Universidade Federal do Rio Grande do Sul

Instituto de Matemática e Estatística

Programa de Pós-Graduação em Matemática

S-Convolução e o Operador de Transferência Generalizado

Tese de Doutorado

Lucas Spillere Barchinski

Porto Alegre, 15 de dezembro de 2016

Tese submetida por Lucas Spillere Barchinski como requisito parcial para a obtenção do grau de Doutor em Ciência Matemática, pelo Programa de Pós Graduação em Matemática, do Instituto de Matemática e Estatística da Universidade Federal do Rio Grande do Sul.

Professor Orientador:
Prof. Dr. Alexandre Tavares Baraviera

Banca examinadora:
Prof. Dr. Alexandre Tavares Baraviera(UFRGS)
Prof. Dr. Diego Marcon Farias (UFRGS)
Prof. Dr. José Afonso Barrionuevo (PPGMAp - UFRGS)
Prof. Dr. Rodrigo Bissacot Proença (IME - USP)

Agradecimentos

Gostaria de agradecer a todos os meus amigos. Vocês foram essenciais em minha caminhada acadêmica. Várias risadas, discussões filosóficas e matemáticas, que me mantiveram firme em meu objetivo. Também agradeço ao professor Alexandre, por ter me aceitado como orientando, e pelas inúmeras vezes me auxiliou neste trabalho. Agradeço também aos membros da banca, pelas importantes colocações realizadas.

Agradeço ao professor Fagner pela parceria no futebol, pelos conselhos pessoais e também pela colaboração na parte do operador generalizado.

Agradeço ao meu colega Bruno, que foi meu grande parceiro nesta tese. Muitos dias elaborando e testando conjecturas, regadas a muito café com paçoca.

Finalmente, agradeço a minha família. Amo vocês. Em especial a minha esposa Aline, que cuidou muito bem de tudo enquanto eu estudava.

Resumo

Nesta tese apresentamos uma variação do conceito de convolução de medidas. Trata-se da S-convolução, uma operação derivada da convolução usual, porém não-associativa e não-comutativa. Exploramos suas principais propriedades e suas relações com caracteres do grupo $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Utilizando tais relações, diagonalizamos algumas matrizes Bloco-Hankel. Na segunda parte da tese, definimos o operador de transferência generalizado, inspirados na definição de subshift generalizado desenvolvida, por exemplo, nos trabalhos de Gromov em [5] e de Friedland em [3]. Nesse contexto, provamos o Teorema de Ruelle-Perron-Frobenius.

Abstract

In this thesis we present a variation of concept of the convolution measure. This is a S-convolution, a derived operation of the usual convolution, but noncommutative and nonassociative. We have explored its main properties and its relationship with characters of the $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ group. Using such relations, we have diagonalized some Bloco-Hankel matrices. In the second part of this thesis, we have defined a generalized transfer operator, inspired by the definition of the generalized subshift developed, for example, in the works of Gromov in [5] and Friedland in [3]. In this context, we have proved the Ruelle-Perron-Frobenius Theorem.

Conteúdo

Introdução

No ano de 1967, Furstenberg, em [4], discutiu várias propriedades da dinâmica $(\mathbb{T}, \sigma_p)$, onde $\mathbb{T}$ é o toro unidimensional e $\sigma_p(x) = px \mod 1$. Neste trabalho, Furstenberg conjecturou que

Conjectura. *(Furstenberg) As únicas medidas invariantes para o semi-grupo dos endomorfismos do círculo gerado por σ_p e σ_q, para p e q multiplicativamente indenpendentes, são a medida de Lebesgue λ, e medidas atômicas suportadas em órbitas periódicas.*

Boa parte das pesquisas envolvendo a dinâmica $(\mathbb{T}, \sigma_p)$ tem relações com esta conjectura. De fato, ela já foi provada se considerarmos medidas com entropia positiva. Primeiramente, para p e q relativamente primos, Rudolph conseguiu o resultado em [15]. O caso em que p e q não são relativamente primos, mas são multiplicativamente independentes, ou seja, quando ambos não são potências do mesmo inteiro, foi demonstrado por Johnson em [9]. Porém, para casos onde a medida tem entropia nula, a conjectura ainda não foi provada.

Motivados pela conjectura de Furstenberg, Lindenstrauss et al. em [10], demonstraram que

Teorema. *Seja $\{\mu_i\}$ uma sequência de medidas ergódicas e p-invariantes em $\mathbb{T}$ cujas entropias normalizadas $h_i = h_{\mu_i}(\sigma_p)/\log p$ satisfazem*

$$\sum_{i=1}^{\infty} \frac{h_i}{|\log h_i|} = \infty.$$

Então

$$h_{\mu_1 * \mu_2 * \cdots * \mu_n}(\sigma) \to \log p, \ \textit{monotonicamente}, \ \textit{quando } n \to \infty.$$

Em particular, $\mu_1 * \mu_2 * \cdots * \mu_n \to \lambda$ *na topologia fraca*.*

Inspirado pelo resultado de Lindenstrauss, mas trocando o grupo $\mathbb{T}$ pelo grupo $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, Bruno B. Uggioni em sua tese de doutorado [18], provou resultados de convergência de convoluções de medidas para a medida de Haar do grupo, a saber, a medida produto de medidas de Bernoulli uniforme em cada uma das coordenadas.

Neste trabalho, continuamos estudando o grupo $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, porém trocamos a operação de convolução pela S-convolução, a saber, se μ e η são probabilidades σ-invariantes em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$

$$\mu *_S \eta(A) = \mu \times \eta(S^{-1}(A)),$$

onde $S : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \times (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ é dada por $S(x, y) = x - y$.

Essa operação entre medidas não é comutativa e nem mesmo associativa, sendo algebricamente mais complexa que a convolução. Por outro lado, se duas medidas de probabilidade são invariantes pelo shift, então sua S-convolução continua invariante pelo shift.

Neste novo cenário, nosso resultado principal da primeira parte do trabalho, exposto nos capítulos 2 e 3, é o seguinte teorema:

Teorema. *Seja* $(\eta_n)_{n\in\mathbb{N}}$ *uma sequência de probabilidades* σ*-invariantes em* $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$*. Se para todo* $m \in \mathbb{N}^*$ *e para todo caracter não trivial* χ *constante em cilindros de tamanho* m*, tivermos*

$$\prod_{n=1}^{\infty} \left| \int \chi \, d\eta_n \right| \to 0,$$

então

$$\eta_1 *_S \eta_2 *_S \cdots *_S \eta_n \to \left(\frac{1}{p}, \frac{1}{p}, \cdots, \frac{1}{p}\right)$$

na topologia fraca.*

A segunda parte deste trabalho, contida no capítulo 4, segue uma direção diferente da primeira parte do texto. Nela tratamos do operador de transferência.

O operador de transferência tem um papel importante em mecânica estatística e formalismo termodinâmico. Quando o sistema dinâmico é dado pelo shift unilateral e pelo espaço simbólico do tipo finito, Ruelle provou em [16] e [17] que o operador de transferência agindo no espaço das funções Hölder contínuas tem um único autovalor positivo maximal com uma autofunção positiva. Isto nos dá a existência e unicidade da medida de Gibbs para este sistema dinâmico.

Em 2014, Lopes et al em [11], generalizaram alguns resultados do formalismo termodinâmico clássico, considerando um espaço métrico compacto M como espaço de estados. Neste artigo, estudaram potenciais que podem depender de infinitas coordenadas em $M^{\mathbb{N}}$.

De outro lado, Gromov em [5], com o objetivo de calcular a entropia topológica de uma aplicação holomórfica, introduziu o conceito de entropia de um gráfico. Em 1995, Friedland em [3], explorou tal conceito. Neste trabalho, X é um espaço métrico compacto e $\Gamma \subset X \times X$ um conjunto fechado. Ainda,

$$\Gamma^{\infty} = \{(x_i)_1^{\infty} : (x_i, x_{i+1}) \in \Gamma,\ i = 1, 2, \cdots, \},$$

e $\sigma : \Gamma^{\infty} \to \Gamma^{\infty}$ é o shift. Assim, definiu $h(\Gamma)$ como sendo a entropia topológica de $\sigma|\Gamma^{\infty}$.

Analisando a estrutura do conjunto acima, percebemos que $\sigma^{-1}(x)$ varia de acordo com x. Desta forma, se A é um espaço métrico compacto, pensamos num operador de transferência em que a medida de probabilidade em $\sigma^{-1}(x)$ usada no operador varia de acordo com $x \in X = A^{\mathbb{N}}$. De maneira mais precisa, definimos o operador de transferência generalizado por $L_g : C(X) \to C(X)$, por

$$L_g\varphi(x) = \int_A e^{g(ax)}\varphi(ax)d\nu_x(a),$$

onde g é Lipschitz em X. O resultado principal desta segunda parte, contida no capítulo 4, é a demonstração de uma versão do Teorema de Ruelle-Perron-Frobenius para este operador.

Esta tese está organizada da seguinte forma:

No capítulo 1, desenvolvemos os conceitos preliminares necessários para a compreensão deste trabalho. Basicamente tratamos da definição de grupos abelianos finitos, seus caracteres e os principais resultados que serão úteis nessa tese. Além disso, mostramos algumas propriedades da topologia fraca* no espaço das medidas de probabilidade em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$.

O capítulo 2 trata da R-convolução, uma generalização do conceito de convolução de medidas, bem como de algumas de suas propriedades.

No terceiro capítulo exploramos a S-convolução. Analisando seu comportamento em cilindros, chegamos a uma definição de matriz de S-convolução que simplifica a análise da mesma. Percebemos que no caso em que tratamos de cilindros de tamanho 1, a matriz originada é uma matriz particular de Hankel. Para cilindros maiores, obtemos matrizes simétricas Bloco-Hankel. Tais matrizes motivam trabalhos na área de matemática aplicada, como em [1], [2] e [14]. Os principais resultados aqui estão relacionados aos autovalores e autovetores destas matrizes. Estes, por sua vez, intimamente ligados aos caracteres do grupo.

Finalmente, no quarto e último capítulo, abordamos uma questão distinta. Inspirados pela definição de subshift generalizado, definimos um operador de transferência generalizado e provamos o resultado análogo ao Teorema de Ruelle-Perron-Frobenius nesse contexto.

Capítulo 1

Preliminares

Neste primeiro capítulo apresentamos os conceitos preliminares para a compreensão do texto. Procuramos colocá-los de forma objetiva e somente demonstramos os resultados que podem servir de inspiração para o restante do trabalho.

1.1 Grupos

Para uma referência básica sobre teoria de grupos, o leitor pode consultar [6].

Definição 1.1.1. Um grupo $(G, *)$ é um conjunto G dotado de uma operação binária

$$\begin{aligned} * : G \times G &\to G \\ (a, b) &\to a * b \end{aligned}$$

satisfazendo as seguintes condições:

G1. (Associatividade) $\forall\ a, b, c \in G$,

$$(a * b) * c = a * (b * c);$$

G2. (Existência do elemento neutro) Existe um elemento $\mathbb{1}_G \in G$ tal que

$$a * \mathbb{1}_G = a = \mathbb{1}_G * a$$

para todo $a \in G$;

G3. (Existência do inverso) Para cada $a \in G$ existe $a^{-1} \in G$ tal que

$$a * a^{-1} = \mathbb{1}_G = a^{-1} * a.$$

Definição 1.1.2. Um grupo $(G, *)$ é abeliano (ou comutativo) se

$$a * b = b * a, \ \forall \ a, b \in G.$$

A ordem $|G|$ de um grupo finito G é a sua cardinalidade, ou seja, $|G| = \#G$. Se G é um grupo finito, definimos a ordem de um elemento $a \in G$ como sendo o menor inteiro positivo n tal que $a^n = \mathbb{1}_G$.

Proposição 1.1.3. *Seja S um subconjunto não-vazio de um grupo $(G, *)$. Se*

S1. $a, b \in S \Rightarrow a * b \in S$, e

S2. $a \in S \Rightarrow a^{-1} \in S$,

então $(S, *)$ é um grupo.

Um subconjunto não-vazio satisfazendo as condições $(S1)$ e $(S2)$ é chamado de **subgrupo** de G. Quando S é finito, $(S1)$ implica $(S2)$. De fato, seja $a \in S$. Então $\{a, a^2, ...\} \subset S$ e como S é finito a tem ordem finita, digamos $a^n = \mathbb{1}_G$. Logo, $a^{-1} = a^{n-1} \in S$.

Note que a intersecção de subgrupos de G também é um subgrupo de G. Também, dado qualquer subconjunto $X \subset G$ existe um menor subgrupo contendo X. Ele é formado por todos os produtos finitos de elementos de X e seus inversos (repetições permitidas). Denotaremos tal conjunto por $\langle X \rangle$ e o chamaremos de **subgrupo gerado por X**. Se X tiver ordem finita, então o conjunto formado por todos os produtos finitos de elementos de X já é um grupo e é igual a $\langle X \rangle$.

Dizemos que X **gera** G se $G = \langle X \rangle$.

Teorema 1.1.4. ***(Teorema Fundamental de Grupos Abelianos Finitos)****. Se G é um grupo abeliano finito com mais de um elemento, então existem inteiros positivos únicos s e $n_1, \ldots, n_s$, com cada $n_j \geq 2$, tal que $n_j | n_{j+1}$, para $j = 1, \ldots, s-1$ e*

$$G \cong \mathbb{Z}/n_1 \times \cdots \times \mathbb{Z}/n_s.$$

Exemplo 1.1.5. Grupos Cíclicos. Um grupo é **cíclico** se ele é gerado por um único elemento $r \in G$, ou seja, $G = \langle r \rangle$. Se G é finito, então r tem ordem finita, digamos n, e

$$G = \{\mathbb{1}_G, r, r^2, \cdots, r^{n-1}\}.$$

Definição 1.1.6. Sejam $(G, *)$ e $(H, \circ)$ grupos. Um homomorfismo de grupos é uma aplicação $h : G \to H$ tal que

$$h(x * y) = h(x) \circ h(y), \ \forall \ x, y \in G.$$

Note que se h é um homomorfismo de grupos, então $h(\mathbb{1}_G) = \mathbb{1}_H$ pois

$$h(x) = h(x * \mathbb{1}_G) = h(x) \circ h(\mathbb{1}_G).$$

Sempre que não tivermos ambiguidades, escreveremos ab ao invés de $a * b$.

1.2 O Grupo dos Caracteres

Nesta seção trataremos de um dos principais objetos desta tese: o grupo dos caracteres. Nosso objetivo aqui é exibir sua definição, apresentar as relações entre um grupo abeliano finito G e o conjunto dos caracteres definidos nesse grupo e calcular explicitamente quem são os caracteres de $(\mathbb{Z}/p\mathbb{Z})^m$, para qualquer $m \in \mathbb{N}$. Para mais detalhes sobre o grupo dos caracteres, sugerimos o livro [12].

Definição 1.2.1. Um **caracter** de um grupo abeliano finito G é um homomorfismo $\chi : G \to S^1$, ou seja, $\chi(ab) = \chi(a)\chi(b)$ e $\chi(\mathbb{1}_G) = 1$.

Quando G tem ordem p e $g \in G$, para qualquer caracter χ de G nós temos $\chi(g)^p = \chi(g^p) = \chi(1) = 1$. Desta forma, $\chi(g)$ é uma das p-ésimas raízes da unidade.

Definição 1.2.2. O **Grupo dual**, ou **grupo dos caracteres**, de um grupo abeliano finito G, é o grupo $(\widehat{G}, \cdot)$ dos homomorfismos $G \to S^1$ sendo $(\cdot)$ a operação de multiplicação pontual de funções.

O caracter trivial $(\chi \equiv 1)$ é a unidade de $\widehat{G}$. Note que $\widehat{G}$ também é abeliano, pois a operação do grupo é a multiplicação pontual de funções.

Definição 1.2.3. Seja χ um caracter de G. O caracter conjugado de χ é a função $\overline{\chi} : G \to S^1$ dada por $\overline{\chi}(g) := \overline{\chi(g)}$.

Note que para um número complexo z, com $|z| = 1$, temos que $\overline{z} = z^{-1}$. Logo $\overline{\chi}(g) = \chi(g)^{-1} = \chi(g^{-1})$.

Teorema 1.2.4. *Se G_1 e G_2 são grupos e $G_1 \cong G_2$, então $\widehat{G_1} \cong \widehat{G_2}$.*

Demonstração. Ver [12], pagina 34. $\square$

Definição 1.2.5. Suponha que G_1 e G_2 são grupos e que χ_1 e χ_2 são caracteres de G_1 e G_2, respectivamente. O *Produto Tensorial* de χ_1 e χ_2 é a função $\chi_1 \otimes \chi_2 : G_1 \times G_2 \to S^1$ definida por

$$\chi_1 \otimes \chi_2(g_1, g_2) = \chi_1(g_1)\chi_2(g_2)$$

Note que $\chi_1 \otimes \chi_2$ é um caracter de $G_1 \times G_2$. Na verdade temos algo mais forte:

Teorema 1.2.6. *Suponha que G_1 e G_2 são grupos. Então χ é um caracter de $G_1 \times G_2$ se e somente se $\chi = \chi_1 \otimes \chi_2$ para algum $\chi_1 \in \widehat{G_1}$ e $\chi_2 \in \widehat{G_2}$.*

Demonstração. Ver [12], página 35. $\square$

Corolário 1.2.7. *Se G_1 e G_2 são grupos, então*

$$\widehat{G_1 \times G_2} = \widehat{G_1} \otimes \widehat{G_2}.$$

Teorema 1.2.8. *Seja G um grupo cíclico finito de ordem p com um gerador escolhido g. Então $\widehat{G}$ também é um grupo cíclico de ordem p, gerado por χ_1, onde $\chi_1(g^k) = e^{\frac{2\pi ik}{p}}$, $1 \leq k \leq p$.*

Demonstração. Ver [8], página 149. $\square$

Teorema 1.2.9. *Se p é um inteiro positivo, então $\mathbb{Z}/p\mathbb{Z} \cong \widehat{\mathbb{Z}/p\mathbb{Z}}$.*

Demonstração. Ver [12], página 37. $\square$

Teorema 1.2.10. *Se G é um grupo abeliano finito, então G é isomorfo a $\widehat{G}$.*

Demonstração. Pelo Teorema 1.1.4, existem únicos inteiros positivos s e $n_1, \ldots, n_s$, com cada $n_j \geq 2$, tal que $n_j | n_{j+1}$, para $j = 1, \ldots, s-1$ e

$$G \cong \mathbb{Z}/n_1 \times \cdots \times \mathbb{Z}/n_s.$$

Utilizando o Teorema 1.2.4, temos que

$$\widehat{G} \cong \widehat{\mathbb{Z}/n_1 \times \cdots \times \mathbb{Z}/n_s}.$$

Agora, aplicando o Corolário 1.2.7, obtemos

$$\widehat{G} \cong \widehat{\mathbb{Z}/n_1} \otimes \cdots \otimes \widehat{\mathbb{Z}/n_s}.$$

Finalmente, usando o Teorema 1.2.9,

$$\widehat{G} \cong \mathbb{Z}/n_1 \times \cdots \times \mathbb{Z}/n_s \cong G.$$

$\square$

Exemplo 1.2.11. Usando a operação + para a operação binária de $\mathbb{Z}/p\mathbb{Z}$, um caracter de $\mathbb{Z}/p\mathbb{Z}$ é uma função $\chi : \mathbb{Z}/p\mathbb{Z} \to U_p$, onde $U_p = \{e^{\frac{2\pi ik}{p}} : 0 \leq k < p\}$, definida por

$$\chi(a+b) = \chi(a)\chi(b)$$

para todo $a, b \in \mathbb{Z}/p\mathbb{Z}$. Defina $\chi_a : \mathbb{Z}/p\mathbb{Z} \to U_p$ por

$$\chi_a(b) = e^{\frac{2\pi iab}{p}}.$$

Então, para $b_1, b_2 \in \mathbb{Z}/p\mathbb{Z}$,temos

$$\begin{aligned} \chi_a(b_1 + b_2) &= e^{\frac{2\pi ia(b_1+b_2)}{p}} \\ &= e^{\frac{2\pi iab_1}{p}} e^{\frac{2\pi iab_2}{p}} = \chi_a(b_1)\chi_a(b_2). \end{aligned}$$

Portanto, χ_a é um caracter de $\mathbb{Z}/p\mathbb{Z}$. Suponha agora que $\chi_a = \chi_b$. Então

$$\chi_a(1) = \chi_b(1) \Leftrightarrow e^{\frac{2\pi i a}{p}} = e^{\frac{2\pi i b}{p}} \Leftrightarrow a = b \ mod(p).$$

Logo, os caracteres de $\mathbb{Z}/p\mathbb{Z}$ são $\chi_0, \chi_1, \ldots, \chi_{p-1}$.

Exemplo 1.2.12. Seja $m, p \in \mathbb{N}^*$, $p \geq 2$, e χ um caracter do grupo $(\mathbb{Z}/p\mathbb{Z})^m = \underbrace{\mathbb{Z}/p\mathbb{Z} \times \cdots \times \mathbb{Z}/p\mathbb{Z}}_{m \ vezes}$. Pelo Teorema 1.2.10 e pelo Teorema 1.2.6, existe um ponto $a = (a_1, \ldots, a_m) \in (\mathbb{Z}/p\mathbb{Z})^m$ tal que

$$\chi_a(b) = \chi_{a_1}(b_1) \ldots \chi_{a_m}(b_m) = \prod_{k=1}^{m} e^{\frac{2\pi i a_k b_k}{p}}$$

para qualquer $b \in (\mathbb{Z}/p\mathbb{Z})^m$.

Definição 1.2.13. Dizemos que um caracter é real se $Im(\chi) \subset \{1, -1\}$.

Note que se χ é real, então $\chi \equiv \overline{\chi}$. Sabemos que existem p^m caracteres em $(\mathbb{Z}/p\mathbb{Z})^m$. Mas quantos deles são reais? Quantos são complexos? Na próxima proposição responderemos esta pergunta.

Proposição 1.2.14. *O número de caracteres reais em $(\mathbb{Z}/p\mathbb{Z})^m$ é 2^m se p é par, ou 1 se p é ímpar.*

Demonstração. Seja χ um caracter real em $(\mathbb{Z}/p\mathbb{Z})^m$. Então pelo Exemplo 1.2.12 temos que existe um ponto $a = (a_1, \ldots, a_m) \in (\mathbb{Z}/p\mathbb{Z})^m$ tal que

$$\chi(b) = \prod_{k=1}^{m} e^{\frac{2\pi i a_k b_k}{p}}$$

para qualquer $b \in (\mathbb{Z}/p\mathbb{Z})^m$. Sendo χ real, temos que $p \mid 2a_k$, para todo $1 \leq k \leq m$. Como $0 \leq a_k \leq p-1$, temos que se p é ímpar, então $p \mid a_k$ o que implica que $a_k = 0$ para todo k. Logo, $\chi \equiv 1$. Se p é par, vamos provar que χ é um caracter real se e somente se $a_k \in \left\{0, \frac{p}{2}\right\}$, para todo $1 \leq k \leq m$. Se p é par, então $\frac{p}{2} \mid a_k$. Assim $a_k = \frac{p}{2}n$. Como $a_k \leq p-1$, n deve ser

0 ou 1. Logo, $a_k \in \left\{0, \frac{p}{2}\right\}$. Reciprocamente, se $a_k \in \left\{0, \frac{p}{2}\right\}$ então $a + a = 0$. Logo, $\chi^2 \equiv 1$. Assim, $Im(\chi) \subset \{1, -1\}$. $\square$

Teorema 1.2.15. *Seja G um grupo abeliano finito. Então*

$$\sum_{g \in G} \chi(g) = \begin{cases} \#G, & se\ \chi \equiv 1, \\ 0, & se\ \chi \not\equiv 1, \end{cases} \qquad \sum_{\chi \in \widehat{G}} \chi(g) = \begin{cases} \#G, & se\ g = \mathbb{1}_G, \\ 0, & se\ g \neq \mathbb{1}_G. \end{cases}$$

Demonstração. Ver [8], páginas 149 e 150. $\square$

1.3 Topologia Fraca*

A Topologia Fraca* é muito utilizada em várias áreas da Matemática. O leitor que deseja uma definição mais geral sobre o conceito pode consultar os ótimos livros [19] e [20]. Porém, como estamos trabalhando com um espaço topológico específico, a saber $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, apresentaremos alguns resultados específicos para este espaço.

Considere os cilindros

$$[m; a_m, \cdots, a_n] = \{(x_j)_{j \in \mathbb{N}} \in (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} : x_m = a_m, \cdots, x_n = a_n\}.$$

Denotaremos por $\mathcal{A}$ a σ-álgebra gerada por todos os cilindros.

Denotaremos por $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ o espaço das medidas de probabilidade em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$.

Dados uma medida $\eta \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$, um conjunto finito $\Phi = \{\phi_1, \cdots, \phi_N\}$ de funções contínuas $\phi_i : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{R}$ e um número $\varepsilon > 0$, definimos o seguinte conjunto

$$V(\eta, \Phi, \varepsilon) = \left\{\mu \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}) : \left|\int \phi_i \, d\mu \; - \int \phi_i \, d\eta\right| < \varepsilon \text{ para todo } i\right\}.$$

Usando essa definição, a família $\{V(\eta, \Phi, \varepsilon)\} : \Phi, \varepsilon\}$ forma uma base de vizinhanças para cada

$\eta \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. A *Topologia fraca** é a topologia definida por estas bases de vizinhanças.

Proposição 1.3.1. *Sejam $\eta, \mu \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Se $\eta(C) = \mu(C)$ para todo cilindro C, então $\eta = \mu$.*

Demonstração. Se $\eta(C) = \mu(C)$ para todo cilindro $C \subset (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, então $\eta(A) = \mu(A)$ para qualquer $A \in \mathcal{A}$. Como η e μ são σ-aditivas, pelo teorema da extensão, elas possuem uma única extensão. Logo, $\eta = \mu$. $\square$

Proposição 1.3.2. *Sejam $(\mu_k)_{k\in\mathbb{N}}$ e η medidas em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Se $\lim_{k\to\infty} \mu_k(C) = \eta(C)$ para todo cilindro C, então μ_k converge para η na topologia fraca*.*

Demonstração. Teríamos que mostrar que $\int f \; d\mu_k \to \int f \; d\eta$ para toda função contínua f. Porém, basta mostrarmos o mesmo para um conjunto D denso em $C((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Escolheremos $D = \{f \in C((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} : f(x) = f(x_1, ..., x_n);\; n \in \mathbb{N}]\}$. Note que qualquer $f \in D$ pode ser escrita como $f(x) = \sum_{i_1,...,i_n=0}^{p-1} f(x_1, ..., x_n, 0, 0, ...)\mathbb{1}_{[i_1,...,i_n]}(x)$. Assim,

$$\begin{aligned}\int f \; d\mu_k &= \sum_{i_1,\cdots,i_n=0}^{p-1} f(i_1, \cdots, i_n, 0, 0, \cdots)\mu_k([i_1, \cdots, i_n]) \underset{k\to\infty}{\rightarrow} \\ &\rightarrow \sum_{i_1,\cdots,i_n=0}^{p-1} f(i_1, \cdots, i_n, 0, 0, \cdots)\eta([i_1, \cdots, i_n]) = \int f \; d\eta.\end{aligned}$$

Logo, μ_k converge para η na topologia fraca*. $\square$

Capítulo 2

R-Convolução

2.1 R-Convolução

Esse capítulo foi feito conjuntamente com meu colega de doutorado Bruno Brogni Uggioni.

A partir de agora, $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Note que, apesar de infinito, este grupo é um produto cartesiano de grupos finitos. Portanto, os resultados do capítulo anterior nos ajudarão a tratá-lo.

Também, σ denota o shift unilateral sobre esse grupo, ou seja,

$$\sigma : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$$
$$(x_n)_{n\in\mathbb{N}} \to (x_{n+1})_{n\in\mathbb{N}}.$$

De outra forma, o shift envia a sequência $(x_0, x_1, x_2, \cdots)$ na sequência $(x_1, x_2, x_3, \cdots)$.

Considere o conjunto das medidas de probabilidade $\mathcal{M}(G)$. Sejam

A convolução usual de duas medidas de probabilidade $\eta, \mu \in \mathcal{M}(G)$ é dada por

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

onde $\eta \times \mu$ é a medida produto de η e μ, A é subconjunto mensurável de G e $T : G \times G \to G$ é a operação de soma em G, $T(g, h) = g + h$. Observe que $\eta * \mu$ é ainda uma medida de probabilidade de $\mathcal{M}(G)$.

Tal operação tem propriedades interessantes, que veremos adiante, tais como:

- se duas medidas de probabilidade são σ invariante, então $\eta * \mu$ também é σ invariante;
- ela não diminui a entropia das medidas, ou seja,

$$h_{\eta*\mu}(\sigma) \geq \max\{h_\eta(\sigma), h_\mu(\sigma)\};$$

- ela conserva o conjunto das medidas de probabilidade de Bernoulli.

Com isso em mente, percebemos que tal conceito poderia ser generalizado. Ou seja, fornecemos uma definição de R-convolução, da qual a convolução usual é um caso particular, e analisaremos algumas de suas propriedades.

Definição 2.1.1. Sejam η e μ probabilidades σ invariantes e $R : G \times G \to G$ uma transformação mensurável qualquer. Definimos a R-convolução $\eta *_R \mu$ em G por

$$\eta *_R \mu(A) := \eta \times \mu(R^{-1}(A)).$$

Observe que $\eta *_R \mu$ também é uma medida em G. De fato, seja $(A_i)_{i \in \mathbb{N}}$ uma sequência de conjuntos disjuntos e mensuráveis em G. Então, temos:

$$\begin{aligned} \eta *_R \mu(\cup_{i=1}^{\infty} A_i) &= \eta \times \mu(R^{-1}(\cup_{i=1}^{\infty} A_i)) \\ &= \eta \times \mu(\cup_{i=1}^{\infty} R^{-1}(A_i)) \\ &= \sum_{i=1}^{\infty} \eta \times \mu(R^{-1}(A_i)) \end{aligned}$$

$$= \sum_{i=1}^{\infty} \eta *_R \mu(A_i),$$

e todas as outras propriedades que uma medida deve satisfazer seguem de forma similar.

2.2 Propriedades da R-convolução

Proposição 2.2.1. *Sejam η e μ probabilidades σ invariantes e $R : G \times G \to G$ uma transformação que satisfaz*

$$R \circ \sigma \times \sigma = \sigma \circ R,$$

Então a medida $\eta *_R \mu$ é uma probabilidade σ invariante.

Demonstração. Que $\eta *_R \mu$ é uma medida de probabilidade em G isso segue diretamente do fato de η e μ o serem. Para a invariância, seja A um conjunto em G. Então:

$$\begin{aligned} \eta *_R \mu(\sigma^{-1}(A)) &= \eta \times \mu(R^{-1}(\sigma^{-1}(A))) \\ &= \eta \times \mu((\sigma \times \sigma)^{-1}(R^{-1}(A))) \\ &= \eta \times \mu(R^{-1}(A)) \\ &= \eta *_R \mu(A) \end{aligned}$$

e provamos a invariância. □

Lema 2.2.2. *Seja $R : G \times G \to G$ uma transformação mensurável qualquer. Então temos:*

1. $H_{\eta *_R \mu}(\mathcal{P}|\mathcal{Q}) = H_{\eta \times \mu}(R^{-1}(\mathcal{P})|R^{-1}(\mathcal{Q}))$;

2. $R^{-1}(\mathcal{P} \vee \mathcal{Q}) = R^{-1}(\mathcal{P}) \vee R^{-1}(\mathcal{Q})$,

para todas $\mathcal{P}$ e $\mathcal{Q}$ partições finitas de G e η e μ medidas de probabilidade σ-invariantes.

Demonstração. Para o primeiro item, considere os elementos tais como elencados na hipótese. Assim, obtemos:

$$\begin{aligned}
H_{\eta *_R \mu}(\mathcal{P}|\mathcal{Q}) &= \sum_{P,Q} \eta *_R \mu(P\cap Q)\log\left(\frac{\eta *_R \mu(P)}{\eta *_R \mu(P\cap Q)}\right) \\
&= \sum_{P,Q} \eta\times\mu(R^{-1}(P)\cap R^{-1}(Q))\log\left(\frac{\eta\times\mu(R^{-1}(P))}{\eta\times\mu(R^{-1}(P)\cap R^{-1}(Q))}\right) \\
&= H_{\eta\times\mu}(R^{-1}(\mathcal{P})|R^{-1}(\mathcal{Q})),
\end{aligned}$$

e provamos o primeiro item. Observe que nas igualdades acima, P representa um elemento genérico da partição $\mathcal{P}$ e Q um elemento genérico da partição $\mathcal{Q}$. O segundo item é de dificuldade e raciocínio similares, por isso omitimos a demonstração. $\square$

Lema 2.2.3. *Seja $R : G\times G\to G$ uma transformação definida da seguinte forma*

$$R(x,y) = (r(x_1,y_1), r(x_2,y_2),\ldots),$$

onde $r : \mathbb{Z}/p\mathbb{Z}\times\mathbb{Z}/p\mathbb{Z}\to\mathbb{Z}/p\mathbb{Z}$ é uma função (não necessariamente homomorfismo de grupos) tal que para todo $i\in\mathbb{Z}/p\mathbb{Z}$, as funções $r(i,\cdot)$ e $r(\cdot,i)$ são bijeções em $\mathbb{Z}/p\mathbb{Z}$. Se tomarmos $\mathcal{P} = \{[0],\ldots,[p-1]\}$ então valem as seguintes igualdades entre partições:

1. $R^{-1}(\mathcal{P})\vee R^{-1}\left(\vee_{j=1}^{n-1}\sigma^{-j}(\mathcal{P})\right)\vee(\mathcal{P}^n\times G) = \mathcal{P}^n\times\mathcal{P}^n$;
2. $R^{-1}\left(\vee_{j=1}^{n-1}\sigma^{-j}(\mathcal{P})\right)\vee(\mathcal{P}^n\times G) = \mathcal{P}^n\times\left(\vee_{j=1}^{n-1}\sigma^{-j}(\mathcal{P})\right)$

Demonstração. Para a parte 1, primeiro note que $R^{-1}(\mathcal{P})\vee R^{-1}\left(\vee_{j=1}^{n-1}\sigma^{-j}(\mathcal{P})\right) = R^{-1}(\mathcal{P}^n)$ e todo conjunto de $R^{-1}(\mathcal{P}^n)$ é da forma $R^{-1}([k_1,\ldots,k_n])$, para algum cilindro $[k_1,\ldots,k_n]$. Também, todo conjunto de $\mathcal{P}^n\times G$ é da forma $[i_1,\ldots,i_n]\times G$ para algum cilindro $[i_1,\ldots,i_n]$. Desse modo, todo conjunto da partição $R^{-1}(\mathcal{P}^n)\vee\mathcal{P}^n\times G$ tem a seguinte forma:

$$R^{-1}([k_1,\ldots,k_n])\bigcap([i_1,\ldots,i_n]\times G) =$$

$$
\begin{aligned}
&= \bigcup_{j_1,\ldots,j_n=0}^{p-1} \big([j_1,\ldots,j_n] \times [r(j_1\cdot)^{-1}(k_1),\ldots,r(j_n\cdot)^{-1}(k_n)]\big) \bigcap ([i_1,\ldots,i_n] \times G) \\
&= [i_1,\ldots,i_n] \times [r(i_1\cdot)^{-1}(k_1),\ldots,r(i_n\cdot)^{-1}(k_n)] \\
&= [i_1,\ldots,i_n] \times [k'_1,\ldots,k'_n],
\end{aligned}
$$

e acabamos de ver que $(\mathcal{P}^n \times \mathcal{P}^n) = R^{-1}(\mathcal{P}^n) \vee (\mathcal{P}^n \times G)$, findando a parte 1.

A segunda parte segue de forma similar. Note que a partição $R^{-1}\left(\vee_{j=1}^{n-1}\sigma^{-j}(\mathcal{P})\right)$ é aquela formada por conjuntos do tipo $R^{-1}(\sigma^{-1}([k_2,\ldots,k_n]))$. Desta forma, tomando um conjunto $[i_1,\ldots,i_n] \times G$ de $(\mathcal{P}^n \times G)$, obtemos:

$$
\begin{aligned}
& R^{-1}(\sigma^{-1}([k_2,\ldots,k_n])) \cap ([i_1,\ldots,i_n] \times G) = \\
&= \bigcup_{a,j_1,\ldots,j_n=0}^{p-1} \big([j_1,j_2\ldots,j_n] \times [a,r(j_2\cdot)^{-1}(k_2),\ldots,r(j_n\cdot)^{-1}(k_n)]\big) \bigcap ([i_1,\ldots,i_n] \times G) \\
&= [i_1,\ldots,i_n] \times \bigcup_{a=0}^{p-1} [a,r(j_2\cdot)^{-1}(k_2),\ldots,r(j_n\cdot)^{-1}(k_n)] \\
&= [i_1,\ldots,i_n] \times \bigcup_{a=0}^{p-1} [a,k'_2,\ldots,k'_n] \\
&= [i_1,\ldots,i_n] \times \sigma^{-1}([k'_2,\ldots,k'_n]),
\end{aligned}
$$

e acabamos de ver que $R^{-1}\left(\vee_{j=1}^{n-1}\sigma^{-j}(\mathcal{P})\right) \vee (\mathcal{P}^n \times G) = \mathcal{P}^n \times \left(\vee_{j=1}^{n-1}\sigma^{-j}(\mathcal{P})\right)$, provando a segunda parte e terminando a demonstração. $\square$

Agora estamos prontos para provar que a entropia não decresce quando consideramos certas R-convoluções.

Teorema 2.2.4. *Seja $R : G \times G \to G$ uma transformação da forma:*

$$R(x,y) = (r(x_1,y_1), r(x_2,y_2),\ldots),$$

em que $r : \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$ é uma função (não necessariamente homomorfismo de *grupos) tal que para todo $i \in \mathbb{Z}/p\mathbb{Z}$, as funções$r(i,\cdot)$ e $r(\cdot,i)$ são bijeções em $\mathbb{Z}/p\mathbb{Z}$. Então, temos as seguintes desigualdades para quaisquer η e μ medidas σ-invariantes em G :*

$$\sup\{h_\eta(\sigma), h_\mu(\sigma)\} \leq h_{\eta *_R \mu}(\sigma) \leq h_\eta(\sigma) + h_\mu(\sigma).$$

Demonstração. A primeira constatação é que toda função R do estilo da hipótese satisfaz:

$$R \circ (\sigma \times \sigma) = \sigma \circ R,$$

e, portanto, $\eta *_R \mu$ será uma medida de probabilidade invariante sempre que η e μ o forem.

Agora, escolhendo $\mathcal{P} = \{[0], [1], \ldots, [p-1]\}$ a partição em cilindros de tamanho 1, obtemos:

$$\begin{aligned} h_{\eta *_R \mu}(\sigma) &= \lim_n H_{\eta *_R \mu}(\mathcal{P} | \vee_{j=1}^n \sigma^{-j}(\mathcal{P})) \\ &= \lim_n H_{\eta \times \mu}(R^{-1}(\mathcal{P}) | \vee_{j=1}^n (\sigma \circ \sigma)^{-j}(R^{-1}(\mathcal{P}))) \quad \text{(Lema 2.2.2)} \\ &\leq \sup_{\mathcal{Q}} \{\lim_n H_{\eta \times \mu}(\mathcal{Q} | \vee_{j=1}^n (\sigma \circ \sigma)^{-j}(\mathcal{Q}))\} \\ &= h_{\eta \times \mu}(\sigma \times \sigma) \\ &= h_\eta(\sigma) + h_\mu(\sigma). \end{aligned}$$

Nas desigualdades e igualdades acima, as partições $\mathcal{Q}$ são de $G \times G$.

Finalmente, usando os Lemas 2.2.2 e 2.2.3, obtemos:

$$\begin{aligned} & H_{\eta *_R \mu}(\mathcal{P} | \vee_{j=1}^n \sigma^{-j}(\mathcal{P})) \\ = \quad & H_{\eta \times \mu}(R^{-1}(\mathcal{P})) | R^{-1}(\vee_{j=1}^n \sigma^{-j}(\mathcal{P}))) \\ \geq \quad & H_{\eta \times \mu}(R^{-1}(\mathcal{P}) | R^{-1}(\vee_{j=1}^n \sigma^{-j}(\mathcal{P})) \vee \mathcal{P}^n \times G)) \\ = \quad & H_{\eta \times \mu}(R^{-1}(\mathcal{P}) | \mathcal{P}^n \times \left(\vee_{j=1}^{n-1} \sigma^{-j}(\mathcal{P})\right)) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{i_k,j_k} \eta\times\mu([i_1,\dots,i_n]\times[j_1,\dots,j_n])\log\left(\frac{\eta\times\mu\left(\bigcup_{a\in(\mathbb{Z}/p\mathbb{Z})}[i_1,\dots,i_n]\times[a,j_2,\dots,j_n]\right)}{\eta\times\mu([i_1,\dots,i_n]\times[j_1,\dots,j_n])}\right)\\
&= \sum_{j_k}\mu([j_1,\dots,j_n])\log\left(\frac{\mu([j_2,\dots,j_n])}{\mu([j_1,\dots,j_n])}\right)\\
&= H_\mu(\mathcal{P}|\vee_{j=1}^n\sigma^{-1}(\mathcal{P})),
\end{aligned}
$$

e tomando o limite em n tendendo ao infinito, obtemos que $h_{\eta*\mu}(\sigma)\geq h_\mu(\sigma)$. De forma análoga também podemos mostrar que $h_{\eta*\mu}(\sigma)\geq h_\eta(\sigma)$, o que termina a demonstração. □

Proposição 2.2.5. *Sejam R uma transformação nos moldes do teorema acima, η e μ medidas de Bernoulli quaisquer. Então, $\eta *_R \mu$ também será uma medida de Bernoulli.*

Demonstração. Para checarmos a tese, basta vermos o comportamento de $\eta *_R \mu$ em cilindros. Temos que:

$$
\begin{aligned}
\eta *_R \mu[i_1,\cdots,i_m] &= \sum_{j_1,\cdots,j_m=0}^{p-1}\eta[j_1,\cdots,j_m]\mu[r(j_1\cdot)^{-1}(i_1),\cdots,r(j_m\cdot)^{-1}(i_m)]\\
&= \sum_{j_1,\cdots,j_m=0}^{p-1}(\prod_{k=1}^m\eta[j_k])(\prod_{k=1}^m\mu[r(j_k\cdot)^{-1}(i_k)])\\
&= \sum_{j_1,\cdots,j_m=0}^{p-1}\prod_{k=1}^m\eta[j_k]\mu[r(j_k\cdot)^{-1}(i_k)]\\
&= \prod_{k=1}^m\sum_{j_k=0}^{p-1}\eta[j_k]\mu[r(j_k\cdot)^{-1}(i_k)]\\
&= \prod_{k=1}^m\eta *_R \mu[i_k],
\end{aligned}
$$

as igualdades acima mostram-nos que $\eta *_R \mu$ é medida de Bernoulli. □

Capítulo 3

S-Convolução de Probabilidades

Percebendo as interessantes propriedades da R-convolução, aprofundaremos nosso estudo em torno da S-convolução, um tipo especícifo de R-convolução. Na convolução usual, usamos a operação $R(x, y) = x + y$ e definimos $\eta * \mu(A) := \eta \times \mu(R^{-1}(A))$. Para a S-convolução usaremos a operação $S(x, y) = x - y$. Mesmo sendo uma troca simples, isto produz mudanças significativas. Podemos citar as duas principais:

1) Considerando $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, mostraremos que, ao contrário da convolução usual, a S-convolução não é associativa para $p \geq 3$;

2) A S-convolução não é comutativa para $p \geq 3$.

Analisando como esta nova medida age em cilindros, definimos uma matriz de S-convolução. Explorando suas propriedades, encontramos uma ligação entre três áreas da matemática: álgebra, sistemas dinâmicos e matemática aplicada.

Quando analisamos cilindros de tamanho 1, por exemplo, a matriz originada é uma matriz específica de *Hankel*. Utlizando caracteres diagonalizaremos tais matrizes, inclusive para casos em que $m > 1$, onde as matrizes não serão mais Hankel, mas uma matriz com características similares.

3.1 S-Convolução

A S-convolução é obtida através da diferença dos elementos do grupo. De maneira mais precisa, temos as seguintes definições:

Definição 3.1.1. Seja $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Definimos $S : G \times G \to G$ por $S(x, y) = x - y$.

Definição 3.1.2. Sejam η e μ probabilidades em $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Definimos a S-convolução de η e μ por

$$\eta *_S \mu(A) = \eta \times \mu(S^{-1}A).$$

Note que se $x = (x_1, x_2, \cdots)$ e $y = (y_1, y_2, \cdots)$, então

$$\begin{aligned} S \circ \sigma \times \sigma(x, y) &= S((x_2, x_3, \cdots), (y_2, y_3, \cdots)) \\ &= (x_2 - y_2, x_3 - y_3, \cdots) \\ &= \sigma(x - y) = \sigma \circ S(x, y), \end{aligned}$$

ou seja, $S \circ \sigma \times \sigma = \sigma \circ S$. Portanto, pela Proposição 2.2.1, temos que a S-convolução preserva as medidas de probabilidade σ-invariantes.
Temos que $s : \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}$, dada por $s(i, j) = i - j$, é uma função tal que para todo $i \in \mathbb{Z}/p\mathbb{Z}$, as funções $s(i, \cdot)$ e $s(\cdot, i)$ são bijeções em $\mathbb{Z}/p\mathbb{Z}$. Desta forma, pelo Teorema 2.2.4

$$\sup\{h_\eta(\sigma), h_\mu(\sigma)\} \leq h_{\eta *_S \mu}(\sigma) \leq h_\eta(\sigma) + h_\mu(\sigma),$$

ou seja, a S-convolução não diminui a entropia das medidas operadas. Além disso, se η e μ são medidas de Bernoulli, pela Proposição 2.2.5, $\eta *_S \mu$ também é uma medida de Bernoulli. Para deixar a notação mais limpa, no restante deste trabalho, escreveremos somente $*$ ao invés de $*_S$ para denotarmos a S-convolução.

Para entendermos melhor como esta operação trabalha, considere os seguintes exemplos:

Exemplo 3.1.3. Tome $p = 2$. Desta forma,

$$\begin{aligned}\eta * \mu[0] &= \eta \times \mu(S^{-1}[0]) \\ &= \eta \times \mu([0] \times [0] \cup [1] \times [1]) \\ &= \eta[0]\mu[0] + \eta[1]\mu[1],\end{aligned}$$

e

$$\begin{aligned}\eta * \mu[1] &= \eta \times \mu(S^{-1}[1]) \\ &= \eta \times \mu([1] \times [0] \cup [0] \times [1]) \\ &= \eta[1]\mu[0] + \eta[0]\mu[1].\end{aligned}$$

Utilizando a forma matricial,

$$\begin{pmatrix} \eta[0] & \eta[1] \\ \eta[1] & \eta[0] \end{pmatrix} \begin{pmatrix} \mu[0] \\ \mu[1] \end{pmatrix} = \begin{pmatrix} \eta * \mu[0] \\ \eta * \mu[1] \end{pmatrix}$$

Exemplo 3.1.4. Tratemos do caso geral. Seja $p \in \mathbb{N}$. Note que a S-convolução mede cilindros com o seguinte padrão

$$\eta * \mu[i_1, \cdots, i_n] = \sum_{j_1, \cdots, j_n = 0}^{p-1} \eta[i_1 + j_1, \cdots, i_n + j_n]\mu[j_1, \cdots, j_n].$$

O comportamento da S-convolução em cilindros permite-nos entender com mais profundidade suas propriedades. Por exemplo, podemos concluir o que acontece com qualquer medida, não necessariamente σ-invariante, quando operada com a medida Bernoulli uniforme:

Exemplo 3.1.5. Sejam $\eta, \lambda \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$, onde $\lambda = \left(\frac{1}{p}, \cdots, \frac{1}{p}\right)$ é a medida Bernoulli uniforme. Então

$$\eta * \lambda = \lambda \text{ e } \lambda * \eta = \lambda.$$

De fato, tome qualquer cilindro $[i_1, \cdots, i_n]$. Então

$$\begin{aligned}\eta * \lambda[i_1, \cdots, i_n] &= \sum_{j_1,\cdots,j_n=0}^{p-1} \eta[i_1+j_1, \cdots, i_n+j_n]\lambda[j_1, \cdots, j_n] \\ &= \frac{1}{p^n} \sum_{j_1,\cdots,j_n=0}^{p-1} \eta[i_1+j_1, \cdots, i_n+j_n] = \frac{1}{p^n}.\end{aligned}$$

De forma semelhante, temos que $\lambda * \eta[i_1, \cdots, i_n] = \frac{1}{p^n}$. Portanto, pela Proposição 1.3.1, $\eta * \lambda = \lambda * \eta = \lambda$.

Como no Exemplo 3.1.3, podemos utilizar a forma matricial para expressar como a S-convolução mede cilindros de tamanho m, com $m \in \mathbb{N}$ fixo.

Definição 3.1.6. Sejam $\eta \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ e $m \in \mathbb{N}$ fixos. Considere somente cilindros de tamanho m. Desta forma, definimos a matriz quadrada B_η^m, de ordem p^m, por

$$b_{ij} = \eta[(i_{m-1} + j_{m-1}) \bmod p, \cdots, (i_0 + j_0) \bmod p],$$

onde $i - 1 = \sum_{k=0}^{m-1} i_k p^k$ e $j - 1 = \sum_{k=0}^{m-1} j_k p^k$.

Para evitar uma notação carregada, omitiremos a expressão "mod p".

Exemplo 3.1.7. Considere $p = 3$ e $m = 1$. Desta forma, temos

$$B_\eta^1 = \begin{pmatrix} \eta[0] & \eta[1] & \eta[2] \\ \eta[1] & \eta[2] & \eta[0] \\ \eta[2] & \eta[0] & \eta[1] \end{pmatrix}.$$

Perceba que tal matriz é simétrica. Na verdade este é um fato geral. Decorre imediatamente da Definição 3.1.6, que para quaisquer $p, m \in \mathbb{N}$,

$$b_{ij} = b_{ji}.$$

Definição 3.1.8. Seja $\eta \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Definimos o vetor v_η^m, de ordem $p^m \times 1$, pela expressão $(v_\eta)_i = \eta[(i_{m-1}, \cdots, i_0]$, para todo $1 \leq i \leq p^m$, onde $i - 1 = \sum_{k=0}^{m-1} i_k p^k$.

Para $m = 1$, por exemplo, temos

$$v_\eta^1 = \begin{pmatrix} \eta[0] \\ \eta[1] \\ \vdots \\ \eta[p-1] \end{pmatrix}.$$

Proposição 3.1.9. *Sejam $\eta, \mu \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Então*

$$B_\eta^1 v_\mu^1 = v_{\eta*\mu}^1.$$

Demonstração. Primeiramente note que, para $1 \leq j \leq p$, temos que $j_0 = j - 1$. Além disso,

$$\begin{aligned} [B_\eta^1 v_\mu^1]_i &= \sum_{j=1}^{p} b_{ij}^1 [v_\mu^1]_j \\ &= \sum_{j=1}^{p} \eta[i_0 + j_0]\mu[j_0] \\ &= [v_{\eta*\mu}^1]_i, \end{aligned}$$

para todo $1 \leq i \leq p$. □

Pretendemos agora analisar a matriz associada a S-convolução de duas medidas com matrizes relacionadas a S-convolução destas medidas. O resultado pretendido seria a multiplicação de tais matrizes. Porém, como veremos a seguir, essa nova matriz depende da matriz da convolução usual de uma das medidas. Para isso, considere a seguinte matriz associada a convolução usual de uma medida, definida por Bruno Brogni Uggioni em sua tese de doutorado [18]:

Definição 3.1.10. Seja $\eta \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ e considere somente cilindros de tamanho m. Desta forma, definimos a matriz quadrada A_η^m, de ordem p^m por

$$a_{ij} = \eta[(j_{m-1} - i_{m-1}) \bmod p, \cdots, (j_0 - i_0) \bmod p],$$

onde $i - 1 = \sum_{k=0}^{m-1} i_k p^k$ e $j - 1 = \sum_{k=0}^{m-1} j_k p^k$.

Com isso em mãos, podemos agora descrever melhor as matrizes B_η^m.

Proposição 3.1.11. *Sejam η e μ probabilidades em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Então*

$$B_{\eta*\mu}^m = B_\eta^m (A_\mu^m)^T,$$

para todo $m \geq 1$.

Demonstração. Temos que

$$\begin{aligned}
(B_\eta^m (A_\mu^m)^T)_{ij} &= \sum_{k=1}^{p^m} (b_\eta)_{ik} (a_\mu)_{kj}^T \\
&= \sum_{k=1}^{p^m} (b_\eta)_{ik} (a_\mu)_{jk} \\
&= \sum_{k=1}^{p^m} \eta[i_{m-1} + k_{m-1}, \cdots, i_0 + k_0] \mu[k_{m-1} - j_{m-1}, \cdots, k_0 - j_0] \\
&= \eta * \mu[i_{m-1} + j_{m-1}, \cdots, i_0 + j_0] = (B_{\eta*\mu}^m)_{ij}.
\end{aligned}$$

□

Indutivamente, se tivermos uma sequência $\eta_1, \eta_2 \cdots, \eta_n$ de probabilidades de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, então

$$B_{\eta_1 * \eta_2 * \cdots * \eta_n}^m = B_{\eta_1}^m (A_{\eta_2}^m)^T \cdots (A_{\eta_n}^m)^T.$$

Na expressão acima, e no resto deste trabalho, pensaremos na expressão $\eta_1 * \eta_2 * \cdots * \eta_n$ sempre sendo resolvida da esquerda para direta. Por exemplo,

$$\eta_1 * \eta_2 * \eta_3 * \eta_4 = ((\eta_1 * \eta_2) * \eta_3) * \eta_4.$$

Isto é necessário, visto que a S-convolução não é associativa. De fato, temos que:

Proposição 3.1.12. *A operação de S-convolução não é associativa. Se η_1, η_2 e η_3 são probabilidades em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, então $(\eta_1 * \eta_2) * \eta_3 = (\eta_1 * \eta_3) * \eta_2$.*

Demonstração. Considere o cilindro $[i_1, \cdots, i_n]$. Então,

$$\begin{aligned}
&(\eta_1 * \eta_2) * \eta_3[i_1, \cdots, i_n] = \\
&= \sum_{j_1,\cdots,j_n=0}^{p-1} (\eta_1 * \eta_2)[i_1 + j_1, \cdots, i_n + j_n]\eta_3[j_1, \cdots, j_n] \\
&= \sum_{j_1,\cdots,j_n=0}^{p-1} \sum_{k_1,\cdots,k_n=0}^{p-1} \eta_1[i_1 + j_1 + k_1, \cdots, i_n + j_n + k_n]\eta_2[k_1, \cdots, k_n]\eta_3[j_1, \cdots, j_n] \\
&= \sum_{k_1,\cdots,k_n=0}^{p-1} \sum_{j_1,\cdots,j_n=0}^{p-1} \eta_1[i_1 + j_1 + k_1, \cdots, i_n + j_n + k_n]\eta_3[j_1, \cdots, j_n]\eta_2[k_1, \cdots, k_n] \\
&= \sum_{k_1,\cdots,k_n=0}^{p-1} \eta_1 * \eta_3[i_1 + k_1, \cdots, i_n + k_n]\eta_2[k_1, \cdots, k_n] \\
&= (\eta_1 * \eta_3) * \eta_2[i_1, \cdots, i_n].
\end{aligned}$$

A não-associatividade da S-convolução também pode ser constatada no Exemplo 3.1.16. $\square$

Exemplo 3.1.13. Ao contrário da convolução usual, a S-convolução, para $p > 2$, não é comutativa. De fato, sejam η e μ medidas de Bernoulli dadas pelos vetores $v_\eta^1 = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{2}\right)$ e

$v_\mu^1 = \left(\frac{2}{3}, \frac{1}{6}, \frac{1}{6}\right)$. Como $v_{\eta*\mu}^1 = B_\eta^1 v_\mu^1$, temos que

$$v_{\eta*\mu}^1 = \begin{pmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \end{pmatrix} \begin{pmatrix} \frac{2}{3} \\ \frac{1}{6} \\ \frac{1}{6} \end{pmatrix} = \begin{pmatrix} \frac{7}{24} \\ \frac{7}{24} \\ \frac{10}{24} \end{pmatrix}$$

Por outro lado,

$$v_{\mu*\eta}^1 = \begin{pmatrix} \frac{2}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \\ \frac{1}{6} & \frac{2}{3} & \frac{1}{6} \end{pmatrix} \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{7}{24} \\ \frac{10}{24} \\ \frac{7}{24} \end{pmatrix}$$

Logo, $\eta * \mu \neq \mu * \eta$. Porém, note que a entropia destas duas novas medidas são iguais. E, de fato, isso vale em geral, como mostramos no Corolário 3.1.15 da Proposição 3.1.14.

Proposição 3.1.14. *Sejam $\eta, \mu \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Então*

$$\eta * \mu[a_1, \cdots, a_n] = \mu * \eta[-a_1, \cdots, -a_n].$$

Demonstração. Note que

$$\begin{aligned} \eta * \mu[a_1, \cdots, a_n] &= \sum_{i_1, \cdots, i_n = 0}^{p-1} \eta[a_1 + i_1, \cdots, a_n + i_n]\mu[i_1, \cdots, i_n] \\ &= \sum_{k_1, \cdots, k_n = 0}^{p-1} \mu[-a_1 + k_1, \cdots, -a_n + k_n]\eta[k_1, \cdots, k_n] \\ &= \mu * \eta[-a_1, \cdots, -a_n]. \end{aligned}$$

□

Corolário 3.1.15. *Sejam $\eta, \mu \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$, σ-invariantes. Então*

$$h_{\eta*\mu}(\sigma) = h_{\mu*\eta}(\sigma).$$

Demonstração. Seja $\mathcal{P}$ a partição por cilindros de tamanho m.

$$\begin{aligned} H_{\eta*\mu}(\sigma) &= \sum_{C\in\mathcal{P}} -\eta*\mu(C)\log\eta*\mu(C) \\ &= \sum_{C\in\mathcal{P}} -\mu*\eta(-C)\log\mu*\eta(-C) = H_{\mu*\eta}(\sigma). \end{aligned}$$

Tomando limites, chegamos ao resultado. $\square$

Portanto, mesmo não sendo comutativa, a entropia da convolução de duas medidas é igual, independente da ordem de operação. Poderíamos pensar que o mesmo deva acontecer com a associatividade. Porém, isso é falso.

Exemplo 3.1.16. Sejam as medidas de Bernoulli definidas pelos vetores de probabilidade $v^1_{\eta_1} = (0.3\ 0.5\ 0.2)$, $v^1_{\eta_2} = (0.4\ 0.4\ 0.2)$ e $v^1_{\eta_3} = (0.1\ 0.6\ 0.3)$. Calcularemos a entropia de $(\eta_1*\eta_2)*\eta_3$ e $\eta_1*(\eta_2*\eta_3)$. Pela Proposição 2.2.5, a S-convolução de medidas de Bernoulli também é uma medida de Bernoulli. Pela Proposição 3.1.9, temos que

$$\begin{aligned} v^1_{\eta_1*\eta_2} &= B^1_{\eta_1}v^1_{\eta_2} \\ &= \begin{pmatrix} 0.3 & 0.5 & 0.2 \\ 0.5 & 0.2 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{pmatrix} \begin{pmatrix} 0.4 \\ 0.4 \\ 0.2 \end{pmatrix} = \begin{pmatrix} 0.36 \\ 0.34 \\ 0.3 \end{pmatrix}. \end{aligned}$$

Assim,

$$\begin{aligned} v^1_{(\eta_1*\eta_2)*\eta_3} &= B^1_{\eta_1*\eta_2}v^1_{\eta_3} \\ &= \begin{pmatrix} 0.36 & 0.34 & 0.3 \\ 0.34 & 0.3 & 0.36 \\ 0.3 & 0.36 & 0.34 \end{pmatrix} \begin{pmatrix} 0.1 \\ 0.6 \\ 0.3 \end{pmatrix} = \begin{pmatrix} 0.330 \\ 0.322 \\ 0.348 \end{pmatrix} \end{aligned}$$

Analogamente,

$$v^1_{\eta_1*(\eta_2*\eta_3)} = \begin{pmatrix} 0.318 \\ 0.340 \\ 0.342 \end{pmatrix}.$$

Portanto,

$$h_{(\eta_1*\eta_2)*\eta_3}(\sigma) = 1.098082... \neq 1.098076... = h_{\eta_1*(\eta_2*\eta_3)}(\sigma).$$

3.2 Matrizes e Caracteres

Nesta seção mostraremos algumas conexões entre caracteres e matrizes da S-convolução. Integrar caracteres é fundamental para a diagonalização de matrizes relacionadas a S-covolução de uma medida.

Definição 3.2.1. Dado um caracter $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}$ constante em cilindros de tamanho m, podemos associar a ele um vetor $v_\chi \in \mathbb{C}^{p^m}$ com entradas complexas definido por:

$$(v_\chi)_i = \chi[i_0, \cdots, i_{m-1}],$$

em que $i - 1 = \sum_{k=0}^{m-1} i_k p^k$ e $i_k \in \{0, 1, \cdots, p-1\}$, ou seja, os i_k's são os índices encontrados na expansão em base p do número $i - 1$.

Proposição 3.2.2. *Sejam χ e ψ caracteres constantes em cilindros de tamanho m tais que $\chi \not\equiv \psi$. Então*

$$< v_\chi, v_\psi >= 0.$$

Demonstração. Como $\chi \not\equiv \psi$, temos que $\chi\overline{\psi} \not\equiv 1$. Note que mesmo que $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ seja infinito, como estes caracteres são constantes em cilindros de tamanho m, podemos aplicar o Teorema 1.2.15. Desta forma, temos que

$$\sum_{i_0,\cdots,i_{m-1}=0}^{p-1} \chi\overline{\psi}[i_0,\cdots,i_{m-1}] = 0.$$

Portanto,

$$\begin{aligned} < v_\chi, v_\psi > &= \sum_{i_0,\cdots,i_{m-1}=0}^{p-1} \chi[i_0,\cdots,i_{m-1}]\overline{\psi}[i_0,\cdots,i_{m-1}] \\ &= \sum_{i_0,\cdots,i_{m-1}=0}^{p-1} \chi\overline{\psi}[i_0,\cdots,i_{m-1}] = 0. \end{aligned}$$

□

Proposição 3.2.3. *Sejam $\eta \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ uma medida de probabilidade e χ um caracter de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ constante em cilindros de tamanho m. Se B_η^m for a matriz associada a η considerando cilindros de tamanho m e v_χ o vetor associado ao caracter χ, então vale a igualdade vetorial:*

$$B_\eta^m v_\chi = \left(\int \chi d\eta\right) v_{\overline{\chi}}.$$

Demonstração. Note que

$$\begin{aligned} \left(\int \chi \; d\eta \; v_{\overline{\chi}}\right)_i &= \sum_{j=1}^{p^m} \eta[j_{m-1},\cdots,j_1,j_0]\chi([j_{m-1},\cdots,j_1,j_0])\overline{\chi}[i_{m-1},\cdots,i_1,i_0] \\ &= \sum_{j=1}^{p^m} \eta[j_{m-1},\cdots,j_1,j_0]\chi([j_{m-1}-i_{m-1},\cdots,j_1-i_1,j_0-i_0]) \\ &= \sum_{l=1}^{p^m} \eta[i_{m-1}+l_{m-1},\cdots,i_1+l_1,i_0+l_0]\chi([l_{m-1},\cdots,l_1,l_0]) \\ &= \sum_{l=1}^{p^m} \left(B_\eta^m\right)_{il}(v_\chi)_l \\ &= \left(B_\eta^m v_\chi\right)_i \end{aligned}$$

$\square$

Proposição 3.2.4. *Sejam $\eta \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ uma medida de probabilidade e χ um caracter de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ constante em cilindros de tamanho m. Se B_η^m for a matriz associada a η considerando cilindros de tamanho m e v_χ o vetor associado ao caracter χ, então valem as seguintes propriedades:*

1) Se χ é real, então

$$B_\eta^m v_\chi = \left(\int \chi d\eta\right) v_\chi.$$

2) Se χ não é real, ou seja, $\chi \not\equiv \overline{\chi}$, e $\int \chi d\eta = 0$, então

$$B_\eta^m v_\chi = 0\cdot v_\chi \quad e \quad B_\eta^m v_{\overline{\chi}} = 0\cdot v_{\overline{\chi}}.$$

3) Se $\chi \not\equiv \overline{\chi}$ e $\int \chi d\eta \neq 0$, então $\pm\left|\int \chi d\eta\right|$ são autovalores de B_η^m associados, respectivamente, aos autovetores reais $v = (\alpha v_\chi + \overline{\alpha} v_{\overline{\chi}}) = 2Re(\alpha v_\chi)$ e $w = \left(\beta v_\chi + \overline{\beta} v_{\overline{\chi}}\right) = 2Re(\beta v_\chi)$, onde $\alpha = \left|\int \chi d\eta\right| + \overline{\int \chi d\eta}$ e $\beta = -\left|\int \chi d\eta\right| + \overline{\int \chi d\eta}$.

Demonstração. Pela Proposição 3.2.3 temos que

$$\begin{cases} B_\eta^m v_\chi = \left(\int \chi d\eta\right) v_{\overline{\chi}}, \\ B_\eta^m v_{\overline{\chi}} = \left(\int \overline{\chi} d\eta\right) v_\chi. \end{cases}$$

Desta forma, para o caso 1), basta notarmos que $\chi = \overline{\chi}$. Para 2) note que $\int \overline{\chi} d\eta = \overline{\int \chi d\eta}$. Agora, vamos demonstrar 3). Perceba que a matriz B_η^m deixa invariante o subespaço vetorial spam$\{v_\chi, v_{\overline{\chi}}\}$. Como essa matriz é simétrica, deve existir um número real λ tal que

$$B_\eta^m (a v_\chi + b v_{\overline{\chi}}) = \lambda(a v_\chi + b v_{\overline{\chi}}) = (\lambda a) v_\chi + (\lambda b) v_{\overline{\chi}}. \tag{3.1}$$

Note que

$$\begin{aligned} B_\eta^m(av_\chi + bv_{\overline{\chi}}) &= aB_\eta^m v_\chi + bB_\eta^m v_{\overline{\chi}} \\ &= \left(a\int \chi d\eta\right) v_{\overline{\chi}} + \left(b\int \overline{\chi} d\eta\right) v_\chi \end{aligned} \tag{3.2}$$

Pela Proposição 3.2.2 o conjunto $\{v_\chi, v_{\overline{\chi}}\}$ é L.I., (3.1) e (3.2) implicam que

$$\begin{cases} \lambda a = b\displaystyle\int \overline{\chi} d\eta, \\ \lambda b = a\displaystyle\int \chi d\eta. \end{cases} \tag{3.3}$$

Desta forma, assumindo $a \neq 0$ e $b \neq 0$,

$$\begin{aligned} \lambda^2 &= \frac{a}{b}\int \chi d\eta \, \frac{b}{a}\int \overline{\chi} d\eta \\ &= \int \chi d\eta \int \overline{\chi} d\eta \\ &= \int \chi d\eta \overline{\int \chi d\eta} = \left|\int \chi d\eta\right|^2. \end{aligned}$$

Assim, $\lambda = \pm\left|\int \chi d\eta\right|$. Para $\lambda = \left|\int \chi d\eta\right|$, uma das soluções do sistema (3.3) é

$$a = \left(\left|\int \chi d\eta\right| + \overline{\int \chi d\eta}\right) =: \alpha, \text{ e } b = \overline{\alpha} = \left(\left|\int \chi d\eta\right| + \int \chi d\eta\right).$$

Portanto, tomando $v = (\alpha v_\chi + \overline{\alpha} v_{\overline{\chi}})$, temos

$$\begin{aligned} B_\eta^m(v) &= B_\eta^m((\alpha v_\chi + \overline{\alpha} v_{\overline{\chi}})) \\ &= \alpha\left(\int \chi d\eta\right) v_{\overline{\chi}} + \overline{\alpha}\left(\int \overline{\chi} d\eta\right) v_\chi \\ &= \left[\left(\left|\int \chi d\eta\right| + \overline{\int \chi d\eta}\right)\left(\int \chi d\eta\right)\right] v_{\overline{\chi}} + \left[\left(\left|\int \chi d\eta\right| + \int \chi d\eta\right)\left(\overline{\int \chi d\eta}\right)\right] v_\chi \\ &= \left(\int \chi d\eta \left|\int \chi d\eta\right| + \left|\int \chi d\eta\right|^2\right) v_{\overline{\chi}} + \left(\overline{\int \chi d\eta}\left|\int \chi d\eta\right| + \left|\int \chi d\eta\right|^2\right) v_\chi \end{aligned}$$

$$= \left|\int \chi d\eta\right| \left(\int \chi d\eta + \left|\int \chi d\eta\right|\right) v_{\overline{\chi}} + \left|\int \chi d\eta\right| \left(\overline{\int \chi d\eta} + \left|\int \chi d\eta\right|\right) v_{\chi}$$
$$= \lambda \left(\alpha v_{\chi} + \overline{\alpha} v_{\overline{\chi}}\right) = \lambda v.$$

Analogamente, tomando $\lambda = -\left|\int \chi d\eta\right|$, encontramos outra solução para o sistema (3.3):

$$a = \left(-\left|\int \chi d\eta\right| + \overline{\int \chi d\eta}\right) =: \beta, \text{ e } b = \overline{\beta} = \left(-\left|\int \chi d\eta\right| + \int \chi d\eta\right).$$

Então tomando $w = \left(\beta v_{\chi} + \overline{\beta} v_{\overline{\chi}}\right)$ temos que

$$B^m_\eta(w) = \lambda w$$

. □

Sabemos que toda matriz simétrica é diagonalizável. Porém, na demonstração do corolário abaixo, iremos explicitar os autovalores e autovetores que compõem as matrizes que diagonalizam B^m_η.

Corolário 3.2.5. *A matriz B^m_η é diagonalizável.*

Demonstração. Pelo Teorema 1.2.10, temos que $(\mathbb{Z}/p\mathbb{Z})^m \cong (\widehat{\mathbb{Z}/p\mathbb{Z}})^m$. Sendo assim, temos p^m caracteres em $(\mathbb{Z}/p\mathbb{Z})^m$, que podem ser estendidos para p^m caracteres constantes em cilindros de tamanho m em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Temos duas possibilidades para estes caracteres:

- χ é real:

Pelo caso 1), v_χ é o autovetor associado a $\int v_\chi \, d\eta$.

- χ não é real:

Note que

$$\langle v_\chi, v_\chi \rangle = \sum_{i_0,\dots,i_{m-1}=0}^{p-1} \chi[i_0,\dots,i_{m-1}]\overline{\chi}[i_0,\dots,i_{m-1}] = \langle v_{\overline{\chi}}, v_{\overline{\chi}} \rangle.$$

Assim, pela Proposição 3.2.2, temos que

$$\begin{aligned}\langle v,w\rangle &= \langle \alpha v_\chi + \overline{\alpha} v_{\overline{\chi}}, \beta v_\chi + \overline{\beta} v_{\overline{\chi}}\rangle \\ &= \alpha\overline{\beta}\langle v_\chi, v_\chi\rangle + \overline{\alpha}\beta\langle v_{\overline{\chi}}, v_{\overline{\chi}}\rangle \\ &= 2\mathrm{Re}\left(\alpha\overline{\beta}\right)\langle v_\chi, v_\chi\rangle.\end{aligned}$$

Mas

$$\begin{aligned}\alpha\overline{\beta} &= \left(\left|\int \chi d\eta\right| + \overline{\int \chi d\eta}\right)\left(-\left|\int \chi d\eta\right| + \int \chi d\eta\right) \\ &= -\left|\int \chi d\eta\right|^2 + \int \chi d\eta\left|\int \chi d\eta\right| - \overline{\int \chi d\eta}\left|\int \chi d\eta\right| + \left|\int \chi d\eta\right|^2 \\ &= \left|\int \chi d\eta\right|\left(\int \chi d\eta - \overline{\int \chi d\eta}\right).\end{aligned}$$

Portanto, $\mathrm{Re}(\alpha\overline{\beta}) = 0$. Logo $\langle v,w\rangle = 0$.

Desta forma, seja o caso 2) ou 3) da Proposição 3.2.4, cada dupla de caracteres $(\chi, \overline{\chi})$ gerará uma dupla de vetores L.I's, a saber, v_χ e $v_{\overline{\chi}}$ no caso 2), e v e w no caso 3).

Finalmente, usando a Proposição 3.2.2, podemos deduzir que todos estes autovetores, gerados pelo caso real e não real, são ortogonais e, portanto, L.I's. Assim obtemos p^m vetores L.I's e B_η^m é diagonalizável. $\square$

Agora estabeleceremos um critério muito útil para diferenciar medidas em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Esse resultado será essencial para demonstrarmos o Teorema 3.4.3, que afirma que a sequência de probabilidades σ-invariantes $\eta_n * \eta_{n-1} * \cdots * \eta_1$ converge, na topologia fraca*, para a medida Bernoulli uniforme $\left(\frac{1}{p}, \frac{1}{p}, \cdots, \frac{1}{p}\right)$.

Corolário 3.2.6. *Sejam η e μ medidas de probabilidade em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ então $\eta = \mu$ se e somente se para todo $m \in \mathbb{N}$ e todo caracter $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to S^1$, constante em cilindros de* *tamanho m,*

$$\int \chi d\eta = \int \chi d\mu.$$

Demonstração. Se $\eta = \mu$, então dados $m \in \mathbb{N}$ e um caracter χ de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ constante em cilindros de tamanho m, temos que

$$\int \chi d\eta = \int \chi d\mu.$$

Reciprocamente, suponha que para todo $m \in \mathbb{N}$ e todo caracter $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to S^1$ constante em cilindros de tamanho m,

$$\int \chi d\eta = \int \chi d\mu.$$

Então, pelo Teorema 3.2.4 e pelo Corolário 3.2.5, B_η^m e B_μ^m são diagonalizadas pelas mesmas matrizes, ou seja, $B_\eta^m = VDV^{-1} = B_\mu^m$, onde V é a matriz formada pelos autovetores e D a matriz diagonal formada pelos respectivos autovalores reais. Logo, $B_\eta^m = B_\mu^m$. Mas isto nos diz que $\eta[i_1, \cdots, i_m] = \mu[i_1, \cdots, i_m]$ para todo cilindro $[i_1, \cdots, i_m]$. Como m é qualquer, temos que $\eta = \mu$. $\square$

Exemplo 3.2.7. Seja λ a medida de Bernoulli uniforme. Então, dado qualquer caracter não-trivial χ, constante em cilindros de tamanho m, temos que

$$\int_{(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}} \chi \, d\lambda = \frac{1}{p^m}\left(\sum_{i_1,\ldots,i_m=0}^{p-1} \chi([i_1, \ldots, i_m])\right).$$

Como $\chi \not\equiv 1$, o Teorema 1.2.15 garante-nos que a soma do lado direito da igualdade acima é 0. Logo

$$\int_{(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}} \chi \, d\lambda = 0, \ \forall \ \chi \not\equiv 1,$$

e se μ é uma medida qualquer tal que $\int \chi \ d\mu = 0$ para todo caracter χ não-trivial, pela Proposição 3.2.6 podemos concluir que μ é a medida Bernoulli uniforme.

3.3 Diagonalização de algumas matrizes bloco Hankel

Agora daremos uma aplicação da teoria desenvolvida neste capítulo. Note que na demonstração da Proposição 3.2.3, não usamos o fato de η ser uma medida de probabilidade. Apenas usamos a definição da matriz

$$b_{ij} = \eta[(i_{m-1} + j_{m-1}) \bmod p, \cdots, (i_0 + j_0) \bmod p],$$

onde $i - 1 = \sum_{k=0}^{m-1} i_k p^k$ e $j - 1 = \sum_{k=0}^{m-1} j_k p^k$. Sendo assim, generalizaremos estes resultados.

Proposição 3.3.1. *Sejam $m \in \mathbb{N}^*$, p um número natural maior ou igual a 2, e A uma matriz quadrada de ordem p^m. Suponha que exista uma função $f : (\mathbb{Z}/p\mathbb{Z})^m \to \mathbb{C}$ tal que*

$$A_{ij} = f((i_0, i_1, \ldots, i_{m-1}) + (j_0, j_1, \ldots, j_{m-1}))$$

onde $1 \leq i, j \leq p^m$, $i - 1 = \sum_{k=0}^{m-1} i_k p^k$, $j - 1 = \sum_{k=0}^{m-1} j_k p^k$. Então A é diagonalizável.

Demonstração. Basta seguir o roteiro a partir da Proposição 3.2.3, substituindo a expressão $\int \chi d\eta$ por $\sum_{i_0,\ldots,i_{m-1}=0}^{p-1} f(i_0, \ldots, i_{m-1})\chi(i_0, \ldots, i_{m-1})$. □

Vamos analisar alguns exemplos para entender melhor este resultado.

Exemplo 3.3.2. Considere a matriz

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 2 & -1 & 3 \\ -1 & 3 & 2 \end{pmatrix}.$$

Nosso objetivo é diagonalizá-la usando caracteres. Nesse caso, tomaremos $p = 3$ e $m = 1$.

Desta forma, devemos encontrar $f : (\mathbb{Z}/3\mathbb{Z}) \to \mathbb{C}$ tal que $A_{ij} = f(i_0 + j_0)$. Note que

$$3 = A_{11} = f(0+0) = f(0), \quad 2 = A_{12} = f(0+1) = f(1) \quad e \quad -1 = A_{13} = f(0+2) = f(2).$$

Portanto, f fica definida pela primeira linha da matriz A.

Agora perceba que temos 3 caracteres χ_j associados a seus respectivos v_{χ_j}:

$$\chi_0 \to v_{\chi_0} = (1,1,1), \; \chi_1 \to v_{\chi_1} = (1, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}) \text{ e } \chi_2 \to v_{\chi_2} = (1, e^{\frac{4\pi i}{3}}, e^{\frac{2\pi i}{3}}).$$

Vamos determinar agora os autovalores:

$$\sum_{j=0}^{2} f(j_0)\chi_0(j_0) = 3\cdot 1 + 2\cdot 1 - 1\cdot 1 = 4$$
$$\sum_{j=0}^{2} f(j_0)\chi_1(j_0) = 3\cdot 1 + 2\cdot e^{\frac{2\pi i}{3}} - 1\cdot e^{\frac{4\pi i}{3}} = \frac{5}{2} + \frac{3\sqrt{3}}{2}i$$
$$\sum_{j=0}^{2} f(j_0)\chi_2(j_0) = 3\cdot 1 + 2\cdot e^{\frac{4\pi i}{3}} - 1\cdot e^{\frac{2\pi i}{3}} = \frac{5}{2} - \frac{3\sqrt{3}}{2}i$$

Como χ_0 é real o primeiro autovalor é 4. Para χ_1 e χ_2 os autovalores são $\pm\left|\frac{5}{2} + \frac{3\sqrt{3}}{2}i\right| = \pm\sqrt{13}$.

Lembrando que $v = 2Re(\alpha v_{\chi_1})$ e $w = 2Re(\beta v_{\chi_1})$, determinamos os autovetores:

$$v_{\chi_0} = (1,1,1)$$
$$v = 2Re\left(\sqrt{13} + \frac{5 - 3\sqrt{3}i}{2}\right)\left(1, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}\right) = \left(5 + 2\sqrt{13}, 2 - \sqrt{13}, -7 - \sqrt{13}\right)$$
$$w = 2Re\left(-\sqrt{13} + \frac{5 - 3\sqrt{3}i}{2}\right)\left(1, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}\right) = \left(5 - 2\sqrt{13}, \sqrt{13} + 2, \sqrt{13} - 7\right)$$

Logo, se

$$V = \begin{pmatrix} 1 & 5+2\sqrt{13} & 5-2\sqrt{13} \\ 1 & 2-\sqrt{13} & 2+\sqrt{13} \\ 1 & -7-\sqrt{13} & \sqrt{13}-7 \end{pmatrix} \text{ e } D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & \sqrt{13} & 0 \\ 0 & 0 & -\sqrt{13} \end{pmatrix}$$

então $A = VDV^{-1}$.

Definição 3.3.3. Dizemos que uma matriz $H = (h_{i,j})_{n\times n}$ é uma matriz de Hankel se existe uma sequência $(s_k)_{k\in\mathbb{N}}$ tal que $h_{i,j} = s_{i+j-1}$, com $i, j = 1, 2, 3, \ldots$. Se s_k são matrizes quadradas, então dizemos que H é uma matriz bloco Hankel.

De outra forma, uma matriz é Hankel se as entradas nas diagonais paralelas à diagonal secundária são iguais. Elas aparecem naturalmente na literatura sobre processamento de sinais. O leitor interessado pode consultar, por exemplo, [1] e [7].

Note que quando $m = 1$ a matriz B^1_η é uma matriz de Hankel. No caso em que $m \geq 2$, teremos uma matriz bloco Hankel simétrica.

Em alguns casos, devemos analisar a matriz para podermos deduzir quem é o p e quem é o m. Note, por exemplo, que uma matriz 4×4 pode ser um caso de $p = 4$ e $m = 1$, ou de $p = 2$ e $m = 2$. No próximo exemplo analisaremos umas matriz 9×9, a fim de esclarecer melhor este parágrafo.

Exemplo 3.3.4. Seja

$$A = \left(\begin{array}{ccc|ccc|ccc} 1 & 0 & -1 & 2 & -3 & 0 & 1 & 2 & -1 \\ 0 & -1 & 1 & -3 & 0 & 2 & 2 & -1 & 1 \\ -1 & 1 & 0 & 0 & 2 & -3 & -1 & 1 & 2 \\ \hline 2 & -3 & 0 & 1 & 2 & -1 & 1 & 0 & -1 \\ -3 & 0 & 2 & 2 & -1 & 1 & 0 & -1 & 1 \\ 0 & 2 & -3 & -1 & 1 & 2 & -1 & 1 & 0 \\ \hline 1 & 2 & -1 & 1 & 0 & -1 & 2 & -3 & 0 \\ 2 & -1 & 1 & 0 & -1 & 1 & -3 & 0 & 2 \\ -1 & 1 & 2 & -1 & 1 & 0 & 0 & 2 & -3 \end{array}\right).$$

Note que a matriz acima não é Hankel, mas bloco Hankel. Portanto, $m \neq 1$. Assim devemos ter $p = 3$ e $m = 2$. Vamos agora encontrar f:

$$\begin{aligned} 1 &= A_{1,1} = f((0,0) + (0,0)) = f(0,0) \\ 0 &= A_{1,2} = f((0,0) + (0,1)) = f(0,1) \\ -1 &= A_{1,3} = f((0,0) + (0,2)) = f(0,2) \\ 2 &= A_{1,4} = f((0,0) + (1,0)) = f(1,0) \\ -3 &= A_{1,5} = f((0,0) + (1,1)) = f(1,1) \\ 0 &= A_{1,6} = f((0,0) + (1,2)) = f(1,2) \\ 1 &= A_{1,7} = f((0,0) + (2,0)) = f(2,0) \\ 2 &= A_{1,8} = f((0,0) + (2,1)) = f(2,1) \\ -1 &= A_{1,9} = f((0,0) + (2,2)) = f(2,2). \end{aligned}$$

Como $(\mathbb{Z}/3\mathbb{Z})^2$ tem 9 elementos, teremos 9 caracteres χ_j associados a seus respectivos v_{χ_j}:

$$\begin{aligned}
\chi_0 &\to v_{\chi_0} = (1,1,1,1,1,1,1,1,1),\\
\chi_1 &\to v_{\chi_1} = (1, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}, 1, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}, 1, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}})\\
\chi_2 &\to v_{\chi_2} = (1, e^{\frac{4\pi i}{3}}, e^{\frac{2\pi i}{3}}, 1, e^{\frac{4\pi i}{3}}, e^{\frac{2\pi i}{3}}, 1, e^{\frac{4\pi i}{3}}, e^{\frac{2\pi i}{3}})\\
\chi_3 &\to v_{\chi_3} = (1,1,1, e^{\frac{2\pi i}{3}}, e^{\frac{2\pi i}{3}}, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}, e^{\frac{4\pi i}{3}}, e^{\frac{4\pi i}{3}})\\
\chi_4 &\to v_{\chi_4} = (1, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}, 1, e^{\frac{4\pi i}{3}}, 1, e^{\frac{2\pi i}{3}})\\
\chi_5 &\to v_{\chi_5} = (1, e^{\frac{4\pi i}{3}}, e^{\frac{2\pi i}{3}}, e^{\frac{2\pi i}{3}}, 1, e^{\frac{4\pi i}{3}}, e^{\frac{4\pi i}{3}}, e^{\frac{2\pi i}{3}}, 1)\\
\chi_6 &\to v_{\chi_6} = (1,1,1, e^{\frac{4\pi i}{3}}, e^{\frac{4\pi i}{3}}, e^{\frac{4\pi i}{3}}, e^{\frac{2\pi i}{3}}, e^{\frac{2\pi i}{3}}, e^{\frac{2\pi i}{3}})\\
\chi_7 &\to v_{\chi_7} = (1, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}, e^{\frac{4\pi i}{3}}, 1, e^{\frac{2\pi i}{3}}, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}, 1)\\
\chi_8 &\to v_{\chi_8} = (1, e^{\frac{4\pi i}{3}}, e^{\frac{2\pi i}{3}}, e^{\frac{4\pi i}{3}}, e^{\frac{2\pi i}{3}}, 1, e^{\frac{2\pi i}{3}}, 1, e^{\frac{4\pi i}{3}})
\end{aligned}$$

Para encontrarmos os autovalores, basta calcularmos, para cada $0 \leq k \leq 8$, a expressão $\left|\sum_{j=0}^{8} f(j_1, j_0)\chi_k(j_1, j_0)\right|$, que nos fornecerá os seguintes autovalores:

$$1, \sqrt{31}, -\sqrt{31}, \sqrt{7}, 2\sqrt{7}, 2\sqrt{7}, -\sqrt{7}, 2\sqrt{7}, -2\sqrt{7}.$$

3.4 Convergência fraca* da S-convolução

Sabemos que a S-convolução não diminui a entropia das medidas σ-invariantes envolvidas. Contudo, esse crescimento é limitado em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$, pois sabemos que a medida λ Bernoulli realiza a entropia topológica. Em outras palavras, como $h_\lambda(\sigma) = \log p$, dada qualquer sequência de S-convoluções $\eta_n * \eta_{n-1} * \cdots * \eta_1$ temos que

$$0 \leq h_{\eta_n * \eta_{n-1} * \cdots * \eta_1}(\sigma) \leq \log p.$$

Sendo uma sequência não-decrescente, o limite quando n tende ao infinito existe. De fato, acreditamos que:

Conjectura. *Seja $(\eta_n)_{n\in\mathbb{N}}$ uma sequência de probabilidades σ-invariantes em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Se para todo $m \in \mathbb{N}^*$ e para todo caracter não trivial χ constante em cilindros de tamanho m, tivermos*

$$\prod_{n=1}^{\infty}\left|\int \chi \ d\eta_n\right| \to 0,$$

então

$$h_{\eta_n*\eta_{n-1}*\cdots*\eta_1}(\sigma) \to \log p.$$

Por outro lado, o limite na convergência fraca* é mais fácil de analisar, visto que essa convergência fica caracterizada através de cilindros e também, pelo Corolário 3.2.6, por caracteres constantes em cilindros de tamanho m. Trataremos deste limite nessa seção.

Proposição 3.4.1. *Sejam $\eta, \mu \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ medidas de probabilidade e χ um caracter de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ constante em cilindros de tamanho m. Então*

$$\int \chi \ d_{\eta*\mu} = \int \chi \ d_\eta \int \overline{\chi} \ d_\mu.$$

Demonstração. Note que $\int \chi \ d_\eta \int \overline{\chi} \ d_\mu$ é

$$\sum_{j=1}^{p^m} \eta[j_{m-1},\cdots,j_1,j_0]\chi([j_{m-1},\cdots,j_1,j_0]) \sum_{k=1}^{p^m} \mu[k_{m-1},\cdots,k_1,k_0]\overline{\chi}([k_{m-1},\cdots,k_1,k_0]) =$$

$$= \sum_{j,k=1}^{p^m} \eta[j_{m-1},\cdots,j_1,j_0]\mu[k_{m-1},\cdots,k_1,k_0]\chi([j_{m-1}-k_{m-1},\cdots,j_1-k_1,j_0-k_0])$$

$$= \sum_{k,l=1}^{p^m} \eta[l_{m-1}+k_{m-1},\cdots,l_1+k_1,l_0+k_0]\mu[k_{m-1},\cdots,k_1,k_0]\chi([l_{m-1},\cdots,l_1,l_0])$$

$$= \sum_{l=1}^{p^m} \eta*\mu[l_{m-1},\cdots,l_1,l_0]\chi([l_{m-1},\cdots,l_1,l_0]) = \int \chi \ d_{\eta*\mu}$$

$\square$

Corolário 3.4.2. *Sejam $\eta_1, \eta_2, \eta_3 \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ medidas de probabilidade e χ um caracter de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ constante em cilindros de tamanho m. Então*

$$\left|\int \chi \, d_{(\eta_1 * \eta_2) * \eta_3}\right| = \left|\int \chi \, d_{\eta_1 * (\eta_2 * \eta_3)}\right|.$$

Demonstração. O corolário segue imediatamente da Proposição 3.4.1. $\square$

Teorema 3.4.3. *Seja $(\eta_n)_{n \in \mathbb{N}}$ uma sequência de probabilidades σ-invariantes em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Se para todo $m \in \mathbb{N}^*$ e para todo caracter não trivial χ constante em cilindros de tamanho m, tivermos*

$$\prod_{n=1}^{\infty} \left|\int \chi \, d\eta_n\right| \to 0,$$

então

$$\eta_n * \eta_{n-1} * \cdots * \eta_1 \to \left(\frac{1}{p}, \frac{1}{p}, \cdots, \frac{1}{p}\right)$$

na topologia fraca.*

Demonstração. Seja χ um caracter não trivial qualquer, constante em cilindros de tamanho m. Pela Proposição 3.4.1 temos

$$\left|\int \chi \, d\eta_n * \eta_{n-1} * \cdots * \eta_1\right| = \prod_{k=1}^{n} \left|\int \chi \, d\eta_k\right|.$$

Mas, por hipótese,

$$\prod_{k=1}^{n} \left|\int \chi \, d\eta_k\right| \underset{n \to \infty}{\to} 0.$$

Logo, pelo Corolário 3.2.6 e pelo Exemplo 3.2.7 , temos que $\eta_n * \eta_{n-1} * \cdots * \eta_1 \to \left(\frac{1}{p}, \frac{1}{p}, \cdots, \frac{1}{p}\right)$. $\square$

Exemplo 3.4.4. Sejam $p = 2$ e as medidas de Bernoulli dadas pelos vetores de probabilidade $v_{\eta_n} = \left(\frac{1}{n}, 1 - \frac{1}{n}\right)$, para todo $n \geq 3$. Por se tratarem de medidas de Bernoulli, só precisamos analisar cilindros de tamanho 1, e de caracteres constantes em cilindros de tamanho 1. Isso porque qualquer medida de Bernoulli em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ fica completamente definida pelos valores que assume em cada cilindro de tamanho 1. Portanto, basta-nos saber que

$$\begin{cases} \eta_n[0] = \frac{1}{n} \\ \eta_n[1] = 1 - \frac{1}{n}. \end{cases}$$

Seja χ um caracter não trivial. Então

$$\prod_{k=3}^{N} \left| \int \chi \, d\eta_k \right| \leq \prod_{k=3}^{N} \left(1 - \frac{1}{k} \right) = \prod_{k=3}^{N} \frac{k-1}{k} = \frac{2}{N} \to 0$$

Logo,

$$\eta_n * \eta_{n-1} * \cdots * \eta_1 \to \left(\frac{1}{2}, \frac{1}{2} \right).$$

Capítulo 4

Operador de Ruelle Generalizado

Neste capítulo exploramos uma generalização do operador de transferência e o mostramos resultado análogo ao Teorema de Ruelle-Perron-Frobenius para tal operador.

Os resultados contidos em [13] pressupõem um conjunto finito $A = \{1, \cdots, d\}$, que podemos imaginar como sendo um espaço amostral de um experimento, e a definição de $X = A^{\mathbb{N}}$, que seria uma sequência de eventos contidas em A. Neste trabalho, e em [11], A pode ser um espaço métrico compacto qualquer. Nesse sentido, na definição do operador de transferência, o somatório dá lugar a integral.

Finalmente, em [11] a medida utilizada nesta integral é sempre a mesma, chamada de medida de probabilidade a priori. Neste trabalho tal medida varia de acordo com $x \in X$, ou seja, para cada $x \in X$ teremos uma medida de probabilidade ν_x associada, definida na pré-imagem de x pelo shift (σ).

Incialmente apresentamos os resultados já obtidos em [13] e [11]. Na segunda parte deste capítulo apresentamos a nova definição e adaptamos tais resultados. Este capítulo terá como forte referência [13].

4.1 O Operador de Ruelle

Nesta seção faremos um resumo das definições e resultados que estão em [13].

Seja $S = \{1, 2, 3, \cdots, k\}$ nosso alfabeto e $X = S^{\mathbb{N}}$, o espaço das sequências $(x_n)_{n\in\mathbb{N}}$ com $x_n \in S$ para todo n. Vamos munir S com a topologia discreta (assim S é compacto pois é finito) e X pela respectiva topologia produto, ou seja, a topologia gerada pelos cilindros $[m; a_m, ..., a_n]$, onde $a_i \in S$, $\forall\ m \leq i \leq n$. Então pelo teorema de Tychonoff, X é compacto. Chamamos de *shift* a aplicação

$$\begin{aligned} \sigma: \quad & X \to X \\ & (x_n)_n \to (x_{n+1})_n, \end{aligned}$$

ou seja, a aplicação que "descarta"o primeiro termo da sequência. Considere $\theta \in (0, 1)$ e defina

$$\begin{aligned} d(.,.): \quad & X \times X \to \mathbb{R} \\ & (x, y) \to d(x, y) = \theta^N, \end{aligned}$$

onde N é o maior natural tal que $x_i = y_i$ para $i < N$. É possível mostrar que esta métrica gera a topologia produto de X.

Definição 4.1.1. Seja $f : X \to \mathbb{C}$ uma função contínua e $n \geq 0$, definimos

$$var_n(f) = \sup\{|f(x) - f(y)| : x_i = y_i,\ i < n\}.$$

Proposição 4.1.2. *Seja $f : X \to \mathbb{C}$ uma função contínua. Então $|f(x) - f(y)| \leq Cd(x, y)$ se e somente se $var_n(f) \leq C\theta^n$, para todo $n \in \mathbb{N}$.*

Demonstração. Suponha que $|f(x) - f(y)| \leq Cd(x, y)$ e escolha $n \in \mathbb{N}$. Se $x_i = y_i$ para

$i < n$, então $|f(x) - f(y)| \leq C\theta^n$. Logo $var_n(f) \leq C\theta^n$.

Reciprocamente, suponha que $var_n(f) \leq C\theta^n$, para todo $n \in \mathbb{N}$. Sejam $x, y \in X$. Suponha que $d(x,y) = \theta^N$. Então $var_N(f) \leq C\theta^N$ e assim $|f(x) - f(y)| \leq Cd(x,y)$. □

Seja $F = \{f : X \to \mathbb{C} : f \text{ é continua}, var_n(f) \leq C\theta^n, n \in \mathbb{N}, \text{para algum } C > 0\}$. Pela Proposição 4.1.2, F é o conjunto das funções Lipschitz. Para $f \in F$ defina $|f|_\infty = \sup\{|f(x)| : x \in X\}$ e $|f|_\theta = \sup\{\frac{var_n(f)}{\theta^n},\ n \in \mathbb{N}\}$.

Definição 4.1.3. Para $f \in F$ nós definimos $\|f\| = |f|_\infty + |f|_\theta$.

Proposição 4.1.4. *$(F, \|\cdot\|)$ é um espaço normado.*

Demonstração. Ver [13], página 13. □

Proposição 4.1.5. *O espaço normado $(F, \|.\|)$ é um espaço de Banach. Além disso, $\|fg\| \leq \|f\||g|_\infty + |f|_\infty\|g\|$ e $\left\|\frac{1}{f}\right\| \leq \left|\frac{1}{f^2}\right|_\infty \cdot \|f\|$.*

Demonstração. Ver [13], página 13. □

Definição 4.1.6. Definimos o operador de Ruelle $L_f : C(X) \to C(X)$ por

$$L_f(w) = \sum_{\sigma(y)=x} e^{f(y)} w(y).$$

Não é difícil ver que o operador L_f é linear e contínuo. Quando f é real e $L_f(1) = 1$, dizemos que o operador está normalizado.

Teorema 4.1.7. *(**Ruelle Perron Frobenius - RPF**) Seja $f \in F$ tomando valores reais. Então:*

(i.) Existe um autovalor simples maximal β de $L_f : C(X) \to C(X)$ correspondente a uma autofunção $h \in F$ estritamente positiva.

(ii.) O espectro de $L_f : F \to F$(excluindo $\beta > 0$) está contido em um disco com um raio estritamente menor que β.

(iii.) Há uma única medida de probabilidade μ tal que $L_f^* \mu = \beta\mu$, ou seja, $\int L_f(w)\, d\mu = \beta \int w\, d\mu,\ \forall\, w \in C(X)$.

(iv.) $\frac{1}{\beta^n} L_f^n(w) \to h \int w\, d\mu$ uniformemente para toda $w \in C(X)$, onde h é a autofunção do item (i) e $\int h\, d\mu = 1$.

Demonstração. A demonstração pode ser encontrada em [13], a partir da página 21. □

Já no contexto de [11], considera-se S como um espaço métrico compacto qualquer (M, d_M). Após isso, os autores definem a métrica em $X = M^{\mathbb{N}}$

$$d(x, y) = \sum_{n=1}^{\infty} \frac{1}{2^n} d_M(x_n, y_n).$$

Usam a notação H_α para as funções Hölder $A : X \to \mathbb{R}$ com a norma dada por

$$\|A\|_\alpha = |A|_\infty + |A|_\alpha$$

onde

$$|A|_\infty = \sup_{x \in X} |A(x)| \ e \ |A|_\alpha = \sup_{x \neq y} \frac{|A(x) - A(y)|}{d(x, y)^\alpha}.$$

Então fixam uma medida de probabilidade a-priori ν sobre a sigma-álgebra de Borel de M, assumindo que o suporte desta medida é M. Fixando $A \in H_\alpha$, o operador de Ruelle $L_A : C(X) \to C(X)$ é definido neste contexto por

$$L_A(\varphi)(x) = \int_M e^{ax} \varphi(ax) d\nu(a),$$

onde $x \in X$ e $ax = (a, x_1, x_2, \cdots)$ denota a pré-imagem de x com $a \in M$.

Desta forma, os autores demonstram um resultado similar ao Teorema 4.1.7.

4.2 Operador de Transferência Generalizado

Quando A é um conjunto finito, podemos definir uma matriz de Transição T e associar a ela um subshifts do tipo finito. Tais objetos já foram amplamente estudados. Entretanto, no caso em que A é um conjunto infinito, os resultados para subshifts são relativamente novos.

Se A um espaço métrico compacto e $\Gamma \subset A \times A$ é um conjunto fechado, podemos definir

$$\Gamma^{\infty} = \{(x_i)_1^{\infty} : (x_i, x_{i+1}) \in \Gamma,\ i = 1, 2, \cdots, \},$$

e considerar a aplicação shift $\sigma : \Gamma^{\infty} \to \Gamma^{\infty}$. Este novo conceito de subshift, bem como sua entropia, já foi explorado em alguns textos, como por exemplo em [3].

Note que, nesse novo contexto, $\sigma^{-1}(x)$ pode variar de acordo com o elemento que tomamos. Assim como a pré-imagem varia, a medida sobre ela também pode variar.

Pensando nisso, o que vamos fazer nessa seção é introduzir um operador de transferência cuja medida varia de acordo com x e demonstrar o Teorema de Ruelle-Perron-Frobenius nesse contexto.

Fixe um espaço métrico compacto (A, d_A) que será nosso alfabeto e tome $X = A^{\mathbb{N}}$. Como A é compacto, seja D_A o diâmetro de A. Definiremos em X a topologia gerada pelos cilindros $[m; A_m, \cdots, A_n]$ onde os conjuntos $A_m, \cdots, A_n$ são abertos de A com respeito a métrica d_A. Então pelo teorema de Tychonoff, X é compacto. Tomando $0 < \theta < \frac{1}{2}$, definimos

$$d(x, y) = \sum_{i=1}^{\infty} \theta^{i-1} d_A(x_i, y_i). \tag{4.1}$$

Proposição 4.2.1. *A métrica dada por (4.1) gera a topologia produto do espaço X.*

Demonstração. De fato, tome um cilindro qualquer $[m; A_m, \cdots, A_n]$. Escolha $x \in X$ tal que $x_m \in A_m, \cdots, x_n \in A_n$. Como os A_i's são abertos, existem $\varepsilon_m, \cdots, \varepsilon_n$ tais que $B_A(x_i, \varepsilon_i) \subset A_i$, para todo $m \leq i \leq n$. Tomando $\varepsilon = \min\{\varepsilon_i : m \leq i \leq n\}$, temos que se $y \in B(x, \theta^n \varepsilon)$

$$\sum_{i=1}^{\infty} \theta^{i-1} d_A(x_i, y_i) < \theta^n \varepsilon$$

e então $\theta^{i-1} d(x_i, y_i) < \theta^n \varepsilon, \forall\, m \leq i \leq n$, o que implica que $d(x_i, y_i) < \varepsilon, \forall\, m \leq i \leq n$. Assim $B(x, \theta^n \varepsilon) \subset [m; A_m, \cdots, A_n]$. Por outro lado, tome $x \in X$, $r > 0$ e considere a bola $B(x, r)$. Tome $s < \frac{r(1-\theta)}{2}$ e n tal que $\frac{\theta^n D_A}{1-\theta} < \frac{r}{2}$, onde D_A é o diâmetro de A. Considere os abertos $A_1 = B(x_1, s), \cdots, A_n = B(x_n, s)$ de A. Desta forma o cilindro $[1; A_1, \cdots, A_n] \subset B(x, r)$. De fato, se $y \in [1; A_1, \cdots, A_n]$ então $d_A(x_i, y_i) < s$ para todo $1 \leq i \leq n$. Assim

$$\begin{aligned} d(x,y) = \sum_{i=1}^{\infty} \theta^{i-1} d_A(x_i, y_i) &= \sum_{i=1}^{n} \theta^{i-1} d_A(x_i, y_i) + \sum_{i=n+1}^{\infty} \theta^{i-1} d_A(x_i, y_i) \\ &\leq \sum_{i=1}^{n} \theta^{i-1} s + \sum_{i=n+1}^{\infty} \theta^{i-1} D_A \\ &\leq \frac{s}{1-\theta} + \frac{\theta^n D_A}{1-\theta} < \frac{r}{2} + \frac{r}{2} = r, \end{aligned}$$

e portanto $y \in B(x, r)$. □

Como anteriormente $\sigma \colon X \to X$ é o shift e $C(X)$ é o espaço das funções contínuas em X. Denotaremos por D_X o diâmetro do conjunto X. Note que agora a pré-imagem de um ponto x é o conjunto $\{ax : a \in A\}$. Desta forma, podemos identificar $\sigma^{-1}(x)$ com o conjunto A.

Observação 4.2.2. Seja L o conjunto das funções Lipschitz em X e $\mathcal{B}$ a σ-álgebra gerada pelos cilindros em X. Durante o restante do texto, se f é uma função Lipschitz, denotaremos por k_f sua constante de Lipschitz.

Fixando $g \in L$, tomaremos um conjunto V_g de funções que leva cada elemento de X em uma probabilidade de A com uma propriedade especial. De maneira mais precisa,

$$V_g = \{\nu : X \to \mathcal{P}(A) | \nu_x(B) \leq \nu_y(B) e^{\frac{\theta}{1-2\theta} k_g d(x,y)}, \forall\ B\ \in\ \mathcal{B}\}.$$

Ainda, assumiremos que ν_x tem suporte total para todo $x \in X$.

Proposição 4.2.3. *Se $\nu \in V_g$, então $\nu_x \approx \nu_y$ para todo $x, y \in X$.*

Demonstração. Suponha que $\nu_y(B) = 0$, para algum $B\ \in\ \mathcal{B}$. Então

$$\nu_x(B) \leq \nu_y(B) e^{\frac{\theta}{1-2\theta} k_g d(x,y)} = 0.$$

Logo, $\nu_x(B) = 0$. Analogamente, se $\nu_x(B) = 0$ temos que $\nu_y(B) = 0$. □

Proposição 4.2.4. *Se $\nu \in V_g$, então existe uma constante $k_\nu \geq 0$ tal que $\forall\ x, y \in X$ temos $|\nu_x - \nu_y| \leq k_\nu d(x,y)$, onde $|\nu_x - \nu_y| = \sup_{B \in \mathcal{B}} |\nu_x(B) - \nu_y(B)|$.*

Demonstração. Seja B em $\mathcal{B}$. Então

$$\begin{aligned}
\nu_x(B) - \nu_y(B) &\leq \nu_y(B) \left(e^{\frac{\theta}{1-2\theta} k_g d(x,y)} - 1 \right) \\
&\overset{*}{\leq} \nu_y(B) \frac{\theta}{1-2\theta} k_g d(x,y) \left(e^{\frac{\theta}{1-2\theta} k_g d(x,y)} \right) \\
&\leq \frac{\theta}{1-2\theta} k_g d(x,y) \left(e^{\frac{\theta}{1-2\theta} k_g d(x,y)} \right) \\
&\leq k_\nu d(x,y),
\end{aligned}$$

onde $k_\nu = \frac{\theta}{1-2\theta} k_g \left(e^{\frac{\theta}{1-2\theta} k_g D_X} \right)$, e em $(*)$ usamos a expansão de Taylor da função e^x. □

Fixando $g \in L$ e $\nu \in V_g$, definimos então o Operador de Transferência Generalizado, $L_g : C(X) \to C(X)$, por

$$L_g \varphi(x) = \int_A e^{g(ax)} \varphi(ax) d\nu_x(a)$$

Note que L_g é linear e também positivo, ou seja, $L_g(\varphi) \geq 0$ sempre que $\varphi \geq 0$. Ainda,

$$|L_g(\varphi)|_\infty = \sup_{x \in X} |L_g \varphi(x)| \leq |\varphi|_\infty \int_A e^{g(ax)} d\nu_x(A) = L_g(1)|\varphi|_\infty$$

Logo, L_g é contínuo.

Lema 4.2.5. *Se $\varphi \in L$ e $v \in V_g$, então $L_g(\varphi) \in L$.*

Demonstração. Como X é compacto, temos que $e^g \varphi \in L$. Então

$$\begin{aligned}
&\frac{|L_g(\varphi)(x) - L_g(\varphi)(y)|}{d(x,y)} = \frac{|\int_A e^{g(ax)}\varphi(ax)d\nu_x(a) - \int_A e^{g(ay)}\varphi(ay)d\nu_y(a)|}{d(x,y)} = \\
&= \frac{|\int_A e^{g(ax)}\varphi(ax)d\nu_x - \int_A e^{g(ax)}\varphi(ax)d\nu_y + \int_A e^{g(ax)}\varphi(ax)d\nu_y - \int_A e^{g(ay)}\varphi(ay)d\nu_y|}{d(x,y)} \\
&= \frac{|\int_A e^{g(ax)}\varphi(ax)d(\nu_x - \nu_y) + \int_A e^{g(ax)}\varphi(ax) - e^{g(ay)}\varphi(ay)d\nu_y|}{d(x,y)} \\
&\leq \frac{\int_A |e^{g(ax)}\varphi(ax)|d|\nu_x - \nu_y| + \int_A |e^{g(ax)}\varphi(ax) - e^{g(ay)}\varphi(ay)|d\nu_y}{d(x,y)} \\
&\leq mk_\nu + \theta k_{e^g\varphi}.
\end{aligned}$$

onde $m = max\{e^{g(x)}\varphi(x) : x \in X\}$. □

Teorema 4.2.6. *Fixados $g \in L$ e $\nu \in V_g$, existem $h \in L$, estritamente positiva, e $\beta > 0$ tais que $L_g(h) = \beta h$.*

Demonstração. Seja

$$\Lambda = \{f \in C(X) : 0 \leq f \leq 1,\ f(x) \leq f(y)e^{\frac{k_g d(x,y)}{1-2\theta}}\}.$$

Observamos que Λ é um conjunto convexo. Agora, seja $\{f_k\}_{k\in\mathbb{N}}$ uma sequência em Λ convergindo para f uniformemente. Como a convergência é uniforme, $f \in C(X)$. Ainda, $0 \leq f \leq 1$ e

$$f(x) = \lim\ f_k(x) \leq \lim\ f_k(y)e^{\frac{k_g d(x,y)}{1-2\theta}} = f(y)e^{\frac{k_g d(x,y)}{1-2\theta}}.$$

Então Λ é uniformemente fechado. Além disso, note que

$$-f(x)\left(e^{\frac{k_g d(x,y)}{1-2\theta}}-1\right) \leq f(x)-f(y) \leq f(y)\left(e^{\frac{k_g d(x,y)}{1-2\theta}}-1\right).$$

Desta forma

$$\begin{aligned}
|f(x)-f(y)| &\leq \max\left\{|f(x)|\left(e^{\frac{k_g d(x,y)}{1-2\theta}}-1\right), |f(y)|\left(e^{\frac{k_g d(x,y)}{1-2\theta}}-1\right)\right\} \\
&\leq \left(e^{\frac{k_g d(x,y)}{1-2\theta}}-1\right) \\
&\leq \frac{k_g d(x,y)}{1-2\theta} e^{\frac{k_g D_X}{1-2\theta}},
\end{aligned}$$

e concluimos que $f \in L$. Então Λ é uma família equicontínua de funções, e visto que $\{f(x) : f \in \Lambda\}$ é relativamente compacto, podemos aplicar o teorema de Arzela-Ascoli e concluir que Λ é compacto com respeito a norma da convergência uniforme.

Para $n \geq 1$, definimos

$$L_n(f) = \frac{L_g(f+\frac{1}{n})}{|L_g(f+\frac{1}{n})|_\infty},$$

para $f \in \Lambda$. Note que $|L_n(f)|_\infty = 1$ e para $x, y \in X$ e $\nu \in V_g$ temos que

$$\begin{aligned}
L_g\left(f+\frac{1}{n}\right)(x) &= \int_A e^{g(ax)}\left(f+\frac{1}{n}\right)(ax)d\nu_x \\
&\leq \int_A e^{g(ax)}\left(f+\frac{1}{n}\right)(ay)e^{\frac{k_g\theta d(x,y)}{1-2\theta}}d\nu_x \\
&\leq \int_A e^{g(ay)}e^{k_g d(x,y)}\left(f+\frac{1}{n}\right)(ay)e^{\frac{k_g\theta d(x,y)}{1-2\theta}}d\nu_x \\
&\leq \int_A e^{g(ay)}e^{k_g d(x,y)}\left(f+\frac{1}{n}\right)(ay)e^{\frac{k_g\theta d(x,y)}{1-2\theta}}e^{\frac{\theta}{1-2\theta}k_g d(x,y)}d\nu_y \\
&= \int_A e^{g(ay)}\left(f+\frac{1}{n}\right)(ay)e^{k_g d(x,y)\left(\frac{1-2\theta+\theta+\theta}{1-2\theta}\right)}d\nu_y \\
&= L_g\left(f+\frac{1}{n}\right)(y)e^{\frac{k_g d(x,y)}{1-2\theta}}.
\end{aligned}$$

Então $L_n : \Lambda \to \Lambda$. Visto que $\Lambda \subset C(X)$ é um conjunto convexo uniformemente compacto, podemos aplicar o teorema do ponto fixo de Schauder-Tychonov para cada L_n, e concluir que existe $h_n \in \Lambda$ com $L_g(h_n + \frac{1}{n}) = \beta_n h_n$, onde $\beta_n = |L_g(h_n + \frac{1}{n})|_\infty$.

Pela compacidade de Λ, existe uma subsequência h_{n_k} de $(h_n)_{n\in\mathbb{N}}$ convergindo para $h \in \Lambda$. Então,

$$L_g(h) = L_g\left(\lim_{k\to\infty} h_{n_k} + \frac{1}{n_k}\right) = \lim_{k\to\infty} L_g\left(h_{n_k} + \frac{1}{n_k}\right) = \lim_{k\to\infty} \beta_{n_k} h_{n_k} = \beta h,$$

onde $\beta = |L_g(h)|_\infty$.

Para mostrar que β é estritamente positivo note que

$$\beta_n h_n(x) = L_g\left(h_n + \frac{1}{n}\right)(x) = \int_A e^{g(ax)}\left(h_n + \frac{1}{n}\right)(ax)d\nu_x \geq \left(\inf_{x\in X}\ h_n(x) + \frac{1}{n}\right)e^{-|g|_\infty}$$

e assim $\beta_n(\inf\limits_{x\in X}\ h_n(x)) \geq \left(\inf\limits_{x\in X}\ h_n(x) + \dfrac{1}{n}\right)e^{-|g|_\infty} > 0$, para todo $n \in \mathbb{N}$. Desse modo, $\inf\limits_{x\in X}\ h_n(x) > 0$ e $\beta_n \geq e^{-|g|_\infty} \Rightarrow \beta \geq e^{-|g|_\infty} > 0$.

Para mostrar que h é estritamente positiva, podemos assumir por contradição que $h(x) = 0$, para algum $x \in X$. Assim

$$\int_A e^{g(ax)} h(ax)d\nu_x = L_g(h)(x) = \beta h(x) = 0$$

e concluímos que a função $a \to h(ax)$ é nula para ν_x quase todo ponto. Agora seja $z \in X$. Então

$$\begin{aligned}
L_g(h)(z) = \int_A e^{g(az)} h(az)d\nu_z &\leq e^{|g|_\infty}\int_A h(az)d\nu_z \\
&\leq e^{|g|_\infty}\int_A h(ax)e^{\frac{k_g\theta d(z,x)}{1-2\theta}}d\nu_z \\
&\leq e^{|g|_\infty}\int_A h(ax)e^{\frac{k_g d(z,x)}{1-2\theta}}e^{\frac{\theta}{1-2\theta}k_g d(z,x)}d\nu_x. \\
&= e^{|g|_\infty}\int_A e^{\frac{(1+\theta)k_g d(z,x)}{1-2\theta}} h(ax)d\nu_x = 0.
\end{aligned}$$

Por conseguinte, $\beta = \|L_g(h)\|_\infty = 0$, o que contradiz a desigualdade $\beta > 0$ que obtivemos anteriormente. Logo h é estritamente positiva. $\square$

No Teorema 4.2.6 podemos garantir a existência da função h considerando o conjunto Λ. O que faremos agora é garantir a unicidade (a menos de multiplicação por constante) de tal função, associada ao autovalor β, em todo o espaço $C(X)$.

Para afimar que $f = g$ precisamos que $f(x) = g(x)$, para todo $x \in X$. Como em teoria da medida as afirmações são feitas em ν quase todo ponto, usaremos a continuidade da função para garantir a igualdade em todo conjunto X. Com esse intuito, temos os três lemas abaixo:

Lema 4.2.7. *Sejam $\nu \in V_g$, $z \in X$ e $\delta > 0$. Suponha que $w \in C(X)$, com $w \geq 0$, satisfaça ${L_g}^N(w)(y) = 0$, para algum $y \in X$ e algum $N \in \mathbb{N}$. Então existem $a_1, \cdots, a_N$ tais que $d(a_i, z_i) < \delta$, para todo $1 \leq i \leq N$, e $w(a_1 \cdots a_N y) = 0$.*

Demonstração. Comecemos com o caso $N = 1$. Existe $y \in X$ tal que

$$0 = L_g(w)(y) = \int e^{g(ay)} w(ay) d_{\nu_y}(a).$$

Assim, $\{a \in A : w(ay) = 0\}$ é conjunto de medida ν_y total. Como ν_y dá peso positivo para todos os abertos de A, vai existir a_1 tal que $w(a_1 y) = 0$ e $d(a_1, z_1) < \delta$.

Para $N = 2$,

$$0 = L_g^2(w)(y) = \int e^{g(ay)} L_g(w)(ay) d_{\nu_y}(a).$$

Assim, temos que $\{a \in A : L_g(w)(ay) = 0\}$ é conjunto de medida ν_y total e, portanto, $\exists\ a_2 \in B(z_2, \delta) \cap \{a \in A : L_g(w)(ay) = 0\}$.

Portanto,

$$0 = L_g(w)(a_2 y) = \int e^{g(aa_2 y)} w(aa_2 y) d_{\nu_{a_2 y}}(a).$$

Assim, $\{a \in A : w(aa_2 y) = 0\}$ é conjunto de medida $\nu_{a_2 y}$ total. Como $\nu_{a_2 y}$ dá peso positivo para todos os abertos, segue que $\exists\ a_1 \in B(z_1, \delta) \cap \{a \in A : w(aa_2 y) = 0\}$.

Agora fixe $N = k$. Suponha que $w \in C(X)$, com $w \geq 0$, satisfaça ${L_g}^k(w)(y) = 0$, para algum $y \in X$ e algum $N \in \mathbb{N}$. Vamos supor, por indução, que existam $a_1, \cdots, a_k$ tais que $d(a_i, z_i) < \delta$, para todo $1 \leq i \leq k$, e $w(a_1 \cdots a_k y) = 0$.

Suponha que $w' \in C(X)$, com $w' \geq 0$, satisfaça ${L_g}^{k+1}(w')(y') = 0$, para algum $y' \in X$.

Note que

$$
{L_g}^k(L_g(w'))(y') = 0.
$$

Pela hipótese de indução, existem $a_2, \cdots, a_{k+1}$ tais que $d(a_i, z_i) < \delta$, para todo $2 \leq i \leq k+1$, e $L_g(w')(a_2 \cdots a_{k+1} y) = 0$. Pelo caso $N = 1$, existe a_1 tal que $d(a_1, z_1) < \delta$ e $w'(a_1, a_2, \cdots, a_{k+1} y) = 0$. □

Lema 4.2.8. *Seja $\nu \in V_g$. Suponha que $w \in C(X)$, com $w \geq 0$, satisfaça ${L_g}^N(w)(y) = 0$, para algum $y \in X$(que depende de N), para todo $N \in \mathbb{N}$. Então $w \equiv 0$.*

Demonstração. Sejam $z \in X$ e $\varepsilon > 0$. Da continuidade de w, vai existir $\delta > 0$ tal que $|w(z) - w(y)| < \varepsilon$ sempre que $y \in X$ satisfizer $d(y, z) < \delta$. Como $\theta \in \left(0, \frac{1}{2}\right)$, seja N um natural tal que

$$
n \geq N \Rightarrow \left[\frac{\delta}{2}\left(\frac{1}{1-\theta}\right) + \left(\sum_{i=n+1}^{\infty} \theta^{i-1}\right) D_X\right] < \delta.
$$

Pelo Lema 4.2.7, vão existir $a_1, \cdots, a_N \in A$ tais que $d(a_i, z_i) < \frac{\delta}{2}$ e $w(a_1 a_2 \cdots a_N y) = 0$. Mas observe que

$$
\begin{aligned}
d(a_1 a_2 \cdots a_N y, z) &= \sum_{i=1}^{N} \theta^{i-1} d(a_i, z_i) + \sum_{i=N+1}^{\infty} \theta^{i-1} d(y_{i-N}, z_i) \\
&\leq \frac{\delta}{2} \sum_{i=1}^{N} \theta^{i-1} + \left(\sum_{i=N+1}^{\infty} \theta^{i-1}\right) D_X \\
&\leq \frac{\delta}{2}\left(\frac{1}{1-\theta}\right) + \left(\sum_{i=N+1}^{\infty} \theta^{i-1}\right) D_X \\
&< \delta.
\end{aligned}
$$

Assim, $|w(z)| = |w(z) - w(a_1 a_2 \cdots a_N y)| < \varepsilon$, e como ε é qualquer, $w(z) = 0$. Portanto, como z é arbitrário, $w \equiv 0$. $\square$

Lema 4.2.9. *Seja $f \in C(X)$, $f \geq 0$, uma autofunção de $L_g : C(X) \to C(X)$ associada a $\beta > 0$. Suponha que $\nu \in V_g$ e que $f(x) = 0$ para algum $x \in X$. Então $f \equiv 0$.*

Demonstração. Note que ${L_g}^N(f)(x) = \beta^N f(x) = 0$, para todo $N \in \mathbb{N}$. Agora, basta utilizarmos o Lema 4.2.8. $\square$

Proposição 4.2.10. *Nas mesmas hipóteses do Teorema* 4.2.6, *podemos concluir que o autovalor β é simples, ou seja, a autofunção h é única (a menos de multiplicação por constante).*

Demonstração. Para mostrar que β é simples, seja $h' \in C(X)$ autofunção de L_g, associada a β. Tome $t = \left\{\inf \frac{h'(x)}{h(x)}\right\} = \frac{h'(y)}{h(y)}$, para algum $y \in X$. Então temos que $h'(y) - th(y) = 0$ e $h'(x) - th(x) \geq 0$, para todo $x \in X$. Se definirmos $f(x) = h'(x) - th(x)$, temos que

$$
\begin{aligned}
(L_g(f))(x) = \int_A e^{g(ax)} f(ax) d\nu_x &= \int_A e^{g(ax)} (h'(ax) - th(ax)) d\nu_x \\
&= \int_A e^{g(ax)} h'(ax) d\nu_x - t \int_A e^{g(ax)} h(ax) d\nu_x \\
&= \beta h'(x) - \beta t h(x) = \beta f(x).
\end{aligned}
$$

Logo, f é uma autofunção de L_g associada a β. Como $f \geq 0$ e $f(y) = 0$, podemos aplicar o Lema 4.2.9 e concluir que $f \equiv 0$. Portanto, $h'(x) = th(x)$, para todo $x \in X$, como queríamos. $\square$

Nem sempre nosso operador de transferência está normalizado, ou seja, nem sempre $L_g(1) = 1$. Mas esta propriedade será útil para a demonstração da unicidade da automedida do operador dual. Portanto, precisamos de uma forma de normalizar o operador de transferência.

Para este propósito, utilizaremos o operador $M_h : C(X) \to C(X)$ que é a multiplicação por h, ou seja,

$$M_h(f)(x) = h(x)f(x),$$

e o seguinte lema:

Lema 4.2.11. *Se* $g \in L$, $f = g - \log h \circ \sigma + \log h - \log \beta$, *então* $L_f = \beta^{-1} M_{h^{-1}} L_g M_h$ *e* $L_f(1) = 1$.

Demonstração. Temos que

$$\begin{aligned}
L_f(w)(x) &= \int_A e^{(g - \log h \circ \sigma + \log h - \log \beta)(ax)} w(ax) d\nu_x \\
&= \int_A \frac{e^{g(ax)} h(ax) w(ax)}{h(x)\beta} d\nu_x \\
&= \frac{1}{h(x)\beta} L_g(hw)(x) \\
&= \left(\beta^{-1} M_{h^{-1}} L_g(M_h w)\right)(x).
\end{aligned}$$

Então

$$\begin{aligned}
L_f(1)(x) &= \left(\beta^{-1} M_{h^{-1}} L_g M_h 1\right)(x) \\
&= \left(\beta^{-1} M_{h^{-1}} L_g(h)\right)(x) \\
&= \left(\beta^{-1} M_{h^{-1}} \beta h\right)(x) \\
&= \frac{\beta h(x)}{\beta h(x)} \\
&= 1,
\end{aligned}$$

para todo $x \in X$. $\square$

Lema 4.2.12. *Seja* $g \in L$*(Observação 4.2.2). Se* $L_g(1) = 1$ *então existe* $C > 0$ *tal que*

$$k_{L_g^n(w)} \leq C|w|_\infty + \theta^n k_w, \text{ para todos } w \in L \text{ e } n \geq 0.$$

Demonstração. Temos que

$$
\begin{aligned}
&|L_g w(x) - L_g w(y)| = \left| \int_A e^{g(ax)} w(ax) d\nu_x - \int_A e^{g(ay)} w(ay) d\nu_y \right| \\
&= | \int_A e^{g(ax)} w(ax) - e^{g(ax)} w(ay) d\nu_x + \int_A e^{g(ax)} w(ay) - e^{g(ay)} w(ay) d\nu_x + \\
&\qquad\qquad + \int_A e^{g(ay)} w(ay) d(\nu_x - \nu_y)| \\
&\leq \int_A e^{g(ax)} |w(ax) - w(ay)| d\nu_x + |w|_\infty \int_A e^{g(ay)} \left(e^{|g(ax)-g(ay)|} - 1 \right) d\nu_x + |e^g w|_\infty \int_A d|\nu_x - \nu_y| \\
&\leq \theta d(x,y) k_w \int_A e^{g(ax)} d\nu_x + |w|_\infty e^{|g|_\infty} (e^{\theta d(x,y) k_g} - 1) \int_A d\nu_x + |e^g w|_\infty d(x,y) k_\nu \\
&\leq \theta d(x,y) k_w + |w|_\infty e^{|g|_\infty} \theta d(x,y) k_g e^{\theta d(x,y) k_g} + |e^g|_\infty |w|_\infty d(x,y) k_\nu \\
&\leq |w|_\infty d(x,y) (\theta e^{|g|_\infty} k_g e^{D_X k_g} + k_\nu |e^g|_\infty) + \theta d(x,y) k_w \\
&= (|w|_\infty (\theta e^{|g|_\infty} k_g e^{D_X k_g} + k_\nu |e^g|_\infty) + \theta k_w) d(x,y)
\end{aligned}
$$

Portanto,

$$k_{L_g(w)} \leq C_0 |w|_\infty + \theta k_w$$

Aplicando a indução: se $k_{L_g^n(w)} \leq C_n |w|_\infty + \theta^n k_w$ então

$$
\begin{aligned}
k_{L_g^{n+1}(w)} = k_{L_g^n(L_g(w))} &\leq C_n |L_g(w)|_\infty + \theta^n k_{L_g(w)} \\
&\leq C_n |w|_\infty + \theta^n \left(C_0 |w|_\infty + \theta k_w \right) \\
&= \left(C_n + \theta^n C_0 \right) |w|_\infty + \theta^{n+1} k_w.
\end{aligned}
$$

Agora note que

$$C_1 = C_0 + \theta C_0$$

$$C_2 = C_1 + \theta^2 C_0 = C_0 + \theta C_0 + \theta^2 C_0$$

$$\vdots$$

$$C_n = C_0 + \theta C_0 + \theta^2 C_0 + \cdots + \theta^n C_0$$

Portanto, tomando $C = \frac{C_0}{1-\theta}$ o resultado está provado. $\square$

Agora que estabelecemos os principais resultados para o operador de transferência, vamos analizar o seu operador dual e encontrar sua automedida.

Seja L_g^* o operador dual de L_g. Então, $L_g^* : \mathcal{M}(X) \to \mathcal{M}(X)$ é definido por

$$\int_X \varphi dL_g^* \mu = \int_X L_g(\varphi(x)) d\mu(x) = \int_X \left(\int_A e^{g(ax)} \varphi(ax) d\nu_x(a) \right) d\mu(x)$$

Este operador é contínuo relativamente a topologia fraca*. De fato, se $(\mu_n) \to \mu$ então para toda função $\varphi \in C(X)$, temos

$$\int_X \varphi dL_g^* \mu_n = \int_X L_g(\varphi) d\mu_n \to \int_X L_g(\varphi) d\mu = \int_X \varphi dL_g^* \mu.$$

Proposição 4.2.13. *Se $\nu \in V_g$, então há uma única medida de probabilidade η tal que ${L_g}^* \eta = \beta \eta$.*

Demonstração. Se L_g não está normalizado, nós definimos $f = g - \log h \circ \sigma + \log h - \log \beta$. Pelo Lema 4.2.11, L_f está normalizado. Então, se μ é uma medida de probabilidade, temos que $L_f^* \mu(X) = 1$. Além disso, L_f^* é linear e positivo. Então o conjunto compacto convexo $\mathcal{M}_1(X)$ é invariante pelo operador contínuo L_f^*. Usando o Teorema de Schauder-Tychonov, L_f^* tem um ponto fixo, ou seja, $L_f^*(\mu) = \mu$. Portanto, para todo $\varphi \in C(X)$,

$$\begin{aligned}
\int_X h\varphi \; d_{{L_g}^*(\frac{\mu}{h})} &= \int_X \frac{L_g(h\varphi)}{h} \; d\mu \\
&= \int_X \int_A \frac{e^{g(ax)} h(ax) \varphi(ax)}{h(x)} \; d\nu_x(a) d\mu(x) \\
&= \int_X \int_A e^{g(ax) - \log h \circ \sigma(ax) + \log h(ax)} \varphi(ax) \; d\nu_x(a) d\mu(x) \\
&= \int_X \int_A \beta e^{g(ax) - \log h \circ \sigma(ax) + \log h(ax) - \log \beta} \varphi(ax) \; d\nu_x(a) d\mu(x)
\end{aligned}$$

$$
\begin{aligned}
&= \int_X \int_A \beta e^{f(ax)} \varphi(ax)\ d\nu_x(a) d\mu(x) \\
&= \int_X \beta L_f(\varphi)\ d\mu \\
&= \int_X \beta\varphi\ dL_f^* \mu \\
&= \beta \int_X \varphi d\mu = \int_X h\varphi\ d_{\beta(\frac{\mu}{h})}.
\end{aligned}
$$

Então, se escolhermos $\eta = \frac{\mu}{h}$, nós temos ${L_g}^* \eta = \beta\eta$.

Vamos provar agora que η é única.

Sejam $x, y \in X$. Usando o Lema 4.2.12, temos para todo $n \geq 1$

$$
\begin{aligned}
|{L_f}^n(w)(x) - {L_f}^n(w)(y)| &\leq k_{{L_f}^n(w)} d(x,y) \\
&\leq d(x,y)C|w|_\infty + d(x,y)\theta^n k_w \\
&\leq d(x,y)C|w|_\infty + d(x,y)k_w \\
&= (C|w|_\infty + k_w)d(x,y).
\end{aligned}
$$

Então $({L_f}^n(w))_{n\in\mathbb{N}}$ é uma família equicontínua de funções Lipschitz. Logo, pelo teorema de Ascoli-Arzelá, $({L_f}^n(w))_{n\in\mathbb{N}}$ é um conjunto relativamente compacto na norma $\|\cdot\|_\infty$. Desta forma, existe alguma subsequência convergente $\{{L_f}^{n_k}(w)\}$. Seja $\overline{w}$ seu limite. Mostraremos, na verdade, que $\overline{w}$ é uma função constante e que a sequência original $({L_f}^n(w))_{n\in\mathbb{N}}$ converge para $\overline{w}$.

Visto que $L_f(1) = 1$, temos que ${L_f}^n(1) = 1$, para todo $n \in \mathbb{N}$. Portanto,

$$
\|w\|_\infty \geq \|L_f(w)\|_\infty \geq \|{L_f}^2(w)\|_\infty \geq \cdots \geq \|\overline{w}\|_\infty.
$$

Como ${L_f}^N$ é contínua, temos

$$
\|{L_f}^N(\overline{w})\|_\infty = \|{L_f}^N(\lim_{k\to\infty} {L_f}^{n_k}(w))\|_\infty
$$

$$
\begin{aligned}
&= \lim_{k\to\infty} \|{L_f}^N({L_f}^{n_k}(w))\|_\infty \\
&\geq \lim_{k\to\infty} \|\overline{w}\|_\infty \\
&= \|\overline{w}\|_\infty.
\end{aligned}
$$

Logo, $\|{L_f}^N(\overline{w})\|_\infty \geq \|\overline{w}\|_\infty$. Mas como $\|{L_f}^N(\overline{w})\|_\infty \leq \|\overline{w}\|_\infty$, temos que $\|\overline{w}\|_\infty = \|{L_f}^N(\overline{w})\|_\infty$, para todo $N \in \mathbb{N}$. Seja $\overline{w}(x_0) = \sup \overline{w} = {L_f}^N(\overline{w})(x_N)$. Defina $v = \overline{w}(x_0) - \overline{w}$. Note que $v \geq 0$ e, para todo $N \in \mathbb{N}$, temos que

$$
\begin{aligned}
{L_f}^N(v)(x_N) &= {L_f}^N(\overline{w}(x_0) - \overline{w})(x_N) \\
&= {L_f}^N(\overline{w}(x_0))(x_N) - {L_f}^N(\overline{w})(x_N) \\
&= \overline{w}(x_0) - \overline{w}(x_0) \\
&= 0.
\end{aligned}
$$

Portanto, pelo Lema 4.2.8, concluímos que $v \equiv 0$. Assim, $\overline{w}$ é constante. Visto que ${L_f}^*\mu = \mu$, temos

$$
\lim {L_f}^{n_k}(w) = \overline{w} = \int_X \overline{w} d\mu = \int_X \lim {L_f}^{n_k}(w) d\mu = \lim \int_X {L_f}^{n_k}(w) d\mu = \int_X w d\mu
$$

Note que $w \in L$, mas como $L \subset C(X)$ é uniformemente denso, nós podemos assumir que $w \in C(X)$. Este argumento funciona para qualquer outra subsequência de ${L_f}^n(w)$. Portanto devemos ter $\lim {L_f}^n(w) = \int_X w d\mu$ uniformemente. Seja λ medida de probabilidade tal que ${L_f}^*\lambda = \lambda$. Então, para todo $w \in C(X)$

$$
\int_X w d\lambda = \overline{w} = \int_X w d\mu.
$$

Logo, $\lambda = \mu$. Visto que μ é única, η é única também. $\square$

A medida de probabilidade μ, tal que ${L_f}^*\mu = \mu$, é σ-invariante. De fato, temos que

$$L_f(v \circ \sigma \cdot w)(x) = \int_A e^{f(ax)} v(\sigma(ax)) w(ax) d\nu_x = [v \cdot L_f(w)](x)$$

para todo $v, w \in C(X)$, e então

$$\int_X v \circ \sigma \ d\mu = \int_X v \circ \sigma \ d{L_f}^*\mu = \int_X L_f(v \circ \sigma \cdot 1) \ d\mu = \int_X v L_f(1) \ d\mu = \int_X v \ d\mu.$$

Corolário 4.2.14. *Temos que* $\frac{1}{\beta^n}{L_g}^n w \to h \int w \ d\eta$ *uniformemente para todo* $w \in C(X)$ *e* $\int h \ d\eta = 1$*, onde* h *é a autofunção obtida no Teorema* 4.2.6.

Demonstração. Pela Proposição 4.2.13, $\lim {L_f}^n(w) = \int_X w d\mu$ uniformemente. Então temos que

$$\begin{aligned} \frac{1}{\beta^n}{L_g}^n w &= \frac{1}{\beta^n}\left(\beta M_h L_f M_{h^{-1}}\right)^n(w) = \\ &= h{L_f}^n\left(\frac{w}{h}\right) \to h\int \frac{w}{h} \ d\mu = h\int w \ d\eta. \end{aligned}$$

Além disso,

$$\int h \ d\eta = \int h \ d\frac{\mu}{h} = \int \ d\mu = 1.$$

□

Proposição 4.2.15. *O restante do espectro de* $L_g : L \to L$ *(excluindo* $\beta > 0$*) está contido em um disco de raio estritamente menor do que* β*.*

Demonstração. A prova é a mesma contida em [13], página 26. □

Exemplo 4.2.16. Sejam $g \in L$, $A = [0, 1]$ e λ a medida de Lebesgue. Desta forma, definimos $\nu : X \to \mathcal{P}(A)$ dada por

$$\nu_x(B) = \int_B e^{g(ax)} d\lambda.$$

Portanto,

$$\begin{aligned}\nu_x(B) &= \int_B e^{g(ax)} d\lambda \\ &\leq \int_B e^{g(ay)} e^{k_g \theta d(x,y)} d\lambda \\ &\leq \int_B e^{g(ay)} e^{\frac{k_g \theta d(x,y)}{1-2\theta}} d\lambda,\end{aligned}$$

pois $1 \leq \frac{1}{1-2\theta}$. Logo,

$$\nu_x(B) \leq \nu_y(B) e^{\frac{k_g \theta d(x,y)}{1-2\theta}}.$$

Podemos então aplicar os resultados desta seção para garantir resultados similares ao Teorema 4.1.7.

Bibliografia

[1] FSV Bazán and Ph L Toint. Conditioning of infinite hankel matrices of finite rank. *Systems & control letters*, 41(5):347–359, 2000.

[2] Sven Feldmann and Georg Heinig. Vandermonde factorization and canonical representations of block hankel matrices. *Linear algebra and its applications*, 241:247–278, 1996.

[3] Shmuel Friedland. Entropy of graphs, semigroups and groups. In *London Mathematical Society Lecture Notes Series*, volume 228, pages 319–343. Cambridge Univ. Press, 1996.

[4] Harry Furstenberg. Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation. *Theory of Computing Systems*, 1(1):1–49, 1967.

[5] Mikhaïl Gromov. On the entropy of holomorphic maps. *preprint*, 1977.

[6] Israel N Herstein. *Topics in algebra*. John Wiley & Sons, 2006.

[7] Pooja Jain and Ram Bilas Pachori. An iterative approach for decomposition of multicomponent non-stationary signals based on eigenvalue decomposition of the hankel matrix. *Journal of the Franklin Institute*, 352(10):4017–4044, 2015.

[8] Graham James Oscar Jameson. *The prime number theorem*, volume 53. Cambridge University Press, 2003.

[9] Aimee SA Johnson. Measures on the circle invariant under multiplication by a nonlacunary subsemigroup of the integers. *Israel Journal of Mathematics*, 77(1-2):211–240, 1992.

[10] Elon Lindenstrauss, David Meiri, and Yuval Peres. Entropy of convolutions on the circle. *Annals of mathematics*, 149:871–904, 1999.

[11] Artur O Lopes, Jairo K Mengue, Joana Mohr, and Rafael R Souza. Entropy and variational principle for one-dimensional lattice systems with a general a priori probability: positive and zero temperature. *Ergodic Theory and Dynamical Systems*, 35(06):1925–1961, 2015.

[12] Bao Luong. *Fourier Analysis on Finite Abelian Groups*. Springer Science & Business Media, 2009.

[13] William Parry and Mark Pollicott. *Zeta functions and the periodic orbit structure of hyperbolic dynamics*. Société mathématique de France, 1990.

[14] J Rissanen. Algorithms for triangular decomposition of block hankel and toeplitz matrices with application to factoring positive matrix polynomials. *Mathematics of computation*, 27(121):147–154, 1973.

[15] Daniel J Rudolph. $\times$ 2 and$\times$ 3 invariant measures and entropy. *Ergodic Theory and Dynamical Systems*, 10(02):395–406, 1990.

[16] David Ruelle. Statistical mechanics of a one-dimensional lattice gas. *Communications in Mathematical Physics*, 9(4):267–278, 1968.

[17] David Ruelle. A measure associated with axiom-a attractors. *American Journal of Mathematics*, pages 619–654, 1976.

[18] Bruno B. Uggioni. *Convergência da convolução de probabilidades invariantes pelo deslocamento*. PhD thesis, Porto Alegre, 2016.

[19] Marcelo Viana and Krerley Oliveira. *Fundamentos da Teoria Ergódica*, volume 90. Rio de Janeiro: SBM, 2014.

[20] Peter Walters. *An introduction to ergodic theory*, volume 79. Springer Science & Business Media, 2000.