



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

JÔNATAS MARINHO SANTOS JÚNIOR

TEORIA ERGÓDICA DE RAMSEY E A CONJECTURA $B+B+t$

FORTALEZA

2025

JÔNATAS MARINHO SANTOS JÚNIOR

TEORIA ERGÓDICA DE RAMSEY E A CONJECTURA $B + B + t$

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

Orientador: Prof. Dr. Yuri Gomes Lima.

FORTALEZA

2025

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

S235t Santos Júnior, Jônatas Marinho.

Teoria ergódica de Ramsey e a conjectura $B + B + t$ / Jônatas Marinho Santos Júnior. – 2025.
84 f.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Matemática, Fortaleza, 2025.

Orientação: Prof. Dr. Yuri Gomes Lima.

1. Teoria ergódica de Ramsey. 2. Fator de Kronecker. 3. Teorema de Roth. 4. Teorema de Hindman. 5. Conjectura $B + B + t$. I. Título.

CDD 510

JÔNATAS MARINHO SANTOS JÚNIOR

TEORIA ERGÓDICA DE RAMSEY E A CONJECTURA $B + B + t$

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

Aprovada em: 20 de janeiro de 2025.

BANCA EXAMINADORA

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

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

Prof. Dr. Alexander Arbieto
Universidade Federal do Rio de Janeiro (UFRJ)

À minha família, atual e futura.

AGRADECIMENTOS

Esta dissertação não teria sido finalizada sem o apoio de muitas pessoas.

Gostaria de agradecer a Deus por ter me sustentado durante os anos de mestrado. Quero agradecer também aos meus pais Jônatas e Joseane pelo apoio recorrente e aos meus irmãos.

Gostaria de agradecer também ao corpo acadêmico da UFC, em particular ao meu orientador Dr. Yuri Lima, Dr. Mauricio Poletti, e aos amigos que fiz nesta fase.

E um agradecimento especial à Talita, por estar sempre presente ao longo destes últimos anos.

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

“Se você soubesse resolver todos os seus problemas, eles não seriam problemas”

(Talita e Jônatas)

RESUMO

Nesta dissertação, exploramos a relação entre duas áreas distintas da Matemática, a Teoria Ergódica e a Teoria de Ramsey. A interação destas áreas é atualmente conhecida como *Teoria Ergódica de Ramsey*. Apresentaremos a prova de dois resultados clássicos: o *Teorema de Roth*, onde utilizaremos ferramentas de teoria ergódica e teoria espectral, e o *Teorema de Hindman*, onde utilizaremos ultrafiltros da teoria de conjuntos. Por fim, estudamos a prova da *Conjectura $B + B + t$* de Erdős, que recentemente obtida utilizando Teoria Ergódica de Ramsey. Um elemento de relevante importância nesta dissertação, e que será utilizado nas provas do Teorema de Roth e da Conjectura $B + B + t$, é o *Fator de Kronecker*, que encapsula uma estrutura de grupo de rotação dentro de um sistema que preserva medida.

Palavras-chave: teoria ergódica de Ramsey; fator de Kronecker; teorema de Roth; teorema de Hindman; conjectura $B + B + t$.

ABSTRACT

In this work, we explore the relation between two distinct areas of Mathematics, namely Ergodic Theory and Ramsey Theory. The interaction between these two areas is nowadays called *Ergodic Ramsey Theory*. We will present the proof of two classical results: *Roth theorem*, where we will use tools from ergodic theory and spectral theory, and *Hindman theorem*, where we will use ultrafilters from set theory. Finally, we study the proof of Erdős $B + B + t$ Conjecture, that was recently obtained using Ergodic Ramsey Theory. An element of great importance in this dissertation, that will be used both in the proof of Roth theorem and in the proof of $B + B + t$ Conjecture, is the *Kronecker factor*, that captures an structure of rotation group inside a system that preserves measure.

Keywords: ergodic Ramsey theory; Kronecker factor; Roth theorem; Hindman theorem. $B + B + t$ conjecture.

SUMÁRIO

1	INTRODUÇÃO	9
2	PRELIMINARES DE TEORIA ERGÓDICA	11
2.1	Recorrência em teoria ergódica	11
2.2	Ergodicidade	16
2.3	Aspectos espectrais de teoria ergódica	18
2.4	Desintegração da medida associada a um fator	21
2.5	Fator de kronecker	22
3	TEORIA ERGÓDICA DE RAMSEY	28
3.1	Elementos de teoria de ramsey	28
3.2	Princípio de correspondência de furstenberg	32
3.3	Teorema de roth	36
4	TEOREMA DE HINDMAN	44
4.1	Ultrafiltros	44
4.2	Construção de p -lim a partir de ultrafiltro	48
4.3	Compactificação de stone-čech	50
4.4	Elementos idempotentes e o teorema de hindman	51
5	A CONJECTURA $B + B + t$	58
5.1	Progressões de erdős	59
5.2	Medidas sobre progressões de erdős	62
5.3	Suporte da medida	65
5.4	Conclusão da prova do teorema	70
5.5	Comentários e consequências da prova da conjectura	71
6	CONCLUSÃO	73
	REFERÊNCIAS	74
	APÊNDICE A - TEORIA DA MEDIDA PARA GRUPOS	75
	APÊNDICE B - APLICAÇÃO SUPORTE	79
	APÊNDICE C - SUPORTE DE UMA DESINTEGRAÇÃO	82

1 INTRODUÇÃO

Ao longo das últimas décadas, a Matemática tem presenciado uma enorme interação entre a Teoria Ergódica e outras áreas, como a Combinatória e a Teoria dos Números. Esta frutífera interrelação tem demonstrado o poder da interdisciplinaridade, com a obtenção de profundos resultados. Nesta dissertação, apresentamos resultados da Teoria Ergódica que possuem aplicações diretas na Teoria de Ramsey. Esta área de pesquisa é hoje conhecida por *Teoria Ergódica de Ramsey*. Mais especificamente, discutimos dois resultados clássicos da área e um resultado extremamente recente, provado em 2022, conhecido como a *Conjectura $B + B + t$ de Erdős*.

No Capítulo 2, fornecemos os pré-requisitos básicos de Teoria Ergódica necessários à compreensão do restante do texto. Destacamos duas seções importantes. Na primeira, Seção 2.1, introduzimos noções importantes como conjuntos IP e conjuntos sindéticos. Mostramos que conjuntos dinamicamente relevantes possuem tal estrutura, exibindo assim a riqueza combinatória obtida pela recorrência. Na segunda, Seção 2.5, introduzimos o *fator de Kronecker* de um sistema que preserva medida, que pode ser visto como sua maior sub-estrutura de grupo de rotação. O fator de Kronecker também pode ser entendido como a parte estruturada do sistema. Esta ferramenta será utilizada de maneira decisiva na Seção 3.3 e no Capítulo 5.

No Capítulo 3, apresentamos elementos de Teoria de Ramsey e um resultado que permite traduzir problemas desta área para problemas de recorrência em Teoria Ergódica. Este resultado é conhecido como *Princípio de Correspondência de Furstenberg*. Em Teoria de Ramsey, destacamos a noção de densidade positiva e mencionamos teoremas clássicos da área, conhecidos como Teorema de Van der Waerden e Teorema de Szemerédi. Este último, provado em 1975 por Endre Szemerédi utilizando ferramentas puramente combinatórias, foi reprovado em 1977 por Hillel Furstenberg utilizando teoria ergódica, no que é considerado o marco inaugural da Teoria Ergódica de Ramsey. A prova de Furstenberg é extremamente profunda e delicada, e vai além do escopo desta dissertação. De todo modo, a fim de demonstrar o poder da técnica, provamos o Teorema de Roth, que representa o caso inicial ($n = 3$) do Teorema de Szemerédi. De fato, provamos sua versão ergódica, representada pelo Teorema 3.3.1. A ideia principal consiste em decompor funções L^2 integráveis em duas componentes, uma “estruturada” e outra “aleatória”. Uma vez provada a versão ergódica, aplicamos o Princípio de Correspondência de Furstenberg

para obter o Teorema de Roth original.

No Capítulo 4, continuando o estudo de Teoria de Ramsey, estudamos o Teorema de Hindman sobre conjuntos IP em partições de \mathbb{N} . Apresentamos uma prova deste teorema que utiliza a teoria de ultrafiltros. Para este fim, fornecemos todas as definições relevantes de ultrafiltros e o fato de que o conjunto de ultrafiltros é um semigrupo compacto e Hausdorff. Em posse deste fato, podemos aplicar um resultado de Dinâmica Topológica, chamado *Teorema de Ellis*, sobre a existência de elementos idempotentes. Por fim, notamos que a estrutura idempotente de um tal elemento induz uma estrutura recorrente de conjunto IP, provando assim o Teorema de Hindman.

O último capítulo desta dissertação, Capítulo 5, fornece uma prova da *Conjectura $B+B+t$ de Erdős*. Esta conjectura, proposta por Erdős na década de 1970, reúne os Teoremas de Szemerédi e Hindman sob a pergunta: *existem transladados de conjuntos de soma $B+B$ contidos em um conjunto de densidade positiva?* A obtenção deste resultado, realizada em 2022 por Bryna Kra, Joel Moreira, Florian Richter e Donald Robertson, é uma prova em Teoria Ergódica de Ramsey. Neste contexto, formulamos o problema dinâmico associado e mostramos como ele implica, via um princípio de correspondência, a conjectura original, mostrando assim que a Teoria Ergódica de Ramsey permanece sendo uma área moderna e relevante. Aqui, novamente utilizamos de modo essencial o fator de Kronecker, que é utilizado para definir medidas que identificam as chamadas *progressões de Erdős*.

Os apêndices fornecem os fatos que dão suporte ao Capítulo 5.

2 PRELIMINARES DE TEORIA ERGÓDICA

Em teoria ergódica, usamos técnicas de teoria da medida para estudar a dinâmica de aplicações mensuráveis que preservam a medida. Nesta seção, nós fornecemos a base necessária desta área para a leitura do trabalho. Por parte do leitor, assumiremos uma base de teoria da medida, e portanto, a menos que não seja um fato conhecido, não trataremos aqui.

No que se segue, (X, β_X, μ) denota um espaço de probabilidade, isto é, X é um conjunto, β_X é uma σ -álgebra sobre X e μ é uma medida de probabilidade.

Definição 2.0.1 (Pushforward de uma medida). *Sejam (X, β_X, μ) um espaço de probabilidade, (Y, β_Y) um espaço de medida e $T : X \rightarrow Y$ uma aplicação mensurável. Para $B \in \beta_Y$, seja*

$$T_*(\mu)(B) = \mu(T^{-1}(B)).$$

A função $T_*(\mu) : \beta_Y \rightarrow [0, 1]$ é uma medida de probabilidade, e a chamamos de pushforward de μ .

Seja (X, β_X, μ) um espaço de medida e $T : X \rightarrow X$ uma aplicação mensurável. Dizemos que T *preserva a medida* μ , ou que a *medida* μ é T -*invariante*, quando $T_*(\mu) = \mu$. Neste caso, denotamos por (X, β_X, μ, T) quando isto acontece, e o chamaremos de sistema que preserva medida. Este é o contexto clássico da Teoria Ergódica.

2.1 Recorrência em teoria ergódica

Como primeira aplicação do estudo da dinâmica sob a ótica da medida, estudamos o seguinte problema de recorrência:

Dado $A \in \beta_X$, quando um ponto $x \in A$ retorna para A pela dinâmica de T ? Se sim, quão frequente é este retorno?

Responderemos à segunda pergunta quando falarmos de ergodicidade mais à frente. Antes, estudaremos a primeira pergunta. Em relação a ela, temos o seguinte resultado.

Teorema 2.1.1 (Teorema de Recorrência de Poincaré). *Seja (X, β_X, μ, T) um sistema que preserva medida, onde μ é medida finita*

$$\mu\{x \in A : T^n x \notin A, \forall n \in \mathbb{N}\} = 0.$$

Além do mais,

$$\mu(\{x \in A : T^n x \in A \text{ para uma quantidade finita de } n \in \mathbb{N}\}) = 0.$$

Demonstração. Seja $A_0 = \{x \in A : x \text{ não retorna a } A\}$. É claro que $A_0 = A \setminus \bigcup_{n=1}^{\infty} T^{-n} A$, que é mensurável. Logo, caso $\mu(A) = 0$, então $\mu(A_0) = 0$, pois $A_0 \subset A$.

Assuma que $\mu(A) > 0$. Afirmamos que $\mu(A_0) = 0$. De fato, suponha que $\mu(A_0) > 0$. Note que $\{T^k A_0\}_{k \geq 1}$ são disjuntos dois a dois: caso contrário, existem $k, l \in \mathbb{N}$ tais que $T^{-k} A_0 \cap T^{-k-l} A_0 \neq \emptyset$. Logo, existe $x \in X$ tal que $T^k x \in A_0$, e $T^l(T^k(x)) \in A_0$. Logo, $T^k x \in A$ retorna para A , o que contradiz a definição de A_0 . Desta forma, por serem disjuntos, e pela medida ser finita,

$$\mu(X) \geq \mu\left(\bigcup_{k=1}^{\infty} T^{-k} A_0\right) = \sum_{k=1}^{\infty} \mu(T^{-k} A_0) = \sum_{k=1}^{\infty} \mu(A_0) = \infty,$$

contradição, o que prova que $\mu(A_0) = 0$.

Agora, note que $x \in A$ retorna uma quantidade finita de vezes a A quando $T^k x \in A_0$ para algum $k \geq 1$ ou equivalentemente $x \in T^{-k} A_0$. Desta forma, $\{x \in A : T^n x \in A \text{ para uma quantidade finita de } n \in \mathbb{N}\} = \bigcup_{k=1}^{\infty} T^{-k} A_0$. Como $\mu(A_0) = 0$, temos que

$$0 \leq \mu\left(\bigcup_{k=1}^{\infty} T^{-k} A_0\right) \leq \sum_{k=1}^{\infty} \mu(T^{-k} A_0) = \sum_{k=1}^{\infty} \mu(A_0) = 0,$$

o que prova o que queremos. □

Usando a mesma notação da prova anterior, note que

$$A = A_0 \cup A \cap \left(\bigcup_{k=1}^{\infty} T^{-k} A\right) = A_0 \cup \left(\bigcup_{k=1}^{\infty} (A \cap T^{-k} A)\right),$$

donde

$$\mu(A) \leq \mu(A_0) + \mu\left(\bigcup_{k=1}^{\infty} (A \cap T^{-k} A)\right) = \mu\left(\bigcup_{k=1}^{\infty} (A \cap T^{-k} A)\right).$$

Desta forma, temos o seguinte corolário.

Corolário 2.1.1. *Seja (X, β_X, μ, T) um sistema que preserva medida e $A \in \beta_X$ tal que $\mu(A) > 0$. Então existe $n \in \mathbb{N}$ tal que $\mu(A \cap T^{-n}A) > 0$.*

Demonstração. Ora, note que

$$\sum_{k=1}^{\infty} \mu(A \cap T^{-k}A) \geq \mu\left(\bigcup_{k=1}^{\infty} (A \cap T^{-k}A)\right) \geq \mu(A) > 0,$$

e daí algum termo $\mu(A \cap T^{-n}A)$ deve ser maior que zero. \square

Com base no teorema acima, μ -quase todo ponto de $A \in \beta_A$ retorna a A infinitas vezes. Apesar disso, em um primeiro momento, não podemos afirmar nada sobre a frequência de retorno. Motivados por esta incógnita, somos levados a estudar o conjunto $\{n \in \mathbb{N} : \mu(A \cap T^{-n}A) > 0\}$, onde assumimos $\mu(A) > 0$. Para este fim, desenvolvemos um estudo de subconjuntos especiais dos números naturais: conjuntos IP e conjuntos sindéticos.

Definição 2.1.1 (Conjunto IP). *Seja $\{n_k\}_k \subset \mathbb{N}$ uma sequência crescente. Chamamos de $FS(n_k)_k$ ao conjunto de todas as somas finitas de termos distintos da sequência $(n_k)_k$. Dizemos que $A \subset \mathbb{N}$ é um conjunto IP quando existe uma sequência $(n_k)_k$ tal que $A = FS(n_k)_k$.*

Exemplos.

1. Claramente, \mathbb{N} é um conjunto IP: basta tomar $n_k = k$, para todo $k \in \mathbb{N}$.
2. A sequência que torna um conjunto como conjunto IP não é única: de fato, tome $(2^k)_{k \in \mathbb{N}}$. Como todo número natural na base 2 escreve-se como uma soma finita de potência de 2, temos $FS(2^k)_k = \mathbb{N}$.
3. Note que $FS(10^k)_k =$ números naturais cujos dígitos são 0 e 1 na base decimal. De maneira análoga, dado $1 \leq d \leq 9$, $FS(d \cdot 10^k)_k =$ números naturais cujos dígitos são 0 e d na base decimal.
4. Todo conjunto IP contém números pares: de fato, dada $(n_k)_k$ uma sequência crescente de números naturais, pelo Princípio da Casa dos Pombos, possui uma subsequência com mesma paridade. Deste modo, a soma deles é divisível por dois. Deste modo, o conjunto dos números ímpares não é IP.

Definição 2.1.2. *Dizemos que $A \subset \mathbb{N}$ é IP^* se $A \cap I \neq \emptyset$ para todo subconjunto IP $I \subset \mathbb{N}$.*

Exemplos.

1. \mathbb{N} é claramente IP^* .
2. Progressões aritméticas são exemplos de conjuntos IP^* : de fato, seja $k \in \mathbb{N}$ e $A = \{kn : n \in \mathbb{N}\}$. Dado $B = FS(n_s)_s$ um conjunto IP , mais uma vez pelo Princípio da Casa dos Pombos, existe uma subsequência $(n_{s_l})_l$ que possui o mesmo resto r módulo k . Logo, $n_{s_1} + \dots + n_{s_k} \in B$ deixa resto 0 módulo k , e portanto escreve-se da forma nk . Isto é, $A \cap B \neq \emptyset$. Pela arbitrariedade da escolha de B , concluímos que A é IP^* .

Definição 2.1.3 (Conjunto sindético). *Dizemos que $A = \{a_1 < a_2 < \dots\} \subset \mathbb{N}$ é sindético quando existe $c \in \mathbb{R}$, $c > 0$, tal que $a_{k+1} - a_k \leq c$.*

Exemplos.

1. Progressões aritméticas são exemplos de conjuntos sindéticos. Neste caso, note que a distância entre elementos consecutivos é constante.
2. $\{n^k\}_{k \in \mathbb{N}}$, $n \geq 2$, são exemplos de conjuntos que não são sindéticos: a distância entre elementos consecutivos é $n^{k+1} - n^k = n^k(n - 1)$, isto é, o crescimento é exponencial e não pode ser limitado.

Conjuntos sindéticos possuem buracos limitados. Todos estes tipos de subconjuntos que foram apresentados aqui são noções importantes que figuram como *estruturas bem organizadas* de \mathbb{N} , no sentido de Teoria de Ramsey. Em particular, conjuntos sindéticos são robustos, onde a noção de robustez é fornecida pela densidade de Banach, que no caso da definição acima, é maior que $1/c$. Retomaremos estas noções nos próximos dois capítulos.

Note que todos os exemplos de conjuntos IP^* são exemplos de conjuntos sindéticos. Isto não é coincidência, conforme o próximo teorema afirma.

Teorema 2.1.2. *Seja $A \subset \mathbb{N}$ um conjunto IP^* . Então A é sindético.*

Demonstração. Provaremos a afirmação contrapositiva: se A não é sindético, então existe um conjunto IP $F = FS(n_k)_k$ tal que $A \cap F = \emptyset$. Começamos tomando $n_1 \notin A$. Como A não é sindético, existe um intervalo $I_2 = [n_2, m_2]$ de tamanho maior que n_1 tal que $A \cap I_2 = \emptyset$ e $n_1 < n_2$. Segue que $n_1, n_2 \notin A$ e $n_1 + n_2 \notin A$ pois $n_1 + n_2 \in I_2$.

Por indução, suponha que construímos $n_1 < n_2 < \dots < n_k$ tais que $\sum_{i \in L} n_i \notin A$ para todo $L \subset \{1, \dots, k\}$ não-vazio. Como A não é sindético, existe $I_{k+1} = [n_{k+1}, m_{k+1}]$ com $n_{k+1} > n_k$ tal que $A \cap I_{k+1} = \emptyset$ e $m_{k+1} - n_{k+1} > n_1 + \dots + n_k$. Logo, para todo

$L = L' \cup \{k+1\} \subset \{1, \dots, k+1\}$ temos que $\sum_{i \in L} n_i \in I_{k+1}$ e portanto $\sum_{i \in L} n_i \notin A$. Isso conclui a construção indutiva de $F = FS(n_k)_k$ tal que $A \cap F = \emptyset$. \square

Como consequência, temos o teorema abaixo.

Teorema 2.1.3. *Seja (X, β_X, μ, T) um sistema que preserva medida e $A \in \beta_X$ tal que $\mu(A) > 0$. Então $\{n \in \mathbb{N} : \mu(A \cap T^{-n}A) > 0\}$ é sindético.*

Demonstração. Seja $L = \{n \in \mathbb{N} : \mu(A \cap T^{-n}A) > 0\}$. Pelo Teorema 2.1.2, basta mostrar que L é IP^* . Seja $(n_k)_{k \in \mathbb{N}} \subset \mathbb{N}$ uma sequência crescente de números naturais e seja $F = FS(n_k)_k$. Seja também $N_l = n_1 + n_2 + \dots + n_l$, que define uma sequência crescente. Afirmamos que existem $r > s$ tais que $\mu(T^{-N_r}A \cap T^{-N_s}A) > 0$. Caso contrário, teríamos $\mu(T^{-N_r}A \cap T^{-N_s}A) = 0$ para todos $r \neq s$ e daí

$$\infty > \mu\left(\bigcup_{l=1}^{\infty} T^{-N_l}A\right) = \sum_{l=1}^{\infty} \mu(T^{-N_l}A) = \sum_{l=1}^{\infty} \mu(A) = \infty,$$

uma contradição. Logo, existem $r > s$ tais que $\mu(T^{-N_r}A \cap T^{-N_s}A) > 0$ e daí $\mu(T^{-(N_r - N_s)}A \cap A) > 0$, donde $N_r - N_s \in L$. Como $N_r - N_s = n_{s+1} + \dots + n_r \in F$ segue que $F \cap L \neq \emptyset$. Pela arbitrariedade da escolha de F , concluímos que L é IP^* . \square

Em outras palavras, a frequência com a qual um ponto retorna a A é limitada.

Progressões aritméticas são exemplos de conjuntos sindéticos com buracos de tamanho uniforme. Neste contexto, é natural perguntar se o conjunto sindético $\{n \in \mathbb{N} : \mu(A \cap T^{-n}A) > 0\}$ possui progressões aritméticas. A resposta é positiva, e este resultado é conhecido como o Teorema de Recorrência Múltipla de Poincaré.

Teorema 2.1.4 (Teorema de Recorrência Múltipla de Poincaré). *Seja (X, β_X, μ, T) um sistema que preserva medida e $A \in \beta_X$ tal que $\mu(A) > 0$. Dado um inteiro positivo $n \in \mathbb{N}$, existe um inteiro positivo $k > 0$ tal que*

$$\mu\left(A \cap T^{-k}A \cap \dots \cap T^{k(n-1)}A\right) > 0.$$

Em outras palavras, existe um conjunto de medida positiva de A que retorna a A com tempos igualmente espaçados. O teorema acima, provado pelo matemático Hillel Furstenberg em 1977, introduziu métodos que foram usados para provar resultados importantes de combinatória com aparato de teoria ergódica, dentre eles diversos resultados

que não haviam sido obtidos utilizando técnicas puramente combinatórias. De fato, até alguns tempos atrás, a única prova conhecida para uma generalização do Teorema de Szemerédi em \mathbb{Z}^d fazia uso da teoria ergódica. Sobre a discussão do Teorema 2.1.4, referenciamos o próximo capítulo.

2.2 Ergodicidade

Na discussão de recorrência deste capítulo, nos perguntamos quais pontos retornavam a um conjunto mensurável fixado e quão frequente ele retorna a este conjunto. Respondemos à primeira pergunta, apresentando uma discussão sobre a estrutura de iterados que retornam a um conjunto dado. Agora, responderemos à segunda pergunta.

Para isto, seja (X, β_X, μ, T) um sistema que preserva medida. Dado $A \in \beta$, considere

$$\tau(A, x) = \lim_{N \rightarrow \infty} \frac{1}{N} |\{1 \leq n \leq N : T^n x \in A\}|$$

o *tempo médio de visita* de x em relação a A , quando o limite existe. Há uma forma conveniente de reescrever a definição acima.

Lembre-se que, dado $A \subset X$, a *função característica* de A é dada por $\mathbb{1}_A : X \rightarrow \mathbb{R}$ definida via a relação $\mathbb{1}_A(x) = 1$ caso $x \in A$ e $\mathbb{1}_A(x) = 0$, quando $x \notin A$. Deste modo,

$$\tau(A, x) = \frac{1}{N} \lim_{N \rightarrow \infty} \sum_{n=0}^{N-1} \mathbb{1}_A \circ T^n(x).$$

Isto nos motiva à seguinte definição.

Definição 2.2.1 (Média temporal). *Seja (X, β_X, μ, T) um sistema que preserva medida, e $f \in L^1_\mu(X)$. Seja $f \in L^1_\mu(X)$. Chamamos de *média temporal de f* a função \tilde{f} definida por*

$$\tilde{f}(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x)),$$

quando o limite existe.

O próximo teorema nos diz que esta função está sempre bem definida a menos de um conjunto de medida nula. Antes, precisamos de uma definição. Dado um espaço

de probabilidade (X, β_X, μ) , lembre-se que duas funções mensuráveis $f, g : X \rightarrow \mathbb{R}$ são identificadas se $f(x) = g(x)$ para μ -quase todo ponto $x \in X$. Neste contexto, $L_\mu^1(X)$ é o conjunto das classes de equivalência das funções $f : X \rightarrow \mathbb{R}$ com valor absoluto integrável.

Definição 2.2.2 (Conjuntos e funções invariantes). *Seja (X, β_X, μ, T) um sistema que preserva medida. Dizemos que uma função mensurável $f : X \rightarrow \mathbb{R}$ é T -invariante ou invariante por T quando $f \circ T = f$. Dizemos que $A \subset X$ é T -invariante ou invariante por T quando $T^{-1}(A) = A$.*

A igualdade de conjuntos é, de modo similar, considerada módulo μ .

Teorema 2.2.1 (Birkhoff). *Seja (X, β_X, μ, T) um sistema que preserva medida e $f \in L_\mu^1(X)$. Então o limite*

$$\tilde{f}(x) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k(x))$$

existe para μ -quase todo ponto $x \in X$. Ademais, $\tilde{f} \in L_\mu^1(X)$, \tilde{f} é T -invariante e

$$\int_X \tilde{f} d\mu = \int_X f d\mu.$$

Para uma prova do Teorema de Birkhoff, consulte (OLIVEIRA; VIANA, 2019, p. 68).

Definição 2.2.3 (Ergodicidade). *Seja (X, β_X, μ, T) um sistema que preserva medida. Dizemos que o sistema é ergódico quando toda $f : X \rightarrow \mathbb{R}$ em $L_\mu^1(X)$ invariante por T é uma função constante.*

A ergodicidade definida acima é um conceito central em Teoria Ergódica. Ela surge com a *Hipótese Ergódica de Boltzmann*, que afirma que a média temporal que uma partícula visita uma certa região é igual à média espacial desta região. A fim de apresentar uma caracterização alternativa de ergodicidade, definimos

Definição 2.2.4 (Conjunto invariante). *Seja (X, β_X, μ, T) um sistema que preserva medida. Dizemos que $A \in \beta_X$ é invariante por T ou T -invariante quando $\mu(T^{-1}A \Delta A) = 0$.*

Teorema 2.2.2. *Seja (X, β_X, μ, T) um sistema que preserva medida. As seguintes afirmações são equivalentes:*

- (a) O sistema é ergódico;
- (b) Para toda $f \in L^1_\mu(X)$, a função de média temporal \tilde{f} é constante e igual a $\int_X f d\mu$;
- (c) Todo subconjunto $A \in \beta_X$ T -invariante é tal que $\mu(A) = 0$ ou $\mu(A) = 1$.

2.3 Aspectos espectrais de teoria ergódica

Seja (X, β_X, μ, T) um sistema que preserva medida. É comum estudarmos tais sistemas em seus aspectos espectrais, no sentido que introduzimos abaixo. No que segue, denotamos por $L^2_\mu(X)$ o conjunto das funções mensuráveis $f : X \rightarrow \mathbb{C}$ que são L^2 -integráveis, isto é, as funções tais que

$$\|f\|_2 := \left(\int_X f(x) \overline{f(x)} d\mu(x) \right)^{\frac{1}{2}} < +\infty,$$

onde \bar{z} é o conjugado do número complexo z . Como sabemos de Análise Funcional, $L^2_\mu(X)$ é um espaço vetorial de dimensão infinita sempre que X é infinito, que admite o produto interno

$$\langle f, g \rangle = \int_X f(x) \overline{g(x)} d\mu(x),$$

que faz de $L^2_\mu(X)$ um espaço de Hilbert.

Definição 2.3.1 (Operador de Koopman). *Seja (X, β_X, μ, T) um sistema que preserva medida. Chamamos de operador de Koopman o operador $U_T : L^2_\mu(X) \rightarrow L^2_\mu(X)$ dada por $U_T(f) = f \circ T$.*

Note que, como T preserva medida, $f \circ T \in L^p_\mu(X)$ para todo $f \in L^p_\mu(X)$, $p \in [1, +\infty]$, e em particular para $p = 2$.

Como o sistema (X, β_X, μ, T) preserva medida, vale

$$\langle U_T(f), U_T(g) \rangle = \int_X f \circ T \cdot \overline{g \circ T} d\mu = \int_X f \circ \overline{g} d\mu = \langle f, g \rangle.$$

Portanto, o operador U_T é unitário. A recíproca é verdadeira: se U_T é unitário e μ é uma medida de probabilidade, então T preserva a medida. De fato, neste caso, seja $f \in L^2_\mu(X)$ qualquer e $g = \mathbf{1}_X$ a função indicadora de X . Temos

$$\int_X f d\mu = \langle f, \mathbf{1}_X \rangle = \langle f \circ T, \mathbf{1}_X \circ T \rangle = \langle f \circ T, \mathbf{1}_X \rangle = \int_X f \circ T d\mu,$$

o que prova a recíproca.

Em Teoria Ergódica, temos algumas formas de decompor sistemas que preservam a medida, e uma delas é o que conhecemos por fatores. Fatores podem ser interpretados como “subsistemas” de um sistema que preserva medida e, dentre as suas aplicações, podem ser utilizados para estudar o sistema original. Antes de apresentar as suas aplicações, vamos à sua definição formal.

Definição 2.3.2 (Fator). *Sejam (X, β_X, μ, T) e (Y, β_Y, ν, S) sistemas que preservam medida. Dizemos que uma aplicação mensurável $\varphi : X \rightarrow Y$ é uma aplicação de fator quando*

- $\varphi \circ T = S \circ \varphi$ em ν -quase todo ponto,
- $\varphi_* \mu = \nu$.

Neste caso, dizemos que (Y, β_Y, ν, S) é fator (mensurável) de (X, β_X, μ, T) .

Uma observação importante sobre fatores é que eles são naturalmente identificados com sub- σ -álgebras invariantes do sistema. Isto será explicado e aplicado na próxima seção.

Podemos ver a aplicação de fator do ponto de vista espectral. Dados (X, β_X, μ, T) e (Y, β_Y, ν, S) sistemas que preservam a medida e $\varphi : X \rightarrow Y$ uma aplicação de fator, definimos o *pullback* de φ como a transformação $\varphi^* : L^2_\nu(Y) \rightarrow L^2_\mu(X)$ definida por $\varphi^*(g) = g \circ \varphi$.

Proposição 2.3.1. *Sejam (X, β_X, μ, T) e (Y, β_Y, ν, S) sistemas que preservam a medida e φ^* como acima. Então:*

- (a) φ^* está bem-definida e é uma isometria linear.
- (b) $\varphi^* \circ U_S = U_T \circ \varphi^*$.

Demonstração. (a) Note que, uma vez verificada a boa definição de φ^* , ela é linear: dados $f, g \in L^2_\nu(Y)$ e $\alpha \in \mathbb{C}$, temos que $\varphi^*(f + \alpha g)(x) = (f + \alpha g)(\varphi(x)) = f(\varphi(x)) + \alpha g(\varphi(x)) = \varphi^* f(x) + \alpha \varphi^* g(x)$, para todo $x \in X$, e daí fica verificado a linearidade.

Para começar, dado $B \in \beta_Y$, vale $\mathbb{1}_{\varphi^{-1}(B)} = \varphi^* \mathbb{1}_B$. Por outro lado, que a propriedade $\varphi_* \mu = \nu$ significa que $\nu(B) = \mu(\varphi^{-1}(B))$ para todo $B \in \beta_Y$. Traduzindo em termos de integral, temos que

$$\|\mathbb{1}_B\|_2^2 \int_Y \mathbb{1}_B^2 d\nu = \int_X \mathbb{1}_{\varphi^{-1}(B)}^2 d\mu = \int_X \varphi^* \mathbb{1}_B d\mu^2 = \|\varphi^*\|_2^2.$$

Logo, $\varphi^* \mathbf{1}_B \in L_\mu^2(X)$. Por outro lado, seja $f = \sum_{k=1}^n a_k \mathbf{1}_{B_k} \in L_\nu^2(Y)$ uma função simples, com $B_i \cap B_j = \emptyset$ para $i \neq j$. Logo, por serem disjuntos dois a dois, temos que $|f| = \sum_{k=1}^n |a_k| \mathbf{1}_{B_k}$. É claro que $\mathbf{1}_{A \cap B} = \mathbf{1}_A \cdot \mathbf{1}_B$, para todo A, B subconjuntos, segue do fato de B_1, \dots, B_n serem dois a dois disjuntos que

$$|f|^2 = \sum_{k,j=1}^n |a_k a_j| \mathbf{1}_{B_k} \mathbf{1}_{B_j} = \sum_{k=1}^n |a_k|^2 \mathbf{1}_{B_k}.$$

Logo, pelo o que provamos para funções características,

$$\begin{aligned} \int_Y |f|^2 d\nu &= \int_X |f| \mathbf{1}_{B_k} \mathbf{1}_{B_j} \\ &= \sum_{k=1}^n |a_k|^2 \nu(B_k) \\ &= \sum_{k=1}^n |a_k|^2 \mu(\varphi^{-1}(B_k)) \\ &= \int_Y |\varphi^* f|^2 d\mu, \end{aligned}$$

e daí φ está bem definida e é isometria para funções simples. Por fim seja $g \in L_\nu^2(Y)$ e $(g_n) \subset L_\nu^2(Y)$ sequências de funções simples tal que $|g_n| \leq |g|$ e que $g_n \rightarrow g$ em $L_\nu^2(Y)$. Isto é, $(g_n)_n$ é de Cauchy. Como φ é isometria nas funções simples, temos que $(\varphi^* g_n)_n \subset L_\mu^2(X)$ é de Cauchy, donde pela completude de $L_\mu^2(X)$ a sequência converge para algum f . Afirimo que $f = \varphi^* g$ e que $\|f\|_2 = \|g\|_2$.

Para a primeira parte: sabemos de Teoria da Medida que existe uma subsequência (g_{n_k}) que converge pontualmente a g em ν -quase todo ponto: existe $M \subset Y$, $\nu(M) = 1$, tal que $\lim_{n \rightarrow \infty} g_n(x) = g(x)$ para todo $x \in M$. Daí, $\varphi^* g_n$ converge pontualmente a $\varphi^* g$ em μ -quase todo ponto: de fato, note que $\mu(\varphi^{-1}(M)) = \nu(M) = 1$, e que para $x \in \varphi^{-1}(M)$, $\varphi(x) \in M$ e vale $\lim_{n \rightarrow \infty} g_n(\varphi(x)) = g(\varphi(x))$.

Ora, por outro lado, temos que $(g_{n_k} \circ \varphi)_k$ continua convergindo a f em $L_\mu^2(X)$, donde a menos de tomar subsequência, concluimos também que $(g_{n_k})_k$ converge pontualmente a f em μ -quase todo ponto. Daí, segue do parágrafo anterior que $f = \varphi^* g$, isto é, $\varphi^* g \in L_\mu^2(X)$.

Por fim, para verificar que é isometria, basta ver que

$$\begin{aligned} \int_X |\varphi^* g|^2 d\mu &= \lim_n \int \int_X |\varphi^* g_n|^2 d\mu \\ &= \int_Y |g_n|^2 d\nu \\ &= \int_Y |g|^2 d\nu, \end{aligned}$$

donde $\|\varphi^*g\|_2 = \|g\|_2$. E isto conclui a prova do item (a).

(b) Lembre-se que $T \circ \varphi = \varphi \circ S$. Desta forma, temos que dado $g \in L^2_\nu(Y)$,

$$\varphi^*(U_S(g)) = \varphi^*(g \circ S) = g \circ (S \circ \varphi) = g \circ (\varphi \circ T) = U_T(\varphi^*g),$$

o que verifica o item (b). □

Com base no que acabamos de provar, pelo item (a), φ^* é um mergulho isométrico, e daí podemos identificar $L^2_\nu(Y)$ com $\varphi^*(L^2_\nu(Y)) \subset L^2_\mu(X)$. Pelo item (b), $\varphi^*(L^2_\nu(Y))$ é U_T -invariante. Desta forma, a nível espectral, fatores traduzem-se em subespaços fechados de $L^2_\mu(X)$ invariantes por U_T .

Destacamos a palavra “mensurável” para atentar que há também a noção de aplicação de fator e fatores para sistemas topológicos. Para isto, chamaremos de sistema topológico ou simplesmente sistema quando estiver claro pelo contexto ao par (X, T) , onde X é um espaço métrico compacto e $T : X \rightarrow X$ é contínua. Diremos que o sistema topológico é *invertível* quando T for invertível.

Definição 2.3.3 (Fatores contínuos). *Sejam (X, T) e (Y, S) sistemas topológicos. Dizemos que uma aplicação contínua $\pi : X \rightarrow Y$ é uma aplicação de fator contínua quando for sobrejetiva e $\pi \circ T = S \circ \pi$. Neste caso, dizemos que (Y, S) é fator do sistema (X, T) .*

Esta distinção será importante no capítulo em que discutiremos sobre a conjectura $B + B + t$.

2.4 Desintegração da medida associada a um fator

Com a ideia de que fatores são “subsistemas” do sistema original, podemos decompor um sistema com relação a um fator, de modo a guardar informações dinâmicas e ergódicas relevantes. Esta é a noção de desintegração de uma medida.

Seja X um espaço métrico compacto e seja um (X, β_X, μ, T) um sistema mensurável. Denote por $\mathcal{M}(X)$ o conjunto das medidas de probabilidade de Borel em X equipadas com a topologia fraca-*. Naturalmente, equipamos $\mathcal{M}(X)$ com a σ -álgebra de Borel associada a esta topologia.

Teorema 2.4.1 (Desintegração de fatores). *Dados (X, β_X, μ, T) e (Y, β_Y, ν, S) sistemas mensuráveis e $\pi : X \rightarrow Y$ um fator mensurável, existe uma função $y \in Y \mapsto \mu_y \in \mathcal{M}(X)$ mensurável, definida em um conjunto de medida total em Y , tal que:*

- (i) *Para toda função limitada e mensurável $f : X \rightarrow \mathbb{R}$ a função $y \in Y \mapsto \int_X f d\mu_y$ está definida em ν -quase todo ponto e é mensurável em Y , satisfazendo*

$$\int_D \left(\int_X f d\mu_y \right) d\nu = \int_{\pi^{-1}(D)} f d\mu,$$

para todo $D \subset Y$ mensurável.

- (ii) *Para ν -quase todo $y \in Y$, temos $\mu_y(\pi^{-1}(y)) = 1$.*

- (iii) *A propriedade (i) determina unicamente a aplicação $y \mapsto \mu_y$, isto é, $y \mapsto \mu'_y$ é outra aplicação mensurável satisfazendo (i), então $\mu_y = \mu'_y$ para ν -quase todo $y \in Y$.*

- (iv) *Input dinâmico: $T_*\mu_y = \mu_{S(y)}$ para ν -quase todo $y \in Y$.*

A prova deste teorema pode ser encontrada em (??, Capítulo 5). A aplicação mensurável $y \mapsto \mu_y$ satisfazendo o teorema acima é chamada de *desintegração* de μ com respeito ao fator π .

2.5 Fator de Kronecker

Nesta seção, introduziremos um fator de grande importância, chamado *fator de Kronecker*.

Definição 2.5.1 (Grupo de Rotação). *Um grupo de rotação é um sistema da forma (Z, β_Z, m, R) onde $Z = (Z, +)$ é um grupo abeliano compacto, m é a medida de Haar de Z e $R : Z \rightarrow Z$ é uma rotação, isto é, uma aplicação da forma $R(z) = z + \alpha$, onde $\alpha \in Z$. Quando (Z, β_Z, m, R) é um grupo de rotação, escolhemos uma métrica em Z tal que para todo $z \in Z$ a transformação $z \mapsto z + w$ é uma isometria.*

Chamamos de *sistema de Kronecker* um sistema ergódico que é conjugado a um grupo de rotação. Provaremos nesta seção que todo sistema ergódico possui um sistema de Kronecker maximal, chamado *fator de Kronecker*. De grande importância em Teoria Ergódica de Ramsey, o fator de Kronecker é uma “simplificação” do sistema original, que apresenta mais estrutura e portanto mais fácil de ser estudado. A fim de construir o fator de Kronecker, precisamos de alguns resultados auxiliares.

Lema 2.5.1. *Se (X, β_X, μ, T) é ergódico então todo autovalor de U_T é simples, e o conjunto de todos os autovalores de U_T é um subgrupo de \mathbb{S}^1 .*

Demonstração. Para começar, suponha que $U_T f = \lambda f$, onde $\lambda \in \mathbb{C}$ e $f \in L^2_\mu(X) \setminus \{0\}$ é um autovalor e uma autofunção. Como U_T é unitário, das propriedades de produto interno segue que

$$\langle f, f \rangle = \langle U_T f, U_T f \rangle = \langle \lambda f, \lambda f \rangle = \lambda \bar{\lambda} \langle f, f \rangle.$$

Como $f \neq 0$, temos $\langle f, f \rangle > 0$ e portanto $\lambda \bar{\lambda} = 1$. Desta forma, o conjunto dos autovalores $\Gamma = \{\lambda \in \mathbb{C} : \lambda \text{ é autovalor de } U_T\}$ está contido em \mathbb{S}^1 .

Por outro lado, note que $U_T |f|(x) = |U_T f(x)| = |\lambda f(x)| = |\lambda| \cdot |f(x)| = |f(x)|$ para μ -quase todo $x \in X$. Por ergodicidade, $|f| \neq 0$ é constante em quase todo ponto.

Mostraremos agora que Γ é subgrupo de \mathbb{S}^1 . Para tal, basta mostrar se $\lambda, \eta \in \Gamma$ então $\lambda \eta^{-1} \in \Gamma$. Para tal, sejam $f, g \in L^2_\mu \setminus \{0\}$ autofunções associadas a λ e η , respectivamente. Então

$$\left(\frac{f}{g}\right) \circ T = \frac{f \circ T}{g \circ T} = \frac{\lambda f}{\eta g} = \frac{\lambda}{\eta} \left(\frac{f}{g}\right).$$

Logo, $\lambda/\eta \in \Gamma$, conforme afirmado.

Para ver que os autovalores são simples, tome $f, g \neq 0$ autofunções com autovalor associado λ . Queremos mostrar que f e g são linearmente dependentes. Para tal, note que f/g é uma autofunção com autovalor associado $\lambda/\lambda = 1$. Por ergodicidade, f/g é constante, e portanto f e g são linearmente dependentes, o que encerra a prova do lema. \square

Definição 2.5.2 (Fator de Kronecker). *Sejam $\mathbb{X} = (X, \beta_X, \mu, T)$ sistema que preserva medida e $\mathbb{Y} = (Y, \beta_Y, \nu, S)$ um fator. Diremos que \mathbb{Y} é o fator de Kronecker de \mathbb{X} quando \mathbb{Y} é sistema de Kronecker e é maximal com esta propriedade, ou seja, dado qualquer outro fator que é sistema de Kronecker, ele também é fator de \mathbb{Y} .*

Teorema 2.5.1 (Construção do fator de Kronecker). *Seja (X, β_X, μ, T) um sistema ergódico invertível que preserva medida, onde μ é uma medida de probabilidade de Borel. Seja \mathcal{A} a σ -álgebra de X gerada pelas autofunções em $L^2_\mu(X)$ de U_T . Então o fator correspondente (Y, β_Y, ν, S) é o maior fator de (X, β_X, μ, T) que é sistema de Kronecker.*

Demonstração. Pelo lema anterior, todos os autovalores de U_T são simples. Para cada autovalor λ , podemos associar uma única autofunção f_λ em $L^2_\mu(X)$ que gera o autoespaço associado a λ . Sem perda de generalidade, suponha que f_λ toma valores em \mathbb{S}^1 , possivelmente

trocando-a por $x \rightarrow \frac{f_\lambda(x)}{|f_\lambda(x)|}$. Isto está bem definido, pois pela prova do Lema 2.5.1 a função $x \rightarrow |f_\lambda(x)|$ é constante em μ -quase todo ponto. Então $\mathcal{L} = \{f_\lambda \in L^2_\mu(X) : \lambda \text{ é autovalor}\}$ é um subconjunto ortonormal. Como $L^2_\mu(X)$ é separável, \mathcal{L} é enumerável. Seja f_1, f_2, \dots uma enumeração de \mathcal{L} , com autovalores associados $\lambda_1, \lambda_2, \dots$, de sorte que $U_T f_i = \lambda_i f_i$, $\forall i \in \mathbb{N}$.

Defina $\varphi : X \rightarrow \mathbb{S}^\mathbb{N}$ por

$$\varphi(x) = (f_i(x))_{i \in \mathbb{N}}.$$

Munindo $\mathbb{S} = \{z \in \mathbb{C} : |z| = 1\}$ da topologia usual e $\mathbb{S}^\mathbb{N}$ da topologia produto, temos que $\mathbb{S}^\mathbb{N}$ é um grupo abeliano compacto, o qual é metrizável. Denote por $R_x : \mathbb{S}^\mathbb{N} \rightarrow \mathbb{S}^\mathbb{N}$ a rotação por $x = (x_i)_{i \in \mathbb{N}} : R_x(y_i)_{i \in \mathbb{N}} = (y_i x_i)_{i \in \mathbb{N}}$. Seja então $a = (\lambda_i)_{i \in \mathbb{N}}$. Note que

$$R_a \circ \varphi(x) = (\lambda_i f_i(x))_{i \in \mathbb{N}} = (f_i(T(x)))_{i \in \mathbb{N}} = \varphi \circ T(x).$$

Defina $\nu = \varphi_* \mu$, que é uma medida sobre a σ -álgebra $\varphi_*(\beta_X)$, definida por $B \in \varphi_*(\beta_X) \iff \varphi^{-1}(B) \in \beta_X$. Noutras palavras, esta é a menor σ -álgebra para a qual φ é mensurável, e por construção de φ , $\mathcal{A} = \varphi^{-1}(\beta_{\mathbb{S}^\mathbb{N}})$. Podemos supor, sem perda de generalidade, que $a \in \varphi(X)$. Ora, caso $a \notin \varphi(X)$, tome $b \in \varphi(X)$ e ponha $\tilde{\varphi} = R_{ab^{-1}} \circ \varphi$. Como $\mathbb{S}^\mathbb{N}$ é abeliano, temos

$$R_a \circ (R_{ab^{-1}} \circ \varphi) = R_{ab^{-1}} \circ (R_a \circ \varphi) = (R_{ab^{-1}} \circ \varphi) \circ T$$

e então $\tilde{\varphi}$ conjuga T e R_a e vale $a \in \tilde{\varphi}(X)$.

Continuando a prova, é consequência de $\varphi \circ T = R_a \circ \varphi$ que $\varphi \circ T^n = R_a^n \circ \varphi$ para todo $n \in \mathbb{Z}$. Como $a \in \varphi(X)$, existe $y \in X$ tal que $\varphi(y) = a$. Logo, $a^n = R_a^{n-1}(a) = R_a^{n-1}(\varphi(y)) = \varphi(T^{n-1}y)$, donde $a^n \in \varphi(X)$ para todo $n \in \mathbb{Z}$. Desta forma, concluímos que o subgrupo cíclico $\{a^n : n \in \mathbb{Z}\}$ gerado por a está contido em $\varphi(X)$. Defina $Y := \overline{\{a^n : n \in \mathbb{Z}\}}$. Por ser o fecho de um subgrupo, dos nossos conhecimentos de topologia sabemos que Y é um subgrupo fechado, e portanto compacto e metrizável. É claro que Y é R_a -invariante, e podemos definir $S := R_a|_Y : Y \rightarrow Y$. Além do mais, o subgrupo cíclico $\{a^n : n \in \mathbb{Z}\}$ é denso em Y . Segue que R_a restrito a Y é unicamente ergódico pelo Teorema 6.0.2, e que a única medida invariante é a medida de Haar em Y .

Por outro lado, por construção, $\nu = \varphi_* \mu$ é R_a -invariante e ergódica. Afirimo que $\nu(Y) = 1$. De fato, observe que ν é ergódica para R_a implica $R_a|_{\text{supp}(\nu)}$ transitiva. Como o fecho das órbitas de R_a tem a forma $\overline{\{R_a^n(x) : n \in \mathbb{Z}\}} = Y + x$, a menos de translação

podemos supor que o ponto cuja órbita é densa em $\text{supp}(\nu)$ é o ponto a . Logo, $Y = \text{supp}(\nu)$ e daí $\nu(Y) = 1$. Então podemos ver ν como uma medida em Y que é ergódica para $R_a|_Y$. Da unicidade ergódica, obtemos que ν é a medida de Haar de Y , e daí $\varphi : X \rightarrow Y$ é uma aplicação de fator. \square

Nos falta mostrar a maximalidade do fator de Kronecker. Para tal, precisamos entender a relação entre sub- σ -álgebras invariantes e fatores. Sabemos na teoria que na verdade há uma bijeção entre eles, mas no nosso trabalho, mostraremos tal bijeção em um contexto simplificado, qual seja, em *sistemas de espectro discreto*. Para entender esta relação e cada um dos objetos citados, invocamos alguns resultados clássicos cujas provas fogem do escopo deste texto (o leitor interessado pode consultá-las em (EINSIEDLER; WARD, 2011, Capítulo 6)).

Definição 2.5.3. *Seja (X, β, μ) um espaço de medida, e sejam $A, B \in \beta$. Dizemos que $A \underset{\mu}{=} B$ quando $\mu(A \Delta B) = 0$.*

É fácil ver que “ $\underset{\mu}{=}$ ” estabelece uma relação de equivalência em β . Seja $\tilde{\beta}$ o conjunto das classes de equivalência. Dado $\alpha \in \beta$, escreveremos $\tilde{\alpha}$ para representar o conjunto das classes de equivalência de α , e $\alpha \underset{\mu}{=} \eta$ quando $\tilde{\alpha} = \tilde{\eta}$. Isto quer dizer que os conjuntos são iguais, a menos de conjuntos de medida nula.

Definição 2.5.4. *Seja (X, β_X, μ, T) um sistema que preserva medida. Dizemos que uma sub- σ -álgebra $\mathcal{A} \subset \beta_X$ é T -invariante quando $T^{-1}(\mathcal{A}) \underset{\mu}{=} \mathcal{A}$.*

Lema 2.5.2. *Sejam (X, β_X, μ, T) e (Y, β_Y, ν, S) sistemas invertíveis que preservam medida, $\varphi : X \rightarrow Y$ uma aplicação de fator e $\mathcal{A} = \varphi^{-1}(\beta_Y)$. Então $T^{-1}\mathcal{A} \underset{\mu}{=} \mathcal{A}$.*

Noutras palavras, todo fator está associado a uma sub- σ -álgebra invariante de β_X . O teorema abaixo nos assegura a recíproca.

Teorema 2.5.2. *Seja (X, β_X, μ, T) um sistema que preserva medida, e seja \mathcal{A} uma sub- σ -álgebra T -invariante. Então existe um sistema (Y, β_Y, ν, S) que preserva medida e $\varphi : X \rightarrow Y$ aplicação de fator tal que $\varphi^{-1}(\beta_Y) \underset{\mu}{=} \mathcal{A}$.*

O Lema 2.5.2 e Teorema 2.5.2 nos dizem que há uma dualidade entre fatores de um sistema que preserva medida e suas sub- σ -álgebras invariantes, quando os espaços considerados são espaços métricos compactos.

Agora discutimos a noção de *espectro discreto*. Para isto, definimos $H_T := \overline{\text{span}\{f \in L_\mu^2(X) : f \text{ é autofunção de } U_T\}}$. É fácil ver que $H_T \subset L_\mu^2(X)$ é completamente invariante por U_T , isto é, $U_T(H_T) = H_T$.

Definição 2.5.5 (Espectro discreto). *Sejam (X, β_X, μ, T) um sistema ergódico e $U_T : L_\mu^2(X) \rightarrow L_\mu^2(X)$ seu operador de Koopman. Dizemos que (X, β_X, μ, T) possui espectro discreto se $H_T = L_\mu^2(X)$.*

Uma pergunta natural de classificação de sistemas é a seguinte: em quais condições dois sistemas ergódicos são conjugados a partir de suas propriedades espectrais, isto é, quando uma propriedade espectral caracteriza completamente um sistema ergódico? O próximo teorema, devido a von Neumann e Halmos, (HALMOS; NEUMANN, 1942), fornece uma resposta completa a esta pergunta para sistemas ergódicos com espectro discreto.

Teorema 2.5.3. *Sejam \mathbb{X} e \mathbb{Y} dois sistemas ergódicos com espectro discreto. Então \mathbb{X} e \mathbb{Y} são conjugados se e somente se possuem o mesmo espectro.*

Acima, dizer que \mathbb{X} e \mathbb{Y} possuem o mesmo espectro significa dizer que os conjuntos de autovalores dos operadores de Koopman coincidem. Dito isto, o teorema abaixo apresenta exemplos de sistemas que possuem espectro discreto: sistemas de Kronecker.

Teorema 2.5.4. *Sistemas de Kronecker possuem espectro discreto.*

Na verdade, sistemas de Kronecker são os únicos exemplos de sistemas com espectro discreto. Para isto, lembre-se que o conjunto de autovalores do operador de Koopman é sempre um subgrupo de \mathbb{S}^1 (Lema 2.5.1). Sob a égide do Teorema 2.5.3, basta mostrarmos que dado qualquer subgrupo Γ de \mathbb{S}^1 , existe um sistema de Kronecker com espectro igual a Γ . Este é o conteúdo do teorema abaixo.

Teorema 2.5.5. *Seja Γ um subgrupo de \mathbb{S}^1 . Então existe um sistema de Kronecker com espectro Γ .*

Findas estas caracterizações, voltemos à prova do Teorema 2.5.1.

Prova da maximalidade do fator de Kronecker. Seja \mathbb{X} um sistema ergódico como no Teorema 2.5.1 e seja \mathbb{Y} seu fator de Kronecker, com aplicação de fator $\varphi : X \rightarrow Y$. Suponha

que $\mathbb{Y}' = (Y', \beta_{Y'}, \nu', S')$ seja um sistema de Kronecker que é fator de X , com aplicação de fator $\varphi' : X \rightarrow Y'$. Considere

$$\varphi^* : L_\nu^2(Y) \rightarrow L_\mu^2(X) \text{ e } (\varphi')^* : L_{\nu'}^2(Y') \rightarrow L_\mu^2(X).$$

Afirmamos que $(\varphi')^*(L_{\nu'}^2(Y')) \subset \varphi^*(L_\nu^2(Y))$. Relembre que, pela construção de \mathbb{Y} , temos $\varphi^{-1}(\beta_Y) = \mathcal{A}$.

Como \mathbb{Y}' tem espectro discreto, temos que $H_{S'} = L_{\nu'}^2(Y')$. Note que se g é autofunção de $U_{S'}$ com autovalor λ , então $(\varphi')^*g$ é autofunção de U_T com o mesmo autovalor, pois:

$$U_T((\varphi')^*g) = U_T(g \circ \pi') = g \circ \pi' \circ T = g \circ S' \circ \pi' = (\varphi')^*(g \circ S') = \lambda(\varphi')^*(g).$$

Portanto, $(\varphi')^*(L_{\nu'}^2(Y')) \subset H_T$.

Afirmamos que $H_T = \varphi^*L_\nu^2(Y)$. Por um lado, a mesma argumentação do parágrafo acima dá que $H_T \supset \varphi^*L_\nu^2(Y)$. Para a outra inclusão, vamos usar que $H_T = \overline{\text{span}\{\mathbb{1}_A \in L_\mu^2(X) : A \in \mathcal{A}\}}$. Como $\mathcal{A} = \varphi^{-1}(\beta_Y)$, então para todo $A \in \mathcal{A}$ a função característica $\mathbb{1}_A$ se escreve como $\varphi^*\mathbb{1}_B$ para algum $B \in \beta_Y$, e portanto $\mathbb{1}_A \in \varphi^*L_\nu^2(Y)$. Tomando fecho, segue que $H_T \subset \varphi^*L_\nu^2(Y)$.

Assim, $(\varphi')^*(L_{\nu'}^2(Y')) \subset H_T = \varphi^*(L_\nu^2(Y))$, o que prova o afirmado. Em particular, podemos mergulhar $L_{\nu'}^2(Y')$ isometricamente em $L_\nu^2(Y)$. Notando que $L_{\nu'}^2(Y')$ é S -invariante pela Proposição 2.3.1, segue do Teorema 2.5.2 que existe um fator $\mathbb{W} = (W, \beta_W, \eta, Q)$ de \mathbb{Y} com aplicação de fator $\psi : Y \rightarrow W$ tal que $L_\eta^2(W) = L_{\nu'}^2(Y')$. Por construção, temos que $U_{S'}$ e U_Q são conjugados. Pelo Teorema 2.5.3, \mathbb{Y}' e \mathbb{W} são conjugados. Como \mathbb{W} é fator de \mathbb{Y} , segue que \mathbb{Y}' é fator de \mathbb{Y} , o que mostra a maximalidade desejada. \square

3 TEORIA ERGÓDICA DE RAMSEY

Em 1936, em um artigo intitulado “On some sequence of integers” (ERDŐS, 1936), Paul Erdős e Paul Turán conjecturaram que um conjunto com densidade positiva dos números naturais possui progressões aritméticas finitas arbitrariamente grandes. O primeiro resultado parcial satisfatório é o Teorema de Roth (ROTH, 1953), provado por Klaus Roth, que fornece uma prova restringindo-se a progressões aritméticas com três termos. Em 1975, com uma prova puramente combinatoria, a conjectura foi estabelecida em sua generalidade por Endre Szemerédi. Dois anos após, em 1977, Hillel Furstenberg forneceu uma prova ergódica do mesmo teorema, via um princípio de correspondência e um profundo resultado de recorrência, atualmente conhecido como Teorema da Recorrência Múltipla de Poincaré. Com esta prova, Furstenberg inaugurou uma nova área conhecida como *Teoria Ergódica de Ramsey*, no qual busca-se resolver problemas de combinatória por meio de um aparato de Teoria Ergódica. Neste capítulo, introduzimos noções de Teoria de Ramsey e da Teoria Ergódica de Ramsey, expondo as principais técnicas. Finalizamos o capítulo provando o Teorema de Roth, que fornece um excelente exemplo da prolífica relação entre Combinatória e Teoria Ergódica.

Este capítulo foi baseado nas notas de aula presentes em (LIMA, Rio de Janeiro, 2009).

3.1 Elementos de teoria de ramsey

Em Teoria de Ramsey, buscamos condições suficientes para que conjuntos discretos possuam estruturas bem organizadas. Por exemplo, quando os conjuntos discretos são subconjuntos dos números naturais, fixamos $F \subset \mathbb{N}$ e perguntamos:

Sob quais condições um subconjunto $A \subset \mathbb{N}$ contém F ? (*)

Neste caso, o subconjunto fixo F pode ser considerado uma estrutura, e nos perguntamos quais $A \subset \mathbb{N}$ possuem esta estrutura. Note que uma estrutura padrão dos naturais é o próprio \mathbb{N} . Mas ao fazer $F = \mathbb{N}$, o único conjunto que o contém é o próprio \mathbb{N} . Então podemos enfraquecer a condição sobre A e toma-lo igual ao conjunto $\{1, 2, \dots, n\}$ dos n primeiros naturais. Tampouco é razoável esperar que conjuntos contêmam A , pois essa

é uma condição forte. Podemos, por outro lado, esperar que conjuntos grandes contenham “cópias” de A , dadas por uma homotetia seguida de translação, igual à expressão $a + kA$.

Observe que $a + kA$ representa uma *progressão aritmética* $F = \{a + k, \dots, a + nk\}$ com n termos. Desta forma, dizemos que progressões aritméticas finitas são *cópias finitas* de \mathbb{N} . Nessa direção, reformulamos a pergunta (*) do seguinte modo:

Quais subconjuntos de \mathbb{N} possuem progressões aritméticas finitas arbitrariamente grandes?

Uma condição suficiente é dada pelo seguinte teorema, devido a van der Waerden (WAERDEN, 1927).

Teorema 3.1.1 (van der Waerden, 1927). *Seja $\mathbb{N} = A_1 \cup \dots \cup A_r$ uma partição finita dos números naturais. Então algum A_i possui progressões aritméticas finitas arbitrariamente grandes.*

O teorema acima é *qualitativo*: ele nos fornece a *existência* de pelo menos um conjunto com progressões aritméticas finitas arbitrariamente longas. É importante ressaltar que há resultados *quantitativos*, no qual a existência é assegurada estimando quantas progressões aritméticas existem. Neste texto, trataremos apenas de resultados qualitativos.

Relembrando que o resultado acima pode ser visto como algum dos conjuntos A_i contendo todas as cópias finitas de \mathbb{N} , enunciemos um princípio da Teoria de Ramsey.

Primeiro Princípio de Ramsey: *Estruturas bem organizadas são preservadas por partições finitas.*

Observamos que o Teorema de van der Waerden é falso se pedirmos por progressões aritméticas infinitas.

Exemplo. Fornecemos uma partição de \mathbb{N} em que nenhum subconjunto possui uma progressão aritmética infinita. Dados $a < b$ em \mathbb{N} , escreva $[a, b) = \{n \in \mathbb{N} : a \leq n < b\}$. Sejam

$$C_1 = \bigcup_{k \geq 1} [2^{2k}, 2^{2k+1} - 1), \quad C_2 = \bigcup_{k \geq 1} [2^{2k-1}, 2^{2k} - 1),$$

que formam uma partição de \mathbb{N} , onde os blocos entre potências de 2 são distribuídos alternadamente entre C_1 e C_2 .

Suponha, por absurdo, que exista uma progressão aritmética infinita $F = \{a + kn : k \in \mathbb{N}\}$ contida em C_1 , para alguma escolha de $a, n \in \mathbb{N}$. Essa progressão intersecta qualquer intervalo de tamanho n que começa a partir de a . Tomando k tal que $2^{2k-1} > \max\{a, n\}$, necessariamente F intersecta o intervalo $[2^{2k-1}, 2^{2k} - 1)$, o que contradiz o fato de F ser um subconjunto de C_1 . Analogamente, C_2 não contém progressões aritméticas infinitas.

A seguir, introduzimos uma noção de grandeza assintótica, conhecida por densidade superior. Relembramos que $|X|$ denota a quantidade de elementos de X .

Definição 3.1.1 (Densidade superior). A **densidade superior** de $A \subset \mathbb{N}$ é definida por

$$\bar{d}(A) = \limsup_{n \rightarrow \infty} \frac{|A \cap \{1, \dots, n\}|}{n},$$

Quando o limite existe, dizemos que $\bar{d}(A)$ é a densidade de A .

Exemplo. Dado $a \in \mathbb{N}$, seja $a\mathbb{N} = \{an : n \in \mathbb{N}\}$ o conjunto dos múltiplos de a . Temos $\bar{d}(A) = 1/a$. De fato, dado $N \in \mathbb{N}$, pelo algoritmo da divisão, escrevemos $N = q(N)a + r(N)$, onde $q(N) \geq 0$ e $0 \leq r(N) < a$. Deste modo, $|A \cap \{1, \dots, N\}|$ é igual à quantidade de múltiplos de a menores ou iguais a N , que é $q(N) = \frac{N-r(N)}{a}$, logo $\frac{|A \cap \{1, \dots, N\}|}{N} = \frac{1}{a} - \frac{r(N)}{N}$. Como $0 \leq r(N) < a$, fazendo $N \rightarrow \infty$ concluímos que $\bar{d}(A) = 1/a$.

Recorde que um conjunto $A \subset \mathbb{N}$ é *sindético* se possui buracos uniformemente limitados.

Proposição 3.1.1. *Seja $A = \{a_1 < a_2 < \dots\}$ um conjunto sindético com $a_{k+1} - a_k \leq c$ para todo $k \in \mathbb{N}$. Então*

$$\liminf_{N \rightarrow \infty} \frac{|A \cap \{1, \dots, N\}|}{N} \geq \frac{1}{c}.$$

Em particular, $\bar{d}(A) \geq 1/c$.

Demonstração. A prova é simples: por indução, verifica-se facilmente que $a_k \leq c(k-1) + a_1$. Para $N > a_1$, escreva pelo algoritmo da divisão que $N - a_1 = q(N)c + r(N)$, com $0 \leq r(N) <$

c. Segue que $|A \cap \{1, \dots, N\}| \geq \frac{N - a_1 - r(N)}{c}$. Como $r(N)$ é limitada, segue que

$$\liminf_{N \rightarrow \infty} \frac{|A \cap \{1, \dots, N\}|}{N} \geq \frac{1}{c}.$$

Em particular, $\bar{d}(A) \geq 1/c$. □

Note que $\bar{d}(A) = \delta$ mede assintoticamente quanto de A temos em intervalos de números naturais começando a partir de 1. Por exemplo, supondo que o limite existe, ele nos diz que para n grande, $|A \cap \{1, \dots, n\}| \approx \delta n$. Desta forma, se $\delta > 0$, então $|A \cap \{1, \dots, n\}|$ tem um crescimento *linear*. Em 1936, Erdős e Turán conjecturaram que se um conjunto possui densidade superior positiva, então este conjunto possui progressões aritméticas finitas arbitrariamente grandes (ERDŐS, 1936). Em 1975, Endre Szemerédi forneceu uma prova que usa argumentos combinatórios a esta conjectura, e hoje ela é conhecida como Teorema de Szemerédi, conforme enunciamos abaixo.

Teorema 3.1.2 (Szemerédi, 1975). *Seja $A \subset \mathbb{N}$ tal que $\bar{d}(A) > 0$. Então A possui progressões aritméticas finitas arbitrariamente grandes.*

Note que se $\mathbb{N} = A_1 \cup \dots \cup A_r$, então $\bar{d}(A_i) \geq \frac{1}{r}$ para algum $i \in \{1, \dots, r\}$. De fato, como a união é disjunta, temos

$$1 = \frac{n}{n} = \sum_{i=1}^r \frac{|A_i \cap \{1, \dots, n\}|}{n},$$

donde tomando o lim sup na expressão acima segue que

$$\sum_{i=1}^r \bar{d}(A_i) \geq 1,$$

e daí pelo menos para um i vale que $\bar{d}(A_i) \geq \frac{1}{r}$. Fica assim claro que o Teorema Szemerédi implica o Teorema de van der Waerden.

Note que ambos teoremas gozam da mesma conclusão: subconjuntos grandes possuem progressões aritméticas finitas arbitrariamente grandes. Isto pode ser condensado no seguinte princípio.

Segundo Princípio de Ramsey: *Conjuntos “grandes” possuem estruturas bem organizadas.*

Obviamente, o conjunto de grandeza é amplo, e pode ser considerado em vários contextos. Ainda há um terceiro princípio, que agora discutiremos. Seja $A \subset \mathbb{N}$ com $\bar{d}(A) > 0$. Pelo Teorema de Szemerédi, A possui infinitas progressões aritméticas finitas de qualquer tamanho. Como uma progressão aritmética de tamanho $m > n$ contém várias progressões aritméticas de tamanho n , de fato este teorema implica que, para todo n , A possui muitas progressões aritméticas de tamanho n . Neste contexto, temos o terceiro princípio de Ramsey.

Terceiro Princípio de Ramsey: *Conjuntos “grandes” possuem estruturas bem organizadas em abundância.*

Os princípios de Ramsey são um guia da Teoria de Ramsey. Adaptações deste princípio são ricas e exploradas, e um exemplo delas é a conjectura $B + B + t$, cuja discussão e prova devidas a (KRA et al., 2024b) serão vistas no último capítulo desta dissertação.

3.2 Princípio de correspondência de furstenberg

O conceito de densidade superior é satisfatório para o contexto do Teorema de Szemerédi. Para a sequência da dissertação, consideramos uma generalização de densidade e do Teorema de Szemerédi. A noção de grandeza é definida a partir de *sequências de Følner*. Dizemos que uma sequência $(\Phi_n)_{n \in \mathbb{N}}$ de subconjuntos de \mathbb{N} forma uma sequência de Følner se para todo $t \in \mathbb{N}$ vale que

$$\lim_{n \rightarrow \infty} \frac{|(\Phi_n + t) \cap \Phi_n|}{|\Phi_n|} = 1.$$

Informalmente, esta condição diz que os conjuntos Φ_n possuem uma proporção baixa de elementos próximos às suas fronteiras. Analogamente, outra forma de escrever o limite acima é que

$$\lim_{n \rightarrow \infty} \frac{|(\Phi_n + t) \Delta \Phi_n|}{|\Phi_n|} = 0$$

para todo $t \in \mathbb{N}$, onde $A \Delta B = (A \setminus B) \cup (B \setminus A)$ denota a *diferença simétrica* de A e B .

Note que necessariamente $|\Phi_n| \rightarrow \infty$ quando $n \rightarrow \infty$. De fato, se $|\Phi_{n_k}| = m$ para uma sequência infinita n_k , então a condição

$$\lim_{n_k \rightarrow \infty} \frac{|(\Phi_{n_k} + t) \cap \Phi_{n_k}|}{|\Phi_{n_k}|} = 1$$

implica que $|(\Phi_{n_k} + t) \cap \Phi_{n_k}| = m$ para k grande. Como $|\Phi_{n_k}| = m$, isso implica que $\Phi_{n_k} + t = \Phi_{n_k}$, o que claramente não ocorre.

Definição 3.2.1 (Densidade de Banach superior). *Seja $A \subset \mathbb{N}$. Definimos a densidade de Banach superior com relação à sequência de Følner $(\Phi_n)_n$ de A como*

$$d^*(A) = \limsup_{n \rightarrow \infty} \frac{|A \cap \Phi_n|}{|\Phi_n|}.$$

A prova dada por Szemerédi ao teorema que leva seu nome é puramente combinatória. Em 1977, Furstenberg provou o mesmo teorema usando um aparato ergódico (FURSTENBERG, 1977), que introduziremos a seguir. Sucintamente, dado $A \subset \mathbb{N}$, iremos associar um elemento $\underline{x} = (x_n)_n \in \{0, 1\}^{\mathbb{N}} = \Sigma$ pondo $x_n = 1$ para $n \in A$ e $x_n = 0$ para $n \notin A$. Reciprocamente, dado $\underline{x} = (x_n)_n \in \{0, 1\}^{\mathbb{N}}$, definimos $A \subset \mathbb{N}$ via a equivalência $n \in A \iff x_n = 1$.

Definição 3.2.2 (Ponto genérico). *Seja (X, β_X, μ, T) um sistema que preserva medida. Dizemos que um ponto $a \in X$ é genérico com relação a uma sequência de Følner $(\Phi_n)_n$ quando*

$$\mu = \lim_{n \rightarrow \infty} \frac{1}{|\Phi_n|} \sum_{k \in \Phi_n} \delta_{T^k a},$$

onde δ_x é a medida de Dirac em x e o limite é tomado na topologia fraca-*

Denotamos o conjunto dos pontos genéricos de (X, β_X, μ, T) com relação a uma sequência de Følner $\Phi = (\Phi_n)_{n \in \mathbb{N}}$ por $\text{gen}(\mu, \Phi)$.

Teorema 3.2.1 (Princípio de correspondência de Furstenberg). *Seja $A \subset \mathbb{N}$ com $d^*(A) > 0$ onde $\Phi = (\Phi_n)_n$ é uma sequência de Følner. existem um sistema que preserva medida (X, β_X, μ, T) , onde X é um espaço métrico compacto, $E \subset X$ subconjunto de Borel aberto e fechado com $\mu(E) = d^*(A) > 0$, uma sequência de Følner $\Psi = (\Psi_n)_n$ e $a \in \text{gen}(\mu, \Psi)$ tais que $A = \{n \in \mathbb{N} : T^n a \in E\}$. Ademais,*

$$d^*(A \cap (A - n_1) \cap \dots \cap (A - n_k)) \geq \mu(E \cap T^{-n_1} E \cap \dots \cap T^{-n_k} E).$$

Demonstração. Seja $A \subset \mathbb{N}$ como no enunciado. Considere $\Sigma = \{0, 1\}^{\mathbb{N}}$. Dados $\underline{x} = (x_n), \underline{y} = (y_n) \in \Sigma$, defina $N_0 = \min\{n; x_n \neq y_n\}$, e ponha $d(\underline{x}, \underline{y}) = 2^{-N_0}$. Sabemos da teoria clássica que d é uma métrica, que torna Σ um espaço métrico compacto. Considere β a σ -álgebra de Borel gerada pela topologia de d .

Seja agora $T : \Sigma \rightarrow \Sigma$ o shift em Σ , isto é, $T(x_n)_n = (x_{n+1})_n$. Para construir a medida, tome uma subsequência $(n_k)_k$ tal que

$$d^*(A) = \lim_{k \rightarrow \infty} \frac{|A \cap \Phi_{n_k}|}{|\Phi_{n_k}|}.$$

e defina $\mu_k = \frac{1}{|\Phi_{n_k}|} \sum_{m \in \Phi_{n_k}} \delta_{T^m \underline{x}}$, onde \underline{x} é definido por A .

Como o conjunto das medidas de probabilidade munidos da topologia fraca-* é compacto e metrizável, a sequência $(\mu_k)_k$ admite uma subsequência convergente. Deste modo, a menos de tomar subsequência, assumamos que $(\mu_k)_k$ converge a μ . Afirmamos que $T_*\mu = \mu$. Para isto, primeiro note que para $E \subset \Sigma$ mensurável e $\underline{y} \in \Sigma$, temos $\delta_{\underline{y}}(T^{-1}(E)) = 1 \iff \underline{y} \in T^{-1}(E) \iff T\underline{y} \in E \iff \delta_{T\underline{y}}(E) = 1$. Isto mostra que

$$T_*\mu_k = \frac{1}{|\Phi_{n_k}|} \sum_{m \in \Phi_{n_k}} T_*\delta_{T^m \underline{x}} = \frac{1}{|\Phi_{n_k}|} \sum_{m \in \Phi_{n_k}} \delta_{T^{m+1} \underline{x}} = \sum_{m \in \Phi_{n_k} - 1} \delta_{T^m \underline{x}}.$$

Logo, para $E \subset \Sigma$ mensurável,

$$|T_*\mu_k(E) - \mu_k(E)| = \left| \frac{1}{|\Phi_{n_k}|} \sum_{m \in \Phi_{n_k} - 1} \delta_{T^m \underline{x}}(E) - \sum_{m \in \Phi_{n_k}} \delta_{T^m \underline{x}}(E) \right| \leq \frac{|(\Phi_{n_k} - 1) \Delta \Phi_{n_k}|}{|\Phi_{n_k}|}.$$

Pela propriedade que define sequências de Følner, temos que

$$|T_*\mu_k(E) - \mu_k(E)| \leq \frac{|(\Phi_{n_k} - 1) \Delta \Phi_{n_k}|}{|\Phi_{n_k}|} \xrightarrow{k \rightarrow \infty} 0.$$

Concluimos que $T_*\mu(E) = \mu(E)$ para todo aberto de Σ (neste momento, usamos o Teorema de Portmanteau). Como Σ é totalmente desconexo, todo aberto é conjunto de continuidade, logo, mais uma vez pelo Teorema de Portmanteau, concluimos que $T_*\mu = \mu$. Além do mais, por construção,

$$\mu = \lim_{k \rightarrow \infty} \frac{1}{|\Phi_{n_k}|} \sum_{m \in \Phi_{n_k}} \delta_{T^m \underline{x}}.$$

Seja $\Psi = (\Psi_k)_k = (\Phi_{n_k})_k$. Então $\underline{x} \in \text{gen}(\mu, \Psi)$.

Agora, seja $E = [1]_0 = \{(x_n)_{n \in \mathbb{N}} : x_0 = 1\}$. Note que $n \in A \iff x_n = 1 \iff T^n \underline{x} \in E \iff \delta_{T^n \underline{x}}(E) = 1$. Logo,

$$\mu_k(E) = \frac{1}{|\Phi_{n_k}|} \sum_{m \in \Phi_{n_k}} \delta_{T^m \underline{x}}(E) = \frac{|A \cap \Phi_{n_k}|}{|\Phi_{n_k}|},$$

donde tomando o limite em ambos os lados e usando que E é conjunto de continuidade, concluimos que $\mu(E) = d^*(A)$. Ainda pelo que provamos acima, $A = \{n \in \mathbb{N} : T^n \underline{x} \in E\}$.

Por fim, sejam $n_1, \dots, n_l \in \mathbb{N}$. Note que

$$\begin{aligned} \delta_{T^m \underline{x}}(E \cap T^{-n_1} E \cap \dots \cap T^{-n_l} E) = 1 &\iff T^m \underline{x}, T^{m+n_1} \underline{x}, \dots, T^{m+n_l} \underline{x} \in E \\ &\iff m, m+n_1, \dots, m+n_l \in A \\ &\iff m \in A \cap (A-n_1) \cap \dots \cap (A-n_l). \end{aligned}$$

Portanto,

$$\begin{aligned} \mu_k(E \cap T^{-n_1} E \cap \dots \cap T^{-n_l} E) &= \frac{1}{|\Phi_{n_k}|} \sum_{m \in \Phi_{n_k}} \delta_{T^m \underline{x}}(E \cap T^{-n_1} E \cap \dots \cap T^{-n_l} E) \\ &= \frac{|(A \cap (A-n_1) \cap \dots \cap (A-n_l)) \cap \Phi_{n_k}|}{|\Phi_{n_k}|}. \end{aligned}$$

Passando o limsup, concluímos que

$$d^*(A \cap (A-n_1) \cap \dots \cap (A-n_k)) \geq \mu(E \cap T^{-n_1} E \cap \dots \cap T^{-n_k} E).$$

Isto conclui a prova. □

Em geral, o sistema que preserva medida obtido pelo teorema acima não é ergódico. Podemos fazê-lo ser ergódico, às custas de perder a relação entre a densidade de Banach superior das interseções de translações de A .

Teorema 3.2.2 (Princípio de correspondência de Furstenberg - versão ergódica). *Seja $A \subset \mathbb{N}$ com $d^*(A) > 0$ para uma sequência de Følner $\Phi = (\Phi_n)_n$. Então existem um sistema ergódico (X, β_X, μ, T) , onde X é um espaço métrico compacto, $E \subset X$ subconjunto de Borel aberto e fechado com $\mu(E) > 0$, uma sequência de Følner $\Psi = (\Psi_n)_n$ e $a \in \text{gen}(\mu, \Psi)$ tais que $A = \{n \in \mathbb{N} : T^n a \in E\}$.*

Demonstração. Do teorema anterior, obtemos um sistema que preserva medida (X, β_X, μ, T) , $E \in \beta_X$ aberto e fechado com $\mu(E) = d^*(A) > 0$ tal que $A = \{n \in \mathbb{N} : T^n a \in E\}$. Como não há razão para que μ seja ergódico, recorremos à decomposição ergódica de μ : existe um sistema de medida $(Y, \beta_Y, \tilde{\mu})$ e uma transformação $y \in Y \mapsto \mu_y \in \mathcal{M}(X)$ mensurável tal que μ_y é invariante por T e ergódica para quase todo y satisfazendo $\mu(B) = \int_Y \mu_y(B) d\tilde{\mu}(y)$ para todo $B \in \beta_X$. Como $\mu(E) > 0$, existe $y \in Y$ tal que $\mu_y(E) > 0$. O sistema ergódico (X, β_X, μ_y, T) satisfaz o enunciado do teorema. □

Com os Teoremas 3.2.1 e 3.2.2, temos uma ferramenta para transportar problemas da Teoria de Ramsey para problemas de recorrência em Teoria Ergódica. Para exemplificar, enunciaremos a versão ergódica do Teorema de Furstenberg, também conhecida como *Teorema de Recorrência Múltipla de Poincaré*.

Teorema 3.2.3 (Teorema de Recorrência de Poincaré). *Seja (X, β_X, μ, T) um sistema que preserva medida e $A \in \beta_X$ tal que $\mu(A) > 0$. Dado um inteiro positivo $n \in \mathbb{N}$, existe um inteiro positivo $k > 0$ tal que*

$$\mu(A \cap T^{-k}A \cap \dots \cap T^{-k(n-1)}A) > 0.$$

Note que o Teorema de Recorrência de Poincaré implica o caso $n = 2$. O teorema acima foi provado por Furstenberg (FURSTENBERG, 1977). Na ocasião, ele mostrou que o Teorema de Szemerédi segue do Teorema de Recorrência Múltipla de Poincaré, via o princípio de correspondência, do seguinte modo. Seja $A \subset \mathbb{N}$ com $d^*(A) > 0$. Pelo Teorema 3.2.1, existe um sistema que preserva medida (X, β_X, μ, T) e $E \in \beta_X$ tal que $\mu(E) = d^*(A) > 0$ e

$$d^*(A \cap (A - n_1) \cap \dots \cap (A - n_k)) \geq \mu(E \cap T^{-n_1}E \cap \dots \cap T^{-n_k}E),$$

para todos $n_1, \dots, n_k \in \mathbb{N}$. Dado $n \in \mathbb{N}$, pelo Teorema de Recorrência Múltipla de Poincaré existe um inteiro $k > 0$ para o qual $\mu(E \cap T^{-k}E \cap \dots \cap T^{-k(n-1)}E) > 0$. Tomando $n_i = ik$ para $i = 1, \dots, n - 1$, segue que

$$d^*(A \cap (A - k) \cap \dots \cap (A - k(n - 1))) \geq \mu(E \cap T^{-k}E \cap \dots \cap T^{-k(n-1)}E) > 0$$

e em particular $A \cap (A - k) \cap \dots \cap (A - k(n - 1)) \neq \emptyset$. É fácil ver que $a \in A \cap (A - k) \cap \dots \cap (A - k(n - 1))$ se e somente se $a, a + k, \dots, a + k(n - 1) \in A$. Pela arbitrariedade da escolha de n , isto prova o Teorema de Szemerédi.

Com esta prova que usa um aparato de Teoria Ergódica, Furstenberg construiu uma ponte entre a Teoria de Ramsey e Teoria Ergódica: problemas em Teoria de Ramsey traduzem-se em problemas de recorrência em Teoria Ergódica, inaugurando a área que hoje conhecemos por *Teoria Ergódica de Ramsey*.

3.3 Teorema de roth

Nesta seção, vamos provar o caso particular do Teorema de Recorrência Múltipla de Poincaré quando $n = 3$. Este caso particular também é conhecido como a versão ergódica do Teorema de Roth (relembra: o Teorema de Roth afirma que conjuntos com densidade de Banach positiva possuem infinitas progressões aritméticas de tamanho 3). Enunciamos abaixo o resultado.

Teorema 3.3.1 (Teorema de Roth, versão ergódica). *Seja (X, β_X, μ, T) um sistema que preserva medida. Para toda $f \in L^\infty_\mu(X)$ com $f \geq 0$ e $\int_X f d\mu > 0$, vale que*

$$\liminf_{n \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \int_X f \cdot U_T^n f \cdot U_T^{2n} f d\mu > 0.$$

Para obter o caso particular $n = 2$ do Teorema de Recorrência Múltipla de Poincaré, dado $A \in \beta_X$ com $\mu(A) > 0$, basta tomar $f = \mathbb{1}_A$. Pelo teorema acima, existe $n > 0$ tal que $\int_X f \cdot U_T^n f \cdot U_T^{2n} f d\mu > 0$. Como $U_T \mathbb{1}_A = \mathbb{1}_{T^{-1}(A)}$, segue que $f \cdot U_T^n f \cdot U_T^{2n} f = \mathbb{1}_A \cdot \mathbb{1}_{T^{-n}(A)} \cdot \mathbb{1}_{T^{-2n}(A)} = \mathbb{1}_{A \cap T^{-n}A \cap T^{-2n}A}$, donde

$$\mu(A \cap T^{-n}A \cap T^{-2n}A) = \int_X f \cdot U_T^n f \cdot U_T^{2n} f d\mu > 0,$$

o que prova o afirmado.

Para provar o teorema, precisaremos de uma decomposição de $L^2_\mu(X)$ em dois subespaços invariantes por U_T .

Definição 3.3.1. *Sejam H um espaço de Hilbert e $U : H \rightarrow H$ um operador unitário.*

Dizemos que $x \in H$ é:

1) **quase-periódico** ou **compacto** quando o fecho de sua órbita por U é compacto, isto é, quando $\{U^n x : n \in \mathbb{Z}\}$ é pré-compacto;

2) **fracamente misturador** quando $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N |\langle U^n x, x \rangle| = 0$.

Definimos também $H_c = \{f \in H : f \text{ é compacto}\}$ e $H_{wm} = \{f \in H : f \text{ é quase misturador}\}$.

A sigla *wm* se refere ao termo “weak mixing”, que significa fracamente misturador em inglês. Quando construímos o fator de Kronecker no capítulo passado, definimos o conjunto $H_T \subset L^2_\mu(X)$. O próximo teorema mostra que H_c é a versão espectral de H_T (vide o comentário que antecede a Definição 2.5.5), caracterizando-o em termos dos autovetores de U .

Proposição 3.3.1. $H_c = H_U := \overline{\text{span}\{f \in H : f \text{ é autovetor de } U\}}$.

Demonstração. Seja $f \in H$ um autovetor de U associada ao autovalor $\lambda \in \mathbb{S}^1$. É claro que $U^n f = \lambda^n f$, $\forall n \in \mathbb{Z}$. Como $\{\lambda^n : n \in \mathbb{Z}\}$ é compacto em \mathbb{S}^1 , segue que $\overline{\{U^n f : n \in \mathbb{Z}\}}$ é compacto e portanto $f \in H_c$. Agora, seja $g = \sum_{i=1}^m x_i f_i$, onde $x_i \in \mathbb{C}$ e f_i é autovetor associado ao autovalor λ_i com $|\lambda_i| = 1$ e $\|f_i\| = 1$ para $i = 1, \dots, m$. Segue que $U^n g = \sum_{i=1}^m x_i \lambda_i^n f_i$, donde $\|U^n g\| \leq \sum_{i=1}^m \|x_i \lambda_i^n f_i\| \leq \sum_{i=1}^m |x_i|$, para todo $n \in \mathbb{Z}$. Então $\overline{\{U^n g : n \in \mathbb{Z}\}}$ é limitado, fechado

e contido no espaço vetorial de dimensão finita $\text{span}\{f_1, \dots, f_m\}$. Daí, $\overline{\{U^n g : n \in \mathbb{Z}\}}$ é compacto, provando que $g \in H_c$. Por fim, como H_c é fechado, obtemos a inclusão $H_U \subset H_c$.

Por outro lado, seja $g \in H_c$ e ponha $\Lambda = \{U^n g : n \in \mathbb{Z}\}$ e $H_0 = \overline{\text{span}(\Lambda)}$. Como $U(\Lambda) = \Lambda$, temos que $U(H_0) = H_0$. Segue que $U|_{H_0}$ é compacto, donde pelo Teorema Espectral, temos que $U|_{H_0}$ é diagonalizável. Como $g \in H_0$, g está no fecho do conjunto gerado pelos autovetores de $U|_{H_0}$, que são autovetores de U . Daí, $g \in H_U$, conforme queríamos provar. \square

Elementos compactos possuem propriedades de recorrência interessantes. Elas seguem de um teorema de dinâmica topológica, dado pela proposição abaixo.

Proposição 3.3.2. *Seja $T : X \rightarrow X$ uma isometria bijetiva em um espaço métrico compacto (X, d) . Seja $a \in X$ tal que $X = \overline{\{T^n a : n \in \mathbb{Z}\}}$. Para qualquer $\varepsilon > 0$, o conjunto $\{n \in \mathbb{Z} : d(T^n a, a) < \varepsilon\}$ é sindético.*

Demonstração. Mostraremos um resultado mais forte: $\{n \in \mathbb{Z} : d(T^n a, a) < \varepsilon\}$ é IP^* . Uma vez provado isto, a proposição seguirá do Teorema 2.1.2. Seja $B = FS(n_k)_{k \in \mathbb{N}} \subset \mathbb{Z}$ um conjunto IP , e ponha $N_l = n_1 + \dots + n_l$ para $l \geq 1$. Note que $\Lambda = \overline{\{T^{N_l} a : l \in \mathbb{N}\}}$ é compacto, e dado $\varepsilon > 0$, $\Lambda \subset \bigcup_{l \in \mathbb{N}} T^{N_l} B(a, \varepsilon/2)$ é uma cobertura aberta de Λ , logo existe uma subcobertura finita de Λ , digamos $\Lambda \subset T^{N_{l_1}} B(a, \varepsilon/2) \cup \dots \cup T^{N_{l_k}} B(a, \varepsilon/2)$. Daí, tomando $s > l_k$, temos $N_s > N_{l_i}$ para $1 \leq i \leq k$. Então temos $T^{N_s} B(a, \varepsilon/2) \cap T^{N_{l_i}} B(a, \varepsilon/2) \neq \emptyset$ para algum i . Por outro lado, como T é isometria, temos $T^i B(a, \varepsilon/2) = B(T^i a, \varepsilon/2)$. Daí, $B(T^{N_s} a, \varepsilon/2) \cap B(T^{N_{l_i}} a, \varepsilon/2) \neq \emptyset$. Tome $y \in B(T^{N_s} a, \varepsilon/2) \cap B(T^{N_{l_i}} a, \varepsilon/2)$. Novamente pelo fato de T ser isometria, temos

$$d(T^{N_s - N_{l_i}} a, a) = d(T^{N_s} a, T^{N_{l_i}} a) \leq d(T^{N_s} a, y) + d(T^{N_{l_i}} a, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Desta forma, $N_s - N_{l_i} = n_s + n_{s-1} + \dots + n_{l_i+1}$ pertence aos conjuntos $\{n \in \mathbb{Z} : d(T^n a, a) < \varepsilon\}$ e $FS(n_k)_{k \geq 1}$. Isto conclui a prova. \square

Corolário 3.3.1. *Seja H um espaço de Hilbert, $U : H \rightarrow H$ unitário e $x \in H$ um elemento compacto. Então o conjunto $\{n \in \mathbb{Z} : \|U^n x - x\| < \varepsilon\}$ é sindético para todo $\varepsilon > 0$.*

Demonstração. Ponha $X = \overline{\{U^n x : n \in \mathbb{Z}\}}$. É claro que $U(X) = X$ e portanto a restrição $U : X \rightarrow X$ é uma isometria. Pela proposição anterior, o resultado segue. \square

É fácil notar que, nas condições da Proposição 3.3.2, todo elemento de X é compacto. Com isto em mãos, conseguimos provar a seguinte proposição.

Proposição 3.3.3. *Seja (X, β_X, T, μ) um sistema que preserva medida com espectro discreto. Sejam $f \in L_\mu^\infty(X) \setminus \{0\}$ e $k \in \mathbb{N}$, onde f assume apenas valores reais. Para todo $\varepsilon > 0$, o conjunto*

$$\left\{ n \in \mathbb{N} : \int_X \prod_{i=0}^k U_T^{ni} f d\mu > \int_X f^{k+1} d\mu - \varepsilon \right\}$$

é sindético.

Demonstração. Assumimos, sem perda de generalidade, que $\|f\|_\infty = 1$. Então

$$\int_X f \cdot U_T^n f \cdots U_T^{kn} f d\mu \leq \|f\|_\infty^k = 1$$

para todos $k, n > 1$. Dado $\varepsilon > 0$, pelo Corolário 3.3.1 sabemos que o conjunto $A = A_{\varepsilon, k} = \{n \in \mathbb{N} : \|U_T^n f - f\| < \varepsilon/k^2\}$ é sindético. Deste modo, para $n \in A$ e $1 \leq i \leq k$, temos, pela desigualdade triangular e pelo fato de U_T ser unitário, que

$$\|U_T^{ni} f - f\| \leq \sum_{j=1}^i \|U_T^{jn} - U_T^{(j-1)n}\| = \sum_{j=1}^i \|U_T^n f - f\| < i \cdot \frac{\varepsilon}{k^2} \leq \frac{\varepsilon}{k}.$$

Provaremos por indução em i que para $2 \leq i \leq k$ e $n \in A$ vale que

$$\int_X \prod_{j=0}^k U_T^{jn} f d\mu > \int_X f^i \prod_{j=1}^{k-(i-1)} U_T^{jn} f d\mu - \frac{\varepsilon(i-1)}{k}.$$

A idéia é utilizar a desigualdade de Cauchy-Schwarz repetidas vezes. De fato, note que

$$\begin{aligned} \int_X \prod_{j=0}^k U_T^{jn} f d\mu &= \int_X f \prod_{j=1}^k U_T^{jn} f d\mu = \int_X (f - U_T^n f + U_T^n f) \prod_{j=1}^k U_T^{jn} f d\mu \\ &= \int_X U_T^n f^2 \prod_{j=2}^k U_T^{jn} f d\mu + \int_X (f - U_T^n f) \prod_{j=1}^k U_T^{jn} f d\mu \\ &\geq \int_X f^2 \prod_{j=1}^{k-1} U_T^{jn} f d\mu - \|U_T^n f - f\| \cdot \left\| \prod_{j=1}^{k-1} U_T^{jn} f \right\| \\ &\geq \int_X f^2 \prod_{j=1}^{k-1} U_T^{jn} f d\mu - \frac{\varepsilon}{k} \prod_{j=1}^{k-1} \|U_T^{jn} f\|_\infty \\ &\geq \int_X f^2 \prod_{j=1}^{k-1} U_T^{jn} f d\mu - \frac{\varepsilon}{k} \|f\|_\infty^k \\ &\geq \int_X f^2 \prod_{j=1}^{k-1} U_T^{jn} f d\mu - \frac{\varepsilon}{k}. \end{aligned}$$

Ademais, para $2 \leq i \leq k$,

$$\begin{aligned}
\int_X f^i \prod_{j=1}^{k-(i-1)} U_T^{jn} f d\mu &= \int_X f^i \left(f - f + U_T^{[k-(i-1)n]} f \right) \prod_{j=1}^{k-i} U_T^{jn} f d\mu \\
&= \int_X f^{i+1} \prod_{j=1}^{k-i} U_T^{jn} f d\mu - \int_X f^i \left(f - U_T^{[k-(i-1)n]} f \right) \prod_{j=1}^{k-i} U_T^{jn} f d\mu \\
&\geq \int_X f^{i+1} \prod_{j=1}^{k-i} U_T^{jn} f d\mu - \|U_T^{[k-(i-1)n]} - f\| \cdot \left\| \prod_{j=1}^{k-i} U_T^{jn} f d\mu \right\| \\
&\geq \int_X f^{i+1} \prod_{j=1}^{k-i} U_T^{jn} f d\mu - \frac{\varepsilon}{k}.
\end{aligned}$$

Deste modo, assumindo a hipótese de indução para $2 \leq i < k$, e o que provamos acima, temos

$$\int_X \prod_{j=0}^k U_T^{jn} f d\mu \geq \int_X f^i \prod_{j=1}^{k-(i-1)} U_T^{jn} f d\mu - \frac{\varepsilon(i-1)}{k} \geq \int_X f^{i+1} \prod_{j=1}^{k-i} U_T^{jn} f d\mu - \frac{\varepsilon \cdot i}{k},$$

o que prova a afirmação.

Para $i = k$, temos $\prod_{j=1}^{k-(i-1)} U_T^{jn} f = f$, e a afirmação que foi provada via indução para $i = k$ nos dá

$$\int_X \prod_{j=0}^k U_T^{jn} f d\mu > \int_X f^k \cdot f d\mu - \frac{\varepsilon(k-1)}{k} = \int_X f^{k+1} d\mu - \varepsilon + \varepsilon/k > \int_X f^{k+1} d\mu - \varepsilon,$$

para todo $n \in A$. Isto conclui a prova da proposição. \square

Enunciamos agora, sem prova, três resultados que nos permitirão concluir o Teorema 3.3.1. Embora as provas não sejam difíceis, preferimos omitir para não alongar o texto. Os dois primeiros resultados são técnicos, enquanto o terceiro é essencial e representa uma manifestação de uma dicotomia muito presente em Teoria Ergódica de Ramsey, entre estruturas regulares e estruturas aleatórias. Aqui, a estrutura regular é representada por elementos compactos e a aleatória por elementos fracamente misturadores.

Lema 3.3.1 (Truque de van der Corput). *Seja $(x_n)_n$ uma sequência limitada em um espaço de Hilbert H tal que*

$$\limsup_{N \rightarrow \infty} \frac{1}{H} \sum_{h=1}^H \frac{1}{N} \sum_{n=1}^N \langle x_{n+h}, x_n \rangle = 0.$$

Então

$$\left\| \frac{1}{N} \sum_{n=1}^N x_n \right\| \rightarrow 0.$$

Abaixo, apresentamos uma caracterização para H_{wm} .

Proposição 3.3.4. *O espaço H_{wm} é igual a*

$$H_{wm} = \left\{ x \in H : \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N |\langle U^n x, y \rangle| = 0, \forall y \in H \right\}.$$

Teorema 3.3.2 (Koopman, von Neumann). $H = H_c \oplus^\perp H_{wm}$.

Seja $H = L_\mu^2(X)$. Com base neste último teorema, dado $f \in L_\mu^\infty(X) \subset L_\mu^2(X)$, podemos escrever $f = f_c + f_{wm}$ com $f_c \in H_c$ e $f_{wm} \in H_{wm}$. Por linearidade,

$$\int_X f \cdot U_T^n f \cdot U_T^{2n} f d\mu = \sum_{i,j,k \in \{c,wm\}} f_i \cdot U_T^n f_j \cdot U_T^{2n} f_k d\mu.$$

A próxima proposição diz que, a fim de provar o Teorema 3.3.1, podemos supor $f = f_c$.

Proposição 3.3.5. *Se pelo menos um dos $f, g, h \in L_\mu^2(X)$ é fracamente misturador, então*

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \int_X f \cdot U_T^n g \cdot U_T^{2n} h d\mu = 0.$$

Demonstração. Assumiremos, sem perda de generalidade, duas hipóteses. A primeira é que μ é ergódica. Se este não for o caso, basta aplicarmos o resultado para cada componente ergódica de μ . A segunda é que podemos assumir g ou h fracamente misturador. De fato, como T preserva a medida, se f é fracamente misturador, então

$$\int_X f \cdot U_T^n g \cdot U_T^{2n} h d\mu = \int_X h \cdot U_T^{-n} g \cdot U_T^{-2n} f d\mu$$

e daí basta provar a proposição para U_T^{-1} . Assumindo estas duas hipóteses, provaremos que

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N U_T^n g U_T^{2n} h = 0 \tag{3.1}$$

em $L_\mu^2(X)$. Uma vez provado isto, seguirá que

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \int_X f \cdot U_T^n g \cdot U_T^{2n} h d\mu = \left\langle f, \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N U_T^n g \cdot U_T^{2n} h \right\rangle = \langle f, 0 \rangle = 0,$$

o que estabelece a proposição.

Para provar (3.1), usaremos o truque de Van der Corput (Lema 3.3.1) e a caracterização de H_{wm} (Proposição 3.3.4). Seja $f_n = U_T^n g U_T^{2n} h$. Temos

$$\begin{aligned} \langle f_{n+m}, f_n \rangle &= \int_X U_T^{n+m} g U_T^{2n+2m} h \cdot \overline{U_T^n g U_T^{2n} h} d\mu \\ &= \int_X U_T^m g U_T^{n+2m} h \cdot \overline{U_T^n h} d\mu \\ &= \int_X \overline{g} U_T^m g \cdot \overline{U_T^{n+2m} h U_T^n h} d\mu \\ &= \int_X \overline{g} U_T^m g \cdot \overline{U_T^n (h U_T^{2m} h)} d\mu \\ &= \langle \overline{g} U_T^m g, U_T^n (h U_T^{2m} h) \rangle. \end{aligned}$$

Pela bilinearidade e continuidade do produto interno,

$$\begin{aligned} y_m &:= \lim_{n \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \langle f_{n+m}, f_n \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \langle \overline{g} U_T^m g, U_T^n (h U_T^{2m} h) \rangle \\ &= \left\langle \overline{g} U_T^m g, \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N U_T^n (h U_T^{2m} h) \right\rangle = \left\langle \overline{g} U_T^m g, \int_X h U_T^{2m} h d\mu \right\rangle \\ &= \int_X h U_T^{2m} h d\mu \cdot \langle \overline{g} U_T^m g, 1 \rangle = \int_X \overline{g} U_T^m g d\mu \cdot \int_X h U_T^{2m} h d\mu = \langle U_T^m g, g \rangle \cdot \langle h, U_T^{2m} h \rangle. \end{aligned}$$

Consideramos dois casos. Primeiro, assuma que g é fracamente misturador. Pela desigualdade de Cauchy-Schwarz, $|\langle h, U_T^{2m} h \rangle| \leq \|h\| \cdot \|U_T^{2m} h\| = \|h\|^2$ para todo m , e daí

$$\begin{aligned} \lim_{M \rightarrow \infty} \left| \frac{1}{M} \sum_{m=1}^M y_m \right| &\leq \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M |y_m| = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M |\langle U_T^m g, g \rangle \cdot \langle h, U_T^{2m} h \rangle| \\ &\leq \|h\|^2 \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M |\langle U_T^m g, g \rangle| = 0, \end{aligned}$$

pela Proposição 3.3.4.

No segundo caso, assuma h fracamente misturador. Por Cauchy-Schwarz, temos

$$|\langle U_T^m g, g \rangle| \leq \|U_T^m g\| \cdot \|g\| = \|g\|^2 \text{ e daí}$$

$$\begin{aligned} \lim_{M \rightarrow \infty} \left| \frac{1}{M} \sum_{m=1}^M y_m \right| &\leq \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M |y_m| = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M |\langle U_T^m g, g \rangle \cdot \langle h, U_T^{2m} h \rangle| \\ &\leq \|g\|^2 \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M |\langle h, U_T^{2m} h \rangle| \leq 2\|g\|^2 \lim_{M \rightarrow \infty} \frac{1}{2M} \sum_{m=1}^{2M} |\langle h, U_T^m h \rangle| = 0, \end{aligned}$$

onde na última passagem utilizamos novamente a Proposição 3.3.4.

Assim, $\lim_{M \rightarrow \infty} \left| \frac{1}{M} \sum_{m=1}^M y_m \right| = 0$ e portanto, pelo truque de van der Corput, obtemos que

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N U_T^n g U_T^{2n} h = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N f_n = 0,$$

o que conclui a prova. \square

Demonstração do Teorema 3.3.1. Seja $f \in L_\mu^\infty(X)$ com $f \geq 0$ e $\int_X f d\mu > 0$. Escreva $f = f_c + f_{wm}$, com $f_c \in H_c$ e $f_{wm} \in H_{wm}$. Pela Proposição 3.3.5,

$$\liminf_{n \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \int_X f \cdot U_T^n f \cdot U_T^{2n} f d\mu = \liminf_{n \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \int_X f_c \cdot U_T^n f_c \cdot U_T^{2n} f_c d\mu.$$

Como f_c é a projeção ortogonal de f em H_T e esta projeção é positiva (pois é uma esperança condicional), segue que $f_c \geq 0$. Ademais, cada função constante é compacta e daí, pelo Teorema 3.3.2, temos que $\int_X f_{wm} d\mu = 0$. Portanto, $\int_X f_c d\mu = \int_X f d\mu > 0$. Em particular, $\int_X f_c^3 d\mu > 0$. Escolha $\varepsilon > 0$ tal que $\int_X f_c^3 d\mu > \varepsilon$.

Pela Proposição 3.3.3, o conjunto

$$A = \left\{ n \in \mathbb{N} : \left| \int_X f_c \cdot U_T^n f_c \cdot U_T^{2n} f_c d\mu - \int_X f_c^3 d\mu \right| < \varepsilon \right\}$$

é sindético. Logo,

$$\begin{aligned} \frac{1}{N} \sum_{n=1}^N \int_X f \cdot U_T^n f \cdot U_T^{2n} f d\mu &\geq \frac{1}{N} \sum_{\substack{1 \leq n \leq N \\ n \in A}} \int_X f \cdot U_T^n f \cdot U_T^{2n} f d\mu \\ &\geq \frac{1}{N} \sum_{\substack{1 \leq n \leq N \\ n \in A}} \left(\int_X f_c^3 d\mu - \varepsilon \right) = \frac{|A \cap \{1, \dots, N\}|}{N} \left(\int_X f_c^3 d\mu - \varepsilon \right) > 0 \end{aligned}$$

e então o limsup das expressões acima são positivos, pois A é sindético. Isto conclui a prova do teorema. \square

4 TEOREMA DE HINDMAN

Seguindo os princípios da Teoria de Ramsey, consideramos outro tipo de estrutura: as somas finitas de uma sequência, ou conjuntos IP. Estes conjuntos foram vistos no Capítulo 2 (veja a Definição 2.1.1), quando estudamos propriedades de recorrência. Em (HINDMAN, 1974), Hindman mostrou que em qualquer partição finita dos números naturais pelo menos um dos membros da partição contém um conjunto IP. O intuito deste capítulo é provar este teorema, utilizando a técnica de *ultrafiltros*. Mais especificamente, provaremos o Teorema de Hindman utilizando um Teorema de Ellis sobre a existência de elementos idempotentes, aplicado à compactificação de Stone-Čech.

No que segue, faremos uma introdução sobre ultrafiltros e p -limites, mostrando uma dualidade entre estes conceitos. Após, provaremos o Teorema de Ellis, e por fim, o Teorema de Hindman.

4.1 Ultrafiltros

Filtros e ultrafiltros são formas de introduzir noções de grandeza em uma família de subconjuntos de \mathbb{N} (ou, mais geralmente, de qualquer conjunto). Introduzimos tal noção a partir de exemplos de sequências.

Dada $(x_n)_{n \geq 1}$ uma sequência de números reais, sabemos que ela converge para $a \in \mathbb{R}$ se para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que $x_n \in (a - \varepsilon, a + \varepsilon)$ para todo $n > N$. Em outras palavras, a menos de uma quantidade finita de índices n , todo x_n está próxima de a . Assim, dizer que $(x_n)_n$ converge para a significa que a é o valor que mais agrada, a menos de qualquer erro $\varepsilon > 0$, a grande maioria (exceto uma quantidade finita) dos termos da sequência.

Por outro lado, a sequência $0, 1, 0, 1, 0, 1, \dots$ não possui limite. Na verdade, ela possui dois “limites”, 0 e 1, pois cada um destes valores agrada metade dos termos da sequência.

Podemos então definir uma noção de “grandeza” associada à existência de limites, que consiste em checar quando um valor agrada todos os termos da sequência com a exceção de uma quantidade finita. Para isso, listamos abaixo propriedades usuais de limites, a fim de formalizar a ideia de grandeza.

Sejam $(x_n), (y_n)$ sequências de números reais e $\lambda \in \mathbb{R}$. A operação de limite

satisfaz as seguintes propriedades:

L.1. Compatibilidade aritmética:

$$\lim(x_n + y_n) = \lim x_n + \lim y_n \quad \text{e} \quad \lim(x_n y_n) = (\lim x_n)(\lim y_n).$$

L.2. $\lim(\lambda x_n) = \lambda \lim x_n$.

L.3. Limite é unitário: $\lim 1 = 1$.

L.4. $\inf\{x_n\}_{n \in \mathbb{N}} \leq \lim x_n \leq \sup\{x_n\}_{n \in \mathbb{N}}$.

L.5. Não-principalidade (a minoria não altera o resultado do limite): se $x_n \neq y_n$ para finitos valores de n , então $\lim x_n = \lim y_n$.

L.6. Invariância pelo shift: $\lim_{n \rightarrow \infty} x_{n+m} = \lim_{n \rightarrow \infty} x_n$.

Outra forma de ver as propriedades acima é a seguinte: considere o conjunto $E = \{(x_n)_n : \lim x_n \text{ existe}\}$ munido da estrutura de espaço vetorial (soma e multiplicação por escalar) e da norma do supremo, que torna E um espaço de Banach. Defina o operador linear $\lim : E \rightarrow \mathbb{R}$ que consiste em calcular o limite da sequência. As propriedades acima nos permitem inferir que \lim é um funcional linear unitário que satisfaz L.4, L.5 e L.6. A propriedade L.4 nos dá uma cota grosseira para o valor do limite, L.5 nos diz que o valor do limite depende apenas da “maioria” dos elementos da sequência e L.6 nos diz que este valor é invariante por shifts.

A partir de agora, definiremos novos operadores abdicando da propriedade L.5 e L.6. Seja então $\mathbb{R}_0^{\mathbb{N}} = l_{\infty}(\mathbb{N}) = \{(x_n)_{n \in \mathbb{N}} \subset \mathbb{R} : (x_n) \text{ é limitada}\}$. Denotaremos por $p\text{-}\lim : \mathbb{R}_0^{\mathbb{N}} \rightarrow \mathbb{R}$ qualquer funcional linear contínuo satisfazendo L.1–L.4. Chamamos este funcional de um $p\text{-limite}$. Quando este funcional satisfaz L.5, dizemos que o $p\text{-limite}$ é *não-principal*. A motivação por trás deste enfraquecimento é para que sequências como $0, 1, 0, 1, \dots$ possam possuir limite. De fato, a exemplo desta sequência, duas cópias de \mathbb{N} , a saber, $2\mathbb{N}$ e $2\mathbb{N} + 1$, são conjuntos grandes, a menos de homotetia. Os $p\text{-limites}$ também levam o nome de *limites de Banach*.

Exemplo. Dada $(x_n)_n \in \mathbb{R}_0^{\mathbb{N}}$, considere

$$f((x_n)_n) = \limsup_{n \rightarrow \infty} \frac{|x_1| + \dots + |x_n|}{n},$$

que define uma seminorma. O limite usual de sequências é dominado por esta seminorma. Pelo Teorema de Hahn-Banach, conseguimos estender o limite para $\mathbb{R}_0^{\mathbb{N}}$, que continua

dominado pela seminorma. Esta extensão é um exemplo de p -limite.

Com esta nova noção de limite, conseguimos introduzir uma noção de grandeza em \mathbb{N} , que de fato caracteriza totalmente o p -limite. Primeiro, definiremos esta noção e, posteriormente, mostraremos como caracterizar o p -limite a partir dela. Dado $A \subset \mathbb{N}$, lembre que $\mathbb{1}_A$ denota a função característica de A , que é uma sequência definida por $\mathbb{1}_A(n) = 1$ se $n \in A$ e 0 caso contrário. Seja

$$p = \{A \subset \mathbb{N} : p\text{-lim}(\mathbb{1}_A) = 1\}.$$

Então p coleciona todos os $A \subset \mathbb{N}$ com uma quantidade de elementos suficientes a ponto de $p\text{-lim}(\mathbb{1}_A)$ ser igual a 1. A seguir, listamos as propriedades que p satisfaz.

F.1 $\emptyset \notin p$.

De fato, $\mathbb{1}_{\emptyset}(n) = 0$ para todo $n \in \mathbb{N}$ e daí L.2 implica que

$$p\text{-lim} \mathbb{1}_{\emptyset} = p\text{-lim}(0 \cdot \mathbb{1}_{\emptyset}) = 0 \cdot (p\text{-lim} \mathbb{1}_{\emptyset}) = 0.$$

F.2 $A \in p$ e $B \supset A \implies B \in p$.

A intuição deste item é: se A é grande então todo $B \supset A$ também é grande. Para provar este item, note que $\mathbb{1}_B = \mathbb{1}_A + \mathbb{1}_{B \setminus A}$. Por L.4, temos que $p\text{-lim} \mathbb{1}_B \leq \sup\{\mathbb{1}_B(n)\} = 1$. Por outro lado, L.1 implica que

$$p\text{-lim} \mathbb{1}_B = p\text{-lim} \mathbb{1}_A + p\text{-lim} \mathbb{1}_{B \setminus A} = 1 + p\text{-lim} \mathbb{1}_{B \setminus A} \geq 1,$$

donde $p\text{-lim} \mathbb{1}_B = 1$.

F.3 $A, B \in p \implies A \cap B \in p$.

Para ver essa propriedade, usaremos que $\mathbb{1}_{A \cup B} + \mathbb{1}_{A \cap B} = \mathbb{1}_A + \mathbb{1}_B$. Por F.2, temos $A \cup B \in p$. Daí, passando o p -limite, concluímos que

$$1 + p\text{-lim} \mathbb{1}_{A \cap B} = 2 \implies p\text{-lim} \mathbb{1}_{A \cap B} = 1.$$

F.4: Se $A \subset \mathbb{N}$ então $A \in p$ ou $A^c \in p$, excludentemente.

Ora, como $\mathbb{1}_A + \mathbb{1}_{A^c} = 1$ e $\mathbb{1}_A \cdot \mathbb{1}_{A^c} = 0$, segue de L.1 que

$$p\text{-lim} \mathbb{1}_A + p\text{-lim} \mathbb{1}_{A^c} = 1 \quad \text{e} \quad (p\text{-lim} \mathbb{1}_A)(p\text{-lim} \mathbb{1}_{A^c}) = 0,$$

o que obviamente implica que $(a, b) = (0, 1)$ ou $(1, 0)$. Note que apenas um dentre os conjuntos A , A^c pertence a p .

F.4' Se $\mathbb{N} = A_1 \cup \dots \cup A_r$, então $A_i \in p$ para exatamente um índice $1 \leq i \leq r$.

Provaremos esta propriedade mostrando que ela equivale a F.4. É claro que F.4' implica F.4. Para a outra implicação, assumamos por contradição que $A_1, \dots, A_r \notin p$. Então $A_1^c, \dots, A_r^c \in p$, e daí F.3 implica que a interseção $\emptyset = A_1^c \cap \dots \cap A_r^c \in p$, o que contraria F.1. Por fim, note que se $A_i, A_j \in p$ para $i \neq j$, então $\emptyset = A_i \cap A_j \in p$ por F.3, novamente uma contradição.

F.4'' Se $A \in p$ e $A = B \cup C$ então $B \in p$ ou $C \in p$.

Assim como em F.4, não podemos ter ambos $B, C \in p$. Para ver esta propriedade, basta notar que $\mathbb{1}_A = \mathbb{1}_B + \mathbb{1}_C$ e $\mathbb{1}_B \cdot \mathbb{1}_C = 0$ e repetir os passos da prova de F.4.

F.5 Não-principalidade: $A \in p$ e $B \subset \mathbb{N}$ finito $\implies A \setminus B \in p$.

Neste contexto, $\mathbb{1}_A(n) \neq \mathbb{1}_{A \setminus B}(n)$ apenas no conjunto finito de índices $n \in B$, e portanto L.5 implica que $p\text{-}\lim \mathbb{1}_{A \setminus B} = p\text{-}\lim \mathbb{1}_A = 1$.

Definição 4.1.1 (Filtro e ultrafiltro). *Dizemos que $p \subset \mathcal{P}(\mathbb{N})$ é um filtro se satisfizer as propriedades F.1, F.2, F.3. Se adicionalmente p satisfizer F.4, dizemos que p é um ultrafiltro. Um ultrafiltro que satisfaz F.5 é chamado ultrafiltro não-principal.*

É importante observar que a noção de ultrafiltro nos dá uma noção de grandeza de conjuntos, por meio das propriedades F.1–F.3. Já a propriedade F.4 é uma propriedade maximal: em uma partição de \mathbb{N} , exatamente um dos conjuntos está em p . Dito isto, é natural considerar a seguinte definição.

Definição 4.1.2 (Conjunto p -grande). *Seja $p \subset \mathcal{P}(\mathbb{N})$ um ultrafiltro. Dizemos que $A \subset \mathbb{N}$ é p -grande quando $A \in p$. Seja \mathcal{P} uma propriedade e $(A_n)_{n \geq 1}$ uma sequência de subconjuntos de \mathbb{N} . Dizemos que \mathcal{P} vale para p -quase todo $n \in \mathbb{N}$ quando $\{n \in \mathbb{N} : A_n \text{ possui } \mathcal{P}\} \in p$.*

Exemplos.

1. Um p -limite não principal define um ultrafiltro não-principal. A recíproca é verdadeira: um ultrafiltro não-principal induz uma noção de p -limite. Veremos isto mais adiante, onde construímos explicitamente a noção de limite.
2. Para cada $n \in \mathbb{N}$, a família $p_n = \{A \subset \mathbb{N} : n \in A\}$ é um ultrafiltro. A prova é simples, e deixamos a cargo do leitor.

Lembre-se de que um ultrafiltro é não-principal quando ele satisfizer F.5. O ultrafiltro p_n do último exemplo é principal, ou seja, não satisfaz F.5. De fato, para todo

$A \in p_n$, temos que $A \setminus \{n\} \notin p_n$. A nomenclatura vem do conceito de ideais principais em anéis, que são aqueles gerados por um único elemento. A próxima proposição mostra que não existem ultrafiltros principais além de p_n , $n \geq 1$.

Proposição 4.1.1. *Um ultrafiltro p é principal se e somente se $p = p_n$ para algum $n \in \mathbb{N}$.*

Demonstração. Basta provar a ida. Como p é principal, existem $A \in p$ e $B \subset \mathbb{N}$ finito tais que $A \setminus B = A \cap B^c \notin p$. Logo, $(A \cap B^c)^c = A^c \cup B \in p$ por F.4, donde $A \cap B = A \cap (A^c \cup B) \in p$ por F.3 e então $B \in p$ por F.2. Noutras palavras, encontramos B finito que está em p . Por F.4", existe $n \in B$ tal que $\{n\} \in p$. Agora, fica fácil concluir que $p = p_n$. Por um lado, F.2 implica que $p_n \subset p$. Por outro lado, se $p \not\subset p_n$, então existe $A \in p$ tal que $n \notin A$, isto é, $n \in A^c$; por F.2, $A^c \in p$ e assim $A, A^c \in p$, o que contradiz F.4. A proposição está provada. \square

4.2 Construção de p -lim a partir de ultrafiltro

Finalizamos esta seção mostrando como definir uma noção de limite a partir de um ultrafiltro. Recordemos da definição clássica de limite para sequências de números reais: dada $(x_n)_n \subset \mathbb{R}$, escrevemos $\lim x_n = x$ quando, para todo aberto U contendo x , existe $N \in \mathbb{N}$ tal que $x_n \in U$ para todo $n \geq N$, ou seja, $\{n \in \mathbb{N} : x_n \in U\} \supset \{N, N+1, \dots\}$. Este último conjunto é "grande" por ser o complementar de um conjunto finito. Nossa ideia é trocar $\{N, N+1, \dots\}$ pelos conjuntos de um ultrafiltro p . Seja então $p \subset \mathcal{P}(\mathbb{N})$ um ultrafiltro. Pela Proposição 4.1.1, p não contém nenhum subconjunto finito quando for não-principal.

Definição 4.2.1 (p -limite definido por ultrafiltro). *Sejam $p \subset \mathbb{N}$ um ultrafiltro e (X, d) um espaço métrico. Uma sequência $(x_n)_n \subset X$ é dita p -convergente a $x \in X$ se, para todo aberto U que contém x , tem-se $\{n \in \mathbb{N} : x_n \in U\} \in p$. Neste caso, escrevemos $p\text{-lim } x_n = x$.*

Note que a definição está bem posta. Para ver isso, assumamos que $p\text{-lim } x_n = x$ e $p\text{-lim } x_n = y$ para $x, y \in X$ distintos. Como X é Hausdorff, existem $U, V \subset X$ abertos disjuntos tais que $x \in U$ e $y \in V$. Mas então, como $A_x = \{n \in \mathbb{N} : x_n \in U\} \in p$ e $A_y = \{n \in \mathbb{N} : x_n \in V\} \in p$, segue que $A_x \cap A_y \in p$ e portanto é não-vazio, uma contradição.

Nos restringimos ao caso $X = \mathbb{R}_0^{\mathbb{N}}$, com métrica induzida pela norma do máximo. Mais uma vez, seja $p \subset \mathbb{N}$ ultrafiltro. O objetivo é mostrar que para todo $(x_n)_n \in \mathbb{R}_0^{\mathbb{N}}$, $p\text{-lim } x_n$ existe.

Proposição 4.2.1. *Seja $p \subset \mathbb{N}$ ultrafiltro e $(x_n)_n \in \mathbb{R}_0^{\mathbb{N}}$. Então $p\text{-}\lim x_n$ existe.*

Demonstração. Se $p = p_m$ é principal, então $p\text{-}\lim x_n = x_m$. Assuma portanto que p é não-principal. Dada $(x_n) \in \mathbb{R}_0^{\mathbb{N}}$, existe um intervalo limitado $I_0 = [a_0, b_0]$ tal que $x_n \in I_0$ para todo $n \in \mathbb{N}$. Considere $I_1^1 = [a_0, (a_0 + b_0)/2]$ e $I_1^2 = ((a_0 + b_0)/2, b_0]$. Então $I_0 = I_1^1 \cup I_1^2$, onde $|I_1^i| = \frac{1}{2}|I_0|$, $i = 1, 2$. Sejam

$$A_1^1 = \{n \in \mathbb{N} : x_n \in I_1^1\}, \quad A_1^2 = \{n \in \mathbb{N} : x_n \in I_1^2\}.$$

Temos $A_1^1 \cup A_1^2 = \mathbb{N}$, e portanto exatamente um destes conjuntos está em p . Sejam $A_1 \in p$ este conjunto $I_1 \in \{I_1^1, I_1^2\}$ o intervalo correspondente. Agora dividimos I_1 em dois intervalos de mesmo comprimento, e encontramos um destes intervalos I_2 tal que $A_2 := \{n \in \mathbb{N} : x_n \in I_2\} \in p$. Procedendo por indução, definimos intervalos $I_0 \supset I_1 \supset \dots$ tais que $|I_i| = \frac{1}{2^i}|I_0|$ e $A_i = \{n \in \mathbb{N} : x_n \in I_i\} \in p$.

Pelo Teorema dos Intervalos encaixados, $\bigcap_{k=0}^{\infty} \overline{I_k} = \{x\}$ para algum $x \in \mathbb{R}$. Afir-mamos que $p\text{-}\lim x_n = x$. De fato, dado U aberto contendo x , para i grande temos $I_i \subset U$ e portanto $A_i \subset \{n \in \mathbb{N} : x_n \in U\}$. Como $A_i \in p$, segue de F.2 que $\{n \in \mathbb{N} : x_n \in U\} \in p$. \square

Por fim, mostramos que $p\text{-}\lim$ definido acima é um $p\text{-}\limite$ no sentido das propriedades L.1–L.4.

Proposição 4.2.2. *Seja p um ultrafiltro e $p\text{-}\lim : \mathbb{R}_0^{\mathbb{N}} \rightarrow \mathbb{R}$ como na Proposição 4.2.1. Então $p\text{-}\lim$ satisfaz as propriedades L.1–L.4.*

Demonstração. Checamos cada propriedade separadamente.

L.1. Sejam $(x_n), (y_n) \in \mathbb{R}_0^{\mathbb{N}}$ e escreva $a = p\text{-}\lim x_n$ e $b = p\text{-}\lim y_n$. Fixe $\varepsilon > 0$. Por definição, cada um dos conjuntos $A = \{n \in \mathbb{N} : |x_n - a| < \varepsilon/2\}$ e $B = \{n \in \mathbb{N} : |y_n - b| < \varepsilon/2\}$ está em p e portanto $A \cap B \in p$. Como $n \in A \cap B \Rightarrow |(x_n + y_n) - (a + b)| < \varepsilon$, segue de F.2 que $\{n \in \mathbb{N} : |(x_n + y_n) - (a + b)| < \varepsilon\} \in p$. Como $\varepsilon > 0$ é arbitrário, concluímos que $p\text{-}\lim(x_n + y_n) = a + b$.

Para ver a compatibilidade com o produto, tome $C > 0$ tal que $|x_n| \leq C$ para todo n . Separamos a verificação em dois casos: $b = 0$ e $b \neq 0$. No primeiro caso, temos da definição que para todo $\varepsilon > 0$ vale que $\{n \in \mathbb{N} : |y_n| < \varepsilon/C\} \in p$. Para n neste conjunto, $|x_n y_n| \leq C|y_n| < \varepsilon$ e portanto $p\text{-}\lim(x_n y_n) = 0$. Assuma agora que $b \neq 0$. Os conjuntos $D = \{n \in \mathbb{N} : |y_n - b| < \varepsilon/2C\}$ e $E = \{n \in \mathbb{N} : |x_n - a| < \varepsilon/2b\}$ estão em p , e portanto

$D \cap E \in p$. Para $n \in D \cap E$, temos

$$|x_n y_n - ab| \leq |x_n| \cdot |y_n - b| + |b| \cdot |x_n - a| \leq C \frac{\varepsilon}{2C} + b \frac{\varepsilon}{2b} = \varepsilon.$$

e portanto, novamente por F.2, $\{n \in \mathbb{N} : |x_n y_n - ab| < \varepsilon\} \in p$.

L.2. Claramente $p\text{-lim } \lambda = \lambda$. Por L.1, segue que $p\text{-lim}(\lambda x_n) = \lambda \cdot p\text{-lim } x_n$.

L.3. Claro.

L.4. Sejam $c = \inf\{x_n : n \in \mathbb{N}\}$ e $C = \sup\{x_n : n \in \mathbb{N}\}$. Se $c = C$, não há nada a fazer. Suponha $c < C$. Repetindo a demonstração da proposição 4.2.1 com $I_0 = [c, C]$, segue que $p\text{-lim } x_n \in [c, C]$. \square

É fácil ver que $p\text{-lim}$ satisfaz L.5 se e somente se p é não-principal. Assim, obtemos uma dualidade entre ultrafiltros e p -limites, que se restringe a uma dualidade entre ultrafiltros não-principais e p -limites satisfazendo L.5.

Exemplo. Considere o p -limite de $\mathbb{R}_0^{\mathbb{N}}$ obtido ao estender o limite usual de sequências para $\mathbb{R}_0^{\mathbb{N}}$ com respeito à seminorma $f((x_n)_n) = \limsup_n \left(\frac{|x_1| + \dots + |x_n|}{n} \right)$, e seja p o ultrafiltro associado a este p -limite. Pela dualidade, este ultrafiltro induz a noção usual de convergência de sequências.

4.3 Compactificação de stone-čech

Definição 4.3.1 (Compactificação de Stone-Čech). *A compactificação de Stone-Čech de \mathbb{N} , denotada por $\beta\mathbb{N}$, é o conjunto dos ultrafiltros $p \subset \mathcal{P}(\mathbb{N})$:*

$$\beta\mathbb{N} = \{p \subset \mathcal{P}(\mathbb{N}) : p \text{ é ultrafiltro}\}.$$

É fácil ver que $\beta\mathbb{N}$ contém naturalmente \mathbb{N} . Para cada $n \in \mathbb{N}$, seja $p_n = \{A \subset \mathbb{N} : n \in A\}$ o ultrafiltro principal. A inclusão $i : \mathbb{N} \rightarrow \beta\mathbb{N}$ dada por $i(n) = p_n$ é claramente injetiva. Note que para todo $(x_k)_k \in \mathbb{R}_0^{\mathbb{N}}$, vale que $p_n\text{-lim}(x_k) = x_n$. Isto diz que $p_n\text{-lim}$ associado depende apenas do termo x_n da sequência.

Para nossos fins, dotamos $\beta\mathbb{N}$ de uma estrutura algébrica e topológica. Começamos definindo soma em $\beta\mathbb{N}$. Dados $p, q \in \beta\mathbb{N}$, defina

$$p + q = \{A \subset \mathbb{N} : \{n : A - n \in p\} \in q\},$$

isto é, $A \in p+q$ se e somente se o conjunto dos $n \in \mathbb{N}$ tais que as translações $A - n$ estão em p está em q . Em particular,

$$\begin{aligned} p_n + p_m &= \{A \subset \mathbb{N} : \{k \in \mathbb{N} : A - k \in p_n\} \in p_m\} = \{A \subset \mathbb{N} : A - n \in p_m\} \\ &= \{A \subset \mathbb{N} : n + m \in A\} = p_{n+m}. \end{aligned}$$

Deste modo, $(\beta\mathbb{N}, +)$ é uma extensão de $(\mathbb{N}, +)$. Por outro lado, $(\beta\mathbb{N}, +)$ é altamente não-comutativo. Este fenômeno pode ser observado pelo modo como os p -limites se comportam por composição.

Proposição 4.3.1. *Sejam $p, q \in \beta\mathbb{N}$ e $(x_n)_n \in \mathbb{R}_0^{\mathbb{N}}$. Então*

$$(q+p) - \lim_n x_n = p - \lim_m \left(q - \lim_l x_{l+m} \right).$$

Como p -limite em geral não é invariante pelo shift, em geral temos que $p+q \neq q+p$.

Demonstração. Chame $\bar{x} = (p+q) - \lim x_n$ e seja U vizinhança de \bar{x} . Então $A_U = \{n \in \mathbb{N} : x_n \in U\} \in p+q$. Pela definição de soma de ultrafiltros, o conjunto

$$B_U = \{t \in \mathbb{N} : A_U - t \in q\} = \{t \in \mathbb{N} : \{n \in \mathbb{N} : x_n \in U\} - t \in q\} = \{t \in \mathbb{N} : \{n \in \mathbb{N} : x_{n+t} \in U\} \in q\}$$

está em p . Queremos mostrar que $\bar{x} = p - \lim_m (q - \lim_l x_{l+m})$. Escrevendo $y_m = q - \lim_l x_{l+m}$, é suficiente provar que $\{m \in \mathbb{N} : y_m \in U\} \in p$. Afirmamos que $\{m \in \mathbb{N} : y_m \in U\} \supset B_U$ o que, por F.2, implicará no desejado.

Tome $m \in B_U$. Temos que $\{l \in \mathbb{N} : x_{l+m} \in U\} \in q$. Por contradição, assumamos que $y_m \notin U$. Então existe uma vizinhança $V \ni y_m$ tal que $V \cap U = \emptyset$. Pela definição de p -limite, temos $\{l \in \mathbb{N} : x_{l+m} \in V\} \in q$. Mas então $\emptyset = \{l \in \mathbb{N} : x_{l+m} \in V\} \cap \{l \in \mathbb{N} : x_{l+m} \in U\} \in q$, o que contradiz F.1. Assim, $y_m \in U$, como requerido. \square

4.4 Elementos idempotentes e o teorema de hindman

Dada a estrutura da soma em $\beta\mathbb{N}$, existem ultrafiltros especiais, que podem fornecer informações combinatórias interessantes. Um exemplo particular notável, que estudaremos agora, consiste nos elementos idempotentes.

Definição 4.4.1 (Elemento idempotente). *Dizemos que $p \in \beta\mathbb{N}$ é idempotente se $p+p=p$.*

Se $p \in \beta\mathbb{N}$ é idempotente, então

$$p - \lim_n x_n = p - \lim_m (p - \lim_l x_{l+m}),$$

isto é, o p -limite de (x_n) pode ser obtido tomando o p -limite dos p -limites das transladas da sequência. Do ponto de vista de conjuntos, temos que para todo $A \in p$ vale que $\{n \in \mathbb{N} : A - n \in p\} \in p$, ou seja, para um conjunto p -grande de naturais n o translado $A - n$ é p -grande, o que indica propriedades de recorrência interessantes. De fato, é a existência de elementos idempotentes que fornecerá a prova do Teorema de Hindman que apresentaremos adiante.

É claro que se p é idempotente, então p é não-principal. Caso contrário, teríamos $p_n = p = p + p = p_n + p_n = p_{2n}$, implicando que $n = 0$.

Não é claro que $\beta\mathbb{N}$ possui elementos idempotentes. Mostraremos inicialmente este fato, em grande generalidade, no contexto de semigrupos. Um semigrupo é um par $(G, *)$, onde G é um conjunto e $*$ é uma operação associativa sobre G . Fixado $a \in G$, denotamos por $L_a : G \rightarrow G$ a “multiplicação à esquerda” $L_a(g) = a * g$. Dada uma topologia em G , dizemos que $(G, *)$ é um semigrupo topológico à esquerda quando L_a é contínuo com respeito a esta topologia para todo $a \in G$. O teorema abaixo, provado por Ellis em (ELLIS, 1958), garante a existência de elementos idempotentes.

Teorema 4.4.1 (Ellis). *Se $(G, *)$ é um semigrupo topológico à esquerda, compacto e Hausdorff, então G possui um elemento idempotente.*

Iremos postergar a prova deste teorema para a parte final da seção, e antes iremos mostrar que $\beta\mathbb{N}$ satisfaz as condições por ele elencadas, introduzindo uma topologia que torna $\beta\mathbb{N}$ um semigrupo topológico à esquerda, compacto e Hausdorff. Dado $A \subset \mathbb{N}$, seja $A^L = \{p \in \beta\mathbb{N} : A \in p\}$. Seja $\tau = \{A^L : A \subset \mathbb{N}\}$.

Proposição 4.4.1. *A família τ é uma base de topologia para $\beta\mathbb{N}$.*

Demonstração. É claro que $\bigcup_{A \subset \mathbb{N}} A^L = \beta\mathbb{N}$. Observe que, dados $A, B \subset \mathbb{N}$ e $p \in \beta\mathbb{N}$, $A, B \in p \Leftrightarrow A \cap B \in p$. De fato, por F.3 temos $A, B \in p \Rightarrow A \cap B \in p$. Por outro lado, por F.2, $A \cap B \in p$ e $A, B \supset A \cap B \Rightarrow A, B \in p$. Desta forma,

$$\begin{aligned} A^L \cap B^L &= \{p \in \beta\mathbb{N} : A \in p\} \cap \{p \in \beta\mathbb{N} : B \in p\} \\ &= \{p \in \beta\mathbb{N} : A, B \in p\} = \{p \in \beta\mathbb{N} : A \cap B \in p\} = (A \cap B)^L, \end{aligned}$$

o que prova que τ é realmente uma base. □

Seja τ^L a topologia gerada por τ . A partir de agora, sempre que nos referenciarmos a $\beta\mathbb{N}$ como um espaço topológico, será em relação a esta topologia. Pela propriedade F.4, temos

$$(A^L)^c = \{p \in \beta\mathbb{N} : A \not\subseteq p\} = \{p \in \beta\mathbb{N} : A^c \in p\} = (A^c)^L,$$

e portanto todo A^L é aberto e fechado. Além do mais, segue das propriedades de base que todo fechado de $\beta\mathbb{N}$ pode ser escrito da forma $\bigcap_{i \in \mathcal{I}} (A_i)^L$.

Com esta topologia, $\beta\mathbb{N}$ é claramente Hausdorff: dados $p \neq q \in \beta\mathbb{N}$, existe $A \in p$ tal que $A \not\subseteq q$ e daí $A^L, (A^c)^L$ são abertos disjuntos de $\beta\mathbb{N}$ com $p \in A^L$ e $q \in (A^c)^L$.

Nosso próximo objetivo será mostrar que $\beta\mathbb{N}$ é compacto. Para isto, faremos uso do teorema abaixo, cuja prova pode ser encontrada em (MUNKRES, 2014, p. 169). Antes, relembremos um conceito usual visto em cursos de topologia.

Definição 4.4.2 (Propriedade de interseção finita). *Seja X um conjunto não-vazio e \mathcal{A} uma família de subconjuntos de X . Dizemos que \mathcal{A} tem a propriedade de interseção finita (PIF) se para qualquer família finita $A_1, \dots, A_n \in \mathcal{A}$ a interseção $A_1 \cap \dots \cap A_n$ for não-vazia.*

Teorema 4.4.2. *Um espaço topológico X é compacto se e somente se para toda coleção \mathcal{F} de fechados com PIF vale que $\bigcap_{F \in \mathcal{F}} F \neq \emptyset$.*

Proposição 4.4.2. *$\beta\mathbb{N}$ é compacto.*

Demonstração. Vamos verificar a condição do Teorema 4.4.2. Seja \mathcal{F} uma família de fechados com PIF. Começamos mostrando que podemos assumir que cada elemento de \mathcal{F} é da forma A^L . Para isso, escreva $\mathcal{F} = \{F_\lambda\}_{\lambda \in \Lambda}$ com $F_\lambda = \bigcap_{\gamma \in \Gamma_\lambda} (A_{\lambda, \gamma})^L$. Isto define naturalmente uma nova família de fechados $\tilde{\mathcal{F}} = \{(A_{\lambda, \gamma})^L : \lambda \in \Lambda \text{ e } \gamma \in \Gamma_\lambda\}$. Afirmamos que $\tilde{\mathcal{F}}$ possui PIF e que $\bigcap_{F \in \mathcal{F}} F = \bigcap_{\tilde{F} \in \tilde{\mathcal{F}}} \tilde{F}$. A segunda afirmação é claramente verdadeira pela associatividade da interseção, então focamos na primeira afirmação. Sejam $(\lambda_1, \gamma_1), \dots, (\lambda_n, \gamma_n)$ elementos com $\lambda_i \in \Lambda$ e $\gamma_i \in \Gamma_{\lambda_i}$. Como $(A_{\lambda_i, \gamma_i})^L \supset F_{\lambda_i}$ e \mathcal{F} possui PIF, segue que

$$(A_{\lambda_1, \gamma_1})^L \cap \dots \cap (A_{\lambda_n, \gamma_n})^L \supset F_{\lambda_1} \cap \dots \cap F_{\lambda_n} \neq \emptyset,$$

provando que $\widetilde{\mathcal{F}}$ possui PIF. Desta forma, se a condição do Teorema 4.4.2 é verdadeira para $\widetilde{\mathcal{F}}$, então $\bigcap_{F \in \mathcal{F}} F = \bigcap_{\widetilde{F} \in \widetilde{\mathcal{F}}} \widetilde{F} \neq \emptyset$, e vale a condição do Teorema 4.4.2 para \mathcal{F} . Assim, podemos assumir que $\mathcal{F} = \{A_\lambda^L : \lambda \in \Lambda\}$.

Resta agora mostrar que se $\mathcal{F} = \{A_\lambda^L : \lambda \in \Lambda\}$ possui PIF, então $\bigcap_{\lambda \in \Lambda} A_\lambda^L \neq \emptyset$. Nossa ideia para mostrar isto será construir um ultrafiltro $p \in \beta\mathbb{N}$ que contém $A_\lambda, \forall \lambda \in \Lambda$. Isso implica que $p \in A_\lambda^L$ para todo $\lambda \in \Lambda$ e daí $p \in \bigcap_{\lambda \in \Lambda} A_\lambda^L$.

Começamos observando que para qualquer família *finita* Λ' , vale que

$$\bigcap_{\lambda' \in \Lambda'} A_{\lambda'}^L = \left(\bigcap_{\lambda'} A_{\lambda'} \right)^L.$$

Procedemos por indução na quantidade de elementos de Λ' . Para $|\Lambda'| = 1$, não há dada a ser provado. Suponha então que vale a igualdade para $|\Lambda'| = m \geq 1$, e tome $\Lambda' = \{1, \dots, m, m+1\}$. Pelo argumento da prova da proposição 4.4.1,

$$\begin{aligned} \bigcap_{\lambda' \in \{1, \dots, m, m+1\}} A_{\lambda'}^L &= \left(\bigcap_{\lambda' \in \{1, \dots, m\}} A_{\lambda'}^L \right) \cap A_{m+1}^L = \left(\bigcap_{\lambda' \in \{1, \dots, m\}} A_{\lambda'} \right)^L \cap A_{m+1}^L \\ &= \left(\left(\bigcap_{\lambda' \in \{1, \dots, m\}} A_{\lambda'} \right) \cap A_{m+1} \right)^L = \left(\bigcap_{\lambda' \in \{1, \dots, m+1\}} A_{\lambda'} \right)^L \end{aligned}$$

e isto encerra a prova da igualdade.

Definindo a família de conjuntos $\mathcal{F}_0 = \{A_\lambda : \lambda \in \Lambda\} \subset \mathcal{P}(\mathbb{N})$, a igualdade acima nos permite concluir que \mathcal{F}_0 possui PIF. Para construir p , defina

$$\mathfrak{G} = \{\mathcal{G} \subset \mathcal{P}(\mathbb{N}) : \mathcal{G} \supset \mathcal{F}_0 \text{ e } \mathcal{G} \text{ possui PIF}\}$$

com a ordem parcial de inclusão. Note que $\mathfrak{G} \neq \emptyset$ pois $\mathcal{F}_0 \in \mathfrak{G}$. Afirmamos que \mathfrak{G} é indutivo, isto é, qualquer subconjunto $\mathcal{A} \subset \mathfrak{G}$ totalmente ordenado possui uma cota superior em \mathfrak{G} . Seja então $\mathcal{A} \subset \mathfrak{G}$ totalmente ordenado, e defina $\mathcal{M} = \bigcup_{\mathcal{L} \in \mathcal{A}} \mathcal{L}$. Afirmamos que \mathcal{M} é a cota superior requerida. É claro que $\mathcal{M} \supset \mathcal{L}$ para qualquer $\mathcal{L} \in \mathcal{A}$, logo basta mostrar que $\mathcal{M} \in \mathfrak{G}$. Dado $\mathcal{L} \in \mathcal{A}$, temos $\mathcal{M} \supset \mathcal{L} \supset \mathcal{F}_0$. Ademais, \mathcal{M} possui PIF: dados $A_1, \dots, A_n \in \mathcal{M}$, existem $\mathcal{L}_1, \dots, \mathcal{L}_n \in \mathcal{A}$ tais que $A_i \in \mathcal{L}_i$ para $1 \leq i \leq n$; como \mathcal{A} é totalmente ordenado, podemos supor que $\mathcal{L}_1 \subset \mathcal{L}_2 \subset \dots \subset \mathcal{L}_n$ e portanto $A_1, \dots, A_n \in \mathcal{L}_n$, o que garante pela PIF gozada por \mathcal{L}_n que $A_1 \cap \dots \cap A_n \neq \emptyset$.

Aplicando o Lema de Zorn, \mathfrak{G} possui um elemento maximal, que denotamos por p . Afirmamos que p é um ultrafiltro.

- F.1. $\emptyset \notin p$, pois $p \supset \mathcal{F}_0 \neq \emptyset$.
- F.2. Sejam $A \in p$ e $B \supset A$. A família $p \cup \{B\}$ também está em \mathfrak{G} . Pela maximalidade de p , segue que $B \in p$.
- F.3. Sejam $A, B \in p$. A família $p \cup \{A \cap B\}$ também está em \mathfrak{G} . Novamente pela maximalidade de p , segue que $A \cap B \in p$.
- F.4. Seja $A \subset \mathbb{N}$. Por contradição, assumamos que $A, A^c \notin p$. Existem $A_1, A_2 \in p$ tais que $A_1 \cap A = \emptyset$, $A_2 \cap A^c = \emptyset$. Para ver que podemos fazer esta escolha, caso não fosse possível, então $B \cap A \neq \emptyset$ para todo $B \in p$ e daí $p \cup \{A\} \in \mathfrak{G}$; pela maximalidade de p , teríamos que $A \in p$, contradição. Desta forma, podemos escolher A_1, A_2 como acima. Observe que $A_1 \cap A = \emptyset \Rightarrow A_1 \subset A^c$. Analogamente, $A_2 \subset A$. Segue que $A_1 \cap A_2 = \emptyset$, o que contradiz F.1 e F.3.

Assim, p é um ultrafiltro, o que conclui a prova da proposição. \square

Proposição 4.4.3. $\beta\mathbb{N}$ é um semigrupo topológico à esquerda.

Demonstração. Queremos mostrar que, dado $p \in \beta\mathbb{N}$, a função $L_p : \beta\mathbb{N} \rightarrow \beta\mathbb{N}$ dada por $L_p(q) = p + q$ é contínua. Para tal, fixe $q_0 \in \beta\mathbb{N}$ e $A^L \supset L_p(q_0)$ aberto básico; devemos mostrar que existe um aberto básico $B^L \ni q_0$ tal que $L_p(B^L) \subset A^L$. Por definição,

$$p + q \in A^L \iff A \in p + q \iff B := \{n \in \mathbb{N} : A - n \in p\} \in q \iff q \in B^L.$$

Pronto: basta tomar B acima, observando que como por hipótese $A \in p + q_0$, então $B \in q_0$. \square

Desta forma, concluímos a prova de que $\beta\mathbb{N}$ é um semigrupo topológico à esquerda, compacto e Hausdorff. Pelo Teorema de Ellis, $\beta\mathbb{N}$ possui um elemento idempotente. Abaixo, fornecemos uma prova do Teorema de Ellis.

Demonstração do Teorema de Ellis. Seja $(G, *)$ semigrupo topológico à esquerda, compacto e Hausdorff. Dados $A, B \subset G$, escreva $A * B = \{a * b : a \in A, b \in B\}$. Construimos um subsemigrupo minimal de G . Para tal, considere a família

$$\mathcal{G} = \{A \subset G : A \text{ é compacto não-vazio e } A * A \subset A\}$$

ordenada pela inclusão. Afirmamos que \mathcal{G} é indutivo. Como $G \in \mathcal{G}$, temos $\mathcal{G} \neq \emptyset$. Agora, seja $\mathcal{F} \subset \mathcal{G}$ totalmente ordenado, e seja $A_0 = \bigcap_{A \in \mathcal{F}} A$. Note que A_0 é compacto não-vazio. De fato, por ser totalmente ordenado, \mathcal{F} possui PIF, e daí pelo Teorema 4.4.2 segue

que $A_0 \neq \emptyset$. Adicionalmente, como G é Hausdorff, A_0 é compacto. Por fim, note que se $x, y \in A_0$ então $x, y \in A$, $\forall A \in \mathcal{F} \implies x * y \in A$, $\forall A \in \mathcal{F} \implies x * y \in A_0$. Assim, $A_0 * A_0 \subset A_0$. Pelo Lema de Zorn, \mathcal{G} possui um elemento minimal $L \subset G$. Afirmamos que todo elemento de L é idempotente. Começamos checando que $x * L = L$ para todo $x \in L$. Para provar isso, utilizamos a minimalidade de L , notando três fatos. Primeiro, pela escolha de L , temos $x * L \subset L$. Segundo, como L é compacto, temos que $L_x(L) = x * L$ é compacto pela continuidade à esquerda. Terceiro,

$$(x * L) * (x * L) \subset (x * L) * L = x * (L * L) \subset x * L.$$

Pela minimalidade de L , segue que $x * L = L$.

Como L é compacto e cada $L_x : G \rightarrow G$ é contínua, a restrição $L_x|_L$ é contínua donde $L_x|_L$ é própria. Como G é Hausdorff, $\{x\}$ é compacto. Estas duas afirmações nos dão que, para $x \in L$, o conjunto

$$K_x = (L_x|_L)^{-1}(\{x\}) = \{y \in L : x * y = x\} \subset L$$

é um compacto não-vazio, visto que $x * L = L$. Ademais,

$$y, z \in K_x \implies x * (y * z) = (x * y) * z = x * z = x,$$

provando que $K_x * K_x \subset K_x$. Pela minimalidade de L , obtemos que $K_x = L$. Desta forma, $x \in L$ implica que $x \in K_x$ e portanto $x * x = x$, como afirmado. \square

Finalizamos este capítulo provando o Teorema de Hindman.

Teorema 4.4.3 (Hindman). *Se $\mathbb{N} = A_1 \cup \dots \cup A_r$ então algum A_i contém um conjunto IP.*

Demonstração. Tome $p \in \beta\mathbb{N}$ idempotente. Por F.4, $A_i \in p$ para exatamente um índice i . Mostraremos que $B_0 := A_i$ contém um conjunto IP. Para isto, construiremos por indução conjuntos $B_0 \supset B_1 \supset \dots$ em p e uma sequência crescente $k_0 < k_1 < \dots$ com $k_n \in B_n$ tais que $\sum_{i \in E} k_i \in B_0$ para todo $E \subset \{0, \dots, n\}$ e

$$B_{n+1} = \bigcap_{E \subset \{0, \dots, n\}} \left(B_0 - \sum_{j \in E} k_j \right) \in p$$

para todo $n \in \mathbb{N}$ (aqui, convencionamos $\sum_{i \in \emptyset} k_i = 0$).

O caso inicial é simples. Ora, como $B_0 \in p = p + p$, então $B_0^* := \{n \in \mathbb{N} : A - n \in p\} \in p$. Por F.3, $B_0 \cap B_0^* \in p$. Tome $k_0 \in B_0 \cap B_0^*$, o que diz que $B_0 - k_0 \in p$. Ponha $B_1 = B_0 \cap (B_0 - k_0) \in p$.

Pelo mesmo argumento acima, $B_1 \cap B_1^* \in p$ e daí existe $k_1 \in B_1 \cap B_1^*$ com $k_1 > k_0$ e $B_0 - k_1 \in p$. Então $k_1 \in B_1 \subset B_0 - k_0 \implies k_0 + k_1 \in B_0$.

Por indução, suponha construídos $B_0, \dots, B_n \in p$ e $k_i \in B_i$ com $k_0 < \dots < k_{n-2} < k_{n-1}$ tais que $\sum_{i \in E} k_i \in B_0$ para todo $E \subset \{0, \dots, n\}$ e

$$B_{i+1} = \bigcap_{E \subset \{0, \dots, i\}} \left(B_0 - \sum_{j \in E} k_j \right) \in p$$

para todo $i = 0, \dots, n-1$.

Seja então $k_n \in B_n \cap B_n^* \in p$. Como $B_n \cap B_n^*$ é infinito, podemos escolher $k_n > k_{n-1}$. Então $k_n \in B_n$. Logo,

$$\begin{aligned} k_n \in \bigcap_{E \subset \{0, \dots, n-1\}} \left(B_0 - \sum_{j \in E} k_j \right) &\implies k_n \in B_0 - \sum_{j \in E} k_j, \forall E \subset \{0, \dots, n-1\} \\ &\implies k_n + \sum_{j \in E} k_j \in B_0, \forall E \subset \{0, \dots, n-1\}. \end{aligned}$$

Por outro lado, $B_n - k_n \in p$. Assim como acima, concluímos que $B_0 - \sum_{j \in E} k_j \in p$ para todo $E \subset \{0, \dots, n\}$. Logo, por F.3,

$$B_{n+1} := \bigcap_{E \subset \{0, \dots, n\}} \left(B_0 - \sum_{j \in E} k_j \right) \in p$$

está bem-definido e isto conclui a prova do passo indutivo.

Assim, construímos, por indução, uma sequência crescente $(k_n)_{n=0}^\infty$ tal que $FS(k_n)_{n=0}^\infty \subset B_0 = A_i$, conforme queríamos provar. \square

5 A CONJECTURA $B + B + t$

Nos anos 70, dois teoremas importantes em combinatória aditiva foram provados, conforme vimos nos capítulos 3 e 4.

Teorema 5.0.1 (Hindman, 1974). *Seja $N = A_1 \cup \dots \cup A_k$ uma partição finita dos números naturais. Então algum A_i contém um conjunto IP.*

Teorema 5.0.2 (Szemerédi, 1975). *Se $d^*(A) > 0$ então A possui progressões aritméticas finitas de qualquer tamanho.*

Escreva $B = \{b_1 < b_2 < \dots\} \subset \mathbb{N}$ e seja

$$\oplus^k B = B \oplus \dots \oplus B = \left\{ \sum_{i \in F} b_i : F \subset \mathbb{N} \text{ com } |F| = k \right\}$$

o conjunto das somas de k elementos distintos de B . Observe que o conjunto IP $FS(b_k)_k$ pode ser reescrito como

$$FS(b_k)_k = \bigcup_{i=1}^{\infty} \oplus^k B.$$

Desta forma, podemos reescrever o Teorema de Hindman como abaixo.

Teorema 5.0.3 (Hindman renunciado). *Seja $N = A_1 \cup \dots \cup A_k$ uma partição finita dos números naturais. Então existem $B \subset \mathbb{N}$ infinito e $i \in \{1, \dots, k\}$ tais que A_i contém $\bigcup_{i=1}^{\infty} \oplus^k B$.*

Seguindo os Princípios de Ramsey vistos no Capítulo 3, é natural conjecturar o seguinte: se $d^*(A) > 0$, então A ou algum de seus transladados contém um conjunto IP. Frisamos a palavra *transladado*, pois sem permitirmos esta operação, então a resposta é claramente negativa, bastando tomar $2\mathbb{Z} + 1$ como contra-exemplo. Infelizmente (ou felizmente), a resposta é não, e o contra-exemplo é fornecido por Ernest G. Straus em (ERDŐS, 1980), segundo Erdős. Então, ao invés de perguntar para todas as ordens k , perguntamos: a conjectura é verdadeira para $\oplus^k B$ para algum k ? Permitindo translações, isto motivou a seguinte conjectura.

Conjectura (Erdős). *Seja $A \subset \mathbb{N}$ com densidade de Banach positiva. Então existem $B \subset \mathbb{N}$ infinito e $t \in \mathbb{N}$ tais que $B \oplus B + t \subset A$.*

É exatamente a prova desta conjectura, tal como no artigo (KRA et al., 2024b), que apresentaremos neste capítulo.

Teorema 5.0.4 (B. Kra, J. Moreira, K. Richter, D. Robertson). *Se $A \subset \mathbb{N}$ possui densidade de Banach superior positiva, então existem $B \subset \mathbb{N}$ infinito e $t \in \mathbb{N}$ tais que $B \oplus B + t \subset A$.*

5.1 Progressões de Erdős

O principal conceito relacionado à prova do Teorema 5.0.4 é o de *progressão de Erdős*. A existência de tais sequências garante que o conjunto dos pontos que visitam uma determinada região do espaço contém conjuntos da forma $B \oplus B$.

Definição 5.1.1 (Progressão de Erdős). *Dados X um espaço métrico compacto e $T : X \rightarrow X$ um homeomorfismo, um ponto $(x_0, x_1, x_2) \in X^3$ é dito ser uma progressão de Erdős de três termos quando existe uma sequência crescente $n_1 < n_2 < n_3 < \dots$ de inteiros tal que $(T \times T)^{n_i}(x_0, x_1) \rightarrow (x_1, x_2)$ quando $i \rightarrow \infty$.*

O teorema abaixo mostra como progressões de Erdős induzem uma estrutura em conjuntos de tempos de retorno.

Teorema 5.1.1. *Sejam X um espaço métrico compacto e $T : X \rightarrow X$ um homeomorfismo, e sejam $E, F \subset X$ abertos. Se existe uma progressão de Erdős $(x_0, x_1, x_2) \in X^3$ com $x_1 \in E$ e $x_2 \in F$, então existe um conjunto infinito $B \subset \{n \in \mathbb{N} : T^n x_0 \in E\}$ tal que $B \oplus B \subset \{n \in \mathbb{N} : T^n x_0 \in F\}$.*

Demonstração. Pela definição de sequência de Erdős, existe uma sequência crescente $c : \mathbb{N} \rightarrow \mathbb{N}$ para a qual

$$(T \times T)^{c(n)}(x_0, x_1) \rightarrow (x_1, x_2) \text{ quando } n \rightarrow \infty. \quad (5.1)$$

Note que $E \times F$ é vizinhança de (x_1, x_2) . Em particular, como $T^{c(n)}(x_0) \rightarrow x_1$, podemos supor, sem perda de generalidade, que $c(\mathbb{N}) \subset \{n \in \mathbb{N} : T^n x_0 \in E\}$.

Construímos $B \subset c(\mathbb{N})$ indutivamente. Por (5.1), $T^{c(n)}(x_1) \rightarrow x_2$. Como F é vizinhança de x_2 , existe $b(1) \in c(\mathbb{N})$ tal que $T^{b(1)}(x_1) \in F$. Logo, $x_1 \in T^{-b(1)}(F)$. Então $(E \cap T^{-b(1)}(F)) \times F$ é vizinhança de (x_1, x_2) . Mais uma vez por (5.1), existe $b(2) > b(1)$ tal que $(T \times T)^{b(2)}(x_0, x_1) \in T^{-b(1)}(F) \times F$. Então $T^{b(2)}x_0 \in T^{-b(1)}F \Rightarrow T^{b(1)+b(2)}x_0 \in F$, e $x_1 \in T^{-b(1)}(F) \cap T^{-b(2)}(F)$.

Suponha definidos $b(1) < b(2) < \dots < b(n) \subset c(\mathbb{N})$ com

$$x_0 \in \bigcap_{1 \leq i < j \leq n} T^{-b(i)-b(j)}(F) \text{ e } x_1 \in \bigcap_{1 \leq i \leq n} T^{-b(i)}(F).$$

Então $\left(\bigcap_{1 \leq i \leq n} T^{-b(i)}(F) \right) \times F$ é vizinhança de (x_1, x_2) . Por (5.1), existe $b(n+1) \in c(\mathbb{N})$ tal que

$$(T \times T)^{b(n+1)}(x_0, x_1) \in \left(\bigcap_{1 \leq i \leq n} T^{-b(i)}(F) \right) \times F.$$

Então

$$T^{b(n+1)}x_0 \in \bigcap_{1 \leq i \leq n} T^{-b(i)}F \implies x_0 \in \bigcap_{1 \leq i \leq n} T^{-b(n+1)-b(i)}F$$

o que, juntamente com o passo indutivo, nos dá que

$$x_0 \in \bigcap_{1 \leq i < j \leq n+1} T^{-b(i)-b(j)}F$$

e

$$T^{b(n+1)}x_1 \in F \implies x_1 \in T^{-b(n+1)}F \implies x_1 \in \bigcap_{1 \leq i \leq n+1} T^{-b(i)}F. \quad (5.2)$$

Isto prova o passo indutivo. É claro ver que $B = \{b(n) : n \in \mathbb{N}\}$ satisfaz as condições do teorema. \square

A existência de progressões de Erdős, no contexto da conjectura, é garantida pelo teorema abaixo.

Teorema 5.1.2. *Seja (X, β_X, μ, T) um sistema ergódico. Seja $a \in \text{gen}(\mu, \Phi)$ para alguma sequência de Følner Φ , e seja $E \subset X$ um conjunto aberto com $\mu(E) > 0$. Então existem $t \in \mathbb{N}$ e uma progressão de Erdős da forma $(a, x_1, x_2) \in X^3$ tal que $x_1 \in E$ e $T^t x_2 \in E$.*

Resta mostrar como este teorema pode ser aplicada para estabelecer a conjectura. Como vimos no Capítulo 3, isto é feito utilizando o Princípio de correspondência de Furstenberg, veja a Seção 3.2.

Prova do Teorema 5.0.4 assumindo o Teorema 5.1.2. Pelo princípio de correspondência de Furstenberg - versão ergódica (Teorema 3.2.2), existem um sistema ergódico (X, β_X, μ, T) , uma sequência de Følner Φ , um elemento $a \in \text{gen}(\mu, \Phi)$ e um conjunto $E \subset X$ aberto e fechado tais que $A = \{n \in \mathbb{N} : T^n a \in E\}$ e $\mu(E) = d^*(A) > 0$.

Pelo Teorema 5.1.2, existem $x_1, x_2 \in X$, $t \in \mathbb{N}$ e $n_1 < n_2 < \dots <$ tais que $x_1 \in E, x_2 \in T^{-t}E$ e (a, x_1, x_2) é progressão de Erdős. Ponha $F = T^{-t}E$ e note que $A - t = \{n \in \mathbb{N} : T^n a \in F\}$. Pelo Teorema 5.1.1, existe um conjunto infinito $B \subset A$ tal que $B \oplus B \subset \{n \in \mathbb{N} : T^n a \in F\} = A - t$, ou seja, $B \oplus B + t \subset A$. Isto prova a conjectura. \square

Dizemos que o fator de Kronecker de um sistema ergódico é contínuo quando a aplicação de fator associada é contínua. É de suma importância que o fator de Kronecker seja contínuo para a prova do Teorema 5.1.2. Em geral, o fator de Kronecker é apenas mensurável. Em algumas situações, podemos assumir que tal fator é contínuo, sem perder a generalidade dos argumentos. Para isto, precisaremos da proposição a seguir, que é chave para a prova do Teorema 5.1.2, conforme veremos nas próximas seções. Para a prova da Proposição 5.1.1, consulte (KRA et al., 2024a, Prop. 3.20).

Proposição 5.1.1. *Seja (X, β_X, T, μ) um sistema ergódico e $a \in \text{gen}(\mu, \Phi)$ para alguma sequência de Følner Φ . Então existe um sistema ergódico $(\tilde{X}, \beta_{\tilde{X}}, \tilde{T}, \tilde{\mu})$, uma sequência de Følner Ψ , um ponto $\tilde{a} \in \tilde{X}$ e um fator contínuo $\tilde{\pi} : \tilde{X} \rightarrow X$ tal que $\tilde{\pi}(\tilde{a}) = a$, $\tilde{a} \in \text{gen}(\tilde{\mu}, \Psi)$ e $(\tilde{X}, \beta_{\tilde{X}}, \tilde{\mu}, \tilde{T})$ tem fator de Kronecker contínuo.*

O seguinte teorema sustenta a importância da proposição acima.

Teorema 5.1.3. *Seja (X, β_X, μ, T) um sistema ergódico e assumamos que seu fator de Kronecker é contínuo. Seja $a \in \text{gen}(\mu, \Phi)$ para alguma sequência de Følner Φ e seja $E \subset X$ aberto tal que $\mu(E) > 0$. Então as seguintes afirmações são verdadeiras:*

- (i) *Existe $t \in \mathbb{N}$ e uma progressão de Erdős da forma $(a, x_1, x_2) \in X \times X \times X$ tal que $x_1 \in E$ e $T^t x_2 \in E$.*
- (ii) *Existe $t \in \mathbb{N}$ e uma progressão de Erdős da forma $(a, x_1, x_2) \in X \times X \times X$ tal que $T^t x_1 \in E$ e $T^t x_2 \in E$.*

Enunciamos o Teorema acima como na referência do artigo, (KRA et al., 2024b). Salientamos que dela usamos apenas o item (i). Discutiremos apenas este item.

Afirmamos que a Proposição 5.1.1 e o Teorema 5.1.3 implicam o Teorema 5.1.2. Para isto, sejam (X, β_X, μ, T) um sistema ergódico, $a \in \text{gen}(\mu, \Phi)$ para alguma sequência de Følner Φ e $E \subset X$ aberto com $\mu(E) > 0$. Tome $(\tilde{X}, \beta_{\tilde{X}}, \tilde{\mu}, \tilde{T})$, \tilde{a} e $\tilde{\pi}$ como na Proposição 5.1.1 e seja $\tilde{E} = \tilde{\pi}^{-1}(E) \subset \tilde{X}$. Como $(\tilde{X}, \beta_{\tilde{X}}, \tilde{\mu}, \tilde{T})$ tem fator de Kronecker contínuo, podemos aplicar o Teorema 5.1.3 e encontrar $t \in \mathbb{N}$ e uma progressão de Erdős

$(\tilde{a}, \tilde{x}_1, \tilde{x}_2) \in \tilde{X} \times \tilde{X} \times \tilde{X}$ com $\tilde{x}_1 \in \tilde{E}$ e $\tilde{T}^t(\tilde{x}_2) \in \tilde{E}$. Então $(\tilde{\pi}(\tilde{a}), \tilde{\pi}(\tilde{x}_1), \tilde{\pi}(\tilde{x}_2)) \in X \times X \times X$ é uma progressão de Erdős tal que $\tilde{\pi}(\tilde{x}_1) \in E$ e $T^t(\tilde{\pi}(\tilde{x}_2)) \in E$ (a prova deste fato é direta).

Portanto, daqui em diante, assumiremos que o fator de Kronecker é contínuo.

5.2 Medidas sobre progressões de erdős

Assumiremos, daqui em diante, as seguintes condições:

- O sistema mensurável (X, β_X, μ, T) é ergódico;
- (Z, β_Z, m, R) é o fator de Kronecker de (X, β_X, μ, T) (veja a Definição 2.5.2) e ”+” representa a operação do grupo Z ;
- Existe uma aplicação de fator $\pi : X \rightarrow Y$ contínua;
- Fixamos $z \mapsto \eta_z$ desintegração de μ sobre o fator $\pi : (X, \beta_X, \mu, T) \rightarrow (Z, \beta_Z, m, R)$.

Com as convenções acima, para $(x_1, x_2) \in X \times X$, definimos

$$\lambda_{(x_1, x_2)} = \int_Z \eta_{z+\pi(x_1)} \times \eta_{z+\pi(x_2)} dm(z)$$

medida de probabilidade em $X \times X$. Note que $\lambda_{(x_1, x_2)}$ não depende da desintegração de fator $z \mapsto \eta_z$, uma vez que a desintegração é única a menos de um conjunto de medida nula. Note também que a medida está bem definida, uma vez que $\eta_{z+\pi(x_i)}$ está definida para m -quase todo $z \in Z$, para $i = 1, 2$. O teorema a seguir diz que, em algum sentido, $(x_1, x_2) \mapsto \lambda_{(x_1, x_2)}$ é uma desintegração *ergódica e contínua* para $\mu \times \mu$.

Teorema 5.2.1. *Seja $\lambda_{(x_1, x_2)} \in \mathcal{M}(X \times X)$ como definida acima. Então:*

1. $(x_1, x_2) \in X \times X \mapsto \lambda_{(x_1, x_2)}$ é contínua;
2. $(x_1, x_2) \mapsto \lambda_{(x_1, x_2)}$ é uma desintegração de $\mu \times \mu$ no sentido de

$$\int_{X \times X} \lambda_{(x_1, x_2)} d(\mu \times \mu)(x_1, x_2) = \mu \times \mu;$$

3. Para $(\mu \times \mu)$ -quase todo $(x_1, x_2) \in X \times X$, o ponto (x_1, x_2) é genérico para $\lambda_{(x_1, x_2)}$ e $\lambda_{(x_1, x_2)}$ é ergódica para $T \times T$;
4. $\lambda_{(x_1, x_2)} = \lambda_{(T x_1, T x_2)}$ para $(\mu \times \mu)$ -quase todo $(x_1, x_2) \in X \times X$.

A prova deste teorema pode ser vista em (KRA et al., 2024a, Proposição 3.11). Lembremos que o nosso objetivo é o de encontrar uma progressão de Erdős nas condições do Teorema 5.1.3, e é aqui que as medidas $\lambda_{(x_1, x_2)}$ desempenham um papel crucial. Assumindo as hipóteses do Teorema 5.1.3, para encontrar uma progressão de Erdős (a, x_1, x_2) , basta exigir que:

1. O par (a, x_1) é genérico para a medida $\lambda_{(a, x_1)}$;
2. $(x_1, x_2) \in \text{supp}(\lambda_{(x_1, x_2)})$;
3. $\lambda_{(a, x_1)} = \lambda_{(x_1, x_2)}$.

Postergamos a prova disto para a última seção deste capítulo. A propriedade 3 é equivalente a (a, x_1, x_2) serem projetadas em uma progressão aritmética pelo fator de Kronecker. Ao fixar a , o conjunto dos (x_1, x_2) tais que isto acontece possui $(\mu \times \mu)$ -medida nula sempre que o fator de Kronecker não possui átomos, que é a situação usual.

Para superar esta dificuldade, *fixado um elemento $a \in X$* , introduzimos uma medida σ em $X \times X$ que dá medida total aos pontos (x_1, x_2) tais que (a, x_1, x_2) é progressão de Erdős, e mostraremos que as propriedades 1 e 2 continuam verdadeiras para σ -quase todo ponto.

Definição 5.2.1 (Medidas σ). *Fixado $a \in X$, definimos a medida σ_a por*

$$\sigma_a = \int_Z \eta_z \times \eta_{2z - \pi(a)} dm(z) = \int_Z \eta_{z + \pi(a)} \times \eta_{2z + \pi(a)} dm(z).$$

A seguir, provamos alguns lemas importantes sobre estas medidas. No que segue, $\pi_1 : X \times X \rightarrow X$ e $\pi_2 : X \times X \rightarrow X$ denotam as projeções na primeira e segunda coordenada respectivamente, isto é, $\pi_1(x_1, x_2) = x_1$ e $\pi_2(x_1, x_2) = x_2$.

Lema 5.2.1. $(\pi_1)_* \sigma = \mu$.

Demonstração. Dada $f \in L^1_\mu(X)$,

$$\begin{aligned} \int_X f d(\pi_1)_* \sigma &= \int_{X \times X} (f \otimes 1) d\sigma = \int_Z \left(\int_{X \times X} (f \otimes 1) d(\eta_z \times \eta_{2z - \pi(a)}) \right) dm(z) \\ &= \int_Z \left(\int_X f d\eta_z \right) dm(z) = \int_X f d\mu, \end{aligned}$$

o que conclui a prova. □

Lema 5.2.2. *Vale a igualdade*

$$\frac{1}{2}(\pi_2)_* \sigma + \frac{1}{2}(T \circ \pi_2)_* \sigma = \mu.$$

Demonstração. Seja $2Z = \{z + z : z \in Z\}$. É claro que $2Z$ é um subgrupo compacto de Z . A primeira afirmação é óbvia. Para ver que é compacto, notamos que $2Z = s(\Delta_Z)$ para a soma $s : Z \times Z \rightarrow Z$, $s(x, y) = x + y$, onde $\Delta_Z = \{(z, z) : z \in Z\}$. Como Δ_Z é compacto em $Z \times Z$ (Z é compacto e Hausdorff e o mesmo vale para $Z \times Z$, logo Δ_Z é compacto) e s é contínua, segue que $2Z$ é compacto.

A ergodicidade de R garante que $Z = 2Z \cup R(2Z)$. De fato, observe que $R^2(2Z) = 2Z$ e portanto $2Z \cup R(2Z)$ é um conjunto fechado R -invariante não-vazio. Como o sistema é ergódico, segue do Teorema 6.0.2 que ele é unicamente ergódico e minimal, e além disso $\text{supp}(m) = Z$, pois m é medida de Haar. Segue assim que $2Z \cup R(2Z) = Z$.

Seja ξ a medida de Haar de $2Z$. Denotando $i : 2Z \rightarrow Z$ a inclusão, seja $\hat{\xi} = i_*\xi$, que é uma medida de probabilidade em Z . Por definição, $\hat{\xi}(A) = \xi(2Z \cap A)$. Afirmamos que

$$m = \frac{1}{2}\hat{\xi} + \frac{1}{2}R_*\hat{\xi}.$$

Para ver isto, escreva $R = R_\alpha$ para algum elemento $\alpha \in Z$. Temos dois casos.

1. $\alpha \in 2Z$: necessariamente $2Z = Z$, então $\hat{\xi} = \xi = m$ e daí $\frac{1}{2}\hat{\xi} + \frac{1}{2}R_*\hat{\xi} = \frac{1}{2}m + \frac{1}{2}m = m$.
2. $\alpha \notin 2Z$: Como $R^2(2Z) = 2Z$, tem-se

$$R_*\left(\frac{1}{2}\hat{\xi} + \frac{1}{2}R_*\hat{\xi}\right) = \frac{1}{2}R_*\hat{\xi} + \frac{1}{2}(R^2)_*\hat{\xi} = \frac{1}{2}\hat{\xi} + \frac{1}{2}R_*\hat{\xi},$$

onde na segunda igualdade usamos que $R^2 = R_{2\alpha}$ é uma rotação em $2Z$ e ξ é medida de Haar em $2Z$. Então a medida é R -invariante. Pela unicidade ergódica de R , tem-se a igualdade afirmada.

Agora, provemos a igualdade enunciada no lema. Recorde que

$$\sigma = \sigma_a = \int_Z \eta_z \times \eta_{2z-\pi(a)} dm(z) = \int_Z \eta_{z+\pi(a)} \times \eta_{2z+\pi(a)} dm(z).$$

A projeção $\pi(a) \in Z = 2Z \cup R(2Z)$, então $\pi(a) = 2w$ ou $\pi(a) = 2w + \alpha$, para algum $w \in Z$. Vamos analisar os dois casos separadamente.

Caso 1: $\pi(a) = 2w$. Note que

$$(\pi_2)_*\sigma(A) = \sigma(X \times A) = \int_Z \eta_{z+\pi(a)}(X) \eta_{2z+\pi(a)}(A) dm(z) = \int_Z \eta_{2z+\pi(a)}(A) dm(z).$$

Realizando uma mudança de variável $l : z \in Z \mapsto (2z + 2w) \in 2Z$ e notando que $l_*m = \hat{\xi}$, a integral acima é igual a $\int_{2Z} \eta_z(A) d\hat{\xi}(z)$. Usando que $\hat{\xi}(2Z) = 1$, segue que

$$(\pi_2)_*\sigma = \int_{2Z} \eta_z d\hat{\xi}(z) = \int_Z \eta_z d\hat{\xi}(z).$$

Por outro lado, como a desintegração satisfaz $T_*\eta_z = \eta_{Rz}$, uma conta direta fazendo uso de mudança de variáveis nos dá

$$\begin{aligned} (T \circ \pi_2)_*\sigma(A) &= (\pi_2)_*\sigma(T^{-1}(A)) = \int_{2Z} \eta_z(T^{-1}(A)) d\hat{\xi}(z) = \int_{2Z} \eta_{Rz}(A) d\hat{\xi}(z) \\ &= \int_{2Z+\alpha} \eta_z dR_*\hat{\xi}(z) = \int_Z \eta_z dR_*\hat{\xi}(z), \end{aligned}$$

onde a última igualdade vale porque $R_*\hat{\xi}(2Z + \alpha) = 1$. Assim,

$$\frac{1}{2}(\pi_2)_*\sigma + \frac{1}{2}(T \circ \pi_2)_*\sigma = \int_Z \eta_z d\left(\frac{1}{2}\hat{\xi} + \frac{1}{2}R_*\hat{\xi}\right)(z) = \int_Z \eta_z dm(z) = \mu.$$

Caso 2: $\pi(a) = 2w + \alpha$. Seja $j : z \in Z \mapsto (2z + \alpha) \in 2Z + \alpha$, que satisfaz $j_*m = R_*\hat{\xi}$. Daí,

$$\begin{aligned} (\pi_2)_*\sigma &= \int_Z \eta_{2(z+w)+\alpha} dm(z) = \int_{2Z+\alpha} \eta_z dj_*m(z) \\ &= \int_{2Z+\alpha} \eta_z d(R_*\hat{\xi})(z) = \int_Z \eta_z d(R_*\hat{\xi})(z), \end{aligned}$$

onde usamos novamente que $R_*\hat{\xi}(2Z + \alpha) = 1$. Como $(R^2)\hat{\xi} = \hat{\xi}$, temos também que

$$\begin{aligned} (T \circ \pi_2)_*\sigma(A) &= (\pi_2)_*\sigma(T^{-1}(A)) = \int_{2Z+\alpha} \eta_z(T^{-1}(A)) d(R_*\hat{\xi})(z) \\ &= \int_{2Z+\alpha} \eta_{Rz}(A) d(R_*\hat{\xi})(z) = \int_{2Z+2\alpha} \eta_z d((R^2)_*\hat{\xi})(z) = \int_{2Z} \eta_z d\hat{\xi}(z) = \int_{2Z} \eta_z d\hat{\xi}(z), \end{aligned}$$

onde usamos que $2Z = 2Z + 2\alpha$ e mais uma vez que $\hat{\xi}(2Z) = 1$. Assim,

$$\frac{1}{2}(\pi_2)_*\sigma + \frac{1}{2}(T \circ \pi_2)_*\sigma = \int_Z \eta_z d\left(\frac{1}{2}\hat{\xi} + \frac{1}{2}R_*\hat{\xi}\right)(z) = \int_Z \eta_z dm(z) = \mu.$$

A prova está completa. □

5.3 Suporte da medida

Lema 5.3.1. *Seja $a \in X$. Para σ -quase todo $(x_1, x_2) \in X \times X$, as medidas $\lambda_{(a, x_1)}$ e $\lambda_{(x_1, x_2)}$ são iguais.*

Demonstração. Relembramos que

$$\lambda_{(x_1, x_2)} = \int_Z \eta_{z+\pi(x_1)} \times \eta_{z+\pi(x_2)} dm(z).$$

Seja

$$\begin{aligned} P &= \{(x_1, x_2) \in X \times X : \exists z \in Z \text{ tal que } \pi(x_1) = z + \pi(a) \text{ e } \pi(x_2) = 2z + \pi(a)\} \\ &= \bigsqcup_{z \in Z} \pi^{-1}(z + \pi(a)) \times \pi^{-1}(2z + \pi(a)). \end{aligned}$$

Para todo $z \in Z$, temos

$$(\eta_{z+\pi(a)} \times \eta_{2z+\pi(a)})(P) \geq (\eta_{z+\pi(a)} \times \eta_{2z+\pi(a)})[\pi^{-1}(z + \pi(a)) \times \pi^{-1}(2z + \pi(a))] = 1$$

e portanto

$$\sigma(P) = \int_Z (\eta_{z+\pi(a)} \times \eta_{2z+\pi(a)})(P) dm(z) = 1.$$

Logo, o resultado estará provado se mostrarmos que $\lambda_{(a,x_1)} = \lambda_{(x_1,x_2)}$ para todo $(x_1, x_2) \in P$. Fixe então um tal elemento. Por definição, temos $\pi(x_1) - \pi(a) = \pi(x_2) - \pi(x_1) =: z_0$. Aplicando a mudança de variáveis $z \mapsto z - z_0$ à expressão de $\lambda_{(x_1,x_2)}$, temos

$$\begin{aligned} \lambda_{(x_1,x_2)} &= \int_Z \eta_{z+\pi(x_1)} \times \eta_{z+\pi(x_2)} dm(z) = \int_Z \eta_{z-z_0+\pi(x_1)} \times \eta_{z-z_0+\pi(x_2)} dm(z) \\ &= \int_Z \eta_{z+\pi(a)} \times \eta_{z+\pi(x_1)} dm(z) = \lambda_{(a,x_1)} \end{aligned}$$

e a prova está completa. \square

Seja $S = \{(x_1, x_2) \in X \times X : (x_1, x_2) \in \text{supp}(\lambda_{(x_1,x_2)})\}$. Como $(x_1, x_2) \mapsto \lambda_{(x_1,x_2)}$ é contínua, pelo Lema 6.0.2 temos que S é Borel mensurável. Nosso objetivo é mostrar que $\sigma(S) = 1$. Precisaremos do Teorema de Lusin, enunciado abaixo. Para a prova, consulte (FOLLAND, 1999, p. 217).

Teorema 5.3.1 (Lusin). *Seja (X, Σ, μ) um espaço de medida de Radon e Y um espaço topológico com base enumerável e munido da σ -álgebra de Borel. Seja $f : X \rightarrow Y$ uma função mensurável. Dado $\varepsilon > 0$, para todo $A \in \Sigma$ com $\mu(A) < +\infty$ existe um fechado E de X em A tal que $\mu(A \setminus E) < \varepsilon$ e $f|_E$ é contínua.*

Proposição 5.3.1. *Seja (X, β_X, μ, T) ergódico com uma transformação de fator contínua π ao seu fator de Kronecker (Z, β_Z, m, R) , e fixe uma desintegração $z \mapsto \eta_z$ de μ com respeito a π . Então existe uma sequência $\delta(j) \rightarrow 0$ tal que para μ -quase todo $x \in X$, a seguinte propriedade é verificada: para toda vizinhança U de x vale que*

$$\lim_{j \rightarrow +\infty} \frac{m(\{z \in Z : \eta_z(U) > 0\} \cap B(\pi(x), \delta(j)))}{m(B(\pi(x), \delta(j)))} = 1.$$

Demonstração. O argumento é baseado no Teorema de Lusin, e em um argumento de \liminf . Considere a transformação $\Phi : Z \rightarrow \mathcal{F}(X)$ dado por $\Phi(z) = \text{supp}(\eta_z)$. Aqui, $\mathcal{F}(X)$ representa a família dos conjuntos fechados e não-vazios de X , munido com a topologia gerada pela métrica de Hausdorff. Comentamos mais sobre isso no Apêndice 6. Sabemos que Φ é Borel mensurável por ser composição de $z \mapsto \eta_z$ e $\nu \mapsto \text{supp}(\nu)$. Como (Z, m, R) é espaço de Radon e $\mathcal{F}(X)$ é Hausdorff com base enumerável, aplicamos o Teorema de Lusin e obtemos, para cada $j \in \mathbb{N}$, um conjunto fechado $Z_j \subset Z$ com $m(Z_j) > 1 - \frac{1}{2^j}$ tal que $\Phi|_{Z_j}$ é contínua. Cada Z_j é de fato compacto (pois Z é compacto) e daí $\Phi|_{Z_j}$ é uniformemente contínua. Logo, existe $\delta(j) < \frac{1}{j}$ tal que para todos $z_1, z_2 \in Z_j$,

$$d(z_1, z_2) < \delta(j) \implies H(\Phi(z_1), \Phi(z_2)) < \frac{1}{j}.$$

Considere agora, para cada $j \in \mathbb{N}$, o conjunto

$$K_j = \left\{ z \in Z_j : m\left(B(z, \delta(j)) \cap Z_j\right) > \left(1 - \frac{1}{j}\right) m\left(B(z, \delta(j))\right) \right\},$$

e defina $\chi_j : Z \rightarrow \mathbb{R}$ via a relação

$$\chi_j(z) = \frac{1}{m(B(0, \delta(j)))} \int_{Z_j} \mathbb{1}_{B(0, \delta(j))}(w - z) dm(w).$$

Como

$$1 = \mathbb{1}_{B(0, \delta(j))}(w - z) \iff w - z \in B(0, \delta(j)) \iff w \in B(z, \delta(j)) \iff 1 = \mathbb{1}_{B(z, \delta(j))}(w),$$

tem-se $\mathbb{1}_{B(0, \delta(j))}(w - z) = \mathbb{1}_{B(z, \delta(j))}(w)$ e daí

$$\chi_j(z) = \frac{1}{m(B(z, \delta(j)))} \int_{Z_j} \mathbb{1}_{B(z, \delta(j))}(w) dm(w) = \frac{m(B(z, \delta(j)) \cap Z_j)}{m(B(z, \delta(j)))}$$

e podemos reescrever K_j como

$$K_j = Z_j \cap \left\{ z \in Z : \chi_j(z) > 1 - \frac{1}{j} \right\}.$$

Note que χ_j é integrável pois $0 \leq \chi_j \leq 1$. Pelo Teorema de Fubini,

$$\begin{aligned} \int_Z \chi_j dm(z) &= \frac{1}{m(B(z, \delta(j)))} \int_Z \left(\int_{Z_j} \mathbb{1}_{B(0, \delta(j))}(w - z) dm(w) \right) dm(z) \\ &= \frac{1}{m(B(z, \delta(j)))} \int_{Z_j} \left(\int_Z \mathbb{1}_{B(0, \delta(j))}(w - z) dm(z) \right) dm(w) \\ &= \frac{1}{m(B(z, \delta(j)))} \int_{Z_j} \left(\int_Z \mathbb{1}_{B(0, \delta(j))}(-z) dm(z) \right) dm(w) \\ &= \frac{1}{m(B(z, \delta(j)))} \int_{Z_j} \left(\int_Z \mathbb{1}_{B(0, \delta(j))}(z) dm(z) \right) dm(w) \\ &= \frac{1}{m(B(z, \delta(j)))} \int_{Z_j} m(B(0, \delta(j))) dm(z) \\ &= \frac{m(Z_j) \cdot m(B(0, \delta(j)))}{m(B(0, \delta(j)))} = m(Z_j) > 1 - \frac{1}{2^j} \end{aligned}$$

e portanto $\int (1 - \chi_j) dm < \frac{1}{2^j}$. Como $0 \leq 1 - \chi_j \leq 1$, a desigualdade de Markov nos dá

$$\begin{aligned} m\left(\left\{z \in Z : \chi_j > 1 - \frac{1}{j}\right\}\right) &= 1 - m\left(\left\{z \in Z : \chi_j \leq 1 - \frac{1}{j}\right\}\right) \\ &= 1 - m\left(\left\{z \in Z : 1 - \chi_j \geq \frac{1}{j}\right\}\right) \geq 1 - \frac{\int (1 - \chi_j) dm}{1/j} \geq 1 - \frac{j}{2^j}. \end{aligned}$$

Temos $Z \setminus K_j = (Z \setminus Z_j) \cup Z \setminus \left\{z \in Z : \chi_j(z) > \left(1 - \frac{1}{j}\right)\right\}$, donde

$$m(Z \setminus K_j) \leq m(Z \setminus Z_j) + m\left(Z \setminus \left\{z \in Z : \chi_j > \left(1 - \frac{1}{j}\right)\right\}\right) \leq \frac{1}{2^j} + \frac{j}{2^j} = \frac{j+1}{2^j}$$

e portanto

$$\sum_{j \in \mathbb{N}} m(Z \setminus K_j) \leq \sum_{j \in \mathbb{N}} \frac{j+1}{2^j} < +\infty.$$

Pelo Lema de Borel-Cantelli, o conjunto

$$K = \liminf_{j \geq 1} K_j = \bigcup_{M \geq 1} \bigcap_{j \geq M} K_j$$

satisfaz $m(K) = 1$, logo $\mu(\pi^{-1}(K)) = 1$. Relembrando do Lema 6.0.3 que $\mu(\{x \in X : x \in \text{supp}(\eta_{\pi(x)})\}) = 1$ (aqui é onde utilizamos a continuidade do fator π), o conjunto

$$L = \{x \in X : x \in \text{supp}(\eta_{\pi(x)})\} \cap \pi^{-1}(K)$$

satisfaz $\mu(L) = 1$. Afirmamos que L satisfaz as condições do enunciado, isto é, para todo $x \in L$ e todo aberto $U \ni x$ vale que

$$\lim_{j \rightarrow +\infty} \frac{m(\{z \in Z : \eta_z(U) > 0\} \cap B(\pi(x), \delta(j)))}{m(B(\pi(x), \delta(j)))} = 1$$

Para isto, fixe $x \in L$ e $U \ni x$ aberto, e ponha $z = \pi(x)$. Como $z \in K$, existe j_0 tal que $z \in K_j$ para todo $j \geq j_0$. Como U é aberto, podemos tomar j_0 grande para que $B(x, 1/j_0) \subset U$. Seja $H = \{z \in Z : \eta_z(U) > 0\}$.

Afirmação. Para todo $j \geq j_0$, temos $B(z, \delta(j)) \cap Z_j \subset H$.

De fato, tome $z' \in B(z, \delta(j)) \cap Z_j$. Queremos mostrar que $\eta_{z'}(U) > 0$. Ora, temos que $d(z, z') < \delta(j)$ e $z' \in Z_j$, donde pela continuidade uniforme em Z_j , temos $H(\Phi(z), \Phi(z')) < 1/j$. Pela definição de H ,

$$\frac{1}{j} > H(\Phi(z), \Phi(z')) = \max \left\{ \sup_{x \in \Phi(z)} d(x, \Phi(z')), \sup_{y \in \Phi(z')} d(y, \Phi(z)) \right\} \implies \frac{1}{j} > \sup_{x \in \Phi(z)} d(x, \Phi(z'))$$

Então existe $x' \in \text{supp}(\eta_{z'})$ tal que $d(x, x') < 1/j$ e daí $x' \in B(x, 1/j) \subset U$. Como $x' \in \text{supp}(\eta_{z'})$, segue que $\eta_{z'}(U) > 0$, conforme queríamos.

Voltando para a prova da proposição, como $z \in K_j$ para todo $j \geq j_0$, segue da afirmação que $\{z \in Z : \eta_z(U) > 0\} \cap B(z, \delta(j)) \supset B(z, \delta(j)) \cap Z_j$ donde

$$\frac{m(\{z \in Z : \eta_z(U) > 0\} \cap B(\pi(x), \delta(j)))}{m(B(\pi(x), \delta(j)))} \geq \frac{m(B(z, \delta(j)) \cap Z_j)}{m(B(z, \delta(j)))} > 1 - \frac{1}{j}.$$

Tomando o limite, concluimos que

$$\lim_{j \rightarrow +\infty} \frac{m(\{z \in Z : \eta_z(U) > 0\} \cap B(\pi(x), \delta(j)))}{m(B(\pi(x), \delta(j)))} = 1$$

conforme queríamos provar. \square

Proposição 5.3.2. $\sigma(S) = 1$.

Demonstração. Seja $\delta(j) \rightarrow 0$ dada pela proposição anterior, e L definido na prova de tal proposição. Como $(\pi_1)_*\sigma = \mu$ (Lema 5.2.1), temos $\sigma(L \times X) = \mu(L) = 1$. Daí, segue do Lema 5.2.2 que

$$1 = \mu(L) = \frac{1}{2}\sigma(X \times L) + \frac{1}{2}(\pi_2)_*\sigma(T^{-1}(L)) = \frac{1}{2}\sigma(X \times L) + \frac{1}{2}\sigma(X \times T^{-1}(L))$$

e portanto $\sigma(X \times L) = \sigma(X \times T^{-1}L) = 1$. Logo, $\sigma(L \times L) = \sigma((X \times L) \cap (L \times X)) = 1$. Assim, o lema estará provado se mostrarmos a inclusão $L \times L \subset S$. Fixe $x_1, x_2 \in L$. Queremos mostrar que $\lambda_{(x_1, x_2)} \in \text{supp}(\lambda_{(x_1, x_2)})$. Para isto, basta mostrar que para todas U_1 e U_2 vizinhanças de x_1 e x_2 vale que $\lambda_{(x_1, x_2)}(U_1 \times U_2) > 0$. Seja $\beta = \pi(x_2) - \pi(x_1)$. Da invariância de m por translações, aplicando a mudança de variáveis $z \mapsto z - \pi(x_1)$, obtemos

$$\lambda_{(x_1, x_2)} = \int_Z \eta_{z+\pi(x_1)} \times \eta_{z+\pi(x_2)} dm(z) = \int_Z \eta_z \times \eta_{z+\beta} dm(z)$$

Como $x_1, x_2 \in L$, a Proposição 5.3.1 implica que

$$\frac{m(\{z \in Z : \eta_z(U_i) > 0\} \cap B(\pi(x_i), \delta(j)))}{m(B(\pi(x_i), \delta(j)))} \geq \frac{3}{4}, \quad i = 1, 2$$

para j suficientemente grande. Sejam $A = \{z \in Z : \eta_z(U_1) > 0\}$, $C = \{z \in Z : \eta_{z+\beta}(U_2) > 0\}$ e $B = B(\pi(x_1), \delta)$. Note que $A \cap B - z = (A - z) \cap (B - z)$ e $\{z \in Z ; \eta_z(U_2) > 0\} - \beta = \{z \in Z ; \eta_{z+\beta}(U_2) > 0\}$.

Pela invariância de m por translações, temos

$$\frac{m(\{z \in Z : \eta_{z+\beta}(U_2) > 0\} \cap B(\pi(x_1), \delta))}{m(B(\pi(x_1), \delta))} = \frac{m(\{z \in Z : \eta_z(U_2) > 0\} \cap B(\pi(x_2), \delta))}{m(B(\pi(x_2), \delta))} \geq \frac{3}{4}.$$

Defina $W = \{z \in Z : \eta_z(U_1) > 0 \text{ e } \eta_{z+\beta}(U_2) > 0\} = A \cap C$. Usando a (sub)aditividade da medida, $m((A \cap C) \cap B) \geq m(A \cap B) + m(C \cap B) - m(B) \Rightarrow \frac{m((A \cap C) \cap B)}{m(B)} \geq \frac{m(A \cap B)}{m(B)} + \frac{m(C \cap B)}{m(B)} - 1$ e daí

$$\frac{m(W \cap B(\pi(x_1), \delta))}{m(B(\pi_1), \delta)} = \frac{m((A \cap C) \cap B)}{m(B)} \geq \frac{m(A \cap B)}{m(B)} + \frac{m(C \cap B)}{m(B)} - 1 \geq \frac{3}{4} + \frac{3}{4} - 1 = \frac{1}{2}.$$

Assim, $m(W) \geq m(W \cap B(\pi(x_1), \delta)) \geq \frac{m(B(\pi(x_1), \delta))}{2} > 0$. Para $z \in W$, $\eta_z \times \eta_{z+\beta}(U_1 \times U_2) = \eta_z(U_1)\eta_{z+\beta}(U_2) > 0$, donde

$$\lambda_{(x_1, x_2)}(U_1 \times U_2) = \int \eta_z \times \eta_{z+\beta}(U_1 \times U_2) dm(z) \geq \int_W \eta_z \times \eta_{z+\beta}(U_1 \times U_2) dm(z) > 0.$$

Isto conclui a prova. \square

5.4 Conclusão da prova do teorema

Lema 5.4.1. *Sejam (X, β_X, μ, T) um sistema ergódico e $a \in \text{gen}(\mu, \Phi)$ para alguma sequência de Følner Φ . Então existe uma sequência de Følner Ψ tal que para μ -quase todo $x_1 \in X$ o ponto $(a, x_1) \in \text{gen}(\lambda_{(a, x_1)}, \Psi)$.*

A prova deste lema pode ser encontrada em (KRA et al., 2024a, Lema 3.18). Vamos agora à prova do Teorema 5.1.3. Fixe um sistema ergódico (X, β_X, μ, T) , $a \in \text{gen}(\mu, \Phi)$ para alguma sequência de Følner Φ , e assuma que o fator de Kronecker (Z, β_Z, m, R) possui uma transformação de fator contínua $\pi : X \rightarrow Z$.

Do Lema 5.4.1, existe uma sequência de Følner Ψ tal que para μ -quase todo ponto $x_1 \in X$ vale que $(a, x_1) \in \text{gen}(\lambda_{(a, x_1)}, \Psi)$. Chame de C o conjunto de tais pontos. Como $(\pi_1)_* \sigma = \mu$, então $\sigma(C \times X) = \mu(C) = 1$ e, para todo $(x_1, x_2) \in C \times X$ temos que $(a, x_1) \in \text{gen}(\lambda_{(a, x_1)}, \Psi)$. Aplicando o Lema 5.3.1 e a Proposição 5.3.2, segue que para σ -quase todo ponto $(x_1, x_2) \in X \times X$ temos as seguintes propriedades:

1. $(a, x_1) \in \text{gen}(\lambda_{(a, x_1)}, \Psi)$.
2. $(x_1, x_2) \in \text{supp}(\lambda_{(x_1, x_2)})$.
3. $\lambda_{(a, x_1)} = \lambda_{(x_1, x_2)}$.

Isto implica que σ -quase todo $(x_1, x_2) \in X \times X$ é tal que (a, x_1, x_2) é progressão de Erdős, conforme a proposição abaixo.

Proposição 5.4.1. *Seja $D = \{(x_1, x_2) \in X \times X : (a, x_1, x_2) \text{ é uma progressão de Erdős}\}$. Então $\sigma(D) = 1$.*

Demonstração. Ora, σ -quase todo $(x_1, x_2) \in X \times X$ satisfaz as propriedades 1, 2, 3 descritas acima. Escreva $\Psi = (\Psi_n)_n$. Por 1 e 3:

$$\lambda_{(a, x_1)} = \lambda_{(x_1, x_2)} = \lim_{n \rightarrow \infty} \frac{1}{|\Psi_n|} \sum_{m \in \Psi_n} \delta_{(T \times T)^m}(a, x_1).$$

Como $(x_1, x_2) \in \text{supp}(\lambda_{(x_1, x_2)})$, dado $B_k = B((x_1, x_2), \frac{1}{k})$, segue do Teorema de Portmanteau que

$$\liminf_{n \rightarrow \infty} \frac{1}{|\Psi_n|} \sum_{m \in \Psi_n} \delta_{(T \times T)^m(a, x_1)} \geq \lambda_{(x_1, x_2)}(B_k) > 0.$$

Como $|\Psi_n| \xrightarrow{n \rightarrow \infty} \infty$, temos que $(T \times T)^n(a, x_1) \in B_k$ para infinitos valores de n . Um argumento indutivo mostra que existe $(n_k)_k \subset \mathbb{N}$ tal que $(T \times T)^{n_k}(a, x_1) \in B_k$ e daí $(T \times T)^{n_k}(a, x_1) \rightarrow (x_1, x_2)$, provando que (a, x_1, x_2) é progressão de Erdős. \square

Agora, tome $E \subset X$ aberto com $\mu(E) > 0$. Provaremos (i) do Teorema 5.1.3, que equivale a encontrar uma progressão de Erdős (a, x_1, x_2) tal que $(x_1, x_2) \in E \times T^{-t}E$ para algum $t \in \mathbb{N}$. Para isto, é suficiente mostrar que $\sigma(E \times T^{-t}(E)) > 0$ para algum $t > 0$. De fato, pela proposição acima, neste caso temos que $\sigma(D \cap (E \times T^{-t}(E))) > 0$ e daí qualquer elemento $(x_1, x_2) \in D \cap (E \times T^{-t}(E))$ define (a, x_1, x_2) satisfazendo o requerido. Portanto, resta mostrar que

$$\sigma\left(E \times \bigcup_{t \in \mathbb{N}} T^{-t}E\right) > 0$$

Ora, como μ é T -invariante e ergódica e $\mu(E) > 0$, $\bigcup_{t \in \mathbb{N}} T^{-t}E$ tem medida total. Escrevendo

$$E \times \left(\bigcup_{t \geq 1} T^{-t}E\right) = (E \times X) \cap \left(X \times \left(\bigcup_{t \geq 1} T^{-t}E\right)\right)$$

temos pelo Lema 5.2.2 que

$$1 = \mu\left(\bigcup_{t \geq 1} T^{-t}E\right) = \frac{1}{2}\sigma\left(X \times \bigcup_{t \geq 1} T^{-t}E\right) + \frac{1}{2}\sigma\left(X \times \bigcup_{t \geq 2} T^{-t}E\right)$$

Como σ é medida de probabilidade, a única forma de valer a igualdade acima é que ambos termos da soma sejam iguais a $\frac{1}{2}$. Em particular, $\sigma\left(X \times \bigcup_{t \geq 1} T^{-t}E\right) = 1$. Além do mais, pelo Lema 5.2.1, $\sigma(E \times X) = \sigma(\pi_1^{-1}(E)) = \mu(E)$, e daí

$$\sigma\left(E \times \bigcup_{t \geq 1} T^{-t}E\right) = \sigma\left((E \times X) \cap \left(X \times \bigcup_{t \geq 1} T^{-t}E\right)\right) = \sigma(E \times X) = \mu(E) > 0$$

o que implica na existência de $t \in \mathbb{N}$ tal que $\sigma(E \times T^{-t}E) > 0$, como requerido.

5.5 Comentários e consequências da prova da conjectura

Note que a conjectura é enunciada sob a condição de uma soma *restrita*: dado um subconjunto $A \subset \mathbb{N}$ de densidade de Banach positiva, existe um subconjunto infinito

$B \subset \mathbb{N}$ e $t \in \mathbb{N}$ tais que $B \oplus B + t \subset A$, e não $B + B + t \subset A$. Esta restrição é necessária, conforme o exemplo a seguir.

Exemplo (Steven Leth). Para $c \in (1, 2)$, defina $A = \mathbb{N} \cap \bigcup_{n \in \mathbb{N}} [4^n, c \cdot 4^n)$. Observamos inicialmente que $\bar{d}(A) > 0$, onde calculamos a densidade de Banach com relação à sequência de Følner clássica $\Phi_n = \{1, \dots, n\}$. Para ver isto, note que $|\mathbb{N} \cap [4^n, c \cdot 4^n)| = \lfloor c \cdot 4^n \rfloor - 4^n \geq (c-1) \cdot 4^n - 1$. Deste modo,

$$\begin{aligned} \frac{|A \cap \{1, 2, \dots, \lfloor c \cdot 4^n \rfloor\}|}{\lfloor c \cdot 4^n \rfloor} &= \sum_{k=1}^n \frac{\lfloor c \cdot 4^k \rfloor - 4^k}{\lfloor c \cdot 4^n \rfloor} \\ &\geq \sum_{k=1}^n \frac{(c-1)4^k - 1}{c \cdot 4^n} \\ &\geq \frac{c-1}{3c} \cdot \frac{4^{n+1} - 1}{4^n} - \frac{n}{c \cdot 4^n} \end{aligned}$$

Tomando o limite em n , concluímos que

$$\limsup_{n \rightarrow \infty} \frac{|A \cap \{1, 2, \dots, \lfloor c \cdot 4^n \rfloor\}|}{\lfloor c \cdot 4^n \rfloor} \geq \lim_{n \rightarrow \infty} \left[\frac{c-1}{c} \cdot \frac{4^{n+1} - 1}{4^n} - \frac{n}{c \cdot 4^n} \right] = \frac{4(c-1)}{3c}$$

donde $\bar{d}(A) > 0$.

É claro que A não contém nenhum conjunto da forma $\{b, 2b\}$, para $b \in \mathbb{N}$. Como qualquer translação de A permanece com densidade de Banach superior positiva, concluímos que nenhum transladado de A pode conter um conjunto da forma $B + B$, $B \subset \mathbb{N}$, uma vez que $\{b, 2b\} \subset B + B$ para todo $b \in B$. O exemplo acima nos informa que densidade positiva não é suficiente para garantir que subconjuntos de \mathbb{N} possuam subconjuntos da forma $B + B + t$, $B \subset \mathbb{N}$ infinito.

Por fim, destacamos algumas consequências da prova da Conjectura $B + B + t$. Em 2024, Ioannis Kousek e Tristán Radić provaram que quando $\bar{d}(A) > \frac{3}{4}$, existem $B \subset \mathbb{N}$ e $t \in \{0, 1\}$ tais que $B + B + t \subset A$. O artigo está disponível em (KOUSSEK; RADIC, 2024). Também em 2024, Dimitrios Charamaras e Andreas Mountakis conseguiram provar uma generalização da conjectura para várias classes de grupos “amenable” (vide (CHARAMARAS; MOUNTAKIS, 2024)).

6 CONCLUSÃO

Ao longo deste trabalho, vimos aplicações da Teoria Ergódica na prova de resultados em Teoria de Ramsey, finalizando com o estudo da prova da Conjectura $B + B + t$. Isto mostra a importância da interdisciplinaridade destas áreas, e demonstra que as técnicas aqui apresentadas são modernas e profundas. Isto fica claro na última seção do capítulo anterior, onde citamos algumas consequências importantes da prova da conjectura $B + B + t$. Endereçamos o leitor a (KRA et al., 2024a, p. 41) para uma lista de problemas em aberto em Combinatória, Dinâmica e Ultrafiltros, conceitos aqui estudados.

REFERÊNCIAS

- ALIPRANTIS, C.; BORDER, K. *Infinite Dimensional Analysis:: A Hitchhiker's Guide*. 3. ed. New York: Springer-Verlag, 2006.
- CHARAMARAS, D.; MOUNTAKIS, A. *Finding Product Sets in Some Classes of Amenable Groups*. 2024. Disponível em: <https://arxiv.org/abs/2402.07779>. Acesso em: 10 de jan. de 2025.
- EINSIEDLER, M.; WARD, T. *Ergodic Theory with a view towards Number Theory*. London: Springer-Verlag, 2011. (Graduate Texts in Mathematics, vol. 259).
- ELLIS, R. Distal transformation groups. *Pacific J. Math.*, v. 8, p. 401–405, 1958.
- ERDŐS, P. A survey of problems in combinatorial number theory. *Annals of Discrete Mathematics*, v. 6, p. 89–115, 1980.
- ERDŐS, P. T. P. On some sequence of integers. *Journal of the London Mathematical Society.*, v. 28, p. 261–264, 1936.
- FOLLAND, G. B. *Real Analysis: Modern Techniques and Their Applications*. 2. ed. New York: John Wiley and Sons, Inc, 1999.
- FURSTENBERG, H. Ergodic behavior of diagonal measures and a theorem of Szemerédi's Theorem. *J. Analyse Math.*, v. 7, n. 3, p. 527–552, 1977.
- HALMOS, P.; NEUMANN, J. von. Operator methods in classical mechanics, ii. *Annals of Mathematics*, v. 43, n. 2, p. 332–350, 1942.
- HINDMAN, N. Finite sums from sequences within cells of a partition of \mathbb{N} . *Journal of Combinatorial Theory*, v. 17, p. 1–11, 1974.
- KOUSEK, I.; RADIĆ, T. *Infinite unrestricted sumsets of the form $B + B$ in sets with large density*. 2024. Disponível em: <https://arxiv.org/pdf/2404.12201>. Acesso em: 10 de jan. de 2025.
- KRA, B. et al. Infinite sumsets in sets with positive density. *J. Amer. Math. Soc.*, v. 37, n. 3, p. 637–682, 2024.
- KRA, B. et al. A proof of Erdős $B + B + t$ conjecture. *Commun. Amer. Math. Soc.*, v. 4, n. 17, p. 480–494, 2024.
- LIMA, Y. *Teoria de Ramsey Ergódica*. Rio de Janeiro, 2009. Notas de aula.
- MUNKRES, J. *Topology*. 2. ed. Upper Saddle River: Pearson Education, 2014.
- OLIVEIRA, K.; VIANA, M. *Fundamentos da Teoria Ergódica*. 2. ed. Rio de Janeiro: IMPA, 2019.
- ROTH, K. On some sequence of integers. *Journal of the London Mathematical Society.*, v. 28, n. 1, p. 104–109, 1953.
- WAERDEN, B. L. van der. Beweis einer baudetschen vermutung. *Nieuw. Arch. Wisk.*, v. 15, p. 212–216, 1927.

APÊNDICE A - TEORIA DA MEDIDA PARA GRUPOS

Neste apêndice, discutimos aspectos importantes de medidas de Borel em grupos topológicos localmente compactos, destacando o caso especial de quando o grupo é métrico e compacto. Definimos uma medida especial, chamada de medida de Haar, a qual possui propriedades ergódicas relevantes. A referência para esta discussão pode ser encontrada em (OLIVEIRA; VIANA, 2019), capítulo 6.

MEDIDAS DE HAAR

Definição 6.0.1. *Um grupo topológico é um grupo (G, \cdot) munido de uma topologia de sorte que o produto $(g, h) \in G \times G \rightarrow g \cdot h \in G$ e a inversão $g \in G \rightarrow g^{-1} \in G$ aplicações contínuas.*

Exemplos.

1. Seja $\mathbb{S}^1 \subset \mathbb{C}$ a esfera unitária em \mathbb{C} com a topologia usual. Com o produto herdado de \mathbb{C} , sabemos que \mathbb{S}^1 é um grupo. Além do mais, o produto e a inversão são contínuas com a topologia induzida por \mathbb{C} .
2. Seja $\mathbb{T}^d = \mathbb{S}^1 \times \cdots \times \mathbb{S}^1$ o produto de \mathbb{S}^1 tomados k vezes. Então \mathbb{T}^d é um grupo com a operação usual de cartesiano de grupos, e claramente as operações de produto e inversão são contínuas na topologia produto.
3. Também com a topologia produto, $\mathbb{S}^{\mathbb{N}} = \{(x_n)_{\mathbb{N}} : n \in \mathbb{N}\}$ é um grupo cuja operação de produto e inversão são contínuas.
4. $GL_n(\mathbb{R}) =$ conjunto das matrizes $n \times n$ com entradas em \mathbb{R} que são invertíveis é um grupo com a operação usual de multiplicação de matrizes. Além do mais, vendo $GL_n(\mathbb{R})$ como um subconjunto de \mathbb{R}^{n^2} , $GL_n(\mathbb{R})$ é um conjunto aberto de \mathbb{R}^{n^2} pois $\det : \mathbb{R}^{n^2} \rightarrow \mathbb{R}$ é contínua e $\det^{-1}(\mathbb{R} \setminus \{0\}) = GL_n(\mathbb{R})$. Além do mais, a operação de produto e inversa são contínuas.
5. Com a topologia induzida de $GL_n(\mathbb{R})$, $O(n) =$ conjunto das matrizes reais e ortogonais, $SL_n(\mathbb{R}) =$ o conjunto das matrizes com determinante igual a 1 e $SO(n) = SL_n(\mathbb{R}) \cap O(n)$ são exemplos de grupos topológicos.

Dado um grupo (G, \cdot) , existem operações naturais que refletem a simetria de um grupo, que são as *translações à esquerda e à direita*.

Definição 6.0.2. *Seja $g \in G$. Definimos $L_g, R_g : G \rightarrow G$ dadas por $L_g(x) = gx, R_g(x) = xg$ e as chamamos de translação à esquerda e à direita por g , respectivamente.*

Note que quando G é um grupo topológico, L_g, R_g são aplicações contínuas $\forall g \in G$. Além do mais, é fácil ver que são aplicações invertíveis com inversas dada por $L_{g^{-1}}$ e $R_{g^{-1}}$, respectivamente. Desta forma, as translações são homeomorfismos. Além do mais, as translações fazem de G um espaço *em homogêneo*: dados $x, y \in G$, tome $g = xy^{-1}$ que vale $L_g y = x$. É neste sentido que dizemos que um grupo é um espaço simétrico, pois a menos de uma simetria, todas as vizinhanças são as mesmas.

Após esta discussão inicial, voltamos a nossa atenção a medidas em G . Antes disso, relembremos a definição de medidas de Borel.

Definição 6.0.3 (Medidas de Borel). *Seja X um espaço topológico. Chamamos de σ -álgebra de Borel a σ -álgebra β gerada pela topologia de X . Todo elemento de β é dito ser um Boreliano. Toda medida definida com respeito a esta σ -álgebra é chamada de medida de Borel.*

A menos de menção do contrário, a σ -álgebra considerada será a σ -álgebra de Borel.

Suponha que G seja um grupo topológico. Uma pergunta natural é a seguinte: existe uma medida de Borel que reflita a simetria do grupo, isto é, que seja invariante por translações. Formalmente, temos a definição abaixo. Introduzimos também algumas notações úteis: dados $a \in G$ e $A \subset G$, escrevemos $gA = \{g \cdot a : a \in A\}$, $Ag = \{a \cdot g : a \in A\}$ e $A^{-1} = \{a^{-1} : a \in A\}$.

Definição 6.0.4. *Seja G um grupo topológico e μ uma medida, não necessariamente de Borel. Dizemos que μ é invariante à esquerda quando $\mu(gA) = \mu(A)$ para todo $g \in G$ e para todo mensurável $A \subset G$. Similarmente, dizemos que μ é invariante à direita quando $\mu(Ag) = \mu(A)$ para todo $g \in G$ e para todo mensurável $A \subset G$.*

A resposta a esta pergunta é positiva, conforme o Teorema abaixo devido ao matemático húngaro Álfred Haar afirma. Além de sua existência, goza de propriedades de regularidade interessantes.

Teorema 6.0.1 (Haar). *Seja G um grupo localmente compacto. Existe uma única medida de Borel μ , a menos de multiplicação por escalar, que satisfaz as seguintes propriedades:*

- (a) μ é invariante à esquerda.
 (b) μ é finita em compactos: $\mu(K) < \infty$ para todo $K \subset G$ compacto.
 (c) μ dá peso positivo a conjuntos abertos: $\mu(U) > 0$ para todo $U \subset G$ aberto e não-vazio.
 (d) μ é regular exteriormente: para qualquer $X \subset G$ boreliano,

$$\mu(X) = \inf\{\mu(A) : A \supset X \text{ aberto}\}.$$

- (e) μ é regular interiormente: Para qualquer $X \subset G$ boreliano,

$$\mu(X) = \inf\{\mu(A) : A \supset X \text{ compacto}\}.$$

Quando o grupo é compacto, a medida do Teorema acima é finita, e é única se exigirmos que ela seja de probabilidade. Então denotamos por μ_G a única medida de probabilidade que satisfaz as propriedades (a)-(e) do Teorema acima e a chamaremos de *medida de Haar de G* .

Propriedades ergódicas das medidas de Haar

Nesta seção, supomos que o grupo G é compacto e metrizável.

Definição 6.0.5. *Seja $d : G \times G \rightarrow [0, +\infty]$ uma métrica em G . Dizemos que d é invariante à esquerda por translações quando $d(gx, gy) = d(x, y) \forall g, x, y \in G$. Similarmente, dizemos que d é invariante à esquerda por translações quando $d(gx, gy) = d(x, y) \forall g, x, y \in G$.*

Quando o grupo possui estas propriedades, podemos selecionar uma métrica compatível com a topologia de G que reflete a sua simetria, isto é, é invariante por translações, conforme a proposição abaixo.

Proposição 6.0.1. *Seja G um grupo compacto metrizável. Então existe alguma métrica $d : G \times G \rightarrow [0, +\infty]$ compatível com a topologia de G que é invariante tanto à esquerda, como à direita. Em outras palavras, as translações à direita e à esquerda são isometrias.*

Seja (X, β_X) um espaço mensurável. Dizemos que uma transformação mensurável $T : X \rightarrow X$ é *unicamente ergódica* quando existe uma única medida de probabilidade invariante μ . Quando X é um espaço topológico e β_X é a σ -álgebra de Borel, uma aplicação contínua $T : X \rightarrow X$ possui propriedades topológicas interessantes.

Proposição 6.0.2. *Seja $T : X \rightarrow X$ unicamente ergódica. Então o suporte da única medida de probabilidade invariante μ é um conjunto minimal. Em particular, T restrito ao suporte é topologicamente misturador.*

Isto nos interessa pelo Teorema abaixo.

Teorema 6.0.2. *Seja G um grupo compacto e metrizável, $g \in G$. São equivalentes:*

- (a) L_g é unicamente ergódica.
- (b) L_g é ergódica com respeito a μ_G .
- (c) o subgrupo $\{g^n : n \in \mathbb{Z}\}$ gerado por g é denso em G .

Em particular, como a medida de Haar dá peso positivo a conjuntos abertos e não-vazios, temos $\text{supp}(\mu_G) = G$. Pela Proposição 6.0.2, μ_G ergódica para L_g implica que G é um conjunto minimal para L_g .

APÊNDICE B - APLICAÇÃO SUPORTE

Neste capítulo, buscamos justificar o porquê da aplicação $\mu \rightarrow \text{supp}(\mu)$ ser mensurável. A referência utilizada é (ALIPRANTIS; BORDER, 2006, Capítulos 17 e 18).

Definição 6.0.6. *Dados X, Y conjuntos, definimos $\mathcal{L} : X \rightrightarrows Y$ uma correspondência uma relação que associa a cada ponto de $x \in X$ um subconjunto $\mathcal{L}(x) \subset Y$.*

Ou seja, uma correspondência é uma função multi-valorada. Outra forma de ver uma correspondência é como uma função de X no conjunto das partes de Y , 2^Y .

Definição 6.0.7. *Dada uma correspondência $\mathcal{L} : X \rightrightarrows Y$ e $A \subset Y$, definimos*

- *A imagem inversa superior \mathcal{L}^u de A como*

$$\mathcal{L}^u(A) = \{x \in X; \mathcal{L}(x) \subset A\}.$$

- *A imagem inversa inferior \mathcal{L}^l de A como*

$$\mathcal{L}^l(A) = \{x \in X; \mathcal{L}(x) \cap A = \emptyset\}.$$

Em correspondências, destacamos algumas noções de continuidade:

Definição 6.0.8. *Sejam X, Y espaços topológicos e $\mathcal{L} : X \rightrightarrows Y$ uma correspondência. Dizemos que a correspondência é*

- *hemicontínua superiormente em um ponto $x \in X$ quando para toda vizinhança U de $\mathcal{L}(x)$, existe uma vizinhança V de x tal que para todo $y \in V$, $\mathcal{L}(y) \subset U$, ou seja, $\mathcal{L}^u(U) \supset V$. Dizemos que \mathcal{L} é hemicontínua superiormente quando o é em todo ponto $x \in X$.*
- *hemicontínua inferiormente em $x \in X$ quando para toda vizinhança U de $\mathcal{L}(x)$ existe vizinhança V de x tal que $y \in V$ implica $\mathcal{L}(y) \cap U \neq \emptyset$, ou seja, $\mathcal{L}^l(U) \supset V$. Dizemos que \mathcal{L} é hemicontínua inferiormente quando o é em todo ponto $x \in X$.*
- *contínua em x quando é hemicontínua superior e inferiormente. Dizemos que \mathcal{L} é contínua quando o é em todo ponto.*

Lema 6.0.1. *sejam X, Y espaços topológicos e $\mathcal{L} : X \rightrightarrows Y$ uma correspondência. As seguintes afirmações são equivalentes:*

1. \mathcal{L} é hemicontínua inferiormente.
2. $\mathcal{L}^l(U)$ é aberto para cada aberto V de Y .
3. $\mathcal{L}^u(F)$ é fechado para cada fechado F de Y .

Seja X espaço métrico compacto. Seja $\mathcal{M}(X)$ o espaço das medidas de probabilidade de Borel, que munido da topologia fraca-* é um espaço compacto e metrizável. A cada $\nu \in \mathcal{M}(X)$, podemos associar o seu suporte, denotado por $\text{supp}(\nu)$. Isto define uma correspondência de $\mathcal{M}(X)$ em X . Seja também $\mathcal{F}(X)$ o espaço dos subconjuntos compactos de X , que munido da métrica de Hausdorff se torna um espaço métrico compacto. Então podemos ver $\nu \rightarrow \text{supp}(\nu)$ como uma aplicação de $\mathcal{M}(X)$ em $\mathcal{F}(X)$.

Proposição 6.0.3. *Seja X um espaço topológico metrizável e separável. Então a correspondência $\nu \mapsto \text{supp}(\nu)$ de X em $\mathcal{F}(X)$ é hemicontínua inferiormente.*

A fim de mostrar a mensurabilidade de $\nu \rightarrow \text{supp}(\nu)$, definimos algumas noções mais fracas de mensurabilidade de correspondências, que darão suporte à prova.

Definição 6.0.9. *Seja (X, β_X) um espaço mensurável e Y um espaço topológico. Dizemos que uma correspondência $\mathcal{L} : X \rightarrow Y$ é*

- fracamente mensurável quando $\mathcal{L}^l(U) \in \beta_X$ para todo U aberto de Y .
- mensurável quando $\mathcal{L}^l(F) \in \beta_X$ para todo F fechado de Y .
- Borel mensurável quando $\mathcal{L}^l(B) \in \beta_X$ para todo subconjunto de borel B de Y .

Desta forma, é consequência do Lema 6.0.1 que correspondências hemicontínuas inferiormente são fracamente mensuráveis.

O que ajuda o nosso caso é o teorema a seguir.

Teorema 6.0.3. *Seja (X, β_X) um espaço mensurável e seja Y um espaço metrizável e separável. Seja $\mathcal{L} : X \rightarrow Y$ uma correspondência com a propriedade de que $\mathcal{L}(x)$ é um compacto não-vazio de Y para todo $x \in X$. Então as seguintes afirmações são equivalentes:*

1. A correspondência \mathcal{L} é fracamente mensurável.
2. A correspondência \mathcal{L} é mensurável.
3. Vendo a correspondência como a função $f : X \rightarrow \mathcal{F}(Y)$ dada por $f(x) = \mathcal{L}(x)$, f é uma aplicação Borel mensurável no sentido de que $f^{-1}(B) \in \beta_X$ para todo $B \in \mathcal{B}_{\mathcal{F}(Y)}$.

Por fim, deste teorema tiramos o corolário que queremos:

Corolário 6.0.1. *Seja X um espaço métrico compacto. A aplicação $f : \mathcal{M}(X) \rightarrow \mathcal{F}(X)$ é mensurável.*

Demonstração. Como observamos, $\nu \rightarrow \text{supp}(\nu)$ é hemicontínua inferiormente, donde fracamente mensurável. Pelo Teorema 6.0.3, temos que f é mensurável. \square

APÊNDICE C - SUPORTE DE UMA DESINTEGRAÇÃO

Relembremos algumas definições.

Seja μ uma medida de Borel de um espaço métrico compacto X . Chamamos de *suporte* de μ ao conjunto dos pontos $x \in X$ tais que $\mu(U) > 0$ para toda vizinhança aberta de x . Denotamos o suporte de μ por $\text{supp}(\mu)$. A ideia é que fora do suporte, topologicamente falando, o espaço é desprezível. Isto está formalizado em essência na proposição abaixo.

Proposição 6.0.4. *Seja $\mathcal{U}_\mu = \{U \subset X \text{ aberto} : \mu(U) = 0\}$. As seguintes afirmações são verdadeiras:*

- (a) $\text{supp}(\mu) = X \setminus \bigcup_{U \in \mathcal{U}_\mu} U$.
- (b) $\text{supp}(\mu)$ é um conjunto fechado.
- (c) $\mu(\text{supp}(\mu)) = 1$.

Demonstração. (a) Dado $U \in \mathcal{U}_\mu$, temos que $\mu(U) = 0$. Então é claro que $U \subset X \setminus \text{supp}(\mu)$, donde $\bigcup_{U \in \mathcal{U}_\mu} U \subset X \setminus \text{supp}(\mu) \Rightarrow \text{supp}(\mu) \subset X \setminus \bigcup_{U \in \mathcal{U}_\mu} U$. Por outro lado, para $x \in X \setminus \bigcup_{U \in \mathcal{U}_\mu} U$ e qualquer vizinhança V de x , devemos ter $\mu(V) > 0$, pois do contrário $\mu(V) = 0$ e $x \in V \subset X \setminus \text{supp}(\mu)$, uma contradição. Disto segue que $X \setminus \bigcup_{U \in \mathcal{U}_\mu} U \subset \text{supp}(\mu)$. Isto prova que $\text{supp}(\mu) = X \setminus \bigcup_{U \in \mathcal{U}_\mu} U$.

(b) Note que $\bigcup_{U \in \mathcal{U}_\mu} U$ é um conjunto aberto por ser uma união arbitrária de conjuntos abertos. Pelo item (a), temos que $\text{supp}(\mu) = X \setminus \bigcup_{U \in \mathcal{U}_\mu} U$ é fechado.

(c) Por ser espaço métrico compacto, X possui base enumerável. Espaços com base enumerável são espaços de Lindelöf: toda cobertura de abertos possui subcobertura enumerável. Desta forma, $\bigcup_{U \in \mathcal{U}_\mu} U$ é um aberto de X e por isso possui base enumerável.

Isto implica que existem $U_1, \dots, U_n, \dots \in \mathcal{U}_\mu$ tais que $\bigcup_{U \in \mathcal{U}_\mu} U = \bigcup_{n=1}^{\infty} U_n$. Como U_n tem medida nula para todo n , é claro que $\bigcup_{U \in \mathcal{U}_\mu} U = \bigcup_{n=1}^{\infty} U_n$ também tem medida nula por ser a união enumerável de conjuntos de medida nula. Desta forma, por (a), $\text{supp}(\mu)$ tem medida total em X .

□

Denote por $\mathcal{M}(X)$ o conjunto das medidas de probabilidade de X .

Sejam (X, β_X, μ, T) e (Y, β_Y, ν, S) sistemas que preservam a medida, onde X e Y são espaços métricos compactos, β_X e β_Y são a σ -álgebra de Borel de X e Y , respectivamente, e $T : X \rightarrow X$ $S : Y \rightarrow Y$ contínuas. Seja $\pi : X \rightarrow Y$ uma aplicação de fator mensurável. Pelo Teorema 2.4.1, existe uma desintegração de fator $z \in Y \rightarrow \eta_z \in \mathcal{M}(X)$.

Exemplo. Seja $X = [0, 1] \times [0, 1]$, $Y = [0, 1]$, $T = \text{Id}_{[0,1] \times [0,1]}$, $S = \text{Id}_{[0,1]}$ e $\pi : X \rightarrow Y$ definida por $\pi(x, y) = x$. Sejam também $\mu = \text{Leb}([0, 1] \times [0, 1])$ e $\nu = \text{Leb}([0, 1])$. Neste setup, fica claro que (Y, β_Y, ν, S) é fator de (X, β_X, μ, T) , com aplicação de fator $\pi : X \rightarrow Y$. Seja $\eta_x = \delta_x \times \text{Leb}([0, 1])$, onde δ_x é a medida de Dirac concentrada em x . Pelo Teorema de Fubini, temos que $x \in Y \rightarrow \mathcal{M}(X)$ é a desintegração associada à aplicação de fator. Ainda no exemplo acima, note que $\text{supp}(\eta_x) = \{x\} \times [0, 1]$. Desta forma, $v \in \text{supp}(\eta_{\pi(v)})$ para todo $v \in X$. Na verdade, este fato é bem geral, quando assumimos que a aplicação de fator $\pi : X \rightarrow Y$ é contínua.

Antes de seguir, introduzimos algumas notações. Seja $\mathcal{F}(X)$ a família de subconjuntos fechados e não-vazios de um espaço métrico compacto (X, d) dado. Defina em $\mathcal{F}(X)$ a *métrica de Hausdorff* H , dada por

$$H(F, G) = \max \left\{ \sup_{x \in F} d(x, G), \sup_{y \in G} d(y, F) \right\},$$

para todo $F, G \in \mathcal{F}(X)$.

Lema 6.0.2. *Sejam W um espaço métrico compacto, $\mathcal{M}(W)$ o espaço das medidas de probabilidade de Borel em W munido da topologia fraca- $*$ e $F(W)$ o espaço dos subconjuntos fechados e não-vazios de W munido da métrica de Hausdorff. Então:*

1. *A transformação $\nu \in \mathcal{M}(W) \mapsto \text{supp}(\nu) \in F(W)$ é Borel-mensurável.*
2. *Se $x \in W \mapsto \rho_x \in \mathcal{M}(W)$ é mensurável, então $\{x \in W : x \in \text{supp}(\rho_x)\}$ é um conjunto de Borel.*

Demonstração. A prova de 1 pode ser encontrada em (ALIPRANTIS; BORDER, 2006, Capítulos 17 e 18). No apêndice 6, resumimos os resultados usados deste livro.

2. Seja $\psi_1 : W \rightarrow \mathcal{F}(W)$ dada por $\psi_1(x) = \{x\}$, que é contínua (de fato, uma isometria). Seja também $\psi_2 : W \rightarrow \mathcal{F}(W)$ dada por $\psi_2(x) = \text{supp}(\rho_x)$. Sabemos da parte 1 que ψ_2 é mensurável, por ser composição de transformações mensuráveis. Logo, $\psi : W \rightarrow \mathcal{F}(W) \times \mathcal{F}(W)$ dado por $\psi(x) = (\{x\}, \text{supp}(\rho_x))$ é mensurável. Por outro lado, note que

$$\Omega = \{(F_1, F_2) \in \mathcal{F}(W) \times \mathcal{F}(W) : F_1 \cap F_2 \neq \emptyset\}$$

é fechado. Para tal, basta verificar que o complementar é aberto, cuja prova é a seguinte: se $(F, Q) \in \mathcal{F}(W) \times \mathcal{F}(W)$ são tais que $F \cap Q \neq \emptyset$, consideramos $d = \min_{x \in F, y \in Q} d(x, y)$ e notamos que se $F', Q' \in \mathcal{F}(W)$ satisfazem $H(F, F') < d/2$ e $H(Q, Q') < d/2$, então $F' \cap Q' \neq \emptyset$. Em particular, $\Omega \subset \mathcal{F}(W) \times \mathcal{F}(W)$ é mensurável, e daí $\psi^{-1}(\Omega)$ também o é. Para encerrar, basta notar que este último conjunto é justamente o conjunto que queríamos estudar, pois

$$\begin{aligned} \psi^{-1}(\Omega) &= \{x \in W : \psi(x) \in \Omega\} = \{x \in W : (\{x\}, \text{supp}(\rho_x)) \in \Omega\} \\ &= \{x \in W : \{x\} \cap \text{supp}(\rho_x) \neq \emptyset\} = \{x \in W : x \in \text{supp}(\rho_x)\}. \end{aligned}$$

□

O lema abaixo é onde utilizamos fortemente a continuidade do fator. Ele é utilizado para a prova da proposição 5.3.1.

Lema 6.0.3. *A desintegração $z \rightarrow \eta_z$ satisfaz*

$$\mu(\{x \in X; x \in \text{supp}(\eta_{\pi(x)})\}) = 1.$$

Demonstração. Por conveniência, escreva $G = \{x \in X; x \in \text{supp}(\eta_{\pi(x)})\}$. Como $x \rightarrow \pi(x)$ e $z \rightarrow \eta_z$ são mensuráveis, a composição $x \rightarrow \eta_{\pi(x)}$ também o é. Pelo lema 6.0.2, G é mensurável, e portanto o enunciado faz sentido: podemos tomar a medida de G . Como $\mu(G) = \int_Y \eta_z(G) d\nu(z)$, basta mostrar que $\eta_z(G) = 1$ para m -quase todo $z \in Z$. Por ser desintegração, temos $\eta_z(\pi^{-1}(z)) = 1$ para m -quase todo $z \in Z$. Isto implica que $\text{supp}(\eta_z) \subset \pi^{-1}(z)$: caso contrário, existe $x \in \text{supp}(\eta_z)$ mas $x \notin \pi^{-1}(z)$. Como π é contínua, $\pi^{-1}(z)$ é fechado, e como estamos em um espaço métrico compacto, $d(x, \pi^{-1}(z)) > 0$. Logo, existe $\varepsilon > 0$ tal que $B(x, \varepsilon) \cap \pi^{-1}(z) \neq \emptyset$. Contudo, $x \in \text{supp}(\eta_z)$, donde $\eta_z(B(x, \varepsilon)) > 0$. Logo, como η_z é medida de probabilidade, temos

$$1 \geq \eta_z(B(x, \varepsilon) \sqcup \pi^{-1}(z)) = \eta_z(B(x, \varepsilon)) + \eta_z(\pi^{-1}(z)) = \eta_z(B(x, \varepsilon)) + 1 > 1,$$

contradição. Então devemos ter $\text{supp}(\eta_z) \subset \pi^{-1}(z)$. Mas então $\text{supp}(\eta_z) \subset G$ para m -quase todo z : de fato, para m -quase todo z , $\text{supp}(\eta_z) \subset \pi^{-1}(z)$. Daí, para todo $x \in \text{supp}(\eta_z) \subset \pi^{-1}(z)$ é tal que $\pi(x) = z$, donde $x \in \text{supp}(\eta_{\pi(x)})$, o que implica $\text{supp}(\eta_z) \subset G$ para m -quase todo z . Pela proposição 6.0.4, temos que $\eta_z(\text{supp}(\eta_z)) = 1$. Daí, $G \supset \text{supp}(\eta_z) \Rightarrow \eta_z(G) \geq \eta_z(\text{supp}(\eta_z)) = 1$, e pois $\eta_z(G) = 1$, para m -quase todo $z \in Z$. E acabou. □