Universidade Federal do Rio Grande do Sul

Instituto de Matemática e Estatística

Programa de Pós-Graduação em Matemática

Convergência da convolução de probabilidades invariantes pelo deslocamento

Tese de Doutorado

Bruno Brogni Uggioni

Porto Alegre, 19 de dezembro de 2016

Tese submetida por Bruno Brogni Uggioni[1], como requisito parcial para a obtenção do grau de Doutor em Ciência Matemática, pelo Programa de Pós Graduação em Matemática, do Instituto de Matemática da Universidade Federal do Rio Grande do Sul.

Professor Orientador:

Prof. Dr. Alexandre Tavares Baraviera

Banca examinadora:

Prof. Dr. Alexandre Tavares Baraviera (UFRGS - Orientador)

Prof. Dr. Marcelo Sobottka (UFSC)

Prof. Dr. Leonardo Fernandes Guidi (UFRGS)

Prof. Dr. Paolo Giulietti (UFRGS)

[1]Bolsista da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

Agradecimentos

Agradeço primeiramente a Deus pela produção da tese. Gostaria de agradecer também aos meus pais Edison Uggioni e Adriane Brogni Uggioni e ao meu irmão Hugo Brogni Uggioni e demais familiares por todo o apoio que me deram. Amo-os muito. Agradeço também a minha namorada Juliana Sanches, por toda paciência, atenção e amor. Eu a amo muito. Aos meus amigos da pós, pelas tardes de estudo e momentos de descontração, aos professores da banca, Leonardo Guidi, pelas sugestões quanto a escrita, Marcelo Sobottka, pela atenção também quanto ao projeto de pós-doutorado e ao professor Paolo Giulietti, por toda sinceridade nas críticas feitas à tese e por toda a assistência para que o doutorado sanduíche ocorresse bem. Serei eternamente grato. Agradeço também ao professor Artur Oscar Lopes, por todas as discussões e atenção para com minhas dúvidas, ao meu orientador Alexandre Tavares Baraviera, por ter aceitado ser meu orientador aqui no Brasil e ao professor Mark Pollicott, por ter aceitado ser meu orientador durante o período na Inglaterra. Finalmente, agradeço a CAPES pela bolsa de doutorado aqui no Brasil e pelo CNPq, pela bolsa de doutorado sanduíche no exterior. Apoios financeiros muito importantes para o desenvolvimento dessa tese.

Resumo

Essa tese foi inspirada no artigo [10] de Lindenstrauss et al. e remete ao trabalho fundamental de Furstenberg [5]. Sejam $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ o produto cartesiano unlilateral de infinitas cópias de $\mathbb{Z}/p\mathbb{Z}$ e σ a função *shift* em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Tal espaço é grupo topológico compacto e espaço mensurável quando munido da operação de soma coordenada-a-coordenada. Nossos principais resultados consistem em apresentar condições suficientes que garantam que uma sequência de medidas σ-invariantes, $(\eta_n)_{n\in\mathbb{N}}$, no espaço de Bernoulli de p símbolos (p primo), convirja em convolução para a medida de Bernoulli uniforme (denotada por $\left(\frac{1}{p},\ldots,\frac{1}{p}\right)$), na topologia fraca*. Ou seja, condições que garantam o seguinte: $\eta_n * \ldots * \eta_1 \to \left(\frac{1}{p},\ldots,\frac{1}{p}\right)$. Provamos também que tais condições não são suficientes para nenhum p não primo. Ainda, conseguimos relacionar essa teoria de convergência na topologia fraca* com diagonalização de certas matrizes.

Abstract

This thesis was inspired by the Lindenstrauss' article [10] and the fundamental work of Furstenberg [5]. Let $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ be the compact group which is the cartesian product of infinite copies of the finite group $\mathbb{Z}/p\mathbb{Z}$ and σ be the shift function on $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Our main results consist in presenting enough conditions to guarantee convergence in convolution of a sequence of shift invariant probability measures, $(\eta_n)_{n\in\mathbb{N}}$, in the Bernoulli space of p symbols (p prime) to the uniform Bernoulli measure (denoted by $\left(\frac{1}{p},\ldots,\frac{1}{p}\right)$), in the weak* topology, i.e., conditions that guarantee the following: $\eta_n * \ldots * \eta_1 \to \left(\frac{1}{p},\ldots,\frac{1}{p}\right)$. We also proved that such conditions are not enough if p is not a prime. And even more, we could see that this theory of convergence in the weak* topology and diagonalization of some matrices are related.

Conteúdo

Introdução

Dado um grupo abeliano compacto G, podemos considerar o conjunto das medidas de probabilidade $\mathcal{M}(G)$ e muni-lo de uma operação, chamada de *convolução*, cuja definição pode ser assim apresentada: sendo $S : G \times G \to G$ a operação de soma em G, $S(g,h) = g+h$, e $\eta, \mu \in \mathcal{M}(G)$ duas medidas de probabilidade, denotamos a medida de convolução de η e μ por $\eta * \mu$ e a definimos como sendo:

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

em que A é subconjunto mensurável de G. Observe que $\eta * \mu$ é ainda uma medida de probabilidade de $\mathcal{M}(G)$.

Nas últimas décadas, as pesquisas relativas a essa operação no espaço de medidas variaram muito quanto a generalidade dos resultados obtidos. Por exemplo, Berg em [2], provou que dado qualquer $T : G \to G$ automorfismo contínuo, em que G é grupo abeliano compacto, a função entropia, h, no espaço das medidas de G, assume seu valor máximo (podendo ser infinito) na medida de Haar m de G. Mais ainda, se T for ergódica com respeito a m e $h(m) < \infty$, então m é unicamente caracterizada como a medida que maximiza a entropia. Ainda nesse artigo, o autor obteve resultados relativos à entropia da convolução de medidas T-invariantes e mostrou que, sob certas circustâncias, a convolução de duas medidas T-invariantes e ergódicas é, ainda, ergódica. Já Cohen, em [3], caracterizou, para qualquer grupo abeliano, as medidas μ que satisfazem a equação $\mu * \mu = \mu$, (chamadas de *idempotentes*) mostrando que as únicas medidas idempotentes em G, grupo abeliano, são as medidas de Haar (normalizadas) de subgrupos compactos e as de forma $\nu := \phi d\mu$, $\mu * \eta$, $\mu + \eta - \mu * \eta$ e $\delta - \mu$, em que μ e η são medidas idempotentes

tais que $\mu * \eta = \eta * \mu$, δ é a unidade da álgebra das medidas de G e ϕ um homomorfismo contínuo do grupo G em $\mathbb{C}$. Outra prova desse fato encontra-se também em [7]. Ainda, Rider, em [14], tece alguns comentários à respeito desse teorema de Cohen e afirma que omitindo a hipótese de o grupo ser abeliano, tal teorema já não é verdadeiro.

Por outro lado, sendo mais específicos, Lindenstrauss, Meiri e Peres, em [10], consideraram $G = \mathbb{S}^1$, o círculo unitário, e, grosso modo, provaram que se $(\mu_n)_{n\in\mathbb{N}}$ for uma sequência de medidas ergódicas em relação a transformação σ_p do círculo nele mesmo $\sigma_p(x) = px \mod 1$ cujas entropias $h_{\mu_n}(\sigma_p)$ formam uma sequência de números reais que não decresce muito rápido para zero, então, a sequência $h_{\mu_n * \ldots * \mu_1}(\sigma_p)$ converge para $\log(p)$.

Ainda no contexto abordado por Lindenstrauss, o resultado sobre convolução e crescimento de entropia liga vários tópicos que à primeira vista não estariam necessiariamente relacionados. Como a convolução de duas medidas σ_p-invariantes resulta em uma medida σ_p-invariante, faz sentido calcular a entropia de uma medida resultante da convolução de outras duas. A convergência para $\log(p)$, como Lindenstrauss conclui no Teorema 1.1 do artigo [10], diz algo mais forte do que convergência fraca*: a sequência de medidas $\mu_n * \ldots * \mu_1$ está se aproximando da medida de Lebesgue relativamente a métrica $\bar{d}$, definida, por exemplo, em [6, p. 137] por Glasner, do seguinte modo:

$$\bar{d}(\mu, \eta) = \inf_{J \in \mathcal{J}(\mu,\eta)} \{J(\cup_{i\neq j}[i] \times [j])\},$$

em que μ e η são medidas de probabilidade σ-invariantes no espaço simbólico de Bernoulli de p-símbolos, $[i]$ representa o conjunto aberto das sequências cuja primeira entrada é i e $\mathcal{J}(\mu, \eta)$ é o espaço dos joinings de μ e η. Isso nos mostra que apesar de a operação de convolução depender somente da estrutura de grupo de G para ser definida, ela traz consequências marcantes para o conjunto das medidas σ_p-invariantes no que se refere aos limites de sequências do tipo μ_1, $\mu_2 * \mu_1$, $\mu_3 * \mu_2 * \mu_1$ e etc. e ao modo como ocorre de tal convergência: em $\bar{d}$, não somente na topologia fraca*. Rudolph, em [16, p. 137], afirma que a topologia gerada por $\bar{d}$ é muito mais rica do que a gerada pela topologia fraca*. Por exemplo, o espaço das medidas de probabilidade σ_p-invariantes é fraco* compacto

(ver [18, p.40],[19, p.152] e [16, p. 130]) porém não o é segundo a métrica $\bar{d}$ ([16, p.137]); outro fato é que enquanto o conjunto das medidas ergódicas é denso no espaço das medidas invariantes na topologia fraca* ([18, p.121]), Rudolph, com o Teorema 7.8 de [16], mostrou que oconjunto das medidas ergódicas é $\bar{d}$-fechado.

Por abordar dinâmicas no círculo, interação de medidas σ_p-invariantes com a medida de probabilidade de Lebesgue e outros assuntos relacionados, a pesquisa de Lindenstrauss [10] e a de muitos outros dinamicistas remetem ao trabalho fundamental [5] de Furstenberg e de sua conjectura. Grosso modo, Fursntenberg conjecturou que não há muitas medidas (ergódicas) que são simultaneamente invariantes por σ_p e σ_q, (em que p e q são multiplicativamente independentes, ou seja, $\log(p)/\log(q) \notin \mathbb{Q}$). Apenas a medida de Lebesgue e as medidas atômicas suportadas em órbitas periódicas seriam os exemplares. Isso tudo no contexto do círculo $G = \mathbb{S}^1$. Tal conjectura está ainda em aberto. Até o momento, apenas resultados parciais foram provados. Por exemplo, Lyons em [13] e D. Rudolph em [15] trataram do caso $p = 2$ e $q = 3$ com hipóteses adicionais sob as medidas e o resultado mais expressivo na direção de resolver tal conjectura é devido a D. Rudoplh e A. Jhonson ([8]). Tal resultado garante que a única medida ergódica em relação a ambas as transformações σ_p e σ_q com entropia positiva em relação a σ_p é a medida de Lebesgue no círculo. Segundo [10], a conjectura de Furstenberg no contexto de medidas com entropia nula, aparentemente, vai exigir novos métodos para solução e, em geral, vem inspirando a maior parte das pesquisas relacionadas às dinâmicas de $(\sigma_p, \mathbb{S}^1)$. Para mais detalhes sobre tal conjectura, consultar a introdução do artigo [1], de Assaf Katz.

Nosso trabalho foi inspirado no Teorema 1.1 de Lindenstrauss [10], e, então, indiretamente, remete também ao trabalho de Furstenberg [5]. Ao invés de considerarmos o grupo do círculo $\mathbb{S}^1$, vamos fixar $p \in \mathbb{N}$ e considerar $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, ou seja, o grupo do cartesiano (enumerável) infinito de cópias de $\mathbb{Z}/p\mathbb{Z}^{\mathbb{N}}$ com estrutura de grupo dada pela soma coordenada-a-coordenada. Ao tomarmos a transformação shift (ou deslocamento), $\sigma : G \to G$, podemos considerar as medidas de probabilidade σ-invariantes e a operação de convolução também preserva tais medidas nesse contexto. Assim, faz sentido fazer as mesmas perguntas que Lindenstrauss fez: será que podemos impor alguma condição

razoável a uma sequência de medidas σ-invariantes $(\mu_n)_{n\in\mathbb{N}}$ em $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ garantindo convergência na topologia fraca* da sequência $\mu_n * \ldots * \mu_1$ para a medida de Bernoulli uniforme, $\left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$?

Essa questão foi tratada no capítulo 2, que é a parte central dessa tese. Ali descrevemos com detalhes que o módulo da integral de caracteres não triviais tende a ser pequeno para medidas que possuem entropia positiva e explicitamos essa dependência na Proposição 2.22, utilizada para obtermos o Teorema 2.27, o qual é parecido com Teorema 1.1 de [10]. A diferença é que os grupos considerados são distintos, enquanto Lindenstrauss et al. tratou do círculo $\mathbb{S}^1$, tratamos de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$.

Finalmente, apresentamos um roteiro de leitura para o presente trabalho. O primeiro capítulo foi reservado para as noções de Teoria da Medida, Topologia e Teoria da Informação necessárias para o entendimento dessa tese. As principiapais referências que serviram de base para as preliminares foram [18] e [19]. Inclusive, nessas obras encontram-se as demonstrações dos resultados que eventualmente enunciamos no capítulo primeiro.

No capítulo 2, encontram-se tanto nosso principal resultado (Teorema 2.27) quanto os preparativos para tal, que inicia na seção 2.1, com a caracterização da topologia fraca* no espaço das probabilidades do grupo $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ via utilização de caracteres de tal grupo, com o Teorema 2.8, Proposição 2.13 e Corolário 2.14. Já na seção 2.2, introduzimos a noção de *caracter especial* (Definição 2.17) e trabalhamos com tais caracteres na Proposição 2.18 e Corolário 2.19, o que permitiu-nos provar a Proposição 2.22, que, dentre outras coisas, diz que entropia positiva implica módulo da integral de caracteres não triviais pequeno. Assim, sendo feitos mais uns detalhes técnicos com os lemas 2.24, 2.25 e Corolário 2.26, o Teorema 2.27 segue suavemente. Nesse sentido, exortamos ao leitor afoito, ávido por simplesmente compreender nosso Teorema 2.27, que, tendo conhecimento prévio de Teoria da Medida e Teoria da Informação, poderá iniciar a leitura no segundo capítulo e fixar-se nele, sem perda substancial de entendimento.

Já no terceiro capítulo, abordamos algumas propriedades da convolução de probabilidades em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ em diversos contextos. Cada seção (da 3.1 a 3.3) aborda um contexto específico. Assim, na 3.1, provamos que a convolução por uma medida η é uma contração

fraca tanto na topologia fraca* (Teorema 3.1) quanto na topologia gerada pela métrica $\bar{d}$ (Teorema 3.10) e que a entropia da convolução de duas medidas σ-invariantes tem valor, pelo menos, igual ao da maior das duas prévias entropias (Teorema 3.6). Na 3.2, falamos que as medidas de Bernoulli são um "laboratório" para conjecturar e/ou verificar resultados relacionados a convolução de medidas. Por exemplo, o Corolário 3.11 ilustra esse ponto relacionando topologia fraca*, métrica $\bar{d}$ e convergência em convolução (Definição 2.15) no contexto de sequências formadas por medidas de Bernoulli. Nesse sentido, na seção 3.3, discorremos sobre equação do tipo $\eta * \mu = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$ (nas variáveis η e μ) apresentando soluções em situações específicas que abrangem, principalmente, medidas de Bernoulli (Proposição 3.12, Observação 3.13 e Teorema 3.14).

Prosseguindo, reservamos o quarto capítulo para relacionar o ato de convoluir medidas de probabilidade em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ com o de multiplicar certas matrizes e/ou aplicar tais matrizes em vetores específicos. E mais ainda, provamos que essas matrizes são diagonalizáveis, cujos autovalores são integrais dos caracteres de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Tais fatos são conclusões da Proposição 4.1 e do Teorema 4.4.

Finalmente, no quinto e último capítulo, elencamos uma série de projetos futuros e conjecturas devido aos resultados obtidos nesse trabalho.

Capítulo 1

Preliminares

1.1 Teoria da Medida para $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$

Nesta seção desenvolveremos os conceitos necessários relacionados à Teoria da Medida, de forma geral, e posteriormente, particularizaremos para o conjunto $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, que para nós, será tanto um espaço mensurável quanto um grupo topológico compacto. A referência básica para essa parte é [19].

Dado um conjunto G qualquer, uma σ-álgebra $\mathcal{G}$ de G é um subconjunto do conjunto das partes de G que satisfaz as seguintes propriedades:

1) G pertence a $\mathcal{A}$;

2) Se A pertencer a $\mathcal{G}$ então A^c também pertencerá;

3) Se $\{A_n\}_{n\in\mathbb{N}}$ é uma coleção enumerável de conjuntos que pertencem a $\mathcal{G}$ então $\bigcup_{n\in\mathbb{N}} A_n$ pertence a $\mathcal{G}$.

Chamaremos o par $(G, \mathcal{G})$ de *espaço mensurável* e os conjuntos pertencentes à σ-álgebra $\mathcal{G}$ serão os conjuntos s conjuntos *mensuráveis*, ou seja, os conjuntos "possíveis de serem medidos".

Finalmente, definimos *medida* como sendo uma função $\eta : \mathcal{G} \to \mathbb{R}^+$ satisfazendo:

1) $\eta(\emptyset) = 0$;

2) $\eta(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \eta(A_n)$, sempre que $\{A_n\}_{n\in\mathbb{N}}$ for uma sequência de conjuntos mensuráveis dois-a-dois disjuntos. Essa propriedade é denominada de σ-aditiva.

E por *medida de probabilidade* η (ou simplesmente *probabilidade*) entendemos uma medida com a hipótese adcional de que $\eta(G) = 1$. Denotaremos por $\mathcal{M}(G)$ o espaço das medidas de probabilidades do espaço G.

Agora, sejam $R : G \times G \to G$ uma transformação mensurável e η e μ medidas em G. O próximo resultado fala da existência de outra medida em G, proveniente de η, μ e R.

Proposição 1.1. *Sejam η e μ medidas de um espaço mensurável X e $R : G \times G \to G$ uma transformação mensurável qualquer. Então a igualdade:*

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

para todo mensurável A de G define uma medida em G, que será de probabilidade sempre que η e μ também o forem.

Demonstração. De fato, seja $(A_i)_{i\in\mathbb{N}}$ uma sequência de conjuntos disjuntos e mensuráveis em G. Então, temos:

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

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

Com a noção de espaço mensurável, podemos introduzir os Espaços de Bernoulli. Para cada $i \in \mathbb{N}$, tome $(G_i, \mathcal{G}_i)$ espaço mensurável e considere o produto cartesiano infinito

$G = \prod_{i\in\mathbb{N}} G_i$. Assim, um ponto de G é uma sequência unilateral infinita $\{x_i\}_{i\geq 0}$, com $x_i \in G_i$, para cada i. Definamos agora a chamada *σ-álgebra produto* $\mathcal{G}$ do espaço G. Fixe um natural n e tome mensuráveis $A_i \in \mathcal{G}_i$. Considere, então, o seguinte conjunto:

$$\prod_{i=0}^{n} A_i \times \prod_{i=n+1}^{\infty} G_i := \{(x)_{i=0}^{\infty} \in G | x_i \in A_i \text{ para } i \leq n\}.$$

Tal conjunto será chamado de *cilindro* e a coleção de tais subconjuntos de G formam uma semi-álgebra, digamos, $\mathcal{A}$. Finalmente, a σ-álgebra $\mathcal{G}$ será aquela gerada por $\mathcal{A}$ e $(G, \mathcal{G})$ será, para nós, um *Espaço de Bernoulli*. Para mais detalhes, consultar as primeiras páginas de [19] e o apêndice de [18].

Em nosso trabalho, estamos particularmente interessados no caso mais simples em que $G_i = \{0, 1, \ldots, p-1\}$ para todo $i \in \mathbb{N}$ em que p é um natural previamente fixado. Como trata-se de um conjunto finito, a σ-álgebra será o conjunto das partes para cada i e podemos ver que nesse caso, a σ-álgebra $\mathcal{G}$ de G será aquela gerada pelos conjuntos do tipo:

$$[x_0, x_1, \ldots, x_{m-1}] := \{g \in G | g_i = x_i \text{ para } i \leq m-1\}.$$

E, finalmente, utilizaremos, também, a seguinte notação: $G = \{0, 1, \ldots, p-1\}^{\mathbb{N}}$.

Tendo tratado da noção de mensurabilidade para $G = \{0, 1, \ldots, p-1\}^{\mathbb{N}}$, vamos munir tal espaço de uma métrica. Defina a aplicação $d : G \times G \to [0, 1]$ por:

$$d(x, y) := \begin{cases} \left(\frac{1}{2}\right)^{N_{xy}}, & \text{se } x \neq y \\ 0, & \text{se } x = y \end{cases},$$

em que N_{xy} é o menor natural i tal que $x_i \neq y_i$. Note que d define uma métrica em G tal que

$$B\left(x, \left(\frac{1}{2}\right)^{N+1}\right) = B\left[x, \left(\frac{1}{2}\right)^{N}\right] = [x_0, x_1, \ldots, x_N],$$

em que $B\left(x, \left(\frac{1}{2}\right)^{N+1}\right)$ e $B\left[x, \left(\frac{1}{2}\right)^{N}\right]$ são, respectivamente, bolas aberta e fechada com centro em x e raios, também respectivos, $\left(\frac{1}{2}\right)^{N+1}$ e $\left(\frac{1}{2}\right)^{N}$. Assim, os cilindros, em G, são conjuntos mensuráveis, abertos e fechados, que geram tanto uma topologia quanto uma σ-álgebra para G.

Note ainda que $\{0, 1, \ldots, p-1\}$ é um espaço compacto (por ser finito) e o produto cartesiano infinito de compactos com a topologia produto é um espaço compacto (Teorema de Tychonov). Mas, tal topologia é a mesma gerada pela métrica que acabamos de definir. Como conclusão, obtemos que G é um espaço métrico compacto e a σ-álgebra que definimos é gerada pelos abertos do espaço. Para mais detalhes, consultar o exercício A.1.11 de [18] e para uma demonstração do Teorema de Tychonov, ver [4].

Finalmente, vamos munir nosso espaço G de mais uma estrutura: a de grupo. Para isso, vamos considerar, com certo abuso de notação, que $\mathbb{Z}/p\mathbb{Z} = \{0, 1, \ldots, p-1\}$, ou seja, G será o produto cartesiano infinito unilateral do grupo dos inteiros módulo-p, cuja estrutura de grupo será dada por:

$$x + y = (x_0 + y_0, x_1 + y_1, \ldots) = (x_i + y_i)_{i \in \mathbb{N}}$$

em que, novamente com certo abuso de notação, $x_i + y_i = x_i + y_i(\mod p)$. Observe que com essa nova estrutura, G torna-se um grupo (abeliano) topológico compacto, visto que as operações de inversão $(x \to -x)$ e de soma $((x, y) \to x + y)$ são contínuas. O seguinte teorema elenca todas as propriedades de G que discutimos até então:

Teorema 1.2. *Seja $p \geq 2$ um natural qualquer. Então o Espaço de Bernoulli $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ é um grupo topológico compacto e espaço mensurável, quando munido da σ-álgebra gerada pelos abertos da métrica d.*

Observação 1.3. Quando G for o Espaço de Bernoulli $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ e $R : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \times (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ for a função

$$R(x, y) = x + y,$$

então $\eta *_R \mu$ será chamada medida de *convolução* das medidas η e μ, sendo denotada, simplesmente, por $\eta * \mu$.

1.2 Estendendo medidas

Trataremos nessa seção sobre o processo de extensão de medidas definidas numa álgebra $\mathcal{A}$ para medidas definidas na σ-álgebra $\mathcal{B}$ gerada por $\mathcal{A}$, principalmente no contexto em que o espaço mensurável em questão G for compacto. Com esse intuito, enunciamos os teoremas A.1.14 e A.1.13 de [18] na sequência.

Teorema 1.4. *(Continuidade no vazio) Seja $\mathcal{A}$ uma álgebra de subconjuntos de G e seja $\mu_0 : \mathcal{A} \to [0, +\infty]$ uma função σ-aditiva com $\mu(X) < \infty$. Então μ é σ-aditiva se e somente se*

$$\lim_n \mu(A_n) = 0,$$

para toda sequência $A_1 \subset \cdots \subset A_j \subset \cdots$ de elementos de $\mathcal{A}$ com $\cap_{j=1}^{\infty} A_j = \emptyset$.

Teorema 1.5. *Seja $\mathcal{A}$ uma álgebra de subconjuntos de G e seja $\mu_0 : \mathcal{A} \to [0, +\infty]$ uma função σ-aditiva com $\mu_0(X) < \infty$. Então existe uma única medida μ definida na σ-álgebra $\mathcal{B}$ gerada por $\mathcal{A}$ que é uma extensão de μ_0, ou seja, tal que $\mu(A) = \mu_0(A)$ para todo $A \in \mathcal{A}$.*

Como esses dois resultados podem nos ajudar? Consideremos $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ o espaço de Bernoulli e $\mathcal{A}$ a álgebra das uniões finitas de cilindros. Como cada cilindro é um compacto, todos os elementos álgebra $\mathcal{A}$ são compactos. E toda sequência decrescente $A_1 \supset A_2 \supset \ldots \supset A_n \ldots$ de conjuntos compactos possui interseção $\cap_n A_n$ não vazia. Então, por vacuidade, o Teorema 1.4 juntamente com o Teorema 1.5 garantem-nos que toda função aditiva $\mu_0 : \mathcal{A} \to [0, \infty)$ na álgebra $\mathcal{A}$ dos cilindros que satisfaz $\mu_0(X) = 1$ é σ-aditiva e estende-se unicamente para uma medida de probabilidade μ na σ-álgebra $\mathcal{B}$ gerada por $\mathcal{A}$.

Exemplo 1.6. Seja $v = (v_0, v_1, \ldots, v_{p-1})$ um *vetor de probabilidade*, ou seja, um vetor com entradas não negativas tais que $\sum_{i=0}^{p-1} v_i = 1$. Dizemos que η é uma *medida de Bernoulli* em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ (associada ao vetor v) se nos cilindros, satisfaz:

$$\eta[i_0, i_1, \ldots, i_{m-1}] = v_{i_0} \cdot v_{i_1} \cdot \ldots \cdot v_{i_{m-1}}.$$

Observe que para tal medida η, temos

$$(\eta[0], \eta[1], \ldots, \eta[p-1]) = (v_0, v_1, \ldots, v_{p-1}) = v,$$

e ainda, η é a única medida que, nos cilindros, satisfaz a fórmula acima. Para essa questão da unicidade e extensão, ver os teoremas 1.4 e 1.5.

No caso em que $v = \left(\frac{1}{p}, \frac{1}{p}, \ldots, \frac{1}{p}\right)$ denominaremos a medida η associada a v de *medida de Bernoulli uniforme.*

1.3 A topologia fraca*

Iniciemos essa parte um tanto generalistas. Sejam X um espaço métrico compacto e $\mathcal{X}$ a σ-álgebra gerada pelos abertos. Seja $C_1(X)$ o espaço das funções complexas contínuas de norma unitária. Por [9], $C_1(X)$ possui um conjunto enumerável denso de funções contínuas $\mathcal{F} = \{f_i\}_{i=1}^{\infty}$. Assim, podemos definir, de acordo com [19, p. 148], a seguinte distância em $\mathcal{M}(X)$:

$$d_{\mathcal{F}}(\eta, \mu) = \sum_{i=1}^{\infty} \frac{1}{2^i} \left| \int_X f_i d\eta - \int_X f_i d\mu \right| \tag{1.1}$$

A seguir, expomos um teorema que caracteriza a topologia fraca* tal como precisamos e cuja demonstração encontra-se em [19]:

Teorema 1.7. *Seja X um espaço métrico compacto. Então a Equação (1.1) define uma distância em $\mathcal{M}(X)$, tornando-o um espaço métrico compacto. Mais ainda, dada uma sequência η_n em $\mathcal{M}(X)$, são equivalentes:*

- $d_{\mathcal{F}}(\eta_n, \eta) \to 0$*;*
- $\left|\int_X f_i d\eta_n - \int_X f_i d\eta\right| \to 0$*, para todo i;*
- $\left|\int_X f d\eta_n - \int_X f d\eta\right| \to 0$*, para toda $f \in C_1(X)$.*

A distância $d_{\mathcal{F}}$ definida anteriormente gera a chamada *topologia fraca-** em $\mathcal{M}(X)$ e o teorema acima apresentou três formas equivalentes de se obter a convergência de uma

sequência de medidas de probabilidade η_n para uma medida de probabilidade η nessa topologia. Agora, vamos particularizar. Tome $X = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Assim, o Teorema 1.7 será válido para $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ se pudermos explicitar um conjunto $\mathcal{F}$. Note que dado um cilindro $[x_1, \ldots, x_n]$ temos que a função indicadora $I_{[x_1,\ldots,x_n]}$ é contínua justamente por $[x_1, \ldots, x_n]$ ser simultaneamente aberto e fechado. Mas, combinações lineares de tais funções (com coeficientes racionais) aproximam qualquer função contínua em norma. Portanto, podemos escolher para $\mathcal{F}$ o conjunto das combinações lineares de funções indicadoras de cilindros de norma unitária e coeficientes racionais.

Observe agora algo ainda mais fundamental: se garantirmos que uma sequência η_n em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ satisfizer:

$$\eta_n[i_m, \ldots, i_0] \to \eta[i_m, \ldots, i_0],$$

qualquer que seja o cilindro, então η_n converge para η na topologia fraca*. E assim, acabamos de esboçar a demonstração de um resultado já conhecido e importante:

Teorema 1.8. *Uma sequência η_n em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ converge na topologia fraca* para uma medida de probabilidade $\eta \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ se e somente se:*

$$\eta_n([i_0, i_1, \ldots, i_m]) \to \eta([i_0, i_1, \ldots, i_m]), \tag{1.2}$$

para todo cilindro $[i_0, i_1, \ldots, i_m]$

Demonstração. Se η_n for uma sequência de medidas de probabilidade convergente à medida de probabilidade η na topologia fraca*, então, como a função indicadora $I_{[i_0,i_1,\ldots,i_m]}$ de um cilindro $[i_0, i_1, \ldots, i_m]$ é contínua, segue que:

$$\eta_n([i_0, i_1 \ldots, i_m]) = \int I_{[i_0,i_1\ldots,i_m]} d\eta_n \to \int I_{[i_0,i_1\ldots,i_m]} d\eta = \eta([i_0, i_1 \ldots, i_m]).$$

Por outro lado, supondo que a sequência η_n de medidas de probabilidade satisfaz a equação 1.2, então dada qualquer função f que seja combinação linear finita de funções indicadoras de cilindros, temos $\int f\eta_n \to \int f d\eta$. Mas, o conjunto das combinações lineares de funções indicadoras de cilindros é denso em norma no conjunto das funções contínuas,

o que garante a convergência na topologia fraca*, como queríamos. □

Observação 1.9. O Teorema 6.2 de [19] diz que, para um espaço métrico compacto X e σ-álgebra de Borel, duas medidas de probabilidade η e μ são a mesma medida se e somente se $\int f d\eta = \int f d\mu$ para toda $f : X \to \mathbb{R}$ contínua. Contudo, novamente por [19] (Teorema 6.4), podemos simplesmente trocar o conjunto das funções contínuas por qualquer subconjunto denso (em norma) de funções contínuas (reais ou complexas) de norma unitária. O fato de podermos determinar uma medida sabendo apenas como ela integra um subconjunto denso de funções contínuas foi a chave para podermos definir a métrica $d_{\mathcal{F}}$ acima. E mais ainda, no caso em que $X = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, combinações lineares de funções características de cilindros (ou funções indicadoras de cilindros) são funções contínuas e ainda formam um subconjunto denso. Novamente, por linearidade da integral, as funções características de cilindros já nos bastam para determinarmos a topologia fraca*, o que permitiu-nos escrever o Teorema 1.8 e ainda concluir que duas medidas η e μ em $\mathcal{M}(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ são iguais se e somente se

$$\eta([i_0, i_1, \ldots, i_m]) = \mu([i_0, i_1, \ldots, i_m]),$$

para todo cilindro $[i_0, i_1, \ldots, i_m]$. E finalmente, utilizando a notação

$$d_n(\eta, \mu) = \sum_{P \in \mathcal{P}^n} |\eta(P) - \mu(P)|,$$

outra possível distância em $\mathcal{M}(G)$, para $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, compatível com a topologia fraca* é dada por:

$$d(\eta, \mu) = \sum_{n=1}^{\infty} \left(\frac{d_n(\eta, \mu)}{p^n} \right),$$

em que $\mathcal{P} = \{[0], [1], \ldots, [p-1]\}$, $\mathcal{P}^n = \mathcal{P} \vee \ldots \vee \sigma^{-n+1}(\mathcal{P})$ e a função $\sigma : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ é a função *shift* (ou "deslocamento"), a ser definida na próxima seção.

Note que d define uma distância apenas porque toda medida de probabilidade fica determinada quando se conhece o peso que a mesma dá aos cilindros e a compatibilidade

com a topologia fraca* segue do Teorema 1.8.

1.4 Teoria da Informação

Dado $p \in \mathbb{N}$, definimos a transformação *shift*, (ou "deslocamento", conforme o capítulo 4 de [18]), $\sigma : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, como sendo a aplicação contínua:

$$\sigma(x_1, x_2, x_3, \ldots) = (x_2, x_3, \ldots).$$

Em Teoria da Informação e Teoria Ergódica, interessa-nos as chamadas *medidas σ-invariantes* e a entropia das mesmas. Veremos que a convolução de duas medidas σ-invariantes resulta numa medida σ-invariante e que a entropia da convolução de duas medidas é, pelo menos, o máximo das duas entropias anteriores.

Definição 1.10. Para $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ uma medida η é dita σ-invariante se, para todo mensurável A de G, satisfizer:

$$\eta(\sigma^{-1}(A)) = \eta(A).$$

O Lema 1.3.1 de [18] diz algo melhor para nós: η é σ-invariante se e somente se

$$\eta(\sigma^{-1}([x_1, \ldots, x_n])) = \eta([x_1, \ldots, x_n])$$

para todo cilindro $[x_1, \ldots, x_n]$. A partir dessa constatação e do fato de $\mathcal{M}(G)$ ser um espaço métrico compacto com a topologia fraca*, obtemos o seguinte resultado:

Teorema 1.11. *O conjunto das medidas de probabilidade σ-invariantes, denotado por $\mathcal{M}_\sigma(G)$, é um conjunto fechado (portanto compacto) em $\mathcal{M}(G)$ na topologia fraca*.*

Demonstração. Seja η_n uma sequência de medidas σ-invariantes em $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ tal que $\eta_n \to \eta$ na topologia fraca*, com $\eta \in \mathcal{M}(G)$. Para provarmos o teorema, basta mostrarmos que η é σ-invariante. Tome $[x_1, \ldots, x_n]$ um cilindro qualquer. Então:

$$\begin{aligned}\eta(\sigma^{-1}([x_1,\ldots,x_n])) &= \lim_n \eta_n(\sigma^{-1}([x_1,\ldots,x_n])) \\ &= \lim_n \eta_n([x_1,\ldots,x_n]) \\ &= \eta([x_1,\ldots,x_n]),\end{aligned}$$

e η é de fato uma medida σ-invariante, como queríamos. □

A relação entre medidas invariantes e convolução será dada pelo resultado abaixo.

Proposição 1.12. *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 $\eta *_R \mu$, definida na Proposição 1.1, 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. □

Observação 1.13. Como já vimos, a convolução é a operação associada a função $R(x, y) = x + y$ e tal função satisfaz a hipótese da proposição acima. Assim, a convolução de duas medidas de probabilidade σ-invariantes é ainda uma medida de probabilidade σ-invariante.

Dizemos que uma medida η é *misturadora com respeito à transformação σ* (ou simplesmente *medida misturadora*) se

$$\lim_n \eta(\sigma^{-n}(A) \cap B) \to \eta(A)\eta(B)$$

quaisquer que sejam os mensuráveis A e B. Analogamente, dizemos que uma medida η é *fracamente misturadora com respeito à transformação* σ (ou simplesmente *medida fracamente misturadora*) se

$$\lim_n \frac{\sum_{j=0}^{n-1} \eta(\sigma^{-j}(A) \cap B)}{n} \to \eta(A)\eta(B)$$

para todos os mensuráveis A e B. Toda medida misturadora é fracamente misturadora. Mais ainda, a operação de convolução preserva ambas as classes de medidas. Isso se deve pelo simples fato de a medida produto de duas medidas misturadoras/fracamente misturadoras ser ainda uma medida misturadora/fracamente misturadora. Para mais detalhes, consultar o sétimo capítulo de [18].

Por outro lado, dizemos que η é uma *medida ergódica com respeito a transformação* σ (ou simplesmente *ergódica*) se toda vez que η se exprimir como a combinação convexa de duas medidas invariantes η_1, η_2, digamos $\eta = \alpha\eta_1 + (1-\alpha)\eta_2$ então $\eta = \eta_1$ ou $\eta = \eta_2$. E a convolução não preserva as medidas ergódicas. Por exemplo, escolha $\eta := \frac{1}{2}\delta_{\bar{01}} + \frac{1}{2}\delta_{\bar{10}}$ a única medida ergódica suportada na órbita periódica do ponto $\bar{01} = (0,1,0,1,0,1,0,\ldots)$ em $(\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$. Então, nesse caso, temos $\eta^2 = \eta * \eta = \frac{1}{2}\delta_{\bar{0}} + \frac{1}{2}\delta_{\bar{1}}$, que não é uma medida ergódica, já que se exprime como a combinação convexa de duas medidas ergódicas distintas. Para mais detalhes sobre ergodicidade, consultar a página 119 de [18].

Definiremos, agora, um número associado às medidas σ-invariantes, chamado de *entropia*. Para falar a verdade, é possível definir tal número para contextos mais gerais, como se vê no capítulo 9 de [18] e no capítulo 4 de [19]. Contudo, como o presente trabalho não visa tal generalidade, elencaremos somente o necessário.

Dadas uma medida de probabilidade η em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ e as partições finitas $\mathcal{Q}_1$ e $\mathcal{Q}_2$ de tal espaço, definimos a *entropia da partição* $\mathcal{Q}_1$ como sendo o seguinte valor

$$H_\eta(\mathcal{Q}_1) = -\sum_{Q \in \mathcal{Q}_1} \eta(Q) \log \eta(Q)$$

e, analogamente, a *entropia condicional da partição* $\mathcal{Q}_1$ *sendo dada a partição* $\mathcal{Q}_2$ como

$$H_\eta(\mathcal{Q}_1|\mathcal{Q}_2) = \sum_{Q_1\in\mathcal{Q}_1, Q_2\in\mathcal{Q}_2} \eta(Q_1\cap Q_2)\log\left(\frac{\eta(Q_2)}{\eta(Q_1\cap Q_2)}\right)$$

Agora, seja $\mathcal{P}_0 = \{[0],[1],\ldots,[p-1]\}$ a partição em cilindros de tamanho 1. Para qualquer patição $\mathcal{P}$, usaremos a seguinte notação $\mathcal{P}^n := \mathcal{P}\vee\ldots\vee\sigma^{-n+1}(\mathcal{P})$ para o refinamento das partições $\{\mathcal{P},\sigma^{-1}(\mathcal{P}),\ldots,\sigma^{-n+1}(\mathcal{P})\}$. Observe que, nesse caso, $\mathcal{P}_0^n$ é justamente a partição em cilindros de tamanho n. Feitas essas considerações, obtemos o seguinte teorema:

Teorema 1.14. *Sejam* η *uma medida de probabilidade* σ*-invariante em* $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ *e* $\mathcal{Q}$ *uma partição. Então, as sequências limitadas* $H_\eta(\mathcal{Q}^n)$ *e* $H_\eta(\mathcal{Q}|\vee_{i=1}^n\sigma^{-i}(\mathcal{Q}))$ *são, respectivamente, sub-aditiva e decrescente. Em particular, para* $a_n := H_\eta(\mathcal{P}_0^n)$ *e* $b_n := H_\eta(\mathcal{P}_0|\vee_{i=1}^n\sigma^{-i}(\mathcal{P}_0))$*, existem os limites* $\inf_n\frac{1}{n}a_n$, $\lim_n b_n$ *e eles ainda satisfazem as seguintes igualdades:*

$$\lim_n\frac{1}{n}a_n = \inf_n\frac{1}{n}a_n =: h_\eta(\sigma) = \lim_n b_n = \sup_{\mathcal{Q}}\left(\frac{1}{n}\lim_n H_\eta(\mathcal{Q}^n)\right).$$

Demonstração. Ver quarto capítulo [19] e o nono capítulo [18]. □

Observação 1.15. Dadas uma partição $\mathcal{P}$ e uma medida de probabilidade σ-invariante η, chamamos de *entropia da partição* $\mathcal{P}$ o número $h_\eta(\mathcal{P}) := \lim_n\frac{1}{n}H_\eta(\mathcal{P}^n)$ e o valor $h_\eta(\sigma)$ é chamado de *entropia da medida* η (ou também *entropia métrica*). É possível provar que $h_\eta(\sigma)\leq\log p$ e que a única medida que atinge entropia $\log p$ é a chamada medida de Bernoulli uniforme $\left(\frac{1}{p},\ldots,\frac{1}{p}\right)$. Para tais informações, consultar o quarto capítulo [19] e o nono capítulo de [18].

1.5 Uma nova topologia para $\mathcal{M}\left((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}\right)$

Um questionamento natural que se faz é sobre a relação entre a topologia com a qual o espaço das probabilidades σ-invariantes está munido e a entropia métrica de tais medidas. Por exemplo, é fato conhecido (ver de [19, p.184] e/ou página [18, p.263]) que se

uma sequência η_n de medidas de probabilidade σ-invariantes converge na topologia fraca* para uma medida de probabilidade η, então, temos a seguinte relação:

$$h_\eta(\sigma) \geq \limsup_n h_{\eta_n}(\sigma). \tag{1.3}$$

Assim, é de se esperar, em muitos casos, que a medida limite tenha entropia maior do que as medidas da sequência convergente. Aliás, muito pior, em [18, p.239], o autor apresenta uma sequência de medidas probabilidade invariantes suportadas em órbitas periódicas (portanto todas com entropia nula) que converge para a medida de máxima entropia. Contudo, será que existem topologias que evitam esse possível salto entre a entropia da medida limite e as entropias das medidas da sequência? E ainda, será possível obter alguma relação entre a convolução (que depende apenas da operação de grupo do espaço mensurável para ser definida) e tais novas topologias? A resposta é sim. Nessa seção apresentaremos uma topologia metrizável mais forte do que a topologia fraca*, cuja distância será denotada por $\bar{d}$. Sua principal propriedade para nós é que a função entropia é $\bar{d}$ contínua, ou seja, se η_n for uma sequência de medidas de probabilidade σ-invariantes em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ que converge para outra probabilidade invariante η em $\bar{d}$, então a sequência das entropias $h_{\eta_n}(\sigma)$ converge para $h_\eta(\sigma)$. De forma resumida, se

$$\bar{d}(\eta_n, \eta) \to 0$$

então

$$h_{\eta_n}(\sigma) \to h_\eta(\sigma).$$

Convém destacar que apesar deste já ser um resultado conhecido (ver [6]), a prova detalhada dessa continuidade é um tanto delicada e trabalhosa. Portanto, nesta seção, apresentaremos as primeiras definições concernentes a teoria e em seguida trabalharemos alguns pontos da demonstração encontrada no décimo quinto capítulo de [6] dessa relação de continuidade entre entropia e $\bar{d}$.

Para os preparativos, comecemos com uma noção de *"joinings"* de probabilidades

σ-invariantes.

Definição 1.16. Sejam η e μ probabilidades σ-invariantes. Um *joining* J de η e μ é uma probabilidade em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \times (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ que é $(\sigma \times \sigma)$-invariante e satisfaz as seguintes igualdades para todo mensurável $A \subset (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$:

$$J(A \times (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}) = \eta(A),$$

$$J((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \times A) = \mu(A).$$

Denotamos o espaço dos joinings de η e μ por $\mathcal{J}(\eta, \mu)$.

Exemplo 1.17. Dadas η e μ medidas de probabilidade σ-invariantes em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, temos que $J_\eta(A \times B) := \eta(A \cap B)$ e $J := \eta \times \mu$ definem probabilidades em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \times (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ tais que $J_\eta \in \mathcal{J}(\eta, \eta)$ e $(\eta \times \mu) \in \mathcal{J}(\eta, \mu)$.

No caso em que η e μ são ambas medidas de Bernoulli, existe o seguinte joining $J_{\eta\mu} \in \mathcal{J}(\eta, \mu)$:

$$J_{\eta\mu}([i_1, \ldots, i_n] \times [j_1, \ldots, j_n]) = \prod_{k=1}^{n} \eta[i_k]\mu[j_k],$$

que Glasner, em [6], chama-o de *joining independente*.

Proposição 1.18. *A função* $\bar{d} : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \times (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to [0, \infty)$ *dada por*

$$\bar{d}(\eta, \mu) := \inf_{J \in \mathcal{J}(\eta,\mu)} J\left(\bigcup_{i \neq j} ([i] \times [j])\right),$$

define uma métrica no espaço das medidas de probabilidade invariantes. Mais ainda, tal ínfimo é sempre atingido.

Demonstração. Ver o Lema 7.6 de [16]. □

Tendo nos familiarizado com *joinings*, vamos agora trilhar um caminho que mostrará que a entropia é $\bar{d}$ contínua. Convém salientar que estamos trabalhando com o espaço simbólico de Bernoulli de p símbolos, as medidas são sempre σ-invariantes e $\mathcal{P}_0$ denota a partição em cilindros de tamanho de 1, $\mathcal{P}_0 = \{[0], [1], \ldots, [p-1]\}$.

Proposição 1.19. *As funções*

$$d_{ent}^{\mu}(\mathcal{P},\mathcal{Q}) = H_{\mu}(\mathcal{P}|\mathcal{Q}) + H_{\mu}(\mathcal{Q}|\mathcal{P})$$

e

$$d_{part}^{\mu}(\mathcal{P},\mathcal{Q}) = \frac{1}{2}\sum_{i=1}^{l}\mu(P_i \triangle Q_i) = \frac{1}{2}\sum_{i=1}^{l}\mu(P_i \cap Q_i^c) + \mu(P_i^c \cap Q_i),$$

são métricas no conjunto das partições de l elementos.

Demonstração. Consultar [6], páginas 274 e 275. □

Observação 1.20. Quando a medida μ estiver subentendida, podemos usar $d_{ent}(\mathcal{P},\mathcal{Q})$ e $d_{part}(\mathcal{P},\mathcal{Q})$ para nos referirmos às distâncias definidas na proposição acima. Ainda, na página 274 de [6], há outra notação que convém destacar: $\mu(\mathcal{P}\Delta\mathcal{Q}) := d_{part}^{\mu}(\mathcal{P},\mathcal{Q})$. Por exemplo, dado $J \in \mathcal{J}(\eta,\mu)$ um *joining* de η e μ, temos:

$$\begin{aligned}
J((\mathcal{P}_0 \times G)\triangle(G \times \mathcal{P}_0)) &= d_{part}^{J}((\mathcal{P}_0 \times G),(G \times \mathcal{P}_0)) \\
&= \frac{1}{2}\sum_{i=1}^{l} J((P_i \times G) \cap (G \times \cup_{j\neq i}P_j) \cup (G \times P_i) \cap (\cup_{j\neq i}P_j \times G)) \\
&= \frac{1}{2}\sum_{i=1}^{l} J((P_i \times \cup_{j\neq i}P_j)) + J((\cup_{j\neq i}P_j \times P_i)) \\
&= \frac{1}{2}\cdot 2\sum_{i\neq j} J_{\eta\mu}(P_i \times P_j) \\
&= \sum_{i\neq j} J_{\eta\mu}([i]\times[j]).
\end{aligned}$$

Assim, a $\bar{d}$, por meio dessa notação, fica assim escrita:

$$\bar{d}(\eta,\mu) = \inf_{J\in\mathcal{J}(\eta,\mu)} J((\mathcal{P}_0 \times G)\Delta(G \times \mathcal{P}_0)).$$

Para prosseguirmos, precisaremos do resultado técnico abaixo.

Lema 1.21. *Dado $\varepsilon > 0$, existe $\delta(\varepsilon) > 0$ tal que para todo $x \in [0,1]$ e todo $0 < \delta \leq \delta(\varepsilon)$ tem-se*

$$x\log\left(\frac{x+\delta}{x}\right) < \varepsilon.$$

Demonstração. Dê $\varepsilon > 0$. Primeiramente, tome $0 < \delta_1 < \frac{1}{2}$ que satifaz

$$\max\{-\delta_1 \log \delta_1,\ \log(1-\delta_1),\ -\log(1-\delta_1)\} < \varepsilon.$$

Agora, vai existir $\delta(\varepsilon) > 0$ tal que $\log\left(1 + \frac{\delta(\varepsilon)}{\delta_1}\right) < \varepsilon$. Assim, para $x \in [0, \delta_1)$, obtemos:

$$\begin{aligned} x \log\left(\frac{x+\delta(\varepsilon)}{x}\right) &\leq x \log\left(\frac{1}{x}\right) \\ &\leq -\delta_1 \log \delta_1 \\ &< \varepsilon. \end{aligned}$$

Analogamente, para $x \in [\delta_1, 1]$, conseguimos:

$$\begin{aligned} x \log\left(\frac{x+\delta(\varepsilon)}{x}\right) &\leq \log\left(\frac{x+\delta(\varepsilon)}{x}\right) \\ &= \log\left(1 + \frac{\delta(\varepsilon)}{x}\right) \\ &\leq \log\left(1 + \frac{\delta(\varepsilon)}{\delta_1}\right) \\ &< \varepsilon. \end{aligned}$$

Finalmente, dado $\delta > 0$ satisfazendo $\delta \leq \delta(\varepsilon)$, obtemos, para todo $x \in [0, 1]$:

$$x \log\left(\frac{x+\delta}{x}\right) \leq x \log\left(\frac{x+\delta(\varepsilon)}{x}\right),$$

o que finaliza a demonstração do lema. □

Proposição 1.22. *Fixe $l \geq 2$. Dado $\varepsilon > 0$, existe $\delta = \delta(\varepsilon, l) > 0$ tal que todas as partições α e β (de l elementos) que satisfizerem $d_{part}(\alpha, \beta) < \delta$, também satisfazem $d_{ent}(\alpha, \beta) < \varepsilon$.*

Demonstração. Dado $\varepsilon > 0$, tome $\delta = \frac{\delta(\varepsilon)}{2}$, em que $\delta(\varepsilon)$ é dado na Proposição 1.21. Assim, se $\mathcal{P}$ e $\mathcal{Q}$ forem partições de l elementos tais que $d_{part}(\mathcal{P}, \mathcal{Q}) < \delta$, vamos obter para $i \neq j$:

$$\begin{aligned}\mu(P_i \cap Q_j) &\leq \mu(P_i \cap Q_i^c) \\ &\leq 2d_{part}(\mathcal{P}, \mathcal{Q}) \\ &< \delta(\varepsilon),\end{aligned}$$

e também

$$\begin{aligned}\mu(P_i \cap Q_i) &= \mu(Q_i) - \mu(Q_i \cap P_i^c) \\ &\geq \mu(Q_i) - 2d_{part}(\mathcal{P}, \mathcal{Q}) \\ &> \mu(Q_i) - \delta(\varepsilon).\end{aligned}$$

Então:

$$\begin{aligned}& H_\mu(\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{Q}) \\ = & \sum_{i \neq j} \mu(P_i \cap Q_j) \log\left(\frac{\mu(\mathcal{Q}_j)}{\mu(\mathcal{P}_i \cap \mathcal{Q}_j)}\right) + \sum_i \mu(P_i \cap Q_i) \log\left(\frac{\mu(\mathcal{Q}_i)}{\mu(\mathcal{P}_i \cap \mathcal{Q}_i)}\right) \\ \leq & \sum_{i \neq j} \mu(P_i \cap Q_j) \log\left(\frac{1}{\mu(\mathcal{P}_i \cap \mathcal{Q}_j)}\right) + \sum_i \mu(P_i \cap Q_i) \log\left(\frac{\mu(P_i \cap Q_i) + \delta(\varepsilon)}{\mu(\mathcal{P}_i \cap \mathcal{Q}_i)}\right) \\ < & (l^2 - l)\varepsilon + l\varepsilon \\ = & l^2\varepsilon.\end{aligned}$$

Para finalizarmos a proposição, basta observar que

$$\begin{aligned}d_{ent}(\mathcal{P}, \mathcal{Q}) &= 2H_\mu(\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{P}) - H_\mu(\mathcal{Q}) \\ &\leq |H_\mu(\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{P})| + |H_\mu(\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{Q})| \\ &< 2l^2\varepsilon\end{aligned}$$

□

A Proposição 1.22 acima mostrou que partições próximas segundo d_{part} tendem a ser próximas segundo d_{ent}. Assim, como a Proposição 1.23 vai nos mostrar que a função que leva cada partição a sua entropia é d_{ent} contínua, tal função será, também, d_{part} contínua, conforme consta no Teorema 1.24.

Proposição 1.23. $|h_\mu(\mathcal{P}) - h_\mu(\mathcal{Q})| \leq d_{ent}(\mathcal{P}, \mathcal{Q})$.

Demonstração. Primeiramente, temos a seguinte constatação, para quaisquer partições $\mathcal{P}$ e $\mathcal{Q}$ finitas (de mesma quantidade de elementos):

$$\begin{aligned} |H_\mu(\mathcal{P}) - H_\mu(\mathcal{Q})| &= |H_\mu(\mathcal{P}) - H_\mu(\mathcal{P} \vee \mathcal{Q}) + H_\mu(\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{Q})| \\ &\leq |H_\mu(\mathcal{P}) - H_\mu(\mathcal{P} \vee \mathcal{Q})| + |H_\mu(\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{Q})| \\ &= H_\mu(\mathcal{P}|\mathcal{Q}) + H_\mu(\mathcal{Q}|\mathcal{P}) = d_{ent}(\mathcal{P}, \mathcal{Q}). \end{aligned}$$

Prosseguindo, a ideia da demonstração é a seguinte: mostraremos que

$$|H_\mu(\mathcal{P}^n) - H_\mu(\mathcal{Q}^n)| \leq d_{ent}(\mathcal{P}^n, \mathcal{Q}^n) \leq n d_{ent}(\mathcal{P}, \mathcal{Q}),$$

para todo n natural. Após, multiplicando por $\frac{1}{n}$ e tomando o limite em n, obtemos a desigualde desejada.

Note que, por definição de d_{ent}, basta verificarmos que $H_\mu(\mathcal{P}^n|\mathcal{Q}^n) \leq n H_\mu(\mathcal{P}|\mathcal{Q})$, para todo n natural. Façamos por indução.

Para $n = 1$ é trivial. Para $n = 2$, temos

$$\begin{aligned} H_\mu(\mathcal{P} \vee \sigma^{-1}(\mathcal{P})|\mathcal{Q} \vee \sigma^{-1}(\mathcal{Q})) &= H_\mu(\mathcal{P}|\mathcal{Q} \vee \sigma^{-1}(\mathcal{Q})) + H_\mu(\sigma^{-1}(\mathcal{P})|\mathcal{P} \vee \mathcal{Q} \vee \sigma^{-1}(\mathcal{Q})) \\ &\leq H_\mu(\mathcal{P}|\mathcal{Q}) + H_\mu(\mathcal{P}|\mathcal{Q}) = 2H_\mu(\mathcal{P}|\mathcal{Q}) \end{aligned}$$

Agora suponha o problema resolvido para algum $n \in \mathbb{N}$. Então, obtemos

$$\begin{aligned} H_\mu(\mathcal{P}^{n+1}|\mathcal{Q}^{n+1}) &= H_\mu(\mathcal{P}^n \vee \sigma^{-n}(\mathcal{P})|\mathcal{Q}^n \vee \sigma^{-n}(\mathcal{Q})) \\ &\leq H_\mu(\mathcal{P}^n|\mathcal{Q}^n \vee \sigma^{-n}(\mathcal{Q})) + H_\mu(\sigma^{-n}(\mathcal{P})|\mathcal{P}^n \vee \mathcal{Q}^n \vee \sigma^{-n}(\mathcal{Q})) \\ &\leq H_\mu(\mathcal{P}^n|\mathcal{Q}^n) + H_\mu(\sigma^{-n}(\mathcal{P})|\sigma^{-n}(\mathcal{Q})) \\ &= H_\mu(\mathcal{P}^n|\mathcal{Q}^n) + H_\mu(\mathcal{P}|\mathcal{Q}) \\ &\leq (n+1)H_\mu(\mathcal{P}|\mathcal{Q}). \end{aligned}$$

Pelo princípio da indução, nosso problema está resolvido. □

Teorema 1.24. *A entropia $\mathcal{P} \mapsto h_\mu(\mathcal{P})$ é uma função d_{part} contínua: para todo $\varepsilon > 0$ existe $\delta > 0$ tal que*

$$d_{part}(\mathcal{P}, \mathcal{Q}) < \delta \Rightarrow |h_\mu(\mathcal{P}) - h_\mu(\mathcal{Q})| \leq \varepsilon.$$

Demonstração. Dê $\varepsilon > 0$ e $\mathcal{P}$ uma partição com l elementos. Então, pela Proposição 1.22, existe $\delta > 0$ tal que

$$d_{part}(\mathcal{P}, \mathcal{Q}) < \delta \Rightarrow d_{ent}(\mathcal{P}, \mathcal{Q}) < \varepsilon,$$

assim, o teorema segue pela Proposição 1.23. □

Agora estamos prontos, finalmente, para mostrar que a função que leva cada medida σ-invariante a sua entropia é contínua relativamente a métrica $\bar{d}$.

Teorema 1.25. *A entropia $\mu \mapsto h_\mu(\sigma)$ é $\bar{d}$ contínua.*

Demonstração. Lembre que $\bar{d}(\eta, \mu) = J((\mathcal{P}_0 \times G)\Delta(G \times \mathcal{P}_0))$ (para algum *joining* J) também pode ser entendida assim:

$$\bar{d}(\eta, \mu) = d^J_{part}(\mathcal{P}_0 \times G, G \times \mathcal{P}_0).$$

Agora, pela Proposição 1.22, existe $\delta > 0$ tal que

$$d^J_{part}(\mathcal{P}_0 \times G, G \times \mathcal{P}_0) < \delta \Rightarrow d^J_{ent}(\mathcal{P}_0 \times G, G \times \mathcal{P}_0) < \varepsilon,$$

também, pela Proposição 1.23, sabemos que

$$|h_J(\mathcal{P}_0 \times G) - h_J(G \times \mathcal{P}_0)| \leq d^J_{ent}(\mathcal{P}_0 \times G, G \times \mathcal{P}_0). \tag{1.4}$$

Analisemos mais de perto. Note que para quaisquer transformações mensuráveis T_1, T_2, temos:

$$\bigvee_{i=1}^{n} (T_1 \times T_2)^{-i}(\mathcal{P}_0 \times G) = \bigvee_{i=1}^{n} (T_1)^{-i}(\mathcal{P}_0) \times G,$$

e vale também

$$H_J(\mathcal{P}_0 \times G) = \sum_P -J(P \times G)\log(J(P \times G)) = \sum_P -\eta(P)\log(\eta(P)) = H_\eta(\mathcal{P}_0).$$

Cálculos todos análogos para μ. Agora, a inequação 1.4 pode ser reescrita assim:

$$|h_\eta(\mathcal{P}_0) - h_\mu(\mathcal{P}_0)| \leq d^J_{ent}(\mathcal{P}_0 \times G, G \times \mathcal{P}_0) < \varepsilon.$$

Para finalizar, lembre que $h_\mu(\sigma) = h_\mu(\mathcal{P}_0)$ para toda medida invariante μ. Isso garante que $\mu \mapsto h_\mu(\sigma)$ é $\bar{d}$ contínua, como queríamos. □

Capítulo 2

Convoluir para convergir

Conforme já trabalhamos no capítulo anterior, dizemos que uma sequência de medidas de probabilidade $\{\eta_n\}_{n\in\mathbb{N}}$ em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ *converge na topologia fraca** para a medida de probabilidade η (e denotamos $\eta_n \to \eta$) se $\int f\eta_n \to \int f d\eta$ para toda função contínua $f : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}$. Agora, atentemos para uma definição fundamental.

Definição 2.1. Sejam $C\left((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}\right)$ o espaço das funções contínuas $f : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}$ e D um subcontjunto de $C\left((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}\right)$. Dizemos que D *determina* a topologia fraca* em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ se a condição $\int g d\eta_n \to \int g d\eta$ para toda função $g \in D$ garantir que η_n converge na tologia fraca* para η, quaisquer que sejam a sequência η_n e o limite η.

À luz dessa definição, o Teorema 1.8 garante que o conjunto das combinações lineares de funções indicadoras de cilindros *determina* a topologia fraca*. Nesse capítulo, falaremos sobre outro conjunto de funções contínuas que também a determina: o grupo dos caracteres de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Nesse embalo, apresentaremos os resultados mais importantes deste trabalho, o Teorema 2.27, que é similar ao Teorema 1.1 de [10] quanto às sequências de medidas de probabilidade que convergem na topologia fraca* para a medida de Haar do grupo em questão ($(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ no nosso caso e o círculo $\mathbb{S}^1$ no caso de Lindenstrauss) e o Corolário 2.29, que discute as discrepâncias entre o caso p primo e não primo quanto a convergência de tais sequências.

Inicialmente, a Proposição 2.2 adiante diz que o peso que a convolução de duas medidas dá aos cilindros depende apenas do peso que as medidas previamente davam a eles.

Proposição 2.2. *Sejam $G = (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ e η e μ medidas de probabilidade em $\mathcal{M}(G)$. Então:*

$$\eta * \mu[a_1, \ldots, a_n] = \sum_{i_1,\ldots,i_n=0}^{p-1} \eta[i_1, \ldots, i_n]\mu[a_1 - i_1, \ldots, a_n - i_n]$$

Demonstração. A proposição segue das seguintes igualdades:

$$\begin{aligned}
\eta * \mu[a_1, \ldots, a_n] &= \eta \times \mu(S^{-1}([a_1, \ldots, a_n])) \\
&= \eta \times \mu(\cup_{i_1,\ldots,i_n=0}^{p-1}[i_1, \ldots, i_n] \times [a_1 - i_1, \ldots, a_n - i_n]) \\
&= \sum_{i_1,\ldots,i_n=0}^{p-1} \eta \times \mu([i_1, \ldots, i_n] \times [a_1 - i_1, \ldots, a_n - i_n]) \\
&= \sum_{i_1,\ldots,i_n=0}^{p-1} \eta[i_1, \ldots, i_n]\mu[a_1 - i_1, \ldots, a_n - i_n]
\end{aligned}$$

□

2.1 O grupo dos caracteres de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$

Definição 2.3. Seja G um grupo. Um *caracter* de G é um homomorfismo de G no grupo multiplicativo dos números complexos. Ou seja, é uma função $\chi : G \to \mathbb{C}^*$ que satisfaz $\chi(gh) = \chi(g)\chi(h)$ para todos $g, h \in G$. O espaço dos caracteres de G será denotado por $\widehat{G}$. No caso em que G for um grupo topológico, exigir-se-á que χ seja homomorfismo contínuo.

Dados dois caracteres $\chi, \chi' \in \widehat{G}$, o produto pontual dessas funções define outro caracter $\chi\chi' : G \to \mathbb{C}^*$, $\chi\chi'(g) := \chi(g)\chi'(g)$. Assim, $\widehat{G}$ é um grupo (abeliano) quando munido do produto pontual, cuja identidade é o *caracter trivial* $1_G \equiv 1$.

Restrinjamos um pouco para tratarmos do grupo finito $(\mathbb{Z}/p\mathbb{Z})^m$. Para conforto do leitor, apresentamos na sequência a Proposição 2.4 e o Teorema 2.5, que caracterizam o grupo $\widehat{(\mathbb{Z}/p\mathbb{Z})^m}$. E, após, enunciaremos o Teorema 2.7, que induz uma relação de ortogonalidade entre caracteres, útil em contas vindouras. Uma referência básica para caracteres de grupos finitos (principalmente abelianos) é [12].

Proposição 2.4. *Se* $(l_{m-1}, l_{m-2}, \ldots, l_0) \in (\mathbb{Z}/p\mathbb{Z})^m$ *então a função* $\chi_l : (\mathbb{Z}/p\mathbb{Z})^m \to \mathbb{C}$ *dada por*

$$\chi_l(j_{m-1}, \ldots, j_0) = \prod_{k=0}^{m-1} e^{\frac{2\pi i(l_k j_k)}{p}}$$

é um caracter de $(\mathbb{Z}/p\mathbb{Z})^m$ *e ainda mais: todos os caracteres de* $(\mathbb{Z}/p\mathbb{Z})^m$ *são dessa forma. Na igualdade acima,* l *é dado por* $l - 1 = \sum_{k=0}^{m-1} l_k p^k$.

Demonstração. Iniciemos verificando que χ_l define um caracter de fato:

$$\begin{aligned}
& \chi_l((j_{m-1}, \ldots, j_0) + (r_{m-1}, \ldots, r_0)) \\
= \; & \chi_l(j_{m-1} + r_{m-1}, \ldots, j_0 + r_0) \\
= \; & \prod_{k=0}^{m-1} e^{\frac{2\pi i(l_k(j_k + r_k))}{p}} \\
= \; & \prod_{k=0}^{m-1} e^{\frac{2\pi i(l_k j_k)}{p}} \cdot e^{\frac{2\pi i(l_k r_k)}{p}} \\
= \; & \prod_{k=0}^{m-1} e^{\frac{2\pi i(l_k j_k)}{p}} \prod_{k=0}^{m-1} e^{\frac{2\pi i(l_k r_k)}{p}} \\
= \; & \chi_l(j_{m-1}, \ldots, j_0) \cdot \chi_l(r_{m-1}, \ldots, r_0).
\end{aligned}$$

Para finalizarmos, note que dado χ caracter de $(\mathbb{Z}/p\mathbb{Z})^m$, se tivermos $\chi(e_k) = e^{\frac{2\pi i l_k}{p}}$ então $\chi = \chi_l$ para $e_k = (\underbrace{0, 0, \ldots, 0}_{k-1}, 1, 0, \ldots, 0)$. □

Na verdade, a Proposição 2.4 acima dá margens para escrevermos o seguinte isomorfismo.

Teorema 2.5. *A aplicação* $\phi : (\mathbb{Z}/p\mathbb{Z})^m \to \widehat{(\mathbb{Z}/p\mathbb{Z})^m}$ *definida por*

$$\phi((i_{m-1}, \ldots, i_0)) = \chi_i,$$

é um isomorfismo de grupos, com $i - 1 = \sum_{k=0}^{m} i_k p^k$.

Demonstração. Que ϕ é uma aplicação sobrejetora, isso foi feito na Proposição 2.4. Faltanos ver que ϕ é um homomorfismo de grupos injetivo. Note que $\phi((i_{m-1}, \ldots, i_0)) = \phi((j_{m-1}, \ldots, j_0))$ se e somente se $e^{\frac{2\pi i(i_k - j_k)}{p}} = 1$ para todo $k \in \{0, 1, \ldots, m-1\}$. E isso é o

mesmo que dizer que $i_k - j_k$ é múltiplo de p para todo $k \in \{0, 1, \ldots, m-1\}$. Mas, como $0 \leq i_k, j_k \leq p-1$ então $|i_k - j_k| < p$ e a única opção que nos resta é termos $i_k - j_k = 0$, para todo $k \in \{0, 1, \ldots, m-1\}$, e ϕ é, então, injetora. Finalmente, para $l^1 = (l^1_{m-1}, \ldots, l^1_0)$ e $l^2 = (l^2_{m-1}, \ldots, l^2_0)$ temos:

$$\begin{aligned}
\phi(l^1 + l^2)(j_{m-1}, \ldots, j_0) &= \chi_{l^1+l^2}(j_{m-1}, \ldots, j_0) \\
&= \prod_{k=0}^{m-1} e^{\frac{2\pi i (l^1_k + l^2_k) j_k}{p}} \\
&= \prod_{k=0}^{m-1} e^{\frac{2\pi i l^1_k j_k}{p}} \prod_{k=0}^{m-1} e^{\frac{2\pi i l^2_k j_k}{p}} \\
&= \chi_{l^1}(j_{m-1}, \ldots, j_0) \cdot \chi_{l^2}(j_{m-1}, \ldots, j_0) \\
&= (\phi(l^1) \cdot \phi(l^2))(j_{m-1}, \ldots, j_0),
\end{aligned}$$

o que garante que $\phi(l^1 + l^2) = \phi(l^1) \cdot \phi(l^2)$ e mostramos que ϕ é homomorfismo de grupos, finalizando a demonstração. □

Observação 2.6. O Teorema 2.5 e a Proposição 2.4 acima levam-nos a concluir uma fórmula ainda mais interessante e sugestiva para um caracter χ_i qualquer (com $i - 1 = \sum_{k=0}^{m} i_k p^k$):

$$\chi_i(j_{m-1}, \ldots, j_0) = \prod_{k=0}^{m-1} \chi_{i_k}(j_k).$$

A igualdade acima diz-nos que todo caracter de $(\mathbb{Z}/p\mathbb{Z})^m$ se escreve como uma espécie de produtório de m caracteres de $(\mathbb{Z}/p\mathbb{Z})$.

Finalmente, apresentamos uma relação de ortognalidade para carcteres de grupos abelianos finitos.

Teorema 2.7. *Sejam G um grupo abeliano finito $\chi \in \widehat{G}$. Então*

$$\sum_{g \in G} \chi(g) = \begin{cases} \#G, & \text{se } \chi \text{ for o carcter trivial } \chi = 1_G; \\ 0, & \text{caso contrário,} \end{cases}$$

Demonstração. Ver o Teorema 3.2.1 de [12]. □

O principal motivo de termos abordado caracteres de grupos finitos abelianos é que tais funções relacionam-se com o grupo $\widehat{(\mathbb{Z}/p\mathbb{Z})}^{\mathbb{N}}$, que será caracterizado com o Teorema 2.8 na sequência, cujo enunciado e demonstração elaboramos. Os resultados básicos sobre grupos abelianos finitos, que necessitamos no decorrer da demonstração, encontram-se em [12].

Teorema 2.8. *Dado $p \in \mathbb{N}$, uma função contínua $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}$ é um caracter não trivial se e somente se satisfaz as seguintes propriedades:*

1. *$Im(\chi) = \left\{e^{\frac{2\pi ikj}{p}} : k = 0, 1, \ldots, p-1\right\}$ para algum $j \neq p$ que divide p;*
2. *$\chi(x+y) = \chi(x) \cdot \chi(y)$, para todos $x, y \in (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$;*
3. *Existe um natural m tal que $\chi^{-1}\left(\left\{e^{\frac{2\pi ikj}{p}}\right\}\right)$ é a união de $\frac{p^m}{\#Im(\chi)}$ cilindros de tamanho m, para todo $k = 0, 1, \ldots, p-1$.*

Demonstração. Por definição, toda função contínua que satisfaz a propriedade 2 acima é, de fato, um caracter. Vejamos agora a outra parte.

Seja χ um caracter. Como para todo $x \in (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ temos

$$\underbrace{x + x + \ldots + x}_{p} = (0, 0, \ldots, 0, \ldots) := \mathbf{0},$$

e assim,

$$1 = \chi(\mathbf{0}) = \chi(\underbrace{x + x + \ldots + x}_{p}) = \chi(x)^p.$$

E, finalmente

$$Im(\chi) \subset \left\{e^{\frac{2\pi ik}{p}} : k = 0, 1, \ldots, p-1\right\}.$$

Agora, pelo Teorema 1.1.2 de [12], como $Im(\chi)$ é um subgrupo do grupo finito cíclico das raízes p-ésimas da unidade, segue que $Im(\chi)$ também será um grupo finito cíclico. Dessa forma, vai existir $g = e^{\frac{2\pi ij}{p}}$, com $j \in \{0, 1, \ldots, p-1\}$, que gera o subgrupo $Im(\chi)$, ou seja, tal que $Im(\chi) = \{g^n : n \in 0, 1, \ldots, p-1\}$. Sem perda de generalidade, podemos assumir que $j = \min\left\{k \geq 0 : e^{\frac{2\pi ik}{p}} \text{ gera } Im(\chi)\right\}$. Nesse caso, j necessariamente divide p e

obtemos a primeira propriedade do teorema. Já a segunda propriedade segue diretamente do fato de χ ser caracter.

Partamos para a terceira propriedade. Como χ é função contínua e $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ é compacto, $\chi^{-1}\left(\left\{e^{\frac{2\pi ikj}{p}}\right\}\right)$ também será compacto para todo $k = 0, 1, \ldots, p-1$. Novamente, pela continuidade de χ e por $Im(\chi)$ ser um grupo discreto, para cada $x \in \chi^{-1}\left(\left\{e^{\frac{2\pi ikj}{p}}\right\}\right)$ existe um cilindro $[x_1, \ldots, x_{m_x}]$ tal que $x \in [x_1, \ldots, x_{m_x}] \subset \chi^{-1}\left(\left\{e^{\frac{2\pi ikj}{p}}\right\}\right)$. Portanto, obtemos a seguinte cobertura aberta:

$$\chi^{-1}\left(\left\{e^{\frac{2\pi ikj}{p}}\right\}\right) = \bigcup_{\left\{x:\ \chi(x)=e^{\frac{2\pi ikj}{p}}\right\}} [x_1, \ldots, x_{m_x}].$$

Por compacidade, para cada $k = 0, 1, \ldots, p-1$, podemos extrair uma subcobertura finita, obtendo, assim, um conjunto $A_k \subset (\mathbb{Z}/p\mathbb{Z})^{m_k}$ tal que: $\chi^{-1}\left(\left\{e^{\frac{2\pi ikj}{p}}\right\}\right) = \bigcup_{(x_1^k,\ldots,x_{m_k}^k)\in A_k} [x_1^k, \ldots, x_{m_k}^k]$. No entanto, podemos definir $m = \max\{m_0, \ldots, m_{p-1}\}$, encontrar outros subconjuntos $A'_k \subset (\mathbb{Z}/p\mathbb{Z})^m$ e assim reescrever, para todo $k \in \{0, 1, \ldots, p-1\}$:

$$\chi^{-1}\left(\left\{e^{\frac{2\pi ikj}{p}}\right\}\right) = \bigcup_{(x_1^k,\ldots,x_m^k)\in A'_k} [x_1^k, \ldots, x_m^k].$$

Observe que acabamos de escrever os conjuntos $\left\{\chi^{-1}\left(\left\{e^{\frac{2\pi ikj}{p}}\right\}\right) : k = 0, 1 \ldots, p-1\right\}$ como uniões disjuntas de cilindros de mesmo tamanho m, o que nos aproxima de findar a terceira propriedade. Agora, só falta mostrar que a quantidade de cilindros independe de k, ou seja, que $\#A_{k_1} = \#A_{k_2}$ quaisquer que sejam $k_1, k_2 \in \{0, 1, \ldots, p-1\}$. Mas, dado qualquer cilindro $[x_1^k, \ldots, x_m^k]$ conforme notação acima, temos a seguinte igualdade:

$$\chi^{-1}\left(\left\{e^{\frac{2\pi ikj}{p}}\right\}\right) = [x_1^k, \ldots, x_m^k] + \chi^{-1}(\{1\}) \tag{2.1}$$

Agora, tomando dois cilindros C_0 e D_0 de tamanho m que estão em $\chi^{-1}(\{1\})$, temos:

$$C_0 + [x_1^k, \ldots, x_m^k] = D_0 + [x_1^k, \ldots, x_m^k] \Leftrightarrow C_0 = D_0.$$

Como somas de cilindros de tamanho m produzem cilindros de tamanho m, segue que o lado direito da igualdade (2.1) exprime-se como uma união finita de cilindros e com a mesma quantidade de cilindros que $\chi^{-1}(\{1\})$ possui. Como k era qualquer, obtemos a terceira propriedade. □

Observação 2.9. O Teorema 2.8 afirma que $\chi^{-1}(\{1\})$ é um subgrupo de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ e todo caracter de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ é uma combinação linear de funções indicadoras de cilindros de mesmo tamanho cujos coeficientes são raízes p-ésimas da unidade.

Observação 2.10. Quando p for primo e χ um caracter não trivial, o Teorema 2.8 diz-nos que $Im(\chi) = \left\{e^{\frac{2\pi ij}{p}} : j = 0, 1, \ldots, p-1\right\}$ e que vai existir um natural m tal que $\chi^{-1}\left(\left\{e^{\frac{2\pi ij}{p}}\right\}\right)$ é a união de p^{m-1} cilindros de tamanho m, para todo $j = 0, 1, \ldots, p-1$.

Convém destacarmos, como conclusão do Teorema 2.8 e Observação 2.9, que o grupo dos caracteres de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ é fundamentalmente uma "união" dos caracteres de $(\mathbb{Z}/p\mathbb{Z})^m$ para todo m natural, haja vista que para cada caracter de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ existe m tal que χ é constante nos cilindros de tamanho m. É por essa razão que, apesar de $\widehat{(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}}$ não ser um grupo abelino finito, também obteremos uma relação de ortogonalidade entre caracteres com ajuda do Teorema 2.7, dada no Lema 2.12 adiante. Antes, contudo, vamos associar a cada caracter um vetor complexo, com a definição seguinte.

Definição 2.11. Seja χ um caracter de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ constante em cilindros de tamanho m. Definimos o *vetor* $v_\chi \in \mathbb{C}^{p^m}$ *associado ao caracter* χ *e a* $m \in \mathbb{N}$ por

$$(v_\chi)_i = \chi[i_{m-1}, i_{m-2}, \ldots, i_0], \tag{2.2}$$

em que $i - 1 = \sum_{k=0}^{m} i_k p^k$.

Lema 2.12. *Sejam* $\chi_1, \chi_2 : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}^*$ *caracteres distintos constantes em cilindros de tamanho* m_1 *e* m*, respectivamente, com* $m_1 \leq m$*. Então* χ_1 *é também constante em cilindros de tamanho* m *e* $\langle v_{\chi_1}, v_{\chi_2}\rangle = 0$*, em que* v_{χ_1} *e* v_{χ_2} *são os respectivos vetores de* $\mathbb{C}^{p^m}$ *associados aos caracteres* χ_1 *e* χ_2*, dados pela Equação 2.2. Em particular, o conjunto*

$$V_m = \{v_\chi : \chi \text{ é caracter de } (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \text{ constante em cilindros de tamanho } m\}$$

é, também, uma base para $\mathbb{C}^{p^m}$, como queríamos.

Demonstração. Dados χ_1 e χ_2 conforme hipótese, como cada cilindro de tamanho m_1 expressa-se como a união finita disjunta de cilindros de tamanho m, segue que χ_1 é também constante em cilindros de tamanho m.

Considere, agora, o seguinte caracter $\overline{\chi_2}(g) := \overline{\chi_2(g)}$. Como $\chi_1 \neq \chi_2$, temos que $\chi_1\bar{\chi_2}$ é caracter não trivial, ou seja, $\chi_1\overline{\chi_2} \neq 1_{(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}}$. Agora, interpretando $\chi_1\overline{\chi_2}$ como um caracter de $(\mathbb{Z}/p\mathbb{Z})^m$, obtemos, pelo Teorema 2.7, que

$$\langle v_{\chi_1}, v_{\chi_2}\rangle = \sum_{i_0,\ldots,i_{m-1}=0}^{p-1} \chi\overline{\chi'}[i_{m-1},\ldots,i_0] = 0,$$

o que garante V_m é um conjunto formado por exatamente $\#\widehat{(\mathbb{Z}/p\mathbb{Z})^m}$ vetores ortogonais entre si. Mas, o Corolário 3.1.2 de [12] diz que todo grupo abeliano finito é isomorfo ao seu grupo de caracteres. Em particular, eles possuem a mesma cardinalidade. Assim, V_m é, de fato, uma base para $\mathbb{C}^{p^m}$ formada por vetores ortogonais entre si, como queríamos. □

Finalmente, com a Proposição 2.13 e Corolário 2.14, vamos caracterizar a topologia fraca* de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ via utilização de caracteres.

Proposição 2.13. *Sejam η e μ medidas de probabilidade em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Então $\eta = \mu$ se e somente se $\int \chi d\eta = \int \chi d\mu$ para todo caracter $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}$.*

Demonstração. É claro que se duas medidas coincidem, elas devem dar a mesma integral para todas as funções integráveis. Assuma, então, que η e μ satisfaçam

$$\int \chi d\eta = \int \chi d\mu \tag{2.3}$$

para todo χ caracter. Seja $\{\chi_1, \ldots, \chi_{p^m}\}$ uma enumeração para os caracteres constantes em cilindros de tamanho m. A condição (2.3) acima implica a seguinte igualdade matricial:

$$\begin{pmatrix} \chi_1[0,\ldots,0] & \ldots & \chi_1[p-1,\ldots,p-1] \\ \chi_2[0,\ldots,0] & \ldots & \chi_2[p-1,\ldots,p-1] \\ \vdots & \ddots & \vdots \\ \chi_{p^m}[0,\ldots,0] & \ldots & \chi_{p^m}[p-1,\ldots,p-1] \end{pmatrix} \begin{pmatrix} \eta[0,\ldots,0] \\ \vdots \\ \vdots \\ \eta[p-1,\ldots,p-1] \end{pmatrix} =$$

$$= \begin{pmatrix} \int \chi_1 d\mu \\ \int \chi_2 d\mu \\ \vdots \\ \int \chi_{p^m} d\mu \end{pmatrix}.$$

Conforme o Lema 2.12, a matriz quadrada do lado esquerdo possui linhas que são ortogonais entre si e portanto ela é uma matriz invertível. Como a seguinte igualdade também é válida:

$$\begin{pmatrix} \chi_1[0,\ldots,0] & \ldots & \chi_1[p-1,\ldots,p-1] \\ \chi_2[0,\ldots,0] & \ldots & \chi_2[p-1,\ldots,p-1] \\ \vdots & \ddots & \vdots \\ \chi_{p^m}[0,\ldots,0] & \ldots & \chi_{p^m}[p-1,\ldots,p-1] \end{pmatrix} \begin{pmatrix} \mu[0,\ldots,0] \\ \vdots \\ \vdots \\ \mu[p-1,\ldots,p-1] \end{pmatrix} =$$

$$= \begin{pmatrix} \int \chi_1 d\mu \\ \int \chi_2 d\mu \\ \vdots \\ \int \chi_{p^m} d\mu \end{pmatrix},$$

concluímos, assim, que η e μ coincidem nos cilindros de tamanho m. Sendo m geral, usando os teoremas 1.4 e 1.5, segue que $\eta = \mu$, como queríamos demonstrar. □

Corolário 2.14. *Uma sequência de medidas de probabilidade μ_n em $\mathcal{M}\left(\mathbb{Z}/p\mathbb{Z}\right)^{\mathbb{N}}\big)$ converge pra uma medida μ na topologia fraca* se e somente se*

$$\int \chi d\mu_n \to \int \chi d\mu, \tag{2.4}$$

para todo caracter $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}^*$.

Demonstração. Suponha que $\mu_n \to \mu$ na topologia fraca* e seja χ um caracter. Como χ é função contínua, a convergência (2.4) é satisfeita.

Agora, suponha que a condição (2.4) ocorra para todo caracter χ. Pela Proposição 2.13 e usando a mesma notação que usamos nela, se $\{\chi_1, \ldots, \chi_{p^m}\}$ for uma enumeração

para os caracteres de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ constantes em cilindros de tamanho m, obtemos que a sequência de vetores

$$w_n := \begin{pmatrix} \chi_1[0,\ldots,0] & \ldots & \chi_1[p-1,\ldots,p-1] \\ \chi_2[0,\ldots,0] & \ldots & \chi_2[p-1,\ldots,p-1] \\ \vdots & \ddots & \vdots \\ \chi_{p^m}[0,\ldots,0] & \ldots & \chi_{p^m}[p-1,\ldots,p-1] \end{pmatrix} \begin{pmatrix} \mu_n[0,\ldots,0] \\ \vdots \\ \vdots \\ \mu_n[p-1,\ldots,p-1] \end{pmatrix}$$

converge para o vetor

$$w := \begin{pmatrix} \chi_1[0,\ldots,0] & \ldots & \chi_1[p-1,\ldots,p-1] \\ \chi_2[0,\ldots,0] & \ldots & \chi_2[p-1,\ldots,p-1] \\ \vdots & \ddots & \vdots \\ \chi_{p^m}[0,\ldots,0] & \ldots & \chi_{p^m}[p-1,\ldots,p-1] \end{pmatrix} \begin{pmatrix} \mu[0,\ldots,0] \\ \vdots \\ \vdots \\ \mu[p-1,\ldots,p-1] \end{pmatrix},$$

e isso garante que $\mu_n[i_1,\ldots,i_m] \to \mu[i_1,\ldots,i_m]$ para todo cilindro $[i_1,\ldots,i_m]$ de tamanho m. Logo, μ_n converge para μ na topologia fraca*, como queríamos. □

2.2 Entropia e convergência em convolução

Iniciemos com um conceito de convergência que permeia toda essa seção.

Definição 2.15. Seja η_n uma sequência de probabilidades em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Dizemos que η_n *converge em convolução* para a probabilidade η se

$$\eta_n * \ldots * \eta_2 * \eta_1 \to \eta,$$

na topologia fraca*.

O objetivo dessa seção é apresentar uma condição suficiente sobre as entropias de uma sequência de medidas de probabilidade σ-invariantes de modo a garantir a convergência em convolução para a medida de Bernoulli uniforme, no caso do espaço de Bernoulli com

p símbolos, p primo. Faremos isso demonstrando nosso principal resultado, o Teorema 2.27, que como já dissemos na introdução desse trabalho, guarda algumas similaridades com o Teorema 1.1 de [10]. Para a demonstração, precisamos entender as relações entre invariância, caracteres e entropia de medidas invariantes. E também, mais adiante, explicaremos essa distinção entre primos e não-primos para o trato da convergência em convolução.

Em prol de nosso intento, elaboramos os três resultados que seguem (Proposição 2.16, Proposição 2.18 e Corolário 2.19) que permitir-nos-ão não só detalhar propriedades de caracteres não triviais mas, também, relacionar a entropia de medidas de probabilidade invariantes com o módulo da integral desses caracteres não triviais. Grosso modo, entropia positiva implica módulo da integral dos caracteres não muito grande. E a convergência em convolução de uma dada sequência de medidas para a Bernoulli uniforme dependerá dessa relação, conforme veremos adiante, com a Proposição 2.22 e Teorema 2.27.

Proposição 2.16. *Sejam p primo e χ um caracter de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ constante nos cilindros de tamanho 1. Então χ será o caracter trivial se e somente existir $j \neq 0$ tal que $\chi[j] = 1$.*

Demonstração. Se χ for o caracter trivial, então $\chi[j] = 1$ para todo $j \in \mathbb{Z}/p\mathbb{Z}$. Por outro lado, suponha que exista $j \in \mathbb{Z}/p\mathbb{Z}$ não nulo, tal que $\chi[j] = 1$. Seja k um elemento de $\mathbb{Z}/p\mathbb{Z}$. Como p é primo, vai existir $n \in \{0, 1, ..., p-1\}$ tal que $k = nj$. Assim, $\chi[k] = \chi[nj] = (\chi[j])^n = 1$ e χ é o caracter trivial, como queríamos. □

Definição 2.17. Sejam p primo e $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}^*$ um caracter constante em cilindros de tamanho $m \geq 2$. Fixe $j \in \{0, 1, \ldots, p-1\}$. Chamamos de *propriedade especial* a propriedade seguinte:

Para toda $(m-1)$-upla $(i_2, \ldots, i_m) \in (\mathbb{Z}/p\mathbb{Z})^{m-1}$, existe único $i_1 \in \mathbb{Z}/p\mathbb{Z}$ tal que $[i_1, i_2, \ldots, i_m] \subset \chi^{-1}(e^{\frac{2\pi ij}{p}})$.

Dizemos também que χ é *caracter especial* se $\chi^{-1}(e^{\frac{2\pi ij}{p}})$ satisfaz a *propriedade especial* para todo $j \in \{0, 1, \ldots, p-1\}$.

Proposição 2.18. *Sejam p um número primo e $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}^*$ um caracter constante nos cilindros de tamanho $m \geq 2$.*

1. *Assim,* $[1,\underbrace{0,\ldots,0}_{m-1}] \nsubseteq \chi^{-1}(1)$ *se e somente se* χ *for um caracter especial (Definição 2.17).*

2. *Agora, se* $[1,\underbrace{0,\ldots,0}_{m-1}] \subseteq \chi^{-1}(1)$*, então vai existir um caracter* $\chi_\sigma : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}^*$ *constante nos cilindros de tamanho* $m-1$ *e satisfazendo a seguinte igualdade para todo* $j \in \{0,1,\ldots,p-1\}$*:*

$$\chi^{-1}\left(e^{\frac{2\pi ij}{p}}\right) = \sigma^{-1}\left(\chi_\sigma^{-1}(e^{\frac{2\pi ij}{p}})\right). \tag{2.5}$$

Em particular,

$$\int \chi d\eta = \int \chi_\sigma d\eta$$

para toda medida de probabilidade σ*-invariante* η*.*

Demonstração. Tome χ caracter constante nos cilindros de tamanho m. Assuma, primeiramente, que $[1,\underbrace{0,\ldots,0}_{m-1}] \nsubseteq \chi^{-1}(1)$. Agora, fixe $j \in \{0,1,\ldots,p-1\}$ e suponha:

$$[i_1,i_2,\ldots,i_m],[i_1',i_2,\ldots,i_m] \subset \chi^{-1}\left(e^{\frac{2\pi ij}{p}}\right).$$

Portanto

$$[i_1-i_1',0,0,\ldots,0] = ([i_1,i_2,\ldots,i_m]-[i_1',i_2,\ldots,i_m]) \subset \chi^{-1}(1),$$

e isso garante-nos que $i_1 = i_1'$, já que caso contrário, por $\chi^{-1}(1)$ ser um grupo, concluiríamos que

$$[1,0,\ldots,0] \subseteq \chi^{-1}(1),$$

e nós estamos supondo justamente o oposto disso. Isso nos diz que a quantidade de cilindros de tamanho m em $\chi^{-1}(e^{\frac{2\pi ij}{p}})$ é, no máximo, p^{m-1}, pois, em outras palavras, acabamos de mostrar que para cada $(m-1)$-upla $(i_2,\ldots,i_m) \subset (\mathbb{Z}/p\mathbb{Z})^{m-1}$ existe *no máximo um* $i_1 \in \mathbb{Z}/p\mathbb{Z}$ tal que $[i_1,i_2,\ldots,i_m] \subset \chi^{-1}(e^{\frac{2\pi ij}{p}})$. E p^{m-1} é justamente a quantidade de $(m-1)$-uplas distintas $(i_2,\ldots,i_m)$ que podemos escrever com $i_2,\ldots,i_m \in \{0,1,\ldots,p-1\}$.

Porém, sabemos de antemão, pelo Teorema 2.8 e Observação 2.10, que $\chi^{-1}(e^{\frac{2\pi ij}{p}})$ é formado por *exatamente* p^{m-1} cilindros de tamanho m, garantindo que χ é um *caracter especial*.

Por outro lado, supondo que o caracter χ seja *especial*, como $[0,\underbrace{0,\ldots,0}_{m-1}] \subseteq \chi^{-1}(1)$, segue que $[1,\underbrace{0,\ldots,0}_{m-1}] \nsubseteq \chi^{-1}(1)$. Com isso, findamos a primeira parte do teorema.

Agora, suponha que $[1,\underbrace{0,\ldots,0}_{m-1}] \subseteq \chi^{-1}(1)$ e tome um cilindro $[i_1,i_2,\ldots,i_m]$ em $\chi^{-1}(e^{\frac{2\pi ij}{p}})$, j qualquer, porém fixo. Assim, se $i_1' \in \mathbb{Z}/p\mathbb{Z}$, obtemos:

$$\begin{aligned}
\chi([i_1',i_2,\ldots,i_m]) &= \chi([i_1'-i_1,0,\ldots,0]+[i_1,i_2,\ldots,i_m]) \\
&= \chi([i_1'-i_1,0,\ldots,0])\cdot\chi([i_1,i_2,\ldots,i_m]) \\
&= \chi([1,\underbrace{0,\ldots,0}_{m-1}])^{i_1'-i_1}\cdot\chi([i_1,i_2,\ldots,i_m]) \\
&= \chi([i_1,i_2,\ldots,i_m]).
\end{aligned}$$

Logo, cada vez que encontrarmos um cilindro $[i_1,i_2,\ldots,i_m]$ em $\chi^{-1}(e^{\frac{2\pi ij}{p}})$, poderemos garantir que $\bigcup_{j_1=0}^{p-1}[j_1,i_2,\ldots,i_m]$ estará também em $\chi^{-1}(e^{\frac{2\pi ij}{p}})$. Mas, lembre que

$$\bigcup_{j_1=0}^{p-1}[j_1,i_2,\ldots,i_m] = \sigma^{-1}[i_2,\ldots,i_m].$$

Agora, defina a função χ_σ por $\chi_\sigma[i_2,\ldots,i_m] = \chi[0,i_2,\ldots,i_m]$, para toda $(m-1)$-upla $(i_2,\ldots,i_m) \in (\mathbb{Z}/p\mathbb{Z})^{m-1}$. Então

$$\begin{aligned}
\chi_\sigma([i_2,\ldots,i_m]+[i_2',\ldots,i_m']) &= \chi_\sigma([i_2+i_2',\ldots,i_m+i_m']) \\
&= \chi[0,i_2+i_2',\ldots,i_m+i_m'] \\
&= \chi([0,i_2,\ldots,i_m]+[0,i_2',\ldots,i_m']) \\
&= \chi([0,i_2,\ldots,i_m])\cdot\chi([0,i_2',\ldots,i_m']) \\
&= \chi_\sigma([i_2,\ldots,i_m])\cdot\chi_\sigma([i_2',\ldots,i_m']),
\end{aligned}$$

e χ_σ trata-se de fato de um caracter constante nos cilindros de tamanho $m-1$. Para finalizar a parte 2. do teorema, temos as seguintes igualdades:

$$\begin{aligned}
\sigma^{-1}\left(\chi_\sigma^{-1}(e^{\frac{2\pi ij}{p}})\right) &= \sigma^{-1}\left(\bigcup_{(i_2,\ldots,i_m):\chi_\sigma[i_2,\ldots,i_m]=e^{\frac{2\pi ij}{p}}} [i_2,\ldots,i_m]\right) \\
&= \bigcup_{(i_2,\ldots,i_m):\chi_\sigma[i_2,\ldots,i_m]=e^{\frac{2\pi ij}{p}}} \sigma^{-1}([i_2,\ldots,i_m]) \\
&= \bigcup_{(i_2,\ldots,i_m):\chi[0,i_2,\ldots,i_m]=e^{\frac{2\pi ij}{p}}} \sigma^{-1}([i_2,\ldots,i_m]) \\
&= \chi^{-1}(e^{\frac{2\pi ij}{p}}),
\end{aligned}$$

o que termina a demonstração. □

Corolário 2.19. *Sejam p primo e $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}^*$ um caracter não trivial constante em cilindros de tamanho $m \geq 2$. Então ocorrerá somente uma das seguintes opções:*

1. *χ é caracter especial (Definição 2.17) ou;*
2. *vai existir um caracter especial χ_σ^E constante em cilindros de tamanho $1 < k \leq m$ tal que $\chi^{-1}(e^{\frac{2\pi ij}{p}}) = \sigma^{-1}((\chi_\sigma^E)^{-1}(e^{\frac{2\pi ij}{p}}))$ ou;*
3. *vai existir um caracter não trivial constante em cilindros de tamanho 1, χ_σ^1, que também satisfaz $\chi^{-1}(e^{\frac{2\pi ij}{p}}) = \sigma^{-1}((\chi_\sigma^1)^{-1}(e^{\frac{2\pi ij}{p}}))$.*

Demonstração. Sejam p e $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}^*$ conforme hipótese. Admita que χ não seja especial. Assim, de acordo com a Proposição 2.18, o cilindro $[1, \underbrace{0,\ldots,0}_{m-1}]$ está contido em $\chi^{-1}(1)$ e podemos construir um caracter χ_σ constante em cilindros de tamanho $m-1$, tal que satisfaz, para todo $j \in \mathbb{Z}/p\mathbb{Z}$:

$$\chi^{-1}(e^{\frac{2\pi ij}{p}}) = \sigma^{-1}(\chi_\sigma^{-1}(e^{\frac{2\pi ij}{p}})).$$

Se χ_σ for especial ou constante em cilindros de tamanho 1, o processo termina. Caso não se enquadre em nenhuma dessas categorias, utilizamos novamente a Proposição 2.18 para

construir um caracter χ_{σ^2} que satisfaz para todo $j \in \mathbb{Z}/p\mathbb{Z}$:

$$\chi_{\sigma}^{-1}(e^{\frac{2\pi ij}{p}}) = \sigma^{-1}(\chi_{\sigma^2}^{-1}(e^{\frac{2\pi ij}{p}})).$$

Novamente, se χ_{σ^2} for especial ou constante em cilindros de tamanho 1, terminamos. Caso contrário, o processo continua. E terminará em, no máximo, $m-1$ passos, com a realização de uma (e somente uma) das três opções elencadas no corolário. □

Observação 2.20. Pela demonstração do Corolário 2.19, dado um caracter não trivial χ constante em cilindros de tamanho $m \geq 2$, existe um processo que vai criando, passo a passo, novos caracteres, até que culmina com a produção ou de um caracter especial ou de um caracter não trivial constante em cilindros de tamanho 1. Sejam χ' um caracter produzido em alguma das etapas do processo iniciado com χ e η uma medida de probabilidade σ-invariante. Da Equação 2.5, concluímos que $\int \chi d\eta = \int \chi' d\eta$. Mais geralmente, se η_n for uma sequência de medidas de probabilidade σ-invariantes, então $\int \chi d\eta_n \to \int \chi d\eta$ se e somente existir um caracter χ' produzido com o Corolário 2.19 (tendo como ponto de partida o próprio χ) que satisfaz $\int \chi' d\eta_n \to \int \chi' d\eta$. Assim, quando falamos que "sem perda de generalidade" podemos tomar ou um caracter especial ou um caracter não trivial constante em cilindros de tamanho 1, estamos nos referindo ao que foi exposto nessa observação.

Observação 2.21. Antes de enunciarmos e provarmos Proposição 2.22, é conveniente uma prévia explicação a respeito das igualdades e desigualdades que faremos uso. Usaremos a mesma notação da referida proposição nessa observação.

Inicialmente, seja $(\varepsilon, \delta := \delta(\varepsilon))$ o par obtido na Proposição 1.21 e suponha a existência de uma medida de probabilidade σ-invariante η e de um *caracter especial* χ (Definição 2.17), constante nos cilindros de tamanho $m \geq 2$, tal que para algum $j \in \{0, 1, \ldots, p-1\}$, satisfaçam $\eta\left(\chi^{-1}\left(e^{\frac{2\pi ij}{p}}\right)\right) \geq 1 - \delta$.

O primeiro fato a ser percebido é que se $[i_1, \ldots, i_m]$ é cilindro não pertencente a $\chi^{-1}(e^{\frac{2\pi ij}{p}})$, então $\eta[i_1, \ldots, i_m] \leq \delta = \delta(\varepsilon)$. Dessa afirmação, da Proposição 1.21 e da Observação 2.10, segue a desigualdade:

$$\sum_{[i_1,\ldots,i_m]\not\subseteq\chi^{-1}(e^{\frac{2\pi ij}{p}})} \eta[i_1,\ldots,i_m]\log\left(\frac{1}{\eta[i_1,\ldots,i_m]}\right) \;<\; (p^m-p^{m-1})\varepsilon.$$

Agora, voltemo-nos para os cilindros $[i'_1,\ldots,i'_m]$ que estão em $\chi^{-1}(e^{\frac{2\pi ij}{p}})$. Esses cilindros são, em geral, os mais "pesados" de acordo com a medida η, visto que eles totalizam uma medida de pelo menos $1-\delta$. Se um cilindro $[i'_1,\ldots,i'_m]$ estiver contido em $\chi^{-1}(e^{\frac{2\pi ij}{p}})$, então, por χ ser especial, o primeiro item da Proposição 2.18 fala-nos que a união dos cilindros $\left(\bigcup_{j\neq i'_1}[j,i'_2,\ldots,i'_m]\right)$ estará toda contida em $\bigcup_{k\neq j}\left(\chi^{-1}\left(e^{\frac{2\pi ik}{p}}\right)\right)$, um conjunto medindo no máximo δ, de acordo com a probabilidade η. Todas essas informações, juntamente com a σ-invariância de η, fazem-nos obter o seguinte:

$$\begin{aligned}\eta[i'_2,\ldots,i'_m] &= \eta(\sigma^{-1}([i'_2,\ldots,i'_m]))\\ &= \eta[i'_1,\ldots,i'_m]+\eta\left(\bigcup_{j\neq i'_1}[j,i'_2,\ldots,i'_m]\right)\\ &\leq \eta[i'_1,\ldots,i'_m]+\delta,\end{aligned}$$

e, finalmente, concluímos as desigualdades seguintes:

$$\begin{aligned}&\sum_{[i'_1,\ldots,i'_m]\subset\chi^{-1}(e^{\frac{2\pi ij}{p}})}\eta[i'_1,\ldots,i'_m]\log\left(\frac{\eta[i'_2,\ldots,i'_m]}{\eta[i'_1,\ldots,i'_m]}\right)\leq\\ \leq &\sum_{[i'_1,\ldots,i'_m]\subset\chi^{-1}(e^{\frac{2\pi ij}{p}})}\eta[i'_1,\ldots,i'_m]\log\left(\frac{\eta[i'_1,\ldots,i'_m]+\delta(\varepsilon)}{\eta[i'_1,\ldots,i'_m]}\right)\\ \leq\; & p^{m-1}\varepsilon.\end{aligned}$$

Feitas essas observações, podemos passar para a proposição que segue.

Proposição 2.22. *Sejam p um número primo, $\varepsilon_0>0$ um número real e $\chi:(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}\to\mathbb{C}^*$ um caracter não trivial. Então, existe $\delta=\delta(\varepsilon_0,\chi)>0$ tal que, para todo $j\in\mathbb{Z}/p\mathbb{Z}$,*

toda medida de probabilidade σ-invariante η munida da propriedade $h_\eta(\sigma) \geq \varepsilon_0$ satisfaz:

$$\eta\left(\chi^{-1}(e^{\frac{2\pi ij}{p}})\right) < 1 - \delta.$$

Demonstração. Sejam p um número primo, $\varepsilon_0 > 0$ e χ um caracter não trivial. Assuma que χ seja constante nos cilindros de tamanho m.

Comecemos por selecionar $0 < \varepsilon < \frac{\varepsilon_0}{p^m}$ e escolher $\delta(\varepsilon_0, \chi) = \delta(\varepsilon)$ da Proposição 1.21. Suponha que exista alguma medida de probabilidade σ-invariante η que satisfaça, para algum $j \in \{0, 1, \ldots, p-1\}$

$$\eta\left(\chi^{-1}(e^{\frac{2\pi ij}{p}})\right) \geq 1 - \delta. \tag{2.6}$$

Mostraremos que $h_\eta(\sigma) < \varepsilon_0$.

Para ganharmos intuição, suponha primeiramente que $m = 1$, ou seja, que χ seja caracter não trivial constante nos cilindros $\{[0], [1], \ldots, [p-1]\}$. Nesse caso, pela Proposição 2.16, existe único $l_j \in \{0, 1, \ldots, p-1\}$ tal que $\chi^{-1}(e^{\frac{2\pi ij}{p}}) = [l_j]$. Assim, pela equação (2.6) e por propriedades do δ descritas na Proposição 1.21, conseguimos:

$$\begin{aligned} h_\eta(\sigma) &\leq -\sum_{l \in (\mathbb{Z}/p\mathbb{Z})} \eta[l] \log \eta[l] \\ &= -\eta[l_j] \log \eta[l_j] - \sum_{l \neq l_j} \eta[l] \log \eta[l] \\ &\leq \varepsilon + (p-1)\varepsilon \\ &< \varepsilon_0, \end{aligned}$$

e terminamos a demonstração para $m = 1$.

A partir de agora, suporemos $m \geq 2$. Pela Observação 2.20, podemos assumir, sem perda de generalidade, que ou χ é não trivial e constante em cilindros de tamanho 1 ou trata-se de um *caracter especial* (Definição 2.17). O caso do tamanho 1 foi recém feito acima. Suponha, então, que χ seja especial. Então, pelo Teorema 1.14 e Observação 2.21,

obtemos:

$$\begin{aligned}
h_\eta(\sigma) &\leq H_\eta\left(\mathcal{P}|\bigvee_{j=1}^{m-1}\sigma^{-j}(\mathcal{P})\right)\\
&= \sum_{[i_1,\ldots,i_m]\not\subseteq\chi^{-1}(e^{\frac{2\pi ij}{p}})}\eta[i_1,\ldots,i_m]\log\left(\frac{\eta[i_2,\ldots,i_m]}{\eta[i_1,\ldots,i_m]}\right)+\\
&\quad+\sum_{[i'_1,\ldots,i'_m]\subset\chi^{-1}(e^{\frac{2\pi ij}{p}})}\eta[i'_1,\ldots,i'_m]\log\left(\frac{\eta[i'_2,\ldots,i'_m]}{\eta[i'_1,\ldots,i'_m]}\right)\\
&\leq \sum_{[i_1,\ldots,i_m]\not\subseteq\chi^{-1}(e^{\frac{2\pi ij}{p}})}\eta[i_1,\ldots,i_m]\log\left(\frac{1}{\eta[i_1,\ldots,i_m]}\right)+\\
&\quad+\sum_{[i'_1,\ldots,i'_m]\subset\chi^{-1}(e^{\frac{2\pi ij}{p}})}\eta[i'_1,\ldots,i'_m]\log\left(\frac{\eta[i'_1,\ldots,i'_m]+\delta(\varepsilon)}{\eta[i'_1,\ldots,i'_m]}\right)\\
&\leq (p^m-p^{m-1})\varepsilon+p^{m-1}\varepsilon\\
&< \varepsilon_0,
\end{aligned}$$

o que demonstra a proposição. □

Estamos quase prontos para expor uma condição que, imposta a uma sequência de medidas σ-invariantes, garante que tal sequência convirja em convolução (Definição 2.15) para a medida de Bernoulli uniforme. Precisamos, ainda, de algumas estimativas para o módulo de combinações convexas de raízes p-ésimas da unidade, motivo pelo qual produzimos os lemas 2.24, 2.25 e o Corolário 2.26 na sequência.

Definição 2.23. Um *vetor estocástico* é um vetor $v=(v_0,\ldots,v_{m-1})$ com entradas reais não negativas tais que $\sum_i v_i=1$.

Lema 2.24. *Sejam $\alpha_0,\alpha_1,\ldots,\alpha_{p-1}$ números reais não negativos e $\theta_0,\theta_1,\ldots,\theta_{p-1}\in[0,2\pi]$ ângulos quaisquer. Então*

$$\left(\sum_{j=0}^{p-1}\alpha_j\cos(\theta_j)\right)^2+\left(\sum_{j=0}^{p-1}\alpha_j\sin(\theta_j)\right)^2\leq\left(\sum_{j=0}^{p-1}\alpha_j\right)^2.$$

Demonstração. Sejam $\{\alpha_j\}_{j=0}^{p-1}$ e $\{\theta_j\}_{j=0}^{p-1}$ conforme hipótese. Então:

$$
\begin{aligned}
& \left(\sum_{j=0}^{p-1} \alpha_j \cos(\theta_j)\right)^2 + \left(\sum_{j=0}^{p-1} \alpha_j \sin(\theta_j)\right)^2 \\
= & \sum_{j=0}^{p-1} \alpha_j^2 \cos^2(\theta_j) + 2\left(\sum_{\substack{k<j \\ 1\leq k\leq p-2 \\ 2\leq j\leq p-1}} \alpha_k\alpha_j \cos(\theta_k)\cos(\theta_j)\right) + \\
& + \sum_{j=0}^{p-1} \alpha_j^2 \sin^2(\theta_j) + 2\left(\sum_{\substack{k<j \\ 1\leq k\leq p-2 \\ 2\leq j\leq p-1}} \alpha_k\alpha_j \sin(\theta_k)\sin(\theta_j)\right) \\
= & \sum_{j=0}^{p-1} \alpha_j^2(\cos^2(\theta_j) + \sin^2(\theta_j)) + 2\left(\sum_{\substack{k<j \\ 1\leq k\leq p-2 \\ 2\leq j\leq p-1}} \alpha_k\alpha_j(\cos(\theta_k)\cos(\theta_j) + \sin(\theta_k)\sin(\theta_j))\right) \\
= & \sum_{j=0}^{p-1} \alpha_j^2 + 2\left(\sum_{\substack{k<j \\ 1\leq k\leq p-2 \\ 2\leq j\leq p-1}} \alpha_k\alpha_j(\cos(\theta_k - \theta_j))\right) \\
\leq & \sum_{j=0}^{p-1} \alpha_j^2 + 2\left(\sum_{\substack{k<j \\ 1\leq k\leq p-2 \\ 2\leq j\leq p-1}} \alpha_k\alpha_j\right) \\
= & \left(\sum_{j=0}^{p-1} \alpha_j\right)^2
\end{aligned}
$$

□

Lema 2.25. *Todo vetor p-dimensional com entradas não negativas* $(\alpha_0, \alpha_1, \ldots, \alpha_{p-1})$ *(p não necessariamente primo) satisfaz:*

$$\left|\sum_{j=0}^{p-1} \alpha_j e^{\frac{2\pi ij}{p}}\right| \leq \left|\alpha_{j_0} + \left(\sum_{j\neq j_0} \alpha_j\right) e^{\frac{2\pi i}{p}}\right|,$$

qualquer que seja $j_0 \in \{0, 1, \ldots, p-1\}$.

Em particular, se o vetor $(\alpha_0, \alpha_1, \ldots, \alpha_{p-1})$ *for estocástico, obtemos:*

$$\left| \sum_{j=0}^{p-1} \alpha_j e^{\frac{2\pi i j}{p}} \right| \leq \left| \alpha_{j_0} + (1 - \alpha_{j_0})\, e^{\frac{2\pi i}{p}} \right|,$$

Demonstração. Fixe $j_0 \in \{0, 1, \ldots, p-1\}$. Usando o Lema 2.24 para $\theta_j = \frac{2\pi(j+p-j_0)}{p}$ e lembrando que $\left| e^{\frac{2\pi i(p-j_0)}{p}} \right| = 1$, obtemos:

$$\begin{aligned}
\left| \sum_{j=0}^{p-1} \alpha_j e^{\frac{2\pi i j}{p}} \right|^2 &= \left| \sum_{j=0}^{p-1} \alpha_j e^{\frac{2\pi i j}{p}} \right|^2 \left| e^{\frac{2\pi i(p-j_0)}{p}} \right|^2 \\
&= \left| \sum_{j=0}^{p-1} \alpha_j e^{\frac{2\pi i(j+p-j_0)}{p}} \right|^2 \\
&= \left| \alpha_{j_0} + \sum_{j \neq j_0} \alpha_j e^{\frac{2\pi i(j+p-j_0)}{p}} \right|^2 \\
&= \left| \alpha_{j_0} + \sum_{j \neq j_0} \alpha_j \left(\cos\left(\frac{2\pi(j+p-j_0)}{p} \right) + i \sin\left(\frac{2\pi(j+p-j_0)}{p} \right) \right) \right|^2 \\
&= \alpha_{j_0}^2 + 2\alpha_{j_0} \left[\sum_{j \neq j_0} \alpha_j \cos\left(\frac{2\pi(j+p-j_0)}{p} \right) \right] + \\
&\quad + \left(\sum_{j \neq j_0} \alpha_j \cos\left(\frac{2\pi(j+p-j_0)}{p} \right) \right)^2 + \left(\sum_{j \neq j_0} \alpha_j \sin\left(\frac{2\pi(j+p-j_0)}{p} \right) \right)^2 \\
&\leq \alpha_{j_0}^2 + 2\alpha_{j_0} \left(\sum_{j \neq j_0} \alpha_j \right) \cos\left(\frac{2\pi}{p} \right) + \left(\sum_{j \neq j_0} \alpha_j \right)^2 \\
&= \left| \alpha_{j_0} + \left(\sum_{j \neq j_0} \alpha_j \right) e^{\frac{2\pi i}{p}} \right|^2,
\end{aligned}$$

como queríamos demonstrar. □

Corolário 2.26. *Fixe* $p \in \mathbb{N}$. *Dados* $0 < \delta < \frac{1}{p}$ *e* $(\alpha_0, \alpha_1, \ldots, \alpha_{p-1})$ *um vetor estocástico*

tal que $0 < \alpha_j \leq 1-\delta$ *para todo* $j \in \{0,1,\ldots,p-1\}$*, então:*

$$\left|\sum_{j=0}^{p-1} \alpha_j e^{\frac{2\pi ij}{p}}\right| \leq \left|\delta + (1-\delta)e^{\frac{2\pi i}{p}}\right|.$$

Demonstração. Dado $(\alpha_0, \alpha_1, \ldots, \alpha_{p-1})$ um vetor estocástico de acordo com as hipóteses, vai existir $j_0 \in \{0,1,\ldots,p-1\}$ tal que $\alpha_{j_0} \in \left[\frac{1}{p}, 1-\delta\right]$. Como a função

$$f : \left[\frac{1}{p}, 1-\delta\right] \to [0,1]$$

dada por $f(x) = \left|x + (1-x)e^{\frac{2\pi i}{p}}\right|$ possui como máximo o valor $f(1-\delta)$ e $f(x) = f(1-x)$ para todo x e $1-x$ no domínio, segue, pelo Lema 2.25, que

$$\begin{aligned}\left|\sum_{j=0}^{p-1} \alpha_j e^{\frac{2\pi ij}{p}}\right| &\leq \left|\alpha_{j_0} + (1-\alpha_{j_0})e^{\frac{2\pi i}{p}}\right| \\ &= f(\alpha_{j_0}) \\ &\leq f(1-\delta) \\ &\leq \left|\delta + (1-\delta)e^{\frac{2\pi i}{p}}\right|,\end{aligned}$$

terminando a demonstração. □

A seguir, apresentamos o resultado mais importante desse trabalho. Conforme já dissemos, o Teorema 1.1 de [10] garante convergência de certas sequências de medidas em $\mathcal{M}(\mathbb{S}^1)$ para a medida de Lebesgue no círculo, enquanto no Teorema 2.27, garantimos a convergência de sequências de medidas em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ para a medida de Bernoulli uniforme.

Teorema 2.27. *Se* p *for primo e* $(\eta_n)_{n\in\mathbb{N}}$ *uma sequência de medidas de probabilidade* σ*-invariantes em* $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$ *satisfizer*

$$\inf_n\{h_{\eta_n}(\sigma)\} > 0, \tag{2.7}$$

então

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

na topologia fraca.*

Demonstração. Tome $\chi : (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}$ um caracter não trivial. Como a inequação (2.7) é válida por hipótese, vai existir $\varepsilon_0 > 0$ tal que $h_{\eta_n}(\sigma) \geq \varepsilon_0$ para todo n natural. Escolha $\delta < \min\left\{\frac{1}{p}, \delta(\varepsilon_0, \chi)\right\}$, em que $\delta(\varepsilon_0, \chi)$ é obtido na Proposição 2.22. Assim, devemos ter, quaisquer que sejam o natural n e $j \in (\mathbb{Z}/p\mathbb{Z})$, que

$$\eta_n\left(\chi^{-1}(e^{\frac{2\pi ij}{p}})\right) < 1 - \delta.$$

Agora, escolhendo $\alpha_j = \eta_n\left(\chi^{-1}(e^{\frac{2\pi ij}{p}})\right)$ para cada $j \in (\mathbb{Z}/p\mathbb{Z})$, o Corolário 2.26 permite-nos escrever:

$$\begin{aligned}\left|\int \chi d\eta_n\right| &= \left|\sum_{j=0}^{p-1} \eta_n\left(\chi^{-1}(e^{\frac{2\pi ij}{p}})\right) e^{\frac{2\pi ij}{p}}\right| \\ &\leq \left|\delta + (1-\delta)e^{\frac{2\pi iK}{p}}\right| =: b_\chi < 1,\end{aligned}$$

e assim

$$\left|\int \chi d\eta_n * \eta_{n-1} * \ldots * \eta_1\right| = \prod_{i=1}^{n}\left|\int \chi d\eta_i\right| \leq b_\chi^n \to 0.$$

Como χ era um caracter não-trivial qualquer, pelo Corolário 2.14 (considerando $\mu_n := \eta_n * \ldots * \eta_1$), concluímos que

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

na topologia fraca*. □

No Teorema 2.27 acima assumimos que p fosse primo não por acaso. Nossas mais importantes contribuições no caso de p não primo são o Teorema 2.28 e Corolário 2.29 na sequência. Com esses resultados, mostramos que dado p não primo, sempre haverá

sequências de medidas de probabilidade σ-invariantes que possuem suporte total e entropias uniformemente afastadas de zero mas que não convergem em convolução (Definição 2.15) para a medida de Bernoulli uniforme.

Teorema 2.28. *Sejam p um número não primo e $p = d_1 \cdot d_2$ uma decomposição em fatores não triviais $d_1, d_2 \geq 2$. Então, toda sequência η_n de medidas de probabilidade em $\mathcal{M}\left((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}\right)$ que satisfaz:*

$$\eta_n[j] = \begin{cases} a_n, & \text{se } j = kd_1, \text{ para algum } k \in \{0, 1, \ldots, d_2 - 1\}; \\ \varepsilon_n, & \text{caso contrário,} \end{cases}$$

também satisfaz:

$$\eta_n * \ldots * \eta_1[j] = \begin{cases} a'_n, & \text{se } j = kd_1, \text{ para algum } k \in \{0, 1, \ldots, d_2 - 1\}; \\ \varepsilon'_n, & \text{caso contrário.} \end{cases}$$

E ainda, temos a seguinte estimativa:

$$\varepsilon'_n \leq \sum_{i=1}^{n} \varepsilon_i.$$

Demonstração. Sejam $(\eta_n, a_n, \varepsilon_n)_{n=1}^{\infty}$ exatamente conforme hipótese. O primeiro passo será provarmos que a medida $\eta_2 * \eta_1$ tem a forma $(\eta_2 * \eta_1; a'_2, \varepsilon'_2)$. Nesse caso, concluiríamos que para todo n, a medida $\eta_n * \ldots * \eta_1$ será também da forma $(\eta_n * \ldots * \eta_1, a'_n, \varepsilon'_n)$.

Para as igualdades abaixo, algumas explicações auxiliares. Um número inteiro n é múltiplo de d_1 se e somente se, em $\mathbb{Z}/p\mathbb{Z}$, ele pertencer a uma das seguintes classes $\{0, d_1, \ldots, (d_2 - 1)d_1\} \subset \mathbb{Z}/p\mathbb{Z}$. Em particular, se $n \in \mathbb{Z}$ e $l, l_0 \in \{1, \ldots, d_1 - 1\} \subset \mathbb{Z}$ com $l \neq l_0$, então nem $(l_0 - l)$, nem $(nd_1 + l_0)$ pertencem a uma daquelas classes em $\mathbb{Z}/p\mathbb{Z}$. Munidos dessas informações, tome primeiramente $j = k_0 d_1 + l_0$, com $k_0 \in \{0, 1, \ldots, d_2 - 1\}$ e $l_0 \in \{1, 2, \ldots, d_1 - 1\}$ e defina $A_{l_0} = \{1, 2, \ldots, l_0 - 1, l_0 + 1, \ldots, d_1 - 1\}$. Assim, temos:

$$\eta_2 * \eta_1[j] = \sum_{k=0}^{p-1} \eta_2[k]\eta_1[j - k]$$

$$
\begin{aligned}
&= \sum_{k=0}^{p-1} \eta_2[k]\eta_1[k_0d_1 + l_0 - k] \\
&= \sum_{k=0}^{d_2-1} \eta_2[kd_1]\eta_1[(k_0-k)d_1 + l_0] + \sum_{k=0}^{d_2-1} \eta_2[kd_1 + l_0]\eta_1[(k_0-k)d_1] + \\
&\quad + \sum_{l\in A_{l_0}} \sum_{k=0}^{d_2-1} \eta_2[kd_1 + l]\eta_1[(k_0-k)d_1 + (l_0 - l)] \\
&= d_2a_2\varepsilon_1 + d_2\varepsilon_2a_1 + (p-2d_2)\varepsilon_2\varepsilon_1 \\
&= \varepsilon_2'
\end{aligned}
$$

Para o caso $j' = k_0d_1$, as contas são similares. A primeira constatação, conforme já anunciamos, é a seguinte:

$$
\eta_n * \ldots * \eta_1[j] = \begin{cases} a_n', & \text{se } j = kd_1, \text{ para algum } k \in \{0, 1, \ldots, d_2 - 1\} \\ \varepsilon_n', & \text{caso contrário.} \end{cases}
$$

Usaremos de indução para provarmos que $\varepsilon_n' \leq \sum_{i=1}^{n} \varepsilon_i$. Para $n = 2$, tem-se:

$$
\begin{aligned}
\varepsilon_2' &= d_2a_2\varepsilon_1 + d_2\varepsilon_2a_1 + (p-2d_2)\varepsilon_2\varepsilon_1 \\
&= \varepsilon_2(d_2a_1 + (p-2d_2)\varepsilon_1) + \varepsilon_1(d_2a_2) \\
&\leq \varepsilon_2 + \varepsilon_1.
\end{aligned}
$$

Agora, supondo o resultado provado até para certo $n \in \mathbb{N}$, obtemos:

$$
\begin{aligned}
\varepsilon_{n+1}' &= d_2a_{n+1}\varepsilon_n' + d_2\varepsilon_{n+1}a_n' + (p-2d_2)\varepsilon_{n+1}\varepsilon_n' \\
&= \varepsilon_{n+1}(d_2a_n' + (p-2d_2)\varepsilon_n') + \varepsilon_n'(d_2a_{n+1}) \\
&\leq \varepsilon_{n+1} + \varepsilon_n' \\
&\leq \varepsilon_{n+1} + \varepsilon_n + \varepsilon_{n-1} + \ldots + \varepsilon_1,
\end{aligned}
$$

e terminamos a demonstração. □

Corolário 2.29. *Dado p não primo, existe uma sequência η_n de medidas de probabilidade*

σ-invariantes em $\mathcal{M}_\sigma((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$, todas de Bernoulli, possuindo suporte total e satisfazendo

$$\inf_n\{h_{\eta_n}(\sigma)\} > 0,$$

mas que não convergem em convolução para a medida de Bernoulli uniforme na topologia fraca.*

Demonstração. Dado p não primo, escreva $p = d_1 \cdot d_2$, uma fatoração qualquer em que $d_1, d_2 > 1$. Para cada n natural, defina a seguinte medida de Bernoulli:

$$\eta_n[j] = \begin{cases} \frac{1}{d_2} - (d_1 - 1)\varepsilon_n =: a_n, & \text{se } j = kd_1, \text{ para algum } k \in \{0, 1, \ldots, d_2 - 1\} \\ \varepsilon_n, & \text{caso contrário.} \end{cases}$$

Suponha, adicionalmente, a seguinte estimativa para o somatório infinito:

$$\sum_{n=1}^{\infty} \varepsilon_n < \frac{1}{p}.$$

Como $\varepsilon_n \to 0$, segue que

$$h_{\eta_n}(\sigma) = -d_2 \left(\frac{1}{d_2} - (d_1 - 1)\varepsilon_n\right) \log\left(\frac{1}{d_2} - (d_1 - 1)\varepsilon_n\right) - (p - d_2)\varepsilon_n \log \varepsilon_n \to \log(d_2),$$

garantindo que $\inf_n\{h_{\eta_n}(\sigma)\} > 0$.

Agora, usando o Teorema 2.28, obtemos para $j \neq kd_1$:

$$\begin{aligned} \eta_n * \ldots * \eta_1[j] &\leq \sum_{i=1}^{n} \varepsilon_i \\ &\leq \sum_{i=1}^{\infty} \varepsilon_i \\ &< \frac{1}{p} \end{aligned}$$

Desse modo, para $j \neq kd_1$, concluímos que

$$\eta_n * \ldots * \eta_1[j] \nrightarrow \frac{1}{p},$$

o que garante que a sequência $\mu_n := \eta_n * \ldots * \eta_1$ não converge na topologia fraca* para a medida de Bernoulli uniforme.

□

Capítulo 3

Generalidades sobre a convolução em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$

A operação de convolução em $\mathcal{M}\left((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}\right)$, apesar de depender apenas da estrutura de grupo de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, traz propriedades marcantes para as novas medidas fabricadas. Por exemplo, conforme vimos no Teorema 2.27, sequências de medidas da forma $\mu_n := \eta_n * \ldots * \eta_1$ convergem na topologia fraca* para a medida de Bernoulli uniforme sempre que $(\eta_n)_{n\in\mathbb{N}}$ for sequência de medidas σ-invariantes satisfazendo $\inf_n\{h_{\eta_n}(\sigma)\} > 0$. Isso dá-nos indícios de que a convolução interage tanto com a entropia de medidas σ-invariantes quanto com a proximidade em relação a medida de Bernoulli uniforme considerando-se a topologia fraca* ou, possivelmente, topologias mais fortes.

3.1 Encurtando distâncias e aumentando a entropia

Nessa seção, vamos mostrar que a convolução por uma medida σ-invariante η é uma contração fraca tanto na topologia fraca* quanto na topologia gerada por $\bar{d}$, com os respectivos teoremas 3.1 e 3.3. Mais adiante, com o Teorema 3.6, mostraremos que a convolução faz com que a entropia não decresça e, com a Proposição 3.9, exibiremos um contexto onde pode-se detectar aumento estrito da entropia devido a convolução.

Teorema 3.1. *Existe uma distância d em $\mathcal{M}\left((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}\right)$, que gera a topologia fraca* nesse*

espaço e ainda satisfaz, para quaisquer probabilidades η, μ e ν:

$$d(\eta * \mu, \eta * \nu) \leq d(\mu, \nu).$$

Em particular, escolhendo-se $\eta = \left(\frac{1}{p}, \ldots \frac{1}{p}\right)$, obtemos:

$$d\left(\eta * \mu, \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)\right) \leq \min\left\{d\left(\eta, \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)\right); d\left(\mu, \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)\right)\right\}.$$

Demonstração. Dadas η, μ e ν medidas de probabilidade, fixe $n \in \mathbb{N}$ e, tendo em vista a notação da Observação 1.9, considere o seguinte valor:

$$d_n(\eta * \mu, \eta * \nu) = \sum_{P \in \mathcal{P}^n} |\eta * \mu(P) - \eta * \nu(P)|.$$

Vamos mostrar que $d_n(\eta * \mu, \eta * \nu) \leq d_n(\mu, \nu)$. De fato:

$$\begin{aligned}
d_n(\eta * \mu, \eta * \nu) &= \sum_{P \in \mathcal{P}^n} |\eta * \mu(P) - \eta * \nu(P)| \\
&= \sum_{P \in \mathcal{P}^n} \left| \sum_{Q \in \mathcal{P}^n} \eta(Q)\mu(P-Q) - \sum_{Q \in \mathcal{P}^n} \eta(Q)\nu(P-Q) \right| \\
&= \sum_{P \in \mathcal{P}^n} \left| \sum_{Q \in \mathcal{P}^n} \eta(Q)(\mu(P-Q) - \nu(P-Q)) \right| \\
&\leq \sum_{P \in \mathcal{P}^n} \sum_{Q \in \mathcal{P}^n} \eta(Q) \left|\mu(P-Q) - \nu(P-Q)\right| \\
&= \left(\sum_{Q \in \mathcal{P}^n} \eta(Q) \right) \sum_{P \in \mathcal{P}^n} |\mu(P-Q) - \nu(P-Q)| \\
&= \left(\sum_{Q \in \mathcal{P}^n} \eta(Q) \right) d_n(\mu, \nu) \\
&= d_n(\mu, \nu).
\end{aligned}$$

Agora, retomando novamente a Observação 1.9 e usando os cálculos acima, segue que

a distância dada por:

$$d(\eta,\mu)=\sum_{n=1}^{\infty}\left(\sum_{P\in\mathcal{P}^n}\frac{|\eta(P)-\mu(P)|}{p^n}\right),$$

é tal que

$$\begin{aligned} d\left(\eta*\mu,\eta*\nu\right) &= \sum_{n=1}^{\infty}\left(\sum_{P\in\mathcal{P}^n}\frac{|\eta*\mu(P)-\eta*\nu(P)|}{p^n}\right)\\ &= \sum_{n=1}^{\infty}\left(\frac{d_n(\eta*\mu,\eta*\nu)}{p^n}\right)\\ &\leq \sum_{n=1}^{\infty}\left(\frac{d_n(\mu,\nu)}{p^n}\right)\\ &= d(\mu,\nu)\end{aligned}$$

como queríamos. □

Passemos para os preparativos da contração fraca no contexto a métrica $\bar{d}$. Lembre que os *joinings* são probabilidades de $G\times G$ em que $G=(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Sendo G um grupo, $G\times G$ também o será, com operação de soma coordenada-a-coordenada. Denotemos por $\mathcal{S}:(G\times G)\times(G\times G)\to(G\times G)$ tal operação de soma. Como vale a relação $(\sigma\times\sigma)\circ\mathcal{S}=\mathcal{S}((\sigma\times\sigma)\times(\sigma\times\sigma))$, podemos concluir que a convolução de probabilidades em $G\times G$ preserva as medidas de probabilidade$(\sigma\times\sigma)$-invariantes. Em particular, dadas as medidas de probabilidade σ-invariantes η_1,μ_1,η_2 e μ_2, se $J_1\in\mathcal{J}(\eta_1,\mu_1)$ e $J_2\in\mathcal{J}(\eta_2,\mu_2)$ então J_1*J_2 é uma medida de probabilidade $(\sigma\times\sigma)$-invariante e elaboramos a Proposição 3.2 adiante, que conta-nos mais sobre J_1*J_2 e é peça fundamental para provarmos o Teorema 3.3, com o qual concluímos que convoluir por uma medida η é uma contração fraca em relação a métrica $\bar{d}$. Na Proposição 3.2, para evitar contas extensas, vamos usar notação ligeiramente diferente da que vínhamos utilizando até então. Ao invés de escrevermos um cilindro de tamanho n em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ da forma $[a_1,\ldots,a_n]$, denotaremos $A:=(a_1,\ldots,a_n)\in(\mathbb{Z}/p\mathbb{Z})^n$ e, com certo abuso de notação, escreveremos $[A]:=[a_1,\ldots,a_n]$.

Proposição 3.2. *Sejam η_1,μ_1,η_2 e μ_2 probabilidades σ-invariantes. Se $J_1\in\mathcal{J}(\eta_1,\mu_1)$ e $J_2\in\mathcal{J}(\eta_2,\mu_2)$, então $(J_1*J_2)\in\mathcal{J}(\eta_1*\eta_2,\mu_1*\mu_2)$.*

Demonstração. Seja $[A]$ um cilindro em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, com $A \in (\mathbb{Z}/p\mathbb{Z})^n$. Então:

$$\begin{aligned}
& J_1 * J_2\left([A] \times (\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}\right) \\
= & \sum_{B \in (\mathbb{Z}/p\mathbb{Z})^n} J_1 * J_2([A] \times [B]) \\
= & \sum_{B \in (\mathbb{Z}/p\mathbb{Z})^n} \left(\sum_{I,K \in (\mathbb{Z}/p\mathbb{Z})^n} J_1([I] \times [K]) J_2([A-I] \times [B-K]) \right) \\
= & \sum_{I,K \in (\mathbb{Z}/p\mathbb{Z})^n} J_1([I] \times [K]) \sum_{B \in (\mathbb{Z}/p\mathbb{Z})^n} J_2([A-I] \times [B-K]) \\
= & \sum_{I,K \in (\mathbb{Z}/p\mathbb{Z})^n} J_1([I] \times [K]) \eta_2[A-I] \\
= & \sum_{I \in (\mathbb{Z}/p\mathbb{Z})^n} \eta_1[I] \eta_2[A-I] \\
= & \ \eta_1 * \eta_2([A]).
\end{aligned}$$

Da mesma forma, prova-se que

$$J_1 * J_2((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}} \times [A]) = \mu_1 * \mu_2([A]).$$

Portanto, $J_1 * J_2 \in \mathcal{J}(\eta_1 * \eta_2, \mu_1 * \mu_2)$. □

Teorema 3.3. *Sejam η, μ e ν probabilidades invariantes. Então*

$$\bar{d}(\eta * \mu, \eta * \nu) \leq \bar{d}(\mu, \nu).$$

Em particular, para $\nu = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$, a Bernoulli uniforme, obtemos

$$\bar{d}\left(\eta * \mu, \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)\right) \leq \min\left\{\bar{d}\left(\mu, \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)\right), \bar{d}\left(\eta, \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)\right)\right\}.$$

Demonstração. Seja $J \in \mathcal{J}(\mu, \nu)$. Tome $J_\eta \in \mathcal{J}(\eta, \eta)$ o joining $J_\eta(A \times B) = \eta(A \cap B)$. Pela Proposição 3.2, sabemos que $J_\eta * J \in \mathcal{J}(\eta * \mu, \eta * \nu)$. Mas

$$
\begin{aligned}
J_\eta * J\left(\bigcup_{a\neq b}[a]\times[b]\right) &= \sum_{a\neq b}\sum_{i} J_\eta([i]\times[i])J([a-i]\times[b-i]) \\
&= \sum_{i}\eta([i])\left(\sum_{a\neq b}J([a-i]\times[b-i])\right) \\
&= \sum_{i}\eta([i])\left(\sum_{a\neq b}J([a]\times[b])\right) \\
&= J\left(\bigcup_{a\neq b}[a]\times[b]\right)
\end{aligned}
$$

Como J era qualquer, o teorema segue quando tomamos o ínfimo nas igualdades sobre todos $J \in \mathcal{J}(\mu,\nu)$, por definição de $\bar{d}$. □

Finalmente, vamos tratar da interação entre convolução de medidas σ-invariantes e entropia. Nesse sentido, formulamos os lemas 3.4 e 3.5 para provarmos o não-decrescimento de entropia de R-convoluções específicas com o Teorema 3.6, que é uma de nossas contribuições para estreitar ainda mais a relação entre entropia e certas R-convoluções. A título de informação, enquanto o Lema 3.4 é válido para qualquer espaço mensurável, tanto o Lema 3.5 quanto o Teorema 3.6 contempla apenas o grupo topológico $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$.

Lema 3.4. *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. O primeiro item segue assim:

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

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. □

Lema 3.5. *Seja $R : G \times G \to G$ uma transformação com o seguinte padrão:*

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

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 $\mathbb{Z}/p\mathbb{Z}$. Se $\mathcal{P} = \{[0], \ldots, [p-1]\}$, 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 tal patição é constituída por conjuntos do tipo $R^{-1}([k_1, \ldots, k_n])$, 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, concluímos as seguintes igualdades para um conjunto genérico da partição $R^{-1}(\mathcal{P}^n) \vee \mathcal{P}^n \times G$:

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

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 $R^{-1}(\sigma^{-1}([k_2, \ldots, k_n]))$. Tomando um conjunto $[i_1, \ldots, i_n] \times G$ de $(\mathcal{P}^n \times G)$, obtemos:

$$
\begin{aligned}
& R^{-1}(\sigma^{-1}([k_2,\ldots,k_n]))\cap([i_1,\ldots,i_n]\times G)= \\
= & \bigcup_{a,j_1,\ldots,j_n=0}^{p-1}\left([j_1,\ldots,j_n]\times[a,r(j_2\cdot)^{-1}(k_2),\ldots,r(j_n\cdot)^{-1}(k_n)]\right)\bigcap([i_1,\ldots,i_n]\times G) \\
= & [i_1,\ldots,i_n]\times\bigcup_{a=0}^{p-1}[a,r(i_2\cdot)^{-1}(k_2),\ldots,r(i_n\cdot)^{-1}(k_n)] \\
= & [i_1,\ldots,i_n]\times\bigcup_{a=0}^{p-1}[a,k_2',\ldots,k_n'] \\
= & [i_1,\ldots,i_n]\times\sigma^{-1}([k_2',\ldots,k_n']),
\end{aligned}
$$

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. □

Agora estamos prontos para provar que a entropia não decresce quando consideramos certas R-convoluções. A convolução usual é um exemplo dessas operações.

Teorema 3.6. *Nas mesmas hipóteses do Lema 3.5, são satisfeitas 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 e usando o Teorema 1.14, obtemos:

$$
\begin{aligned}
h_{\eta *_R\mu}(\sigma) &= \lim_n H_{\eta *_R\mu}(\mathcal{P}|\vee_{j=1}^n\sigma^{-j}(\mathcal{P})) \quad \text{(Teorema 1.14)} \\
&= \lim_n H_{\eta\times\mu}(R^{-1}(\mathcal{P})|\vee_{j=1}^n(\sigma\circ\sigma)^{-j}(R^{-1}(\mathcal{P}))) \quad \text{(Lema 3.4)} \\
&\leq \sup_{\mathcal{Q}}\{\lim_n H_{\eta\times\mu}(\mathcal{Q}|\vee_{j=1}^n(\sigma\circ\sigma)^{-j}(\mathcal{Q}))\}
\end{aligned}
$$

$$
\begin{aligned}
&= h_{\eta\times\mu}(\sigma) \quad \text{(Página 255 de [18])} \\
&= h_\eta(\sigma) + h_\mu(\sigma).
\end{aligned}
$$

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

Finalmente, usando os lemas 3.4 e 3.5, obtemos:

$$
\begin{aligned}
& H_{\eta *_R \mu}(\mathcal{P}| \vee_{j=1}^{n} \sigma^{-j}(\mathcal{P})) \\
= \; & H_{\eta\times\mu}(R^{-1}(\mathcal{P}))|R^{-1}(\vee_{j=1}^{n}\sigma^{-j}(\mathcal{P}))) \\
\geq \; & H_{\eta\times\mu}(R^{-1}(\mathcal{P})|R^{-1}(\vee_{j=1}^{n}\sigma^{-j}(\mathcal{P})) \vee \mathcal{P}^n \times G)) \\
= \; & H_{\eta\times\mu}(R^{-1}(\mathcal{P})|\mathcal{P}^n \times \left(\vee_{j=1}^{n-1}\sigma^{-j}(\mathcal{P})\right)) \\
= \; & \sum_{i_k,j_k} \eta \times \mu([i_1,\ldots,i_n]\times[j_1,\ldots,j_n]) \log \left(\frac{\eta \times \mu\left(\bigcup\limits_{a\in(\mathbb{Z}/p\mathbb{Z})} [i_1,\ldots,i_n]\times[a,j_2,\ldots,j_n]\right)}{\eta\times\mu([i_1,\ldots,i_n]\times[j_1,\ldots,j_n])} \right) \\
= \; & \sum_{j_k} \mu([j_1,\ldots,j_n]) \log\left(\frac{\mu([j_2,\ldots,j_n])}{\mu([j_1,\ldots,j_n])}\right) \\
= \; & H_\mu(\mathcal{P}| \vee_{j=1}^{n} \sigma^{-1}(\mathcal{P})),
\end{aligned}
$$

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. □

Exemplo 3.7. Consideremos a transformação $R : G \times G \to G$ dada por $R(x,y) = x + y$. Conforme já falamos, $\eta *_R \mu = \eta * \mu$ é a convolução usual das medidas η e μ. Vejamos alguns exemplos para ilustrar o Teorema 3.6.

Note que se $h_\eta(\sigma) = 0$ então $h_{\eta*\mu} = h_\mu(\sigma)$. Há ainda casos em que a entropia da convolução atinge exatamente o valor da soma das prévias entropias. Por exemplo, se $p = 6$, para $\eta = \left(\frac{1}{2}, 0, 0, \frac{1}{2}, 0, 0\right)$ e $\mu = \left(\frac{1}{3}, 0, \frac{1}{3}, 0, \frac{1}{3}, 0\right)$, (medidas de Bernoulli), obtemos:

$$
\eta * \mu = \left(\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}\right),
$$

e assim:

$$h_{\eta*\mu}(\sigma) = \log(6) = \log(2) + \log(3) = h_\eta(\sigma) + h_\mu(\sigma).$$

Ainda, em outras situações, a convolução de duas medidas com entropia positiva pode não gerar aumento estrito da entropia. Para $p = 4$, consideremos $\eta = \mu = \left(\frac{1}{2}, 0, \frac{1}{2}, 0\right)$ então $\eta * \mu = \eta$ e segue que $h_{\eta*\mu}(\sigma) = h_\eta(\sigma) = \log(2)$.

3.2 Convolução de medidas de Bernoulli

O conjunto das medidas de Bernoulli é fechado em $\mathcal{M}(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ na topologia fraca* e, por serem medidas de probabilidade mais simples, merecem atenção especial. Nessa seção vamos demonstrar, por exemplo, que a convolução de duas medidas de Bernoulli resulta numa medida de Bernoulli, com a Proposição 3.8. Esse resultado, aparentemente desimportante, suporta resultados mais profundos, como a Proposição 3.9, em que mostramos haver aumento estrito na entropia da convolução de duas medidas de Bernoulli com suporte total, usando fortemente que convolução preserva as medidas de Bernoulli. E também, para p primo, com o Teorema 3.10 e Corolário 3.11, mostramos que sequências do tipo $(\mu_n := \eta_n * \ldots * \eta_1)_{n\in\mathbb{N}}$, em que η_n são medidas de Bernoulli tais que $\inf_n\{h_{\eta_n}(\sigma)\} > 0$, satisfazem que a sequência das entropias das μ_n converge para $\log p$, conclusão igual a do Teorema 1.1 do trabalho [10] de Lindenstrauss, sendo respeitados os contextos.

Proposição 3.8. *Sejam R uma transformação nos moldes do Teorema 3.6 e 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, assim, que:

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

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

e concluímos que $\eta *_R \mu$ é medida de Bernoulli. □

O resultado anterior diz, também, que a convolução de duas medidas de Bernoulli é ainda uma medida de Bernoulli. Agora, é sabido que a função $\phi(x) = -x \log x$ é côncava e que

$$\sum_{i=0}^{p-1} t_i \phi(x_i) < \phi\left(\sum_{i=0}^{p-1} t_i x_i\right), \tag{3.1}$$

sempre que $x_i, y_i > 0$ com $\sum_{i=0}^{p-1} t_i = 1$, para todo $i \in \{0, 1, \ldots, p-1\}$. Tal afirmação consta na página 228 de [18]. A partir dessas verdades, conseguimos o seguinte resultado, que é um melhoramento do obtido no Teorema 3.6, porém para uma classe menor de medidas.

Proposição 3.9. *Sejam p um número natural qualquer, η e μ medidas de Bernoulli em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ com suporte total (ou seja, tais que $\eta[i], \mu[i] > 0$ para todo $i \in \{0, 1, ..., p-1\}$) e ambas distintas da medida de Bernoulli uniforme. Então, temos*

$$h_{\eta*\mu}(\sigma) > \max\{h_\eta(\sigma), h_\mu(\sigma)\}.$$

Demonstração. Considere as medidas η e μ conforme hipótese. A intenção aqui é usar o argumento da desigualdade estrita (3.1). Assim, fixe $j \in \{0, 1, \ldots, p-1\}$ e defina, para todo $i \in \{0, 1, \ldots, p-1\}$, $t_i := \mu[j-i]$ e $x_i := \eta[i]$. Por conseguinte:

$$
\begin{aligned}
h_{\eta*\mu}(\sigma) &= \sum_{j=0}^{p-1} (-\eta * \mu[j]) \log(\eta * \mu[j]) \\
&= \sum_{j=0}^{p-1} - \left(\sum_{i=0}^{p-1} \mu[j-i]\eta[i]\right) \log\left(\sum_{i=0}^{p-1} \mu[j-i]\eta[i]\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{j=0}^{p-1} \phi\left(\sum_{i=0}^{p-1} \mu[j-i]\eta[i]\right) \\
&> \sum_{j=0}^{p-1} \left(\sum_{i=0}^{p-1} \mu[j-i]\phi\left(\eta[i]\right)\right) \\
&= \sum_{i=0}^{p-1} \left(\sum_{j=0}^{p-1} \mu[j-i]\phi\left(\eta[i]\right)\right) \\
&= \sum_{i=0}^{p-1} \phi(\eta[i]) \\
&= h_\eta(\sigma).
\end{aligned}
$$

□

Vamos agora discorrer sobre a relação entre as medidas de Bernoulli, métrica $\bar{d}$ (cuja definição encontra-se na Proposição 1.18) e aumento estrito de entropia. Em geral, não se tem uma fórmula explícita para se calcular a distância $\bar{d}$ de duas medidas de probabilidade σ-invariantes quaisquer. Em [11], o autor fala que apenas resultados parciais foram alcançados quando se assume, por exemplo, que as medidas são markovianas. Contudo, utilizando o Teorema 3 de [11] e ideias do Lema 15.26 de [6], apresentamos, no teorema que segue, a fórmula (talvez já conhecida, aparentemente não publicada) da distância $\bar{d}$ entre duas medidas de Bernoulli em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$.

Teorema 3.10. *Sejam η e μ medidas de Bernoulli em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Então*

$$\bar{d}(\eta,\mu) = \frac{1}{2}\sum_{i=0}^{p-1} |\eta[i] - \mu[i]|$$

Demonstração. Sejam η e μ medidas de Bernoulli em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. O primeiro passo será escrever a fórmula da $\bar{d}$ convenientemente:

$$
\begin{aligned}
\bar{d}(\eta,\mu) &= \inf_{J\in\mathcal{J}(\eta,\mu)} J\left(\bigcup_{i\neq j}([i]\times[j])\right) \\
&= \inf_{J\in\mathcal{J}(\eta,\mu)} \sum_{i=0}^{p-1}\left(\sum_{j\neq i} J\left([i]\times[j]\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \inf_{J\in\mathcal{J}(\eta,\mu)} \sum_{i=0}^{p-1} (1 - J([i]\times[i])) \\
&= \inf_{J\in\mathcal{J}(\eta,\mu)} \sum_{i=0}^{p-1} \eta[i] - J([i]\times[i]).
\end{aligned}
$$

Também, todo $J \in \mathcal{J}(\eta,\mu)$ satisfaz, por definição:

$$J([i]\times[i]) \leq \min\{\eta[i],\mu[i]\},$$

para todo $i \in \{0,1,\ldots,p-1\}$. Portanto:

$$
\begin{aligned}
\bar{d}(\eta,\mu) &\geq \sum_{i=0}^{p-1} \eta[i] - \min\{\eta[i],\mu[i]\} \\
&= \sum_{i=0}^{p-1} \left(\eta[i] + \frac{|\eta[i]-\mu[i]| - \eta[i] - \mu[i]}{2}\right) \\
&= \frac{1}{2}\sum_{i=0}^{p-1} |\eta[i]-\mu[i]|.
\end{aligned}
$$

Para finalizarmos com a igualdade, vamos encontrar um joining $J_{\eta\mu}$ tal que:

$$J_{\eta\mu}\left(\bigcup_{i\neq j}[i]\times[j]\right) = \frac{1}{2}\sum_{i=0}^{p-1}|\eta[i]-\mu[i]|.$$

Assim, é suficiente que tal *joining* satisfaça, para todo $i \in \{0,1,\ldots,p-1\}$:

$$J_{\eta\mu}([i]\times[i]) = \min\{\eta[i],\mu[i]\}.$$

Se $p = 2$ e $\eta[0] \leq \mu[0]$, basta considerarmos o joining independente $J_{\eta\mu}$ (ver o Exemplo 1.17) dado por:

$$
\begin{pmatrix} J_{\eta\mu}(00) & J_{\eta\mu}(01) \\ J_{\eta\mu}(10) & J_{\eta\mu}(11) \end{pmatrix} = \begin{pmatrix} \eta([0]) & 0 \\ \mu([0]) - \eta([0]) & \mu([1]) \end{pmatrix},
$$

em que $J_{\eta\mu}(ij) := J_{\eta\mu}([i] \times [j])$.

Prossigamos por indução. Suponha verdadeiro o resultado até para certo $p \in \mathbb{N}$, ou seja, suponha que dados $p' \in \{2, 3, \ldots, p\}$ e η', μ' medidas de Bernoulli em $(\mathbb{Z}/p'\mathbb{Z})^{\mathbb{N}}$, sempre existe um *joining* $J_{\eta'\mu'}$ tal que $J_{\eta'\mu'}([i]\times[i]) = \min\{\eta'[i], \mu'[i]\}$ para $i \in \{0, 1, \ldots, p'-1\}$.

Sejam, então, η e μ medidas de Bernoulli em $(\mathbb{Z}/(p+1)\mathbb{Z})^{\mathbb{N}}$. Sem perda de generalidade, podemos supor que existe $k \leq p$ tal que

$$\min\{\eta[i], \mu[i]\} = \begin{cases} \eta[i], & \text{se } i < k \\ \mu[i], & \text{se } i \geq k. \end{cases}$$

No caso em que $k = p$, podemos obter o *joining* independente $J_{\eta\mu}$ (ver o Exemplo 1.17) dado por:

$$J_{\eta\mu}([i] \times [j]) = \begin{cases} \eta[i], & \text{se } i = j \leq p-1 \\ \mu[j] - \eta[j], & \text{se } i = p \text{ e } 0 \leq j \leq p-1. \\ \mu[p], & \text{se } i = j = p. \\ 0, & \text{caso contrário} \end{cases}$$

Para facilitar o entendimento, se tomássemos $p + 1 = 3$, teríamos:

$$\begin{pmatrix} J_{\eta\mu}(00) & J_{\eta\mu}(01) & J_{\eta\mu}(02) \\ J_{\eta\mu}(10) & J_{\eta\mu}(11) & J_{\eta\mu}(12) \\ J_{\eta\mu}(20) & J_{\eta\mu}(21) & J_{\eta\mu}(22) \end{pmatrix} = \begin{pmatrix} \eta([0]) & 0 & 0 \\ 0 & \eta([1]) & 0 \\ \mu([0]) - \eta([0]) & \mu([1]) - \eta([1]) & \mu([2]) \end{pmatrix}$$

Agora, consideremos o caso $k < p$ e defina as seguintes medidas de Bernoulli η' e μ' em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$:

$$\begin{aligned} \mu' &= \frac{1}{1-\mu[p]}(\mu[0], \mu[1], \ldots, \mu[p-1]); \\ \eta' &= \frac{1}{1-\mu[p]}(\eta[0], \eta[1], \ldots, \eta[p-2], \eta[p-1] + \eta[p] - \mu[p]). \end{aligned}$$

Pela hipótese de indução, seja $J_{\eta'\mu'}$ um *joining* de η' e μ' que satisfaça, para todo $i \in$

$\{0, 1, \ldots, p-1\}$:

$$J_{\eta'\mu'}([i] \times [i]) = \min\{\eta'[i], \mu'[i]\}.$$

A partir de $J_{\eta'\mu'}$, construiremos um *joining* $J_{\eta\mu}$ de η e μ para finalizar o teorema. Para $a_{ij} := (1 - \mu[p])J_{\eta'\mu'}([i] \times [j])$, $i, j \in \{0, 1, \ldots, p-1\}$, obtemos:

$$\sum_{j=0}^{p-1} a_{ij} = \begin{cases} \eta[i], & \text{se } i \neq p-1 \\ \eta[p-1] + \eta[p] - \mu[p], & \text{se } i = p-1, \end{cases}$$

e, para todo $j \in \{0, 1, \ldots, p-1\}$,

$$\sum_{i=0}^{p-1} a_{ij} = \mu[j].$$

De fato, seja $i \neq p-1$. Então:

$$\begin{aligned} \sum_{j=0}^{p-1} a_{ij} &= (1 - \mu[p]) \sum_{j=0}^{p-1} J_{\eta'\mu'}([i] \times [j]) \\ &= (1 - \mu[p])\eta'[i] \\ &= (1 - \mu[p]) \left(\frac{\eta[i]}{1 - \mu([p])} \right) \\ &= \eta[i]. \end{aligned}$$

Agora, se $i = p-1$, é válido o seguinte:

$$\begin{aligned} \sum_{j=0}^{p-1} a_{ij} &= (1 - \mu[p]) \sum_{j=0}^{p-1} J_{\eta'\mu'}([i] \times [j]) \\ &= (1 - \mu[p]) \sum_{j=0}^{p-1} J_{\eta'\mu'}([p-1] \times [j]) \\ &= (1 - \mu[p])\eta'[p-1] \\ &= (1 - \mu[p]) \left(\frac{\eta[p-1] + \eta[p] - \mu[p]}{1 - \mu([p])} \right) \\ &= \eta[p-1] + \eta[p] - \mu[p]. \end{aligned}$$

E, finalmente:

$$\begin{aligned}\sum_{i=0}^{p-1} a_{ij} &= (1-\mu[p])\sum_{i=0}^{p-1} J_{\eta'\mu'}([i]\times[j]) \\ &= (1-\mu[p])\mu'([j]) \\ &= (1-\mu[p])\left(\frac{\mu[j]}{1-\mu[p]}\right) \\ &= \mu[j].\end{aligned}$$

Note que $\eta[p-1]+\eta[p]-\mu[p] \geq \eta[p-1]$. Então, escolha quaisquer números reais não negativos x_l, $l \in \{0,1,\ldots,p-2\}$, tais que satisfazem:

$$\sum_{l=0}^{p-2} x_l = \eta[p]-\mu[p]$$

e

$$x_l \leq a_{(p-1)l}.$$

Para finalizar, defina o seguinte joining independente $J_{\eta\mu}$ de η e μ por

$$J_{\eta\mu}([i]\times[j]) = \begin{cases} a_{ij}, & \text{se } 0 \leq i \leq p-2 \text{ e } 0 \leq j \leq p-1; \\ a_{(p-1)j} - x_j, & \text{se } i = p-1 \text{ e } 0 \leq j \leq p-2; \\ a_{p-1p-1}, & \text{se } i = j = p-1; \\ x_j, & \text{se } i = p \text{ e } 0 \leq j \leq p-2; \\ 0, & \text{se } i = p \text{ e } j = p-1; \\ \mu[p], & \text{se } i = j = p; \\ 0, & \text{se } 0 \leq i \leq p-1 \text{ e } j = p \end{cases}$$

Agora vamos ver que $J_{\eta\mu}$ é de fato um joining de η e μ que satisfaz

$$J_{\eta\mu}([i]\times[i]) = \min\{\eta[i],\mu[i]\}.$$

Para $0 \leq i \leq p-2$, temos

$$\begin{aligned}\sum_{j=0}^{p} J_{\eta\mu}([i]\times[j]) &= \sum_{j=0}^{p-1} a_{ij} + J_{\eta\mu}([i]\times[p]) \\ &= \eta[i].\end{aligned}$$

Se $i = p-1$, vale o seguinte:

$$\begin{aligned}\sum_{j=0}^{p} J_{\eta\mu}([p-1]\times[j]) &= \sum_{j=0}^{p-1} a_{p-1}j - \sum_{j=0}^{p-2} x_j + J_{\eta\mu}[p-1]\times[p] \\ &= \eta[p-1] + \eta[p] - \mu[p] - (\eta[p] - \mu[p]) \\ &= \eta[p-1].\end{aligned}$$

E, para $i = p$, obtemos:

$$\begin{aligned}\sum_{j=0}^{p} J_{\eta\mu}([p]\times[j]) &= \sum_{j=0}^{p-2} J_{\eta\mu}([p]\times[j]) + J_{\eta\mu}[p]\times[p-1] + J_{\eta\mu}[p]\times[p] \\ &= \sum_{j=0}^{p-2} x_j + 0 + \mu[p] \\ &= \eta[p] - \mu[p] + \mu[p] \\ &= \eta[p].\end{aligned}$$

Precisamos agora fazer todo o raciocínio para a medida μ. Se $j \in \{0, 1, \ldots, p-2\}$, então:

$$\begin{aligned}\sum_{i=0}^{p} J_{\eta\mu}([i]\times[j]) &= \sum_{i=0}^{p-2} J_{\eta\mu}([i]\times[j]) + J_{\eta\mu}([p-1]\times[j]) + J_{\eta\mu}([p]\times[j]) \\ &= \sum_{i=0}^{p-2} a_{ij} + a_{p-1j} - x_j + x_j \\ &= \sum_{i=0}^{p-1} a_{ij} \\ &= \mu[j].\end{aligned}$$

Similarmente, para $j = p-1$ tem-se:

$$\begin{aligned}\sum_{i=0}^{p} J_{\eta\mu}([i]\times[p-1]) &= \sum_{i=0}^{p-1} J_{\eta\mu}([i]\times[p-1]) + J_{\eta\mu}([p]\times[p-1]) \\ &= \sum_{i=0}^{p-1} a_{ip-1} + 0 \\ &= \mu[p-1].\end{aligned}$$

Por último, se $j = p$, as duas últimas linhas da definição de $J_{\eta\mu}$ garantem-nos:

$$\begin{aligned}\sum_{i=0}^{p} J_{\eta\mu}([i]\times[p]) &= J_{\eta\mu}([p]\times[p]) \\ &= \mu[p].\end{aligned}$$

Apenas para uma melhor visualização, no caso em que $k = 2$ e $p+1 = 4$, temos:

$$\begin{pmatrix} J_{\eta\mu}(00) & J_{\eta\mu}(01) & J_{\eta\mu}(02) & J_{\eta\mu}(03) \\ J_{\eta\mu}(10) & J_{\eta\mu}(11) & J_{\eta\mu}(12) & J_{\eta\mu}(13) \\ J_{\eta\mu}(20) & J_{\eta\mu}(21) & J_{\eta\mu}(22) & J_{\eta\mu}(23) \\ J_{\eta\mu}(30) & J_{\eta\mu}(31) & J_{\eta\mu}(32) & J_{\eta\mu}(33) \end{pmatrix} = \begin{pmatrix} a_{00} & a_{01} & a_{02} & 0 \\ a_{10} & a_{11} & a_{12} & 0 \\ a_{20}-x_0 & a_{21}-x_1 & a_{22} & 0 \\ x_0 & x_1 & 0 & \mu([3]) \end{pmatrix}.$$

Finalmente, vamos calcular $J_{\eta\mu}([i]\times[i])$ para todo $i \in \{0, 1, \ldots, p\}$. Seja k da definição de $J_{\eta\mu}$. Se $i < k$, então:

$$\begin{aligned}J_{\eta\mu}([i]\times[i]) &= a_{ii} \\ &= (1-\mu[p])J_{\eta'\mu'}([i]\times[i]) \\ &= (1-\mu[p])\min\{\eta'[i], \mu'[i]\} \\ &= (1-\mu[p])\frac{\eta[i]}{1-\mu[p]} \\ &= \eta[i] \\ &= \min\{\eta[i], \mu[i]\}.\end{aligned}$$

Se $k \leq i < p$, obtemos:

$$\begin{aligned} J_{\eta\mu}([i] \times [i]) &= a_{ii} \\ &= (1-\mu[p]) J_{\eta'\mu'}([i] \times [i]) \\ &= (1-\mu[p]) \min\{\eta'[i], \mu'[i]\} \\ &= (1-\mu[p]) \frac{\mu[i]}{1-\mu[p]} \\ &= \mu[i] \\ &= \min\{\eta[i], \mu[i]\}. \end{aligned}$$

Agora, para $i = p$, tem-se:

$$\begin{aligned} J_{\eta\mu}([p] \times [p]) &= \mu[p] \\ &= \min\{\eta[p], \mu[p]\}, \end{aligned}$$

como queríamos. □

Corolário 3.11. *Se p for primo e $(\eta_n)_{n\in\mathbb{N}}$ uma sequência de medidas de Bernoulli em $\mathcal{M}\left((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}\right)$ satisfizer*

$$\inf_n \{h_{\eta_n}(\sigma)\} > 0,$$

então $\bar{d}\left(\eta_n * \ldots * \eta_1, \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)\right) \to 0$.

Em particular, $h_{\eta_n * \ldots * \eta_1}(\sigma) \to \log p$.

Demonstração. Pelo Teorema 2.27, obtemos

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

na topologia fraca*. Em particular, para cada cilindro de tamanho 1, $[i]$, conseguimos

$$\eta_n * \ldots * \eta_1[i] \to \frac{1}{p}.$$

Assim, segue pelo Teorema 3.10 que

$$\bar{d}\left(\eta_n * \ldots * \eta_1, \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)\right) = \sum_{i=0}^{p-1} \frac{1}{2} \left|\eta_n * \ldots * \eta_1[i] - \frac{1}{p}\right| \to_n 0.$$

Finalmente, lembrando que a entropia é $\bar{d}$ contínua (conforme consta no décimo quinto capítulo de [6] e também em nossa seção 1.5), concluímos que $h_{\eta_n * \ldots * \eta_1}(\sigma) \to \log p$ e terminamos o corolário. □

3.3 A equação $\eta * \mu = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$

Conforme já escrevemos na introdução desse trabalho, a seguinte equação envolvendo convolução, $\eta * \eta = \eta$, já foi estudada por Cohen em [3]. Ali, ele caracterizou todas as medidas η que satisfazem tal equação, as chamadas *medidas idempotentes*. Muitos outros trabalhos tratam dessas e de outras indagações similares, como por exemplo [7] e [14]. Assim, é um tanto frequente o estudo de equações envolvendo convolução e cujas variáveis são as medidas.

Nesse sentido, fixando $p \in \mathbb{N}$, podemos nos perguntar quando que a equação

$$\eta * \mu = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right), \tag{3.2}$$

nas variáveis η e μ, é satisfeita. Já sabemos, por exemplo, que $\eta * \left(\frac{1}{p}, \ldots, \frac{1}{p}\right) = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$, para toda medida de probabilidade η em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Chamaremos as soluções do tipo $\left(\eta, \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)\right)$ de *soluções triviais*. Assim, o objetivo dessa seção é encontrar contextos em que a equação (3.2) possui soluções *não* triviais, ou seja, soluções do tipo (η, μ) em que ambas as medidas são distintas da Bernoulli uniforme.

Para aquecermos, apresentamos a Proposição 3.12, que fala da relação entre medidas de Bernoulli e a equação (3.2).

Proposição 3.12. *Sejam $p \in \{2, 3\}$ e η e μ medidas de Bernoulli em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Se*

$$\eta * \mu = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right),$$

então ou $\eta = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$ *ou* $\mu = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$.

Demonstração. Primeiramente, tomemos $p = 2$. Se η e μ são medidas de Bernoulli tais que $\eta * \mu = \left(\frac{1}{2}, \frac{1}{2}\right)$ então, para χ_1 o caracter não trivial constante em cilindros de tamanho 1, $\chi_1([1]) = -1$, temos

$$\begin{aligned} \int \chi_1 d\eta * \mu &= \left(\int \chi_1 d\eta\right)\left(\int \chi_1 d\mu\right) \\ &= (\eta[0] - \eta[1])\,(\mu[0] - \mu[1]) \\ &= 0, \end{aligned}$$

e assim, temos $\eta[0] = \eta[1]$ ou $\mu[0] = \mu[1]$, ou seja, η ou μ é a medida de Bernoulli uniforme em $(\mathbb{Z}/2\mathbb{Z})^{\mathbb{N}}$.

O caso $p = 3$ é similar. Se η e μ são medidas de Bernoulli tais que $\eta * \mu = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$, então, para $\chi_1 : (\mathbb{Z}/3\mathbb{Z})^{\mathbb{N}} \to \mathbb{C}$ o caracter constante em cilindros de tamanho 1 tal que $\chi_1([1]) = e^{\frac{2\pi i}{3}}$, devemos ter, sem perda de generalidade, que $\int \chi_1 d\eta = 0$. As informações contidas nessa igualdade e no fato de η ser uma medida de probabilidade podem ser traduzidas, matricialmente, da seguinte forma:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & \cos\left(\frac{2\pi}{3}\right) & \cos\left(\frac{4\pi}{3}\right) \\ 1 & \sin\left(\frac{2\pi}{3}\right) & \sin\left(\frac{4\pi}{3}\right) \end{pmatrix} \begin{pmatrix} \eta[0] \\ \eta[1] \\ \eta[2] \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Agora, a única possibilidade para η é termos:

$$\begin{pmatrix} \eta[0] \\ \eta[1] \\ \eta[2] \end{pmatrix} = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix},$$

pois a matriz

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & \cos\left(\frac{2\pi}{3}\right) & \cos\left(\frac{4\pi}{3}\right) \\ 1 & \sin\left(\frac{2\pi}{3}\right) & \sin\left(\frac{4\pi}{3}\right) \end{pmatrix}$$

é invertível, já que seu determinante é $\frac{3\sqrt{3}}{2} \neq 0$. E assim, η é a medida de Bernoulli uniforme, como queríamos. □

Observação 3.13. A seguinte constatação

$$\begin{pmatrix} \frac{1}{\sqrt{5}} & 0 & \frac{2}{5+\sqrt{5}} & \frac{2}{5+\sqrt{5}} & 0 \\ 0 & \frac{1}{\sqrt{5}} & 0 & \frac{2}{5+\sqrt{5}} & \frac{2}{5+\sqrt{5}} \\ \frac{2}{5+\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} & 0 & \frac{2}{5+\sqrt{5}} \\ \frac{2}{5+\sqrt{5}} & \frac{2}{5+\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} & 0 \\ 0 & \frac{2}{5+\sqrt{5}} & \frac{2}{5+\sqrt{5}} & 0 & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \frac{2}{5+\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ \frac{2}{5+\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \\ \frac{1}{5} \end{pmatrix},$$

diz que para o caso $p = 5$, a equação $\eta * \mu = \left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right)$ possui, por exemplo, a seguinte solução não trivial:

$$(\eta, \mu) \;=\; \left(\left(\frac{1}{\sqrt{5}}, 0, \frac{2}{5+\sqrt{5}}, \frac{2}{5+\sqrt{5}}, 0\right); \left(0, 0, \frac{2}{5+\sqrt{5}}, \frac{1}{\sqrt{5}}, \frac{2}{5+\sqrt{5}}\right)\right).$$

No caso, η e μ são as medidas de Bernoulli especificadas pelos vetores acima, conforme notação sugerida já no Exemplo 1.6. E isso é uma constatação muito diferente da que ocorre com os primos $p = 2$ e $p = 3$, conforme mostramos na Proposição 3.12.

O caso p não primo é tratado no teorema que segue.

Teorema 3.14. *Seja $p \in \mathbb{N}$ não primo. Então a equação*

$$\eta * \mu = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right),$$

nas variáveis $\eta, \mu \in \mathcal{M}(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$, apresenta uma solução não trivial (e portanto infinitas).

Demonstração. Sendo p não primo, seja $d \geq 2$ um divisor de p e defina a seguinte medida de Bernoulli:

$$\eta[j] := \begin{cases} \frac{d}{p}, & \text{se } j = kd,\ k = 0, 1, \ldots, \frac{p}{d} - 1 \\ 0, & \text{cc} \end{cases}.$$

Agora, para μ, escolha alguma medida, também de Bernoulli, que satisfaça

$$\sum_{k=0}^{\frac{p}{d}-1} \mu[kd+j] = \frac{1}{d}, \tag{3.3}$$

qualquer que seja $j \in \{0, 1, 2, \ldots, d-1\}$.

Devido a estrutura de $\mathbb{Z}/p\mathbb{Z}$, a equação (3.3) é automaticamente válida para qualquer $j \in \mathbb{Z}/p\mathbb{Z}$, não somente para $j \in \{0, 1, 2, \ldots, d-1\}$ e também, como o inverso de kd em $\mathbb{Z}/p\mathbb{Z}$ é $\left(\frac{p}{d} - k\right) d$ para qualquer $k \in \left\{0, 1, \ldots, \frac{p}{d} - 1\right\}$, podemos escrever, para todo $j \in \{0, 1, \ldots, p-1\}$:

$$\sum_{k=0}^{\frac{p}{d}-1} \mu[kd+j] = \sum_{k=0}^{\frac{p}{d}-1} \mu[j-kd].$$

Nessas condições, fixado $j_0 \in (\mathbb{Z}/p\mathbb{Z})$, obtemos:

$$\begin{aligned}
\eta * \mu[j_0] &= \sum_{i \in (\mathbb{Z}/p\mathbb{Z})} \eta[i]\mu[j_0 - i] \\
&= \sum_{k=0}^{\frac{p}{d}-1} \eta[kd]\mu[j_0 - kd] \\
&= \frac{d}{p} \sum_{k=0}^{\frac{p}{d}-1} \mu[j_0 - kd] \\
&= \frac{d}{p}\frac{1}{d} \\
&= \frac{1}{p}.
\end{aligned}$$

Segue, assim, que $\eta * \mu = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$. Como há várias medidas de Bernoulli que satisfazem a equação (3.3), terminamos a demonstração do teorema. □

Capítulo 4

Matrizes e caracteres

Reservamos esse capítulo para construir pontes entre dinâmica simbólica, álgebra linear e caracteres do grupo $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Considerando a operação de convolução de probabilidades (não necessariamente σ-invariantes), construiremos matrizes associadas às medidas que se relacionam tanto com a operação de convolução em si, quanto com os caracteres, cilindros e topologia fraca* e ainda mostraremos, com Teorema 4.3, que tais matrizes são diagonalizáveis via uma linguagem de sistemas dinâmicos e dinâmica simbólica. E finalmente, estendemos os resultados sobre diagonalização para uma classe maior de matrizes, não somente àquelas associadas às medidas, com o Teorema 4.4.

Para início de conversa, nessa construção retomaremos a Proposição 2.2 bem como a notação ali utilizada. No que segue, i e j são tais que $1 \leq i, j \leq p^m$, e $i-1 = \sum_{k=0}^{m-1} i_k p^k$ e $j-1 = \sum_{k=0}^{m-1} j_k p^k$ são as expansões em base p de $i-1$ e $j-1$.

Dada uma medida de probabilidade $\eta \in \mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$, a matriz A_η^m de ordem p^{2m} cujos elementos $(A_\eta^m)_{ij}$ são dados por:

$$\begin{aligned} (A_\eta^m)_{ij} &= \eta[(j_{m-1} - i_{m-1}) \mod p, \ldots, (j_0 - i_0) \mod p] && (4.1) \\ &= \eta[j_{m-1} - i_{m-1}, \ldots, j_0 - i_0], && (4.2) \end{aligned}$$

é a *matriz associada a η considerando cilindros de tamanho m*. No caso, omitimos da

segunda igualdade acima a escrita mod p para não poluir as contas, mas ela continua subentendida. E, de forma similar, definimos o *vetor associado a η considerando cilindros de tamanho m*, v_η^m, como sendo a primeira coluna da matriz A_η^m, ou seja,

$$(v_\eta^m)_i = A_{i1}.$$

A relação dessas matrizes com a operação de convolução de medidas de probabilidade é explicada na proposição abaixo.

Proposição 4.1. *Sejam η e μ medidas em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$. Então para todo $m \in \mathbb{N}$, A_η^m é biestocástica e vale a seguinte igualdade:*

$$A_\eta^m A_\mu^m = A_{\eta * \mu}^m$$

Demonstração. Que a matriz A_η^m é biestocástica, isso segue das seguintes igualdades:

$$\sum_{i_0,\ldots,i_{m-1}=0}^{p-1} \eta[j_0 - i_0, \ldots, j_{m-1} - i_{m-1}] = 1 = \sum_{j_0,\ldots,j_{m-1}=0}^{p-1} \eta[j_0 - i_0, \ldots, j_{m-1} - i_{m-1}].$$

Em relação a segunda parte, para o cálculo de $(A_\eta^m A_\mu^m)_{ij}$, obtemos:

$$\begin{aligned}
(A_\eta^m A_\mu^m)_{ij} &= \sum_{l=1}^{p^m} (A_\eta^m)_{il} (A_\mu^m)_{lj} \\
&= \sum_{l=1}^{p^m} \eta[l_{m-1} - i_{m-1}, \ldots, l_0 - i_0] \mu[j_{m-1} - l_{m-1}, \ldots, j_0 - l_0] \\
&= \sum_{l'=1}^{p^m} \eta[l'_{m-1}, \ldots, l'_0] \mu[j_{m-1} - i_{m-1} - l'_{m-1}, \ldots, j_0 - i_0 - l'_0] \\
&= \eta * \mu[j_{m-1} - i_{m-1}, \ldots, j_0 - i_0] \\
&= (A_{\eta*\mu}^m)_{ij},
\end{aligned}$$

e a proposição está provada. □

Observação 4.2. Em particular, de acordo com a proposição acima, obtemos também:

$$A_\eta^m v_\mu^m = v_{\eta*\mu}^m,$$

para qualquer m natural e quaisquer medidas de probabilidade η e μ.

Teorema 4.3. *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 A_η^m for a matriz associada a η considerando cilindros de tamanho m e v_χ o vetor associado ao caracter χ (definido em 2.2), então vale a igualdade vetorial:*

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

Em particular, para toda medida de probabilidade η e para todo $m \in \mathbb{N}$, a matriz A_η^m é diagonalizável e se $\{\chi_1, \ldots, \chi_{p^m}\}$ forem os caracteres constantes em cilindros de tamanho m e $V_m = \{v_{\chi_1}, \ldots, v_{\chi_{p^m}}\}$ os respectivos vetores associados, então V_m é uma base de autovetores ortogonais para A_η^m cujos respectivos autovalores são $\{\int \chi_1 d\eta, \ldots, \int \chi_{p^m} d\eta\}$.

Demonstração. Sejam χ um caracter constante em cilindros de tamanho m, η medida de probabilidade em $\mathcal{M}((\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}})$, A_η^m e v_χ conforme hipótese. Lembrando que para $i-1 = \sum_{k=0}^{m-1} i_k p^k$ temos

$$(v_\chi)_i = \chi[i_{m-1}, \ldots, i_1, i_0],$$

portanto (fazendo $i-1 = \sum_{k=0}^{m-1} i_k p^k$, $j-1 = \sum_{k=0}^{m-1} j_k p^k$ e $l-1 = \sum_{k=0}^{m-1} l_k p^k$) obtemos:

$$\begin{aligned}
\left(\int \chi d\eta \, v_\chi \right)_i &= \sum_{l_{m-1},\ldots,l_0=0}^{p-1} \eta[l_{m-1}, \ldots, l_0] \chi[l_{m-1}, \ldots, l_0] \chi[i_{m-1}, \ldots, i_0] \\
&= \sum_{l_{m-1},\ldots,l_0=0}^{p-1} \eta[l_{m-1}, \ldots, l_0] \chi[l_{m-1} + i_{m-1}, \ldots, l_0 + i_0] \\
&= \sum_{j_{m-1},\ldots,j_0=0}^{p-1} \eta[j_{m-1} - i_{m-1}, \ldots, j_0 - i_0] \chi[j_{m-1}, \ldots, j_0]
\end{aligned}$$

$$= \quad (A_\eta^m v_\chi)_i,$$

nas igualdades acima, fizemos uma troca de variáveis $j = l+i$. Assim, cada v_χ é autovetor de A_η^m associado ao autovelor $\int \chi d\eta$. Mas, como consequência do Teorema 3.2.1 de [12], conforme já discutimos nessa seção, V_m é uma base ortogonal para $\mathbb{C}^{p^m}$, o que nos faz findar o presente teorema. □

Agora, para finalizar a seção, vamos estender o Teorema 4.3 para uma classe maior de matrizes a serem diagonalizadas pelos mesmos vetores v_χ. Em primeiro lugar, note que as matrizes da forma A_η^m são diagonalizáveis não por estarem associadas a uma medida de probabilidade, mas sim, porque existe uma função real (ou complexa) f_η com domínio em $(\mathbb{Z}/p\mathbb{Z})^m$ tal que

$$(A_\eta^m)_{ij} = f_\eta((j_{m-1}, \ldots, j_0) - (i_{m-1}, \ldots, i_0)),$$

no caso, a função f_η é definida por:

$$f_\eta(i_{m-1}, \ldots, i_0) = \eta[i_{m-1}, i_{m-2}, \ldots, i_0].$$

Essas informações estão bem detalhadas no teorema que segue, que é o mais importante desta seção (e nossa mais importante contribuição) no que se refere a diagonalização de uma classe especial de matrizes.

Teorema 4.4. *Seja A uma matriz quadrada de ordem p^m tal que existe uma função $f : (\mathbb{Z}/p\mathbb{Z})^m \to \mathbb{C}$ que satisfaz, para $i - 1 = \sum_{k=0}^{m-1} i_k p^k$ e $j - 1 = \sum_{k=0}^{m-1} j_k p^k$:*

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

Então A é diagonalizável, $B = \{v_{\chi_1}, \ldots, v_{\chi_{p^m}}\}$ é uma base de autovetores ortogonais de

A associados aos autovalores $C = \{\langle f, v_{\chi_1}\rangle, \langle f, v_{\chi_2}\rangle, \ldots, \langle f, v_{\chi_{p^m}}\rangle\}$, *em que*

$$\langle f, v_{\chi_i}\rangle = \sum_{j=1}^{p^m} f(j_{m-1}, \ldots, j_0) \cdot \chi_i([j_{m-1}, \ldots, j_0]).$$

Demonstração. A demonstração é praticamente uma mimetização daquela apresentada no Teorema 4.3. De fato (usando mais uma vez que $i-1 = \sum_{k=0}^{m} i_k p^k$, $j-1 = \sum_{k=0}^{m} j_k p^k$ e $l-1 = \sum_{k=0}^{m} l_k p^k$), temos:

$$\begin{aligned}
(\langle f, v_{\chi_j}\rangle \cdot v_{\chi_j})_i &= \left(\sum_{l=1}^{p^m} f(l_{m-1}, \ldots, l_0) \cdot \chi_j[l_{m-1}, \ldots, l_0] \cdot v_{\chi_j}\right)_i \\
&= \sum_{l=1}^{p^m} f(l_{m-1}, \ldots, l_0) \cdot \chi_j([l_{m-1}, \ldots, l_0])\chi_j[i_{m-1}, \ldots, i_0] \\
&= \sum_{l=1}^{p^m} f(l_{m-1}, \ldots, l_0) \cdot \chi_j([l_{m-1} + i_{m-1}, \ldots, l_0 + i_0]) \\
&= \sum_{l=1}^{p^m} f((l_{m-1} - i_{m-1}, \ldots, l_0 - i_0)) \cdot \chi_j([l_{m-1}, \ldots, l_0]) \\
&= (Av_{\chi_j})_i,
\end{aligned}$$

e isso mostra que v_{χ_j} é autovetor de A associado ao autovalor $\langle f, v_{\chi_j}\rangle$, como queríamos demonstrar. □

Exemplo 4.5. Considere $p = 3$ e seja η a medida de Bernoulli associada ao vetor $\left(\frac{1}{10}, \frac{7}{10}, \frac{2}{10}\right)$. Assim, a matriz A_η^1 associada a η considerando cilindros de tamanho 1 é a matriz circulante dada por:

$$A_\eta^1 = \begin{pmatrix} \frac{1}{10} & \frac{7}{10} & \frac{2}{10} \\ \frac{2}{10} & \frac{1}{10} & \frac{7}{10} \\ \frac{7}{10} & \frac{2}{10} & \frac{1}{10} \end{pmatrix}.$$

Observe que tal matriz é associada a função $f_\eta^1 : \mathbb{Z}/3\mathbb{Z} \to \mathbb{C}$, $(f_\eta^1(0), f_\eta^1(1), f_\eta^1(2)) = \left(\frac{1}{10}, \frac{7}{10}, \frac{2}{10}\right)$, conforme com o Teorema 4.4. De acordo com esse mesmo teorema e também com [17], sendo $\{\chi_0, \chi_1, \chi_2\}$ os caracteres constantes em cilindros de tamanho 1 (dados

por $\chi_j([l]) = e^{\frac{2\pi ijl}{3}}$), obtemos os autovalores $\left\{\int \chi_0 d\eta, \int \chi_1 d\eta, \int \chi_2 d\eta\right\}$ dessa matriz:

$$\begin{aligned}\langle f_\eta^1, v_{\chi_0}\rangle &= \int \chi_0 d\eta \\ &= \sum_{j=0}^{2} \eta[j]\chi_0[j] \\ &= 1;\end{aligned}$$

$$\begin{aligned}\langle f_\eta^1, v_{\chi_1}\rangle &= \int \chi_1 d\eta \\ &= \sum_{j=0}^{2} \eta[j]\chi_1[j] \\ &= \eta[0] + \eta[1]e^{\frac{2\pi i}{3}} + \eta[2]e^{\frac{2\pi i 2}{3}} \\ &= \eta[0] + \eta[1]\cos\left(\frac{2\pi}{3}\right) + \eta[2]\cos\left(\frac{4\pi}{3}\right) + i\left(\eta[1]\sin\left(\frac{2\pi}{3}\right) + \eta[2]\sin\left(\frac{4\pi}{3}\right)\right) \\ &= \frac{1}{10} + \frac{7}{10}\cos\left(\frac{2\pi}{3}\right) + \frac{2}{10}\cos\left(\frac{4\pi}{3}\right) + i\left(\frac{7}{10}\sin\left(\frac{2\pi}{3}\right) + \frac{2}{10}\sin\left(\frac{4\pi}{3}\right)\right)\end{aligned}$$

$$\begin{aligned}\langle f_\eta^1, v_{\chi_2}\rangle &= \int \chi_2 d\eta \\ &= \sum_{j=0}^{2} \eta[j]\chi_2[j] \\ &= \eta[0] + \eta[1]e^{\frac{2\pi i 2}{3}} + \eta[2]e^{\frac{2\pi i}{3}} \\ &= \eta[0] + \eta[1]\cos\left(\frac{4\pi}{3}\right) + \eta[2]\cos\left(\frac{2\pi}{3}\right) + i\left(\eta[1]\sin\left(\frac{4\pi}{3}\right) + \eta[2]\sin\left(\frac{2\pi}{3}\right)\right) \\ &= \frac{1}{10} + \frac{7}{10}\cos\left(\frac{4\pi}{3}\right) + \frac{2}{10}\cos\left(\frac{2\pi}{3}\right) + i\left(\frac{7}{10}\sin\left(\frac{4\pi}{3}\right) + \frac{2}{10}\sin\left(\frac{2\pi}{3}\right)\right),\end{aligned}$$

associados, respectivamente, aos autovetores $\{v_{\chi_0}, v_{\chi_1}, v_{\chi_2}\}$.

Agora, vamos considerar a matriz A_η^2 associada a η relativa aos cilindros de tamanho 2 e perceber certas relações entre ela e a matriz A_η^1 descrita anteriormente. Temos que A_η^2 é a seguinte matriz bloco circulante:

$$100 \cdot A_\eta^2 = \left(\begin{array}{ccc|ccc|ccc} 1 & 3 & 7 & 2 & 4 & 14 & 7 & 14 & 49 \\ 7 & 1 & 3 & 14 & 2 & 4 & 49 & 7 & 14 \\ 3 & 7 & 1 & 4 & 14 & 2 & 14 & 49 & 7 \\ \hline 7 & 14 & 49 & 1 & 3 & 7 & 2 & 4 & 14 \\ 49 & 7 & 14 & 7 & 1 & 3 & 14 & 2 & 4 \\ 14 & 49 & 7 & 3 & 7 & 1 & 4 & 14 & 2 \\ \hline 2 & 4 & 14 & 7 & 14 & 49 & 1 & 3 & 7 \\ 14 & 2 & 4 & 49 & 7 & 14 & 7 & 1 & 3 \\ 3 & 7 & 1 & 14 & 49 & 7 & 3 & 7 & 1 \end{array}\right).$$

Dividimos a matriz em blocos para facilitar sua leitura. Não calcularemos os nove autovalores e autovetores dessa matriz, já que fizemos isso no Teorema 4.4 para todos os casos. Note, porém, que os autovalores

$$\{\langle f_\eta^1, v_{\chi_0}\rangle; \langle f_\eta^1, v_{\chi_1}\rangle; \langle f_\eta^1, v_{\chi_2}\rangle\}$$

da matriz A_η^1 também o são da matriz A_η^2, simplesmente pelo fato de que todo caracter constante em cilindros de tamanho 1 é automaticamente constante em cilindros de tamanho 2.

Capítulo 5

Considerações finais

Conforme pudemos ver ao longo desse trabalho, apesar de a convolução de duas medidas de probabilidade de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ depender apenas da estrutura de grupo para ser definida, as consequências e implicações dessa operação transcendem o aspecto algébrico. Por exemplo, no aspecto ergódico, vimos que a convolução preserva as medidas de probabilidade σ-invariantes, as misturadoras e as fracamente misturadoras (Proposição 1.12 e Observação 1.13). Podemos também citar Lindenstrauss, que em [10] mostrou que uma sequência η_n de medidas de probabilidade p-invariantes (e ergódicas) no círculo é tal que $\eta_n * \ldots * \eta_1$ converge, na topologia fraca* (e em $\bar{d}$), para medida de Lebesgue no círculo, sempre que a sequência satisfizer uma condição especial em relação às entropias (em particular, toda sequência η_n de medidas ergódicas que possuir entropia uniformemente afastada de zero satisfaz tal condição). Nosso Teorema 2.27 é uma versão desse resultado para o grupo $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$. Assim, a convolução de medidas traz consequências algébricas, dinâmicas, topológicas, ergódicas e etc. que pudemos constatar ao longo desse trabalho e elas, ainda, induzem-nos a fazer conjecturas. E é justamente de possíveis trabalhos futuros que trataremos nos parágrafos que seguem.

Apesar de termos estabelecido convergência em nível de topologia fraca* com o Teorema 2.27, não obtivemos resultados expressivos sobre crescimento de entropia. No Teorema 3.6 demonstramos apenas que a entropia da convolução de duas medidas de probabilidade σ-invariantes em $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ é maior ou igual ao máximo das duas prévias

entropias. Nossa conjectura é que no caso em que p é primo e se tenha uma sequência de medidas de probabilidade σ-invariantes η_n de $(\mathbb{Z}/p\mathbb{Z})^{\mathbb{N}}$ tal que $\inf_n h_{\eta_n}(\sigma) > 0$, poderemos concluir que $h_{\eta_n * \ldots * \eta_1}(\sigma) \to \log p$. Talvez um aprofundamento sobre a noção de *joinings* (muito bem trabalhada, por exemplo, no sexto capítulo de [6]) venha a calhar nessas questões que tratam de aumento estrito da entropia.

Há ainda questões um pouco mais periféricas, porém não desimportantes, que referem-se às convoluções. Por exemplo, iniciamos no capítulo terceiro um estudo sobre a equação $\eta * \mu = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$ e encontramos soluções não triviais para $p = 5$ e para qualquer p não primo (Proposição 3.12, Observação 3.13 e Teorema 3.14). No caso, solução trivial é toda solução em que $\eta = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$ ou $\mu = \left(\frac{1}{p}, \ldots, \frac{1}{p}\right)$. Apesar desses avanços, estamos longe de caracterizar completamente as soluções não triviais desse tipo de equação. Uma futura investigação poderia se dar no sentido de detalhar as diferenças entre as soluções no caso em que p é ou não primo, de perceber como a entropia influencia e etc.

Em resumo, a convolução de medidas é uma operação um tanto surpreendente, já que uma análise mais aprofundada de medidas resultantes da convolução de outras duas pode revelar, dentre outras consequências dinâmicas e ergódicas, aumento de entropia e encurtamento da distância em relação à medida de Bernoulli uniforme em algumas métricas.

Bibliografia

[1] Asaf, Katz: *Generalizations of Furstenberg's Diophantine Result.* Ergodic Theory and Dynamical Systems, 2016.

[2] Berg, Kenneth R: *Convolution of invariant measures, maximal entropy.* Theory of Computing Systems, 3(2):146–150, 1969.

[3] Cohen, Paul J: *On a conjecture of Littlewood and idempotent measures.* American Journal of Mathematics, 82(2):191–212, 1960.

[4] Dugundji, James: *Topology.* Allyn and Bacon, 1966.

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

[6] Glasner, Eli: *Ergodic theory via joinings.* Número 101 em *Mathematical Surveys and Monographs.* American Mathematical Soc., 2003.

[7] Itô, Takashi e Ichiro Amemiya: *A simple proof of the theorem of PJ Cohen.* Bulletin of the American Mathematical Society, 70(6):774–776, 1964.

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

[9] Kelley, John L: *General topology.* Springer Science & Business Media, 1975.

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

[11] Lopes, Artur O.: *An introduction to Coupling.* Springer, 2016.

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

[13] Lyons, Russell: *On measures simultaneously 2-and 3-invariant.* Israel Journal of Mathematics, 61(2):219–224, 1988.

[14] Rider, Daniel: *Central idempotent measures on SIN groups.* Duke Math. J, 38:187–191, 1971.

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

[16] Rudolph, Daniel J: *Fundamentals of measurable dynamics.* The Clarendon Press, copublished in the United States with John Wiley &, 1991.

[17] Tee, Garry J: *Eigenvectors of block circulant and alternating circulant matrices.* New Zealand Journal of Mathematics, 36:195–211, 2007.

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

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