UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE MATEMÁTICA
CURSO EM MATEMÁTICA BACHARELADO
TRABALHO DE CONCLUSÃO DE CURSO

GLEYDSON SANTOS DA SILVA

PROPRIEDADES ERGÓDICAS DO SHIFT DE BERNOULLI

MACEIÓ
2021

GLEYDSON SANTOS DA SILVA

PROPRIEDADES ERGÓDICAS DO SHIFT DE BERNOULLI

Trabalho de Conclusão de Curso submetido ao Colegiado do Curso de Graduação em Matemática da Universidade Federal de Alagoas, como requisito parcial para obtenção do título de bacharel em Matemática.

Orientador: Prof. Dr. Rafael Nóbrega de Oliveira Lucena

MACEIÓ

2021

Catalogação na Fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecário: Jorge Raimundo da Silva – CRB-4 – 1528

S586p Silva, Gleydson Santos da.
Propriedades ergódicas do Shift de Bernoulli / Gleydson Santos da Silva. 2021.
74 f. : il.

Orientador: Rafael Nóbrega de Oliveira Lucena.
Monografia (Trabalho de Conclusão de Curso em Matemática : Bacharelado) Universidade Federal de Alagoas. Instituto de Matemática. Maceió, 2021.

Bibliografia: f. 74.

1. Sistemas Dinâmicos - ergódicos. 2. Shift de Bernoulli. I. Título.

CDU: 514.7

9

UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE MATEMÁTICA
COORDENAÇÃO DO CURSO DEMATEMÁTICA BACHARELADO
Fone: 3214-1405 / E-mail: coordenacao.mat@im.ufal.br

DECLARAÇÃO DE AVALIAÇÃO DE TCC

Informamos à Coordenação do Curso de Graduação em Matemática Bacharelado que o Trabalho de Conclusão de Curso do(a) aluno(a) **GLEYDSON SANTOS DA SILVA**, matrícula nº **17110779**, do curso de **PROPRIEDADES ERGÓDICAS DO SHIFT DE BERNOULLI**", recebeu da Banca Examinadora a seguinte nota: **__9,75__** (**nove e setenta e cinco**), média obtida a partir das seguintes notas atribuídas pelos componentes da Banca Examinadora:

Prof. Dr. Rafael Nóbrega de Oliveira Lucena (Ufal): **___10_____**
Prof. Dr. Davi dos Santos Lima (Ufal): ___10_____
Prof. Dr. Dione Andrade Lara (UFLA): **___9,5_____**
Prof. Dr. Wagner Ranter Gouveia da Silva (Ufal): **___9,5_____**

Maceió, __7___ de março de 2021.

gov.br Documento assinado digitalmente
Rafael Nobrega de Oliveira Lucena
Data: 10/03/2022 10:34:20-0300
Verifique em https://verificador.iti.br

____
Prof. Dr. Rafael Nóbrega de Oliveira Lucena

gov.br Documento assinado digitalmente
DIONE ANDRADE LARA
Data: 10/03/2022 11:19:13-0300
Verifique em https://verificador.iti.br

____
Prof. Dr. Dione Andrade Lara

gov.br Documento assinado digitalmente
DAVI DOS SANTOS LIMA
Data: 10/03/2022 10:58:40-0300
Verifique em https://verificador.iti.br

____
Prof. Dr. Davi dos Santos Lima

gov.br Documento assinado digitalmente
WAGNER RANTER GOUVEIA DA SILVA
Data: 10/03/2022 11:35:20-0300
Verifique em https://verificador.iti.br

____
Prof. Dr. Wagner Ranter Gouveia da Silva

A todos que já quiseram uma obra dedicada a si
e nunca tiveram.

AGRADECIMENTOS

Agradeço a todas as vezes em que, por mais que quisesse, não tive a coragem para desistir;

Agradeço a todas as vezes em que, por mais que não quisesse, tive a coragem para desistir;

Agradeço à família, aos amigos, às irmãs e aos irmãos que estiveram comigo (total ou parcialmente) durante esta minha (e nossa) jornada na graduação;

Agradeço ao meu orientador, Rafael Lucena, por todo o apoio, acolhimento, aconselhamento e incentivo que me proporcionou desde o princípio da graduação;

Agradeço ao professor Dione Lara por aturar minhas importunações algébricas, por me inspirar e por fazer eu me sentir parte do meio matemático;

Sou eternamente grato a todos os meus irmãos e irmãs, especialmente a Murilo Santos, Leandro Lucena, Lucas Honorato e Kaique Cabral, Maxmilian Siqueira e Monique Melo, Victor Ferreira e Lucas Bulandeira, Talvanes Araújo e Caio Philipe, Hegel Marinho e Davi Matheus, Rebeca Alves e Daniel Silva, George Tavares, Symon Igor e Weverson Clayton, Janaine Santos, Wanessa Amancio e Mariana Bispo.

“Posso fazer de tudo, mas não tudo.”

(Greg McKeown)

RESUMO

O presente trabalho consiste no estudo do shift de Bernoulli e algumas de suas propriedades ergódicas.
O shift de Bernoulli é um sistema dinâmico de tempo discreto com medida (a medida de Bernoulli) invariante sob a ação da dinâmica (o shift à esquerda), esta que é uma aplicação mensurável e que atua sobre um espaço de símbolos.

Palavras-chave: Sistemas Dinâmicos, Sistemas Ergódicos, Shift de Bernoulli.

ABSTRACT

The present work consists of the study of the Bernoulli's shift and some ergodic properties of it. The Bernoulli's shift is a dynamical system of discrete time with a measure (the Bernoulli's measure) that is invariant under the action of the dynamic (the shift), this is a measurable application acting over a space of symbols.

Keywords: Dynamical Systems, Ergodic Systems, Bernoulli's Shift.

SUMÁRIO

1 INTRODUÇÃO

Sistemas dinâmicos são modelos matemáticos que representam um conjunto de objetos que (possivelmente) se movimentam ou mudam de estado com o passar do tempo. Lidaremos muito com dois conceitos fundamentais: o conceito de ergodicidade e o conceito de mistura. Abordaremos superficialmente estes conceitos a seguir para fins de interpretá-los, estudá-los-emos com mais detalhes nos próximos capítulos.

Formalmente um sistema (T,μ) - onde $T: X \to X$ é uma função mensurável no conjunto X que preserva a probabilidade μ (definida numa σ-álgebra de subconjuntos de X) - é ergódico, se todo mensurável A que satisfaz $T^{-1}(A) = A$, necessariamente tem $\mu(A) = 1$ ou $\mu(A) = 0$. Para entendermos a importância desta definição, vamos considerar um mensurável invariante A qualquer. Então, $T^{-1}(A) = A$. Ora, neste caso, se considerarmos $x \in A$, então teremos $x \in T^{-1}(A)$ e, portanto, $T(x) \in A$. Isto nos diz, então, que $T(A) \subseteq A$. Portanto, podemos considerar a função mensurável $T|A : A \to A$. Além disto, vemos também que $T^{-1}(X-A) = X - T^{-1}(A) = X - A$. Ou seja, $X - A$ é também um mensurável invariante. E - pelo argumento acima - podemos considerar a função mensurável $T|X-A : X-A \to X-A$. Em outras palavras, um ponto num mensurável invariante está "preso"nele. A sua órbita pela dinâmica T está inteiramente contida neste conjunto. Então conseguimos decompor o sistema em dois sub-sistemas "isolados". A definição de sistema ergódico nos diz então que se há um mensurável invariante ele deve possuir medida total ou medida nula (ou seja, ou ele tem medida nula ou o seu complementar tem medida nula), o que em termos de medida nos diz que estes sub-sistemas são triviais. Em outras palavras, podemos dizer que um sistema ergódico é "indecomponível"em sub-sistemas "isolados".

Algo sobre o qual todos temos uma noção intuitiva é a mistura, o ato de misturar algo é muito familiar a todos, pois comumente se vê no dia a dia. Muitos já tiveram a oportunidade de misturar um pouco de leite numa boa xícara de café, por exemplo. Se pensarmos um pouco, mesmo que intuitivamente, poderemos nos convencer de que se misturamos um sistema (o café e o leite, por exemplo) os pontos (as partículas do leite, por exemplo) vão percorrer todo o espaço possível da xícara ao longo do tempo, conforme misturamos o sistema. Ora, se os pontos percorrem todo o espaço, então - naturalmente - não há pontos que fiquem "confinados"numa região do espaço, ou seja, não há sub-sistemas "isolados". Então é razoável pensar que um sistema misturador seja um sistema ergódico. Bem, isso não se reduz apenas à intuição. De fato, mostraremos neste trabalho que o shift de Bernoulli é um sistema misturador e, por tal motivo, ele é um sistema ergódico.

Estudar o shift de Bernoulli é importante, pois o sistema é muito utilizado em diversas aplicações já que é um modelo adequado para representar vários fenômenos que evoluam no tempo de forma discreta, ou seja, durante certos intervalos de tempo o sistema se encontra em um estado fixo. De tal forma que pode-se considerar a passagem de tempo de maneira discreta (como se o tempo não fosse contínuo). Além disto, outros sistemas dinâmicos podem ser também codificados por meio do shift de Bernoulli o que permite, devido à sua simplicidade, descobrir resultados importantes a cerca do sistema original de maneira simples.

O objetivo deste trabalho é expor alguns resultados básicos da teoria ergódica a respeito do shift de Bernoulli. Sabidamente, queremos definir um espaço adequado para que possamos trabalhar com o shift de Bernoulli, então defini-lo apropriadamente. Depois exploraremos algumas propriedades iniciais do sistema e mostraremos que ele é, de fato, um sistema misturador (e, portanto, ergódico). Por fim, mostraremos uma consequência dos resultados que obteremos.

No capítulo 2 veremos alguns conceitos básicos da teoria da medida que serão imprescindíveis para o nosso estudo. Veremos os conceitos de σ-álgebra, medida (juntamente com alguns exemplos importantes para ilustrar os conceitos), espaço mensurável, espaço de medida e espaço de probabilidade. Daremos motivações a estes conceitos para que fiquem o mais claro possível.

No capítulo 3, portando os conceitos definidos no capítulo 2, daremos início à construção do nosso espaço de probabilidade de interesse, o espaço simbólico. Para construí-lo precisaremos de alguns resultados sobre σ-álgebras e medidas aos quais não demonstraremos, mas faremos as devidas menções de referências que o leitor pode consultar. Depois construiremos o shift de Bernoulli de maneira totalmente natural a partir da definição do shift e da medida de Bernoulli.

No capítulo 4 mostraremos resultados iniciais básicos a cerca do shift de Bernoulli. Mostraremos que o shift é uma aplicação mensurável e mostraremos que a medida de Bernoulli é invariante sob a ação do shift. Finalmente mostraremos o nosso resultado principal em que o shift de Bernoulli é um sistema mixing (ou misturador).

Por fim, no capítulo 5, veremos uma aplicação do fato do shift de Bernoulli ser um sistema mixing para mostrar que todos os autovalores do operador de Perron-Fröbenius associado ao shift de Bernoulli são iguais a 1. Usaremos vários resultados de teoria da medida e análise funcional para isto, vários dos quais não demonstraremos.

2 ESPAÇOS DE MEDIDA

Para definir o shift de Bernoulli primeiro precisaremos construir um espaço apropriado, para isto temos que entender o que são espaços de medida. Começamos com um conjunto X qualquer e a partir de X construiremos um espaço. Queremos ser capazes de "medir"os subconjuntos de X (medir a área, medir o volume, medir a massa, medir o comprimento, etc.), mas talvez nem todos os subconjuntos nos sejam interessantes. Então vamos tentar obter uma coleção $\mathbb{X}$ de subconjuntos de X que possua os subconjuntos que "nos sejam interessantes".

Nos parece razoável imaginar que possamos "medir"o conjunto X (que é o nosso maior subconjunto) e o conjunto vazio $\varnothing$ (que é o nosso menor subconjunto). Ou seja, é razoável exigir que $\varnothing, X \in \mathbb{X}$. Então esta será uma primeira exigência para a coleção $\mathbb{X}$.

Agora, suponhamos que A é um subconjunto de X ao qual sabemos "medir". Ou seja, suponhamos que $A \in \mathbb{X}$. Então nos parece razoável, também, que saibamos "medir"o complemento $X - A$ de A. Intuitivamente, devemos ter que a "medida"de X é a soma das "medidas"de A e de $X - A$. Vamos exigir também que se $A \in \mathbb{X}$, então $X - A \in \mathbb{X}$. Desta forma não é mais necessário exigir que X e $\varnothing$, ambos, sejam elementos de $\mathbb{X}$. Basta exigir que $X \in \mathbb{X}$, pois assim $\varnothing = X - X$ deverá pertencer a $\mathbb{X}$ também graças à segunda exigência.

Suponhamos que A e B sejam elementos de $\mathbb{X}$. Então sabemos "medir"A e também sabemos "medir"B. Nos parece razoável imaginar que também sejamos capazes de "medir"o conjunto $A \cup B$. E a "medida"de $A \cup B$ deve ser no máximo a soma das "medidas"de A e de B. Faz sentido porque se A e B possuírem alguma intersecção, então a "medida"de $A \cup B$ seria um pouco menor do que a soma das "medidas"de A e de B, já que a "medida"da intersecção estaria sendo contabilizada tanto na "medida"de A quanto na "medida"de B. Assim vamos exigir que se $A, B \in \mathbb{X}$, então $A \cup B \in \mathbb{X}$. Decorre daí que se tivermos uma quantidade finita qualquer $A_1, A_2, \ldots, A_n \in \mathbb{X}$, então $\bigcup_{i=1}^{n} A_i \in \mathbb{X}$, basta unir A_1 com A_2, e depois unir $A_1 \cup A_2$ com A_3. E prosseguir assim, sempre adicionando mais um conjunto à união.

Se $A, B \in \mathbb{X}$, então pela segunda exigência temos que $X - A \in \mathbb{X}$ e $X - B \in \mathbb{X}$. Logo, pela terceira exigência temos que $(X - A) \cup (X - B) \in \mathbb{X}$. Mas então pela segunda exigência novamente teremos que $X - [(X-A) \cup (X-B)] \in \mathbb{X}$. E pelas leis de De Morgan, obtemos que

$$X - [(X-A) \cup (X-B)] = [X - (X-A)] \cap [X - (X-B)]$$
$$= A \cap B.$$

Ou seja, obtemos que se $A, B \in \mathbb{X}$, então $A \cap B \in \mathbb{X}$. Prosseguindo indutivamente como antes vamos obter que se $A_1, A_2, ..., A_n \in \mathbb{X}$, então $\bigcap_{i=1}^{n} A_i \in \mathbb{X}$.

Então vamos ter uma coleção $\mathbb{X}$ "fechada para uniões e intersecções finitas". No entanto, será interessante para nós que $\mathbb{X}$ seja "fechada para uniões e intersecções infinitas enumeráveis"também. Ou seja, queremos que se $A_1, A_2, ..., A_n, ...$ é uma sequência de elementos de $\mathbb{X}$, então $\bigcup_{i=1}^{+\infty} A_i \in \mathbb{X}$. Exigindo isto, pelas leis de De Morgan como antes, obtemos que $\bigcap_{i=1}^{+\infty} A_i \in \mathbb{X}$ também. Portanto, com tudo isto chegamos à seguinte definição.

Definição 1. Sejam X um conjunto qualquer e $P(X)$ a coleção de todos os subconjuntos de X. Dizemos que $\mathbb{X} \subseteq P(X)$ é uma **σ-álgebra de subconjuntos de** X se, e somente se:

- $X \in \mathbb{X}$;
- Se $A \in \mathbb{X}$, então $X - A \in \mathbb{X}$;
- Se $(A_n)_{n \in \mathbb{N}}$ é uma sequência em $\mathbb{X}$, então

$$\bigcup_{n=1}^{+\infty} A_n \in \mathbb{X}.$$

Dizemos que $(X, \mathbb{X})$ é um **espaço mensurável** e todo $A \in \mathbb{X}$ é dito ser um **conjunto $\mathbb{X}$-mensurável**.

Vamos considerar um exemplo simples e muito importante de uma σ-álgebra para clarear as ideias.

Exemplo 1. Sejam X um conjunto qualquer e $\mathbb{X} = P(X)$ a coleção de todos os subconjuntos de X. Vamos mostrar que $\mathbb{X}$ é uma σ-álgebra de subconjuntos de X. Começamos notando que

$$X \subseteq X.$$

Portanto, $X \in P(X)$. Em outras palavras, $X \in \mathbb{X}$. Agora, seja $A \in \mathbb{X}$. Então

$$A \subseteq X.$$

Daí,

$$X - A \subseteq X.$$

Ou seja, $X - A \in P(X)$. O que significa que $X - A \in \mathbb{X}$ sempre que $A \in \mathbb{X}$.

Só falta mostrar que $\mathbb{X}$ é fechado para uniões enumeráveis. Seja $(A_n)_{n\in\mathbb{N}}$ uma sequência de elementos de $\mathbb{X}$. Isto significa que

$$A_n \subseteq X, \forall n \in \mathbb{N}.$$

Portanto,

$$\bigcup_{n\in\mathbb{N}} A_n \subseteq X.$$

O que significa que $\bigcup\limits_{n\in\mathbb{N}} A_n \in \mathbb{X}$. Isto conclui a demonstração de que a coleção $P(X)$ de todos os subconjuntos de X é uma σ-álgebra de subconjuntos de X, não importa qual seja o conjunto X.

Agora que temos a coleção dos subconjuntos de X que queremos "medir", vamos falar do que é "medir". Queremos associar a cada $A \in \mathbb{X}$ um número real $\mu(A)$ que será a sua "medida". Então faz sentido pensar que a "medida"seja uma função $\mu : \mathbb{X} \to \mathbb{R}$. Para nós é interessante que a "medida"de um conjunto não seja negativa, pois área, volume, comprimento, massa são números que só nos fazem sentido se forem positivos. Desta forma, nossa primeira exigência à função μ é que $\mu(A) \geq 0$ para todo $A \in \mathbb{X}$.

Faz todo o sentido do mundo que o conjunto $\varnothing$ tenha volume nulo, área nula, comprimento nulo, massa nula, enfim independentemente da interpretação, é natural pensar que devemos ter $\mu(\varnothing) = 0$. Mas não queremos que $\varnothing$ seja o único conjunto de "medida"nula, pois todo quadrado possui volume nulo, intuitivamente, por exemplo.

Queremos que X seja o "maior"conjunto e se pensarmos sobre o volume do espaço tridimensional em relação ao volume das figuras geométricas comuns, nos parece fazer sentido que o volume do espaço inteiro seja infinito. Então vamos permitir que a função μ assuma o valor $+\infty$ e assim trabalharemos com o conjunto dos números reais estendido $\overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$.

Discutimos antes sobre o conjunto $A \cup B$ e concluímos que a sua "medida"deve ser no máximo a soma das "medidas"de A e de B. Se exigirmos que $A \cap B = \varnothing$, isto é, que os conjuntos A e B sejam disjuntos, então é razoável pensar que valha $\mu(A \cup B) = \mu(A) + \mu(B)$. Imagine um trapézio retângulo (Figura 1), ele pode ser dividido em um retângulo e um triângulo retângulo disjuntos. Faz sentido que a área do trapézio seja a soma das áreas do retângulo e do triângulo.

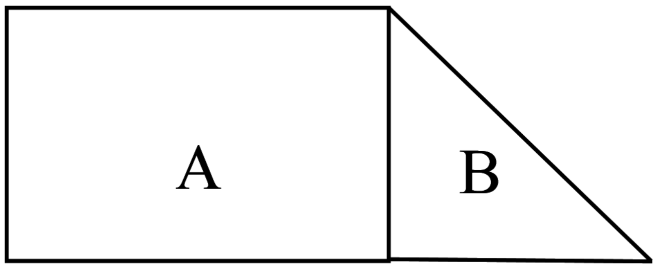

Figura 1 – A área do Trapézio Retângulo $A \cup B$ é a soma das áreas do Retângulo A e do Triângulo B, pois eles são disjuntos.

Por isto, vamos exigir que $\mu(A \cup B) = \mu(A) + \mu(B)$ sempre que $A \cap B = \varnothing$. Por indução, como antes, isto vai nos fornecer que se $A_1, \ldots, A_n \in \mathbb{X}$ e $A_i \cap A_j = \varnothing$ para $i \neq j$, então

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

Mas queremos que isto também seja válido para uniões infinitas enumeráveis de conjuntos dois-a-dois disjuntos. Reunindo tudo isto obtemos a seguinte definição.

Definição 2. Seja $(X,\mathbb{X})$ um espaço mensurável. Uma **medida** em $(X,\mathbb{X})$ é uma função $\mu:\mathbb{X}\to\overline{\mathbb{R}}$ tal que:

- $\mu(A)\geq 0, \forall A\in\mathbb{X}$;

- $\mu(\varnothing)=0$;

- Se $(A_n)_{n\in\mathbb{N}}$ é uma sequência em $\mathbb{X}$ de conjuntos dois-a-dois disjuntos, então

$$\mu\left(\bigcup_{n=1}^{+\infty}A_n\right)=\sum_{n=1}^{+\infty}\mu(A_n).$$

Observamos que na definição exigimos que a série no lado direito da igualdade na terceira condição seja convergente (lembramos que ela pode assumir o valor $+\infty$).
Dizemos que $(X,\mathbb{X},\mu)$ é um **espaço de medida**. Além disto, dizemos que μ é uma **probabilidade** em $(X,\mathbb{X})$, se $\mu(X)=1$. Neste caso, $(X,\mathbb{X},\mu)$ é dito um **espaço de probabilidade**.

Vamos ver um exemplo clássico de uma medida para ilustrar o conceito.

Exemplo 2. Sejam $X=\mathbb{N}$ o conjunto de todos os números naturais e $\mathbb{X}=P(\mathbb{N})$ a coleção de todos os subconjuntos de X. Já vimos antes que $(X,\mathbb{X})$, assim definidos, é um espaço mensurável.
Definimos a função

$$c:\mathbb{X}\to\overline{\mathbb{R}}$$

por

$$c(A)=\begin{cases}\#A, \text{ se A é finito}\\ \infty, \text{ se A é infinito}\end{cases},$$

onde $\#A$ denota o número de elementos do conjunto A.

Vejamos como a função c atua sobre alguns conjuntos simples:

Como o conjunto $\{1,2,3\}$ é finito e possui três elementos, então

$$c(\{1,2,3\}) = \#\{1,2,3\} = 3.$$

Como o conjunto $\{2,7,89,65\}$ é finito e possui quatro elementos, então

$$c(\{2,7,89,65\}) = \#\{2,7,89,65\} = 4.$$

Como o conjunto $\varnothing$ é finito e possui zero elementos (não possui elementos), então

$$c(\varnothing) = \#\varnothing = 0.$$

Como o conjunto $\{1,3,5,7,9,11,13,...\}$ de todos os números naturais ímpares é infinito, então

$$c(\{1,3,5,7,9,11,13,...\}) = \infty.$$

Como o conjunto $\{2,3,5,7,11,13,...\}$ de todos os números naturais primos é infinito, então

$$c(\{2,3,5,7,11,13,...\}) = \infty.$$

Bom, daí nós vimos que $c(\varnothing) = 0$ e é claro da sua definição que

$$c(A) \geq 0, \forall A \in \mathbb{X}.$$

Resta mostrar apenas a σ-aditividade para que c seja uma medida. Para isto seja $(A_n)_{n\in\mathbb{N}}$ uma sequência de elementos de $\mathbb{X}$ dois-a-dois disjuntos. Só existem dois casos:

- Ou todos os A_n's são finitos;

- Ou nem todos os A_n's são finitos.

Se nem todos os A_n's são finitos, então pelo menos um deles, digamos A_j, é infinito. E, portanto, a união de todos eles é um conjunto infinito. E, por conseguinte,

$$c\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \infty.$$

Por outro lado, como A_j é infinito, então

$$c\left(A_j\right) = \infty.$$

Portanto, se somarmos os valores de c em todos os A_n's teremos necessariamente o valor infinito, por causa do valor de c em A_j. Ou seja,

$$\sum_{n=1}^{+\infty} c\left(A_n\right) = \infty.$$

O que mostra que

$$c\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n=1}^{+\infty} c\left(A_n\right).$$

Agora, se, por outro lado, todos os A_n's forem finitos devemos observar se a união de todos é finita (se só uma quantidade finita dos A_n's for não-vazia) ou infinita (se uma quantidade infinita dos A_n's for não-vazia).

Se $\bigcup\limits_{n\in\mathbb{N}} A_n$ é infinito, então existe pelo menos um A_j não-vazio, pois caso contrário todos os A_n's seriam vazios e, portanto, a união de todos eles seria o conjunto vazio (que não é infinito). Desta forma podemos considerar

$$\bigcup_{n\in\mathbb{N}} A_n = \left(\bigcup_{n\in\mathbb{N}-\{j\}} A_n\right) \cup A_j.$$

Já que $\bigcup\limits_{n\in\mathbb{N}} A_n$ é infinito e A_j é finito, então $\bigcup\limits_{n\in\mathbb{N}-\{j\}} A_n$ é infinito e disjunto de A_j (já que A_j é disjunto de todos os A_n's). Portanto, o problema se reduz a uma união de dois conjuntos disjuntos sendo um deles infinito, pelo que vimos acima a função c é σ-aditiva neste caso.

O último caso que resta analisar é o caso em que todos os A_n's são finitos e $\bigcup_{n\in\mathbb{N}} A_n$ também é finito. Neste caso, somente uma quantidade finita de A_n's é finita. Digamos $A_{n_1},\ldots,A_{n_j}$. Portanto,

$$\bigcup_{n\in\mathbb{N}} A_n = A_{n_1}\cup A_{n_2}\cup\cdots\cup A_{n_j}.$$

Logo, por ser finito, temos

$$c\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \#\left(A_{n_1}\cup A_{n_2}\cup\cdots\cup A_{n_j}\right).$$

Como os conjuntos A_n's são dois-a-dois disjuntos ficamos com

$$c\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \#A_{n_1}+\#A_{n_2}+\cdots+\#A_{n_j}.$$

E como cada A_n é finito, então isso significa que

$$c\left(\bigcup_{n\in\mathbb{N}} A_n\right) = c\left(A_{n_1}\right)+c\left(A_{n_2}\right)+\cdots+c\left(A_{n_j}\right).$$

Portanto, em todos os casos a função c satisfaz a σ-aditividade. Consequentemente, c é uma medida em $(X,\mathbb{X})$ como queríamos demonstrar. A medida c é chamada **medida de contagem**.

3 O ESPAÇO SIMBÓLICO E O SHIFT DE BERNOULLI

Vamos construir o nosso espaço. Consideremos $(X,\mathbb{X},P)$ um espaço de probabilidade. Definimos

$$\Sigma = X^{\mathbb{N}} = X \times X \times \cdots \times X \times \cdots = \{(x_n)_{n\in\mathbb{N}} | x_n \in X, \forall \in \mathbb{N}\}.$$

Σ é a coleção de todas as sequências de elementos de X. Ou seja, dado $x \in \Sigma$ temos $x : \mathbb{N} \to X$ e $x(n) = x_n$ para cada $n \in \mathbb{N}$. Podemos definir Σ como $X^{\mathbb{Z}}$ também se quisermos. Mas este não será o nosso caso.

O nosso espaço será um espaço de probabilidade cujo conjunto é Σ. Precisamos agora encontrar uma σ-álgebra e uma probabilidade para completá-lo. Primeiro vamos provar o seguinte resultado bastante útil.

Teorema 1. Sejam X um conjunto e Π uma família não-vazia de σ-álgebras de subconjuntos de X. Então a coleção

$$\mathbb{A} = \bigcap_{\mathbb{X}\in\Pi} \mathbb{X}$$

é uma σ-álgebra de subconjuntos de X.

Demonstração. Primeiro observamos que $X \in \mathbb{A}$. De fato, para cada $\mathbb{X} \in \Pi$, $\mathbb{X}$ é uma σ-álgebra de subconjuntos de X. Portanto, $X \in \mathbb{X}$. Então $X \in \bigcap\limits_{\mathbb{X}\in\Pi} \mathbb{X} = \mathbb{A}$.

Agora seja $A \in \mathbb{A}$ arbitrário. Então, pela definição de $\mathbb{A}$, $A \in \mathbb{X}$ para todo $\mathbb{X} \in \Pi$. Como $\mathbb{X}$ é uma σ-álgebra de subconjuntos de X, então $X - A \in \mathbb{X}$, isto para cada $\mathbb{X} \in \Pi$. E então, $X - A \in \bigcap\limits_{\mathbb{X}\in\Pi} \mathbb{X} = \mathbb{A}$.

Por fim, seja $(A_n)_{n\in\mathbb{N}}$ uma sequência em $\mathbb{A}$. Então, $(A_n)_{n\in\mathbb{N}}$ é uma sequência em $\mathbb{X}$ para todo $\mathbb{X} \in \Pi$. Como cada $\mathbb{X}$ é uma σ-álgebra de subconjuntos de X, então

$$\bigcup_{n=1}^{+\infty} A_n \in \mathbb{X},$$

para todo $\mathbb{X} \in \Pi$. Por conseguinte,

$$\bigcup_{n=1}^{+\infty} A_n \in \bigcap_{\mathbb{X} \in \Pi} \mathbb{X} = \mathbb{A}.$$

Consequentemente, $\mathbb{A}$ é uma σ-álgebra de subconjuntos de X.

Q. E. D.

Com este resultado em mãos, podemos definir algo que nos será importante. Primeiro vamos considerar Γ uma família qualquer de subconjuntos de X. Não é difícil ver que a coleção $P(X)$ de todos os subconjuntos de X é uma σ-álgebra de subconjuntos de X. E é claro que $\Gamma \subseteq P(X)$, pois todo elemento de Γ é um subconjunto de X.

Então se considerarmos

$$\Pi = \{\mathbb{X} \subseteq P(X) | \mathbb{X} \text{ é uma } \sigma\text{-álgebra de subconjuntos de } X \text{ e } \Gamma \subseteq \mathbb{X}\},$$

teremos $\Pi \neq \varnothing$, pois acabamos de ver que $P(X) \in \Pi$. Ora, então se aplicarmos o teorema que acabamos de provar para esta coleção Π obteremos uma σ-álgebra $\sigma(\Gamma)$ de subconjuntos de X que é a "menor"σ-álgebra de subconjuntos de X que contém Γ. Condensamos esta discussão na seguinte definição.

Definição 3. Sejam X um conjunto qualquer e Γ uma família qualquer de subconjuntos de X. A **σ-álgebra gerada por** Γ é

$$\sigma(\Gamma) = \bigcap_{\mathbb{X} \in \Pi} \mathbb{X},$$

onde Π é a coleção de todas as σ-álgebras de subconjuntos de X que contém Γ.

Com esta noção de σ-álgebra gerada podemos considerar uma σ-álgebra importantíssima que tem muita utilidade na matemática. Vamos ver um pouco dela no exemplo a seguir.

Exemplo 3. Seja $\mathbb{R}$ o conjunto de todos os números reais. Vamos considerar a família $\mathscr{F}$ de todos os intervalos abertos de $\mathbb{R}$. A σ-álgebra $\mathbb{S}$ gerada pela família $\mathscr{F}$ é chamada **σ-álgebra de Borel do conjunto** $\mathbb{R}$. Praticamente todos os subconjuntos usuais que conhecemos de $\mathbb{R}$ são elementos de $\mathbb{S}$. Todos os conjuntos abertos, todos os conjuntos fechados e todos os intervalos de $\mathbb{R}$, os subconjuntos numéricos $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R} - \mathbb{Q}$. É uma tarefa um tanto quanto difícil encontrar um subconjunto de $\mathbb{R}$ que não seja elemento de $\mathbb{S}$, mas tais conjuntos existem. Um exemplo são os conjuntos de Vitali (veja [8]).

Para motivar a seguinte definição vamos ver um exemplo para ilustrar bem tudo o que já vimos até agora.

Exemplo 4. Vamos considerar $X = \{1,2,3,4,5,6\}$ o conjunto dos possíveis valores que se pode obter após o lançamento de um dado de seis faces. Imaginemos que lançamos este dado infinitas vezes consecutivas. Após cada lançamento anotamos os valores obtidos. Digamos que:

- no primeiro lançamento o resultado obtido foi 5;
- no segundo laçamento o resultado obtido foi 4;
- no terceiro laçamento o resultado obtido foi 5;
- no quarto lançamento o resultado obtido foi 1;
- no quinto lançamento o resultado obtido foi 3;
- no sexto lançamento o resultado obtido foi 3;
- . . .

Assim obtemos uma sequência $(5,4,5,1,3,3,...)$. Porém esta é somente uma das inúmeras sequências possíveis que poderíamos obter se realizássemos infinitos lançamentos consecutivos. Por exemplo, poderíamos ter a sequência $(1,3,4,6,1,...)$. E existem inúmeras outras possibilidades de sequências.

O conjunto Σ é a coleção de todas estas sequências possíveis. Um fato interessante sobre isto é que não importa quais sejam os resultados que obteremos nos lançamentos, sempre podemos ter certeza que a sequência destes valores obtidos estará em Σ. Se fossemos capazes de identificá-la, então estaríamos prevendo todos os resultados dos infinitos lançamentos.

Se fizéssemos uma aposta de que no quarto lançamento o resultado seria 6, então a coleção de todas as sequências possíveis que nos permitiriam vencer a aposta seria

$$[4;6] = \{(r_1, r_2, r_3, 6, r_5, ...)| r_i \in X, \forall i \in \mathbb{N} - \{4\}\}.$$

Os $r_1, r_2, r_3, r_5, ...$ são os resultados obtidos no primeiro, segundo, terceiro, quinto, ... lançamentos, respectivamente. Não nos interessa que resultado obteremos nestes lançamentos, contanto que no quarto lançamento o resultado seja 6.

Conjuntos como $[4;6]$ serão muito importantes para nós e, portanto, vamos defini-los apropriadamente. No entanto, para nós será necessário generalizar esta definição um pouco mais.

Definição 4. Se A é um elemento de $\mathbb{X}$, então denotamos

$$[m:A] = \{(x_n)_{n\in\mathbb{N}} \in \Sigma | x_m \in A\} = X \times X \times \cdots \times X \times A \times X \times \cdots.$$

Se $m \leq n$ e $A_m, ..., A_n$ são elementos de $\mathbb{X}$, então denotamos

$$\begin{aligned}
[m:A_m,...,A_n] &= [m:A_m] \cap [m+1:A_{m+1}] \cap ... \cap [n:A_n] \\
&= X \times X \times \cdots \times X \times A_m \times \cdots \times A_n \times X \times X \times \cdots \\
&= \{(x_n)_{n\in\mathbb{N}} \in \Sigma | x_i \in A_i, i \in \{m, m+1, ..., n\}\}
\end{aligned}$$

Os conjuntos da forma $[m;A_m,...,A_n]$ serão chamados de **cilindros** de Σ por nós.

Denotaremos por $\mathbb{B}$ a σ-álgebra gerada pela família de todos os cilindros de Σ. E assim obtemos um espaço mensurável $(\Sigma, \mathbb{B})$. Nos resta apenas encontrar agora a probabilidade para o nosso espaço. Conforme a página 99 de [7], $\mathbb{B}$ é também a **σ-álgebra produto** (para uma definição formal consulte as páginas 408 e 409 no apêndice de [7]).

Existe uma única probabilidade μ (que denotamos $\mu = P^{\mathbb{N}}$ e dizemos que é a **medida produto**) em $(\Sigma, \mathbb{B})$ tal que

$$\mu([m;A_m,...,A_n]) = P(A_m) \cdot ... \cdot P(A_n),$$

para qualquer cilindro $[m;A_m,...,A_n]$ de Σ (conforme página 99 de [7] e - para uma demonstração deste fato veja -, de acordo com o que há na página 409 de [7], o teorema 38.B de [6]). Esta probabilidade μ chama-se **medida de Bernoulli**. Observe que μ não depende do número m, depende apenas dos mensuráveis $A_m,...,A_n \in \mathbb{X}$ e da probabilidade P.

Definição 5. Sejam $(X,\mathbb{X},P)$ um espaço de probabilidade, $\Sigma = X^{\mathbb{N}}, \mathbb{B}$ a σ-álgebra gerada pela família de todos os cilindros de Σ e $\mu = P^{\mathbb{N}}$. O espaço de probabilidade $(\Sigma,\mathbb{B},\mu)$ chama-se **espaço simbólico**.

Com o nosso espaço definido, temos tudo o que precisamos para definir o shift de Bernoulli.

Definição 6. Definimos $\sigma : \Sigma \to \Sigma$ por $\sigma((x_n)_{n\in\mathbb{N}}) = (x_{n+1})_{n\in\mathbb{N}}$. A função σ chama-se **shift** (ou **deslocamento**) à esquerda. O **deslocamento de Bernoulli** ou **shift de Bernoulli** é a dupla (σ,μ) formada pelo shift e pela medida de Bernoulli.

Vejamos um exemplo para melhor compreensão.

Exemplo 5. Sejam $X = \mathbb{N}, \Sigma = X^{\mathbb{N}}$ e $(1,2,3,4,5,6,7,8,...) \in \Sigma$. Então

$$\sigma((1,2,3,4,5,6,7,8,...)) = (2,3,4,5,6,7,8,...).$$

Ou seja, na prática o shift apenas "apaga"o primeiro termo da sequência. Mais ainda, é importante observar que na sequência original $(1,2,3,4,5,6,7,8,...)$ o número 2 ocupa a segunda posição, o número 3 ocupa a terceira posição, o número $n+1$ ocupa a $(n+1)$-ésima posição, já na sequência $(2,3,4,5,6,7,8,...)$ o número 2 ocupa a primeira posição, o número 3 ocupa a segunda posição e o número $n+1$ ocupa a n-ésima posição. Em outras palavras, cada termo da sequência é deslocado uma posição à esquerda. O que justifica o nome deslocamento (à esquerda) para o shift.

4 PROPRIEDADES DO SHIFT DE BERNOULLI

Começamos esta seção com as duas definições importantes a seguir.

Definição 7. Sejam $(X,\mathbb{X})$ e $(A,\mathbb{A})$ dois espaços mensuráveis. Uma função $f : X \to A$ é **mensurável com respeito às σ-álgebras $\mathbb{X}$ e $\mathbb{A}$** se, e somente se,

$$f^{-1}(H) \in \mathbb{X}, \forall H \in \mathbb{A}.$$

Ou seja, se a pré-imagem por f de qualquer conjunto $\mathbb{A}$-mensurável for um conjunto $\mathbb{X}$-mensurável. Caso $X = A$ e $\mathbb{X} = \mathbb{A}$, então dizemos apenas que f é **mensurável com respeito à σ-álgebra $\mathbb{X}$.**

Definição 8. Sejam $(X,\mathbb{X},m)$ um espaço de medida e $T : X \to X$ uma função mensurável com respeito à σ-álgebra $\mathbb{X}$. Dizemos que m é uma **medida T-invariante** (ou invariante sob a ação de T), se e somente se,

$$m\left(T^{-1}(H)\right) = m(H),$$

para todo $H \in \mathbb{X}$.

As seguintes definições e resultados serão importantes para demonstrar dois importantes resultados a cerca do shift de Bernoulli:

Definição 9. Seja X um conjunto qualquer. Uma coleção $\mathbb{A}$ de subconjuntos de X é uma **álgebra de subconjuntos de X**, se satisfizer:

A1) $X \in \mathbb{A}$;

A2) Se $E \in \mathbb{A}$, então $X - E \in \mathbb{A}$;

A3) Se E e F são elementos de $\mathbb{A}$, então $E \cup F \in \mathbb{A}$.

Definição 10. Seja X um conjunto qualquer. Uma coleção C de subconjuntos de X é uma **classe monótona**, se satisfizer:

CM1) $X \in C$;

CM2) Se $(A_n)_{n\in\mathbb{N}}$ é uma sequência em C e $A_1 \subseteq A_2 \subseteq ... \subseteq A_n \subseteq ...$, então $\bigcup\limits_{n=1}^{+\infty} A_n \in C$.

CM3) Se $(A_n)_{n\in\mathbb{N}}$ é uma sequência em C e $A_1 \supseteq A_2 \supseteq ... \supseteq A_n \supseteq ...$, então $\bigcap\limits_{n=1}^{+\infty} A_n \in C$.

Teorema 2. Seja X um conjunto qualquer e seja Γ uma coleção não-vazia de classes monótonas (formadas por subconjuntos de X). Então $\mathscr{C} = \bigcap\limits_{C\in\Gamma} C$ é uma classe monótona.

Demonstração. Para cada $C \in \Gamma$ como C é uma classe monótona, então $X \in C$. Portanto, $X \in \bigcap\limits_{C\in\Gamma} C$. Ou seja, $X \in \mathscr{C}$.

Agora, se $(A_n)_{n\in\mathbb{N}}$ é uma sequência em $\mathscr{C}$ tal que $A_1 \subseteq A_2 \subseteq ... \subseteq A_n \subseteq ...$, então como $A_n \in \mathscr{C}$ para cada $n \in \mathbb{N}$, a definição de $\mathscr{C}$ nos diz que $A_n \in C$ para cada $C \in \Gamma$ e para cada $n \in \mathbb{N}$. Logo, como C é uma classe monótona, $\bigcup\limits_{n=1}^{+\infty} A_n \in C$, para cada $C \in \Gamma$. Portanto, $\bigcup\limits_{n=1}^{+\infty} A_n \in \bigcap\limits_{C\in\Gamma} C$. Ou seja, $\bigcup\limits_{n=1}^{+\infty} A_n \in \mathscr{C}$.

Por fim, se $(A_n)_{n\in\mathbb{N}}$ é uma sequência em $\mathscr{C}$ tal que $A_1 \supseteq A_2 \supseteq ... \supseteq A_n \supseteq ...$, então como $A_n \in \mathscr{C}$ para cada $n \in \mathbb{N}$, a definição de $\mathscr{C}$ nos diz que $A_n \in C$ para cada $C \in \Gamma$ e para cada $n \in \mathbb{N}$. Logo, como C é uma classe monótona, $\bigcap\limits_{n=1}^{+\infty} A_n \in C$, para cada $C \in \Gamma$. Portanto, $\bigcap\limits_{n=1}^{+\infty} A_n \in \bigcap\limits_{C\in\Gamma} C$. Ou seja, $\bigcap\limits_{n=1}^{+\infty} A_n \in \mathscr{C}$. O que prova que realmente $\mathscr{C}$ é uma classe monótona.

Q.E.D.

Com este resultado em mãos e observando que a coleção $\mathscr{P}(X)$ de todas as partes de X é uma classe monótona chegamos à seguinte definição:

Definição 11. Sejam X um conjunto qualquer, F uma família de subconjuntos de X e

$$\Gamma_F = \{C \subseteq \mathscr{P}(X) \mid C \text{ é uma classe monótona e } F \subseteq C\}.$$

A menor classe monótona que contém a família F é a classe monótona $\mathscr{C} = \bigcap\limits_{C\in\Gamma_F} C$.

Teorema 3. (Teorema das classes monótonas)
A menor classe monótona que contém uma álgebra $\mathbb{A}$ coincide com a σ-álgebra $\sigma(\mathbb{A})$ gerada por $\mathbb{A}$.

Demonstração. Este é o teorema A.1.18 de [7], página 397, uma demonstração deste teorema pode ser encontrada na página 36 de [13].

Podemos então mostrar o seguinte resultado.

Teorema 4. O shift $\sigma : \Sigma \to \Sigma$ é uma função mensurável com respeito à σ-álgebra $\mathbb{B}$ e, além disto, a medida de Bernoulli $\mu : \mathbb{B} \to \mathbb{R}$ é invariante sob a ação do shift.

Demonstração. Primeiro observamos que a pré-imagem de qualquer cilindro pelo shift ainda é um cilindro:

$$\begin{aligned}\sigma^{-1}\left([m : A_m, ..., A_n]\right) &= \sigma^{-1}\left(\{(x_k)_k \in \Sigma | x_m \in A_m, ..., x_n \in A_n\}\right)\\ &= \{(y_k)_k \in \Sigma | y_{m+1} \in A_m, ..., y_{n+1} \in A_n\}\\ &= [m+1 : A_m, ..., A_n].\end{aligned}$$

Seja $\mathbb{A}$ a coleção de todas as uniões finitas de cilindros dois-a-dois disjuntos. Pode-se mostrar que $\mathbb{A}$ é uma álgebra. Consideremos $B \in \mathbb{A}$. Então existem $C_1, C_2, ..., C_n$ cilindros dois-a-dois disjuntos tais que

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

Donde,

$$\begin{aligned}\sigma^{-1}(B) &= \sigma^{-1}\left(\bigcup_{i=1}^{n} C_i\right)\\ &= \bigcup_{i=1}^{n} \sigma^{-1}(C_i).\end{aligned}$$

Como cada C_i é um cilindro, então cada $\sigma^{-1}(C_i)$ é um cilindro também e já que $C_i \cap C_j = \varnothing$, se $i \neq j$, então

$$\begin{aligned}\sigma^{-1}(C_i) \cap \sigma^{-1}(C_j) &= \sigma^{-1}(C_i \cap C_j)\\ &= \sigma^{-1}(\varnothing)\\ &= \varnothing.\end{aligned}$$

Portanto, concluímos que $\sigma^{-1}(B)$ é uma união finita de cilindros dois-a-dois disjuntos. Ou seja, $\sigma^{-1}(B) \in \mathbb{A}, \forall B \in \mathbb{A}$.

Seja agora $C = \left\{A \in \mathbb{B} \mid \sigma^{-1}(A) \in \mathbb{B}\right\}$. Pelo que acabamos de mostrar temos que $\mathbb{A} \subseteq C$. Queremos mostrar que $\mathbb{B} \subseteq C$. Pode-se mostrar que $\mathbb{A}$ gera a σ-álgebra $\mathbb{B}$. Portanto, se mostrarmos que C é uma classe monótona, o teorema das classes monótonas vai nos fornecer que $\mathbb{B} \subseteq C$, pois $\mathbb{B}$ é a menor classe monótona que contém $\mathbb{A}$ e C é uma classe monótona que contém $\mathbb{A}$. Vamos então mostrar que C é uma classe monótona.

Como $\mathbb{A} \subseteq C$, temos que $\Sigma \in C$. Seja agora $(A_n)_{n\in\mathbb{N}}$ uma sequência monótona não-decrescente em C e seja $A = \bigcup_{n=1}^{+\infty} A_n$. Como cada $A_n \in C$, então cada $A_n \in \mathbb{B}$. Logo, $A \in \mathbb{B}$. Vamos mostrar que $A \in C$. Para isto, notamos que

$$
\begin{aligned}
\sigma^{-1}(A) &= \sigma^{-1}\left(\bigcup_{n=1}^{+\infty} A_n\right) \\
&= \bigcup_{n=1}^{+\infty} \sigma^{-1}(A_n).
\end{aligned}
$$

Como $A_n \in C, \forall n \in \mathbb{N}$, então $\sigma^{-1}(A_n) \in \mathbb{B}, \forall n \in \mathbb{N}$. Logo, $\bigcup_{n=1}^{+\infty} \sigma^{-1}(A_n) \in \mathbb{B}$. Ou seja, $\sigma^{-1}(A) \in \mathbb{B}$, o que nos diz que $A \in C$, ou em outras palavras, C é fechado para uniões enumeráveis monótonas.

Agora seja $(A_n)_{n\in\mathbb{N}}$ uma sequência monótona não-crescente em C e seja $A = \bigcap_{n=1}^{+\infty} A_n$. Ora, como cada $A_n \in C$, então cada $A_n \in \mathbb{B}$. Portanto, A, por ser uma interseção enumerável de conjuntos $\mathbb{B}$-mensuráveis, é um conjunto $\mathbb{B}$-mensurável. Queremos mostrar que $A \in C$, para isto consideremos:

$$
\begin{aligned}
\sigma^{-1}(A) &= \sigma^{-1}\left(\bigcap_{n=1}^{+\infty} A_n\right) \\
&= \bigcap_{n=1}^{+\infty} \sigma^{-1}(A_n).
\end{aligned}
$$

Como cada $A_n \in C$, então $\sigma^{-1}(A_n) \in \mathbb{B}$, para cada $n \in \mathbb{N}$. Logo, $\bigcap_{n=1}^{+\infty} A_n \in \mathbb{B}$. Ou seja, $\sigma^{-1}(A) \in \mathbb{B}$ e, portanto, $A \in C$. O que nos diz que C é fechada para intersecções enumeráveis monótonas. Consequentemente, C é uma classe monótona e, como discutimos antes, o teorema das classes monótonas nos garante que $\mathbb{B} \subseteq C$. Ou seja, $\sigma^{-1}(A) \in \mathbb{B}, \forall A \in \mathbb{B}$. O que mostra que o shift é mensurável com respeito à σ-álgebra $\mathbb{B}$.

Além disso,

$$\begin{aligned}\mu(\sigma^{-1}([m : A_m, ..., A_n])) &= \mu([m+1 : A_m, ..., A_n]) \\ &= P(A_m) \cdot ... \cdot P(A_n) \\ &= \mu([m : A_m, ..., A_n]).\end{aligned}$$

Seja $\mathbb{A}$ a álgebra das uniões finitas de cilindros dois-a-dois disjuntos. Dado $B \in \mathbb{A}$ temos que $B = \bigcup_{i=1}^{n} C_i$ onde C_i são cilindros dois-a-dois disjuntos. Logo,

$$\begin{aligned}\mu\left(\sigma^{-1}(B)\right) &= \mu\left(\sigma^{-1}\left(\bigcup_{i=1}^{n} C_i\right)\right) \\ &= \mu\left(\bigcup_{i=1}^{n} \sigma^{-1}(C_i)\right).\end{aligned}$$

Como C_i são dois-a-dois disjuntos, então $\sigma^{-1}(C_i)$ são dois-a-dois disjuntos e, pela σ-aditividade da medida de Bernoulli, obtemos:

$$\mu\left(\sigma^{-1}(B)\right) = \sum_{i=1}^{n} \mu\left(\sigma^{-1}(C_i)\right).$$

Como C_i são cilindros e como já mostramos que μ é σ-invariante para cilindros, então

$$\mu\left(\sigma^{-1}(C_i)\right) = \mu\left(C_i\right), \forall i = 1, 2, ..., n.$$

Logo,

$$\begin{aligned}\mu\left(\sigma^{-1}(B)\right) &= \sum_{i=1}^{n} \mu\left(C_i\right) \\ &= \mu\left(\bigcup_{i=1}^{n} C_i\right) \\ &= \mu\left(B\right).\end{aligned}$$

Usamos o fato de que os C_i são dois-a-dois disjuntos e a σ-aditividade de μ na penúltima igualdade. Com isto mostramos que μ é σ-invariante para elementos de $\mathbb{A}$.

Agora seja $C = \left\{A \in \mathbb{B} \mid \mu\left(\sigma^{-1}(A)\right) = \mu\left(A\right)\right\}$. Pelo que acabamos de mostrar, temos que $\mathbb{A} \subseteq C$. Vamos mostrar que C é uma classe monótona e o teorema das classes monótonas nos fornecerá que $\mathbb{B} \subseteq C$ e teremos o desejado.

Bem, como $\mathbb{A} \subseteq C$, então $\Sigma \in C$. Agora seja $(A_n)_{n\in\mathbb{N}}$ uma sequência monótona não-decrescente em C e seja $A = \bigcup_{n=1}^{+\infty} A_n$.

Como cada $A_n \in C$, então cada $A_n \in \mathbb{B}$ e, portanto, $A \in \mathbb{B}$. Queremos mostrar que $A \in C$, para isto observamos:

$$\begin{aligned}
\mu\left(\sigma^{-1}(A)\right) &= \mu\left(\sigma^{-1}\left(\bigcup_{n=1}^{+\infty} A_n\right)\right) \\
&= \mu\left(\bigcup_{n=1}^{+\infty} \sigma^{-1}(A_n)\right).
\end{aligned}$$

Como $A_n \subseteq A_{n+1}, \forall n \in \mathbb{N}$, então $\sigma^{-1}(A_n) \subseteq \sigma^{-1}(A_{n+1}), \forall n \in \mathbb{N}$. Portanto, o lema 3.4(a) de [1], página 21, nos fornece:

$$\begin{aligned}
\mu\left(\sigma^{-1}(A)\right) &= \lim_{n\to+\infty} \mu\left(\sigma^{-1}(A_n)\right) \\
&= \lim_{n\to+\infty} \mu(A_n) \\
&= \mu\left(\bigcup_{n=1}^{+\infty} A_n\right) \\
&= \mu(A).
\end{aligned}$$

Usamos na segunda igualdade o fato de que cada $A_n \in \mathbb{A}$ e que μ é σ-invariante nos elementos de $\mathbb{A}$ e usamos na terceira igualdade o lema 3.4(a) de [1], novamente. Isto nos mostra que $A \in C$ e, portanto, C é fechada para uniões enumeráveis monótonas.

Agora seja $(A_n)_{n\in\mathbb{N}}$ uma sequência monótona não-crescente em C e seja $A = \bigcap_{n=1}^{+\infty} A_n$.

Como cada $A_n \in C$, então cada $A_n \in \mathbb{B}$ e, portanto, $A \in \mathbb{B}$. Queremos mostrar que $A \in C$, para isto observamos:

$$\begin{aligned}
\mu\left(\sigma^{-1}(A)\right) &= \mu\left(\sigma^{-1}\left(\bigcap_{n=1}^{+\infty} A_n\right)\right) \\
&= \mu\left(\bigcap_{n=1}^{+\infty} \sigma^{-1}(A_n)\right).
\end{aligned}$$

Como $A_n \supseteq A_{n+1}, \forall n \in \mathbb{N}$, então $\sigma^{-1}(A_n) \supseteq \sigma^{-1}(A_{n+1}), \forall n \in \mathbb{N}$. Portanto, o lema 3.4(b) de [1], página 21, nos fornece:

$$\begin{aligned} \mu\left(\sigma^{-1}(A)\right) &= \lim_{n\to+\infty} \mu\left(\sigma^{-1}(A_n)\right) \\ &= \lim_{n\to+\infty} \mu(A_n) \\ &= \mu\left(\bigcap_{n=1}^{+\infty} A_n\right) \\ &= \mu(A). \end{aligned}$$

Usamos na segunda igualdade o fato de que cada $A_n \in \mathbb{A}$ e que μ é σ-invariante nos elementos de $\mathbb{A}$ e usamos na terceira igualdade o lema 3.4(b) de [1], novamente. Isto nos mostra que $A \in C$ e, portanto, C é fechada para intersecções enumeráveis monótonas. Consequentemente, C é uma classe monótona e, pelo que discutimos antes, o teorema das classes monótonas nos garante que $\mathbb{B} \subseteq C$, ou seja,

$$\mu\left(\sigma^{-1}(A)\right) = \mu(A)$$

para cada $A \in \mathbb{B}$.

Mais ainda, por indução, seja $j \in \mathbb{N}$ tal que

$$\mu(\sigma^{-j}(A)) = \mu(A),$$

para todo $A \in \mathbb{B}$, então

$$\mu(\sigma^{-(j+1)}(A)) = \mu(\sigma^{-j}(\sigma^{-1}(A))).$$

Como $\sigma^{-1}(A) \in \mathbb{B}$, a nossa hipótese de indução nos fornece:

$$\begin{aligned} \mu(\sigma^{-(j+1)}(A)) &= \mu(\sigma^{-1}(A)) \\ &= \mu(A). \end{aligned}$$

Decorre do princípio de indução finita que

$$\mu(\sigma^{-j}(A)) = \mu(A),$$

para todo $A \in \mathbb{B}$ e todo $j \in \mathbb{N}$.

Q.E.D.

Agora apresentaremos uma definição e mostraremos um lema importantíssimos para o nosso objetivo.

Definição 12. Um sistema $(\sigma,\ \mu)$, onde σ é uma função mensurável e μ é uma medida σ-invariante, é **mixing** (ou misturador), se para cada B e C mensuráveis tem-se

$$\lim_{j\to+\infty}\mu(B\cap\sigma^{-j}(C))=\mu(B)\mu(C).$$

A seguir mostraremos que o shift de Bernoulli satisfaz uma propriedade mais forte até do que a propriedade mixing.

Lema 1. Se B e C são uniões finitas de cilindros dois-a-dois disjuntos, então tem-se

$$\mu(B\cap\sigma^{-j}(C))=\mu(B)\mu(\sigma^{-j}(C))=\mu(B)\mu(C),$$

para todo j suficientemente grande.

Demonstração. Inicialmente, iremos supor que B e C são cilindros, sabidamente $B=[k:B_k,...,B_l]$ e $C=[m:C_m,...,C_n]$. Então,

$$\sigma^{-j}(C)=[m+j:C_m,...,C_n],$$

para cada $j\in\mathbb{N}$.

Como $\mathbb{N}$ é ilimitado superiormente, então existe um $r_0\in\mathbb{N}$ tal que $r_0>l-m$. Daí, se $j\in\mathbb{N}$ é tal que $j>r_0>l-m$, então $j>l-m$, ou seja, $m+j>l$. Logo,

$$\begin{aligned}
B\cap\sigma^{-j}(C) &= [k:B_k,...,B_l]\cap\sigma^{-j}([m:C_m,...,C_n])\\
&= [k:B_k,...,B_l]\cap[m+j:C_m,...,C_n]\\
&= \{(x_r)_r|x_k\in B_k,...,x_l\in B_l\}\cap\{(x_i)_i|x_{m+j}\in C_m,...,x_{n+j}\in C_n\}\\
&= \{(x_t)_t|x_k\in B_k,\ldots,x_l\in B_l,x_{m+j}\in C_m,\ldots,x_{n+j}\in C_n\}\\
&= \{(x_t)_t|x_k\in B_k,\ldots,x_l\in B_l,x_{l+1}\in X,\ldots,x_{m+j-1}\in X,x_{m+j}\in C_m,\ldots,x_{n+j}\in C_n\}\\
&= [k:B_k,...,B_l,X,...,X,C_m,...,C_n],
\end{aligned}$$

onde X aparece $m+j-l-1$ vezes.

Portanto, obtemos que

$$\begin{aligned}
\mu(B\cap\sigma^{-j}(C)) &= \mu([k:B_k,...,B_l,X,...,X,C_m,...,C_n])\\
&= P(B_k)\cdot ...\cdot P(B_l)\cdot P(X)\cdot ...\cdot P(X)\cdot P(C_m)\cdot ...\cdot P(C_n)\\
&= P(B_k)\cdot ...\cdot P(B_l)\cdot 1^{m+j-l-1}\cdot P(C_m)\cdot ...\cdot P(C_n)\\
&= P(B_k)\cdot ...\cdot P(B_l)\cdot P(C_m)\cdot ...\cdot P(C_n)\\
&= \mu([k:B_k,...,B_l])\mu([m:C_m,...,C_n])\\
&= \mu(B)\mu(C),
\end{aligned}$$

para todo j suficientemente grande (ou seja, para todo $j>r_0$). Isto mostra o lema para o caso particular em que B e C são cilindros.

Agora, suponhamos que

$$B=\bigcup_{k=1}^{r}B_k \text{ e } C=\bigcup_{i=1}^{s}C_i,$$

onde $B_1,...,B_r,C_1,...,C_s$ são todos cilindros e $B_i\cap B_j=\varnothing$, se $i\neq j$ e $C_i\cap C_j=\varnothing$, se $i\neq j$.

Assim,

$$\begin{aligned}
\mu(B\cap\sigma^{-j}(C)) &= \mu\left(\left(\bigcup_{k=1}^{r}B_k\right)\cap\sigma^{-j}\left(\bigcup_{i=1}^{s}C_i\right)\right)\\
&= \mu\left(\bigcup_{k=1}^{r}\left[B_k\cap\sigma^{-j}\left(\bigcup_{i=1}^{s}C_i\right)\right]\right)\\
&= \mu\left(\bigcup_{k=1}^{r}\left[B_k\cap\left(\bigcup_{i=1}^{s}\sigma^{-j}(C_i)\right)\right]\right)\\
&= \mu\left(\bigcup_{k=1}^{r}\left[\bigcup_{i=1}^{s}(B_k\cap\sigma^{-j}(C_i))\right]\right)\\
&= \mu\left(\bigcup_{(k,i)\in\Gamma}(B_k\cap\sigma^{-j}(C_i))\right).
\end{aligned}$$

Onde $\Gamma=\{1,...,r\}\times\{1,...,s\}$.

Agora, perceba que se $k \neq t$, então $B_k \cap B_t = \varnothing$. Logo,

$$\begin{aligned}
[B_k \cap \sigma^{-j}(C_i)] \cap [B_t \cap \sigma^{-j}(C_o)] &= [B_k \cap B_t] \cap [\sigma^{-j}(C_i) \cap \sigma^{-j}(C_o)] \\
&= \varnothing \cap [\sigma^{-j}(C_i) \cap \sigma^{-j}(C_o)] \\
&= \varnothing.
\end{aligned}$$

Além disto, se $i \neq o$, então $C_i \cap C_o = \varnothing$. Donde,

$$\begin{aligned}
\sigma^{-j}(C_i) \cap \sigma^{-j}(C_o) &= \sigma^{-j}(C_i \cap C_o) \\
&= \sigma^{-j}(\varnothing) \\
&= \varnothing.
\end{aligned}$$

Portanto, $B_k \cap \sigma^{-j}(C_i)$ são dois-a-dois disjuntos. Pois se $B_k \cap \sigma^{-j}(C_i)$ e $B_t \cap \sigma^{-j}(C_o)$ são tais que $(k,i) \neq (t,o)$, então $k \neq t$ ou $i \neq o$ e, em ambos os casos, temos

$$\left[B_k \cap \sigma^{-j}(C_i)\right] \cap \left[B_t \cap \sigma^{-j}(C_o)\right] = [B_k \cap B_t] \cap \left[\sigma^{-j}(C_i) \cap \sigma^{-j}(C_o)\right] = \varnothing.$$

O que nos fornece:

$$\begin{aligned}
\mu(B \cap \sigma^{-j}(C)) &= \mu\left(\bigcup_{(k,i)\in\Gamma} (B_k \cap \sigma^{-j}(C_i))\right) \\
&= \sum_{(k,i)\in\Gamma} \left(\mu(B_k \cap \sigma^{-j}(C_i))\right) \\
&= \sum_{(k,i)\in\Gamma} (\mu(B_k)\mu(C_i)), \text{ para j suficientemente grande,} \\
&= \sum_{k=1}^{r}\left(\sum_{i=1}^{s} \mu(B_k)\mu(C_i)\right) \\
&= \left(\sum_{k=1}^{r} \mu(B_k)\right)\left(\sum_{i=1}^{s} \mu(C_i)\right) \\
&= \left(\sum_{k=1}^{r} \mu(B_k)\right)\mu\left(\bigcup_{i=1}^{s} C_i\right) \\
&= \left(\sum_{k=1}^{r} \mu(B_k)\right)\mu(C) \\
&= \mu\left(\bigcup_{k=1}^{r} B_k\right)\mu(C) \\
&= \mu(B)\,\mu(C).
\end{aligned}$$

Concluímos assim que $\mu(B \cap \sigma^{-j}(C)) = \mu(B)\,\mu(C)$, para todo j suficientemente grande.

Q. E. D.

Vamos agora generalizar o lema e assim mostraremos que o shift de Bernoulli é de fato mixing. A demonstração a seguir foi retirada da demonstração do Lema 7.1.2 de [7], páginas 170 e 171:

Lema 2. Sejam f uma função mensurável e μ uma probabilidade f-invariante e suponha que

$$\lim_{n\to\infty} \mu(f^{-n}(A)\cap B) = \mu(A)\mu(B),$$

para todo par de conjuntos A e B em alguma álgebra $\mathbb{A}$ geradora da σ-álgebra dos conjuntos mensuráveis (a coleção de todas as uniões finitas de cilindros dois-a-dois disjuntos é uma tal álgebra para $\mathbb{B}$). Então (f,μ) é misturador (mixing).

Demonstração. Seja C a família de todos os conjuntos mensuráveis A tais que

$$\mu(f^{-n}(A)\cap B) \to \mu(A)\mu(B)$$

para todo $B \in \mathbb{A}$. Por hipótese, C contém $\mathbb{A}$. Afirmamos que C é uma classe monótona. De fato, sejam $A_1 \subseteq ... \subseteq A_k \subseteq ...$ elementos de C e seja $A = \bigcup_{k=1}^{+\infty} A_k$.

Como $A_k \to A$, então $\mu(A_k) \to \mu(A)$. Logo, dado $\varepsilon > 0$ existe $k_0 \geq 1$ tal que

$$|\mu(A) - \mu(A_k)| = |\mu(A - A_k)| = \mu(A - A_k) < \varepsilon$$

para todo $k \geq k_0$.

Além disso, para todo $n \geq 1$,

$$\begin{aligned} |\mu(f^{-n}(A)\cap B) - \mu(f^{-n}(A_k)\cap B)| &= |\mu(f^{-n}(A-A_k)\cap B)| \\ &= \mu(f^{-n}(A-A_k)\cap B) \\ &\leq \mu(f^{-n}(A-A_k)) \\ &= \mu(A-A_k) \\ &< \varepsilon. \end{aligned}$$

Para todo $k \geq k_0$ fixado, o fato de que $A_k \in C$ garante que existe um $n(k) \geq 1$ tal que

$$|\mu(f^{-n}(A_k)\cap B)-\mu(A_k)\mu(B)|<\varepsilon \text{ para todo } n\geq n(k).$$

Daí, concluímos que

$$\begin{aligned}\left|\mu(f^{-n}(A)\cap B)-\mu(A)\mu(B)\right| &= |\mu(f^{-n}(A)\cap B)-\mu(f^{-n}(A_k)\cap B)+\mu(f^{-n}(A_k)\cap B)-\mu(A_k)\mu(B)\\ &\quad +\mu(A_k)\mu(B)-\mu(A)\mu(B)|\\ &\leq \left|\mu(f^{-n}(A)\cap B)-\mu(f^{-n}(A_k)\cap B)\right|+\left|\mu(f^{-n}(A_k)\cap B)-\mu(A_k)\mu(B)\right|\\ &\quad +|\mu(A_k)\mu(B)-\mu(A)\mu(B)|\\ &< \varepsilon+\varepsilon+\varepsilon\cdot\mu(B)\\ &= \varepsilon\cdot(2+\mu(B)).\end{aligned}$$

Como $\varepsilon>0$ é arbitrário, isto mostra que $\mu(f^{-n}(A)\cap B)\to\mu(A)\mu(B)$. O que nos diz que $A\in C$ e, portanto, C é fechada para uniões enumeráveis monótonas.

Sejam, agora, $A_1\supseteq\ldots\supseteq A_k\supseteq\ldots$ elementos de C e seja $A=\bigcap_{k=1}^{+\infty}A_k$.

Como $A_k\to A$, então $\mu(A_k)\to\mu(A)$. Logo, dado $\varepsilon>0$ existe $k_0\geq 1$ tal que

$$|\mu(A)-\mu(A_k)|=|\mu(A-A_k)|=\mu(A-A_k)<\varepsilon$$

para todo $k\geq k_0$.

Além disso, para todo $n\geq 1$,

$$\begin{aligned}|\mu(f^{-n}(A)\cap B)-\mu(f^{-n}(A_k)\cap B)| &= |\mu(f^{-n}(A-A_k)\cap B)|\\ &= \mu(f^{-n}(A-A_k)\cap B)\\ &\leq \mu(f^{-n}(A-A_k))\\ &= \mu(A-A_k)\\ &< \varepsilon.\end{aligned}$$

Para todo $k \geq k_0$ fixado, o fato de que $A_k \in C$ garante que existe um $n(k) \geq 1$ tal que

$$|\mu(f^{-n}(A_k) \cap B) - \mu(A_k)\mu(B)| < \varepsilon \text{ para todo } n \geq n(k).$$

Daí, concluímos que

$$\begin{aligned}
\left|\mu(f^{-n}(A) \cap B) - \mu(A)\mu(B)\right| &= |\mu(f^{-n}(A) \cap B) - \mu(f^{-n}(A_k) \cap B) + \mu(f^{-n}(A_k) \cap B) - \mu(A_k)\mu(B) \\
&+ \mu(A_k)\mu(B) - \mu(A)\mu(B)| \\
&\leq \left|\mu(f^{-n}(A) \cap B) - \mu(f^{-n}(A_k) \cap B)\right| + \left|\mu(f^{-n}(A_k) \cap B) - \mu(A_k)\mu(B)\right| \\
&+ |\mu(A_k)\mu(B) - \mu(A)\mu(B)| \\
&< \varepsilon + \varepsilon + \varepsilon \cdot \mu(B) \\
&= \varepsilon \cdot (2 + \mu(B)).
\end{aligned}$$

Como $\varepsilon > 0$ é arbitrário, isto mostra que $\mu(f^{-n}(A) \cap B) \to \mu(A)\mu(B)$. Ou seja, $A \in C$ e, portanto, C é fechada para intersecções enumeráveis monótonas.

Portanto, C é uma classe monótona, tal como afirmamos. Pelo teorema das classes monótonas, segue que C contém todo conjunto mensurável: para todo conjunto mensurável A tem-se

$$\lim_{n \to +\infty} \mu(f^{-n}(A) \cap B) = \mu(A)\mu(B) \text{ para todo } B \in \mathbb{A}.$$

Resta deduzir que esta propriedade vale para todo conjunto mensurável B.

Seja, agora, C a família de todos os conjuntos mensuráveis B tais que

$$\mu(f^{-n}(A) \cap B) \to \mu(A)\mu(B)$$

para todo A mensurável. Pelo que acabamos de mostrar, C contém $\mathbb{A}$. Afirmamos que C é uma classe monótona. De fato, sejam $B_1 \subseteq \ldots \subseteq B_k \subseteq \ldots$ elementos de C e seja $B = \bigcup_{k=1}^{+\infty} B_k$.

Como $B_k \to B$, então $\mu(B_k) \to \mu(B)$. Logo, dado $\varepsilon > 0$ existe $k_0 \geq 1$ tal que

$$|\mu(B)-\mu(B_k)| = |\mu(B-B_k)| = \mu(B-B_k) < \varepsilon$$

para todo $k \geq k_0$.

Além disso, para todo $n \geq 1$,

$$\begin{aligned}
|\mu(f^{-n}(A)\cap B)-\mu(f^{-n}(A)\cap B_k)| &= |\mu(f^{-n}(A)\cap(B-B_k))| \\
&= \mu(f^{-n}(A)\cap(B-B_k)) \\
&\leq \mu(B-B_k)) \\
&< \varepsilon.
\end{aligned}$$

Para todo $k \geq k_0$ fixado, o fato de que $B_k \in C$ garante que existe um $n(k) \geq 1$ tal que

$$|\mu(f^{-n}(A)\cap B_k)-\mu(A)\mu(B_k)| < \varepsilon \text{ para todo } n \geq n(k).$$

Daí, concluímos que

$$\begin{aligned}
\left|\mu(f^{-n}(A)\cap B)-\mu(A)\mu(B)\right| &= |\mu(f^{-n}(A)\cap B)-\mu(f^{-n}(A)\cap B_k)+\mu(f^{-n}(A)\cap B_k)-\mu(A)\mu(B_k) \\
&\quad +\mu(A)\mu(B_k)-\mu(A)\mu(B)| \\
&\leq \left|\mu(f^{-n}(A)\cap B)-\mu(f^{-n}(A)\cap B_k)\right|+\left|\mu(f^{-n}(A)\cap B_k)-\mu(A)\mu(B_k)\right| \\
&\quad +|\mu(A)\mu(B_k)-\mu(A)\mu(B)| \\
&< \varepsilon+\varepsilon+\varepsilon\cdot\mu(A) \\
&= \varepsilon\cdot(2+\mu(A)).
\end{aligned}$$

Como $\varepsilon > 0$ é arbitrário, isto mostra que $\mu(f^{-n}(A)\cap B) \to \mu(A)\mu(B)$. Ou seja, $B \in C$ e, portanto, C é fechada para uniões enumeráveis monótonas.

Sejam, agora, $B_1 \supseteq \ldots \supseteq B_k \supseteq \ldots$ elementos de C e seja $B = \bigcap_{k=1}^{+\infty} B_k$.

Como $B_k \to B$, então $\mu(B_k) \to \mu(B)$. Logo, dado $\varepsilon > 0$ existe $k_0 \geq 1$ tal que

$$|\mu(B)-\mu(B_k)| = |\mu(B-B_k)| = \mu(B-B_k) < \varepsilon$$

para todo $k \geq k_0$.

Além disso, para todo $n \geq 1$,

$$\begin{aligned}
|\mu(f^{-n}(A)\cap B)-\mu(f^{-n}(A)\cap B_k)| &= |\mu(f^{-n}(A)\cap(B-B_k))| \\
&= \mu(f^{-n}(A)\cap(B-B_k)) \\
&\leq \mu(B-B_k) \\
&< \varepsilon.
\end{aligned}$$

Para todo $k \geq k_0$ fixado, o fato de que $B_k \in C$ garante que existe um $n(k) \geq 1$ tal que

$$|\mu(f^{-n}(A)\cap B_k)-\mu(A)\mu(B_k)| < \varepsilon \text{ para todo } n \geq n(k).$$

Daí, concluímos que

$$\begin{aligned}
\left|\mu(f^{-n}(A)\cap B)-\mu(A)\mu(B)\right| &= |\mu(f^{-n}(A)\cap B)-\mu(f^{-n}(A)\cap B_k)+\mu(f^{-n}(A)\cap B_k)-\mu(A)\mu(B_k) \\
&\quad +\mu(A)\mu(B_k)-\mu(A)\mu(B)| \\
&\leq \left|\mu(f^{-n}(A)\cap B)-\mu(f^{-n}(A)\cap B_k)\right|+\left|\mu(f^{-n}(A)\cap B_k)-\mu(A)\mu(B_k)\right| \\
&\quad +|\mu(A)\mu(B_k)-\mu(A)\mu(B)| \\
&< \varepsilon+\varepsilon+\varepsilon\cdot\mu(A) \\
&= \varepsilon\cdot(2+\mu(A)).
\end{aligned}$$

Como $\varepsilon > 0$ é arbitrário, isto mostra que $\mu(f^{-n}(A)\cap B) \to \mu(A)\mu(B)$. Ou seja, $B \in C$ e, portanto, C é fechada para interseções enumeráveis monótonas.

Portanto, C é uma classe monótona, tal como afirmamos. Pelo teorema das classes monótonas, segue que C contém todo conjunto mensurável: para todo conjunto mensurável B tem-se

$$\lim_{n\to+\infty} \mu(f^{-n}(A)\cap B) = \mu(A)\mu(B) \text{ para todo } A \text{ mensurável.}$$

Q.E.D.

E isto nos diz que o shift de Bernoulli é um sistema mixing. Consideremos agora a seguinte definição:

Definição 13. Consideremos um espaço de probabilidade $(X, \mathbb{X}, m)$ com uma função $T : X \to X$ $\mathbb{X}$-mensurável e cuja probabilidade m seja T-invariante. Dizemos que $A \in \mathbb{X}$ é um **mensurável invariante** (sob a ação de T), se $T^{-1}(A) = A$.

Agora fornecemos a definição de sistema ergódico:

Definição 14. Consideremos um espaço de probabilidade $(X, \mathbb{X}, m)$ com uma função $T : X \to X$ $\mathbb{X}$-mensurável e cuja probabilidade m seja T-invariante. Dizemos que o sistema (T, m) é **ergódico**, se sempre que $A \in \mathbb{X}$ é um mensurável invariante temos necessariamente que $m(A) = 0$ ou $m(A) = 1$.

Como consequência dos lemas anteriores, mostraremos que o deslocamento de Bernoulli (σ, μ) é um sistema ergódico.

Teorema 5. O shift de Bernoulli (σ, μ) é ergódico.

Demonstração. De fato, se $A \in \mathbb{B}$ é tal que $\sigma^{-1}(A) = A$, então $\sigma^{-j}(A) = A, \forall j \in \mathbb{N}$. Por um lado, a propriedade mixing nos dá que

$$\lim_{j \to +\infty} \mu\left(A \cap \sigma^{-j}(A)\right) = \mu(A) \cdot \mu(A) = \mu(A)^2.$$

Por outro lado, o fato de A ser um mensurável invariante nos diz que

$$\lim_{j \to +\infty} \mu\left(A \cap \sigma^{-j}(A)\right) = \lim_{j \to +\infty} \mu(A \cap A) = \lim_{j \to +\infty} \mu(A) = \mu(A).$$

Portanto,

$$\mu(A) = \mu(A)^2.$$

O que nos fornece $\mu(A) = 0$ ou $\mu(A) = 1$.

Q.E.D.

5 UMA APLICAÇÃO

Alguns resultados presentes aqui nesta seção não serão demonstrados, mas serão adequadamente referenciados. De agora em diante fixaremos o espaço de probabilidade $(X, \mathbb{X}, P)$ com $X = \{1, ..., d\}$, onde $d \in \mathbb{N}$ e $\mathbb{X} = \mathscr{P}(X)$ e consideraremos o espaço simbólico $(\Sigma, \mathbb{B}, \mu)$ construído a partir deste espaço de probabilidade.

Iniciamos nossa discussão com a seguinte definição importante.

Definição 15. Seja X um conjunto qualquer. Uma função $d : X \times X \to \mathbb{R}$ é uma métrica para X, se:

(M1) $d(x,y) \geq 0, \forall x,y \in X$;

(M2) $d(x,y) = 0 \Leftrightarrow x = y$;

(M3) $d(x,y) = d(y,x), \forall x,y \in X$;

(M4) $d(x,y) \leq d(x,z) + d(z,y), \forall x,y,z \in X$.

Queremos construir uma métrica para Σ, para isto fixaremos um número $\theta \in (0,1)$. A partir deste número vamos definir a seguinte função $d_\theta : \Sigma \times \Sigma \to \mathbb{R}$ por

$$d_\theta(x,y) = \sum_{k=1}^{+\infty} \theta^k \cdot (1 - \delta(x_k, y_k)),$$

onde $x, y \in \Sigma$, $x = (x_1, x_2, x_3, x_4, ...)$, $y = (y_1, y_2, y_3, y_4, ...)$ e

$$\delta(x_k, y_k) = \begin{cases} 1, \text{ se } x_k = y_k \\ 0, \text{ se } x_k \neq y_k \end{cases}$$

é o delta de Kronecker.

Teorema 6. A função $d_\theta : \Sigma \times \Sigma \to \mathbb{R}$ é uma métrica para Σ.

Demonstração. Primeiro, percebamos que esta função está bem-definida, isto é, a série apresentada é convergente, pois $0 < \theta < 1$ e, portanto, $0 < \theta^k < 1, \forall k \in \mathbb{N}$. Além disto, da definição do delta de Kronecker $\delta(x_k, y_k) \in \{0,1\}, \forall k \in \mathbb{N}$. Assim, $1 - \delta(x_k, y_k) \in \{0,1\}$. O que nos diz que

$$0 \leq 1 - \delta(x_k, y_k) \leq 1, \forall k \in \mathbb{N}.$$

Logo,

$$0 \leq \theta^k \cdot (1 - \delta(x_k, y_k)) \leq \theta^k, \forall k \in \mathbb{N}.$$

Assim,

$$\sum_{k=1}^{m} \theta^k \cdot (1 - \delta(x_k, y_k)) \leq \sum_{k=1}^{m} \theta^k, \forall m \in \mathbb{N}.$$

Como $0 < \theta < 1$, então a série geométrica $\sum_{k=1}^{+\infty} \theta^k$ é convergente (ver exemplo 24 de [9], páginas 135 e 136) e, portanto, via critério de comparação (corolário do teorema 16 de [9], página 137) obtemos que $\sum_{k=1}^{+\infty} \theta^k \cdot (1 - \delta(x_k, y_k))$ é convergente e mostramos também que (M1) é valida, isto é, $d_\theta(x, y) \geq 0, \forall x, y \in \Sigma$.

Vemos que para cada $x \in \Sigma$, se $x = (x_1, x_2, x_3, \ldots)$, então

$$d_\theta(x, x) = \sum_{k=1}^{+\infty} \theta^k \cdot (1 - \delta(x_k, x_k)).$$

Por definição do delta de Kronecker, temos que

$$\delta(x_k, x_k) = 1, \forall k \in \mathbb{N}.$$

Portanto,

$$1 - \delta(x_k, x_k) = 0, \forall k \in \mathbb{N}.$$

O que nos fornece

$$\begin{aligned} \sum_{k=1}^{+\infty} \theta^k \cdot (1 - \delta(x_k, x_k)) &= \sum_{k=1}^{+\infty} \theta^k \cdot 0 \\ &= \sum_{k=1}^{+\infty} 0 \\ &= 0. \end{aligned}$$

Ou seja, $d_\theta(x, x) = 0, \forall x \in \Sigma$. Por outro lado, sejam $x, y \in \Sigma, x = (x_1, x_2, x_3, \ldots), y = (y_1, y_2, y_3, \ldots)$ tais que

$$d_\theta(x,y)=0.$$

Então,

$$\sum_{k=1}^{+\infty}\theta^k\cdot(1-\delta(x_k,y_k))=0.$$

Ou ainda,

$$\lim_{m\to+\infty}\sum_{k=1}^{m}\theta^k\cdot(1-\delta(x_k,y_k))=0. \tag{5.1}$$

Agora, lembramos que $\theta^k>0$ e $1-\delta(x_k,y_k)\geq 0, \forall k\in\mathbb{N}$. Portanto,

$$\theta^k\cdot(1-\delta(x_k,y_k))\geq 0, \forall k\in\mathbb{N}. \tag{5.2}$$

Ora, (5.2) nos fornece que a sequência $\left(\sum_{k=1}^{m}\theta^k\cdot(1-\delta(x_k,y_k))\right)_{m\in\mathbb{N}}$ é uma sequência monótona não-decrescente, enquanto que (5.1) nos fornece que a mesma sequência é limitada (vide teorema 3 de [9], página 110) e, portanto, via teorema 4 de [9], página 111, temos que

$$\lim_{m\to+\infty}\sum_{k=1}^{m}\theta^k\cdot(1-\delta(x_k,y_k))=\sup\left\{\sum_{k=1}^{m}\theta^k\cdot(1-\delta(x_k,y_k))\in\mathbb{R}\mid m\in\mathbb{N}\right\}. \tag{5.3}$$

Portanto, de (5.1) e (5.3) obtemos que

$$\sup\left\{\sum_{k=1}^{m}\theta^k\cdot(1-\delta(x_k,y_k))\in\mathbb{R}\mid m\in\mathbb{N}\right\}=0.$$

Daí, pela definição de supremo (cota superior) e lembrando de (5.2) obtemos

$$0\leq\sum_{k=1}^{m}\theta^k\cdot(1-\delta(x_k,y_k))\leq 0, \forall m\in\mathbb{N}.$$

O que nos diz que

$$\sum_{k=1}^{m}\theta^k\cdot(1-\delta(x_k,y_k))=0, \forall m\in\mathbb{N}.$$

Consequentemente, por (5.2) (uma soma de números não-negativos resulta em 0 se, e somente se, cada número for 0) ficamos com

$$\theta^k \cdot (1 - \delta(x_k, y_k)) = 0, \forall k \in \mathbb{N}.$$

Lembramos que $\theta^k > 0, \forall k \in \mathbb{N}$ e obtemos, então, que

$$1 - \delta(x_k, y_k) = 0, \forall k \in \mathbb{N}.$$

Ou ainda,

$$\delta(x_k, y_k) = 1, \forall k \in \mathbb{N}.$$

Mas, da definição do delta de Kronecker, isto significa que $x_k = y_k, \forall k \in \mathbb{N}$. Em outras palavras, $x = y$. Provamos (M2), $d_\theta(x,y) = 0 \Leftrightarrow x = y$.

Agora, percebamos que para $x, y \in \Sigma, x = (x_1, x_2, x_3, \ldots), y = (y_1, y_2, y_3, \ldots)$ temos que

$$\begin{aligned} d_\theta(x,y) &= \sum_{k=1}^{+\infty} \theta^k \cdot (1 - \delta(x_k, y_k)) \\ &= \sum_{k=1}^{+\infty} \theta^k \cdot (1 - \delta(y_k, x_k)) \\ &= d_\theta(y,x). \end{aligned}$$

O que prova (M3). Consideremos agora $x, y, z \in \Sigma, x = (x_1, x_2, x_3, \ldots), y = (y_1, y_2, y_3, \ldots), z = (z_1, z_2, z_3, \ldots)$ e analisaremos se a inequação a seguir pode ser falsa

$$1 - \delta(x_k, y_k) \leq [1 - \delta(x_k, z_k)] + [1 - \delta(z_k, y_k)].$$

Como o delta de Kronecker só resulta em 0 ou 1, então o termo do lado esquerdo da inequação só pode ser 0 ou 1. Já a soma no lado direito só pode ser 0, 1 ou 2. Se o termo no lado esquerdo for 0, não importa o que ocorra no lado direito, a inequação sempre será válida. Agora, se o termo no lado esquerdo for 1 o único caso problemático é o caso em que a soma no lado direito é igual a 0. Mas a soma no lado direito só é 0 se os dois somandos forem iguais a 0. Ou seja, neste caso problemático devemos ter:

$$\begin{cases} 1-\delta\left(x_k,y_k\right)=1 \\ 1-\delta\left(x_k,z_k\right)=0 \\ 1-\delta\left(z_k,y_k\right)=0 \end{cases}.$$

Mas isto é o mesmo que

$$\begin{cases} \delta\left(x_k,y_k\right)=0 \\ \delta\left(x_k,y_k\right)=1 \\ \delta\left(x_k,y_k\right)=1 \end{cases}.$$

Ou seja,

$$\begin{cases} x_k \neq y_k \\ x_k = z_k \\ z_k = y_k \end{cases}.$$

Mas isto é um absurdo, pois se $x_k = z_k$ e $z_k = y_k$, então $x_k = y_k$. Assim, devemos ter $x_k \neq y_k$ e $x_k = y_k$. Portanto, este caso problemático nunca ocorre. Concluímos assim que

$$1-\delta\left(x_k,y_k\right) \leq \left[1-\delta\left(x_k,z_k\right)\right]+\left[1-\delta\left(z_k,y_k\right)\right], \forall k \in \mathbb{N}.$$

Logo, se multiplicamos ambos os membros desta desigualdade pelo número positivo θ^k obtemos

$$\theta^k \cdot \left(1-\delta\left(x_k,y_k\right)\right) \leq \theta^k \cdot \left(1-\delta\left(x_k,z_k\right)\right)+\theta^k \cdot \left(1-\delta\left(z_k,y_k\right)\right), \forall k \in \mathbb{N}.$$

E, portanto, ficamos com

$$\sum_{k=1}^{+\infty} \theta^k \cdot \left(1-\delta\left(x_k,y_k\right)\right) \leq \sum_{k=1}^{+\infty} \theta^k \cdot \left(1-\delta\left(x_k,z_k\right)\right)+\sum_{k=1}^{+\infty} \theta^k \cdot \left(1-\delta\left(z_k,y_k\right)\right).$$

Ou, em outras palavras,

$$d_\theta(x,y) \leq d_\theta(x,z) + d_\theta(z,y), \forall x,y,z \in \Sigma.$$

(M4) está provada. Consequentemente, $d_\theta : \Sigma \times \Sigma \to \mathbb{R}$ é uma métrica em Σ, como afirmamos.

Q.E.D.

Vamos considerar agora o conjunto

$$F_\theta = \{f : \Sigma \to \mathbb{R} \mid \exists k \in \mathbb{R}, |f(x) - f(y)| \leq k \cdot d_\theta(x,y), \forall x,y \in \Sigma\}$$

de todas as funções (de valores reais) de Lipschitz definidas em Σ. Vamos mostrar que F_θ com a soma usual de funções e o produto de uma função por um escalar real usual definem um subespaço vetorial do espaço vetorial real $(\mathscr{F}(\Sigma;\mathbb{R}), +, \cdot, \mathbb{R})$, onde $\mathscr{F}(\Sigma;\mathbb{R})$ é o conjunto de todas as funções $f : \Sigma \to \mathbb{R}$.

Teorema 7. F_θ é um subespaço vetorial de $\mathscr{F}(\Sigma;\mathbb{R})$.

Demonstração. Primeiro, notamos que a função nula $n : \Sigma \to \mathbb{R}$ é um elemento de F_θ. De fato, dados $x,y \in \Sigma$ temos que

$$|n(x) - n(y)| = |0 - 0| = 0 \leq d_\theta(x,y) = 1 \cdot d_\theta(x,y).$$

Portanto, com $n \in F_\theta$. Agora, sejam $f,g \in F_\theta$ e $\alpha \in \mathbb{R}$. Então, existem $H,K \geq 0$ tais que

$$\begin{cases} |f(x) - f(y)| \leq H \cdot d_\theta(x,y), \forall x,y \in \Sigma \\ |g(x) - g(y)| \leq K \cdot d_\theta(x,y), \forall x,y \in \Sigma \end{cases}.$$

Desta forma, vemos que para quaisquer $x,y \in \Sigma$ valem

$$
\begin{aligned}
|(f+\alpha\cdot g)(x)-(f+\alpha\cdot g)(y)| &= |f(x)+\alpha\cdot g(x)-f(y)-\alpha\cdot g(y)| \\
&\leq |f(x)-f(y)|+|\alpha\cdot(g(x)-g(y))| \\
&\leq |f(x)-f(y)|+|\alpha|\cdot|(g(x)-g(y))| \\
&\leq H\cdot d_\theta(x,y)+|\alpha|\cdot K\cdot d_\theta(x,y) \\
&= (H+|\alpha|\cdot K)\, d_\theta(x,y).
\end{aligned}
$$

Como $H,K,|\alpha|\geq 0$, então $H+|\alpha|\cdot K\geq 0$ e, assim, $f+\alpha\cdot g\in F_\theta, \forall f,g\in F_\theta, \forall\alpha\in\mathbb{R}$. Consequentemente, $(F_\theta,+,\cdot,\mathbb{R})$ é um subespaço vetorial de $(\mathscr{F}(\Sigma;\mathbb{R}),+,\cdot,\mathbb{R})$.

Q.E.D.

Se $f\in F_\theta$, então existe um $k\in\mathbb{N}$ tal que

$$|f(x)-f(y)|\leq k\cdot d_\theta(x,y), \forall x,y\in\Sigma.$$

Logo, se $x\neq y$, ficamos com

$$\frac{|f(x)-f(y)|}{d_\theta(x,y)}\leq k.$$

Ou seja, k é uma cota superior para o conjunto $A_f=\left\{\frac{|f(x)-f(y)|}{d_\theta(x,y)}\in\mathbb{R}\mid x,y\in\Sigma, x\neq y\right\}$. E pela completude de $\mathbb{R}$ (veja [9] página 80, logo abaixo do exemplo 14) existe o número $\sup A_f$. Desta forma, podemos definir a seguinte função $|\cdot|_\theta : F_\theta\to\mathbb{R}$ por

$$|f|_\theta=\sup\left\{\frac{|f(x)-f(y)|}{d_\theta(x,y)}\in\mathbb{R}\mid x,y\in\Sigma, x\neq y\right\}, \forall f\in F_\theta.$$

O número $|f|_\theta$ é chamado de **a melhor constante de Lipschitz da função** f.

Por outro lado, se fixamos $y\in\Sigma$, então ficamos com

$$|f(x)-f(y)|\leq k\cdot d_\theta(x,y), \forall x\in\Sigma.$$

Como

$$|f(x)| - |f(y)| \leq |f(x) - f(y)|, \forall x \in \Sigma$$

então

$$|f(x)| - |f(y)| \leq k \cdot d_\theta(x,y), \forall x \in \Sigma.$$

O que nos fornece:

$$|f(x)| \leq k \cdot d_\theta(x,y) + |f(y)|, \forall x \in \Sigma.$$

Lembramos que

$$d_\theta(x,y) = \sum_{k=1}^{+\infty} \theta^k \cdot (1 - \delta(x_k, y_k)).$$

Como $1 - \delta(x_k, y_k) \leq 1$, então

$$d_\theta(x,y) \leq \sum_{k=1}^{+\infty} \theta^k = \frac{1}{1-\theta}.$$

Ou seja,

$$|f(x)| \leq k \cdot \frac{1}{1-\theta} + |f(y)|, \forall x \in \Sigma.$$

Como $y \in \Sigma$ está fixado, então o termo $k \cdot \dfrac{1}{1-\theta} + |f(y)|$ é um número real constante e é uma cota superior para o conjunto $B_f = \{|f(x)| \in \mathbb{R} \mid x \in \Sigma\}$. Novamente pela completude de $\mathbb{R}$ existe o número real $\sup B_f$. Isto nos permite definir a seguinte função $|\cdot|_\infty : F_\theta \to \mathbb{R}$ por

$$|f|_\infty = \sup\{|f(x)| \in \mathbb{R} \mid x \in \Sigma\}, \forall f \in F_\theta.$$

Vamos dar uma definição importante para o que se seguirá em nossa discussão.

Definição 16. Seja X um espaço vetorial sobre $\mathbb{R}$. Uma função $n : X \to \mathbb{R}$ é uma **norma em** X, se:

(N1) $n(x) \geq 0, \forall x \in X$;

(N2) $n(x) = 0 \Leftrightarrow x = 0$;

(N3) $n(\alpha x) = |\alpha| \cdot n(x), \forall x \in X, \alpha \in \mathbb{R}$;

(N4) $n(x+y) \leq n(x) + n(y), \forall x, y \in X$.

Queremos construir uma norma para o espaço F_θ, então vamos definir a função $||\cdot||_\theta : F_\theta \to \mathbb{R}$ por

$$||f||_\theta = |f|_\theta + |f|_\infty, \forall f \in F_\theta$$

e mostraremos a seguir que a função $||\cdot||_\theta$ é realmente uma norma em F_θ.

Teorema 8. $||\cdot||_\theta$ é uma norma em F_θ.

Demonstração. Para isto, começamos notando que para cada $f \in F_\theta$ temos que

$$|f|_\theta = \sup\left\{ \frac{|f(x) - f(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x, y \in \Sigma, x \neq y \right\} \geq \frac{|f(x) - f(y)|}{d_\theta(x,y)} \geq 0,$$

$\forall x, y \in \Sigma, x \neq y$ e,

$$|f|_\infty = \sup\{|f(x)| \in \mathbb{R} \mid x \in \Sigma\} \geq |f(x)| \geq 0, \forall x \in \Sigma.$$

Portanto,

$$||f||_\theta = |f|_\theta + |f|_\infty \geq 0, \forall f \in F_\theta.$$

O que prova (N1). Além disto, se $f \in F_\theta$ for tal que

$$||f||_\theta = 0,$$

então

$$|f|_\theta + |f|_\infty = 0.$$

Logo, devemos ter

$$\begin{cases} |f|_\theta = 0 \\ |f|_\infty = 0 \end{cases}.$$

O que vai nos fornecer (pela definição de supremo - cota superior) que

$$\begin{cases} \dfrac{|f(x) - f(y)|}{d_\theta(x,y)} = 0, \forall x, y \in \Sigma, x \neq y \\ |f(z)| = 0, \forall z \in \Sigma \end{cases}.$$

Ou ainda,

$$\begin{cases} f(x) = f(y), \forall x, y \in \Sigma \\ f(z) = 0, \forall z \in \Sigma \end{cases}.$$

Portanto, se $f \in F_\theta$ é tal que $||f||_\theta = 0$, então $f \equiv 0$. Por outro lado, se $f \equiv 0$, então

$$\begin{aligned} |f|_\theta &= \sup\left\{\frac{|f(x) - f(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x, y \in \Sigma, x \neq y\right\} \\ &= \sup\left\{\frac{|0-0|}{d_\theta(x,y)} \in \mathbb{R} \mid x, y \in \Sigma, x \neq y\right\} \\ &= \sup\left\{\frac{0}{d_\theta(x,y)} \in \mathbb{R} \mid x, y \in \Sigma, x \neq y\right\} \\ &= \sup\{0\} \\ &= 0. \end{aligned}$$

E

$$\begin{aligned}|f|_\infty &= \sup\{|f(x)| \in \mathbb{R} \mid x \in \Sigma\}\\ &= \sup\{|0| \in \mathbb{R} \mid x \in \Sigma\}\\ &= \sup\{0\}\\ &= 0.\end{aligned}$$

Portanto,

$$||f||_\theta = |f|_\theta + |f|_\infty = 0 + 0 = 0.$$

Ou seja, mostramos que $||f||_\theta = 0 \Leftrightarrow f \equiv 0$, o que prova (N2). Agora, seja $\varepsilon \in \mathbb{R}$. Então (utilizando o lema 2 de [11], página 118):

$$\begin{aligned}||\varepsilon \cdot f||_\theta &= |\varepsilon \cdot f|_\theta + |\varepsilon \cdot f|_\infty\\ &= \sup\left\{\frac{|\varepsilon \cdot f(x) - \varepsilon \cdot f(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y \in \Sigma, x \neq y\right\} + \sup\{|\varepsilon \cdot f(x)| \in \mathbb{R} \mid x \in \Sigma\}\\ &= \sup\left\{\frac{|\varepsilon \cdot (f(x) - f(y))|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y \in \Sigma, x \neq y\right\} + \sup\{|\varepsilon| \cdot |f(x)| \in \mathbb{R} \mid x \in \Sigma\}\\ &= \sup\left\{|\varepsilon| \cdot \frac{|f(x) - f(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y \in \Sigma, x \neq y\right\} + |\varepsilon| \cdot \sup\{|f(x)| \in \mathbb{R} \mid x \in \Sigma\}\\ &= |\varepsilon| \cdot \sup\left\{\frac{|f(x) - f(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y \in \Sigma, x \neq y\right\} + |\varepsilon| \cdot \sup\{|f(x)| \in \mathbb{R} \mid x \in \Sigma\}\\ &= |\varepsilon| \cdot \left[\sup\left\{\frac{|f(x) - f(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y \in \Sigma, x \neq y\right\} + \sup\{|f(x)| \in \mathbb{R} \mid x \in \Sigma\}\right]\\ &= |\varepsilon| \cdot [|f|_\theta + |f|_\infty]\\ &= |\varepsilon| \cdot ||f||_\theta\,.\end{aligned}$$

Portanto, $||\varepsilon \cdot f||_\theta = |\varepsilon| \cdot ||f||_\theta\,, \forall f \in F_\theta, \forall \varepsilon \in \mathbb{R}$ e (N3) está provada. Por fim, sejam $f, g \in F_\theta$. Vemos que (utilizando os lemas 1 e 2 de [11], páginas 117 e 118):

$$\begin{aligned}
||f+g|| &= |\varepsilon \cdot f+g|_\theta + |\varepsilon \cdot f+g|_\infty \\
&= \sup\left\{\frac{|(f+g)(x)-(f+g)(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y\in\Sigma, x\neq y\right\} + \sup\left\{|(f+g)(x)| \in \mathbb{R} \mid x\in\Sigma\right\} \\
&= \sup\left\{\frac{|f(x)+g(x)-f(y)-g(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y\in\Sigma, x\neq y\right\} + \sup\left\{|f(x)+g(x)| \in \mathbb{R} \mid x\in\Sigma\right\} \\
&\leq \sup\left\{\frac{|f(x)-f(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y\in\Sigma, x\neq y\right\} + \sup\left\{\frac{|g(x)-g(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y\in\Sigma, x\neq y\right\} \\
&+ \sup\left\{|f(x)| \in \mathbb{R} \mid x\in\Sigma\right\} + \sup\left\{|g(x)| \in \mathbb{R} \mid x\in\Sigma\right\} \\
&= \left[\sup\left\{\frac{|f(x)-f(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y\in\Sigma, x\neq y\right\} + \sup\left\{|f(x)| \in \mathbb{R} \mid x\in\Sigma\right\}\right] \\
&+ \left[\sup\left\{\frac{|g(x)-g(y)|}{d_\theta(x,y)} \in \mathbb{R} \mid x,y\in\Sigma, x\neq y\right\} + \sup\left\{|g(x)| \in \mathbb{R} \mid x\in\Sigma\right\}\right] \\
&= \left[|f|_\theta + |f|_\infty\right] + \left[|g|_\theta + |g|_\infty\right] \\
&= ||f||_\theta + ||g||_\theta .
\end{aligned}$$

Ou seja, $||f+g|| \leq ||f||_\theta + ||g||_\theta, \forall f,g \in F_\theta$. Mostramos (N4). Consequentemente, $||\cdot||_\theta$ é uma norma para o espaço vetorial $(F_\theta, +, \cdot, \mathbb{R})$.

Q.E.D.

Teorema 9. $(F_\theta, +, \cdot, \mathbb{R}, ||\cdot||)$ é um espaço de Banach.

Demonstração. Confira [10], teorema 5.4, páginas 42 e 43.

Q.E.D.

Vamos considerar também os espaços $L^1(\Sigma, \mathbb{B}, \mu)$ e $L^\infty(\Sigma, \mathbb{B}, \mu)$ de todas as classes de μ-equivalência de funções integráveis e de todas as classes de funções $\mathbb{B}$-mensuráveis limitadas em μ-quase todos os pontos (consulte [1], definições 6.6 e 6.15, páginas 54 e 60), respectivamente.

Teorema 10. $L^1(\Sigma, \mathbb{B}, \mu)$ é um espaço de Banach.

Demonstração. Consulte [1], teorema 6.14, páginas 59 e 60.

Q.E.D.

Teorema 11. $L^{\infty}(\Sigma, \mathbb{B}, \mu)$ é um espaço de Banach.

Demonstração. Consulte [10], teorema 5.2, páginas 40 e 41.

Q.E.D.

Nós agora daremos mais algumas definições importantes.

Definição 17. Sejam X e Y espaços vetoriais normados e $T : X \to Y$ um operador linear. Dizemos que T é um operador linear compacto, se para cada $M \subseteq X$ limitado tem-se que o fecho $\overline{T(M)}$ da imagem de M é um subconjunto compacto de Y.

Definição 18. Dizemos que um espaço normado F é compactamente imerso num espaço normado L, se $F \subseteq L$ e a inclusão $j : F \to L$, $j(x) = x, \forall x \in F$, é um operador linear compacto.

Se definirmos a aplicação $i : F_\theta \to L^{\infty}(\Sigma, \mathbb{B}, \mu)$ por

$$i(f) = [f], \forall f \in F_\theta,$$

então podemos definir também:

(E1) $||[f]||_\theta = \inf\{||g||_\theta \in \mathbb{R} \mid g \in [f]\}, \forall f \in F_\theta$;
(E2) $[f] + [g] = [f+g], \forall f, g \in F_\theta$;
(E3) $\alpha \cdot [f] = [\alpha \cdot f], \forall f \in F_\theta, \forall \alpha \in \mathbb{R}$;
(E4) $[F_\theta] = \{[f] \in L^{\infty}(\Sigma, \mathbb{B}, \mu) \mid f \in F_\theta\}$.

E a partir de todas estas definições, pode-se mostrar que $[F_\theta]$, com as operações definidas acima por (E2) e (E3), é um subespaço vetorial de $L^{\infty}(\Sigma, \mathbb{B}, \mu)$, normado com norma definida por (E1), e vale o seguinte resultado.

Teorema 12. $[F_\theta]$ é compactamente imerso em $L^{\infty}(\Sigma, \mathbb{B}, \mu)$.

Demonstração. Consulte [10], teorema 5.9, páginas 44 e 45. Consulte também teorema 8.1-3 de [12], página 407.

Q.E.D.

Daqui em diante identificaremos F_θ com $[F_\theta]$ assim como identificaremos $f \in F_\theta$ com $[f] \in [F_\theta]$, e o faremos sem mais comentários.

Definição 19. Seja agora $P_\sigma : L^1(\Sigma, \mathbb{B}, \mu) \to L^1(\Sigma, \mathbb{B}, \mu)$ o operador de Perron-Frobenius associado ao shift de Bernoulli, isto é, o único operador linear tal que

$$\int (\varphi \circ \sigma)\psi d\mu = \int \varphi P_\sigma(\psi) d\mu,$$

para cada $\varphi \in L^\infty(\Sigma, \mathbb{B}, \mu)$, $\psi \in L^1(\Sigma, \mathbb{B}, \mu)$. Para uma demonstração da existência de tal operador consulte o teorema 3.7 de [10], páginas 23 - 26.

Sabemos que F_θ é P_σ-invariante (consequência da proposição 5.11 de [10], páginas 45, 46 e 47), portanto faz sentido estudarmos a sua restrição ao conjunto F_θ, isto é, $P_\sigma : F_\theta \to F_\theta$.

Teorema 13. (Desigualdade de Lasota-Yorke)
Existem $\alpha \in [0,1)$ e $C \geq 0$ tais que para cada $\varphi \in F_\theta$ vale a desigualdade

$$\|P_\sigma \varphi\|_\theta \leq \alpha \|\varphi\|_\theta + C \|\varphi\|_\infty.$$

Demonstração. Consulte a proposição 5.11 de [10], páginas 45 - 47.

Q.E.D.

Um corolário deste teorema nos será importante.

Teorema 14. (Corolário de Lasota-Yorke)
Existem $\alpha \in [0,1)$ e $D \in \mathbb{R}$ tais que para cada $n \in \mathbb{N}$ vale a desigualdade

$$\|P_\sigma^n \varphi\|_\theta \leq \alpha^n \|\varphi\|_\theta + D \|\varphi\|_\infty.$$

Demonstração. Consulte o corolário 5.12 de [10], página 47.

Q.E.D.

Teorema 15. (Teorema da limitação uniforme de Banach-Steinhaus)
Sejam C um espaço de Banach e D um espaço normado e seja $(T_n)_{n\in\mathbb{N}}$ uma sequência de operadores lineares limitados $T_n : C \to D, \forall n \in \mathbb{N}$. Se para cada $x \in C$ existe um $C_x > 0$ tal que

$$\sup\{\|T_n(x)\| \in \mathbb{R} \mid n \in \mathbb{N}\} \leq C_x,$$

então existe $K > 0$ tal que

$$\sup\{\|T_n\| \in \mathbb{R} \mid n \in \mathbb{N}\} \leq K.$$

Demonstração. Consulte o teorema 4.7-3 de [12], página 249.

Q.E.D.

Teorema 16. (Teorema de Ionescu-Tulcea e Marinescu)
Sejam $(F, \|\cdot\|_F)$ e $(L, \|\cdot\|_L)$ espaços de Banach tais que $F \subseteq L$ e seja $P : F \to L$ um operador linear limitado sobre as normas $\|\cdot\|_F$ e $\|\cdot\|_L$. Suponhamos:

A) Se $(f_n)_{n\in\mathbb{N}}$ é uma sequência em $F, f \in L, \lim_{n\to\infty} \|f_n - f\|_L = 0$ e existe $C > 0$ tal que $\|f_n\|_F \leq C, \forall n \in \mathbb{N}$, então $f \in F$ e $\|f\|_F \leq C$;

B) O conjunto $\{\|P^n f\|_L \in \mathbb{R} \mid n \in \mathbb{N} \cup \{0\}, f \in F, \|f\|_L \leq 1\}$ é limitado;

C) Existem $n_0 \in \mathbb{N}, \alpha \in [0,1), \beta \in \mathbb{R}$ tais que para cada $f \in F$ vale

$$\|P^{n_0} f\|_F < \alpha \|f\|_F + \beta \|f\|_L;$$

D) Se V é um conjunto limitado no espaço $(F, \|\cdot\|_F)$, então para cada $n \in \mathbb{N}$ o fecho $\overline{P^n V}$ de sua imagem pelo operador $P^n : F \to L$ é compacto no espaço $(L, \|\cdot\|_L)$.

Se estas hipóteses são satisfeitas, então existe um número finito de autovalores $\lambda_1, ..., \lambda_k \in \mathbb{R}$ de $P : F \to F$ e subespaços vetoriais de dimensão finita $E_{\lambda_1}, ..., E_{\lambda_k}$ de F tais que $E_{\lambda_i} = \{\varphi \in F \mid P(\varphi) = \lambda_i \varphi\}, \forall i = 1, ..., k$. Além disto, $|\lambda_i| = 1, \forall i = 1, ..., k$. E também existem operadores lineares $P_i : F \to E_{\lambda_i}, i = 1, ..., k$ tais que para cada $n \in \mathbb{N}$ temos

$$P^n = \sum_{i=1}^{k} \lambda_i^n P_i + N^n,$$

onde P_i é uma projeção sobre $E_{\lambda_i}, \|P_i\| \leq 1, \forall i = 1, ..., k$ e $N : F \to F$ é um operador linear com raio espectral $\rho(N) < 1$. Por fim, vale que $P_iP_j = P_jP_i = 0, i \neq j, P_i^2 = P_i, \forall i = 1, ..., k$ e $P_iN = NP_i = 0, \forall i = 1, ..., k$.

Demonstração. Este resultado é demonstrado no artigo original [5].

Q.E.D.

Teorema 17. O operador P_σ de Perron-Frobenius associado ao shift de Bernoulli satisfaz as hipóteses A), B), C) e D) do Teorema de Ionescu-Tulcea e Marinescu. Onde o operador está definido como $P_\sigma : F_\theta \to F_\theta, P_\sigma : L^\infty(\Sigma, \mathbb{B}, \mu) \to L^\infty(\Sigma, \mathbb{B}, \mu)$ e $P_\sigma : F_\theta \to L^\infty(\Sigma, \mathbb{B}, \mu)$.

Demonstração. A) é satisfeita pelo teorema 9, em que F_θ está compactamente imerso em $L^\infty(\Sigma, \mathbb{B}, \mu)$. Agora para mostrar B) veremos que P_σ é uniformemente limitado. Para isto, vemos que se $n \in \mathbb{N}$ e $\varphi \in F_\theta$ satisfaz $\|\varphi\|_\infty \leq 1$, então

$$\begin{aligned} \|P_\sigma^n \varphi\|_\infty &\leq \|P_\sigma^n\| \cdot \|\varphi\|_\infty \\ &\leq \|P_\sigma^n\| \cdot 1 \\ &= \|P_\sigma^n\|, \end{aligned}$$

percebamos que para cada $\varphi \in F_\theta$ o número $\|P_\sigma^n\|$ é uma constante que depende de $n \in \mathbb{N}$ e não depende de φ, portanto, o teorema da limitação uniforme de Banach-Steinhaus nos garante que existe um número $K > 0$ que para todo $n \in \mathbb{N}$ satisfaz

$$\|P_\sigma^n\| \leq K.$$

Consequentemente, obtemos que

$$\|P_\sigma^n \varphi\|_\infty \leq K,$$

para todo $n \in \mathbb{N}$ e toda $\varphi \in F_\theta$ que satisfaça $\|\varphi\|_\infty \leq 1$. O que significa que o conjunto

$$\{\|P_\sigma^n \varphi\|_\infty \in \mathbb{R} \mid n \in \mathbb{N} \cup \{0\}, \varphi \in F_\theta, \|\varphi\|_\infty \leq 1\}$$

é limitado. O que prova B).

O item C) decorre da desigualdade de Lasota-Yorke. Agora para o item D) consideremos V um subconjunto limitado de F_θ. Vamos mostrar que $\overline{P_\sigma^n V}$ é compacto em $L^\infty(\Sigma, \mathbb{B}, \mu)$. Para isto, consideremos uma sequência $(P_\sigma^n \varphi_k)_{k\in\mathbb{N}}$ em $P_\sigma^n V$.

Como V é limitado em F_θ e F_θ está compactamente imerso em $L^\infty(\Sigma, \mathbb{B}, \mu)$, então $(\varphi_k)_{k\in\mathbb{N}}$ admite uma subsequência $(\varphi_{k_j})_{j\in\mathbb{N}}$ que converge para alguma φ em $L^\infty(\Sigma, \mathbb{B}, \mu)$. Logo, como P_σ é uma contração fraca, e portanto é um operador linear contínuo, então $(P_\sigma^n \varphi_{k_j})_{j\in\mathbb{N}}$ é uma subsequência convergente de $(P_\sigma^n \varphi_k)_{k\in\mathbb{N}}$ e converge para $P_\sigma^n \varphi$.

Mostramos assim que toda sequência em $P_\sigma^n V$ possui uma subsequência convergente para um elemento de $P_\sigma^n V$. Pois bem, seja agora uma sequência $(\psi_m)_{m\in\mathbb{N}}$ em $\overline{P_\sigma^n V}$. Então para cada ψ_m existe uma sequência $(P_\sigma \varphi_j)_{j\in\mathbb{N}}$ em $P_\sigma^n V$ que converge para ψ_m. Pelo que acabamos de mostrar, alguma subsequência de $(P_\sigma \varphi_j)_{j\in\mathbb{N}}$ converge para um elemento de $P_\sigma^n V$, mas como $(P_\sigma \varphi_j)_{j\in\mathbb{N}}$ converge para ψ_m, então todas as suas subsequências também convergem para ψ_m. Logo, $\psi_m \in P_\sigma^n V$. Ou seja, $(\psi_m)_{m\in\mathbb{N}}$ é uma sequência em $P_\sigma^n V$ e, portanto, possui uma subsequência convergente para um elemento de $P_\sigma^n V \subseteq \overline{P_\sigma^n V}$. Então mostramos que toda sequência em $\overline{P_\sigma^n V}$ admite uma subsequência convergente em $\overline{P_\sigma^n V}$. O que significa que $\overline{P_\sigma^n V}$ é sequencialmente compacto e, como $L^\infty(\Sigma, \mathbb{B}, \mu)$ é um espaço métrico, $\overline{P_\sigma^n V}$ é compacto em $L^\infty(\Sigma, \mathbb{B}, \mu)$. O que prova o item D).

Q.E.D.

Teorema 18. Seja P_σ o operador de Perron-Frobenius associado ao shift de Bernoulli e sejam $\lambda_1, \lambda_2, ..., \lambda_k \in \mathbb{R}$ os seus autovalores. Então, por (σ, μ) ser mixing, temos que $\lambda_1 = ... = \lambda_k = 1$.

Demonstração. Sabemos que para quaisquer mensuráveis $A, B \in \mathbb{B}$ tem-se

$$\lim_{j\to+\infty} \mu\left(A \cap \sigma^{-j}(B)\right) = \mu(A)\,\mu(B).$$

Ou seja, por definição de integral de Lebesgue,

$$\lim_{j\to+\infty} \int \chi_{A\cap\sigma^{-j}(B)} d\mu = \int \chi_A d\mu \int \chi_B d\mu.$$

$$\Rightarrow \lim_{j\to+\infty} \int \chi_A \cdot \chi_{\sigma^{-j}(B)} d\mu = \int \chi_A d\mu \int \chi_B d\mu.$$

$$\Rightarrow \lim_{j\to+\infty} \int \chi_A \cdot \left(\chi_B \circ \sigma^j\right) d\mu = \int \chi_A d\mu \int \chi_B d\mu.$$

Se considerarmos agora $B = \Sigma$, ficaremos com:

$$\lim_{j\to+\infty}\int \chi_A\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu = \int \chi_A d\mu \int \chi_\Sigma d\mu.$$

$$\Rightarrow \lim_{j\to+\infty}\int \chi_A\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu = \int \chi_A d\mu\cdot\mu\left(\Sigma\right).$$

$$\Rightarrow \lim_{j\to+\infty}\int \chi_A\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu = \int \chi_A d\mu\cdot 1.$$

$$\Rightarrow \lim_{j\to+\infty}\int \chi_A\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu = \int \chi_A d\mu.$$

Isto vale para cada mensurável $A \in \mathbb{B}$. Sejam $A, B \in \mathbb{B}$ dois mensuráveis quaisquer e sejam $\alpha, \beta \in \mathbb{R}$ constantes quaisquer. Então valem (aplicando o que acabamos de obter):

$$\begin{cases} \lim\limits_{j\to+\infty}\int \chi_A\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu & = \int \chi_A d\mu, \\ \lim\limits_{j\to+\infty}\int \chi_B\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu & = \int \chi_B d\mu. \end{cases}$$

Logo, multiplicando a primeira identidade por α e a segunda por β, temos:

$$\begin{cases} \alpha\cdot\lim\limits_{j\to+\infty}\int \chi_A\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu & = \alpha\cdot\int \chi_A d\mu, \\ \beta\cdot\lim\limits_{j\to+\infty}\int \chi_B\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu & = \beta\cdot\int \chi_B d\mu. \end{cases}$$

Pela linearidade do limite e da integral de Lebesgue ficamos com:

$$\begin{cases} \lim\limits_{j\to+\infty}\int \alpha\chi_A\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu & = \int \alpha\chi_A d\mu, \\ \lim\limits_{j\to+\infty}\int \beta\chi_B\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu & = \int \beta\chi_B d\mu. \end{cases}$$

Somamos estas equações e, novamente pela linearidade do limite e da integral, obtemos:

$$\lim_{j\to+\infty}\int \alpha\chi_A\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu + \lim_{j\to+\infty}\int \beta\chi_B\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu = \int \alpha\chi_A d\mu + \int \beta\chi_B d\mu.$$

$$\Rightarrow \lim_{j\to+\infty}\int \left(\alpha\chi_A\cdot\left(\chi_\Sigma\circ\sigma^j\right) + \beta\chi_B\cdot\left(\chi_\Sigma\circ\sigma^j\right)\right)d\mu = \int \left(\alpha\chi_A + \beta\chi_B\right)d\mu.$$

$$\Rightarrow \lim_{j\to+\infty}\int \left[\left(\alpha\chi_A + \beta\chi_B\right)\cdot\left(\chi_\Sigma\circ\sigma^j\right)\right]d\mu = \int \left(\alpha\chi_A + \beta\chi_B\right)d\mu.$$

Prosseguimos da mesma forma e, por indução, mostramos que:

$$\lim_{j\to+\infty}\int\left[\left(\sum_{i=1}^{k}\alpha_i\chi_{A_i}\right)\cdot\left(\chi_\Sigma\circ\sigma^j\right)\right]d\mu=\int\left(\sum_{i=1}^{k}\alpha_i\chi_{A_i}\right)d\mu.$$

Ou seja,

$$\lim_{j\to+\infty}\int\varphi\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu=\int\varphi d\mu,$$

para toda função simples φ (consulte a definição 4.1 de [1], página 27).

Seja $\psi:\Sigma\to\mathbb{R}$ uma função não-negativa e mensurável com respeito aos espaços mensuráveis $(\Sigma,\mathbb{B})$ e $(\mathbb{R},\mathbb{S})$, onde $\mathbb{S}$ é a σ-álgebra de Borel de $\mathbb{R}$. Então (via lema 2.11 de [1], página 13) existe uma sequência monótona não-decrescente $(\varphi_n)_{n\in\mathbb{N}}$ de funções simples não-negativas tal que

$$\psi=\lim_{n\to+\infty}\varphi_n.$$

Para cada uma destas funções simples vale:

$$\lim_{j\to+\infty}\int\varphi_n\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu=\int\varphi_n d\mu.$$

$$\Rightarrow\lim_{n\to+\infty}\left(\lim_{j\to+\infty}\int\varphi_n\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu\right)=\lim_{n\to+\infty}\int\varphi_n d\mu.$$

Logo, pelo teorema da convergência monótona (teorema 4.6 de [1], página 31), temos que

$$\begin{aligned}\lim_{n\to+\infty}\left(\lim_{j\to+\infty}\int\varphi_n\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu\right)&=\lim_{n\to+\infty}\int\varphi_n d\mu\\&=\int\lim_{n\to+\infty}\varphi_n d\mu\\&=\int\psi d\mu.\end{aligned}$$

Ou seja, temos que

$$\int\psi d\mu=\lim_{n\to+\infty}\left(\lim_{j\to+\infty}\int\varphi_n\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu\right).\tag{5.4}$$

Por outro lado, como $\psi=\lim_{n\to+\infty}\varphi_n$, vemos que

$$\lim_{j\to+\infty}\int\psi\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu=\lim_{j\to+\infty}\int\lim_{n\to+\infty}\varphi_n\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu.$$

Agora, como $\chi_\Sigma \circ \sigma^j : \Sigma \to \mathbb{R}$ é definida por

$$\left(\chi_\Sigma \circ \sigma^j\right)(\mathbf{x}) = \chi_\Sigma\left(\sigma^j(\mathbf{x})\right) = 1, \forall \mathbf{x} \in \Sigma,$$

pois $\sigma^j : \Sigma \to \Sigma$ (portanto, $\sigma^j(x) \in \Sigma, \forall x \in \Sigma$), então $\chi_\Sigma \circ \sigma^j \equiv 1$ e, assim,

$$\begin{aligned}
\lim_{j\to+\infty} \int \psi \cdot \left(\chi_\Sigma \circ \sigma^j\right) d\mu &= \lim_{j\to+\infty} \int \lim_{n\to+\infty} \varphi_n \cdot \left(\chi_\Sigma \circ \sigma^j\right) d\mu \\
&= \lim_{j\to+\infty} \int \lim_{n\to+\infty} \varphi_n \cdot 1 d\mu \\
&= \lim_{j\to+\infty} \int \lim_{n\to+\infty} \varphi_n d\mu \\
&= \int \lim_{n\to+\infty} \varphi_n d\mu.
\end{aligned}$$

Agora, aplicamos o teorema da convergência monótona (teorema 4.6 de [1], página 31) e obtemos:

$$\lim_{j\to+\infty} \int \psi \cdot \left(\chi_\Sigma \circ \sigma^j\right) d\mu = \lim_{n\to+\infty} \int \varphi_n d\mu. \tag{5.5}$$

Assim, como queremos que

$$\lim_{j\to+\infty} \int \psi \cdot \left(\chi_\Sigma \circ \sigma^j\right) d\mu = \int \psi d\mu,$$

então, por meio de (5.4) e (5.5), precisamos mostrar que

$$\lim_{n\to+\infty} \left(\lim_{j\to+\infty} \int \varphi_n \cdot \left(\chi_\Sigma \circ \sigma^j\right) d\mu \right) = \lim_{n\to+\infty} \int \varphi_n d\mu.$$

Lembramos que $\chi_\Sigma \circ \sigma^j \equiv 1$. Com isto em mãos, obtemos

$$\begin{aligned}
\lim_{n\to+\infty} \left(\lim_{j\to+\infty} \int \varphi_n \cdot \left(\chi_\Sigma \circ \sigma^j\right) d\mu \right) &= \lim_{n\to+\infty} \left(\lim_{j\to+\infty} \int \varphi_n \cdot 1 d\mu \right) \\
&= \lim_{n\to+\infty} \left(\lim_{j\to+\infty} \int \varphi_n d\mu \right).
\end{aligned}$$

Percebamos que $\int \varphi_n d\mu$ não depende em nada da variável j e, por tal motivo, temos que

$$\lim_{j\to+\infty} \int \varphi_n d\mu = \int \varphi_n d\mu.$$

Daí, segue que

$$\lim_{n\to+\infty}\left(\lim_{j\to+\infty}\int \varphi_n\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu\right)=\lim_{n\to+\infty}\left(\lim_{j\to+\infty}\int \varphi_n d\mu\right)$$
$$=\lim_{n\to+\infty}\int \varphi_n d\mu.$$

Mostramos o que precisávamos. Consequentemente,

$$\lim_{j\to+\infty}\int \psi\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu=\int \psi d\mu,$$

para cada função $\psi:\Sigma\to\mathbb{R}$ não-negativa e mensurável com respeito aos espaços mensuráveis $(\Sigma,\mathbb{B})$ e $(\mathbb{R},\mathbb{S})$.

Agora seja $\psi:\Sigma\to\mathbb{R}$ uma função integrável (consulte a definição 5.1 de [1], página 41) qualquer e sejam ψ^+ e ψ^- a sua parte positiva e a sua parte negativa, respectivamente. Então temos, por definição de integral, que:

$$\int \psi d\mu=\int \psi^+ d\mu-\int \psi^- d\mu.$$

Como $\psi^+:\Sigma\to\mathbb{R}$ e $\psi^-:\Sigma\to\mathbb{R}$ são funções não-negativas e mensuráveis com respeito aos espaços mensuráveis $(\Sigma,\mathbb{B})$ e $(\mathbb{R},\mathbb{S})$, então valem:

$$\begin{cases}\displaystyle\lim_{j\to+\infty}\int \psi^+\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu=\int \psi^+ d\mu,\\[2em]\displaystyle\lim_{j\to+\infty}\int \psi^-\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu=\int \psi^- d\mu.\end{cases}$$

Logo, multiplicando a segunda equação por -1, obtemos

$$\begin{cases}\displaystyle\lim_{j\to+\infty}\int \psi^+\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu=\int \psi^+ d\mu,\\[2em]\displaystyle-\lim_{j\to+\infty}\int \psi^-\cdot\left(\chi_\Sigma\circ\sigma^j\right)d\mu=-\int \psi^- d\mu.\end{cases}$$

Graças à linearidade da integral de Lebesgue e do limite, temos:

$$\begin{cases} \lim\limits_{j\to+\infty} \displaystyle\int \psi^{+} \cdot \left(\chi_{\Sigma} \circ \sigma^{j}\right) d\mu = \int \psi^{+} d\mu, \\[2em] \lim\limits_{j\to+\infty} \displaystyle\int -\psi^{-} \cdot \left(\chi_{\Sigma} \circ \sigma^{j}\right) d\mu = \int -\psi^{-} d\mu. \end{cases}$$

Somamos estas equações e obtemos assim:

$$\lim_{j\to+\infty} \int \psi^{+} \cdot \left(\chi_{\Sigma} \circ \sigma^{j}\right) d\mu + \lim_{j\to+\infty} \int -\psi^{-} \cdot \left(\chi_{\Sigma} \circ \sigma^{j}\right) d\mu = \int \psi^{+} d\mu + \int -\psi^{-} d\mu.$$

$$\Rightarrow \lim_{j\to+\infty} \int \left(\psi^{+} - \psi^{-}\right) \cdot \left(\chi_{\Sigma} \circ \sigma^{j}\right) d\mu = \int \left(\psi^{+} - \psi^{-}\right) d\mu.$$

$$\Rightarrow \lim_{j\to+\infty} \int \psi \cdot \left(\chi_{\Sigma} \circ \sigma^{j}\right) d\mu = \int \psi d\mu.$$

Portanto,

$$\lim_{j\to+\infty} \int \psi \cdot \left(\chi_{\Sigma} \circ \sigma^{j}\right) d\mu = \int \psi d\mu,$$

para cada função $\psi : \Sigma \to \mathbb{R}$ integrável.

Agora, aplicamos a relação de dualidade do operador P_{σ} e ficamos com:

$$\lim_{j\to+\infty} \int \chi_{\Sigma} \cdot P_{\sigma}^{j}(\psi)\, d\mu = \int \psi d\mu.$$

$$\Rightarrow \lim_{j\to+\infty} \int 1 \cdot P_{\sigma}^{j}(\psi)\, d\mu = \int \psi d\mu.$$

$$\Rightarrow \lim_{j\to+\infty} \int P_{\sigma}^{j}(\psi)\, d\mu = \int \psi d\mu.$$

Como isto vale para qualquer função integrável ψ, podemos fazer o mesmo com um autovetor

φ_i de P_σ associado ao autovalor λ_i. Obtemos assim:

$$\lim_{j\to+\infty}\int P_\sigma^j(\varphi_i)\,d\mu = \int \varphi_i d\mu.$$

$$\Rightarrow \lim_{j\to+\infty}\int \lambda_i^j \varphi_i d\mu = \int \varphi_i d\mu.$$

$$\Rightarrow \left(\lim_{j\to+\infty}\lambda_i^j\right)\cdot\left(\int \varphi_i d\mu\right) = \int \varphi_i d\mu.$$

$$\Rightarrow \lim_{j\to+\infty}\lambda_i^j = 1.$$

Como $\lambda_i \in \mathbb{R}$ e $|\lambda_i| = 1, \forall i \in \{1,2,...,k\}$, então $\lambda_i = 1$ ou $\lambda_i = -1$. Se $\lambda_i = -1$, então

$$\lambda_i^j = (-1)^j,$$

que não converge quando $j \to +\infty$. Portanto, como

$$\lim_{j\to+\infty}\lambda_i^j = 1,$$

então não podemos ter $\lambda_i = -1$. Ou seja, $\lambda_i = 1, \forall i \in \{1,2,...,k\}$.

Q. E. D.

O teorema de Ionescu-Tulcea e Marinescu juntamente com este resultado nos diz que

$$\begin{aligned} P_\sigma^n &= \sum_{i=1}^{k} 1^n \cdot P_i + N^n \\ &= \sum_{i=1}^{k} 1 \cdot P_i + N^n \\ &= \sum_{i=1}^{k} P_i + N^n, \end{aligned}$$

onde N possui raio espectral $\rho(N) < 1$. Observe que como cada P_i é uma projeção e $P_iP_j = P_jP_i = 0$, sempre que $i \neq j$, então vale que

$$(P_1 + \cdots + P_k)^2 = P_1 + \cdots + P_k.$$

Ou seja, $P := P_1 + \cdots + P_k$ é uma projeção e, portanto,

$$P_\sigma^n = P + N^n. \tag{5.6}$$

Teorema 19. Para cada $\varphi \in F_\theta$ vale que $P(\varphi) = \int \varphi d\mu$.

Demonstração. Do teorema de Ionescu-Tulcea e Marinescu cada P_i é a projeção sobre o autoespaço associado ao autovalor λ_i, mas como todos os autovalores são iguais a 1, então cada P_i é uma projeção sobre o autoespaço associado ao autovalor 1. Portanto, P é também uma projeção sobre o autoespaço $E_1 = \{\varphi \in F_\theta \mid P_\sigma(\varphi) = \varphi\}$. Pode-se mostrar que $dim E_1 = 1$, consulte a proposição 6.5 de [10], páginas 51 - 53. Portanto, $E_1 = span\{I\}$, onde $I : \Sigma \to \mathbb{R}$ denota a função constante igual a 1 (para verificar que I é um elemento de E_1 consulte a observação imediatamente acima da proposição 6.5 de [10], página 51).

Então para cada $\varphi \in F_\theta$ existe um $\lambda_\varphi \in \mathbb{R}$ tal que

$$P(\varphi) = \lambda_\varphi \cdot I.$$

Logo,

$$\begin{aligned}
\int P(\varphi) d\mu &= \int \lambda_\varphi \cdot I d\mu \\
&= \lambda_\varphi \int I d\mu \\
&= \lambda_\varphi \int \chi_\Sigma d\mu \\
&= \lambda_\varphi \cdot \mu(\Sigma) \\
&= \lambda_\varphi \cdot 1 \\
&= \lambda_\varphi.
\end{aligned}$$

Portanto, para encontrar uma expressão para $P(\varphi)$ temos que conhecer λ_φ. Para isto, observamos o seguinte

$$|N^n(\varphi)| \le |N^n(\varphi)|_\infty.$$

Portanto,

$$\int |N^n(\varphi)| d\mu \le \int |N^n(\varphi)|_\infty d\mu.$$

Perceba que $N: F_\theta \to F_\theta$, portanto, $N^n(\varphi) \in F_\theta$, ou seja é uma função. Já $|N^n(\varphi)|_\infty$ é um número real constante, então

$$\begin{aligned}
\int |N^n(\varphi)| d\mu &\leq \int |N^n(\varphi)|_\infty d\mu \\
&= |N^n(\varphi)|_\infty \cdot \int 1 d\mu \\
&= |N^n(\varphi)|_\infty \cdot 1 \\
&= |N^n(\varphi)|_\infty \\
&\leq \|N^n(\varphi)\|_\theta \\
&\leq \|N^n\| \cdot \|\varphi\|_\theta \,.
\end{aligned}$$

Agora, lembramos que N possui raio espectral $\rho(N) < 1$ e, portanto, existe $D > 0$ tal que

$$\|N^n\| \leq D \cdot \rho(N)^n.$$

Desta forma, ficamos com

$$\int |N^n(\varphi)| d\mu \leq D \cdot \rho(N)^n \cdot \|\varphi\|_\theta \,.$$

Agora, lembramos que

$$\left| \int N^n(\varphi) d\mu \right| \leq \int |N^n(\varphi)| d\mu \leq D \cdot \rho(N)^n \cdot \|\varphi\|_\theta \,.$$

O que nos diz, finalmente, que

$$0 \leq \left| \int N^n(\varphi) d\mu \right| \leq D \cdot \rho(N)^n \cdot \|\varphi\|_\theta \,.$$

Agora, como $0 < \rho(N) < 1$, então $\lim\limits_{n\to\infty} \rho(N)^n = 0$. Portanto, $\lim\limits_{n\to\infty} D \cdot \rho(N)^n \cdot \|\varphi\|_\theta = 0$. Logo, o teorema do confronto nos garante que

$$\lim_{n\to\infty} \left| \int N^n(\varphi) d\mu \right| = 0.$$

Consequentemente,

$$\lim_{n\to\infty} \int N^n(\varphi) d\mu = 0.$$

Integramos a equação 5.6 e obtemos

$$\begin{aligned}
\int P_\sigma^n(\varphi) d\mu &= \int (P(\varphi) + N^n(\varphi)) d\mu \\
&= \int P(\varphi) d\mu + \int N^n(\varphi) d\mu.
\end{aligned}$$

Fazemos $n \to \infty$, e, por um lado, obtemos

$$\begin{aligned}
\lim_{n\to\infty}\int P_\sigma^n(\varphi)d\mu &= \lim_{n\to\infty}\left(\int P(\varphi)d\mu + \int N^n(\varphi)d\mu\right)\\
&= \lim_{n\to\infty}\int P(\varphi)d\mu + \lim_{n\to\infty}\int N^n(\varphi)d\mu\\
&= \lim_{n\to\infty}\int P(\varphi)d\mu + 0\\
&= \lim_{n\to\infty}\int P(\varphi)d\mu\\
&= \int P(\varphi)d\mu.
\end{aligned}$$

Por outro lado, a relação de dualidade nos diz que

$$\begin{aligned}
\lim_{n\to\infty}\int P_\sigma^n(\varphi)d\mu &= \lim_{n\to\infty}\int P_\sigma^n(\varphi)\cdot 1 d\mu\\
&= \lim_{n\to\infty}\int \varphi\cdot(1\circ\sigma^n)d\mu\\
&= \lim_{n\to\infty}\int \varphi\cdot 1 d\mu\\
&= \lim_{n\to\infty}\int \varphi d\mu\\
&= \int \varphi d\mu.
\end{aligned}$$

Com estas duas igualdades obtemos que

$$\int P(\varphi)d\mu = \int \varphi d\mu.$$

Lembramos que

$$\lambda_\varphi = \int P(\varphi)d\mu.$$

Portanto,

$$\lambda_\varphi = \int \varphi d\mu.$$

Agora, lembramos que

$$P(\varphi) = \lambda_\varphi \cdot I$$

e, obtemos

$$P(\varphi) = \int \varphi d\mu \cdot I = \int \varphi d\mu \cdot 1.$$

Concluindo assim que

$$P(\varphi) = \int \varphi d\mu.$$

Q.E.D.

Uma consequência interessante destes fatos é o decaimento exponencial de correlações que enunciamos a seguir.

Teorema 20. Seja (σ,μ) o shift de Bernoulli. Então existe um $t\in(0,1)$ tal que para cada $\psi\in F_\theta$ e para cada $\varphi\in L^\infty(\Sigma,\mathbb{B},\mu)$ existe uma constante $C>0$ que satisfaz:

$$\left|\int(\varphi\circ\sigma^n)\psi d\mu-\int\varphi d\mu\int\psi d\mu\right|\leq C\cdot t^n,\forall n\in\mathbb{N}.$$

Demonstração. Vemos que $\int\psi d\mu$ é um número real constante e, então

$$\left|\int(\varphi\circ\sigma^n)\psi d\mu-\int\varphi d\mu\int\psi d\mu\right|=\left|\int(\varphi\circ\sigma^n)\psi d\mu-\int\left[\varphi\cdot\left(\int\psi d\mu\right)\right]d\mu\right|.$$

Utilizamos a relação de dualidade e obtemos:

$$\begin{aligned}\left|\int(\varphi\circ\sigma^n)\psi d\mu-\int\varphi d\mu\int\psi d\mu\right|&=\left|\int\varphi\cdot P_\sigma^n(\psi)d\mu-\int\left[\varphi\cdot\left(\int\psi d\mu\right)\right]d\mu\right|\\&=\left|\int\left[\varphi\cdot P_\sigma^n(\psi)-\varphi\cdot\left(\int\psi d\mu\right)\right]d\mu\right|\\&=\left|\int\left[\varphi\cdot\left(P_\sigma^n(\psi)-\int\psi d\mu\right)\right]d\mu\right|\\&\leq\int\left|\varphi\cdot\left(P_\sigma^n(\psi)-\int\psi d\mu\right)\right|d\mu\\&=\int|\varphi|\cdot\left|\left(P_\sigma^n(\psi)-\int\psi d\mu\right)\right|d\mu.\end{aligned}$$

Como $|\varphi|\leq|\varphi|_\infty$ e $|\varphi|_\infty$ é um número real constante, então

$$\begin{aligned}\left|\int(\varphi\circ\sigma^n)\psi d\mu-\int\varphi d\mu\int\psi d\mu\right|&\leq\int|\varphi|_\infty\cdot\left|\left(P_\sigma^n(\psi)-\int\psi d\mu\right)\right|d\mu\\&=|\varphi|_\infty\cdot\int\left|\left(P_\sigma^n(\psi)-\int\psi d\mu\right)\right|d\mu.\end{aligned}$$

Agora, lembramos que $\int\psi d\mu=P(\psi)$ e como

$$P_\sigma^n(\psi)=P(\psi)+N^n(\psi),$$

então

$$P_\sigma^n(\psi)-P(\psi)=N^n(\psi).$$

Portanto,

$$P_\sigma^n(\psi) - \int \psi d\mu = N^n(\psi).$$

Logo,

$$\begin{aligned}\left|\int (\varphi\circ\sigma^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu\right| &\leq |\varphi|_\infty \cdot \int \left|\left(P_\sigma^n(\psi) - \int \psi d\mu\right)\right| d\mu \\ &= |\varphi|_\infty \cdot \int |N^n(\psi)|\, d\mu.\end{aligned}$$

Como $|N^n(\psi)| \leq |N^n(\psi)|_\infty \leq \|N^n(\psi)\|_\theta$ e como $\|N^n(\psi)\|_\theta$ é um número real constante, então

$$\begin{aligned}\left|\int (\varphi\circ\sigma^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu\right| &\leq |\varphi|_\infty \cdot \int \|N^n(\psi)\|_\theta\, d\mu \\ &= |\varphi|_\infty \cdot \|N^n(\psi)\|_\theta \cdot \int 1 d\mu \\ &= |\varphi|_\infty \cdot \|N^n(\psi)\|_\theta \cdot 1 \\ &= |\varphi|_\infty \cdot \|N^n(\psi)\|_\theta \\ &\leq |\varphi|_\infty \cdot \|N^n\| \cdot \|\psi\|_\theta\,.\end{aligned}$$

Como N tem raio espectral $0 < \rho(N) < 1$, então existe $D > 0$ tal que $\|N^n\| \leq D\cdot\rho(N)^n$. Logo,

$$\left|\int (\varphi\circ\sigma^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu\right| \leq |\varphi|_\infty \cdot D \cdot \rho(N)^n \cdot \|\psi\|_\theta\,.$$

Seja $C = |\varphi|_\infty \cdot D \cdot \|\psi\|_\theta$ e seja $t = \rho(N)$. Assim, obtemos que para cada $n \in \mathbb{N}$, para cada $\psi \in F_\theta$ e para cada $\varphi \in L^\infty(\Sigma, \mathbb{B}, \mu)$ vale

$$\left|\int (\varphi\circ\sigma^n)\psi d\mu - \int \varphi d\mu \int \psi d\mu\right| \leq C \cdot t^n.$$

Q.E.D.

Este resultado é interessante, pois ele nos diz a velocidade da convergência (que a propriedade mixing garante) $\int (\varphi\circ\sigma^n)\psi d\mu \to \int \varphi d\mu \int \psi d\mu$. No caso, a convergência tem velocidade exponencial.

6 CONCLUSÃO

Neste trabalho construímos um espaço de probabilidade, o espaço simbólico, que é apropriado para a construção do sistema dinâmico de tempo discreto (σ, μ), o shift de Bernoulli. Vimos alguns conceitos importantes para a teoria ergódica como o de medida invariante e provamos que o shift de Bernoulli é um sistema cuja função σ é mensurável e cuja probabilidade μ é σ-invariante.

Com este ambiente propício para estudar propriedades ergódicas vimos que o shift de Bernoulli é um sistema misturador, ou seja, para cada $A, B \in \mathbb{B}$ vale a propriedade

$$\lim_{n\to\infty} \mu\left(A \cap \sigma^{-n}(B)\right) = \mu(A)\mu(B).$$

Partindo deste conceito, demos uma demonstração de que o shift de Bernoulli é um sistema ergódico. Ou seja, não pode ser decomposto em subsistemas que não sejam os triviais.

Se o conjunto estudado for um espaço métrico compacto, como por exemplo o que estudamos no capítulo 5, pode-se mostrar que um sistema que satisfaz uma propriedade mais fraca que a propriedade mixing é um sistema ergódico através de um conceito equivalente ao de ergodicidade (consulte o lema 6.11(i) de [14], página 154, com a mudança apropriada de notação) que é o seguinte

$$\frac{1}{n}\sum_{i=0}^{n-1} \int f(\sigma^i(x))g(x)d\mu(x) \to \int f(x)d\mu(x) \int g(x)d\mu(x), \forall f \in C(\Sigma), \forall g \in L^1(\Sigma, \mathbb{B}, \mu).$$

Enquanto que sistemas misturadores satisfazem (consulte o lema 6.11(ii) de [14], página 154):

$$\int f(\sigma^n(x))g(x)d\mu(x) \to \int f(x)d\mu(x) \int g(x)d\mu(x), \forall f \in C(\Sigma), \forall g \in L^1(\Sigma, \mathbb{B}, \mu).$$

Ou seja, a ergodicidade nada mais é do que a mesma convergência do mixing, porém com uma convergência mais fraca, que é a convergência em média ou soma de Cesàro. É sabido da Análise Real que a convergência de uma sequência implica a sua convergência em média. Portanto, sob este ponto de vista, um sistema misturador é, devido a este resultado, um sistema ergódico.

Concluímos o nosso trabalho apresentando a seguinte definição que nos permite dizer que dois sistemas são "idênticos"sob o ponto de vista da teoria ergódica.

Definição 20. Sejam μ e ν probabilidades invariantes por transformações mensuráveis $f : M \to M$ e $g : N \to N$, respectivamente. Dizemos que os sistemas (f, μ) e (g, ν) são **ergodicamente**

equivalentes, se podemos escolher conjuntos mensuráveis $X \subseteq M$ e $Y \subseteq N$ com $\mu(X) = 1$ e $\nu(Y) = 1$, e uma função mensurável $\varphi : X \to Y$ bijetora com inversa mensurável, de tal forma que

$$\varphi * \mu = \nu \qquad \text{e} \qquad \varphi \circ f = g \circ \varphi.$$

Onde $\varphi * \mu$ é definida para cada mensurável $A \subseteq Y$ por

$$\varphi * \mu(A) = \mu(\varphi^{-1}(A)).$$

É com esta definição que podemos "codificar"certos sistemas através do shift de Bernoulli. Ou seja, certos sistemas dinâmicos são ergodicamente equivalentes ao shift de Bernoulli e, portanto, podemos obter propriedades sobre o sistema estudando-as no shift de Bernoulli. Vejamos um exemplo a seguir.

Consideremos $f : [0,1] \to [0,1]$ definida por $f(x) = 10x - [10x]$, onde $[10x]$ denota a parte inteira do número real $10x$. É sabido que esta transformação preserva a medida de Lebesgue λ em $[0,1]$ (consulte a seção 1.3.1 de [7]). Escrevendo um número $x \in [0,1]$ em sua expansão decimal

$$x = 0, a_1 a_2 a_3 \cdots,$$

a transformação f corresponde simplesmente a deslocar os dígitos de x uma unidade para a esquerda. Isso nos motiva a considerar:

$$\varphi : \{0,1,...,9\}^{\mathbb{N}} \to [0,1], \qquad \varphi\left((a_n)_{n\in\mathbb{N}}\right) = \sum_{n=1}^{+\infty} \frac{a_n}{10^n} = 0, a_1 a_2 a_3 \cdots.$$

É claro que φ é sobrejetiva. Por outro lado, ela não é injetiva, uma vez que certos números reais possuem mais que uma expansão decimal: por exemplo, $0,1000000\cdots = 0,099999\cdots$. De fato, isso acontece somente se o número admite uma expansão decimal finita, ou seja, tal que todos os dígitos a partir de certa ordem são nulos. Esses números formam um conjunto enumerável (por ser um subconjunto de $\mathbb{Q}$) e, portanto, são irrelevantes do ponto de vista da medida de Lebesgue (pois o conjunto tem medida nula). Mais precisamente, consideremos o conjunto $X \subseteq \{0,1,...,9\}^{\mathbb{N}}$ das sequências com um número infinito de símbolos diferentes de zero e o conjunto $Y \subseteq [0,1]$ dos números cuja expansão decimal é infinita (logo, única). Então a restrição de φ a X é uma bijeção sobre Y.

É fácil verificar que tanto φ quanto a sua inversa são mensuráveis: use o fato de que a imagem da interseção de X com cada cilindro $[0; a_1, ..., a_m]$ é a interseção de Y com um intervalo de comprimento $10 - m$. Esta observação também mostra que $\varphi * \lambda = \mu$, onde μ representa a medida de Bernoulli em $\{0, 1, ..., 9\}^{\mathbb{N}}$ que dá igual peso a todos os dígitos. Além disso, se denotarmos por σ o deslocamento à esquerda em $\{0, 1, ..., 9\}^{\mathbb{N}}$, temos que

$$\varphi \circ \sigma\left((a_n)_{n\in\mathbb{N}}\right) = 0, a_2 a_3 a_4 \cdots = f \circ \varphi\left((a_n)_{n\in\mathbb{N}}\right)$$

para todo $(a_n)_{n\in\mathbb{N}} \in X$. Isto prova que (f, λ) é ergodicamente equivalente ao deslocamento de Bernoulli (σ, μ).

ÍNDICE

REFERÊNCIAS

[1] BARTLE, R. G., *The Elements of Integration and Lebesgue Measure*. New York: Wiley Classics Library, 1995.

[2] LUCENA, R. N. O, *The Transfer and Perron-Frobenius Operators*. 12 de Outubro de 2015. Disponível em: https://sites.google.com/im.ufal.br/rafaellucena. Acesso em: 18 de novembro de 2021.

[3] LUCENA, R. N. O, *Spectral Gap and statistical properties for piecewise expanding maps*. Maceió, 2014.

[4] BALADI, V., *Positive Transfer Operators and Decay of Correlations*. 1a Ed. Singapore: World Scientific, 2000. v.16.

[5] TULCEA C. T. I; MARINESCU G, *Theorie ergodique pour des classes d'operations non completement continues*. Annals of Mathematics, v. 52, n. 1, p. 140-147, Washington, 1950.

[6] HALMOS, P., *Measure Theory*. D. Van Nostrand Company, 1950.

[7] VIANA, M.; OLIVEIRA, K., *Fundamentos da Teoria Ergódica*. Rio de Janeiro, SBM, 2014.

[8] VITALI, G., *Sul problema della misura dei gruppi di punti di una retta*. Bologna, Tip. Gamberini e Parmeggiani, 1905.

[9] LIMA, E. L., *Curso de análise vol. 1*. 1a Ed. Rio de Janeiro, IMPA, 2014.

[10] LIMA, S. I. P. S. *Gap Espectral para o Operador de Perron-Frobenius associado ao Shift de Bernoulli*. Maceió, 2021.

[11] LIMA, E. L. *Análise Real vol. 1. Funcões de uma Variável*. 8 Ed. Rio de Janeiro, IMPA, 2006.

[12] KREYSZIG, E. *Introductory functional analysis with applications*. 1 Ed. John Wiley & Sons. Inc., 1978.

[13] JACOD, J.; PROTTER, P. *Probability Essentials*. 2a Ed. Springer, Berlin, Heidelberg, 2004.

[14] WALTERS, P. *An Introduction to Ergodic Theory*. 1a Ed. Springer, Berlin, Heidelberg, 2000.