



Lino Ramada Ferreira Junior

Otimização Ergódica para
Difeomorfismos de Anosov

CAMPINAS
2015



Universidade Estadual de Campinas

INSTITUTO DE MATEMÁTICA, ESTATÍSTICA
E COMPUTAÇÃO CIENTÍFICA

Lino Ramada Ferreira Junior

Otimização Ergódica para Difeomorfismos de Anosov

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Matemática.

Orientador: Eduardo Garibaldi

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA PELO ALUNO LINO RAMADA FERREIRA JUNIOR E ORIENTADA PELO PROF. DR. EDUARDO GARIBALDI.

Assinatura do Orientador

A handwritten signature in black ink, appearing to be "Eduardo Garibaldi", is written over a horizontal line. The signature is fluid and cursive, with a long horizontal stroke extending to the right.

**Campinas
2015**

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Maria Fabiana Bezerra Muller - CRB 8/6162

F413o Ferreira Junior, Lino Ramada, 1991-
Otimização ergódica para difeomorfismos de Anosov / Lino Ramada Ferreira Junior. – Campinas, SP : [s.n.], 2015.

Orientador: Eduardo Garibaldi.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Otimização ergódica. 2. Anosov, Difeomorfismos de. I. Garibaldi, Eduardo, 1977-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Ergodic optimization for Anosov diffeomorphisms

Palavras-chave em inglês:

Ergodic optimization

Anosov diffeomorphisms

Área de concentração: Matemática

Titulação: Mestre em Matemática

Banca examinadora:

Eduardo Garibaldi [Orientador]

Pedro José Catuogno

Ricardo dos Santos Freire Júnior

Data de defesa: 27-02-2015

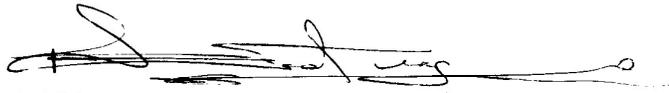
Programa de Pós-Graduação: Matemática

Dissertação de Mestrado defendida em 27 de fevereiro de 2015 e aprovada

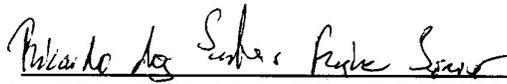
Pela Banca Examinadora composta pelos Profs. Drs.



Prof.(a). Dr(a). EDUARDO GARIBALDI



Prof.(a). Dr(a). PEDRO JOSÉ CATUOGNO



Prof.(a). Dr(a). RICARDO DOS SANTOS FREIRE JÚNIOR

Abstract. In this master’s thesis, we study ergodic optimization techniques in the context of an Anosov dynamical system. We present different approaches to the problem of maximization of the integral of Hölder potentials on a compact metric space in the presence of a hyperbolic dynamics. We discuss the thermodynamical formalism in an expansive model, obtaining maximizing probabilities at zero temperature. In the hyperbolic case, we determine a cohomological inequality in an amphidynamical system, from which follows a Lipschitz subaction for Lipschitz potentials associated with Anosov diffeomorphisms. Finally, we argue that periodic probabilities are maximizing for open sets of functions in the Lipschitz topology.

Keywords: Ergodic Optimization, Anosov Diffeomorphisms, Maximizing Probabilities, Transfer Operator, Sub-actions.

Resumo. Nesta dissertação, estudamos técnicas de otimização ergódica no contexto de uma dinâmica do tipo Anosov. Mostramos diferentes maneiras de abordar o problema de maximização da integral de potenciais holderianos definidos sobre um espaço métrico compacto na presença de uma dinâmica hiperbólica. Discutimos o formalismo termodinâmico sobre modelo expansivo, obtendo probabilidades maximizantes em temperatura nula. No caso hiperbólico, determinamos uma desigualdade coomológica em um sistema anfidinâmico, da qual resulta subação lipschitziana para potenciais lipschitzianos associados a difeomorfismos de Anosov. Finalmente, argumentamos que probabilidades periódicas são maximizantes para abertos de funções na topologia lipschitziana.

Palavras-chave: Otimização Ergódica, Difeomorfismos de Anosov, Probabilidades Maximizantes, Operador de Transferência, Subações.

Sumário

Introdução	1
1 Princípios de Otimização Ergódica	5
1.1 Preliminares	5
1.2 Maximizando a integral de uma função	15
1.2.1 Probabilidades maximizantes	16
1.2.2 Analisando o valor máximo da integral como um funcional	22
1.3 Cociclos e o conjunto de Aubry	31
1.3.1 O conjunto de Aubry	31
1.3.2 Cociclos	35
2 Formalismo Termodinâmico em Temperatura Zero e o Operador de Transferência	43
2.1 Formalismo termodinâmico e pressão	43
2.2 O operador de transferência	54
2.2.1 <i>Subshift</i> de tipo finito	62
3 Subações para Difeomorfismos de Anosov	67
3.1 Coberturas markovianas para difeomorfismos de Anosov	67
3.2 Sistemas anfidinâmicos	80
3.3 Probabilidades maximizantes associadas a difeomorfismos de Anosov	92
Referências Bibliográficas	97

Dedicatória

Para meu pai, Lino; minha mãe, Laura; e minha companheira, Grazi; por todo o amor, carinho e apoio em todos os momentos.

Agradecimento

Gostaria de agradecer, primeiramente, ao professor Eduardo Garibaldi, pelos ensinamentos, pela orientação, por todas as discussões de ideias e paciência, tendo motivado-me na pesquisa e proporcionado um amadurecimento como matemático.

Agradeço à Fapesp¹ pelo apoio financeiro através da concessão de bolsas de estudo que contribuem para a pesquisa, e sem o qual não seria possível elaborar este trabalho.

Gostaria de agradecer ainda à UNICAMP, que dispõe da infraestrutura necessária para a realização de um ótimo curso de graduação e mestrado, e também a todos os seus professores, que muito me ajudaram com todos os seus ensinamentos.

A minha família, em especial minha mãe Laura, por todo esforço, amor, dedicação e carinho, por ter sempre apoiado meus estudos e estar sempre por perto, principalmente nos maus momentos.

Aos amigos Cléber e João, com quem lado a lado pude aprender muito e são de grande inspiração como jovens matemáticos.

À Bianca, que nos quatro anos de graduação e mais dois de mestrado esteve sempre presente, compartilhando estudos, sorrisos e muitos abraços. Foi uma pessoa que me ensinou muito e a quem espero ter podido também ensinar.

Aos amigos Clarissa e Daniel, que me proporcionaram bons momentos, inclusive nas divergências de ideias.

Um agradecimento especial à minha namorada, esposa, amiga, companheira, confiante... Grazi, mulher com quem compartilho meu dia a dia, que tem paciência de me ver horas estudando, que me suporta, me incentiva, motiva, me alegra e me faz realizado, independente do resto. Também a toda sua família que me acolheu com muito carinho, inclusive para discutir nos churrascos sobre a não exatidão da matemática.

¹ Suporte financeiro: Bolsa de Mestrado da Fundação de Amparo à Pesquisa do Estado de São Paulo (Fapesp) 03/2013 – 02/2015, processo n° 2012/19977-1.

Epígrafe

*"If people do not believe that
mathematics is simple, it is only
because they do not realize how
complicated life is."*

John von Neumann

Introdução

Nesta dissertação de mestrado, abordaremos um tópico de sistemas dinâmicos que vem ganhando destaque e rendendo interessantes resultados nos últimos anos: trata-se da otimização ergódica para funções holderianas. Apresentaremos, no decorrer do texto, algumas das ferramentas mais importantes a serem estudadas nesta subárea da teoria ergódica, dando enfoque principalmente a resultados sobre difeomorfismos de Anosov.

Visando uma melhor compreensão do texto, principalmente para o leitor que não é completamente familiarizado com sistemas dinâmicos, iniciaremos o primeiro capítulo da dissertação recordando alguns fatos importantes sobre integração de medidas abstratas e fixando determinadas notações relacionadas. Apresentaremos alguns resultados bem conhecidos de teoria ergódica, tais como o teorema de recorrência de Poincaré e o teorema ergódico de Birkhoff, além de outros resultados relevantes que serão utilizados nas demais seções. Indicamos ao leitor a eventual consulta das referências [26, 29, 36] para uma ampla base de resultados em teoria ergódica.

Ainda no primeiro capítulo, introduziremos o problema central de otimização ergódica, o qual consiste em determinar quais as medidas de probabilidade μ , invariantes pela dinâmica topológica $T: X \rightarrow X$, maximizam a integral da função contínua $f: X \rightarrow \mathbb{R}$. Denotaremos por $\beta(f)$ o valor máximo que pode ser obtido ao se maximizar a integral $\int f d\mu$ com respeito a tais probabilidades μ . Mostraremos que o valor $\beta(f)$ pode ser calculado de diferentes maneiras, em particular, pela avaliação do ínfimo de máximos globais sobre a classe de coomologia da função f , isto é, sobre funções da forma

$$g = f + h \circ T - h,$$

onde $h: X \rightarrow \mathbb{R}$ é também uma função contínua. Como de costume na literatura, funções

do tipo $h \circ T - h$ serão denominadas cobordos – como veremos, estas funções desempenham importante papel no estudo de otimização ergódica. De fato, uma função $v \in C(X)$ que produz a desigualdade $f + v \circ T - v \leq \beta(f)$ sobre X será chamada de subação, e sua obtenção contribuirá para elucidação do problema de maximização da integral de f .

Ainda nesta parte introdutória, utilizaremos técnicas de análise funcional para demonstrar que o valor $\beta(f)$ pode ser visto como uma avaliação particular de funcional linear β sobre o espaço de Banach de funções contínuas $C(X)$. Este ponto de vista permitirá concluir que há um aberto denso de funções em $C(X)$ para o qual cada elemento admite uma única probabilidade maximizante. Mais ainda, utilizando um resultado geral sobre recorrência de cociclos, aliado à hipótese de existência de subação, poderemos obter informações cruciais a respeito do suporte das probabilidades maximizantes para f .

No segundo capítulo, abordaremos propriedades do formalismo termodinâmico sobre uma dinâmica expansiva. Recordaremos não só as definições mas também as principais propriedades da entropia topológica $h(T)$ e da entropia métrica $h_\mu(T)$. Veremos que a entropia topológica é um caso particular de um conceito mais amplo, denominado pressão topológica, a qual pode ser representada pela igualdade

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

onde o supremo é tomado sobre as probabilidades T -invariantes μ .

Este conceito, em particular, fornecerá nova abordagem em otimização ergódica ao permitir calcular o valor máximo da integral de uma função f em um processo que chamaremos de congelamento do sistema. Ao considerar um parâmetro t (interpretado como o inverso da temperatura) multiplicando a função f , discutiremos o limite quando $t \rightarrow \infty$ de

$$\frac{P_T(tf)}{t} = \sup \left\{ \frac{h_\mu(T)}{t} + \int f \, d\mu \right\}.$$

Utilizando o conceito de estados de equilíbrio para funções tf , que são as probabilidades que verificam o supremo da pressão $P_T(tf)$, o processo de congelamento do sistema possibilitará estudo de convergência de certas probabilidades para uma probabilidade que é maximizante para f . Neste contexto, quando a aplicação $T: X \rightarrow X$ for expansiva e f for lipschitziana, o operador de transferência, definido por

$$\mathcal{L}_f g(x) = \sum_{y \in T^{-1}x} e^{f(y)} g(y), \quad \forall g \in C(X),$$

fornecerá, via o conhecido teorema de Ruelle-Perron-Frobenius, uma probabilidade μ_t que é um estado de equilíbrio para tf .

É fato que a grande maioria dos resultados presentes na literatura de otimização ergódica até a presente data encontra-se ambientada no contexto de uma dinâmica expansiva. Dedicaremos o último capítulo desta dissertação a apresentar alguns resultados em otimização ergódica para dinâmicas hiperbólicas, as quais, é bem sabido, incluem os sistemas dinâmicos expansivos. O principal modelo de dinâmica que será estudado são os difeomorfismos de Anosov, isto é, difeomorfismos de uma variedade compacta M que admitem uma decomposição contínua do espaço tangente $T_x M = \mathbb{E}^s(x) \oplus \mathbb{E}^u(x)$ tal que a derivada DT_x produz contração equilimitada em $\mathbb{E}^s(x)$ e expansão também equilimitada em $\mathbb{E}^u(x)$.

Detalharemos como podem ser construídas subações para funções holderianas na presença de um difeomorfismo de Anosov. Essa construção será definida localmente sobre uma cobertura markoviana da variedade M e estendida para toda a variedade via colagem por uma partição da unidade. Esse processo requererá hipótese sobre a regularidade do difeomorfismo T para que seja possível obter um lema de sombreamento de pseudo-órbitas. A construção de tais subações baseia-se na referência [22].

Seguindo [6], obteremos uma inequação coomológica para potenciais 1-lipschitzianos no contexto de um sistema anfidinâmico $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$, no qual X_0 e X_1 são espaços métricos compactos e $T_0, T_1: X_1 \rightarrow X_0$ são aplicações contínuas. Isto significa que, dada função lipschitziana $u: X_1 \rightarrow \mathbb{R}$, queremos encontrar função também lipschitziana $v: X_0 \rightarrow \mathbb{R}$ tal que $u(x) \geq v \circ T_1(x) - v \circ T_0(x)$, $\forall x \in X_1$. Desta desigualdade resultará subação para funções lipschitzianas ao considerar $u = \beta(f) - f$ e o sistema anfidinâmico que generaliza o sistema dinâmico topológico (X, T) , ou seja, $X_0 = X_1 = X$, $T_1 = T$ e $T_0 = Id$. Utilizaremos enfim essas construções para mostrar que probabilidades periódicas são sempre maximizantes para conjuntos abertos de funções na topologia lipschitziana.

Princípios de Otimização Ergódica

1.1 Preliminares

Nesta seção inicial, relembremos alguns resultados elementares de teoria da medida, teoria ergódica, análise *etc.*, os quais servirão ao leitor como base de consulta, visando um melhor desenvolvimento do texto. O leitor que possuir maior maturidade pode, sem peso na consciência, avançar direto à próxima seção. Começamos fixando algumas das notações que serão utilizadas ao longo do texto.

Definição. *Damos o nome de sistema dinâmico topológico (SDT) ao par (X, T) , onde X é um espaço métrico compacto munido de uma distância d , e $T : X \rightarrow X$ é uma aplicação contínua e sobrejetiva.*

Como nosso propósito é avaliar valores otimais de integrais, recordamos inicialmente o conceito de uma medida sobre (X, T) .

Definição. *Uma σ -álgebra \mathcal{B} de X é um subconjunto das partes de X , satisfazendo:*

- $\emptyset, X \in \mathcal{B}$;
- $E \in \mathcal{B} \Rightarrow E^c \in \mathcal{B}$;
- $\{E_i\}_{i \in \mathbb{N}} \subseteq \mathcal{B} \Rightarrow \bigcup_{i \in \mathbb{N}} E_i \in \mathcal{B}$.

(Lembre que a σ -álgebra de Borel de X é a σ -álgebra gerada pelos abertos da topologia de X , ou seja, a menor σ -álgebra contendo os abertos de X .) Uma medida sobre a σ -álgebra \mathcal{B} é uma função $\mu : \mathcal{B} \rightarrow \mathbb{R}_+$ tal que

- $\mu(\emptyset) = 0$;
- para $E_1, E_2 \in \mathcal{B}$ tais que $E_1 \subseteq E_2$, temos $\mu(E_1) \leq \mu(E_2)$;
- se $E = \cup_{i \in \mathbb{N}} E_i$, com $E, E_i \in \mathcal{B}$, então, $\mu(E) \leq \sum_i \mu(E_i)$ e, caso $E = \sqcup_{i \in \mathbb{N}} E_i$, obtemos a igualdade.

A medida μ é dita finita se $\mu(X) < \infty$, e, em particular, é dita de probabilidade quando $\mu(X) = 1$. (Dizemos que μ é uma medida boreliana se está definida sobre a σ -álgebra de Borel de X .)

O trio (X, \mathcal{B}, μ) é usualmente chamado de espaço de medida. No decorrer da dissertação, pouco usaremos essa notação por dois motivos: será muito frequente trabalharmos com conjuntos de probabilidades em vez de considerá-las separadamente, e, a menos que seja dito o contrário, estaremos considerando a σ -álgebra de Borel de X .

Vejamos um resultado geral sobre medidas que nos será útil futuramente.

Lema 1.1.1. *Seja (X, \mathcal{B}, μ) um espaço de medida. Considere a sequência de conjuntos mensuráveis $\{B_n\}_{n \in \mathbb{N}} \subseteq \mathcal{B}$ tal que $B_n \subseteq B_{n+1}$ e $\cup_{n \in \mathbb{N}} B_n = B \in \mathcal{B}$, com $\mu(B) < \infty$. Então, $\mu\left(\cup_{n \in \mathbb{N}} B_n\right) = \lim_{n \rightarrow \infty} \mu(B_n)$ e, dado $\varepsilon > 0$, há $N = N(\varepsilon) \in \mathbb{N}$ tal que $\mu(B) - \varepsilon < \mu(B_n)$, para todo $n \geq N$.*

Demonstração. Denotamos $C_1 = B_1$ e $C_k = B_k \setminus C_{k-1}$, para $k \geq 2$. Claramente esses conjuntos são dois a dois disjuntos e, para todo $n \in \mathbb{N}$, ocorre

$$\bigcup_{i=1}^n C_i = \bigcup_{i=1}^n B_i = B_n.$$

Assim, $\cup_{n \in \mathbb{N}} C_n = \cup_{n \in \mathbb{N}} \cup_{k=1}^n B_k = \cup_{n \in \mathbb{N}} B_n = B$, donde

$$\mu\left(\bigcup_{n \in \mathbb{N}} C_n\right) = \sum_{k=1}^{\infty} \mu(C_k) = \lim_{n \rightarrow \infty} \sum_{k=1}^n \mu(C_k) = \lim_{n \rightarrow \infty} \mu\left(\bigcup_{k=1}^n C_k\right) = \lim_{n \rightarrow \infty} \mu(B_n) = \mu(B).$$

Por outro lado, como a série $\sum_{k=1}^{\infty} \mu(C_k)$ é convergente, dado $\varepsilon > 0$, há $N = N(\varepsilon)$ tal que $\sum_{k=n}^{\infty} \mu(C_k) < \varepsilon$, para todo $n \geq N$. Assim,

$$\mu(B_n) = \sum_{k=1}^n \mu(C_k) > \mu(B) - \varepsilon, \quad \forall n \geq N.$$

□

Para discutir integrabilidade, recordemos primeiro que uma função $f: X \rightarrow \mathbb{R}$ é mensurável se $f^{-1}(-\infty, \alpha) \in \mathcal{B}$, para todo $\alpha \in \mathbb{R}$. Além disso, a função característica de um conjunto $E \in \mathcal{B}$, denotada por χ_E , é a função mensurável que assume valores 1 sobre E e 0 sobre E^c . Uma função $f: X \rightarrow \mathbb{R}$ é dita simples se há uma partição de $X = E_1 \sqcup E_2 \sqcup \dots \sqcup E_n$, com $E_1, \dots, E_n \in \mathcal{B}$, e constantes $c_1, \dots, c_n \in \mathbb{R}$ tais que

$$f(x) = \sum_{i=1}^n c_i \chi_{E_i}(x), \quad \forall x \in X.$$

Observe que uma função f é mensurável se é pontualmente o limite de uma sequência de funções simples, isto é, $f(x) = \lim_n f_n(x)$, $\forall x \in X$. Para mais detalhes sobre essa propriedade, o leitor pode consultar, por exemplo, [4].

Dada uma função f mensurável, considere $\{f_n = \sum_{i=1}^{k_n} c_i \chi_{E_i}\}$ sequência de funções simples que aproximam f . Podemos definir a integral de f com respeito a μ como o limite

$$\int f \, d\mu := \lim_{n \rightarrow \infty} \int f_n \, d\mu = \sum_{i=1}^{k_n} c_i \mu(E_i),$$

Dizemos então que f é integrável se f for mensurável e $\int |f| \, d\mu < \infty$.

Na verdade, por trás desta proposta de definição de integral, estamos omitindo vários resultados de teoria da medida, como, por exemplo, o lema de Fatou, o teorema da convergência monótona e o teorema da convergência dominada de Lebesgue. Tais resultados também podem ser consultados em [4].

Na maior parte do texto, não lidamos com funções apenas mensuráveis. Dado um SDT (X, T) , estamos interessados em estudar propriedades do espaço das funções contínuas $C(X)$ – é claro que toda função contínua é mensurável, pois a imagem inversa de um aberto de \mathbb{R} é um aberto de X . Note ainda que, como X é compacto, toda função contínua é limitada, de modo que faz sentido considerar sobre $C(X)$ a norma do supremo

$$\|f\|_0 := \sup_{x \in X} |f(x)|, \quad \forall f \in C(X).$$

Com isto, temos a nossa disposição um espaço vetorial normado completo, ou melhor, o espaço de Banach $(C(X), \|\cdot\|_0)$, o qual é separável (para provas desses fatos, veja, por exemplo, [34]). Quando for conveniente, este será considerado como subespaço fechado de um espaço maior, o qual passamos a introduzir.

A abreviatura q.t.p. (significando “quase todo ponto”) é muito comum na literatura

e será empregada sempre que nos referirmos a alguma propriedade válida fora de um conjunto de medida nula. Dada μ uma medida boreliana sobre X , consideremos, sobre o conjunto das funções mensuráveis $f: X \rightarrow \mathbb{R}$ que são limitadas q.t.p., a relação de equivalência

$$f \sim g \iff \mu(\{x \in X : f(x) \neq g(x)\}) = 0.$$

Quando equivalentes, dizemos que $f = g$ μ -q.t.p., isto é, que f e g coincidem em quase todo ponto com respeito a μ .

O conjunto das classes de equivalência acima definidas é denotado $L^\infty(\mu)$. É comum identificarmos a classe de equivalência com seus representantes. Esse abuso de notação permite-nos escrever $f \in L^\infty(\mu)$. Definimos, então, a seguinte norma sobre $L^\infty(\mu)$

$$\|f\|_\infty = \inf \left\{ c \in \mathbb{R} : \mu(\{x \in X : |f(x)| > c\}) = 0 \right\}, \quad \forall f \in L^\infty(\mu).$$

Fica a cargo do leitor verificar que, para qualquer $f \in C(X)$, temos $f \in L^\infty(\mu)$ e $\|f\|_0 = \|f\|_\infty$. Fato fundamental que não demonstraremos aqui, mas que o leitor pode encontrar, por exemplo, em [4], é que $(L^\infty(\mu), \|\cdot\|_\infty)$ é um espaço de Banach.

Os conceitos acima também nos permitem explicar um abuso de notação que será cometido algumas vezes durante o texto. Se μ é medida finita, dada $f \in C(X)$, escrevemos $\mu(f)$ para representar $\int f d\mu$. Esse abuso justifica-se via o teorema da representação de Riesz, pois podemos identificar o espaço dual de $C(X)$ com um subespaço fechado das medidas borelianas com sinal sobre X . Portanto, *grosso modo*, a notação $\mu(f)$ significa a aplicação do funcional μ no elemento f . Em particular, é usual considerar o espaço de probabilidades borelianas munido da topologia fraca-*, o qual vem a ser compacto, uma vez que X é compacto. Para mais detalhes, o leitor pode consultar, por exemplo, [10, 32].

Uma vez lembrado o conceito de medida e recordados rapidamente aspectos da integração, a dinâmica passa a ter nossa atenção. Considere as seguintes definições.

Definição. Dado um SDT (X, T) , a órbita de um ponto $x \in X$ é definida por $Orb_T(x) := \{T^n x : n \in \mathbb{N}\}$, quando T não for inversível, e por $Orb_T(x) := \{T^n x : n \in \mathbb{Z}\}$, no caso em que T é inversível.

Dizemos que um conjunto $A \in \mathcal{B}$ é T -invariante se $T^{-1}(A) \subseteq A$.

Definição. Sejam (X, T) um SDT e μ uma medida em X . Dizemos que μ é T -invariante, ou que T preserva μ , ou ainda que T é invariante por μ , se para todo conjunto $A \in \mathcal{B}$ ocorrer $\mu(T^{-1}(A)) = \mu(A)$.

Repare que, para um SDT (X, T) e μ uma medida qualquer em X , podemos definir uma nova medida em X , denominada *pullback* de μ por T , da seguinte maneira:

$$T_*\mu(A) := \mu(T^{-1}(A)), \quad \forall A \in \mathcal{B}.$$

Em particular, quando T preserva μ , é claro que $T_*\mu = \mu$, implicando assim que, para toda função f integrável, ocorre

$$\int f \, d\mu = \int f \, dT_*\mu = \int f \circ T \, d\mu.$$

Uma questão pertinente é se sempre existe probabilidade invariante. De fato, como nosso SDT é composto por um espaço métrico compacto X e uma aplicação contínua T , o teorema de Krylov-Bogoliubov fornece resposta para essa pergunta.

Teorema (Krylov-Bogoliubov). *Sejam (X, T) um SDT e $x \in X$ um ponto qualquer. Denotando por δ_y a probabilidade que atribui valor 1 ao conjunto $\{y\}$ e 0 ao seu complementar, considere as probabilidades*

$$\mu_n := \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i x_n},$$

onde $\{x_n\}$ é uma sequência arbitrária de pontos em X . Então, qualquer ponto de acumulação de $\{\mu_n\}$ na topologia fraca-* é uma probabilidade invariante por T .

Para uma prova desse clássico teorema, o leitor pode consultar, por exemplo, [26].

Um dos primeiros resultados a ser discutido em um curso de teoria ergódica é o teorema de recorrência de Poincaré, o qual atesta que μ -q.t.p. tudo retorna.

Teorema 1.1.2 (Recorrência de Poincaré). *Sejam (X, T) um SDT e μ uma medida finita T -invariante. Então, para todo $E \in \mathcal{B}$, o conjunto*

$$F = \{x \in E : \text{há } n_0 \in \mathbb{N} \text{ tal que } T^n x \notin E, \forall n \geq n_0\}$$

possui medida nula.

A prova desse teorema pode ser vista em [26]. Demonstramos uma versão paralela, a qual fornece uma estimativa para o limite superior da sequência $\{\mu(A \cap T^n(A))\}_{n \in \mathbb{N}}$. Para tanto, precisamos do conceito abaixo.

Definição. Chamamos de isomorfismo mensurável de X toda aplicação $T: X \rightarrow X$ inversível tal que T e T^{-1} são mensuráveis. Dizemos ainda que o isomorfismo mensurável T preserva uma medida μ se T e T^{-1} preservam μ .

Teorema 1.1.3 (Recorrência Múltipla de Poincaré). *Sejam (X, \mathcal{B}, μ) um espaço de probabilidade e $T: X \rightarrow X$ um isomorfismo mensurável que preserva μ . Se $A \in \mathcal{B}$ é um conjunto de medida positiva, então,*

$$\limsup_{n \rightarrow \infty} \mu(A \cap T^n(A)) \geq \mu(A)^2.$$

Demonstração. Considere um conjunto $A \in \mathcal{B}$ com $\mu(A) > 0$. Observe que, dado $n \in \mathbb{N}$,

$$\int \sum_{i=0}^{n-1} \chi_{T^i(A)} d\mu = \sum_{i=0}^{n-1} \mu(T^i(A)) = n\mu(A),$$

pois $\mu(T(B)) = \mu(B)$, para todo $B \in \mathcal{B}$. Pela desigualdade de Cauchy, temos

$$n\mu(A) = \int \sum_{i=0}^{n-1} \chi_{T^i(A)} d\mu \leq \left(\int \left| \sum_{i=0}^{n-1} \chi_{T^i(A)} \right|^2 d\mu \right)^{\frac{1}{2}} \left(\int 1^2 d\mu \right)^{\frac{1}{2}}.$$

Elevando ambos os lados ao quadrado, resulta que

$$n^2 \mu(A)^2 \leq \int \left| \sum_{i=0}^{n-1} \chi_{T^i(A)} \right|^2 d\mu. \quad (1.1)$$

Utilizamos agora que o produto de funções características de dois conjuntos é a função característica da interseção destes dois conjuntos para reescrever o somatório em (1.1):

$$\left| \sum_{i=0}^{n-1} \chi_{T^i(A)} \right|^2 = \sum_{i=0}^{n-1} \chi_{T^i(A)}^2 + 2 \sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} \chi_{T^i(A) \cap T^j(A)}.$$

Sendo assim, integrando o somatório em (1.1) obtemos

$$\begin{aligned} n^2 \mu(A)^2 &\leq \int \left| \sum_{i=0}^{n-1} \chi_{T^i(A)} \right|^2 d\mu = n\mu(A) + 2 \sum_{i=0}^{n-2} \sum_{j=i+1}^{n-1} \mu(T^i(A) \cap T^j(A)) \\ &= n\mu(A) + 2 \sum_{i=1}^{n-1} (n-i) \mu(A \cap T^i(A)). \end{aligned}$$

A última igualdade ocorre, pois, para $j > i$, $\mu(T^i(A) \cap T^j(A)) = \mu(A \cap T^{j-i}(A))$. Logo,

$$\mu(A)^2 \leq \frac{\mu(A)}{n} + \sum_{i=1}^{n-1} \frac{2(n-i)}{n^2} \mu(A \cap T^i(A)).$$

Veja que, como $\sum_{i=0}^{n-1} \frac{2(n-i)}{n^2} = \frac{n+1}{n}$, tomando o limite superior na desigualdade acima, resulta que

$$\mu(A)^2 \leq \limsup_{n \rightarrow \infty} \left[\frac{\mu(A)}{n} + \sum_{i=1}^{n-1} \frac{2(n-i)}{n^2} \mu(A \cap T^i(A)) \right] = \limsup_{n \rightarrow \infty} \sum_{i=1}^{n-1} \frac{2(n-i)}{n^2} \mu(A \cap T^i(A)).$$

Então, dado $\varepsilon > 0$, há $M \in \mathbb{N}$ tal que $\mu(A \cap T^n(A)) < \limsup_{n \rightarrow \infty} \mu(A \cap T^n(A)) + \varepsilon$, $\forall n \geq M$. Portanto, temos

$$\begin{aligned} \limsup_{n \rightarrow \infty} \sum_{i=1}^{n-1} \frac{2(n-i)}{n^2} \mu(A \cap T^i(A)) &= \\ &= \limsup_{n \rightarrow \infty} \left[\sum_{i=1}^M \frac{2(n-i)}{n^2} \mu(A \cap T^i(A)) + \sum_{i=M+1}^{n-1} \frac{2(n-i)}{n^2} \mu(A \cap T^i(A)) \right] \\ &\leq \limsup_{n \rightarrow \infty} \left(\mu(A \cap T^n(A)) + \varepsilon \right) \frac{n+1}{n} \\ &= \limsup_{n \rightarrow \infty} \mu(A \cap T^n(A)) + \varepsilon. \end{aligned}$$

Como ε é arbitrário, resulta que $\limsup_{n \rightarrow \infty} \mu(A \cap T^n(A)) \geq \mu(A)^2$. \square

Esses resultados fornecem pistas de que o conjunto das medidas T -invariantes possui propriedades relevantes a serem estudadas. Sendo assim, dado um SDT (X, T) , denotamos por $\mathcal{M}_T(X)$ (ou, por simplicidade, apenas \mathcal{M}_T) o conjunto das probabilidades T -invariantes de X .

Dados μ medida sobre X e $U \in \mathcal{B}$ conjunto tal que $\mu(U) = \mu(X)$, então, se f é uma função definida sobre U , dizemos que f está μ -q.t.p. definida em X . Nesse caso, consideramos $f(x)$ na mesma classe de equivalência em $L^\infty(\mu)$ do representante $g(x)$, tal que $g(x) = f(x)$, se $x \in U$ e $g(x) = 0$, caso contrário. É exatamente nesse contexto que se enuncia o resultado a seguir.

Teorema (Tempo Médio de Retorno). *Sejam (X, T) um SDT e $\mu \in \mathcal{M}_T$. Dado $A \in \mathcal{B}$,*

a função tempo médio de retorno $\tau_A(x)$ é definida μ -q.t.p. sobre A :

$$\tau_A(x) := \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} \chi_A \circ T^i(x), \quad \mu - q.t.p., \quad x \in A. \quad (1.2)$$

Esta função é mensurável e cumpre $\int \tau_A d\mu = \mu(A)$.

Esse resultado é um caso particular do famoso teorema de Birkhoff (aplicado à função característica χ_A), o qual será enunciado a seguir.

Uma vez mais identificando o conjunto das medidas de X com o dual das funções contínuas, supomos \mathcal{M}_T munido da topologia fraca-*. Apenas para relembrar, a topologia fraca-* em \mathcal{M}_T possui base de vizinhanças (em $\mu \in \mathcal{M}_T$) da forma

$$B_\varepsilon(\mu, m, f_1, \dots, f_m) = \{\nu \in \mathcal{M}_T : |\mu(f_i) - \nu(f_i)| < \varepsilon, \forall i = 1, \dots, m\},$$

onde $\varepsilon > 0$, $m \in \mathbb{N}$ e $f_1, \dots, f_m: X \rightarrow \mathbb{R}$ são funções contínuas. Podemos mencionar, então, que, além de ser convexo, \mathcal{M}_T é compacto nesta topologia. O leitor pode consultar, por exemplo, em [26] prova da compacidade de \mathcal{M}_T na topologia fraca-*. Já a convexidade de \mathcal{M}_T é evidente, pois, para quaisquer $\mu, \nu \in \mathcal{M}_T$, $t \in (0, 1)$ e $A \in \mathcal{B}$,

$$T_*[t\mu + (1-t)\nu](A) = t\mu(T^{-1}(A)) + (1-t)\nu(T^{-1}(A)) = [t\mu + (1-t)\nu](A).$$

Além disso, a convexidade está intrinsecamente ligada ao conceito de probabilidades ergódicas.

Definição 1.1.4. *Seja (X, T) um SDT. Dizemos que uma probabilidade $\mu \in \mathcal{M}_T$ é ergódica ou que T é ergódica se todos os conjuntos invariantes por T possuem medida 0 ou 1, ou seja, se $T^{-1}(A) \subseteq A$ implicar $\mu(A) = 0$ ou $\mu(A) = 1$.*

Antes de explicar a relação entre a convexidade de \mathcal{M}_T e suas probabilidades ergódicas, enunciaremos o clássico teorema ergódico de Birkhoff. Para tanto, dados $f \in C(X)$ e $x \in X$, definimos a n -ésima soma de Birkhoff de f no ponto x por

$$S_n f(x) := \sum_{i=0}^{n-1} f \circ T^i(x).$$

Teorema 1.1.5 (Teorema Ergódico de Birkhoff). *Sejam (X, T) um SDT, $\mu \in \mathcal{M}_T$ e $f: X \rightarrow \mathbb{R}$ uma função integrável com respeito a μ . Então,*

(i) existe o limite $\bar{f}(x) := \lim_{n \rightarrow \infty} \frac{1}{n} S_n f(x)$ para μ -q.t.p. $x \in X$;

(ii) $\bar{f} \circ T = \bar{f}$;

(iii) $\int \bar{f} d\mu = \int f d\mu$.

(iv) μ ergódica implica $\bar{f} \equiv \int f d\mu$.

A estratégia da prova desse teorema passa por uma astuciosa decomposição das somas de Birkhoff em subconjuntos de \mathbb{N} , nos quais é possível controlar o somatório. Devido à quantidade de passos técnicos necessários para a prova do teorema, deixamos indicadas as referências [26, 35] para que o leitor possa compreender seus detalhes. O teorema de Birkhoff garante que se f é integrável, para quase todo ponto $x \in X$, a sua média temporal (média dos valores de f avaliada ao longo da órbita de x) é igual à média global de f (sua integral). *Grosso modo*, isto significa que as órbitas relevantes para μ encontram-se bem distribuídas sobre seu suporte em X , de modo que é possível avaliar integrais apenas tomando médias sobre essas órbitas.

Enunciamos abaixo resultado fundamental da teoria ergódica que relaciona a convexidade de \mathcal{M}_T e as probabilidades ergódicas.

Teorema (Teorema da Decomposição Ergódica). *Sejam (X, T) um SDT e $\mu \in \mathcal{M}_T$. Então, existem $\tilde{X} \in \mathcal{B}$ com $\mu(\tilde{X}) = 1$, uma partição \mathcal{P} de \tilde{X} , uma família de probabilidades $\{\mu_P : P \in \mathcal{P}\}$ e uma probabilidade $\hat{\mu}$ sobre \mathcal{P} tais que*

- $\mu_P(P) = 1$ para $\hat{\mu}$ -q.t.p. $P \in \mathcal{P}$;
- $P \mapsto \mu_P(A)$ é mensurável para todo conjunto $A \in \mathcal{B}$;
- μ_P é invariante e ergódica $\hat{\mu}$ -q.t.p. $P \in \mathcal{P}$;
- $\mu(A) = \int \mu_P(A) d\hat{\mu}$ para todo $A \in \mathcal{B}$.

Em sua versão geral, o teorema da decomposição ergódica garante que, asseguradas separabilidade e completude do espaço (ou das peças de uma possível descomposição do mesmo), para o estudo de qualquer propriedade linear, sem perda de generalidade podemos nos restringir apenas às probabilidades ergódicas. Na verdade, dessa decomposição resulta que as probabilidades ergódicas de \mathcal{M}_T são extremais, isto é, qualquer medida $\mu \in \mathcal{M}_T$ está no envoltório convexo das probabilidades ergódicas e, caso μ seja probabilidade

ergódica, μ só admite combinações convexas triviais (como ocorre, por exemplo, com os vértices de um poliedro convexo). Para uma prova desse teorema, o leitor pode consultar, por exemplo, [26].

Gostaríamos de registrar aqui um resultado geral de análise na reta que é bastante aplicado em teoria ergódica. Recorde que uma sequência real $\{a_n\}$ é dita ser subaditiva se cumpre, para quaisquer $m, n \in \mathbb{N}$, a desigualdade $a_{m+n} \leq a_m + a_n$.

Lema 1.1.6 (Lema Subaditivo). *Seja $\{a_n\}_{n \in \mathbb{N}}$ uma sequência subaditiva. Então, é bem definido o limite*

$$\lim_{n \rightarrow \infty} \frac{1}{n} a_n = \inf_{n \in \mathbb{N}} \frac{1}{n} a_n,$$

podendo este assumir valor $-\infty$.

Destacamos, por ora, que o supremo das somas de Birkhoff, ou seja, $\sup_{x \in X} S_n f(x)$, fornece exemplo de sequência subaditiva – voltaremos a discutir tal particularidade na próxima seção. A demonstração do lema acima pode ser consultada, por exemplo, em [29].

Por fim, fornecemos abaixo uma prova curta de mais um relevante resultado para o estudo de probabilidades ergódicas.

Lema 1.1.7 (Kakutani-Rokhlin). *Sejam (X, \mathcal{B}, μ) um espaço de probabilidade não atômico (μ não dá massa positiva para pontos) e $T: X \rightarrow X$ uma isomorfismo mensurável ergódico que preserva a medida μ . Dados $N \in \mathbb{N}$ e $\varepsilon > 0$, existe um conjunto $E \in \mathcal{B}$ tal que os conjuntos $\{E, T(E), \dots, T^{N-1}(E)\}$ são dois a dois disjuntos e $\mu\left(\bigsqcup_{i=0}^{N-1} T^i(E)\right) > 1 - \varepsilon$.*

Demonstração. Dados $N \in \mathbb{N}$ e $\varepsilon > 0$, tome $A \subseteq X$ cumprindo $0 < \mu(A) < \varepsilon/N$. Definimos os conjuntos

$$A_0 = \{x \in A : T^n(x) \notin A, \forall n \geq 1\} \quad \text{e}$$

$$A_k = \{x \in A : T^i(x) \notin A \text{ para } 1 \leq i \leq k-1, \text{ mas } T^k(x) \in A\}.$$

Esses conjuntos formam uma partição de A . Pelo teorema de recorrência de Poincaré, $\mu(A_0) = 0$. Além disso, note que $\mu\left(Y = \bigcup_{i=0}^{\infty} T^i(A)\right) = 1$, pois $\mu(A) > 0$ e Y é um conjunto invariante por T , a qual é uma aplicação ergódica. Definimos, então,

$$E = \bigcup_{i \geq 1} \bigcup_{k \geq iN} T^{k-iN}(A_k).$$

Repare que os conjuntos $E, T(E), \dots, T^{N-1}(E)$ são dois a dois disjuntos, pois E é definido como união de pontos $x \in A$ tais que $T^i x \notin A$ para $i = 1, \dots, N-1$, mas $T^N x \in A$. Logo, uma interseção não vazia entre dois desses conjuntos implicaria a existência de um ponto $x \in E$ com retorno ao conjunto A antes de N iteradas, o que é impossível. Por fim, temos que $\mu\left(\bigcup_{j=0}^{N-1} T^j(E)\right) > 1 - \varepsilon$, pois $\bigcup_{j=0}^{N-1} T^j(E)$ cobre Y a menos de um conjunto $W \subseteq \bigcup_{k \geq 1} \bigcup_{i=0}^{N-1} T^i(A_k)$, de modo que

$$\mu\left(\bigcup_{j=0}^{N-1} T^j(E)\right) = 1 - \mu(W) \geq 1 - N \sum_{k \geq 1} \mu(A_k) \geq 1 - N\mu(A) > 1 - \varepsilon.$$

□

1.2 Maximizando a integral de uma função

Nesta seção, queremos motivar o leitor ao estudo de otimização ergódica em um SDT apresentando uma abordagem do problema de maximizar o valor da integral de uma função $f: X \rightarrow \mathbb{R}$ sobre o conjunto \mathcal{M}_T . As principais referências adotadas para esta introdução à otimização ergódica são [15, 3].

Considere uma função $f \in C(X)$. Definimos o intervalo de rotação de f como o menor intervalo que contém todos os valores de aderência de $\{\frac{1}{n}S_n f(x), x \in X\}$, e o denotamos por $[\alpha(f), \beta(f)]$. Note que, se $f = g + h \circ T - h$, para alguma função h mensurável e limitada, então, para todo $x \in X$, temos

$$\frac{1}{n}S_n f(x) = \frac{1}{n}S_n(g + h \circ T - h)(x) = \frac{1}{n}S_n g(x) + \frac{h(x) - h(T^n x)}{n}.$$

Como h é limitada, para uma subsequência $\{n_j\}$ para a qual o limite existe, segue que

$$\lim_{n_j \rightarrow \infty} \frac{1}{n_j}S_{n_j} f(x) = \lim_{n_j \rightarrow \infty} \frac{1}{n_j}S_{n_j} g(x),$$

ou seja, neste caso, f e g partilham o mesmo intervalo de rotação.

No caso em que o sistema é unicamente ergódico, isto é, que \mathcal{M}_T é formado por uma única probabilidade (ergódica), este intervalo reduz-se a um ponto, pois, para todo $x \in X$, temos a convergência

$$\lim_n \frac{1}{n}S_n f(x) = \int f d\mu.$$

Como indica o próprio título da seção, queremos maximizar a integral de funções $f \in C(X)$. Na subseção 1.2.1, dedicamo-nos a provar o seguinte teorema.

Teorema 1.2.1. *Sejam (X, T) um SDT e $f \in C(X)$. Dado $\beta(f)$ extremo superior do intervalo de rotação de f como acima, temos a igualdade*

$$\beta(f) = \sup_{\mu \in \mathcal{M}_T} \int f \, d\mu, \quad (1.3)$$

sendo o supremo realizado por uma medida $\mu_0 \in \mathcal{M}_T$, dita probabilidade maximizante.

Já na subseção 1.2.2, avaliaremos $\beta(f)$ de maneira dual, vendo-o como o valor obtido ao se aplicar funcional não linear β sobre a função real contínua f , de modo a concluir que genericamente $f \in C(X)$ admite uma única probabilidade maximizante.

1.2.1 Probabilidades maximizantes

Sejam (X, T) um SDT e $f \in C(X)$. Para demonstrar o teorema 1.2.1, considere primeiramente a sequência $\{\beta_n f\}_{n \in \mathbb{N}}$ definida por

$$\beta_n f = \sup_{x \in X} S_n f(x).$$

Note que a sequência $\{\beta_n f\}$ é subaditiva:

$$\begin{aligned} \beta_{m+n} f &= \sup_{x \in X} \left(\sum_{k=0}^{m-1} f \circ T^k(x) + \sum_{k=m}^{m+n-1} f \circ T^k(x) \right) \\ &\leq \sup_{x \in X} \sum_{k=0}^{m-1} f \circ T^k(x) + \sup_{x \in X} \sum_{k=0}^{n-1} f \circ T^k(x) = \beta_m f + \beta_n f. \end{aligned}$$

Logo, pelo lema subaditivo, existe o limite $\lim_{n \rightarrow \infty} \frac{1}{n} \beta_n f = \inf_n \frac{1}{n} \beta_n f$. Esse limite ajuda-nos a demonstrar que $\beta(f) = \sup_{\mu \in \mathcal{M}_T} \int f \, d\mu$.

Lema 1.2.2. *Dados um SDT (X, T) e $f \in C(X)$, temos a igualdade*

$$\sup_{\mu \in \mathcal{M}_T} \int f \, d\mu = \lim_{n \rightarrow \infty} \frac{1}{n} \beta_n f = \beta(f).$$

Demonstração. Note que podemos escrever $\beta(f) = \sup_{x \in X} \limsup_{n \rightarrow \infty} \frac{1}{n} S_n f(x)$. Pela definição

de supremo, dado $\varepsilon > 0$, existe $x_0 \in X$ tal que

$$\beta(f) - \varepsilon \leq \limsup_{n \rightarrow \infty} \frac{1}{n} S_n f(x_0) \leq \limsup_{n \rightarrow \infty} \frac{1}{n} \beta_n f = \lim_{n \rightarrow \infty} \frac{1}{n} \beta_n f.$$

Tomamos, então, $\varepsilon \rightarrow 0$ para obter $\beta(f) \leq \lim_{n \rightarrow \infty} \frac{1}{n} \beta_n f$. Seja $\mu \in \mathcal{M}_T$ uma probabilidade ergódica. Assim, pelo teorema ergódico de Birkhoff, para μ -q.t.p. $x_0 \in X$, ocorre

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \frac{1}{n} S_n f(x_0) = \limsup_{n \rightarrow \infty} \frac{1}{n} S_n f(x_0) \leq \beta(f).$$

Recordando que as probabilidades ergódicas são extremais do convexo \mathcal{M}_T , resulta que

$$\sup_{\mu \in \mathcal{M}_T} \int f \, d\mu \leq \beta(f).$$

Por fim, sendo X compacto e f contínua, para cada $n \in \mathbb{N}$, existe $x_n \in X$ tal que $\beta_n f = S_n f(x_n)$. Então, para $\mu_n = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{T^i x_n}$, temos

$$\frac{1}{n} \beta_n f = \frac{1}{n} S_n f(x_n) = \int f \, d\mu_n.$$

Mas, como X é espaço métrico compacto e T é aplicação contínua, o teorema de Krylov-Bogoliubov garante que, na topologia fraca-*, qualquer ponto de acumulação μ da sequência $\{\mu_n\}$ é uma probabilidade invariante. Sem perda de generalidade, supomos que $\mu_n \rightharpoonup^* \mu$. Utilizando essa convergência, resulta que

$$\lim_{n \rightarrow \infty} \frac{1}{n} \beta_n f = \lim_{n \rightarrow \infty} \int f \, d\mu_n = \int f \, d\mu \leq \sup_{\mu \in \mathcal{M}_T} \int f \, d\mu,$$

donde conclui-se a igualdade entre os três valores. □

Apresentamos, a seguir, algumas propriedades desse valor maximal $\beta(f)$.

Teorema 1.2.3. *Sejam (X, T) um SDT e $f \in C(X)$. Temos que:*

- i. há probabilidade ergódica $\mu_0 \in \mathcal{M}_T$ tal que $\beta(f) = \mu_0(f)$;*
- ii. existe $x_0 \in X$ tal que $\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x_0) = \beta(f)$;*
- iii. para qualquer função estritamente crescente e convexa $\varphi: \mathbb{R} \rightarrow \mathbb{R}$, com inversa φ^{-1} ,*

ocorre

$$\lim_{n \rightarrow \infty} \frac{1}{n} \varphi^{-1} \left(\int \varphi \left(\sum_{i=0}^{n-1} f \circ T^i \right) d\mu_0 \right) = \beta(f),$$

em particular,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \ln \left(\int \exp \left(\sum_{i=0}^{n-1} f \circ T^i \right) d\mu_0 \right) = \beta(f),$$

Demonstração. (i) Assim como na prova do lema anterior, tomemos $\{x_n\}$ sequência, que satisfaz $S_n f(x_n) = \beta_n f$. Considere $\{g_j\}_{j \in \mathbb{N}}$ um conjunto denso em $C(X)$ – recorde que $C(X)$ é um espaço separável. Há uma subsequência $\{x_{n_{k,1}}\}$ de $\{x_n\}$ e um valor que denotamos por $\mu_0(g_1)$ tais que

$$\frac{1}{n_{k,1}} \sum_{i=0}^{n_{k,1}-1} g_1 \circ T^i(x_{n_{k,1}}) \rightarrow \mu_0(g_1), \quad \text{quando } k \rightarrow \infty.$$

De igual maneira, há subsequência $\{x_{n_{k,2}}\}$ de $\{x_{n_{k,1}}\}$ e valor $\mu_0(g_2)$ para os quais

$$\frac{1}{n_{k,2}} \sum_{i=0}^{n_{k,2}-1} g_2 \circ T^i(x_{n_{k,2}}) \rightarrow \mu_0(g_2), \quad \text{quando } k \rightarrow \infty.$$

Repetindo esse processo, construímos uma subsequência $\{n_k\}$ tal que, para toda função g_j do conjunto acima escolhido, existe o limite

$$\lim_{k \rightarrow \infty} \frac{1}{n_k} S_{n_k} g_j(x_{n_k}) = \mu_0(g_j).$$

Porém, como o conjunto $\{g_j\}_{j \in \mathbb{N}}$ é denso em $C(X)$, para qualquer $g \in C(X)$, dado $\varepsilon > 0$, há índice i tal que, para todo $k \in \mathbb{N}$, ocorre

$$\left| \frac{1}{n_k} (S_{n_k} g(x_{n_k}) - S_{n_k} g_i(x_{n_k})) \right| \leq \|g - g_i\|_0 < \frac{\varepsilon}{3}.$$

Mas, como a sequência $\{\frac{1}{n_k} S_{n_k} g_i(x_{n_k})\}$ é de Cauchy, há $k_0 > 0$ tal que, para todos $p, q > k_0$, temos

$$\left| \frac{1}{n_p} S_{n_p} g(x_{n_p}) - \frac{1}{n_q} S_{n_q} g_i(x_{n_q}) \right| < \frac{\varepsilon}{3}.$$

Sendo assim, para $p, q > k_0$, resulta que

$$\begin{aligned}
\left| \frac{1}{n_p} S_{n_p} g(x_{n_p}) - \frac{1}{n_q} S_{n_q} g(x_{n_q}) \right| &\leq \\
&\leq \left| \frac{1}{n_p} (S_{n_p} [g - g_i](x_{n_p})) \right| + \left| \frac{1}{n_p} S_{n_p} g_i(x_{n_p}) - \frac{1}{n_q} S_{n_q} g_i(x_{n_q}) \right| + \left| \frac{1}{n_q} (S_{n_q} [g_i - g](x_{n_q})) \right| \\
&< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.
\end{aligned}$$

Logo, $\left\{ \frac{1}{n_k} S_{n_k} g(x_{n_k}) \right\}$ define uma sequência de Cauchy, a qual converge para um limite que denotamos por $\mu_0(g)$. Por fim, a aplicação $g \mapsto \mu_0(g)$ define um funcional linear sobre $C(X)$. Consequentemente, como $g \geq 0$ implica $\mu \geq 0$, o teorema da representação de Riesz garante que $g \mapsto \mu_0(g)$ define uma medida tal que, por construção, $\mu_0(f) = \beta(f)$. É fácil ver que μ_0 é invariante por T e atribui valor 1 à função constante $g(x) \equiv 1$. Portanto, $\mu_0 \in \mathcal{M}_T$. Como mencionado na seção anterior, os pontos extremais do convexo \mathcal{M}_T são probabilidades ergódicas. Então, caso a definição de μ_0 não resulte em uma probabilidade ergódica, podemos passar a considerar uma das probabilidades μ envolvidas em sua decomposição ergódica, pois todas elas verificam a igualdade $\mu(f) = \beta(f)$, uma vez que já provamos a desigualdade $\mu(f) \leq \beta(f)$ para toda $\mu \in \mathcal{M}_T$. Logo, supomos μ_0 ergódica.

(ii) A existência de um ponto x_0 tal que $\lim_n \frac{1}{n} S_n f(x_0) = \beta(f)$ resulta do teorema de Birkhoff aplicado a μ_0 . Como μ_0 é probabilidade ergódica, o limite é satisfeito μ_0 -q.t.p. $x \in X$.

(iii) Considere μ_0 como no item *i*. Como φ é função convexa, pela desigualdade de Jensen, vale

$$\int \varphi \left(\sum_{k=0}^{n-1} f \circ T^k \right) d\mu_0 \geq \varphi \left(\int \sum_{k=0}^{n-1} f \circ T^k d\mu_0 \right) = \varphi \left(n \int f d\mu_0 \right).$$

Assim,

$$\liminf_{n \rightarrow \infty} \frac{1}{n} \varphi^{-1} \left(\int \varphi \left(\sum_{k=0}^{n-1} f \circ T^k \right) d\mu_0 \right) \geq \int f d\mu_0 = \mu_0(f) = \beta(f).$$

Por outro lado, para cada $n \in \mathbb{N}$, a monotonicidade de φ implica

$$\varphi^{-1} \left(\int \varphi \left(\sum_{k=0}^{n-1} f \circ T^k \right) d\mu_0 \right) \leq \varphi^{-1} \left(\int \varphi(\beta_n f) d\mu_0 \right) \leq \beta_n f,$$

donde segue que

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \varphi^{-1} \left(\int \varphi \left(\sum_{k=0}^{n-1} f \circ T^k \right) d\mu_0 \right) \leq \beta(f),$$

o que conclui a prova. \square

Para continuar a discussão a respeito do valor $\beta(f)$, fixamos a notação do conjunto que contém as probabilidades maximizantes de f , denotando-o por $\mathcal{M}_{max}(f)$. Sendo assim, $\mu \in \mathcal{M}_{max}(f) \Leftrightarrow \mu(f) = \beta(f)$. Mostramos, a seguir, como calcular o valor $\beta(f)$ por meio de um processo de dualidade.

Definição 1.2.4. *Uma função f é chamada de cobordo se $f = h \circ T - h$, para alguma função $h \in C(X)$. Além disso, duas funções são ditas coomólogas se sua diferença é um cobordo.*

Para cada $f \in C(X)$, definimos o seguinte valor:

$$\gamma(f) := \inf_{h \in C(X)} \sup_{x \in X} [f(x) + h \circ T(x) - h(x)]. \quad (1.4)$$

Note que esse valor $\gamma(f)$ é o ínfimo dos supremos das funções na classe de coomologia de f . Portanto, é óbvio que para $g = f + h \circ T - h$, com $h \in C(X)$, temos $\gamma(g) = \gamma(f)$. Observe ainda que a definição de $\gamma(f)$ depende apenas da classe de coomologia de f , sem mencionar qualquer propriedade de \mathcal{M}_T . A seguir, mostramos a relação de $\gamma(f)$ com as probabilidades invariantes.

Teorema 1.2.5. *Dada $f \in C(X)$, temos a igualdade $\beta(f) = \gamma(f)$.*

Demonstração. Escrevemos primeiramente $\gamma(f) = \inf_{h \in C(X)} \gamma(f, h)$, onde

$$\gamma(f, h) = \sup_{x \in X} [f(x) + h \circ T(x) - h(x)].$$

Segue que, para quaisquer $\mu \in \mathcal{M}_{max}(f)$ e $h \in C(X)$,

$$\beta(f) = \int f d\mu = \int (f + h \circ T - h) d\mu \leq \gamma(f, h).$$

Como $\beta(f)$ é limitado por $\gamma(f, h)$ e não depende de h , então, $\beta(f) \leq \inf_h \gamma(f, h) = \gamma(f)$. A desigualdade oposta pode ser obtida, por exemplo, de astúcia ao se calcular médias.

Dado $L \in \mathbb{N}$, defina a função $h_L = \frac{1}{L} \sum_{n=1}^L S_n f$. Temos que

$$\begin{aligned} \frac{1}{L} \sum_{n=1}^L f \circ T^n + h_L - h_L \circ T &= \\ &= \frac{1}{L} \left(\sum_{n=1}^L f \circ T^n + \sum_{n=1}^L S_n f - \sum_{n=1}^L S_n (f \circ T) \right) \\ &= \frac{1}{L} \left(\sum_{n=1}^L f \circ T^n + \sum_{n=1}^L \sum_{k=0}^{n-1} f \circ T^k - \sum_{n=1}^L \sum_{k=0}^{n-1} f \circ T^{k+1} \right) \\ &= \frac{1}{L} \left(\sum_{n=1}^L f \circ T^n + \sum_{n=1}^L (f - f \circ T^n) \right) = f. \end{aligned}$$

Portanto, para todo $L \in \mathbb{N}$,

$$\gamma(f) \leq \gamma(f, h_L) = \sup_{x \in X} [f(x) + h_L \circ T(x) - h_L(x)] = \sup_{x \in X} \frac{1}{L} \sum_{n=1}^L f \circ T^n(x) = \frac{1}{L} \beta_L f,$$

donde, ao fazer $L \rightarrow \infty$, obtém-se $\gamma(f) \leq \beta(f)$. \square

Esta última caracterização do valor máximo da integral de uma função $f \in C(X)$ sobre as probabilidades invariantes permite-nos aplicar técnicas de análise funcional para estudar o conjunto $\mathcal{M}_{max}(f)$. Em particular, estudaremos na próxima subseção $\beta(\cdot)$ como um funcional sobre o espaço de Banach $C(X)$.

Observação. Antes de estudar as propriedades de β como funcional, gostaríamos de observar que, se existe função $h \in C(X)$ tal que $\gamma(f) = \gamma(f, h)$, então, a função g definida por $g(x) := f(x) + h \circ T(x) - h(x) - \beta(f)$ possui as seguintes propriedades:

$$g \in C(X), \quad g \leq 0 \quad \text{e} \quad \mu \in \mathcal{M}_{max}(f) \Leftrightarrow g|_{supp(\mu)} \equiv 0.$$

A continuidade é óbvia, pois f e h são contínuas. O fato de g não ser positiva também é claro, pois $g = [f + h \circ T - h] - \sup_{x \in X} [f(x) + h \circ T(x) - h(x)]$. Além disso, como para toda $\mu \in \mathcal{M}_T$

$$\int g \, d\mu = \int f \, d\mu - \beta(f),$$

se $\mu \in \mathcal{M}_{max}(f)$, a continuidade de g determina seu valor nulo sobre o suporte de μ . Por outro lado, se há probabilidade $\nu \in \mathcal{M}_T$ com $g|_{supp(\nu)} \equiv 0$, obrigatoriamente, $\int f \, d\nu = \beta(f)$, donde $\nu \in \mathcal{M}_{max}(f)$.

1.2.2 Analisando o valor máximo da integral como um funcional

Como já mencionamos no texto, graças ao teorema da representação de Riesz, o conjunto das medidas com sinal sobre X pode ser identificado ao espaço dual de $C(X)$. Estudamos aqui algumas relações entre $\beta: C(X) \rightarrow \mathbb{R}$ e o conjunto de probabilidades invariantes \mathcal{M}_T . Recorde que

$$\begin{aligned} \beta: C(X) &\rightarrow \mathbb{R} \\ f &\mapsto \beta(f) = \sup_{\mu \in \mathcal{M}_T} \int f \, d\mu. \end{aligned}$$

Vejam primeiro algumas propriedades básicas.

Proposição 1.2.6. *A aplicação β é lipschitziana, positivamente homogênea, convexa e subaditiva.*

Demonstração. Demonstremos primeiro que β é lipschitziana. Sejam $\mu_f, \mu_g \in \mathcal{M}_T$ probabilidades tais que $\beta(f) = \mu_f(f)$ e $\beta(g) = \mu_g(g)$. Temos que

$$\beta(f) - \beta(g) = \mu_f(f) - \mu_g(g) \leq \mu_f(f) - \mu_f(g) = \mu_f(f - g) \leq \|f - g\|_\infty.$$

Como f e g possuem papéis simétricos na desigualdade acima, podemos concluir que $|\beta(f) - \beta(g)| \leq \|f - g\|_\infty$.

A homogeneidade é óbvia, pois, se μ_f maximiza a integral de f e a é uma constante não negativa, então μ_f também maximiza a integral de af (lembre que integrais são homogêneas).

Para provar a subaditividade, considere $f, g \in C(X)$ e μ_{f+g}, μ_f e μ_g probabilidades que maximizam as integrais de $f + g$, f e g , respectivamente. Segue, então, que

$$\beta(f + g) = \mu_{f+g}(f + g) = \mu_{f+g}(f) + \mu_{f+g}(g) \leq \mu_f(f) + \mu_g(g) = \beta(f) + \beta(g).$$

Por fim, se $t \in (0, 1)$ e $f, g \in C(X)$, pela subaditividade e homogeneidade, temos

$$\beta(tf + (1 - t)g) \leq \beta(tf) + \beta((1 - t)g) = t\beta(f) + (1 - t)\beta(g).$$

□

Para uma rápida revisão de conceitos de análise funcional, relembramos o que são

funcionais tangentes.

Definição 1.2.7. Fixada $f \in C(X)$, chamamos de funcionais tangentes a β em f as probabilidades $\mu \in \mathcal{M}_T$ tais que

$$\beta(f + g) \geq \beta(f) + \mu(g), \quad \forall g \in C(X).$$

O lema a seguir caracteriza os funcionais tangentes a β em f .

Lema 1.2.8. Dada $f \in C(X)$, os funcionais tangentes a β em f são exatamente as probabilidades pertencentes ao conjunto $\mathcal{M}_{max}(f)$, isto é, as probabilidades T -invariantes que maximizam a integral de f .

Demonstração. Argumentamos que, como $C(X)$ é espaço vetorial, $f \in C(X)$ implica que $-f \in C(X)$. Seja, então, $\mu \in \mathcal{M}_T$ uma probabilidade tangente a β em f . Temos

$$0 = \beta(0) = \beta(f - f) \geq \beta(f) - \mu(f).$$

Logo, $\mu(f) \geq \beta(f)$, o que é verificado se, e somente se, $\mu \in \mathcal{M}_{max}(f)$.

Ademais, se $\mu \in \mathcal{M}_{max}(f)$, para qualquer $g \in C(X)$ ocorre

$$\beta(f) + \mu(g) = \mu(f) + \mu(g) = \mu(f + g) \leq \beta(f + g),$$

verificando a condição de μ ser tangente a β em f . □

Gostaríamos de ressaltar para o leitor que não traremos muitos detalhes a respeito das técnicas de análise funcional aqui utilizadas. O foco estará em aplicar tais técnicas para se obter alguns resultados em otimização ergódica.

Dedicamos o resto desta subseção a estabelecer um critério para assegurar que uma determinada função $f \in C(X)$ possua apenas uma probabilidade maximizante.

Notação. Para que não fiquem dúvidas, ao leitor, como os resultados que apresentamos a seguir são propriedades gerais dos espaços de Banach, usamos a notação E para o espaço considerado e, como de costume, E^* para seu dual. Além disso, permitimo-nos escrever $\xi(x)$ ou $\langle \xi, x \rangle$ para denotar a aplicação do funcional linear $\xi \in E^*$ no ponto $x \in E$. As funções, por ora, serão preferencialmente denotadas por $\phi: E \rightarrow \mathbb{R}$ e continuam com a notação usual $\phi(x)$, $\forall x \in E$.

Recordamos primeiramente dois teoremas clássicos de análise funcional: o teorema de Hahn-Banach e o teorema de Banach-Alaoglu. Para demonstrações desses teoremas, consulte, por exemplo, [10].

Teorema de Hahn-Banach. *Sejam E um espaço de Banach, $U \subseteq E$ um subespaço vetorial e $\phi: E \rightarrow \mathbb{R}$ um funcional sublinear, ou seja, positivamente homogêneo e subaditivo. Seja $\varphi: U \rightarrow \mathbb{R}$ um funcional linear que é majorado por ϕ em U , isto é, $\varphi(x) \leq \phi(x)$, $\forall x \in U$. Então, há um funcional linear $\psi: E \rightarrow \mathbb{R}$ satisfazendo*

$$\psi(x) = \varphi(x), \quad \forall x \in U \quad e \quad \psi(x) \leq \phi(x), \quad \forall x \in E.$$

Teorema de Banach-Alaoglu. *Sejam E um espaço de Banach e B^* a bola unitária do espaço dual de E . Então, B^* é relativamente compacta, isto é, toda sequência de B^* admite subsequência convergente.*

Definimos, a seguir, o que vem a ser a derivada de Gateaux e o subdiferencial de uma função convexa $\phi: E \rightarrow \mathbb{R}$ do espaço de Banach E .

Definição 1.2.9. *Sejam E um espaço de Banach real, $D \subseteq E$ um subconjunto aberto, convexo e não vazio, e $\phi: E \rightarrow \mathbb{R}$ uma função convexa. Dizemos que ϕ é diferenciável segundo Gateaux em $x_0 \in D$ se existe o limite*

$$\lim_{t \rightarrow 0} \frac{\phi(x_0 + tx) - \phi(x_0)}{t} = d\phi(x_0)(x), \quad \forall x \in E.$$

A função $d\phi(x_0)$ é chamada derivada de Gateaux de ϕ em x_0 . Também podemos definir as funções $d\phi^+(x_0)(x)$ e $d\phi^-(x_0)(x)$, tomando o limite em t tendendo a zero pela direita ou esquerda, respectivamente. É importante frisar que esses limites laterais sempre existem devido à convexidade de ϕ .

O subdiferencial de ϕ em um ponto x_0 , denotado por $\partial\phi(x_0)$, é definido como o conjunto de todos os $\xi \in E^*$, tais que

$$\langle \xi, y - x_0 \rangle + \phi(x_0) \leq \phi(y), \quad \text{para todo } y \in E.$$

Cada $\xi \in \partial\phi(x_0)$ é, por definição, um funcional tangente a ϕ em x_0 .

Observe que, se $\phi: D \rightarrow \mathbb{R}$ é uma função convexa, então sua derivada lateral à direita $d\phi^+(x)(\cdot)$ é sublinear. De fato, para qualquer constante $c \geq 0$, $d\phi^+(x_0)(cy) = cd\phi^+(x_0)(y)$.

Além disso, considerando $x_0 \in D$, $t > 0$ e $y, z \in E$, com $x_0 + ty \in D$ e $x_0 + tz \in D$, obtemos

$$\frac{\phi(x_0 + ty) + \phi(x_0 + tz) - 2\phi(x_0)}{t} \geq \frac{\phi\left(x_0 + \frac{t}{2}(y + z)\right) - \phi(x_0)}{t/2},$$

donde, ao passar ao limite quando $t \rightarrow 0^+$, resulta

$$d\phi^+(x_0)(y + z) \leq d\phi^+(x_0)(y) + d\phi^+(x_0)(z), \quad \forall y, z \in E.$$

Argumentação análoga garante que $d\phi^-(x_0)(\cdot)$ é função superlinear.

O lema e as proposições a seguir são baseados em [28]. Repare que estamos utilizando o fato de $C(X)$ ser um espaço de Banach.

Lema 1.2.10. *Sejam $D \subseteq \mathbb{R}$ um intervalo aberto e $f : D \rightarrow \mathbb{R}$ uma função convexa. Então, $f'(x)$ existe a menos de um conjunto enumerável de pontos em D .*

Demonstração. Começamos mostrando que a função $df^+(x)(1)$ (derivada lateral à direita), denotada simplesmente por $df^+(x)$, é não decrescente em x . Mais precisamente, para pontos $x, y \in D$, com $x > y$, queremos mostrar que $df^+(x) \geq df^+(y)$. Assumiremos, sem perda de generalidade, que $y = 0 \in D$ e que $f(0) = 0$ (isso pode ser feito via translações, que não afetam o comportamento da derivada). Considere um ponto $x \in D$, com $x > 0$. A convexidade de f garante que

$$df^+(0) = \lim_{t \rightarrow 0^+} \frac{f(t)}{t} = \lim_{t \rightarrow 0^+} \frac{f\left(\left(1 - \frac{t}{x}\right)0 + \frac{t}{x}x\right)}{t} \leq \frac{f(x)}{x}.$$

Logo, é suficiente mostrar que

$$\frac{f(x)}{x} \leq \frac{f(x+t) - f(x)}{t}, \quad \forall t > 0.$$

Tomando $\kappa = x/(x+t)$, note que podemos escrever $x = \kappa(x+t) + (1-\kappa)0$. Mais uma vez utilizando a convexidade de f , temos

$$\frac{f(x)}{x} \leq \frac{1}{x} (\kappa f(x+t) + (1-\kappa)f(0)) = \frac{f(x+t)}{x+t}.$$

Observe que a desigualdade acima é equivalente à desigualdade desejada.

Como $df^+(x)(\cdot)$ é função sublinear, ocorre

$$df^+(x)(1) \geq -df^+(x)(-1) = -\lim_{t \rightarrow 0^+} \frac{f(x-t) - f(x)}{t} = df^-(x)(1).$$

De posse dessa desigualdade, e sabendo que f é diferenciável em x_0 se, e só se, as derivadas laterais coincidem, é óbvio que f não possui derivada em um ponto x_0 se, e só se, $df^-(x_0) < df^+(x_0)$. A estratégia então consiste em mostrar que

$$df^-(x_0) \geq \lim_{x \rightarrow x_0^-} df^+(x), \quad (1.5)$$

ficando claro que a função f não é diferenciável segundo Gateaux em x_0 se, e somente se, df^+ possuir um salto em x_0 .

Assumindo por ora a desigualdade (1.5), note que uma função não decrescente φ definida em um intervalo fechado $[a, b]$ possui, no máximo, uma quantidade enumerável de saltos. Com efeito, sejam $s = \varphi(b)$ e $r = \varphi(a)$. Considere os subconjuntos de $[a, b]$ definidos por

$$A_k := \left\{ x \in [a, b] : \varphi(x) - \lim_{y \rightarrow x^-} \varphi(y) \geq \frac{s-r}{k} \right\}.$$

Como φ é função não decrescente, A_k possui no máximo k elementos. Definindo então o conjunto $A = \cup_{k \in \mathbb{N}} A_k$, sabemos que A é enumerável (pois é uma união enumerável de conjuntos enumeráveis) e contém todos os pontos $x \in [a, b]$ para os quais φ possui um salto. Para aplicar esta observação à função não decrescente df^+ , basta escrever o intervalo aberto D como união enumerável de intervalos compactos e aplicar a conclusão para df^+ restrita a cada um deles, resultando, então, que $f'(x)$ existe a menos de um conjunto enumerável.

Para agora demonstrar (1.5), observe primeiramente que, quando $t \rightarrow 0^+$, temos

$$\frac{f(x_0+t) - f(x_0)}{t} \downarrow df^+(x_0) \quad \text{e} \quad -\frac{f(x_0-t) - f(x_0)}{t} \uparrow df^-(x_0).$$

Logo, para $x < x_0$, apenas é preciso mostrar que há $t_0 > 0$ tal que

$$\frac{f(x+t_0) - f(x)}{t_0} \leq -\frac{f(x_0-t_0) - f(x_0)}{t_0}.$$

Porém, a desigualdade acima é obtida ao considerar $t_0 = \frac{1}{2}(x_0 - x)$ e aplicar a convexidade de f no ponto $\frac{1}{2}(x_0 + x)$. \square

Proposição 1.2.11. *Sejam E um espaço de Banach, $D \subseteq E$ um aberto convexo e*

$\varphi: D \rightarrow \mathbb{R}$ uma função convexa que é contínua em $x_0 \in D$. Então, φ é localmente lipschitziana em x_0 , ou seja, há constantes $K, \delta > 0$ tais que, se $y, z \in B_\delta(x_0) \subseteq D$, então,

$$|\varphi(y) - \varphi(z)| \leq K\|y - z\|.$$

Demonstração. Sendo φ contínua em x_0 , há $\delta > 0$ e $K_1 > 0$ tais que $|\varphi(y)| \leq K_1$, $\forall y \in B_{2\delta}(x_0) \subseteq D$. Considere $z, y \in B_\delta(x_0)$ dois pontos distintos e denote por d a distância entre eles. Tome ainda $w := y + \frac{\delta}{d}(y - z)$. Por definição, $w \in B_{2\delta}(x_0)$. Utilizando a convexidade de φ , temos que

$$\varphi(y) \leq \frac{d}{d+\delta}\varphi(w) + \frac{\delta}{d+\delta}\varphi(z),$$

o que produz

$$\varphi(y) - \varphi(z) \leq \frac{d}{d+\delta}[\varphi(w) - \varphi(z)] \leq \frac{2dK_1}{\delta} = \frac{2K_1}{\delta}\|y - z\|.$$

O resultado segue ao notar que y e z cumprem papéis simétricos e ao tomar $K = 2K_1/\delta$. \square

Proposição 1.2.12. *Sejam E um espaço de Banach e $D \subseteq E$ um aberto convexo. Uma função contínua e convexa $\phi: D \rightarrow \mathbb{R}$ é diferenciável segundo Gateaux em um ponto $x_0 \in D$ se, e somente se, o subdiferencial $\partial\phi(x_0)$ é composto por um único funcional $\xi \in E^*$.*

Demonstração. Primeiramente, suponha que ϕ é diferenciável segundo Gateaux em um ponto $x_0 \in D$ e considere $\xi \in \partial\phi(x_0)$. Para qualquer $y \in D$, a definição de subdiferencial, ao ser aplicada no ponto $ty + x_0 \in D$, fornece a desigualdade abaixo

$$\phi(ty + x_0) - \phi(x_0) \geq \langle \xi, ty \rangle.$$

Sendo assim, por um lado, temos

$$\lim_{t \rightarrow 0^+} \frac{\phi(ty + x_0) - \phi(x_0)}{t} \geq \langle \xi, y \rangle,$$

e, por outro, verificamos

$$\lim_{t \rightarrow 0^-} \frac{\phi(ty + x_0) - \phi(x_0)}{t} \leq \langle \xi, y \rangle.$$

Como os limites acima coincidem, para qualquer $\xi \in \partial\phi(x_0)$, temos a igualdade $d\phi(x_0)(y) =$

$\langle \xi, y \rangle, \forall y \in D$, isto é, $\partial\phi(x_0)$ é um conjunto unitário contendo apenas $d\phi(x_0)$.

Suponha agora que $\partial\phi(x_0)$ é formado de um único elemento. Por hipótese, há único funcional linear $\xi \in E^*$ que é majorado por $d\phi^+(x_0)$:

$$d\phi^+(x_0)(y) \geq \langle \xi, y \rangle, \quad \forall y \in E.$$

Recorde que $d\phi^+(x_0)$ é um funcional sublinear. Pela proposição anterior, resulta também que $d\phi^+(x_0)(y)$ é contínuo na variável y . Essas duas propriedades sobre $d\phi^+(x_0)$ são suficientes para garantir que este é de fato um funcional linear. Com efeito, se $d\phi^+(x_0)$ não fosse linear, haveria $y, z \in E$ tais que

$$d\phi^+(x_0)(y+z) < d\phi^+(x_0)(y) + d\phi^+(x_0)(z).$$

Dessa desigualdade estrita podemos concluir que ocorre

$$-d\phi^+(x_0)(-y) < d\phi^+(x_0)(y) \quad \text{ou} \quad -d\phi^+(x_0)(-z) < d\phi^+(x_0)(z),$$

pois, se ocorresse $d\phi^+(x_0)(y) + d\phi^+(x_0)(-y) = 0$ e $d\phi^+(x_0)(z) + d\phi^+(x_0)(-z) = 0$, obteríamos a contradição

$$\begin{aligned} 0 &= d\phi^+(x_0)(y) + d\phi^+(x_0)(-y) + d\phi^+(x_0)(z) + d\phi^+(x_0)(-z) > \\ &> d\phi^+(x_0)(y+z) + d\phi^+(x_0)(-y-z) \geq d\phi^+(x_0)(0) = 0. \end{aligned}$$

Supomos $d\phi^+(x_0)(y) > -d\phi^+(x_0)(-y)$. Note agora que, sobre a reta $U := \{ry : r \in \mathbb{R}\}$, os funcionais lineares

$$\zeta_1(ry) := rd\phi^+(x_0)(y) \quad \text{e} \quad \zeta_2(ry) := -rd\phi^+(x_0)(-y)$$

são majorados por $d\phi^+(x_0)$. De fato, para $r \geq 0$, temos

$$\zeta_1(ry) = rd\phi^+(x_0)(y) = d\phi^+(x_0)(ry)$$

e, por outro lado,

$$\zeta_1(-ry) = -rd\phi^+(x_0)(y) < rd\phi^+(x_0)(-y) = d\phi^+(x_0)(-ry).$$

Argumentação análoga produz a majoração para ζ_2 . Basta agora aplicar o teorema de Hahn-Banach para estender os funcionais lineares ζ_1 e ζ_2 para todo E , preservando a majoração por $d\phi^+(x_0)$.

Logo, a suposição de que $d\phi^+(x_0)$ é não linear, leva-nos a concluir, de forma contraditória, que há mais de um funcional linear majorado por $d\phi^+(x_0)$. Segue, então, da linearidade, que

$$d\phi^+(x_0)(y) + d\phi^+(x_0)(-y) = 0, \quad \forall y \in E,$$

donde obtemos que $d\phi^+(x_0)(y) = -d\phi^+(x_0)(-y) = d\phi^-(x_0)(y)$, de modo que ϕ é diferenciável segundo Gateaux no ponto x_0 . Em particular, $d\phi(x_0) = \xi$. \square

Apresentamos agora o teorema devido a Mazur [24]. Relembre que um subconjunto de um espaço topológico é dito ser G_δ se for uma interseção enumerável de abertos densos, e F_σ se for o complementar de um conjunto G_δ . Recorde ainda que todo espaço de Banach possui a propriedade de Baire, ou seja, que todo conjunto G_δ é denso.

Teorema 1.2.13 (Mazur). *Sejam E um espaço de Banach separável e $\phi: D \rightarrow \mathbb{R}$ uma função convexa e contínua definida em um aberto convexo $D \subseteq E$. Então, a coleção dos pontos $x \in D$ tais que a derivada $d\phi(x)$ existe forma um conjunto G_δ em D .*

Demonstração. Basta argumentar que o conjunto dos pontos $x \in D$ para os quais $d\phi(x)$ não existe é um conjunto F_σ , isto é, uma união enumerável de fechados com interior vazio.

Seja $\{x_n\}$ uma sequência densa na esfera unitária de E . Dados $n, m \geq 1$, seja $A_{n,m}$ o conjunto dos pontos $x \in D$ para os quais existem $\xi, \zeta \in \partial\phi(x)$, satisfazendo $\langle \xi - \zeta, x_n \rangle \geq \frac{1}{m}$. Como a existência de $d\phi(x)$ está associada à unitariedade do conjunto $\partial\phi(x)$, vemos que $d\phi(x)$ não existe se, e só se, $x \in \cup_{n,m \geq 1} A_{n,m}$.

Afirmamos que cada $A_{n,m}$ é fechado. Com efeito, dados $m, n \in \mathbb{N}$, tome uma sequência $\{z_k\}_{k \in \mathbb{N}} \subseteq A_{n,m}$ tal que $z_k \rightarrow z \in D$. Então, para cada $k \in \mathbb{N}$ escolhemos $\xi_k, \zeta_k \in \partial\phi(z_k)$, tais que $\langle \xi_k - \zeta_k, x_n \rangle \geq \frac{1}{m}$. Perceba que as sequências $\{\xi_k\}_{k \in \mathbb{N}}$ e $\{\zeta_k\}_{k \in \mathbb{N}}$ são limitadas, pois, para todo $k \in \mathbb{N}$, temos

$$\|\xi_k\| = \sup_{\|w\|=1} \langle \xi_k, w \rangle = \sup_{\|y-z_k\|=1} \langle \xi_k, y - z_k \rangle \leq \sup_{y \in E} \phi(y) - \phi(z_k) \leq 2\|\phi\|_0.$$

O mesmo ocorre para ζ_k . Logo, pelo teorema de Banach-Alaoglu, podemos assumir, sem perda de generalidade, que $\xi_k \rightarrow \xi \in E^*$ e $\zeta_k \rightarrow \zeta \in E^*$. Segue então que, para todo

$y \in E$, temos:

$$\langle \xi, y - z \rangle = \lim_{k \rightarrow \infty} \langle \xi_k, y - z_k \rangle \leq \lim_{k \rightarrow \infty} [\phi(y) - \phi(z_k)] = \phi(y) - \phi(z).$$

Mas a desigualdade acima implica que $\xi \in \partial\phi(z)$. Da mesma forma, temos $\zeta \in \partial\phi(z)$. Então, como $\xi, \zeta \in \partial\phi(z)$ e

$$\langle \xi - \zeta, x_n \rangle = \lim_{k \rightarrow \infty} \langle \xi_k - \zeta_k, x_n \rangle \geq \frac{1}{m},$$

concluimos que $z \in A_{n,m}$. Logo, $A_{n,m}$ é um conjunto fechado.

Veremos agora que $D \setminus A_{n,m}$ é denso em D . Sejam $n \geq 1$ e $x_0 \in D$. Considere um intervalo I contendo a origem tal que $tx_n + x_0 \in D, \forall t \in I$. Definimos sobre I a função

$$\tilde{\phi}(t) = \phi(x_0 + tx_n).$$

Pelo lema 1.2.10, $\tilde{\phi}$ é diferenciável a menos de uma quantidade enumerável de pontos. Em particular, podemos aproximar x_0 por pontos do tipo $\hat{x} = x_0 + tx_n$, para os quais a derivada $\tilde{\phi}'(t)$ está definida. Logo, dados $\xi, \zeta \in \partial\phi(\hat{x})$, consideramos suas restrições sobre o segmento $R := \{x_0 + tx_n : t \in I\}$, denotando-as por $\tilde{\xi}$ e $\tilde{\zeta}$, respectivamente. Segue, então, que $\tilde{\xi}, \tilde{\zeta} \in \partial\tilde{\phi}(\hat{x})$. Mas lembre que $\tilde{\phi}'$ está definida em \hat{x} , de modo que $\tilde{\phi}$ possui derivada de Gateaux em \hat{x} . Logo, pela proposição 1.2.10, $\partial\tilde{\phi}(\hat{x})$ contém apenas um elemento. Assim, $\tilde{\xi} = \tilde{\zeta}$ sobre R , em particular, $\langle \tilde{\xi}, x_0 \rangle = \langle \tilde{\zeta}, x_0 \rangle$. Então, pela linearidade dos funcionais ξ e ζ , obtemos $\langle \xi, x_n \rangle = \langle \zeta, x_n \rangle$, donde segue que $\hat{x} \in D \setminus A_{n,m}, \forall m \in \mathbb{N}$. Com isto, vemos que um ponto arbitrário $x_0 \in D$ pode ser aproximado por elementos de $D \setminus A_{n,m}$, ou seja, que $D \setminus A_{n,m}$ é denso em D . Portanto, cada conjunto $A_{n,m}$, com n e m fixos, possui o interior vazio. \square

Nosso próximo passo consiste em aplicar o teorema de Mazur ao funcional β .

Corolário 1.2.14. *Seja (X, T) um SDT. Há um subconjunto G_δ em $C(X)$ tal que toda função a este pertencente admite uma única probabilidade maximizante.*

Demonstração. Veja que todas as hipóteses do teorema 1.2.13 são satisfeitas para o funcional β . Temos que $C(X)$ é um espaço de Banach separável e $\beta: C(X) \rightarrow \mathbb{R}$ é uma função contínua e convexa. Portanto, há uma interseção de abertos densos $V \subseteq C(X)$ tal que, para toda $f \in V$, β é diferenciável segundo Gateaux em f . Mas, então, a

proposição 1.2.12 garante que $\partial\beta(f)$ é constituído de um único elemento. Pelo lema 1.2.8, este elemento é exatamente a probabilidade $\mu \in \mathcal{M}_{max}(f)$. \square

Para dinâmicas expansiva, Contreras [14] obteve um importante avanço na teoria ao provar, em 2013, a existência de um subconjunto genérico das funções lipschitzianas de X , no qual cada função admite uma única probabilidade periódica maximizante.

1.3 Cociclos e o conjunto de Aubry

Nesta seção, abordamos um conceito bastante usual ao estudar otimização ergódica: o conjunto de Aubry. Veremos que toda medida maximizante para a integral de uma função $f \in C(X)$ deve estar suportada neste conjunto. O leitor encontrará mais informações sobre o conjunto de Aubry, por exemplo, em [18, 22].

Adotamos, de agora em diante, a nomenclatura de potencial para a função $f \in C(X)$, a qual é objeto de estudo de otimização. Esta nomenclatura é usual na literatura de otimização ergódica, possuindo motivações em resultados da física.

1.3.1 O conjunto de Aubry

Começamos definindo o que vem a ser o conjunto de Aubry.

Definição 1.3.1. *Seja (X, T) um SDT. Dado um potencial $f \in C(X)$, dizemos que um ponto $x \in X$ é ponto de Aubry de f se, para todo $\varepsilon > 0$, existir ponto $y \in X$ e inteiro $n \geq 1$ tais que*

$$d(x, y) < \varepsilon, \quad d(x, T^n y) < \varepsilon \quad e \quad |S_n(f - \beta(f))(y)| < \varepsilon.$$

Denotamos o conjunto de Aubry de f por $\Omega(f)$, o conjunto constituído de todos os pontos de Aubry.

O conjunto de Aubry é um subconjunto dos pontos não errantes de X , ou seja, o conjunto dos pontos $x \in X$ tais que, para qualquer aberto U contendo x , há $n \in \mathbb{N}$ com $T^n(U) \cap U \neq \emptyset$. Vejamos algumas propriedades básicas do conjunto de Aubry.

Lema 1.3.2. *Sejam (X, T) um SDT e $f \in C(X)$. Então, o conjunto de Aubry $\Omega(f)$ satisfaz:*

- $\Omega(f)$ é compacto;
- $T(\Omega(f)) \subseteq \Omega(f)$;
- se $g = f + h \circ T - h + c$, com $h \in C(X)$ e $c \in \mathbb{R}$, então, $\Omega(f) = \Omega(g)$.

Mais ainda, se T é uma aplicação aberta, temos

- $\Omega(f) = T(\Omega(f))$.

Demonstração. Como X é compacto, basta mostrar que $\Omega(f)$ é fechado para obter sua compacidade. Seja $\{x_n\} \subseteq \Omega(f)$ uma seqüência que converge para um ponto $x \in X$. Dado $\varepsilon > 0$, pela definição de $\Omega(f)$, para cada x_n , há $y_n \in X$ e $k_n \in \mathbb{N}$ com $d(x_n, y_n) < \varepsilon/2$, $d(x_n, T^{k_n}y_n) < \varepsilon/2$ e $|S_{k_n}(f - \beta(f))(y_n)| < \varepsilon/2$. Uma vez que $x_n \rightarrow x$, há N suficientemente grande tal que $d(x_N, x) < \varepsilon/2$. Logo, utilizando a desigualdade triangular, $d(x, y_N) < \varepsilon$, $d(x, T^{k_N}y_N) < \varepsilon$ e $|S_{k_N}(f - \beta(f))(y_N)| < \varepsilon$. Portanto, $x \in \Omega(f)$.

Provamos agora que $T(\Omega(f)) \subseteq \Omega(f)$. Sejam $x \in \Omega(f)$ e $\varepsilon > 0$. Utilizando a continuidade de T e do potencial f , podemos escolher $0 < \delta < \varepsilon/2$, satisfazendo as duas condições a seguir

$$T(B_\delta(x)) \subseteq B_\varepsilon(Tx) \quad \text{e} \quad f \text{ possui oscilação limitada por } \frac{\varepsilon}{2} \text{ em } B_\delta(x).$$

Ao tomar $y \in X$ e $n \in \mathbb{N}$ que satisfazem a definição de $\Omega(f)$ para o ponto x , com constante δ , encontramos

$$d(x, y) < \delta \Rightarrow d(Tx, Ty) < \varepsilon, \quad d(x, T^n y) < \delta \Rightarrow d(Tx, T^{n+1}y) < \varepsilon \quad \text{e}$$

$$|S_n(f - \beta(f))(Ty)| \leq |S_n(f - \beta(f))(y)| + |f \circ T^n(y) - f(y)| < \delta + \frac{\varepsilon}{2} < \varepsilon,$$

mostrando que $Tx \in \Omega(f)$.

Vejamos então a inclusão $\Omega(f) \subseteq T(\Omega(f))$, com a hipótese adicional que T é uma aplicação aberta. Dado $x \in \Omega(f)$, considere seqüências $\{y_n\} \subseteq X$ e $\{k_n\} \subseteq \mathbb{N}$ tais que k_n é crescente e y_n satisfaz a definição de $\Omega(f)$ para x , com $\varepsilon = 1/n$ e iterada k_n . Tal escolha é possível devido à definição de $\Omega(f)$. Note que ambas as seqüências $\{y_n\}$ e $\{T^{k_n}y_n\}$ convergem para x . Então, a menos de considerar uma subsequência, podemos assumir que $\{T^{k_n-1}y_n\}$ converge para um ponto z tal que $T(z) = x$. Precisamos mostrar que $z \in \Omega(f)$. Dado $\varepsilon > 0$, podemos escolher um aberto U contendo z , de diâmetro menor que ε , sobre o qual f tem oscilação limitada por $\varepsilon/2$. Como T é uma aplicação aberta, $V := T(U)$ é

um aberto contendo x . Fixamos n suficientemente grande tal que $1/n < \varepsilon/2$, $y_n \in V$ e $T^{k_n-1}y_n \in U$. Então, ao considerar $w_n \in T^{-1}y_n \cap U$, temos

$$d(z, w_n) < \varepsilon \quad \text{e} \quad d(z, T^{k_n}w_n) = d(z, T^{k_n-1}y_n) < \varepsilon.$$

Note ainda que, como f tem oscilação limitada por $\varepsilon/2$ em U , ocorre

$$|S_{k_n}(f - \beta(f))(w_n)| = |S_{k_n}(f - \beta(f))(y_n) - (f(T^{k_n-1}y_n) - f(w_n))| < \frac{1}{n} + \frac{\varepsilon}{2} < \varepsilon,$$

donde w_n e k_n cumprem as condições para que z seja um ponto de Aubry.

Finalmente, note que, para $c \in \mathbb{R}$, temos $\beta(f + c) = \beta(f) + c$, donde

$$S_n(f + c - \beta(f + c))(x) = S_n(f - \beta(f))(x), \quad \forall n \in \mathbb{N} \text{ e } \forall x \in X.$$

Assim, a definição de um ponto x pertencer a $\Omega(f)$ é a mesma que a $\Omega(f + c)$. Por outro lado, dada função $h \in C(X)$, podemos escolher δ pequeno suficiente para que h possua oscilação limitada por $\varepsilon/2$ em $B_\delta(x)$, resultando que um ponto $y \in X$ que realiza a definição de $x \in \Omega(f)$, com constante $\min\{\delta, \varepsilon/2\}$, satisfaz

$$|S_n(f + h \circ T - h - \beta(f + h \circ T - h))(y)| = |S_n(f - \beta(f))(y) + h(T^n y) - h(y)| < \varepsilon,$$

garantindo que $x \in \Omega(f + h \circ T - h)$. □

Antes de apresentar ao leitor a relação entre os conjuntos $\Omega(f)$ e $\mathcal{M}_{max}(f)$, definimos um conceito fundamental da teoria, que são as subações para o potencial f .

Definição 1.3.3. *Sejam (X, T) um SDT e $f \in C(X)$. É denominada uma subação para o potencial f qualquer função $u \in C(X)$ que cumpra a desigualdade*

$$f + u \circ T - u \leq \beta(f).$$

Há dois tipos de subações que merecem destaque por suas propriedades: calibradas e separantes. Uma subação u para o potencial f é dita ser calibrada se para todo $x \in X$ ocorrer

$$u(x) = \min_{Ty=x} [u(y) - f(y) + \beta(f)].$$

Por outro lado, uma subação u é dita ser separante se verifica

$$(f + u \circ T - u - \beta(f))^{-1}(0) = \Omega(f).$$

Como vemos, subações são funções $u \in C(X)$ tais que os cobordos $u \circ T - u$, somados ao potencial f , são limitados pela constante $\beta(f)$. Recorde que, sutilmente, introduzimos o conceito de subação ao estudar o valor $\gamma(f) = \inf_{h \in C(X)} \sup_{x \in X} [f(x) + h \circ T(x) - h(x)]$. Note que toda função $h \in C(X)$ que satisfaz esse ínfimo é uma subação para f , pois para todo $x \in X$, $\beta(f) \geq f(x) + h \circ T(x) - h(x)$. Por outro lado, se u é uma subação para o potencial f , obrigatoriamente $\sup_{x \in X} [f(x) + u \circ T(x) - u(x)] = \beta(f)$, uma vez que, para toda probabilidade $\mu \in \mathcal{M}_{max}(f)$, ocorre

$$\beta(f) = \int (f + h \circ T - h) d\mu \quad \text{e} \quad \beta(f) \geq f(x) + u \circ T(x) - h(x).$$

A existência de uma subação $u \in C(X)$ para o potencial f claramente faz com que a função $g := f + u \circ T - u - \beta(f)$ seja não positiva. Mas, uma vez que $\mathcal{M}_{max}(f) = \mathcal{M}_{max}(g)$ (cobordos e constantes não alteram o conjunto de probabilidades maximizantes), então o suporte de qualquer probabilidade maximizante para g (e, portanto, também para f) deve estar contido em $g^{-1}(0) = (f + u \circ T - u - \beta(f))^{-1}(0)$, sendo que isso ocorre para qualquer subação u . Em outros termos, a existência de uma subação u garante que $\mu \in \mathcal{M}_T$ é maximizante para f se, e só se, o suporte de μ está contido em $(f + u \circ T - u - \beta(f))^{-1}(0)$. Em particular, se f possui subação separante, é imediato que este último conjunto é exatamente $\Omega(f)$.

Com essa breve introdução sobre subações, esperamos mostrar ao leitor a importância dessa ferramenta no estudo de otimização ergódica. Garantir a existência de uma subação é um tema central de muitos artigos recentes da literatura de otimização ergódica. Um exemplo prático de construções de subações calibradas e de separantes pode ser encontrado em [19], no qual o problema central é otimizar custo médio de percurso sobre grafos orientados.

Mostraremos, na próxima subseção, um resultado sobre recorrência de cociclos (teorema de Atkinson [1]), o qual nos ajudará a provar que toda probabilidade maximizante para o potencial f possui seu suporte contido em $\Omega(f)$ e *vice-versa*.

1.3.2 Cociclos

Nesta subseção apresentamos alguns resultados sobre recorrência de cociclos sobre um SDT (X, T) . Mostramos que a recorrência de um cociclo para uma probabilidade $\mu \in \mathcal{M}_T$ está associada a μ ser probabilidade maximizante para f . A partir desta propriedade, com o auxílio do conceito de subação, demonstramos que uma probabilidade $\mu \in \mathcal{M}_T$ é maximizante para o potencial $f \in C(X)$ se, e somente se, $\text{supp}(\mu) \subseteq \Omega(f)$.

Começamos definindo o que é um cociclo.

Definição 1.3.4. *Seja (X, T) um SDT. Dada função $f \in C(X)$, definimos o cociclo de T induzido por f como sendo a aplicação $a_f(n, x): \mathbb{N} \times X \rightarrow \mathbb{R}$ tal que*

$$\begin{aligned} a_f(0, x) &= 0 & \forall x \in X; \\ a_f(n, x) &= \sum_{i=0}^{n-1} f \circ T^i(x) & \text{se } n \neq 0, \forall x \in X. \end{aligned}$$

Como o leitor pode perceber, o cociclo $a_f(n, x)$ é a n -ésima soma de Birkhoff de f no ponto x . Há versão de cociclo $a_f: \mathbb{Z} \times X \rightarrow \mathbb{R}$ quando a aplicação $T: X \rightarrow X$ for inversível. Neste caso, para $n \geq 0$ a definição de $a_f(n, x)$ é a mesma que a apresentada acima e, para o caso $n < 0$, consideramos

$$a_f(n, x) = -a_f(-n, T^n x) \quad \forall x \in X.$$

É de fácil verificação que, para quaisquer $n, m \in \mathbb{N}$ e para todo $x \in X$, vale a igualdade

$$a_f(n + m, x) = a_f(m, x) + a_f(n, T^m x).$$

Com efeito, dados $m, n \in \mathbb{N}$ e $x \in X$, temos

$$\begin{aligned} a_f(n + m, x) &= \sum_{k=0}^{n+m-1} f \circ T^k(x) = \sum_{k=0}^{m-1} f \circ T^k(x) + \sum_{k=m}^{n+m-1} f \circ T^k(x) \\ &= \sum_{k=0}^{m-1} f \circ T^k(x) + \sum_{k=0}^{n-1} f \circ T^k(T^m(x)) = a_f(m, x) + a_f(n, T^m x). \end{aligned}$$

Cociclos possuem relação com passeios aleatórios em \mathbb{R} . Um espaço de probabilidade (X, \mathcal{B}, μ) e uma aplicação $T: X \rightarrow X$ modelam um passeio aleatório em \mathbb{R} no qual, a cada

$x \in X$, temos definido o passeio $a_f(n, x)$ com $n \in \mathbb{N}$. Para detalhes sobre esta relação, o leitor pode consultar, por exemplo, [1]. Estamos interessados em estudar propriedades da recorrência de um cociclo a_f , a qual está subordinada a uma probabilidade $\mu \in \mathcal{M}_T$. Definimos a seguir o que vem a ser a recorrência do cociclo a_f .

Definição 1.3.5. *Sejam (X, T) um SDT e $f \in C(X)$. Dada probabilidade $\mu \in \mathcal{M}_T$, o cociclo a_f é dito ser recorrente para μ caso, para todo conjunto $A \in \mathcal{B}$ com $\mu(A) > 0$ e para qualquer $\varepsilon > 0$, existir um inteiro $n > 0$ tal que*

$$\mu(A \cap T^n(A) \cap \{x \in X : |a_f(n, x)| < \varepsilon\}) > 0.$$

Caso contrário, dizemos que o cociclo a_f é transiente.

Repare que a definição de recorrência de um cociclo guarda semelhanças com a definição de conjunto de Aubry $\Omega(f)$. Na verdade, se $a_{f-\beta(f)}$ for cociclo recorrente com respeito a μ probabilidade maximizante para f , então todo ponto $x \in \text{supp}(\mu)$ é tal que $x \in \Omega(f)$. Isto é fácil de ser provado: se $x \in \text{supp}(\mu)$, então a bola de centro x e raio ε possui medida positiva, logo, pela definição de recorrência de $a_{f-\beta(f)}$, há $n > 0$ tal que $B_\varepsilon(x) \cap T^n(B_\varepsilon(x)) \cap \{z \in X : |S_n(f - \beta(f))(z)| < \varepsilon\} \neq \emptyset$, de modo que $x \in \Omega(f)$. A recorrência de cociclos está, portanto, intimamente ligada à caracterização de probabilidades maximizantes.

Apresentamos agora o teorema de Atkinson, conforme demonstrado em [20]. Este é um resultado bastante interessante sobre cociclos quando a probabilidade $\mu \in \mathcal{M}_T$ é ergódica.

Teorema 1.3.6 (Atkinson). *Sejam (X, T) um SDT, $f \in C(X)$ e $\mu \in \mathcal{M}_T$ uma probabilidade ergódica. Seja $A \in \mathcal{B}$ um boreliano tal que $\mu(A) > 0$. Denotamos por $\Xi(f, A)$ o conjunto*

$$\left\{ x \in A : \forall \varepsilon > 0, \exists n \geq 1 \text{ com } T^n x \in A \text{ e } \left| a_f(n, x) - n \int f d\mu \right| < \varepsilon \right\}.$$

Então, $\Xi(f, A)$ e A diferem por um conjunto de medida nula para μ .

Demonstração. Suponha primeiramente que $\int f d\mu = 0$ e considere $\varepsilon > 0$. Definimos

$$\Xi_\varepsilon(f, A) := \{x \in A : \exists n \geq 1 \text{ com } T^n x \in A \text{ e } |a_f(n, x)| < \varepsilon\}.$$

Veja que $\Xi(f, A) = \bigcap_{k \in \mathbb{N}} \Xi_{k^{-1}}(f, A)$, pois dado $\varepsilon > 0$, há $k \in \mathbb{N}$ tal que $k^{-1} < \varepsilon$. Então, é

suficiente mostrar que $\mu(\Xi_\varepsilon(f, A)) = \mu(A)$ para um ε arbitrário. Suponha, por contradição, que o conjunto $\tilde{C} := A \setminus \Xi_\varepsilon(f, A)$ tem medida positiva. Considere $C \subseteq \tilde{C}$, o conjunto dos pontos recorrentes de \tilde{C} , isto é, o conjunto dos pontos de \tilde{C} que retornam a \tilde{C} para infinitas iteradas de T . Pelo teorema de recorrência de Poincaré, $\mu(C) = \mu(\tilde{C})$. Consideramos, então, $\rho: C \rightarrow \mathbb{N}_*$ a aplicação de primeiro retorno para C (para cada $x \in C$, $\rho(x)$ denota o menor inteiro positivo i tal que $T^i x \in C$). Definimos ainda a aplicação $T_C: C \rightarrow C$, denominada mapa de primeiro retorno, ou seja, $T_C(x) = T^{\rho(x)}(x)$, $\forall x \in C$, e a função $f_C(x) = a_f(\rho(x), x)$. Consideramos enfim, para todo $x \in C$, o cociclo de T_C induzido por f_C dado por $a_{f_C}(n, x) = \sum_{k=0}^{n-1} f_C \circ T_C^k(x)$. Fixe $x \in C$ para o qual os limites

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f \circ T^k(x) = \int f d\mu \quad e \quad \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_C \circ T^k(x) = \mu(C)$$

são satisfeitos. Note que é possível tomar tal x , pois ambos os limites acima são satisfeitos μ -q.t.p. em X e supomos $\mu(C) > 0$. Pela definição de C , temos que

$$\begin{aligned} |a_{f_C}(n, x) - a_{f_C}(m, x)| &= \left| \sum_{k=0}^{n-1} f_C \circ T_C^k(x) - \sum_{k=0}^{m-1} f_C \circ T_C^k(x) \right| \\ &= \left| \sum_{k=m}^{n-1} f_C \circ T_C^k(x) \right| = \left| \sum_{k=0}^{n-m-1} f_C \circ T_C^k(T_C^m(x)) \right| \\ &= |a_{f_C}(n-m, T_C^m(x))| \geq \varepsilon, \quad \forall n > m \geq 0. \end{aligned}$$

A desigualdade decorre de se mostrar que $a_{f_C}(n-m, T_C^m(x))$ equivale a $a_f(l, x)$, para algum $l \in \mathbb{N}_*$. Denotando $\rho_k = \sum_{i=0}^{k-1} \rho \circ T^i(x)$, para $k \geq 1$ e $\rho_0 = 0$, temos

$$\begin{aligned} |a_{f_C}(n-m, T_C^m(x))| &= \left| \sum_{k=0}^{n-m-1} f_C \circ T_C^k(x) \right| = \left| \sum_{k=0}^{n-m-1} f_C(T^{\rho_k}(x)) \right| \\ &= \left| \sum_{k=0}^{n-m-1} \sum_{i=0}^{\rho_{k+1}-1} f \circ T^{\rho_k+i}(x) \right| = \left| \sum_{k=0}^{\rho_{n-m}-1} f \circ T^k(x) \right| \\ &= |a_f(\rho_{n-m}, x)|. \end{aligned} \tag{1.6}$$

Tome então um inteiro $N \geq 1$. Subdividimos o intervalo $I_N = [-N\frac{\varepsilon}{2}, N\frac{\varepsilon}{2})$ em N intervalos de tamanho ε , como segue

$$\left[-N\frac{\varepsilon}{2} + i\varepsilon, -N\frac{\varepsilon}{2} + (i+1)\varepsilon \right), \quad \text{para } i = 0, \dots, N-1.$$

Cada um desses subintervalos contém no máximo um valor da forma $a_{f_C}(n, x)$. Como vimos acima, dois valores desse tipo estão sempre a uma distância maior ou igual a ε . Há, portanto, $k(N) \in \{0, \dots, N\}$, tal que $f_C(k(N), x) \notin I_N$ (tome $k(N)$ como sendo o menor dentre os índices $\{0, \dots, N\}$ que satisfaz essa propriedade). Assim,

$$|a_{f_C}(k(N), x)| \geq N \frac{\varepsilon}{2} \geq k(N) \frac{\varepsilon}{2}, \quad \forall N \geq 1.$$

Veja que a sequência $\{k(N)\}_{N \geq 1}$ não pode ser limitada, porque, para todo $N \geq 1$, temos $|a_{f_C}(k(N), x)| \geq N \frac{\varepsilon}{2}$, e caso $\{k(N)\}_{N \geq 1}$ fosse limitada, digamos, por $K \in \mathbb{N}$, ocorreria

$$N \frac{\varepsilon}{2} \leq |a_{f_C}(k(N), x)| \leq \max_{1 \leq i \leq K} |a_{f_C}(i, x)|,$$

o que não é satisfeito para N arbitrariamente grande. Logo,

$$\limsup_{n \rightarrow \infty} \frac{|a_{f_C}(n, x)|}{n} \geq \frac{\varepsilon}{2}.$$

Por outro lado, de (1.6) temos que

$$\frac{a_{f_C}(n, x)}{n} = \frac{a_f(\rho_n, x)}{n} = \frac{\sum_{l=0}^{\rho_n-1} f \circ T^l(x)}{\sum_{i=0}^{\rho_n-1} \chi_C \circ T^i(x)}.$$

Esta última igualdade é verificada, pois há exatamente $n - 1$ índices i entre 1 e $\rho_n - 1$ para os quais $T^i(x) \in C$. Assim, pela escolha de $x \in C$, segue que

$$\lim_{n \rightarrow \infty} \frac{|a_{f_C}(n, x)|}{n} = \lim_{n \rightarrow \infty} \frac{|\sum_{l=0}^{\rho_n-1} f \circ T^l(x)|}{\sum_{i=0}^{\rho_n-1} \chi_C \circ T^i(x)} = \frac{|\int f d\mu|}{\mu(C)}.$$

Mas, como vimos acima,

$$\limsup_{n \rightarrow \infty} \frac{|a_{f_C}(n, x)|}{n} \geq \frac{\varepsilon}{2} \quad \Rightarrow \quad \left| \int f d\mu \right| \geq \mu(C) \frac{\varepsilon}{2} > 0,$$

contradizendo a suposição que $\int f d\mu = 0$. Portanto, $\mu(C) = 0$, fornecendo assim a igualdade $\mu(A) = \mu(\Xi_\varepsilon(f, A))$.

No caso em que $\int f d\mu \neq 0$, considere $g = f - \int f d\mu$. Claramente, $\int g d\mu = 0$, podemos, então, aplicar o raciocínio acima. \square

Este resultado sobre cociclos revela algo interessante a respeito das somas de Birkhoff

$S_n f$. Lembre que o cociclo a_f está definido em $\mathbb{N} \times X$, sendo que $a_f(n, x) = S_n f(x)$. Quando μ é uma probabilidade ergódica, sabemos que μ -q.t.p. $x \in X$ existe o limite $\lim_{n \rightarrow \infty} \frac{1}{n} S_n f(x) = \int f d\mu$. O teorema nos revela que, sobre um conjunto de probabilidade total $\bar{X} \subseteq X$, para todo elemento $x \in \bar{X}$, desde que n seja suficientemente grande, a soma de Birkhoff $S_n f(x)$ é aproximadamente $n \int f d\mu$.

Agora vamos apresentar um segundo teorema sobre cociclos, também provado por Atkinson [1], o qual nos mostra que a recorrência de um cociclo com relação a uma probabilidade μ está diretamente ligada ao valor da integral $\int f d\mu$.

Teorema 1.3.7 (Atkinson). *Seja (X, T) um SDT no qual T é um isomorfismo mensurável. Considere uma função $f \in C(X)$ e uma probabilidade ergódica não atômica $\mu \in \mathcal{M}_T$. Então, a_f é recorrente para μ se, e somente se, $\int f d\mu = 0$.*

Demonstração. (\Rightarrow) Suponha primeiramente que $\int f d\mu = C > 0$. Como μ é ergódica, há conjunto $\bar{X} \subseteq X$, com $\mu(\bar{X}) = 1$, para o qual

$$\lim_{n \rightarrow \infty} \frac{a_f(n, x)}{n} = \int f d\mu = C > 0, \quad \forall x \in \bar{X}.$$

Fixe um ponto $x \in \bar{X}$. Como o limite acima existe para tal x , então, para $\varepsilon = \frac{C}{2}$, há $N_x \in \mathbb{N}$ tal que, $\forall n \geq N_x$, ocorre

$$\left| \frac{a_f(n, x)}{n} - C \right| < \varepsilon = \frac{C}{2}, \quad \text{ou seja,} \quad \frac{nC}{2} < a_f(n, x) < \frac{3nC}{2}.$$

Assim, para n suficientemente grande, temos $a_f(n, x) \geq 1$. No caso $C < 0$, podemos utilizar um raciocínio análogo e concluir que, para n suficientemente grande, temos $a_f(n, x) \leq -1$. Definimos os conjuntos

$$B_k = \{x \in \bar{X} : |a_f(n, x)| \geq 1, \forall n \geq k\}.$$

É evidente que $B_k \subseteq B_{k+1}$. Como para cada $x \in \bar{X}$ há $N_x \in \mathbb{N}$ tal que $x \in B_{N_x}$, segue que $\bar{X} = \cup_{k \geq 1} B_k$. Deste modo, pelo lema 1.1.1, para $\varepsilon = \frac{1}{4}$, há $N \in \mathbb{N}$ tal que $\mu(B_N) > \frac{3}{4}$. Denotamos tal B_N por B . Pelo lema 1.1.7, existe conjunto $E \in \mathcal{B}$ satisfazendo

$$\mu \left(\bigcup_{n=0}^{4N} T^n(E) \right) > \frac{3}{4},$$

com $T^i(E) \cap T^j(E) = \emptyset$, se $i \neq j$, para $0 \leq i, j \leq 4N$. Como μ é invariante para o isomorfismo T , então

$$\mu \left(\bigcup_{n=N}^{3N} T^n(E) \right) > \frac{1}{4},$$

obrigando ao menos um dos conjuntos $T^N(E), \dots, T^{3N}(E)$ a interceptar B em um subconjunto de medida positiva. Tomando $i \in \{N, \dots, 3N\}$ o menor índice com tal propriedade, denotamos essa interseção por A . Como $A \subseteq T^i(E)$, então

$$A \cap T^n(A) = \emptyset, \quad \text{para } n < N, \quad n \neq 0.$$

Por outro lado, como $A \subseteq B$, para cada $x \in A$ temos

$$|a_f(n, x)| \geq 1, \quad \text{sempre que } n \geq N.$$

Logo, se $\varepsilon < 1$, resulta que, para todo inteiro positivo n ,

$$\mu(A \cap T^n(A) \cap \{x \in X : |a_f(n, x)| < \varepsilon\}) = \mu(\emptyset) = 0.$$

Isto implica que a_f não pode ser recorrente.

(\Leftarrow) A outra implicação é produzida pelo teorema 1.3.6. Suponha que $\int f \, d\mu = 0$. Considere um conjunto $A \in \mathcal{B}$, com $\mu(A) > 0$. Então, dado $\varepsilon > 0$, definimos os conjuntos

$$\Xi_\varepsilon^{[m]}(f, A) := \{x \in A : \exists n \in \{1, \dots, m\}, \text{ com } T^n(x) \in A \text{ e } |a_f(n, x)| < \varepsilon\}$$

$$\text{e } \Xi_\varepsilon^m(f, A) := \{x \in A : T^m(x) \in A \text{ e } |a_f(m, x)| < \varepsilon\}.$$

Temos que $\Xi_\varepsilon(f, A) = \bigcup_{m \geq 1} \Xi_\varepsilon^{[m]}(f, A)$. Novamente usando o lema 1.1.1, existe $M \in \mathbb{N}$ tal que $\mu(\Xi_\varepsilon^{[M]}(f, A)) \geq \mu(\Xi_\varepsilon(f, A))/2$. Porém, note que

$$\Xi_\varepsilon^{[M]}(f, A) = \bigcup_{i=0}^M \Xi_\varepsilon^i(f, A).$$

Desse modo, há $j \in \{1, \dots, M\}$ com $\mu(\Xi_\varepsilon^j(f, A)) \geq \mu(\Xi_\varepsilon(f, A))/2M > 0$. O conjunto $\Xi_\varepsilon^j(f, A)$ pode ser escrito como $A \cap T^j(A) \cap \{x \in A : |a_f(j, x)| < \varepsilon\}$, concluindo assim que a_f satisfaz a condição de ser recorrente, pois A é tomado arbitrariamente. \square

O teorema de Atkinson será utilizado para demonstrar que toda probabilidade maximi-

zante para o potencial f possui suporte contido em $\Omega(f)$. Reciprocamente, a noção de subação nos permitirá mostrar que $\text{supp}(\mu) \subseteq \Omega(f)$ implica $\mu \in \mathcal{M}_{\max}(f)$.

Teorema 1.3.8. *Sejam (X, T) um SDT e $f \in C(X)$. Então,*

- (i) *toda probabilidade $\mu \in \mathcal{M}_{\max}$ possui o suporte contido em $\Omega(f)$;*
- (ii) *se existir subação $u \in C(X)$ associada ao potencial f , toda probabilidade $\mu \in \mathcal{M}_T$ suportada em $\Omega(f)$ é maximizante para f .*

Demonstração. (i) Sem perda de generalidade, considere uma probabilidade ergódica $\mu \in \mathcal{M}_{\max}(f)$. Sejam $x \in \text{supp}(\mu)$ e $\eta > 0$. Como, pela definição de suporte, $\mu(B_\eta(x)) > 0$, o teorema 1.3.6 garante que $\mu(\Xi(f, B_\eta(x))) > 0$, ou seja,

$$\mu\left(\left\{y \in B_\eta(x) : \forall \varepsilon > 0, \exists n \geq 1 \text{ com } T^n y \in B_\eta(x) \text{ e } \left|a_f(n, y) - n \int f \, d\mu\right| < \varepsilon\right\}\right) > 0.$$

Logo, como $\Xi(f, B_\eta(x)) \neq \emptyset$, ao tomar $\varepsilon = \eta$, qualquer $y \in \Xi(f, B_\varepsilon(x))$ possui iterada $n > 0$ tal que a condição necessária para x pertencer a $\Omega(f)$ é satisfeita.

(ii) Suponha que $u \in C(X)$ é uma subação. Recorde observação feita após a definição de subação, a qual garante que uma probabilidade $\mu \in \mathcal{M}_T$ é maximizante se, e somente se, $\text{supp}(\mu) \subseteq (f + u \circ T - u - \beta(f))^{-1}(0)$. Basta então mostrar que $\Omega(f)$ está contido neste conjunto. Sejam $x \in \Omega(f)$ e $\varepsilon > 0$. Pelo lema 1.3.2, $x \in \Omega(f + h \circ T - h - \beta(f))$. Portanto, há $y_\varepsilon \in X$ e $n_\varepsilon \in \mathbb{N}$ tais que,

$$d(x, y_\varepsilon) < \varepsilon, \quad d(x, T^{n_\varepsilon} y_\varepsilon) < \varepsilon \quad \text{e} \quad |S_{n_\varepsilon}(f + u \circ T - u - \beta(f))(y_\varepsilon)| < \varepsilon.$$

Lembramos que $f + u \circ T - u - \beta(f)$ é uma função não positiva. Assim, certamente,

$$0 \geq S_{n_\varepsilon}(f + u \circ T - u - \beta(f))(y_\varepsilon) > -\varepsilon.$$

Em particular, como todas as parcelas somadas são não positivas, temos

$$0 \geq (f + u \circ T - u - \beta(f))(y_\varepsilon) > -\varepsilon.$$

Então, ao fazer $\varepsilon \rightarrow 0$, como $y_\varepsilon \rightarrow x$, a continuidade das funções f e u garante que

$$(f + u \circ T - u - \beta(f))(x) = 0,$$

concluindo a demonstração. □

No terceiro capítulo, veremos, em detalhes, como obter subações para potenciais holderianos no contexto de difeomorfismos de Anosov.

Formalismo Termodinâmico em Temperatura Zero e o Operador de Transferência

Caro leitor, neste capítulo abordamos o problema de maximização da integral de um potencial $f \in C(X)$ através de propriedades do formalismo termodinâmico. Fazemos uma breve revisão dos conceitos de entropia topológica e de entropia métrica associados à dinâmica $T: X \rightarrow X$, introduzimos ainda a noção de pressão topológica e recordamos o chamado princípio variacional. Com essas ferramentas, utilizamos um processo denominado congelamento do sistema para obter o valor maximal da integral de f e informações sobre as probabilidades maximizantes. Ademais, mostramos propriedades do operador de transferência, as quais fornecem condições para se estudar limites de estados de equilíbrio obtidos no processo de congelamento do sistema. Os resultados aqui abordados são modelados sobre uma dinâmica expansiva, tal como definimos adiante, porém podem ser estendidos para outros tipos de dinâmica por meio de semi-conjugação, o que permite que o formalismo termodinâmico seja, como é bem conhecido na literatura, objeto de estudo para difeomorfismos de Anosov.

2.1 Formalismo termodinâmico e pressão

Como mencionado, nesta seção abordamos clássico conteúdo da teoria ergódica, mais precisamente, o tópico referente à entropia topológica e também à entropia métrica, este

segundo item sendo muito usual também em teoria da informação. Mostramos que o conceito de entropia topológica pode ser generalizado para o de pressão topológica, o qual possui estreita relação com as probabilidades maximizantes que estamos interessados em estudar. Para esta seção baseamo-nos nas obras [8], [26] e [36].

Fixamos primeiramente o SDT (X, T) . Recorde que X é um espaço métrico compacto e $T: X \rightarrow X$ é uma aplicação contínua. Seja α uma cobertura de X por abertos. Para cada inteiro $n > 0$, podemos definir o refinamento dinâmico de α , como segue:

$$\alpha^n := \{A_0 \cap T^{-1}(A_1) \cap \dots \cap T^{-n+1}(A_{n-1}) : A_i \in \alpha, i = 0, \dots, n-1\}.$$

Uma outra forma de denotar esse refinamento é $\alpha^n = \bigvee_{k=0}^{n-1} T^{-k}(\alpha)$.

Como nosso espaço métrico X é compacto, toda cobertura de X por abertos admite uma subcobertura de cardinalidade finita. Então, se γ é cobertura de X , $N(\gamma)$ denota a menor cardinalidade entre subcoberturas de γ . Evidentemente, $N(\gamma) < \infty$ para qualquer cobertura de abertos γ . Com isso, já podemos definir o que vem a ser a entropia topológica.

Definição 2.1.1. *Seja α uma cobertura de X por abertos e considere α^n seu refinamento dinâmico. Definimos a entropia de T com relação à cobertura α como*

$$h(T, \alpha) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log N(\alpha^n).$$

Então, a entropia topológica de T é definida como o supremo de $h(T, \alpha)$ com respeito às coberturas α

$$h(T) := \sup_{\alpha} h(T, \alpha).$$

Como o leitor pode perceber, a definição da entropia topológica $h(T)$ depende apenas das coberturas α finitas de X , uma vez que o cálculo é feito a partir da cardinalidade mínima de uma subcobertura de α , que é sempre um número finito.

Antes de apresentar algumas propriedades, mostramos uma forma equivalente de se definir a entropia topológica, que é via conjuntos (n, ε) -geradores e conjuntos (n, ε) -separados.

Definição 2.1.2. *Um conjunto $G \subseteq X$ é dito (n, ε) -gerador para T se, para qualquer $y \in X$, há $x \in G$ tal que*

$$d(T^k(x), T^k(y)) < \varepsilon, \quad \forall k = 0, \dots, n-1.$$

Por outro lado, um conjunto $F \subseteq X$ é dito (n, ε) -separado para T se para quaisquer $x, y \in F$ ocorrer

$$d(T^k(x), T^k(y)) \geq \varepsilon, \quad \text{para algum } k \in \{0, \dots, n-1\}.$$

A nomenclatura desses conjuntos é bastante intuitiva, dadas suas definições. Denotamos por $r_n(T, \varepsilon)$ a menor cardinalidade de um conjunto (n, ε) -gerador para T e $s_n(T, \varepsilon)$ a maior cardinalidade de um conjunto (n, ε) -separado para T . Vamos, então, considerar os seguintes valores

$$r(T, \varepsilon) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log r_n(T, \varepsilon) \quad \text{e}$$

$$s(T, \varepsilon) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log s_n(T, \varepsilon).$$

É fácil perceber que $r(T, \varepsilon)$ e $s(T, \varepsilon)$, enquanto funções de ε , são positivas e não crescentes. Com efeito, se $0 < \varepsilon_1 < \varepsilon_2$, todo conjunto (n, ε_1) -gerador também é (n, ε_2) -gerador, e, como buscamos a cardinalidade mínima, $r(T, \varepsilon_1) \geq r(T, \varepsilon_2)$. De mesmo modo, os conjuntos (n, ε_2) -separados também são conjuntos (n, ε_1) -separados, donde $s(T, \varepsilon_1) \geq s(T, \varepsilon_2)$, pois aqui queremos a cardinalidade máxima de conjuntos (n, ε) -separados. Logo, ficam bem definidos os limites quando $\varepsilon \rightarrow 0$ (ainda que eventualmente sejam $+\infty$):

$$r(T) := \lim_{\varepsilon \rightarrow 0} r(T, \varepsilon) \quad \text{e} \quad s(T) = \lim_{\varepsilon \rightarrow 0} s(T, \varepsilon).$$

Os lemas que apresentamos a seguir são clássicos na literatura de teoria ergódica e relacionam os valores $r(T)$ e $s(T)$ com a entropia topológica $h(T)$.

Lema 2.1.3. *Seja (X, T) um SDT. Então, para todo $n \geq 1$ e $\varepsilon > 0$, ocorrem as desigualdades*

$$r_n(T, \varepsilon) \leq s_n(T, \varepsilon) \leq r_n(T, \varepsilon/2).$$

Sendo assim, é claro que $r(T) = s(T)$.

Demonstração. A primeira desigualdade é bastante simples de se provar. Considere um conjunto (n, ε) -separado F , com $\#F = s(n, \varepsilon)$, ou seja, de cardinalidade máxima. Então, para qualquer $x \in X \setminus F$, o conjunto $\{x\} \cup F$ não é (n, ε) -separado, pois F é de cardinalidade máxima. Mas isto quer dizer que há $y \in F$ tal que $d(T^i x, T^i y) < \varepsilon$, para $i = 0, \dots, n-1$, ou seja, o conjunto F é (n, ε) -gerador, não necessariamente de cardinalidade mínima, resultando na primeira desigualdade.

Para provar a outra desigualdade, considere F conjunto (n, ε) -separado e G conjunto $(n, \varepsilon/2)$ -gerador. Definimos uma aplicação $\xi: F \rightarrow G$, que para cada $x \in F$ associa $y \in G$ tal que $d(T^i x, T^i y) < \varepsilon/2$, para todo $i = 0, \dots, n-1$. A aplicação ξ é necessariamente injetiva, pois, caso contrário, se $\xi(x) = \xi(z) = y$, da desigualdade triangular resultaria que $d(T^i x, T^i z) < \varepsilon$, para todo $i = 0, \dots, n-1$, o que contradiz a definição de F ser conjunto (n, ε) -separado. Portanto $\#F \leq \#G$. Como G e F são arbitrários, $s(T, \varepsilon) \leq r(T, \varepsilon/2)$, provando a segunda desigualdade. \square

Lema 2.1.4. *Seja (X, T) um SDT. Então, a entropia topológica de T coincide com os limites $r(T)$ e $s(T)$.*

Demonstração. Dados $\varepsilon > 0$ e $n \in \mathbb{N}$, sejam F um conjunto (n, ε) -separado de cardinalidade maximal ($\#F = s_n(T, \varepsilon)$) e α uma cobertura finita de X , com diâmetro menor que ε . Observe que cada elemento de α^n contém no máximo um ponto de F , pois, dados $x, y \in F$ verificando

$$x, y \in A_{i_0} \cap T^{-1}(A_{i_1}) \cap \dots \cap T^{-(n-1)}(A_{i_{n-1}}),$$

para $A_{i_0}, \dots, A_{i_{n-1}} \in \alpha$, então obrigatoriamente $d(T^i x, T^i y) < \varepsilon$, para $i = 0, \dots, n-1$, o que é proibido por F ser conjunto (n, ε) -separado. Mas, então, $s_n(T, \varepsilon) \leq N(\alpha^n)$, donde, tomando limites, temos $s(T) \leq h(T)$.

Agora, sejam α uma cobertura finita de X com número de Lebesgue $\varepsilon > 0$ (toda cobertura finita admite número de Lebesgue) e G um conjunto (n, ε) -gerador de cardinalidade minimal ($\#G = r_n(T, \varepsilon)$). A cada $x \in G$, associamos conjuntos $A_{x,i} \in \alpha$ tais que $B_\varepsilon(T^i x) \subseteq A_{x,i}$. Então,

$$B_\varepsilon(x, n) := \{y \in X : d(T^i x, T^i y) < \varepsilon, i = 0, \dots, n-1\} \subseteq \bigcap_{i=0}^{n-1} T^{-i}(A_{x,i}).$$

Como G é conjunto (n, ε) -gerador, todo $y \in X$ pertence a $B_\varepsilon(x, n)$, para algum $x \in G$. Logo, $\tilde{\alpha} := \{\bigcap_{i=0}^{n-1} T^{-i}(A_{x,i}) : x \in G\} \subseteq \alpha^n$ é também uma cobertura de X . Sendo assim, $N(\tilde{\alpha}) \leq r_n(T, \varepsilon)$, donde segue que $h(T) \leq r(T)$.

Pelo lema anterior, vimos que $s(T) = r(T)$ e, portanto, segue a tríplice igualdade. \square

Estas três formas de caracterizar a entropia topológica da aplicação T são canônicas na literatura (vide, por exemplo, referências já citadas no início da seção), sendo que, no

processo do cálculo numérico da entropia, a abordagem por conjuntos (n, ε) -geradores e (n, ε) -separados pode ser mais palpável.

Temos algumas propriedades básicas a serem destacadas sobre a entropia topológica.

Proposição 2.1.5. *Sejam X , X_1 e X_2 espaços métricos compactos, $T: X \rightarrow X$ e $T_i: X_i \rightarrow X_i$ aplicações contínuas. Temos as seguintes propriedades da entropia topológica:*

- (i) *se T_2 é semiconjugada a T_1 , isto é, há aplicação $\varphi: X_1 \rightarrow X_2$ tal que $\varphi \circ T_1 = T_2 \circ \varphi$, então, $h(T_2) \leq h(T_1)$;*
- (ii) *caso T_1 e T_2 sejam topologicamente equivalentes ($T_1 = \varphi \circ T_2 \circ \varphi^{-1}$), temos a igualdade $h(T_1) = h(T_2)$;*
- (iii) *se $(\beta_k)_{k \in \mathbb{N}}$ é uma sequência de coberturas de X tal que $\text{diam}(\beta_k) \rightarrow 0$ quando $k \rightarrow \infty$, resulta,*

$$h(T) = \lim_{k \rightarrow \infty} h(T, \beta_k) = \sup_{k \in \mathbb{N}} h(T, \beta_k);$$

- (iv) *para uma cobertura α de X tal que seu refinamento dinâmico por T possui diâmetro convergindo a zero, temos*

$$h(T) = h(T, \alpha) = \lim_{n \rightarrow \infty} \frac{1}{n} \log N(\alpha^n);$$

- (v) *para todo $k \in \mathbb{N}$ vale a igualdade $h(T^k) = kh(T)$;*
- (vi) *caso T seja um homeomorfismo, vale que $h(T^{-1}) = h(T)$ e $h(T^k) = |k|h(T)$, $\forall k \in \mathbb{Z}$.*

O leitor pode consultar a demonstração dessas propriedades da entropia topológica, por exemplo, em [26]. O intuito da dissertação neste tópico é esboçar alguns conceitos da teoria para que o leitor saiba o básico a respeito das ferramentas utilizadas, e que fique instigado a pesquisar mais sobre o assunto. Contudo, a propriedade de conjugação é muito importante, pois, se estudamos a dinâmica do SDT (X, T) e a aplicação $T: X \rightarrow X$ é conjugada a uma aplicação $S: Y \rightarrow Y$, a qual sabemos calcular a entropia, então obtemos também a entropia da aplicação T .

Paralelamente ao conceito de entropia topológica, podemos definir a entropia métrica de uma probabilidade $\mu \in \mathcal{M}_T$. Note que, para o cálculo da entropia métrica, usamos partições de X em vez de coberturas.

Definição 2.1.6. *Seja (X, T) um SDT. Dada uma partição \mathcal{P} de X , definimos sua entropia com respeito a $\mu \in \mathcal{M}_T$ por*

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

Assim como definimos o refinamento de coberturas, podemos definir o refinamento de partições. Dadas duas partições \mathcal{P} e \mathcal{Q} de X , temos o refinamento $\mathcal{P} \vee \mathcal{Q}$ dado por

$$\mathcal{P} \vee \mathcal{Q} := \{P \cap Q \mid P \in \mathcal{P} \text{ e } Q \in \mathcal{Q}\}.$$

Também denotamos por \mathcal{P}^n o refinamento dinâmico $\bigvee_{k=0}^{n-1} T^{-k}(\mathcal{P})$. Apontamos, então, a seguinte propriedade sobre a entropia do refinamento de partições.

Lema 2.1.7. *Sejam \mathcal{P} e \mathcal{Q} duas partições de X . Dada $\mu \in \mathcal{M}_T$, temos que*

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

Demonstração. A prova do lema é bastante simples, basta usar que $\mu(P \cap Q) = \mu(P)\mu(Q)$ e aplicar a essa igualdade a propriedade de logaritmo ($\log(a \times b) = \log(a) + \log(b)$). Após isto, resta fazer o somatório sobre os elementos de \mathcal{P} e de \mathcal{Q} . \square

Considere então \mathcal{P} uma partição de X . Dada $\mu \in \mathcal{M}_T$, a propriedade acima garante que a sequência $a_n = H_\mu(\mathcal{P}^n)$ é subaditiva, isto é, $a_{n+m} \leq a_n + a_m$. Sendo assim, pelo lema subaditivo, é bem definido o limite

$$\lim_{n \rightarrow \infty} \frac{1}{n} H_\mu(\mathcal{P}^n) = \inf_{n \in \mathbb{N}} \frac{1}{n} H_\mu(\mathcal{P}^n).$$

Definição 2.1.8. *A entropia de T com respeito a uma probabilidade μ e à partição \mathcal{P} será o limite a seguir:*

$$h_\mu(T, \mathcal{P}) := \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu(\mathcal{P}^n).$$

Então, a entropia métrica de T com respeito à μ é dada pelo supremo

$$h_\mu(T) := \sup_{\mathcal{P}} h_\mu(T, \mathcal{P}).$$

O principal teorema sobre as entropias métricas é o princípio variacional. Para esse resultado, vamos assumir uma hipótese adicional sobre a aplicação T , que é a hipótese de

expansividade.

Definição 2.1.9. *Seja (X, T) um SDT. Dizemos que a aplicação T é expansiva se há constante $\varepsilon_0 > 0$ tal que, para quaisquer $x, y \in X$,*

$$d(T^n x, T^n y) < \varepsilon_0, \quad \forall n \in \mathbb{N} \quad \Rightarrow \quad x = y.$$

A constante ε_0 é denominada constante de expansividade de T .

Enunciamos, a seguir, o princípio variacional das entropias.

Teorema 2.1.10. *Seja (X, T) um SDT no qual T é uma aplicação expansiva. Então,*

$$h(T) = \sup_{\mu \in \mathcal{M}_T} \{h_\mu(T)\}.$$

Para provar esse teorema, são necessários vários lemas técnicos. Enunciamos, a seguir, alguns dos lemas necessários para demonstrar o princípio variacional. Para as demonstrações dos lemas enunciados a seguir, sugerimos que o leitor consulte, por exemplo, [8]. Primeiramente definimos a entropia condicional de duas partições com relação a uma probabilidade μ .

Definição 2.1.11. *Sejam \mathcal{P} e \mathcal{Q} partições de X e $\mu \in \mathcal{M}_T$. Definimos a entropia condicional*

$$H_\mu(\mathcal{P}|\mathcal{Q}) := H_\mu(\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{Q}).$$

Como mencionado, enunciamos a seguir algumas propriedades e lemas a respeito da entropia métrica.

Proposição 2.1.12. *Sejam $\mathcal{P}, \mathcal{Q}, \mathcal{R}$ partições finitas de X e $\mu \in \mathcal{M}_T$. Então,*

- $H_\mu(\mathcal{P}|\mathcal{Q}) \leq H_\mu(\mathcal{P}|\mathcal{R})$, se $\mathcal{R} \subseteq \mathcal{Q}$;
- $H_\mu(\mathcal{P}|\mathcal{Q}) = 0$, se $\mathcal{P} \subseteq \mathcal{Q}$;
- $H_\mu(\mathcal{P} \vee \mathcal{Q}|\mathcal{R}) \leq H_\mu(\mathcal{P}|\mathcal{R}) + H_\mu(\mathcal{Q}|\mathcal{R})$;
- $H_\mu(\mathcal{P}) \leq H_\mu(\mathcal{Q}) + H_\mu(\mathcal{P}|\mathcal{Q})$;
- $H_\mu(T^{-k}(\mathcal{P})|T^{-k}(\mathcal{Q})) = H_\mu(\mathcal{P}|\mathcal{Q})$, para todo $k \geq 0$;

- $h_\mu(T, \mathcal{P}) \leq h_\mu(T, \mathcal{Q}) + H_\mu(\mathcal{P}|\mathcal{Q});$
- $h_\mu(T, \bigvee_{k=0}^{n-1} T^{-k}(\mathcal{P})) = h_\mu(T, \mathcal{P}).$

Lema 2.1.13. *Sejam (X, T) um SDT e $\mu \in \mathcal{M}_T$. Dados $\varepsilon > 0$ e \mathcal{P} uma partição boreliana finita de X , há $\delta > 0$ tal que $H_\mu(\mathcal{P}|\mathcal{Q}) < \varepsilon$, sempre que \mathcal{Q} é uma partição boreliana finita de X com $\text{diam}(\mathcal{Q}) < \delta$.*

Lema 2.1.14. *Sejam $\{\mathcal{P}_n\}$ partições borelianas finitas de X tais que $\text{diam}(\mathcal{P}_n) \rightarrow 0$, então,*

$$h_\mu(T) = \lim_{n \rightarrow \infty} h_\mu(T, \mathcal{P}_n).$$

Caso a aplicação T seja expansiva com constante de expansividade ε_0 , $h_\mu(T) = h_\mu(T, \mathcal{P})$, sempre que \mathcal{P} for partição boreliana finita de X com $\text{diam}(\mathcal{P}) < \varepsilon_0$. Além disso, para todo $n > 0$,

$$h_\mu(T^n) = nh_\mu(T).$$

Fazemos, adiante, um breve esboço da demonstração do princípio variacional.

Ideias da prova do princípio variacional da pressão: Dados $\varepsilon > 0$ e $n \in \mathbb{N}$, consideramos um conjunto (n, ε) -separado F_n , de cardinalidade $s_n(T, \varepsilon)$. Definimos as probabilidades $\mu_n := \frac{1}{s_n(T, \varepsilon)} \sum_{x \in F_n} \delta_x$ e $\nu_n = \frac{1}{n} \sum_{i=0}^{n-1} \mu_n \circ T^{-i}$. Pelo teorema de Krylov-Bogoliubov, todo ponto de aderência da sequência ν_n é uma probabilidade invariante por T . Mostra-se, então, que há probabilidade ν , a qual é ponto de acumulação da sequência ν_n , e verifica-se $h_\nu(T) \geq \limsup_{n \rightarrow \infty} \frac{1}{n} \log(s_n(T, \varepsilon))$. Com isso, resulta que $h_\nu(T) \geq h(T)$. Em particular, a probabilidade ν construída é de entropia maximal.

A outra desigualdade é obtida ao se estimar que, para toda $\mu \in \mathcal{M}_T$, temos a desigualdade $h_\mu(T) \leq h(T)$. Para isso, dada $\mu \in \mathcal{M}_T$ e $\eta > 0$, escolhemos partição $\mathcal{P} = \{P_1, \dots, P_k\}$ de X cumprindo $h_\mu(T, \mathcal{P}) \geq h_\mu(T) - \eta$. Consideramos, então, valor $\varepsilon \in (0, \frac{1}{k} \log k)$ e compactos $B_i \subseteq P_i$ tais que $\mu(P_i \setminus B_i) < \varepsilon$. Além disso, definimos $B_0 := X \setminus \bigcup_{i=1}^k B_i$. Assim, utilizando a concavidade da função $\varphi(x) = -x \log(x)$, segue que

$$H_\mu(\mathcal{P}|\mathcal{B}) = \mu(B_0) \sum_{i=1}^k \frac{\mu(P_i \cap B_0)}{\mu(B_0)} \log \left(\frac{\mu(P_i \cap B_0)}{\mu(B_0)} \right) \leq \mu(B_0) \log(k) \leq \varepsilon k \log(k) \leq 1.$$

Logo,

$$h_\mu(T, \mathcal{P}) \leq h_\mu(T, \mathcal{B}) + H_\mu(\mathcal{P}|\mathcal{B}) \leq h_\mu(T, \mathcal{B}) + 1.$$

Ao considerar a cobertura $\mathcal{U} := \{U_1 = B_0 \cup B_1, \dots, U_k = B_0 \cup B_k\}$, podemos concluir que

$$\# \left(\bigvee_{i=0}^{k-1} T^{-i}(\mathcal{B}) \right) \leq 2^n N \left(\bigvee_{i=0}^{k-1} T^{-i}(\mathcal{U}) \right),$$

donde obtemos

$$h_\mu(T, \mathcal{B}) \leq \log 2 + H(T, \mathcal{U}).$$

Sendo assim,

$$h_\mu(T) - \eta \leq 1 + \log 2 + H(T, \mathcal{U}) \leq 1 + \log 2h(T).$$

Pela arbitrariedade de η , temos $h_\mu(T) \leq 1 + \log 2 + h(T)$. Porém, utilizando as igualdades $h_\mu(T^m) = mh_\mu(T)$ e $h(T^m) = mh(T)$, e repetindo esse processo para T^m no lugar de T , segue que

$$h_\mu(T) \leq \frac{1 + \log 2}{m} + h(T), \quad \forall m \in \mathbb{N},$$

donde $h_\mu(T) \leq h(T)$.

□

Para mais detalhes sobre a demonstração, o leitor pode consultar, por exemplo, [8].

Como mencionado, o conceito de entropia topológica pode ser generalizado pela pressão topológica, ferramenta esta que permite estender o princípio variacional. Definimos aqui a pressão topológica apenas por conjuntos (n, ε) -separados e conjuntos (n, ε) -geradores, porém, assim como a entropia topológica, há equivalente definição via coberturas abertas.

Definição 2.1.15. *Seja (X, T) um SDT. Dada $f \in C^0(X)$, para cada $n \in \mathbb{N}$ e $\varepsilon > 0$, definimos o valor*

$$r_T(f, n, \varepsilon) := \inf \left\{ \sum_{x \in G} e^{S_n f(x)} \mid G \text{ conjunto } (n, \varepsilon)\text{-gerador} \right\}.$$

Por um processo análogo à definição da entropia topológica, consideramos

$$r_T(f, \varepsilon) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log r_T(f, n, \varepsilon),$$

e, por fim, definimos a pressão topológica de f pelo limite

$$P_T(f) := \lim_{\varepsilon \rightarrow 0} r_T(f, \varepsilon).$$

Esta é a definição da pressão topológica via conjuntos (n, ε) -geradores. Podemos também defini-la via conjuntos (n, ε) -separados, escrevendo

$$s_T(f, n, \varepsilon) := \sup \left\{ \sum_{x \in F} e^{S_n f(x)} \mid F \text{ conjunto } (n, \varepsilon)\text{-separado} \right\},$$

e tomando os limites

$$s_T(f, \varepsilon) := \limsup_{n \rightarrow \infty} \frac{1}{n} \log s_T(f, n, \varepsilon) \quad \text{e}$$

$$P_T(f) = \lim_{\varepsilon \rightarrow 0} s_T(f, \varepsilon).$$

A validade desta equivalência é garantida pela proposição a seguir, similar a uma demonstrada no cálculo da entropia topológica.

Proposição 2.1.16. *Se $|f(x) - f(y)| < \eta$ sempre que $d(x, y) < \varepsilon/2$, então,*

$$r_T(f, n, \varepsilon) \leq s_T(f, n, \varepsilon) \leq e^{m\eta} r_T(f, n, \varepsilon/2).$$

Repare que, ao aplicar a pressão topológica ao potencial nulo, obtemos exatamente a entropia topológica de T :

$$\begin{aligned} P_T(0) &= \lim_{\varepsilon \rightarrow 0} r_T(0, \varepsilon) \\ &= \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log (\inf \{ \#G \mid G \text{ é conjunto } (n, \varepsilon)\text{-gerador} \}) \\ &= \lim_{\varepsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{1}{n} \log (r_n(T, \varepsilon)) = h(T). \end{aligned}$$

Apontamos, a seguir, algumas propriedades básicas da pressão topológica. Tais propriedades são de fácil convencimento do leitor, sendo que uma demonstração pode ser consultada, por exemplo, em [26].

Proposição 2.1.17. *Sejam (X, T) um SDT, $c \in \mathbb{R}$ e $f, g \in C(X)$. Temos que*

- P_T comuta com constantes, isto é, $P_T(f + c) = P_T(f) + c$;
- se $f \leq g$, então, $P_T(f) \leq P_T(g)$;
- para qualquer $m \geq 1$, vale que $P_{T^m}(S_m f) = mP_T(f)$;
- caso T seja um homeomorfismo, temos $P_{T^{-1}}(f) = P_T(f)$;

- se f e g são potenciais coomólogos, resulta que $P_T(f) = P_T(g)$;
- o funcional $P_T(\cdot)$ é lipschitziano, ou seja, $P_T(f - g) \leq \|f - g\|_\infty$;
- o funcional $P_T(\cdot)$ é convexo: para todo $t \in (0, 1)$, vale que

$$P_T(tf + (1 - t)g) \leq tP_T(f) + (1 - t)P_T(g).$$

Enunciamos agora o princípio variacional da pressão, muitas vezes chamado somente de princípio variacional, uma vez que generaliza o princípio variacional das entropias.

Teorema 2.1.18 (Princípio variacional da pressão). *Sejam (X, T) um SDT e $f \in C(X)$. Então,*

$$P_T(f) = \sup_{\mu \in \mathcal{M}_T} \left\{ h_\mu(T) + \int f \, d\mu \right\}.$$

Para a demonstração desse teorema, indicamos novamente [26]. É evidente que o princípio variacional das entropias é obtido como corolário do princípio variacional da pressão. Para isto, basta aplicar o princípio variacional da pressão ao potencial nulo: como vimos, $P_T(0) = h(T)$, logo,

$$h(T) = P_T(0) = \sup_{\mu \in \mathcal{M}_T} \{h_\mu(T)\}.$$

O que vamos discutir agora é como a pressão topológica nos ajuda no cálculo da maximização da integral de um potencial $f \in C(X)$. A ferramenta utilizada é a caracterização do princípio variacional e um processo que denominamos congelamento do sistema. Dado potencial $f \in C(X)$, considere um parâmetro $t \in \mathbb{R}^+$. Pelo princípio variacional, temos a igualdade

$$P_T(tf) = \sup_{\mu \in \mathcal{M}_T} \left\{ h_\mu(T) + t \int f \, d\mu \right\}.$$

Supondo que $h(T) < \infty$, se multiplicarmos ambos os lados da equação por $\frac{1}{t}$ e tomarmos o limite $t \rightarrow \infty$, podemos estudar propriedades do limite

$$\limsup_{t \rightarrow \infty} \frac{P_T(tf)}{t} = \limsup_{t \rightarrow \infty} \sup_{\mu \in \mathcal{M}_T} \left\{ \frac{1}{t} h_\mu(T) + \int f \, d\mu \right\} = \max_{\mu \in \mathcal{M}_T} \int f \, d\mu.$$

Em dinâmica estatística, esse processo é chamado de congelamento do sistema, onde o parâmetro t representa o inverso da temperatura (a nomenclatura possui motivações na

física). Utilizando ainda tal raciocínio, seria bastante interessante obter não só o valor $\beta(f)$, mas também probabilidades invariantes que sejam maximizantes. Isto pode ser obtido pelo estudo das probabilidades que satisfazem a igualdade do princípio variacional. Tais probabilidades possuem nomenclatura especial.

Definição 2.1.19. *Dado um sistema dinâmico topológico (X, T) , dizemos que $\mu \in \mathcal{M}_T$ é um estado de equilíbrio associado a f se*

$$P_T(f) = h_\mu(T) + \int f \, d\mu.$$

Suponha que, para cada $t > 0$, exista um estado de equilíbrio μ_t associado ao potencial tf , isto é,

$$P_T(tf) = h_{\mu_t}(T) + t \int f \, d\mu_t.$$

Logo, vale a igualdade

$$\limsup_{t \rightarrow \infty} \frac{P_T(tf)}{t} = \limsup_{t \rightarrow \infty} \int f \, d\mu_t = \max_{\mu \in \mathcal{M}_T} \int f \, d\mu.$$

Em síntese, se uma probabilidade μ_∞ é ponto de aderência de $\{\mu_t\}_{t \in \mathbb{R}}$, então, μ_∞ é probabilidade maximizante para f . Ou seja, se tivermos alguma ferramenta que nos permita construir estados de equilíbrio para os potenciais tf , então, o processo descrito acima fornecerá probabilidade maximizante para o potencial f . Um exemplo da aplicação do processo de congelamento do sistema e do estudo de propriedades da convergência de μ_t pode ser visto em [3].

Na próxima seção mostraremos um processo canônico para se obter estados de equilíbrio para potenciais lipschitzianos, que é a aplicação do operador de transferência.

2.2 O operador de transferência

Após ter visto alguns conceitos básicos sobre formalismo termodinâmico e mais uma abordagem do problema de maximização da integral de potenciais $f \in C(X)$, vamos estudar propriedades do operador de transferência e sua importância para otimização ergódica. Por meio de um conhecido teorema de ponto fixo, chamado teorema de Ruelle-Perron-Frobenius, obtemos estados de equilíbrio para os potenciais estudados. O leitor poderá encontrar diferentes definições na literatura para esse operador, como, por exemplo,

em [27, 2, 5], mas, em linhas gerais, o interesse é obter pontos fixos do operador.

Relembramos ao leitor que estamos trabalhando no contexto de um SDT (X, T) , isto é, X é um espaço métrico compacto e $T: X \rightarrow X$ uma aplicação contínua. Além disso, assumimos que a aplicação T é expansiva, como descrito na definição 2.1.9. Antes de estudar o operador de transferência, definimos uma outra propriedade de dilatação para a aplicação T , a qual ajudar-nos-á estabelecer alguns resultados.

Definição 2.2.1. *Seja (X, T) um SDT. Dizemos que a aplicação T é **expansora** se existir constantes $\eta > 0$ e $0 < \lambda < 1$ tais que*

$$d(x, y) < \eta \quad \Rightarrow \quad \lambda d(Tx, Ty) \geq d(x, y).$$

Na literatura, a hipótese mais abordada para T é a expansividade, porém, sob nosso contexto de SDT, as definições de aplicação expansiva e aplicação expansora são equivalentes. Isto é garantido pela proposição a seguir.

Proposição 2.2.2. *Seja (X, T) um SDT. A aplicação $T: X \rightarrow X$ é expansiva se, e somente se, for também expansora.*

Demonstração. (\Leftarrow) Escolhemos $\varepsilon_0 = \eta$ como constante de expansividade. Por hipótese, $d(Tx, Ty) \geq \lambda^{-1}d(x, y)$ sempre que $x \neq y$ e $d(x, y) < \eta$. Então, se há $x, y \in X$ tais que, para todo natural n ocorrer $d(T^n x, T^n y) < \varepsilon_0$, é porque $\varepsilon_0 > \lambda^{-n}d(x, y) \forall n \in \mathbb{N}$. Logo, a única possibilidade de isso ocorrer é se $d(x, y) = 0$, pois $\lambda^{-n} \rightarrow \infty$ quando $n \rightarrow \infty$.

(\Rightarrow) A prova desta implicação passa por alguns lemas técnicos. É preciso argumentar que T é expansora se, e só se, T^n é expansora para algum $n \in \mathbb{N}$. Também será necessário construir vizinhanças abertas, encaixadas, sobre a diagonal do cartesiano $X \times X$, tais que sua interseção é a própria diagonal, e uma métrica auxiliar no cartesiano – devido ao lema de metrização de Frink – sobre a qual é possível mostrar que T^N é expansora para algum $N \in \mathbb{N}$. Para os detalhes da prova veja, por exemplo, [30]. \square

Uma vez provada esta equivalência de definições, continuamos chamando a aplicação $T: X \rightarrow X$ de expansiva, mas, quando necessário, fazemos uso da definição de aplicação expansora.

No estudo de otimização ergódica, muitas vezes lidamos com funções e operadores tais que sua definição em um ponto $x \in X$ depende das pré-imagens $y \in T^{-1}x$. Geralmente, operadores desta forma atuam como reguladores de funções. Então, se queremos obter

estimativas a respeito de regularidade, é interessante que pontos suficientemente próximos tenham o mesmo número de pré-imagens – em alguns casos, é necessário assumir essa hipótese para a aplicação T . A seguir, mostramos que a compacidade de X e a expansividade de T são suficientes para garantir que o número de pré-imagens por T é localmente constante.

Lema 2.2.3. *Seja (X, T) um SDT em que T é uma aplicação expansiva. Então, o número de pré-imagens de cada ponto $x \in X$ é finito e localmente constante.*

Demonstração. Finitude: Seja $x \in X$. Se o conjunto $T^{-1}x := \{y \in X : Ty = x\}$ for infinito, pela compacidade de X é possível encontrar $y_1 \neq y_2 \in T^{-1}x$ tais que $d(y_1, y_2) < \varepsilon_0$. Como $T^n y_1 = T^n y_2 \forall n \geq 1$, a definição de expansividade fornece $y_1 = y_2$, o que é um absurdo, uma vez que os escolhemos diferentes.

Localmente constante: Fixe $x \in X$ e tome $y \in T^{-1}x$. Como a cardinalidade de $T^{-1}x$ é finita, podemos escolher $0 < \delta < \eta$ tal que, para todo $y \in T^{-1}x$, temos

$$T^{-1}x \cap B_{\frac{\delta}{2}}(y) = \{y\}.$$

Da expansividade, é possível concluir que

$$B_{\frac{\delta}{2\lambda}}(x) \subseteq T\left(B_{\frac{\delta}{2}}(y)\right).$$

Considere, agora, $z \in B_{\frac{\delta}{2\lambda}}(x)$. A inclusão acima garante que há $w \in B_{\frac{\delta}{2}}(y)$ tal que $T(w) = z$. Isso nos garante que $\#T^{-1}z \geq \#T^{-1}x$. Por outro lado, perceba que $y \in B_{\frac{\delta}{2}}(w)$ e que $T^{-1}z \cap B_{\frac{\delta}{2}}(w) = \{w\}$, pois $\delta < \varepsilon_0 = \varepsilon_0$. Logo, $\#T^{-1}x \geq \#T^{-1}z$, o que conclui a demonstração. \square

O lema acima nos permite definir $r \in \mathbb{N}$ tal que $\#T^{-1}x \leq r, \forall x \in X$. Para isso provar, considere a cobertura aberta $\mathcal{A} := \left\{B_{\frac{\delta}{2\lambda}}(x) : x \in X\right\}$. Pelo lema, a cardinalidade do conjunto de pré-imagens é constante sobre cada uma dessas bolas. A compacidade de X , por sua vez, permite refinar tal cobertura para uma subcobertura finita $\{B_{\frac{\delta}{2\lambda}}(x_1), \dots, B_{\frac{\delta}{2\lambda}}(x_k)\}$. Definimos, então,

$$r := \max_{1 \leq i \leq k} \#T^{-1}x_i.$$

Após fornecer essas propriedades da dinâmica que estamos estudando, vamos definir o operador de transferência.

Definição 2.2.4. Dado um potencial $f \in C(X)$, o operador $\mathcal{L}_f: C(X) \rightarrow C(X)$ é definido como segue:

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

É direto da definição que \mathcal{L}_f é um operador linear e limitado, pois, para $g \in C(X)$, $\|\mathcal{L}_f(g)\|_\infty \leq r e^{\|f\|_\infty} \|g\|_\infty$ (recorde que r denota a cardinalidade máxima do conjunto de pré-imagens de um ponto $x \in X$). Portanto, o operador \mathcal{L}_f é contínuo. Quando o operador de transferência fixa as funções constantes, dizemos que \mathcal{L}_f está normalizado. Note que fixar funções constantes significa

$$\mathcal{L}_f 1(x) = \sum_{y \in T^{-1}x} e^{f(y)} = 1, \quad \forall x \in X,$$

onde $1(x)$ denota a função constante e igual a 1.

Mostramos abaixo que \mathcal{L}_f preserva regularidade de funções, quando f é um potencial lipschitziano.

Lema 2.2.5. *Seja $v: [0, \infty) \rightarrow [0, \infty)$ função contínua na origem, não decrescente, tal que $v(0) = 0$. Consideramos o cone das funções contínuas e positivas tais que v é módulo de continuidade para $\log(g)$, ou seja,*

$$\mathcal{C}_v := \left\{ 0 < g \in C(X) : \left| \log \left(\frac{g(x)}{g(z)} \right) \right| \leq v(d(x, z)), \text{ se } d(x, z) \leq \frac{\eta}{2\lambda} \right\} \cup \{g \equiv 0\}.$$

Se f é função lipschitziana, podemos estabelecer módulo de continuidade local u tal que $\mathcal{L}_f(\mathcal{C}_v) \subseteq \mathcal{C}_u$. Em particular, para $\tilde{v}(s) = \text{Lip}(f) \frac{\lambda s}{1-\lambda}$, onde $\text{Lip}(f)$ denota a constante Lipschitz de f , verificamos que o cone $\mathcal{C}_{\tilde{v}}$ é invariante por \mathcal{L}_f .

Demonstração. Considere uma função qualquer $g \in \mathcal{C}_v$. Por definição,

$$\mathcal{L}_f g(x) = \sum_{Ty=x} e^{f(y)} g(y).$$

Pelo lema 2.2.3, dado δ_1 suficientemente pequeno (ou seja, $\delta_1 < \frac{\eta}{2}$), se $x, z \in X$ satisfazem $d(x, z) \leq \delta_1$, então, a cardinalidade do conjunto de pré-imagens de x e z é a mesma. Além disso, para cada pré-imagem $y \in T^{-1}x$ há $w = w(y) \in T^{-1}z$ tal que $d(y, w) \leq \lambda d(x, z)$.

Nessas condições, podemos escrever as seguintes desigualdades:

$$\begin{aligned}
 \mathcal{L}_f g(x) = \sum_{Ty=x} e^{f(y)} g(y) &\leq \sum_{\substack{Ty=x, w=w(y) \\ d(y,w) \leq \lambda d(x,z), Tw=z}} e^{f(w)+Lip(f)d(y,w)} g(y) \\
 &\leq \sum_{Tw=z} e^{f(w)+Lip(f)d(y,w)} g(w) e^{v(d(w,y))} \\
 &\leq \mathcal{L}_f g(z) e^{Lip(f)\lambda d(x,z)+v(\lambda d(x,z))}.
 \end{aligned}$$

Por simetria, podemos estabelecer também a desigualdade

$$\mathcal{L}_f g(z) \leq \mathcal{L}_f g(x) e^{Lip(f)\lambda d(x,z)+v(\lambda d(x,z))}.$$

Portanto, se $d(x, z) \leq \delta_1$, obtemos

$$|\log(\mathcal{L}_f g(x)) - \log(\mathcal{L}_f g(z))| \leq Lip(f)\lambda d(x, z) + v(\lambda d(x, z)) = u(d(x, z)).$$

No caso particular $\tilde{v}(s) = Lip(f) \frac{\lambda s}{1-\lambda}$, vemos que

$$\begin{aligned}
 |\log(\mathcal{L}_f g(x)) - \log(\mathcal{L}_f g(z))| &\leq Lip(f)\lambda d(x, z) + Lip(f) \frac{\lambda^2 d(x, z)}{1-\lambda} \\
 &= Lip(f) \frac{\lambda d(x, z)}{1-\lambda} \\
 &= \tilde{v}(d(x, z)).
 \end{aligned}$$

Além disso, como $g > 0$, obviamente $\mathcal{L}_f g > 0$, de modo que $\mathcal{C}_{\tilde{v}}$ é preservado por \mathcal{L}_f . \square

O lema acima revela que funções com boa regularidade possuem a ótima propriedade de ser invariantes por \mathcal{L}_f . Sendo assim, vamos trabalhar sobre o espaço de funções lipschitzianas, denotado por $\mathcal{H}_\lambda^1(X)$. A constante Lipschitz de g , como mencionada anteriormente, é denotada por $Lip(g)$, isto é,

$$Lip(g) := \inf\{k \in \mathbb{R} : |g(x) - g(y)| \leq kd(x, y), \forall x, y \in X\}.$$

Munimos $\mathcal{H}_\lambda^1(X)$ da seguinte norma

$$\|g\|_\lambda := \|g\|_\infty + Lip(g).$$

Com esta definição, $(\mathcal{H}_\lambda^1, \|\cdot\|_\lambda)$ é um espaço de Banach de funções lipschitzianas. Reparem que uma definição análoga pode ser considerada para tratar do espaço de funções α -Holder; neste caso, a notação vem a ser $\mathcal{H}_\lambda^\alpha(X)$ e a norma considerada é $\|\cdot\|_\lambda = \|\cdot\|_\infty + Hold_\alpha(\cdot)$, onde $Hold_\alpha$ denota a constante holderiana da função. É fácil perceber que, como X é um espaço métrico compacto, uma simples alteração da distância definida em X transforma funções holderianas em funções lipschitzianas. Além disso, a constante λ nesta definição é devido à expansividade da aplicação T .

Dando sequência no estudo das propriedades do operador de transferência, a proposição a seguir (que pode ser vista, por exemplo, em [27, 2]) é uma propriedade chave sobre a norma $\|\cdot\|_\lambda$.

Proposição 2.2.6. *Dado um potencial $f \in \mathcal{H}_\lambda^1(X)$, suponha que o operador \mathcal{L}_f está normalizado. Então, para qualquer $g \in \mathcal{H}_\lambda^1(X)$ há $C = C(g) > 0$ tal que*

$$\|\mathcal{L}_f^n(g)\|_\lambda \leq C\|g\|_\infty + \lambda^n\|g\|_\lambda.$$

Demonstração. Pelo lema 2.2.3, o número de pré-imagens de cada ponto $x \in X$ é finito e localmente constante, sendo possível encontrar $0 < \delta < \eta$ tal que $\#T^{-1}x = \#T^{-1}z$ quando $d(x, z) \leq \delta$. Além disso, para cada $y \in T^{-1}x$ há um único $w \in T^{-1}z$ cumprindo $\lambda d(x, z) \geq d(y, w)$. A prova dar-se-á por indução. Para $n = 1$, tome $x, z \in X$ com $d(x, z) < \delta < \frac{\eta}{2}$. Temos que

$$\begin{aligned} |\mathcal{L}_f g(x) - \mathcal{L}_f g(z)| &= \left| \sum_{y \in T^{-1}x} e^{f(y)} g(y) - \sum_{w \in T^{-1}z} e^{f(w)} g(w) \right| \\ &\leq \sum_{\substack{y \in T^{-1}x \\ w \in T^{-1}z}} \left\{ |e^{f(y)} - e^{f(w)}| |g(y)| + e^{f(w)} |g(y) - g(w)| \right\} \\ &\leq r\lambda d(x, z) Lip(e^f) \|g\|_\infty + \lambda d(x, z) Lip(g), \end{aligned}$$

onde r denota o número máximo de pré-imagens de um ponto $x \in X$. Note que foi utilizada a hipótese de normalização na segunda parcela do somatório. Sendo assim, como X é compacto, há constante C_0 , dependendo apenas de δ e X tal que

$$Lip(\mathcal{L}_f g) \leq \lambda C_0 \|g\|_\infty + \lambda Lip(g).$$

De fato, para garantir a existência de C_0 , basta considerar dois casos: se $d(x, z) < \delta$,

prossequimos como acima, majorando $rLip(e^f)$ por C_0 . Caso contrário, para $d(x, z) \geq \delta$, como X é compacto e $\mathcal{L}_f g$ contínua, basta tomar C_0 tal que

$$|\mathcal{L}_f g(x) - \mathcal{L}_f g(z)| \leq 2\|\mathcal{L}_f g\|_\infty \leq (\lambda C_0 \|g\|_\infty + \lambda Lip(g))\delta.$$

Por outro lado, como o operador está normalizado, temos que $\|\mathcal{L}_f g\|_\infty \leq \|g\|_\infty$, donde

$$\begin{aligned} \|\mathcal{L}_f g\|_\lambda &= \|\mathcal{L}_f g\|_\infty + Lip(\mathcal{L}_f g) \\ &\leq \|g\|_\infty + \lambda C_0 \|g\|_\infty + \lambda Lip(g) \\ &= (\lambda C_0 + 1 - \lambda) \|g\|_\infty + \lambda \|g\|_\lambda. \end{aligned}$$

Logo, tomamos $C_1 = (\lambda C_0 + 1 - \lambda)$ para concluir este primeiro caso.

Agora, por indução, supomos que $\|\mathcal{L}_f^n g\|_\lambda \leq C_n \|g\|_\infty + \lambda^n \|g\|_\lambda$. Segue que

$$\begin{aligned} \|\mathcal{L}_f^{n+1} g\|_\lambda &= \|\mathcal{L}_f^n(\mathcal{L}_f g)\|_\lambda \\ &\leq C_n \|\mathcal{L}_f g\|_\infty + \lambda^n \|\mathcal{L}_f g\|_\lambda \\ &\leq (C_n + \lambda^n C_1) \|g\|_\infty + \lambda^{n+1} \|g\|_\lambda. \end{aligned}$$

Basta então definir $C_{n+1} = C_n + \lambda^n C_1$. Para concluir, note que

$$C_n = \sum_{i=0}^{n-1} \lambda^i C_1 \leq C_1 \left(\frac{1}{1 - \lambda} \right),$$

donde $C := C_1 \frac{1}{1 - \lambda}$ é a constante procurada, verificando a desigualdade para todo $n \in \mathbb{N}$. \square

Tal propriedade sugere uma contração na norma $\|\cdot\|_\lambda$, ao iterarmos o operador \mathcal{L}_f . Ela é importante para a prova do teorema a seguir, o clássico teorema de Ruelle-Perron-Frobenius. Note que, para o teorema de Ruelle-Perron-Frobenius, será necessário exigir uma hipótese a mais sobre a aplicação T a qual é bastante usual em teoria ergódica.

Definição 2.2.7. *Seja (X, T) um SDT. Dizemos que a aplicação T é transitiva se há $x \in X$ tal que $orb(x)$ é um conjunto denso de X .*

Teorema 2.2.8 (Ruelle-Perron-Frobenius). *Sejam (X, T) um SDT no qual T é aplicação expansiva e transitiva, $f \in H_\lambda^1(X)$ e \mathcal{L}_f o operador de transferência. Então:*

- (i) *há um autovalor θ de \mathcal{L}_f que é real, simples e positivo, com autofunção $h > 0$;*

- (ii) os demais autovalores de \mathcal{L}_f possuem módulo estritamente menor que θ ;
- (iii) há uma única medida de probabilidade μ_0 tal que $\mathcal{L}_f^* \mu_0 = \theta \mu_0$. Em outras palavras, $\int \mathcal{L}_f g \, d\mu_0 = \theta \int g \, d\mu_0$, para toda $g \in C(X)$;
- (iv) $\theta^{-n} \mathcal{L}_f^n g \rightarrow \int h g \, d\mu_0$, para qualquer $g \in C(X)$, onde h é a autofunção correspondente a θ tal que $\int h \, d\mu_0 = 1$;
- (v) $\mu = \frac{h\mu_0}{\theta} \in \mathcal{M}_T$.

Como mencionamos acima, este teorema é bastante abordado na literatura de otimização ergódica. A maioria das demonstrações passa por demonstrar que o operador \mathcal{L}_f age como uma contração fraca em cones de funções com algumas propriedades de regularidade (assim como mostramos acima), e, então, se utilizam teoremas de ponto fixo para encontrar a autofunção. Deixamos indicadas algumas referências nas quais o leitor pode consultar diferentes provas desse teorema [2, 5, 15, 27].

Gostaríamos de discutir sobre como tal teorema nos ajuda a obter informações a respeito da maximização da integral de um potencial f . Bem, caso o operador de transferência não seja normalizado para $f \in \mathcal{H}_\lambda^1$, através do teorema de Ruelle-Perron-Frobenius é possível obter uma normalização associada.

Proposição 2.2.9. *Sejam (X, T) um SDT transitivo e expansivo, e $f \in H_\lambda^1(X)$. Sejam θ e h_0 o autovalor e a autofunção de \mathcal{L}_f , obtidos na aplicação do teorema. Então, para*

$$\tilde{f} = f + \log(h) - \log(h \circ T) - \log(\theta),$$

o operador de transferência $\mathcal{L}_{\tilde{f}}$ está normalizado.

Demonstração. A prova é direta. Basta ver que para todo $x \in X$, temos

$$\begin{aligned} \mathcal{L}_{\tilde{f}} 1(x) &= \sum_{y \in T^{-1}x} e^{f(y) + \log(h_0(y)) - \log(h_0 \circ T(y)) - \log(\theta)} \\ &= \sum_{y \in T^{-1}x} \frac{e^{f(y)} h_0(y)}{\theta h_0(x)} \\ &= \frac{1}{\mathcal{L}_f h(x)} \sum_{y \in T^{-1}x} e^{f^*(y)} h_0(y) \\ &= \frac{1}{\mathcal{L}_f h(x)} \mathcal{L}_f h(x) = 1. \end{aligned}$$

□

Chamamos $\mathcal{L}_{\tilde{f}}$ de normalização associada de \mathcal{L}_f . Repare ainda que, quando o operador \mathcal{L}_f está normalizado, como $\mathcal{L}_f 1(x) = 1 \forall x \in X$, temos a autofunção $h_0 \equiv 1$ associada ao autovalor $\theta = 1$. Logo, a probabilidade μ_0 obtida no teorema como automedida do operador dual \mathcal{L}_f^* é invariante por T .

Lembramos que, apesar de utilizar as hipóteses de expansividade da aplicação $T: X \rightarrow X$, se há um outro SDT (Y, S) e uma conjugação entre T e S , isto é, uma aplicação $U: X \rightarrow Y$ tal que $U \circ T = S \circ U$, o processo pode ser aplicado para (X, T) e transportado para (Y, S) por meio da conjugação U . Na próxima seção vamos exemplificar a aplicação do teorema de Ruelle-Perron-Frobenius sob uma dinâmica de *subshift* de tipo finito. Mostramos que a probabilidade obtida no teorema é o único estado de equilíbrio para f .

2.2.1 Subshift de tipo finito

Começamos introduzindo nosso espaço e a dinâmica sobre ele. Dado um natural $r \geq 2$, considere uma matriz M de dimensão $r \times r$ com entradas 0 ou 1 tal que toda linha e toda coluna de M possuem ao menos uma entrada não nula. Tais matrizes são chamadas matrizes de transição. Esta matriz é a base da dinâmica, uma transição do índice i para o índice j é permitida se $M_{ij} = 1$, do contrário ($M_{ij} = 0$), a transição não será permitida. Vamos supor que nossa matriz de transição M é aperiódica, como na definição abaixo.

Definição 2.2.10. *Uma matriz M é dita ser aperiódica se há $k_0 \in \mathbb{N}$ tal que M^k possui todas as entradas positivas, para todo $k \geq k_0$.*

A teoria de matrizes não negativas pode ser associada à teoria de grafos. Neste caso, nossa matriz M modela um grafo orientado conexo tal que, para todo $k \geq k_0$ e vértices i e j , há um caminho de tamanho k que começa em i e termina em j . Podemos, então, considerar cada passeio infinito sobre o grafo associado à matriz M como um elemento de um espaço, e introduzir neste espaço a aplicação de caminho sobre o passeio. Essa breve divagação feita é para ilustrar o que vem a ser um *subshift* de tipo finito.

Definição 2.2.11. *Chamamos de subshift de tipo finito gerado por M – uma matriz de transição e aperiódica – o espaço de sequências*

$$\Sigma_M := \left\{ \bar{x} = (x_0, x_1, x_2, \dots) : x_i \in \{1, \dots, r\} \text{ e } M_{x_i x_{i+1}} = 1, \forall i \in \mathbb{N} \right\}. \quad (2.1)$$

Dada uma constante $0 < \lambda < 1$, definimos uma distância em Σ_M como segue:

$$d(\bar{x}, \bar{y}) = \lambda^k, \quad \text{onde} \quad k := \min\{i \in \mathbb{N} : x_i \neq y_i\}.$$

Já a dinâmica, como o próprio nome sugere, é induzida por um shift nas sequências, isto é,

$$\sigma(x_0, x_1, x_2, \dots) = (x_1, x_2, \dots).$$

Os abertos básicos desse espaço são constituídos de cilindros de tamanho n , ou seja,

$$[i_0, \dots, i_n] := \{\bar{x} \in \Sigma_M : x_0 = i_0, x_1 = i_1, \dots, x_n = i_n\}.$$

Tais abertos também formam uma partição de Σ_M . Fica a cargo do leitor verificar tais afirmações, bem como a afirmação de que σ é uma aplicação transitiva. Além disso, se $x_0 \neq y_0$, então $d(\bar{x}, \bar{y})\lambda = d(\sigma\bar{x}, \sigma\bar{y})$, ou seja, σ é uma aplicação expansiva, com constante de expansividade $\varepsilon_0 = 1$.

Considere, então, $f: \Sigma_M \rightarrow \mathbb{R}$ uma função lipschitziana. Como as pré-imagens de um ponto $\bar{x} \in \Sigma_M$ são exatamente os pontos $i\bar{x} = (i, x_0, x_1, \dots)$ tais que $M_{ix_0} = 1$, então, a definição de \mathcal{L}_f pode ser escrita como

$$\mathcal{L}_f g(\bar{x}) = \sum_{i=1}^r M_{ix_0} e^{f(i\bar{x})} g(i\bar{x}), \quad \forall g \in C(\Sigma_M),$$

onde $i\bar{x} = (i, x_0, x_1, \dots)$. Supomos que \mathcal{L}_f está normalizado para obter o seguinte resultado.

Teorema 2.2.12. *Sejam (Σ_M, σ) um subshift de tipo finito, onde M é matriz aperiódica e $f: \Sigma_M \rightarrow \mathbb{R}$ uma função lipschitziana tal que \mathcal{L}_f está normalizado. Então, a probabilidade $\mu_0 \in \mathcal{M}_\sigma$ tal que $\mathcal{L}_f^* \mu_0 = \mu_0$ é o único estado de equilíbrio para f .*

Para provar esse lema são necessários alguns resultados complementares.

Lema 2.2.13. *Sejam $p_1, \dots, p_r, q_1, \dots, q_r$ números reais tais que*

$$0 < p_i, q_i < 1, \quad e \quad \sum_{i=1}^r p_i = \sum_{i=1}^r q_i = 1.$$

Então, temos a desigualdade

$$-\sum_{i=1}^r q_i \log(q_i) + \sum_{i=1}^r q_i \log(p_i) \leq 0,$$

sendo que a igualdade se dá apenas quando $p_i = q_i, \forall i = 1, \dots, r$.

Demonstração. Basta reescrever o somatório usando a propriedade de soma do logaritmo e utilizar a concavidade da função $\phi(x) = -x \log(x)$ para mostrar a desigualdade. \square

Lema 2.2.14. *Nas condições do teorema acima, considere μ_0 a automedida de \mathcal{L}_f^* , $\mu \in \mathcal{M}_T$ uma probabilidade qualquer e $g \in C(\Sigma_M)$. Então, temos as igualdades*

$$\lim_{n \rightarrow \infty} \sum_{i=1}^r \int g(i\bar{x}) \frac{\mu[i, x_1, \dots, x_n]}{\mu[x_1, \dots, x_n]} d\mu(\bar{x}) = \int g d\mu, \quad (\star)$$

$$\lim_{n \rightarrow \infty} \sum_{i=1}^r \int \frac{\mu[i, x_1, \dots, x_n]}{\mu[x_1, \dots, x_n]} \log \left(\frac{\mu_0[i, x_1, \dots, x_n]}{\mu_0[x_1, \dots, x_n]} \right) d\mu(\bar{x}) = \int f d\mu(\bar{x}). \quad (\star\star)$$

Para uma prova desse lema o leitor pode consultar, por exemplo, [27].

Prova do Teorema: Como um cilindro $[x_1, \dots, x_n] \subseteq \Sigma_M$ pode ser escrito como

$$[x_1, \dots, x_n] = \sqcup_{i=1}^r [i, x_1, \dots, x_n],$$

o primeiro lema fornece a seguinte desigualdade

$$-\sum_{i=1}^r \frac{\mu[i, x_1, \dots, x_n]}{\mu[x_1, \dots, x_n]} \log \left(\frac{\mu[i, x_1, \dots, x_n]}{\mu[x_1, \dots, x_n]} \right) + \sum_{i=1}^r \frac{\mu[i, x_1, \dots, x_n]}{\mu[x_1, \dots, x_n]} \log \left(\frac{\mu_0[i, x_1, \dots, x_n]}{\mu_0[x_1, \dots, x_n]} \right) \leq 0. \quad (2.2)$$

O segundo lema, por sua vez, garante que

$$\lim_{n \rightarrow \infty} \sum_{i=1}^r \int \frac{\mu[i, x_1, \dots, x_n]}{\mu[x_1, \dots, x_n]} \log \left(\frac{\mu_0[i, x_1, \dots, x_n]}{\mu_0[x_1, \dots, x_n]} \right) d\mu(\bar{x}) = \int f d\mu(\bar{x}).$$

Fica a cargo do leitor verificar que $\mathcal{P} = \{[i] : i = 1, \dots, r\}$ constitui uma partição de Σ_M tal que $\text{diam}(\mathcal{P}^n) \rightarrow 0$, quando $n \rightarrow \infty$. Com isso, é válida a seguinte igualdade para a entropia métrica

$$h_\mu(\sigma) = h_\mu(\sigma, \mathcal{P}) = \lim_{n \rightarrow \infty} H_\mu(\mathcal{P} | \bigvee_{i=1}^n \sigma^{-i}(\mathcal{P})).$$

Escrevendo em termos dos cilindros temos

$$\begin{aligned} H_\mu(\mathcal{P} | \bigvee_{i=1}^n \sigma^{-i}(\mathcal{P})) &= -\sum_{i=1}^r \sum_{x_1, \dots, x_n} \mu[i, x_1, \dots, x_n] \log \left(\frac{\mu[i, x_1, \dots, x_n]}{\mu[x_1, \dots, x_n]} \right) \\ &= -\sum_{i=1}^r \frac{\mu[i, x_1, \dots, x_n]}{\mu[x_1, \dots, x_n]} \log \left(\frac{\mu[i, x_1, \dots, x_n]}{\mu[x_1, \dots, x_n]} \right). \end{aligned}$$

Logo, tomando o limite quando $n \rightarrow \infty$ em (2.2), obtemos que

$$h_\mu(\sigma) + \int f \, d\mu \leq 0.$$

Falta verificar que a igualdade ocorre apenas para $\mu = \mu_0$.

Iremos utilizar um resultado de convergência de martingales que garante a existência do limite $\mu_i(\bar{x}) = \lim_{n \rightarrow \infty} \mu[i, x_1, \dots, x_n] \mu[x_1, \dots, x_n]^{-1}$ μ -q.t.p., $\bar{x} \in \Sigma_M$. Além disso, note que

$$(\mu_0)_i(\bar{x}) = \lim_{n \rightarrow \infty} \frac{\mu_0[i, x_1, \dots, x_n]}{\mu_0[x_1, \dots, x_n]} = e^{f(i, x_1, x_2, \dots)}.$$

Esta afirmação é obtida ao se mostrar que

$$e^{-Lip(f)\lambda^n} \leq \frac{\mu_0[x_0, x_1, \dots, x_n]}{\mu_0[x_1, \dots, x_n]} e^{-f(\bar{x})} \leq e^{Lip(f)\lambda^n}.$$

O leitor pode consultar a prova dessas afirmações, por exemplo, em [27]. Tais propriedades produzem a seguinte desigualdade em (2.2):

$$-\sum_{i=1}^r \int \mu_i(\bar{x}) \, d\mu(\bar{x}) + \sum_{i=1}^r \int \mu_i(\bar{x}) f(i, x_1, x_2, \dots) \, d\mu(\bar{x}) \leq 0,$$

sendo que, pelo primeiro lema, a igualdade ocorrerá somente quando $\mu_i(\bar{x}) = e^{f(i, x_1, x_2, \dots)}$, μ -q.t.p., $\bar{x} \in \Sigma_M$. Logo, por (\star) no segundo lema, pelo fato de μ ser σ -invariante e pela definição de \mathcal{L}_f , para toda $g \in C(\Sigma_M)$ temos

$$\begin{aligned} \int \mathcal{L}_f g \, d\mu &= \int \mathcal{L}_f g \circ \sigma \, d\mu \\ &= \sum_{i=1}^r \int e^{f(i, x_1, x_2, \dots)} g(i, x_1, x_2, \dots) \, d\mu(\bar{x}) \\ &= \sum_{i=1}^r \int g(i, x_1, \dots) \mu_i(\bar{x}) \, d\mu(\bar{x}) \quad (\star) \\ &= \int g \, d\mu. \end{aligned}$$

Mas, então, $\mathcal{L}_f^* \mu = \mu$, e isto ocorre se, e somente se, $\mu = \mu_0$. Ou seja, μ_0 é o único estado de equilíbrio para f .

□

Observação. Note que, quando \mathcal{L}_f não está normalizado, consideramos sua normalização

$\mathcal{L}_{\tilde{f}}$, onde $\tilde{f} = f + \log(h) - \log(h \circ \sigma) - \log(\theta)$. Aplicando o teorema acima, resulta que $P_\sigma(f) = \log(\theta)$ e que $\mu = \frac{h\mu_0}{\theta}$ é o único estado de equilíbrio de f .

Relembremos agora o processo de congelamento do sistema ao se considerar

$$\lim_{t \rightarrow \infty} \frac{P_\sigma(tf)}{t} = \sup_{\mu \in \mathcal{M}_\sigma} \left\{ \frac{h_\mu(\sigma)}{t} + \int f d\mu \right\}.$$

Se $f \in \mathcal{H}_\lambda^1(\Sigma)$, para cada $t > 0$, podemos aplicar o teorema de Ruelle-Perron-Frobenius ao potencial tf , de modo a encontrar um único estado de equilíbrio que chamamos de μ_t . Lembramos que \mathcal{M}_σ é compacto na topologia fraca-*, de modo que a sequência $\{\mu_t\}_{t>0}$ admite ponto de acumulação μ_∞ .

Proposição 2.2.15. *A probabilidade μ_∞ é maximizante para o potencial f .*

Demonstração. Seja $\mu \in \mathcal{M}_\sigma$ uma probabilidade qualquer. Então, como $h_\mu(\sigma) \geq 0$, temos $t \int f d\mu \leq h_\mu(\sigma) + t \int f d\mu$. Mas tf possui apenas um estado de equilíbrio, que é μ_t . Logo,

$$h_\mu(\sigma) + t \int f d\mu \leq h_{\mu_t}(\sigma) + t \int f d\mu_t.$$

Mas, lembre que, pelo princípio variacional da entropia, $h_{\mu_t}(\sigma) \leq h(\sigma)$, donde concluímos que

$$t \int f d\mu \leq h(\sigma) + t \int f d\mu_t.$$

Dividindo ambos os lados da desigualdade acima por t e tomando o limite quando $t \rightarrow \infty$, obtemos

$$\int f d\mu \leq \int f d\mu_\infty, \quad \forall \mu \in \mathcal{M}_\sigma.$$

□

Observação. O fato de μ_∞ ser um ponto de acumulação de $\{\mu_t\}$ não significa que $\mu_t \xrightarrow{*} \mu_\infty$. Há exemplos na literatura de casos em que há a convergência de μ_t como, por exemplo, [9] para potenciais localmente constantes, e de não convergência, como, por exemplo, [12] para um potencial holderiano. A referência já citada [3] estuda como se dá a convergência de μ_t , para um potencial localmente constante que possui mais de uma probabilidade maximizante.

Subações para Difeomorfismos de Anosov

Caro leitor, enfim chegamos ao capítulo final da presente dissertação, no qual discutimos conceitos de otimização ergódica para difeomorfismos de Anosov em variedades compactas. Definimos o que é um difeomorfismo de Anosov, seguindo a definição de [22]. Através de uma partição markoviana, mostramos uma importante propriedade dos difeomorfismos de Anosov, que é a semiconjugação com um *subshift* de tipo finito. Apresentamos, então, dois métodos de construção de subações para potenciais holderianos associados ao difeomorfismo de Anosov, segundo [22] e [7]. Por fim, discutimos sobre propriedades das probabilidades maximizantes em uma dinâmica do tipo Anosov.

3.1 Coberturas markovianas para difeomorfismos de Anosov

Começamos esta seção pela definição de um difeomorfismo de Anosov.

Definição 3.1.1. *Sejam M uma variedade compacta, orientável e sem fronteira e a aplicação $T: M \rightarrow M$ um difeomorfismo. Dizemos que T é um difeomorfismo de Anosov se para todo ponto $p \in M$ existe uma decomposição do espaço tangente $T_p M = \mathbb{E}^s(p) \oplus \mathbb{E}^u(p)$, em dois subespaços transversais $\mathbb{E}^s(p)$ e $\mathbb{E}^u(p)$, tais que*

(i) *esta decomposição é contínua;*

(ii) a decomposição é invariante pela derivada de T :

$$DT(p)(\mathbb{E}^s(p)) = \mathbb{E}^s(T(p)), \quad DT(p)(\mathbb{E}^u(p)) = \mathbb{E}^u(T(p));$$

(iii) existem constantes $C_0(M) > 0$ e $0 < \gamma_s < \Gamma_s < 1 < \gamma_u < \Gamma_u$ tais que

$$\begin{aligned} C_0(M)\gamma_s^n \|v_p\| &\leq \|DT^n(p) \cdot v_p\| \leq C_0(M)\Gamma_s^n \|v_p\|, \quad \forall v_p \in \mathbb{E}^s(p), \\ C_0(M)\gamma_u^n \|v_p\| &\leq \|DT^{-n}(p) \cdot v_p\| \leq C_0(M)\Gamma_u^n \|v_p\|, \quad \forall v_p \in \mathbb{E}^u(p). \end{aligned}$$

Essa definição é motivada pelo trabalho de Lopes e Thieulen [22] e requer um bom controle sobre as constantes de dilatação e contração da aplicação T (tais constantes são fundamentais para a primeira construção de subação para potenciais holderianos). É comum encontrar na literatura definição de difeomorfismo de Anosov semelhante a esta, mas envolvendo apenas uma constante de dilatação λ tal que, para todo $n \geq 1$, $DT^n(p)$ contrai $\mathbb{E}^s(p)$ por fator λ^n e $DT^{-n}(p)$ contrai $\mathbb{E}^u(p)$, também por fator λ^n .

No capítulo anterior, vimos alguns resultados que necessitavam da hipótese de expansividade da aplicação T . Quando T não é um homeomorfismo, ser uma aplicação expansiva implica não existir direções de contração. Como agora estudamos difeomorfismos, podemos admitir a existência de direções de contração e expansão. Para entender o que são tais direções de contração e expansão, introduzimos a noção de conjunto estável e instável.

Definição 3.1.2. *Seja (X, T) um SDT em que T é um homeomorfismo. Para cada $\varepsilon > 0$ definimos os conjuntos ε -estável e ε -instável de um ponto $x \in X$, respectivamente, por*

$$\begin{aligned} W_\varepsilon^s(x) &= \{y \in X : d(T^n x, T^n y) \leq \varepsilon, \forall n \geq 0\}, \\ W_\varepsilon^u(x) &= \{y \in X : d(T^{-n} x, T^{-n} y) \leq \varepsilon, \forall n \geq 0\}. \end{aligned}$$

A expansividade, neste caso, passa a considerar toda a órbita $T^n x$, com $n \in \mathbb{Z}$, ou seja, T é uma aplicação expansiva se há $\varepsilon_0 > 0$ tal que

$$d(T^n x, T^n y) < \varepsilon_0, \quad \forall n \in \mathbb{Z} \quad \Rightarrow \quad x = y.$$

Contudo, aplicações expansivas representam caso particular de dinâmica hiperbólica. Para que uma aplicação $T: X \rightarrow X$ seja hiperbólica, é necessário que exista um certo controle sobre os conjuntos $W_\varepsilon^s(x)$ e $W_\varepsilon^u(x)$ limitados a pequenas vizinhanças dos pontos, isto é:

Definição 3.1.3. *Seja (X, T) um SDT no qual T é um homeomorfismo. Dizemos que T é uma aplicação hiperbólica se para todo $x \in X$ há constantes $\varepsilon_0, C > 0$ e $0 < \theta < 1$ tais que*

$$\begin{aligned} d(T^n y, T^n z) &\leq C\theta^n, & \forall n \leq 0, \forall y, z \in W_{\varepsilon_0}^s(x), \\ d(T^{-n} y, T^{-n} z) &\leq C\theta^n, & \forall n \leq 0, \forall y, z \in W_{\varepsilon_0}^u(x). \end{aligned}$$

É fácil perceber que todo difeomorfismo de Anosov é uma aplicação hiperbólica. Com efeito, para cada $p \in M$, as estimativas da aplicação $DT(p): T_p M \rightarrow T_p M$ garantem que $W_\varepsilon^s(p)$ e $W_\varepsilon^u(p)$ são localmente as projeções de $\mathbb{E}^s(p)$ e $\mathbb{E}^u(p)$ via, por exemplo, o mapa exponencial, interceptadas por uma vizinhança $U_p \subseteq M$. O leitor pode encontrar detalhes a respeito das ferramentas de geometria diferencial aqui empregadas, por exemplo, em [11]. Podemos notar também que todo difeomorfismo de Anosov é expansivo. De fato, a não existência de uma constante de expansividade ε_0 implicaria que, para cada $\varepsilon > 0$, há pontos $x \neq y \in M$ com $d(T^n x, T^n y) < \varepsilon$, $\forall n \in \mathbb{Z}$, resultando que $y \in W_\varepsilon^s(x) \cap W_\varepsilon^u(x)$. Esta conclusão é impossível de se obter, uma vez que a definição de T produz uma estrutura de produto local para todo ponto $x \in M$, ou seja, que há $\varepsilon > 0$ tal que $W_\varepsilon^u(x) \cap W_\varepsilon^s(x) = \{x\}$.

A constante de expansividade ε_0 é fundamental para definir unicamente o ponto de interseção de $W_{\varepsilon_0}^s(x)$ e $W_{\varepsilon_0}^u(y)$, sempre que $d(x, y) \leq \varepsilon_0$:

$$[x, y] = W_{\varepsilon_0}^s(x) \cap W_{\varepsilon_0}^u(y).$$

A partir deste conceito definimos retângulos sobre variedade M . Note que o interior de um conjunto $A \subseteq M$ é denotado por $\text{int}(A)$. Temos então a definição de retângulo.

Definição 3.1.4. *Um retângulo $R \subseteq M$ é um conjunto fechado que cumpre as seguintes propriedades*

- $R = \overline{\text{int}(R)}$;
- $x, y \in R \Rightarrow [x, y] \in R$.

Observe que, em particular, se $R \neq M$, a fronteira de R , denotada por ∂R , está localmente contida em $W^s(p) \cup W^u(p)$, para todo $p \in \partial R$.

Estamos interessados em construir uma cobertura markoviana de M . Mais ainda, queremos que nossa cobertura markoviana seja formada por retângulos. Para tanto, utilizamos uma cobertura $\mathcal{U} = \{U_1, \dots, U_l\}$ de M como base para a cobertura markoviana. Lembramos que as construções feitas nesta seção são baseadas em [22].

Definição 3.1.5. *Seja $\mathcal{U} = \{U_1, \dots, U_l\}$ uma cobertura de M por abertos. Dizemos que $\mathcal{R} = \{R_1, \dots, R_l\}$ é cobertura markoviana por retângulos associada a \mathcal{U} se cumprir*

- $U_i \subseteq R_i$;
- $x \in U_i \cap T^{-1}(U_j) \Rightarrow T(W^s(x, R_i)) \subseteq W^s(Tx, R_j)$;
- $y \in T(U_i) \cap U_j \Rightarrow T^{-1}(W^u(y, R_j)) \subseteq W^u(T^{-1}y, R_i)$;

onde $W^s(x, R_i) = W_{\varepsilon_0}^s(x) \cap R_i$ e $W^u(y, R_j) = W_{\varepsilon_0}^u(y) \cap R_j$.

Há na literatura, também, o conceito de partição markoviana, muito similar a esse de coberturas, com a ressalva de que a única interseção possível neste caso é realizada na fronteira dos retângulos.

Mostramos que é sempre possível construir uma cobertura markoviana sobre o difeomorfismo de Anosov $T: M \rightarrow M$. Para tanto, precisamos de um lema de sombreamento, introduzido por Bowen em [8]. Explicamos a seguir o que isto vem a ser.

Definição 3.1.6. *Um conjunto $\{x_i\}_{i=0}^{n-1}$ é dito ser uma δ -pseudo-órbita se $d(Tx_i, x_{i+1}) < \delta$, para todo $i = 0, \dots, n-2$. Além disso, dizemos que o trecho de órbita $\{T^i x\}_{i=0}^{n-1}$ ε -sombreia a δ -pseudo-órbita $\{x_i\}_{i=1}^n$ se*

$$d(T^i x, x_i) < \varepsilon, \quad \forall i = 0, \dots, n-1.$$

Note que as δ -pseudo-órbitas podem ser finitas ou infinitas, isto é, podemos considerar trechos $\{x_i\}_{i=0}^{n-1}$ ou, então, $\{x_i\}_{i \in \mathbb{Z}}$. Ora, os lemas de sombreamento são resultados que garantem existência de órbita sombreante como na definição acima. O lema a seguir é uma versão particular do lema sombreamento de Bowen [8].

Lema 3.1.7 (Sombreamento de Bowen). *Sejam M variedade compacta, sem fronteira de dimensão maior ou igual a 2, e $T: M \rightarrow M$ um difeomorfismo de Anosov transitivo de classe C^2 com constante de expansividade ε_0 . Então, há constante positiva K_0 tal que para toda δ -pseudo-órbita $\{x_i\}_{i \in \mathbb{Z}}$ com $K_0\delta < \varepsilon_0$, há um único $x \in M$ tal que $d(T^i x, x_i) < \varepsilon$, onde $\varepsilon = K_0\delta$.*

Caro leitor, repare que o lema acima fornece um forte corolário quando consideramos uma δ -pseudo-órbita $\{x_i\}_{i=0}^{n-1}$ tal que $d(Tx_{n-1}, x_0) < \delta$, com $\delta K_0 < \varepsilon_0$. Ao aplicar o

lema para a δ -pseudo-órbita dada pelos pontos $y_j = x_{j \pmod n}$, concluímos que a órbita verdadeira que ε -sombreia $\{y_j\}$ é, na verdade, uma órbita periódica de período n .

Usamos esse lema de sombreamento como ferramenta para construir coberturas markovianas em M . A partir de uma cobertura \mathcal{U} com diâmetro pequeno (podemos supor finita, uma vez que M é compacta), construímos sua cobertura markoviana associada.

Lema 3.1.8. *Dada cobertura de M por abertos $\mathcal{U} = \{U_1, \dots, U_l\}$ com $\text{diam}(U_i) < \varepsilon_0/K_0^2$, existe uma cobertura markoviana associada $\mathcal{R} = \{R_1, \dots, R_l\}$ tal que $\text{diam}(R_i) < \varepsilon_0/K_0$.*

Demonstração. Considere o *subshift* bilateral de tipo finito, indexado por todas as transições possíveis do tipo $U_i \cap T^{-1}(U_j)$:

$$\Sigma := \{w = (\dots, w_{-2}, w_{-1} | w_0, w_1, \dots) : U_{w_n} \cap T^{-1}(U_{w_{n+1}}) \neq \emptyset, \forall n \in \mathbb{Z}\}.$$

Dado $w \in \Sigma$, para cada $n \in \mathbb{Z}$, escolhemos um ponto $x_n \in U_{w_n}$ tal que $Tx_n \in U_{w_{n+1}}$. Pela restrição no diâmetro dos abertos U_i , $\{x_n\}$ é uma ε_0/K_0^2 -pseudo-órbita. Logo, o lema (3.1.7) fornece uma única órbita verdadeira $\{T^n x\}$, a qual satisfaz $d(T^n x, U_{w_n}) < \varepsilon_0/K_0$. Denotamos tal ponto x por $\Pi(w)$. Sendo assim, o mapa $\Pi: \Sigma \rightarrow M$ é sobrejetivo, contínuo e comuta com o *shift* ($T \circ \Pi = \Pi \circ \sigma$), ou seja, Π é uma semiconjugação entre T e σ .

É fácil se convencer de que os cilindros $[i]_0 := \{w \in \Sigma : w_0 = i\}$ formam uma cobertura markoviana de Σ por retângulos: para cada $w \in \Sigma$, os conjuntos estável e instável de w são, respectivamente,

$$W^s(w) = \{w' \in \Sigma : w'_i = w_i, \forall i \geq 0\} \quad \text{e} \quad W^u(w) = \{w' \in \Sigma : w'_i = w_i, \forall i \leq 0\}.$$

Logo, para $w, w' \in [i]_0$, o ponto $[w, w']$ vem a ser $(\dots, w'_{-2}, w'_{-1} | i, w_1, w_2, \dots) \in [i]_0$. Note que as iteradas positivas $\sigma^n([w, w'])$ aproximam-se do ponto w , enquanto as negativas $\sigma^{-n}([w, w'])$ aproximam-se de w' . As demais propriedades de cobertura markoviana também são facilmente verificadas.

Sendo assim, definimos os retângulos em M como $R_i := \Pi([i]_0)$. É claro que $U_i \subseteq R_i$, pois cada $x \in U_i$ corresponde a uma órbita verdadeira representada por $w \in \Sigma$, com $w_0 = i$. Além disso, pela propriedade de sombreamento, $\text{diam}(R_i) < \varepsilon_0/K_0$. Resta apenas mostrar que

$$x \in U_i \cap T^{-1}(U_j) \quad \Rightarrow \quad T(W^s(x, R_i)) \subseteq W^s(Tx, R_j),$$

uma vez que a outra inclusão se calcula de maneira análoga. Se $x = \Pi(w)$, para

$w = (\dots, w_{-1}|i, j, w_2, \dots)$, pela definição dos conjuntos estável e instável em Σ , temos que $\sigma(W^s(w, [i]_0)) \subseteq W^s(\sigma(w), [j]_0)$. Portanto, ao compor esta inclusão com a aplicação Π e usar a conjugação $\Pi \circ \sigma = T \circ \Pi$, obtemos

$$\begin{aligned} \Pi \circ \sigma (W^s(w, [i]_0)) &\subseteq \Pi (W^s(\sigma(w), [j]_0)) \\ T \circ \Pi (W^s(w, [i]_0)) &\subseteq W^s(\Pi \circ \sigma(w), R_j) \\ T (W^s(\Pi(w), R_i)) &\subseteq W^s(T \circ \Pi(w), R_j) \\ T (W^s(x, R_i)) &\subseteq W^s(Tx, R_j). \end{aligned}$$

□

O lema acima fornece, além da cobertura markoviana de M , uma semiconjugação entre o difeomorfismo de Anosov e um *subshift* de tipo finito bilateral. Esta construção é peça chave da construção de subação para potencial holderiano, pois esta será feita sobre cada retângulo R_i . Considere $f: M \rightarrow \mathbb{R}$ um potencial α -Hölder. Como de costume, denotamos a constante de Hölder de f por

$$\text{Hold}_\alpha(f) := \sup_{x, y \in M} \frac{|f(x) - f(y)|}{d(x, y)^\alpha}.$$

Definimos a seguir o objeto candidato à subação para f .

Definição 3.1.9. *Sejam $\mathcal{R} = \{R_1, \dots, R_l\}$ uma partição markoviana de M por retângulos e $\{\theta_1, \dots, \theta_l\}$ uma partição da unidade associada à \mathcal{R} . Definimos sobre cada R_i a função a seguir:*

$$V_i(x) := \inf_{\substack{n \geq 0 \\ y \in W^s(x, R_i)}} \{S_n(\beta(f) - f)(T^{-n}y) + \Delta^s(x, y)\},$$

onde $\beta(f)$ é o valor máximo da integral de f assim como definida em (1.2.1) e

$$\Delta^s(x, y) := \sum_{k \geq 0} (f \circ T^k(x) - f \circ T^k(y)).$$

Por fim, definimos sobre M a colagem das funções $V_i(x)$

$$V(x) = \sum_{i=1}^l \theta_i(x) V_i(x). \quad (3.1)$$

Demonstramos abaixo que a função $V(x)$ logo acima definida corresponde a uma subação para o potencial f .

Teorema 3.1.10. *Sejam M uma variedade compacta sem fronteira, de dimensão maior ou igual a 2 e $T: M \rightarrow M$ um difeomorfismo de Anosov transitivo de classe C^2 . Então, para todo potencial α_1 -holderiano $f: M \rightarrow \mathbb{R}$, a função $V(x)$ definida como acima (3.1) é uma subação para f , de regularidade α_2 -holderiana, tal que*

$$\alpha_2 = \alpha_1 \frac{\ln(1/\gamma_s)}{\ln(\Gamma_u/\gamma_s)} \quad e \quad \text{Hold}_{\alpha_2}(V) \leq \text{Hold}_{\alpha_1}(f) \frac{C(M)}{\min\{1 - \gamma_u^{-\alpha_1}, 1 - \gamma_s^{\alpha_1}\}^2}.$$

Demonstração. Assumimos inicialmente que cada função V_i é α_2 -holderiana em R_i . Considere, então, $x \in U_i \cap T^{-1}(U_j)$ e $y \in W^s(x, R_i)$. Temos que

$$\begin{aligned} S_n(\beta(f) - f)(T^{-n}y) + \Delta^s(x, y) + (\beta(f) - f)(x) &= \\ &= \sum_{k=0}^{n-1} (\beta(f) - f) \circ T^{k-n}(y) + \sum_{k \geq 0} (f \circ T^k(x) - f \circ T^k(y)) + (\beta(f) - f)(x) \\ &= \sum_{k=0}^n (\beta(f) - f) \circ T^{k-n-1}(Ty) + \sum_{k \geq 0} (f \circ T^k(Tx) - f \circ T^k(Ty)) \\ &= S_{n+1}(\beta(f) - f) \circ T^{-(n+1)}(Ty) + \Delta^s(T(x), T(y)) \\ &\geq V_j \circ T(x). \end{aligned}$$

Ressaltamos algumas informações: como $y \in W^s(x, R_i)$ e $Tx \in U_j$, então, $Ty \in W^s(Tx, R_j)$. Além disso, lembre que V_j é definido através de um ínfimo, donde decorre a desigualdade acima. Mas, tomando o ínfimo também na primeira linha da desigualdade, obtemos

$$V_i(x) + (\beta(f) - f)(x) \geq V_j(x), \quad \forall x \in U_i \cap T^{-1}(U_j). \quad (3.2)$$

Por fim, note que, ao multiplicarmos a desigualdade (3.2) por $\theta_i(x)\theta_j(Tx)$ e fazer a soma em i e depois em j , obtemos:

$$V(x) + (\beta(f) - f)(x) \geq V \circ T(x),$$

ou ainda,

$$\beta(f) \geq f(x) + V \circ T(x) - V(x),$$

de modo que V é uma subação para f . A prova de que cada função V_i é α_2 -holderiana

sobre R_i é feita no lema a seguir. □

Lema 3.1.11. *Sejam $f: M \rightarrow \mathbb{R}$, uma função α_1 -holderiana assim como no teorema anterior, R um retângulo de M e $V: R \rightarrow \mathbb{R}$ a função definida por*

$$V(x) = \inf_{\substack{n \geq 0 \\ y \in W^s(x, R)}} \{S_n(\beta(f) - f)(T^{-n}y) + \Delta^s(x, y)\}.$$

Então, V é função α_2 -holderiana em R , com constante $Hold_{\alpha_2}(V)$ tal que

$$\alpha_2 = \alpha_1 \frac{\ln(1/\gamma_s)}{\ln(\Gamma_u/\gamma_s)} \quad e \quad Hold_{\alpha_2}(V) \leq Hold_{\alpha_1}(f) \frac{C(M)}{\min\{1 - \gamma_u^{-\alpha_1}, 1 - \gamma_s^{\alpha_1}\}^2}.$$

Demonstração. Lembramos que ε_0 denota a constante de expansividade da aplicação T , e K_0 é constante obtida no lema de sombreamento, a qual relaciona a δ -pseudo-órbita com o ε sombreamento. Note ainda que, análoga à definição de $\Delta^s(x, y)$, podemos definir $\Delta^u(x, y) := \sum_{k \leq 0} (f \circ T^k(x) - f \circ T^k(y))$, quando x e y estão em uma mesma folha instável. Sendo assim, é fácil ver que se $x, x' \in R$ estão em uma mesma folha estável e cumprem $d(x, x') < \varepsilon_0$, então,

$$\Delta^s(x, x') \leq \sum_{k \geq 0} |f \circ T^k(x) - f \circ T^k(x')| \leq Hold_{\alpha}(f) \sum_{k \geq 0} d(T^k x, T^k x')^{\alpha}.$$

Mas o fato de x e x' pertencerem a uma mesma folha estável implica haver $C_1 > 0$ tal que $d(T^k x, T^k x') \leq C_1^{1/\alpha_1} \gamma_s^k d(x, x')$. Logo, concluímos que

$$\Delta^s(x, x') \leq C_1 Hold_{\alpha_1}(f) d(x, x')^{\alpha_1} \sum_{k \geq 0} (\gamma_s^{\alpha_1})^k = C_1 \frac{Hold_{\alpha_1}(f)}{1 - \gamma_s^{\alpha_1}} d(x, x')^{\alpha_1}. \quad (3.3)$$

De maneira análoga, podemos concluir que

$$\Delta^u(x, x') \leq C_1 \frac{Hold_{\alpha_1}(f)}{1 - \gamma_u^{-\alpha_1}} d(x, x')^{\alpha_1}, \quad (3.4)$$

quando x e x' estão em uma mesma folha instável e $d(x, x') < \varepsilon_0$. Por outro lado, se há $n \geq 1$, tal que os pontos $x, x' \in R$ satisfazem $d(T^k x, T^k x') < \varepsilon_0/K_0$, para todo $0 \leq k \leq n$, então há uma constante $C_3 > 0$ tal que

$$\sum_{k=0}^n |f \circ T^k(x) - f \circ T^k(x')| \leq C_3 \max\{d(x, x')^{\alpha_1}, d(T^n x, T^n x')^{\alpha_1}\}. \quad (3.5)$$

Para isto ver, definimos $y = [x, x']$ (lembrando, $y = W_{\varepsilon_0}^s(x) \cap W_{\varepsilon_0}^u(x')$). Aplicando o lema de sombreamento de Bowen para a ε_0/K_0 -pseudo-órbita

$$x_i = \begin{cases} T^i x, & i \geq 0 \\ T^i x', & i < 0 \end{cases},$$

obtemos um ponto $y' \in M$ tal que $d(T^i y', T^i x) \leq K_0 d(x, x') < \varepsilon_0, \forall i \geq 0$. Com isto resulta que $y' \in W^s(x, R)$. Porém, y' também verifica $d(T^i y', T^i x') \leq K_0 d(x, x') < \varepsilon_0, \forall i < 0$, o que implica que $Ty' \in W^u(Tx', R)$, o que, por sua vez, implica que $y' \in W^u(x', R)$. Portanto, pela unicidade de $[x, x']$, segue que $y' = y$. Disso, concluímos que $d(x, y) \leq k_0 d(x, x')$. Um raciocínio análogo (recordando que $y = [x, x'] \Rightarrow T^n y = [T^n x, T^n x']$) nos leva a concluir que $d(T^n y, T^n x') \leq K_0 d(T^n x, T^n x')$. Dessas majorações, temos

$$\begin{aligned} \sum_{k=0}^n |f \circ T^k(x) - f \circ T^k(x')| &\leq \sum_{k=0}^n |f \circ T^k(x) - f \circ T^k(y)| + \sum_{k=0}^n |f \circ T^k(y) - f \circ T^k(x')| \\ &\leq \sum_{k \geq 0} |f \circ T^k(x) - f \circ T^k(y)| + \sum_{k \leq 0} |f \circ T^k(T^n y) - f \circ T^k(T^n x')| \\ &\leq C_1 \text{Hold}_{\alpha_1}(f) \left(\frac{d(x, y)^{\alpha_1}}{1 - \gamma_s^{\alpha_1}} + \frac{d(T^n y, T^n x')^{\alpha_1}}{1 - \gamma_u^{-\alpha_1}} \right) \\ &\leq 2C_1 \text{Hold}_{\alpha_1}(f) K_0^{\alpha_1} \frac{\max\{d(x, x')^{\alpha_1}, d(T^n x, T^n x')^{\alpha_1}\}}{\min\{1 - \gamma_s^{\alpha_1}, 1 - \gamma_u^{-\alpha_1}\}}. \end{aligned}$$

A constante K_0 está diretamente relacionada com a máxima dilatação de T , de modo que podemos estimá-la em

$$K_0 = \frac{C_0(M)}{\min\{1 - \gamma_u^{-1}, 1 - \gamma_s\}},$$

resultando na seguinte majoração

$$\sum_{k=0}^n |f \circ T^k(x) - f \circ T^k(x')| \leq \frac{2C_0(M)^{\alpha_1} C_1 \text{Hold}_{\alpha_1}(f) \max\{d(x, x')^{\alpha_1}, d(T^n x, T^n x')^{\alpha_1}\}}{\min\{1 - \gamma_u^{-1}, 1 - \gamma_s\} \min\{1 - \gamma_s^{\alpha_1}, 1 - \gamma_u^{-\alpha_1}\}}.$$

Escrevemos $C_2 = 2C_0(M)^{\alpha_1} C_1$ e minoramos o denominador da seguinte maneira: se $a \in (0, 1)$, então, $(1 - a)^{-\alpha_1} < (1 - a)^{-1} < (1 - a^{\alpha_1})^{-1}$, de modo a obter

$$\sum_{k=0}^n |f \circ T^k(x) - f \circ T^k(x')| \leq \frac{C_2 \text{Hold}_{\alpha_1}(f)}{\min\{1 - \gamma_s^{\alpha_1}, 1 - \gamma_u^{-\alpha_1}\}^2} \max\{d(x, x')^{\alpha_1}, d(T^n x, T^n x')^{\alpha_1}\}.$$

Denotamos a constante do lado direito da desigualdade por C_3 .

Com as estimativas feitas até agora é possível mostrar que $V(x) > -\infty$. Para tanto, usamos o fato de que todo difeomorfismo de Anosov é topologicamente *mixing*, o que garante a existência de um inteiro $\tau_0 \geq 1$ tal que, para todo $n \in \mathbb{N}$ e todo trecho de órbita $\{T^{-n}y, \dots, y\}$, ocorre

$$T^{-n}y \in T^{\tau_0} \left(B_{\varepsilon_0/K_0}(y) \right).$$

Logo, por consequência do lema 3.1.7 aplicado ao trecho de órbita $\{T^{-n}y, \dots, y\}$, para cada $n \in \mathbb{N}$ há uma órbita periódica $\{T^k z\}_{k=0}^{\tau_0+n-1}$, de período $\tau_0 + n$, satisfazendo

$$d(T^{-k}y, T^{-k}z) \leq \varepsilon_0, \quad \forall k = 0, \dots, n.$$

Sendo assim, a estimativa (3.5) garante que $\sum_{k=0}^n (f \circ T^k(y) - f \circ T^k(z))$ é uniformemente limitada por alguma constante C_4 . Além disso, como $\beta(f) \geq \frac{1}{n + \tau_0} \sum_{k=1}^{n+\tau_0} f \circ T^k(z)$, obtemos a seguinte majoração

$$\begin{aligned} \sum_{k=1}^n f \circ T^{-k}(y) &\leq C_4 + \sum_{k=1}^n f \circ T^{-k}(z) + \sum_{k=n+1}^{\tau_0+n} f \circ T^{-k}(z) + \tau_0 \|f\|_\infty \\ &\leq C_4 + (n + \tau_0)\beta(f) + \tau_0 \|A\|_\infty \end{aligned}$$

donde concluímos que

$$S_n(\beta(f) - f)(T^{-n}y) \geq -C_4 - 2\tau_0 \|A\|_\infty. \quad (3.6)$$

Utilizando esta minoração e as estimativas (3.3) e (3.4), podemos concluir que $V(x)$ toma valores reais, $\forall x \in R$. Para completar a prova, mostremos que V é função α_2 -holderiana. Com efeito, considere pontos $x, x' \in R$, $y \in W^s(x, R)$ e $y' = [x', y]$. Seja $N \in \mathbb{N}$ um valor a ser escolhido. Temos que

$$S_n(\beta(f) - f)(T^{-n}y) + \Delta^s(x, y) \leq S_n(\beta(f) - f)(T^{-n}y') + \Delta(x', y') + \sum_{k=-n}^N |f \circ T(y') - f \circ T(y)| \quad (3.7)$$

$$+ \sum_{k=0}^N |f \circ T(x) - f \circ T(x')| \quad (3.8)$$

$$+ |\Delta^s(T^N x, T^N y)| \quad (3.9)$$

$$+ |\Delta^s(T^N x', T^N y')|. \quad (3.10)$$

Por (3.4), estimamos

$$\sum_{k=-n}^N |f \circ T(y') - f \circ T(y)| \leq \frac{C_1 \text{Hold}_{\alpha_1}(f)}{1 - \gamma_u^{-\alpha_1}} d(T^N y, T^N y')^{\alpha_1}.$$

Por (3.5), segue que

$$\sum_{k=0}^N |f \circ T(x) - f \circ T(x')| \leq \frac{C_2 \text{Hold}_{\alpha_1}(f)}{\min\{1 - \gamma_s^{\alpha_1}, 1 - \gamma_u^{-\alpha_1}\}^2} \max\{d(x, x')^{\alpha_1}, d(T^N x, T^N x')^{\alpha_1}\}.$$

E, por (3.3), resulta

$$|\Delta^s(T^N x, T^N y)| \leq \frac{C_1 \text{Hold}_{\alpha_1}(f)}{1 - \gamma_s^{\alpha_1}} d(T^N y, T^N x)^{\alpha_1} \quad \text{e}$$

$$|\Delta^s(T^N x', T^N y')| \leq \frac{C_1 \text{Hold}_{\alpha_1}(f)}{1 - \gamma_s^{\alpha_1}} d(T^N y', T^N x')^{\alpha_1}.$$

A conclusão da prova depende de uma boa escolha de N . Seja $t \in \mathbb{R}$ o real que produz a igualdade

$$\gamma_s^t \varepsilon_0 = \Gamma_u^t d(x, x').$$

Definimos $N = N(x, x') := \lfloor t \rfloor + 1$, o menor inteiro maior que t . Repare que, como $\Gamma_u > 1 > \gamma_s$,

$$\varepsilon_0 \leq \left(\frac{\Gamma_u}{\gamma_s} \right)^N d(x, x').$$

Seguem, então, as desigualdades abaixo:

$$\begin{aligned} d(T^N x, T^N x') &\leq C_0(M) \Gamma_u^N d(x, x') \\ d(T^N x, T^N y) &\leq C_0(M) \gamma_s^N \varepsilon_0 \leq C_0(M) \Gamma_u^N d(x, x') \\ d(T^N x', T^N y') &\leq C_0(M) \gamma_s^N \varepsilon_0 \leq C_0(M) \Gamma_u^N d(x, x'), \end{aligned}$$

sendo que a primeira desigualdade é devido à dilatação máxima possível por T , e as demais pela contração mínima possível. Assim, pela desigualdade triangular,

$$d(T^N y, T^N y') \leq 3C_0(M) \Gamma_u^N d(x, x').$$

Temos agora todos os majorantes necessários para (3.7), (3.8), (3.9) e (3.10), concluindo que

$$V(x) \leq V(x') + K(M, T, f) \left(\Gamma_u^N d(x, x') \right)^{\alpha_1}, \quad (3.11)$$

onde

$$K(M, T, f) = \frac{C(M) \text{Hold}_{\alpha_1}(f)}{\min\{1 - \gamma_s^{\alpha_1}, 1 - \gamma_u^{-\alpha_1}\}^2}, \quad \text{e,}$$

$$C(M) = C_0(M)^{\alpha_1} \left[\frac{3^{\alpha_1} C_1 \min\{1 - \gamma_s^{\alpha_1}, 1 - \gamma_u^{-\alpha_1}\}^2}{1 - \gamma_u^{-\alpha_1}} + C_2 + \frac{2C_1 \min\{1 - \gamma_s^{\alpha_1}, 1 - \gamma_u^{-\alpha_1}\}^2}{1 - \gamma_s^{\alpha_1}} \right].$$

Ainda falta um detalhe, que é relacionar Γ_u^N com a distância $d(x, x')$. Note que, *a priori*, não temos garantia de que $\Gamma_u^{N\alpha_1} d(x, x')^{\alpha_1}$ seja limitado. Suponha que exista $\eta > 0$, que satisfaz

$$\Gamma_u^{N\alpha_1} d(x, x')^\eta \leq 1.$$

Ao aplicar a função logarítmica \ln , em ambos os lados da desigualdade, devemos ter

$$N \leq -\frac{\eta \ln(d(x, x'))}{\alpha_1 \ln(\Gamma_u)}. \quad (3.12)$$

Por outro lado, como $\varepsilon_0 \gamma_s^N \leq \Gamma_u^N d(x, x')$, temos uma outra desigualdade

$$\frac{\ln(\varepsilon_0) - \ln(d(x, x'))}{\ln(\Gamma_u) - \ln(\gamma_s)} \leq N. \quad (3.13)$$

Como ε_0 pode ser suposto, desde o início, menor que 1, então, para se obter (3.13), é

suficiente estimar

$$\frac{-\ln(d(x, x'))}{\ln(\Gamma_u) - \ln(\gamma_s)} \leq N. \quad (3.14)$$

Unindo agora (3.12) e (3.14), obtemos

$$\frac{-\ln(d(x, x'))}{\ln(\Gamma_u) - \ln(\gamma_s)} \leq -\frac{\eta}{\alpha_1} \frac{\ln(d(x, x'))}{\ln(\Gamma_u)}. \quad (3.15)$$

Logo, como $\ln(d(x, x')) < 0$, segue que

$$\eta \geq \alpha_1 \frac{\ln(\Gamma_u)}{\ln(\Gamma_u/\gamma_s)}.$$

Considerando η o menor possível, que é atingido na igualdade acima, resulta que

$$(\Gamma_u^N d(x, x'))^{\alpha_1} \leq d(x, x')^{\alpha_1 \left[1 - \frac{\ln(\Gamma_u)}{\ln(\Gamma_u/\gamma_s)}\right]} = d(x, x')^{\alpha_1 \frac{\ln(1/\gamma_s)}{\ln(\Gamma_u/\gamma_s)}},$$

donde obtemos que $V(x)$ é função α_2 -holderiana com

$$\alpha_2 = \alpha_1 \frac{\ln(1/\gamma_s)}{\ln(\Gamma_u/\gamma_s)}.$$

□

Apesar de este teorema ter uma demonstração bastante longa, ele fornece não simplesmente a existência de uma subação para o potencial f mas também seu método de construção. Contudo, é preciso notar que a subação possui uma perda de regularidade em relação ao potencial. Adiantamos que isto não será um problema, uma vez que, com um método bastante similar a este, porém em um contexto mais amplo, será possível construir subação de igual regularidade que o potencial em questão.

Recomendamos fortemente ao leitor que teve seu primeiro contato com os difeomorfismos de Anosov nesta dissertação para que consulte os trabalhos de Contreras [13] e Pollicott e Yuri [29]. Neles são estudados resultados para automorfismos lineares do toro. Este vem a ser o exemplo mais simples e ilustrativo de difeomorfismo de Anosov sobre uma variedade compacta. Os automorfismos lineares do toro são induzidos de aplicações $A: \mathbb{R}^2 \rightarrow \mathbb{R}^2$, onde A é uma matriz pertencente a $GL(2, \mathbb{Z})$, de determinante $\det(A) = 1$ e autovalores de módulo não unitário. Mais ainda, segundo os clássicos resultados de Franks [17] e Manning [23], todo difeomorfismo de Anosov em um toro \mathbb{T}^n , $n \geq 2$, é conjugado ao automorfismo

linear induzido pela parte linear de sua diferencial. Logo, todo o estudo de otimização ergódica para difeomorfismos de Anosov em toros pode ser simplificado ao estudo dos automorfismos lineares.

3.2 Sistemas anfidinâmicos

Apresentamos nesta seção o conceito bastante inovador denominado sistema anfidinâmico, o qual foi introduzido por T. Bousch em [6] e generaliza os sistemas dinâmicos topológicos. Mostramos como construir subação para potenciais lipschitzianos (podendo também considerá-los α -holderianos) no contexto de um sistema anfidinâmico. A partir desta construção, obtemos subação para potenciais lipschitzianos associados a um difeomorfismo de Anosov. Mais uma vez relembramos que a passagem do estudo de um caso α -holderiano para um lipschitziano pode ser realizada considerando uma pequena alteração na distância.

Definição 3.2.1. *Sejam X_0, X_1 dois espaços métricos compactos e $T_0, T_1: X_1 \rightarrow X_0$ duas aplicações contínuas. Um sistema anfidinâmico é definido pela dinâmica conjunta das aplicações T_0 e T_1 , representadas no diagrama abaixo,*

$$X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0,$$

no qual os trechos de órbita são conjuntos $\{x_1, \dots, x_n\} \subseteq X_1$ tais que $T_1 x_i = T_0 x_{i+1}$, para todo $i = 1, \dots, n - 1$.

Repare que é bem fácil representar um SDT (X, T) por um sistema anfidinâmico. Basta tomar $X_0 = X_1 = X$, $T_0 = Id$ e $T_1 = T$. Obtemos o sistema anfidinâmico $X \xleftarrow{Id} X \xrightarrow{T} X$, no qual os trechos de órbitas são conjuntos $\{x_i\}_{i=1}^n \subseteq X$ tais que $T x_i = x_{i+1}$, os quais são exatamente as órbitas do SDT (X, T) .

Apresentamos aqui as ideias abordadas em [6]. Para facilitar a escrita, em vez de aplicar o raciocínio ao potencial $\beta(f) - f$ de modo a obter uma subação v para f , fazemos o raciocínio para um potencial $u: X_1 \rightarrow \mathbb{R}$, obtendo função $v: X_0 \rightarrow \mathbb{R}$, a qual satisfaz a desigualdade coomológica

$$u(x) \geq v(T_1 x) - v(T_0 x), \quad \forall x \in X_1. \quad (3.16)$$

Note que, na representação de um SDT (X, T) como sistema anfidinâmico, tomando

$u = \beta(f) - f$, a desigualdade acima representa exatamente a condição para v ser subação para f .

A melhor ideia do raciocínio que apresentamos a seguir é considerar não simplesmente o potencial u mas também a distância entre a imagem dos pontos $t \in X_1$ por T_0 e T_1 . Isso fica ilustrado no resultado a seguir.

Lema 3.2.2. *Sejam (X_1, d_1) , (X_0, d_0) dois espaços métricos compactos e $T_0, T_1: X_1 \rightarrow X_0$ duas aplicações contínuas. Considere $u: X_1 \rightarrow \mathbb{R}$ uma função 1-lipschitziana. Então, são equivalentes:*

(i) *para qualquer sequência de pontos $\{t_i\}_{i \in \mathbb{Z}_n}$ (não necessariamente uma órbita de $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$), temos a desigualdade*

$$\sum_{i \in \mathbb{Z}_n} u(t_i) + \sum_{i \in \mathbb{Z}_n} d_0(T_1 t_i, T_0 t_{i+1}) \geq 0; \quad (3.17)$$

(ii) *existe função 1-lipschitziana $v: X_0 \rightarrow \mathbb{R}$ tal que*

$$u(t) \geq v(T_1 t) - v(T_0 t), \quad \forall t \in X_1.$$

Demonstração. (ii) \Rightarrow (i) A volta é bastante direta de ser demonstrada. Considere uma sequência $\{t_i\}_{i \in \mathbb{Z}_n} \subseteq X_1$. Se (ii) é válido, então,

$$\sum_{i \in \mathbb{Z}_n} u(t_i) - \sum_{i \in \mathbb{Z}_n} (v(T_1 t_i) - v(T_0 t_i)) \geq 0.$$

Como a soma é finita, podemos efetuar uma troca de índices, obtendo

$$\sum_{i \in \mathbb{Z}_n} u(t_i) - \sum_{i \in \mathbb{Z}_n} (v(T_1 t_i) - v(T_0 t_{i+1})) \geq 0.$$

Considerando, então, que v é 1-lipschitziana, o que fornece $v(T_1 t_i) - v(T_0 t_{i+1}) \geq -d(T_1 t_i, T_0 t_{i+1})$, resulta que

$$\sum_{i \in \mathbb{Z}_n} u(t_i) - \left(- \sum_{i \in \mathbb{Z}_n} d_0(T_1 t_i, T_0 t_{i+1})\right) \geq 0,$$

demonstrando (i).

(i) \Rightarrow (ii) Nessa implicação construímos a função $v: X_0 \rightarrow \mathbb{R}$. Fixe $p \in X_0$ um ponto

qualquer e defina as funções

$$v_0(x) = d_0(x, p), \quad \forall x \in X_0,$$

$$v_{n+1}(x) = \inf_{z \in X_1} \{v_n(T_0 z) + u(z) + d_0(T_1 z, x)\}, \quad \forall x \in X_0.$$

Além disso, defina $v(x) = \inf_{n \in \mathbb{N}} v_n(x)$ (*a priori*, não sabemos se v é função a valores reais).

Note que, para todo $n \in \mathbb{N}$, v_n é função 1-lipschitziana: dados $x, y \in X_0$

$$\begin{aligned} d_0(x, y) + v_n(x) &= d_0(x, y) + \inf_{z \in X_1} \{v_{n-1}(T_0 z) + u(z) + d_0(T_1 z, x)\} \\ &\leq \inf_{z \in X_1} \{v_{n-1}(T_0 z) + u(z) + d_0(T_1 z, y)\} = v_n(y). \end{aligned}$$

Por indução, provamos que para todo $n \geq 1$ podemos escrever

$$v_n(x) = \inf_{z_1, \dots, z_n \in X_1} \left[d_0(T_0 z_1, p) + \sum_{i=1}^{n-1} d_0(T_1 z_i, T_0 z_{i+1}) + \sum_{i=1}^n u(z_i) + d_0(T_1 z_n, x) \right].$$

Para $n = 1$, a própria definição de v_1 verifica a condição, uma vez que $v_0(z) = d_0(z, p)$.

Supondo, então válida a hipótese para algum $n \geq 1$, temos:

$$\begin{aligned} v_{n+1}(x) &= \inf_{z_{n+1} \in X_0} [v_n(T_0 z_{n+1}) + u(z_{n+1}) + d_0(T_1 z_{n+1}, x)] \\ &= \inf_{z_{n+1} \in X_0} \left[\inf_{z_1, \dots, z_n \in X_1} \left(d_0(T_0 z_1, p) + \sum_{i=1}^{n-1} d_0(T_1 z_i, T_0 z_{i+1}) + \right. \right. \\ &\quad \left. \left. + \sum_{i=1}^n u(z_i) + d_0(T_1 z_n, T_0 z_{n+1}) \right) + u(z_{n+1}) + d_0(T_1 z_{n+1}, x) \right] \\ &= \inf_{z_1, \dots, z_{n+1} \in X_1} \left[d_0(T_0 z_1, p) + \sum_{i=1}^n d_0(T_1 z_i, T_0 z_{i+1}) + \sum_{i=1}^{n+1} u(z_i) + d_0(T_1 z_{n+1}, x) \right]. \end{aligned}$$

A igualdade acima é exatamente a tese da indução para $n + 1$. Em particular, segue da hipótese (3.17) e das estimativas acima que $v_n(p) \geq 0$, $\forall n \in \mathbb{N}$, e, portanto, também temos $v(p) \geq 0$. Sendo assim, $v(x)$ é ínfimo de funções 1-lipschitzianas que possuem um ponto em comum no domínio tal que seu valor é uniformemente limitado. Isto implica que v também é função 1-lipschitziana. Logo,

$$0 \leq v(p) \leq v(x) + d(x, p),$$

donde segue que $v(x)$ toma valores reais para todo $x \in X_0$.

Para concluir, seja $z \in X_1$ um ponto qualquer. Então, para todo $x \in X_0$ ocorre

$$v(x) \leq v_{n+1}(x) \leq v_n(T_0z) + u(z) + d_0(T_1z, x).$$

Tomando ínfimo em n no lado direito da desigualdade acima, resulta que

$$v(x) \leq v(T_0z) + u(z) + d_0(T_1z, x).$$

Como isto vale para qualquer $x \in X_0$, escolhemos $x = T_1z$, resultando em

$$v(T_1z) - v(T_0z) \leq u(z), \quad \forall z \in X_1.$$

□

Este lema nos diz que a hipótese (3.17), até então um pouco mística, é suficiente para a construção de uma subação. Esta hipótese, traduzida para um SDT seria

$$\sum_{i \in \mathbb{Z}_n} u(x_i) + \sum_{i \in \mathbb{Z}_n} d(Tx_i, x_{i+1}) \geq 0,$$

e isto parece requerer muito mais que simplesmente a desigualdade acima sobre órbitas periódicas, que é a hipótese usual encontrada na literatura. Mostraremos que, em nosso contexto de difeomorfismos de Anosov, esta hipótese surge de forma natural.

Assim como no outro processo de obtenção de subação, usamos um lema de sombreamento, o qual requer algumas hipóteses adicionais sobre o sistema anfidinâmico (lembre que foi solicitado anteriormente que o difeomorfismo de Anosov seja transitivo e de classe C^2). Fazemos então as definições necessárias.

Definição 3.2.3. *Seja $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$ um sistema anfidinâmico. Para $L \in (0, \infty]$, dizemos que o sistema anfidinâmico é L -retificável se verifica a seguinte propriedade: para quaisquer $x, y \in X_1$ tais que $d_0(T_1x, T_0y) < L$, existem $x', y' \in X_1$ com $T_1x' = T_0y'$ e*

$$d_1(x, x') + d_1(y, y') + d_0(T_0x, T_0x') + d_0(T_1y, T_1y') \leq d_0(T_1x, T_0y). \quad (3.18)$$

Observação. A todo momento estamos analisando as condições impostas para um sistema anfidinâmico, considerando um SDT representado como tal. Então, nesse caso, a definição

de L -retificabilidade traduz-se da seguinte maneira: o sistema anfidinâmico $X \xleftarrow{d} X \xrightarrow{T} X$ é L -retificável se verificar que para quaisquer $x, y \in X$, com $d(Tx, y) < L$, exista $y' \in X$ e $x' \in T^{-1}y$ tais que

$$2d(x, x') + d(y, y') + d(Ty, Ty') \leq d(Tx, y). \quad (3.19)$$

A hipótese de retificabilidade é parte fundamental para se obter o seguinte lema de sombreamento.

Lema 3.2.4. *Seja $X_0 \xrightarrow{T_0} X_1 \xrightarrow{T_1} X_0$ um sistema anfidinâmico L -retificável para algum $L \in (0, \infty]$. Então, para todo $n \geq 2$ e toda sequência $\{x_i\}_{i \in \mathbb{Z}_n} \subseteq X_1$, há sequência $\{y_i\}_{i \in \mathbb{Z}_n} \subseteq X_1$ tal que*

$$\sum_{i \in \mathbb{Z}_n} d_1(x_i, y_i) + \sum_{i \in \mathbb{Z}_n} d_0(T_1 y_i, T_0 y_{i+1}) \leq \sum_{i \in \mathbb{Z}_n} d_0(T_1 x_1, T_0 x_{i+1}), \quad (3.20)$$

onde ou $T_1 y_i = T_0 y_{i+1}$ (formando uma órbita periódica), ou $d_0(T_1 y_i, T_0 y_{i+1}) \geq L$.

Repare, caro leitor, que a afirmação final do lema garante que se $\sum d_0(T_1 x_1, T_0 x_{i+1}) < L$, então a sequência $\{y_i\}_{i \in \mathbb{Z}_n}$ será de fato uma órbita periódica.

Demonstração. Considere o conjunto $X := (X_1)^{\mathbb{Z}_n}$, munido da distância de Manhattan

$$d(s, t) = \sum_{i \in \mathbb{Z}_n} d_1(s_i, t_i)$$

e uma função $\Phi: X \rightarrow \mathbb{R}$, definida por

$$\Phi(t) = \sum_{i \in \mathbb{Z}_n} d_0(T_1 t_i, T_0 t_{i+1}).$$

Seja $x = (x_i) \in X$ e considere $\Psi: X \rightarrow \mathbb{R}$ a função que leva $t \in X$ em $\Phi(t) + d(x, t)$. Então, o conjunto $\{t \in X : \Psi(t) = \inf \Psi\}$ é um compacto não vazio de X , sobre o qual a restrição da função Φ atinge seu mínimo em um ponto y . Logo, podemos dizer que para qualquer $z \in X$ temos

$$\Psi(z) > \Psi(y) \quad \text{ou} \quad \Psi(z) = \Psi(y) \text{ e } \Phi(z) \geq \Phi(y). \quad (3.21)$$

Em particular, a desigualdade (3.20) é exatamente o mesmo que escrever $\Psi(x) \geq \Psi(y)$.

Encontrado o ponto $y = (y_i)$ que realiza o sombreamento, falta mostrar que $T_1 y_i = T_0 y_{i+1}$ ou, então, que $d_0(T_1 y_i, T_0 y_{i+1}) \geq L$. Suponha que há índice $i \in \mathbb{Z}_n$ tal que $0 < d(T_1 y_i, T_0 y_{i+1}) < L$. Como o sistema anfidinâmico é L -retificável, há $w, z \in X_1$ tais que $T_1 w = T_0 z$ e

$$d_1(y_i, w) + d_1(y_{i+1}, z) + d_0(T_1 y_{i+1}, T_1 z) + d_0(T_0 y_i, T_0 w) \leq d_0(T_1 y_i, T_0 y_{i+1}). \quad (3.22)$$

Definimos o ponto $t \in X$ da seguinte maneira:

$$t_j = \begin{cases} y_j, & j \notin \{i, i+1\} \\ w, & j = i \\ z, & j = i+1 \end{cases}.$$

Por definição, $d(y, t) = d_1(y_1, w) + d_1(y_{i+1}, z) > 0$. Também da definição decorre que

$$\begin{aligned} \Phi(y) - \Phi(t) &= \begin{pmatrix} d_0(T_1 y_{i-1}, T_0 y_i) + d_0(T_1 y_i, T_0 y_{i+1}) + d_0(T_1 y_{i+1}, T_0 y_{i+2}) \\ -d_0(T_1 y_{i-1}, T_0 w) - d_0(T_1 w, T_0 z) - d_0(T_1 z, T_0 y_{i+2}) \end{pmatrix} \\ &\geq d_0(T_1 y_i, T_0 y_{i+1}) - d_0(T_1 y_{i+1}, T_1 z) - d_0(T_0 y_i, T_0 w). \end{aligned}$$

Combinando agora a desigualdade acima com (3.22), segue que

$$\Phi(y) - \Phi(t) \geq d_1(y_i, w) + d_1(y_{i+1}, z) = d(y, t).$$

Utilizando na desigualdade acima que $d(y, t) \geq d(x, t) - d(x, y)$, obtemos $\Psi(y) \geq \Psi(t)$. Por outro lado, $d(y, t) > 0$ implica que $\Phi(y) > \Phi(t)$. Mas estas duas conclusões simultaneamente contradizem (3.21), ou seja, não pode haver índice $i \in \mathbb{Z}_n$ com $0 < d_0(T_1 y_i, T_0 y_{i+1}) < L$. \square

Mais uma vez fazendo menção aos SDTs, note que se $X \xleftarrow{Ld} X \xrightarrow{T} X$ é ε -retificável e $\{x_i\}_{i \in \mathbb{Z}_n}$ é uma δ -pseudo-órbita, com $\delta < \frac{\varepsilon}{n}$, então, ao aplicar o lema acima, obtemos um trecho de órbita verdadeiro $\{T^i y\}_{i \in \mathbb{Z}_n}$ que ε -sombreia a δ -pseudo-órbita $\{x_i\}_{i \in \mathbb{Z}_n}$. Mais ainda, o sombreamento é um pouco mais forte do que aquele visto anteriormente, pois a diferença de ε é obtida no total da soma das distâncias $d(T^i y, x_i)$ e não somente a cada comparação.

Vejamos, então, como é garantida a existência de subação para potenciais lipschitzianos em um sistema anfidinâmico L -retificável.

Teorema 3.2.5. *Seja $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$ um sistema anfidinâmico L -retificável para algum $L \in (0, \infty]$. Considere a uma função 1-lipschitziana $u: X_1 \rightarrow \mathbb{R}$ que verifica para todo $n \geq 1$ e toda órbita periódica $\{x_i\}_{i \in \mathbb{Z}_n}$, de comprimento n , a desigualdade*

$$\sum_{i \in \mathbb{Z}_n} u(x_i) \geq 0. \quad (3.23)$$

Então, existe uma função C -lipschitziana $v: X_0 \rightarrow \mathbb{R}$ tal que $u(x) \geq v \circ T_1(x) - v \circ T_0(x)$, com

$$C = (2N - 1) \max(1, \sup(-L/u)), \quad (3.24)$$

onde N é a cardinalidade máxima de um subconjunto de X_0 no qual todos os seus elementos distam de pelo menos L entre si.

Antes de provar este teorema, simplifiquemos um pouco a notação. Escreveremos $B = \max(1, \sup(-u/L))$. Então, é óbvio que $C = (2N - 1)B \geq B \geq 1$ e, além disso, $u \geq -BL$. Provemos, ainda, um resultado auxiliar.

Lema 3.2.6. *Sejam $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$ um sistema anfidinâmico e $u: X_1 \rightarrow \mathbb{R}$ uma função nas mesmas condições do teorema 3.2.5. Dados os pontos $t_0, t_1, \dots, t_{n-1} \in X_1$, com $n \geq 2$, verificando $T_1 t_i = T_0 t_{i+1}$, $\forall i = 0, \dots, n-2$, temos o seguinte:*

(i) se $d_0(T_1 t_{n-1}, T_0 t_0) < L$, então,

$$\sum_{i \in \mathbb{Z}_n} u(t_i) \geq -d_0(T_1 t_{n-1}, T_0 t_0) > -L \geq -BL; \quad (3.25)$$

(ii) caso contrário,

$$\sum_{i \in \mathbb{Z}_n} u(t_i) \geq -CL. \quad (3.26)$$

Demonstração. (i) Como $d_0(T_1 t_{n-1}, T_0 t_0) < L$, pelo lema (3.2.4) há órbita periódica $\{z_i\}_{i \in \mathbb{Z}_n}$ tal que

$$\sum_{i \in \mathbb{Z}_n} d_1(t_i, z_i) \leq d_0(T_1 t_{n-1}, T_0 t_0).$$

Então, utilizando a hipótese sobre órbitas periódicas e o fato de que u é 1-lipschitziana, obtemos

$$0 \leq \sum_{i \in \mathbb{Z}_n} u(z_i) \leq \sum_{i \in \mathbb{Z}_n} u(t_i) + \sum_{i \in \mathbb{Z}_n} d_1(t_i, z_i) \leq \sum_{i \in \mathbb{Z}_n} u(t_i) + d_0(T_1 t_{n-1}, T_0 t_0),$$

donde pode ser obtido (3.25).

(ii) Sejam p_0, \dots, p_n os pontos de X_0 tais que $T_0 t_i = p_i$ e $T_1 t_i = p_{i+1}$, para $i = 0, \dots, n-1$. Seja $\omega: \mathbb{N} \rightarrow [0, n]$ a função definida por

$$\omega(0) = 0 \quad e \quad \omega(k+1) = \min\{1 + \eta(\omega(k)), n\},$$

onde $\eta: [0, n] \rightarrow [0, n]$ é a função que a cada índice i associa o maior índice j tal que $d_0(p_i, p_j) < L$.

Seja $s \geq 1$ o menor inteiro para o qual $\eta(\omega(s-1)) = n$. Então,

$$0 = \omega(0) < \omega(1) < \dots < \omega(s-1) \leq \omega(s) = n.$$

Repare que se $\eta(i) < j$, então, $d_0(p_i, p_j) \geq L$, pois $\eta(i)$ é definido como o maior índice k tal que $d_0(p_i, p_k) < L$. Em particular,

$$d_0(p_{\omega(k)}, p_{\omega(l)}) \geq L, \quad \forall k, l \in \{0, \dots, s-1\} \text{ tais que } k < l,$$

pois $\omega(l) \geq \omega(k+1) = 1 + \eta(\omega(k))$. Disto resulta que $s \leq N$. Assim sendo, podemos escrever

$$\sum_{i \in \mathbb{Z}_n} u(t_i) = \sum_{k=0}^{s-1} \left[\sum_{i=\omega(k)}^{\eta(\omega(k))-1} u(t_i) \right] + \sum_{k=0}^{s-2} u(t_{\eta(\omega(k))}).$$

Por (i), o somatório em cada colchete é minorado por $-BL$, pois a soma é feita sobre um trecho de órbita no qual o início e o fim distam menos de L . Igual minoração é obtida para cada $u(t_{\eta(\omega(k))})$. Logo,

$$\sum_{i \in \mathbb{Z}_n} u(t_i) \geq -(2s-1)BL \geq -(2N-1)BL = -CL.$$

□

Podemos, então, provar o teorema.

Demonstração do teorema 3.2.5: Pelo lema (3.2.2), é suficiente mostrar que

$$\sum_{i \in \mathbb{Z}_n} u(x_i) + C \sum_{i \in \mathbb{Z}_n} d_0(T_1 x_i, T_0 x_{i+1}) \geq 0, \quad (3.27)$$

para todo $n \geq 1$ e toda sequência de pontos $\{x_i\}_{i \in \mathbb{Z}_n} \subseteq X_1$. Como o lema auxiliar nos dá

ferramentas para trabalhar com $n \geq 2$, deixamos o caso $n = 1$ para depois, a ser provado como consequência dos demais. Considere, então, um inteiro $n \geq 2$ e uma sequência $\{x_i\}_{i \in \mathbb{Z}_n} \subseteq X_1$. Seja $\{y_i\}_{i \in \mathbb{Z}_n}$ a sequência obtida ao se aplicar o lema de sombreamento 3.2.4 à $\{x_i\}_{i \in \mathbb{Z}_n}$. É direto do lema de sombreamento que

$$\frac{1}{C} \sum_{i \in \mathbb{Z}_n} d_1(x_i, y_i) \leq \sum_{i \in \mathbb{Z}_n} d_1(x_i, y_i) \leq \sum_{i \in \mathbb{Z}_n} d_0(T_1 x_i, T_0 x_{i+1}) - \sum_{i \in \mathbb{Z}_n} d_0(T_1 y_i, T_0 y_{i+1}).$$

Consequentemente, pela desigualdade acima e a propriedade de u ser 1-lipschitziana, temos

$$\begin{aligned} \sum_{i \in \mathbb{Z}_n} u(x_i) + C \sum_{i \in \mathbb{Z}_n} d_0(T_1 x_i, T_0 x_{i+1}) &\geq \\ &\geq \sum_{i \in \mathbb{Z}_n} (u(y_i) - d_1(x_i, y_i)) + C \left[\sum_{i \in \mathbb{Z}_n} d_0(T_1 y_i, T_0 y_{i+1}) + \frac{1}{C} \sum_{i \in \mathbb{Z}_n} d_1(x_i, y_i) \right] \\ &= \sum_{i \in \mathbb{Z}_n} u(y_i) + C \sum_{i \in \mathbb{Z}_n} d_0(T_1 y_i, T_0 y_{i+1}). \end{aligned}$$

Logo, para demonstrar (3.27), basta demonstrar que

$$\sum_{i \in \mathbb{Z}_n} u(y_i) + C \sum_{i \in \mathbb{Z}_n} d_0(T_1 y_i, T_0 y_{i+1}) \geq 0. \quad (3.28)$$

Dizemos, então, que $i \in \mathbb{Z}_n$ é um bom índice se $T_1 y_i = T_0 y_{i+1}$, do contrário (isto é, $d_0(T_1 y_i, T_0 y_{i+1}) \geq L$), i é chamado de mau índice. Denotamos por M a cardinalidade de maus índices. Caso M seja nulo, a demonstração de (3.28) reduz-se a

$$\sum_{i \in \mathbb{Z}_n} u(y_i) \geq 0,$$

o que é verificado por hipótese, uma vez que $\{y_i\}_{i \in \mathbb{Z}_n}$ é uma órbita verdadeira. Caso M seja não nulo, obrigatoriamente temos $L < \infty$, uma vez que há pontos $y_i, y_{i+1} \in X_1$ com $d(T_1 y_{i+1}, T_0 y_i) \geq L$. Denotamos, assim, por $b(0), \dots, b(M-1)$ os maus índices. Logo,

$$0 \leq b(0) < b(1) < \dots < b(M-1) \leq n-1.$$

Escrevemos, ainda, $b(M) = n + b(0)$. Deste modo,

$$\sum_{i \in \mathbb{Z}_n} u(y_i) + C \sum_{i \in \mathbb{Z}_n} d_0(T_1 y_i, T_0 y_{i+1}) = \sum_{0 \leq j < M} \left[\sum_{b(j) < i \leq b(j+1)} u(y_i) + C d_0(T_1 y_{b(j)}, T_0 y_{b(j)+1}) \right].$$

Ora, pelo item (ii) do lema anterior,

$$\sum_{b(j) < i \leq b(j+1)} u(y_i) \geq -CL,$$

por se tratarem de trechos de órbita. Como $d_0(T_1 y_{b(j)}, T_0 y_{b(j)+1}) \geq L$, cada colchete acima é não negativo, donde resulta prova de (3.28).

Apenas retomando, o caso $n = 1$ é obtido ao se aplicar a conclusão acima obtida para a sequência $x_1 = x_2 = x$. Temos

$$\begin{aligned} 0 &\leq \frac{1}{2} [u(x_1) + u(x_2) + d_0(T_1 x_1, T_0 x_2) + d_0(T_1 x_2, T_0 x_2)] \\ &= u(x) + d_0(T_1 x, T_0 x). \end{aligned}$$

□

Temos, enfim, a desigualdade coomólica para um sistema anfidinâmico L -retificável. Vamos agora fazer a passagem de um sistema anfidinâmico para um SDT usual e relacionar as ferramentas obtidas com os conceitos de subação para os difeomorfismos de Anosov. Recordamos que estamos seguindo quase que *ipsis litteris* as ideias apresentadas em [6].

Definição 3.2.7. *Seja (X_0, d_0) um espaço métrico compacto. É chamada de correspondência de (X_0, d_0) qualquer subconjunto fechado $X_1 \subseteq X_0^2$.*

Uma correspondência X_1 de X_0 define um sistema anfidinâmico $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$, onde T_0 e T_1 são as restrições a X_1 das projeções canônicas $\text{proj}_i: X_0^2 \rightarrow X_0$, $i = 1, 2$. Contudo, pode ser preciso ajustar a métrica em X_1 para que o sistema anfidinâmico seja L -retificável. Consideramos a distância induzida pela distância de Manhattan em X_0^2 : dados $x, y \in X_1$, definimos

$$d_1(x, y) := \eta [d_0(T_0 x, T_0 y) + d_0(T_1 x, T_1 y)], \quad (3.29)$$

com $0 < \eta \leq 1$. Dizemos assim que a correspondência X_1 é L -retificável com coeficiente η se o sistema anfidinâmico $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$ for L -retificável, onde X_1 está munido da distância (3.29).

Uma correspondência bastante interessante é do tipo gráfico. Seja (X_0, T) um SDT (recordando, X_0 é um espaço métrico compacto, e $T: X_0 \rightarrow X_0$ uma aplicação contínua). Considere o gráfico de T :

$$X_1 = \{(x, Tx) \in X_0^2 : x \in X_0\} \subseteq X_0^2.$$

Suponha, primeiramente, que T é uma aplicação expansiva, ou seja, que há constantes $0 < L \leq \infty$ e $0 < \lambda \leq 1$ tais que $\forall x, y \in X_0$ verificando $d_0(Tx, y) < L$, há $x' \in T^{-1}y$ com $d_0(x, x') \leq \lambda d_0(Tx, Tx') = \lambda d_0(Tx, y)$. Neste caso, T é L -retificável com coeficiente $\eta = \frac{1-\lambda}{1+\lambda}$. Isto é fácil de verificar: dados $(x, Tx), (y, Ty) \in X_1$ que verificam $d_0(T_0(y, Ty), T_1(x, Tx)) < L$, pela expansividade há $x' \in T^{-1}y$ com $d_0(x, x') \leq \lambda d_0(Tx, y)$. Tomamos, então, (x', y) e (y, Ty) como os pontos a verificar a hipótese de retificabilidade. Logo, a desigualdade (3.18) é verificada pois, já utilizando a igualdade $y' = y$, temos

$$\begin{aligned} d_1[(x, Tx), (x', y)] + d_0[T_0(x, Tx), T_0(x', y)] &= \\ &= \frac{1-\lambda}{1+\lambda} [d_0(x, x') + d_0(Tx, y)] + d_0(x, x') \\ &= \frac{2d_0(x, x') + (1-\lambda)d_0(Tx, y)}{1+\lambda} \\ &\leq \frac{2\lambda d_0(Tx, y) + (1-\lambda)d_0(Tx, y)}{1+\lambda} \\ &= d_0(Tx, y) = d_0[T_1(x, Tx), T_0(y, Ty)]. \end{aligned}$$

No caso em que T é um homeomorfismo, podemos estabelecer condição de retificabilidade similar ao caso expansivo. Temos que T é L -retificável com coeficiente η se, e somente se, para quaisquer $x, y \in X_0$ com $d_0(x, y) < L$, há $z \in X_0$ tal que

$$(1 + \eta) [d_0(T^{-1}x, T^{-1}z) + d_0(Tz, Ty)] + \eta [d_0(x, z) + d_0(z, y)] \leq d_0(x, y). \quad (3.30)$$

A existência de tal z deixa clara a condição de L -retificabilidade, uma vez que (3.30) é exatamente a condição (3.18) escrita para os pontos $(T^{-1}x, x), (y, Ty), (T^{-1}z, z)$ e (z, Tz) pertencentes à X_1 , onde $(T^{-1}z, z)$ e (z, Tz) fazem papel de x' e y' , respectivamente. Por outro lado, a condição de L -retificabilidade implica que, para $(x, Tx), (y, Ty) \in X_1$, com $d_0(Tx, y) < L$, há $(x', Tx'), (y', Ty') \in X_1$ com $Tx' = y'$, verificando (3.18). Basta, agora, nomear este trecho de órbita de $\{T^{-1}z, z, Tz\}$.

Repare, então, caro leitor, que a condição (3.30) pode ser simplificada caso haja expansão de distâncias em dois sentidos, isto é, caso $d_0(T^{-1}x, T^{-1}z) \leq \lambda d_0(x, z)$ e $d_0(Tz, Ty) \leq \lambda d_0(z, y)$. Porém, isto é exatamente o que ocorre nos difeomorfismos de Anosov para $L \leq \varepsilon_0$ e λ como a menor das constantes relacionadas com a expansão de distâncias, para o ponto $z = [y, x]$ (na mesma nomenclatura adotada na primeira prova de existência de subação). Assim, se T é difeomorfismo de Anosov, a condição de L -retificabilidade com coeficiente η fica satisfeita se obtemos

$$d_0(x, z) + d_0(z, y) \leq \frac{1}{(1 + \eta)\lambda + \eta} d_0(x, y), \quad (3.31)$$

onde $z = [y, x]$.

Observação. Precisamos fazer uma ressalva aqui. É claro que o lado esquerdo de (3.31) majora $d_0(x, y)$, o que nos permite obter uma desigualdade apenas com η e λ . Mas, tal desigualdade poderia não ser suficiente para garantir os resultados. Precisamos então de uma distância que atue como a distância de Manhattan sobre pequenas vizinhanças, somando distâncias sobre os conjuntos estável e instável. Lembre que $z \in W_{\varepsilon_0}^s(y) \cap W_{\varepsilon_0}^u(x)$. Bousch propõe utilizarmos então um tipo de distância denominado distância Finsler, a qual é equivalente à distância usual da variedade e proporciona a igualdade desejada. Temos, localmente, $d_0(x, z) + d_0(z, y) = d_0(x, y)$, resultando na seguinte condição para η :

$$(1 + \eta)\lambda + \eta \leq 1, \quad \text{ou, ainda,}$$

$$\eta \leq \frac{1 - \lambda}{1 + \lambda}.$$

Logo, concluímos que se $u: X_1 \rightarrow \mathbb{R}$ é uma função 1-lipschitziana na métrica adequada, com soma positiva sobre órbitas periódicas, há função $v: X_0 \rightarrow \mathbb{R}$, a qual é C -lipschitziana, e que satisfaz a desigualdade coomológica. Porém, como X_1 nada mais é que o gráfico do difeomorfismo de Anosov $T: X_0 \rightarrow X_0$, na verdade, u depende apenas de uma coordenada (uma vez que a segunda coordenada é unicamente determinada pela primeira). Portanto, podemos escrever $u: X_0 \rightarrow \mathbb{R}$.

Sendo assim, se M é uma variedade compacta sem fronteira munida de uma distância adequada e $T: M \rightarrow M$ um difeomorfismo de Anosov, ao aplicar o teorema de Bousch à função $u = \beta(f) - f$, para um potencial 1-lipschitziano f , concluímos que há subação C -lipschitziana $v: M \rightarrow \mathbb{R}$ para o potencial f . Note que, neste segundo teorema de

obtenção de subação, não foi preciso supor que T seja de classe C^2 , apenas precisamos da hipótese de retificabilidade, a qual apareceu de forma natural para difeomorfismos de Anosov na métrica adequada.

3.3 Probabilidades maximizantes associadas a difeomorfismos de Anosov

Caro leitor, após ter visto como podemos obter subações para potenciais lipschitzianos associados a uma dinâmica do tipo difeomorfismo de Anosov, gostaríamos de apresentar algumas propriedades das probabilidades maximizantes sob este contexto de dinâmica hiperbólica. Lembramos que uma medida de probabilidade $\mu_0 \in \mathcal{M}_T$ é maximizante para o potencial $f: X \rightarrow \mathbb{R}$ se $\int f d\mu_0 = \sup_{\mu \in \mathcal{M}_T} \{ \int f d\mu \}$.

O primeiro resultado que queremos discutir é que se uma probabilidade $\mu \in \mathcal{M}_T$ é maximizante para um aberto de funções (na topologia lipschitziana), então, μ é uma probabilidade periódica, isto é, o suporte de μ é uma órbita periódica de T . Este resultado é devido a Yuan e Hunt [37]. Bousch, em [7], apresentou uma nova demonstração desse resultado, utilizando sistemas anfidinâmicos. Uma vez que já discutimos aqui a respeito dos sistemas anfidinâmicos, fazemos a demonstração neste contexto, o qual é mais geral que um SDT.

Um resultado bastante conhecido de [25] fornece condições sobre o SDT (X, T) para que toda função contínua $f: X \rightarrow \mathbb{R}$, cumprindo $\int f d\mu = 0, \forall \mu \in \mathcal{M}_T$, seja, na verdade, um cobordo, isto é $f = h \circ T - h$, para alguma função $h \in C(X)$. Temos um resultado auxiliar, de característica dual a este, que pode ser escrito para sistemas anfidinâmicos.

Teorema 3.3.1. *Sejam $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$ um sistema anfidinâmico, μ_0 e μ_1 probabilidades sobre X_0 e X_1 , respectivamente, tais que $T_{0*}\mu_1 = T_{1*}\mu_1 = \mu_0$. Suponha que para toda função lipschitziana $f: X_1 \rightarrow \mathbb{R}$ que verifica $\int f d\mu_1 = 0$, exista função $h: X_1 \rightarrow \mathbb{R}$ tal que $f = h \circ T_1 - h \circ T_0$, μ_1 -q.t.p. Então, há $N \in \mathbb{N}$ e uma órbita periódica $\{x_i\}_{i \in \mathbb{Z}_N} \subseteq X_1$ (com imagem $\{y_i\}_{i \in \mathbb{Z}_N} \subseteq X_0$) tal que*

$$\mu_1 = \frac{1}{N} \sum_{i \in \mathbb{Z}_N} \delta_{x_i}, \quad e, \quad \mu_0 = \frac{1}{N} \sum_{i \in \mathbb{Z}_N} \delta_{y_i}.$$

Damos apenas uma ideia geral da demonstração do teorema. São definidos os conjuntos

$$\mathcal{E}_0 := \{h: \text{supp}(\mu_0) \rightarrow \mathbb{R} : h \text{ é função Lipschitz}\}/\mathbb{R} \quad \text{e}$$

$$\mathcal{E}_1 := \{f: \text{supp}(\mu_1) \rightarrow \mathbb{R} : f \text{ é função Lipschitz}\},$$

e uma aplicação $\varsigma: \mathcal{E}_0 \rightarrow \mathcal{E}_1$ tal que $\varsigma(f) = f \circ T_1 - f \circ T_0$. As hipóteses exigidas garantem que esta aplicação é sobrejetiva, o que proporciona boas estimativas quando avaliamos $f \in \mathcal{E}_1$ sobre órbitas periódicas. Ademais, é empregada a distância de Wasserstein para garantir que μ_1 e μ_0 são probabilidades periódicas. Deixamos os detalhes da demonstração deste teorema para que o leitor consulte, por exemplo, em [7].

Como vimos na seção anterior, se $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$ é um sistema anfidinâmico L -retificável, então, para toda função 1-lipschitziana $u: X_1 \rightarrow \mathbb{R}$, com soma não negativa sobre órbitas periódicas, há função C -lipschitziana $v: X_0 \rightarrow \mathbb{R}$ tal que $u \geq \varsigma(v) = v \circ T_1 - v \circ T_0$. Este resultado e o teorema enunciado acima garantem a propriedade de maximização sobre abertos para uma determinada probabilidade $\mu \in \mathcal{M}_T$.

Teorema 3.3.2. *Sejam $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$ um sistema anfidinâmico L -retificável e μ_0, μ_1 probabilidades sobre X_0 e X_1 , respectivamente, tais que $T_{0*}\mu_1 = T_{1*}\mu_1 = \mu_0$. Suponha que exista uma função 1-lipschitziana $f: X_1 \rightarrow \mathbb{R}$, tal que $\beta(f)$ é realizado por μ_1 . Suponha ainda que há $\eta > 0$ tal que μ_1 também maximiza toda função $g: X_1 \rightarrow \mathbb{R}$ verificando $\text{Lip}(f - g) < \eta$. Então, μ_1 e μ_0 são probabilidades periódicas.*

Demonstração. Sem perda de generalidade, podemos supor que $\beta(f) = 0$. Seja então $h: X_1 \rightarrow \mathbb{R}$ uma função κ -lipschitziana, com $\kappa < \eta$, tal que $\int h d\mu_1 = 0$. Por hipótese, μ_1 maximiza uma bola de centro f raio η na norma lipschitziana, de modo que o valor da integral de $f + h$ é maximizado por μ_1 . Mas, o teorema da seção anterior garante que há $v_1: X_0 \rightarrow \mathbb{R}$ tal que $\varsigma(v_1) \leq f + h$ e $\text{Lip}(v_1) = C \times \text{Lip}(f + h) \leq C(1 + \eta)$. Porém, é evidente que $\int \varsigma(v_1) d\mu_1 = 0$, pois $\varsigma(v_1)$ é um cobordo, e, assim, obrigatoriamente temos $\varsigma(v_1) = f + h$ sobre o suporte de μ_1 . Aplicando o mesmo raciocínio para a função $f - h$, obtemos função $v_2: X_0 \rightarrow \mathbb{R}$ com $\text{Lip}(v_2) \leq C(1 + \eta)$ e $\varsigma(v_2) = f - h$ sobre o suporte de μ_1 . Logo, tomando $v = (v_1 - v_2)/2$, resulta que $\varsigma(v) = h$ sobre o suporte de μ_1 .

Obtivemos, então, a hipótese do teorema 3.3.1, de que toda função lipschitziana h definida sobre o suporte de μ_1 tal que $\int h d\mu_1 = 0$ é, na verdade, um cobordo. Logo, por este teorema, μ_1 e μ_0 são probabilidades periódicas. \square

Note que a existência de subação para potenciais lipschitzianos com constante Lipschitz controlada foi importante para se obter esse resultado. Repare também que os resultados foram obtidos para probabilidades μ_0 e μ_1 tais que $T_{0*}\mu_1 = T_{1*}\mu_1 = \mu_0$. A probabilidade μ_1 é, nesse caso, chamada de equiprojativa no sistema anfidinâmico – quando analisamos um SDT como sistema anfidinâmico, considerando $X_1 = \text{graf}(T)$, as probabilidades produto $\mu \times \mu$, com $\mu \in \mathcal{M}_T$, são, claramente, equiprojativas. Lembre que $\mu(T^{-1}(A)) = \mu(A)$, para toda $\mu \in \mathcal{M}_T$ e $A \in \mathcal{B}$.

Para o resultado anterior, há uma recíproca. Mostramos que, dadas as hipóteses, se μ é probabilidade que maximiza um aberto de funções lipschitzianas, então, μ é probabilidade periódica. Mostremos agora que toda probabilidade periódica maximiza um aberto de funções. Lembrando, estamos utilizando a linguagem de um sistema anfidinâmico, como em [7].

Teorema 3.3.3. *Seja $X_0 \xleftarrow{T_0} X_1 \xrightarrow{T_1} X_0$ um sistema anfidinâmico no qual T_0 e T_1 são aplicações lipschitzianas. Considere uma órbita periódica $\{x_i\}_{i \in \mathbb{Z}_N} \subseteq X_1$, e probabilidades $\mu_1 = \sum_{i \in \mathbb{Z}_N} \delta_{x_i}$ e $\mu_0 = T_{0*}\mu_1$. Defina a função 1-lipschitziana $f: X_1 \rightarrow \mathbb{R}$ dada por*

$$f(x) = -d_1(x, \text{supp}(\mu_1)).$$

Então, μ_1 é probabilidade f -maximizante e há $\eta > 0$ tal que μ_1 maximiza toda função $g: X_1 \rightarrow \mathbb{R}$ com $\text{Lip}(g - f) \leq \eta$.

Repare que um difeomorfismo de Anosov cumpre a hipótese de ser aplicação lipschitziana, uma vez que há cotas que limitam a dilatação local.

Demonstração. Como μ_0 e μ_1 são probabilidades periódicas, a aplicação $\varsigma: \mathcal{E}_0 \rightarrow \mathcal{E}_1$, definida na demonstração de teorema anterior, que leva funções lipschitzianas módulo constante definidas sobre o suporte de μ_0 em funções lipschitzianas de integral nula para μ_1 , via operação de cobordismo, é um isomorfismo. De fato, a função nula sobre o suporte de μ_1 só será atingida por função que seja constante sobre o suporte de μ_0 , a qual é o mesmo que zero, já que o espaço é considerado módulo constantes. Sendo assim, sua inversa também é contínua e, para toda função de integral nula $h: \text{supp}(\mu_1) \rightarrow \mathbb{R}$, há função $h_0: \text{supp}(\mu_0) \rightarrow \mathbb{R}$ com $\varsigma(h_0) = h$ e $\text{Lip}(h_0) \leq C\text{Lip}(h)$. É claro que $\int f d\mu_1 = 0$, então, podemos supor que a integral de g também é nula (lembre que g é função a mostrar que é maximizada por μ_1). Definimos a função $h = g - f$, a qual possui integral nula para μ_1 e, pelo que foi mencionado acima, há $h_0: \text{supp}(\mu_0) \rightarrow \mathbb{R}$ tal que $h = \varsigma(h_0)$ sobre

$\text{supp}(\mu_1)$, com $Lip(h_0) \leq CLip(h)$. Estendendo $\varsigma(h_0)$ para todo X_1 , podemos estabelecer a seguinte desigualdade

$$\begin{aligned} Lip(\varsigma(h_0) - h) &\leq Lip(h) + Lip(h_0 \circ T_1) + Lip(h_0 \circ T_0) \\ &\leq Lip(h) + (Lip(T_0) + Lip(T_1))Lip(h_0) \\ &\leq (1 + C(Lip(T_0) + Lip(T_1)))Lip(h). \end{aligned}$$

Assim, podemos escrever

$$\begin{aligned} (\varsigma(h_0) - h)(x) \leq Lip(\varsigma(h_0) - h)d(x, \text{supp}(\mu_1)) &\leq Lip(\varsigma(h_0) - h)d(x, \text{supp}(\mu_1)) \\ &\leq \frac{Lip(h)}{\eta} f(x), \end{aligned}$$

onde $\eta = (1 + C(Lip(T_0) + Lip(T_1)))^{-1}$. Repare que a comparação de x é feita com um ponto do suporte de μ_1 . Logo, escrevendo $g - \varsigma(h_0) = f - (\varsigma h_0 - h)$, temos

$$(g - \varsigma(h_0))(x) \leq \left(1 - \frac{Lip(h)}{\eta}\right) f(x) \leq 0,$$

sempre que $Lip(h) \leq \eta$. Neste caso, resulta que

$$\beta(g) = \beta(g - \varsigma(h_0)) \leq \sup(g - \varsigma(h_0)) \leq 0 = \int g \, d\mu_1,$$

o que prova que μ_1 é maximizante para g . □

Referências Bibliográficas

- [1] G. Atkinson, “Recurrence of co-cycles and random walks”, *The Journal of the London Mathematical Society* **13** (1976), 486-488.
- [2] V. Baladi, *Positive transfer operators and decay of correlations*, Advanced Series in Nonlinear Dynamics, World Scientific Publishing Company, Cingapura, 2000.
- [3] A. Baraviera, R. Leplaideur e A. O. Lopes, *Ergodic optimization, zero temperature limits and the max-plus algebra*, 29º Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro, 2013.
- [4] O. Bartle, *The Elements of Integration and Lebesgue Measure*, Wiley Classics Library, Wiley Interscience, New York, 1995.
- [5] T. Bousch, “La condition de Walters”, *Annales Scientifiques de l’École Normale Supérieure* **34** (2001), 287-311.
- [6] T. Bousch, “Le lemme de Mañé-Conze-Guivarc’h pour les systèmes amphidynamiques rectifiables”, *Annales de la Faculté des Sciences de Toulouse* **20** (2007), 1-14.
- [7] T. Bousch, “Nouvelle preuve d’un théorème de Yuan et Hunt”, *Bulletin de la Société Mathématique de France* **136** (2006), 227-242.
- [8] R. Bowen, *Equilibrium states and the ergodic theory of the Anosov diffeomorphisms*, Lecture Notes in Mathematics **470**, Springer, Berlin, 2008.
- [9] J. Brémont, “Gibbs measures at temperature zero”, *Nonlinearity* **16** (2003), 419-426.

- [10] H. Brézis, *Functional analysis, Sobolev spaces and partial differential Equations*, Universitext, Springer, New York, 2010.
- [11] M. P. do Carmo, *Geometria Diferencial de Curvas e Superfícies*, Sociedade Brasileira de Matemática - IMPA, Rio de Janeiro, 2005.
- [12] J. R. Chazottes e M. Hochman, “On the zero temperature limit of Gibbs states”, *Communications in Mathematical Physics* **297**, Springer, (2010), 265-281.
- [13] G. Contreras, Automorfismos lineares do toro e dinâmica simbólica, pré-publicação.
- [14] G. Contreras, Ground states are generically a periodic orbit, arXiv:1307.0559 (2013).
- [15] J. P. Conze e Y. Guivarc’h, Croissance des sommes ergodiques et principe variationnel, pré-publicação (1993).
- [16] L. C. Evans e R. F. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics **5**, CRC Press, Boca Raton, 1991.
- [17] J. Franks, “Anosov diffeomorphisms on tori”, *Transaction of the American Mathematical Society* **145** (1969), 117-124.
- [18] E. Garibaldi, Introduction to expanding ergodic optimization, pré-publicação (2014).
- [19] E. Garibaldi e J. T. A. Gomes, *Otimização de médias sobre grafos orientados*, 29º Colóquio Brasileiro de Matemática, IMPA, Rio de Janeiro, 2013.
- [20] E. Garibaldi e P. Thieullen, “Minimizing orbits in the discrete Aubry-Mather model”, *Nonlinearity* **24** (2011), 563-611.
- [21] O. Jenkinson, “Ergodic optimization”, *Discrete & Continuous Dynamical Systems*, **15** (2006), 197-224.
- [22] A. O. Lopes e P. Thieullen, “Subactions for Anosov diffeomorphisms”, *Astérisque* **287** (2003), 115-130.
- [23] A. Manning, “There are no new Anosov diffeomorphisms on tori”, *American Journal of Mathematics* **96** (1974), 422-429.
- [24] S. Mazur, “Über konvexe mengen in linearen normierten räumen”, *Studia Mathematica* **4** (1933), 70-84.

- [25] I. D. Morris, “A sufficient condition for the subordination principle in ergodic optimization”, *The Journal of the London Mathematical Society* **39** (2007), 214-220.
- [26] K. Oliveira e M. Viana, *Fundamentos da teoria ergódica*, Sociedade Brasileira de Matemática - IMPA, Rio de Janeiro, 2014.
- [27] W. Parry e M. Pollicott, *Zeta functions and the periodic orbit structure of hyperbolic dynamics*, Astérisque **187-188**, Société Mathématique de France, Paris, 1990.
- [28] R. R. Phelps, *Convex functions, monotone operators and differentiability*, Lecture Notes in Mathematics **1364**, Springer, Berlin, 1993.
- [29] M. Pollicott e M. Yuri, *Dynamical systems and ergodic theory*, London Mathematical Society Student Texts **40**, Cambridge University Press, Cambridge, 1998.
- [30] W. L. Reddy, “Expanding maps on compact metric spaces”, *Topology and its Applications* **13** (1982), 327-334.
- [31] R. T. Rockafellar, *Convex analysis*, Princeton Landmarks in Mathematics and Physics, Princeton University Press, New Jersey, 1970.
- [32] W. Rudin, *Principles of Mathematical Analysis*, International Series in Pure and Applied Mathematics, McGraw-Hill, Inc, New York, 1976.
- [33] D. Ruelle, *Thermodynamic formalism: the mathematical structure of equilibrium statistical mechanics*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2004.
- [34] G. F. Simmons, *Introduction to Topology and Modern Analysis*, Robert E. Krieger Publishing Company, Malabar, Florida, 1963.
- [35] N. S. R. Soares, *Teoremas ergódicos*, Dissertação de Mestrado, Faculdade de Ciências da Universidade do Porto, 2001.
- [36] P. Walters, *An introduction to ergodic theory*, Graduate Texts in Mathematics **79**, Springer, New York, 1982.
- [37] G. Yuan e B. R. Hunt, “Optimal orbits of hyperbolic systems”, *Nonlinearity* **12** (1999), 1207-1224.