Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Princípio Variacional e Entropia de Endomorfismos de Grupos de Lie

por

André Caldas de Souza

2012

Universidade de Brasília
Instituto de Ciências Exatas
Departamento de Matemática

Princípio Variacional e Entropia de Endomorfismos de Grupos de Lie

por

André Caldas de Souza *

Tese apresentada ao Departamento de Matemática da Universidade de Brasília, como parte dos requisitos para obtenção do grau de

DOUTOR EM MATEMÁTICA

Brasília, 05 de dezembro de 2012.

Comissão Examinadora:

Prof. Dr. Mauro M. A. Patrão - MAT/UnB (Orientador)

Prof. Dr. Luiz A. B. San Martin - UNICAMP/SP

Prof. Dr. Ryuichi Fukuoka - UEM/PR

Prof. Dr. Josiney Alves de Souza - UEM/PR

Prof. Dr. Lucas Conque Seco Ferreira - MAT/UnB

*O autor foi bolsista do CNPq durante a elaboração desta tese.

Resumo

Na presente tese, estendemos para sistemas dinâmicos não-compactos, o conceito de entropia topológica formulado por Adler, Konheim e McAndrew. Estendemos o *princípio variacional*, que relaciona as entropias topológica, de Kolmogorov-Sinai, e de Bowen, demonstrando que para todo sistema dinâmico contínuo $T : X \to X$, toda medida de Radon μ, e toda métrica d compatível com a topologia de X,

$$h_\mu (T) \leq h (T) \leq h^d (T).$$

Mostramos também que vale a igualdade

$$\sup_\mu h_\mu (T) = h (T) = \min_d h^d (T)$$

para uma classe de aplicações contínuas, definidas sobre o produto enumerável de espaços localmente compactos separáveis que chamamos de *sistemas dinâmicos do tipo produto*. Esta classe incluí aplicações contínuas quaisquer definidas em espaços localmente compactos separáveis, generalizando resultados de [HKR95] e [Pat10].

Na segunda parte, utilizamos o princípio variacional para mostrar que para um endomorfismo sobrejetivo $\phi : G \to G$ de um grupo de Lie G nilpotente ou redutível, a entropia topológica de ϕ coincide com a entropia de ϕ restrita ao maior subgrupo compacto e conexo do centro de G, $T(G)$. Ou seja,

$$h (\phi) = h \left(\phi|_{T(G)}\right),$$

generalizando o resultado em [Pat10].

Palavras chave: Entropia, Entropia topológica, Dinâmica topológica, Endomorfismo de grupo de Lie.

Abstract

In this thesis, we extend for non-compact dynamical systems the concept of topological entropy formulated by Adler, Konheim e McAndrew. We extend the *variational principle*, that relates the topological entropy, the Kolmogorov-Sinai entropy and the Bowen entropy, showing that for every continuous system $T : X \to X$, every Radon measure μ, and every metric d compatible with the topology of X,

$$h_\mu (T) \leq h (T) \leq h^d (T).$$

We also show that the equality

$$\sup_\mu h_\mu (T) = h (T) = \min_d h^d (T)$$

holds for a class of continuous systems defined over the countable product of locally compact separable spaces that we have called *product type dynamical system*. This class contains the continuous applications defined over locally compact separable spaces, generalizing the results in [HKR95] and [Pat10].

In the second part, we use the variational principle to show that for a surjective endomorphism $\phi : G \to G$ of a nilpotent or reducible Lie group G, the topological entropy of ϕ is equal to the topological entropy of ϕ restricted to the maximal connected and compact subgroup of the center of G. That is,

$$h (\phi) = h \left(\phi|_{T(G)}\right),$$

generalizing the result in [Pat10].

Keywords: Entropy, Topological entropy, Topological dynamics, Lie group endomorphism.

Sumário

Sumário **v**

Introdução **1**

1 Preliminares **5**
1.1 Sistemas Dinâmicos 5
1.2 Definições de Entropia 7

2 Sistemas Dinâmicos *do Tipo Produto* **32**
2.1 Definição 32
2.2 Propriedades dos Sistemas *do Tipo Produto* 36

3 Princípio Variacional **46**
3.1 Demonstração do Princípio Variacional 46

4 Pares de Li-Yorke **57**
4.1 Definição e Propriedades 57

5 Dinâmica de Endomorfismos **65**
5.1 Caso Linear 65
5.2 Endomorfismos de Grupos de Lie Nilpotentes 68
5.3 Endomorfismos de Grupos de Lie Redutíveis 78

A Apêndice **85**
A.1 Esperança Condicional 85
A.2 Teoremas de Recorrência de Poincaré 86
A.3 Entropia Condicional 87
A.4 Teorema de Representação de Riesz 93
A.5 Entropia de Bowen em Grupos de Lie 97

Referências Bibliográficas **99**

Introdução

Esta tese pode ser dividida em duas partes. Na primeira, desenvolvemos o conceito de entropia topológica para sistemas dinâmicos não necessariamente compactos. Os resultados são bastante gerais, pois não são impostas maiores restrições às hipóteses dos resultados.

O conceito de *entropia* nasceu da observação dos sistemas termodinâmicos. Sua formulação estatística (combinatorial) é fruto do trabalho de Boltzmann em mecânica estatística, e posteriormente do trabalho de Shannon em teoria de informação (veja [Ell85]). Adler, Konheim e McAndrew criaram o conceito de *entropia topológica* para sistemas dinâmicos topológicos compactos. Dinaburg e Bowen deram uma definição de entropia para sistemas munidos de uma métrica. Para sistemas metrizáveis compactos, a definição de entropia de Dinaburg e Bowen é facilmente mostrada ser independente da métrica adotada, e equivalente à entropia topológica definida por Adler, Konheim e McAndrew. É sabido que para os sistemas compactos metrizáveis $T : X \to X$, vale o *princípio variacional*

$$\sup_{\mu} h_{\mu}(T) = h(T) = h_d(T),$$

onde o supremo é tomado em todas as medidas T-invariantes, e d é uma métrica qualquer compatível com a topologia de X. As maiores inspirações para esta tese, foram

- a demonstração clássica de Misiurewicz para o princípio variacional no caso compacto (veja o teorema 8.6 de [Wal00]); e

- o trabalho [Pat10] de Patrão sobre o princípio variacional em espaços métricos não-compactos — principalmente a idéia de *coberturas admissíveis* (veja a Definição 1.2.7).

Estendemos o princípio variacional (Teorema 3.1.2) a uma classe de sistemas dinâmicos topológicos definidos em um produto enumerável de espaços metrizáveis localmente compactos separáveis. A este sistema, demos o nome de *sistema dinâmico do tipo produto* (Definição 2.1.1). Estendemos a Proposição 1.4 de [HKR95], de modo que a desigualdade no Lema 3.1.1 é válida para todos os sistemas dinâmicos topológicos, e não apenas homeomorfismos de sistemas metrizáveis. Graças à definição de entropia topológica utilizando o conceito de *cobertura admissível*, nossa demonstração pôde ser feita de maneira muito mais simples, utilizando argumentos elementares, enquanto que a demonstração em [HKR95] apela para teoremas mais avançados como o *Teorema de Shannon-McMillan-Breiman* e o *Teorema Ergódico de Birkhoff*. O princípio variacional apresentado nesta tese, generaliza também os Lemas 1.5 e 16 de [HKR95].

A segunda parte é uma aplicação da teoria no estudo dos endomorfismos de grupos de Lie. Utilizamos o princípio variacional para estender resultados de [CP13], e calcular a entropia de endomorfismos de grupos de Lie conexos. Em, [Bow71], Bowen apresenta uma fórmula para sua entropia de endomorfismos quando a métrica considerada é invariante sob a ação do grupo. No entanto, no caso não-compacto, o valor da entropia depende da métrica escolhida. E, em [Pat10], Patrão mostrou que a *fórmula de Bowen* fornece apenas uma cota superior para as entropias de Kolmogorov-Sinai, e que não coincide necessariamente com o supremo dessas entropias. O argumento em [Pat10] nos levou a conjecturar se a entropia de um endomorfismo T de um grupo de Lie conexo não seria simplesmente a entropia da restrição a sua *componente toral* (o maior subgrupo compacto e conexo do centro). De fato, demonstramos que este fato é verdadeiro para um endomorfismo sobrejetivo de um grupo de Lie nilpotente ou redutível. O caso em que T é uma aplicação linear já havia sido feito em [Pat10].

No Capítulo 1, relembramos as definições de entropia de Kolmogorov-Sinai, de entropia topológica de sistemas compactos, de entropia de Bowen e de d-entropia. Além de estendermos a definição de entropia topológica, também estabelecemos algumas relações simples entre todos esses conceitos. O resultado mais importante do capítulo é a Proposição 1.2.31, que corresponde a uma das desigualdades do princípio variacional. A simplicidade das demonstrações se deve à forma como foi definida a entropia topológica, e à forma como foi reformulada a entropia de Bowen, permitindo que relações elementares poderosas fossem estabelecidas entre as diferentes formas de entropia.

No Capítulo 2, definimos os *sistemas do tipo produto*, e determinamos as propriedades desses sistemas que nos permitirão demonstrar o princípio variacional (Teorema 3.1.2). Os sistemas do tipo produto são tais que permitem uma noção de convergência semelhante ao que se consegue em sistemas compactos com a topologia fraca-$*$ e a compacidade do conjunto das medidas de probabilidade. A idéia é conseguir condições mínimas — Lemas 2.2.7 e 2.2.9 — que permitam que a demonstração clássica de Misiurewicz fosse adaptada aos sistemas do tipo produto.

O princípio variacional é finalmente demonstrado no Capítulo 3. É importante notar que na demonstração do Teorema 3.1.2, foi assumida a hipótese de que o sistema em questão é *do tipo produto*, mas que o Lema 3.1.1 não impõe ao sistema $T : X \to X$ nenhuma restrição além da continuidade da aplicação T.

No Capítulo 4, a existência dos pares de Li-Yorke (Definição 4.1.1) é demonstrada para qualquer sistema metrizável localmente compacto separável com entropia positiva. Este resultado é usado posteriormente no Capítulo 5 para mostrar que grupos semi-simples não podem ter entropia positiva, por não possuirem pares de Li-Yorke.

E no Capítulo 5, mostramos que a entropia de endomorfismos de grupos de Lie lineares, e a entropia de endomorfismos sobrejetivos de grupos de Lie nilpotentes conexos e dos grupos de Lie redutíveis conexos é igual à entropia da restrição a sua componente toral. Além do próprio princípio variacional, a *decomposição de Jordan multiplicativa* tem um papel fundamental nessas demonstrações. Cada componente da decomposição de Jordan é mais fácil de ser estudada a partir do ponto de vista topológico, métrico ou de medida (veja a Proposição 5.1.1 e o Lema 5.3.3). Por este motivo, os resultados dos artigos [FPS10], [PSS11] e [Cal10] foram importantes. Por exemplo, no caso linear, o fato de a componente elíptica ser uma isometria para uma métrica adequada d, nos permite concluir que $h^d(T_E) = 0$. E a caracterização do conjunto recorrente em termos das componentes de Jordan nos permite concluir que

$$h(T) = h(T_E) \leq h^d(T_E) = 0.$$

Da mesma forma, a caracterização do conjunto recorrente das conjugações de um grupo semi-simples em termos de suas componentes de Jordan também foi fundamental para concluirmos que a entropia topológica de um endomorfismo sobrejetivo ϕ de um grupo de Lie semi-simples G é igual à entropia de ϕ restrito à componente toral de G. O resultado principal são o Teorema 5.2.8 e o Corolário 5.3.12, que afirmam que para um endomorfismo sobrejetivo $\phi : G \to G$ de um grupo de Lie G nilpotente ou redutível,

$$h(\phi) = h\left(\phi|_{T(G)}\right),$$

onde $T(G)$ é a componente toral de G. Uma observação interessante, é que mesmo no caso compacto foi feito progresso, pois nesse caso também ficou demonstrado que a entropia topológica de T é igual à entropia topológica da restrição de T à componente toral do grupo.

Propriedades elementares de conceitos conhecidos que foram utilizados na tese, como o de *esperança condicional* e o de *entropia condicional* se encontram no apêndice. O *teorema de recorrência de Poincaré*, que relaciona a topologia de um sistema dinâmico com as medidas invariantes, o *teorema de representação de Riesz*, que nos permite associar às medidas finitas a topologia fraca-$*$, e a *fórmula de Bowen* também estão no apêndice.

CAPÍTULO 1

Preliminares

1.1 Sistemas Dinâmicos

Um *sistema dinâmico* $T : X \to X$ é uma aplicação contínua T definida sobre um espaço topológico X. Uma compactificação de T é um sistema dinâmico $\widetilde{T} : \widetilde{X} \to \widetilde{X}$ tal que $\widetilde{X}$ é compacto, $X \subset \widetilde{X}$ é denso em $\widetilde{X}$ e $\widetilde{T}|_X = T$.

Uma família $\mathcal{A}$ de subconjuntos de X é uma *cobertura de* X, ou simplesmente uma *cobertura*, quando

$$X = \bigcup_{A \in \mathcal{A}} A.$$

Se os conjuntos em $\mathcal{A}$ são disjuntos, então dizemos que $\mathcal{A}$ é uma *partição de* X. Uma *subcobertura* de $\mathcal{A}$ é uma família $\mathcal{B} \subset \mathcal{A}$ que também é uma cobertura de X. Se $Y \subset X$ e $\mathcal{A}$ é uma cobertura de X, então denotamos por $Y \cap \mathcal{A}$ a cobertura de Y dada por

$$Y \cap \mathcal{A} = \left\{ A \cap Y \,\middle|\, A \in \mathcal{A} \right\}.$$

Denotamos por $N(\mathcal{A})$ o ínfimo das cardinalidades dentre as subcoberturas de $\mathcal{A}$. Para $Y \subset X$, usamos $N_Y(\mathcal{A})$ para denotar $N(Y \cap \mathcal{A})$.

Dadas as coberturas $\mathcal{A}$ e $\mathcal{B}$ de um conjunto arbitrário X, dizemos que $\mathcal{A}$ é *mais fina que* $\mathcal{B}$ ou que $\mathcal{A}$ *refina* $\mathcal{B}$ — e escrevemos $\mathcal{B} \prec \mathcal{A}$ — quando todo elemento de $\mathcal{A}$ é subconjunto de algum elemento de $\mathcal{B}$. Neste caso, também dizemos que $\mathcal{B}$ é mais grossa que $\mathcal{A}$. A relação $\prec$ é uma *pré-ordem*, e se identificarmos os conjuntos *simétricos* (i.e.: os conjuntos A e B tais que $A \prec B$ e $B \prec A$), temos um *reticulado*. Como de costume, $\mathcal{A} \vee \mathcal{B}$ denota o

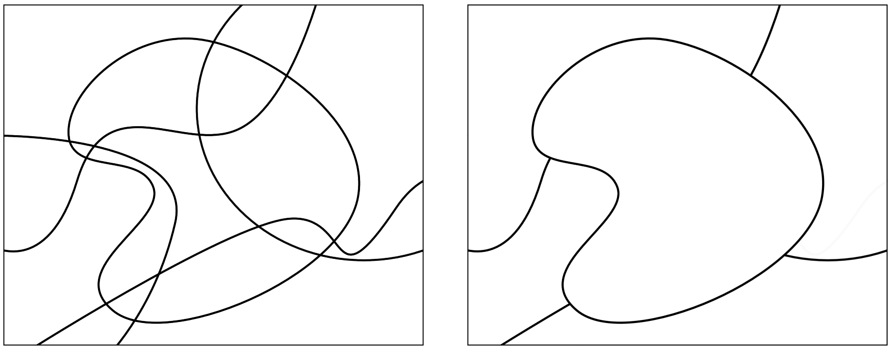

Figura 1.1: Uma cobertura e uma partição.

representante dentre as coberturas mais grossas de X que refinam ambos, $\mathcal{A}$ e $\mathcal{B}$. Explicitamente,

$$\mathcal{A} \vee \mathcal{B} = \left\{ A \cap B \,\middle|\, A \in \mathcal{A}, B \in \mathcal{B}, A \cap B \neq \emptyset \right\}.$$

Dado um sistema dinâmico $T : X \to X$ e uma cobertura $\mathcal{A}$, então, para cada $n \in \mathbb{N}$, definimos

$$\mathcal{A}^n = \mathcal{A} \vee \cdots \vee T^{-(n-1)}(\mathcal{A}),$$

onde

$$T^{-j}(\mathcal{A}) = \left\{ T^{-j}(A) \,\middle|\, A \in \mathcal{A} \right\}.$$

Se quisermos enfatizar que o sistema dinâmico em questão é T, então escrevemos $\mathcal{A}_T^n$.

Exemplo 1.1.1 (Endomorfimso de Grupo de Lie)**.** Se G é um grupo de Lie, e $T : G \to G$ um endomorfismo de G, então T é um sistema dinâmico. Dada uma vizinhança aberta da identidade, A, então $\mathcal{A} = \left\{ gA \,\middle|\, g \in G \right\}$ é uma cobertura aberta de G. E, por exemplo, $\mathcal{B} = \{A\} \cup \left\{ gA \,\middle|\, g \in G \setminus A \right\}$ é uma subcobertura de $\mathcal{A}$. Se $C \subset A$ também é uma vizinhança aberta da identidade, então, fazendo $\mathcal{C} = \left\{ gC \,\middle|\, g \in G \right\}$, temos que $\mathcal{C}$ é uma cobertura aberta tal que $\mathcal{A} \prec \mathcal{C}$.

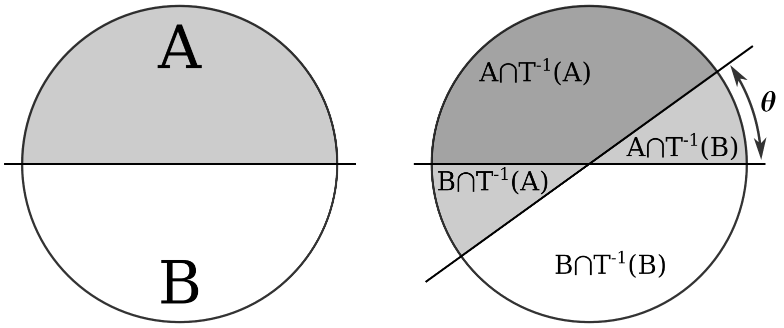

Figura 1.2: Dada a partição $\mathcal{C} = \{A, B\}$ e a rotação do disco por um ângulo $-\theta$, com $0 < \theta < \pi$, então $\mathcal{C}^2$ possui quatro "fatias".

1.2 Definições de Entropia

Entropia com Medida

A estrutura que provavelmente melhor se encaixa com as ideias por trás do conceito de entropia, é um sistema dinâmico $T : X \to X$ onde está definida também uma medida de probabilidade μ que é T-invariante. Ou seja, uma medida de probabilidade definida nos borelianos de X, tal que $\mu = \mu \circ T^{-1}$. No entanto, para as definições que se seguem, vamos evitar restrições desnecessárias. A maior parte das definições que seguem não exigem que o sistema seja contínuo ou que a medida seja invariante, ou mesmo que seja uma medida de probabilidade.

Considere o espaço de medida finito (X, μ) e uma partição mensurável $\mathcal{C}$. A *entropia da partição* $\mathcal{C}$ é

$$H_\mu(\mathcal{C}) = \sum_{A \in \mathcal{C}} \mu(A) \log \frac{1}{\mu(A)}.$$

Se a aplicação $T : X \to X$ é mensurável, a *entropia de T com respeito a partição $\mathcal{C}$* é

$$h_\mu\big(T \,\big|\, \mathcal{C}\big) = \limsup_{n\to\infty} \frac{1}{n} H_\mu(\mathcal{C}^n).$$

Quando μ é T-invariante, a sequência do lado direito é convergente e portanto podemos tomar o limite ao invés do $\limsup$. A *entropia de Kolmogorov-Sinai de T* é

$$h_\mu(T) = \sup_{\substack{\mathcal{C}:\text{ partição} \\ \text{mensurável finita}}} h_\mu\big(T \,\big|\, \mathcal{C}\big).$$

Note que em nossa definição para entropia de Kolmogorov-Sinai não assumimos que μ seja uma medida de probabilidade. Isso será útil para o caso em que X é metrizável localmente compacto (mas não necessariamente compacto), pois neste caso, o conjunto das medidas de probabilidade não é compacto na topologia fraca-$*$, enquanto que o conjunto das medidas finitas com $0 \leq \mu(X) \leq 1$ é de fato compacto (veja o Lema 2.2.7).

A seguinte proposição nos permite estimar a diferença entre a entropia de duas partições. Ou então, aproximar uma partição por outra "mais adequada" sem alterar significativamente o valor da entropia. Utilizaremos esta aproximação na demonstração do *princípio variacional* (Teorema 3.1.2). Dado $\delta > 0$, diremos que $\mathcal{D}$ é um δ-refinamento de $\mathcal{C}$, quando $N(\mathcal{D}) \leq N(\mathcal{C}) + 1$, e para cada $C \in \mathcal{C}$, existir $D_C \in \mathcal{D}$ tal que

$$\mu(C \cap D_C) \geq (1 - \delta)\mu(C).$$

Quando $\mathcal{D}$ δ-refina $\mathcal{C}$ e vice-versa, dizemos que $\mathcal{D}$ é uma δ-aproximação de $\mathcal{C}$.

Proposição 1.2.1. *Seja $T : X \to X$ um sistema dinâmico mensurável, com medida de probabilidade T-invariante μ. Dados $\varepsilon > 0$ e $N \in \mathbb{N}$, então, existe $\delta > 0$ tal que para toda cobertura mensurável $\mathcal{C}$ com $N(\mathcal{C}) < N$ e todo δ-refinamento $\mathcal{D}$ de $\mathcal{C}$,*

$$h_\mu(T \mid \mathcal{C}) \leq h_\mu(T \mid \mathcal{D}) + \varepsilon.$$

Em particular, se $\mathcal{D}$ for uma δ-aproximação de $\mathcal{C}$, então,

$$\left| h_\mu(T \mid \mathcal{C}) - h_\mu(T \mid \mathcal{D}) \right| \leq \varepsilon.$$

Demonstração. Veja no apêndice, a demonstração da Proposição A.3.6 □

Corolário 1.2.2. *Seja $T : X \to X$ um sistema dinâmico mensurável, com medida de probabilidade T-invariante μ. Se $\mathcal{F}$ é uma álgebra geradora dos borelianos de X, então*

$$h_\mu(T) = \sup_{\substack{\mathcal{D} \subset \mathcal{F} \\ \text{partição finita}}} h_\mu(T \mid \mathcal{D}).$$

Demonstração. Dada uma partição mensurável finita $\mathcal{C}$ e $\varepsilon > 0$, escolha δ tal que todo δ-refinamento $\mathcal{D}$ de $\mathcal{C}$ é tal que

$$h_\mu(T \mid \mathcal{C}) \leq h_\mu(T \mid \mathcal{D}) + \varepsilon.$$

É fácil mostrar, usando o fato de que todo conjunto mensurável pode ser *"μ-aproximado"* por um elemento de $\mathcal{F}$, que existe um δ-refinamento $\mathcal{D} \subset \mathcal{F}$. Tomando o supremo, em todas as partições finitas $\mathcal{D}$,

$$h_\mu\big(T \,\big|\, \mathcal{C}\big) \leq \sup_{\substack{\mathcal{D} \subset \mathcal{F} \\ \text{partição finita}}} h_\mu\big(T \,\big|\, \mathcal{D}\big) + \varepsilon \leq h_\mu\left(T\right) + \varepsilon.$$

Agora, basta usar o fato de que ε é arbitrário e então tomar o supremo em $\mathcal{C}$ para finalizar a demonstração do Corolário. $\square$

Exemplo 1.2.3 (Shift de Markov)**.** Considere o sistema dinâmico

$$\begin{array}{rccc} T: & S^{\mathbb{N}} & \to & S^{\mathbb{N}} \\ & (x_j)_{j\in\mathbb{N}} & \mapsto & (x_{j+1})_{j\in\mathbb{N}} \end{array},$$

onde S é um conjunto finito. Seja $\pi_k : S^{\mathbb{N}} \to S^k$ a projeção nas primeiras k coordenadas, e seja μ uma medida T-invariante tal que o sistema seja uma cadeia de Markov, com probabilidades iniciais p_s e probabilidades de transição p_{st}, onde $s, t \in S$. Isso significa que, dado $(x_1, \ldots, x_k) \in S^k$,

$$\mu\left(\pi_k^{-1}(x_1, \ldots, x_k)\right) = p_{x_1} p_{x_1,x_2} \cdots p_{x_{k-1},x_k}.$$

Pelo Corolário 1.2.2, basta calcular a entropia referente às partições finitas formadas por conjuntos da forma $\pi_k^{-1}(B)$, onde $k \in \mathbb{N}$ e $B \subset S^k$ é um boreliano. Mas dada uma tal partição $\mathcal{C}$, podemos escolher k tal que $\left\{\pi_k^{-1}(x) \,\middle|\, x \in S^k\right\}$ refina $\mathcal{C}$. Assim, podemos assumir que

$$\mathcal{C} = \mathcal{C}_k = \left\{\pi_k^{-1}(x) \,\middle|\, x \in S^k\right\}$$

para algum k.

Agora, note que

$$\mathcal{C}_k^n = \mathcal{C}_{k+n}.$$

E portanto,

$$h_\mu\big(T \,\big|\, \mathcal{C}_k\big) = \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\mathcal{C}_{k+n}\right).$$

Fazendo $m = n + k$,

$$
\begin{aligned}
h_\mu\big(T \bigm| \mathcal{C}_k\big) &= \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\mathcal{C}_m\right) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{s\in S^m} \mu\left(\pi_m^{-1}(x)\right) \log \frac{1}{\mu\left(\pi_m^{-1}(x)\right)} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{(s_1,\ldots,s_m)\in S^m} p_{s_1}p_{s_1,s_2}\cdots p_{s_{m-1},s_m} \log \frac{1}{p_{s_1}p_{s_1,s_2}\cdots p_{s_{m-1},s_m}} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{s_1\in S} \left[p_{s_1} \left(\sum_{(s_2,\ldots,s_k)\in S^{k-1}} p_{s_1,s_2}p_{s_2,s_3}\cdots p_{s_{k-1},s_k} \right) \log \frac{1}{p_{s_1}} \right] \\
&\quad + \lim_{n\to\infty} \frac{1}{n} \sum_{(s_1,\ldots,s_m)\in S^m} p_{s_1}p_{s_1,s_2}\cdots p_{s_{m-1},s_m} \left(\sum_{j=1}^{m-1} \log \frac{1}{p_{s_j,s_{j+1}}} \right) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{s\in S} p_s \log \frac{1}{p_s} \\
&\quad + \lim_{n\to\infty} \frac{1}{n} \sum_{(s_1,\ldots,s_m)\in S^m} p_{s_1}p_{s_1,s_2}\cdots p_{s_{m-1},s_m} \left(\sum_{j=1}^{m-1} \log \frac{1}{p_{s_j,s_{j+1}}} \right) \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{(s_1,\ldots,s_m)\in S^m} p_{s_1}p_{s_1,s_2}\cdots p_{s_{m-1},s_m} \left(\sum_{j=1}^{m-1} \log \frac{1}{p_{s_j,s_{j+1}}} \right).
\end{aligned}
$$

Note que dado $j = 1, \ldots, m-1$, a probabilidade de $x_j = s$ e $x_{j+1} = t$ é $p_s p_{s,t}$. Portanto, se agruparmos os termos $\log \frac{1}{p_{s,t}}$, teremos

$$
\begin{aligned}
h_\mu\big(T \bigm| \mathcal{C}_k\big) &= \lim_{n\to\infty} \frac{1}{n} \sum_{s,t\in S} \left(\sum_{j=1}^{m-1} \sum_{\substack{(s_1,\ldots,s_m)\in S^m \\ s_j=s, s_{j+1}=t}} p_{s_1}p_{s_1,s_2}\cdots p_{s_{m-1},s_m} \right) \log \frac{1}{p_{s,t}} \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{s,t\in S} \left(\sum_{j=1}^{m-1} p_s p_{s,t} \right) \log \frac{1}{p_{s,t}} \\
&= \lim_{n\to\infty} \frac{1}{n} (m-1) \sum_{s,t\in S} p_s p_{s,t} \log \frac{1}{p_{s,t}} \\
&= \sum_{s,t\in S} p_s p_{s,t} \log \frac{1}{p_{s,t}}.
\end{aligned}
$$

Pelo Corolário 1.2.2,

$$h_\mu(T) = \sum_{s,t \in S} p_s p_{s,t} \log \frac{1}{p_{s,t}}.$$

Em particular, se $\mu = \nu^{\mathbb{N}}$ para uma medida ν em S, então $p_{s,t} = p_t$, e $\sum_{s \in S} p_s p_{s,t} = p_t \sum_{s \in S} p_s = p_t$. Portanto, neste caso,

$$h_\mu(T) = \sum_{t \in S} p_t \log \frac{1}{p_t}.$$

Na definição de entropia de Kolmogorov-Sinai, não exigimos que a medida em questão fosse uma medida de probabilidade. A medida pode até ser igual a zero. O lema a seguir serve para "normalizar" a entropia de medidas finitas, substituindo-as por medidas de probabilidade.

Lema 1.2.4. *Dado um sistema dinâmico $T : X \to X$ e uma medida finita T-invariante μ nos borelianos de X, então, para $\alpha \geq 0$,*

$$h_{\alpha\mu}(T) = \alpha h_\mu(T).$$

Demonstração. Podemos assumir que $\alpha \neq 0$, pois $h_0(T) = 0$.

Para qualquer partição mensurável finita $\mathcal{C}$,

$$\begin{aligned}
\frac{1}{n} \sum_{C \in \mathcal{C}^n} \alpha\mu(C) \log \frac{1}{\alpha\mu(C)} &= \frac{\alpha}{n} \sum_{C \in \mathcal{C}^n} \mu(C) \log \frac{1}{\alpha\mu(C)} \\
&= \frac{\alpha}{n} \left(\sum_{C \in \mathcal{C}^n} \mu(C) \log \frac{1}{\mu(C)} \right) + \frac{\alpha}{n} \mu(X) \log \frac{1}{\alpha}.
\end{aligned}$$

Agora, basta tomar o limite para $n \to \infty$. □

Entropia Topológica

Inspirados pela entropia de Kolmogorov-Sinai, Adler, Konheim e McAndrew introduziram em [AKM65] um conceito puramente topológico de entropia. Quando X é compacto, Dinaburg e Goodman provaram (veja [Din69, Goo71]) o *princípio variacional*, que afirma que a *entropia topológica* é igual ao supremo das entropias de Kolmogorov-Sinai tomadas sobre todas as medidas de probabilidade de Radon T-invariantes. Uma medida finita μ é de Radon se para todo conjunto mensurável A,

$$\mu(A) = \sup_{\substack{K \subset A \\ K:\ \text{compacto}}} \mu(K).$$

Em [Pat10], Patrão percebeu que quando o sistema dinâmico admite uma compactificação por um ponto, o princípio variacional continua valendo desde que a definição original seja adaptada. Neste trabalho, fornecemos uma definição de *entropia topológica* que estende as definições anteriores e nos permite demonstrar o princípio variacional em um contexto mais amplo.

Exemplo 1.2.5 (Sistema que não admite compactificação por um ponto)**.** Seja $X = (0,1)$. Considere o sistema

$$T(x) = \begin{cases} x, & x \leq \frac{1}{2} \\ \frac{1}{2}, & x > \frac{2}{2} \end{cases} .$$

Esse sistema não admite compactificação por um ponto. De fato, se $(0,1) \cup \{\infty\}$ é uma compactificação por um ponto de X, então $1 - \frac{1}{n} \to \infty$ e $0 + \frac{1}{n} \to \infty$. No entanto,

$$\lim_{n\to\infty} T\left(1 - \frac{1}{n}\right) = \frac{1}{2} \neq \infty = \lim_{n\to\infty} T\left(0 + \frac{1}{n}\right)$$

(veja a Figura 1.3).

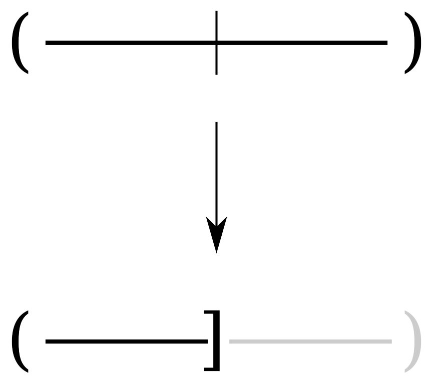

Figura 1.3: Sistema dinâmico que não admite compactificação por um ponto.

Definição 1.2.6 (Entropia da Cobertura)**.** *Dada uma cobertura $\mathcal{A}$ de um conjunto X, a* entropia da cobertura $\mathcal{A}$ *é*

$$H(\mathcal{A}) = \log N(\mathcal{A}).$$

Motivados pela definição apresentada em [Pat10], vamos nos restringir a coberturas abertas de um tipo específico.

Definição 1.2.7 (Cobertura Admissível)**.** *Em um espaço topológico X, uma cobertura de abertos $\mathcal{A}$ é chamada de* cobertura admissível *quando ao menos um de seus elementos tem complemento compacto.*

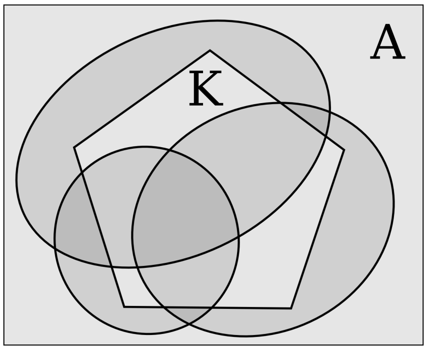

Figura 1.4: Em uma cobertura admissível $\mathcal{A}$, temos que os elementos de $\mathcal{A}$ cobrem um conjunto K compacto tal que $K^c = A \in \mathcal{A}$.

Observação 1.2.8. Se $\widetilde{X}$ é uma compactificação de X, então, para qualquer cobertura admissível $\mathcal{A}$ de X, existe uma cobertura aberta $\widetilde{\mathcal{A}}$ de $\widetilde{X}$, tal que $\mathcal{A} = X \cap \widetilde{\mathcal{A}}$. De fato, $\widetilde{\mathcal{A}}$ pode ser construída escolhendo, para cada $A \in \mathcal{A}$, um aberto $\widetilde{A} \subset \widetilde{X}$ tal que $A = X \cap \widetilde{A}$. A princípio, nada garante que a família de abertos

$$\widetilde{\mathcal{A}} = \left\{ \widetilde{A} \,\middle|\, A \in \mathcal{A} \right\}$$

cobre $\widetilde{X}$. Seja $K \subset X$ um compacto tal que $A_K = X \setminus K \in \mathcal{A}$. Tomando o cuidado de escolher $\widetilde{A_K} = \widetilde{X} \setminus K$, garantimos que $\widetilde{\mathcal{A}}$ seja de fato uma cobertura para $\widetilde{X}$.

Uma outra alternativa, menos construtiva, é simplesmente tomar para $\widetilde{\mathcal{A}}$, a família de todos os abertos $\widetilde{A} \subset \widetilde{X}$ tais que $X \cap \widetilde{A} \in \mathcal{A}$. Novamente, $\widetilde{\mathcal{A}}$ será uma cobertura de $\widetilde{X}$ pelo simples fato de $\widetilde{X} \setminus K$ pertencer a $\widetilde{\mathcal{A}}$, além de K ser coberto por elementos de $\widetilde{\mathcal{A}}$.

Definição 1.2.9 (Entropia Topológica)**.** *Para um sistema dinâmico* $T : X \to X$ *e uma cobertura* $\mathcal{A}$*, a* entropia topológica de T com respeito a $\mathcal{A}$ *é*

$$h\big(T \,\big|\, \mathcal{A}\big) = \lim_{n\to\infty} \frac{1}{n} H\left(\mathcal{A}^n\right).$$

A entropia topológica de T *é*

$$h\left(T\right) = \sup_{\substack{\mathcal{A}\text{: cobertura} \\ \text{admissível}}} h\big(T \,\big|\, \mathcal{A}\big).$$

Neste trabalho, o termo *entropia de AKM* se refere à definição original de entropia dada por Adler, Konheim and McAndrew, onde o supremo é tomado sobre todas as coberturas abertas, enquanto que o termo *entropia topológica* se refere a nossa definição modificada, onde o supremo é tomado sobre todas as coberturas admissíveis.

Observação 1.2.10. Como no caso da entropia de AKM e da entropia de Kolmogorov-Sinai, o limite na Definição 1.2.9 existe graças à desigualdade

$$N\left(\mathcal{A} \vee \mathcal{B}\right) \leq N\left(\mathcal{A}\right) N\left(\mathcal{B}\right)$$

(veja o Teorema 4.10 de [Wal00]).

Vamos estabelecer algumas propriedades elementares da entropia topológica. O lema a seguir, apesar de muito simples, contém a essência dessas propriedades.

Lema 1.2.11. *Dado um sistema dinâmico* $T : X \to X$*, sejam* $\mathcal{A}$ *e* $\mathcal{B}$ *coberturas de* X *tais que* $\mathcal{B} \prec \mathcal{A}$*. Então, para todo* $k \in \mathbb{N}$*, e todo subconjunto* $Y \subset X$*,*

1. $\mathcal{B}^k \prec \mathcal{A}^k$.
2. $N_Y\left(\mathcal{B}^k\right) \leq N_Y\left(\mathcal{A}^k\right)$.
3. $N_{T^{-1}(Y)}\left(T^{-1}(\mathcal{A})\right) \leq N_Y\left(\mathcal{A}\right)$.
4. $h\big(T \,\big|\, \mathcal{B}\big) \leq h\big(T \,\big|\, \mathcal{A}\big)$.

Demonstração. Os itens (1), (2) e (3) são imediatos. O item (4) é consequência do item (2). □

Como uma simples aplicação do Lema 1.2.11, vamos calcular a entropia topológica de uma rotação do círculo.

Exemplo 1.2.12 (Entropia topológica da rotação)**.** Seja $X = S^1 \subset \mathbb{C}$ o círculo, e

$$\begin{array}{rccc} T: & X & \rightarrow & X \\ & z & \mapsto & z\,\mathrm{e}^{-i\theta} \end{array}$$

a rotação no sentido horário por um ângulo θ. Vamos mostrar que $h(T) = 0$.

Tome uma cobertura aberta $\mathcal{A}$ qualquer. Como todo aberto de X é uma união enumerável de arcos (*intervalos*), podemos assumir que a família $\mathcal{A}$ é composta por arcos, e como X é compacto, também podemos assumir que $\mathcal{A}$ é um conjunto finito. É evidente que existe uma partição finita $\mathcal{C}$, composta por arcos (não necessariamente abertos), tal que $\mathcal{A} \prec \mathcal{C}$. Sendo assim, resta apenas mostrar que $h\big(T \,\big|\, \mathcal{C}\big) = 0$. Denote por $E_{\mathcal{C}}$, o conjunto dos extremos de cada intervalo de $\mathcal{C}$ (veja a Figura 1.5). Neste caso,

$$E_{\mathcal{C}^n} = \bigcup_{j=0}^{n-1} T^{-j}\left(E_{\mathcal{C}}\right).$$

Em especial,

$$N\left(\mathcal{C}^n\right) \leq n \# E_{\mathcal{C}}.$$

Ou seja,

$$h\big(T \,\big|\, \mathcal{C}\big) \leq \lim_{n\to\infty} \frac{1}{n} \log\left(n \# E_{\mathcal{C}}\right) = 0.$$

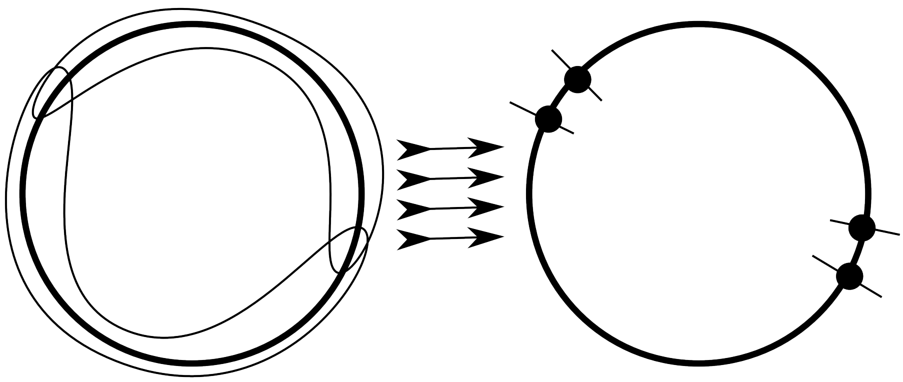

Figura 1.5: A cobertura $\mathcal{A}$ determina os extremos dos intervalos da partição $\mathcal{C}$.

O mesmo argumento pode ser utilizado para mostrar que um sistema dinâmico $T : \mathbb{R} \rightarrow \mathbb{R}$, onde T é um homeomorfismo de $\mathbb{R}$, possui entropia topológica nula. Neste caso, a família $\mathcal{A}$ será uma cobertura admissível, garantindo que $\mathcal{C}$ seja finita.

Lema 1.2.13. *Considere o sistema dinâmico $T : X \to X$, e seja $k \in \mathbb{N}$. Então,*

$$h\left(T^k\right) \leq kh\left(T\right).$$

Demonstração. Seja $\mathcal{A}$ uma cobertura admissível de X. Note que $(\mathcal{A}_T^k)_{T^k}^n = \mathcal{A}_T^{kn}$. Assim,

$$k\frac{1}{kn}H\left(\mathcal{A}_T^{kn}\right) = \frac{1}{n}H\left((\mathcal{A}_T^k)_{T^k}^n\right).$$

Tomando o limite para $n \to \infty$, temos que

$$kh\big(T \bigm| \mathcal{A}\big) = h\Big(T^k \Bigm| \mathcal{A}_T^k\Big).$$

Assim, como $\mathcal{A} \prec \mathcal{A}_T^k$,

$$h\Big(T^k \Bigm| \mathcal{A}\Big) \leq h\Big(T^k \Bigm| \mathcal{A}_T^k\Big) = kh\big(T \bigm| \mathcal{A}\big).$$

Agora, basta tomar o supremo para todas as coberturas admissíveis $\mathcal{A}$ para concluir que

$$h\left(T^k\right) \leq kh\left(T\right).$$

$\square$

Observação 1.2.14. O fato de $\mathcal{A}_T^k$ não ser necessariamente uma família admissível nos impediu de concluir que $h\left(T^k\right) = kh\left(T\right)$. No caso da entropia de Kolmogorov-Sinai, o supremo de $h_\mu\big(T \bigm| \mathcal{C}\big)$ é tomado para todas as partições mensuráveis finitas $\mathcal{C}$. Neste caso, a demonstração do Lema 1.2.13 é facilmente adaptada para mostrar que $h_\mu\left(T^k\right) = kh_\mu\left(T\right)$. Basta notar que, como $\mathcal{C}_T^k$ é mensurável e finita,

$$kh_\mu\big(T \bigm| \mathcal{C}\big) = h_\mu\Big(T^k \Bigm| \mathcal{C}_T^k\Big) \leq h_\mu\left(T^k\right).$$

Para os *sistemas dinâmicos do tipo produto* (Definição 2.1.1), a igualdade $h\left(T^k\right) = kh\left(T\right)$ será consequência do princípio variacional (Corolário 3.1.4).

Se $T : X \to X$ é um sistema dinâmico e $\pi : X \to Y$ é uma aplicação sobrejetiva, podemos "transportar" o sistema dinâmico em X para um sistema dinâmico em Y, de modo que $\pi(x) \in Y$ seja levado em $\pi(Tx)$.

Definição 1.2.15. *Se $T : X \to X$ e $S : Y \to Y$ são sistemas dinâmicos, e $\pi : X \to Y$ é uma aplicação contínua sobrejetiva tal que o diagrama*

$$\begin{array}{ccc} X & \xrightarrow{T} & X \\ {\scriptstyle\pi}\downarrow & & \downarrow{\scriptstyle\pi} \\ Y & \xrightarrow{S} & Y \end{array}$$

comuta, então dizemos que o sistema dinâmico S é semi-conjugado de T pela aplicação π*, ou que S (ou Y) é um* fator *de T (ou X).*

O exemplo não trivial mais simples de semi-conjugação é o fator de um sistema produto.

Exemplo 1.2.16 (Fator de um produto)**.** Sejam $T_1 : X_1 \to X_1$ e $T_2 : X_2 \to X_2$ sistemas dinâmicos. Denotando por $T = T_1 \times T_2$ o produto dos dois sistemas, e por $\pi_i : X_1 \times X_2 \to X_i$, a projeção canônica na primeira coordenada, temos que T_i é semi-conjugado de T pela projeção π_i. A comutatividade do diagrama

$$\begin{tikzcd} X_1 \times X_2 \arrow[r, "T_1 \times T_2"] \arrow[d, "\pi_i"'] & X_1 \times X_2 \arrow[d, "\pi_i"] \\ X_i \arrow[r, "T_i"] & X_i \end{tikzcd}$$

é evidente.

Exemplo 1.2.17 (Quociente de um grupo)**.** Seja $T : G \to G$ um endomorfismo de um grupo de Lie. E seja $H \subset G$ um subgrupo qualquer, invariante por T. Denotando por $\pi : G \to G/H$ a projeção canônica, e por

$$\begin{array}{rccc} S: & G/H & \to & G/H \\ & gH & \mapsto & T(g)H \end{array}$$

o sistema induzido por T no quociente, então S é semi-conjugado de T.

Exemplo 1.2.18. Seja G um grupo de Lie e $H \subset G$ um subgrupo qualquer. Se $S : H \to H$ é um sistema dinâmico qualquer definido em H, e se π_G e π_H são as projeções canônicas de $G \times H$ na primeira e na segunda coordenadas respectivamente, então, definindo o sistema

$$\begin{array}{rccc} T: & G \times H & \to & G \times H \\ & (g, h) & \mapsto & (S(h)g, S(h)) \end{array},$$

temos que π_H faz de S um sistema semi-conjugado de T, mas em geral, π_G não induz uma semi-conugação em G.

Quando $S : Y \to Y$ é semi-conjugado de $T : X \to X$, é de se esperar que a entropia topológica de S nunca seja superior à entropia topológica de T. No caso de sistemas dinâmicos compactos é exatamente o que acontece. Mas para o caso não compacto, precisamos de alguma hipótese adicional.

Proposição 1.2.19. *Se o sitema dinâmico $S : Y \to Y$ é semi-conjugado de $T : X \to X$ por uma aplicação própria π, então*

$$h\,(S) \leq h\,(T).$$

Demonstração. Assim como no caso compacto, a essência da demonstração está no fato de que π^{-1} leva coberturas admissíveis de Y em coberturas admissíveis de X. De fato, coberturas abertas de Y são levadas em coberturas abertas de X, e se $K \subset Y$ é um conjunto compacto, então como π é própria, $\pi^{-1}(K)$ é um conjunto compacto. Assim,

$$\pi^{-1}\,(K^c) = \pi^{-1}(K)^c$$

é o complementar de um compacto (veja a Figura 1.6).

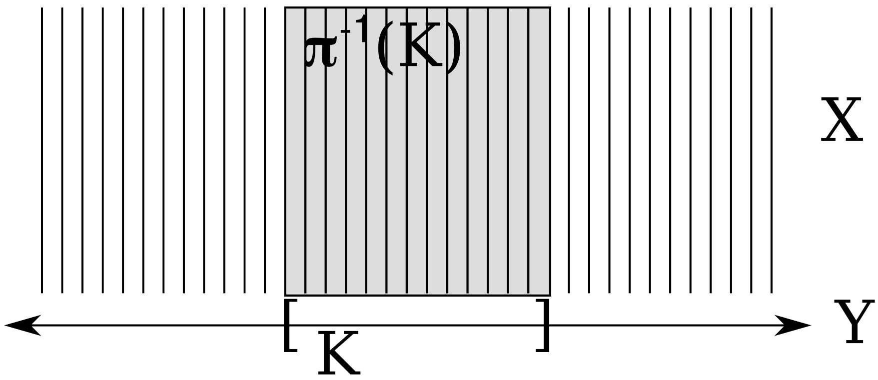

Figura 1.6: $\pi^{-1}(K^c) = \pi^{-1}(K)^c$ é o complemento de um compacto quando π é própria.

Assim, para toda cobertura admissível $\mathcal{A}$ de Y, $\mathcal{D} = \pi^{-1}\,(\mathcal{A})$ é uma cobertura admissível de X. Como Y é semi-conjugado de X, $\pi^{-1}\,(\mathcal{A}^n) = \mathcal{D}^n$. Sabendo que $\mathcal{A}_n \subset \mathcal{A}^n$, então $\mathcal{A}_n$, é uma subcobertura de $\mathcal{A}^n$ se, e somente se, $\pi^{-1}\,(\mathcal{A}_n)$ for uma subcobertura de $\mathcal{D}^n$. Concluímos que

$$N\,(\mathcal{A}^n) = N\,(\mathcal{D}^n).$$

Em particular,

$$h\big(S \bigm| \mathcal{A}\big) = h\big(T \bigm| \mathcal{D}\big) \leq h\,(T).$$

Tomando o supremo em todas as coberturas admissíveis $\mathcal{A}$, chegamos ao resultado desejado. $\square$

O exemplo a seguir mostra que se π não é uma aplicação própria, o resultado da Proposição 1.2.19 pode não ser verdadeiro.

Exemplo 1.2.20. Considere $X = \mathbb{R}$ e seja $T(x) = 2x$. A aplicação T é um homeomorfismo, e portanto, de acordo com o parágrafo final do Exemplo 1.2.12, $h(T) = 0$.

Por outro lado, apesar de $S(z) = z^2$ ser semi-conjugado de T pela aplicação $\pi(x) = \exp(ix)$ (veja a Figura 1.7), é fato conhecido que $h(S) = \log 2$ (veja o Exemplo 1.2.35). Note que como π não é uma aplicação própria, este exemplo não contradiz a Proposição 1.2.19.

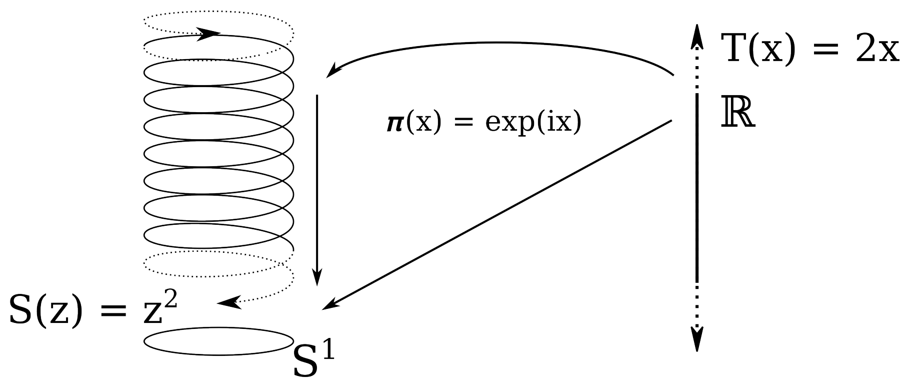

Figura 1.7: A transformação T é bastante simples, e possui entropia topológica nula (veja o comentário ao final do Exemplo 1.2.12). Por outro lado, S possui entropia topológica positiva.

Entropia de Bowen

Bowen introduziu em [Bow71] uma definição de entropia que coincide com a entropia de AKM quando o sistema dinâmico é compacto metrizável. Neste trabalho, não vamos nos restringir à definição que Bowen deu para o caso não-compacto. Ao invés disso, seguimos a ideia de Patrão de restringir nossa atenção a métricas oriundas de alguma compactificação $\widetilde{X}$ de X (veja o conceito de d-entropia na Definição 1.2.23).

O seguinte resultado foi extraído do Teorema 6.1 de [Pat07], e mostra que todo sistema dinâmico separável e metrizável admite uma compactificação metrizável.

Proposição 1.2.21. *Seja $T : X \to X$ um sistema dinâmico onde X é metrizável e separável. Então T admite uma compactificação metrizável $\widetilde{T} : \widetilde{X} \to \widetilde{X}$.*

Demonstração. Por ser metrizável e separável, X é tal que existe uma família enumerável de aplicações

$$\mathcal{F} \subset C(X, I),$$

onde $C(X, I)$ são as aplicações contínuas de X no intervalo $I = [0, 1]$, tal que $\mathcal{F}$ separa pontos e conjuntos fechados. Em particular, a família enumerável

$$\mathcal{F}_T = \left\{ f \circ T^n \,\middle|\, f \in \mathcal{F},\, n = 0, 1, \ldots \right\},$$

além de separar pontos e fechados, é tal que pela Proposição 4.53 de [Fol99], a aplicação

$$\begin{array}{rccl} e_T : & X & \rightarrow & I^{\mathcal{F}_T} \\ & x & \mapsto & (f(x))_{f \in \mathcal{F}_T} \end{array},$$

definida de modo que para cada $f \in \mathcal{F}_T$, a coordenada f de $e_T(x)$ é exatamente $f(x)$, é um homeomorfismo sobre sua imagem. Ou seja, podemos identificar X com o subconjunto $e_T(X)$ do espaço compacto metrizável $I^{\mathcal{F}_T}$. Note que, com esta identificação, a aplicação T age em $e_T(X)$ como um shift. Mais precisamente, T é a restrição a $e_T(X)$ da aplicação

$$\widetilde{T}\left((a_f)_{f \in \mathcal{F}_T}\right) = (a_{f \circ T})_{f \in \mathcal{F}_T},$$

e $\widetilde{X}$ é o fecho de $e_T(X)$. É evidente que $\widetilde{T}$ é contínua. □

Vamos ilustrar como fica a compactificação da Proposição 1.2.21 para um caso bem simples.

Exemplo 1.2.22. Vamos considerar o sistema dinâmico do Exemplo 1.2.5. Seja $X = (0, 1)$, e

$$\begin{array}{rccl} T : & X & \rightarrow & X \\ & x & \mapsto & \begin{cases} x, & x \leq \frac{1}{2} \\ \frac{1}{2}, & x > \frac{1}{2} \end{cases} \end{array}.$$

Tome para $\mathcal{F}_T$, o conjunto $\{i, i \circ T\}$, onde $i : X \rightarrow [0, 1]$ é a inclusão. Note que $i \circ T^2 = i \circ T$. Neste caso, a imagem de e_T é o conjunto

$$e_T(X) = \left\{ (x, T(x)) \,\middle|\, x \in X \right\},$$

que é exatamente o gráfico de T. A compactificação neste caso é igual ao gráfico de T acrescido dos pontos $(0, 0)$ e $\left(1, \frac{1}{2}\right)$ (veja a Figura 1.8).

Escolha uma métrica d em X e, dado $\varepsilon > 0$, denote por

$$\mathcal{B}_d(\varepsilon) = \left\{ B_d(x; \varepsilon) \,\middle|\, x \in X \right\}$$

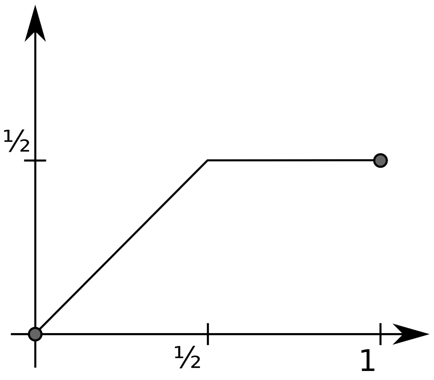

Figura 1.8: Uma compactificação de (0,1) construída como o fecho do gráfico de T. A compactificação consiste em acrescentar ao gráfico, os pontos $(0,0)$ e $\left(1, \frac{1}{2}\right)$.

a família composta pelas bolas de raio ε. E para $n \in \mathbb{N}$, defina

$$d_n(x,y) = \max_{0 \leq j < n} d(T^j x, T^j y).$$

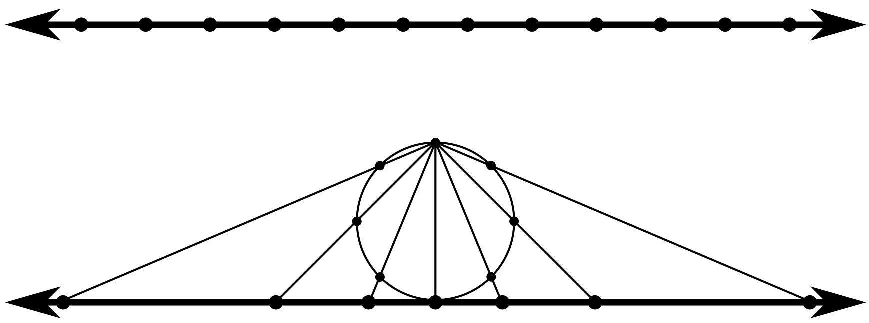

Figura 1.9: O conceito de (n,ε)-separado depende da métrica escolhida. Os $(1,\varepsilon)$-separados maximais de $\mathbb{R}$ são sempre infinitos com a métrica euclidiana, mas são sempre finitos com métricas vindas de uma compactificação.

Definição 1.2.23 (d-Entropia)**.** *Seja $T : X \to X$ um sistema dinâmico e d uma métrica em X. Dado $\varepsilon > 0$, definimos*

$$h^d(T,\varepsilon) = \lim_{n\to\infty} \frac{1}{n} \log N(\mathcal{B}_{d_n}(\varepsilon)),$$

e

$$h^d\left(T\right) = \sup_{\varepsilon>0} h^d\left(T,\varepsilon\right).$$

Denotamos a entropia de Bowen usual por

$$h_d\left(T\right) = \sup_{\substack{\varepsilon>0 \\ K:\ compacto}} \lim_{n\to\infty} \frac{1}{n} \log N_K\left(\mathcal{B}_{d_n}\left(\varepsilon\right)\right).$$

Observação 1.2.24. Exatamente como no caso da entropia topológica e da entropia de Kolmogorov-Sinai, o limite na Definição 1.2.23 existe graças à desigualdade

$$N\left(\mathcal{B}_{d_n}\left(\varepsilon\right)\right) \leq N\left(\mathcal{B}_{d_q}\left(\varepsilon\right)\right) N\left(\mathcal{B}_{d_{n-q}}\left(\varepsilon\right)\right)$$

para todo $q \in \mathbb{N}$ tal que $0 \leq q \leq n$ (veja o Teorema 4.10 de [Wal00]).

Uma outra maneira de caracterizar a entropia de Bowen é através dos *conjuntos separados*. Nossa demonstração do princípio variacional utiliza esta caracterização.

Seja $T : X \to X$ um sistema dinâmico com métrica d. Dados $\varepsilon > 0$ e $n \in \mathbb{N}$, dizemos que o conjunto $S \subset X$ é (n,ε)*-separado* quando todos os pares de pontos distintos $x, y \in S$ forem tais que $d_n(x,y) \geq \varepsilon$. Para um subconjunto $Y \subset X$ e $\varepsilon > 0$, escrevemos $s\left(n,\varepsilon,Y\right)$ para denotar o supremo dentre as cardinalidades dos conjuntos (n,ε)-separados que são subconjuntos de Y.

Lema 1.2.25. *Em um sistema dinâmico $T : X \to X$, onde (X,d) é um espaço métrico, dado $Y \subset X$ e $\varepsilon > 0$, temos que para todo $n \in \mathbb{N}$,*

$$N_Y\left(\mathcal{B}_{d_n}\left(\varepsilon\right)\right) \leq s\left(n,\varepsilon,Y\right) \leq N_Y\left(\mathcal{B}_{d_n}\left(\frac{\varepsilon}{2}\right)\right).$$

Demonstração. A primeira desigualdade segue da afirmação seguinte, juntamente com a existência (através do lema de Zorn) de conjuntos (n,ε)-separados maximais.

Afirmação. *Se $E \subset Y$ é um conjunto (n,ε)-separado maximal, então*

$$\mathcal{B}_E = \left\{ Y \cap B_{d_n}\left(x;\varepsilon\right) \,\middle|\, x \in E \right\}$$

é uma cobertura de Y.

Se $\mathcal{B}_E$ não for uma cobertura, então, tomando $y \in Y \setminus \bigcup_{x\in E} B_{d_n}(x;\varepsilon)$, temos que o conjunto $E \cup \{y\}$ é (n,ε)-separado, infringindo a maximalidade de E.

Para a segunda desigualdade, se $S \subset Y$ é um conjunto (n,ε)-separado, e $\mathcal{B} \subset \mathcal{B}_{d_n}\left(\frac{\varepsilon}{2}\right)$ uma cobertura de Y de cardinalidade $N_Y\left(\mathcal{B}_{d_n}\left(\frac{\varepsilon}{2}\right)\right)$, então, para cada $s \in S$, existe $B(s) \in \mathcal{B}$ com $s \in B(s)$, pois $\mathcal{B}$ cobre Y. Esta correspondência é injetiva, pois se $B(s_1) = B(s_2)$, então $d_n(s_1,s_2) < \varepsilon$. Neste caso, como S é (n,ε)-separado, devemos ter $s_1 = s_2$. Como a aplicação é injetiva, a cardinalidade de S é menor ou igual à cardinalidade de $\mathcal{B}$. $\square$

Proposição 1.2.26. *Para um sistema dinâmico $T : X \to X$, onde (X,d) é um espaço métrico,*

$$h^d(T) = \sup_{\varepsilon>0} \lim_{n\to\infty} \frac{1}{n} \log s(n,\varepsilon,X).$$

Demonstração. É imediato do Lema 1.2.25. Basta tomar o log, dividir por n, tomar o limite em

$$N(\mathcal{B}_{d_n}(\varepsilon)) \leq s(n,\varepsilon,X) \leq N\left(\mathcal{B}_{d_n}\left(\frac{\varepsilon}{2}\right)\right)$$

e depois o supremo para $\varepsilon > 0$. $\square$

Para a definição da d-entropia, utilizamos as famílias $\mathcal{B}_{d_n}(\varepsilon)$. Note que estas famílias não são o mesmo que $[\mathcal{B}_d(\varepsilon)]^n$. O lema a seguir mostra que poderíamos ter utilizado as famílias $[\mathcal{B}_d(\varepsilon)]^n$ para definir a d-entropia, tornando-a mais fácil de comparar com a entropia topológica.

Proposição 1.2.27. *Seja $T : X \to X$ um sistema dinâmico com métrica d. Então,*

$$h^d(T) = \sup_{\varepsilon>0} h\left(T \,\middle|\, \mathcal{B}_d(\varepsilon)\right).$$

Demonstração. É suficiente mostrar que

$$[\mathcal{B}_d(\varepsilon)]^n \prec \mathcal{B}_{d_n}(\varepsilon) \prec \left[\mathcal{B}_d\left(\frac{\varepsilon}{2}\right)\right]^n.$$

De fato, isso implicaria que

$$\frac{1}{n}\log N([\mathcal{B}_d(\varepsilon)]^n) \leq \frac{1}{n}\log N(\mathcal{B}_{d_n}(\varepsilon)) \leq \frac{1}{n}\log N\left(\left[\mathcal{B}_d\left(\frac{\varepsilon}{2}\right)\right]^n\right).$$

Agora, basta fazer $n \to \infty$ e tomar o supremo para $\varepsilon > 0$.

É imediato que

$$\mathcal{B}_{d_n}(\varepsilon) \subset [\mathcal{B}_d(\varepsilon)]^n,$$

pois toda bola na métrica d_n é da forma

$$B_d(x;\varepsilon) \cap \cdots \cap T^{-(n-1)}\left(B_d\left(T^{n-1}x;\varepsilon\right)\right)$$

(veja a Figura 1.10). Por outro lado, um conjunto $A \in \left[\mathcal{B}_d\left(\frac{\varepsilon}{2}\right)\right]^n$ é da forma

$$A = B_d\left(x_0;\frac{\varepsilon}{2}\right) \cap \cdots \cap T^{-(n-1)}\left(B_d\left(x_{n-1};\frac{\varepsilon}{2}\right)\right).$$

Em particular, quando $A \neq \emptyset$, podemos tomar $y \in A$. Neste caso, temos que $d(T^j y, x_j) < \varepsilon$. Assim,

$$B_d\left(x_j;\frac{\varepsilon}{2}\right) \subset B_d\left(T^j y;\varepsilon\right),$$

pois $z \in B_d\left(x_j;\frac{\varepsilon}{2}\right)$ implica que $d(z,T^j y) \leq d(z,x_j) + d(x_j,T^j y) < \varepsilon$. Portanto,

$$A \subset B_d(y;\varepsilon) \cap \cdots \cap T^{-(n-1)}\left(B_d\left(T^{n-1}y;\varepsilon\right)\right) \in \mathcal{B}_{d_n}(\varepsilon),$$

provando que $\left[\mathcal{B}_d\left(\frac{\varepsilon}{2}\right)\right]^n$ refina $\mathcal{B}_{d_n}(\varepsilon)$.

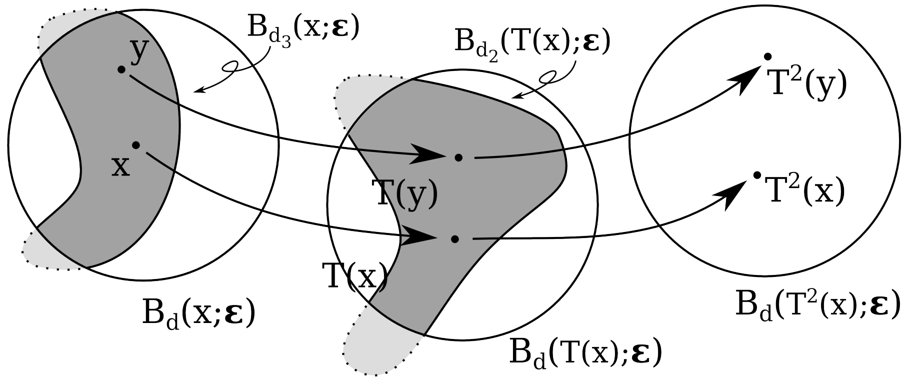

Figura 1.10: Dizer que $d_3(x,y) < \varepsilon$, é o mesmo que dizer que $y \in B_d(x;\varepsilon)$, $T(y) \in B_d(T(x);\varepsilon)$ e $T^2(y) \in B_d(T^2(x);\varepsilon)$. A bola centrada em x e de raio ε na métrica d_3 está representada na figura. Note que $B_{d_3}(x;\varepsilon) = B_d(x;\varepsilon) \cap T^{-1}(B_{d_2}(T(x);\varepsilon))$.

□

Observação 1.2.28. Do ponto de vista deste trabalho, seria mais natural que a d entropia fosse definida como caracterizada na Proposição 1.2.27. Não o fizemos para não divergir da literatura atual. Daqui por diante, usaremos a fórmula

$$h^d(T) = \sup_{\varepsilon>0} h\Big(T \,\Big|\, \mathcal{B}_d(\varepsilon)\Big).$$

como se esta fosse a definição de d-entropia, sem fazer nenhuma referência à Proposição 1.2.27.

Observação 1.2.29. Da mesma forma que na Proposição 1.2.27, a entropia de Bowen também pode ser caracterizada pela equação

$$h_d(T) = \sup_{\substack{\varepsilon>0 \\ K:\ \text{compacto}}} \limsup_{n\to\infty} \frac{1}{n} \log N_K\left(\mathcal{B}_d(\varepsilon)^n\right),$$

pois assim como na Proposição 1.2.27, basta notar que

$$K \cap [\mathcal{B}_d(\varepsilon)]^n \prec K \cap \mathcal{B}_{d_n}(\varepsilon) \prec K \cap \left[\mathcal{B}_d\left(\frac{\varepsilon}{2}\right)\right]^n.$$

O lema a seguir enuncia a existência do chamado *número de Lebesgue* em uma forma fácil de aplicar na construção de refinamentos de coberturas. Essencialmente, toda cobertura admissível pode ser refinada por uma cobertura dada por bolas com um certo raio fixado. Note que pela observação 1.2.8, quando $\widetilde{X}$ é uma compactificação de X, toda cobertura admissível é da forma $X \cap \widetilde{\mathcal{A}}$, onde $\widetilde{\mathcal{A}}$ é uma cobertura aberta de $\widetilde{X}$.

Lema 1.2.30 (Número de Lebesgue)**.** *Seja (X, d) um espaço métrico que admite uma compactificação $\left(\widetilde{X}, \widetilde{d}\right)$. Então, para qualquer cobertura da forma $\mathcal{A} = X \cap \widetilde{\mathcal{A}}$, onde $\widetilde{\mathcal{A}}$ é uma cobertura aberta de $\widetilde{X}$, existe $\varepsilon > 0$ tal que*

$$\mathcal{A} \prec \mathcal{B}_d(\varepsilon).$$

Em particular, o resultado vale para qualquer cobertura admissível de X.

Demonstração. Seja

$$C_\varepsilon = \left\{x \in \widetilde{X} \,\middle|\, \exists A \in \widetilde{\mathcal{A}},\ B_{\widetilde{d}}(x;\varepsilon) \subset A\right\}$$

o conjunto de todos os $x \in \widetilde{X}$ tais que a bola centrada em x com raio ε está contida em algum elemento de $\widetilde{\mathcal{A}}$. Vamos encontrar $\varepsilon > 0$ tal que $\widetilde{X} = C_\varepsilon$.

Afirmação. $\widetilde{X} = \bigcup_{\varepsilon>0} C_\varepsilon$.

Para cada $x \in \widetilde{X}$, existe um $A \in \widetilde{\mathcal{A}}$ que contém x. Como A é aberto, isso significa que existe $\varepsilon > 0$ com $B_{\widetilde{d}}(x;\varepsilon) \subset A$.

Afirmação. $C_\varepsilon \subset \operatorname{int}\left(C_{\frac{\varepsilon}{2}}\right)$.

Basta notar que se $x \in C_\varepsilon$, então

$$B_{\widetilde{d}}\left(x;\frac{\varepsilon}{2}\right) \subset C_{\frac{\varepsilon}{2}}.$$

De fato, se $\widetilde{d}(x,y) < \frac{\varepsilon}{2}$, temos que $B_{\widetilde{d}}\left(y;\frac{\varepsilon}{2}\right) \subset B_{\widetilde{d}}(x;\varepsilon)$. Esta última bola está contida em algum elemento de $\widetilde{\mathcal{A}}$, pois $x \in C_\varepsilon$. Portanto, $y \in C_{\frac{\varepsilon}{2}}$.

Juntando ambas as afirmações, temos que

$$\widetilde{X} = \bigcup_{\varepsilon>0} \operatorname{int}\left(C_\varepsilon\right).$$

Como $\widetilde{X}$ é compacto, existe $\varepsilon > 0$ tal que

$$\widetilde{X} = C_\varepsilon.$$

Tomando a interseção com X,

$$\mathcal{A} = X \cap \widetilde{\mathcal{A}} \prec X \cap \mathcal{B}_{\widetilde{d}}(\varepsilon).$$

Agora, basta observar que

$$\mathcal{B}_d\left(\varepsilon\right) \subset X \cap \mathcal{B}_{\widetilde{d}}(\varepsilon),$$

pois

$$B_d\left(x;\varepsilon\right) = X \cap B_{\widetilde{d}}(x;\varepsilon).$$

E portanto,

$$\mathcal{A} \prec \mathcal{B}_d\left(\varepsilon\right).$$

□

Finalmente, relacionamos a entropia topológica e a d-entropia.

Proposição 1.2.31. *Seja $T : X \to X$ um sistema dinâmico e $\widetilde{X}$ uma compactificação metrizável de X. Se d é a restrição a X de uma métrica $\widetilde{d}$ em $\widetilde{X}$, então*

$$h^d\left(T\right) = \sup_{\widetilde{\mathcal{A}}} h\left(T \,\middle|\, X \cap \widetilde{\mathcal{A}}\right),$$

onde o supremo é tomado sobre todas as coberturas abertas $\widetilde{\mathcal{A}}$ de $\widetilde{X}$.

Demonstração. A família $\mathcal{D} = \left\{ B_{\widetilde{d}}(x;\varepsilon) \,\middle|\, x \in X \right\}$ cobre $\widetilde{X}$, e é tal que $\mathcal{B}_d(\varepsilon) = X \cap \mathcal{D}$. O que implica que

$$h^d(T) \leq \sup_{\widetilde{\mathcal{A}}:\text{ abertos}} h\left(T \,\middle|\, X \cap \widetilde{\mathcal{A}}\right).$$

Por outro lado, para uma cobertura aberta $\widetilde{\mathcal{A}}$ qualquer de $\widetilde{X}$, podemos usar o Lema 1.2.30 para obter $\varepsilon > 0$ tal que

$$X \cap \widetilde{\mathcal{A}} \prec \mathcal{B}_d(\varepsilon).$$

Isso significa que

$$h\left(T \,\middle|\, X \cap \widetilde{\mathcal{A}}\right) \leq h\left(T \,\middle|\, \mathcal{B}_d(\varepsilon)\right).$$

Tomando o supremo em ε, concluímos que

$$h\left(T \,\middle|\, X \cap \widetilde{\mathcal{A}}\right) \leq h^d(T).$$

Tomando o supremo para todas as coberturas abertas $\widetilde{\mathcal{A}}$ de $\widetilde{X}$,

$$\sup_{\widetilde{\mathcal{A}}:\text{ abertos}} h\left(T \,\middle|\, X \cap \widetilde{\mathcal{A}}\right) \leq h^d(T).$$

□

Quando o sistema dinâmico é compacto, sabemos que a d-entropia independe da métrica d escolhida. O seguinte corolário da Proposição 1.2.31 estende este resultado.

Corolário 1.2.32. *Seja $T : X \to X$ um sistema dinâmico com X admitindo uma compactificação metrizável $\widetilde{X}$. Se d e c são a restrição a X de métricas $\widetilde{d}$ e $\widetilde{c}$ em $\widetilde{X}$, então*

$$h^d(T) = h^c(T).$$

Na demonstração a seguir, vamos utilizar uma técnica semelhante à demonstração do Lema 1.2.30 para mostrar que independente da métrica d, escolhida para X, dada uma cobertura admissível $\mathcal{A}$, sempre existe $\varepsilon > 0$ tal que $\mathcal{A} \prec \mathcal{B}_d(\varepsilon)$.

Proposição 1.2.33. *Seja $T : X \to X$ um sistema dinâmico com uma métrica d. Então,*

$$h(T) \leq h_d(T) \leq h^d(T).$$

Demonstração. A entropia de Bowen será calculada conforme descrito na Observação 1.2.29. Dada uma cobertura admissível $\mathcal{A}$, tome $K \subset X$ tal que $K^c \in \mathcal{A}$. Vamos utilizar uma técnica semelhante à demonstração do Lema 1.2.30. Seja

$$C_\varepsilon = \left\{ x \in K \;\middle|\; \exists A \in \mathcal{A},\ B_d(x; 2\varepsilon) \subset A \right\}$$

o conjunto de todos os $x \in K$ tais que a bola centrada em x com raio 2ε está contida em algum elemento de $\mathcal{A}$. Assim como no Lema 1.2.30, podemos tomar $\varepsilon > 0$ tal que $K = C_\varepsilon$.

Afirmação. $\mathcal{A} \prec \mathcal{B}_d(\varepsilon)$.

Se $B \in \mathcal{B}_d(\varepsilon)$ não intersecta K, então, $B \subset K^c$. E se B intersecta K, então tomando $x \in B \cap K$, pela escolha de ε, existe $A_x \in \mathcal{A}$ tal que

$$B \subset B_d(x; 2\varepsilon) \subset A_x.$$

Ou seja, $\mathcal{A} \prec \mathcal{B}_d(\varepsilon)$.

Assim, definindo

$$\mathcal{D} = \mathcal{B}_d(\varepsilon) \cup \{K^c\},$$

temos que $\mathcal{A} \prec \mathcal{D}$. Vamos particionar X em $K_0, \ldots, K_{n-1}$ e $\widetilde{K}$, com

$$\widetilde{K} = K^c \cap T^{-1}(K^c) \cap \cdots \cap T^{-(n-1)}(K^c) \in \mathcal{D}^n$$

e, para $m = 0, \ldots, n-1$,

$$K_m = K^c \cap T^{-1}(K^c) \cap \cdots \cap T^{-(m-1)}(K^c) \cap T^{-m}(K).$$

(veja a Figura 1.11). Note que, de fato, $X = \widetilde{K} \cup \bigcup_{m=0}^{n-1} K_m$. Assim,

$$N(\mathcal{D}^n) \leq N_{\widetilde{K}}(\mathcal{D}^n) + \sum_{m=0}^{n-1} N_{K_m}(\mathcal{D}^n).$$

Como $\widetilde{K} \in \mathcal{D}^n$,

$$N_{\widetilde{K}}(\mathcal{D}^n) = 1.$$

Para $m = 0, \ldots, n-1$, denote por $\mathcal{B}_m$ a família de conjuntos da forma

$$A_0 \cap T^{-1}(A_1) \cap \cdots \cap T^{-(n-1)}(A_{n-1}),$$

onde $A_j = K^c$ para $j = 0, \ldots, m-1$, e $A_j \in \mathcal{B}_d(\varepsilon)$ para $j \geq m$. De forma mais abreviada,

$$\mathcal{B}_m = K^c \cap T^{-1}(K^c) \cap \cdots \cap T^{-(m-1)}(K^c) \cap T^{-m}\left(\mathcal{B}_d(\varepsilon)^{n-m}\right) \subset \mathcal{D}^n.$$

Se $\mathcal{F} \subset \mathcal{B}_d(\varepsilon)^{n-m}$ é uma cobertura de K, então,

$$K^c \cap T^{-1}(K^c) \cap \cdots \cap T^{-(m-1)}(K^c) \cap T^{-m}(\mathcal{F}) \subset \mathcal{B}_m$$

é uma cobertura de K_m. Assim,

$$\begin{aligned} N_{K_m}(\mathcal{D}^n) &\leq N_{K_m}(\mathcal{B}_m) \\ &\leq N_K\left(\mathcal{B}_d(\varepsilon)^{n-m}\right) \\ &\leq N_K\left(\mathcal{B}_d(\varepsilon)^n\right). \end{aligned}$$

E portanto,

$$\begin{aligned} N(\mathcal{A}^n) &\leq N(\mathcal{D}^n) \\ &\leq N_{\widetilde{K}}(\mathcal{D}^n) + \sum_{m=0}^{n-1} N_{K_m}(\mathcal{D}^n) \\ &\leq 1 + \sum_{m=0}^{n-1} N_{K_m}(\mathcal{D}^n) \\ &\leq (n+1) N_K\left(\mathcal{B}_d(\varepsilon)^n\right). \end{aligned}$$

Tomando o logarítmo, dividindo por n, e fazendo $n \to \infty$, chegamos em

$$\begin{aligned} h\big(T \bigm| \mathcal{A}\big) &\leq \limsup_{n\to\infty} \frac{1}{n}\left(\log(n+1) + \log N_K\left(\mathcal{B}_d(\varepsilon)^n\right)\right) \\ &= \limsup_{n\to\infty} \frac{1}{n} \log N_K\left(\mathcal{B}_d(\varepsilon)^n\right) \\ &\leq h_d(T), \end{aligned}$$

pois $\frac{1}{n}\log(n+1) \to 0$. Tomando o supremo em $\mathcal{A}$,

$$h(T) \leq h_d(T).$$

E é evidente que $h_d(T) \leq h^d(T)$.

$\square$

Observação 1.2.34. Quando as famílias $\mathcal{B}_d(\varepsilon)$ formam coberturas admissíveis, é evidente que

$$h^d(T) \leq h(T).$$

No entanto, mesmo quando d é a restrição de uma métrica proveniente de alguma compactificação do sistema, pode acontecer que $\mathcal{B}_d(\varepsilon)$ não seja admissível. Como exemplo, tome $\varepsilon \in (0, \frac{1}{2})$, e considere $\mathcal{B}_d(\varepsilon)$ em $(0,1)$ com a métrica euclideana. Nenhum intervalo em $\mathcal{B}_d(\varepsilon)$ terá complemento compacto, e portanto, $\mathcal{B}_d(\varepsilon)$ não é admissível (veja a Figura 1.12).

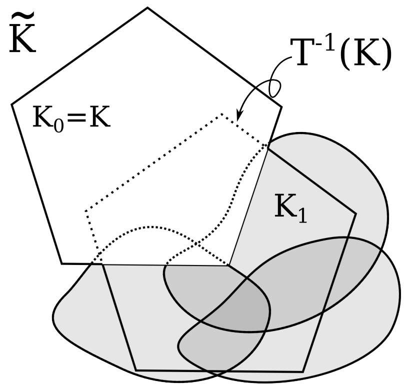

Figura 1.11: Elementos da família $\mathcal{B}_1$ que cobrem $K_1 = K^c \cap T^{-1}(K)$. Aqui, $n = 2$. Cada elemento da cobertura é a interseção de K^c com a imagem inversa de bolas de raio ε. O conjunto X é a união disjunta de K_0, K_1 e $\widetilde{K}$.

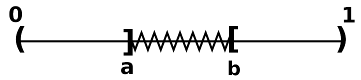

Figura 1.12: Nenhum intervalo (a, b) contido propriamente em $(0, 1)$ possui complemento compacto, pois $(a, b)^c = (0, a] \cup [b, 1)$ só é compacto se for vazio, ou seja, se $a = 0$ e $b = 1$.

Exemplo 1.2.35 (Entropia de z^2)**.** Vamos mostrar que a entropia de

$$\begin{array}{rccc} f: & S^1 & \to & S^1 \\ & z & \mapsto & z^2 \end{array}$$

é $\log 2$. Como S^1 é compacto, a Proposição 1.2.33 juntamente com a Observação 1.2.34 mostram que basta calcularmos a d-entropia de f para uma métrica d qualquer. Vamos tomar para d o comprimento de arco, ou seja, o ângulo em radianos. Para tanto, a Proposição 1.2.26 nos permite fazer o cálculo utilizando os valores de $s\left(n, \varepsilon, S^1\right)$. O efeito da transformação f é o de dobrar o comprimento de arco para pontos suficientemente próximos. Assim, o valor de $s\left(n, \varepsilon, S^1\right)$ praticamente dobra a cada iteração do sistema. De fato, particionando o S^1 em arcos de comprimento $\varepsilon > 0$, podemos ver que $s\left(n, \varepsilon, S^1\right) = \left\lfloor 2^n \frac{2\pi}{\varepsilon} \right\rfloor$, a parte inteira de $2^n \frac{2\pi}{\varepsilon}$ (veja a Figura

1.13). Assim,

$$\frac{2^{n+1}\pi}{\varepsilon} - 1 \leq s\left(n, \varepsilon, S^1\right) \leq \frac{2^{n+1}\pi}{\varepsilon}.$$

E portanto,

$$\lim \frac{1}{n} \log\left(\frac{2^{n+1}\pi}{\varepsilon} - 1\right) \leq h\left(f\right) \leq \lim \frac{1}{n} \log\left(\frac{2^{n+1}\pi}{\varepsilon}\right).$$

Ou seja, $h\left(f\right) = \log 2$.

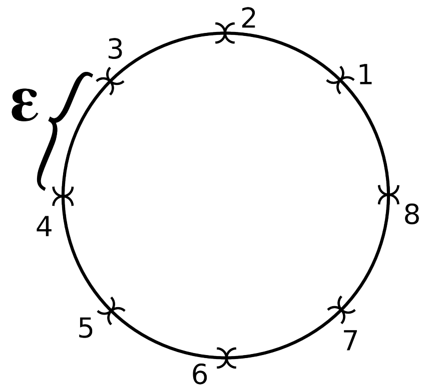

Figura 1.13: Para estimar o valor de $s\left(n, \varepsilon, S^1\right)$, basta dividir o círculo em arcos de comprimento ε. No desenho, $s\left(1, \varepsilon, S^1\right) = 8$, mas seria 7, se ε fosse um pouco maior que $\frac{2\pi}{8}$.

CAPÍTULO 2

Sistemas Dinâmicos *do Tipo Produto*

Neste trabalho, estamos interessados em uma generalização de sistemas dinâmicos metrizáveis localmente compactos, que chamamos de sistemas dinâmicos *do tipo produto*. Neste capítulo, apresentamos a definição dos sistemas *do tipo produto* e demonstramos algumas propriedades que nos permitirão estabelecer a validade do *princípio variacional* para estes sistemas.

2.1 Definição

Considere o espaço $X = \prod_{i=1}^{\infty} X_j$, onde X_j são espaços topológicos. Para simplificar, vamos escrever simplesmente $X = \prod X_j$. Denotamos por X^n o espaço $X_1 \times \cdots \times X_n$, e por π_n a projeção canônica

$$\begin{array}{rccc} \pi_n : & X & \to & X^n \\ & (x_j) & \mapsto & (x_1, \ldots, x_n) \end{array} .$$

Dizemos que um conjunto $A \subset X$ é π_n-mensurável quando existe um boreliano $B \subset X^n$ tal que $A = \pi_n^{-1}(B)$. Dizemos que uma aplicação $f : X \to Y$ é π_n-mensurável quando para todo conjunto mensurável $A \subset Y$, $f^{-1}(A)$ for π_n-mensurável.

Definição 2.1.1 (Sistema Dinâmico do Tipo Produto). *Seja $T : X \to X$ um sistema dinâmico com $X = \prod X_j$. Dizemos que $T : X \to X$ é um* sistema dinâmico do tipo produto *quando*

1. *Cada X_j é localmente compacto separável.*

2. *Para todo $n \in \mathbb{N}$ existe $m = m(n) = m_T(n) \in \mathbb{N}$, com $m \geq n$, tal que $\pi_n \circ T$ é π_m-mensurável.*

Cada X_j admite ao menos uma compactificação metrizável: a compactificação por um ponto. Para um sistema dinâmico do tipo produto $T : X \to X$, vamos supor que temos fixada uma compactificação metrizável $\widetilde{X}_j$ de cada X_j. Vamos denotar por

$$\widetilde{X} = \prod \widetilde{X}_j$$

a compactificação de X formada pelo produto das compactificações de X_j. Nem sempre $\widetilde{X}_j$ será tomada como sendo a compactificação por um ponto de X_j. Por vezes, queremos que o sistema T possa ser estendido a

$$\widetilde{T} : \widetilde{X} \to \widetilde{X}$$

(veja a Proposição 1.2.21).

Um sistema $T : X \to X$, onde X é localmente compacto separável, é isomorfo ao sistema do tipo produto cujo isomorfismo é dado pela identificação entre X e $X \times \{0\}^{\mathbb{N}}$. O produto enumerável de sistemas localmente compactos separáveis $T_j : X_j \to X_j$, denotado por $\prod T_j$ é um exemplo de sistema do tipo produto. Um outro exemplo, é o *shift*

$$\begin{array}{rccc} T: & Z^{\mathbb{N}} & \to & Z^{\mathbb{N}} \\ & (z_n) & \mapsto & (z_{n+1}) \end{array},$$

onde $X_j = Z$ é localmente compacto separável (veja o Exemplo 2.1.2). Note que a classe de sistemas dinâmicos do tipo produto é fechada por composição. Ou seja, se $T : X \to X$ e $S : X \to X$ são do tipo produto, então, $T \circ S$ também é do tipo produto. Neste caso, podemos tomar $m_{T \circ S} = m_T \circ m_S$.

Exemplo 2.1.2 (O *shift*)**.** Seja Z um espaço topológico localmente compacto e separável. Então o shift

$$\begin{array}{rccc} T: & Z^{\mathbb{N}} & \to & Z^{\mathbb{N}} \\ & (z_n) & \mapsto & (z_{n+1}) \end{array}$$

é um sistema dinâmico do tipo produto. De fato, se $\widetilde{Z}$ é uma compactificação qualquer de Z, então o shift é fácilmente estendido para

$$\begin{array}{rccc} \widetilde{T}: & \widetilde{Z}^{\mathbb{N}} & \to & \widetilde{Z}^{\mathbb{N}} \\ & (z_n) & \mapsto & (z_{n+1}) \end{array},$$

e para $m(n)$, basta tomar $m(n) = n + 1$, pois para todo boreliano $B \subset Z^n$,

$$T^{-1}\left(\pi_n^{-1}(B)\right) = \pi_{n+1}^{-1}(Z \times B).$$

A instância não trivial mais simples para esta construção, é quando $\widetilde{Z} = Z \cup \{\infty\}$ é uma compactificação por um ponto de Z. Note que esta construção *não* é a compactificação por um ponto de $Z^{\mathbb{N}}$.

Exemplo 2.1.3 (Derivação de séries formais)**.** Considere o conjunto das séries formais reais $\left\{\sum_{j=0}^{\infty} a_j x^j \,\middle|\, a_j \in \mathbb{R}\right\}$, que pode ser identificado canonicamente com $\mathbb{R}^\infty$. A derivada formal, dada por

$$\frac{d}{dx}\sum_{j=0}^{\infty} a_j x^j = \sum_{j=0}^{\infty}(j+1)a_{j+1}x^j$$

é um sistema dinâmico do tipo produto. De fato, a derivada formal é a composição do shift com o sistema $\prod T_j$, onde

$$T_j a_j x^j = j a_j x^j.$$

As derivadas formais de ordens mais elevadas também são sistemas dinâmicos do tipo produto, pois são composições das derivadas formais de primeira ordem.

Exemplo 2.1.4 (Produto enumerável de *sistemas do tipo produto*)**.** Seja $T_i : X_i \to X_i$, $(i = 0, 1, \dots)$ uma família de sistemas dinâmicos do tipo produto. Então, o sistema $T = T_0 \times T_1 \times \cdots$ também é do tipo produto. De fato, escrevendo

$$X_i = X_{i,1} \times X_{i,2} \times \dots,$$

temos que $X = X_0 \times X_1 \times \dots$ é produto dos enumeráveis fatores $X_{i,j}$ que, por hipótese, satisfazem o item (1) da definição, pois cada T_i é do tipo produto.

Para verificar o item (2), escreva X como um produto dos fatores $X_{i,j}$ ordenado como na Figura 2.1. O k-ésimo termo do produto é da forma $X_{i,j}$, tal que

$$k = k(i,j) = \frac{(i+j+1)(i+j)}{2} - i.$$

De fato, usando a Figura 2.1 como referência, o termo $\frac{(i+j+1)(i+j)}{2}$ equivale ao número de fatores contidos no triângulo. Enquanto que i é o número de termos que "faltam" para preencher o triângulo. O número de "linhas" envolvidas é $i + j$. Ou seja, são as linhas de 0 a $i + j - 1$.

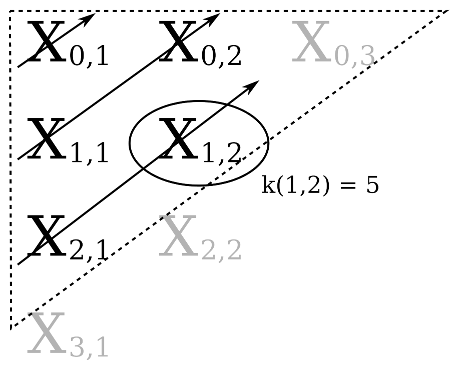

Figura 2.1: Escrevendo X como um produto ordenado dos fatores $X_{i,j}$.

Vamos escrever $m(n)$ para $m_T(n)$ e $m_j(n)$ para $m_{T_j}(n)$. Seja $\pi_{k(i,j)}$ a projeção de X nos $k = k(i,j)$ primeiros termos $X_{0,1} \times X_{1,1} \times \ldots \times X_{i,j}$. e $\pi_{i,j}$ a projeção de X nos j primeiros termos $X_{i,1} \times \ldots \times X_{i,j}$. Vamos encontrar $m(k) = m(i,j)$ tal que $\pi_{k(i,j)} \circ T$ seja $\pi_{m(i,j)}$-mensurável. Note que, reordenando os fatores,

$$\pi_{k(i,j)} = \pi_{0,b_0} \times \cdots \times \pi_{i+j-1,b_{i+j-1}}$$

para b_k adequados. Na Figura 2.1, por exemplo, reordenando os fatores,

$$\pi_5 = \pi_{0,2} \times \pi_{1,2} \times \pi_{2,1}.$$

Neste caso específico, pela Figura 2.2 vemos que para borelianos $B_0 \subset X_0$, $B_1 \subset X_1$ e $B_2 \subset X_2$,

$$\pi_5^{-1}(B_0 \times B_1 \times B_2) = \pi_{0,2}^{-1}(B_0) \times \pi_{1,2}^{-1}(B_1) \times \pi_{2,1}^{-1}(B_2),$$

e portanto,

$$\begin{aligned}(\pi_5 \circ T)^{-1}(B_0 \times B_1 \times B_2) &= \\ &= (\pi_{0,2} \circ T)^{-1}(B_0) \times (\pi_{1,2} \circ T)^{-1}(B_1) \times (\pi_{2,1} \circ T)^{-1}(B_2)\end{aligned}$$

é π_{19}-mensurável, pois cada um dos $(\pi_{a,b_a} \circ T)^{-1}(B_a)$ é $\pi_{a,m_a(b_a)}$-mensurável, e $19 \geq k(a, m_a(b_a))$. Como os conjuntos da forma $B_0 \times B_1 \times B_2$ geram os borelianos de $X_{0,1} \times \cdots \times X_{1,2}$, contra-domínio de $\pi_5 \circ T$, temos que $\pi_5 \circ T$ é π_{19}-mensurável.

Para o caso geral, basta escolhermos $m(i,j) = k(i',j')$ para i' e j' tais que a sequência $X_{0,1}, X_{1,1}, \ldots, X_{i',j'}$ inclua todos os termos $X_{s,t}$ com $s =$

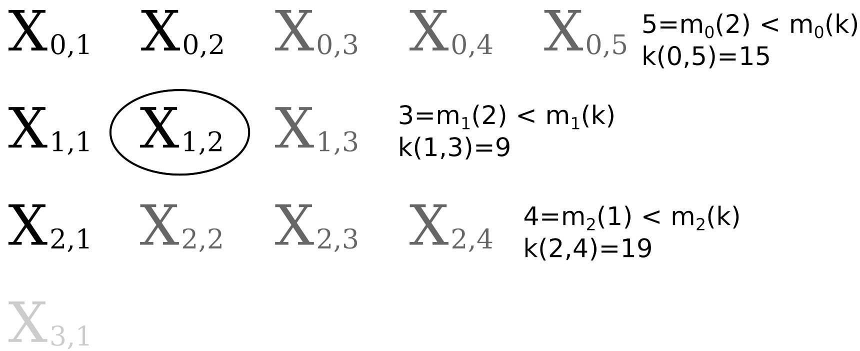

Figura 2.2: Fatores envolvidos na mensurabilidade de $\pi_{k(i,j)} \circ T$. No desenho, $m_0(2) = 5$, $m_1(2) = 3$ e $m_2(1) = 4$. Portanto, basta fazer $m(i,j) = k(2,4) = 19$.

$0, \dots, i+j-1$ e $t = 1, \dots, m_s(b_s)$ (veja a Figura 2.2). Por exemplo, podemos escolher

$$m(i,j) = \max \left\{ k\left(s, m_s(b_s)\right) \;\middle|\; s = 0, 1, \dots, i+j-1 \right\}.$$

Neste caso, para todo conjunto $\pi_{k(i,j)}$-mensurável da forma

$$\begin{aligned} B &= \pi_{0,b_0}^{-1}(B_0) \cap \pi_{1,b_1}^{-1}(B_1) \cap \cdots \cap \pi_{i,b_i}^{-1}(B_i) \\ &= \pi_{k(i,j)}^{-1}(B_0 \times \cdots \times B_i) \end{aligned}$$

temos que $T^{-1}(B)$ é $\pi_{m(i,j)}$-mensurável. De fato, para cada $s = 0, \dots, i+j-1$,

$$T^{-1}\left(\pi_{s,b_s}^{-1}(B_s)\right) = \pi_{s,m_s(b_s)}^{-1}(C_s)$$

para algum $C_s \in X_{s,1} \times \cdots \times X_{s,m_s(b_s)}$ mensurável. Mas esses conjuntos são todos $\pi_{m(i,j)}$ mensuráveis. Como a família dos conjuntos $\pi_{k(i,j)}$-mensuráveis é a σ-álgebra gerada pelos conjuntos da forma de B, temos que $\pi_{k(i,j)} \circ T$ é $\pi_{m(i,j)}$-mensurável.

2.2 Propriedades dos Sistemas *do Tipo Produto*

Vamos mostrar algumas propriedades que os sistemas dinâmicos tipo produto possuem. Primeiro, vamos constatar algumas propriedades elementares, como a metrizabilidade do sistema.

Os sistemas dinâmicos do tipo produto são definidos sobre o espaço $X = \prod_j X_j$, onde cada X_j possui uma compactificação metrizável $\widetilde{X}_j$. A proposição a seguir, cuja demonstração será omitida, nos da uma métrica para a compactificação $\widetilde{X}$ que nos será útil posteriormente.

Proposição 2.2.1. *Seja $T : X \to X$ um sistema dinâmico do tipo produto. Então existe uma métrica $\widetilde{d}_j$ para cada $\widetilde{X}_j$ tal que $\widetilde{X}_j$ tem diâmetro menor que* 1 *e*

$$\widetilde{d}\,((x_j),(y_j)) = \sup_{j=1}^{\infty} \frac{1}{j} \widetilde{d}_j(x_j, y_j)$$

é compatível com a topologia produto em $\widetilde{X} = \prod \widetilde{X}_j$.

Vamos chamar a métrica $\widetilde{d}$ da Proposição 2.2.1 (ou sua restrição a X) de *métrica do tipo produto.*

Lema 2.2.2. *Seja $B \subset X^n$ um boreliano. Então,*

$$\partial\left(\pi_n^{-1}(B)\right) = \pi_n^{-1}(\partial B).$$

Demonstração.

Afirmação.

$$\overline{\pi_n^{-1}(B)} = \pi_n^{-1}(\overline{B}).$$

Como π_n é contínua,

$$\overline{\pi_n^{-1}(B)} \subset \pi_n^{-1}(\overline{B}).$$

Por outro lado, se $(x_j) \in \pi_n^{-1}(\overline{B})$, então existe $(b_1^k, \ldots, b_n^k) \in B$ convergindo para $(x_1, \ldots, x_n)$. Fazendo $b_j^k = x_j$ para $j > n$, temos que $(b_j^k) \in \pi_n^{-1}(B)$ converge para x. Isso implica que $x \in \overline{\pi_n^{-1}(B)}$. Ou seja,

$$\pi_n^{-1}(\overline{B}) \subset \overline{\pi_n^{-1}(B)}.$$

Usando a afirmação anterior para B e para B^c, segue que

$$\begin{aligned}\partial\left(\pi_n^{-1}(B)\right) &= \overline{\pi_n^{-1}(B)} \cap \overline{\pi_n^{-1}(B^c)} \\ &= \pi_n^{-1}(\overline{B}) \cap \pi_n^{-1}(\overline{B^c}) \\ &= \pi_n^{-1}(\partial B).\end{aligned}$$

□

Definição 2.2.3 (Medida de Radon)**.** *Uma medida finita μ, definida nos borelianos de um espaço topológico X, é chamada de* medida de Radon *quando para todo boreliano A,*

$$\mu(A) = \sup_{\substack{K \subset A \\ compacto}} \mu(K).$$

É fato conhecido que toda medida finita nos borelianos de um espaço compacto metrizável é de Radon (Proposição A.4.3).

Proposição 2.2.4. *Seja $T : X \to X$ um sistema dinâmico do tipo produto. Então X é um boreliano de $\widetilde{X}$. Em particular, todas as medidas nos borelianos de X são medidas de Radon.*

Demonstração.

Afirmação. *Cada X^n é um boreliano de $\widetilde{X}^n$.*

Como X^n é um espaço de Haudorff localmente compacto separável, existe uma base enumerável para a topologia composta por vizinhanças com fecho compacto. Em particular, X^n pode ser escrito como uma união enumerável de compactos. Por sua vez, os compactos de X^n são também compactos em $\widetilde{X}^n$. Por ser uma união enumerável de compactos de $\widetilde{X}^n$, X^n é um boreliano de $\widetilde{X}^n$.

Utilizando a afirmação anterior, temos que os conjuntos $\widetilde{\pi}_n^{-1}(X^n)$ são borelianos de $\widetilde{X}$. Como

$$\widetilde{\pi}_n^{-1}(X^n) \downarrow X,$$

segue que X é também um boreliano de $\widetilde{X}$.

A topologia de X é dada por conjuntos da forma $A \cap X$, com $A \subset \widetilde{X}$ aberto. Assim, os borelianos de X também serão conjuntos da forma $B \cap X$, onde $B \subset \widetilde{X}$ é um boreliano de $\widetilde{X}$. Em particular, toda medida μ definida nos borelianos de X pode ser estendida a uma medida nos borelianos de $\widetilde{X}$ dada por

$$\widetilde{\mu}(B) = \mu(B \cap X).$$

Como $\widetilde{\mu}$ é uma medida de Radon e como os borelianos de X são também borelianos de $\widetilde{X}$, temos que todos eles podem ser aproximados "por dentro" por conjuntos compactos de $\widetilde{X}$. Agora, basta notar que os compactos de $\widetilde{X}$ contidos em X são também compactos de X. □

A principal dificuldade em se estudar sistemas dinâmicos que não são localmente compactos decorre do fato de que o Teorema de Representação de Riesz não pode ser utilizado diretamente. No restante desta seção, vamos afirmar e demonstrar propriedades dos sistemas dinâmicos do tipo produto que se assemelham àquelas garantidas pelo Teorema de Representação de Riesz, que nos permitirão demonstrar a validade do *princípio variacional* de maneira semelhante à demonstração dada por Misiurewicz (veja Teorema 8.6 de [Wal00]).

Lema 2.2.5. *As aplicações $f : X \to \mathbb{R}$ que são π_n-mensuráveis estão em bijeção com as aplicações mensuráveis $g : X^n \to \mathbb{R}$ através da correspondência*

$$g \mapsto g \circ \pi_n.$$

Demonstração. Como evidentemente $g \circ \pi_n$ é π_n-mensurável, a parte não trivial é a demonstração de que qualquer aplicação π_n-mensurável $f : X \to \mathbb{R}$ é da forma $g \circ \pi_n$ para alguma aplicação mensurável g.

A aplicação g fica bem definida através da relação $g(\pi_n(x)) = f(x)$. De fato, se $\pi_n(x) = \pi_n(y)$, então, fazendo $\alpha = f(x)$, como f é π_n-mensurável, existe um boreliano $B \subset X^n$, tal que

$$f^{-1}(\alpha) = \pi_n^{-1}(B).$$

Isso implica que $\pi_n(y) = \pi_n(x) \in B$. E portanto, $f(y) = \alpha$. Note também, que $g : X^n \to \mathbb{R}$ é mensurável. De fato, como f é π_n-mensurável, dado um boreliano $A \subset \mathbb{R}$, existe um boreliano $B \subset X^n$, tal que

$$\pi_n^{-1}\left(g^{-1}(A)\right) = f^{-1}(A) = \pi_n^{-1}(B).$$

Como π_n é sobrejetiva, isso implica que $g^{-1}(A) = B$. $\square$

Exemplo 2.2.6. Vamos examinar o Lema 2.2.5 para o caso específico das aplicações π_1-mensuráveis

$$f : \{0,1\}^{\mathbb{N}} \to \mathbb{R}.$$

As aplicações $g : \{0,1\} \to \mathbb{R}$, podem ser associadas a $f_g = g \circ \pi_1$. Neste caso,

$$f_g : \begin{array}{ccc} \{0,1\}^{\mathbb{N}} & \to & \mathbb{R} \\ (x_n)_{n\in\mathbb{N}} & \mapsto & g(x_1) \end{array}.$$

É evidente que $f_g = g \circ \pi_1$ é π_1-mensurável.

O que o Lema 2.2.5 afirma, é que toda aplicação π_1-mensurável $f : \{0,1\}^{\mathbb{N}} \to \mathbb{R}$ assume no máximo dois valores, e que o valor de $f(x)$ depende apenas de $x_1 = \pi_1(x)$. Neste caso, associamos a f, a aplicação $g = g_f$ dada, por exemplo, por

$$g(\alpha) = f(\alpha, 0, 0, \dots).$$

O Lema 2.2.7 é utilizado na demonstração do princípio variacional 3.1.2. No Lema, a proposição A.4.1 é utilizada na construção de uma medida T-invariante. No contexto dos sistemas dinâmicos do tipo produto, existem dois obstáculos a serem transpostos. Primeiramente, $X = \prod X_j$ não é localmente compacto, mas sim o produto enumerável de espaços de Hausdorff localmente compactos.

Lema 2.2.7. *Seja $T : X \to X$ um sistema dinâmico do tipo produto e σ_k uma sequência de medidas de probabilidade nos borelianos de X. Então, fazendo*

$$\mu_k = \frac{1}{k} \sum_{j=0}^{k-1} \sigma_k \circ T^{-j},$$

temos que existe uma medida μ nos borelianos de X tal que

1. *$0 \leq \mu(X) \leq 1$.*
2. *A medida μ é T-invariante.*
3. *Para $n \in \mathbb{N}$, se $C \subset X$ é π_n-mensurável e satisfaz $\mu(\partial C) = 0$, então*

$$\liminf_{k\to\infty} \mu_k(C) \leq \mu(C).$$

Demonstração. Note que cada X^n é um espaço de Hausdorff localmente compacto e metrizável. E portanto, pela Proposição A.4.1, o conjunto das medidas positivas em X^n de medida total menor ou igual a 1 é compacto na topologia fraca-$*$. Vamos utilizar este fato para primeiro construímos a medida μ, e somente depois mostramos que ela possui as propriedades mencionadas. Defina a sequência $s(0, j) = j$. Se $s(n-1, j)$ está definida, podemos escolher $s(n, j)$, uma subsequência de $s(n-1, j)$, e uma medida positiva μ^n com $0 \leq \mu^n(X^n) \leq 1$ tal que

$$\mu_{s(n,j)} \circ \pi_n^{-1} \to \mu^n$$

na topologia fraca-$*$. Usando o Teorema de consistência de Kolmogorov (Teorema A.4.4), existe uma medida μ definida nos borelianos de X, tal que

$$\mu \circ \pi_n^{-1} = \mu^n.$$

Vamos mostrar que μ possui as propriedades desejadas.

— Item (1).

Note que $0 \leq \mu^1(X_1) \leq 1$. Mas, $\mu(X) = \mu(\pi_1^{-1}(X_1)) = \mu^1(X_1)$.

— Item (2).

Afirmação. *Para toda aplicação limitada contínua $f : X^n \to \mathbb{R}$,*

$$\int f \,\mathrm{d}(\mu \circ \pi_n^{-1}) = \int f \,\mathrm{d}(\mu \circ T^{-1} \circ \pi_n^{-1}).$$

Ou seja,

$$\mu \circ \pi_n^{-1} = \mu \circ T^{-1} \circ \pi_n^{-1}.$$

Seja $m = m(n)$ (Definição 2.1.1). Note que $m \geq n$. Como f é limitada, é fácil verificar através da própria definição de μ_k, que

$$\int f \,\mathrm{d}(\mu_{s(m,j)} \circ T^{-1} \circ \pi_n^{-1}) \to \int f \,\mathrm{d}\mu^n = \int f \,\mathrm{d}(\mu \circ \pi_n^{-1}).$$

Por outro lado, temos que $f \circ \pi_n \circ T$ é π_m-mensurável. Utilizando o Lema 2.2.5, $f \circ \pi_n \circ T$ pode ser escrita como $g \circ \pi_m$ para alguma aplicação mensurável $g : X^m \to \mathbb{R}$. Assim,

$$\begin{aligned}
\int f \,\mathrm{d}(\mu_{s(m,j)} \circ T^{-1} \circ \pi_n^{-1}) &= \int f \circ \pi_n \circ T \,\mathrm{d}\mu_{s(m,j)} \\
&= \int g \circ \pi_m \,\mathrm{d}\mu_{s(m,j)} \\
&= \int g \,\mathrm{d}(\mu_{s(m,j)} \circ \pi_m^{-1}) \\
&\to \int g \,\mathrm{d}(\mu \circ \pi_m^{-1}) \\
&= \int g \circ \pi_m \,\mathrm{d}\mu \\
&= \int f \circ \pi_n \circ T \,\mathrm{d}\mu \\
&= \int f \,\mathrm{d}(\mu \circ T^{-1} \circ \pi_n^{-1}).
\end{aligned}$$

Ou seja, para todo n, e toda aplicação contínua limitada $f : X^n \to \mathbb{R}$,

$$\int f \,\mathrm{d}(\mu \circ \pi_n^{-1}) = \int f \,\mathrm{d}(\mu \circ T^{-1} \circ \pi_n^{-1}).$$

Portanto,

$$\mu \circ \pi_n^{-1} = \mu \circ T^{-1} \circ \pi_n^{-1}.$$

A afirmação anterior, juntamente com a unicidade do Teorema de consistência de Kolmogorov implica que $\mu = \mu \circ T^{-1}$.

— Item (3).

Existe um $B \subset X^n$ tal que $C = \pi_n^{-1}(B)$. Note que pelo Lema 2.2.2, $\partial C = \pi_n^{-1}(\partial B)$. Portanto, pela Observação (3) da página 149 de [Wal00], como $\mu \circ \pi_n^{-1}(\partial B) = 0$, temos que

$$\mu_{s(n,j)}(C) = \mu_{s(n,j)} \circ \pi_n^{-1}(B) \to \mu \circ \pi_n^{-1}(B) = \mu(C).$$

Em particular,

$$\liminf_{k\to\infty} \mu_k(C) \leq \mu(C).$$

□

Exemplo 2.2.8. Considere o sistema dinâmico $T : \{0,1\}^{\mathbb{N}} \to \{0,1\}^{\mathbb{N}}$, onde T é o shift. No Lema 2.2.7, tome

$$\sigma_k = \delta_x,$$

com $x \in \{0,1\}^{\mathbb{N}}$ sendo um ponto periódico com período t. Ou seja, $x = (x_1, x_2, \dots, x_t, x_1, x_2, \dots)$. Neste caso,

$$\begin{aligned}
\mu_{kt} &= \frac{1}{kt} \sum_{j=0}^{kt-1} \delta_x \circ T^{-j} \\
&= \frac{1}{t} \sum_{j=0}^{t-1} \delta_x \circ T^{-j}. \\
&= \mu_t
\end{aligned}$$

Ou seja, μ_{kt} é a medida de probabilidade com distribuição uniforme na órbita de x. Neste caso, a medida μ do Lema 2.2.7 será necessariamente igual a μ_t. De fato, note que

$$\begin{aligned}
\liminf_{j\to\infty} \mu_j(A) &= \lim_{j\to\infty} \mu_j(A) \\
&= \lim_{k\to\infty} \mu_{kt}(A) \\
&= \mu_t(A).
\end{aligned}$$

Assim, para qualquer y na órbita de x,

$$\begin{aligned}
\frac{1}{t} &= \liminf_{j\to\infty} \mu_j\left(\pi_n^{-1}\left(\pi_n(y)\right)\right) \\
&= \mu_t\left(\pi_n^{-1}\left(\pi_n(y)\right)\right)
\end{aligned}$$

para todo $n \geq t$. Mas como $\pi_n^{-1}(\pi_n(y)) \downarrow \{y\}$, temos que $\mu(\{y\}) \geq \frac{1}{t}$. E como existem t pontos na órbita de x, a condição $\mu \leq 1$ implica na igualdade.

Lema 2.2.9. *Seja $T : X \to X$ um sistema dinâmico do tipo produto, μ uma medida de Radon T-invariante sobre X e $\varepsilon > 0$. Então, existe $n \in \mathbb{N}$, e uma partição mensurável finita $\mathcal{C}$ formada por conjuntos π_n-mensuráveis de diâmetro menor que ε com respeito à métrica do tipo produto, e tais que para cada $C \in \mathcal{C}$*

$$\mu(\partial C) = 0.$$

Além disso, para todo $k \in \mathbb{N}$, cada membro de $\mathcal{C}^k$ é $\pi_{m^k(n)}$-mensurável com bordo de medida nula, onde $m^k(n)$ é a função m do item (2) da Definição 2.1.1 aplicada k vezes a n.

Demonstração. Escolha $n > \frac{1}{\varepsilon}$. Seja $\widetilde{c}$ a métrica em $\widetilde{X}^n$ dada por

$$\widetilde{c}((a_1, \ldots, a_n), (b_1, \ldots, b_n)) = \max_{j=1}^{n} \widetilde{d}_j(a_j, b_j).$$

E seja c a restrição de $\widetilde{c}$ a X^n. Para cada $x \in X^n$, existe ε_x com $\frac{\varepsilon}{2} \leq \varepsilon_x \leq \varepsilon$, tal que $B_c(x; \varepsilon_x) \subset X$ possui bordo de medida $\mu \circ \pi_n^{-1}$ nula. De fato, para diferentes $\varepsilon' > 0$, os bordos de $B_c(x; \varepsilon')$ são disjuntos. A família dos bordos das bolas centradas em x com raio entre $\frac{\varepsilon}{2}$ e ε é uma família não enumerável de conjuntos mensuráveis disjuntos. Assim, uma infinidade deles possui medida $\mu \circ \pi_n^{-1}$ igual a 0.

A família $B_{\widetilde{c}}\left(x; \frac{\varepsilon}{2}\right)$ com $x \in X^n$, cobre $\widetilde{X}_1 \times \cdots \times \widetilde{X}_n$, e portanto possui subcobertura finita. Usando o fato de que

$$B_c\left(x; \frac{\varepsilon}{2}\right) = X^n \cap B_{\widetilde{c}}\left(x; \frac{\varepsilon}{2}\right),$$

obtemos uma cobertura finita para X^n da forma

$$B_c\left(x_1; \frac{\varepsilon}{2}\right), \ldots, B_c\left(x_k; \frac{\varepsilon}{2}\right).$$

Em particular, as bolas $B_c(x_1; \varepsilon_{x_1}), \ldots, B_c(x_k; \varepsilon_{x_k})$ cobrem X. Podemos então assumir que esta subcobertura é minimal. Faça

$$A_j = \pi_n^{-1}(B_c(x_1; \varepsilon_{x_1})).$$

Usando o Lema 2.2.2, temos que o bordo de A_j tem medida μ nula.

Agora, vamos construir $\mathcal{C}$. Faça

$$C_j = A_j \setminus (A_1 \cup \cdots \cup A_{j-1})$$

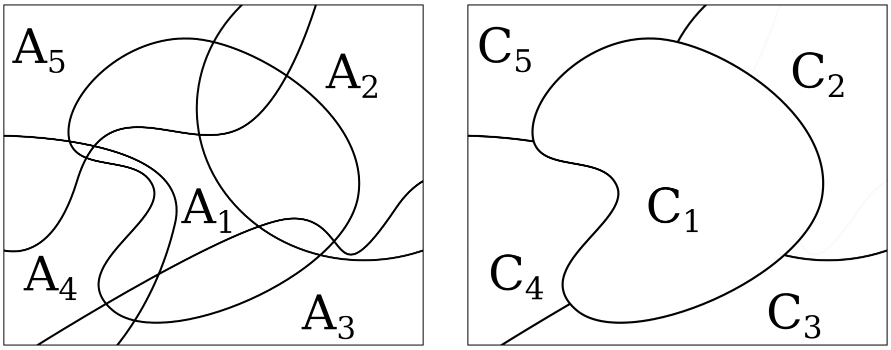

Figura 2.3: Partição $\mathcal{C}$ construída a partir da cobertura $\mathcal{A} = \{A_1, \ldots, A_5\}$. Note que $\partial C_k \subset \bigcup_{j=1}^{k} \partial A_j$.

(veja a Figura 2.3).

A família $\mathcal{C} = \{C_1, \ldots, C_k\}$ é uma partição tal que todos os seus elementos possuem diâmetro menor que ε em qualquer métrica do tipo produto, pois escolhemos $n > \frac{1}{\varepsilon}$. Cada C_j é tal que

$$C_j{}^c = A_j{}^c \cup A_1 \cup \cdots \cup A_{j-1}.$$

O bordo de um conjunto é o mesmo que o bordo de seu complemento. Para concluir que o bordo de cada um dos C_j tem medida μ igual a 0, basta mostrar que o bordo de uma união de finitos conjuntos está contido na união dos bordos de cada um deles.

Afirmação. *Para uma família $Y_1, \ldots, Y_s \subset X^n$,*

$$\partial \left(Y_1 \cup \cdots \cup Y_s\right) \subset \partial Y_1 \cup \cdots \cup \partial Y_s.$$

Como

$$\overline{Y_1 \cup \cdots \cup Y_s} = \overline{Y_1} \cup \cdots \cup \overline{Y_s}$$

e

$$\operatorname{int}\left(Y_1 \cup \cdots \cup Y_s\right) \supset \operatorname{int}\left(Y_1\right) \cup \cdots \cup \operatorname{int}\left(Y_s\right),$$

segue que

$$\begin{aligned}
\partial \left(Y_1 \cup \cdots \cup Y_s\right) &= \overline{Y_1} \cup \cdots \cup \overline{Y_s} \setminus \left(\operatorname{int}\left(Y_1 \cup \cdots \cup Y_s\right)\right) \\
&\subset \overline{Y_1} \cup \cdots \cup \overline{Y_s} \setminus \left(\operatorname{int}\left(Y_1\right) \cup \cdots \cup \operatorname{int}\left(Y_s\right)\right) \\
&\subset \big(\overline{Y_1} \setminus \operatorname{int}\left(Y_1\right)\big) \cup \cdots \cup \big(\overline{Y_s} \setminus \operatorname{int}\left(Y_s\right)\big) \\
&= \partial Y_1 \cup \cdots \cup \partial Y_s.
\end{aligned}$$

Usando a afirmação anterior, não apenas o bordo de cada elemento de $\mathcal{C}$ está contido em uma união finita de conjuntos de medida μ nula, mas também segue que o mesmo vale para o bordo de qualquer elemento de $\mathcal{C}^k$, pois μ é T-invariante. Pela definição de sistema dinâmico do tipo produto, é imediato que os elementos de $\mathcal{C}^k$ são $\pi_{m^k(n)}$-mensuráveis. $\square$

CAPÍTULO 3

Princípio Variacional

3.1 Demonstração do Princípio Variacional

A demonstração que será apresentada do *princípio variacional* foi inspirada na demonstração dada por Misiurewicz para o princípio variacional, no caso em que o sistema dinâmico é compacto. A demonstração de Misiurewicz é apresentada, por exemplo, em [Wal00].

Vamos começar com um lema que generaliza a Proposição 1.4 de [HKR95], pois a única hipótese que impomos ao sistema $T : X \to X$ é que T seja contínua. O princípio variacional fala sobre o supremo das entropias de Kolmogorov-Sinai tomado dentre todas as medidas de probabilidade de Radon T-invariantes. Se não houver nenhuma medida de probabilildade Radon T-invariante, então, convencionamos que

$$\sup_{\mu} h_{\mu}(T) = 0.$$

Lema 3.1.1. *Seja $T : X \to X$ um sistema dinâmico topológico. Então,*

$$\sup_{\mu} h_{\mu}(T) \leq h(T),$$

onde o supremo é tomado em todas as medidas de probabilidade de Radon T-invariantes. Se X é metrizável, então

$$\sup_{\mu} h_{\mu}(T) \leq h(T) \leq \inf_{d} h_{d}(T).$$

Demonstração. A última afirmação segue da Proposição 1.2.33.

Vamos mostrar que, dada uma medida de probabilidade de Radon μ definida nos borelianos de X,

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

Seja $\mathcal{C} = \{C_1, \ldots, C_k\}$ uma partição mensurável. É suficiente mostrar que existe uma cobertura admissível $\mathcal{A} = \{A_1, \ldots, A_k\}$ tal que

$$h_\mu\big(T \bigm| \mathcal{C}\big) \leq h\big(T \bigm| \mathcal{A}\big).$$

Utilizando a Proposição 1.2.1, escolha $\delta > 0$ tal que todo δ-refinamento $\mathcal{D}$ de $\mathcal{C}$ é tal que $h_\mu\big(T \bigm| \mathcal{C}\big) \leq h_\mu\big(T \bigm| \mathcal{D}\big) + 1$. Como μ é uma medida de Radon, podemos tomar $K_j \subset C_j$ compacto, com $\mu(C_j \setminus K_j) \leq \delta\mu(C_j)$. Faça

$$K_0 = (K_1 \cup \cdots \cup K_k)^c.$$

Então, denote a família de tais conjuntos por

$$\mathcal{D} = \{K_0, \ldots, K_k\}.$$

Afirmação. $h_\mu\big(T \bigm| \mathcal{C}\big) \leq h_\mu\big(T \bigm| \mathcal{D}\big) + 1.$

De fato, os compactos $K_1, \ldots, K_k$ foram escolhidos de modo que $\mathcal{D}$ seja um δ-refinamento de $\mathcal{C}$, pois para $j = 1, \ldots, k$,

$$\begin{aligned}
\mu(C_j \cap K_j) &= \mu(C_j) - \mu(C_j \setminus K_j) \\
&\geq \mu(C_j) - \delta\mu(C_j) \\
&= (1 - \delta)\mu(C_j).
\end{aligned}$$

Mas δ foi escolhido através da Proposição 1.2.1, de modo que a desigualdade da afirmação seja verdadeira.

Seja $\mathcal{A} = \{K_0 \cup K_1, \ldots, K_0 \cup K_k\}$. Note que $\mathcal{A}$ é uma cobertura admissível.

Afirmação. $N(\mathcal{D}^n) \leq 2^n N(\mathcal{A}^n)$.

Seja $\Lambda \subset \{1, \ldots, k\}^n$ um conjunto com $\#\Lambda = N(\mathcal{A}^n)$, tal que

$$X = \bigcup_{\lambda \in \Lambda} (K_0 \cup K_{\lambda_0}) \cap \cdots \cap T^{-(n-1)}(K_0 \cup K_{\lambda_{n-1}}),$$

onde $\lambda = (\lambda_0, \ldots, \lambda_{n-1})$. Considere a aplicação

$$f : \Lambda \times \{0,1\}^n \to \mathcal{D}^n,$$

que leva (λ, x) em $Y_0 \cap \cdots \cap Y_{n-1}$, onde $Y_j = T^{-j} K_{\lambda_j}$ quando $x_j = 1$, e $Y_j = T^{-j} K_0$ quando $x_j = 0$. Como,

$$\begin{aligned} X &= \bigcup_{\lambda \in \Lambda} T^0(K_{\lambda_0} \cup K_0) \cap \cdots \cap T^{-(n-1)}(K_{\lambda_{n-1}} \cup K_0) \\ &= \bigcup f(\Lambda \times \{0,1\}^n), \end{aligned}$$

e $\mathcal{D}^n$ é uma partição, temos que a imagem de f contém todos os elementos não vazios de $\mathcal{D}^n$. Assim,

$$N(\mathcal{D}^n) \leq 2^n \# \Lambda.$$

Afirmação. *Para todo $n \in \mathbb{N}$,*

$$h_\mu\big(T \bigm| \mathcal{D}\big) \leq h\big(T \bigm| \mathcal{A}\big) + \log 2.$$

Basta tomar o log na desigualdade da afirmação anterior, dividir por n e fazer $n \to \infty$.

Juntando todas as afirmações, temos que

$$h_\mu\big(T \bigm| \mathcal{C}\big) \leq h(T) + 1 + \log 2. \tag{3.1}$$

E como a partição $\mathcal{C}$ havia sido escolhida arbitrariamente, podemos tomar o supremo para concluir que

$$h_\mu(T) \leq h(T) + 1 + \log 2. \tag{3.2}$$

E isso também é verdade se utilizarmos T^n ao invés de T:

$$h_\mu(T^n) \leq h(T^n) + 1 + \log 2.$$

Agora, pela Observação 1.2.14 e pelo Lema 1.2.13,

$$h_\mu(T^n) = n h_\mu(T) \quad \text{ e } \quad h(T^n) \leq n h(T).$$

E portanto,

$$\begin{aligned} h_\mu(T) &= \frac{1}{n} h_\mu(T^n) \\ &\leq \frac{1}{n} h(T^n) + \frac{1}{n}(1 + \log 2) \\ &\leq h(T) + \frac{1}{n}(1 + \log 2). \end{aligned}$$

Fazendo $n \to \infty$, obtemos o resultado desejado. □

O teorema a seguir é o princípio variacional para sistemas dinâmicos do tipo produto.

Teorema 3.1.2. *Seja $T : X \to X$ um sistema dinâmico do tipo produto. Então,*

$$\sup_{\mu} h_{\mu}(T) = h(T) = \min_{d} h_d(T).$$

O mínimo é atingido quando d é qualquer métrica que possa ser estendida a uma compactificação $\widetilde{X}$ de X.

Demonstração. Do Lema 3.1.1, sabemos que

$$\sup_{\mu} h_{\mu}(T) \leq h(T) \leq \inf_{d} h_d(T) \leq \inf_{d} h^d(T),$$

onde o ínfimo é tomado em todas as métricas d compatíveis com a topologia de X. Basta então mostrar que se r é a restrição a X de uma métrica qualquer em $\widetilde{X}$, então, vale que

$$h^r(T) \leq \sup_{\mu} h_{\mu}(T).$$

De fato, isso implica que

$$h^r(T) \leq \sup_{\mu} h_{\mu}(T) \leq h(T) \leq \inf_{d} h_d(T) \leq \inf_{d} h^d(T) \leq h^r(T).$$

Para mostrar que $h^r(T) \leq \sup_{\mu} h_{\mu}(T)$, o Corolário 1.2.32 nos permite assumir que r é uma *métrica do tipo produto*, ou seja, r é a métrica dada pela Proposição 2.2.1.

Pela Proposição 1.2.26, resta mostrar que na métrica r, para cada $\varepsilon > 0$ fixado, e cada sequência de conjuntos (n, ε)-separados, E_n, existe uma medida de Radon μ que é T-invariante e tem medida total menor ou igual a 1, e também que existe uma partição mensurável $\mathcal{C}$ tal que

$$\log \# E_n \leq H_{\mu}(\mathcal{C}^n).$$

Vamos primeiro construir a medida μ e a partição $\mathcal{C}$, para somente então mostrar que elas satisfazem as condições especificadas. Seja σ_n a medida de probabilidade uniforme em E_n. Ou seja,

$$\sigma_n = \frac{1}{\# E_n} \sum_{x \in E_n} \delta_x,$$

onde δ_x é a medida de Dirac com suporte em x. Defina também

$$\mu_n = \frac{1}{n}\sum_{j=0}^{n-1} \sigma_n \circ T^{-j}.$$

Agora, tome a medida de Radon T-invariante μ do Lema 2.2.7, e então escolha $\mathcal{C}$ como no Lema 2.2.9.

Afirmação. $H_{\sigma_n}(\mathcal{C}^n) = \log \#E_n$.

Seja $C \in \mathcal{C}^n$. Se $x, y \in C$, então, exitem $C_0, \ldots, C_{n-1} \in \mathcal{C}$ tais que $T^j x, T^j y \in C_j$ para $j = 0, \ldots, n-1$. Como cada elemento de $\mathcal{C}$ tem diâmetro menor que ε, temos que $r_n(x, y) < \varepsilon$. Assim, C pode conter no máximo um elemento de E_n. Ou seja, $\sigma_n(C) = 0$ ou $\sigma_n(C) = \frac{1}{\#E_n}$. E portanto,

$$H_{\sigma_n}(\mathcal{C}^n) = \log \#E_n.$$

Para passar de σ_n para μ_n, utilizaremos o Lema A.3.3. Dados $n, q \in \mathbb{N}$ com $1 < q < n$, vamos calcular $H_{\mu_n}(\mathcal{C}^q)$. Tome um inteiro m tal que $mq \geq n$. Então, para todo $j = 0, \ldots, q-1$,

$$\begin{aligned}
\mathcal{C}^n &\prec \mathcal{C}^j \vee T^{-j}(\mathcal{C}^{qm}) \\
&= \mathcal{C}^j \vee T^{-j}(\mathcal{C}^q) \vee T^{-(j+q)}(\mathcal{C}^q) \vee \cdots \vee T^{-(j+(m-1)q)}(\mathcal{C}^q).
\end{aligned}$$

Assim, obtemos

$$\begin{aligned}
H_{\sigma_n}(\mathcal{C}^n) &\leq H_{\sigma_n}\left(\mathcal{C}^j\right) + H_{\sigma_n \circ T^{-(j+0q)}}(\mathcal{C}^q) + \cdots + H_{\sigma_n \circ T^{-(j+(m-1)q)}}(\mathcal{C}^q) \\
&\leq H_{\sigma_n}(\mathcal{C}^q) + H_{\sigma_n \circ T^{-(j+0q)}}(\mathcal{C}^q) + \cdots + H_{\sigma_n \circ T^{-(j+(m-1)q)}}(\mathcal{C}^q) \\
&\leq \log \#\mathcal{C}^q + H_{\sigma_n \circ T^{-(j+0q)}}(\mathcal{C}^q) + \cdots + H_{\sigma_n \circ T^{-(j+(m-1)q)}}(\mathcal{C}^q),
\end{aligned}$$

onde a última desigualdade segue do Lema A.3.2. Somando em $j = 0, \ldots, q-1$, e utilizando o Lema A.3.3,

$$\begin{aligned}
q \log \#E_n &= q H_{\sigma_n}(\mathcal{C}^n) \\
&\leq q \log \#\mathcal{C}^q + \sum_{j=0}^{q-1}\sum_{a=0}^{m-1} H_{\sigma_n \circ T^{-(j+aq)}}(\mathcal{C}^q) \\
&= q \log \#\mathcal{C}^q + \sum_{p=0}^{n-1} H_{\sigma_n \circ T^{-p}}(\mathcal{C}^q) \\
&= q \log \#\mathcal{C}^q + n \sum_{p=0}^{n-1} \frac{1}{n} H_{\sigma_n \circ T^{-p}}(\mathcal{C}^q) \\
&\leq q \log \#\mathcal{C}^q + n H_{\mu_n}(\mathcal{C}^q).
\end{aligned}$$

Como cada elemento de $\mathcal{C}^q$ é π_m-mensurável para algum m, e cada um tem bordo com medida μ nula, segue do item (3) do Lema 2.2.7 que dividindo por qn, e fazendo $n \to \infty$,

$$\begin{aligned}\limsup_{n\to\infty} \frac{1}{n} \log \# E_n &\leq 0 + \frac{1}{q} \liminf_{n\to\infty} H_{\mu_n}\left(\mathcal{C}^q\right) \\ &\leq \frac{1}{q} H_\mu\left(\mathcal{C}^q\right)\end{aligned}$$

para todo q. Agora, basta tomar o limite com $q \to \infty$ para obter

$$\limsup_{n\to\infty} \frac{1}{n} \log \# E_n \leq h_\mu\big(T \bigm| \mathcal{C}\big).$$

□

O resultado seguinte, um caso particular do teorema anterior, generaliza o princípio variacional para sistemas dinâmicos localmente compactos apresentado em [Pat10], pois não assume que a aplicação que define o sistema seja própria.

Corolário 3.1.3. *Seja $T : X \to X$ um sistema dinâmico metrizável, localmente compacto e separável. Então,*

$$\sup_\mu h_\mu\left(T\right) = h\left(T\right) = \min_d h_d\left(T\right).$$

O mínimo é atingido quando d é qualquer métrica que possa ser estendida a uma compactificação $\widetilde{X}$ de X.

O Lema 1.2.13 foi necessário para a demonstração do princípio variacional. Agora, o princípio variacional nos permite irmos um pouco além.

Corolário 3.1.4. *Considere o sistema dinâmico do tipo produto $T : X \to X$, e seja $k \in \mathbb{N}$. Então,*

$$h\left(T^k\right) = kh\left(T\right).$$

Demonstração. A igualdade vale para a entropia de Kolmogorov-Sinai (veja a Observação 1.2.14), e portanto, pelo princípio variacional, vale para a entropia topológica. □

Juntamente com o princípio variacional, a Proposição 1.2.31 fornece uma caracterização alternativa da entropia topológica.

Corolário 3.1.5. *Seja $T : X \to X$ um sistema dinâmico do tipo produto, e $\widetilde{X}$ uma compactificação metrizável de X. Então,*

$$h(T) = \sup_{\substack{\widetilde{\mathcal{A}}:\ \text{cobertura aberta} \\ \text{de } \widetilde{X}}} h\Big(T \,\Big|\, X \cap \widetilde{\mathcal{A}}\Big),$$

onde o supremo é tomado sobre todas as coberturas abertas de $\widetilde{X}$.

O princípio variacional relaciona medida e topologia, assim como a versão topológica do *Teorema de Recorrência de Poincaré*. Uma implicação simples desta relação, é o seguinte corolário.

Corolário 3.1.6. *Se $T : X \to X$ é um sistema dinâmico onde X é localmente compacto separável, então, denotando por $\mathcal{R}(T)$ o conjunto dos pontos recorrentes de T, temos que*

$$h(T) = h\left(T|_{\overline{\mathcal{R}(T)}}\right)$$

Demonstração. O fecho do conjunto recorrente, $R = \overline{\mathcal{R}(T)}$ é invariante e localmente compacto. Assim, $T : R \to R$ é um sistema dinâmico do tipo produto. Toda medida de probabilidade $T|_R$-invariante sobre R é a restrição a R de uma medida de probabilidade T-invariante sobre X com suporte contido em R. Por outro lado, pela versão topológica do *Teorema de Recorrência de Poincaré* (Corolário A.2.2), para toda medida T-invariante μ, $\operatorname{supp}(\mu) \subset \overline{\mathcal{R}(T)}$ (Definição 4.1.3), e portanto,

$$\begin{aligned}
h_\mu(T) &= \sup_{\substack{\mathcal{C}:\ \text{partição mensurável} \\ \text{finita}}} h_\mu\Big(T \,\Big|\, \mathcal{C} \vee \{R, R^c\}\Big) \\
&= \sup_{\substack{\mathcal{C}:\ \text{partição mensurável} \\ \text{finita}}} h_{\mu|_R}\Big(T|_R \,\Big|\, \mathcal{R}(T) \cap \mathcal{C}\Big) \\
&= h_{\mu|_R}(T|_R).
\end{aligned}$$

Assim,

$$\begin{aligned}
h(T|_R) &= \sup_\mu h_{\mu|_R}(T|_R) \\
&= \sup_\mu h_\mu(T) \\
&= h(T).
\end{aligned}$$

$\square$

Por vezes, precisamos considerar apenas as medidas ergódicas de um sistema dinâmico. O corolário a seguir é o princípio variacional para as medidas ergódicas. Lembre-se que uma medida T-invariante é ergódica quando todos os conjuntos invariantes tiverem medida 0 ou 1.

Ou seja, vamos mostrar que no caso dos sistemas dinâmicos do tipo produto, o supremo das entropias das medidas invariantes coincide com o supremo das entropias das medidas ergódicas.

Corolário 3.1.7. *Seja* $\widetilde{T} : \widetilde{X} \to \widetilde{X}$ *um sistema dinâmico compacto, e* $X \subset \widetilde{X}$ *um boreliano invariante. Então, denotando por* $T : X \to X$ *a restrição do sistema a* X*, temos que o valor de* $\sup_\mu h_\mu(T)$ *é o mesmo se o supremo for tomado nas medidas ergódicas ou nas medidas* T*-invariantes.*

Demonstração. A demonstração do Teorema 1, ao final da página 310 de [PP84], mostra que dada um medida T-invariante μ, existe uma medida T-ergódica μ_s tal que $h_\mu(T) \leq h_{\mu_s}(T)$. Isso implica que o supremo tomado nas medidas ergódicas é maior ou igual ao supremo tomado nas medidas invariantes. Como as medidas ergódicas são também invariantes, temos a igualdade. $\square$

Exemplo 3.1.8. Seja

$$\begin{array}{rccc} T: & \mathbb{R} & \to & \mathbb{R} \\ & x & \mapsto & 2x \end{array}.$$

De acordo com o parágrafo final do Exemplo 1.2.12, $h(T) = 0$. Pelo Princípio Variacional, para a métrica $d(x,y)$ em $\mathbb{R}$ dada pelo ângulo θ da Figura 3.1, devemos ter $h_d(T) = 0$. Vamos verificar este fato diretamente. O argumento é muito parecido com o do Exemplo 1.2.12.

Primeiramente, note que na métrica d, as bolas serão sempre intervalos. Para $\varepsilon > 0$, tome $\mathcal{A} \subset \mathcal{B}_d(\varepsilon)$ finito. Assim como no Exemplo 1.2.12, é evidente que existe uma partição finita $\mathcal{C}$, composta por arcos (não necessariamente abertos), tal que $\mathcal{A} \prec \mathcal{C}$. Sendo assim, resta apenas mostrar que $h\big(T \bigm| \mathcal{C}\big) = 0$. Denote por $E_{\mathcal{C}}$, o conjunto dos extremos de cada intervalo de $\mathcal{C}$. Agora note, que a transformação T também leva intervalos em intervalos. Portanto, assim como no Exemplo 1.2.12,

$$E_{\mathcal{C}^n} = \bigcup_{j=0}^{n-1} T^{-1}(E_{\mathcal{C}}).$$

Em especial,

$$N(\mathcal{C}^n) \leq n\# E_{\mathcal{C}}.$$

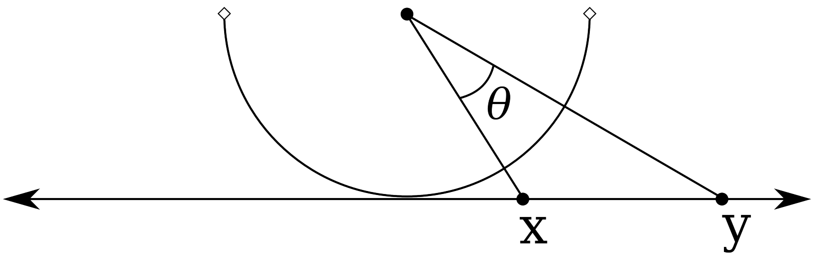

Figura 3.1: A métrica onde $d(x, y)$ é igual ao ângulo θ é uma métrica em $\mathbb{R}$ que admite uma compactificação.

Ou seja,

$$h\big(T \,\big|\, \mathcal{C}\big) \leq \lim_{n\to\infty} \frac{1}{n} \log\left(n \# E_{\mathcal{C}}\right) = 0.$$

Exemplo 3.1.9 (Entropia topológica do shift em $\mathbb{N}$)**.** Seja

$$\begin{array}{rccc} T: & \mathbb{N}^{\mathbb{N}} & \to & \mathbb{N}^{\mathbb{N}} \\ & (x_n)_{n\in\mathbb{N}} & \mapsto & (x_{n+1})_{n\in\mathbb{N}} \end{array}$$

o shift nos naturais. Utilizando o Princípio Variacional, vamos mostrar que a entropia topológica de T é infinita. Basta mostrar que existe uma medida μ T-invariante tal que $h_\mu(T)$ é tão grande quanto queiramos. Considere em $\mathbb{N}$ a medida

$$\nu = \frac{1}{n} \sum_{j=1}^{n} \delta_j,$$

e faça $\mu = \nu^{\mathbb{N}}$, a medida produto. Agora, tome a partição finita

$$\mathcal{C} = \left\{ \{j\} \times \mathbb{N}^{\mathbb{N}} \,\middle|\, j = 1, \ldots, n \right\} \cup \left\{ \{n+1, n+2, \ldots\} \times \mathbb{N}^{\mathbb{N}} \right\}.$$

Note que cada elemento de $\mathcal{C}^k$ da forma $C = \{(c_1, \ldots, c_k)\} \times \mathbb{N}^{\mathbb{N}}$, com $c_j \in \{1, \ldots, n\}$ tem medida $\nu(C_j) = \frac{1}{n^k}$. Enquanto que os demais elementos de $\mathcal{C}^k$ são subconjuntos de $\{n+1, n+2, \ldots\} \times \mathbb{N}^{\mathbb{N}}$, que tem medida ν nula.

Assim,

$$\begin{aligned} h(T) &\geq h_\mu(T) \\ &\geq h_\mu\big(T \bigm| \mathcal{C}\big) \\ &= \lim_k \frac{1}{k} H_\mu\left(\mathcal{C}^k\right) \\ &= \lim_k \frac{1}{k} \log n^k \\ &= \log n. \end{aligned}$$

Fazendo $n \to \infty$, concluímos que $h(T) = \infty$.

É bastante sabido que no caso compacto, a entropia do produto de dois sistemas dinâmicos é igual à soma das entropias de cada um deles. O princípio variacional nos permite estender este resultado.

Proposição 3.1.10. *Seja $T_1 : X_1 \to X_1, T_2 : X_2 \to X_2, \ldots$ uma família enumerável de sistemas dinâmicos do tipo produto. Considere o sistema $T = \prod_{j=1}^{\infty} T_j$. Então, T é do tipo produto, e*

$$h(T) = \sum h(T_n).$$

Demonstração. O sistema T é do tipo produto pelo Exemplo 2.1.4.

Considere a métrica $d(x,y) = \sup \frac{1}{j} d_j(x_j, y_j)$ em $\prod_{j=1}^{\infty} X_j$, onde cada d_j é a restrição a X_j de uma métrica em $\widetilde{X}_j$. Então, para cada $\varepsilon > 0$,

$$\mathcal{B}_d(\varepsilon) = \mathcal{B}_{d_1}(\varepsilon) \times \mathcal{B}_{d_2}(2\varepsilon) \times \mathcal{B}_{d_3}(3\varepsilon) \times \cdots .$$

Note também que para $n > \frac{1}{\varepsilon}$, $\mathcal{B}_{d_n}(n\varepsilon) = \{X_n\}$. De modo mais geral, como T age em cada coordenada, para cada $k \in \mathbb{N}$,

$$[\mathcal{B}_d(\varepsilon)]^k = [\mathcal{B}_{d_1}(\varepsilon)]^k \times [\mathcal{B}_{d_2}(2\varepsilon)]^k \times [\mathcal{B}_{d_3}(3\varepsilon)]^k \times \cdots .$$

Novamente, para $n > \frac{1}{\varepsilon}$, $[\mathcal{B}_{d_n}(n\varepsilon)]^k = \{X_n\}$. Segue que

$$N\left([\mathcal{B}_d(\varepsilon)]^k\right) = N\left([\mathcal{B}_{d_1}(\varepsilon)]^k\right) \cdots N\left([\mathcal{B}_{d_n}(n\varepsilon)]^k\right),$$

e $N\left([\mathcal{B}_{d_m}(m\varepsilon)]^k\right) = 1$ para $m > n$. Assim, para $n > \frac{1}{\varepsilon}$,

$$\begin{aligned} h\Big(T \Bigm| \mathcal{B}_d(\varepsilon)\Big) &= \sum_{j=1}^{n} h\Big(T_j \Bigm| \mathcal{B}_{d_j}(j\varepsilon)\Big) \\ &= \sum_{j=1}^{\infty} h\Big(T_j \Bigm| \mathcal{B}_{d_j}(j\varepsilon)\Big). \end{aligned}$$

Utilizando o princípio variacional, a igualdade anterior implica que

$$\begin{aligned} h(T) &= \sup_{\varepsilon>0} h\Big(T \,\Big|\, \mathcal{B}_d(\varepsilon)\Big) \\ &= \lim_{\varepsilon\downarrow 0} h\Big(T \,\Big|\, \mathcal{B}_d(\varepsilon)\Big) \\ &= \lim_{\varepsilon\downarrow 0} \sum_{j=1}^{\infty} h\Big(T_j \,\Big|\, \mathcal{B}_{d_j}(j\varepsilon)\Big) \\ &= \sum_{j=1}^{\infty} \lim_{\varepsilon\downarrow 0} h\Big(T_j \,\Big|\, \mathcal{B}_{d_j}(j\varepsilon)\Big) \\ &= \sum_{j=1}^{\infty} h(T_j). \end{aligned}$$

Onde a quarta equação é devida ao Teorema da Convergência Monótona.

□

CAPÍTULO 4

Pares de Li-Yorke

4.1 Definição e Propriedades

De certa forma, a entropia de um sistema dinâmico determina seu grau de caoticidade. Vamos mostrar que entropia topológica positiva implica, por exemplo, na existência dos chamados *pares de Li-Yorke*. A definição usual de *par de Li-Yorke* utiliza uma métrica (veja [BGKM02]). No caso de sistemas não compactos, esta definição é dependente da métrica adotada. Neste trabalho, vamos optar por uma definição puramente topológica.

Definição 4.1.1 (Par de Li-Yorke Topológico)**.** *Dado um conjunto X, e um subconjunto $W \subset X$, denotamos por Δ_W a* diagonal

$$\Delta_W = \{(w, w) \in W \times W\}.$$

Se $W = X$, omitimos o índice e escrevemos apenas Δ no lugar de Δ_X.

Em um sistema dinâmico $T : X \to X$, dizemos que $(x, y) \in X \times X$ é um par de Li-Yorke topológico *quando existirem $a, b, c \in X$, com $a \neq b$, tal que para toda vizinhança $V \subset X \times X$ de (a, b), $(T \times T)^n(x, y) \in V$ para infinitos valores de n, e para toda vizinhança $U \subset X \times X$ de (c, c), $(T \times T)^n(x, y) \in U$ para infinitos valores de n (veja a Figura 4.1).*

Observação 4.1.2. Quando (x, y), com $x \neq y$ for um ponto recorrente de $T \times T$, podemos simplesmente tomar x e y no lugar de a e b na Definição 4.1.1. Neste caso, resta mostrar a existência de $c \in X$ tal que para toda vizinhança $U \subset X \times X$ de (c, c), $(T \times T)^n(x, y) \in U$ para infinitos valores de n.

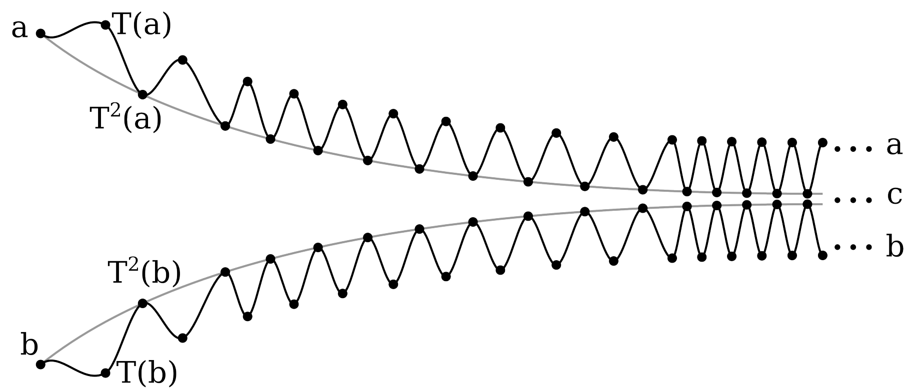

Figura 4.1: O par (x, y) forma um par de Li-Yorke. Os pares de Li-Yorke podem ser ainda mais complicados, com $(T \times T)^n(x, y)$ possuindo subsequências que convergem para diferentes pontos de $X \times X$.

Definição 4.1.3 (Suporte). *Se μ é uma medida de probabilidade de Radon em um espaço separável, o* suporte *de μ é o menor fechado que tem medida total. Ou seja,*

$$\operatorname{supp}(\mu) = \bigcap_{\substack{F:\ \textit{fechado} \\ \mu(F)=1}} F.$$

Proposição 4.1.4. *Seja μ uma medida de probabilidade de Radon em um espaço métrico separável X. Então, para todo aberto $A \subset X$,*

$$\mu(A) > 0 \Leftrightarrow A \cap \operatorname{supp}(\mu) \neq \emptyset.$$

Demonstração. Note que $A \cap \operatorname{supp}(\mu) = \emptyset$ se, e somente se, A^c é um fechado que contém $\operatorname{supp}(\mu)$. Mas isso acontece se, e somente se $\mu(A^c) = 1$. □

Lema 4.1.5. *Se $T : X \to X$ é um sistema dinâmico e μ uma medida T-ergódica, então existe um conjunto $B \subset X$ de medida 1 tal que a órbita de cada um de seus pontos é densa em* $\operatorname{supp}(\mu)$*.*

Demonstração. Seja A_λ $(\lambda \in \Lambda)$ uma base enumerável da topologia. Considere a família $\Gamma \subset \Lambda$ dada por

$$\Gamma = \left\{ \gamma \in \Lambda \,\middle|\, A_\gamma \cap \operatorname{supp}(\mu) \neq \emptyset \right\}.$$

Vamos denotar por B_γ os pontos $x \in X$ tais que

$$\frac{\#\left(\{T^0(x), \dots, T^{n-1}(x)\} \cap A_\gamma\right)}{n} \to \mu(A_\gamma). \tag{4.1}$$

Pelo teorema ergódico de Birkhoff (Teorema 1.14 de [Wal00]), $\mu(B_\gamma) = 1$. Como Γ é enumerável, fazendo $B = \bigcap_{\gamma \in \Gamma} B_\gamma$,

$$\mu(B) = 1.$$

Os pontos de B satisfazem a equação (4.1), e em particular suas órbitas passam infinitas vezes por qualquer A_γ de medida positiva. Ou seja, qualquer A_γ com $\gamma \in \Gamma$. Seja $x \in \operatorname{supp}(\mu)$ e $A \subset X$ uma vizinhança aberta de x. Então existe $\gamma \in \Gamma$, tal que $A_\gamma \subset A$. Em particular, a órbita dos pontos de B passa infinitas vezes por A. Isso implica que x está no fecho da órbita de cada ponto de B. □

Vamos mostrar que os sistemas dinâmicos do tipo produto com entropia positiva possuem pares de Li-Yorke. Em particular, isso implica que um sistema do tipo produto que não admite nenhum par de Li-Yorke não pode ter entropia positiva. Para investigar os pares de Li-Yorke de um sistema dinâmico $T : X \to X$, é natural olhar para o sistema produto $(T \times T) : X \times X \to X \times X$. Dada uma medida de probabilidade T-ergódica μ, vamos introduzir uma técnica para construir uma medida $(T \times T)$-ergódica λ em $X \times X$. Para tanto, vamos utilizar a chamada *σ-álgebra de Pinsker* do sistema dinâmico T com respeito à medida μ.

Proposição 4.1.6. *Seja $T : X \to X$ um sistema dinâmico e μ uma medida T-invariante. Então a família*

$$\mathcal{P} = \left\{ P\text{: boreliano} \,\middle|\, h_\mu\Big(T \,\Big|\, \{P, P^c\}\Big) = 0 \right\}$$

é uma σ-álgebra.

Demonstração. É evidente que $\emptyset \in \mathcal{P}$, e que $\mathcal{P}$ é uma família fechada por complementação. Para ver que $\mathcal{P}$ é fechada por interseção, note que fazendo $\mathcal{C} = \{P, P^c\}$ e $\mathcal{D} = \{Q, Q^c\}$, temos que

$$\{P \cap Q, (P \cap Q)^c\} \prec \mathcal{C} \vee \mathcal{D}.$$

Portanto, se P e Q estão em $\mathcal{P}$, então, pelo item (7) do Lema A.3.5,

$$\begin{aligned} h_\mu\Big(T \,\Big|\, \{P \cap Q, (P \cap Q)^c\}\Big) &\leq h_\mu\big(T \,\big|\, \mathcal{C} \vee \mathcal{D}\big) \\ &\leq h_\mu\big(T \,\big|\, \mathcal{C}\big) + h_\mu\big(T \,\big|\, \mathcal{D}\big) = 0. \end{aligned}$$

Está provado que $\mathcal{P}$ é uma álgebra. Vamos verificar que se $A_n \in \mathcal{P}$, com $A_n \uparrow A$, então $A \in \mathcal{P}$. Note que $\mu(A \setminus A_n) \to 0$, e que

$$\{A, A^c\} \prec \{A_n, {A_n}^c\} \vee \{A \setminus A_n, (A \setminus A_n)^c\}.$$

Assim, pelos itens (8) e (4) do Lema A.3.5,

$$\begin{aligned}
h_\mu\Big(T \,\Big|\, \{A, A^c\}\Big) &\leq h_\mu\Big(T \,\Big|\, \{A_n, {A_n}^c\}\Big) + H_\mu\Big(\{A \setminus A_n, (A \setminus A_n)^c\} \,\Big|\, \{A_n, {A_n}^c\}\Big) \\
&\leq h_\mu\Big(T \,\Big|\, \{A_n, {A_n}^c\}\Big) + H_\mu\left(\{A \setminus A_n, (A \setminus A_n)^c\}\right) \\
&= 0 + H_\mu\left(\{A \setminus A_n, (A \setminus A_n)^c\}\right) \\
&= \varepsilon \log \frac{1}{\varepsilon} + (1-\varepsilon) \log \frac{1}{1-\varepsilon} \to 0,
\end{aligned}$$

onde $\varepsilon = \mu(A \setminus A_n) \to 0$. □

Definição 4.1.7 (σ-Álgebra de Pinsker)**.** *Seja $T : X \to X$ um sistema dinâmico e μ uma medida T-invariante. A σ-álgebra da proposição 4.1.6 é chamada de* σ-álgebra de Pinsker *do sistema dinâmico T com respeito à medida μ, e é denotada por $\mathcal{P}_\mu$ ou simplesmente $\mathcal{P}$.*

Proposição 4.1.8. *Dado um sistema dinâmico $T : X \to X$ e uma medida T-invariante μ, a σ-álgebra de Pinsker $\mathcal{P}_\mu$ é T-invariante. Ou seja, $T^{-1}(\mathcal{P}_\mu) \subset \mathcal{P}_\mu$.*

Demonstração. Basta observar que

$$h_\mu\Big(T \,\Big|\, \{P, P^c\}\Big) = h_\mu\Big(T \,\Big|\, \big\{T^{-1}(P), T^{-1}(P^c)\big\}\Big),$$

pois μ é T-invariante. □

A proposição a seguir servirá para construirmos uma medida (ergódica). Para tanto, utilizaremos o conceito de *esperança condicional*. Mais detalhes e propriedades da esperança condicional podem ser encontrados no Apêndice A.1.

Proposição 4.1.9. *Seja μ uma medida finita em uma σ-álgebra $\mathcal{F}$ qualquer, e $\mathcal{G}$ uma sub-σ-álgebra de $\mathcal{F}$. Então existe uma medida λ na* σ-álgebra produto *$\mathcal{F} \times \mathcal{F}$ que satisfaz*

$$\lambda(A \times B) = \int E\big(I_A \,\big|\, \mathcal{G}\big) E\big(I_B \,\big|\, \mathcal{G}\big) \,\mathrm{d}\mu.$$

Demonstração. Pela linearidade da integral e da esperança condicional, é evidente que λ é aditiva e portanto pode ser estendida a uma medida aditiva na álgebra das uniões disjuntas finitas dos retângulos mensuráveis. Pela Proposição 2.4.1 de [Fer96], é suficiente mostrar que λ é contínua superiormente no vazio quando restrita à álgebra dos retângulos mensuráveis. Considere então a sequência

$$A_j = \bigcup_{k=1}^{n_j} B_{jk} \times C_{jk},$$

com $A_j \downarrow \emptyset$. Então,

$$\begin{aligned}
\lambda(A_j) &= \sum_{k=1}^{n_j} \int E\big(I_{B_{jk}} \mid \mathcal{G}\big) E\big(I_{C_{jk}} \mid \mathcal{G}\big) \, \mathrm{d}\mu \\
&\leq \sum_{k=1}^{n_j} \big\| E\big(I_{B_{jk}} \mid \mathcal{G}\big) \big\|_2 \big\| E\big(I_{C_{jk}} \mid \mathcal{G}\big) \big\|_2 \\
&\leq \sum_{k=1}^{n_j} \left(\int I_{B_{jk}} \, \mathrm{d}\mu \right)^2 \left(\int I_{C_{jk}} \, \mathrm{d}\mu \right)^2 \\
&= \sum_{k=1}^{n_j} \mu(B_{jk})^2 \mu(C_{jk})^2 \\
&= \sum_{k=1}^{n_j} (\mu \times \mu)(B_{jk} \times C_{jk})^2 \\
&= (\mu \times \mu)(A_j)^2.
\end{aligned}$$

Mas como $\mu \times \mu$ é uma medida e $A_j \downarrow \emptyset$, temos que $(\mu \times \mu)(A_j) \downarrow 0$, e portanto, $(\mu \times \mu)(A_j)^2 \downarrow 0$. □

Definição 4.1.10 (Medida Produto ao Longo de uma σ-Álgebra)**.** *Seja μ uma medida em uma σ-álgebra $\mathcal{F}$ de X, e $\mathcal{G}$ uma sub-σ-álgebra de $\mathcal{F}$. Definimos a* medida produto $\mu \times_{\mathcal{G}} \mu$ ao longo de $\mathcal{G}$ *como sendo a medida λ da proposição 4.1.9.*

Proposição 4.1.11. *Se $T : X \to X$ é um sistema dinâmico, μ é uma medida T-ergódica e $\mathcal{P}$ é a σ-álgebra de Pinsker de T, então, fazendo $\lambda = \mu \times_{\mathcal{P}} \mu$,*

1. *A medida λ é ergódica.*
2. *Se $h_\mu(T) > 0$, então $\lambda(\Delta) = 0$.*
3. *Se $h_\mu(T) > 0$, e se $B \subset X \times X$ é o conjunto do lema 4.1.5 para o sistema $(T \times T)$ dotado da medida λ, então*

$$\operatorname{supp}(\lambda) \subset \overline{B \setminus \Delta}.$$

4. Se $h_\mu(T) > 0$, então $\Delta_{\mathrm{supp}(\mu)} \subset \mathrm{supp}(\lambda)$. Em particular, $\Delta \cap \mathrm{supp}(\lambda) \neq \emptyset$.

Demonstração. — Item (1).

Este é o conteúdo do Corolário 18.20 de [Gla03] (veja também a Definição 9.22 em [Gla03]).

— Item (2).

Este item também está enunciado — mas não está demonstrado — no Corolário 18.20 de [Gla03]. Suponha que $\lambda(\Delta) \neq 0$. A diagonal é evidentemente invariante. Pela ergodicidade de λ, dada pelo item (1), temos que $\lambda(\Delta) = 1$. Então, dado um conjunto mensurável $A \subset X$ qualquer, como $(A \times A^c) \cap \Delta = \emptyset$, temos que

$$0 = \lambda(A \times A^c) = \int E\big(I_A \,\big|\, \mathcal{P}\big) E\big(I_{A^c} \,\big|\, \mathcal{P}\big) \,\mathrm{d}\mu.$$

Se notarmos que para um conjunto mensurável C qualquer, $0 \leq E\big(I_C \,\big|\, \mathcal{P}\big)$, a igualdade acima implica que

$$E\big(I_A \,\big|\, \mathcal{P}\big) E\big(I_{A^c} \,\big|\, \mathcal{P}\big) = 0 \quad [\mu - \text{qtp}]. \tag{4.2}$$

Seja

$$B = \left\{ x \in X \,\middle|\, E\big(I_A \,\big|\, \mathcal{P}\big) > 0 \right\}.$$

Então, $B \in \mathcal{P}$, pois $E\big(I_A \,\big|\, \mathcal{P}\big)$ é $\mathcal{P}$-mensurável. Assim,

$$\mu(A \setminus B) = \int_{B^c} I_A \,\mathrm{d}\mu = \int_{B^c} E\big(I_A \,\big|\, \mathcal{P}\big) \,\mathrm{d}\mu = 0.$$

Por outro lado, a equação (4.2) implica que para quase todo ponto $x \in B$, $E\big(I_{A^c} \,\big|\, \mathcal{P}\big)(x) = 0$. Assim,

$$\mu(B \setminus A) = \int_B I_{A^c} \,\mathrm{d}\mu = \int_B E\big(I_{A^c} \,\big|\, \mathcal{P}\big) \,\mathrm{d}\mu = 0.$$

Como os conjuntos de medida nula estão em $\mathcal{P}$, $\mathcal{P}$ é uma σ-álgebra e $B \in \mathcal{P}$, concluímos que

$$A = (A \cup B) \setminus (B \setminus A) = ((A \setminus B) \cup B) \setminus (B \setminus A) \in \mathcal{P}.$$

Ou seja, todo boreliano está em $\mathcal{P}$. Mas isto implica que para uma partição mensurável finita qualquer $\mathcal{C}$, como

$$\mathcal{C} = \bigvee_{C \in \mathcal{C}} \{C, C^c\},$$

temos que

$$h(T \mid \mathcal{C}) \leq \sum_{C \in \mathcal{C}} h\Big(T \,\Big|\, \{C, C^c\}\Big) = 0.$$

Ou seja, $h_\mu(T) = 0$.

— Item (3).

É imediato dos itens (1) e (2) que $\lambda(B \setminus \Delta) = 1$. Em particular, o conjunto fechado $\overline{B \setminus \Delta}$ possui medida 1. Pela definição de suporte, isso implica que

$$\operatorname{supp}(\lambda) \subset \overline{B \setminus \Delta}.$$

— Item (4).

Considere $(z, z) \in \Delta_{\operatorname{supp}(\mu)}$ e uma vizinhança aberta $A \subset X \times X$ de (z, z). Como os conjuntos da forma $U \times U$, com $U \subset X$ vizinhança aberta de z compõem uma base de vizinhanças do ponto (z, z), então existe um tal $U \subset X$ com $U \times U \subset A$. Assim,

$$\lambda(A) \geq \lambda(U \times U) = \int E\big(I_U \mid \mathcal{P}\big)^2 \,\mathrm{d}\mu.$$

Mas como, $0 \leq E\big(I_U \mid \mathcal{P}\big) \leq 1$, temos que

$$\int E\big(I_U \mid \mathcal{P}\big)^2 \,\mathrm{d}\mu = 0 \Leftrightarrow \int E\big(I_U \mid \mathcal{P}\big) \,\mathrm{d}\mu = 0 \Leftrightarrow \mu(U) = 0.$$

Em particular, se $\lambda(A) = 0$, então $\mu(U) = 0$. No entanto, se $z \in \operatorname{supp}(\mu)$, temos que $\mu(U) \neq 0$. $\square$

Os resultados deste capítulo estão intimamente relacionados com a ergodicidade das medidas consideradas. Para chegarmos no nosso resultado principal, vamos relacionar a entropia topológica com as entropias de Kolmogorov-Sinai das medidas ergódicas. Note que pela Proposição 2.2.4, um sistema dinâmico do tipo produto satisfaz as hipóteses da primeira parte do Corolário 3.1.7.

Teorema 4.1.12. *Seja $T : X \to X$ um sistema dinâmico do tipo produto que admite uma compactificação $\widetilde{T} : \widetilde{X} \to \widetilde{X}$, onde $\widetilde{X} = \prod \widetilde{X}_j$. Suponha que $h(T) > 0$. Então, existe uma medida T-ergódica μ, possuindo pares de Li-Yorke em todo aberto $A \subset X$ que intersecta* $\operatorname{supp}(\mu)$. *Em particular, o resultado vale para todo sistema dinâmico metrizável, separável e localmente compacto.*

Demonstração. Pelo *princípio variacional* (Teorema 3.1.2), juntamente com o Corolário 3.1.7, existe uma medida de probabilidade de Radon T-ergodica μ tal que $h_\mu(T) > 0$.

Note que como $A \times A$ intercepta $\operatorname{supp}(\lambda)$, então, pelo item (3) da Proposição 4.1.11, $A \times A$ também intersecta o conjunto $B \setminus \Delta$, onde B é o conjunto do Lema 4.1.5. Agora, é só notar que todo elemento de $B \setminus \Delta$ é um par de Li-Yorke. De fato, pelo lema 4.1.5, os pontos desse conjunto são recorrentes, e pelo mesmo lema, os pontos de B possuem órbita que passa por qualquer aberto que tenha interseção com $\Delta \cap \operatorname{supp}(\lambda)$, que pelo item (4) da Proposição 4.1.11, é não vazio.

Um sistema localmente compacto é do tipo produto. E pela Proposição 1.2.21, todo sistema dinâmico metrizável separável admite uma compactificação. Assim, o resultado vale para todo sistema dinâmico metrizável separável localmente compacto. $\square$

CAPÍTULO 5

Dinâmica de Endomorfismos

5.1 Caso Linear

Os grupos de Lie lineares são casos particulares de grupos de Lie. Além de motivar as técnicas que serão utilizadas para os grupos de Lie que não são necessariamente lineares, o objetivo desta seção é determinar a entropia das conjugações de grupos lineares, ou seja, determinar a entropia de transformações da forma

$$\begin{array}{rccc} C_g : & G & \to & G \\ & x & \mapsto & gxg^{-1} \end{array},$$

onde $G \subset \mathrm{Gl}(n)$ é um grupo linear, $\mathrm{Gl}(n)$ é o grupo das transformações lineares inversíveis de $\mathbb{R}^n$, e $g \in G$. Com a entropia das conjugações, determinaremos a entropia dos endomorfismos de grupos de Lie nilpotentes e redutíveis (veja as Seções 5.2 e 5.3).

Aplicações Lineares

Seja V um espaço vetorial real de dimensão finita e

$$T : V \to V$$

um automorfismo linear. Vamos mostrar que a entropia topológica de T é zero. Utilizaremos o Corolário 3.1.6, e portanto, o primeiro passo é determinar o conjunto recorrente de T. Considere a decomposição de Jordan multiplicativa

$$T = T_E T_H T_U,$$

onde T_E, T_H e T_U comutam entre si, e são, as componentes elíptica, hiperbólica e unipotente de T. Ou seja, $T_U = I + N$ para alguma aplicação nilpotente N, T_H é diagonalizável com autovalores estritamente positivos, e T_E é diagonalizável na complexificação $\mathbb{C}V$ de V, com autovalores complexos de norma 1. A proposição a seguir é essencialmente a Proposição 4.2 de [Pat10].

Proposição 5.1.1. *Denotando por* $\mathrm{fix}(T_H)$ *e* $\mathrm{fix}(T_U)$ *os pontos fixos de* T_H *e* T_U*,*

$$\mathcal{R}(T) = \mathrm{fix}(T_H) \cap \mathrm{fix}(T_U).$$

Além disso, $h\,(T) = 0$*.*

Demonstração. Considere a aplicação

$$\begin{array}{rccc} \mathbb{P}T : & \mathbb{P}V & \to & \mathbb{P}V \\ & [v] & \mapsto & [Tv] \end{array}$$

induzida por T no espaço projetivo. Pelo Teorema 2.2.11 de [Cal10], ou pelo Teorema A.8 de [FPS10], sabemos que

$$\mathcal{R}(\mathbb{P}T) = \mathrm{fix}(\mathbb{P}T_H) \cap \mathrm{fix}(\mathbb{P}T_U).$$

Note que os pontos fixos de uma transformação linear projetivizada são justamente a projetivização dos os auto-vetores da transformação original. Como $[\mathcal{R}(T)] \subset \mathcal{R}(\mathbb{P}T)$, temos que

$$\mathcal{R}(T) \subset \mathrm{eig}(T_H) \cap \mathrm{eig}(T_U) = \mathrm{eig}(T_H) \cap \mathrm{fix}(T_U),$$

onde $\mathrm{eig}(\cdot)$ denota o conjunto dos auto-vetores da aplicação linear. A última igualdade segue do fato de que como T_U é unipotente, seus auto-valores são sempre 1, e consequentemente, $\mathrm{eig}(T_U) = \mathrm{fix}(T_U)$.

A parte elíptica T_E é diagonalizável na complexificação $\mathbb{C}V$ de V, com auto-valores de módulo 1. Dada uma base $x_1, \dots, x_n$ que diagonaliza T_E, considere a métrica d induzida pela norma

$$\begin{array}{rccc} \|\cdot\| : & \mathbb{C}V & \to & \mathbb{C}V \\ & \sum \alpha_i x_i & \mapsto & \max |\alpha_i| \end{array}.$$

Com esta métrica, T_E é uma isometria. Assim, se $v \in \mathrm{eig}(T_H) \cap \mathrm{fix}(T_U)$ é tal que $T_H v = \lambda v$, então,

$$\|T^n v\| = \|T_E^n T_H^n T_U^n v\| = \|T_H^n v\| = |\lambda|^n \|v\|.$$

Portanto, se v for um ponto recorrente de T, notando que $\lambda > 0$, temos que $\lambda = 1$. Ou seja, $v \in \mathrm{fix}(T_H)$. Sendo assim,

$$\mathcal{R}(T) \subset \mathrm{fix}(T_H) \cap \mathrm{fix}(T_U).$$

Por outro lado, quando restrita a $\mathrm{fix}(T_H) \cap \mathrm{fix}(T_U)$, a transformação T é elíptica, e portanto, todos os pontos são recorrentes. Assim,

$$\mathrm{fix}(T_H) \cap \mathrm{fix}(T_U) \subset \mathcal{R}(T).$$

Para ver que $h\,(T) = 0$, note que $\mathcal{R}(T)$ é fechado. Da caracterização de $\mathcal{R}(T)$ que acabamos de demonstrar, temos que

$$T|_{\mathcal{R}(T)} = T_E|_{\mathcal{R}(T)}.$$

Com a métrica d induzida pela norma $\|\cdot\|$ que torna T_E uma isometria, $T|_{\mathcal{R}(T)} = T_E|_{\mathcal{R}(T)}$ também é isometria. Assim, $h_d\left(T|_{\mathcal{R}(T)}\right) = 0$. Pelo princípio variacional,

$$h_\mu\left(T|_{\mathcal{R}(T)}\right) \leq h_d\left(T|_{\mathcal{R}(T)}\right) = 0.$$

□

Conjugações

Seja o sistema dinâmico $T_g : G \to G$, onde $G \subset \mathrm{Gl}(n)$ é um grupo linear, g é um elemento de G, e

$$\begin{array}{rccc} T_g : & G & \to & G \\ & x & \mapsto & gxg^{-1} \end{array}$$

é uma conjugação.

A conjugação pode ser estendida a $\mathfrak{gl}(n)$:

$$\begin{array}{rccc} \tilde{T}_g : & \mathfrak{gl}(n) & \to & \mathfrak{gl}(n) \\ & X & \mapsto & gXg^{-1} \end{array},$$

onde $\mathfrak{gl}(n)$ é a álgebra de Lie de G. Ou seja, $\mathfrak{gl}(n)$ é o espaço vetorial formado por todas as transformações lineares de $\mathbb{R}^n$. Por ser uma aplicação linear,

$$h\left(\tilde{T}_g\right) = 0.$$

Como a entropia da restrição não pode ser maior que a da extensão, temos que

$$h\,(T_g) = 0.$$

5.2 Endomorfismos de Grupos de Lie Nilpotentes

Dado um sistema dinâmico $T : X \to X$ com base enumerável, o teorema de recorrência de Poincaré nos garante que para qualquer medida de probabilidade μ definida nos borelianos de X que seja T-invariante, $\operatorname{supp}(\mu) \subset \overline{\mathcal{R}}$. Sendo assim, uma importante consequência do princípio variacional é o fato de que para calcular a entropia de sistemas dinâmicos do tipo produto, podemos nos restringir ao conjunto $\overline{\mathcal{R}}$. As técnicas empregadas neste capítulo para determinar o conjunto recorrente foram desenvolvidas para o caso de grupos de Lie semi-simples agindo em variedades *flag*, no artigo [FPS10]. Em [Fer07], Ferraiol utiliza uma técnica semelhante para demonstrar que a entropia de translações em variedades *flag* é sempre zero. E em [Pat10], Patrão também utiliza esta técnica para demonstrar que a entropia de automorfismos de grupos de Lie nilpotentes conexos é sempre zero (Veja o Teorema 4.4 em [Pat10]). Neste capítulo, faremos o mesmo que Patrão em [Pat10].

Caso Abeliano

Como o caso abeliano já estar incluso no caso nilpotente (que será feito logo em seguida), vamos analisá-lo aqui, apenas a título de ilustração, sem muitos detalhes. Os grupos de Lie Abelianos conexos são difeomorfos a $\mathbb{T}^m \times \mathbb{R}^n$, onde $\mathbb{T}^m$ é o toro m-dimensional. Os endomorfismos de $\mathbb{T}^m \times \mathbb{R}^n$, podem ser representados por matrizes da forma

$$\widetilde{T} = \begin{pmatrix} S & * \\ 0 & T \end{pmatrix}, \tag{5.1}$$

pois todo elemento de $\mathbb{T}^m$ tem que ser levado em $\mathbb{T}^m$. Vamos mostrar que a entropia de $\widetilde{T}$ é exatamente igual à entropia de S. Como S é a restrição de $\widetilde{T}$ a $\mathbb{T}^m$, é evidente que $h(S) \leq h\left(\widetilde{T}\right)$.

Note que para que um ponto $(a, b) \in \mathbb{T}^m \times \mathbb{R}^n$ seja recorrente, é necessário que $b \in \mathcal{R}(T)$. Ou seja, $\mathcal{R}\left(\widetilde{T}\right) \subset \mathbb{T}^m \times \mathcal{R}(T)$. Já mostramos que $\mathcal{R}(T)$, é fechado; e que T restrito a $\mathcal{R}(T)$ é o mesmo que T_E, a parte elíptica de T. Assim, fazendo

$$\widetilde{T}_E = \begin{pmatrix} S & * \\ 0 & T_E \end{pmatrix},$$

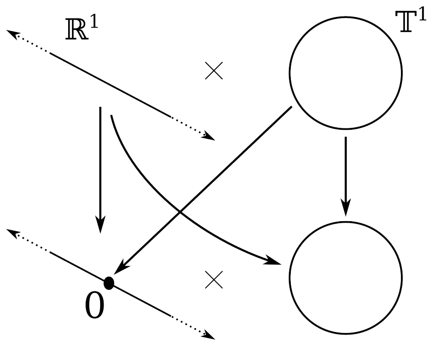

Figura 5.1: A imagem de $\mathbb{R}^n \times \{0\}$ pode ter componente em $\mathbb{T}^m$. No entanto, $\{0\} \times \mathbb{T}^m$, por ser compacto, pode ser levado apenas em elementos com componente $\mathbb{R}^n$ nula.

sabemos que

$$\begin{aligned} h\left(\widetilde{T}\right) &= h\left(\widetilde{T}|_{\overline{\mathcal{R}(\widetilde{T})}}\right) \\ &\leq h\left(\widetilde{T}|_{\mathbb{T}^m \times \mathcal{R}(T)}\right) \\ &= h\left(\widetilde{T}_E|_{\mathbb{T}^m \times \mathcal{R}(T)}\right) \\ &\leq h\left(\widetilde{T}_E\right). \end{aligned}$$

Pela fórmula de Bowen (Proposição A.5.1) e pelo princípio variacional,

$$h\left(\widetilde{T}_E\right) \leq \sum_{\substack{\lambda:\ \text{auto-valor de } \widetilde{T}_E \\ |\lambda|>1}} \log|\lambda| = \sum_{\substack{\lambda:\ \text{auto-valor de } S \\ |\lambda|>1}} \log|\lambda|.$$

Pela fórmula de Bowen (Proposição A.5.1), e pelo princípio variacional do caso compacto metrizável, o último termo é igual a $h(S)$. Assim,

$$h\left(\widetilde{T}\right) \leq h(S).$$

Caso Nilpotente

Vamos começar mostrando que todo endomorfismo sobrejetivo ϕ de um grupo de Lie nilpotente conexo G é uma aplicação própria. Como ϕ' é um endomorfismo sobrejetivo da álgebra de Lie $\mathfrak{g}$ de G, ϕ' é um isomorfismo linear e portanto uma aplicação própria. Se G é nilpotente simplesmente conexo, temos que ϕ é conjugado a ϕ' e o seguinte diagrama comuta

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\phi'"] \arrow[d, "\exp"'] & \mathfrak{g} \arrow[d, "\exp"] \\ G \arrow[r, "\phi"'] & G \end{tikzcd} .$$

De fato, como G é nilpotente, a exponencial é bijetiva, e pelo Teorema 3.32 de [War83], o diagrama comuta. Assim, ϕ é uma aplicação própria. Podemos reduzir o caso nilpotente conexo ao caso simplesmente conexo. Para tanto, precisaremos do Lema 5.2.1 e da Proposição 5.2.3.

Lema 5.2.1. *Seja $\phi : G \to G$ um endomorfismo sobrejetivo de um grupo de Lie G. Então,*

$$\phi \textit{ é própria} \Leftrightarrow \ker(\phi) \textit{ é compacto.}$$

Demonstração. Sendo ϕ uma aplicação própria, $\ker(\phi)$ é evidentemente compacto, pois é a imagem inversa do conjunto compacto $\{1_G\}$. Por outro lado, um endomorfismo sobrejetivo é sempre uma aplicação aberta. Se $\ker(\phi)$ é compacto, então ϕ é uma sobrejeção contínua com fibras compactas. É um fato conhecido que qualquer sobrejeção contínua e aberta com fibras compactas é uma aplicação própria. □

De posse de um critério para que um endomorfismo sobrejetivo seja uma aplicação própria, vamos mostrar que quando G é um grupo de Lie nilpotente conexo e $T(G)$ sua componente toral (o maior subgrupo compacto e conexo do centro $Z(G)$), então $G/T(G)$ é simplesmente conexo. É fácil ver que $T(G)$ existe, pois $Z(G)$ é abeliano, e portanto sua componente conexa é homeomorfa a $\mathbb{T}^m \times \mathbb{R}^n$.

Definição 5.2.2 (Componente Toral)**.** *Seja G um grupo de Lie. A* componente toral *de G é o maior subgrupo compacto e conexo de seu centro $Z(G)$.*

Proposição 5.2.3. *Seja G um grupo de Lie nilpotente conexo, e $T(G)$ sua componente toral. Então $G/T(G)$ é simplesmente conexo.*

Demonstração. Seja $r : \widetilde{G} \to G$ o recobrimento universal de G. Vamos utilizar o seguinte diagrama

$$
\begin{tikzcd}
\mathfrak{g} \arrow[r, leftrightarrow, "\exp"] \arrow[d] & \widetilde{G} \arrow[d, "\pi"] \arrow[dr, "r"] & \\
\dfrac{\mathfrak{g}}{\mathfrak{z}} \arrow[r, leftrightarrow, "\exp(\,\cdot\,+\mathfrak{z})"] \arrow[d] & \dfrac{\widetilde{G}}{\exp(\mathfrak{z})} \arrow[r, leftrightarrow, "p"] \arrow[d] & G \arrow[d] \\
\dfrac{\mathfrak{g}}{\langle\mathfrak{z}\rangle} \arrow[r, leftrightarrow, "\exp(\,\cdot\,+\langle\mathfrak{z}\rangle)"] & \dfrac{\widetilde{G}}{\exp(\langle\mathfrak{z}\rangle)} \arrow[r, leftrightarrow, "q"] & \dfrac{G}{T(G)}
\end{tikzcd}
\quad ,
$$

onde $\mathfrak{z} = (r \circ \exp)^{-1}(1_G)$ são os elementos da álgebra de Lie tais que sua exponencial em G é 1_G, e $\langle\mathfrak{z}\rangle$ é o subespaço vetorial gerado por $\mathfrak{z}$.

A aplicação exp é um difeomorfismo entre $\mathfrak{g}$ e $\widetilde{G}$. A diferença entre ambos, é que $\mathfrak{g}$ possui estrutura de espaço vetorial, enquanto que $\widetilde{G}$ possui estrutura de grupo de Lie. Mas como variedades diferenciáveis, são idênticas.

Para demonstrar a proposição, precisamos mostrar que p, q, $\exp(\cdot + \mathfrak{z})$ e $\exp(\cdot + \langle\mathfrak{z}\rangle)$ estão bem definidas e são homeomorfismos, pois $\mathfrak{g}/\langle\mathfrak{z}\rangle$ é um espaço vetorial, e portanto, é simplesmente conexo.

Afirmação. $Z(\widetilde{G}) = \exp(\ker(\mathrm{ad}))$.

De fato,

$$
\begin{aligned}
\exp(X) \in Z(\widetilde{G}) &\Leftrightarrow \exp(X)\exp(Y)\exp(-X) = \exp(Y) && (\forall Y \in \mathfrak{g}) \\
&\Leftrightarrow \exp\left(\mathrm{Ad}\left(\exp(X)\right)Y\right) = \exp(Y) && (\forall Y \in \mathfrak{g}) \\
&\Leftrightarrow \exp\left(\mathrm{e}^{\mathrm{ad}(X)}\,Y\right) = \exp(Y) && (\forall Y \in \mathfrak{g}) \\
&\Leftrightarrow \mathrm{e}^{\mathrm{ad}(X)}\,Y = Y && (\forall Y \in \mathfrak{g}) \\
&\Leftrightarrow \mathrm{e}^{\mathrm{ad}(X)} = \mathrm{id} \\
&\Leftrightarrow \mathrm{ad}(X) = 0 \\
&\Leftrightarrow X \in \ker(\mathrm{ad}),
\end{aligned}
$$

onde o fato de $\mathrm{e}^{\mathrm{ad}(x)} = \mathrm{id}$ implicar em $\mathrm{ad}(x) = 0$ segue do fato de que como $\mathfrak{g}$ é nilpotente, então $\mathrm{ad}(\mathfrak{g})$ também é nilpotente, e a exponencial é um difeomorfismo em álgebras de Lie nilpotentes (veja Proposição 1.32 de [Kna02]).

Afirmação. *As aplicações* $\exp(\cdot + \mathfrak{z})$ *e* $\exp(\cdot + \langle \mathfrak{z} \rangle)$ *estão bem definidas e são homeomorfismos.*

Vamos mostrar que as aplicações estão bem definidas e são bijetivas. Neste caso, a comutatividade do diagrama e a definição de topologia quociente implicarão que são homeomorfismos. Para tanto, basta mostrar que $\langle \mathfrak{z} \rangle \subset \ker(\mathrm{ad})$. De fato, neste caso, $\exp(X + \mathfrak{z}) = \exp(X)\exp(\mathfrak{z})$ e $\exp(X + \langle \mathfrak{z} \rangle) = \exp(X)\exp(\langle \mathfrak{z} \rangle)$.

Note que $\exp(\mathfrak{z}) = r^{-1}(1_G)$ é discreto e normal (Proposição 1.101 de [Kna02]). Por ser normal, para $Z \in \mathfrak{z}$,

$$\begin{array}{rccc} F_Z : & \widetilde{G} & \to & \exp(\mathfrak{z}) \\ & g & \mapsto & g\exp(Z)g^{-1} \end{array}$$

está definida e é contínua. Como $\widetilde{G}$ é conexo e $\exp(\mathfrak{z})$ é discreto, temos que F_Z é constante igual a $F_Z(1_G) = \exp(Z)$. Ou seja, $\exp(Z) \in Z(\widetilde{G})$. E portanto, $\exp(\mathfrak{z}) \subset Z(\widetilde{G})$. Pela Afirmação anterior e pela injetividade de exp,

$$\mathfrak{z} \subset \ker(\mathrm{ad}).$$

Mas como ad é linear, temos que $\langle \mathfrak{z} \rangle \subset \ker(\mathrm{ad})$.

Note que $\widetilde{G}/\exp(\mathfrak{z})$ é um espaço homogêneo, pois $\exp(\mathfrak{z})$ é um subgrupo fechado. A aplicação

$$\begin{array}{rccc} p : & \widetilde{G}/\exp(\mathfrak{z}) & \to & G \\ & \exp(X)\exp(\mathfrak{z}) & \mapsto & (r \circ \exp)(X) \end{array}$$

está bem definida e é uma bijeção, pois $\exp(\mathfrak{z})$ é o kernel de r. E é contínua pela definição de topologia quociente. Pelo Teorema 3.58 de [War83], que garante a unicidade da estrutura de variedade nos espaços homogêneos, p é um difeomorfismo. Para que q esteja bem definida e seja um homeomorfismo, basta que a afirmação seguinte seja verdadeira, pois neste caso, a definição de topologia quociente, e o fato de p ser homeomorfismo (difeomorfismo) implicará na continuidade de q e q^{-1}.

Afirmação. $T(G) = r\left(\exp(\langle \mathfrak{z} \rangle)\right)$.

Note que p leva $\pi(\exp(\langle \mathfrak{z} \rangle))$ em $r(\exp(\langle \mathfrak{z} \rangle))$, que são, portanto, difeomorfos. Por sua vez, o homeomorfismo $\exp(\cdot + \mathfrak{z})$ mostra que $\pi(\exp(\langle \mathfrak{z} \rangle))$ é homeomorfo a $\langle \mathfrak{z} \rangle / \mathfrak{z}$, que é compacto. Ou seja, $r(\exp(\langle \mathfrak{z} \rangle))$ é compacto. Já sabemos que $\exp(\langle \mathfrak{z} \rangle) \subset Z(\widetilde{G})$. Portanto, pela sobrejetividade de r,

$$r(\exp(\langle \mathfrak{z} \rangle)) \subset r(Z(\widetilde{G})) \subset Z(G).$$

Assim,

$$r(\exp(\langle \mathfrak{z} \rangle)) \subset T(G).$$

Por outro lado, se

$$r(\exp(X)) \in T(G),$$

então, pela compacidade de $T(G)$, existem $n_k \to \infty$ e $Y \in \mathfrak{g}$ tais que

$$r\left(\exp(n_k X)\right) \to r\left(\exp(Y)\right).$$

Ou seja,

$$\exp(n_k X + \mathfrak{z}) \to \exp(Y + \mathfrak{z}).$$

E portanto, como $\exp(\cdot + \mathfrak{z})$ é um homeomorfismo,

$$n_k X + \mathfrak{z} \to Y + \mathfrak{z}.$$

Em particular,

$$n_k X + \langle \mathfrak{z} \rangle \to Y + \langle \mathfrak{z} \rangle.$$

Mas isso só pode ocorrer quando $X \in \langle \mathfrak{z} \rangle$ (e $Y \in \langle \mathfrak{z} \rangle$), pois caso contrário, em uma norma qualquer em $\mathfrak{g}/\langle \mathfrak{z} \rangle$, $\|n_k X + \langle \mathfrak{z} \rangle\| \to \infty$ Assim,

$$T(G) \subset r(\exp(\langle \mathfrak{z} \rangle)).$$

□

Exemplo 5.2.4 (Grupo de Heisenberg)**.** O grupo de Heisenberg sobre os números reais é o subgrupo de Gl(3) formado pelas matrizes da forma

$$\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix},$$

onde a, b, c são números reais. Sua álgebra de Lie é formada pelas matrizes da forma

$$\begin{pmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix},$$

e o centro da álgebra é dado pelas matrizes tais que $a = b = 0$. A aplicação exponencial — $(r \circ \exp)$ na notação da Proposição 5.2.3 — é dada por

$$\begin{aligned} r \circ \exp \begin{pmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix} &= I + \begin{pmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix}^2 \\ &= \begin{pmatrix} 1 & a & c + \frac{1}{2}ab \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

Portanto, $\mathfrak{z} = (r \circ \exp)^{-1}(1_G)$ é trivial. Ou seja, o grupo de Heisenberg é difeomorfo a seu recobrimento universal. Portanto, o grupo é simplesmente conexo e sua componente toral é trivial.

Exemplo 5.2.5. Considere o grupo formado pelas matrizes da forma

$$\begin{pmatrix} 1 & a + \mathbb{Z} & c + \mathbb{Z} \\ 0 & 1 & b + \mathbb{Z} \\ 0 & 0 & 1 \end{pmatrix},$$

onde a, b, c são números reais. Sua álgebra de Lie é formada pelas matrizes da forma

$$\begin{pmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix},$$

e o centro da álgebra é dado pelas matrizes tais que $a = b = 0$. A aplicação exponencial — $(r \circ \exp)$ na notação da Proposição 5.2.3 — é dada por

$$r \circ \exp \begin{pmatrix} 0 & a & c \\ 0 & 0 & b \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & a + \mathbb{Z} & c + \frac{1}{2}ab + \mathbb{Z} \\ 0 & 1 & b + \mathbb{Z} \\ 0 & 0 & 1 \end{pmatrix}.$$

Os elementos de $(r \circ \exp)^{-1}(1_G)$ são as matrizes tais que $a, b, c \in \mathbb{Z}$. Essas matrizes geram a álgebra de Lie, e portanto, este grupo não é simplesmente conexo, e sua componente toral é todo o grupo. O que de fato é evidente, dado que o grupo é compacto e conexo.

Finalmente, mostramos que todo endomorfismo sobrejetivo de um grupo de Lie nilpotente conexo é uma aplicação própria.

Proposição 5.2.6. *Seja $\phi : G \to G$ um endomorfismo sobrejetivo de um grupo de Lie nilpotente conexo G. Então ϕ é uma aplicação própria.*

Demonstração. Pela Proposição 5.2.3, temos que $\widetilde{G} = G/T(G)$ é simplesmente conexo. Além disso, para qualquer endomorfismo sobrejetivo $\phi : G \to G$, existe um outro endomorfismo sobrejetivo $\widetilde{\phi} : \widetilde{G} \to \widetilde{G}$, tal que o seguinte diagrama comuta

$$\begin{tikzcd} G \arrow[r, "\phi"] \arrow[d, "\pi"'] & G \arrow[d, "\pi"] \\ \widetilde{G} \arrow[r, "\widetilde{\phi}"'] & \widetilde{G} \end{tikzcd} ,$$

onde π é a projeção canônica. De fato, como $\phi\,(T(G))$ é compacto, conexo e abeliano, está necessariamente contido em $T(G)$. Assim, podemos definir

$$\widetilde{\phi}(\pi(x)) = \pi(\phi(x)) = \phi(x)T(G).$$

Como $\widetilde{G}$ é simplesmente conexo, temos que $\widetilde{\phi}$ é um automorfismo de $\widetilde{G}$. Afirmamos que $\ker(\phi)$ é um subconjunto fechado do compacto $T(G)$. De fato, temos que

$$\begin{aligned} \phi^{-1}\left(1_G\right) &\subset \phi^{-1}\left(T(G)\right) \\ &= \phi^{-1}\left(\pi^{-1}(1_{\widetilde{G}})\right) \\ &= \pi^{-1}\left(\widetilde{\phi}^{-1}(1_{\widetilde{G}})\right) \\ &= \pi^{-1}\left(1_{\widetilde{G}}\right) \\ &= T(G). \end{aligned}$$

Portanto, $\ker(\phi)$ é compacto e, pelo Lema 5.2.1, sabemos que ϕ é uma aplicação própria. □

Vamos mostrar que, como no caso abeliano, a entropia de um endomorfismo sobrejetivo de um grupo de Lie nilpotente G coincide com a entropia de sua restrição à componente toral de G. Antes, vamos a uma proposição referente ao caso compacto.

Proposição 5.2.7. *Seja G um grupo de Lie e $T(G) \subset G$ sua componente toral. Denote por $\pi : G \to G/T(G)$ a projeção canônica. Então, se*

$\phi : G \to G$ é um endomorfismo de G e $K \subset G$ um subconjunto compacto invariante da forma $\pi^{-1}(B)$ para algum $B \subset G/T(G)$, temos que

$$h\left(\phi|_K\right) = h\left(\widetilde{\phi}|_B\right) + h\left(\phi|_{T(G)}\right),$$

onde $\widetilde{\phi} : G/T(G) \to G/T(G)$ é o endomorfismo induzido por ϕ em $G/T(G)$.

Demonstração. Note que B é compacto, pois $B = \pi(K)$. Note também, que os grupos de Lie são metrizáveis, e portanto, $T(G)$, K e B são compacto metrizáveis. Então, basta utilizar a Proposição 3.3.3 de [Fer07], substituindo G por $T(G)$, X por K, T por $\phi|_K$, S por $\widetilde{\phi}|_B$, τ por id, e definindo $p : K \times T(G) \to K$ por $p(x, g) = xg$, para concluir que

$$h_d\left(\phi|_K\right) = h_{\widetilde{d}}\left(\widetilde{\phi}|_B\right) + h_d\left(\phi|_{T(G)}\right),$$

onde d é uma métrica em $T(G)$, e $\widetilde{d}$ é uma métrica em $G/T(G)$. Mas no caso compacto, as entropias de Bowen e a entropia topológica coincidem. Portanto,

$$h\left(\phi|_K\right) = h\left(\widetilde{\phi}|_B\right) + h\left(\phi|_{T(G)}\right).$$

□

Teorema 5.2.8. *Seja $\phi : G \to G$ um endomorfismo sobrejetivo de um grupo de Lie nilpotente conexo G. Então,*

$$h(\phi) = h\left(\phi|_{T(G)}\right).$$

Demonstração. Seja $\widetilde{G} = G/T(G)$, e denote por $\pi : G \to G/T(G)$ a projeção canônica. Pela Proposição 5.2.3, temos que $\widetilde{G}$ é um grupo de Lie nilpotente simplesmente conexo e, como na demonstração da Proposição 5.2.6, podemos considerar o endomorfismo induzido $\widetilde{\phi} : \widetilde{G} \to \widetilde{G}$, que é conjugado através da exponencial a $\widetilde{\phi}' : \mathfrak{g} \to \mathfrak{g}$, sua diferencial na identidade.

Por outro lado, temos que $\widetilde{\phi}'$ é uma aplicação linear e, pela Proposição 4.2 de [Pat10], seu conjunto recorrente $\mathcal{R}(\widetilde{\phi}')$ é fechado. Também sabemos que existe uma norma em $\widetilde{\mathfrak{g}}$ tal que $\widetilde{\phi}'|_{\mathcal{R}(\widetilde{\phi}')}$ é uma isometria. Em particular, para qualquer bola fechada $B \subset \widetilde{\mathfrak{g}}$ com centro em 0, $B \cap \mathcal{R}(\widetilde{\phi}')$ é compacto e $\widetilde{\phi}'$-invariante. Pela conjugação dada pela exponencial, existe uma métrica em $\mathcal{R}(\widetilde{\phi})$ tal que qualquer bola fechada $B \subset \mathcal{R}(\widetilde{\phi})$ centrada na identidade é compacta, com B e B^c $\widetilde{\phi}$-invariantes.

Seja $R = \pi^{-1}(\mathcal{R}(\widetilde{\phi}))$. Então, como R é fechado, segue que $\overline{\mathcal{R}(\phi)} \subset R$. Para $\varepsilon > 0$, existe uma cobertura admissível $\mathcal{A} = \{A_0, \ldots, A_k\}$ de R, tal que A_0 tem complemento compacto, e tal que

$$h(\phi) - \varepsilon = h(\phi|_R) - \varepsilon \leq h\Big(\phi|_R \,\Big|\, \mathcal{A}\Big).$$

Se $B \subset \mathcal{R}(\widetilde{\phi})$ é uma bola compacta $\widetilde{\phi}$-invariante tal que $\pi({A_0}^c) \subset B$. Fazendo $K = \pi^{-1}(B)$, temos que K é compacto, já que π é própria. E também, $N_{R \setminus K}(\mathcal{A}^n) = 1$, pois

$$\begin{aligned} R \setminus K &= (R \setminus K) \cap \cdots \cap T^{-n+1}(R \setminus K) \\ &\subset A_0 \cap \cdots \cap T^{-n+1}(A_0). \end{aligned}$$

Veja a Figura 5.2.

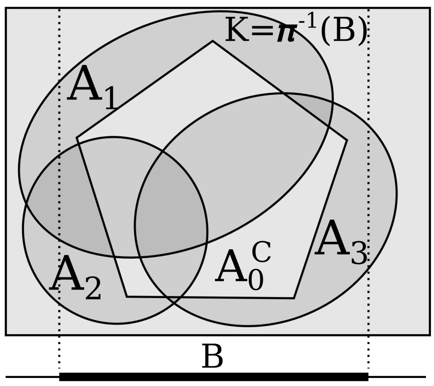

Figura 5.2: Dada uma cobertura admissível $\mathcal{A} = \{A_0, \ldots, A_n\}$, com ${A_0}^c$ compacto, encontramos K compacto invariante com $N_{R\setminus K}(\mathcal{A}) = 1$.

Assim,

$$\begin{aligned} h\Big(\phi|_R \,\Big|\, \mathcal{A}\Big) &= \lim \frac{1}{n} \log N_R(\mathcal{A}^n) \\ &\leq \lim \frac{1}{n} \log \big(N_K(\mathcal{A}^n) + N_{R\setminus K}(\mathcal{A}^n)\big) \\ &= \lim \frac{1}{n} \log (N_K(\mathcal{A}^n) + 1) \\ &= \lim \frac{1}{n} \log N_K(\mathcal{A}^n) \\ &= h\Big(\phi|_K \,\Big|\, \mathcal{A} \cap K\Big). \end{aligned}$$

E portanto,

$$h\,(\phi) - \varepsilon \leq h\Big(\phi|_R \,\Big|\, \mathcal{A}\Big) \leq h\Big(\phi|_K \,\Big|\, \mathcal{A} \cap K\Big) \leq h\,(\phi|_K).$$

Por outro lado, aplicando a Proposição 5.2.7 para o fibrado compacto $T(G)$-principal $\pi|_K : K \to B$, como $\widetilde{\phi}$ é uma isometria quando restrita a B, concluímos que

$$h\,(\phi|_K) = h\left(\widetilde{\phi}|_B\right) + h\left(\phi|_{T(G)}\right) = h\left(\phi|_{T(G)}\right).$$

Assim,

$$h\,(\phi) - \varepsilon \leq h\left(\phi|_{T(G)}\right) \leq h\,(\phi).$$

E como ε é arbitrário, segue que

$$h\,(\phi) = h\left(\phi|_{T(G)}\right).$$

□

5.3 Endomorfismos de Grupos de Lie Redutíveis

Caso Semi-Simples

Vamos mostrar que todo endomorfismo sobrejetivo ϕ de um grupo de Lie conexo semi-simples G é uma aplicação própria. De fato, mostramos que neste caso, o endomorfismo sobrejetivo será um automorfismo.

Proposição 5.3.1. *Seja $\phi : G \to G$ um endomorfismo sobrejetivo de um grupo de Lie conexo semi-simples G. Então existe $k \in \mathbb{N}$ tal que $\phi^k = C_g$ para algum $g \in G$, onde C_g é a conjugação por g. Em particular, ϕ é um automorfismo.*

Demonstração. Note que

$$\phi^k\,(\exp(X)) = \exp\left[(\phi')^k X\right].$$

Mas como $\mathfrak{g}$ é semi-simples, sabemos que existe $k \in \mathbb{N}$ tal que $(\phi')^k$ é um endomorfismo interno de $\mathfrak{g}$ (veja o Teorema 5.4, página 423 de [Hel78]). Ou seja, existe $g \in G$ tal que $(\phi')^k = \mathrm{Ad}(g)$, e portanto,

$$\phi^k\,(\exp(X)) = \exp\left((\phi')^k X\right) = \exp\left(\mathrm{Ad}(g)X\right) = C_g(\exp X).$$

Como G é gerado por elementos da forma $\exp(X)$, segue que $\phi^k = C_g$.

Assim, concluímos que ϕ é um automorfismo, pois $\phi^k = C_g$ é automorfismo. □

Exemplo 5.3.2. Seja $G \subset \mathrm{Gl}(n)$ a componente conexa de $\mathrm{Gl}(n)$ que contém a identidade. Para $h \in G$, defina

$$\begin{array}{rccc} T: & G & \to & G \\ & g & \mapsto & hg^{-t}h^{-1} \end{array},$$

onde $g^{-t} = \left(g^{-1}\right)^t = \left(g^t\right)^{-1}$. Essa notação se beneficia do fato de as operações de transposição e inversão de uma matriz comutarem. Vamos mostrar que $T^2 = C_{hh^{-t}}$. De fato,

$$\begin{aligned} T^2(g) &= h\left(hg^{-t}h^{-1}\right)^{-t}h^{-1} \\ &= h\left(hg^{t}h^{-1}\right)^{t}h^{-1} \\ &= h\left(h^{-t}gh^{t}\right)h^{-1} \\ &= \left(hh^{-t}\right)g\left(h^{t}h^{-1}\right) \\ &= \left(hh^{-t}\right)g\left(hh^{-t}\right)^{-1} \\ &= C_{hh^{-t}}(g). \end{aligned}$$

Agora, caracterizaremos o conjunto recorrente de um endomorfismo sobrejetivo ϕ de um grupo de Lie linear conexo semi-simples G.

A Proposição 5.3.1 mostra que para algum k, ϕ^k é, de fato, igual à conjugação por algum elemento de G. Assim, vamos primeiro considerar a dinâmica das conjugações e posteriormente vamos reduzir o caso geral a este caso. Para $g \in G$, denotamos por G_g o centralizador de g em G. Este é o conjunto dos pontos fixos de C_g.

Lema 5.3.3. *Seja G um grupo de Lie linear semi-simples conexo, e seja $C_g : G \to G$ a conjugação por $g \in G$. Então,*

$$\mathcal{R}(C_g) = G_h \cap G_u,$$

onde $g = ehu$ é a decomposição de Jordan multiplicativa de g. Em particular, restrita a sua componente recorrente, C_g é uma isometria para uma métrica adequada.

Demonstração. Note que $C_g = \mathrm{Ad}(g)|_G$, onde Ad é a representação adjunta de $\mathrm{Gl}(n)$, com $G \leq \mathrm{Gl}(n)$. Note também que o Lema 3.6 de [PSS11] mostra que

$$\mathrm{Ad}(g) = \mathrm{Ad}(e)\,\mathrm{Ad}(h)\,\mathrm{Ad}(u)$$

é a decomposição de Jordan de $\mathrm{Ad}(g)$.

Como o conjunto recorrente se comporta bem com respeito a restrições, sabemos por [Pat10] que

$$\begin{aligned}\mathcal{R}(C_g) &= \mathcal{R}(\mathrm{Ad}(g)) \cap G \\ &= \mathrm{fix}(\mathrm{Ad}(h)) \cap \mathrm{fix}(\mathrm{Ad}(u)) \cap G \\ &= \mathrm{fix}(C_h) \cap \mathrm{fix}(C_u) \\ &= G_h \cap G_u.\end{aligned}$$

□

Vamos agora, utilizar os resultados anteriores para mostrar que todo endomorfismo sobrejetivo de um grupo de Lie semi-simples conexo possui entropia topológica nula.

Teorema 5.3.4. *Seja $\phi : G \to G$ um endomorfismo sobrejetivo de um grupo de Lie semi-simples conexo G. Então,*

$$h\left(\phi\right) = h\left(\phi|_{T(G)}\right) = 0.$$

Demonstração. Como $h\left(\phi^k\right) = kh\left(\phi\right)$, temos que $h\left(\phi\right) = 0$ se, e somente se, $h\left(\phi^k\right) = 0$. Como G é conexo e semi-simples, existe $k > 0$ tal que $\phi^k = C_g$, para algum $g \in G$. Portanto, é suficiente provar que $h\left(C_g\right) = 0$.

Do Teorema 4.1.12, como G é metrizável separável localmente compacto, sabemos que se $h\left(C_g\right) > 0$, então existe um par de Li-Yorke para C_g. Ou seja, existem dois elementos distintos $a, b \in G$, tais que $(C_g^{n_k}(a), C_g^{n_k}(b)) \to (a, b)$ e $(C_g^{m_k}(a), C_g^{m_k}(b)) \to (c, c)$, para algum $c \in G$. Considere $C_{\mathrm{Ad}(g)}$, e note que $\mathrm{Ad} \circ C_g = C_{\mathrm{Ad}(g)} \circ \mathrm{Ad}$. Agora, como $a, b \in \mathcal{R}(C_g)$, também temos que $\mathrm{Ad}(a), \mathrm{Ad}(b) \in \mathcal{R}(C_{\mathrm{Ad}(g)})$. Mas $C_{\mathrm{Ad}(g)}|_{\mathcal{R}(C_{\mathrm{Ad}(g)})} = C_{\mathrm{Ad}(e)}|_{\mathcal{R}(C_{\mathrm{Ad}(g)})}$ é uma isometria para uma distância adequada em $\mathrm{Ad}(G)$. Assim, o fato de que $\left(C_{\mathrm{Ad}(g)}^{m_k}(\mathrm{Ad}(a)), C_{\mathrm{Ad}(g)}^{m_k}(\mathrm{Ad}(b))\right)$ converge para $(\mathrm{Ad}(c), \mathrm{Ad}(c))$ implica que $\mathrm{Ad}(a) = \mathrm{Ad}(b)$.

Portanto, sabemos que $a = wu$ e $b = wv$ para algum $w \in G$ e algum $u, v \in Z(G)$. Também temos que

$$\begin{aligned} C_g^{m_k}(w)u = C_g^{m_k}(a) \to c \\ C_g^{m_k}(w)v = C_g^{m_k}(b) \to c. \end{aligned}$$

Mas isso significa que $u = v$. E assim, $a = b$, contradizendo o fato de que (a, b) é um par de Li-Yorke. □

Caso Redutível

Seja G um grupo de Lie redutível e conexo. Será útil considerar o homomorfismo sobrejetivo

$$\begin{array}{rrcl} \pi: & Z(G)_0 \times G' & \to & G \\ & (z,g) & \mapsto & zg \end{array},$$

onde $Z(G)_0$ é a componente conexa do centro que contém a identidade, e $G' = [G,G]$ é o grupo derivado. Note que como G é redutível, G' é semi-simples. É fácil ver que G' também é conexo.

Lema 5.3.5. *Se G é um grupo de Lie conexo, então G' também é conexo.*

Demonstração. De fato, a aplicação

$$\begin{array}{rrcl} f: & G \times G & \to & G \\ & (g,h) & \mapsto & [g,h] \end{array}$$

é contínua. E como $G \times G$ é conexo, $f(G \times G)$ também é conexo. Se

$$\begin{array}{rrcl} p_n: & G^n & \to & G \\ & (g_1, \ldots, g_N) & \mapsto & g_1 \cdots g_n \end{array}$$

é a aplicação contínua que multiplica n elementos de G, então $P_n = p_n\left(f(G \times G), \ldots, f(G \times G)\right)$ também é um conjunto conexo, e contém a identidade. Mas

$$G' = \bigcup_{n>1} P_n$$

é uma união de conexos que possuem ao menos um ponto em comum. Portanto, G' é conexo. □

Note que, por G ser redutível, G e $Z(G)_0 \times G'$ possuem a mesma álgebra de Lie $\mathfrak{g} = \mathfrak{z} \oplus \mathfrak{g}'$, onde $\mathfrak{g}$ é a álgebra de Lie de G, $\mathfrak{z}$ é a álgebra de Lie de $Z(G)_0$ e $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ é a álgebra de Lie de G' (Corolário 1.56 de [Kna02]).

Lema 5.3.6. *Seja $\phi: G \to G$ um endomorfismo sobrejetivo de um grupo de Lie G redutível e conexo. Então, ϕ induz o homomorfismo sobrejetivo*

$$\widetilde{\phi} = \phi|_{Z(G)_0} \times \phi|_{G'}$$

em $Z(G)_0 \times G'$, tal que o diagrama

$$\begin{tikzcd} Z(G)_0 \times G' \arrow[r, "\widetilde{\phi}"] \arrow[d, "\pi"'] & Z(G)_0 \times G' \arrow[d, "\pi"] \\ G \arrow[r, "\phi"'] & G \end{tikzcd}$$

comuta.

Demonstração. É evidente que

$$\phi(Z(G)_0) \subset Z(G)_0,$$

pois $\phi(Z(G)_0)$ é um subgrupo conexo de $Z(G)$ que contém a identidade, e portanto está contido em $Z(G)_0$. Acontece que

$$\phi(G') \subset G'.$$

O que implica que $\phi'(\mathfrak{z}) = \mathfrak{z}$ e $\phi'(\mathfrak{g}') = \mathfrak{g}'$. E como ambos, $Z(G)_0$ e G' são conexos, temos que $\phi|_{Z(G)_0} : Z(G)_0 \to Z(G)_0$ e $\phi|_{G'} : G' \to G'$ são sobrejetivos.

A comutatividade do diagrama é uma consequência imediata do fato de ϕ ser um homomorfismo. $\square$

Proposição 5.3.7. *Seja $\phi : G \to G$ um endomorfismo sobrejetivo de um grupo de Lie redutível e conexo G. Se π é uma aplicação própria, então ϕ também é própria.*

Demonstração. Primeiro, observe que o endomorfismo $\widetilde{\phi}$ apresentado no Lema 5.3.6 é próprio, pois é o produto de duas aplicações próprias. De fato, temos que $\phi|_{Z(G)_0}$ e $\phi|_{G'}$ são endomorfismos próprios pelas Proposições 5.2.6 e 5.3.1, respectivamente. Considerando o diagrama no Lema 5.3.6, temos que se $K \subset G$ é compacto, então $\phi^{-1}(K) = \pi \circ \widetilde{\phi}^{-1} \circ \pi^{-1}(K)$ também é compacto, pois $\widetilde{\phi}$ e π são aplicações próprias. $\square$

Entropia Topológica

Iniciamos calculando a entropia topológica do endomorfismo associado $\widetilde{\phi}$.

Proposição 5.3.8. *Seja $\phi : G \to G$ um endomosfismo sobrejetivo de um grupo de Lie redutível e conexo G. E seja $\widetilde{\phi}$ o endomorfismo associado. Então,*

$$h\left(\widetilde{\phi}\right) = h\left(\phi|_{T(G)}\right).$$

Demonstração. Usando 3.1.10, temos que

$$h\left(\widetilde{\phi}\right) = h\left(\phi|_{Z(G)_0}\right) + h\left(\phi|_{G'}\right).$$

O resultado é então consequência dos casos abeliano e semi-simples, pois sabemos que $h\left(\phi|_{Z(G)_0}\right) = h\left(\phi|_{T(G)}\right)$ e que $h\left(\phi|_{G'}\right) = 0$. $\square$

Corolário 5.3.9. *Para qualquer endomorfismo sobrejetivo $\phi : G \to G$ de um grupo de Lie redutível e simplesmente conexo G,*

$$h(\phi) = h\left(\phi|_{T(G)}\right).$$

Demonstração. Como G é um recobrimento universal, e como as álgebras de Lie de G e $Z(G)_0 \times G'$ coincidem, o homomorfismo π é uma conjugação entre $\widetilde{\phi}$ e ϕ. □

Agora, vamos considerar o caso em que G não é homeomorfo a $Z(G)_0 \times G'$.

Proposição 5.3.10. *Seja $\phi : G \to G$ um endomorfismo sobrejetivo de um grupo de Lie redutível e conexo G. Se a projeção $\pi : Z(G)_0 \times G' \to G$ for própria, então*

$$h(\phi) = h\left(\phi|_{T(G)}\right).$$

Demonstração. Considere o endomorfismo $\tilde{\phi}$ do Lema 5.3.6. Então, use as Proposições 1.2.19 e 5.3.8 para concluir que

$$h(\phi) \leq h\left(\tilde{\phi}\right) = h\left(\phi|_{T(G)}\right) \leq h(\phi).$$

□

Uma consequência imediata, é a seguinte solução do caso redutível linear.

Corolário 5.3.11. *Seja $\phi : G \to G$ um endomorfismo sobrejetivo de um grupo de Lie linear redutível e conexo G. Então,*

$$h(\phi) = h\left(\phi|_{T(G)}\right).$$

Demonstração. Usando a Proposição 5.3.10, precisamos apenas mostrar que $\pi : Z(G)_0 \times G' \to G$ é própria. Mas $\ker(\pi) = \left\{(x, x^{-1}) \;\middle|\; x \in Z(G)_0 \cap G'\right\}$. E, como G é um grupo linear redutível, $Z(G)_0 \cap G'$ está contido no centro de G'. Mas o centro de um grupo de Lie linear semi-simples conexo é sempre finito. Agora, basta usar o Lema 5.2.1. □

Outra consequência imediata é que a fórmula para a entropia de um endomorfismo de um grupo compacto se reduz à fórmula para a entropia de um endomorfismo de um toro.

Corolário 5.3.12. *Seja* $\phi : G \to G$ *um endomorfismo sobrejetivo de um grupo de Lie compacto e conexo* G*. Então,*

$$h(\phi) = h\left(\phi|_{T(G)}\right).$$

Demonstração. Como todo grupo de Lie compacto é redutível (veja a Proposição 6.6 na página 132 de [Hel78]), precisamos apenas mostrar que $\pi : Z(G)_0 \times G' \to G$ é própria. Mas $Z(G)_0$ e G' são subgrupos compactos do grupo compacto G (veja o Teorema 6.9 na página 133 de [Hel78]). Então $Z(G)_0 \times G'$ é compacto e π é própria. □

APÊNDICE A

Apêndice

A.1 Esperança Condicional

Vamos definir *esperança condicional* e enunciar algumas de suas propriedades, que precisaremos no Capítulo 4. Mais informações sobre esperança condicional podem ser encontradas em [Chu01].

Definição A.1.1 (Esperança Condicional)**.** *Seja $\mu : \mathcal{F} \to [0,1]$ uma medida de probabilidade na σ-álgebra $\mathcal{F}$ definida sobre X. E seja $\mathcal{G} \subset \mathcal{F}$ uma sub-σ-álgebra. Dada uma função mensurável $f : X \to \mathbb{R}$ com integral definida, a* esperança condicional *$E\big(f \,\big|\, \mathcal{G}\big)$ de f com relação a $\mathcal{G}$ é a única [μ-qtp] função $\mathcal{G}$-mensurável*

$$E\big(f \mid \mathcal{G}\big) : X \to \mathbb{R}$$

tal que, para todo $A \in \mathcal{G}$,

$$\int_A f \, \mathrm{d}\mu = \int_A E\big(f \mid \mathcal{G}\big) \, \mathrm{d}\mu.$$

A existência da esperança condicional é garantida pelo *Teorema de Radon-Nikodym*. As propriedades que precisamos da esperança condicional são enunciadas a seguir. As igualdades e desigualdades são μ-qtp.

Lema A.1.2. *A esperança condicional satisfaz:*

1. *Linearidade: $E\big(\alpha f + \beta g \mid \mathcal{G}\big) = \alpha E\big(f \mid \mathcal{G}\big) + \beta E\big(g \mid \mathcal{G}\big)$.*

2. *Se $0 \leq f \leq 1$, então,*
$$0 \leq E\big(f \mid \mathcal{G}\big) \leq 1.$$

3. Se $0 \leq f \leq 1$, *então,*

$$\left\|E(f \mid \mathcal{G})\right\|_2 \leq \left(\int f \, \mathrm{d}\mu\right)^{\frac{1}{2}}.$$

Demonstração. Vamos mostrar apenas os itens (2) e (3). denote por h a esperança condicional de f. Seja $A = h^{-1}((1, \infty))$. Como h é $\mathcal{G}$-mensurável, $A \in \mathcal{G}$. Assim, pela definição de esperança condicional, se $\mu(A) \neq 0$,

$$\mu(A) < \int_A h \, \mathrm{d}\mu = \int_A f \, \mathrm{d}\mu \leq \mu(A).$$

Portanto, $\mu(A) = 0$. Da mesma forma, $B = h^{-1}((-\infty, 0))$ também tem medida nula.

Vamos calcular $\|h\|_2$. Pelo parágrafo anterior, como $0 \leq h \leq 1$,

$$0 \leq \|h\|_2{}^2 = \int h^2 \, \mathrm{d}\mu \leq \int h \, \mathrm{d}\mu = \int f \, \mathrm{d}\mu.$$

□

A.2 Teoremas de Recorrência de Poincaré

O primeiro teorema é uma propriedade dos sitemas dinâmicos onde está definida uma medida de probabilidade invariante. Nenhuma hipótese é feita sobre a topologia do sistema. Em seguida, para o caso de um sistema dinâmico, essencialmente mostramos que para efeitos de cálculo da entropia de Kolmogorov-Sinai, sempre podemos nos restringir ao fecho do conjunto dos pontos recorrentes, independentemente da medida de probabilidade considerada.

Teorema A.2.1 (Teorema de Recorrência de Poincaré)**.** *Seja* μ *uma medida finita sobre* X, *e* $T : X \to X$ *uma aplicação mensurável* μ*-invariante. Então, para qualquer conjunto mensurável* $E \subset X$,

$$\mu\left(E \setminus \limsup_{n \to \infty} T^{-n}(E)\right) = 0.$$

Ou seja, a órbita de quase todo ponto de E *passa por* E *infinitas vezes.*

Demonstração. Considere os conjuntos

$$E_n = \bigcup_{j=n}^{\infty} T^- j(E).$$

E considere o conjunto $E_\infty = \lim E_n = \limsup_{n\to\infty} T^{-n}(E)$. Pela T-invariância de μ, $\mu(E_n) = \mu(E_0)$ para todo $n \in \mathbb{N}$. Como μ é finita, $E_n \downarrow E_\infty$ implica que $\mu(E_n) \downarrow \mu(E_\infty)$. Ou seja, $\mu(E_\infty) = \mu(E_0)$. Assim, $\mu(E_0 \setminus E_\infty) = 0$. Mas como $E \subset E_0$, isso implica que

$$\mu(E \setminus E_\infty) = 0.$$

□

Corolário A.2.2 (Recorrência de Poincaré: versão topológica)**.** *Seja $T : X \to X$ um sistema dinâmico separável, e μ uma medida T-invariante nos borelianos de X. Então, o conjunto $\mathcal{R}(T)$ dos pontos recorrentes de T tem medida total. Ou seja,*

$$\mu(\mathcal{R}(T)) = 1.$$

Demonstração. Seja $A_1, A_2, \dots$ uma base enumerável para a topologia de X. Pelo Teorema A.2.1, o conjunto B_n dos pontos de A_n que não retornam a A_n infinitas vezes tem medida nula. Agora, basta mostrar que

$$\mathcal{R}(T)^c \subset \bigcup_{n=1}^{\infty} B_n.$$

De fato, se $x \notin \mathcal{R}(T)$, então existe n tal que $x \in A_n$, mas a órbita de x não passa infinitas vezes por A_n. Ou seja, existe n tal que $x \in B_n$. □

A.3 Entropia Condicional

A demonstração do princípio variacional que apresentamos neste trabalho depende de podermos aproximar uma partição mensurável finita com n elementos por uma outra partição formada por n compactos e 1 aberto, de modo que diferença entre a entropia dessas duas partições seja tão pequena quanto necessário. Esta construção é o objeto do Teorema A.3.6. O conceito que utilizamos para demonstrá-lo é o de *entropia condicional*.

Se μ é uma medida de probabilidade, e $C \subset X$ é um conjunto mensurável de medida não nula, então a *probabilidade condicional* dado o evento C é a medida de probabilidade dada por

$$\mu(\cdot \mid C) = \frac{\mu(\cdot \cap C)}{\mu(C)}.$$

Dadas duas partições mensuráveis finitas $\mathcal{C}$ e $\mathcal{D}$, a *entropia condicional de $\mathcal{D}$ com relação a $\mathcal{C}$*, denotada por $H_\mu\big(\mathcal{D} \,\big|\, \mathcal{C}\big)$, é a média das entropias $H_{\mu(\cdot\,|\,C)}(\mathcal{D})$. Ou seja,

$$H_\mu\big(\mathcal{D} \,\big|\, \mathcal{C}\big) = \sum_{C \in \mathcal{C}} \mu(C) H_{\mu(\cdot\,|\,C)}(\mathcal{D}).$$

Vamos enunciar alguns lemas básicos relacionados com a função $x \log \frac{1}{x}$, parte essencial da definição de entropia de Kolmogorov-Sinai.

Lema A.3.1. *A aplicação*

$$\begin{array}{rccc} f: & [0,\infty) & \to & \mathbb{R} \\ & x & \mapsto & x \log \frac{1}{x} \end{array}$$

é concava. Ou seja, se $\alpha_1, \ldots, \alpha_n \in [0,1]$ e $x_1, \ldots, x_n \in [0,\infty)$, com $\sum_{j=1}^n \alpha_j = 1$, então,

$$\sum_{j=1}^n \alpha_j f(x_j) \le f\left(\sum_{j=1}^n \alpha_j x_j\right).$$

Demonstração. É consequência direta do fato de $f'(x) = -1 - \log x$ ser decrescente. Para maiores detalhes, veja o Teorema 4.2 de [Wal00]. $\square$

Lema A.3.2. *Seja μ uma medida de probabilidade e $\mathcal{C}$ uma partição mensurável finita. Então,*

$$H_\mu(\mathcal{C}) \le \log \#\mathcal{C}.$$

Demonstração. Segue do Lema A.3.1, pois

$$\begin{aligned} H_\mu(\mathcal{C}) &= \sum_{C \in \mathcal{C}} f(\mu(C)) \\ &= \#\mathcal{C} \sum_{C \in \mathcal{C}} \frac{1}{\#\mathcal{C}} f(\mu(C)) \\ &\le \#\mathcal{C} f\left(\sum_{C \in \mathcal{C}} \frac{1}{\#\mathcal{C}} \mu(C)\right) \\ &= \#\mathcal{C} f\left(\frac{1}{\#\mathcal{C}}\right) \\ &= \#\mathcal{C} \frac{1}{\#\mathcal{C}} \log \#\mathcal{C} \\ &= \log \#\mathcal{C}. \end{aligned}$$

$\square$

Lema A.3.3. *Sejam $\nu_1, \ldots, \nu_n$ medidas de probabilidade nos Borelianos de X. E seja*

$$\mu = \sum_{j=1}^{n} \alpha_j \nu_j$$

uma combinação convexa das medidas $\nu_1, \ldots, \nu_n$. Então, para qualquer partição mensurável finita $\mathcal{C}$,

$$H_\mu(\mathcal{C}) \geq \sum_{j=1}^{n} \alpha_j H_{\nu_j}(\mathcal{C}).$$

Demonstração. Basta mostrar que para qualquer $C \in \mathcal{C}$,

$$\mu(C) \log \frac{1}{\mu(C)} \geq \sum_{j=1}^{n} \alpha_j \nu_j(C) \log \frac{1}{\nu_j(C)}.$$

Mas isso é consequência da concavidade da função $f(x) = x \log \frac{1}{x}$, pois

$$\begin{aligned}
\mu(C) \log \frac{1}{\mu(C)} &= f(\mu(C)) \\
&= f\left(\sum_{j=1}^{n} \alpha_j \nu_j(C)\right) \\
&\geq \sum_{j=1}^{n} \alpha_j f(\nu_j(C)) \\
&= \sum_{j=1}^{n} \alpha_j \nu_j(C) \log \frac{1}{\nu_j(C)}.
\end{aligned}$$

Veja o Teorema 4.2 de [Wal00]. $\square$

Observação A.3.4. Note que dadas duas partições mensuráveis finitas $\mathcal{C}$ e $\mathcal{D}$, com $\mathcal{C} \prec \mathcal{D}$, então, para cada $C \in \mathcal{C}$, os elementos $D_1, \ldots, D_n \in \mathcal{D}$ contidos em C, são tais que $\mu(\cdot \mid C)$ é dada pela combinação convexa

$$\mu(\cdot \mid C) = \sum_{j=1}^{n} \mu(D_j \mid C) \mu(\cdot \mid D_j).$$

Lema A.3.5. *Em um sistema dinâmico mensurável $T : X \to X$, com medida de probabilidade μ, valem as seguintes afirmações.*

1. *Para μ T-invariante, $H_\mu(T^{-1}(\mathcal{C})) = H_\mu(\mathcal{C})$.*

2. *$H_\mu\big(\mathcal{D} \,\big|\, \mathcal{C}\big) = H_\mu(\mathcal{C} \vee \mathcal{D}) - H_\mu(\mathcal{C})$.*

3. *Se $\mathcal{D} \prec \mathcal{E}$, então $H_\mu\big(\mathcal{D} \,\big|\, \mathcal{C}\big) \leq H_\mu\big(\mathcal{E} \,\big|\, \mathcal{C}\big)$.*

4. *Se $\mathcal{C} \prec \mathcal{E}$, então $H_\mu\big(\mathcal{D} \,\big|\, \mathcal{E}\big) \leq H_\mu\big(\mathcal{D} \,\big|\, \mathcal{C}\big)$.*
   Em particular, $H_\mu\big(\mathcal{D} \,\big|\, \mathcal{E}\big) \leq H_\mu(\mathcal{D})$.

5. *$H_\mu(\mathcal{C} \vee \mathcal{D}) \leq H_\mu(\mathcal{C}) + H_\mu(\mathcal{D})$.*

6. *$h_\mu\big(T \,\big|\, \mathcal{C}\big) \leq H_\mu(\mathcal{C})$.*

7. *Para μ T-invariante, $h_\mu\big(T \,\big|\, \mathcal{C} \vee \mathcal{D}\big) \leq h_\mu\big(T \,\big|\, \mathcal{C}\big) + h_\mu\big(T \,\big|\, \mathcal{D}\big)$.*

8. *Para μ T-invariante, $h_\mu\big(T \,\big|\, \mathcal{C} \vee \mathcal{D}\big) \leq h_\mu\big(T \,\big|\, \mathcal{C}\big) + H_\mu\big(\mathcal{D} \,\big|\, \mathcal{C}\big)$.*

Demonstração. O item (1) é imediato da invariância μ por T. O item (2) segue por manipulação algébrica:

$$\begin{aligned}
H_\mu(\mathcal{C} \vee \mathcal{D}) - H_\mu(\mathcal{C}) &= \sum_{C \in \mathcal{C}, D \in \mathcal{D}} \mu(C \cap D) \left(\log \frac{1}{\mu(C \cap D)} - \log \frac{1}{\mu(C)} \right) \\
&= \sum_{C \in \mathcal{C}, D \in \mathcal{D}} \mu(C \cap D) \log \frac{\mu(C)}{\mu(C \cap D)} \\
&= \sum_{C \in \mathcal{C}} \sum_{D \in \mathcal{D}} \mu(C) \mu\big(D \,\big|\, C\big) \log \frac{1}{\mu\big(D \,\big|\, C\big)} \\
&= \sum_{C \in \mathcal{C}} \mu(C) \sum_{D \in \mathcal{D}} \mu\big(D \,\big|\, C\big) \log \frac{1}{\mu\big(D \,\big|\, C\big)} = H_\mu\big(\mathcal{D} \,\big|\, \mathcal{C}\big).
\end{aligned}$$

Para o item (3), o resultado segue se notarmos que cada elemento $D \in \mathcal{D}$ é a união disjunta de $E_1, \ldots, E_k \in \mathcal{E}$. Note que para cada E_i,

$$\log \frac{1}{\mu(D)} \leq \log \frac{1}{\mu(E_i)}.$$

Em particular,

$$\log \frac{1}{\mu(D)} = \sum_{j=1}^{k} \frac{\mu(E_i)}{\mu(D)} \log \frac{1}{\mu(D)} \leq \sum_{j=1}^{k} \frac{\mu(E_i)}{\mu(D)} \log \frac{1}{\mu(E_i)}.$$

Ou seja,

$$\mu(D) \log \frac{1}{\mu(D)} \leq \sum_{j=1}^{k} \mu(E_i) \log \frac{1}{\mu(E_i)}.$$

Agora é só somar para todos os elementos $D \in \mathcal{D}$.

Para o item (4), note que pela Observação A.3.4, para cada $C \in \mathcal{C}$ os elementos $E_1, \dots, E_n \in \mathcal{E}$ contidos em C, são tais que

$$\mu\big(\cdot \mid C\big) = \sum_{j=1}^{n} \mu\big(E_j \mid C\big)\mu\big(\cdot \mid E_j\big).$$

Pelo Lema A.3.3,

$$H_{\mu\left(\cdot \mid C\right)}(\mathcal{D}) \geq \sum_{j=1}^{n} \mu\big(E_j \mid C\big) H_{\mu\left(\cdot \mid E_j\right)}(\mathcal{D}).$$

Ou seja,

$$\begin{aligned}
\mu(C) H_{\mu\left(\cdot \mid C\right)}(\mathcal{D}) &\geq \sum_{j=1}^{n} \mu(C)\mu\big(E_j \mid C\big) H_{\mu\left(\cdot \mid E_j\right)}(\mathcal{D}) \\
&= \sum_{j=1}^{n} \mu(E_j) H_{\mu\left(\cdot \mid E_j\right)}(\mathcal{D}).
\end{aligned}$$

Agora, é só somar em $C \in \mathcal{C}$ para obter o resultado desejado. O caso particular segue tomando $\mathcal{C} = \{X\}$, pois

$$H_\mu\big(\mathcal{D} \mid \mathcal{E}\big) \leq H_\mu\Big(\mathcal{D} \,\Big|\, \{X\}\Big) = H_\mu(\mathcal{D}).$$

O item (5) segue dos itens (2) e (4). O item (6) segue de (1) e (5), pois

$$H_\mu(\mathcal{C}^n) \leq \sum_{j=0}^{n-1} H_\mu\left(T^{-j}(\mathcal{C})\right) = n H_\mu(\mathcal{C}).$$

Agora, é só dividir por n e tomar o limite.

O item (7) segue do item (5), se notarmos que

$$(\mathcal{C} \vee \mathcal{D})^n = \mathcal{C}^n \vee \mathcal{D}^n.$$

O item (8) é imediato dos itens (6) e (7), pois

$$\begin{aligned}
h_\mu\big(T \mid \mathcal{C} \vee \mathcal{D}\big) &\leq h_\mu\big(T \mid \mathcal{C}\big) + h_\mu\big(T \mid \mathcal{D}\big) \\
&\leq h_\mu\big(T \mid \mathcal{C}\big) + H_\mu(\mathcal{D}).
\end{aligned}$$

□

A entropia condicional nos permite estimar o *"erro"* quando *"aproximamos"* uma partição $\mathcal{C}$ por uma outra $\mathcal{D}$. Utilizaremos esta aproximação na demonstração do *princípio variacional* (Teorema 3.1.2). Dado $\delta > 0$, diremos que $\mathcal{D}$ é um δ-refinamento de $\mathcal{C}$, quando $N(\mathcal{D}) \leq N(\mathcal{C}) + 1$, e para cada $C \in \mathcal{C}$, existir $D_C \in \mathcal{D}$ tal que

$$\mu(C \cap D_C) \geq (1 - \delta)\mu(C).$$

Quando $\mathcal{D}$ δ-refina $\mathcal{C}$ e vice-versa, dizemos que $\mathcal{D}$ é uma δ-aproximação de $\mathcal{C}$.

Proposição A.3.6. *Seja $T : X \to X$ um sistema dinâmico mensurável, com medida de probabilidade T-invariante μ. Dados $\varepsilon > 0$ e $N \in \mathbb{N}$, então, existe $\delta > 0$ tal que para toda cobertura mensurável $\mathcal{C}$ com $N(\mathcal{C}) < N$ e todo δ-refinamento $\mathcal{D}$ de $\mathcal{C}$,*

$$h_\mu\big(T \,\big|\, \mathcal{C}\big) \leq h_\mu\big(T \,\big|\, \mathcal{D}\big) + \varepsilon.$$

Em particular, se $\mathcal{D}$ for uma δ-aproximação de $\mathcal{C}$, então,

$$\big|h_\mu\big(T \,\big|\, \mathcal{C}\big) - h_\mu\big(T \,\big|\, \mathcal{D}\big)\big| \leq \varepsilon.$$

Demonstração. Escolha $\delta > 0$ tal que para todo x δ-próximo de 0 ou 1, $x \log \frac{1}{x} \leq \frac{\varepsilon}{N^2}$. Note que, para $\mu(C) > 0$, temos que $\mu\big(D_C \,\big|\, C\big) \geq 1 - \delta$, e que, para $D \neq D_C$, temos que $\mu\big(D \,\big|\, C\big) \leq \delta$. Ou seja, pela escolha de δ, para todo $C \in \mathcal{C}$ e todo $D \in \mathcal{D}$,

$$\mu\big(D \,\big|\, C\big) \log \frac{1}{\mu\big(D \,\big|\, C\big)} \leq \frac{\varepsilon}{N^2}.$$

Seja $M = \max_{x \in [0,1]}(x \log x)$. Então,

$$\begin{aligned}
\mu(C) H_{\mu(\cdot \,|\, C)}(\mathcal{D}) &\leq H_{\mu(\cdot \,|\, C)}(\mathcal{D}) \\
&= \sum_{D \in \mathcal{D}} \mu\big(D \,\big|\, C\big) \log \frac{1}{\mu\big(D \,\big|\, C\big)} \\
&\leq N(\mathcal{D}) \frac{\varepsilon}{N^2} \\
&\leq N \frac{\varepsilon}{N^2} \\
&= \frac{\varepsilon}{N}.
\end{aligned}$$

Portanto,

$$H_\mu\big(\mathcal{D} \,\big|\, \mathcal{C}\big) \leq \varepsilon.$$

Agora, basta usar os itens (3) e (8) do Lema A.3.5, pois

$$\begin{aligned} h_\mu(T \,|\, \mathcal{C}) &\leq h_\mu(T \,|\, \mathcal{C} \vee \mathcal{D}) \\ &\leq h_\mu(T \,|\, \mathcal{D}) + H_\mu(\mathcal{D} \,|\, \mathcal{C}) \\ &\leq h_\mu(T \,|\, \mathcal{D}) + \varepsilon. \end{aligned}$$

Para a última afirmação basta trocar os papéis de $\mathcal{C}$ e $\mathcal{D}$. □

A.4 Teorema de Representação de Riesz

Os espaços de fases dos sistemas dinâmicos do tipo produto são da forma $\prod_j X_j$, onde X_j são métricos localmente compactos separáveis. Consequentemente, o espaço formado pelas primeiras n coordenadas, $X^n = \prod_{j=1}^n X_j$, também é métrico localmente compacto separável. A demonstração do princípio variacional envolve a construção, através de um processo de limite, de uma medida μ^n em cada um desses X^n. Assim sendo, nesta seção, fixamos um espaço de Hausdorff localmente compacto X, para descrevermos e mostrarmos propriedades importantes da topologia fraca-* no conjunto das medidas finitas definidas nos borelianos de X.

Denote por $\mathcal{M}(X)$ a família de todas as medidas finitas definidas nos borelianos de X. Denotamos por $C_b(X)$ a família de todas as funções contínuas em X que são limitadas. Com a norma do supremo $\|\cdot\|_\infty$, $C_b(X)$ é um espaço de Banach. A família $C_c(X) \subset C_b(X)$ é a sub-família de $C_b(X)$ das funções contínuas com suporte compacto. O fecho de $C_c(X)$ em $C_b(X)$ na topologia da norma $\|\cdot\|_\infty$ é $C_0(X) = \overline{C_c(X)}$.

Note que para cada $\mu \in \mathcal{M}(X)$,

$$\begin{array}{rccc} \Lambda_\mu : & C_0(X) & \to & \mathbb{R} \\ & f & \mapsto & \int f \, \mathrm{d}\mu \end{array}$$

é um funcional linear contínuo na topologia da norma $\|\cdot\|_\infty$, ou seja, $\Lambda_\mu \in C_0(X)^*$. Note também, que se $\mu \neq \nu$, então $\Lambda_\mu \neq \Lambda_\nu$. Ou seja, $\mathcal{M}(X)$ pode ser identificado com subconjunto de $C_0(X)^*$. Em particular, podemos induzir em $\mathcal{M}(X)$ a topologia fraca-* de $C_0(X)^*$. Ou seja, $\mu_n \to \mu$ quando para todo $f \in C_0(X)$, $\Lambda_{\mu_n}(f) \to \Lambda_\mu(f)$. Explicitamente, quando

$$\int f \, \mathrm{d}\mu_n \to \int f \, \mathrm{d}\mu$$

para todo $f \in C_0(X)$. Poderíamos ter colocado outras topologias em $\mathcal{M}(X)$. A vantagem da topologia fraca-* de $C_0(X)^*$ é ser fraca o suficiente para que o conjunto $\mathcal{B}(X) = \left\{\mu \in \mathcal{M}(X) \,\middle|\, \mu(X) \leq 1\right\}$ seja compacto (Proposição A.4.1), e forte o suficiente para que a convergência tenha implicações significativas (Lema A.4.2).

O primeiro passo é determinar quais são os elementos de $C_0(X)^*$ que são da forma Λ_μ. Este é o conteúdo do Teorema de Representação de Riesz. A seguir, mostramos que o conjunto $\mathcal{B}(X)$ é compacto na topologia fraca-*.

Proposição A.4.1. *Seja X um espaço de Hausdorff localmente compacto, então*

$$\mathcal{B}(X) = \left\{\mu \in \mathcal{M}(X) \,\middle|\, \mu(X) \leq 1\right\}$$

é compacto na topologia fraca-.*

Demonstração. Pelo Teorema de Representação de Riesz (Teorema 6.3.4 de [Ped89]), $\mathcal{M}(X)$ está em bijeção com o conjunto dos funcionais lineares positivos de $C_0(X)^*$, ou seja, os elementos $\phi \in C_0(X)^*$ tais que para $f \geq 0$, $\phi(f) \geq 0$. Pelo mesmo teorema, $\mu(X) = \|\Lambda_\mu\|$.

Na topologia fraca-*, o conjunto $\mathcal{M}(X)$ é fechado em $C_0(X)^*$. De fato, se $\Lambda_{\mu_n}(f) \to \phi(f)$, para toda função $f \in C_0(X)$, então para $f \geq 0$, $\phi(f) = \lim \int f \,\mathrm{d}\mu_n \geq 0$, e pelo Teorema de Representação de Riesz, existe $\mu \in \mathcal{M}(X)$, tal que $\phi = \Lambda_\mu$.

Seja

$$\mathcal{B} = \left\{\phi \in C_0(X)^* \,\middle|\, \|\phi\| \leq 1\right\}.$$

Pelo teorema de Alaoglu (Teorema 2.5.2 de [Ped89]), $\mathcal{B}$ é compacto. Portanto,

$$\mathcal{B}(X) = \mathcal{M}(X) \cap \mathcal{B}$$

é compacto. □

A Proposição A.4.1 garante que a topologia fraca-* é fraca o suficiente para que seja fácil encontrar (sub-)sequências convergentes. Resta mostrar que tal topologia é forte o suficiente para garantir que propriedades interessantes sejam preservadas pelo processo de limite. É importante notar que $\mu_n \to \mu$ não implica que $\mu_n(C) \to \mu(C)$ para um conjunto mensurável qualquer.

Lema A.4.2. *Seja uma sequência de medidas de Radon μ_n que converge para μ na topologia fraca-*. Então, para um conjunto mensurável C, se $\mu(\partial C) = 0$, teremos que $\mu_n(C) \to \mu(C)$.*

Demonstração. Veja o "remark" 3 na seção 6.1 de [Wal00]. □

Se μ_n é uma sequência de medidas com uma subsequência μ_{n_k} que converge para μ, e se $\mu(\partial C) = 0$,

$$\liminf_n \mu_n(C) \leq \mu(C).$$

Para garantir a existência de uma medida em X com as propriedades necessárias à nossa demonstração do princípio variacional (T-invariância e satisfazendo o Lema A.4.2 em X^n), utilizamos o Teorema de Consistência de Kolmogorov. A estratégia consiste em, dada uma sequência μ_k de medidas de probabilidade em X, obter subsequências convergentes a uma medida μ^n, quando μ_k é vista como uma medida de probabilidade no espaço $X^n = \prod_{j=1}^n X_j$. Vamos começar com uma proposição que garante que para uma medida de probabilidade em um espaço métrico separável localmente compacto, a medida de um boreliano E pode ser aproximada pela medida de um compacto $K \subset E$.

Proposição A.4.3. *Seja X um espaço topológico de Hausdorff, localmente compacto e com base enumerável, e seja μ uma medida de probabilidade nos borelianos de X. Então, para todo boreliano $E \subset X$,*

$$\mu(E) = \sup_{\substack{K \subset E \\ \textit{compacto}}} \mu(K).$$

Demonstração. Seja

$$\mathcal{F} = \left\{ E\text{: boreliano} \;\middle|\; \mu(E) = \sup_{\substack{K \subset E \\ \text{compacto}}} \mu(K),\ \mu(E) = \inf_{\substack{A \supset E \\ \text{aberto}}} \mu(A) \right\}$$

a família de todos os borelianos que podem ser aproximados internamente por conjuntos compactos e externamente por abertos. É evidente que $\mathcal{F}$ contém todos os compactos. Vamos mostrar que $\mathcal{F}$ é uma σ-álgebra que contém todos os abertos, e portanto, provando que $\mathcal{F}$ é formado por todos os borelianos.

É fácil ver que se $E_1, E_2 \in \mathcal{F}$, então $E_1 \cup E_2 \in \mathcal{F}$. Suponha que $E_1, E_2, \ldots \in \mathcal{F}$ seja uma sequência crescente. Então, $E = \bigcup E_j \in \mathcal{F}$. De fato, para todo $\varepsilon > 0$, existem compactos $K_n \subset E_n$ com $K_n \subset K_{n+1}$, e abertos $A_n \supset E_n$ com $A_n \subset A_{n+1}$, tais que

$$\mu(A_n) - \varepsilon \leq \mu(E_n) \leq \mu(\lim K_n) + \frac{\varepsilon}{2} = \lim \mu(K_n) + \frac{\varepsilon}{2}.$$

Assim, fazendo $A = \bigcup A_n$ e tomando N suficientemente grande,

$$\mu(A) - \varepsilon \leq \mu(E) \leq \mu(K_N) + \varepsilon.$$

Ou seja, $E \in \mathcal{F}$. Como X é localmente compacto com base enumerável, todo aberto é união enumerável de compactos, e portanto, todo aberto está em $\mathcal{F}$. Da mesma forma, se $F \subset X$ é fechado, então tomando $K_1, K_2, \ldots$ compactos com $X = \bigcup_j K_j$, temos que $F = \bigcup_j F \cap K_j$ é uma união enumerável de compactos. Portanto, todo fechado está em $\mathcal{F}$.

Para concluir que $\mathcal{F}$ é composto por todos os borelianos falta apenas mostrar que $\mathcal{F}$ é uma família fechada por complementação. Para tanto, seja $E \in \mathcal{F}$ e $\varepsilon > 0$. Então, existe um fechado (compacto) $F \subset E$ e um aberto $A \supset E$ tais que

$$\mu(A) - \frac{\varepsilon}{2} \leq \mu(E) \leq \mu(F) + \varepsilon.$$

Ou seja,

$$\mu(F^c) - \varepsilon \leq \mu(E^c) \leq \mu(A^c) + \frac{\varepsilon}{2}.$$

Como $A^c \in \mathcal{F}$, existe um compacto $K \subset A^c \subset E^c$ tal que

$$\mu(A^c) \leq \mu(K) + \frac{\varepsilon}{2}$$

Assim,

$$\mu\left(F^c\right) - \varepsilon \leq \mu\left(E^c\right) \leq \mu\left(A^c\right) + \frac{\varepsilon}{2} \leq \mu(K) + \varepsilon.$$

Ou seja, $E^c \in \mathcal{F}$. □

Vamos apresentar o Teorema de Consistência de Kolmogorov em um formato fácil de ser aplicado na demonstração do princípio variacional.

Teorema A.4.4. *Sejam $X_1, X_2, \ldots$ espaços topológicos localmente compactos com base enumerável. Denote por X^n o produto finito $X_1 \times \cdots \times X_n$, e por X o produto enumerável $X_1 \times X_2 \times \cdots$.*

$$\begin{array}{rccc} \pi_n : & X & \to & X^n \\ & (x_j)_{j=1}^{\infty} & \mapsto & (x_j)_{j=1}^{n} \end{array}.$$

Suponha que μ_n sejam medidas de probabilidade nos borelianos de X^n, satisfazendo a condição de consistência

$$\pi_n^{-1}(E) = \pi_m^{-1}(F) \Rightarrow \mu_n(E) = \mu_m(F).$$

Então existe uma única medida de probabilidade μ nos borelianos de X tal que $\mu_n = \mu \circ \pi_n^{-1}$.

Demonstração. Note que pela condição de consistência, μ está bem definida em toda a álgebra composta pelos conjuntos da forma $\pi_n^{-1}(E)$, onde E é um boreliano de X^n. Considere a família

$$\mathcal{C} = \left\{ \pi_n^{-1}(K) \;\middle|\; n = 1, 2, \dots \; K \subset X^n \text{ é compacto} \right\}.$$

Suponha que a família $\mathcal{C}$ seja uma *classe compacta* (Definição 2.4.2 de [Fer96]). Então, a Proposição A.4.3 garante que as hipóteses do Teorema 2.4.1 de [Fer96] são satisfeitas, e que portanto, μ é σ-aditiva. Basta então aplicar o Teorema de Extensão de Carathéodory (Teorema 2.2.1 de [Fer96]) para garantir que μ pode ser estendida a toda a σ-álgebra pelos conjuntos da forma $\pi_n^{-1}(E)$.

Para mostrar que $\mathcal{C}$ é uma classe compacta, precisamos mostrar que para uma sequência de compactos não vazios $K_j \subset X^{n_j}$ tais que $\pi_j^{-1}(K_j) \supset \pi_{j+1}^{-1}(K_{j+1})$, então $\bigcap_{j=1}^{\infty} \pi_j^{-1}(K_j) \neq \emptyset$. Vamos encontrar $x \in \bigcap_{j=1}^{\infty} \pi_j^{-1}(K_j)$.

Suponha que a sequência finita $x_1, \dots, x_n$ (possivelmente com $n = 0$) tenha sido escolhida de modo que para qualquer m, $\pi_n^{-1}(x_1, \dots, x_n) \cap \pi_m^{-1}(K_m) \neq \emptyset$. Denote por

$$K_{m,n} = \left\{ (y_1, \dots, y_m) \in K_m \;\middle|\; y_1 = x_1, \dots, y_{\max(n,m)} = x_{\max(n,m)} \right\}$$

o subconjunto de K_m composto pelas sequências que começam com $x_1, \dots, x_n$. Os conjuntos $K_{m,n}$ são compactos com $\pi_m^{-1}(K_{m,n})$ decrescentes. Então, $\pi_{n+1}\left(\pi_m^{-1}(K_{m,n})\right)$ é compacto para $m > n$, pois é a projeção em X^{n+1} das primeiras coordenadas do compacto K_m. Agora, escolha x_{n+1} tal que

$$(x_1, \dots, x_n, x_{n+1}) \in \bigcap_{j=1}^{\infty} \pi_{n+1}\left(\pi_j^{-1}(K_{m,n})\right) \neq \emptyset.$$

Para ver que $x = (x_1, x_2, \dots)$ está na interseção dos conjuntos $\pi_m^{-1}(K_m)$, note que para qualquer m, $\pi_m(x) = \pi_m(a^{m+1}) \in K_{m,m} \subset K_m$. Ou seja, $\mathcal{C}$ é uma classe compacta.

A unicidade é consequência do fato de os conjuntos da forma $\pi_n^{-1}(C)$, para $C \subset X^n$ mensurável formam uma álgebra geradora. $\square$

Por fim, construída a medida μ, o que garante sua T-invariância é o item 2 da própria definição de sistema dinâmico do tipo produto (Definição 2.1.1). Veja o item 2 do Lema 2.2.7.

A.5 Entropia de Bowen em Grupos de Lie

Pelo princípio variacional (Teorema 3.1.2), independentemente da métrica escolhida em um sistema do tipo produto, a entropia de Bowen nos fornece uma cota superior para a entropia topológica. Para transformações lineares em $\mathbb{R}^n$, a entropia de Bowen pode ser calculada para a métrica euclideana.

Proposição A.5.1. *Seja* $S : \mathbb{R}^n \times \mathbb{T}^m \to \mathbb{R}^n \times \mathbb{T}^m$ *um sistema dinâmico induzido por uma aplicação linear*

$$T : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}.$$

E sejam $\lambda_1, \ldots, \lambda_m$ *os autovalores de* T *(possivelmente repetidos). Então,*

$$h_d\,(S) = \sum_{|\lambda_i|>1} \log|\lambda_i|,$$

onde d *é a métrica induzida em* $\mathbb{R}^n \times \mathbb{T}^m$ *pela métrica euclideana em* $\mathbb{R}^{n+m}$*.*

Demonstração. Veja o Corolário 3.2.9 de [Fer07]. □

Referências Bibliográficas

[AKM65] R. Adler, A. Konheim, and M. McAndrew, *Topological entropy*, Transactions of the American Mathematical Society **114** (1965), no. 2, 309–319.

[BGKM02] F. Blanchard, E. Glasner, S. Kolyada, and A. Maass, *On Li-Yorke pairs*, J. Reine Angew. Math **547** (2002), 51–68.

[Bow71] R. Bowen, *Entropy for group endomorphisms and homogeneous spaces*, Trans. Americ. Math. Soc. **153** (1971), 401–414.

[Cal10] A. Caldas, *Dinâmica em órbitas projetivas compactas e a decomposição de Jordan*, Dissertação de Mestrado, Departamento de Matemática — Universidade de Brasília, 2010.

[Chu01] K. Chung, *A course in probability theory*, Academic Press, 2001.

[CP13] A. Caldas and M. Patrão, *Dynamics of endomorphisms of Lie groups*, Discrete and Continuous Dynamical Systems **33** (2013), (a ser publicado).

[Din69] E. Dinaburg, *The relation between topological entropy and metric entropy*, Soviet Math. Dokl. **11** (1969), 13–16.

[Ell85] R. Ellis, *Entropy, large deviations, and statistical mechanics*, A Series of Comprehensive Studies in Mathematics, vol. 271, Springer-Verlag, 1985.

[Fer96] P. Fernandez, *Medida e integração*, Projeto Euclides, Insituto de Matemática Pura e Aplicada, CNPq, 1996.

[Fer07] T. Ferraiol, *Entropia e ações de grupos de Lie*, Dissertação de Mestrado, Imecc — Unicamp, 2007.

[Fol99] G. Folland, *Real analysis: modern techniques and their application*, John Wiley & Sons Inc., 1999.

[FPS10] T. Ferraiol, M. Patrão, and L. Seco, *Jordan decomposition and dynamics on flag manifolds*, Discrete and Continuous Dynamical Systems **26** (2010), 923–947.

[Gla03] E. Glasner, *Ergodic theory via joinings*, American Mathematical Society, 2003.

[Goo71] T. Goodman, *Relating topological entropy to measure entropy*, Bull. London. Math. Soc. **3** (1971), 176–180.

[Hel78] S. Helgason, *Differential geometry, Lie groups, and symmetric spaces*, Academic Press, Inc., 1978.

[HKR95] M. Handel, B. Kitchens, and D. Rudolph, *Metrics and entropy for non-compact spaces*, Israel Journal of Mathematics **91** (1995), 253–271.

[Kna02] A. Knapp, *Lie groups*, Progress in Mathematics, vol. 140, Birkhäuser, 2002.

[Pat07] M. Patrão, *Morse decomposition of semiflows on topological spaces*, Journal of Dynamics and Differential Equations **19** (2007), 181–198.

[Pat10] ______, *Entropy and its variational principle for non-compact metric spaces*, Ergodic Theory and Dynamical Systems **30** (2010), 1529–1542.

[Ped89] G. Pedersen, *Analysis now*, Graduate Texts in Mathematics, vol. 118, Springer, 1989.

[PP84] Ya. Pesin and B. Pitskel, *Topological pressure and the variational principle for noncompact sets*, Functional Analysis and Its Applications **18** (1984), 307–318.

[PSS11] M. Patrão, L. Santos, and L. Seco, *A note on the Jordan decomposition*, Proyecciones Journal of Mathematics **30** (2011), 123–136.

[Wal00] P. Walters, *An introduction to ergodic theory*, Springer-Verlag, Berlin, 2000.

[War83] F. Warner, *Foundations of differentiable manifolds and Lie groups*, Graduate Texts in Mathematics, vol. 94, Springer, 1983.