



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIAS
DEPARTAMENTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA

BRENO CAVALCANTE DE SOUSA

ÁLGEBRA DE PINSKER E A PROPRIEDADE K

FORTALEZA

2024

BRENO CAVALCANTE DE SOUSA

ÁLGEBRA DE PINSKER E A PROPRIEDADE K

Dissertação apresentada ao Programa de Pós-Graduação em Matemática, do Centro de Ciências, da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Matemática. Área de Concentração: Sistemas Dinâmicos.

Orientador: Prof. Dr. Yuri Gomes Lima.

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas
Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

S696Á Sousa, Breno Cavalcante de.

Álgebra de Pinsker e a propriedade K / Breno Cavalcante de Sousa. – 2024.
72 f.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Matemática, Fortaleza, 2024.

Orientação: Prof. Dr. Yuri Gomes Lima.

1. Álgebra de Pinsker. 2. Entropia. 3. K-automorfismo. 4. K-mixing. 5. Partições mensuráveis. I. Título.
CDD 510

BRENO CAVALCANTE DE SOUSA

ÁLGEBRA DE PINSKER E A PROPRIEDADE K

Dissertação apresentada ao Programa de Pós-Graduação em Matemática, do Centro de Ciências, da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Matemática. Área de Concentração: Sistemas Dinâmicos.

Aprovada em: 25/06/24.

BANCA EXAMINADORA

Prof. Dr. Yuri Gomes Lima (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Mauricio José Poletti Merlo
Universidade Federal do Ceará (UFC)

Prof. Dr. Jiagang Yang
Universidade Federal Fluminense (UFF)

Aos meus pais Clara de Assis e Raimunda Garcia, por serem minha inspiração para nunca desistir diante das dificuldades.

AGRADECIMENTOS

Agradeço primeiro a Deus por sempre me dar força e saúde para que eu pudesse chegar até aqui. Agradeço aos meus pais Clara de Assis e Raimundo Garcia e a minha esposa Claudiane Gomes, que estiveram comigo em todo esse percurso, me apoiando e motivando, não importasse a dificuldade, sempre me encorajando a não desistir dos meus objetivos.

Agradeço ao meu orientador Yuri Lima, que me orientou neste trabalho e em toda a minha trajetória no mestrado. Agradeço a todos os meus professores, em especial Yuri Lima e Mauricio Poletti, que estiveram comigo compartilhando conhecimentos para me edificar como mestre pesquisador em matemática.

Agradeço a cada um dos meus colegas Alisson, Jônatas e Ernandes, por várias horas incansáveis de estudos tentando solucionar diversos problemas matemáticos. Esses estudos foram fundamentais para que eu possa estar aqui, concluindo mais esta etapa, na minha caminhada acadêmica.

Agradeço a todos da UFC, que de forma direta ou indireta me ajudaram a concluir esse mestrado. Em especial, agradeço a todos do Departamento da Pós-Graduação em Matemática.

Agradeço aos professores Yuri Lima, Maurício Poletti e Jiagang Yang, por participarem da minha Banca Examinadora.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

“A resolução de um problema matemático contém um processo inconsciente. Mas a situação do inconsciente jamais tem lugar, se você não fizer um esforço consciente prolongado, persistente” (Elon Lages Lima).

RESUMO

Nesse trabalho, estudamos sobre a Álgebra de Pinsker e a Propriedade K em teoria ergódica, explicitando a conexão entre esses dois conceitos. Mais precisamente, o nosso objetivo será estabelecer, por meio da Álgebra de Pinsker, condições necessárias e suficientes para que um automorfismo seja um K-automorfismo. Provaremos ainda que a propriedade K é equivalente à propriedade K-mixing.

Inicialmente, estudaremos sobre a entropia métrica de uma partição de partições finitas e enumeráveis e, posteriormente, passaremos a considerar a entropia métrica envolvendo partições mensuráveis. Em seguida, utilizaremos a noção de entropia métrica para definir e caracterizar a Álgebra de Pinsker. Por fim, passaremos ao estudo da Propriedade K, apresentando sua caracterização por meio da Álgebra de Pinsker, bem como definições equivalentes, e ainda apresentaremos três exemplos clássicos de K-automorfismos.

Palavras-chave: álgebra de Pinsker; entropia; k-automorfismo; k-mixing; partições mensuráveis.

ABSTRACT

In this work, we study the ergodic theoretical notions of Pinsker Algebra and the K Property, stressing the relationship between these two concepts. More specifically, our goal is to establish, via the Pinsker Algebra, necessary and sufficient conditions for an automorphism to be K. We also prove that the K property is equivalent to K-mixing. Firstly, we study the metric entropy of finite and countable partitions and, later, we consider the metric entropy involving measurable partitions. Later, we use the notion of metric entropy to define and characterize the Pinsker algebra. Finally, we study the K property, presenting a characterization in terms of the Pinsker algebra, as well as equivalent definitions, and we also present three classical examples of K-automorphisms.

Keywords: Pinsker algebra; entropia; k-automorphism; k-mixing; measurable partitions.

SUMÁRIO

1	INTRODUÇÃO	9
2	ENTROPIA DE PARTIÇÕES FINITAS E ENUMERÁVEIS .	11
2.1	Entropia de uma partição	11
2.2	Entropia de um sistema dinâmico	25
3	PARTIÇÕES MENSURÁVEIS	36
3.1	Partições mensuráveis e σ -álgebras	36
3.2	Entropia de uma partição mensurável	41
3.3	Fórmula de Abramov-Rokhlin	46
4	O FATOR DE PINSKER	49
4.1	Álgebra de Pinsker	49
4.2	Tail σ -álgebras	52
5	A PROPRIEDADE K	62
5.1	K-automorfismos	62
5.2	Sistemas K-mixing	64
6	EXEMPLOS	66
6.1	Shifts de Bernoulli	66
6.2	Deslocamentos de Markov	68
6.3	Automorfismo linear do toro	71
	REFERÊNCIAS	72

1 INTRODUÇÃO

Em 1959, ao provar que os sistemas 2-shift e 3-shift não são metricamente isomorfos, Kolmogorov introduziu uma classe de sistemas chamada de sistemas “quasi-regulares”. Atualmente, tais sistemas são chamados de K -automorfismos. Posteriormente, em seus estudos em 1960, Pinsker construiu o Fator de Pinsker e introduziu a noção de sistemas com entropia completamente positiva. Além disso, Pinsker provou que um sistema tem a propriedade K se e somente se tem entropia completamente positiva. Esse trabalho tem como objetivo estabelecer uma caracterização da Propriedade K por meio da Álgebra de Pinsker. Mais precisamente, nosso objetivo é de estabelecer, por meio da Álgebra de Pinsker, condições necessárias e suficientes para que um automorfismo seja um K -automorfismo. Além disso, provaremos que a propriedade K é equivalente a propriedade K -mixing.

No segundo capítulo, estudaremos sobre entropia métrica de partições finitas e enumeráveis, aonde seguiremos de perto a aparência original da Teoria de Entropia Métrica apresentada por Andrey Kolmogorov e Yakov Sinai na década de 1950, motivada pela Teoria da Informação. Ao longo do capítulo apresentaremos os conceitos de entropia de uma partição finita/enumerável, esperança condicional, entropia condicional (via esperança condicional) e entropia de um sistema dinâmico. Além de apresentar tais conceitos, apresentaremos algumas propriedades que serão úteis ao longo de nossos estudos.

No Terceiro capítulo, estudaremos sobre partições mensuráveis que, possivelmente, possuem uma quantidade não-enumerável de átomos. Iniciaremos o capítulo estabelecendo algumas relações entre σ -álgebras e partições mensuráveis. Logo após, apresentaremos a definição de Entropia Condicional para partições mensuráveis via Desintegração de Rokhlin (Teorema 3.1.7), que generaliza as definições apresentadas no Capítulo 2. Além disso, apresentaremos alguns resultados que serão fundamentais em nossos estudos do Fator de Pinsker que será feito no Capítulo 4. Terminaremos o capítulo apresentando o Teorema de Abramov-Rokhlin, que relaciona a entropia de um automorfismo e seus fatores.

No quarto capítulo, definiremos a Álgebra de Pinsker e a Tail sigma-álgebra de um automorfismo, que serão fundamentais para a caracterização de um K -automorfismo que será apresentada no Capítulo 5. Ao longo do capítulo, apresentaremos as definições e algumas propriedades da Álgebra de Pinsker e da Tail sigma-álgebra. Além disso,

estabeleceremos algumas relações entre a Álgebra de Pinsker e a Tail σ -álgebra de um automorfismo.

No quinto capítulo, estudaremos automorfismos com a propriedade K, ou seja, estudaremos os K-automorfismos. Inicialmente introduziremos a definição de K-automorfismo. Logo em seguida, com o auxílio da Álgebra de Pinsker, apresentaremos uma caracterização de K-automorfismos e, como consequência, obteremos que todo fator de um K-automorfismo também é um K-automorfismo. Finalizaremos o capítulo vendo que um sistema é um K-automorfismo se, e somente se, é K-mixing.

Finalmente, no sexto capítulo, finalizaremos este trabalho apresentando alguns exemplos de K-automorfismos, aonde veremos que os Shifts de Bernoulli, os deslocamentos de Markov e os automorfismos lineares hiperbólicos do Toro são K-automorfismos.

2 ENTROPIA DE PARTIÇÕES FINITAS E ENUMERÁVEIS

O conceito de entropia pode ser descrito de diversas formas. A grosso modo, a entropia pode ser entendida como uma medida de caoticidade de sistemas. É importante frisar que existem vários tipos de entropia, mas em nosso estudo focaremos apenas no conceito de **entropia métrica**, que foi introduzido pelos matemáticos russos Andrey Kolmogorov e Yakov Sinai na década de 1950. A discussão abaixo se aproxima de sua aparência original, motivada pela teoria da informação. As referências deste capítulo são [3], [1], [9], [6] e [4].

2.1 Entropia de uma partição

Seja $(X, \mathcal{B}, \mu)$ um espaço de probabilidade de Lebesgue¹. No que segue, neste e nos demais capítulos, escrevemos que dois conjuntos mesuráveis A e B são iguais “módulo μ ” ou que $A = B \pmod{\mu}$ sempre que a igualdade for verdadeira a menos de um conjunto de medida nula. Analogamente, dizemos que uma propriedade $\mathbb{P}$ vale para quase todo ponto $x \in X$ se existe um conjunto $E \subseteq X$ com $\mu(X \setminus E) = 0$ tal que a propriedade $\mathbb{P}$ vale para todo ponto de E . Neste caso, dizemos que $\mathbb{P}$ vale em quase todo ponto e escrevemos “ $\mathbb{P}$ vale para μ -q.t.p. $x \in X$ ”, ou simplesmente que “ $\mathbb{P}$ vale μ -q.t.p.”.

Definição 2.1.1 (Partição finita/enumerável) *Uma partição de X é uma coleção α disjunta de conjuntos $A \in \mathcal{B}$ não-vazios que cobre X módulo μ . Cada elemento $A \in \alpha$ é chamado de átomo de α . Chamamos α de partição finita se possuir finitos átomos, e de enumerável se possuir uma quantidade enumerável de átomos.*

Dados $x \in X$ e uma partição enumerável α de X , denote por $\alpha(x)$ o átomo de α que contém x . O átomo $\alpha(x)$ é chamado de α -nome de x .

Embora, algumas das propriedades que apresentaremos neste capítulo possam ser trabalhadas de forma mais geral, estaremos sempre trabalhando no caso de partições finitas ou enumeráveis. Assim, durante este capítulo, as partições α e β sempre serão entendidas como partições finitas ou enumeráveis.

Existe uma relação entre partições e sigma-álgebras. Dada uma partição α , podemos considerar $\mathcal{F} := \hat{\alpha}$ como sendo a menor sigma-álgebra (módulo μ) tal que todo

¹ Aqui, consideraremos um espaço de Lebesgue como sendo um espaço de probabilidade isomorfo ao espaço de Lebesgue canônico do intervalo unitário.

átomo de α é mensurável. Assim, a sigma-álgebra $\mathcal{F}$ é a interseção, de todas as sigma-álgebras $\mathcal{A}$ tais que $\alpha \subseteq \mathcal{A}$, a menos de medida nula. Neste caso, dizemos que $\mathcal{F}$ é a sigma-álgebra gerada por α . Reciprocamente, dada uma sigma-álgebra $\mathcal{F}$, a partição gerada por $\mathcal{F}$ é a partição $\alpha = \mathcal{F}^\wedge$ cujos átomos são

$$\alpha(x) := \bigcap_{\substack{F \in \mathcal{F} \\ x \in F}} F.$$

Observe que cada $\alpha(x) \in \mathcal{F}$. Não é difícil ver que as aplicações $\alpha \mapsto \hat{\alpha}$ e $\mathcal{F} \mapsto \mathcal{F}^\wedge$ são inversas uma da outra, isto é, que $(\hat{\alpha})^\wedge = \alpha$ e $\widehat{\mathcal{F}^\wedge} = \mathcal{F}$.

Observe que se $\mathcal{F} := \hat{\alpha}$, onde α é uma partição enumerável, então $\mathcal{F}$ consiste do conjunto $\emptyset$ e de todos os subconjuntos de X que podem ser obtidos a partir de uma união, módulo μ , de átomos de α . Uma consequência imediata dessa observação é que toda função $\mathcal{F}$ -mensurável $f : X \rightarrow \mathbb{R}$ é constante em cada átomo de α .

Definição 2.1.2 (Refinamento) *Dadas duas partições α, β de X , dizemos que β é um refinamento de α se todo átomo de α é igual à união módulo μ de átomos de β . Neste caso, escrevemos $\alpha \leq \beta$ e dizemos que β é mais fina que α ou, equivalentemente, que α é mais grossa que β .*

Claramente, pelas definições de refinamento e de σ -álgebra gerada por uma partição, $\alpha \leq \beta$ se e somente se $\hat{\alpha} \subseteq \hat{\beta}$. Além disso, também por definição, vale que $\alpha \leq \beta$ se e somente se $\beta(x) \subseteq \alpha(x)$ para μ -q.t.p. $x \in X$.

Observe que se $\alpha \leq \beta$ então todo átomo de β está contido em algum átomo de α . Reciprocamente, se todo átomo de β está contido em algum átomo de α , temos que

$$\alpha(x) = \bigcup_{y \in \alpha(x)} \beta(y),$$

donde concluímos que $\alpha \leq \beta$. Assim, podemos reformular a definição de refinamento: β é mais fina do que α quando todo átomo de β está contido em algum átomo de α , a menos de medida nula.

Definição 2.1.3 (Partição soma) *Dadas duas partições α, β de X , a partição soma de α e β é a partição $\alpha \vee \beta := \{A \cap B : A \in \alpha \text{ e } B \in \beta\}$. Esta é a partição mais grosseira que refina α e β .*

Definição 2.1.4 (Independência) Dizemos que dois conjuntos $A, B \in \mathcal{B}$ são independentes se $\mu(A \cap B) = \mu(A)\mu(B)$. Duas partições α, β são independentes se todo par de átomos $A \in \alpha$ e $B \in \beta$ são independentes.

Definição 2.1.5 (Função informação) A informação de um átomo $A \in \alpha$ é $I_\mu(A) = -\log \mu(A)$. A função informação de α é a função $I_\mu(\alpha) : X \rightarrow [0, +\infty]$ definida por

$$I_\mu(\alpha)(x) = \sum_{A \in \alpha} I_\mu(A) 1_A(x) = \sum_{A \in \alpha} -\log \mu(A) 1_A(x).$$

Intuitivamente, $I_\mu(A)$ mede a incerteza da afirmação “ $x \in A$ ”, ou seja, mede o quanto é preciso “pagar” para saber se $x \in A$ ou não. Existem algumas justificativas para definir I_μ como acima. A principal delas é que I_μ definida assim está unicamente determinada, a menos de seus múltiplos, pelas cinco propriedades abaixo:

1. $I_\mu(A)$ é uma função não-negativa de $\mu(A)$.
2. A aplicação $A \mapsto I_\mu(A)$ é contínua.
3. Se $A \subseteq B$ então $I_\mu(A) \geq I_\mu(B)$.
4. Se A e B são independentes então $I_\mu(A \cap B) = I_\mu(A) + I_\mu(B)$.
5. Se $\mu(A) = 1$ então $I_\mu(A) = 0$.

A terceira propriedade significa que quanto menor o conjunto A , mais devemos pagar para saber se $x \in A$. A quarta propriedade significa que se A e B são independentes então as informações $x \in A$ e $x \in B$ não estão relacionadas. Por fim, a última propriedade significa que $x \in X$ é sempre verdadeira.

Definição 2.1.6 (Entropia de partição) A entropia de uma partição α , denotada por $H_\mu(\alpha)$, é dada por

$$H_\mu(\alpha) := \int_X I_\mu(\alpha)(x) d\mu(x) = \sum_{A \in \alpha} -\mu(A) \log \mu(A).$$

Aqui, como é usual na teoria da integral de Lebesgue, usamos a convenção $0 \log 0 = \lim_{x \rightarrow 0} x \log x = 0$. Dizemos que uma partição α tem entropia finita se $H_\mu(\alpha) < \infty$, e denotamos por $\mathcal{Z}$ o espaço de todas as partições finitas/enumeráveis de entropia finita.

Intuitivamente, a entropia de α é o preço médio que temos que pagar para saber o átomo de α que contém x . Se considerarmos α como o conjunto dos resultados possíveis de um experimento, então a entropia $H_\mu(\alpha)$ também pode ser interpretada como a quantidade de incerteza sobre o resultado do experimento: quanto maior $H_\mu(\alpha)$, menos

podemos prever o resultado. Uma vez realizado o experimento, $H_\mu(\alpha)$ pode ser visto como a incerteza removida ou a informação gerada pelo experimento. Segue-se direto da sua definição que a entropia satisfaz algumas propriedades básicas:

1. $H_\mu(\alpha) \geq 0$.

De fato, como $-\mu(A) \log \mu(A) \geq 0$ para todo $A \in \alpha$, temos que $H_\mu(\alpha) \geq 0$.

2. Se $\alpha \leq \beta$ então $H_\mu(\alpha) \leq H_\mu(\beta)$.

De fato, uma vez que $\alpha \leq \beta$, para cada $A \in \alpha$ vale que $A = \bigcup_{\substack{B \in \beta \\ B \subseteq A}} B$. Assim,

$$\begin{aligned} -\mu(A) \log \mu(A) &= -\mu \left(\bigcup_{\substack{B \in \beta \\ B \subseteq A}} B \right) \log \mu(A) = \sum_{\substack{B \in \beta \\ B \subseteq A}} -\mu(B) \log \mu(A) \\ &\leq \sum_{\substack{B \in \beta \\ B \subseteq A}} -\mu(B) \log \mu(B), \end{aligned}$$

onde na última desigualdade usamos o fato de que a função $-\log(x)$ é decrescente e, portanto, $-\log \mu(A) \leq -\log \mu(B)$ para todo $B \subseteq A$. Como todo átomo de β está contido em um átomo de α , a desigualdade acima implica que $H_\mu(\alpha) \leq H_\mu(\beta)$.

Considere a função contínua $\varphi: [0, 1] \rightarrow \mathbb{R}$ dada por $\varphi(t) = -t \log(t)$ para $t \neq 0$ e $\varphi(0) = 0$. Derivando duas vezes φ , vemos que $\varphi''(t) = -\frac{1}{t} < 0$ para todo $t \in (0, 1)$. Logo φ é côncava, isto é,

$$a_1 \varphi(t_1) + \cdots + a_k \varphi(t_k) \leq \varphi(a_1 t_1 + \cdots + a_k t_k) \quad (2.1)$$

para todos $t_1, \dots, t_k > 0$ e $a_1, \dots, a_k \geq 0$ com $a_1 + \cdots + a_k = 1$. Além disso, a concavidade é estrita: a igualdade em (2.1) ocorre se, e somente se, $t_1 = \cdots = t_k$. Observe que

$$H_\mu(\alpha) = \sum_{A \in \alpha} \varphi(\mu(A)).$$

Uma consequência útil do fato de φ ser côncava é que toda partição finita tem entropia finita. De fato, seja $\alpha = \{A_1, A_2, \dots, A_n\}$ uma partição finita de X e considere os números $a_i = \frac{1}{n}$ e $t_i = \mu(A_i)$ para $i = 1, 2, \dots, n$. Pela propriedade de concavidade (2.1), temos

$$\frac{1}{n} H_\mu(\alpha) = \sum_{i=1}^n a_i \varphi(t_i) \leq \varphi \left(\sum_{i=1}^n a_i t_i \right) = \varphi \left(\frac{1}{n} \right) = \frac{\log n}{n}$$

e portanto $H_\mu(\alpha) \leq \log n$.

Visando discutir *entropia condicional*, iremos apresentar agora o conceito de *esperança condicional* e algumas de suas propriedades. A esperança condicional é definida

por meio de uma aplicação do Teorema de Radon-Nikodym, e é um dos tipos mais importantes de processos de cálculo de médias.

Sejam $f \in L^1(X, \mathcal{B}, \mu)$ e $\mathcal{F} \subseteq \mathcal{B}$ uma sub σ -álgebra. Então a aplicação $\nu_f : \mathcal{F} \rightarrow [0, \infty]$ definida por

$$\nu_f(A) = \int_A f d\mu, \quad A \in \mathcal{F},$$

define uma medida absolutamente contínua com respeito à restrição de μ em $\mathcal{F}$. Pelo Teorema de Radon-Nikodym, existe uma função $g \in L^1(X, \mathcal{F}, \mu)$ tal que

$$\nu_f(A) = \int_A g d\mu, \quad \forall A \in \mathcal{F}.$$

Além disso, g é única para μ -q.t.p. $x \in X$. Usamos a notação $\mathbb{E}(f|\mathcal{F})$ para g , e chamamos $\mathbb{E}(f|\mathcal{F})$ de esperança condicional de f em relação a $\mathcal{F}$.

Como vimos acima, o Teorema de Radon-Nikodym caracteriza bem a esperança condicional. Sua definição formal é a seguinte.

Definição 2.1.7 (Esperança condicional) A esperança condicional de $f \in L^1(X, \mathcal{B}, \mu)$ com respeito à σ -álgebra $\mathcal{F}$ é a única função $\mathbb{E}(f|\mathcal{F})$ tal que:

- (i) $\mathbb{E}(f|\mathcal{F}) \in L^1(X, \mathcal{F}, \mu)$;
- (ii) $\int_X \varphi \mathbb{E}(f|\mathcal{F}) d\mu = \int_X \varphi f d\mu$, para toda função $\mathcal{F}$ -mensurável limitada $\varphi : X \rightarrow \mathbb{R}$.

Observe que, quando $\mathcal{F} = \{\emptyset, X\}$, as únicas funções $\mathcal{F}$ -mensuráveis são as funções constantes. Assim, pela Definição 2.1.7, temos que $\mathbb{E}(1_A|\mathcal{F}) = \mu(A)$ para todo $A \in \mathcal{B}$. Abaixo, veremos mais algumas propriedades da esperança condicional.

Proposição 2.1.8 Sejam $f, g \in L^1(X, \mathcal{B}, \mu)$. Então:

- (a) Dados $a, b \in \mathbb{R}$, temos que $\mathbb{E}(af + bg|\mathcal{F}) = a\mathbb{E}(f|\mathcal{F}) + b\mathbb{E}(g|\mathcal{F})$.
- (b) Se $f \geq 0$ então $\mathbb{E}(f|\mathcal{F}) \geq 0$.
- (c) Se $f \geq g$ então $\mathbb{E}(f|\mathcal{F}) \geq \mathbb{E}(g|\mathcal{F})$.
- (d) Se f é $\mathcal{F}$ -mensurável, então $\mathbb{E}(f|\mathcal{F}) = f$.
- (e) $\int_X \mathbb{E}(f|\mathcal{F}) d\mu = \int_X f d\mu$.
- (f) Se $\mathcal{F}, \mathcal{G} \subseteq \mathcal{B}$ são σ -álgebras tais que $\mathcal{G} \subseteq \mathcal{F}$, então $\mathbb{E}(\mathbb{E}(f|\mathcal{F})|\mathcal{G}) = \mathbb{E}(f|\mathcal{G})$.
- (g) Se $h : X \rightarrow \mathbb{R}$ é $\mathcal{F}$ -mensurável, então $\int_X h \mathbb{E}(f|\mathcal{F}) d\mu = \int_X hf d\mu$.

Demonstração: (a) Por definição, $\mathbb{E}(f|\mathcal{F})$ e $\mathbb{E}(g|\mathcal{F})$ estão em $L^1(X, \mathcal{F}, \mu)$, logo $a\mathbb{E}(f|\mathcal{F}) + b\mathbb{E}(g|\mathcal{F}) \in L^1(X, \mathcal{F}, \mu)$. Além disso, para toda função $\mathcal{F}$ -mensurável limitada $\varphi : X \rightarrow \mathbb{R}$, temos

$$\begin{aligned} \int_X \varphi \cdot (a\mathbb{E}(f|\mathcal{F}) + b\mathbb{E}(g|\mathcal{F})) d\mu &= \int_X a\varphi \cdot \mathbb{E}(f|\mathcal{F}) d\mu + \int_X b\varphi \mathbb{E}(g|\mathcal{F}) d\mu \\ &= \int_X \varphi \cdot (af + bg) d\mu. \end{aligned}$$

Assim, pela caracterização da esperança condicional, temos que $\mathbb{E}(af + bg|\mathcal{F}) = a\mathbb{E}(f|\mathcal{F}) + b\mathbb{E}(g|\mathcal{F})$.

(b) Assuma que $f \geq 0$ em quase todo ponto. Suponha, por absurdo, que não vale que $\mathbb{E}(f|\mathcal{F}) \geq 0$. Então, existe um conjunto $E \in \mathcal{F}$ com $\mu(E) > 0$ tal que $\mathbb{E}(f|\mathcal{F}) < 0$ em E . Assim, vale que

$$0 \leq \int_E f d\mu = \int_X 1_E f d\mu = \int_X 1_E \mathbb{E}(f|\mathcal{F}) d\mu = \int_E \mathbb{E}(f|\mathcal{F}) d\mu < 0,$$

o que é um absurdo. Logo, se $f \geq 0$ μ -q.t.p. então $\mathbb{E}(f|\mathcal{F}) \geq 0$ μ -q.t.p.

(c) De fato, se $f \geq g$ em quase todo ponto, então $f - g \geq 0$ em quase todo ponto. Assim, pelos itens (a) e (b), vemos que $\mathbb{E}(f|\mathcal{F}) - \mathbb{E}(g|\mathcal{F}) = \mathbb{E}(f - g|\mathcal{F}) \geq 0$, donde concluímos que $\mathbb{E}(f|\mathcal{F}) \geq \mathbb{E}(g|\mathcal{F})$.

(d) É imediato da caracterização de esperança condicional.

(e) Com efeito, como 1_X é uma função $\mathcal{F}$ -mensurável limitada, temos

$$\int_X \mathbb{E}(f|\mathcal{F}) d\mu = \int_X 1_X \mathbb{E}(f|\mathcal{F}) d\mu = \int_X 1_X f d\mu = \int_X f d\mu,$$

como desejado.

(f) Seja $\varphi : X \rightarrow \mathbb{R}$ uma função $\mathcal{G}$ -mensurável limitada. Então, como $\mathcal{G} \subseteq \mathcal{F}$, temos que φ é $\mathcal{F}$ -mensurável. Assim, obtemos

$$\int_X \varphi \mathbb{E}(\mathbb{E}(f|\mathcal{F})|\mathcal{G}) d\mu = \int_X \varphi \mathbb{E}(f|\mathcal{F}) d\mu = \int_X \varphi f d\mu,$$

donde concluímos que $\mathbb{E}(\mathbb{E}(f|\mathcal{F})|\mathcal{G}) = \mathbb{E}(f|\mathcal{G})$.

(g) Seja $h : X \rightarrow \mathbb{R}$ uma função $\mathcal{F}$ -mensurável. Como $h = h^+ - h^-$ onde $h^+(x) = \max\{h(x), 0\}$ e $h^-(x) = \max\{-h(x), 0\}$, pela linearidade da integral de Lebesgue, podemos assumir que h é não-negativa. Se h for uma função limitada, pela definição de Esperança condicional, vemos que

$$\int_X h \mathbb{E}(f|\mathcal{F}) d\mu = \int_X h f d\mu.$$

Agora, suponha que h é uma função $\mathcal{F}$ -mensurável qualquer. Tome uma sequência $(h_n)_n$ de funções $\mathcal{F}$ -mensuráveis limitadas tais que $h_n(x) \xrightarrow{n \rightarrow +\infty} h(x)$ para μ -q.t.p. $x \in X$ e $h_n(x) \leq h_{n+1}(x)$ para μ -q.t.p. $x \in X$ e todo $n \geq 1$. Então, usando o teorema da Convergência Monótona duas vezes, temos

$$\int_X h \mathbb{E}(f|\mathcal{F}) d\mu = \lim_{n \rightarrow +\infty} \int_X h_n \mathbb{E}(f|\mathcal{F}) d\mu = \lim_{n \rightarrow +\infty} \int_X h_n f d\mu = \int_X h f d\mu.$$

□

Um caso particular da Definição 2.1.7 é quando $f = 1_A$ é uma função característica de $A \in \mathcal{B}$. Nesse caso, usamos a notação

$$\mu(A|\mathcal{F}) := \mathbb{E}(1_A|\mathcal{F})$$

e chamamos $\mu(A|\mathcal{F})$ de probabilidade condicional de A dada $\mathcal{F}$. Pela Proposição 2.1.8, fixado $A \in \mathcal{B}$, temos que $0 \leq \mu(A|\mathcal{F})(x) \leq \mu(X|\mathcal{F})(x) = 1$ μ -q.t.p. Por conveniência, usaremos a notação $\mu(A|\beta)$ para representar $\mu(A|\hat{\beta})$, onde α e β são partições enumeráveis de X .

A fim de obtermos uma intuição do que $\mathbb{E}(f|\mathcal{F})$ representa, tomemos $A \in \mathcal{B}$ e $\mathcal{F} = \{\emptyset, B, X \setminus B, X\}$ onde B é um conjunto mensurável com medida positiva. Então, pela propriedade (ii) da Definição 2.1.7, devemos ter que

$$\mathbb{E}(1_A|\mathcal{F})(x) = \begin{cases} \frac{\mu(A \cap B)}{\mu(B)} = \mu(A|B) & , \text{ se } x \in B \\ \frac{\mu(A \cap (X \setminus B))}{\mu(X \setminus B)} = \mu(A|X \setminus B) & , \text{ se } x \in X \setminus B. \end{cases}$$

Logo, $\mathbb{E}(1_A|\mathcal{F})$ é constante igual a $\mu(A|B)$ em B e constante igual a $\mu(A|X \setminus B)$ em $X \setminus B$. Assim, vemos que $\mathbb{E}(1_A|\mathcal{F})(x)$ nos fornece a probabilidade da ocorrência de A uma vez que sabemos se x pertence a B ou não. O mesmo vale para o caso geral: $\mathbb{E}(f|\mathcal{F})(x)$ representa o valor esperado de f se, para cada $B \in \mathcal{F}$, soubermos se x está ou não em B .

Pelo que vimos no parágrafo anterior, se $\beta = \{B, X \setminus B\}$ onde $B \in \mathcal{B}$ possui medida positiva, então $\mu(A|\beta) = \mu(A|B) \cdot 1_B + \mu(A|X \setminus B) \cdot 1_{X \setminus B}$. De fato, esse resultado é mais geral, como afirma o lema abaixo.

Lema 2.1.9 *Seja $\beta \in \mathcal{L}$. Então, para qualquer $A \in \mathcal{B}$, temos*

$$\mu(A|\beta) = \sum_{B \in \beta} \mu(A|B) 1_B.$$

Demonstração: Como $\beta \in \mathcal{Z}$ podemos supor, sem perda de generalidade, que todos os átomos de β possuem medida positiva. Seja $A \in \mathcal{B}$ arbitrário. Então, por definição, temos que $\mu(A|\beta)$ é $\hat{\beta}$ -mensurável, logo constante em cada átomo de β . Assim, $\mu(A|\beta) = \sum_{B \in \beta} \gamma_B 1_B$, onde $\gamma_B \in \mathbb{R}$ para todo $B \in \beta$. Resta determinarmos as constantes γ_B . Usando a condição (ii) da Definição 2.1.7, temos

$$\gamma_{B_0} \mu(B_0) = \int_X 1_{B_0} \mu(A|\beta) d\mu = \int_X 1_{B_0} 1_A d\mu = \mu(A \cap B_0),$$

para todo $B_0 \in \beta$. Isso significa que $\gamma_{B_0} = \frac{\mu(A \cap B_0)}{\mu(B_0)} = \mu(A|B)$ para todo $B_0 \in \beta$, o que prova o lema. □

Definição 2.1.10 (Função informação condicional) *Sejam $\alpha \in \mathcal{Z}$ e $\mathcal{F} \subseteq \mathcal{B}$ uma σ -álgebra. A função informação de um conjunto $A \in \mathcal{B}$ dado $\mathcal{F}$ é $I_\mu(A|\mathcal{F})(x) := -\log \mu(A|\mathcal{F})(x)$. A função informação condicional de α dado $\mathcal{F}$ é*

$$I_\mu(\alpha|\mathcal{F})(x) := \sum_{A \in \alpha} I_\mu(A|\mathcal{F})(x) 1_A(x).$$

Observe que a função $I_\mu(A|\mathcal{F})(x) = -\log \mu(A|\mathcal{F})(x)$ é $\mathcal{F}$ -mensurável não-negativa para todo $A \in \alpha$, pois a função $\mu(A|\mathcal{F}) : X \rightarrow [0, 1]$ é uma função $\mathcal{F}$ -mensurável e a função $-\log : [0, 1] \rightarrow [0, \infty]$ é uma função contínua decrescente. Portanto, o mesmo ocorre para $I_\mu(\alpha|\mathcal{F})$.

Definição 2.1.11 (Entropia condicional) *Sejam $\alpha \in \mathcal{Z}$ e $\mathcal{F} \subseteq \mathcal{B}$ uma σ -álgebra. A entropia condicional de α dado $\mathcal{F}$ é*

$$H_\mu(\alpha|\mathcal{F}) := \int_X I_\mu(\alpha|\mathcal{F})(x) d\mu(x).$$

Por conveniência, dados $\alpha, \beta \in \mathcal{Z}$, denotaremos $I_\mu(\alpha|\hat{\beta})$ e $H_\mu(\alpha|\hat{\beta})$ por $I_\mu(\alpha|\beta)$ e $H_\mu(\alpha|\beta)$ respectivamente. O lema baixo apresenta duas definições alternativas para $H_\mu(\alpha|\mathcal{F})$.

Lema 2.1.12 *Sejam $\alpha, \beta \in \mathcal{Z}$ e $\mathcal{F} \in \mathcal{B}$ uma σ -álgebra. Então:*

- (a) $H_\mu(\alpha|\mathcal{F}) = \int_X \sum_{A \in \alpha} -\mu(A|\mathcal{F}) \log \mu(A|\mathcal{F}) d\mu.$
- (b) $H_\mu(\alpha|\beta) = \sum_{A \in \alpha} \sum_{B \in \beta} -\mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(B)}.$

Demonstração: (a) Usando a Proposição 2.1.8(g) e o fato de que $-\log \mu(A|\mathcal{F})$ é uma função $\mathcal{F}$ -mensurável não-negativa para todo $A \in \alpha$, obtemos que

$$\begin{aligned}
 H_\mu(\alpha|\mathcal{F}) &= \int_X I_\mu(\alpha|\mathcal{F}) d\mu \\
 &= \int_X \sum_{A \in \alpha} -\log \mu(A|\mathcal{F}) 1_A d\mu \\
 &= \sum_{A \in \alpha} \int_X -\log \mu(A|\mathcal{F}) 1_A d\mu \\
 &= \sum_{A \in \alpha} \int_X -\log \mu(A|\mathcal{F}) \mu(A|\mathcal{F}) d\mu \\
 &= \int_X \sum_{A \in \alpha} -\mu(A|\mathcal{F}) \log \mu(A|\mathcal{F}) d\mu.
 \end{aligned}$$

(b) Usando a parte (a) e o Lema 2.1.9, temos

$$\begin{aligned}
 H_\mu(\alpha|\beta) &= \int_X \sum_{A \in \alpha} -\mu(A|\beta) \log \mu(A|\beta) d\mu \\
 &= \sum_{A \in \alpha} \int_X -\left(\sum_{B \in \beta} \mu(A|B) 1_B \right) \log \left(\sum_{B' \in \beta} \mu(A|B') 1_{B'} \right) d\mu \\
 &= \sum_{A \in \alpha} \int_X -\sum_{B \in \beta} \mu(A|B) 1_B \log \mu(A|B) d\mu \\
 &= \sum_{A \in \alpha} \sum_{B \in \beta} -\mu(B) \mu(A|B) \log \mu(A|B) \\
 &= \sum_{A \in \alpha} \sum_{B \in \beta} -\mu(A \cap B) \log \frac{\mu(A \cap B)}{\mu(B)}.
 \end{aligned}$$

□

Pelo Lema 2.1.12(a), considerando a função côncava $\varphi(t) = -t \log t$, temos que

$$H_\mu(\alpha|\mathcal{F}) = \sum_{A \in \alpha} \int_X \varphi(\mu(A|\mathcal{F})) d\mu.$$

Por outro lado, Pelo Lema 2.1.12(b), temos

$$H_\mu(\alpha|\beta) = \sum_{B \in \beta} \mu(B) \sum_{A \in \alpha} -\mu(A|B) \log \mu(A|B).$$

Assim, podemos ver $H_\mu(\alpha|\beta)$ como uma média ponderada de expressões do tipo

$$-\sum_{A \in \alpha} \mu(A|B) \log \mu(A|B).$$

Além disso, cada uma destas expressões pode ser considerada como a entropia condicional de α dado B .

Proposição 2.1.13 *Sejam $\alpha, \beta \in \mathcal{L}$ e $\mathcal{F} \subseteq \mathcal{B}$ uma σ -álgebra. Então:*

- (a) $H_\mu(\alpha|\mathcal{F}) = 0$ se e somente se $\alpha \subset \mathcal{F}$ a menos de conjuntos de medida nula.
- (b) $H_\mu(\alpha|\beta) = 0$ se, e somente se, $\alpha \leq \beta$.

Demonstração: (a) Se $\alpha \subset \mathcal{F}$, então $\mu(A|\mathcal{F}) = 1_A$ μ -q.t.p. para cada $A \in \alpha$. Segue-se daí que $I_\mu(\alpha|\mathcal{F}) = 0$ μ -q.t.p. e, portanto, que $H_\mu(\alpha|\mathcal{F}) = 0$. Reciprocamente, suponha que $H_\mu(\alpha|\mathcal{F}) = 0$. Então pelo Lema 2.1.12(a), temos que $\log \mu(A|\mathcal{F}) = 0$ μ -q.t.p. para todo $A \in \alpha$. Fixe $A \in \alpha$ arbitrariamente. Mostraremos que $A \in \mathcal{F}$ a menos de um conjunto de medida nula. Como $\log \mu(A|\mathcal{F}) = 0$ em quase todo ponto, obtemos que $\mu(A|\mathcal{F}) \in \{0, 1\}$. Isto significa que $\mu(A|\mathcal{F}) = 1_F$ μ -q.t.p. para algum $F \in \mathcal{F}$. Assim, obtemos

$$\mu(F) = \int_X 1_F d\mu = \int_X \mu(A|\mathcal{F}) d\mu = \mu(A).$$

Por outro lado, também vale que

$$\mu(A \cap F) = \int_X 1_F \cdot 1_A d\mu = \int_X 1_F \mu(A|\mathcal{F}) d\mu = \int_X 1_F d\mu = \mu(F).$$

Logo, como $\mu(A \cap F) = \mu(F) = \mu(A)$, concluímos que $A = F$ (mod μ). Portanto, $A \in \mathcal{F}$ a menos de um conjunto de medida nula.

(b) Pelo Lema 2.1.12(b) vemos que $H_\mu(\alpha|\beta) = 0$ se, e somente se, para quaisquer que sejam $A \in \alpha$ e $B \in \beta$ vale que

$$\mu(A \cap B) = 0 \quad \text{ou, então} \quad \frac{\mu(A \cap B)}{\mu(B)} = 1.$$

Em outras palavras, ou B é disjunto de A (a menos de medida nula) ou B está contido em A (a menos de medida nula). Isto significa que $H_\mu(\alpha|\beta) = 0$ se, e somente, se $\alpha \leq \beta$. □

Definição 2.1.14 *Seja $\{\mathcal{F}_\lambda\}_{\lambda \in \Lambda}$ uma família de σ -álgebras de X . Denote por $\bigvee_{\lambda \in \Lambda} \mathcal{F}_\lambda$ a σ -álgebra $\mathcal{F}$ que é unicamente definida pelas seguintes propriedades:*

- (i) $\mathcal{F}_\lambda \subseteq \mathcal{F}$ para toda $\lambda \in \Lambda$;
- (ii) se $\mathcal{G}$ é uma σ -álgebra tal que $\mathcal{F}_\lambda \subseteq \mathcal{G}$ para toda $\lambda \in \Lambda$, então $\mathcal{F} \subseteq \mathcal{G}$.

Observe que, nas condições da definição acima, $\mathcal{F}$ é a menor σ -álgebra que contém todas as σ -álgebras $\mathcal{F}_\lambda$ com $\lambda \in \Lambda$. Neste caso, dizemos que $\mathcal{F}$ é a σ -álgebra mais grossa que refina todas as σ -álgebras $\mathcal{F}_\lambda$, $\lambda \in \Lambda$.

Claramente, se β é partição enumerável e $\mathcal{F}$ é uma σ -álgebra, então a σ -álgebra $\mathcal{F} \vee \hat{\beta}$ é gerada pela família $\mathcal{A} = \{B \cap F : B \in \beta \text{ e } F \in \mathcal{F}\}$. Assim, os elementos de $\mathcal{F} \vee \hat{\beta}$ são da forma $\bigcup_{i \in I} B_i \cap F_i$, onde I é um subconjunto enumerável dos naturais, $B_i \in \beta$ e $F_i \in \mathcal{F}$.

Proposição 2.1.15 *Sejam $\alpha, \beta, \gamma \in \mathcal{L}$ e $\mathcal{F} \in \mathcal{B}$ uma σ -álgebra. Então:*

- (a) $I_\mu(\alpha \vee \beta | \mathcal{F}) = I_\mu(\beta | \mathcal{F}) + I_\mu(\alpha | \mathcal{F} \vee \hat{\beta})$. Em particular, vale $I_\mu(\alpha \vee \beta | \gamma) = I_\mu(\beta | \gamma) + I_\mu(\alpha | \gamma \vee \beta)$.
- (b) $H_\mu(\alpha \vee \beta | \mathcal{F}) = H_\mu(\beta | \mathcal{F}) + H_\mu(\alpha | \mathcal{F} \vee \hat{\beta})$. Em particular, vale $H_\mu(\alpha \vee \beta | \gamma) = H_\mu(\beta | \gamma) + H_\mu(\alpha | \gamma \vee \beta)$.

Demonstração: Observe que (b) segue-se de (a) por integração. Assim, é suficiente provarmos apenas (a). Mas, para tanto, precisamos da seguinte afirmação.

Afirmção: Se $A \in \mathcal{B}$, então $\mu(A | \mathcal{F} \vee \hat{\beta}) = \sum_{B \in \beta} \frac{\mu(A \cap B | \mathcal{F})}{\mu(B | \mathcal{F})} 1_B$.

Fixe $A \in \mathcal{B}$ arbitrário. Inicialmente, observe que a razão $\frac{\mu(A \cap B | \mathcal{F})}{\mu(B | \mathcal{F})}$ está bem definida para todo $B \in \beta$, com a conversão de que $\frac{0}{0} = 0$. Com efeito, pela Proposição 2.1.8, vemos que $0 \leq 1_{A \cap B} \leq 1_B$ implica que $0 \leq \mu(A \cap B | \mathcal{F}) \leq \mu(B | \mathcal{F})$. Para provar a afirmação, temos que verificar que a expressão do lado direito é a esperança condicional de 1_A com respeito a $\mathcal{F} \vee \hat{\beta}$. Como a σ -álgebra $\mathcal{F} \vee \hat{\beta}$ é gerada pelos conjuntos $F \cap B$ com $F \in \mathcal{F}$ e $B \in \beta$, é suficiente garantir que as integrais de $\mu(A | \mathcal{F} \vee \hat{\beta})$ e de $\sum_{B \in \beta} \frac{\mu(A \cap B | \mathcal{F})}{\mu(B | \mathcal{F})} 1_B$ coincidam nesses conjuntos. Fixe $F_0 \in \mathcal{F}$ e $B_0 \in \beta$. Por um lado,

$$\int_{F_0 \cap B_0} \mu(A | \mathcal{F} \vee \hat{\beta}) d\mu = \int_{F_0 \cap B_0} 1_A d\mu = \mu(A \cap B_0 \cap F_0).$$

Por outro lado, como a função $\frac{\mu(A \cap B_0 | \mathcal{F})}{\mu(B_0 | \mathcal{F})}$ é $\mathcal{F}$ -mensurável limitada, temos

$$\begin{aligned} \int_{F_0 \cap B_0} \left[\sum_{B \in \beta} \frac{\mu(A \cap B | \mathcal{F})}{\mu(B | \mathcal{F})} 1_B \right] d\mu &= \int_{F_0 \cap B_0} \frac{\mu(A \cap B_0 | \mathcal{F})}{\mu(B_0 | \mathcal{F})} 1_{B_0} d\mu \\ &= \int_{F_0} \frac{\mu(A \cap B_0 | \mathcal{F})}{\mu(B_0 | \mathcal{F})} 1_{B_0} d\mu \\ &= \int_{F_0} \frac{\mu(A \cap B_0 | \mathcal{F})}{\mu(B_0 | \mathcal{F})} \mu(B_0 | \mathcal{F}) d\mu \\ &= \int_{F_0} \mu(A \cap B_0 | \mathcal{F}) d\mu \\ &= \int_{F_0} \mu(A \cap B_0) d\mu \\ &= \mu(A \cap B_0 \cap F_0), \end{aligned}$$

onde na terceira igualdade tomámos a esperança condicional de 1_{B_0} com respeito a σ -álgebra $\mathcal{F}$.

Agora, concluiremos a prova de (a). Tomando log na equação da afirmação, temos

$$\sum_{B \in \beta} \log \mu(A \cap B | \mathcal{F}) 1_B = \sum_{B \in \beta} \log \mu(B | \mathcal{F}) 1_B + \log \mu(A | \mathcal{F} \vee \hat{\beta})$$

para todo $A \in \alpha$. Multiplicando por -1_A e somando em $A \in \alpha$, obtemos que

$$\sum_{A \in \alpha, B \in \beta} -\log \mu(A \cap B | \mathcal{F}) 1_B 1_A = \sum_{A \in \alpha, B \in \beta} -\log \mu(B | \mathcal{F}) 1_B 1_A + \sum_{A \in \alpha} -\mu(A | \mathcal{F} \vee \hat{\beta}) 1_A.$$

Agora, como $\sum_{A \in \alpha} 1_A = 1$ em quase todo ponto, concluimos que

$$\sum_{A \in \alpha, B \in \beta} -\log \mu(A \cap B | \mathcal{F}) 1_{B \cap A} = \sum_{B \in \beta} -\log \mu(B | \mathcal{F}) 1_B + \sum_{A \in \alpha} -\mu(A | \mathcal{F} \vee \hat{\beta}) 1_A$$

e, portanto, que

$$I_\mu(\alpha \vee \beta | \mathcal{F}) = I_\mu(\beta | \mathcal{F}) + I_\mu(A | \mathcal{F} \vee \hat{\beta})$$

o que prova a proposição. □

A proposição acima expressou $H_\mu(\alpha \vee \beta)$ em termos de $H_\mu(\alpha)$ e $H_\mu(\beta | \alpha)$. Assim, tomamos emprestada a intuição da teoria da informação: isto significa que a entropia de $\alpha \vee \beta$ deve ser igual à entropia de α mais a entropia condicional de β dado α . Na verdade, isso continua sendo verdade, mesmo no nível das funções de informação condicional.

Relembre a desigualdade de Jensen para funções côncavas.

Teorema 2.1.16 (Desigualdade de Jensen) *Seja $(X, \mathcal{B}, \mu)$ um espaço de medida e $\mathcal{F} \subseteq \mathcal{B}$ uma σ -álgebra. Se $f : X \rightarrow \mathbb{R}$ é integrável e $\psi : \mathbb{R} \rightarrow \mathbb{R}$ é côncava, então*

$$\psi(\mathbb{E}(f | \mathcal{F})) \geq \mathbb{E}(\psi \circ f | \mathcal{F}) \quad \mu\text{-q.t.p.}$$

Se ψ for estritamente côncava, então a igualdade ocorre μ -q.t.p. se e somente se f for $\mathcal{F}$ -mensurável.

Proposição 2.1.17 *Sejam $\alpha, \beta, \gamma \in \mathcal{Z}$ com $\alpha \leq \beta$ e $\mathcal{G}, \mathcal{F} \subseteq \mathcal{B}$ σ -álgebras tais que $\mathcal{G} \subseteq \mathcal{F}$. Então:*

- (a) $H_\mu(\alpha|\mathcal{F}) \leq H_\mu(\beta|\mathcal{F})$. Em particular, $H_\mu(\alpha|\gamma) \leq H_\mu(\beta|\gamma)$.
- (b) $H_\mu(\gamma|\mathcal{F}) \leq H_\mu(\gamma|\mathcal{G})$. Em particular, $H_\mu(\gamma|\beta) \leq H_\mu(\gamma|\alpha)$.
- (c) $H_\mu(\gamma|\beta) \leq H_\mu(\gamma)$.
- (d) $H_\mu(\beta \vee \gamma) \leq H_\mu(\beta) + H_\mu(\gamma)$.

Demonstração: (a) Como $H_\mu(\beta|\mathcal{F} \vee \hat{\alpha}) \geq 0$ e $\beta \vee \alpha = \beta$ (pois $\alpha \leq \beta$), temos que

$$H_\mu(\beta|\mathcal{F}) = H_\mu(\beta \vee \alpha|\mathcal{F}) = H_\mu(\alpha|\mathcal{F}) + H_\mu(\beta|\mathcal{F} \vee \hat{\alpha}) \geq H_\mu(\alpha|\mathcal{F}).$$

Em particular, tomando $\mathcal{F} = \hat{\gamma}$, obtemos que $H_\mu(\alpha|\gamma) \leq H_\mu(\beta|\gamma)$.

(b) Fixe $C \in \gamma$ arbitrário e considere a função côncava $\varphi(t) = -t \log t$. Pela desigualdade de Jensen, temos

$$\varphi(\mathbb{E}(1_C|\mathcal{G})) = \varphi(\mathbb{E}(\mathbb{E}(1_C|\mathcal{F})|\mathcal{G})) \geq \mathbb{E}(\varphi(\mathbb{E}(1_C|\mathcal{F})|\mathcal{G})) \quad \mu\text{-q.t.p.}$$

Integrando, obtemos

$$\int_X \varphi(\mathbb{E}(1_C|\mathcal{G})) d\mu \geq \int_X \mathbb{E}(\varphi(\mathbb{E}(1_C|\mathcal{F})|\mathcal{G})) d\mu = \int_X \varphi(\mathbb{E}(1_C|\mathcal{F})) d\mu.$$

Somando sobre $C \in \gamma$, conclui-se que

$$H_\mu(\gamma|\mathcal{G}) = \sum_{C \in \gamma} \int_X \varphi(\mathbb{E}(1_C|\mathcal{G})) d\mu \geq \sum_{C \in \gamma} \int_X \varphi(\mathbb{E}(1_C|\mathcal{F})) d\mu = H_\mu(\gamma|\mathcal{F}).$$

Em particular, tomando $\mathcal{F} = \hat{\beta}$ e $\mathcal{G} = \hat{\alpha}$ obtemos que $H_\mu(\gamma|\beta) \leq H_\mu(\gamma|\alpha)$.

(c) Tomando $\alpha = \{X\}$ e usando a caracterização fornecida pelo Lema 2.1.12(b), obtemos que $H_\mu(\gamma|\alpha) = H_\mu(\gamma)$. Assim, pelo item (b), concluímos que $H_\mu(\gamma|\beta) \leq H_\mu(\gamma)$.

(d) Usando a proposição 2.1.15(b) e o item (c), temos

$$H_\mu(\beta \vee \gamma) = H_\mu(\beta) + H_\mu(\gamma|\beta) \leq H_\mu(\beta) + H_\mu(\gamma),$$

donde $H_\mu(\beta \vee \gamma) \leq H_\mu(\beta) + H_\mu(\gamma)$.

□

Intuitivamente, o item (a) diz que partições mais finas codificam mais informação, o item (b) diz que quanto mais condicionamos menos incerteza temos, o item (c) diz que a incerteza de um condicionamento nunca supera a incerteza inicial e o item (d) diz que a incerteza de um refinamento não supera a soma das incertezas.

Usando a Proposição 2.1.15(b) e a Proposição 2.1.17, obtemos

$$H_\mu(\alpha|\beta) \leq H_\mu(\alpha \vee \gamma|\beta) = H_\mu(\gamma|\beta) + H_\mu(\alpha|\beta \vee \gamma) \leq H_\mu(\gamma|\beta) + H_\mu(\alpha|\gamma),$$

quaisquer que sejam $\alpha, \beta, \gamma \in \mathcal{Z}$. A propriedade acima, que é muito útil para a realização de cálculos, é uma espécie de regra da cadeira para a entropia de partições enumeráveis com entropia finita. Na realidade, tal propriedade é verdadeira quaisquer que sejam as partições enumeráveis α, β, γ .

O próximo resultado apresenta uma propriedade de continuidade para a entropia condicional de partições finitas.

Proposição 2.1.18 *Dados $k \geq 1$ e $\varepsilon > 0$ existe $\delta > 0$ tal que, para quaisquer partições $\alpha = \{A_1, A_2, \dots, A_k\}$ e $\beta = \{B_1, B_2, \dots, B_k\}$,*

$$\mu(A_i \Delta B_i) < \delta \text{ para todo } i = 1, \dots, k \implies H_\mu(\alpha|\beta) < \varepsilon.$$

Demonstração: Inicialmente, fixe $k \geq 1$ e $\varepsilon > 0$ de forma arbitrária. Pela continuidade da função $\varphi : [0, 1] \rightarrow \mathbb{R}$, $\varphi(t) = -t \log(t)$, existe $\rho > 0$ tal que $\varphi(t) < \varepsilon/k^2$ para todo $t \in [0, \rho) \cup (1 - \rho, 1]$. Tome $\delta = \rho/k$. Sejam $\alpha = \{A_1, A_2, \dots, A_k\}$ e $\beta = \{B_1, B_2, \dots, B_k\}$ duas partições tais que $\mu(A_i \Delta B_i) < \delta$ para todo $i = 1, \dots, k$. Agora, denote por η a partição de X cujos átomos são as interseções $A_i \cap B_j$ com $i \neq j$ juntamente com o conjunto $\cup_{i=1}^k A_i \cap B_i$. Observe que $\mu(A_i \cap B) \leq \mu(A_i \Delta B_i) < \delta$ para todo $i \neq j$ e

$$\mu\left(\bigcup_{i=1}^k A_i \cap B_i\right) \geq \sum_{i=1}^k (\mu(A_i) - \mu(A_i \Delta B_i)) > \sum_{i=1}^k (\mu(A_i) - \delta) = 1 - \rho.$$

Assim,

$$H_\mu(\eta) = \sum \varphi(\mu(R)) < \# \eta \frac{\varepsilon}{k^2} \leq \varepsilon.$$

Pela definição da partição η , tem-se $\alpha \vee \beta = \alpha \vee \eta$. Então, usando Proposição 2.1.17, obtemos

$$\begin{aligned} H_\mu(\alpha|\beta) &= H_\mu(\alpha \vee \beta) - H_\mu(\alpha) \\ &= H_\mu(\alpha \vee \eta) - H_\mu(\alpha) \\ &= H_\mu(\eta|\alpha) \leq H_\mu(\eta) < \varepsilon, \end{aligned}$$

o que prova a proposição.

□

2.2 Entropia de um sistema dinâmico

Seja $T : X \rightarrow X$ uma transformação mensurável no espaço de probabilidade de Lebesgue $(X, \mathcal{B}, \mu)$, isto é, $T^{-1}(A) \in \mathcal{B}$ para todo $A \in \mathcal{B}$. Dizemos que T preserva μ ou que μ é uma medida T -invariante se $\mu(T^{-1}(A)) = \mu(A)$ para todo $A \in \mathcal{B}$. Quando T preserva μ a quádrupla $(X, \mathcal{B}, \mu, T)$ é chamada de *endomorfismo*. Por indução, vemos que se T preserva μ , então T^n preserva μ para todo $n \geq 1$. Isso mostra que se $(X, \mathcal{B}, \mu, T)$ é um endomorfismo, então $(X, \mathcal{B}, \mu, T^n)$ é um endomorfismo para todo $n \geq 1$.

Se $(X, \mathcal{B}, \mu, T)$ é um endomorfismo com T bimensurável, isto é T e T^{-1} são aplicações mensuráveis, então T^{-1} também preserva medida. De fato, dado $A \in \mathcal{B}$ temos que $T(A) = (T^{-1})^{-1}(A) \in \mathcal{B}$, pois T^{-1} é mensurável. Assim,

$$\mu((T^{-1})^{-1}(A)) = \mu(T(A)) = \mu(T^{-1}(T(A))) = \mu(A),$$

o que mostra que T^{-1} preserva μ . Quando T é uma aplicação bimensurável que preserva μ , a quádrupla $(X, \mathcal{B}, \mu, T)$ é chamada de *automorfismo*. Pelo mesmo argumento do parágrafo acima, vemos que se $(X, \mathcal{B}, \mu, T)$ é um automorfismo, então $(X, \mathcal{B}, \mu, T^n)$ é um automorfismo para todo $n \in \mathbb{Z}$.

Proposição 2.2.1 *Sejam $T : X \rightarrow X$ uma transformação mensurável e μ uma medida em X . Então T preserva μ se, e somente se,*

$$\int_X \phi d\mu = \int_X (\phi \circ T) d\mu \quad (2.2)$$

para toda função μ -integrável $\phi : X \rightarrow \mathbb{R}$.

Demonstração: Suponha que a medida μ é T -invariante. Devemos mostrar que a igualdade (2.2) é verdadeira para toda função μ -integrável $\phi : X \rightarrow \mathbb{R}$. Por hipótese, temos que $\mu(A) = \mu(T^{-1}A)$ para todo conjunto mensurável. Assim, como

$$\int_X 1_A d\mu = \mu(A) = \mu(T^{-1}A) = \int_X 1_A \circ T d\mu,$$

vemos que a igualdade (2.2) para as funções características. Então, pela linearidade da integral, vemos que (2.2) é válida para as funções simples. Agora, dada uma função integrável $\phi : X \rightarrow \mathbb{R}$ arbitrária, considere uma sequência $(s_n)_n$ de funções simples convergindo para ϕ e tal que $|s_n| \leq |\phi|$ para todo natural n (a existência da sequência

$(s_n)_n$ é uma propriedade clássica de espaços de medida). Então, usando o teorema da convergência dominada duas vezes, obtemos

$$\int_X \phi d\mu = \lim_n \int_X s_n d\mu = \lim_n \int_X s_n \circ T d\mu = \int_X \phi \circ T d\mu.$$

Isso mostra que (2.2) vale para toda função integrável. A recíproca é imediata da observação de que $1_{T^{-1}A} = 1_A \circ T$.

□

Seja α uma partição enumerável de X . Para cada $n \in \mathbb{Z}$, defina $T^{-n}\alpha := \{T^{-n}(A) : A \in \alpha\}$. Suponha que $(X, \mathcal{B}, \mu, T)$ é um endomorfismo e fixe $n \geq 0$. Então, $T^{-n}\alpha$ é uma partição mensurável de X , pois T^{-n} é mensurável. Além disso, como T^n pereserva μ , temos que $H_\mu(T^{-n}\alpha) = H_\mu(\alpha)$. Em particular, se $\alpha \in \mathcal{Z}$, então $T^{-n}\alpha \in \mathcal{Z}$ para todo $n \geq 0$. Pela mesma razão, supondo que $(X, \mathcal{B}, \mu, T)$ é um automorfismo e tomando $n \in \mathbb{Z}$, temos que $H_\mu(T^n\alpha) = H_\mu(\alpha)$. Em particular, se $\alpha \in \mathcal{Z}$ então $T^n\alpha \in \mathcal{Z}$ para todo $n \in \mathbb{Z}$.

Dados $m \geq n \geq 0$ e $\alpha \in \mathcal{Z}$, defina $\alpha_n^m := \bigvee_{i=n}^m T^{-i}\alpha$. Quando T é bimensurável, podemos tomar inteiros $m \geq n$ quaisquer na definição de α_n^m . Segue-se imediatamente da definição de soma de partições que $\alpha_n^m = T^{-n}\alpha_0^{m-n}$. Além disso, usando a Proposição 2.1.17(d), concluímos que $\alpha_n^m \in \mathcal{Z}$. Intuitivamente, $H_\mu(\alpha_n^m)$ mede a complexidade/imprevisibilidade das potências $T^n, \dots, T^m$ na resolução fornecida por α .

Definição 2.2.2 *Sejam $(X, \mathcal{B}, \mu, T)$ um endomorfismo e $\alpha \in \mathcal{Z}$. A entropia métrica de T com relação à partição α é*

$$h_\mu(T, \alpha) := \lim_{n \rightarrow +\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) = \lim_{n \rightarrow +\infty} \frac{1}{n} H_\mu(\alpha_0^n). \quad (2.3)$$

A entropia $h_\mu(T, \alpha)$ mede o crescimento de incerteza da partição α por iterações de T .

Mostremos que (2.3) está bem definida. É claro que se um dos dois limites de (2.3) existe, então ambos existem e coincidem. Assim, é suficiente mostrar que a sequência $\{\frac{1}{n} H_\mu(\alpha_0^{n-1})\}_{n \geq 1}$ converge. Fixe $m, n \geq 1$ arbitrários. Pela Proposição 2.1.17(d) e pela invariância de T^n por μ , temos

$$H_\mu(\alpha_0^{m+n-1}) \leq H_\mu(\alpha_0^{n-1}) + H_\mu(\alpha_n^{m+n-1}) = H_\mu(\alpha_0^{n-1}) + H_\mu(\alpha_0^{m-1}),$$

portanto $\{H_\mu(\alpha_0^{n-1})\}_{n \geq 1}$ é uma sequência subaditiva não-negativa. Assim, a existência do limite segue do lema abaixo.

Lema 2.2.3 (Fekete) *Seja $\{a_n\}_{n \geq 1}$ uma sequência subaditiva de números reais não-negativos. Então*

$$\lim_{n \rightarrow +\infty} \frac{a_n}{n} = \inf_{n \geq 0} \frac{a_n}{n}.$$

Demonstração: Seja $L = \inf_{n \geq 0} \frac{a_n}{n} \in [0, +\infty)$. Tome $M \in \mathbb{R}$ maior do que L . Então, usando a definição de ínfimo de um conjunto, podemos encontrar $k \geq 1$ tal que

$$\frac{a_k}{k} < M.$$

Para $n > k$, podemos escrever $n = pk + q$, onde p e q são números inteiros tais que $p \geq 1$ e $1 \leq q \leq k$. Assim, pela subaditividade de $\{a_n\}_{n \geq 1}$, temos

$$a_n \leq a_{kp} + a_q \leq pa_k + c,$$

onde $c = \max\{a_j : 1 \leq j \leq k\}$. Logo,

$$\frac{a_n}{n} \leq \frac{pk}{n} \frac{a_k}{k} + \frac{c}{n}.$$

Fazendo $n \rightarrow +\infty$, vemos que $\frac{pk}{n}$ converge para 1 e $\frac{c}{n}$ converge para zero. Assim, uma vez que $\frac{a_k}{k} < M$, temos que

$$L \leq \frac{a_n}{n} < M$$

para todo n suficientemente grande. Em particular, isso mostra que

$$L \leq \liminf_n \frac{a_n}{n} \leq \limsup_n \frac{a_n}{n} \leq B.$$

Agora, fazendo $M \rightarrow L$, obtemos

$$\lim_{n \rightarrow +\infty} \frac{a_n}{n} = L = \inf_n \frac{a_n}{n},$$

o que completa a demonstração. □

Uma das muitas interpretações de $h_\mu(T, \alpha)$ é que ela mede a incerteza de $\alpha(x)$ dado que conhecemos todo o futuro $\alpha(Tx), \alpha(T^2x), \dots$ de x . Claramente, $h_\mu(T, \alpha)$ é monótona sob refinamentos, isto é, se $\alpha \leq \beta$ então $h_\mu(T, \alpha) \leq h_\mu(T, \beta)$.

Observe que, em automorfismos, temos

$$\lim_{n \rightarrow +\infty} \frac{1}{n} H_\mu(\alpha_{-(n-1)}^0) = \lim_{n \rightarrow +\infty} \frac{1}{n} H_\mu(T^{n-1} \alpha_0^{n-1}) = \lim_{n \rightarrow +\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) = h_\mu(T, \alpha).$$

Isto mostra que, em automorfismos, podemos recuperar a incerteza $h_\mu(T, \alpha)$ de $\alpha(x)$ conhecendo o passado $\alpha(T^{-1}x), \alpha(T^{-2}x), \dots$ de x .

Uma propriedade muito útil, para fins operacionais de cálculos, é a invariância da entropia condicional.

Proposição 2.2.4 *Sejam $\alpha \in \mathcal{L}$ e $\mathcal{F} \subseteq \mathcal{B}$ uma σ -álgebra. Se $(X, \mathcal{B}, \mu, T)$ é um endomorfismo, então $H_\mu(T^{-1}\alpha|T^{-1}\mathcal{F}) = H_\mu(\alpha|\mathcal{F})$. Adicionalmente, se T for uma aplicação invertível bimensurável, vale ainda que $H_\mu(T\alpha|T\mathcal{F}) = H_\mu(\alpha|\mathcal{F})$.*

Demonstração: Inicialmente, provaremos a primeira afirmação. Segue-se diretamente das propriedades de esperança condicional que $\mu(T^{-1}A|T^{-1}\mathcal{F})(x) = \mu(A|\mathcal{F})(Tx)$ para μ -q.t.p. $x \in X$. Assim, como μ é T -invariante, temos

$$\begin{aligned} H_\mu(T^{-1}\alpha|T^{-1}\mathcal{F}) &= \int_X \sum_{A \in \alpha} -\mu(T^{-1}A|T^{-1}\mathcal{F})(x) \log \mu(T^{-1}A|T^{-1}\mathcal{F})(x) d\mu \\ &= \int_X \sum_{A \in \alpha} -\mu(A|\mathcal{F})(Tx) \log \mu(A|\mathcal{F})(Tx) d\mu \\ &= \int_X \sum_{A \in \alpha} -\mu(A|\mathcal{F})(x) \log \mu(A|\mathcal{F})(x) d\mu \\ &= H_\mu(\alpha|\mathcal{F}), \end{aligned}$$

o que prova primeira afirmação. A segunda afirmação decorre da primeira, uma vez que $(T^{-1})^{-1} = T$ e $(X, \mathcal{B}, \mu, T^{-1})$ é um endomorfismo. □

Seja $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ uma família de sub- σ -álgebras de $\mathcal{B}$. Escreveremos $\mathcal{F}_n \uparrow \mathcal{F}$ para expressar que $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots \subseteq \mathcal{F}$ e que $\mathcal{F}$ é a σ -álgebra gerada por $\cup_{n \in \mathbb{N}} \mathcal{F}_n$ (isto é, $\mathcal{F} = \sigma(\cup_{n \in \mathbb{N}} \mathcal{F}_n)$). Similarmente, escreveremos $\mathcal{F}_n \downarrow \mathcal{F}$ para indicar que $\mathcal{F}_1 \supseteq \mathcal{F}_2 \supseteq \dots \supseteq \mathcal{F}$ e que $\mathcal{F} = \cap_{n \in \mathbb{N}} \mathcal{F}_n$.

Teorema 2.2.5 (Teorema de convergência de martingais) *Sejam $\{\mathcal{F}_n\}_{n \in \mathbb{N}}$ uma sequência de sub- σ -álgebras de $\mathcal{B}$ e $f \in L^1(X, \mathcal{B}, \mu)$. Se valer que $\mathcal{F}_n \uparrow \mathcal{F}$ ou que $\mathcal{F}_n \downarrow \mathcal{F}$, então*

$$\mathbb{E}(f|\mathcal{F}_n) \xrightarrow{n \rightarrow +\infty} \mathbb{E}(f|\mathcal{F})$$

em quase todo ponto e em $L^1(X, \mathcal{B}, \mu)$.

Demonstração: Veja [3, Teoremas 5.5 e 5.8]. □

Teorema 2.2.6 (Lema de Chung-Neveu) Se $\alpha \in \mathcal{L}$ e $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots$ são sub- σ -álgebras de $\mathcal{B}$, então a função $f^* = \sup_{n \geq 0} I_\mu(\alpha | \mathcal{F}_n)$ é integrável.

Demonstração: Dados $A \in \alpha$, $n \geq 0$ em $\mathbb{Z}$ e um número real $\lambda \geq 0$, defina

$$B_n(\lambda; A) := \{x \in X : n \text{ é o menor natural tal que } -\log \mu(A | \mathcal{F}_n)(x) > \lambda\}.$$

Claramente, $B_n(\lambda; A)$ é $\mathcal{F}_n$ -mensurável com $\mu(B_n(\lambda; A)) \leq \mu(A)$. Além disso, dados $A \in \alpha$ e $\lambda \geq 0$, temos que $B_n(\lambda; A) \cap B_m(\lambda; A) = \emptyset$ para $n \neq m$ em $\mathbb{Z}$. Fixe $A \in \alpha$ e $\lambda \geq 0$ arbitrariamente. Então,

$$\begin{aligned} \mu(A \cap B_n(\lambda; A)) &= \int_X 1_A 1_{B_n(\lambda; A)} d\mu \\ &= \int_X \mathbb{E}(1_A 1_{B_n(\lambda; A)} | \mathcal{F}_n) d\mu \\ &= \int_X 1_{B_n(\lambda; A)} \mathbb{E}(1_A | \mathcal{F}_n) d\mu \\ &= \int_X 1_{B_n(\lambda; A)} \mu(1_A | \mathcal{F}_n) d\mu \\ &= \int_X 1_{B_n(\lambda; A)} e^{\log \mu(1_A | \mathcal{F}_n)} d\mu \\ &\leq \int_X 1_{B_n(\lambda; A)} e^{-\lambda} d\mu \\ &= e^{-\lambda} \mu(B_n(\lambda; A)), \end{aligned}$$

para todo inteiro $n \geq 0$. Agora, como $f^* = \sup_{n \geq 0} I_\mu(\alpha | \mathcal{F}_n)$ e $I_\mu(\alpha | \mathcal{F}_n)(x) = -\log \mu(A | \mathcal{F}_n)(x)$ em A , temos

$$A \cap [f^* > \lambda] = \bigcup_{n \geq 0} (A \cap B_n(\lambda; A)).$$

Portanto,

$$\mu(A \cap [f^* > \lambda]) = \sum_{n \geq 0} \mu(A \cap B_n(\lambda; A)) \leq e^{-\lambda} \sum_{n \geq 0} \mu(B_n(\lambda; A)) \leq e^{-\lambda} \mu(A).$$

Isso mostra que $\mu(A \cap B_n(\lambda; A)) \leq \min\{\mu(A), e^{-\lambda}\}$ para todo $A \in \alpha$ e todo $\lambda \geq 0$. Assim,

como $f^* \geq 0$, temos

$$\begin{aligned}
\int_X f^* d\mu &= \int_X \int_0^\infty 1_{[0 \leq \lambda < f^*(x)]}(x, \lambda) d\lambda d\mu(x) \\
&= \int_0^\infty \int_X 1_{[0 \leq \lambda < f^*(x)]}(x, \lambda) d\mu(x) d\lambda \\
&= \int_0^\infty \mu([f^* > \lambda]) d\lambda \\
&= \sum_{A \in \alpha} \int_0^\infty \mu(A \cap [f^* > \lambda]) d\lambda \\
&\leq \sum_{A \in \alpha} \int_0^\infty \min\{\mu(A), e^{-\lambda}\} d\lambda \\
&= \sum_{A \in \alpha} \left[\int_0^{-\log \mu(A)} \mu(A) d\lambda + \int_{-\log \mu(A)}^\infty e^{-\lambda} d\lambda \right] \\
&= \sum_{A \in \alpha} [-\mu(A) \log \mu(A) + \mu(A)] \\
&= H_\mu(\alpha) + 1 < \infty,
\end{aligned}$$

o que prova o teorema. □

Corolário 2.2.7 *Se $\alpha \in \mathcal{L}$ e $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \dots$ são sub- σ -álgebras de $\mathcal{B}$, então*

$$H_\mu(\alpha|\mathcal{F}_n) \xrightarrow{n \rightarrow +\infty} H_\mu(\alpha|\mathcal{F}_\infty)$$

onde $\mathcal{F}_\infty = \bigvee_{n \geq 0} \mathcal{F}_n$.

Demonstração: Fixe $A \in \alpha$ qualquer. Pelo Teorema de convergência de martingais, $\mu(A|\mathcal{F}_n)(x) \xrightarrow{n \rightarrow +\infty} \mu(A|\mathcal{F}_\infty)(x)$ para μ -q.t.p. $x \in X$ (e em L^1), portanto por continuidade obtemos que $-\log \mu(A|\mathcal{F}_n) \xrightarrow{n \rightarrow +\infty} -\log \mu(A|\mathcal{F}_\infty)$ para μ -q.t.p. $x \in X$ (e em L^1). Assim, obtemos que

$$I_\mu(\alpha|\mathcal{F}_n) = \sum_{A \in \alpha} -\log \mu(A|\mathcal{F}_n) 1_A \xrightarrow{n \rightarrow +\infty} \sum_{A \in \alpha} -\log \mu(A|\mathcal{F}_\infty) 1_A = I_\mu(\alpha|\mathcal{F}_\infty)$$

para μ -q.t.p. $x \in X$. Logo, pelo Lema de Chung-Neveu e pelo Teorema da Convergência Dominada, concluímos que $H_\mu(\alpha|\mathcal{F}_n) \xrightarrow{n \rightarrow +\infty} H_\mu(\alpha|\mathcal{F}_\infty)$. □

Dados uma partição enumerável α do conjunto X e um número inteiro $n \geq 0$, seja $\alpha_n^\infty = \bigvee_{i=n}^\infty T^{-i}\alpha$. Por definição, α_n^∞ é a partição mais grosseira que refina toda $T^{-i}\alpha$ com $i \geq n$.

Teorema 2.2.8 *Sejam $(X, \mathcal{B}, \mu, T)$ um endomorfismo e $\alpha \in \mathcal{X}$. Então,*

$$h_\mu(T, \alpha) = \lim_{n \rightarrow +\infty} H_\mu(\alpha | \alpha_1^n) = H_\mu(\alpha | \alpha_1^\infty). \quad (2.4)$$

Demonstração: A segunda igualdade é o Corolário 2.2.7. Provaremos que $h_\mu(T, \alpha) = H_\mu(\alpha | \alpha_1^\infty)$. Fixe $n \geq 1$ arbitrariamente. Então, utilizamos repetidamente a Proposição 2.1.15(b) juntamente com o fato de T ser μ -invariante, vemos que

$$\begin{aligned} H_\mu(\alpha_0^n) &= H_\mu(\alpha_1^n) + H_\mu(\alpha | \alpha_1^n) \\ &= H_\mu(T^{-1}(\alpha_0^{n-1})) + H_\mu(\alpha | \alpha_1^n) \\ &= H_\mu(\alpha_0^{n-1}) + H_\mu(\alpha | \alpha_1^n) \\ &= H_\mu(\alpha_0^{n-2}) + H_\mu(\alpha | \alpha_1^{n-1}) + H_\mu(\alpha | \alpha_1^n) \\ &= \dots \\ &= H_\mu(\alpha | \alpha_1^1) + H_\mu(\alpha | \alpha_1^2) + \dots + H_\mu(\alpha | \alpha_1^n). \end{aligned}$$

Assim, dado $n \geq 1$, temos

$$\frac{1}{n} H_\mu(\alpha_0^n) = \frac{1}{n} \sum_{i=1}^n H_\mu(\alpha | \alpha_1^i).$$

Como $\{H_\mu(\alpha | \alpha_1^i)\}_{n \geq 1}$ é uma sequência não-crescente de números reais não-negativos que converge para $H_\mu(\alpha | \alpha_1^\infty)$, suas médias Cesàro converge para o mesmo limite. Logo,

$$h_\mu(T, \alpha) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=1}^n H_\mu(\alpha | \alpha_1^i) = H_\mu(\alpha | \alpha_1^\infty).$$

Isso completa a demonstração do teorema. □

Proposição 2.2.9 *Seja $(X, \mathcal{B}, \mu, T)$ um endomorfismo. Então*

$$h_\mu(T, \alpha) \leq h_\mu(T, \beta) + H_\mu(\alpha | \beta)$$

para quaisquer que sejam $\alpha, \beta \in \mathcal{X}$.

Demonstração: Fixe $\alpha, \beta \in \mathcal{X}$ arbitrariamente. Usando as Proposições 2.1.15 e 2.1.17 juntamente com a invariância de μ por T , obtemos

$$\begin{aligned} H_\mu(\alpha_0^{n-1} | \beta_0^{n-1}) &\leq H_\mu(\alpha | \beta_0^{n-1}) + H_\mu(T^{-1} \alpha | \beta_0^{n-1}) + \dots + H_\mu(T^{-n+1} \alpha | \beta_1^{n-1}) \\ &\leq H_\mu(\alpha | \beta) + H_\mu(T^{-1} \alpha | T^{-1} \beta) + \dots + H_\mu(T^{-n+1} \alpha | T^{-n+1} \beta) \\ &= n H_\mu(\alpha | \beta), \end{aligned}$$

para todo $n \in \mathbb{N}$. Assim, usando novamente a Proposição 2.1.17, obtemos

$$\begin{aligned} H_\mu(\alpha_0^{n-1}) &\leq H_\mu(\alpha_0^{n-1} \vee \beta_0^{n-1}) \\ &= H_\mu(\beta_0^{n-1}) + H_\mu(\alpha_0^{n-1} | \beta_0^{n-1}) \\ &\leq H_\mu(\beta_0^{n-1}) + nH_\mu(\alpha | \beta), \end{aligned}$$

para todo $n \in \mathbb{N}$. Portanto,

$$\begin{aligned} h_\mu(T, \alpha) &= \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}) \\ &\leq \lim_{n \rightarrow \infty} \frac{1}{n} H_\mu(\beta_0^{n-1}) + H_\mu(\alpha | \beta) \\ &= h_\mu(T, \alpha) + H_\mu(\alpha | \beta). \end{aligned}$$

□

Definição 2.2.10 (Entropia métrica) *Sejam $(X, \mathcal{B}, \mu, T)$ um endomorfismo. A entropia métrica de T (também chamada de entropia de Kolmogorov-Snail) é*

$$h_\mu(T) := \sup\{h_\mu(T, \alpha) : \alpha \in \mathcal{Z}\}.$$

Teorema 2.2.11 *Seja $(X, \mathcal{B}, \mu, T)$ um endomorfismo. Então*

$$h_\mu(T) := \sup\{h_\mu(T, \alpha) : \alpha \text{ é uma partição finita}\}.$$

Demonstração: Seja $\alpha = \{A_i : i \in \mathbb{N}\}$ uma partição enumerável de X com entropia finita. Para cada $k \in \mathbb{N}$ defina $\alpha_k = \{A_1, \dots, A_k, \cup_{j>k} A_j\}$. Temos que $\alpha_k \leq \alpha_{k+1} \leq \alpha$ para todo $k \in \mathbb{N}$. Assim, pelo Proposição 2.2.9, tem-se

$$h_\mu(T, \alpha_k) \leq h_\mu(T, \alpha) \leq h_\mu(T, \alpha_k) + H_\mu(\alpha | \alpha_k), \quad (2.5)$$

para todo $k \in \mathbb{N}$.

Lema 2.2.12 *Sejam $\alpha = \{A_i : i \in \mathbb{N}\} \in \mathcal{Z}$ e $\alpha_k = \{A_1, \dots, A_k, \cup_{j>k} A_j\}$ para $k \geq 1$. Então,*
 $\lim_{k \rightarrow \infty} H_\mu(\alpha | \alpha_k) = 0$.

Demonstração: De fato, fixado $k \in \mathbb{N}$, temos

$$\begin{aligned}
H_\mu(\alpha|\alpha_k) &= \sum_{i=1}^{+\infty} \sum_{A \in \alpha_k} -\mu(A_i \cap A) \log \frac{\mu(A_i \cap A)}{\mu(A)} \\
&= \sum_{i=1}^{+\infty} \sum_{j=1}^k -\mu(A_i \cap A_j) \log \frac{\mu(A_i \cap A_j)}{\mu(A_j)} \\
&\quad + \sum_{i=1}^{+\infty} -\mu(A_i \cap (\cup_{j>k} A_j)) \log \frac{\mu(A_i \cap (\cup_{j>k} A_j))}{\mu(\cup_{j>k} A_j)} \\
&= \sum_{i>k} -\mu(A_i) \log \frac{\mu(A_i)}{\mu(\cup_{j>k} A_j)} \\
&\leq \sum_{i>k} -\mu(A_i) \log \mu(A_i).
\end{aligned}$$

Agora, sendo $H_\mu(A) = \sum_{i \geq 1} -\mu(A_i) \log \mu(A_i) < \infty$, segue-se

$$\lim_{k \rightarrow +\infty} H_\mu(\alpha|\alpha_k) \leq \lim_{k \rightarrow +\infty} \sum_{i>k} -\mu(A_i) \log \mu(A_i) = 0.$$

□

Usando o Lema 2.2.12 anterior e a desigualdade (2.5), vemos que

$$h_\mu(T, \alpha) = \lim_{k \rightarrow +\infty} h_\mu(T, \alpha_k) = \sup_k h_\mu(T, \alpha_k).$$

Segue-se daí, que $h_\mu(T, \alpha)$ é o supremo de uma sequencia de partições finitas de X , qualquer que seja $\alpha \in \mathcal{Z}$. Portanto, o supremo de $h_\mu(T, \alpha)$ sobre todas partições finitas coincide com o supremo de $h_\mu(T, \alpha)$ sobre todas as partições enumeráveis com entropia finita.

□

Proposição 2.2.13 *Se $(X, \mathcal{B}, \mu, T)$ é um endomorfismo, então $h_\mu(T^n) = nh_\mu(T)$, para todo $n \geq 0$. Se $(X, \mathcal{B}, \mu, T)$ é um automorfismo, então $h_\mu(T^n) = |n|h_\mu(T)$, para todo $n \in \mathbb{Z}$.*

Demonstração: Se $n = 0$ ambas afirmações valem trivialmente, uma vez que $f^0 = id_X$ e a igualdade $h_\mu(id_X) = 0$ é óbvia. No que se segue, supomos que $n \neq 0$.

Suponha que $(X, \mathcal{B}, \mu, T)$ é um endomorfismo. Dado $\alpha \in \mathcal{Z}$, temos

$$\frac{1}{k} H_\mu \left(\bigvee_{i=0}^{k-1} T^{-ni} \left(\bigvee_{j=0}^{n-1} T^{-j} \alpha \right) \right) = n \frac{1}{nk} H_\mu \left(\bigvee_{i=0}^{nk-1} T^{-i} \alpha \right),$$

para todo $k \in \mathbb{N}$. Fazendo $k \rightarrow +\infty$, obtemos que $h_\mu(T^n, \alpha_0^{n-1}) = nh_\mu(T, \alpha)$. Assim, tomando o supremo em $\alpha \in \mathcal{Z}$, vemos que $h_\mu(T^n) \geq nh_\mu(T)$. Reciprocamente, dado

$\alpha \in \mathcal{Z}$, temos

$$\frac{1}{k} H_\mu \left(\bigvee_{i=0}^{k-1} T^{-in} \alpha \right) \leq n \frac{1}{nk} H_\mu \left(\bigvee_{i=0}^{nk-1} T^{-i} \alpha \right),$$

para todo $k \in \mathbb{N}$. Fazendo $k \rightarrow +\infty$, obtemos que $h_\mu(T^n, \alpha) \leq n h_\mu(T, \alpha)$. Logo, tomando o supremo em $\alpha \in \mathcal{Z}$, obtemos que $h_\mu(T^n) \leq n h_\mu(T)$. Isso mostra que $h_\mu(T^n) = n h_\mu(T)$.

Agora, suponha que $(X, \mathcal{B}, \mu, T)$ é um automorfismo. Seja $\alpha \in \mathcal{Z}$ arbitrária.

Então, para qualquer $n \in \mathbb{N}$,

$$H_\mu \left(\bigvee_{i=0}^{n-1} T^{-i} \alpha \right) = H_\mu \left(T^{-n+1} \left(\bigvee_{j=0}^{n-1} T^j \alpha \right) \right) = H_\mu \left(\bigvee_{j=0}^{n-1} T^j \alpha \right)$$

uma vez que T é μ -invariante. Dividindo por n e fazendo $n \rightarrow +\infty$, obtemos que $h_\mu(T, \alpha) = h_\mu(T^{-1}, \alpha)$. Tomando o supremo sobre $\alpha \in \mathcal{Z}$, concluímos que $h_\mu(T) = h_\mu(T^{-1})$. Substituindo T por T^n e usando a primeira afirmação da proposição, concluímos que $h_\mu(T^{-n}) = h_\mu(T^n) = n h_\mu(T)$ para todo $n \in \mathbb{N}$. □

Proposição 2.2.14 *Sejam $(X, \mathcal{B}, \mu, T)$ um automorfismo e $\alpha \in \mathcal{Z}$. Então, para cada $n \geq 1$, vale que $h_\mu(T, \alpha) = h_\mu(T, \alpha_{-n}^n)$.*

Demonstração: Com efeito, dados $n, k \in \mathbb{N}$, temos

$$\bigvee_{j=0}^k T^{-j}(\alpha_{-n}^n) = \bigvee_{j=0}^k T^{-j} \left(\bigvee_{i=-n}^n T^{-i} \alpha \right) = \bigvee_{l=-n}^{n+k} T^{-l} \alpha = T^n \left(\bigvee_{i=0}^{2n+k} T^{-i} \alpha \right) = T^n(\alpha_0^{2n+k}).$$

Assim, como μ é T^{-1} -invariante, obtemos

$$h_\mu(T, \alpha_{-n}^n) = \lim_{k \rightarrow +\infty} \frac{1}{k} H_\mu(T^n(\alpha_0^{2n+k})) = \lim_{k \rightarrow +\infty} \frac{2n+k}{k(2n+k)} H_\mu(\alpha_0^{2n+k}) = h_\mu(T, \alpha),$$

para todo $n \in \mathbb{N}$. □

Definição 2.2.15 (Gerador e gerador forte) *Sejam $(X, \mathcal{B}, \mu, T)$ um automorfismo e $\alpha \in \mathcal{Z}$. Quando $\widehat{(\alpha_{-\infty}^{+\infty})} = \mathcal{B}$ módulo μ , dizemos que α é gerador para T . Quando $\widehat{(\alpha_0^{+\infty})} = \mathcal{B}$ módulo μ , dizemos que α é um gerador forte para T .*

Enquanto a noção de geradores só faz sentido para automorfismos, a noção de geradores fortes também faz sentido para endomorfismos. Claramente, geradores fortes são geradores. Uma propriedade importante dos geradores é que podemos calcular a entropia métrica a partir de uma única partição geradora.

Teorema 2.2.16 (Kolmogorov-Sinai) *Seja $(X, \mathcal{B}, \mu, T)$ um automorfismo e $\alpha \in \mathcal{Z}$ um gerador para T . Então, $h_\mu(T) = h_\mu(T, \alpha)$.*

Demonstração: Fixe $\beta \in \mathcal{Z}$ arbitráriamente. Usando a Proposição 2.2.9 e a Proposição 2.2.14, obtemos

$$h_\mu(T, \beta) \leq h_\mu(T, \alpha_{-n}^n) + H_\mu(\beta | \alpha_{-n}^n) = h_\mu(T, \alpha) + H_\mu(\beta | \alpha_{-n}^n), \quad (2.6)$$

para todo $n \in \mathbb{N}$. Agora, pela Proposição 2.1.13(a) e pelo Corolário 2.2.7, vemos que

$$0 = H_\mu(\beta | \mathcal{B}) = H_\mu(\beta | \alpha_{-\infty}^{\infty}) = \lim_{n \rightarrow +\infty} H_\mu(\beta | \alpha_{-n}^n),$$

pois $\widehat{(\alpha_{-\infty}^{\infty})} = \mathcal{B} \pmod{\mu}$. Assim, fazendo $n \rightarrow +\infty$ em (2.6), concluímos que

$$h_\mu(T, \beta) \leq h_\mu(T, \alpha) + \lim_{n \rightarrow +\infty} H_\mu(\beta | \alpha_{-n}^n) = h_\mu(T, \alpha),$$

donde $h_\mu(T, \beta) \leq h_\mu(T, \alpha)$. Logo, $h_\mu(T) = h_\mu(T, \alpha)$, uma vez que $\beta \in \mathcal{Z}$ é arbitrário. $\square$

É mais difícil encontrar geradores fortes do que geradores. E, de fato, os únicos automorfismos com geradores fortes são os que têm entropia zero.

Corolário 2.2.17 *Todo automorfismo com um gerador forte tem entropia zero.*

Demonstração: Seja $(X, \mathcal{B}, \mu, T)$ um automorfismo, e seja $\alpha \in \mathcal{Z}$ um gerador forte para T . Pelo Teorema de Kolmogorov-sinai e o Teorema 2.2.8, temos $h_\mu(T) = h_\mu(T, \alpha) = H_\mu(\alpha | \alpha_1^\infty)$. Agora, usando o Lema 2.1.12(b) e o fato de que μ é T -invariante, obtemos que $H_\mu(\alpha | \alpha_1^\infty) = H_\mu(T\alpha | T\alpha_0^\infty) = H_\mu(T\alpha | \alpha_0^\infty) = H_\mu(T\alpha | \mathcal{B})$. Assim, obtemos que $h_\mu(T) = H_\mu(T\alpha | \mathcal{B}) = 0$, pois $T\alpha \subset \mathcal{B}$. $\square$

3 PARTIÇÕES MENSURÁVEIS

Seja $(X, \mathcal{B}, \mu)$ um espaço de probabilidade de Lebesgue completo. A noção de espaço de Lebesgue continua significando que $(X, \mathcal{B}, \mu)$ é isomorfo ao espaço de Lebesgue canônico do intervalo $[0, 1]$. A noção de completo refere-se a dizer que a σ -álgebra $\mathcal{B}$ é completa, isto é, se $A \in \mathcal{B}$ é um conjunto de medida nula e $B \subseteq A$ então $B \in \mathcal{B}$.

Enfatizamos que, embora não escrito de forma explícita, todas as propriedades que estabeleceremos entre partições e σ -álgebras são verdadeiras sempre módulo μ . Por exemplo, dadas duas σ -álgebras $\mathcal{F}$ e $\mathcal{G}$, escreveremos simplesmente $\mathcal{F} \subseteq \mathcal{G}$ para expressar que $\mathcal{F} \subseteq \mathcal{G} \pmod{\mu}$, isto é, a notação $\mathcal{F} \subseteq \mathcal{G}$ significa que dado $F \in \mathcal{F}$ existe $G \in \mathcal{G}$ tal que $F = G \pmod{\mu}$. Em particular, escreveremos que $\mathcal{F} = \mathcal{G}$ quando $\mathcal{F} \subseteq \mathcal{G} \pmod{\mu}$ e $\mathcal{G} \subseteq \mathcal{F} \pmod{\mu}$.

Neste capítulo estaremos seguindo [5], [8], [2, Capítulos 1, 2] e [1] como material base de referência.

3.1 Partições mensuráveis e σ -álgebras

Relembre a Definição 2.1.1 de uma partição de X . Como anteriormente, sempre assumiremos que os elementos de uma partição são mensuráveis. A partir de agora, lidaremos com partições possivelmente não-enumeráveis. Um elemento $A \in \xi$ continua sendo chamado de átomo de ξ . Como no caso enumerável, escrevemos $\xi \leq \eta$ (e dizemos que η é mais fina que ξ ou que ξ é mais grossa que η) se cada átomo $A_\xi \in \xi$ é uma união, módulo μ , de átomos de η . Dizemos que duas partições ξ e η são iguais, e escrevemos $\xi = \eta$, quando $\xi \leq \eta$ e $\eta \leq \xi$.

Definição 3.1.1 (Conjuntos ξ -mensuráveis) *Seja ξ uma partição de X . Os conjuntos $B \in \mathcal{B}$ que são uniões de elementos $A \in \xi$ são chamados de ξ -mensuráveis.*

Definição 3.1.2 (Partição mensurável) *Uma partição ξ é dita mensurável se existe uma família enumerável de conjuntos ξ -mensuráveis $\{B_n\}_{n \geq 1}$ com a seguinte propriedade separadora: dados $A, B \in \xi$ com $A \neq B$, existe $n \geq 1$ tal que $A \subseteq B_n$ e $B \subseteq X \setminus B_n$ ou $A \subseteq X \setminus B_n$ e $B \subseteq B_n$. Neste caso, a família $\{B_n\}_{n \geq 1}$ é chamada de família separadora para ξ .*

Denotamos por $\mathcal{M}$ o conjunto das partições mensuráveis. Segue-se diretamente da definição que qualquer partição finita ou enumerável é uma partição mensurável. Embora as partições finitas ou enumeráveis possuem átomos com medida positiva, existem partições mensuráveis em que cada um de seus átomos tem medida nula. Vejamos um exemplo com tal propriedade. Seja $(X, \mathcal{B}, \mu)$ o espaço de Lebesgue canônico no quadrado $[0, 1]^2$ e seja ξ a partição de X em linhas verticais. Note que os átomos de ξ tem medida nula para a medida de Lebesgue μ . Além disso, observe que os conjuntos ξ -mensuráveis são da forma $A \times [0, 1]$, onde $A \subseteq [0, 1]$ é um boreliano. Considere a família

$$\mathcal{A} := \left\{ \left[\frac{j}{2^n}, \frac{j+1}{2^n} \right] \times [0, 1] : n \in \mathbb{N} \text{ e } 0 \leq j \leq 2^n - 1 \right\}.$$

Claramente, $\mathcal{A}$ é enumerável e separa os elementos de ξ . Assim, vemos que $\xi \in \mathcal{M}$.

Definição 3.1.3 *Dada uma família de partições mensuráveis $\{\xi_\lambda\}_{\lambda \in \Lambda}$, seja $\xi := \bigvee_{\lambda \in \Lambda} \xi_\lambda$ a partição mensurável unicamente definida pelas seguintes propriedades:*

- (i) $\xi_\lambda \leq \xi$ para todo $\lambda \in \Lambda$;
- (ii) se $\eta \in \mathcal{M}$ é tal que $\xi_\lambda \leq \eta$ para todo $\lambda \in \Lambda$, então $\xi \leq \eta$.

Assim, por definição, dada uma família de partições mensuráveis $\{\xi_\lambda\}_{\lambda \in \Lambda}$, temos que $\xi = \bigvee_{\lambda \in \Lambda} \xi_\lambda$ é a partição mensurável mais grossa que refina todas as partições ξ_λ . Além disso, não é difícil provar que, quando $\Lambda \subset \mathbb{N}$, vale $\xi = \{\bigcap_{\lambda \in \Lambda} A_\lambda : A_\lambda \in \xi_\lambda\}$.

Definição 3.1.4 *Dada uma família de partições mensuráveis $\{\xi_\lambda\}_{\lambda \in \Lambda}$, seja $\xi := \bigwedge_{\lambda \in \Lambda} \xi_\lambda$ a partição mensurável unicamente definida pelas seguintes propriedades:*

- (i) $\xi \leq \xi_\lambda$ para todo $\lambda \in \Lambda$;
- (ii) se $\eta \in \mathcal{M}$ é tal que $\eta \leq \xi_\lambda$ para todo $\lambda \in \Lambda$, então $\eta \leq \xi$.

Assim, por definição, dada uma família de partições mensuráveis $\{\xi_\lambda\}_{\lambda \in \Lambda}$ temos que $\xi = \bigwedge_{\lambda \in \Lambda} \xi_\lambda$ é a partição mensurável mais fina que engrossa todas as partições ξ_λ .

Sejam $\{\xi_n\}_{n \in \mathbb{N}} \subset \mathcal{M}$ e $\xi \in \mathcal{M}$. Usaremos a notação $\xi_n \uparrow \xi$ para expressar que $\xi_1 \leq \xi_2 \leq \dots \leq \xi$ e que ξ é a partição mais grossa que refina todas as partições ξ_n (isto é, $\xi = \bigvee_{n \in \mathbb{N}} \xi_n$). De modo análogo, usaremos a notação $\xi_n \downarrow \xi$ para expressar que $\xi \leq \dots \leq \xi_2 \leq \xi_1$ e que ξ é a partição mais fina que engrossa todas as partições ξ_n (isto é, $\xi = \bigwedge_{n \in \mathbb{N}} \xi_n$).

Se $\{B_n\}_{n \geq 1}$ é uma família separadora para a partição mensurável ξ e $\beta_n = \{B_n, X \setminus B_n\}$, então a partição $\zeta_n = \bigvee_{i=1}^n \beta_i$ forma uma sequência crescente tal que $\bigvee_{n \geq 1} \zeta_n =$

ξ . Assim, para qualquer partição mensurável ξ existe uma sequência de partições finitas ζ_n tal que $\zeta_n \uparrow \xi$.

Seja $\xi \in \mathcal{M}$. Defina a aplicação $\pi : X \rightarrow \xi$ pondo $\pi(x) = \xi(x)$, onde $\xi(x)$ é o átomo de ξ que contém o ponto x . Como ξ é uma partição, a aplicação π está bem definida. Essa aplicação induz em ξ uma estrutura de espaço de probabilidade, onde $\mathcal{B}_\xi = \{C \subseteq \xi : \pi^{-1}(C) \in \mathcal{B}(\text{mod } \mu)\}$ é a σ -álgebra e $\mu_\xi = \mu \circ \pi^{-1}$ é a probabilidade induzida por π em ξ respectivamente. Dado $Q \in \mathcal{B}_\xi$, note que $\pi^{-1}(Q)$ é a união módulo μ de todos os elementos $A \in \xi$ que estão em Q . Assim, é fácil ver que as definições de $\mathcal{B}_\xi$ e μ_ξ são coerentes, isto é, $\mathcal{B}_\xi$ é uma σ -álgebra em ξ e μ_ξ uma probabilidade em ξ .

O teorema abaixo, que é um resultado clássico da Teoria da Medida, garante que o espaço de probabilidade $(\xi, \mathcal{B}_\xi, \mu_\xi)$ é de Lebesgue.

Teorema 3.1.5 (Rokhlin) *Se $\xi \in \mathcal{M}$ então $(\xi, \mathcal{B}_\xi, \mu_\xi)$ é um espaço de Lebesgue.*

A σ -álgebra $\mathcal{B}_\xi$ permite considerar operações teóricas definidas entre átomos de ξ , e μ_ξ permite escrever “quase todos os átomos de ξ ”. Se ξ for finita ou enumerável então $(\xi, \mathcal{B}_\xi, \mu_\xi)$ é um espaço discreto. Agora, se ξ é a partição do quadrado $[0, 1]^2$ em linhas verticais, então $(\xi, \mathcal{B}_\xi, \mu_\xi)$ é o espaço de Lebesgue canônico do intervalo $[0, 1]$.

Existem definições que são análogas às Definições 3.1.3 e 3.1.4 para sub- σ -álgebras de $\mathcal{B}$.

Dada uma família $\{F_\lambda\}_{\lambda \in \Lambda} \subseteq \mathcal{B}$, denote por $\sigma(\{F_\lambda\}_{\lambda \in \Lambda})$ a σ -álgebra completa mais grossa que contém $\bigcup_{\lambda \in \Lambda} F_\lambda$, ou seja, $\sigma(\{F_\lambda\}_{\lambda \in \Lambda})$ é a interseção de todas as σ -álgebras completas que contém a família $\{F_\lambda\}_{\lambda \in \Lambda}$. Em outras palavras, $\sigma(\{F_\lambda\}_{\lambda \in \Lambda})$ é o complemento do corpo de Borel gerado por todos os F_λ , $\lambda \in \Lambda$.

Seja $\{\mathcal{F}_\lambda\}_{\lambda \in \Lambda}$ uma família de sub- σ -álgebras de $\mathcal{B}$. A soma $\bigvee_{\lambda \in \Lambda} \mathcal{F}_\lambda$ é definida como a menor σ -álgebra completa que contém todas as σ -álgebras $\mathcal{F}_\lambda$, isto é, que satisfaz:

- (i) $\mathcal{F}_\lambda \subseteq \mathcal{F}$ para toda $\lambda \in \Lambda$;
- (ii) se $\mathcal{G} \subseteq \mathcal{B}$ é tal que $\mathcal{F}_\lambda \subseteq \mathcal{G}$ para toda $\lambda \in \Lambda$, então $\mathcal{F} \subseteq \mathcal{G}$.

É claro que a interseção $\bigcap_{\lambda \in \Lambda} \mathcal{F}_\lambda$ é uma sub- σ -álgebra de $\mathcal{B}$. Ademais, se cada σ -álgebra $\mathcal{F}_\lambda$ é completa, então $\bigcap_{\lambda \in \Lambda} \mathcal{F}_\lambda$ é completa.

Escreveremos $\mathcal{F}_n \uparrow \mathcal{F}$ para indicar que $\mathcal{F}_1 \subseteq \mathcal{F}_2 \subseteq \dots \subseteq \mathcal{F}$ e $\mathcal{F} = \bigvee_{n \in \mathbb{N}} \mathcal{F}_n$. Analogamente, usaremos a notação $\mathcal{F}_n \downarrow \mathcal{F}$ para indicar que $\mathcal{F} \subseteq \dots \subseteq \mathcal{F}_2 \subseteq \mathcal{F}_1$ e $\mathcal{F} = \bigcap_{n \in \mathbb{N}} \mathcal{F}_n$.

Agora veremos que partições mensuráveis de X induzem sub- σ -álgebras completas de $\mathcal{B}$ e que sub- σ -álgebras de $\mathcal{B}$ induzem partições mensuráveis de X .

Dada uma partição mensurável ξ , seja $\hat{\xi} \subseteq \mathcal{B}$ a σ -álgebra completa que consiste de todos os conjuntos mensuráveis $A \in \mathcal{B}$ que coincide, módulo μ , com um conjunto ξ -mensurável. Em outras palavras, temos que

$$\hat{\xi} := \{A \in \mathcal{B} : \exists B \text{ } \xi\text{-mensurável com } \mu(A \Delta B) = 0\}.$$

É fácil ver que, de fato, $\hat{\xi}$ é uma σ -álgebra. Mostraremos que $\hat{\xi}$ é completa. Seja $A \in \hat{\xi}$ não-vazio com $\mu(A) = 0$. Então, existe um conjunto ξ -mensurável, digamos A_ξ , tal que $\mu(A \Delta A_\xi) = 0$. Como $\mu(A_\xi) \leq \mu(A \cap A_\xi) + \mu(A \Delta A_\xi)$, temos que $\mu(A_\xi) = 0$. Assim, dado qualquer conjunto $B \subseteq A$, temos que $\mu(B \Delta A_\xi) \leq \mu(A) + \mu(A_\xi) = 0$. Logo $B \in \hat{\xi}$, o que prova que $\hat{\xi}$ é completa. Agora, observe que sendo $\hat{\xi}$ uma sub- σ -álgebra de $\mathcal{B}$, vemos que $\hat{\xi}$ é gerada de forma enumerável. Mais explicitamente, se $\{B_n\}_{n \geq 1}$ é uma família separadora para ξ , temos que $\hat{\xi} = \sigma(\{B_n\}_{n \geq 1})$.

Seja $\mathcal{F} \subseteq \mathcal{B}$ uma σ -álgebra completa. Temos que $\mathcal{F}$ é gerada de forma enumerável (por ser uma sub- σ -álgebra de $\mathcal{B}$, que é gerada por uma família enumerável), digamos $\mathcal{F} = \sigma(\{F_n\}_{n \geq 1})$. Defina $\xi = \mathcal{F}^\wedge$ pondo

$$\xi(x) := \bigcap_{F_n \ni x} F_n.$$

Claramente $\xi(x) \in \mathcal{B}$ para todo $x \in X$. Além disso, ξ é de fato uma partição por ser o conjunto das classes de equivalência da relação

$$x \sim y \iff (x \in F_n \iff y \in F_n, \forall n \geq 1).$$

Além disso, $\{F_n\}_{n \geq 1}$ é uma família separadora para ξ pois:

- $F_n \cap \xi(x) \neq \emptyset \iff \xi(x) \subseteq F_n$; de fato, a implicação $\Leftarrow$ é clara, para a implicação $\Rightarrow$ note que se $y \in F_n \cap \xi(x)$ então $\xi(x) = \xi(y) = \bigcap_{F_k \ni y} F_k \subseteq F_n$.
- $\xi(x) \subseteq F_n \iff x \in F_n$; de fato, a implicação $\Rightarrow$ é clara, já a implicação $\Leftarrow$ segue-se do item anterior, pois $x \in F_n \Rightarrow F_n \cap \xi(x) \neq \emptyset \Rightarrow \xi(x) \subseteq F_n$.
- Portanto, $x \in F_n \iff F_n \cap \xi(x) \neq \emptyset \iff \xi(x) \subseteq F_n$.
- Se $\xi(x) \neq \xi(y)$, então existe $n \in \mathbb{N}$ tal que $x \in F_n$ e $y \notin F_n$, logo $x \in F_n$ e $y \in X \setminus F_n$.

A aplicação $\wedge$, definida acima, é uma bijeção entre $(\xi, \{F_n\}_{n \geq 1})$ e $(\mathcal{F}, \{F_n\}_{n \geq 1})$, portanto podemos pensar em partições mensuráveis como sub- σ -álgebras completas de

$\mathcal{B}$ e vice-versa. Observe que, assim como no caso enumerável, $\xi \leq \eta$ ($\xi, \eta \in \mathcal{M}$) se e somente se $\hat{\xi} \subseteq \hat{\eta}$. Em particular, $\xi \leq \eta$ ($\xi, \eta \in \mathcal{M}$) se e somente se $\eta(x) \subseteq \xi(x)$ para μ -q.t.p. $x \in X$. Por conveniência, em algumas ocasiões usaremos ξ para denotar tanto a partição mensurável quanto a sub- σ -álgebra completa de $\mathcal{B}$.

Seja $\{\xi_n\}_{n \in \mathbb{N}}$ uma família de partições mensuráveis. Observe que

$$\bigvee_{n \in \mathbb{N}} \hat{\xi}_n = \left(\bigvee_{n \in \mathbb{N}} \xi_n \right)^\wedge \quad \text{e} \quad \bigwedge_{n \in \mathbb{N}} \hat{\xi}_n = \bigcap_{n \in \mathbb{N}} \hat{\xi}_n.$$

Em particular, $\xi_n \uparrow \xi$ equivale a $\hat{\xi}_n \uparrow \hat{\xi}$ e $\xi_n \downarrow \xi$ equivale a $\hat{\xi}_n \downarrow \hat{\xi}$.

Podemos identificar qualquer sub- σ -álgebra de um espaço de probabilidade como um fator deste espaço.

Definição 3.1.6 *Sejam $(X, \mathcal{B}, \mu)$ e $(Y, \mathcal{C}, \nu)$ espaços de probabilidade. Dizemos que $(Y, \mathcal{C}, \nu)$ é um fator de $(X, \mathcal{B}, \mu)$ se existe uma aplicação mensurável sobrejetiva, não necessariamente invertível, $\pi : X \rightarrow Y$ tal que $\nu = \mu \circ \pi^{-1}$. Nesse caso, π é chamado de aplicação fator.*

Claramente, se $\mathcal{F} \subseteq \mathcal{B}$ é uma sub- σ -álgebra completa então $(X, \mathcal{F}, \mu)$ é um espaço de probabilidade de Lebesgue completo. Em particular, tem-se que $(X, \mathcal{F}, \mu)$ é um fator de $(X, \mathcal{B}, \mu)$, com a identidade como aplicação fator. Reciprocamente, seja $\pi : (X, \mathcal{B}, \mu) \rightarrow (Y, \mathcal{C}, \nu)$ um fator entre espaços de probabilidade de Lebesgue completos, então $\pi^{-1}\mathcal{C}$ é uma sub- σ -álgebra de $\mathcal{B}$ (que podemos tomar como completa). Assim, $(Y, \mathcal{C}, \nu)$ é isomorfo ao espaço $(X, \pi^{-1}(\mathcal{C}), \mu)$.

Agora, veremos que do ponto de vista dinâmico, partições mensuráveis induzem medidas condicionais nos átomos.

Teorema 3.1.7 (Rokhlin) *Seja $(X, \mathcal{B}, \mu)$ um espaço de probabilidade de Lebesgue completo, e $\xi \in \mathcal{M}$. Então, existe uma família $\{\mu_A^\xi\}_{A \in \xi}$ com as seguintes propriedades:*

- (a) μ_A^ξ é definida em uma σ -álgebra $\mathcal{B}_A$.
- (b) $(X, \mathcal{B}_A, \mu_A^\xi)$ é um espaço de probabilidade de Lebesgue.
- (c) Para qualquer $B \in \mathcal{B}$, $B \cap A \in \mathcal{B}_A$ para μ_ξ -q.t.p. $A \in \xi$, a função $A \in \xi \mapsto \mu_A^\xi(B \cap A)$ é mensurável, e

$$\mu(B) = \int_\xi \mu_A^\xi(B \cap A) d\mu_\xi(A).$$

Além disso, se $\{\nu_A\}_{A \in \xi}$ é outra família de probabilidades que satisfaz (a)-(c), então $\mu_A^\xi = \nu_A$ para μ_ξ -q.t.p. $A \in \xi$.

A família $\{\mu_A^\xi\}_{A \in \xi}$ é chamada de sistema de medidas condicionais de ξ . Por exemplo, se $(X, \mathcal{B}, \mu)$ é o espaço de Lebesgue canônico de $[0, 1]^2$ e ξ é a partição em linhas verticais, então $\mu_{\{x\} \times [0, 1]}^\xi$ é a medida de Lebesgue em $\{x\} \times [0, 1]$. Observe que a propriedade (c) é um teorema do tipo Fubini. Assim, as partições mensuráveis são partições em que o Teorema de Fubini é válido.

Se $\mathcal{F} = \widehat{\xi}$, denote $\mu_x^\mathcal{F} = \mu_{\xi(x)}^\xi$ e chame a família $\{\mu_x^\mathcal{F}\}_{x \in X}$ de sistema de medidas condicionais de $\mathcal{F}$. Usando esse ponto de vista, podemos dar uma interpretação alternativa para as medidas condicionais. Dada $f \in L^1(X, \mathcal{B}, \mu)$, seja $\mathbb{E}(f|\mathcal{F})$ a esperança condicional de f com respeito a $\mathcal{F}$. Observando que $\mathbb{E}(f|\mathcal{F})$ é constante em cada átomo $A \in \xi$, seja $\mu_A^\xi(f)$ o valor de $\mathbb{E}(f|\mathcal{F})$ em A . Fixado $A \in \xi$, a aplicação $f \in C_c(X, \mathcal{B}, \mu) \mapsto \mu_A^\xi(f) \in \mathbb{R}$ é um funcional linear positivo limitado. Então, pelo Teorema de representação de Riesz-Markov-Kakutani, existe uma única medida $\hat{\mu}_A^\xi$ em X tal que $\mu_A^\xi(f) = \int_X f d\hat{\mu}_A^\xi$ para toda $f \in C_c(X, \mathcal{B}, \mu)$. Além disso, como $\mu_A^\xi(1) = 1$, temos que $\hat{\mu}_A^\xi$ é uma probabilidade. Esta medida é necessariamente μ_A^ξ , pois $\mu_A^\xi(f) = \int_X f d\mu_A^\xi$ para todo $f \in C_c(X, \mathcal{B}, \mu)$ (condição (c) do Teorema 3.1.7) e $\mu_A^\xi(A) = 1$ (condições (a)-(b) do Teorema 3.1.7).

3.2 Entropia de uma partição mensurável

Seja $\mathcal{F} \subseteq \mathcal{B}$ uma σ -álgebra completa. Sabemos que existe $\xi \in \mathcal{M}$ tal que $\mathcal{F} = \widehat{\xi}$. Dado $x \in X$ denotaremos o átomo $\xi(x)$ simplesmente por $\mathcal{F}(x)$. Além disso, lembre-se que neste caso usamos a notação $\mu_x^\mathcal{F}$ para representar a medida condicional $\mu_{\xi(x)}^\xi$ fornecida pelo Teorema de Rokhlin.

Definição 3.2.1 *Sejam $\mathcal{F}, \mathcal{G} \subseteq \mathcal{B}$ σ -álgebras completas. A função informação condicional de $\mathcal{F}$ dado $\mathcal{G}$ é a função $I_\mu : X \rightarrow [0, +\infty]$ definida por*

$$I_\mu(\mathcal{F}|\mathcal{G})(x) := -\log \mu_x^\mathcal{G}(\mathcal{F}(x)).$$

Observe que quando ξ é uma partição enumerável a definição acima coincide com a que foi dada no Capítulo 2. De fato, se $\xi(x) = A$ então $\mu_x^\mathcal{G}(\mathcal{F}(x)) = \mu_x^\mathcal{G}(A) = \mathbb{E}(1_A|\mathcal{G})(x)$.

Claramente $I_\mu(\mathcal{F}|\mathcal{G})(x)$ pode ser infinito. Por exemplo, se $\mathcal{G} = \{\emptyset, X\}$ e $\mu(\mathcal{F}(x)) = 0$ para todo $x \in X$ então $I_\mu(\mathcal{F}|\mathcal{G})(x) = -\log(\mu(\mathcal{F}(x))) = +\infty$ para todo $x \in X$.

O mesmo também ocorre quando $\mathcal{F}$ e $\mathcal{G}$ não estão correlacionada, isto é, $\mu_x^{\mathcal{G}}(\mathcal{F}(x)) = 0$ em um conjunto de medida positiva. Por exemplo, se $(X, \mathcal{B}, \mu)$ é o espaço de Lebesgue canônico em $[0, 1]^2$, $\mathcal{F}$ é a σ -álgebra definida por linhas verticais e $\mathcal{G}$ é a σ -álgebra definida por linhas horizontais então $I_\mu(\mathcal{F}|\mathcal{G})(x) = +\infty$ para todo $x \in X$.

Lema 3.2.2 *Sejam $\mathcal{F}_n, \mathcal{F}, \mathcal{G} \subseteq \mathcal{B}$ σ -álgebras. Se $\mathcal{F}_n \uparrow \mathcal{F}$, então $I_\mu(\mathcal{F}_n|\mathcal{G}) \xrightarrow{n \rightarrow +\infty} I_\mu(\mathcal{F}|\mathcal{G})$ μ -q.t.p. Em particular, a aplicação $x \mapsto I_\mu(\mathcal{F}|\mathcal{G})(x)$ é mensurável.*

Demonstração: Como $\mathcal{F}_n \uparrow \mathcal{F}$ temos que $\mathcal{F}_n(x) \downarrow \mathcal{F}(x)$ para μ -q.t.p. $x \in X$ e, portanto, $\mu_x^{\mathcal{G}}(\mathcal{F}_n(x)) \downarrow \mu_x^{\mathcal{G}}(\mathcal{F}(x))$ para μ -q.t.p. $x \in X$. Assim, pelo Teorema da Convergência Monótona, vemos que

$$I_\mu(\mathcal{F}_n|\mathcal{G})(x) = \mu_x^{\mathcal{G}}(\mathcal{F}_n(x)) \xrightarrow{n \rightarrow +\infty} \mu_x^{\mathcal{G}}(\mathcal{F}(x)) = I_\mu(\mathcal{F}|\mathcal{G})(x)$$

para μ -q.t.p. $x \in X$, o que prova a primeira afirmação. A segunda afirmação segue da primeira tomando uma sequência de σ -álgebras finitas $\mathcal{F}_n \subseteq \mathcal{B}$ tal que $\mathcal{F}_n \uparrow \mathcal{F}$ e observando que cada $I_\mu(\mathcal{F}_n|\mathcal{G})$ é mensurável.

□

Definição 3.2.3 *Sejam $\mathcal{F}, \mathcal{G} \subseteq \mathcal{B}$ σ -álgebras completas. A entropia condicional de $\mathcal{F}$ dado $\mathcal{G}$ é*

$$H_\mu(\mathcal{F}|\mathcal{G}) := \int_X I_\mu(\mathcal{F}|\mathcal{G})(x) d\mu = \int_X -\log \mu_x^{\mathcal{G}}(\mathcal{F}(x)) d\mu \in [0, +\infty].$$

Em particular, a entropia de $\mathcal{F}$ é $H_\mu(\mathcal{F}) := H_\mu(\mathcal{F}|\{\emptyset, X\})$.

Novamente, $H_\mu(\mathcal{F}|\mathcal{G}) = +\infty$ a menos que $\mathcal{F}$ e $\mathcal{G}$ sejam correlacionadas. Além disso, se $\mu(\mathcal{F}(x)) = 0$ em um conjunto de medida positiva, então $H_\mu(\mathcal{F}) = +\infty$.

Note que, se ξ e η são partições mensuráveis tais que $\xi \leq \eta$ então $\mu_{\eta(x)}^\eta(\xi(x)) = 1$ para μ -q.t.p. $x \in X$ e, portanto, $H_\mu(\xi|\eta) = 0$. A recíproca também é verdadeira. De fato, assumindo que

$$0 = H_\mu(\xi|\eta) = \int_X -\log \mu_{\eta(x)}^\eta(\xi(x)) d\mu,$$

obtemos que $\mu_{\eta(x)}^\eta(\xi(x)) = 1$ para μ -q.t.p. $x \in X$. Logo, $\eta(x) \subseteq \xi(x)$ para μ -q.t.p. $x \in X$ e, portanto, $\xi \leq \eta$.

O próximo lema que enunciaremos, sem apresentar sua demonstração, é uma generalização do Corolário 2.2.7 para partições mensuráveis. Tal lema, que nomearemos de

Lema de Chung-Neveu, apresenta uma espécie de continuidade para a entropia condicional, no que diz respeito à convergência de partições mensuráveis.

Lema 3.2.4 (Chung-Neveu) *Sejam $\xi, \eta \in \mathcal{M}$ e $\{\xi_n\}_{n \in \mathbb{N}}$ uma família de partições mensuráveis. Temos que:*

- (a) *Se $\xi_n \uparrow \xi$ e $H_\mu(\eta|\xi_n) < +\infty$ para algum n , então $H_\mu(\eta|\xi_n) \xrightarrow{n \rightarrow +\infty} H_\mu(\eta|\xi)$.*
- (b) *Se $\xi_n \uparrow \xi$ então $H_\mu(\xi_n|\eta) \xrightarrow{n \rightarrow +\infty} H_\mu(\xi|\eta)$.*
- (c) *Se $\xi_n \downarrow \xi$ então $H_\mu(\eta|\xi_n) \xrightarrow{n \rightarrow +\infty} H_\mu(\eta|\xi)$.*

Demonstração: Ver [5, Teorema 5.5].

□

Proposição 3.2.5 *Sejam $\mathcal{F}, \mathcal{G}, \mathcal{H} \subseteq \mathcal{B}$ σ -álgebras. Temos que:*

- (a) $H_\mu(\mathcal{F} \vee \mathcal{G}|\mathcal{H}) = H_\mu(\mathcal{G}|\mathcal{H}) + H_\mu(\mathcal{F}|\mathcal{G} \vee \mathcal{H})$. Em particular, $H_\mu(\mathcal{F} \vee \mathcal{G}) = H_\mu(\mathcal{G}) + H_\mu(\mathcal{F}|\mathcal{G})$.
- (b) *Se $\mathcal{F} \subseteq \mathcal{G}$, então:*
 - (i) $H_\mu(\mathcal{F}|\mathcal{H}) \leq H_\mu(\mathcal{G}|\mathcal{H})$;
 - (ii) $H_\mu(\mathcal{H}|\mathcal{F}) \geq H_\mu(\mathcal{H}|\mathcal{G})$.

Demonstração: Provaremos apenas o item (a), pois o item (b) segue por um argumento semelhante ao usado no item (a).

Inicialmente, assuma que $\mathcal{G}$ é finita e tome uma sequência de σ -álgebras finitas $\mathcal{F}_n \subset \mathcal{B}$ tal que $\mathcal{F}_n \uparrow \mathcal{F}$. Pela Proposição 2.1.15(b), temos $H_\mu(\mathcal{F}_n \vee \mathcal{G}|\mathcal{H}) = H_\mu(\mathcal{G}|\mathcal{H}) + H_\mu(\mathcal{F}_n|\mathcal{H} \vee \mathcal{G})$ para todo $n \in \mathbb{N}$. Assim, pelo Lema 3.2.4(b), obtemos que $H_\mu(\mathcal{F} \vee \mathcal{G}|\mathcal{H}) = H_\mu(\mathcal{G}|\mathcal{H}) + H_\mu(\mathcal{F}|\mathcal{H} \vee \mathcal{G})$.

Agora, suponha que $\mathcal{G}$ é arbitrária e tome uma sequência de σ -álgebras finitas $\mathcal{G}_n \subset \mathcal{B}$ tal que $\mathcal{G}_n \uparrow \mathcal{G}$. Como cada $\mathcal{G}_n$ é finita, temos que $H_\mu(\mathcal{F} \vee \mathcal{G}_n|\mathcal{H}) = H_\mu(\mathcal{G}_n|\mathcal{H}) + H_\mu(\mathcal{F}|\mathcal{H} \vee \mathcal{G}_n)$ para todo $n \in \mathbb{N}$. Assim, usando os itens (a) e (b) do Lema 3.2.4, concluímos que $H_\mu(\mathcal{F} \vee \mathcal{G}|\mathcal{H}) = H_\mu(\mathcal{G}|\mathcal{H}) + H_\mu(\mathcal{F}|\mathcal{H} \vee \mathcal{G})$.

□

Já sabemos que $H_\mu(\xi|\zeta) \leq H_\mu(\xi|\eta) + H_\mu(\eta|\zeta)$ quaisquer que sejam as partições $\xi, \eta, \zeta \in \mathcal{Z}$. Pela Proposição 3.2.5, usando o mesmo argumento do caso enumerável, vemos que essa propriedade ainda é verdadeira quando $\xi, \eta, \zeta \in \mathcal{M}$.

Seja $(X, \mathcal{B}, \mu, T)$ um automorfismo, onde $(X, \mathcal{B}, \mu)$ é um espaço de probabilidade de Lebesgue completo. Dados $m, n \in \mathbb{N} \cup \{\pm\infty\}$ e uma partição $\xi \in \mathcal{M}$, defina a partição mensurável ξ_m^n pela seguinte igualdade

$$\xi_m^n := \bigvee_{i=m}^n T^{-i}\xi,$$

onde $T^{-i}\xi = \{T^{-i}A : A \in \xi\}$. Como no caso enumerável, diremos que uma partição $\xi \in \mathcal{M}$ é geradora para a transformação T se $(\xi_{-\infty}^{+\infty})^\wedge = \mathcal{B}$.

Dadas $\xi, \eta \in \mathcal{M}$ arbitrárias de X , afirmamos que $H_\mu(T^{-1}\xi|T^{-1}\eta) = H_\mu(\xi|\eta)$. Com efeito, usando a unicidade do Teorema 3.1.7, obtemos que $\mu_{T^{-1}A}^{T^{-1}\xi} = \mu_A^\xi \circ T$ para todo átomo $A \in \xi$. Assim, para quase todo ponto $x \in X$, temos

$$\mu_{(T^{-1}\xi)(x)}^{T^{-1}\xi}((T^{-1}\eta)(x)) = \mu_{\xi(Tx)}^\xi(\eta(Tx)),$$

pois $T((T^{-1}\eta)(x)) = \eta(Tx)$. Daí, vemos que $I_\mu(T^{-1}\xi|T^{-1}\eta) = I_\mu(\xi|\eta) \circ T$ e, pela invariância de μ por T , concluímos que $H_\mu(T^{-1}\xi|T^{-1}\eta) = H_\mu(\xi|\eta)$ como afirmado. De modo análogo, usando que $T = (T^{-1})^{-1}$, concluímos que $H_\mu(T\xi|T\eta) = H_\mu(\xi|\eta)$. Em geral, dado um número inteiro $n \neq 0$, vale que $H_\mu(T^n\xi|T^n\eta) = H_\mu(\xi|\eta)$. Esta observação é extremamente útil na demonstração do lema abaixo.

Lema 3.2.6 *Sejam $(X, \mathcal{B}, \mu, T)$ um automorfismo e $\xi, \eta \in \mathcal{M}$.*

(a) *Se $\eta \leq \xi$ e $H_\mu(\xi|\eta_1^\infty) < +\infty$, então $\frac{1}{n}H_\mu(\xi_{-n}^0|\eta_1^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\xi|\xi_1^\infty)$.*

(b) *Se $\xi \leq \eta$ e $H_\mu(\eta|\xi_1^\infty) < +\infty$, então $\frac{1}{n}H_\mu(\xi_{-n}^0|\eta_1^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\xi|\xi_1^\infty)$.*

Em particular, se $H_\mu(\xi|\xi_1^\infty) < +\infty$, então $\frac{1}{n}H_\mu(\xi_{-n}^0|\xi_1^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\xi|\xi_1^\infty)$.

Demonstração: (a) Como $\xi_{-n}^0 = \xi_{-(n-1)}^0 \vee T^n\xi$, pela Proposição 3.2.5(b), obtemos

$$\begin{aligned} H_\mu(\xi_{-n}^0|\eta_1^\infty) &= H_\mu(\xi_{-(n-1)}^0|\eta_1^\infty) + H_\mu(T^n\xi|\xi_{-(n-1)}^0 \vee \eta_1^\infty) \\ &= H_\mu(\xi_{-(n-1)}^0|\eta_1^\infty) + H_\mu(T^n\xi|T^n(\xi_1^n \vee \eta_{n+1}^\infty)) \\ &= H_\mu(\xi_{-(n-1)}^0|\eta_1^\infty) + H_\mu(\xi|\xi_1^n \vee \eta_{n+1}^\infty). \end{aligned}$$

Seguindo esse processo recursivamente, obtemos que

$$\frac{1}{n}H_\mu(\xi_{-n}^0|\eta_1^\infty) = \frac{1}{n}H_\mu(\xi|\eta_1^\infty) + \frac{1}{n} \sum_{k=1}^n H_\mu(\xi|\xi_1^k \vee \eta_{k+1}^\infty). \quad (3.1)$$

Agora, como $\eta_1^\infty \leq \xi_1^k \vee \eta_{k+1}^\infty$ para todo $k \in \mathbb{N}$, temos que $H_\mu(\xi|\xi_1^k \vee \eta_{k+1}^\infty) < +\infty$ para todo $k \in \mathbb{N}$. Além disso, usando que $\xi_1^k \vee \eta_{k+1}^\infty \uparrow \xi_1^\infty$ e o Lema de Chung-Neveu, concluímos que

a sequência $\{H_\mu(\xi|\xi_1^k \vee \eta_{k+1}^\infty)\}$ converge para $H_\mu(\xi|\xi_1^\infty)$ e, por consequência, obtemos

$$\frac{1}{n} \sum_{k=1}^n H_\mu(\xi|\xi_1^k \vee \eta_{k+1}^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\xi|\xi_1^\infty).$$

Assim, fazendo $n \rightarrow \infty$ na igualdade (3.1), obtemos que $\frac{1}{n} H_\mu(\xi_{-n}^0|\eta_1^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\xi|\xi_1^\infty)$.

Em particular, se $H_\mu(\xi|\xi_1^\infty) < +\infty$, então $\frac{1}{n} H_\mu(\xi_{-n}^0|\xi_1^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\xi|\xi_1^\infty)$.

(b) Pelo item (a), temos que $\limsup \frac{1}{n} H_\mu(\xi_{-n}^0|\eta_1^\infty) \leq H_\mu(\xi|\xi_1^\infty)$, pois $\frac{1}{n} H_\mu(\xi_{-n}^0|\eta_1^\infty) \leq \frac{1}{n} H_\mu(\xi_{-n}^0|\xi_1^\infty)$ para todo $n \in \mathbb{N}$ e $H_\mu(\xi|\xi_1^\infty) \leq H_\mu(\eta|\xi_1^\infty) < +\infty$. Por outro lado, observe que a sequência não-negativa $\{\frac{1}{n} H_\mu(\eta_{-n}^0|\xi_{-n}^0 \vee \eta_1^\infty)\}_{n \geq 1}$ é limitada superiormente pela sequência $\{\frac{1}{n} H_\mu(\eta_{-n}^0|\xi_1^\infty)\}_{n \geq 1}$, que pelo item (a), converge para $H_\mu(\eta|\eta_1^\infty) < +\infty$. Pela Proposição 3.2.5, vemos que $H_\mu(\xi_{-n}^0|\eta_1^\infty) = H_\mu(\eta_{-n}^0|\eta_1^\infty) - H_\mu(\eta_{-n}^0|\xi_{-n}^0 \vee \eta_1^\infty)$ para todo n . Agora, usando repetidas vezes o item (a), obtemos

$$\begin{aligned} \liminf \frac{1}{n} H_\mu(\xi_{-n}^0|\eta_1^\infty) &= \liminf \left[\frac{1}{n} H_\mu(\eta_{-n}^0|\eta_1^\infty) - \frac{1}{n} H_\mu(\eta_{-n}^0|\xi_{-n}^0 \vee \eta_1^\infty) \right] \\ &= H_\mu(\eta|\eta_1^\infty) - \limsup \frac{1}{n} H_\mu(\eta_{-n}^0|\xi_{-n}^0 \vee \eta_1^\infty) \\ &\geq H_\mu(\eta|\eta_1^\infty) - \limsup \frac{1}{n} H_\mu(\eta_{-n}^0|\xi_{-n}^0 \vee \xi_1^\infty) \\ &= \liminf \left[\frac{1}{n} H_\mu(\eta_{-n}^0|\xi_1^\infty) - \frac{1}{n} H_\mu(\eta_{-n}^0|\xi_{-n}^0 \vee \xi_1^\infty) \right] \\ &= \liminf \frac{1}{n} H_\mu(\xi_{-n}^0|\xi_1^\infty) \\ &= H_\mu(\xi|\xi_1^\infty). \end{aligned}$$

Portanto, $\frac{1}{n} H_\mu(\xi_{-n}^0|\eta_1^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\xi|\xi_1^\infty)$.

□

Proposição 3.2.7 *Sejam $(X, \mathcal{B}, \mu, T)$ um automorfismo e $\xi, \eta \in \mathcal{M}$ tais que $H_\mu(\xi \vee \eta|\xi_1^\infty) < +\infty$. Então,*

$$H_\mu(\xi \vee \eta|\xi_1^\infty \vee \eta_1^\infty) = H_\mu(\xi|\xi_1^\infty) + H_\mu(\eta|\eta_1^\infty \vee \xi_{-\infty}^\infty).$$

Demonstração: Note que $H_\mu(\xi \vee \eta|\xi_1^\infty \vee \eta_1^\infty) \leq H_\mu(\xi \vee \eta|\xi_1^\infty) < +\infty$. Pela parte final do Lema 3.2.6, temos

$$\begin{aligned} H_\mu(\xi \vee \eta|\xi_1^\infty \vee \eta_1^\infty) &= \lim \frac{1}{n} H_\mu(\xi_{-n}^0 \vee \eta_{-n}^0|\xi_1^\infty \vee \eta_1^\infty) \\ &= \lim \frac{1}{n} H_\mu(\xi_{-n}^0|\xi_1^\infty \vee \eta_1^\infty) + \lim \frac{1}{n} H_\mu(\eta_{-n}^0|\xi_{-n}^\infty \vee \eta_1^\infty). \end{aligned}$$

Usando o Lema 3.2.6(b), vemos que $\frac{1}{n}H_\mu(\xi_{-n}^0|\xi_1^\infty \vee \eta_1^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\xi|\xi_1^\infty)$. Resta mostrar que $\frac{1}{n}H_\mu(\eta_{-n}^0|\xi_{-n}^\infty \vee \eta_1^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\eta|\eta_1^\infty \vee \xi_{-\infty}^\infty)$. E de fato, fixado $n \geq 1$, temos $\eta_{-n}^0 = \eta \vee \eta_{-n}^{-1}$. Daí, vemos que

$$\begin{aligned} H_\mu(\eta_{-n}^0|\eta_1^\infty \vee \xi_{-n}^\infty) &= H_\mu(\eta|\eta_1^\infty \vee \xi_{-n}^\infty) + H_\mu(\eta_{-n}^{-1}|\eta_0^\infty \vee \xi_{-n}^\infty) \\ &= H_\mu(\eta|\eta_1^\infty \vee \xi_{-n}^\infty) + H_\mu(\eta_{-(n-1)}^0|\eta_1^\infty \vee \xi_{-(n-1)}^\infty). \end{aligned}$$

Por indução, obtemos que

$$\frac{1}{n}H_\mu(\eta_{-n}^0|\eta_1^\infty \vee \xi_{-n}^\infty) = \frac{1}{n} \sum_{k=0}^n H_\mu(\eta|\eta_1^\infty \vee \xi_{-k}^\infty).$$

Agora, observe que $H_\mu(\eta|\eta_1^\infty \vee \xi_{-k}^\infty) \leq H_\mu(\eta|\xi_1^\infty) < \infty$ para todo $k \geq 0$. Assim, pelo Lema 3.2.4(a), obtemos que a sequência de números reais $\{H_\mu(\eta|\eta_1^\infty \vee \xi_{-k}^\infty)\}_{k \geq 1}$ converge para $H_\mu(\eta|\eta_1^\infty \vee \xi_{-\infty}^\infty)$ e, portanto, suas médias Cesàro também convergem para o mesmo valor. Logo, concluímos que $\frac{1}{n}H_\mu(\eta_{-n}^0|\eta_1^\infty \vee \xi_{-n}^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\eta|\eta_1^\infty \vee \xi_{-\infty}^\infty)$, o que conclui a prova. $\square$

Corolário 3.2.8 *Sejam $(X, \mathcal{B}, \mu, T)$ um automorfismo e $\xi, \eta \in \mathcal{X}$. Então $h_\mu(T, \xi \vee \eta) = h_\mu(T, \xi) + H_\mu(\eta|\eta_1^\infty \vee \xi_{-\infty}^{+\infty})$. Em particular, $h_\mu(T, \xi \vee \eta) \leq h_\mu(T, \xi) + h_\mu(T, \eta)$.*

Demonstração: Pelo Teorema 2.2.8 e a Proposição 3.2.7, temos

$$\begin{aligned} h_\mu(T, \xi \vee \eta) &= H_\mu(\xi \vee \eta|\xi_1^\infty \vee \eta_1^\infty) \\ &= H_\mu(\xi|\xi_1^\infty) + H_\mu(\eta|\eta_1^\infty \vee \xi_{-\infty}^{+\infty}) \\ &= h_\mu(T, \xi) + H_\mu(\eta|\eta_1^\infty \vee \xi_{-\infty}^{+\infty}). \end{aligned}$$

Em particular, vemos que

$$h_\mu(T, \xi \vee \eta) \leq h_\mu(T, \xi) + h_\mu(T, \eta),$$

pois $H_\mu(\eta|\eta_1^\infty \vee \xi_{-\infty}^{+\infty}) \leq H_\mu(\eta|\eta_1^\infty) = h_\mu(T, \eta)$. $\square$

3.3 Fórmula de Abramov-Rokhlin

Sejam $(X, \mathcal{B}, \mu, T)$ um endomorfismo e $\xi \in \mathcal{M}$. Assuma que ξ é T -invariante, ou seja, que $T^{-1}\xi = \xi$ (alternativamente, a σ -álgebra $\widehat{\xi}$ é T -invariante).

Definição 3.3.1 *Seja $\alpha \in \mathcal{X}$. A entropia de T com respeito a α e condicionada a ξ é*

$$h_\mu(T, \alpha|\xi) := \lim_{n \rightarrow +\infty} \frac{1}{n} H_\mu(\alpha_0^{n-1}|\xi) = H_\mu(\alpha|\alpha_1^\infty \vee \xi).$$

Observe que, na definição acima, é crucial que ξ seja T -invariante. De fato, isso garante que

$$\begin{aligned} H_\mu(\alpha_0^{m+n-1}|\xi) &= H_\mu(\alpha_0^{m-1}|\xi) + H_\mu(\alpha_m^{m+n-1}|\xi \vee \alpha_0^{m-1}) \\ &\leq H_\mu(\alpha_0^{m-1}|\xi) + H_\mu(\alpha_m^{m+n-1}|\xi) \\ &= H_\mu(\alpha_0^{m-1}|\xi) + H_\mu(\alpha_0^{n-1}|\xi). \end{aligned}$$

Assim, pela subaditividade (Lema 2.2.3), o limite $h_\mu(T, \alpha|\xi)$ está bem definido. Usando um argumento análogo ao usado na prova do Teorema 2.2.8, concluímos que $\frac{1}{n} H_\mu(\alpha_0^{n-1}|\xi) = \frac{1}{n} \sum_{i=1}^{n-1} H_\mu(\alpha|\xi \vee \alpha_1^i)$ e, portanto, que $h_\mu(T, \alpha|\xi) = H(\alpha|\alpha_1^\infty \vee \xi)$.

Definição 3.3.2 *A entropia de T condicionada a ξ é*

$$h_\mu(T|\xi) := \sup\{h_\mu(T, \alpha|\xi) : \alpha \in \mathcal{X}\}.$$

Note que $h_\mu(T|\xi)$ é monótono sob refinamentos, ou seja, se $\xi \leq \eta$ então $h_\mu(T|\xi) \geq h_\mu(T|\eta)$. Com efeito, dados $\alpha \in \mathcal{X}$ e $n \geq 0$ qualquer, temos que $H_\mu(\alpha_0^{n-1}|\xi) \geq H_\mu(\alpha_0^{n-1}|\eta)$, donde vemos que $h_\mu(T, \alpha|\xi) \geq h_\mu(T, \alpha|\eta)$. Assim, pela definição de supremo, obtemos que $h_\mu(T|\xi) \geq h_\mu(T|\eta)$.

Definição 3.3.3 (Fator de um automorfismo) *Sejam $(X, \mathcal{B}, \mu, T)$ e $(Y, \mathcal{C}, \nu, S)$ automorfismos. Dizemos que $(Y, \mathcal{C}, \nu, S)$ é um fator de $(X, \mathcal{B}, \mu, T)$ se existe uma aplicação mensurável sobrejetiva, não necessariamente invertível, $\pi : X \rightarrow Y$ tal que $\pi \circ T = S \circ \pi$ e $\nu = \mu \circ \pi^{-1}$.*

O Teorema de Abramov-Rokhlin, apresentado abaixo, nos permite relacionar a entropia de um automorfismo e de seus fatores.

Teorema 3.3.4 (Abramov-Rokhlin) *Se $\pi : (X, \mathcal{B}, \mu, T) \rightarrow (Y, \mathcal{C}, \nu, S)$ é um fator entre automorfismos, então*

$$h_\mu(T) = h_\nu(S) + h_\mu(T|\pi^{-1}\mathcal{C}). \quad (3.2)$$

Demonstração: Inicialmente, observe que se η é uma partição de Y , então $\pi^{-1}\eta = \{\pi^{-1}(A) : A \in \eta\}$ é uma partição de X . Além disso, temos

$$H_\mu(\pi^{-1}(\eta)) = \sum_{A \in \eta} -\mu(\pi^{-1}(A)) \log \mu(\pi^{-1}(A)) = \sum_{A \in \eta} -\nu(A) \log \nu(A) = H_\nu(\eta).$$

Assim, η tem entropia finita se e somente se $\pi^{-1}\eta$ tem entropia finita. Fixe $n \in \mathbb{N}$ arbitrário. Então, usando que $T^{-i} \circ \pi^{-1} = (\pi \circ T^i)^{-1} = (S^i \circ \pi)^{-1} = \pi^{-1} \circ S^{-i}$ para todo i , obtemos

$$(\pi^{-1}\eta)_0^{n-1} = \bigvee_{i=0}^{n-1} T^{-i}(\pi^{-1}\eta) = \bigvee_{i=0}^{n-1} \pi^{-1}(S^{-i}\eta) = \pi^{-1}\eta_0^{n-1},$$

donde vemos que $\frac{1}{n}H_\mu((\pi^{-1}\eta)_0^{n-1}) = \frac{1}{n}H_\nu(\eta_0^{n-1})$ para todo n natural. Fazendo n tender para $+\infty$, concluímos que $h_\mu(T, \pi^{-1}\eta) = h_\nu(S, \eta)$. Agora, provaremos a igualdade (3.2) em duas etapas.

Etapla 1: $h_\mu(T) \geq h_\nu(S) + h_\mu(T|\pi^{-1}\mathcal{C})$.

Sejam ξ e η partições enumeráveis de entropia finita de X e Y , respectivamente. Por definição, $H_\mu(\xi|\xi_1^\infty \vee (\pi^{-1}\eta)_{-\infty}^{+\infty}) = h_\mu(T, \xi|(\pi^{-1}\eta)_{-\infty}^{+\infty})$. Assim, pelo Corolário 3.2.8, obtemos

$$h_\mu(T) \geq h_\mu(T, \xi \vee \pi^{-1}\eta) = h_\nu(S, \eta) + h_\mu(T, \xi|(\pi^{-1}\eta)_{-\infty}^{+\infty}),$$

pois $h_\mu(T, \pi^{-1}\eta) = h_\nu(S, \eta)$. Tomando o supremo em ξ , obtemos $h_\mu(T) \geq h_\nu(S, \eta) + h_\mu(T|(\pi^{-1}\eta)_{-\infty}^{+\infty})$. Como $(\pi^{-1}\eta)_{-\infty}^{+\infty} \leq \pi^{-1}\mathcal{C}$, temos $h_\mu(T) \geq h_\nu(S, \eta) + h_\mu(T|\pi^{-1}\mathcal{C})$. Por fim, tomando o supremo em η , concluímos que $h_\mu(T) \geq h_\nu(S) + h_\mu(T|\pi^{-1}\mathcal{C})$.

Etapla 2: $h_\mu(T) \leq h_\nu(S) + h_\mu(T|\pi^{-1}\mathcal{C})$.

Seja ξ uma partição enumerável de X com entropia finita. Tome uma sequência $(\eta_k)_{k \in \mathbb{N}}$ de partições finitas de Y tais que $\eta_k \uparrow \mathcal{C}$. Fixe $k \in \mathbb{N}$ arbitrário. Pelo corolário 3.2.8, temos

$$h_\mu(T, \xi) \leq h_\mu(T, \xi \vee \pi^{-1}\eta_k) = h_\nu(S, \eta_k) + H_\mu(\xi|\xi_1^\infty \vee (\pi^{-1}\eta_k)_{-\infty}^{+\infty}),$$

pois $h_\mu(T, \pi^{-1}\eta_k) = h_\nu(S, \eta_k)$. Em particular, temos que

$$h_\mu(T, \xi) \leq h_\nu(S) + H_\mu(\xi|\xi_1^\infty \vee (\pi^{-1}\eta_k)_{-\infty}^{+\infty}).$$

Passando o limite quando $k \rightarrow +\infty$, obtemos

$$h_\mu(T, \xi) \leq h_\nu(S) + H_\mu(\xi|\xi_1^\infty \vee \pi^{-1}\mathcal{C}),$$

pois $\pi^{-1}\eta_k \uparrow \pi^{-1}\mathcal{C}$ (na igualdade acima usamos o Corolário 2.2.7). Por fim, tomando o supremo em ξ , concluímos que $h_\mu(T) \leq h_\nu(S) + h_\mu(T|\pi^{-1}\mathcal{C})$.

□

4 O FATOR DE PINSKER

Neste capítulo apresentaremos os conceitos de Álgebra de Pinsker e Tail σ -álgebra de uma partição mensurável e, além disso, apresentaremos algumas relações entre tais conceito. Estes conceitos são ferramentas extremamente úteis para determinar se certos automorfismos são K -automorfismos. Esta discussão é baseada em [1], [5], [2, Chapter 2] e [8].

4.1 Álgebra de Pinsker

Definição 4.1.1 (Álgebra de Pinsker) *Dado um automorfismo $(X, \mathcal{B}, \mu, T)$, sua álgebra de Pinsker é o conjunto*

$$\mathcal{P}(T) := \{A \in \mathcal{B} : h_\mu(T, \{A, X \setminus A\}) = 0\}.$$

Dado um conjunto $A \in \mathcal{B}$, defina $\xi_A := \{A, X \setminus A\}$. Então, $A \in \mathcal{P}(T) \Leftrightarrow h_\mu(T, \xi_A) = 0 \Leftrightarrow H_\mu(\xi_A | (\xi_A)_1^\infty) = 0$. Alternativamente, $A \in \mathcal{P}(T)$ se, e somente se, as informações de que $T^n x \in A$ ou não para todo $n \geq 1$ determinam completamente se $x \in A$ ou não.

Observe que a álgebra de Pinsker contém todo conjunto $A \in \mathcal{B}$ que é T -invariante. Com efeito, se $A \in \mathcal{B}$ é T -invariante, temos $T^n \xi_A = \{T^n A, T^n(X \setminus A)\} = \{A, X \setminus A\} = \xi_A$ para todo n . Assim, como $H_\mu(\xi_A) < \infty$, temos

$$h_\mu(T, \xi_A) = \lim_n \frac{1}{n} H_\mu((\xi_A)_0^n) = \lim_n \frac{1}{n} H_\mu(\xi_A) = 0,$$

donde obtemos que $A \in \mathcal{P}(T)$.

Além disso, se $h_\mu(T, \xi) = 0$ então todo conjunto ξ -mensurável é um elemento da álgebra de Pinsker. De fato, se A é ξ -mensurável então $\xi_A \leq \xi$ e, portanto, $h_\mu(T, \xi_A) \leq h_\mu(T, \xi) = 0$.

A proposição abaixo caracteriza a álgebra de Pinsker, mostrando que ela é uma σ -álgebra e apresentando algumas de suas propriedades. O fator $(X, \mathcal{P}(T), \mu, T)$ é chamado de *Fator de Pinsker*.

Proposição 4.1.2 *Seja $(X, \mathcal{B}, \mu, T)$ um automorfismo. Então:*

- (a) $\mathcal{P}(T)$ é uma sub- σ -álgebra.
- (b) $\mathcal{P}(T)$ é T -invariante.

- (c) O fator $(X, \mathcal{P}(T), \mu, T)$ tem entropia nula.
- (d) $(X, \mathcal{P}(T), \mu, T)$ é o maior fator com entropia nula de $(X, \mathcal{B}, \mu, T)$.
- (e) $h_\mu(T|\mathcal{P}(T)) = h_\mu(T)$.
- (f) $\mathcal{P}(T^n) = \mathcal{P}(T), \forall n \in \mathbb{Z} \setminus \{0\}$.

Demonstração: (a) Observe que $\emptyset, X \in \mathcal{P}(T)$ (por serem T -invariantes) e, além disso, segue-se diretamente da definição de $\mathcal{P}(T)$ que $A \in \mathcal{P}(T) \Leftrightarrow X \setminus A \in \mathcal{P}(T)$. Resta provar apenas que $\mathcal{P}(T)$ é fechado para uniões enumeráveis. Dada $\{A_n\}_{n \geq 1}$ uma família de elementos de $\mathcal{P}(T)$, vamos mostrar que $A := \bigcup_{n \geq 1} A_n \in \mathcal{P}(T)$, ou seja, que $h_\mu(T, \xi_A) = 0$. Fixe $\varepsilon > 0$ arbitrário. Como $\xi_A \leq \bigvee_{n \geq 1} \xi_{A_n}$, temos que

$$0 = H_\mu(\xi_A | \bigvee_{n \geq 1} \xi_{A_n}) = \lim_n H_\mu(\xi_A | \bigvee_{i=1}^n \xi_{A_i}),$$

donde obtemos que

$$H_\mu(\xi_A | \bigvee_{i=1}^n \xi_{A_i}) < \varepsilon$$

para n suficientemente grande. Assim, para n suficientemente grande, temos

$$h_\mu(T, \xi_A) \leq h_\mu(T, \bigvee_{i=1}^n \xi_{A_i}) + H_\mu(\xi_A | \bigvee_{i=1}^n \xi_{A_i}) < \sum_{i=1}^n h_\mu(T, \xi_{A_i}) + \varepsilon = \varepsilon.$$

Como ε é arbitrário, concluímos que $h_\mu(T, \xi_A) = 0$. Logo, $A \in \mathcal{P}(T)$.

(b) De fato, dado $A \in \mathcal{P}(T)$ temos

$$T^{-1}\xi_A = \{T^{-1}A, X \setminus T^{-1}A\} = \xi_{T^{-1}A}$$

e, portanto, $h_\mu(T, \xi_{T^{-1}A}) = h_\mu(T, T^{-1}\xi_A) = h_\mu(T, \xi_A)$. Segue-se daí que $A \in \mathcal{P}(T)$ se e somente se $T^{-1}A \in \mathcal{P}(T)$. Assim, concluímos que $T^{-1}\mathcal{P}(T) = \mathcal{P}(T)$.

(c) Seja $\xi \in \mathcal{Z}$, com $\xi \subset \mathcal{P}(T)$. Pelo Teorema 2.2.11, podemos supor que ξ é finita, digamos $\xi = \{A_1, \dots, A_n\}$. Como $\xi \leq \bigvee_{i=1}^n \xi_{A_i}$, obtemos que

$$h_\mu(T, \xi) \leq h_\mu(T, \bigvee_{i=1}^n \xi_{A_i}) \leq \sum_{i=1}^n h_\mu(T, \xi_{A_i}) = 0.$$

Portanto, $(X, \mathcal{P}(T), \mu, T)$ tem entropia nula.

(d) Seja $\mathcal{F} \subseteq \mathcal{B}$ uma σ -álgebra tal que $(X, \mathcal{F}, \mu, T)$ tem entropia nula. Então, $h_\mu(T, \xi_A) = 0$ para todo $A \in \mathcal{F}$, donde concluímos que $\mathcal{F} \subseteq \mathcal{P}(T)$.

(e) Pelo Teorema de Abramov-Rokhlin (Teorema 3.3.4), temos

$$h_\mu(T) = h_\mu(T \upharpoonright_{\mathcal{P}(T)}) + h_\mu(T|\mathcal{P}(T)) = h_\mu(T|\mathcal{P}(T)),$$

donde concluimos que $h_\mu(T) = h_\mu(T|\mathcal{P}(T))$.

(f) Suponha que $n > 0$ (o caso $n < 0$ é análogo) e tome $\xi \in \mathcal{Z}$ arbitrária. Dado $n \in \mathbb{N}$, temos que $h_\mu(T^n, \bigvee_{i=0}^{n-1} T^{-i}\xi) = nh_\mu(T, \xi)$ (veja a prova da Proposição 2.2.13). Assim, como T preserva μ , temos

$$nh_\mu(T, \xi) = h_\mu(T^n, \bigvee_{i=0}^{n-1} T^{-i}\xi) \leq \sum_{i=0}^{n-1} h_\mu(T^n, T^{-i}\xi) = nh_\mu(T^n, \xi),$$

donde $h_\mu(T, \xi) \leq h_\mu(T^n, \xi)$. Por outro lado, note que

$$h_\mu(T^n, \xi) \leq h_\mu(T^n, \bigvee_{j=0}^{n-1} T^{-j}\xi) = nh_\mu(T, \xi),$$

donde $h_\mu(T^n, \xi) \leq nh_\mu(T, \xi)$. Agora, seja $\xi = \xi_A$. Então:

◦ se $A \in \mathcal{P}(T)$ então $h_\mu(T^n, \xi) \leq nh_\mu(T, \xi) = 0$, donde obtemos que $A \in \mathcal{P}(T^n)$.

◦ se $A \in \mathcal{P}(T^n)$ então $h_\mu(T, \xi) \leq h_\mu(T^n, \xi) = 0$, donde obtemos que $A \in \mathcal{P}(T)$.

Portanto, $\mathcal{P}(T^n) = \mathcal{P}(T), \forall n \in \mathbb{Z} \setminus \{0\}$.

□

Na demonstração do item (c) acima, vimos que se uma partição $\xi \in \mathcal{Z}$ é tal que $\xi \subset \mathcal{P}(T)$ então $h_\mu(T, \xi) = 0$. Assim, dado $\xi \in \mathcal{Z}$, temos que $h_\mu(T, \xi) = 0$ se e somente se $\xi \subset \mathcal{P}(T)$.

Proposição 4.1.3 *Se ξ é uma partição mensurável tal que $H_\mu(\xi|\xi_1^\infty) < +\infty$, então $H_\mu(\xi|\xi_1^\infty \vee \mathcal{P}(T)) = H_\mu(\xi|\xi_1^\infty)$.*

Demonstração: Inicialmente, considere $\eta \subset \mathcal{P}(T)$ uma partição enumerável com entropia finita. Então $H_\mu(\eta|\xi_1^\infty \vee \eta_1^\infty \vee \xi), H_\mu(\eta|\eta_1^\infty \vee \xi_{-\infty}^\infty) \leq H_\mu(\eta|\eta_1^\infty) = 0$. Como

$$H_\mu(\xi \vee \eta|\xi_1^\infty) \leq H_\mu(\xi|\xi_1^\infty) + H_\mu(\eta) < +\infty,$$

podemos aplicar a Proposição 3.2.7 e obter que

$$\begin{aligned} H_\mu(\xi|\xi_1^\infty \vee \eta_1^\infty) &= H_\mu(\xi|\xi_1^\infty \vee \eta_1^\infty) + H_\mu(\eta|\xi_1^\infty \vee \eta_1^\infty \vee \xi) \\ &= H_\mu(\xi \vee \eta|\xi_1^\infty \vee \eta_1^\infty) \\ &= H_\mu(\xi|\xi_1^\infty) + H_\mu(\eta|\eta_1^\infty \vee \xi_{-\infty}^\infty) \\ &= H_\mu(\xi|\xi_1^\infty), \end{aligned}$$

donde $H_\mu(\xi|\xi_1^\infty \vee \eta_1^\infty) = H_\mu(\xi|\xi_1^\infty)$ (a Proposição 3.2.7 foi utilizada na terceira igualdade). Agora, considere uma sequência de partições $\eta_n \in \mathcal{Z}$ tal que $\eta_n \uparrow \mathcal{P}(T)$. Então $H_\mu(\xi|\xi_1^\infty \vee (\eta_n)_1^\infty) = H_\mu(\xi|\xi_1^\infty)$ para todo $n \in \mathbb{N}$. Tomando o limite em $n \in \mathbb{N}$, segue do Corolário 2.2.7 que $H_\mu(\xi|\xi_1^\infty \vee \mathcal{P}(T)) = H_\mu(\xi|\xi_1^\infty)$, o que prova a proposição. $\square$

4.2 Tail σ -álgebras

Seja $(X, \mathcal{B}, \mu, T)$ um automorfismo.

Definição 4.2.1 *Seja $\xi \in \mathcal{M}$. A tail σ -álgebra de ξ é a σ -álgebra gerada pela partição mensurável $\bigwedge_{n \geq 0} \xi_n^\infty$. Um elemento da tail σ -álgebra de ξ é chamado de ξ -evento tail.*

Por conveniência, denotaremos tanto a σ -álgebra quanto a partição por $\text{Tail}(\xi)$. A partir da noção de tail σ -álgebra de uma partição, podemos definir o que é a tail σ -álgebra de $(X, \mathcal{B}, \mu, T)$.

Definição 4.2.2 *A tail σ -álgebra de $(X, \mathcal{B}, \mu, T)$ é a soma $\bigvee_{\xi \in \mathcal{Z}} \text{Tail}(\xi)$ de todas as tail σ -álgebras com $\xi \in \mathcal{Z}$.*

Essas noções foram introduzidas por Kolmogorov, no contexto da teoria da probabilidade do seguinte modo. Dadas as variáveis aleatórias $X_1, X_2, \dots$ em um espaço de probabilidade $(\Omega, \mathcal{B}, \mathbb{P})$, chamamos $A \in \mathcal{B}$ de um evento tail se A for independente da σ -álgebra $\sigma(X_1, X_2, \dots, X_n)$ para todo n natural, isto é, se para cada $n \in \mathbb{N}$ valer que $\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$ para todos $B \in \sigma(X_1, X_2, \dots, X_n)$.

Adiante, veremos algumas relações entre a tail σ -álgebra e a álgebra de Pinsker. Começaremos mostrando que $\mathcal{P}(T)$ é a tail σ -álgebra de T .

Proposição 4.2.3 *Seja $(X, \mathcal{B}, \mu, T)$ um automorfismo. Então:*

- (a) $\text{Tail}(\xi) \subseteq \mathcal{P}(T)$ para toda partição $\xi \in \mathcal{Z}$.
- (b) Se $\xi \in \mathcal{Z}$ gera $\mathcal{B}$, então $\text{Tail}(\xi) = \mathcal{P}(T)$.
- (c) $\mathcal{P}(T) = \bigvee_{\xi \in \mathcal{Z}} \text{Tail}(\xi)$.

Demonstração: (a) Seja $\xi \in \mathcal{Z}$ qualquer. Fixe $A \in \text{Tail}(\xi)$ qualquer e defina $\eta = \xi_A$. Vamos provar que $h_\mu(T, \eta) = 0$. Primeiro, afirmamos que $\eta_{-\infty}^{+\infty} \leq \xi_1^\infty$. De fato, observe que:
 o Se $n \geq 0$ então $\eta \leq \xi_{n+1}^\infty = T^{-n}(\xi_1^\infty) \Rightarrow T^n(\eta) \leq \xi_1^\infty$.

◦ Se $n \geq 1$ então $\eta \leq \xi_{1-n}^\infty = T^n(\xi_1^\infty) \Rightarrow T^{-n}(\eta) \leq \xi_1^\infty$.

Portanto $\eta_{-\infty}^{+\infty} \leq \xi_1^\infty$, como afirmamos. Assim, pelo Corolário 3.2.8, obtemos:

$$h_\mu(T, \xi \vee \eta) = h_\mu(T, \xi) + H_\mu(\eta | \eta_1^\infty \vee \xi_{-\infty}^{+\infty}) = h_\mu(T, \xi),$$

pois $\eta \leq \xi_1^\infty \leq \eta_1^\infty \vee \xi_{-\infty}^{+\infty}$. Por outro lado, usando novamente o Corolário 3.2.8 e a inclusão $\eta_{-\infty}^{+\infty} \leq \xi_1^\infty$, obtemos

$$\begin{aligned} h_\mu(T, \xi \vee \eta) &= h_\mu(T, \eta) + H_\mu(\xi | \xi_1^\infty \vee \eta_{-\infty}^{+\infty}) \\ &= h_\mu(T, \eta) + H_\mu(\xi | \xi_1^\infty) \\ &= h_\mu(T, \eta) + h_\mu(T, \xi). \end{aligned}$$

Comparando as duas igualdades, segue que $h_\mu(T, \eta) = 0$.

(b) Suponha que $\xi \in \mathcal{Z}$ é uma partição que gera $\mathcal{B}$, isto é, $(\xi_{-\infty}^{+\infty})^\wedge = \mathcal{B}$. Inicialmente, observe que pelo item anterior é suficiente provar que $\mathcal{P}(T) \subseteq \text{Tail}(\xi)$. Afirmamos que é suficiente provar que $A \in \mathcal{P}(T) \Rightarrow \xi_A \leq \xi_0^\infty$. De fato, se isto vale então, para cada $A \in \mathcal{P}(T)$, temos que $\xi_A = T^{-n}\xi_{T^n A} \leq T^{-n}\xi_0^\infty = \xi_n^\infty$ para todo $n \geq 0$ (pois $T^n A \in \mathcal{P}(T)$ para todo $n \geq 0$), donde vemos que $\xi_A \leq \text{Tail}(\xi)$. Assim, concluimos que $\xi_A \subseteq \text{Tail}(\xi)$ e, em particular, que $A \in \text{Tail}(\xi)$. Fixe $A \in \mathcal{P}(T)$ arbitrário e tome $\eta = \xi_A$. Expressaremos $H_\mu(\eta \vee \xi_{-n}^{-1} | \xi_0^\infty)$ de duas maneiras diferentes. Primeiro,

$$\begin{aligned} H_\mu(\eta \vee \xi_{-n}^{-1} | \xi_0^\infty) &= H_\mu(\xi_{-n}^{-1} | \xi_0^\infty) + H_\mu(\eta | \xi_{-n}^\infty) \\ &= h_\mu(T^n, \xi_{-n}^{-1}) + H_\mu(\eta | \xi_{-n}^\infty) \\ &= h_\mu(T^n) + H_\mu(\eta | \xi_{-n}^\infty) \quad \because \text{Teo. 2.2.16.} \end{aligned}$$

Por outro lado,

$$\begin{aligned} H_\mu(\eta \vee \xi_{-n}^{-1} | \xi_0^\infty) &= H_\mu(\eta | \xi_0^\infty) + H_\mu(\xi_{-n}^{-1} | \xi_0^\infty \vee \eta) \\ &= H_\mu(\eta | \xi_0^\infty) + h_\mu(T^n, \xi_{-n}^{-1} | \eta) \\ &\geq H_\mu(\eta | \xi_0^\infty) + h_\mu(T^n, \xi_{-n}^{-1} | \mathcal{P}(T^n)) \quad \because \text{Prop.. 4.1.2(f)} \\ &= H_\mu(\eta | \xi_0^\infty) + h_\mu(T^n) \quad \because \text{Teo. 2.2.16 e Prop.. 4.1.2(e).} \end{aligned}$$

Assim, obtemos que $H_\mu(\eta | \xi_0^\infty) \leq \lim H_\mu(\eta | \xi_{-n}^\infty) = H_\mu(\eta | \mathcal{B}) = 0$, pois ξ gera $\mathcal{B}$ (aqui, usamos o Corolário 2.2.7). Agora, pela Proposição 2.1.13(b), segue que $\eta \leq \xi_0^\infty$ como

queríamos demonstrar.

(c) Pelo item (a), temos que

$$\bigvee_{\xi \in \mathcal{Z}} \text{Tail}(\xi) \subseteq \mathcal{P}(T)$$

pois $\text{Tail}(\xi) \subseteq \mathcal{P}(T)$ para toda partição $\xi \in \mathcal{Z}$. Agora, provaremos a inclusão contrária. Seja $\xi_k \in \mathcal{Z}$ tais que $\xi_k \uparrow \mathcal{P}(T)$. Então, para cada $k \in \mathbb{N}$, temos

$$H_\mu(\xi_k | (\xi_k)_1^\infty) = h_\mu(T, \xi_k) = 0$$

pois $\xi_k \subset \mathcal{P}(T)$. Daí, segue-se que $\xi_k \leq (\xi_k)_1^\infty$ para todo $k \in \mathbb{N}$ e, portanto, $(\xi_k)_m^\infty = (\xi_k)_n^\infty$ quaisquer que sejam $k \in \mathbb{N}$ e $m, n \geq 0$. Assim, obtemos que $\text{Tail}(\xi_k) = (\xi_k)_0^\infty$ para todo $k \in \mathbb{N}$, donde

$$\mathcal{P}(T) = \bigvee_{k=1}^{\infty} (\xi_k)_0^\infty = \bigvee_{k=1}^{\infty} \text{Tail}(\xi_k) \leq \bigvee_{\xi \in \mathcal{Z}} \text{Tail}(\xi).$$

Logo, $\mathcal{P}(T) \subseteq \bigvee_{\xi \in \mathcal{Z}} \text{Tail}(\xi)$ o que completa a prova da proposição.

□

Para aplicações dinâmicas precisamos considerar $\text{Tail}(\xi)$ onde a partição ξ é mensurável não necessariamente enumerável. Neste caso, a relação entre $\mathcal{P}(T)$ e ξ é mais sutil.

Teorema 4.2.4 *Sejam $(X, \mathcal{B}, \mu, T)$ um automorfismo e ξ uma partição mensurável (não necessariamente enumerável). Se $T^n \xi \uparrow \mathcal{B}$, então $\mathcal{P}(T) \subseteq \text{Tail}(\xi)$.*

Antes de demonstrar o teorema acima, apresentaremos dois lemas que serão necessários e que são interessantes por si só.

Lema 4.2.5 *Seja $\{\mathcal{B}_n\}_{n \in \mathbb{N}}$ uma família de sub- σ -álgebras de $\mathcal{B}$ tais que $\mathcal{B}_n \uparrow \mathcal{B}$. Então, para todo $\varepsilon > 0$ e para toda partição $\xi \in \mathcal{Z}$, existe $n_0 \in \mathbb{N}$ e uma partição finita $\eta \subset \mathcal{B}_{n_0}$ tal que $H_\mu(\xi | \eta) + H_\mu(\eta | \xi) < \varepsilon$.*

Demonstração: Fixemos $\xi = \{A_n\}_{n \in \mathbb{N}} \in \mathcal{Z}$ e $\varepsilon > 0$ arbitrário. Pelo Lema 2.2.12, podemos tomar $N > 1$ tal que, pondo

$$\xi' = \{A_1, \dots, A_N, \cup_{n \geq N} A_n\},$$

tem-se $H_\mu(\xi|\xi') < \varepsilon/4$. Tome $\delta > 0$ qualquer. Como $\mathcal{B}_n \uparrow \mathcal{B}$, existe um número natural n_i e um conjunto $B_i \in \mathcal{B}_{n_i}$ tal que

$$\mu(A_i \Delta B_i) < \frac{\delta}{2N^2} \quad (4.1)$$

para cada $i = 1, \dots, N$. Indutivamente, defina $P_1 = B_1$, $P_i = B_i \setminus \bigcup_{j=1}^{i-1} B_j$ para $i = 2, \dots, N$ e $P_0 = X \setminus \bigcup_{i=1}^N B_i$. Agora, considere $n_0 = \max\{n_i : 1 \leq i \leq N\}$ e $\eta = \{P_1, \dots, P_N, P_0\}$. Claramente η é uma partição de X e $\eta \subset \mathcal{B}_{n_0}$. Observe que, por definição, $P_i \Delta B_i = B_i \setminus P_i = B_i \cap (\bigcup_{j=1}^{i-1} B_j)$ para todo $i = 1, \dots, N$. Fixe $i = 1, \dots, N$ arbitrário. Dado $x \in P_i \Delta B_i$, existe $j < i$ tal que $x \in B_i \cap B_j$. Daí, como $A_i \cap A_j = \emptyset$, segue que $x \in (B_i \setminus A_i) \cup (B_j \setminus A_j)$. Isto mostra que

$$P_i \Delta B_i \subset \bigcup_{j=1}^i A_j \Delta B_j$$

e, portanto, $\mu(P_i \Delta B_i) \leq i\delta/2N^2 \leq \delta/2N$. Usando a desigualdade anterior juntamente com (4.1), obtemos

$$\mu(P_i \Delta A_i) < \frac{\delta}{2N} + \frac{\delta}{2N^2} \leq \frac{\delta}{N} \quad \text{para todo } i = 1, \dots, N, \quad (4.2)$$

uma vez que $P_i \Delta A_i \subset (P_i \Delta B_i) \cup (B_i \Delta A_i)$ para todo $i = 1, \dots, N$. Além disso, como η e ξ' são partições de X , temos que

$$P_0 \Delta \left(\bigcup_{n \geq N} A_n \right) \subset \bigcup_{i=1}^N (P_i \Delta A_i).$$

Assim, usando (4.2), concluímos que

$$\mu \left(P_0 \Delta \left(\bigcup_{n \geq N} A_n \right) \right) < \delta. \quad (4.3)$$

Conforme vimos na Proposição 2.1.18, supondo $\delta > 0$ suficientemente pequeno, as relações (4.2) e (4.3) implicam que $H_\mu(\xi'|\eta) < \varepsilon/4$ e $H_\mu(\eta|\xi') < \varepsilon/4$. Logo,

$$H_\mu(\xi|\eta) \leq H_\mu(\xi|\xi') + H_\mu(\xi'|\eta) < \frac{\varepsilon}{2}$$

e

$$H_\mu(\eta|\xi) \leq H_\mu(\eta|\xi') < \frac{\varepsilon}{4},$$

donde concluímos que $H_\mu(\xi|\eta) + H_\mu(\eta|\xi) < \varepsilon$.

□

Lema 4.2.6 *Sejam $\zeta, \eta \in \mathcal{M}$ tais que $\zeta \leq \eta$. Se $H_\mu(\xi|\zeta) = H_\mu(\xi|\eta)$ para toda partição $\xi \in \mathcal{Z}$, então $\zeta = \eta$.*

Demonstração: Tome $A \in \eta$. Vamos mostrar que $A \in \zeta$. Considere a partição finita $\xi := \{A, X \setminus A\}$. Por hipótese, temos que

$$H_\mu(\xi|\zeta) = H_\mu(\xi|\eta) = 0.$$

Assim, pela Proposição 2.1.13(a), obtemos que $\xi \subset \zeta$, logo $A \in \zeta$. Isso mostra que $\eta \subseteq \zeta$. De modo análogo, obtemos a inclusão reversa. □

Demonstração do Teorema 4.2.4: Seja $\zeta = \text{Tail}(\xi)$. Então, $\zeta = \bigwedge_{n \geq 0} \xi_n^\infty = \bigwedge_{n \geq 0} T^{-n}\xi$ e $T\zeta = \zeta$, uma vez que $\xi \leq T\xi$. Queremos mostrar que $\mathcal{P}(T) \leq \zeta$. Para tal, é suficiente provar que $H_\mu(\rho|\zeta) = H_\mu(\rho|\zeta \vee \mathcal{P}(T))$ para toda partição $\rho \in \mathcal{Z}$. De fato, uma vez que isso é verdade, então, pelo Lema 4.2.6, temos que $\mathcal{P}(T) \leq \zeta \vee \mathcal{P}(T) = \zeta$. Seja $\rho \in \mathcal{Z}$ arbitrária. Usando a Proposição 3.2.5(a) e a invariância de ζ por T , obtemos que

$$H_\mu(\rho \vee \zeta | \rho_1^\infty \vee \zeta_1^\infty) = H_\mu(\rho | \rho_1^\infty \vee \zeta_1^\infty) + H_\mu(\zeta | \rho_0^\infty \vee \zeta_1^\infty) = H_\mu(\rho | \rho_1^\infty \vee \zeta_1^\infty) < \infty$$

e, pela Proposição 4.1.3, tem-se

$$\begin{aligned} H_\mu(\rho | \rho_1^\infty \vee \zeta) &= H_\mu(\rho \vee \zeta | \rho_1^\infty \vee \zeta_1^\infty) \\ &= H_\mu(\rho \vee \zeta | \rho_1^\infty \vee \zeta_1^\infty \vee \mathcal{P}(T)) \\ &= H_\mu(\rho | \rho_1^\infty \vee \zeta \vee \mathcal{P}(T)), \end{aligned}$$

onde na primeira e na última igualdade usamos a Proposição 3.2.5(a) e a invariância de ζ por T simultaneamente. Observe que, usando que $\mathcal{P}(T) = \mathcal{P}(T^p)$ e substituindo T por T^p no cálculo anterior, obtemos

$$H_\mu\left(\rho \left| \left(\bigvee_{n=1}^{\infty} T^{-np} \rho \right) \vee \zeta \right.\right) = H_\mu\left(\rho \left| \left(\bigvee_{n=1}^{\infty} T^{-np} \rho \right) \vee \zeta \vee \mathcal{P}(T) \right.\right), \quad (4.4)$$

para todo $p \in \mathbb{N}$.

Suponha que $\rho \leq T^{k_0}\xi$ para algum $k_0 \in \mathbb{N}$. Então, usando que $T^{-1}\xi \leq \xi$, obtemos que $T^{-nk_0}\rho \leq T^{-m}\xi$ quaisquer que sejam os inteiros $n \geq 1$ e $0 \leq m \leq n-1$. Isto implica que $\bigvee_{i \geq 1} T^{-i(nk_0)}\rho \leq \bigwedge_{j=0}^{n-1} T^{-j}\xi$ para todo $n \in \mathbb{N}$. Daí, obtemos que

$$\zeta \leq \left(\bigvee_{i=1}^{\infty} T^{-i(nk_0)}\rho \right) \vee \zeta \leq \left(\bigvee_{i=1}^{\infty} T^{-i(nk_0)}\rho \right) \vee \left(\bigwedge_{j=0}^{n-1} T^{-j}\xi \right) = \bigwedge_{j=0}^{n-1} T^{-j}\xi$$

para todo $n \in \mathbb{N}$ e, por consequência,

$$\bigwedge_{n=1}^{\infty} \left(\left(\bigvee_{i=1}^{\infty} T^{-i(nk_0)} \rho \right) \vee \zeta \right) = \zeta.$$

Assim, usando o Lema 3.2.4(c) (sob todas as potências de T que são múltiplas de k_0) na igualdade (4.4), concluimos que $H_\mu(\rho|\zeta) = H_\mu(\rho|\zeta \vee \mathcal{P}(T))$.

Agora, suponha apenas que $\rho \in \mathcal{Z}$ e fixe $\varepsilon > 0$ arbitrário. Então, pelo Lema 4.2.5, existem $k_0 \in \mathbb{N}$ e uma partição finita $\eta_\varepsilon \subset T^{k_0}\xi$ tal que $H_\mu(\rho|\eta_\varepsilon) + H_\mu(\eta_\varepsilon|\rho) < \varepsilon$. Assim, pelo que vimos anteriormente, temos

$$\begin{aligned} H_\mu(\rho|\zeta) &\leq H_\mu(\rho|\eta_\varepsilon) + H_\mu(\eta_\varepsilon|\zeta) \\ &= H_\mu(\rho|\eta_\varepsilon) + H_\mu(\eta_\varepsilon|\zeta \vee \mathcal{P}(T)) \\ &\leq H_\mu(\rho|\eta_\varepsilon) + H_\mu(\eta_\varepsilon|\rho) + H_\mu(\rho|\zeta \vee \mathcal{P}(T)) \\ &< \varepsilon + H_\mu(\rho|\zeta \vee \mathcal{P}(T)). \end{aligned}$$

Sendo $\varepsilon > 0$ arbitrário, obtemos que $H_\mu(\rho|\zeta) \leq H_\mu(\rho|\zeta \vee \mathcal{P}(T))$ e portanto $H_\mu(\rho|\zeta) = H_\mu(\rho|\zeta \vee \mathcal{P}(T))$. Isso conclui a prova do teorema.

□

Independentemente da existência ou não de partições geradoras $\xi \in \mathcal{Z}$, sempre existe uma partição mensurável ξ tal que $\text{Tail}(\xi) = \mathcal{P}(T)$.

Teorema 4.2.7 *Dado um automorfismo $(X, \mathcal{B}, \mu, T)$, existe uma partição mensurável ξ tal que:*

- (a) $T^{-1}\xi \leq \xi$,
- (b) $T^n\xi \uparrow \mathcal{B}$, e
- (c) $\text{Tail}(\xi) = \mathcal{P}(T)$.

Demonstração: Para provar esse teorema vamos precisar do seguinte lema.

Lema 4.2.8 *Se $\alpha, \beta, \gamma \in \mathcal{Z}$ e $\alpha \leq \beta$, então*

$$H_\mu(\alpha|\beta_1^\infty \vee \gamma_n^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\alpha|\beta_1^\infty).$$

Demonstração: Para provar esse lema, notemos primeiro que $H_\mu(\alpha|\beta_1^\infty \vee \gamma_n^\infty)\}_{n \geq 1}$ é não-decrescente. Afirmamos que é suficiente provar o lema para o caso em que $\alpha = \beta$.

Com efeito, como

$$\begin{aligned}
H_\mu(\beta|\beta_1^\infty \vee \gamma_n^\infty) &= H_\mu(\alpha \vee \beta|\beta_1^\infty \vee \gamma_n^\infty) \\
&= H_\mu(\alpha|\beta_1^\infty \vee \gamma_n^\infty) + H_\mu(\beta|\beta_1^\infty \vee \gamma_n^\infty \vee \alpha) \\
&\leq H_\mu(\alpha|\beta_1^\infty) + H_\mu(\beta|\beta_1^\infty \vee \alpha) \\
&= H_\mu(\alpha \vee \beta|\beta_1^\infty) \\
&= H_\mu(\beta|\beta_1^\infty),
\end{aligned}$$

se é verdade que $H_\mu(\beta|\beta_1^\infty \vee \gamma_n^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\beta|\beta_1^\infty)$, então também será verdade que $H_\mu(\alpha|\beta_1^\infty \vee \gamma_n^\infty) \xrightarrow{n \rightarrow +\infty} H_\mu(\alpha|\beta_1^\infty)$ pois $H_\mu(\alpha|\beta_1^\infty \vee \gamma_n^\infty) \leq H_\mu(\alpha|\beta_1^\infty)$ e $H_\mu(\beta|\beta_1^\infty \vee \gamma_n^\infty \vee \alpha) \leq H_\mu(\beta|\beta_1^\infty \vee \alpha)$. (Intuitivamente isso significa que, sob condição de monotonicidade, basta provar o lema para o menor β possível.)

Assuma que $\alpha = \beta$. Afirmamos que

$$\frac{1}{n} \sum_{k=0}^n H_\mu(\beta|\beta_1^\infty \vee \gamma_{k+1}^\infty) = \frac{1}{n} H_\mu(\beta_{-n}^0|\beta_1^\infty \vee \gamma_1^\infty).$$

De fato, como $\beta_{-n}^0 = \beta_{-(n-1)}^0 \vee T^n \beta$, temos

$$\begin{aligned}
H_\mu(\beta_{-n}^0|\beta_1^\infty \vee \gamma_1^\infty) &= H_\mu(\beta_{-(n-1)}^0|\beta_1^\infty \vee \gamma_1^\infty) + H_\mu(T^n \beta|\beta_{-(n-1)}^\infty \vee \gamma_1^\infty) \\
&= H_\mu(\beta_{-(n-1)}^0|\beta_1^\infty \vee \gamma_1^\infty) + H_\mu(T^n \beta|T^n(\beta_1^\infty \vee \gamma_{n+1}^\infty)) \\
&= H_\mu(\beta_{-(n-1)}^0|\beta_1^\infty \vee \gamma_1^\infty) + H_\mu(\beta|\beta_1^\infty \vee \gamma_{n+1}^\infty).
\end{aligned}$$

Seguindo o procedimento acima de forma indutiva, obtemos que

$$H_\mu(\beta_{-n}^0|\beta_1^\infty \vee \gamma_1^\infty) = \sum_{k=0}^n H_\mu(\beta|\beta_1^\infty \vee \gamma_{k+1}^\infty),$$

como afirmamos. Agora, pelo Lema 3.2.6(a), vemos que

$$\lim_n \frac{1}{n} \sum_{k=0}^n H_\mu(\beta|\beta_1^\infty \vee \gamma_{k+1}^\infty) = \lim_n \frac{1}{n} H_\mu(\beta_{-n}^0|\beta_1^\infty \vee \gamma_1^\infty) = H_\mu(\beta|\beta_1^\infty).$$

Assim, como a sequência $\{H_\mu(\beta|\beta_1^\infty \vee \gamma_{k+1}^\infty)\}_{k \geq 0}$ é não-decrescente, concluímos que

$$\lim_n \frac{1}{n} H_\mu(\beta|\beta_1^\infty \vee \gamma_n^\infty) = \lim_n \frac{1}{n} \sum_{k=0}^n H_\mu(\beta|\beta_1^\infty \vee \gamma_{k+1}^\infty) = H_\mu(\beta|\beta_1^\infty),$$

o que prova o lema.

□

Agora, retornaremos à demonstração do teorema e procederemos na construção de ξ . Primeiro, tome uma sequência $(\alpha_k)_{k \geq 1}$ em $\mathcal{X}$ tal que $\alpha_k \uparrow \mathcal{B}$ quando $k \rightarrow +\infty$. Fixe um número natural $n_1 \geq 1$ qualquer e tome $\eta_1 = T^{-n_1}\alpha_1$. Pelo Lema 4.2.8, temos

$$H_\mu(T^{-n_1}\alpha_1|(T^{-n_1}\alpha_1)_1^\infty) - H_\mu(T^{-n_1}\alpha_1|(T^{-n_1}\alpha_1)_1^\infty \vee (T^{-k}\alpha_2)_1^\infty) \xrightarrow{k \rightarrow +\infty} 0.$$

Escolha $n_2 > n_1$ tal que

$$H_\mu(T^{-n_1}\alpha_1|(T^{-n_1}\alpha_1)_1^\infty) - H_\mu(T^{-n_1}\alpha_1|(T^{-n_1}\alpha_1 \vee T^{-n_2}\alpha_2)_1^\infty) < \frac{1}{2}.$$

Defina $\eta_2 = T^{-n_1}\alpha_1 \vee T^{-n_2}\alpha_2$. Novamente, pelo Lema 4.2.8, temos

$$H_\mu(\eta_p|(\eta_2)_1^\infty) - H_\mu(\eta_p|(\eta_2)_1^\infty \vee (T^{-k}\alpha_3)_1^\infty) \xrightarrow{k \rightarrow +\infty} 0, \quad p = 1, 2.$$

Escolha $n_3 > n_2 > n_1$ tal que

$$H_\mu(\eta_p|(\eta_2)_1^\infty) - H_\mu(\eta_p|(\eta_2 \vee T^{-n_3}\alpha_3)_1^\infty) < \frac{1}{p2^{3-p}}, \quad p = 1, 2.$$

Defina $\eta_3 = \eta_2 \vee T^{-n_3}\alpha_3 = T^{-n_1}\alpha_1 \vee T^{-n_2}\alpha_2 \vee T^{-n_3}\alpha_3$. Repetindo esse procedimento indefinidamente, podemos definir $\eta_p = \bigvee_{k=1}^p T^{-n_k}\alpha_k$ para todo $p \in \mathbb{N}$, onde os naturais $n_1 < n_2 < n_3 < \dots$ são definidos indutivamente satisfazendo

$$H_\mu(\eta_p|(\eta_{q-1})_1^\infty) - H_\mu(\eta_p|(\eta_q)_1^\infty) < \frac{1}{p2^{q-p}}, \quad p = 1, 2, \dots, q-1. \quad (4.5)$$

Seja $\eta = \bigvee_{k \geq 1} T^{-n_k}\alpha_k$. Como η não satisfaz (a), definamos $\xi = \eta_1^\infty$. Assim, ξ satisfaz (i). Claramente, ξ satisfaz (b), pois

$$T^{-n_q}\alpha_q \leq T\xi \implies \alpha_q \leq T^{n_q+1}\xi \implies T^{n_q+1}\xi \uparrow \mathcal{B}.$$

Resta provar que ξ satisfaz (c). Antes disso, vamos obter uma consequência de (4.5): dado $k > p$, some sob q , de $p+1$ até k em (4.5), para obter que

$$H_\mu(\eta_p|(\eta_p)_1^\infty) - H_\mu(\eta_p|(\eta_k)_1^\infty) < \frac{1}{p}.$$

Fixando p e fazendo $k \rightarrow +\infty$ obtemos, por Chung-Neveu, que

$$H_\mu(\eta_p|(\eta_p)_1^\infty) - H_\mu(\eta_p|\xi) < \frac{1}{p}.$$

Portanto, passando o limite em p , obtemos

$$\lim_{p \rightarrow +\infty} [H_\mu(\eta_p|(\eta_p)_1^\infty) - H_\mu(\eta_p|\xi)] = 0. \quad (4.6)$$

Pelo Teorema 4.2.4, temos que $\mathcal{P}(T) \subseteq \text{Tail}(\xi)$. Para obtermos a inclusão reversa, provaremos que $H_\mu(\zeta|\zeta_1^\infty) = 0$ para toda partição $\zeta \in \mathcal{Z}$ tal que $\zeta \leq \text{Tail}(\xi)$. Primeiro, note que $T^{-(n+1)}\xi \leq T^{-n}\xi$ para todo $n \geq 0$, pois $T^{-1}\xi \leq \xi$. Segue-se daí que $\text{Tail}(\xi) \leq T^{-n}\xi$ para todo $n \geq 0$. Assim, dado $n \geq 0$ temos que

$$\zeta \leq \text{Tail}(\xi) \leq \xi \implies T^{-n}\zeta \leq \xi$$

e

$$\zeta \leq \text{Tail}(\xi) \leq T^{-n}\xi \implies T^n\zeta \leq \xi,$$

donde concluímos que $\zeta_{-\infty}^{+\infty} \leq \xi$. Além disso, pela definição de ξ , temos que $(\eta_p)_1^\infty \leq \xi$ para todo $p \geq 1$. Agora, pela Proposição 3.2.7, temos

$$\begin{aligned} H_\mu(\zeta|\zeta_1^\infty) &= H_\mu(\zeta \vee \eta_p|\zeta_1^\infty \vee (\eta_p)_1^\infty) - H_\mu(\eta_p|(\eta_p)_1^\infty \vee \zeta_{-\infty}^{+\infty}) \\ &= H_\mu(\eta_p|(\eta_p)_1^\infty) + H_\mu(\zeta|\zeta_1^\infty \vee (\eta_p)_{-\infty}^{+\infty}) - H_\mu(\eta_p|(\eta_p)_1^\infty \vee \zeta_{-\infty}^{+\infty}) \\ &\leq H_\mu(\eta_p|(\eta_p)_1^\infty) + H_\mu(\zeta|(\eta_p)_1^\infty) - H_\mu(\eta_p|\xi) \end{aligned}$$

para todo $p \geq 1$, pois $(\eta_p)_1^\infty \leq \zeta_1^\infty \vee (\eta_p)_{-\infty}^{+\infty}$ e $(\eta_p)_1^\infty \vee \zeta_{-\infty}^{+\infty} \leq \xi$. Fazendo $p \rightarrow +\infty$ e usando igualdade (4.6) e o Lema de Chung-Neveu na desigualdade acima, obtemos que $h_\mu(T, \zeta) = H_\mu(\zeta|\zeta_1^\infty) \leq H_\mu(\zeta|\xi) = 0$. Em particular, tomando $\zeta = \xi_A$ com $A \in \text{Tail}(\xi)$, vemos que $\text{Tail}(\xi) \subseteq \mathcal{P}(T)$, o que completa a nossa demonstração.

□

Definição 4.2.9 (Lei zero-um de Kolmogorov) *Um automorfismo $(X, \mathcal{B}, \mu, T)$ satisfaz a Lei zero-um de Kolmogorov se para cada $\xi \in \mathcal{Z}$ tem-se que $\mu(A) \in \{0, 1\}$ para todo $A \in \text{Tail}(\xi)$. Alternativamente, isso significa que $\text{Tail}(\xi) = \xi_\emptyset \pmod{\mu}$ para todo $\xi \in \mathcal{Z}$, onde $\xi_\emptyset = \{\emptyset, X\}$.*

Proposição 4.2.10 *Um automorfismo $(X, \mathcal{B}, \mu, T)$ satisfaz a Lei zero-um de Kolmogorov se, e somente se, $\mathcal{P}(T) = \xi_\emptyset$.*

Demonstração: Com efeito, se o automorfismo $(X, \mathcal{B}, \mu, T)$ satisfaz a Lei zero-um de Kolmogorov, então $\text{Tail}(\xi) = \xi_\emptyset$ para toda partição $\xi \in \mathcal{Z}$. Pela Proposição 4.2.3(c), segue que

$$\mathcal{P}(T) = \bigvee_{\xi \in \mathcal{Z}} \text{Tail}(\xi) = \xi_\emptyset.$$

Reciprocamente, assumamos que $\mathcal{P}(T) = \xi_\emptyset$. Pela Proposição 4.2.3(a), vemos que $\text{Tail}(\xi) \subseteq \mathcal{P}(T) = \xi_\emptyset$ para toda partição $\xi \in \mathcal{Z}$, donde concluímos que $\text{Tail}(\xi) = \xi_\emptyset$ para toda partição $\xi \in \mathcal{Z}$.

□

5 A PROPRIEDADE K

Neste capítulo, discutiremos sobre os K-automorfismos. Estes automorfismos foram introduzidos inicialmente por Kolmogorov, quando ele provou que sistemas 2-shift e 3-shift não são metricamente isomorfos. Em seus estudos, partindo dos argumentos da teoria da probabilidade, Kolmogorov chamou tais automorfismos de sistemas “quasi-regulares”. Hoje em dia, esses sistemas passaram a ser chamados de automorfismos de Kolmogorov, K-automorfismos ou simplesmente K -sistemas. As principais referências que seguiremos neste capítulo são [1], [5], [7], [10] e [8].

5.1 K-automorfismos

No que se segue, faremos um abuso de notação e usaremos a notação $\xi_\emptyset$ tanto para denotar a partição trivial $\{X\}$ quanto a σ -álgebra trivial $\{\emptyset, X\}$.

Definição 5.1.1 *Um automorfismo $(X, \mathcal{B}, \mu, T)$ é chamado de K-automorfismo (ou automorfismo de Kolmogorov) se existe uma partição mensurável ξ , chamada de K -partição, tal que:*

- (K1) $T^{-1}\xi \leq \xi$;
- (K2) $T^n\xi \uparrow \mathcal{B}$ quando $n \rightarrow +\infty$;
- (K3) $T^{-n}\xi \downarrow \xi_\emptyset$ quando $n \rightarrow +\infty$.

Claramente, mantida a propriedade (K1), as propriedades (K2) e (K3) podem ser substituídas por

$$(K2)' \bigvee_{n=0}^{+\infty} T^n\xi = \mathcal{B} \quad \text{e} \quad (K3)' \bigwedge_{n=0}^{+\infty} T^{-n}\xi = \xi_\emptyset$$

respectivamente. A álgebra de Pinsker caracteriza bem os K-automorfismos, como veremos agora.

Teorema 5.1.2 (Rokhlin-Sinai) *Seja $(X, \mathcal{B}, \mu, T)$ um automorfismo. As seguintes afirmações são equivalentes:*

- (a) $(X, \mathcal{B}, \mu, T)$ é um K -automorfismo.
- (b) $\mathcal{P}(T) = \xi_\emptyset$.
- (c) $(X, \mathcal{B}, \mu, T)$ satisfaz a Lei zero-um de Kolmogorov.
- (d) Todo fator não-trivial de $(X, \mathcal{B}, \mu, T)$ tem entropia positiva.

Demonstração: $(a) \Rightarrow (b)$: Se $(X, \mathcal{B}, \mu, T)$ é um K-automorfismo então existe uma partição mensurável ξ que satisfaz (K1)–(K3). Daí, pelo Teorema 4.2.4 temos

$$\mathcal{P}(T) \subseteq \text{Tail}(\xi) = \bigwedge_{n \geq 0} \xi_n^\infty = \bigwedge_{n \geq 0} T^{-n} \xi = \xi_\emptyset,$$

donde concluímos que $\mathcal{P}(T) = \xi_\emptyset$.

$(b) \Rightarrow (a)$: Segue do Teorema 4.2.7.

$(b) \Leftrightarrow (c)$: Essa equivalência é a Proposição 4.2.10.

$(b) \Leftrightarrow (d)$: Segue da Proposição 4.1.2(d).

□

Definição 5.1.3 Dizemos que um automorfismo $(X, \mathcal{B}, \mu, T)$ tem entropia completamente positiva se todo fator não-trivial de $(X, \mathcal{B}, \mu, T)$ tem entropia positiva.

Pelo Teorema 5.1.2, um automorfismo $(X, \mathcal{B}, \mu, T)$ tem entropia completamente positiva se, e somente se, ele é um K-automorfismo. Além disso, por definição, todo fator não-trivial de um sistema com entropia completamente positiva também tem entropia completamente positiva. Assim obtemos o seguinte corolário do Teorema 5.1.2.

Corolário 5.1.4 Todo fator não trivial de um K-automorfismo é um K-automorfismo.

Corolário 5.1.5 $(X, \mathcal{B}, \mu, T)$ é um K-automorfismo se, e somente se, $h_\mu(T, \xi) > 0$ para toda partição finita $\xi \neq \xi_\emptyset$.

Demonstração: Suponha que $(X, \mathcal{B}, \mu, T)$ é um K-automorfismo. Dada uma partição finita $\xi \neq \xi_\emptyset$, considere a σ -álgebra $\mathcal{F}$ que consiste do conjunto $\emptyset$ e de todos os subconjuntos de X que podem ser obtidos a partir de uma união, módulo μ , de átomos de ξ . Então, $(X, \mathcal{F}, \mu, T)$ é um fator não-trivial de $(X, \mathcal{B}, \mu, T)$. Pelo Teorema 5.1.2(d), obtemos que

$$0 < h_\mu(T \upharpoonright \mathcal{F}) = \sup_{\eta \leq \xi} h_\mu(T, \eta) \leq h_\mu(T, \xi),$$

donde concluímos que $h_\mu(T, \xi) > 0$. Agora, suponha que $h_\mu(T, \xi) > 0$ para toda partição finita $\xi \neq \xi_\emptyset$. Seja $(Y, \mathcal{C}, \nu, S)$ um fator não-trivial arbitrário de $(X, \mathcal{B}, \mu, T)$, com aplicação fator π . Fixe uma partição finita não-trivial η do conjunto Y . Como $\pi^{-1}\eta$ é uma partição finita não-trivial de X , temos $h_\nu(S, \eta) = h_\mu(T, \pi^{-1}\eta) > 0$ (veja a demonstração do Teorema 3.3.4). Logo, $h_\nu(S) \geq h_\nu(S, \eta) > 0$. Como o fator $(Y, \mathcal{C}, \nu, S)$ é arbitrário, seguem novamente pelo Teorema 5.1.2(d) que $(X, \mathcal{B}, \mu, T)$ é um K-automorfismo.

□

5.2 Sistemas K-mixing

A propriedade K-mixing, introduzida logo abaixo, é conveniente porque fornece uma propriedade uniforme de mistura, mais forte do que as noções usuais de ergodicidade, fracamente misturador (weak mixing) e misturador (mixing) vistas usualmente em um curso básico de teoria ergódica.

Definição 5.2.1 (K-mixing) *Um automorfismo $(X, \mathcal{B}, \mu, T)$ é K-mixing se para quaisquer conjuntos $A_0, A_1, \dots, A_r \in \mathcal{B}$, $r \geq 0$, temos*

$$\lim_{n \rightarrow \infty} \left(\sup_{B \in \mathcal{B}_n^\infty(A_1, \dots, A_r)} |\mu(A_0 \cap B) - \mu(A_0)\mu(B)| \right) = 0, \quad (5.1)$$

onde $\mathcal{B}_n^\infty(A_1, \dots, A_r)$ é a menor σ -álgebra gerada pelos conjuntos $T^{-k}A_i$ com $k \geq n$, $i = 1, \dots, r$.

Olhando para a Definição 5.2.1 com uma visão probabilística, a propriedade K-mixing significa que o movimento do conjunto A_0 não depende de quaisquer eventos que ocorram na parte suficientemente distante da trajetória dos conjuntos A_i com $1 \leq i \leq r$. É fácil ver que se $(X, \mathcal{B}, \mu, T)$ é K-mixing então $(X, \mathcal{B}, \mu, T)$ é um sistema misturador. Mais do que isso, a propriedade K-mixing implica que o sistema é misturador de qualquer ordem.

O próximo teorema garante que a propriedade K-mixing caracteriza bem os K-automorfismos.

Teorema 5.2.2 *Um automorfismo $(X, \mathcal{B}, \mu, T)$ é um K-automorfismo se, e somente se, é K-mixing.*

Demonstração: Inicialmente, suponha que $(X, \mathcal{B}, \mu, T)$ é um K-automorfismo. Sejam $A_0, A_1, \dots, A_r \in \mathcal{B}$ arbitrários. Sem perda de generalidade, podemos assumir que os conjuntos $A_1, \dots, A_r$ constituem uma partição η do espaço X . Como $(X, \mathcal{B}, \mu, T)$ é um K-automorfismo, temos que $\text{Tail}(\eta) = \xi_\emptyset$. Fixado $n \in \mathbb{N}$, tome $B \in \mathcal{B}_n$ qualquer, onde $\mathcal{B}_n$ denota a σ -álgebra gerada pela partição η_n^∞ . Usando a definição de esperança condicional, obtemos

$$\begin{aligned} |\mu(A_0 \cap B) - \mu(A_0)\mu(B)| &= \left| \int_B \mathbb{E}(1_{A_0} | \mathcal{B}_n) d\mu(x) - \int_B \mu(A_0) d\mu(x) \right| \\ &\leq \int_B |\mathbb{E}(1_{A_0} | \mathcal{B}_n) - \mu(A_0)| d\mu(x) \\ &\leq \|\mathbb{E}(1_{A_0} | \mathcal{B}_n) - f\|_{L^1}, \end{aligned}$$

onde $f = \int 1_{A_0} d\mu = \mathbb{E}(1_{A_0} | \{\emptyset, X\})$ é a função constante. Como $B \in \mathcal{B}_n$ é arbitrário, temos

$$\sup_{B \in \mathcal{B}_n} |\mu(A_0 \cap B) - \mu(A_0)\mu(B)| \leq \|\mathbb{E}(1_{A_0} | \mathcal{B}_n) - \mu(A_0)\|_{L^1}.$$

Por fim, como $\mathcal{B}_n \downarrow \{\emptyset, X\}$, segue do Teorema de convergência de martingais (Teorema 2.2.5) que $\lim_{n \rightarrow \infty} \|\mathbb{E}(1_{A_0} | \mathcal{B}_n) - f\|_{L^1} = 0$ e portanto

$$\lim_{n \rightarrow \infty} \sup_{B \in \mathcal{B}_n} |\mu(A_0 \cap B) - \mu(A_0)\mu(B)| = 0,$$

provando que $(X, \mathcal{B}, \mu, T)$ é K-mixing. Reciprocamente, assuma que $(X, \mathcal{B}, \mu, T)$ é K-mixing. Seja η uma partição finita arbitraria. Dado $A \in \text{Tail}(\eta)$ qualquer, temos que $A \in \mathcal{B}_n$ para todo $n \in \mathbb{N}$. Assim, pela igualdade (5.1), obtemos

$$|\mu(A \cap A) - \mu(A)\mu(A)| = 0,$$

donde vemos que $\mu(A) \in \{0, 1\}$. Isso mostra que $\text{Tail}(\eta) = \xi_\emptyset$ para toda partição finita η . Usando a Proposição 4.2.3(c), segue que $\mathcal{P}(T) = \xi_\emptyset$ e daí o Teorema 5.1.2(b) implica que $(X, \mathcal{B}, \mu, T)$ é um K-automorfismo.

□

6 EXEMPLOS

Neste capítulo, apresentaremos três exemplos importantes de K-automorfismos. Vale salientar que usaremos algumas ferramentas Dinâmica Hiperbólica e da Teoria da Medida (que estão além deste estudo) para concluir que tais sistemas são K-automorfismos. As principais referências que seguiremos neste capítulo são [4], [10] e [7].

6.1 Shifts de Bernoulli

Fixe um inteiro $d > 0$ e considere o alfabeto $\mathcal{A} = \{1, \dots, d\}$. Seja

$$\Sigma = \{\underline{x} = (x_n)_{n \in \mathbb{Z}} : x_n \in \mathcal{A}, \forall n \in \mathbb{Z}\}$$

dotado da topologia produto, e seja $\mathcal{B}$ a σ -álgebra de Borel de Σ . Dados $v_m, v_{m+1}, \dots, v_{m+k-1} \in \{1, \dots, d\}$, podemos considerar o conjunto

$${}_m[v_m, \dots, v_{m+k-1}] = \{\underline{x} = (x_n)_{n \in \mathbb{Z}} \in \Sigma : x_n = v_n \text{ para } n = m, m+1, \dots, m+k-1\},$$

que é chamado de um *cilindro* na posição m e tamanho k . Note que os cilindros são conjuntos clopen (abertos e fechados). Seja $\mathcal{B}_0$ a álgebra formada por uniões finitas de cilindros disjuntos. Temos que σ -álgebra $\mathcal{B}$ é gerada pela álgebra $\mathcal{B}_0$.

Seja $p = (p_1, \dots, p_d)$ um vetor de probabilidade com entradas não-nulas. A medida

$$\mu = \prod_{n \in \mathbb{Z}} p$$

é chamada de *medida de Bernoulli*. Equivalentemente, temos que

$$\mu({}_m[v_m, \dots, v_k]) = p_{v_m} \cdots p_{v_k}.$$

Seja $\sigma : \Sigma \rightarrow \Sigma$ o deslocamento à esquerda em Σ , isto é, $\sigma((x_n)_{n \in \mathbb{Z}}) = (x_{n+1})_{n \in \mathbb{Z}}$. Observe que a pré-imagem de um cilindro ainda é um cilindro: $\sigma^{-1}({}_m[v_m, \dots, v_k]) = {}_{m+1}[v_m, \dots, v_k]$. Daí, segue-se que σ é mensurável relativamente a $\mathcal{B}$. Além disso,

$$\mu(\sigma^{-1}({}_m[v_m, \dots, v_k])) = p_{v_m} \cdots p_{v_k} = \mu({}_m[v_m, \dots, v_k]).$$

Assim, a classe monótona $\mathcal{C} = \{A \in \mathcal{B} : \mu(\sigma^{-1}(A)) = \mu(A)\}$ contém a álgebra $\mathcal{B}_0$ das uniões finitas de cilindros disjuntos. Pelo Teorema das Classes Monótonas (veja [4],

Teorema A.1.18), concluímos que $\mathcal{C} = \mathcal{B}$. Isso mostra que μ é σ -invariante. Usando argumentos análogos, vemos que a transformação inversa σ^{-1} é mensurável relativamente a $\mathcal{B}$ e preserva a medida μ . Portanto, $(\Sigma, \mathcal{B}, \mu, \sigma)$ é um automorfismo.

Proposição 6.1.1 *O automorfismo $(\Sigma, \mathcal{B}, \mu, \sigma)$ é um K -automorfismo.*

Demonstração: Pelo Teorema 5.1.2, é suficiente mostrar que $\mathcal{P}(\sigma)$ é trivial. Para tal, considere $\xi = \{0[i] : i = 1, \dots, d\}$ a partição de Σ em cilindros na posição zero. Sabemos que ξ gera $\mathcal{B}$, pois $\xi_{-\infty}^{\infty}$ é a partição de Σ em átomos unitários. Assim, pela Proposição 4.2.3(b), obtemos que $\mathcal{P}(\sigma) = \text{Tail}(\xi)$. Afirmamos que $\text{Tail}(\xi) = \xi_{\emptyset}$. Com efeito, fixe

$$A \in \text{Tail}(\xi) = \bigcap_{n=0}^{\infty} \left(\bigvee_{k=-\infty}^{-n} \sigma^k \xi \right),$$

arbitrário. Seja $B \in \bigvee_{k=j}^{\infty} \sigma^k \xi$ para algum $j \in \mathbb{Z}$ fixo. Como $A \in \text{Tail}(\xi) \subset \bigvee_{k=-\infty}^{-n} \sigma^k \xi$ para todo $n < j$, temos que A e B são independentes¹, ou seja, $\mu(A \cap B) = \mu(A)\mu(B)$. Considere o conjunto $\mathcal{C}_A = \{B \in \mathcal{B} : \mu(A \cap B) = \mu(A)\mu(B)\}$. Claramente, $\mathcal{C}_A$ é uma classe monótona e, além disso, pelo que vimos acima $\mathcal{C}_A$ contém

$$\mathcal{B}_0 \subset \bigcup_{j=-\infty}^{\infty} \bigvee_{k=j}^{\infty} \sigma^k \xi.$$

Daí, pelo Teorema das Classes Monótonas, tem-se $\mu(A \cap B) = \mu(A)\mu(B)$ para todo $B \in \mathcal{B}$. Em particular, colocando $B = A$ obtemos que $\mu(A) = \mu(A)^2$ e, desta forma, vemos que $\mu(A) \in \{0, 1\}$. Como $A \in \text{Tail}(\xi)$ é arbitrário, obtemos que $\text{Tail}(\xi) = \xi_{\emptyset}$ como afirmamos. Portanto, $\mathcal{P}(T) = \text{Tail}(\xi) = \xi_{\emptyset}$ o que prova a proposição. □

Seja ξ a partição de Σ em cilindros na posição zero como na demonstração da Proposição 6.1.1. Considerando a partição $\xi^- = \xi_0^{\infty}$, não é difícil notar que $\sigma^{-1}\xi^- \leq \xi^-$ e que $\sigma^n \xi^- \uparrow \mathcal{B}$. Além disso, uma vez que $\sigma^{-1}\xi^- \leq \xi^-$, temos que

$$\text{Tail}(\xi^-) = \bigcap_{n=0}^{\infty} \left(\bigvee_{k=n}^{\infty} \sigma^{-k} \xi^- \right) = \bigcap_{n=0}^{\infty} \sigma^{-n} \xi^- = \text{Tail}(\xi) = \xi_{\emptyset},$$

donde vemos que $\sigma^{-n}\xi^- \downarrow \xi_{\emptyset}$. Isso mostra que a partição ξ^- é uma K -partição.

¹ Para ver isso, observe que se A e B estão em $\mathcal{B}_0$ então A e B são independentes. Para o caso geral, aproxime A e B por elementos de $\mathcal{B}_0$.

6.2 Deslocamentos de Markov

Fixe um inteiro $d > 0$, e seja $P = (p_{ij})_{1 \leq i, j \leq d}$ uma matriz estocástica, isto é, $p_{ij} \geq 0$ para todo $1 \leq i, j \leq d$ e $\sum_j p_{ij} = 1$ para todo $1 \leq i \leq d$. Consideremos $\Sigma = \Sigma_P$ o espaço das sequências P -admissíveis, ou seja,

$$\Sigma = \{\underline{x} = (x_n)_{n \in \mathbb{Z}} \in \{1, \dots, d\}^{\mathbb{Z}} : p_{x_n, x_{n+1}} > 0, \forall n \in \mathbb{Z}\}.$$

A matriz P pode ser considerada como a matriz de incidência de um grafo direcionado, e Σ como o espaço dos caminhos bilaterais no grafo. Definimos cilindros de modo análogo à seção anterior. Observe que, ${}_m[v_m, \dots, v_k] \neq \emptyset$ se e somente se $p_{v_m, v_{m+1}}, \dots, p_{v_{k-1}, v_k} > 0$. Muniremos Σ com a σ -álgebra de Borel $\mathcal{B}_\Sigma$. Seja $\mathcal{B}_0$ a álgebra formada por uniões finitas de cilindros disjuntos de Σ . Temos que $\mathcal{B}_\Sigma$ é gerada pela álgebra $\mathcal{B}_0$.

Seja $p = (p_i)_{1 \leq i \leq d}$ um vetor de probabilidade com $p_i > 0$ para todo $i = 1, \dots, d$. Assuma que o vetor p é estacionário para a matriz P , isto é, $pP = p$. Neste caso, podemos definir uma probabilidade μ definindo-a nos cilindros como $\mu({}_m[v_m, \dots, v_k]) := p_{v_m} p_{v_m, v_{m+1}} \cdots p_{v_{k-1}, v_k}$ e estendendo-a pelo Teorema de Carathéodory. O espaço de probabilidade $(\Sigma, \mathcal{B}_\Sigma, \mu)$ pode ser considerado como uma cadeia de Markov, em que p denota a probabilidade inicial dos estados e P denota as probabilidades de transições entre os estados. Seja $\sigma : \Sigma \rightarrow \Sigma$ o shift à esquerda em Σ . Observando que a pré-imagem de um cilindro ainda é um cilindro, concluímos que σ é mensurável relativamente a $\mathcal{B}_\Sigma$. Além disso, pela estacionariedade de p , obtemos que μ é σ -invariante. Usando argumentos análogos, vemos que a transformação inversa σ^{-1} é mensurável relativamente a $\mathcal{B}_\Sigma$ e preserva a medida μ . Portanto $(\Sigma, \mathcal{B}_\Sigma, \mu, \sigma)$ é um automorfismo, o qual chamaremos de *deslocamento de Markov* ou simplesmente de *shift de Markov*.

Lembre-se que, para cada $n \in \mathbb{N}$, usamos a notação $(p_{i,j}^n)_{1 \leq i, j \leq d}$ para denotar a matriz P^n . Dizemos que a matriz P é aperiódica se existe $N \in \mathbb{N}$ tal que $p_{i,j}^N > 0$ quaisquer que sejam $i, j \in \{1, \dots, d\}$. Sabemos que a matriz P é aperiódica se, e somente se,

$$\lim_{n \rightarrow \infty} P_{i,j}^n = p_j$$

quaisquer que sejam $i, j \in \{1, \dots, d\}$ (Veja [4, Lema 7.2.12]). Além disso, o deslocamento de Markov é mixing se, e somente se, a matriz de transição P é uma matriz aperiódica (Veja [4, Teorema 7.2.12]).

Observe que, se um deslocamento de Markov é um K -automorfismo então a matriz de transição é aperiódica. Com efeito, se o sistema $(\Sigma, \mathcal{B}_\Sigma, \mu, \sigma)$ é um K -

automorfismos então, pelo Teorema 5.2.2, tem-se que $(\Sigma, \mathcal{B}_\Sigma, \mu, \sigma)$ é K -mixing. Em particular, o sistema $(\Sigma, \mathcal{B}_\Sigma, \mu, \sigma)$ é mixing. Assim, a matriz de transição P é necessariamente aperiódica.

A recíproca também é verdadeira, como mostra a proposição abaixo.

Proposição 6.2.1 *Se a matriz de transição P é aperiódica, então o Deslocamento de Markov $(\Sigma, \mathcal{B}_\Sigma, \mu, \sigma)$ é um K -automorfismo.*

Demonstração: Nesta demonstração, procederemos de forma análoga ao que fizemos na demonstração da Proposição 6.1.1. Seja $\xi = \{0[i] : i = 1, \dots, d\}$ a partição de Σ em cilindros na posição zero. Temos que ξ gera $\mathcal{B}$, pois $\xi_{-\infty}^\infty$ é a partição de Σ em átomos unitários. Mostraremos que $\text{Tail}(\xi) = \xi_\emptyset$. Para tal, precisaremos do seguinte lema.

Lema 6.2.2 *Dados a e $r \geq 0$ inteiros, denote por $\mathcal{A}(\xi_a^{a+r})$ a álgebra gerada pela partição ξ_a^{a+r} . Então, para $A \in \mathcal{A}(\xi_a^{a+r})$ e $N > a + r$ fixos, tem-se*

$$\mu(A)\mu(B) \inf_{\substack{i,j \\ b \geq N}} \left(\frac{P_{i,j}^{(b-a-r)}}{p_j} \right) \leq \mu(A \cap B) \leq \mu(A)\mu(B) \sup_{\substack{i,j \\ b \geq N}} \left(\frac{P_{i,j}^{(b-a-r)}}{p_j} \right) \quad (6.1)$$

qualquer que seja o conjunto B pertencente a σ -álgebra ξ_N^∞ .

Demonstração: Inicialmente, observe que $\mathcal{A}(\xi_a^{a+r})$ consiste precisamente de todos os subconjuntos de Σ que podem ser escritos como uma união finita disjunta de cilindros do tipo ${}_a[v_a, \dots, v_{a+r}]$. Assim, pela aditividade da medida μ , é suficiente provar a desigualdade (6.1) para o caso em que $A = {}_a[v_a, \dots, v_{a+r}]$. Fixados $b \geq N$ e $s \geq 0$ inteiros, seja $B = {}_b[v_b, \dots, v_{b+s}]$ com $N \leq b$. Então,

$$\begin{aligned} \mu(A \cap B) &= \sum_{v_{a+r+1}, \dots, v_{b-1}} \mu({}_a[v_a, \dots, v_{a+r}, \dots, v_b, \dots, v_{b+s}]) \\ &= \sum_{v_{a+r+1}, \dots, v_{b-1}} p_{v_a} P_{v_a, v_{a+1}} \cdots P_{v_{b+s-1}, v_{b+s}} \\ &= \mu(A)\mu(B) \frac{1}{p_{v_b}} \left(\sum_{v_{a+r+1}, \dots, v_{b-1}} P_{v_{a+r}, v_{a+r+1}} \cdots P_{v_{b-1}, v_b} \right) \\ &= \mu(A)\mu(B) \frac{P_{v_{a+r}, v_b}^{(b-a-r)}}{p_{v_b}}. \end{aligned}$$

Segue-se daí,

$$\mu(A)\mu(B) \min_{i,j} \left(\frac{P_{i,j}^{(b-a-r)}}{p_j} \right) \leq \mu(A \cap B) \leq \mu(A)\mu(B) \max_{i,j} \left(\frac{P_{i,j}^{(b-a-r)}}{p_j} \right).$$

Pela aditividade da medida μ , a mesma desigualdade acima vale para todo conjunto $B \in \mathcal{A}(\xi_b^{b+s})$. Considere a coleção $\mathcal{C}_A$ de todos os conjuntos $B \in \mathcal{B}_\Sigma$ tais que

$$\mu(A)\mu(B) \inf_{\substack{i,j \\ b \geq N}} \left(\frac{P_{i,j}^{(b-a-r)}}{p_j} \right) \leq \mu(A \cap B) \leq \mu(A)\mu(B) \sup_{\substack{i,j \\ b \geq N}} \left(\frac{P_{i,j}^{(b-a-r)}}{p_j} \right).$$

Claramente a coleção $\mathcal{C}_A$ é uma classe monótona e, pelo que vimos acima, ela contém a álgebra

$$\bigcup_{b=N}^{\infty} \bigcup_{s=0}^{\infty} \mathcal{A}(\xi_b^{b+s}).$$

Assim, pelo Teorema das Classes Monótonas, a classe $\mathcal{C}_A$ contém a σ -álgebra ξ_N^∞ , o que prova o lema. □

Agora, fixe $\varepsilon > 0$ arbitrariamente. Como a matriz P é aperiódica, existe $n_0 \in \mathbb{N}$ tal que

$$\left| \frac{P_{i,j}^n}{p_j} - 1 \right| < \varepsilon$$

quaisquer que sejam $n \geq n_0$ e $i, j \in \{1, \dots, d\}$. Pelo lema anterior, dado um conjunto $A \in \mathcal{A}(\xi_a^{a+r})$ com a e $r \geq 0$ inteiros, temos $|\mu(A \cap B) - \mu(A)\mu(B)| < \varepsilon$ onde $n \in \mathbb{N}$ é suficientemente grande e $B \in \xi_n^\infty$. Em particular, vale que $|\mu(A \cap B) - \mu(A)\mu(B)| < \varepsilon$ para todo $B \in \bigwedge_{n=0}^\infty \xi_n^\infty$. Fixado $B \in \bigwedge_{n=0}^\infty \xi_n^\infty$ arbitrário, considere o conjunto

$$\mathcal{R} := \{A \in \mathcal{B}_\Sigma : |\mu(A \cap B) - \mu(A)\mu(B)| < \varepsilon\}.$$

Observe que $\mathcal{R}$ é uma classe monótona e, além disso, $\mathcal{R}$ contém $\mathcal{A}(\xi_a^{a+r})$ quaisquer que sejam os inteiros a e $r \geq 0$. Assim, $\mathcal{B}_0 \subset \mathcal{R}$ e, pelo Teorema das Classes Monótonas, tem-se $\mathcal{R} = \mathcal{B}_\Sigma$. Isso mostra que $|\mu(A \cap B) - \mu(A)\mu(B)| < \varepsilon$ para todo $B \in \bigwedge_{n=0}^\infty \xi_n^\infty$ e para todo $A \in \mathcal{B}_\Sigma$. Como ε é arbitrário, concluímos que $\mu(A \cap B) = \mu(A)\mu(B)$ para todo $B \in \bigwedge_{n=0}^\infty \xi_n^\infty$ e para todo $A \in \mathcal{B}_\Sigma$. Daí, segue-se que $\mu(B) = \mu(B)^2$ para todo conjunto $B \in \text{Tail}(\xi) = \bigwedge_{n=0}^\infty \xi_n^\infty$ e, por consequência, tem-se $\text{Tail}(\xi) = \xi_\emptyset$. Pela Proposição 4.2.3(b) obtemos que $\mathcal{P}(T) = \text{Tail}(\xi) = \xi_\emptyset$. Logo, pelo Teorema 5.1.2, o Deslocamento de Markov $(\Sigma, \mathcal{B}_\Sigma, \mu, \sigma)$ é um K-automorfismo. □

Se ξ é a partição de Σ em cilindros na posição zero então, usando a prova da Proposição 6.2.1, podemos mostrar que a partição $\xi^- = \xi_0^\infty$ é uma K-partição para o sistema $(\Sigma, \mathcal{B}_\Sigma, \mu, \sigma)$.

6.3 Automorfismo linear do toro

Seja $\mathbb{T}^2$ o toro de dimensão 2, isto é, o quociente $\mathbb{R}^2/\mathbb{Z}^2$ da relação de equivalência $\sim$ em $\mathbb{R}^2$ definida por $x \sim y \Leftrightarrow x - y \in \mathbb{Z}^2$. Sabemos que $\mathbb{T}^2$ herda, do plano $\mathbb{R}^2$, a estrutura de variedade diferenciável de dimensão 2. Além disso, podemos munir $\mathbb{T}^2$ com a métrica riemanniana plana, o que o torna localmente isométrico ao plano $\mathbb{R}^2$. Sejam $\mathcal{B}$ e μ , respectivamente, a σ -álgebra de Borel e a medida de Lebesgue de $\mathbb{T}^2$, que estão associadas a esta métrica riemanniana.

Sejam $A \in \text{SL}(2, \mathbb{R})$ e $T : \mathbb{T}^2 \rightarrow \mathbb{T}^2$ a aplicação induzida por A (isto é, $T([x]) = [Ax]$). Observe que T é diferenciável e o jacobiano $\det(DT([x]))$ é localmente constante igual a 1, uma vez que a derivada $DT([x])$ em cada ponto está canonicamente identificada com A . Claramente, a transformação T é invertível, onde a sua inversa T^{-1} é a transformação induzida pela matriz $A^{-1} \in \text{SL}(2, \mathbb{R})$. Afirmamos que T preserva a medida de Lebesgue em $\mathbb{T}^2$. De fato, sendo T um difeomorfismo local, a pré-imagem de qualquer conjunto mensurável B , com diâmetro suficientemente pequeno, é formada por $|\det(A)| = \text{grau}(T)$ conjuntos disjuntos D_i difeomorfos ao conjunto B . Pela fórmula de mudança de variável, temos que $\mu(B) = |\det(A)| \cdot \mu(D_i)$ para todo conjunto mensurável B com diâmetro suficientemente pequeno. Isto mostra que $\mu(B) = \mu(T^{-1}(B))$ para todo conjunto mensurável B com diâmetro suficientemente pequeno. Logo, T preserva a medida μ , como afirmado. O automorfismo $(\mathbb{T}^2, \mathcal{B}, \mu, T)$ é chamado de *automorfismo linear do toro*.

Suponha que a transformação T é um difeomorfismo de Anosov, isto é, suponha que a matriz A é uma matriz hiperbólica. Neste caso, chamamos $(\mathbb{T}^2, \mathcal{B}, \mu, T)$ de automorfismos linear hiperbólico do toro.

Todo automorfismo linear hiperbólico é ergódico. Mais ainda, ele é um K-automorfismo. Este fato é um corolário imediato do Teorema 6.3.1, que é proveniente da Teoria de Dinâmica Hiperbólica, apresentado abaixo.

Teorema 6.3.1 *Seja $f : M \rightarrow M$ um difeomorfismo Anosov de classe $C^{1+\alpha}$ ($0 < \alpha < 1$) que preserva o volume de uma variedade Riemanniana compacta M . Então f é um K-automorfismo.*

Demonstração: Ver [7, Theorem 4.8].

□

REFERÊNCIAS

- [1] CORNFELD, I. P.; FOMIN, S. V.; SINAI, Ya. G. **Ergodic theory**. Traduzido do russo para inglês por A. B. Sosinskii. New York: Springer-Verlag, 1982. [Russian Mathematical Surveys, v. 245].
- [2] EINSIEDLER, M.; LINDENSTRAUSS, E.; WARD, T. **Entropy in Ergodic Theory and Topological Dynamics**. [S. l.], 2023. Disponível em: <https://tbward0.wixsite.com/books/entropy>. Acesso em: 19 set. 2023.
- [3] EINSIEDLER, M.; WARD, T. **Ergodic theory with a view towards number theory**. Londres: Springer Verlag, 2011. (Graduate Texts in Mathematics, v. 259).
- [4] OLIVEIRA, K.; VIANA, M. **Fundamentos da teoria ergódica**. 2 ed. Rio de Janeiro: IMPAR, 2019.
- [5] PARRY, W. **Entropy and generators in ergodic theory**. Amsterdam: W. A. Benjamin, 1969.
- [6] PETERSEN, K. **Ergodic theory**. Cambridge: Cambridge University Press, 1989. Cambridge Studies in Advanced Mathematics, v. 2). Reedição corrigida do original de 1983.
- [7] PONCE, G.; VARÃO, R.; **An introduction to the Kolmogorov-Bernoulli equivalence**. Zug, Chem: Springer, 2019.
- [8] ROKHLIN, V. A. Lectures on the entropy theory of measure-preserving transformations. **Russian Mathematical Surveys**, v. 22, n. 5, p.1-52, out. 1967.
- [9] SARIG, O. **Lecture Notes on Ergodic Theory**, 2023. Disponível em: <https://www.weizmann.ac.il/math/sarigo/sites/math.sarigo/files/uploads/ergodicnotes.pdf>. Acesso em: 04 ago. 2023.
- [10] WALTERS, P. **An Introduction to Ergodic Theory**. New York: Springer-Verlag, 1982. (Graduate Texts in Mathematics, v. 79).