_\nu(\mathcal{P}) + H_\nu(\mathcal{Q}^n/\mathcal{P}^n) \leq H_\nu(\mathcal{P}^n) + n H_\nu(\mathcal{Q}/\mathcal{P}).$$

Dividindo por n e passando o limite quando $n \to \infty$ obtemos a conclusão do lema. □

Lema 1.2.7. $h_\nu(f,\mathcal{P}) = \lim_n H_\nu(\mathcal{P}/\vee_{i=1}^n f^{-i}(\mathcal{P}))$ *para qualquer partição* $\mathcal{P}$ *com entropia finita.*

Demonstração. Usando o Lema 1.7 (a) e o fato de que a medida ν é invariante:

$$\begin{aligned} H_\nu(\bigvee_{i=0}^{n-1} f^{-i}(\mathcal{P})) &= H_\nu(\bigvee_{j=1}^{n-1} f^{-j}(\mathcal{P})) + H_\nu(\mathcal{P}/\bigvee_{j=1}^{n-1} f^{-j}(\mathcal{P})) \\ &= H_\nu(\bigvee_{j=0}^{n-2} f^{-j}(\mathcal{P})) + H_\nu(\mathcal{P}/\bigvee_{j=1}^{n-1} f^{-j}(\mathcal{P})) \end{aligned}$$

para todo n. Por recorrência, segue que

$$H_\nu(\bigvee_{j=0}^{n-1} f^{-j}(\mathcal{P})) = H_\nu(\mathcal{P}) + \sum_{k=1}^{n-1} H_\nu(\mathcal{P}/\bigvee_{j=1}^{k} f^{-j}(\mathcal{P})).$$

Portanto, $h_\nu(f,\mathcal{P})$ é dada pelo limite Cesaro

$$\lim_n \frac{1}{n} H_\nu(\bigvee_{j=0}^{n-1} f^{-j}(\mathcal{P})) = \lim_n \frac{1}{n} \sum_{k=1}^{n-1} H_\nu(\mathcal{P}/\bigvee_{j=1}^{k} f^{-j}(\mathcal{P})).$$

Por outro lado, o item (b) do Lema 1.7, garante que a sequência $H_\nu(\mathcal{P}/\vee_{j=1}^n f^{-j}(\mathcal{P}))$ é decrescente. Em particular, $\lim_n \vee_{j=1}^n f^{-j}(\mathcal{P})$ existe e, consequentemente, coincide com o limite Cesaro na igualmente anterior. $\square$

1.2.3 Teorema de Kolmogorov-Sinai

Em geral, a principal dificuldade no cálculo da entropia reside no cálculo do supremo da definição (1.10). Os métodos que vamos desenvolver nesta seção permitem simplificar a tarefa em muitos casos de interesse, identificamos certas partições $\mathcal{P}$ que realizam o supremo, isto é, tais que $h_\nu(f,\mathcal{P}) = h_\nu(f)$. O resultado principal é o seguinte:

Teorema 1.2.8 (Kolmogorov-Sinai)**.** *Seja $\mathcal{P}_1 \prec \cdots \prec \mathcal{P}_n \prec \cdots$ uma sequência não-decrescente de partições com entropia finita tais que $\cup_{n=1}^\infty \mathcal{P}_n$ gera a σ-álgebra dos conjuntos mensuráveis a menos de medidas nula. Então*

$$h_\nu(f) = \lim_n h_\nu(f,\mathcal{P}_n).$$

Demonstração. O limite sempre existe, pois a propriedade (1.9) implica que a sequência $h_\nu(f,\mathcal{P}^n)$ é não decrescente. Para provar tal teorema usaremos o seguinte fato:

Lema 1.2.9. $\lim_n H_\nu(\mathcal{Q}/\mathcal{P}^n) = 0$ *para qualquer partição finita $\mathcal{Q}$.*

Demonstração. Escreva $\mathcal{Q} = \{Q_1,...,Q_n\}$. Dado qualquer $\varepsilon > 0$, existe $\delta > 0$ como no Lema 1.2.2. Seja $\mathcal{A}$ a álgebra formada pelas uniões finitas de elementos de $\cup_n \mathcal{P}^n$. Por hipótese, $\mathcal{A}$ gera a σ-álgebra de todos os conjuntos mensuráveis. Logo, pelo teorema de aproximação (Apêndice 4.4.7), para cada $i = 1,...,k$ existe $A_i \in \mathcal{A}$ tal que

$$\nu(Q_i \Delta A_i) < \delta/(4k). \tag{1.12}$$

O fato de que os Q_i formam uma cobertura de M garante que os A_i estão perto de serem, também, uma cobertura:

$$\nu(A_i \cap (\cup_{j\neq i} A_j)) \quad \leq \quad \nu(\cup_{j=1}^k (A_j \backslash Q_j)) < \delta/4 \ \text{ para todo } \ i \tag{1.13}$$

e

$$\nu(M \backslash \cup_{j=i}^{k} A_i) \;\; \leq \;\; \nu(\cup_{i=1}^{k}(Q_i \backslash A_i)) < \delta/4. \tag{1.14}$$

A seguir, defina

$$Q_i = \begin{cases} A_1, & \text{para } i = 1 \\ A_i \backslash \cup_{j=1}^{i-1} A_j, & \text{para } 1 < i < k \\ M \backslash \cup_{j=1}^{k-1} A_j, & \text{para } i = k \end{cases}$$

Então $\mathcal{Q}' = \{Q'_1, ..., Q'_k\}$ é uma partição de M. Afirmamos que

$$\nu(A_i \Delta Q'_i) < \delta/2 \;\; \text{para todo} \;\; i = 1, ..., k. \tag{1.15}$$

Isto é trivial para $i = 1$. Para $i > 1$ temos que $A_i \backslash Q'_i$ esta contido em $A_i \cap (\cup_{j<i} A_j)$. Logo, usando 1.13, obtemos que $\nu(A_i \backslash Q'_i) < \delta/4$. Isso prova a afirmação para $1 < i < k$, uma vez que nesse caso $Q'_i \backslash A_i = \emptyset$. Finalmente, para $i = k$, temos que $Q'_k \backslash A_k$ está contido no complementar de $\cup_{i=1}^{k} A_i$. Logo, usando (1.14), vemos que $\nu(Q'_k \backslash A_k) < \delta/4$. Somando essa estimativa com a anterior, vem que $\nu(A_k \Delta Q'_k) < \delta/2$. Isso completa a prova da afirmação (1.15).

Combinando as desigualdades (1.13) e (1.15), obtemos que $\nu(Q_i \Delta Q'_i) < \delta$ para todo $i = 1, ..., k$. Agora, é claro que $Q'_i \in \mathcal{A}$ para todo i. Então, como se trata de uma família finita, podemos encontrar $m \geq q$ tal que todo Q'_i é uma união de elementos de $\mathcal{P}_m$. Em outras palavras, a partição $\mathcal{Q}' = \{Q'_1, ..., Q'_k\}$ é menos fina do que $\mathcal{P}_m$. Então, pelos Lema 1.2.1 e Lema 1.2.2,

$$H_\nu(\mathcal{Q}/\mathcal{P}_n) \leq H_\nu(\mathcal{Q}/\mathcal{P}_m) \leq H_\nu(\mathcal{Q}/\mathcal{Q}') < \varepsilon \;\; \text{para todo} \;\; n \geq m.$$

Isso completa a demonstração do lema.

□

Pelo Lema 1.2.6, também temos que

$$h_\nu(f, \mathcal{Q}) \leq h_\nu(f, \mathcal{P}^n) + H_\nu(\mathcal{Q}/\mathcal{P}^n) \;\; \text{para todo} \;\; n.$$

Passando o limite quando $n \to \infty$ e, então, tomando o supremo sobre todas as partições $\mathcal{Q}$, obtemos a conclusão do teorema. □

Corolário 1.2.9.1. *Seja $\mathcal{P}_1 \prec \cdots \prec \mathcal{P}_n \prec \cdots$ uma sequência de partições finitas tais que* *o* $\text{diam}\mathcal{P}_n(x) \to 0$ *para* ν*-quase todo ponto* $x \in M$*. Então*

$$h_\nu(f) = \lim_n h_\nu(f, \mathcal{P}_n).$$

Demonstração. Seja U um aberto qualquer de M. A hipótese garante que para cada x tal que o conjunto $P_x = \mathcal{P}_{n(x)}(x)$ está contido em U. É claro que P_x pertence à álgebra $\mathcal{A}$ gerada por $\cup_n \mathcal{P}_n$. Observe também que esta álgebra é enumerável, já que ela está formada pelas uniões finitas de elementos das partições $\mathcal{P}_n$. Em particular, o conjunto dos valores tomados por P_x é enumerável. Segue que $U = \cup_{x \in U} P_x$ também está na σ-álgebra gerada por $\mathcal{A}$. Isso prova que σ-álgebra gerada por $\mathcal{A}$ contém todos os abertos e, portanto, contém todos os conjuntos borelianos. Agora, a conclusão segue de uma aplicação direta do Teorema 1.2.8. □

1.3 Fórmula de Rokhlin

O fórmula de Rokhlin é uma ferramenta importante no estudo de sistemas dinâmicos, ela dá uma relação entre entropia métrica do sistema e a integral de uma função especifica chamada de *jacobiano*, quando tal jacobiano existe.

1.3.1 Jacobiano

Agora daremos a noção de *jacobiano* de uma medida. Essencialmente, o jacobiano de uma medida ν (quando existir) é uma função que satisfaz as condições do Teorema de Mudança de variáveis. Vamos a definição precisa.

Definição 1.3.1. Uma função $\xi : M \to \mathbb{R}$ é um jacobiano para a medida ν, se é localmente integrável com respeito à ν e para todo mensurável A tal que $f|_A$ é injetora, então

$$\nu(f(A)) = \int_A \xi d\nu.$$

Em outras palavras, o jacobiano é definido por $\xi = d(\nu(f))/d\nu$.

Exemplo 1.3.1. *Seja* M *um aberto do* $\mathbb{R}^n$ *e* f *é uma transformação de classe* C^1*, o Teorema de Mudança de Variável nos garante que se* $f|_A$ *é uma difeomorfismo, então*

$$m(f(A)) = \int_A |\det Df(x)| dx.$$

Em outras palavras, isso significa que a função $\xi(x) = |\det Df(x)|$ *é um jacobiano para a medida de Lebesgue* m*.*

Definição 1.3.2. Dizemos que uma medida ν é não-singular com respeito à f. Se dado um conjunto A com $\nu(A) = 0$, então $\nu(f(A)) = 0$.

A seguir mostrar que uma medida não-singular para uma transformação f localmente injetora possui um único jacobiano.

Proposição 1.3.3. *Seja M um espaço métrico separável e f uma transformação localmente injetiva. Então, dada uma medida boreliana ν não-singular com respeito à f, existe um único jacobiano $\xi : M \to \mathbb{R}$. Se, além disso, existir $k \in \mathbb{N}$ de modo que cada ponto de M possui no máximo k pré-imagens, então $\xi \in L^1(\nu)$.*

Demonstração. **Existência:** Para cada $p \in M$ podemos encontrar um aberto A_p de modo que $f|_{A_p}$ é injetora. Tomando uma subcobertura de $\{A_p\}_{p\in M}$ se necessário, podemos supor que existe uma subcobertura enumerável $\{A_{p_i}\}_{i\in\mathbb{N}} \subset \{A_p\}_{p\in M}$ tal que $\nu(A_{p_i}) > 0$ para todo $i \in \mathbb{N}$.

Podemos decompor M em subconjuntos mensuráveis $B_i \subset A_{p_i}$ de modo que $\{B_i\}_{i\in\mathbb{N}}$ é uma partição de M. Para tanto, basta definir $B_1 = A_1$ e por indução $B_k = A_k \setminus \cup_{i=1}^{k-1} B_i$. Novamente,renumerando os B_i's se necessário, podemos assumir que $\nu(B_i) > 0$. Assim, cada ponto $x \in M$ admite, no máximo, uma quantidade enumerável de pré-imagens e $f_i = f|_{B_i} : B_i \to f(B_i)$ é uma bijeção.

Fixado $i \in \mathbb{N}$, vamos denotar por ν_i a medida em definida em B_i por $\nu_i(A) = \nu(f_i(A))$. Observe que ν_i é absolutamente contínua com respeito à ν em B_i, já que se $\nu(A) = 0$, então $\nu_i(A) = \nu(f(A)) = 0$. Portanto, pelo teorema de Rodón-Nykodim existe uma função $\xi_i \in L^1(\nu)$ tal que

$$\nu_i(A) = \int_A \xi_i d\nu,$$

Para cada $A \subset B_i$.

Seja ξ a função definida por

$$\xi(x) = \sum_{i\in\mathbb{N}} \xi_i(x)\mathcal{X}_{B_i}(x).$$

Afirmamos que ξ é um jacobiano para ν e f. De fato, se f_A é injetora, podemos escrever A como a união disjunta $A = \cup(A \cap B_i)$, onde segue-se que

$$\nu(f(A)) = \sum_{i\geq 1} \nu_i(A \cap B_i) = \sum_{i\geq 1} \int_{A\cap B_i} \xi_i d\nu = \int_A \xi d\nu,$$

como queríamos demostrar.

Unicidade: Se ξ é um jacobiano para ν e f, pelo Teorema da Derivação 4.4.10(Apêndice), existe um conjunto A de medida ν total tal que se $x \in A$ vale

$$\xi(x) = \lim_{r\to 0} \frac{1}{\nu(B_r(x))} \int_{B_r(x)} \xi(y) d\nu.$$

Por outro lado, assumindo que r é suficientemente pequeno de modo que $f|_{B_r(x)}$ é injetora, temos que

$$\int_{B_r(x)} \xi(y) d\nu = \nu(f(B_r(x))).$$

Portanto, temos que para um conjunto de medida ν total vale

$$\xi(x) = \lim_{r\to 0} \frac{\nu(f(B_r(x))}{\nu(B_r(x))},$$

que mostra que ξ é único.

Assuma agora que existe $k \in \mathbb{N}$ tal que todo ponto tem no máximo k pré-imagens. Para mostrar que $\xi \in L^1(\nu)$, basta observa que

$$\int \xi d\nu = \sum_i \int_{B_i} \xi d\nu = \sum_i \nu(f(B_i)) \leq k\nu(M),$$

já que todo ponto tem no máximo k pré-imagens. □

Iremos denotar o jacobiano de f com respeito à ν por $J_\nu f$. Usando a definição, podemos verificar que $J_\nu f^n(x) = \prod_{i=0}^{n-1} J_\nu f(f^i(x))$ é um jacobiano para f^n.

Teorema 1.3.4 (Fórmula de Rohklin)**.** *Seja $f : M \to M$ uma transformação mensurável localmente invertível num espaço métrico separável e seja ν uma probabilidade invariante por f. Suponha que existe uma partição finita ou enumerável $\mathcal{P}$ tal que cada elemento $P \in \mathcal{P}$ é domínio de injetividade de f e* $\operatorname{diam}\mathcal{P}^n(x) \to 0$ *para ν-quase todo x. Então*

$$h_\nu(f) = \int \log J_\nu f d\nu.$$

Demonstração. Consideremos a sequência da partições $\mathcal{Q}_n = \vee_{j=1}^n f^{-j}(\mathcal{P})$. Pelo Corolário 1.2.9.1 e pelo Lema 1.2.7,

$$h_\nu(f) = h_\nu(f, \mathcal{P}) = \lim_n H_\nu(\mathcal{P}/\mathcal{Q}_n). \tag{1.16}$$

Por definição,

$$\begin{aligned} H_\nu(\mathcal{P}/\mathcal{Q}_n) &= \sum_{P\in\mathcal{P}}\sum_{Q_n\in\mathcal{Q}_n} -\nu(P\cap Q_n)\log\frac{\nu(P\cap Q_n)}{\nu(Q_n)} \\ &= \sum_{P\in\mathcal{P}}\sum_{Q_n\in\mathcal{Q}_n} -\nu(Q_n)\phi\Big(\frac{\nu(P\cap Q_n)}{\nu(Q_n)}\Big). \end{aligned}$$

Seja $e_n(\psi,x)$ a esperança condicional relativamente à partição $\mathcal{Q}_n$, definida na seção 4.6 (Apêndice), e seja $e(\psi,x)$ o seu limite quando $n\to\infty$. É claro da definição que

$$\frac{\nu(P\cap Q_n)}{\nu(Q_n)} = e_n(\mathcal{X}_P,x) \quad \text{para todo} \quad x\in Q_n \quad \text{e todo} \quad Q_n\in\mathcal{Q}_n.$$

Portanto,

$$\sum_{P\in\mathcal{P}}\sum_{Q_n\in\mathcal{Q}_n} -\nu(Q_n)\phi\Big(\frac{\nu(P\cap Q_n)}{\nu(Q_n)}\Big) = \sum_{P\in\mathcal{P}}\int\phi(e_n(\mathcal{X}_P,x))d\nu(x). \tag{1.17}$$

Pelo Lema 4.6.1, o limite $e(\mathcal{X}_P,x)=\lim_n e_n(\mathcal{X}_P,x)$ existe para ν-quase todo ponto $x\in M$. Então, observado que a função ϕ é não negativa limitada, podemos usar o teorema da convergência dominada para deduzir das relações (1.16) e (1.17) que

$$h_\nu(f) = \sum_{P\in\mathcal{P}}\int\phi(e(\mathcal{X}_P,x))d\nu(x). \tag{1.18}$$

Resta relacionar o integrando do lado direito com o jacobiano. Isso será feito por meio do seguinte lema.

Lema 1.3.5. *Para toda função mensurável limitada $\psi: M\to\mathbb{R}$ e ν-quase todo $x\in M$,*

$$e(\psi,x)=\hat\psi(f(x)) \quad \textit{onde} \quad \hat\psi(y)=\sum_{z\in f^{-1}(y)}\frac{\psi}{J_\nu f}(z).$$

Demonstração. Lembre-se que $\mathcal{Q}_n=\vee_{j=1}^n f^{-j}(\mathcal{P})$. Também usaremos a sequência de partições $\mathcal{P}^n=\vee_{j=0}^{n-1}f^{-j}(\mathcal{P})$. Representaremos por $e'_n(\phi,y)$ a esperança condicional relativa à partição $\mathcal{P}^n$ e seja $e(\phi,y)$ o seu limite quando $n\to\infty$. Afirmamos que

$$e_n(\psi,x)=e'_{n-1}(\hat\psi,f(x)) \quad \text{para todo} \quad x \text{e todo} \quad n>1. \tag{1.19}$$

Passando o limite, $e(\psi, x) = e'(\hat{\psi}, f(x))$ para ν-quase todo $x \in M$. Por outro lado, pelo do Corolário 1.2.9.1, a hipótese do teorema implica que a sequência de partições $\mathcal{P}^n$ é geradora. Então pelo item (e) do Teorema 4.5.1, $e'(\hat{\psi}, y) = \hat{\psi}(y)$ para ν-quase todo $y \in M$. Logo, para provar o lema basta provar a afirmação (1.19).

Observe que $\mathcal{Q}_n(x) = f^{-1}(\mathcal{P}^{n-1}(f(x)))$ para todo n e todo x. Por definição de esperança condicional,

$$e_n(\psi, x) = \frac{1}{\nu(\mathcal{Q}_n(x))} \int_{\mathcal{Q}_n(x)} \psi d\nu = \sum_{P \in \mathcal{P}} \frac{1}{\nu(\mathcal{Q}_n(x))} \int_{P \cap \mathcal{Q}_n(x)} \psi d\nu,$$

para toda função mensurável limitada ψ. Por definição de jacobiano e pelo Lema 4.6.2, a expressão do lado direito é igual a

$$\sum_{P \in \mathcal{P}} \frac{1}{\nu(\mathcal{Q}_n(x))} \int_{f(P \cap \mathcal{Q}_n(x))} \frac{\psi}{J_\nu f} \circ (f|_P)^{-1} d\nu = \frac{1}{\nu(\mathcal{Q}_n(x))} \int_{\mathcal{P}^{n-1}(f(x))} \hat{\psi} d\nu.$$

Lembrando que a medida ν é invariante, também segue que a expressão do lado direito é igual a

$$\frac{1}{\nu(\mathcal{P}^{n-1}(f(x)))} \int_{\mathcal{P}^{n-1}(f(x))} \hat{\psi} d\nu = e'_{n-1}(\hat{\psi}, f(x)).$$

Combinando essa três igualdade obtemos (1.19). $\square$

Vamos aplicar esse resultado a $\psi = \mathcal{X}_P$. Como f é injetiva em todo elemento de $\mathcal{P}$, cada interseção $P \cap f^{-1}(y)$ ou é vazio ou contém exatamente um ponto. Portanto, segue do Lema 1.3.5 que $e(\mathcal{X}_P, x) = \hat{\mathcal{X}}_P(f(x))$, com

$$\hat{\mathcal{X}}_P(y) = \begin{cases} 1/J_\nu f((f|_P)^{-1}) & \text{se } y \in f(P) \\ 0 & \text{se } y \notin f(P). \end{cases}$$

Então, como a medida ν é invariante,

$$\begin{aligned} \int \phi(e(\mathcal{X}_P, x)) d\nu(x) &= \int \phi(\hat{\mathcal{X}}_P(y)) d\nu(y) \\ &= \int_{f(P)} (\frac{1}{J_\nu f} \log J_\nu f) \circ (f|_P)^{-1} d_\nu \\ &= \int_P \log J_\nu f d\nu \end{aligned}$$

(a última igualdade segue da definição do jacobiano, veja o Lema 4.6.2).
Substituindo esta expressão em (1.18), vem que

$$h_\nu(f) = \sum_{P \in \mathcal{P}} \int_P \log J_\nu f d\nu = \int \log J_\nu f d\nu,$$

tal como afirmado no teorema. □

1.4 Entropia topológica

A entropia topológica de um sistema dinâmico é um número real não negativo que mede a complexidade do sistema. A primeira introdução de entropia topológica foi dada em 1965 por Adler, Konheim e McAndrew. Essa definição foi modelada após a definição de Kolmogorov-Sinai ou entropia métrica. Mais tarde, Dinaburg (8),em 1970, e Rufus Bowen (9), em 1971, deram independentemente uma diferente definição equivalente.

Nesse trabalho, estamos interessados apenas em espaços métricos compactos e suas coberturas abertas.

1.4.1 Definição via Coberturas

A definição de entropia topológica que daremos a seguir foi introduzida por Adler, Konheim e McAndrew (7), em 1965. Sua ideia de atribuir um número a uma cobertura aberta para medir a sua complexidade foi inspirada por Kolmogorov e Tihomirov (6), em 1961.

A seguir, definiremos a noção de coberturas abertas que será de fundamental importância na definição da entropia topológica.

Definição 1.4.1. Seja X um subconjunto de um espaço métrico M. Uma cobertura de X uma família $C = (A_\lambda)_{\lambda \in L}$ de subconjuntos de M tal que $X \subset \bigcup_{\lambda \in L} A_\lambda$. Dizemos que a cobertura é aberta se cada A_λ, $\lambda \in L$, é aberto em M.

Definiremos agora uma operação entre coberturas que será útil no cálculo da entropia topológica.

Definição 1.4.2. Sejam α e β coberturas abertas de um espaço métrico X. Então,

- $\alpha \vee \beta = \{A_i \cap B_j; A_i \in \alpha, B_j \in \beta\}$ também é uma cobertura aberta de X.
- A cobertura β refina α se, para todo $B \in \beta$, existir $A \in \alpha$ tal que $B \subset A$. Denotaremos por $\alpha \prec \beta$.
- Se α é uma cobertura aberta de um espaço métrico X e $f : X \to X$ é uma aplicação contínua então $f^{-1}(\alpha) = \{f^{-1}(A); A \in \alpha\}$ é uma cobertura de X.

Temos, $f^{-1}(\alpha \vee \beta) = f^{-1}(\alpha) \vee f^{-1}(\beta)$ e, se $\alpha \prec \beta$, então $f^{-1}(\alpha) \prec f^{-1}(\beta)$. Em particular, denotamos $\alpha \vee f^{-1}(\alpha) \vee ... \vee f^{-n+1}(\alpha)$ por $\vee_{i=0}^{n-1} f^{-i}(\alpha)$ ou simplesmente, α^n.

Definição 1.4.3. Sejam X um espaço métrico compacto e α uma cobertura aberta de X. Definimos $N(\alpha)$ como sendo o número de conjuntos de uma subcobertura de α com a menor cardinalidade, ou seja,

$$N(\alpha) = \inf\{\#\beta; \beta \subset \alpha \text{ subcobertura de} X\}.$$

A entropia de α é o número $H(\alpha) = \log N(\alpha)$.

Exemplo 1.4.1. *Seja $X = [0,1] \subset R$. Considere a cobertura $\alpha = \{(-1, \frac{1}{2}); (-\frac{1}{5}, \frac{2}{5}); (\frac{1}{3}, \frac{4}{3}); (\frac{2}{3}, \frac{3}{2})\}$. Daí segui que $N(\alpha) = 2$, portanto, $H(\alpha) = \log 2$.*

Os lemas que veremos a seguir serão úteis para provarmos a proposição 2.1.

Lema 1.4.4. *Sejam α, β coberturas abertas de X e $f : X \to X$ uma aplicação contínua. Então,*

(i) $H(\alpha) \geq 0$;

(ii) $\alpha \prec \beta \Rightarrow H(\alpha) \leq H(\beta)$;

(iii) $H(\alpha \vee \beta) \leq H(\alpha) + H(\beta)$;

(iv) $H(f^{-1}(\alpha)) \leq H(\alpha)$. *Se f sobrejetiva então $H(f^{-1}(\alpha)) = H(\alpha)$.*

Demonstração. Sejam α, β coberturas abertas de X e $f : X \to X$ uma aplicação contínua.

(i) segue direto da definição.

Para (ii), seja $\{B_1, B_2, ..., B_{N(\beta)}\}$ uma subcobertura de β com menor cardinalidade. Para cada B_i existe $A_i \in \alpha$ tal que $B_i \subset A_i$. Portanto, $\{A_1, A_2, ..., A_{N(\beta)}\}$ é uma subcobertura de α. Logo, $N(\alpha) \leq N(\beta)$, assim, $H(\alpha) \leq H(\beta)$.

(iii) Sejam $\{A_1, A_2, ..., A_{N(\alpha)}\}$ e $\{B_1, B_2, ..., B_{N(\beta)}\}$ subcoberturas de α e β, respectivamente, de menor cardinalidade. Então o conjunto $\{A_i \cap B_j; A_i \in \alpha, B_j \in \beta\}$ é uma subcobertura de $\alpha \vee \beta$. Assim, $N(\alpha \vee \beta) \leq N(\alpha)N(\beta)$.

Finalmente para (iv), seja $\{A_1, A_2, ..., A_{N(\alpha)}\}$ uma subcobertura de α de menor cardinalidade, então $\{f^{-1}(A_1), f^{-1}(A_2), ..., f^{-1}(A_{N(\alpha)})\}$ é uma subcobertura de $f^{-1}(\alpha)$. Portanto $N(f^{-1}(\alpha)) \leq N(\alpha)$ o que nos leva $H(f^{-1}(\alpha)) \leq H(\alpha)$.

Se f sobrejetiva e $\{f^{-1}(A_1), ..., f^{-1}(A_{N(f^{-1}(\alpha))})\}$ uma subcobertura de $f^{-1}(\alpha)$ de menor cardinalidade, então $\{A_1, ..., A_{N(f^{-1}(\alpha))}\}$ também uma subcobertura de α. Assim, $N(f^{-1}(\alpha)) \geq N(\alpha)$ o que acarreta $N(f^{-1}(\alpha)) = N(\alpha)$. Portanto, $H(f^{-1}(\alpha)) = H(\alpha)$. □

Proposição 1.4.5. *Se α é uma cobertura de X e $f : X \to X$ é contínua, então*

$$\lim_{n\to+\infty} \frac{1}{n} H(\alpha^n) \text{ existe.}$$

Demonstração. Considere $a_n = H(\alpha^n)$. Logo, $a_{n+m} = H(\alpha^{n+m-1}) = H(\vee_{i=0}^{n+m-1} f^{-i}(\alpha)) \leq H(\vee_{i=0}^{n-1} f^{-i}(\alpha)) + H(\vee_{i=n}^{n+m-1} f^{-i}(\alpha)) \leq H(\vee_{i=0}^{n-1} f^{-i}(\alpha)) + H(f^{-n}(\vee_{i=0}^{m-1} f^{-i}(\alpha)) \leq H(\vee_{i=0}^{n-1} f^{-i}(\alpha)) + H(\vee_{i=0}^{m-1} f^{-i}(\alpha)) = a_n + a_m$. Portanto, a_n subaditiva. Logo $\lim\limits_{n\to+\infty} \dfrac{a_n}{n} = \lim\limits_{n\to+\infty} \dfrac{1}{n} H(\vee_{i=0}^{n-1} f^{-i}(\alpha))$ existe. □

Como consequência da proposição anterior, podemos definir o seguinte.

Definição 1.4.6. Seja α uma cobertura aberta de X e $f : X \to X$ uma aplicação contínua, então a entropia de f com relação à α é dada por

$$h(f, \alpha) = \lim_{n\to+\infty} \frac{1}{n} H(\alpha^n).$$

Exemplo 1.4.2. *Seja $f : [0,1] \to [0,1]$, dada por $f(x) = 2x\,(mod 1)$. Tome a cobertura aberta $\alpha = \{[0,1/2],[1/2,1]\}$. Então, temos $f^{-1}(\alpha) = \{[0,1/4]; [1/4,1/2]; [1/2,3/4]; [3/4,1]\}, ..., f^{-n+1}$ $\{[j/2^n, (j+1)/2^n]; j = 0, 1, ..., 2^n - 1\}$. Logo $H(\alpha^n) = \log 2^n$. Daí segui que,*

$$h(f, \alpha) = \lim_{n\to+\infty} \frac{1}{n} H(\alpha^n) = \lim_{n\to+\infty} \frac{1}{n} n \log 2 = \log 2.$$

Agora, mostraremos algumas propriedades técnicas da entropia de uma aplicação f com relação a uma cobertura.

Lema 1.4.7. *Se α, β são coberturas de X. Então*

(v) $h(f, \alpha) \leq H(\alpha)$;

(vi) $\alpha \prec \beta \Rightarrow h(f, \alpha) \leq h(f, \beta)$.

Demonstração. Para (v), temos

$$H(\alpha^n) \leq \sum_{i=0}^{n-1} H(f^{-1}(\alpha)) \leq nH(\alpha) \Rightarrow \frac{1}{n} H(\alpha^n) \leq H(\alpha) \Rightarrow h(f, \alpha) \leq H(\alpha).$$

Para (vi), como $\alpha \prec \beta \Rightarrow \alpha^n \prec \beta^n$, temos pelo Lema 1.4.4 (item (ii)) que

$$H(\alpha^n) \leq H(\beta^n) \Rightarrow h(f,\alpha^n) \leq h(f,\beta^n).$$

□

A definição de entropia topológica que daremos a seguir está bem definida, pelo que vimos anteriormente.

Definição 1.4.8. Seja $f : X \to X$ uma aplicação contínua, a entropia topológica de f é dada por

$$h(f) = \sup_{\alpha} h(f,\alpha)$$

onde α varia ao longo de toda cobertura aberta de X.

O teorema abaixo diz que, se $f : X \to X$ é um homeomorfismo, a entropia topológica de f e f^{-1} coincidem. Assim, só estaremos interessados nos iterados positivos.

Teorema 1.4.9. *Seja $f : X \to X$ um homeomorfismo de X, então $h(f) = h(f^{-1})$.*

Demonstração. Do item (iv) acima, segue que

$$\begin{aligned} h(f,\alpha) &= \lim_{n\to+\infty} \frac{1}{n} H(\alpha^n) = \lim_{n\to+\infty} \frac{1}{n} H(f^{n-1}(\alpha^n)) \\ &= \lim_{n\to+\infty} \frac{1}{n} H(\vee_{i=0}^{n-1} f^i(\alpha)) = \lim_{n\to+\infty} \frac{1}{n} H(\alpha^{-n}) \\ &= h(f^{-1},\alpha). \end{aligned}$$

□

1.4.2 Definição de Rufus Bowen e Dinaburg

Daremos a seguinte definição para introduzir uma outra noção de entropia topológica.

Definição 1.4.10. Seja $f : X \longrightarrow X$ uma aplicação contínua no espaço métrico compacto.

(a) Para $n \geq 1$ e $\varepsilon > 0$, chamamos o conjunto $S \subset X$ de um conjunto $(n,\varepsilon)-separado$ se, para cada $x, y \in S$ com $x \neq y$, tivermos que $d(f^i(x), f^i(y)) \geq \varepsilon$ para algum $i = 0, 1, ..., n-1$.

Seja $s_n(\varepsilon) \in \mathbb{N}$ o máximo das cardinalidades de todos os conjuntos $(n,\varepsilon)-separado$.

(b) Para $n \geq 1$ e $\varepsilon > 0$ chamamos um conjunto $R \subset X$ de um conjunto $(n,\varepsilon)-$*gerador* se, para todo $x \in X$, existir $y \in R$ tal que $d(f^i(x), f^i(y)) < \varepsilon$ para todo $i = 0, 1, ..., n-1$.

Seja $r_n(\varepsilon) \in \mathbb{N}$ a menor das cardinalidades de todos os conjuntos $(n,\varepsilon)-$*gerador*.

O lema abaixo será usado para provarmos a equivalência entre a Definição 1.4.8 e a Definição 1.4.10.

Lema 1.4.11. (i) *Para $\varepsilon > \varepsilon'$, temos que $s_n(\varepsilon') \geq s_n(\varepsilon)$ e $r_n(\varepsilon') \geq r_n(\varepsilon)$. Em particular,*

$$\limsup_{n\to+\infty} \frac{1}{n}\log(s_n(\varepsilon')) \geq \limsup_{n\to+\infty} \frac{1}{n}\log(s_n(\varepsilon)) \quad e \quad \limsup_{n\to+\infty} \frac{1}{n}\log(r_n(\varepsilon')) \geq \limsup_{n\to+\infty} \frac{1}{n}\log(r_n(\varepsilon));$$

(ii) *Para $\varepsilon > 0$ e $n \geq 1$ temos que*

$$r_n(\varepsilon) \leq s_n(\varepsilon) \leq r_n(\varepsilon/2). \tag{1.20}$$

Demonstração. (i) Segue direto da definição.

Para (ii), observemos o conjunto S que é um $(n,\varepsilon)-$*separado* de maior cardinalidade é, também um $(n,\varepsilon)-$*gerador*. Caso contrário, existe $x_0 \in X$ tal que para todo $y \in S$ vale que $d(f^i(x_0), f^i(y)) \geq \varepsilon$ para algum $i = 0, ..., n-1$. Assim, $S \cup \{x_0\}$ é um $(n,\varepsilon)-$*separado* com cardinalidade maior que a cardinalidade de S, absurdo. Logo, $r_n(\varepsilon) \leq s_n(\varepsilon)$.

Para mostrarmos que $s_n(\varepsilon) \leq r_n(\varepsilon/2)$, considere S $(n,\varepsilon)-$*separado*. Seja R um $(n,\varepsilon/2)-$*gerador* de menor cardinalidade. Defina $\varphi : S \to R$ de forma que, para cada $x \in S$, $\varphi(x) \in R$ com $d(f^i(x), f^i(\varphi(x))) < \varepsilon/2$ para todo $i = 0, 1, ..., n-1$.

Afirmamos que φ é injetiva.

Com efeito, suponha que $x, x' \in S$ tais que $\varphi(x), \varphi(x') \in R$, com $\varphi(x) = \varphi(x')$. Logo,

$$d(f^i(x), f^i(x')) \leq d(f^i(x), f^i(\varphi(x))) + d(f^i(x'), f^i(\varphi(x'))) < \varepsilon$$

para todo $i = 0, ..., n-1$. Daí segue que x, x' são iguais.

□

A definição de entropia topológica que veremos a seguir foi introduzida por Rufus Bowen(9) e Dinaburg(8), em 1971 e 1970, respectivamente, ambos utilizaram a seguinte noção.

Definição 1.4.12. Seja $f : X \to X$ uma aplicação contínua no espaço métrico compacto. A entropia topológica de f via $(n,\varepsilon)-$*gerador* é dada por:

$$\bar{h}(f) = \lim_{\varepsilon\to 0}\limsup_{n\to+\infty}\frac{1}{n}\log(r_n(\varepsilon)).$$

Segue de (1.20) que:

$$\bar{h}(f) = \lim_{\varepsilon\to 0}\limsup_{n\to+\infty}\frac{1}{n}\log(s_n(\varepsilon)) = \lim_{\varepsilon\to 0}\limsup_{n\to+\infty}\frac{1}{n}\log(r_n(\varepsilon)). \tag{1.21}$$

1.4.3 A equivalência entre as definições

Apresentamos agora, alguns resultados técnicos para mostrarmos a equivalência entre as definições de entropia topológica.

Definição 1.4.13. Diz-se que o número $\delta > 0$ é o número de Lebesgue de uma cobertura α de um espaço métrico X, quando todo subconjunto $U \subset X$ com diâmetro menor que δ está contido em algum elemento de α.

Proposição 1.4.14. *Seja X um espaço métrico compacto, então toda cobertura aberta α possui um número de Lebesgue.*

Demonstração. Suponha, por absurdo, que a cobertura dada não possua número de Lebesgue. Obtemos, para cada $n \in \mathbb{N}$, um conjunto $S_n \subset X$ com $\operatorname{diam} S_n = \sup\{d(x,y); x, y \in S_n\} < 1/n$ tal que nenhum A_λ contenha S_n. Escolhamos em cada S_n um ponto x_n. Passando a uma subsequência se necessário, podemos admitir que $\lim x_n = a \in X$. Temos $a \in A_\lambda$ para algum $\lambda \in L$. Existe $\varepsilon > 0$ tal que $B(a;\varepsilon) \subset A_\lambda$. Tomemos $n \in \mathbb{N}$ tal que $1/n < \varepsilon/2$ e $d(a,x_n) < \varepsilon/2$. Então

$$y \in S_n \Rightarrow d(a,y) \le d(a,x_n) + d(x_n,y) < \frac{1}{n} + \frac{\varepsilon}{2} < \varepsilon.$$

Logo $S_n \subset B(a;\varepsilon) \subset A_\lambda$, uma contradição. $\square$

Lema 1.4.15. *Seja $f : X \to X$ contínua em um espaço métrico compacto.*

(vii) *Seja α uma cobertura aberta de X com número de Lebesgue $\delta > 0$, então*

$$N(\vee_{i=0}^{n-1} f^{-i}(\alpha)) \le r_n(\delta)$$

para todo $n \ge 1$;

(viii) *Seja $\varepsilon > 0$ e $\gamma = \{B_1, B_2, ..., B_k\}$ uma cobertura aberta com $\max\limits_{1\le i\le k} \operatorname{diam} B_i < \varepsilon$, então*

$$s_n(\varepsilon) \leq N(\vee_{i=0}^{n-1} f^{-i}(\gamma)).$$

Demonstração. (vii) Seja R um $(n,\delta)-$*gerador* de cardinalidade $r_n(\delta)$. Então $X = \bigcup_{x\in R} D(x,n,\delta)$, onde $D(x,n,\delta) = \bigcap_{i=0}^{n-1} f^{-i}B(f^i(x),\delta)$. Portanto, para cada $x \in R$, temos $B(f^i(x),\delta) \subset A_{k_i}$, com $A_{k_i} \in \alpha$ para cada $i = 0, ..., n-1$, i. e, $D(x,n,\delta) \subset \bigcap_{i=0}^{n-1} f^{-i}(A_{k_i}) \in \vee_{i=0}^{n-1} f^{-i}(\alpha)$. Em particular, $\bigcup_{x\in R} \bigcap_{i=0}^{n-1} f^{-i}(A_{k_i})$ cobertura de X. Daí, temos que

$$N(\vee_{i=0}^{n-1} f^{-i}(\alpha)) \leq r_n(\delta).$$

s (viii) Seja S um $(n,\varepsilon)-$*separado* de cardinalidade $s_n(\varepsilon)$. Cada ponto do conjunto S pertence a diferentes elementos de $\vee_{i=0}^{n-1} f^{-i}(\gamma)$ (desde que $x, y \in \bigcap_{j=0}^{n-1} f^{-j}(B_{i_j}) \in \vee_{i=0}^{n-1} f^{-i}(\gamma)$, temos $d(f^r(x), f^r(y)) \leq \operatorname{diam} B_{i_r} < \varepsilon$). Em particular,

$$s_n(\varepsilon) \leq N(\vee_{i=0}^{n-1} f^{-i}(\gamma)).$$

$\square$

O seguinte resultado mostra a equivalência entre a definição de Adler, Konheim e McAndrew e a definição de Dinaburg e Rufus Bowen.

Teorema 1.4.16. *A entropia topológica de uma aplicação contínua $f : X \to X$ em um espaço métrico compacto é dada por:*

$$h(f) = \lim_{\varepsilon\to 0} \limsup_{n\to+\infty} \frac{1}{n} \log(r_n(\varepsilon))$$

e

$$h(f) = \lim_{\varepsilon\to 0} \limsup_{n\to+\infty} \frac{1}{n} \log(s_n(\varepsilon)).$$

Demonstração. Sendo $h(f) = \sup_{\alpha} h(f,\alpha)$. Dado qualquer $\eta > 0$ existe uma cobertura aberta de X, seja α essa cobertura, tal que

$$h(f,\alpha) + \eta \geq h(f) \geq h(f,\alpha).$$

Seja $\delta > 0$ o número de Lebesgue de α. Sabemos, pela de (1.21), que

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

Em particular,

$$\begin{aligned}\lim_{\varepsilon\to 0}\limsup_{n\to+\infty}\frac{1}{n}\log(r_n(\varepsilon)) &\geq \limsup_{n\to+\infty}\frac{1}{n}\log(r_n(\delta))\\ &\geq \limsup_{n\to+\infty}\frac{1}{n}\log(N(\vee_{i=0}^{n-1}f^{-i}(\alpha)))\\ &= h(f,\alpha)\geq h(f)-\eta.\end{aligned}$$

Por outro lado, dado $\varepsilon>0$, podemos escolher uma cobertura aberta $\beta=\{B_1,...,B_k\}$ de X com $\max\limits_{1\leq i\leq k}\operatorname{diam} B_i<\varepsilon$. Assim, temos

$$\limsup_{n\to+\infty}\frac{1}{n}\log(s_n(\varepsilon))\leq\limsup_{n\to+\infty}\frac{1}{n}\log(N(\vee_{i=0}^{n-1}f^{-i}(\beta)))=h(f,\beta)\leq h(f).$$

Fazendo $\varepsilon\to 0$, temos

$$\lim_{\varepsilon\to 0}\limsup_{n\to+\infty}\frac{1}{n}\log(s_n(\varepsilon))\leq h(f).$$

Portanto,

$$h(f)-\eta\leq\lim_{\varepsilon\to 0}\limsup_{n\to+\infty}\frac{1}{n}\log(s_n(\varepsilon))\leq h(f).$$

Isso completa a prova. □

1.5 Princípio variacional

Agora daremos uma importante relação entre entropia métrica e entropia topológica.

Teorema 1.5.1 (Princípio variacional). *Se $f:M\to M$ é uma transformação contínua num espaço métrico compacto então a sua entropia topológica $h_{top}(f)$ coincide com o supremo das entropias métricas $h_\nu(f)$ da transformação f relativamente a todas as probabilidades invariantes. Mais precisamente,*

$$h_{top}(f)=\sup_\nu h_\nu(f). \tag{1.22}$$

Mostraremos agora que o supremo em (1.22) pode ser tomado sobre todas probabilidades ergódicas.

Com efeito, considere $\bar{h}=\sup h_\nu(f)$ o supremo com relação as probabilidades ergódicas. Claramente, vemos que $h_{top}(f)\geq\bar{h}$. Por outro lado, dado $\varepsilon>0$ existe uma probabilidade η tal que

$$h_{top}(f) - \varepsilon \leq h_\eta(f) \leq h_{top}(f).$$

Como, pelo Teorema de Jacobs 4.5.2, temos que

$$h_\nu(f) = \int h_{\nu_P}(f) d\hat{\nu}(P).$$

Logo, alguma destas probabilidade ν_P temos $h_{\nu_P}(f) \geq h_\eta(f)$. Assim,

$$h_{top}(f) \leq \bar{h}(f).$$

Portanto,

$$h_{top}(f) = \sup_\nu h_\nu(f) \quad \text{com } \nu \text{ probabilidade ergódica.} \tag{1.23}$$

Denotaremos entropia topológica de uma transformação $f : M \to M$ por $h_{top}(f)$.

2 PRELIMINARES

O objetivo deste capítulo é apresentar alguns resultados que será de fundamental importância no capítulo seguinte, onde mostraremos que sob algumas hipóteses garantimos a existência de medidas maximizantes.

2.1 Operador de Ruelle-Perron-Frobenius

Sejam (X,d) um espaço métrico compacto e $f: X \to X$ uma aplicação contínua. O par (X,f) é chamado de sistema dinâmico. Seja $\psi: X \to \mathbb{R}^+_*$ uma função contínua estritamente positiva que chamaremos de *potencial*. O *operador de Ruelle-Perron-Frobenius* $\mathcal{L} = \mathcal{L}_{f,\psi}$, simplesmente chamado de *operador de Ruelle*, é definido como

$$\mathcal{L}g(x) = \sum_{y:f(y)=x} \psi(y)g(y)$$

para g em um espaço adequado de funções em X. O *operador de Ruelle* é uma importante ferramenta no estude de sistemas dinâmicos.

Considere M uma variedade riemanniana de classe C^1, compacta, conexa. Sejam $f: M \to M$ um difeomorfismo local de classe C^1 e o potencial $\psi \equiv 1$. Seja $p \geq 1$ o grau de f, que é, o número $\# f^{-1}(x)$ das pré-imagens de qualquer x em M. Definiremos *operador de Ruelle-Perron-Frobenius* no espaço das funções contínua $C(M)$. i. e, $\mathcal{L}: C(M) \to C(M)$ definido por,

$$\mathcal{L}g(x) = \sum_{y:f(y)=x} g(y).$$

Agora mostraremos que $\mathcal{L}$ está bem definido. De fato, como f é um difeomorfismo local existem V com $x \in V$ e $U_1, U_2, ..., U_p$ abertos em M tal que $f: U_i \to V$ é difeomorfismo, $i = 1, 2, ..., p$. Logo,

$$\mathcal{L}g(x) = \sum_{i=1}^{p} g \circ f^{-1}(x), x \in V.$$

Portanto, $\mathcal{L}$ está bem definido. É fácil ver que se $g \geq 0$ então $\mathcal{L}g \geq 0$. Além disso, $||\mathcal{L}|| = p$.

Agora daremos algumas definições que será útil no decorrer desta teoria.

Seja E um espaço de Banach. Um subconjunto fechado e convexo C é chamado de *cone* de E, se ele satisfaz:

$$\lambda C \subset C \text{ para todo } \lambda \geq 0 \text{ e } C \cap (-C) = \{0\}.$$

Dizemos que o cone C é *normal* quando

$$\inf\{||x+y|| : x, y \in C \text{ tais que } ||X|| = ||y|| = 1\} > 0.$$

Fixemos um cone C de E. Dado um operador linear contínuo $T : E \to E$, diremos que T é um *operador positivo* sobre C se $T(C) \subset C$. Dado um funcional linear contínuo $\phi : E \to \mathbb{R}$, diremos que ϕ é um *funcional positivo* sobre C se $\phi(v) \geq 0$ para todo $v \in C$. Chamaremos de *cone dual* conjunto C^* de todos os funcionais positivos sobre C.

Exemplo 2.1.1. *Seja M espaço métrico compacto. Denote por $C^0_+(M)$ conjunto de todas as funções contínuas positivas. Observe que $C^0_+(M)$ é um cone e que é preservado pelo operador de Ruelle $\mathcal{L}$. O cone dual*

$$C^0_+(M)^* = \{\eta \in C^0(M)^* : \eta(g) \geq 0, \forall g \in C^0_+(M)\}.$$

Seja $\mathcal{L}^* : \mathcal{M}(M) \to \mathcal{M}(M)$ o operador dual de $\mathcal{L}$ agindo sobre o espaço das medidas de Borel de M, sabemos por definição que à ação do dual satisfaz,

$$\int g d\mathcal{L}^*(\nu) = \mathcal{L}^*\nu(g) = \nu(\mathcal{L}g) = \int \mathcal{L}g d\nu.$$

Esse operador linear é positivo, no sentido de que se η é uma medida positiva então $\mathcal{L}^*\eta$ é também positiva.

Daremos agora uma definição que será útil mais adiante.

Definição 2.1.1. Seja $T : V \to V$ um operador linear contínuo. O raio espectral é dado por:

$$r(T) = \lim_{n\to\infty} \sqrt[n]{||T^n||}.$$

É fácil ver que os espectros de $\mathcal{L}$ e $\mathcal{L}^*$ estão contido no disco fechado de raio p. Chamamos automedida maximal qualquer medida de probabilidade μ que satisfaz

$$\mathcal{L}^*\mu = p\mu.$$

Agora para mostraremos à existência de uma medida maximal. Para isso, defina $G : \mathcal{M}(M) \to \mathcal{M}(M)$ no espaço das medidas $\mathcal{M}(M)$ por

$$G(\nu) = \frac{1}{p}\mathcal{L}^*(\nu).$$

Observe que G está bem definida e é contínua na topologia fraca* em $\mathcal{M}_1$. Com efeito, é fácil ver que $\mathcal{L}^*\nu(M) = p\nu(M)$, portanto, bem definida. Agora dado $\nu_k \xrightarrow{*} \nu$, temos

$$\int gd\mathcal{L}^*\nu_k = \int \mathcal{L}gd\nu_k \to \int \mathcal{L}gd\nu = \int gd\mathcal{L}^*\nu, \forall g \in C(M).$$

Isso mostra a continuidade de G na topologia fraca* em $\mathcal{M}(M)$. Desde de que $\mathcal{M}_1$ é um espaço compacto e convexo, usaremos o seguinte teorema para concluir que G possui ponto fixo.

Teorema 2.1.2 (Tychonoff-Schauder)**.** *Seja $T : V \to V$ uma transformação contínua num V um espaço vetorial topológico. Suponha que exita $K \subset V$ compacto, convexo tal que $T(K) \subset K$. Então T possui ponto fixo.*

Assim, tomando $V = \mathcal{M}(M)$ e $K = \mathcal{M}_1$, pelo teorema de Tychonoff-Schauder, existe alguma probabilidade μ tal que $G(\mu) = \mu$. Em outras palavras, μ é uma automedida maximal. Observe também que μ é invariante para f. De fato, para toda g contínua temos que $\mathcal{L}(g \circ f)(x) = pg(x)$ e

$$\int (g \circ f)d\mu = \frac{1}{p}\int (g \circ f)d\mathcal{L}^*\mu = \frac{1}{p}\int \mathcal{L}(g \circ f)d\mu = \int gd\mu.$$

Logo, pelo seguinte lema concluímos que μ é invariante para f.

Lema 2.1.3. *Sejam M um espaço métrico compacto, $f : M \to M$ uma transformação mensurável e μ uma medida em M. Então μ é invariante para f se, e somente se,*

$$\int gd\mu = \int (g \circ f)d\mu,$$

para toda $g : M \to \mathbb{R}$ contínua.

2.2 Medidas com entropia grande

Nesta seção provaremos que, sob nossas hipóteses, qualquer medida com entropia grande tem todos os seus expoentes de Lyapunov positivos.

2.2.1 Produto exterior

Seja V é um espaço vetorial de dimensão d com um produto interno. Para cada inteiro k tal que $1 \leq k \leq d$, seja $V^{\wedge_k}$ é o espaço das k-linear formas alternadas em V, isto é,

$$\{v_1 \wedge \cdots \wedge v_k; v_i \in V \quad \text{para} \quad 1 \leq j \leq k\}$$

e seus elementos satisfazem:

(a) $v_1 \wedge \cdots \wedge (xu_i + yw_i) \wedge \cdots \wedge v_k = x v_1 \wedge \cdots \wedge u \wedge \cdots \wedge v_k + y v_1 \wedge \cdots \wedge w \wedge \cdots \wedge v_k$ para todo $x, y \in \mathbb{R}$;

(b) $v_1 \wedge \cdots \wedge u \wedge w \wedge \cdots \wedge v_k = -v_1 \wedge \cdots \wedge w \wedge u \wedge \cdots \wedge v_k$.

Para qualquer transformação linear A de V, o k-ésimo produto exterior $A^{\wedge_k}$ de A é a única transformação linear de $V^{\wedge_k}$ tal que

$$A^{\wedge_k} v_1 \wedge \cdots \wedge v_k = Av_1 \wedge \cdots \wedge Av_k$$

para quaisquer $v_1, ..., v_k \in V^{\wedge_1} \equiv V$.

Se $\{e_1, ..., e_d\}$ é uma base de V, então $\{e_{i_1} \wedge \cdots \wedge e_{i_k} : 1 \leq i_1 < \cdots < i_k \leq d\}$ é uma base de $V^{\wedge_k}$, por isso $\dim V^{\wedge_k} = \binom{d}{k}$.

Podemos definir um produto interno em $V^{\wedge_k}$ dado por

$$\langle v_1 \wedge \cdots \wedge v_k, w_1 \wedge \cdots \wedge w_k \rangle = \det(\langle v_i, w_j \rangle), 1 \leq i, j \leq k,$$

para quaisquer $v_1 \wedge \cdots \wedge v_k, w_1 \wedge \cdots \wedge w_k \in V^{\wedge_k}$. A norma induzida satisfaz as seguintes propriedades:

1. $||v_1 \wedge \cdots \wedge v_k|| \leq ||v_1 \wedge \cdots \wedge v_l|| ||v_{l+1} \wedge \cdots \wedge v_k||$ para qualquer $l < k$,

2. Para quaisquer transformações A, B de V e $1 \leq k, l \leq d$, os operadores induzidos satisfazem:

   (i) $||(AB)^{\wedge_k}|| \leq ||A^{\wedge_k}|| ||B^{\wedge_k}||$;

   (ii) $||A^{\wedge_{(k+l)}}|| \leq ||A^{\wedge_k}|| ||A^{\wedge_l}|| \leq ||A||^{k+l}$;

   (iii) $(A^{\wedge_k})^* = (A^*)^{\wedge_k}$;

   (iv) $||A^{\wedge_k}|| = \prod_{j=1}^{k} \lambda_j$, onde $\lambda_1 \leq \cdots \leq \lambda_d$ são os autovalores de $(A^* A)^{1/2}$.

(v) Os autovalores de $A^{\wedge k}$ são da forma $\lambda_{i_1}\cdots\lambda_{i_k}$ com $1 \leq i_j \leq d$, onde os λ_j são autovalores de A.

Dessas propriedades segue que dada $(A_n)_{n\geq 1}$ tal que $A_n \to A$ então $A_n^{\wedge k} \to A^{\wedge k}$.

2.2.2 Expoente de Lyapunov

Seja M um espaço métrico compacto. Considere $(M, \mathcal{B}, \rho)$ o completamento de um espaço de probabilidade, i. e, $\mathcal{B}$ contém todos os subconjuntos de X tal que $\rho(X) = 0$. $(M, \mathcal{B}, \rho)$ é chamado de *espaço de Lebesgue*. Seja $T : M \to GL_d$ uma função mensurável. Dada $\tau : M \to M$ uma função que preserva ρ, definimos

$$T_x^n = T(\tau^{n-1}(x))\cdots T(\tau(x))T(x).$$

Denotaremos por T^* a matriz transposta de T.

Teorema 2.2.1 (Ergódico multiplicativo)**.** *Se*

$$\log^+ ||T(\cdot)|| \in L^1(M, \rho).$$

Então existe $\Gamma \subset M$ tal que $\tau\Gamma \subset \Gamma$, $\rho(\Gamma) = 1$, e as seguintes propriedades valem se $x \in \Gamma$:

(a) $\lim((T_x^n)^* T_x^n)^{1/2n} = \Lambda_x$

(b) *Sejam $\exp\lambda_x^{(1)} < \cdots < \exp\lambda_x^{(s)}$ os autovalores de Λ_x (com $s = s(x)$, os $\lambda_x^{(r)}$ são reais, e $\lambda_x^{(1)}$ pode ser $-\infty$), e $U_x^{(1)}, ..., U_x^{(s)}$ os autoespaços correspondentes. Seja $m_x^{(r)} = \dim U_x^{(r)}$. As funções $x \mapsto \lambda_x^{(r)}, m_x^{(r)}$ são τ- invariante. Escrevendo $V_x^{(0)} = \{0\}$ e $V_x^{(r)} = U_x^{(1)} \oplus \cdots \oplus U_x^{(r)}$, temos:*

$$\lim_{n\to\infty} \frac{1}{n} \log^+ ||T_x^n u|| = \lambda_x^{(r)} \quad \textit{quando} \quad u \in V_x^{(r)} \backslash V_x^{(r-1)}$$

para $r = 1, ..., s$.

Os números $\lambda_x^{(r)}$ são chamados de expoente de Lyapunov de T no ponto $x \in M$. Também escreveremos todos os expoentes de Lyapunov como

$$\lambda_x^{(1)} \geq \lambda_x^{(2)} \geq \cdots \geq \lambda_x^{(d)},$$

onde cada expoente de Lyapunov é repetido de acordo com sua multiplicidade. Chamaremos $\lambda_\rho^{(i)}$ de expoente de Lyapunov de ρ que é a *integral do i-ésimo expoente de Lyapunov* com respeito à medida ρ, i. e,

$$\lambda_\rho^{(i)} = \int \max\{0, \lambda_x^{(i)}\} d\rho, \quad \text{para} \quad i = 1, ..., d.$$

Mostraremos que o expoente de Lyapunov de $T^{\wedge k}$ são as somas de k expoentes de Lyapunov de T.

Seja $T^{\wedge k} : M \to GL_{\binom{d}{k}}$ o k-ésimo produto exterior de T. Temos:

$$T^{\wedge k}(\tau^{n-1}(x)) \cdots T^{\wedge k}(x) = (T_x^n)^{\wedge k}$$

e

$$\lim_{n\to\infty}((T_x^n)^{\wedge k})^*(T_x^n)^{\wedge k})^{1/2n} = \lim_{n\to\infty}((T_x^n)^*(T_x^n))^{1/2n})^{\wedge k} = \Lambda_x^{\wedge k}. \tag{2.1}$$

De fato, observe que

$$\begin{aligned}
(T^{\wedge k})_x^n(u_1 \wedge \cdots \wedge u_k) &= T^{\wedge k}(\tau^{n-1}(x)) \cdots T^{\wedge k}(x)(u_1 \wedge \cdots \wedge u_k) \\
&= T(\tau^{n-1}(x)) \cdots T(x) u_1 \wedge \cdots \wedge T(\tau^{n-1}(x)) \cdots T(x) u_k \\
&= T_x^n u_1 \wedge \cdots \wedge T_x^n u_k \\
&= (T_x^n)^{\wedge k}(u_1 \wedge \cdots \wedge u_k)
\end{aligned}$$

e

$$\begin{aligned}
((T_x^n)^{\wedge k})^*(T_x^n)^{\wedge k}(u_1 \wedge \cdots \wedge u_k) &= (T_x^n)^*(T_x^n) u_1 \wedge \cdots \wedge (T_x^n)^*(T_x^n) u_k \\
&= ((T_x^n)^*(T_x^n))^{\wedge k}(u_1 \wedge \cdots \wedge u_k).
\end{aligned}$$

Isso mostra à afirmação (2.1).

Assim, pelo Teorema 2.2.1 e da Propriedade (v) acima, segue que os autovalores de $\Lambda_x^{\wedge k}$ são da forma $\exp(\sum_{j=1}^k \lambda_x^{(i_j)})$ e $U_x^{(i_1)} \wedge \cdots \wedge U_x^{(i_k)}$ é autoespaço correspondente à $(i_1, ..., i_k)$ com $1 \leq i_1 < \cdots < i_k \leq s$. Com isso, temos

$$\lim_{n\to\infty} \frac{1}{n} \log ||(T_x^n)^{\wedge k}(u_{i_1} \wedge \cdots \wedge u_{i_k})|| = \sum_{j=1}^k \lambda_x^{(i_j)}. \tag{2.2}$$

Agora, considere $f : M \to M$ um difeomorfismo local de classe C^1 em uma variedade riemaniana compacta. Sejam $(M, \mathcal{B}, \mu)$ um espaço de Lebesgue e μ uma medida

f-invariante. Tomando, $T = Df$ e $\tau = f$, por (2.2), temos

$$\lambda_x^{i_1} + \cdots + \lambda_x^{i_k} = \lim_{n\to\infty} \frac{1}{n} \log ||(Df_x^n)^{\wedge_k}(u_{i_1} \wedge \cdots \wedge u_{i_k})||. \tag{2.3}$$

Suponha que,

$$\max_{1\le k\le d-1} \max_{x\in M} \limsup_{n\to\infty} \frac{1}{n} \log ||Df_x^{\wedge_k}|| < \log p.$$

Assim, a soma (2.3) é estritamente menor que $\log p$, para todo $k < d$.

Lema 2.2.2. *Se μ é uma medida de probabilidade invariante com alguma integral do expoente de Lyapunov menor que*

$$c(f) = \log p - \max_{1\le k\le d-1} \max_{x\in M} \limsup_{n\to\infty} \frac{1}{n} \log ||Df_x^{\wedge_k}||, \tag{2.4}$$

então $h_\mu(f) < \log p$.

Demonstração. Seja μ uma probabilidade invariante, e suponha $\lambda^d(\mu) = \int \max\{0, \lambda_x^d\} d\mu < c(f)$. Por (2.3), temos $\sum_{1\le j\le k} \lambda_x^j \le \max_{x\in M} \limsup_{n\to\infty} \frac{1}{n} \log ||Df_x^{\wedge_k}||$ para todo $1 \le k < d$. Então, usando a desigualdade de Ruelle 4.8.1,

$$h_\mu(f) \le \int \sum_{j:\lambda_x^j>0} \lambda_x^j d\mu < c(f) + \max_{1\le k\le d-1} \max_{x\in M} \limsup_{n\to\infty} \frac{1}{n} \log ||Df_x^{\wedge_k}|| \le \log p.$$

Isso prova o lema. □

2.3 Tempos hiperbólicos

Nesta introduziremos a noção de tempos hiperbólicos que será útil na construção da prova do nosso resultado principal.

Definição 2.3.1. Dado $c > 0$, dizemos que $n \in \mathbb{N}$ é um *c-tempo hiperbólico* para $x \in M$ se

$$\prod_{k=0}^{j-1} ||Df(f^{n-k}(x))^{-1}|| \le e^{-2cj} \quad \text{para todo} \quad 1 \le j \le n.$$

No que se segue fixamos $c = c(f)/10$ e falamos, simplesmente, de *tempos hiperbólicos.* Dizemos que f tem *densidade de tempos hiperbólicos positiva* para x se o conjunto H_x dos inteiros que são tempos hiperbólicos de f para x satisfaz

$$\liminf_n \frac{1}{n} \#(H_x \cup [1,n]) > 0. \tag{2.5}$$

O próximo lema mostra que há uma infinidade de tempos hiperbólicos.

Proposição 2.3.2. *Se um ponto x satisfaz*

$$\limsup_{n\to+\infty}\frac{1}{n}\sum_{i=0}^{n-1}\log||Df(f^i(x))^{-1}|| < -4c < 0,$$

então f tem densidade de tempos hiperbólicos positiva para x.

Para mostrarmos à proposição precisamos do seguinte lema.

Lema 2.3.3. *Dados $A \geq c_2 > c_1 > 0$, seja $\theta_0 = (c_2 - c_1)/(A - c_1)$. Então, dados quaisquer números reais $a_1, \ldots, a_N$ tais que*

$$\sum_{j=1}^{N} a_j \geq c_2 N \quad e \quad a_j \leq A \quad para \quad todo \quad 1 \leq j \leq N,$$

existem $l > \theta_0 N$ e $1 < n_1 < \cdots < n_l \leq N$ de modo que

$$\sum_{j=n+1}^{n_i} a_j \geq c_1(n_i - n) \quad para \quad todo \quad 0 \leq n \leq n_i - 1 \quad e \quad i = 1, \ldots, l.$$

Demonstração. Defina $S(n) = \sum_{j=1}^{n}(a_j - c_1)$, para cada $1 \leq n \leq N$, e também $S(0) = 0$. Então defina $1 < n_1 < \cdots < n_l \leq N$ para ser a sequencia maximal tal que $S(n_i) \geq S(n)$ para todo $0 \leq n \leq n_i - 1$ e $i = 1, ..., l$. Note que l não pode ser zero, desde que $S(N) > S(0)$. Além disso, a definição significa que

$$\sum_{j=n+1}^{n_i} a_j \geq c_1(n_i - n) \quad \text{para} \quad 0 \leq n \leq n_i - 1 \quad \text{e} \quad i = 1, ..., l.$$

Assim, só temos de verificar que $l > \theta_0 N$. Observe que, por definição,

$$S(n_i - 1) < S(n_{i-1}) \quad \text{e assim} \quad S(n_i) < S(n_{i-1}) + (A - c_1)$$

para todo $1 < i \leq l$. Além disso, $S(n_1) \leq A - c_1$ e $S(n_l) \geq S(N) \geq N(c_2 - c_1)$. Isso implica,

$$N(c_2 - c_1) \leq S(n_l) = \sum_{i=2}^{l}(S(n_i) - S(n_{i-1}) + S(n_1) < l(A - c_1),$$

que completa a prova.

$\square$

Agora provaremos a proposição anterior.

Demonstração. Como por hipótese

$$\limsup_n \frac{1}{n}\sum_{i=0}^{n-1} \log ||Df(f^i(x))^{-1}|| < -4c < 0$$

segue-se

$$\liminf_n \frac{1}{n}\sum_{i=0}^{n-1} -\log ||Df(f^i(x))^{-1}|| > 4c > 0.$$

Tomando, $x_n = \frac{1}{n}\sum_{i=0}^{n-1} -\log ||Df(f^i(x))^{-1}||$ e $b_n = \inf\{x_1,...,x_n\}$, temos

$$b_1 \geq b_2 \geq \cdots \geq b_n \geq \cdots \geq \lim b_n > 4c. \tag{2.6}$$

Assim, se tomarmos $a_i = -\log ||Df(f^i(x))^{-1}||$, $c_2 = 4c$, $c_1 = 2c$ e $A \geq c_2$ suficientemente grande, de modo que, $A \geq a_i$ para todo $i \in \mathbb{N}$, como por exemplo, $A = \max_{x\in M}\{-\log ||Df(x)^{-1}||\}$.

Logo, por (2.6) temos que, dado $N \in \mathbb{N}$ segue que

$$\sum_{i=0}^{N-1} a_i \geq 4cN = c_2N,$$

então pelo lema anterior, existem $l > \theta_0 N$ e $1 < n_1 < \cdots < n_l \leq N$ tal que

$$\sum_{i=n+1}^{n_j} a_i \geq c_1(n_j - n), \quad \text{para} \quad 0 \leq n \leq n_j - 1 \quad \text{e} \quad j = 1,...,l.$$

Assim, para todo $0 \leq n \leq n_j - 1$,

$$\sum_{i=n+1}^{n_j} a_i = -\log \prod_{i=n+1}^{n_j} ||Df(f^i(x))^{-1}|| \geq 2c(n_j - n) \tag{2.7}$$

fazendo $k = n_j - n$, temos $1 \leq k \leq n_j$ e por (2.7) obtemos

$$\prod_{m=0}^{k-1} ||Df(f^{n_j-m}(x))^{-1}|| \leq e^{-2ck} \quad \text{para todo} \quad 1 \leq k \leq n_j.$$

Logo, $n_1,...,n_l$ são tempos hiperbólicos. Portanto, $\frac{1}{n}\#(H_x \cup [1,n]) > \theta_0 > 0$ para todo $n \in \mathbb{N}$, com isso, segue que a densidade dos tempos hiperbólicos é positiva. □

Observação 2.3.4. A densidade que é, o lim inf em (2.5), é limitado por baixo pela constante θ_0 que depende apenas de f e nossa escolha de c.

No próximo lema, temos que os ramos inversos de iterados de f são contrações.

Lema 2.3.5. *Existe $\delta_0 > 0$, dependendo apenas de f e c, tal que dado qualquer tempo hiperbólico $n \geq 1$ para um ponto x em uma variedade riemaniana compacta M, e dado* *qualquer* $1 \leq j \leq n$, *o ramo inverso* $f_{x,n}^{-j}$ *de* f^j *que leva* $f^n(x)$ *à* $f^{n-j}(x)$ *é definido na bola de raio* δ_0 *centrada em* $f^n(x)$, *e satisfaz*

$$d(f_{x,n}^{-j}(z), f_{x,n}^{-j}(w)) \leq e^{-cj} d(z,w)$$

para todo z,w *na bola de raio* δ_0 *centrada em* $f^n(x)$.

Antes de começarmos a prova do lema, faremos a seguinte afirmação.

Como f é um difeomorfismo de classe C^1 e M é compacta, podemos tomar $\varepsilon_0 > 0$ de modo que para todo ponto $x \in M$ tenhamos $f|_{B_{\varepsilon_0}(x)}$ um difeomorfismo. Além disso, dados $x, y \in M$ com $y \in B_{\varepsilon_0}(x)$, temos

$$\frac{||Df(y)^{-1}||}{||Df(x)^{-1}||} \leq e^c.$$

De fato, defina $f : M \times M \to \mathbb{R}$ por

$$f(x,y) = \frac{||Df(y)^{-1}||}{||Df(x)^{-1}||},$$

facilmente vemos que f está bem definida e contínua, logo uniformemente contínua. Assim, dado $\theta > 0$, existe $\varepsilon > 0$ tal que se $d(x,y) + d(r,s) = d((x,r);(y,s)) < \varepsilon$ então

$$|f(x,r) - f(y,s)| < \theta.$$

Daí tomando ε_0 suficientemente pequeno, segue que

$$d(x,y) = d((x,y);(y,y)) < \varepsilon \quad \Rightarrow \quad 1 - \theta < f(x,y) < 1 + \theta.$$

Portanto, a afirmação é verdadeira.

Agora demostraremos o lema acima.

Demonstração. Seja $n \geq 1$ qualquer tempo hiperbólico para um ponto $x \in M$, então dado qualquer $1 \leq j \leq n$, temos

$$\prod_{k=1}^{j} ||Df(f^{n-k}(x))^{-1}|| \leq e^{-2cj}.$$

A prova será por indução sobre j. Logo, para $j = 1$ note que, dados $z, w \in B_{\delta_0}(f^n(x))$, onde δ_0 o mínimo entre o δ dado pela *Observação* 4.3.2, e o ε_0 da afirmação acima, segue que

$$d(f_{x,n}^{-1}(z), f_{x,n}^{-1}(w)) \leq e^{-c} d(z,w).$$

Com efeito, seja γ a geodésica ligando z a w com $\gamma([0,1]) \subset B_{\delta_0}(f^n(x))$. Assim, $\beta(t) = f_{x,n}^{-1}(\gamma(t))$ é uma curva diferenciável que liga $f_{x,n}^{-1}(z)$ a $f_{x,n}^{-1}(w)$ tal que

$$d(f_{x,n}^{-1}(z), f_{x,n}^{-1}(w)) \leq l(\beta),$$

mas

$$l(\beta) = \int_0^1 ||\beta'(t)||dt \leq \int_0^1 ||Df(\gamma(t))^{-1}||||\gamma'(t)||dt,$$

como $\gamma(t) \in B_{\delta_0}(f^n(x))$ para todo $t \in [0,1]$, temos

$$||Df(\gamma(t))^{-1}|| \leq e^c||Df(f^n(x))^{-1}|| \leq e^{-c}.$$

Assim vale para $j = 1$. Suponha que valha para $j-1$, e usando a mesma ideia veremos que

$$d(f_{x,n}^{-j}(z), f_{x,n}^{-j}(w)) \leq e^{-c}d(f_{x,n}^{j-1}(z), f_{x,n}^{j-1}(w)) \leq e^{-cj}d(z,w).$$

Isso prova o lema.

□

O próximo lema será útil quando provamos que f admiti uma partição geradora.

Lema 2.3.6. *Dada uma medida ergódica invariante ν cujos os expoentes de Lyapunov são todos maiores que $8c$, existe $N \in \mathbb{N}$ tal que f^N tem densidade de tempos hiperbólicos positiva para ν-quase todo ponto.*

Demonstração. Desde que todos expoente de Lyapunov de ν são maiores do que $8c$, para quase todo $x \in M$ existe $n_0(x) \geq 1$ tal que

$$||Df^n(x)v|| \geq e^{6cn}, \text{para todo } \ v \in T_xM, \text{com } \ ||v|| = 1, \text{e } \ n \geq n_0(x).$$

Em particular,

$$||Df^n(x)w|| \geq e^{6cn}||w||, \ \ \text{para} \ \ \text{todo} \ \ w \in T_xM \ \ \text{e} \ \ n \geq n_0(x).$$

Em outras palavras,

$$||Df^n(x)^{-1}|| \leq e^{-6cn}, \ \ \text{para} \ \ \text{todo} \ \ n \geq n_0(x).$$

Defina $\alpha_n = \nu(\{x; n_0(x) > n\})$. Desde que f é um difeomorfismo local de classe C^1, podemos também fixa uma constante $K > 0$ tal que $\log||Df(x)^{-1}|| \leq K$ para todo $x \in M$.

Então

$$\begin{aligned}\int_M \log||Df^n(x)^{-1}||d\nu &= \int_{\{x:n\geq n_0(x)\}} \log||Df^n(x)^{-1}||d\nu + \int_{\{x:n_0(x)>n\}} \log||Df^n(x)^{-1}||d\nu \\ &\leq -(6c - K\alpha_n)n.\end{aligned}$$

Desde $\alpha_n \to 0$ quando $n \to +\infty$, escolhendo N suficientemente grande garantimos que

$$\int_M \frac{1}{N} \log||Df^N(x)^{-1}||d\nu < -4c < 0.$$

Então, como ν é ergódica para f^N, pois seja A um conjunto f^N-invariante com medida positiva, então considere $A_0 = \{x \in A : f^k(x) \in A \text{ frenquência infinita}\}$ que, pelo Teorema 4.4.9, tem medida igual a medida de A. Como A_0 é f-invariante e ν é para f, segue que A tem medida total, portanto, A também tem. Logo,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \frac{1}{N} \log||Df^N(f^{kN}(y))^{-1}|| = \int_M \frac{1}{N} \log||Df^N(x)^{-1}||d\nu < -4c.$$

Isso significa que podemos aplicar Proposição 2.3.2 para concluir. □

Segundo a *Observação* 2.3.4, temos que a densidade de tempos hiperbólicos é limitada por baixo por uma constante que depende apenas de f^N e nossa escolha de c.

Lema 2.3.7. *Sejam $B \subset M, \theta > 0$ e $g : M \to M$ um difeomorfismo local tal que g tem densidade de tempos hiperbólicos maior que 2θ para todo $x \in B$. Então, dado qualquer medida de probabilidade ν em B e qualquer $m \geq 1$, existe $n > m$ tal que*

$$\nu(\{x \in B : n \textit{ é tempo hiperbólico de } g \textit{ para } x\}) > \theta.$$

Demonstração. Defina H como o conjunto dos pares $(x,n) \in B \times \mathbb{N}$ tal que n é um tempo hiperbólico para x. Ou seja, $H \cap (\{x\} \times \mathbb{N}) = H_x$. Para cada $k \geq 1$, seja $\mathcal{X}_k$ a medida de contagem no intervalo $[m+1, m+k]$. A hipótese implica que, dado quaisquer $x \in B$, temos

$$\mathcal{X}_k(H_x) > 2\theta \tag{2.8}$$

para todo k suficientemente grande. De fato,

$$\frac{1}{n}\#(H_x \cap [1,n]) = \frac{1}{n}\#(H_x \cap [1,m]) + \frac{1}{n}\#(H_x \cap [m+1,n]).$$

Tomando $k = n - m$, segue que

$$\frac{1}{n}\#(H_x \cap [1,n]) = \frac{1}{n}\#(H_x \cap [1,m]) + \frac{1}{n}\#(H_x \cap [m+1,n])$$

$$= \frac{1}{n}\#(H_x \cap [1,m]) + \frac{n-m}{n}\mathcal{X}_k(H_x).$$

Logo,

$$2\theta < \liminf \frac{1}{n}\#(H_x \cap [1,n])) = \liminf \mathcal{X}_k(H_x). \tag{2.9}$$

Isso pelo seguinte lema.

Lema 2.3.8. *Se $a_n = b_n + c_n$, onde $b_n \to 0$ com a_n e c_n limitadas inferiormente. Então $\liminf a_n = \liminf c_n$.*

Com efeito, como $\liminf a_n = \liminf(b_n + c_n) \geq \liminf c_n$. Por outro lado, $c_n = a_n - b_n$ segue que $\liminf c_n \geq \liminf a_n - \limsup b_n = \liminf a_n$.
Isso demonstra o lema.

De (2.9), temos que existe $k_0 = k_0(x)$ tal que para todo $k > k_0$ valha

$$\mathcal{X}_k(H_x) > 2\theta.$$

Agora, fixe $k \geq 1$ de modo que que valha (2.8) para um subconjunto C de pontos $x \in B$ com $\nu(C) > 1/2$, isso porque, considerando $B_k = \{x \in B; \mathcal{X}_k(H_x) > 2\theta\}$ segue que $B_k \subset B_{k+1}$ e $B = \cup_k B_k$, portanto $\lim_{k\to\infty} \nu(B_k) \to 1$. Considere, $\mathcal{X}_H : B \times \mathbb{N} \to \mathbb{R}$. Logo, pelo teorema de Fubini,

$$(\nu \times \mathcal{X}_k)(H) > \int_C \int \mathcal{X}_H(x,n) d\mathcal{X}_k d\nu = \int_C \mathcal{X}_k(H_x) d\nu > \int_C 2\theta d\nu > \theta,$$

isso implica que

$$\nu(H \cap (B \times \{n\})) > \theta$$

para algum $n \in [m+1, m+k]$. Caso contrário, temos

$$\iint \mathcal{X}_H(x,n) d\nu d\mathcal{X}_k = \int \nu(H \cap (B \times \{n\})) d\mathcal{X}_k \leq \theta.$$

Absurdo, pois $(\nu \times \mathcal{X}_k)(H) = \iint \mathcal{X}_H(x,n) d\nu d\mathcal{X}_k$. Portanto, existe $n \in [m+1, m+k]$ tal que $\nu(H \cap (B \times \{n\})) > \theta$. $\square$

2.4 Partições geradoras

Andes de provarmos a existência de uma partição geradora, daremos algumas informações úteis para o que segue.

Definição 2.4.1. Uma medida boreliana ν num espaço topológico é regular se para todo subconjunto B e todo $\varepsilon > 0$ existe um conjunto fechado F e um conjunto aberto A tais que $F \subset B \subset A$ e $\nu(F \backslash A) < \varepsilon$.

Lema 2.4.2. *Toda medida de probabilidade em um espaço métrico M é regular.*

Demonstração. Seja $\mathcal{B}_0$ a família dos subconjunto borelianos B tais que para todo $\varepsilon > 0$ existe um fechado F_ε e um aberto A_ε satisfazendo $F_\varepsilon \subset B \subset A_\varepsilon$ e $\nu(F_\varepsilon \backslash A_\varepsilon) < \varepsilon$. Mostraremos que $\mathcal{B}_0$ é uma σ-álgebra. Claramente $M \in \mathcal{B}_0$. Seja $B \in \mathcal{B}_0$; mostraremos que $B^c \in \mathcal{B}_0$. se $\varepsilon > 0$ existem F_ε fechado, A_ε aberto com $F_\varepsilon \subset B \subset A_\varepsilon$ tal que $\nu(F_\varepsilon \backslash A_\varepsilon) < \varepsilon$. Assim $A_\varepsilon^c \subset B^c \subset F_\varepsilon^c$ e $A_\varepsilon^c \backslash F_\varepsilon^c = A_\varepsilon \backslash F_\varepsilon$, então $\nu(A_\varepsilon^c \backslash F_\varepsilon^c) = \nu(A_\varepsilon \backslash F_\varepsilon) < \varepsilon$. Portanto $B \in \mathcal{B}_0$.

Agora mostraremos que $\mathcal{B}_0$ é fechado sobre a união enumerável. Sejam $B_1, B_2, ... \in \mathcal{B}_0$ e seja $B = \cup_i^\infty B_i$. Seja $\varepsilon > 0$ dado. Existem abertos $A_{\varepsilon,i}$, fechados $F_{\varepsilon,i}$ tais que $F_{\varepsilon,i} \subset B_i \subset A_{\varepsilon,i}$ e $\nu(A_{\varepsilon,i} \backslash F_{\varepsilon,i}) < \varepsilon/3^i$. Sejam $A_\varepsilon = \cup_{i=1}^\infty A_{\varepsilon,i}$(que é aberto), $\tilde{F}_\varepsilon = \cup_{i=1}^\infty F_{\varepsilon,i}$, e escolhendo k tal que $\nu(\tilde{F}_\varepsilon \backslash \cup_{i=1}^k F_{\varepsilon,i}) < \varepsilon/2$. Seja $F_\varepsilon = \cup_{i=1}^k F_{\varepsilon,i}$(que é fechado). Temos $F_\varepsilon \subset B \subset A_\varepsilon$. Assim,

$$
\begin{aligned}
\nu(A_\varepsilon \backslash F_\varepsilon) &\leq \nu(A_\varepsilon \backslash \tilde{F}_\varepsilon) + \nu(\tilde{F}_\varepsilon \backslash F_\varepsilon) \\
&\leq \sum_{i=1}^\infty \nu(A_{\varepsilon,i} \backslash F_{\varepsilon,i}) + \nu(\tilde{F}_\varepsilon \backslash F_\varepsilon) \\
&\leq \sum_{i=1}^\infty \frac{\varepsilon}{3^i} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}
$$

Portanto, $\mathcal{B}_0$ é uma σ-álgebra.

Para completar a prova mostraremos que $\mathcal{B}_0$ contém todos os subconjuntos fechados de M. Seja B um conjunto fechado e $\varepsilon > 0$. Considere $A_n = \{x \in M : d(B,x) < 1/n\}$ um conjunto aberto, então $A_1 \supset A_2 \supset \cdots \supset A_n \supset \cdots$ e $\cap_{i=1}^\infty A_i = B$. Escolha k tal que $\nu(A_k \backslash B) < \varepsilon$, seja $F_\varepsilon = B$ e $A_\varepsilon = A_k$. Isso mostra que $B \in \mathcal{B}_0$. □

Em todo que segue a constante $\delta_0 > 0$ está fixada como dada no Lema 2.3.5.

Lema 2.4.3. *Se ν é uma medida ergódica invariante tal que todos os seus expoente de Lyapunov são maiores que $8c$, e α é uma partição com diâmetro menor que δ_0. Então para ν-quase todo ponto $x \in M$, o diâmetro de $\alpha_n(x)$ vai a zero quando n vai a infinito. Em particular, α é uma partição f-geradora com respeito à ν.*

Demonstração. Pelo Lema 2.3.6 existe $N \geq 1$ tal que f^N tem densidade de tempos hiperbólicos positivo para ν-quase todo ponto. Defina

$$\gamma_k = \bigvee_{j=0}^{k-1} f^{-jN}(\alpha) \quad \text{para} \quad \text{cada} \quad k \geq 1.$$

Pelo Lema 2.3.5, se k é um tempo hiperbólico de f^N para x então $\operatorname{diam}\gamma_k(x) \leq e^{-ck}$. Em particular, uma vez que os conjuntos $\gamma_k(x)$ são não-crescentes com k, o diâmetro de $\gamma_k(x)$ vai a zero quando $k \to \infty$. Visto que $\alpha_{kN}(x) \subset \gamma_k(x)$ e a sequência $\operatorname{diam}\alpha_n(x)$ é não-crescente, isso imediatamente nos dá que o diâmetro de $\alpha_n(x)$vai a zero quando n vai ao infinito, para ν-quase todo ponto $x \in M$. Pela prova do Corolário 1.2.9.1, temos que a álgebra $\mathcal{A}$ gerada por $\cup_n \alpha_n$ gera a σ-álgebra de Borel. Isso completa a prova. □

2.5 Fórmula de Rokhlin

O próximo teorema será útil para mostrarmos a existência do jacobiano para um difeomorfismo local em uma variedade riemaniana compacta.

Teorema 2.5.1 (Teorema fundamental de seções transversais de Rokhlin)**.** *Seja M um espaço métrico completo e separável. Suponha que $\mathcal{A}$ e $\mathcal{B}$ são duas partições mensuráveis de um espaço de Lebesgue $(M, \mathcal{F}, \nu)$ tal que para ν_β-quase toda fibras B em $\mathcal{B}$, $\mathcal{A} \cap B = \{A \cap B : A \in \mathcal{A}\}$ é enumerável para ν_B . Então existe uma partição enumerável $\gamma = \{\gamma_1, \gamma, ...\}$ de M tal que cada $\gamma_j \in \gamma$ intersecta quase todo B em no máximo um ponto* $(\mod 0)$ *para ν_B, que é então um átomo de ν_B.*

Demonstração. A prova desse fato está em (4). □

Lema 2.5.2. *Se $f : M \to M$ é um difeomorfismo local num espaço métrico compacto, ν uma medida de probabilidade f-invariante. Então ν é não-singular com respeito à f.*

Demonstração. Sejam $\mathcal{A} = \varepsilon$ e $\mathcal{B} = f^{-1}(\varepsilon)$ duas partições mensurável, onde ε representa a partição por pontos. Como $\mathcal{A} \cap B$ é enumerável, pois $f^{-1}(x)$ é finita, temos, pelo Teorema 2.5.1, que existe uma partição $\gamma = \{\gamma_1, \gamma_2, ...\}$ tal que cada $\gamma_j \in \gamma$ intersecta $\nu_{\mathcal{B}}$-quase todo B em no máximo um ponto para ν_B, que é então um átomo de ν_B. Considere $Y \subset M$,

$$Y = \bigcap_{n \geq 0} f^{-n}(\cup_{j \geq 1} \gamma_j).$$

Note que, $f(Y) \subset Y$ e para todo $x \in Y$, $f^{-1}(x) \cap Y$ consiste apenas de átomos da medida condicional $\nu_{f^{-1}(x)}$. Considere $(Y_\mathcal{A}, \mathcal{F}_\mathcal{A}, f_\mathcal{A})$ (como no Apêndice). Assim, para $B \subset Y$

mensurável, o conjunto

$$\{A \in \mathcal{A} : \nu_A(B \cap A) \neq 0\} = \{A \in \mathcal{A} : A \cap B \neq \emptyset\} \tag{2.10}$$

é mensurável em $\mathcal{F}_{\mathcal{A}}$ e assim sua imagem por $f_{\mathcal{A}}$, igual $f(B)$, é mensurável. Se $\nu(B) = 0$, então o conjunto em (2.10) tem medida $\nu_{\mathcal{A}}$ igual a 0, como $f_{\mathcal{A}}$ é isomorfismo, obtemos que $f(B)$ é mensurável e tem medida 0. Portanto, ν é não-singular. □

Seja $f : M \to M$ uma transformação mensurável, ν é uma probabilidade invariante. Suponha que existe uma partição finita ou enumerável α de M tal que

(a) f é localmente injetiva, o que significa que é injetiva em todos os átomos de α;

(b) α é f-geradora com respeito à ν, no sentido que $\operatorname{diam} \alpha(x) \to 0$ para ν-quase todo ponto $x \in M$.

Teorema 2.5.3 (Versão fraca da Fórmula de Rokhlin)**.** *Seja $f : M \to M$ uma difeomorfismo local num espaço métrico compacto e seja ν uma probabilidade invariante por f. Suponha que existe uma partição finita ou enumerável α satisfazendo* (a) *e* (b)*. Então*

$$h_\nu(f) = \int \log J_\nu f \, d\nu.$$

Demonstração. Basta observar que a partição α satisfaz a hipótese do Teorema 1.3.4. □

3 RESULTADO PRINCIPAL

Aqui sempre consideraremos M uma variedade riemaniana de classe C^1, compacta, conexa e d-dimensional.

Teorema 3.0.4. *Seja $f: M \to M$ um difeomorfismo local de classe C^1 e grau $p \geq 1$. Se f satisfaz:*

$$\max_{1 \leq k \leq d-1} \max_{x \in M} \limsup_{n \to \infty} \frac{1}{n} \log ||(Df_x^n)^{\wedge_k}|| < \log p. \tag{3.1}$$

Então $h_{top}(f) = \log p$, e qualquer automedida maximal μ do dual do operador de transferência $\mathcal{L}$ é uma medida maximizante. Em particular, existe alguma medida maximizante para f. Se f é topologicamente mixing então a medida maximizante é única e positiva nos abertos.

3.1 Existência

Aqui provaremos que toda automedida maximal é uma medida maximizante. O primeiro passo é:

Lema 3.1.1. *Se μ é uma automedida maximal então $J_\mu f$ é constante igual à p.*

Demonstração. Seja A qualquer conjunto mensurável tal que $f|_A$ é injetiva. Tome uma sequência $\{g_n\} \in C(M)$ tal que $g_n \to \mathcal{X}_A$ em μ-quase todo ponto e $\sup |g_n| \leq 2$ para todo n. Por definição,

$$\mathcal{L}g_n(x) = \sum_{y: f(y) = x} g_n(y).$$

Observe que a expressão passada converge para $\mathcal{X}_{f(A)}$ em μ-quase todo ponto. Assim, pelo teorema da convergência dominada,

$$\int p g_n d\mu = \int g_n d(\mathcal{L}^* \mu) = \int \mathcal{L} g_n d\mu \to \mu(f(A)).$$

Por outro lado, temos que $\int_A p d\mu$, concluímos que

$$\mu(f(A)) = \int_A p d\mu.$$

□

O seguinte lema será útil para concluirmos que a integral do expoente de Lyapunov é positivo.

Lema 3.1.2. *Se μ é uma automedida maximal então $h_\mu \geq \log p$.*

Demonstração. Definimos a bola dinâmica $\mathcal{B}_\epsilon(n,x)$ por

$$\mathcal{B}_\epsilon(n,x) = \{y \in M; d(f^i(x), f^i(y)) < \epsilon \;\text{ para }\; i = 0,...n-1\}.$$

Tome ϵ é pequeno o suficiente para que $f|_{f^j(\mathcal{B}_\epsilon(n,x))}$ seja injetiva para $j = 0,...,n-1$ e para todo n. Isso é possível, pois basta tomar ϵ de modo que 2ϵ seja menor que o número de Lebesgue de uma cobertura $\{U_k\}_{k=1}^N$ de M cuja $f|_{U_k}$ é injetiva para todo $k = 1,...,N$. Assim, como $f^j(\mathcal{B}_\epsilon(n,x))$ tem diâmetro menor que 2ϵ, temos $f^j(\mathcal{B}_\epsilon(n,x)) \subset U_k$ para algum $1 \leq k \leq N$. Então tomando ϵ como acima e pelo fato de que $J_\mu f = p$, segue-se

$$1 = \mu(M) \geq \mu(f^n(\mathcal{B}_\epsilon(n,x))) = p^n \mu(\mathcal{B}_\epsilon(n,x)).$$

Em particular, podemos concluir que

$$\log p \leq -\frac{1}{n} \log \mu(\mathcal{B}_\epsilon(n,x)), \forall n \in \mathbb{N}$$

$$\log p \leq \liminf_n (-1/n) \log \mu(\mathcal{B}_\epsilon(n,x)) = -\limsup_n (1/n) \log \mu(\mathcal{B}_\epsilon(n,x)),$$

para todo ϵ pequeno. Pela formula local da entropia de Brin-Katok 4.7.1(Apêndice), temos

$$h_\mu(f) = -\int \lim_{\epsilon \to 0} \limsup_n (1/n) \log \mu(\mathcal{B}_\epsilon(n,x)) d\mu(x) \geq \log p.$$

□

Corolário 3.1.2.1. *Toda automedida ergódica maximal μ tem entropia igual a $\log p$.*

Demonstração. Pelo Lema 3.1.2, a entropia é no mínimo $\log p$. Então, podemos aplicar o Lema 2.2.2, para concluir que todos os expoentes de Lyapunov de μ são positivos. Segue-se, pelo Lema 2.4.3, que μ admite uma partição geradora com diâmetro pequeno.

Portanto, podemos aplicar a Teorema 2.5.3 e o Lema 3.1.1, para encontrar que

$$h_\mu(f) = \int \log J_\mu d\mu = \log p.$$

□

No seguinte lema mostraremos que a entropia topológica é igual a $\log p$.

Lema 3.1.3. *A entropia topológica $h_{top}(f) = \log p$. Além disso, se η é alguma medida ergódica maximizante então $J_\eta f$ é constante igual a p.*

Demonstração. Usando a afirmação (1.23), basta apenas considerar as probabilidades ergódicas. Seja η uma probabilidade ergódica tal que $h_\eta(f) \geq \log p$. Pelo Lema 2.2.2 todos expoentes de Lyapunov de η são maiores que $c(f)$. Então, pelo Lema 2.4.3, existe uma partição geradora com diâmetro arbitrariamente pequeno. Isso garante que podemos aplicar Teorema 2.5.3 para η, e então obter que

$$h_\eta(f) = \int \log J_\eta(f) d\eta. \tag{3.2}$$

Vamos escrever $g_\eta = 1/(J_\eta(f))$. Como por hipótese η é f-invariante, segue-se que

$$\sum_{y:f(y)=x} g_\eta(y) = 1$$

para η-quase todo ponto $x \in M$. Da igualdade (3.2) e usando o Lema 4.4.11, temos

$$0 \leq h_\eta(f) - \log p = \int \log \frac{p^{-1}}{g_\eta} d\eta = \int \sum_{y:f(y)=x} g_\eta(y) \log \frac{p^{-1}}{g_\eta(y)} d\eta. \tag{3.3}$$

Agora, usaremos o seguinte fato elementar de cálculo:

Lema 3.1.4. *Sejam p_i, x_i números reais positivos para, $i = 1,...n$, tal que $\sum_{i=1}^{n} p_i = 1$. Então $\sum_{i=1}^{n} p_i \log x_i \leq \log\Big(\sum_{i=1}^{n} p_i x_i\Big)$ e a igualdade vale se, e somente se, os x_i são todos iguais.*

Tome $p_i = g_\eta(y)$ e $x_i = p^{-1}/g_\eta$, obtemos

$$\sum_{y:f(y)=x} g_\eta(y) \log \frac{p^{-1}}{g_\eta(y)} \leq \log\Big(\sum_{y:f(y)=x} g_\eta(y) \frac{p^{-1}}{g_\eta(y)}\Big) = \log\Big(\sum_{y:f(y)=x} p^{-1}\Big) = 0 \tag{3.4}$$

em η-quase todo ponto. Desde que a integral é não-negativa, por (3.3), a igualdade é valida η-quase toda parte, e $h_\eta(f) - \log p = 0$. Como η é arbitrário, isso mostra que $h_{top}(f) = \log p$.

De (3.3) e (3.4), obtemos a seguinte igualdade:

$$\sum_{y:f(y)=x} g_\eta(y) \log \frac{p^{-1}}{g_\eta(y)} = \log\Big(\sum_{y:f(y)=x} g_\eta(y) \frac{p^{-1}}{g_\eta(y)} \Big) = 0 \quad \text{para} \quad \eta-\text{q.t.p.} \quad x \in M. \tag{3.5}$$

Assim, pelo Lema 3.1.4, segue que $p^{-1}/g_\eta(y)$ são o mesmo para todos $y \in f^{-1}(x)$. Em outras palavras, para η-quase todo $x \in M$ existe um número $c(x)$ tal que $p^{-1}/g_\eta(y) = c(x)$ para todo $y \in f^{-1}(x)$. Então

$$\frac{1}{c(x)} = \sum_{y \in f^{-1}(x)} \frac{p^{-1}}{c(x)} = \sum_{y \in f^{-1}(x)} g_\eta(y) = 1$$

para η-quase todo ponto $x \in M$. Isso significa, precisamente, que $J_\eta f(y) = p$ para todo $y \in f^{-1}(x)$ e x pertencente a um conjunto de medida η total. $\square$

Isso prova a existência de uma medida maximizante.

3.2 Unicidade

Nesta seção assumiremos f topologicamente mixing, e concluiremos que a medida maximizante é única e suportada em todo ambiente M. È suficiente considerar as medidas ergódicas. Porque, pela observação feita do Teorema de Jacobs 4.5.2(Apêndice), as componentes ergódicas de medidas maximizante também são medidas maximizantes.

Lema 3.2.1. *Qualquer medida maximizante ergódica μ é suportada nos abertos de M.*

Demonstração. Suponha $\mu(U) = 0$ para algum conjunto aberto não vazio U. Pela hipótese mixing, existe $N \geq 1$ tal que $f^N(U) = M$. Partindo U em subconjuntos $U_1, ..., U_k$ de modo que $f|_{U_j}$ seja injetiva, obtemos que

$$\mu(f^N(U_j)) = \int_{U_j} J_\mu f^N d\mu = 0$$

para todo $j = 1, ..., k$. Pelo Lema 3.1.3, temos que $J_\mu f^N$ é constante. Isso implica que $\mu(f^N(U)) = 0$, que é uma contradição. $\square$

Isso tem a seguinte consequência útil: dada qualquer $\delta > 0$ existe $b = b(\delta)$ tal que

$$\mu(B(x,\delta)) \geq b \quad \text{para todo} \quad x \in M. \tag{3.6}$$

De fato, suponha que existem $\delta > 0$ e uma sequência $(x_n)_{n\geq 1}$ tal que $\mu(B(x_n,\delta)) < 1/n$ então, considerando um ponto de acumulação x e uma bola $B(x,\delta/2)$, temos que passando a uma subsequência se necessário todas as bolas $B(x_n,\delta)$ contém a bola $B(x,\delta/2)$. Daí segue, $\mu(B(x,\delta)) = 0$. O que contradiz o Lema 3.2.1.

Agora sejam μ_1 e μ_2 quaisquer duas medidas maximizantes ergódicas. Nosso objetivo é provar que as duas medidas coincidem. Como um primeiro passo provaremos que elas são equivalentes. Para isso, fixemos quaisquer partição $\mathcal{P}$ finita de M tal que os elementos $P \in \mathcal{P}$ tem interior não-vazio, e o bordo ∂P tem medida zero para ambas μ_1 e μ_2. Isso é possível, pois dado $x \in M$ temos uma quantidade não enumerável para $\partial B(x,\varepsilon)$. Logo, $\mu(\partial B(x,\varepsilon)) > 0$ no máximo para uma quantidade enumerável de ε, portanto, podemos pegar tal partição e cada elemento dessa partição tenha diâmetro tão pequeno quanto desejado. Fixando $\delta > 0$ pequeno o suficiente de modo que todo $P \in \mathcal{P}$ contenha alguma bola de raio δ, usando (3.6) para ambas medidas, concluímos que existe $a > 0$ tal que

$$\mu_1(P) \leq a\mu_2(P) \quad \text{e} \quad \mu_2(P) \leq a\mu_1(P) \quad \text{para todo} \quad P \in \mathcal{P}. \tag{3.7}$$

Agora seja g um ramo inverso de qualquer iterado f^n, $n \geq 1$. Usando Lema 3.1.3, obtemos

$$\mu_i(P) = \mu_i(f^n(g(P))) = \int_{g(P)} J_{\mu_i} f^n d\mu_i = p^n \mu_i(g(P))$$

para $i = 1, 2$. Por (3.7), segue que

$$\mu_1(g(P)) \leq a\mu_2(g(P)) \quad \text{e} \quad \mu_2(g(P)) \leq a\mu_1(g(P)) \tag{3.8}$$

para todo $P \in \mathcal{P}$ e todo ramo inverso g de f^n, para quaisquer $n \geq 1$. Denotamos por $\mathcal{Q}$ a família de todos os $g(P)$.

Lema 3.2.2. *Dados quaisquer conjunto mensurável $E \subset M$ e $\varepsilon > 0$ existe uma família $\mathcal{E}$ de elementos disjuntos de $\mathcal{Q}$ tal que*

$$\mu_i\Big(E \setminus \bigcup_{g(P)\in\mathcal{E}} g(P)\Big) = 0 \ \ e \ \ \mu_i\Big(\bigcup_{g(P)\in\mathcal{E}} g(P) \setminus E\Big) \leq \varepsilon \ \ para \ i = 12.$$

Demonstração. Pelo Lema 2.2.2, todos expoentes de Lyapunov de μ_i maiores que $c(f)$.

Pelo Lema 2.3.6, existem $N \geq 1$ e $\theta > 0$, com θ dependendo apenas de f^N e da escolha de c, tal que μ_i-quase todo ponto tem densidade de tempos hiperbólicos para f^N maior que 2θ.

Sejam U_1 um conjunto aberto e K_1 um conjunto compacto tais que $K_1 \subset E \subset U_1$ e $\mu_i(U_1 \backslash E) \leq \varepsilon$ para $i = 1, 2$, com $\mu_i(K_1) \geq \frac{1}{2}\mu_i(E)$. Usando Lema 2.3.7, com $B = K_1$ e $\nu = \mu_i/\mu_i(K_1)$, podemos encontrar $n_1 \geq 1$ tal que $e^{-cn_1} < d(K_1, U_1^c)$ e o subconjunto L_1 dos pontos $x \in K_1$ para os quais n_1 é um tempo hiperbólico para f^N satisfazendo $\mu_i(L_1) \geq \theta\mu_i(K_1) \geq (\theta/2)\mu_i(E)$. Seja $\mathcal{E}_1$ a família de todos $g(P)$ que intersecta L_1, com $P \in \mathcal{P}$ e g um ramo inverso de f^{Nn_1}. Note que os elementos de $\mathcal{E}_1$ são dois a dois disjuntos, porque os elementos de $\mathcal{P}$ são dois a dois disjuntos. Além disso, pelo Lema 2.3.5, seus diâmetros são menores que e^{-cn_1}. Assim, a união E_1 de todos os elementos de $\mathcal{E}_1$ está contido em U_1. Por construção, temos

$$\mu_i(E_1 \cap E) \geq \mu_i(L_1) \geq \theta\mu_i(K_1) \geq (\theta/2)\mu_i(E).$$

Agora, considere o conjunto aberto $U_2 = U_1 \backslash \bar{E}_1$, seja $K_2 \subset E \backslash \bar{E}_1$ um conjunto compacto tal que $\mu_i(K_2) \geq (1/2)\mu_i(E \backslash E_1)$. Observe que $\mu_i(\bar{E}_1 \backslash E_1) = 0$, pois as fronteiras dos átomos de $\mathcal{P}$ tem medida zero e elas são preservadas pelos ramos inversos, desde de que μ_1, μ_2 são invariantes. Seguindo o mesmo raciocínio, podemos encontrar $n_2 > n_1$ tais que $e^{-cn_2} < d(K_2, U_2^c)$ e um conjunto $L_2 \subset K_2$ dos pontos x tal que n_2 é tempo hiperbólico para f^N com $\mu_i(L_2) \geq \theta\mu_i(K_2)$. Denote por $\mathcal{E}_2$ a família de imagens inversas $g(P)$ que intersecta L_2, com $P \in \mathcal{P}$ e g um ramo inverso de f^{Nn_2}. Como antes, os elementos de $\mathcal{E}_2$ são dois a dois disjuntos, seus diâmetros são menores que e^{-cn_2}. Logo, sua união E_2 está contida em U_2. Consequentemente, os elementos da união $\mathcal{E}_1 \cup \mathcal{E}_2$ são também dois a dois disjuntos. Além disso,

$$\mu_i(E_2 \cap [E \backslash E_1]) \geq \mu_i(L_2) \geq \theta\mu_i(K_2) \geq (\theta/2)\mu_i(E \backslash E_1).$$

Repetindo esse processo, construímos a família $\mathcal{E}_k, k \geq 1$, de elementos de $\mathcal{Q}$ tais que seus elementos são todos dois a dois disjuntos, estão contidos em U_1 e

$$\mu_i(E_{k+1} \cap [E \backslash (E_1 \cup \cdots \cup E_k)]) \geq (\theta/2)\mu_i(E \backslash (E_1 \cup \cdots \cup E_k)) \tag{3.9}$$

para cada $k \geq 1$, onde $E_j = \bigcup_{g(P) \in \mathcal{E}_j} g(P)$. Assim, $\mu_i(\cup_{k=1}^{\infty} E_k \backslash E) \leq \mu_i(U_1 \backslash E) \leq \varepsilon$ para $i = 1, 2$, e (3.9) implica que

$$\mu_i(E \backslash \cup_{k=1}^{\infty} E_k) = 0.$$

Desde que,

$$[E\backslash(E_1\cup\cdots\cup E_{n+1})]\cup[E_{n+1}\cap(E\backslash(E_1\cup\cdots\cup E_n)]=E\backslash(E_1\cup\cdots\cup E_n).$$

Logo,

$$\mu_i\Big(E\backslash(E_1\cup\cdots\cup E_{n+1})\Big)\leq\Big(1-\frac{\theta}{2}\Big)^{n+1}\mu_i(E).$$

Isso completa a prova do lema, com $\mathcal{E}=\cup_{k=1}^{\infty}\mathcal{E}_k$.

$\square$

Logo, μ_1,μ_2 são probabilidades ergódicas absolutamente contínuas, pois se $\mu_1(E)=0$, pelo Lema 3.2.2, temos que

$$\begin{aligned}\mu_2(E)\leq\mu_2(\bigcup_{g(P)\in\mathcal{E}}g(P)) &\leq B\mu_1(\bigcup_{g(P)\in\mathcal{E}}g(P))\\ &= B(\mu_1(\bigcup_{g(P)\in\mathcal{E}}g(P)\bigcap E)+\mu_i(\bigcup_{g(P)\in\mathcal{E}}g(P)\backslash E))<\varepsilon.\end{aligned}$$

Como $\varepsilon>0$ é arbitrário, isso prova o desejado. Portanto, segue-se $\mu_1=\mu_2$. Isso prova a unicidade.

4 APÊNDICE

4.1 Variedades diferenciáveis; espaços tangentes

Definição 4.1.1. Uma variedade diferenciável de dimensão n é um conjunto M e uma família de aplicações injetivas $x_\alpha : U_\alpha \subset \mathbb{R}^n \to M$ de abertos U_α de $\mathbb{R}^n$ em M tais que:

(1) $\bigcup\limits_{\alpha} x_\alpha(U_\alpha) = M$;

(2) Para todo par α, β, com $x_\alpha(U_\alpha) \cap x_\beta(U_\beta) = W \neq \emptyset$, os conjuntos $x_\alpha^{-1}(W)$ e $x_\beta^{-1}(W)$ são abertos de $\mathbb{R}^n$ e as aplicações $x_\beta^{-1} \circ x_\alpha$ são diferenciáveis;

(3) A família $\{(U_\alpha, x_\alpha)\}$ é máxima relativamente às condições (1) e (2).

O par (U_α, x_α) com $p \in x_\alpha(U_\alpha)$ é chamado uma parametrização de M em p. Uma família $\{(U_\alpha, x_\alpha)\}$ satisfazendo (1) e (2) é chamada uma estrutura diferenciável em M.

Observação 4.1.2. Uma estrutura diferenciável em M induz de maneira natural uma topologia em M. $A \subset M$ é aberto de M se $x_\alpha^{-1}(A \cap x_\alpha(U_\alpha))$ é aberto em $\mathbb{R}^n$ para todo α. Denotaremos por M^n uma variedade de dimensão n. Dizemos que uma variedade é de classe $C^r, r \geq 1$, se existe uma família satisfazendo (1), (2) e (3) tal que $x_\beta^{-1} \circ x_\alpha$ é de classe C^r.

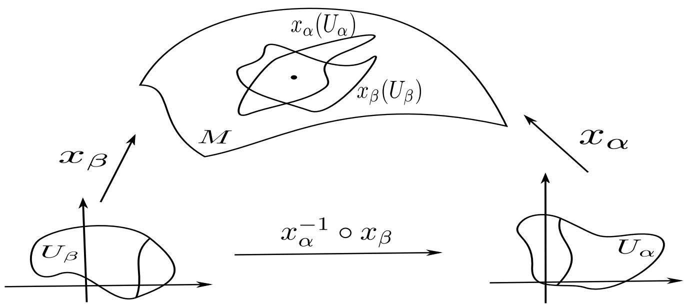

Figura 1: Variedade diferenciável

Definição 4.1.3. Sejam M_1^n e M_2^m variedades diferenciáveis. Uma aplicação $\varphi : M_1 \to M_2$ é diferenciável em $p \in M_1$ se dada uma parametrização $y : V \subset \mathbb{R}^m \to M_2$ em $\varphi(p)$ existe uma parametrização $x : U \to M_1$ em p tal que $\varphi(U) \subset y(V)$ e a aplicação

$$y^{-1} \circ \varphi \circ x : U \subset \mathbb{R}^n \to \mathbb{R}^m \tag{4.1}$$

é diferenciável em $x^{-1}(p)$.

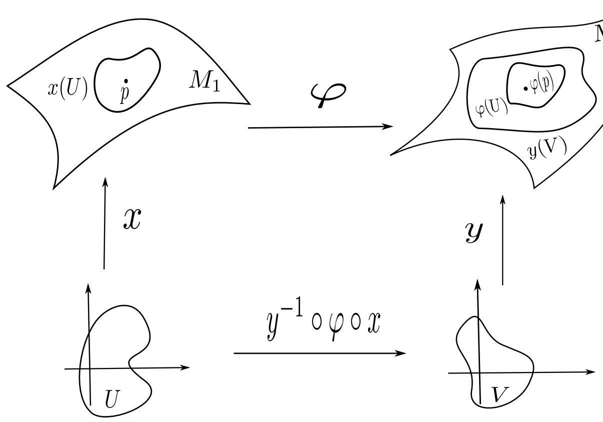

Figura 2: φ diferenciável em p

φ é diferenciável em um aberto de M_1 se é diferenciável em todos os pontos desse aberto.

Decorre, da condição (2) da Definição (4.1.1), que a definição acima independe das escolhas das parametrizações. A aplicação (4.1) é chamada expressão de φ nas parametrizações x e y. Sejam M_1 e M_2 variedades de classes $C^r, \geq 1$, a aplicação φ é dita de classe $C^k, k \leq r$ se sua expressão em x e y é de classe C^k.

Daremos agora uma noção de vetor tangente às variedades diferenciáveis.

Definição 4.1.4. Seja M uma variedade diferenciável. Uma aplicação diferenciável $\alpha : (-\varepsilon, \varepsilon) \to M$ é chamada uma curva (diferenciável) em M. Suponha $\alpha(0) = p \in M$, e seja $\mathcal{D}$ o conjunto das funções de M diferenciáveis em p. O vetor tangente à curva em $t = 0$ é a função $\alpha'(0) : \mathcal{D} \to \mathbb{R}$ dada por

$$\alpha'(0)f = \frac{d}{dt}(f \circ \alpha)(t)\bigg|_{t=0} \quad f \in \mathcal{D}.$$

Um vetor tangente em p é o vetor tangente em $t = 0$ de alguma curva $\alpha : (-\varepsilon, \varepsilon) \to M$ com $\alpha(0) = p$. O conjunto dos vetores à M em p será indicado por T_pM.

Se escolhermos uma parametrização $x : U \to M$ em $p = x(0)$, podemos exprimir a função f e a curva α nessa parametrização por

$$f \circ x(q) = f(x_1, ..., x_n), \quad q = (x_1, ..., x_n) \in U,$$

e

$$x^{-1} \circ \alpha(t) = (x_1(t), ..., x_n(t)),$$

respectivamente. Portanto, restringindo f à α, obteremos

$$\alpha'(0)f = \frac{d}{dt}(f \circ \alpha)(t)\Big|_{t=0} = \frac{d}{dt}f(x_1(t), ..., x_n(t))\Big|_{t=0}$$
$$= \sum_{i=1}^{n} x_i'(0)\Big(\frac{\partial f}{\partial x_i}\Big) = \Big(ds\sum_{i=1}^{n} x_i'(0)\Big(\frac{\partial}{\partial x_i}\Big)_0\Big)f.$$

Em outras palavras, o vetor $\alpha'(0)$ pode ser expresso na parametrização x por

$$\alpha'(0) = \sum_{i=1}^{n} x_i\Big(\frac{\partial}{\partial x_i}\Big)_0. \tag{4.2}$$

Observe que $\Big(\frac{\partial}{\partial x_i}\Big)_0$ é o vetor tangente em p à curva "curva coordenada":

$$x_i \to x(0, ..., 0, x_i, 0, ..., 0).$$

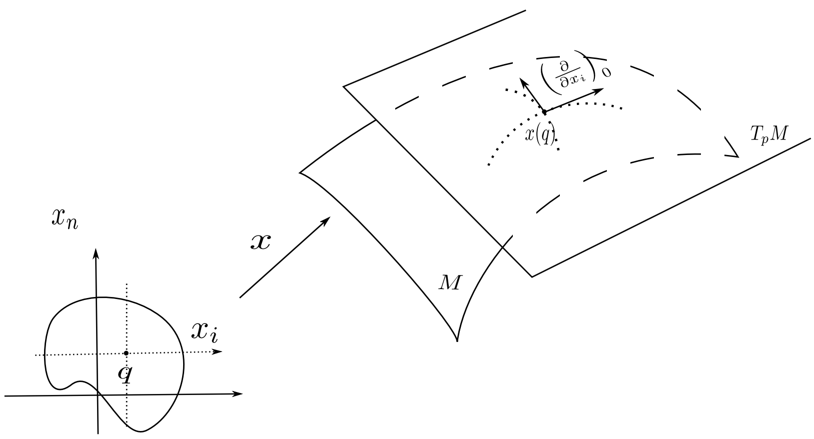

Figura 3: Curvas coordenadas e seus vetores tangentes

A expressão (4.2) mostra que o vetor tangente a uma curva α em p depende apenas das derivadas de α em um sistema de coordenadas. Decorre também de (4.2) que o conjunto T_pM, com as operações usuais de funções, forma um espaço vetorial de dimensão n, e que a escolha de uma parametrização $x : U \to M$ determina uma base associada

$\left\{\left(\frac{\partial}{\partial x_1}\right)_0,...,\left(\frac{\partial}{\partial x_n}\right)_0\right\}$ em T_pM (Fig.3). É imediato que a estrutura linear em T_pM assim definida não depende da parametrização x. O espaço vetorial TpM é chamado o espaço tangente de M em p.

Com a noção de espaço tangente podemos estender às variedades diferenciável a noção de diferencial de uma aplicação diferenciável.

Proposição 4.1.5. *Sejam M_1^n e M_2^m variedades diferenciáveis e seja $\varphi : M_1 \to M_2$ uma aplicação diferenciável. Para cada $p \in M_1$ e cada $v \in T_pM$, escolha uma curva diferenciável $\alpha : (-\varepsilon, \varepsilon) \to M_1$ com $\alpha(0) = p$, $\alpha'(0) = v$. Faça $\beta = \varphi \circ \alpha$. A aplicação $d\varphi_p : T_pM_1 \to T_{\varphi(p)}M_2$ dada por $d\varphi_p(v) = \beta'(0)$ é uma aplicação linear que não depende da escolha de α.*

Demonstração. Sejam $x : U \to M_1$ e $y : V \to M_2$ parametrizações em p e $\varphi(p)$, respectivamente. Exprimindo φ nestas parametrizações, podemos escrever

$$y^{-1} \circ \varphi \circ x(q) = (y_1(x_1,...,x_n),...,y_m(x_1,...,x_n))$$

$$q = (x_1,...,x_n) \in U, \quad (y_1,...,y_m) \in V.$$

Por outro lado, exprimindo α na parametrização x, obtemos

$$x^{-1} \circ \alpha(t) = (x_1(t),...,x_n(t)).$$

Portanto,

$$y^{-1} \circ \beta(t) = (y_1(x_1(t),...,x_n(t)),...,y_m(x_1(t),...,x_n(t))).$$

Decorre daí que a expressão de $\beta'(0)$ na base $\left\{\left(\frac{\partial}{\partial y_i}\right)_0\right\}$ de $T_{\varphi(p)}M_2$, associada à parametrização y, é dada por

$$\beta'(0) = \left(\sum_{i=1}^{n} \frac{\partial y_1}{\partial x_i} x_i'(0),...,\sum_{i=1}^{n} \frac{\partial y_m}{\partial x_i} x_i'(0)\right). \tag{4.3}$$

A relação (4.3) mostra que $\beta'(0)$ não depende da escolha de α. Além disso, (4.3) pode ser escrito como

$$\beta'(0) = d\varphi_p(v) = \left(\frac{\partial y_i}{\partial x_j}\right)(x_j'(0)), \quad i = 1,...,m; \quad j = 1,...,n,$$

onde $\left(\frac{\partial y_i}{\partial x_j}\right)$ indica uma matriz $m \times n$ e $x_j'(0)$ indica uma matriz coluna com n elementos. Portanto, $d\varphi_p$ é uma aplicação linear de T_pM_1 em $T_{\varphi(p)}M_2$ cuja matriz nas bases associadas às parametrizações x e y é precisamente a matriz $\left(\frac{\partial y_i}{\partial x_j}\right)$. $\square$

A aplicação $d\varphi_p$ é chamada diferencial de φ em p. Agora daremos à definição de difeomorfismo local.

Definição 4.1.6. Sejam M_1 e M_2 variedades diferenciáveis. Uma aplicação $\varphi : M_1 \to M_2$ é um difeomorfismo se ela é diferenciável, bijetiva e sua inversa φ^{-1} é diferenciável. φ é um difeomorfismo local se para todo ponto $p \in M_1$ existem vizinhanças U de p e V de $\varphi(p)$ tais que $\varphi : U \to V$ é um difeomorfismo.

Sejam M_1, M_2 variedades diferenciáveis de classes C^r. $\varphi : M_1 \to M_2$ é um difeomorfismo de classe $C^k, k \leq r$, se φ é de classe C^k e é um difeomorfismo. Analogamente, φ é um difeomorfismo local de classe C^k se para todo ponto $p \in M_1$ existem vizinhanças U de p e V de $\varphi(p)$ tais que $\varphi : U \to V$ é um difeomorfismo de classe C^k. Segue do teorema da aplicação inversa em $\mathbb{R}^n$, ver (3), que sendo φ um difeomorfismo de classe C^k então φ^{-1} também é de classe C^k.

4.2 Variedades Riemanniana

Uma métrica riemanniana numa variedade diferenciável M é uma correspondência que associa a cada ponto $p \in M$ um produto interno no espaço tangente T_pM.

Seja g uma métrica riemanniana em M. Indicamos por $\langle u,v \rangle_p$ ou simplesmente $\langle u,v \rangle$ o produto interno dos vetores $u,v \in T_pM$.

O comprimento ou norma do vetor tangente $u \in T_pM$ é definido obviamente por

$$|u| = \sqrt{\langle u,v \rangle}.$$

Uma variedade diferenciável onde está definida uma métrica riemanniana chama-se uma *variedade riemanniana.* Ou seja, trata-se de um par $(M, \langle \cdot,\cdot \rangle)$ onde $\langle \cdot,\cdot \rangle$ é uma métrica riemanniana na variedade M.

Observação 4.2.1. É possível definir uma métrica riemanniana de classe C^{k-1} em qualquer variedade $M \in C^k$.

4.3 A distância intrínseca

Numa variedade riemanniana M, faz sentido falar em muitos conceitos geométricos. Por exemplo, podemos definir o comprimento de um caminho $\alpha : [0,1] \to M$, de classe

C^1, pondo $l(\alpha) = \int_0^1 |\alpha'(t)|dt$. Nessa expressão, $|\alpha'(t)| = \sqrt{\langle \alpha'(t), \alpha'(t) \rangle}$ é a norma do verto tangente $\alpha'(t) \in T_{\alpha(t)}M$, segundo o produto interno definido pela métrica de M.

Um caminho $\alpha : [0,1] \to M$ diz-se seccionalmente de classe C^1 se α é contínua e existe uma partição $0 = t_0 < t_1 < \cdots < t_m = 1$ tal que $\alpha_i = \alpha|_{[t_i, t_{i-1}]}$ é de classe C^1 para todo $i = 1, ..., m$. Assim, definimos o comprimento de α por

$$l(\alpha) = l(\alpha_1) + \cdots + l(\alpha_m)$$

A aditividade da integral mostra que $l(\alpha)$ não depende da escolha da partição.

Seja M uma variedade conexa, de classe C^k. Dados $p, q \in M$, existe um caminho $\alpha : [0,1] \to M$ seccionalmente de classe C^k, tal que $\alpha(0) = p$ e $\alpha(1) = q$, ver (3).

Podemos então definir a distância intrínseca $d(p,q)$ entre dois pontos p, q de uma variedade riemanniana conexa por

$$d(p,q) = \inf\{l(\alpha) : \alpha \text{ é uma curva } C^1 \text{ por partes ligando } p \text{ a } q\}.$$

Seja M uma variedade diferenciável, com uma métrica riemanniana contínua. A distância intrínseca acima definida satisfaz os axiomas que definem um espaço métrico. Além disso, a topologia dessa métrica coincide com a inicial, ver (3).

A seguir daremos algumas definições e enunciaremos uma proposição que será útil mais adiante, mas não daremos à prova desse fato que pode ser encontrada em (3).

Uma curva γ tal que $d(\gamma(0), \gamma(1)) = l(\gamma)$ é chamada de geodésica minimizante. Diremos também que um conjunto $S \subset M$ é fortemente convexo se para quaisquer dois pontos q_1, q_2 do fecho $\overline{S}$ de S existe uma única geodésica minimizante γ ligando q_1 a q_2 cujo interior está contido em S.

Proposição 4.3.1 (vizinhanças convexa)**.** *Para cada $p \in M$ existe um número $\delta > 0$ tal que a bola geodésica $B_{\delta(p)}$ é fortemente convexa.*

Observação 4.3.2. Como M é compacta podemos na Proposição 4.3.1 tomar $\delta > 0$ uniforme.

4.4 Isomorfismo e conjugação de transformações

Definição 4.4.1. Suponha $(X_1, \mathcal{B}_1, \nu_1, T_1)$ e $(X_2, \mathcal{B}_2, \nu_2, T_2)$ são espaços de probabilidades junto com transformações que preservem as medidas. Dizemos que T_1 é isomorfo à T_2 se existe $M_1 \in \mathcal{B}_1$ e $M_2 \in \mathcal{B}_2$ com $\nu_1(M_1) = 1$ e $\nu_2(M_2) = 1$ tal que

(i) $T_1(M_1) \subset M_1, T_2(M_2) \subset M_2$, e

(ii) Existe uma transformação invertível $\phi : M_1 \to M_2$ tal que

$$\phi T_1(x) = T_2\phi(x), \forall x \in M_1 \text{ e } \nu_1(\cdot) = \nu_2(\phi^{-1}(\cdot)).$$

Denotaremos por $T_1 \simeq T_2$.(em (ii) os conjuntos $M_i\,(i = 1,2)$ estão equipados com a σ-álgebra induzida por $\mathcal{B}_i$ e ν_i restrita as σ-álgebras induzidas, respectivamente).

Note que $\simeq$ é uma relação de equivalência.

Daremos agora outra noção de equivalência, só que de medidas. Antes, considere a seguinte relação de equivalência:

$$\forall A, B \in \mathcal{B}_i,\, A \sim B \Longleftrightarrow \nu_i(A\Delta B) = 0.$$

Denote por $\tilde{\mathcal{B}_i}$ a coleção das classe de equivalência e ν_i induz uma medida $\tilde{\nu_i}$ em $\tilde{\mathcal{B}_i}$ por $\tilde{\nu_i}(\tilde{B}) = \nu_i(B)$. O par $(\tilde{\mathcal{B}_i}, \tilde{\nu_i})$ é chamado álgebra de medida.

Seja $(X_1, \mathcal{B}_1, \nu_1)$, $(X_2, \mathcal{B}_2, \nu_2)$ espaços de probabilidades com suas álgebras de medidas $(\tilde{\mathcal{B}_1}, \tilde{\nu_1})$ e $(\tilde{\mathcal{B}_2}, \tilde{\nu_2})$, respectivamente. Se $\phi : X_1 \to X_2$ preserva à medida, i. e, $\nu_2(A) = \nu_1(\phi^{-1}(A)), \forall A \in \mathcal{B}_2$. Então temos uma aplicação $\tilde{\phi}^{-1} : (\mathcal{B}_2, \tilde{\nu_2}) \to (\tilde{\mathcal{B}_1}, \tilde{\nu_1})$ definida por $\tilde{\phi}^{-1}(\tilde{B_2}) = \tilde{B_1}$, onde $B_1 = \phi^{-1}(B_2)$. Essa aplicação está bem definida desde de que ϕ preserva à medida. A aplicação $\tilde{\phi}^{-1}$ preserva complementares e uniões enumeráveis (e disso, interseções enumeráveis). Também $\nu_1(\tilde{\phi}^{-1}(\tilde{B}) = \nu_2(\tilde{B}), \forall \tilde{B} \in \tilde{\mathcal{B}_2}$. Portanto $\tilde{\phi}^{-1}$ pode ser considera um homomorfismo de álgebras de medidas. Observe que $\tilde{\phi}^{-1}$ é injetiva.

Um isomorfismo entre álgebras de medidas é uma aplicação $\Phi : (\tilde{\mathcal{B}_2}, \tilde{\nu_2}) \to (\tilde{\mathcal{B}_1}, \tilde{\nu_1})$ bijetiva tal que preserva complementares, uniões enumeráveis e satisfaz $\nu_1(\Phi(\tilde{B}) = \nu_2(\tilde{B}), \forall \tilde{B} \in \tilde{\mathcal{B}_2}$.

Definição 4.4.2. Seja T_i uma transformação que preserva à medida do espaço de probabilidade $(X_i, \mathcal{B}_i, \nu_i), i = 1,2$. Dizemos que T_1 é conjugado a T_2 se existe um isomorfismo Φ entre as álgebras de medidas $(\tilde{\mathcal{B}_1}, \tilde{\nu_1})$ e $(\tilde{\mathcal{B}_2}, \tilde{\nu_2})$ tal que $\Phi\tilde{T}_2^{-1} = \tilde{T}_1^{-1}\Phi$.

Note que se $T_1 \simeq T_2$ então T_1 é conjugada a T_2. Pois, $\phi : M_1 \to M_2$ determina um isomorfismo nas álgebras de medidas. O próximo resultado dá uma condição suficiente para a reciproca.

Teorema 4.4.3. *Sejam $(X_1,\mathcal{B}_1,\nu_1),(X_2,\mathcal{B}_2,\nu_2)$ espaços de probabilidades que são espaços de Lebesgue ou X_i é um espaço métrico separável e $\mathcal{B}_i$ é sua σ-álgebra de conjuntos de Borel. Se T_1 é conjugada a T_2 então T_1 é isomorfa a T_2.*

Daremos noção mais fracado que a definição de isomorfismo e conjugação.

Definição 4.4.4. Suponha $(X_i,\mathcal{B}_i,\nu_i)$ é uma espaço de probabilidade e $T_i : X_i \to X_i$ é uma transformação que preserva a medida $\nu_i, i = 1,2$. Dizemos que T_2 é um fator de T_1 se existe $M_i \in \mathcal{B}_i$ com $\nu_i(M_i) = 1$ e $T_i(M_i) \subset M_i\,(i = 1,2)$ e existe uma transformação que preserva à medida $\phi : M_1 \to M_2$ com $\phi T_1(x) = T_2\phi(x), \forall x \in M_1$.

A diferença entre isso e isomorfismo é que a transformação ϕ pode não ser invertível.

A seguinte definição é uma noção mais fraca de conjugação.

Definição 4.4.5. Seja T_i uma transformação que preserva à medida de um espaço de probabilidade $X_i,\mathcal{B}_i,\nu_i), i = 1,2$. Dizemos que T_2 é semi-conjugada de T_1 se existe um homomorfismo das álgebra de medida $\Phi : (\tilde{\mathcal{B}_2},\tilde{\nu_2}) \to (\tilde{\mathcal{B}_1},\tilde{\nu_1})$ tal que $\Phi\tilde{T}_2^{-1} = \tilde{T}_1^{-1}\Phi$.

A aplicação $\Phi : (\tilde{\mathcal{B}_2},\tilde{\nu_2}) \to (\tilde{\mathcal{B}_1},\tilde{\nu_1})$ é um homomorfismo das álgebra de medida se $\Phi(\tilde{Y}\backslash\tilde{B}) = \tilde{X}\backslash\Phi(\tilde{B}), \forall B \in \mathcal{B}_2, \Phi(\cup_{n=1}^{\infty}\tilde{B}_n)$, sempre que $\tilde{B}_n \in \mathcal{B}_2$, e $\tilde{\nu_1}(\Phi(\tilde{B})) = \tilde{\nu_2}(\tilde{B}), \forall\tilde{B} \in \mathcal{B}_2$.

Daremos agora um teorema que será útil no que segue.

Teorema 4.4.6. *Assuma que $T : X \to X'$ é uma aplicação mensurável injetiva de um espaço de Lebesgue $(X,\mathcal{F},\nu)$ em um espaço separável $(X',\mathcal{F}',\nu')$ e pré-imagens dos conjuntos de medida 0 (ou positivo) são de medida 0 (ou positiva). Então o espaço $(X',\mathcal{F}',\nu')$ é Lebesgue e T^{-1} é uma aplicação mensurável.*

Sejam $(X,\mathcal{B},\nu)$ um espaço de Lebesgue e seja ζ uma partição arbitrária de X, ζ não é necessariamente enumerável e consistindo de conjuntos mensuráveis. Por $\hat{\zeta}$ denotaremos a sub-σ-álgebra de $\mathcal{B}$ consistindo dos conjuntos $\mathcal{B}$-mensuráveis que são união de elementos inteiros (fibra) de ζ. Obviamente $\hat{\zeta} \subseteq \{\emptyset, X\}$. Chamamos uma partição de ζ-mensurável se ela possui a seguinte propriedade de separação.

Existe $\{B_n\} \subset \hat{\zeta}$ tal que para quaisquer C_1,C_2 em ζ com $C_1 \neq C_2$, existe um inteiro n tal que ou

$$C_1 \subset B_n \;\text{ e }\; C_2 \subset X \setminus B_n$$

ou

$$C_2 \subset B_n \;\text{ e }\; C_1 \subset X \setminus B_n.$$

Sejam $(X,\mathcal{B},\nu)$ um espaço de Lebesgue. Podemos a partir de uma partição mensurável ζ formar um novo espaço de Lebesgue, para isso, tomemos o conjunto X_ζ cujos os pontos são elementos de ζ. Então existe uma projeção natural $\pi : X \to X_\zeta$ que leva um ponto no elemento da partição que o contém, e podemos considerar $\mathcal{F}_\zeta = \pi(\hat{\zeta})$ como uma σ-álgebra de subconjuntos de X_ζ. Também definimos $\nu_\zeta(\cdot) = \nu(\pi^{-1}(\cdot))$ que é uma medida de probabilidade em $(X_\zeta,\mathcal{F}_\zeta)$, o espaço $(X_\zeta,\mathcal{F}_\zeta,\nu_\zeta)$ é chamado de espaço fator de $(X,\mathcal{B},\nu)$ com respeito à ζ. Seja $T : X \to X$ uma transformação que preserva à medida ν. Considere a partição mensurável $\zeta = T^{-1}(\varepsilon)$, onde ε uma partição por pontos, então definimos $T_\zeta : X_\zeta \to X$ por $T_\zeta(\pi(x)) = T(x)$ é mensurável e injetiva. Portanto, pelo Teorema 4.4.6, segue que T_ζ é um isomorfismo.

Agora suponha novamente que ζ é uma partição mensurável de X. Usando a medida ν_ζ, é possível falar de quase toda fibra de ζ. Agora, pode ser demonstrado que para quase todo $C \in \zeta$, existe uma σ-álgebra $\mathcal{B}_C$ e uma medida ν_C definida em C tal que $(C,\mathcal{B}_C,\nu_C)$ é um espaço de Lebesgue e além disso,

(i) Se $B \in \mathcal{B}$, então $B \cap C \in \mathcal{B}_C$ para quase todo C;

(ii) A função $\nu_C(B \cap C)$ é $\mathcal{B}_\zeta$-mensurável se $B \in \mathcal{B}$;

(iii) Para $B \in \mathcal{B}$, $\nu(B) = \int \nu_C(B \cap C) d\nu_\zeta$.

Além disso, pode ser mostrado que o espaço $(C,\mathcal{B}_C,\nu_C)$ são unicamente definido mod 0 para ν_ζ. O sistema de medidas, assim definido, é chamado um sistema canônico com respeito à ζ.

Para mais informações ler (1) ou (4).

Teorema 4.4.7 (Teorema de aproximação)**.** *Seja $(M,\mathcal{B},\nu)$ uma espaço de probabilidade e seja $\mathcal{A}$ uma álgebra que gera a σ-álgebra $\mathcal{B}$. Então para todo $\varepsilon > 0$ e todo $B \in \mathcal{B}$ existe $A \in \mathcal{A}$ tal que $\nu(A \Delta B) < \varepsilon$.*

Demonstração. Pode encontrada em (1) ou (10). □

Teorema 4.4.8 (Extensão)**.** *Seja $\mathcal{A}$ uma álgebra de subconjuntos de M e seja $\nu_0 : \mathcal{A} \to \mathbb{R}$ uma função σ-aditiva com $\nu_0(M) < \infty$. Então existe uma única medida ν definida na σ-álgebra gerada por $\mathcal{A}$ que é uma extensão de ν_0, ou seja, tal que $\nu(A) = \nu_0(A)$ para todo $A \in \mathcal{A}$.*

Demonstração. Pode encontrada em (10). □

Teorema 4.4.9 (Recorrência de Poincaré)**.** *Seja $f : M \to M$ uma transformação mensurável e seja ν uma medida finita invariante por f. Seja $E \subset M$ qualquer conjunto mensurável com $\nu(E) > 0$. Então, para ν-quase todo ponto $x \in E$ temos que $f^k(x) \in E$ com frequência infinita, i. e, existem infinitos valores de k tal que $f^k(x)$ também está em E.*

Demonstração. Mais detalhes pode ser encontrada em (10). □

Teorema 4.4.10. *Seja $(M,\mathcal{B},\nu)$ um espaço de Lebesgue e M um espaço métrico completo, separável. Sejam $B \subset M$ um mensurável, e $f : B \to \mathbb{R}$ uma função integrável. Então a seguinte fórmula vale para ν-quase todo $x \in B$:*

$$\lim_{\varepsilon \to 0} \frac{1}{\nu(B(x,\varepsilon))} \int_{B(x,\varepsilon)} f d\nu = f(x),$$

onde $B(x,\varepsilon) = \{y \in M : d(x,y) \leq \varepsilon\}$

Demonstração. Pode encontrada em (11). □

Lema 4.4.11. *Seja M um espaço métrico compacto. Suponha que $f : M \to M$ admita um jacobiano $J_\nu f$ com relação a uma de probabilidade invariante ν. Então*

$$\int \psi d\nu = \int \sum_{x \in f^{-1}(y)} \frac{\psi(x)}{J_\nu f(x)} d\nu(x),$$

para toda função mensurável limitada $\psi : M \to \mathbb{R}$.

Para mais informações consulte (10)

4.5 Decomposição ergódica

Antes de enunciarmos o teorema de decomposição ergódica vamos analisar um exemplo que ajudará a motivar e delimitar o seu enunciado.

Exemplo 4.5.1. *Considere $f : [0,1] \to [0,1]$ dada por $f(x) = x^2$. As medidas de δ_0 e δ_1 são invariantes e ergódicas para f. Também é claro que $x = 0$ e $x = 1$ são os únicos pontos recorrentes por f e portanto toda probabilidade invariante ν satisfaz $\nu(\{0,1\}) = 1$. Em particular, $\nu = \nu(\{0\})\delta_0 + \nu(\{1\})\delta_1$ é uma combinação convexa (finita) de medidas ergódicas.*

Seja $(M,\mathcal{B},\nu)$ um espaço de probabilidade e $\mathcal{P}$ uma partição de M em conjuntos mensuráveis. Considere $\pi : M \to \mathcal{P}$ a projeção natural que associa a cada $x \in M$ o elemento

$\mathcal{P}(x)$ da partição que o contém. Essa projeção permite munir $\mathcal{P}$ de uma estrutura de espaço de probabilidade, da seguinte forma. Primeiramente, dizemos que um subconjunto $\mathcal{Q}$ de $\mathcal{P}$ é mensurável se, e somente se, a pré-imagem

$$\pi^{-1}(\mathcal{Q}) = \text{união dos elementos } \ P \in \mathcal{P} \ \text{ que pertencem a } \ \mathcal{Q}$$

é um subconjunto mensurável de M. É fácil ver que essa definição está correta: a família $\hat{\mathcal{B}}$ dos subconjunto mensuráveis é uma σ-álgebra em $\mathcal{P}$. Em seguida, definimos a *medida quociente* $\hat{\nu}$ por

$$\hat{\nu}(\mathcal{Q}) = \nu(\pi^{-1}(\mathcal{Q})) \ \text{ para todo } \ \mathcal{Q} \in \hat{\mathcal{B}}.$$

Teorema 4.5.1 (Decomposição ergódica)**.** *Seja M um espaço métrico completo separável, $f: M \to M$ uma transformação mensurável e ν uma probabilidade invariante. Então existe um conjunto mensurável $M_0 \subset M$ com $\nu(M_0) = 1$, uma partição $\mathcal{P}$ de M_0 em subconjuntos mensuráveis e uma família de probabilidades $\{\nu_P : P \in \mathcal{P}\}$ em M, satisfazendo*

(a) $\nu_P(P) = 1$ *para $\hat{\nu}$-quase todo $P \in \mathcal{P}$;*

(b) $P \mapsto \nu_P(E)$ *é mensurável, para todo conjunto mensuráveis $E \subset M$;*

(c) ν_P *é invariante e ergódica para $\hat{\nu}$-quase todo $P \in \mathcal{P}$;*

(d) $\nu(E) = \int \nu_P(E) d\hat{\nu}(P)$, *para todo conjunto mensurável $E \subset M$;*

(e) *Para toda função integrável $\psi : M \to \mathbb{R}$ vale*

$$\int \psi d\nu = \int \Big(\int \psi d\nu_P \Big) d\hat{\nu}.$$

Teorema 4.5.2 (Jacobs)**.** *Suponha que M é um espaço métrico separável. Dada qualquer probabilidade invariante ν, seja $\{\nu_P : P \in \mathcal{P}\}$ a sua decomposição ergódica. Então $h_\nu(f) = \int h_{\nu_P}(f) d\hat{\nu}(P)$(quando um dos lados da igualdade é infinito o outro também é).*

Com isso, temos que se ν é uma medida maximizante, segue que suas componentes ergódicas também são maximizantes.

De fato, seja ν uma medida de probabilidade maximizante. Então, pelos Teorema de Jacobs 4.5.2 e o princípio variacional 1.5.1, temos que $h_{top}(f) = h_\nu(f) = \int h_{\nu_P}(f) d\hat{\nu}(P)$. Considere o conjunto $A = \{\nu_P : \nu_P \ \text{ é maximizante}\}$, logo

$$\begin{aligned} h_{top}(f) = h_\nu(f) &= \int_A h_{\nu_P}(f) d\hat{\nu}(P) + \int_{A^c} h_{\nu_P}(f) d\hat{\nu}(P) \\ &= h_{top}(f)\nu(A) + \int_{A^c} h_{\nu_P}(f) d\hat{\nu}(P) \\ &< h_{top}(f)\hat{\nu}(A) + h_{top}(f)\hat{\nu}(A^c) = h_{top}(f). \end{aligned}$$

Portanto, absurdo. Isso prova que as ν_P são medidas maximizantes.

Para mais informações consulte (10).

4.6 Esperanças condicionais

Fixe uma sequência crescente qualquer $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \cdots \prec \mathcal{P}_n \prec \cdots$ de partições enumeráveis tal que $\mathcal{P} = \bigvee_{n=1}^{\infty} \mathcal{P}_n$ restrito a algum conjunto $M_0 \subset M$ com medida total. Usaremos $\mathcal{P}_n(x)$ para denotar o elemento de $\mathcal{P}_n$ que contém um dado ponto $x \in M_0$.

Seja $\psi : M \to \mathbb{R}$ uma função mensurável limitada qualquer. Para cada $n \geq 1$, defina $e_n(\psi) : M \to \mathbb{R}$ da seguinte forma:

$$e_n(\psi, x) = \begin{cases} \frac{1}{\nu(\mathcal{P}_n(x))} \int_{\mathcal{P}_n(x)} \psi d\nu & \text{se } \nu(\mathcal{P}_n(x)) > 0 \\ 0 & \text{caso contrário.} \end{cases} \tag{4.4}$$

Como as partições $\mathcal{P}_n$ são enumeráveis, o segundo caso da definição se aplica somente num conjunto de pontos com medida ν igual zero. Observe também que $e_n(\psi)$ é contante em cada $P_n \in \mathcal{P}_n$; denotemos por $E_n(\psi, P_n)$ o valor desta constante. Então

$$\int \psi d\nu = \sum_{P_n \in \mathcal{P}_n} \int_{P_n} \psi d\nu = \sum_{P_n \in \mathcal{P}_n} \nu(P_n) E_n(\psi, P_n) = \int e_n(\psi) d\nu \tag{4.5}$$

para todo $n \in \mathbb{N}$(as somas envolvem somente os $P_n \in \mathcal{P}_n$ com medida positiva).

Lema 4.6.1. *Dada qualquer função mensurável limitada $\psi : M \to \mathbb{R}$, existe um subconjunto M_ψ de M com $\nu(M_\psi) = 1$ tal que*

(a) $e(\psi, x) = \lim_n e_n(\psi, x)$ *existe para ν-quase todo $x \in M_\psi$;*

(b) $e(\psi) : M_\psi \to \mathbb{R}$ *é mensurável e é constante em cada $P \in \mathcal{P}$;*

(c) $\int \psi d\nu = \int e(\psi) d\nu$.

O próximo lema será útil para provarmos a Fórmula de Rokhlin.

Lema 4.6.2. *Seja $f : M \to M$ uma função que admite o jacobiano $J_\nu f$ de f com relação a uma probabilidade ν. Então vale a fórmula de mudança de variável:*

$$\int_{f(U_x)} \psi d\nu = \int_{U_x} (\psi \circ f) J_\nu f d\nu$$

para função $\psi : M \to \mathbb{R}$ mensurável e limitada (U_x é qualquer vizinhança de x onde f é invertível).

Mais informações consulte (10).

4.7 Entropia local

Suponhamos que $f: M \to M$ é uma aplicação contínua num espaço métrico compacto. Dado $x \in M, n \geq 1$ e $\varepsilon > 0$, chamamos *bola dinâmica* de comprimento n e raio ε em torno de x ao conjunto:

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

Em outras palavras, $B(x,n,\varepsilon) = \cap_{i=0}^{n-1} f^{-i}(B(f^i(x),\varepsilon))$. Defina:

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

Teorema 4.7.1 (Brin-Katok)**.** *Seja ν uma medida invariante por f. Os limites*

$$\lim_{\varepsilon\to 0} h_\nu^+(f,\varepsilon,x) \quad e \quad \lim_{\varepsilon\to 0} h_\nu^-(f,\varepsilon,x)$$

existem e s ao iguais para ν-quase todo ponto. Denotemos por $h_\nu(f,x)$ o seu valor comum, a função $x \mapsto h_\nu(f,x)$ é mensurável e tem-se

$$h_\nu(f) = \int h_\nu(f,x)d\nu.$$

Para mais informações consulte (11).

4.8 Desigualdade de Ruelle

Esta desigualdade é uma importante estimativa superior para entropia, onde relaciona a entropia de uma medida com os expoente de Lyapunov.

Teorema 4.8.1 (Desigualdade de Ruelle)**.** *Seja $f: M \to M$ uma transformação de classe C^1 numa variedade riemaniana compacta e ν uma probabilidade f-invariante. Então*

$$h_\nu(f) \leq \int \sum_{j:\lambda_x^j>0} m_x^j \lambda_x^j d\nu.$$

Quando ν é ergódica, temos $h_\nu(f) \leq \sum_{j:\lambda_x^j>0} m_x^j\lambda_x^j$.

Demonstração. Pode ser encontrada em (5) □

BIBLIOGRAFIA

1 WALTERS, P. An introduction to ergodic theory. Springer-Verlag. New York. 1981.

2 OLIVEIRA, K.; VIANA, M. Existence and uniquiness of maximizing measures for robust classes of local diffeomorphisms,Journal American Institute of Mathematical Sciences,v.15,n.1,p.225-236, 2006.

3 do CARMO, M. P. Geometria riemanniana. 4ª ed. Rio de Janeiro: IMPA, 2008. (Projeto Euclides).

4 PARRY, W. Entropy and generators in ergodic theory. W.A. Benjamin, INC. New York. 1969.

5 RUELLE, D. An inequality for the entropy of differentiable maps, Journal: Boletim da Sociedade Brasileira de Matemática. v.9, n.1, p. 83-87. 1978.

6 KOMOGOROV, A. N. TIHOMIROV, V. M. ϵ-Entropy and ϵ-capacity in function spaces.Journal:Translations of the American Mathematical Society. n.2, v.17, p. 277-364, 1961.

7 ADLER, R. L. A.G. KONHEIM, A. G. McANDREW, M. H. Topological entropy.Journal: Translation of the American Mathematical Society. v. 144, p. 309-319. 1965.

8 DINABURG, E. I. The relation between topological entropy and metric entropy.Journal: Dokl. Akad. Nauk SSSR. v. 190, p.19-22.1970. Journal: Soviet Math. Dokl. v.11, p.13-16.1969.

9 BOWEN, R. Entropy for group endmorphisms and homogeneous spaces. Journal: Translation of the American Mathematical Society. v. 153, p. 401-414. 1971.

10 OLIVEIRA, K. VIANA, M. Teoria Ergódica Um Curso Introdutório - Impa
URL: www.impa.br/ viana/out/ite.pdf

11 MAÑE, R. Ergodic theory and differential dynamics. Springer-Verlag. New York. 1987.