UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS FÍSICAS E MATEMÁTICAS
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA PURA E APLICADA

Luiz Guilherme de Carvalho Lopes

Entropia completamente positiva e K-automorfismos

Florianópolis
2024

Luiz Guilherme de Carvalho Lopes

Entropia completamente positiva e K-automorfismos

Dissertação submetida ao Programa de Pós-Graduação em Matemática Pura e Aplicada da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Análise.
Orientador: Prof. Rômulo Maia Vermersch, Dr.

Florianópolis
2024

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Lopes, Luiz Guilherme de Carvalho
Entropia completamente positiva e K-automorfismos Florianópolis / Luiz Guilherme de Carvalho Lopes ; orientador, Rômulo Maia Vermersch, 2024.
96 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Ciências Físicas e Matemáticas, Programa de Pós-Graduação em Matemática Pura e Aplicada, Florianópolis, 2024.

Inclui referências.

1. Matemática Pura e Aplicada. 2. Entropia e K automorfismos. 3. Sistemas dinâmicos. 4. Espaços de probabilidade separáveis. I. Vermersch, Rômulo Maia . II. Universidade Federal de Santa Catarina. Programa de Pós Graduação em Matemática Pura e Aplicada. III. Título.

Luiz Guilherme de Carvalho Lopes

Entropia completamente positiva e K-automorfismos

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Nilson da Costa Bernardes Junior, Dr.
Universidade Federal do Rio de Janeiro

Prof. Paulo Mendes de Carvalho Neto, Dr.
Universidade Federal de Santa Catarina

Prof. Rômulo Maia Vermersch, Dr.
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de Mestre em Análise.

ICP-Edu
Documento assinado digitalmente
Douglas Soares Goncalves
Data: 14/02/2024 11:15:47-0300
CPF: ***.566.318-**
Verifique as assinaturas em https://v.ufsc.br

Coordenação do Programa de Pós-Graduação

GOV.BR
Documento assinado digitalmente
ROMULO MAIA VERMERSCH
Data: 14/02/2024 10:36:17-0300
CPF: ***.515.047-**
Verifique as assinaturas em https://v.ufsc.br

Prof. Rômulo Maia Vermersch, Dr.
Orientador

Florianópolis, 2024.

AGRADECIMENTOS

Inicialmente a minha família, em especial minha mãe Ligia Maria dos Santos Carvalho Lopes, meu falecido pai Luiz Lopes e meu irmão Lucas, que sempre mantiveram o apoio, carinho, compreensão e sempre se fizeram presentes durante toda a minha vida acadêmica.

A todos os amigos que fiz durante essa trajetória e a todos meus colegas da pós-graduação pela diversão, risadas, jogos de tabuleiro e bons serviços realizados.

Aos meus colegas da UFSC, que sempre incentivaram o bom estudo e a troca de conhecimento. À minha amiga Gabrielle Goulart Bortolotto, por nossas conversas sobre matemática sempre acompanhadas de um bom café, que proporcionaram tardes agradáveis e boas risadas. Aos meus amigos da graduação João Guilherme Fritsche Colombo, Gabriel Lese Pereira, Ciro Igor Gonçalves da Silva, Andrieli Pscheidt, João Victor Brochier Gonçalves e Fernanda Meenken Berti que fizeram meus dias mais alegres.

A todos os professores que estiveram presentes no meu amadurecimento intelectual e profissional, destaco especialmente o professor Paulo Mendes de Carvalho Neto, sempre disposto a contribuir com seu conhecimento e admirar o rigor matemático.

Finalmente, expresso minha gratidão aos professores que aceitaram compor a banca avaliadora, ao CNPq por fomentar financeiramente o meu mestrado acadêmico e ao meu orientador Rômulo Maia Vermersch, que me introduziu à noção de pesquisa e, graças a ele, fez este trabalho possível.

RESUMO

Neste trabalho apresentamos a caracterização, devida a Rokhlin e Sinai, dos K-automorismos em termos de sua entropia relativamente a partições mensuráveis finitas.

Palavras-chave: Entropia, K-automorfismos, espaços de probabilidade separáveis.

ABSTRACT

In this work we present the characterization, due to Rokhlin and Sinai, of the K-automorphisms in terms of its entropy relatively to finite measurable partitions.

Keywords: Entropy, K-automorphisms, separable probability spaces.

SUMÁRIO

INTRODUÇÃO

Na natureza, muitos fenômenos podem ser descritos utilizando-se o formalismo de um sistema dinâmico. Geralmente, em tais fenômenos naturais há um número muito grande de partículas envolvidas, o que torna impraticável escrever e resolver as equações diferenciais correspondentes. Por outro lado, quando há uma medida finita (por exemplo, uma probabilidade) invariante pela ação do sistema, pode-se obter informações de longo prazo sobre o sistema de um ponto de vista estatístico que prescinde da resolução das equações diferenciais envolvidas. Este é o caso nos sistemas hamiltonianos, os quais descrevem a evolução de sistemas conservativos em mecânica newtoniana.

No presente trabalho um sistema dinâmico é uma quádrupla $(X,\mathcal{X},\mu,T)$, onde $(X,\mathcal{X},\mu)$ é um espaço de probabilidade e $T : X \to X$ é uma transformação mensurável bijetiva cuja inversa também é mensurável e $\mu(T^{-1}A) = \mu(A)$ para cada $A \in \mathcal{X}$. Em outras palavras, a ação de T deixa invariante a medida μ, ou ainda, μ é T-invariante.

Em nosso trabalho abordaremos o importante conceito de entropia, originalmente introduzido pelo físico e matemático alemão Rudolf Clausius em 1854, no contexto da Termodinâmica, como uma medida da "desordem" do sistema.

Em 1948 o engenheiro elétrico americano Claude Shannon iniciou um estudo matemático sistemático sobre transmissão de informação e o conceito principal utilizado foi a entropia de uma fonte emissora de informação, desta vez como uma medida de incerteza sobre a informação emitida por tal fonte.

Cerca de uma década depois do trabalho de Shannon, o matemático russo Andrey Kolmogorov introduziu o conceito de entropia em Teoria Ergódica como um invariante de sistemas isomorfos (isto é, equivalentes). Com tal invariante à disposição, Kolmogorov foi capaz de mostrar que o 2-deslocamento de Bernoulli e o 3-deslocamento de Bernoulli não são isomorfos, o que consistia em um problema básico da Teoria Ergódica que resistiu mesmo aos métodos espectrais de von Neumann e Halmos.

É importante salientar que o matemático russo (e aluno de Kolmogorov) Yakov Sinai, entre outras enormes contribuições, simplificou e sistematizou o conceito de entropia em Teoria Ergódica e, após isto, tal conceito ficou conhecido como entropia de Kolmogorov-Sinai.

Essa dissertação possui quatro capítulos. No primeiro capítulo apresentamos alguns preliminares da Teoria da Medida e Integração que são necessários à leitura do presente trabalho. Não apresentamos as demonstrações de todos os fatos preliminares, uma vez que isso tornaria o texto muito longo e, além disso, existem boas referências para as demonstrações que não apresentamos, as quais indicamos ao longo do texto. No capítulo 2, o mais curto, apresentamos a construção do sistema dinâmico com entropia positiva mais importante da Teoria Ergódica: o deslocamento de Bernoulli. Dada a sua relevância, apresentamos a sua construção matemática completa. No capítulo 3 apresentamos os elementos da Teoria de Entropia de Kolmogorov-Sinai que são fundamentais para o entendimento do último capítulo. Aqui cabe salientar que alguns resultados clássicos importantes da Teoria de Entropia, por não serem indispensáveis ao andamento do texto, ficaram de fora da presente exposição; o mais notável deles talvez seja o Teorema de Shannon-McMillan-Breiman, resultado fundamental da teoria que fornece uma interpretação da entropia de sistemas ergódicos em termos combinatoriais. No quarto capítulo, o último do presente trabalho, apresentamos os K-automorfismos e uma demonstração para o belo Teorema de Rokhlin-Sinai, o qual fornece uma caracterização para tais transformações em termos de sua entropia relativa a partições mensuráveis finitas do espaço sobre o qual elas agem. Mais precisamente, o Teorema de Rokhlin-Sinai garante que um sistema dinâmico $(X,\mathcal{X},\mu,T)$ é um K-automorfismo se, e somente se, a entropia de T relativa a qualquer partição mensurável finita do espaço X é positiva. Por fim observamos que, por questões de espaço, as propriedades de mistura dos K-automorfismos não constam na presente exposição.

1 PRELIMINARES

Neste capítulo apresentaremos a terminologia básica que será utilizada neste trabalho.

1.1 MEDIDAS E TRANSFORMAÇÕES MENSURÁVEIS

No que segue $(X,\mathcal{X},\mu)$ denotará um espaço de probabilidade. Para esta e demais noções básicas sobre a Teoria da Medida, veja por exemplo (RUDIN, 1987) ou (HALMOS, Paul, 1974). Para as noções básicas sobre espaços métricos e também espaços topológicos recomendamos (DUGUNDJI, 1966).

Definição 1.1 . Dizemos que uma transformação $T : X \to X$ é *mensurável* se $T^{-1}(A) \in \mathcal{X}$ sempre que $A \in \mathcal{X}$.

Definição 1.2 . Dizemos que uma transformação mensurável $T : X \to X$ *preserva* a medida μ ou, equivalentemente, que μ é T-invariante, se $\mu(T^{-1}(A)) = \mu(A)$ para todo $A \in \mathcal{X}$.

Definição 1.3 . Um *automorfismo* sobre $(X,\mathcal{X},\mu)$ é uma bijeção mensurável $T : X \to X$ que preserva μ cuja inversa T^{-1} também é mensurável (e preserva μ).

Definição 1.4 . Um *sistema dinâmico* é uma quádrupla $(X,\mathcal{X},\mu,T)$, onde $(X,\mathcal{X},\mu)$ é um espaço de probabilidade e $T : X \to X$ é um automorfismo.

Definição 1.5 . Uma *medida de Borel* sobre um espaço métrico X é uma medida μ definida sobre uma σ-álgebra que contém todos os subconjuntos abertos (ou, equivalentemente, fechados) do espaço X.

Vamos agora enunciar um resultado que remonta ao início do estudo das transformações que preservam uma dada medida. Tal resultado mostra, em particular, que é possível encontrar uma infinidade dessas transformações.

Teorema 1.6 . (Liouville) Seja $F : U \to \mathbb{R}^n$ campo vetorial de classe C^1 definido no aberto $U \subset \mathbb{R}^n$. Considere a equação diferencial autônoma $x' = F(x)$ e denote por $\phi^t : U \to U$ o fluxo de classe C^1 associado. Suponha que ϕ^t esteja definido para todo $t \in \mathbb{R}$.

Então, o fluxo $(\phi^t)_{t\in\mathbb{R}}$ preserva o volume em U se, e somente se, div $F(x) = 0$ para todo $x \in U$. Ou seja, se μ denota a medida de volume (isto é, a medida de Lebesgue), então $\mu(\phi^t(A)) = \mu(A)$ para todo conjunto Lebesgue mensurável $A \subset U$ e para todo $t \in \mathbb{R}$.

Em particular, a transformação de tempo um $\phi^1 : U \to U$ preserva o volume sempre que div F é identicamente nulo em U.

Demonstração. Veja a página 149 de (NADKARNI, 2013).

□

Definição 1.7 . Dizemos que uma álgebra $\mathcal{A}$ de subconjuntos de X *gera* a σ-álgebra $\mathcal{X}$ se $\mathcal{X}$ é a menor σ-álgebra de subconjuntos de X que contém $\mathcal{A}$. Neste caso, denotamos isso por $\sigma(\mathcal{A}) = \mathcal{X}$. Mais geralmente, dada qualquer coleção $\mathcal{C}$ de subconjuntos de X denotamos por $\sigma(\mathcal{C})$ a menor σ-álgebra sobre X que contém $\mathcal{C}$.

Abaixo temos um resultado que nos permite verificar a invariância de uma probabilidade no caso em que existe uma álgebra geradora.

Proposição 1.8 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade e $\mathcal{A}$ uma álgebra que gera $\mathcal{X}$. Se $T : X \to X$ é uma transformação mensurável tal que $\mu(T^{-1}A) = \mu(A)$ para todo $A \in \mathcal{A}$, então $\mu(T^{-1}B) = \mu(B)$ para todo $B \in \mathcal{X}$.

Demonstração. Veja o Lemma 1.3.1 da página 11 de (VIANA; OLIVEIRA, 2016). □

Definição 1.9 . Agora recorde que, dados dois conjuntos A e B, sua *diferença simétrica* é dada por $A \triangle B := (A \setminus B) \cup (B \setminus A)$.

Seja $(X,\mathcal{X},\mu)$ um espaço de probabilidade e defina $d(A,B) := \mu(A \triangle B) \ \ (A,B \in \mathcal{X})$. Usando a identificação $A \simeq B \iff \mu(A \triangle B) = 0$ não é difícil verificar que d é uma métrica sobre $\mathcal{X}$. Ou seja, neste caso, $(\mathcal{X},d)$ é um espaço métrico. Também usaremos a notação $A = B \pmod \mu$ quando $\mu(A \triangle B) = 0$ e, mais geralmente, dadas duas sub-σ-álgebras $\mathcal{A}$ e $\mathcal{B}$, escreveremos $\mathcal{A} \subset \mathcal{B} \pmod \mu$ quando para cada $A \in \mathcal{A}$ existir $B \in \mathcal{B}$ tal que $\mu(A \triangle B) = 0$. Além disso, escreveremos $\mathcal{A} = \mathcal{B} \pmod \mu$ quando $\mathcal{A} \subset \mathcal{B} \pmod \mu$ e $\mathcal{B} \subset \mathcal{A} \pmod \mu$. Não havendo perigo de confusão, usaremos apenas o símbolo de inclusão ou de igualdade em lugar do símbolo de inclusão ou igualdade $(\mathrm{mod}\ \mu)$.

Definição 1.10 . Dizemos que um espaço de probabilidade $(X,\mathcal{X},\mu)$ é *separável* se o espaço métrico $(\mathcal{X},d)$ é separável.

Exemplo 1.11 . Dados o conjunto $[0,1]$, $\mathcal{L}$ a σ-álgebra dos subconjuntos Lebesgue mensuráveis de $[0,1]$ e λ a medida de Lebesgue sobre $[0,1]$, tem-se que $([0,1],\mathcal{L},\lambda)$ é um espaço de probabilidade separável.

Mais geralmente, suponha que $(X,\mathcal{X},\mu)$ é um espaço de probabilidade para o qual existem subconjuntos mensuráveis $X_0 \subset X$ e $L_0 \subset [0,1]$ com $\mu(X_0) = 1$ e $\lambda(L_0) = 1$, e um automorfismo $\phi : X_0 \to L_0$ que satisfaz $\lambda(\phi(A)) = \mu(A)$ para todo subconjunto mensurável $A \subset X_0$ e $\mu(\phi^{-1}(B)) = \lambda(B)$ para todo subconjunto mensurável $B \subset L_0$. Neste caso, dizemos que $(X,\mathcal{X},\mu)$ é *isomorfo mod 0* a $([0,1],\mathcal{L},\lambda)$, ou ainda, um *espaço de Lebesgue*. Observe que, como $([0,1],\mathcal{L},\lambda)$ é um espaço de probabilidade separável, segue que todo espaço de Lebesgue é um espaço de probabilidade separável.

O próximo resultado garante que a maioria dos espaços de probabilidade com os quais trabalhamos são espaços de probabilidade separáveis.

Teorema 1.12 . Se X é um espaço métrico separável completo, $\mathcal{X}$ é a σ-álgebra de Borel sobre X e μ é uma medida de Borel probabilística não-atômica, então $(X,\mathcal{X},\mu)$ é um espaço de Lebesgue.

Demonstração. Veja o Theorem 8.5.4 na página 237 de (VIANA; OLIVEIRA, 2016). □

Abaixo temos um importante resultado que nos garante que, no caso de um espaço de probabilidade é possível aproximar, na métrica d, um conjunto mensurável qualquer por um conjunto de uma álgebra geradora. Mais precisamente, vale:

Teorema 1.13 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade e $\mathcal{A}$ uma álgebra geradora. Então, para cada $B \in \mathcal{X}$ e $\varepsilon > 0$ existe $A \in \mathcal{A}$ tal que $\mu(B \triangle A) < \varepsilon$.

Demonstração. Veja Teorema D na página 56 de (HALMOS, Paul, 1956). □

Seja $(A_n)_{n=1}^{\infty}$ uma sequência de conjuntos mensuráveis. Então, definimos $\liminf A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k$ e $\limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$.

Proposição 1.14 . Sejam $(X,\mathcal{X},\mu)$ espaço de probabilidade e $(A_n)_{n=1}^{\infty}$ sequência de conjuntos mensuráveis. Então,

$$\limsup \mu(A_n) \leqslant \mu(\limsup A_n).$$

Demonstração. Defina $F_n = \bigcup_{i=n}^{\infty} A_i$. Pela definição tem-se, $F_{n+1} \subset F_n$ para todo $n \geqslant 1$, donde a sequência $(F_n)_{n=1}^{\infty}$ é monótona decrescente. Daí, como μ é medida finita,

$$\begin{aligned} \mu(\limsup A_n) &= \mu\Big(\bigcap_{n=1}^{\infty} F_n\Big) \\ &= \lim_{n\to\infty} \mu(F_n). \end{aligned} \tag{1}$$

Seja $x_n = \sup_{m \geqslant n} \mu(A_m) \quad (n \geqslant 1)$. Note que $(x_n)_{n\geqslant 1}$ é uma sequência monótona decrescente e, portanto,

$$\begin{aligned} \limsup \mu(A_n) &= \inf_{n\geqslant 1} \sup_{m\geqslant n} \mu(A_m) \\ &= \lim_{n\to\infty} x_n. \end{aligned} \tag{2}$$

Note que $\mu(F_n) = \mu\Big(\bigcup_{m=n}^{\infty} A_m\Big) \geqslant \mu(A_m)$ para todo $m \geqslant n$ e, portanto, $x_n \leqslant \mu(F_n)$ para todo $n \geqslant 1$. Disto e das relações (1) e (2) obtém-se:

$$\begin{aligned} \limsup \mu(A_n) &= \lim_{n\to\infty} x_n \\ &\leqslant \lim_{n\to\infty} \mu(F_n) \\ &= \mu(\limsup A_n). \end{aligned}$$

□

Proposição 1.15 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade e $(A_n)_{n=1}^{\infty}$ uma sequência de conjuntos mensuráveis tal que $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. Então, $\mu(\limsup A_n) = 0$.

Demonstração. Dado $\varepsilon > 0$ arbitrário existe $n_0 \geqslant 1$ tal que,

$$\sum_{n \geqslant n_0} \mu(A_n) < \varepsilon.$$

Seja $F_m = \bigcup_{n=m}^{\infty} A_n$. Note que a sequência $(F_m)_{m=1}^{\infty}$ é monótona decrescente. Como μ é finita, temos:

$$\begin{aligned}
\mu(\limsup A_n) &= \mu(\bigcap_{m=1}^{\infty} F_m) \\
&= \lim_{m\to\infty} \mu(F_m) \\
&= \lim_{m\to\infty} \mu(\bigcup_{n=m}^{\infty} A_n) \\
&\leqslant \sum_{n\geqslant n_0} \mu(A_n) \\
&< \varepsilon.
\end{aligned}$$

□

Proposição 1.16 . Se $(X,\mathcal{X},\mu)$ é um espaço de probabilidade, $B \in \mathcal{X}$ e $\{A_n\}_{n=1}^{\infty} \subset \mathcal{X}$ são tais que $\mu(B\triangle A_n) < 2^{-n}$ para todo $n \geqslant 1$, então

$$\mu(B\triangle \liminf A_n) = 0.$$

Demonstração. Escreva $A = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty} A_k = \liminf A_n$.

Note que $B \setminus A = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}(B \setminus A_k)$ e $B \setminus A_k \subset B\triangle A_k$ para todo $k \geqslant 1$. Portanto,

$$\mu(B \setminus A_k) < 2^{-k} \text{ para todo } k \geqslant 1,$$

Daí, segue da Proposição 1.15 que $\mu(B \setminus A) = \mu(\limsup(B \setminus A_n)) = 0$.

Por outro lado, temos $A \setminus B = \bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}(A_k \setminus B)$ e $\mu(A_n\triangle B) < 2^{-n}$ para todo $n \geqslant 1$, donde novamente pela Proposição 1.15 temos

$$\mu(A \setminus B) = \mu(\liminf(A_n \setminus B)) \leqslant \mu(\limsup(A_n \setminus B)) = 0.$$

Concluímos que $\mu(B\triangle A) = 0$, como desejado.

□

Definição 1.17 . Dadas duas σ-álgebras $\mathcal{A}_1$ e $\mathcal{A}_2$ definimos $\mathcal{A}_1 \vee \mathcal{A}_2 := \sigma(\mathcal{A}_1 \cup \mathcal{A}_2)$ e, mais geralmente, dada $(\mathcal{A}_n)_{n\geqslant 1}$ uma sequência de σ-álgebras, definimos

$$\bigvee_{n=1}^{\infty} \mathcal{A}_n = \sigma\Big(\bigcup_{n\geqslant 1} \mathcal{A}_n\Big)$$

1.2 PARTIÇÕES E σ-ÁLGEBRAS

Em nossa exposição a noção de *partição* desempenhará um papel crucial. Vejamos sua definição:

Definição 1.18 . Uma *partição* do espaço de probabilidade $(X,\mathcal{X},\mu)$ é uma coleção $\mathcal{P} \subset \mathcal{X}$ que satisfaz:

(i) $\mu(P) > 0$ para cada $P \in \mathcal{P}$;

(ii) $\mu(P \cap Q) = 0$ para cada $P,Q \in \mathcal{P}$ com $P \neq Q$;

(iii) $\mu(\bigcup_{P\in\mathcal{P}} P) = 1$.

Observação 1.19 . Se $\mathcal{P}$ é uma partição de um espaço de probabilidade, então $\mathcal{P}$ é necessariamente uma coleção enumerável. Portanto, ao excluirmos um conjunto com medida nula, se necessário, podemos supor que $P \cap Q = \varnothing$ sempre que $P \neq Q$ são elementos de $\mathcal{P}$.

Observação 1.20 . Se $\mathcal{P}$ é uma partição do espaço de probabilidade $(X,\mathcal{X},\mu)$ e $T : X \to X$ é uma transformação que preserva a medida μ, então $T^{-1}\mathcal{P} := \{T^{-1}(P) : P \in \mathcal{P}\}$ é uma partição de $(X,\mathcal{X},\mu)$. Além disso, se T é um automorfismo, então $T\mathcal{P} := \{T(P) : P \in \mathcal{P}\}$ também é uma partição de $(X,\mathcal{X},\mu)$.

Definição 1.21 . Seja $\mathcal{P}$ uma partição. Dizemos que $P \in \mathcal{P}$ é um *átomo* da partição $\mathcal{P}$. Além disso, denotamos por $\widehat{P}$ a menor σ-álgebra que contém os átomos de $\mathcal{P}$. Ou seja, $\widehat{P} = \sigma(\mathcal{P})$.

Observação 1.22 . Se $\mathcal{P}$ é uma partição, então $\widehat{\mathcal{P}}$ é precisamente a coleção de todas as uniões de átomos de $\mathcal{P}$ (incluindo o conjunto vazio $\varnothing$).

Definição 1.23 . Sejam $\mathcal{P}$ e $\mathcal{Q}$ partições. Dizemos que $\mathcal{P}$ é *mais fina* do que $\mathcal{Q}$, o que denotaremos por $\mathcal{Q} \prec \mathcal{P}$, quando todo átomo de $\mathcal{Q}$ puder ser escrito, a menos de um conjunto de medida nula, como uma união de átomos de $\mathcal{P}$. Note que $\mathcal{Q} \prec \mathcal{P} \iff \widehat{\mathcal{Q}} \subset \widehat{\mathcal{P}} \pmod{\mu}$.

Denotamos por $\mathcal{P} \vee \mathcal{Q}$ a coleção de todos os elementos da forma $P \cap Q$ com $P \in \mathcal{P}$ e $Q \in \mathcal{Q}$ tais que $\mu(P \cap Q) > 0$. Neste caso, $\mathcal{P} \vee \mathcal{Q}$ é uma partição, dita o *menor refinamento comum* de $\mathcal{P}$ e $\mathcal{Q}$.

Mais geralmente, dadas partições $\mathcal{P}_1, \ldots ,\mathcal{P}_k$ definimos similarmente a partição

$$\bigvee_{i=1}^{k} \mathcal{P}_i := \mathcal{P}_1 \vee \ldots \vee \mathcal{P}_k.$$

Proposição 1.24 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade, $\mathcal{P}$ e $\mathcal{Q}$ partições e $T : X \to X$ uma transformação que preserva μ. Valem:

(i) $\widehat{\mathcal{P} \vee \mathcal{Q}} = \widehat{\mathcal{P}} \vee \widehat{\mathcal{Q}}$;

(ii) $\widehat{T^{-1}\mathcal{P}} = T^{-1}\widehat{\mathcal{P}}$;

(iii) $\widehat{T^{-1}(\mathcal{P} \vee \mathcal{Q})} = T^{-1}(\widehat{\mathcal{P} \vee \mathcal{Q}}) = T^{-1}\widehat{\mathcal{P}} \vee T^{-1}\widehat{\mathcal{Q}}$.

Além disso, se T é um automorfismo, então $\widehat{T\mathcal{P}} = T\widehat{\mathcal{P}}$ e $\widehat{T(\mathcal{P} \vee \mathcal{Q})} = T(\widehat{\mathcal{P} \vee \mathcal{Q}}) = T\widehat{\mathcal{P}} \vee T\widehat{\mathcal{Q}}$.

Demonstração. (i): Como $\mathcal{P} \subset \widehat{\mathcal{P}}$ e $\mathcal{Q} \subset \widehat{\mathcal{Q}}$, temos $\mathcal{P} \cup \mathcal{Q} \subset \widehat{\mathcal{P}} \cup \widehat{\mathcal{Q}}$. Daí, como $\sigma(\mathcal{P} \cup \mathcal{Q}) = \sigma(\mathcal{P} \vee \mathcal{Q}) = \widehat{\mathcal{P} \vee \mathcal{Q}}$ e $\sigma(\widehat{\mathcal{P}} \cup \widehat{\mathcal{Q}}) = \widehat{\mathcal{P}} \vee \widehat{\mathcal{Q}}$, concluímos que $\widehat{\mathcal{P} \vee \mathcal{Q}} \subset \widehat{\mathcal{P}} \vee \widehat{\mathcal{Q}}$. Reciprocamente, tome $R \in \widehat{\mathcal{P}} \cup \widehat{\mathcal{Q}}$ arbitrário, digamos $R \in \widehat{\mathcal{P}}$. Como $\widehat{\mathcal{P}} \subset \widehat{\mathcal{P} \vee \mathcal{Q}}$, temos $R \in \widehat{\mathcal{P} \vee \mathcal{Q}}$. Consequentemente, como $R \in \widehat{\mathcal{P}} \cup \widehat{\mathcal{Q}}$ é arbitrário, temos $\widehat{\mathcal{P}} \vee \widehat{\mathcal{Q}} = \sigma(\widehat{\mathcal{P}} \cup \widehat{\mathcal{Q}}) \subset \widehat{\mathcal{P} \vee \mathcal{Q}}$.

(ii): Apenas note que $\widehat{T^{-1}\mathcal{P}} = \sigma(T^{-1}\mathcal{P}) = T^{-1}(\sigma(\mathcal{P})) = T^{-1}\widehat{\mathcal{P}}$. Se T é um automorfismo, então aplicamos o que acabamos de ver com T^{-1} em lugar de T e obtemos $\widehat{T\mathcal{P}} = T\widehat{\mathcal{P}}$.

(iii): Usando (i) e (ii) obtém-se:

$$\begin{aligned}
\widehat{T^{-1}(\mathcal{P} \vee \mathcal{Q})} &= T^{-1}(\widehat{\mathcal{P} \vee \mathcal{Q}}) \\
&= T^{-1}(\widehat{\mathcal{P}} \vee \widehat{\mathcal{Q}}) \\
&= T^{-1}(\sigma(\widehat{\mathcal{P}} \cup \widehat{\mathcal{Q}})) \\
&= \sigma(T^{-1}(\widehat{\mathcal{P}} \cup \widehat{\mathcal{Q}})) \\
&= \sigma(T^{-1}\widehat{\mathcal{P}} \cup T^{-1}\widehat{\mathcal{Q}}) \\
&= T^{-1}\widehat{\mathcal{P}} \vee T^{-1}\widehat{\mathcal{Q}}.
\end{aligned}$$

Novamente, se T é um automorfismo, então aplicando o que acabamos de ver com T^{-1} em lugar de T obtemos $\widehat{T(\mathcal{P} \vee \mathcal{Q})} = T(\widehat{\mathcal{P} \vee \mathcal{Q}}) = T\widehat{\mathcal{P}} \vee T\widehat{\mathcal{Q}}$. $\square$

Seja $(\mathcal{P}_n)_{n \geqslant 1}$ uma sequência de partições. Definimos $\bigvee\limits_{n \geqslant 1} \mathcal{P}_n := \sigma\Big(\bigcup\limits_{n \geqslant 1} \mathcal{P}_n\Big)$. Como $\widehat{\mathcal{P}_n}$ é precisamente a coleção de todas as uniões de átomos de $\mathcal{P}_n$ (incluindo o conjunto vazio $\varnothing$), segue que $\bigvee\limits_{n \geqslant 1} \mathcal{P}_n = \sigma\Big(\bigcup\limits_{n \geqslant 1} \mathcal{P}_n\Big) = \sigma\Big(\bigcup\limits_{n \geqslant 1} \widehat{\mathcal{P}_n}\Big) = \bigvee\limits_{n \geqslant 1} \widehat{\mathcal{P}_n}$. Daí, embora $\bigvee\limits_{n \geqslant 1} \mathcal{P}_n$ não necessariamente seja uma partição no sentido da Definição 1.18, usaremos o seguinte abuso de notação: $\widehat{\bigvee\limits_{n \geqslant 1} \mathcal{P}_n} = \bigvee\limits_{n \geqslant 1} \widehat{\mathcal{P}_n} = \bigvee\limits_{n \geqslant 1} \mathcal{P}_n$. Em outras palavras, $\widehat{\bigvee\limits_{n \geqslant 1} \mathcal{P}_n}$ denotará a menor σ-álgebra que contém os átomos de $\mathcal{P}_n$ para todo $n \geqslant 1$.

Além disso, se $T : X \to X$ é uma transformação que preserva μ, então $T^{-1}\mathcal{P}_n$ é uma partição para cada $n \geqslant 1$, donde pelo abuso de notação acima e de (ii) da proposição anterior obtemos:

$$\begin{aligned}
\widehat{\bigvee_{n \geqslant 1} T^{-1}\mathcal{P}_n} &= \bigvee_{n \geqslant 1} \widehat{T^{-1}\mathcal{P}_n} \\
&= \bigvee_{n \geqslant 1} T^{-1}\widehat{\mathcal{P}_n} \\
&= \sigma\Big(\bigcup_{n \geqslant 1} T^{-1}\widehat{\mathcal{P}_n}\Big) \\
&= \sigma\Big(\bigcup_{n \geqslant 1} \widehat{T^{-1}\mathcal{P}_n}\Big) \\
&= \sigma\Big(\bigcup_{n \geqslant 1} T^{-1}\mathcal{P}_n\Big) \\
&= \sigma\Big(T^{-1}\Big(\bigcup_{n \geqslant 1} \mathcal{P}_n\Big)\Big)
\end{aligned}$$

$$= T^{-1}\Big(\sigma\Big(\bigcup_{n\geqslant 1}\mathcal{P}_n\Big)\Big)$$
$$= T^{-1}\Big(\bigvee_{n\geqslant 1}\mathcal{P}_n\Big)$$
$$= T^{-1}\Big(\widehat{\bigvee_{n\geqslant 1}\mathcal{P}_n}\Big)$$

Consequentemente, como $\widehat{\bigvee_{n\geqslant 1} T^{-1}\mathcal{P}_n}$ denota a menor σ-álgebra que contém os átomos de $T^{-1}\mathcal{P}_n$ para todo $n \geqslant 1$, também usaremos o seguinte abuso de notação:

$$\widehat{T^{-1}\Big(\bigvee_{n\geqslant 1}\mathcal{P}_n\Big)} = T^{-1}\Big(\widehat{\bigvee_{n\geqslant 1}\mathcal{P}_n}\Big) = \widehat{\bigvee_{n\geqslant 1} T^{-1}\mathcal{P}_n}.$$

Similarmente, se T é um automorfismo, então

$$\widehat{T\Big(\bigvee_{n\geqslant 1}\mathcal{P}_n\Big)} = T\Big(\widehat{\bigvee_{n\geqslant 1}\mathcal{P}_n}\Big) = \widehat{\bigvee_{n\geqslant 1} T\mathcal{P}_n}.$$

Definição 1.25 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade, $(\mathcal{A}_n)_{n=1}^{\infty}$ uma sequência de sub-σ-álgebras e $\mathcal{A}$ uma sub-σ-álgebra. Escrevemos $\mathcal{A}_n \uparrow \mathcal{A}$ se $\mathcal{A}_n \subset \mathcal{A}_{n+1} \subset \mathcal{A}$ para todo $n \geqslant 1$ e $\mathcal{A} = \bigvee_{n=1}^{\infty} \mathcal{A}_n \pmod{\mu}$.

Observação 1.26 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade, $\mathcal{P}$ uma partição e $T : X \to X$ uma transformação que preserva μ. Como $\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}}$ é a menor σ-álgebra que contém os átomos de $T^{-j}\mathcal{P}$ para cada $j \geqslant 1$, segue que $\bigvee_{j=1}^{k} \widehat{T^{-j}\mathcal{P}} \uparrow \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}}$. Além disso, se T é um automorfismo, então também vale $\bigvee_{j=1}^{k} \widehat{T^{j}\mathcal{P}} \uparrow \bigvee_{j=1}^{\infty} \widehat{T^{j}\mathcal{P}}$.

Proposição 1.27 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade, $(\mathcal{A}_n)_{n=1}^{\infty}$ uma sequência de sub-σ-álgebras e $\mathcal{A}$ uma sub-σ-álgebra tais que $\mathcal{A}_n \subset \mathcal{A}_{n+1} \subset \mathcal{A}$ para todo $n \geqslant 1$. Então, $\mathcal{A}_n \uparrow \mathcal{A}$ se, e somente se, dados $A \in \mathcal{A}$ e $\varepsilon > 0$ existe $k \geqslant 1$ tal que $\mu(A \triangle A_k) < \varepsilon$ para algum $A_k \in \mathcal{A}_k$.

Demonstração. Suponha $\mathcal{A}_n \uparrow \mathcal{A}$. Pelas definições, temos $\mathcal{A} = \bigvee_{n=1}^{\infty} \mathcal{A}_n = \sigma\Big(\bigcup_{n\geqslant 1}\mathcal{A}_n\Big)$. Como $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ para todo $n \geqslant 1$, segue que $\bigcup_{n\geqslant 1}\mathcal{A}_n$ é uma álgebra. Logo, segue do Teorema 1.13 que, dados $A \in \mathcal{A}$ e $\varepsilon > 0$ existe $B \in \bigcup_{n\geqslant 1}\mathcal{A}_n$ tal que $\mu(A \triangle B) < \varepsilon$. Portanto, $B \in \mathcal{A}_k$ para algum $k \geqslant 1$ e $\mu(A \triangle B) < \varepsilon$.

Reciprocamente, fixe $A \in \mathcal{A}$ arbitrário. Vamos mostrar que $A \in \bigvee_{n=1}^{\infty} \mathcal{A}_n \pmod{\mu}$. Por hipótese, para cada $j \geqslant 1$ existe $k_j \geqslant 1$ tal que $\mu(A \triangle A_{k_j}) < 2^{-j}$ para algum $A_{k_j} \in \mathcal{A}_{k_j}$. Daí, segue da Proposição 1.16 que $\mu(A \triangle \liminf A_{k_j}) = 0$.

Logo, como $\liminf A_{k_j} \in \bigvee_{n \geqslant 1} \mathcal{A}_n$, segue que $A \in \bigvee_{n=1}^{\infty} \mathcal{A}_n \pmod{\mu}$, como queríamos. $\square$

Proposição 1.28 . Se $\mathcal{A}_n \uparrow \mathcal{A}$ e $\mathcal{B}_n \uparrow \mathcal{B}$, então $\mathcal{A}_n \vee \mathcal{B}_n \uparrow \mathcal{A} \vee \mathcal{B}$.

Demonstração. Suponha que $\mathcal{A}_n \uparrow \mathcal{A}$ e $\mathcal{B}_n \uparrow \mathcal{B}$. Como $\mathcal{A}_n \subset \mathcal{A}_{n+1} \subset \mathcal{A}$ e $\mathcal{B}_n \subset \mathcal{B}_{n+1} \subset \mathcal{B}$ para todo $n \geqslant 1$, temos $\mathcal{A}_n \vee \mathcal{B}_n \subset \mathcal{A}_{n+1} \vee \mathcal{B}_{n+1} \subset \mathcal{A} \vee \mathcal{B}$ para todo $n \geqslant 1$. Além disso, para todo $n \geqslant 1$ temos $\mathcal{A}_n \cup \mathcal{B}_n \subset \mathcal{A} \cup \mathcal{B}$ e, portanto, $\bigvee_{n=1}^{\infty} \mathcal{A}_n \vee \mathcal{B}_n \subset \mathcal{A} \vee \mathcal{B}$. Por outro lado, note que $\mathcal{A} = \bigvee_{n=1}^{\infty} \mathcal{A}_n \subset \bigvee_{n=1}^{\infty} (\mathcal{A}_n \vee \mathcal{B}_n)$ e, similarmente, $\mathcal{B} \subset \bigvee_{n=1}^{\infty} \mathcal{A}_n \vee \mathcal{B}_n$, donde obtemos $\mathcal{A} \vee \mathcal{B} \subset \bigvee_{n=1}^{\infty} \mathcal{A}_n \vee \mathcal{B}_n$. Portanto, $\mathcal{A}_n \vee \mathcal{B}_n \uparrow \mathcal{A} \vee \mathcal{B}$. $\square$

Proposição 1.29 . Se $\mathcal{A}_n \uparrow \mathcal{A}$ e $T : X \to X$ preserva μ, então para cada $j \geqslant 1$ fixado tem-se $T^{-j}\mathcal{A}_n \uparrow T^{-j}\mathcal{A}$. Além disso, se T é um automorfismo, então para cada $j \geqslant 1$ fixado também tem-se $T^{j}\mathcal{A}_n \uparrow T^{j}\mathcal{A}$.

Demonstração. Fixe $j \geqslant 1$ arbitrário. Como $\mathcal{A}_n \subset \mathcal{A}_{n+1} \subset \mathcal{A}$ para todo $n \geqslant 1$, temos $T^{-j}\mathcal{A}_n \subset T^{-j}\mathcal{A}_{n+1} \subset T^{-j}\mathcal{A}$ para todo $n \geqslant 1$. Agora, fixe $T^{-j}A \in T^{-j}\mathcal{A}$ e $\varepsilon > 0$ arbitrários. Como $\mathcal{A}_n \uparrow \mathcal{A}$, pela Proposição 1.27 existe $n_o \geqslant 1$ tal que $\mu(A \triangle A_{n_0}) < \varepsilon$ para algum $A_{n_0} \in \mathcal{A}_{n_0}$. Mas, como T preserva μ, temos $\mu(T^{-j}A \triangle T^{-j}A_{n_0}) = \mu(A \triangle A_{n_0}) < \varepsilon$. Portanto, novamente pela Proposição 1.27 temos $T^{-j}\mathcal{A}_n \uparrow T^{-j}\mathcal{A}$. Por fim, se T é um automorfismo, aplicando o que acabamos de ver para T^{-1} em lugar de T também obtemos $T^{j}\mathcal{A}_n \uparrow T^{j}\mathcal{A}$. $\square$

Proposição 1.30 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade, $T : X \to X$ uma transformação que preserva μ e $(\mathcal{A}_n)_{n \geqslant 1}$ uma sequência de sub-σ-álgebras tais que $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ para cada $n \geqslant 1$. Então, $T^{-1}\Big(\bigvee_{n=1}^{\infty} \mathcal{A}_n\Big) = \bigvee_{n=1}^{\infty} T^{-1}\mathcal{A}_n$. Além disso, se T é um automorfismo, então também tem-se $T\Big(\bigvee_{n=1}^{\infty} \mathcal{A}_n\Big) = \bigvee_{n=1}^{\infty} T\mathcal{A}_n$.

Demonstração. Primeiro, recorde que $\bigvee_{n \geqslant 1} \mathcal{A}_n = \sigma\Big(\bigcup_{n \geqslant 1} \mathcal{A}_n\Big)$, donde tem-se $T^{-1}\Big(\bigvee_{n \geqslant 1} \mathcal{A}_n\Big) = T^{-1}\Big(\sigma\Big(\bigcup_{n \geqslant 1} \mathcal{A}_n\Big)\Big)$. Agora, tome $A := T^{-1}B \in T^{-1}\Big(\bigvee_{n \geqslant 1} \mathcal{A}_n\Big)$ arbitrário. Como $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ para todo $n \geqslant 1$, segue que $\bigcup_{n \geqslant 1} \mathcal{A}_n$ é uma álgebra que gera $\bigvee_{n \geqslant 1} \mathcal{A}_n$. Daí, pelo Teorema 1.13 existe uma sequência crescente de números naturais $n_j \to \infty$ tal que $\mu(B \triangle B_{n_j}) < 2^{-j}$ para algum $B_{n_j} \in \mathcal{A}_{n_j}$ $(j \geqslant 1)$. Como T preserva μ, temos $\mu(A \triangle T^{-1}(B_{n_j})) < 2^{-j}$ para todo $j \geqslant 1$, donde pela Proposição 1.16 segue que $A = \liminf T^{-1}(B_{n_j}) \in \bigvee_{n\geqslant 1} T^{-1}\mathcal{A}_n$. Isso mostra que $T^{-1}\Big(\bigvee_{n=1}^{\infty} \mathcal{A}_n\Big) \subset \bigvee_{n=1}^{\infty} T^{-1}\mathcal{A}_n$.

Reciprocamente, tome $A \in \bigvee_{n\geqslant 1} T^{-1}\mathcal{A}_n = \sigma\Big(\bigcup_{n\geqslant 1} T^{-1}\mathcal{A}_n\Big)$ arbitrário. Como $T^{-1}\mathcal{A}_n \subset T^{-1}\mathcal{A}_{n+1}$ para todo $n \geqslant 1$, segue que $\bigcup_{n\geqslant 1} T^{-1}\mathcal{A}_n$ é uma álgebra que gera $\bigvee_{n\geqslant 1} T^{-1}\mathcal{A}_n$. Daí, novamente pelo Teorema 1.13 existe uma sequência crescente de números naturais $n_j \to \infty$ tal que $\mu(A \triangle A_{n_j}) < 2^{-j}$ para algum $A_{n_j} \in T^{-1}\mathcal{A}_{n_j}$ $(j \geqslant 1)$. Mas, como $T^{-1}\mathcal{A}_{n_j} \subset T^{-1}\Big(\bigvee_{n\geqslant 1} \mathcal{A}_n\Big)$ para cada $j \geqslant 1$, vemos que $A_{n_j} \in T^{-1}\Big(\bigvee_{n\geqslant 1} \mathcal{A}_n\Big)$ para cada $j \geqslant 1$, donde pela Proposição 1.16 obtemos $A = \liminf A_{n_j} \in T^{-1}\Big(\bigvee_{n\geqslant 1} \mathcal{A}_n\Big)$. Isso mostra que $\bigvee_{n\geqslant 1} T^{-1}\mathcal{A}_n \subset T^{-1}\Big(\bigvee_{n\geqslant 1} \mathcal{A}_n\Big)$. Por fim, se T é um automorfismo, então aplicando o que acabamos de ver com T^{-1} em lugar de T também obtemos $T\Big(\bigvee_{n=1}^{\infty} \mathcal{A}_n\Big) = \bigvee_{n=1}^{\infty} T\mathcal{A}_n$. Isso conclui a demonstração. $\square$

Proposição 1.31 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade separável e $\mathcal{A}$ uma sub-σ-álgebra. Então, existe uma sequência $(\mathcal{P}_n)_{n=1}^{\infty}$ de partições finitas de X tal que $\widehat{P_n} \subset \mathcal{A}$ para todo $n \geqslant 1$ e $\widehat{\mathcal{P}_n} \uparrow \mathcal{A}$.

Demonstração. Como $\mathcal{A} \subset \mathcal{X}$, segue que $(\mathcal{A},d)$ é separável, donde existe uma sequência $(A_n)_{n=1}^{\infty} \subset \mathcal{A}$ densa em $(\mathcal{A},d)$.

Defina $\mathcal{P}_1 := \{A_1, X \setminus A_1\}$. É claro que pelo menos um dos conjuntos de $\mathcal{P}_1$ tem medida positiva. Agora, defina $\mathcal{P}_2 := \{A_1 \cap A_2, A_1 \setminus (A_1 \cap A_2), A_2 \setminus (A_1 \cap A_2), X \setminus (A_1 \cup A_2)\}$. Novamente, pelo menos um dos conjuntos de $\mathcal{P}_2$ tem medida positiva. Mais geralmente, para $n \geqslant 3$ vamos definir recursivamente os elementos de $\mathcal{P}_n$ através de n passos, a saber:

(1) Escreva $(j_1^1, \ldots, j_n^1) = (1, \ldots, n)$ e considere o conjunto $A_1 \cap \cdots \cap A_n$.

(2) Para cada escolha $j_1^2, \ldots, j_{n-1}^2 \in \{j_1^1, \ldots, j_n^1\}$ considere o conjunto

$$A^2_{j_1^2,\ldots,j_{n-1}^2} = \bigcap_{i=1}^{n-1} A_{j_i^2} \setminus \bigcap_{i=1}^{n} A^1_{j_i}.$$

(3) Para cada escolha $j_1^3, \ldots, j_{n-2}^3 \in \{j_1^2, \ldots, j_{n-1}^2\}$ considere o conjunto

$$A^3_{j_1^3,\ldots,j_{n-2}^3} = \bigcap_{i=1}^{n-2} A_{j_i^3} \setminus \bigcap_{i=1}^{n-1} A^2_{j_i}$$

$(n-1)$ Para cada escolha $j_1^{n-1}, j_2^{n-1} \in \{j_1^{n-2}, j_2^{n-2}, j_3^{n-2}\}$ considere o conjunto

$$A^{n-1}_{j_1^{n-1},j_2^{n-1}} = \bigcap_{i=1}^{2} A^{n-1}_{j_i} \setminus \bigcap_{i=1}^{3} A^{n-2}_{j_i}$$

(n) Para cada $j \in \{1, \ldots, n\}$ considere o conjunto

$$A_j \setminus \bigcup_{\substack{i=1 \\ i\neq j}}^{n} A_i.$$

A partição $\mathcal{P}_n$ é a partição de X cujos átomos são precisamente os conjuntos de medida positiva que pertencem à coleção formada pelo conjunto $X \setminus \bigcup_{i=1}^{n} A_i$ e os conjuntos dados nesses n passos.

Note que $A_1, \ldots, A_n \in \widehat{\mathcal{P}_n}$ e $\mathcal{P}_n \prec \mathcal{P}_{n+1}$ para cada $n \geqslant 1$. Portanto, segue que $\widehat{\mathcal{P}_n} \subset \widehat{\mathcal{P}_{n+1}} \subset \mathcal{A}$ para todo $n \geqslant 1$, donde $\bigvee_{n\geqslant 1} \widehat{\mathcal{P}_n} \subset \mathcal{A}$.

Por fim, seja $A \in \mathcal{A}$ arbitrário. Para cada $l \geqslant 1$ existe $A_{n_l} \in (A_n)_{n=1}^{\infty}$ tal que $\mu(A \triangle A_{n_l}) < 2^{-l}$ donde, pelo Teorema 1.16, temos

$$A = \liminf A_{n_l} = \bigcup_{i=1}^{\infty} \bigcap_{j=i}^{\infty} A_{n_j} \in \bigvee_{n=1}^{\infty} \widehat{\mathcal{P}_n} \pmod{\mu}.$$

Portanto, $\widehat{\mathcal{P}_n} \uparrow \mathcal{A}$.

□

Proposição 1.32 . Sejam $(X, \mathcal{X}, \mu)$ um espaço de probabilidade separável e $\mathcal{A}$ uma álgebra tais que $\sigma(\mathcal{A}) = \mathcal{X}$. Então, existe uma sequência $(\mathcal{P}_n)_{n=1}^{\infty}$ de partições finitas de X tal que $\widehat{\mathcal{P}_n} \subset \mathcal{A}$ para todo $n \geqslant 1$ e $\widehat{\mathcal{P}_n} \uparrow \mathcal{X}$.

Demonstração. Seja $(B_n)_{n=1}^{\infty} \subset \mathcal{X}$ uma sequência densa. Como $\sigma(\mathcal{A}) = \mathcal{X}$, segue do Teorema 1.13 que, para todos $m, n \geqslant 1$ existe $A_n^m \in \mathcal{A}$ tal que

$$\mu(B_n \triangle A_n^m) < (1/m)2^{-n}.$$

Sejam $\varepsilon > 0$ e $A \in \mathcal{X}$ arbitrários. Existe $n \geqslant 1$ tal que $\mu(A \triangle B_n) < \varepsilon/2$. Existe também $m \geqslant 1$ tal que $\mu(B_n \triangle A_n^m) < \varepsilon/2$ e, portanto, $\mu(A \triangle A_n^m) < \varepsilon$. Vemos assim que a coleção $\{A_n^m : n, m \geqslant 1\} \subset \mathcal{A}$ é enumerável e densa em $\mathcal{X}$, donde podemos aplicar o processo recursivo da demonstração da Proposição 1.31 para obter uma sequência $(\mathcal{P}_n)_{n=1}^{\infty}$ de partições finitas de X tal que $\widehat{\mathcal{P}_n} \subset \mathcal{A}$ para todo $n \geqslant 1$ e $\widehat{\mathcal{P}_n} \uparrow \mathcal{X}$. Isso conclui a demonstração.

□

O próximo resultado nos dá uma condição suficiente para que uma sequência crescente de partições $(\mathcal{P}_n)_{n\geqslant 1}$ satisfaça $\bigvee_{n=1}^{\infty} \mathcal{P}_n = \mathcal{X}$, no caso em que μ é uma medida de Borel probabilística sobre um espaço métrico X.

Lema 1.33 . Suponha $(X, \mathcal{X}, \mu)$ um espaço de probabilidade, onde X é metrizável, $\mathcal{X}$ é a σ-álgebra de Borel, μ é uma medida de Borel probabilística e $\mathcal{P}_1 \prec \mathcal{P}_2 \prec \mathcal{P}_3 \prec \ldots$ é uma sequência de partições. Para cada $x \in X$ denote por $\mathcal{P}_n(x)$ o elemento da partição $\mathcal{P}_n$ que contém x. Se para μ-q.t.p. $x \in X$ tem-se $\operatorname{diam} \mathcal{P}_n(x) \to 0$, então $\bigvee_{n=1}^{\infty} \mathcal{P}_n = \mathcal{X}$.

Demonstração. Sejam U um aberto não-vazio e $x \in U$ arbitrários. A hipótese garante que existem $n_x \geqslant 1$ e $P_{n_x} \in \mathcal{P}_{n_x}$ tais que $x \in P_{n_x} \subset U$. É imediato que $P_{n_x} \in \mathcal{A}\Big(\bigcup_{n=1}^{\infty} \mathcal{P}_n\Big) =: \mathcal{A}$. Agora, como os elementos de $\mathcal{A}$ são uniões finitas de elementos das partições $\mathcal{P}_n$ juntamente com o complementar dessas uniões, segue que $\mathcal{A}$ é enumerável. Em particular, a aplicação $\phi : X \to \mathcal{A}$ definida por $\phi(x) = P_{n_x}$ tem imagem enumerável. Portanto, $U = \bigcup_{x \in U} P_{n_x}$ é uma união enumerável, donde $U \in \sigma(\mathcal{A})$. Isso mostra que $\sigma(\mathcal{A})$ contém os abertos. Dessa forma, $\mathcal{X} \subset \sigma(\mathcal{A})$. Por fim, como cada elemento de $\mathcal{A}$ está contido em $\mathcal{X}$, temos $\mathcal{X} = \sigma(\mathcal{A}) = \bigvee_{n=1}^{\infty} \widehat{\mathcal{P}_n}$, o que conclui a demonstração. □

O próximo resultado exemplifica uma ideia importante que será bastante utilizada em nosso trabalho e que justifica o fato de trabalharmos em espaços de probabilidade separáveis, a saber: um resultado válido no contexto de partições finitas pode valer também no contexto de sub-σ-álgebras gerais.

Proposição 1.34 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade separável, $\mathcal{A}_1$ e $\mathcal{A}_2$ duas sub-σ-álgebras e $T : X \to X$ uma transformação que preserva μ. Então, $T^{-1}(\mathcal{A}_1 \vee \mathcal{A}_2) = T^{-1}\mathcal{A}_1 \vee T^{-1}\mathcal{A}_2$. Além disso, se T é um automorfismo, então $T(\mathcal{A}_1 \vee \mathcal{A}_2) = T\mathcal{A}_1 \vee T\mathcal{A}_2$.

Demonstração. Como o espaço $(X,\mathcal{X},\mu)$ é separável, segue da Proposição 1.31 a existência de duas sequências de partições finitas $(\mathcal{P}_n)_{n=1}^{\infty}$ e $(\mathcal{Q}_n)_{n=1}^{\infty}$ tais que $\widehat{\mathcal{P}_n} \uparrow \mathcal{A}_1$ e $\widehat{\mathcal{Q}_n} \uparrow \mathcal{A}_2$.

Pelas Proposição 1.28 e Proposição 1.29 temos

$$
\begin{aligned}
&\widehat{\mathcal{P}_n} \vee \widehat{\mathcal{Q}_n} \uparrow \mathcal{A}_1 \vee \mathcal{A}_2 \\
&T^{-1}\widehat{\mathcal{P}_n} \uparrow T^{-1}\mathcal{A}_1 \\
&T^{-1}\widehat{\mathcal{Q}_n} \uparrow T^{-1}\mathcal{A}_2 \\
&T^{-1}\widehat{\mathcal{P}_n} \vee T^{-1}\widehat{\mathcal{Q}_n} \uparrow T^{-1}\mathcal{A}_1 \vee T^{-1}\mathcal{A}_2 \\
&T^{-1}(\widehat{\mathcal{P}_n} \vee \widehat{\mathcal{Q}_n}) \uparrow T^{-1}(\mathcal{A}_1 \vee \mathcal{A}_2).
\end{aligned}
$$

Além disso, pela Proposição 1.24 tem-se $T^{-1}(\widehat{\mathcal{P}_n} \vee \widehat{\mathcal{Q}_n}) = T^{-1}\widehat{\mathcal{P}_n} \vee T^{-1}\widehat{\mathcal{Q}_n}$ para todo $n \geqslant 1$. Portanto, conclui-se:

$$
T^{-1}(\mathcal{A}_1 \vee \mathcal{A}_2) = \bigvee_{n=1}^{\infty} T^{-1}(\widehat{\mathcal{P}_n} \vee \widehat{\mathcal{Q}_n}) = \bigvee_{n=1}^{\infty} (T^{-1}\widehat{\mathcal{P}_n} \vee T^{-1}\widehat{\mathcal{Q}_n}) = T^{-1}\mathcal{A}_1 \vee T^{-1}\mathcal{A}_2.
$$

Por fim, se T é um automorfismo, aplicando o que acabamos de ver para T^{-1} em lugar de T também obtemos $T(\mathcal{A}_1 \vee \mathcal{A}_2) = T\mathcal{A}_1 \vee T\mathcal{A}_2$.

□

Finalizamos esta seção com a importante noção de *independência*, tanto no contexto de partições como também no contexto de sub-σ-álgebras gerais.

Definição 1.35 . Sejam $(X\mathcal{X},\mu)$ um espaço de probabilidade e $\mathcal{P}$ e $\mathcal{Q}$ partições. Dizemos que $\mathcal{P}$ e $\mathcal{Q}$ são *independentes* se, $\mu(P \cap Q) = \mu(P)\mu(Q)$ para todo $P \in \mathcal{P}$ e todo $Q \in \mathcal{Q}$. Similarmente, dadas duas sub-σ-álgebras $\mathcal{A}_1$ e $\mathcal{A}_2$, dizemos que $\mathcal{A}_1$ e $\mathcal{A}_2$ são *independentes* se $\mu(A_1 \cap A_2) = \mu(A_1)\mu(A_2)$ para todo $A_1 \in \mathcal{A}_1$ e todo $A_2 \in \mathcal{A}_2$.

Observação 1.36 . Dadas $\mathcal{P}$ e $\mathcal{Q}$ partições, temos que $\mathcal{P}$ e $\mathcal{Q}$ são independentes se, e somente se, $\widehat{\mathcal{P}}$ e $\widehat{\mathcal{Q}}$ são independentes.

1.3 INTEGRAÇÃO

Nessa seção apresentamos os resultados da Teoria de Integração que serão necessários à nossa exposição.

Teorema 1.37 . Seja $f : X \to [0,\infty]$ mensurável. Existe uma sequência de funções simples não-negativas $(s_n)_{n\geqslant 1}$ tal que $0 \leqslant s_n \leqslant s_{n+1} \leqslant f$ para todo $n \geqslant 1$ e $s_n \to f$ pontualmente.

Demonstração. Veja o 1.17 Theorem na página 15 de (RUDIN, 1987). □

Teorema 1.38 . (Teorema da Convergência Monótona) Suponha que $(f_n)_{n\geqslant 1}$ é uma sequência de funções mensuráveis tal que $0 \leqslant f_n(x) \leqslant f_{n+1}(x)$ para todo $n \geqslant 1$, para todo $x \in X$ e $f_n(x) \to f(x)$ para cada $x \in X$. Então, f é mensurável e $\displaystyle\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu$.

Demonstração. Veja o 1.26 Theorem na página 21 de (RUDIN, 1987). □

Teorema 1.39 . (Teorema da Convergência Dominada) Suponha que $(f_n)_{n\geqslant 1}$ é uma sequência de funções mensuráveis tal que $\displaystyle\lim_{n\to\infty} f_n(x) = f(x)$ existe para cada $x \in X$ e para a qual existe uma função $g \in L^1(X,\mathcal{X},\mu)$ satisfazendo $|f_n(x)| \leqslant g(x)$ para todo $n \geqslant 1$, para todo $x \in X$. Então, $f \in L^1(X,\mathcal{X},\mu)$, $\displaystyle\lim_{n\to\infty} \int_X |f_n - f| d\mu = 0$ e $\displaystyle\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu$.

Demonstração. Veja o 1.34 Theorem na página 26 de (RUDIN, 1987). □

Definição 1.40 . Sejam $(X,\mathcal{X})$ um espaço mensurável e μ e ν duas medidas sobre $(X,\mathcal{X})$. Dizemos que ν é *absolutamente contínua com respeito a* μ se $\nu(A) = 0$ sempre que $\mu(A) = 0$.

O próximo resultado é fundamental em muitas aplicações da Teoria de Integração, particularmente na Teoria das Probabilidades.

Teorema 1.41 . (Radon-Nikodym) Sejam $(X,\mathcal{X})$ espaço mensurável, μ e ν duas medidas σ-finitas sobre $(X,\mathcal{X})$ tais que ν é absolutamente contínua com respeito a μ. Então, existe $f \in L^1(X,\mathcal{X},\mu)$ tal que $\displaystyle\nu(A) = \int_A f d\mu$ para todo $A \in \mathcal{X}$ e f é essencialmente única no seguinte sentido: se $g \in L^1(X,\mathcal{X},\mu)$ satisfaz $\displaystyle\nu(A) = \int_A g d\mu$ para todo $A \in \mathcal{X}$, então $f(x) = g(x)$ para μ-q.t.p. $x \in X$.

Demonstração. Veja Teorema B na página 128 de (HALMOS, Paul, 1956). □

Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade e $\mathcal{A} \subset \mathcal{X}$ uma sub-σ-álgebra. Dada uma função $f \in L^1(X,\mathcal{X},\mu)$ definimos uma medida ν sobre o espaço de medida $(X,\mathcal{A})$ por

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

Note que a restrição $\mu_{|\mathcal{A}}$ e ν são medidas sobre $(X,\mathcal{A})$ e, além disso, ν é absolutamente contínua com respeito a μ. Daí, pelo Teorema 1.41 existe essencialmente uma única função $g \in L^1(X,\mathcal{A},\mu_{|\mathcal{A}})$ tal que

$$\nu(A) = \int_A g d\mu \text{ para todo } A \in \mathcal{A}.$$

Definição 1.42 . A função g dada acima é dita *função valor esperado de f dada $\mathcal{A}$* e é denotada por $\mathbb{E}(f|\mathcal{A})$. Além disso, definimos a *probabilidade condicional de $A \in \mathcal{X}$ dada $\mathcal{A}$* por $\mu(A|\mathcal{A}) := \mathbb{E}(1_A|\mathcal{A})$.

Observação 1.43 . Sejam $\mathcal{A} \subset \mathcal{X}$ uma sub-σ-álgebra e $f \in L^1(X,\mathcal{X},\mu)$. Então, decorre do Teorema 1.41 que $\mathbb{E}(f|\mathcal{A})$ fica completamente caracterizado pelas seguintes propriedades:

(i) $\mathbb{E}(f|\mathcal{A})$ é $\mathcal{A}$-mensurável;

(ii) $\int_A \mathbb{E}(f|\mathcal{A})d\mu = \int_A f d\mu$ para todo $A \in \mathcal{A}$.

Mais especificamente: se $h \in L^1(X,\mathcal{X},\mu)$ satisfaz (i) e (ii), então $h = \mathbb{E}(f|\mathcal{A})$ μ-q.t.p. Neste caso, uma tal h é dita uma *versão* de $\mathbb{E}(f|\mathcal{A})$.

Proposição 1.44 . Sejam $f \in L^1(X,\mathcal{X},\mu)$ e $\mathcal{A} \subset \mathcal{X}$ uma sub-σ-álgebra. Valem as seguintes propriedades:

(i) Se $f \geqslant 0$ μ-q.t.p. então $E(f|\mathcal{A}) \geqslant 0$ μ-q.t.p.;

(ii) Se f é $\mathcal{A}$-mensurável, então $E(f|\mathcal{A}) = f$ μ-q.t.p.;

(iii) $\int_A \mathbb{E}(\mathbb{E}(f|\mathcal{A})|\mathcal{A})d\mu = \int_A \mathbb{E}(f|\mathcal{A})d\mu = \int_A f d\mu$ para todo $A \in \mathcal{A}$;

(iv) $E(\cdot|\mathcal{A}) : L^1(X,\mathcal{X},\mu) \to L^1(X,\mathcal{X},\mu)$ é um operador linear limitado;

(v) Se $g : X \to \mathbb{R}$ é uma função $\mathcal{A}$-mensurável tal que $gf \in L^1(X,\mathcal{X},\mu)$, então $\mathbb{E}(gf|\mathcal{A}) = g\mathbb{E}(f|\mathcal{A})$ μ-q.t.p.

Demonstração. (i): Suponha que $f \geqslant 0$ para μ-q.t.p. e considere $A = \{x \in X : \mathbb{E}(f|\mathcal{A}) < 0\}$. Como $\mathbb{E}(f|\mathcal{A})$ é $\mathcal{A}$-mensurável, segue que $A \in \mathcal{A}$. Agora, como

$$0 \geqslant \int_A \mathbb{E}(f|\mathcal{A})d\mu = \int_A f d\mu,$$

temos $\mu(A) = 0$, o que implica que $\mathbb{E}(f|\mathcal{A}) \geqslant 0$ μ-q.t.p.

(ii): Por definição, para todo $A \in \mathcal{A}$ tem-se:

$$\int_A f d\mu = \int_A \mathbb{E}(f|\mathcal{A})d\mu.$$

Daí, pela Observação 1.43 segue que $\mathbb{E}(f|\mathcal{A}) = f$ μ-q.t.p.

(iii): Decorre imediatamente do item (ii) juntamente com a Observação 1.43.

(iv): Sejam $f,g \in L^1(X,\mathcal{X},\mu)$ e $\lambda \in \mathbb{R}$ arbitrários. Dado $A \in \mathcal{A}$, temos:

$$\begin{aligned}
\int_A [\mathbb{E}(f|\mathcal{A}) + \lambda\mathbb{E}(g|\mathcal{A})]d\mu &= \int_A \mathbb{E}(f|\mathcal{A})d\mu + \lambda \int_A \mathbb{E}(g|\mathcal{A})d\mu \\
&= \int_A f d\mu + \lambda \int_A g d\mu \\
&= \int_A (f + \lambda g) d\mu.
\end{aligned}$$

Logo, pela Observação 1.43 tem-se $\mathbb{E}(f + \lambda g|\mathcal{A}) = \mathbb{E}(f|\mathcal{A}) + \lambda\mathbb{E}(g|\mathcal{A})$.

Por fim, observe que $||E(f|\mathcal{A})|| = \displaystyle\int_X E(f|\mathcal{A})d\mu = \int_X f d\mu = ||f||$.

Portanto, $E(\cdot|\mathcal{A})$ é um operador linear limitado.

(v): Suponha inicialmente que $g = 1_E$ para algum $E \in \mathcal{A}$, ou seja, $g(x) = 1$ se $x \in E$ e $g(x) = 0$ caso contrário, para algum $E \in \mathcal{A}$. Então, g é $\mathcal{A}$-mensurável tal que $gf \in L^1(X,\mathcal{X},\mu)$ e, para cada $A \in \mathcal{A}$ tem-se:

$$\begin{aligned}
\int_A g\mathbb{E}(f|\mathcal{A})d\mu &= \int_A 1_E\mathbb{E}(f|\mathcal{A})d\mu \\
&= \int_{E\cap A} \mathbb{E}(f|\mathcal{A})d\mu \\
&= \int_{E\cap A} f d\mu \\
&= \int_A 1_E f d\mu \\
&= \int_A g f d\mu.
\end{aligned}$$

Daí, segue da Observação 1.43 que $\mathbb{E}(gf|\mathcal{A}) = g\mathbb{E}(f|\mathcal{A})$ μ-q.t.p.

Agora, suponha que $g = \displaystyle\sum_{i=1}^{k} c_i 1_{E_i}$ é uma função mensurável simples não-negativa $\mathcal{A}$-mensurável, ou seja, $c_i \geqslant 0$ e $E_i \in \mathcal{A}$ para todo $1 \leqslant i \leqslant k$. Então, $gf \in L^1(X,\mathcal{X},\mu)$ e, para cada $A \in \mathcal{A}$, pela parte já provada obtém-se:

$$\begin{aligned}
\int_A g\mathbb{E}(f|\mathcal{A})d\mu &= \int_A \Big(\sum_{i=1}^{k} c_i 1_{E_i}\Big)\mathbb{E}(f|\mathcal{A})d\mu \\
&= \sum_{i=1}^{k} c_i \int_A 1_{E_i}\mathbb{E}(f|\mathcal{A})d\mu \\
&= \sum_{i=1}^{k} c_i \int_A 1_{E_i} f d\mu \\
&= \int_A \Big(\sum_{i=1}^{k} c_i 1_{E_i}\Big) f d\mu
\end{aligned}$$

$$= \int_A gf d\mu.$$

Agora, suponha que g é uma função real $\mathcal{A}$-mensurável não-negativa tal que $gf \in L^1(X,\mathcal{X},\mu)$ e suponha também que $f \geqslant 0$. Pelo Teorema 1.37 existe uma sequência crescente $(s_n)_{n\geqslant 1}$ de funções $\mathcal{A}$-mensuráveis simples não-negativas tal que $s_n \uparrow g$ pontualmente. Pela parte já provada, para cada $A \in \mathcal{A}$ temos

$$\int_A s_n \mathbb{E}(f|\mathcal{A}) d\mu = \int_A s_n f d\mu \ \text{ para todo } n \geqslant 1.$$

Como $f \geqslant 0$, segue do item (i) que $\mathbb{E}(f|\mathcal{A}) \geqslant 0$ μ-q.t.p. Daí, segue que $s_n\mathbb{E}(f|\mathcal{A}) \uparrow g\mathbb{E}(f|\mathcal{A})$ pontualmente, donde tomando o limite quando $n \to \infty$ em ambos os lados, o Teorema 1.38 (Teorema da Convergência Monótona) nos dá:

$$\int_A g\mathbb{E}(f|\mathcal{A}) d\mu = \int_A gf d\mu.$$

Como $A \in \mathcal{A}$ é arbitrário, pela Observação 1.43 segue que $\mathbb{E}(gf|\mathcal{A}) = g\mathbb{E}(f|\mathcal{A})$ μ-q.t.p., no caso em que $g,f \geqslant 0$.

Para o caso geral, decompomos $g = g^+ - g^-$ e $f = f^+ - f^-$ através de suas partes positiva e negativa (recorde que $h^+ := \max\{h,0\}$ e $h^- := \max\{-h,0\}$). Daí, como g^+,g^-,f^+,f^- são não-negativas, pela parte já provada juntamente com a linearidade do valor esperado condicional dada no item (iv), para cada $A \in \mathcal{A}$ temos:

$$\begin{aligned}
\int_A g\mathbb{E}(f|\mathcal{A}) d\mu &= \int_A (g^+ - g^-)\mathbb{E}(f^+ - f^-|\mathcal{A}) d\mu \\
&= \int_A g^+\mathbb{E}(f^+|\mathcal{A}) d\mu - \int_A g^+\mathbb{E}(f^-|\mathcal{A}) d\mu \\
&- \int_A g^-\mathbb{E}(f^+|\mathcal{A}) d\mu + \int_A g^-\mathbb{E}(f^-|\mathcal{A}) d\mu \\
&= \int_A g^+ f^+ d\mu - \int_A g^+ f^- d\mu \\
&- \int_A g^- f^+ d\mu + \int_A g^- f^- d\mu \\
&= \int_A (g^+ - g^-) f^+ d\mu - \int_A (g^+ - g^-) f^- d\mu \\
&= \int_A g f^+ d\mu - \int_A g f^- d\mu \\
&= \int_A g(f^+ - f^-) d\mu \\
&= \int_A gf d\mu.
\end{aligned}$$

Como $A \in \mathcal{A}$ é arbitrário, pela Observação 1.43 segue que $\mathbb{E}(gf|\mathcal{A}) = g\mathbb{E}(f|\mathcal{A})$ μ-q.t.p.

Isso conclui a demonstração.

$\square$

Observação 1.45 . Segue de (iii) e (iv) da Proposição 1.44 que o operador $\mathbb{E}(\cdot|\mathcal{A})$ é uma projeção limitada de $L^1(X,\mathcal{X},\mu)$.

Proposição 1.46 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade, $\mathcal{G} \subset \mathcal{X}$ uma sub-σ-álgebra, $\mathcal{A}$ uma álgebra que gera $\mathcal{G}$, $g \in L^1(X,\mathcal{X},\mu)$ não-negativa e $C \in \mathcal{X}$.

Se g é $\mathcal{G}$-mensurável e $\displaystyle\int_A g d\mu = \mu(C \cap A)$ para todo $A \in \mathcal{A}$, então g é uma versão de $\mu(C|\mathcal{G})$.

Demonstração. Primeiro, observe que as funções

$$A \in \mathcal{G} \longmapsto \int_A g d\mu$$

e

$$A \in \mathcal{G} \longmapsto \mu(C \cap A)$$

são medidas sobre $\mathcal{G}$. Por hipótese, elas coincidem sobre a álgebra $\mathcal{A}$, donde elas têm de coincidir sobre $\mathcal{G} = \sigma(\mathcal{A})$. Ou seja, para cada $G \in \mathcal{G}$ tem-se:

$$\begin{aligned}\int_G g d\mu &= \mu(C \cap G)\\ &= \int_G 1_C d\mu,\end{aligned}$$

donde o resultado segue da Observação 1.43. □

Teorema 1.47 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade, $\mathcal{F} \subset \mathcal{X}$ uma sub-σ-álgebra e $\mathcal{P}$ uma partição. Para cada $C \in \mathcal{X}$, tem-se:

$$\mu(C|\widehat{\mathcal{P}} \vee \mathcal{F}) = \sum_{P \in \mathcal{P}} 1_P \frac{\mu(C \cap P|\mathcal{F})}{\mu(P|\mathcal{F})} \quad \mu\text{-}q.t.p.$$

Demonstração. Seja $\mathcal{A}$ a coleção de todas as uniões disjuntas finitas de elementos da forma $P \cap F$ com $P \in \mathcal{P}$ e $F \in \mathcal{F}$. Não é difícil verificar que $\mathcal{A}$ é uma álgebra que gera $\widehat{\mathcal{P}} \vee \mathcal{F}$. Consequentemente, pela Proposição 1.46 é suficiente mostrar que

$$\int_{P \cap F} \sum_{P' \in \mathcal{P}} 1_{P'} \frac{\mu(C \cap P'|\mathcal{F})}{\mu(P'|\mathcal{F})} d\mu = \mu(C \cap P \cap F),$$

para cada $P \in \mathcal{P}$ e cada $F \in \mathcal{F}$. Fixemos $P \in \mathcal{P}$ e $F \in \mathcal{F}$ arbitrários. Pela Observação 1.19 podemos supor que os $P' \in \mathcal{P}$ são dois a dois disjuntos. Portanto,

$$\begin{aligned}\int_{P \cap F} \sum_{P' \in \mathcal{P}} 1_{P'} \frac{\mu(C \cap P'|\mathcal{F})}{\mu(P'|\mathcal{F})} d\mu &= \int_F 1_P \frac{\mathbb{E}(1_{C \cap P}|\mathcal{F})}{\mathbb{E}(1_P|\mathcal{F})} d\mu\\ &= \int_F \mathbb{E}\Big(1_P \frac{\mathbb{E}(1_{C \cap P}|\mathcal{F})}{\mathbb{E}(1_P|\mathcal{F})}|\mathcal{F}\Big) d\mu\\ &= \int_F \frac{\mathbb{E}(1_{C \cap P}|\mathcal{F})}{\mathbb{E}(1_P|\mathcal{F})} \mathbb{E}(1_P|\mathcal{F}) d\mu\end{aligned}$$

$$
\begin{aligned}
&= \int_F \mathbb{E}(1_{C\cap P}|\mathcal{F})d\mu \\
&= \int_X 1_F\mathbb{E}(1_{C\cap P}|\mathcal{F})d\mu \\
&= \int_X \mathbb{E}(1_F 1_{C\cap P}|\mathcal{F})d\mu \\
&= \int_X \mathbb{E}(1_{C\cap P\cap F}|\mathcal{F})d\mu \\
&= \int_X 1_{C\cap P\cap F}d\mu \\
&= \mu(C\cap P\cap F),
\end{aligned}
$$

como queríamos demonstrar.

□

Teorema 1.48 . (Desigualdade de Jensen) Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade, $\phi : [0,1] \to \mathbb{R}$ uma função côncava, ou seja, tal que $s\phi(a) + t\phi(b) \leqslant \phi(sa + tb)$ para todos $a,b \in [0,1]$, $0 \leqslant s,t \leqslant 1$ com $s + t = 1$, e $f \in L^1(X,\mathcal{X},\mu)$ tal que $0 \leqslant f(x) \leqslant 1$ para todo $x \in X$. Então, para qualquer sub-σ-álgebra $\mathcal{G} \subset \mathcal{X}$ tem-se $\mathbb{E}(\phi \circ f|\mathcal{G}) \leqslant \phi(\mathbb{E}(f|\mathcal{G}))$ μ-q.t.p.

Demonstração. Veja o 4.Theorem na página 35 de (PARRY, 1981). □

Teorema 1.49 . (Teorema de Convergência de Martingale) Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade, $(\mathcal{A}_n)_{n=1}^{\infty}$ uma sequência de sub-σ-álgebras e $f \in L^1\Big(X, \bigvee_{n=1}^{\infty} \mathcal{A}_n,\mu\Big)$. Valem:

(i) Se $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ para todo $n \geqslant 1$, então

$$\mathbb{E}(f|\mathcal{A}_n) \uparrow \mathbb{E}\Big(f|\bigvee_{n=1}^{\infty} \mathcal{A}_n\Big) \ \mu\text{-q.t.p.}$$

(ii) Se $\mathcal{A}_n \supset \mathcal{A}_{n+1}$ para todo $n \geqslant 1$, então

$$\mathbb{E}(f|\mathcal{A}_n) \downarrow \mathbb{E}\Big(f|\bigcap_{n\geqslant 1} \mathcal{A}_n\Big) \ \mu\text{-q.t.p.}$$

Demonstração. Veja o Theorem 4.3 na página 331 de (DOOB, 1990). □

Terminamos este capítulo de preliminares com dois fatos simples de análise na reta que serão fundamentais em nossa exposição. O primeiro deles será usado de forma extensiva no último capítulo e o segundo será utilizado no capítulo 3 para garantir a existência da entropia de um sistema dinâmico.

Proposição 1.50 . Sejam $(s_n)_{n=1}^{\infty}$ uma sequência em $\mathbb{R}$ e $s \in \mathbb{R}$ tais que $s_n \to s$.

Então,

$$\frac{1}{n}\sum_{j=1}^{n-1} s_j \to s.$$

Demonstração. Uma vez que a sequência $(s_n)_{n=1}^{\infty}$ converge para $s \in \mathbb{R}$, dado $\varepsilon > 0$ existe $n_0 \geqslant 1$ tal que $|s_n - s| < \varepsilon$ sempre que $n \geqslant n_0$. Tome $M > 0$ tal que $|s_n - s| < M$ para todo $n \geqslant 1$. Note que, fixado $n_1 \geqslant \max\left\{n_0, \dfrac{|s| + M(n_0-1)}{\varepsilon}\right\}$ arbitrariamente, se $n \geqslant n_1$, então

$$\begin{aligned}
\Big|\frac{1}{n}\sum_{j=1}^{n-1} s_j - s\Big| &= \frac{1}{n}\Big|\sum_{j=1}^{n-1} s_j - ns\Big| \\
&= \frac{1}{n}\Big|\sum_{j=1}^{n-1} s_j - \sum_{k=1}^{n} s\Big| \\
&\leqslant \frac{1}{n}\Big(\Big(\sum_{j=1}^{n-1} |s_j - s|\Big) + |s|\Big) \\
&= \frac{1}{n}(|s| + \sum_{j=1}^{n_0-1} |s_j - s|) + \frac{1}{n}\sum_{i=n_0}^{n-1} |s_i - s| \\
&< \frac{1}{n}(|s| + M(n_0-1)) + \frac{1}{n}\varepsilon(n - n_0) \\
&< 2\varepsilon.
\end{aligned} \tag{3}$$

□

Definição 1.51 . Uma sequência $(a_n)_{n\geqslant 1} \subset \mathbb{R}$ é dita *subaditiva* se $a_{m+n} \leqslant a_m + a_n$ para todos $m,n \geqslant 1$.

Lema 1.52 . (Lema de Fekete) Se $(a_n)_{n\geqslant 1} \subset \mathbb{R}$ é uma sequência subaditiva, então

$$\lim_n \frac{a_n}{n} = \inf_n \frac{a_n}{n},$$

onde a existência do limite é parte da conclusão.

Demonstração. Escreva $\gamma := \inf_{n\geqslant 1}(a_n/n)$ e suponha inicialmente que $\gamma > -\infty$. Dado $\varepsilon > 0$ existe $k \geqslant 1$ tal que $a_k \leqslant (\gamma+\varepsilon)k$. Pela subaditividade da sequência $(a_n)_{n\geqslant 1}$, para cada $n \geqslant 1$ temos $a_{nk} \leqslant na_k \leqslant nk(\gamma+\varepsilon)$, ou seja, $(a_{nk}/nk) \leqslant \gamma+\varepsilon$. Daí, vemos que $\liminf_n (a_n/n) \leqslant \gamma+\varepsilon$. Portanto, como $\varepsilon > 0$ é arbitrário, concluímos que $\liminf_n (a_n/n) \leqslant \gamma = \inf_{n\geqslant 1}(a_n/n)$, donde $\liminf_n (a_n/n) = \gamma$. Agora, dado $m \geqslant 1$ arbitrário podemos escrever $m = nk+j$ com $0 \leqslant j < k$, donde novamente pela subaditividade obtemos:

$$\begin{aligned}
a_m &= a_{nk+j} \\
&\leqslant a_{nk} + a_j \\
&\leqslant (\gamma+\varepsilon)nk + \max_{0\leqslant j<k} a_j
\end{aligned}$$

Dividindo em ambos os lados por $m \geqslant 1$ e fazendo $m \to \infty$ obtém-se $\limsup_m (a_m/m) \leqslant \gamma+\varepsilon$. Como $\varepsilon > 0$ é arbitrário, concluímos que $\limsup_m (a_m/m) \leqslant \gamma = \liminf_m (a_m/m)$. Portanto, $\lim_n (a_n/n) = \inf_{n\geqslant 1}(a_n/n)$ no caso em que $\inf_{n\geqslant 1}(a_n/n) > -\infty$.

Por fim, suponha que $\inf_{n \geqslant 1}(a_n/n) = -\infty$. Dado $M > 0$ arbitrário, existe $k \geqslant 1$ tal que $a_k < -Mk$. Pela subaditividade da sequência $(a_n)_{n \geqslant 1}$, para cada $n \geqslant 1$ temos $a_{nk} \leqslant na_k < nk(-M)$, ou seja, $(a_{nk}/nk) < -M$. Daí, vemos que $\liminf_n (a_n/n) \leqslant -M$. Como $M > 0$ é arbitrário, obtemos $\liminf_n (a_n/n) = -\infty$. O resto da demonstração segue exatamente os mesmos passos do caso anterior com $-M$ em lugar de $\gamma + \varepsilon$. $\square$

2 O DESLOCAMENTO DE BERNOULLI

Considere um experimento aleatório com um número finito de resultados igualmente prováveis, digamos $a \geqslant 2$. Por exemplo, girar uma roleta "honesta" com $a \geqslant 2$ saídas possíveis. Denotamos por $\{1,2,\ldots,a\}$ o conjunto de todos os resultados possíveis e por $X = \{1,2,\ldots,a\}^{\mathbb{Z}}$ o conjunto de todas as sequências bilaterais infinitas $x = \ldots x_{-2}x_{-1}x_0x_1x_2\ldots$ de resultados possíveis. É comum denotarmos por $x(i)$ a i-ésima coordenada de x ($i \in \mathbb{Z}$).

Agora, considere a transformação $T : \{1,\ldots,a\}^{\mathbb{Z}} \to \{1,\ldots,a\}^{\mathbb{Z}}$ dada por

$$T((x(i))_{i\in\mathbb{Z}}) = (x(i+1))_{i\in\mathbb{Z}}.$$

Ou seja, a transformação T *desloca à esquerda* cada coordenada de todo elemento $(x(i))_{i\in\mathbb{Z}} \in X$. Claramente, T é invertível e sua inversa é precisamente a transformação de *deslocamento à direita* $T^{-1}((x(i))_{i\in\mathbb{Z}}) = (x(i-1))_{i\in\mathbb{Z}}$.

Vamos mostrar que existem uma σ-álgebra $\mathcal{X}$ sobre X e uma medida de probabilidade μ sobre $\mathcal{X}$ que é invariante pela transformação T. Em outras palavras, $(X,\mathcal{X},\mu,T)$ é um sistema dinâmico, dito *deslocamento de Bernoulli*. O deslocamento de Bernoulli fornece um modelo matemático para o experimento aleatório que consiste em girar uma roleta com um número finito de saídas igualmente prováveis, "desde o passado remoto até o eterno futuro".

Considere o conjunto $\{1,\ldots,a\}$ munido da métrica discreta, ou seja, dados $x,y \in \{1,\ldots,a\}$, tem-se $d(x,y) = 0$ se $x = y$ e $d(x,y) = 1$ se $x \neq y$.

Observe que a métrica $\rho(x,y) := \sum_{i=-\infty}^{\infty} 2^{-|i|}d(x(i),y(i))$ é compatível com a topologia produto de $X = \{1,\ldots,a\}^{\mathbb{Z}}$ e, além disso, X é compacto pelo Teorema de Tychonoff (veja o item (4) do 1.4 Theorem na página 224 de (DUGUNDJI, 1966)).

Suponha que $x(i) = y(i)$ e $x(-i) = y(-i)$ para todo $0 \leqslant i \leqslant k$. Então,

$$\begin{aligned}
\rho(x,y) &= \sum_{i=-\infty}^{\infty} 2^{-|i|}d(x(i),y(i)) \\
&= \sum_{i<-k} 2^{-|i|}d(x(i),y(i)) + \sum_{i>k} 2^{-|i|}d(x(i),y(i)) \\
&= \sum_{i>k} 2^{-i}d(x(-i),y(-i)) + \sum_{i>k} 2^{-i}d(x(i),y(i)) \\
&\leqslant 2\sum_{i=k+1}^{\infty} 2^{-i} = 2^{-(k-1)}.
\end{aligned} \tag{4}$$

Por outro lado, se $x(i) \neq y(i)$ ou $x(-i) \neq y(-i)$ para algum $0 \leqslant i \leqslant k$, então $\rho(x,y) = \sum_{i=-\infty}^{\infty} 2^{-|i|}d(x(i),y(i)) \geqslant 2^{-k}$.

Em outras palavras, se $\rho(x,y) < 2^{-k}$, então

$$x(i) = y(i) \text{ e } x(-i) = y(-i) \text{ para todo } 0 \leqslant i \leqslant k. \tag{5}$$

Agora, considere $n = \min\{i \geqslant 0 : x(i) \neq y(i) \text{ ou } x(-i) \neq y(-i)\}$ e a métrica $\rho'(x,y) := 2^{-n}$.

Note que, se $x(i) = y(i)$ e $x(-i) = y(-i)$ para todo $0 \leqslant i \leqslant k$, então tem-se $n = \min\{i \geqslant 0 : x(i) \neq y(i) \text{ ou } x(-i) \neq y(-i)\} > k$, o que nos dá $\rho'(x,y) < 2^{-k}$. Daí, por (5) vemos que $\rho'(x,y) \leqslant \rho(x,y)$.

Por outro lado, se $\rho'(x,y) = 2^{-n}$ com $n > 0$, então $x(i) = y(i)$ e $x(-i) = y(-i)$ para todo $0 \leqslant i \leqslant n-1$. Logo, por (4) segue que $\rho(x,y) \leqslant 2^{-(n-2)} = 4\rho'(x,y)$. Além disso, se $\rho'(x,y) = 2^{-0} = 1$, então é claro que $\rho(x,y) \leqslant 4$.

Portanto, em qualquer caso temos $\rho'(x,y) \leqslant \rho(x,y) \leqslant 4\rho'(x,y) \quad (x,y \in \{1,\ldots,a\}^{\mathbb{Z}})$.

Isso mostra que ρ e ρ' são *uniformemente equivalentes* sobre $\{1,\ldots,a\}^{\mathbb{Z}}$.

Afirmação: T é contínua.

De fato, se $\rho'(x,y) = 2^{-0} = 1$, então $\rho'(T(x),T(y)) \leqslant 1 = \rho'(x,y)$. Além disso, se $\rho'(x,y) = 2^{-n}$ e $m = \min\{i \geqslant 0 : x(i+1) \neq y(i+1) \text{ ou } x(-(i+1)) \neq y(-(i+1))\}$, então $m \geqslant n-1$, donde $\rho'(T(x),T(y)) = 2^{-m} \leqslant 2^{-n+1} = 2\rho'(x,y)$.

Isso mostra que T é contínua e, similarmente, vemos que T^{-1} é contínua. Portanto, T e T^{-1} são funções mensuráveis com respeito à σ-álgebra de Borel de X, a qual denotaremos por $\mathcal{X}$.

Vamos agora mostrar que existe uma medida de Borel probabilística μ que é T-invariante.

Com efeito, dados $i \in \mathbb{Z}$ e $j \in \{1,\ldots,a\}$, considere o conjunto $A_i^j := \{x : x(i) = j\}$.

Note que, se $k \in \mathbb{N}$ é suficientemente grande tal que $|i| < k$, então segue de (5) que $y(i) = x(i) = j$ sempre que $\rho(x,y) < 2^{-k}$. Ou seja, A_i^j é um conjunto aberto. Além disso, note que $\{1,\ldots,a\}^{\mathbb{Z}} \setminus A_i^j = \bigcup_{j' \neq j} A_i^{j'}$. Ou seja, A_i^j também é um conjunto fechado.

Recorde que um conjunto que é simultaneamente aberto e fechado é dito *clopen*.

Dados $m \leqslant n$ em $\mathbb{Z}$ e $(u_m,\ldots,u_n) \in \{1,\ldots,a\}^{n-m+1}$, dizemos que o conjunto $A = \{x \in \{1,\ldots,a\}^{\mathbb{Z}} : (x(m),\ldots,x(n)) = (u_m,\ldots,u_n)\}$ é um *cilindro elementar de comprimento* $r = n - m + 1$.

Afirmação: O conjunto de todos os cilindros elementares forma uma base para a topologia de $\{1,\ldots,a\}^{\mathbb{Z}}$.

Com efeito, note inicialmente que todo cilindro elementar é *clopen*. Reciprocamente, considere uma bola aberta $B_\rho(x;r)$ arbitrária e fixe $k \in \mathbb{N}$ suficientemente grande tal que $B_\rho(x;2^{-k}) \subset B_\rho(x;r)$.

Considere $y \in B_\rho(x;2^{-k})$ qualquer. Por (5) acima vemos que $x(i) = y(i)$ e $x(-i) = y(-i)$ para todo $0 \leqslant i \leqslant k$. Considere o seguinte cilindro elementar

$$A = \{z \in \{1,\ldots,a\}^{\mathbb{Z}} : (z(-k-1),\ldots,z(k+1)) = (y(-k-1),\ldots,y(k+1))\}.$$

Então, $x \in A$ e, por (4), obtemos $A \subset B_\rho(x;2^{-k}) \subset B_\rho(x;r)$; isto conclui a demonstração da afirmação.

Agora, dados $m \leqslant n$ em $\mathbb{Z}$ e $U \subset \{1,\ldots,a\}^{n-m+1}$, dizemos que o conjunto

$A = \{x \in \{1, \dots, a\}^{\mathbb{Z}} : (x(m), \dots, x(n)) \in U\}$ é um *cilindro de posto* $r = n - m + 1$.

Note que todo cilindro elementar de comprimento $r = n - m + 1$ é um cilindro de posto $r = n - m + 1$ e, reciprocamente, todo cilindro de posto $r = n - m + 1$ é uma união disjunta finita de cilindros elementares de comprimento $r = n - m + 1$.

Seja $\mathcal{C}_r$ a coleção de todos os cilindros de posto $r \geqslant 1$ e considere a coleção $\mathcal{S} = \{\varnothing\} \cup \mathcal{C}_1 \cup \mathcal{C}_2 \cup \dots$

Afirmação: $\mathcal{S}$ é uma álgebra de subconjuntos de $\{1, \dots, a\}^{\mathbb{Z}}$.

Observe que $\{1, \dots, a\}^{\mathbb{Z}} = \{x \in \{1, \dots, a\} : x(0) \in \{1, \dots, a\}\} \in \mathcal{S}$.

Sejam $C = \{x \in \{1, \dots, a\}^{\mathbb{Z}} : (x(m), \dots, x(n)) \in U\}$ com $U \subset \{1, \dots, a\}^{n-m+1}$, $m \leqslant n$ e $D = \{x \in \{1, \dots, a\}^{\mathbb{Z}} : (x(k), \dots, x(l)) \in U\}$ com $V \subset \{1, \dots, a\}^{l-k+1}$, $k \leqslant l$ cilindros arbitrários. Sejam $\alpha = \min\{m, k\}$ e $\beta = \max\{n, l\}$.

Considere U' o conjunto de todas as sequências em $\{1, \dots, a\}^{\beta-\alpha+1}$ cuja restrição a $\{1, \dots\ a\}^{n-m+1}$ pertencem a U e, similarmente, considere V' o conjunto de todas as sequências em $\{1, \dots, a\}^{\beta-\alpha+1}$ cuja restrição a $\{1, \dots, a\}^{l-k+1}$ pertencem a V.

Então, $C = \{x : (x(\alpha), \dots, x(\beta)) \in U'\}$ e $D = \{x : (x(\alpha), \dots, x(\beta)) \in V'\}$.

Daí, $C \cup D = \{x : (x(\alpha), \dots, x(\beta)) \in U' \cup V'\} \in \mathcal{S}$.

Note também que $C^c = \{x \in \{1, \dots, a\}^{\mathbb{Z}} : (x(m), \dots, x(n)) \in U^c\} \in S$; isto conclui a demonstração da afirmação.

Definimos uma função de conjunto $P : \mathcal{S} \to [0,1]$ da seguinte maneira: $P(\varnothing) = 0$ e $P(\{x : (x(m), \dots, x(n)) \in U\}) := \displaystyle\sum_{(u_m, \dots, u_n) \in U \subset \{1, \dots, a\}^{n-m+1}} a^{-(n-m+1)}$.

Para simplificar a notação, vamos escrever

$$P(\{x : (x(m), \dots, x(n)) \in U\}) := \sum_U a^{-(n-m+1)},$$

onde admite-se implicitamente que a soma é feita sobre todos os $(u_m, \dots, u_n) \in U \subset \{1, \dots, a\}^{n-m+1}$.

Note que

$$P(\{x : (x(m), \dots, x(n)) = (u_m, \dots, u_n)\}) = a^{-(n-m+1)}$$

para todo $(u_m, \dots, u_n) \in U$.

Afirmação: P está bem definida.

Com efeito, suponha que $\{x : (x(m), \dots, x(n)) \in U)\} = \{x : (x(k), \dots, x(l)) \in V)\}$ com $U \subset \{1, \dots, a\}^{n-m+1}$ e $V \subset \{1, \dots, a\}^{l-k+1}$.

Se $n - m + 1 = l - k + 1$, então $U = V$ e não há nada a demonstrar. Então, podemos supor, por simetria, que $n - m + 1 < l - k + 1$, donde $V = U \times \{1, \dots, a\}^{(l+m)-(k+n)}$.

Daí,

$$\begin{aligned}\sum_V a^{-(l-k+1)} &= \sum_U a^{-(n-m+1)} \sum_{\{1,\dots,a\}^{(l+m)-(k+n)}} a^{-[(l+m)-(k+n)]} \\ &= \sum_U a^{-(n-m+1)},\end{aligned}$$

o que mostra que P está bem definida.

Afirmação: P é finitamente aditiva.

De fato, suponha $C = \{x : (x(m), \ldots, x(n)) \in U)\}$ com $U \subset \{1, \ldots, a\}^{n-m+1}$ e $D = \{x : (x(k), \ldots, x(l)) \in V)\}$ com $V \subset \{1, \ldots, a\}^{l-k+1}$ tais que $C \cap D = \varnothing$. Sejam $\alpha = \min\{m, k\}$, $\beta = \max\{n, l\}$.

Considere U' o conjunto de todas as sequências em $\{1, \ldots, a\}^{\beta-\alpha+1}$ cuja restrição a $\{1, \ldots, a\}^{n-m+1}$ pertence a U e, similarmente, considere V' o conjunto de todas as sequências em $\{1, \ldots, a\}^{\beta-\alpha+1}$ cuja restrição a $\{1, \ldots, a\}^{l-k+1}$ pertence a V.

Daí, $C = \{x : (x(\alpha), \ldots, x(\beta)) \in U'\}$, $D = \{x : (x(\alpha), \ldots, x(\beta)) \in V'\}$, $C \cup D = \{x : (x(\alpha), \ldots, x(\beta)) \in U' \cup V'\}$ e $U' \cap V' = \varnothing$ pois $C \cap D = \varnothing$.

Consequentemente, como P está bem definida, temos:

$$\begin{aligned} P(C \cup D) &= \sum_{U' \cup V'} a^{-(\beta-\alpha+1)} \\ &= \sum_{U} a^{-(n-m+1)} + \sum_{V} a^{-(l-k+1)} \\ &= P(C) + P(D). \end{aligned}$$

Isto conclui a demonstração da afirmação.

Como toda união infinita contável de cilindros é um conjunto clopen e não podemos escrever um conjunto clopen não-vazio como uma união infinita contável de conjuntos clopen não-vazios, segue que P é também *contavelmente aditiva* sobre $\mathcal{S}$.

Portanto, segue do Teorema de extensão da Teoria da Medida (veja o Teorema A na página 54 de (HALMOS, Paul, 1956)) que P se estende a uma única medida μ definida sobre uma σ-álgebra $\mathcal{A}$ que contém $\mathcal{S}$.

Como os cilindros elementares formam uma base para a topologia de $\{1, \ldots, a\}^{\mathbb{Z}}$, temos que todo subconjunto aberto de $\{1, \ldots, a\}^{\mathbb{Z}}$ pode ser escrito como uma união contável de elementos de $\mathcal{S}$, donde segue que a σ-álgebra $\mathcal{A}$ contém os subconjuntos de Borel de $\{1, \ldots, a\}^{\mathbb{Z}}$, ou seja, $\mathcal{X} \subset \mathcal{A}$.

Desse modo, μ é uma medida de Borel.

Além disso, como $\mu(\{x : (x(m), \ldots, x(n)) \in \{1, \ldots, a\}^{n-m+1}\}) = 1$, segue que μ é uma medida de Borel *probabilística*, dita *medida de Bernoulli*.

Afirmação: T preserva μ.

De fato, escreva $L := \{1, \ldots, a\}$ e note que, para qualquer $U \subset L^{n-m+1}$ com $m \leqslant n$

em $\mathbb{Z}$ tem-se:

$$\begin{aligned}
\mu(T^{-1}(\{x : (x(m),\ldots,x(n)) \in U\})) &= \mu(\{x : (x(m+1),\ldots,x(n+1)) \in U\}) \\
&= \mu(\{x : (x(m),x(m+1),\ldots,x(n+1)) \in L \times U\}) \\
&= \sum_L a^{-1} \sum_U a^{-(n-m+1)} \\
&= \sum_U a^{-(n-m+1)} \\
&= \mu(\{x : (x(m),\ldots,x(n)) \in U\})
\end{aligned}$$

Como os cilindros geram a σ-álgebra $\mathcal{X}$, segue pelo Teorema 1.8 que T preserva μ. Isso prova a afirmação.

Portanto, vemos que $(X,\mathcal{X},\mu,T)$ é um sistema dinâmico, dito *deslocamento de Bernoulli*. Como dissemos anteriormente, o deslocamento de Bernoulli fornece um modelo matemático para o experimento aleatório que consiste em girar uma roleta com um número finito de saídas igualmente prováveis, "desde o passado remoto até o eterno futuro".

Agora, vejamos uma propriedade satisfeita pelo deslocamento de Bernoulli que será importante à frente.

Proposição 2.1 . Dados quaisquer cilindros disjuntos $C = \{x : (x(m),\ldots,x(n)) \in U\}$ com $U \subset \{1,\ldots,a\}^{n-m+1}$ e $D = \{x : (x(k),\ldots,x(l)) \in V\}$ com $V \subset \{1,\ldots,a\}^{l-k+1}$, temos

$$\mu(C \cap T^{-j}(D)) = \mu(C)\mu(T^{-j}(D)) = \mu(C)\mu(D)$$

sempre que $j \in \mathbb{N}$ é suficientemente grande.

Demonstração. Com efeito, sejam $\alpha = \min\{m,k\}$ e $\beta = \max\{n,l\}$. Considere U' o conjunto de todas as sequências em $\{1,\ldots,a\}^{\beta-\alpha+1}$ cuja restrição a $\{1,\ldots,a\}^{n-m+1}$ pertencem a U e, similarmente, considere V' o conjunto de todas as sequências em $\{1,\ldots,a\}^{\beta-\alpha+1}$ cuja restrição a $\{1,\ldots,a\}^{l-k+1}$ pertencem a V. Observe que $C = \{x : (x(\alpha),\ldots,x(\beta)) \in U'\}$ e $D = \{x : (x(\alpha),\ldots,x(\beta)) \in V'\}$. Além disso, note que $U' \cap V' = \varnothing$ pois C e D são disjuntos. Também, dado qualquer $j \in \mathbb{N}$ temos $T^{-j}(D) = \{x : (x(\alpha+j),\ldots,x(\beta+j)) \in V'\}$. Agora, escreva $L := \{1,\ldots,a\}$ e tome $j \in \mathbb{N}$ suficientemente grande tal que $\alpha + j > \beta$. Então,

$$C \cap T^{-j}(D) = \{x : (x(\alpha),\ldots,x(\beta),x(\beta+1),\ldots,x(\alpha+j),\ldots,x(\beta+j)) \in U' \times L^{\alpha+j-\beta} \times V'\}.$$

Daí,

$$\begin{aligned}
\mu(C \cap T^{-j}(D)) &= \mu(\{x : (x(\alpha),\ldots,x(\beta+j)) \in U' \times L^{\alpha+j-\beta} \times V'\}) \\
&= \sum_{U'} a^{-(\beta-\alpha+1)} \sum_{L^{\alpha+j-\beta}} a^{-(\alpha+j-\beta)} \sum_{V'} a^{-(\beta-\alpha+1)} \\
&= \sum_U a^{-(n-m+1)} \sum_V a^{-(l-k+1)} \\
&= \mu(C)\mu(D),
\end{aligned}$$

como queríamos demonstrar. Note que o mesmo é válido se C e D são, cada um, uma união finita de cilindros dois a dois disjuntos. $\square$

Tendo em vista a construção acima, podemos considerar a seguinte definição:

Definição 2.2 . Um sistema dinâmico $(X,\mathcal{X},\mu,T)$ é dito *deslocamento de Bernoulli* se existe uma partição finita $\mathcal{P} := \{P_1,\ldots,P_a\}$ satisfazendo:

(i) $\displaystyle\bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{P} = \mathcal{X} \ (\mathrm{mod}\ \mu)$;

(ii) $\displaystyle\mu\Big(\bigcap_{i=m}^{n} T^{i}P_{j_i}\Big) = \prod_{i=m}^{n}\mu(P_{j_i}),\ j_i \in \{1,\ldots,a\},\ m \leqslant i \leqslant n \text{ e } m \leqslant n \text{ em } \mathbb{Z}.$

Observação 2.3 . Na construção acima, uma partição $\mathcal{P} = \{P_1,\ldots,P_a\}$ que satisfaz as condições (i) e (ii) da Definição 2.2 é dada por $P_j := \{x : x(0) = j\} \text{ para cada } j \in \{1,\ldots,a\}$, onde a σ-álgebra $\mathcal{X}$ é a σ-álgebra de Borel sobre $\{1,\ldots,a\}^{\mathbb{Z}}$. Note que, neste caso, tem-se $\mu(P_1) = \ldots = \mu(P_a) = 1/a$. Evidentemente, para deslocamentos de Bernoulli gerais pode-se ter $\mu(P_i) \neq \mu(P_j)$ para algum par de índices $i \neq j$.

Terminamos este capítulo introduzindo uma notação para os deslocamentos de Bernoulli que é bastante utilizada. Suponha que $(X,\mathcal{X},\mu,T)$ é um deslocamento de Bernoulli e considere $\mathcal{P} = \{P_1,\ldots,P_a\}$ como na Definição 2.2. Para cada $j \in \{1,\ldots,a\}$ escreva $p_j := \mu(P_j)$. Sem perda de generalidade, podemos supor $p_1 \leqslant \ldots \leqslant p_a$. Neste caso, denotamos $(X,\mathcal{X},\mu,T)$ simplesmente por $B(p_1,\ldots,p_a)$.

3 ENTROPIA

Neste capítulo apresentamos o conceito de entropia em Teoria Ergódica, cuja introdução é devida a Andrey Kolmogorov (KOLMOGOROV, 1958) e cuja simplificação e sistematização é devida a Yakov Sinai (SINAI, 1959, 1989).

Seja $(X,\mathcal{X},\mu,T)$ um sistema dinâmico e considere $\mathcal{P} := \{P_1,\ldots,P_k\}$ uma partição finita *ordenada*. É conveniente pensarmos em $\mathcal{P}$ como sendo um *experimento aleatório* no seguinte sentido: para cada $x \in X$ selecionado de acordo com a distribuição de probabilidade μ, $\mathcal{P}$ fornece como resultado o átomo P_j que contém x. Ou seja, se $\mathcal{P} := \{P_1,\ldots,P_k\}$, então consideramos $\mathcal{P}$ como sendo a *variável aleatória*

$$\mathcal{P} : X \to \{1,\ldots,k\}$$

definida por $\mathcal{P}(x) = j \iff x \in P_j$ $(1 \leqslant j \leqslant k)$.

Definição 3.1 . O *conteúdo de informação* obtido ao realizarmos o experimento $\mathcal{P}$ é a função dada por $I_{\mathcal{P}}(x) := -\ln(\mu(P_j))$ para todo $x \in P_j$ $(1 \leqslant j \leqslant k)$.

Ou seja,

$$I_{\mathcal{P}}(x) = -\sum_{P\in\mathcal{P}} 1_P(x)\ln(\mu(P)).$$

Como $\lim\limits_{x\to 0} x\ln x = 0$, adotamos a seguinte convenção: $0\ln 0 = 0$.

Definição 3.2 . A *entropia* da partição $\mathcal{P}$ é, por definição, a *quantidade média de informação* obtida ao realizarmos o experimento $\mathcal{P}$, ou seja,

$$H(\mathcal{P}) = \int_X I_{\mathcal{P}}(x)d\mu = \int_X -\sum_{P\in\mathcal{P}} 1_P(x)\ln(\mu(P))d\mu.$$

Além disso, note que:

$$\begin{aligned}
\int_X -\sum_{P\in\mathcal{P}} 1_P(x)\ln(\mu(P))d\mu &= \int_X \mathbb{E}\Big(-\sum_{P\in\mathcal{P}} 1_P(x)\ln(\mu(P))\Big)d\mu \\
&= -\sum_{P\in\mathcal{P}} \ln\mu(P)\int_X \mathbb{E}(1_P)d\mu \\
&= -\sum_{P\in\mathcal{P}} \ln\mu(P)\int_X 1_P d\mu \\
&= \sum_{P\in\mathcal{P}} -\mu(P)\ln(\mu(P))
\end{aligned}$$

Ou seja,

$$H(\mathcal{P}) = \sum_{P\in\mathcal{P}} -\mu(P)\ln(\mu(P))$$

Dada a expressão de H é conveniente introduzirmos a função $\phi : [0,1] \to \mathbb{R}$ dada por $\phi(s) = -s\ln s$. Como $\phi'' \leqslant 0$, segue que ϕ é côncava, ou seja, $a\phi(s) + b\phi(t) \leqslant \phi(as + bt)$ para quaisquer $0 \leqslant a,b \leqslant 1$ tais que $a + b = 1$ e quaisquer $s,t \in [0,1]$. Mais geralmente, tem-se

$\sum_{i=1}^{k} p_i\phi(t_i) \leqslant \phi\Big(\sum_{i=1}^{k} p_i t_i\Big)$ para quaisquer $0 \leqslant p_1,\ldots,p_k \leqslant 1$ tais que $\sum_{i=1}^{k} p_i = 1$ e quaisquer $t_1,\ldots,t_k \in [0,1]$. Além disso, vale a igualdade se, e somente se, todos os $t_i's$ associados a $p_i's$ não-nulos são iguais.

Proposição 3.3 . Seja $\mathcal{P} = \{P_1,\ldots,P_k\}$ uma partição finita. Então, $H(\mathcal{P}) \leqslant \ln k$ e ocorre a igualdade se, e somente se, $\mu(P) = 1/k$ para todo $P \in \mathcal{P}$.

Demonstração. Para cada $i = 1,\ldots,k$ considere $p_i = 1/k$ e $t_i = \mu(P_i)$. Pela concavidade de ϕ, temos:

$$\frac{1}{k}H(\mathcal{P}) = \sum_{i=1}^{k} p_i\phi(t_i) \leqslant \phi\Big(\sum_{i=1}^{k} p_i t_i\Big) = \phi\Big(\frac{1}{k}\Big) = \frac{\ln(k)}{k}.$$

Portanto $H(\mathcal{P}) \leqslant \ln(k)$ e ocorre a igualdade se, e somente se, $\mu(P_i) = 1/k$ para todo $i = 1,\ldots,k$.

□

Agora, considere $\mathcal{P}$ e $\mathcal{Q}$ dois experimentos e observe que:

$$\begin{aligned}
H(\mathcal{P}\vee\mathcal{Q}) - H(\mathcal{Q}) &= \sum_{P\in\mathcal{P}}\sum_{Q\in\mathcal{Q}} -\mu(P\cap Q)\ln(\mu(P\cap Q)) - \sum_{Q\in\mathcal{Q}} -\mu(Q)\ln(\mu(Q)) \\
&= \sum_{P\in\mathcal{P}}\sum_{Q\in\mathcal{Q}} -\mu(P\cap Q)\ln\Big(\mu(P\cap Q)/\mu(Q)\Big) \\
&+ \sum_{P\in\mathcal{P}}\sum_{Q\in\mathcal{Q}} -\mu(P\cap Q)\ln(\mu(Q)) - \sum_{Q\in\mathcal{Q}} -\mu(Q)\ln(\mu(Q)) \\
&= \sum_{P\in\mathcal{P}}\sum_{Q\in\mathcal{Q}} -\mu(Q)\Big(\mu(P\cap Q)/\mu(Q)\Big)\ln\Big(\mu(P\cap Q)/(\mu(Q))\Big) \\
&+ \sum_{Q\in\mathcal{Q}}\sum_{P\in\mathcal{P}} -\mu(P\cap Q)\ln(\mu(Q)) + \sum_{Q\in\mathcal{Q}} \mu(Q)\ln(\mu(Q)) \\
&= -\sum_{Q\in\mathcal{Q}}\mu(Q)\sum_{P\in\mathcal{P}}\Big(\mu(P\cap Q)/\mu(Q)\Big)\ln\Big(\mu(P\cap Q)/\mu(Q)\Big) \\
&- \sum_{Q\in\mathcal{Q}}\mu(Q)\ln(\mu(Q)) + \sum_{Q\in\mathcal{Q}}\mu(Q)\ln(\mu(Q)) \qquad (6) \\
&= \sum_{Q\in\mathcal{Q}}\mu(Q)\sum_{P\in\mathcal{P}} -\Big(\mu(P\cap Q)/\mu(Q)\Big)\ln\Big(\mu(P\cap Q)/\mu(Q)\Big) \\
&= \sum_{P\in\mathcal{P}}\sum_{Q\in\mathcal{Q}}\mu(Q)\phi\Big(\mu(P\cap Q)/\mu(Q)\Big) \\
&\leqslant \sum_{P\in\mathcal{P}}\phi\Big(\sum_{Q\in\mathcal{Q}}\mu(Q)\Big(\mu(P\cap Q)/\mu(Q)\Big)\Big) \\
&= \sum_{P\in\mathcal{P}}\phi\Big(\sum_{Q\in\mathcal{Q}}\mu(P\cap Q)\Big) \\
&= \sum_{P\in\mathcal{P}}\phi(\mu(P)) \\
&= H(\mathcal{P}),
\end{aligned}$$

onde na desigualdade acima usamos o fato de a função ϕ ser côncava. Portanto, vemos que

$$H(\mathcal{P} \vee \mathcal{Q}) - H(\mathcal{Q}) \leqslant H(\mathcal{P}) \tag{7}$$

e ocorre a igualdade se, e somente se, $\mu(P) = \mu(P \cap Q)/\mu(Q)$ para todos $P \in \mathcal{P}$ e $Q \in \mathcal{Q}$, ou seja, se e somente se $\mathcal{P}$ e $\mathcal{Q}$ são independentes.

Observação 3.4 . Pela equação (6) temos que

$$\begin{aligned} H(\mathcal{P} \vee \mathcal{Q}) - H(\mathcal{Q}) &= \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} \mu(Q)\phi\Big(\mu(P \cap Q)/\mu(Q)\Big) \\ &= \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} -\mu(P \cap Q) \ln \Big(\mu(P \cap Q)/\mu(Q)\Big). \end{aligned}$$

Definição 3.5 . Dadas $\mathcal{P}$ e $\mathcal{Q}$ duas partições, a diferença $H(\mathcal{P} \vee \mathcal{Q}) - H(\mathcal{Q})$ é dita *entropia condicional* de $\mathcal{P}$ dado $\mathcal{Q}$ e é denotada por $H(\mathcal{P}|\mathcal{Q})$. Ou seja,

$$H(\mathcal{P}|\mathcal{Q}) = \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} -\mu(P \cap Q) \ln \Big(\mu(P \cap Q)/\mu(Q)\Big).$$

Intuitivamente, a entropia condicional $H(\mathcal{P}|\mathcal{Q})$ é a quantidade média de informação que se obtém ao realizar o experimento $\mathcal{P}$ dado que o experimento $\mathcal{Q}$ foi realizado.

Definição 3.6 . Dadas $\mathcal{P}$ e $\mathcal{Q}$ duas partições, o *conteúdo de informação condicional* de $\mathcal{P}$ dado $\mathcal{Q}$ é a função dada por:

$$I_{\mathcal{P}|\mathcal{Q}}(x) := \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} -1_{P \cap Q}(x) \ln \Big(\mu(P \cap Q)/\mu(Q)\Big).$$

Note que:

$$\begin{aligned} \int_X \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} 1_{P \cap Q}(x) \ln(\mu(P \cap Q)/\mu(Q)) d\mu &= \int_X \mathbb{E}\Big(\sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} 1_{P \cap Q}(x) \ln \Big(\frac{\mu(P \cap Q)}{\mu(Q)}\Big)\Big) d\mu \\ &= \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} \ln \Big(\frac{\mu(P \cap Q)}{\mu(Q)}\Big) \int_X \mathbb{E}(1_{P \cap Q}) d\mu \\ &= \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} \ln \mu(P \cap Q)/\mu(Q) \int_X 1_{P \cap Q} d\mu \\ &= \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} \mu(P \cap Q) \ln \Big(\mu(P \cap Q)/\mu(Q)\Big) \end{aligned}$$

Ou seja,

$$H(\mathcal{P}|\mathcal{Q}) = \int_X I_{\mathcal{P}|\mathcal{Q}} d\mu.$$

Além disso, pelo que vimos acima, tem-se $H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P})$ se, e somente se, $\mathcal{P}$ e $\mathcal{Q}$ são independentes. Em outras palavras, tem-se $H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P})$ precisamente quando a realização do experimento $\mathcal{Q}$ não modifica a quantidade média de informação obtida ao realizar o experimento $\mathcal{P}$. Ou seja,

$$H(\mathcal{P} \vee \mathcal{Q}) = H(\mathcal{P}) + H(\mathcal{Q}) \iff H(\mathcal{P}|\mathcal{Q}) = H(\mathcal{P}) \iff \mathcal{P} \text{ e } \mathcal{Q} \text{ são independentes.} \tag{8}$$

Abaixo temos outras propriedades da função H, a saber:

Proposição 3.7 . Sejam $\mathcal{P}$,$\mathcal{Q}$ e $\mathcal{R}$ partições. Valem:

(i) $H(\mathcal{P} \vee \mathcal{Q}|R) = H(\mathcal{P}|\mathcal{R}) + H(\mathcal{Q}|\mathcal{P} \vee \mathcal{R})$;

(ii) Se $\mathcal{P} \prec \mathcal{Q}$, então $H(\mathcal{P}|\mathcal{R}) \leqslant H(\mathcal{Q}|\mathcal{R})$ e $H(\mathcal{R}|\mathcal{Q}) \leqslant H(\mathcal{R}|\mathcal{P})$. Em particular, se $\mathcal{P} \prec \mathcal{Q}$ e $\mathcal{R} = \{X\}$, então $H(\mathcal{P}) = H(\mathcal{P}|\{X\}) \leqslant H(\mathcal{Q}|\{X\}) = H(\mathcal{Q})$;

(iii) $\mathcal{P} \prec \mathcal{Q}$ se, e somente se, $H(\mathcal{P}|\mathcal{Q}) = 0$;

(iv) $H(\mathcal{P}|\mathcal{R}) \leqslant H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}|\mathcal{R})$.

Demonstração. (i):Observe que

$$\begin{aligned}
H(\mathcal{P} \vee \mathcal{Q}|\mathcal{R}) &= \sum_{P,Q,R} -\mu(P \cap Q \cap R) \ln \Big(\frac{\mu(P \cap Q \cap R)}{\mu(R)}\Big) \\
&= \sum_{P,Q,R} -\mu(P \cap Q \cap R)\Big(\ln \Big(\frac{\mu(P \cap Q \cap R)}{\mu(P \cap R)}\Big) + \ln \Big(\frac{\mu(P \cap R)}{\mu(R)}\Big)\Big).
\end{aligned}$$

Além disso,

$$\begin{aligned}
\sum_{P,Q,R} -\mu(P \cap Q \cap R) \ln \Big(\frac{\mu(P \cap R)}{\mu(R)}\Big) &= \sum_{P,R} \sum_{Q} -\mu(P \cap Q \cap R) \ln \Big(\frac{\mu(P \cap R)}{\mu(R)}\Big) \\
&= \sum_{P,R} -\mu(P \cap R \cap \sum_{Q \in \mathcal{Q}} Q) \ln \Big(\frac{\mu(P \cap R)}{\mu(R)}\Big) \\
&= \sum_{P,R} -\mu(P \cap R) \ln \Big(\frac{\mu(P \cap R)}{\mu(R)}\Big) \\
&= H(\mathcal{P}|\mathcal{R}).
\end{aligned}$$

Por outro lado, escrevendo $P \cap R = S$ temos:

$$\begin{aligned}
\sum_{P,Q,R} -\mu(P \cap Q \cap R) \ln \Big(\frac{\mu(P \cap Q \cap R)}{\mu(P \cap R)}\Big) &= \sum_{S \in \mathcal{P} \vee \mathcal{R}, Q \in \mathcal{Q}} -\mu(S \cap Q) \ln \Big(\frac{\mu(S \cap Q)}{\mu(S)}\Big) \\
&= H(\mathcal{Q}|\mathcal{P} \vee \mathcal{R}),
\end{aligned}$$

donde concluímos (i).

(ii): Suponha $\mathcal{P} \prec \mathcal{Q}$. Neste caso, temos:

$$\begin{aligned}
H(\mathcal{P}|\mathcal{R}) &= \sum_{P \in \mathcal{P}} \sum_{R \in \mathcal{R}} -\mu(P \cap R) \ln \Big(\frac{\mu(P \cap R)}{\mu(R)}\Big) \\
&= \sum_{P} \sum_{R} \sum_{Q \subset P} -\mu(Q \cap R) \ln \Big(\frac{\mu(P \cap R)}{\mu(R)}\Big) \\
&\leqslant \sum_{P} \sum_{R} \sum_{Q \subset P} -\mu(Q \cap R) \ln \Big(\frac{\mu(Q \cap R)}{\mu(R)}\Big) \\
&= H(\mathcal{Q}|\mathcal{R}),
\end{aligned}$$

onde a desigualdade decorre do fato de a função $-\ln(x)$ ser decrescente juntamente com a relação $\mu(Q \cap R) \leqslant \mu(P \cap R)$ válida sempre que $Q \subset P$.

Agora observe que, dados $P \in \mathcal{P}$ e $R \in \mathcal{R}$, tem-se:

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

e, além disso, $\displaystyle\sum_{Q \subset P} \frac{\mu(Q)}{\mu(P)} = 1.$

Logo, pela concavidade da função ϕ obtemos:

$$\phi\Big(\frac{\mu(R \cap P)}{\mu(P)}\Big) \geqslant \sum_{Q \subset P} \frac{\mu(Q)}{\mu(P)} \phi\Big(\frac{\mu(R \cap Q)}{\mu(Q)}\Big).$$

Consequentemente,

$$\begin{aligned}
H(\mathcal{R}|\mathcal{P}) &= \sum_{P,R} -\mu(P \cap R) \ln\Big(\frac{\mu(P \cap R)}{\mu(P)}\Big) \\
&= \sum_{P,R} -\mu(P) \frac{\mu(P \cap R)}{\mu(P)} \ln\Big(\frac{\mu(P \cap R)}{\mu(P)}\Big) \\
&= \sum_{P,R} \mu(P) \phi\Big(\frac{\mu(P \cap R)}{\mu(P)}\Big) \\
&\geqslant \sum_{P,R} \mu(P) \sum_{Q \subset P} \frac{\mu(Q)}{\mu(P)} \phi\Big(\frac{\mu(R \cap Q)}{\mu(Q)}\Big) \\
&= \sum_{Q,R} \mu(Q) \phi\Big(\frac{\mu(R \cap Q)}{\mu(Q)}\Big) \\
&= \sum_{Q,R} \mu(R \cap Q) \ln\Big(\frac{\mu(R \cap Q)}{\mu(Q)}\Big) = H(\mathcal{R}|\mathcal{Q}),
\end{aligned}$$

o que termina a demonstração do item (ii).

(iii): Segue da definição de entropia condicional que $H(\mathcal{P}|\mathcal{Q}) = 0$ se, e somente se, para todos $P \in \mathcal{P}$ e $Q \in \mathcal{Q}$ tem-se $\mu(P \cap Q) = 0$ ou $\mu(P \cap Q) = \mu(Q)$. Portanto, temos $H(\mathcal{P}|\mathcal{Q}) = 0$ se, e somente se, Q é disjunto de P a menos de um conjunto de medida nula, ou Q está contido em P a menos de um conjunto de medida nula. Em outras palavras, temos $H(\mathcal{P}|\mathcal{Q}) = 0$ se, e somente se, $\mathcal{P} \prec \mathcal{Q}$, o que termina a demonstração do item (iii).

(iv): Utilizando os itens (i) e (ii) obtemos:

$$H(\mathcal{P}|\mathcal{R}) \leqslant H(\mathcal{P} \vee \mathcal{Q}|\mathcal{R}) = H(\mathcal{Q}|\mathcal{R}) + H(\mathcal{P}|\mathcal{Q} \vee \mathcal{R}) \leqslant H(\mathcal{Q}|\mathcal{R}) + H(\mathcal{P}|\mathcal{Q}).$$

Isso termina a demonstração da proposição. □

A próxima proposição garante que a entropia condicional possui uma certa propriedade de continuidade.

Proposição 3.8 . Dados $k \geqslant 1$ e $\varepsilon > 0$ existe $\delta > 0$ tal que, para quaisquer partições finitas $\mathcal{P} = \{P_1, \ldots, P_k\}$ e $\mathcal{Q} = \{Q_1, \ldots, Q_k\}$ que satisfazem $\mu(P_i \triangle Qi) < \delta$ para todo $i \in \{1, \ldots, k\}$ tem-se $H(\mathcal{Q}|\mathcal{P}) < \varepsilon$.

Demonstração. Sejam $\varepsilon > 0$ e $k \geqslant 1$ arbitrários. Como ϕ é contínua em $[0,1]$, existe $\rho > 0$ tal que $\phi(x) < \varepsilon/k^2$, sempre que $x \in [0,\rho) \cup (1-\rho,1]$. Considere $\delta = \rho/k$ e suponha que $\mathcal{P} = \{P_1, \ldots, P_k\}$ e $\mathcal{Q} = \{Q_1, \ldots, Q_k\}$ são partições finitas tais que $\mu(P_i \triangle Q_i) < \delta$ para todo $i \in \{1, \ldots, k\}$. Agora, denote por $\mathcal{R}$ a partição formada pelos elementos da forma $P_i \cap Q_j$ com $i \neq j$ juntamente com o conjunto $\displaystyle\bigcup_{i=1}^{k} P_i \cap Q_i$.

Como $P_i \cap Q_j \subset P_i \triangle Q_i$ sempre que $i \neq j$, temos $\mu(P_i \cap Q_j) \leqslant \mu(P_i \triangle Q_i) < \delta$. Além disso,

$$\mu\Big(\bigcup_{i=1}^{k} P_i \cap Q_i\Big) \geqslant \sum_{i=1}^{k} (\mu(P_i) - \mu(P_i \triangle Q_i)) > \sum_{i=1}^{k} \mu(P_i - \delta) > 1 - \rho.$$

Portanto,

$$H(\mathcal{R}) = \sum_{R \in \mathcal{R}} \phi(\mu(R)) < \#\mathcal{R}\frac{\varepsilon}{k^2} \leqslant \varepsilon.$$

Pela definição da partição $\mathcal{R}$ temos que $\mathcal{P} \vee \mathcal{Q} = \mathcal{P} \vee \mathcal{R}$ donde, por (7), obtemos:

$$\begin{aligned} H(\mathcal{Q}|\mathcal{P}) &= H(\mathcal{P} \vee \mathcal{Q}) - H(\mathcal{P}) = H(\mathcal{P} \vee \mathcal{R}) - H(P) \\ &= H(\mathcal{R}|\mathcal{P}) \leqslant H(\mathcal{R}) < \varepsilon, \end{aligned}$$

como queríamos demonstrar.

□

Agora, dado um experimento $\mathcal{P} = \{P_1, \ldots, P_k\}$, como μ é T-invariante podemos pensar no experimento

$$T^{-1}\mathcal{P} = \{T^{-1}P_1, \ldots, T^{-1}P_k\}$$

como sendo uma *repetição* do experimento $\mathcal{P}$. Observe que os experimentos $\mathcal{P}$ e $T^{-1}\mathcal{P}$ não necessariamente são independentes. Além disso, para cada $x \in X$ selecionado de acordo com μ, o experimento aleatório $\mathcal{P} \vee T^{-1}\mathcal{P}$ nos diz o átomo de $\mathcal{P} \vee T^{-1}\mathcal{P}$ que contém x. Equivalentemente, para cada $x \in X$ selecionado de acordo com a distribuição de probabilidade μ, o experimento aleatório $\mathcal{P} \vee T^{-1}\mathcal{P}$ nos diz o átomo de $\mathcal{P}$ que contém x e também o átomo de $\mathcal{P}$ que contém $T(x)$. Mais geralmente, dado $n \geqslant 1$ arbitrário, para cada $x \in X$ selecionado de acordo com a distribuição de probabilidade μ, o experimento aleatório

$$\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} = \mathcal{P} \vee T^{-1}\mathcal{P} \vee \ldots \vee T^{-(n-1)}\mathcal{P}$$

nos diz o átomo P_{j_i} que contém o ponto $T^i x$ para todo $0 \leqslant i \leqslant n-1$. Ou seja, para cada $n \geqslant 1$ fixado, $\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} : X \to \{1, \ldots ,k\}^n$ é a variável aleatória dada por:

$$\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big)(x) = (j_0,j_1 \ldots ,j_{n-1}) \iff T^i(x) \in P_{j_i} \text{ para cada } 0 \leqslant i \leqslant n-1.$$

De modo equivalente, $\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}$ corresponde a n repetições consecutivas - não necessariamente independentes entre si - do experimento aleatório $\mathcal{P}$.

Além disso, como T é um automorfismo, segue que o experimento $\mathcal{P}$ induz um *processo estocástico bilateral* (DOOB, 1990) $\Lambda : X \to \{1, \ldots ,k\}^{\mathbb{Z}}$ de modo natural, a saber:

$$\Lambda(x) = (\ldots ,j_{-1},j_0,j_1, \ldots) \iff T^i(x) \in P_{j_i} \text{ para cada } i \in \mathbb{Z}.$$

Com efeito, observe que Λ induz uma medida de Borel probabilística ν sobre o espaço compacto $\{1, \ldots ,k\}^{\mathbb{Z}}$ de todas as sequências infinitas bilaterais de símbolos $1, \ldots ,k$, a saber:

$$\nu(B) := \mu(\Lambda^{-1}(B)) \text{ para todo conjunto de Borel } B \subset \{1, \ldots ,k\}^{\mathbb{Z}}.$$

Também, se $S : \{1, \ldots ,k\}^{\mathbb{Z}} \to \{1, \ldots ,k\}^{\mathbb{Z}}$ é a transformação de deslocamento à esquerda $S((a_i)_{i\in\mathbb{Z}}) = (a_{i+1})_{i\in\mathbb{Z}}$, então $S \circ \Lambda = \Lambda \circ T$. Daí, ν é S-invariante pois, para cada conjunto de Borel $B \subset \{1, \ldots ,k\}^{\mathbb{Z}}$ tem-se:

$$\begin{aligned} \nu(S^{-1}(B)) &= \mu(\Lambda^{-1}(S^{-1}(B))) \\ &= \mu((S \circ \Lambda)^{-1}(B)) \\ &= \mu((\Lambda \circ T)^{-1}(B)) \\ &= \mu(T^{-1}(\Lambda^{-1}(B))) \\ &= \mu(\Lambda^{-1}(B)) \\ &= \nu(B). \end{aligned}$$

Por esta razão é comum, em Teoria Ergódica, denominar o par $(T,\mathcal{P})$ (ou $(\mathcal{P},T)$) de *processo*.

Observação 3.9 . Na construção acima, a medida ν não necessariamente é a medida de Bernoulli. No caso em que a medida ν é a medida de Bernoulli dizemos que o processo $(T,\mathcal{P})$ é um *processo Bernoulli*. Ou ainda, $(T,\mathcal{P})$ é dito um *processo independente e identicamente distribuído*.

Observação 3.10 . Seja $(X,\mathcal{X},\mu,T)$ um sistema dinâmico que modela algum fenômeno físico. Na prática, para estudarmos o fenômeno em questão realizamos uma *medição* a cada instante de interesse a qual possui, necessariamente, *precisão finita*. Tal precisão finita das medições induz uma partição finita $\mathcal{P} := \{P_1, \ldots ,P_k\}$ do espaço de configurações X em conjuntos de medida positiva; este é o nosso "experimento". O processo $(T,\mathcal{P})$ pode ser interpretado como a "história completa" das medições do fenômeno em estudo e é, de fato, a única maneira de obter informações, na prática, sobre o fenômeno a partir do modelo idealizado $(X,\mathcal{X},\mu,T)$.

Agora, sejam $m,n \geqslant 1$ e $\mathcal{P} = \{P_1, \ldots, P_k\}$ uma partição finita. Sem perda de generalidade, podemos supor que $m \geqslant n$. Daí, observe que:

$$\begin{aligned} H\Big(\bigvee_{i=0}^{m+n-1} T^{-i}\mathcal{P}\Big) &= H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} \vee \bigvee_{i=n}^{m+n-1} T^{-i}\mathcal{P}\Big) \\ &\leqslant H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) + H\Big(\bigvee_{i=n}^{m+n-1} T^{-i}\mathcal{P}\Big) \\ &= H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) + H\Big(\bigvee_{i=0}^{m-1} T^{-i}\mathcal{P}\Big), \end{aligned}$$

onde na desigualdade usamos (7) e na última igualdade usamos o fato de μ ser T-invariante.

Portanto, vemos que a sequência $\Big(H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big)\Big)_{n\geqslant 1}$ é subaditiva. Consequentemente, segue do Lema 1.52 que existe

$$\lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big)\Big)$$

e este coincide com $\inf_{n\geqslant 1} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big)\Big)$.

Definição 3.11 . A entropia de T com respeito à partição $\mathcal{P}$ é a *média assintótica* da quantidade média de informação obtida ao realizarmos n repetições consecutivas do experimento $\mathcal{P}$ quando $n \to \infty$. Ou seja,

$$H(T,\mathcal{P}) = \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big).$$

Esta é a entropia do processo $(T,\mathcal{P})$.

Observe que, dados $\mathcal{P}$ e $\mathcal{Q}$ dois experimentos tais que $\mathcal{P} \prec \mathcal{Q}$ tem-se

$$\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} \prec \bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q} \text{ para todo } n \geqslant 1.$$

Portanto, pelo item (ii) da Proposição 3.7, temos $H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) \leqslant H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q}\Big)$ para todo $n \geqslant 1$, donde conclui-se:

$$\mathcal{P} \prec \mathcal{Q} \Rightarrow H(T,\mathcal{P}) \leqslant H(T,\mathcal{Q}).$$

Observação 3.12 . Excetuando-se a Proposição 3.8, tudo o que foi exposto anteriormente se estende para partições *infinitas enumeráveis* $\mathcal{P} := \{P_1,P_2,\ldots\}$ tais que $H(\mathcal{P}) < +\infty$. Neste caso, em particular, nas definições sobre entropia temos *séries* em lugar de somas finitas.

Definição 3.13 . A entropia do sistema dinâmico $(X,\mathcal{X},\mu,T)$ é o supremo das entropias dos processos $(T,\mathcal{P})$ tomado sobre todas as partições finitas $\mathcal{P}$. Ou seja,

$$H(T) := \sup_{\mathcal{P}} H(T,\mathcal{P}),$$

onde o supremo é tomado sobre todas as partições finitas $\mathcal{P}$.

É praticamente impossível calcular a entropia de um sitema dinâmico a partir da definição. Desse modo, nosso próximo objetivo é demonstrar o Teorema de Kolmogorov-Sinai, o qual fornece uma maneira de a calcularmos.

Para tal, vamos precisar de alguns resultados auxiliares.

Lema 3.14 . Sejam $\mathcal{P}$ e $\mathcal{Q}$ partições com entropia finita. Vale:

$$H(T,\mathcal{P} \vee \mathcal{Q}) \leqslant H(T,\mathcal{P}) + H(T,\mathcal{Q}).$$

Demonstração. Pela relação (7) temos

$$\begin{aligned}
H(T,\mathcal{P} \vee \mathcal{Q}) &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} T^{-i}(\mathcal{P} \vee \mathcal{Q})\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P} \vee \bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q}\Big) \\
&\leqslant \lim_{n\to\infty} \frac{1}{n} \Big(H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) + H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q}\Big)\Big) \\
&= H(T,\mathcal{P}) + H(T,\mathcal{Q}).
\end{aligned}$$

□

Lema 3.15 . Se $\mathcal{P}$ e $\mathcal{Q}$ são partições com entropia finita, então

$$H(T,\mathcal{Q}) \leqslant H(T,\mathcal{P}) + H(\mathcal{Q}|\mathcal{P}).$$

Demonstração. Pelo item (i) da Proposição 3.7, para todo $n \geqslant 1$ temos:

$$\begin{aligned}
H\Big(\bigvee_{i=0}^{n} T^{-i}\mathcal{Q} | \bigvee_{j=0}^{n} T^{-j}\mathcal{P}\Big) &= H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q} \vee T^{-n}\mathcal{Q} | \bigvee_{j=0}^{n-1} T^{-j}\mathcal{P} \vee T^{-n}\mathcal{P}\Big) \\
&= H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q} | \bigvee_{j=0}^{n-1} T^{-j}\mathcal{P} \vee T^{-n}\mathcal{P}\Big) \\
&+ H\Big(T^{-n}\mathcal{Q} | \bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q} \vee \bigvee_{j=0}^{n} T^{-j}\mathcal{P}\Big) \\
&\leqslant H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q} | \bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}\Big) + H(T^{-n}\mathcal{Q}|T^{-n}\mathcal{P}) \\
&= H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q} | \bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}\Big) + H(\mathcal{Q}|\mathcal{P}).
\end{aligned}$$

Daí, obtemos

$$H\Big(\bigvee_{i=0}^{n} T^{-i}\mathcal{Q} | \bigvee_{j=0}^{n} T^{-j}\mathcal{P}\Big) \leqslant nH(\mathcal{Q}|\mathcal{P}) \text{ para todo } n \geqslant 1.$$

Finalmente, para todo $n \geqslant 1$ tem-se:

$$\begin{aligned}\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q}\Big) &\leqslant \frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q} \vee \bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}\Big) \\ &= \frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) + \frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{Q}|\bigvee_{j=0}^{n-1} T^{-j}\mathcal{P}\Big) \\ &\leqslant \frac{1}{n}H\Big(\bigvee_{i=0}^{n} T^{-i}\mathcal{P}\Big) + \frac{n-1}{n}H(\mathcal{Q}|\mathcal{P}),\end{aligned}$$

donde tomando o limite quando $n \to \infty$ concluímos:

$$H(T,\mathcal{Q}) \leqslant H(T,\mathcal{P}) + H(\mathcal{Q}|\mathcal{P}).$$

□

Lema 3.16 . Seja $\mathcal{P} = \{P_n : n \in \mathbb{N}\}$ uma partição com entropia finita e considere, para cada $k \geqslant 1$, a partição finita $\mathcal{P}_k = \Big\{P_1, \ldots, P_k, X \setminus \bigcup_{i=1}^{n} P_i\Big\}$. Vale:

$$\lim_{k\to\infty} H(\mathcal{P}|\mathcal{P}_k) = 0.$$

Demonstração. Escreva $P_0 := X \setminus \bigcup_{i=1}^{k} P_i$. Por definição, temos

$$H(\mathcal{P}|\mathcal{P}_k) = \sum_{i=0}^{k}\sum_{j\geqslant 1} -\mu(P_i \cap P_j)\ln\Big(\frac{\mu(P_i \cap P_j)}{\mu(P_i)}\Big).$$

Agora, observe que todos os termos com $i \geqslant 1$ desaparecem uma vez que, neste caso, se $i \neq j$, então $\mu(P_i \cap P_j) = 0$ e, se $i = j$, então $\ln\Big(\frac{\mu(P_i \cap P_j)}{\mu(P_i)}\Big) = 0$.

Agora, para $i = 0$ temos $\mu(P_i \cap P_j) = 0$ se $j \leqslant k$ e, se $j > k$, então $\mu(P_i \cap P_j) = \mu(P_j)$. Portanto,

$$H(\mathcal{P}|\mathcal{P}_k) = \sum_{j>k} -\mu(P_j)\ln\Big(\frac{\mu(P_j)}{\mu(P_0)}\Big) \leqslant \sum_{j>k} -\mu(P_j)\ln(\mu(P_j)) \leqslant H(\mathcal{P}) < \infty.$$

Pela convergência da série $\sum -\mu(P_j)\ln\Big(\frac{\mu(P_j)}{\mu(P_0)}\Big)$ obtém-se $\lim_{k\to\infty} H(\mathcal{P}|\mathcal{P}_k) = 0$, como queríamos demonstrar.

□

Observação 3.17 . Tendo em vista o Lema 3.15 e o Lema 3.16 poderíamos ter definido $H(T)$ como sendo o supremo das entropias dos processos $(T,\mathcal{P})$ tomado sobre todas as partições *infinitas enumeráveis* $\mathcal{P}$ tais que $H(\mathcal{P}) < +\infty$. Com efeito, sejam $\mathcal{P}$ uma partição infinita enumerável com $H(\mathcal{P}) < +\infty$ e $\varepsilon > 0$ arbitrários. Utilizando a notação do Lema 3.16 existem $k \geqslant 1$ e uma partição finita $\mathcal{P}_k$ tais que $H(\mathcal{P}|\mathcal{P}_k) < \varepsilon$. Logo, pelo Lema 3.15 tem-se:

$$H(T,\mathcal{P}) - \varepsilon < H(T,\mathcal{P}_k).$$

Proposição 3.18 . Seja $\mathcal{P}$ uma partição com entropia finita. Então, para todo $k \geqslant 0$ temos $H(T,\mathcal{P}) = H\Big(T, \bigvee_{i=0}^{k} T^{-i}\mathcal{P}\Big)$. Além disso, se T é um automorfismo, então

$$H(T,\mathcal{P}) = H\Big(T, \bigvee_{i=-k}^{k} T^{i}\mathcal{P}\Big).$$

Demonstração. Dados $n \geqslant 1$ e $k \geqslant 0$ temos:

$$\bigvee_{j=0}^{n-1} T^{-j}\Big(\bigvee_{i=0}^{k} T^{-i}\mathcal{P}\Big) = \bigvee_{i=0}^{k} T^{-i}\mathcal{P} \vee \bigvee_{j=1}^{k+1} T^{-j}\mathcal{P} \vee \cdots \vee \bigvee_{l=n-1}^{n+k-1} T^{-l}\mathcal{P} = \bigvee_{i=0}^{n+k-1} T^{-i}\mathcal{P}.$$

Logo,

$$\begin{aligned}
H\Big(T, \bigvee_{j=0}^{k} T^{-j}\mathcal{P}\Big) &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{j=0}^{n-1} T^{-j}\Big(\bigvee_{i=0}^{k} T^{-i}\mathcal{P}\Big)\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n+k-1} T^{-i}\mathcal{P}\Big) \\
&= \lim_{n\to\infty} \frac{n+k-1}{n} \frac{1}{n+k-1} H\Big(\bigvee_{i=0}^{n+k-1} T^{-i}\mathcal{P}\Big) \\
&= H(T,\mathcal{P}).
\end{aligned}$$

Agora, suponha que T é um automorfismo. Para cada $n \geqslant 1$ e $k \geqslant 0$ tem-se:

$$\bigvee_{j=0}^{n-1} T^{-j}\Big(\bigvee_{i=-k}^{k} T^{-i}\mathcal{P}\Big) = \bigvee_{i=-k}^{k} T^{-i}\mathcal{P} \vee \bigvee_{j=-k+1}^{k+1} T^{-j}\mathcal{P} \vee \cdots \vee \bigvee_{l=n-k-1}^{n+k-1} T^{-l}\mathcal{P} = \bigvee_{i=-k}^{n+k-1} T^{-i}\mathcal{P}.$$

Portanto, temos:

$$\begin{aligned}
H\Big(T, \bigvee_{j=-k}^{k} T^{-j}\mathcal{P}\Big) &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{j=0}^{n-1} T^{-j}\Big(\bigvee_{i=-k}^{k} T^{-i}\mathcal{P}\Big)\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=-k}^{n+k-1} T^{-i}\mathcal{P}\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(T^{k}\Big(\bigvee_{i=0}^{n+2k-1} T^{-i}\mathcal{P}\Big)\Big) \\
&= \lim_{n\to\infty} \frac{n+2k-1}{n} \frac{1}{n+2k-1} H\Big(\bigvee_{i=0}^{n+2k-1} T^{-i}\mathcal{P}\Big) \\
&= H(T,\mathcal{P}).
\end{aligned}$$

□

Proposição 3.19 . Seja $\mathcal{P}$ uma partição com entropia finita. Então, para todo $k \geqslant 1$ temos $H\Big(T^{k}, \bigvee_{j=0}^{k-1} T^{-j}\mathcal{P}\Big) = kH(T,\mathcal{P})$.

Demonstração. □

Realmente,

$$
\begin{aligned}
H\Big(T^k, \bigvee_{j=0}^{k-1} T^{-j}\mathcal{P}\Big) &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{j=0}^{n-1} T^{-kj}\Big(\bigvee_{i=0}^{k-1} T^{-i}\mathcal{P}\Big)\Big)\\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{j=0}^{n-1} T^{-kj}\Big(\mathcal{P} \vee T^{-1}\mathcal{P} \vee \ldots \vee T^{-(k-1)}\mathcal{P}\Big)\Big)\\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{j=0}^{n-1} T^{-kj}\mathcal{P} \vee T^{-kj-1}\mathcal{P} \vee \ldots \vee T^{-kj-k+1}\mathcal{P}\Big)\\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{j=0}^{nk-1} T^{-j}\mathcal{P}\Big)\\
&= \lim_{n\to\infty} \frac{k}{nk} H\Big(\bigvee_{j=0}^{nk-1} T^{-j}\mathcal{P}\Big)\\
&= k \lim_{n\to\infty} \frac{1}{nk} H\Big(\bigvee_{j=0}^{nk-1} T^{-j}\mathcal{P}\Big)\\
&= kH(T,\mathcal{P}).
\end{aligned} \tag{9}
$$

Proposição 3.20 . Para todo $k \geqslant 0$ vale $H(T^k) = kH(T)$. Além disso, se T é um automorfismo, então tem-se $H(T^k) = |k|H(T)$ para todo $k \in \mathbb{Z}$.

Demonstração. Primeiro, se $k = 0$, então $T^0 = id$ e o resultado segue neste caso, pois $H(id) = 0$.

Agora, sejam $k \geqslant 1$ e $\mathcal{P}$ uma partição com entropia finita arbitrários.

A Proposição 3.19 nos diz que $H(T,\mathcal{P}) = \dfrac{1}{k} H\Big(T^k, \bigvee_{j=0}^{k-1} T^{-j}\mathcal{P}\Big)$.

Daí, $H(T,\mathcal{P}) = \dfrac{1}{k} H\Big(T^k, \bigvee_{j=0}^{k-1} T^{-j}\mathcal{P}\Big) \leqslant \dfrac{1}{k} H(T^k)$ e, portanto, $H(T) \leqslant \dfrac{1}{k} H(T^k)$. Por outro lado,

$$
H(T^k,\mathcal{P}) \leqslant H\Big(T^k, \bigvee_{i=0}^{k-1} T^{-i}\mathcal{P}\Big) = kH(T,\mathcal{P}) \leqslant kH(T), \tag{10}
$$

donde $H(T^k) \leqslant kH(T)$. Portanto, $H(T^k) = kH(T)$.

Agora, suponha que T é um automorfismo e observe que, como μ é T-invariante e $\bigvee_{i=0}^{n-1} T^i\mathcal{P} = \mathcal{P} \vee \cdots \vee T^{n-1}\mathcal{P} = T^{n-1}(\mathcal{P} \vee \cdots \vee T^{-n+1}\mathcal{P})$, temos:

$$
H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}\Big) = H\Big(T^{n-1} \bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) = H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big).
$$

Logo,

$$
H(T^{-1},\mathcal{P}) = H(T,\mathcal{P}) \tag{11}
$$

e, portanto, $H(T^{-1}) = H(T)$.

Finalmente, se $k \in \mathbb{Z}$ com $k < 0$, então temos $H(T^k) = H((T^{-1})^{-k}) = -kH(T^{-1}) = |k|H(T)$.

$\square$

Proposição 3.21 . Sejam $(X,\mathcal{X},\mu,T)$ um sistema dinâmico e $\mathcal{P}$ uma partição. Tem-se

$$H(T,\mathcal{P}) = \frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) \text{ para cada } n \geqslant 1.$$

Demonstração. De fato, fixe $n \geqslant 1$ arbitrário e observe que

$$\begin{aligned}
H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) &= H\Big(\mathcal{P}\vee\ldots\vee T^{n-1}\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) \\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) + H\Big(T\mathcal{P}|\mathcal{P}\vee\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) \\
&+ \ldots + H\Big(T^{n-1}\mathcal{P}|\mathcal{P}\vee T\mathcal{P}\vee\ldots\vee T^{n-2}\mathcal{P}\vee\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) \\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) + H\Big(\mathcal{P}|T^{-1}\mathcal{P}\vee\bigvee_{j=2}^{\infty} T^{-j}\mathcal{P}\Big) \\
&+ \ldots + H\Big(\mathcal{P}|T^{-(n-1)}\mathcal{P}\vee\ldots\vee T^{-1}\mathcal{P}\vee\bigvee_{j=n}^{\infty} T^{-j}\mathcal{P}\Big) \\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) + \ldots + H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) \\
&= nH(T,\mathcal{P}),
\end{aligned}$$

como queríamos demonstrar. $\square$

Proposição 3.22 . Seja $\mathcal{A}$ uma álgebra que gera $\mathcal{X}$. Para toda $\mathcal{Q}$ partição com $H(\mathcal{Q}) < +\infty$ e para todo $\varepsilon > 0$ existe uma partição finita $\mathcal{P} \subset \mathcal{A}$ tal que $H(\mathcal{Q}|\mathcal{P}) + H(\mathcal{P}|\mathcal{Q}) < 2\varepsilon$.

Demonstração. Para cada $k \geqslant 1$ considere $\mathcal{Q}_k = \Big\{Q_1,\ldots,Q_k, X\setminus\bigcup_{i=1}^{k} Q_i\Big\}$.

Seja $\varepsilon > 0$ arbitrário. Pelo Lema 3.16 existe $k \geqslant 1$ tal que $H(\mathcal{Q}|\mathcal{Q}_k) < \varepsilon/2$.

Seja $\delta > 0$ a ser especificado. Pelo Teorema 1.13, para cada $i = 1,\ldots,k$ existe $A_i \in \mathcal{A}$ tal que

$$\mu(Q_i\triangle A_i) < \frac{\delta}{2k^2}. \tag{12}$$

Considere $P_1 = A_1$ e, para cada $i = 2,\ldots,k$, $P_i = A_i\setminus\bigcup_{j=1}^{i-1} P_j$. Considere também $P_o = X\setminus\bigcup_{j=1}^{k} P_j$.

Observe que $\mathcal{P} = \{P_1,\ldots,P_k,P_0\}$ é uma partição finita de X tal que $\mathcal{P} \subset \mathcal{A}$.

Note que $P_1 \triangle A_1 = \varnothing$. Além disso, para cada $i = 2, \ldots, k$ temos:

$$P_i \triangle A_i = A_i \setminus P_i = A_i \cap (A_i \setminus \bigcup_{j=1}^{i-1} P_j)^c = A_i \cap (A_i^c \cup (\bigcup_{j=1}^{i-1} P_j)) = A_i \cap (\bigcup_{j=1}^{i-1} P_j) = \bigcup_{j=1}^{i-1} (A_i \cap P_J).$$

Portanto, se $x \in P_i \triangle A_i$, então existe $1 \leqslant j < i$ tal que $x \in A_i \cap P_j$, donde segue que $x \in A_i \cap A_j$. Como $Q_i \cap Q_j = \varnothing$, segue que $x \in (A_i \setminus Q_i) \cup (A_j \setminus Q_j)$. Daí, vemos que $P_i \triangle A_i \subset \bigcup_{j=1}^{i} (A_j \setminus Q_j) \subset \bigcup_{j=1}^{i} (A_j \triangle Q_j)$. Assim, por (12) obtemos $\mu(P_i \triangle A_i) < k\delta/2k^2 = \delta/2k$. Logo,

$$\mu(P_i \triangle Q_i) \leqslant \mu(P_i \triangle A_i) + \mu(A_i \triangle Q_i) < \delta/2k + \delta/2k^2 \leqslant \delta/k \text{ para cada } i = 1, \ldots, k. \quad (13)$$

Além disso, como $\mathcal{P}$ e $\mathcal{Q}_k$ são partições de X, temos $P_0 \triangle Q_0 \subset \bigcup_{i=1}^{k} P_i \triangle Q_i$. Assim, (13) implica

$$\mu(P_0 \triangle Q_0) < \delta. \quad (14)$$

Pela Proposição 3.8 e pelas relações (13) e (14) concluímos que, se $\delta > 0$ é tomado suficientemente pequeno, então $H(\mathcal{Q}_k|\mathcal{P}) < \varepsilon/2$ e $H(\mathcal{P}|\mathcal{Q}_k) < \varepsilon/2$.

Finalmente, do item (iv) da Proposição 3.7 temos:

$$H(\mathcal{Q}|\mathcal{P}) \leqslant H(\mathcal{Q}|\mathcal{Q}_k) + H(\mathcal{Q}_k|\mathcal{P}) < \varepsilon$$

e

$$H(\mathcal{P}|\mathcal{Q}) \leqslant H(\mathcal{P}|\mathcal{Q}_k) + H(\mathcal{Q}_k|\mathcal{Q}) < \varepsilon,$$

o que conclui a demonstração. □

Lema 3.23 . Sejam $\mathcal{P}$ uma partição de X com entropia finita e $(\mathcal{A}_n)_{n=1}^{\infty}$ uma sequência de sub-σ-álgebras tais que $\mathcal{A}_n \uparrow \mathcal{X}$. Então, para todo $\varepsilon > 0$ existem $n \geqslant 1$ e uma partição com entropia finita $\mathcal{Q} \subset \mathcal{A}_n$ tais que $H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}|\mathcal{P}) < \varepsilon$.

Demonstração. Seja $\varepsilon > 0$ arbitrário. Utilizando a notação do Lema 3.16 temos $H(\mathcal{P}|\mathcal{P}_k) < \varepsilon/4$ e $H(\mathcal{P}_k|\mathcal{P}) = 0$ para algum $k \geqslant 1$ suficientemente grande. Como $\mathcal{A}_n \subset \mathcal{A}_{n+1}$ para todo $n \geqslant 1$, temos que $\bigcup_{n=1}^{\infty} \mathcal{A}_n$ é uma álgebra, a qual gera $\mathcal{X}$ pois $\mathcal{A}_n \uparrow \mathcal{X}$ por hipótese. Logo, pela Proposição 3.22 existe $\mathcal{Q}$ partição finita de X tal que $\mathcal{Q} \subset \bigcup_{n=1}^{\infty} \mathcal{A}_n$ e $H(\mathcal{Q}|\mathcal{P}_k) + H(\mathcal{P}_k|\mathcal{Q}) < \varepsilon/4$.

Como $\mathcal{Q}$ é finita, existe n_0 tal que $\mathcal{Q} \subset \mathcal{A}_{n_0}$.

Pelo item (iv) da Proposição 3.7 tem-se:

$$H(\mathcal{P}|\mathcal{Q}) + H(\mathcal{Q}|\mathcal{P}) \leqslant H(\mathcal{P}|\mathcal{P}_k) + H(\mathcal{P}_k|\mathcal{Q}) + H(\mathcal{Q}|\mathcal{P}_k) + H(\mathcal{P}_k|\mathcal{P}) < \varepsilon.$$

□

Nesse ponto será conveniente introduzir as noções de conteúdo de informação condicional e entropia condicional no caso em que o dado é uma *sub-σ-álgebra*, generalizando o caso em que o dado é uma partição. Mais precisamente, temos:

Definição 3.24 . Sejam $\mathcal{P}$ uma partição e $\mathcal{F}$ uma sub-σ-álgebra. O *conteúdo de informação condicional* de $\mathcal{P}$ dada $\mathcal{F}$ é a função definida por:

$$I_{\mathcal{P}|\mathcal{F}}(x) = \sum_{P\in\mathcal{P}} -1_P(x)\ln(\mu(P|\mathcal{F})(x)). \tag{15}$$

Definição 3.25 . Sejam $\mathcal{P}$ uma partição e $\mathcal{F}$ uma sub-σ-álgebra. A *entropia condicional* de $\mathcal{P}$ dada $\mathcal{F}$ é definida por:

$$H(\mathcal{P}|\mathcal{F}) = \int_X I_{\mathcal{P}|\mathcal{F}} d\mu. \tag{16}$$

Observe que:

$$\begin{aligned}
H(\mathcal{P}|\mathcal{F}) &= \int_X I_{\mathcal{P}|\mathcal{F}} d\mu \\
&= \int_X \sum_{P\in\mathcal{P}} -1_P(x)\ln(\mu(P|\mathcal{F})(x)) d\mu \\
&= \int_X \mathbb{E}\Big(\sum_{P\in\mathcal{P}} -1_P(\cdot)\ln(\mu(P|\mathcal{F})(\cdot))|\mathcal{F}\Big) d\mu \\
&= \int_X \sum_{P\in\mathcal{P}} -\ln(\mu(P|\mathcal{F}))\mathbb{E}(1_P|\mathcal{F}) d\mu \\
&= \sum_{P\in\mathcal{P}} \int_X -\mu(P|\mathcal{F})\ln(\mu(P|\mathcal{F})) d\mu \\
&= \int_X \phi(\mu(P|\mathcal{F})) d\mu.
\end{aligned}$$

Proposição 3.26 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade e $\mathcal{P}$ uma partição. Tem-se:

(i) $I_{\mathcal{P}|\{\varnothing,X\}} = I_{\mathcal{P}}$ μ-q.t.p e $H(\mathcal{P}|\{\varnothing,X\}) = H(\mathcal{P})$.

(ii) $I_{\mathcal{P}|\mathcal{X}} = 0$ μ-q.t.p. e $H(\mathcal{P}|\mathcal{X}) = 0$.

Demonstração. Pelo Teorema 1.47, para todo $P \in \mathcal{P}$ tem-se:

$$\mu(P|\{\varnothing,X\}) = E(1_P|\{\varnothing,X\}) = \frac{\mu(P\cap X)}{\mu(X)} = \mu(P) \text{ } \mu\text{-q.t.p.}$$

Daí,

$$\begin{aligned}
I_{\mathcal{P}|\{\varnothing,X\}}(x) &= \sum_{P\in\mathcal{P}} -1_P(x)\ln(\mu(P|\{\varnothing,X\}(x)) \\
&= \sum_{P\in\mathcal{P}} -1_P(x)\ln(\mu(P)) \\
&= I_{\mathcal{P}}(x) \text{ } \mu\text{-q.t.p.}
\end{aligned}$$

Integrando em ambos os lados, concluímos a demonstração de (i). Por fim, o item (ii) decorre do fato de que, para cada $P \in \mathcal{P}$, tem-se $\mu(P|\mathcal{X})(x) = 1$ para μ-q.t.p. $x \in X$. Isso conclui a demonstração.

□

O próximo resultado garante que as definições acima generalizam o caso em que o dado é uma partição.

Lema 3.27 . Sejam $\mathcal{P}$ e $\mathcal{Q}$ duas partições de X. Então,

$$I_{\mathcal{P}|\widehat{\mathcal{Q}}} = I_{\mathcal{P}|\mathcal{Q}} \ \mu\text{-q.t.p. e } H(\mathcal{P}|\widehat{\mathcal{Q}}) = H(\mathcal{P}|\mathcal{Q}). \tag{17}$$

Demonstração. Sejam $P \in \mathcal{P}$ e $Q \in \mathcal{Q}$ arbitrários. Pelo Teorema 1.47 (com $\mathcal{F} = \{\varnothing, X\}$), para μ-q.t.p. $x \in P \cap Q$ temos:

$$\begin{aligned} I_{\mathcal{P}|\widehat{\mathcal{Q}}}(x) &= -\ln(\mu(P|\widehat{\mathcal{Q}})(x)) \\ &= -\ln\Big(\sum_{Q' \in \mathcal{Q}} 1_{Q'}(x)\frac{\mu(P \cap Q')}{\mu(Q')}\Big) \\ &= -\ln\Big(\frac{\mu(P \cap Q)}{\mu(Q)}\Big). \end{aligned}$$

Portanto, $I_{\mathcal{P}|\widehat{\mathcal{Q}}}(x) = \sum_{P \in \mathcal{P}} \sum_{Q \in \mathcal{Q}} -1_{P \cap Q}(x) \ln\Big(\frac{\mu(P \cap Q)}{\mu(Q)}\Big) = I_{\mathcal{P}|\mathcal{Q}}(x)$ μ-q.t.p. Por fim, integrando em ambos os lados conclui-se a demonstração.

□

Proposição 3.28 . Sejam $\mathcal{F}$ e $\mathcal{G}$ sub-σ-álgebras e $\mathcal{P}$ e $\mathcal{Q}$ partições tais que $H(\mathcal{P}) < +\infty$ e $H(\mathcal{Q}) < +\infty$. Valem as seguintes afirmações:

(i) $I_{\mathcal{P} \vee \mathcal{Q}|\mathcal{F}} = I_{\mathcal{P}|\mathcal{F}} + I_{\mathcal{Q}|\widehat{\mathcal{P}} \vee \mathcal{F}}$ para μ-q.t.p. e $H(\mathcal{P} \vee \mathcal{Q}|\mathcal{F}) = H(\mathcal{P}|\mathcal{F}) + H(\mathcal{Q}|\widehat{\mathcal{P}} \vee \mathcal{F})$;

(ii) Se $\mathcal{F} \subset \mathcal{G}$, então $H(\mathcal{P}|\mathcal{G}) \leqslant H(\mathcal{P}|\mathcal{F})$;

(iii) $H(T^{-1}\mathcal{P}|T^{-1}\mathcal{F}) = H(\mathcal{P}|\mathcal{F})$;

(iv) Se $\mathcal{P} \prec \mathcal{Q}$, então $H(\mathcal{P}|\mathcal{F}) \leqslant H(\mathcal{Q}|\mathcal{F})$;

(v) $H(\mathcal{P} \vee \mathcal{Q}|\mathcal{F}) \leqslant H(\mathcal{P}|\mathcal{F}) + H(\mathcal{Q}|\mathcal{F})$.

Demonstração. (i): Pela Proposição 1.47, para cada $Q \in \mathcal{Q}$ temos

$$\mu(Q|\widehat{P} \vee \mathcal{F}) = \sum_{P \in \mathcal{P}} 1_P \frac{\mu(Q \cap P|\mathcal{F})}{\mu(P|\mathcal{F})} \ \mu\text{-q.t.p.}$$

Agora note que, para qualquer $x \in X$, apenas um dos termos da soma acima é não-nulo. Daí,

$$\ln \mu(Q|\widehat{P} \vee \mathcal{F})(x) = \sum_{P \in \mathcal{P}} 1_P(x) \ln \frac{\mu(Q \cap P|\mathcal{F})(x)}{\mu(P|\mathcal{F})(x)} \ \text{para } \mu\text{-q.t.p.}$$

Consequentemente, para μ-q.t.p. tem-se:

$$\begin{aligned} \sum_{P \in \mathcal{P}, Q \in \mathcal{Q}} -1_{P \cap Q} \ln \mu(P \cap Q|\mathcal{F}) &= \sum_{P \in \mathcal{P}, Q \in \mathcal{Q}} -1_P 1_Q \ln \mu(P|\mathcal{F}) \\ &\quad + \sum_{P \in \mathcal{P}, Q \in \mathcal{Q}} -1_P 1_Q \ln \frac{\mu(P \cap Q|\mathcal{F})}{\mu(P|\mathcal{F})} \\ &= \sum_{P \in \mathcal{P}} -1_P \ln \mu(P|\mathcal{F}) + \sum_{Q \in \mathcal{Q}} -1_Q \ln \mu(Q|\widehat{P} \vee \mathcal{F}) \end{aligned}$$

Ou seja, $I_{\mathcal{P}\vee\mathcal{Q}|\mathcal{F}} = I_{\mathcal{P}|\mathcal{F}} + I_{\mathcal{Q}|\widehat{\mathcal{P}}\vee\mathcal{F}}$ para μ-q.t.p. Integrando em ambos os lados, obtemos também $H(\mathcal{P}\vee\mathcal{Q}|\mathcal{F}) = H(\mathcal{P}|\mathcal{F}) + H(\mathcal{Q}|\widehat{\mathcal{P}}\vee\mathcal{F})$.

(ii): Considere novamente a função $\phi : [0,1] \to \mathbb{R}$ definida por $\phi(x) = -t\log t$.

Para cada $P \in \mathcal{P}$ considere $f_P(x) = \mathbb{E}(1_P|\mathcal{G})(x)$.

Afirmação: $\mathbb{E}(\mathbb{E}(1_P|\mathcal{G})|\mathcal{F}) = \mathbb{E}(1_P|\mathcal{F})$ para todo $P \in \mathcal{P}$.

Com efeito, pela definição do valor esperado condicional temos que $\mathbb{E}(\mathbb{E}(1_P|\mathcal{G})|\mathcal{F})$ é $\mathcal{F}$-mensurável.

Agora, observe que, dado $F \in \mathcal{F} \subset \mathcal{G}$ arbitrário, temos:

$$\int_F \mathbb{E}(\mathbb{E}(1_P|\mathcal{G})|\mathcal{F})d\mu = \int_F \mathbb{E}(1_P|\mathcal{G})d\mu = \int_F 1_P d\mu.$$

Portanto, a afirmação está demonstrada.

Agora, pelo Teorema 1.48 (Desigualdade de Jensen) e da afirmação acima temos:

$$\begin{aligned}\mathbb{E}(\phi\circ\mathbb{E}(1_P|\mathcal{G})|\mathcal{F}) &\leqslant \phi(\mathbb{E}(\mathbb{E}(1_P|\mathcal{G})|\mathcal{F}))\\ &= \phi(\mathbb{E}(1_P|\mathcal{F}))\\ &= -\mathbb{E}(1_P|\mathcal{F})\ln(\mathbb{E}(1_P|\mathcal{F})).\end{aligned}$$

Daí, para todo $P \in \mathcal{P}$ temos:

$$\mathbb{E}(-\mathbb{E}(1_P|\mathcal{G})\ln(\mathbb{E}(1_P|\mathcal{G})|\mathcal{F})) \leqslant -\mathbb{E}(1_P|\mathcal{F})\ln(\mathbb{E}(1_P|\mathcal{F})).$$

Integrando, obtém-se:

$$\begin{aligned}\int_X (-\mathbb{E}(1_P|\mathcal{G})\ln(\mathbb{E}(1_P|\mathcal{G}))d\mu &= \int_X \mathbb{E}(-\mathbb{E}(1_P|\mathcal{G})\ln(\mathbb{E}(1_P|\mathcal{G})|\mathcal{F}))d\mu\\ &\leqslant \int_X -\mathbb{E}(1_P|\mathcal{F})\ln(\mathbb{E}(1_P|\mathcal{F}))d\mu.\end{aligned}$$

Por fim, somando sobre todos os $P \in \mathcal{P}$, obtemos $H(\mathcal{P}|\mathcal{G}) \leqslant H(\mathcal{P}|\mathcal{F})$.

(iii): **Afirmação:** $\mathbb{E}(1_P|\mathcal{F})\circ T = \mathbb{E}(1_P\circ T|T^{-1}\mathcal{F})$ para todo $P \in \mathcal{P}$.

De fato, sejam $B \in T^{-1}\mathcal{F}$ e $P \in \mathcal{P}$ arbitrários. Então, $B = T^{-1}A$ para algum $A \in \mathcal{F}$.

Como μ é T-invariante, temos:

$$\begin{aligned}\int_B \mathbb{E}(1_P|\mathcal{F})\circ T d\mu &= \int_X (\mathbb{E}(1_P|\mathcal{F})\circ T)1_{T^{-1}A}d\mu\\ &= \int_X (\mathbb{E}(1_P|\mathcal{F})1_A)\circ T d\mu\\ &= \int_X \mathbb{E}(1_P|\mathcal{F})1_A d\mu\\ &= \int_A 1_P d\mu\\ &= \int_X (1_P 1_A)\circ T d\mu\\ &= \int_X (1_P\circ T)(1_A\circ T)d\mu\\ &= \int_{T^{-1}A} 1_P\circ T d\mu.\end{aligned}$$

Como $\mathbb{E}(1_P|\mathcal{F})\circ T$ é $\mathcal{F}$-mensurável a afirmação segue.

Por fim, note que:

$$\begin{aligned}
H(T^{-1}\mathcal{P}|T^{-1}\mathcal{F}) &= \int_X \sum_{P\in\mathcal{P}} -\ln(\mu(T^{-1}P|T^{-1}\mathcal{F}))1_{T^{-1}P}d\mu\\
&= \int_X \sum_{P\in\mathcal{P}} -\ln(\mu(P|\mathcal{F})\circ T)1_P\circ T d\mu\\
&= \int_X \sum_{P\in\mathcal{P}} -\ln(\mu(P|\mathcal{F}))1_P d\mu\\
&= H(\mathcal{P}|\mathcal{F}).
\end{aligned}$$

(iv): Se $\mathcal{P}\prec\mathcal{Q}$, então pelo item (i) temos:

$$\begin{aligned}
H(\mathcal{P}|\mathcal{F}) &\leqslant H(\mathcal{P}|\mathcal{F}) + H(\mathcal{Q}|\widehat{\mathcal{P}}\vee\mathcal{F})\\
&= H(\mathcal{P}\vee\mathcal{Q}|\mathcal{F})\\
&= H(\mathcal{Q}|\mathcal{F}).
\end{aligned}$$

(v): Utilizando os itens (i) e (ii) tem-se:

$$\begin{aligned}
H(\mathcal{P}\vee\mathcal{Q}|\mathcal{F}) &= H(\mathcal{P}|\mathcal{F}) + H(\mathcal{Q}|\widehat{\mathcal{P}}\vee\mathcal{F})\\
&\leqslant H(\mathcal{P}|\mathcal{F}) + H(\mathcal{Q}|\mathcal{F}).
\end{aligned}$$

□

Teorema 3.29 . Sejam $\mathcal{P}$ uma partição e $\mathcal{A}$ uma sub-σ-álgebra. Então, $H(\mathcal{P}|\mathcal{A})=0$ se, e somente se, $\widehat{\mathcal{P}}\subset\mathcal{A}$.

Demonstração. Suponha $\widehat{\mathcal{P}}\subset\mathcal{A}$. Então, para cada $P\in\mathcal{P}$ a função 1_P é $\mathcal{A}$-mensurável, donde $\mathbb{E}(1_P|\mathcal{A})=1_P$ μ-q.t.p. Daí,

$$I_{\mathcal{P}|\mathcal{A}}(x)=\sum_{P\in\mathcal{P}} -1_P(x)\ln(\mu(P|\mathcal{A}))=\sum_{P\in\mathcal{P}} -1_P(x)\ln(1_P(x))=0 \ \mu\text{-q.t.p. } x\in X.$$

Logo, $H(\mathcal{P}|\mathcal{A})=0$.

Reciprocamente, suponha $H(\mathcal{P}|\mathcal{A})=0$.

Então,

$$\begin{aligned}
0=H(\mathcal{P}|\mathcal{A}) &= \int_X \mathbb{E}(I_{\mathcal{P}|\mathcal{A}}|\mathcal{A})d\mu\\
&= \int_X \mathbb{E}\Big(\sum_{P\in\mathcal{P}} -1_P\ln(\mu(\mathcal{P}|\mathcal{A}))|\mathcal{A}\Big)d\mu\\
&= \int_X \sum_{P\in\mathcal{P}} \mathbb{E}\Big(-1_P\ln(\mu(\mathcal{P}|\mathcal{A}))|\mathcal{A}\Big)d\mu\\
&= \int_X \sum_{P\in\mathcal{P}} -\ln(\mu(\mathcal{P}|\mathcal{A}))\mathbb{E}(1_P|\mathcal{A})d\mu\\
&= \int_X \sum_{P\in\mathcal{P}} -\mu(P|\mathcal{A})\ln(\mu(P|\mathcal{A}))d\mu\\
&= \sum_{P\in\mathcal{P}}\int_X \phi(\mu(P|\mathcal{A}))d\mu.
\end{aligned} \tag{18}$$

Consequentemente, obtemos $\phi(\mu(P|\mathcal{A})) = 0$ para todo $P \in \mathcal{P}$, ou seja, $\mu(P|\mathcal{A})(x) = 0$ ou 1 para μ-q.t.p. $x \in X$.

Agora, fixe $P \in \mathcal{P}$ arbitrário e sejam $B_1 = \mu(P|\mathcal{A})^{-1}(\{0\})$ e $B_2 = \mu(P|\mathcal{A})^{-1}(\{1\})$. Então, B_1 e B_2 são $\mathcal{A}$-mensuráveis e $\mu(B_1 \cup B_2) = 1$. Note que:

$$\int_{B_2} \mu(P|\mathcal{A}) d\mu = \int_{B_2} \mathbb{E}(1_P|\mathcal{A}) d\mu = \int_{B_2} 1_P d\mu = \mu(B_2 \cap P).$$

Por outro lado,

$$\int_{B_2} \mu(P|\mathcal{A}) d\mu = \int_{B_2} 1 d\mu = \mu(B_2).$$

Portanto, $\mu(B_2 \cap P) = \mu(B_2)$. Além disso,

$$\mu(B_2) = \int_X 1_{B_2} d\mu = \int_X \mu(P|\mathcal{A}) d\mu = \int_X \mathbb{E}(1_P|\mathcal{A}) d\mu = \int_X 1_P d\mu = \mu(P).$$

Logo, $\mu(B_2) = \mu(B_2 \cap P) = \mu(P)$. Portanto, segue que $\mu(B_2 \cap P)/\mu(P) = 1$ e $\mu(B_2 \cap P)/\mu(B_2) = 1$. Consequentemente, $P = B_2 \in \mathcal{A}$ exceto possivelmente por um conjunto de medida nula. Como $P \in \mathcal{P}$ é arbitrário, obtemos $\widehat{\mathcal{P}} \subset \mathcal{A}$, o que conclui a demonstração.

□

Lema 3.30 . (Lema de Chung) Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade e $\mathcal{P}$ uma partição com $H(\mathcal{P}) < +\infty$. Valem as seguintes afirmações:

(i) Se $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \ldots$ é uma sequência crescente de sub-σ-álgebras tal que $\mathcal{F}_n \uparrow \mathcal{F}$, ou seja, $\mathcal{F} = \bigvee_{n \geqslant 1} \mathcal{F}_n$, então a função $f(x) := \sup_{n \geqslant 1} I_{\mathcal{P}|\mathcal{F}_n}(x)$ é integrável e

$$\int_X f d\mu \leqslant H(\mathcal{P}) + 1.$$

(ii) Se $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \ldots$ é uma sequência decrescente de sub-σ-álgebras tal que $\mathcal{F}_n \downarrow \mathcal{F}$, ou seja, $\mathcal{F} = \bigcap_{n \geqslant 1} \mathcal{F}_n$, então a função $f(x) := \sup_{n \geqslant 1} I_{\mathcal{P}|\mathcal{F}_n}(x)$ é integrável e

$$\int_X f d\mu \leqslant H(\mathcal{P}) + 1.$$

Demonstração. (i): Fixe $P \in \mathcal{P}$ e $t \geqslant 0$ arbitrários e considere o conjunto

$$C := \{x \in X; f(x) > t\}.$$

Como para cada $x \in P$ tem-se $I_{\mathcal{P}|\mathcal{F}_n}(x) = -\ln \mu(P|\mathcal{F}_n)(x)$, vemos que

$$f(x) = \sup_{n \geqslant 1} I_{\mathcal{P}|\mathcal{F}_n}(x) > t$$

equivale a dizer $\mu(P|\mathcal{F}_n)(x) < e^{-t}$ para algum $n \geqslant 1$. Agora, para cada $n \geqslant 1$ considere o conjunto

$$C_n := \{x \in X; \mu(P|\mathcal{F}_n)(x) < e^{-t} \text{ e } \mu(P|\mathcal{F}_k)(x) \geqslant e^{-t} \text{ para cada } k < n\}.$$

Observe que $C_n \in \mathcal{F}_n$, $C_m \cap C_n = \varnothing$ sempre que $m \neq n$ e $C \cap P \subset \bigcup_{n \geqslant 1} C_n$. Além disso,

$$\begin{aligned}\mu(P \cap C_n) &= \int_{C_n} 1_P d\mu \\ &= \int_{C_n} \mathbb{E}(1_P | \mathcal{F}_n) d\mu \\ &\leqslant e^{-t}\mu(C_n).\end{aligned}$$

Daí, como os C_n's são dois a dois disjuntos, somando sobre todos os $n \geqslant 1$ obtém-se:

$$\begin{aligned}\mu(\{x \in P; f(x) > t\}) &= \mu(P \cap C) \\ &= \sum_{n=1}^{\infty} \mu(P \cap C_n) \\ &\leqslant e^{-t} \sum_{n=1}^{\infty} \mu(C_n) \\ &= e^{-t}\mu\Big(\cup_{n=1}^{\infty} C_n\Big) \\ &\leqslant e^{-t}.\end{aligned}$$

Consequentemente, denotando $F(t) := \mu(\{x \in X; f(x) > t\})$ vemos que

$$F(t) \leqslant \sum_{P \in \mathcal{P}} \min\{\mu(P), e^{-t}\}$$

e, daí, obtemos:

$$\begin{aligned}\int_X f d\mu &= -\int_0^{\infty} t dF(t) \\ &= -tF(t)|_0^{\infty} + \int_0^{\infty} F(t) dt \\ &\leqslant \int_0^{\infty} F(t) dt \\ &= \int_0^{\infty} \mu(\{x \in X; f(x) > t\}) dt \\ &\leqslant \int_0^{\infty} \sum_{P \in \mathcal{P}} \min\{\mu(P), e^{-t}\} dt \\ &= \sum_{P \in \mathcal{P}} \Big[\int_0^{-\ln \mu(P)} \mu(P) dt + \int_{-\ln \mu(P)}^{\infty} e^{-t} dt \Big] \\ &= H(\mathcal{P}) + \sum_{P \in \mathcal{P}} \mu(P) \\ &= H(\mathcal{P}) + 1,\end{aligned}$$

o que conclui a demonstração do item (i).

(ii): Para cada $n \geqslant 1$ a sequência finita decrescente $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \ldots \supset \mathcal{F}_n$ pode ser vista como uma sequência crescente $\mathcal{F}_n \subset \ldots \subset \mathcal{F}_2 \subset \mathcal{F}_1$. Logo, pelo item (i) segue que a função $f_n(x) := \sup\limits_{1\leqslant j\leqslant n} I_{\mathcal{P}|\mathcal{F}_j}(x)$ é integrável e satisfaz

$$\int_X f_n d\mu \leqslant H(\mathcal{P}) + 1.$$

Além disso, $(f_n)_{n\geqslant 1}$ é uma sequência crescente de funções mensuráveis não-negativas tal que $f_n \uparrow f = \sup\limits_{n\geqslant 1} I_{\mathcal{P}|\mathcal{F}_n}$. Portanto, pelo Teorema 1.38 (Teorema da Convergência Monótona) obtemos:

$$\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu \leqslant H(\mathcal{P}) + 1.$$

□

O próximo resultado será fundamental no último capítulo deste trabalho.

Teorema 3.31 . Sejam $(\mathcal{F}_n)_{n=1}^{\infty}$ uma sequência de sub-σ-álgebras e $\mathcal{P}$ uma partição com $H(\mathcal{P}) < +\infty$. Valem as seguintes afirmações:

(i) Se $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ para todo $n \geqslant 1$, então $H(\mathcal{P}|\mathcal{F}_n) \downarrow H\Big(\mathcal{P}|\bigvee\limits_{n=1}^{\infty} \mathcal{F}_n\Big)$.

(ii) Se $\mathcal{F}_n \supset \mathcal{F}_{n+1}$ para todo $n \geqslant 1$, então $H(\mathcal{P}|\mathcal{F}_n) \uparrow H\Big(\mathcal{P}|\bigcap\limits_{n\geqslant 1} \mathcal{F}_n\Big)$.

Demonstração. (i) e (ii): Fixe $P \in \mathcal{P}$ arbitrário. Para μ-q.t.p. $x \in P$, temos $I_{\mathcal{P}|\mathcal{F}_n}(x) = -\ln \mathbb{E}(1_P|\mathcal{F}_n)(x)$ e $I_{\mathcal{P}|\mathcal{F}}(x) = -\ln \mathbb{E}(1_P|\mathcal{F})(x)$, donde pela continuidade de $\ln$ o Teorema 1.49 implica que $I_{\mathcal{P}|\mathcal{F}_n}$ converge para $I_{\mathcal{P}|\mathcal{F}}$ em μ-q.t.p. Além disso, pelo Lema 3.30 a sequência de funções $(I_{\mathcal{P}|\mathcal{F}_n})_{n\geqslant 1}$ é dominada pela função integrável $\sup\limits_{n\geqslant 1} \mathbb{E}(1_P|\mathcal{F}_n)$. Portanto, a sequência $(I_{\mathcal{P}|\mathcal{F}_n})_{n\geqslant 1}$ é uniformemente integrável, donde a convergência μ-q.t.p. implica convergência em $L^1(X,\mathcal{X},\mu)$. Consequentemente, pelo item (ii) da Proposição 3.28 obtemos $H(\mathcal{P}|\mathcal{F}_n) \downarrow H(\mathcal{P}|\mathcal{F})$ se $\mathcal{F}_n \subset \mathcal{F}_{n+1}$ para todo $n \geqslant 1$ e $H(\mathcal{P}|\mathcal{F}_n) \uparrow H(\mathcal{P}|\mathcal{F})$ se $\mathcal{F}_n \supset \mathcal{F}_{n+1}$ para todo $n \geqslant 1$. Isto conclui a demonstração.

□

Lema 3.32 . Para toda $\mathcal{P}$ partição com $H(\mathcal{P}) < +\infty$ tem-se

$$H(T,\mathcal{P}) = \lim_{n\to\infty} H\Big(\mathcal{P}|\bigvee_{i=1}^{n} T^{-i}\mathcal{P}\Big) = H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}\Big).$$

Demonstração. Primeiramente observe que, como $\bigvee\limits_{i=1}^{n} T^{-i}\mathcal{P} = \bigvee\limits_{i=1}^{n} \widehat{T^{-i}\mathcal{P}} \uparrow \bigvee\limits_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}$, a segunda igualdade decorre do Teorema 3.31. Agora, para todo $n \geqslant 1$ temos:

$$\begin{aligned} H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) &= H\Big(\bigvee_{i=1}^{n-1} T^{-i}\mathcal{P} \vee \mathcal{P}\Big) \\ &= H\Big(\bigvee_{i=1}^{n-1} T^{-i}\mathcal{P}\Big) + H\Big(\mathcal{P}|\bigvee_{i=1}^{n-1} T^{-i}\mathcal{P}\Big) \qquad (19) \\ &= H\Big(\bigvee_{i=0}^{n-2} T^{-i}\mathcal{P}\Big) + H\Big(\mathcal{P}|\bigvee_{i=1}^{n-1} T^{-i}\mathcal{P}\Big). \end{aligned}$$

Utilizando recursivamente a equação (19) segue que

$$H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) = H(\mathcal{P}) + \sum_{k=1}^{n-1} H\Big(\mathcal{P}|\bigvee_{j=1}^{k} T^{-j}\mathcal{P}\Big).$$

Portanto,

$$\begin{aligned} H(T,\mathcal{P}) &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) \\ &= \lim_{n\to\infty} \frac{1}{n} H(\mathcal{P}) + \frac{1}{n}\sum_{k=1}^{n-1} H\Big(\mathcal{P}|\bigvee_{j=1}^{k} T^{-j}\mathcal{P}\Big) \\ &= \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n-1} H\Big(\mathcal{P}|\bigvee_{j=1}^{k} T^{-j}\mathcal{P}\Big) \\ &= \lim_{n\to\infty} H\Big(\mathcal{P}|\bigvee_{j=1}^{n} T^{-j}\mathcal{P}\Big), \end{aligned} \tag{20}$$

onde na última igualdade usamos a Proposição 1.50.

□

Observação 3.33 . A sub-σ-álgebra $\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}$ é usualmente denominada *o passado completo* do processo $(T,\mathcal{P})$. Tendo em vista o Lema 3.32, vemos que $H(T,\mathcal{P})$ é precisamente *a quantidade média de informação obtida ao realizarmos o experimento $\mathcal{P}$ dado o passado completo do processo* $(T,\mathcal{P})$. Além disso, observe que decorre do Teorema 3.29 e do Lema 3.32 que

$$H(T,\mathcal{P}) = 0 \iff \widehat{\mathcal{P}} \subset \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}.$$

Ou seja, a entropia do processo $(T,\mathcal{P})$ é nula se, e somente se, "o experimento $\mathcal{P}$ está contido no passado completo do processo $(T,\mathcal{P})$“. Em outras palavras, $H(T,\mathcal{P}) = 0$ se, e somente se, “o passado determina o presente”. Neste caso, o processo $(T,\mathcal{P})$ é dito *estocasticamente determinístico*, ou simplesmente *determinístico*.

Agora, observe que $\bigvee_{i=k}^{\infty} \widehat{T^{-i}\mathcal{P}} \supset \bigvee_{i=k+1}^{\infty} \widehat{T^{-i}\mathcal{P}}$ para todo $k \geqslant 1$. A sub-σ-álgebra $\bigcap_{k=1}^{\infty}\bigvee_{i=k}^{\infty} \widehat{T^{-i}\mathcal{P}}$ é dita *o passado remoto* do processo $(T,\mathcal{P})$.

Por fim, a sub-σ-álgebra $\bigvee_{i=-\infty}^{+\infty} T^{-i}\mathcal{P}$ é dita *a história completa* do processo $(T,\mathcal{P})$.

Definição 3.34 . Sejam $(X,\mathcal{X},\mu,T)$ um sistema dinâmico e $\mathcal{P}$ uma partição com $H(\mathcal{P}) < +\infty$. Dizemos que $\mathcal{P}$ é um *gerador* de $(X,\mathcal{X},\mu,T)$ se $\bigvee_{i=-\infty}^{+\infty} T^{-i}\mathcal{P} = \mathcal{X} \pmod{\mu}$. Em outras palavras, todo evento $A \in \mathcal{X}$ é determinado pela história completa do processo $(T,\mathcal{P})$. Além disso, dizemos que $\mathcal{P}$ é um *gerador forte* de $(X,\mathcal{X},\mu,T)$ se $\bigvee_{i=1}^{+\infty} T^{-i}\mathcal{P} = \mathcal{X} \pmod{\mu}$. Em outras palavras, neste caso, todo evento $A \in \mathcal{X}$ é determinado pelo passado do processo $(T,\mathcal{P})$.

Teorema 3.35 . Sejam $(X,\mathcal{X},\mu,T)$ um sistema dinâmico e $\mathcal{P}_1 \prec \cdots \prec \mathcal{P}_n \prec \ldots$ uma sequência não-decrescente de partições tal que $H(\mathcal{P}_n) < +\infty$ para todo $n \geqslant 1$ e $\bigvee_{n=1}^{\infty} \mathcal{P}_n = \mathcal{X}$. Vale:

$$H(T) = \lim_{n\to\infty} H(T,\mathcal{P}_n).$$

Demonstração. Com efeito, seja $\mathcal{Q}$ uma partição finita arbitrária. Pelo Lema 3.15, para cada $n \geqslant 1$ temos:

$$\begin{aligned} H(T,\mathcal{Q}) &\leqslant H(T,\mathcal{P}_n) + H(\mathcal{Q}|\mathcal{P}_n) \\ &\leqslant H(T) + H(\mathcal{Q}|\mathcal{P}_n). \end{aligned}$$

Agora, como $\widehat{\mathcal{P}_n} \uparrow \mathcal{X}$, segue do Teorema 3.29 e do Teorema 3.31:

$$\lim_{n\to\infty} H(\mathcal{Q}|\widehat{\mathcal{P}_n}) = H(\mathcal{Q}|\mathcal{X}) = 0,$$

donde vemos que

$$H(T,\mathcal{Q}) \leqslant \lim_{n\to\infty} H(T,\mathcal{P}_n) \leqslant H(T).$$

Como $\mathcal{Q}$ foi tomada arbitrariamente, obtemos $H(T) = \lim_{n\to\infty} H(T,\mathcal{P}_n)$, como queríamos demonstrar. □

Teorema 3.36 . (Kolmogorov-Sinai) Se $\mathcal{P}$ é um gerador do sistema dinâmico $(X,\mathcal{X},\mu,T)$, então

$$H(T) = H(T,\mathcal{P}).$$

Demonstração. Seja $\mathcal{Q}$ uma partição finita arbitrária. Pelo Lema 3.15 e pelo Lema 3.18, para cada $k \geqslant 0$ vemos que:

$$\begin{aligned} H(T,\mathcal{Q}) &\leqslant H(T, \bigvee_{i=-k}^{k} T^{-i}\mathcal{P}) + H(\mathcal{Q}| \bigvee_{i=-k}^{k} T^{-i}\mathcal{P}) \\ &= H(T,\mathcal{P}) + H(\mathcal{Q}| \bigvee_{i=-k}^{k} T^{-i}\mathcal{P}). \end{aligned}$$

Agora, como $\bigvee_{i=-k}^{k} T^{-i}\mathcal{P} \uparrow \bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{P}$ e $\bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{P} = \mathcal{X}$ por hipótese, fazendo $k \to \infty$ segue do Teorema 3.29 e do Teorema 3.31:

$$H(T,\mathcal{Q}) \leqslant H(T,\mathcal{P}) \leqslant H(T).$$

Como $\mathcal{Q}$ foi tomada arbitrariamente, segue que $H(T) = H(T,\mathcal{P})$, como queríamos demonstrar. □

Corolário 3.37 . Seja $(X,\mathcal{X},\mu,T)$ um sistema dinâmico e suponha que $\mathcal{P}$ seja um gerador forte. Então, $H(T) = 0$.

Demonstração. Por hipótese, $\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} = \mathcal{X}$. Daí, pelos Teorema 3.36, Teorema 3.29 e Lema 3.32, tem-se:

$$H(T) = H(T,\mathcal{P}) = H(\mathcal{P}|\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}) = H(\mathcal{P}|\mathcal{X}) = 0,$$

o que conclui a demonstração. □

Observação 3.38 . Suponha que $\mathcal{P}$ é um gerador para o sistema dinâmico $(X,\mathcal{X},\mu,T)$. Pelo Lema 3.18, para cada $k \geqslant 0$ temos $H(T, \bigvee_{i=-k}^{k} T^{-i}\mathcal{P}) = H(T,\mathcal{P}) \leqslant H(\mathcal{P}) < +\infty$. Daí, como $\bigvee_{i=-k}^{k} T^{-i}\mathcal{P} \uparrow \bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{P}$ vemos que a conclusão do Teorema de Kolmogorov-Sinai decorre do Teorema 3.35 (observe também a semelhança entre as respectivas demonstrações).

Definição 3.39 . Dizemos que os sistemas dinâmicos $(X,\mathcal{X},\mu,T)$ e $(Y,\mathcal{Y},\nu,S)$ são *isomorfos* se existem $X_0 \in \mathcal{X}$, $Y_0 \in \mathcal{Y}$ e uma transformação bimensurável $\kappa : X_0 \to Y_0$ satisfazendo:

(i) $\mu(X_0) = \nu(Y_0) = 1$;

(ii) $T(X_0) = X_0$ e $S(Y_0) = Y_0$;

(iii) $\mu(A) = \nu(\kappa(A))$ para todo conjunto mensurável $A \subset X_0$ e $\nu(B) = \mu(\kappa^{-1}(B))$ para todo conjunto mensurável $B \subset Y_0$;

(iv) $\kappa \circ T(x) = S \circ \kappa(x)$ para todo $x \in X_0$.

Observação 3.40 . É usual estender o isomorfismo $\kappa : X_0 \to Y_0$ arbitrariamente aos conjuntos de medida nula $X \setminus X_0$ e $Y \setminus Y_0$ e considerar $\kappa : X \to Y$.

Teorema 3.41 . Se $(X,\mathcal{X},\mu,T)$ e $(Y,\mathcal{Y},\nu,S)$ são isomorfos, então $H(T) = H(S)$.

Em outras palavras, a entropia é um *invariante* de sistemas dinâmicos.

Demonstração. Seja $\kappa : X \to Y$ um isomorfismo. Dada qualquer $\mathcal{P}$ partição finita de X, segue que $\kappa\mathcal{P}$ é uma partição finita de Y e, além disso,

$$\begin{aligned}
H(T,\mathcal{P}) &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} (\kappa^{-1} \circ S \circ \kappa)^{-i}(\mathcal{P})\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} (\kappa^{-1} \circ S^{-i} \circ \kappa)(\mathcal{P})\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\kappa^{-1} \bigvee_{i=0}^{n-1} S^{-i}(\kappa\mathcal{P})\Big)
\end{aligned}$$

$$= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n-1} S^{-i}(\kappa\mathcal{P})\Big)$$
$$= H(S,\kappa\mathcal{P}).$$

□

Exemplo 3.42 . (Rotação irracional) Considere o sistema dinâmico $(S^1,\mathcal{B},\lambda,R_\alpha)$, onde $R_\alpha(x) = x+\alpha \pmod 1$ é a rotação sobre o círculo unitário fechado S^1 pelo ângulo $\alpha \in \mathbb{R}\setminus\mathbb{Q}$ e λ é a restrição da medida de Lebesgue à σ-álgebra de Borel $\mathcal{B}$ sobre S^1. Identifique $S^1 \sim [0,1]$ e considere a partição finita $\mathcal{P} = \{[0,1/2),[1/2,1]\}$. Agora, note que $\bigvee_{i=1}^{n} R_\alpha^{-i}\mathcal{P} \prec \bigvee_{i=1}^{n+1} R_\alpha^{-i}\mathcal{P}$ para todo $n \geqslant 1$ e, para λ-q.t.p. $x \in S^1$, tem-se $\operatorname{diam}\Big(\bigvee_{i=1}^{n} R_\alpha^{-i}\mathcal{P}\Big)(x) \to 0$ quando $n \to \infty$. Daí, pelo Lema 1.33 segue que $\bigvee_{i=1}^{\infty} R_\alpha^{-i}\mathcal{P} = \mathcal{B} \pmod\lambda$, donde pelo Corolário 3.37 obtém-se $H(R_\alpha) = 0$.

Exemplo 3.43 . (Deslocamento de Bernoulli) Seja $(X,\mathcal{X},\mu,T)$ um deslocamento de Bernoulli e considere $\mathcal{P} = \{P_1,\dots,P_a\}$ uma partição finita satisfazendo:

(i) $\bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{P} = \mathcal{X} \pmod\mu$;

(ii) $\mu\Big(\bigcap_{i=m}^{n} T^iP_{j_i}\Big) = \prod_{i=m}^{n}\mu(P_{j_i}),\ j_i \in \{1,\dots,a\},\ m \leqslant i \leqslant n \text{ e } m \leqslant n \text{ em } \mathbb{Z}.$

Pelo Teorema 3.36 (Kolmogorov-Sinai) segue que $H(T) = H(T,\mathcal{P})$. Agora observe que, como $T^{-i}\mathcal{P}$ e $T^{-j}\mathcal{P}$ são independentes sempre que $i \neq j$ (pela invariância de μ juntamente com a condição (ii) acima), pela relação (8) obtemos:

$$H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) = H(\mathcal{P}\vee T^{-1}\mathcal{P}\vee\ldots\vee T^{-(n-1)}\mathcal{P})$$
$$= \sum_{i=0}^{n-1} H(T^{-i}\mathcal{P})$$
$$= nH(\mathcal{P})$$
$$= n\sum_{j=1}^{a} -\mu(P_j)\ln\mu(P_j).$$

Consequentemente,

$$H(T) = H(T,\mathcal{P}) = \lim_{n\to\infty}\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^{-i}\mathcal{P}\Big) = -\sum_{j=1}^{a}\mu(P_j)\ln\mu(P_j).$$

Observação 3.44 . Pelo Exemplo 3.43 vemos que o deslocamento de Bernoulli $B(1/2,1/2)$ tem entropia $\ln 2$ e o deslocamento de Bernoulli $B(1/3,1/3,1/3)$ tem entropia $\ln 3$. Portanto, pelo Teorema 3.41 vemos que $B(1/2,1/2)$ e $B(1/3,1/3,1/3)$ *não* são isomorfos. Foi dessa maneira que Kolmogorov (KOLMOGOROV, 1958) respondeu a este problema básico da teoria, que persistiu mesmo depois dos resultados espectrais de von Neumann e Halmos (NEUMANN, 1932),(NEUMANN; HALMOS, P., 1942), (HALMOS, Paul, 1956). Por outro lado, em (ORNSTEIN, 1970) o matemático americano Donald Ornstein demonstrou que, se dois deslocamentos de Bernoulli $(X,\mathcal{X},\mu,T)$ e $(Y,\mathcal{Y},\nu,S)$ têm a mesma entropia, então eles são isomorfos. Consequentemente, dois deslocamentos de Bernoulli são isomorfos se, e somente se, têm a mesma entropia. Para estes e outros aspectos sobre o *problema do isomorfismo* em Teoria Ergódica recomendamos (WEISS, 1972).

O próximo resultado aborda a relação entre a noção de *independência relativa* e a entropia condicional, fornecendo uma generalização da relação (8).

Proposição 3.45 . Sejam $\mathcal{P}$,$\mathcal{Q}$ e $\mathcal{R}$ partições de X com entropia finita.

Então, tem-se

$$H(\mathcal{P}|\widehat{\mathcal{Q}\vee\mathcal{R}}) = H(\mathcal{P}|\widehat{\mathcal{Q}})$$

se, e somente se, $\mu(P\cap Q\cap R)\mu(Q) = \mu(P\cap Q)\mu(Q\cap R)$ para todos $P\in\mathcal{P}$, $Q\in\mathcal{Q}$ e $R\in\mathcal{R}$. Em outras palavras, tem-se

$$H(\mathcal{P}|\widehat{\mathcal{Q}\vee\mathcal{R}}) = H(\mathcal{P}|\widehat{\mathcal{Q}})$$

se, e somente se, $\mathcal{P}$ é independente de $\mathcal{R}$ sobre $\mathcal{Q}$.

Demonstração. Observe que tem-se $H(\mathcal{P}|\widehat{\mathcal{Q}\vee\mathcal{R}}) = H(\mathcal{P}|\widehat{\mathcal{Q}})$ se, e somente se,

$$\sum_{P\in\mathcal{P}}\sum_{Q\in\mathcal{Q}}\sum_{R\in\mathcal{R}} -\mu(P\cap Q\cap R)\ln\left(\frac{\mu(P\cap Q\cap R)}{\mu(Q\cap R)}\right)$$
$$= \sum_{P\in\mathcal{P}}\sum_{Q\in\mathcal{Q}}\sum_{R\in\mathcal{R}} -\mu(P\cap Q\cap R)\ln\left(\frac{\mu(P\cap Q)}{\mu(Q)}\right),$$

o que por sua vez ocorre se, e somente se, $\mu(P\cap Q\cap R)\mu(Q) = \mu(P\cap Q)\mu(Q\cap R)$ para todos $P\in\mathcal{P}$, $Q\in\mathcal{Q}$ e $R\in\mathcal{R}$.

□

Proposição 3.46 . Sejam $\mathcal{P}$, $\mathcal{P}'$, $\mathcal{Q}$ e $\mathcal{R}$ partições de X com entropia finita tais que $\mathcal{P}'\prec\mathcal{P}$ e $H(\mathcal{P}|\widehat{\mathcal{Q}\vee\mathcal{R}}) = H(\mathcal{P}|\widehat{\mathcal{Q}})$. Então, $H(\mathcal{P}'|\widehat{\mathcal{Q}\vee\mathcal{R}}) = H(\mathcal{P}'|\widehat{\mathcal{Q}})$

Demonstração. Sejam $P'\in\mathcal{P}'$, $Q\in\mathcal{Q}$ e $R\in\mathcal{R}$ arbitrários. Como $\mathcal{P}'\prec\mathcal{P}$, segue que P' é uma união de átomos de $\mathcal{P}$, digamos $P' = \bigcup_{j\in J} P_j$.

Daí, observe que

$$\begin{aligned}\mu(P'\cap Q\cap R)\mu(Q) &= \mu\Big(\bigcup_{j\in J} P_j\cap Q\cap R\Big)\mu(Q)\\ &= \sum_{j\in J}\mu(P_j\cap Q\cap R)\mu(Q)\\ &= \sum_{j\in J}\mu(P_j\cap Q)\mu(Q\cap R)\\ &= \mu(Q\cap R)\sum_{j\in J}\mu(P_j\cap Q)\\ &= \mu(Q\cap R)\mu(P'\cap Q),\end{aligned}$$

donde o resultado segue pela Proposição 3.45.

□

O próximo resultado generaliza a proposição anterior no caso em que o espaço de probabilidade é separável.

Proposição 3.47 . Sejam $(X,\mathcal{X},\mu)$ um espaço de probabilidade separável e $\mathcal{P}$ e $\mathcal{P}'$ partições de X com entropia finita tais que $\mathcal{P}' \prec \mathcal{P}$. Se $\mathcal{A}_1$ e $\mathcal{A}_2$ são sub-σ-álgebras tais que

$$H(\mathcal{P}|\mathcal{A}_1\vee\mathcal{A}_2) = H(\mathcal{P}|\mathcal{A}_1),$$

então

$$H(\mathcal{P}'|\mathcal{A}_1\vee\mathcal{A}_2) = H(\mathcal{P}'|\mathcal{A}_1).$$

Demonstração. Seja $\mathcal{Q}$ partição de X arbitrária tal que $\mathcal{Q}\subset\mathcal{A}_2$.

Observe que $H(\mathcal{P}|\mathcal{A}_1) \geqslant H(\mathcal{P}|\mathcal{A}_1\vee\widehat{\mathcal{Q}}) \geqslant H(\mathcal{P}|\mathcal{A}_1\vee\mathcal{A}_2) = H(\mathcal{P}|\mathcal{A}_1)$ e, portanto, $H(\mathcal{P}|\mathcal{A}_1\vee\widehat{\mathcal{Q}}) = H(\mathcal{P}|\mathcal{A}_1)$. Consequentemente,

$$H(\mathcal{P}\vee\mathcal{Q}|\mathcal{A}_1) = H(\mathcal{Q}|\mathcal{A}_1) + H(\mathcal{P}|\mathcal{A}_1\vee\widehat{\mathcal{Q}}) = H(\mathcal{Q}|\mathcal{A}_1) + H(\mathcal{P}|\mathcal{A}_1).$$

Agora, note que

$$\begin{aligned}H(\mathcal{P}'\vee\mathcal{P}\vee\mathcal{Q}|\mathcal{A}_1) &= H(\mathcal{P}\vee\mathcal{Q}|\mathcal{A}_1)\\ &= H(\mathcal{Q}|\mathcal{A}_1) + H(\mathcal{P}|\mathcal{A}_1)\\ &= H(\mathcal{Q}|\mathcal{A}_1) + H(\mathcal{P}'\vee\mathcal{P}|\mathcal{A}_1)\\ &= H(\mathcal{Q}|\mathcal{A}_1) + H(\mathcal{P}'|\mathcal{A}_1) + H(\mathcal{P}|\mathcal{A}_1\vee\widehat{\mathcal{P}'}).\end{aligned} \tag{21}$$

Além disso,

$$H((\mathcal{P}'\vee\mathcal{Q})\vee\mathcal{P}|\mathcal{A}_1) = H(\mathcal{P}'\vee\mathcal{Q}|\mathcal{A}_1) + H(P|\mathcal{A}_1\vee\widehat{\mathcal{P}'}\vee\widehat{\mathcal{Q}}) \tag{22}$$

Por outro lado,

$$H(\mathcal{P}'\vee\mathcal{Q}\vee\mathcal{P}|\mathcal{A}_1) = H(\mathcal{P}'|\mathcal{A}_1) + H(\mathcal{P}\vee\mathcal{Q}|\mathcal{A}_1\vee\widehat{\mathcal{P}'}) \tag{23}$$

Pelas relações (21) e (23) e de $\mathcal{P}' \prec \mathcal{P}$, temos

$$\begin{aligned}H(\mathcal{Q}|\mathcal{A}_1) + H(\mathcal{P}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'}) &= H(\mathcal{P} \vee \mathcal{Q}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'}) \\ &= H(\mathcal{P}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'}) + H(\mathcal{Q}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'} \vee \widehat{\mathcal{P}}) \\ &= H(\mathcal{P}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'}) + H(\mathcal{Q}|\mathcal{A}_1 \vee \widehat{\mathcal{P}}).\end{aligned}$$

Consequentemente, $H(\mathcal{Q}|\mathcal{A}_1) = H(\mathcal{Q}|\mathcal{A}_1 \vee \widehat{\mathcal{P}})$, donde obtemos

$$H(\mathcal{Q}|\mathcal{A}_1) \geqslant H(\mathcal{Q}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'}) \geqslant H(\mathcal{Q}|\mathcal{A}_1 \vee \widehat{\mathcal{P}}) = H(\mathcal{Q}|\mathcal{A}_1).$$

Portanto,

$$H(\mathcal{Q}|\mathcal{A}_1) = H(\mathcal{Q}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'}). \tag{24}$$

Pelas relações (22), (23) e (24) obtém-se

$$\begin{aligned}H(\mathcal{P}' \vee \mathcal{Q}|\mathcal{A}_1) + H(\mathcal{P}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'} \vee \widehat{\mathcal{Q}}) &= H(\mathcal{P}'|\mathcal{A}_1) + H(\mathcal{P} \vee \mathcal{Q}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'}) \\ &= H(\mathcal{P}'|\mathcal{A}_1) + H(\mathcal{Q}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'}) + H(\mathcal{P}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'} \vee \widehat{\mathcal{Q}}) \\ &= H(\mathcal{P}'|\mathcal{A}_1) + H(\mathcal{Q}|\mathcal{A}_1) + H(\mathcal{P}|\mathcal{A}_1 \vee \widehat{\mathcal{P}'} \vee \widehat{\mathcal{Q}}).\end{aligned} \tag{25}$$

Portanto, $H(\mathcal{P}' \vee \mathcal{Q}|\mathcal{A}_1) = H(\mathcal{P}'|\mathcal{A}_1) + H(\mathcal{Q}|\mathcal{A}_1)$.

Agora, note que

$$H(\mathcal{P}' \vee \mathcal{Q}|\mathcal{A}_1) = H(\mathcal{Q}|\mathcal{A}_1) + H(\mathcal{P}'|\mathcal{A}_1 \vee \widehat{\mathcal{Q}}),$$

donde obtemos

$$H(\mathcal{P}'|\mathcal{A}_1) = H(\mathcal{P}'|\mathcal{A}_1 \vee \widehat{\mathcal{Q}}). \tag{26}$$

Pela Proposição 1.31 existe $(\mathcal{Q}_n)_{n=1}^{\infty}$ sequência de partições de X com entropia finita tal que $\widehat{\mathcal{Q}_n} \uparrow \mathcal{A}_2$. Por (26) temos

$$H(\mathcal{P}'|\mathcal{A}_1) = H(\mathcal{P}'|\mathcal{A}_1 \vee \widehat{\mathcal{Q}_n}) \text{ para cada } n \geqslant 1. \tag{27}$$

Mas, pela Proposição 1.28 temos $\mathcal{A}_1 \vee \widehat{\mathcal{Q}_n} \uparrow \mathcal{A}_1 \vee \mathcal{A}_2$, donde fazendo $n \to \infty$ em (27) o Teorema 3.31 nos dá $H(\mathcal{P}'|\mathcal{A}_1) = H(\mathcal{P}'|\mathcal{A}_1 \vee \mathcal{A}_2)$. Isso conclui a demonstração.

□

4 O TEOREMA DE ROKHLIN-SINAI

A fim de motivarmos a noção de K-automorfismo e prepararmo-nos para o Teorema de Rokhlin-Sinai - objetivo último do presente trabalho - vejamos o seguinte resultado devido a Kolmogorov (KOLMOGOROV, 1958):

Teorema 4.1 . (Lei 0-1 de Kolmogorov) Se $(X,\mathcal{X},\mu,T)$ é um deslocamento de Bernoulli e $\mathcal{P} = \{P_1, \ldots, P_a\}$ é uma partição finita satisfazendo

(i) $\displaystyle\bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{P} = \mathcal{X} \pmod{\mu}$;

(ii) $\displaystyle\mu\Big(\bigcap_{i=m}^{n} T^{i}P_{j_i}\Big) = \prod_{i=m}^{n} \mu(P_{j_i}),\ j_i \in \{1, \ldots, a\},\ m \leqslant i \leqslant n \text{ e } m \leqslant n \text{ em } \mathbb{Z},$

então

$$\bigcap_{k\geqslant 1} \bigvee_{i\geqslant k} T^{-i}\mathcal{P} = \{\varnothing, X\}.$$

Demonstração. Tome $A \in \displaystyle\bigcap_{k\geqslant 1} \bigvee_{i\geqslant k} T^{-i}\mathcal{P}$ arbitrário. Temos de mostrar que $\mu(A) = 0$ ou $\mu(A) = 1$. Realmente, como $A \in \displaystyle\bigcap_{k\geqslant 1} \bigvee_{i\geqslant k} T^{-i}\mathcal{P} \subset \bigvee_{i=-\infty}^{+\infty} T^{-i}\mathcal{P}$, para cada $k \geqslant 1$ existem números inteiros $m_k \leqslant n_k$ e $A_k \in \displaystyle\bigvee_{i=m_k}^{n_k} T^{-i}\mathcal{P}$ tais que $|m_k| \to \infty$, $|n_k| \to \infty$ e $\mu(A_k \triangle A) \to 0$ quando $k \to \infty$. Então, $\mu(A_k) \to \mu(A)$ e $\mu(A_k \cap A) \to \mu(A)$ quando $k \to \infty$. Por outro lado, como $A \in \displaystyle\bigvee_{i\geqslant n_k+1} T^{-i}\mathcal{P}$, $A_k \in \displaystyle\bigvee_{i=m_k}^{n_k} T^{-i}\mathcal{P}$ e, $\displaystyle\bigvee_{i\geqslant n_k+1} T^{-i}\mathcal{P}$ e $\displaystyle\bigvee_{i=m_k}^{n_k} T^{-i}\mathcal{P}$ são independentes (pela condição (ii) acima juntamente com a invariância de μ), obtemos

$$\mu(A_k \cap A) = \mu(A_k)\mu(A) \text{ para cada } k \geqslant 1.$$

Agora, fazendo $k \to \infty$ conclui-se que $\mu(A) = \mu(A)^2$, o que implica $\mu(A) = 0$ ou $\mu(A) = 1$. Isso conclui a demonstração. $\square$

Definição 4.2 . Um sistema dinâmico $(X,\mathcal{X},\mu,T)$ é dito um *deslocamento de Kolmogorov* se existe uma partição finita $\mathcal{P} = \{P_1, \ldots, P_a\}$ satisfazendo:

(i) $\displaystyle\bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{P} = \mathcal{X} \pmod{\mu}$;

(ii) $\displaystyle\bigcap_{k\geqslant 1} \bigvee_{i\geqslant k} T^{-i}\mathcal{P} = \{\varnothing, X\}$.

Observação 4.3 . Se $(X,\mathcal{X},\mu,T)$ é um deslocamento de Bernoulli, então a Lei 0-1 de Kolmogorov nos diz que $(X,\mathcal{X},\mu,T)$ é um deslocamento de Kolmogorov. Cabe mencionar que Kolmogorov conjecturou (KOLMOGOROV, 1958) que, se dois deslocamentos de Kolmogorov têm a mesma entropia, então eles são isomorfos; em particular, todo deslocamento de Kolmogorov é isomorfo a um deslocamento de Bernoulli. Tal conjectura mostrou-se falsa, com o primeiro contra-exemplo tendo sido construído por Ornstein (ORNSTEIN, 1973).

Lema 4.4 . Sejam $(X,\mathcal{X},\mu,T)$ um sistema dinâmico e $\mathcal{P}$ e $\mathcal{Q}$ partições com entropia finita. Então,

$$H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\Big)=H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\vee\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n}T^{-j}\mathcal{Q}\Big).$$

Em palavras, se conhecemos o "passado completo" do processo $(T,\mathcal{P})$, então o "passado remoto" do processo $(T,\mathcal{Q})$ não fornece ganho de informação ao "presente" do processo $(T,\mathcal{P})$.

Demonstração. Para cada $n\geqslant 1$ a Proposição 3.21 nos fornece:

$$\begin{aligned}H(T,\mathcal{P}\vee\mathcal{Q})&=H\Big(\mathcal{P}\vee\mathcal{Q}|\bigvee_{j=1}^{\infty}T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)\\&=\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1}T^{i}(\mathcal{P}\vee\mathcal{Q})|\bigvee_{j=1}^{\infty}T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)\\&\leqslant\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1}T^{i}(\mathcal{P}\vee\mathcal{Q})|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\Big)\\&\leqslant\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1}T^{i}(\mathcal{P}\vee\mathcal{Q})\Big)\\&\to H(T^{-1},\mathcal{P}\vee\mathcal{Q})=H(T,\mathcal{P}\vee\mathcal{Q}),\end{aligned}$$

onde na última igualdade utilizamos a relação (11). Daí, vemos que

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1}T^{i}(\mathcal{P}\vee\mathcal{Q})|\bigvee_{j=1}^{\infty}T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)&=H(T,\mathcal{P}\vee\mathcal{Q})\\&=\lim_{n\to\infty}\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1}T^{i}(\mathcal{P}\vee\mathcal{Q})|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\Big)\end{aligned}\tag{28}$$

Agora, note que

$$\begin{aligned}\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1}T^{i}(\mathcal{P}\vee\mathcal{Q})|\bigvee_{j=1}^{\infty}T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)&=\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1}T^{i}\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)\\&\quad+\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1}T^{i}\mathcal{Q}|\bigvee_{i=0}^{n-1}T^{i}\mathcal{P}\vee\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\vee\bigvee_{j=1}^{\infty}T^{-j}\mathcal{Q}\Big)\\&=\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1}T^{i}\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)\\&\quad+\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1}T^{i}\mathcal{Q}|\bigvee_{j=-\infty}^{n-1}T^{j}\mathcal{P}\vee\bigvee_{j=1}^{\infty}T^{-j}\mathcal{Q}\Big)\end{aligned}\tag{29}$$

Similarmente,

$$\begin{aligned}
\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i(\mathcal{P}\vee\mathcal{Q})|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) &= \frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big)\\
&+\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{Q}|\bigvee_{i=0}^{n-1} T^i\mathcal{P}\vee\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big)\\
&=\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big)\\
&+\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{Q}|\bigvee_{j=-\infty}^{n-1} T^{j}\mathcal{P}\Big)
\end{aligned}\tag{30}$$

Além disso, observe que os dois termos do lado direito de (30) são, respectivamente, maiores ou iguais aos correspondentes dois termos do lado direito de (29).

Portanto, tem-se

$$\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) \geqslant \frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)\tag{31}$$

e

$$\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{Q}|\bigvee_{j=-\infty}^{n-1} T^{j}\mathcal{P}\Big) \geqslant \frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{Q}|\bigvee_{j=-\infty}^{n-1} T^{j}\mathcal{P}\vee\bigvee_{j=1}^{\infty} T^{-j}\mathcal{Q}\Big)\tag{32}$$

Consequentemente, por (28), (29), (30), (31) e (32) obtemos:

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big) &= \lim_{n\to\infty}\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big)\\
&= H(T,\mathcal{P}),
\end{aligned}\tag{33}$$

onde na última igualdade usamos a Proposição 3.21.

Ou seja,

$$H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) = \lim_{n\to\infty}\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big).\tag{34}$$

Finalmente, pelo item (i) da Proposição 3.28 podemos escrever:

$$\begin{aligned}
H\Big(\bigvee_{i=0}^{n-1} T^i\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big) &= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big) + H\Big(T\mathcal{P}|\mathcal{P}\vee\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)\\
&+\ldots+H\Big(T^{n-1}\mathcal{P}|\mathcal{P}\vee T\mathcal{P}\vee\ldots\vee T^{n-2}\mathcal{P}\vee\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big) + H\Big(\mathcal{P}|T^{-1}\mathcal{P}\vee\bigvee_{j=2}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)\\
&+\ldots+H\Big(\mathcal{P}|T^{-(n-1)}\mathcal{P}\vee\ldots\vee T^{-1}\mathcal{P}\vee\bigvee_{j=n}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)
\end{aligned}$$

$$\begin{aligned}
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big) + H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\bigvee_{j\geqslant 2} T^{-j}\mathcal{Q}\Big)\\
&+\ldots + H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\bigvee_{j\geqslant n} T^{-j}\mathcal{Q}\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big) + \sum_{l=1}^{n-1} H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\bigvee_{j\geqslant l+1} T^{-j}\mathcal{Q}\Big).
\end{aligned} \tag{35}$$

Portanto, como $\lim\limits_{n\to\infty}\frac{1}{n}H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big) = 0$, segue de (35), (34) e (33):

$$\begin{aligned}
H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) &= \lim_{n\to\infty}\frac{1}{n}H\Big(\bigvee_{i=0}^{n-1} T^{i}\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}(\mathcal{P}\vee\mathcal{Q})\Big)\\
&= \lim_{n\to\infty}\frac{1}{n}\sum_{l=1}^{n-1} H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\bigvee_{j\geqslant l+1} T^{-j}\mathcal{Q}\Big)\\
&= \lim_{n\to\infty} H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\bigvee_{j\geqslant n+1} T^{-j}\mathcal{Q}\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{Q}\Big),
\end{aligned}$$

onde na penúltima igualdade utilizamos a Proposição 1.50 e na última igualdade utilizamos o Teorema 3.31. Isso conclui a demonstração. □

Teorema 4.5 . Sejam $(X,\mathcal{X},\mu,T)$ um sistema dinâmico e $\mathcal{R}$ uma partição com entropia finita. Então, $H(T,\mathcal{R}) = 0$ se, e somente se, $\widehat{\mathcal{R}} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{Q}$ para alguma partição com entropia finita $\mathcal{Q}$.

Demonstração. Suponha $H(T,\mathcal{R}) = 0$. Então, pela Observação 3.33 temos $\widehat{\mathcal{R}} \subset \bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{R}}$. Daí, para cada $n \geqslant 0$ temos

$$\widehat{T^{-n}\mathcal{R}} \subset \bigvee_{j=n+1}^{\infty}\widehat{T^{-j}\mathcal{R}} \subset \bigvee_{j=0}^{\infty}\widehat{T^{-j}\mathcal{R}},$$

o que implica

$$\widehat{\mathcal{R}} \subset \bigvee_{j\geqslant n}\widehat{T^{-j}\mathcal{R}}$$

para todo $n \geqslant 1$. Consequentemente, obtemos

$$\widehat{\mathcal{R}} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n}\widehat{T^{-j}\mathcal{R}}.$$

Reciprocamente, suponha que $\widehat{\mathcal{R}} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{Q}$ para alguma partição com entropia finita $\mathcal{Q}$. Pelo Lema 4.4, para qualquer partição com entropia finita $\mathcal{P}$ temos

$$H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\Big)=H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\vee\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n}T^{-j}\mathcal{Q}\Big),$$

o que implica

$$H(T,\mathcal{P})=H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\Big)=H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\vee\mathcal{R}\Big),$$

pois

$$\begin{aligned}H(T,\mathcal{P})&=H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\Big)\\&=H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\vee\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n}T^{-j}\mathcal{Q}\Big)\\&\leqslant H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\vee\mathcal{R}\Big)\\&\leqslant H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{P}\Big)\\&=H(T,\mathcal{P}).\end{aligned}$$

Em particular, para $\mathcal{P}=\mathcal{R}$ obtemos:

$$H(T,\mathcal{R})=H\Big(\mathcal{R}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{R}\Big)=H\Big(\mathcal{R}|\bigvee_{j=1}^{\infty}T^{-j}\mathcal{R}\vee\mathcal{R}\Big)=0.$$

□

Lema 4.6 . Seja $(X,\mathcal{X},\mu,T)$ um sistema dinâmico e considere $\mathcal{P}$ e $\mathcal{Q}$ partições com entropia finita. Para cada $k\geqslant 0$ tem-se:

$$H\Big(\bigvee_{i=-k}^{k}T^{i}\mathcal{P}|\bigvee_{j=k+1}^{\infty}T^{-j}\mathcal{P}\Big)=H\Big(\bigvee_{i=-k}^{k}T^{i}\mathcal{P}|\bigvee_{j=k+1}^{\infty}T^{-j}\mathcal{P}\vee\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n}T^{-j}\mathcal{Q}\Big).$$

Demonstração. Usaremos indução sobre $k\geqslant 0$. Note que o caso $k=0$ é precisamente o Lema 4.4. Suponhamos o resultado válido para algum $k\geqslant 0$ e provemo-lo para $k+1$.

Além disso, para simplificar a notação, escreva $\mathcal{R}:=\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n}T^{-j}\mathcal{Q}$.

Agora, vamos utilizar repetidas vezes o item (i) da Proposição 3.28.

Realmente,

$$\begin{aligned}H\Big(\bigvee_{i=-(k+1)}^{k+1}T^{i}\mathcal{P}|\bigvee_{j=k+2}^{\infty}T^{-j}\mathcal{P}\Big)&=H\Big(T^{-(k+1)}\mathcal{P}\vee T^{k+1}\mathcal{P}\vee\bigvee_{i=-k}^{k}T^{i}\mathcal{P}|\bigvee_{j=k+2}^{\infty}T^{-j}\mathcal{P}\Big)\\&=H\Big(T^{-(k+1)}\mathcal{P}|\bigvee_{j=k+2}^{\infty}T^{-j}\mathcal{P}\Big)\end{aligned}$$

$$
\begin{aligned}
&+ H\Big(T^{k+1}\mathcal{P} \vee \bigvee_{i=-k}^{k} T^i\mathcal{P} | T^{-(k+1)}\mathcal{P} \vee \bigvee_{j=k+2}^{\infty} T^{-j}\mathcal{P}\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) + H\Big(\bigvee_{i=-k}^{k} T^i\mathcal{P}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P}\Big)\\
&+ H\Big(T^{k+1}\mathcal{P}| \bigvee_{i=-k}^{k} T^i\mathcal{P} \vee \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P}\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) + H\Big(\bigvee_{i=-k}^{k} T^i\mathcal{P}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P}\Big)\\
&+ H\Big(T^{k+1}\mathcal{P}| \bigvee_{j=-k}^{\infty} T^{-j}\mathcal{P}\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) + H\Big(\bigvee_{i=-k}^{k} T^i\mathcal{P}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P}\Big)\\
&+ H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big) + H\Big(\bigvee_{i=-k}^{k} T^i\mathcal{P}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big)\\
&+ H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big) + H\Big(\bigvee_{i=-k}^{k} T^i\mathcal{P}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big)\\
&+ H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big)\\
&+ H\Big(\bigvee_{i=-k}^{k} T^i\mathcal{P}| T^{-(k+1)}\mathcal{P} \vee \bigvee_{j=k+2}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big)\\
&+ H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big),
\end{aligned}
$$

onde na penúltima igualdade usamos o caso $k = 0$ e na antepenúltima igualdade usamos a hipótese de indução.

Daí, vemos que

$$
H\Big(\bigvee_{i=-(k+1)}^{k+1} T^i\mathcal{P}| \bigvee_{j=k+2}^{\infty} T^{-j}\mathcal{P}\Big) = H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big)
$$

$$
\begin{aligned}
&+ H\Big(\bigvee_{i=-k}^{k} T^{i}\mathcal{P}|T^{-(k+1)}\mathcal{P}\vee\bigvee_{j=k+2}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big)\\
&+ H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big)\\
&+ H\Big(\bigvee_{i=-k}^{k} T^{i}\mathcal{P}\vee T^{-(k+1)}\mathcal{P}|\bigvee_{j=k+2}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big)\\
&- H\Big(T^{-(k+1)}\mathcal{P}|\bigvee_{j=k+2}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big)\\
&+ H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big)\\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big)\\
&+ H\Big(\bigvee_{i=-k}^{k} T^{i}\mathcal{P}\vee T^{-(k+1)}\mathcal{P}|\bigvee_{j=k+2}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big)\\
&= H\Big(\bigvee_{i=-(k+1)}^{k+1} T^{i}\mathcal{P}|\bigvee_{j=k+2}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big),
\end{aligned}
$$

onde na penúltima igualdade usamos o fato que

$$H\Big(T^{-(k+1)}\mathcal{P}|\bigvee_{j=k+2}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big) = H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty} T^{-j}\mathcal{P}\vee\mathcal{R}\Big).$$

Isso conclui a demonstração. □

Teorema 4.7 . Seja $(X,\mathcal{X},\mu,T)$ um sistema dinâmico, onde $(X,\mathcal{X},\mu)$ é um espaço de probabilidade separável. Se $\mathcal{P}$ e $\mathcal{Q}$ são partições finitas tais que $\widehat{\mathcal{Q}} \subset \bigvee_{i=-\infty}^{\infty} \widehat{T^{i}\mathcal{P}}$, então $\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{Q} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P}$.

Em palavras, dados dois "experimentos" $\mathcal{P}$ e $\mathcal{Q}$, se o "presente" do processo $(T,\mathcal{Q})$ pode ser descrito pela "história completa" do processo $(T,\mathcal{P})$, então o "passado remoto" do processo $(T,\mathcal{Q})$ pode ser descrito pelo "passado remoto" do processo $(T,\mathcal{P})$.

Demonstração. Pela Proposição 1.31 juntamente com o Teorema 3.29 e do fato de a função $H(F|G)$ ser contínua em F, é suficiente mostrar que

$$H\Big(\mathcal{R}|\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P}\Big) = 0$$

para cada partição finita $\mathcal{R} \subset \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{Q}$.

Para tal, fixemos $k \geqslant m \geqslant 1$ arbitrariamente e tomemos uma partição finita arbitrária $\mathcal{S} \subset \bigvee_{i=-m}^{m} T^i\mathcal{P} \subset \bigvee_{i=-k}^{k} T^i\mathcal{P}$. Pelo Lema 4.6 temos:

$$H\Big(\bigvee_{i=-k}^{k} T^i\mathcal{P}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P}\Big) = H\Big(\bigvee_{i=-k}^{k} T^i\mathcal{P}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P} \vee \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{Q}\Big).$$

Como $\mathcal{S} \subset \bigvee_{i=-k}^{k} T^i\mathcal{P}$, pela Proposição 3.47 temos:

$$H\Big(\mathcal{S}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P}\Big) = H\Big(\mathcal{S}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P} \vee \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{Q}\Big).$$

Daí, se $\mathcal{R} \subset \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{Q}$ é uma partição finita arbitrária, então

$$\begin{aligned} H\Big(\mathcal{S}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big) &\leqslant H\Big(\mathcal{S}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P}\Big) \\ &= H\Big(\mathcal{S}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P} \vee \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{Q}\Big) \\ &\leqslant H\Big(\mathcal{S}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big), \end{aligned}$$

donde $H\Big(\mathcal{S}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P} \vee \mathcal{R}\Big) = H\Big(\mathcal{S}| \bigvee_{j=k+1}^{\infty} T^{-j}\mathcal{P}\Big)$. Agora, fazendo $k \to +\infty$ o Teorema 3.31 nos dá

$$H\Big(\mathcal{S}| \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{P} \vee \mathcal{R}\Big) = H\Big(\mathcal{S}| \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{P}\Big) \tag{36}$$

para cada partição finita $\mathcal{R} \subset \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{Q}$. Além disso, como

$$\mathcal{R} \subset \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{Q} \subset \bigvee_{i=-\infty}^{\infty} \widehat{T^i\mathcal{P}},$$

$\mathcal{S} \subset \bigvee_{i=-m}^{m} T^i\mathcal{P}$ e $\bigvee_{i=-m}^{m} T^i\mathcal{P} \uparrow \bigvee_{i=-\infty}^{\infty} T^i\mathcal{P}$, o Lema 3.23 nos permite substituir $\mathcal{S}$ por $\mathcal{R}$ em (36), o que nos dá:

$$0 = H\Big(\mathcal{R}| \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{P} \vee \mathcal{R}\Big) = H\Big(\mathcal{R}| \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{P}\Big),$$

onde na primeira igualdade utilizamos o Teorema 3.29. Portanto, usando o Teorema 3.29 na segunda igualdade obtém-se $\mathcal{R} \subset \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{P}$, o que conclui a demonstração. $\square$

O próximo resultado estende o Teorema 4.7.

Teorema 4.8 . Seja $(X,\mathcal{X},\mu,T)$ um sistema dinâmico, onde $(X,\mathcal{X},\mu)$ é um espaço de probabilidade separável. Se $\mathcal{Q}$ é uma partição com entropia finita e $\mathcal{A}$ é uma sub-σ-álgebra que satisfazem $\widehat{\mathcal{Q}} \subset \bigvee_{i=-\infty}^{\infty} T^i\mathcal{A}$, então $\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{Q} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{A}$.

Demonstração. Pela Proposição 1.31 existe uma sequência $(\mathcal{P}_l)_{l\geqslant 1}$ de partições finitas tal que $\widehat{\mathcal{P}_l} \uparrow \mathcal{A}$. Fixe $l \geqslant 1$ arbitrariamente e tome $\mathcal{S} \subset \bigvee_{i=-\infty}^{\infty} T^i\mathcal{P}_l$ uma partição finita arbitrária. Pela demonstração do Teorema 4.7, tem-se:

$$H\Big(\mathcal{S}|\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P}_l \vee \mathcal{R}\Big) = H\Big(\mathcal{S}|\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P}_l\Big)$$

para cada partição finita $\mathcal{R} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{Q}$. Agora, para cada $m \geqslant l$ temos $\mathcal{S} \subset \bigvee_{i=-\infty}^{\infty} T^i\mathcal{P}_l \subset \bigvee_{i=-\infty}^{\infty} T^i\mathcal{P}_m$ e, portanto, fazendo $m \to \infty$ a Proposição 1.29 nos diz que existe uma sub-σ-álgebra $\mathcal{G} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{A}$ tal que $\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P}_m \uparrow \mathcal{G}$. Daí, o Teorema 3.31 nos dá:

$$H\Big(\mathcal{S}|\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P}_m \vee \mathcal{R}\Big) \downarrow H(\mathcal{S}|\mathcal{G} \vee \mathcal{R}) \tag{37}$$

e

$$H\Big(\mathcal{S}|\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P}_m\Big) \downarrow H(\mathcal{S}|\mathcal{G}). \tag{38}$$

Logo, como

$$H\Big(\mathcal{S}|\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P}_m \vee \mathcal{R}\Big) = H\Big(\mathcal{S}|\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P}_m\Big)$$

para cada partição finita $\mathcal{R} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{Q}$ e para todo $m \geqslant 1$, fazendo $m \to \infty$ obtemos pelas relações (37) e (38):

$$H(\mathcal{S}|\mathcal{G} \vee \mathcal{R}) = H(\mathcal{S}|\mathcal{G}). \tag{39}$$

Agora, como $\widehat{\mathcal{P}_l} \uparrow \mathcal{A}$, segue da Proposição 1.29 que as partições $\mathcal{S} \subset \bigvee_{i=-\infty}^{\infty} T^i\mathcal{P}_l$ $(l \geqslant 1)$ são densas em $\bigvee_{i=-\infty}^{\infty} T^i\mathcal{A}$. Consequentemente, como a função $H(F|\mathcal{G})$ é contínua em F e $\mathcal{R} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{Q} \subset \bigvee_{i=-\infty}^{\infty} T^i\mathcal{A}$, podemos substituir $\mathcal{S}$ por $\mathcal{R}$ na relação (39) e, assim, obter:

$$0 = H(\mathcal{R}|\mathcal{G} \vee \mathcal{R}) = H(\mathcal{R}|\mathcal{G}).$$

Portanto, $\mathcal{R} \subset \mathcal{G} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{A}$. Daí, como $\mathcal{R} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{Q}$ é partição finita arbitrária, novamente pela Proposição 1.31 obtemos

$$\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{Q} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{A},$$

como queríamos demonstrar.

□

Teorema 4.9 . Se $(X,\mathcal{X},\mu,T)$ é um deslocamento de Kolmogorov, então

$$\bigcap_{k\geqslant 1}\bigvee_{i\geqslant k} T^{-i}\mathcal{Q} = \{\varnothing,X\}$$

para cada partição finita $\mathcal{Q}$.

Demonstração. Como $(X,\mathcal{X},\mu,T)$ é um deslocamento de Kolmogorov, por definição existe uma partição finita $\mathcal{P}$ satisfazendo:

(i) $\bigvee_{i=-\infty}^{\infty} T^{-i}\mathcal{P} = \mathcal{X} \pmod{\mu}$;

(ii) $\bigcap_{k\geqslant 1}\bigvee_{i\geqslant k} T^{-i}\mathcal{P} = \{\varnothing,X\}$.

Daí, se $\mathcal{Q}$ é uma partição finita arbitrária, então pelo item (i) acima temos trivialmente $\mathcal{Q} \subset \bigvee_{i=-\infty}^{\infty} T^{i}\mathcal{P}$. Consequentemente, pelo Teorema 4.9 segue que $\bigcap_{k\geqslant 1}\bigvee_{i\geqslant k} T^{-i}\mathcal{Q} \subset \bigcap_{k\geqslant 1}\bigvee_{i\geqslant k} T^{-i}\mathcal{P}$. Portanto, pelo item (ii) conclui-se $\bigcap_{k\geqslant 1}\bigvee_{i\geqslant k} T^{-i}\mathcal{Q} = \{\varnothing, X\}$. □

Proposição 4.10 . Se $(X,\mathcal{X},\mu,T)$ é um sistema dinâmico e $\mathcal{P}$ é uma partição com entropia finita tal que $\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P} = \{\varnothing, X\}$, então

$$\lim_{n\to\infty} H(T^n,\mathcal{P}) = H(\mathcal{P}).$$

Em particular, $H(T,\mathcal{P}) > 0$.

Demonstração. Se $\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P} = \{\varnothing, X\}$, então o Teorema 3.31 nos dá:

$$\begin{aligned} H(\mathcal{P}) &= H\Big(\mathcal{P}|\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P}\Big) \\ &= \lim_{n\to\infty} H\Big(\mathcal{P}|T^{-n}\Big(\bigvee_{j\geqslant 0} T^{-j}\mathcal{P}\Big)\Big) \\ &\leqslant \lim_{n\to\infty} H\Big(\mathcal{P}|T^{-n}\Big(\bigvee_{j\geqslant 0} T^{-nj}\mathcal{P}\Big)\Big) \\ &= \lim_{n\to\infty} H\Big(\mathcal{P}|\bigvee_{j\geqslant 0}(T^{n})^{-(j+1)}\mathcal{P}\Big) \end{aligned}$$

$$
\begin{aligned}
&= \lim_{n\to\infty} H\Big(\mathcal{P}|\bigvee_{j\geqslant 1}(T^n)^{-j}\mathcal{P}\Big)\\
&= \lim_{n\to\infty} H(T^n,\mathcal{P})\\
&\leqslant H(\mathcal{P}).
\end{aligned}
$$

Portanto, $\lim\limits_{n\to\infty} H(T^n,\mathcal{P}) = H(\mathcal{P}) > 0$. Daí, pela relação (10) segue que $nH(T,\mathcal{P}) > 0$ para todo $n \geqslant 1$ suficientemente grande, donde $H(T,\mathcal{P}) > 0$. □

Observação 4.11 . Segue do Teorema 4.9 juntamente com a Proposição 4.10 que, se $(X,\mathcal{X},\mu,T)$ é um deslocamento de Kolmogorov, então $H(T,\mathcal{P}) > 0$ para toda partição finita não-trivial $\mathcal{P}$. O Teorema de Rokhlin-Sinai, como veremos à frente, nos dará uma generalização desse fato e também a sua recíproca em versão também geral.

Agora, suponhamos que $(X,\mathcal{X},\mu,T)$ é um deslocamento de Kolmogorov e considere uma partição finita $\mathcal{P}$ satisfazendo:

(i) $\bigvee\limits_{i=-\infty}^{\infty} T^{-i}\mathcal{P} = \mathcal{X} \pmod{\mu}$;

(ii) $\bigcap\limits_{n\geqslant 1}\bigvee\limits_{j\geqslant n} T^{-j}\mathcal{P} = \{\varnothing,X\}$.

Então, $\mathcal{A} := \bigvee\limits_{j\geqslant 0} T^{-j}\mathcal{P}$ é uma sub-σ-álgebra que satisfaz:

(i) $T^{-1}\mathcal{A} \subset \mathcal{A}$;

(ii) $T^n\mathcal{A} \uparrow \mathcal{X}$;

(iii) $T^{-n}\mathcal{A} \downarrow \{\varnothing, X\}$.

Isso motiva a seguinte definição fundamental:

Definição 4.12 . (K-automorfismo) Dizemos que o sistema dinâmico $(X,\mathcal{X},\mu,T)$ é um *K-automorfismo* se existe uma sub-σ-álgebra $\mathcal{A} \subset \mathcal{X}$ satisfazendo:

(i) $T^{-1}\mathcal{A} \subset \mathcal{A}$;

(ii) $T^n\mathcal{A} \uparrow \mathcal{X}$, ou seja, $\bigvee\limits_{i=-\infty}^{\infty} T^{i}\mathcal{A} = \mathcal{X}$;

(iii) $T^{-n}\mathcal{A} \downarrow \{\varnothing, X\}$, ou seja, $\bigcap\limits_{n\geqslant 1}\bigvee\limits_{j\geqslant n} T^{-j}\mathcal{A} = \{\varnothing, X\}$.

Observação 4.13 . Segue imediatamente da Definição 4.12 que todo deslocamento de Kolmogorov é um K-automorfismo. Além disso, se $(X,\mathcal{X},\mu,T)$ é um K-automorfismo, então segue do Teorema 4.8 que todo processo $(T,\mathcal{P})$ advindo de uma partição finita $\mathcal{P}$ é um *processo de Kolmogorov*. Em outras palavras, se $(X,\mathcal{X},\mu,T)$ é um K-automorfismo e $\mathcal{P}$ é uma partição finita arbitrária, então a restrição de T à sub-σ-álgebra $\bigvee\limits_{i=-\infty}^{\infty} T^{i}\mathcal{P}$ é um deslocamento de Kolmogorov. Realmente, se $\mathcal{A} \subset \mathcal{X}$ é uma sub-σ-álgebra que satisfaz às condições da Definição 4.12 e $\mathcal{P}$ é uma partição finita arbitrária, então $\widehat{\mathcal{P}} \subset \bigvee\limits_{i=-\infty}^{\infty} T^{i}\mathcal{A}$, donde pelo Teorema 4.8 segue que

$$\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{P} \subset \bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{A} = \{\varnothing, X\}.$$

Daí, pela Proposição 4.10 obtemos $H(T,\mathcal{P}) > 0$ para toda $\mathcal{P}$ partição de X com entropia finita.

Definição 4.14 . Dizemos que um sistema dinâmico $(X,\mathcal{X},\mu,T)$ tem *entropia completamente positiva* (CPE) se $H(T,\mathcal{P}) > 0$ para cada partição não-trivial com entropia finita $\mathcal{P}$.

O Teorema de Rokhlin-Sinai garante que um sistema dinâmico sobre um espaço de probabilidade separável tem entropia completamente positiva se, e somente se, é um K-automorfismo.

A partir desse momento fixemos um sistema dinâmico $(X,\mathcal{X},\mu,T)$, onde $(X,\mathcal{X},\mu)$ é um espaço de probabilidade separável.

A fim de atingirmos o nosso objetivo principal nesse trabalho - o Teorema de Rokhlin-Sinai - será importante analisarmos a "parte determinística maximal" do sistema $(X,\mathcal{X},\mu,T)$. Vejamos isso em termos precisos:

Definição 4.15 . Seja $\mathcal{D}(T)$ a coleção de todos os elementos $A \in \mathcal{X}$ para os quais existe uma partição $\mathcal{P}$ de X tal que $A \in \mathcal{P}$, $H(\mathcal{P}) < \infty$ e $H(T,\mathcal{P}) = 0$, incluindo o conjunto $A = \varnothing$.

Proposição 4.16 . $\mathcal{D}(T)$ é uma álgebra.

Demonstração. Primeiro, é claro que $X \in \mathcal{D}(T)$, pois a partição trivial $\mathcal{P} := \{\varnothing, X\}$ satisfaz $X \in \mathcal{P}$, $H(\mathcal{P}) < \infty$ e $H(T,\mathcal{P}) = 0$.

Agora, sejam $A,B \in \mathcal{D}(T)$ arbitrários. Pela definição de $\mathcal{D}(T)$ existem partições $\mathcal{P}$ e $\mathcal{Q}$ de X com entropia finita tais que $A \in \mathcal{P}, B \in \mathcal{Q}$ e $H(T,\mathcal{P}) = H(T,\mathcal{Q}) = 0$.

Note que $A \cap B \in \mathcal{P} \vee \mathcal{Q}$. Pelo item (iv) da Proposição 3.7 e pelo Lema 3.14 temos $H(\mathcal{P} \vee \mathcal{Q}) \leqslant H(\mathcal{P}) + H(\mathcal{Q}) < \infty$ e $H(T,\mathcal{P} \vee \mathcal{Q}) \leqslant H(T,\mathcal{P}) + H(T,\mathcal{Q}) = 0$. Portanto, $A \cap B \in \mathcal{D}(T)$.

Por fim, como $A \in \mathcal{P}$, temos que $\mathcal{Q} := \{A,A^c\} \prec \mathcal{P}$ é uma partição de X tal que $H(\mathcal{Q}) \leqslant H(\mathcal{P}) < +\infty$ e $H(T,\mathcal{Q}) \leqslant H(T,\mathcal{P}) = 0$. Daí, $A^c \in \mathcal{D}(T)$ e, como $A \cup B = (A^c \cap B^c)^c$, conclui-se que $A \cup B \in \mathcal{D}(T)$. Portanto, $\mathcal{D}(T)$ é uma álgebra, como queríamos demonstrar.

□

Definição 4.17 . (σ-álgebra de Pinsker) A *σ-álgebra de Pinsker*, denotada por $\mathcal{P}(T)$, é a σ-álgebra gerada pela álgebra $\mathcal{D}(T)$, ou seja, é a menor σ-álgebra que contém a álgebra $\mathcal{D}(T)$.

Observação 4.18 . Segue imediatamente da definição de $\mathcal{P}(T)$ que, se $\mathcal{P}$ é uma partição de X com entropia finita que satisfaz $H(T,\mathcal{P}) = 0$, então $\widehat{\mathcal{P}} \subset \mathcal{P}(T)$.

Veremos agora algumas propriedades básicas da σ-álgebra de Pinsker $\mathcal{P}(T)$:

Proposição 4.19 . $T^{-1}\mathcal{P}(T) = \mathcal{P}(T)$.

Demonstração. Primeiro vamos mostrar que $T^{-1}\mathcal{P}(T) \subset \mathcal{P}(T)$. Tome $A \in \mathcal{D}(T)$ arbitrário. Por definição, existe uma partição $\mathcal{P}$ tal que $A \in \mathcal{P}$, $H(\mathcal{P}) < \infty$ e $H(T,\mathcal{P}) = 0$. Daí, como $T^{-1}\mathcal{P}$ é uma partição tal que $H(T^{-1}\mathcal{P}) = H(\mathcal{P}) < \infty$ e $H(T,T^{-1}\mathcal{P}) = H(T,\mathcal{P}) = 0$, pela Observação 3.33 obtemos:

$$\begin{aligned}
\widehat{T^{-1}\mathcal{P}} &\subset \bigvee_{j=1}^{\infty} \widehat{T^{-j}(T^{-1}\mathcal{P})} \\
&= \bigvee_{j=2}^{\infty} \widehat{T^{-j}\mathcal{P}} \\
&\subset \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \\
&\subset \bigvee_{j=0}^{\infty} \widehat{T^{-j}\mathcal{P}}.
\end{aligned}$$

Portanto, vemos que $T^{-1}A \in \widehat{T^{-1}\mathcal{P}} \subset \bigvee_{j=0}^{\infty} \widehat{T^{-j}\mathcal{P}}$. Daí, pela Proposição 1.27 temos que, para cada $l \geqslant 1$ existem $n_l \geqslant 1$ e $C_{n_l} \in \bigvee_{j=0}^{n_l} \widehat{T^{-j}\mathcal{P}}$ tais que $\mu(T^{-1}A \triangle C_{n_l}) < 2^{-l}$. Agora, pela Proposição 3.18 tem-se:

$$0 = H(T,\mathcal{P}) = H\Big(T, \bigvee_{j=0}^{n_l} T^{-j}\mathcal{P}\Big) \text{ para cada } l \geqslant 1.$$

Além disso, pela sub-aditividade de H, temos:

$$H\Big(\bigvee_{j=0}^{n_l} T^{-j}\mathcal{P}\Big) \leqslant (n_l + 1)H(\mathcal{P}) < \infty \text{ para cada } l \geqslant 1.$$

Isto implica que cada átomo de $\bigvee_{j=0}^{n_l} T^{-j}\mathcal{P}$ pertence a $\mathcal{D}(T)$ e, portanto, a $\mathcal{P}(T)$. Daí, obtemos $C_{n_l} \in \mathcal{P}(T)$ para cada $l \geqslant 1$. Portanto, fazendo $l \to \infty$ em $\mu(T^{-1}A \triangle C_{n_l}) < 2^{-l}$ obtemos pelo Teorema 1.16:

$$T^{-1}A = \liminf C_{n_l} \in \mathcal{P}(T).$$

Consequentemente, $T^{-1}\mathcal{D}(T) \subset \mathcal{P}(T)$, o que por sua vez implica $T^{-1}\mathcal{P}(T) \subset \mathcal{P}(T)$. Agora, vamos mostrar a inclusão oposta: $\mathcal{P}(T) \subset T^{-1}\mathcal{P}(T)$. Tome $A \in \mathcal{D}(T)$ arbitrário. Novamente pela definição existe uma partição $\mathcal{P}$ tal que $A \in \mathcal{P}$, $H(\mathcal{P}) < \infty$ e $H(T,\mathcal{P}) = 0$. Novamente, como $H(T,\mathcal{P}) = 0 \iff \widehat{\mathcal{P}} \subset \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}}$ (Observação 3.33), para cada $l \geqslant 1$ existem $n_l \geqslant 1$ e $C_{n_l} \in \bigvee_{j=0}^{n_l} \widehat{T^{-j}\mathcal{P}}$ tais que $\mu(A \triangle C_{n_l}) < 2^{-l}$. Mas, observe que, como $H\Big(\bigvee_{j=0}^{n_l-1} T^{-j}\mathcal{P}\Big) \leqslant (n_l - 1)H(\mathcal{P}) < \infty$ (pela sub-aditividade de H) e $H\Big(T, \bigvee_{j=0}^{n_l-1} T^{-j}\mathcal{P}\Big) = H(T,\mathcal{P}) = 0$ (pela

Proposição 3.18), segue que $\bigvee_{j=0}^{n_l-1} T^{-j}\mathcal{P} \subset \mathcal{D}(T)$ e, portanto, $\bigvee_{j=0}^{n_l-1} T^{-j}\mathcal{P} \subset \mathcal{P}(T)$ para cada $l \geqslant 1$. Consequentemente, vemos que:

$$\begin{aligned}
C_{n_l} \in \bigvee_{j=1}^{n_l} T^{-j}\mathcal{P} &= T^{-1}\Big(\bigvee_{j=0}^{n_l-1} T^{-j}\mathcal{P}\Big)\\
&\subset T^{-1}\mathcal{D}(T)\\
&\subset T^{-1}\mathcal{P}(T), \text{ para cada } l \geqslant 1.
\end{aligned}$$

Daí, como $T^{-1}\mathcal{P}(T)$ é uma σ-álgebra, $C_{n_l} \in T^{-1}\mathcal{P}(T)$ e $\mu(A\triangle C_{n_l}) < 2^{-l}$ para todo $l \geqslant 1$, fazendo $l \to \infty$ obtemos $A = \liminf C_{n_l} \in T^{-1}\mathcal{P}(T)$. Logo, como $A \in \mathcal{D}(T)$ foi tomado arbitrariamente, segue que $\mathcal{D}(T) \subset T^{-1}\mathcal{P}(T)$, o que implica $\mathcal{P}(T) \subset T^{-1}\mathcal{P}(T)$. Isso conclui a demonstração.

□

Proposição 4.20 . $\mathcal{P}(T^k) = \mathcal{P}(T)$ para cada $k \in \mathbb{Z}$.

Demonstração. Primeiro observe que, como $T^{-k} = (T^{-1})^k$, é suficiente mostrar que $\mathcal{P}(T^k) = \mathcal{P}(T)$ para cada $k \geqslant 1$ e $\mathcal{P}(T^{-1}) = \mathcal{P}(T)$. Suponha $k \geqslant 1$ e tome $A \in \mathcal{D}(T^k)$ arbitrário. Por definição existe uma partição $\mathcal{P}$ tal que $A \in \mathcal{P}$, $H(\mathcal{P}) < \infty$ e $H(T^k,\mathcal{P}) = 0$. Mas, pela Observação 3.33 tem-se:

$$H(T^k,\mathcal{P}) = 0 \iff \widehat{\mathcal{P}} \subset \bigvee_{j=1}^{\infty} \widehat{(T^k)^{-j}\mathcal{P}}.$$

Daí, como $\bigvee_{j=1}^{\infty} \widehat{(T^k)^{-j}\mathcal{P}} \subset \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}}$, obtemos $\widehat{\mathcal{P}} \subset \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}}$ donde, novamente pela Observação 3.33, tem-se $H(T,\mathcal{P}) = 0$. Isso implica que $A \in \mathcal{D}(T)$, donde $A \in \mathcal{P}(T)$. Daí, como $A \in \mathcal{D}(T^k)$ foi tomado arbitrariamente obtemos $\mathcal{D}(T^k) \subset \mathcal{P}(T)$, o que implica $\mathcal{P}(T^k) \subset \mathcal{P}(T)$.

Reciprocamente, tome $A \in \mathcal{D}(T)$ arbitrário. Por definição existe uma partição $\mathcal{P}$ tal que $A \in \mathcal{P}$, $H(\mathcal{P}) < \infty$ e $H(T,\mathcal{P}) = 0$. Pela Proposição 3.19 obtemos:

$$\begin{aligned}
H(T^k,\mathcal{P}) &\leqslant H\Big(T^k, \bigvee_{j=0}^{k-1} T^{-j}\mathcal{P}\Big)\\
&= kH(T,\mathcal{P}) = 0,
\end{aligned}$$

donde $A \in \mathcal{D}(T^k) \subset \mathcal{P}(T^k)$. Daí, como $A \in \mathcal{D}(T)$ foi tomado arbitrariamente, segue que $\mathcal{D}(T) \subset \mathcal{P}(T^k)$, o que implica $\mathcal{P}(T) \subset \mathcal{P}(T^k)$. Por fim, temos $\mathcal{P}(T^{-1}) = \mathcal{P}(T)$ pois, para cada partição $\mathcal{P}$ com entropia finita, tem-se $H(T^{-1},\mathcal{P}) = H(T,\mathcal{P})$. Isso conclui a demonstração. □

O próximo resultado formaliza a ideia de que a σ-álgebra de Pinsker $\mathcal{P}(T)$ corresponde à "parte determinística maximal" do sistema dinâmico $(X,\mathcal{X},\mu,T)$.

Teorema 4.21 . Seja $\mathcal{P}$ partição de X com entropia finita. Então, $\widehat{\mathcal{P}} \subset \mathcal{P}(T)$ se, e somente se, $H(T,\mathcal{P}) = 0$.

Demonstração. Suponha $\widehat{\mathcal{P}} \subset \mathcal{P}(T)$. Como $\mathcal{D}(T)$ é uma álgebra que gera $\mathcal{P}(T)$, pela Proposição 1.32 existe uma sequência de partições finitas $(\mathcal{P}_n)_{n=1}^{\infty}$ tal que $\mathcal{P}_n \subset \mathcal{D}(T)$ para todo $n \geqslant 1$ e $\widehat{\mathcal{P}_n} \uparrow \mathcal{P}(T)$.

Pelo Lema 3.15, para todo $n \geqslant 1$ temos:

$$\begin{aligned} H(T,\mathcal{P}) &\leqslant H(T,\mathcal{P}_n) + H(\mathcal{P}|\mathcal{P}_n) \\ &= H(T,\mathcal{P}_n) + H(\mathcal{P}|\widehat{\mathcal{P}_n}) \\ &= H(\mathcal{P}|\widehat{\mathcal{P}_n}), \end{aligned} \tag{40}$$

pois $H(T,\mathcal{P}_n) = 0$, uma vez que $\mathcal{P}_n \subset \mathcal{D}(T)$ para todo $n \geqslant 1$.

Além disso, como $\widehat{\mathcal{P}_n} \uparrow \mathcal{P}(T)$, pelo Teorema 3.31 tem-se:

$$\lim_{n\to\infty} H(\mathcal{P}|\widehat{\mathcal{P}_n}) = H(\mathcal{P}|\mathcal{P}(T)) = 0,$$

onde a última igualdade decorre do Teorema 3.29, uma vez que $\widehat{\mathcal{P}} \subset \mathcal{P}(T)$ por hipótese.

Portanto, fazendo $n \to \infty$ na relação (40) obtemos $H(T,\mathcal{P}) = 0$.

Reciprocamente, se $\mathcal{P}$ é uma partição com entropia finita tal que $H(T,\mathcal{P}) = 0$, então $\mathcal{P} \subset \mathcal{D}(T)$, donde $\widehat{\mathcal{P}} \subset \mathcal{P}(T)$. Isso conclui a demonstração.

$\square$

Lema 4.22 . Sejam $\mathcal{P}$ e $\mathcal{Q}$ partições de X com entropia finita tais que $\mathcal{P} \prec \mathcal{Q}$ ou $\mathcal{Q} \prec \mathcal{P}$ e $\mathcal{A} \subset \mathcal{X}$ uma sub-σ-álgebra tal que $T^{-1}\mathcal{A} = \mathcal{A}$. Então,

$$\frac{1}{n}H\Big(\bigvee_{i=0}^{n} T^i\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}} \vee \mathcal{A}\Big) \to H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}} \vee \mathcal{A}\Big) \text{ quando } n \to \infty.$$

Em particular, se $\mathcal{A} := \{\varnothing, X\}$ e $\mathcal{P} = \mathcal{Q}$ é tal que $H(\mathcal{P}) < \infty$, então:

$$\frac{1}{n}H\Big(\bigvee_{i=0}^{n} T^i\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\Big) \to H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\Big) \text{ quando } n \to \infty.$$

Demonstração. Suponha $\mathcal{Q} \prec \mathcal{P}$. Observe que

$$\begin{aligned} T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}} \vee \bigvee_{i=0}^{n-1}\widehat{T^i\mathcal{P}}\Big) \vee \mathcal{A} &\subset T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}} \vee \bigvee_{i=0}^{n-1}\widehat{T^i\mathcal{P}}\Big) \vee \mathcal{A} \\ &= \bigvee_{j=1}^{\infty}\widehat{T^{-j-n}\mathcal{P}} \vee \bigvee_{i=0}^{n-1}\widehat{T^{i-n}\mathcal{P}} \vee \mathcal{A} \\ &= \bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}} \vee \mathcal{A}. \end{aligned}$$

Afirmação 1: Para cada $n \geqslant 1$, tem-se:

$$T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\subset T^{-(n+1)}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}. \tag{41}$$

Realmente, observe que, para cada $n \geqslant 1$ temos:

$$\begin{aligned}
T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A} &= T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\Big)\vee T^{-n}\Big(\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\\
&= T^{-(n+1)}\Big(\bigvee_{j=0}^{\infty}\widehat{T^{-j}\mathcal{Q}}\Big)\vee T^{-(n+1)}\Big(\bigvee_{i=1}^{n}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\\
&\subset T^{-(n+1)}\Big(\bigvee_{j=0}^{\infty}\widehat{T^{-j}\mathcal{Q}}\Big)\vee T^{-(n+1)}\Big(\bigvee_{i=0}^{n}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\\
&= T^{-(n+1)}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\Big)\vee\widehat{T^{-(n+1)}\mathcal{Q}}\vee T^{-(n+1)}\Big(\bigvee_{i=0}^{n}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\\
&\subset T^{-(n+1)}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\Big)\vee\widehat{T^{-(n+1)}\mathcal{P}}\vee T^{-(n+1)}\Big(\bigvee_{i=0}^{n}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\\
&= T^{-(n+1)}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A},
\end{aligned}$$

onde na última inclusão utilizamos a hipótese $\mathcal{Q}\prec\mathcal{P}$. Isso prova a Afirmação 1.

Afirmação 2:

$$\begin{gathered}
\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee\mathcal{A}\text{ é a menor }\sigma\text{-álgebra que contém}\\
T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\text{ para todo }n\geqslant 1.
\end{gathered}\tag{42}$$

De fato, tome $A\in\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}$ arbitrário. Para cada $l\geqslant 1$ existe $m_l\geqslant 1$ tal que $\mu(A\triangle C_l)<2^{-l}$ para algum $C_l\in\bigvee_{j=1}^{m_l}\widehat{T^{-j}\mathcal{P}}$.

Note que $C_l\in\bigvee_{j=1}^{m_l}\widehat{T^{-j}\mathcal{P}}=T^{-m_l}\Big(\bigvee_{i=0}^{m_l-1}\widehat{T^{i}\mathcal{P}}\Big)\subset T^{-m_l}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{m_l-1}\widehat{T^{i}\mathcal{P}}\Big)$ para cada $l\geqslant 1$.

Daí, como $\mu(A\triangle C_l)<2^{-l}$ para todo $l\geqslant 1$ implica que $A=\liminf C_l$ (Teorema 1.16), vemos que A pertence à menor σ-álgebra que contém $T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)$ para todo $n\geqslant 1$, o que prova a Afirmação 2.

Consequentemente, pelas relações (41) e (42) obtém-se:

$$T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\uparrow\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee\mathcal{A}. \tag{43}$$

Logo, pelo Teorema 3.31 obtemos:

$$H\Big(\mathcal{P}|T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\Big)\to H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee\mathcal{A}\Big). \tag{44}$$

Além disso, usando repetidas vezes o item (i) da Proposição 3.28 obtemos:

$$\begin{aligned}
\frac{1}{n}H\Big(\bigvee_{i=0}^{n}T^{i}\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big) &= \frac{1}{n}H\Big(\mathcal{P}\vee T\mathcal{P}\vee\ldots\vee T^{n}\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\\
&= \frac{1}{n}H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)+\frac{1}{n}H\Big(T\mathcal{P}|\widehat{\mathcal{P}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\\
&+\ldots+\frac{1}{n}H(T^{n}\mathcal{P}|\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\\
&= \frac{1}{n}H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)+\frac{1}{n}H\Big(\mathcal{P}|T^{-1}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\widehat{\mathcal{P}}\Big)\vee\mathcal{A}\Big)\\
&+\ldots+\frac{1}{n}H\Big(\mathcal{P}|T^{-(n-1)}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-2}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\Big)\\
&+\frac{1}{n}H\Big(\mathcal{P}|T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\Big)\\
&= \frac{1}{n}H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\\
&+\frac{1}{n}\sum_{k=1}^{n-1}H\Big(\mathcal{P}|T^{-k}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{k-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\Big)\\
&+\frac{1}{n}H\Big(\mathcal{P}|T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\Big)
\end{aligned} \tag{45}$$

Daí, como

$$0\leqslant\frac{1}{n}H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\leqslant\frac{1}{n}H(\mathcal{P})\to 0,$$

$$0\leqslant\frac{1}{n}H\Big(\mathcal{P}|T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\Big)\leqslant\frac{1}{n}H(\mathcal{P})\to 0$$

e, além disso, pela Proposição 1.50 temos

$$\frac{1}{n}\sum_{k=1}^{n-1}H\Big(\mathcal{P}|T^{-k}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n-1}\widehat{T^{i}\mathcal{P}}\Big)\vee\mathcal{A}\Big)\to H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee\mathcal{A}\Big),$$

fazendo $n \to \infty$ em (45) obtemos:

$$\frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{P}\Big) \to H\Big(\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \mathcal{A}\Big), \tag{46}$$

o que prova o resultado no caso $\mathcal{Q} \prec \mathcal{P}$.

Agora, suponha $\mathcal{P} \prec \mathcal{Q}$.

Como $\mathcal{P} \prec \mathcal{P}$, pela parte já demonstrada temos

$$\frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \mathcal{A}\Big) \to H\Big(\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \mathcal{A}\Big).$$

Daí, como $\mathcal{P} \prec \mathcal{Q}$, temos:

$$\lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big) \leqslant \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}}\vee\mathcal{A}\Big) = H\Big(\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}}\vee\mathcal{A}\Big). \tag{47}$$

Por outro lado, como a relação $\mathcal{P} \prec \mathcal{Q}$ implica

$$\bigvee_{i=0}^{n} T^i\mathcal{P} \vee \bigvee_{i=0}^{n} T^i\mathcal{Q} = \bigvee_{i=0}^{n} T^i\mathcal{Q},$$

pelo item (i) da Proposição 3.28 obtém-se:

$$\begin{aligned}
\frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) &= \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{Q}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) \\
&- \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{Q}| \bigvee_{i=0}^{n} \widehat{T^{i}\mathcal{P}} \vee \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big)
\end{aligned}$$

Daí, tomando o limite quando $n \to \infty$ obtemos:

$$\begin{aligned}
\lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{Q}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) \\
&- \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{Q}| \bigvee_{i=0}^{n} \widehat{T^{i}\mathcal{P}} \vee \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) \\
&\geqslant \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{Q}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) \\
&- \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{Q}| \bigvee_{i=0}^{n} \widehat{T^{i}\mathcal{P}} \vee \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \mathcal{A}\Big) \\
&= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{Q}| \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \mathcal{A}\Big) \\
&- \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{Q}| \bigvee_{i=0}^{n} \widehat{T^{i}\mathcal{P}} \vee \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \mathcal{A}\Big)
\end{aligned}$$

$$= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{P} | \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \mathcal{A}\Big)$$

$$= H\Big(\mathcal{P} | \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \mathcal{A}\Big), \tag{48}$$

onde na última igualdade utilizamos a parte já demonstrada (pois $\mathcal{P} \prec \mathcal{P}$), na penúltima igualdade utilizamos o item (i) da Proposição 3.28, na antepenúltima igualdade utilizamos utilizamos a parte já demonstrada ($\mathcal{P} \prec \mathcal{Q}$ por hipótese) e, na desigualdade, utilizamos a hipótese $\mathcal{P} \prec \mathcal{Q}$. Portanto, por (47) e (48) obtemos o resultado.

□

O próximo resultado estende o Lema 4.4 e será usado extensivamente durante a demonstração do Teorema de Rokhlin-Sinai.

Lema 4.23 . Sejam $\mathcal{P}$,$\mathcal{Q}$ e $\mathcal{R}$ partições de X com entropia finita tais que $\mathcal{P} \prec \mathcal{Q}$ e $\mathcal{A}$ uma sub-σ-álgebra tal que $T^{-1}\mathcal{A} = \mathcal{A}$. Então,

$$H\Big(\mathcal{P} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{Q}} \vee T^{-n}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{R}}\Big) \vee \mathcal{A}\Big) \to H\Big(\mathcal{P} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{Q}} \vee \mathcal{A}\Big) \text{ quando } n \to \infty.$$

Em particular, se $\mathcal{Q} = \mathcal{P}$ e $\mathcal{A} = \{\varnothing, X\}$, então

$$H\Big(\mathcal{P} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} \vee T^{-n}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{R}}\Big)\Big) \to H\Big(\mathcal{P} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}\Big) = H(T,\mathcal{P}) \text{ quando } n \to \infty.$$

Demonstração. Como $\mathcal{P} \prec \mathcal{Q}$, temos $\mathcal{P} \vee \mathcal{Q} = \mathcal{Q}$, donde pelo item (i) da Proposição 3.28 temos:

$$\begin{aligned} H\Big(\mathcal{P} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{Q}} \vee T^{-n}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{R}}\Big) \vee \mathcal{A}\Big) &= H\Big(\mathcal{Q} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{Q}} \vee T^{-n}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{R}}\Big) \vee \mathcal{A}\Big) \\ &\quad - H\Big(\mathcal{Q} | \widehat{\mathcal{P}} \vee \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{Q}} \vee T^{-n}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{R}}\Big) \vee \mathcal{A}\Big) \end{aligned} \tag{49}$$

Agora, aplicando repetidas vezes o item (i) da Proposição 3.28 tem-se, para todo $n \geqslant 1$:

$$\begin{aligned} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{Q} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}(\mathcal{Q} \vee \mathcal{R})} \vee \mathcal{A}\Big) &= H\Big(\mathcal{Q} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}(\mathcal{Q} \vee \mathcal{R})} \vee \mathcal{A}\Big) \\ &\quad + \sum_{k=1}^{n-1} H\Big(T^k\mathcal{Q} | T^k\Big(\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{Q}}\Big) \vee \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{R}} \vee \mathcal{A}\Big) \\ &\quad + H\Big(T^n\mathcal{Q} | T^n\Big(\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{Q}}\Big) \vee \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{R}} \vee \mathcal{A}\Big) \\ &= H\Big(\mathcal{Q} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}(\mathcal{Q} \vee \mathcal{R})} \vee \mathcal{A}\Big) \end{aligned}$$

$$+\sum_{k=1}^{n-1} H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}} \vee T^{-k}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{R}}\Big) \vee \mathcal{A}\Big)$$
$$+ H\Big(\mathcal{Q}|\Big(\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}}\Big) \vee T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{R}}\Big) \vee \mathcal{A}\Big). \quad (50)$$

Como $\mathcal{Q} \prec \mathcal{Q} \vee \mathcal{R}$, o Lema 4.22 nos dá:

$$H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}} \vee \mathcal{A}\Big) = \lim_{n\to\infty}\frac{1}{n}H\Big(\bigvee_{i=0}^{n} T^{i}\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}(\mathcal{Q}\vee\mathcal{R})} \vee \mathcal{A}\Big). \quad (51)$$

Agora, como $H(\mathcal{Q}) < \infty$, temos

$$\lim_{n\to\infty}\frac{1}{n}H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}(Q\vee\mathcal{R})} \vee \mathcal{A}\Big) = 0, \quad (52)$$

$$\lim_{n\to\infty}\frac{1}{n}H\Big(\mathcal{Q}|\Big(\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}}\Big) \vee T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{R}}\Big) \vee \mathcal{A}\Big) = 0 \quad (53)$$

Para simplificar a notação denote $\widehat{\mathcal{R}}^{-} := \bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{R}}$. Pela Proposição 1.50, também temos

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1} H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}} \vee T^{-k}(\widehat{\mathcal{R}}^{-}) \vee \mathcal{A}\Big) = \lim_{n\to\infty} H\Big(\mathcal{Q}|\Big(\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}}\Big) \vee T^{-n}(\widehat{\mathcal{R}}^{-}) \vee \mathcal{A}\Big). \quad (54)$$

Daí, pelas relações (49), (50), (51), (52), (53) e (54) obtemos:

$$\begin{aligned}
\lim_{n\to\infty} H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}} \vee T^{-n}(\widehat{\mathcal{R}}^{-}) \vee \mathcal{A}\Big) &= \lim_{n\to\infty} H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}} \vee T^{-n}(\widehat{\mathcal{R}}^{-}) \vee \mathcal{A}\Big) \\
&\quad - \lim_{n\to\infty} H\Big(\mathcal{Q}|\widehat{\mathcal{P}} \vee \bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}} \vee T^{-n}(\widehat{\mathcal{R}}^{-}) \vee \mathcal{A}\Big) \\
&= H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}} \vee \mathcal{A}\Big) \\
&\quad - \lim_{n\to\infty} H\Big(\mathcal{Q}|\widehat{\mathcal{P}} \vee \bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}} \vee T^{-n}(\widehat{\mathcal{R}}^{-}) \vee \mathcal{A}\Big) \\
&\geqslant H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}} \vee \mathcal{A}\Big) - H\Big(\mathcal{Q}|\widehat{\mathcal{P}} \vee \bigvee_{l=1}^{\infty}\widehat{T^{-l}\mathcal{Q}} \vee \mathcal{A}\Big) \\
&= H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big),
\end{aligned} \quad (55)$$

onde na última igualdade usamos o fato de que $\mathcal{P} \vee \mathcal{Q} = \mathcal{Q}$, uma vez que temos $\mathcal{P} \prec \mathcal{Q}$ por hipótese.

Por outro lado é claro que, para todo $n \geqslant 1$, temos

$$H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}T^{-i}\widehat{\mathcal{Q}}\vee T^{-n}\Big(\bigvee_{j=1}^{\infty}T^{-j}\widehat{\mathcal{R}}\Big)\vee\mathcal{A}\Big)\leqslant H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}T^{-i}\widehat{\mathcal{Q}}\vee\mathcal{A}\Big). \tag{56}$$

Consequentemente, pelas relações (55) e (56) obtemos o resultado.

□

Teorema 4.24 . Sejam $\mathcal{P}$ e $\mathcal{Q}$ partições de X com entropia finita e $\mathcal{A}\subset\mathcal{X}$ uma sub-σ-álgebra tal que $T^{-1}\mathcal{A}=\mathcal{A}$. Então,

$$H\Big(\mathcal{P}\vee\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}(\mathcal{P}\vee\mathcal{Q})}\vee\mathcal{A}\Big)=H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}}\vee\mathcal{A}\Big)+H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\vee\bigvee_{j=-\infty}^{\infty}\widehat{T^{j}\mathcal{Q}}\vee\mathcal{A}\Big).$$

Em particular, se $\mathcal{A}=\{\varnothing,X\}$, então

$$H\Big(\mathcal{P}\vee\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}(\mathcal{P}\vee\mathcal{Q})}\Big)=H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}}\Big)+H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\vee\bigvee_{j=-\infty}^{\infty}\widehat{T^{j}\mathcal{Q}}\Big).$$

Demonstração. Para cada $n\geqslant 1$ temos:

$$\begin{aligned}
&\frac{1}{n}H\Big(\bigvee_{i=0}^{n}T^{i}\mathcal{P}\vee\bigvee_{i=0}^{n}T^{i}\mathcal{Q}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\\
&=\frac{1}{n}H\Big(\bigvee_{i=0}^{n}T^{i}\mathcal{Q}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\\
&+\frac{1}{n}H\Big(\bigvee_{i=0}^{n}T^{i}\mathcal{P}|\bigvee_{i=0}^{n}\widehat{T^{i}\mathcal{Q}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\\
&=\frac{1}{n}H\Big(\bigvee_{i=0}^{n}T^{i}\mathcal{Q}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\\
&+\frac{1}{n}H\Big(\mathcal{P}|\bigvee_{i=0}^{n}\widehat{T^{i}\mathcal{Q}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\\
&+\frac{1}{n}\sum_{k=1}^{n-1}H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee T^{-k}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n}\widehat{T^{i}\mathcal{Q}}\Big)\vee\mathcal{A}\Big)\\
&+\frac{1}{n}H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n}\widehat{T^{i}\mathcal{Q}}\Big)\vee\mathcal{A}\Big).
\end{aligned}\tag{57}$$

Agora, observe que

$$\frac{1}{n}H\Big(\mathcal{P}|\bigvee_{i=0}^{n}\widehat{T^{i}\mathcal{Q}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\mathcal{A}\Big)\leqslant\frac{1}{n}H(\mathcal{P})\to 0, \tag{58}$$

$$\frac{1}{n}H\Big(\mathcal{P}|\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{P}}\vee T^{-n}\Big(\bigvee_{j=1}^{\infty}\widehat{T^{-j}\mathcal{Q}}\vee\bigvee_{i=0}^{n}\widehat{T^{i}\mathcal{Q}}\Big)\vee\mathcal{A}\Big)\leqslant\frac{1}{n}H(\mathcal{P})\to 0 \tag{59}$$

e, como $\mathcal{Q} \prec \mathcal{P} \vee \mathcal{Q}$, pelo Lema 4.22 tem-se

$$\frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i \mathcal{Q} | \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) \to H\Big(\mathcal{Q} | \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) \text{ quando } n \to \infty. \quad (60)$$

Além disso note que, para cada $1 \leqslant k \leqslant n-1$, temos:

$$\begin{aligned} T^{-k}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \bigvee_{i=0}^{n} \widehat{T^{i}\mathcal{Q}}\Big) \vee \mathcal{A} &= \bigvee_{j=1}^{\infty} \widehat{T^{-k-j}\mathcal{Q}} \vee \bigvee_{i=0}^{n} \widehat{T^{-k+i}\mathcal{A}} \vee \mathcal{A} \\ &= \bigvee_{j=k-n}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}, \end{aligned}$$

donde

$$T^{-k}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \bigvee_{i=0}^{n} \widehat{T^{i}\mathcal{Q}}\Big) \vee \mathcal{A} \downarrow \bigvee_{j=-\infty}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A} \text{ quando } n \to \infty. \quad (61)$$

Além disso, também vale:

$$\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \bigvee_{j=k-n}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A} \downarrow \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \bigvee_{j=-\infty}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A},$$

donde, pelo Teorema 3.31 tem-se:

$$H\Big(\mathcal{P} | T^{-k}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \bigvee_{i=0}^{n} \widehat{T^{i}\mathcal{Q}}\Big) \vee \mathcal{A}\Big) \to H\Big(\mathcal{P} | \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \bigvee_{j=-\infty}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) \text{ quando } n \to \infty. \quad (62)$$

Para simplificar a notação, denote $\widehat{\mathcal{P}^- \vee \mathcal{Q}^-} := \bigvee_{i=1}^{\infty} T^{-i}(\widehat{\mathcal{P} \vee \mathcal{Q}})$.

Como $\mathcal{Q} \prec \mathcal{P} \vee \mathcal{Q}$, o Lema 4.22 juntamente com as relações (57), (58), (59), (60), (61), (62) nos dão:

$$\begin{aligned} H\Big(\mathcal{P} \vee \mathcal{Q} | \widehat{\mathcal{P}^- \vee \mathcal{Q}^-} \vee \mathcal{A}\Big) &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i\mathcal{P} \vee \bigvee_{i=0}^{n} T^i \mathcal{Q} | \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) \\ &= \lim_{n\to\infty} \frac{1}{n} H\Big(\bigvee_{i=0}^{n} T^i \mathcal{Q} | \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) \\ &+ \lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n-1} H\Big(\mathcal{P} | \bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{P}} \vee T^{-k}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}} \vee \bigvee_{i=0}^{n} \widehat{T^{i}\mathcal{Q}}\Big) \vee \mathcal{A}\Big) \\ &= H\Big(\mathcal{Q} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{Q}} \vee \mathcal{A}\Big) \\ &+ H\Big(\mathcal{P} | \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} \vee \bigvee_{j=-\infty}^{\infty} \widehat{T^{j}\mathcal{Q}} \vee \mathcal{A}\Big), \end{aligned}$$

onde na última igualdade utilizamos a Proposição 1.50. Isso conclui a demonstração do resultado.

$\square$

O próximo resultado nos dá uma descrição explícita da σ-álgebra de Pinsker $\mathcal{P}(T)$.

Teorema 4.25 .

$$\mathcal{P}(T) = \bigvee_{H(\mathcal{P})<\infty} \Big(\bigcap_{n=1}^{\infty} T^{-n}\Big(\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}\Big)\Big).$$

Ou seja, $\mathcal{P}(T)$ é a menor σ-álgebra que contém o "passado remoto" de todos os processos $(T,\mathcal{P})$ oriundos de partições com entropia finita.

Demonstração. Pela Proposição 1.31 existe uma sequência $(\mathcal{P}_k)_{k=1}^{\infty}$ de partições finitas de X tal que $\widehat{\mathcal{P}_k} \uparrow \mathcal{P}(T)$. Como $\widehat{\mathcal{P}_k} \subset \mathcal{P}(T)$ para todo $k \geqslant 1$, pelo Teorema 4.21 temos que $H(T,\mathcal{P}_k) = 0$ para todo $k \geqslant 1$ e, portanto, pela Observação 3.33 juntamente com o mesmo argumento contido na demonstração do Teorema 4.5 segue que $\widehat{\mathcal{P}_k} \subset \bigcap_{n=1}^{\infty} T^{-n}\Big(\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}_k}\Big)$ para todo $k \geqslant 1$.

Daí,

$$\mathcal{P}(T) = \bigvee_{k=1}^{\infty} \widehat{\mathcal{P}_k} \subset \bigvee_{k=1}^{\infty} \Big(\bigcap_{n=1}^{\infty} T^{-n}\Big(\bigvee_{i=1}^{\infty} T^{-i}\widehat{\mathcal{P}_k}\Big)\Big) \subset \bigvee_{H(\mathcal{P})<\infty} \Big(\bigcap_{n=1}^{\infty} T^{-n}\Big(\bigvee_{i=1}^{\infty} T^{-i}\widehat{\mathcal{P}}\Big)\Big).$$

Reciprocamente, sejam $\mathcal{P}$ e $\mathcal{Q}$ partições arbitrárias de X com entropia finita tais que $\widehat{\mathcal{Q}} \subset \bigcap_{n=1}^{\infty} T^{-n}\Big(\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}\Big)$.

Afirmação: $\bigvee_{i=-\infty}^{\infty} \widehat{T^{-i}\mathcal{Q}} \subset \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}$.

Em particular, tem-se $\bigvee_{i=0}^{\infty} \widehat{T^{-i}\mathcal{Q}} \subset \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}$.

Realmente, para cada $n \geqslant 1$ temos $\widehat{\mathcal{Q}} \subset T^{-n}\Big(\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}\Big)$, donde $T^{n}\widehat{\mathcal{Q}} \subset \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}$. Daí, $\bigvee_{i=1}^{\infty} \widehat{T^{i}\mathcal{Q}} \subset \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}$.

Além disso, para cada $l \geqslant 1$ temos

$$\widehat{T^{-l}\mathcal{P}} \subset \bigvee_{i=1}^{\infty} \widehat{T^{-l-i}\mathcal{Q}} \subset \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{Q}},$$

donde $\bigvee_{i=-\infty}^{1} \widehat{T^{i}\mathcal{Q}} \subset \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}$

Logo, $\bigvee_{i=-\infty}^{\infty} \widehat{T^{-i}\mathcal{Q}} \subset \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}$, o que prova a afirmação.

Consequentemente,

$$\begin{aligned} H\Big(\mathcal{P}\vee\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}(\mathcal{P}\vee\mathcal{Q})}\Big) &= H\Big(\mathcal{P}\vee\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\Big)\\ &= H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\Big)+H\Big(\mathcal{Q}|\widehat{\mathcal{P}}\vee\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\Big)\\ &= H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\Big), \end{aligned} \tag{63}$$

onde na primeira igualdade usamos o fato que $\bigvee_{i=1}^{\infty}\widehat{T^{-i}(\mathcal{P}\vee\mathcal{Q})}=\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}$ e na terceira usamos o fato $\widehat{\mathcal{Q}}\subset\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}$, ambos contidos na Afirmação acima.

Por outro lado, pelo caso particular do Lema 4.24 temos:

$$\begin{aligned} H\Big(\mathcal{P}\vee\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}(\mathcal{P}\vee\mathcal{Q})}\Big) &= H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}}\Big)+H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\vee\bigvee_{j=-\infty}^{\infty}\widehat{T^{j}\mathcal{Q}}\Big)\\ &= H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}}\Big)+H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\Big), \end{aligned} \tag{64}$$

onde na última igualdade usamos a Afirmação acima.

Portanto, das relações (63) e (64) segue que $H\Big(\mathcal{Q}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{Q}}\Big)=0$.

Daí, pelo Teorema 4.21 temos $\widehat{\mathcal{Q}}\subset\mathcal{P}(T)$. Portanto, como $\mathcal{P}$ e $\mathcal{Q}$ são partições arbitrárias de X com entropia finita tais que $\widehat{\mathcal{Q}}\subset\bigcap_{n=1}^{\infty}T^{-n}\Big(\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\Big)$, obtemos

$$\bigvee_{H(\mathcal{P})<\infty}\Big(\bigcap_{n=1}^{\infty}T^{-n}\Big(\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\Big)\Big)\subset\mathcal{P}(T).$$

Isso conclui a demonstração.

□

Tendo em vista o Teorema 4.25, o próximo resultado nos fornece uma extensão do Lema 4.4.

Teorema 4.26 . Sejam $\mathcal{P}$ uma partição de X com entropia finita e $\mathcal{A}\subset\mathcal{X}$ uma sub-σ-álgebra tal que $T^{-1}\mathcal{A}=\mathcal{A}$. Então,

$$H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\vee\mathcal{P}(T)\vee\mathcal{A}\Big)=H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\vee\mathcal{A}\Big).$$

Em particular, se $\mathcal{A}=\{\varnothing,X\}$, então

$$H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\vee\mathcal{P}(T)\Big)=H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty}\widehat{T^{-i}\mathcal{P}}\Big).$$

Demonstração. Seja $\mathcal{Q}$ uma partição arbitrária de X com entropia finita tal que $\widehat{\mathcal{Q}} \subset \mathcal{P}(T)$. Pelo Teorema 4.21 tem-se $H(T,\mathcal{Q}) = 0$ donde, pela Observação 3.33, tem-se $\widehat{\mathcal{Q}} \subset \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{Q}}$. Além disso, pelo mesmo argumento da primeira parte da demonstração do Teorema 4.5, temos $\widehat{\mathcal{Q}} \subset \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} \widehat{T^{-j}\mathcal{Q}}$. Agora observe que, como $T^{-n}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}}\Big) \downarrow \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} \widehat{T^{-j}\mathcal{Q}}$, pelo Teorema 3.31, obtém-se:

$$\begin{aligned} H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} \vee \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} \widehat{T^{-j}\mathcal{Q}} \vee \mathcal{A}\Big) &= \lim_{n\to\infty} H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} \vee T^{-n}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}}\Big) \vee \mathcal{A}\Big) \\ &= H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} \vee \mathcal{A}\Big), \end{aligned} \tag{65}$$

onde na última igualdade utilizamos o Lema 4.23.

Agora, pela Proposição 1.31, existe uma sequência $(\mathcal{Q}_{k\geqslant 1})$ de partições finitas de X tal que $\widehat{\mathcal{Q}_k} \uparrow \mathcal{P}(T)$. Pela relação (65) temos:

$$\begin{aligned} H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} \vee \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} \widehat{T^{-j}\mathcal{Q}_k} \vee \mathcal{A}\Big) &= \lim_{n\to\infty} H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} \vee T^{-n}\Big(\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}_k}\Big) \vee \mathcal{A}\Big) \\ &= H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} \vee \mathcal{A}\Big) \end{aligned} \tag{66}$$

para todo $k \geqslant 1$. Além disso, temos $\widehat{\mathcal{Q}_k} \subset \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} \widehat{T^{-j}\mathcal{Q}_k} \subset \mathcal{P}(T)$ para todo $k \geqslant 1$, onde na segunda inclusão utilizamos o Teorema 4.25. Daí, tem-se $\bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} \widehat{T^{-j}\mathcal{Q}_k} \uparrow \mathcal{P}(T)$ donde, fazendo $k \to \infty$ na relação (66), obtemos pelo Teorema 3.31:

$$H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} \vee \mathcal{P}(T) \vee \mathcal{A}\Big) = H\Big(\mathcal{P}|\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}} \vee \mathcal{A}\Big).$$

Isso conclui a demonstração do resultado. □

Suponha que $\mathcal{A} \subset \mathcal{X}$ é uma sub-σ-álgebra tal que $T^{-1}\mathcal{A} \subset \mathcal{A}$.

Denote $\mathcal{A}_{-\infty} := \bigcap_{n\geqslant 1} T^{-n}\mathcal{A}$ e $\mathcal{A}_{\infty} := \bigvee_{n\geqslant 1} T^{n}\mathcal{A}$.

Ou seja, $\mathcal{A}_{-\infty}$ é a maior σ-álgebra contida em todas as σ-álgebras da forma $T^{-n}\mathcal{A}$ e $\mathcal{A}_{\infty}$ é a menor σ-álgebra que contém todas as σ-álgebras da forma $T^{n}\mathcal{A}$ $(n \geqslant 1)$.

Observe que $\mathcal{A}_{-\infty} = \bigcap_{n\geqslant 1} \bigvee_{j\geqslant n} T^{-j}\mathcal{A}$ e, além disso,

$$\mathcal{A}_{-\infty} \subset \ldots \subset T^{-n}\mathcal{A} \subset \ldots \subset \mathcal{A} \subset \ldots \subset T^{n}\mathcal{A} \subset \ldots \subset \mathcal{A}_{\infty} \quad (n \geqslant 1). \tag{67}$$

Teorema 4.27 . Sejam $\mathcal{A} \subset \mathcal{X}$ uma sub-σ-álgebra tal que $T^{-1}\mathcal{A} \subset \mathcal{A}$ e $\mathcal{P}$ uma partição de X com entropia finita tal que $\widehat{\mathcal{P}} \subset \mathcal{A}_{\infty}$. Então,

$$H(\mathcal{P}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) = H(\mathcal{P}|\mathcal{A}_{-\infty}).$$

Demonstração. Suponha inicialmente que $\widehat{\mathcal{P}} \subset \mathcal{A}$. Note que:

$$\begin{aligned} H(\mathcal{P}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) &\geqslant \lim_{p\to\infty} H\Big(\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-jp}\mathcal{P}} \vee \mathcal{P}(T) \vee \mathcal{A}_{-\infty}\Big) \\ &= \lim_{p\to\infty} H\Big(\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{(T^p)^{-j}\mathcal{P}} \vee \mathcal{P}(T) \vee \mathcal{A}_{-\infty}\Big) \\ &= \lim_{p\to\infty} H\Big(\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{(T^p)^{-j}\mathcal{P}} \vee \mathcal{P}(T^p) \vee \mathcal{A}_{-\infty}\Big) \\ &= \lim_{p\to\infty} H\Big(\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{(T^p)^{-j}\mathcal{P}} \vee \mathcal{A}_{-\infty}\Big) \\ &= \lim_{p\to\infty} H\Big(\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-jp}\mathcal{P}} \vee \mathcal{A}_{-\infty}\Big) \\ &\geqslant \lim_{p\to\infty} H\Big(\mathcal{P}| \bigvee_{j=1}^{\infty} \widehat{T^{-jp}\mathcal{A}} \vee \mathcal{A}_{-\infty}\Big) \\ &= H\Big(\mathcal{P}| \bigcap_{p\geqslant 1} \bigvee_{j=1}^{\infty} \widehat{T^{-jp}\mathcal{A}} \vee \mathcal{A}_{-\infty}\Big) \\ &= H(\mathcal{P}|\mathcal{A}_{-\infty}), \end{aligned} \tag{68}$$

onde na segunda igualdade usamos a Proposição 4.20, na terceira igualdade usamos o Teorema 4.26 aplicado à T^p (pois $(T^p)^{-1}\mathcal{A}_{-\infty} = \mathcal{A}_{-\infty}$ para cada $p \geqslant 1$), na segunda desigualdade usamos a hipótese $\widehat{\mathcal{P}} \subset \mathcal{A}$, na penúltima igualdade usamos o Teorema 3.31 e na última igualdade usamos o fato que $\bigcap_{p\geqslant 1} \bigvee_{j=1}^{\infty} \widehat{T^{-jp}\mathcal{P}} = \mathcal{A}_{-\infty}$.

Portanto, o resultado é verdadeiro no caso em que $\widehat{\mathcal{P}} \subset \mathcal{A}$.

Veremos agora que a mesma conclusão é verdadeira no caso em que $\widehat{\mathcal{P}} \subset T^n\mathcal{A}$ para algum $n \geqslant 1$. Realmente, note que

$$\widehat{\mathcal{P}} \subset T^n\mathcal{A} \iff T^{-n}\widehat{\mathcal{P}} \subset \mathcal{A}.$$

Daí, como $T^{-n}\mathcal{P}$ é uma partição com entropia finita, pela parte já demonstrada temos:

$$H(T^{-n}\mathcal{P}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) = H(T^{-n}\mathcal{P}|\mathcal{A}_{-\infty}),$$

o que equivale a dizer

$$H(\mathcal{P}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) = H(\mathcal{P}|\mathcal{A}_{-\infty}),$$

uma vez que $\mathcal{P}(T)$ e $\mathcal{A}_{-\infty}$ são T-invariantes (Proposição 4.19).

Portanto, tem-se

$$H(\mathcal{P}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) = H(\mathcal{P}|\mathcal{A}_{-\infty})$$

sempre que $\widehat{\mathcal{P}} \subset T^n\mathcal{A}$ $(n \geqslant 0)$.

Por fim, suponha que $\mathcal{P}$ é uma partição de X com entropia finita tal que $\widehat{\mathcal{P}} \subset \mathcal{A}_\infty$. Como $T^n\mathcal{A} \uparrow \mathcal{A}_\infty$, pela Proposição 3.23 dado $\varepsilon > 0$ existem $n \in \mathbb{Z}$ e uma partição $\mathcal{Q}$ com entropia finita tais que $\widehat{\mathcal{Q}} \subset T^n\mathcal{A}$ e $H(\mathcal{P}|\widehat{\mathcal{Q}}) + H(\mathcal{Q}|\widehat{\mathcal{P}}) < \varepsilon$.

Logo,

$$\begin{aligned} H(\mathcal{P}|\mathcal{A}_{-\infty}) &\leqslant H(\mathcal{P} \vee \mathcal{Q}|\mathcal{A}_{-\infty}) \\ &= H(\mathcal{Q}|\mathcal{A}_{-\infty}) + H(\mathcal{P}|\widehat{\mathcal{Q}} \vee \mathcal{A}_{-\infty}) \\ &\leqslant H(\mathcal{Q}|\mathcal{A}_{-\infty}) + H(\mathcal{P}|\widehat{\mathcal{Q}}) \\ &< H(\mathcal{Q}|\mathcal{A}_{-\infty}) + \varepsilon \\ &= H(\mathcal{Q}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) + \varepsilon \\ &\leqslant H(\mathcal{P} \vee \mathcal{Q}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) + \varepsilon \\ &= H(\mathcal{P}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) + H(\mathcal{Q}|\widehat{\mathcal{P}} \vee \mathcal{P}(T) \vee \mathcal{A}_{-\infty}) + \varepsilon \\ &\leqslant H(\mathcal{P}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) + H(\mathcal{Q}|\widehat{\mathcal{P}}) + \varepsilon \\ &< H(\mathcal{P}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) + 2\varepsilon \\ &\leqslant H(\mathcal{P}|\mathcal{A}_{-\infty}) + 2\varepsilon. \end{aligned} \tag{69}$$

Consequentemente, como $\varepsilon > 0$ é arbitrário obtemos

$$H(\mathcal{P}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) = H(\mathcal{P}|\mathcal{A}_{-\infty}),$$

como queríamos demonstrar. □

Corolário 4.28 . Seja $\mathcal{A} \subset \mathcal{X}$ uma sub-σ-álgebra tal que $T^{-1}\mathcal{A} \subset \mathcal{A}$ e $T^n\mathcal{A} \uparrow \mathcal{X}$.

Então, $\mathcal{P}(T) \subset \mathcal{A}_{-\infty}$.

Demonstração. Seja $\mathcal{P}$ uma partição arbitrária de X com entropia finita. Então, $\widehat{\mathcal{P}} \subset \mathcal{X} = \mathcal{A}_\infty$, donde pelo Teorema 4.27 tem-se:

$$H(\mathcal{P}|\mathcal{P}(T) \vee \mathcal{A}_{-\infty}) = H(\mathcal{P}|\mathcal{A}_{-\infty}).$$

Daí, pelo Teorema 3.29 vemos que $\widehat{\mathcal{P}} \subset \mathcal{P}(T) \vee \mathcal{A}_{-\infty} \iff \widehat{\mathcal{P}} \subset \mathcal{A}_{-\infty}$ para toda $\mathcal{P}$ partição de X com entropia finita. Consequentemente, obtemos $\mathcal{P}(T) \subset \mathcal{A}_{-\infty}$, como desejado. □

Teorema 4.29 . (Rokhlin-Sinai) Seja $(X,\mathcal{X},\mu,T)$ um sistema dinâmico, onde $(X,\mathcal{X},\mu)$ é um espaço de probabilidade separável.

Então, $(X,\mathcal{X},\mu,T)$ é um K-automorfismo se, e somente se, tem entropia completamente positiva.

Ou seja, $(X,\mathcal{X},\mu,T)$ é um K-automorfismo se, e somente se, tem-se $H(T,\mathcal{P}) > 0$ para toda $\mathcal{P}$ partição de X tal que $H(\mathcal{P}) < \infty$.

Demonstração. Suponha que $(X,\mathcal{X},\mu,T)$ é um K-automorfismo. Segue da Observação 4.13 que $H(T,\mathcal{P}) > 0$ para cada $\mathcal{P}$ partição de X com entropia finita, ou seja, $(X,\mathcal{X},\mu,T)$ tem entropia completamente positiva (CPE).

Reciprocamente, suponha que $(X,\mathcal{X},\mu,T)$ tem entropia completamente positiva (CPE). Observe que, pelo Teorema 4.21 isto significa dizer que $\mathcal{P}(T) = \{\varnothing, X\}$. Portanto, o resultado restará demonstrado se provarmos que, dado um sistema dinâmico $(X,\mathcal{X},\mu,T)$ arbitrário, onde $(X,\mathcal{X},\mu)$ é um espaço de probabilidade separável, existe uma sub-σ-álgebra $\mathcal{A} \subset \mathcal{X}$ que satisfaz:

(i) $T^{-1}\mathcal{A} \subset \mathcal{A}$;

(ii) $T^n\mathcal{A} \uparrow \mathcal{X}$, ou seja, $\displaystyle\bigvee_{i=-\infty}^{\infty} T^i\mathcal{A} = \mathcal{X}$;

(iii) $T^{-n}\mathcal{A} \downarrow \mathcal{P}(T)$, ou seja, $\displaystyle\bigcap_{n\geqslant 1}\bigvee_{j\geqslant n} T^{-j}\mathcal{A} = \mathcal{P}(T)$.

Em outras palavras, sempre existe uma sub-σ-álgebra cuja "história completa" descreve todos os eventos e cujo "passado remoto" corresponde à "parte determinística maximal" do sistema dinâmico.

Realmente, seja $(\mathcal{P}_n)_{n\geqslant 1}$ uma sequência de partições finitas tal que $\widehat{\mathcal{P}_n} \uparrow \mathcal{X}$ (Proposição 1.31). A fim de simplificarmos a notação no que segue, denote por $\widehat{\mathcal{R}^-} := \displaystyle\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{R}}$ o "passado" do processo $(T,\mathcal{R})$. Agora, considere $\mathcal{Q}_1 := \mathcal{P}_1$ e $n_1 := 1$. Pelo caso particular do Lema 4.23, existe $n_2 \geqslant n_1$ tal que

$$H(\mathcal{Q}_1|\widehat{\mathcal{Q}_1^-}) - H(\mathcal{Q}_1|\widehat{\mathcal{Q}_1^-} \vee T^{-n_2}\widehat{\mathcal{P}_2^-}) < \frac{1}{2}.$$

Agora, considere $\mathcal{Q}_2 := \mathcal{Q}_1 \vee T^{-n_2}\mathcal{P}_2$. Então, temos $\mathcal{Q}_1 \prec \mathcal{Q}_2$ e $\widehat{\mathcal{Q}_2^-} = \widehat{\mathcal{Q}_1^-} \vee T^{-n_2}\widehat{\mathcal{P}_2^-}$.

Novamente pelo caso particular do Lema 4.23, existe $n_3 \geqslant n_2$ tal que

$$H(\mathcal{Q}_2|\widehat{\mathcal{Q}_2^-}) - H(\mathcal{Q}_2|\widehat{\mathcal{Q}_2^-} \vee T^{-n_3}\widehat{\mathcal{P}_3^-}) < \frac{1}{2^2}.$$

Considere $\mathcal{Q}_3 := \mathcal{Q}_2 \vee T^{-n_3}\mathcal{P}_3$. Então, temos $\mathcal{Q}_2 \prec \mathcal{Q}_3$ e $\widehat{\mathcal{Q}_3^-} = \widehat{\mathcal{Q}_2^-} \vee T^{-n_3}\widehat{\mathcal{P}_3^-}$. Como $\mathcal{Q}_1 \prec \mathcal{Q}_2$, pelo caso geral do Lema 4.23 podemos supor que n_3 é suficientemente grande tal que

$$H(\mathcal{Q}_1|\widehat{\mathcal{Q}_2^-}) - H(\mathcal{Q}_1|\widehat{\mathcal{Q}_3^-}) < \frac{1}{2^2}.$$

Novamente pelo caso particular do Lema 4.23, existe $n_4 \geqslant n_3$ tal que

$$H(\mathcal{Q}_3|\widehat{\mathcal{Q}_3^-}) - H(\mathcal{Q}_3|\widehat{\mathcal{Q}_3^-} \vee T^{-n_4}\widehat{\mathcal{P}_4^-}) < \frac{1}{2^3}.$$

Considere $\mathcal{Q}_4 := \mathcal{Q}_3 \vee T^{-n_4}\mathcal{P}_4$. Então, temos $\mathcal{Q}_3 \prec \mathcal{Q}_4$ e $\widehat{\mathcal{Q}_4^-} = \widehat{\mathcal{Q}_3^-} \vee T^{-n_4}\widehat{\mathcal{P}_4^-}$.

Como $\mathcal{Q}_1 \prec \mathcal{Q}_2 \prec \mathcal{Q}_3$, pelo caso geral do Lema 4.23 podemos supor que n_4 é suficientemente grande tal que

$$H(\mathcal{Q}_1|\widehat{\mathcal{Q}_3^-}) - H(\mathcal{Q}_1|\widehat{\mathcal{Q}_4^-}) < \frac{1}{2^3}$$

e também

$$H(\mathcal{Q}_2|\widehat{\mathcal{Q}_3^-}) - H(\mathcal{Q}_2|\widehat{\mathcal{Q}_4^-}) < \frac{1}{2^3}.$$

Daí, temos:

$$H(\mathcal{Q}_1|\widehat{\mathcal{Q}_3^-}) - H(\mathcal{Q}_1|\widehat{\mathcal{Q}_4^-}) < \frac{1}{2^3},$$
$$H(\mathcal{Q}_2|\widehat{\mathcal{Q}_3^-}) - H(\mathcal{Q}_2|\widehat{\mathcal{Q}_4^-}) < \frac{1}{2^3},$$
$$H(\mathcal{Q}_3|\widehat{\mathcal{Q}_3^-}) - H(\mathcal{Q}_3|\widehat{\mathcal{Q}_4^-}) < \frac{1}{2^3}.$$

Continuando dessa maneira obtemos uma sequência crescente de números naturais $n_1 \leqslant n_2 \leqslant \ldots$ e uma sequência crescente de partições finitas $\mathcal{Q}_1 \prec \mathcal{Q}_2 \prec \ldots$ tais que: $\mathcal{Q}_1 := \mathcal{P}_1$, $\mathcal{Q}_p := \mathcal{Q}_{p-1} \vee T^{-n_p}\mathcal{P}_p$ $(p \geqslant 1)$ e

$$H(\mathcal{Q}_1|\widehat{\mathcal{Q}_q^-}) - H(\mathcal{Q}_1|\widehat{\mathcal{Q}_{q+1}^-}) < \frac{1}{2^q}\ (q \geqslant 1),$$
$$H(\mathcal{Q}_2|\widehat{\mathcal{Q}_q^-}) - H(\mathcal{Q}_2|\widehat{\mathcal{Q}_{q+1}^-}) < \frac{1}{2^q}\ (q \geqslant 2)$$
$$\vdots$$
$$H(\mathcal{Q}_p|\widehat{\mathcal{Q}_q^-}) - H(\mathcal{Q}_p|\widehat{\mathcal{Q}_{q+1}^-}) < \frac{1}{2^q}\ (q \geqslant p).$$

Consequentemente, temos

$$H(\mathcal{Q}_p|\widehat{\mathcal{Q}_q^-}) - H(\mathcal{Q}_p|\widehat{\mathcal{Q}_{q+1}^-}) < \frac{1}{p}.\frac{1}{2^{q-p}}\ (p \leqslant q).$$

Daí, se fixamos $p \geqslant 1$ arbitrariamente e tomamos qualquer $n \geqslant p$, ao somarmos sobre todos os q's tais que $p \leqslant q \leqslant n-1$ obtemos:

$$\begin{aligned} H(\mathcal{Q}_p|\widehat{\mathcal{Q}_p^-}) - H(\mathcal{Q}_p|\widehat{\mathcal{Q}_n^-}) &= H(\mathcal{Q}_p|\widehat{\mathcal{Q}_p^-}) - H(\mathcal{Q}_p|\widehat{\mathcal{Q}_{p+1}^-}) \\ &+ H(\mathcal{Q}_p|\widehat{\mathcal{Q}_{p+1}^-}) - H(\mathcal{Q}_p|\widehat{\mathcal{Q}_{p+2}^-}) \\ &+ \ldots \\ &\vdots \\ &+ H(\mathcal{Q}_p|\widehat{\mathcal{Q}_{n-1}^-}) - H(\mathcal{Q}_p|\widehat{\mathcal{Q}_n^-}) \\ &< \frac{1}{p}\Big(1 + \frac{1}{2} + \ldots + \frac{1}{2^{n-p-1}}\Big) \\ &< \frac{2}{p}. \end{aligned}$$

Portanto, vemos que

$$H(\mathcal{Q}_p|\widehat{\mathcal{Q}_p^-}) - H(\mathcal{Q}_p|\widehat{\mathcal{Q}_n^-}) < \frac{2}{p}, \text{ sempre que } n \geqslant p. \tag{70}$$

Agora, considere $\mathcal{C} := \bigvee_{n=1}^{\infty} \widehat{\mathcal{Q}_n}$. Observe que, como $\mathcal{Q}_1 \prec \mathcal{Q}_2 \prec \ldots$, temos $\widehat{\mathcal{Q}_n} \uparrow \mathcal{C}$, o que implica $\widehat{\mathcal{Q}_n^-} \uparrow \mathcal{C}^-$, onde $\mathcal{C}^- := \bigvee_{j=1}^{\infty} T^{-j}\mathcal{C}$ denota o "passado" da sub-σ-álgebra $\mathcal{C}$.

Considere $\mathcal{A} := \mathcal{C}^-$.

Afirmação 1: $T^{-1}\mathcal{A} \subset \mathcal{A}$.

De fato, observe que

$$\begin{aligned}
T^{-1}\mathcal{A} &= T^{-1}\Big(\bigvee_{n=1}^{\infty} \widehat{\mathcal{Q}_n^-}\Big) \\
&= T^{-1}\Big(\bigvee_{n=1}^{\infty}\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}_n}\Big) \\
&= \bigvee_{n=1}^{\infty}\bigvee_{j=2}^{\infty} \widehat{T^{-j}\mathcal{Q}_n} \\
&\subset \bigvee_{n=1}^{\infty}\bigvee_{j=1}^{\infty} \widehat{T^{-j}\mathcal{Q}_n} \\
&= \mathcal{A},
\end{aligned}$$

o que prova a Afirmação 1.

Afirmação 2: $T^m\mathcal{A} \uparrow \mathcal{X}$.

Com efeito, como $\widehat{\mathcal{Q}_k} \supset T^{-n_k}\widehat{\mathcal{P}_k}$ para cada $k \geqslant 2$, para cada $m \geqslant 1$ temos:

$$\begin{aligned}
T^m\mathcal{A} &= \bigvee_{j=1}^{\infty} T^{-j+m}\Big(\bigvee_{k=1}^{\infty} \widehat{\mathcal{Q}_k}\Big) \\
&= \bigvee_{j=-m+1}^{\infty} T^{-j}\Big(\bigvee_{k=1}^{\infty} \widehat{\mathcal{Q}_k}\Big) \\
&\supset \bigvee_{j=-m+1}^{\infty} T^{-j}\Big(\bigvee_{k=1}^{\infty} \widehat{T^{-n_k}\mathcal{P}_k}\Big) \\
&= \bigvee_{k=1}^{\infty}\bigvee_{j=-m+1+n_k}^{\infty} \widehat{T^{-j}\mathcal{P}_k},
\end{aligned}$$

o que implica

$$\widehat{\mathcal{P}_k} \subset \bigvee_{j=-\infty}^{\infty}\bigvee_{k=1}^{\infty} \widehat{T^{-j}\mathcal{P}_k} \subset \bigvee_{m=1}^{\infty} T^m\mathcal{A} \subset \bigvee_{m=-\infty}^{\infty} T^m\mathcal{A} \subset \mathcal{X},$$

para cada $k \geqslant 2$. Daí como $\widehat{\mathcal{P}_k} \uparrow \mathcal{X}$ quando $k \to \infty$, conclui-se que $\bigvee\limits_{m=-\infty}^{\infty} T^m\mathcal{A} = \mathcal{X}$, ou seja, $T^m\mathcal{A} \uparrow \mathcal{X}$ quando $m \to \infty$. Isso prova a Afirmação 2.

Daí, como $T^{-1}\mathcal{A} \subset \mathcal{A}$ (Afirmação 1) e $T^n\mathcal{A} \uparrow \mathcal{X}$ (Afirmação 2), pelo Corolário 4.28 segue que $\mathcal{P}(T) \subset \bigcap\limits_{n\geqslant 1} T^{-n}\mathcal{A}$.

Portanto, falta apenas mostrar que $\bigcap\limits_{n\geqslant 1} T^{-n}\mathcal{A} \subset \mathcal{P}(T)$. Para tal, seja $\mathcal{P}$ uma partição arbitrária de X com entropia finita tal que $\widehat{\mathcal{P}} \subset \bigcap\limits_{n\geqslant 1} T^{-n}\mathcal{A}$. Note que, como $T^{-1}\mathcal{A} \subset \mathcal{A}$, temos $\bigvee\limits_{j=-\infty}^{\infty} \widehat{T^{-j}\mathcal{P}} \subset \bigcap\limits_{n\geqslant 1} T^{-n}\mathcal{A} \subset \mathcal{A}$. Daí, dados $n \geqslant p$, temos:

$$
\begin{aligned}
H(\mathcal{P}|\widehat{\mathcal{P}^-}) &= H(\mathcal{Q}_p \vee \mathcal{P}|\widehat{\mathcal{Q}_p^-} \vee \widehat{\mathcal{P}^-}) - H\Big(\mathcal{Q}_p|\widehat{\mathcal{Q}_p^-} \vee \bigvee_{j=-\infty}^{\infty} \widehat{T^{-j}\mathcal{P}}\Big) \\
&= H(\mathcal{Q}_p|\widehat{\mathcal{Q}_p^-} \vee \widehat{\mathcal{P}^-}) + H(\mathcal{P}|\widehat{\mathcal{Q}_p} \vee \widehat{\mathcal{Q}_p^-} \vee \widehat{\mathcal{P}^-}) \\
&\quad - H\Big(\mathcal{Q}_p|\widehat{\mathcal{Q}_p^-} \vee \bigvee_{j=-\infty}^{\infty} \widehat{T^{-j}\mathcal{P}}\Big) \\
&\leqslant H(\mathcal{Q}_p|\widehat{\mathcal{Q}_p^-}) - H(\mathcal{Q}_p|\mathcal{A}) + H(\mathcal{P}|\widehat{\mathcal{Q}_p^-}) \\
&= \Big[H(\mathcal{Q}_p|\widehat{\mathcal{Q}_p^-}) - H(\mathcal{Q}_p|\widehat{\mathcal{Q}_n^-})\Big] \\
&\quad + \Big[H(\mathcal{Q}_p|\widehat{\mathcal{Q}_n^-}) - H(\mathcal{Q}_p|\mathcal{A})\Big] + H(\mathcal{P}|\widehat{\mathcal{Q}_p^-}) \\
&< \frac{2}{p} + \Big[H(\mathcal{Q}_p|\widehat{\mathcal{Q}_n^-}) - H(\mathcal{Q}_p|\mathcal{A})\Big] + H(\mathcal{P}|\widehat{\mathcal{Q}_p^-}),
\end{aligned}
\tag{71}
$$

onde na primeira igualdade utilizamos o Teorema 4.24, na segunda igualdade utilizamos o item (i) da Proposição 3.28, na primeira desigualdade utilizamos as inclusões $\widehat{\mathcal{Q}_p^-} \subset \widehat{\mathcal{Q}_p^-} \vee \widehat{\mathcal{P}^-}$ e $\widehat{\mathcal{Q}_p^-} \vee \bigvee_{j=-\infty}^{\infty} \widehat{T^{-j}\mathcal{P}} \subset \mathcal{A}$ e na última desigualdade utilizamos a relação (70).

Mas, fixando $p \geqslant 1$ e fazendo $n \to \infty$, pelo Teorema 3.31 obtém-se

$$\lim_{n\to\infty} \Big[H(\mathcal{Q}_p|\widehat{\mathcal{Q}_n^-}) - H(\mathcal{Q}_p|\mathcal{A})\Big] = 0,$$

pois $\widehat{\mathcal{Q}_n^-} \uparrow \mathcal{A}$. Além disso, como $\widehat{\mathcal{P}} \subset \mathcal{A}$ e $\widehat{\mathcal{Q}_p^-} \uparrow \mathcal{A}$ quando $p \to \infty$, novamente pelo Teorema 3.31, juntamente com o Teorema 3.29, vemos que

$$\lim_{p\to\infty} H(\mathcal{P}|\widehat{\mathcal{Q}_p^-}) = 0.$$

Consequentemente, fazendo primeiro $n \to \infty$ e depois $p \to \infty$ em (71), obtemos

$$H(\mathcal{P}|\widehat{\mathcal{P}^-}) = 0,$$

ou seja,

$$H(T,\mathcal{P}) = 0.$$

Logo, pelo Teorema 4.21 segue que $\widehat{\mathcal{P}} \subset \mathcal{P}(T)$. Portanto, como $\widehat{\mathcal{P}} \subset \bigcap_{n\geqslant 1} T^{-n}\mathcal{A}$ foi tomada arbitrariamente, conclui-se $\bigcap_{n\geqslant 1} T^{-n}\mathcal{A} \subset \mathcal{P}(T)$, como queríamos. Isso conclui a demonstração do Teorema de Rokhlin-Sinai.

Observação 4.30 . (Comentário final): Vimos na Observação 3.33 que

$$H(T,\mathcal{P}) = 0 \iff \widehat{\mathcal{P}} \subseteq \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}.$$

Equivalentemente, tem-se

$$H(T,\mathcal{P}) > 0 \iff \widehat{\mathcal{P}} \nsubseteq \bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}.$$

Em outras palavras, a entropia do processo $(T,\mathcal{P})$ é positiva se, e somente se, seu "presente" não pode ser integralmente determinado pelo seu "passado". Consequentemente, o Teorema de Rokhlin-Sinai nos diz que um sistema dinâmico $(X,\mathcal{X},\mu,T)$ é um K-automorfismo se, e somente se, para cada "observação" $\mathcal{P}$ do fenômeno em questão, o "passado completo das observações", ou seja, $\bigvee_{i=1}^{\infty} \widehat{T^{-i}\mathcal{P}}$, não pode descrever integralmente a "observação do presente", ou seja, $\widehat{\mathcal{P}}$. Dito de outra maneira, não é possível descrever integralmente o "momento seguinte". Nesse sentido, o Teorema de Rokhlin-Sinai nos diz que um K-automorfismo é um sistema dinâmico com comportamento *completamente aleatório*.

$\square$

BIBLIOGRAFIA

[1] P. Billingsley, Ergodic Theory and Information, John Wiley and Sons, Inc., 1965.

[2] J. Dugundji, Topology, Allyn and Bacon, Inc., Boston, 1966.

[3] E. Glasner, Ergodic Theory via Joinings, Mathematical Surveys and Monographs, American Mathematical Society, 2003.

[4] P. Halmos, Lectures on Ergodic Theory, Chelsea Publishing Company, 1956.

[5] P. Halmos, Measure Theory, Springer-Verlag, 1974.

[6] S. Kalikow e R. McCutcheon, An Outline of Ergodic Theory, Cambridge University Press, 2010.

[7] A. Kolmogorov, A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces, Dokl. Akad. Nauk. SSSR. **119**, 861-864, 1958 (Russian).

[8] M. Nadkarni, Basic Ergodic Theory, Texts and Readings in Mathematics, Hindustan Book Agency, 2013.

[9] J. von Neumann, Zur Operatorenmethode in der klassischen Mechanik, Ann. Math **33**, 587-642, 1932.

[10] J. von Neumann e P. Halmos, Operator methods in classical mechanics II, Ann. Math **43**, 332-350, 1942.

[11] D. Ornstein, An example of a Kolmogorov Automorphism that is not a Bernoulli Shift, Advances in Mathematics **10**, 49-62, 1973.

[12] D. Ornstein, Bernoulli shifts with the same entropy are isomorphic, Advances in Mathematics **4**, 337-352, 1970.

[13] W. Parry, Topics in Ergodic Theory, Cambridge University Press, 1981.

[14] K. Petersen, Ergodic Theory, Cambridge University Press, University of North Carolina, Chapel Hill, 1983.

[15] M. Pinsker, Information and information stability of random variables and processes, Izdat. Akad. Nauk. SSSR., 1960 (Russian) - Holden-Day, San Francisco, 1964.

[16] V. Rokhlin e Ja. Sinai, Constructions and properties of invariant measurable partitions, Dokl. Acad. Nauk. SSSR. $\mathbf{141}$, 1038-1041, 1961 (Russian). Sov. Math. 2 $(\mathbf{6})$, 1611-1614, 1961.

[17] W. Rudin, Real and Complex Analysis, Mc-Graw Hill International Editions, 1987.

[18] Ja. Sinai, Introduction to Ergodic Theory, Princeton University Press, 1977.

[19] Ja. Sinai, Kolmogorov's Work on Ergodic Theory, The Annals of Probability, $(\mathbf{17})$, no.3, 833-839, 1989.

[20] Ja. Sinai, On the concept of entropy for a dynamic system, Dokl. Acad. Nauk. SSSR. $\mathbf{124}$, 768-771, 1959 (Russian).

[21] M. Smorodinsky, Ergodic Theory and Entropy, Springer, 1971.

[22] M. Viana e K. Oliveira, Foundations on Ergodic Theory, Cambridge Studies in Advanced Mathematics, Cambridge, 2016.

[23] B. Weiss, The isomorphism problem in Ergodic Theory, Bulletin of the American Mathematical Society, $(\mathbf{78})$, no.5, 668-684, 1972.