Universidade Federal de Santa Catarina
Curso de Pós-Graduação em Matemática e Computação Científica

Operador de Ruelle-Perron-Frobenius e Transformações Expansoras

Anderson Luiz Maciel
Orientador: Prof. Dr. Aldrovando L. A. Araújo

Florianópolis
Março de 2005

Universidade Federal de Santa Catarina
Curso de Pós-Graduação em Matemática e Computação Científica

Operador de Ruelle-Perron-Frobenius e Transformações Expansoras

Dissertação apresentada ao Curso de Pós-Graduação em Matemática e Computação Científica, do Centro de Ciências Físicas e Matemáticas da Universidade Federal de Santa Catarina, para a obtenção do grau de Mestre em Matemática, com Área de Concentração em Geometria e Topologia.

Anderson Luiz Maciel
Florianópolis
Março de 2005

Operador de Ruelle-Perron-Frobenius e Transformações Expansoras

por

Anderson Luiz Maciel

Esta Dissertação foi julgada adequada para a obtenção do Título de "Mestre", Área de Concentração em Geometria e Topologia, e aprovada em sua forma final pelo Curso de Pós-Graduação em Matemática e Computação Científica.

Igor Mozolevski
Coordenador

Comissão Examinadora

Prof. Dr. Aldrovando L. Azeredo Araújo (UFSC-Orientador)

Prof. Dr. Artur O. Lopes (UFRGS)

Prof. Dr. Celso Melchiades Doria (UFSC)

Florianópolis, Março de 2005.

À Deus

À minha família

SUMÁRIO

Agradecimentos

Agradeço inicialmente aos meus familiares, em especial a minha namorada Divane e o meu irmão Cristiano que me aguentaram nas horas difíceis.

Os professores do departamento de matemática e os colegas da pós graduação com quem mantive contato durante esses dois anos, principalmente os professores Igor Mozolevski, Gustavo da Costa, Celso Doria, Eliezer Batista, José Pinho, Joel Souza entre outros. Ao professor Artur Lopes (UFRGS) que além de ter participado da banca da defesa, me auxiliou em muitas questões relativas a minha formação.

Ao meu orientador e amigo Aldrovando Araújo, a quem guardo um sincero respeito e admiração, que me ajudou de várias maneiras diferentes, não só na matemática.

Agradeço também ao suporte financeiro oferecido pela CAPES.

Resumo

Seja $T : X \to X$ uma aplicação expansora em um espaço métrico compacto X.

Demonstramos o teorema de Ruelle para potenciais na classe de Schwarz, que é uma classe um pouco mais geral do que a classicamente analisada (apenas Hölder). A demonstração de alguns itens do teorema de Ruelle fica trivial quando usamos o teorema de Ledrappier [W1], que caracteriza as g-medidas. Para o teorema sobre a existência e unicidade das g-medidas, T além de ser expansora deve ser, também, topologicamente mixing.

Por fim, estudamos funções expansoras por partes (também conhecidas como aplicações monotônicas, ou monótonas por partes). Além do espectro do operador de Ruelle relacionado com estas funções, apresentamos o estado de equilíbrio para tais funções.

Introdução

Durante as últimas três décadas a compreensão do comportamento assintótico de órbitas genéricas para inúmeras classes de sistemas dinâmicos se consolidou consideravelmente. De um ponto de vista inicialmente mais determinístico, os estudos evoluíram no sentido de considerar propriedades com caráter estatístico. Este procedimento se mostrou extremamente frutífero, permitindo-se avançar o entendimento dos sistemas uniformemente hiperbólicos que começara a consolidar-se nos anos oitenta para os não-uniformemente hiperbólicos que tem sido o grande desafio dos últimos anos. Essas técnicas foram desenvolvidas para sistemas uniformemente hiperbólicos mas, felizmente, puderam ser estendidas a uma classe mais ampla de transformações. Uma pequena parte desse desenvolvimento será apresentada neste trabalho. Para isso vamos tornar nossas idéias um pouco mais precisas.

Genérico é entendido aqui no sentido da teoria da medida. Assuma que $(X, \mathcal{A})$ é um espaço mensurável, com X um espaço métrico compacto, e $T : X \to X$ uma aplicação $\mathcal{A}$-mensurável (i. e., $T^{-1}(A) \in \mathcal{A}$ para qualquer $A \in \mathcal{A}$). Por $\mathcal{M}_T(X)$ denotamos o espaço das probabilidades invariantes por T no espaço de medida $(X, \mathcal{A}, \mu)$ (i. e., para todo $A \in \mathcal{A}$ vale que $\mu(A) = \mu(T^{-1}(A))$). No que segue, X será um espaço métrico compacto, $\mathcal{A}$ a σ-álgebra de Borel dada pela métrica de X e $C(X)$ denotará o espaço das funções de X em $\mathbb{R}$ contínuas. Considere uma função positiva $g : X \to \mathbb{R}$ usualmente chamada de potencial. Estamos interessados nas medidas de probabilidade que maximizam a pressão topológica, i. e., medidas $\mu \in \mathcal{M}_T(X)$ tais que

$$h_\mu(T) + \int_X \log g \, d\mu = \sup \left\{ h_\nu(T) + \int_X \log g \, d\nu; \nu \in \mathcal{M}_T(X) \right\},$$

onde $h_\mu(T)$ representa a entropia da transformação relativamente à medida μ, cuja definição é demasiadamente longa para ser apresentada nesta introdução, mas que está apresentada no apêndice deste trabalho. Pode-se definir a pressão topológica do potencial g via:

$$P(T, g) = \sup \left\{ h_\nu(T) + \int_X \log g \, d\nu; \nu \in \mathcal{M}_T(X) \right\}.$$

Uma medida que satisfaz a primeira igualdade é dita um estado de equilíbrio para o potencial g, caso $g \equiv 0$ então esta medida maximiza a entropia. Esses estados de equilíbrio têm propriedades ergódicas fortes e são relevantes em aplicações físicas. Para se obter uma tal medida uma técnica foi desenvolvida por Ruelle em seu trabalho pioneiro [R], para funções $g : X \to \mathbb{R}$ Hölder, que consiste no estudo das propriedades de convergências para iteradas do operador $\mathcal{L}_g : C(X) \to C(X)$, agindo no espaço das funções contínuas em um espaço métrico compacto a valores reais, sendo $T : X \to X$ uma aplicação expansora definimos

$$\mathcal{L}_g(\varphi)(x) = \sum_{y \in T^{-1}(x)} g(y)\varphi(y)$$

onde a quantidade de pré-imagens, por T, para cada $x \in X$ é finita, uma vez que X é compacto. Tal operador é conhecido usualmente como operador de Ruelle-Perron-Frobenius (ou operador de transferência) e é apresentado em várias formas permitindo, em algumas delas, ser estendido ao espaço das funções integráveis com respeito a alguma medida especial. Por exemplo, se o espaço ambiente é uma variedade riemanniana pode-se definir para uma aplicação expansora o operador de Ruelle-Perron-Frobenius, $\mathcal{L}$, como a aplicação em $L^1(\mu)$ tal que para toda $f \in L^1(\mu)$ e para todo boreliano A satisfaça

$$\int_A \mathcal{L}(f) d\mu = \int_{T^{-1}(A)} f d\mu,$$

onde μ aqui denota a medida de Lebesgue da variedade.

Para se obter boas propriedades de convergências das iteradas deste operador devemos fazer algumas restrições sobre o sistema dinâmico em questão, bem como sobre o potencial que começamos. Falando informalmente, o ambiente geral que garante a validade da técnica depende da exigência de um certo tipo de hiperbolicidade fraca (expande distâncias) no sistema dinâmico e na restrição do potencial para uma classe de funções boas onde o operador age como uma tranformação que essencialmente contrai.

Para vários casos de sistemas dinâmicos uma classe de potenciais g deve ser escolhida de modo a se obter um único estado de equilíbrio. O procedimento também é aplicado a sistemas descontínuos, em particular, para aplicações monótonas por partes do intervalo onde o potencial é assumido ser de variação limitada, [LY2],[Wo].

No início dos anos oitenta Keller, Hofbauer e outros começaram o estudo das propriedades espectrais do operador de Ruelle-Perron-Frobenius no caso de transformações monótonas por partes do intervalo, e conseguiram um enfoque unificado na maioria dos casos onde as propriedades de convergência foram obtidas das propriedades espectrais deste operador. Esta abordagem teve, após isto, um desenvolvimento importante trazendo à tona as condições espectrais necessárias que devemos exigir para obtermos as propriedades de convergência desejadas [B]. Finalmente novas técnicas foram desenvolvidas em parte por Lai-Sang Young, que se mostraram aplicáveis a outras situações onde menos hiperbolicidade é assumida sobre a aplicação.

Esta mesma técnica começada com Ruelle, que permite a obtenção de estados de equilíbrio, pode encontrar medidas SRB e estados de Gibbs ou até mesmo medidas absolutamente contínuas com respeito à medida de Lebesgue, quando o espaço é uma variedade riemanniana, desde que se escolha corretamente o potencial inicialmente estudado.

Neste trabalho vamos apresentar esta teoria para dois casos. O primeiro segue de perto a apresentação de Peter Walters em [W1] exceto que não é desenvolvido no caso de subshifts unilateral, abordada em Bowen [B1] por exemplo, mas em um ambiente mais geral que são as aplicações chamadas expansoras em espaços métricos [C1] que incluem transformações espansoras de variedades e subshifts unilaterais. Além disso o potencial é escolhido em uma classe diferente que chamamos de classe de Schwarz, cujas definição foi usada inicialmente por Schwarz na sua bem conhecida prova do teorema de Denjoy sobre a não existência de intervalos não-errantes para difeomorfismos de classe $C^{1+\varepsilon}$ do círculo. Na segunda parte desse trabalho apresentaremos o enfoque espectral, devido a Keller e Hofbauer [HK], para aplicações monótonas por partes do intervalo.

Capítulo 1

Transformações expansoras

Neste capítulo apresentaremos uma breve introdução à teoria ergódica e logo após definiremos transformações expansoras em variedades compactas e em espaços métricos compactos, e mostraremos que toda transformação expansora em variedade é expansora no sentido métrico. Na seção seguinte enunciaremos o teorema de Ruelle e faremos uma aplicação do teorema para um potencial pré-definido. Na última seção deste capítulo definiremos funções homólogas, enunciaremos o Shadowing Lemma e provaremos quando duas funções Hölder são homólogas. Por fim, provaremos que se duas funções Hölder são homólogas então elas possuem os mesmos estados de equilíbrio.

1.1 Elementos de Teoria Ergódica

Iniciaremos esta seção com uma breve revisão sobre teoria da medida, onde apenas enunciaremos algumas definições básicas. Para resultados em teoria da medida as referências são [Ca], [Co] ou [Ru]. Logo após apresentaremos algumas definições e resultados básicos de um curso de teoria ergódica, onde citamos [M1] e [W2].

Definição 1.1. *Seja X um conjunto arbitrário. Uma família $\mathcal{A}$ de subconjuntos de X é uma σ-álgebra se*

i) $X \in \mathcal{A}$

ii) se $A \in \mathcal{A}$ *então* $A^c = X - A \in \mathcal{A}$

iii) se $A_i \in \mathcal{A}$ *para* $i = 1, 2, 3, \ldots$ *então* $\bigcup_{i \geq 1} A_i \in \mathcal{A}$.

Definição 1.2. *Se $\mathcal{A}$ é uma σ-álgebra de subconjuntos de X, dizemos que $\mu : \mathcal{A} \to [0, +\infty]$ é uma medida se*

a) $\mu(\emptyset) = 0$

b) para toda família $\{A_i\}_i$, $i = 1, 2, 3, \ldots$, de conjuntos disjuntos em $\mathcal{A}$ vale que

$$\mu\left(\bigcup_{i\geq 1} A_i\right) = \sum_{i\geq 1} \mu(A_i).$$

Definição 1.3. *Um espaço de medida é uma terna $(X, \mathcal{A}, \mu)$ onde X é um conjunto arbitrário, $\mathcal{A}$ é uma σ-álgebra de subconjuntos de X e μ uma medida.*

Dizemos que o espaço de medida $(X, \mathcal{A}, \mu)$ é um espaço de probabilidade se $\mu(X) = 1$, neste caso a medida μ é dita ser uma medida de probabilidade ou, simplesmente, uma probabilidade.

Definição 1.4. *Seja $(X, \mathcal{A}, \mu)$ um espaço de probabilidade. Se $\mathcal{A}_0$ é uma família de subconjuntos de X, dizemos que $\mathcal{A}$ é gerada por $\mathcal{A}_0$ se $\mathcal{A}_0 \subseteq \mathcal{A}$ e toda σ-álgebra $\mathcal{A}'$ de subconjuntos de X tal que $\mathcal{A}_0 \subseteq \mathcal{A}'$ satisfaz $\mathcal{A} \subseteq \mathcal{A}'$. Se $\mathcal{A}_n$ é uma família de subconjuntos de X, $n \geq 1$, denotamos por $\bigvee_{n\geq 1} \mathcal{A}_n$ a σ-álgebra gerada por $\bigcup_{n\geq 1} \mathcal{A}_n$.*

Obs.: Se X é um espaço topológico, a σ-álgebra de Borel de X é a σ-álgebra gerada pela família dos conjuntos abertos. Neste caso, os conjuntos em $\mathcal{A}$ denominam-se borelianos de X.

Definição 1.5. *Seja $(X, \mathcal{A}, \mu)$ um espaço de medida. Definimos o suporte de μ pelo conjunto*

$$sup(\mu) = \overline{\{x \in X; \forall V_x, \mu(V_x) > 0\}}$$

onde V_x é uma vizinhança do ponto x.

Definição 1.6. *Se $(X, \mathcal{A}, \mu)$ é um espaço de medida e para todo $A \in \mathcal{A}$, $A \subseteq \bigcup_{i=1}^{\infty} A_i$, com $\mu(A_i) < +\infty$ para todo $i \geq 1$, dizemos que μ é uma medida σ-finita.*

Um dos principais objetivos da teoria ergódica é o estudo da dinâmica das transformações que preservam medida. Assim, vamos à definição de uma medida invariante.

Definição 1.7. *Sejam $(X, \mathcal{A}, \mu)$ e $(Y, \mathcal{B}, \nu)$ espaços de medida.*

a) Dizemos que uma transformação $T : (X, \mathcal{A}, \mu) \to (Y, \mathcal{B}, \nu)$ *é mensurável se para todo* $A \in \mathcal{B}$ *tivermos que* $T^{-1}(A) \in \mathcal{A}$.

b) A transformação $T : (X, \mathcal{A}, \mu) \to (Y, \mathcal{B}, \nu)$ *preserva medida se é mensurável e para todo* $A \in \mathcal{B}$ *temos que* $\mu(T^{-1}(A)) = \nu(A)$.

Obs.: Estaremos interessados principalmente em casos onde $(X, \mathcal{A}, \mu) = (Y, \mathcal{B}, \nu)$. Se $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ é uma transformação que preserva medida, também dizemos que T preserva μ ou que μ é T-invariante, ou ainda, que μ é invariante por T.

Definição 1.8. *Seja* $T : X \to X$ *uma aplicação contínua em um espaço métrico compacto. O conjunto das probabilidades sobre os borelianos de* X *é denotado por* $\mathcal{M}(X)$ *e denotamos por* $\mathcal{M}_T(X)$ *o conjunto das* $\mu \in \mathcal{M}(X)$ *invariantes por* T.

Obs.: Seja $T : X \to X$ uma transformação contínua em um espaço métrico compacto. Como consequência do teorema de Krylov e Bogolioubov [W2], temos que o espaço $\mathcal{M}_T(X)$ não é vazio e é um subconjunto convexo de $\mathcal{M}(X)$ e ainda é fracamente compacto.

Definição 1.9. *Uma transformação* $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ *em um espaço de probabilidade que preserva medida é chamada ergódica se os únicos membros* $A \in \mathcal{A}$ *tais que* $T^{-1}(A) = A$ *satisfazem* $\mu(A) = 0$ *ou* $\mu(A) = 1$. *Neste caso, dizemos que* μ *é uma medida ergódica.*

Definição 1.10. *Seja* $(X, \mathcal{A}, \mu)$ *um espaço de probabilidade. Dizemos que um conjunto* $A \subseteq X$ *é de medida nula se existe* $A_1 \in \mathcal{A}$ *tal que* $A \subseteq A_1$ *e* $\mu(A_1) = 0$. *Dizemos que dois conjuntos* $A_1, A_2 \in \mathcal{A}$ *coincidem* μ *mod* 0, *e o denotamos por* $A_1 = A_2$ *mod* 0 *se* $A_1 \Delta A_2 := (A_1 - A_2) \cup (A_2 - A_1)$ *é de medida nula.*

Definição 1.11. *Seja* $(X, \mathcal{A}, \mu)$ *um espaço de probabilidade. Uma propriedade aplicada a pontos de um subconjunto* $S \subseteq X$ *vale em* μ *quase todo ponto (abreviadamente* μ *q.t.p., ou simplesmente, q.t.p.), ou quase sempre (*μ *q.s., ou simplesmente, q.s.) se o conjunto dos pontos de* S *onde a propriedade é falsa tem medida nula.*

Definição 1.12. *Seja* $(X, \mathcal{A}, \mu)$ *um espaço de probabilidade. Se* $\mathcal{B}$ *é uma família de subconjuntos de* X, *escrevemos* $A \in \mathcal{B}$ *mod* 0 *se* $A = A_0$ *mod* 0 *para algum* $A_0 \in \mathcal{B}$ *e definimos*

$$\mathcal{B} \text{ mod } 0 = \{A \subseteq X; A \in \mathcal{B} \text{ mod } 0\}.$$

Dizemos que $\mathcal{B}$ *gera* $\mathcal{A}$ *mod* 0 *se* $\mathcal{A} = \mathcal{A}_0$ *mod* 0, *onde* $\mathcal{A}_0$ *é a* σ*-álgebra gerada por* $\mathcal{B}$.

Definição 1.13. *Seja $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ uma transformação que preserva medida em um espaço de probabilidade. Dizemos que T é uma transformação exata ou que μ é uma medida exata com respeito a T, se para todo $A \in \bigcap_{n\geq 0} T^{-n}(\mathcal{A})$ mod 0, (ou seja, A pertence a $\bigcap_{n\geq 0} T^{-n}(\mathcal{A})$ a menos de um conjunto de medida nula) tivermos que $\mu(A) = 0$ ou $\mu(A) = 1$.*

Precisaremos de alguns teoremas de integração, os mais utilizados são os seguintes.

Definição 1.14. *Seja X um espaço métrico compacto. Definimos por $C(X)$ o espaço de todas as funções contínuas de X em $\mathbb{R}$. Sobre $C(X)$ usaremos a seguinte norma $\|f\|_C = \sup\{|f(x)|; x \in X\}$.*

Teorema 1.1. *Sejam ν, μ duas medidas de probabilidade de Borel sobre o espaço métrico X. Se*

$$\int_X f d\mu = \int_X f d\nu, \qquad \forall f \in C(X)$$

então $\mu = \nu$.

Demonstração. [W2] □

Teorema 1.2 (Representação de Riesz). *Sejam X um espaço métrico compacto e $L : C(X) \to \mathbb{R}$ uma aplicação linear contínua que satisfaça $L(1) = 1$ e $L(f) \geq 0$ para toda $f \geq 0$. Então existe $\mu \in \mathcal{M}(X)$ tal que*

$$L(f) = \int_X f d\mu, \qquad \forall f \in C(X).$$

Demonstração. [Ca] □

Obs.: Em [W2] temos a prova de que $\mathcal{M}(X)$ é identificado com um subconjunto convexo da bola unitária em $C(X)^*$, o dual de $C(X)$. Por este motivo podemos obter uma topologia para $\mathcal{M}(X)$ através da topologia fraca-* em $C(X)^*$.

Definição 1.15. *A topologia fraca-* sobre $\mathcal{M}(X)$, onde X é um espaço métrico compacto, é a menor topologia onde cada aplicação $\mu \to \int_X f d\mu$, para $f \in C(X)$, é contínua. Uma base para tal topologia é dada pela coleção de todos os conjuntos da forma*

$$V_\mu(f_1, \ldots, f_k; \varepsilon) = \left\{ \nu \in \mathcal{M}(X); \left| \int_X f_i d\nu - \int_X f_i d\mu \right| < \varepsilon, 1 \leq i \leq k \right\}$$

onde $\mu \in \mathcal{M}(X)$, $k \geq 1$, $\varepsilon > 0$ e para todo $i \in \{1, 2, \ldots, k\}$, $f_i \in C(X)$.

Teorema 1.3. *Se X é um espaço métrico compacto então $\mathcal{M}(X)$ é compacto na topologia fraca-*.*

Demonstração. [W2] $\square$

Teorema 1.4. *Sejam $T : X \to X$ contínua em um espaço métrico compacto e $\mu \in \mathcal{M}(X)$. Então, $\mu \in \mathcal{M}_T(X)$ se e somente se*

$$\int_X f \circ T d\mu = \int_X f d\mu, \qquad \forall f \in C(X).$$

Demonstração. [W2] $\square$

Definição 1.16. *Uma transformação entre espaços métricos $\psi : (X, d_X) \to (Y, d_Y)$ é Hölder contínua, ou simplesmente Hölder, se existem $M > 0$ e $\gamma > 0$ tais que*

$$d_Y(\psi(x), \psi(y)) \leq M d_X(x, y)^\gamma$$

para $x, y \in X$. A constante γ é o expoente de Hölder. No caso $Y = \mathbb{R}$ dizemos que $\psi : X \to \mathbb{R}$ é uma função Hölder.

Definição 1.17. *Dizemos que $\psi : X \to X$ em uma variedade riemanniana é de classe $C^{k+\gamma}$, para $\gamma > 0$ e $k \geq 1$, se ψ é de classe C^k e a derivada $D^k(\psi)$ é Hölder contínua com γ sendo o expoente de Hölder.*

Por fim, caso $T : X \to X$ seja uma transformação contínua em um espaço métrico compacto, temos que $\mathcal{M}_T(X) \neq \emptyset$. Então cabe perguntar se $\mathcal{M}_T(X)$ contém elementos ergódicos.

Até o final desta seção estaremos trabalhando com $T : X \to X$ uma transformação mensurável em um espaço métrico compacto. Para maiores informações veja [M1].

Definição 1.18. *Definimos $\Sigma_0(T)$ como o conjunto dos pontos $x \in X$ tais que, para toda $f : X \to \mathbb{R}$ contínua, exista o limite*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x)).$$

Se $x \in \Sigma_0(T)$ definimos $L_x : C(X) \to \mathbb{R}$ por

$$L_x(f) = \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(T^j(x))$$

que é linear e positivo donde, pelo teorema 1.2, existe uma única probabilidade μ_x sobre os borelianos de X tal que

$$L_x(f) = \int_X f d\mu_x.$$

Definição 1.19. *Definimos $\Sigma_1(T)$ como o conjunto dos $x \in \Sigma_0(T)$ tais que μ_x é T-invariante.*

Obs.: Quando T é contínua, $\Sigma_0(T) = \Sigma_1(T)$ [M1].

Definição 1.20. *Definimos $\Sigma_2(T)$ como o conjunto dos $x \in \Sigma_1(T)$ tais que μ_x é ergódica, e também definimos $\Sigma(T)$ como o conjunto dos $x \in \Sigma_2(T)$ tais que $x \in sup(\mu_x)$.*

Obs.: Os conjuntos $\Sigma(T)$, $\Sigma_i(T)$, $i = 0, 1, 2$ são borelianos [M1].

Definição 1.21. *Um conjunto $A \subseteq X$ é de probabilidade total se $\mu(A^c) = 0$ para toda $\mu \in \mathcal{M}_T(X)$.*

Por fim, enunciamos o teorema da decomposição ergódica das medidas invariantes.

Teorema 1.5. *Se $\mathcal{M}_T(X) \neq \emptyset$, $\Sigma(T)$ é um conjunto de probabilidade total.*

Demonstração. [M1] □

1.2 Transformações Expansoras

As transformações expansoras em variedades são aplicações de grande importância na teoria de sistemas dinâmicos e, consequentemente, na teoria ergódica. Seja, portanto, a definição.

Definição 1.22. *Seja X uma variedade riemanniana compacta sem bordo. Dada a transformação $T : X \to X$ de classe C^r, $r \geq 1$, dizemos que T é expansora se existe $\alpha > 1$ tal que*

$$\|DT(x)v\| \geq \alpha \|v\|$$

para todo $x \in X$ e $v \in T_xX$.

O exemplo mais imediato deste tipo de transformação é:

Exemplo1. Se $X = S^1 = \{z \in \mathbb{C}; |z| = 1\}$, seja $T : S^1 \hookleftarrow$ dada por

$$T(z) = z^n, \qquad\qquad \text{com } n \geq 2.$$

Então S^1 com a métrica induzida de $\mathbb{C} \simeq \mathbb{R}^2$ satisfaz

$$T'(z) = nz^{n-1}$$

$$\|T'(z)\| = n\|z^{n-1}\| = n$$

e então se fizermos $\alpha = n$ teremos o resultado. □

Exemplo 2. Seja A uma matriz 2×2 de coeficientes inteiros satisfazendo $|\alpha| > 1$ para qualquer $\alpha \in \text{esp}(A)$, onde $\text{esp}(A)$ é o espectro de A. Temos que $A(\mathbb{Z}^2) \subseteq \mathbb{Z}^2$. Deste fato segue que A induz sobre o toro $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ a aplicação $T : \mathbb{T}^2 \to \mathbb{T}^2$ tal que se $\pi : \mathbb{R}^2 \to \mathbb{T}^2$ é a projeção canônica, i. e., $\pi(x) = [x] = x + \mathbb{Z}^2$ então $T \circ \pi = \pi \circ A$.

Do fato que A tem apenas autovalores com módulo maior que um, segue que se $1 < \alpha < \inf\{|\gamma|; \gamma \in \text{ esp } (A)\}$ então

$$\|D_x T v\| \geq \alpha \|v\|.$$

□

Estão caracterizadas todas as variedades que admitem uma transformação expansora (ver [S1] e [G1]). A existência de medidas especiais para as transformações expansoras em variedades está provada quando a classe de diferenciabilidade é $C^{1+\varepsilon}$, $\forall \varepsilon > 0$. Em particular elas admitem medidas SRB, medidas de equilíbrio, etc. Um dos objetivos deste trabalho é a apresentação destas propriedades em um contexto mais geral que engloba o conceito de transformação expansora de variedades, mas que atinge também os shifts unilaterais e outros exemplos em espaços métricos. É importante observar que a condição $\alpha > 1$, na definição de transformação expansora, é claramente uma condição aberta na topologia C^1 e que portanto, toda aplicação suficientemente C^1-próxima de uma aplicação destas também é expansora (para uma apresentação sobre a topologia C^1 veja [S2]).

Um resultado importante é o seguinte:

Teorema 1.6. *Dada $T : X \to X$ expansora em um espaço métrico compacto X existe uma única probabilidade, μ, sobre os borelianos de X, invariante por T e que seja* *absolutamente contínua relativamente à medida de Lebesgue, m, de X. Além disso, μ pode ser obtida pelo limite*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} T_*^j(m) = \mu$$

onde o limite acima é considerado na topologia fraca-$*$ e $T_*^j(m)$ denota a medida

$$T_*^j(m)(A) = m(T^{-j}(A))$$

para qualquer $A \in \mathcal{A}$ onde $\mathcal{A}$ é a σ-álgebra dos borelianos de X.

Vamos generalizar o conceito de transformação expansora que engloba o anterior e inclui os shifts unilaterais.

Definição 1.23. *Seja X um espaço métrico compacto e d sua métrica. Uma transformação expansora é uma aplicação $T : X \to X$ contínua satisfazendo:*

existem $r > 0$, $0 < \alpha < 1$ e $c > 0$ tais que

i) $x \neq y$, $T(x) = T(y) \Rightarrow d(x, y) > c$

ii) $\forall x \in X$ *e* $a \in T^{-1}(x)$ *existe uma função contínua* $\varphi : B_r(x) \to X$ *tal que*

$$\varphi(x) = a$$

$$(T \circ \varphi)(z) = z, \quad \forall z \in B_r(x)$$

$$d(\varphi(z), \varphi(z')) \leq \alpha d(z, z'), \quad \forall z, z' \in B_r(x).$$

Obs.: A função φ é denominada de ramo da inversa de T.

No que segue, iremos definir subshift unilateral do tipo finito.

Para $n \geq 1$ seja $I_n = \{1, 2, \cdots, n\}$. Seja $B^+(n)$ o conjunto das sequências de n símbolos, ou seja,

$$B^+(n) = \{\theta; \theta : \mathbb{N} \to I_n\}.$$

Definimos um cilindro de largura m com início em j como sendo o conjunto

$$C(j; i_0, i_1, \ldots, i_{m-1}) = \{\theta \in B^+(n); \theta(j + l) = i_l, i_l \in I_n, l = 0, 1, \ldots, m - 1\}.$$

Podemos munir $B^+(n)$ da topologia produto, associada à topologia discreta de I_n, donde pelo teorema de Tychonoff (veja [Ke]) segue que $B^+(n)$ é um espaço topológico compacto. Os cilindros formam uma base para a topologia produto de $B^+(n)$, veja [B].

Sobre $B^+(n)$ definimos

$$d'(\theta,\eta) = \begin{cases} \frac{1}{2^{k+1}}, & \theta \neq \eta \\ 0, & \theta = \eta \end{cases}$$

onde k é o menor índice tal que $\theta(k) \neq \eta(k)$. Temos que d' é uma métrica e $(B^+(n), d')$ é um espaço métrico compacto.

Seja $A = (a_{ij})$ uma matriz $n \times n$ formada por 0 e 1, chamada matriz de transição. Considere $B^+(A) \subseteq B^+(n)$ o subconjunto dado por

$$B^+(A) = \{\theta \in B^+(n); a_{\theta(i)\theta(i+1)} = 1,\ \forall i \in \mathbb{N}\}.$$

Sobre $B^+(A)$ consideremos a aplicação $\sigma : B^+(A) \to B^+(A)$ dada por

$$\sigma(\theta)(i) = \theta(i+1), \qquad i = 0, 1, \ldots$$

tal aplicação é denominada de shift.

Definição 1.24. *Ao sistema dinâmico*

$$(B^+(A), \sigma)$$

como definido acima, denominamos de subshift unilateral de tipo finito associado à matriz A.

Lema 1.1. *A aplicação $\sigma : B^+(A) \to B^+(A)$ é expansora.*

Demonstração. Sejam θ e γ dois elementos de $B^+(A)$ tais que

$$\theta \neq \gamma \qquad \text{e} \qquad \sigma(\theta) = \sigma(\gamma).$$

Então,

$$\sigma(\theta)(i) = \sigma(\gamma)(i), \qquad \forall i = 0, 1, \ldots$$

$$\theta(i+1) = \gamma(i+1), \qquad \forall i = 0, 1, \ldots$$

$$\theta(i) = \gamma(i), \qquad \forall i = 1, 2, \ldots$$

Como $\theta \neq \gamma \Rightarrow \theta(0) \neq \gamma(0)$ e portanto

$$d'(\theta, \gamma) = \frac{1}{2}.$$

Fazendo $0 < c < 1/2$ provamos i).

Seja $\theta \in B^+(A)$. Então

$$\sigma^{-1}(\theta) = \{(i, \theta(0), \theta(1), \ldots); 1 \leq i \leq n \text{ e } a_{i\theta(0)} = 1\}.$$

Sejam $r = 1/2$ e $\gamma = (i, \theta(0), \theta(1), \ldots)$ com $a_{i\theta(0)} = 1$. Assim, se $\omega \in B_r(\theta)$ então ω satisfaz $\omega(0) = \theta(0)$. Definimos a aplicação $\varphi : B_r(\theta) \to B^+(A)$ dada por $\varphi(\omega) = i\omega$, onde $i\omega$ denota o elemento de $B^+(A)$ definido por

$$(i\omega)(j) = \begin{cases} i, & j = 0 \\ \omega(j-1), & j \geq 1. \end{cases}$$

Então, para quaisquer $\omega_1, \omega_2 \in B_r(\theta)$ temos

$$d'(\varphi(\omega_1), \varphi(\omega_2)) = d'(i\omega_1, i\omega_2) = \frac{1}{2} d'(\omega_1, \omega_2).$$

Basta pois tomar $\alpha = 1/2$, terminando assim a prova do lema. $\square$

Lema 1.2. *Seja $T : X \to X$ uma aplicação expansora em uma variedade compacta. Então T é expansora no sentido métrico.*

Demonstração. Sejam $T : X \to X$ uma aplicação expansora de variedade e $(U_i)_i$ uma cobertura finita de X tal que $\forall i$

$$f^{-1}(U_i) = \bigcup_{j=1}^{k} V_{ij}$$

onde $f : V_{ij} \to U_i$ é um difeomorfismo. Seja r_1 o número de Lebesgue desta cobertura. Seja $r_2 > 0$ tal que se $p, q \in X$ e $d(p,q) < 2r_2$ então, da compacidade de X, existe uma geodésica minimizante ligando p e q, i. e., $\gamma : [0,1] \to X$ geodésica com $\gamma(0) = p$, $\gamma(1) = q$ e $d(p,q) = l(\gamma)$.

Tome $r = \inf\{r_1, r_2\}$. Seja $p \in X$, então existe um aberto U_i tal que $B_r(p) \subset U_i$, $\gamma = T \circ \beta$ onde β é uma curva (que pode ser ou não uma geodésica) que liga $\varphi(p)$ a $\varphi(q)$, onde $\varphi : B_r(p) \to X$ é um ramo da inversa de T. Assim,

$$\int_X \|\dot{\gamma}(t)\| dt = \int_X \|T'(\beta(t))\dot{\beta}(t)\| dt \geq \int_X \alpha \|\dot{\beta}(t)\| dt = \alpha l(\beta)$$

portanto $l(\gamma) \geq \alpha l(\beta)$ donde $d(p,q) \geq \alpha l(\beta) \geq \alpha d(\varphi(p), \varphi(q))$ e, por fim, temos que

$$d(\varphi(p), \varphi(q)) \leq \frac{1}{\alpha} d(p,q).$$

Como $\alpha > 1$, $0 < 1/\alpha < 1$ e a prova de ii) está concluída.

Para provarmos i), como X é compacto e supondo i) falso teríamos sequências $(p_n)_n$ e $(q_n)_n$ em X tais que

$$p_n \to p \qquad\qquad q_n \to q$$

$$T(p_n) = T(q_n)$$

e temos que

$$\varphi : B_r(T(p)) \to X.$$

Como φ é um homeomorfismo local, chegamos a uma contradição. $\square$

Isto mostra que a definição de transformação expansora engloba a grande maioria dos exemplos conhecidos.

Definição 1.25. *Seja $T : X \to X$ uma tranformação bijetora em um espaço métrico. A órbita de um ponto $x \in X$ é o conjunto $\{T^n(x); n \in \mathbb{Z}\}$.*

Definição 1.26. *Seja $T : X \to X$ uma transformação contínua em um espaço métrico compacto. Dizemos que um ponto $x \in X$ é um ponto periódico para T se $T^n(x) = x$ para algum $n \geq 1$. O menor $n \in \mathbb{N}$ cuja igualdade é válida, é o período do ponto x. Caso $n = 1$ dizemos que o ponto é fixo.*

Definição 1.27. *Dizemos que uma transformação $T : X \to X$ contínua em um espaço topológico X é topologicamente mixing se para todo par de abertos U e V em X existe $N > 0$ tal que*

$$T^{-n}(U) \cap V \neq \emptyset$$

para todo $n \geq N$.

Obs.: Pode-se provar que se uma transformação $T : X \to X$ é topologicamente mixing em um espaço topológico então é transitiva, ou seja, existe uma órbita densa em X, ver [M1].

Sejam $T : X \to X$ expansora no espaço métrico compacto X e $\mathrm{Per}(T)$ o conjunto dos pontos periódicos de T, denotaremos por Λ o fecho dos pontos periódicos de T, ou seja, $\Lambda = \overline{\mathrm{Per}(T)}$.

Teorema 1.7. *Sejam $T : X \to X$ uma transformação expansora em um espaço métrico compacto X e $\Lambda = \overline{Per(T)}$. Então existem compactos disjuntos únicos $\Lambda_i^{(m)} \subseteq X$ onde $i = 1, \ldots, n_m$, $m = 1, \ldots, N$ tais que* *(a)* $T(\Lambda_i^{(m)}) = \Lambda_{i+1}^{(m)}$, $1 \leq i \leq n_m$, $1 \leq m \leq N$

$T(\Lambda_{n_m}^{(m)}) = \Lambda_1^{(m)}, \qquad\qquad 1 \leq m \leq N$

(b) $\bigcup_{i,m} \Lambda_i^{(m)} = \Lambda$,

(c) $T^{n_m}\Big|_{\Lambda_i^{(m)}} : \Lambda_i^{(m)} \to \Lambda_i^{(m)}$ *é topologicamente mixing*

Além disso, valem as seguintes propriedades:

(d) $T^{n_m}\Big|_{\Lambda_i^{(m)}} : \Lambda_i^{(m)} \to \Lambda_i^{(m)}$ *é uma transformação expansora*

(e) para todo aberto U *de* $\Lambda_i^{(m)}$*, existe* $M > 0$ *tal que*

$$(T^{n_m})^M(U) = \Lambda_i^{(m)}.$$

Demonstração. [C1] □

Lema 1.3. *Seja* $T : X \to X$ *uma transformação expansora topologicamente mixing em um espaço métrico compacto, então*

a) para qualquer aberto $U \subseteq X$ *existe um* $M > 0$ *tal que* $T^M(U) = X$.

b) dado $\varepsilon > 0$ *existe* $M > 0$ *tal que para qualquer* $x \in X$, $T^{-M}(x)$ *é* ε*-denso em* X.

Demonstração. a) Sendo T topologicamente mixing, a decomposição de $\Lambda = \overline{\mathrm{Per}(T)}$, dada pelo teorema 1.7, é formada por um único compacto [M1] donde no teorema anterior temos que $N = 1$ e, pelo item (e) do citado teorema para todo aberto U de Λ existe um $M > 0$ tal que $T^M(U) = \Lambda = X$.

A igualdade $X = \Lambda$ depende do Shadowing Lemma e será provada na seção 1.4.

b) Da compacidade de X é possível escolher uma cobertura de abertos $\mathcal{U}$ de X formada por um número finito de bolas de raio $\varepsilon/2$. Já sabemos, por a), que existe um $M > 0$ tal que

$$T^M(U) = X$$

para toda $U \in \mathcal{U}$. Agora, escolha qualquer $x \in X$ e considere a bola $B(y, \varepsilon)$ para algum $y \in X$. Então, existe $U \in \mathcal{U}$ tal que $y \in U$ mas como $\mathrm{diam}(U) < \varepsilon$ segue que $U \subseteq B(y, \varepsilon)$. Além disso,

$$T^M(B(y, \varepsilon)) \supseteq T^M(U) = X.$$

Em particular, existe $z \in U \subseteq B(y,\varepsilon)$ tal que $T^M(z) = x$, provando que $T^{-M}(x)$ é ε-denso. $\square$

1.3 O teorema de Ruelle e algumas aplicações

Para as próximas definições precisaremos do conceito de entropia de uma transformação e da prova da existência de partições geradoras para uma transformação expansora, que estão apresentados no apêndice deste trabalho.

Definição 1.28 (Estados de Equilíbrio). *Sejam $\psi : X \to \mathbb{R}$ uma aplicação contínua em um espaço métrico compacto X, $T : X \to X$ uma aplicação expansora e $\nu \in \mathcal{M}(X)$ (uma medida invariante por T). A pressão topológica de ψ é dada por*

$$P(T,\psi) = \sup\left\{h_\nu(T) + \int_X \psi d\nu; \nu \in \mathcal{M}_T(X)\right\}$$

onde $h_\nu(T)$ denota a entropia de T com relação à medida ν. Dizemos que $\mu \in \mathcal{M}_T(X)$ é um estado de equilíbrio associado a ψ se satisfaz

$$P(T,\psi) = h_\mu(T) + \int_X \psi d\mu.$$

Definição 1.29 (Medidas SRB). *Seja $T : (X,\mathcal{A},\mu) \to (X,\mathcal{A},\mu)$ uma transformação mensurável em uma variedade riemanniana. A bacia de μ é o conjunto $B(\mu)$ dos pontos $y \in X$ tais que*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi(T^j(y)) = \int_X \varphi d\mu$$

para qualquer função contínua $\varphi : X \to \mathbb{R}$. Dizemos que a medida μ é física ou SRB (Sinai-Ruelle-Bowen) para T se sua bacia $B(\mu)$ tem medida de Lebesgue positiva.

É possível provar, veja [V], que para $T : X \to X$ expansora em variedade temos que μ é SRB se para os conjuntos $A \subseteq X$, de medida de Lebesgue total, i. e., $m(X - A) = 0$ tivermos que para todo $x \in A$

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=0}^{n-1} \varphi(T^j(x)) = \int_X \varphi d\mu$$

onde $\varphi : X \to \mathbb{R}$ é uma função contínua.

Medidas SRB podem ser entendidas como aquelas medidas invariantes compatíveis com o volume quando este não é preservado pela transformação. Uma observação a ser feita a respeito das medidas SRB para transformações expansoras é que essas medidas são únicas, e sempre existem [V].

Finalmente apresentamos a noção de medida de Gibbs que são medidas invariantes definidas sobre os shifts. Para isto seja a função $\varphi : B^+(n) \to \mathbb{R}$ contínua e defina sua variação por

$$\mathrm{var}_k(\varphi) = \sup\{|\varphi(\theta) - \varphi(\eta)|; \theta(i) = \eta(i), \forall i \leq k\}$$

para $k \geq 1$.

Definição 1.30 (Medidas de Gibbs). *Uma medida de Gibbs para o potencial φ é uma medida de probabilidade invariante μ tal que existem constantes $c_1 > 0$, e $P > 0$ satisfazendo: para todo cilindro $C(0; i_0, \ldots, i_{m-1})$ vale*

$$c_1 \leq \frac{\mu(C(0; i_0, \ldots, i_{m-1}))}{\exp\left(-mP + \sum\limits_{k=0}^{m-1} \varphi(\sigma^k(\theta))\right)} \leq c_1^{-1}.$$

Definição 1.31. *Sejam $\psi : X \to \mathbb{R}$ uma função contínua numa variedade compacta X e $T : X \to X$ expansora. O operador de Ruelle-Perron-Frobenius (ou operador de transferência) $\mathcal{L}_\psi : C(X) \to C(X)$ é definido por*

$$\mathcal{L}_\psi \varphi(x) = \sum_{y \in T^{-1}(x)} e^{\psi(y)} \varphi(y)$$

para $\varphi \in C(X)$ e $x \in X$.

Para as várias propriedades do operador de Ruelle-Perron-Frobenius citamos [R], [AB], [LY2] e [B].

Obs.: 1) Sejam X uma variedade compacta e k o grau topológico de $T : X \to X$, i. e., o número de pré-imagens da aplicação T. Se a aplicação for um difeomorfismo local, prova-se que este número é constante e independe dos pontos $x \in X$, veja [E]. Observe que a definição de transformação expansora em variedade implica que T é um difeomorfismo local.

Obs.: 2) Em espaços métricos compactos vamos supor que o número de pré-imagens de um ponto seja um conjunto contável (finito ou enumerável) e que o operador de Ruelle-Perron-Frobenius seja somável.

Teorema 1.8 (Ruelle). *Sejam $T : X \to X$ expansora em uma variedade riemanniana compacta e $\psi : X \to \mathbb{R}$ Hölder contínua. Então existem $h : X \to \mathbb{R}$ Hölder contínua e estritamente positiva, $\nu \in \mathcal{M}(X)$ e $\lambda > 0$ tais que*

a) $\mathcal{L}_\psi h = \lambda h$

b) $\mathcal{L}_\psi^* \nu = \lambda \nu$

c) $\int_X h d\nu = 1$

d) $\forall \varphi \in C(X)$,

$$\left\| \frac{1}{\lambda^n} \mathcal{L}_\psi^n \varphi - h \int_X \varphi d\nu \right\|_C \to 0$$

e) h é a única auto-função positiva de $\mathcal{L}_\psi$, a menos de multiplicação por escalar

f) A probabilidade $\mu = h\nu$ é T-invariante, exata, positiva sobre abertos e satisfaz

$$\log \lambda = h_\mu(T) + \int_X \psi d\mu$$

g) $\forall \eta \in \mathcal{M}_T(X), \eta \neq \mu$

$$\log \lambda > h_\eta(T) + \int_X \psi d\eta.$$

Vejamos como este teorema pode ser aplicado às transformações expansoras de variedade na obtenção de medidas absolutamente contínuas relativamente à medida de Lebesgue bem como de medidas que maximizam a entropia.

No restante desta seção estaremos trabalhando sempre com $T : X \to X$ uma aplicação de classe $C^{1+\gamma}$, com $\gamma > 0$, expansora em uma variedade riemanniana compacta e com a função

$$\psi(x) = -\log|\det D_x T|$$

e é fácil ver que a função ψ é Hölder. Assim, o operador de Ruelle-Perron-Frobenius usado no restante desta seção é

$$\mathcal{L}_\psi \varphi(x) = \sum_{y \in T^{-1}(x)} \frac{\varphi(y)}{|\det D_y T|}.$$

A σ-álgebra dos borelianos de X será denotada por $\mathcal{A}$. E vamos assumir que a medida de Lebesgue em X, m, é uma probabilidade, ou seja, $m(X) = 1$.

Afirmação 1. *Para* $\psi(x) = -\log|\det D_x T|$ *temos que se* m *é a medida de Lebesgue em* X *então* $\mathcal{L}_\psi^* m = m$.

Demonstração. Seja m a medida de Lebesgue de X, queremos calcular $\mathcal{L}_\psi^* m$. Observe que $(\mathcal{L}_\psi^* m)\varphi = m(\mathcal{L}_\psi \varphi)$ e assim

$$(\mathcal{L}_\psi^* m)\varphi = \int_X \varphi d\mathcal{L}_\psi^* m = \int_X \mathcal{L}_\psi \varphi dm = \int_X \sum_{y \in T^{-1}(x)} \frac{\varphi(y)}{|\det D_y T|} dm.$$

Para calcularmos esta integral subdividimos a variedade compacta X em abertos disjuntos $(U_i)_i$ tais que $\forall i$

$$f^{-1}(U_i) = \bigcup_{j=1}^{k} V_{ij}$$

onde $f : V_{ij} \to U_i$ é um difeomorfismo local. Assim,

$$\int_X \sum_{y \in T^{-1}(x)} \frac{\varphi(y)}{|\det D_y T|} dm(x) = \sum_{i=1}^{n} \int_{U_i} \sum_{y \in T^{-1}(x)} \frac{\varphi(y)}{|\det D_y T|} dm(x).$$

Sejam $g_{ij} : U_i \to V_{ij}$ os ramos da inversa de T em U_i. Então

$$\begin{aligned}
\int_{U_i} \sum_{y \in T^{-1}(x)} \frac{\varphi(y)}{|\det D_y T|} dm(x) &= \int_{U_i} \sum_{j=1}^{k} \frac{\varphi(g_{ij}(x))}{|\det D_{g_{ij}(x)} f|} dm(x) \\
&= \sum_{j=1}^{k} \int_{f(V_{ij})} \frac{\varphi(g_{ij}(x))}{|\det D_{g_{ij}(x)} f|} dm(x) \\
&= \sum_{j=1}^{k} \int_{V_{ij}} \frac{\varphi(g_{ij}(f(z)))}{|\det D_{g_{ij}(f(z))} f|} |\det D_z f| dm(z) \\
&= \sum_{j=1}^{k} \int_{V_{ij}} \varphi(z) dm(z) \\
&= \int_{\bigcup_{j=1}^{k} V_{ij}} \varphi(z) dm(z).
\end{aligned}$$

logo, $\forall \varphi \in C(X)$

$$\begin{aligned}
\int_X \sum_{y \in T^{-1}(x)} \frac{\varphi(y)}{|\det D_y f|} dm(x) &= \sum_{i=1}^{n} \int_{U_i} \sum_{y \in T^{-1}(x)} \frac{\varphi(y)}{|\det D_y f|} dm(x) \\
&= \sum_{i,j} \int_{V_{ij}} \varphi(x) dm(x) \\
&= \int_X \varphi(x) dm(x)
\end{aligned}$$

ou seja,

$$\mathcal{L}_\psi^* m = m.$$

□

Afirmação 2. *Sejam ν, h e μ dadas pelo teorema de Ruelle então $\mu = hm$ onde m é a medida de Lebesgue.*

Demonstração. Do item d) do teorema de Ruelle temos que

$$\frac{1}{\lambda^n}\mathcal{L}_\psi^n \varphi - h\int_X \varphi d\nu \to 0$$

integrando relativamente a m tem-se

$$\begin{aligned}
\frac{1}{\lambda^n}\int_X \mathcal{L}_\psi^n \varphi dm &\to \int_X hdm \int_X \varphi d\nu \\
\frac{1}{\lambda^n}\int_X \varphi d\mathcal{L}_\psi^{n*} m &\to \int_X hdm \int_X \varphi d\nu \\
\frac{1}{\lambda^n}\int_X \varphi dm &\to \int_X hdm \int_X \varphi d\nu
\end{aligned}$$

donde $\lambda = 1$ e segue que $\forall \varphi \in C(X)$,

$$\int_X \varphi dm = \int_X \left(\int_X hdm\right)\varphi d\nu$$

fazendo $\varphi \equiv 1$ temos

$$\begin{aligned}
\int_X dm &= \int_X hdm\nu(X) \\
\int_X hdm &= 1 \\
\text{i. e., } \int_X \varphi dm &= \int_X \varphi d\nu
\end{aligned}$$

$\forall \varphi \in C(X)$, implicando que

$$m = \nu$$

como $\mu = h\nu$ segue que

$$\mu = hm.$$

□

Afirmação 3. *Seja* $\psi(x) = -\log|\det D_xT|$. *Seja* μ *a medida dada pelo teorema de Ruelle. A entropia de* T *com relação a* μ *vale* $\int_X \log|\det D_xT|d\mu$ *e, além disso, para outra medida invariante* η, *diferente de* μ, *temos que*

$$h_\eta(T) + \int_X \log|\det D_xT|d\eta \leq 0.$$

Demonstração. Do item f) do teorema temos que

$$0 = h_\mu(T) + \int_X -\log|\det D_xT|d\mu$$
$$\Rightarrow h_\mu(T) = \int_X \log|\det D_xT|d\mu.$$

Agora de g) se η é uma medida invariante por T diferente de μ então

$$h_\eta(T) + \int_X \log|\det D_xT|d\eta \leq 0.$$

□

Afirmação 4. *Para qualquer* $A \in \mathcal{A}$ *vale que*

$$m(T^{-n}(A)) \to \mu(A)$$

onde m *é a medida de Lebesgue em* X, *com* $m(X) = 1$, *e* μ *a medida dada pelo teorema de Ruelle.*

Demonstração. Seja $\varphi \in C(X)$ e considere

$$\int_X \varphi \circ T^n dm = \int_X \varphi \circ T^n d\mathcal{L}_\psi^{n*} m = \int_X \mathcal{L}_\psi^n(\varphi \circ T^n) dm$$

vamos analisar o comportamento de $\mathcal{L}_\psi^n \varphi(x)$. Para $n = 1$ temos

$$\mathcal{L}_\psi \varphi(x) = \sum_{y \in T^{-1}(x)} \frac{\varphi(y)}{|\det D_yT|}$$

já para $n = 2$ vale que

$$\begin{aligned}
\mathcal{L}_\psi^2 \varphi(x) = \mathcal{L}_\psi(\mathcal{L}_\psi \varphi)(x) &= \sum_{y \in T^{-1}(x)} \frac{(\mathcal{L}_\psi \varphi)(y)}{|\det D_yT|} \\
&= \sum_{y \in T^{-1}(x)} \frac{1}{|\det D_yT|} \sum_{z \in T^{-1}(y)} \frac{\varphi(z)}{|\det D_zT|} \\
&= \sum_{z \in T^{-2}(x)} \frac{\varphi(z)}{|\det D_zT^2|}
\end{aligned}$$

segue, por indução, que para qualquer $n \in \mathbb{N}$

$$\mathcal{L}_\psi^n \varphi(x) = \sum_{y \in T^{-n}(x)} \frac{\varphi(y)}{|\det D_y T^n|}$$

donde

$$\mathcal{L}_\psi^n (\varphi \circ T^n)(x) = \sum_{y \in T^{-n}(x)} \frac{\varphi \circ T^n(y)}{|\det D_y T^n|} = \varphi(x) \sum_{y \in T^{-n}(x)} \frac{1}{|\det D_y T^n|}.$$

Assim,

$$\begin{aligned}
\int_X \mathcal{L}_\psi^n (\varphi \circ T^n) dm &= \int_X \varphi(x) \sum_{y \in T^{-n}(x)} \frac{1}{|\det D_y T^n|} dm \\
&= \int_X \varphi(x) (\mathcal{L}_\psi^n 1)(x) dm \\
&\to \int_X \varphi(x) h(x) dm \\
&= \int_X \varphi d\mu
\end{aligned}$$

ou seja,

$$\int_X \varphi \circ T^n dm \to \int_X \varphi d\mu, \qquad\qquad \forall \varphi \in C(X).$$

Logo, se $A \in \mathcal{A}$ então

$$\int_X \chi_A \circ T^n dm \to \int_X \chi_A d\mu$$

onde χ_A é a função característica do conjunto A, assim

$$m(T^{-n}(A)) \to \mu(A).$$

□

Afirmação 5. *Se a aplicação ψ usada no teorema de Ruelle for identicamente nula então teremos outra medida de probabilidade μ_{max} que maximiza a entropia entre as medidas invariantes.*

Demonstração. Para obtermos outra medida de probabilidade invariante toma-se

$$\psi(x) \equiv 0$$

e o operador de Ruelle-Perron-Frobenius se escreve

$$\mathcal{L}_\psi \varphi(x) = \sum_{y \in T^{-1}(x)} \varphi(y)$$

acarretando a existência de λ, h e $\nu \in \mathcal{M}(X)$ nas condições do teorema. Sendo k o grau topológico de T segue que

$$\mathcal{L}_\psi 1(x) = \sum_{y \in T^{-1}(x)} 1(y) = k = k1(x)$$

i. e., $\mathcal{L}_\psi 1 = k1$ donde tiramos que

$$\lambda = k \qquad \text{e} \qquad h = 1.$$

Além disso,

$$\frac{1}{k^n} \sum_{y \in T^{-n}(x)} \varphi(y) \to \int_X \varphi d\mu_{max}, \qquad \forall \varphi \in C(X).$$

Esta é a medida que maximiza a entropia pelo princípio variacional. Sobre princípio variacional veja [M1] ou [W2]. □

A afirmação anterior motiva a definição de entropia topológica para a aplicação $T : X \to X$ contínua no espaço compacto X, que é o supremo das entropias métricas, ou seja,

$$h_{\text{top}}(T) := \sup\{h_\mu(T); \mu \in \mathcal{M}_T(X)\}.$$

1.4 Funções homólogas

Pode acontecer que dois potenciais diferentes gerem o mesmo conjunto de estados de equilíbrio. Vamos desenvolver nesta seção um critério que garanta quando dois potenciais possuem o mesmo conjunto de estados de equilíbrio.

Definição 1.32 (Critério de Homologia). *Supondo $T : X \to X$ expansora em um espaço métrico compacto X e topologicamente mixing, dizemos que ψ e $\varphi \in C(X)$ são homólogas, $\psi \sim \varphi$, se existir uma função $u \in C(X)$ tal que*

$$\psi = \varphi + u \circ T - u.$$

Um corolário do próximo teorema apresenta um método de construção da função u quando ψ e φ são funções Hölder.

Teorema 1.9. *Suponhamos $\psi : X \to \mathbb{R}$ Hölder contínua em um espaço métrico compacto X e $T : X \to X$ expansora e topologicamente mixing. Então $\psi \sim 0$ se, e somente se,*

$$T^n(x) = x \quad \textit{implica} \quad S_n\psi(x) = \sum_{j=0}^{n-1} \psi(T^j(x)) = 0.$$

Na demonstração deste teorema usamos alguns resultados de dinâmica de transformações expansoras que podem ser encontrados em [C1]. Iremos demonstrar apenas um destes lemas, devido à sua importância, que é uma versão do Shadowing Lemma para transformações expansoras em espaços métricos.

Definição 1.33. *Sejam (X, d) um espaço métrico compacto, $T : X \to X$ uma aplicação expansora, $\varepsilon > 0$ e $\delta > 0$. Dizemos que:*

i) a sequência $\{x_n\}_{n\geq 0}$ de pontos em X é uma ε-pseudo-órbita para T se acontecer que $d(x_{n+1}, T(x_n)) < \varepsilon$, $n \geq 0$.

ii) a sequência $\{x_n\}_{n\geq 0}$ de pontos em X é uma ε-pseudo-órbita periódica de período $N > 0$ se é uma ε-pseudo-órbita e $x_{i+N} = x_i$, $\forall i \geq 0$.

iii) a sequência $\{x_n\}_{n\geq 0}$ de pontos em X é δ-sombreada pela órbita de um ponto $x \in X$ se $d(x_n, T^n(x)) < \delta$, $n \geq 0$.

Antes de enunciar e provar o Shadowing Lemma vamos lembrar que se $T : X \to X$ é expansora no espaço métrico compacto X, a função $\varphi : S \subseteq X \to X$ é ramo contrativo de T^{-n} se $T^n(\varphi(x)) = x$, $\forall x \in S$ e

$$d(T^j(\varphi(x)), T^j(\varphi(y))) \leq \alpha^{n-j} d(x, y), \qquad \forall x, y \in S, \qquad 0 \leq j \leq n$$

onde $0 < \alpha < 1$ é a constante da definição de T expansora.

Lema 1.4 (Shadowing Lemma). *Seja $T : X \to X$ expansora em espaço métrico compacto. Para todo $\varepsilon > 0$ existe $\delta > 0$ tal que se $\{x_n\}_{n\geq 0} \subseteq X$ é uma δ-pseudo-órbita então existe um único $x \in X$ cuja órbita forma uma ε-sombra sobre $\{x_n\}_{n\geq 0}$.*

Demonstração. Sejam $\varepsilon > 0$, $T : X \to X$ expansora e $r > 0, 0 < \alpha < 1$ da definição de T. Seja

$$\delta < \min\left\{\frac{1-\alpha}{\alpha}\varepsilon, \frac{r}{\alpha} - r\right\}$$

e considere a δ-pseudo-órbita $\{x_n\}_{n\geq 0}$.

Considere os ramos contrativos, $\varphi_n : B_r(x_n) \to X$, de T^{-n} com $\varphi_n(T(x_{n-1})) = x_{n-1}$. Se $z \in \overline{B_r(x_n)}$ então

$$d(z, T(x_{n-1})) \leq d(z, x_n) + d(x_n, T(x_{n-1})) < r + \delta$$
$$d(\varphi_n(z), x_{n-1}) = d(\varphi_n(z), \varphi_n(T(x_{n-1}))) < \alpha(r + \delta) = \alpha r + \alpha\delta < r$$

resultando que $\varphi_n(\overline{B_r(x_n)}) \subseteq B_r(x_{n-1})$, $\forall n \geq 1$.

Consideremos $\{\varphi_1 \circ \cdots \circ \varphi_n(\overline{B_r(x_n)})\}_{n\geq 1}$ que é uma sequência decrescente de conjuntos compactos com diâmetro tendendo a zero. Logo, existe um único ponto $x \in X$ tal que $x \in \bigcap_{n\geq 1} \varphi_1 \circ \cdots \circ \varphi_n(\overline{B_r(x_n)})$.

Seja $m \in \mathbb{N}$ qualquer, então

$$\begin{aligned} d(T^m(x), x_m) &\leq \alpha d(T^{m+1}(x), T(x_m)) \leq \alpha d(T^{m+1}(x), x_{m+1}) + \alpha d(x_{m+1}, T(x_m)) \\ &\leq \alpha^2 d(T^{m+2}(x), T(x_{m+1})) + \alpha d(x_{m+1}, T(x_m)) \leq \cdots \\ &\leq \alpha d(x_{m+1}, T(x_m)) + \alpha^2 d(x_{m+2}, T(x_{m+1})) + \cdots + \alpha^{m+j} d(x_{m+j}, T(x_{m+j})) \\ &< \delta \sum_{i=1}^{m+j} \alpha^i. \end{aligned}$$

Fazendo $j \to \infty$ obtemos

$$d(T^m(x), x_m) < \delta \frac{\alpha}{1-\alpha} < \varepsilon.$$

Quanto à unicidade sejam $x, y \in X$ pontos distintos cujas órbitas formam uma ε-sombra sobre $\{x_n\}_{n\geq 0}$. Assim, $\forall n \geq 0$ temos que

$$\begin{aligned} &d(x_n, T^n(x)) < \varepsilon \qquad \text{e} \qquad d(x_n, T^n(y)) < \varepsilon \\ &T^n(x), T^n(y) \in B(x_n, \varepsilon), \qquad \forall n \geq 0 \\ &x, y \in T^{-n}(B(x_n, \varepsilon)), \qquad \forall n \geq 0 \end{aligned}$$

e pela definição de T expansora, temos que $\operatorname{diam}(T^{-n}(B(x_n, \varepsilon))) \leq \alpha^n \varepsilon$ com $0 < \alpha < 1$ e $\varepsilon > 0$ arbitrário, provando assim que $x = y$. □

Corolário 1.4.1. *Sejam $T : X \to X$ uma transformação expansora em um espaço métrico compacto, $\varepsilon > 0$ e $\delta > 0$ dado pelo Shadowing Lemma. Seja $\{x_n\}_{n\geq 0}$ uma δ-pseudo-órbita periódica de período N. Então, existe um único ponto periódico de período N, $x \in X$, cuja órbita forma uma ε-sombra sobre $\{x_n\}_{n\geq 0}$.*

Demonstração. Sejam $\varepsilon > 0$ e $\{x_n\}_{n\geq 0}$ uma δ-pseudo-órbita periódica de período $N > 0$ para $\delta > 0$. Seja $x \in X$ o ponto cuja órbita forma uma ε-sombra sobre $\{x_n\}_{n\geq 0}$, obtido pelo Shadowing Lemma. Assim, para $0 \leq i \leq n$, $d(T^i(x), x_i) < \varepsilon$ e também $d(T^{i+kN}(x), x_i) < \varepsilon$ para $0 \leq i \leq n$ e $k > 0$. Logo, $z = T^N(x)$ também é uma ε-sombra e, por unicidade da sombra, $z = x$ mostrando que $T^N(x) = x$. □

Corolário 1.4.2. *Seja $T : X \to X$ uma transformação expansora topologicamente mixing em um espaço métrico compacto. Então $X = \overline{Per(T)}$.*

Demonstração. Sejam $x \in X$ qualquer e $z \in X$ um ponto cuja órbita é densa em X. Assim, dado $\varepsilon > 0$ seja $\delta > 0$ dado pelo Shadowing Lemma. Seja $B(x, \delta/2)$ então existem $0 < k < j$ naturais tais que $T^k(z)$ e $T^j(z) \in B(x, \delta/2)$ (use a densidade) portanto $\{T^k(z), T^{k+1}(z), \ldots, T^{j-1}(z), T^k(z), T^{k+1}(z), \ldots\}$ é uma δ-pseudo-órbita pois $d(T^k(z), T^j(z)) < \delta/2 + \delta/2 = \delta$. Logo, pelo corolário anterior, existe uma ε-sombra periódica $w \in \mathrm{Per}(T)$ e $d(w, x) \leq d(x, T^k(z)) + d(T^k(z), w) = 2\varepsilon$. $\square$

Obs.: Se $x_{j+1} = T(x_j)$, $0 \leq j \leq N-2$ e $d(T(x_{N-1}), x_0) < \delta$ resulta que, para $0 \leq j \leq N-1$,

$$d(T^j(x), T^j(x_0)) = d(T^j(x), x_j) \leq \alpha^{N-j} d(T^N(x), T(x_{N-1}))$$

veja [C1].

Agora vamos provar o teorema 1.9.

Demonstração. [$\Rightarrow$] Se $\psi \sim 0$ então existe uma função $u \in C(X)$ tal que $\psi = u \circ T - u$. Seja $x \in X$ tal que $T^n(x) = x$, então

$$\begin{aligned} S_n\psi(x) &= \sum_{j=0}^{n-1} \psi(T^j(x)) = \sum_{j=0}^{n-1} \big(u \circ T(T^j(x)) - u(T^j(x))\big) \\ &= \sum_{j=0}^{n-1} \big(u(T^{j+1}(x) - u(T^j(x))\big) = u(T^n(x)) - u(x) \end{aligned}$$

implicando que $S_n\psi(x) = 0$.

[$\Leftarrow$] Seja $S_n\psi(x) = 0$ para todo $x \in X$ tal que $T^n(x) = x$. Como T é topologicamente mixing também é transitiva [M1], ou seja, existe um $a \in X$ tal que a órbita de a é densa em X. Agora, definimos u na órbita de a como

$$u(T^n(a)) = u(a) + S_n\psi(a)$$

onde $u(a)$ é um valor qualquer.

Para provar a continuidade uniforme sejam $\delta > 0$ e $d(T^m(a), T^{m+n}(a)) < \delta$, e podemos tomar a δ-pseudo-órbita

$$T^m(a), T^{m+1}(a), \ldots, T^{m+n-1}(a), T^m(a), \ldots.$$

Dado $\varepsilon > 0$, seja δ tal que existe x periódico de período n cuja órbita ε-sombreia a δ-pseudo-órbita acima. E temos que $d(T^j(x), T^{m+j}(a)) \leq \alpha^{n-j} d(x, T^{m+n}(a))$, $0 \leq j \leq n-1$. Assim, se $C > 0$ e $\gamma > 0$ são as constantes da definição de ψ Hölder

$$\begin{aligned}
|u(T^{m+n}(a)) - u(T^n(a))| &= |S_{m+n}\psi(a) - S_m\psi(a)| = |S_n\psi(T^m(a))| \\
&= |S_n\psi(T^m(a)) - S_n\psi(x)| \leq \sum_{j=0}^{n-1} |\psi(T^{m+j}(a)) - \psi(T^j(x))| \\
&\leq C\sum_{j=0}^{n-1} d(T^{m+j}(a), T^j(x))^\gamma \leq C\sum_{j=0}^{n-1} (\alpha^{n-j}\varepsilon)^\gamma \leq C_1\varepsilon^\gamma
\end{aligned}$$

o que prova a continuidade uniforme de u. Assim, podemos estender u a uma função contínua em X. Seja $y \in X$ com $y = \lim_{j\to\infty} T^{n_j}(a)$. Então,

$$\begin{aligned}
u(T(y)) - u(y) &= \lim_{j\to\infty} u(T^{n_j+1}(a)) - \lim_{j\to\infty} u(T^{n_j}(a)) \\
&= \lim_{j\to\infty} S_{n_j+1}\psi(a) - S_{n_j}\psi(a) = \lim_{j\to\infty} \psi(T^{n_j}(a)) \\
&= \psi(y).
\end{aligned}$$

Logo, $\psi = u \circ T - u$. □

Corolário 1.9.1. *Sejam ψ e φ Hölder contínuas definidas no espaço métrico X. Então $\psi \sim \varphi$ se, e somente se, $T^n(x) = x$ implica que $S_n\psi(x) = S_n\varphi(x)$.*

Demonstração. É uma consequência imediata do teorema 1.9. □

Teorema 1.10. *Sejam $\psi, \varphi : X \to \mathbb{R}$ funções homólogas e $T : X \to X$ expansora em espaço métrico compacto. Então $E_\psi = E_\varphi$ onde*

$$E_\varphi = \left\{\mu \in \mathcal{M}_T(X); h_\mu(T) + \int_X \varphi d\mu = \sup_{\nu\in\mathcal{M}_T(X)} \left\{h_\nu(T) + \int_X \varphi d\nu\right\}\right\}$$

é o conjunto dos estados de equilíbrio para φ.

Demonstração. Se $\mu \in \mathcal{M}_T(X)$ então

$$\int_X \psi d\mu = \int_X (\varphi + u\circ T - u) d\mu = \int_X \varphi d\mu$$

donde

$$\int_X \psi d\mu = \int_X \varphi d\mu.$$

Assim, se $\mu \in E_\psi$, $h_\mu(T) + \int_X \varphi d\mu = h_\mu(T) + \int_X \psi d\mu$ e portanto $\mu \in E_\varphi$. □

Obs.: O teorema acima é valido para a seguinte generalização de homologia. As funções $\varphi, \psi \in C(X)$ são homólogas com constante c se existe uma aplicação $u \in C(X)$ tal que para algum $c \in \mathbb{R}$ vale $\varphi = \psi + u \circ T - u + c$.

Lema 1.5. *Seja $T : X \to X$ expansora em um espaço métrico compacto. Consideremos $\varphi, \psi : X \to \mathbb{R}$ Hölder contínuas tal que $\psi \sim \varphi$. Então,*

$$\mu_\psi = h\nu_\psi \qquad e \qquad \mu_\varphi = e^{-u}h\nu_\varphi$$

onde μ_ψ e μ_φ denotam o estado de equilíbrio para ψ e φ, respectivamente, e h é a função dada pelo teorema de Ruelle.

Demonstração. Sejam $T : X \to X$ e ψ satisfazendo as hipóteses do teorema de Ruelle e suponha que $\psi \sim \varphi$. Assim, para qualquer $f \in C(X)$, e $x \in X$

$$\begin{aligned}\mathcal{L}_\psi(f)(x) &= \sum_{y \in T^{-1}(x)} e^{\psi(y)} f(y) = \sum_{y \in T^{-1}(x)} e^{\varphi(y) + u \circ T(y) - u(y)} f(y) \\ &= e^{u(x)} \sum_{y \in T^{-1}(x)} e^{\varphi(y)} f(y) e^{-u(y)}.\end{aligned}$$

Sejam λ e h autovalor e autovetor, respectivamente, do operador de Ruelle-Perron-Frobenius dados pelo teorema de Ruelle. Assim, $\mathcal{L}_\psi(h) = \lambda h$ e

$$\mathcal{L}_\psi(h)(x) = \sum_{y \in T^{-1}(x)} e^{\psi(y)} h(y) = \left(\sum_{y \in T^{-1}(x)} e^{\varphi(y)} h(y) e^{-u(y)} \right) e^{u(x)} = \lambda h(x)$$

para $\mathcal{L}_\varphi(e^{-u}h)$ temos, usando que $\varphi = \psi - u \circ T + u$

$$\begin{aligned}\mathcal{L}_\varphi(e^{-u}h)(x) &= \sum_{y \in T^{-1}(x)} e^{\varphi(y)} e^{-u(y)} h(y) = e^{-u(x)} \sum_{y \in T^{-1}(x)} e^{\psi(y)} h(y) \\ &= e^{-u(x)} \mathcal{L}_\psi(h)(x) = \lambda e^{-u} h(x).\end{aligned}$$

Isto é, se $\mathcal{L}_\psi(h) = \lambda h$ então

$$\mathcal{L}_\varphi(e^{-u}h) = \lambda e^{-u} h$$

e, portanto, $e^{-u}h$ é o autovetor associado ao autovalor λ. Assim, se μ_ψ (μ_φ) denota o estado de equilíbrio para ψ (φ) temos que

$$\begin{aligned}\mu_\psi &= h\nu_\psi \\ \mu_\varphi &= e^{-u}h\nu_\varphi.\end{aligned}$$

$\square$

Capítulo 2

Jacobiano

Estritamente relacionado ao operador de Ruelle-Perron-frobenius está a idéia de Jacobiano de uma medida relativamente a uma transformação localmente injetiva. Esta noção é claramente uma extensão da noção usual de Jacobiano de um difeomorfismo com respeito ao volume no $\mathbb{R}^n$. Neste capítulo iremos definir o Jacobiano de uma medida e calculá-lo para alguns casos. Também provaremos sua existência e o teorema da mudança de variável e uma variação do teorema da função inversa, ambos utilizando essa noção de Jacobiano. No próximo capítulo iremos explicitar a importância do Jacobiano.

2.1 Definição e existência do Jacobiano para medidas

No núcleo da nossa apresentação sobre o assunto desta seção está o conceito de Jacobiano de uma transformação relativamente a uma medida, não necessariamente invariante por esta transformação. Sendo esta noção uma clara generalização do conceito de Jacobiano para um difeomorfismo de um conjunto aberto do $\mathbb{R}^n$, com respeito à medida de Lebesgue, vamos primeiramente recordar esta definição.

Para uma aplicação diferenciável que é localmente um difeomorfismo, a seguinte identidade é bem conhecida de um curso de análise

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

onde $f : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ é um difeomorfismo local C^1, m é a medida de Lebesgue em

$\mathbb{R}^n$, e A é um conjunto de Borel tal que $f|_A$ é injetiva. A função

$$J_m(f) : U \to \mathbb{R}$$

definida por $J_m(f) = |\det Df(x)|$ é usualmente chamada de determinante do Jacobiano de f no ponto x. Vamos estender esta definição de Jacobiano para isto, seja a seguinte definição.

Definição 2.1. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de medida e $p \geq 1$. Denota-se $L^p(X, \mathcal{A}, \mu)$, ou $L^p(\mu)$ o conjunto das funções $f : X \to \mathbb{R}$ tais que $|f|^p$ é integrável, identificando funções que coincidem em q.t.p., i. e., a menos de um conjunto de medida nula. Em $L^p(\mu)$ definimos a norma $\|\cdot\|_p$ por*

$$\|f\|_p = \left(\int_X |f|^p d\mu \right)^{1/p}.$$

Se $p = \infty$, define-se $L^\infty(X, \mathcal{A}, \mu)$, ou $L^\infty(\mu)$ como o conjunto das funções $f : X \to \mathbb{R}$ tais que existe $K > 0$ onde $|f(x)| \leq K$ para q.t.p. $x \in X$, identificando funções que coincidem em q.t.p. O ínfimo dos K com esta propriedade denota-se $\|f\|_\infty$ e define uma norma em $L^\infty(\mu)$.

Definição 2.2. *Seja $T : X \to X$ uma função contínua localmente injetiva definida em um espaço métrico compacto X, e seja μ uma medida sobre os borelianos em X. Uma função $J_\mu \in L^1(\mu)$ onde*

$$J_\mu(T) : X \to \mathbb{R}^+$$

é um Jacobiano para T, com respeito à medida μ, se vale

$$\mu(T(A)) = \int_A J_\mu(T) d\mu$$

para qualquer boreliano $A \subseteq X$ tal que $T|_A$ é injetiva.

Obs.: Quando um Jacobiano existe ele deve ser único, μ quase sempre.

Definição 2.3. *Seja $(X, \mathcal{A}, \mu)$ um espaço de medida. Dizemos que $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ é duplamente mensurável se para todo $A \in \mathcal{A}$ vale que*

$$T(A) \in \mathcal{A} \qquad e \qquad T^{-1}(A) \in \mathcal{A}.$$

Definição 2.4. *Sejam $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ duplamente mensurável e μ uma medida σ-finita, não necessariamente T-invariante. Dizemos que T é μ absolutamente contínua (para frente) se $\forall A \in \mathcal{A}$ com $\mu(A) = 0$ tivermos que $\mu(T(A)) = 0$.*

Para a prova da existência do Jacobiano, para transformações duplamente mensuráveis, vamos precisar do seguinte lema.

Lema 2.1. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de medida e $T : X \to X$ localmente injetiva em um espaço métrico separável. Para todo $A \in \mathcal{A}$ tal que $T|_A$ é injetiva definimos*

$$\nu(A) = \mu(T(A)).$$

Então, ν é uma medida sobre $\mathcal{A}$.

Demonstração. Seja X um espaço métrico separável. Então existe um conjunto denso e enumerável em X, seja $S = \{x_1, x_2, x_3, \ldots\}$ esse conjunto.

Como $T : X \to X$ é localmente injetiva, para cada $x \in X$ existe $\varepsilon_x > 0$ tal que $T|_{B_{\varepsilon_x}(x)}$ é injetiva. Note que $X = \bigcup_{x \in X} B_{\varepsilon_x}(x)$.

Sejam, para todo $i \geq 1$, $B_i = B_{\varepsilon_x}(x)$ tal que $x_i \in B_{\varepsilon_x}(x)$ para todo $x_i \in S$. Assim, temos que $\{B_i\}_i$ é uma família enumerável de bolas abertas que cobrem X. Agora, consideremos os conjuntos

$$\begin{aligned} A_1 &= B_1 \\ A_2 &= B_2 - A_1 \\ &\vdots \\ A_n &= B_n - \bigcup_{j=1}^{n-1} A_j \\ &\vdots \end{aligned}$$

Assim, $\{A_j\}_j$ é uma família de borelianos disjuntos cuja união é X e tal que $T|_{A_j}$, para todo $j \geq 1$, é injetiva.

Se $B \in \mathcal{A}$ então $B = \bigcup_{j=1}^{\infty} (B \cap A_j)$. Agora definimos $\nu : \mathcal{A} \to [0, +\infty]$ por

$$\nu(B) = \sum_{j=1}^{\infty} \mu(T(B \cap A_j))$$

para todo $B \in \mathcal{A}$. Vamos provar que ν é uma medida.

Inicialmente é trivial que $\nu(\emptyset) = 0$. Caso $B = \bigcup_{j=1}^{\infty}(B_i) \in \mathcal{A}$ é uma união de conjuntos disjuntos, então

$$\begin{aligned}\nu(B) = \nu\left(\bigcup_{i=1}^{\infty}(B_i)\right) &= \sum_{j=1}^{\infty}\mu\left(T\left(\bigcup_{i=1}^{\infty}(B_i)\cap A_j\right)\right)\\ &= \sum_{j=1}^{\infty}\mu\left(\bigcup_{i=1}^{\infty}T(B_i\cap A_j)\right)\end{aligned}$$

caso $\{T(B_i \cap A_j)\}_i$ não for uma família de conjuntos dois a dois disjuntos, deve existir um ponto $y \in X$ tal que para $i \neq k$, quaisquer

$$y \in T(B_i \cap A_j) \qquad \text{e} \qquad y \in T(B_k \cap A_j)$$

donde devem existir $x_{ij} \in B_i \cap A_j$ e $x_{kj} \in B_k \cap A_j$ tais que $y = T(x_{ij})$ e $y = T(x_{kj})$. Como $\{B_i\}_i$ é uma família de subconjuntos disjuntos, temos que $x_{ij} \neq x_{kj}$. Mas, $x_{ij} \in A_j$ e $x_{kj} \in A_j$ e sendo $T|_{A_j}$ injetora temos uma contradição. Logo, $\{T(B_i \cap A_j)\}_i$ é uma família de conjuntos dois a dois disjuntos. Assim,

$$\begin{aligned}\nu(B) = \sum_{j=1}^{\infty}\mu\left(\bigcup_{i=1}^{\infty}T(B_i\cap A_j)\right) &= \sum_{j=1}^{\infty}\sum_{i=1}^{\infty}\mu\left(T(B_i\cap A_j)\right)\\ &= \sum_{i=1}^{\infty}\nu(B_i).\end{aligned}$$

Provando assim, que ν é uma medida sobre $\mathcal{A}$. $\square$

Teorema 2.1. *Seja $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ uma aplicação duplamente mensurável e localmente injetiva em um espaço métrico compacto. Seja μ uma medida de probabilidade. Suponha que T seja μ absolutamente contínua (para frente). Então, existe o Jacobiano de T com respeito à medida μ.*

Demonstração. Para todo $A \in \mathcal{A}$ tal que $T|_A$ é injetiva seja, pelo lema anterior, a medida $\nu : \mathcal{A} \to [0, +\infty]$ dada por

$$\nu(A) = \mu(T(A))$$

que é finita sobre $\mathcal{A}$.

Assim, se $A \in \mathcal{A}$ é tal que $\mu(A) = 0$ então $\nu(A) = \mu(T(A)) = 0$, por hipótese. Logo, $\nu \ll \mu$ e pelo teorema de Radon-Nikodym [DS] existe uma função

$f \in L^1(\mu)$ tal que para todo $A \in \mathcal{A}$,

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

Chamamos $J_\mu(T) = f$ o Jacobiano de T com respeito à medida μ. $\square$

2.2 Propriedades do Jacobiano

Lema 2.2. *Sejam $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ uma função contínua localmente injetiva definida em um espaço métrico X e μ uma medida σ-finita e T-invariante então $J_\mu(T)$ satisfaz*

$$\sum_{y \in T^{-1}(x)} \frac{1}{J_\mu(T)(y)} = 1, \qquad \mu \text{ q.s.}, \forall x$$

sempre que $J_\mu(T) > 0$.

Demonstração. Pelo teorema de Radon-Nikodym, veja [DS]

$$\lim_{\varepsilon \to 0} \frac{\mu(T(B_\varepsilon(x)))}{\mu(B_\varepsilon(x))} = J_\mu(T)(x), \qquad \mu \text{ q.s. }, \forall x.$$

Seja $T^{-1}(x) = \{x_1, \ldots, x_n\}$ e tome $\varepsilon > 0$ tão pequeno que

$$T^{-1}(B_\varepsilon(x)) = \bigcup_{i=1}^{n} A_i$$

onde $x_i \in A_i$ e $A_i \cap A_j = \emptyset$ para $i \neq j$, $1 \leq i, j \leq n$. Então,

$$\begin{aligned}
\sum_{i=1}^{n} \frac{1}{J_\mu(T)(x_i)} &= \lim_{\varepsilon \to 0} \sum_{i=1}^{n} \frac{\mu(A_i)}{\mu(T(A_i))} = \lim_{\varepsilon \to 0} \sum_{i=1}^{n} \frac{\mu(A_i)}{\mu(B_\varepsilon(x))} \\
&= \lim_{\varepsilon \to 0} \frac{1}{\mu(B_\varepsilon(x))} \sum_{i=1}^{n} \mu(A_i) = \lim_{\varepsilon \to 0} \frac{\mu(T^{-1}(B_\varepsilon(x)))}{\mu(B_\varepsilon(x))} \\
&= 1
\end{aligned}$$

pois $\mu(T^{-1}(B_\varepsilon(x))) = \mu(B_\varepsilon(x))$ porque μ é T-invariante. $\square$

Vejamos como o operador de Ruelle-Perron-Frobenius se relaciona com a noção de Jacobiano.

Definição 2.5. *Seja $T : X \to \mathbb{R}$ uma função definida em um espaço métrico. O suporte de T é o conjunto*

$$supp(T) = \overline{\{x \in X; T(x) \neq 0\}}.$$

Caso $supp(T)$ seja um conjunto compacto, dizemos que T tem suporte compacto.

Lema 2.3. *Seja $T : X \to X$ expansora em espaço métrico compacto. Seja ν a medida obtida pelo teorema de Ruelle para $\psi : X \to \mathbb{R}$ Hölder contínua onde*

$$\mathcal{L}_\psi^*\nu = \lambda\nu, \qquad\qquad \lambda > 0.$$

Então,

$$J_\nu(T) = \lambda e^{-\psi}.$$

Demonstração. Seja A um boreliano de X tal que $A \subseteq S$ onde $\varphi : S \to X$ é um ramo da inversa de T. Considere uma sequência $\varphi_n : X \to \mathbb{R}$, $\varphi_n \in C(X)$ com suporte compacto em A e tal que o suporte não intersecta os outros ramos da inversa. Suponha, ainda, que

$$\begin{aligned} &\varphi_n \to \chi_A \qquad \text{em } L^1, q.t.p. \\ &\|\varphi_n\|_C < 2. \end{aligned}$$

Então, pelo teorema da Convergência Dominada de Lebesgue [DS]

$$\int_X \lambda e^{-\psi}\chi_A d\nu = \lim_{n\to\infty}\int_X \lambda e^{-\psi}\varphi_n d\nu$$

mas

$$\begin{aligned} \int_X \lambda e^{-\psi}\varphi_n d\nu &= \int_X \lambda e^{-\psi}\varphi_n \frac{1}{\lambda} d\mathcal{L}_\psi^*\nu = \int_X e^{-\psi}\varphi_n d\mathcal{L}_\psi^*\nu \\ &= \int_X \mathcal{L}_\psi(e^{-\psi}\varphi_n)d\nu \end{aligned}$$

$$\mathcal{L}_\psi(e^{-\psi}\varphi_n)(x) = \sum_{y\in T^{-1}(x)} e^{-\psi(y)}e^{\psi(y)}\varphi_n(y) = \sum_{y\in T^{-1}(x)}\varphi_n(y).$$

Afirmação: definindo $g_n(x) = \sum_{y\in T^{-1}(x)}\varphi_n(y)$ segue que $g_n(x) \to \chi_{T(A)}$.

De fato, se $x \in T(A)$ e como $T|_A$ é injetiva, existe um único $y_* \in A$ tal que $T(y_*) = x$ e observe que $\varphi_n(z) = 0$ para todo $z \in T^{-1}(x) \cap A^c$. Logo,

$$\sum_{y\in T^{-1}(x)}\varphi_n(y) = \varphi_n(y_*) \to 1, \qquad\quad \nu \text{ q. s.}$$

$$\Rightarrow g_n(x) \to 1.$$

Se $x \notin T(A)$ então $\forall y \in T^{-1}(x)$, $y \notin A$ e portanto

$$g_n(x) \to 0 \qquad\qquad \nu \text{ q. s.}$$

isto é,

$$g_n \to \chi_{T(A)}.$$

Segue que

$$\begin{aligned}
\int_A \lambda e^{-\psi} d\nu &= \lim_{n\to\infty} \int_X \mathcal{L}_\psi(e^{-\psi}\varphi_n) d\nu = \lim_{n\to\infty} \int_X \sum_{y \in T^{-1}(x)} \varphi_n(y) d\nu(x) \\
&= \lim_{n\to\infty} \int_X g_n(x) d\nu(x) = \int_X \chi_{T(A)} d\nu \\
&= \nu(T(A)),
\end{aligned}$$

donde, pela unicidade q.t.p. do jacobiano,

$$J_\nu(T) = \lambda e^{-\psi}.$$

□

Corolário 2.3.1. *Seja $T : X \to X$ expansora em espaço métrico compacto. Se $\mu = h\nu$ dado pelo teorema de Ruelle então*

$$J_\mu(T) = \lambda e^{-\psi} \frac{h \circ T}{h}.$$

Demonstração. Idêntica a do lema anterior. □

O próximo lema servirá de motivação para apresentar o lema da distorção limitada. O lema da distorção, além de mostrar que a distorção de iteradas da T e dos ramos da inversa são limitadas, auxilia na definição da classe de Schwarz, que enunciaremos no próximo capítulo.

Lema 2.4. *Seja $T : X \to X$ uma transformação expansora em um espaço métrico compacto. Se $J_\mu(T) > 0$ é Hölder contínua então existe $A > 0$ tal que para todo n, se $\varphi : S \subseteq X \to X$ é ramo contrativo de T^{-n} então para quaisquer $x, y \in \varphi(S)$,*

$$\frac{J_\mu(T^n)(x)}{J_\mu(T^n)(y)} \leq A.$$

Demonstração. [C1] □

Lema 2.5 (Distorção). *Nas hipóteses do lema anterior, existe* $B > 0$ *tal que para* S_1, S_2 *subconjuntos de* S *e* $\varphi : S \subseteq X \to X$ *ramo da inversa de* T^n *temos que*

$$\frac{1}{B}\frac{\mu(\varphi(S_1))}{\mu(\varphi(S_2))} \leq \frac{\mu(S_1)}{\mu(S_2)} \leq B\frac{\mu(\varphi(S_1))}{\mu(\varphi(S_2))}.$$

Demonstração. A idéia da prova é, fixado $x \in \varphi(S)$, temos

$$\mu(S_1) = \int_{\varphi(S_1)} J_\mu(T^n)d\mu \leq AJ_\mu(T^n)(x)\mu(\varphi(S_1))$$

$$\mu(S_2) = \int_{\varphi(S_2)} J_\mu(T^n)d\mu \geq \frac{1}{A}J_\mu(T^n)(x)\mu(\varphi(S_2))$$

e o resultado sai da manipulação entre essas duas desigualdades [C1]. □

Lema 2.6. *Se* $T : X \to X$ *é uma aplicação expansora em um espaço métrico compacto e topologicamente mixing com* $J_\mu(T) > 0$, *então* μ *é positiva sobre abertos.*

Demonstração. Assuma, por contradição, que existe um conjunto aberto $U \subseteq X$ tal que $\mu(U) = 0$. Por indução, e usando a hipótese $J_\mu(T) > 0$, podemos provar que $\mu(T^n(U)) = 0$ para qualquer $n > 0$. Mas sabemos que para algum $M > 0$

$$T^M(U) = X$$

e então $\mu(X) = 0$ absurdo. Logo, μ é positiva sobre abertos. □

Obs.: Uma consequência do lema 2.6 é que o suporte da medida μ é todo o espaço X, veja [C1].

Lema 2.7 (Fórmula da Mudança de Variável). *Sejam* $T : X \to X$ *uma aplicação contínua definida em um espaço métrico, localmente injetiva e* μ *uma medida de Borel. Se* $J_\mu(T) : X \to \mathbb{R}$ *é o Jacobiano para* T*, com respeito à medida* μ*, e* $J_\mu(T) \in L^\infty(\mu)$ *então para qualquer* $f \in L^1(\mu)$ *vale*

$$\int_{T(A)} f d\mu = \int_A (f \circ T)J_\mu(T)d\mu$$

sempre que A *é um conjunto boreliano tal que* $T|_A$ *é injetiva.*

Demonstração. Sabemos que

$$\mu(T(A)) = \int_A J_\mu(T) d\mu$$

assim, se $f = \chi_B$ para algum conjunto boreliano B então

$$\begin{aligned}\int_{T(A)} \chi_B d\mu = \mu(B \cap T(A)) &= \mu(T(A \cap T^{-1}(B))) \\ &= \int_{A \cap T^{-1}(B)} J_\mu(T) d\mu = \int_A J_\mu(T) \chi_{T^{-1}(B)} d\mu \\ &= \int_A (\chi_B \circ T) J_\mu(T) d\mu\end{aligned}$$

isto é,

$$\int_{T(A)} \chi_B d\mu = \int_A (\chi_B \circ T) J_\mu(T) d\mu$$

para qualquer boreliano A tal que $T|_A$ seja injetivo.

Como este resultado é válido para qualquer função simples, também será válido para qualquer $f \in L^1(\mu)$. $\square$

Lema 2.8. *Sejam $T : X \to X$ uma aplicação expansora definida em um espaço métrico compacto e μ uma probabilidade a Borel. Se $J_\mu(T) : X \to \mathbb{R}$ é o Jacobiano para T, com respeito à medida μ, e $J_\mu(T) > 0$ então*

$$J_\mu(T)(x_j) = \frac{1}{J_\mu(\varphi_j)(x)}, \qquad \forall x \in X$$

onde $x_j = \varphi_j(x)$, e φ_j é um ramo da inversa de T.

Demonstração. Sejam $\varphi_j : S_j \to X$ um ramo da inversa de T e $L : S_j \to X$ dada por $L = T \circ \varphi_j$ que é a identidade em S_j. Assim, para todo boreliano A em S_j temos que $\mu(L(A)) = \mu(A)$ e pela definição de Jacobiano temos

$$\int_A 1 d\mu = \mu(A) = \int_A J_\mu(L) d\mu$$

donde $J_\mu(L) = 1$ para μ q.t.p.

Para terminar a demonstração basta provar que para todo $x \in X$ vale

$$J_\mu(T)(x_j) J_\mu(\varphi_j)(x) = 1.$$

Para isto vamos verificar que

$$J_\mu(T \circ \varphi_j) = (J_\mu(T) \circ \varphi_j) J_\mu(\varphi_j).$$

Assim, seja A um boreliano qualquer em S_j então

$$\mu(T \circ \varphi_j(A)) = \int_A J_\mu(T \circ \varphi_j) d\mu$$

por outro lado, usando o teorema anterior temos que

$$\mu(T \circ \varphi_j(A)) = \int_{\varphi_j(A)} J_\mu(T) d\mu = \int_A (J_\mu(T) \circ \varphi_j) J_\mu(\varphi_j) d\mu.$$

Logo, para quase todo ponto temos que

$$J_\mu(T \circ \varphi_j) = (J_\mu(T) \circ \varphi_j) J_\mu(\varphi_j).$$

Assim, como φ_j é qualquer e pela unicidade do Jacobiano segue o resultado. $\square$

Capítulo 3

g-medidas

Neste capítulo além de explicitarmos a importância do Jacobiano, são dadas definições e teoremas importantes como a representação integral do operador de Ruelle seguindo Lasota e Yorke. Também definimos a classe de Schwarz que será utilizada no teorema de Ruelle no próximo capítulo. No final desta seção será demonstrado o teorema de Ledrappier que caracteriza os estados de equilíbrio para potenciais na classe das funções $\mathcal{G}$. Por fim provaremos um teorema de existência e unicidade de g-medidas quando g está em $\mathcal{G}$ e na classe de Schwarz.

3.1 Representação integral do operador de Ruelle-Perron-Frobenius

No que segue, vamos mostrar que o procedimento geral para a contrução do operador de Ruelle pode ser visto como um mecanismo de achar uma medida associada a um Jacobiano pré-conhecido. Em outras palavras, começando com uma função dada tentamos achar uma medida cujo Jacobiano é esta função. Para ser um Jacobiano a condição que a função deve satisfazer é exatamente o que define a classe das g-funções como apresentada por Peter Walters em [W1].

O operador de Ruelle-Perron-Frobenius pode ser representado de várias maneiras, vamos definir e unificar duas dessas representações. A forma mais comum do operador de Ruelle-Perron-Frobenius, que chamaremos de definição clássica, é a seguinte:

Sejam $T : X \to X$ uma transformação expansora em um espaço métrico compacto e $\psi : X \to \mathbb{R}$ uma função contínua. O operador de Ruelle-Perron-Frobenius $\mathcal{L}_\psi : C(X) \to C(X)$ é dado por

$$\mathcal{L}_\psi(f)(x) = \sum_{y \in T^{-1}(x)} e^{\psi(y)} f(y)$$

para $f \in C(X)$ e $x \in X$.

Nos restringiremos ao caso $X = M$ uma variedade compacta, T uma aplicação expansora C^1 diferenciável,

$$\psi(x) = -\log|\det D_x T|$$

e m a medida de Lebesgue em M. Uma propriedade interessante para este caso é que para $f, g \in L^2(m)$

$$\int_M \mathcal{L}_\psi(f) g dm = \int_M f U_T(g) dm$$

onde $U_T : L^2(m) \to L^2(m)$ é o operador

$$U_T(f) = f \circ T.$$

Assim, se olharmos para ambos os operadores agindo em $L^2(m)$ teremos

$$< \mathcal{L}_\psi(f), g >=< f, U_T(g) >$$

mostrando que $\mathcal{L}_\psi = U_T^t$. Podemos facilmente provar esta propriedade usando o argumento na prova de $\mathcal{L}^* m = m$ dado no primeiro capítulo.

Alguns autores usam esta forma integral para definir o operador de Ruelle-Perron-frobenius no contexto acima especificado, citamos [Kl] e [AB]. Assumindo que T preserva uma medida de probabilidade μ que é invariante e absolutamente contínua com respeito a medida de Lebesgue m, eles escrevem $\mathcal{L}^\mu$ para o operador de Ruelle-Perron-Frobenius com respeito a essa medida. Esse operador satisfaz

$$\int_M \mathcal{L}^\mu(f) g d\mu = \int_M f U_T(g) d\mu$$

para todas as $f \in L^1(\mu)$ e $g \in L^\infty(\mu)$ normalizadas, i. e., $\mathcal{L}^\mu 1 = 1$. Se μ tem densidade h, ou seja, $\mu = hm$ onde $h : X \to \mathbb{R}$ é uma função contínua então $\mathcal{L}^\mu$ e $\mathcal{L}^m$ são relacionadas pela equação $\mathcal{L}^{\mu}(f) = \frac{\mathcal{L}^{m}(hf)}{h}$. De fato,

$$\begin{aligned}
\int_M \mathcal{L}^{\mu}(f) g d\mu &= \int_M f U_T(g) d\mu = \int_M f U_T(g) h dm \\
&= \int_M \mathcal{L}^{m}(fh) g dm = \int_M \mathcal{L}^{m}(fh) \frac{g}{h} h dm \\
&= \int_M \mathcal{L}^{m}(fh) \frac{g}{h} d\mu \\
&\Rightarrow \mathcal{L}^{\mu}(f) = \frac{\mathcal{L}^{m}(fh)}{h}.
\end{aligned}$$

Uma bela representação integral do operador de Ruelle-Perron-Frobenius é apresentada por Lasota e Yorke no artigo [LY1]. Vamos seguir essa representação no que segue.

Definição 3.1. *Seja $(X, \mathcal{A}, \mu)$ um espaço de probabilidade. Seja $T : X \to X$ uma transformação duplamente mensurável. Dizemos que T é não-singular, ou que é μ absolutamente contínua (para trás), se*

$$\mu(T^{-1}(A)) = 0, \qquad \textit{sempre que} \qquad \mu(A) = 0.$$

Definição 3.2. *Se $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ é uma transformação duplamente mensurável e não-singular em um espaço de medida, com μ uma medida σ-finita. Definimos o operador de Ruelle-Perron-Frobenius $P : L^1(\mu) \to L^1(\mu)$ por Lasota e Yorke como*

$$\int_A P(f) d\mu = \int_{T^{-1}(A)} f d\mu$$

para quaisquer $A \in \mathcal{A}$ e $f \in L^1(\mu)$.

Dado $f \in L^1(\mu)$ considere a medida ν dada por

$$\nu(A) = \int_{T^{-1}(A)} f d\mu.$$

Se $A \in \mathcal{A}$ e $\mu(A) = 0$ então, como T é não-singular, $\mu(T^{-1}(A)) = 0$ e assim

$$\nu(A) = \int_{T^{-1}(A)} f d\mu = 0$$

mostrando que $\nu \ll \mu$. Logo, pelo teorema de Radon-Nikodym [DS] temos que

$$P(f) = \frac{d\nu}{d\mu}$$

onde $d/d\mu$ denota a derivada de Radon-Nikodym. Este operador é claramente linear, preserva a integral e é uma contração em $L^1(\mu)$, a saber,

$$\|P(f)\|_{L^1} \leq \|f\|_{L^1}.$$

Vamos ver como as representações clássica e por Lasota e Yorke, do operador de Ruelle-Perron-Frobenius, estão relacionadas.

Lema 3.1. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de medida, com μ uma medida σ-finita, e $P : L^1(\mu) \to L^1(\mu)$ o operador de Ruelle-Perron-Frobenius, de acordo com a definição de Lasota e Yorke, e assuma que $T : X \to X$ é localmente injetiva e admite um Jacobiano $J_\mu(T) > 0$. Se $\mathcal{L}$ é o operador de Ruelle-Perron-Frobenius, de acordo com a definição clássica, então $P(f)(x) = \mathcal{L}_{-\log J_\mu(T)}(f)(x)$ para toda $f \in L^1(\mu)$.*

Demonstração. Se $f \in L^1(\mu)$ temos que

$$\begin{aligned}
P(f)(x) &= \lim_{\varepsilon\to 0} \frac{\nu(B_\varepsilon(x))}{\mu(B_\varepsilon(x))} = \lim_{\varepsilon\to 0} \frac{\int_{T^{-1}(B_\varepsilon(x))} f d\mu}{\mu(B_\varepsilon(x))} \\
&= \lim_{\varepsilon\to 0} \frac{\sum_j \int_{A_j} f d\mu}{\mu(B_\varepsilon(x))} = \lim_{\varepsilon\to 0} \frac{\sum_j \int_{\varphi_j(B_\varepsilon(x))} f d\mu}{\mu(B_\varepsilon(x))} \\
&= \lim_{\varepsilon\to 0} \frac{\sum_j \int_{B_\varepsilon(x)} (f\circ\varphi_j) J_\mu(\varphi_j) d\mu}{\mu(B_\varepsilon(x))} \\
&= \sum_j \lim_{\varepsilon\to 0} \frac{\int_{B_\varepsilon(x)} (f\circ\varphi_j) J_\mu(\varphi_j) d\mu}{\mu(B_\varepsilon(x))} \\
&= \sum_j (f\circ\varphi_j)(x) J_\mu(\varphi_j)(x) \\
&= \sum_{y\in T^{-1}(x)} \frac{f(y)}{J_\mu(T)(y)}, \qquad \mu \text{ q.s., } x \in X
\end{aligned}$$

onde $\varphi_j : B_\varepsilon(x) \to X$ são os ramos da inversa de T e usamos também a definição de diferencial de uma medida [Ru] e os lemas 2.7 e 2.8.

Assim, temos que $\forall f \in L^1(\mu)$

$$P(f)(x) = \sum_{y\in T^{-1}(x)} \frac{f(y)}{J_\mu(T)(y)} = \mathcal{L}_{-\log J_\mu(T)}(f)(x).$$

$\square$

É importante notar que, na definição do operador de Ruelle-Perron-Frobenius, P, não estamos exigindo que μ seja T-invariante. Por este fato necessitamos da condição de não-singularidade. Quando μ é T-invariante a condição de não-singularidade é trivial e pode ser omitida.

3.2 O teorema de Ledrappier

A definição do operador de Ruelle-Perron-Frobenius, $\mathcal{L}_\psi$, no nosso contexto será a clássica onde $T : X \to X$ é uma aplicação expansora topologicamente mixing de um espaço métrico compacto e $\psi : X \to \mathbb{R}$ uma função contínua.

Para provar propriedades espectrais relevantes de $\mathcal{L}_\psi$ precisamos restringir as funções ψ para uma classe menor de funções. Na literatura são propostas diferentes classes de funções, mas do nosso ponto de vista nenhuma parece culminar em uma apresentação unificada e transparente. Assim, para preencher esta lacuna propomos a seguinte classe de funções:

Definição 3.3. *Seja $T : X \to X$ uma aplicação expansora em um espaço métrico compacto, sendo d sua métrica, e $r > 0$ dado pela definição de T. Dizemos que uma aplicação $\psi : X \to \mathbb{R}$ satisfaz a propriedade de Schwarz com respeito a T se:*

a) $\psi(x) \neq 0$, *para todo* $x \in X$

b) existem $A(r) > 0$ e $B > 0$ tais que se $n \geq 1$ e $s : U \to X$ é um ramo contrativo de T^{-n} com diam$(U) \leq r$ vale

$$\prod_{j=0}^{n-1} \frac{\psi(T^j(x))}{\psi(T^j(y))} \leq A(r) exp \left(B \sum_{j=0}^{n-1} d(T^j(x), T^j(y)) \right)$$

para todos $x, y \in s(U)$

c) $\lim_{r \to 0^+} A(r) = 1$.

Esta classe de funções é uma leve modificação da classe de funções primeiramente introduzida por Schwarz na sua famosa demonstração do teorema de Denjoy. A função ψ que aparece na definição clássica do operador de Ruelle-Perron-Frobenius está fortemente relacionada com o Jacobiano das medidas invariantes provenientes do operador. Bons Jacobianos irão constituir o que na literatura é chamado de "distorção limitada." E isto é o que temos quando requeremos que certa função pertença à classe de Schwarz. Nos parece, e o leitor poderá se convencer a medida que avança no texto, que esta classe de funções é a mais apropriada para esta teoria.

Exemplo. A qualquer função $\psi : X \to \mathbb{R}$ Hölder em um espaço métrico compacto, que não se anula em nenhum ponto, podemos adicionar uma constante para que ψ passe a pertencer a classe de Schwarz. $\square$

Lema 3.2. *Seja $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ uma transformação mensurável em um espaço de medida. Para todo $n \geq 0$ seja $\mathcal{A}_n = T^{-n}(\mathcal{A})$. Então, $\Phi : X \to \mathbb{R}$ é $\mathcal{A}_n$-mensurável se, e somente se, existe $\psi : X \to \mathbb{R}$ $\mathcal{A}$-mensurável tal que*

$$\Phi = \psi \circ T^n$$

para todo $n \geq 0$.

Demonstração. Para todo $n \geq 0$ seja $\mathcal{A}_n = T^{-n}(\mathcal{A})$. Claramente, $\mathcal{A}_n$ é uma σ-álgebra e

$$\mathcal{A} \supseteq \mathcal{A}_1 \supseteq \mathcal{A}_2 \supseteq \cdots \supseteq \mathcal{A}_n \supseteq \ldots$$

Assim, se

$$L^2(\mathcal{A}_n, \mu) = \{\Phi \in L^2(\mu) : \Phi \text{ é } \mathcal{A}_n \text{ mensurável}\}$$

para qualquer $n \geq 0$ teremos, também, que

$$L^2(\mathcal{A}, \mu) \supseteq L^2(\mathcal{A}_1, \mu) \supseteq \cdots \supseteq L^2(\mathcal{A}_n, \mu) \supseteq \ldots$$

Para descrever $L^2(\mathcal{A}_n, \mu)$, seja $\Phi : X \to \mathbb{R}$ uma aplicação $\mathcal{A}_n$-mensurável para um $n \geq 0$ qualquer. Então, para cada $y \in \mathbb{R}$

$$\Phi^{-1}(y) \in \mathcal{A}_n$$

e portanto $\Phi^{-1}(y) = T^{-1}(z_y)$, ou seja, $\Phi|_{T^{-n}(z_y)} = y$ com $z_y \in \mathcal{A}$. Em particular, para cada $x \in X$

$$\Phi|_{T^{-n}(x)} = \text{cte.}$$

Defina $\psi : X \to \mathbb{R}$ por $\psi(x) = \Phi|_{T^{-n}(x)}$. Então

$$\psi(T^n(x)) = \Phi|_{T^{-n}(T^n(x))} = \Phi(x)$$

ou, em outras palavras

$$\Phi = \psi \circ T^n.$$

Além disso, ψ é $\mathcal{A}$-mensurável. De fato, dado A um boreliano qualquer de $\mathbb{R}$

$$\Phi^{-1}(A) = T^{-n}(\psi^{-1}(A))$$

desde que Φ seja $\mathcal{A}_n$-mensurável. Então,

$$\Phi^{-1}(A) = T^{-n}(B) \qquad \text{com} \qquad B \in \mathcal{A}.$$

Deste argumento segue que $\Phi : X \to \mathbb{R}$ é $\mathcal{A}_n$-mensurável se, e somente se, existe $\psi : X \to \mathbb{R}$ $\mathcal{A}$-mensurável tal que

$$\Phi = \psi \circ T^n.$$

$\square$

Definição 3.4. *Uma transformação T que preserva medida de um espaço de probabilidade $(X, \mathcal{A}, \mu)$ é um endomorfismo exato (μ é exata) se*

$$\mathcal{A}_\infty := \bigcap_{j \geq 0} T^{-j}(\mathcal{A}) \ mod\ 0 = \{X, \emptyset\}.$$

Obs.: A definição anterior é equivalente a

$$\mu \text{ é exata} \iff \bigcap_{n \geq 0} L^2(\mathcal{A}_n, \mu) = \{\text{cte}\}$$

veja [M1].

Definição 3.5. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade e $\mathcal{B} \subseteq \mathcal{A}$ uma sub-σ-álgebra. O operador esperança condicional é*

$$E(.|\mathcal{B}) : L^1(X, \mathcal{A}, \mu) \to L^1(X, \mathcal{B}, \mu)$$

que a cada $f \in L^1(X, \mathcal{A}, \mu)$ associa uma única função, μ q.s., denotada por $E(f|\mathcal{B}) \in L^1(X, \mathcal{B}, \mu)$ tal que

$$\int_B f d\mu = \int_B E(f|\mathcal{B}) d\mu$$

para todo $B \in \mathcal{B}$.

Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade, e $\mathcal{B} \subseteq \mathcal{A}$ uma sub-σ-álgebra. Para $f \in L^1(X, \mathcal{A}, \mu)$ positiva seja

$$\nu_f(B) = \int_B f d\mu, \qquad \forall B \in \mathcal{B}$$

ν_f é uma medida em $\mathcal{B}$, veja [W2], e $\nu_f \ll \mu$. Então, pelo teorema de Radon-Nikodym [DS], existe um único elemento $E(f|\mathcal{B}) \in L^1(X, \mathcal{B}, \mu)$ tal que

$$\nu_f(B) = \int_B E(f|\mathcal{B}) d\mu$$

para todo $B \in \mathcal{B}$. Usando as partes positiva e negativa de uma $f \in L^1(X, \mathcal{A}, \mu)$ arbitrária temos que $E(.|\mathcal{B})$ age como um operador linear para todo $L^1(X, \mathcal{A}, \mu)$. Logo, $E(.|\mathcal{B})$ está bem definida.

Lema 3.3. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade e $\mathcal{B}$ uma sub-σ-álgebra de $\mathcal{A}$. A restrição do operador esperança condicional $E(.|\mathcal{B})$ a $L^2(X, \mathcal{A}, \mu)$ é a projeção ortogonal de $L^2(X, \mathcal{A}, \mu)$ em $L^2(X, \mathcal{B}, \mu)$.*

Demonstração. Para $f \in L^1(X, \mathcal{A}, \mu)$ sabemos que $E(f|\mathcal{B})$ é a única função $\mathcal{B}$-mensurável tal que

$$\int_B E(f|\mathcal{B}) d\mu = \int_B f d\mu, \qquad \forall B \in \mathcal{B}.$$

Seja P a projeção ortogonal de $L^2(X, \mathcal{A}, \mu)$ sobre um subespaço fechado $L^2(X, \mathcal{B}, \mu)$. Se $f \in L^2(X, \mathcal{A}, \mu)$

$$\begin{aligned} \int_B f d\mu = \int_X f \chi_B d\mu =< f, \chi_B >=< f, P\chi_B > \\ =< Pf, \chi_B >= \int_B Pf d\mu, \qquad \forall B \in \mathcal{B}. \end{aligned}$$

Portanto,

$$Pf = E(f|\mathcal{B}).$$

□

Outras propriedades sobre esperança condicional podem ser encontradas no apêndice deste trabalho.

Definição 3.6. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de medida, onde X é um espaço métrico compacto, e $T : X \to X$ uma transformação expansora. Definimos*

$$\mathcal{G} = \left\{ g : X \to \mathbb{R} \ ; \ g \in C(X), g > 0 \ e \sum_{y \in T^{-1}(x)} g(y) = 1, \forall x \in X \right\}.$$

Assim, dado $g \in \mathcal{G}$ seu operador de Ruelle-Perron-Frobenius associado é definido por

$$\mathcal{L}_{\log g}(f)(x) = \sum_{y \in T^{-1}(x)} g(y) f(y)$$

para $f \in C(X)$ e $x \in X$. Note que $\mathcal{L}_{\log g}$ satisfaz

$$\begin{aligned}\mathcal{L}_{\log g}(U_T(f))(x) &= \sum_{y \in T^{-1}(x)} g(y) f(T(y)) = \sum_{y \in T^{-1}(x)} g(y) f(x) \\ &= f(x) \sum_{y \in T^{-1}(x)} g(y) = f(x).\end{aligned}$$

Portanto, $\forall f \in C(X)$

$$\mathcal{L}_{\log g}(U_T(f)) = f$$

ou, de outro modo,

$$\mathcal{L}_{\log g}(U_T) = Id$$

onde $Id : C(X) \to C(X)$ é a aplicação identidade.

Teorema 3.1 (Ledrappier). *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade, onde X é um espaço métrico compacto, $T : X \to X$ uma transformação expansora e $g \in \mathcal{G}$. Se $\mathcal{L}$ denota $\mathcal{L}_{\log g}$ então são equivalentes*

i) $\mathcal{L}^* \mu = \mu$

ii) $\mu \in \mathcal{M}_T(X)$ *e para qualquer* $f \in L^1(\mu)$

$$E(f | T^{-1}(\mathcal{A}))(x) = \sum_{z \in T^{-1}(T(x))} g(z) f(z) \quad \mu \text{ q.s., } x \in X$$

iii) $\mu \in \mathcal{M}_T(X)$ *e* μ *é um estado de equilíbrio para* $\log g$.

Demonstração. *i)* $\Rightarrow$ *ii)* Seja $f \in C(X)$ e suponha que $\mathcal{L}^* \mu = \mu$. Pelo teorema 1.4, para provar que $\mu \in \mathcal{M}_T(X)$ basta mostrar que $\int_X f \circ T d\mu = \int_X f d\mu$. Mas isto é válido pois

$$\begin{aligned}\int_X f \circ T d\mu &= \int_X f \circ T d\mathcal{L}^* \mu = \int_X \mathcal{L}(f \circ T) d\mu \\ &= \int_X \mathcal{L}(U_T(f)) d\mu = \int_X f d\mu, \qquad \forall f \in C(X)\end{aligned}$$

mostrando assim que $\mu \in \mathcal{M}_T(X)$.

Agora, seja $A \in T^{-1}(\mathcal{A})$ donde

$$\begin{aligned}\int_A f d\mu = \int_X f\chi_A d\mu &= \int_X \mathcal{L}(f\chi_A) d\mu \\ = \int_X \mathcal{L}(f\chi_A)\circ T d\mu &= \int_X \sum_{y\in T^{-1}(T(x))} g(y)f(y)\chi_A(y)d\mu.\end{aligned}$$

Como $A \in T^{-1}(\mathcal{A})$, para algum $B \in \mathcal{A}$ temos que $A \in T^{-1}(B)$. Então, se $x \in A$ vale que $T(x) \in B$ e, portanto $T^{-1}(T(x)) \subseteq A$. Disto segue que

$$\chi_A(y) = \chi_A(x), \qquad \forall y \in T^{-1}(T(x)).$$

Portanto,

$$\begin{aligned}\int_X \sum_{y\in T^{-1}(T(x))} g(y)f(y)\chi_A(y)d\mu(x) &= \int_X \chi_A(x) \sum_{y\in T^{-1}(T(x))} g(y)f(y)d\mu(x) \\ &= \int_A \sum_{y\in T^{-1}(T(x))} g(y)f(y)d\mu(x)\end{aligned}$$

i. e., $\forall A \in T^{-1}(\mathcal{A})$

$$\int_A f d\mu = \int_A \sum_{y\in T^{-1}(T(x))} g(y)f(y)d\mu.$$

Já que $\sum_{y\in T^{-1}(T(x))} g(y)f(y)$ é claramente uma função $T^{-1}(\mathcal{A})$-mensurável, segue que

$$E(f|T^{-1}(\mathcal{A}))(x) = \sum_{y\in T^{-1}(T(x))} g(y)f(y).$$

ii) $\Rightarrow$ *iii)* A prova da existência de partições geradoras para T expansora se encontra no apêndice. Assim, seja $\mathcal{P} = \{P_1, \ldots, P_n\}$ uma partição de X tal que

$$\bigvee_{j=0}^{\infty} T^{-j}(\mathcal{P}) = \mathcal{A} \bmod 0$$

e portanto, para qualquer $\nu \in \mathcal{M}_T(X)$,

$$h_\nu(T) = -\int_X \log g_\nu(x) d\nu,$$

onde $g_\nu : X \to \mathbb{R}$ é a função definida, ν q.s., por

$$E_\nu(f|T^{-1}(\mathcal{A}))(x) = \sum_{y\in T^{-1}(T(x))} g_\nu(y)f(y).$$

Esta definição de entropia e a prova da existência da função g_ν se encontram no apêndice do trabalho. Assim,

$$\begin{aligned} h_\nu(T) &= -\int_X \log g_\nu(x) d\nu = -\int_X E_\nu(\log g_\nu | T^{-1}(\mathcal{A})) d\nu \\ &= -\int_X \sum_{y \in T^{-1}(T(x))} g_\nu(y) \log g_\nu(y) d\nu. \end{aligned}$$

Logo, usando que $\log x \leq x - 1$ para todo x temos

$$\begin{aligned} h_\nu(T) + \int_X \log g d\nu = -\int_X \log g_\nu d\nu + \int_X \log g d\nu &= \int_X \log\left(\frac{g}{g_\nu}\right) d\nu \\ &\leq \int_X \left(\frac{g}{g_\nu} - 1\right) d\nu \\ &= \int_X E\left(\frac{g}{g_\nu} - 1 \middle| T^{-1}(\mathcal{A})\right) d\nu \\ &= \int_X \sum_{y \in T^{-1}(T(x))} g_\nu(y) \left(\frac{g(y)}{g_\nu(y)} - 1\right) d\nu \\ &= \int_X \sum_{y \in T^{-1}(T(x))} (g(y) - g_\nu(y)) d\nu = 0 \end{aligned}$$

i. e.,

$$h_\nu(T) + \int_X \log g d\nu \leq 0.$$

Mas,

$$h_\mu(T) + \int_X \log g d\mu = 0.$$

Então ν é um estado de equilíbrio para $\log g$ se, e somente se,

$$\log \frac{g}{g_\nu} = \frac{g}{g_\nu} - 1, \qquad \nu \text{ q.s.}$$

i. e.,

$$g = g_\nu, \qquad \nu \text{ q. s.}$$

Como ν satisfaz $ii)$, $g_\nu = g$ portanto ν é um estado de equilíbrio para $\log g$.

$iii) \Rightarrow i)$ Sendo $\nu \in \mathcal{M}(X)$ um estado de equilíbrio para $\log g$ então pelo argumento acima $g_\nu = g$. Assim, se $f \in C(X)$

$$\begin{aligned} \int_X \mathcal{L}(f) d\nu = \int_X \mathcal{L}(f) \circ T d\nu &= \int_X \sum_{y \in T^{-1}(T(x))} g(y) f(y) d\nu \\ &= \int_X E(f | T^{-1}(\mathcal{A})) d\nu = \int_X f d\nu \end{aligned}$$

i. e.,

$$\mathcal{L}^*\nu = \nu.$$

□

Definição 3.7. *Dado $g \in \mathcal{G}$, uma medida satisfazendo as condições equivalentes do teorema de Ledrappier é dita de g-medida.*

Obs.: Nas hipóteses do teorema de Ledrappier, segue do teorema de Schauder-Tychonoff [DS] que $\mathcal{L}^*$ sempre tem um ponto fixo em $\mathcal{M}(X)$ donde uma g-medida sempre existe. Além disso, se μ é uma g-medida vale que

$$h_\mu(T) + \int_X \log g d\mu = 0$$

e g-medidas são estados de equilíbrios para $\log g$.

Lema 3.4. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade, onde X é um espaço métrico compacto, $T : X \to X$ uma transformação expansora que admite um jacobiano $J_\mu(T) > 0$, seja ainda $g \in \mathcal{G}$. Se μ é uma g-medida então*

i) $J_\mu(T) = 1/g$

ii) *μ é positiva sobre abertos quando T é topologicamente mixing*

iii) *se g_1, $g_2 \in \mathcal{G}$ e alguma g_1-medida coincide com outra g_2-medida qualquer então $g_1 = g_2$.*

Demonstração. i) Temos que

$$\begin{aligned}\int_A \frac{1}{g} d\mu = \int_X \frac{1}{g}\chi_A d\mu = \int_X \mathcal{L}\left(\frac{1}{g}\chi_A\right) d\mu \\ = \int_X \sum_{y \in T^{-1}(x)} g(y)\frac{1}{g(y)}\chi_A(y) d\mu \\ = \int_X \chi_{T(A)}(x) d\mu(x) = \mu(T(A))\end{aligned}$$

para qualquer $A \in \mathcal{B}$ tal que $T|_A$ é injetiva. Portanto,

$$J_\mu(T) = \frac{1}{g}.$$

ii) Segue imediatamente do lema 2.6 pois $J_\mu(T) > 0$.

iii) Segue imediatamente da unicidade do Jacobiano e de i). □

O próximo teorema afirma que sobre apropriadas restrições na função $g \in \mathcal{G}$ existe uma única g-medida μ. Além disso, para qualquer $f \in C(X)$ temos a convergência uniforme

$$\mathcal{L}^n(f) \rightrightarrows \int_X f d\mu$$

onde usaremos $\rightrightarrows$ para denotar a convergência uniforme. Para simplificar a notação vamos denotar $\mu(f) = \int_X f d\mu$.

Teorema 3.2. *Sejam $T : X \to X$ uma transformação expansora topologicamente mixing de um espaço métrico compacto e $g \in \mathcal{G}$ na classe de Schwarz. Então, existe uma única g-medida μ. Além disso, se $\mathcal{L}$ denota $\mathcal{L}_{\log g}$ vale que $\mathcal{L}^n(f)$ converge uniformemente para $\mu(f)$, $\forall f \in C(X)$.*

Demonstração. Dado $f \in C(X)$ primeiro vamos provar que

$$(\mathcal{L}^n(f))_n$$

é um subconjunto equicontínuo de $C(X)$ para obtermos, então, a convergência uniforme [DS]. Um cálculo direto nos dá que para todo $n \geq 1$ e $x \in X$

$$\mathcal{L}^n(f)(x) = \sum_{y \in T^{-n}(x)} g(y)g(T(y)) \cdots g(T^{n-1}(y))f(y).$$

Dado $\varepsilon > 0$ escolha $\delta > 0$ tal que se $d(x,y) < \delta$ possamos escrever, para todo $n \geq 1$

$$T^{-n}(x) = \{x_1, \dots, x_{r_n}\}$$
$$T^{-n}(y) = \{y_1, \dots, y_{r_n}\}$$

tal que para todos $1 \leq j \leq n$ e $1 \leq i \leq r_n$,

$$d(T^j(x_i), T^j(y_i)) < \alpha^{n-j} d(x,y)$$

onde $0 < \alpha < 1$ vem da definição da transformação expansora T. Como g está na classe de Schwarz, existem $A(\delta) > 0$, $B > 0$ e $\forall x_i \in T^{-n}(x)$, $\forall y_i \in T^{-n}(y)$ vale que

$$\begin{aligned}\prod_{j=0}^{n-1} \frac{g(T^j(x_i))}{g(T^j(y_i))} &\leq A(\delta) \exp\left(B \sum_{j=0}^{n-1} d(T^j(x_i), T^j(y_i))\right) \\ &\leq A(\delta) \exp\left(B \sum_{j=0}^{n-1} \alpha^{n-j} d(x,y)^\gamma\right) \\ &\leq A(\delta) \exp(C d(x,y)^\gamma)\end{aligned}$$

onde $C = B\sum_{n=0}^{\infty} \alpha^n < +\infty$, $\gamma > 0$. Diminuindo $\delta > 0$, se necessário, podemos assumir que se $d(x,y) < \delta$ então

$$d(f(x), f(y)) < \frac{\varepsilon}{2}.$$

Então,

$$\begin{aligned}
|\mathcal{L}^n(f)(x) - \mathcal{L}^n(f)(y)| &= \left| \sum_{i=1}^{r_n} g(x_i)\cdots g(T^{n-1}(x_i))f(x_i) \right. \\
&\quad \left. - \sum_{i=1}^{r_n} g(y_i)\cdots g(T^{n-1}(y_i))f(y_i) \right| \\
&\leq \sum_{i=1}^{r_n} g(x_i)\cdots g(T^{n-1}(x_i))|f(x_i) - f(y_i)| \\
&\quad + \sum_{i=1}^{r_n} |f(y_i)||g(x_i)\cdots g(T^{n-1}(x_i)) - g(y_i)\cdots g(T^{n-1}(y_i))| \\
&\leq \frac{\varepsilon}{2}\sum_{i=1}^{r_n} g(x_i)\cdots g(T^{n-1}(x_i)) + \|f\|_C \sum_{i=1}^{r_n} \left| \prod_{j=0}^{n-1} g(T^j(x_i)) - \prod_{j=0}^{n-1} g(T^j(y_i)) \right|
\end{aligned}$$

mas,

$$\begin{aligned}
&\left| \prod_{j=0}^{n-1} g(T^j(x_i)) - \prod_{j=0}^{n-1} g(T^j(y_i)) \right| = \prod_{j=0}^{n-1} g(T^j(y_i)) \left| \prod_{j=0}^{n-1} \frac{g(T^j(x_i))}{g(T^j(y_i))} - 1 \right| \\
&\leq \prod_{j=0}^{n-1} g(T^j(y_i)) \sup\{A(\delta)\exp(Cd(x,y)^\gamma) - 1, 1 - A(\delta)\exp(Cd(x,y)^\gamma)\} \\
&= \prod_{j=0}^{n-1} g(T^j(y_i))K.
\end{aligned}$$

Assim,

$$\begin{aligned}
|\mathcal{L}^n(f)(x) - \mathcal{L}^n(f)(y)| &\leq \frac{\varepsilon}{2} + \|f\|_C K \sum_i \prod_{j=0}^{n-1} g(T^j(y_i)) \\
&= \frac{\varepsilon}{2} + \|f\|_C K
\end{aligned}$$

já que

$$\sum_{i=1}^{r_n} \prod_{j=0}^{n-1} g(T^j(y_i)) = 1.$$

Como $K = \sup\{A(\delta)\exp(Cd(x,y)^\gamma) - 1, 1 - A(\delta)\exp(Cd(x,y)^\gamma)\}$ então se diminuirmos o valor de $\delta > 0$, mais uma vez, tal que

$$K < \frac{\varepsilon}{2\|f\|_C}$$

obtemos que, para $d(x, y) < \delta$

$$|\mathcal{L}^n(f)(x) - \mathcal{L}^n(f)(y)| < \varepsilon$$

provando que $(\mathcal{L}^n(f))_n$ é equicontínua. Já que $\|\mathcal{L}^n(f)\|_C \leq \|f\|_C$ segue do teorema de Arzelá-Ascoli [DS] a existência de uma sequência $(n_i)_i$ de números naturais e $f_\infty \in C(X)$ tais que

$$\mathcal{L}^{n_i}(f) \rightrightarrows f_\infty.$$

Agora, vamos provar que f_∞ é uma função constante. Para cada $f \in C(X)$ sejam

$$I(f) = \inf\{f(x); x \in X\}$$

$$S(f) = \sup\{f(x); x \in X\}.$$

Note que $I(f) \leq I(\mathcal{L}(f))$ e $S(f) \geq S(\mathcal{L}(f))$. De fato, para qualquer $x \in X$

$$\mathcal{L}(f)(x) = \sum_{y \in T^{-1}(x)} g(y) f(y) \geq I(f) \sum_{y \in T^{-1}(x)} g(y) = I(f)$$

e então $I(\mathcal{L}(f)) \geq I(f)$. Analogamente provamos a desigualdade para S.

Além disso, $I(f_\infty) = I(\mathcal{L}(f_\infty))$ e $S(f_\infty) = S(\mathcal{L}(f_\infty))$. Vamos provar a primeira igualdade, a segunda é análoga. Como $\mathcal{L}^{n_i}(f) \rightrightarrows f_\infty$ temos que $\mathcal{L}^{n_i+1}(f) \rightrightarrows f_\infty$ e portanto

$$I(f) \leq I(\mathcal{L}(f)) \leq \cdots \leq I(\mathcal{L}^n(f)) \leq \cdots \leq I(f_\infty).$$

Agora como

$$I(\mathcal{L}^{n_i+1}(f)) \leq I(f_\infty)$$

concluímos que

$$I(\mathcal{L}(f_\infty)) \leq I(f_\infty)$$

e como já vimos que

$$I(f_\infty) \leq I(\mathcal{L}(f_\infty))$$

segue o resultado.

Finalmente, se $x \in X$ é tal que

$$I(f_\infty) = I(\mathcal{L}(f_\infty)) = \mathcal{L}(f_\infty)(x)$$

então para qualquer $y \in T^{-1}(x)$ vale que $f_\infty(y) = I(f_\infty)$. Se não for este o caso,

$$\begin{aligned} I(f_\infty) = I(\mathcal{L}(f_\infty)) = \mathcal{L}(f_\infty)(x) = &\sum_{y \in T^{-1}(x)} g(y) f_\infty(y) \\ &> \sum_{y \in T^{-1}(x)} g(y) I(f_\infty) \\ &= I(f_\infty) \end{aligned}$$

um absurdo. Então, $\forall y \in T^{-1}(x)$

$$f_\infty(y) = I(f_\infty).$$

Se $U \subseteq X$ é um conjunto aberto qualquer, pelo item a) do lema 1.3, existe um $M > 0$ tal que

$$T^M(U) = X.$$

Isto significa que existe $y \in U$ tal que $T^M(y) = x$, i. e., $y \in T^{-M}(x)$ e portanto

$$f_\infty(y) = I(f_\infty).$$

Além disso, f_∞ assume seu valor mínimo (global) em qualquer conjunto aberto e como f_∞ é contínua segue que é constante. Disto segue facilmente que

$$\mathcal{L}^n(f) \rightrightarrows f_\infty.$$

Usando o teorema de representação de Riesz, teorema 1.2, existe $\mu \in \mathcal{M}(X)$ uma probabilidade tal que

$$\mathcal{L}^n(f) \rightrightarrows \mu(f)$$

para toda $f \in C(X)$. Se $f \in C(X)$ então

$$\begin{aligned} \int_X f d\mathcal{L}^*\mu &= \int_X \mathcal{L}(f) d\mu = \lim_{n\to\infty} \mathcal{L}^n(\mathcal{L}(f)) = \lim_{n\to\infty} \mathcal{L}^{n+1}(f) \\ &= \mu(f) \end{aligned}$$

provando assim que μ é uma g-medida.

Finalmente, se ν é outra g-medida então para toda $f \in C(X)$

$$\begin{aligned} \int_X f d\nu = \int_X f d\mathcal{L}^*\nu = \int_X \mathcal{L}(f) d\nu = \int_X \mathcal{L}^n(f) d\nu \\ \to \int_X \left(\int_X f d\mu \right) d\nu = \int_X f d\mu \end{aligned}$$

e então, pelo teorema 1.1, $\mu = \nu$ finalizando assim a prova do teorema. $\square$

Exemplo. Neste exemplo definiremos o subshift de Bernoulli. Para isto, seja $B^+(n)$ o espaço das sequências unilaterais no conjunto $X = \{1, \ldots, n\}$. Seja μ_0 uma probabilidade de X definida por $\mu_0(\{i\}) = p_i$ onde $0 \leq p_i \leq 1$ e $\sum_{i=1}^n p_i = 1$. Pelo teorema de extensão de medidas existe uma única medida de probabilidade μ, denominada medida produto associada a μ_0, tal que

$$\mu(C(j; A_0, \ldots, A_m)) = \prod_{j=0}^{m} \mu_0(A_j).$$

Denotamos por $B_\mu^+(n)$ ao espaço $B^+(n)$ munido da medida μ, e o shift $\sigma : B_\mu^+(n) \to B_\mu^+(n)$ é denominado shift de Bernoulli.

Se definirmos a função $g : B_\mu^+(n) \to \mathbb{R}$ como $g(\theta) = p_j$ onde $\theta(0) = j$ então a medida produto μ é a única g-medida para o sistema. $\square$

Exemplo. Sejam $X = \mathbb{R}/\mathbb{Z}$ e $Tx = 2x(\mod 1)$. Se $g \in \mathcal{G}$ é dada por $g(0) = g(1/3) = g(2/3) = 1$ então é fácil provar que as medidas

$$\mu_0 = \delta_0 \qquad \text{e} \qquad \mu_1 = \frac{\delta_{1/3} + \delta_{2/3}}{2}$$

onde δ_p é a medida de Dirac do ponto p, satisfazem $\mathcal{L}_{\log g}^* \mu = \mu$ portanto, são ambas g-medidas. $\square$

Exemplo. Sejam $X = \mathbb{R}/\mathbb{Z}$ e $Tx = 2x(\mod 1)$. Vamos considerar uma função g dada por

$$g(x) = \begin{cases} \frac{1+\cos 2\pi x}{2}, & x \neq 0, 1/2 \\ 1/2, & x = 0, 1/2. \end{cases}$$

Um cálculo simples nos dá que para qualquer $f \in C(X)$ temos que $\mathcal{L}_{\log g}^n f$ converge a $f(0)$ tome, por exemplo, $f \equiv 1$ e $x = 0$. Porém, $\mu = \delta_0$ não é uma g-medida pois g é descontínua em zero. $\square$

Capítulo 4

Prova do teorema de Ruelle

Neste capítulo provaremos o principal teorema deste trabalho, o teorema de Ruelle para potenciais na classe de Schwarz.

4.1 Prova do teorema

Definição 4.1. *Seja $T : X \to X$ uma transformação expansora em um espaço métrico compacto (X, d). Para qualquer $n \geq 1$ definimos a métrica d_n por*

$$d_n(x, y) = \sup\{d(T^j(x), T^j(y)); 0 \leq j \leq n-1\}$$

$\forall x, y \in X$.

Obs.: É trivial provar que d_n é uma métrica.

Agora podemos enunciar e provar o principal teorema deste trabalho que pode ser visto como uma generalização do teorema 3.2.

Teorema 4.1 (Ruelle). *Sejam $T : X \to X$ uma aplicação expansora topologicamente mixing em um espaço métrico compacto, $\psi \in C(X)$ uma aplicação tal que e^{ψ} esteja na classe de Schwarz e $\mathcal{L}_\psi$ o operador de Ruelle-Perron-Frobenius dado por*

$$\mathcal{L}_\psi(f)(x) = \sum_{y \in T^{-1}(x)} e^{\psi(y)} f(y)$$

para toda $f \in C(X)$. Então, existem um número $\lambda > 0$, uma função $h \in C(X)$ e uma medida $\nu \in \mathcal{M}(X)$ tais que

i) $\mathcal{L}_\psi^* \nu = \lambda \nu$

ii) $h > 0$ *e* $\mathcal{L}_\psi h = \lambda h$

iii) h *é a única autofunção positiva de* $\mathcal{L}_\psi$ *exceto por multiplicação por escalar*

iv) $\int_X h d\nu = 1$

v) para toda $f \in C(X)$ *vale*

$$\left\| \lambda^{-n} \mathcal{L}_\psi^n f - h \int_X f d\nu \right\|_C \to 0$$

vi) $\mu = h\nu$ *é uma probabilidade* T*-invariante, exata, positiva sobre abertos e*

$$\log \lambda = h_\mu(T) + \int_X \psi d\mu$$

vii) para qualquer $\eta \in \mathcal{M}_T(X)$, $\eta \neq \mu$ *vale*

$$\log \lambda > h_\eta(T) + \int_X \psi d\eta.$$

Demonstração. Primeiro, vamos provar o item i) e a existência de $\nu \in \mathcal{M}(X)$ e $\lambda > 0$.

Seja o operador $L : \mathcal{M}(X) \to \mathcal{M}(X)$ definido por

$$L(\nu) = \frac{\mathcal{L}_\psi^* \nu}{(\mathcal{L}_\psi^* \nu)(1)}.$$

Como $\mathcal{M}(X)$ é um conjunto convexo e compacto, o teorema do ponto fixo de Schauder-Tychonoff [DS] nos garante a existência de um ponto fixo, $\nu \in \mathcal{M}(X)$, para L. Portanto,

$$\mathcal{L}_\psi^* \nu = \lambda \nu$$

onde $\lambda = (\mathcal{L}_\psi^* \nu)(1) > 0$.

Para achar h e provar o item ii) vamos, novamente, fazer uso do teorema de Schauder-Tychonoff para $\mathcal{L}_\psi$ agindo em um subconjunto de $C(X)$ convexo e compacto. Para definir este conjunto fixemos $\varepsilon_0 > 0$ menor que a constante de expansividade de T, $0 < \alpha < 1$, e tal que para quaisquer $x, y \in X$ com $d(x, y) < \varepsilon_0$ e $n > 0$ possamos escrever

$$T^{-n}(x) = \{x_1, \ldots, x_s\}$$

$$T^{-n}(y) = \{y_1, \ldots, y_s\},$$

onde s é o número de pré-imagens de T, satisfazendo para $0 \leq j \leq n$ e $0 \leq i \leq s$

$$d(T^j(x_i), T^j(y_i)) \leq \alpha^{n-j} d(x,y).$$

Para cada $k \geq 1$, $n > 0$ e $x, y \in X$ com $d_k(x,y) \leq \varepsilon_0$ definimos

$$B_k^n(\psi)(x,y) = \sup \left\{ \prod_{j=0}^{n-1} \frac{e^{\psi(T^j(x_i))}}{e^{\psi(T^j(y_i))}}; 1 \leq i \leq k_n \right\}$$

onde $T^{-n}(x) = \{x_1, \ldots, x_{k_n}\}$, $T^{-n}(y) = \{y_1, \ldots, y_{k_n}\}$ e

$$d(T^j(x_i), T^j(y_i)) \leq \alpha^{n-j} d(x,y)$$

para $0 \leq j \leq n$ e $1 \leq i \leq k_n$.

Finalmente, definimos

$$B_k(\psi)(x,y) = \sup\{B_k^n(\psi)(x,y); n \geq 1\}.$$

É fácil ver que, quando e^{ψ} pertence à classe de Schwarz, B_k é uma função contínua bem definida em $X \times X$.

Agora, consideremos o conjunto

$$\begin{aligned} \Lambda = \{ f \in C(X); \nu(f) = 1, f > 0 \text{ e } \forall k \geq 1, \\ f(x) \leq B_k(\psi)(x,y) f(y) \text{ se } d_k(x,y) \leq \varepsilon_0 \} \end{aligned}$$

temos que Λ é um subconjunto convexo e compacto de $C(X)$.

De fato Λ é um subconjunto convexo pois se $\alpha + \beta = 1$ com $\alpha, \beta > 0$ e $f, h \in \Lambda$ então quando $d_k(x,y) \leq \varepsilon_0$

$$\begin{aligned} \alpha f(x) + \beta h(x) &= \frac{f(x)}{f(y)} \alpha f(y) + \frac{h(x)}{h(y)} \beta h(y) \\ &\leq \sup \left\{ \frac{f(x)}{f(y)}, \frac{h(x)}{h(y)} \right\} (\alpha f(y) + \beta h(y)) \\ &\leq B_k(\psi)(x,y)(\alpha f(y) + \beta h(y)) \end{aligned}$$

provando que $\alpha f + \beta h \in \Lambda$.

Para provar que Λ é limitado, escolha $M > 0$, dado pelo item b) do lema 1.3, tal que para qualquer ponto $x \in X$, $T^{-M}(x)$ é ε_0-denso. Então considere quaisquer pontos $x, z \in X$ e escolha $y \in T^{-M}(z)$ com $d(x,y) \leq \varepsilon_0$. Se $f \in \Lambda$, $f > 0$ então

$$\begin{aligned}(\mathcal{L}_\psi^M f)(z) &= \sum_{s \in T^{-M}(z)} e^{\psi(s)+\cdots+\psi(T^{M-1}(s))} f(s) \\ &\geq e^{\psi(y)+\cdots+\psi(T^{M-1}(y))} f(y) \\ &\geq e^{-M\|\psi\|} B_1^{-1}(\psi)(y,x) f(x) \\ &\geq e^{-M\|\psi\|} B_1^{-1} f(x)\end{aligned}$$

onde $B_1 = \sup\{B_1(\psi)(x,y); x, y \in X\}$. Portanto,

$$f(x) \leq \frac{(\mathcal{L}_\psi^M f)(z)}{\lambda^M} e^{M\|\psi\|} B_1 \lambda^M$$

para todo z, e integrando esta desigualdade com respeito a ν e usando o fato que $\mathcal{L}^*\nu = \lambda\nu$ obtemos

$$\begin{aligned}f(x) &\leq \lambda^M e^{M\|\psi\|} B_1 \nu\left(\frac{\mathcal{L}^M f}{\lambda^M}\right) \\ &= \lambda^M e^{M\|\psi\|} B_1 = K.\end{aligned}$$

Assim, f é limitada.

Para provar a equicontinuidade seja $\varepsilon > 0$ e escolha $k \geq 1$ tal que

$$K|B_k - 1| < \varepsilon.$$

Agora escolha $\delta > 0$ tal que

$$d(x,y) < \delta \qquad \Rightarrow \qquad d_k(x,y) \leq \varepsilon_0.$$

Então, se $d(x,y) < \delta$, $d_k(x,y) \leq \varepsilon_0$ e

$$\begin{aligned}|f(x) - f(y)| &\leq f(x)|B_k - 1| \\ &\leq K|B_k - 1| < \varepsilon.\end{aligned}$$

Como Λ é claramente fechado segue do teorema de Arzelá-Ascoli [DS] que Λ é compacto.

Defina $L : C(X) \to C(X)$ por

$$L = \lambda^{-1}\mathcal{L}_\psi$$

e vamos verificar que $L(\Lambda) \subset \Lambda$.

Se $f \in \Lambda$, $L(f) > 0$ e $\nu(L(f)) = \nu(f) = 1$. Além disso, se $d_k(x,y) \leq \varepsilon_0$ podemos escrever $T^{-1}(x) = \{x_1, \ldots, x_s\}$, $T^{-1}(y) = \{y_1, \ldots, y_s\}$ com $d(x_i, y_i) \leq \alpha d(x,y)$. Então, $d_{k+1}(x_i, y_i) \leq \varepsilon_0$ e

$$\begin{aligned}
(L(f))(x) &= \lambda^{-1} \sum_i e^{\psi(x_i)} f(x_i) \\
&\leq \lambda^{-1} \sum_i e^{\psi(x_i)} B_{k+1}(\psi)(x_i, y_i) f(y_i) \\
&= \lambda^{-1} \sum_i \frac{e^{\psi(x_i)}}{e^{\psi(y_i)}} B_{k+1}(\psi)(x_i, y_i) e^{\psi(y_i)} f(y_i) \\
&\leq \lambda^{-1} \sum_i B_k(\psi)(x,y) e^{\psi(y_i)} f(y_i) \\
&= B_k(\psi)(x,y)(L(f))(y)
\end{aligned}$$

provando que $L(f) \in \Lambda$, i. e.,$L(\Lambda) \subset \Lambda$.

O teorema do ponto fixo de Schauder-Tychonoff [DS] nos garante então, a existência de um ponto fixo para o operador L, i. e., uma função $h \in \Lambda$ tal que

$$L(h) = h$$

ou, de outro modo,

$$\mathcal{L}_\psi h = \lambda h.$$

Para provar que $h > 0$, assuma por contradição que $h(x) = 0$ para algum $x \in X$. Como $L(h) = h$ segue que

$$L(h)(x) = \lambda^{-1} \sum_{y \in T^{-1}(x)} e^{\psi(y)} h(y) = h(x) = 0$$

e então, $h(y) = 0$ para qualquer $y \in T^{-1}(x)$. Por indução, $h(z) = 0$ para todo $z \in T^{-n}(x)$ e todo $n \geq 1$. Como $T^{-n}(x)$ é ε-denso para n suficientemente grande e $h \in C(X)$, segue que $h \equiv 0$ contradizendo o fato que $\nu(h) = 1$.

Os items i) e ii) já estão provados. Como $h \in \Lambda$ então $\nu(h) = 1$ que é o item iv).

Agora vamos provar o item v). Para isto seja

$$g(x) = \frac{e^{\psi(x)} h(x)}{\lambda h(T(x))}$$

e vale que g é contínua, $g > 0$ e

$$\begin{aligned}
\sum_{y\in T^{-1}(x)} g(y) &= \sum_{y\in T^{-1}(x)} \frac{e^{\psi(y)}h(y)}{\lambda h(T(y))} = \frac{1}{\lambda h(x)} \sum_{y\in T^{-1}(x)} e^{\psi(y)}h(y) \\
&= \frac{1}{\lambda}\frac{(\mathcal{L}_\psi h)(x)}{h(x)} = 1.
\end{aligned}$$

Além disso,

$$\begin{aligned}
L(f)(x) &= \frac{1}{\lambda} \sum_{y\in T^{-1}(x)} e^{\psi(y)} f(y) \\
&= \sum_{y\in T^{-1}(x)} \frac{e^{\psi(y)}h(y)}{\lambda h(T(y))} \frac{f(y)}{h(y)} h(T(y)) \\
&= h(x) \sum_{y\in T^{-1}(x)} g(y) \left(\frac{f}{h}\right)(y) \\
&= h(x)\mathcal{L}_{\log g}\left(\frac{f}{h}\right)(x)
\end{aligned}$$

ou seja, $L(f) = h\mathcal{L}_{\log g}(f/h)$. Para a $(n+1)$-ésima iterada temos que

$$\begin{aligned}
L^{n+1}(f) = L(L^n(f)) &= L\left(h\mathcal{L}^n_{\log g}(f/h)\right) \\
&= h\mathcal{L}_{\log g}\left(h\mathcal{L}^n_{\log g}(f/h)/h\right) \\
&= h\mathcal{L}^{n+1}_{\log g}(f/h)
\end{aligned}$$

provando, por indução, que para toda $f \in C(X)$

$$\lambda^{-n}\mathcal{L}^n_\psi(f) = h\mathcal{L}^n_{\log g}(f/h).$$

Precisamos aplicar o teorema 3.2 e para isto precisamos provar que

$$g(x) = \frac{e^{\psi(x)}h(x)}{\lambda h(T(x))}$$

está na classe de Schwarz. Assim, para qualquer $r > 0$ defina

$$K(r) = \sup\left\{\frac{h(x)}{h(y)}; d(x,y) \le r\right\}.$$

Como e^ψ está na classe de Schwarz, existe uma constante $B > 0$ e $A(r) > 0$ tal que sempre que $S : W \to X$ é um ramo contrativo de T^{-n}, $n \ge 1$ e $x, y \in S(W)$ então

$$\prod_{j=0}^{n-1} \frac{e^{\psi(T^j(x))}}{e^{\psi(T^j(y))}} \le A_\psi(r)\exp\left(B\sum_{j=0}^{n-1} d(T^j(x), T^j(y))\right).$$

Assim,

$$
\begin{aligned}
\prod_{j=0}^{n-1} \frac{g(T^j(x))}{g(T^j(y))} &= \prod_{j=0}^{n-1} \frac{e^{\psi(T^j(x))}}{e^{\psi(T^j(y))}} \frac{h(T^j(x))\lambda h(T^{j+1}(y))}{\lambda h(T^{j+1}(x))h(T^j(y))} \\
&= \left(\prod_{j=0}^{n-1} \frac{e^{\psi(T^j(x))}}{e^{\psi(T^j(y))}} \right) \frac{h(x)}{h(y)} \frac{h(T^n(y))}{h(T^n(x))} \\
&\leq A_\psi(r) K(r)^2 \exp\left(B \sum_{j=0}^{n-1} d(T^j(x), T^j(y)) \right) \\
&= A_g(r) \exp\left(B \sum_{j=0}^{n-1} d(T^j(x), T^j(y)) \right).
\end{aligned}
$$

Claramente, $\lim_{r\to 0} A_g(r) = 1$, provando que g pertence à classe de Schwarz. Então aplicando o teorema 3.2 obtemos

$$
\frac{1}{\lambda^n} \mathcal{L}_\psi^n f = h(x) \mathcal{L}_{\log g}^n (f/h)(x) \to h\mu(f/h)
$$

onde μ é a única g-medida. Assim,

$$
\frac{1}{\lambda^n} \mathcal{L}_\psi^n f \to h \int_X \frac{f}{h} d\mu.
$$

Defina $\eta = h d\nu$ e como $\mathcal{L}_\psi^* \nu = \lambda \nu$ então

$$
\begin{aligned}
\int_X f d\mathcal{L}_{\log g}^* \eta = \int_X \mathcal{L}_{\log g} f d\eta &= \int_X \mathcal{L}_{\log g} f \frac{h}{h} d\eta \\
&= \int_X \frac{1}{\lambda h} \mathcal{L}_\psi(fh) d\eta = \int_X \frac{1}{\lambda h} \mathcal{L}_\psi(fh) h d\nu \\
&= \frac{1}{\lambda} \int_X \mathcal{L}_\psi(fh) d\nu = \frac{1}{\lambda} \int_X fh d\mathcal{L}_\psi^* \nu \\
&= \frac{1}{\lambda} \int_X fh d\lambda\nu = \int_X f d\eta
\end{aligned}
$$

i. e., para toda $f \in C(X)$

$$
\int_X f d\mathcal{L}_{\log g}^* \eta = \int_X f d\eta
$$

e então

$$
\mathcal{L}_{\log g}^* \eta = \eta.
$$

Portanto da unicidade das g-medidas temos que

$$
\eta = \mu
$$

i. e., $\mu = hd\nu$ e provamos que para qualquer $f \in C(X)$

$$\frac{1}{\lambda^n}\mathcal{L}_\psi^n f \rightrightarrows h\int_X f d\nu$$

quando $n \to \infty$, ou seja

$$\left\|\frac{1}{\lambda^n}\mathcal{L}_\psi^n f - h\int_X f d\nu\right\|_C \to 0.$$

Pelo teorema de Ledrappier 3.1 segue que a medida $\mu = h\nu$ é probabilidade T-invariante, um estado de equilíbrio para $\log g$, é positiva sobre conjuntos abertos e é exata. E ainda temos que $h_\mu(T) + \int_X \psi d\mu = \log\lambda$ pois

$$\begin{aligned}
h_\mu(T) &= \int_X -\log g d\mu = \int_X -\log e^{\psi(x)}\frac{h(x)}{\lambda h(T(x))}d\mu \\
&= -\int_X \psi d\mu - \int_X \log h d\mu + \log\lambda + \int_X \log h(T(x))d\mu \\
&= -\int_X \psi d\mu + \log\lambda - \int_X \log h d\mu + \int_X \log h \circ T d\mu \\
&= -\int_X \psi d\mu + \log\lambda.
\end{aligned}$$

O que prova vi). Além disso, se η é outra probabilidade T-invariante então

$$\log\lambda > h_\eta(T) + \int_X \psi d\eta.$$

Terminando, assim, a prova do teorema de Ruelle. $\square$

Para finalizar este capítulo vamos apresentar dois corolários do teorema de ruelle que envolve funções homólogas.

Corolário 4.1.1. *Sejam $T : X \to X$ uma aplicação expansora e topologicamente mixing em um espaço métrico compacto, e $\phi \in C(X)$ tal que e^ϕ pertence à classe de Schwarz. Então, para $\nu \in \mathcal{M}(X)$, $\lambda > 0$ e $h \in C(X)$ dados pelo teorema de Ruelle,*

i) ϕ tem um único estado de equilíbrio, μ_ϕ, e $\mu_\phi(f) = \nu(hf)$, $\forall f \in C(X)$

ii) μ_ϕ é a única g-medida para $g = \frac{e^\phi h}{\lambda h T}$

iii) μ_ϕ é positiva sobre abertos não-vazios

iv) a pressão topológica de ϕ com relação a T vale $\log\lambda$ em símbolos, $P(T,\phi) = \log\lambda$ e λ é o raio espectral de $\mathcal{L}_\phi$ e vale a convergência $(\log\mathcal{L}_\phi^n 1)/n \to P(T,\phi)$.

Demonstração. i) Calculando o logaritmo de $g = \frac{e^{\phi}h}{\lambda h T}$ obtemos que $\phi \sim \log g$ pois

$$\phi = \log g + \log h \circ T - \log h + \log \lambda$$

assim, pelo teorema 1.10, $E_{\phi} = E_{\log g}$. Vimos que $\mathcal{L}^n_{\log g}(f)$ converge a $\nu(fh)$ donde temos que $\log g$ tem um único estado de equilíbrio dado por $\mu_{\phi}(f) = \nu(hf), \forall f \in C(X)$. Assim μ_{ϕ} é o único estado de equilíbrio para ϕ.

ii) Segue do teorema de Ruelle.

iii) É imediato do item *ii*) do lema 3.4.

iv) Como μ_{ϕ} é um estado de equilíbrio para $\phi = \log g + \log h \circ T - \log h + \log \lambda$ temos, por um cálculo direto, que

$$P(T, \phi) = h_{\mu_{\phi}}(T) + \mu_{\phi}(\phi) = h_{\mu_{\phi}}(T) + \mu_{\phi}(\log g) + \log \lambda = \log \lambda.$$

O raio espectral para $\mathcal{L}_{\phi}$ é dado por

$$\lim_{n\to\infty} \|\mathcal{L}^n_{\phi}\|^{1/n} = \lim_{n\to\infty} \|\mathcal{L}^n_{\phi} 1\|^{1/n} = \lambda$$

pelo teorema de Ruelle. Como $\mathcal{L}^n_{\phi} 1/\lambda^n \to h$ segue que $(\log \mathcal{L}^n_{\phi} 1)/n \to P(T, \phi)$.

Note que λ, ν e $h > 0$ são todos unicamente determinados por $\mathcal{L}^n_{\phi} f/\lambda^n \to h\nu(f)$ visto que

$$\log \lambda = \lim_{n\to\infty} \frac{1}{n} \log \mathcal{L}^n_{\phi} 1$$

e

$$h = \lim_{n\to\infty} \frac{\mathcal{L}^n_{\phi} 1}{\lambda^n}.$$

□

Corolário 4.1.2. *Sejam $T : X \to X$ expansora e topologicamente mixing em um espaço métrico compacto e as aplicações $\phi, \psi \in C(X)$ tais que e^{ϕ} e e^{ψ} pertencem à classe de Schwarz. Então, $\mu_{\phi} = \mu_{\psi}$ se e somente se $\phi - \psi = f \circ T - f + c$ para alguma $f \in C(X)$ e $c \in \mathbb{R}$.*

Demonstração. [$\Rightarrow$] Vamos supor que $\mu_{\phi} = \mu_{\psi}$. Pelo corolário anterior temos que μ_{ϕ} é a única g_1-medida para alguma $g_1 = e^{\phi} h_1/(\lambda_1 h_1 T)$ e é a única g_2-medida para alguma $g_2 = e^{\phi} h_2/(\lambda_2 h_2 T)$. Assim, pelo item *iii*) do lema 3.4 temos que $g_1 = g_2$. Portanto,

$$\phi - \psi = \log \frac{h_1 \circ T}{h_2} - \log \frac{h_1}{h_2} + \log \frac{\lambda_1}{\lambda_2}.$$

[$\Leftarrow$] Foi provado no teorema 1.10. □

Capítulo 5

Aplicações Expansoras por Partes e Teoria Espectral

Primeiramente vamos expor o ambiente de trabalho para este capítulo.

Para definições complementares em topologia geral veja [Ke]. Seja $(X,<)$ um conjunto totalmente ordenado, fechado, limitado pela ordem e completo pela ordem, i. e., todo conjunto não-vazio com uma cota superior tem um supremo. A topologia da ordem para X é a menor topologia onde a ordem é contínua no sentido que se $a, b \in X$ com $a < b$ então existem vizinhanças U de a e V de b tais que para quaisquer $x \in U$ e $y \in V$ tenhamos que $x < y$.

Consideraremos $T : X \to X$ uma aplicação monótona por partes e contínua pela ordem. Uma função monótona por partes $T : X \to X$ é tal que para $N > 1$, $X = \bigcup_{i=1}^{N} I_i$, onde os I_i são intervalos disjuntos e $T|_{I_i}$ é contínua e preserva ou inverte a ordenação de modo que $T(I_i)$ ainda seja um intervalo.

Em X consideramos uma medida que é uma probabilidade a Borel, l, e denotamos por $B_l(X)$ o espaço das funções mensuráveis a Borel e limitadas em X.

Definição 5.1. *Sejam $(X,<)$ um conjunto totalmente ordenado, fechado, limitado pela ordem e completo pela ordem e $f : X \to \mathbb{R}$. A variação de f em X é dada por*

$$var_X(f) = \sup\left\{\sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|; n \geq 2, x_0 < x_1 < \ldots < x_n\right\}.$$

A função f é dita de variação limitada se $var_X(f) < \infty$.

Dada uma função densidade (ou potencial) $g : X \to (0, d]$ onde $d < 1$, que assumimos ser de variação limitada denotamos, como nos capítulos anteriores, o operador de Ruelle-Perron-Frobenius por

$$\begin{aligned}\mathcal{L} :& B_l(X) \to B_l(X)\\ \mathcal{L}(f)(x) &= \sum_{y \in T^{-1}(x)} g(y)f(y)\end{aligned}$$

para $f \in B_l(X)$ e $x \in X$.

Quando $T : X \to X$ é uma aplicação $C^{1+\gamma}$ expansora por partes e

$$g(x) = \frac{1}{|DT(x)|}$$

é fácil verificar que $\mathcal{L}$ satisfaz

$$m(\mathcal{L}f) = m(f)$$

onde m é a medida de Lebesgue no intervalo. Assim, adicionamos essa hipótese à nossa medida, i. e.,

$$l(\mathcal{L}f) = l(f)$$

para toda $f \in B_l(X)$.

Definição 5.2. *Definimos por $BV(X)$ o espaço das funções de variação limitada em X com a norma*

$$\|f\|_{var} = \int_X |f|dl + var_X(f).$$

Precisaremos da variação para funções em $L^1(l)$, o conjunto de todas as classes de equivalência das funções a valores complexos, l-integráveis em X. Assim, definimos para $f \in L^1(l)$

$$\text{var}(f) = \inf\{\text{var}_X(\widetilde{f}) : \widetilde{f} \sim f\}.$$

5.1 Lemas Básicos

Vamos recordar nossas definições usuais. Para $f \in B_l(X)$ o operador de Ruelle-Perron-Frobenius $\mathcal{L}$ é dado por

$$\begin{aligned}\mathcal{L}(f)(x) &= \sum_{y \in T^{-1}(x)} g(y)f(y)\\ &= \sum_{i=1}^{N} g(T_i^{-1}(x))f(T_i^{-1}(x))\chi_{T(I_i)}(x)\end{aligned}$$

onde $T_i = T|_{I_i}$. Também assumimos que l é uma medida que é uma probabilidade a Borel em X e que satisfaz

$$l(\mathcal{L}f) = l(f), \qquad \forall f \in B_l(X).$$

Gostaríamos de mudar nossa função de forma que o suporte de l se torne invariante, e então estudar a aplicação restrita ao suporte da medida l. Para isto precisamos garantir que o operador $\mathcal{L}$ não sofra mudanças ao agir em $L^1(l)$. Para fazer isto precisaremos dos seguintes lemas.

Lema 5.1. *Seja B um boreliano de X com $l(B) = 0$ então $l(T(B)) = 0$.*

Demonstração. Podemos assumir que $B \subseteq I_i$ para algum i. Se $l(T(B)) > 0$ então

$$\begin{aligned} l(B) = l(\chi_B) = l(\mathcal{L}\chi_B) &= l\left(\sum_{j=1}^{N} g(T_j^{-1}(x))\chi_B(T_j^{-1}(x))\chi_{T(I_j)}(x)\right) \\ &= l\left(g(T_i^{-1}(x))\chi_B(T_i^{-1}(x))\chi_{T(I_i)}(x)\right) = l\left(g(T_i^{-1}(x))\chi_{T(B)}(x)\right) \end{aligned}$$

pois $\chi_B(T_i^{-1}(x)) = 0$ se $x \notin T(B)$.

Agora considere o conjunto

$$A_k = \left\{x \in T(B) : g(T_i^{-1}(x)) \geq \frac{1}{k}\right\}$$

para todo $k > 1$, como $l(T(B)) > 0$, $\int_{T(B)} gdl > 0$ e então existe um k_0 tal que

$$l(A_{k_0}) > 0.$$

Assim, temos que

$$l(B) = l\left(g(T_i^{-1}(x))\chi_{T(B)}(x)\right) \geq \frac{1}{k_0} l(T(B) \cap A_{k_0}) > 0$$

o que prova o lema. $\square$

Definição 5.3. *Seja $\mathcal{A} = \{I_1, \ldots, I_N\}$ uma partição de X em intervalos onde a função T é contínua e monótona. Então, para todo $k \geq 1$, $\mathcal{A}^k = \bigvee_{i=0}^{k-1} T^{-i}(\mathcal{A})$ é uma partição finita de X em intervalos onde T ainda é contínua e monótona.*

Lema 5.2. *Para todo $k \geq 1$ e $A \in \mathcal{A}^k$ temos que $l(A) \leq d^k$ onde $d < 1$ é tal que $g : X \to (0, d]$. Em particular, $l(\{x\}) = 0$ para todo $x \in X$.*

Demonstração. Vamos provar por indução. Se $A \in \mathcal{A}^k$ podemos assumir que $A \subseteq I_j$ para algum $1 \leq j \leq N$. Então,

$$\begin{aligned} l(A) = l(\chi_A) = l(\mathcal{L}\chi_A) &= l\left(g(T_j^{-1}(x))\chi_A(T_j^{-1}(x))\chi_{T(I_j)}(x)\right) \\ &= l\left(g(T_j^{-1}(x))\chi_{T(A)}(x)\right) \leq dl(T(A)). \end{aligned}$$

Como $A \in \mathcal{A}^k$ temos que $T(A) \in \mathcal{A}^{k-1}$ e pela hipótese de indução $l(T(A)) \leq d^{k-1}$. Assim,

$$l(A) \leq l(T(A))d \leq d^{k-1}d = d^k.$$

□

Lema 5.3. *Seja* $Y = \bigcap\{F; F \subseteq X, F$ *é fechado e* $l(F) = 1\}$ *o suporte de* l. *A aplicação* T *pode ser mudada em um número finito de pontos de tal modo que* $T(Y) \subseteq Y$ *e* $T^{-1}(\{x\}) \subseteq Y$ *para todo, exceto por um número finito,* $x \in Y$.

Demonstração. Começamos provando que se $J \subseteq X$ é um intervalo então $l(J) = 0$ se e somente se $l(T(J)) = 0$. Se $l(J) = 0$ então pelo lema 5.1 temos que $l(T(J)) = 0$. Agora se $l(T(J)) = 0$ então

$$\begin{aligned} l(J) \leq l(T^{-1}(T(J))) &= l(\chi_{T(J)} \circ T) = l(\mathcal{L}(\chi_{T(J)} \circ T)) \\ &= l(\chi_{T(J)}\mathcal{L}(1)) \leq l(T(J))Nd \end{aligned}$$

pois $g \leq d$ e $\#T^{-1}(x) \leq N$ para todo $x \in X$.

Este resultado implica que se $x \in \text{int}(I_i)$ então $x \in Y$ se e somente se $T(x) \in Y$. Quando x é um ponto extremal do intervalo, sabemos do lema 5.2 que x não pode ser isolado em Y e assim redefinimos $T(x)$ de modo que seja o limite unilateral de $T(y)$ para $y \in Y$. Então, $T(Y) \subseteq Y$ o que finaliza a prova do lema. □

Teorema 5.1 (Helly, [V]). *Seja* $\varphi_n : X \to \mathbb{R}$, $n \geq 1$ *uma sequência de funções em* $BV[X]$ *e assuma que existam constantes* $K_1 > 0$ *e* $K_2 > 0$ *tais que* $sup|\varphi_n| \leq K_1$ *e* $var_X(\varphi_n) \leq K_2$ *para todos os* $n \geq 1$. *Então, existe uma subsequência* $(\varphi_{n_k})_k$ *e uma função* $\varphi_0 : X \to \mathbb{R}$ *com* $sup(\varphi_0) \leq K_1$ *e* $var_X(\varphi_0) \leq K_2$ *tal que* $(\varphi_{n_k})_k$ *converge a* φ_0 *quando* $k \to \infty$ l*-quase sempre em* $L^1(l)$.

Demonstração. Como $(\varphi_n)_n$ são funções de variação limitada podemos escrevê-las na forma $\varphi_n = \varphi_n^+ - \varphi_n^-$ onde $\varphi_n^\pm$ são sequências uniformemente limitadas de funções não-decrescentes [Ru]. Usando o processo da diagonal [V] podemos escolher $(n_k)_k$ tal que

$$\varphi^{\pm}_{n_k}(q) \to \varphi^{\pm}_0(q)$$

quando $k \to \infty$, para cada q em um conjunto denso S em X. Se $q < q'$ em S então

$$\varphi^{\pm}_{n_k}(q) \leq \varphi^{\pm}_{n_k}(q') \leq K_1$$

e assim, $\varphi^{\pm}_0(q) \leq \varphi^{\pm}_0(q') \leq K_1$ mostrando que $\varphi^{\pm}_0$ é uma função não-decrescente em S. Portanto podemos estender $\varphi^{\pm}_0$ a todo o conjunto X definindo

$$\varphi^{\pm}_0(x) = \inf_{q \in S, q > x} \varphi^{\pm}_0(q).$$

Vamos agora provar que $\varphi^{\pm}_{n_k}(x)$ converge a $\varphi^{\pm}_0(x)$ quando $k \to \infty$, para todo ponto de continuidade de $\varphi^{\pm}_0$. De fato, dados x um ponto de continuidade de $\varphi^{\pm}_0$ e $\varepsilon > 0$ existe $\delta > 0$ tal que $|x - x'| < \delta$ implica que $|\varphi^{\pm}_0(x) - \varphi^{\pm}_0(x')| < \varepsilon/2$. Escolha $q, q' \in S$ tal que $q' < x < q$ e $|q - q'| < \delta/2$. Então,

$$\varphi^{\pm}_0(q') < \varphi^{\pm}_0(x) < \varphi^{\pm}_0(q).$$

Agora seja $k_0 > 0$ tal que se $k > k_0$

$$\varphi^{\pm}_0(q') - \frac{\delta}{2} < \varphi^{\pm}_{n_k}(q') < \varphi^{\pm}_0(x) < \varphi^{\pm}_{n_k}(q) < \varphi^{\pm}_0(q) + \frac{\delta}{2}.$$

Então, se $k > k_0$

$$\varphi^{\pm}_{n_k}(q') < \varphi^{\pm}_{n_k}(x) < \varphi^{\pm}_{n_k}(q)$$

e portanto

$$|\varphi^{\pm}_{n_k}(x) - \varphi^{\pm}_0(x)| < \varepsilon, \qquad \forall k \geq k_0.$$

Assim, $\varphi^{\pm}_{n_k}$ converge para $\varphi^{\pm}_0$ exceto em um conjunto contável. Como $l(\{x\}) = 0$ segue que $\varphi^{\pm}_{n_k}$ converge para $\varphi^{\pm}_0$ l-quase sempre. Seja $\widetilde{\varphi^{\pm}_0}$ uma função contínua pela direita coincidindo com $\varphi^{\pm}_0$ em todo ponto de continuidade de $\varphi^{\pm}_0$. Logo, φ_n converge para $\varphi_0 = \widetilde{\varphi^{+}_0} - \widetilde{\varphi^{-}_0}$ l-quase sempre em $L^1(l)$. Além disso,

$$|\varphi_0(x)| \leq \lim_{k \to \infty} |\varphi_{n_k}(x)| \leq K_1$$

$$\sum_{j=1}^{N} |\varphi_0(a_j) - \varphi_0(a_{j-1})| = \lim_{k \to \infty} \sum_{j=1}^{N} |\varphi_{n_k}(a_j) - \varphi_{n_k}(a_{j-1})|$$

$$\leq \sup_k \mathrm{var}_X(\varphi_{n_k}) \leq K_2$$

sempre que x e $a_1 < a_2 < \ldots < a_N$ estão em X exceto para um conjunto contável. Como φ_0 é contínua pela direita segue que

$$|\varphi_0(x)| \leq K_1, \qquad \forall x \in X,$$

e

$$\mathrm{var}_X(\varphi_0) \leq K_2$$

$\square$

Lema 5.4. *i) Para todo $c > 0$ o conjunto $E = \{f \in L^1(l, X) : \|f\|_{var} \leq c\}$ é compacto em $(L^1(l), \|\cdot\|_1)$;*
ii) $(BV[X], \|\cdot\|_{var})$ é um espaço de Banach;
iii) $BV[X]$ é denso em $(L^1(l), \|\cdot\|_1)$.

Demonstração. i) Seja $(f_n)_n \subseteq E$. Então, para cada $\varepsilon > 0$ existe $(\widetilde{f}_n)_n$ uma sequência de funções tal que

$$\mathrm{var}(\widetilde{f}_n) \leq \mathrm{var}(f_n) + \varepsilon, \qquad \forall n.$$

Assim, $\widetilde{f}_n : X \to \mathbb{R}$ é uma função em $L^1(l)$ tal que

$$\mathrm{var}(\widetilde{f}_n) + \|\widetilde{f}_n\|_1 \leq c + \varepsilon.$$

Afirmamos que

$$\sup|\widetilde{f}_n| \leq \mathrm{var}(\widetilde{f}_n) + \|\widetilde{f}_n\|_1.$$

Para provar isto, sejam $\varepsilon_0 > 0$ qualquer e $\bar{x} \in X$ tal que $|\widetilde{f}_n(\bar{x})| \geq \sup|\widetilde{f}_n| + \varepsilon_0$. Agora seja $y \in X$ tal que $|\widetilde{f}_n(y)| \leq \int_X |\widetilde{f}_n| dl$ então

$$\sup|\widetilde{f}_n| + \varepsilon_0 \leq |\widetilde{f}_n(\bar{x})| \leq |\widetilde{f}_n(\bar{x}) - \widetilde{f}_n(y)| + |\widetilde{f}_n(y)| \leq \mathrm{var}(\widetilde{f}_n) + \int_X |\widetilde{f}_n| dl$$

e como isto vale para qualquer ε_0 arbitrário, segue o resultado.

Assim,

$$\sup|\widetilde{f}_n| \leq \mathrm{var}(\widetilde{f}_n) + \int_X |\widetilde{f}_n| dl \leq c + \varepsilon.$$

Aplicando o teorema de Helly a $(\widetilde{f}_n)_n$ obtemos uma subsequência de funções $(\widetilde{f}_{n_k})_k : X \to \mathbb{R}$ e uma função $f_0 : X \to \mathbb{R}$ tal que $\sup|f_0| \leq c+\varepsilon$ e $\mathrm{var}(f_0) \leq c+\varepsilon$ além disso,

$$\widetilde{f}_{n_k} \to f_0, \qquad \text{quando } k \to \infty$$

l-quase sempre em $L^1(l)$. Da prova do teorema de Helly segue que

$$\operatorname{var}(f_0) = \lim_{k\to\infty} \operatorname{var}(f_{n_k})$$

$$\|f_0\|_1 = \lim_{k\to\infty} \|\widetilde{f}_{n_k}\|_1$$

e portanto,

$$\|f_0\|_{\text{var}} \leq c + \varepsilon, \qquad \forall \varepsilon.$$

Como $\varepsilon > 0$ é arbitrário existe $\widetilde{f}_0 : X \to \mathbb{R}$ tal que $\|\widetilde{f}_0\|_{\text{var}} \leq c$ e $\|f_0\|_1 = \|\widetilde{f}_0\|$ provando que E é um conjunto compacto.

ii) Exceto pelo completamento, as outras propriedades de um espaço de Banach são óbvias. Assim, seja $(f_n)_n \subseteq BV[X]$ uma sequência de Cauchy. Então para cada $\varepsilon > 0$ existe $N > 0$ tal que se $m, n > N$

$$\|f_n - f_m\|_{\text{var}} < \varepsilon.$$

Disto segue que $\|f_n - f_m\|_1 < \varepsilon$ se $n, m > N$ e portanto $(f_n)_n$ é uma sequência de Cauchy em $L^1(l, X)$. Assim, existe uma constante $c > 0$ tal que

$$f_n \in E_c = \{f \in L^1(l, X); \|f\|_{\text{var}} \leq c\}.$$

Como por i) E_c é compacto em $L^1(l, X)$, existe uma $f \in E_c$ tal que

$$f_n \to f \qquad \text{em} \qquad L^1(l, X)$$

l-quase sempre. Usando o mesmo argumento da prova do teorema de Helly podemos provar que

$$\operatorname{var}(f_n - f) \to 0 \qquad \text{quando} \qquad n \to \infty$$

e então

$$\|f_n - f\|_{\text{var}} \to 0 \qquad \text{quando} \qquad n \to \infty$$

provando que $BV[X]$ é completo, sendo portanto um espaço de Banach.

iii) Como $\bigcup_{n=1}^{\infty} \mathcal{A}^n$ gera a σ-álgebra de Borel, l-quase sempre, então as funções simples em $\mathcal{A}^n$ são densas em $(L^1(l, X), \|\cdot\|_1)$. $\square$

5.2 Propriedades Espectrais de $\mathcal{L}$ e U_T

Vamos iniciar com a seguinte fórmula importante

$$\begin{aligned}\mathcal{L}(f(\varphi\circ T))(x) &= \sum_{i=1}^{n} g(T_i^{-1}(x))f(T_i^{-1}(x))\varphi\circ T(T_i^{-1}(x))\chi_{T(I_i)}(x)\\ &= \left(\sum_{i=1}^{n} g(T_i^{-1}(x))f(T_i^{-1}(x))\chi_{T(I_i)}(x)\right)\varphi(x)\\ &= \varphi(x)\mathcal{L}(f)(x)\end{aligned}$$

para $f,\varphi\in B_l(X)$, ou seja,

$$\mathcal{L}(f(\varphi\circ T)) = \varphi\mathcal{L}(f).$$

No próximo lema vamos estender $\mathcal{L}$ ao espaço $L^1(l,Y)$.

Lema 5.5. *Seja $\mathcal{F}$ o conjunto de todas as funções a valores complexos limitadas em Y e que são mensuráveis com respeito à σ-álgebra de Borel $\mathcal{B}$. Para toda $f\in\mathcal{F}$ vale*

$$\|\mathcal{L}(f)\|_1 \leq \|f\|_1.$$

Demonstração. Se $f\in\mathcal{F}$ defina $\varphi(x) = \mathcal{L}(f)(x)/|\mathcal{L}(f)(x)|$ quando $\mathcal{L}(f)(x)\neq 0$ e $\varphi(x)=1$ caso contrário. Então, para todo $x\in Y$ temos que $|\varphi(x)|=1$. Assim,

$$\begin{aligned}\|\mathcal{L}(f)\|_1 &= \int_Y |\mathcal{L}(f)|dl = \int_Y \mathcal{L}(f)\varphi dl = \int_Y \mathcal{L}(f\varphi\circ T)dl\\ &= \int_Y f\varphi\circ T dl \leq \int_Y |f||\varphi\circ T|dl \leq \int_Y |f|dl\\ &= \|f\|_1\end{aligned}$$

□

De acordo com o lema 5.1 qualquer mudança de f em um conjunto de medida nula (com respeito à medida l) mudará $\mathcal{L}(f)$ em um conjunto de medida também nula (com respeito à medida l). Assim, $\mathcal{L}$ pode ser visto como um operador agindo no conjunto de funções limitadas em $L^1(l,Y)$. Fazendo uso do lema 5.5 podemos estender $\mathcal{L}$ para que atue no espaço $L^1(l,Y)$. De agora em diante, vamos assumir que

$$\mathcal{L}: L^1(l,Y)\to L^1(l,Y).$$

Defina $g_n(x) = \prod_{j=0}^{n-1} g(T^j(x))$. Como $0<g(x)\leq d<1$, $\forall x$, podemos escolher $M\in\mathbb{N}$ tal que

$$\|g_M\|_\infty = \sup\{g_M(x); x\in Y\} < \frac{1}{3}.$$

O lema seguinte é o resultado principal sobre o operador $\mathcal{L}$ e expressa o fato que o operador de Ruelle-Perron-Frobenius melhora a regularidade de funções de variação limitada com respeito a suas variações.

Lema 5.6. *Existem* $0 < \alpha < 1$ *e* $\beta > 0$ *tal que para toda* $f \in L^1(l, Y)$

$$var(\mathcal{L}^M(f)) \leq \alpha var(f) + \beta \|f\|_1.$$

Demonstração. Escrevendo Q para $\mathcal{L}^M$ e U para T^M temos que

$$Q(f)(x) = \sum_{i=1}^{N} g_M(U_i^{-1}(x)) f(U_i^{-1}(x)) \chi_{U(J_i)}(x)$$

onde $(J_1, J_2, \ldots, J_N)$ é uma partição de Y em intervalos onde $U|_{J_i}$ é monótona, contínua e $\mathrm{var}(g_M|_{J_i}) \leq \|g_M\|_\infty$. Para achar a última exigência precisaremos subdividir a partição de U em intervalos onde U é monótona e contínua. Para $\varepsilon > 0$ associe $\widetilde{f}$ na classe de equivalência definida por $f \in L^1(l, Y)$ tal que

$$\mathrm{var}(\widetilde{f}) \leq \mathrm{var}(f) + \varepsilon.$$

Seja $K_i = U(J_i) = U([a_i, b_i])$. Então

$$\begin{aligned}
\mathrm{var}_Y(Q(\widetilde{f})) &= \mathrm{var}_Y \left(\sum_{i=1}^{N} g_M(U_i^{-1}(x)) \widetilde{f}(U_i^{-1}(x)) \chi_{K_i}(x) \right) \\
&\leq \sum_{i=1}^{N} \mathrm{var}_Y \left(g_M(U_i^{-1}(x)) \widetilde{f}(U_i^{-1}(x)) \chi_{K_i}(x) \right) \\
&\leq \sum_{i=1}^{N} \mathrm{var}_{K_i}(g_M \widetilde{f}) + \|g_M\|_\infty \sum_{i=1}^{N} (|\widetilde{f}(b_i)| + |\widetilde{f}(a_i)|).
\end{aligned}$$

No sentido de calcular o segundo termo, para $x_i \in (a_i, b_i)$

$$\begin{aligned}
|\widetilde{f}(b_i)| + |\widetilde{f}(a_i)| &\leq |\widetilde{f}(b_i) - \widetilde{f}(x_i)| + |\widetilde{f}(a_i) - \widetilde{f}(x_i)| + 2|\widetilde{f}(x_i)| \\
&\leq \mathrm{var}_{[a_i, b_i]}(\widetilde{f}) + 2|\widetilde{f}(x_i)|.
\end{aligned}$$

Como esta desigualdade vale para todo $x_i \in (a_i, b_i)$ segue que

$$|\widetilde{f}(b_i)| + |\widetilde{f}(a_i)| \leq \mathrm{var}_{[a_i, b_i]}(\widetilde{f}) + 2d_i$$

onde $d_i = \inf\{|\widetilde{f}(x)|; x \in J_i\}$. Do teorema do valor médio para integrais [HS]

$$d_i \leq l([a_i, b_i])^{-1} \int_{a_i}^{b_i} |\widetilde{f}| dl$$

e portanto,

$$d_i \leq b^{-1} \int_{J_i} |\widetilde{f}| dl$$

onde $b = \min\{l([a_i, b_i]); 1 \leq i \leq N\}$. Assim,

$$\sum_{i=1}^{N} (|\widetilde{f}(b_i)| + |\widetilde{f}(a_i)|) \leq \sum_{i=1}^{N} \mathrm{var}_{J_i}(\widetilde{f}) + 2b^{-1} \int_{Y} |\widetilde{f}| dl.$$

Do primeiro termo se $x_1 < y_1 < x_2 < \ldots < y_r$ é uma partição de J_i então

$$\begin{aligned}
&\sum_{k=1}^{r} |\widetilde{f}(y_k) g_M(y_k) - \widetilde{f}(x_k) g_M(x_k)| \\
&\qquad \leq \sum_{k=1}^{r} |\widetilde{f}(y_k) g_M(y_k) - \widetilde{f}(y_k) g_M(x_k)| + \sum_{k=1}^{r} |\widetilde{f}(y_k) g_M(x_k) - \widetilde{f}(x_k) g_M(x_k)| \\
&\qquad \leq \sum_{k=1}^{r} |\widetilde{f}(y_k)| |g_M(y_k) - g_M(x_k)| + \sum_{k=1}^{r} |g_M(x_k)| |\widetilde{f}(y_k) - \widetilde{f}(x_k)| \\
&\qquad \leq \sum_{k=1}^{r} |\widetilde{f}(y_k)| \mathrm{var}_{[x_k, y_k]}(g_M) + \|g_M\|_\infty \mathrm{var}_{J_i}(\widetilde{f}).
\end{aligned}$$

Mas,

$$\begin{aligned}
\sum_{k=1}^{r} |\widetilde{f}(y_k)| \mathrm{var}_{[x_k, y_k]}(g_M) &\leq \sum_{k=1}^{r} (|\widetilde{f}(y_k) - d_i| + |d_i|) \mathrm{var}_{[x_k, y_k]}(g_M) \\
&\leq \sum_{k=1}^{r} (\mathrm{var}_{J_i}(\widetilde{f}) + d_i) \mathrm{var}_{[x_k, y_k]}(g_M) \\
&\leq \sum_{k=1}^{r} \left(\mathrm{var}_{J_i}(\widetilde{f}) + b^{-1} \int_{a_i}^{b_i} |\widetilde{f}| dl \right) \mathrm{var}_{[x_k, y_k]}(g_M) \\
&\leq \left(\mathrm{var}_{J_i}(\widetilde{f}) + b^{-1} \int_{a_i}^{b_i} |\widetilde{f}| dl \right) \mathrm{var}_{J_i}(g_M) \\
&\leq \|g_M\|_\infty \left(\mathrm{var}_{J_i}(\widetilde{f}) + b^{-1} \int_{J_i} |\widetilde{f}| dl \right).
\end{aligned}$$

Disto segue que

$$\begin{aligned}
\sum_{k=1}^{r} |\widetilde{f}(y_k) g_M(y_k) - \widetilde{f}(x_k) g_M(x_k)| \leq \|g_M\|_\infty &\left(\mathrm{var}_{J_i}(\widetilde{f}) + b^{-1} \int_{J_i} |\widetilde{f}| dl \right) \\
&+ \|g_M\|_\infty \mathrm{var}_{J_i}(\widetilde{f})
\end{aligned}$$

e então

$$\mathrm{var}_{J_i}(\widetilde{f} g_M) \leq \|g_M\|_\infty \left(\mathrm{var}_{J_i}(\widetilde{f}) + b^{-1} \int_{J_i} |\widetilde{f}| dl \right) + \|g_M\|_\infty \mathrm{var}_{J_i}(\widetilde{f})$$

somando sobre os valores de i temos

$$\sum_{i=1}^{N} \operatorname{var}_{J_i}(\widetilde{f} g_M) \leq \|g_M\|_\infty \left(\operatorname{var}_Y(\widetilde{f}) + b^{-1} \int_Y |\widetilde{f}| dl \right) + \|g_M\|_\infty \operatorname{var}_Y(\widetilde{f}).$$

Juntando essas estimativas temos

$$\begin{aligned} \operatorname{var}(Q(\widetilde{f})) &\leq 2\|g_M\|_\infty \left(\operatorname{var}_Y(\widetilde{f}) + b^{-1} \int_Y |\widetilde{f}| dl \right) + \|g_M\|_\infty \operatorname{var}_Y(\widetilde{f}) + 2b^{-1} \int_Y |\widetilde{f}| dl \\ &= 3\|g_M\|_\infty \operatorname{var}_Y(\widetilde{f}) + 3b^{-1}\|g_M\|_\infty \|\widetilde{f}\|_1 \\ &= \alpha \operatorname{var}_Y(\widetilde{f}) + \beta|\widetilde{f}\|_1 \end{aligned}$$

onde $\alpha = 3\|g_M\|_\infty < 1$ e $\beta = 3b^{-1}\|g_M\|_\infty$ para $b = \min\{l(J_i); 1 \leq i \leq N\}$.

Como, claramente, $Q(\widetilde{f})$ é equivalente a $Q(f)$ segue que

$$\begin{aligned} \operatorname{var}(Q(f)) \leq \operatorname{var}(Q(\widetilde{f})) &\leq \alpha \operatorname{var}_Y(\widetilde{f}) + \beta|\widetilde{f}\|_1 \\ &= \alpha \operatorname{var}_Y(f) + \beta\|f\|_1 + \varepsilon \end{aligned}$$

sendo ε arbitrátio segue o resultado. □

O próximo teorema é uma consequência imediata do teorema ergódico de Ionescu Tulcea e Marinescu [ITM] que enunciamos, sem prova. Para isto sejam E um espaço de Banach linear sobre o corpo dos complexos e B um subespaço de Banach de E.

Seja $C(B, E)$ a classe dos operadores lineares $T : B \to T(B) \subseteq B$ limitados satisfazendo

i)existe uma constante positiva H tal que para todo $n \geq 1$, $\|T^n\|_B \leq H$;

ii) existem duas constantes positivas R e $p < 1$ tais que $\|T(x)\|_B \leq p\|x\|_B + R\|x\|_E$ para qualquer que seja $x \in B$;

iii) se $P \subseteq B$ é um subconjunto limitado então $T(P)$ é um compacto em E.

Teorema 5.2 (Ionescu Tulcea e Marinescu). *Seja T um operador em $C(B, E)$. Então,*

i) não existe mais que um número finito de autovalores, $\lambda_1, \lambda_2, \ldots, \lambda_r$, de módulo 1 de T;

ii) *podemos escrever para todo* $n \geq 1$,

$$T^n = \sum_{i=1}^{r} (1/\lambda_i^n) T_i + S^n$$

onde T_i, *para* $1 \leq i \leq r$, *são operadores limitados de* B *em uma parte linear de* B *com dimensão finita e* $S : B \to S(B) \subseteq B$ *um operador linear limitado;*

iii) *para todo* i, $T_i^2 = T_i$, *para* $i \neq j$ *vale que* $T_i T_j = 0$ *e para todo* i, $T_i S = S T_i = 0$;

iv) *para todo* $n \geq 1$, $\|S^n\| \leq M/(1+h)^n$ *onde* M *e* h *são duas constantes positivas.*

Teorema 5.3. *Com as hipóteses do teorema de Ionescu Tulcea e Marinescu para* $\mathcal{L}$, T, X *e* l *temos*

i) $\mathcal{L} : L^1(l, X) \to L^1(l, X)$ *tem apenas um número finito de autovalores* $\lambda_1, \lambda_2, \ldots, \lambda_r$ *de módulo* 1;

ii) *seja* $E_i = \{f \in L^1(l, X); \mathcal{L}(f) = \lambda_i f\}$ *para* $1 \leq i \leq r$. *Então,* $E_i \subset BV[Y]$ *e* $\dim E_i < +\infty$;

iii) *o operador* $\mathcal{L}$ *pode ser representado por*

$$\mathcal{L} = \sum_{i=1}^{r} \lambda_i \phi_i + \psi$$

onde ϕ_i *são projeções sobre os autoespaços* E_i, $\|\phi_i\| \leq 1$ *e* ψ *é um operador linear em* $L^1(l, X)$ *com* $\sup\{\|\psi^n\|_1; n \geq 1\} < +\infty$. *Além disso,* $\phi_i \phi_j = 0$ *para* $i \neq j$ *e* $\phi_i \psi = 0$ *para todo* i;

iv) $\psi(BV[Y]) \subset BV[Y]$ *e* ψ *visto como um operador linear em* $(BV, \|\cdot\|_{var})$ *satisfaz* $\|\psi^n\|_{var} \leq H q^n$ $(n \geq 1)$ *para constantes* H *e* q *com* $H > 0$ *e* $0 < q < 1$;

v) 1 *é um autovalor de* $\mathcal{L}$. *Vamos assumir que* $\lambda_1 = 1$;

vi) *Escrevendo* $h = \phi_1(1)$ *e* $\mu = hl$ *temos* $h \geq 0$, $\int hdl = 1$ *e* $\mathcal{L}(h) = h$. *Então* μ *é uma probabilidade* T*-invariante em* Y *e portanto em* X. *Se* μ' *é* T*-invariante e* $\mu' \ll l$ *então* $\mu' \ll \mu$.

Demonstração. Do item i) ao iv) é imediato do teorema de Ionescu Tulcea e Marinescu. De iii) segue que

$$\begin{aligned}
\mathcal{L}^2(f) &= \sum_{i=1}^{r} \lambda_i \phi_i(\mathcal{L}(f)) + \psi(\mathcal{L}(f)) \\
&= \sum_{i=1}^{r} \lambda_i \phi_i \left(\sum_{j=1}^{r} \lambda_j \phi_j(f) + \psi(f) \right) + \psi \left(\sum_{j=1}^{r} \lambda_j \phi_j(f) + \psi(f) \right) \\
&= \sum_{i,j=1}^{r} \lambda_i \lambda_j \phi_i \phi_j(f) + \lambda_i \phi_i \psi(f) + \sum_{j=1}^{r} \lambda_j \psi \phi_j(f) + \psi^2(f) \\
&= \sum_{i=1}^{r} \lambda_i^2 \phi_i(f) + \psi^2(f)
\end{aligned}$$

onde usamos que $\phi_i^2 = \phi_i$ e $\phi_i \phi_j = 0$, $i \neq j$ e $\phi_i \psi = 0$, $\forall i$. Indutivamente temos que $\forall k \in \mathbb{N}$

$$\mathcal{L}^k(f) = \sum_{i=1}^{r} \lambda_i^k \phi_i(f) + \psi^k(f).$$

Seja $h_n = 1/n \sum_{k=0}^{n-1} \mathcal{L}^k(1)$. Então $h_n \geq 0$ e $\int_Y h_n dl = 1$

$$h_n = \frac{1}{n} \sum_{k=0}^{n-1} \left(\sum_{i=1}^{r} \lambda_i^k \phi_i(1) + \psi^k(1) \right) = \sum_{i=1}^{r} \left(\frac{1}{n} \sum_{k=0}^{n-1} \lambda_i^k \right) \phi_i(1) + \frac{1}{n} \sum_{k=0}^{n-1} \psi^k(1)$$

donde

$$h_n = \sum_{i=1}^{r} \gamma_n(i) \phi_i(1) + \frac{1}{n} \sum_{k=0}^{n-1} \psi^k(1)$$

onde $\gamma_n(i) = \frac{1}{n} \sum_{k=0}^{n-1} \lambda_i^k$. De iv), como $1 \in BV(Y)$ segue que

$$\mathrm{var}(\psi^k(1)) + \|\psi^k(1)\|_1 \leq H q^k$$

e portanto

$$\begin{aligned}
\left\| \frac{1}{n} \sum_{k=0}^{n-1} \psi^k(1) \right\|_1 &\leq \frac{1}{n} \sum_{k=0}^{n-1} \|\psi^k(1)\|_1 \leq \frac{1}{n} \sum_{k=0}^{n-1} H q^k \\
&= \left(H \frac{1-q^n}{1-q} \right) \frac{1}{n} \to 0 \qquad \text{quando } n \to \infty.
\end{aligned}$$

Agora vamos analisar o primeiro termo de h_n. Para cada i tal que $\lambda_i \neq 1, -1$ temos

$$\gamma_n(i) = \frac{1}{n} \sum_{k=0}^{n-1} \lambda_i^k = \frac{1}{n} \left(\frac{1-\lambda^n}{1-\lambda} \right) \to 0 \qquad \text{quando } n \to \infty.$$

O mesmo vale quando $\lambda_i = -1$, i. e., $\gamma_n(i) \to 0$.

Assim, se $1 \notin \{\lambda_1, \ldots, \lambda_r\}$ então em $L^1(l, Y)$ vale que $h_n \to 0$. Pois $\int h_n dl = 1$ para todo n que não valha. Assim, $1 \in \{\lambda_1, \ldots, \lambda_r\}$ e assumimos que $\lambda_1 = 1$. Do argumento acima vemos que $\gamma_n(1) = 1$ e $\gamma_n(i) \to 0$ para $i \neq 1$. Logo, em $L^1(l, Y)$

$$h_n \to \phi_1(1).$$

Vamos denotar $h = \phi_1(1)$. Então, $h \geq 0$ e $\int_Y hdl = 1$. Além disso, h é invariante por $\mathcal{L}$. De fato,

$$\begin{aligned}
\mathcal{L}(h) = \mathcal{L}\left(\lim_{n\to\infty} h_n\right) &= \lim_{n\to\infty} \mathcal{L}(h_n) = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} \mathcal{L}^{k+1}(1) \\
&= \lim_{n\to\infty} \frac{1}{n}(1 + \mathcal{L}(1) + \cdots + \mathcal{L}^{n-1}(1) + \mathcal{L}^n(1) - 1) \\
&= \lim_{n\to\infty} h_n + \frac{\mathcal{L}^n(1) - 1}{n} = h
\end{aligned}$$

pois $\|\mathcal{L}^n(1)\| \leq \|1\|$. Além disso,

$$\begin{aligned}
\mu(f \circ T) &= l(hf \circ T) = l(\mathcal{L}(hf \circ T)) = l(f\mathcal{L}(h)) \\
&= l(fh) = \mu(f)
\end{aligned}$$

isto é, $\mu = hl$ é T-invariante. Finalmente, assuma que μ' é T-invariante e $\mu' \ll l$ então $\mu' = h'l$ para alguma $h' \in L^1(l, Y)$, e

$$\begin{aligned}
\mu'(f \circ T) &= \mu'(f) \\
\mu'(f \circ T) &= l(f\mathcal{L}(h')) \\
\mu'(f) &= l(fh')
\end{aligned}$$

então

$$\int f\mathcal{L}(h')dl = \int fh'dl$$

para todo $f \in L^1(l, Y)$ implicando que

$$\mathcal{L}(h') = h'.$$

Daí segue que $h' \in BV[Y]$ donde existe $c > 0$ tal que $h' \leq c$. Assim, $\forall n \geq 1$

$$h' = \frac{1}{n}\sum_{k=0}^{n-1} \mathcal{L}^k(h') \leq \frac{1}{n}\sum_{k=0}^{n-1} c\mathcal{L}^k(1) \leq c\frac{1}{n}\sum_{k=0}^{n-1} \mathcal{L}^k(1).$$

Portanto, $h' \leq ch'$ ou seja, $\mu' \ll \mu$ completando a prova do teorema. $\square$

Um modo de garantir a existência de medidas invariantes absolutamente contínuas com respeito à medida de Lebesgue m, para uma aplicação monótona por partes $T : [0,1] \to [0,1]$ é dada pelo teorema de Lasota e Yorke [LY2], que pode ser visto como um corolário do teorema 5.3 onde o potencial g é o inverso do determinante da derivada de T. Como T é expansora segue que $\sup(g) < 1$ e portanto, o potencial g satisfaz as condições do teorema de Lasota e Yorke. Além disso, o teorema de mudança de variável acarreta que $m(\mathcal{L}f) = m(f)$ donde as condições do teorema 5.3 são satisfeitas.

Antes de enunciar o teorema vamos relembrar a definição do operador de Ruelle-Perron-Frobenius. Sendo $T : [0,1] \to [0,1]$ não-singular definimos o operador $P_T : L_1 \to L_1$, por

$$P_T f(x) = \frac{d}{dx} \int_{T^{-1}([0,x])} f(s) ds.$$

Algumas propriedades de P_T são

a) P_T é positiva, i. e., $f \geq 0$ implica que $P_T f \geq 0$

b) $\int_0^1 P_T f dm = \int_0^1 f dm$, $\forall f \in L_1$

c) $P_{T^n} = P_T^n$.

d) $P_T f = f$ se e somente se a medida $d\mu = f dm$ é invariante por T.

Agora vamos enunciar, sem prova, o resultado de Lasota e Yorke [LY2]

Corolário 5.3.1. *Seja $T : [0,1] \to [0,1]$ uma aplicação C^2 por partes tal que* $\inf\{|T'(x)|; x \in [0,1]\} > 1$. *Então, para qualquer $f \in L_1$ a sequência*

$$\frac{1}{n} \sum_{j=0}^{n-1} P_T^n f$$

converge, em norma, para $f^* \in L_1$. A função limite tem as seguintes propriedades

1) $f \geq 0 \Rightarrow f^* \geq 0$

2) $\int_0^1 f^* dm = \int_0^1 f dm$

3) $P_T f^* = f^*$ *e, consequentemente, a medida $d\mu^* = f^* dm$ é invariante por T*

4) a função f^* é de variação limitada e existe uma constante c, independente da escolha inicial de f, tal que $var_{[0,1]} f^* \leq c\|f\|$.

Vamos apresentar um exemplo, que pode ser encontrado em [LY2], onde a hipótese $|T'(x)| > 1$ falha em apenas um ponto do domínio da T. Seja

$$T(x) = \begin{cases} x/(1-x), & 0 \leq x < 1/2 \\ 2x - 1, & 1/2 \leq x \leq 1 \end{cases}$$

onde $T'(0) = 1$. A sequência $P_T^n(f)$, $\forall f \in L^1$, converge em medida para zero, para uma definição de convergência em medida veja [DS] ou [Ca]. Também, $P_T(f) = f$ tem como única solução a trivial e não existe uma medida não trivial e absolutamente contínua com respeito à medida de Lebesgue que seja invariante por T.

A prova da convergência em medida se faz do seguinte modo, primeiro prova-se que para $f_0 \equiv 1$ a sequência $g_n(x) = xf_n(x)$ converge para uma constante c, onde $f_n = P_T^n(f_0)$. Então, usando que $\|f_n\| = 1$ obtemos $c = 0$ donde $f_n \to 0$ em medida. Por fim, por um argumento de aproximações podemos estender este resultado para uma sequência arbitrária $P_T^n(f)$ com $f \in L^1$.

5.3 Propriedades de (U_T, μ)

Os próximos lemas descrevem a adjunta, U_T^*, de U_T em termos de $\mathcal{L}$. Para simplificar a notação vamos utilizar L^2_μ no lugar de $L^2(\mu, X)$.

Lema 5.7. *Para toda $f \in L^2_\mu$ e $\lambda \in \mathbb{C}$ temos que*

i) $U_T^* f = \mathcal{L}(fh)/h$

ii) $U_T^* f = \lambda f$ *se e somente se* $\mathcal{L}(fh) = \lambda fh$.

Demonstração. i) segue das identidades

$$\begin{aligned} \langle U_T^* f, g\rangle = \langle f, U_T g\rangle &= \int f\overline{g} \circ T d\mu = \int f\overline{g} \circ Thdl = \int \mathcal{L}(f\overline{g} \circ Th)dl \\ &= \int \overline{g}\mathcal{L}(fh)dl = \int \overline{g}\frac{\mathcal{L}(fh)}{h}hdl = \int \overline{g}\frac{\mathcal{L}(fh)}{h}d\mu \\ &= \left\langle \frac{\mathcal{L}(fh)}{h}, g \right\rangle. \end{aligned}$$

O item ii) segue imediatamente de i). □

Lema 5.8. *Valem as seguintes afirmações* *i)* $(U_T^*)^n U_T^n = Id$ *para cada* $n \in \mathbb{N}$

ii) $U_T^n \circ (U_T^*)^n = \mathcal{L}_n$ *onde* $\mathcal{L}_n$ *denota a projeção ortogonal sobre o subespaço* $U_T^n(L_\mu^2)$.

Demonstração. Para provar i) note que

$$\langle U_T^* U_T f, g\rangle = \langle U_T f, U_T g\rangle = \langle f, g\rangle$$

e então $U_T^* U_T = Id$. Portanto, para cada $n \in \mathbb{N}$ vale que $(U_T^*)^n U_T^n = Id$.

Para provar o item ii) vamos provar primeiro que $U_T^n (U_T^*)^n$ é uma projeção

$$U_T^n \circ \underbrace{(U_T^*)^n \circ U_T^n}_{Id} \circ (U_T^*)^n = U_T^n \circ (U_T^*)^n$$

$$(U_T^n \circ (U_T^*)^n)^* = (U_T^*)^{n*} (U_T^n)^* = U_T^n \circ (U_T^*)^n.$$

Seja $E = U_T^n \circ (U_T^*)^n (L_\mu^2)$. É claro que $U_T^n(L_\mu^2) \subset E$. Agora, vamos assumir que $f \in E$ então existe $g \in L_\mu^2$ tal que

$$U_n^T \circ U_T^{*n} g = f$$

ou seja,

$$f = U_T^n \widehat{g}$$

onde $\widehat{g} = U_T^{*n} g$, i. e.,

$$f = U_T^n \widehat{g}$$

onde $\widehat{g} \in L_\mu^2$ e então

$$E \subset U_T^n(L_\mu^2).$$

Logo,

$$E = U_T^n(L_\mu^2).$$

$\square$

Teorema 5.4. *Seja* $esp(U_T) \subseteq \{z \in \mathbb{C}; |z| = 1\}$ *o conjunto dos autovalores de* $U_T f$. *Então,*

i) $esp(U_T) \subseteq \{\lambda_1, \ldots, \lambda_r\}$;

ii) $U_T f = \lambda f \iff U_T^* \overline{f} = \lambda \overline{f}$ *e* $|\lambda| = 1$;

iii) $\dim E_{(U_T,\mu)}(\lambda) < +\infty$ *para cada* $\lambda \in esp(U_T)$ *onde* $E_{(U_T,\mu)}(\lambda) = \{f \in L_\mu^2; U_T f = \lambda f\}$.

Demonstração. Vamos provar o item ii) primeiro.

[$\Rightarrow$] Se $U_T(f) = \lambda f$ então $U_T\overline{f} = \overline{\lambda} f$ e

$$U_T^* f = \lambda\overline{\lambda} U_T^* f = \lambda U_T^* \overline{\lambda f} = \lambda U_T^* U_T \overline{f} = \lambda \overline{f}.$$

[$\Leftarrow$] Agora, se $U_T^* f = \lambda f$ com $|\lambda| = 1$ então

$$\langle f, f \rangle = \lambda\overline{\lambda} \langle f, f \rangle = \langle \lambda f, \lambda f \rangle = \langle U_T^* f, \lambda f \rangle = \langle f, \lambda U_T f \rangle$$

e

$$\langle \lambda U_T f, \lambda U_T f \rangle = \langle U_T f, U_T f \rangle = \langle f, f \rangle$$

ou seja,

$$\langle f, f - \lambda U_T f \rangle = 0$$

e

$$\langle f - \lambda U_T f, \lambda U_T f \rangle = 0.$$

Após subtrairmos temos

$$\langle f - \lambda U_T f, f - \lambda U_T f \rangle = 0$$

e portanto $f = \lambda U_T f$, $U_T \overline{f} = \lambda \overline{f}$.

Para provar i) assuma $U_T f = \lambda f$. Usando o lema e o item iii) temos

$$U_T^* \overline{f} = \lambda \overline{f}, \qquad \mathcal{L}(h\overline{f}) = \lambda h \overline{f}$$

implicando que λ é um autovalor de $\mathcal{L}$, ou seja,

$$\operatorname{esp}(U_T) \subseteq \{\lambda_1, \ldots, \lambda_r\}.$$

Por outro lado, se $\mathcal{L}(g) = \lambda g$ para algum i, então $g \in BV$ e pelo teorema

$$|g| = |\lambda_i^n g| = |\mathcal{L}^n(g)| \leq \|g\|_\infty \mathcal{L}^n(1)$$
$$|g| \leq \|g\|_\infty \frac{1}{n} \sum_{j=0}^{n-1} \mathcal{L}^j(1) \to \|g\|_\infty h$$

isto dá $|g|/h \leq \|g\|_\infty$ e então $|g|/h \in L^2_\mu$. Usando o lema temos

$$U_T^* \left(\frac{g}{h} \right) = \frac{\mathcal{L}(\frac{g}{h} h)}{h} = \frac{\mathcal{L}(g)}{h} = \lambda_i \frac{g}{h}.$$

Finalmente, usando novamente o lema

$$U_T\overline{\left(\frac{g}{h}\right)} = \lambda_i\overline{\left(\frac{g}{h}\right)}$$

ou seja, $\lambda_i \in \text{esp}(U_T)$ donde

$$\{\lambda_1, \ldots, \lambda_r\} \subseteq \text{esp}(U_T)$$

terminando assim a prova de i). O mesmo argumento prova iii). $\square$

Definição 5.4. *Seja $T : X \to X$ uma transformação que preserva medida de um espaço de probabilidade $(X, \mathcal{A}, \mu)$. Dizemos que (T, μ) é fracamente mixing se para quaisquer $A, B \in \mathcal{A}$*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} |\mu(T^{-i}(A) \cap B) - \mu(A)\mu(B)| = 0.$$

Obs.: Temos que (T, μ) é fracamente mixing se e somente se T tem espectro contínuo, ou seja, 1 é seu único autovalor e suas autofunções são constantes. Uma prova deste fato pode ser encontrada em [W2].

Teorema 5.5. *Seja $\mathcal{B}_\infty = \bigcap_{n=1}^{\infty} T^{-n}(\mathcal{B})$ onde $\mathcal{B}$ denota a σ-álgebra de Borel em Y. Então,*

i) $\left\{f \in L^2_\mu; f \in \mathcal{B}_\infty \text{ mensurável}\right\} = \bigcap_{n=1}^{\infty} T^n(L^2_\mu) = \sum_{i=1}^{r} E_{(U_T,\mu)}(\lambda_i)$. Em particular $\mathcal{B}_\infty$ é finito μ-quase sempre.

ii) Se (T, μ) é fracamente mixing então $\mathcal{B}_\infty$ é trivial, i. e., (T, μ) é exata.

Demonstração. Para provar i) seja $\mathcal{B}_n = T^{-n}(\mathcal{B})$ e $E_n = \{f \in L^2_\mu; f \in \mathcal{B}_n \text{ mensurável}\}$. Provamos anteriormente que

$$E_n = \left\{f \in L^2_\mu; f = g \circ T^n \text{ onde } g \text{ é } B\text{-mensurável}, g \in L^2_\mu\right\}$$

e portanto $E_n = U_T^n(L^2_\mu)$ donde segue a primeira identidade.

Para provar a segunda igualdade tome $f \in \sum_{i=1}^{r} E_{(U_T,\mu)}(\lambda_i)$. Então,

$$f = \sum_{i=1}^{r} f_i, \qquad f_i \in E_{(U_T,\mu)}(\lambda_i).$$

Se $g = \sum_{i=1}^{r} \overline{\lambda_i^n} f_i$, então teremos que

$$\begin{aligned} U_T^n(g) &= \sum_{i=1}^{r} \overline{\lambda_i^n} T^n f_i = \sum_{i=1}^{r} \overline{\lambda_i^n} \lambda_i^n f_i \\ &= \sum_{i=1}^{r} f_i = f. \end{aligned}$$

Daí, segue que $f \in U_T^n(L_\mu^2)$ e assim

$$\sum_{i=1}^{r} E_{(U_T,\mu)}(\lambda_i) \subseteq U_T^n(L_\mu^2)$$

para todo $n \in \mathbb{N}$, e portanto

$$\sum_{i=1}^{r} E_{(U_T,\mu)}(\lambda_i) \subseteq \bigcap_{n=1}^{\infty} U_T^n(L_\mu^2).$$

ii) Se (T,μ) é fracamente mixing então 1 é o único autovalor e o autoespaço de autovetores associados a 1 contém apenas funções constantes. Assim,

$$\begin{aligned} \sum_{i=1}^{r} E_{(U_T,\mu)}(\lambda_i) &= E_{(U_T,\mu)}(1) = \{f \in L_\mu^2; f = \text{cte. } \mu \text{ q. s.}\} \\ &= \bigcap_{n=1}^{\infty} U_T^n(L_\mu^2) = \{f \in L_\mu^2; f \in \mathcal{B}_\infty \text{ mensurável}\} \end{aligned}$$

i. e., $\{f \in L_\mu^2; f \text{ é } B_\infty \text{ mensurável}\} = \{f = c\}$.

Assim, $\mathcal{B}_\infty = \{X, \emptyset\}$ provando que (T,μ) é exata. □

Obs.: As definições para esta observação se encontram no final da primeira seção do primeiro capítulo. A medida μ pode não ser ergódica, mas se μ' é uma componente ergódica de μ então $\text{supp}(\mu') = W \subseteq Y$ é um conjunto invariante por T, ou seja, $T^{-1}(W) = W$ e $\chi_W \circ T = \chi_W$. Definindo $l'(A) = l(A \cap W)/l(W)$ obtemos

$$\begin{aligned} l'(\mathcal{L}f) &= \frac{l(\chi_W \circ \mathcal{L}f)}{l(W)} = \frac{l(\mathcal{L}(f\chi_W \circ T))}{l(W)} \\ \frac{l(f\chi_W \circ T)}{l(W)} &= \frac{l(f\chi_W)}{l(W)} = l'(f) \end{aligned}$$

além disso, temos que

$$\mathcal{L}(h\chi_W) = \mathcal{L}(h\chi_W \circ T) = \chi_W \mathcal{L}(h) = \chi_W h = h'.$$

Assim, toda componente ergódica de μ é da forma $\mu' = h'l'$ para l' satisfazendo $l'(\mathcal{L}f) = l'(f)$ e h' tal que $\mathcal{L}(h') = h'$.

Considerando T^k, onde k é tal que $\lambda_i^k = 1$ para todo i, então uma componente ergódica de μ é fracamente mixing para T^k. De agora em diante vamos considerar T^k ao invés de T e g_k no lugar de g na definição de $\mathcal{L}$. Portanto podemos assumir que μ é fracamente mixing.

5.4 Estados de Equilíbrio

Vamos provar que g satisfaz

$$h_\mu(T) + \int \log g d\mu = \sup \left\{ h_\nu(T) + \int \log g d\nu \right\}$$

onde o supremo é tomado sobre todas as probabilidades ν que são T-invariantes em Y. Provar isto é o mesmo que provar que μ é um estado de equilíbrio para o potencial $\log g$. Para isto, vamos precisar de alguns lemas.

Definição 5.5. *Seja $f : Y \to \mathbb{R}$ definimos a parte positiva de f por* $f^+(x) = \max\{f(x), 0\}$.

Lema 5.9. *Se $f : Y \to \mathbb{R}$ é mensurável e ν é uma probabilidade T-invariante em Y, então $(f - f \circ T)^+ \in L^1_\nu$ implica que $f - f \circ T \in L^1_\nu$ e*

$$\int_Y (f - f \circ T) d\nu = 0.$$

Demonstração. Seja $F = (f - f \circ T)^+$ então $0 \leq F - (f - f \circ T) = F + f \circ T - f$ se definirmos $f_n : Y \to \mathbb{R}$ por $f_n(x) = \min\{f(x), n\}$ obtemos $0 \leq F + f_n \circ T - f_n$ e também

$$\lim_{n \to \infty} f_n \circ T - f_n = f \circ T - f.$$

Agora, usando o lema de Fatou [DS]

$$\begin{aligned}
\int (F + f \circ T - f) d\nu &= \int \liminf (F + f_n \circ T - f_n) d\nu \\
&\leq \liminf \int (F + f_n \circ T - f_n) d\nu \\
&= \int F d\nu < +\infty.
\end{aligned}$$

Daí, segue que $F + f \circ T - f \in L^1_\nu$ e portanto

$$f \circ T - f \in L^1_\nu.$$

Mas $|f_n \circ T - f_n| \leq |f \circ T - f|$ e pelo teorema da convergência dominada de Lebesgue [DS]

$$\int_Y (f \circ T - f) d\nu = \lim_{n\to\infty} \int_Y (f_n \circ T - f_n) d\nu = 0.$$

$\square$

Lema 5.10. *Se ν é uma probabilidade T-invariante em Y e*

$$\overline{g} = \frac{gh}{h \circ T}$$

então

$$\int \log \overline{g} d\nu = \int \log g d\nu.$$

Demonstração. Sabemos que $\mathcal{L}h = h$, ou seja, $\sum_{y \in T^{-1}(x)} g(y)h(y) = h(x)$. Então,

$$\begin{aligned}\sum_{y\in T^{-1}(x)} \overline{g}(y) &= \sum_{y \in T^{-1}(x)} \frac{g(y)h(y)}{h \circ T(y)} \\ &= \frac{1}{h(x)} \sum_{y\in T^{-1}(x)} g(y)h(y) = \frac{h(x)}{h(x)} = 1\end{aligned}$$

donde para todo $x \in Y$, $\sum_{y \in T^{-1}(x)} \overline{g}(y) = 1$. Desta igualdade e do fato que $\overline{g} \geq 0$ obtemos que $\log \overline{g} \leq 0$.

Se $\int \log \overline{g} d\nu = \int \log g d\nu = -\infty$ então não há nada para provar.

Se $\int \log g d\nu > -\infty$ então $\log g \in L^1_\nu$ e

$$\log \overline{g} = \log g + \log h - \log h \circ T$$

i. e.,

$$\log h - \log h \circ T = \log \overline{g} - \log g.$$

Assim,

$$(\log h - \log h \circ T)^+ = (\log \overline{g} - \log g)^+ \leq |\log g|$$

pois $\log g \leq 0$ ($g(x) \leq d < 1$.) Agora aplicamos o lema para obter

$$\log h - \log h \circ T \in L^1_\nu.$$

Então,

$$\log \overline{g} \in L^1_\nu$$

e usando novamente o lema

$$0 = \int (\log h - \log h \circ T) d\nu = \int \log \overline{g} d\nu - \int \log g d\nu.$$

Se $\int \log \overline{g} d\nu > -\infty$, usando o argumento acima para

$$\log h^{-1} - \log h^{-1} \circ T$$

temos o resultado. □

Lema 5.11. *A medida μ é um estado de equilíbrio para* $\log g$ *em* Y.

Demonstração. Defina o operador

$$\overline{\mathcal{L}}_{\overline{g}} f(x) = \sum_{y \in T^{-1}(x)} \overline{g}(y) f(y)$$

queremos aplicar o teorema de Ledrappier 3.1. Para isto, primeiro note que $\overline{g} \in \mathcal{G}$ e vamos verificar que $\overline{\mathcal{L}}^* \mu = \mu$. Esta última igualdade é válida pois

$$\begin{aligned} \overline{\mathcal{L}}^* \mu &= \int f d\overline{\mathcal{L}}^* \mu = \int \overline{\mathcal{L}} f d\mu \\ &= \int \sum_{y \in T^{-1}(x)} \frac{g(y) h(y)}{h \circ T(y)} f(y) h dl \\ &= \int \mathcal{L}(fh) dl = \int f h dl = \int f d\mu. \end{aligned}$$

□

5.5 Existência e exemplos

Para provarmos a existência de uma medida l e uma função g que satisfaçam

i) $l(\mathcal{L} f) = l(f)$, para qualquer $l \in \mathcal{F}$

ii) $g : X \to (0, d]$ com $d < 1$ e tal que $\mathrm{var}_X(g) < \infty$

precisaremos que o espaço X onde a transformação T monótona por partes está definida seja totalmente ordenado, munido da topologia da ordem, compacto e separável.

A existência da medida l será obtida ao aplicarmos o teorema de Schauder-Tychonoff à aplicação contínua

$$l \to \frac{\widetilde{\mathcal{L}}^* l}{\widetilde{\mathcal{L}}^* l(1)}$$

onde $\widetilde{\mathcal{L}}$ é o operador de Ruelle-Perron-Frobenius definido em X', que é uma modificação de X. Vamos iniciar pela definição de X'.

Seja $\mathcal{A} = \{I_1, \dots, I_N\}$ uma partição de X. Além disso, para $k \geq 1$ definimos $\mathcal{A}^k$ como no início deste capítulo. Sejam $\varphi \in C(X)$ de variação limitada satisfazendo $\sum_{i=1}^{\infty} \mathrm{var}_i(\varphi) < \infty$ onde $\mathrm{var}_i(\varphi) = \sup\{|\varphi(x) - \varphi(y)|; x, y \in A, A \in \mathcal{A}^i\}$. Seja $\widetilde{W} = \bigcup_{j=0}^{\infty} T^{-j}(E)$ onde

$$E = \left\{ \lim_{t \uparrow x} T^k(t), \lim_{t \downarrow x} T^k(t); k \geq 1 \text{ e } x \in \partial(I_i), 1 \leq i \leq N \right\},$$

onde $\partial(I_i)$ é a fronteira do intervalo I_i.

Seja W o conjunto, aberto e fechado, dos pontos de $\widetilde{W}$ que não são pontos extremos de intervalos. Considere $X' = X \cup W'$ onde W' é uma cópia de W disjunta de X. Em X' vamos estender a relação de ordem da seguinte forma, seja $y < x < z$ em X com $x \in W$ e seja $x' \in W'$ associado a x então, definimos $y < x < x' < z$ ou $y < x' < x < z$ tal que possamos estender T continuamente em X'. Sendo T contínua em X' e definindo $\widetilde{W}$ para X', todos os elementos de $\widetilde{W}$ são extremos de intervalos, que são abertos e fechados. A topologia da ordem de X' é compacta e separável. Também estendemos φ continuamente em X' para então definirmos o opeador $\widetilde{\mathcal{L}}$ por

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

para $x \in X'$.

Como os intervalos $T(I_i)$ para $1 \leq i \leq N$ são abertos e fechados em X', $\widetilde{\mathcal{L}}$ leva $C(X')$ em $C(X')$. Agora podemos aplicar o teorema de Schauder-Tychonoff na aplicação

$$l \to \frac{\widetilde{\mathcal{L}}^* l}{\widetilde{\mathcal{L}}^* l(1)}$$

definida no conjunto das probabilidades que é um subconjunto compacto e convexo do dual de $C(X')$ com a topologia fraca-*. Assim, obtemos um ponto fixo l tal que

$$l(\widetilde{\mathcal{L}} f) = \lambda l(f)$$

para todo $f \in C(X')$ onde $\lambda = \widetilde{\mathcal{L}}^* l(1)$.

Agora seja $\widetilde{g}(x) = e^{\varphi(x)}/\lambda$, queremos mostrar que $\|\widetilde{g}_n\|_\infty < 1$, para algum n, onde

$$\widetilde{g}_n(x) = \widetilde{g}(x)\widetilde{g}(T(x)) \cdots \widetilde{g}(T^{n-1}(x))$$

e também provar que $l(W') = 0$. Para isto sejam os lemas

Lema 5.12. *Seja A um boreliano de X' tal que $T^n|_A$ é monótona para todo n. Suponha que para todo $j \geq 1$ exista um boreliano B_j que é um subconjunto de algum elemento de $\mathcal{A}^{k_j}$ (definido no início deste capítulo) e que satisfaça $T^{k_j}(B_j) = A$, para algum k_j. Além disso, suponha que $B_i \cap B_j = \emptyset$ para $i \neq j$ e que existe um $r \in \mathbb{N}$ tal que B_j e A estejam contidos no mesmo elemento de $\mathcal{A}^{k_j - r}$ para todo $j \geq 1$. Então, $l(A) = 0$.*

Demonstração. Temos que

$$l(B_j) = \lambda^{-k_j} l(\widetilde{\mathcal{L}}^{k_j} \chi_{B_j}) \geq \inf\{\widetilde{g}_{k_j}(y); y \in B_j\} l(A)$$

e, similarmente,

$$l(A) \leq \sup\{\widetilde{g}_{k_j}(y); y \in A\} l(T^{k_j}(A)).$$

Isto implica que

$$l(B_j) l(T^{k_j}(A)) \geq \frac{\inf\{\widetilde{g}_{k_j}(y); y \in B_j\}}{\sup\{\widetilde{g}_{k_j}(y); y \in A\}} l(A)^2 \geq \exp\left(-\sum_{i=0}^{\infty} \mathrm{var}_i \varphi - 2r\|\varphi\|_\infty\right) l(A)^2$$

pois A e B_j estão contidos no mesmo elemento de $\mathcal{A}^{k_j - r}$. Como os B_j são disjuntos, temos que $l(B_j) \to 0$ quando $j \to \infty$. Claramente, $l(T^{k_j}(A)) \leq 1$ donde $l(A) = 0$. □

Lema 5.13. *Seja J um intervalo fechado que é um subconjunto de algum elemento de $\mathcal{A}^n$. Então,*

$$\sup\{\widetilde{g}_n; J\} \leq \frac{K l(J)}{l(T^n(J))}$$

onde $K = \exp\left(\sum_{i=1}^{\infty} var_i \varphi\right)$.

Demonstração. Temos que

$$\begin{aligned} l(J) = \lambda^{-n} l(\widetilde{\mathcal{L}}^n \chi_J) &\geq \inf\{\widetilde{g}_n; J\} l(T^n(J)) \\ &\geq \sup\{\widetilde{g}_n; J\} \exp\left(-\sum_{i=1}^{n} \mathrm{var}_i \varphi\right) l(T^n(J)) \\ &\geq K^{-1} \sup\{\widetilde{g}_n; J\} l(T^n(J)). \end{aligned}$$

□

Para finalizar, vamos aplicar os lemas anteriores em dois exemplos que são, o subshift unilateral do tipo finito e uma aplicação monótona por partes.

Exemplo 1. Neste exemplo mostraremos, para qualquer $\varphi : X \to \mathbb{R}$ de variação limitada que satisfaça $\sum_{i=1}^{\infty} \mathrm{var}_i(\varphi) < \infty$, a existência de uma probabilidade μ, um $\lambda > 0$ e $n \in \mathbb{N}$ tal que $l(\mathcal{L}f) = l(f)$ e $\|g\|_\infty < 1$ é satisfeita para $g(x) = \exp\left(\sum_{i=0}^{n-1} \varphi(T^i(x))\right)/\lambda^n$. Além disso, μ é um estado de equilíbrio para φ.

Seja $X = B^+(M)$ um subshift unilateral do tipo finito com N símbolos, vamos supor que X seja topologicamente transitivo, ou seja, a matriz de transição M é irredutível [M1]. Além disso, assumimos que a entropia topológica de X é positiva ou, em outras palavras, que X não é formado apenas de órbitas periódicas.

Se introduzimos em X a ordem lexicográfica, ou seja, dizemos que $\theta < \eta$ em X se ao menor $i \geq 0$ tal que $\theta(i) \neq \eta(i)$ tivermos que $\theta(i) < \eta(i)$. A transformação shift $\sigma : X \to X$ então se torna uma transformação monótona por partes. Para isto basta considerar os cilindros disjuntos $C(0; i)$, para $1 \leq i \leq N$, onde $X = \cup_{i=1}^{N} C(0; i)$ e σ restrito a cada cilindro é contínua.

Temos que $W = \emptyset$ e, consequentemente, $X' = X$. Vamos aplicar o lema 5.12 ao conjunto $A = \{\theta\}$ para $\theta \in X$. Como X não tem só órbitas periódicas, podemos achar uma imagem inversa $\eta_0, \eta_1, \ldots, \eta_{m-1}, \theta_0, \theta_1, \ldots$ de $\theta = \theta_0, \theta_1, \ldots$ que não é periódica. Sendo M irredutível existe um k_0 tal que para todo θ_j, $\theta_0 \in \{1, 2, \cdots, N\}$ existem $i_0 = \theta_j$, $i_1, i_1, \ldots, i_k = \eta_0$ com $k \leq k_0$ e $M(i_r, i_{r+1}) = 1$ para $0 \leq r \leq k-1$. Então, existe uma infinidade de pontos da forma

$$\theta_0, \theta_1, \ldots, \theta_j, i_1, i \ldots, i_{k-1}, \eta_0, \eta_1, \ldots, \eta_{m-1}, \theta_0, \theta_1, \ldots \in X$$

para $j \geq 1$. Eles podem ser usados no lema 5.12 como B_j com $r = m + k_0$ para obter $l(\{\theta\}) = 0$.

Seja $q_k = \max\{l(A) : A \in \mathcal{A}^k\}$, neste caso $\mathcal{A}^k$ é o conjunto de todos os cilindros $C(0; \eta_0, \eta_1, \ldots, \eta_{k-1})$. Temos que $q_k \to 0$ quando $k \to \infty$ pois, caso contrário, teríamos $\theta \in X$ e $\varepsilon > 0$ tal que cada vizinhança de θ contém um $A \in \mathcal{A}^k$, para algum k com $l(A) \geq \varepsilon$. Mas isto implicaria que $l(\{\theta\}) \geq \varepsilon$, uma contradição.

Agora vamos provar que o suporte, Y, de l é o espaço todo X. Existe, pelo menos, um i entre 1 e N tal que $l(C(0; i)) > 0$ pois, caso contrário, teríamos $l(X) = 0$. Escolha algum cilindro $C(0; j_0, \ldots, j_{n-1})$ e, como argumentado acima, podemos escolher $i_0 = j_{n-1}$, $i_1, \ldots, i_k = i$ com $M(i_r, i_{r+1}) = 1$. Então,

$$l(C(0; j_0, \ldots, j_{n-1}, i_1, \ldots, i_k)) \geq (\inf \widetilde{g})^{k+n-1} \, l(C(0; i_k)) > 0$$

donde $l(C(0; j_0, \ldots, j_{n-1})) > 0$. Isto implica que $Y = X$. Em particular, $l(C(0; i)) > 0$ para $1 \leq i \leq N$.

Agora seja $J = C(0; \theta_0, \ldots, \theta_{n-1})$ então $\sigma^n(J) \supseteq C(0; i)$ para algum i, donde

$$l(\sigma^n(J)) \geq \min\{l(C(0; i)); 1 \leq i \leq N\} = c > 0.$$

O lema 5.13 implica que $\sup\{\widetilde{g}_n; J\} \leq Kq_n/c$. Como J é um elemento arbitrário de $\mathcal{A}^n$, existe n com $\|\widetilde{g}_n\|_\infty < 1$. Tomando $g = \widetilde{g}_n$, temos que $\mathcal{L} = \widetilde{\mathcal{L}}^n/\mathcal{L}$ e os itens i) e ii) da seção anterior estão satisfeitos, assumindo que φ, $\widetilde{g}$ e g são de variação limitada.

Como $Y = X$ e do fato que $h_\mu(\sigma^n) = nh_\mu(\sigma)$ segue, pelo lema 5.11, que μ é um estado de equilíbrio para φ. $\square$

Exemplo 2. Alguns resultados usados neste exemplo estão provados no artigo [H1] que considera transformações monótonas por partes em $[0, 1]$ que tenham entropia topológica positiva. Assim, sejam $X = [0, 1]$, $T : X \to X$ monótona por partes com entropia topológica positiva. Seja φ a função identicamente nula então $\widetilde{g} = e^{\varphi(x)}/\lambda = \lambda^{-1}$. Se fizermos $g = \widetilde{g}$ e se $\lambda > 1$ então os itens i) e ii) do início desta seção serão satisfeitos.

Consideremos os subintervalos $I_i \subseteq X'$, $1 \leq i \leq N$, onde T é monótona. Dizemos que um intervalo K é um sucessor de um intervalo $J \subseteq I_i$ se K é um dos intervalos, não-vazios, da forma $T(J) \cap I_i$, $1 \leq i \leq N$. Definimos D como o conjunto que contém I_i, $1 \leq i \leq N$, e cada intervalo com seus respectivos sucessores. Como as imagens dos subintervalos I_i são subconjuntos abertos e fechados de X', todos os elementos de D são abertos e fechados. Definimos, também, M uma matriz com valores 0 ou 1 dada por $M(J, K) = 1$ se e somente se K é um sucessor para J onde $J, K \in D$. Elementos de D são sucessores de mais que um elemento de D. Se a matriz M não é irredutível podemos dividí-la em submatrizes irredutíveis L_l, $l \in \mathbb{N}$, tal que a cada L_l está associado um subconjunto $\Omega \subseteq [0, 1]$ fechado e invariante por T [H1].

Se $F \subseteq D$ é um conjunto de índices para L_l, ou seja, se $L_l = M|F$ então $I = \cup\{J; J \in \widetilde{F}\}$, onde

$$\widetilde{F} = \{J \in D; \exists J_0, J_1, \ldots, J_n = J \text{ com } J_0 \in F \text{ e } M(J_i, J_{i+1}) = 1\},$$

e $I' = \cup\{J; J \in \widetilde{F} - F\}$ são conjuntos invariantes por T que são uniões finitas de intervalos e

$$\Omega = \bigcap_{i=0}^{\infty} T^{-i}\overline{(I - I')}.$$

Vamos escolher Ω tal que $r(L) = r(M)$, onde r denota o raio espectral de L e M vistos como operadores agindo no espaço l^1. Sendo I invariante por T, $T|_I$ é ainda uma transformação monótona por partes. Assim, podemos considerar I como sendo X e assumir que a medida de Lebesgue l está concentrada em I.

Lema 5.14. *(a) Se $M(J_i, J_{i+1}) = 1$ para $0 \leq i \leq k-1$ então $T^k(J_0 \cap \cdots \cap T^{-k}(J_k)) = J_k$.*
(b) Para todo $J \in F$, temos que $l(J) > 0$.

Demonstração. (a) A prova é por indução. O processo de indução é dado por

$$\begin{aligned} T^k(J_0 \cap T^{-1}(J_1) \cap \cdots \cap T^{-k}(J_k)) &= T^{k-1}(T(J_0) \cap J_1 \cap \cdots \cap T^{-k+1}(J_k)) \\ &= T^{k-1}(J_1 \cap \cdots \cap T^{-k+1}(J_k)) \end{aligned}$$

pois $T(J_0) \supseteq J_1$, visto que J_1 é sucessor para J_0.

(b) Note que existe $K \in \widetilde{F}$ com $l(K) > 0$, pois $l(I) = 1$. Podemos achar $J_0 = J$, $J_1, \ldots, J_k = K$ com $M(J_i, J_{i+1}) = 1$ e $J \in F$. Como L é irredutível, podemos fazer isto para todos os $J \in F$. Seja $Z = J_0 \cap T^{-1}(J_1) \cap \ldots \cap T^{-k}(J_k) \subseteq J$. Então, usando o item anterior

$$l(Z) = \lambda^{-k} l(\widetilde{\mathcal{L}}^k \chi_Z) = \lambda^{-k} l(T^k(Z)) = \lambda^{-k} l(K) > 0.$$

Portanto, $l(J) > 0$ o que prova o item. □

Outro resultado do artigo [H1] é que $r(M) = \exp h_{top}(X)$. Assim, $r(L) = r(M) > 1$ e L consiste não somente de um ciclo. Entretanto podemos achar $J_0, J_1, \ldots, J_n = J_0 \in F$ e $K_0 = J_0$, $K_1, \ldots, K_k = J_r \in F$ para alguns $n, k \geq 0$ e $0 \leq r \leq n-1$ tal que $M(J_i, J_{i+1}) = M(K_i, K_{i+1}) = 1$. Agora seja

$$A_k = \bigcap_{i=0}^{k-1} T^{-ni} \left(\bigcap_{j=0}^{n-1} T^{-j}(J_j) \right), \qquad A = \bigcap_{k=0}^{\infty} A_k,$$

$$B_j = \bigcap_{i=0}^{jn} T^{-i}(J_{i (\text{mod } n)}) \cap \bigcap_{i=1}^{k} T^{-jn-i}(K_i) \cap \bigcap_{i=1}^{\infty} T^{-jn-k-i}(J_{r+i (\text{mod } n)}).$$

Os conjuntos B_j são disjuntos pois $T^{jn}(B_j) \subseteq K_1$, $T^{jn}(B_{j+m}) \subseteq J_1$ e K_1 e J_1 são intervalos disjuntos pois são sucessores distintos para J_0. Além disso, $T^{jn+k+n-r}(B_j) = A$ e A e B_j estão contidos no mesmo elemento de $\mathcal{A}^{jn}$. E segue do lema 5.12 que $l(A) = 0$. Agora, aplicando o lema 5.13 obtemos que

$$\sup_{A_k} \widetilde{g}_{kn} \leq \frac{K l(A_k)}{l(T^{kn}(A_k))}. \tag{5.1}$$

Como $T^{kn}(A_k) = T(J_{n-1}) \supseteq J_0$, pelo item (a) do lema 5.14, e $l(J_0) > 0$, pelo item (b) do lema 5.14, os lados direito e esquerdo da desigualdade 5.1 tendem a zero. Logo, existe um k com $\sup_{A_k} \widetilde{g}_{kn} < 1$. E, finalmente, como $\widetilde{g}_{kn} \equiv \lambda^{-kn}$ obtemos que $\lambda > 1$. Segue do lema 5.2 que $l(x) = 0$ para todo $x \in X'$, donde $l(W') = 0$.

Uma prova similar àquela do item (b) do lema 5.14 mostra que o suporte Y de l contém Ω. Como $h_{\text{top}}(\Omega) = \log r(L) = \log r(M) = h_{\text{top}}(X)$, segue do lema 5.11 que μ é uma medida com entropia máxima. $\square$

Capítulo 6

Apêndice

As cinco primeiras seções deste apêndice estão inteiramente contidas em [PY], a não ser pela construção da partição geradora para uma transformação expansora, a última seção deste apêndice é importante pois trata do cálculo da entropia para transformações expansoras e provamos alguns resultados usados na demonstração do teorema de Ledrappier. Outros textos que abordam entropia num sentido mais geral são [M1] e [W2].

6.1 Partições e Esperança Condicional

Começaremos com a definição de partição de um espaço de probabilidade. Para isto, usaremos a definição 1.10 de um conjunto de medida nula.

Definição 6.1. *Seja $\mathcal{P} = \{P_i\}_{i\in I}$ uma partição mensurável contável de um espaço de probabilidade $(X, \mathcal{A}, \mu)$, ou seja, $\mathcal{P}$ satisfaz*

i) $X = \bigcup\limits_{i\in I} P_i$ *a menos de um conjunto de medida nula, com respeito à μ;*

ii) $P_i \cap P_j = \emptyset$ *para quaisquer $i \neq j$, a menos de um conjunto de medida nula, com respeito à μ.*

De agora em diante $\mathcal{P} = \{P_i\}_{i\in I}$ denotará uma partição de um espaço de probabilidade. Os elementos P_i de uma partição são chamados de átomos da partição.

Definição 6.2. *Definimos a função informação* $I(\mathcal{P}) : X \to \mathbb{R}$ *em um espaço de probabilidade* $(X, \mathcal{A}, \mu)$ *por*

$$I(\mathcal{P})(x) = -\sum_{i \in I} \log \mu(P_i) \chi_{P_i}(x),$$

ou seja, $I(\mathcal{P})(x) = -\log \mu(P_i)$ *se* $x \in P_i$, $\forall i \in I$.

A esperança condicional $E(.|\mathcal{B}) : L^1(X, \mathcal{A}, \mu) \to L^1(X, \mathcal{B}, \mu)$ em um espaço de probabilidade com respeito à σ-álgebra $\mathcal{A}$ e a sub-σ-álgebra $\mathcal{B}$, definida em 3.5, tem algumas propriedades interessantes que enunciaremos agora.

Lema 6.1. *Sejam* $(X, \mathcal{A}, \mu)$ *um espaço de probabilidade e* $\mathcal{B}$ *uma sub-*σ*-álgebra. As principais propriedades de* $E(f|\mathcal{B})$ *são:*

i) $\int_B E(f|\mathcal{B}) d\mu = \int_B f d\mu$, *para todo* $B \in \mathcal{B}$ *e* $f \in L^1(X, \mathcal{A}, \mu)$

ii) $E(f|\mathcal{B}) \in L^1(X, \mathcal{B}, \mu)$

iii) se $f \in L^1(X, \mathcal{A}, \mu)$ *e* $g \in L^\infty(X, \mathcal{B}, \mu)$ *então* $E(fg|\mathcal{B}) = gE(f|\mathcal{B})$

iv) se $f \in L^1(X, \mathcal{A}, \mu)$ *e* $\mathcal{B}_2 \subseteq \mathcal{B}_1 \subseteq \mathcal{A}$ *são sub-*σ*-álgebras então* $E(E(f|\mathcal{B}_1)|\mathcal{B}_2) = E(f|\mathcal{B}_2)$

v) se $f \in L^1(X, \mathcal{A}, \mu)$ *então* $|E(f|\mathcal{B})| \leq E(|f||\mathcal{B})$, *e se* $f, g \in L^2(X, \mathcal{A}, \mu)$ *e* $1/p + 1/q = 1$ *então* $E(|fg||\mathcal{B}) \leq E(|f|^p|\mathcal{B})^{1/p} E(|g|^q|\mathcal{B})^{1/q}$

vi) se T *preserva* μ *então* $E(f|\mathcal{B})T = E(f \circ T|T^{-1}(\mathcal{B}))$, *onde* $T^{-1}(\mathcal{B}) = \{T^{-1}(B); B \in \mathcal{B}\}$.

Demonstração. Os itens i) e ii) são essencialmente a definição de esperança condicional e já foram apresentados no capítulo 3. A prova dos outros itens pode ser encontrada em [W2]. □

Definição 6.3. *Sejam* $(X, \mathcal{A}, \mu)$ *um espaço de probabilidade e* $\mathcal{P}$ *uma partição. Dada qualquer sub-*σ*-álgebra* $\mathcal{B} \subseteq \mathcal{A}$ *definimos a função informação condicional* $I(\mathcal{P}|\mathcal{B}) : X \to \mathbb{R}$ *por*

$$I(\mathcal{P}|\mathcal{B}) = -\sum_{i \in I} \log \mu(P_i|\mathcal{B})(x) \chi_{P_i}(x)$$

onde $\mu(P_i|\mathcal{B})(x) = E(\chi_{P_i}|\mathcal{B})(x)$ *é chamada de medida condicional.*

As propriedades seguintes são uma consequência imediata da definição.

Lema 6.2. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade, $\mathcal{P}$ uma partição e $\mathcal{B} \subseteq \mathcal{A}$ uma sub-σ-álgebra.*

(1) Quando $\mathcal{B} = \{\emptyset, X\}$ então $E(P_i|\{\emptyset, X\})(x) = \mu(P_i)$, para todo $i \in I$, e também $I(\mathcal{P}|\{\emptyset, X\})(x) = I(\mathcal{P})(x)$.

(2) Se $T : X \to X$ preserva a medida μ então $I(\mathcal{P}|\mathcal{B})(T(x)) = I(T^{-1}(\mathcal{P})|T^{-1}(\mathcal{B}))(x)$, onde $T^{-1}(\mathcal{P}) = \{T^{-1}(P_i); i \in I\}$.

(3) Se $\mathcal{P} \subseteq \mathcal{B}$ então $I(\mathcal{P}|\mathcal{B}) = 0$ quase sempre.

Podemos associar a cada partição $\mathcal{P}$ a σ-álgebra $\hat{\mathcal{P}}$ gerada por $\mathcal{P}$.

Definição 6.4. *Dadas duas partições $\mathcal{P} = \{P_i\}_{i \in I}$, $\mathcal{Q} = \{Q_j\}_{j \in J}$ de um espaço de probabilidade $(X, \mathcal{A}, \mu)$. Definimos seu refinamento por*

$$\mathcal{P} \vee \mathcal{Q} = \{P_i \cap Q_j; i \in I, j \in J\}.$$

Dadas duas σ-álgebras $\mathcal{A}$, $\mathcal{B}$ denotamos por $\mathcal{A} \vee \mathcal{B}$ a σ-álgebra gerada por $\{A \cap B; A \in \mathcal{A}$ e $B \in \mathcal{B}\}$.

Lema 6.3 (Identidade Básica da Informação). *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade, $\mathcal{P}$, $\mathcal{Q}$, $\mathcal{R}$ partições de X e $\hat{\mathcal{R}}$ a σ-álgebra gerada por $\mathcal{R}$. Então temos que*

$$I(\mathcal{P} \vee \mathcal{Q}|\hat{\mathcal{R}})(x) = I(\mathcal{P}|\hat{\mathcal{Q}} \vee \hat{\mathcal{R}})(x) + I(\mathcal{Q}|\hat{\mathcal{R}})(x)$$

μ q.s.

Demonstração. Observe que para qualquer função $g \in L^1(X, \mathcal{A}, \mu)$ temos que

$$E(g|\hat{\mathcal{R}})(x) = \sum_{R \in \mathcal{R}} \chi_R(x) \frac{\int_R g(x) d\mu}{\mu(R)}.$$

Em particular, para $Q \in \mathcal{Q}$ podemos fazer $g(x) = \chi_Q(x)$ e então

$$\mu(Q|\hat{\mathcal{R}})(x) = \sum_{R \in \mathcal{R}} \chi_R(x) \frac{\mu(Q \cap R)}{\mu(R)}$$

e, portanto,

$$I(\mathcal{Q}|\hat{\mathcal{R}}) = -\sum_{R\in\mathcal{R},Q\in\mathcal{Q}} \chi_{R\cap Q}(x)\log\left(\frac{\mu(Q\cap R)}{\mu(R)}\right). \tag{6.1}$$

A partição $\mathcal{Q}\vee\mathcal{R}$ nos dá que

$$I(\mathcal{P}|\hat{\mathcal{R}}\vee\hat{\mathcal{Q}})(x) = -\sum_{R\in\mathcal{R},Q\in\mathcal{Q},P\in\mathcal{P}} \chi_{P\cap Q\cap R}(x)\log\left(\frac{\mu(P\cap Q\cap R)}{\mu(Q\cap R)}\right). \tag{6.2}$$

Somando 6.1 e 6.2 obtemos

$$\begin{aligned}
&I(\mathcal{Q}|\hat{\mathcal{R}})(x) + I(\mathcal{P}|\hat{\mathcal{Q}}\vee\hat{\mathcal{R}})(x)\\
&= -\sum_{R\in\mathcal{R},Q\in\mathcal{Q},P\in\mathcal{P}} \chi_{P\cap Q\cap R}(x)\left(\log\left(\frac{\mu(Q\cap R)}{\mu(R)}\right)+\log\left(\frac{\mu(P\cap Q\cap R)}{\mu(Q\cap R)}\right)\right)\\
&= -\sum_{R\in\mathcal{R},Q\in\mathcal{Q},P\in\mathcal{P}} \chi_{P\cap Q\cap R}(x)\log\left(\frac{\mu(P\cap Q\cap R)}{\mu(R)}\right)\\
&= I(\mathcal{P}\vee\mathcal{Q}|\hat{\mathcal{R}})(x).
\end{aligned}$$

□

Definição 6.5. *Sejam $\mathcal{P}$ e $\mathcal{Q}$ partições de $(X,\mathcal{A},\mu)$ um espaço de probabilidade. Escrevemos $\mathcal{P}<\mathcal{Q}$ se todo elemento de $\mathcal{P}$ for uma união de elementos de $\mathcal{Q}$. Neste caso $\mathcal{P}\vee\mathcal{Q}=\mathcal{Q}$.*

Corolário 6.3.1. *Nas hipóteses do lema anterior, se $\mathcal{P}<\mathcal{Q}$ então $I(\mathcal{P}|\hat{\mathcal{R}})(x) \leq I(\mathcal{Q}|\hat{\mathcal{R}})(x)$.*

6.2 Entropia de uma Partição

Nesta seção $(X,\mathcal{A},\mu)$ será um espaço de probabilidade, e toda partição mencionada é referente a este espaço X. A aplicação $T: X\to X$ será uma transformação que preserva a medida μ.

Definição 6.6. *Definimos a entropia da partição $\mathcal{P}$ por*

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

Dada uma sub-σ-álgebra $\mathcal{B}\subseteq\mathcal{A}$ definimos a entropia condicional por

$$H(\mathcal{P}|\mathcal{B}) = \int I(\mathcal{P}|\mathcal{B})d\mu.$$

Lema 6.4. *Sejam $\mathcal{P}$ uma partição e $\mathcal{B}$ uma sub-σ-álgebra.*

(1) Quando $\mathcal{B} = \{\emptyset, X\}$ então $H(\mathcal{P}|\{\emptyset, X\}) = H(\mathcal{P})$.

(2) Se $T : X \to X$ preserva a medida μ então $H(\mathcal{P}|\mathcal{B}) = H(T^{-1}(\mathcal{P})|T^{-1}(\mathcal{B}))$.

(3) Se $\mathcal{P} \subseteq \mathcal{B}$ então $H(\mathcal{P}|\mathcal{B}) = 0$.

(4) Dados $\mathcal{P}$ e $\mathcal{R}$ partições temos que $H(\mathcal{P}|\mathcal{R}) \leq H(\mathcal{P})$.

Demonstração. Os itens (1), (2) e (3) seguem da integração dos correspondentes resultados para a função informação no lema 6.2.

Vamos provar o item (4)

$$
\begin{aligned}
H(\mathcal{P}|\mathcal{R}) &= - \sum_{R\in\mathcal{R}, P\in\mathcal{P}} \mu(P\cap R) \log\left(\frac{\mu(P\cap R)}{\mu(R)}\right) \\
&= - \sum_{R\in\mathcal{R}} \mu(R) \left[\sum_{P\in\mathcal{P}} \frac{\mu(P\cap R)}{\mu(R)} \log\left(\frac{\mu(P\cap R)}{\mu(R)}\right)\right] \\
&\leq - \sum_{P\in\mathcal{P}} \left[\sum_{R\in\mathcal{R}} \mu(P\cap R)\right] \log\left[\sum_{R\in\mathcal{R}} \mu(P\cap R)\right] \\
&\leq - \sum_{P\in\mathcal{P}} \mu(P) \log \mu(P) = H(\mathcal{P})
\end{aligned}
$$

onde para $P \in \mathcal{P}$ fixado usamos que

$$
- \sum_{R\in\mathcal{R}} \frac{\mu(P\cap R)}{\mu(R)} \log\left(\frac{\mu(P\cap R)}{\mu(R)}\right) \leq - \left[\sum_{R\in\mathcal{R}} \mu(P\cap R)\right] \log\left[\sum_{R\in\mathcal{R}} \mu(P\cap R)\right]
$$

para provar este fato usamos a concavidade da função $t \to -t\log t$, veja [W2]. $\square$

Lema 6.5 (Identidade Básica para Entropia). *Dadas as partições $\mathcal{P}$, $\mathcal{Q}$ e uma terceira $\mathcal{R}$ que gera a σ-álgebra $\hat{\mathcal{R}}$, temos que*

$$
H(\mathcal{P} \vee \mathcal{Q}|\hat{\mathcal{R}}) = H(\mathcal{P}|\hat{\mathcal{Q}} \vee \hat{\mathcal{R}}) + H(\mathcal{Q}|\hat{\mathcal{R}}).
$$

Demonstração. Segue imediatamente do lema 6.3. $\square$

Uma consequência da identidade básica da entropia é o seguinte corolário.

Corolário 6.5.1 (Monotonicidade da Entropia para Partições). *Dadas duas partições $\mathcal{P}$ e $\mathcal{Q}$ com $\mathcal{P} < \mathcal{Q}$ temos que*

$$H(\mathcal{P}|\hat{\mathcal{R}}) \leq H(\mathcal{Q}|\hat{\mathcal{R}})$$

e, em particular, $H(\mathcal{P}) \leq H(\mathcal{Q})$.

Definição 6.7. *Assuma que $T : X \to X$ preserve a medida μ. Dada uma partição $\mathcal{P} = \{P_i\}_{i \in I}$ escrevemos, para $n \geq 1$*

$$\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}) = \{A_{r_0} \cap T^{-1}(A_{r_1}) \cap \cdots \cap T^{-n+1}(A_{r_{n-1}}); A_{r_i} \in \mathcal{P}, 0 \leq i \leq n-1\}.$$

Obs.: Frequentemente iremos denotar $\hat{\mathcal{R}}$ por $\mathcal{R}$, donde se escrevermos $H(\mathcal{P}|\mathcal{R})$ entendemos que isto significará $H(\mathcal{P}|\hat{\mathcal{R}})$.

Seja o seguinte lema sobre sequências numéricas.

Lema 6.6. *Seja $\{a_n\}_{n \geq 1}$ uma sequência de números reais tal que $a_{m+n} \leq a_m + a_n$, $\forall m, n$, então existe o limite $\lim_{n \to \infty} a_n/n$ e vale $\inf\{a_n/n; n \geq 1\}$. (O limite pode valer $-\infty$ mas se a_n for limitada por baixo então o limite é positivo.)*

Demonstração. [W2] □

Obs.: Uma sequência de números reais que satisfaça as hipóteses do lema 6.6 é dita de sequência subaditiva.

Para $n \geq 1$ escrevemos $H_n(\mathcal{P}) = H(\vee_{i=0}^{n-1} T^{-i}(\mathcal{P}))$. Assim,

$$\begin{aligned}
H_{m+n}(\mathcal{P}) &= H\left(\mathcal{P} \vee T^{-1}(\mathcal{P}) \vee \cdots \vee T^{-m}(\mathcal{P}) \vee T^{-(m+1)}(\mathcal{P}) \vee \cdots \vee T^{-(n+m-1)}(\mathcal{P})\right) \\
&\leq H\left(\mathcal{P} \vee T^{-1}(\mathcal{P}) \vee \cdots \vee T^{-(m-1)}(\mathcal{P})\right) \\
&+ H\left(T^{-m}(\mathcal{P}) \vee T^{-(m+1)}(\mathcal{P}) \vee \cdots \vee T^{-(n+m-1)}(\mathcal{P})\right) \\
&= H\left(\mathcal{P} \vee T^{-1}(\mathcal{P}) \vee \cdots \vee T^{-(m-1)}(\mathcal{P})\right) \\
&+ H\left(T^{-m}\left(\mathcal{P} \vee T^{-1}(\mathcal{P}) \vee \cdots \vee T^{-(n-1)}(\mathcal{P})\right)\right) \\
&= H_m(\mathcal{P}) + H_n(\mathcal{P})
\end{aligned}$$

provando assim, que a sequência $H_n(\mathcal{P})$, $n \geq 1$ é subaditiva donde temos a existncia do limite na próxima definição.

Definição 6.8. *Definimos a entropia da partição $\mathcal{P}$ relativa à transformação $T : X \to X$ pelo limite*

$$h(T, \mathcal{P}) = \lim_{n \to \infty} \frac{H_n(\mathcal{P})}{n}.$$

Proposição 6.1 (Definição Alternativa de $h(T,\mathcal{P})$).

$$h(T,\mathcal{P}) = \lim_{n\to\infty} H\left(\mathcal{P}\,\middle|\,\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right).$$

Ás vezes, é conveniente escrevermos este limite como $H(\mathcal{P}|\bigvee_{i=1}^{\infty} T^{-i}(\mathcal{P}))$.

Demonstração. Pelo lema 6.5 temos que, para todo $n \geq 1$

$$\begin{aligned}
H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) &= H\left(\mathcal{P}\,\middle|\,\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right) + H\left(\bigvee_{i=0}^{n-2} T^{-i}(\mathcal{P})\right) \\
&= H\left(\mathcal{P}\,\middle|\,\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right) + H\left(\mathcal{P}\,\middle|\,\bigvee_{i=1}^{n-2} T^{-i}(\mathcal{P})\right) + H\left(\bigvee_{i=0}^{n-3} T^{-i}(\mathcal{P})\right) \\
&\cdots \\
&= \sum_{r=2}^{n} H\left(\mathcal{P}\,\middle|\,\bigvee_{i=1}^{r-1} T^{-i}(\mathcal{P})\right) + H(\mathcal{P}).
\end{aligned}$$

Vemos então que

$$\lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) = \lim_{n\to\infty} H\left(\mathcal{P}\,\middle|\,\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P})\right).$$

Aqui usamos o fato que se uma sequência de números reais é tal que $a_n \to a$ então $\frac{a_1+a_2+\cdots+a_n}{n} \to a$. $\square$

A entropia de uma transformação relativa a duas partiços é descrita pela seguinte desigualdade.

Lema 6.7. *Para partições com entropia finita $\mathcal{P}$ e $\mathcal{Q}$ temos*

$$h(T,\mathcal{P}) \leq h(T,\mathcal{Q}) + H(\mathcal{P}|\mathcal{Q}).$$

Demonstração. Como $(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})) \vee (\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})) > \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})$ temos, usando o lema 6.5, que

$$\begin{aligned}
H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) &\leq H\left(\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) \vee \left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right)\right) \\
&= H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\,\middle|\,\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right).
\end{aligned}$$

Agora, seja

$$\begin{aligned}
&H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}) \middle| \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right)\\
&= H\left(\mathcal{P} \middle| \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + H\left(\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P}) \middle| \mathcal{P} \vee \left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right)\right)\\
&\leq H(\mathcal{P}|\mathcal{Q}) + H\left(\bigvee_{i=1}^{n-1} T^{-i}(\mathcal{P}) \middle| \bigvee_{i=1}^{n-1} T^{-i}(\mathcal{Q})\right)\\
&\leq H(\mathcal{P}|\mathcal{Q}) + H\left(\bigvee_{i=0}^{n-2} T^{-i}(\mathcal{P}) \middle| \bigvee_{i=0}^{n-2} T^{-i}(\mathcal{Q})\right).
\end{aligned}$$

Indutivamente temos que

$$H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P}) \middle| \bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) \leq nH(\mathcal{P}|\mathcal{Q}).$$

Finalmente, podemos perceber que

$$\frac{1}{n}H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P})\right) \leq \frac{1}{n}H\left(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{Q})\right) + H(\mathcal{P}|\mathcal{Q}).$$

E o lema está provado ao tomarmos $n \to \infty$. □

Corolário 6.0.1. *Para partições com entropia finita $\mathcal{P}$ e $\mathcal{Q}$ temos que*

$$|h(T,\mathcal{Q}) - h(T,\mathcal{P})| \leq H(\mathcal{Q}|\mathcal{P}) + H(\mathcal{P}|\mathcal{Q}).$$

Demonstração. Basta mudar $\mathcal{P}$ por $\mathcal{Q}$ no lema 6.7 onde obtemos que $h(T,\mathcal{Q}) \leq h(T,\mathcal{P}) + H(\mathcal{Q}|\mathcal{P})$. □

6.3 Entropia de uma Transformação

Nesta seção consideraremos $T : X \to X$ uma transformação que preserva medida em um espaço de probabilidade $(X, \mathcal{A}, \mu)$.

Definição 6.9. *Definimos a entropia métrica para a transformação T por*

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

onde o supremo é tomado sobre todas as partições finitas de X.

Pode ser muito difícil calcular a entropia métrica para uma transformação T. Agora, vamos enunciar dois resultados que auxiliam nesse cálculo.

Lema 6.8 (Abramov). *Seja*

$$\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \subseteq \mathcal{A}_k \subseteq \mathcal{A}$$

uma sequência crescente de partições com entropia finita e tal que $\bigcup_{k\geq 1} \mathcal{A}_k$ gera a σ-álgebra $\mathcal{A}$. Então,

$$h_\mu(T) = \lim_{k\to\infty} h(T, \mathcal{A}_k).$$

Demonstração. [PY] □

A seguinte definição nos dá um modo de generalizar partições crescentes.

Definição 6.10. *Dizemos que uma partição $\mathcal{P}$ com entropia finita é um gerador para o espaço de probabilidade $(X, \mathcal{A}, \mu)$ se $\bigvee_{i=0}^{\infty} T^{-i}(\mathcal{P}) = \mathcal{A}(mod\ 0)$.*

Caso T seja inversível, dizemos que uma partição $\mathcal{P}$ com entropia finita é um gerador para o espaço de probabilidade $(X, \mathcal{A}, \mu)$ se $\bigvee_{i=-\infty}^{\infty} T^{-i}(\mathcal{P}) = \mathcal{A}(mod\ 0)$.

Lema 6.9 (Sinai). *Se $\mathcal{P}$ é uma partição geradora então $h_\mu(T) = h(T, \mathcal{P})$.*

Demonstração. [PY] □

No que segue, iremos provar a existência de uma partição geradora para uma transformação expansora em um espaço métrico compacto.

Teorema 6.1. *Seja $T : (X\mathcal{A}, \mu) \to (X\mathcal{A}, \mu)$ uma transformação expansora definida em um espaço métrico compacto X. Então existe uma partição geradora para $\mathcal{A}$.*

Demonstração. Seja $T : X \to X$ uma transformação expansora em um espaço métrico compacto. Sejam $0 < \alpha < 1$, $r > 0$ e $c > 0$ constantes dadas pela definição de T expansora.

Considere a cobertura de X dada por

$$\bigcup_{x\in X} B(x, r^*)$$

onde $B(x, r^*)$ é a bola aberta de centro em x e raio r^* e $r^* < \min\{r/2, c/5\}$. Como X é compacto existem $x_1, x_2, \ldots, x_m \in X$ tais que

$$\bigcup_{i=1}^{m} B(x_i, r^*) \supseteq X.$$

Seja $\mathcal{P}$ a partição de X dada por

$$\mathcal{P} = \{P_1, \dots, P_m\}$$

onde

$$\begin{aligned}
P_1 &= B(x_1, r^*) \\
P_2 &= B(x_2, r^*) - B(x_1, r^*) \\
&\vdots \\
P_m &= B(x_m, r^*) - \bigcup_{i=1}^{m-1} B(x_i, r^*).
\end{aligned}$$

Definimos $\mathcal{P}^n = \mathcal{P} \vee T^{-1}(\mathcal{P}) \vee \cdots \vee T^{-n+1}(\mathcal{P})$. Queremos mostrar que $\mathcal{P}$ é geradora, isto é, que

$$\mathcal{A} \subseteq \bigvee_{j=0}^{\infty} T^{-j}(\mathcal{P})(\mu \pmod 0).$$

Agora, seja $A \in \mathcal{A}$ um aberto qualquer. Para $n \geq 1$ definimos os conjuntos

$$A_n = \{P \in \mathcal{P}^n; P \subseteq A\}$$

$$B_n = \bigcup_{P \in A_n} P$$

note que $B_1 \subseteq B_2 \subseteq \cdots \subseteq B_n \subseteq \cdots \subseteq A$. Agora, vamos provar que $A = \bigcup_{n \geq 1} B_n$.

Sejam $x \in A$ e $\varepsilon > 0$ tais que $d(x, \partial(A)) = \varepsilon$, onde $\partial(A)$ é a fronteira de A, então pelo lema 6.10 (enunciado e provado logo após o término desta demonstração) existe um $n \in \mathbb{N}$ tal que diam $(\mathcal{P}^n) \leq \varepsilon/2$. Assim, deve existir $P \in \mathcal{P}^n$ tal que $x \in P$, uma vez que diam $(P) \leq \varepsilon/2$ donde $P \subseteq A$. Como isto vale para todo $x \in A$ e todo $\varepsilon > 0$, temos que todo ponto de A está contido em um B_n e podemos concluir que

$$A = \bigcup_{n \geq 1} B_n.$$

Assim, para uma medida qualquer μ temos que

$$\mu(A \Delta B_n) \to 0$$

quando $n \to \infty$, ou seja, $A \in \bigvee_{n=0}^{\infty} \mathcal{P}^n$. Provando, assim, que T possui uma partição geradora. $\square$

Lema 6.10. *Sejam $T : X \to X$ uma transformação expansora em um espaço métrico compacto, e $\mathcal{P}$ a partição de X definida na prova do teorema 6.1. Considere, para $n \geq 1$, $\mathcal{P}^n = \mathcal{P} \vee T^{-1}(\mathcal{P}) \vee \cdots \vee T^{-n+1}(\mathcal{P})$. Se diam$(\mathcal{P}^n) = \sup\{$diam$(P); P \in \mathcal{P}^n\}$ então diam$(\mathcal{P}^n)$ converge a zero quando $n \to \infty$.*

Demonstração. Seja $T : X \to X$ uma transformação expansora em um espaço métrico compacto. Sejam $0 < \alpha < 1$, $r > 0$ e $c > 0$ constantes dadas pela definição de T expansora e $r^* < \min\{r/2, c/5\}$.

Seja Q um átomo qualquer de $\mathcal{P}^n$, então

$$Q = P_{i_1} \cap T^{-1}(P_{i_2}) \cap \cdots \cap T^{-n+1}(P_{i_n}).$$

Vamos provar, por indução em n, que

$$\text{diam } (Q) \leq \alpha^n r^*.$$

Se $n = 0$ então da própria definição de $\mathcal{P}$ temos que se $Q \in \mathcal{P}$ então diam $(Q) < r^*$.

Suponhamos a afirmação verdadeira para um $n > 0$. Agora vamos verificar para $n + 1$. Observe que

$$Q = P_{i_1} \cap T^{-1}(P_{i_2}) \cap \cdots \cap T^{-n+1}(P_{i_n}) \cap T^{-n}(P_{i_{n+1}}).$$

Então, $Q = P_{i_1} \cap T^{-1}(Q')$ onde $Q' = P_{i_2} \cap \cdots \cap T^{-n+1}(P_{i_{n+1}}) \in \mathcal{P}^n$. Segue por indução que diam $(Q') < \alpha^n r^*$. Agora, $T^{-1}(Q')$ é formada por k pré-imagens disjuntas de Q'. Afirmamos que duas destas pré-imagens quaisquer não podem interceptar o mesmo átomo de $\mathcal{P}$ pois, caso contrário, dado que diam $(P_{i_1}) < c/5$ segue que quaisquer dois pontos destas duas pré-imagens que são levados, por T, em um mesmo ponto estariam a uma distância menor que c, contradizendo o fato de T ser expansora.

Logo, existe $\varphi : Q' \to X$ um ramo contrativo de T^{-1} tal que

$$\varphi(Q') \cap P_{i_1} = T^{-1}(Q') \cap P_{i_1}$$

e portanto

$$\begin{aligned} \text{diam } (Q) &= \text{diam } (T^{-1}(Q') \cap P_{i_1}) = \text{diam } (\varphi(Q') \cap P_{i_1}) \\ &< \text{diam } (Q')\alpha < r^* \alpha^n \alpha = r^* \alpha^{n+1}. \end{aligned}$$

Provando assim o lema. $\square$

6.4 O Teorema de Martingale

Seja $T : (X, \mathcal{A}, \mu) \to (X, \mathcal{A}, \mu)$ uma transformação em um espço de probabilidade. Sabemos que se $f \in L^1(X, \mathcal{A}, \mu)$ e $\mathcal{B} \subseteq \mathcal{A}$ é uma sub-σ-álgebra então podemos associar a esperança condicional $E(f|\mathcal{B}) \in L^1(X, \mathcal{A}, \mu)$. O teorema de Martingale descreve como $E(f|\mathcal{B})$ depende da σ-álgebra $\mathcal{B}$.

Lema 6.11. *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade,*

$$\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \subseteq \mathcal{A}_N \subseteq \mathcal{A}$$

uma sequência de σ-álgebras e $\lambda > 0$. Se

$$E = \left\{ x \in X; \max_{1 \leq n \leq N} E(f|\mathcal{A}_n)(x) > \lambda) \right\}$$

então temos uma cota superior para sua medida, i. e.,

$$\mu(E) \leq \frac{1}{\lambda} \int |f| d\mu, \qquad \forall f \in L^1(X, \mathcal{A}, \mu).$$

Demonstração. Sem perda de generalidade podemos assumir que $f \geq 0$, pois caso contrário mudamos f por $\max_{x \in X}\{f(x), 0\}$.

Consideremos a partição $E = E_1 \cup E_2 \cup \cdots \cup E_N$ onde para $1 \leq n \leq N$

$$E_n = \{x \in X; E(f|\mathcal{A}_n)(x) > \lambda, E(f|\mathcal{A}_i)(x) \leq \lambda, 1 \leq i \leq n-1\}$$

e observe que $E_n \in \mathcal{A}_n$. Assim, $E_i \cap E_j = \emptyset$ para todo $i \neq j$. Então, podemos escrever

$$\begin{aligned} \int_E f d\mu &= \sum_{n=1}^{N} \int_{E_n} f d\mu = \sum_{n=1}^{N} \int_{E_n} E(f|\mathcal{A}_n) d\mu \geq \sum_{n=1}^{N} \lambda \mu(E_n) \\ &= \lambda \mu(E). \end{aligned}$$

Assim, $\mu(E) \leq \frac{1}{\lambda} \int f d\mu = \frac{1}{\lambda} \int |f| d\mu$. □

Obs.: O lema 6.11 é muito parecido com a desigualdade de Chebyshev para $f \in L^1(X, \mathcal{A}, \mu)$ e $\lambda > 0$ que afirma que

$$\mu\{x \in X; f(x) > \lambda\} \leq \frac{\int |f| d\mu}{\lambda}.$$

Teorema 6.2 (Teorema Crescente de Martingale). *Sejam $(X, \mathcal{A}, \mu)$ um espaço de probabilidade e $f \in L^1(X, \mathcal{A}, \mu)$. Assuma que*

$$\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \subseteq \mathcal{A}_N \subseteq \mathcal{A}$$

seja uma sequência crescente de σ-álgebras e que a união $\bigcup_{n=1}^{\infty} \mathcal{A}_n$ gere a σ-álgebra $\mathcal{A}$, que denotamos por $\mathcal{A}_n \to \mathcal{A}$. Então, $E(f|\mathcal{A}_n) \to f$ em $L^1(X, \mathcal{A}, \mu)$ e $E(f|\mathcal{A}_n)(x) \to f(x)$, μ q.s.

Demonstração. O teorema é claramente verdadeiro no subespaço $\bigcup_{n=1}^{\infty} L^1(X, \mathcal{A}_n, \mu)$ pois se $g \in L^1(X, \mathcal{A}_n, \mu)$ então $E(g|\mathcal{A}_n) = g$ para $n \geq k$. Além disso, esse subespaço é denso em $L^1(X, \mathcal{A}, \mu)$ na norma L^1.

Dada $f \in L^1(X, \mathcal{A}, \mu)$ podemos escolher $\varepsilon > 0$ e $g \in L^1(X, \mathcal{A}_n, \mu)$ com, digamos, $\int |f - g| d\mu < \varepsilon$. Assim, para qualquer $n \geq k$ temos que

$$\begin{aligned}
&\int |E(f|\mathcal{A}_n) - f| d\mu \\
&\quad \leq \int |E(f|\mathcal{A}_n) - E(g|\mathcal{A}_n)| d\mu + \int |E(g|\mathcal{A}_n) - g| d\mu + \int |g - f| d\mu \\
&\quad \leq 2 \int |g - f| d\mu
\end{aligned}$$

onde $\int |E(g|\mathcal{A}_n) - g| d\mu = 0$ pois $E(g|\mathcal{A}_n) = g$ e usamos que $E(\cdot|\mathcal{A}_n)$ é uma contração em $L^1(X, \mathcal{A}, \mu)$. Em particular, $\limsup_{n\to\infty} \int |E(f|\mathcal{A}_n) - f| d\mu \leq 2\varepsilon$. Como $\varepsilon > 0$ é arbitrário, temos a convergência em L^1.

Para mostrar que também temos a convergência μ quase sempre, seja

$$\begin{aligned}
&\mu \left\{ x \in X; \limsup_{n\to\infty} |E(f|\mathcal{A}_n)(x) - f(x)| > \varepsilon^{1/2} \right\} \\
&\quad \leq \mu \left\{ x \in X; \limsup_{n\to\infty} \left(|E(f - g|\mathcal{A}_n)(x) - (f - g)(x)| \right.\right. \\
&\qquad\qquad\qquad \left.\left. + |E(g|\mathcal{A}_n)(x) - g(x)| \right) > \varepsilon^{\frac{1}{2}} \right\} \\
&\quad \leq \mu \left\{ x \in X; \limsup_{n\to\infty} |E(f - g|\mathcal{A}_n)(x)| + |(f - g)(x)| > \varepsilon^{\frac{1}{2}} \right\} \\
&\quad \leq \mu \left\{ x \in X; \limsup_{n\to\infty} |E(f - g|\mathcal{A}_n)(x)| > \frac{1}{2}\varepsilon^{\frac{1}{2}} \right\} \\
&\quad + \mu \left\{ x \in X; |(f - g)(x)| > \frac{1}{2}\varepsilon^{\frac{1}{2}} \right\} \\
&\quad \leq 2 \left(\frac{1}{\frac{1}{2}\varepsilon^{1/2}} \right) \int |f - g| d\mu \leq 2 \left(\frac{1}{\frac{1}{2}\varepsilon^{1/2}} \right) \varepsilon \leq 4\varepsilon^{1/2}
\end{aligned}$$

onde usamos o lema 6.11 e a desigualdade de Chebyshev. Sendo $\varepsilon > 0$ arbitrário, isto mostra a convergência μ quase sempre. $\square$

6.5 Entropia e σ-álgebras

Vamos aplicar o teorema crescente de Martingale para uma sequência de funções informação. Como anteriormente, $(X, \mathcal{A}, \mu)$ será um espaço de probabilidade.

Lema 6.12. *Sejam $\mathcal{P}$ uma partição com $H(\mathcal{P}) < \infty$ e as sub-σ-álgebras $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \subseteq \mathcal{A}$. Então,*

$$\int \left(\sup_{n \geq 1} I(\mathcal{P}|\mathcal{A}_n) \right) d\mu \leq H(\mathcal{P}) + 1$$

e, em particular, $f(x) = \sup_{n\geq 1} I(\mathcal{P}|\mathcal{A}_n)(x) \in L^1(X, \mathcal{A}, \mu)$.

Demonstração. Podemos escrever, para $f \in L^1(X, \mathcal{A}, \mu)$

$$\int f(x) d\mu(x) = \int_0^\infty F(t) dt \tag{6.3}$$

onde $F(t) = \mu\{x \in X; f(x) > t\}$ supondo que o lado direito de 6.1 seja finito. Assim, sendo $P \in \mathcal{P}$ disjuntos

$$\begin{aligned} F(t) &= \mu \left\{ x \in X; \sup_{n\geq 1} I(\mathcal{P}|\mathcal{A}_n)(x) > t \right\} \\ &= \mu \left\{ x \in X; \sup_{n\geq 1} \left(-\sum_{P\in\mathcal{P}} \chi_P(x) \log \mu(P|\mathcal{A}_n)(x) \right) > t \right\} \\ &= \sum_{P\in\mathcal{P}} \mu \left(P \cap \left\{ x \in X; \sup_{n\geq 1} \left(-\log \mu(P|\mathcal{A}_n)(x) \right) > t \right\} \right). \end{aligned}$$

Simplificando temos

$$\begin{aligned} &\left\{ x \in X; \sup_{n\geq 1} \left(-\log \mu(P|\mathcal{A}_n)(x) \right) > t \right\} = \left\{ x \in X; \inf_{n\geq 1} \left(\log \mu(P|\mathcal{A}_n)(x) \right) < -t \right\} \\ &= \left\{ x \in X; \inf_{n\geq 1} \left(\mu(P|\mathcal{A}_n)(x) \right) < e^{-t} \right\} = \bigcup_{n\geq 1} A_n \end{aligned}$$

onde para todo $n \geq 1$,

$$A_n = \left\{ x \in X; \mu(P|\mathcal{A}_n)(x) < e^{-t}, \mu(P|\mathcal{A}_i)(x) \geq e^{-t}, 1 \leq i \leq n-1 \right\}$$

são conjuntos disjuntos. Se escrevermos

$$F(t) = \sum_{P \in \mathcal{P}} \mu \left(P \cap \bigcup_{n \geq 1} A_n \right) = \sum_{P \in \mathcal{P}} \sum_{n \geq 1} \mu(P \cap A_n)$$

então podemos usar a estimativa

$$\mu(P \cap A_n) = \int_{A_n} \chi_P d\mu = \int_{A_n} E(\chi_P|\mathcal{A}_n) d\mu \leq \int_{A_n} e^{-t} d\mu = e^{-t}\mu(A_n).$$

Assim, temos duas possíveis cotas superiores para a mesma soma, ou seja,

$$\begin{aligned} \sum_{n \geq 1} \mu(P \cap A_n) &\leq \sum_{n \geq 1} e^{-t}\mu(A_n) = e^{-t} \\ \sum_{n \geq 1} \mu(P \cap A_n) &\leq \mu(P). \end{aligned}$$

Logo, $F(t) \leq \sum_{P \in \mathcal{P}} \min\{e^{-t}, \mu(P)\}$. Finalmente, podemos usar esta cota para a estimativa seguinte

$$\begin{aligned} \int_0^\infty F(t) dt &\leq \int_0^\infty \left(\sum_{P \in \mathcal{P}} \min\{e^{-t}, \mu(P)\} \right) dt \\ &= -\sum_{P \in \mathcal{P}} \left(\mu(P) \log \mu(P) - \int_{-\log \mu(P)}^\infty e^{-t} dt \right) \\ &= -\sum_{P \in \mathcal{P}} (\mu(P) \log \mu(P) - \mu(P)) \\ &= H(\mathcal{P}) + 1. \end{aligned}$$

□

Teorema 6.3. *Se $\mathcal{P}$ é uma partição com $H(\mathcal{P}) < \infty$ e $\mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \to \mathcal{A}$ é uma sequência crescente de sub-σ-álgebras então $I(\mathcal{P}|\mathcal{A}_n)(x) \to I(\mathcal{P}|\mathcal{A})(x)$ quase sempre e em L^1. Assim, $H(\mathcal{P}|\mathcal{A}_n) \to H(\mathcal{P}|\mathcal{A})$ quando $n \to \infty$.*

Demonstração. Pelo teorema 6.2 temos que $\mu(P|\mathcal{A}_n) \to \mu(P|\mathcal{A})$ quase sempre, para qualquer $P \in \mathcal{P}$. Isto implica que $I(\mathcal{P}|\mathcal{A}_n)(x) \to I(\mathcal{P}|\mathcal{A})(x)$ quase sempre.

Pelo lema 6.12 temos que $I(\mathcal{P}|\mathcal{A}_n)$ são dominadas pela função integrável $\sup_{n \geq 1} I(\mathcal{P}|\mathcal{A}_n)(x)$. Assim, pelo teorema da convergência dominada de Lebesgue [DS] temos que $I(\mathcal{P}|\mathcal{A}_n)(x) \to I(\mathcal{P}|\mathcal{A})(x)$ μ em L^1, ou seja,

$$\int |I(\mathcal{P}|\mathcal{A}_n)(x) - I(\mathcal{P}|\mathcal{A})(x)| \, d\mu(x) \to 0$$

quando $n \to \infty$.

Por integração temos que $H(\mathcal{P}|\mathcal{A}_n) \to H(\mathcal{P}|\mathcal{A})$ quando $n \to \infty$. □

Corolário 6.3.1. *Seja $T : X \to X$ uma transformação que preserva a medida μ. Sejam $\mathcal{B} \subseteq \mathcal{A}$ uma sub-σ-álgebra e $\mathcal{P}$ uma partição de X com entropia finita e tal que*

$$\bigvee_{j=0}^{\infty} T^{-j}(\mathcal{P}) = \mathcal{B} \ mod\ 0.$$

Então,

$$h(T, \mathcal{P}) = -\int_X \sum_{i \geq 1} \log E(\chi_{P_i} | T^{-1}(\mathcal{B}))(x) \chi_{P_i}(x) d\mu(x)$$

onde $E(\chi_{P_i} | T^{-1}(\mathcal{B}))$ denota a esperança condicional de χ_{P_i} relativa a σ-álgebra $T^{-1}(\mathcal{B})$.

Demonstração. Da proposição 6.1 temos que

$$\begin{aligned} h(T, \mathcal{P}) &= \lim_{n\to\infty} H\left(\mathcal{P} \left| \bigvee_{j=1}^{n-1} T^{-j}(\mathcal{P}) \right.\right) = \lim_{n\to\infty} H\left(\mathcal{P} \left| T^{-1}\left(\bigvee_{j=0}^{n-2} T^{-j}(\mathcal{P})\right)\right.\right) \\ &= \lim_{n\to\infty} H(\mathcal{P} | T^{-1}(\mathcal{P}_n)) \end{aligned}$$

onde $\mathcal{P}_n = \bigvee_{j=0}^{n-2} T^{-j}(\mathcal{P})$. Como $\lim_{n\to\infty} \mathcal{P}_n = \mathcal{B}$, temos que

$$\lim_{n\to\infty} T^{-1}(\mathcal{P}_n) = T^{-1}(\mathcal{B}),$$

e pelo teorema 6.3 segue que

$$\begin{aligned} h(T, \mathcal{P}) &= \lim_{n\to\infty} H(\mathcal{P} | T^{-1}(\mathcal{P}_n)) = H(\mathcal{P} | T^{-1}(\mathcal{B})) \\ &= \int_X I(\mathcal{P} | T^{-1}(\mathcal{B}))(x) d\mu(x) \\ &= -\int_X \sum_{i \geq 1} \log E(\chi_{P_i} | T^{-1}(\mathcal{B}))(x) \chi_{P_i}(x) d\mu(x). \end{aligned}$$

□

6.6 Cálculo da Entropia das Transformações Expansoras

Lema 6.13. *Sejam $T : X \to X$ uma transformação expansora em um espaço métrico compacto e μ uma probabilidade T-invariante sobre os borelianos $\mathcal{A}$ de X. Então existe*

$$g_\mu : X \to \mathbb{R}^+$$

uma aplicação μ-integrável tal que para toda $f \in L^1(X, \mathcal{A}, \mu)$ e $x \in X$

$$E(f|T^{-1}(\mathcal{A}))(x) = \sum_{y \in T^{-1}(T(x))} g_\mu(y) f(y).$$

Demonstração. Dado $\varphi : B \to A$ um ramo contrativo de T^{-1} definimos ν a medida definida em $T(B)$ dada por

$$\nu(C) = \mu(\varphi(C))$$

então $\nu \ll \mu$ pois, se $\mu(C) = 0$ então $\mu(T^{-1}(C)) = 0$, uma vez que μ é invariante por T, e como

$$\varphi(C) \subseteq T^{-1}(C)$$

segue que $\mu(\varphi(c)) = 0$ e portanto, $\nu(C) = \mu(\varphi(C)) = 0$. Assim, segue do teorema de Radon-Nikodym que existe

$$J_\mu(\varphi) : T(B) \to \mathbb{R}^+$$

tal que

$$\nu(C) = \mu(\varphi(C)) = \int_C J_\mu(\varphi) d\mu.$$

Seja $X = \bigcup_{i=1}^N A_i$ tal que $T|_{A_i} : A_i \to T(A_i)$ é um homeomorfismo e A_i é um conexo contido em uma bola de raio $r > 0$, e tal que a fronteira de A_i tenha μ medida nula. Observe que $(T|_{A_i})^{-1} : T(A_i) \to A_i$ é um ramo contrativo de T^{-1}. Definimos $g_\mu : X \to \mathbb{R}^+$ via, se $x \in A_i$ então

$$g_\mu(x) = J_\mu(\varphi_i)(T(x))$$

que é uma aplicação μ-integrável. Onde $\varphi_i : T(A_i) \to A_i$ é um ramo contrativo de T^{-1}.

Agora, devemos mostrar que

$$\int_A \sum_{y \in T^{-1}(T(x))} g_\mu(y) f(y) d\mu = \int_A f d\mu$$

para todo $A = T^{-1}(B)$ com $B \in \mathcal{A}$. Basta verificar esta fórmula para $A \in T^{-1}(\mathcal{A})$ tais que $A = T^{-1}(B)$ com $B \subseteq B_r(y)$ de modo a termos ramos contrativos definidos em B, i. e., $T^{-1}(B) = \bigcup_{i=1}^k A_i$ com $\varphi_i : B \to A_i$, $i = 1, 2, \ldots, k$ ramos contrativos de T^{-1}.

Assim,

$$\begin{aligned}
\int_A \sum_{y\in T^{-1}(T(x))} g_\mu(y)f(y)d\mu(x) &= \int_A \sum_{i=1}^{k} g_\mu(\varphi_i(T(x)))f(\varphi_i(T(x)))d\mu(x) \\
&= \int_X \sum_{i=1}^{k} g_\mu(\varphi_i(T(x)))f(\varphi_i(T(x)))\chi_A(x)d\mu(x) \\
&= \int_X \sum_{i=1}^{k} g_\mu(\varphi_i(T(x)))f(\varphi_i(T(x)))\chi_{T(A)}(T(x))d\mu(x)
\end{aligned}$$

nesta última igualdade usamos que $\chi_A(x) = \chi_{T(A)}(T(x))$ e usando que μ é invariante por T temos

$$\begin{aligned}
&\int_X \sum_{i=1}^{k} g_\mu(\varphi_i(T(x)))f(\varphi_i(T(x)))\chi_{T(A)}(T(x))d\mu(x) \\
&\qquad = \int_X \sum_{i=1}^{k} g_\mu(\varphi_i(y))f(\varphi_i(y))\chi_{T(A)}(y)d\mu(y) \\
&\qquad = \sum_{i=1}^{k} \int_X g_\mu(\varphi_i(y))f(\varphi_i(y))\chi_{T(A)}(y)d\mu(y) \\
&\qquad = \sum_{i=1}^{k} \int_{T(A)} g_\mu(\varphi_i(y))f(\varphi_i(y))d\mu(y) \\
&\qquad = \sum_{i=1}^{k} \int_B g_\mu(\varphi_i(y))f(\varphi_i(y))d\mu(y).
\end{aligned}$$

Mas, também temos que

$$\int_{\varphi_i(B)=A_i} f d\mu = \int_B J_\mu(\varphi_i)(y)f(\varphi_i(y))d\mu(y).$$

Logo, para $x = \varphi_i(y)$ temos que

$$g_\mu(x) = J_\mu(\varphi)(T(x)).$$

Então,

$$\begin{aligned}
\sum_{i=1}^{k} \int_B g_\mu(\varphi_i(y))f(\varphi_i(y))d\mu(y) &= \sum_{i=1}^{k} \int_B J_\mu(\varphi_i)(y)f(\varphi_i(y))d\mu(y) \\
&= \sum_{i=1}^{k} \int_{\varphi_i(B)=A_i} f d\mu \\
&= \sum_{i=1}^{k} \int_{A_i} f d\mu = \int_A f d\mu
\end{aligned}$$

provando a igualdade desejada. $\square$

Corolário 6.13.1. *Sejam $T : X \to X$ uma transformação expansora em um espaço métrico compacto X, que preserva a medida de probabilidade μ e admite um Jacobiano $J_\mu(T) > 0$. Seja $g_\mu : X \to \mathbb{R}$ tal que*

$$E(f|T^{-1}(\mathcal{A}))(x) = \sum_{y \in T^{-1}(T(x))} g_\mu(y) f(y).$$

Então para μ q.t.p. $x \in X$

$$g_\mu(x) = \frac{1}{J_\mu(T)(x)}.$$

Teorema 6.4. *Sejam $T : X \to X$ uma transformação expansora, que preserva a medida de probabilidade μ, $\mathcal{P}$ uma partição geradora para T e seja $g_\mu : X \to \mathbb{R}$ tal que*

$$E(f|T^{-1}(\mathcal{A}))(x) = \sum_{y \in T^{-1}(T(x))} g_\mu(y) f(y).$$

Então,

$$h_\mu(T) = -\int_X \log g_\mu(x) d\mu(x).$$

Demonstração. Como $\mathcal{P}$ é geradora, segue que

$$\bigvee_{j=0}^{\infty} T^{-j}(\mathcal{P}) = \mathcal{A} \bmod 0.$$

Assim, pelo lema 6.9 e o corolário 6.3.1 segue que

$$\begin{aligned}
h_\mu(T) = h(T, \mathcal{P}) &= -\int_X \sum_{i \geq 1} \log E(\chi_{P_i}|T^{-1}(\mathcal{A}))(x) \chi_{P_i}(x) d\mu(x) \\
&= -\sum_{i \geq 1} \int_X \log E(\chi_{P_i}|T^{-1}(\mathcal{A}))(x) \chi_{P_i}(x) d\mu(x) \\
&= -\sum_{i \geq 1} \int_{P_i} \log E(\chi_{P_i}|T^{-1}(\mathcal{A}))(x) d\mu(x).
\end{aligned}$$

Como os átomos de $\mathcal{P}$ foram escolhidos de modo que $T|_{P_i}$ seja injetiva, ver prova de 6.1, segue que se $x \in P_i$ para algum $i \geq 1$

$$\log E(\chi_{P_i}|T^{-1}(\mathcal{A}))(x) = \log \left(\sum_{y \in T^{-1}(T(x))} g_\mu(y) \chi_{P_i}(y) \right) = \log g_\mu(x)$$

pois, se $T^{-1}(T(x)) = \{x_1, \ldots, x_n\}$ e $x_i \neq x$ então $x_i \notin P_i$ pois $x \in P_i$ e $T|_{P_i}$ é injetora. Logo, $\chi_{P_i}(x_i) = 0$. Portanto,

$$h_\mu(T) = -\sum_{i \geq 1} \int_{P_i} \log g_\mu(x) d\mu(x) = -\int_X \log g_\mu(x) d\mu(x)$$

$\square$

Referências Bibliográficas

[AB] A. Broise. *Transformations Dilatantes de L'Intervalle et Théorèmes Limites.* Études Spectrales D'Operateurs de Transfert et Applications, Astérisque, vol. 238, Soc. Mat. de France, 1996.

[B] V. Baladi. *Positive Transfer Operators and Decay of Correlations.* Advanced Series in Nonlinear Dynamics, Volume 16, World Scientific Publishing Company, 2000.

[B1] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms.* Lecture Notes in Mathematics, 470, 1975.

[Ca] A. Armando de Castro Jr. *Curso de Teoria da Medida.* Projeto Euclides, IMPA.

[Co] Donald L. Cohn. *Measure Theory.* Birkhäuser Boston, 1980.

[C1] M. Craizer. *Teoria Ergódica das Transformações Expansoras.* Informes de Matemática - 018/85 Série E, IMPA, 1985.

[DS] N. Dunford e J. T. Schwartz. *Linear Operators, Part I.* Interscience Publishers Inc., New York 1957.

[E] E. L. Lima. *Introdução à Topologia Diferencial.* Monografias de matemática n. 60, IMPA.

[G1] M. Gromov. *Expanding Maps and Groups of Polinomial Growth.* IHES, 53:53-73, 1979.

[H1] F. Hofbauer. *On intrinsic ergodicity of piecewise monotonic transformations with positive entropy.* Israel Journal of Mathematics, Vol. 34, n. 3, 213-237, 1979.

[H2] F. Hofbauer. *Examples for the nonuniqueness of the equilibrium states.* Transactions of the AMS, Vol. 228, 223-241, 1977.

[HK] F. Hofbauer e G. Keller. *Ergodic Properties of Invariant Measures for Piecewise Monotonic Transformations.* Mathematische Zeitschrift, 180, 119-140, 1982.

[HS] E. Hewitt e K. Stromberg. *Real and Abstract Analysis.* Springer-Verlag, Berlim, 1965.

[ITM] C. T. Ionescu Tulcea e G. Marinescu *Theorie Ergodique Pour des Classes D'Operations non Completement Continues.* Annals of Mathemtics, Vol. 52, N. 1,July 1950.

[K] M. Keane. *Strongly Mixing g-measures.* Invent. Math. 16, 309-324, MR 46#9295, 1972.

[Ke] J. Kelley. *General Topology.* Springer-Verlag, 1955.

[Kl] G. Keller. *Un théorème de la limite centrale pour une classe de transformations monotones par morceaux.* C. R. Acad. Sc. Paris, t. 291, 1980.

[LY1] A. Lasota, J. A. Yorke. *Exact Dynamical Systems and the Frobenius-Perron Operator.* Transactions of the AMS, Vol. 273, número 1, 1982.

[LY2] A. Lasota, J. A. Yorke. *On the Existence of Invariant Measures for Piecewise Monotonic Transformations.* Transactions of the AMS, Vol. 186, December, 1973.

[L1] F. Ledrappier. *Principe variationnel et systèmes dynamiques symboliques.* Z. Wahrscheinlichkeitstheorie verw. Gebiete 30, 185-202, 1974.

[M1] R. Mañé. *Introdução à Teoria Ergódica.* Projeto Euclides, IMPA, 1983.

[PY] M.Pollicott e M. Yuri. *Dynamical Systems and Ergodic Theory.* Versão online em http://www.maths.man.ac.uk/%7Emp/book.html

[R] D. Ruelle.*Statistical Mechanics of a One-dimensional Lattice Gas.* Comm. Math. Phys. 9, 267-278, MR 38# 3013, 1968.

[Ru] W. Rudin.*Real and complex analysis.* McGraw-Hill, 3rd edition, 1987.

[S1] M. Shub. *Endomorphisms of compact differentiable manifolds.* American Journal of Mathematics 91, 129-155, 1969.

[S2] M. Shub. *Global Stability of Dynamical Systems.* Springer-Verlag, New York, 1987.

[Si1] G. L. da Silva. *g-Medidas e Aplicações Expansoras Não-Ergódicas.* Dissertação de Mestrado em Matemática e Computação Científica, UFSC, 1997.

[V] M. Viana. *Stochastic Dynamics of Deterministic Systems.* Lecture Notes XXI Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro, 1997.

[W1] P. Walters. *Ruelle's Operator Theorem and g-Measures.* Transactions of the AMS, Vol. 214, 1975.

[W2] P. Walters. *An Introduction to Ergodic Theory.* Springer-Verlag, New york, 1982.

[Wo] S. Wong. *Some Metric Properties of Piecewise Monotonic Mappings of the Unit Interval.* Transactions of the AMS, Vol. 246, December 1978.