UNIVERSIDADE FEDERAL DE PERNAMBUCO
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

RICARDO FREIRE DA SILVA

# ESTADOS DE EQUILÍBRIO PARA TRANSFORMAÇÕES EXPANSORAS, DINÂMICA SIMBÓLICA, DIFEOMORFISMOS AXIOMA A E HOMEOMORFISMOS COM ESPECIFICAÇÃO

Recife
2020

RICARDO FREIRE DA SILVA

# ESTADOS DE EQUILÍBRIO PARA TRANSFORMAÇÕES EXPANSORAS, DINÂMICA SIMBÓLICA, DIFEOMORFISMOS AXIOMA A E HOMEOMORFISMOS COM ESPECIFICAÇÃO

Dissertação apresentada ao programa de Pós-graduação do Departamento de Matemática da Universidade Federal do Pernambuco, como requisito parcial para a obtenção do título de Mestrado em Matemática.

Área de concentração: Análise

Orientador (a): Ricardo Turolla Bortolotti

Recife
2020

Catalogação na fonte
Bibliotecária Monick Raquel Silvestre da S. Portes, CRB4-1217

S586e Silva, Ricardo Freire da
Estados de equilíbrio para tansformações expansoras, dinâmica simbólica, defeomorfismos axioma A e homeomorfismo com especificação / Ricardo Freire da Silva. – 2020.
212 f.: il., fig.

Orientador: Ricardo Turolla Bortolotti.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CCEN, Matemática, Recife, 2020.
Inclui referências.

1. Análise matemática. 2. Estados de equilíbrio. I. Bortolotti, Ricardo Turolla (orientador). II. Título.

515 CDD (23. ed.) UFPE- CCEN 2020 - 48

RICARDO FREIRE DA SILVA

**ESTADOS DE EQUILÍBRIO PARA TRANSFORMAÇÕES EXPANSORAS, DINÂMICA SIMBÓLICA, DIFEOMORFISMOS AXIOMA A E HOMEOMORFISMOS COM ESPECIFICAÇÃO**

Dissertação apresentada ao programa de Pós-graduação do Departamento de Matemática da Universidade Federal do Pernambuco, como requisito parcial para a obtenção do título de Mestrado em Matemática

Aprovado em: 18/02/2020

**BANCA EXAMINADORA**

Prof. Dr. Ricardo Turolla Bortolotti (Orientador)
Universidade Federal do Pernambuco

Prof. Dr. Augusto de Castro Júnior (Examinador Externo)
Universidade Federal da Bahia

Prof. Dr. Yuri Gomes de Lima (Examinador Externo)
Universidade Federal do Ceará

## AGRADECIMENTOS

Agradeço a Deus pela vida que Ele me concedeu.

Agradeço aos meus pais Rosa e Rivaldo por todo o esforço investido na minha educação e apesar de não estarem mais presentes fisicamente, acredito que eles estejam orgulhosos por esta conquista. Agradeço a minha família, em especial ao meu avô Manoel e minha irmã Ruth.

Agradeço à minha namorada Úrsula que sempre esteve ao meu lado durante o meu percurso acadêmico.

Sou grato pela confiança depositada na minha proposta de projeto pelo meu orientador, o professor Ricardo. Obrigado por me manter motivado durante todo o processo!

Também quero agradecer aos meus amigos, bons amigos, que fiz em Recife, pelos momentos divertidos que vivemos "isso ajudou a manter a sanidade em dias", Ada Azevedo, Ellen, Geovani, Igor, Jackellyny, Junior, Júlio, Micael, Micaela, Mirelle, etc e também ao meu bom amigo Matheus de Juazeiro do Norte-CE, que não ajudou em nada, porém é um bom amigo, e não posso deixar de citar os professores da Urca: Braga, Valéria, Paulo César, Bárbara, Luciana, etc todos contribuíram imensamente para minha formação inicial e para que eu pudesse vir cursar o mestrado. Por fim agradeço ao CNPQ pelo apoio financeiro.

## RESUMO

Neste trabalho estudamos existência e unicidade de medidas invariantes denominadas estados de equilíbrio. Primeiro estudamos a existência e unicidade para transformações expansoras por meio de propriedades espectrais do operador de transferência de Ruelle-Perron-Frobenius, em seguida para a dinâmica simbólica e para difeomorfismos que satisfazem o Axioma A, usamos neste caso a noção de partições de Markov e da semiconjugação entre a dinâmica simbólica e o difeomorfismo restrito aos conjuntos chamados de básicos. Por fim, para homeomorfismos expansivos com a propriedade de especificação com uma abordagem do ponto de vista topológico. Para fazer isso, estudamos conceitos fundamentais de Teoria Ergódica, como entropia métrica e topológica, transformações expansoras e expansivas, pressão topológica, o princípio variacional, o operador de transferência de Ruelle-Perron-Frobenius e de Dinâmica Hiperbólica como conjuntos hiperbólicos, em particular construiremos uma ferradura de Smale como exemplo de conjunto hiperbólico, Teorema da Variedade Estável, Teorema da Decomposição Espectral, Lema de Sombreamento, partições de Markov e Teorema da Especificação. No final estudamos um exemplo onde não há unicidade do estado de equilíbrio devido a não regularidade Hölder contínua do potencial.



**Palavras-chave:** Estados de equilíbrio. Transformações expansoras. Difeomorfismos Axioma A. Homeomorfismos com a propriedade de especificação. Teoria Ergódica.

## ABSTRACT

In this work, we study the existence and uniqueness of invariant measures called equilibrium states. First we study the existence and uniqueness for expanding maps by means of spectral properties of the Ruelle-Perron-Frobenius transfer operator, then for symbolic dynamics and for diffeomorphisms that satisfy Axiom A, in this case we use the notion of Markov and the semiconjugation between symbolic dynamics and diffeomorphism restricted to the so-called basic sets. Finally, for expansive homeomorphisms with the property of specification with an approach from the topological point of view. To do it, we study fundamental concepts of Ergodic Theory, such as metric and topological entropy, topological pressure, the variational principle and the Ruelle-Perron-Frobenius transfer operator and Hyperbolic Dynamics as hyperbolic sets, in particular we will build a Smale horseshoe as an example of a hyperbolic set, The Stable Manifold Theorem, Spectral Decomposition Theorem, Shadowing Lemma, Markov partitions and Specification Theorem.. In the end, we study one example, where there is no uniqueness of the equilibrium state due to the Hölder continuous non-regularity of the potential.



**Keywords:** Equilibrium states. Expanding maps. Axiom A diffeomorphisms. Homeomorphism with specification property. Ergodic Theory.

## LISTA DE FIGURAS

## SUMÁRIO

# 1 INTRODUÇÃO

Teoria Ergódica é a disciplina matemática que estuda sistemas dinâmicos munidos de medidas invariantes. Neste trabalho, consideraremos um sistema dinâmico como sendo uma transformação $f : M \to M$, definida num espaço $M$ que na maioria das vezes será um espaço métrico compacto e que associa a cada estado $x \in M$ do sistema o estado $f(x) \in M$, no qual o sistema se encontrará uma unidade de tempo depois, que por conseguinte associa o estado $f^2(x) \in M$ em que o sistema se encontrará duas unidades de tempo depois, e assim por diante. Trata-se portanto de um modelo de dinâmica com tempo discreto. Também na maioria dos casos suporemos que o sistema dinâmico é mensurável, ou seja, que o espaço $M$ está munido de uma $\sigma$-álgebra de subconjuntos mensuráveis (geralmente a $\sigma$-álgebra de Borel).

O principal intuito deste trabalho é fazer um estudo amplo sobre objetos matemáticos chamados estados de equilíbrio, que são medidas de probabilidade invariantes $\mu$ tais que

$$h_\mu(f) + \int \phi d\mu = \sup \left\{ h_\nu(f) + \int \phi d\nu \right\},$$

onde $h_\mu(f)$ é a entropia da transformação $f$ com respeito à medida $\mu$, $\phi$ é uma função contínua usualmente chamada de potencial e o supremo é tomado sobre todas as medidas de probabilidade invariantes por $f$. O supremo do lado direito é chamado de pressão topológica do potencial $\phi$. Os aspectos que estamos interessados em estudar dos estados de equilíbrio são existência e unicidade. Vamos mostrar ao longo do texto que para algumas classes de sistemas dinâmicos a existência e unicidade do estado de equilíbrio para potenciais com certas propriedades (Hölder contínuo, diferenciável, com variação limitada, etc) é válida.

O capítulo 1 deste trabalho traz os principais resultados de Teoria Ergódica que serão usados ao longo do texto. As demonstrações de tais resultados são omitidas pois existem excelentes bibliografias com as demonstrações de tais resultados. No capítulo 2 é que de fato iniciamos o nosso estudo. Este capítulo traz as definições de entropia métrica, entropia topológica, Jacobianos, pressão topológica e finalmente de estados de equilíbrio. O capítulo também traz exemplos concretos do cálculo da entropia métrica, da entropia topológica e exemplos de estados de equilíbrio.

O capítulo 3 aborda o primeiro caso em que há existência e unicidade do estado de equilíbrio para a classe das transformações chamadas expansoras. Isto será mostrado no Teorema de Ruelle (Teorema 4.17). Além disso, outras propriedades importantes serão obtidas, como por exemplo, que esta medida é um estado de Gibbs, com suporte igual a todo espaço.

A ideia agora é ampliar a classe das transformações para as quais podemos garantir que existe um único estado de equilíbrio. Isto será feito no capítulo 4, no qual mostraremos que existe um único estado de equilíbrio para dinâmicas simbólicas (Teorema 5.4), com o auxílio de um outro importante resultado, que é o Teorema de Ruelle-Perron-Frobenius (Teorema 5.5). No mesmo capítulo, mostramos que existe único estado de equilíbrio para um difeomorfismo Axioma A quando restrito a um de seus conjuntos básicos (Teorema 5.55). Para fazer isso será necessário conhecer várias definições e resultados de Dinâmica Hiperbólica, a definição de conjunto hiperbólico, o Teorema da Variedade Estável, o Lema do Sombreamento, o Teorema da decomposição espectral, partições de Markov entre outros. A existência de partições de Markov terá papel fundamental, pois permitirá associarmos uma dinâmica simbólica à dinâmica do conjunto hiperbólico.

O capítulo 5 traz um outro caso para o qual há existência e unicidade de estados de equilíbrio. É o caso de homeomorfismos expansivos que possuem a propriedade de especificação, como mostramos no Teorema 6.6. Ainda neste capítulo demonstramos um resultado importante chamado Teorema da especificação, para provar que um difeomorfismo Axioma A possui um iterado que, quando restrito à decomposição dos conjuntos básicos, satisfaz a propriedade de especificação. Finalmente, no capítulo 6 damos um exemplo de dinâmica simbólica que possui dois estados de equilíbrio: o ponto chave deste caso é a construção de um potencial que não é Hölder contínuo.

# 2 RESULTADOS PRELIMINARES

Neste capítulo, estabeleceremos algumas definições e resultados básicos de Teoria Ergódica essenciais para o entendimento da teoria dos capítulos seguintes. As demonstrações de tais resultados foram omitidas, porém todas podem ser encontradas na referência [7].

**Definição 2.1.** Seja $(M, \mathcal{B}, \mu)$ um espaço de medida e seja $f : M \to M$ uma transformação mensurável. Dizemos que a medida $\mu$ é **invariante** por $f$ se

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

Nesse caso também dizemos que $f$ preserva $\mu$. Uma definição equivalente a esta é que $f$ preserva $\mu$ se, e somente se,

$$\int \phi d\mu = \int (\phi \circ f) d\mu$$

para toda função $\mu$-integrável $\phi : M \to \mathbb{R}$.

**Teorema 2.2** (**Desigualdade de Jensen**)**.** *Seja $f : M \to I$ uma função em $L^1(\mu)$ e $\phi : I \to \mathbb{R}$ uma função convexa. Se $\mu$ é uma probabilidade em $M$ e $f$ é tal que $\int f d\mu \in I$, então:*

$$\phi\left(\int f d\mu\right) \leq \int (\phi \circ f) d\mu \tag{2.1}$$

**Corolário 2.3.** *Seja $\phi : \mathbb{R} \to \mathbb{R}$ uma função convexa, seja $(\lambda_i)_i$ uma sequência de números reais não negativos satisfazendo $\sum_{i=1}^{\infty} \lambda_i \leq 1$ e seja $(a_i)_i$ uma sequência limitada de números reais. Então,*

$$\phi\left(\sum_{i=1}^{\infty} \lambda_i a_i\right) \leq \sum_{i=1}^{\infty} \lambda_i \phi(a_i) \tag{2.2}$$

.

A $\sigma$-álgebra gerada por uma família $\xi$ de subconjuntos de $X$ é a menor $\sigma$-álgebra $\sigma(\xi)$ que contém a família $\xi$, ou seja, é a interseção de todas as $\sigma$-álgebras que contêm $\xi$.

Um resultado importante sobre $\sigma$-álgebras, que nos será útil mais tarde, afirma que todo elemento $B$ da $\sigma$-álgebra gerada por uma álgebra é aproximado por algum elemento $A$ da álgebra, no sentido de que a medida da diferença simétrica pode ser tão pequena quanto se queira.

**Teorema 2.4 (Aproximação).** *Seja $(M, \mathcal{B}, \mu)$ um espaço de probabilidade e seja $\mathcal{A}$ uma álgebra que gera a $\sigma$-álgebra $\mathcal{B}$. Então para todo $\varepsilon > 0$ e todo $B \in \mathcal{B}$ existe $A \in \mathcal{A}$ tal que $\mu(A \Delta B) < \varepsilon$.*

**Definição 2.5.** Suponhamos que $\mu$ é uma medida de probabilidade invariante por uma transformação mensurável $f : M \to M$. Diremos que o sistema $(f, \mu)$ é **ergódico** se, para todo conjunto mensurável $E$ invariante (isto é $\mu(E \Delta f^{-1}(E)) = 0$), tem-se $\mu(E) = 0$ ou $\mu(E) = 1$. Uma função mensurável $\phi : M \to \mathbb{R}$ é dita **invariante** para $\mu$ se $\phi = \phi \circ f$ em $\mu$-quase todo ponto. Uma condição equivalente à definição de ergodicidade é a seguinte: toda função integrável invariante $\phi : M \to \mathbb{R}$ é constante em $\mu$-quase todo ponto. Dizemos que o sistema $(f, \mu)$ é **misturador** se

$$\lim_{n \to \infty} \mu(f^{-n}(A) \cap B) = \mu(A)\mu(B)$$

para quaisquer conjuntos mensuráveis $A, B \subset M$. É imediato que um sistema misturador é necessariamente ergódico: de fato, se não é ergódico, existe algum conjunto invariante $A \subset M$ com $0 < \mu(A) < 1$. Tomando $B = M \setminus A$ vem que $f^{-n}(A) \cap B = A \cap (M \setminus A) = \emptyset$ para todo $n$. Logo, $\mu(f^{-n}(A) \cap B) = 0$ para todo $n$, enquanto que $\mu(A)\mu(B) \neq 0$. Logo $\lim_n \mu(f^{-n}(A) \cap B) \neq \mu(A)\mu(B)$ e o sistema não é misturador.

**Observação 2.6.** *Existe definição para medida ergódica não invariante. Também é interessante saber que o Teorema ergódico de Birkhoff pode ser obtido sem que a medida seja invariante, por exemplo, veja o trabalho de Witold Hurewicz: Ergodic Theorem Without Invariant Measure.*

**Lema 2.7.** *Se $\mu$ e $\nu$ são probabilidades invariantes tais que $\mu$ é ergódica e $\nu$ é absolutamente contínua com relação a $\mu$, então $\mu = \nu$.*

**Teorema 2.8** (**Decomposição Ergódica**). *Seja $M$ um espaço completo separável, $f : M \to M$ uma transformação mensurável e $\mu$ uma probabilidade invariante. Então existe um conjunto mensurável $M_0 \subset M$ com $\mu(M_0) = 1$, uma partição $\mathscr{P}$ de $M_0$ em subconjuntos mensuráveis e uma família de probabilidades $\{\mu_P : P \in \mathscr{P}\}$ em $M$, satisfazendo:*

*(a)* $\mu_P(P) = 1$ *para $\hat{\mu}$-quase todo $P \in \mathscr{P}$;*

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

*(c)* $\mu_P$ *é invariante e ergódica para $\hat{\mu}$-quase todo $P \in \mathscr{P}$;*

*(d)* $\mu(E) = \int \mu_P(E)d\hat{\mu}(P)$, *para todo conjunto mensurável $E \subset M$ ($\hat{\mu}$ é a medida quociente, definida por $\hat{\mu}(\mathcal{Q}) = \mu(\pi^{-1}(\mathcal{Q}))$ para cada $Q$ na $\sigma$-álgebra $\hat{\mathcal{B}}$ de $\mathscr{P}$ e $\pi : M \to \mathscr{P}$ é a projeção natural que associa a cada $x \in M$ o elemento $\mathscr{P}(x)$ da partição que o contém).*

**Definição 2.9** (**Esperanças Condicionais**). Fixe uma sequência crescente qualquer $\mathscr{P}_1 \prec \mathscr{P}_2 \prec \cdots \prec \mathscr{P}_n \prec \cdots$ de partições enumeráveis tal que $\mathscr{P} = \bigvee_{n=1}^{\infty} \mathscr{P}_n$ restrito a algum conjunto $M_0 \subset M$ com medida total. Usaremos $\mathscr{P}_n(x)$ para denotar o elemento de $\mathscr{P}_n$ que contém um dado ponto $x \in M$. Seja $\psi : M \to \mathbb{R}$ uma função mensurável limitada qualquer. Para cada $n \geq 1$, defina a esperança condicional $e_n(\psi) : M \to \mathbb{R}$ da seguinte forma:

$$e_n(\psi, x) = \begin{cases} \dfrac{1}{\mu(\mathscr{P}_n(x))} \displaystyle\int_{\mathscr{P}_n(x)} \psi d\mu & \text{se } \mu(\mathscr{P}_n(x)) > 0 \\ 0 & \text{caso contrário.} \end{cases}$$

Se $\psi = \chi_A$ para algum $A \subset M$ mensurável, então a definição significa que

$$e_n(\psi, x) = \frac{\mu(\mathscr{P}_n(x) \cap A)}{\mu(\mathscr{P}_n(x))}. \tag{2.3}$$

**Lema 2.10.** *Dada qualquer função $\psi : M \to \mathbb{R}$ mensurável limitada, existe um subconjunto $M_\psi$ de $M$ com $\mu(M_\psi) = 1$ tal que:*

*(a)* $e(\psi, x) = \lim_n e_n(\psi, x)$ *existe para todo $x \in M_\psi$.*

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

*(c)* $\int \psi d\mu = \int e(\psi) d\mu$.

**Teorema 2.11.** *Seja* $\mathscr{P}_1 \prec \mathscr{P}_2 \prec \cdots \prec \mathscr{P}_n \prec \cdots$ *uma sequência crescente de partições enumeráveis tais que* $\cup_n \mathscr{P}_n$ *gera a* $\sigma$*-álgebra* $\mathcal{B}$ *dos conjuntos mensuráveis a menos de medida nula. Então, a esperança condicional* $e(\psi) = \lim_n e_n(\psi)$ *coincide com* $\psi$ *em quase todo ponto, para toda função mensurável limitada.*

**Definição 2.12** (**Deslocamentos de Bernoulli**)**.** Seja $(X, \mathcal{C}, \nu)$ um espaço de probabilidade qualquer. Consideramos o espaço produto $\Sigma = X^{\mathbb{N}}$, munido da $\sigma$-álgebra produto $\mathcal{B} = \mathcal{C}^{\mathbb{N}}$ e da medida produto $\mu = \nu^{\mathbb{N}}$. Isto quer dizer que $\Sigma$ é o conjunto de todas as sequências $(x_n)_{n\in\mathbb{N}}$ com $x_n \in X$ para todo $n$. Por definição, $\mathcal{B}$ é a $\sigma$-álgebra gerada pelos cilindros

$$[m; A_m, \ldots, A_n] = \{(x_i)_{i\in\mathbb{N}} : x_i \in A_i \quad \text{para} \quad m \leq i \leq n\}$$

onde $m \leq n$ e cada $A_i$ é um elemento de $\mathcal{C}$. Além disso, $\mu$ é caracterizada por

$$\mu([m; A_m, \ldots, A_n]) = \prod_{i=m}^{n} \nu(A_i). \tag{2.4}$$

Podemos pensar nos elementos de $\Sigma$ como representando os resultados de sequências de experimentos regidos por uma mesma distribuição de probabilidade $\nu$: dado qualquer conjunto mensurável $A \subset X$, a probabilidade de obtermos $x_i \in A$ é igual a $\nu(A)$, qualquer que seja $i$. Além disso, os resultados dos sucessivos experimentos são independentes: de fato a relação (2.4) significa que a probabilidade de $x_i \in A_i$ para todo $m \leq i \leq n$ é o produto das probabilidades de cada um dos eventos $x_i \in A_i$ separadamente.

O **deslocamento (ou "shift") de Bernoulli** é a dupla $(\sigma, \mu)$, onde a dinâmica $\sigma : \Sigma \to \Sigma$ preserva a medida $\mu$ e é definida por

$$\sigma((x_n)_n) = (x_{n+1})_n.$$

Convém observar que é possível substituir $\mathbb{N}$ por $\mathbb{Z}$ em toda a construção, ou seja, podemos considerar $\Sigma = X^{\mathbb{Z}}$ como sendo o espaço das sequências bilaterais $(\ldots, x_{-n}, \ldots, x_0, \ldots, x_n, \ldots)$. A dupla $(\sigma, \mu)$ é um **deslocamento de Bernoulli bilateral** e $\sigma$ nesse caso é uma aplicação invertível.

**Definição 2.13** (**Deslocamentos de Markov**). Consideremos $X$ um conjunto finito, digamos $X = \{1, \dots, d\}$ para algum $d \geq 2$ munido da topologia discreta e da respectiva $\sigma$-álgebra de Borel. Considere $\Sigma = X^{\mathbb{N}}$ munido da $\sigma$-álgebra produto e o deslocamento $\sigma : \Sigma \to \Sigma$, $\sigma((x_n)_n) = \sigma((x_{n+1})_n)$. Suponha que é dada uma família $\{P(x, *) : x \in X\}$ de probabilidades em $X$, chamadas **probabilidades de transição**. Como $X$ é finito, essas probabilidades ficam completamente caracterizadas pelos valores

$$P_{i,j} = P(i, \{j\}) \quad \text{para todo } 1 \leq i, j \leq d. \tag{2.5}$$

Uma probabilidade $p$ em $X$ é chamada uma **medida estacionária** relativamente à família de probabilidades de transição se ela satisfaz

$$\int P(x, E) dp(x) = p(E) \quad \text{para todo conjunto mensurável } E \subset X. \tag{2.6}$$

Como $X$ é finito, uma medida estacionária $p$ em $X$ fica completamente caracterizada pelos valores $p_i = p(\{i\}), 1 \leq i \leq d$. Com esta notação, (2.6) traduz-se por

$$\sum_{i=1}^{d} p_i P_{i,j} = p_j \quad \text{para todo } 1 \leq j \leq d. \tag{2.7}$$

Fixada uma medida estacionária $p$ qualquer (elas sempre existem, veja a observação 2.15 abaixo), defina

$$\mu([m; a_m, \dots, a_n]) = p_{a_m} P_{a_m, a_{m+1}} \dots P_{a_{n-1}, a_n} \tag{2.8}$$

para todo cilindro $[m; a_m, \dots, a_n]$ em $\Sigma$. Esta função $\mu$ se estende a uma probabilidade na $\sigma$-álgebra gerada pelos cilindros e é invariante pelo deslocamento $\sigma$. Toda probabilidade $\mu$ obtida desta forma é chamada **medida de Markov**; além disso, o sistema $(\sigma, \mu)$ é chamado **deslocamento de Markov**. Note que, quando $P_{i,j} = p_j$, temos exatamente o deslocamento de Bernoulli. Convém observar que toda a construção é valida se considerarmos $\Sigma = X^{\mathbb{Z}}$. Nesse caso, temos um **deslocamento de Markov bilateral**.

**Definição 2.14.** Uma matriz $P$ é dita **matriz estocástica** se satisfaz as seguintes condições:

1. $P_{i,j} \geq 0$ para todo $1 \leq i, j \leq d$;

2. $\sum_{j=1}^{d} P_{i,j} = 1$ para todo $1 \leq i \leq d$.

**Observação 2.15.** *Se denotarmos $p = (p_1, \ldots, p_d)$, a relação (2.7) corresponde a*

$$P^* p = p \tag{2.9}$$

*onde $P^*$ representa a transposta da matriz $P$. Em outras palavras: medidas estacionárias correspondem precisamente aos autovalores da matriz transposta para o autovalor 1. O famoso Teorema de Perron-Fröbenius permite mostrar que tais autovalores sempre existem.*

**Teorema 2.16** (**Teorema de Perron-Fröbenius**)**.** *Seja $A$ uma matriz $d \times d$ com entradas não negativas. Então existe $\lambda \geq 0$ e existe algum vetor $v \neq 0$ com entradas não negativas tal que $Av = \lambda v$ e $\lambda \geq |\gamma|$ para todo autovalor $\gamma$ de $A$. Se $A$ admite alguma potência cujas entradas são positivas então $\lambda > 0$ e existe algum autovetor $v$ com entradas positivas. Ademais, $\lambda > |\gamma|$ para qualquer outro autovalor $\gamma$ de $A$. Além disso, o autovalor $\lambda$ tem multiplicidade 1 e é o único autovalor de $A$ que admite algum autovetor com entradas não negativas.*

*Aplicando este teorema à matriz $A = P^*$, concluímos que existem $\lambda \geq 0$ e $p \neq 0$ com $p_i \geq 0$ para todo $i$, tais que*

$$\sum_{i=1}^{d} p_i P_{i,j} = \lambda p_j \quad \textit{para todo } 1 \leq j \leq d,$$

*somando sobre todos os $i = 1, \ldots, d$ obtemos que*

$$\sum_{j=1}^{d} \sum_{i=1}^{d} p_i P_{i,j} = \lambda \sum_{j=1}^{d} p_j,$$

*usando a propriedade 2 da matriz estocástica, o lado esquerdo desta desigualdade pode ser reescrito como*

$$\sum_{i=1}^{d} p_i \sum_{j=1}^{d} P_{i,j} = \lambda \sum_{i=1}^{d} p_i.$$

*Lembrando que a soma das entradas de $p$ é um número positivo, concluímos que $\lambda = 1$. Então, sempre existem vetores $p \neq 0$ satisfazendo (2.9).*

**Lema 2.17.** *Seja $P$ uma matriz estocástica e seja $p = (p_1, \ldots, p_d)$ uma solução de $P^* p = p$. Para cada $n \geq 0$, denote por $P^n_{i,j}$, $1 \leq i, j \leq d$ as entradas da matriz $P^n$. Então:*

*(a)* $\sum_{j=1}^{d} P^n_{i,j} = 1$ *para todo* $1 \leq i \leq d$ *e todo* $n \geq 1$*;*

*(b)* $\sum_{i=1}^{d} p_i P^n_{i,j} = p_j$ *para todo* $1 \leq j \leq d$ *e todo* $n \geq 1$*;*

*(c) o hiperplano* $H = \{(h_1, \ldots, h_d) : h_1 + \cdots + h_d = 0\}$ *é invariante por* $P^*$*.*

**Definição 2.18** (**Equivalência Ergódica**)**.** Sejam $\mu$ e $\nu$ probabilidades invariantes por transformações $f : M \to M$ e $g : N \to N$, respectivamente. Dizemos que os sistemas $(f, \mu)$ e $(g, \nu)$ são **ergodicamente equivalentes** se podemos escolher conjuntos mensuráveis $X \subset M$ e $Y \subset N$ com $\mu(M \setminus X) = 0$ e $\nu(N \setminus Y) = 0$, e uma bijeção mensurável $\phi : X \to Y$ com inversa mensurável, de tal forma que

$$\phi_* \mu = \nu \;\; \text{e} \;\; \phi \circ f = g \circ \phi. \tag{2.10}$$

**Observação 2.19.** *Os conjuntos $X$ e $Y$ na definição de equivalência ergódica podem ser escolhidos invariantes por $f$ e $g$, respectivamente. De fato, considere $X_0 = \cap_{n=0}^{\infty} f^{-n}(X)$. É claro da definição que $X_0 \subset X$ e $X_0 \subset f^{-1}(X_0)$. Como $\mu(X) = 1$ e a interseção é enumerável, temos que $\mu(X_0) = 1$. Analogamente, $Y_0 = \cap_{n=0}^{\infty} g^{-n}(Y)$ é subconjunto mensurável de $Y$ tal que $\nu(Y_0) = 1$ e $Y_0 \subset g^{-1}(Y_0)$. Além disso, por construção, $Y_0 = \phi(X_0)$. Portanto, a restrição de $\phi$ a $X_0$ ainda é uma bijeção sobre $Y_0$.*

**Teorema 2.20.** *Suponha que $(f, \mu)$ e $(g, \nu)$ são ergodicamente equivalentes. Então*

*(i) $(f, \mu)$ é ergódico se e somente se $(g, \nu)$ é ergódico.*

*(ii) $(f, \mu)$ é misturador se e somente se $(g, \nu)$ é misturador.*

**Definição 2.21.** As propriedades de mistura e ergodicidade são ditas **invariantes de equivalência ergódica**.

# 3 ENTROPIA E PRESSÃO, PRINCÍPIO VARIACIONAL E ESTADOS DE EQUILÍBRIO

## 3.1 Entropia Métrica

Seja $(M, \mathcal{B}, \mu)$ um espaço de probabilidade.

**Definição 3.1.** Uma **partição** de $M$ é uma família finita ou enumerável $\mathscr{P}$ de subconjuntos mensuráveis de $M$ disjuntos dois a dois e cuja união tem medida total. Denotamos por $\mathscr{P}(x)$ o elemento da partição que contém um ponto $x$. Dada qualquer família enumerável de partições $\mathscr{P}_n$, definimos a **soma** das partições da família $\mathscr{P}_n$ como sendo

$$\bigvee_n \mathscr{P}_n = \left\{ \bigcap_n P_n : P_n \in \mathscr{P}_n \text{ para cada } n \right\}.$$

Associamos a cada partição $\mathscr{P}$ a função mensurável $\mathscr{I}_{\mathscr{P}} : M \to \mathbb{R}$ chamada **função de informação**, definida por

$$\mathscr{I}_{\mathscr{P}}(x) = -\log \mu(\mathscr{P}(x)).$$

Então chamamos **entropia da partição** $\mathscr{P}$ ao número

$$H_\mu(\mathscr{P}) = \int \mathscr{I}_{\mathscr{P}} d\mu = \sum_{P \in \mathscr{P}} -\mu(P) \log \mu(P).$$

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

$$H_\mu(\mathscr{P} \setminus \mathscr{Q}) = \sum_{P \in \mathscr{P}} \sum_{Q \in \mathscr{Q}} -\mu(P \cap Q) \log \frac{\mu(P \cap Q)}{\mu(Q)}. \tag{3.1}$$

Ressaltamos que trabalharemos com partições com entropia finita e que também convencionaremos que $0 \log 0 = 0$.

**Definição 3.2.** Dadas duas partições $\mathscr{P}$ e $\mathscr{Q}$, dizemos que $\mathscr{P}$ é menos fina que $\mathscr{Q}$, e escrevemos $\mathscr{P} \prec \mathscr{Q}$, se todo elemento de $\mathscr{Q}$ está contido em algum elemento de $\mathscr{P}$, a menos de medida nula. Em geral, dadas duas partições $\mathscr{P}$ e $\mathscr{Q}$, a soma $\mathscr{P} \bigvee \mathscr{Q}$ é, precisamente, a menos fina de todas as partições $\mathscr{R}$ tais que $\mathscr{P} \prec \mathscr{R}$ e $\mathscr{Q} \prec \mathscr{R}$.

**Lema 3.3.** *Sejam $\mathscr{P}, \mathscr{Q}$ e $\mathscr{R}$ partições com entropia finita. Então,*

*(i)* $H_\mu((\mathscr{P} \bigvee \mathscr{Q}) \setminus \mathscr{R}) = H_\mu(\mathscr{P} \setminus \mathscr{R}) + H_\mu(\mathscr{Q} \setminus (\mathscr{P} \bigvee \mathscr{R}))$*;*

*(ii) se* $\mathscr{P} \prec \mathscr{Q}$ *então* $H_\mu(\mathscr{P} \setminus \mathscr{R}) \leq H_\mu(\mathscr{Q} \setminus \mathscr{R})$ *e* $H_\mu(\mathscr{R} \setminus \mathscr{P}) \geq H_\mu(\mathscr{R} \setminus \mathscr{Q})$*;*

*(iii)* $\mathscr{P} \prec \mathscr{Q}$ *se, e somente se,* $H_\mu(\mathscr{P} \setminus \mathscr{Q}) = 0$*.*

*Demonstração.* Por definição,

$$H_\mu((\mathscr{P} \bigvee \mathscr{Q}) \setminus \mathscr{R}) = \sum_{X \in \mathscr{P} \bigvee \mathscr{Q}} \sum_{R \in \mathscr{R}} -\mu(X \cap R) \log \frac{\mu(X \cap R)}{\mu(R)}$$

.

Mas, $X \in \mathscr{P} \bigvee \mathscr{Q}$ é, por definição, da forma $P \cap Q$ com $P \in \mathscr{P}$ e $Q \in \mathscr{Q}$. Então

$$\begin{aligned} H_\mu((\mathscr{P} \bigvee \mathscr{Q}) \setminus \mathscr{R}) &= \sum_{P,Q,R} -\mu(P \cap Q \cap R) \log \frac{\mu(P \cap Q \cap R)}{\mu(R)} \\ &= \sum_{P,Q,R} -\mu(P \cap Q \cap R) \log \frac{\mu(P \cap Q \cap R)}{\mu(P \cap R)} \\ &\quad + \sum_{P,Q,R} -\mu(P \cap Q \cap R) \log \frac{\mu(P \cap R)}{\mu(R)}. \end{aligned}$$

A soma do lado direito pode ser reescrita como

$$\sum_{Q \in \mathscr{Q}} \sum_{S \in \mathscr{P} \bigvee \mathscr{R}} -\mu(Q \cap S) \log \frac{\mu(Q \cap S)}{\mu(S)} + \sum_{P \in \mathscr{P}} \sum_{R \in \mathscr{R}} \sum_{Q \in \mathscr{Q}} -\mu((P \cap Q) \cap R) \log \frac{\mu(P \cap R)}{\mu(R)}$$

$$= \sum_{Q\in\mathscr{Q}} \sum_{S\in\mathscr{P}\bigvee\mathscr{R}} -\mu(Q\cap S)\log\frac{\mu(Q\cap S)}{\mu(S)} + \sum_{P\in\mathscr{P}}\sum_{R\in\mathscr{R}} -\mu(P\cap R)\log\frac{\mu(P\cap R)}{\mu(R)}$$

$$= H_\mu(\mathscr{Q}\setminus(\mathscr{P}\bigvee\mathscr{R})) + H_\mu(\mathscr{P}\setminus\mathscr{R}).$$

Isto demonstra o item *(i)*. Agora, observe que se $\mathscr{P} \prec \mathscr{Q}$ , então cada $P\in\mathscr{P}$ se escreve como $\bigcup_{Q\subset P} Q$, onde cada $Q\in\mathscr{Q}$. Portanto,

$$\begin{aligned}
H_\mu(\mathscr{P}\setminus\mathscr{R}) &= \sum_{P\in\mathscr{P}}\sum_{R\in\mathscr{R}} -\mu(P\cap R)\log\frac{\mu(P\cap R)}{\mu(R)}\\
&= \sum_{P\in\mathscr{P}}\sum_{R\in\mathscr{R}}\sum_{Q\subset P} -\mu(Q\cap R)\log\frac{\mu(P\cap R)}{\mu(R)}\\
&\le \sum_{P\in\mathscr{P}}\sum_{R\in\mathscr{R}}\sum_{Q\subset P} -\mu(Q\cap R)\log\frac{\mu(Q\cap R)}{\mu(R)} = H_\mu(\mathscr{Q}\setminus\mathscr{R}).
\end{aligned}$$

Note ainda que, se $P\in\mathscr{P}$ e $R\in\mathscr{R}$, temos

$$\frac{\mu(R\cap P)}{\mu(P)} = \sum_{Q\subset P}\frac{\mu(Q)}{\mu(P)}\frac{\mu(R\cap Q)}{\mu(Q)}.$$

Uma vez que $\sum_{Q\subset P}\dfrac{\mu(Q)}{\mu(P)} = 1$, aplicamos o Corolário 2.3 à função convexa $-\phi$, onde $\phi(x) = x\log x$, à sequência $(\lambda_Q)_{Q\subset P} = \dfrac{\mu(Q)}{\mu(P)}$ e à sequência $(a_Q)_{Q\subset P} = \dfrac{\mu(R\cap Q)}{\mu(Q)}$. Assim, obtemos

$$\phi\left(\frac{\mu(R\cap P)}{\mu(P)}\right) \ge \sum_{Q\subset P}\frac{\mu(Q)}{\mu(P)}\phi\left(\frac{\mu(R\cap Q)}{\mu(Q)}\right)$$

para todo $P\in\mathscr{P}$ e $R\in\mathscr{R}$. Consequentemente

$$\begin{aligned}
H_\mu(\mathscr{R}\setminus\mathscr{P}) &= \sum_{P,R}\mu(P)\phi\left(\frac{\mu(R\cap P)}{\mu(P)}\right) \ge \sum_{P,R}\mu(P)\sum_{Q\subset P}\frac{\mu(Q)}{\mu(P)}\phi\left(\frac{\mu(R\cap Q)}{\mu(Q)}\right)\\
&= \sum_{Q,R}\mu(Q)\phi\left(\frac{\mu(R\cap Q)}{\mu(Q)}\right) = H_\mu(\mathscr{R}\setminus\mathscr{Q}).
\end{aligned}$$

Isto prova o item *(ii)*. Finalmente, usando a definição de entropia condicional dada em (3.1), observamos o seguinte: $H_\mu(\mathscr{P}\setminus\mathscr{Q}) = 0$ se, e somente se, para todo $P\in\mathscr{P}$ e $Q\in\mathscr{Q}$ tem-se $\mu(P\cap Q) = 0$ ou $\dfrac{\mu(P\cap Q)}{\mu(Q)} = 1$.

Em outras palavras, ou $Q$ é disjunto de $P$ (a menos de medida nula) ou $Q$ está contido em $P$ (a menos de medida nula). Isto quer dizer que $H_\mu(\mathscr{P}\setminus\mathscr{Q}) = 0$ se, e somente se, $\mathscr{P} \prec \mathscr{Q}$. $\square$

**Definição 3.4.** Seja $f : M \to M$ uma transformação mensurável, não necessariamente invertível preservando uma medida de probabilidade $\mu$. Dada uma partição $\mathscr{P}$ de $M$ com entropia finita, denotamos

$$\mathscr{P}^n = \bigvee_{i=0}^{n-1} f^{-i}(\mathscr{P}) \text{ para cada } n \geq 1. \tag{3.2}$$

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

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

Podemos notar que a sequência $\mathscr{P}^n$ é não-decrescente, isto é, $\mathscr{P}^n \prec \mathscr{P}^{n+1}$ para todo $n$, e portanto, pelo Lema 3.3 item *(ii)* aplicado à sequência $\mathscr{P}^n$ e com a partição $\mathscr{R}$ sendo tomada como vazia, concluímos que a sequência das entropias $H_\mu(\mathscr{P}^n)$ também é não-decrescente.

Chamamos **entropia de $f$ com respeito à medida $\mu$ e à partição $\mathscr{P}$** o limite

$$h_\mu(f, \mathscr{P}) = \lim_n \frac{1}{n} H_\mu(\mathscr{P}^n) = \inf_n \frac{1}{n} H_\mu(\mathscr{P}^n). \tag{3.3}$$

Esta entropia é tanto maior quanto mais fina for a partição. De fato, se $\mathscr{P} \prec \mathscr{Q}$ então $\mathscr{P}^n \prec \mathscr{Q}^n$ para todo $n$. Usando o Lema 3.3 item *(ii)*, segue que $H_\mu(\mathscr{P}^n) \leq H_\mu(\mathscr{Q}^n)$ para todo $n$. Consequentemente,

$$\mathscr{P} \prec \mathscr{Q} \quad \Rightarrow \quad h_\mu(f, \mathscr{P}) \leq h_\mu(f, \mathscr{Q}). \tag{3.4}$$

Finalmente, **a entropia do sistema dinâmico** $(f, \mu)$ é definida por

$$h_\mu(f) = \sup_{\mathscr{P}} h_\mu(f, \mathscr{P}) \tag{3.5}$$

onde o supremo é tomado sobre todas as partições com entropia finita.

Mostraremos agora que o limite em (3.3) existe sempre, e de fato é o ínfimo.

**Lema 3.5.** *Se $(a_n)_n$ é uma sequência de números reais tais que $a_{n+p} \leq a_n + a_p$ para todo $n$ e $p$, então*

$$\lim_n \frac{a_n}{n} = \inf_n \frac{a_n}{n} \in [-\infty, \infty).$$

*Demonstração.* Fixe $p > 0$. Cada $n > 0$ pode ser escrito na forma $n = kp + i$, com $0 \leq i < p$. Então

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

Se $n \to \infty$, então $k \to \infty$. Assim

$$\limsup_n \frac{a_n}{n} \leq \frac{a_p}{p}$$

e portanto

$$\limsup_n \frac{a_n}{n} \leq \inf \frac{a_p}{p}.$$

Mas,

$$\inf \frac{a_p}{p} \leq \liminf_n \frac{a_n}{n}.$$

Comparando as últimas duas desigualdades, concluímos que $\lim_n \dfrac{a_n}{n} = \inf_n \dfrac{a_n}{n}$. Se $a_m = -\infty$ para algum $m$ então, pela subaditividade, temos que $a_n = -\infty$ para todo $n > m$. Daí, $\lim_n \dfrac{a_n}{n} = -\infty = \inf_n \dfrac{a_n}{n}$. □

**Lema 3.6.** *Seja $f : M \to M$ uma transformação mensurável preservando uma probabilidade $\mu$. Então $H_\mu(\mathscr{P}^{m+n}) \leq H_\mu(\mathscr{P}^m) + H_\mu(\mathscr{P}^n)$ para todo $m, n \geq 1$.*

*Demonstração.* Por definição, $\mathscr{P}^{m+n} = \bigvee_{i=0}^{m+n-1} f^{-i}(\mathscr{P}) = \mathscr{P}^m \bigvee f^{-m}(\mathscr{P}^n)$. Portanto, usando o item *(i)* do Lema 3.3 e considerando a partição $\mathscr{R} = \mathscr{M} = \{M\}$, vem que

$$\begin{aligned}
H_\mu(\mathscr{P}^{m+n}) = H_\mu(\mathscr{P}^m \bigvee f^{-m}(\mathscr{P}^n)) &= H_\mu((\mathscr{P}^m \bigvee f^{-m}(\mathscr{P}^n)) \setminus \mathscr{M}) \\
&= H_\mu(\mathscr{P}^m \setminus \mathscr{M}) + H_\mu(f^{-m}(\mathscr{P}^n) \setminus (\mathscr{P}^m \bigvee \mathscr{M})) \\
&\leq H_\mu(\mathscr{P}^m) + H_\mu(f^{-m}(\mathscr{P}^n)) \\
&= H_\mu(\mathscr{P}^m) + H_\mu(\mathscr{P}^n).
\end{aligned}$$

□

Levando em conta o Lema 3.6 e aplicando o Lema 3.5 a $H_\mu(\mathscr{P}^n)$, concluímos que o limite em (3.3) sempre existe e é o ínfimo.

**Observação 3.7.** *Na prova do Lema 3.6 usamos os seguintes fatos gerais:*

*1.* $H_\mu(\mathscr{P} \setminus \mathscr{M}) = H_\mu(\mathscr{P})$ *para toda partição* $\mathscr{P}$*, onde* $\mathscr{M}$ *denota a partição trivial* $\mathscr{M} = \{M\}$*. A prova disso é simplesmente usar a definição de entropia condicional:*

$$H_\mu(\mathscr{P} \setminus \mathscr{M}) = \sum_{P \in \mathscr{P}} -\mu(P \cap M) \log \frac{\mu(P \cap M)}{\mu(M)} = \sum_{P \in \mathscr{P}} -\mu(P) \log \mu(P) = H_\mu(\mathscr{P}).$$

*2.* $H_\mu(\mathscr{R} \setminus \mathscr{Q}) \leq H_\mu(\mathscr{R})$ *para quaisquer partições* $\mathscr{Q}$ *e* $\mathscr{R}$*. Para provar este fato, basta usar a segunda desigualdade dada no item (ii) do Lema 3.3, tomando a partição* $\mathscr{P} = \mathscr{M} = \{M\}$*, que claramente satisfaz* $\mathscr{P} \prec \mathscr{Q}$ *e usar o item 1 desta observação.*

*3.* $H_\mu(f^{-m}(\mathscr{P}^n)) = H_{f_*\mu}(\mathscr{P}^n)$ *sempre que* $f : M \to N$ *é mensurável e* $\mu$ *é uma probabilidade em* $M$*. Faremos o caso em que* $m = 1$*. O caso geral segue por indução. Começe notando que* $f_*\mu$ *é uma probabilidade em* $N$*, e se* $\mathscr{P}$ *é uma partição em* $N$*, então* $f^{-1}(\mathscr{P}) = \{f^{-1}(P) : P \in \mathscr{P}\}$*é uma partição em* $M$*. Segue da definição de entropia da partição que:*

$$H_\mu(f^{-1}(\mathscr{P})) = \sum_{P \in \mathscr{P}} -\mu(f^{-1}(P)) \log \mu(f^{-1}(P)) = \sum_{P \in \mathscr{P}} -f_*\mu(P) \log f_*\mu(P) = H_{f_*\mu}(\mathscr{P}).$$

*Em particular, se* $M = N$ *e a medida* $\mu$ *é invariante por* $f$*, então* $f_*\mu = \mu$*, e portanto*

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

**Observação 3.8.** *Uma outra desigualdade que será útil é a seguinte:*

$$H_\mu(\mathscr{P} \bigvee \mathscr{Q}) = H_\mu(\mathscr{P}) + H_\mu(\mathscr{Q} \setminus \mathscr{P}) \leq H_\mu(\mathscr{P}) + H_\mu(\mathscr{Q}). \tag{3.6}$$

*A igualdade, vem do item (i) do Lema 3.3 quando tomamos a partição* $\mathscr{R} = \mathscr{M} = \{M\}$*, junto com o item 1 da observação 3.7 e notando que* $\mathscr{P} \bigvee \mathscr{M} = \mathscr{P}$*. A desigualdade vem da observação 3.7 item 2.*

Em geral, a principal dificuldade no cálculo da entropia reside no cálculo do supremo em (3.5). Os resultados seguintes permitem simplificar esta tarefa em muitos casos, identificando certas partições $\mathscr{P}$ que realizam o supremo. O resultado principal é o seguinte:

**Teorema 3.9 (Kolmogorov-Sinai).** *Seja* $\mathscr{P}_1 \prec \mathscr{P}_2 \prec \cdots \prec \mathscr{P}_n \prec \cdots$ *uma sequência não-decrescente de partições com entropia finita tais que* $\cup_{n=1}^{\infty} \mathscr{P}_n$ *gera a* $\sigma$*-álgebra dos conjuntos mensuráveis a menos de medida nula. Então*

$$h_\mu(f) = \lim_n h_\mu(f, \mathscr{P}_n). \tag{3.7}$$

Precisaremos de três lemas técnicos para demonstrar o Teorema de Kolmogorov-Sinai:

**Lema 3.10.** *Dado* $k \geq 1$ *e* $\varepsilon > 0$ *existe* $\delta > 0$ *tal que, para quaisquer partições finitas* $\mathscr{P} = \{P_1, \ldots, P_k\}$ *e* $\mathscr{Q} = \{Q_1, \ldots, Q_k\}$,

$$\mu(P_i \Delta Q_i) < \delta \ \ \textit{para todo} \ \ i = 1, \ldots, k \quad \Rightarrow \quad H_\mu(\mathscr{Q} \setminus \mathscr{P}) < \varepsilon.$$

*Demonstração.* Fixe $k \geq 1$ e $\varepsilon > 0$. Pela continuidade da função $\phi : [0,1] \to \mathbb{R}$, $\phi(x) = -x \log x$, existe $\gamma > 0$ tal que $\phi(x) < \dfrac{\varepsilon}{k^2}$ para todo $x \in [0, \gamma) \cup (1-\gamma, 1]$. Tome $\delta = \dfrac{\gamma}{k}$. Dadas partições $\mathscr{P}$ e $\mathscr{Q}$ como no enunciado, denote por $\mathscr{R}$ 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 $\mu(P_i \cap Q_j) \leq \mu(P_i \Delta Q_i) < \delta$ para todo $i \neq j$, pois $Q_j \subset Q_i^c$, e portanto $P_i \cap Q_j \subset (P_i \cap Q_i^c) \cup (Q_i \cap P_i^c)$. Então

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

Portanto,

$$H_\mu(\mathscr{R}) = \sum_{R \in \mathscr{R}} -\mu(R) \log \mu(R) = \sum_{R \in \mathscr{R}} \phi(\mu(R)) < \sum_{R \in \mathscr{R}} \frac{\varepsilon}{k^2} = \#\mathscr{R} \frac{\varepsilon}{k^2} \leq \varepsilon.$$

Vem da definição de $\mathscr{R}$ que $\mathscr{P} \bigvee \mathscr{Q} = \mathscr{P} \bigvee \mathscr{R}$. Usando a igualdade dada na equação (3.6) da observação 3.8

$$H_\mu(\mathscr{Q} \setminus \mathscr{P}) = H_\mu(\mathscr{P} \bigvee \mathscr{Q}) - H_\mu(\mathscr{P}) = H_\mu(\mathscr{P} \bigvee \mathscr{R}) - H_\mu(\mathscr{P}) = H_\mu(\mathscr{R} \setminus \mathscr{P}).$$

Finalmente, aplicando o item 2 da observação 3.7 a igualdade acima, obtemos

$$H_\mu(\mathscr{Q} \setminus \mathscr{P}) = H_\mu(\mathscr{R} \setminus \mathscr{P}) \leq H_\mu(\mathscr{R}) < \varepsilon.$$

□

**Lema 3.11.** *Seja* $\mathscr{P}_1 \prec \cdots \prec \mathscr{P}_n \prec \cdots$ *como no Teorema de Kolmogorov-Sinai. Então* $\lim_n H_\mu(\mathscr{Q} \setminus \mathscr{P}_n) = 0$ *para qualquer partição finita* $\mathscr{Q}$*.*

*Demonstração.* Escreva $\mathscr{Q} = \{Q_1, \ldots, Q_k\}$. Dado qualquer $\varepsilon > 0$, fixe $\delta > 0$ como no Lema 3.10. Seja $\mathcal{A}$ a álgebra formada pelas uniões finitas de elementos de $\cup_n \mathscr{P}_n$. Por hipótese, $\mathcal{A}$ gera a $\sigma$-álgebra de todos os conjuntos mensuráveis. Logo, pelo Teorema de aproximação (Teorema 2.4), para cada $i = 1, \ldots, k$ existe $A_i \in \mathcal{A}$ tal que

$$\mu(Q_i \Delta A_i) < \frac{\delta}{4k}. \tag{3.8}$$

O fato de que os $Q_i$ são cobertura de $M$ garante que os $A_i$ estão perto de o serem também:

$$\mu(A_i \cap (\cup_{j \neq i} A_j)) \leq \mu(\cup_{j=1}^{k} (A_j \setminus Q_j)) < \frac{\delta}{4} \text{ para todo } i \tag{3.9}$$

e

$$\mu(M \setminus \cup_{i=1}^{k} A_i) \leq \mu(\cup_{i=1}^{k} (Q_i \setminus A_i)) < \frac{\delta}{4}. \tag{3.10}$$

Em seguida, defina

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

Então $\mathscr{Q}' = \{Q_1', \cdots, Q_k'\}$ é uma partição de $M$. Afirmamos que

$$\mu(A_i \Delta Q_i') < \frac{\delta}{2} \text{ para todo } i = 1, \cdots, k. \tag{3.11}$$

Isto é imediato para $i = 1$. Para $i > 1$ temos que $A_i \setminus Q_i'$ está contido em $A_i \cap (\cup_{j<i} A_j)$.Logo, usando (3.9), obtemos que $\mu(A_i \setminus Q_i') < \frac{\delta}{4}$. Isto prova a afirmação para todo $1 < i < k$, uma vez que nesse caso $Q_i' \setminus A_i =$. Finalmente, para $i = k$, temos que $Q_k' \setminus A_k$ está contido no complementar de $\cup_{i=1}^{k} A_i$. Logo, usando (3.10), vemos que $\mu(Q_k' \setminus A_k) < \frac{\delta}{4}$. Por outro lado $A_k \setminus Q_k' = A_k \cap (\cup_{j<k} A_j)$ e portanto pela estimativa em (3.9) $\mu(A_k \setminus Q_k') < \frac{\delta}{4}$. Somando esta estimativa com a anterior, vem que $\mu(A_i \Delta Q_i') < \frac{\delta}{2}$. Isto completa a prova da afirmação (3.11).

Combinando as desigualdades (3.8) e (3.11), obtemos que

$$\mu(Q_i \Delta Q_i') \leq \mu(Q_i \Delta A_i) + \mu(A_i \Delta Q_i') \leq \delta \text{ para todo } i = 1, \cdots, k.$$

Agora, é claro que $Q'_i \in \mathcal{A}$ para todo $i$. Então, como se trata de uma família finita, podemos encontrar $m \geq 1$ tal que todo $Q'_i$ é uma união de elementos de $\mathscr{P}_m$. Em outras palavras, a partição $\mathscr{Q}' = \{Q'_1, \cdots, Q'_k\}$ é menos fina do que $\mathscr{P}_m$. Então, notando que $\mathscr{Q}' \prec \mathscr{P}_m \prec \mathscr{P}_n$ para $n \geq m$, podemos usar o Lema 3.3 item *(ii)* duas vezes para obter

$$H_\mu(\mathscr{Q} \setminus \mathscr{P}_n) \leq H_\mu(\mathscr{Q} \setminus \mathscr{P}_m) \leq H_\mu(\mathscr{Q} \setminus \mathscr{Q}') \text{ para todo } n \geq m.$$

Usando o Lema 3.10 para as partições $\mathscr{Q}$ e $\mathscr{Q}'$

$$H_\mu(\mathscr{Q} \setminus \mathscr{P}_n) \leq H_\mu(\mathscr{Q} \setminus \mathscr{P}_m) \leq H_\mu(\mathscr{Q} \setminus \mathscr{Q}') < \varepsilon \text{ para todo } n \geq m.$$

Isto completa a demonstração.

□

**Lema 3.12.** *$h_\mu(f, \mathscr{Q}) \leq h_\mu(f, \mathscr{P}) + H_\mu(\mathscr{Q} \setminus \mathscr{P})$ para quaisquer partições $\mathscr{P}$ e $\mathscr{Q}$ com entropia finita.*

*Demonstração.* Recordando da definição de $\mathscr{P}^n$ dada em (3.2) e pelo Lema 3.3 item *(i)*, para todo $n \geq 1$ vale que

$$\begin{aligned} H_\mu(\mathscr{Q}^{n+1} \setminus \mathscr{P}^{n+1}) &= H_\mu((\mathscr{Q}^n \bigvee f^{-n}(\mathscr{Q})) \setminus (\mathscr{P}^n \bigvee f^{-n}(\mathscr{P}))) \\ &\leq H_\mu(\mathscr{Q}^n \setminus \mathscr{P}^n) + H_\mu(f^{-n}(\mathscr{Q}) \setminus f^{-n}(\mathscr{P})). \end{aligned}$$

O último termo é igual a $H_\mu(\mathscr{Q} \setminus \mathscr{P})$, porque a medida $\mu$ é invariante por $f$. Portanto a relação anterior prova que

$$H_\mu(\mathscr{Q}^n \setminus \mathscr{P}^n) \leq n H_\mu(\mathscr{Q} \setminus \mathscr{P}) \text{ para todo } n \geq 1. \tag{3.12}$$

Lembrando que $\mathscr{Q}^n \prec \mathscr{P}^n \bigvee \mathscr{Q}^n$ (Definição 3.2 ) e usando novamente o Lema 3.3 item *(i)*, segue que

$$H_\mu(\mathscr{Q}^n) \leq H_\mu(\mathscr{P}^n \bigvee \mathscr{Q}^n) = H_\mu(\mathscr{P}^n) + H_\mu(\mathscr{Q}^n \setminus \mathscr{P}^n) \leq H_\mu(\mathscr{P}^n) + n H_\mu(\mathscr{Q} \setminus \mathscr{P}).$$

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

***Demonstração do Teorema de Kolmogorov-Sinai.*** O limite em (3.7) sempre existe, pois a propriedade (3.4) implica que a sequência $h_\mu(f, \mathscr{P}_n)$ é não decrescente. Portanto, já temos que

$$h_\mu(f) = \sup_{\mathscr{Q}} h_\mu(f, \mathscr{Q}) \geq \lim_n h_\mu(f, \mathscr{P}_n).$$

Por outro lado, pelo Lema 3.12

$$h_\mu(f, \mathscr{Q}) \leq h_\mu(f, \mathscr{P}_n) + H_\mu(\mathscr{Q} \setminus \mathscr{P}_n) \quad \text{para todo} \quad n.$$

Passando ao limite quando $n \to \infty$ e usando o Lema 3.11

$$h_\mu(f, \mathscr{Q}) \leq \lim_n h_\mu(f, \mathscr{P}_n).$$

Finalmente, tomando o supremo sobre todas as partições finitas $\mathscr{Q}$, obtemos

$$h_\mu(f) = \sup_{\mathscr{Q}} h_\mu(f, \mathscr{Q}) \leq \lim_n h_\mu(f, \mathscr{P}_n).$$

Isto conclui a prova. $\square$

**Corolário 3.13.** *Seja $\mathscr{P}$ uma partição com entropia finita tal que a união dos seus iterados $\mathscr{P}^n = \bigvee_{i=0}^{n-1} f^{-i}(\mathscr{P})$ para cada $n \geq 1$ gera a $\sigma$-álgebra dos conjuntos mensuráveis. Então $h_\mu(f) = h_\mu(f, \mathscr{P})$.*

*Demonstração.* Primeiro provamos que se $\mathscr{P}$ é uma partição com entropia finita, então

$$h_\mu(f, \mathscr{P}) = h_\mu(f, \mathscr{P}^n) \text{ para todo } n \geq 1. \tag{3.13}$$

De fato, observe que, dado qualquer $k \geq 1$

$$\bigvee_{i=0}^{k-1} f^{-i}(\mathscr{P}^n) = \bigvee_{i=0}^{k-1} f^{-i}(\bigvee_{j=0}^{n-1} f^{-j}(\mathscr{P})) = \bigvee_{l=0}^{n+k-2} f^{-l}(\mathscr{P}) = \mathscr{P}^{n+k-1}.$$

Portanto,

$$h_\mu(f, \mathscr{P}^n) = \lim_k \frac{1}{k} H_\mu(\mathscr{P}^{n+k-1}) = \lim_k \frac{1}{k} H_\mu(\mathscr{P}^k) = h_\mu(f, \mathscr{P}).$$

Agora a prova do corolário se resume a aplicar o Teorema de Kolmogorov-Sinai à sequência $\mathscr{P}^n$, observando que a equação (3.13) vale para todo $n$. $\square$

Uma partição $\mathscr{P}$ como no Corolário 3.13 é chamada **partição geradora**, ou um **gerador** do sistema. Outras consequências do Teorema de Kolmogorov-Sinai são:

**Corolário 3.14.** *Seja* $\mathscr{P}_1 \prec \mathscr{P}_2 \prec \cdots \prec \mathscr{P}_n \prec \cdots$ *uma sequência não decrescente de partições com entropia finita tais que* $\operatorname{diam} \mathscr{P}_n(x) \to 0$ *para* $\mu$*-quase todo* $x \in M$*. Então*

$$h_\mu(f) = \lim_n h_\mu(f, \mathscr{P}_n)$$

**Corolário 3.15.** *Seja* $\mathscr{P}$ *uma partição com entropia finita tal que, para* $\mu$*-quase todo* $x \in M$*, tem-se* $\operatorname{diam} \mathscr{P}^n(x) \to 0$*. Então* $h_\mu(f) = h_\mu(f, \mathscr{P})$*.*

## 3.2 Exemplos: entropia de alguns sistemas

**Exemplo 3.16.** Suponhamos que a medida invariante $\mu$ está suportada numa órbita periódica. Em outras palavras, existe $x$ em $M$ e $k \geq 1$ tal que $f^k(x) = x$ e a medida $\mu$ é dada por

$$\mu = \frac{1}{k}(\delta_x + \delta_{f(x)} + \cdots + \delta_{f^{k-1}(x)}).$$

Neste caso a medida só toma um número finito de valores. Consequentemente, a entropia $H_\mu(\mathscr{P})$ também só toma um número finito de valores quando consideramos todas as partições enumeráveis $\mathscr{P}$. Em particular, $\lim_n \dfrac{1}{n} H_\mu(\mathscr{P}^n) = 0$ para toda partição $\mathscr{P}$. Isto prova que neste caso $h_\mu(f) = 0$.

**Exemplo 3.17.** Considere a transformação expansão decimal $f : [0,1] \to [0,1]$, dada por $f(x) = 10x - \lfloor 10x \rfloor$. Sabe-se que $f$ preserva a medida de Lebesgue $\mu$ no intervalo $[0,1]$. Seja $\mathscr{P}$ a partição de $[0,1]$ nos intervalos da forma $\left(\dfrac{i-1}{10}, \dfrac{i}{10}\right]$ com $i = 1, \cdots, 10$. Então $\mathscr{P}^n$ é a partição nos intervalos da forma $\left(\dfrac{i-1}{10^n}, \dfrac{i}{10^n}\right]$ com $i = 1, \cdots, 10^n$.

Consequentemente

$$
\begin{aligned}
H_\mu(\mathscr{P}^n) &= \sum_{(\frac{i-1}{10^n}, \frac{i}{10^n}] \in \mathscr{P}^n} -\mu(\left(\frac{i-1}{10^n}, \frac{i}{10^n}\right]) \log \mu(\left(\frac{i-1}{10^n}, \frac{i}{10^n}\right]) \\
&= \sum_{i=0}^{10^n} -\frac{1}{10^n} \log \frac{1}{10^n} \\
&= \sum_{i=0}^{10^n} \frac{1}{10^n} \log 10^n \\
&= n \log 10.
\end{aligned}
$$

Portanto,

$$
h_\mu(f, \mathscr{P}) = \lim_n \frac{1}{n} H_\mu(\mathscr{P}^n) = \log 10.
$$

Agora, como $\mathscr{P}$ é uma partição geradora, segue do Corolário 3.13 que $_\mu(f) = h_\mu(f, \mathscr{P}) = \log 10$.

**Exemplo 3.18.** Considere o deslocamento $\sigma : \Sigma \to \Sigma$ em $\Sigma = \{1, \ldots, d\}^{\mathbb{N}}$, munido de uma medida de Bernoulli $\mu = \nu^{\mathbb{N}}$, e denotemos $\nu(\{i\}) = p_i$ para cada $i \in \{1, \ldots, d\}$. Seja $\mathscr{P}$ a partição de $\Sigma$ em cilindros $[0; a]$ com $a = 1, \ldots, d$. Então $\mathscr{P}^n$ é a partição em cilindros $[0; a_1, \ldots, a_n]$ de comprimento $n$. A entropia da partição $\mathscr{P}^n$ é

$$
\begin{aligned}
H_\mu(\mathscr{P}^n) &= \sum_{[0;a_1,\ldots,a_n] \in \mathscr{P}^n} -\mu([0; a_1, \ldots, a_n]) \log \mu([0; a_1, \ldots, a_n]) \\
&= \sum_{a_1,\ldots,a_n} -p_{a_1} \ldots p_{a_n} \log(p_{a_1} \ldots p_{a_n}) \\
&= \sum_{a_1,\ldots,a_n} -p_{a_1} \ldots p_{a_n} \sum_j \log(p_{a_j}) \\
&= \sum_j \sum_{a_1,\ldots,a_n} -p_{a_1} \ldots p_{a_n} \log(p_{a_j}) \\
&= \sum_j \sum_{a_j} -p_{a_j} \log(p_{a_j}) \sum_{a_i : i \neq j} p_{a_1} \ldots p_{a_{j-1}} p_{a_{j+1}} \ldots p_{a_n}.
\end{aligned}
$$

Temos que $\sum_{a_i : i \neq j} p_{a_1} \ldots p_{a_{j-1}} p_{a_{j+1}} \ldots p_{a_n} = 1$ uma vez que $\sum_i p_i = 1$. Portanto

$$
H_\mu(\mathscr{P}^n) = \sum_{j=1}^{n} \sum_{a_j=1}^{d} -p_{a_j} \log(p_{a_j}) = \sum_{j=1}^{n} \sum_{r=1}^{d} -p_r \log(p_r) = -n \sum_{r=1}^{d} p_r \log(p_r)
$$

e

$$h_\mu(\sigma, \mathscr{P}) = \lim_n \frac{1}{n} H_\mu(\mathscr{P}^n) = -\sum_{r=1}^{d} p_r \log(p_r).$$

Agora, como $\mathscr{P}$ é uma partição geradora, segue do Corolário 3.13 que $h_\mu(\sigma) = h_\mu(\sigma, \mathscr{P}) = -\sum_{r=1}^{d} p_r \log(p_r)$.

**Exemplo 3.19** (**Entropia do Deslocamento de Markov**)**.** Seja $\Sigma = \{1, \ldots, d\}^{\mathbb{N}}$ e seja $\sigma : \Sigma \to \Sigma$ a transformação deslocamento. Seja $\mu$ uma medida de Markov associada a uma matriz estocástica $P = (P_{i,j})_{i,j}$ e um vetor de probabilidade $p = (p_i)_i$. Provaremos neste exemplo que:

$$h_\mu(\sigma) = \sum_{a=1}^{d} p_a \sum_{b=1}^{d} -P_{a,b} \log(P_{a,b}).$$

Considere $\mathscr{P}$ a partição de $\Sigma$ em cilindros $[0; a]$ com $a = 1, \ldots, d$. Então, para cada $n$, o iterado $\mathscr{P}^n$ é a partição em cilindros $[0; a_1, \ldots, a_n]$ de comprimento $n$. Lembrando que $\mu([0; a_1, \ldots, a_n]) = p_{a_1} P_{a_1,a_2} \ldots P_{a_{n-1},a_n}$, Temos o seguinte:

$$\begin{aligned}
H_\mu(\mathscr{P}^n) &= \sum_{[0;a_1,\ldots,a_n]\in\mathscr{P}^n} -\mu([0; a_1, \ldots, a_n]) \log \mu([0; a_1, \ldots, a_n]) \\
&= \sum_{a_1,\ldots,a_n} -p_{a_1} P_{a_1,a_2} \ldots P_{a_{n-1},a_n} \log(p_{a_1} P_{a_1,a_2} \ldots P_{a_{n-1},a_n}) \\
&= \sum_{a_1,\ldots,a_n} -p_{a_1} P_{a_1,a_2} \ldots P_{a_{n-1},a_n} \log(p_{a_1}) + \sum_{a_1,\ldots,a_n} -p_{a_1} P_{a_1,a_2} \ldots P_{a_{n-1},a_n} \log(P_{a_1,a_2}) + \\
&\qquad \ldots + \sum_{a_1,\ldots,a_n} -p_{a_1} P_{a_1,a_2} \ldots P_{a_{n-1},a_n} \log(P_{a_{n-1},a_n}) \\
&= \sum_{a_1} -p_{a_1} \log(p_{a_1}) \sum_{a_2,\ldots,a_n} P_{a_1,a_2} \ldots P_{a_{n-1},a_n} + \\
&\qquad \sum_{a_1,a_2} -\log(P_{a_1,a_2}) \sum_{a_3,\ldots,a_n} p_{a_1} P_{a_1,a_2} \ldots P_{a_{n-1},a_n} + \\
&\qquad \ldots + \sum_{a_{n-1},a_n} -\log(P_{a_{n-1},a_n}) \sum_{a_1,\ldots,a_{n-2}} p_{a_1} P_{a_1,a_2} \ldots P_{a_{n-1},a_n}.
\end{aligned}$$

Agora, vamos trabalhar sobre essa última igualdade. Começando com

$$\sum_{a_1} -p_{a_1} \log(p_{a_1}) \sum_{a_2,\ldots,a_n} P_{a_1,a_2} \ldots P_{a_{n-1},a_n}.$$

Usamos que $P$ é matriz estocástica (veja a definição 2.14) para concluir que

$\sum_{a_2,\dots,a_n} P_{a_1,a_2} \dots P_{a_{n-1},a_n} = 1$. Logo

$$\sum_{a_1} -p_{a_1} \log(p_{a_1}) \sum_{a_2,\dots,a_n} P_{a_1,a_2} \dots P_{a_{n-1},a_n} = \sum_{a_1} -p_{a_1} \log(p_{a_1}). \tag{3.14}$$

A soma que restou pode ser abordada da seguinte forma: tome $2 \leq j \leq n$ e estudemos

$$\sum_{a_{j-1},a_j} -\log(P_{a_{j-1},a_j}) \sum p_{a_1} P_{a_1,a_2} \dots P_{a_{n-1},a_n} \tag{3.15}$$

onde a última soma é sobre todos os valores de $a_1, \dots, a_{j-2}, a_{j+1} \dots, a_n$. Vamos então reescrever (3.15) como

$$\sum_{a_{j-1},a_j} -\log(P_{a_{j-1},a_j}) \sum \left( \sum_{a_1} p_{a_1} P_{a_1,a_2} \right) P_{a_2,a_3} \dots P_{a_{n-1},a_n}. \tag{3.16}$$

Lembrando também que $P^* p = p$, usamos o Lema 2.17 item b) e obtemos

$$\sum_{a_{j-1},a_j} -\log(P_{a_{j-1},a_j}) \sum p_{a_2} P_{a_2,a_3} \dots P_{a_{n-1},a_n} \tag{3.17}$$

onde a última soma é sobre todos os valores de $a_2, \dots, a_{j-2}, a_{j+1} \dots, a_n$. Fazemos sucessivas vezes os passos (3.16) e (3.17), porém, variando os $a_i$ até chegarmos em

$$\sum_{a_{j-1},a_j} -\log(P_{a_{j-1},a_j}) \sum p_{a_{j-1}} P_{a_{j-1},a_j} \dots P_{a_{n-1},a_n}.$$

onde a última soma é feita sobre todos os valores de $a_{j+1} \dots, a_n$. Então esta última equação pode ser reescrita como

$$\sum_{a_{j-1},a_j} -p_{a_{j-1}} P_{a_{j-1},a_j} \log(P_{a_{j-1},a_j}) \sum_{a_{j+1} \dots, a_n} P_{a_j,a_{j+1}} \dots P_{a_{n-1},a_n}.$$

Mas já vimos que a última soma é 1. Consequentemente, para todo $2 \leq j \leq n$ temos que

$$\sum_{a_{j-1},a_j} -\log(P_{a_{j-1},a_j}) \sum p_{a_1} P_{a_1,a_2} \dots P_{a_{n-1},a_n} = \sum_{a_{j-1},a_j} -p_{a_{j-1}} P_{a_{j-1},a_j} \log(P_{a_{j-1},a_j}). \tag{3.18}$$

Portanto, usando (3.14) e (3.18) podemos reescrever $H_\mu(\mathscr{P}^n)$ como

$$
\begin{aligned}
H_\mu(\mathscr{P}^n) &= \sum_{a_1} -p_{a_1}\log(p_{a_1}) + \sum_{j=2}^{n}\sum_{a_{j-1},a_j} -p_{a_{j-1}}P_{a_{j-1},a_j}\log(P_{a_{j-1},a_j}) \\
&= \sum_{a_1} -p_{a_1}\log(p_{a_1}) + (n-1)\sum_{a,b} -p_a P_{a,b}\log(P_{a,b}) \\
&= \sum_{a_1} -p_{a_1}\log(p_{a_1}) + (n-1)\sum_{a=1}^{d} p_a \sum_{b=1}^{d} -P_{a,b}\log(P_{a,b}).
\end{aligned}
$$

Dividindo esta última igualdade por $n$ e fazendo $n \to \infty$, obtemos que

$$
h_\mu(\sigma,\mathscr{P}) = \sum_{a=1}^{d} p_a \sum_{b=1}^{d} -P_{a,b}\log(P_{a,b}).
$$

Agora, como $\mathscr{P}$ é uma partição geradora, segue do Corolário 3.13 que $h_\mu(\sigma) = h_\mu(\sigma,\mathscr{P}) = \sum_{a=1}^{d} p_a \sum_{b=1}^{d} -P_{a,b}\log(P_{a,b})$.

## 3.3 Semicontinuidade da função entropia

Seja $\mathcal{M}_1(f)$ o conjunto das medidas de probabilidades invariantes por $f$. Vamos analisar agora a função $\nu \mapsto h_\nu(f)$ que associa a cada $\nu \in \mathcal{M}_1(f)$ a entropia $h_\nu(f)$. Em geral essa função não é contínua, porém, veremos que sob algumas hipóteses, ela é semicontínua superiormente em $\mu$: dado qualquer $\varepsilon > 0$, tem-se que $h_\nu(f) \le h_\mu(f) + \varepsilon$ para toda $\nu$ suficientemente próxima (na topologia fraca*) de $\mu$.

**Teorema 3.20** (**Semicontinuidade da Entropia**). *Seja $\mathscr{P}$ uma partição finita tal que $\mu(\partial\mathscr{P}) = 0$. Então, a função $\nu \mapsto h_\nu(f,\mathscr{P})$ é semicontínua superiormente em $\mu$. Se além disso, $\cup_n \mathscr{P}^n$ gera a $\sigma$-álgebra dos subconjuntos mensuráveis de $M$ a menos de medida nula. Então a função $\nu \mapsto h_\nu(f)$ é semicontínua superiormente em $\mu$.*

*Demonstração.* Considere qualquer partição finita $\mathscr{P}$ de $M$ cujo bordo $\partial\mathscr{P} = \bigcup_{P\in\mathscr{P}} \partial P$ satisfaz $\mu(\partial\mathscr{P}) = 0$. Devido ao Teorema Portmanteau (veja a seção 2.1 de [7]), a função $\nu \mapsto \nu(P)$ é contínua no ponto $\mu$, para todo $P \in \mathscr{P}$. Consequentemente, a função

$$
\nu \mapsto H_\nu(\mathscr{P}) = \sum_{P\in\mathscr{P}} -\nu(P)\log\nu(P)
$$

também é contínua em $\mu$. A hipótese sobre $\mathscr{P}$ também implica que $\mu(\partial\mathscr{P}^n) = 0$ para todo $n \geq 1$, uma vez que $\partial\mathscr{P}^n \subset \partial\mathscr{P} \cup f^{-1}(\partial\mathscr{P}) \cup \ldots \cup f^{-n+1}(\partial\mathscr{P})$. Assim, a função $\nu \mapsto H_\nu(\mathscr{P}^n)$ é contínua em $\mu$ para todo $n$. Lembrando de (3.3), temos que

$$h_\nu(f, \mathscr{P}) = \inf_n \frac{1}{n} H_\nu(\mathscr{P}^n)$$

é semicontínua superiormente em $\mu$, porque o ínfimo de qualquer família de funções contínuas é uma função semicontínua superiormente. Agora, suponha que $\cup_n \mathscr{P}^n$ gera a $\sigma$-álgebra dos subconjuntos mensuráveis de $M$ a menos de medida nula. Então, dada $\nu$ próxima de $\mu$ e $\varepsilon > 0$, pela primeira parte da demonstração, $h_\nu(f, \mathscr{P}) \leq h_\mu(f, \mathscr{P}) + \varepsilon$. Como a entropia é não decrescente $h_\mu(f, \mathscr{P}) \leq h_\mu(f)$. Por outro lado, pelo Corolário 3.13, $h_\nu(f, \mathscr{P}) = h_\nu(f)$. Portanto, $h_\nu(f) \leq h_\mu(f) + \varepsilon$. $\square$

**Corolário 3.21.** *Suponha que $M$ é um espaço métrico compacto e que existe $\varepsilon_0 > 0$ tal que toda partição finita $\mathscr{P}$ com* $\operatorname{diam} \mathscr{P} < \varepsilon_0$ *satisfaz* $\lim_n \operatorname{diam} \mathscr{P}^n = 0$. *Então, a função $\nu \mapsto h_\nu(f)$ é semicontínua superiormente. Consequentemente, essa função é limitada e o seu supremo é atingido para alguma medida $\mu$.*

**Definição 3.22.** Uma transformação contínua $f : M \to M$ num espaço métrico compacto é dita **expansiva** se existe $\varepsilon_0 > 0$ tal que $d(f^j(x), f^j(y)) < \varepsilon_0$ para todo $j \in \mathbb{N}$ implica que $x = y$. Quando $f : M \to M$ é invertível, dizemos que ela é expansiva no sentido bilateral, se existe $\varepsilon_0 > 0$ tal que $d(f^j(x), f^j(y)) < \varepsilon_0$ para todo $j \in \mathbb{Z}$ implica que $x = y$. Nos dois casos, $\varepsilon_0$ é chamado **constante de expansividade** de $f$.

**Teorema 3.23.** *Seja $f : M \to M$ uma transformação expansiva num espaço métrico compacto e seja $\varepsilon_0 > 0$ uma constante de expansividade. Então tem-se* $\lim_n \operatorname{diam} \mathscr{P}^n = 0$ *para toda partição finita $\mathscr{P}$ com* $\operatorname{diam} \mathscr{P} < \varepsilon_0$.

Resulta do Teorema acima e do Corolário 3.21 que a função entropia de toda transformação expansiva num espaço métrico compacto é semicontínua superiormente e existem probabilidades invariantes $\mu$ cuja entropia $h_\mu(f)$ é máxima entre todas as probabilidades invariantes de $f$.

## 3.4 Entropia, Equivalência e Decomposição Ergódica

A entropia foi introduzida em Teoria Ergódica com o objetivo principal de distinguir sistemas que não são ergodicamente equivalentes, especialmente no caso de sistemas que são espectralmente equivalentes e que, portanto, não podem ser distinguidos por meio de invariantes espectrais. Provaremos abaixo que a entropia é, de fato, um invariante de equivalência ergódica. A definição de equivalência ergódica, pode ser consultada nas preliminares, definição 2.18.

**Teorema 3.24.** *Sejam $f : M \to M$ e $g : N \to N$ transformações preservando probabilidades $\mu$ em $M$ e $\nu$ em $N$. Se $(f,\mu)$ é ergodicamente equivalente a $(g,\nu)$, então $h_\mu(f) = h_\nu(g)$.*

*Demonstração.* Seja $\phi : M \to N$ uma equivalência ergódica entre os dois sistemas. Isto significa que $\phi_*\mu = \nu$ e existem conjuntos $X \subset M$ e $Y \subset N$ com medida total nos respectivos espaços, tais que $\phi$ é uma bijeção mensurável de $X$ em $Y$, com inversa mensurável. Além disso, vem da observação 2.19 que os conjuntos $X$ e $Y$ podem ser escolhidos invariantes. Seja $\mathscr{P}$ uma partição de $M$ com entropia finita para $\mu$. A sua restrição a $X$ é uma partição de $(X,\mu)$.

A respectiva imagem $\mathscr{Q} = \phi(\mathscr{P})$ é uma partição de $(Y,\nu)$ que, naturalmente, também podemos ver como uma partição de $(N,\nu)$. Note que

$$\begin{aligned} H_\nu(\mathscr{Q}) &= \sum_{Q\in\mathscr{Q}} -\nu(Q)\log\nu(Q) \\ &= \sum_{\phi(P)\in\mathscr{Q}} -\nu(\phi(P))\log\nu(\phi(P)) \\ &= \sum_{P\in\mathscr{P}} -\phi_*\nu(P)\log\phi_*\nu(P) \\ &= \sum_{P\in\mathscr{P}} -\mu(P)\log\mu(P) \\ &= H_\mu(\mathscr{P}). \end{aligned}$$

Como

$$\mathscr{Q}^n = \bigvee_{j=0}^{n-1} g^{-j}(\mathscr{Q}) = \bigvee_{j=0}^{n-1} g^{-j}(\phi(\mathscr{P})) = \bigvee_{j=0}^{n-1} \phi(f^{-j}(\mathscr{P})) = \phi(\bigvee_{j=0}^{n-1}(f^{-j}(\mathscr{P}))) = \phi(\mathscr{P}^n)$$

para todo $n$, segue também que

$$h_\nu(g,\mathscr{Q}) = \lim_n \frac{1}{n} H_\nu(\mathscr{Q}^n) = \lim_n \frac{1}{n} H_\mu(\mathscr{P}^n) = h_\mu(f,\mathscr{P}).$$

Tomando o supremo sobre todas as partições $\mathscr{P}$, concluímos que $h_\nu(g) \geq h_\mu(f)$. Repetindo os mesmos cálculos para a também equivalência ergódica entre os sistemas $\psi = \phi^{-1} : N \to M$, obtemos $h_\nu(g) \leq h_\mu(f)$. Isto conclui a prova. $\square$

**Exemplo 3.25.** Kolmogorov mostrou que a entropia dos deslocamentos unilaterais $\left(\frac{1}{2},\frac{1}{2}\right)$ e $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right)$ são respectivamente, $\log 2, \log 3$. Usando o teorema acima, foi possível concluir que nem todos os deslocamentos de Bernoulli unilaterais são ergodicamente equivalentes.

No entanto, um resultado notável devido a Donald Ornstein afirma que a entropia é um invariante completo para os deslocamentos de Bernoulli bilaterais:

**Teorema 3.26** (**Ornstein**). *Dois deslocamentos de Bernoulli bilaterais em espaços de Lebesgue*[1] *são ergodicamente equivalentes se, e somente se, as suas entropias são iguais.*

A entropia $h_\mu(f)$ é sempre uma função afim da medida invariante $\mu$. Mais ainda, a propriedade de afinidade se estende para a decomposição ergódica (ver Teorema 2.8) de $\mu$:

**Teorema 3.27** (**Jacobs**). *Suponha que $M$ é um espaço métrico separável. Dada qualquer probabilidade invariante $\mu$, seja $\{\mu_P : P \in \mathscr{P}\}$ a sua decomposição ergódica. Então $h_\mu(f) = \int h_{\mu_P}(f) d\hat{\mu}(P)$ (quando um dos lados da igualdade é infinito o outro também é).*

## 3.5 Jacobianos e fórmula de Rokhlin

Seja $U$ um aberto do $\mathbb{R}^d$, seja $m$ a medida de Lebesgue e seja $f : U \to U$ um difeomorfismo local. Pela fórmula de mudança de variáveis,

$$m(f(A)) = \int_A |\det Df(x)| dx \tag{3.19}$$

[1] Se $M$ é um espaço métrico completo separável e $\mu$ é uma probabilidade boreliana sem átomos então $(M, \mathcal{B}, \mu)$ é um espaço de Lebesgue.

para todo $A$ contido numa bola restrita à qual $f$ é injetivo. Apresentaremos agora uma noção de jacobiano, que estende este tipo de relação para transformações e medidas muito mais gerais.

**Definição 3.28.** Seja $f : M \to M$ uma transformação mensurável. Diremos que $f$ é **localmente invertível** se existe alguma cobertura enumerável $\{U_k : k \geq 1\}$ de $M$ por conjuntos mensuráveis tais que a restrição de $f$ a cada $U_k$ é uma bijeção sobre a sua imagem, a qual é um conjunto mensurável, e a inversa dessa bijeção também é mensurável.

Os subconjuntos mensuráveis destes conjuntos $U_k$ serão chamados **domínios de invertibilidade**. Se $A$ é domínio de invertibilidade então $f(A)$ é um conjunto mensurável. Observe, igualmente, que se $f$ é localmente invertível então a pré-imagem $f^{-1}(y)$ de qualquer $y \in M$ é enumerável: ela contém no máximo um ponto em cada $U_k$.

Seja $\eta$ uma probabilidade em $M$, não necessariamente invariante por $f$. Uma função mensurável $\xi : M \to [0, \infty)$ é um **jacobiano** de $f$ relativamente a $\eta$ se a restrição de $\xi$ a qualquer domínio de invertibilidade $A$ é integrável com relação a $\eta$ e satisfaz

$$\eta(f(A)) = \int_A \xi d\eta. \tag{3.20}$$

**Proposição 3.29.** *A definição de jacobiano não depende da escolha da cobertura* $\{U_k : k \geq 1\}$.

*Demonstração.* Seja $\{V_m : m \geq 1\}$ uma outra cobertura. Todo subconjunto $B$ de algum $V_m$ pode ser escrito como união disjunta de conjuntos mensuráveis $A_k \subset U_k, k \geq 1$. Então, como os $A_k$ são domínios de invertibilidade, $\eta(f(A_k)) = \int_{A_k} \xi d\eta, k \geq 1$. Portanto, $\eta(f(B)) = \sum_k \eta(f(A_k)) = \sum_k \int_{A_k} \xi d\eta = \int_B \xi d\eta$. $\square$

**Exemplo 3.30.** Seja $\sigma : \Sigma \to \Sigma$ o deslocamento em $\Sigma = \{1, 2, \ldots, d\}^{\mathbb{N}}$ e seja $\mu$ a medida de Bernoulli associada a um vetor de probabilidade $p = (p_1, \ldots, p_d)$. Consideramos $\Sigma$ munido da distância $d((x_i)_{i\in\mathbb{N}}, (y_i)_{i\in\mathbb{N}}) = \theta^N$, onde $N = N((x_i)_{i\in\mathbb{N}}, (y_i)_{i\in\mathbb{N}}) = \inf\{n : x_n = y_n\}$ e $\theta \in (0, 1)$ é arbitrário. A restrição de $\sigma$ a cada cilindro $[0; a]$ é uma transformação invertível (Portanto $\{[0; a] : a \in \{1, 2, \ldots, d\}\}$ é a cobertura como na definição 3.28). Além disso, dado qualquer cilindro $[0; a, a_1, \ldots, a_n] \subset [0; a]$,

$$\mu(\sigma([0; a, a_1, \ldots, a_n])) = p_{a_1} \ldots p_{a_n} = \frac{1}{p_a}\mu([0; a, a_1, \ldots, a_n]) = \int_{[0;a,a_1,\ldots,a_n]} \frac{1}{p_a} d\mu.$$

Como a $\sigma$-álgebra gerada pelos cilindros elementares $[0; a, a_1, \ldots, a_n]$ coincide com a $\sigma$-álgebra gerada por todos os cilindros, então, naturalmente dado $A \subset [0; a]$ conjunto mensurável (domínio de invertibilidade), vale $\mu(\sigma(A)) = \int_A \frac{1}{p_a} d\mu$. Portanto $\xi((x_n)_n)) = \dfrac{1}{p_{x_0}}$ é um jacobiano de $\sigma$ relativamente a $\mu$.

**Definição 3.31.** Dizemos que uma medida $\eta$ é **não singular** com relação à transformação $f$ se a imagem de qualquer domínio de invertibilidade com medida nula também tem medida nula: se $\eta(A) = 0$ então $\eta(f(A)) = 0$. Por exemplo, se $f : U \to U$ é um difeomorfismo local num aberto de $\mathbb{R}^d$ e $\eta$ é a medida de Lebesgue, então $\eta$ é não singular, basta usar a fórmula de mudança de variáveis (3.19). Também toda probabilidade invariante é não singular.

Segue imediatamente da definição (3.20) que se $f$ admite jacobiano com relação a uma medida $\eta$ então essa medida é não singular.A recíproca também é verdadeira:

**Proposição 3.32.** *Seja $f : M \to M$ uma transformação localmente invertível e seja $\eta$ uma medida boreliana em $M$, não singular com relação a $f$. Então, existe algum jacobiano de $f$ com relação a $\eta$ e ele é essencialmente único: dois jacobianos quaisquer coincidem em $\eta$-quase todo ponto.*

*Demonstração.* Começemos por provar a existência. Por definição, $f$ localmente invertível implica que existe uma cobertura enumerável $\{U_k : k \geq 1\}$ de $M$ por domínios de invertibilidade de $f$. Defina, $P_1 = U_1$ e para cada $k > 1$, $P_k = U_k \setminus (U_1 \cup \ldots \cup U_{k-1})$. Então $\mathscr{P} = \{P_k : k \geq 1\}$ é uma partição de $M$ formada por domínios de invertibilidade. Para cada $P_k \in \mathscr{P}$, represente por $\eta_k$ a medida em definida em $P_k$ por $\eta_k(A) = \eta(f(A))$ para todo conjunto mensurável $A \subset P_k$. A hipótese de que $\eta$ é não singular implica que cada $\eta_k$ é absolutamente contínua com relação a $\eta$ restrita a $P_k$. Seja $\xi_k = \dfrac{d\eta_k}{d(\eta|_{P_k})}$ a derivada de Radón-Nykodim. Pelo Teorema de Radón-Nykodim $\xi_k$ é uma função definida em $P_k$, integrável com relação a $\eta$ e satisfazendo

$$\eta(f(A)) = \eta_k(A) = \int_A \xi_k d\eta \tag{3.21}$$

para todo conjunto mensurável $A \subset P_k$. Considere a função $\xi : M \to [0, \infty)$ cuja restrição a cada $P_k \in \mathscr{P}$ está dada por $\xi_k$. Todo subconjunto de $U_k$ pode ser escrito como união disjunta de subconjuntos de $P_1, \ldots, P_k$. Aplicando (3.21) a cada um desses subconjuntos e somando as respectivas igualdades, obtemos que $\eta(f(A)) = \int_A \xi d\eta$ para todo conjunto mensurável $A \subset U_k,\ k \geq 1$. Isto prova que $\xi$ é um jacobiano de $f$ relativamente a $\eta$.

Agora, provemos a unicidade. Suponha que $\xi$ e $\zeta$ são jacobianos de $f$ relativamente a $\eta$ e que existe $B \subset M$ com $\eta(B) > 0$ tal que $\xi(x) \neq \zeta(x)$ para todo $x \in B$. A menos de substituir $B$ por um subconjunto adequado, e permutar os papéis de $\xi$ e $\zeta$ se necessário, podemos supor que $\xi(x) < \zeta(x)$ para todo $x \in B$. De modo similar, podemos supor que $B$ está contido em algum $U_k$. Então,

$$\eta(f(B)) = \int_B \xi d\eta < \int_B \zeta d\eta = \eta(f(B)).$$

Esta contradição prova que o jacobiano é essencialmente único. □

**Observação 3.33.** *Usaremos a notação $J_\eta f$ para representar o (essencialmente único) jacobiano de $f$ com relação a $\eta$, quando exista.*

O próximo Teorema exprime a entropia de uma medida invariante explicitamente em termos do jacobiano.

**Teorema 3.34 (Fórmula de Rokhlin).** *Seja $f : M \to M$ uma transformação localmente invertível e seja $\mu$ uma probabilidade invariante por $f$. Suponha que existe alguma partição finita ou enumerável $\mathscr{P}$ tal que $\cup_n \mathscr{P}^n$ gera a $\sigma$-álgebra de $M$ e todo $P \in \mathscr{P}$ é domínio de invertibilidade de $f$. Então*

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

*Demonstração.* Consideremos a sequência de partições $\mathscr{Q}_n = \bigvee_{j=1}^n f^{-j}(\mathscr{P})$. Como $\mathscr{P}$ é partição geradora, podemos usar o Corolário 3.13 e o fato que $h_\mu(f, \mathscr{P}) = \lim_n H_\mu(\mathscr{P} \setminus \bigvee_{j=1}^n f^{-j}(\mathscr{P}))$ para qualquer partição $\mathscr{P}$ com entropia finita para concluir que

$$h_\mu(f) = h_\mu(f, \mathscr{P}) = \lim_n H_\mu(\mathscr{P} \setminus \mathscr{Q}_n) \tag{3.22}$$

Por definição,

$$
\begin{aligned}
H_\mu(\mathscr{P} \setminus \mathscr{Q}_n) &= \sum_{P \in \mathscr{P}} \sum_{Q_n \in \mathscr{Q}_n} -\mu(P \cap Q_n) \log \frac{\mu(P \cap Q_n)}{\mu(Q_n)} \\
&= \sum_{P \in \mathscr{P}} \sum_{Q_n \in \mathscr{Q}_n} \mu(Q_n) \left( -\frac{\mu(P \cap Q_n)}{\mu(Q_n)} \log \frac{\mu(P \cap Q_n)}{\mu(Q_n)} \right) \\
&= \sum_{P \in \mathscr{P}} \sum_{Q_n \in \mathscr{Q}_n} \mu(Q_n) \phi \left( \frac{\mu(P \cap Q_n)}{\mu(Q_n)} \right). \qquad (3.23)
\end{aligned}
$$

Como anteriormente, $\phi(x) = -x \log x$. Seja $e_n(\psi, x)$ a esperança condicional de uma função $\psi$ relativamente à partição $\mathscr{Q}_n$ e seja $e(\psi, x)$ o seu limite quando $n$ vai para infinito (estas noções foram introduzidas na definição 2.9 e no Lema 2.10). É claro da definição (2.3) que

$$
e_n(\chi_P, x) = \frac{\mu(P \cap Q_n)}{\mu(Q_n)} \text{ para todo } x \in Q_n \text{ e todo } Q_n \in \mathscr{Q}_n.
$$

Portanto,

$$
\sum_{P \in \mathscr{P}} \sum_{Q_n \in \mathscr{Q}_n} \mu(Q_n) \phi \left( \frac{\mu(P \cap Q_n)}{\mu(Q_n)} \right) = \sum_{P \in \mathscr{P}} \int \phi(e_n(\chi_P, x)) d\mu(x). \qquad (3.24)
$$

Pelo Lema 2.10, o limite $e(\chi_P, x) = \lim_n e_n(\chi_P, x)$ existe para $\mu$-quase todo $x$. Então, observando que a função $\phi \circ e_n(\chi_P, x)$ é limitada, podemos usar o teorema da convergência dominada para deduzir das relações (3.22), (3.23) e (3.24) que

$$
\begin{aligned}
h_\mu(f) &= \lim_n \sum_{P \in \mathscr{P}} \int \phi(e_n(\chi_P, x)) d\mu(x) \\
&= \sum_{P \in \mathscr{P}} \int \lim_n \phi(e_n(\chi_P, x)) d\mu(x) \\
&= \sum_{P \in \mathscr{P}} \int \phi(e(\chi_P, x)) d\mu(x). \qquad (3.25)
\end{aligned}
$$

Resta relacionar o integrando do lado direito com o jacobiano. Isso será feito por meio dos seguintes lemas:

**Lema 3.35 (Fórmulas de mudança de variáveis).** *Seja $f : M \to M$ uma transformação localmente invertível e $\eta$ uma probabilidade boreliana em $M$ não singular com relação a $f$. Valem as seguintes fórmulas de mudança de variáveis:*

*(a)* $\int_{f(A)} \varphi d\eta = \int_A (\varphi \circ f) J_\eta f d\eta$ *para todo domínio de invertibilidade* $A \subset M$ *e toda função mensurável* $\varphi : f(A) \to \mathbb{R}$ *tal que as integrais estão definidas (podendo ser* $\pm\infty$*).*

*(b)* $\int_A \psi d\eta = \int_{f(A)} \left(\frac{\psi}{J_\eta f}\right) \circ (f|_A)^{-1} d\eta$ *para qualquer função mensurável* $\psi : A \to \mathbb{R}$ *tal que as integrais estão definidas (podendo ser* $\pm\infty$*).*

*Demonstração. (a)* Pela proposição 3.32 existe algum jacobiano de $f$ com relação a $\eta$ essencialmente único. Ele satisfaz

$$\eta(f(A)) = \int_A J_\eta f d\eta \tag{3.26}$$

conforme a equação (3.20). Note que se $\varphi = \chi_B$ para algum conjunto $B$ mensurável então, usando a equação (3.26)

$$\begin{aligned} \int_{f(A)} \chi_B d\eta &= \eta(B \cap f(A)) = \eta(f(A \cap f^{-1}(B))) \\ &= \int_{A \cap f^{-1}(B)} J_\eta f d\eta = \int_A \chi_{f^{-1}(B)} J_\eta f d\eta \\ &= \int_A (\chi_B \circ f) J_\eta f d\eta. \end{aligned} \tag{3.27}$$

Logo a afirmação do item *(a)* vale para $\varphi = \chi_B$. Pela linearidade da integral, o resultado vale para funções simples. Agora, suponha que $\varphi : f(A) \to \mathbb{R}$ é mensurável e não negativa; existe uma sequência de funções simples $\varphi_n : f(A) \to \mathbb{R}$ tal que $0 \leq \varphi_n \leq \varphi_{n+1}$ para todo $n$ e $\varphi = \lim_n \varphi_n$ para todo ponto $x$. Além disso, a sequência $(\varphi_n \circ f) J_\eta f$ também é monótona e $\lim_n (\varphi_n \circ f) J_\eta f = (\varphi \circ f) J_\eta f$. Pelo Teorema da Convergência monótona:

$$\int_{f(A)} \varphi d\eta = \lim_n \int_{f(A)} \varphi_n d\eta = \lim_n \int_A (\varphi_n \circ f) J_\eta f d\eta = \int_A (\varphi \circ f) J_\eta f d\eta.$$

Finalmente, para $\varphi : f(A) \to \mathbb{R}$ mensurável arbitrária, escreva $\varphi = \varphi^+ - \varphi^-$ onde $\varphi^+$, $\varphi^-$ são não negativas, use a linearidade da integral e o resultado anterior.

*(b)* Dada uma função mensurável $\psi : A \to \mathbb{R}$ qualquer, note que $\varphi = \left(\frac{\psi}{J_\eta f}\right) \circ (f|_A)^{-1} : f(A) \to \mathbb{R}$ é mensurável. Aplicando a esta o item anterior, obtemos

$$\int_{f(A)} \left(\frac{\psi}{J_\eta f}\right) \circ (f|_A)^{-1} d\eta = \int_A \left(\left(\left(\frac{\psi}{J_\eta f}\right) \circ (f|_A)^{-1}\right) \circ f\right) J_\eta f d\eta = \int_A \psi d\eta$$

como queríamos mostrar. $\square$

**Lema 3.36.** *Para toda função mensurável limitada $\psi : M \to \mathbb{R}$ e toda probabilidade boreliana $\eta$ invariante por $f$,*

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

*Demonstração.* Lembre que $\mathscr{Q}_n = \bigvee_{j=1}^{n} f^{-j}(\mathscr{P})$. Também usaremos a sequência de partições $\mathscr{P}^n = \bigvee_{j=0}^{n-1} f^{-j}(\mathscr{P})$. Observe que $\mathscr{Q}_n(x) = f^{-1}(\mathscr{P}^{n-1}(f(x)))$ e $\mathscr{P}^n(x) = \mathscr{P}(x) \cap \mathscr{Q}_n(x)$ para todo $n$ e todo $x$. Então,

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

Usando a fórmula de mudança de variáveis dada pelo Lema 3.35 item (b), podemos reescrever a expressão do lado direito como

$$\sum_{P \in \mathscr{P}} \int_{P \cap \mathscr{Q}_n(x)} \psi(z) d\eta)(z) = \int_{\mathscr{Q}_n(x)} \psi d\eta.$$

Portanto,

$$\int_{\mathscr{P}^{n-1}(f(x))} \hat{\psi} d\eta = \int_{\mathscr{Q}_n(x)} \psi d\eta. \tag{3.28}$$

A hipótese de que $\eta$ é invariante dá que $\eta(\mathscr{P}^{n-1}(f(x))) = \eta(\mathscr{Q}_n(x))$. Dividindo ambos os lados de (3.28) por este número, obtemos que

$$e_n(\psi, x) = e'_{n-1}(\hat{\psi}, f(x)) \ \ \textit{para todo } x \textit{ e todo } n > 1. \tag{3.29}$$

Então passando ao limite, $e(\psi, x) = e'(\hat{\psi}, f(x))$ para $\eta$-quase todo $x$. Por outro lado, de acordo com o Teorema 2.11, a hipótese implica que $e'(\hat{\psi}, y) = \hat{\psi}(y)$ para $\eta$-quase todo $y \in M$. $\square$

Vamos aplicar o Lema 3.36 a $\psi = \chi_P$ e $\eta = \mu$. Como $f$ é injetiva em todo elemento de $\mathscr{P}$, cada interseção $P \cap f^{-1}(y)$ ou é vazia ou contém exatamente um ponto. Portanto, segue do Lema 3.36 que $e(\chi_P, x) = \hat{\chi_P}(f(x))$, com

$$\hat{\chi_P}(y) = \begin{cases} \dfrac{1}{J_\mu f((f|_P)^{-1}(y))} & \text{se } y \in f(P) \\ 0 & \text{se } y \notin f(P) \end{cases}$$

Então, lembrando que a medida $\mu$ é invariante,

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

(a última igualdade usa o Lema 3.35 item (b)). Substituindo esta expressão em (3.25), vem que

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

tal como afirmado no Teorema. □

**Lema 3.37.** *Seja $f : M \to M$ uma transformação localmente invertível e seja $\eta$ uma probabilidade boreliana em $M$ não singular com relação a $f$. Então para toda função mensurável limitada $\psi : M \to \mathbb{R}$,*

$$\int \psi d\eta = \int \sum_{z\in f^{-1}(x)} \frac{\psi}{J_\eta f}(z)d\eta(x).$$

*Demonstração.* Seja $\mathscr{P}$ uma partição enumerável em domínios de invertibilidade. Usando a fórmula de mudança de variáveis dada pelo Lema 3.35 item (b) e lembrando que a pré-imagem $f^{-1}(x)$ de qualquer $x \in M$ é enumerável, segue que

$$\begin{aligned}\int \psi d\eta = \sum_{P\in\mathscr{P}} \int_P \psi d\eta &= \sum_{P\in\mathscr{P}} \int_{f(P)} \frac{\psi}{J_\eta f} \circ (f|_P)^{-1}(x)d\eta(x) \\ &= \sum_{P\in\mathscr{P}} \int_{f(P)} \sum_{z\in f^{-1}(x)} \frac{\psi}{J_\eta f}(z)d\eta(x) \\ &= \int_{f(M)} \sum_{z\in f^{-1}(x)} \frac{\psi}{J_\eta f}(z)d\eta(x) \\ &= \int \sum_{z\in f^{-1}(x)} \frac{\psi}{J_\eta f}(z)d\eta(x).\end{aligned}$$

□

**Lema 3.38.** *Seja $f : M \to M$ uma transformação localmente invertível e seja $\eta$ uma probabilidade boreliana em $M$ não singular com relação a $f$. Então $\eta$ é invariante por $f$ se, e somente se,*

$$\sum_{z\in f^{-1}(x)} \frac{1}{J_\eta f(z)} = 1 \text{ para } \eta\text{-quase todo } x \in M.$$

*Demonstração.* Usando a igualdade dada pelo Lema anterior considerando ao invés de $\psi$ a função $\psi \circ f$,

$$\int (\psi \circ f) d\eta = \int \sum_{z \in f^{-1}(x)} \frac{(\psi \circ f)}{J_\eta f}(z) d\eta(x).$$

Mas $(\psi \circ f)(z) = \psi(x)$ uma vez que $z \in f^{-1}(x)$. Portanto

$$\int (\psi \circ f) d\eta = \int \sum_{z \in f^{-1}(x)} \frac{(\psi \circ f)}{J_\eta f}(z) d\eta(x) = \int \psi(x) \left( \sum_{z \in f^{-1}(x)} \frac{1}{J_\eta f}(z) \right) d\eta(x). \tag{3.30}$$

Então, se $\eta$ é invariante por $f$, o lado esquerdo de (3.30) é igual a $\int \psi d\eta$. Logo $\sum_{z \in f^{-1}(x)} \dfrac{1}{J_\eta f(z)} = 1$ para $\eta$-quase todo $x \in M$.

Por outro lado, se $\sum_{z \in f^{-1}(x)} \dfrac{1}{J_\eta f(z)} = 1$ para $\eta$-quase todo $x \in M.$, então o lado direito de (3.30) é igual a $\int \psi d\eta$ e isto mostra que $\eta$ é invariante por $f$. $\square$

**Lema 3.39.** *Seja $f : M \to M$ uma transformação localmente invertível e seja $\eta$ uma probabilidade boreliana em $M$ não singular com relação a $f$. Então para todo $k \geq 1$, existe jacobiano de $f^k$ com relação a $\eta$ e ele é dado por*

$$J_\eta f^k(x) = \prod_{j=0}^{k-1} J_\eta f(f^j(x)) \ \textit{para } \eta\textit{-quase todo } x \in M. \tag{3.31}$$

*Demonstração.* Se $f$ é localmente invertível, então $f^k$ também é: Da hipótese que $f$ é localmente invertível existe alguma cobertura enumerável $\{U_k : k \geq 1\}$ de $M$ por conjuntos mensuráveis tais que a restrição de $f$ a cada $U_k$ é uma bijeção sobre a sua imagem, a qual é um conjunto mensurável, e a inversa dessa bijeção também é mensurável. Daí note que $\{f^{-k}(U_k) : k \geq 1\}$ é cobertura enumerável de $M$ por conjuntos mensuráveis e que a restrição de $f^k$ a cada $f^{-k}(U_k)$ é uma bijeção sobre a sua imagem. De acordo com a proposição 3.32 existe algum jacobiano de $f^k$ com relação a $\eta$ essencialmente único.

Então se $A$ é um domínio de invertibilidade de $f^k$, usando a fórmula de mudança de variáveis dada pelo Lema 3.35 item (a), obtemos

$$\int_A J_\eta f^k d\eta = \eta(f^k(A)) = \eta(f^{k-1}(f(A))) = \int_{f(A)} J_\eta f^{k-1} d\eta = \int_A (J_\eta f^{k-1} \circ f) J_\eta f d\eta. \tag{3.32}$$

Portanto $J_\eta f^k = (J_\eta f^{k-1} \circ f) J_\eta f$ para todo $k$ e para $\eta$-quase todo $x \in M$. A fórmula (3.31) é obtida por indução em $k$: para $k = 1$ o resultado é imediato. Supondo que (3.31) vale para $k$ arbitrário, então pelo que provamos em (3.32) e pela hipótese de indução

$$J_\eta f^{k+1}(x) = (J_\eta f^k \circ f)(x) J_\eta f(x) = \prod_{j=0}^{k-1} J_\eta f(f^{j+1}(x)) J_\eta f(x)$$

$$= \prod_{j=1}^{k} J_\eta f(f^j(x)) J_\eta f(x) = \prod_{j=0}^{k} J_\eta f(f^j(x))$$

como queríamos mostrar. □

## 3.6 Entropia Topológica

**Definição 3.40** (**R. Adler, A. Konheim e M. McAndrew** ). Seja $M$ um espaço topológico compacto. Chamamos cobertura aberta de $M$ qualquer família $\Upsilon$ de abertos cuja união é todo o $M$. Por compacidade, toda cobertura aberta admite uma subcobertura finita (isto é, uma subfamília que ainda é uma cobertura) com um número finito de elementos. Chamamos **entropia da cobertura** $\Upsilon$ ao número

$$H(\Upsilon) = \log N(\Upsilon) \tag{3.33}$$

onde $N(\Upsilon)$ é o menor número tal que $\Upsilon$ admite alguma subcobertura finita com esse número de elementos. Chamaremos essa subcobertura de cardinalidade $N(\Upsilon)$ de **subcobertura ótima** de $\Upsilon$ e a denotaremos por $\Upsilon_o$.

Dadas duas coberturas abertas $\Upsilon$ e $\Lambda$, dizemos que $\Upsilon$ é **menos fina** que $\Lambda$, e escrevemos $\Upsilon \prec \Lambda$, se todo elemento de $\Lambda$ está contido em algum elemento de $\Upsilon$. Se $\Upsilon \prec \Lambda$, então $N(\Upsilon) \leq N(\Lambda)$, porque se $\Lambda_o$ é a subcobertura ótima de $\Lambda$, então para cada $U \in \Lambda_o$ existe $V_U \in \Upsilon$ tal que $U \subset V_U$. Mas note que $\Upsilon_{\Lambda_o} = \{V_U : U \in \Lambda_o\}$ forma uma subcobertura finita de $M$ por elementos de $\Upsilon$ e sua cardinalidade não é maior que a de $\Lambda_o$. Em particular, a cardinalidade de $\Upsilon_o$ não é maior que a cardinalidade de $\Upsilon_{\Lambda_o}$. Portanto, $N(\Upsilon) \leq N(\Lambda)$. Isto em especial mostra que

$$\Upsilon \prec \Lambda \quad \Rightarrow \quad H(\Upsilon) \leq H(\Lambda). \tag{3.34}$$

Dadas coberturas $\Upsilon_1, \ldots, \Upsilon_n$, denotamos por $\Upsilon_1 \bigvee \cdots \bigvee \Upsilon_n$ a sua **soma**, isto é, a cobertura cujos elementos são as interseções $U_1 \cap \cdots \cap U_n$ com $U_j \in \Upsilon_j$ para cada $j$. Note que $\Upsilon_j \prec \Upsilon_1 \bigvee \cdots \bigvee \Upsilon_n$ para todo $j$.

Seja $f : M \to M$ uma transformação contínua. Se $\Upsilon$ é uma cobertura aberta de $M$ então $f^{-j}(\Upsilon) = \{f^{-j}(A) : A \in \Upsilon\}$ também é uma cobertura aberta. Para cada $n \geq 1$, denotamos

$$\Upsilon^n = \Upsilon \bigvee f^{-1}(\Upsilon) \bigvee \cdots \bigvee f^{-n+1}(\Upsilon).$$

Observemos dois fatos importantes. Primeiro, note que $N(\Upsilon \bigvee \Lambda) \leq N(\Upsilon)N(\Lambda)$. De fato, sejam $\Upsilon_o, \Lambda_o$ as subcoberturas ótimas de $\Upsilon$ e $\Lambda$, respectivamente. Então, $\Upsilon_o \bigvee \Lambda_o$ é subcobertura de $\Upsilon \bigvee \Lambda$ (podendo não ser a subcobertura ótima $\Upsilon \bigvee \Lambda$) com cardinalidade não maior que o produto das cardinalidades de $\Upsilon_o$ e $\Lambda_o$. Logo a cardinalidade de $(\Upsilon \bigvee \Lambda)_o$ não é maior que o produto das cardinalidades de $\Upsilon_o$ e $\Lambda_o$. Consequentemente, $N(\Upsilon \bigvee \Lambda) \leq N(\Upsilon)N(\Lambda)$. Também, $N(f^{-1}(\Upsilon)) \leq N(\Upsilon)$. De fato, basta observar que se $\Upsilon_o$ é a subcobertura ótima de $\Upsilon$, então $f^{-1}(\Upsilon_o)$ é subcobertura de $f^{-1}(\Upsilon)$ (podendo não ser a ótima). Logo a cardinalidade de $(f^{-1}(\Upsilon))_o$ não é maior que a cardinalidade de $\Upsilon_o$. Resulta do que foi dito acima que a sequência $H(\Upsilon^n)$ é subaditiva:

$$\begin{aligned} H(\Upsilon^{m+n}) &= H(\Upsilon^m \bigvee f^{-m}(\Upsilon^n)) \\ &= \log N(\Upsilon^m \bigvee f^{-m}(\Upsilon^n)) \\ &\leq \log N(\Upsilon^m)N(f^{-m}(\Upsilon^n)) \\ &= \log N(\Upsilon^m) + \log N(f^{-m}(\Upsilon^n)) \\ &= H(\Upsilon^m) + H(\Upsilon^n) \end{aligned}$$

para todo $m, n \geq 1$. Em virtude do Lema 3.5,

$$h(f, \Upsilon) = \lim_n \frac{1}{n} H(\Upsilon^n) = \inf_n \frac{1}{n} H(\Upsilon^n) \tag{3.35}$$

sempre existe. Ele é chamado **entropia de $f$ com respeito à cobertura** $\Upsilon$. A relação (3.34) implica que

$$\Upsilon \prec \Lambda \quad \Rightarrow \quad h(f, \Upsilon) \leq h(f, \Lambda). \tag{3.36}$$

Finalmente, definimos a **entropia topológica de $f$** como sendo

$$h(f) = \sup\{h(f, \Upsilon) : \ \Upsilon \ \text{é cobertura aberta de } \ M\}. \tag{3.37}$$

Em particular, se $\Lambda$ é subcobertura de $\Upsilon$ então $h(f,\Upsilon) \leq h(f,\Lambda)$. Portanto, a definição (3.37) não muda se restringirmos o supremo às coberturas abertas finitas. Assim definida, a entropia $h(f)$ é um número não negativo, podendo ser infinito.

**Definição 3.41** (**Efim Dinaburg e Rufus Bowen**). Seja $f : M \to M$ contínua num espaço métrico $M$, não necessariamente compacto, e seja $K \subset M$ um subconjunto compacto qualquer.

Dados $\varepsilon > 0$ e $n \in \mathbb{N}$, dizemos que um conjunto $E \subset M$ é um $(n,\varepsilon)$-**gerador** de $K$, se para todo $x \in K$ existe $a \in E$ tal que $d(f^i(x), f^i(a)) < \varepsilon$ para todo $i \in \{0,\ldots,n-1\}$. Em outras palavras,

$$K \subset \bigcup_{a\in E} B(a,n,\varepsilon),$$

onde $B(a,n,\varepsilon) = \{x \in M : d(f^i(x), f^i(a)) < \varepsilon \ \text{para} \ i = 0,\ldots,n-1\}$ é chamada de **bola dinâmica** de centro $a$, comprimento $n$ e raio $\varepsilon$. Note que $\{B(x,n,\varepsilon) : x \in K\}$ é uma cobertura aberta de $K$. Logo, por compacidade, sempre existem conjuntos $(n,\varepsilon)$-geradores finitos. Denotamos $g_n(f,\varepsilon,K) = \min\{\#E : E \ \text{é} \ (n,\varepsilon)\text{-gerador de} \ K\}$. Definimos

$$g(f,\varepsilon,K) = \limsup_n \frac{1}{n} \log g_n(f,\varepsilon,K). \tag{3.38}$$

Observe que a função $\varepsilon \mapsto g(f,\varepsilon,K)$ é monótona não crescente. De fato, é claro da definição que se $\varepsilon_1 < \varepsilon_2$ então todo conjunto $(n,\varepsilon_1)$-gerador também é $(n,\varepsilon_2)$-gerador. Portanto, $g_n(f,\varepsilon_1,K) \geq g_n(f,\varepsilon_2,K)$ para todo $n \geq 1$ e, passando ao limite, $g(f,\varepsilon_1,K) \geq g(f,\varepsilon_2,K)$. Isto garante, em particular, que

$$g(f,K) = \lim_{\varepsilon\to 0} g(f,\varepsilon,K) \tag{3.39}$$

existe. Finalmente, definimos

$$g(f) = \sup\{g(f,K) : K \subset M \ \text{compacto}\}. \tag{3.40}$$

Também introduzimos a seguinte noção dual. Dados $\varepsilon > 0$ e $n \in \mathbb{N}$, dizemos que um conjunto $E \subset K$ é $(n,\varepsilon)$-**separado**, se dados $x, y \in E$, existe $j \in \{0,\ldots,n-1\}$ tal que $d(f^j(x), f^j(y)) \geq \varepsilon$. Em outras palavras, se $x \in E$ então $B(x,n,\varepsilon)$ não contém nenhum outro ponto de $E$. Denotamos $s_n(f,\varepsilon,K) = \max\{\#E : E \ \text{é} \ (n,\varepsilon)\text{-separado}\}$.

Definimos

$$s(f, \varepsilon, K) = \limsup_n \frac{1}{n} \log s_n(f, \varepsilon, K). \tag{3.41}$$

Observe que a função $\varepsilon \mapsto s(f, \varepsilon, K)$ é monótona não crescente. De fato, é claro da definição que se $0 < \varepsilon_1 < \varepsilon_2$ então todo conjunto $(n, \varepsilon_2)$-separado também é $(n, \varepsilon_1)$-separado. Portanto, $s_n(f, \varepsilon_1, K) \geq s_n(f, \varepsilon_2, K)$ para todo $n \geq 1$ e, passando ao limite, $s(f, \varepsilon_1, K) \geq s(f, \varepsilon_2, K)$. Isto garante, em particular, que

$$s(f, K) = \lim_{\varepsilon \to 0} s(f, \varepsilon, K) \tag{3.42}$$

existe. Finalmente, definimos

$$s(f) = \sup\{s(f, K) : K \subset M \text{ compacto}\}. \tag{3.43}$$

É claro que $g(f, K_1) \leq g(f, K_2)$ e $s(f, K_1) \leq s(f, K_2)$ se $K_1 \subset K_2$. Em particular

$$g(f) = g(f, M) \quad e \quad s(f) = s(f, M) \;\text{ se } M \text{ é compacto}. \tag{3.44}$$

**Proposição 3.42.** *Tem-se $g(f, K) = s(f, K)$ para todo compacto $K \subset M$. Consequentemente, $g(f) = s(f)$.*

**Definição 3.43.** Definimos a **entropia topológica de uma transformação contínua** $f : M \to M$ **num espaço métrico** $M$ como sendo $g(f) = s(f)$.

**Definição 3.44.** Quando $M$ é um espaço métrico, chamamos **diâmetro de uma cobertura aberta** $\Upsilon$ e denotamos por $\operatorname{diam}(\Upsilon)$ ao supremo dos diâmetros dos seus elementos. Também, para cada cobertura aberta $\Upsilon$, um **número de Lebesgue** para $\Upsilon$, é um número real $\varepsilon > 0$, tal que cada bola de raio $\varepsilon$ está contida em algum elemento de $\Upsilon$. Denotaremos o supremo desses números $\varepsilon$ por $\operatorname{Leb}(\Upsilon)$.

**Proposição 3.45.** *Suponha que $M$ é um espaço métrico compacto. Seja $(\Upsilon_k)_k$ qualquer sequência de coberturas abertas de $M$ tal que* $\operatorname{diam} \Upsilon_k$ *converge para zero. Então,*

$$h(f) = \sup_k h(f, \Upsilon_k) = \lim_k h(f, \Upsilon_k).$$

*Demonstração.* Dada qualquer cobertura aberta $\Lambda$, seja $\varepsilon > 0$ um número de Lebesgue de $\Lambda$. Tome $n \geq 1$ tal que $\operatorname{diam} \Upsilon_k < \varepsilon$ para todo $k \geq n$. Pela definição de número de Lebesgue, segue que todo elemento de $\Upsilon_k$ está contido em algum elemento de $\Lambda$. Em outras palavras, $\Lambda \prec \Upsilon_k$ e, portanto, $h(f, \Upsilon_k) \geq h(f, \Lambda)$. Lembrando a definição (3.37), isto prova que

$$\liminf_k h(f, \Upsilon_k) \geq h(f).$$

Também é claro das definições que $h(f) \geq \sup_k h(f, \Upsilon_k) \geq \limsup_k h(f, \Upsilon_k)$. Combinando estas duas observações obtemos a conclusão da proposição. $\square$

**Teorema 3.46** (**Bowen, 1971**). *Se $M$ é espaço métrico compacto, então $h(f) = g(f) = s(f)$.*

*Demonstração.* Um conjunto $E$ é $(n, \varepsilon)$-gerador se, e somente se, a família $B(x, n, \varepsilon)$, $x \in E$, é uma subcobertura da cobertura $\Upsilon_{(n,\varepsilon)}$ de todas as $B(x, n, \varepsilon)$ as quais, por sua vez, é subcobertura de $\Upsilon^n_{(1,\varepsilon)} = \Upsilon_{(1,\varepsilon)} \bigvee f^{-1}(\Upsilon_{(1,\varepsilon)}) \bigvee \cdots \bigvee f^{-n+1}(\Upsilon_{(1,\varepsilon)})$. Por definição de número de Lebesgue, $\Upsilon \prec \Upsilon_{(1,\varepsilon)}$ sempre que $\varepsilon \leq \mathrm{Leb}(\Upsilon)$. Usando que: $\Upsilon \prec \Lambda \;\Rightarrow\; N(\Upsilon) \leq N(\Lambda)$ e $\Upsilon^n \prec \Lambda^n$, resulta que

$$g_n(f, \varepsilon, M) = N(\Upsilon_{(n,\varepsilon)}) \geq N(\Upsilon^n_{(1,\varepsilon)}) \geq N(\Upsilon^n). \tag{3.45}$$

Note que qualquer conjunto $(n, \varepsilon)$-separado $E$ com cardinalidade maximal deve ser $(n, \varepsilon)$-gerador; do contrário, existiria um ponto $x_0 \in M$, tal que $\mathrm{dist}(x_0, x) \geq \varepsilon$ para todo $x \in E$. Mas assim $E \cup \{x_0\}$ seria um conjunto $(n, \varepsilon)$-separado, contradizendo a maximalidade de $E$. Isto implica que

$$s_n(f, \varepsilon, M) = \#E \geq g_n(f, \varepsilon, M). \tag{3.46}$$

Agora, tome uma cobertura $\Lambda$ com $\mathrm{diam}(\Lambda) < \varepsilon$. Seja $E$ um conjunto $(n, \varepsilon)$-separado. Então cada elemento de $\Lambda^n$ contém no máximo um elemento de $E$. Por outro lado, qualquer subcobertura de $\Lambda^n$ cobre todos os elementos de $E$. Portanto

$$N(\Lambda^n) \geq s_n(f, \varepsilon, M). \tag{3.47}$$

Combinando as fórmulas (3.45), (3.46) e (3.47), vem que

$$N(\Lambda^n) \geq s_n(f, \varepsilon, M) \geq g_n(f, \varepsilon, M) \geq N(\Upsilon^n)$$

e portanto,

$$\limsup_n \frac{1}{n}\log N(\Lambda^n) \geq \limsup_n \frac{1}{n}\log s_n(f,\varepsilon,M)$$
$$\geq \limsup_n \frac{1}{n}\log g_n(f,\varepsilon,M) \geq \limsup_n \frac{1}{n}N(\Upsilon^n).$$

A prova é completada passando o limite sobre uma sequência de refinamentos $\Upsilon_k$, pondo $\varepsilon_k = \mathrm{Leb}(\Upsilon_k)$ e escolhendo uma sequência de coberturas $\Lambda_k$ com $\mathrm{diam}(\Lambda_k) \leq \varepsilon_k$. □

## 3.7 Exemplos: Cálculo da Entropia Topológica

**Exemplo 3.47 (Transformações expansivas).** Se $f$ é expansiva com constante de expansividade $\varepsilon_0 > 0$, então

(a) $h(f) = h(f,\Upsilon)$ para toda cobertura aberta $\Upsilon$ com diâmetro menor que $\varepsilon_0$;

(b) $\limsup_n \frac{1}{n}\log \#\,\mathrm{Fix}(f^n) \leq h(f)$, onde $\mathrm{Fix}(f^n)$ denota o conjunto dos pontos $x \in M$ tais que $f^n(x) = x$. *Isto prova em especial que, para transformações expansivas a entropia topológica é uma cota superior para a taxa de crescimento do número de pontos periódicos.*

Seja $\Upsilon$ qualquer cobertura aberta de $M$ com $\mathrm{diam}\,\Upsilon < \varepsilon_0$. Afirmamos que $\lim_k \mathrm{diam}\,\Upsilon^k = 0$. De fato, suponha que isso não é verdade. É claro que a sequência dos diâmetros é não crescente. Então existe $\delta > 0$ e para cada $k \geq 1$ existem pontos $x_k$ e $y_k$ num mesmo elemento de $\Upsilon^k$ tais que $d(x_k, y_k) \geq \delta$. Por compacidade, podemos escolher uma subsequência $(k_j)_j$ tal que existem $x = \lim_j x_{k_j}$ e $y = \lim_j y_{k_j}$ . Observe que $x \neq y$, de fato $d(x,y) \geq \delta$. Por outro lado, a escolha de $x_k$ e $y_k$ num mesmo elemento de $\Upsilon^k$ implica que

$$d(f^i(x_k), f^i(y_k)) \leq \mathrm{diam}\,\Upsilon \text{ para todo } 0 \leq i \leq k.$$

Passando ao limite, vem que $d(f^i(x), f^i(y)) \leq \mathrm{diam}\,\Upsilon < \varepsilon_0$ para todo $i \geq 0$. Isto contradiz a hipótese de que $\varepsilon_0$ é uma constante de expansividade para $f$. Esta contradição prova a nossa afirmação. Usando a Proposição 3.45, segue que $h(f) = h(f,\Upsilon)$, tal como afirma o item (a).

Para provar o item (b), seja $\Upsilon$ uma cobertura de $M$ com $\operatorname{diam}\Upsilon < \varepsilon_0$, onde $\varepsilon_0$ é uma constante de expansividade de $f$. Afirmamos que cada elemento de $\Upsilon^n$ contém no máximo um ponto de $\operatorname{Fix}(f^n)$. De fato, se $x, y \in \operatorname{Fix}(f^n)$ estão no mesmo elemento de $\Upsilon^n$, então $d(f^i(x), f^i(y)) < \operatorname{diam}\Upsilon < \varepsilon_0$ para todo $i = 0, \dots, n-1$. Como $f^n(x) = x$ e $f^n(y) = y$, segue que $d(f^i(x), f^i(y)) < \varepsilon_0$ para todo $i \geq 0$. Por expansividade, isso implica que $x = y$, o que prova a nossa afirmação. Segue que

$$\limsup_n \frac{1}{n} \log \# \operatorname{Fix}(f^n) \leq \limsup_n \frac{1}{n} \log N(\Upsilon^n) = h(f, \Upsilon).$$

Agora, basta tomar o limite quando o diâmetro de $\Upsilon$ vai para zero.

**Exemplo 3.48** (**Deslocamentos de tipo finito**). Seja $X = \{1, \dots, d\}$ um conjunto finito e seja $A = (A_{i,j})_{i,j}$ uma **matriz de transição**: matriz quadrada de dimensão $d$ cujos coeficientes tomam apenas os valores 0 ou 1 e tal que nenhuma linha é identicamente nula. Considere o subconjunto $\Sigma_A^+$ de $\Sigma = X^{\mathbb{N}}$ das sequências $(x_n)_n$ que são $A$**-admissíveis**, ou seja, tais que

$$A_{x_n, x_{n+1}} = 1 \text{ para todo } n \in \mathbb{N}. \tag{3.48}$$

A restrição do deslocamento $\sigma_A^+ : \Sigma_A^+ \to \Sigma_A^+$ ao espaço métrico compacto invariante $\Sigma_A^+$ é chamado de **deslocamento unilateral de tipo finito** associado a $A$. O **deslocamento bilateral de tipo finito** associado a uma matriz de transição $A$ é definido de maneira análoga, considerando $\Sigma = X^{\mathbb{Z}}$ e exigindo (3.48) para todo $n \in \mathbb{Z}$. Neste caso também exigimos, como parte da definição de matriz de transição, que as colunas (não apenas as linhas) de $A$ sejam não nulas. Calcularemos a entropia topológica destas transformações. Para o enunciado precisamos de alguns comentários prévios sobre grafos de matrizes de transição. Geralmente associamos a uma matriz de transição $A$ o grafo orientado

$$G_A = \{(a, b) \in X \times X : A_{a,b} = 1\}.$$

Por exemplo, se quisermos esboçar o grafo da matriz de transição

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix},$$

pomos como vértices do grafo os pontos de $X = \{1, 2, 3, 4\}$, e escrevemos uma seta ligando os vértices $a$ e $b$ sempre que $A_{a,b} = 1$ para todo $a, b \in X$. A figura 1 descreve o grafo associado a matriz $A$.

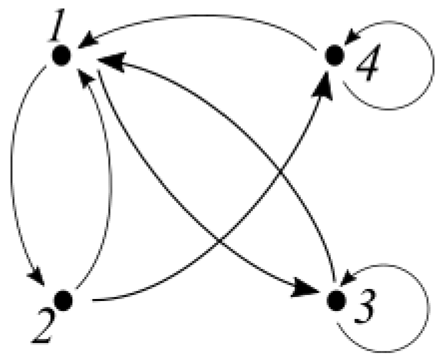


Figura 1: Grafo associado a matriz $A$

Chamamos **caminho de comprimento** $l \geq 1$ no grafo $G_A$ a qualquer sequência $a_0, \ldots, a_l$ em $X$ tal que $A_{a_i-1,a_i} = 1$ para todo $i$, isto é, tal que sempre existe uma aresta ligando $a_{i-1}$ a $a_i$ (por exemplo, no grafo da Figura 1 $1, 2, 4$ é um caminho).

Lembre que o raio espectral $\rho(B)$ de uma aplicação linear $B : \mathbb{R}^d \to \mathbb{R}^d$ (isto é, o máximo dos valores absolutos dos autovalores de $B$) é dado por

$$\rho(B) = \lim_n ||B^n||^{\frac{1}{n}} = \lim_n |\operatorname{trc} B^n|^{\frac{1}{n}}, \tag{3.49}$$

onde $||.||$ representa uma norma qualquer no espaço vetorial das aplicações lineares e trc designa o traço da matriz. Todas as normas são equivalentes, já que estamos em dimensão finita. Em geral, consideramos a norma de operador $||B|| = \sup\{\frac{||Bv||}{||v||} : v \neq 0\}$ mas também será útil considerar a norma definida por

$$||B||_s = \sum_{i,j=1}^{d} |B_{i,j}|.$$

Agora suponha que $A$ é uma matriz de transição. Como os coeficientes de $A$ são não negativos, podemos usar o teorema de Perron-Frobenius (Teorema 2.16), para concluir que $A$ admite um autovalor não negativo $\lambda_A$ que é igual ao raio espectral. Pela definição de matriz de transição, também temos que as linhas da matriz $A$ são não nulas. Então o mesmo vale para $A^n$, qualquer que seja $n \geq 1$. Isto implica que todos os coeficientes do vetor $A_n(1,\ldots,1)$ são (inteiros) positivos e, portanto,

$$||A_n|| \geq \frac{||A_n(1,\ldots,1)||}{||(1,\ldots,1)||} \geq 1 \text{ para todo } n \geq 1.$$

Usando (3.49), obtemos que $\lambda_A = \rho(A) \geq 1$ para toda matriz de transição $A$. Vamos provar agora que: A entropia topológica de um deslocamento de tipo finito $\sigma_A : \Sigma_A \to \Sigma_A$ é dada por $h(\sigma_A) = \log \lambda_A$ onde $\lambda_A$ é o maior autovalor da matriz de transição $A$.

Trataremos o caso de deslocamentos unilaterais. Consideremos a cobertura aberta $\Upsilon$ de $\Sigma_A^+$ formada pelas restrições

$$[0; a]_A = \{(x_j)_j \in \Sigma_A^+ : x_0 = a\}$$

dos cilindros $[0; a]$ de $M$. Para cada $n \geq 1$, a cobertura aberta $\Upsilon^n$ está formada pelas restrições

$$[0; a_0, ..., a_{n-1}]_A = \{(x_j)_j \in \Sigma_A^+ : x_j = a_j \text{ para } j = 0,\ldots,n-1\}.$$

dos cilindros de comprimento $n$. Observe que este conjunto é não vazio se, e somente se, $a_0,\ldots,a_{n-1}$ é um caminho (de comprimento $n-1$) no grafo $G_A$: se o conjunto é não vazio, então possui uma sequência $(a_0, ..., a_{n-1}, x_n, \ldots) \in \Sigma_A^+$, e isto significa que $A_{a_{i-1},a_i} = 1$ para $i = 1,\ldots,n-1$, portanto $a_0,\ldots,a_{n-1}$ é um caminho de comprimento $n-1$ no grafo $G_A$; e reciprocamente, se $A_{a_{i-1},a_i} = 1$ para $i = 1,\ldots,n-1$, podemos montar uma sequência em $[0; a_0, ..., a_{n-1}]_A$ pela própria definição de $\Sigma_A^+$. Como os cilindros são disjuntos dois-a-dois, esta observação mostra que $N(\Upsilon^n)$ é igual ao número total de caminhos de comprimento $n-1$ no grafo $G_A$, ou seja,

$$N(\Upsilon^n) = \sum_{i,j=1}^{d} A_{i,j}^{n-1} = ||A^{n-1}||_s.$$

Pela fórmula do raio espectral em (3.49), segue que

$$h(\sigma_A^+, \Upsilon) = \lim_n \frac{1}{n} \log N(\Upsilon^n) = \lim_n \frac{1}{n} \log ||A^{n-1}||_s = \log \rho(A) = \log \lambda_A.$$

Finalmente, como $\operatorname{diam} \Upsilon^n \to 0$, $h(\sigma_A^+) = h(\sigma_A^+, \Upsilon)$.

**Observação 3.49.** *Ao final do exemplo anterior usamos a seguinte observação: Se $M$ é um espaço métrico compacto e $\Upsilon$ é uma cobertura aberta tal que* $\operatorname{diam}(\Upsilon^n)$ *converge para zero quando $n \to \infty$, então $h(f) = h(f, \Upsilon)$.*

**Observação 3.50.** *Se $\sigma_A : \Sigma_A \to \Sigma_A$ é um deslocamento de tipo finito então*

$$\lim_n \frac{1}{n} \log \# \operatorname{Fix}({\sigma_A}^n) = h(\sigma_A).$$

## 3.8 Pressão Topológica

**Definição 3.51.** Seja $f : M \to M$ uma transformação contínua num espaço métrico compacto. Chamamos **potencial** em $M$ a qualquer função contínua $\phi : M \to \mathbb{R}$. Para cada $n \in \mathbb{N}$, definimos $\phi_n : M \to \mathbb{R}$ por $\phi_n = \sum_{i=0}^{n-1} \phi \circ f^i$. Além disso, dado qualquer conjunto não vazio $C \subset M$, denotamos

$$\phi_n(C) = \sup\{\phi_n(x) : x \in C\}. \tag{3.50}$$

Dada uma cobertura aberta $\Upsilon$ de $M$ definimos

$$P_n(f, \phi, \Upsilon) = \inf\{\sum_{U \in \gamma} e^{\phi_n(U)} : \gamma \text{ é subcobertura finita de } \Upsilon^n\}. \tag{3.51}$$

Note que

$$\sum_{U \in \gamma} e^{\phi_{m+n}(U)} = \sum_{U \in \gamma} e^{\phi_n(U) + \phi_m(f^n(U))} \leq \sum_{U \in \gamma} e^{\phi_n(U)} \sum_{V \in f^n(\gamma)} e^{\phi_m(V)}. \tag{3.52}$$

Portanto a sequência $\log P_n(f, \phi, \Upsilon)$ é subaditiva e, portanto, o limite

$$P(f, \phi, \Upsilon) = \lim_n \frac{1}{n} \log P_n(f, \phi, \Upsilon). \tag{3.53}$$

existe (Lema 3.5). Finalmente, chamamos **pressão do potencial** $\phi$ relativamente a $f$ ao limite $P(f, \phi)$ de $P(f, \phi, \Upsilon)$ quando o diâmetro de $\Upsilon$ vai para zero. A existência deste limite é garantida pelo seguinte lema:

**Lema 3.52.** *Existe* $\lim_{\operatorname{diam} \Upsilon \to 0} P(f, \phi, \Upsilon)$*, ou seja, existe $P(f, \phi) \in [0, \infty]$ tal que*

$$\lim_k P(f, \phi, \Upsilon_k) = P(f, \phi)$$

*para toda sequência $(\Upsilon_k)_k$ de coberturas abertas com* $\operatorname{diam} \Upsilon_k \to 0$.

*Demonstração.* Sejam $(\Upsilon_k)_k$ e $(\Lambda_k)_k$ sequências quaisquer de coberturas abertas com diâmetros convergindo para zero. Dado qualquer $\varepsilon > 0$ fixe $\delta > 0$ tal que $|\phi(x) - \phi(y)| \leq \varepsilon$ sempre que $d(x,y) \leq \delta$. Por hipótese, $\operatorname{diam} \Upsilon_k < \delta$ para todo $k$ suficientemente grande. Para $k$ fixado, seja $\alpha > 0$ um número de Lebesgue para $\Upsilon_k$. Por hipótese, $\operatorname{diam} \Lambda_m < \alpha$ para todo $m$ suficientemente grande. Pela definição de número de Lebesgue, todo $B \in \Lambda_m$ está contido em algum $A \in \Upsilon_k$. Observe que $\phi_n(A) \leq n\varepsilon + \phi_n(B)$ para todo $n \geq 1$, uma vez que $\operatorname{diam} \Upsilon_k < \delta$. Isto implica que

$$P_n(f, \phi, \Upsilon_k) \leq e^{n\varepsilon} P_n(f, \phi, \Lambda_m)$$

para todo $n \geq 1$ e, portanto, $P(f, \phi, \Upsilon_k) \leq \varepsilon + P(f, \phi, \Lambda_m)$. Fazendo $m \to \infty$ e depois $k \to \infty$, obtemos que

$$\limsup_k P(f, \phi, \Upsilon_k) \leq \varepsilon + \liminf_m P(f, \phi, \Lambda_m).$$

Como $\varepsilon > 0$ é arbitrário, segue que $\limsup_k P(f, \phi, \Upsilon_k) \leq \liminf_m P(f, \phi, \Lambda_m)$. Permutando os papéis das duas sequências de coberturas, concluímos que os limites $\lim_k P(f, \phi, \Upsilon_k)$ e $\lim_m P(f, \phi, \Lambda_m)$ existem e são iguais. □

**Observação 3.53.** *São consequências imediatas da definição de pressão:*

1. *A pressão do potencial nulo coincide com a entropia topológica. De fato, dada uma cobertura aberta $\Upsilon$, note que*

$$P_n(f, 0, \Upsilon) = \inf\{\sum_{U \in \gamma} 1 : \gamma \text{ é subcobertura finita de } \Upsilon^n\} = N(\Upsilon^n)$$

*para todo $n \geq 1$. Portanto, $h(f, \Upsilon) = P(f, 0, \Upsilon)$. Finalmente, fazendo o diâmetro de cobertura $\Upsilon$ tender a zero, obtemos $h(f) = P(f, 0)$.*

2. *A pressão é um invariante de equivalência topológica: Sejam $f : M \to M$ e $g : N \to N$ transformações contínuas em espaços métricos compactos. Se existe um homeomorfismo $h : M \to N$ tal que $h \circ f = g \circ h$ então $P(g, \phi) = P(f, \phi \circ h)$ para todo potencial $\phi$ em $N$. Para provar este fato, começamos notando que $g^i = h \circ f^i \circ h^{-1}$ e portanto, $\phi_n = \sum_{i-0}^{n-1} \phi \circ g^i = \sum_{i-0}^{n-1} (\phi \circ h) \circ f^i \circ h^{-1}$. Então, se $\Upsilon$ é cobertura*

*aberta de* $N$,

$$\begin{aligned} P_n(g,\phi,\Upsilon) &= \inf\{\sum_{U\in\gamma} e^{\phi_n(U)} : \gamma \text{ é subcobertura finita de } \Upsilon^n\} \\ &= \inf\{\sum_{V\in h^{-1}(\gamma)} e^{(\phi\circ h)_n(V)} : h^{-1}(\gamma) \text{ é subcobertura finita de } h^{-1}(\Upsilon^n)\}. \end{aligned}$$

*Observamos também que*

$$h^{-1}(\Upsilon^n) = \bigvee_{i=0}^{n-1} h^{-1}\circ g^{-i}(\Upsilon) = \bigvee_{i=0}^{n-1} f^{-i}\circ h^{-1}(\Upsilon) = (h^{-1}(\Upsilon))^n.$$

*Daí,*

$$P_n(g,\phi,\Upsilon) = P_n(f,\phi\circ h, h^{-1}(\Upsilon)).$$

*implica que*

$$P(g,\phi,\Upsilon) = P(f,\phi\circ h, h^{-1}(\Upsilon)).$$

*Fazendo o diâmetro de* $\Upsilon$ *tender a zero, o diâmetro de* $h^{-1}(\Upsilon)$ *também tende a zero, assim:*

$$P(g,\phi) = P(f,\phi\circ h).$$

3. *Sejam* $f, g$ *como no item acima. Se* $h : M \to N$ *é contínua, sobrejetiva (este caso não exige que* $h$ *seja injetiva) e satisfaz* $h\circ f = g\circ h$ *então* $P(g,\phi) \le P(f,\phi\circ h)$ *para todo potencial* $\phi$ *em* $N$*. A prova é semelhante a do item acima. Tome* $\Upsilon$ *é cobertura aberta de* $N$ *e conclua que*

$$P_n(g,\phi,\Upsilon) \le P_n(f,\phi\circ h, h^{-1}(\Upsilon)).$$

*A desigualdade acontece porque a subcobertura finita de* $h^{-1}(\Upsilon^n)$ *de* $M$ *pode ter mais abertos do que a de* $\Upsilon$ *devido à sobrejetividade de* $h$*. Concluímos então que* $P(g,\phi,\Upsilon) \le P(f,\phi\circ h, h^{-1}(\Upsilon))$*. Fazendo o diâmetro de* $h^{-1}(\Upsilon)$ *tender a zero, o diâmetro de* $\Upsilon$ *também tende a zero e assim* $P(g,\phi) \le P(f,\phi\circ h)$.

4. *Se* $h(f)$ *é finita então* $P(f,\phi) < \infty$ *para todo potencial* $\phi$ *e caso contrário* $P(f,\phi) = \infty$ *para todo potencial* $\phi$*. A prova dessa afirmação decorre dos seguintes fatos: dada qualquer constante* $c \in \mathbb{R}$*, temos* $P_n(f,\phi+c,\Upsilon) = e^{nc}P_n(f,\phi,\Upsilon)$ *para*

*todo $n \geq 1$ e, portanto, $P(f, \phi + c, \Upsilon) = P(f, \phi, \Upsilon) + c$ para toda cobertura aberta $\Upsilon$. Logo,*

$$P(f, \phi + c) = P(f, \phi) + c. \tag{3.54}$$

*Também, é claro que*

$$\phi \leq \psi \Rightarrow P(f, \phi) \leq P(f, \psi). \tag{3.55}$$

*Em particular, como $\inf \phi \leq \phi \leq \sup \phi$, segue de (3.54) e (3.55) que*

$$h(f) + \inf \phi \leq P(f, \inf \phi) \leq P(f, \phi) \leq P(f, \sup \phi) \leq h(f) + \sup \phi$$

*para todo potencial $\phi$. Caso $h(f) = \infty$, temos $P(f, \phi) = \infty$ como afirmado.*

**Definição 3.54** (**Definição de Pressão em termos de conjuntos geradores e separados**). Seja $f : M \to M$ contínua num espaço métrico compacto e $\phi : M \to \mathbb{R}$ uma função contínua.

Dados $n \geq 1$ e $\varepsilon > 0$, defina

$$\begin{aligned} G_n(f, \phi, \varepsilon) &= \inf\{\sum_{x \in E} e^{\phi_n(x)} : E \text{ é conjunto } (n, \varepsilon)\text{-gerador de } M\} \\ S_n(f, \phi, \varepsilon) &= \sup\{\sum_{x \in E} e^{\phi_n(x)} : E \text{ é conjunto } (n, \varepsilon)\text{-separado de } M\}. \end{aligned} \tag{3.56}$$

Em seguida, defina

$$\begin{aligned} G(f, \phi, \varepsilon) &= \limsup_n \frac{1}{n} \log G_n(f, \phi, \varepsilon) \\ S(f, \phi, \varepsilon) &= \limsup_n \frac{1}{n} \log S_n(f, \phi, \varepsilon) \end{aligned} \tag{3.57}$$

e também

$$\begin{aligned} G(f, \phi) &= \lim_{\varepsilon \to 0} G(f, \phi, \varepsilon) \\ S(f, \phi) &= \lim_{\varepsilon \to 0} S(f, \phi, \varepsilon) \end{aligned} \tag{3.58}$$

(os limites existem, porque as funções são monótonas em $\varepsilon$).

Note que $G_n(f, 0, \varepsilon) = g_n(f, \varepsilon)$ e $S_n(f, 0, \varepsilon) = s_n(f, \varepsilon)$ para todo $n \geq 1$ e todo $\varepsilon > 0$. Portanto, pelo Teorema 3.46, $G(f, 0) = g(f)$ e $S(f, 0) = s(f)$ são iguais a entropia topológica $h(f)$, uma vez que $M$ é espaço métrico compacto. Em geral, temos:

**Teorema 3.55.** *$P(f,\phi) = G(f,\phi) = S(f,\phi)$ para todo potencial $\phi$ em $M$.*

*Demonstração.* Considere $n \geq 1$ e $\varepsilon > 0$. Vimos na prova do Teorema 3.46 que todo conjunto $(n,\varepsilon)$-separado maximal é $(n,\varepsilon)$-gerador. Então,

$$\begin{aligned} S_n(f,\phi,\varepsilon) &= \sup\{\sum_{x\in E} e^{\phi_n(x)} : E \text{ é conjunto } (n,\varepsilon)\text{-separado de } M\} \\ &= \sup\{\sum_{x\in E} e^{\phi_n(x)} : E \text{ é conjunto } (n,\varepsilon)\text{-separado maximal de } M\} \\ &\geq \inf\{\sum_{x\in E} e^{\phi_n(x)} : E \text{ é conjunto } (n,\varepsilon)\text{-gerador de } M\} \ = \ G_n(f,\phi,\varepsilon) \end{aligned} \tag{3.59}$$

Isto implica que $G(f,\phi,\varepsilon) \leq S(f,\phi,\varepsilon)$. Passando ao limite quando $\varepsilon \to 0$, obtemos $G(f,\phi) \leq S(f,\phi)$.

Em seguida provamos que $S(f,\phi) \leq P(f,\phi)$. Sejam $\varepsilon$ e $\delta$ números positivos tais que $d(x,y) \leq \delta$ implica $|\phi(x) - \phi(y)| \leq \varepsilon$. Seja $\Upsilon$ qualquer cobertura aberta de $M$ com $\operatorname{diam}\Upsilon < \delta$. Seja $E \subset M$ qualquer conjunto $(n,\delta)$-separado. Dada qualquer subcobertura $\gamma$ de $\Upsilon^n$, é claro que todo ponto de $E$ está contido em algum elemento de $\gamma$. Por outro lado, a hipótese de que $E$ é $(n,\delta)$-separado implica que cada elemento de $\gamma$ contém no máximo um elemento de $E$. Portanto,

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

Tomando o supremo em $E$ e o ínfimo em $\gamma$, obtemos que

$$S_n(f,\phi,\delta) \leq P_n(f,\phi,\Upsilon). \tag{3.60}$$

Segue que $S(f,\phi,\delta) \leq P(f,\phi,\Upsilon)$. Fazendo $\delta \to 0$ (logo $\operatorname{diam}\Upsilon \to 0$), concluímos que $S(f,\phi) \leq P(f,\phi)$, conforme afirmado.

Finalmente, provamos que $P(f,\phi) \leq G(f,\phi)$. Sejam $\varepsilon$ e $\delta$ números positivos tais que $d(x,y) \leq \delta$ implica $|\phi(x) - \phi(y)| \leq \varepsilon$. Seja $\Upsilon$ qualquer cobertura aberta de $M$ com $\operatorname{diam}\Upsilon < \delta$. Seja $\rho > 0$ um número de Lebesgue de $\Upsilon$ e seja $E \subset M$ um conjunto $(n,\rho)$-gerador qualquer de $M$. Para cada $x \in E$ e $i = 0,\dots,n-1$, existe $A_{x,i} \in \Upsilon$ tal que $B(f^i(x),\rho)$ está contida em $A_{x,i}$. Denotamos,

$$\gamma(x) = \bigcap_{i=0}^{n-1} f^{-i}(A_{x,i}).$$

Note que $\gamma(x) \in \Upsilon^n$ e que $B(x,n,\rho) \subset \gamma(x)$. Logo, a hipótese de que $E$ é $(n,\rho)$-gerador implica que $\gamma = \{\gamma(x) : x \in E\}$ é uma subcobertura de $\Upsilon$. Note também que

$$\phi_n(\gamma(x)) \leq n\varepsilon + \phi_n(x) \text{ para todo } x \in E,$$

uma vez que $\operatorname{diam} A_{x,i} < \delta$ para todo $i$. Segue que

$$\sum_{U \in \gamma} e^{\phi_n(U)} \leq e^{n\varepsilon} \sum_{x \in E} e^{\phi_n(x)}.$$

Isto prova que $P_n(f,\phi,\Upsilon) \leq e^{n\varepsilon} G_n(f,\phi,\rho)$ para todo $n \geq 1$ e, por consequência,

$$P(f,\phi,\Upsilon) \leq \varepsilon + \liminf_n \frac{1}{n} G_n(f,\phi,\rho) \leq \varepsilon + G(f,\phi,\rho).$$

Fazendo $\rho \to 0$ vem que $P(f,\phi,\Upsilon) \leq \varepsilon + G(f,\phi)$. Então, fazendo $\varepsilon, \delta$ e $\operatorname{diam} \Upsilon$ ir para zero, $P(f,\phi) \leq G(f,\phi)$. □

**Teorema 3.56.** *Seja $f : M \to M$ uma transformação contínua num espaço métrico compacto e seja $\phi$ um potencial em $M$. Então:*

$$P(f^k, \phi_k) = kP(f,\phi) \text{ para todo } k \geq 1.$$

*Demonstração.* Lembre que

$$S_n(f,\phi,\varepsilon) = \sup\{\sum_{x \in E} e^{\phi_n(x)} : E \text{ é conjunto } (n,\varepsilon)\text{-separado de } M\}.$$

Se $E$ for um conjunto $(n,\varepsilon)$-separado para $f^k$, então, dados quaisquer $x,y \in E$, existe $j \in \{0,\ldots,n\}$ tal que $d(f^j(x), f^j(y)) \geq \varepsilon$. Isto significa que para algum $i \in \{0,\ldots,nk\}$, $d(f^i(x), f^i(y)) \geq \varepsilon$. Portanto $E$ é $(nk,\varepsilon)$-separado para $f$. Consequentemente,

$$S_n(f^k, \phi_k, \varepsilon) \leq S_{nk}(f,\phi,\varepsilon) \leq P_{nk}(f,\phi,\Upsilon).$$

A última desigualdade vem da equação (3.60) do Teorema 3.55 e $\Upsilon$ é uma cobertura aberta de $M$ com $\operatorname{diam} \Upsilon < \varepsilon$. Usando a subaditividade de $P_{nk}(f,\phi,\Upsilon)$, vem

$$S(f^k,\phi_k,\varepsilon) = \limsup_n \frac{1}{n} \log S_n(f^k,\phi_k,\varepsilon) \leq k \lim_n \frac{1}{n} \log P_n(f,\phi,\Upsilon) \leq kP(f,\phi,\Upsilon).$$

Se fizermos $\varepsilon \to 0$, $\operatorname{diam} \Upsilon \to 0$. Portanto

$$S(f^k, \phi_k) \leq kP(f, \phi).$$

Usando o Teorema 3.55

$$P(f^k, \phi_k) \leq kP(f, \phi).$$

Por outro lado, dado $\varepsilon > 0$, escolha $\delta > 0$ de modo que $d(x, y) < \delta$ implique

$$\max_{1 \leq i \leq k-1} d(f^i(x), f^i(y)) < \varepsilon.$$

Se $E$ é $(n, \delta)$-gerador de $M$ para $f^k$, então $E$ é $(nk, \varepsilon)$-gerador de $M$ para $f$. De fato, dado $x \in M$, existe $a \in E$ tal que $d(f^{ki}(x), f^{ki}(a)) < \delta$ para todo $i \in \{0, \ldots, n\}$. Mas por hipótese $d(f^{ki}(x), f^{ki}(a)) < \delta$ para todo $i \in \{0, \ldots, n\}$ implica que $d(f^i(x), f^i(a)) < \varepsilon$ para todo $i \in \{0, \ldots, nk\}$, ou seja, $E$ é $(nk, \varepsilon)$-gerador de $M$ para $f$. Consequentemente,

$$G_n(f^k, \phi_k, \delta) \geq G_{nk}(f, \phi, \varepsilon).$$

Portanto,

$$G(f^k, \phi_k, \delta) \geq kG(f, \phi, \varepsilon).$$

Fazendo $\varepsilon, \delta \to 0$,

$$G(f^k, \phi_k) \geq kG(f, \phi).$$

Usando o Teorema 3.55

$$P(f^k, \phi_k) \geq kP(f, \phi).$$

$\square$

## 3.9 Princípio Variacional

**Teorema 3.57** (**Princípio Variacional**)**.** *Seja $f : M \to M$ uma transformação contínua num espaço métrico compacto e seja $\mathcal{M}_1(f)$ o conjunto das medidas de probabilidade invariantes por $f$. Então, para toda função contínua $\phi : M \to \mathbb{R}$,*

$$P(f, \phi) = \sup \left\{ h_\nu(f) + \int \phi d\nu : \nu \in \mathcal{M}_1(f) \right\}.$$

*Demonstração.* Inicialmente, provamos que para toda probabilidade invariante $\nu$, vale

$$h_\nu(f) + \int \phi d\nu \leq P(f,\phi). \tag{3.61}$$

Para isto, o seguinte Lema será útil:

**Lema 3.58.** *Sejam $a_1,\ldots,a_k$ números reais e sejam $p_1,\ldots,p_k$ números não negativos tais que $p_1+\ldots+p_k=1$. Seja $A=\sum_{i=1}^k e^{a_i}$. Então,*

$$\sum_{i=1}^k p_i(a_i - \log p_i) \leq \log A. \tag{3.62}$$

*Além disso, vale a igualdade se, e somente se, $p_j = \dfrac{e^{a_j}}{A}$ para todo $j$.*

*Demonstração.* Escreva $t_i = \dfrac{e^{a_i}}{A}$ e $x_i = \dfrac{p_i}{e^{a_i}}$. Note que $\sum_{i=1}^k t_i = 1$. Como $\phi(x) = -x\log x$ é côncava,

$$\sum_{i=1}^k t_i\phi(x_i) \leq \phi\left(\sum_{i=1}^k t_i x_i\right).$$

Note que $t_i\phi(x_i) = -\dfrac{e^{a_i}}{A}\dfrac{p_i}{e^{a_i}}\log\dfrac{p_i}{e^{a_i}} = \dfrac{p_i}{A}(a_i - \log p_i)$ e que $\sum_{i=1}^k t_i x_i = \sum_{i=1}^k \dfrac{p_i}{A} = 1$. Portanto, a desigualdade anterior pode ser reescrita na seguinte forma:

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

Multiplicando por $A$ obtemos a desigualdade do enunciado do lema. Além disso, vale a igualdade se, e somente se, os $x_i$ são todos iguais, ou seja se existe $c$ tal que $p_i = ce^{a_i}$ para todo $i$. Somando sobre $i=1,\ldots,k$ vemos que nesse caso $c = \dfrac{1}{A}$, conforme afirmado no enunciado. $\square$

Seja $\mathscr{P} = \{P_1,\ldots,P_s\}$ uma partição finita qualquer. Como toda probabilidade invariante num espaço métrico é regular (veja a Proposição A.3.2 de [7]), segue que $\nu$ é regular. Portando, dado $\varepsilon > 0$ podemos encontrar compactos $Q_i \subset P_i$ tais que $\nu(P_i \setminus Q_i) < \varepsilon$ para todo $i=1,\ldots,s$. Seja $Q_0$ o complementar de $\bigcup_{i=1}^s Q_i$ e seja $P_0 = \emptyset$. Então $\mathscr{Q} = \{Q_0, Q_1, \ldots, Q_s\}$ é uma partição finita de $M$ satisfazendo $\nu(P_i \Delta Q_i) \leq s\varepsilon$ para todo $i = 0,1,\cdots,s$. Logo, pelo Lema 3.10,

$$H_\nu(\mathscr{P} \setminus \mathscr{Q}) \leq \log 2$$

desde que $\varepsilon > 0$ seja suficientemente pequeno (dependendo apenas de $s$). Fixe $\varepsilon$ e $\mathscr{Q}$ a partir daqui e suponha que $\Upsilon$ é uma cobertura aberta de $M$ satisfazendo

$$\operatorname{diam}\Upsilon < \min\{d(Q_i, Q_j) : 1 \leq i < j \leq s\}. \tag{3.63}$$

Pelo Lema 3.12, temos que $h_\nu(f, \mathscr{P}) \leq h_\nu(f, \mathscr{Q}) + H_\nu(\mathscr{P} \setminus \mathscr{Q}) \leq h_\nu(f, \mathscr{Q}) + \log 2$. Portanto

$$h_\nu(f, \mathscr{P}) + \int \phi d\nu \leq h_\nu(f, \mathscr{Q}) + \int \phi d\nu + \log 2. \tag{3.64}$$

Provaremos agora que

$$h_\nu(f, \mathscr{Q}) + \int \phi d\nu \leq \log 2 + P(f, \phi, \Upsilon). \tag{3.65}$$

Observe que

$$\begin{aligned} H_\nu(\mathscr{Q}^n) + \int \phi_n d\nu &\leq \sum_{Q \in \mathscr{Q}_n} -\nu(Q) \log \nu(Q) + \sum_{Q \in \mathscr{Q}_n} \nu(Q) \sup\{\phi_n(x) : x \in Q\} \\ &= \sum_{Q \in \mathscr{Q}_n} \nu(Q)(-\log \nu(Q) + \phi_n(Q)) \end{aligned}$$

para todo $n \geq 1$. Usando o Lema 3.58 para $A = \sum_{Q \in \mathscr{Q}_n} e^{\phi_n(Q)}$ e $\sum_{Q \in \mathscr{Q}_n} \nu(Q) = 1$, obtemos

$$H_\nu(\mathscr{Q}^n) + \int \phi_n d\nu \leq \sum_{Q \in \mathscr{Q}_n} \nu(Q)(-\log \nu(Q) + \phi_n(Q)) \leq \log \left( \sum_{Q \in \mathscr{Q}_n} e^{\phi_n(Q)} \right). \tag{3.66}$$

Seja $\gamma$ uma subcobertura finita qualquer de $\Upsilon^n$. Para cada $Q \in \mathscr{Q}^n$ considere qualquer ponto $x_Q$ no fecho de $Q$ tal que $\phi_n(x_Q) = \phi_n(Q)$ (lembre que $\phi_n(Q)$ denota o supremo de $\phi_n$ no conjunto $Q$). Considere $U_Q \in \gamma$ tal que $x_Q \in U_Q$. Então, $\phi_n(Q) \leq \phi_n(U_Q)$ para todo $Q \in Q_n$. A condição (3.63) implica que cada elemento de $\Upsilon$ intersecta o fecho de não mais que dois elementos de $\mathscr{Q}$. Portanto, cada elemento de $\Upsilon^n$ intersecta o fecho de, no máximo, $2^n$ elementos de $\mathscr{Q}^n$. Em particular, para cada $U \in \gamma$ existem não mais que $2^n$ elementos $Q$ de $\mathscr{Q}^n$ tais que $U_Q = U$. Portanto:

$$\sum_{Q \in \mathscr{Q}_n} e^{\phi_n(Q)} \leq 2^n \sum_{U \in \gamma} e^{\phi_n(U)} \tag{3.67}$$

para qualquer subcobertura finita $\gamma$ de $\Upsilon^n$. Substituindo (3.67) em (3.66), obtemos

$$H_\nu(\mathscr{Q}^n) + \int \phi d\nu \leq n \log 2 + \log P_n(f, \phi, \Upsilon).$$

Dividindo por $n$ e passando ao limite quando $n \to \infty$, obtemos a equação (3.65) exatamente como queríamos provar. Agora, substituindo (3.65) em (3.64), obtemos

$$h_\nu(f, \mathscr{P}) + \int \phi d\nu \leq 2\log 2 + P(f, \phi, \Upsilon). \tag{3.68}$$

Fazendo $\operatorname{diam} \Upsilon \to 0$, segue que $h_\nu(f, \mathscr{P}) + \int \phi d\nu \leq 2\log 2 + P(f, \phi)$ para toda partição finita $\mathscr{P}$. Logo, $h_\nu(f) + \int \phi d\nu \leq 2\log 2 + P(f, \phi)$. Agora substitua $f$ por $f^k$ e o potencial $\phi$ por $\phi_k$. Note que $\int \phi_k d\nu = \int \sum_{i=0}^{k-1} \phi \circ f^i d\nu = k \int \phi d\nu$, uma vez que $\nu$ é invariante por $f$. Usando também o Teorema 3.56 e o fato que $h_\nu(f^k) = kh_\nu(f)$ para todo $k \in \mathbb{N}$, vem que

$$kh_\nu(f) + k\int \phi d\nu \leq 2\log 2 + kP(f, \phi)$$

para todo $k \geq 1$. Dividindo por $k$ e passando ao limite quando $k \to \infty$ obtemos, a desigualdade (3.61).

Mostraremos agora que para todo $\varepsilon > 0$ existe uma probabilidade $\mu$ invariante por $f$ e tal que

$$h_\mu(f) + \int \phi d\mu \geq S(f, \phi, \varepsilon). \tag{3.69}$$

Claramente, isto implica que o supremo dos valores de $h_\nu(f) + \int \phi d\nu$ quando $\nu$ varia em $\mathcal{M}_1(f)$ é maior ou igual que $S(f, \phi) = P(f, \phi)$. Para cada $n \geq 1$, seja $E$ um conjunto $(n, \varepsilon)$-separado tal que

$$\sum_{y \in E} e^{\phi_n(y)} \geq \frac{1}{2} S_n(f, \phi, \varepsilon). \tag{3.70}$$

Denotaremos por $A$ a expressão no lado esquerdo desta desigualdade. Considere as medidas de probabilidade $\nu_n$ e $\mu_n$ definidas em $M$ por:

$$\nu_n = \frac{1}{A} \sum_{x \in E} e^{\phi_n(x)} \delta_x \ \text{ e } \ \mu_n = \frac{1}{n} \sum_{j=0}^{n-1} f_*^j \nu_n.$$

Como o espaço das probabilidades é compacto, e lembrando a definição (3.57), podemos escolher uma subsequência $(n_j)_j \to \infty$ tal que:

1. $\dfrac{1}{n_j} \log S_{n_j}(f, \phi, \varepsilon)$ converge para $S(f, \phi, \varepsilon)$ e

2. $\mu_{n_j}$ converge para alguma probabilidade $\mu$ na topologia fraca$^*$.

Vamos mostrar que tal probabilidade $\mu$ é invariante por $f$ e satisfaz (3.69). Dividimos o argumento em quatro passos.

**Passo 1:** Provamos que $\mu$ é invariante. Seja $\varphi : M \to \mathbb{R}$ uma função contínua qualquer. Para cada $n \geq 1$,

$$\int \varphi d(f_*\mu_n) = \frac{1}{n}\sum_{j=1}^{n}\int \varphi \circ f^j d\nu_n = \int \varphi d\mu_n + \frac{1}{n}\left(\int \varphi \circ f^n d\nu_n - \int \varphi d\nu_n\right)$$

e, por consequência,

$$\left|\int \varphi d(f_*\mu_n) - \int \varphi d\mu_n\right| \leq \frac{1}{n}\sup|\varphi|.$$

Restringindo a $n = n_j$ e passando ao limite quando $j \to \infty$, vemos que $\int \varphi d(f_*\mu_n) = \int \varphi d\mu_n$ para toda função contínua $\varphi : M \to \mathbb{R}$. Mas sabemos que probabilidades cujas integrais coincidem para toda função contínua num espaço métrico compacto são iguais. Portanto, $f_*\mu = \mu$ tal como foi afirmado.

**Passo 2:** Estimamos a entropia relativamente a $\nu_n$. Seja $\mathscr{P}$ qualquer partição finita de $M$ tal que $\operatorname{diam}\mathscr{P} < \varepsilon$ e $\mu(\partial\mathscr{P}) = 0$. A primeira condição implica que cada elemento de $\mathscr{P}^n$ contém no máximo um elemento de $E$. Por outro lado, todo elemento de $E$ está contido em algum elemento de $\mathscr{P}^n$. Logo,

$$\begin{aligned} H_{\nu_n}(\mathscr{P}^n) &= \sum_{x\in E} -\nu_n(\{x\})\log\nu_n(\{x\}) = \sum_{x\in E} -\frac{1}{A}e^{\phi_n(x)}\log(\frac{1}{A}e^{\phi_n(x)}) \\ &= \log A - \frac{1}{A}\sum_{x\in E} e^{\phi_n(x)}\phi_n(x) = \log A - \int \phi_n d\nu_n. \end{aligned} \tag{3.71}$$

A última igualdade decorre diretamente da definição de $\nu_n$.

**Passo 3:** Calculamos a entropia relativamente a $\mu_n$. Considere $1 \leq k < n$. Para cada $r \in \{0, \ldots, k-1\}$, seja $q_r \geq 0$ o maior número inteiro tal que $r + kq_r \leq n$. Em outras palavras, $q_r = \left\lfloor \dfrac{(n-r)}{k} \right\rfloor$. Então,

$$\mathscr{P}^n = \mathscr{P}^r \bigvee (\bigvee_{j=0}^{q_r-1} f^{-(kj+r)}(\mathscr{P}^k)) \bigvee f^{-(kq_r+r)}(\mathscr{P}^{n-(kq_r+r)})$$

(o primeiro termo não existe se $r = 0$ e o terceiro não existe se $n = kq_r + r$). Portanto,

$$H_{\nu_n}(\mathscr{P}^n) \leq \sum_{j=0}^{q_r-1} H_{\nu_n}(f^{-(kj+r)}(\mathscr{P}^k)) + H_{\nu_n}(\mathscr{P}^r) + H_{\nu_n}(f^{-(kq_r+r)}(\mathscr{P}^{n-(kq_r+r)})).$$

Claro que $\#\mathscr{P}^r \leq (\#\mathscr{P})^k$. Usando que: *para toda partição finita $\mathscr{Q}$, $H_\mu(\mathscr{Q}) \leq \log\#\mathscr{Q}$*, segue que $H_\mu(\mathscr{P}^r) \leq \log\#\mathscr{P}^r \leq \log(\#\mathscr{P})^k = k\log\#\mathscr{P}$. Pela mesma razão, o último termo na desigualdade anterior também é limitado por $k\log\#\mathscr{P}$. Então, usando também a observação 3.7 item (3),

$$H_{\nu_n}(\mathscr{P}^n) \leq \sum_{j=0}^{q_r-1} H_{f_*^{(kj+r)}\nu_n}(\mathscr{P}^k) + 2k\log\#\mathscr{P} \tag{3.72}$$

para todo $r \in \{0,\ldots,k-1\}$. Agora, é claro que todo número $i \in \{0,\ldots,n-1\}$ se escreve de maneira única na forma $i = kj + r$ com $0 \leq j \leq q_r - 1$. Então, somando (3.72) sobre todo os valores de $r$,

$$kH_{\nu_n}(\mathscr{P}^n) \leq \sum_{i=0}^{n-1} H_{f_*^i\nu_n}(\mathscr{P}^k) + 2k^2\log\#\mathscr{P} \tag{3.73}$$

A propriedade da concavidade da função $\phi(x) = -x\log x$ implica que

$$\frac{1}{n}\sum_{i=0}^{n-1} H_{f_*^i\nu_n}(\mathscr{P}^k) \leq H_{\mu_n}(\mathscr{P}^k).$$

Combinando esta última desigualdade com (3.73), vemos que

$$\frac{1}{n}H_{\nu_n}(\mathscr{P}^n) \leq \frac{1}{k}H_{\mu_n}(\mathscr{P}^k) + \frac{2k}{n}\log\#\mathscr{P}. \tag{3.74}$$

Por outro lado, pela definição de $\mu_n$,

$$\frac{1}{n}\int\phi_n d\nu_n = \frac{1}{n}\sum_{i=0}^{n-1}\int\phi\circ f^i d\nu_n = \int\phi d\mu_n. \tag{3.75}$$

Somando (3.74) e (3.75), obtemos

$$\frac{1}{n}H_{\nu_n}(\mathscr{P}^n) + \frac{1}{n}\int\phi_n d\nu_n \leq \frac{1}{k}H_{\mu_n}(\mathscr{P}^k) + \int\phi d\mu_n + \frac{2k}{n}\log\#\mathscr{P}. \tag{3.76}$$

**Passo 4:** Traduzimos as estimativas anteriores para a medida limite $\mu$. Combinando (3.71) e (3.76), vem que

$$\frac{1}{k}H_{\mu_n}(\mathscr{P}^k) + \int\phi d\mu_n \geq \frac{1}{n}\log A - \frac{2k}{n}\log\#\mathscr{P}.$$

Pela escolha de $E$ em (3.70), segue que

$$\frac{1}{k}H_{\mu_n}(\mathscr{P}^k) + \int\phi d\mu_n \geq \frac{1}{n}\log S_n(f,\phi,\varepsilon) - \frac{1}{n}\log 2 - \frac{2k}{n}\log\#\mathscr{P}. \tag{3.77}$$

A escolha da partição $\mathscr{P}$ implica que $\mu(\partial\mathscr{P}^k) = 0$ para todo $k \geq 1$, já que

$$\partial\mathscr{P}^k \subset \bigcup_{i=0}^{k-1} f^{-i}(\partial\mathscr{P}).$$

Então, $\mu(P) = \lim_j \mu_{n_j}(P)$ para todo $P \in \mathscr{P}^k$ e, portanto,

$$H_\mu(\mathscr{P}^k) = \lim_j H_{\mu_{n_j}}(\mathscr{P}^k).$$

Como a função $\phi$ é contínua, também temos $\int \phi d\mu = \lim_j \int \phi d\mu_{n_j}$ . Portanto, restringindo (3.77) à subsequência $(n_j)_j$ e passando ao limite quando $j \to \infty$,

$$\frac{1}{k} H_\mu(\mathscr{P}^k) + \int \phi d\mu \geq S(f, \phi, \varepsilon).$$

Passando o limite quando $k \to \infty$, vem que

$$h_\mu(f, \mathscr{P}) + \int \phi d\mu \geq S(f, \phi, \varepsilon).$$

Agora, fazendo $\varepsilon \to 0$ (e consequentemente, $\operatorname{diam} \mathscr{P} \to 0$), obtemos (3.69). Isto completa a demonstração do Princípio Variacional. □

Abaixo mencionamos algumas consequências diretas do Princípio Variacional.

**Corolário 3.59** (**Princípio Variacional para entropia Topológica**). *Se $f : M \to M$ é uma transformação contínua num espaço métrico compacto então a sua entropia topológica $h(f)$ coincide com o supremo das entropias $h_\mu(f)$ da transformação $f$ relativamente a todas as probabilidades invariantes.*

*Demonstração.* Basta usar o item (1) da observação a 3.53 e o Princípio Variacional (Teorema 3.57) para ver que

$$h(f) = P(f, 0) = \sup\left\{h_\nu(f) + \int 0 d\nu : \nu \in \mathcal{M}_1(f)\right\} = \sup\{h_\nu(f) : \nu \in \mathcal{M}_1(f)\}.$$

□

**Corolário 3.60.** *Seja $f : M \to M$ uma transformação contínua num espaço métrico compacto e seja $\mathcal{M}_e(f)$ o conjunto das medidas de probabilidade invariantes e ergódicas. Então,*

$$P(f, \phi) = \sup\left\{h_\nu(f) + \int \phi d\nu : \nu \in \mathcal{M}_e(f)\right\}.$$

*Em particular,*

$$h(f) = \sup\{h_\nu(f) : \nu \in \mathcal{M}_e(f)\}.$$

*Demonstração.* Dada qualquer $\nu \in \mathcal{M}_1(f)$, seja $\nu_P : P \in \mathscr{P}$ a sua decomposição ergódica. Pelo Teorema da decomposição ergódica (Teorema 2.8) e pelo Teorema de Jacobs (Teorema 3.27),

$$h_\nu(f) + \int \phi d\nu = \int \left( h_{\nu_P}(f) + \int \phi d\nu_P \right) d\hat{\mu}(P).$$

Isto implica que

$$\sup\left\{h_\nu(f) + \int \phi d\nu : \nu \in \mathcal{M}_1(f)\right\} \leq \sup\left\{h_\nu(f) + \int \phi d\nu : \nu \in \mathcal{M}_e(f)\right\}.$$

A desigualdade recíproca é trivial, uma vez que $\mathcal{M}_e(f) \subset \mathcal{M}_1(f)$. Agora basta aplicar o Teorema 3.57. □

Outra consequência interessante é que, para transformações com entropia topológica finita, a função pressão determina o conjunto das medidas de probabilidade invariantes:

**Corolário 3.61** (**Walters**). *Seja $f : M \to M$ uma transformação contínua num espaço métrico compacto, com entropia topológica $h(f) < \infty$. Seja $\eta$ uma medida com sinal finita em $M$. Então $\eta$ é uma probabilidade invariante por $f$ se, e somente se, $\int \phi d\eta \leq P(f, \phi)$ para toda função contínua $\phi : M \to \mathbb{R}$.*

*Demonstração.* Veja o Corolário 10.4.3 de [7]. □

## 3.10 Estados de Equilíbrio

**Definição 3.62.** Seja $f : M \to M$ uma transformação contínua num espaço métrico compacto. Uma medida de probabilidade invariante $\mu$ é um **estado de equilíbrio** para um potencial $\phi : M \to M$ se ela realiza o supremo no princípio variacional, ou seja, se

$$h_\mu(f) + \int \phi d\mu = P(f, \phi) = \sup\left\{h_\nu(f) + \int \phi d\nu : \nu \in \mathcal{M}_1(f)\right\}.$$

No caso particular em que $\phi \equiv 0$, também dizemos que $\mu$ é uma **medida de máxima entropia**. De agora em diante, denotaremos por $\mathcal{E}(f,\phi)$ o conjunto dos estados de equilíbrio.

O resultado abaixo mostra que o conjunto dos estados de equilíbrio é convexo e que seus elementos extremais são ergódicos.

**Teorema 3.63.** *Suponha que $h(f) < \infty$. Então o conjunto dos estados de equilíbrio para qualquer potencial $\phi : M \to \mathbb{R}$ é um subconjunto convexo de $\mathcal{M}_1(f)$: mais precisamente, dado $t \in (0,1)$ e dadas $\mu_1, \mu_2 \in \mathcal{M}_1(f)$,*

$$(1-t)\mu_1 + t\mu_2 \in \mathcal{E}(f,\phi) \Leftrightarrow \{\mu_1, \mu_2\} \subset \mathcal{E}(f,\phi).$$

*Além disso, uma probabilidade invariante $\mu$ está em $\mathcal{E}(f,\phi)$ se, e somente se, quase toda componente ergódica de $\mu$ está em $\mathcal{E}(f,\phi)$.*

*Demonstração.* A hipótese de que a entropia topológica é finita assegura que $P(f,\phi) < \infty$ para todo potencial $\phi$, como vimos na observação 3.53 item (3). Consideremos o funcional $\Psi : \mathcal{M}_1(f) \to \mathbb{R}$ definido por $\Psi(\mu) = h_\mu(f) + \int \phi d\mu$. Este funcional é linear e convexo, uma vez que $h_{(1-t)\mu_1+t\mu_2}(f) = (1-t)h_{\mu_1}(f)t + h_{\mu_2}(f)$ (veja o Teorema 8.1 de [8] ). Portanto

$$\Psi((1-t)\mu_1 + t\mu_2) = (1-t)\Psi(\mu_1) + t\Psi(\mu_2)$$

para todo $t \in (0,1)$ e quaisquer $\mu_1, \mu_2 \in \mathcal{M}_1(f)$. Então $\Psi((1-t)\mu_1 + t\mu_2)$ é igual ao supremo de $\Psi$ se, e somente se,$\Psi(\mu_1)$ e $\Psi(\mu_2)$ são iguais a esse supremo. Isto prova a primeira parte da proposição.

A prova da segunda parte é análoga: para cada probabilidade invariante $\mu$, seja $\{\mu_P : P \in \mathscr{P}\}$ a sua respectiva decomposição ergódica. Segue do Teorema de Jacobs (Teorema 3.27) que

$$\Psi(\mu) = \int \Psi(\mu_P) d\hat{\mu}(P). \tag{3.78}$$

A relação (3.78) dá que $\Psi(\mu) = \sup \Psi$ se, e somente se, $\Psi(\mu_P) = \sup \Psi$ para $\hat{\mu}$-quase todo $P$. □

**Corolário 3.64.** *Se $\mathcal{E}(f,\phi)$ é não vazio então ele contém probabilidades invariantes ergódicas. Além disso, os elementos extremais do convexo $\mathcal{E}(f,\phi)$ são precisamente as medidas ergódicas contidas nele.*

*Demonstração.* Para provar a primeira afirmação basta considerar as componentes ergódicas de qualquer elemento de $\mathcal{E}(f,\phi)$. Passemos a provar a segunda afirmação. Sabemos que $\mathcal{M}_1(f)$ é convexo e que as medidas ergódicas são os elementos extremais deste convexo. Portanto, se $\mu \in \mathcal{E}(f,\phi)$ é ergódica então $\mu$ é um elemento extremal de $\mathcal{M}_1(f)$. Com maior razão, $\mu$ é um elemento extremal de $\mathcal{E}(f,\phi)$. Reciprocamente, se

$$\mu \in \mathcal{E}(f,\phi)$$

não é ergódica então podemos escrever

$$\mu = (1-t)\mu_1 + t\mu_2, \ \text{ com } \ 0 < t < 1 \ \text{ e } \ \mu_1, \mu_2 \in \mathcal{M}_1(f).$$

Pelo Teorema 3.63 temos que $\mu_1, \mu_2 \in \mathcal{E}(f,\phi)$ e, portanto, $\mu$ não é elemento extremal do conjunto $\mathcal{E}(f,\phi)$. □

Um fato muito relevante é que o conjunto $\mathcal{E}(f,\phi)$ pode ser vazio, conforme o seguinte exemplo.

**Exemplo 3.65.** Seja $f_n : M_n \to M_n$ uma sequência de homeomorfismos em espaços métricos compactos tal que a sequência $(h(f_n))_n$ é crescente e limitada. A ideia é construir um espaço métrico $M$ e um homeomorfismo $f : M \to M$ com as seguintes características:

1. $M$ é a união de todos os $M_n$ com um ponto adicional, que representamos por $\infty$, com uma função distância tal que $(M_n)_n$ converge para $\infty$.

2. $f$ fixa o ponto $\infty$ e a sua restrição a cada $M_n$ coincide com $f_n$.

Veremos então que $f : M \to M$ não tem nenhuma medida de máxima entropia. Pelo Teorema 3.60, basta mostrar que nenhuma probabilidade invariante e ergódica realiza $h(f)$.

Passemos aos detalhes. Seja $d_n$ a distância em cada espaço métrico $M_n$. Não é restrição supor que $d_n \leq 1$ para todo $n$. Defina $M = \cup_n M_n \cup \{\infty\}$ e considere a distância $d$ definida em $M$ por:

$$d(x,y) = \begin{cases} n^{-2} d_n(x,y) & \text{se } \ x \in M_n \ \text{ e } \ y \in M_n \ \text{ com } \ n \geq 1 \\ \sum_{i=n}^{m} i^{-2} & \text{se } \ x \in M_n \ \text{ e } \ y \in M_m \ \text{ com } \ n \geq m \\ \sum_{i=n}^{\infty} i^{-2} & \text{se } \ x \in M_n \ \text{ e } \ y = \infty. \end{cases}$$

Seja $\beta = \sup_n h(f_n)$. Como os conjuntos $\{\infty\}$ e $M_n$, $n \geq 1$ são invariantes e cobrem todo o $M$, qualquer probabilidade ergódica $\mu$ de $f$ satisfaz $\mu(\{\infty\}) = 1$, ou então $\mu(M_n) = 1$ para algum $n \geq 1$. No primeiro caso, $h_\mu(f) = 0$. No segundo, $\mu$ pode ser vista como uma probabilidade invariante de $f_n$ e, portanto, $h_\mu(f) \leq h(f_n)$. Em particular, $h_\mu(f) \leq \beta$ para toda probabilidade $\mu$ invariante e ergódica para $f$. A observação anterior também mostra que

$$\begin{aligned}&\sup\{h_\mu(f) : \mu \text{ invariante e ergódica para } f\}\\ &\qquad = \sup_n \sup\{h_\mu(f) : \mu \text{ invariante e ergódica para } f_n\}.\end{aligned}$$

De acordo com o Teorema 3.60, isto significa que $h(f) = \sup_n h(f_n) = \beta$. Desta forma, fica provado que nenhuma medida invariante e ergódica de $f$ realiza a entropia topológica.

Contudo, existe uma classe ampla de transformações para as quais a existência de estados de equilíbrio está garantida para todo potencial:

**Lema 3.66.** *Se a função entropia de $f$ é semicontínua superiormente então $\mathcal{E}(f,\phi)$ é compacto (na topologia fraca$^*$) e não vazio, para qualquer potencial $\phi$.*

*Demonstração.* Seja $(\mu_n)_n$ uma sequência em $\mathcal{M}_1(f)$ tal que $h_{\mu_n}(f) + \int \phi d\mu_n$ converge para $P(f,\phi)$. Por compacidade de $\mathcal{M}_1(f)$, a sequência admite algum ponto de acumulação $\mu$. A semicontinuidade da entropia, juntamente com a continuidade da integral, implica que

$$h_\mu(f) + \int \phi d\mu \geq \liminf_n h_{\mu_n}(f) + \int \phi d\mu_n = P(f,\phi).$$

Portanto, $\mu$ é um estado de equilíbrio, tal como afirmado. Analogamente, tomando uma sequência qualquer $(\nu_n)_n$ em $\mathcal{E}(f,\phi)$ vemos que qualquer ponto de acumulação $\nu$ é um estado de equilíbrio. Isto mostra que $\mathcal{E}(f,\phi)$ é fechado e, consequentemente, compacto. □

Sabemos que a função entropia de toda transformação expansiva num espaço métrico compacto é semicontínua superiormente devido ao Teorema 3.23. Isto juntamente com o última Lema dá que:

**Corolário 3.67.** *Seja $f : M \to M$ uma transformação expansiva num espaço métrico compacto. Então todo potencial $\phi$ admite algum estado de equilíbrio.*

A unicidade é um problema mais complexo. No capítulo 2 definiremos a classe das transformações expansoras, para as quais tem-se unicidade do estado de equilíbrio para todo potencial Hölder. Nos capítulos 3 e 4 também estudaremos dinâmicas para as quais tem-se unicidade para potenciais Hölder. Para finalizar, daremos dois exemplos simples.

**Exemplo 3.68.** Se $f : M \to M$ tem entropia topológica nula, toda probabilidade invariante $\mu$ é medida de máxima entropia, já que $h_\mu(f) = 0 = h(f)$. Para um potencial qualquer $\phi : M \to \mathbb{R}$,

$$P(f,\phi) = \sup\left\{\int \phi d\nu : \nu \in \mathcal{M}_1(f)\right\}.$$

Logo, $\nu$ é estado de equilíbrio se, e somente se, $\nu$ maximiza a integral de $\phi$. Como a função $\nu \to \int \phi d\nu$ é contínua e $\mathcal{M}_1(f)$ é compacto, relativamente à topologia fraca$^*$, máximos existem para todo potencial $\phi$.

**Exemplo 3.69.** Seja $\sigma : \Sigma \to \Sigma$ o deslocamento em $\Sigma = \{1, \ldots, d\}^{\mathbb{N}}$ e seja $\mu$ a medida de Bernoulli dada por um vetor de probabilidade $p = (p_1, \ldots, p_d)$. Como vimos no Exemplo 3.18,

$$h_\mu(\sigma) = \sum_{i=1}^{d} -p_i \log(p_i).$$

Afirmamos que esta função atinge o seu máximo exatamente quando os $p_i$ são todos iguais a $\dfrac{1}{d}$. De fato, usando o Lema 3.58 com os $a_i$ todos iguais a zero, temos que $A = d$ e a desigualdade (3.62) dada pelo Lema pode ser reescrita da seguinte forma:

$$h_\mu(\sigma) = \sum_{i=1}^{d} -p_i \log p_i \leq \log d$$

para todas as medidas de Bernoulli $\mu$. Para finalizar lembre que $\log d$ é exatamente a entropia do deslocamento com relação ao vetor de probabilidade $\left(\dfrac{1}{d}, \ldots, \dfrac{1}{d}\right)$. Além disso, pode ser verificado que $h(\sigma) = \log d$. Portanto, a medida de Bernoulli dada pelo vetor $p = \left(\dfrac{1}{d}, \ldots, \dfrac{1}{d}\right)$ é a única medida de máxima entropia entre todas as medidas de Bernoulli.

# 4 DINÂMICA DAS TRANSFORMAÇÕES EXPANSORAS E TEOREMA DE RUELLE

Estudaremos nesse capítulo a dinâmica das transformações expansoras, com o objetivo de demonstrar a existência e unicidade de estados de equilíbrio para potenciais Hölder contínuos. Para isso, introduzimos o operador de Ruelle-Perron-Frobenius e demonstraremos o Teorema de Ruelle. Também daremos condições em termos dos potenciais que garantem quando os respectivos estados de equilíbrio são iguais (tal resultado corresponde ao Teorema de Livsič).

## 4.1 Transformações Expansoras e o Operador de Ruelle-Perron-Frobenius

**Definição 4.1.** Uma transformação contínua $f : M \to M$ num espaço métrico compacto $M$ é dita **expansora** (ou $\sigma$-expansora) se existem constantes $\sigma > 1$ e $\rho > 0$ tais que para todo $p \in M$, a imagem da bola $B(p, \rho)$ contém uma vizinhança do fecho de $B(f(p), \rho)$ e

$$d(f(x), f(y)) \geq \sigma d(x, y) \ \text{ para todo } \ x, y \in B(p, \rho). \tag{4.1}$$

**Exemplo 4.2.** Seja $\sigma_A^+ : \Sigma_A^+ \to \Sigma_A^+$ o deslocamento unilateral de tipo finito associado a uma matriz de transição $A$ (veja também o exemplo 3.48). Consideramos em $\Sigma_A^+$ a distância definida por

$$d((x_n)_n, (y_n)_n) = 2^{-N}, \quad N = \inf\{n \in \mathbb{N} : x_n \neq y_n\}.$$

Então $\sigma_A^+$ é uma transformação expansora. De fato, fixe $\rho \in \left(\frac{1}{2}, 1\right)$ e $\sigma = 2$. A bola de raio $\rho$ em torno de qualquer ponto $(p_n)_n \in \Sigma_A^+$ é o cilindro $[0; p_0]_A$ que contém esse ponto. A definição da distancia $d$ dá que

$$d(\sigma_A^+((x_n)_n), \sigma_A^+((y_n)_n)) = d((x_{n+1})_n, (y_{n+1})_n) = 2d((x_n)_n, (y_n)_n)$$

para quaisquer $(x_n)_n$ e $(y_n)_n$ no cilindro $[0; p_0]_A$. Além disso, $\sigma_A^+([0; p_0]_A)$ é a união de todos os cilindros $[0; q]$ tais que $A_{p_0,q} = 1$. Em particular, ela contém o cilindro $[0; p_1]_A$. Como os cilindros são abertos e fechados de $\Sigma_A^+$, isto mostra que a imagem da bola de raio $\rho$ em torno de $(p_n)_n$ contém uma vizinhança do fecho da bola de raio $\rho$ em torno de $(p_{n+1})_n$. Isto mostra que todo deslocamento de tipo finito é uma transformação expansora.

**Definição 4.3.** Seja $f : M \to M$ uma transformação expansora. A equação (4.1) mostra que a restrição de $f$ a cada bola $B(p, \rho)$ é injetiva, além disso, $f(B(p, \rho))$ contém o fecho de $B(f(p), \rho)$. Portanto, a restrição de $f$ a $B(p, \rho) \cap f^{-1}(B(f(p), \rho))$ é um homeomorfismo sobre $B(f(p), \rho)$. Chamamos **ramo inverso de** $f$ **em** $p$ à inversa

$$h_p : B(f(p), \rho) \to B(p, \rho)$$

da restrição de $f$ a $B(p, \rho) \cap f^{-1}(B(f(p), \rho))$. Segue da definição que $h_p(f(p)) = p$ e que $f \circ h_p = id$. A condição (4.1) implica que $h_p$ é $\sigma^{-1}$-contratora:

$$d(h_p(z), h_p(w)) \leq \sigma^{-1} d(f(h_p(z)), f(h_p(w))) = \sigma^{-1} d(z, w) \text{ para todo } z, w \in B(f(p), \rho). \tag{4.2}$$

Em geral, dado qualquer $n \geq 1$, chamamos **ramo inverso de** $f^n$ **em** $p$ à composição

$$h_p^n = h_p \circ h_{f(p)} \circ \cdots \circ h_{f^{n-1}(p)} : B(f^n(p), \rho) \to B(p, \rho)$$

dos ramos inversos de $f$ nos iterados de $p$. Como anteriormente $h_p^n(f^n(p)) = p$ e $f^n \circ h_p^n = id$, porque $h_{f^j(p)}(f(f^j(p))) = f^j(p)$ e $f \circ h_{f^j(p)} = id$ para todo $0 \leq j \leq n-1$.

Também, $f^j \circ h_p^n = f^j \circ h_p^j \circ h_{f^j(p)}^{n-j} = id \circ h_{f^j(p)}^{n-j} = h_{f^j(p)}^{n-j}$ para cada $0 \leq j \leq n$. Além disso, usando a condição (4.1) para o iterado $f^{n-j}$, vem que

$$\begin{aligned} d(f^j(h_p^n(z)), f^j(h_p^n(w))) &\leq \sigma^{-(n-j)} d(f^{n-j}(f^j(h_p^n(z))), f^{n-j}(f^j(h_p^n(w)))) \\ &= \sigma^{-(n-j)} d(f^n(h_p^n(z)), f^n(h_p^n(w))) \qquad (4.3) \\ &= \sigma^{j-n} d(z, w) \end{aligned}$$

para todo $z, w \in B(f^n(p), \rho)$ e todo $0 \leq j \leq n$.

Vamos descrever alguns fatos sobre a dinâmica das transformações expansoras que utilizaremos mais adiante.

**Lema 4.4.** *Se $f : M \to M$ é expansora então, para todo $y \in M$,*

$$f^{-1}(B(y, \rho)) = \bigcup_{x \in f^{-1}(y)} h_x(B(y, \rho)).$$

*Demonstração.* A relação $f \circ h_x = id$ implica que $h_x(B(y, \rho))$ está contido na pré-imagem de $B(y, \rho)$ para todo $x \in f^{-1}(y)$. Portanto $f^{-1}(B(y, \rho)) \supset \bigcup_{x \in f^{-1}(y)} h_x(B(y, \rho))$. Para provar a outra inclusão, seja $z$ qualquer ponto tal que $f(z) \in B(y, \rho)$. Pela definição de transformação expansora, $f(B(z, \rho))$ contém $B(f(z), \rho)$ e, portanto, contém $y$. Seja $h_z : B(f(z), \rho) \to M$ o ramo inverso de $f$ que envia $f(z)$ em $z$ e seja $x = h_z(y)$. Tanto $z$ quanto $h_x(f(z))$ estão em $B(x, \rho)$. Como $f$ é injetiva em cada bola de raio $\rho$ e $f(z) = f(h_x(f(z)))$, segue que $z = h_x(f(z))$. Portanto $f^{-1}(B(y, \rho)) \subset \bigcup_{x \in f^{-1}(y)} h_x(B(y, \rho))$. $\square$

**Lema 4.5.** *Se $f : M \to M$ é expansora, $f^n(B(p, n+1, \varepsilon)) = B(f^n(p), \varepsilon)$ para todo $p \in M$, $n \geq 0$ e $\varepsilon \in (0, \rho]$.*

*Demonstração.* A inclusão $f^n(B(p, n+1, \varepsilon)) \subset B(f^n(p), \varepsilon)$ é consequência imediata da definição de bola dinâmica. Para provar a recíproca, considere o ramo inverso $h_p^n : B(f^n(p), \rho) \to B(p, \rho)$. Dado qualquer $y \in B(f^n(p), \varepsilon)$, considere $x = h_p^n(y)$. Então, $f^n(x) = f^n(h_p^n(y)) = id(y) = y$ e, pela propriedade (4.3),

$$d(f^j(x), f^j(p)) = d(f^j(h_p^n(y)), f^j(h_p^n(p))) \leq \sigma^{j-n} d(f^n(x), f^n(p)) \leq d(y, f^n(p)) < \varepsilon$$

para todo $0 \leq j \leq n$. Isto mostra que $x \in B(p, n+1, \varepsilon)$. Portanto $f^n(B(p, n+1, \varepsilon)) \supset B(f^n(p), \varepsilon)$. $\square$

**Corolário 4.6.** *Toda transformação expansora é expansiva (veja a definição 3.22).*

*Demonstração.* Suponha que existam $z$ e $w$ tais que $d(f^n(z), f^n(w)) < \rho$ para todo $n \geq 0$. Isso implica $f^n(z) \in B(f^n(w), \rho)$ para todo $n \geq 0$. Portanto o mesmo ramo inverso que envia $f^n(w)$ em $w$ envia $f^n(z)$ em $z$. Logo $z = h_w^n(f^n(z))$ para todo $n \geq 0$. Então, a propriedade (4.3) dá que

$$d(z, w) = d(h_w^n(f^n(z)), h_w^n(f^n(w))) \leq \sigma^{-n} d(f^n(z), f^n(w)) < \rho\sigma^{-n}.$$

Pela definição de transformação expansora $\sigma > 1$. Portanto $\sigma^{-1} < 1$. Fazendo $n \to \infty$, concluímos que $z = w$. Portanto, $\rho$ é uma constante de expansividade para $f$. $\square$

Enunciaremos abaixo o Lema de sombreamento em sua versão para transformações expansoras. A demonstração neste momento será omitida, contudo, demonstraremos este Teorema na sua versão para Difeomorfismos Axioma A no Lema 5.37. Começaremos com a seguinte definição:

**Definição 4.7.** Dado $\delta > 0$, chamamos $\delta$-pseudo-órbita da transformação $f : M \to M$ qualquer sequência $(x_n)_{n \geq 0}$ tal que $d(f(x_n), x_{n+1}) < \delta$ para todo $n \geq 0$. Dizemos que a $\delta$-pseudo-órbita é periódica se existe $\kappa \geq 1$ tal que $x_n = x_{n+\kappa}$ para todo $n \geq 0$.

**Lema 4.8 (Lema do Sombreamento).** *Suponha que $f : M \to M$ é uma transformação expansora. Para todo $\varepsilon > 0$ existe $\delta > 0$ tal que para toda $\delta$-pseudo-órbita $(x_n)_n$ existe $x \in M$ tal que $d(f^n(x), x_n) < \varepsilon$ para todo $n \geq 0$. Se $\varepsilon$ é suficientemente pequeno para que $2\varepsilon$ seja uma constante de expansividade de $f$ então o ponto $x$ é único. Se, além disso, a pseudo-órbita é periódica então $x$ é ponto periódico.*

Vamos definir agora um importante objeto para o estudo dos estados de equilíbrio, que é o operador de transferência.

**Definição 4.9.** Suponha que $f : M \to M$ é transformação expansora num espaço métrico compacto e que $\varphi : M \to \mathbb{R}$ é um potencial. O **operador de transferência**, ou **operador de Ruelle-Perron-Frobenius** é o operador linear $\mathcal{L} : C^0(M) \to C^0(M)$

definido no espaço $C^0(M)$ das funções contínuas complexas por

$$\mathcal{L}g(y) = \sum_{x \in f^{-1}(y)} e^{\varphi(x)} g(x). \tag{4.4}$$

Note que $\mathcal{L}$ está bem definido: $\mathcal{L}g \in C^0(M)$ sempre que $g \in C^0(M)$. De fato, como vimos no Lema 4.4, para cada $y \in M$ existem ramos inversos $h_i : B(y, \rho) \to M$, $i = 1, \cdots, k$ da transformação $f$ tais que $\cup_{i=1}^k h_i(B(y, \rho))$ coincide com a pré-imagem da bola $B(y, \rho)$. Então

$$\mathcal{L}g = \sum_{i=1}^{k} (e^{\varphi} g) \circ h_i. \tag{4.5}$$

restrito a $B(y, \rho)$ e, claramente, esta expressão define uma função contínua.

Segue da definição que $\mathcal{L}$ é um operador positivo: se $g(x) \geq 0$ para todo $x \in M$ então $\mathcal{L}g(y) \geq 0$ para todo $y \in M$. Notando que o número de pré-imagens de uma transformação expansora é sempre finito: basta considerar uma cobertura finita do compacto $M$ por bolas de raio $\rho$ e observar que qualquer ponto tem no máximo uma pré-imagem em cada uma dessas bolas, verificamos que $\mathcal{L}$ é um operador contínuo: de fato,

$$||\mathcal{L}g|| = \sup |\mathcal{L}g| \leq \text{grau}(f) e^{\sup \varphi} \sup |g| = \text{grau}(f) e^{\sup \varphi} ||g||, \tag{4.6}$$

onde $\text{grau}(f) = \max\{\# f^{-1}(y) : y \in M\}$. Como isso vale para todo $g \in C^0(M)$, então $||\mathcal{L}|| \leq \text{grau}(f) e^{\sup \varphi}$.

De acordo com o teorema de Riesz-Markov, o dual do espaço de Banach $C^0(M)$ se identifica com o espaço vetorial $\mathcal{M}(M)$ das medidas borelianas complexas. Então, podemos considerar o seu operador dual.

**Definição 4.10. O dual do operador de transferência** é o operador linear $\mathcal{L}^* : \mathcal{M}(M) \to \mathcal{M}(M)$ definido por

$$\int g d(\mathcal{L}^* \eta) = \int (\mathcal{L}g) d\eta \quad \text{para todo} \quad g \in C^0(M) \quad \text{e} \quad \eta \in \mathcal{M}(M). \tag{4.7}$$

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

Estamos interessados em estudar propriedades espectrais de $\mathcal{L}$ e $\mathcal{L}^*$. Como são operadores positivos, o raio espectral $r(\mathcal{L}^*) = r(\mathcal{L})$ é autovalor de $\mathcal{L}^*$.

**Definição 4.11.** Um subconjunto fechado e convexo $C$ de um espaço de Banach $E$ é chamado de cone se

$$\alpha C \subset C \text{ para todo } \alpha \geq 0 \text{ e } C \cap (-C) = \{0\}. \tag{4.8}$$

Dizemos que o cone $C$ é normal quando

$$\inf\{||x+y|| : x, y \in C \text{ tais que } ||x|| = ||y|| = 1\} > 0. \tag{4.9}$$

O resultado a seguir é uma consequência do Teorema de Banach-Mazur e pode ser encontrado no livro de Deimling [2], capítulo 6, exercício 12.

**Lema 4.12.** *Seja $C$ um cone normal num espaço de Banach $E$ e seja $T : E \to E$ um operador linear positivo sobre $C$, isto é $T(C) \subset C$. Então, o raio espectral $r(T^*) = r(T)$ é autovalor do operador dual $T^* : E^* \to E^*$ e admite algum autovetor $v^* \in C^*$.*

## 4.2 Teorema de Ruelle para Transformações Expansoras

Vimos até agora, condições que garantem a existência de algum estado de equilíbrio para um potencial (Lema 3.66 e Corolário 3.67). Mas nada foi dito a respeito da unicidade deste objeto. O Teorema de Ruelle, que provaremos a seguir, garante existência e unicidade do estado de equilíbrio para a classe das transformações expansoras e para todo potencial Hölder. Daremos primeiros algumas definições.

**Definição 4.13.** Uma transformação $f : M \to M$ num espaço métrico compacto é **topologicamente exata** se para todo aberto $U \subset M$ existe $N \geq 1$ tal que $f^N(U) = M$. Note que o deslocamento (exemplo 4.2) é uma transformação expansora topologicamente exata

**Definição 4.14.** Uma medida, $\nu$ é um **estado de Gibbs** se existem constantes $K \geq 1$ e $P \in \mathbb{R}$ tais que

$$K^{-1} \leq \frac{\nu(B(x,n,\varepsilon))}{\exp(\varphi_n(x) - nP)} \leq K \tag{4.10}$$

para todo $x \in M$ e todo $n$. Recorde que $\varphi_n(x) = \sum_{i=0}^{n-1} \varphi(f^i(x))$ e $\varphi$ é um potencial no espaço métrico compacto $M$.

**Definição 4.15.** O sistema $(f, \mu)$ é dito **exato** se dado $B \subset M$ tal que existem conjuntos mensuráveis $B_n$ satisfazendo $B = f^{-n}(B_n)$ para todo $n \geq 1$, então $\mu(B) = 0$ ou $\mu(B) = 1$.

**Definição 4.16.** O **suporte** de uma medida $\mu$ é o conjunto dos pontos $x \in M$ tais que $\mu(V) > 0$ para toda vizinhança $V$ de $x$. Denotamos o suporte de $\mu$ por $\operatorname{supp} \mu$.

Agora vem o enunciado do Teorema.

**Teorema 4.17 (Ruelle).** *Seja $f : M \to M$ uma transformação expansora topologicamente exata num espaço métrico compacto e seja $\varphi : M \to \mathbb{R}$ um potencial Hölder. Então existe um único estado de equilíbrio $\mu$ para $\varphi$. Além disso, a medida $\mu$ é um estado de Gibbs, é exata e* $\operatorname{supp} \mu = M$.

Antes de passarmos aos detalhes da demonstração, faremos alguns comentários breves. Primeiramente, note que a existência de estado de equilíbrio segue imediatamente do Corolário 3.67, já que vimos no Corolário 4.6 que toda transformação expansora é expansiva. No entanto este fato não será usado na demonstração.

O outro comentário é que a fórmula de Rokhlin, que usaremos mais adiante, é válida sob as hipóteses do Teorema de Ruelle: Se $\mathscr{P}$ é uma partição finita de $M$ com $\operatorname{diam} \mathscr{P} < \rho$ (daqui em diante $\rho$ será o mesmo da definição de transformação espansora), então para cada $n \geq 1$, todo elemento de $\mathscr{P}^n$ está contido na imagem $h^{n-1}(P)$ de algum $P \in \mathscr{P}$ por algum ramo inverso $h^{n-1}$ do iterado $f^{n-1}$. Logo $\mathscr{P}$ é uma partição em domínios de invertibilidade. Como $\operatorname{diam} \mathscr{P}^n < \sigma^{-n+1}\rho$ para todo $n$, segue que $\cup_n \mathscr{P}^n$ gera a $\sigma$-álgebra de $M$. Logo, a fórmula de Rokhlin vale para toda probabilidade invariante de $f$.

***Demonstração do Teorema de Ruelle***. A demonstração está dividida em cinco passos para facilitar a leitura. Ao início de cada passo explicamos o que pretendemos provar.

**Passo 1:** Mostraremos que existe alguma probabilidade $\nu$ em $M$ tal que $\mathcal{L}^*\nu = \lambda\nu$ (Lema 4.18), a qual chamaremos de medida de referência. Vamos provar que existe um

Jacobiano de $f$ com respeito a $\nu$ (Lema 4.20), que seu suporte é o espaço $M$ inteiro (Corolário 4.22) e que é um estado de Gibbs (Lema 4.25). Para esta última propriedade provaremos dois resultados de controle da distorção limitada (Lema 4.23 e Corolário 4.24), e para estes a hipótese de que o potencial $\varphi$ é Hölder é fundamental.

**Lema 4.18.** *Considere o raio espectral $\lambda = r(\mathcal{L}^*) = r(\mathcal{L})$. Então existe alguma probabilidade $\nu$ em $M$ tal que $\mathcal{L}^*\nu = \lambda\nu$.*

*Demonstração.* O conjunto $C^0_+(M) \subset C^0(M)$ das funções contínuas positivas é um cone normal, e é preservado pelo operador de transferência $\mathcal{L}$. Portanto, podemos aplicar o Lema 4.12 com $E = C^0(M)$, $C = C^0_+(M)$ e $T = \mathcal{L}$. A conclusão do Lema 4.12 significa que $\mathcal{L}^*$ admite algum autovetor (uma automedida) $\nu \in C^0_+(M)^* = \{\nu \in C^0(M)^* : \nu(\psi) \geq 0 \text{ para todo } \psi \in C^0_+(M)\}$ no cone das medidas borelianas positivas finitas correspondente ao autovalor $\lambda$. Normalizando $\nu$, podemos supor que se trata de uma probabilidade. $\square$

**Definição 4.19.** A partir daqui sempre suporemos que $\nu$ é uma **medida de referência**, ou seja, uma probabilidade satisfazendo $\mathcal{L}^*\nu = \lambda\nu$ para algum $\lambda > 0$. De fato, concluiremos na observação 4.28 que $\lambda$ deve ser igual a $r(\mathcal{L}^*) = r(\mathcal{L})$.

**Lema 4.20.** *A transformação $f : M \to M$ admite jacobiano relativamente a $\nu$, dado por $J_\nu f = \lambda e^{-\varphi}$.*

*Demonstração.* Seja $A$ um domínio de invertibilidade qualquer de $f$. Seja $(g_n)_n$ uma sequência de funções contínuas convergindo em $\nu$-quase todo ponto para a função característica de $A$ e tal que $\sup |g_n| \leq \sup \chi_A = 1$ para todo $n$. Observe que

$$\mathcal{L}(e^{-\varphi} g_n)(y) = \sum_{x \in f^{-1}(y)} e^{\varphi(x)} e^{-\varphi(x)} g_n(y) = \sum_{x \in f^{-1}(y)} g_n(y).$$

A expressão do lado direito é menor ou igual que $\sup |g_n| \operatorname{grau}(f) \leq \operatorname{grau}(f)$, e

$$\lim_n \mathcal{L}(e^{-\varphi} g_n)(y) = \lim_n \sum_{x \in f^{-1}(y)} g_n(y) = \sum_{x \in f^{-1}(y)} \chi_A(y) = \chi_{f(A)}(y)$$

em $\nu$-quase todo ponto. Logo, usando o Teorema da convergência dominada, por um lado temos

$$\lim_n \int \mathcal{L}(e^{-\varphi} g_n) d\nu = \int \lim_n \mathcal{L}(e^{-\varphi} g_n) d\nu = \int \chi_{f(A)} d\nu = \nu(f(A)),$$

mas por outro lado, segue da definição do operador dual do operador de transferência e do Lema anterior que

$$\lim_n \int \mathcal{L}(e^{-\varphi} g_n) d\nu = \lim_n \int (e^{-\varphi} g_n) d(\mathcal{L}^* \nu) = \lim_n \int (e^{-\varphi} g_n) d(\lambda \nu) = \lim_n \int (\lambda e^{-\varphi} g_n) d\nu.$$

Novamente, pelo Teorema da convergência dominada

$$\lim_n \int (\lambda e^{-\varphi} g_n) d\nu = \int \lim_n (\lambda e^{-\varphi} g_n) d\nu = \int (\lambda e^{-\varphi} \chi_A) d\nu = \int_A (\lambda e^{-\varphi}) d\nu.$$

Segue que $\nu(f(A)) = \int_A (\lambda e^{-\varphi}) d\nu$, e portanto $J_\nu f = \lambda e^{-\varphi}$ como afirmado. □

**Lema 4.21.** *Seja $f : M \to M$ uma transformação expansora topologicamente exata e seja $\eta$ qualquer probabilidade boreliana tal que existe $J_\eta f$. Então* $\operatorname{supp} \eta = M$.

*Demonstração.* Suponha, por absurdo, que existe algum aberto $U \subset M$ tal que $\eta(U) = 0$. Note que $f$ é uma aplicação aberta, uma vez que é um homeomorfismo local. Então, a imagem $f(U)$ também é um aberto. Além disso, podemos particionar $U$ em uma união finita de domínios de invertibilidade $A$. Para cada um deles,

$$\eta(f(A)) = \int_A J_\eta f d\eta = 0, \text{ pois } \eta(A) = 0.$$

Portanto, $\eta(f(U)) = 0$. Por indução, segue que $\eta(f^n(U)) = 0$ para todo $n \geq 0$. Por outro lado, como supomos que $f$ é topologicamente exata, existe $N \geq 1$ tal que $f^N(U) = M$. Como $\eta(M) = 1$, isto gera uma contradição. □

**Corolário 4.22.** $\operatorname{supp} \nu = M$.

*Demonstração.* Segue do Lema anterior, uma vez que o Jacobiano de $f$ com respeito a $\nu$ existe e é igual a $\lambda e^{-\varphi}$, que é não negativo. □

Nos próximos lemas, vamos estudar a variação de $\varphi_n(x) = \sum_{i=0}^{n-1} \varphi(f^i(x))$ dentro de uma bola dinâmica em $x$. Para isso utilizaremos a regularidade Hölder do potencial. Fixe constantes $K_0 > 0$ e $\alpha > 0$ tais que $|\varphi(z) - \varphi(w)| \leq K_0 d(z,w)^\alpha$ para quaisquer $z, w \in M$.

**Lema 4.23.** *Existe $K_1 > 0$ tal que para todo $n \geq 1$, todo $x \in M$ e todo $y \in B(x, n+1, \rho)$,*

$$|\varphi_n(x) - \varphi_n(y)| \leq K_1 d(f^n(x), f^n(y))^\alpha.$$

*Demonstração.* Por hipótese, $d(f^j(x), f^j(y)) < \rho$ para todo $0 \leq j \leq n$. Então, para cada $j = 1, \cdots, n$, o ramo contrativo $h^j : B(f^n(x), \rho) \to M$ de $f^j$ que envia $f^n(x)$ em $f^{n-j}(x)$ também envia $f^n(y)$ em $f^{n-j}(y)$. Logo, lembrando da fórmula em (4.3), temos que

$$d(f^{n-j}(x), f^{n-j}(y)) = d(f^{n-j}(h_x^n(f^n(x))), f^{n-j}(h_x^n(f^n(y)))) \leq \sigma^{-j} d(f^n(x), f^n(y))$$

para todo $j = 1, \cdots, n$. Então,

$$\begin{aligned}
|\varphi_n(x) - \varphi_n(y)| &\leq \sum_{j=0}^{n-1} |\varphi(f^j(x)) - \varphi(f^j(y))| \\
&= \sum_{j=1}^{n} |\varphi(f^{n-j}(x)) - \varphi(f^{n-j}(y))| \\
&\leq \sum_{j=1}^{n} K_0 d(f^{n-j}(x), f^{n-j}(y))^\alpha \\
&\leq \sum_{j=1}^{n} K_0 \sigma^{-j\alpha} d(f^n(x), f^n(y))^\alpha.
\end{aligned}$$

Portanto, tomando $K_1 = K_0 \sum_{j=0}^{\infty} \sigma^{-j\alpha}$, segue o desejado. $\square$

O lema acima permite compararmos o valor de $J_\nu f^n$ de dois pontos numa mesma bola dinâmica.

**Corolário 4.24.** *Existe $K_2 > 0$ tal que para todo $n \geq 1$, todo $x \in M$ e todo $y \in B(x, n+1, \rho)$,*

$$K_2^{-1} \leq \frac{J_\nu f^n(x)}{J_\nu f^n(y)} \leq K_2.$$

*Demonstração.* Pelo Lema 4.20, $J_\nu f = \lambda e^{-\varphi}$. Pelo Lema 3.39

$$J_\nu f^n(z) = \prod_{i=0}^{n-1} J_\nu f(f^i(z)) = \prod_{i=0}^{n-1} (\lambda e^{-\varphi(f^i(z))}) = \lambda^n e^{-\sum_{i=0}^{n-1} \varphi(f^i(z))} = \lambda^n e^{-\varphi_n(z)} \tag{4.11}$$

para todo $z \in M$e todo $n \geq 1$. Então, pelo Lema anterior

$$\begin{aligned}
\left| \log \frac{J_\nu f^n(x)}{J_\nu f^n(y)} \right| &= |n \log \lambda - \varphi_n(x) - n \log \lambda + \varphi_n(y)| \\
&= |\varphi_n(x) - \varphi_n(y)| \\
&\leq K_1 d(f^n(x), f^n(y))^\alpha.
\end{aligned}$$

Por hipótese, $d(f^n(x), f^n(y)) < \rho$. Portanto, tomamos $\log K_2 = K_1 \rho^\alpha$. $\square$

Concluímos que $\nu$ é um estado de Gibbs para $f$.

**Lema 4.25.** *Para todo $\varepsilon > 0$ suficientemente pequeno existe $K_3 = K_3(\varepsilon) > 0$ tal que, escrevendo $P = \log \lambda$, vale*

$$K_3^{-1} \leq \frac{\nu(B(x,n,\varepsilon))}{\exp(\varphi_n(x) - nP)} \leq K_3$$

*para todo $x \in M$ e todo $n \geq 1$.*

*Demonstração.* Considere $\varepsilon < \rho$. Então $f|_{B(y,\varepsilon)}$ é injetiva para todo $y \in M$ e, consequentemente, $f^n|_{B(x,n,\varepsilon)}$ é injetiva para todo $x \in M$ e todo $n$. Então,

$$\nu(f^n(B(x,n,\varepsilon))) = \int_{B(x,n,\varepsilon)} J_\nu f^n(y) d\nu(y).$$

Pelo Corolário anterior,

$$K_2^{-1} J_\nu f^n(y) \leq J_\nu f^n(x) \leq K_2 J_\nu f^n(y).$$

Então, resulta das duas últimas equações que

$$K_2^{-1}\nu(f^n(B(x,n,\varepsilon))) \leq J_\nu f^n(x)\nu((B(x,n,\varepsilon))) \leq K_2\nu(f^n(B(x,n,\varepsilon))). \tag{4.12}$$

Pelo Lema 4.5 também temos que $f^n(B(x,n,\varepsilon)) = f(B(f^{n-1}(x),\varepsilon))$ e, portanto, $\nu(f^n(B(x,n,\varepsilon))) = \int_{B(f^{n-1}(x),\varepsilon)} J_\nu f d\nu$. Além disso, temos

$$1 \geq \nu(f^n(B(x,n,\varepsilon))) = \int_{B(f^{n-1}(x),\varepsilon)} J_\nu f d\nu \geq \min(\lambda e^{-\varphi})\nu(B(f^{n-1}(x),\varepsilon)) \geq a > 0$$

para todo $x \in M$ e todo $n$. A expressão acima é limitada de zero porque $\min(\lambda e^{-\varphi}) > 0$ e $\operatorname{supp} \nu = M$. Portanto, Usando estas observações em (4.12), obtemos

$$K_2^{-1} a \leq J_\nu f^n(x)\nu((B(x,n,\varepsilon))) \leq K_2. \tag{4.13}$$

Agora, basta notar que $J_\nu f^n(x) = \lambda^n e^{-\varphi_n(x)} = \exp(nP - \varphi_n(x))$, como vimos em (4.11) e tomar $K_3 = \max\left\{\frac{K_2}{a}, K_2\right\}$. $\square$

**Passo 2:** Mostraremos que existe alguma autofunção positiva $h$ associada ao autovalor $\lambda > 0$ do operador $\mathcal{L}$. Ela será construída como um ponto de acumulação Cesaro da sequência de funções $\lambda^{-n}\mathcal{L}^n 1$. Para mostrar a existência deste ponto de acumulação (Lema 4.30), mostramos no (Corolário 4.27) que esta sequência é uniformemente limitada e no (Lema 4.29) que é equicontínua.

**Lema 4.26.** *Existe $K_4 > 0$ tal que*

$$K_4 d(y_1, y_2)^\alpha \leq \log \frac{\mathcal{L}^n 1(y_1)}{\mathcal{L}^n 1(y_2)} \leq K_4 d(y_1, y_2)^\alpha$$

*para todo $n \geq 1$ e quaisquer $y_1, y_2 \in M$ com $d(y_1, y_2) < \rho$.*

*Demonstração.* Segue de iterar a equação (4.5) que, dada qualquer função contínua $g$,

$$\mathcal{L}^n g = \sum_i (e^{\varphi_n} g) \circ h_i^n \text{ restrito a cada bola } B(y, \rho),$$

onde a soma é sobre os ramos inversos $h_i^n : B(y, \rho) \to M$ do iterado $f^n$. Em particular,

$$\frac{\mathcal{L}^n 1(y_1)}{\mathcal{L}^n 1(y_2)} = \frac{\sum_i (e^{\varphi_n} 1) \circ h_i^n(y_1)}{\sum_i (e^{\varphi_n} 1) \circ h_i^n(y_2)} = \frac{\sum_i e^{\varphi_n(h_i^n(y_1))}}{\sum_i e^{\varphi_n(h_i^n(y_2))}}.$$

Pelo Lema 4.23, para cada um desses ramos inversos $h_i^n$ tem-se

$$|\varphi_n(h_i^n(y_1)) - \varphi_n(h_i^n(y_2))| \leq K_1 d(f^n(h_i^n(y_1)), f^n(h_i^n(y_2)))^\alpha = K_1 d(y_1, y_2)^\alpha.$$

Consequentemente,

$$e^{-K_1 d(y_1, y_2)^\alpha} \leq \frac{\mathcal{L}^n 1(y_1)}{\mathcal{L}^n 1(y_2)} \leq e^{K_1 d(y_1, y_2)^\alpha}.$$

Portanto, basta tomar $K_4 \geq K_1$. □

**Corolário 4.27.** *Existe $K_5 > 0$ tal que ${K_5}^{-1} \leq \lambda^{-n} \mathcal{L}^n 1(x) \leq K_5$ para todo $n \geq 1$ e todo $x \in M$.*

*Demonstração.* Observe que, para todo $n \geq 1$,

$$\int \lambda^{-n} \mathcal{L}^n 1 d\nu = \int \lambda^{-n} d(\mathcal{L}^{*n} \nu) = \int \lambda^{-n} \lambda^n d\nu = 1. \tag{4.14}$$

Em particular, para todo $n \geq 1$,

$$\min_{y \in M} \lambda^{-n} \mathcal{L}^n 1(y) \leq 1 \leq \max_{y \in M} \lambda^{-n} \mathcal{L}^n 1(y). \tag{4.15}$$

Como $f$ é topologicamente exata, existe $N \geq 1$ tal que $f^N(B(x, \rho)) = M$ para todo $x \in M$. Agora, dados $x, y \in M$ quaisquer, podemos encontrar $x' \in B(x, \rho)$ tal que $f^N(x') = y$. Então, por um lado,

$$\mathcal{L}^{n+N} 1(y) = \mathcal{L}^N(\mathcal{L}^n 1)(y) = \sum_{z \in f^{-N}(y)} e^{\varphi_n(z)} \mathcal{L}^n 1(z) \geq e^{\varphi_n(x')} \mathcal{L}^n 1(x') \geq e^{-N \sup |\varphi|} \mathcal{L}^n 1(x').$$

Por outro lado, o Lema 4.26 dá que $\mathcal{L}^n 1(x') \geq \mathcal{L}^n 1(x)\exp(-K_4\rho^\alpha)$. Tome $c = \sup|\varphi|$ e $K \geq \exp(K_4\rho^\alpha)e^{cN}\lambda^N$. Combinando as desigualdades anteriores vem que

$$\mathcal{L}^{n+N}1(y) \geq \exp(-K_4\rho^\alpha)e^{-cN}\mathcal{L}^n 1(x) \geq K^{-1}\lambda^N\mathcal{L}^n 1(x)$$

para todo $x, y \in M$. Portanto, para todo $n \geq 1$,

$$\min \lambda^{-(n+N)}\mathcal{L}^{n+N}1 \geq K^{-1}\max \lambda^{-n}\mathcal{L}^n 1. \tag{4.16}$$

Combinando (4.15) e (4.16) obtemos,

$$\max \lambda^{-n}\mathcal{L}^n 1 \leq K \min \lambda^{-(n+N)}\mathcal{L}^{n+N}1 \leq K \text{ para todo } n \geq 1$$

e fazendo $m = n + N$, também resulta de (4.15) e (4.16) que

$$\min \lambda^{-m}\mathcal{L}^m 1 \geq K^{-1}\max \lambda^{N-m}\mathcal{L}^{m-N}1 \geq K^{-1} \text{ para todo } m > N.$$

Para terminar a demonstração, só falta estender esta última estimativa para os valores $m = 1, \cdots, N$. Para isso, observe que cada $\mathcal{L}^m 1$ é uma função contínua e positiva. Logo, pela compacidade de $M$, o mínimo de $\mathcal{L}^m 1$ é positivo para todo $m$. Então, podemos tomar $K_5 \geq K$ tal que $\min \lambda^{-m}\mathcal{L}^m 1 \geq {K_5}^{-1}$ para todo $m = 1, \cdots, N$. Então, para todo $n \geq 1$

$${K_5}^{-1} \leq K^{-1} \leq \min \lambda^{-n}\mathcal{L}^n 1 \leq \lambda^{-n}\mathcal{L}^n 1(x) \leq \max \lambda^{-n}\mathcal{L}^n 1 \leq K \leq K_5$$

para todo $x \in M$. □

**Observação 4.28.** *Segue imediatamente do Corolário 4.27 que o autovalor $\lambda$ está unicamente determinado (em vista do Lema 4.18, isso quer dizer que $\lambda$ é necessariamente igual ao raio espectral de $\mathcal{L}$ e $\mathcal{L}^*$)*

**Lema 4.29.** *Existe $K_6 > 0$ tal que*

$$|\lambda^{-n}\mathcal{L}^n 1(x) - \lambda^{-n}\mathcal{L}^n 1(y)| \leq K_6 d(x, y)^\alpha$$

*para quaisquer $n \geq 1$ e $x, y \in M$. Em particular, a sequência $\lambda^{-n}\mathcal{L}^n 1$ é equicontínua.*

*Demonstração.* Suponha inicialmente que $d(x, y) < \rho$. Pelo Lema 4.26,

$$\mathcal{L}^n 1(x) \leq \mathcal{L}^n 1(y)\exp(K_4 d(x, y)^\alpha)$$

e, portanto,

$$\lambda^{-n}\mathcal{L}^n 1(x) - \lambda^{-n}\mathcal{L}^n 1(y) \leq [\exp(K_4 d(x,y)^\alpha) - 1]\lambda^{-n}\mathcal{L}^n 1(y).$$

Tome $K > 0$ tal que $|\exp(K_4 t) - 1| \leq K|t|$ sempre que $|t| \leq \rho^\alpha$. Então, usando o Corolário 4.27,

$$\lambda^{-n}\mathcal{L}^n 1(x) - \lambda^{-n}\mathcal{L}^n 1(y) \leq K d(x,y)^\alpha \lambda^{-n}\mathcal{L}^n 1(y) \leq K K_5 d(x,y)^\alpha.$$

Refazendo o cálculo anterior invertendo os papéis de $x$ e $y$, obtemos $\lambda^{-n}\mathcal{L}^n 1(y) - \lambda^{-n}\mathcal{L}^n 1(x) \leq K K_5 d(x,y)^\alpha$. Portanto

$$|\lambda^{-n}\mathcal{L}^n 1(x) - \lambda^{-n}\mathcal{L}^n 1(y)| \leq K K_5 d(x,y)^\alpha \text{ sempre que } d(x,y) < \rho.$$

Quando $d(x,y) \geq \rho$, então $\dfrac{d(x,y)^\alpha}{\rho^\alpha} \geq 1$. Segue disso, da desigualdade triangular e do Corolário 4.27 que

$$|\lambda^{-n}\mathcal{L}^n 1(x) - \lambda^{-n}\mathcal{L}^n 1(y)| \leq 2K_5 \leq 2K_5 \frac{d(x,y)^\alpha}{\rho^\alpha}.$$

Logo, basta tomar $K_6 \geq \max\{K K_5, 2K_5\rho^{-\alpha}\}$. □

**Lema 4.30.** *O operador $\mathcal{L}$ admite alguma autofunção $h$ associada ao autovalor $\lambda$, isto é, a função $h$ satisfaz $\mathcal{L}h = \lambda h$. Além disso,*

- $\int h d\nu = 1$;
- ${K_5}^{-1} \leq h(x) \leq K_5$;
- $|h(x) - h(y)| \leq K_6 d(x,y)^\alpha$ *para todo* $x, y \in M$.

*Demonstração.* Considere a sequência das médias

$$h_n = \frac{1}{n}\sum_{i=0}^{n-1} \lambda^{-i}\mathcal{L}^i 1.$$

O corolário 4.27 mostra que esta sequência é equilimitada, enquanto que o Lema 4.29 mostra que esta sequência é equicontínua. Então, pelo teorema de Arzelá-Ascoli, existe alguma subsequência $(h_{n_i})_i$ convergindo uniformemente para uma função contínua $h$. Mostraremos agora que esta função $h$ satisfaz todas as condições propostas pelo Lema.

Primeiramente, como o operador $\mathcal{L}$ é contínuo

$$\begin{aligned}\mathcal{L}h = \mathcal{L}\left(\lim_i \frac{1}{n_i}\sum_{j=0}^{n_i-1}\lambda^{-j}\mathcal{L}^j 1\right) &= \lim_i \frac{1}{n_i}\sum_{j=0}^{n_i-1}\lambda^{-j}\mathcal{L}^{j+1}1 = \lim_i \frac{\lambda}{n_i}\sum_{j=0}^{n_i-1}\lambda^{-(j+1)}\mathcal{L}^{j+1}1 \\ &= \lim_i \frac{\lambda}{n_i}\sum_{l=1}^{n_i}\lambda^{-l}\mathcal{L}^l 1 = \lim_i \frac{\lambda}{n_i}\sum_{l=0}^{n_i-1}\lambda^{-l}\mathcal{L}^l 1 + \lim_i \frac{\lambda}{n_i}\left(\lambda^{-n_i}\mathcal{L}^{n_i}1 - 1\right).\end{aligned}$$

O primeiro termo do lado direito converge para $\lambda h$ e o segundo converge para zero, uma vez que a sequência $\lambda^{-n}\mathcal{L}^n 1$ é limitada. Portanto, $\mathcal{L}h = \lambda h$ tal como afirmamos.

Vimos em (4.14) que $\int \lambda^{-n}\mathcal{L}^n 1 d\nu = 1$ para todo $n$. Segue que $\int h_n d\nu = \int (1/n)\sum_{i=0}^{n-1}\lambda^{-i}\mathcal{L}^i 1 d\nu = 1$ para todo $n$. Como $\lambda^{-n}\mathcal{L}^n 1$ é limitada, o teorema da convergência dominada implica que, $\int h d\nu = 1$. As demais afirmações seguem de tomar o limite Cesaro nas expressões dadas pelo Corolário 4.27 e pelo Lema 4.29. $\square$

**Passo 3:** Agora que temos a automedida $\nu$ de $\mathcal{L}^*$ e a autofunção $h$ de $\mathcal{L}$, mostraremos que $\mu = h\nu$ é um estado de equilíbrio para $\varphi$. Para isso, mostraremos no (Lema 4.31) que $\mu$ é uma probabilidade invariante por $f$, equivalente a medida de referência $\nu$ e suportada em todo $M$. Além disso, $\mu$ é um estado de Gibbs e existe um Jacobiano de $f$ relativamente a $\mu$. Grande parte dessas propriedades de $\mu$ resultarão do que já provamos nos passos anteriores para a medida de referência $\nu$ e para a autofunção $h$. Mostraremos que $\mu$ é um estado de equilíbrio para o potencial $\varphi$ no (Lema 4.33) juntamente com o (Corolário 4.32).

**Lema 4.31.** *Considere a medida definida por $\mu = h\nu$, ou seja $\mu(A) = \int_A h d\nu$ para cada conjunto mensurável $A \subset M$, onde $\nu$ é a medida de referência do passo 1 e $h$ é a autofunção do passo 2. Então $\mu$ é uma probabilidade invariante por $f$, equivalente a medida de referência $\nu$ e suportada em todo $M$. Além disso, $\mu$ é um estado de Gibbs e existe Jacobiano de $f$ relativamente a $\mu$, dado por $J_\mu f = \lambda e^{-\varphi}(h \circ f)/h$.*

*Demonstração.* Do Lema 4.30 vem que $\mu(M) = \int h d\nu = 1$ e, portanto, $\mu$ é uma medida de probabilidade. Para mostrar a invariância, note que para quaisquer funções contínuas

$g_1, g_2 : M \to \mathbb{R}$ e para todo $y \in M$,

$$\begin{aligned}\mathcal{L}((g_1 \circ f)g_2)(y) &= \sum_{x \in f^{-1}(y)} e^{\varphi(x)} g_1(f(x)) g_2(x) \\ &= g_1(y) \sum_{x \in f^{-1}(y)} e^{\varphi(x)} g_2(x) \\ &= g_1(y) \mathcal{L} g_2(y).\end{aligned} \tag{4.17}$$

Então, para toda função contínua $g : M \to \mathbb{R}$,

$$\begin{aligned}\int (g \circ f) d\mu &= \int (g \circ f) d(h\nu) = \lambda^{-1} \int (g \circ f) h d(\lambda \nu) = \lambda^{-1} \int (g \circ f) h d(\mathcal{L}^* \nu) \\ &= \lambda^{-1} \int \mathcal{L}((g \circ f)h) d\nu = \lambda^{-1} \int g(\mathcal{L}h) d\nu = \lambda^{-1} \int g(\lambda h) d\nu \\ &= \lambda^{-1} \lambda \int g h d\nu = \int g d(h\nu) = \int g d\mu.\end{aligned} \tag{4.18}$$

Isto prova que a probabilidade $\mu$ é invariante por $f$. Do Lema 4.30, ${K_5}^{-1} \leq h(x) \leq K_5$ para todo $x \in M$. Portanto

$${K_5}^{-1} \nu(A) \leq \mu(A) \leq K_5 \nu(A) \tag{4.19}$$

para cada conjunto mensurável $A \subset M$. Logo $\mu$ é equivalente à medida de referência $\nu$.

O fato acima, juntamente com o Corolário 4.22 dá que $\operatorname{supp} \mu = M$. Também segue da relação (4.19), juntamente com o Lema 4.25, que $\mu$ é um estado de Gibbs:

$${K_5}^{-1} K_3^{-1} \leq {K_5}^{-1} \frac{\nu(B(x,n,\varepsilon))}{\exp(\varphi_n(x) - nP)} \leq \frac{\mu(B(x,n,\varepsilon))}{\exp(\varphi_n(x) - nP)} \leq K_5 \frac{\nu(B(x,n,\varepsilon))}{\exp(\varphi_n(x) - nP)} \leq K_5 K_3$$

para todo $x \in M$, $n \geq 1$, $P = \log \lambda$ e $\varepsilon > 0$.

Finalmente para mostrar que existe Jacobiano de $f$ relativamente a $\mu$, considere qualquer domínio de invertibilidade $A$ de $f$. Então, pela fórmula de mudança de variáveis dada pelo Lema 3.35 item (a),

$$\mu(f(A)) = \int_{f(A)} 1 d\mu = \int_{f(A)} h d\nu = \int_A J_\nu f (h \circ f) d\nu = \int_A J_\nu f \frac{(h \circ f)}{h} d\mu.$$

Pelo Lema 4.20, isto significa que

$$J_\mu f = J_\nu f \frac{(h \circ f)}{h} = \lambda e^{-\varphi} \frac{(h \circ f)}{h},$$

tal como foi afirmado. Isto conclui a prova do Lema. $\square$

**Corolário 4.32.** *A probabilidade invariante* $\mu = h\nu$ *satisfaz* $h_\mu(f) + \int \varphi d\mu = P$.

*Demonstração.* Combinando a fórmula de Rokhlin (Teorema 3.34) com a expressão de $J_\mu f$ dada pelo lema anterior,

$$h_\mu(f) = \int \log J_\mu f d\mu = \int \log\left(\lambda e^{-\varphi}\frac{(h\circ f)}{h}\right) d\mu = \log\lambda - \int \varphi d\mu + \int (\log h\circ f - \log h) d\mu.$$

Como $\mu$ é invariante e $\log h$ é limitada (Lema 4.30), a última parcela é igual a zero. Lembrando que $P = \log\lambda$, segue que $h_\mu(f) + \int \varphi d\mu = P$ conforme enunciado. □

**Lema 4.33.** $P(f,\varphi) = P = \log r(\mathcal{L})$. *Portanto* $\mu = h\nu$ *é um estado de equilíbrio para* $\varphi$.

*Demonstração.* Mostraremos inicialmente que se $\eta$ é uma probabilidade invariante por $f$ qualquer, satisfazendo

$$h_\eta(f) + \int \varphi d\eta \geq P \tag{4.20}$$

(por exemplo: a probabilidade $\mu$), então $h_\eta(f) + \int \varphi d\eta = P$.

Seja $g_\eta = 1/J_\eta f$ e considere também a função $g = \lambda^{-1}e^{\varphi}h/(h\circ f)$. Observe que

$$\begin{aligned}\sum_{x\in f^{-1}(y)} g(x) = \sum_{x\in f^{-1}(y)} \lambda^{-1}e^{\varphi(x)}\frac{h(x)}{h(f(x))} &= \frac{1}{\lambda h(y)}\sum_{x\in f^{-1}(y)} e^{\varphi(x)}h(x)\\ &= \frac{\mathcal{L}h(y)}{\lambda h(y)} = \frac{\lambda h(y)}{\lambda h(y)} = 1. \end{aligned} \tag{4.21}$$

para todo $y \in M$. Além disso, como $\eta$ é invariante por $f$, segue do Lema 3.38 que

$$\sum_{x\in f^{-1}(y)} g_\eta(x) = 1 \text{ para } \eta\text{-quase todo } y \in M. \tag{4.22}$$

Usando (4.20) e a fórmula de Rokhlin (Teorema 3.34)

$$0 \leq h_\eta(f) + \int \varphi d\eta - P = \int \log J_\eta f d\eta + \int \varphi d\eta - \log\lambda = \int (\log\frac{1}{g_\eta} + \varphi - \log\lambda) d\eta. \tag{4.23}$$

Observe que pela definição de $g$ e a hipótese de que $\eta$ é invariante por $f$

$$\begin{aligned}
\int(\log\frac{1}{g_\eta}+\varphi-\log\lambda)d\eta &= \int(-\log g_\eta+\varphi-\log\lambda)d\eta+0\\
&= \int(-\log g_\eta+\varphi-\log\lambda)d\eta+\int\log h\circ f d\eta-\int\log h d\eta\\
&= \int(-\log g_\eta+\log e^{\varphi}+\log\lambda^{-1}+\log h\circ f-\log h)d\eta\\
&= \int\left(-\log g_\eta+\log\lambda^{-1}e^{\varphi}\frac{h}{(h\circ f)}\right)d\eta\\
&= \int\left(-\log g_\eta+\log g\right)d\eta=\int\log\frac{g}{g_\eta}d\eta. \qquad (4.24)
\end{aligned}$$

Lembrando que $g_\eta = 1/J_\eta f$, segue do Lema 3.37 que

$$\int\log\frac{g}{g_\eta}d\eta=\int\sum_{x\in f^{-1}(y)}\frac{\log\frac{g}{g_\eta}}{J_\eta f}(x)d\eta(y)=\int\left(\sum_{x\in f^{-1}(y)}g_\eta(x)\log\frac{g}{g_\eta}(x)\right)d\eta(y). \qquad (4.25)$$

A esta altura precisamos do seguinte resutado:

**Lema 4.34.** *Sejam $p_i, b_i, i = 1,\cdots,k$ números reais positivos tais que $\sum_{i=1}^k p_i = 1$. Então $\sum_{i=1}^k p_i\log b_i \le \log\left(\sum_{i=1}^k p_i b_i\right)$ e a igualdade acontece se, e somente se, os números $b_j$ forem todos iguais a $\sum_{i=1}^k p_i b_i$.*

*Demonstração.* Tome $a_i = \log(p_i b_i)$ no Lema 3.58. Então a desigualdade no Lema 3.58 corresponde exatamente à desigualdade no presente lema:

$$\sum_{i=1}^k p_i\log b_i=\sum_{i=1}^k p_i(a_i-\log p_i)\le\log(\sum_{i=1}^k e^{a_i})=\log(\sum_{i=1}^k e^{\log(p_ib_i)})=\log(\sum_{i=1}^k p_ib_i).$$

Além disso, vale a igualdade se, e somente se,

$$p_j=\frac{e^{a_j}}{\sum_{i=1}^k e^{a_i}} \iff p_j=\frac{p_jb_j}{\sum_{i=1}^k p_ib_i} \iff b_j=\sum_{i=1}^k p_ib_i$$

para todo $j = 1,\cdots,n$. □

Para cada $y \in M$, tome $p_i = g_\eta(x_i)$ e $b_i = \log(g(x_i)/g_\eta(x_i))$, onde os $x_i$ pertencem a $f^{-1}(y)$ (A quantidade dos $x_i$ é finita porque $f$ é expansora). A igualdade (4.22) significa que $\sum_i p_i = \sum_{x\in f^{-1}(y)} g_\eta(x) = 1$ para $\eta$-quase todo $y \in M$. Então podemos aplicar o

Lema 4.34 e a igualdade (4.21) para obtermos

$$\begin{aligned}\sum_{x\in f^{-1}(y)} g_\eta(x)\log\frac{g}{g_\eta}(x) &\le \log\sum_{x\in f^{-1}(y)} g_\eta(x)\frac{g}{g_\eta}(x)\\ &= \log\sum_{x\in f^{-1}(y)} g(x) = \log 1 = 0 \end{aligned}\tag{4.26}$$

para $\eta$-quase todo $y$. Combinando as relações (4.23),(4.24),(4.25) e (4.26) obtemos:

$$0 \le h_\eta(f) + \int \varphi d\eta - P = \int \log\frac{g}{g_\eta} d\eta = \int\left(\sum_{x\in f^{-1}(y)} g_\eta(x)\log\frac{g}{g_\eta}(x)\right) d\eta(y) \le 0. \tag{4.27}$$

Portanto, se $\eta$ é uma probabilidade invariante por $f$ qualquer, satisfazendo $h_\eta(f)+\int\varphi d\eta \ge P$, então $h_\eta(f)+\int\varphi d\eta = P$. Em particular isto vale para a probabilidade $\mu$ devido ao Corolário 4.32. Então pelo princípio variacional (Teorema 3.57), segue que $P(f,\varphi) = P = h_\mu(f) + \int\varphi d\mu$. Que $P = \log r(\mathcal{L})$, segue da observação 4.28. Isto prova que a probabilidade $\mu = h\nu$ é um estado de equilíbrio para o potencial $\varphi$. $\square$

**Passo 4:** Mostraremos que a probabilidade $\mu = h\nu$ é o único estado de equilíbrio para o potencial $\varphi$, isto será feito no Corolário 4.38. O argumento depende fortemente do seguinte resultado que provaremos no Lema 4.37: *Todos os estados de equilíbrio de $\varphi$ são equivalentes*. Com esse resultado e algumas noções sobre as propriedades de medidas ergódicas, decomposição ergódica e lembrando que o conjunto dos estados de equilíbrio é convexo poderemos provar o Corolário 4.38. O Lema 4.35 e o Corolário 4.36 são o ponto de partida dessas ideias.

**Lema 4.35.** *Se $\eta$ é estado de equilíbrio para $\varphi$ então* $\operatorname{supp}\eta = M$ *e*

$$J_\eta f = \lambda e^{-\varphi}\frac{(h\circ f)}{h} \quad e \quad \mathcal{L}^*\left(\frac{\eta}{h}\right) = \lambda\left(\frac{\eta}{h}\right).$$

*Demonstração.* A igualdade (4.27) também implica que $\log\dfrac{g}{g_\eta} = 0$ para quase todo $y \in M$. Logo vale a igualdade em (4.26) para quase todo $y \in M$. De acordo com o Lema 4.34, isso acontece se, e somente se, os números $b_i = \log(g(x_i)/g_\eta(x_i))$ onde os $x_i$ pertencem a $f^{-1}(y)$ são todos iguais. Em outras palavras, para $\eta$-quase todo $y \in M$ existe um número $c(y)$ tal que

$$\frac{g(x)}{g_\eta(x)} = c(y) \text{ para todo } x \in f^{-1}(y).$$

Além disso, lembrando as igualdades (4.22) e (4.21),

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

para $\eta$-quase todo $y$. Segue que $g_\eta = g$ em quase todo ponto, ou seja, a função $\dfrac{1}{g} = \lambda e^{-\varphi} \dfrac{(h \circ f)}{h}$ é um jacobiano de $f$ relativamente a $\eta$. Isto prova a segunda afirmação.

Para provar a terceira afirmação, seja $\xi : M \to \mathbb{R}$ uma função contínua qualquer. Por um lado, usando a definição do operador de transferência

$$\begin{aligned} \int \xi d\mathcal{L}^* \left(\frac{\eta}{h}\right) = \int \frac{1}{h}(\mathcal{L}\xi) d\eta = \int \frac{1}{h(y)} \left( \sum_{x \in f^{-1}(y)} e^{\varphi(x)} \xi(x) \right) d\eta(y) \\ = \int \left( \sum_{x \in f^{-1}(y)} \frac{e^{\varphi(x)}}{h(y)} \xi(x) \right) d\eta(y). \end{aligned} \tag{4.28}$$

Pela definição da função $g$ e para todo $x \in f^{-1}(y)$,

$$g(x) = \lambda^{-1} e^{\varphi(x)} \frac{h(x)}{h(f(x))} \;\Rightarrow\; g(x) = \lambda^{-1} e^{\varphi(x)} \frac{h(x)}{h(y)} \;\Rightarrow\; \lambda \frac{g(x)}{h(x)} = \frac{e^{\varphi(x)}}{h(y)}.$$

Substituindo a última igualdade dada pelas implicações acima em (4.28), obtemos

$$\int \xi d\mathcal{L}^* \left(\frac{\eta}{h}\right) = \int \left( \sum_{x \in f^{-1}(y)} \lambda \frac{g(x)}{h(x)} \xi(x) \right) d\eta(y). \tag{4.29}$$

Então, lembrando que $g = g_\eta = \dfrac{1}{J_\eta f}$, podemos usar o Lema 3.37 para concluir que

$$\begin{aligned} \int \xi d\mathcal{L}^* \left(\frac{\eta}{h}\right) &= \int \left( \sum_{x \in f^{-1}(y)} \lambda \frac{g(x)}{h(x)} \xi(x) \right) d\eta(y) \\ &= \int \left( \sum_{x \in f^{-1}(y)} \lambda \frac{\xi(x)}{h(x)} \left(\frac{1}{J_\eta f}\right) \right) d\eta(y) \\ &= \int \frac{\lambda \xi}{h} d\eta = \int \xi d\left(\frac{\lambda \eta}{h}\right). \end{aligned}$$

Como a função contínua $\xi$ é arbitrária, isto mostra que $\mathcal{L}^*\left(\dfrac{\eta}{h}\right) = \left(\dfrac{\lambda\eta}{h}\right)$, como afirmamos. A primeira afirmação do enunciado é uma consequência imediata da existência do Jacobiano com respeito a $\eta$ feita na segunda afirmação e do Lema 4.21. $\square$

**Corolário 4.36.** *Existe $K_7 > 0$ tal que para todo estado de equilíbrio $\eta$, todo $n \geq 1$, todo $x \in M$ e todo $y \in B(x, n+1, \rho)$,*

$$K_7^{-1} \leq \frac{J_\eta f^n(x)}{J_\eta f^n(y)} \leq K_7.$$

*Demonstração.* Pelos Lemas 3.39 e 4.35,

$$\begin{aligned}
J_\eta f^n(x) = \prod_{i=0}^{n-1} J_\eta f(f^i(x)) &= \prod_{i=0}^{n-1} \lambda e^{-\varphi(f^i(x))} \frac{(h \circ f)}{h}(f^i(x)) \\
&= \lambda^n e^{-\varphi(x)} \frac{h(f(x))}{h(x)} e^{-\varphi(f(x))} \frac{h(f^2(x))}{h(f(x))} \cdots e^{-\varphi(f^{n-1}(x))} \frac{h(f^n(x))}{h(f^{n-1}(x))} \\
&= \lambda^n e^{-\sum_{i=0}^{n-1} \varphi(f^i(x))} \frac{(h \circ f^n)}{h}(x) = \lambda^n e^{-\varphi_n(x)} \frac{(h \circ f^n)}{h}(x) \\
&= \left( J_\nu f^n \frac{(h \circ f^n)}{h} \right)(x).
\end{aligned}$$

para cada $n \geq 1$ (A última igualdade decorre da equação (4.11) do Corolário 4.24). Então, usando o Corolário 4.24 e a desigualdade ${K_5}^{-1} \leq h(x) \leq K_5$ dada pelo Lema 4.30,

$$K_2{}^{-1} K_5{}^{-4} \leq \frac{J_\eta f^n(x)}{J_\eta f^n(y)} = \frac{J_\nu f^n(x) \dfrac{h(f^n(x))}{h(x)}}{J_\nu f^n(y) \dfrac{h(f^n(y))}{h(y)}} = \frac{J_\nu f^n(x) h(f^n(x)) h(y)}{J_\nu f^n(y) h(f^n(y)) h(x)} \leq K_2 K_5{}^4$$

Portanto, basta tomar $K_7 = K_2 K_5{}^4$ □

**Lema 4.37.** *Todos os estados de equilíbrio de $\varphi$ são equivalentes.*

*Demonstração.* Considere uma partição finita $\mathscr{P}$ de $M$ tal que todo $P \in \mathscr{P}$ tem interior não vazio e diâmetro menor que $\rho$. Como $\operatorname{supp} \eta_1 = \operatorname{supp} \eta_2 = M$ (pelo Lema 4.35) temos que $\{\eta_i(P) : i = 1, 2 \text{ e } P \in \mathscr{P}\}$ é limitado de zero. Consequentemente, existe $C_1 > 0$ tal que

$$\frac{1}{C_1} \leq \frac{\eta_1(P)}{\eta_2(P)} \leq C_1 \text{ para todo } P \in \mathscr{P}. \tag{4.30}$$

Vamos mostrar que esta relação ainda vale para todo subconjunto mensurável de $M$, a menos de substituirmos $C_1$ por uma constante conveniente $C_2 > C_1$.

Para cada $n \geq 1$, seja $\mathscr{Q}_n$ a partição de $M$ formada pelas imagens $h^n(P)$ dos elementos de $P$ pelos ramos inversos $h^n$ do iterado $f^n$. Pela definição de jacobiano,

$\eta_i(P) = \int_{h^n(P)} J_{\eta_i} f^n d\eta_i$. Logo, integrando a desigualdade dada pelo Corolário 4.36,

$$K_7^{-1} \int_{h^n(P)} J_{\eta_i} f^n(x) d\eta_i(y) \leq \int_{h^n(P)} J_{\eta_i} f^n(y) d\eta_i(y) \leq K_7 \int_{h^n(P)} J_{\eta_i} f^n(x) d\eta_i(y).$$

Portanto

$$K_7^{-1} J_{\eta_i} f^n(x) \leq \frac{\eta_i(P)}{\eta_i(h^n(P))} \leq K_7 J_{\eta_i} f^n(x)$$

para qualquer $x \in h^n(P)$. Lembrando que $J_{\eta_1} f = J_{\eta_2} f$ (O Lema 4.35 diz que o Jacobiano de $f$ de qualquer estado de equilíbrio $\eta$ de $\varphi$ é $J_\eta f = \lambda e^{-\varphi}(h \circ f)/h$ ), segue que

$$K_7^{-2} = K_7^{-2} \frac{J_{\eta_2} f^n(x)}{J_{\eta_1} f^n(x)} \leq \frac{\eta_2(P)\eta_1(h^n(P))}{\eta_1(P)\eta_2(h^n(P))} \leq K_7^2 \frac{J_{\eta_2} f^n(x)}{J_{\eta_1} f^n(x)} = K_7^2. \tag{4.31}$$

Combinando (4.30) e (4.31) e tomando $C_2 = C_1 K_7^2$ vem que

$$\frac{1}{C_2} = \frac{1}{C_1 K_7^2} \leq \frac{\eta_1(P)}{\eta_2(P)} \frac{\eta_2(P)\eta_1(h^n(P))}{\eta_1(P)\eta_2(h^n(P))} = \frac{\eta_1(h^n(P))}{\eta_2(h^n(P))} \leq C_1 K_7^2 = C_2. \tag{4.32}$$

Logo

$$\frac{1}{C_2} \leq \frac{\eta_1(h^n(P))}{\eta_2(h^n(P))} \leq C_2 \tag{4.33}$$

para todo $P \in \mathscr{P}$, todo ramo inverso $h^n$ de $f^n$ e todo $n \geq 1$. Em outras palavras, (4.30) vale para todo elemento de $\mathscr{Q}_n$, com $C_2$ no lugar de $C_1$.

Agora observe que $\operatorname{diam} \mathscr{Q}_n < \sigma^{-n}\rho$ para todo $n$, porque $d(z,w) = d(h_w^n(f^n(z)), h_w^n(f^n(w))) \leq \sigma^{-n} d(f^n(z), f^n(w)) < \sigma^{-n}\rho$ para todo $z, w \in \mathscr{Q}_n$ uma vez que $f$ é expansora. Usando que toda medida de probabilidade num espaço métrico é regular, para qualquer conjunto mensurável $B$ e qualquer $\delta$, podemos encontrar um compacto $F \subset B$ e um aberto $A \supset B$ tais que $\eta_i(A \setminus F) < \delta$ para $i = 1$ e para $i = 2$. Seja $\mathscr{R}_n$ a união de todos os elementos da partição $\mathscr{Q}_n$ que intersectam $F$. Temos que $\mathscr{R}_n \supset F$ e, supondo que $n$ é suficientemente grande, $\mathscr{R}_n \subset A$. Então,

$$\eta_1(B) \leq \eta_1(A) < \eta_1(\mathscr{R}_n) + \delta \quad \text{e} \quad \eta_2(B) \geq \eta_2(F) > \eta_2(\mathscr{R}_n) - \delta.$$

A relação (4.33) dá que $\eta_1(\mathscr{R}_n) \leq C_2 \eta_2(\mathscr{R}_n)$, uma vez que $\mathscr{R}_n$ é uma união (disjunta) de elementos de $\mathscr{Q}_n$. Combinando estas três últimas desigualdades, obtemos

$$\eta_1(B) \leq \eta_1(\mathscr{R}_n) + \delta \leq C_2 \eta_2(\mathscr{R}_n) + \delta \leq C_2(\eta_2(B) + \delta) + \delta.$$

Como $\delta$ é arbitrário, concluímos que $\eta_1(B) \leq C_2\eta_2(B)$ para todo conjunto mensurável $B \subset M$. Permutando os papéis das duas medidas e repetindo o último argumento, também obtemos que $\eta_2(B) \leq C_2\eta_1(B)$ para todo conjunto mensurável $B \subset M$. Portanto, $\leq C_2{}^{-1}\eta_2(B) \leq \eta_1(B) \leq C_2\eta_2(B)$ para todo conjunto mensurável $B \subset M$. Isto mostra que quaisquer dois estados de equilíbrio de $\varphi$ são equivalentes. □

**Corolário 4.38.** *A probabilidade $\mu = h\nu$ é o único estado de equilíbrio para $\varphi$. Além disso, a medida de referência $\nu$ também é única.*

*Demonstração.* Sabemos que medidas ergódicas e equivalentes são iguais. Juntando este fato ao Lema 4.37, obtemos que todos os estados de equilíbrio ergódicos são iguais. Por outro lado, como vimos no Teorema 3.63 e no Corolário 3.64, as componentes ergódicas de um estado de equilíbrio são estados de equilíbrio ergódicos. Segue que existe um único estado de equilíbrio (igual a $\mu$), tal como afirmamos.

Como consequência, a medida de referência $\nu$ também é única: se existissem duas medidas de referência distintas $\nu_1$ e $\nu_2$ então $\mu_1 = h\nu_1$ e $\mu_2 = h\nu_2$ seriam estados de equilíbrio distintos. Analogamente, a autofunção positiva $h$ é única a menos de produto por constante positiva. □

**Passo 5:** Mostraremos que o sistema $(f, \mu)$ é exato (Veja a definição 4.15 ).

Seja $B \subset M$ tal que existem conjuntos mensuráveis $B_n$ satisfazendo $B = f^{-n}(B_n)$ para todo $n \geq 1$ e suponha que $\mu(B) > 0$. Provaremos que $\mu(B) = 1$. Considere uma partição finita $\mathscr{P}$ de $M$ tal que todo $P \in \mathscr{P}$ tem interior não vazio e diâmetro menor que $\rho$. Para cada $n \geq 1$, seja $\mathscr{Q}_n$ a partição de $M$ cujos elementos são as imagens $h^n(P)$ dos conjuntos $P \in \mathscr{P}$ pelos ramos inversos $h^n$ do iterado $f^n$.

**Lema 4.39.** *Para todo $\varepsilon > 0$ e todo $n \geq 1$ suficientemente grande existe algum $h^n(P) \in \mathscr{Q}_n$ tal que*

$$\mu(B \cap h^n(P)) > (1 - \varepsilon)\mu(h^n(P)). \tag{4.34}$$

*Demonstração.* Fixe $\varepsilon > 0$. Como a medida $\mu$ é regular, dado qualquer $\delta > 0$ existe algum compacto $F \subset B$ e algum aberto $A \supset B$ satisfazendo $\mu(A \setminus F) < \delta$. Como supomos que $\mu(B) > 0$, esta desigualdade implica que $\mu(F) > (1 - \varepsilon)\mu(A)$, desde que $\delta$

seja suficientemente pequeno. Fixe $\delta > 0$ nessas condições. Note que diam $\mathscr{Q}_n < \sigma^{-n}\rho$ (vimos isso na prova do Lema 4.37). Então, para todo $n$ suficientemente grande, qualquer elemento $h^n(P)$ de $\mathscr{Q}_n$ que intersecta $F$ está contido em $A$. Suponha que (4.34) fosse falsa para todo $h^n(P)$. Então, somando sobre todos $h^n(P)$ que intersectam $F$,

$$\mu(F) \leq \sum_{h^n(P)} \mu(F \cap h^n(P)) \leq \sum_{h^n(P)} \mu(B \cap h^n(P)) \leq (1-\varepsilon) \sum_{h^n(P)} \mu(h^n(P)) \leq (1-\varepsilon)\mu(A).$$

Esta contradição prova que (4.34) é válida para algum $h^n(P) \in \mathscr{Q}_n$. $\square$

Considere qualquer $h^n(P) \in \mathscr{Q}_n$ satisfazendo (4.34). Como $B = f^{-n}(B_n)$ e $f^n \circ h^n = id$ no seu domínio, temos que $f^n(h^n(P) \setminus B) = P \setminus B_n$. Então, aplicando o Corolário 4.36 à medida $\eta = \mu$,

$$\mu(P \setminus Bn) = \int_{h^n(P)\setminus B} J\mu f^n d\mu \leq K_7 \mu(h^n(P) \setminus B) J_\mu f^n(x) \tag{4.35}$$

e

$$\mu(P) = \int_{h^n(P)} J_\mu f^n d\mu \geq {K_7}^{-1} \mu(h^n(P)) J_\mu f^n(x) \tag{4.36}$$

para qualquer $x \in h^n(P)$. Note que a equação (4.34) pode ser escrita como $\mu(h^n(P)) - \mu(B \cap h^n(P)) < \varepsilon \mu(h^n(P))$ e que o lado esquerdo desta última desigualdade é precisamente $\mu(h^n(P) \setminus B)$. Então, combinando esta última desigualdade com (4.35) e (4.36), obtemos

$$\frac{\mu(P \setminus B_n)}{\mu(P)} \leq {K_7}^2 \frac{\mu(h^n(P) \setminus B)}{\mu(h^n(P))} \leq {K_7}^2 \frac{\varepsilon \mu(h^n(P))}{\mu(h^n(P))} = {K_7}^2 \varepsilon.$$

Resumindo, mostramos que, dado qualquer $\varepsilon > 0$ e qualquer $n \geq 1$ suficientemente grande existe algum $P \in \mathscr{P}$ tal que $\mu(P \setminus B_n) \leq {K_7}^2 \varepsilon \mu(P)$.

Como a partição $\mathscr{P}$ é finita, segue que existe algum $P \in \mathscr{P}$ e alguma sequência $(n_j)_j \to \infty$ tal que

$$\mu(P \setminus B_{n_j}) \to 0 \quad \text{quando} \quad j \to \infty. \tag{4.37}$$

Fixemos $P$ a partir daqui. Como $P$ tem interior não vazio e $f$ é topologicamente exata, por hipótese, existe $N \geq 1$ tal que $f^N(P) = M$. Seja $P = P_1 \cup \cdots \cup P_s$ uma partição finita de $P$ em domínios de invertibilidade de $f^N$. O Corolário 4.27 e o Lema 4.35 dão que $J_\mu f^N = \lambda^N e^{-\varphi_N}(h \circ f^N)/h$ é uma função limitada de zero e infinito. Note também que $f^N(P_i \setminus B_{n_j}) = f^N(P_i) \setminus B_{n_j+N}$ , uma vez que $f^{-n}(B_n) = B$ para todo $n$. Combinando estas duas observações com (4.37), obtemos que, dado qualquer $i = 1, \cdots, s$,

$$\mu(f^N(P_i) \setminus B_{n_j+N}) = \mu(f^N(P_i \setminus B_{n_j})) = \int_{P_i \setminus B_{n_j}} J_\mu f^N d\mu$$

converge para zero quando $j \to \infty$. Agora, $\{f^N(P_i) : i = 1, \cdots, s\}$ é uma cobertura finita de $M$ por conjuntos mensuráveis. Portanto, esta última conclusão implica que $\mu(M \setminus B_{n_j+N})$ converge para zero, ou seja, que $\mu(B) = \mu(B_{n_j+N})$ converge para 1 quando $j \to \infty$. Isto significa que $\mu(B) = 1$.

Com isso, a demonstração do Teorema de Ruelle (Teorema 4.17) está completa. $\square$

## 4.3 Aplicação do Teorema de Ruelle: medidas absolutamente contínuas

**Definição 4.40.** Dada uma probabilidade invariante $\mu$, chamamos bacia de $\mu$ o conjunto $B(\mu)$ dos pontos $x \in M$ tais que

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

para toda função contínua $\varphi : M \to \mathbb{R}$. É possível provar que a bacia é sempre um conjunto invariante. Além disso, se $\mu$ é ergódica então $B(\mu)$ tem $\mu$-medida total.

O que discutiremos nesta seção é motivado pelo seguinte resultado:

**Teorema 4.41.** *Seja $f : M \to M$ uma transformação diferenciável expansora numa variedade compacta conexa $M$. Suponha que o jacobiano $x \mapsto \det Df(x)$ é Hölder. Então $f$ admite uma única probabilidade invariante $\mu$ absolutamente contínua com relação à medida de Lebesgue $m$. Além disso, $\mu$ é ergódica, seu suporte coincide com $M$ e a sua bacia tem medida de Lebesgue total na variedade.*

A demonstração do teorema acima pode ser encontrada na referência [7]. No teorema abaixo, mostramos que a única probabilidade invariante obtida no teorema anterior é na realidade o único estado de equilíbrio dado pelo Teorema de Ruelle.

**Teorema 4.42.** *A probabilidade absolutamente contínua invariante de $f$ coincide com o estado de equilíbrio $\mu$ do potencial $\varphi = -\log|\det Df|$. Consequentemente, ela é equivalente à medida de Lebesgue $m$, com densidade $\frac{d\mu}{dm}$ Hölder e limitada de zero e infinito, e ela é exata.*

*Demonstração.* Primeiramente, devemos notar que o se $\det Df$ é Hölder, então logaritmo $\varphi = -\log|\det Df|$ também é Hölder: digamos que $C_0 > 0$ e $\nu > 0$ são as constantes da definição Hölder de $\det Df$. Então

$$|\log|\det Df(x)| - \log|\det Df(y)|| \leq |\det Df(x) - \det Df(y)| \leq C_0 d(x,y)^\nu$$

para quaisquer $x, y \in M$.

Considere o operador de transferência $\mathcal{L}$ associado ao potencial $\varphi = -\log|\det Df|$. A medida de Lebesgue $m$ de $M$ é uma automedida do operador dual do operador de transferência, correspondente ao autovalor $\lambda = 1$, ou seja

$$\mathcal{L}^* m = m.$$

Para verificar esse fato, basta mostrar que $\mathcal{L}^* m(E) = m(E)$ para todo conjunto mensurável $E$ contido na imagem de uma bola $B(y,\rho)$ por algum ramo inverso $h_i : B(y,\rho) \to M$, $i = 1, \cdots, k$ (pois, pela compacidade de $M$, todo conjunto mensurável pode ser escrito como união finita disjunta de subconjuntos $E$ deste tipo). Usando a definição do operador de transferência, escrevemos,

$$\begin{aligned}
\mathcal{L}^* m(E) &= \int \chi_E d(\mathcal{L}^* m) \\
&= \int (\mathcal{L}\chi_E) dm \\
&= \int \sum_{i=1}^{k} (e^{-\log|\det Df|}\chi_E) \circ h_i dm \\
&= \int \sum_{i=1}^{k} \frac{\chi_E}{|\det Df|} \circ h_i dm.
\end{aligned}$$

Então, pela escolha de $E$ e pela Fórmula de mudança de variáveis do Lema 3.35 item (b),

$$\mathcal{L}^* m(E) = \int \sum_{i=1}^{k} \frac{\chi_E}{|\det Df|} \circ h_i dm. = \int \left(\frac{\chi_E}{J_m f}\right) \circ (f)^{-1} dm = \int \chi_E dm = m(E).$$

Isto prova que $m$ é, de fato, um ponto fixo de $\mathcal{L}^*$.

Agora basta aplicar a teoria anterior (do Lema 4.20 em diante) para $\lambda = 1$ e $\nu = m$. Encontraremos uma função Hölder $h : M \to \mathbb{R}$, limitada de zero e infinito, tal que $\mathcal{L}h = h$ e a medida $\mu = hm$ é o único estado de equilíbrio para potencial $\varphi$. Como a relação $\mu = hm$ diz que $\mu$ é absolutamente contínua com relação a $m$, em vista do Teorema 4.41, segue que $\mu$ também é a única probabilidade invariante de $f$ absolutamente contínua com relação a $m$. O fato de que $h = \dfrac{d\mu}{dm}$ é positiva implica que $\mu$ e $m$ são equivalentes. A exatidão foi o último passo da prova do Teorema de Ruelle. □

## 4.4 Teorema de Livsič

Foi visto no Teorema de Ruelle que se $f$ é expansora e topologicamente exata num espaço métrico compacto e $\varphi$ é um potencial Hölder, então existe um único estado de equilíbrio $\mu$ para $\varphi$. A partir de agora, denotaremos este estado de equilíbrio por $\mu_\varphi$. Nesta seção vamos discutir a seguinte questão: quando é que os estados de equilíbrio $\mu_\phi$ e $\mu_\psi$ de dois potenciais Hölder $\phi$ e $\psi$ são iguais?

**Definição 4.43.** Dizemos que dois potenciais $\phi,\ \psi : M \to \mathbb{R}$ são **cohomólogos** se existe alguma função contínua $u : M \to \mathbb{R}$ tal que $\phi - \psi = u \circ f - u$. Em particular, dizemos que um potencial $\phi$ é cohomólogo a uma constante $c \in \mathbb{R}$, quando $\phi - c = u \circ f - u$.

**Definição 4.44.** Uma transformação $f : M \to M$ definida num espaço topológico $M$ é dita **topologicamente misturadora** se dados quaisquer abertos não-vazios $U, V \subset M$, existe $n_0 \in \mathbb{N}$ tal que $f^{-n}(U) \cap V$ é não-vazio para todo $n \geq n_0$.

**Observação 4.45.** *Se uma transformação expansora num espaço métrico compacto é topologicamente misturadora então ela é topologicamente exata.*

**Teorema 4.46.** *Um potencial Hölder $\varphi : M \to \mathbb{R}$ é cohomólogo a zero se, e somente se, $\varphi_n(x) = \sum_{i=0}^{n-1} \varphi(f^i(x)) = 0$ para todo $x \in \mathrm{Fix}(f^n)$ e $n \geq 1$.*

*Demonstração.* Se $\varphi - 0 = u \circ f - u$ para algum $u$ e $x \in \mathrm{Fix}(f^n)$ é qualquer, então

$$\begin{aligned}\varphi_n(x) &= \sum_{i=0}^{n-1}[u(f^{i+1}(x)) - u(f^i(x))]\\ &= \sum_{i=1}^{n} u(f^i(x)) - \sum_{i=0}^{n-1} u(f^i(x))\\ &= \sum_{i=0}^{n-1} u(f^i(x)) - \sum_{i=0}^{n-1} u(f^i(x)) = 0.\end{aligned}$$

A igualdade das duas somas na última linha, vem de ser $f^n(x) = x$. Agora a recíproca.

Suponha que $\varphi_n(x) = 0$ para todo $x \in \mathrm{Fix}(f^n)$ e todo $n \geq 1$. Considere qualquer ponto $z \in M$ cuja órbita é densa em $M$; tal ponto existe porque $f$ é topologicamente misturadora e, consequentemente, transitiva. Defina $u$ na órbita de $z$ por meio da relação

$$u(f^n(z)) = u(z) + \varphi_n(z), \tag{4.38}$$

onde $u(z)$ é arbitrário. Observe que

$$u(f^{n+1}(z)) - u(f^n(z)) = \varphi_{n+1}(z) - \varphi_n(z) = \varphi(f^n(z)) \tag{4.39}$$

para todo $n \geq 0$, ou seja

$$u(f(f^n(z))) - u(f^n(z)) = \varphi(f^n(z)) - 0.$$

Em outras palavras, o potencial $\varphi$ é cohomólogo a zero sobre a órbita de $z$. Para estender esta relação a $M$, usaremos o seguinte fato:

**Lema 4.47.** *A função $u$ é uniformemente contínua na órbita de $z$.*

*Demonstração.* Dado qualquer $\varepsilon \in (0, \rho)$, tome $\delta > 0$ dado pelo lema de sombreamento (Lema 4.8). Suponha que $k \geq 0$ e $l \geq 1$ são tais que $d(f^k(z), f^{k+l}(z)) < \delta$. Então a sequência periódica $(x_n)_n$ de período $l$ dada por

$$x_0 = f^k(z),\, x_1 = f^{k+1}(z), \cdots,\, x^{l-1} = f^{k+l-1}(z),\, x_l = f^k(z)$$

é uma $\delta$-pseudo-órbita. Logo, pelo lema de sombreamento, existe algum $x \in \mathrm{Fix}(f^l)$ tal que $d(f^j(x), f^{k+j}(z)) < \varepsilon$ para todo $j \geq 0$. Como tomamos $\varepsilon < \rho$, isto também implica que $x = h_l(f^l(x))$, onde $h_l : B(f^{k+l}(z), \rho) \to M$ representa o ramo inverso de $f^l$ que envia $f^{k+l}(z)$ em $f^k(z)$. Pela desigualdade dada em (4.3), segue que

$$d(f^j(x), f^{k+j}(z)) \le \sigma^{j-l} d(f^l(x), f^{k+l}(z)) \text{ para todo } 0 \le j \le l. \tag{4.40}$$

Pela definição (4.38),

$$u(f^{k+l}(z)) - u(f^k(z)) = \varphi_{k+l}(z) - \varphi_k(z) = \varphi_l(f^k(z)). \tag{4.41}$$

Fixe constantes $C > 0$ e $\nu > 0$ tais que $|\phi(x) - \phi(y)| \le C d(x, y)^\nu$ para quaisquer $x, y \in M$. Então,

$$|\varphi_l(f^k(z)) - \varphi_l(x)| \le \sum_{j=0} l - 1 |\varphi(f^{k+j}(z)) - \varphi(f^j(x))| \le \sum_{j=0} l - 1 C d(f^j(x), f^{k+j}(z))^\nu.$$

Usando (4.40), segue que

$$|\varphi_l(f^k(z)) - \varphi_l(x)| \le \sum_{j=0} l - 1 C \sigma^{\nu(j-l)} d(f^l(x), f^{k+l}(z))^\nu \le C_1 \varepsilon^\nu \tag{4.42}$$

onde $C_1 = \sum_{i=0}^{\infty} \sigma^{-i\nu}$. Lembre também que, por hipótese, $\varphi_l(x) = 0$. Logo, combinando (4.41) e (4.42), obtemos que $|u(f^{k+l}(z)) - u(f^k(z))| \le C_1 \varepsilon^\nu$. Isto conclui a prova do lema. $\square$

Segue do Lema 4.47 que $u$ admite uma única extensão ao fecho da órbita de $z$, ou seja, ao espaço ambiente $M$. Então, pela continuidade de $\varphi$ e $u$, a relação de cohomologia se estende a todo o $M$. Isto prova o Teorema 4.46. $\square$

O seguinte Teorema, responde a questão dada no início da seção.

**Teorema 4.48 (Livsič).** *Seja $f : M \to M$ uma transformação expansora topologicamente misturadora num espaço métrico compacto e sejam $\phi$ e $\psi$ dois potenciais Hölder em $M$. As seguintes condições são equivalentes:*

*(a)* $\mu_\phi = \mu_\psi$*;*

*(b) Existe $c \in \mathbb{R}$ e uma função qualquer $u : M \to \mathbb{R}$ tal que $\phi - \psi = c + u \circ f - u$;*

*(c) O potencial $\phi - \psi$ é cohomólogo a alguma constante $c \in \mathbb{R}$;*

*(d) Existe* $c \in \mathbb{R}$ *e uma função Hölder* $u : M \to \mathbb{R}$ *tal que* $\phi - \psi = c + u \circ f - u$*;*

*(e) Existe* $c \in \mathbb{R}$ *tal que* $\phi_n(x) - \psi_n(x) = \sum_{i=0}^{n-1} \phi(f^i(x)) - \sum_{i=0}^{n-1} \psi(f^i(x)) = cn$ *para todo* $x \in \mathrm{Fix}(f^n)$ *e* $n \geq 0$ *;*

*Além disso, as constantes c em (b), (c), (d) e (e) coincidem.*

*Demonstração.* É imediato da definição 4.43 que (d) implica (c) e (c) implica (b). Supondo que (b) seja verdadeira, temos que $\phi - \psi = c + u \circ f - u$. Então, dado $x \in \mathrm{Fix}(f^n)$ qualquer, vale que $f^n(x) = x$ e

$$\begin{aligned}
\phi_n(x) - \psi_n(x) &= \sum_{i=0}^{n-1} \phi(f^i(x)) - \sum_{i=0}^{n-1} \psi(f^i(x)) = \sum_{i=0}^{n-1} (\phi - \psi)(f^i(x)) \\
&= \sum_{i=0}^{n-1} (c + u \circ f - u)(f^i(x)) = cn + \sum_{i=0}^{n-1} (u \circ f - u)(f^i(x)) \\
&= cn + u(f^n(x)) - u(x) = cn + u(x) - u(x) = cn.
\end{aligned}$$

Portanto, $\phi_n(x) - \psi_n(x) = cn$. Isto prova que (b) implica (e).

Supondo que (e) seja verdadeira, temos que $\phi_n(x) - \psi_n(x) = cn$ para todo $x \in \mathrm{Fix}(f^n)$ e todo $n \geq 0$. Isto significa que a função $\varphi = \phi - \psi - c$ satisfaz $\varphi_n = \phi_n(x) - \psi_n(x) - cn = 0$ para todo $x \in \mathrm{Fix}(f^n)$ e todo $n \geq 0$. Note também que $\varphi$ é Hölder: Se $\beta_1, \beta_2, \alpha_1, \alpha_2$ são constantes positivas fixas com $\alpha_1 \geq \alpha_2$ tais que $|\phi(z) - \phi(w)| \leq \beta_1 d(z,w)^{\alpha_1}$ e $|\psi(z) - \psi(w)| \leq \beta_2 d(z,w)^{\alpha_2}$ para quaisquer $z, w \in M$, então $|\varphi(z) - \varphi(w)| \leq (\beta_1 d(z,w)^{\alpha_1 - \alpha_2} + \beta_2) d(z,w)^{\alpha_2}$ . Logo, pelo Teorema 4.46, o potencial Hölder $\varphi : M \to \mathbb{R}$ é cohomólogo a zero. Em outras palavras, $\phi - \psi$ é cohomólogo a $c$. Isto mostra que (e) implica (c).

Se (c) é verdadeira, então $\phi - \psi$ é cohomólogo a $c$. Logo $\phi$ é cohomólogo a $\psi + c$, e nesse caso, segue do princípio variacional que

$$\begin{aligned}
P(f, \psi + c) &= \sup\{h_\nu(f) + \int (\psi + c) d\nu : \nu \in \mathcal{M}_1(f)\} \\
&= \sup\{h_\nu(f) + \int \phi d\nu + \int (\psi + c - \phi) d\nu : \nu \in \mathcal{M}_1(f)\} \\
&= \sup\{h_\nu(f) + \int \phi d\nu + \int (u \circ f - u) d\nu : \nu \in \mathcal{M}_1(f)\} \\
&= \sup\{h_\nu(f) + \int \phi d\nu : \nu \in \mathcal{M}_1(f)\} = P(f, \phi)
\end{aligned} \tag{4.43}$$

e segue da equação (3.54) na observação 3.53 que $P(f, \psi + c) = P(f, \psi) + c$. Juntando isso a (4.43), resulta que

$$P(f, \phi) = P(f, \psi) + c. \tag{4.44}$$

Por outro lado, dada qualquer probabilidade invariante $\nu$, repetimos os cálculos feitos dentro do supremo em (4.43), e obtemos

$$h_\nu(f) + \int \phi d\nu = h_\nu(f) + \int (\psi + c) d\nu = h_\nu(f) + \int \psi d\nu + c. \tag{4.45}$$

Das igualdades em (4.44) e (4.45) fica claro que $\nu$ é estado de equilíbrio para $\phi$ se, e somente se, $\nu$ é estado de equilíbrio para $\psi$. Isto mostra que (c) implica (a).

Se (a) é verdadeira, então $\mu_\phi = \mu_\psi$ e isto implica que essas medidas têm o mesmo jacobiano. Em vista do Lema 4.31, isto quer dizer que

$$\lambda_\phi e^{-\phi} \frac{(h_\phi \circ f)}{h_\phi} = \lambda_\psi e^{-\psi} \frac{(h_\psi \circ f)}{h_\psi}. \tag{4.46}$$

Seja $c = \log \lambda_\phi - \log \lambda_\psi$ e seja $u = \log h_\phi - \log h_\psi$. Estes objetos estão bem definidos, uma vez que $\lambda_\phi, \lambda_\psi, h_\phi$ e $h_\psi$ são positivos. Além disso, como as funções $h_\phi$ e $h_\psi$ são Hölder e limitadas de zero e de infinito (Lema 4.30), segue que a função $u$ é Hölder.

Aplicando o log em ambos os membros de (4.46),

$$\begin{aligned} \log \lambda_\phi - \phi + \log h_\phi \circ f - \log h_\phi = \log \lambda_\psi - \psi + \log h_\psi \circ f - \log h_\psi \Rightarrow \\ \log \lambda_\phi - \log \lambda_\psi + \log h_\phi \circ f - \log h_\psi \circ f + \log h_\psi - \log h_\phi = \phi - \psi. \end{aligned} \tag{4.47}$$

Finalmente, de (4.47), podemos reescrever (4.46) na seguinte forma:

$$c + \log u \circ f - u = \phi - \psi.$$

Isto mostra que (a) implica (d). A prova do teorema está completa. $\square$

# 5 PARTIÇÕES DE MARKOV, DINÂMICA SIMBÓLICA E TEOREMA DE RUELLE PARA DIFEOMORFISMOS AXIOMA A

## 5.1 Dinâmica simbólica e o Teorema de Ruelle-Perron-Frobenius

Primeiramente, considere o espaço produto

$$\Sigma_n = \{1, 2, \ldots, n\}^{\mathbb{Z}} = \prod_{\mathbb{Z}} \{1, 2, \ldots, n\}$$

formado por todas as sequências bilaterais $(\ldots, x_{-i}, \ldots, x_0, \ldots, x_i, \ldots)$ com $x_i \in \{1, 2, \ldots, n\}$ para todo $i \in \mathbb{Z}$. Uma sequência em $\Sigma_n$ será denotada simplesmente por $(x_n)_n$, onde $(x_n)_n = \{x_i\}_{i=-\infty}^{\infty}$. Definimos o deslocamento (ou "shift") bilateral $\sigma : \Sigma_n \to \Sigma_n$ por

$$\sigma((x_n)_n) = \sigma(\{x_i\}_{i=-\infty}^{\infty}) = \{x_{i+1}\}_{i=-\infty}^{\infty} = (x_{n+1})_n.$$

Nesse caso, $\sigma$ é invertível. Equipando $\{1, 2, \ldots, n\}$ com a topologia discreta e $\Sigma_n$ com a topologia produto, resulta que $\sigma$ e $\sigma^{-1}$ são aplicações contínuas. Nesse caso $\sigma$ é um homeomorfismo. (Estas noções foram introduzidas na definição 2.12).

**Definição 5.1.** Para $\phi : \Sigma_n \to \mathbb{R}$ contínua definimos

$$\operatorname{var}_k \phi = \sup\{|\phi((x_n)_n) - \phi((y_n)_n)| \ : \ x_i = y_i \ \text{ para todo } \ i \ \text{ com } \ |i| \leq k\}.$$

Note que $\Sigma_n$ é compacto: De fato $\{1, 2, \ldots, n\}$ é um conjunto finito equipado com a topologia discreta, portanto é compacto. O Teorema de Tychonoff afirma que afirma que o produto de qualquer coleção de espaços topológicos compactos é compacto com respeito à topologia produto. Logo $\Sigma_n$ é compacto. Segue que $\phi$ é uniformemente contínua, e assim $\lim_{k\to\infty} \operatorname{var}_k \phi = 0$.

**Observação 5.2.** *Podemos munir $\Sigma_n$ com a distância*

$$d((x_n)_n, (y_n)_n) = \beta^N, \tag{5.1}$$

*onde $\beta \in (0,1)$ é qualquer e $N$ é o maior inteiro não negativo tal que $x_i = y_i$ para cada $i$ com $|i| < N$ e tornar $\Sigma_n$ um espaço métrico compacto.*

Seja $A = (A_{i,j})_{i,j}$ uma matriz de transição (lembre do exemplo 3.48). Considere o subconjunto $\Sigma_A$ de $\Sigma_n$ das sequências $(x_i)_i$ que são $A$-admissíveis, ou seja, tais que

$$A_{x_i, x_{i+1}} = 1 \ \text{ para todo } \ i \in \mathbb{Z}. \tag{5.2}$$

A restrição do deslocamento $\sigma_A : \Sigma_A \to \Sigma_A$ é chamado de **deslocamento bilateral de tipo finito** associado a $A$.

Observe que $\Sigma_A$ é invariante e fechado (no sentido de $\sigma_A(\Sigma_A) = \Sigma_A$). Então, sendo um subconjunto fechado do espaço métrico compacto $\Sigma_n$, segue que $\Sigma_A$ é também compacto.

**Definição 5.3.** Definiremos $\mathcal{F}_A$ como sendo a família de todas as funções contínuas $\phi : \Sigma_A \to \mathbb{R}$ para as quais $\operatorname{var}_k \phi \leq b\alpha^k$ para todo $k \geq 0$, para alguma constante positiva $b$ e $\alpha \in (0,1)$.

Note que se $\Sigma_A$ estiver munido com a distância dada na observação 5.2, então para qualquer $\phi \in \mathcal{F}_A$ e algum $k \geq 0$

$$|\phi((x_n)_n) - \phi((y_n)_n)| \leq \operatorname{var}_k \phi \leq b\alpha^k = b\left(\frac{\alpha}{\beta}\right)^k \beta^k = b\left(\frac{\alpha}{\beta}\right)^k d((x_n)_n, (y_n)_n)^k.$$

Então $\mathcal{F}_A$ é justamente o conjunto das funções que têm expoente de Hölder positivo com respeito à distância $d$. De fato, se $\phi$ é Hölder, então $|\phi((x_n)_n) - \phi((y_n)_n)| \leq Cd((x_n)_n, (y_n)_n)^\theta = C\beta^{\theta k}$ e portanto $\operatorname{var}_k \phi \leq b\alpha^k$, com $b = C$ e $\alpha = \beta^\theta$.

Denotaremos a partir de agora $\Sigma_A^+$ como o conjunto das sequências $\{x_{i+1}\}_{i=0}^\infty \in \{1, 2, \ldots, n\}^{\mathbb{N}} = \Sigma_n^+$ que são $A$-admissíveis, ou seja, tais que

$$A_{x_i, x_{i+1}} = 1 \text{ para todo } i \in \mathbb{N}. \tag{5.3}$$

Também denotaremos as sequências de $\Sigma_A^+$ por $(x_n)_n = \{x_{i+1}\}_{i=0}^\infty$ (mas sempre ficará explícito se esta sequência está em $\Sigma_A^+$ ou em $\Sigma_A$). A restrição do deslocamento $\sigma_A^+ : \Sigma_A^+ \to \Sigma_A^+$ é chamado de **deslocamento unilateral de tipo finito** associado a $A$. Geralmente chamamos a restrição dos deslocamentos $\sigma_A$ ou $\sigma_A^+$ a quaisquer subconjuntos fechados e invariantes, respectivamente de $\Sigma_n$ ou $\Sigma_n^+$ de **sistema dinâmico simbólico**.

Esta seção e a próxima serão dedicadas a provar o seguinte Teorema, que nada mais é que uma versão do Teorema de Ruelle do capítulo anterior para deslocamentos.

**Teorema 5.4** (**Existência e Unicidade de estados de equilíbrio**). *Suponha que $\sigma_A : \Sigma_A \to \Sigma_A$ é topologicamente misturador e $\phi \in \mathcal{F}_A$. Existe uma única medida de probabilidade $\mu$ em $\Sigma_A$ que é $\sigma_A$-invariante e é um estado de Gibbs, ou seja, existem constantes $c_1 > 0$, $c_2 > 0$ e $P$ tais que*

$$c_1 \leq \frac{\mu(\{(y_n)_n \in \Sigma_A : x_i = y_i \text{ para todo } i = 0, \ldots, m-1\})}{\exp(-mP + \phi_m((x_n)_n))} \leq c_2.$$

*Além disso, $\mu$ é também o único estado de equilíbrio para $\phi \in \mathcal{F}_A$.*

O primeiro passo da demonstração do Teorema 5.4 consiste em provar o importante Teorema de Ruelle-Perron-Frobenius. Ele dará a medida de referência $\nu$, o autovalor $\lambda$ e a autofunção $h$, semelhante ao que ocorreu nos Lemas 4.18, 4.30 da prova do Teorema 4.17. A automedida $\nu$ e a autofunção $h$ serão objetos essenciais para o prosseguimento da demonstração do Teorema 5.4 que só acontecerá na próxima seção, pois durante o restante desta, estaremos provando o Teorema de Ruelle-Perron-Frobenius.

**Teorema 5.5** (**Ruelle-Perron-Frobenius**). *Suponha que $\sigma_A : \Sigma_A \to \Sigma_A$ é topologicamente misturador (lembre da definição 4.44), $\phi \in \mathcal{F}_A \cap \mathcal{C}(\Sigma_A^+)$ onde $\mathcal{C}(\Sigma_A^+)$*

*denota o espaço de todas as funções contínuas* $\phi : \Sigma_A^+ \to \mathbb{R}$ *e* $\mathcal{L} = \mathcal{L}_\phi : \mathcal{C}(\Sigma_A^+) \to \mathcal{C}(\Sigma_A^+)$ *é o operador de transferência como definimos em* (4.4)*:*

$$\mathcal{L}g((y_n)_n) = \mathcal{L}_\phi g((y_n)_n) = \sum_{(x_n)_n \in (\sigma_A^+)^{-1}((y_n)_n)} e^{\phi((x_n)_n)} g((x_n)_n). \tag{5.4}$$

*Então existem* $\lambda > 0$*,* $h \in \mathcal{C}(\Sigma_A^+)$ *com* $h > 0$ *e* $\nu \in \mathcal{M}(\Sigma_A^+)$*, onde* $\mathcal{M}(\Sigma_A^+)$ *é o conjunto das medidas de probabilidade em* $\Sigma_A^+$ *satisfazendo* $\mathcal{L}h = \lambda h$*,* $\mathcal{L}^*\nu = \lambda\nu$*,* $\int hd\nu = 1$ *e*

$$\lim_{m\to\infty} ||\lambda^{-m}\mathcal{L}^m g - h\int gd\nu|| = 0 \;\; \textit{para toda} \;\; g \in \mathcal{C}(\Sigma_A^+).$$

*Demonstração.* Começaremos provando que existe $\nu$ tal que $\mathcal{L}^*\nu = \lambda\nu$.

**Lema 5.6.** *Existe uma medida* $\nu \in \mathcal{M}(\Sigma_A^+)$ *tal que* $\mathcal{L}^*\nu = \lambda\nu$*. Chamaremos essa tal medida de medida de referência, como fizemos na definição 4.19.*

*Demonstração.* Vimos na definição 4.9 que o operador $\mathcal{L}$ é positivo, portanto $\mathcal{L}1 > 0$. Então temos que

$$G(\mu) = \frac{\mathcal{L}^*\mu}{\mathcal{L}^*\mu(1)} = \frac{\mathcal{L}^*\mu}{\int 1d(\mathcal{L}^*\mu)} \in \mathcal{M}(\Sigma_A^+) \;\text{ para toda }\; \mu \in \mathcal{M}(\Sigma_A^+).$$

O operador $G(\mu)$ é contínuo na topologia fraca*: Seja $\varepsilon > 0$ e $\Phi = \{\phi_1, \ldots, \phi_n\}$ uma família qualquer de funções contínuas limitadas. Como $\mathcal{L}$ é contínuo (veja a definição 4.9), a família $\Theta = \{\mathcal{L}\phi_1, \ldots, \mathcal{L}\phi_n\}$ também consiste de funções contínuas limitadas. Então pela definição do operador dual do operador de transferência $\mathcal{L}^*$ (veja a definição 4.10),

$$\left|\int \phi_i d(\mathcal{L}^*\eta) - \int \phi_i d(\mathcal{L}^*\zeta)\right| = \left|\int (\mathcal{L}\phi_i)d\eta - \int (\mathcal{L}\phi_i)d\zeta\right|$$

e portanto o lado esquerdo é menor que $\varepsilon$ se o lado direito for menor que $\varepsilon$. Isto quer dizer que

$$\begin{aligned}\mathcal{L}^*&\left(\{\zeta \in \mathcal{M}(\Sigma_A^+) : \left|\int (\mathcal{L}\phi_i)d\eta - \int (\mathcal{L}\phi_i)d\zeta\right| < \varepsilon\}\right)\\ &\subset \{\mathcal{L}^*\zeta \in \mathcal{M}(\Sigma_A^+) : \left|\int \phi_i d(\mathcal{L}^*\eta) - \int \phi_i d(\mathcal{L}^*\zeta)\right| < \varepsilon\} \;\text{ para todo }\; \eta, \Phi \;\text{ e }\; \varepsilon\end{aligned}$$

e este último fato mostra que $\mathcal{L}^*$ é contínuo. Consequentemente $G(\mu) : \mathcal{M}(\Sigma_A^+) \to \mathcal{M}(\Sigma_A^+)$ é contínuo. Sabemos que $\Sigma_A^+$ é um espaço métrico compacto. Então $\mathcal{M}(\Sigma_A^+)$ munido da topologia fraca* é compacto e além disso é convexo. Podemos usar o seguinte teorema:

**Teorema 5.7 (Schauder-Tychonoff).** *Seja $F : V \to V$ uma transformação contínua num espaço vetorial topológico $V$. Suponha que existe um conjunto compacto convexo $K \subset V$ tal que $F(K) \subset K$. Então $F(v) = v$ para algum $v \in K$.*

Então concluímos que existe uma automedida $\nu$ tal que $G(\nu) = \nu$. Então se $\lambda = \int 1 d(\mathcal{L}^*\nu) > 0$, vale que $\mathcal{L}^*\nu = \lambda G(\nu) = \lambda\nu$. □

Provaremos o Teorema 5.5 por meio da sequência dos lemas seguintes. Sejam $b > 0$ e $\alpha \in (0,1)$ constantes quaisquer de modo que $\operatorname{var}_k \phi \leq b\alpha^k$ para todo $k \geq 0$. Ponha $B_m = \exp(\sum_{k=m+1}^{\infty} 2b\alpha^k)$ e defina o espaço de funções

$$\Lambda = \{f \in \mathcal{C}(\Sigma_A^+) : f \geq 0, \int f d\nu = 1, \; f((x_n)_n) \leq B_m f((\bar{x}_n)_n) \\ \text{se } x_i = \bar{x}_i \text{ para todo } i \in \{0, \dots, m\}\}$$

**Lema 5.8.** *Se $A$ é uma matriz de transição então, para todo $M \geq 1$, nenhuma linha de $A^M$ é identicamente nula.*

*Demonstração.* Seja $S_i = A_{i,1} + A_{i,2} + \cdots + A_{i,n}$ a soma das entradas da linha $i$ da matriz $A$. Por hipótese $S_i > 0$ para todo $i$. Defina $P_{i,j} = A_{i,j}/S_i$. Note que $P = P_{i,j}$ é uma matriz estocástica: De fato $P_{i,j} = 0$ se, e somente se, $A_{i,j} = 0$. Logo $P_{i,j} \geq 0$ e $\sum_{j=1}^{n} P_{i,j} = \sum_{j=1}^{n} A_{i,j}/S_i = S_i/S_i = 1$. Pelo Lema 2.17 item (a), temos que $\sum_{j=1}^{n} P_{i,j}^M = 1$ para todo $1 \leq i \leq n$ e todo $M \geq 1$. Resulta disso que a linha $i$ de $A^M$ é não nula, para todo $i$. □

**Observação 5.9.** *O Lema acima vale para as colunas de $A^M$, $M \geq 1$, se supusermos que $A$ é matriz de transição no sentido bilateral.*

Considerando a sequência $\{\lambda^{-n}\mathcal{L}^n f\}_{n \geq 0}$, podemos encontrar uma autofunção $h$.

**Lema 5.10.** *Existe uma função $h \in \Lambda$ tal que $\mathcal{L}h = \lambda h$ e $h > 0$.*

*Demonstração.* Primeiramente checamos que $\lambda^{-1}\mathcal{L}f \in \Lambda$ quando $f \in \Lambda$. Como $\mathcal{L}$ é um operador positivo $\lambda^{-1}\mathcal{L}f \geq 0$ e

$$\int \lambda^{-1}\mathcal{L}f d\nu = \int \lambda^{-1} f d(\mathcal{L}^*\nu) = \int \lambda^{-1} f d(\lambda\nu) = \int f d\nu = 1. \tag{5.5}$$

Assumindo que $x_i = \bar{x}_i$ para todo $i \in \{0, \ldots, m\}$, segue que

$$\mathcal{L}f((x_n)_n) = \sum_{(z_n)_n \in (\sigma_A^+)^{-1}((x_n)_n)} e^{\phi((z_n)_n)} f((z_n)_n) = \sum_{\{(y_n)_n:\ A_{y_0,x_0}=1\}} e^{\phi((y_n)_n)} f((y_n)_n)$$

e

$$\mathcal{L}f((\bar{x}_n)_n) = \sum_{(\bar{z}_n)_n \in (\sigma_A^+)^{-1}((\bar{x}_n)_n)} e^{\phi((\bar{z}_n)_n)} f((\bar{z}_n)_n) = \sum_{\{(\bar{y}_n)_n:\ A_{\bar{y}_0,\bar{x}_0}=1\}} e^{\phi((\bar{y}_n)_n)} f((\bar{y}_n)_n).$$

Nesse caso $\{(y_n)_n :\ A_{y_0,x_0} = 1\} = \{(\bar{y}_n)_n :\ A_{\bar{y}_0,\bar{x}_0} = 1\}$ porque $x_i = \bar{x}_i$ para todo $i \in 0, \ldots, m$. Também devido a isso, $y_j = \bar{y}_j$ para todo $j \in 0, \ldots, m+1$. Então como $f \in \Lambda$

$$\begin{aligned} e^{\phi((y_n)_n)} f((y_n)_n) &\leq e^{\phi((\bar{y}_n)_n)} B_{m+1} f((\bar{y}_n)_n) \\ &\leq e^{\phi((\bar{y}_n)_n)} e^{b\alpha^{m+1}} B_{m+1} f((\bar{y}_n)_n) \\ &\leq e^{\phi((\bar{y}_n)_n)} B_m f((\bar{y}_n)_n). \end{aligned}$$

Relacionando essa desigualdade com as duas igualdades dadas acima, resulta que

$$\mathcal{L}f((x_n)_n) \leq B_m \mathcal{L}f((\bar{x}_n)_n). \tag{5.6}$$

Considere qualquer $(x_n)_n, (z_n)_n \in \Sigma_A^+$. Uma vez que $A^M > 0$ (Veja o Lema 5.8) existe $(\bar{y}_n)_n \in (\sigma_A^+)^{-M}((x_n)_n)$ com $\bar{y}_0 = z_0$. Para $f \in \Lambda$

$$\mathcal{L}^M f((x_n)_n) = \sum_{(y_n)_n \in (\sigma_A^+)^{-M}((x_n)_n)} e^{\phi_M((y_n)_n)} f((y_n)_n). \tag{5.7}$$

Recordando a definição de $\phi_M$, notamos que $\phi_M((y_n)_n) = \sum_{j=0}^{M-1} \phi((\sigma_A^+)^j((y_n)_n)) \geq -M||\phi||$. Desse fato, de (5.7) e observando que $(\bar{y}_n)_n \in (\sigma_A^+)^{-M}((x_n)_n)$, resulta que

$$\mathcal{L}^M f((x_n)_n) \geq e^{-M||\phi||} f((\bar{y}_n)_n). \tag{5.8}$$

Seja $K = \lambda^M e^{M||\phi||} B_0$. Então fazendo indução na relação (5.5) e em seguida usando a relação (5.8),

$$\begin{aligned} 1 = \int \lambda^{-M} \mathcal{L}^M f d\nu &\geq \int \lambda^{-M} e^{-M||\phi||} f((\bar{y}_n)_n) d\nu \\ &\geq \int \lambda^{-M} e^{-M||\phi||} (B_0)^{-1} f((z_n)_n) d\nu = K^{-1} f((z_n)_n). \end{aligned} \tag{5.9}$$

Portanto, $||f|| \leq K$ uma vez que $(z_n)_n$ é arbitrária. Como $\int f d\nu = 1$, então $f((z_n)_n) \geq 1$ para alguma sequência $(z_n)_n$ e assim obtemos de (5.9) que $\inf \lambda^{-M}\mathcal{L}^M f \geq K^{-1}$.

Se $x_i = \bar{x}_i$ para todo $i \in 0, \ldots, m$. e $f \in \lambda$, temos que

$$\begin{aligned}|f((x_n)_n) - f((\bar{x}_n)_n)| &\leq |B_m f((\bar{x}_n)_n) - f((\bar{x}_n)_n)| \\ &= (B_m - 1)|f((\bar{x}_n)_n)| \leq (B_m - 1)||f|| \leq (B_m - 1)K \to 0\end{aligned}$$

quando $m \to \infty$, porque $B_m = \exp(\sum_{k=m+1}^{\infty} 2b\alpha^k) \to 1$. Então $\Lambda$ é uma família equicontínua e compacta devido ao Teorema de Arzelá-Ascoli. Notando que $\Lambda$ é convexo é não vazio porque $1 \in \Lambda$, aplicamos o Teorema de Schauder-Tychonoff (Teorema 5.7) a $\lambda^{-1}\mathcal{L} : \Lambda \to \Lambda$ e conseguimos uma função $h \in \Lambda$ com $\lambda^{-1}\mathcal{L}h = h$, ou seja, $\mathcal{L}h = \lambda h$. Além disso, $\inf h = \inf \lambda^{-M}\mathcal{L}^M h \geq K^{-1}$. □

**Lema 5.11.** *Existe $\delta \in (0,1)$ de modo que para $f \in \Lambda$, temos $\lambda^{-M}\mathcal{L}^M f = \delta h + (1-\delta)\tilde{f}$ com $\tilde{f} \in \Lambda$.*

*Demonstração.* Seja $g = \lambda^{-M}\mathcal{L}^M f - \delta h$ onde $\delta$ é um número a ser determinado. Supondo que $\delta||h|| \leq K^{-1}$, teremos que $g \geq 0$. Assuma que $x_i = \bar{x}_i$ para todo $i \in \{0, \ldots, m\}$. Queremos escolher $\delta$ de modo que $g((x_n)_n) \leq B_m g((\bar{x}_n)_n)$, ou equivalentemente

$$\begin{aligned}\lambda^{-M}\mathcal{L}^M f((x_n)_n) - \delta h((x_n)_n) &\leq B_m \lambda^{-M}\mathcal{L}^M f((\bar{x}_n)_n) - B_m \delta h((\bar{x}_n)_n) \Leftrightarrow \\ \delta(B_m h((\bar{x}_n)_n) - h((x_n)_n)) &\leq B_m \lambda^{-M}\mathcal{L}^M f((\bar{x}_n)_n) - \lambda^{-M}\mathcal{L}^M f((x_n)_n) \quad (5.10)\end{aligned}$$

Dissemos na equação (5.6) que para qualquer $f_1 \in \Lambda$

$$\mathcal{L}f_1((x_n)_n) \leq B_m \mathcal{L}f_1((\bar{x}_n)_n) = B_{m+1} e^{b\alpha^{m+1}} \mathcal{L}f_1((\bar{x}_n)_n) \leq B_{m+1} e^{b\alpha^m} \mathcal{L}f_1((\bar{x}_n)_n). \quad (5.11)$$

Aplicando a desigualdade (5.11) para $f_1 = \lambda^{-M+1}\mathcal{L}^{M-1} f$, obtemos

$$\lambda^{-M}\mathcal{L}^M f((x_n)_n) \leq B_{m+1} e^{b\alpha^m} \lambda^{-M}\mathcal{L}^M f((\bar{x}_n)_n). \quad (5.12)$$

Temos que

$$h((x_n)_n) \geq (B_m)^{-1} h((\bar{x}_n)_n) \quad (5.13)$$

uma vez que $h \in \Lambda$. Então para obtermos (5.10), é suficiente que mostrarmos que

$$\delta(B_m h((\bar{x}_n)_n) - (B_m)^{-1} h((\bar{x}_n)_n)) \leq (B_m - B_{m+1} e^{b\alpha^m}) \lambda^{-M} \mathcal{L}^M f((\bar{x}_n)_n), \tag{5.14}$$

porque assumindo (5.14) e usando as equações (5.13) e (5.12), concluímos (5.10):

$$\begin{aligned}
\delta(B_m h((\bar{x}_n)_n) - h((x_n)_n)) &\leq \delta(B_m h((\bar{x}_n)_n) - (B_m)^{-1} h((\bar{x}_n)_n)) \\
&\leq B_m \lambda^{-M} \mathcal{L}^M f((\bar{x}_n)_n) - B_{m+1} e^{b\alpha^m} \lambda^{-M} \mathcal{L}^M f((\bar{x}_n)_n) \\
&\leq B_m \lambda^{-M} \mathcal{L}^M f((\bar{x}_n)_n) - \lambda^{-M} \mathcal{L}^M f((x_n)_n).
\end{aligned}$$

Ainda, notando que $\delta||h|| \leq K^{-1} \leq \inf \lambda^{-M} \mathcal{L}^M f$, para concluirmos (5.14) é suficiente mostrar que

$$\delta(B_m - (B_m)^{-1})||h|| \leq (B_m - B_{m+1} e^{b\alpha^m}) K^{-1}. \tag{5.15}$$

Existe algum $L$ de tal forma que os logaritmos de $B_m, (B_m)^{-1}$ e $B_{m+1} e^{b\alpha^m}$ pertencem a $[-L, L]$ para todo $m$. Sejam $u_1, u_2$ constantes positivas tais que

$$u_1(x - y) \leq e^x - e^y \leq u_2(x - y) \text{ para todo } x; y \in [-L, L]; x > y. \tag{5.16}$$

De (5.15) e (5.16), vemos que para que (5.10) seja válida, é suficiente que $\delta > 0$ satisfaça

$$\begin{aligned}
\delta||h||u_1(\log B_m + \log(B_m)) &= \delta||h||u_1(\log B_m - \log(B_m)^{-1}) \\
&\leq \delta||h||u_1(e^{\log B_m} - e^{\log(B_m)^{-1}}) \\
&= \delta||h||(B_m - (B_m)^{-1}) \leq K^{-1}(B_m - B_{m+1} e^{b\alpha^m}) \\
&= K^{-1}(e^{\log B_m} - e^{\log B_{m+1} e^{b\alpha^m}}) \\
&\leq K^{-1} u_2(\log B_m - \log B_{m+1} e^{b\alpha^m})
\end{aligned}$$

ou seja,

$$
\begin{aligned}
\delta||h||u_1\left(\frac{4b\alpha^{m+1}}{1-\alpha}\right) &= \delta||h||u_1(4b\sum_{j=0}^{\infty}\alpha^{m+1+j}) \\
&= \delta||h||u_1(\log(\exp(\sum_{k=m+1}^{\infty} 2b\alpha^k)) + \log(\exp(\sum_{k=m+1}^{\infty} 2b\alpha^k))) \\
&= \delta||h||u_1(\log B_m + \log(B_m)) \leq K^{-1}u_2(\log B_m - \log B_{m+1}e^{b\alpha^m}) \\
&= K^{-1}u_2(2b\sum_{j=0}^{\infty}\alpha^{m+1+j} - b\alpha^m - 2b\sum_{j=0}^{\infty}\alpha^{m+2+j}) \\
&= K^{-1}u_2(2b\sum_{j=0}^{\infty}(\alpha^{m+1+j} - \alpha^{m+2+j}) - b\alpha^m) \\
&= K^{-1}u_2(2b\alpha^{m+1} - b\alpha^m) \leq K^{-1}u_2(2b\alpha^m - b\alpha^m) = K^{-1}u_2 b\alpha^m
\end{aligned}
$$

ou ainda

$$0 < \delta \leq u_2(1-\alpha)(4\alpha u_1||h||K)^{-1}.$$

Portanto se $\delta$ é determinado dessa forma e se definirmos $\tilde{f} = \dfrac{g}{(1-\delta)}$, temos que $\tilde{f}((x_n)_n) \leq B_m\tilde{f}((\bar{x}_n)_n)$ sempre que $x_i = \bar{x}_i$ para todo $i \in 0,\ldots,m$ e

$$\int \tilde{f}d\nu = \int \frac{g}{(1-\delta)}d\nu = \int \frac{(\lambda^{-M}\mathcal{L}^M f - \delta h)}{(1-\delta)}d\nu = \frac{1-\delta}{1-\delta} = 1.$$

Isto prova que $\tilde{f} \in \Lambda$. □

**Lema 5.12.** *Existem constantes $A > 0$ e $\beta \in (0,1)$ tais que*

$$||\lambda^{-n}\mathcal{L}^n f - h|| \leq A\beta^n$$

*para toda função $f \in \Lambda$ e $n \geq 0$.*

*Demonstração.* Seja $n = Mq + r$ com $0 \leq r < M$. Verificamos por indução em $q$ que

$$\lambda^{-Mq}\mathcal{L}^{Mq}f = (1-(1-\delta)^q)h + (1-\delta)^q\tilde{f}_q \tag{5.17}$$

onde $\tilde{f}_q \in \Lambda$. De fato, o caso em $q = 1$ corresponde ao resultado do Lema 5.11. Admitindo que (5.17) seja verdadeira para $1 \leq n \leq q$. Então para $q+1$

$$\begin{aligned}
\lambda^{-M(q+1)}\mathcal{L}^{M(q+1)}f &= \lambda^{-M}\mathcal{L}^{M}(\lambda^{-Mq}\mathcal{L}^{Mq}f)\\
&= \lambda^{-M}\mathcal{L}^{M}((1-(1-\delta)^q)h + (1-\delta)^q\tilde{f}_q)\\
&= (1-(1-\delta)^q)\lambda^{-M}\mathcal{L}^{M}h + (1-\delta)^q\lambda^{-M}\mathcal{L}^{M}\tilde{f}_q\\
&= (1-(1-\delta)^q)h + (1-\delta)^q\lambda^{-M}\left(\frac{\lambda^{-M}\mathcal{L}^{2M}f}{1-\delta} - \frac{\delta}{1-\delta}\lambda^{M}h\right)\\
&= (1-(1-\delta)^q - \delta(1-\delta)^{q-1})h + (1-\delta)^{q-1}\lambda^{-2M}\mathcal{L}^{2M}f\\
&= (1-(1-\delta)^q - \delta(1-\delta)^{q-1})h + (1-\delta)^{q-1}((1-(1-\delta)^2)h + (1-\delta)^2\tilde{f}_2)\\
&= (1-(1-\delta)^q - \delta(1-\delta)^{q-1} + (1-\delta)^{q-1} - (1-\delta)^{q+1})h + (1-\delta)^{q+1}\tilde{f}_2\\
&= (1-((1-\delta)+1)(1-\delta)^{q-1} - \delta(1-\delta)^{q-1} - (1-\delta)^{q+1})h + (1-\delta)^{q+1}\tilde{f}_2\\
&= (1-(1-\delta)^{q+1})h + (1-\delta)^{q+1}\tilde{f}_2.
\end{aligned}$$

Isto finaliza a prova de de (5.17). Como $||\tilde{f}_q|| \leq K$ (Veja o Lema 5.10), temos que

$$||\lambda^{-Mq}\mathcal{L}^{Mq}f - h|| = ||(1-\delta)^q h + (1-\delta)^q\tilde{f}_q|| \leq (1-\delta)^q(||h|| + K)$$

e

$$\begin{aligned}
||\lambda^{-n}\mathcal{L}^{n}f - h|| = ||\lambda^{-Mq-r}\mathcal{L}^{Mq+r}f - h|| &= ||\lambda^{-r}\mathcal{L}^{r}(\lambda^{-Mq}\mathcal{L}^{Mq}f - h)||\\
&\leq \sup_{0\leq r\leq M} ||\lambda^{-r}\mathcal{L}^{r}||\,(1-\delta)^q\,(||h|| + K)\\
&= \sup_{0\leq r\leq M} ||\lambda^{-r}\mathcal{L}^{r}||\frac{(||h|| + K)}{(1-\delta)}\,(1-\delta)^{q+1}.
\end{aligned}$$

Agora tome

$$A = \sup_{0\leq r\leq M} ||\lambda^{-r}\mathcal{L}^{r}||\frac{(||h|| + K)}{(1-\delta)} \quad \text{e} \quad \beta^M = 1-\delta.$$

Segue que

$$||\lambda^{-n}\mathcal{L}^{n}f - h|| \leq A\beta^{Mq+M} \leq A\beta^{Mq+r} = A\beta^{n}.$$

$\square$

**Lema 5.13.** *Seja* $\mathcal{C}_r = \{f \in \mathcal{C}(\Sigma_A^+) : \mathrm{var}_r\, f = 0\}$. *Se* $F \in \Lambda$, $f \in \mathcal{C}_r$, $f \geq 0$ *e* $fF \neq 0$, *então* $\dfrac{\lambda^{-r}\mathcal{L}^{r}(fF)}{\int(fF)d\nu} \in \Lambda$.

*Demonstração.* Supondo que $x_i = \bar{x}_i$ para todo $i \in 0, \ldots, m.$, segue que

$$\mathcal{L}^r(fF)((x_n)_n) = \sum_{(y_n)_n \in (\sigma_A^+)^{-r}((x_n)_n)} e^{\phi_r((y_n)_n)}(fF)((y_n)_n). \tag{5.18}$$

onde $\phi_r((y_n)_n) = \sum_{t=0}^{r-1} \phi((\sigma_A^+)^t((y_n)_n))$ e a soma é tomada sobre todas as sequências $(y_n)_n$ do tipo $(j_1, \ldots, j_r, x_0, x_1, \ldots)$ tais que $A^r_{j_1,x_0} = 1$. Note que a soma na expressão de $\mathcal{L}^r(fF)((\bar{x}_n)_n)$ é tomada sobre todas as sequências $(\bar{y}_n)_n$ cujos os $r$ primeiros símbolos percorrem o mesmo conjunto que os $r$ primeiros símbolos das sequências $(y_n)_n$ da soma de $\mathcal{L}^r(fF)((x_n)_n)$, e isto porque $x_i = \bar{x}_i$ para todo $i \in 0, \ldots, m$. Como $f \in \mathcal{C}_r$, então $\operatorname{var}_r f = 0$ implica que $|f((z_n)_n) - f((w_n)_n)| = 0$, ou seja $f((z_n)_n) = f((w_n)_n)$ para todas as sequências tais que $z_i = w_i$ com $i = 0, \ldots, r$. Mas como $x_i = \bar{x}_i$ para todo $i \in 0, \ldots, m$, obtemos em especial que $f(j_1, \ldots, j_r, x_0, x_1, \ldots) = f(j_1, \ldots, j_r, \bar{x}_0, \bar{x}_1, \ldots)$ porque estas sequências tem os $r + m$ primeiro símbolos iguais. Daí como $F \in \Lambda$, $F(j_1, \ldots, j_r, x_0, x_1, \ldots) \leq B_{m+r} F(j_1, \ldots, j_r, \bar{x}_0, \bar{x}_1, \ldots)$ e

$$\phi((\sigma_A^+)^t(j_1, \ldots, j_r, x_0, x_1, \ldots)) - \phi((\sigma_A^+)^t(j_1, \ldots, j_r, \bar{x}_0, \bar{x}_1, \ldots)) \leq \operatorname{var}_{m+r-t} \phi.$$

Além disso:

$$\begin{aligned} B_{m+r} \exp\left(\sum_{t=0}^{r-1} \operatorname{var}_{m+r-t} \phi\right) &\leq B_{m+r} \exp\left(\sum_{t=0}^{r-1} b\alpha^{m+r-t}\right) \\ &= B_{m+r} \exp\left(\sum_{l=m+1}^{m+r} b\alpha^l\right) \\ &\leq B_m \exp\left(\sum_{l=m+1}^{m+r} (b\alpha^l - 2b\alpha^l)\right) \\ &= B_m \exp\left(\sum_{l=m+1}^{m+r} -b\alpha^l\right) \leq B_m. \end{aligned}$$

Reunindo todas as expressões obtidas, conseguimos a seguinte desigualdade

$$
\begin{aligned}
\mathcal{L}^r(fF)((x_n)_n) &= \sum_{(y_n)_n \in (\sigma_A^+)^{-r}((x_n)_n)} e^{\phi_r((y_n)_n)}(fF)((y_n)_n) \\
&= \sum_{(j_1,\ldots,j_r,x_0,x_1,\ldots)} \exp\left(\sum_{t=0}^{r-1} \phi(((\sigma_A^+)^t)(j_1,\ldots,j_r,x_0,x_1,\ldots))\right) \\
&\quad (fF)(j_1,\ldots,j_r,x_0,x_1,\ldots) \\
&\leq \sum_{(j_1,\ldots,j_r,\bar{x}_0,\bar{x}_1,\ldots)} B_{m+r} \exp\left(\sum_{t=0}^{r-1}(\phi((\sigma_A^+)^t(j_1,\ldots,j_r,\bar{x}_0,\bar{x}_1,\ldots))\right. \\
&\quad \left. + \operatorname{var}_{m+r-t}\phi)\right)(fF)(j_1,\ldots,j_r,\bar{x}_0,\bar{x}_1,\ldots) \\
&= \sum_{(j_1,\ldots,j_r,\bar{x}_0,\bar{x}_1,\ldots)} \exp\left(\sum_{t=0}^{r-1}(\phi((\sigma_A^+)^t(j_1,\ldots,j_r,\bar{x}_0,\bar{x}_1,\ldots))\right) \\
&\quad B_{m+r}\exp\left(\sum_{t=0}^{r-1}\operatorname{var}_{m+r-t}\phi\right)(fF)(j_1,\ldots,j_r,\bar{x}_0,\bar{x}_1,\ldots) \\
&\leq B_m \sum_{(j_1,\ldots,j_r,\bar{x}_0,\bar{x}_1,\ldots)} \exp\left(\sum_{t=0}^{r-1}(\phi((\sigma_A^+)^t(j_1,\ldots,j_r,\bar{x}_0,\bar{x}_1,\ldots))\right) \\
&\quad (fF)(j_1,\ldots,j_r,\bar{x}_0,\bar{x}_1,\ldots) \\
&= B_m \sum_{(\bar{y}_n)_n \in (\sigma_A^+)^{-r}((\bar{x}_n)_n)} e^{\phi_r((\bar{y}_n)_n)}(fF)((\bar{y}_n)_n) \\
&= B_m \mathcal{L}^r(fF)((\bar{x}_n)_n).
\end{aligned}
$$

Agora mostraremos que $\int(fF)d\nu > 0$. Usando a equação (5.9) do Lema 5.10, porém trocando $f$ por $\mathcal{L}^r(fF)$, obtemos

$$
\begin{aligned}
\lambda^r \int (fF)d\nu &= \int \lambda^{M+r}\lambda^{-M}(fF)d\nu \\
&= \int \lambda^{-M}(fF)d(\lambda^{M+r}\nu) \\
&= \int \lambda^{-M}(fF)d(\mathcal{L}^{*M+r}\nu) \\
&= \int \lambda^{-M}\mathcal{L}^{M+r}(fF)d\nu \\
&= \int \lambda^{-M}\mathcal{L}^{M}\mathcal{L}^r(fF)d\nu \\
&\geq K^{-1}\mathcal{L}^r(fF)((z_n)_n)
\end{aligned}
$$

para qualquer sequência $(z_n)_n$. Mas $(fF)((w_n)_n) > 0$ (porque $f$, $F \geq 0$ e $fF \neq 0$ por hipótese) implica que $\mathcal{L}^r(fF)((\sigma_A^+)^r((w_n)_n)) > 0$. Como também $K^{-1} > 0$, temos que $\int (fF)d\nu > 0$. Finalmente, se $g = \dfrac{\lambda^{-r}\mathcal{L}^r(fF)}{\int (fF)d\nu}$ então

$$\begin{aligned}
\int g d\nu &= \int \frac{\lambda^{-r}\mathcal{L}^r(fF)}{\int (fF)d\nu} d\nu \\
&= \frac{1}{\int (fF)d\nu} \int \lambda^{-r}(fF)d(\mathcal{L}^{*r}\nu) \\
&= \frac{1}{\int (fF)d\nu} \int \lambda^{-r}(fF)d(\lambda^r\nu) \\
&= \frac{\int (fF)d\nu}{\int (fF)d\nu} = 1.
\end{aligned}$$

Com isso concluímos que $g = \dfrac{\lambda^{-r}\mathcal{L}^r(fF)}{\int (fF)d\nu} \in \Lambda$. □

**Lema 5.14.** *Para $f \in \mathcal{C}_r$, $F \in \Lambda$ e $n \geq 0$*

$$||\lambda^{-n-r}\mathcal{L}^{n+r}(fF) - h\int (fF)d\nu|| \leq A\beta^n \int |fF|d\nu.$$

*Para $g \in \mathcal{C}(\Sigma_A^+)$, temos que* $\lim_{m\to\infty} ||\lambda^{-m}\mathcal{L}^m g - h\int g d\nu|| = 0$.

*Demonstração.* Para $f \in \mathcal{C}_r$ escreva $f = f^+ - f^-$ com $f^+$, $f^- \geq 0$ e $f^+, f^- \in \mathcal{C}_r$. Então

$$||\lambda^{-n-r}\mathcal{L}^{n+r}(f^\pm F) - h\int (f^\pm F)d\nu|| \leq A\beta^n \int |f^\pm F|d\nu.$$

Isto é claro se $f^\pm F = 0$. Para $f^\pm F \neq 0$, aplicamos os Lemas 5.12 e 5.13:

$$\begin{aligned}
||\lambda^{-n-r}\mathcal{L}^{n+r}(f^\pm F) - h\int (f^\pm F)d\nu|| &= ||\int (f^\pm F)d\nu \left(\frac{\lambda^{-n-r}\mathcal{L}^{n+r}(f^\pm F)}{\int (f^\pm F)d\nu} - h\right)|| \\
&= ||\int (f^\pm F)d\nu||\, ||\lambda^{-n}\mathcal{L}^n\left(\frac{\lambda^{-r}\mathcal{L}^r(f^\pm F)}{\int (f^\pm F)d\nu}\right) - h|| \\
&\leq \left(\int |f^\pm F|d\nu\right) A\beta^n.
\end{aligned} \tag{5.19}$$

A última desigualdade ocorre porque $\dfrac{\lambda^{-r}\mathcal{L}^r(f^\pm F)}{\int (f^\pm F)d\nu} \in \Lambda$ (Veja o Lema 5.13) e o Lema 5.12 diz que para toda $g \in \Lambda$ e $n \geq 0$, $||\lambda^{-n}\mathcal{L}^n g - h|| \leq A\beta^n$. Para concluir a primeira parte do lema, basta repetir o cálculo feito em (5.19) para $f = f^+ - f^-$, usar a desigualdade triangular e usar que $F$, $f^+$, $f^- \geq 0$.

Dada uma função $g$ e $\varepsilon > 0$, podemos encontrar $r$ e funções $f_1, f_2 \in \mathcal{C}_r$ de maneira que $f_1 \leq g \leq f_2$ e $0 \leq f_2 - f_1 \leq \varepsilon$. Como $-\varepsilon \leq f_1 - f_2 \leq g - f_2 \leq g - f_1 \leq f_2 - f_1 \leq \varepsilon$, segue que $|\int f_i \, d\nu - \int g d\nu| < \varepsilon$. A primeira parte do Lema com $F = 1$ implica que

$$\begin{aligned}
||\lambda^{-m}\mathcal{L}^m(f_i) - h\int gd\nu|| &= ||\lambda^{-m}\mathcal{L}^m(f_i) - h\int (f_i)d\nu + h\int (f_i)d\nu - h\int gd\nu|| \\
&\leq ||\lambda^{-m}\mathcal{L}^m(f_i) - h\int (f_i)d\nu|| \,+\, ||h\int (f_i)d\nu - h\int gd\nu|| \\
&\leq ||\lambda^{-m}\mathcal{L}^m(f_i) - h\int (f_i)d\nu|| \,+\, ||h||\,|\int (f_i)d\nu - \int gd\nu| \\
&\leq \varepsilon + \varepsilon||h|| = \varepsilon(1 + ||h||)
\end{aligned}$$

para $m$ suficientemente grande (note que $||\lambda^{-m}\mathcal{L}^m(f_i) - h\int (f_i)d\nu|| \leq A\beta^{m+r}\int |f_i|d\nu \leq \varepsilon$ para $m$ grande ). Uma vez que $\lambda^{-m}\mathcal{L}^m(f_1) \leq \lambda^{-m}\mathcal{L}^m g \leq \lambda^{-m}\mathcal{L}^m(f_2)$, segue que

$$||\lambda^{-m}\mathcal{L}^m g - h\int gd\nu|| \leq \varepsilon(1 + ||h||)$$

para $m$ suficientemente grande. Isto demonstra a segunda parte do Lema. $\square$

A prova do Teorema de Ruelle-Perron-Frobenius (Teorema 5.5) está completa. $\square$

## 5.2 Demonstração do Teorema 5.4: a construção do estado de equilíbrio

Assumiremos que $\phi \in \mathcal{F}_A \cap \mathcal{C}(\Sigma_A^+)$ e $\nu$, $h$, $\lambda$ são como no Teorema de Ruelle-Perron-Frobenius (Teorema 5.5). Consideremos a medida $\mu = h\nu$, a qual vamos provar que é um estado de equilíbrio.

**Lema 5.15.** *$\mu$ é uma medida de probabilidade em $\Sigma_A^+$, e é invariante por $\sigma_A^+ : \Sigma_A^+ \to \Sigma_A^+$.*

*Demonstração.* Que $\mu = h\nu$ é uma medida de probabilidade em $\Sigma_A^+$, é consequência do Teorema de Ruelle-Perron-Frobenius, pois $\int_{\Sigma_A^+} hd\nu = 1$. Além disso $f\mu = fh\nu$, ou seja, $f\mu(\Sigma_A^+) = \int_{\Sigma_A^+} fhd\nu = \int fhd\nu$. Resta mostrar que $\int fd\mu = \int (f \circ \sigma_A^+)d\mu$ para toda função $f \in \mathcal{C}(\Sigma_A^+)$. Fazendo um cálculo semelhante ao que é apresentado em (4.17) do Lema 4.31, vemos que para quaisquer $g_1, g_2 \in \mathcal{C}(\Sigma_A^+)$

$$\mathcal{L}((g_1 \circ \sigma_A^+)g_2)((y_n)_n) = g_1((y_n)_n)\mathcal{L}g_2((y_n)_n).$$

Em seguida, usando isto e imitando o cálculo feito em (4.18) do Lema 4.31, segue que $\int f d\mu = \int (f \circ \sigma_A^+) d\mu$. $\square$

Como $\mu$ é $\sigma_A^+$ invariante em $\Sigma_A^+$ existe uma maneira natural de estender $\mu$ a uma medida em todo $\Sigma_A$.

**Lema 5.16.** *A medida $\mu$ que é $\sigma_A^+$-invariante em $\Sigma_A^+$ pode ser estendida a uma medida $\tilde{\mu}$ que é $\sigma_A$-invariante em $\Sigma_A$.*

*Demonstração.* Para $f \in \mathcal{C}(\Sigma_A)$ defina $f^* \in \mathcal{C}(\Sigma_A^+)$ por

$$f^*(\{x_i\}_{i=0}^{\infty}) = \min\{f((y_n)_n) : (y_n)_n \in \Sigma_A,\, y_i = x_i \text{ para todo } i \geq 0\}.$$

Note que para $m$, $n \geq 0$, temos que

$$\begin{aligned}
((f \circ \sigma_A^n)^* \circ (\sigma_A^+)^m)((y_p)_p) &= (f \circ \sigma_A^n)^*((y_{m+p})_p) \\
&= \min\{(f \circ \sigma_A^n)((u_j)_j) : (u_j)_j \in \Sigma_A,\, y_{m+i} = u_i \;\; \forall \; i \geq 0\} \\
&= \min\{f((u_{n+j})_j) : (u_j)_j \in \Sigma_A,\, y_{m+n+i} = u_{n+i} \;\; \forall \; i \geq 0 \text{ e} \\
&\qquad u_0 = y_m, \ldots, u_{n-1} = y_{m+n-1}\} \\
&= \min\{f((b_j)_j) : (b_j)_j := (u_{n+j})_j \in \Sigma_A,\, y_{m+n+i} = u_{n+i} \\
&\qquad \forall \; i \geq 0 \text{ e } u_0 = y_m, \ldots, u_{n-1} = y_{m+n-1}\}.
\end{aligned}$$

Por outro lado

$$\begin{aligned}
(f \circ \sigma_A^{n+m})^*((y_p)_p) &= \min\{(f \circ \sigma_A^{n+m})((t_j)_j) : (t_j)_j \in \Sigma_A,\, y_i = t_i \;\; \forall \; i \geq 0\} \\
&= \min\{f((t_{n+m+j})_j) : (t_j)_j \in \Sigma_A,\, y_{m+n+i} = t_{m+n+i} \;\; \forall \; i \geq 0 \\
&\qquad \text{e } t_0 = y_0, \ldots, t_{n+m-1} = y_{m+n-1}\} \\
&= \min\{f((c_j)_j) : (c_j)_j := (t_{n+m+j})_j \in \Sigma_A,\, y_{m+n+i} = t_{m+n+i} \;\; \forall \; i \geq 0 \\
&\qquad \text{e } t_0 = y_0, \ldots, t_{n+m-1} = y_{m+n-1}\}.
\end{aligned}$$

Agora repare que se tomarmos sucessivamente $j = -n, -n+1, \ldots, 0, 1, \ldots n,$

teremos

$$\begin{aligned}
b_{-n} &= u_{n-n} = u_0 = y_m = t_m = t_{n+m-n} = c_{-n}\\
b_{-n+1} &= u_{n-n+1} = u_1 = y_{m+1} = t_{m+1} = t_{n+m-n+1} = c_{-n+1}\\
&\vdots\\
b_0 &= u_{n+0} = u_n = y_{m+n} = t_{m+n} = t_{n+m+0} = c_0\\
&\vdots\\
b_{n-1} &= u_{n+n-1} = u_{2n-1} = y_{m+2n-1} = t_{m+2n-1} = t_{n+m+n-1} = c_{n-1}\\
b_n &= u_{n+n} = u_{2n} = y_{2n} = t_{2n} = t_{n+m+n} = c_n.
\end{aligned}$$

Dessas três igualdades, segue que

$$\begin{aligned}
||(f \circ \sigma_A^n)^* \circ (\sigma_A^+)^m - (f \circ \sigma_A^{n+m})^*|| &= \sup_{(y_p)_p \in \Sigma_A^+} |((f \circ \sigma_A^n)^* \circ (\sigma_A^+)^m)((y_p)_p) - (f \circ \sigma_A^{n+m})^*((y_p)_p)|\\
&\leq |f((b_j)_j) - f((c_j)_j)|.
\end{aligned}$$

Mas essas sequências $(b_j)_j$, $(c_j)_j$ coincidem para $|j| < n$ como vimos acima. Portanto, segue da definição de variação que

$$||(f \circ \sigma_A^n)^* \circ (\sigma_A^+)^m - (f \circ \sigma_A^{n+m})^*|| \leq |f((b_j)_j) - f((c_j)_j)| \leq \operatorname{var}_n f.$$

Consequentemente,

$$\begin{aligned}
|\int (f \circ \sigma_A^n)^* d\mu - \int (f \circ \sigma_A^{n+m})^* d\mu| &= |\int ((f \circ \sigma_A^n)^* \circ (\sigma_A^+)^m) d\mu - \int (f \circ \sigma_A^{n+m})^* d\mu|\\
&= |\int ((f \circ \sigma_A^n)^* \circ (\sigma_A^+)^m - (f \circ \sigma_A^{n+m})^*) d\mu|\\
&\leq ||(f \circ \sigma_A^n)^* \circ (\sigma_A^+)^m - (f \circ \sigma_A^{n+m})^*|| \leq \operatorname{var}_n f
\end{aligned}$$

e $\operatorname{var}_n f \to 0$ quando $n \to \infty$, porque $f$ é contínua. Consequentemente $\int f d\tilde{\mu} = \lim_{n\to\infty} \int (f \circ \sigma_A^n)^* d\mu$ existe pelo critério de Cauchy. Também $\tilde{\mu} \in \mathcal{C}(\Sigma_A)^*$, onde $\mathcal{C}(\Sigma_A)^*$ denota o espaço dual de $\mathcal{C}(\Sigma_A)$. Pelo Teorema da Representação de Riesz, $\tilde{\mu}$ é uma medida de probabilidade em $\Sigma_A$, porque

$$\tilde{\mu}(\Sigma_A) = \int 1 d\tilde{\mu} = \lim_{n\to\infty} \int (1 \circ \sigma_A^n)^* d\mu = \lim_{n\to\infty} \int 1 d\mu = \mu(\Sigma_A^+) = 1.$$

Além disso, note que

$$\int (f \circ \sigma_A) d\tilde{\mu} = \lim_{n\to\infty} \int ((f \circ \sigma_A) \circ \sigma_A^n)^* d\mu = \lim_{n\to\infty} \int (f \circ \sigma_A^{n+1})^* d\mu = \int f d\tilde{\mu}$$

portanto, $\tilde{\mu}$ é $\sigma_A$-invariante. Também para $f \in \mathcal{C}(\Sigma_A^+)$,

$$\int f d\tilde{\mu} = \lim_{n\to\infty} \int (f \circ \sigma_A^n)^* d\mu = \lim_{n\to\infty} \int (f \circ \sigma_A^n) d\mu = \lim_{n\to\infty} \int f d\mu = \int f d\mu.$$

□

A partir de agora, iremos denotar a extensão $\tilde{\mu}$ também por $\mu$.

**Proposição 5.17.** *$\mu$ é misturadora para $\sigma_A : \Sigma_A \to \Sigma_A$.*

*Demonstração.* Começamos destacando que é suficiente provar a condição de mistura para conjuntos $E$, $F$ na álgebra gerada pelos cilindros, uma vez que esta álgebra gera a $\sigma$-álgebra dos conjuntos mensuráveis de $\Sigma_A$, conforme o seguinte lema:

**Lema 5.18.** *Suponha que $\lim_n \mu(f^{-n}(A) \cap B) = \mu(A)\mu(B)$ para todo par de conjuntos $A$ e $B$ em alguma álgebra $\mathcal{A}$ geradora da $\sigma$-álgebra dos conjuntos mensuráveis. Então $(f, \mu)$ é misturador.*

A demonstração desse lema pode ser encontrada na referência [7]. Agora passemos a prova da proposição. Escreva $\phi_m((x_n)_n) = \sum_{i=0}^{m-1} \phi(\sigma_A^i((x_n)_n))$. Vamos checar por indução que

$$\mathcal{L}^m f((x_n)_n) = \sum_{(y_n)_n \in \sigma_A{}^{-m}((x_n)_n)} e^{\phi_m((y_n)_n)} f((y_n)_n) \tag{5.20}$$

para $f \in \mathcal{C}(\Sigma_A^+)$. O caso em que $m = 1$ é a definição do operador de transferência.

Supondo que (5.20) vale para algum $m$, temos o seguinte para $m+1$,

$$\begin{aligned}
\mathcal{L}^{m+1}f((x_n)_n) &= \mathcal{L}(\mathcal{L}^m f)((x_n)_n) \\
&= \sum_{(y_n)_n \in {\sigma_A}^{-1}((x_n)_n)} e^{\phi((y_n)_n)} (\mathcal{L}^m f)((y_n)_n) \\
&= \sum_{(y_n)_n \in {\sigma_A}^{-1}((x_n)_n)} e^{\phi((y_n)_n)} \left( \sum_{(z_n)_n \in {\sigma_A}^{-m}((y_n)_n)} e^{\phi_m((z_n)_n)} f((z_n)_n) \right) \\
&= \sum_{(y_n)_n \in {\sigma_A}^{-1}((x_n)_n)} e^{\phi({\sigma_A}^m((z_n)_n))} \left( \sum_{(z_n)_n \in {\sigma_A}^{-m}((y_n)_n)} e^{\phi_m((z_n)_n)} f((z_n)_n) \right) \\
&= \sum_{(z_n)_n \in {\sigma_A}^{-m-1}((x_n)_n)} e^{\phi({\sigma_A}^m((z_n)_n))} e^{\phi_m((z_n)_n)} f((z_n)_n) \\
&= \sum_{(z_n)_n \in {\sigma_A}^{-m-1}((x_n)_n)} e^{\phi_{m+1}((z_n)_n)} f((z_n)_n).
\end{aligned}$$

Isto prova o desejado. Agora, é imediato que para $g \in \mathcal{C}(\Sigma_A^+)$

$$((\mathcal{L}^m f).g)((x_n)_n) = \sum_{(y_n)_n \in {\sigma_A}^{-m}((x_n)_n)} e^{\phi_m((y_n)_n)} f((y_n)_n) g(\sigma_A^m((y_n)_n)) = \mathcal{L}^m(f.(g \circ \sigma_A^m))((x_n)_n). \tag{5.21}$$

Sejam

$$E = \{(y_n)_n \in \Sigma_A : y_i = a_i,\ r \leq i \leq s\}$$

e

$$F = \{(y_n)_n \in \Sigma_A : y_i = b_i,\ u \leq i \leq v\}$$

dois cilindros em $\Sigma_A$. Podemos ainda supor que $r = u = 0$, porque a família de tais cilindros é geradora da $\sigma$-álgebra. Calculamos

$$\begin{aligned}
\mu(E \cap \sigma_A^{-n} F) &= \int \chi_E \ \chi_{\sigma_A^{-n} F} d\mu \\
&= \int \chi_E \ (\chi_F \circ \sigma_A^n) d\mu \\
&= \int \chi_E \ (\chi_F \circ \sigma_A^n) d(h\nu) \\
&= \int h \, \chi_E \ (\chi_F \circ \sigma_A^n) d\nu \\
&= \int \lambda^{-n} h \, \chi_E \ (\chi_F \circ \sigma_A^n) d(\lambda^n \nu) \\
&= \int \lambda^{-n} h \, \chi_E \ (\chi_F \circ \sigma_A^n) d(\mathcal{L}^{*n} \nu) \\
&= \int \lambda^{-n} \mathcal{L}^n (h \, \chi_E \ (\chi_F \circ \sigma_A^n)) d\nu \\
&= \int \lambda^{-n} (\mathcal{L}^n (h \, \chi_E) \, \chi_F) d\nu
\end{aligned}$$

A última igualdade se dá pela relação (5.21). Agora,

$$\begin{aligned}
|\mu(E \cap {\sigma_A}^{-n} F) - \mu(E)\mu(F)| &= |\int \lambda^{-n} (\mathcal{L}^n (h \, \chi_E) \, \chi_F) d\nu - \mu(E)\mu(F)| && (5.22) \\
&= |\int \lambda^{-n} (\mathcal{L}^n (h \, \chi_E) \, \chi_F) d\nu - \int \chi_E d\mu \int \chi_F d\mu| && (5.23) \\
&= |\int \lambda^{-n} (\mathcal{L}^n (h \, \chi_E) \, \chi_F) d\nu - \int h\chi_E d\nu \int h\chi_F d\nu| && (5.24) \\
&= |\int \left( \lambda^{-n} \mathcal{L}^n (h \, \chi_E) - h \int h\chi_E d\nu \right) \chi_F d\nu| && (5.25) \\
&\leq ||\lambda^{-n} \mathcal{L}^n (h \, \chi_E) - h \int h\chi_E d\nu|| \ \nu(F). && (5.26)
\end{aligned}$$

Lembre que $\mathcal{C}_s = \{f \in \mathcal{C}(\Sigma_A^+) : \mathrm{var}_s \, f = 0\}$ e note que $\chi_E \in \mathcal{C}_s$, pois $E$ é um cilindro de tamanho $s$, logo $\mathrm{var}_s \, \chi_E = \sup\{|\chi_E((x_n)_n) - \chi_E((y_n)_n)| : x_i = y_i$ para todo $i$ com $0 \leq i \leq s\}$ e a diferença $|\chi_E((x_n)_n) - \chi_E((y_n)_n)|$ só pode ser $|0-0|$ ou $|1-1|$, uma vez que ou $(x_n)_n$, $(x_n)_n \in E$ já que $x_i = y_i$ para $0 \leq i \leq s$ ou $(x_n)_n$, $(x_n)_n \notin E$. Com isso o supremo da definição de variação é zero e $\mathrm{var}_s \, \chi_E = 0$, daí $\chi_E \in \mathcal{C}_s$. Lembrando que a autofunção $h$ é positiva, resulta do Lema 5.14 que, para $n \geq s$

$$||\lambda^{-n}\mathcal{L}^n(h\,\chi_E) - h\int h\chi_E d\nu|| \leq A\beta^{n-s}\int |h\,\chi_E| d\nu \tag{5.27}$$

$$= A\beta^{n-s}\int \chi_E d(h\nu) \tag{5.28}$$

$$= A\beta^{n-s}\int \chi_E d\mu \tag{5.29}$$

$$= A\beta^{n-s}\mu(E) \tag{5.30}$$

onde $\beta \in (0,1)$. Juntando as desigualdades (5.27) e (5.22), obtemos

$$|\mu(E \cap {\sigma_A}^{-n}F) - \mu(E)\mu(F)| \leq A'\beta^{n-s}\mu(E)\mu(F)$$

para $n \geq s$, onde $A' = A(\inf h)^{-1}$. Portanto, fazendo $n \to \infty$, vemos que $|\mu(E \cap {\sigma_A}^{-n}F) - \mu(E)\mu(F)| \to 0$, ou seja, $\mu$ é misturadora para $\sigma_A$. $\square$

**Lema 5.19.** *Seja* $a = \sum_{k=0}^{\infty} \mathrm{var}_k\,\phi < \infty$. *Se* $(x_n)_n, (y_n)_n \in \Sigma_A$ *com* $x_i = y_i$, *para* $i \in \{0, \ldots, m-1\}$, *então*

$$|\phi_m((x_n)_n) - \phi_m((y_n)_n)| \leq a,$$

*onde* $\phi_m((x_n)_n) = \sum_{k=0}^{m-1} \phi(\sigma_A^k((x_n)_n))$.

*Demonstração.* Sejam $(x_n)_n, (y_n)_n \in \Sigma_A$ com $x_i = y_i$, para $i \in \{0, \ldots, m-1\}$. Defina $(y'_n)_n$ por

$$y'_i = \begin{cases} y_i & \text{para } i \geq 0 \\ x_i & \text{para } i \leq 0. \end{cases}$$

Uma vez que $\phi \in \mathcal{C}(\Sigma_A^+)$, $\phi(\sigma_A^k((y'_n)_n)) = \phi(\sigma_A^k((y_n)_n))$ para $k \geq 0$. Consequentemente,

$$|\phi_m((x_n)_n) - \phi_m((y_n)_n)| \leq \sum_{k=0}^{m-1} |\phi(\sigma_A^k((x_n)_n)) - \phi(\sigma_A^k((y'_n)_n))|$$

$$\leq \sum_{k=0}^{m-1} \mathrm{var}_{m-1-k}\,\phi \leq a$$

como queríamos mostrar. $\square$

Agora podemos verificar que $\mu$ é um estado de Gibbs.

**Lema 5.20.** *$\mu$ é o único estado de Gibbs para $\phi \in \mathcal{F}_A \cap \mathcal{C}(\Sigma_A^+)$.*

*Demonstração.* Seja $E = \{(y_n)_n \in \Sigma_A : x_i = y_i \text{ para todo } i = 0, \ldots, m-1\}$ como no numerador na definição de estado de Gibbs. Para qualquer $(z_n)_n \in \Sigma_A^+$ existe no máximo uma sequência $(y'_n)_n \in \sigma_A^{-m}((z_n)_n)$ com $(y'_n)_n \in E$ (de fato $(y'_n)_n = (y'_0 = x_0,\, y'_1 = x_1, \ldots, y'_{m-1} = x_{m-1},\, y'_m = z_0,\, y'_{m+1} = z_1, \ldots)$ e é única porque os símbolos das sequências em $E$ correspondentes aos índices $0, \ldots, m$ sempre são os mesmos). Vamos provar primeiramente que $\mu$ é um estado de Gibbs usando o Lema 5.19:

$$\begin{aligned}
\mathcal{L}^m(h\chi_E)((z_n)_n) &= \sum_{(y_n)_n \in \sigma_A{}^{-m}((z_n)_n)} e^{\phi_m((y_n)_n)} h((y_n)_n)\chi_E((y_n)_n) \\
&\leq \sum_{(y_n)_n \in \sigma_A{}^{-m}((z_n)_n)} e^{\phi_m((x_n)_n)} e^{\phi_m((y_n)_n)} e^{-\phi_m((x_n)_n)} h((y_n)_n)\chi_E((y_n)_n) \\
&\leq \sum_{(y_n)_n \in \sigma_A{}^{-m}((z_n)_n)} e^{\phi_m((x_n)_n)} e^{|\phi_m((y_n)_n) - \phi_m((x_n)_n)|} h((y_n)_n)\chi_E((y_n)_n) \\
&\leq e^{\phi_m((x_n)_n)} e^a ||h|| \sum_{(y_n)_n \in \sigma_A{}^{-m}((z_n)_n)} \chi_E((y_n)_n) \\
&= e^{\phi_m((x_n)_n)} e^a ||h||
\end{aligned}$$

e assim

$$\begin{aligned}
\mu(E) = \int \chi_E d\mu = \int h\chi_E d\nu \\
&= \lambda^{-m} \int h\chi_E d(\lambda^m \nu) \\
&= \lambda^{-m} \int h\chi_E d(\mathcal{L}^{*m} \nu) \\
&= \lambda^{-m} \int \mathcal{L}^m(h\chi_E) d\nu \\
&\leq \lambda^{-m} e^{\phi_m((x_n)_n)} e^a ||h||.
\end{aligned}$$

Tomamos $c_2 = e^a||h||$. Por outro lado para qualquer $(z_n)_n \in \Sigma_A^+$ existe pelo menos uma sequência $(y'_n)_n \in \sigma_A^{-m-M}((z_n)_n)$ com $(y'_n)_n \in E$. Então

$$\begin{aligned}
\mathcal{L}^{m+M}(h\chi_E)((z_n)_n) &= \sum_{(y_n)_n \in \sigma_A{}^{-m-M}((z_n)_n)} e^{\phi_{m+M}((y_n)_n)} h((y_n)_n)\chi_E((y_n)_n) \\
&\geq e^{\phi_{m+M}((y'_n)_n)} h((y'_n)_n) \\
&= e^{\phi_{m+M}((y'_n)_n)} e^{-\phi_m((x_n)_n)} h((y'_n)_n) e^{\phi_m((x_n)_n)} \\
&= e^{\sum_{k=m+1}^{m+M} \phi(\sigma_A^k((y'_n)_n))} e^{\phi_m((y'_n)_n)-\phi_m((x_n)_n)} h((y'_n)_n) e^{\phi_m((x_n)_n)} \\
&\geq e^{\sum_{k=m+1}^{m+M} \phi(\sigma_A^k((y'_n)_n))} e^{-a} h((y'_n)_n) e^{\phi_m((x_n)_n)} \\
&\geq e^{-M||\phi||} e^{-a} (\inf h) e^{\phi_m((x_n)_n)}
\end{aligned}$$

e

$$\begin{aligned}
\mu(E) &= \int h\chi_E d\nu \\
&= \lambda^{-m-M} \int h\chi_E d(\lambda^{m+M}\nu) \\
&= \lambda^{-m-M} \int h\chi_E d(\mathcal{L}^{*m+M}\nu) \\
&= \lambda^{-m-M} \int \mathcal{L}^{m+M}(h\chi_E) d\nu \\
&\geq \lambda^{-m-M} e^{-M||\phi||} e^{-a} e^{\phi_m((x_n)_n)} \int d\mu \\
&= c_1 \lambda^{-m} e^{\phi_m((x_n)_n)},
\end{aligned}$$

onde $c_1 = \lambda^{-M} e^{-M||\phi||-a}$. Portanto,

$$c_1 \lambda^{-m} e^{\phi_m((x_n)_n)} \leq \mu(E) \leq c_2 \lambda^{-m} e^{\phi_m((x_n)_n)},$$

ou

$$c_1 \leq \frac{\mu(E)}{\lambda^{-m} e^{\phi_m((x_n)_n)}} \leq c_2,$$

ou ainda,

$$c_1 \leq \frac{\mu(E)}{\exp(-m\log\lambda + \phi_m((x_n)_n))} \leq c_2, \tag{5.31}$$

que é exatamente a condição de Gibbs com $P = \log\lambda$.

Agora veremos que este estado de Gibbs é único. Suponha que $\mu'$ é uma outra medida satisfazendo

$$c'_1 \leq \frac{\mu'(E)}{\exp(-mP' + \phi_m((x_n)_n))} \leq c'_2. \tag{5.32}$$

Para $(x_n)_n \in \Sigma_A$ seja $E_m((x_n)_n) = \{(y_n)_n \in \Sigma_A : x_i = y_i \text{ para todo } i = 0, \ldots, m-1\}$. Pela compacidade de $\Sigma_A$, considere $T_m$ um subconjunto finito de $\Sigma_A$ de modo que $\Sigma_A = \bigcup_{(x_n)_n \in T_m} E_m((x_n)_n)$ seja uma união disjunta. Então

$$\begin{aligned} c_1' e^{-P'm} \sum_{(x_n)_n \in T_m} e^{\phi_m((x_n)_n)} &\leq \sum_{(x_n)_n \in T_m} \mu'(E_m((x_n)_n)) \\ &= \mu'(\bigcup_{(x_n)_n \in T_m} E_m((x_n)_n)) = \mu'(\Sigma_A) = 1 \\ &\leq c_2' e^{-P'm} \sum_{(x_n)_n \in T_m} e^{\phi_m((x_n)_n)}. \end{aligned}$$

Disso segue que

$$c_1' \sum_{(x_n)_n \in T_m} e^{\phi_m((x_n)_n)} \leq e^{P'm} \leq c_2' \sum_{(x_n)_n \in T_m} e^{\phi_m((x_n)_n)}.$$

Então

$$\log c_1' + \log\left(\sum_{(x_n)_n \in T_m} e^{\phi_m((x_n)_n)}\right) \leq\ P'm\ \leq \log c_2' + \log\left(\sum_{(x_n)_n \in T_m} e^{\phi_m((x_n)_n)}\right)$$

e vemos que $P' = \lim_{m\to\infty} \dfrac{1}{m} \log\left(\sum_{(x_n)_n \in T_m} e^{\phi_m((x_n)_n)}\right)$. O mesmo raciocínio pode ser aplicado a $\mu$ e encontraremos $P = \lim_{m\to\infty} \dfrac{1}{m} \log\left(\sum_{(x_n)_n \in T_m} e^{\phi_m((x_n)_n)}\right)$, ou seja, $P = P'$.

Das estimativas em (5.31) e (5.32), e sabendo que $P = P'$, temos:

$$\mu'(E_m((x_n)_n)) \leq c_2' e^{-P'm} e^{\phi_m((x_n)_n)} = \frac{c_2'}{c_1} c_1 e^{-Pm} e^{\phi_m((x_n)_n)} \leq \frac{c_2'}{c_1} \mu(E_m((x_n)_n)).$$

Como a álgebra das uniões finitas e disjuntas dos cilindros $E_m((x_n)_n)$ é geradora da $\sigma$-álgebra dos conjuntos mensuráveis, tomando limites na desigualdade acima, conseguimos que $\mu'(E) \leq \dfrac{c_2'}{c_1} \mu(E)$ para qualquer Boreliano $E$. Em particular, se $\mu(E) = 0$, então $\mu'(E) = 0$, ou seja $\mu'$ é absolutamente contínua com respeito a $\mu$. Pelo Teorema de Radón-Nikodym $\mu' = f\mu$ para alguma função $f$ $\mu$-integrável. Tomando $E$ um conjunto mensurável qualquer e usando que $\mu$ é invariante por $\sigma_A$,

$$\int_E f\mu = \mu'(E) = \int_E d\mu' = \int_E f d\mu = \int_E f \circ \sigma_A d\mu.$$

Como a derivada de Radón-Nikodym é única em quase todo ponto, $f = f \circ \sigma_A$ em $\mu$-quase todo ponto. Como $\mu$ é misturadora, segue que é ergódica e isto dá que $f$

é constante em $\mu$-quase todo ponto, por ser uma função invariante por $\sigma_A$ (lembre da definição equivalente de ergodicidade dada na definição 2.5). Seja $c$ esta constante. Então

$$1 = \mu'(\Sigma_A) = \int f d\mu = \int c d\mu = c.$$

Assim, $\mu = \mu'$. $\square$

É importante notar que no enunciado do Teorema 5.4, $\phi$ é tomada em $\mathcal{F}_A$, enquanto que, até agora, consideramos $\phi \in \mathcal{F}_A \cap \mathcal{C}(\Sigma_A^+)$ como no Teorema de Ruelle-Perron-Frobenius. Vamos mostrar agora que $\mu$ é o único estado de Gibbs para $\phi \in \mathcal{F}_A$:

**Lema 5.21.** *$\mu$ é o único estado de Gibbs para $\phi \in \mathcal{F}_A$.*

*Demonstração.* A estratégia é mostrar que, qualquer $\phi \in \mathcal{F}_A$ é cohomóloga a uma $\psi \in \mathcal{F}_A \cap \mathcal{C}(\Sigma_A^+)$. Então como $\mu$ é único estado de Gibbs para $\psi \in \mathcal{F}_A \cap \mathcal{C}(\Sigma_A^+)$ pelo Lema 5.20, é possível estender o resultado a $\phi$ usando a relação de cohomologia sem alterar a medida $\mu$. Faremos isso por meio dos dois lemas seguintes.

**Lema 5.22 (Sinai).** *Se $\phi \in \mathcal{F}_A$, então $\phi$ é cohomóloga a alguma $\psi \in \mathcal{F}_A$ com $\psi(\{x_i\}_{i=-\infty}^{\infty}) = \psi(\{y_i\}_{i=-\infty}^{\infty})$ sempre que $x_i = y_i$ para $i \geq 0$.(Ou seja, $\psi \in \mathcal{F}_A \cap \mathcal{C}(\Sigma_A^+)$).*

*Demonstração.* Note que a condição que $\psi((x_i)_{i=-\infty}^{\infty}) = \psi((y_i)_{i=-\infty}^{\infty})$ sempre que $x_i = y_i$ para $i \geq 0$ diz exatamente que $\psi \in \mathcal{F}_A \cap \mathcal{C}(\Sigma_A^+)$, pois o valor de $\psi(\{x_i\}_{i=-\infty}^{\infty})$ depende apenas da sequência $\{x_i\}_{i=0}^{\infty}$.

Para cada $1 \leq t \leq n$, escolha $\{a_{k,t}\}_{k=-\infty}^{\infty} \in \Sigma_A$ com $a_{0,t} = t$, defina $r : \Sigma_A \to \Sigma_A$ por $r((x_k)_k) = (x_k^*)_k$ onde

$$x_k^* = \begin{cases} x_k & \text{para } k \geq 0 \\ a_{k,x_0} & \text{para } k \leq 0. \end{cases}$$

Seja

$$u((x_k)_k) = \sum_{j=0}^{\infty} (\phi(\sigma_A^j((x_k)_k)) - \phi(\sigma_A^j(r(x_k)_k))).$$

Uma vez que $\sigma_A^j((x_k)_k)$ e $\sigma_A^j(r(x_k)_k)$ coincidem de $-j$ a $= \infty$,

$$|\phi(\sigma_A^j((x_k)_k)) - \phi(\sigma_A^j(r((x_k)_k)))| \leq \operatorname{var}_j \phi \leq b\alpha^j.$$

Como $\sum_{j=0}^{\infty} b\alpha^j < \infty$, $u$ está bem definida e é contínua. Se $x_i = y_i$ para todo $|i| \leq n$, então, para $j \in \{0, \ldots, n\}$,

$$|\phi(\sigma_A^j((x_k)_k)) - \phi(\sigma_A^j((y_k)_k))| \leq \operatorname{var}_{n-j} \phi \leq b\alpha^{n-j}$$

e

$$|\phi(\sigma_A^j(r((x_k)_k))) - \phi(\sigma_A^j(r((y_k)_k)))|\phi \leq b\alpha^{n-j}.$$

Consequentemente

$$\begin{aligned}
|u((x_k)_k) - u((y_k)_k)| &\leq \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} |\phi(\sigma_A^j((x_k)_k)) - \phi(\sigma_A^j((y_k)_k)) + \phi(\sigma_A^j(r((x_k)_k))) \\
&\qquad - \phi(\sigma_A^j(r((y_k)_k)))| + 2b \sum_{j > \lfloor \frac{n}{2} \rfloor} \alpha^j \\
&\leq 2b \left( \sum_{j=0}^{\lfloor \frac{n}{2} \rfloor} \alpha^{n-j} + \sum_{j > \lfloor \frac{n}{2} \rfloor} \alpha^j \right) \\
&\leq \frac{4b\alpha^{\lfloor \frac{n}{2} \rfloor}}{1-\alpha}.
\end{aligned}$$

Isto mostra que $u \in \mathcal{F}_A$. Consequentemente, $\psi := \phi - u + (u \circ \sigma_A) \in \mathcal{F}_A$. Além disso,

$$\begin{aligned}
\psi((x_k)_k) &= \phi((x_k)_k) + \sum_{j=-1}^{\infty} (\phi(\sigma_A^{j+1}(r((x_k)_k))) - \phi(\sigma_A^{j+1}((x_k)_k))) \\
&\qquad + \sum_{j=0}^{\infty} (\phi(\sigma_A^{j+1}((x_k)_k)) - \phi(\sigma_A^j(r(\sigma_A((x_k)_k))))) \\
&= \phi(r((x_k)_k)) + \sum_{j=0}^{\infty} (\phi(\sigma_A^{j+1}(r((x_k)_k))) - \phi(\sigma_A^j(r(\sigma_A((x_k)_k))))).
\end{aligned}$$

A expressão final depende apenas de $\{x_i\}_{i=0}^{\infty}$, como desejávamos. □

O seguinte lema afirma que para $\psi$ cohomólogo a $\phi$, a medida $\mu$ que construímos para $\phi$ é a mesma que a que resultaria se construíssemos $\mu$ para $\psi$.

**Lema 5.23.** *Suponha que $\phi$ cohomólogo a $\psi$ e que o Lema 5.21 é verdadeiro para $\phi$. Então este deve valer para $\psi$ e $\mu_\phi = \mu_\psi$.*

*Demonstração.* A relação de cohomologia diz que $\phi - \psi = u - (u \circ \sigma_A)$. Então

$$\begin{aligned}|\phi_m((x_n)_n) - \psi_m((x_n)_n)| &= \left|\sum_{k=0}^{m-1} \phi(\sigma_A^k((x_n)_n)) - \sum_{k=0}^{m-1} \psi(\sigma_A^k((x_n)_n))\right| \\ &= \left|\sum_{k=0}^{m-1} u(\sigma_A^{k+1}((x_n)_n)) - u(\sigma_A^k((x_n)_n))\right| \\ &= |u(\sigma_A^m((x_n)_n)) - u((x_n)_n)| \leq 2||u||.\end{aligned}$$

Daí,

$$\frac{e^{-2||u||}}{\exp(-mP + \phi_m((x_n)_n))} \leq \frac{1}{\exp(-mP + \psi_m((x_n)_n))} \leq \frac{e^{2||u||}}{\exp(-mP + \phi_m((x_n)_n))}$$

e o Lema 5.21 vale para $\psi$, pois da definição de estado de Gibbs para $\phi$,

$$\tilde{c}_1 = c_1 e^{-2||u||} \leq \frac{\mu(E)}{\exp(-mP + \phi_m((x_n)_n))} \leq c_2 e^{2||u||} = \tilde{c}_2.$$

Note as constantes da definição de estado de Gibbs mudam, porém para $P$ e $\mu$ isso não ocorre. □

O Lema 5.21 fica demonstrado, porque ele é verdadeiro para $\phi \in \mathcal{F}_A \cap \mathcal{C}(\Sigma_A^+)$. Pelo Lema 5.22, existe $\psi \in \mathcal{F}_A$ que é cohomólogo a $\phi$ e finalmente pelo Lema 5.23, concluímos que o Lema 5.21 vale para $\psi \in \mathcal{F}_A$ e a medida permanece a mesma. □

**Lema 5.24.** *$\mu = \mu_\phi$ é o único estado de equilíbrio para $\phi \in \mathcal{F}_A$.*

*Demonstração.* Segue do princípio variacional que

$$P(\sigma_A, \phi) = \sup\left\{h_\nu(\sigma_A) + \int \phi d\nu : \nu \in \mathcal{M}_1(\sigma_A)\right\} \geq h_\mu(\sigma_A) + \int \phi d\mu. \tag{5.33}$$

Provaremos agora que $P(\sigma_A, \phi) \leq h_\mu(\sigma_A) + \int \phi d\mu$.

Considere a cobertura aberta $\Upsilon$ de $\Sigma_A$ formada pelos cilindros $E_i = \{(y_n)_n \in \Sigma_A : i = y_0\}$ para $0 \leq i < n$. Como esta cobertura é ótima,

$$P_m(\sigma_A, \phi, \Upsilon) = \sum_{C \in \Upsilon^m} e^{\phi_m(C)}$$

onde $C$ é o cilindro, $C = \{(y_n)_n \in \Sigma_A : a_i = y_i \text{ para todo } 0 \leq i < m\}$, os $a_i$ são números inteiros entre 1 e $n$, satisfazendo $A_{a_k, a_{k+1}} = 1$ e

$$\phi_m(C) = \sup\{\phi_m((x_n)_n) : (x_n)_n \in C\}.$$

Daí, relembre que a pressão é o número

$$P(\sigma_A, \phi) = P(\sigma_A, \phi, \Upsilon) = \lim_{m\to\infty} \frac{1}{m} \log P_m(\sigma_A, \phi, \Upsilon).$$

Dado o cilindro $C$, escolha $(x_n)_n$ com $x_i = a_i$ $(i = 0, \ldots, m-1)$ e

$$\phi_m((x_n)_n) = \phi_m(C).$$

Como $\mu = \mu_\phi$ é estado de Gibbs,

$$c_1 \leq \frac{\mu(C)}{\exp(-Pm + \phi_m((x_n)_n))} \leq c_2.$$

Então somando, sobre todos os $C \in \Upsilon^m$,

$$\begin{aligned}
c_1 \leq \frac{\sum_{C\in\Upsilon^m} \mu(C)}{\exp(-Pm) \sum_{C\in\Upsilon^m} \exp(\phi_m((x_n)_n)))} \leq c_2 \;\Rightarrow \\
c_1 \leq \frac{\mu(\Sigma_A)}{\exp(-Pm) P_m(\sigma_A, \phi, \Upsilon)} \leq c_2 \;\Rightarrow \\
c_1 \leq \frac{1}{\exp(-Pm) P_m(\sigma_A, \phi, \Upsilon)} \leq c_2 \;\Rightarrow \\
c_1 \leq \frac{\exp(Pm)}{P_m(\sigma_A, \phi, \Upsilon)} \leq c_2 \;\Rightarrow \\
c_2^{-1} \leq \frac{P_m(\sigma_A, \phi, \Upsilon)}{\exp(Pm)} \leq c_1^{-1} \;\Rightarrow \\
c_2^{-1}\exp(Pm) \leq P_m(\sigma_A, \phi, \Upsilon) \leq c_1^{-1}\exp(Pm).
\end{aligned}$$

Portanto

$$\begin{aligned}
P = \lim_m \frac{1}{m} \left(\log c_2^{-1} + Pm\right) &= \lim_m \frac{1}{m} \left(\log c_2^{-1}\exp(Pm)\right) \\
&\leq \lim_m \frac{1}{m} \log P_m(\sigma_A, \phi, \Upsilon) \leq \lim_m \frac{1}{m} \left(\log c_1^{-1}\exp(Pm)\right) \\
&= \lim_m \frac{1}{m} \left(\log c_1^{-1} + Pm\right) = P
\end{aligned}$$

e assim, $P(\sigma_A, \phi) = \lim_m \dfrac{1}{m} \log P_m(\sigma_A, \phi, \Upsilon) = P$.

Se $y_i = x_i$ para todo $i = 0, \ldots, m-1$, então, lembrando que $\phi \in \mathcal{F}_A$ implica que

$\operatorname{var}_k \phi = b\alpha^k$, onde $k \geq 0$ e $b$, $\alpha \in (0,1)$, segue que

$$\begin{aligned}
|\phi_m((y_n)_n) - \phi_m((x_n)_n)| &\leq \sum_{k=0}^{m-1} |\phi(\sigma_A^k((y_n)_n)) - \phi(\sigma_A^k((x_n)_n))| \\
&\leq \operatorname{var}_0 \phi + \operatorname{var}_1 \phi + \cdots + \operatorname{var}_{\lfloor \frac{m}{2} \rfloor} \phi \\
&\qquad + \operatorname{var}_{\lfloor \frac{m}{2} \rfloor + 1} \phi + \cdots + \operatorname{var}_{m-1} \phi \\
&\leq \operatorname{var}_0 \phi + \operatorname{var}_1 \phi + \cdots + \operatorname{var}_{\lfloor \frac{m}{2} \rfloor} \phi \\
&\qquad + \operatorname{var}_{m-\lfloor \frac{m}{2} \rfloor} \phi + \cdots + \operatorname{var}_0 \phi \\
&\leq \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor} b\alpha^k + \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor} b\alpha^k \\
&= 2b \sum_{k=0}^{\lfloor \frac{m}{2} \rfloor} \alpha^k \leq \frac{2b}{1-\alpha} := p.
\end{aligned} \tag{5.34}$$

Consequentemente, se $(x_n)_n \in C$, pelo que vimos acima, $\phi_m((x_n)_n) - p \leq \phi_m((y_n)_n)$ e

$$\begin{aligned}
-\mu(C)\log\mu(C) + \int_C \phi_m d\mu &\geq -\mu(C)\log\mu(C) + \mu(C)(\phi_m((x_n)_n) - p) \\
&= -\mu(C)[\log\mu(C) - \phi_m((x_n)_n) + p] \\
&\geq -\mu(C)[\log\left(c_2 e^{-Pm + \phi_m((x_n)_n)}\right) - \phi_m((x_n)_n) + p] \\
&\geq -\mu(C)[\log c_2 - Pm + \phi_m((x_n)_n) - \phi_m((x_n)_n) + p] \\
&\geq \mu(C)[Pm - \log c_2 - p].
\end{aligned}$$

Notando que $\mathscr{P} = \{E_1, \ldots, E_n\}$ é uma partição geradora, escrevemos

$$\begin{aligned}
H_\mu(\mathscr{P}^m) + \int \phi_m d\mu &= \sum_{C \in \mathscr{P}^m} -\mu(C)\log\mu(C) + \sum_{C \in \mathscr{P}^m} \int_C \phi_m d\mu \\
&= \sum_{C \in \mathscr{P}^m} \left(-\mu(C)\log\mu(C) + \int_C \phi_m d\mu\right) \\
&\geq \sum_{C \in \mathscr{P}^m} \mu(C)[Pm - \log c_2 - p] \\
&= Pm - \log c_2 - p.
\end{aligned}$$

Então

$$h_\mu(\sigma_A) + \int \phi d\mu = \lim_{m\to\infty} \frac{1}{m}\left(H_\mu(\mathscr{P}^m) + \int \phi_m d\mu\right)$$
$$\geq \lim_{m\to\infty} \frac{1}{m}\left(Pm - \log c_2 - p\right) = P = P(\sigma_A, \phi).$$

Destacamos que

$$\int \phi d\mu = \lim_{m\to\infty} \frac{1}{m} \int \phi_m d\mu$$

é consequência de $f$ preservar a medida $\mu$. Juntando a desigualdade acima com (5.33), Concluímos que

$$P(\sigma_A, \phi) = h_\mu(\sigma_A) + \int \phi d\mu.$$

Logo $\mu$ é um estado de equilíbrio para $\phi$.

Na continuação, para provar a unicidade, precisaremos de alguns lemas.

**Lema 5.25.** *Seja $X$ um espaço métrico compacto, $\nu$ uma probabilidade em $X$ e $\mathscr{Q} = \{Q_1, \ldots, Q_n\}$ uma partição de $X$. Suponha que $\{\mathscr{P}_m\}_{m=1}^{\infty}$ é uma sequência de partições de modo que* $\operatorname{diam}(\mathscr{P}_m) = \max_{P\in\mathscr{P}_m} \operatorname{diam}(P) \to 0$ *quando $m \to \infty$. Então existem partições $\{E_1^m, \ldots, E_n^m\}$ tais que cada $E_i^m$ é uma união de membros de $\mathscr{P}_m$ e* $\lim_{m\to\infty} \mu(E_i^m \Delta Q_i) = 0$ *para cada $i$.*

*Demonstração.* Escolha compactos $K_1, \ldots, K_n$ com $K_i \subset Q_i$ e $\mu(Q_i \setminus K_i) < \varepsilon$. Seja $\delta = \inf_{i\neq j} d(K_i, K_j)$ e considere $m$ tal que $\operatorname{diam}(\mathscr{P}_m) \leq \frac{\delta}{2}$. Divida os elementos $P \in \mathscr{P}_m$ em grupos cujas uniões são $\{E_1^m, \ldots, E_n^m\}$ de modo que

$$P \subset E_i^m \text{ se } P \cap K_i \neq \emptyset.$$

Como $\operatorname{diam}(\mathscr{P}_m) \leq \frac{\delta}{2}$, qualquer $P \in \mathscr{P}_m$ pode intersectar no máximo um $K_i$. Ponha os elementos $P$ que não intersectam nenhum $K_i$ em qualquer $E_i^m$. Então $E_i^m \supset K_i$ e

$$\mu(E_i^m \Delta Q_i) = \mu(Q_i \setminus E_i^m) + \mu(E_i^m \setminus Q_i)$$
$$\leq \mu(Q_i \setminus K_i) + \sum_{i=1}^{n} \mu(E_i^m \setminus K_i)$$
$$\leq \varepsilon + \mu\left(X \setminus \bigcup_{i=1}^{n} K_i\right) \leq (n+1)\varepsilon.$$

$\square$

**Lema 5.26.** *Sejam $a_1, \ldots, a_m$ números reais e sejam $p_1, \ldots, p_m$ números não negativos tais que $s = p_1 + \cdots + p_m \leq 1$. Então,*

$$\sum_{i=1}^{m} p_i(a_i - \log p_i) \leq s\left(\log \sum_{i=1}^{m} e^{a_i} - \log s\right). \tag{5.35}$$

*Demonstração.* Escreva $t_i = \dfrac{p_i}{s}$ e note que $\sum_{i=1}^{m} t_i = 1$. Use a equação (3.62) do Lema 3.58 para escrever

$$\sum_{i=1}^{m} t_i(a_i - \log t_i) \leq \log \sum_{i=1}^{m} e^{a_i}.$$

Trocando $t_i$ por $\dfrac{p_i}{s}$,

$$\sum_{i=1}^{m} \frac{p_i}{s}(a_i - \log p_i + \log s) \leq \log \sum_{i=1}^{m} e^{a_i}$$

ou ainda

$$\sum_{i=1}^{m} p_i(a_i - \log p_i) \leq s \log \sum_{i=1}^{m} e^{a_i} - \sum_{i=1}^{m} p_i \log s = s \log \sum_{i=1}^{m} e^{a_i} - s \log s$$

como queríamos. $\square$

Agora demonstramos a unicidade de $\mu = \mu_\phi$.

Seja $\eta \in \mathcal{M}_1(\sigma_A)$ satisfazendo $h_\eta(\sigma_A) + \int \phi d\eta = P$. Primeiro suponha que $\eta$ seja singular em relação a $\mu$. Então existe um conjunto mensurável $B$ com $\sigma_A(B) = B$, $\mu(B) = 0$ e $\eta(B) = 1$. Seja $\mathscr{P} = \{P_1, \ldots, P_n\}$, onde $P_i = \{(x_n)_n \in \Sigma_A : x_0 = i\}$ e seja $\mathscr{Q}_m = \sigma_A^{-\lfloor \frac{m}{2} \rfloor + 1}(\mathscr{P}) \vee \cdots \vee \mathscr{P} \vee \cdots \vee \sigma_A^{\lfloor \frac{m}{2} \rfloor}(\mathscr{P})$. Então $\operatorname{diam}(\mathscr{Q}_m) \to 0$ (usamos a métrica que definimos na observação 5.2). Aplicando o Lema 5.25 a partição $\{B, \Sigma_A \backslash B\}$, encontramos conjuntos $E^m$ que são formados por uniões de elementos de $\mathscr{Q}_m$ e que satisfazem $(\mu + \eta)(B \Delta E^m) \to 0$ quando $m \to \infty$. Como $\mu + \eta$ é $\sigma_A$-invariante e $\sigma_A^{-m + \lfloor \frac{m}{2} \rfloor}(B) = B$, temos que $(\mu + \eta)(B \Delta F^m) \to 0$ onde $F^m = \sigma_A^{-m + \lfloor \frac{m}{2} \rfloor}(E^m)$ é uma união de membros de $\mathscr{P} \vee \cdots \vee \sigma_A^{-m+1}(\mathscr{P})$. Como $h_\eta(\sigma_A) = \lim_m \dfrac{1}{m} H_\eta(\mathscr{P}^m) = \inf_m \dfrac{1}{m} H_\eta(\mathscr{P} \vee \cdots \vee \sigma_A^{-m+1}(\mathscr{P}))$, temos que

$$P = P(\sigma_A, \phi) = h_\eta(\sigma_A) + \int \phi d\eta \leq \frac{1}{m}\left(H_\eta(\mathscr{P}^m) + \int \phi_m d\eta\right)$$

ou

$$Pm \leq \sum_{B\in\mathscr{P}^m}\left(-\eta(B)\log\eta(B)+\int_B \phi_m d\eta\right).$$

Resulta de (5.34) que para $(x_n)_n \in B \in \mathscr{P}^m$, $\phi_m \leq \phi_m((x_n)_n)+p$ em $B$ e assim

$$\begin{aligned} Pm &\leq p + \sum_{B\in\mathscr{P}^m}\eta(B)\left(-\log\eta(B)+\phi_m((x_n)_n)\right)\\ &\leq p+\sum_{B\subset F^m}\eta(B)\left(-\log\eta(B)+\phi_m((x_n)_n)\right)\\ &\qquad +\sum_{B\subset\Sigma_A\setminus F^m}\eta(B)\left(-\log\eta(B)+\phi_m((x_n)_n)\right). \end{aligned}$$

Aplicando o Lema 5.26 às somas da última desigualdade,

$$\begin{aligned} Pm - p &\leq \eta(F^m)\sum_{B\subset F^m}\left(\phi_m((x_n)_n)-\log\eta(B)\right)\\ &\qquad + \eta(\Sigma_A\setminus F^m)\sum_{B\subset\Sigma_A\setminus F^m}\left(\phi_m((x_n)_n)-\log\eta(B)\right)\\ &\leq \eta(F^m)\log\left(\sum_{B\subset F^m}\exp\phi_m((x_n)_n)\right)\\ &\qquad + \eta(\Sigma_A\setminus F^m)\log\left(\sum_{B\subset\Sigma_A\setminus F^m}\exp\phi_m((x_n)_n)\right)+2K^* \end{aligned}$$

onde $K^*=\sup_{0\leq s\leq 1}(-s\log s)$. Rearranjando os termos e usando que $\mu$ é um estado de

Gibbs

$$
\begin{aligned}
-2K^* - p \leq\ & \eta(F^m) \log \left( \sum_{B \subset F^m} \exp \phi_m((x_n)_n) \right) \\
& + \eta(\Sigma_A \setminus F^m) \log \left( \sum_{B \subset \Sigma_A \setminus F^m} \exp \phi_m((x_n)_n) \right) + (-Pm) \\
=\ & \eta(F^m) \log \left( \sum_{B \subset F^m} \exp \phi_m((x_n)_n) \right) \\
& + \eta(\Sigma_A \setminus F^m) \log \left( \sum_{B \subset \Sigma_A \setminus F^m} \exp \phi_m((x_n)_n) \right) \\
& + \eta(F^m)(-Pm) + \eta(\Sigma_A \setminus F^m)(-Pm) \\
=\ & \eta(F^m) \log \left( \sum_{B \subset F^m} \exp(\phi_m((x_n)_n) - Pm) \right) \\
& + \eta(\Sigma_A \setminus F^m) \log \left( \sum_{B \subset \Sigma_A \setminus F^m} \exp(\phi_m((x_n)_n) - Pm) \right) \\
\leq\ & \eta(F^m) \log \left( \sum_{B \subset F^m} c_2^{-1} \mu(B) \right) + \eta(\Sigma_A \setminus F^m) \log \left( \sum_{B \subset \Sigma_A \setminus F^m} c_2^{-1} \mu(B) \right) \\
\leq\ & \log c_2^{-1} + \eta(F^m) \log \mu(F^m) + \eta(\Sigma_A \setminus F^m) \log \mu(\Sigma_A \setminus F^m).
\end{aligned}
$$

Fazendo $m \to \infty$, $\eta(F^m) \to 1$, $\mu(F^m) \to 0$. Assim a desigualdade acima é contrariada.

Em geral, para $\eta' \in \mathcal{M}_1(\sigma_A)$, escreva $\eta' = \beta\eta + (1-\beta)\mu'$ onde $\beta \in (0,1)$, $\eta \in \mathcal{M}_1(\sigma_A)$ é singular com respeito a $\mu$ e $\mu' \in \mathcal{M}_1(\sigma_A)$ é absolutamente contínua com respeito a $\mu$. Como $\eta$ e $\mu'$ são suportadas em conjuntos de disjuntos, a pressão de $\phi$ com respeito a medida $\eta'$ é a soma das pressões com respeito às medidas $\eta$ e $\mu'$ respectivamente:

$$
P_{\eta'}(\sigma_A, \phi) = \beta P_\eta(\sigma_A, \phi) + (1-\beta) P_{\mu'}(\sigma_A, \phi)
$$

onde o $P_\eta(\sigma_A, \phi) = h_\eta(\sigma_A) + \int \phi d\eta$. Suponha que $P_{\eta'}(\sigma_A, \phi) = P$. Como $P_\eta(\sigma_A, \phi) \leq P$ e $P_{\mu'}(\sigma_A, \phi) \leq P$ pelo princípio variacional, segue que $P_\eta(\sigma_A, \phi) = P$ ou $\beta = 0$. Mas provamos no início que $P_\eta(\sigma_A, \phi)$ não pode ser $P$. Assim $\beta = 0$ e $\eta' = \mu'$. Como $\eta' = \mu'$ é absolutamente contínua com respeito a $\mu$ por hipótese, escreva $\eta' = \dfrac{d\eta'}{d\mu}\mu$. Então $\dfrac{d\eta'}{d\mu}$ é

$\sigma_A$-invariante porque $\eta'$ e $\mu$ o são. Como a derivada de Radón-Nikodym é única em quase todo ponto, segue da ergodicidade de $\mu$ que $\dfrac{d\eta'}{d\mu}$ é constante em quase todo ponto e essa constante é 1:

$$1 = \eta'(\Sigma_A) = \int \frac{d\eta'}{d\mu} d\mu = c \int d\mu = c.$$

Portanto, $\eta' = \mu$. Isto finaliza a prova da unicidade do estado de equilíbrio $\mu = \mu_\phi$. □

Temos também uma versão do Teorema de Livsič para $\sigma_A$:

**Teorema 5.27.** *Seja $\sigma_A : \Sigma_A \to \Sigma_A$ uma transformação topologicamente misturadora e sejam $\phi$, $\psi \in \mathcal{F}_A$. As seguintes condições são equivalentes:*

*(a)* $\mu_\phi = \mu_\psi$*;*

*(b) existe* $c \in \mathbb{R}$ *tal que* $\phi_m((x_n)_n) - \psi_m((x_n)_n) = \sum_{i=0}^{m-1} \phi(\sigma_A^i((x_n)_n)) - \sum_{i=0}^{m-1} \psi(\sigma_A^i((x_n)_n)) = cm$ *para todo* $(x_n)_n \in \mathrm{Fix}(\sigma_A^m)$ *;*

*(c) existe* $c \in \mathbb{R}$ *e uma função qualquer* $u \in \mathcal{F}_A$ *tal que*

$$\phi - \psi = c + u \circ \sigma_A - u;$$

*(d) existem constantes* $c$, $L \in \mathbb{R}$ *de modo que* $|\phi_m((x_n)_n) - \psi_m((x_n)_n) - cm| \leq L$ *para todo* $(x_n)_n \in \Sigma_A$ *e* $m > 0$.

*Se estas condições valem, então* $c = P(\sigma_A, \phi) = P(\sigma_A, \psi)$.

*Demonstração.* Veja o Teorema 1.28 de [9]. □

Veja que pelos resultados acima, ampliamos a classe de funções para as quais vale a existência e unicidade do estado de equilíbrio. No Teorema de Ruelle do capítulo anterior, provamos que o estado de equilíbrio existe e é único para a classe das transformações expansoras. Com o último resultado, provamos que o deslocamentos bilaterais de tipo finito (que não são expansores) possuem um único estado de equilíbrio para potenciais em $\mathcal{F}_A$.

No resto deste capítulo, vamos demonstrar a existência e unicidade de estados de equilíbrio para difeomorfismos Axioma A em conjuntos hiperbólicos, para potenciais Hölder contínuos.

## 5.3 Difeomorfismos Axioma A

Suporemos agora que $f : M \to M$ é um difeomorfismo $\mathcal{C}^\infty$ de uma variedade Riemaniana compacta $M$. Então a derivada de $f$ pode ser considerada como uma aplicação $df : TM \to TM$ onde $TM = \cup_{x\in M} T_xM$ é o fibrado tangente de $M$ e $df_x : T_xM \to T_{f(x)}M$.

**Definição 5.28.** Um subconjunto fechado $\Lambda \subset M$ é dito **hiperbólico** se $f(\Lambda) = \Lambda$ e cada espaço tangente $T_xM$ com $x \in \Lambda$ pode ser escrito como soma direta

$$T_xM = E_x^u \oplus E_x^s$$

de subespaços $E_x^u, E_x^s \subset TM$ de modo que

(a) $df(E_x^s) = E_{f(x)}^s, df(E_x^u) = E_{f(x)}^u$;

(b) Existem constantes $c > 0$ e $\lambda \in (0, 1)$ tais que

$$||df^n(v)|| \leq c\lambda^n ||v|| \text{ quando } v \in E_x^s \ , \ n \geq 0$$

e

$$||df^{-n}(v)|| \leq c\lambda^n ||v|| \text{ quando } v \in E_x^u \ , \ n \geq 0;$$

(c) $E_x^u, E_x^s$ variam continuamente com $x$.

Um ponto $x \in M$ é **não errante** se para toda vizinhança $U$ de $x$, existe $k$ positivo tal que $U \cap f^k(U)$ é não vazio. Caso contrário, dizemos que $x$ é **errante**.

O conjunto $\Omega = \Omega(f)$ de todos os pontos não errantes é fechado e como $f$ é difeomorfismo $f(\Omega) = \Omega$: de fato, o conjunto dos pontos errantes é aberto, pois se $x$ é errante, então existe vizinhança $U$ de $x$ tal que $U \cap \bigcup_{n>0} f^n(U) = \emptyset$. Afirmamos que todos os pontos de $U$ são errantes, dado $y \in U$ tome uma vizinhança $V$ de $y$ pequena o suficiente de forma que esteja contida em $U$. Assim $\left(V \cap \bigcup_{n>0} f^n(V)\right) \subset \left(U \cap \bigcup_{n>0} f^n(U)\right) = \emptyset$. Portanto, $\Omega$ é fechado. Agora $f(\Omega) \subset \Omega$: se $x \in \Omega$ e $U$ é uma vizinhança de $f(x)$, então $f^{-1}(U)$ é uma vizinhança de $x$. Portanto, existe $k$ tal que $f^k(f^{-1}(U)) \cap f^{-1}(U) \neq \emptyset$ e a imagem desta interseção por $f$ está contida em $f^k(U) \cap U$ que é portanto não vazio.

Como $f$ é difeomorfismo $f^{-1}(\Omega) \subset \Omega$: se $x \in \Omega$ e $U$ é uma vizinhança de $f^{-1}(x)$, então $f(U)$ é uma vizinhança de $x$. Portanto, existe $k$ tal que $f^k(f(U)) \cap f(U) \neq \emptyset$ e a imagem desta interseção por $f^{-1}$ está contida em $f^k(U) \cap U$ que é portanto não vazio. Assim,

$$\Omega = f(f^{-1}(\Omega)) \subset f(\Omega) \subset \Omega.$$

Um ponto $x$ é **periódico** se $f^n(x) = x$ para algum $n > 0$. Denotemos por $\mathrm{Per}(f)$ o conjunto dos pontos periódicos de $f$. Todo ponto periódico é não errante, portando $\mathrm{Per}(f) \subset \Omega(f)$. Como o conjunto $\Omega(f)$ é fechado $\Omega(f) \supset \overline{\mathrm{Per}(f)}$.

**Definição 5.29** (**Smale**)**.** Diremos que um difeomorfismo $f$ satisfaz o **Axioma A** se $\Omega(f)$ é hiperbólico e $\Omega(f) = \overline{\mathrm{Per}(f)}$, ou seja, o conjunto dos pontos não errantes do difeomorfismo $f$ é hiperbólico e é igual ao fecho do conjunto dos pontos periódicos de $f$.

A métrica Riemaniana de $M$ é usada para descrever a condição (b) da definição de conjunto hiperbólico. Na verdade esta condição não depende de qual métrica é usada, embora as constantes $c$ e $\lambda$ dependam. Uma métrica é **adaptada** (para um difeomorfismo Axioma A $f$) se $\Omega(f)$ é hiperbólico com respeito a esta com constante $c = 1$.

**Lema 5.30.** *Todo difeomorfismo Axioma A admite uma métrica adaptada.*

*Demonstração.* Este lema é devido a Mather. Uma prova pode ser encontrada na proposição 4.2 da referência [11]. □

De agora em diante, usaremos sempre a métrica adaptada. Para $x \in M$ definimos os seguintes conjuntos:

$$W^s(x) = \{y \in M : d(f^n x, f^n y) \to 0 \ \text{ quando } \ n \to \infty\}$$
$$W^s_\varepsilon(x) = \{y \in M : d(f^n x, f^n y) \leq \varepsilon \ \text{ para todo } \ n \geq 0\}$$
$$W^u(x) = \{y \in M : d(f^{-n} x, f^{-n} y) \to 0 \ \text{ quando } \ n \to \infty\}$$
$$W^u_\varepsilon(x) = \{y \in M : d(f^{-n} x, f^{-n} y) \leq \varepsilon \ \text{ para todo } \ n \geq 0\}.$$

Chamamos os conjuntos $W^s(x)$ e $W^u(x)$, respectivamente, de **variedades estável e instável do ponto** $x$. Chamamos os conjuntos $W^s_\varepsilon(x)$ e $W^u_\varepsilon(x)$, respectivamente, de **variedades estável e instável de raio $\varepsilon$ do ponto $x$.** Às vezes chamamos $W^s_\varepsilon(x)$, $W^u_\varepsilon(x)$ de **variedades estável e instável locais.**

O Teorema da variedade estável é o principal fato por trás do comportamento dos difeomorfismos Axioma A.

**Teorema 5.31** (**Variedade Estável**). *Seja $\Lambda$ um conjunto hiperbólico para um difeomorfismo $f$ de classe $\mathcal{C}^r$. Para $\varepsilon > 0$ pequeno são válidas as seguintes conclusões:*

*(a)* $W^s_\varepsilon(x)$, $W^u_\varepsilon(x)$ *são discos* $\mathcal{C}^r$ *para* $x \in \Lambda$ *com* $T_xW^s_\varepsilon(x) = E^s_x$ *e* $T_xW^u_\varepsilon(x) = E^u_x$*;*

*(b)* $d(f^n(x), f^n(y)) \leq \lambda^n d(x,y)$ *para* $y \in W^s_\varepsilon(x)$, $n \geq 0$, *e* $d(f^{-n}(x), f^{-n}(y)) \leq \lambda^n d(x,y)$ *para* $y \in W^u_\varepsilon(x)$, $n \geq 0$*;*

*(c)* $W^s_\varepsilon(x)$, $W^u_\varepsilon(x)$ *variam continuamente com* $x$ *(na topologia* $\mathcal{C}^r$*).*

*Demonstração.* Veja o Teorema 6.2 da referência [11]. □

**Teorema 5.32** (**Coordenadas canônicas**). *Suponha que $f$ é um difeomorfismo Axioma A. Para qualquer $\varepsilon > 0$ pequeno, existe $\delta > 0$ de modo que $W^s_\varepsilon(x) \cap W^u_\varepsilon(y)$ consiste de um único ponto que será denotado por $[x,y]$ sempre que $x, y \in \Omega(f)$ e $d(x,y) \leq \delta$. Além disso $[x,y] \in \Omega(f)$ e a função*

$$[\cdot,\cdot] : \{(x,y) \in \Omega(f) \times \Omega(f) \ : \ d(x,y) \leq \delta\} \to \Omega(f)$$

*é contínua.*

*Demonstração.* Veja a proposição 7.2 da referência [11]. □

**Corolário 5.33.** *Seja $\Lambda$ um conjunto hiperbólico. Então existe $\varepsilon > 0$ tal que $\Lambda$ é expansivo em $M$: se $x \in \Lambda$ e $y \in M$ com $y \neq x$, então*

$$d(f^k(x), f^k(y)) > \varepsilon \ \ \textit{para algum} \ \ k \in \mathbb{Z}.$$

*Demonstração.* Escolha $\varepsilon$ como no Teorema 5.32. Se for $d(f^k(x), f^k(y)) \leq \varepsilon$ para $x \in \Lambda$, $y \in M$ e todo $k \in \mathbb{Z}$, então devemos ter $y \in W^s_\varepsilon(x)$ e $y \in W^u_\varepsilon(x)$. Portanto $y \in W^s_\varepsilon(x) \cap W^u_\varepsilon(x)$. Como $d(x,x) = 0 < \delta$, segue do Teorema 5.32 que $W^s_\varepsilon(x) \cap W^u_\varepsilon(x)$ consiste de um único ponto. Como $x$ também pertence a essa interseção, então $x = y$. □

**Teorema 5.34 (Decomposição Espectral).** *Se $f$ é um difeomorfismo Axioma A, então podemos escrever $\Omega(f) = \Omega_1 \cup \Omega_2 \cup \cdots \cup \Omega_s$ onde os $\Omega_i$ são conjuntos fechados e dois a dois disjuntos com:*

*(a) $f(\Omega_i) = \Omega_i$ e $f|_{\Omega_i}$ é* ***topologicamente transitivo****: Para todo par de abertos não vazios $U$ e $V$ existe $k \geq 1$ tal que $f^{-k}(U)$ intersecta $V$;*

*(b) $\Omega_i = X_{1,i} \cup X_{2,i} \cup \cdots \cup X_{n_i,i}$ com os $X_{j,i}$ fechados e dois a dois, $f(X_{j,i}) = X_{j+1,i}$ $(X_{n_j+1,i} = X_{1,i})$ e $f^{n_i}|_{X_{j,i}}$ topologicamente misturadora.*

*Demonstração.* Veja o Teorema 3.5 da referência [9]. □

Os conjuntos $\Omega_i$ do Teorema da Decomposição Espectral de $\Omega(f)$ são chamados de **conjuntos básicos** de $f$. Note que se $g = f^n$ e $n$ é um múltiplo de cada $n_i$, então os conjuntos básicos de $g$ são os conjuntos $X_{j,i}$ e $g|_{X_{j,i}}$ é misturadora.

**Exemplo 5.35 (Ferradura de Smale).** Descreveremos um importante exemplo de Difeomorfismo Axioma A descoberto por Smale em 1960, quando este visitava o Rio de Janeiro. Começaremos esticando o quadrado $I \times I$ em uma barra (figura 2) por meio de um mapa linear, em seguida dobramos esta barra em formato de ferradura e a colocamos sobre $I \times I$ (figura 2).

Aqui, assumimos que as restrições de $f$ à $B_1$ e $B_2$ são mapas afins que esticam verticalmente e contraem horizontalmente. Agora, estendemos $f$ a uma mapeamento do disco $D^2$ em si mesmo adicionando dois semidiscos a $I \times I$ (figura 3)

A imagem de $f$ do semidisco inferior $T$ está contida em $T$ e a extensão de $f$ para $T$ pode ser escolhida contrativa. Portanto pelo Teorema do Ponto Fixo de Brower, $f|_T$ possui um único ponto fixo $p_1$, que é atrator, uma vez que se $x$ pertence a $T$, $S$ ou a $A \cup C \cup D$ então $\lim_{n\to\infty} f^n(x) = p_1$. De fato os pontos que podem não serem atraídos para $p_1$ residem em $B_1 \cup B_2$. Veremos que a ferradura de Smale é exatamente o conjunto dos pontos que não são atraídos para $P_1$. Agora adicionamos um ponto repulsor no pólo norte de $S^2$ e extendemos $f$ a um difeomorfismo de $S^2$ que continuaremos a chamar de

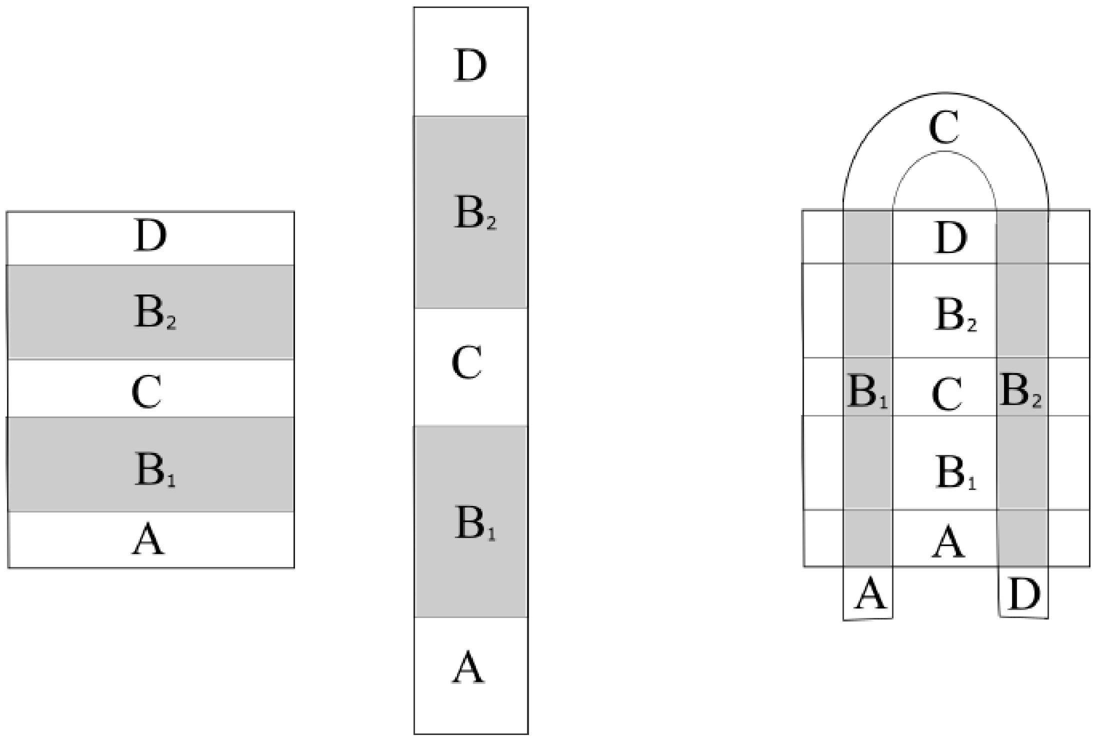


Figura 2: Primeira parte da construção da ferradura

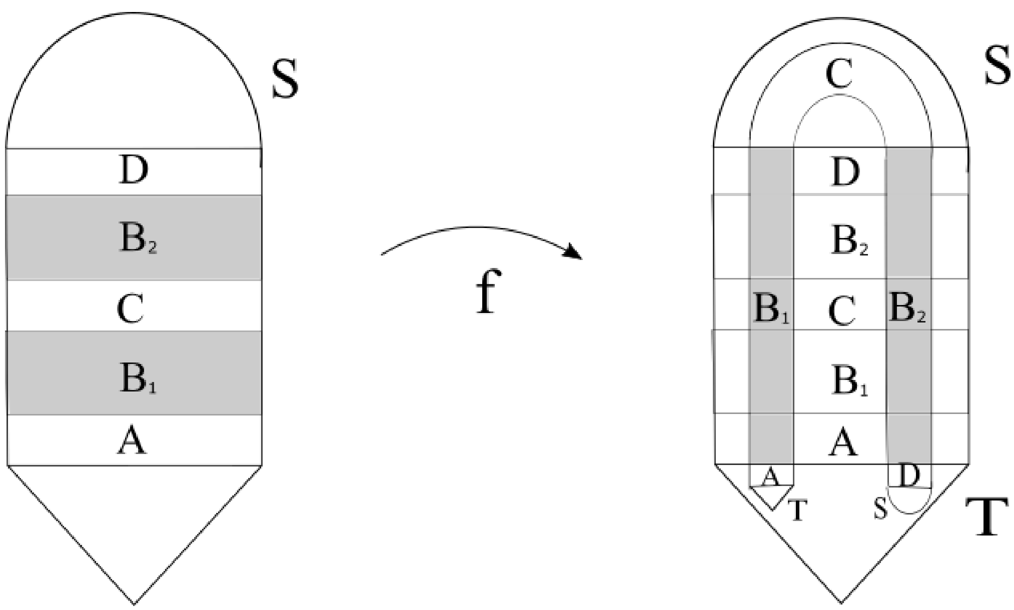


Figura 3: Extensão ao disco $D^2$

$f$. Defina as faixas horizontais

$$U_{11} = B_1 \cap f^{-1}(B_1)$$

$$U_{12} = B_1 \cap f^{-1}(B_2)$$

$$U_{21} = B_2 \cap f^{-1}(B_1)$$

$$U_{22} = B_2 \cap f^{-1}(B_2).$$

Recursivamente definimos $U_{i_0 \dots i_{n-1}} = B_{i_0} \cap f^{-1}(B_{i_1}) \cap \dots \cap f^{-(n-1)}(B_{i_{n-1}})$ (isto dá $2^n$ faixas horizontais) onde $i_0, i_1, \dots, i_{n-1} \in \{1, 2\}$. Analogamente defina as faixas verticais

$$V_{11} = B_1 \cap f(B_1)$$

$$V_{12} = B_1 \cap f(B_2)$$

$$V_{21} = B_2 \cap f(B_1)$$

$$V_{22} = B_2 \cap f(B_2)$$

e $V_{i_0 \dots i_{n-1}} = B_{i_0} \cap f(B_{i_1}) \cap \dots \cap f^{(n-1)}(B_{i_{n-1}})$ (isto dá $2^n$ faixas verticais) onde $i_0, i_1, \dots, i_{n-1} \in \{1, 2\}$. A figura 4 mostra os conjuntos $V_{ij}$ e $U_{ij}$

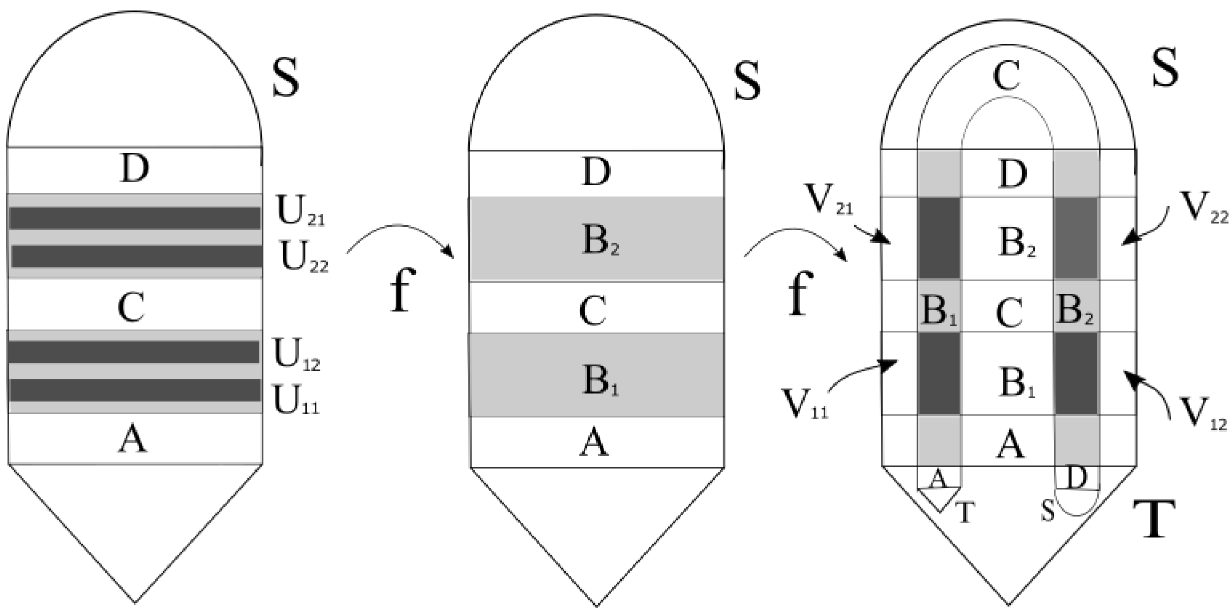


Figura 4: Esboço dos conjuntos $V_{ij}$ e $U_{ij}$

Tome

$$\Lambda = \left( \bigcap_{n \geq 0} \left[ \bigcap_{i_0, \dots, i_{n-1}} U_{i_0 \dots i_{n-1}} \right] \right) \cap \left( \bigcap_{n \geq 0} \left[ \bigcap_{i_0, \dots, i_{n-1}} V_{i_0 \vdots i_{n-1}} \right] \right) = \{ z : f^n(z) \in B_1 \cup B_2 \ \forall n \in \mathbb{Z} \}.$$

Este conjunto $\Lambda$ é a ferradura de Smale. Note que $\Lambda$ é interseção de dois conjuntos de Cantor. Por construção $\Lambda$ é compacto e $f$-invariante. Além disso, faixas horizontais são contraídas enquanto que as verticais são expandidas (as fibras $E^u$ são as verticais, enquanto que as fibras $E^s$ são horizontais), isto implica que $\Lambda$ é um conjunto hiperbólico para $f$. Veremos que $f|_\Lambda$ é um difeomorfismo Axioma A. Para cada $x \in \Lambda$, associamos uma sequência infinita de zeros e uns $(a_n)_n$ como se segue

$$a_n(x) = 0 \ \text{ se } \ f^n(x) \in V_1$$

$$a_n(x) = 1 \ \text{ se } \ f^n(x) \in V_2$$

onde $V_1$, $V_2$ estão na figura 5.

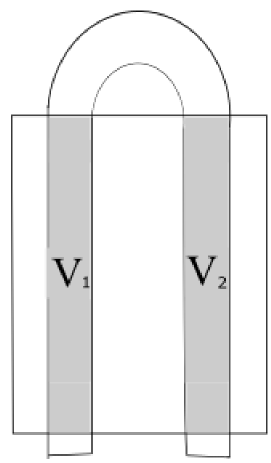


Figura 5: Conjuntos $V_1$ e $V_2$

Defina o mapa

$$\Phi : \Lambda \to \Sigma_2$$

tal que $\Phi(x) = (a_n(x))_n$. O leitor poderá ver em detalhes na referência [11] que o mapa $\Phi$ dá uma conjugação entre $f|_\Lambda$ e $\sigma$ (isto é, faz o diagrama da figura 6 comutar), onde $\sigma$ é a função deslocamento, dada por $\sigma((a_n)_n) = (a_{n+1})_n$.

$$\begin{tikzcd}
\Lambda \arrow[r, "f|_\Lambda"] \arrow[d, "\Phi"'] & \Lambda \arrow[d, "\Phi"] \\
\Sigma_2 \arrow[r, "\sigma"] & \Sigma_2
\end{tikzcd}$$

Figura 6: Conjugação entre $f|_\Lambda$ e $\sigma$

Sabemos que os pontos periódicos de $\Sigma_2$ são densos em $\Sigma_2$. Além disso, se $(a_n)_n$ é um ponto periódico para $\sigma$, então $\Phi^{-1}((a_n)_n)$ é um ponto periódico para $f|_\Lambda$. De fato, se $\sigma^k((a_n)_n) = (a_n)_n$, então $(f|_\Lambda)^k(\Phi^{-1}((a_n)_n)) = \Phi^{-1}(\sigma^k((a_n)_n)) = \Phi^{-1}((a_n)_n)$. Logo os pontos periódicos de $\Lambda$ são densos em $\Lambda$. Logo $\Lambda = \overline{\mathrm{Per}(f|_\Lambda)}$. Vimos no início da seção que $\Omega(f|_\Lambda) \supset \overline{\mathrm{Per}(f|_\Lambda)}$, por outro lado, temos que $\Omega(f|_\Lambda) \subset \Lambda = \overline{\mathrm{Per}(f|_\Lambda)}$. Como $\Lambda$ é hiperbólico, resulta que $\Omega(f|_\Lambda)$ também é hiperbólico e $\Omega(f|_\Lambda) = \overline{\mathrm{Per}(f|_\Lambda)}$. Resulta disso que $f|_\Lambda$ é um difeomorfismo Axioma A. Na verdade, $\Omega(f) = \Lambda \cup \{p_1\} \cup \{p_2\} = \overline{\mathrm{Per}(f)}$ é um conjunto hiperbólico (união de três conjuntos hiperbólicos) então o resultado é mais forte: $f$ satisfaz o Axioma A.

**Definição 5.36.** Uma sequência $(x_n)_n = \{x_i\}_{i=a}^b$ ($a = -\infty$ ou $b = +\infty$ são permitidos) de pontos de $M$ é uma **$\alpha$-pseudo-órbita** se

$$d(f(x_i), x_{i+1}) < \alpha \quad \text{para todo} \quad i = a, \ldots, b-1.$$

Um ponto $x \in M$ **$\beta$-sombreia** $(x_n)_n$ se

$$d(f^i(x), x_i) \leq \beta \quad \text{para todo} \quad i = a, \ldots, b.$$

**Lema 5.37** (**Lema do Sombreamento**)**.** *Para cada $\beta > 0$ existe $\alpha > 0$ de modo que cada $\alpha$-pseudo-órbita $\{x_i\}_{i=a}^b$ em $\Omega$ (isto é, cada $x_i \in \Omega$) é $\beta$-sombreada por um ponto $x \in \Omega$.*

*Demonstração.* Seja $\varepsilon > 0$ um número pequeno a ser determinado mais adiante e escolha $\delta \in (0, \varepsilon)$ como no Teorema 5.32, isto é, $W^s_\varepsilon(x) \cap W^u_\varepsilon(y) \cap \Omega \neq \emptyset$ sempre que $x, y \in \Omega$ com $d(x, y) \leq \delta$. Tome $L \in \mathbb{N}$ suficientemente grande de modo a ser $\lambda^L \varepsilon < \dfrac{\delta}{2}$, onde $\lambda$ é dado pelo Teorema da Variedade estável. Além disso, tome $\alpha > 0$ com a propriedade que se $\{y_i\}_{i=0}^L$ é uma $\alpha$-pseudo-órbita em $\Omega$, então

$$d(f^j(y_0), y_j) < \frac{\delta}{2} \quad \text{para todo} \quad j = 0, \ldots, L.$$

Considere primeiro a $\alpha$-pseudo-órbita $\{x_i\}_{i=0}^{rL}$ com $r > 0$. Defina $x'_{kL}$ recursivamente para $k \in \{0, \ldots, r\}$ por $x'_0 = x_0$ e

$$x'_{(k+1)L} = W^u_\varepsilon(f^L(x'_{kL})) \cap W^s_\varepsilon(x_{(k+1)L}) \in \Omega. \tag{5.36}$$

Repare que isto faz sentido, porque supondo que $x'_{kL} \in W^s_\varepsilon(x_{kL})$, $d(f^L(x'_{kL}), f^L(x_{kL})) \leq \lambda^L\varepsilon < \dfrac{\delta}{2}$ (usamos aqui o item (b) do Teorema da Variedade Estável) e pela escolha de $\alpha$, $d(f^L(x_{kL}), x_{(k+1)L}) = d(f^L(x_{kL}), x_{kL+L}) < \dfrac{\delta}{2}$. Assim, $d(f^L(x'_{kL}), x_{(k+1)L}) < \dfrac{\delta}{2} + \dfrac{\delta}{2} = \delta$ e $x'_{(k+1)L}$ está bem definido. Agora ponha $x = f^{-rL}(x'_{rL})$. Para $i \in \{0, \dots, rL\}$, escolha $s$ tal que $i \in \{sL, \dots, (s+1)L-1\}$. Então

$$\begin{aligned} d(f^i(x), f^{i-sL}(x'_{sL})) &= d(f^{i-rL}(x'_{rL}), f^{i-sL}(x'_{sL})) \\ &\leq d(f^{i-rL}(x'_{rL}), f^{i-(r-1)L}(x'_{(r-1)L})) + \\ &\qquad \cdots + d(f^{i-(s+1)L}(x'_{(s+1)L}), f^{i-sL}(x'_{sL})) \\ &= \sum_{t=s+1}^{r} d(f^{i-tL}(x'_{tL}), f^{i-tL+L}(x'_{(t-1)L})) \\ &= \sum_{t=s+1}^{r} d(f^{i-tL}(x'_{tL}), f^{i-tL}(f^L(x'_{(t-1)L}))) \\ &\leq \sum_{t=s+1}^{r} \varepsilon\lambda^{tL-i} \leq \frac{\varepsilon\,\lambda}{1-\lambda}. \end{aligned} \tag{5.37}$$

Na penúltima passagem, usamos que $x'_{tL} \in W^u_\varepsilon(f^L(x'_{(t-1)L}))$ e usamos o item (b) do Teorema da Variedade estável. Por outro lado, como $x'_{sL} \in W^s_\varepsilon(x_{sL})$ (olhe para (5.36)),

$$d(f^{i-sL}(x'_{sL}), f^{i-sL}(x_{sL})) \leq \varepsilon. \tag{5.38}$$

Mas pela escolha de $\alpha$ temos que

$$d(f^{i-sL}(x_{sL}), x_i) \leq \frac{\delta}{2}. \tag{5.39}$$

Pela desigualdade triangular e pelas equações (5.37), (5.38) e (5.39)

$$\begin{aligned} d(f^i(x), x_i) &\leq d(f^i(x), f^{i-sL}(x'_{sL})) + d(f^{i-sL}(x'_{sL}), f^{i-sL}(x_{sL})) + d(f^{i-sL}(x_{sL}), x_i) \\ &\leq \frac{\varepsilon\,\lambda}{1-\lambda} + \varepsilon + \frac{\delta}{2}. \end{aligned}$$

Agora escolha $\varepsilon < \beta$ pequeno de modo a satisfazer $\dfrac{\varepsilon\,\lambda}{1-\lambda} + \varepsilon + \dfrac{\delta}{2} \leq \beta$. Lembre que o $\beta$ é dado. Isto mostra que a $\alpha$-pseudo-órbita $\{x_i\}_{i=0}^{rL}$ é $\beta$-sombreada pelo ponto $x = f^{-rL}(x'_{rL}) \in \Omega$.

Agora qualquer $\alpha$-pseudo-órbita $\{x_i\}_{i=0}^{n}$ em $\Omega$ se estende para $\{x_i\}_{i=0}^{rL}$ quando $rL > n$ fazendo o seguinte: ponha $x_i = f^{i-n}(x_n)$ para $i \in \{n, \ldots, rL\}$ e observe que essa nova sequência também é uma $\alpha$-pseudo-órbita, pois

$$d(f(x_i), x_{i+1}) = d(f(f^{i-n}(x_n)), f^{(i+1)-n}(x_n)) = d(f^{(i+1)-n}(x_n), f^{(i+1)-n}(x_n)) = 0 < \alpha.$$

Claramente, $x \in \Omega$ que sombreia essa pseudo-órbita estendida irá sombrear a original. Em adição, se $\{x_i\}_{i=a}^{b}$ é uma $\alpha$-pseudo-órbita finita, então o mesmo para $\{x_{j+a}\}_{j=0}^{b-a}$, e se $x$ sombreia esta última, $f^{-a}(x)$ sombreia a $\alpha$-pseudo-órbita original. Logo o Lema vale para $\alpha$-pseudo-órbitas finitas.

Finalmente, se $\{x_i\}_{i=-\infty}^{+\infty}$ é uma $\alpha$-pseudo-órbita infinita em $\Omega$, então encontre uma sequência $x^{(m)} \in \Omega$ que $\beta$-sombreia $\{x_i\}_{i=-m}^{m}$ e seja $x = \lim_m x^{(m)}$. Claro que $x \in \Omega$, pois $\Omega$ é fechado, e $x$ $\beta$-sombreia $\{x_i\}_{i=-\infty}^{+\infty}$. Isto conclui a prova. $\square$

## 5.4 Partições de Markov

Para estudar a dinâmica de $f$ num conjunto básico $\Omega_s$, vamos considerar subconjuntos que se comportam bem com respeito à dinâmica. Isso permitirá verificarmos que a dinâmica de um difeomorfismo $f$ que satisfaz o Axioma A em $\Omega_s$ é essencialmente a mesma de um deslocamento de tipo finito.

Um subconjunto $R \subset \Omega_s$ é chamado de **retângulo** se este tem diâmetro pequeno e

$$[x, y] \in R \text{ sempre que } x,\, y \in R.$$

Um retângulo $R$ é chamado de **próprio** se $R$ é fechado e $R = \overline{\operatorname{int}(R)}$ ($\operatorname{int}(R)$ é o interior de $R$ como subconjunto de $\Omega_s$). Para $x \in R$, seja

$$W^s(x, R) = W^s_\varepsilon(x) \cap R \;\text{ e }\; W^u(x, R) = W^u_\varepsilon(x) \cap R$$

onde $\varepsilon$ é como no Teorema 5.32 e o diâmetro de $R$ é pequeno quando comparado a $\varepsilon$.

**Lema 5.38.** *Seja $R$ um retângulo fechado. Como um subconjunto de $\Omega_s$, $R$ tem fronteira*

$$\partial R = \partial^s R \cup \partial^u R$$

*onde*

$$\partial^s R = \{x \in R : x \notin \mathrm{int}(W^u(x,R))\} \quad e \quad \partial^u R = \{x \in R : x \notin \mathrm{int}(W^s(x,R))\}$$

*e os interiores de* $W^u(x,R)$, $W^s(x,R)$ *são subconjuntos de* $W^u_\varepsilon(x) \cap \Omega$, $W^s_\varepsilon(x) \cap \Omega$.

*Demonstração.* Veja o Lema 3.11 da referência [9]. □

**Definição 5.39** (**Partições de Markov**). Uma **partição de Markov** de $\Omega_s$ é uma cobertura finita $\mathscr{R} = \{R_1, \ldots, R_m\}$ de $\Omega_s$ por retângulos próprios com

(a) $\mathrm{int}(R_i) \cap \mathrm{int}(R_j) = \emptyset$ para $i \neq j$;

(b) $f(W^u(x,R_i)) \supset W^u(f(x),R_j)$ e $f(W^s(x,R_i)) \subset W^s(f(x),R_j)$ quando $x \in \mathrm{int}(R_i)$ e $f(x) \in \mathrm{int}(R_j)$.

Note que, uma partição de Markov não é exatamente uma partição de $\Omega_s$ porque os $R_i$'s não são necessariamente disjuntos.

**Exemplo 5.40.** Os conjuntos $V_1 \cap \Lambda$, $V_2 \cap \Lambda$, onde $V_1$ e $V_2$ são os da na figura 5 e $\Lambda$ é a ferradura de Smale do exemplo 5.35 formam uma partição de Markov de $\Lambda$.

**Exemplo 5.41** (**Construção de uma partição de Markov**). Tome a matriz

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

Seja $X$ o espaço de fase sendo dado pelo toro $\mathbb{T}^2$ e $\phi$ dada por $A$: isto é,

$$\phi(x,y) = (\{x+y\}, x).$$

A matriz $A$ tem dois autovalores: $\lambda = (1+\sqrt{5})/2$ e $\mu = (1-\sqrt{5})/2$. Observe que $\lambda > 1$ e $-1 < \mu < 0$. Associados a esses autovalores estão os autovetores $v_\lambda$ apontando para o primeiro quadrante e $v_\mu$ para o segundo. Na figura 7, temos o esboço das duas linhas a partir da origem nas direções dos autovetores $v_\lambda$ e $v_\mu$. A ação de $A$ em um vetor é contrair sua $v_\mu$-componente por $|\mu|$ e expandir sua $v_\lambda$-componente por $\lambda$. Note que $\mu$ é negativo, o que causa uma inversão de direção além de uma contração na $v_\mu$-componente.

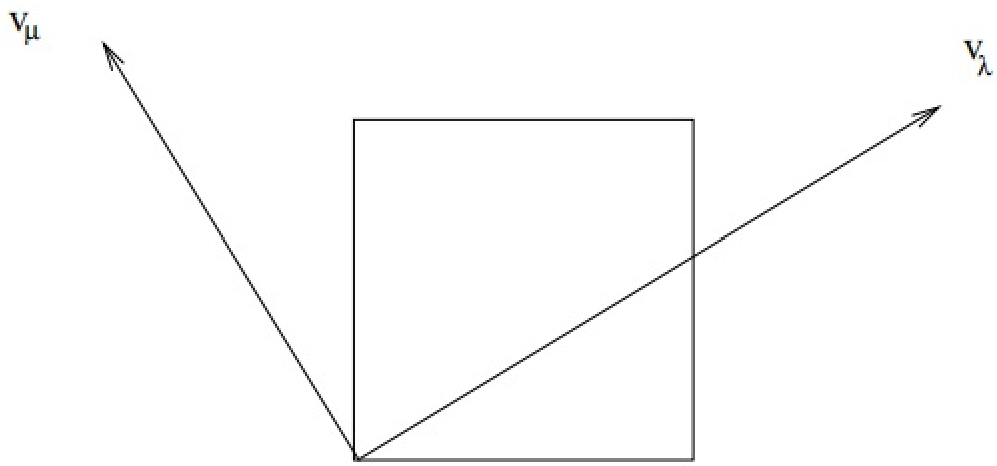


Figura 7: O toro $\mathbb{T}^2$ e as linhas com as direções dos autovetores $v_\lambda$ e $v_\mu$

Nós nos referimos à direção de $v_\lambda$ como a direção expansora e o de $v_\mu$ como a direção contratora. Na figura 8, desenhamos outra região com lados paralelos às direções expansora e contratora. A região da figura 8 é chamada de região fundamental, nela desenhamos a coleção de retângulos abertos $\mathscr{R} = \{R_i : i = 1, 2, 3\}$ como mostrado na figura 9. Esta família de retângulos é um exemplo de partição de Markov.

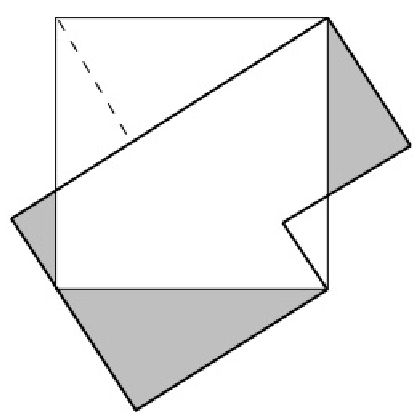

Figura 8: Região fundamental

**Teorema 5.42.** *Seja $\Omega_s$ um conjunto básico para o difeomorfismo $f$ que satisfaz o Axioma A. Então $\Omega_s$ possui partições de Markov $\mathscr{R}$ com diâmetros arbitrariamente pequenos.*

*Demonstração.* Seja $\beta > 0$ muito pequeno e escolha $\alpha > 0$ pequeno como no Lema do

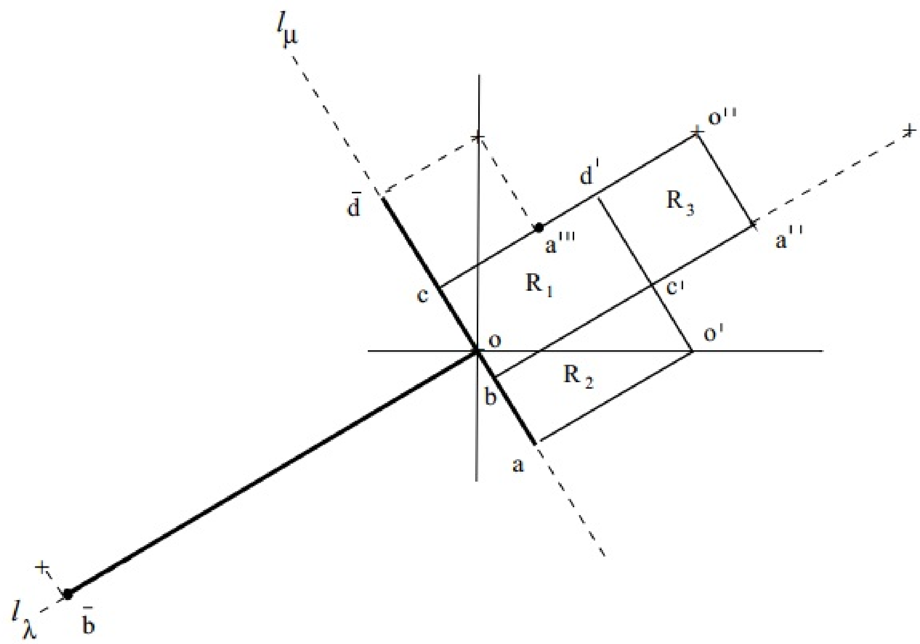


Figura 9: Partição de Markov

Sombreamento (Lema 5.37), ou seja, cada $\alpha$-pseudo-órbita em $\Omega_s$ é $\beta$-sombreada em $\Omega_s$. Escolha $\gamma < \dfrac{\alpha}{2}$ de modo que

$$d(f(x), f(y)) < \frac{\alpha}{2} \text{ quando } d(x, y) < \gamma. \tag{5.40}$$

Seja $P = \{p_1, \ldots, p_r\}$ um subconjunto $\gamma$-**denso** de $\Omega_s$, ou seja, $\Omega_s$ é uma $\gamma$-**vizinhança**[1] de $P$ e

$$\Sigma(P) = \left\{ (q_n)_n \in \prod_{-\infty}^{\infty} P \, : \, d(f(q_j), q_{j+1}) < \alpha \text{ para todo } j \right\}.$$

Repare que toda $(q_n)_n \in \Sigma(P)$ é uma $\alpha$-pseudo-órbita. Logo o Lema do sombreamento (Lema 5.37) garante que existe $x := \theta((q_n)_n) \in \Omega_s$ que $\beta$-sombreia $(q_n)_n$. Afirmamos ainda que este $\theta((q_n)_n)$ é único: se $z$ também $\beta$-sombreia $(q_n)_n$, então pela desigualdade triangular

$$d(f^i(z), f^i(\theta((q_n)_n))) \leq d(f^i(z), (q_n)_n) + d(f^i(\theta((q_n)_n)), (q_n)_n) < \beta + \beta = 2\beta$$

para todo $i \in \mathbb{Z}$. Logo $z \in W^u_{2\beta}(\theta((q_n)_n)) \cap W^s_{2\beta}(\theta((q_n)_n))$ e portanto pelo Teorema 5.32,

[1] $\Omega_s$ é uma $\gamma$-vizinhança de $P$ se estiver contido na união de todas as bolas abertas de raio $\gamma$ centradas em pontos de $P$.

concluímos que $z = \theta((q_n)_n)$. Então fica definida uma função

$$\theta : \Sigma(P) \to \Omega_s$$

que a cada sequência $(q_n)_n \in \Sigma(P)$ associa um ponto $\theta((q_n)_n)$ que $\beta$-sombreia $(q_n)_n$.

Esta função é sobrejetiva. De fato, seja $x \in \Omega_s$. Queremos encontrar $(q_n)_n \in \Sigma(P)$ tal que $\theta((q_n)_n) = x$, ou melhor, queremos encontrar uma sequência $(q_n)_n$ que é $\beta$-sombreada pelo ponto $x$. Considere a sequência $(q_n)_n \in \Sigma(P)$ escolhida da seguinte forma: para cada $i \in \mathbb{Z}$ tome como $q_i$ um elemento de $P$ que diste menos que $\gamma$ de $f^i(x)$, ou seja, $d(q_i, f^i(x)) \leq \gamma$ para todo $i \in \mathbb{Z}$. A sequência $(q_n)_n$ construída dessa forma é $\beta$-sombreada em $x$, para ver isto, basta mostrar que ela é um $\alpha$-pseudo-órbita:

$$\begin{aligned}
d(f(q_j), q_{j+1}) &\leq d(f(q_j), f^{j+1}(x)) + d(f^{j+1}(x), q_{j+1}) \\
&= d(f(q_j), f(f^j(x))) + d(f^{j+1}(x), q_{j+1}) \\
&< \frac{\alpha}{2} + \gamma < \alpha.
\end{aligned}$$

A penúltima desigualdade veio do fato de $d(q_j, f^j(x)) \leq \gamma$ para todo $j \in \mathbb{Z}$ e de (5.40). A última desigualdade veio da escolha de $\gamma < \dfrac{\alpha}{2}$. Logo $\theta$ é sobrejetiva.

A demonstração do Teorema 5.42 continua por meio da sequência de lemas a seguir:

**Lema 5.43.** *A função $\theta : \Sigma(P) \to \Omega_s$ é contínua. Além disso, para quaisquer $(q_n)_n$, $(q'_n)_n \in \Sigma(P)$ com $q_0 = q'_0$ vale que $\theta[(q_n)_n, (q'_n)_n] = [\theta((q_n)_n), \theta((q'_n)_n)]$. Com isso, segue que $T_s = \{\theta((q_n)_n) : (q_n)_n \in \Sigma(P),\, q_0 = p_s\}$ é um retângulo fechado para cada $s \in \{1, \ldots, r\}$.*

*Demonstração.* Se $\theta$ não fosse contínua, existiria $\gamma > 0$ de modo que para cada $N$, poderíamos encontrar sequências $(q_N)_N$, $(q'_N)_N \in \Sigma(P)$ com $q_{N_j} = q'_{N_j}$ para todo $j \in [-N, N]$ (isto diz que $(q_N)_N$, $(q'_N)_N$ tem o mesmo limite) sem que $d(\theta(q_N), \theta(q'_N)) < \gamma$. Como $(q_N)_N$, $(q'_N)_N \in \Sigma(P)$, então são $\beta$-sombreadas respectivamente por $\theta((q_N)_N)$ e $\theta((q'_N)_N)$. Portanto

$$\begin{gathered}
d(f^j(\theta((q_N)_N)), \theta((q_{N_j})_j)) \leq \beta \\
\text{e} \\
d(f^j(\theta((q'_N)_N)), \theta((q'_{N_j})_j)) \leq \beta \quad \text{para todo} \quad j \in [-N, N].
\end{gathered} \tag{5.41}$$

Então usando que $\theta(q_{N_j}) = \theta(q'_{N_j})$ para todo $j \in [-N, N]$, (5.41) e a desigualdade triangular, temos

$$d(f^j(\theta((q_N)_N)), f^j(\theta((q'_N)_N))) \leq 2\beta \text{ para todo } j \in [-N, N].$$

Tomando subsequências se necessário, podemos assumir que $\theta((q_N)_N) \to x$ e $\theta((q'_N)_N) \to y$ quando $N \to \infty$. Então

$$d(f^j(x), f^j(y)) \leq 2\beta \text{ para todo } j \in \mathbb{Z} \text{ e } d(x, y) \geq \gamma. \tag{5.42}$$

Como $\Omega_s$ é um conjunto hiperbólico, o Corolário 5.33 diz que $f|_{\Omega_s}$ é expansiva. Em nosso caso (5.42) diz que $x, y \in W^s_{2\beta}(x) \cap W^u_{2\beta}(y)$ e que $x \neq y$. Então o correto conforme o Corolário 5.33 seria que $d(f^j(x), f^j(y)) > 2\beta$. Temos então uma contradição sobre a expansividade de $f|_{\Omega_s}$. Portanto $\theta$ deve ser contínua.

Para provar a segunda parte tome quaisquer $(q_n)_n$, $(q'_n)_n \in \Sigma(P)$ com $q_0 = q'_0$, defina $(q^*_n)_n = [(q_n)_n, (q'_n)_n] \in \Sigma(P)$ por

$$q^*_j = \begin{cases} q_j & \text{para } j \geq 0 \\ q'_j & \text{para } j \leq 0. \end{cases}$$

Então veja que

$$d(f^j(\theta((q^*_n)_n)), \theta((q_n)_n)) \leq 2\beta \text{ para } j \geq 0 \text{ e } d(f^j(\theta((q^*_n)_n)), \theta((q'_n)_n) \leq 2\beta \text{ para } j \leq 0.$$

Assim, $\theta((q^*_n)_n) \in W^s_{2\beta}(\theta((q_n)_n) \cap W^u_{2\beta}(\theta((q'_n)_n)$, isto é,

$$\theta[(q_n)_n, (q'_n)_n] = \theta((q^*_n)_n) = [\theta((q_n)_n), \theta((q'_n)_n)],$$

como havíamos afirmado.

Agora, podemos mostrar que $T_s = \{\theta((q_n)_n) : (q_n)_n \in \Sigma(P), q_0 = p_s\}$ é um retângulo: para $x, y \in T_s$, podemos escrever $x = \theta((q_n)_n)$ e $y = \theta((q'_n)_n)$ com $q_0 = p_s = q'_0$ pois $\theta$ é sobrejetiva. Então

$$[x, y] = [\theta((q_n)_n), \theta((q'_n)_n)] = \theta[(q_n)_n, (q'_n)_n] \in T_s.$$

Além disso, $T_s$ é fechado para cada $s$: repare que

$$T_s = \{\theta((q_n)_n) : (q_n)_n \in \Sigma(P), q_0 = p_s\} = \theta(\{(q_n)_n : q_0 = p_s\}) = \theta([0; p_s]).$$

onde $\theta([0; p_s])$ é a imagem do cilindro $[0; p_s]$ que é formado por todas as sequências que começam em $p_s$. Mas $[0; p_s]$ é um compacto, portanto $T_s$ é a imagem de um compacto por uma função contínua $\theta$, logo é fechado. □

Note que os retângulos $T_s$ têm diâmetros arbitrariamente pequenos (tomando $\alpha$, $\beta$ e $\gamma$ arbitrariamente pequenos).

**Lema 5.44.** *$f(W^s(x, T_s)) \subset W^s(f(x), T_t)$ e $f(W^u(x, T_s)) \supset W^u(f(x), T_t)$.*

*Demonstração.* Escreva $x = \theta((q_n)_n)$ com $q_0 = p_s$ e $q_1 = p_t$. Considere $y \in W^s(x, T_s) = W^s_\varepsilon(x) \cap T_s$, $y = \theta((q'_n)_n)$ e $q'_0 = p_s$. Então, notando que $[x, y] \in W^u_\varepsilon(y) \cap W^s_\varepsilon(x) \cap T_s$ usamos o Teorema 5.32 para obtermos

$$y = [x, y] = [\theta((q_n)_n), \theta((q'_n)_n)] = \theta[(q_n)_n, (q'_n)_n] = \theta((q^*_n)_n).$$

Mas se $y$ $\beta$-sombreia $(q^*_n)_n$, então $d(f^j(y), q^*_j) \leq \beta$ para todo $j \in \mathbb{Z}$, de outra forma, $d(f^j(f(y)), q^*_{j+1}) = d(f^{j+1}(y), q^*_{j+1}) \leq \beta$ para todo $j \in \mathbb{Z}$, ou seja, $f(y)$ $\beta$-sombreia $(q^*_{n+1})_n$, ou seja $f(y)$ $\beta$-sombreia $\sigma((q^*_n)_n)$ onde $\sigma$ é o deslocamento. Logo

$$f(y) = f(\theta((q^*_n)_n)) = \theta(\sigma((q^*_n)_n)) = \theta(\sigma[(q_n)_n, (q'_n)_n]) \in T_t$$

porque a sequência $\sigma((q^*_n)_n)$ tem $p_t$ como elemento em sua posição zero. Uma vez que $f(y) \in W^s_\varepsilon(f(x))$ (isto é imediato porque $y \in W^s_\varepsilon(x)$), temos que $f(y) \in W^s_\varepsilon(f(x)) \cap T_t = W^s(f(x), T_t)$.

Similarmente, pode-se demonstrar que $f^{-1}(W^u(f(x), T_t)) \subset W^u(x, T_s)$ e isto dá que $f(W^u(x, T_s)) \supset W^u(f(x), T_t)$. □

Como $\theta$ é sobrejetiva, então $\tau = \{T_1, \cdots, T_r\}$ é uma cobertura de $\Omega_s$ por retângulos e o Lema anterior dá exatamente a condição de Markov do item (b) da definição 5.39 para estes retângulos. Contudo nada assegura que os $T_j$ são disjuntos e próprios. Por isso $\tau$ ainda não é a partição de Markov procurada. Para encontrá-la faremos mais argumentos para decompor os retângulos $T_j$ em retângulos com os interiores disjuntos que ainda satisfazem a condição de Markov. Para cada $x \in \Omega_s$ sejam

$$\tau(x) = \{T_j \in \tau : x \in T_j\}$$

e

$$\tau^*(x) = \{T_k \in \tau : T_k \cap T_j \neq \emptyset \text{ para algum } T_j \in \tau(x)\}.$$

Como $\tau$ é uma cobertura fechada de $\Omega_s$, $Z = \Omega_s \setminus \bigcup_j \partial T_j$ é um subconjunto aberto e denso de $\Omega_s$. Também

$$Z^* = \{x \in \Omega_s : W^s_\varepsilon(x) \cap \partial^s T_k = \emptyset \text{ e } W^u_\varepsilon(x) \cap \partial^u T_k = \emptyset \text{ para todo } T_k \in \tau^*(x)\}$$

é aberto e denso em $\Omega_s$: de fato, podemos escrever $Z^* = D \cap E$, onde

$$D = \{x \in \Omega_s : W^s_\varepsilon(x) \cap \partial^s T_k = \emptyset \text{ para todo } T_k \in \tau^*(x)\}$$
$$E = \{x \in \Omega_s : W^u_\varepsilon(x) \cap \partial^u T_k = \emptyset \text{ para todo } T_k \in \tau^*(x)\}$$

e é suficiente provarmos que esses dois últimos são abertos e densos (a interseção de dois abertos denso é denso). Como as demonstrações são semelhantes faremos isto apenas para $D$. Que é $D$ é aberto segue da continuidade da função distância $\varpi(x) = d(W^s_\varepsilon(x), \partial^s T_k)$, e notando que $D = \varpi^{-1}((0,\infty))$ (pré-imagem de um aberto por uma função contínua). Quanto à densidade, dado qualquer $x \in \Omega_s$, queremos mostrar que existe um ponto $a$ arbitrariamente próximo de $x$ com

$$W^s_\varepsilon(a) \cap \partial^s T_k = \emptyset. \tag{5.43}$$

Seja $\delta \leq \dfrac{\varepsilon}{2}$ e $y \in W^s_\delta(x) \cap \partial^s T_k$. É consequência do Lema 5.38 que $y$ pertence a fronteira de $W^u(y, T_k)$ em $W^u_\varepsilon(y) \cap \Omega_s$. Daí $x = [y, x]$ e $d(x, y) < \delta$. Assim, pela continuidade de $[\cdot,\cdot]$ podemos encontrar $z \in W^u_\varepsilon(y) \cap (\Omega_s \setminus T_k)$ tão próximo de $y$ de modo que

$$d(x, z) < \delta \text{ e } d([z, x], y) < \delta. \tag{5.44}$$

Afirmamos que $[z, x]$ é um dos pontos $a$ que satisfaz (5.43). Suponha por contradição que $W^s_\varepsilon([z, x]) \cap \partial^s T_k \neq \emptyset$. Se $t \in W^s_\varepsilon([z, x]) \cap \partial^s T_k$, então $z = [[z, x], y]$ devido a (5.44) e $z = [[z, x], y] = [t, y] \in T_k$ pela escolha de $t \in W^s_\varepsilon([z, x]) \cap \partial^s T_k$. Mais isso é uma contradição porque $z$ foi escolhido fora de $T_k$. Logo $W^s_\varepsilon([z, x]) \cap \partial^s T_k = \emptyset$ e $d([z, x], x) \leq d([z, x], y) + d(x, y) < 2\delta < \varepsilon$. Logo existem pontos arbitrariamente próximos de $x$ satisfazendo (5.43). Portanto $D$ é denso em $\Omega_s$.

Para $T_j \cap T_k \neq \emptyset$, defina (veja a figura 10)

$$T^1_{j,k} = \{x \in T_j : W^u(x, T_j) \cap T_k \neq \emptyset,\ W^s(x, T_j) \cap T_k \neq \emptyset\} = T_j \cap T_k$$
$$T^2_{j,k} = \{x \in T_j : W^u(x, T_j) \cap T_k \neq \emptyset,\ W^s(x, T_j) \cap T_k = \emptyset\}$$
$$T^3_{j,k} = \{x \in T_j : W^u(x, T_j) \cap T_k = \emptyset,\ W^s(x, T_j) \cap T_k \neq \emptyset\}$$
$$T^4_{j,k} = \{x \in T_j : W^u(x, T_j) \cap T_k = \emptyset,\ W^s(x, T_j) \cap T_k = \emptyset\}.$$

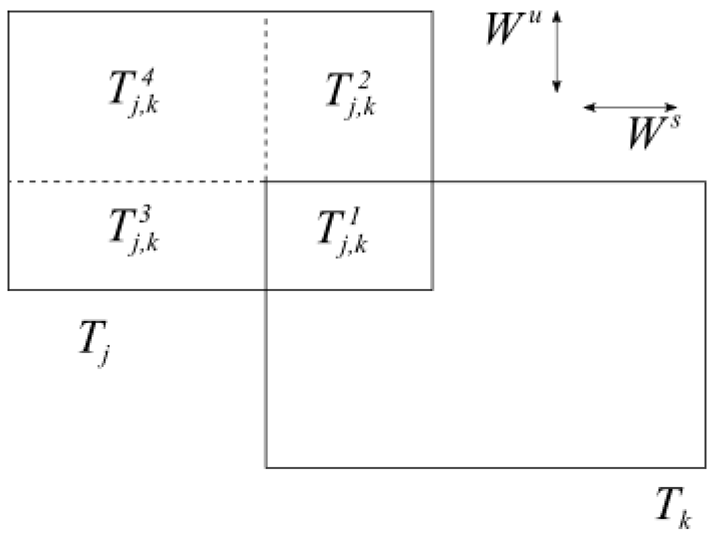


Figura 10: Representação dos conjuntos $T^n_{j,k}$

Para $x \in Z^*$ defina

$$R(x) = \bigcap \{\text{int}(T^n_{j,k}) : x \in T_j,\ T_k \cap T_j \neq \emptyset \text{ e } x \in T^n_{j,k}\}.$$

**Lema 5.45.** *Para $x \in Z^*$, $R(x)$ é um retângulo aberto. Além disso, se $y \in Z^*$, ou $R(x) = R(y)$, ou $R(x) \cap R(y) = \emptyset$.*

*Demonstração.* De fato é retângulo: é suficiente mostrar que cada $\text{int}(T^n_{j,k})$ é um retângulo aberto. Se $a, b \in T^n_{j,k}$ então $W^s([a,b], T_j) = W^s(a, T_j)$ e $W^u([a,b], T_j) = W^u(b, T_j)$ porque $\text{diam}(T^n_{j,k}) \leq \text{diam}(T_j) \leq 2\beta$ é pequeno quando comparado com $\varepsilon$. Portanto quando $a, b \in T^n_{j,k}$, $[a,b] \in T^n_{j,k}$. Como $\text{int}(T^n_{j,k}) \subset T^n_{j,k}$, para todo $a, b \in \text{int}(T^n_{j,k})$, temos que $[a,b] \in \text{int}(T^n_{j,k})$. Segue que $R(x)$ é um retângulo aberto.

Agora, suponha que $y \in R(x) \cap Z^*$. Vamos mostrar que sempre que isso acontece $R(x) = R(y)$, para isto, começamos notando que neste caso $\tau(x) = \tau(y)$: como $R(x) \subset \tau(x)$ e $R(x) \cap T_j = \emptyset$ para $T_j \notin \tau(x)$, segue que os $T_j$ que contém $x$, necessariamente contém $y$. Logo $\tau(x) \subset \tau(y)$. Por outro lado não pode haver $T_j$ que contenha $y$ e não contenha $x$, se isso acontecesse $x \in T^n_{j,k}$ para algum $n \in 2, 3, 4$ e nesse caso $y \notin R(x)$. Logo $\tau(x) \supset \tau(y)$. Para $T_j \in \tau(x) = \tau(y)$ e $T_k \cap T_j \neq \emptyset$, $y$ pertence ao mesmo $T^n_{j,k}$ que $x$, porque $R(x) \subset T^n_{j,k}$, consequentemente $R(y) = \bigcap\{\text{int}(T^n_{j,k}) : y \in T_j,\, T_k \cap T_j \neq \emptyset \text{ e } y \in T^n_{j,k}\} = \bigcap\{\text{int}(T^n_{j,k}) : x \in T_j,\, T_k \cap T_j \neq \emptyset \text{ e } x \in T^n_{j,k}\} = R(x)$.

Finalmente ou $R(x) \cap R(y) = \emptyset$, ou se $R(x) \cap R(y) \neq \emptyset$, existe um $x' \in R(x) \cap R(y) \cap Z^*$, porque esta interseção é aberta e $Z^*$ é denso. Pelo que mostramos acima $R(x) = R(y) = R(x')$. $\square$

Como a quantidade dos $T^n_{j,k}$ é finita, a quantidade dos $R(x)$ também é finita. Então definimos

$$\mathscr{R} = \{\overline{R(x)} : x \in Z^*\} = \{R_1, \ldots, R_m\}.$$

Afirmamos que $\mathscr{R}$ é a partição de Markov que buscamos.

Primeiro $\mathscr{R} = \{R_1, \ldots, R_m\}$ é uma cobertura finita de $\Omega_s$ e cada $R_i$ é um retângulo fechado. Mais ainda, esses retângulos são próprios: para $x' \in Z^*$, $R(x') = R(x)$ ou $R(x') \cap R(x) = \emptyset$ como mostramos no último Lema. Consequentemente

$$(\overline{R(x)} \setminus R(x)) \cap Z^* = \emptyset. \tag{5.45}$$

Mas sendo $Z^*$ denso em $\Omega_s$, (5.45) implica que $\overline{R(x)} \setminus R(x)$ tem interior vazio em $\Omega_s$. Portanto $R(x) = \text{int}(R(x)) = \text{int}(\overline{R(x)})$. Agora os elementos de $\mathscr{R}$ são os fechos dos $R(x)$, consequentemente para cada $i \in \{1, \ldots, m\}$,

$$R_i = \overline{R(x)} = \overline{\text{int}(\overline{R(x)})} = \overline{\text{int}(R_i)}$$

para $x \in Z^*$. Isto mostra que os retângulos $R_i$ são próprios.

Vamos verificar o item (a) da definição de partição de Markov: para $R(x) \neq R(x')$ correspondem conjuntos $R_i$, $R_j$ com $i \neq j$ tais que $R_i = \overline{R(x)}$ e $R_j = \overline{R(x')}$. Então

$$\text{int}(R_i) \cap \text{int}(R_j) = \text{int}(\overline{R(x)}) \cap \text{int}(\overline{R(x')}) = R(x) \cap R(x') = \emptyset.$$

Vamos verificar o item (b) da definição de partição de Markov.

**Lema 5.46.** *Suponha que* $x, y \in Z^* \cap f^{-1}(Z^*)$, $R(x) = R(y)$ *e* $y \in W^s_\varepsilon(x)$. *Então* $R(f(x)) = R(f(y))$.

*Demonstração.* Mostraremos primeiro que $\tau(f(x)) = \tau(f(y))$. Se isto não fosse verdade, teríamos $f(x) \in T_j$ e $f(y) \notin T_j$. Seja $f(x) = \theta(\sigma((q_n)_n))$ com $q_1 = p_j$ e $q_0 = p_s$. Então $x = \theta((q_n)_n) \in T_s$ e pelo Lema 5.44

$$f(y) \in f(W^s(x, T_s)) \subset W^s(f(x), T_j),$$

contradizendo $f(y) \notin T_j$. Agora seja $f(x)$, $f(y) \in T_j$ e $T_k \cap T_j \neq \emptyset$. Nós queremos mostrar que $f(x)$, $f(y)$ pertencem ao mesmo $T^n_{j,k}$. Como $y \in W^s_\varepsilon(x)$, então pelo Lema 5.44, $f(y) \in W^s_\varepsilon(f(x))$, temos portanto que $W^s(f(y), T_j) = W^s(f(x), T_j)$. Se supusermos que $f(x)$, $f(y)$ não pertencem ao mesmo $T^n_{j,k}$, iremos obter uma contradição de

$$W^u(f(y), T_j) \cap T_k = \emptyset, \;\; f(z) \in W^u(f(x), T_j) \cap T_k. \tag{5.46}$$

Recorde que $f(x) = \theta(\sigma((q_n)_n))$, $q_1 = p_j$, $q_0 = p_s$. Então pelo Lema 5.44

$$f(z) \in W^u(f(x), T_j) \subset f(W^u(x, T_s)) \;\text{ ou ainda }\; z \in W^u(x, T_s).$$

Seja $f(z) = \theta(\sigma((q'_n)_n))$ $q'_1 = p_k$ e $q'_0 = p_t$. Então $z \in T_t$ e $f(W^s(z, T_t)) \subset W^s(f(z), T_k)$. Agora, $T_s \in \tau(x) = \tau(y)$ e $z \in T_t \cap T_s \neq \emptyset$. Daí $z \in W^u(x, T_s) \cap T_t$ e assim existe algum $z' \in W^u(y, T_s) \cap T_t$ quando $x$, $y$ estão no mesmo $T^n_{s,t}$. Então

$$z'' = [z, y] = [z, z'] \in W^s(z, T_t) \cap W^u(y, T_s),$$

e $f(z'') = [f(z), f(y)] \in W^s(f(z), T_k) \cap W^u(f(y), T_j) = W^s_\varepsilon(f(x)) \cap T_k \cap W^u(f(y), T_j) = \emptyset$ (usando que $f(z)$, $f(y) \in T_j$ e que este é um retângulo), e isto é uma contradição em (5.46) . Logo $R(f(x)) = R(f(y))$. □

Para $\delta > 0$ pequeno, os conjuntos

$$Y_1 = \bigcup \left\{ W^s_\delta(z) : \, z \in \bigcup_j \partial^s T_j \right\} \;\;\text{ e }\;\; Y_2 = \bigcup \left\{ W^s_\delta(z) : \, z \in \bigcup_j \partial^u T_j \right\}$$

são fechados e **não são densos em lugar algum** (isso significa que $\mathrm{int}(\overline{Y_1}) = \emptyset$ e $\mathrm{int}(\overline{Y_2}) = \emptyset$). Agora $Z^* \supset \Omega_s \setminus (Y_1 \cup Y_2)$ é aberto e denso. Além disso, se $x \notin (Y_1 \cup Y_2) \cap f^{-1}(Y_1 \cup Y_2)$ então $x \in Z^* \cap f^{-1}(Z^*)$ e o conjunto dos $y \in W^s(x, R(x))$ com $y \in Z^* \cap f^{-1}(Z^*)$ é aberto e denso em $W^s(x, \overline{R(x)})$ como um subconjunto de $W^s_\varepsilon(x) \cap \Omega_s$. Pelo que provamos no último Lema, $R(f(y)) = R(f(x))$ para tal $y$. Então pela continuidade de $f$

$$f(W^s(x, \overline{R(x)})) \subset W^s_\varepsilon(f(x)) \cap f(\overline{R(x)}) \subset \overline{R(f(x))}.$$

A desigualdade acima também mostra que $f(W^s(x, \overline{R(x)})) \subset W^s_\varepsilon(f(x))$. Daí $f(W^s(x, \overline{R(x)})) \subset W^s(f(x), \overline{R(f(x))})$. Se $\mathrm{int}(R_i) \cap f^{-1}(\mathrm{int}(R_j)) \neq \emptyset$, então este é um subconjunto aberto de $\Omega_s$ e contém algum ponto $x$ satisfazendo as condições acima, com $R_i = \overline{R(x)}$ e $R_j = \overline{R(f(x))}$. Para qualquer $x' \in R_i \cap f^{-1}(R_j)$, temos que $W^s(x', R_i) = \{[x', y] : y \in W^s(x, R_i)\}$ e

$$\begin{aligned} f(W^s(x', R_i)) &= \{[f(x'), f(y)] : y \in W^s(x, R_i)\} \\ &\subset \{[f(x'), z] : z \in W^s(f(x), R_j)\} \\ &\subset f(W^s(f(x'), R_j)). \end{aligned}$$

Isto prova a segunda inclusão do item (b) da definição de partição de Markov. Para provar a primeira, basta repetir o que foi feito para $f^{-1}$, notando que $W^u_f(x) = W^s_{f^{-1}}(x)$.

A prova do Teorema 5.42 está completa. $\square$

### 5.4.1 Uma aplicação das partições de Markov: a semiconjugação entre $\Sigma_A$ e $\Omega_s$

Uma das vantagens da existência de partições de Markov para difeomorfismos Axioma A é que podemos considerar um deslocamento de tipo finito que possui essencialmente a mesma estrutura de órbitas da $f$.

Suponha que $\mathscr{R} = \{R_1, \ldots, R_m\}$ é uma partição de Markov de um conjunto básico $\Omega_s$. Definimos a **matriz de transição** $A = A(\mathscr{R})$ por

$$A_{i,j} = \begin{cases} 1 & \text{se } \mathrm{int}(R_i) \cap f^{-1}(\mathrm{int}(R_j)) \neq \emptyset \\ 0 & \text{caso contrário.} \end{cases}$$

Considere o deslocamento $\sigma_A : \Sigma_A \to \Sigma_A$ associado a essa matriz $A$. Então $\sigma_A$ e $f|_{\Omega_s}$ são semiconjugados, conforme o seguinte teorema.

**Teorema 5.47.** *Para cada $(a_m)_m \in \Sigma_A$ o conjunto $\bigcap_{j\in\mathbb{Z}} f^{-j}R_{a_j}$ consiste de um único ponto denotado por $\pi((a_m)_m)$. A função $\pi : \Sigma_A \to \Omega_s$ é uma sobrejeção contínua, $\pi \circ \sigma_A = f \circ \pi$, e $\pi$ é injetiva no conjunto residual $Y = \Omega_s \setminus \bigcup_{j\in\mathbb{Z}} f^j(\partial^s\mathscr{R} \cup \partial^u\mathscr{R})$.*

**Lema 5.48.** *Suponha que $x \in R_i$, $f(x) \in R_j$, $A_{i,j} = 1$. Então $f(W^s(x, R_i)) \subset W^s(f(x), R_j)$ e $f(W^u(x, R_i)) \supset W^u(f(x), R_j)$.*

*Demonstração.* Consequência do Lema 5.44. $\square$

Definimos os conjuntos $\partial^s\mathscr{R} = \bigcup_j \partial^s R_j$ e $\partial^u\mathscr{R} = \bigcup_j \partial^u R_j$.

**Proposição 5.49.** *$f(\partial^s\mathscr{R}) \subset \partial^s\mathscr{R}$ e $f^{-1}(\partial^u\mathscr{R}) \subset \partial^u\mathscr{R}$.*

*Demonstração.* Tome $x \in \partial^s R_i$. O conjunto $\bigcup_j(\mathrm{int}(R_i) \cap f^{-1}(\mathrm{int}(R_j)))$ é denso em $R_i$. Logo para qualquer $x \in R_i$ podemos encontrar algum $j$ e $(x_n)_n \in \mathrm{int}(R_i) \cap f^{-1}(\mathrm{int}(R_j))$ com $\lim_{n\to\infty} x_n = x$. Então como $\mathrm{int}(R_i) \cap f^{-1}(\mathrm{int}(R_j)) \neq \emptyset$, $A_{i,j} = 1$, $x \in R_i$ e $f(x) \in R_j$, segue do Lema 5.48 que $f(W^u(x, R_i)) \supset W^u(f(x), R_j)$. Se $f(x) \notin \partial^s\mathscr{R}$, então $W^u(f(x), R_j)$ é uma vizinhança de $f(x)$ em $W^u_\varepsilon(f(x)) \cap \Omega$. Mas assim, $x \in f^{-1}(W^u(f(x), R_j)) \subset W^u(x, R_i)$. Logo $W^u(x, R_i)$ é uma vizinhança de $x$ em $W^u_\varepsilon \cap \Omega_s$, que contraria a hipótese de que $x \in \partial^s R_i$. Assim, mostramos que $f(\partial^s\mathscr{R}) \subset \partial^s\mathscr{R}$. O outro resultado é obtido de maneira similar. $\square$

**Lema 5.50.** *Sejam $A \subset W^s_\delta(x) \cap \Omega$ e $B \subset W^u_\delta(x) \cap \Omega$. Então o retângulo $[B, A]$ é próprio se, e somente se $A = \overline{\mathrm{int}(A)}$ e $B = \overline{\mathrm{int}(B)}$ são subconjuntos de $W^s_\delta(x) \cap \Omega$ e $W^u_\delta(x) \cap \Omega$ respectivamente.*

*Demonstração.* Se $[B, A]$ é próprio, então $[B, A] = \overline{\mathrm{int}[B, A]} = \overline{[\mathrm{int}(B), \mathrm{int}(A)]} = [\overline{\mathrm{int}(B)}, \overline{\mathrm{int}(A)}]$, logo $A = \overline{\mathrm{int}(A)}$ e $B = \overline{\mathrm{int}(B)}$. É imediato da hipótese do enunciado que estes são subconjuntos de $W^s_\delta(x) \cap \Omega$ e $W^u_\delta(x) \cap \Omega$. Reciprocamente, $[B, A] = [\overline{\mathrm{int}(B)}, \overline{\mathrm{int}(A)}] = \overline{\mathrm{int}[B, A]}$, portanto é fechado e $\mathrm{int}[B, A]$ é um subconjunto de $\Omega$. $\square$

**Definição 5.51.** Sejam $R$, $S$ dois retângulos. Dizemos $S$ é um $u$-**sub-retângulo** de $R$ se

(a) $S \neq \emptyset$, $S \subset R$, $S$ é próprio, e

(b) $W^u(y,S) = W^u(y,R)$ para $y \in S$.

**Lema 5.52.** *Suponha que $S$ é um $u$-sub-retângulo de $R_i$ e $A_{i,j} = 1$. Então $f(S) \cap R_j$ é um $u$-sub-retângulo de $R_j$.*

*Demonstração.* Escolha $x \in R_i \cap f^{-1}(R_i)$ e $D = W^s(x, R_i) \cap S$. Como $S$ é um $u$-sub-retângulo, o item (b) da definição 5.51 dá que

$$S = \bigcup_{y \in D} W^u(y, R_i) = [W^u(x, R_i), D].$$

Como $S$ é próprio e não vazio, pelo Lema 5.50, $D \neq \emptyset$ e $D = \overline{\text{int}(D)}$. Agora

$$f(S) \cap R_j = \bigcup_{y \in D} (f(W^u(y, R_i)) \cap R_j).$$

Pelo Lema 5.48, $f(y) \in R_j$ e $f(W^u(y, R_i)) \cap R_j = W^u(f(y), R_j)$. Assim,

$$f(S) \cap R_j = \bigcup_{y' \in f(D)} W^u(y', R_j) = [W^u(f(x), R_j), f(D)].$$

Uma vez que $R_j = [W^u(f(x), R_j), W^s(f(x), R_j)]$ é próprio, temos que $W^u(f(x), R_j)$ é próprio pelo Lema 5.50. Como $f$ mapeia $W^s_\varepsilon(x) \cap \Omega$ homeomorficamente em uma vizinhança de $W^s_\varepsilon(f(x)) \cap \Omega$, $f(D) = \overline{\text{int}(f(D))}$ e assim $f(S) \cap R_j$ é próprio pelo Lema 5.50, $f(S) \cap R_j \neq \emptyset$ quando $f(D) \neq \emptyset$; se $y'' \in f(S) \cap R_j$, então $y'' \in W^u(y', R_j)$ para algum $y' \in f(D)$ e $W^u(y'', R_j) = W^u(y', R_j) \subset f(S) \cap R_j$. Assim $f(S) \cap R_j$ é um $u$-sub-retângulo de $R_j$. $\square$

Agora podemos demonstrar o Teorema 5.47.

***Demonstração do Teorema 5.47***. Se $a_1 a_2 \cdots a_n$ é uma sequência admissível, ou seja, $A_{a_j, a_{j+1}} = 1$ para todo $j = 1, \ldots, n-1$, então, podemos usar recursivamente o Lema 5.52 da seguinte forma: $R_{a_1}$ é um $u$-sub-retângulo de $R_{a_1}$, portanto pelo Lema 5.52, $f(R_{a_1}) \cap R_{a_2}$ é um $u$-sub-retângulo de $R_{a_2}$, portanto pelo Lema 5.52, $f(f(R_{a_1}) \cap R_{a_2}) \cap R_{a_3} = f^2(R_{a_1}) \cap f(R_{a_2}) \cap R_{a_3}$ é um $u$-sub-retângulo de $R_{a_3}$ e repetimos este argumento $n$ vezes para concluirmos que

$$f^n(R_{a_1}) \cap f^{n-1}(R_{a_2}) \cap \cdots \cap f(R_{a_{n-1}}) \cap R_{a_n} = \bigcap_{j=0}^{n} f^{n-j}(R_{a_j})$$

é um $u$-sub-retângulo de $R_{a_n}$. Resulta disso que

$$\pi_n((a_m)_m) = \bigcap_{j=-n}^{n} f^{-j}(R_{a_j})$$

é não-vazio e igual ao fecho de seu interior. Como a sequência $\pi_n(((a_m)_m)$ é decrescente ($n < n+1 < n+2 < \cdots$ implica que $\pi_n((a_m)_m) \supset \pi_{n+1}((a_m)_m) \supset \pi_{n+2}((a_m)_m) \supset \cdots$) e é formada por compactos não vazios, temos que

$$\pi((a_m)_m) = \bigcap_{j=-\infty}^{\infty} f^{-j}(R_{a_j}) = \bigcap_{n=1}^{\infty} \bigcap_{j=-n}^{n} f^{-j}(R_{a_j}) \neq \emptyset.$$

Se $x$, $y \in \pi((a_m)_m)$, então $f^j(x)$, $f^j(y) \in R_{a_j}$ e

$$d(f^j(x), f^j(y)) \leq \operatorname{diam}(R_{a_j}) < \varepsilon \quad \text{para todo} \quad j \in \mathbb{Z}$$

onde $\varepsilon$ é a constante de expansividade de $f|_{\Omega_s}$, portanto $x = y$ e $\pi : \Sigma_A \to \Omega_s$ está bem definida. Como

$$\pi(\sigma_A((a_m)_m)) = \bigcap_j f^{-j} R_{a_{j+1}} = f\left(\bigcap_j f^{-j-1} R_{a_{j+1}}\right) = f\left(\bigcap_k f^{-k} R_{a_k}\right) = f(\pi((a_m)_m)),$$

temos que $\pi \circ \sigma_A = f \circ \pi$. Também $\pi$ é contínua: seja $(a_m)_m \in \Sigma_A$. Suponha que $(b_m)_m^{(j)} \in \Sigma_A$ é uma sequência que converge para $(a_m)_m$ quando $j \to \infty$. Queremos mostrar que $\pi((b_m)_m^{(j)})$ converge para $\pi((a_m)_m)$. Entendemos que a distância que definimos em $\Sigma_A$ é a mesma da observação 5.2. Então $((b_m)_m^{(j)})_j$ convergir para $(a_m)_m$ significa que

$$d(((b_m)_m^{(j)})_j, (a_m)_m) = \beta^{N(j)} \to 0, \tag{5.47}$$

onde $\beta \in (0,1)$ e $N(j)$ é o maior inteiro não negativo tal que $b_i^{(j)} = a_i$ para cada $i$ com $|i| < N(j)$. Por outro lado $\pi_n((a_m)_m) = \bigcap_{p=-n}^{n} f^{-p}(R_{a_p})$ é uma sequência encaixada decrescente e que converge para $\pi((a_m)_m)$: isto significa que para $n > n_0$, $\pi_n((a_m)_m)$ está contida dentro de uma pequena vizinhança $U$ de $\pi((a_m)_m)$. Agora tome, $N(j) > n > n_0$, então $\pi_n((a_m)_m) = \bigcap_{p=-n}^{n} f^{-p}(R_{a_p}) \supset \bigcap_{p=-N(j)}^{N(j)} f^{-p}(R_{a_p}) = \bigcap_{p=-N(j)}^{N(j)} f^{-p}(R_{b_p^{(j)}}) = \pi_{N(j)}((b_m)_m^{(j)})_j)$, logo $\pi_{N(j)}((b_m)_m^{(j)})_j)$ também está contido em $U$. Fazendo $n \to \infty$, $N(j) \to \infty$ e observamos que $\pi((b_m)_m^{(j)})_j)$ converge para $\pi((a_m)_m)$.

Como $\partial^s\mathscr{R} \cup \partial^u\mathscr{R}$ é fechado e não é denso em lugar algum, segue do Teorema de Baire que $Y = \Omega_s \setminus \bigcup_{j\in\mathbb{Z}} f^j(\partial^s\mathscr{R} \cup \partial^u\mathscr{R})$ é residual.

Agora a injetividade sobre o conjunto residual: para $x \in Y$ escolha $a_j$ com $f^j(x) \in R_{a_j}$. Como $x \in Y$, $f^j(x) \in \text{int}(R_{a_j})$ e assim $A_{a_j,a_{j+1}} = 1$. Portanto a sequência $(a_j)_j$ pertence a $\Sigma_A$ e $x = \pi((a_j)_j)$. Se $x = \pi((b_m)_m)$, então $f^j(x) \in R_{b_j}$ e $b_j = a_j$ porque $\text{int}(R_{b_j}) \cap R_{a_j} = \emptyset$ se $a_j \neq b_j$. Assim $\pi$ é injetiva sobre $Y$. Como $\pi$ é contínua e $\Sigma_A$ é compacto, segue que $\pi(\Sigma_A)$ é um subconjunto compacto de $\Omega_s$ contendo o conjunto denso $Y$, logo $\pi(\Sigma_A) = \Omega_s$. □

Mais à frente, usaremos o seguinte Lema.

**Proposição 5.53.** *$\sigma_A : \Sigma_A \to \Sigma_A$ é topologicamente transitivo. Se $f|_{\Omega_s}$ é topologicamente misturadora, o mesmo vale para $\sigma_A : \Sigma_A \to \Sigma_A$.*

*Demonstração.* Veja o Lema 3.19 da referência [9]. □

É interessante observar que em alguns casos $\pi$ é uma bijeção, por exemplo quando o conjunto hiperbólico $\Omega_s$ é a ferradura de Smale (que é um conjunto de Cantor). Temos a seguinte proposição sobre isto.

**Proposição 5.54.** *$\pi$ é uma conjugação se $\Omega_s$ é totalmente desconexo.*

*Demonstração.* Veja a proposição 18.7.8 da referência [5]. □

## 5.5 Estados de equilíbrio para Difeomorfismos Axioma A

Recorde que uma função é Hölder contínua se existem constantes $a > 0$ e $\theta > 0$ tais que $|\phi(z) - \phi(w)| \leq a d(z,w)^\theta$. O principal teorema deste capítulo que enunciamos e provamos abaixo é devido a Bowen, e é uma versão do Teorema de Ruelle para difeomorfismos Axioma A.

**Teorema 5.55 (Bowen).** *Seja $\Omega_s$ um conjunto básico para o difeomorfismo Axioma A $f$ e seja $\phi : \Omega_s \to \mathbb{R}$ Hölder contínua. Então $\phi$ possui um único estado de equilíbrio $\mu_\phi$ com respeito a $f|_{\Omega_s}$.*

*Demonstração.* Começaremos enunciando um lema auxiliar.

**Lema 5.56.** *Existem $\varepsilon > 0$ e $\alpha \in (0,1)$ para os quais o seguinte é válido: Se $x \in \Omega_s$, $y \in M$, e $d(f^k(x); f^k(y)) \leq \varepsilon$ para todo $k \in [-N, N]$, então $d(x,y) < \alpha^N$.*

*Demonstração.* Sejam $(x_i, U_i)$ cobertura de $\Omega_s$ por cartas locais e seja $\varepsilon$ o número de Lebesgue dessa cobertura. Se $d(f^n(x), f^n(y)) \leq \varepsilon$, então $f^n(x)$ e $f^n(y)$ pertencem ao aberto domínio de uma mesma carta local $(x_{i_n}, U_{i_n})$. Defina $F_m := x_{i_{m+1}} \circ f \circ x_{i_m}^{-1} : \mathbb{R}^n \to \mathbb{R}^n$. Denote $x^m = x_{i_m}(f^m(x))$ e $y^m = x_{i_m}(f^m(y))$. Pela compacidade de $\Omega_s$, existe $L$ tal que $L^{-1} \leq \dfrac{d(x,y)}{||x_i(x) - x_i(y)||} \leq L$ para todo $x \neq y$ com $x, y \in U_i$.

Sejam $w = (x_1, x_2)$ e $z = (y_1, y_2)$ em $E^s \oplus E^u$. Se $||x_2 - y_2|| \geq ||x_1 - y_1||$, então:

$$||F_m(w)_2 - F_m(z)_2|| \geq (\lambda^{-1} - \zeta)||x_2 - y_2|| \geq (\lambda + \zeta)||x_1 - y_1|| \geq ||F_m(w)_1 - F_m(z)_1||. \tag{5.48}$$

De fato, usemos em $\mathbb{R}^n$ a norma $||x_1 \oplus x_2|| = \max\{||x_1||, ||x_2||\}$. Numa vizinhança de 0,

$$F_m = DF_m(0) + \varphi_m, \text{ com } \varphi_m(0) = 0\,,\ ||D\varphi_m|| \leq \zeta \text{ e } 0 < \lambda + \zeta < 1,\ \lambda^{-1} - \zeta > 1.$$

Então

$$\begin{aligned}
||F_m(w)_2 - F_m(z)_2|| &= ||DF_m(0)_2(x_2 - y_2) + \varphi_m(w)_2 - \varphi_m(z)_2|| \\
&\geq \lambda^{-1}||x_2 - y_2|| - \zeta||w - z|| \\
&= (\lambda^{-1} - \zeta)||x_2 - y_2|| \geq (\lambda + \zeta)||x_2 - y_2||
\end{aligned}$$

e

$$\begin{aligned}
||F_m(w)_1 - F_m(z)_1|| &= ||DF_m(0)_1(x_1 - y_1) + \varphi_m(w)_1 - \varphi_m(z)_1|| \\
&\leq \lambda||x_1 - y_1|| + \zeta||w - z|| \\
&\leq (\lambda + \zeta)||x_2 - y_2||.
\end{aligned}$$

Agora, se $||(x^0)_2 - (y^0)_2|| \geq ||(x^0)_1 - (y^0)_1||$, usando (5.48), temos que $||(x^j)_2 - (y^j)_2|| \geq ||(x^j)_1 - (y^j)_1||$ para todo $j = 0, \cdots, N$ e assim,

$$L\varepsilon \geq ||(x^N)_2 - (y^N)_2|| \geq \cdots \geq (\lambda^{-1} - \zeta)^N||(x^0)_2 - (y^0)_2|| = (\lambda^{-1} - \zeta)^N||(x^0) - (y^0)||.$$

Portanto, $d(x,y) \leq L^2\varepsilon(\lambda^{-1}-\zeta)^{-N} < \alpha^N$ para algum $\alpha \in (0,1)$. Se $||(x^0)_2 - (y^0)_2|| \geq ||(x^0)_1 - (y^0)_1||$, usamos (5.48) para $F_{-1}, F_{-2}, \cdots, F_{-N}$ e procedemos de modo análogo ao que fizemos acima. $\square$

Sejam $\mathscr{R}$ uma partição de Markov de $\Omega_s$ de diâmetro no máximo $\varepsilon$, $A$ a matriz de transição para $\mathscr{R}$ e $\pi : \Sigma_A \to \Omega_s$ como no Teorema 5.47. Seja $\phi^* = \phi \circ \pi$. Se $(x_n)_n$, $(y_n)_n \in \Sigma_A$ com $x_k = y_k$ para $k \in [-N, N]$, então

$$f^k(\pi((x_n)_n)) = \pi(\sigma_A^k((x_n)_n)) = \pi((x_{n+k})_n) \in R_{x_k} \text{ para } k \in [-N, N]$$

e

$$f^k(\pi((y_n)_n)) = \pi(\sigma_A^k((y_n)_n)) = \pi((y_{n+k})_n) \in R_{y_k} \text{ para } k \in [-N, N].$$

Como $R_{x_k} = R_{y_k}$ para todo $k \in [-N, N]$ e o diâmetro dos elementos da partição de Markov são todos menores que $\varepsilon$, segue que $d(f^k(\pi((x_n)_n)), f^k(\pi((y_n)_n))) < \varepsilon$. Pelo Lema 5.56, $d(\pi((x_n)_n), \pi((y_n)_n)) < \alpha^N$. Como $\phi$ Hölder contínua e $\phi^* = \phi \circ \pi$,

$$|\phi^*((x_n)_n) - \phi^*((y_n)_n)| \leq a d(\pi((x_n)_n), \pi((y_n)_n))^\theta < a(\alpha^\theta)^N. \tag{5.49}$$

Isso significa que $\phi^* \in \mathcal{F}_A$, porque (5.49) diz que $\operatorname{var}_N \phi^* \leq a(\alpha^\theta)^N$. Primeiro assumimos que $f|_{\Omega_s}$ é topologicamente misturadora. Pelo Lema 5.53, $\sigma_A : \Sigma_A \to \Sigma_A$ é topologicamente misturadora. Logo temos a medida de Gibbs $\mu_{\phi^*}$ dada pelo Teorema 5.4. Sejam $D_s = \pi^{-1}(\partial^s\mathscr{R})$ e $D_u = \pi^{-1}(\partial^u\mathscr{R})$. Então $D_s$ e $D_u$ são subconjuntos fechados de $\Sigma_A$, cada um deles é menor que $\Sigma_A$, $\sigma_A(D_s) \subset D_s$ e $\sigma_A^{-1}(D_u) \subset D_u$. Como $\mu_{\phi^*}$ é $\sigma_A$-invariante (pelo Teorema 5.4), $\mu_{\phi^*}(\sigma_A^n(D_s)) = \mu_{\phi^*}(D_s)$. Usando que $\sigma_A^{n+1}(D_s) \subset \sigma_A^n(D_s)$, temos que

$$\mu_{\phi^*}\left(\bigcap_{n\geq 0} \sigma_A^n(D_s)\right) = \lim_{n\to\infty} \mu_{\phi^*}(\sigma_A^n(D_s)) = \mu_{\phi^*}(D_s).$$

Agora, como $\mu_{\phi^*}$ é ergódica e $\bigcap_{n\geq 0} \sigma_A^n(D_s)$ é $\sigma_A$ invariante, sua medida é ou 0 ou 1. Uma vez que seu complementar é um conjunto aberto não vazio e tem medida positiva, temos que $\mu_{\phi^*}(D_s) = \mu_{\phi^*}\left(\bigcap_{n\geq 0} \sigma_A^n(D_s)\right) = 0$. De modo análogo, obtemos que $\mu_{\phi^*}(D_u) = 0$.

Agora seja $\mu_\phi = \pi_*\mu_{\phi^*}$, isto é, $\mu_\phi(E) = \mu_{\phi^*}(\pi^{-1}(E))$. Então $\mu_\phi$ é $f$-invariante, porque $\mu_\phi(f^{-1}(E)) = \mu_{\phi^*}(\pi^{-1}(f^{-1}(E))) = \mu_{\phi^*}(\sigma_A^{-1}(\pi^{-1}(E))) = \mu_{\phi^*}(\pi^{-1}(E)) = \mu_\phi(E)$

uma vez que $\pi \circ \sigma_A = f \circ \pi$ como vimos no Teorema 5.47 e $\pi : \Sigma_A \to \Omega_s$ é uma bijeção exceto no conjunto de medida nula $\bigcup_{n\in\mathbb{Z}} \sigma_A^n(D_s \cup D_u)$. Mas isso significa que os sistemas $(\sigma_A, \mu_{\phi^*})$ e $(f, \mu_\phi)$ são ergodicamente equivalentes (veja a definição 2.18). Em particular, $\mu_\phi$ é ergódica (Veja o Teorema 2.20) e o Teorema 3.24 diz que

$$h_{\mu_\phi}(f) = h_{\mu_{\phi^*}}(\sigma_A).$$

E assim, devido ao estado de Gibbs $\mu_{\phi^*}$ ser também estado de equilíbrio (veja a construção feita na seção 4.2)

$$\begin{aligned} h_{\mu_\phi}(f) + \int \phi\, d\mu_\phi &= h_{\mu_{\phi^*}}(\sigma_A) + \int \phi\, d\mu_\phi \\ &= h_{\mu_{\phi^*}}(\sigma_A) + \int \phi\, d\pi_*\mu_{\phi^*} \\ &= h_{\mu_{\phi^*}}(\sigma_A) + \int \phi \circ \pi\, d\mu_{\phi^*} \\ &= h_{\mu_{\phi^*}}(\sigma_A) + \int \phi^*\, d\mu_{\phi^*} \\ &= P(\sigma_A, \phi^*) \end{aligned} \tag{5.50}$$

Usando o item (3) da observação 3.53 (porque $\pi$ é sobrejetiva), temos que

$$P(\sigma_A, \phi^*) = P(\sigma_A, \phi \circ \pi) \geq P(f, \phi). \tag{5.51}$$

Pelo princípio variacional

$$P(f, \phi) \geq h_{\mu_\phi}(f) + \int \phi\, d\mu_\phi. \tag{5.52}$$

Juntando (5.52), (5.50) e (5.51), concluímos que $P(\sigma_A, \phi^*) = P(f, \phi)$. Consequentemente $\mu_\phi$ é um estado de equilíbrio para $\phi$ pelo que vimos na seção 4.2.

Sabemos que $\mu_{\phi^*}$ é único estado de equilíbrio de $\phi^*$ pela construção que fizemos na seção 4.2, contudo isso não é suficiente para garantir que $\mu_\phi = \pi_*\mu_{\phi^*}$ é único estado de equilíbrio para $\phi$. Para mostrar que $\mu_\phi$ é único, precisamos do seguinte Lema

**Lema 5.57.** *Para qualquer $\mu \in \mathcal{M}_1(f)$ existe $\nu \in \mathcal{M}_1(\sigma_A)$ com $\pi_*\nu = \mu$.*

*Demonstração.* Este fato bem conhecido é provado da seguinte maneira. $F(g\circ\pi) = \int g d\mu$ define um funcional linear positivo em um subespaço de $\mathcal{C}(\Sigma_A)$. Por uma modificação do

Teorema de Hahn-Banach (veja [1]), $F$ se estende ainda positivo a todo $\mathcal{C}(\Sigma_A)$. Como $F(1) = F(1 \circ \pi) = 1$, $F$ é identificado com alguma medida $\beta \in \mathcal{M}(\Sigma_A)$. Pela compacidade, seja $\nu = \lim_{k\to\infty} \dfrac{1}{n_k}(\beta + (\sigma_A)_*\beta + \cdots + (\sigma_A^{n_k-1})_*\beta)$. Então $(\sigma_A)_*\nu = \nu$ e

$$\begin{aligned}
\pi_*\nu &= \lim_{k\to\infty} \frac{1}{n_k}(\pi_*(\beta) + \pi_*((\sigma_A)_*\beta) + \cdots + \pi_*((\sigma_A^{n_k-1})_*\beta)) \\
&= \lim_{k\to\infty} \frac{1}{n_k}(\pi_*(\beta) + f_*(\pi_*\beta) + \cdots + (f^{n_k-1})_*(\pi_*\beta)) \\
&= \lim_{k\to\infty} \frac{1}{n_k}(\mu + f_*(\mu) + \cdots + (f^{n_k-1})_*(\mu) \\
&= \lim_{k\to\infty} \frac{1}{n_k} n_k\, \mu = \mu.
\end{aligned}$$

□

Tendo em vista esse Lema, para qualquer estado de equilíbrio $\mu \in \mathcal{M}_1(f)$ para $\phi$ existe $\nu \in \mathcal{M}_1(\sigma_A)$ com $\pi_*\nu = \mu$. Então imitando a prova do Teorema 3.24, com $\mathscr{P}$ partição de $\Omega_s$, $\mathscr{Q} = \pi^{-1}(\mathscr{P})$ partição de $\Sigma_A$ (por causa da sobrejetividade de $\pi$), conclui-se que $h_\nu(\sigma_A) \geq h_\mu(f)$ e

$$\begin{aligned}
h_\nu(\sigma_A) + \int \phi^* d\nu \geq h_\mu(f) + \int \phi \circ \pi \, d\nu = h_\mu(f) + \int \phi \, d\, \pi_*\nu& \\
= h_\mu(f) + \int \phi d\mu = P(f, \phi) = P(\sigma_A, \phi^*).& \quad (5.53)
\end{aligned}$$

Mas o princípio variacional diz que

$$P(\sigma_A, \phi^*) \geq h_\nu(\sigma_A) + \int \phi^* d\nu. \quad (5.54)$$

Juntando (5.53) e (5.54), resulta que $\nu$ é um estado de equilíbrio para $\phi^*$. Daí $\nu = \mu_{\phi^*}$ já que $\mu_{\phi^*}$ é o único estado de equilíbrio para $\phi^*$ conforme a construção que fizemos na seção 4.2. Então $\mu = \pi_*\nu = \pi_*\mu_{\phi^*} = \mu_\phi$. Como $\mu$ escolhida no início é qualquer, $\mu_\phi$ é o único estado de equilíbrio para $\phi$.

Resta fazer o caso em que $\Omega_s = X_1 \cup \cdots \cup X_m$ com $f(X_k) = X_{K+1}$ e $f^m|_{X_1}$ topologicamente misturadora (Veja o Teorema 5.34 da decomposição Espectral). Para $\mu \in \mathcal{M}_1(f)$, temos que $\mu(X_1) = \dfrac{1}{m}$ e, portanto, $\mu' = m\mu|_{X_1} \in \mathcal{M}_1(f^m|_{X_1})$. Por outro lado, se $\mu' \in \mathcal{M}_1(f^m|_{X_1})$, então $\mu \in \mathcal{M}_1(f)$ onde

$$\mu(E) = \frac{1}{m} \sum_{k=0}^{m-1} \mu'(X_1 \cap f^k(E)).$$

Verifica-se que a aplicação que associa $\mu \in \mathcal{M}_1(f)$ a cada $\mu' \in \mathcal{M}_1(f^m|_{X_1})$ define uma bijeção, $h_{\mu'}(f^m|_{X_1}) = mh_{\mu'}(f|_{X_1}) = mh_\mu(f)$ e $\int \phi_m d\mu' = m \int \phi d\mu$. Encontrar $\mu$ maximizando $h_\mu(f) + \int \phi d\mu$ é equivalente, portanto, a encontrar $\mu'$ maximizando $h_{\mu'}(f^m|_{X_1}) + \int \phi_m d\mu'$. Se $\phi$ é Hölder em $\Omega_s$, então $\phi_m$ será Hölder em $X_1$ e, portanto, como a conclusão do teorema 5.55 é válida para $f^m|_{X_1}$ topologicamente misturadora, o mesmo vale para $f|_{\Omega_s}$. □

Nesse caso também temos uma versão do Teorema de Livsič para $f$ difeomorfismo Axioma A:

**Teorema 5.58.** *Sejam $\phi, \psi : \Omega_s \to \mathbb{R}$ duas funções Hölder contínuas. As seguintes condições são equivalentes:*

*(a)* $\mu_\phi = \mu_\psi$;

*(b) Existem constantes $c, L \in \mathbb{R}$ de modo que $|\phi_m(x) - \psi_m(x) - cm| \leq L$ para todo $x \in \Omega_s$ e $m \geq 0$.*

*(c) Existe $c \in \mathbb{R}$ tal que $\phi_m(x) - \psi_m(x) = \sum_{i=0}^{m-1} \phi(\sigma_A^i((x_n)_n)) - \sum_{i=0}^{m-1} \psi(\sigma_A^i((x_n)_n)) = cm$ para todo $x \in \mathrm{Fix}(f^m)$ ;*

*(d) Existe $c \in \mathbb{R}$ e uma função Hölder $u : \Omega_s \to \mathbb{R}$ tal que*

$$\phi - \psi = c + u \circ f - u;$$

*Se estas condições valem, então $c = P(f, \phi) = P(f, \psi)$.*

*Demonstração.* Veja a proposição 4.5 da referência [9]. □

# 6 ESTADOS DE EQUILÍBRIO PARA HOMEOMORFISMOS COM ESPECIFICAÇÃO

Neste capítulo, vamos estudar uma outra abordagem feita por Bowen [10] para garantir a unicidade de estados de equilíbrio. A propriedade fundamental de tais sistemas é chamada de especificação, a qual definimos logo no início deste capítulo (Definição 6.1).

Bowen [10] demonstrou que todo homeomorfismo expansivo satisfazendo a propriedade de especificação admite um único estado de equilíbrio para potenciais com variação pequena em bolas dinâmicas (ver definição 6.2). Esta abordagem dá outra demonstração para o Teorema 5.55, conforme verificamos na seção 5.2.

## 6.1 Definições e enunciado

**Definição 6.1.** Suponhamos que $f : M \to M$ é um homeomorfismo no espaço métrico compacto $M$. Lembre que $f$ é dita **expansiva** se existe $\varepsilon > 0$ (chamada constante de expansividade) tal que, dados $x, y \in M$ com $x \neq y$ existe $n \in \mathbb{Z}$ tal que $d(f^n(x), f^n(y)) > \varepsilon$. Dizemos que $f$ possui a **propriedade de especificação** se para cada $\delta > 0$ existe um número inteiro $p(\delta)$ para o qual o seguinte é válido: se $I_1, \ldots, I_n$ são intervalos de números inteiros $I_i = [a_i, b_i] \subset \mathbb{Z}$, todos contidos em um intervalo $[a, b] \subset \mathbb{Z}$, com $d(I_i, I_j) \geq p(\delta)$ para $i \neq j$ (ou seja o espaçamento entre cada um dos intervalos $I_i$ é maior ou igual a $p(\delta)$) e $x_1, \ldots, x_n \in M$, então, existe um ponto $x \in M$ com $f^{b-a+p(\delta)}(x) = x$ e $d(f^k(x), f^k(x_i)) < \delta$ para $k \in I_i$, $i = 1, \ldots, n$.

Intuitivamente, $f$ atender a propriedade de especificação significa que existe um ponto $x \in M$ que $\delta$-sombreia parte das órbitas dos $x_i$, ou seja, se pensarmos que os $I_i$ são intervalos de tempo, a órbita de $x$ fica próxima da órbita de cada um dos $x_i$, em determinados intervalos de tempo. Além disso esse ponto $x$ é periódico com período $b - a + p(\delta)$.

Para o potencial $\varphi \in C^0(M)$ e $n \geq 1$, continuaremos a utilizar a notação $\varphi_n(x) = \sum_{i=0}^{n-1} \varphi(f^i(x))$.

**Definição 6.2.** Chamaremos de $V(f)$ o conjunto dos potenciais $\varphi \in C^0(M)$ para os quais existem $\varepsilon > 0$ e $L > 0$ de modo que o seguinte é válido: se $d(f^k(x), f^k(y)) \leq \varepsilon$ para todo $0 \leq k < n$, então $|\varphi_n(x) - \varphi_n(y)| \leq L$.

Seja $\delta > 0$ uma constante de expansividade qualquer para $f$. Se $\varphi \in V(f)$, definimos

$$|\varphi|_f := \sup\{|\varphi_n(x) - \varphi_n(y)| : \; n \geq 1 \;\text{ e }\; d(f^k(x), f^k(y)) \leq \delta \;\text{ para todo }\; 0 \leq k < n\}$$

e

$$|||\varphi||| := |\varphi|_f + ||\varphi||,$$

onde $||\cdot||$ denota a norma uniforme.

**Proposição 6.3.** *$(V(f), |||\cdot|||)$ é um espaço de Banach.*

*Demonstração.* De fato, $V(f)$ é um espaço vetorial, pois $\varphi$, $\psi \in V(f)$ e $d(f^k(x), f^k(y)) \leq \varepsilon$ para todo $0 \leq k < n$, implicam que $|\varphi_n(x) - \varphi_n(y)| \leq L$ e $|\psi_n(x) - \psi_n(y)| \leq L$. Usando a estrutura de espaço vetorial de $C^0(M)$:

$$|(\varphi + \psi)_n(x) - (\varphi + \psi)_n(y)| \leq |\varphi_n(x) - \varphi_n(y)| + |\psi_n(x) - \psi_n(y)| \leq 2L,$$

ou seja, $\varphi + \psi \in V(f)$. De modo semelhante, prova-se que $c\varphi \in V(f)$ para todo $c \in \mathbb{R}$.

Agora note que $|\varphi|_f$ é finito. De fato, sejam $\varepsilon > 0$ e $L > 0$ são como na definição 6.2. Por expansividade, existe $N \in \mathbb{N}$ tal que se $d(f^j(x), f^j(y)) \leq \delta$ para todo $|j| \leq N$, então $d(x, y) \leq \varepsilon$, pois do contrário, para todo $n$ existiriam $x_n$, $y_n \in M$ tais que $d(f^j(x_n), f^j(y_n)) \leq \delta$ para todo $|j| < n$ e $d(x_n, y_n) > \varepsilon$. Por compacidade, existe $(n_j)_j \to \infty$ tal que $(x_{n_j})_j$ e $(y_{n_j})_j$ convergem para pontos $x$ e $y$, respectivamente. Então,

$d(x,y) > \varepsilon$ diz que $x \neq y$, mas $d(f^j(x), f^j(y)) \leq \delta$ para todo $j$, por expansividade, diz que $x = y$. Esta contradição prova que $d(x,y) \leq \varepsilon$. Agora, se $\delta < \varepsilon$, temos imediatamente que $|\varphi|_f \leq L$. Mas, se $\delta > \varepsilon$ e $n \leq N$, então

$$\sup\{|\varphi_N(x) - \varphi_N(y)|\} \leq 2N||\varphi||.$$

Por último, se $\delta > \varepsilon$ e $n \geq N$, então como $d(x,y) \leq \varepsilon$, por expansividade, $d(f^k(x), f^k(y)) \leq \varepsilon$ para todo $0 \leq k \leq n-N$. Daí $|\varphi_{n-N}(x) - \varphi_{n-N}(y)| \leq L$. Nesse caso,

$$\begin{aligned}
|\varphi|_f &= \sup\{|\varphi_n(x) - \varphi_n(y)|\} \\
&\leq \sup\{|\varphi_{n-N}(x) - \varphi_{n-N}(y)|\} + \sup\{\sum_{i=n-N}^{n} |\varphi(f^i(x))| + \sum_{i=n-N}^{n} |\varphi(f^i(y))|\} \\
&< L + 2N||\varphi||.
\end{aligned}$$

A questão agora é que $|\varphi|_f$ é apenas uma seminorma em $V(f)$, já que $|\varphi|_f$ de qualquer potencial constante é zero. Este problema é resolvido quando definimos $|||\varphi|||$ como norma de $V(f)$. Logo $V(f)$ é um espaço de Banach. $\square$

Agora, daremos uma condição sobre o potencial $\varphi$ para que ele pertença a $V(f)$.

**Definição 6.4.** Para $\varepsilon > 0$ e $m \geq 0$, definimos a seguinte variação:

$$\mathrm{var}_m(f, \varphi, \varepsilon) := \sup\{|\varphi(x) - \varphi(y)| : d(f^j(x), f^j(y)) \leq \varepsilon \text{ para todo } |j| \leq m\}.$$

**Proposição 6.5.** *O potencial* $\varphi \in V(f)$, *se* $\sum_{m=0}^{\infty} \mathrm{var}_m(f, \varphi, \varepsilon) < \infty$.

*Demonstração.* Suponha $d(f^k(x), f^k(y)) < \varepsilon$ para todo $0 \leq k < n$. Queremos mostrar que $|\varphi_n(x) - \varphi_n(y)|$ é limitado por um certo $L$. O primeiro passo é definir $m_k = \min\{k, n-k-1\}$. O motivo desta definição é fazer com que $d(f^j(f^k(x)), f^j(f^k(y))) < \varepsilon$ para todo $|j| \leq m_k$, para que possamos fazer estimativas com a variação. Então

$$\begin{aligned}
\mathrm{var}_{m_k}(f, \varphi, \varepsilon) &\geq \sup\{|\varphi(f^k(x)) - \varphi(f^k(y))| : d(f^j(f^k(x)), f^j(f^k(y))) < \varepsilon \ \forall\ |j| \leq m_k\} \\
&\geq |\varphi(f^k(x)) - \varphi(f^k(y))|.
\end{aligned}$$

Portanto

$$\begin{aligned}|\varphi_n(x)-\varphi_n(y)| &\leq \sum_{k=0}^{n-1} |\varphi(f^k(x))-\varphi(f^k(y))| \\ &\leq \sum_{k=0}^{n-1} \operatorname{var}_{m_k}(f,\varphi,\varepsilon) \\ &\leq 2\sum_{m=0}^{\infty} \operatorname{var}_m(f,\varphi,\varepsilon).\end{aligned}$$

A última desigualdade segue porque $m_k = m$ no máximo para dois índices $k$, por exemplo, se fosse $0 \leq n < 5$, então $m_0 = 0$, $m_1 = 1$, $m_2 = 2$, $m_3 = 2$, $m_4 = 1$, $m_5 = 0$. A prova termina tomando $L = 2\sum_{m=0}^{\infty} \operatorname{var}_m(f,\varphi,\varepsilon) < \infty$. □

Este capítulo será dedicado a provar o seguinte Teorema devido a Bowen [10]:

**Teorema 6.6** (**Bowen**). *Seja $f : M \to M$ um homeomorfismo expansivo no espaço métrico compacto $M$ com a propriedade de especificação. Então cada potencial $\varphi \in V(f)$, possui um único estado de equilíbrio, que denotaremos por $\mu = \mu_\varphi$. Além disso $\mu$ é misturador com respeito a $f$ e*

$$\mu = \lim_{n\to\infty} \frac{1}{S_n^{\mathrm{Fix}}(f,\varphi)} \sum_{x\in \mathrm{Fix}\,(f^n)} e^{\varphi_n(x)}\delta_x$$

*onde*

$$S_n^{\mathrm{Fix}}(f,\varphi) = \sum_{x\in \mathrm{Fix}\,(f^n)} e^{\varphi_n(x)}.$$

O potencial nulo pertence a $V(f)$, portanto o Teorema 6.6 afirma que existe uma única medida de máxima entropia. Antes de passarmos aos detalhes da demonstração, mostraremos que tanto o conjunto das $f$ que possuem a propriedade de especificação quanto o conjunto $V(f)$ são não vazios. Faremos isso por meio do exemplo abaixo, no qual mostraremos que um iterado de um difeomorfismo de uma variedade compacta que satisfaz o Axioma A, quando restrito a decomposição de um de seus conjuntos básicos, satisfaz a propriedade de especificação, e que potenciais diferenciáveis ou Hölder, quando restritos a um conjunto básico $\Omega_s$, pertencem a $V(f|_{\Omega_s})$. Aproveitamos para discutir existência e unicidade de estados de equilíbrio nesse contexto.

Isto em particular dará outra demonstração para o teorema principal do capítulo 4 (Teorema 5.55).

## 6.2 Difeomorfismos Axioma A satisfazem a propriedade de especificação

**Exemplo 6.7.** Seja $f : M \to M$ difeomorfismo Axioma A de uma variedade compacta e $\Omega_s \subset M$ um conjunto básico para $f$. Então pelo Teorema de Decomposição Espectral (Teorema 5.34), $f(\Omega_s) = \Omega_s$ e $f|_{\Omega_s}$ é expansivo pelo Corolário 5.33 (pois $\Omega_s$ é um subconjunto fechado de um conjunto hiperbólico, portanto é hiperbólico), contudo $f|_{\Omega_s}$ pode não possuir a propriedade de especificação. Para contornar isso, usamos novamente o Teorema de Decomposição Espectral para escrever $\Omega_s = X_1 \cup \cdots \cup X_n$ como união disjunta de conjuntos compactos invariantes de modo que $f(X_i) = X_{i+1}$, $f(X_n) = X_1$ e $f^n|_{X_i}$ é topologicamente misturadora.

Mostraremos que, nessas condições, $f^n|_{X_i}$ possui a propriedade de especificação.

**Teorema 6.8** (**Teorema da especificação**). *Seja $\Lambda$ um conjunto hiperbólico compacto localmente maximal para um difeomorfismo $f$ tal que $f|_\Lambda$ é topologicamente misturadora. Então $f|_\Lambda$ tem a propriedade de especificação.*

Antes de demonstrarmos o Teorema da especificação, explicamos porque $X_i$ é um conjunto localmente maximal para o difeomorfismo $f^n$.

**Definição 6.9.** Um conjunto hiperbólico $\Lambda$ é dito **localmente maximal** (ou isolado) se existe uma vizinhança $U$ of $\Lambda$ em $M$ tal que $\Lambda = \bigcap_{j\in\mathbb{Z}} f^j(U)$.

**Definição 6.10.** Um conjunto hiperbólico $\Lambda$ tem **estrutura de produto local** se para $\varepsilon > 0$ suficientemente pequeno e quaisquer $x, y \in \Lambda$ tais que $d(x,y) < \varepsilon$, temos que $W^s_\varepsilon(x) \cap W^u_\varepsilon(y) \in \Lambda$.

**Teorema 6.11.** *Um conjunto hiperbólico $\Lambda$ tem estrutura de produto local se, e somente se, é localmente maximal.*

*Demonstração.* Veja a proposição 6.4.21 em [5]. $\square$

É consequência do Teorema da Decomposição Espectral (Teorema 5.34) que $X_i$ tem estrutura de produto local e consequentemente, é localmente maximal, conforme o

Teorema 6.11. O seguinte resultado também é consequência do Teorema da Decomposição Espectral.

**Lema 6.12.** *Se $\Lambda$ é um conjunto hiperbólico compacto localmente maximal para um difeomorfismo $f$ e $f|_\Lambda$ é topologicamente misturadora, então os pontos periódicos de $f$ são densos em $\Lambda$ e a variedade instável de cada ponto periódico é densa em $\Lambda$.*

*Demonstração.* Veja o Corolário 18.3.2 de [5]. $\square$

***Demonstração do Teorema da especificação.*** Começaremos mostrando que o conjunto de todas as variedades instáveis são uniformemente densas em um conjunto hiperbólico compacto localmente maximal para um difeomorfismo $f$, com $f|_\Lambda$ topologicamente misturadora.

**Lema 6.13.** *Seja $\Lambda$ um conjunto hiperbólico compacto localmente maximal para um difeomorfismo $f$ e $f : \Lambda \to \Lambda$ topologicamente misturadora. Se $\alpha > 0$, então existe $N \in \mathbb{N}$ tal que para $x, y \in \Lambda$ e $n > N$, temos $f^n(W^u_\alpha(x)) \cap W^s_\alpha(y) \neq \emptyset$.*

*Demonstração.* Um conjunto $Y$ em um espaço métrico $X$ é chamado $\varepsilon$-denso se $X$ é uma $\varepsilon$-vizinhança de $Y$. Temos estrutura de produto local, dada pelo Teorema 6.11, então existe $\delta^* > 0$ tal que $W^s_\delta(x) \cap W^u_\delta(y)$ consiste de no máximo um ponto sempre que $0 < \delta < \delta^*$ e uma função $\varepsilon(\delta)$ tal que $d(x, y) < \varepsilon(\delta)$ implica que $W^s_\delta(x) \cap W^u_\delta(y) \neq \emptyset$. Seja $\delta := \min\left\{\delta^*, \dfrac{\alpha}{2}, \dfrac{\varepsilon(\alpha/2)}{4}\right\}$. Para escolher $N$, tome um conjunto $\dfrac{\varepsilon(\alpha/2)}{2}$-denso $\{p_k : k = 1, \ldots, r\}$ formado por pontos periódicos (com períodos $t_k$). Pelo Lema 6.12 a variedade instável de cada ponto periódico de $\Lambda$ é densa em $\Lambda$, assim, existe $m_k$ tal que $f^{mt_k}(W^u_\delta(p_k))$ é $\varepsilon(\delta)$-denso para todo $m \geq m_k$ e todo $k$. Seja $N = \prod_{k=1}^r m_k t_k$ e note que $f^N(W^u_\delta(p_k))$ é $\varepsilon(\delta)$-denso para todo $k$.

Agora mostraremos que $N$ satisfaz o que desejamos: para $x, y \in \Lambda$, tome $j$ tal que $d(x, p_j) < \dfrac{\varepsilon(\alpha/2)}{2}$, $z \in f^N(W^u_\delta(p_j))$ tal que $d(y, z) \leq \varepsilon(\delta)$, e $w \in W^u_\delta(z) \cap W^s_\delta(y)$. Então

$$f^{-N}(w) \in W^u_\delta(f^{-N}(z)) \subset W^u_{2\delta}(p_j) \subset W^u_{\frac{\varepsilon(\alpha/2)}{2}}(p_j).$$

Assim,

$$d(f^{-N}(w), x) \leq d(f^{-N}(w), p_j) + d(x, p_j) < \frac{\varepsilon(\alpha/2)}{2} + \frac{\varepsilon(\alpha/2)}{2} = \varepsilon(\alpha/2).$$

Assim, existe $v \in W^s_{\alpha/2}(f^{-N}(w)) \cap W^u_{\alpha/2}(x)$ e

$$f^N(v) \in f^N(W^u_{\alpha/2}(x)) \cap W^s_{\alpha/2}(w) \subset f^N(W^u_\alpha(x)) \cap W^s_\alpha(y) \neq \emptyset$$

uma vez que $\delta < \dfrac{\alpha}{2}$. Para $x, y \in \Lambda$ e $n > N$, $f^n(W^u_\alpha(x)) \cap W^s_\alpha(y) \supset f^N(W^u_\alpha(f^{n-N}(x))) \cap W^s_\alpha(y) \neq \emptyset$. □

Prossigamos com a prova do Teorema da especificação. Para $\beta < \min\{\varepsilon, \delta^*\}$ e $\alpha = \dfrac{\beta}{3}$, tomamos o $N$ obtido com o Lema 6.13. Seja $J > N$ tal que $\lambda^J < \dfrac{1}{2}$, onde $\lambda$ é dado na definição 5.28 de conjunto hiperbólico. Sejam $I_1, \ldots, I_m$ intervalos de números inteiros $I_i = [a_i, b_i] \subset \mathbb{Z}$, com $a_{k+1} - b_k > J > N$ e $P : \bigcup_{i=1}^m I_i \to M$ uma função tal que para $t_1, t_2 \in I_i$, tem-se que $f^{t_2 - t_1}(P(t_1)) = P(t_2)$. Denote $x_1 = f^{-a_1}(P(a_1))$ e defina $x_2, \ldots, x_m$ da seguinte forma: supondo que $x_k$ é dado, pelo Lema 6.13 existe $x_{k+1}$ tal que

$$f^{a_{k+1}}(x_{k+1}) \in f^{a_{k+1} - b_k}(W^u_\alpha(f^{b_k}(x_k))) \cap W^s_\alpha(P(a_{k+1})) \tag{6.1}$$

porque $a_{k+1} - b_k > N$.

Para mostrar que $x := x_m$ possui a órbita desejada, note que, uma vez que $f^{a_k}(x_k) \in W^s_\alpha(P(a_k))$ por construção, $d(f^n(x_k), P(n)) = d(f^n(x_k), f^{n-a_k}(P(a_k))) \le \alpha = \dfrac{\beta}{3}$ e portanto concluímos pela desigualdade triangular que $d(f^n(x), P(n)) < \beta$, uma vez que esteja provado o seguinte Lema.

**Lema 6.14.** *$d(f^n(x), f^n(x_k)) \le \dfrac{2\beta}{3}$ para todo $n \in I_k$, $k \in 1, \ldots, m$.*

*Demonstração.* Mostraremos que $f^{b_k}(x) \in W^u_{\frac{2\beta}{3}}(f^{b_k}(x_k))$. Uma vez que $\sum_{i=0}^\infty \lambda^{Ji} < 2$ e $\alpha = \dfrac{\beta}{3}$, é suficiente mostrarmos que $f^{b_k}(x_{k+r}) \in W^u_{\alpha + \alpha\lambda^J + \cdots + \alpha\lambda^{J(r-1)}}(f^{b_k}(x_k))$. Se $r = 1$, isto segue da equação (6.1). Agora, suponha que $f^{b_k}(x_{k+r}) \in W^u_{\alpha + \alpha\lambda^J + \cdots + \alpha\lambda^{J(r-1)}}(f^{b_k}(x_k))$. Como $f^{b_{k+r}}(x_{k+r+1}) \in W^u_\alpha(f^{b_{k+r}}(x_{k+r}))$ e $b_{k+r} - b_k \ge rJ$ por construção, segue do Teorema da variedade estável que

$$\begin{aligned} d(f^{b_k}(x_{k+r+1}), f^{b_k}(x_{k+r})) &= d(f^{-rJ + rJ + b_k}(x_{k+r+1}), f^{-rJ+rJ+b_k}(x_{k+r})) \\ &\le \lambda^{rJ} d(f^{b_{k+r}}(x_{k+r+1}), f^{b_{k+r}}(x_{k+r})) \\ &\le \lambda^{rJ}\alpha \end{aligned}$$

e portanto $f^{b_k}(x_{k+r+1}) \in W^u_{\alpha\lambda^{Jr}}(f^{b_k}(x_{k+r}))$. Mas usando a hipótese de indução, resulta que $f^{b_k}(x_{k+r+1}) \in W^u_{\alpha + \alpha\lambda^J + \cdots + \alpha\lambda^{J(r-1)} + \alpha\lambda^{Jr}}(f^{b_k}(x_k))$. □

Resta mostrar que a órbita de $x$ é periódica. Precisamos do seguinte resultado.

**Teorema 6.15** (**Anosov Closing Lemma**). *Seja $\Lambda$ um conjunto hiperbólico para $f : U \to M$. Então existe uma vizinhança aberta $V \supset \Lambda$, $C > 0$ e $\varepsilon_0 > 0$ tal que para $\varepsilon < \varepsilon_0$ e qualquer $\varepsilon$-pseudo-órbita $(x_0, \ldots, x_m) \subset V$ existe um ponto $y \in U$ tal que $f^m(y) = y$ e $d(f^k(y), x_k) < C\varepsilon$ para $k = 0, \ldots, m-1$.*

*Demonstração.* veja o Teorema 6.4.15 de [5]. □

Iremos assumir que $\beta \leq \min\{\varepsilon/2C, \varepsilon, 2\delta^*\}/2$, onde $C$ é dado pelo Teorema 6.15 e tomaremos $J > N$ como antes. Se $q \geq b_m - a_1 + J$ é o período desejado, consideramos $I_1, \ldots, I_m, \{a_1 + q\}$, consideramos $P' = P$ se $t_j \in I_i$ e $P'(a_1 + q) = P(a_1)$. Da forma que foram construídos, estes intervalos têm espaçamento no mínimo $J$ entre si. Portanto, obtemos um ponto $x' := f^{a_1}(x) \in \Lambda$ tal que $d(x', f^q(x')) \leq d(x', P(a_1)) + d(f^q(x'), P(a_1)) \leq 2\beta \leq \dfrac{\varepsilon}{2C}$. Então temos uma $\dfrac{\varepsilon}{2C}$-pseudo-órbita e consequentemente, pelo Teorema 6.15, existe um ponto $z \in \Lambda$ periódico de período $q$, tal que $d(f^{n+a_1}(z), f^n(x')) < C\dfrac{\varepsilon}{2C} = \dfrac{\varepsilon}{2}$ para todo $n \in [0, q]$. Resulta que $d(f^{n+a_1}(z), P'(n)) \leq d(f^{n+a_1}(z), f^n(x')) + d(P'(n), f^n(x')) < \varepsilon$.

Isto prova o Teorema da especificação. □

Segue diretamente do Teorema 6.8 que $f^n|_{X_i}$ possui a propriedade de especificação.

**Exemplo 6.16.** Mostramos no exemplo 5.35, que a Ferradura de Smale satisfaz o Axioma A. Conforme o que mostramos acima, $f^n|_{X_i}$ possui a propriedade de especificação.

**Proposição 6.17.** *Se $\psi : M \to \mathbb{R}$ é diferenciável (ou é Hölder de expoente positivo), então $\varphi = \psi|_{\Omega_s} \in V(f|_{\Omega_s})$.*

*Demonstração.* Primeiramente, como $f$ é um difeomorfismo que satisfaz o Axioma A, resulta do Lema 5.56 que existem $\varepsilon > 0$ e $\alpha > 1$ de modo que se $x, y \in \Omega_s$ e $d(f^k(x); f^k(y)) \leq \varepsilon$ para todo $|k| \leq n$, então $d(x, y) < \alpha^{-n}$. Se $\psi$ é diferenciável, então tem uma constante de Lipschitz $L$. Então

$$\mathrm{var}_n(\varphi, f|_{\Omega_s}, \varepsilon) = \mathrm{var}_n(\psi|_{\Omega_s}, f|_{\Omega_s}, \varepsilon) \leq Ld(x, y) < L\alpha^{-n}.$$

Como $\alpha^{-1} \in (0,1)$, $\sum_{n=1}^{\infty} \mathrm{var}_n(\psi|_{\Omega_s}, f|_{\Omega_s}, \varepsilon) < \infty$ e assim $\varphi = \psi|_{\Omega_s} \in V(f|_{\Omega_s})$.

Igualmente, se $\psi$ é Hölder de expoente positivo, então

$$\mathrm{var}_n(\varphi, f|_{\Omega_s}, \varepsilon) \leq C\, d(x,y)^{\theta} < L\alpha^{-n\theta}$$

e também nesse caso, $\sum_{n=1}^{\infty} \mathrm{var}_n(\psi|_{\Omega_s}, f|_{\Omega_s}, \varepsilon) < \infty$. □

Com a proposição acima e o Teorema da Especificação, o Teorema 6.6 dá outra demonstração para o Teorema 5.55. De fato, se $\varphi$ é Hölder de expoente positivo, então $\varphi \in V(f|_{\Omega_s})$, então $\psi = \varphi + \varphi \circ f + \cdots + \varphi \circ f^{n-1} \in V(f^n|_{X_i})$: de fato $\psi \in C^0(M)$, e para todo $\varepsilon > 0$ tal que $d((f^n)^k(x), (f^n)^k(y)) \leq \varepsilon$ para todo $0 \leq k < n$, temos

$$\begin{aligned}
|\psi_n(x) - \psi_n(y)| &= \left|\sum_{j=0}^{n-1} [\psi((f^n)^j(x)) - \psi((f^n)^j(y))]\right| \\
&\leq \sum_{j=0}^{n-1} \left|\sum_{l=0}^{n-1} \varphi(f^l(f^{nj}(x))) - \varphi(f^l(f^{nj}(y)))\right| \\
&\leq nL
\end{aligned}$$

onde $L$ é tal que $|\varphi_n(x) - \varphi_n(y)| \leq L$, como na definição 6.2. Assim, se $\mu_i$ é o único estado de equilíbrio de $f^n|_{X_i}$ com o potencial $\psi|_{X_i}$, podemos mostrar de modo semelhante ao que fizemos no final da demonstração do Teorema 5.55 que $\mu = \dfrac{1}{n}(\mu_1 + \cdots + \mu_n) \in \mathcal{M}_1(f)$ é o único estado de equilíbrio para $f|_{\Omega_s}$ e $\varphi$.

## 6.3 Estimativas da pressão topológica

Relembre que um conjunto $E \subset M$ é $(n,\varepsilon)$-separado (com respeito a $f$) se dados $x, y \in E$ pontos distintos, existe um inteiro $k$ tal que $0 \leq k \leq n$ e $d(f^k(x), f^k(y)) > \varepsilon$. Definimos $S_n(f,\varphi,\varepsilon) = \sup\{\sum_{x\in E} e^{\varphi_n(x)} : E$ é conjunto $(n,\varepsilon)$-separado$\}$. (Lembre que $\varphi_n(x) = \sum_{i=0}^{n-1} \varphi(f^i(x))$).

Denotemos $P_\nu(f,\varphi) = h_\nu(f) + \int \varphi d\nu$. Pelo Princípio Variacional (Teorema 3.57), sabemos que $P(f,\varphi) = \sup\{P_\nu(f,\varphi) : \nu \in \mathcal{M}_1(f)\}$. Seja $\mathrm{Fix}(f^n) = \{x \in M : f^n(x) = x\}$ e

$$S_n^{\mathrm{Fix}}(f,\varphi) = \sum_{x \in \mathrm{Fix}\,(f^n)} e^{\varphi_n(x)}. \tag{6.2}$$

Definimos

$$\mu_{\varphi,n} = \frac{1}{S_n^{\mathrm{Fix}}(f,\varphi)} \sum_{x\in \mathrm{Fix}\,(f^n)} e^{\varphi_n(x)}\delta_x$$

onde $\delta_x$ é a medida de Dirac, tal que $\delta_x(\{x\}) = 1$. Note que para todo $n$, $\mu_{\varphi,n} \in \mathcal{M}_1(f)$, pois de fato é uma probabilidade:

$$\mu_{\varphi,n}(M) = \frac{1}{S_n^{\mathrm{Fix}}(f,\varphi)} \sum_{x\in \mathrm{Fix}\,(f^n)} e^{\varphi_n(x)}\delta_x(M) = \frac{S_n^{\mathrm{Fix}}(f,\varphi)}{S_n^{\mathrm{Fix}}(f,\varphi)} = 1,$$

e é invariante

$$\begin{aligned}\mu_{\varphi,n}(f^{-1}(E)) =& \frac{1}{S_n^{\mathrm{Fix}}(f,\varphi)} \sum_{x\in \mathrm{Fix}\,(f^n)\cap f^{-1}(E)} e^{\varphi_n(x)}\delta_x \\ =& \frac{1}{S_n^{\mathrm{Fix}}(f,\varphi)} \sum_{f(x)\in \mathrm{Fix}\,(f^n)\cap E} e^{\varphi_n(f(x))}\delta_{f(x)} \\ =& \mu_{\varphi,n}(E).\end{aligned}$$

A segunda igualdade ocorre porque $x \in \mathrm{Fix}\,(f^n)$ implica que $\varphi_n(x) = \varphi_n(f(x))$.

Agora, como $\mathcal{M}_1(f)$ é um espaço métrico compacto com a topologia fraca$^*$, existe uma subsequência $\mu_{\varphi,n_k}$, com $n_k \to \infty$ tal que,

$$\mu = \lim_{k\to\infty} \mu_{\varphi,n_k} \tag{6.3}$$

existe e $\mu \in \mathcal{M}_1(f)$.

Durante todo este capítulo, a sequência $\{n_k\}_{k\in\mathbb{N}}$ estará fixada e $\mu$ será o limite como na equação (6.3). Note que, em virtude do corolário 3.67, existe pelo menos um estado de equilíbrio $\nu$ para qualquer potencial $\varphi$. Então nossa preocupação é lidar com a unicidade desse estado de equilíbrio. O que mostraremos é o seguinte: se $P_\nu(f,\varphi) = P(f,\varphi)$, então $\nu = \mu$.

Observe ainda que isto implica que se $\mu' = \lim_{k\to\infty} \mu_{\varphi,n'_k}$ então $P_{\mu'}(f,\varphi) = P(f,\varphi)$ e portanto $\mu' = \mu$. Resultará disso que $\mu = \lim_{n\to\infty} \mu_{\varphi,n}$. Daremos início a prova do Teorema estimando os valores de $S_n(f,\varphi,\varepsilon)$.

***Demonstração do Teorema 6.6***.

**Lema 6.18.** *Para quaisquer $\varepsilon\,, \delta > 0$, suficientemente pequenos, existe uma constante $C_{\delta,\varepsilon}$ tal que $S_n(f,\varphi,\delta) \leq C_{\delta,\varepsilon}\, S_n(f,\varphi,\varepsilon)$ para todo $n \geq 0$.*

*Demonstração.* Assumiremos que $\varepsilon$ é suficientemente pequeno de modo que $2\varepsilon$ é uma constante de expansividade. Então por expansividade, existe $N$ de modo que

$$\text{se } d(f^k(x), f^k(y)) \leq 2\varepsilon \text{ para todo } |k| \leq N, \text{ então } d(x,y) \leq \delta \tag{6.4}$$

(provamos isso na proposição 6.3). Escolha $\alpha > 0$ pequeno de modo que se $d(x,y) \leq \alpha$, então

$$d(f^k(x), f^k(y)) \leq \delta \text{ para todo } k \text{ com } |k| \leq N. \tag{6.5}$$

Escolha $n > 2N$ e sejam $F$ um conjunto $(n,\varepsilon)$-separado maximal e $E$ um conjunto $(n,\delta)$-separado. Da maximalidade de $F$, segue que para cada $x \in E$, existe um $l(x) \in F$ de modo que $d(f^i(x), f^i(l(x))) \leq \varepsilon$ para todo $i \in \{0,\ldots,n-1\}$ (caso contrário, $F \cup \{x\}$ seria um conjunto $(n,\varepsilon)$-separado, contradizendo a maximalidade de $F$). Para $a \in F$, seja $E_a = \{x \in E : l(x) = a\}$. Se $x,\, y \in E_a$, então

$$d(f^i(x), f^i(y)) \leq d(f^i(x), f^i(l(x))) + d(f^i(l(x)), f^i(y)) \leq \varepsilon + \varepsilon = 2\varepsilon$$

para todo $i \in \{0,\ldots,n-1\}$. Assim, por indução na equação (6.4):

$$\begin{gathered}
d(f^k(f^N(x)), f^k(f^N(y))) \leq 2\varepsilon \text{ para todo } |k| \leq N \implies d(f^N(x), f^N(y)) \leq \delta \\
d(f^k(f^{N+1}(x)), f^k(f^{N+1}(y))) \leq 2\varepsilon \text{ para todo } |k| \leq N \implies d(f^{N+1}(x), f^{N+1}(y)) \leq \delta \\
\vdots \\
d(f^k(f^{n-N}(x)), f^k(f^{n-N}(y))) \leq 2\varepsilon \text{ para todo } |k| \leq N \implies d(f^{n-N}(x), f^{n-N}(y)) \leq \delta.
\end{gathered}$$

Portanto, $d(f^i(x), f^i(y)) \leq \delta$ para todo $i \in \{N,\ldots,n-N\}$. Como $x, y \in E_a$, e portanto a $E$ que é $(n,\delta)$-separado, segue que $\{x,\, y\}$ é $(n,\delta)$-separado. Juntando isto a equação (6.5), segue que, ou $d(x,y) > \alpha$ ou $d(f^n(x), f^n(y)) > \alpha$. De fato, se acontecesse $d(x,y) \leq \alpha$ e $d(f^n(x), f^n(y)) \leq \alpha$, teríamos $d(f^k(x), f^k(y)) \leq \delta$ para todo $k \in \{0,\ldots,N\}$ e $d(f^k(f^n(x)), f^k(f^n(y))) \leq \delta$ para todo $k \in \{-N,\ldots,0\}$. Como já temos que $d(f^i(x), f^i(y)) \leq \delta$ para todo $i \in \{N,\ldots,n-N\}$, $\{x,\, y\}$ não seria $(n,\delta)$-separado.

Seja $Q$ o número máximo de pontos que podem ser escolhidos do espaço métrico compacto $M \times M$ de modo que, quaisquer dois pontos estejam separados por uma distância maior ou igual que $\alpha$ (é possível fazer isso, pois, pela compacidade é possível cobrir $M$ com uma quantidade finita de bolas de diâmetro $\alpha$). Note que $\{(x, f^n(x)) : x \in E_a\}$ é um tal conjunto com essa propriedade, pois, equipando $M \times M$ com a distância $d_2$, de modo que $d_2((x, f^n(x)), (y, f^n(y))) = \max\{d(x, y), d(f^n(x), f^n(y))\}$, vemos que $d_2((x, f^n(x)), (y, f^n(y))) > \alpha$. Assim, $\#E_a \leq Q$. Além disso, para $x \in E_a$, $d(f^i(x), f^i(a)) = d(f^i(x), f^i(l(x))) \leq \varepsilon$ e então, como $\varphi \in V(f)$, $|\varphi_n(x) - \varphi_n(a)| \leq L$. Daí

$$\sum_{x\in E} e^{\varphi_n(x)} \leq \sum_{a\in F} (\#E_a) e^{L+\varphi_n(a)} \leq Qe^L \sum_{a\in F} e^{\varphi_n(a)} \leq Qe^L S_n(f, \varphi, \varepsilon) \tag{6.6}$$

Tomando o supremo sobre todos os conjuntos $E$ $(n, \delta)$-separados na equação (6.6), segue o resultado com $C_{\delta,\varepsilon} = Qe^L$. □

**Lema 6.19.** *Para $\varepsilon$ pequeno existem constantes positivas $E_\varepsilon$ e $D_\varepsilon$ de modo que*

$$\prod_{j=1}^{k} E_\varepsilon S_{n_j}(f, \varphi, \varepsilon) \leq S_{n_1+\cdots+n_k}(f, \varphi, \varepsilon) \leq \prod_{j=1}^{k} D_\varepsilon S_{n_j}(f, \varphi, \varepsilon)$$

*sempre que $n_1, \ldots, n_k \geq 1$.*

*Demonstração.* Seja $E$ um conjunto $(n_1 + \cdots + n_k, \varepsilon)$-separado e, para cada $j$, $F_j$ um conjunto $(n_j, \frac{\varepsilon}{2})$-separado maximal. Para $x \in E$, escolha $g(x) := (g_1(x), \cdots, g_k(x)) \in F_1 \times \cdots \times F_k$ de modo que $d(f^{n_1+\cdots+n_{j-1}+i}(x), f^i(g_j(x))) \leq \frac{\varepsilon}{2}$ para todo $0 \leq i < n_j$ (isto é possível, pois se fosse a desigualdade para algum $j$ e $i$ maior que $\frac{\varepsilon}{2}$, então $F_j \cup \{f^{n_1+\cdots+n_{j-1}}(x)\}$ seria $(n_j, \frac{\varepsilon}{2})$-separado, contrariando a maximalidade de $F_j$). O mapa $g : E \to F_1 \times \cdots \times F_k$ é injetivo: $g(x) = g(y) \Longrightarrow g_j(x) = g_j(y)$. Daí

$$\begin{aligned} d(f^{n_1+\cdots+n_{j-1}+i}(x), f^{n_1+\cdots+n_{j-1}+i}(y)) &\leq d(f^{n_1+\cdots+n_{j-1}+i}(x), f^i(g_j(x))) \\ &\quad + d(f^{n_1+\cdots+n_{j-1}+i}(y), f^i(g_j(y))) \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned}$$

para todo $0 \leq i < n_j$ e $1 \leq j \leq k$. Isto implica que $x = y$, já que $E$ é $(n_1 + \cdots + n_k, \varepsilon)$-

separado. Agora, como $\varphi \in V(f)$,

$$\begin{aligned}
|\varphi_{n_1+\cdots+n_k}(x) - \sum_{j=1}^{k} \varphi_{n_j}(g_j(x))| &= |\sum_{i=0}^{n_1+\cdots+n_k-1} \varphi(f^i(x)) - \sum_{j=1}^{k} \varphi_{n_j}(g_j(x))| \\
&= |\sum_{j=1}^{k} \sum_{i=0}^{n_j-1} \varphi(f^i(f^{n_1+\cdots+n_{j-1}}(x))) - \sum_{j=1}^{k} \varphi_{n_j}(g_j(x))| \\
&\leq \sum_{j=1}^{k} |\sum_{i=0}^{n_j-1} \varphi(f^i(f^{n_1+\cdots+n_{j-1}}(x))) - \varphi_{n_j}(g_j(x))| \\
&= \sum_{j=1}^{k} |\varphi_{n_j}(f^{n_1+\cdots+n_{j-1}}(x)) - \varphi_{n_j}(g_j(x))| \\
&\leq \sum_{j=1}^{k} L = kL.
\end{aligned}$$

Consequentemente,

$$\begin{aligned}
\sum_{x\in E} e^{\varphi_{n_1+\cdots+n_k}(x)} &\leq \sum_{x\in E} e^{kL+\sum_{j=1}^{k} \varphi_{n_j}(g_j(x))} \\
&\leq \sum_{j=1}^{k} \sum_{y_j\in F_j} e^{kL+\sum_{j=1}^{k} \varphi_{n_j}(y_j)} \\
&\leq \prod_{j=1}^{k} e^{L} \sum_{y_j\in F_j} e^{\varphi_{n_j}(y_j)} \\
&\leq \prod_{j=1}^{k} e^{L} S_{n_j}(f,\varphi,\frac{\varepsilon}{2}).
\end{aligned} \tag{6.7}$$

Pelo Lema 6.18, existe uma constante $C_{\varepsilon/2,\varepsilon}$ tal que $S_{n_j}(f,\varphi,\frac{\varepsilon}{2}) \leq C_{\varepsilon/2,\varepsilon}\, S_{n_j}(f,\varphi,\varepsilon)$ para todo $n_j \geq 0$. Então, tomamos $D_\varepsilon = e^L C_{\varepsilon/2,\varepsilon}$ e tomamos o supremo sobre todos os conjuntos $E$ $(n_1+\cdots+n_k,\varepsilon)$-separados na equação (6.7). Assim conseguimos a segunda desigualdade $S_{n_1+\cdots+n_k}(f,\varphi,\varepsilon) \leq \prod_{j=1}^{k} D_\varepsilon S_{n_j}(f,\varphi,\varepsilon)$.

Agora, seja $E_j$ um conjunto $(n_j, 3\varepsilon)$-separado para cada $j$. Seja $I_j = [a_j, a_j+n_j-1]$, onde $a_j = n_1+\cdots+n_{j-1}+(j-1)p(\varepsilon)$ e $p(\varepsilon)$ é como na definição 6.1. Então para cada $z = (z_1,\ldots,z_k) \in E_1\times\cdots\times E_k$, podemos encontrar $x = x(z)$ tal que $d(f^{a_j+i}(x), f^i(z_j)) < \varepsilon$ para $0 \leq i < n_j$. Para ver isto, aplique a propriedade de especificação para $x_1 = f^{-a_1}(z_1),\ldots,x_k = f^{-a_k}(z_k)$. Existirá $x = x(z)$ tal que $d(f^r(x), f^r(x_j)) < \varepsilon$ para todo $r \in I_j$. Mas $r \in I_j \implies r = a_j+i$ com $0 \leq i < n_j$, ou seja $d(f^{a_j+i}(x), f^i(z_j)) = d(f^r(x), f^i(f^{a_j}(x_j))) = d(f^r(x), f^r(x_j)) < \varepsilon$.

O conjunto $E = \{x(z) : z \in E_1 \times \cdots \times E_k\}$ é $(m,\varepsilon)$-separado onde $m = n_1 + \cdots + n_k + (k-1)p(\varepsilon)$: de fato, se $x(z), x(w) \in E$ são quaisquer, então como $z, w \in E_1 \times \cdots \times E_k$, existem números $t_j$, com $0 \leq t_j < n_j$ para $j = 1, \ldots, k$ tais que $d(f^{t_j}(z_j), f^{t_j}(w_j)) > 3\varepsilon$. Por outro lado, pelo que fizemos acima, $d(f^{a_j+i}(x(z)), f^i(z_j)) < \varepsilon$ para $0 \leq i < n_j$ e $d(f^{a_j+i}(x(w)), f^i(w_j)) < \varepsilon$ para $0 \leq i < n_j$. Em particular, quando $i = t_j$ para algum $j \in \{1, \ldots, k\}$,

$$\begin{aligned}
d(f^{a_j+t_j}(x(z)), f^{a_j+t_j}(x(z))) &\geq d(f^{t_j}(z_j), f^{t_j}(w_j)) - d(f^{a_j+i}(x(z)), f^i(z_j)) \\
&\quad - d(f^{a_j+i}(x(w)), f^i(w_j)) \\
&> 3\varepsilon - 2\varepsilon = \varepsilon.
\end{aligned}$$

Além disso, vale que

$$\begin{aligned}
\varphi_m(x(z)) &= \sum_{i=0}^{n_1+\cdots+n_k+(k-1)p(\varepsilon)-1} \varphi(f^i(x(z))) \\
&= \sum_{j=1}^{k} \sum_{i=0}^{n_j-1} \varphi(f^i(f^{n_1+\cdots+n_{j-1}+(j-1)p(\varepsilon)}(x(z))))+ \\
&\quad + \sum_{j=1}^{k} \left( \sum_{i=n_1+\cdots+n_j+(j-1)p(\varepsilon)}^{n_1+\cdots+n_j+jp(\varepsilon)-1} \varphi(f^i(x(z))) \right) \\
&= \sum_{j=1}^{k} \varphi_{n_j}(f^{a_j}(x(z))) + \sum_{j=1}^{k} \left( \sum_{i=n_1+\cdots+n_j+(j-1)p(\varepsilon)}^{n_1+\cdots+n_j+jp(\varepsilon)-1} \varphi(f^i(x(z))) \right) \\
&\geq \sum_{j=1}^{k} \big(\varphi_{n_j}(z_j) - L\big) + \sum_{j=1}^{k} (-p(\varepsilon)||\varphi||) \\
&= \sum_{j=1}^{k} \varphi_{n_j}(z_j) - kL - kp(\varepsilon)||\varphi||.
\end{aligned} \tag{6.8}$$

A penúltima desigualdade ocorre porque $\varphi \in V(f)$ e $d(f^{a_j+i}(x), f^i(z_j)) < \varepsilon$ para $0 \leq i < n_j$ e porque $\varphi \geq -||\varphi||$. Resulta da equação (6.8) que

$$\sum_{x(z)\in E} e^{\varphi_m(x(z))} \geq e^{-kL-kp(\varepsilon)||\varphi||} \sum_{x\in E_j} e^{\sum_{j=1}^{k} \varphi_{n_j}(z_j)} = e^{-k(L+p(\varepsilon)||\varphi||)} \prod_{j=1}^{k} \sum_{x\in E_j} e^{\varphi_{n_j}(z_j)}$$

e tomando o supremo sobre todos os conjuntos $E, E_j$, $(m,\varepsilon)$, $(n_j, 3\varepsilon)$-separados, respectivamente,

$$S_m(f,\varphi,\varepsilon) \geq e^{-k(L+p(\varepsilon)||\varphi||)} \prod_{j=1}^{k} S_{n_j}(f,\varphi,3\varepsilon). \tag{6.9}$$

Pela primeira parte deste Lema,

$$\begin{aligned} S_m(f,\varphi,\varepsilon) &= S_{(n_1+\cdots+n_k)+(k-1)p(\varepsilon)} \\ &\leq D_\varepsilon\, S_{n_1+\cdots+n_k}(f,\varphi,\varepsilon) \prod_{j=1}^{k-1} D_\varepsilon S_{p(\varepsilon)}(f,\varphi,\varepsilon) \qquad (6.10) \\ &= D_\varepsilon^k\, S_{n_1+\cdots+n_k}(f,\varphi,\varepsilon)\, \left(S_{p(\varepsilon)}(f,\varphi,\varepsilon)\right)^{k-1}. \end{aligned}$$

Portanto, usando as equações (6.9) e (6.10),

$$S_{n_1+\cdots+n_k}(f,\varphi,\varepsilon) \geq \frac{e^{-k(L+p(\varepsilon)||\varphi||)}}{D_\varepsilon^k \left(S_{p(\varepsilon)}(f,\varphi,\varepsilon)\right)^{k-1}} \prod_{j=1}^{k} S_{n_j}(f,\varphi,3\varepsilon).$$

Pelo Lema 6.18, existe uma constante $C_{\varepsilon,3\varepsilon}$ tal que $S_{n_j}(f,\varphi,\varepsilon) \leq C_{\varepsilon,3\varepsilon}\, S_{n_j}(f,\varphi,3\varepsilon)$ para todo $n_j \geq 0$. Então, tomamos

$$E_\varepsilon = \frac{e^{-(L+p(\varepsilon)||\varphi||)}}{D_\varepsilon \max\{1, S_{p(\varepsilon)}(f,\varphi,\varepsilon)\} C_{\varepsilon,3\varepsilon}}$$

e concluímos a prova do lema. $\square$

**Lema 6.20.** *Para $\varepsilon$ pequeno e todo $n$,*

$$\frac{1}{D_\varepsilon} e^{nP} \leq S_n(f,\varphi,\varepsilon) \leq \frac{1}{E_\varepsilon} e^{nP}$$

*onde $P = P(f,\varphi)$ é a pressão do potencial $\varphi$.*

*Demonstração.* Recorde que a $P = P(f,\varphi) = \lim_{\varepsilon\to 0} \limsup_{n\to\infty} \dfrac{1}{n} \log S_n(f,\varphi,\varepsilon)$ (veja a definição 3.54 e o Teorema 3.55). Suponha que $S_n(f,\varphi,\varepsilon) > \dfrac{e^{nP}}{E_\varepsilon}$. Então, pelo Lema 6.19, $S_{kn}(f,\varphi,\varepsilon) \geq (E_\varepsilon S_n(f,\varphi,\varepsilon))^k$. Daí, resulta que

$$\begin{aligned} P &= \lim_{\varepsilon\to 0} \limsup_{k\to\infty} \frac{1}{kn} \log S_{kn}(f,\varphi,\varepsilon) \\ &\geq \lim_{\varepsilon\to 0} \limsup_{k\to\infty} \frac{1}{kn} \log \left(E_\varepsilon S_n(f,\varphi,\varepsilon)\right)^k \\ &= \lim_{\varepsilon\to 0} \frac{1}{n} \log E_\varepsilon S_n(f,\varphi,\varepsilon) \\ &> \lim_{\varepsilon\to 0} \frac{1}{n} nP = P. \end{aligned}$$

Esta contradição prova que a desigualdade do lado direito é verdadeira. A outra desigualdade é verificada de maneira similar. □

O próximo Lema diz que a entropia é a taxa exponencial do crescimento da quantidade de órbitas periódicas.

**Lema 6.21.** *Existem números positivos $d_1$ e $d_2$ de modo que*

$$d_1 e^{nP} \leq S_n^{\mathrm{Fix}}(f,\varphi) \leq d_2 e^{nP}$$

*para $n$ suficientemente grande. Aqui $P = P(f,\varphi)$ e $S_n^{\mathrm{Fix}}(f,\varphi)$ é como na equação* (6.2).

*Demonstração.* Para $\varepsilon$ pequeno (menor que uma constante de expansividade para $f$), o conjunto $\mathrm{Fix}(f^n)$ é $(n,\varepsilon)$-separado. De fato, se $x, y \in \mathrm{Fix}(f^n)$ e $d(f^i(x), f^i(y)) \leq \varepsilon$ para todo $i = 0, \ldots, n-1$, então, $d(f^i(x), f^i(y) < \varepsilon$ para todo $i \geq 0$, porque $f^n(x) = x$ e $f^n(y) = y$. Por expansividade, isso implica que $x = y$. Consequentemente,

$$S_n^{\mathrm{Fix}}(f,\varphi) = \sum_{x \in \mathrm{Fix}\,(f^n)} e^{\varphi_n(x)} \leq \sup \left\{ \sum_{x \in E} e^{\varphi_n(x)} : E \text{ é } (n,\varepsilon)\text{-separado} \right\} = S_n(f,\varphi,\varepsilon).$$

Mas o Lema 6.20, dá que $S_n(f,\varphi,\varepsilon) \leq \dfrac{1}{E_\varepsilon} e^{nP}$, então tome $d_2 = \dfrac{1}{E_\varepsilon}$.

Considere agora $n \geq p(\varepsilon)$ e $E$ um conjunto $(n - p(\varepsilon), 3\varepsilon)$-separado. Pela propriedade de especificação aplicada ao único intervalo de números inteiros $[0, n - p(\varepsilon))$, para cada $z \in E$, encontramos um ponto $x = x(z)$ tal que $f^{n-p(\varepsilon)-0+p(\varepsilon)}(x) = x$, ou seja, $x = x(z) \in \mathrm{Fix}(f^n)$ e $d(f^i(z), f^i(x(z))) \leq \varepsilon$ para $i = 0, \ldots, n - p(\varepsilon) - 1$. Então o mapa $x(\cdot)$ é injetivo : $x(z) \neq x(z')$ para $z \neq z'$. De fato, se $z$, $z' \in E$, então existe $l \in \{0, \ldots, n - p(\varepsilon) - 1\}$ tal que $d(f^l(z), f^l(z')) > 3\varepsilon$. Agora $x(z), x(z') \in \mathrm{Fix}(f^n)$ pelo que mostramos acima. Daí, $3\varepsilon < d(f^l(z), f^l(z')) \leq d(f^l(x(z)), f^l(x(z'))) + d(f^l(z), f^l(x(z))) + d(f^l(z'), f^l(x(z'))) \leq 2\varepsilon + d(f^l(x(z)), f^l(x(z')))$. Ou seja, $d(f^l(x(z)), f^l(x(z'))) > \varepsilon$. Como mostramos no início $\mathrm{Fix}(f^n)$ é $(n,\varepsilon)$-separado, logo $x(z) \neq x(z')$.

Além disso, como $d(f^i(z), f^i(x(z))) \leq \varepsilon$ para $k = 0, \ldots, n - p(\varepsilon) - 1$, $\varphi \in V(f)$ e

$\varphi(f^i(x(z))) \geq -||\varphi||$,

$$\begin{aligned}
\varphi_n(x(z)) &= \sum_{i=0}^{n-p(\varepsilon)+p(\varepsilon)} \varphi(f^i(x(z))) \\
&= \sum_{i=0}^{n-p(\varepsilon)-1} \varphi(f^i(x(z))) + \sum_{i=n-p(\varepsilon)}^{n-p(\varepsilon)+p(\varepsilon)} \varphi(f^i(x(z))) \\
&\geq \left( \sum_{i=0}^{n-p(\varepsilon)-1} \varphi(f^i(z)) \right) - L - p(\varepsilon)||\varphi|| \\
&= \varphi_{n-p(\varepsilon)}(z) - L - p(\varepsilon)||\varphi||.
\end{aligned}$$

Consequentemente,

$$S_n^{\text{Fix}}(f,\varphi) = \sum_{x \in \text{Fix}\,(f^n)} e^{\varphi_n(x)} \geq \sum_{x(z) \in \text{Fix}\,(f^n)} e^{\varphi_n(x(z))} \geq \sum_{z \in E} e^{\varphi_{n-p(\varepsilon)}(z) - L - P(\varepsilon)||\varphi||}.$$

Tomando o supremo sobre todos os conjuntos $E$ que são $(n-p(\varepsilon), 3\varepsilon)$-separado, vem que

$$S_n^{\text{Fix}}(f,\varphi) \geq e^{-L-P(\varepsilon)||\varphi||} S_{n-p(\varepsilon)}(f,\varphi,3\varepsilon).$$

Agora, pelo Lema 6.20,

$$e^{-L-P(\varepsilon)||\varphi||} S_{n-p(\varepsilon)}(f,\varphi,3\varepsilon) \geq \frac{1}{D_{3\varepsilon}} e^{-L-P(\varepsilon)||\varphi||} e^{(n-p(\varepsilon))P} = \frac{e^{-L-P(\varepsilon)||\varphi||}}{e^{p(\varepsilon)P}\, D_{3\varepsilon}} e^{nP}.$$

Finalmente defina $d_1 = \dfrac{e^{-L-P(\varepsilon)||\varphi||}}{e^{p(\varepsilon)P}\, D_{3\varepsilon}}$. Assim a demonstração está completa. □

## 6.4 $\mu$ é um estado de Gibbs

Com as estimativas anteriores, podemos provar que $\mu$ é um estado de Gibbs (ver definição 4.14).

**Lema 6.22.** *Para cada $\varepsilon$ pequeno, existem constantes $A_\varepsilon > 0$ e $B_\varepsilon > 0$ de modo que, para cada $y \in M$ e $n \geq 1$, temos que*

$$A_\varepsilon e^{\varphi_n(y)-nP} \leq \mu(B(y,n,\varepsilon)) \leq B_\varepsilon e^{\varphi_n(y)-nP}.$$

*O conjunto $B(y,n,\varepsilon) = \{x \in M : d(f^i(x), f^i(y)) \leq \varepsilon$ para $i = 0, \ldots, n-1\}$ é a bola dinâmica, como definimos em 3.41.*

*Demonstração.* Seja $E_m$ um conjunto $(m, 3\varepsilon)$-separado e $r = n + m + 2p(\varepsilon)$. Pela propriedade de especificação aplicada aos intervalos de números inteiros $I_1 = [0, n)$ e $I_2 = [n+p(\varepsilon), n+p(\varepsilon)+m)$, para cada $z \in E_m$ e $y \in M$ (aqui $x_1 = y$ e $x_2 = f^{-n-p(\varepsilon)}(z)$ como na definição de especificação), encontramos um ponto $x = x(z)$ tal que $f^{n+p(\varepsilon)+m-0+p(\varepsilon)}(x) = x$, ou seja, $x = x(z) \in \mathrm{Fix}(f^r)$ e $d(f^i(x(z)), f^i(y)) \leq \varepsilon$ para $i = 0, \cdots, n-1$. Também $\varepsilon > d(f^{j+n+p(\varepsilon)}(x(z)), f^{j+n+p(\varepsilon)}(x_2)) = d(f^{j+n+p(\varepsilon)}(x(z)), f^{j+n+p(\varepsilon)}(f^{-n-p(\varepsilon)}(z))) = d(f^{j+n+p(\varepsilon)}(x(z)), f^j(z))$ para $j = 0, \ldots, m-1$. Note que o mapa $x(\cdot)$ é injetivo em $E_m$, como mostramos no Lema anterior: $x(z) \neq x(z')$ para $z \neq z'$. Além disso, $x(z) \in B(y, n, \varepsilon) \cap \mathrm{Fix}\,(f^r)$ e

$$\begin{aligned}
|\varphi_r(x(z)) - \varphi_n(y) - \varphi_m(z)| &= |(\varphi_n(x(z)) - \varphi_n(y)) + \sum_{j=0}^{r-n} \varphi(f^j(f^n(x(z)))) - \varphi_m(z)| \\
&\leq L + |\varphi_{r-n}(f^n(x(z))) - \varphi_m(z)| \\
&= L + |\varphi_{m+2p(\varepsilon)}(f^n(x(z))) - \varphi_m(z)| \\
&= L + |(\varphi_m(f^{n+p(\varepsilon)}(x(z))) - \varphi_m(z)) \\
&\qquad + \varphi_{p(\varepsilon)}(f^n(x(z))) + \varphi_{p(\varepsilon)}(f^{m+n+p(\varepsilon)}(x(z)))| \\
&\leq 2L + |\varphi_{p(\varepsilon)}(f^n(x(z))) + \varphi_{p(\varepsilon)}(f^{m+n+p(\varepsilon)}(x(z)))| \\
&\leq 2L + 2p(\varepsilon)||\varphi||.
\end{aligned}$$

A primeira desigualdade da equação acima é consequência de $d(f^i(x(z)), f^i(y)) \leq \varepsilon$ para $i = 0, \ldots, n-1$ e de $\varphi \in V(f)$, enquanto que a segunda é consequência de $d(f^{j+n+p(\varepsilon)}(x(z)), f^j(z)) < \varepsilon$ para todo $j = 0, \ldots, m-1$ e de $\varphi \in V(f)$. Consequentemente,

$$-2L - 2p(\varepsilon)||\varphi|| + \varphi_n(y) + \varphi_m(z) \leq \varphi_r(x(z))$$

e portanto,

$$\begin{aligned}
\mu_{\varphi,r}(B(y,n,\varepsilon)) &= \frac{1}{S_r^{\mathrm{Fix}}(f,\varphi)} \sum_{x \in \mathrm{Fix}\,(f^r) \cap B(y,n,\varepsilon)} e^{\varphi_r(x)} \\
&\geq \frac{1}{S_r^{\mathrm{Fix}}(f,\varphi)} \sum_{x \in \mathrm{Fix}\,(f^r) \cap B(y,n,\varepsilon)} e^{-2L-2p(\varepsilon)||\varphi||+\varphi_n(y)+\varphi_m(z)} \\
&\geq \frac{1}{S_r^{\mathrm{Fix}}(f,\varphi)} e^{-2L-2p(\varepsilon)||\varphi||+\varphi_n(y)} \sum_{z \in E_m} e^{\varphi_m(z)}.
\end{aligned}$$

Tomando o supremo do lado direito sobre todos os conjuntos $E_m$ que são $(m, 3\varepsilon)$-separados, temos que

$$\mu_{\varphi,r}(B(y,n,\varepsilon)) \geq \frac{1}{S_r^{\mathrm{Fix}}(f,\varphi)} e^{-2L-2p(\varepsilon)||\varphi||+\varphi_n(y)} S_m(f,\varphi,3\varepsilon).$$

Pelo Lema, 6.21,

$$\frac{1}{S_r^{\mathrm{Fix}}(f,\varphi)} \geq \frac{1}{d_2 e^{rP}}.$$

Pelo Lema, 6.20,

$$S_m(f,\varphi,3\varepsilon) \geq \frac{1}{D_{3\varepsilon}} e^{mP}.$$

Combinando as três últimas desigualdades, vem que

$$\begin{aligned}
\mu_{\varphi,r}(B(y,n,\varepsilon)) &\geq \frac{1}{D_{3\varepsilon}} e^{mP} \frac{1}{d_2 e^{rP}} e^{-2L-2p(\varepsilon)||\varphi||+\varphi_n(y)} \\
&= \frac{1}{D_{3\varepsilon}} \frac{1}{d_2 e^{(r-m)P}} e^{-2L-2p(\varepsilon)||\varphi||} e^{\varphi_n(y)} \\
&= \frac{1}{D_{3\varepsilon}} \frac{1}{d_2 e^{(n+2p(\varepsilon))P}} e^{-2L-2p(\varepsilon)||\varphi||} e^{\varphi_n(y)} \\
&= \frac{1}{D_{3\varepsilon}} \frac{1}{d_2 e^{2p(\varepsilon)P}} e^{-2L-2p(\varepsilon)||\varphi||} e^{\varphi_n(y)-nP} \\
&= A_\varepsilon e^{\varphi_n(y)-nP}
\end{aligned}$$

onde $A_\varepsilon = \dfrac{1}{D_{3\varepsilon}} \dfrac{1}{d_2 e^{2p(\varepsilon)P}} e^{-2L-2p(\varepsilon)||\varphi||}$. Isto vale para todo $r \geq n + 2p(\varepsilon)$. Então seja $r = n_k$ com $k \to \infty$. Como $\mu_{\varphi,n_k} \to \mu$ na topologia fraca$^*$ e $B(y,n,\varepsilon)$ é um conjunto fechado,

$$\mu(B(y,n,\varepsilon)) \geq \limsup_{k\to\infty} \mu_{\varphi,n_k}(B(y,n,\varepsilon)) \geq A_\varepsilon e^{\varphi_n(y)-nP}. \tag{6.11}$$

Para obter a outra desigualdade, note que para $x \in \mathrm{Fix}(f^r) \cap B(y,n,\varepsilon)$ temos

$$\begin{aligned}
|\varphi_r(x) - \varphi_n(y) - \varphi_{r-n}(f^n(x))| &= |\varphi_n(x) + \varphi_{r-n}(f^n(x)) - \varphi_n(y) - \varphi_{r-n}(f^n(x))| \\
&= |\varphi_n(x) - \varphi_n(y)| \leq L
\end{aligned}$$

e note que $f^n(\mathrm{Fix}(f^r)) \cap B(y,n,\varepsilon)$ é $(r-n,\varepsilon)$-separado se $\varepsilon$ é pequeno (de modo a ser menor que uma constante de expansividade para $f$). Portanto,

$$\begin{aligned}
\mu_{\varphi,r}(B(y,n,\varepsilon)) &= \frac{1}{S_r^{\mathrm{Fix}}(f,\varphi)} \sum_{x\in \mathrm{Fix}\,(f^r)\cap B(y,n,\varepsilon)} e^{\varphi_r(x)} \\
&\leq \frac{1}{S_r^{\mathrm{Fix}}(f,\varphi)} \sum_{x\in \mathrm{Fix}\,(f^r)\cap B(y,n,\varepsilon)} e^{L+\varphi_n(y)+\varphi_{r-n}(f^n(x))} \\
&\leq \frac{1}{S_r^{\mathrm{Fix}}(f,\varphi)} e^{L+\varphi_n(y)} \sum_{x\in \mathrm{Fix}\,(f^r)\cap B(y,n,\varepsilon)} e^{\varphi_{r-n}(f^n(x))}.
\end{aligned}$$

Tomando o supremo do lado direito sobre todos os conjuntos $(r-n,\varepsilon)$-separados, temos que

$$\mu_{\varphi,r}(B(y,n,\varepsilon)) \le \frac{1}{S_r^{\mathrm{Fix}}(f,\varphi)} e^{L+\varphi_n(y)} S_{r-n}(f,\varphi,\varepsilon).$$

Pelo Lema, 6.21,

$$\frac{1}{S_r^{\mathrm{Fix}}(f,\varphi)} \le \frac{1}{d_1 e^{rP}}.$$

Pelo Lema, 6.20,

$$S_{r-n}(f,\varphi,\varepsilon) \le \frac{1}{E_\varepsilon} e^{(r-n)P}.$$

Combinando as três últimas desigualdades, vem que

$$\mu_{\varphi,r}(B(y,n,\varepsilon)) \le \frac{1}{d_1 e^{rP}} e^{L+\varphi_n(y)} \frac{1}{E_\varepsilon} e^{(r-n)P} = \frac{e^L}{d_1 E_\varepsilon} e^{\varphi_n(y)-nP} = B_\varepsilon e^{\varphi_n(y)-nP}$$

onde $B_\varepsilon = \dfrac{e^L}{d_1 E_\varepsilon}$. Então seja $r = n_k$ com $k \to \infty$. Como $\mu_{\varphi,n_k} \to \mu$ na topologia fraca* e $\mathrm{int}(B(y,n,\varepsilon))$ é um conjunto aberto,

$$\mu(\mathrm{int}(B(y,n,\varepsilon))) \le \liminf_{k\to\infty} \mu_{\varphi,n_k}(\mathrm{int}(B(y,n,\varepsilon))) \le B_\varepsilon e^{\varphi_n(y)-nP}. \tag{6.12}$$

Resulta das desigualdades (6.11) e (6.12) que

$$A_\varepsilon e^{\varphi_n(y)-nP} \le \mu(B(y,n,\varepsilon)) \le \mu(\mathrm{int}(B(y,n,2\varepsilon))) \le \tilde{B}_{2\varepsilon} e^{\varphi_n(y)-nP} := B_\varepsilon e^{\varphi_n(y)-nP},$$

e portanto a desigualdade afirmada no Lema é verdadeira. □

## 6.5 $\mu$ é medida misturadora

Nesta seção, vamos começar mostrando que $\mu$ é parcialmente misturadora, isto é, que o valor de $\mu(P \cap f^{-n}(Q))$ é limitado por múltiplos de $\mu(P)\mu(Q)$.

**Lema 6.23.** *Existem constantes $c_1 > 0$ e $c_2 > 0$ tais que*

$$c_1\mu(P)\mu(Q) \le \liminf_n \mu(P \cap f^{-n}(Q)) \le \limsup_n \mu(P \cap f^{-n}(Q)) \le c_2\mu(P)\mu(Q) \tag{6.13}$$

*para todos os conjuntos mensuráveis $P$, $Q \subset M$.*

*Demonstração.* Vamos demonstrar inicialmente a primeira desigualdade. Fixe algum $\varepsilon$ pequeno de modo que $3\varepsilon$ é inferior a uma constante de expansividade de $f$. Sejam $A$ e $B$ compactos com fronteira de medida nula e $U$, $V$ $\delta$-vizinhanças de $A$, $B$ respectivamente para algum $\delta > 0$. Por expansividade, existe $N(\delta)$ tal que

$$d(x,y) < \delta \text{ sempre que } d(f^j(x), f^j(y)) \leq \varepsilon \text{ para todo } |j| < N(\delta) \tag{6.14}$$

(provamos isso na proposição 6.3). Considere agora $n \geq 2N(\delta)$ e quaisquer outros inteiros $s$ e $t$. Seja $E_s$ um conjunto $(s, 3\varepsilon)$-separado e $E_t$ um conjunto $(t, 3\varepsilon)$-separado.

Defina os intervalos $I_j = [a_j, b_j] \subset \mathbb{Z}$ para $j = 1, 2, 3, 4$ da seguinte forma

$$\begin{aligned} a_1 &= -\left\lfloor \frac{n}{2} \right\rfloor & b_1 &= -\left\lfloor \frac{n}{2} \right\rfloor + n \\ a_2 &= b_1 + p(\varepsilon) & b_2 &= a_2 + s \\ a_3 &= b_2 + p(\varepsilon) & b_3 &= a_3 + n \\ a_4 &= b_3 + p(\varepsilon) & b_4 &= a_4 + t. \end{aligned}$$

Para cada $z = (z_1, z_2, z_3, z_4)$ em

$$f^{-\left\lfloor \frac{n}{2} \right\rfloor}(\mathrm{Fix}(f^n) \cap A) \times E_s \times f^{-\left\lfloor \frac{n}{2} \right\rfloor}(\mathrm{Fix}(f^n) \cap B) \times E_t$$

usamos a propriedade de especificação (para os pontos $x_j = f^{-a_j}(z_j)$) para encontrar um ponto $x = x(z)$ de modo que $f^{b_4 - a_1 + p(\varepsilon)}(x) = x$ e $d(f^{a_j+k}(x), f^k(z_j)) \leq \varepsilon$ para $0 \leq k < b_j - a_j$ e $j = 1, 2, 3, 4$. Agora afirmamos que $x \in U$ e que $f^{s+n+2p(\varepsilon)}(x) \in V$. De fato, $z_1 \in f^{-\left\lfloor \frac{n}{2} \right\rfloor}(\mathrm{Fix}(f^n) \cap A)$ implica que $f^{\left\lfloor \frac{n}{2} \right\rfloor}(z_1) = f^{-a_1}(z_1) \in A$. Portanto, para provar a primeira afirmação, é suficiente mostrar que $d(f^j(x), f^j(f^{-a_1}(z_1))) \leq \varepsilon$ para todo $|j| < N(\delta)$. Isto segue imediatamente de $d(f^{a_1+k}(x), f^k(z_1)) \leq \varepsilon$ para $0 \leq k \leq b_1 - a_1 = n$, pois reescrevendo $d(f^{a_1+k}(x), f^k(z_1)) = d(f^{a_1+k}(x), f^{a_1+k}(f^{-a_1}(z_1)))$ e tomando $j = a_1 + k$, vemos que $a_1 \leq j \leq a_1 + n$. Mas pela escolha de $N(\delta)$, $\left\lfloor \frac{n}{2} \right\rfloor \leq N(\delta)$. Assim, $a_1 \leq -N(\delta) \leq j \leq N(\delta) < a_1 + n$ e vale $d(f^j(x), f^j(f^{-a_1}(z_1))) \leq \varepsilon$ para todo $|j| < N(\delta)$. Mas por (6.14), segue que $d(x, f^{-a_1}(z_1)) \leq \delta$. Como $f^{-a_1}(z_1) \in A$ e $U$ é $\delta$-vizinhança de $A$, segue que $x \in U$. A segunda afirmação é provada de maneira semelhante: $f^{-a_1}(z_3) \in B \cap \mathrm{Fix}(f^n)$ e $d(f^{a_3+k}(x), f^k(z_3)) \leq \varepsilon$ para $0 \leq k \leq b_3 - a_3 = n$. Mas reescrevendo $d(f^{a_3+k}(x), f^k(z_3)) = d(f^{a_1+k}(f^{n+s+2p(\varepsilon)}(x)), f^{a_1+k}(f^{-a_1}(z_3)))$ e tomando

$j = a_1 + k$, vemos que $a_1 \leq j \leq a_1 + n$. Mas pela escolha de $N(\delta)$, $\left\lfloor \frac{n}{2} \right\rfloor \leq N(\delta)$. Assim, $a_1 \leq -N(\delta) \leq j \leq N(\delta) < a_1 + n$ e vale $d(f^j(f^{n+s+2p(\varepsilon)}(x)), f^j(f^{-a_1}(z_3))) \leq \varepsilon$ para todo $|j| < N(\delta)$. Mas por (6.14), segue que $d(f^{n+s+2p(\varepsilon)}(x), f^{-a_1}(z_3)) \leq \delta$. Como $f^{-a_1}(z_3) \in B$ e $V$ é $\delta$-vizinhança de $B$, segue que $f^{n+s+2p(\varepsilon)}(x) \in V$.

Portanto, $x \in U \cap f^{-n-s-2p(\varepsilon)}(V)$ e, como $E_s$ e $E_t$ são respectivamente $(s, 3\varepsilon)$ e $(t, 3\varepsilon)$ separados, a aplicação $x(\cdot)$ é injetiva. Seja $m := b_4 - a_1 + p(\varepsilon) = t + s + 2n + 4p(\varepsilon)$. Então

$$\begin{aligned}
\varphi_m(x(z)) &= \sum_{j=0}^{b_4-a_1+p(\varepsilon)} \varphi(f^j(x(z))) \\
&= \sum_{j=a_1}^{a_1+n-1} \varphi(f^j(x(z))) + \sum_{j=b_1}^{b_1+p(\varepsilon)-1} \varphi(f^j(x(z))) + \sum_{j=a_2}^{a_2+s-1} \varphi(f^j(x(z))) \\
&\quad + \sum_{j=b_2}^{b_2+p(\varepsilon)-1} \varphi(f^j(x(z))) + \sum_{j=a_3}^{a_3+n-1} \varphi(f^j(x(z))) + \sum_{j=b_3}^{b_3+p(\varepsilon)-1} \varphi(f^j(x(z))) \\
&\quad + \sum_{j=a_4}^{a_4+t-1} \varphi(f^j(x(z))) + \sum_{j=b_4}^{b_4+p(\varepsilon)-1} \varphi(f^j(x(z))) \\
&\geq \sum_{j=0}^{n-1} \varphi(f^j(f^{a_1}(x(z)))) - p(\varepsilon)||\varphi|| + \sum_{j=0}^{s-1} \varphi(f^j(f^{a_2}(x(z)))) - p(\varepsilon)||\varphi|| \\
&\quad + \sum_{j=0}^{n-1} \varphi(f^j(f^{a_3}(x(z)))) - p(\varepsilon)||\varphi|| + \sum_{j=0}^{t-1} \varphi(f^j(f^{a_4}(x(z)))) - p(\varepsilon)||\varphi|| \\
&\geq \varphi_n(z_1) - L + \varphi_s(z_2) - L + \varphi_n(z_3) - L + \varphi_t(z_4) - L - 4p(\varepsilon)||\varphi|| \\
&= \varphi_n(z_1) + \varphi_s(z_2) + \varphi_n(z_3) + \varphi_t(z_4) - 4L - 4p(\varepsilon)||\varphi||.
\end{aligned}$$

Consequentemente,

$$\sum_{z} e^{\varphi_m(x(z))} \geq e^{-4L-4p(\varepsilon)||\varphi||} \sum_{z_1} e^{\varphi_n(z_1)} \sum_{z_3} e^{\varphi_n(z_3)} S_s(f, \varphi, 3\varepsilon) S_t(f, \varphi, 3\varepsilon). \tag{6.15}$$

Como $f^n(f^{\left\lfloor \frac{n}{2} \right\rfloor}(z_1)) = f^{\left\lfloor \frac{n}{2} \right\rfloor}(z_1) \implies f^{\left\lfloor \frac{n}{2} \right\rfloor}(f^n(z_1)) = f^{\left\lfloor \frac{n}{2} \right\rfloor}(z_1) \implies f^n(z_1) = z_1$, resulta que $\varphi_n(z_1) = \varphi_n(f^{\lfloor \frac{n}{2} \rfloor}(z_1))$: de fato,

$$\begin{aligned}
\varphi_n(f^{\lfloor \frac{n}{2} \rfloor}(z_1)) &= \varphi(f^{\lfloor \frac{n}{2} \rfloor}(z_1)) + \cdots + \varphi(f^{\lfloor \frac{n}{2} \rfloor + n - 1}(z_1)) \\
&= \varphi(f^{\lfloor \frac{n}{2} \rfloor}(z_1)) + \cdots + \varphi(f^{n-1}(z_1)) + \varphi(f^{(n-1)+1}(z_1)) + \cdots \varphi(f^{\lfloor \frac{n}{2} \rfloor + n - 1}(z_1)) \\
&= \varphi(f^{\lfloor \frac{n}{2} \rfloor}(z_1)) + \cdots + \varphi(f^{n-1}(z_1)) + \varphi(z_1) + \cdots \varphi(f^{\lfloor \frac{n}{2} \rfloor - 1}(z_1)) \\
&= \varphi_n(z_1).
\end{aligned}$$

Assim,

$$\begin{aligned}
\sum_{z_1} e^{\varphi_n(z_1)} &= \sum_{z_1 \in f^{-\lfloor \frac{n}{2} \rfloor}(\mathrm{Fix}(f^n) \cap A)} e^{\varphi_n(z_1)} \\
&= \sum_{f^{\lfloor \frac{n}{2} \rfloor}(z_1) \in \mathrm{Fix}(f^n) \cap A} e^{\varphi_n(f^{\lfloor \frac{n}{2} \rfloor}(z_1))} \\
&= \sum_{y \in \mathrm{Fix}(f^n) \cap A} e^{\varphi_n(y)} \\
&= \mu_{\varphi,n}(A) S_n^{\mathrm{Fix}}(f, \varphi).
\end{aligned} \tag{6.16}$$

De modo análogo, provamos que

$$\sum_{z_3} e^{\varphi_n(z_3)} = \mu_{\varphi,n}(B) S_n^{\mathrm{Fix}}(f, \varphi). \tag{6.17}$$

Substituindo as equações (6.16) e (6.17) na equação (6.15), temos

$$\sum_z e^{\varphi_m(x(z))} \geq e^{-4L - 4p(\varepsilon)||\varphi||} \mu_{\varphi,n}(A) \mu_{\varphi,n}(B) [S_n^{\mathrm{Fix}}(f, \varphi)]^2 S_s(f, \varphi, 3\varepsilon) S_t(f, \varphi, 3\varepsilon). \tag{6.18}$$

Agora,

$$\begin{aligned}
\mu_{\varphi,m}(U \cap f^{-n-s-2p(\varepsilon)}(V)) &= [S_m^{\mathrm{Fix}}(f, \varphi)]^{-1} \sum_{x \in \mathrm{Fix}(f^m) \cap (U \cap f^{-n-s-2p(\varepsilon)})} e^{\varphi_m(x)} \\
&\geq [S_m^{\mathrm{Fix}}(f, \varphi)]^{-1} \sum_z e^{\varphi_m(x(z))}
\end{aligned} \tag{6.19}$$

porque $x(\cdot)$ é injetiva. Substituindo a equação (6.18) em (6.19), e usando os Lemas 6.20, 6.21, obtemos

$$\begin{aligned}
\mu_{\varphi,m}(U \cap f^{-n-s-2p(\varepsilon)}(V)) &\geq \frac{b[S_n^{\mathrm{Fix}}(f,\varphi)]^2}{S_m^{\mathrm{Fix}}(f,\varphi)} \mu_{\varphi,n}(A)\mu_{\varphi,n}(B) S_s(f,\varphi,3\varepsilon) S_t(f,\varphi,3\varepsilon) \\
&\geq \frac{b(d_1)^2 e^{2nP}}{d_2 e^{mP}} \mu_{\varphi,n}(A)\mu_{\varphi,n}(B) \frac{e^{sP}e^{tP}}{(D_{3\varepsilon})^2} \\
&= \frac{b(d_1)^2}{d_2(D_{3\varepsilon})^2} e^{(2n+s+t-m)P} \mu_{\varphi,n}(A)\mu_{\varphi,n}(B) \\
& \frac{b(d_1)^2}{d_2(D_{3\varepsilon})^2} e^{-4p(\varepsilon)P} \mu_{\varphi,n}(A)\mu_{\varphi,n}(B),
\end{aligned}$$

onde $b = e^{-4L-4p(\varepsilon)||\varphi||}$. Definimos $c_1 = \dfrac{b(d_1)^2}{d_2(D_{3\varepsilon})^2} e^{-4p(\varepsilon)P}$. Agora fixe $s$ e $n \geq N(\delta)$. Fazendo $t \to \infty$ e tomando $m = n_k$, temos que

$$\mu(\overline{U} \cap f^{-n-s-2p(\varepsilon)}(\overline{V})) \geq \limsup_{k\to\infty} \mu_{\varphi,n_k}(U \cap f^{-n-s-2p(\varepsilon)}(V)) \geq c_1 \mu_{\varphi,n}(A)\mu_{\varphi,n}(B).$$

Fazendo $s \to \infty$, temos que

$$\liminf_{r\to\infty} \mu(\overline{U} \cap f^{-r}(\overline{V})) \geq c_1 \mu_{\varphi,n}(A)\mu_{\varphi,n}(B).$$

Finalmente, fazendo $n \to \infty$, temos que

$$\liminf_{r\to\infty} \mu(\overline{U} \cap f^{-r}(\overline{V})) \geq c_1 \limsup_{k\to\infty} \mu_{\varphi,n_k}(A)\mu_{\varphi,n_k}(B).$$

Isto vale para quaisquer compactos $A$, $B$ e quaisquer vizinhanças $U \supset A$ e $V \supset B$. O lado direito converge para $c_1\mu(A)\mu(B)$. Fazendo $\delta \to 0$, encontramos a primeira desigualdade do Lema. Ressaltamos que é suficiente trabalhar com conjuntos compactos com fronteira de medida nula pois estes aproximam qualquer conjunto mensurável. Logo a propriedade acima vale para quaisquer $P$, $Q$ mensuráveis.

Agora provemos a outra desigualdade de maneira similar. Para $m = s + t + 2n + 4p(\varepsilon)$ e $x \in \mathrm{Fix}(f^m) \cap A \cap f^{-n-p(\varepsilon)}(B)$ existe único $z(x) = (z_1, z_2, z_3, z_4) \in \mathrm{Fix}(f^{n+p(\varepsilon)}) \times \mathrm{Fix}(f^{n+p(\varepsilon)}) \times \mathrm{Fix}(f^{s+p(\varepsilon)}) \times \mathrm{Fix}(f^{t+p(\varepsilon)})$, com $d(f^j(f^{a_i}(x)), f^j(f^{a_i}(z_i))) \leq \varepsilon$ e $0 \leq j < b_i - a_i$ para $i = 1, 2, 3, 4$, onde

$$\begin{aligned}
a_1 &= -\left\lfloor \frac{n}{2} \right\rfloor & \quad b_1 &= -\left\lfloor \frac{n}{2} \right\rfloor + n \\
a_2 &= b_1 + p(\varepsilon) & b_2 &= a_2 + n \\
a_3 &= b_2 + p(\varepsilon) & b_3 &= a_3 + s \\
a_4 &= b_3 + p(\varepsilon) & b_4 &= a_4 + t.
\end{aligned}$$

Cada $z_i$ é obtido aplicando a propriedade de especificação para $x$ nos intervalos $[a_1, a_1+n]$, $[a_2, a_2+n]$, $[a_3, a_3+s], [a_4, a_4+t]$, a unicidade de $z$ segue do fato de que o conjunto dos pontos fixos é separado. Como na primeira parte da demonstração desse Lema, $d(f^j(f^{a_i}(x)), f^j(f^{a_i}(z_i))) \leq \varepsilon$ para $0 \leq j < b_i - a_i$ e para $i = 1, 2$, implica que $d(x, z_1) < \delta$ e $d(f^{j+a_1}(f^{n+p(\varepsilon)}(x)), f^{j+a_1}((f^{n+p(\varepsilon)}(z_2))) < \delta$. Como $x \in A$ e $f^{n+p(\varepsilon)}(x) \in B$, isto implica que $z_1 \in U$ e $z_2 = f^{n+p(\varepsilon)}(z_2) \in V$. Além disso,

$$\begin{aligned}
\varphi_m(x) = & \sum_{j=0}^{b_4-a_1+p(\varepsilon)} \varphi(f^j(x)) \\
= & \sum_{j=a_1}^{a_1+n-1} \varphi(f^j(x)) + \sum_{j=b_1}^{b_1+p(\varepsilon)-1} \varphi(f^j(x)) + \sum_{j=a_2}^{a_2+n-1} \varphi(f^j(x)) \\
& + \sum_{j=b_2}^{b_2+p(\varepsilon)-1} \varphi(f^j(x)) + \sum_{j=a_3}^{a_3+s-1} \varphi(f^j(x)) + \sum_{j=b_3}^{b_3+p(\varepsilon)-1} \varphi(f^j(x)) \\
& + \sum_{j=a_4}^{a_4+t-1} \varphi(f^j(x)) + \sum_{j=b_4}^{b_4+p(\varepsilon)-1} \varphi(f^j(x)) \\
\leq & \sum_{j=0}^{n-1} \varphi(f^j(f^{a_1}(x))) + p(\varepsilon)||\varphi|| + \sum_{j=0}^{n-1} \varphi(f^j(f^{a_2}(x))) + p(\varepsilon)||\varphi|| \\
& + \sum_{j=0}^{s-1} \varphi(f^j(f^{a_3}(x))) + p(\varepsilon)||\varphi|| + \sum_{j=0}^{t-1} \varphi(f^j(f^{a_4}(x))) + p(\varepsilon)||\varphi|| \\
\leq & \varphi_n(f^{a_1}(z_1)) + L + \varphi_n(f^{a_2}(z_2)) + L + \varphi_s(f^{a_3}(z_3)) + L + \varphi_t(f^{a_4}(z_4)) + L + 4p(\varepsilon)||\varphi|| \\
= & \varphi_n(f^{a_1}(z_1)) + \varphi_n(f_2^a(z_2)) + \varphi_s(f^{a_3}(z_3)) + \varphi_t(f^{a_4}(z_4)) + 4L + 4p(\varepsilon)||\varphi||.
\end{aligned}$$

(usamos que $d(f^j(f^{a_i}(x)), f^j(f^{a_i}(z_i))) \leq \varepsilon$ e $0 \leq j < b_i - a_i$ para $i = 1, 2, 3, 4$ e que

$\varphi \in V(f))$. Assim,

$$\sum_{x\in \mathrm{Fix}(f^m)\cap A\cap f^{-n-p(\varepsilon)}(B)} e^{\varphi_m(x)} \leq \frac{1}{b} \sum_{z_1\in \mathrm{Fix}(f^{n+p(\varepsilon)})\cap U} e^{\varphi_n(f^{a_1}(z_1))} \sum_{z_2\in \mathrm{Fix}(f^{n+p(\varepsilon)})\cap V} e^{\varphi_n(f^{a_2}(z_2))}$$
$$\sum_{z_3\in \mathrm{Fix}(f^{s+p(\varepsilon)})} e^{\varphi_s(f^{a_3}(z_3))} \sum_{z_4\in \mathrm{Fix}(f^{t+p(\varepsilon)})} e^{\varphi_t(f^{a_4}(z_4))}$$
$$\leq \frac{1}{b} \sum_{z_1\in \mathrm{Fix}(f^{n+p(\varepsilon)})\cap U} e^{\varphi_{n+p(\varepsilon)}(f^{a_1}(z_1))} \sum_{z_2\in \mathrm{Fix}(f^{n+p(\varepsilon)})\cap V} e^{\varphi_{n+p(\varepsilon)}(f^{a_2}(z_2))}$$
$$\sum_{z_3\in \mathrm{Fix}(f^{s+p(\varepsilon)})} e^{\varphi_{s+p(\varepsilon)}(f^{a_3}(z_3))} \sum_{z_4\in \mathrm{Fix}(f^{t+p(\varepsilon)})} e^{\varphi_{t+p(\varepsilon)}(f^{a_4}(z_4))}$$
$$= \frac{1}{b} \sum_{z_1\in \mathrm{Fix}(f^{n+p(\varepsilon)})\cap U} e^{\varphi_{n+p(\varepsilon)}(z_1)} \sum_{z_2\in \mathrm{Fix}(f^{n+p(\varepsilon)})\cap V} e^{\varphi_{n+p(\varepsilon)}(z_2)}$$
$$\sum_{z_3\in \mathrm{Fix}(f^{s+p(\varepsilon)})} e^{\varphi_{s+p(\varepsilon)}(z_3)} \sum_{z_4\in \mathrm{Fix}(f^{t+p(\varepsilon)})} e^{\varphi_{t+p(\varepsilon)}(z_4)}$$
$$\leq \frac{1}{b} \sum_{z_1} e^{\varphi_{n+p(\varepsilon)}(z_1)} \sum_{z_2} e^{\varphi_{n+p(\varepsilon)}(z_2)} S_{s+p(\varepsilon)}(f,\varphi,\varepsilon) S_{t+p(\varepsilon)}(f,\varphi,\varepsilon).$$

onde $b = e^{-4L-4p(\varepsilon)||\varphi||}$. Agora,

$$\sum_{z_1\in \mathrm{Fix}(f^{n+p(\varepsilon)})\cap U} e^{\varphi_{n+p(\varepsilon)}(z_1)} = \mu_{\varphi,n+p(\varepsilon)}(U) S^{\mathrm{Fix}}_{n+p(\varepsilon)}(f,\varphi)$$

e

$$\sum_{z_2\in \mathrm{Fix}(f^{n+p(\varepsilon)})\cap V} e^{\varphi_{n+p(\varepsilon)}(z_2)} = \mu_{\varphi,n+p(\varepsilon)}(V) S^{\mathrm{Fix}}_{n+p(\varepsilon)}(f,\varphi).$$

Resulta que

$$\mu_{\varphi,m}(A\cap f^{-n-p(\varepsilon)}(B)) \leq \frac{1}{b} \frac{[S^{\mathrm{Fix}}_{n+p(\varepsilon)}(f,\varphi)]^2}{S^{\mathrm{Fix}}_m(f,\varphi)} \mu_{\varphi,n+p(\varepsilon)}(U)\mu_{\varphi,n+p(\varepsilon)}(V)$$
$$S_{s+p(\varepsilon)}(f,\varphi,\varepsilon) S_{t+p(\varepsilon)}(f,\varphi,\varepsilon).$$

Usando os Lemas 6.20 e 6.21, obtemos

$$\mu_{\varphi,m}(A\cap f^{-n-p(\varepsilon)}(B)) \leq \frac{1}{b} \frac{(d_2)^2 e^{2(n+p(\varepsilon))P}}{d_1 e^{mP}} \frac{e^{(s+t)P}}{(E_\varepsilon)^2} \mu_{\varphi,n+p(\varepsilon)}(U)\mu_{\varphi,n+p(\varepsilon)}(V).$$

Mas $s+t+2n-m = -4p(\varepsilon)$. Então a estimativa acima é

$$\mu_{\varphi,m}(A\cap f^{-n-p(\varepsilon)}(B)) \leq \frac{(d_2)^2 e^{-2p(\varepsilon)P}}{bd_1(E_\varepsilon)^2} \mu_{\varphi,n+p(\varepsilon)}(U)\mu_{\varphi,n+p(\varepsilon)}(V).$$

Definimos $c_2 = \dfrac{(d_2)^2 e^{-2p(\varepsilon)P}}{bd_1(E_\varepsilon)^2}$. Agora fixe $n \geq N(\delta)$. Fazendo $s, t \to \infty$ e tomando $m = n_k$, temos que

$$\mu(A \cap f^{-n-p(\varepsilon)}(B)) \leq \liminf_{k\to\infty} \mu_{\varphi,n_k}(A \cap f^{-s-n}(B)) \leq c_2 \mu_{\varphi,n+p(\varepsilon)}(U)\mu_{\varphi,n+p(\varepsilon)}(V).$$

Finalmente, fazendo $n \to \infty$, temos que

$$\limsup_{n\to\infty} \mu(A \cap f^{-n-p(\varepsilon)}(B)) \leq c_2 \limsup_{k\to\infty} \mu_{\varphi,n_k+p(\varepsilon)}(\overline{U})\mu_{\varphi,n_k+p(\varepsilon)}(\overline{V}).$$

Isto vale para quaisquer compactos $A$, $B$ e quaisquer vizinhanças $U \supset A$ e $V \supset B$. O lado direito converge para $c_2\mu(\overline{U})\mu(\overline{V})$. Fazendo $\delta \to 0$, encontramos a segunda desigualdade do Lema. □

Com mais um pouco de esforço, podemos mostrar que as desigualdades em (6.13) implicam que a medida $\mu$ é misturadora.

**Lema 6.24.** *Se $f$ for um homeomorfismo de um espaço métrico compacto $M$ e $\mu$ uma medida Boreliana de probabilidade $f$-invariante, tal que para quaisquer conjuntos Borelianos $P$, $Q$ as desigualdades em* (6.13) *valem, então o sistema $(f, \mu)$ é misturador.*

*Demonstração.* Mostraremos inicialmente que, a inequação do lado esquerdo em (6.13) implica que o sistema $(f \times f, \mu \times \mu)$ é ergódico. Sejam $A, B, C, D \subset M$ conjuntos Borelianos. Então

$$\begin{aligned}\liminf_{n\to\infty}(\mu \times \mu)((f \times f)^n(A \times C) \cap (B \times D)) &= \liminf_{n\to\infty}[\mu(f^n(A) \cap B)\mu(f^n(C) \cap D)] \\ &\geq (c_1)^2 \mu(A)\mu(B)\mu(C)\mu(D) \\ &= (c_1)^2 (\mu \times \mu)(A \times C)(\mu \times \mu)(B \times D).\end{aligned}$$

A mesma desigualdade vale se trocarmos $A \times C$ e $B \times D$ por uniões finitas e disjuntas de produto de conjuntos. Uma vez que tais conjuntos aproximam cada $P$, $Q \subset M \times M$ mensurável, temos que

$$\liminf_{n\to\infty}(\mu \times \mu)((f \times f)^n(P) \cap Q) \geq (c_1)^2 (\mu \times \mu)(P)(\mu \times \mu)(Q),$$

e $f \times f$ é ergódico com respeito a $\mu \times \mu$. (se não fosse ergódico existiria $P$ invariante com $0 < (\mu \times \mu)(P) < 1$. Tome $Q = M \setminus P$, teríamos $(\mu \times \mu)((f \times f)^n(P) \cap Q) = (\mu \times \mu)(P \cap Q) = 0$, enquanto que $(c_1)^2(\mu \times \mu)(P)(\mu \times \mu)(Q) > 0$ uma contradição).

Agora, seja $\nu$ a **medida diagonal** em $M \times M$ definida por $\nu(E) = \mu(\pi_1(E \cap \Delta))$, onde $\Delta = \{(x, x) : x \in M\}$ e $\pi_1 : M \times M \to M$ é a projeção na primeira coordenada. Considere também $\nu_n$ definida por $\nu_n(A \times B) = \mu(f^n(A) \cap B)$. A medida $\nu$, bem como $\nu_n$ são $(f \times f)$-invariantes. De fato, é suficiente verificar isto para os retângulos porque estes geram a $\sigma$-álgebra produto:

$$\begin{aligned}\nu((f \times f)^{-1}(A \times B)) &= \nu(f^{-1}(A) \times f^{-1}(B)) \\ &= \mu(f^{-1}(A) \cap f^{-1}(B)) = \mu(A \cap B) = \nu(A \times B)\end{aligned}$$

e

$$\begin{aligned}\nu_n((f \times f)^{-1}(A \times B)) &= \nu_n(f^{-1}(A) \times f^{-1}(B)) \\ &= \mu(f^{n-1}(A) \cap f^{-1}(B)) = \mu(f^{-1}(f^n(A) \cap B)) \\ &= \mu(f^n(A) \cap B) = \nu_n(A \times B).\end{aligned}$$

Pela desigualdade do lado direito em (6.13), temos que

$$\limsup_{n\to\infty} \nu_n(A \times B) = \limsup_{n\to\infty} \mu(f^n(A) \cap B) \leq c_2\mu(A)\mu(B) = c_2(\mu \times \mu)(A \times B) \tag{6.20}$$

Queremos mostrar que $\nu_n(A \times B)$ converge para $\nu(A \times B)$. Seja $\eta$ um ponto de acumulação qualquer da sequência $\nu_n$ na topologia fraca*. Se $A$, $B \subset M$ são conjuntos abertos, então

$$\eta(A \times B) \leq \liminf_{n\to\infty} \nu_n(A \times B) \leq \limsup_{n\to\infty} \nu_n(A \times B) \leq c_2(\mu \times \mu)(A \times B)$$

por (6.20). Tomando uniões finitas e disjuntas de produto de conjuntos abertos e notando que estes conjuntos aproximam cada $P \subset M \times M$ mensurável, deduzimos que $\eta(P) \leq c_2(\mu \times \mu)(P)$. Consequentemente, $\eta$ é absolutamente contínua com respeito a $\mu \times \mu$. Uma vez que $\eta$ é $(f \times f)$-invariante e $\mu \times \mu$ é ergódica, usamos o Lema 2.7 e concluímos que $\eta = \mu \times \mu$. Assim para quaisquer abertos $A$, $B$ com $\mu(\partial A) = \mu(\partial B) = 0$, temos

$$\lim_{n\to\infty} \mu(f^n(A) \cap B) = \lim_{n\to\infty} \nu_n(A \times B) = (\mu \times \mu)(A \times B) = \mu(A)\mu(B).$$

Concluímos, pelo Lema 5.18 que $f$ é misturadora com respeito a $\mu$. $\square$

## 6.6 Conclusão da prova do Teorema 6.6

**Lema 6.25.** *Se $\nu \in \mathcal{M}_1(f)$ e $P_\nu(f,\varphi) = P(f,\varphi)$, então $\mu = \nu$.*

*Demonstração.* Começaremos provando que, se $\nu$ é singular com respeito a $\mu$, então $P_\nu(f,\varphi) < P(f,\varphi)$.

Suponha que $\nu \perp \mu$. Então existe um conjunto $B \subset M$ $(\mu+\nu)$-mensurável tal que $f(B) = B$, $\mu(B) = 0$ e $\nu(B) = 1$. Para cada $n > 1$, definimos uma partição $\mathscr{P}_n$ da seguinte forma. Primeiro fixe $\varepsilon > 0$ pequeno (de modo que $2\varepsilon$ seja uma constante de expansividade para $f$). Seja $E_n$ um conjunto $(n, 2\varepsilon)$-separado maximal. Note que $B(x,n,\varepsilon) \cap B(y,n,\varepsilon) = \emptyset$ para $x$, $y \in E_n$ com $x \neq y$, pois se existisse elemento nessa interseção, digamos $z$, teríamos que $d(f^j(z), f^j(x)) \leq \varepsilon$ e $d(f^j(z), f^j(y)) \leq \varepsilon$ para todo $0 \leq j < n$. Consequentemente $d(f^j(y), f^j(x)) \leq 2\varepsilon$ para todo $0 \leq j < n$ e isto implica que $x = y$, contradição. Mas sendo $E_n$ um conjunto $(n, 2\varepsilon)$-separado maximal, então é necessariamente $(n, 2\varepsilon)$-gerador de $M$, portanto, para todo $a \in M$, existe $x \in E_n$ tal que $d(f^j(a), f^j(x)) \leq 2\varepsilon$ para todo $0 \leq j < n$, ou seja, $M \subset \bigcup_{x \in E_n} B(x, n, 2\varepsilon)$.

Escolha conjuntos borelianos $P_x$ tal que $B(x,n,\varepsilon) \subset P_x \subset B(x,n,2\varepsilon)$ e $\mathscr{P}_n = \{P_x : x \in E_n\}$ é uma partição de $M$.

Agora diam $f^{-\lfloor \frac{n}{2} \rfloor}(\mathscr{P}_n) \to 0$ quando $n \to \infty$ por expansividade. Nessas condições, existem conjuntos $C_n$, cada um formado pela união de elementos de $\mathscr{P}_n$ de modo que $(\mu+\nu)(f^{-\lfloor \frac{n}{2} \rfloor}(C_n)\Delta B) \to 0$ quando $n \to \infty$ (veja o Teorema 2.4). Como $B$ é $f$-invariante, $(\mu+\nu)(C_n \Delta B) \to 0$.

Como $2\varepsilon$ é uma constante de expansividade para $f$, os iterados de $\mathscr{P}_n$ por $f^n$ geram a $\sigma$-álgebra de Borel e assim, pelo Corolário 3.13, $h_\nu(f^n) = h_\nu(f^n, \mathscr{P}_n) \leq H_\nu(\mathscr{P}_n)$. Assim $h_\nu(f) = n^{-1} h_\nu(f^n) \leq n^{-1} H_\nu(\mathscr{P}_n)$ e

$$\frac{1}{n}\int \varphi_n d\nu = \frac{1}{n}\int \sum_{j=0}^{n-1} \varphi \circ f^j d\nu = \frac{1}{n}\sum_{j=0}^{n-1}\int \varphi \circ f^j d\nu = \int \varphi d\nu$$

porque $\nu \in \mathcal{M}_1(f)$. Agora, vamos supor por contradição que $P = P(f,\varphi) = P_\nu(f,\varphi)$.

Então

$$P = h_\nu(f) + \int \varphi d\nu \leq \frac{1}{n}[H_\nu(\mathscr{P}_n) + \int \varphi_n d\nu]$$

e assim

$$nP \leq \sum_{P_x \in \mathscr{P}_n} [-\nu(P_x)\log\nu(P_x) + \int_{P_x} \varphi_n d\nu].$$

Agora, $\varphi_n \leq L + \varphi_n(x)$ em $P_x$, porque $\varphi \in V(f)$ e $x \in E_n$. Assim

$$\int_{P_x} \varphi_n d\nu \leq L \int_{P_x} d\nu + \int_{P_x} \varphi_n(x) d\nu = L\nu(P_x) + \varphi_n(x)\nu(P_x)$$

e

$$\begin{aligned} nP &\leq \sum_{P_x \in \mathscr{P}_n} [-\nu(P_x)\log\nu(P_x) + L\nu(P_x) + \varphi_n(x)\nu(P_x)] \\ &= L + \sum_{P_x \subset C_n} \nu(P_x)[\varphi_n(x) - \log\nu(P_x)] + \sum_{P_x \cap C_n = \emptyset} \nu(P_x)[\varphi_n(x) - \log\nu(P_x)]. \end{aligned}$$

Aplicando o Lema 5.26 a cada soma, obtemos

$$\begin{aligned} nP - L &\leq \nu(C_n)\log\left(\sum_{P_x \subset C_n} e^{\varphi_n(x)}\right) + \nu(C_n^c)\log\left(\sum_{P_x \cap C_n = \emptyset} e^{\varphi_n(x)}\right) \\ &\qquad - \nu(C_n)\log\nu(C_n) - \nu(C_n^c)\log\nu(C_n^c) \\ &\leq \nu(C_n)\log\left(\sum_{P_x \subset C_n} e^{\varphi_n(x)}\right) + \nu(C_n^c)\log\left(\sum_{P_x \cap C_n = \emptyset} e^{\varphi_n(x)}\right) + 2K^* \end{aligned}$$

onde $K^* = \max_{t \in [0,1]}(-t\log t)$. Rearranjando os termos,

$$\begin{aligned} -2K^* - L &\leq [\nu(C_n) + \nu(C_n^c)](-nP) + \nu(C_n)\log\left(\sum_{P_x \subset C_n} e^{\varphi_n(x)}\right) \\ &\qquad + \nu(C_n^c)\log\left(\sum_{P_x \cap C_n = \emptyset} e^{\varphi_n(x)}\right) \\ &\leq \nu(C_n)\log\left(\sum_{P_x \subset C_n} e^{\varphi_n(x) - nP}\right) + \nu(C_n^c)\log\left(\sum_{P_x \cap C_n = \emptyset} e^{\varphi_n(x) - nP}\right) \\ &\leq \nu(C_n)\log\left(\sum_{P_x \subset C_n} A_\varepsilon^{-1}\mu(B(x,n,\varepsilon))\right) + \nu(C_n^c)\log\left(\sum_{P_x \cap C_n = \emptyset} A_\varepsilon^{-1}\mu(B(x,n,\varepsilon))\right) \\ &\leq \nu(C_n)\log\left(A_\varepsilon^{-1}\sum_{P_x \subset C_n} \mu(P_x)\right) + \nu(C_n^c)\log\left(A_\varepsilon^{-1}\sum_{P_x \cap C_n = \emptyset} \mu(P_x)\right) \\ &\leq \nu(C_n)\log\left(A_\varepsilon^{-1}\mu(C_n)\right) + \nu(C_n^c)\log\left(A_\varepsilon^{-1}\mu(C_n^c)\right). \end{aligned}$$

(usamos o Lema 6.22). Quando $n \to \infty$, $\nu(C_n) \to 1$ e $\mu(C_n) \to 0$, consequentemente, $\nu(C_n) \log\left(A_\varepsilon^{-1}\mu(C_n)\right) \to -\infty$ e $\nu(C_n^c) \log\left(A_\varepsilon^{-1}\mu(C_n^c)\right) \to 0$. Isto contraria a desigualdade acima. Portanto deve ser $P_\nu(f, \varphi) < P(f, \varphi)$.

Em geral, se $\nu' \in \mathcal{M}_1(f)$, então $\nu' = \alpha\nu + (1-\alpha)\mu'$, onde $\nu \in \mathcal{M}_1(f)$ é singular com respeito a $\mu$ e $\mu' \in \mathcal{M}_1(f)$ é absolutamente contínua com respeito a $\mu$. Como $\mu'$ e $\nu$ estão suportadas em conjuntos disjuntos,

$$P_{\nu'}(f, \varphi) = \alpha P_\nu(f, \varphi) + (1-\alpha)P_{\mu'}(f, \varphi). \tag{6.21}$$

Suponha agora que $P_{\nu'}(f, \varphi) = P = P(f, \varphi)$. Nossa intenção é provar que $\nu' = \mu$. Ora, uma vez que $\nu \perp \mu$, vale que $P_\nu(f, \varphi) < P$. Mas, pelo princípio variacional, $P_\nu(f, \varphi) \leq P$ e $P_{\mu'}(f, \varphi) \leq P$ e portanto, pela equação (6.21)

$$P = P_{\nu'}(f, \varphi) = \alpha P_\nu(f, \varphi) + (1-\alpha)P_{\mu'}(f, \varphi) \leq P.$$

Com isso, concluímos que ou $P_\nu(f, \varphi) = P$ ou $\alpha = 0$. Mas não podemos ter $P_\nu(f, \varphi) = P$, logo $\alpha = 0$ e temos $\nu' = \mu'$ que é absolutamente contínua com respeito a $\mu$. Como $\nu'$ é invariante e $\mu$ é invariante e ergódica, por ser misturadora, segue que $\nu' = \mu$ em virtude do Lema 2.7. $\square$

Do Lema anterior segue, em particular, que $\mu$ é um estado de equilíbrio, pois o homeomorfismo com especificação admite algum estado de equilíbrio (por expansividade) e pelo Lema anterior este deve coincidir com $\mu$ e além disso este é único. A prova do Teorema 6.6 está completa. $\square$

# 7 UM EXEMPLO DE NÃO UNICIDADE DO ESTADO DE EQUILÍBRIO

Nos capítulos 3, 4 e 5 foram exigidos algumas hipóteses sobre a regularidade do potencial ($\varphi$ Hölder ou $\varphi \in V(f)$) para garantir a unicidade do estado de equilíbrio. É natural questionar se existem potenciais apenas contínuos para os quais ocorrem ou não a unicidade. Essa pergunta for respondida positivamente por Hofbauer [4].

Neste capítulo, vamos descrever o exemplo dado por Hofbauer [4] da não unicidade do estado de equilíbrio e alguns teoremas do artigo [4] sobre a condição de Ruelle-Perron-Frobenius e medidas homogêneas, estes últimos serão citados sem a demonstração, uma vez que nosso foco é descrever o exemplo de não unicidade do estado de equilíbrio.

## 7.1 Descrição da família de potenciais

Para estudarmos um exemplo onde não há unicidade no estado de equilíbrio, trabalharemos com potenciais especiais, os quais definiremos agora.

**Definição 7.1.** Um cilindro é um conjunto do tipo

$$[m; x_m \cdots x_n] = \{(y_i)_{i\in\mathbb{N}} \in \Sigma_A^+ : x_i = y_i \ \text{ para } \ m \leq i \leq n\}.$$

Defina $M_0 = \Sigma_n^+ \setminus [0;1] = \Sigma_n^+ \setminus \{(x_i)_{i\in\mathbb{N}} \in \Sigma_A^+ : x_0 = 1\}$, ou seja, $M_0$ é o conjunto de todas as sequências que não começam em 1. Para $k > 0$, defina

$M_k = \{(x_i)_{i\in\mathbb{N}}\Sigma_A^+ : x_i = 1 \text{ para } 0 \le i \le k-1 \text{ e } x_k \neq 1\}$. Então os conjuntos $M_k$, junto com a sequência $\{11\cdots\}$ formam uma partição de $\Sigma_n^+$. Seja $(a_n)$ uma sequência de números reais com $\lim a_n = 0$. Ponha $s_k = a_0 + \cdots + a_k$. Defina $g \in C(\Sigma_n^+)$ por

$$g((x_n)_n) = a_k \text{ para } (x_n)_n \in M_k \text{ e } g(\{11\cdots\}) = 0 \tag{7.1}$$

Enfatizamos que a função $g$ é contínua (pois é constante em cilindros $M_k$ e se $(x_n)_n \to \{11\cdots\}$, então $g((x_n)_n) = a_k \to 0 = g(\{11\cdots\})$ mas pode não ser Hölder. De fato, considere a sequência

$$a_k = -3\log\left(\frac{k+2}{k+1}\right) \quad \text{para } k \ge 1.$$

Note que para todo $\theta > 0$,

$$\frac{|g((x_n)_n) - g(\{11\cdots\})|}{d((x_n)_n, \{11\cdots\})^\theta} = \frac{3\log\left(\frac{k+2}{k+1}\right)}{\frac{1}{2^{\theta k}}} \to +\infty \text{ quando } k \to \infty \tag{7.2}$$

quando $(x_n)_n \in M_k$. (O denominador na fórmula (7.2) vem da métrica que usamos no exemplo 4.2 e a conclusão que este limite vai a infinito resulta da aplicação da regra de L' Hospital duas vezes).

## 7.2 A condição de Ruelle-Perron-Frobenius e medidas homogêneas

Denotaremos por $\Sigma_A^+$ o conjunto das sequências $(x_n)_n = \{x_i\}_{i=0}^\infty \in \{1,2,\cdots,n\}^{\mathbb{N}} = \Sigma_n^+$ que são A-admissíveis e $\sigma_A^+ : \Sigma_A^+ \to \Sigma_A^+$ o deslocamento unilateral de tipo finito associado a $A$. Denotaremos $\mathcal{M}(\Sigma_A^+)$ como o conjunto das medidas de probabilidade em $\Sigma_A^+$ e $\mathcal{M}_{\sigma_A^+}(\Sigma_A^+) \subset \mathcal{M}(\Sigma_A^+)$ como o conjunto das medidas de probabilidade em $\Sigma_A^+$ que são $\sigma_A^+$-invariantes. Para cada $\varphi \in \mathcal{C}(\Sigma_A^+)$, onde $\mathcal{C}(\Sigma_A^+)$ denota o espaço de todas as funções contínuas $\phi : \Sigma_A^+ \to \mathbb{R}$ com a norma do supremo $||\cdot||$, definimos o operador de transferência $\mathcal{L} = \mathcal{L}_\varphi : \mathcal{C}(\Sigma_A^+) \to \mathcal{C}(\Sigma_A^+)$ por

$$\mathcal{L}g((y_n)_n) = \mathcal{L}_\phi g((y_n)_n) = \sum_{(x_n)_n \in (\sigma_A^+)^{-1}((y_n)_n)} e^{\phi((x_n)_n)} g((x_n)_n). \tag{7.3}$$

**Definição 7.2.** Dizemos que $\varphi$ satisfaz a **condição de Ruelle-Perron-Frobenius** se existe $\lambda > 0$, $h \in \mathcal{C}(\Sigma_A^+)$ com $h > 0$ e $\nu \in \mathcal{M}(\Sigma_A^+)$ satisfazendo $\mathcal{L}h = \lambda h$, $\mathcal{L}^*\nu = \lambda\nu$, $\int hd\nu = 1$ e

$$\lim_{m\to\infty} ||\lambda^{-m}\mathcal{L}^m g - h\int gd\nu|| = 0 \quad \text{para toda} \quad g \in \mathcal{C}(\Sigma_A^+).$$

A medida $\mu$ definida por $\mu = \int hgd\nu$ será chamada de **medida de Ruelle-Perron-Frobenius**

Provamos na seção 4.2 que se $\varphi \in \mathcal{F}_A \cap \mathcal{C}(\Sigma_A^+)$, então satisfaz a condição de Ruelle-Perron-Frobenius e $\mu = h\nu$ é o único estado de equilíbrio para $\varphi$. Isto é generalizado com o seguinte Teorema.

**Teorema 7.3 (Hofbauer).** *Seja $\varphi$ satisfazendo a condição de Ruelle-Perron-Frobenius. Então $\mu = h\nu$ é o único estado de equilíbrio para $\varphi$.*

*Demonstração.* Veja Hofbauer [4]. □

**Teorema 7.4.** *Se $\sum_{k=0}^{\infty} e^{s_k} > \dfrac{1}{(n-1)}$, então $g$ satisfaz a condição de Ruelle-Perron-Frobenius.*

*Demonstração.* Veja Hofbauer [4]. □

Observe que, se o potencial $g$ em (7.1) está definido em $C(\Sigma_2^+)$, então a condição do Teorema 7.4 é $\sum_{k=0}^{\infty} e^{s_k} > \dfrac{1}{(2-1)} = 1$.

**Definição 7.5.** Dizemos que $g$ admite uma **medida homogênea** $\mu \in \mathcal{M}_{\sigma_A^+}(\Sigma_n^+)$ se existem $c_1$, $c_2 > 0$ e $\lambda > 0$ tal que

$$c_1 \leq \frac{\mu([0; x_0 \cdots x_{m-1}])}{\lambda^{-m} e^{\sum_{i=0}^{m-1} g((\sigma_A^+)^i((x_n)_n))}} = \frac{\mu([0; x_0 \cdots x_{m-1}])}{\lambda^{-m} e^{g_m((x_n)_n))}} = Q \leq c_2. \tag{7.4}$$

para cada $(x_n)_n \in \Sigma_n^+$ para cada $m \geq 0$.

Repare que essa definição é equivalente à definição de estado de Gibbs.

**Teorema 7.6.** *A função $g$ definida em (7.1) tem uma medida homogênea se, e somente se $\sum_{k=0}^{\infty} a_k$ é convergente.*

*Demonstração.* Veja Hofbauer [4]. □

## 7.3 $\delta_{\{111\cdots\}}$ é um estado de equilíbrio para $g$

Mostraremos agora que as funções $g$, que definimos em (7.1) dão exemplos de não unicidade do estado de equilíbrio. Isto de certa forma poderia ser esperado, uma vez que ela não é Hölder. Ao longo desta seção, consideraremos $g \in C(\Sigma_2^+) = C(\{1,2\}^{\mathbb{N}})$.

Veremos que sob a condição $\sum_{k=0}^{\infty} e^{s_k} \leq 1$, a medida $\delta_{\{111\cdots\}}$ é um estado de equilíbrio para $g$.

**Lema 7.7.** *Se $\sum_{k=0}^{\infty} e^{s_k} \leq 1$, então a pressão $P(\sigma_A^+, g) \leq 0$.*

*Demonstração.* Veja Hofbauer [4]. □

**Teorema 7.8.** *Se $\sum_{k=0}^{\infty} e^{s_k} \leq 1$, então $\delta_{\{111\cdots\}}$ é um estado de equilíbrio para $g$. (aqui, $\delta_{\{111\cdots\}}$ é a medida de Dirac suportada na sequência composta toda de $1's$.)*

*Demonstração.* Primeiro note que $\int g d\delta_{\{111\cdots\}} = g(\{111\cdots\}) = 0$ e que a entropia métrica $h_{\delta_{\{111\cdots\}}}(\sigma_A^+)$ é nula: de fato a medida $\delta_{\{111\cdots\}}$ só assume os valores 0 ou 1. Temos que $\mathscr{P} = \{[0;1],[0;2]\}$ é uma partição geradora. Então a partição $\mathscr{P}^n$ é formada pelos cilindros $[0, x_0 \cdots x_n]$ com $x_i = 1, 2$ e $0 \leq i \leq n$. Daí

$$\begin{aligned} h_{\delta_{\{111\cdots\}}}(\sigma_A^+) &= h_{\delta_{\{111\cdots\}}}(\sigma_A^+, \mathscr{P}) \\ &= \lim_n \frac{1}{n} H_{\delta_{\{111\cdots\}}}(\mathscr{P}^n) \\ &= \lim_n \frac{1}{n}\left( \sum_{[x_0\cdots x_n]\in\mathscr{P}^n} -\delta_{\{111\cdots\}}([x_0\cdots x_n]) \log \delta_{\{111\cdots\}}([x_0\cdots x_n]) \right). \end{aligned}$$

Mas o termo dentro do somatório da última equação é sempre zero, (as possibilidades são: $1\log(1)$ e $\lim_{x\to 0} -x\log(x)$). Portanto $h_{\delta_{\{111\cdots\}}}(\sigma_A^+) = 0$ como afirmado.

Pelo Lema 7.7 temos que

$$P(\sigma_A^+, g) \leq 0 = h_{\delta_{\{111\cdots\}}}(\sigma_A^+) + \int g d\delta_{\{111\cdots\}}.$$

Por outro lado,

$$P(\sigma_A^+, g) \geq h_\mu(\sigma_A^+) + \int g d\mu \tag{7.5}$$

para toda $\mu \in \mathcal{M}_{\sigma_A^+}(\Sigma_2^+)$. Em particular $P(\sigma_A^+, g) \geq h_{\delta_{\{111\cdots\}}}(\sigma_A^+) + \int g d\delta_{\{111\cdots\}}$ e consequentemente $P(\sigma_A^+, g) = h_{\delta_{\{111\cdots\}}}(\sigma_A^+) + \int g d\delta_{\{111\cdots\}}$. □

## 7.4 Outro estado de equilíbrio para $g$

**Lema 7.9.** *Sejam $\varphi$, $(\varphi_n)_n \in C(\Sigma_2^+)$ com $\varphi_n \to \varphi$ quando $n \to \infty$ e seja $\mu_n$ um estado de equilíbrio para $\varphi_n$. Como $\mathcal{M}_{\sigma_A^+}(\Sigma_2^+)$ é um espaço métrico compacto com a topologia fraca*, existe uma subsequência convergente $(\mu_{n_i})_i$ de $\mu_n$. Então $\mu = \lim_{i\to\infty} \mu_{n_i}$ é um estado de equilíbrio para $\varphi$.*

*Demonstração.* Começamos provando a seguinte afirmação: A função pressão é Lipschitz, com constante de Lipschitz igual a 1:

$$|P(f,\phi) - P(f,\psi)| \leq ||\phi - \psi|| \tag{7.6}$$

para quaisquer potenciais $\phi$ e $\psi$. A demonstração desse fato é observar que $\phi \leq \psi + ||\phi - \psi||$. Logo, pelas equações (3.55) e (3.54) da observação 3.53, temos $P(f,\phi) \leq P(f,\psi) + ||\phi - \psi||$. Trocando os papéis de $\phi$ e $\psi$ obtemos a outra desigualdade. Aplicando isto ao nosso caso, temos

$$\begin{aligned} P(\sigma_A^+, \varphi) &= \lim_{n\to\infty} P(\sigma_A^+, \varphi_n) \\ &= \lim_{n\to\infty} \left( h_{\mu_n}(\sigma_A^+) + \int \varphi_n d\mu_n \right) \\ &\leq \lim_{n\to\infty} \left( h_{\mu_n}(\sigma_A^+) + \int \varphi d\mu_n \right) + \lim_{n\to\infty} ||\varphi_n - \varphi|| \qquad (7.7) \\ &= \lim_{i\to\infty} h_{\mu_{n_i}}(\sigma_A^+) + \lim_{i\to\infty} \int \varphi d\mu_{n_i} \\ &\leq h_\mu(\sigma_A^+) + \int \varphi d\mu. \end{aligned}$$

A primeira desigualdade de (7.7) vem de utilizar (7.6) para os potenciais $\varphi$ e $\varphi_n$, a igualdade que vem em seguida ocorre porque $\lim_{n\to\infty} ||\varphi_n - \varphi|| = 0$ e a última desigualdade ocorre porque $\lim_{i\to\infty} h_{\mu_{n_i}}(\sigma_A^+) \leq h_\mu(\sigma_A^+)$ uma vez que $\mu \mapsto h_\mu(\sigma_A^+)$ é semicontínua superiormente. (Veja o Teorema 3.20.)

Agora, como $P(\sigma_A^+, \varphi) \geq h_\mu(\sigma_A^+) + \int \varphi d\mu$ para toda $\mu \in \mathcal{M}_{\sigma_A^+}(\Sigma_2^+)$, segue que $P(\sigma_A^+, \varphi) = h_\mu(\sigma_A^+) + \int \varphi d\mu$ e isto finaliza a prova do Lema. □

**Lema 7.10.** *Seja $\mu_n \in \mathcal{M}_{\sigma_A^+}(\Sigma_2^+)$ e suponha que $\lim_{n\to\infty} \mu_n([0; x_0 \cdots x_{m-1}])$ existe para cada cilindro $[0; x_0 \cdots x_{m-1}]$. Então $\mu_n$ converge na topologia fraca* para $\mu \in \mathcal{M}_{\sigma_A^+}(\Sigma_2^+)$,* *a qual é unicamente determinada por*

$$\mu([0;x_0\cdots x_{m-1}])=\lim_{n\to\infty}\mu_n([0;x_0\cdots x_{m-1}])$$

O principal Teorema deste capítulo é o seguinte.

**Teorema 7.11** (**Exemplo de não unicidade do estado de equilíbrio**). *Suponha que $g$ é a função que definimos em* (7.1) *satisfazendo*

$$\sum_{k=0}^{\infty}e^{s_k}=1 \quad e \quad \sum_{k=0}^{\infty}(k+1)e^{s_k}=u^{-1}<\infty. \tag{7.8}$$

*Então $g$ tem dois estados de equilíbrio.*

*Demonstração.* Já vimos que $\delta_{\{111\cdots\}}$ é um estado de equilíbrio para $g$, logo nosso objetivo será descrever a construção de um estado de equilíbrio distinto para $g$.

Para um bloco dado $x_0\cdots x_{m-1}\neq 11\cdots 1$ escolha $t$ tal que $x_0=\cdots=x_{t-1}=1$ e $x_t=2$ e $r$ tal que $x_{r-1}=2$ e $x_r=\cdots=x_{m-1}=1$. Mostraremos que a medida $\mu$ definida por

$$\begin{aligned}\mu([0;x_0\cdots x_{m-1}])&=h^t e^{g_r(x_0\cdots x_{r-2}2\cdots)}\sum_{i=m-r}^{\infty}e^{s_i}\ \text{ para }\ [0;x_0\cdots x_{m-1}]\neq[0;11\cdots 1]\\ \mu([0;11\cdots 1])&=u\sum_{i=m}^{\infty}(i-m+1)e^{s_i},\end{aligned}\tag{7.9}$$

onde $h^t=ue^{-s_t}\sum_{i=t}^{\infty}e^{s_i}$, $u$ é definido em (7.8) e $e^{g_r(x_0\cdots x_{r-2}2\cdots)}=e^{\sum_{i=0}^{r-1}g((\sigma_A^+)^i(x_0\cdots x_{r-2}2\cdots))}$ é um estado de equilíbrio para $g$.

Começaremos observando que $\mu$ realmente define uma medida. De fato, como os cilindros geram a $\sigma$-álgebra de $\Sigma_2^+$ e como

$$[m;x_m\cdots x_n]=[m;x_m\cdots x_n,1]\cup[m;x_m\cdots x_n,2]$$

é uma união disjunta, é suficiente verificar que definição da medida de cada cilindro é $[m;x_m\cdots x_n]$ é compatível com as definições das medidas dos seus subcilindros, no sentido de que $\mu([m;x_m\cdots x_n])=\mu([m;x_m\cdots x_n,1])+\mu([m;x_m\cdots x_n,2])$. De forma ainda mais simples, verificaremos isso para os cilindros que têm termo inicial na posição 0. Isso é possível porque os cilindros $[m;x_m\cdots x_n]$ são uniões finitas de cilindros que têm termo inicial na posição 0. Comecemos com um cilindro $[0; x_0 \cdots x_{m-1}] \neq [0; 11 \cdots 1]$. Temos que $[0; x_0 \cdots x_{m-1}] = [0; x_0 \cdots x_{m-1}, 1] \cup [0; x_0 \cdots x_{m-1}, 2]$ e o $t$ tal que $x_t = 2$ e $x_0 = \cdots = x_{t-1} = 1$ para o cilindro do lado esquerdo é o mesmo para os dois cilindros do lado direito. Assim não $h^t$ é constante nesse caso. Então

$$\begin{aligned} \mu([0; x_0 \cdots x_{m-1}]) &= h^t e^{g_r(x_0 \cdots x_{r-2} 2 \cdots)} \sum_{i=m-r}^{\infty} e^{s_i} \\ \mu([0; x_0 \cdots x_{m-1}, 1]) &= h^t e^{g_r(x_0 \cdots x_{r-2} 2 \cdots)} \sum_{i=m+1-r}^{\infty} e^{s_i} \\ \mu([0; x_0 \cdots x_{m-1}, 2]) &= h^t e^{g_{m+1}(x_0 \cdots x_{m-1} 2 \cdots)} \sum_{i=0}^{\infty} e^{s_i} = h^t e^{g_{m+1}(x_0 \cdots x_{m-1} 2 \cdots)}. \end{aligned} \tag{7.10}$$

A única alteração na equação da segunda linha é que o somatório parte de $m+1-r$ uma vez que o cilindro aumentou em uma unidade seu comprimento. O $r$ nesse caso se manteve o mesmo. Por isso a expressão de $g_r$ é a mesma da primeira linha. Porém na segunda linha, a principal alteração é que $r' = m+1$ ($r'$ no lugar de $r$ na definição (7.9)), pois é na posição $m$ que o último dígito 2 do cilindro aparece. Daí a soma parte de 0, porque esse cilindro também tem comprimento $m+1$ e por (7.8) $\sum_{i=0}^{\infty} e^{s_i} = 1$. Agora, Notamos que

$$\begin{aligned} g_{m+1}(x_0 \cdots x_{m-1} 2 \cdots) &= g_r(x_0 \cdots x_{m-1} 2 \cdots) + g_{(m+1)-r}((\sigma_A^+)^r (x_0 \cdots x_{m-1} 2 \cdots)) \\ &= g_r(x_0 \cdots x_{r-2} 2 \cdots) + g_{(m+1)-r}(111 \cdots 2 \cdots) \\ &= g_r(x_0 \cdots x_{r-2} 2 \cdots) + a_0 + a_1 + \cdots + a_m \\ &= g_r(x_0 \cdots x_{r-2} 2 \cdots) + s_m. \end{aligned} \tag{7.11}$$

Chamamos atenção aos seguintes detalhes: $(\sigma_A^+)^r (x_0 \cdots x_{m-1} 2 \cdots) = 111 \cdots 2 \cdots$ porque $x_r = \cdots x_{m-1} = 1$ e $g_{m-r}(111 \cdots 2 \cdots) = a_0 + \cdots + a_m$ porque $g((\sigma_A^+)^j (111 \cdots 2 \cdots)) = g((y_n)_n)$ com $(y_n)_n \in M_{m-j}$ para $0 \leq j \leq m$. Então juntando (7.11) à terceira equação de (7.10), temos

$$\mu([0; x_0 \cdots x_{m-1}, 2]) = h^t e^{g_r(x_0 \cdots x_{r-2} 2 \cdots)} e^{s_m}.$$

Agora é só observar que somando isto a segunda equação de (7.10), obtemos a primeira equação de (7.10).

Analogamente, escreva $[0; 1 \cdots 1] = [0; 1 \cdots 1, 1] \cup [0; 1 \cdots 1, 2]$, onde o cilindro do lado esquerdo tem comprimento $m$ e os do lado direito comprimento $m+1$. Queremos mostrar que nesse caso $\mu([0;11\cdots 1]) = \mu([0;11\cdots 11]) + \mu([0;11\cdots 12])$. Note que

$$\begin{aligned}
\mu([0;11\cdots 1]) &= u\sum_{i=m}^{\infty}(i-m+1)e^{s_i},\\
\mu([0;11\cdots 11]) &= u\sum_{i=m+1}^{\infty}(i-m)e^{s_i} = u\sum_{i=m}^{\infty}(i-m)e^{s_i},\\
\mu([0;11\cdots 12]) &= ue^{s_m}\sum_{i=m}^{\infty}e^{s_i}e^{g_{m+1}(1\cdots 12\cdots)}\sum_{i=0}^{\infty}e^{s_i} = u\sum_{i=m}^{\infty}e^{s_i},
\end{aligned}$$

porque nesse caso, $t = m$ e $r-1 = m$. Isto mostra que $\mu([0;11\cdots 1]) = \mu([0;11\cdots 11]) + \mu([0;11\cdots 12])$ como desejávamos. Logo a medida é aditiva.

Para usar o Lema 7.9, vamos verificar que $\mu$ é um limite de estados de equilíbrio. Considere $g_n$ definida por:

$$\begin{aligned}
&g_n((x_k)_k) = g((x_k)_k) \;\text{ para }\; (x_k)_k \notin M_0,\\
&g_n((x_k)_k) = a_0^{(n)} \;\text{ para }\; (x_k)_k \in M_0,\;\text{ onde }\; a_0^{(n)} \to a_0,\; a_0^{(n)} > a_0 \;\text{ e }\; a_0^{(n)} \geq a_0^{(n+1)}\;\; \forall\, n.
\end{aligned} \tag{7.12}$$

Então $\sum_{k=0}^{\infty} e^{s_k^{(n)}} > 1$, onde $s_k^{(n)} = a_0^{(n)} + a_1 + \cdots + a_k$. Consequentemente, $g_n$ satisfaz a condição de Ruelle-Perron-Frobenius pelo Teorema 7.4. Pelo Teorema 7.3 $\mu_n = \nu_n h_n$ é um estado de equilíbrio para $g_n$, onde $\nu_n$ e $h_n$ são como na condição de Ruelle-Perron-Frobenius para $g_n$. O leitor poderá observar em [4] que a autofunção $h_n$ construída no Teorema 7.4 é da seguinte forma:

$$\begin{aligned}
&h_n((x_k)_k) = u^{(n)}\nu(M_t)^{-1}\sum_{i=t}^{\infty}\nu(M_i) := h_n^t \;\text{ para }\; (x_k)_k \in M_t\,,\\
&h_n(\{11\cdots\}) = u^{(n)}\sum_{i=0}^{\infty}\lambda_n^{-i}\,,\\
&u^{(n)} = \frac{1}{\sum_{i=0}^{\infty}(i+1)\nu_n(M_i)}.
\end{aligned}$$

Note que $h_n$ é constante em $[0; x_0 \cdots x_{m-1}] \neq [0; 11 \cdots 1]$. Assim,

$$
\begin{aligned}
\mu_n([0; x_0 \cdots x_{m-1}]) &= \int h_n(\underline{z}) \chi_{[0;x_0 \cdots x_{m-1}]}(\underline{z}) d\nu_n \\
&= h_n^t \int \chi_{[0;x_0 \cdots x_{m-1}]}(\underline{z}) d\nu_n \\
&= h_n^t \lambda_n^{-r} \int \mathcal{L}^r(\chi_{[0;x_0 \cdots x_{m-1}]})(\underline{z}) d\nu_n \\
&= h_n^t \lambda_n^{-r} \int \sum_{y_0 \cdots y_{r-1}} e^{(g_n)_r(y_0 \cdots y_{r-1}\underline{z})} \chi_{[0;x_0 \cdots x_{m-1}]}(y_0 \cdots y_{r-1}\underline{z}) d\nu_n \\
&= h_n^t \lambda_n^{-r} e^{(g_n)_r(x_0 \cdots x_{r-2} 2 \cdots)} \int \chi_{[0;x_r \cdots x_{m-1}]}(\underline{z}) d\nu_n \\
&= h_n^t \lambda_n^{-r} e^{(g_n)_r(x_0 \cdots x_{r-2} 2 \cdots)} \int \chi_{[0;1 \cdots 1]}(\underline{z}) d\nu_n \\
&= h_n^t \lambda_n^{-r} e^{(g_n)_r(x_0 \cdots x_{r-2} 2 \cdots)} \sum_{i=m-r}^{\infty} \nu_n(M_i)
\end{aligned}
\tag{7.13}
$$

Aqui $\underline{z}$ é uma sequência em $\Sigma_2^+$, $r$ tal que $x_{r-1} = 2$ e $x_r = \cdots = x_{m-1} = 1$. Além disso

$$
\mu_n([0; 11 \cdots 1]) = u^{(n)} \sum_{i=m}^{\infty} (i - m + 1) \nu_n(M_i).
$$

De fato, notando que $\nu_n(\{111 \cdots\}) = 0$,

$$
\begin{aligned}
\mu_n([0; 11 \cdots 1]) &= \mu_n([0; 11 \cdots 12]) + \mu_n([0; 11 \cdots 112]) + \cdots \\
&= \sum_{r=m}^{\infty} h_n^r \lambda_n^{-r-1} e^{(g_n)_{r+1}(x_0 \cdots x_{r-2} 2 \cdots)} \sum_{i=0}^{\infty} \nu_n(M_i) \\
&= \sum_{r=m}^{\infty} u^{(n)} \nu_n(M_r)^{-1} \sum_{i=r}^{\infty} \nu_n(M_i) \lambda_n^{-r-1} e^{a_r + a_{r-1} + \cdots + a_0^{(n)}}.
\end{aligned}
\tag{7.14}
$$

Note que em geral, são válidas as seguintes igualdades para $\nu \in \mathcal{M}(\Sigma_n^+)$:

$$
\begin{aligned}
\lambda \nu(M_0) &= \int \chi_{M_0} d(\lambda \nu) = \int \chi_{M_0} d(\mathcal{L}^* \nu) \\
&= \int \mathcal{L} \chi_{M_0} d\nu \\
&= \int \sum_{i=1}^{n} e^{g(i\underline{z})} \chi_{M_0}(i\underline{z}) d\nu = (n-1) e^{a_0}
\end{aligned}
\tag{7.15}
$$

e

$$\begin{aligned}
\lambda\nu(M_k) &= \int \chi_{M_k} d(\lambda\nu) = \int \chi_{M_k} d(\mathcal{L}^*\nu) \\
&= \int \mathcal{L}\chi_{M_k} d\nu \\
&= \int \sum_{i=1}^{n} e^{g(i\underline{z})}\chi_{M_k}(i\underline{z}) d\nu \\
&= \int \sum_{i=1}^{n} e^{g(1\underline{z})}\chi_{M_{k-1}}(\underline{z}) d\nu = e^{a_k}\int \chi_{M_{k-1}}(\underline{z}) d\nu = e^{a_k}\nu(M_{k-1})
\end{aligned} \tag{7.16}$$

onde $\underline{z}$ é uma sequência em $\Sigma_n^+$. Fazendo indução na relação (7.16) para $\nu_n$ e juntando com (7.15) para $n = 2$, temos que

$$\nu_n(M_r) = \lambda_n^{-r-1} e^{a_r + a_{r-1} + \cdots + a_0^{(n)}}.$$

Substituindo isso em (7.14), temos

$$\begin{aligned}
\mu_n([0; 11\cdots 1]) &= \sum_{r=m}^{\infty} u^{(n)}\nu_n(M_r)^{-1}\nu_n(M_r)\sum_{i=r}^{\infty}\nu_n(M_i) \\
&= u^{(n)}\sum_{r=m}^{\infty}\sum_{i=r}^{\infty}\nu_n(M_i) \\
&= u^{(n)}\sum_{i=m}^{\infty}(i-m+1)\nu_n(M_i)
\end{aligned}$$

como queríamos provar.

Pelo Lema 7.10 é suficiente mostrar que $\mu_n([0; x_0 \cdots x_{m-1}]) \to \mu([0; x_0 \cdots x_{m-1}])$. Faremos isso abaixo.

Defina

$$\begin{aligned}
f(x) &= \sum_{i=0}^{\infty} e^{s_i} x^{i+1} \quad \text{e} \\
f_n(x) &= \sum_{i=0}^{\infty} e^{s_i^{(n)}} x^{i+1} \\
&= \sum_{i=0}^{\infty} e^{a_0^{(n)} + a_1 + \cdots + a_i} x^{i+1} \\
&= e^{a_0^{(n)} - a_0}\sum_{i=0}^{\infty} e^{a_0 + a_1 + \cdots + a_i} x^{i+1} = e^{a_0^{(n)} - a_0} f(x).
\end{aligned}$$

Então $f$ e $f_n$ são estritamente crescentes em $[0,1]$, $f(1) = \sum_{i=0}^{\infty} e^{s_i} = 1$ e $\lambda_n$ é unicamente determinado por $f_n(\lambda_n^{-1}) = 1$.

Temos que $f_n \geq f_{n+1}$ porque $a_0^{(n)} \geq a_0^{(n+1)}$. Isto implica que $\lambda_n \geq \lambda_{n+1} \geq 1$. Fixe $c < 1$. Então existe $N$ com $f_n(c) = e^{a_0^{(n)} - a_0} f(c) < 1$ para cada $n \geq N$ (como $f(c) < 1$, escolha $N$ tal que $|a_0^{(n)} - a_0| < \log\left(\frac{1}{f(c)}\right)$ para cada $n \geq N$). Portanto, uma vez que $f_n$ é crescente, $\lambda_n \leq c^{-1}$ para cada $n \leq N$. Isto significa que $\lambda_n \to 1$ quando $n \to \infty$.

Note que

$$\lambda_n \nu_n(M_0) = \int \mathcal{L}(\chi_{M_0}) d\nu_n = \int \sum_{i=1}^{2} e^{g_n(i\underline{z})} \chi_{M_0}(i\underline{z}) d\nu_n = e^{g_n(2\underline{z})} \nu_n(\Sigma_2^+) = e^{a_0^{(n)}}. \tag{7.17}$$

e

$$\begin{aligned} \lambda_n \nu_n(M_k) = \int \mathcal{L}(\chi_{M_k}) d\nu_n &= \int \sum_{i=1}^{2} e^{g_n(i\underline{z})} \chi_{M_k}(i\underline{z}) d\nu_n \\ &= e^{g(1\underline{z})} \int \chi_{M_{k-1}}(\underline{z}) d\nu_n = e^{a_k} \nu_n(M_{k-1}). \end{aligned} \tag{7.18}$$

Juntando as equações (7.17) e (7.18), obtemos

$$\nu_n(M_i) = \lambda_n^{-i-1} e^{s_i^{(n)}} = \lambda_n^{-i-1} e^{a_0^{(n)} - a_0} e^{s_i} \to e^{s_i} \quad \text{quando } n \to \infty. \tag{7.19}$$

Assim,

$$\sum_{i=t}^{\infty} \nu_n(M_i) = e^{a_0^{(n)} - a_0} \sum_{i=t}^{\infty} \lambda_n^{-i-1} e^{s_i} \to \sum_{i=t}^{\infty} e^{s_i} \quad \text{quando } n \to \infty. \tag{7.20}$$

pelo Teorema de Abel para série de potências.

Usando (7.19) e o Teorema de Abel para série de potências

$$\begin{aligned} \sum_{i=m}^{\infty} (i-m+1) \nu_n(M_i) &= e^{a_0^{(n)} - a_0} \sum_{i=t}^{\infty} (i-m+1) \lambda_n^{-i-1} e^{s_i} \\ &\to \sum_{i=m}^{\infty} (i-m+1) e^{s_i} \quad \text{quando } n \to \infty. \end{aligned}$$

Em particular, quando $m = 0$,

$$\frac{1}{u^{(n)}} = \sum_{i=0}^{\infty} (i+1) \nu_n(M_i) \to \sum_{i=0}^{\infty} (i+1) e^{s_i} = \frac{1}{u} \quad \text{quando } n \to \infty.$$

(a última igualdade por 7.8). Consequentemente $u^{(n)} \to u$, quando $n \to \infty$. Segue disso, junto com (7.19) e (7.20) que

$$h_n^t = u^{(n)}(\nu_n(M_t))^{-1} \sum_{i=t}^{\infty} \nu_n(M_i) \to ue^{-s_t} \sum_{i=t}^{\infty} e^{s_i} = h^t \quad \text{quando } n \to \infty.$$

Por fim, note que

$$\begin{aligned}(g_n)_r(x_0 \cdots x_{r-2}2 \cdots) &= K[a_0^{(n)} - a_0] + g_r(x_0 \cdots x_{r-2}2 \cdots) \\ &\to g_r(x_0 \cdots x_{r-2}2 \cdots) \quad \text{quando } n \to \infty\end{aligned}$$

onde $K$ é a quantidade de dígitos 2 que aparece no bloco $x_0 \cdots x_{r-2}2 \cdots$ .

Assim fica provado que $\mu_n([0; x_0 \cdots x_{m-1}]) \to \mu([0; x_0 \cdots x_{m-1}])$. Pelo Lema 7.10 $\mu_n \to \mu$ na topologia fraca*. Pelo Lema 7.9 $\mu$ é um estado de equilíbrio para $g$. que satisfaz as condições da equação (7.8). Mas para essa mesma $g$ com $\sum_{k=0}^{\infty} e^{s_k} = 1$, $\delta_{\{111\cdots\}}$ é um estado de equilíbrio conforme o Teorema 7.8 e $\mu \neq \delta_{\{111\cdots\}}$, porque $0 = \mu(\{111\cdots\}) \neq \delta_{\{111\cdots\}}(\{111\cdots\}) = 1$. Consequentemente o Teorema está provado. $\square$

## 7.5 Um exemplo explícito

Por fim, explicitamos um potencial $g$ que atende as condições da equação (7.8).

Precisamos exibir uma sequência $(a_k)_k$ de números reais satisfazendo

(a) $(a_k)_k \to 0$;

(b) $\sum_{k=0}^{\infty} e^{s_k} = 1$;

(c) $\sum_{k=0}^{\infty}(k+1)e^{s_k} < \infty$.

Escolha

$$a_k = -3\log\left(\frac{k+2}{k+1}\right) \quad \text{para } k \geq 1$$

e

$$a_0 = -\log\left(1 + \sum_{k=1}^{\infty} e^{a_1+\cdots+a_k}\right).$$

De fato:

(a) Segue de $\log\left(\dfrac{k+2}{k+1}\right) \to 0$;

(b) $\sum_{k=0}^{\infty} e^{s_k} = e^{a_0} + e^{a_0}\sum_{k=1}^{\infty} e^{a_1+\cdots+a_k} = e^{a_0}\left(1 + \sum_{k=1}^{\infty} e^{a_1+\cdots+a_k}\right) = e^{a_0}e^{-a_0} = 1$;

(c) $\sum_{k=0}^{\infty}(k+1)e^{s_k} \leq e^{a_0} + \sum_{k=1}^{\infty}(k+1)(k+2)^{-3} < \infty$.

# REFERÊNCIAS


[1] Baggett, L.W. **Functional Analysis: A Primer**, Taylor & Francis, 1991.

[2] Deimling, Klaus. **Nonlinear Functional Analysis**, Springer-Verlag, 1943.

[3] Downarowicz, T. **Entropy in Dynamical Systems**, New Mathematical Monographs, Cambridge University Press, 2011.

[4] F. Hofbauer. **Examples for the nonuniqueness of the equilibrium states**, Transactions of the AMS, Vol. 228, 223-241, 1977.

[5] Katok, A. and Hasselblatt, B.. **Introduction to the Modern Theory of Dynamical Systems**, Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1997.

[6] M. Hirsch and C. Pugh: **Stable manifolds and hyperbolic sets**, Proc. Sympos. Pure Math., Vol. XIV, Berkeley, Calif., 1968, pp. 133-163, Amer. Math. Soc.,Providence, R.I., 1970.

[7] Oliveira, Krerley, Viana, Marcelo. **Fundamentos da Teoria Ergódica**, SBM, 2014.

[8] P. Walters. **An Introduction to Ergodic Theory**,Springer-Verlag, New york, 1982.

[9] R. Bowen. **Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms**, Lecture Notes in Mathematics, 470, 1975.

[10] R. Bowen. **Some systems with unique equilibrium states**, Math. Systems Theory 8 (1974/75), no.3, 193-202.

[11] Shub, Michael. **Global Stability of Dynamical Systems**, Springer-Verlag, 1987.