Universidade Federal Fluminense

Programa de Pós-Graduação em Matemática

O teorema de Szemerédi

Atila Pianca Guidolini

Niterói, Dezembro de 2007

O TEOREMA DE SZEMERÉDI

Atila Pianca Guidolini

COMISSÃO EXAMINADORA

Prof.ª Dr.ª Isabel Lugão Rios - Orientadora
UFF - Universidade Federal Fluminense

Prof. Dr. Carlos Gustavo Moreira - Membro
IMPA - Instituto de Matemática Pura e Aplicada

Prof. Dr. Samuel Santi - Membro
UFRJ - Universidade Federal do Rio de Janeiro

Prof. Dr. Renaud Leplaideur - Membro
Université de Bretagne Ocidentale

Prof.ª Dr.ª Anne Michelle Dysman - Membro
UFF - Universidade Federal Fluminense

Dissertação apresentada por Atila Pianca Guidolini ao Curso de Mestrado em Matemática - da Universidade Federal Fluminense, como requisito parcial para a obtenção do Grau de Mestre. Linha de Pesquisa: Sistemas Dinâmicos.

Niterói (RJ), 11 de Dezembro de 2007

G948 Guidolini, Atila Pianca

O teorema de Szemerédi/Atila Pianca Guidolini - Niterói, RJ : [s. n.], 2007.

68 f.

Orientadora: Isabel Lugão Rios.

Dissertação (Mestrado en Matemática)- Universidade Federal Fluminense, 2007.

1. Sistmeas Dinâmicos. 2. Teoria Ergódica. I. Título. II. Universidade Federal Fluminense. Instituto de Matemática. Coordenação de Pós-Graduação em Matemática.

CDD:515:42

Agradecimentos

Gostaria de começar agradecendo a minha professora orientadora Isabel Lugão Rios, pelas sempre sinceras palavras de incentivo nas horas de dificuldades e por me confiar tão agradável tema.

Ao Colegiado de Pós-Graduação, em especial à pessoa do professor Dinamérico Pombo Júnior.

Ao sempre presente amigo Fábio Luiz.

Àqueles que na ausência e distância da família, tiveram fundamental importância na minha vida. Os grandes amigos de república; Gabriel, Estevão, Vinicius e Luciano.

E ainda, as demais amizades que conquistei ou fortifiquei nesses últimos anos; Laura, Camila, Mônica, Fernanda, Glauber e alguns outros.

Aos amigos distantes que tanto me fazem falta; Cássio, Leomar, Lidiane, Carlos, Ângela, Sandra, Joana, e outros tantos.

Aos amigos do Movimento Escoteiro.

Aos meus familiares que tanto torceram pela realização desse trabalho; Tios, Primos e Avós.

À Bruna por sua doçura, paciência e amor.

Aos meus afilhados, Paulo Octávio e Loise Helena, que sempre alegram a minha vida.

Aos meus amados Irmãos, Eric e Icaro, que são tão importantes na minha vida.

Aos meus pais, Antônio e Lourdes, a quem tanto amo. Talvez esse trabalho, mais que um sonho meu, seja um sonho deles.

Por fim, agradeço a Deus. Por sempre me abençoar, e por tornar minha página de agradecimentos tão valorosa.

Sumário

Introdução

O tema para esta dissertação foi escolhido com o propósito de apresentar aplicações da teoria de sistemas dinâmicos à teoria dos números. Dentre os possíveis problemas foi escolhido o Teorema de Szemerédi. O fato interessante na história deste problema é que após conjecturado levou-se décadas até uma prova completa e depois uma generalização com o uso de Teoria Ergódica.

Em 1975, E. Szemerédi [14] provou que subconjuntos de números inteiros com densidade superior positiva contêm progressões aritméticas finitas com comprimento arbitrariamente grande, problema que havia sido conjecturado por Erdös e Turan [2], em 1936. Em 1976, H. Furstenberg [4] notou que o argumento demonstrado por Szemerédi era equivalente a um argumento de "recorrência múltipla" de transformações preservando medida, tornando possível uma prova ergódica para o teorema. Na verdade, Furstenberg foi além, demonstrando uma versão do Teorema de Szemerédi para subconjuntos de $\mathbb{Z}^r$.

Vários elementos da prova original de Furstenberg foram simplificados por Y. Katznelson e D. Ornstein no artigo "The ergodic theoretical proof of Szemerédi's theorem" [6], artigo este em que é baseada essa dissertação.

Essa dissertação começa com algumas definições e resultados básicos no Capítulo 1 como pré-requisitos à compreensão do desenvolvimento dos capítulos seguintes.

No Capítulo 2, são introduzidos alguns resultados sobre extensões de espaços de medida. Seu conteúdo é uma compilação simplificada sobre extensões, extraído do livro "Ergodic behavior and a theorem and combinatorial number theory" de Furstenberg [5]. Esse capítulo tem papel fundamental na compreensão dos resultados apresentados no final da dissertação.

O Teorema de Szemerédi é apresentado no Capítulo 3. Nesse capítulo é enunciado o Teorema de Recorrência Múltipla de Furstenberg e mostra-se que tomando este resultado válido é possível demonstrar o Teorema de Szemerédi. No fim do capítulo é mostrado que esses resultados são equivalentes.

O Capítulo 4 trata do Teorema de Recorrência Múltipla de Furstenberg. Primeiro este resultado é demosntrado para duas classes de espaços de medidas. Depois é apresentado um esquema da demonstração geral.

Capítulo 1

Preliminares

Começamos nosso trabalho apresentando algumas definições e resultados em Teoria da Medida e Teoria Ergódica que serão utilizados aqui. As demonstrações são omitidas por não possuirem relevância direta no desenvolvimento dos demais capítulos. Para grande parte dos resultados, indicamos algumas referências para a consulta das mencionadas demonstrações.

1.1 Teoria da Medida

Nesta seção recordaremos algumas noções e resultados básicos da Teoria da Medida. As demonstrações podem ser encontradas em [1], [3] e [13].

1.1.1 Espaços mensuráveis

Começamos por introduzir as noções de Álgebra e σ-álgebra de subconjuntos.

Definição 1.1. *Uma álgebra de subconjuntos de X é uma família $\mathcal{B}$ de subconjuntos tal que:*

- $X \in \mathcal{B}$,
- $A \in \mathcal{B} \Rightarrow A^c \in \mathcal{B}$,
- $A \in \mathcal{B}$ *e* $A' \in \mathcal{B} \Rightarrow A \cup A' \in \mathcal{B}$.

Segue destas propriedades que:

$$A \in \mathcal{B} \text{ e } B \in \mathcal{B} \Rightarrow A \cap B = (A^c \cup B^c)^c \in \mathcal{B}, A \in \mathcal{B} \text{ e } B \in \mathcal{B} \Rightarrow A - B = A \cap B^c \in \mathcal{B}.$$

Definição 1.2. *Uma álgebra $\mathcal{B}$ diz-se uma σ-álgebra de X se*

$$A_j \in \mathcal{B}, j = 1, 2, \ldots, n, \ldots \Rightarrow \bigcup_{j=1}^{\infty} A_j \in \mathcal{B}.$$

Observação 1.1. *Temos que $\mathcal{B}$ também é fechada para interseções enumeráveis: se $A_j \in \mathcal{B}$ para $j = 1, 2, \ldots, n, \ldots$ então $\bigcap_{j=1}^{\infty} A_j = \left(\bigcup_{j=1}^{\infty} {A_j}^c \right)^c \in \mathcal{B}$.*

Definição 1.3. *Um espaço mensurável é uma dupla $(X, \mathcal{B})$ onde X é um conjunto e $\mathcal{B}$ é uma σ-álgebra de subconjuntos de X. Os elementos de $\mathcal{B}$ são chamados conjuntos mensuráveis.*

Definição 1.4. *A σ-álgebra gerada por uma família $\mathcal{E}$ de subconjuntos de X é a menor σ-álgebra que contém a família $\mathcal{E}$.*

No caso em que X vem munido da estrutura de espaço topológico, há uma escolha natural para $\mathcal{E}$, nomeadamente, o conjunto dos subconjuntos abertos. Isto nos conduz à noção de *σ-álgebra de Borel*.

Definição 1.5. *Seja (X, τ) um espaço topológico, então a σ-álgebra de Borel de M é a σ-álgebra gerada por τ, ou seja, a menor σ-álgebra que contém todos os subconjuntos abertos.*

1.1.2 Espaços de medida

Introduziremos o conceito de medida e em seguida apresentaremos alguns resultados sobre a construção de medidas.

Definição 1.6. *Uma medida num espaço mensurável $(X, \mathcal{B})$ é uma função $\mu : \mathcal{B} \to [0, +\infty]$ que satisfaz:*

1. $\mu(\emptyset) = 0$;

2. $\mu(\bigcup_{j=1}^{\infty} A_j) = \sum_{j=1}^{\infty} \mu(A_j)$ *para quaisquer* $A_j \in \mathcal{B}$ *disjuntos dois-a-dois.*

A tripla $(X, \mathcal{B}, \mu)$ é chamada de *espaço de medida.* Quando $\mu(X) = 1$ dizemos que μ é uma medida de *probabilidade* e $(X, \mathcal{B}, \mu)$ é um *espaço de probabilidade.*

A segunda propriedade na definição de medida é chamada a *σ-aditividade.* Dizemos que uma função $\mu : \mathcal{B} \to [0, +\infty]$ é *finitamente aditiva* se:

$$\mu\left(\bigcup_{j=1}^{N} A_j\right) = \sum_{j=1}^{N} \mu\left(A_j\right)$$

para qualquer família finita $A_1, \ldots, A_N \in \mathcal{B}$ de subconjuntos disjuntos dois-a-dois. Note que toda medida é, automaticamente, finitamente aditiva.

Exemplo 1.1. *Seja X um conjunto e consideremos a σ-álgebra $\mathcal{B} = 2^X$, onde 2^X é o conjunto das partes de X. Dado qualquer $p \in X$, consideremos a função $\delta_p : 2^X \to [0, +\infty]$ definida por:*

$$\delta_p(A) = \begin{cases} 1 & se\ p \in A, \\ 0 & se\ p \in A. \end{cases}$$

Temos que δ_p é uma medida, que é usualmente designada por delta de Dirac no ponto p.

Se $(X, \mathcal{B}, \mu)$ é um espaço de medida, dizemos que um conjunto $A \subset X$ é de medida nula se existe $A_1 \in \mathcal{B}$ tal que $A \subset A_1$ e $\mu(A_1) = 0$. Uma propriedade aplicável a pontos de um conjunto $S \subset X$ vale em *quase todo ponto* $x \in S$ (abreviadamente q.t.p.), se o conjunto dos pontos de S onde a propriedade é falsa for um conjunto de medida nula.

Teorema 1.1 (Extensão)**.** *Seja $\mathcal{B}_0$ uma álgebra de subconjuntos de X e seja $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ uma função finita, finitamente aditiva. Então existe uma única função finita, finitamente aditiva $\mu : \mathcal{B} \to [0, +\infty]$ que é uma extensão de μ_0 à σ-álgebra $\mathcal{B}$ gerada por $\mathcal{B}_0$. Se μ_0 é σ-aditiva então μ também o é.*

Teorema 1.2 (σ-aditividade)**.** *Seja $\mathcal{B}_0$ uma álgebra e seja $\mu_0 : \mathcal{B}_0 \to [0, +\infty]$ uma função finitamente aditiva com $\mu_0(X) = 1$. Suponha que que*

$$\lim_{n \to \infty} \mu_0\left(\bigcap_{j=1}^{n} A_j\right) = 0$$

para toda a sequência $A_1 \supset \ldots \supset A_j \supset \ldots$ de conjuntos mensuráveis tal que $\cap_{j=1}^{\infty} A_j = \emptyset$. Então μ_0 é σ-aditiva.

O resultado seguinte nos diz que todo elemento A de uma σ-álgebra $\mathcal{B}$ gerada por uma álgebra $\mathcal{B}_0$ é aproximado por um elemento A_0 da álgebra, no sentido em que a medida da diferença simétrica, $\mu(A\Delta A_0) = \mu(A \cup A_0 - A \cap A_0)$, é pequena.

Teorema 1.3 (Aproximação)**.** *Seja $(X, \mathcal{B}, \mu)$ um espaço de probabilidade e seja $\mathcal{B}_0$ uma álgebra que gera a σ-álgebra $\mathcal{B}$. Então para todo $\epsilon > 0$ e todo $A \in \mathcal{B}$ existe $A_0 \in \mathcal{B}_0$ tal que $\mu(A\Delta A_0) < \epsilon$.*

A seguir definiremos uma medida, chamada medida produto, no espaço das seqüências.

Consideremos os espaços de probabilidade $(X_i, \mathcal{B}_i, \mu_i)$, com $i \in \mathbb{Z}$. Vamos construir uma probabilidade μ no conjunto

$$X = \prod_{i=-\infty}^{\infty} X_i$$

das sequências bilaterais $(x_i)_{i=-\infty}^{\infty}$, com $x_i \in X_i$ para cada i. Mais precisamente, a medida μ será definida na *σ-álgebra produto* $\mathcal{B}$ das σ-álgebras $\mathcal{B}_i$, que é caracterizada do seguinte modo: dados inteiros $m \leq n$ e conjuntos $A_j \in \mathcal{B}_j$ para $m \leq j \leq n$, consideremos

$$[m; A_m, \ldots, A_n] = \left\{(x_i)_{i\in\mathbb{Z}} : x_j \in A_j \text{ para } m \leq j \leq n\right\}.$$

Estes subconjuntos de X são chamados *cilindros*. A família $\mathcal{B}_0$ das uniões finitas de cilindros disjuntos dois-a-dois é uma álgebra. Por definição, a σ-álgebra produto $\mathcal{B}$ é a σ-álgebra gerada por $\mathcal{B}_0$. Para construir a medida μ procedemos do seguinte modo: primeiramente, consideramos a aplicação τ definida na família dos cilindros por

$$\tau\left([m; A_m, \ldots, A_n]\right) = \prod_{j=m}^{n} \mu_j(A_j)$$

Em seguida estendemos τ à álgebra $\mathcal{B}_0$, estipulando que a imagem de qualquer união finita de cilindros disjuntos dois-a-dois é igual à soma das imagens dos cilindros. Esta extensão está bem definida e é finitamente aditiva. Então, recorrendo aos Teoremas 1.1 e 1.2, obtemos uma medida de probabilidade μ em $(X, \mathcal{B})$ que estende τ.

Definição 1.7. *O espaço de probabilidade* $(X, \mathcal{B}, \mu)$ *construído acima é designado produto direto dos espaços* $(X_i, \mathcal{B}_i, \mu_i)$*.*

1.1.3 Integração em espaços de medida

Esta seção apresentará noções da definição de integral de funções em espaços de medida, como feito em [10].

Definição 1.8. *Seja* $\mathcal{B}(\mathbb{R})$ *a* σ*-álgebra de Borel de* $\mathbb{R}$*. Uma função* $f : X \to \mathbb{R}$ *é mensurável se* $f^{-1}(D) \in \mathcal{B}$ *para todo* $D \in \mathcal{B}(\mathbb{R})$*.*

De modo análogo, sejam X e Y dois conjuntos e $\mathcal{B}$ e $\mathcal{D}$ σ-álgebras de subconjuntos de X e Y respectivamente. Uma aplicação $T : X \to Y$ é mensurável com respeito às σ-álgebras $\mathcal{B}$ e $\mathcal{D}$ se $T^{-1}(B) \in \mathcal{B}$ para todo $B \in \mathcal{D}$.

Seja $(X, \mathcal{B}, \mu)$ um espaço de medida. Se $A \in \mathcal{B}$ denotamos 1_A à *função característica* de A, isto é, $1_A(x)$ é igual a 1 se $x \in A$ e zero caso contrário.

Dizemos que uma função $f : X \to \mathbb{R}$ é *simples* se existem constantes $\lambda_1, \ldots, \lambda_N \in \mathbb{R}$ e conjuntos $A_1, \ldots, A_N \in \mathcal{B}$ disjuntos dois-a-dois tais que $f = \sum_{i=1}^{N} \lambda_i 1_{A_i}$. Dizemos que uma função simples é integrável se

$$\sum_{i=1}^{N} \lambda_i \mu(A_i) < +\infty$$

e neste caso definimos a integral de f como:

$$\int_X f d\mu = \sum_{i=1}^{N} \lambda_i \mu(A_i)$$

Pode-se provar que este número independe da forma de escrever f como combinação linear de funções características. Dizemos que uma função $f : X \to \mathbb{R}$ é *integrável* se existe uma sequência $f_n : X \to \mathbb{R}$, $n = 1, 2, \ldots$ de funções simples integráveis tais que

$$\lim_{n\to\infty} f_n(x) = f(x) \quad \text{q.t.p.} \quad x \in X \tag{1.1}$$

Neste caso definimos a *integral* de f com

$$\int_X f d\mu = \lim_{n\to\infty} \int_X f_n d\mu \tag{1.2}$$

Prova-se que o limite em (1.2) existe e que independe da sequência f_n. Dizemos que $f : X \to \mathbb{R}$ é integrável sobre o conjunto $A \in \mathcal{B}$ se $f 1_A$ é integrável e definimos:

$$\int_X f d\mu = \int_X f 1_A d\mu.$$

As funções integráveis não precisam ser mensuráveis, mas sempre coincidem com uma mensurável em q.t.p..

Teorema 1.4. *Se $f : X \to \mathbb{R}$ é integrável, para todo $\epsilon > 0$ existe $\delta > 0$ tal que:*

$$\int_A |f|\, d\mu \leq \epsilon$$

se $A \in \mathcal{B}$ e $\mu(A) \leq \delta$.

Seja $p \geq 1$. Denota-se $L^p(X, \mathcal{B}, \mu)$ o conjunto das funções $f : X \to \mathbb{R}$ tais que $|f|^p$ é integrável, identificando funções que coincidem em q.t.p.. Em $L^p(X, \mathcal{B}, \mu)$ definimos a norma $\|.\|_p$ como:

$$\|f\|_p = \left(\int_X |f|^p d\mu \right)^{1/p}.$$

O fato de $\|.\|_p$ ser uma norma segue da *desigualdade de Minkowski* que estabelece que, se f e g pertencem a $L^p(X, \mathcal{B}, \mu)$, então $f + g \in L^p(X, \mathcal{B}, \mu)$ e

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p.$$

Proposição 1.1. *O espaço $L^p(X, \mathcal{B}, \mu)$ é completo com a norma $\|.\|_p$, $1 \leq p < +\infty$.*

Proposição 1.2. *O espaço $L^p(X, \mathcal{B}, \mu)$ é um espaço de Hilbert se, e somente se, $p = 2$.*

Define-se $L^\infty(X, \mathcal{B}, \mu)$ como o conjunto das $f : X \to \mathbb{R}$ tais que, existe $K > 0$ tal que $|f(x)| \leq K$ para q.t.p. $x \in X$ identificando funções que coincidem em q.t.p.. O ínfimo dos K com esta propriedade denota-se $\|f\|_\infty$ e define uma norma em $L^\infty(X, \mathcal{B}, \mu)$. Para todo $1 \leq p \leq \infty$, $L^p(X, \mathcal{B}, \mu)$

provido da norma $\|.\|_p$ é um espaço de Banach. A *desigualdade de Hölder* estabelece que, se $1 \leq p \leq \infty$ e q satisfaz $1/p+1/q = 1$, se $p < \infty$, ou $q = 1$ e $p = \infty$, então se $f \in L^p(X, \mathcal{B}, \mu)$ e $g \in L^q(X, \mathcal{B}, \mu)$, o produto fg pertence a $L^1(X, \mathcal{B}, \mu)$ e satisfaz

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

1.2 Teoria Ergódica

Nesta seção apresentaremos os elementos básicos de Teoria Ergódica utilizados ao longo do trabalho. Esta seção está dividida em três subseções. Na primeira apresentamos um Teorema de Recorrência como um instrumento motivacional ao interesse pelo Teorema de Recorrência Múltipla apresentado no capítulo 4. Na seguinte um teorema, cuja relevância o concedeu o nome de Teorema Ergódico. Por último, alguns resultados úteis sobre convergência de medidas.

1.2.1 Teorema de Recorrência de Poincaré

O resultado que apresentaremos nesta seção, enunciado por Poincaré perto do fim do século XIX, afirma que *quase todo ponto retorna a um conjunto* quando iterado várias vezes por uma determinada transformação. Uma Demonstração pode ser encontrada em [15].

Diremos que T é uma *transformação preservando medida* de $(X, \mathcal{B}, \mu)$ se T é uma transformação de X em X de modo que se $B \in \mathcal{B}$, então $T^{-1}B \in \mathcal{B}$ e $\mu(T^{-1}B) = \mu(B)$. Diremos que a quadrupla $(X, \mathcal{B}, \mu, T)$ é um *sistema preservando medida (s.p.m)*.

Teorema 1.5 (Teorema de Recorrência de Poincaré)**.** *Seja $T : X \to X$ uma transformação preservando medida de um espaço de probabilidade $(X, \mathcal{B}, \mu)$. Seja $E \in \mathcal{B}$ com $\mu(E) > 0$. Então μ-quase todo ponto de E retorna infinitas vezes a E sobre iterações positivas por T, isto é, existe $F \subset E$ com $\mu(F) = \mu(E)$ tal que para cada $x \in F$ existe uma seqüência $n_1 < n_2 < n_3 < \ldots$ de números naturais com $T^{n_i}(x) \in F$ para cada i.*

No Capítulo 4, tratamos do *Teorema de Recorrência Múltipla de Furstenberg*, que afirma que dado um inteiro k e um conjunto A de medida positiva, existe algum tempo n tal que os iterados T^{in} dos elementos de um subconjunto de medida positiva de A regressam a A para todo $i = 0, \ldots, k-1$ (ou seja, tais elementos retornam *simultaneamente* k vezes por iterações de T^n). Alguns autores chamam este teorema de *Teorema de Recorrência Simultânea de Poincaré* pelo fato deste ser uma generalização do teorema acima.

1.2.2 Teorema Ergódico de Birkhoff

A seguir temos uma proposição caracterizando quando uma medida é invariante, a prova pode ser encontrada em [7].

Proposição 1.3. *Seja $T : X \to X$ uma transformação e μ uma medida. Então T preserva μ se, e somente se, para toda função integrável $f : X \to \mathbb{R}$ vale:*

$$\int f d\mu = \int f \circ T d\mu.$$

Definição 1.9. *O sistema de probabilidade $(X, \mathcal{B}, \mu, T)$ é um sistema ergódico preservando medida se $T^{-1}B = B$ para $B \in \mathcal{B}$ implicar que $\mu(B) = 0$ ou $\mu(B) = 1$. Neste caso diremos que T é uma transformação ergódica.*

Um resultado fundamental é o Teorema Ergódico de Birkhoff, cuja prova pode ser encontrado também em [7].

Teorema 1.6 (Teorema Ergódico de Birkhoff)**.** *Seja $f \in L^1(X, \mathcal{B}, \mu)$ e T uma transformação preservando medida de $(X, \mathcal{B}, \mu)$, então para quase todo x,*

$$\frac{f(x) + f(Tx) + \ldots + f(T^n x)}{n+1} \to \bar{f}(x)$$

onde $\bar{f}(x)$ é uma função T-invariante, isto é, $\bar{f} = \bar{f} \circ T$. Além disso,

$$\int \bar{f}(x) d\mu(x) = \int f(x) d\mu(x).$$

Se $(X, \mathcal{B}, \mu, T)$ é ergódico, então uma função satisfazendo $\bar{f} = \bar{f} \circ T$ é necessariamente quase sempre constante. De fato, a pré-imagem $\bar{f}^{-1}(I)$ para qualquer intervalo $I \subset \mathbb{R}$ é um conjunto invariante, portanto, pela ergodicidade, essa pré-imagem tem medida zero ou um. Como o intervalo I é qualquer, a função $\bar{f}$ é constante num conjunto de medida total. Disto segue que quando T é ergódico, $\bar{f} = \int f d\mu$, e que

$$\frac{f(x) + f(Tx) + \ldots + f(T^n x)}{n+1} \to \int f(y) d\mu(y).$$

Em [15] encontramos o seguinte corolário:

Corolário 1.1 (Teorema L^p Ergódico de Von Neumann)**.** *Seja $1 \leq p < \infty$ e seja $(X, \mathcal{B}, \mu, T)$ um s.p.m.. Se $f \in L^p(X, \mathcal{B}, \mu)$ existe $\bar{f} \in L^p(X, \mathcal{B}, \mu)$ com $\bar{f} = \bar{f} \circ T$ em quase todo ponto, e*

$$\frac{f(x) + f(Tx) + \ldots + f(T^n x)}{n+1} \to \bar{f}(x)$$

na norma de $L^p(X, \mathcal{B}, \mu)$.

1.2.3 Topologia fraca*

Nesta seção vamos introduzir uma topologia importante no conjunto $\mathcal{M}(X)$ das probabilidades borelianas do espaço X, chamada topologia fraca*. Para um estudo das demonstrações aqui omitidas consultar [7].

Dada uma medida $\mu \in \mathcal{M}(X)$, um conjunto finito $F = \{\phi_1, ..., \phi_N\}$ de funções contínuas $\phi_j : X \to \mathbb{R}$, e um número $\epsilon > 0$, definimos

$$V(\mu, F, \epsilon) = \{\eta \in \mathcal{M}(X) : \left| \int \phi_j d\eta - \int \phi_j d\mu \right| < \epsilon \text{ para todo } \phi_j \in F\}.$$

Então a topologia fraca* é definida estipulando que estes conjuntos $V(\mu, F, \epsilon)$ com F e ϵ variável, constituem uma base de vizinhanças da medida μ.

Proposição 1.4. *Uma sequência $(\mu_n)_{n \in \mathbb{N}}$ em $\mathcal{M}(X)$, converge para uma medida $\mu \in \mathcal{M}(X)$ na topologia fraca* se, e somente se,*

$$\int \phi d\mu_n \to \int \phi d\mu \text{ para toda função contínua } \phi : X \to \mathbb{R}.$$

Outra proposição útil que caracteriza a convergência de medidas é dada a seguir:

Proposição 1.5. *Assuma que a sequência μ_n converge para μ na topologia fraca*. Então:*

1. $\limsup_{n\to\infty} \mu_n(K) \leq \mu(K)$ *para cada conjunto compacto* $K \subset X$;

2. $\liminf_{n\to\infty} \mu_n(U) \geq \mu(U)$ *para cada conjunto aberto* $U \subset X$;

Em particular, se o bordo de A tem medida zero, temos que $\lim_{n\to\infty} \mu_n(A) = \mu(A)$

Teorema 1.7. *O Espaço $\mathcal{M}(X)$ munido da topologia fraca* é metrizável e compacto.*

Capítulo 2

Fatores de Espaços de Medida

A intenção deste capítulo é introduzir o conceito de extensão de espaços de medida. Assim, poderemos encontrar algumas propriedades de um espaço comparando-o com outro espaço (em particular, obteremos nossos resultados dos próximos capítulos comparando o espaço em questão com um subespaço próprio). Ao segundo espaço chamaremos fator do primeiro. Apresentamos ainda um operador especial que transforma funções de uma extensão em funções de seu fator, que será uma ferramenta fortíssima no entendimento do último capítulo.

O conteúdo apresentado nas seções deste capítulo, foi extraído de [5] onde Furstenberg traz resultados mais gerais.

2.1 Fatores de espaços de medida

Sejam $(X, \mathcal{B}, \mu)$ e $(Y, \mathcal{D}, \nu)$ probabilidades.

Uma aplicação $\phi : X \to Y$ é uma aplicação preservando medida se:

1. É mensurável;

2. Para cada $A \in \mathcal{D}$, $\mu(\phi^{-1}(A)) = \nu(A)$.

Podemos identificar conjuntos que "diferem" por conjuntos de medida nula. Definimos a σ-álgebra associada $\tilde{\mathcal{B}}$ consistindo da classe de equivalência dos conjuntos em $\mathcal{B}$ onde $A \sim A'$ se, e somente se, $\mu(A \cup A' - A \cap A') = 0$.

Definição 2.1. *Um homomofirsmo $\alpha : (X, \mathcal{B}, \mu) \to (Y, \mathcal{D}, \nu)$ entre dois espaços de medida é dado por uma injeção $\alpha^{-1} : \tilde{\mathcal{D}} \to \tilde{\mathcal{B}}$ satisfazendo:*

(i) $\alpha^{-1}(\tilde{A}_1 \cup \tilde{A}_2) = \alpha^{-1}(\tilde{A}_1) \cup \alpha^{-1}(\tilde{A}_2), \tilde{A}_1,\ \tilde{A}_2 \in \tilde{\mathcal{D}}$.

(ii) $\alpha^{-1}(\tilde{Y} - \tilde{A}) = \tilde{X} - \alpha^{-1}(\tilde{A}),\ \tilde{A} \in \tilde{\mathcal{D}}$.

(iii) $\mu(\alpha^{-1}(\tilde{A})) = \nu(\tilde{A}),\ \tilde{A} \in \tilde{\mathcal{D}}$.

Observamos que as propriedades acima forçam α^{-1} ser injetiva. De fato, $\alpha^{-1}(\tilde{A}) = \alpha^{-1}(\tilde{B}) \Rightarrow \alpha^{-1}(\tilde{A} \cup \tilde{B}) = \alpha^{-1}(\tilde{A} \cap \tilde{B}) \Rightarrow \nu(\tilde{A} \cup \tilde{B}) = \nu(\tilde{A} \cap \tilde{B}) \Rightarrow \tilde{A} = \tilde{B}$.

Além disso α^{-1} preserva operações de σ-álgebras:

$$\tilde{A} = \bigcup_{n=1}^{\infty} A_n,$$

então

$$\nu\left(\tilde{A} - \bigcup_{n=1}^{N} A_n\right) \to 0,$$

onde

$$\mu\left(\alpha^{-1}(\tilde{A}) - \bigcup_{n=1}^{N} \alpha^{-1}(\tilde{A}_n)\right) \to 0.$$

Daí,

$$\alpha^{-1}(\tilde{A}) = \bigcup_{n=1}^{\infty} \alpha^{-1}(\tilde{A}_n).$$

Dado um homomorfismo determinado por uma aplicação espacial $\phi : X \to Y$, nós podemos obter a partir de uma função mensurável em $(Y, \mathcal{D}, \nu)$ uma função mensurável em $(X, \mathcal{B}, \mu)$ por $f \mapsto f \circ \phi$. Esta correspondência aplica $L^p(Y, \mathcal{D}, \nu)$ isometricamente em $L^p(X, \mathcal{B}, \mu)$. Para um homomorfismo geral $\alpha : (X, \mathcal{B}, \mu) \to (Y, \mathcal{D}, \nu)$ podemos associar para cada classe de equivalência de funções mensuráveis em $(Y, \mathcal{D}, \nu)$ uma classe de equivalência de funções mensuráveis em $(X, \mathcal{B}, \mu)$. De fato, dada uma função mensurável f, criamos a família de conjuntos $A_t(f) = \{y : f(y) < t\}$, para t racional. Reciprocamente, dada uma família de conjuntos mensuráveis A_t, $t \in$ racionais, com $t_1 \leq t_2 \Rightarrow A_{t_1} \subset A_{t_2}$, definimos $f(y) = \inf\{t : y \in A_t\}$. Começando com uma função mensurável $f(y)$ em $(Y, \mathcal{D}, \nu)$, formamos $\{A_t(f)\}$ e transformamos estes para $\{\alpha^{-1}(A_t(f))\} \subset \tilde{\mathcal{B}}$. Por tomarmos representantes destes conjuntos em $\mathcal{B}$, construimos uma função f^α em $(X, \mathcal{B}, \mu)$ que satisfaz

$$\alpha^{-1}(f^{-1}(B)) = (f^\alpha)^{-1}(B), \tag{2.1}$$

para cada conjunto de Borel $B \subset \mathbb{R}$. Por (2.1) nós deduzimos que se $f \in L^p(Y, \mathcal{D}, \nu)$, então $f^\alpha \in L^p(X, \mathcal{B}, \mu)$.

Definição 2.2. *Dizemos que o espaço de medida $(Y, \mathcal{D}, \nu)$ é um fator de $(X, \mathcal{B}, \mu)$ se existir um homomorfismo de $(X, \mathcal{B}, \mu)$ para $(Y, \mathcal{D}, \nu)$. Neste caso, dizemos ainda que $(X, \mathcal{B}, \mu)$ é uma extensão de $(Y, \mathcal{D}, \nu)$. Dois espaços $(X, \mathcal{B}, \mu)$ e $(Y, \mathcal{D}, \nu)$ são equivalentes se existir um homomorfismo $\alpha : (X, \mathcal{B}, \mu) \to (Y, \mathcal{D}, \nu)$ com $\alpha^{-1}(\tilde{\mathcal{D}}) = \tilde{\mathcal{B}}$.*

Se $(X, \mathcal{B}, \mu)$ é um espaço de medida, então $\tilde{\mathcal{B}}$, com a métrica $d(\tilde{A}, \tilde{B}) = \mu(\tilde{A} \Delta \tilde{B})$, onde $\tilde{A} \Delta \tilde{B} = \tilde{A} \cup \tilde{B} - \tilde{A} \cap \tilde{B}$, é espaço métrico. Ainda, se nós identificarmos $\tilde{A}$ com $1_A \in L^1(X, \mathcal{B}, \mu)$ a métrica em questão coincide com a métrica de L^1. Temos pela Proposição 1.1 que $\tilde{\mathcal{B}}$ é um espaço métrico completo.

Se $\tilde{\mathcal{B}}_0$ é uma subálgebra de $\tilde{\mathcal{B}}$ pode-se mostrar que $\tilde{\mathcal{B}}_0$ é denso em $\tilde{\mathcal{B}}$ se, e somente se, $\tilde{\mathcal{B}}_0$ gera $\tilde{\mathcal{B}}$ como σ-álgebra.

Proposição 2.1. *Dois espaços de medida $(X, \mathcal{B}, \mu)$ e $(Y, \mathcal{D}, \nu)$ são equivalentes se, e somente se, $\tilde{\mathcal{B}}$ contém uma subálgebra densa $\tilde{\mathcal{B}}_0$ e $\tilde{\mathcal{D}}$ contém uma subálgebra densa $\tilde{\mathcal{D}}_0$ tal que existe um isomorfismo preservando medida entre as subálgebras $\tilde{\mathcal{B}}_0$ e $\tilde{\mathcal{D}}_0$.*

Demonstração: Se os dois espaços são equivalentes podemos tomar $\tilde{\mathcal{B}}_0 = \tilde{\mathcal{B}}$ e $\tilde{\mathcal{D}}_0 = \tilde{\mathcal{D}}$.

Reciprocamente, suponha que $\tilde{\mathcal{B}}_0 \cong \tilde{\mathcal{D}}_0$. Se $\tilde{A} \in \tilde{\mathcal{D}}$, como $\tilde{\mathcal{D}}_0$ é denso em $\tilde{\mathcal{D}}$, tomo $\tilde{A}_n \in \tilde{\mathcal{D}}_0$ com $\mu(\tilde{A}\Delta\tilde{A}_n) \to 0$. Assim $\{\tilde{A}_n\}$ forma uma seqüência de Cauchy e como α, isomorfismo entre $\tilde{\mathcal{B}}_0$ e $\tilde{\mathcal{D}}_0$, preserva medida temos que $\{\alpha^{-1}(\tilde{A}_n)\}$ forma uma seqüência de Cauchy e daí podemos fazer $\alpha^{-1}(\tilde{A}) = \lim \alpha^{-1}(\tilde{A}_n)$. Notamos que α^{-1} cumpre as condições de homomorfismo e como $\tilde{\mathcal{B}}_0$ é denso em $\tilde{\mathcal{B}}$ temos que α^{-1} é sobrejetivo. ■

Abordaremos a partir de agora, sistemas preservando medida $(X, \mathcal{B}, \mu, T)$ exigindo que a transformação T seja invertível. Assim, em um sistema preservando medida $(X, \mathcal{B}, \mu, T)$, a transformação T define um automorfismo de $\mathcal{B}$ com $A \mapsto T^{-1}A$, e isto induz um automorfismo $T^{-1} : \tilde{\mathcal{B}} \to \tilde{\mathcal{B}}$. Nós usamos isto na definição de homomorfismo de sistemas preservando medida.

Definição 2.3. *Um homomorfismo de sistemas preservando medida* $\alpha : (X, \mathcal{B}, \mu, T) \to (Y, \mathcal{D}, \nu, S)$ *é dado por um homomorfismo* $\alpha : (X, \mathcal{B}, \mu) \to (Y, \mathcal{D}, \nu)$ *satisfazendo as condições de (i) à (iii) da Definição 2.1 e ainda:*

(iv) $\alpha^{-1}(S^{-1}\tilde{A}) = T^{-1}\alpha^{-1}(\tilde{A})$, $\tilde{A} \in \tilde{\mathcal{D}}$.

Neste caso dizemos que $(Y, \mathcal{D}, \nu, S)$ *é um fator de* $(X, \mathcal{B}, \mu, T)$ *e que* $(X, \mathcal{B}, \mu, T)$ *é uma extensão de* $(Y, \mathcal{D}, \nu, S)$. *Dois sistemas preservando medida são equivalentes se* $\alpha^{-1}(\tilde{\mathcal{D}}) = \tilde{\mathcal{B}}$.

Por analogia à Proposição 2.1 nós temos o seguinte critério para sistemas preservando medida.

Proposição 2.2. *Dois s.p.m.* $(X, \mathcal{B}, \mu, T)$ *e* $(Y, \mathcal{D}, \nu, S)$ *são equivalentes se, e somente se,* $\tilde{\mathcal{B}}$ *contém uma subálgebra densa* T*-invariante* $\tilde{\mathcal{B}}_0$ *e* $\tilde{\mathcal{D}}$ *contém uma subálgebra densa* S*-invariante* $\tilde{\mathcal{D}}_0$ *tal que existe um isomorfismo preservando medida entre as subálgebras* $\tilde{\mathcal{B}}_0$ *e* $\tilde{\mathcal{D}}_0$ *que é compatível com a ação das transformações.*

Demonstração: Como na Proposição 2.1, observamos que $\alpha^{-1}(S^{-1}\tilde{A}) = \lim \alpha^{-1}(S^{-1}\tilde{A}_n) = T^{-1} \lim \alpha^{-1}(\tilde{A}_n) = T^{-1}\alpha^{-1}(\tilde{A})$. ■

A partir de agora, para uma melhor leitura, muitas vezes denotaremos tanto o espaço de medida $(X, \mathcal{B}, \mu)$ como o s.p.m. $(X, \mathcal{B}, \mu, T)$ por $\mathbf{X}$. O espaço $L^p(X, \mathcal{B}, \mu)$ será denotado por $L^p(\mathbf{X})$. Homomorfismos de sistemas serão escritos $\alpha : \mathbf{X} \to \mathbf{Y}$. Mesmo ϕ sendo uma aplicação espacial, $\phi : X \to Y$, poderemos escrever $\phi : \mathbf{X} \to \mathbf{Y}$. Quando α é um homomorfismo de sistemas $\alpha : \mathbf{X} \to \mathbf{Y}$, vimos que podemos levar funções mensuráveis de Y em funções mensuráveis de X, por $f \mapsto f^\alpha$. Estenderemos a notação para aplicações espaciais $f \mapsto f^\phi$.

Se T é a transformação de um s.p.m. $(X, \mathcal{B}, \mu, T)$ denotaremos $f \circ T$ por Tf.

Pelo Teorema de Mudança de Variáveis dado em [1], se ϕ é uma aplicação mensurável em $(X, \mathcal{B})$, então $\phi\mu$ denotará a medida

$$\phi\mu(A) = \mu(\phi^{-1}(A)), \tag{2.2}$$

e podemos verificar que $f \mapsto f^\phi$ aplica $L^1(Y, \mathcal{D}, \phi\mu)$ em $L^1(X, \mathcal{B}, \mu)$ e

$$\int f \circ \phi d\mu = \int f d\phi\mu. \tag{2.3}$$

2.2 Espaços regulares de medida

Definição 2.4. *Um espaço de medida $(X, \mathcal{B}, \mu)$ é regular se X é um espaço métrico compacto e $\mathcal{B}$ consiste da σ-álgebra de Borel em X. Um s.p.m. $(X, \mathcal{B}, \mu, T)$ é regular se o espaço de medida subordinado é regular.*

Definição 2.5. *Um espaço de medida $(X, \mathcal{B}, \mu)$ é separável se a σ-álgebra associada $\tilde{\mathcal{B}}$ é gerada por um subconjunto enumerável. O s.p.m. $(X, \mathcal{B}, \mu, T)$ é separável se $(X, \mathcal{B}, \mu)$ é separável.*

Uma condição equivalente para a separabilidade do espaço de medida $(X, \mathcal{B}, \mu)$ é a separabilidade do espaço métrico $\tilde{\mathcal{B}}$.

Como um espaço métrico compacto possui uma base enumerável de conjuntos abertos e estes geram a σ-álgebra de Borel, é claro que um espaço de medida regular é separável.

Definição 2.6. *Um homomorfismo $\alpha : (X, \mathcal{B}, \mu, T) \to (Y, \mathcal{D}, \nu, S)$ é regular se $(X, \mathcal{B}, \mu, T)$ é s.p.m. regular e α é uma aplicação espacial, $\alpha : X \to Y$.*

Teorema 2.1. *Seja $\alpha : \boldsymbol{X} \to \boldsymbol{Y}$ um homomorfismo de s.p.m. separáveis. Existe um sistema $\boldsymbol{X'}$ equivalente a $\boldsymbol{X}$ e um sistema $\boldsymbol{Y'}$ equivalente a $\boldsymbol{Y}$ tal que o homomorfismo correspondente $\alpha : \boldsymbol{X'} \to \boldsymbol{Y'}$ é regular.*

Demonstração: Escreveremos $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ e $\mathbf{Y} = (Y, \mathcal{D}, \nu, S)$. Contruiremos os espaços $\mathbf{X}'$ e $\mathbf{Y}'$. Seja $\mathcal{D}_0$ uma álgebra enumerável de conjuntos em $\mathcal{D}$, com $\mathcal{D}_0$, S-invariante e denso sobre $\tilde{\mathcal{D}}_0$. Escreva $\mathcal{D}_0 = \{A_n\}$. Seja $Y' = \{0,1\}^{\mathbb{N}}$ e $\eta_m : Y' \to \{0,1\}$ a função que assume o valor da m-ésima coordenada de $y' \in Y'$. Seja

$$A'_m = \{y' \in Y' : \eta_m(y') = 1\}.$$

Afirmamos que existe uma medida ν' em Y' com

$$\nu'(A'_{i_1} \cap A'_{i_2} \cap \ldots \cap A'_{i_k}) = \nu(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k})$$

para toda k-upla $i_1 \leq i_2 \leq \ldots \leq i_k$.

Para provarmos isto, primeiro encontramos uma medida ν'_N que satisfaça esta condição para $i_k \leq N$. Como, nestas condições, tratamos de uma álgebra finita de conjuntos é fácil ver que ν'_N existe. Tomando o limite de alguma subseqüência $\nu' = \lim \nu'_N$, obtemos ν' como queremos.

Agora, seja $\mathcal{D}'$ a σ-álgebra dos conjuntos de Borel em Y'. Assumimos os conjuntos $\{A_n\}$ sendo todos distintos. Temos que S define uma permutação $\sigma : \mathbb{N} \to \mathbb{N}$ com $S^{-1}A_m = A_{\sigma(m)}$. Defina $S' : Y' \to Y'$ por $S'y'(m) = y'(\sigma(m))$. Além disso, $S'^{-1}A'_m = A'_{\sigma(m)}$. Disto segue que S' preserva medida em $(Y', \mathcal{D}', \nu')$. Por construção, o sistema $\mathbf{Y}'$ é regular, então usando a Proposição 2.1, concluimos que $(Y', \mathcal{D}', \nu', S')$ é equivalente a $(Y, \mathcal{D}, \nu, S)$.

Por construção similar, nós obtemos uma álgebra $\mathcal{B}_0 \subset \mathcal{B}$ que é T-invariante e com $\tilde{\mathcal{B}}_0$ denso em $\tilde{\mathcal{B}}$ para obter $(X', \mathcal{B}', \mu', T')$ regular e equivalente a $(X, \mathcal{B}, \mu, T)$. Escreva $\mathcal{B}_0 = \{B_m\}$ e a ação de T em $\mathcal{B}_0$ induz permutações $T^{-1}B_m = B_{\tau(m)}$. Assumimos ainda que $\alpha^{-1}(\tilde{\mathcal{D}}_0) \subset \tilde{\mathcal{B}}_0$. Definimos $\beta : \mathbb{N} \to \mathbb{N}$ por $\alpha^{-1}(A_m) = B_{\beta(m)}$ e definimos $\phi : X' \to Y'$ por $\phi(x')(m) = x'(\beta(m))$. Então

$$S'\phi(x')(m) = \phi(x')(\sigma(m)) = x'(\beta \circ \sigma(m)),$$

$$\phi T'(x')(m) = T'x'(\beta(m)) = x'(\tau \circ \beta(m)).$$

Estas expressões são iguais por acordarem com $\alpha^{-1}(S^{-1}A_m) = T^{-1}\alpha^{-1}(A_m)$. Isto completa a prova do Teorema. ∎

Corolário 2.1. *Todo espaço de medida separável é equivalente a um espaço de medida regular. Todo s.p.m. separável é equivalente a um s.p.m. regular.*

Na Análise da estrutura de um dado sistema preservando medida, nós devemos considerar junto com o dado sistema, todos os seus fatores. Temos inicialmente que fatores de um sistema separável é separável. De acordo com o Corolário acima, qualquer tal fator é equivalente a um sistema regular que será ele mesmo um fator do sistema dado. Deste modo, ao estudar a estrutura de um s.p.m. separável, nós justificamos nossa atenção aos fatores regulares.

2.3 Esperança condicional

Seja $\alpha : (X, \mathcal{B}, \mu) \to (Y, \mathcal{D}, \nu)$ um homomorfismo. A aplicação $f \mapsto f^\alpha$ identifica $L^2(Y, \mathcal{D}, \nu)$ com um subespaço fechado $L^2(Y, \mathcal{D}, \nu)^\alpha \subset L^2(X, \mathcal{B}, \mu)$. Se P é a projeção ortogonal de $L^2(X, \mathcal{B}, \mu)$ em $L^2(Y, \mathcal{D}, \nu)^\alpha$, então definimos $E(f|\mathbf{Y})$ para $f \in L^2(X, \mathcal{B}, \mu)$ por

$$E(f|\mathbf{Y}) \in L^2(Y, \mathcal{D}, \nu), E(f|\mathbf{Y})^\alpha = Pf. \tag{2.4}$$

Note que na notação $E(f|\mathbf{Y})$ nós suprimimos o homomorfismo $\alpha : \mathbf{X} \to \mathbf{Y}$ que esta implicito na definição.

Proposição 2.3. *O operador esperança condicional $f \mapsto E(f|\boldsymbol{Y})$ definido para $f \in L^2(\boldsymbol{X})$ por (2.4) possui as seguintes propriedades:*

(i) *$f \mapsto E(f|\boldsymbol{Y})$ é um operador linear de $L^2(\boldsymbol{X})$ em $L^2(\boldsymbol{Y})$.*

(ii) *Se $f \geq 0$, $E(f|\boldsymbol{Y}) \geq 0$.*

(iii) *Se $f \in L^2(\boldsymbol{Y})$, $E(f^\alpha|\boldsymbol{Y}) = f$. Em particular $E(1|\boldsymbol{Y}) = 1$.*

(iv) *Se $g \in L^\infty(\boldsymbol{Y})$, $E(g^\alpha f|\boldsymbol{Y}) = gE(f|\boldsymbol{Y})$.*

(v) *Em particular, $\int f \, d\mu = \int E(f|\boldsymbol{Y}) \, d\nu$.*

Demonstração: A propriedade *(i)* vem direto da definição.

(...)

Para provar a propriedade *(iii)*, lembramos que se $f \in L^2(\mathbf{Y})$ então, $f^\alpha \in L^2(Y, \mathcal{D}, \nu)^\alpha$. Por (2.4) temos então, que $E(f^\alpha|\mathbf{Y})^\alpha = Pf^\alpha = f^\alpha$. Pela injetividade de α^{-1} temos $E(f^\alpha|\mathbf{Y}) = f$.

Para provar as propriedades *(iv)* e *(v)* usamos o fato de que em um espaço de Hilbert a projeção ortogonal preserva o produto interno. Assim

$$\int fh^\alpha \, d\mu = \int Pfh^\alpha \, d\mu \quad \text{para todo} \quad h \in L^2(\mathbf{Y}).$$

Transferindo para $E(f|\mathbf{Y})$ por (2.4), temos

$$\int fh^\alpha \, d\mu = \int E(f|\mathbf{Y})^\alpha h^\alpha \, d\mu = \int E(f|\mathbf{Y})h \, d\nu \quad \text{para todo} \quad h \in L^2(\mathbf{Y}). \tag{2.5}$$

Tendo que (2.5) vale para todo $h \in L^2(\mathbf{Y})$, substituimos h por gh com $g \in L^\infty(\mathbf{Y})$. Então

$$\int f(gh)^\alpha \, d\mu = \int E(f|\mathbf{Y})gh \, d\nu$$

ou

$$\int (fg^\alpha)\, h^\alpha \, d\mu = \int (E(f|\mathbf{Y})g)\, h \, d\nu, \quad \text{para todo } h \in L^2(\mathbf{Y}).$$

Substituindo f por fg^α em (2.5), temos

$$\int (fg^\alpha)\, h^\alpha \, d\mu = \int E(fg^\alpha|\mathbf{Y})h \, d\nu.$$

Isto prova a propriedade *(iv)* e ao mesmo tempo a *(v)*. ∎

Da demonstração acima temos o seguinte corolário.

Corolário 2.2. *Se $f \in L^2(\boldsymbol{X})$, a esperança condicional $E(f|\,\boldsymbol{Y})$ é caracterizada por*

$$\int E(f|\,\boldsymbol{Y})h \, d\nu = \int fh^\alpha \, d\mu$$

para todo $h \in L^2(\,\boldsymbol{Y})$.

Proposição 2.4. *Seja $\psi(t)$ uma função convexa não negativa de uma variável real. Se f e $\psi \circ f$ estão em $L^2(\boldsymbol{X})$, então*

$$\psi\,(E(f|\,\boldsymbol{Y})) \;\leq\; E(\psi \circ f|\,\boldsymbol{Y}). \tag{2.6}$$

Demonstração: A função convexa ψ pode ser tomada como sendo o supremo de enumeráveis funções lineares: $\psi(x) = \sup L_n(x), \quad L_n(x) = a_n x + b_n.$ ·

Em virtude de *(i)* e *(iii)* da Proposição 2.3, $E(L_n(f)|\mathbf{Y}) = L_n(E(f|\mathbf{Y}))$.

Daí, $\psi(E(f|\mathbf{Y})) = \sup L_n(E(f|\mathbf{Y})) = \sup E(L_n(f)|\mathbf{Y})$.

De acordo com *(ii)*, $f_1 \leq f_2 \Rightarrow E(f_1|\mathbf{Y}) \leq E(f_2|\mathbf{Y})$ e como $L_n(f) \leq \psi \circ f$ temos que $E(L_n(f)|\mathbf{Y}) \leq E(\psi \circ f|\mathbf{Y})$. ■

O próximo teorema estenderá a definição de esperança condicional a uma aplicação do espaço de funções $L^1(\mathbf{X})$ em $L^1(\mathbf{Y})$.

Teorema 2.2. *A aplicação esperança condicional, $f \mapsto E(f|\boldsymbol{Y})$, estende-se a uma aplicação de $L^1(\boldsymbol{X})$ em $L^1(\boldsymbol{Y})$ satisfazendo (i)-(v) da Proposição 2.3, e que aplica cada $L^p(\boldsymbol{X})$ em $L^p(\boldsymbol{Y})$, $1 \leq p \leq \infty$, com $\|E(f|\boldsymbol{Y})\|_p \leq \|f\|_p$.*

Demonstração: Tomando $\psi(t) = |t|$ na proposição anterior em conjunto com a propriedade *(v)* na Proposição 2.3, temos que a aplicação esperança condicional é uniformemente continua na norma L^1. De fato,

$$\int |E(f|\mathbf{Y}) - E(g|\mathbf{Y})|\, d\nu = \int |E(f-g|\mathbf{Y})|\, d\nu \leq \int E(|f-g|\,|\mathbf{Y}) d\nu = \int |f-g|\, d\mu.$$

Disto e do fato de $L^2(\mathbf{X})$ ser denso em $L^1(\mathbf{X})$, podemos estender a esperança condicional para $L^1(\mathbf{X})$. Assim as propriedades *(i)* a *(v)* valem para esta extensão.

Agora estendemos a Proposição 2.4, substituindo a condição $f, \phi \circ f \in L^2$ por $f, \phi \circ f \in L^1$.

Por fim, suponha que $f \in L^p(\mathbf{X})$. Aplique a Proposição 2.4 com $\psi(t) = |t|^p$. Conluimos que $E(f|\mathbf{Y}) \in L^p(\mathbf{Y})$ e que a aplicação $f \mapsto E(f|\mathbf{Y})$ é uma contração para cada $p < \infty$. Para $p = \infty$ basta usarmos *(ii)* e *(iii)*. ■

Terminamos esta seção com um resutado pertinente aos sistemas preservando medida.

Proposição 2.5. *Seja $(Y, \mathcal{D}, \nu, S)$ um fator de $(X, \mathcal{B}, \mu, T)$. Então para cada $f \in L^1(X, \mathcal{B}, \mu)$, $E(Tf|\boldsymbol{Y}) = SE(f|\boldsymbol{Y})$.*

Demonstração: É suficiente mostrarmos para $f \in L^2(\mathbf{X})$. Usando *(v)* da Proposição 2.3 e o fato de T ser um operador unitário em $L^2(\mathbf{X})$, temos:

$$\int E(Tf|\mathbf{Y})h \, d\nu = \int (Tf)h^\alpha \, d\mu = \int f(T^{-1}h^\alpha) \, d\mu.$$

Retornando a (2.1), vemos que $g \in L^2(\mathbf{Y})$,

$$\begin{aligned} \alpha^{-1}\{y : Sg(y) \in B\} &= \alpha^{-1}S^{-1}\{y : g(y) \in B\} = T^{-1}\alpha^{-1}\{y : g(y) \in B\} = T^{-1}\{x : g^\alpha(x) \in B\} \\ &= \{x : Tg^\alpha(x) \in B\}. \end{aligned}$$

Assim $(Sg)^\alpha = T(g^\alpha)$.

Consequentemente,

$$\begin{aligned} \int E(Tf|\mathbf{Y})h \, d\nu &= \int f(T^{-1}h^\alpha) \, d\mu = \int f(S^{-1}h)^\alpha \, d\mu = \int E(f|\mathbf{Y})S^{-1}h \, d\nu \\ &= \int SE(f|\mathbf{Y})h \, d\nu. \end{aligned}$$

Usando o corolário da Proposição 2.3, $E(Tf|\mathbf{Y}) = SE(f|\mathbf{Y})$. ■

2.4 Desintegração de medida

Seja $(X, \mathcal{B}, \mu)$ um espaço regular de medida, e seja $\alpha : (X, \mathcal{B}, \mu) \to (Y, \mathcal{D}, \nu)$ um homomorfismo entre espaços de medida. Aqui X é um espaço métrico compacto, e lembremos que $\mathcal{M}(X)$ denota o espaço métrico compacto de probabilidades mensuráveis em X.

Teorema 2.3 (Desintegração de medida)**.** *Existe uma aplicação mensurável de Y em $\mathcal{M}(X)$ que devemos denotar $y \mapsto \mu_y$ que satisfaz:*

(i) *Para todo $f \in L^1(X, \mathcal{B}, \mu), f \in L^1(X, \mathcal{B}, \mu_y)$ para quase todo $y \in Y$, e ainda, $E(f|\,\boldsymbol{Y}) = \int f d\mu_y$ para quase todo $y \in Y$.*

(ii) *$\int \left\{\int f d\mu_y\right\} d\nu(y) = \int f d\mu$ para todo $f \in L^1(X, \mathcal{B}, \mu)$.*

A aplicação $y \mapsto \mu_y$ é caracterizado pela condição *(i)*. Devemos escrever $\mu = \int \mu_y d\nu$ e nos referimos a isto como a *desintegração de* μ *com respeito ao fator* $(Y, \mathcal{D}, \nu)$.

Demosntração: Escolha um conjunto denso mensurável de funções em $\mathscr{C}(X)$, o espaço das funções contínuas de valor real em X, assuma que 1 esteja neste conjunto, e seja $\mathcal{A}$ um conjunto formado por todas as combinações lineares finitas destes com coeficientes racionais. O conjunto $\mathcal{A}$ é um subconjunto enumerável de $\mathscr{C}(X)$.

Seja $L(f)$ um funcional linar em $\mathcal{A}$ satisfazendo : *(i)* $L(af_1+bf_2) = aL(f_1)+bL(f_2)$ para $f_1, f_2 \in \mathcal{A}$ e a, b racionais; *(ii)* $|L(f)| \leq \max|f(x)|$; *(iii)* $f \geq 0 \Rightarrow L(f) \geq 0$ *(iv)* $L(1) = 1$. Podemos, então, estender L a um funcional linear em $\mathscr{C}(X)$ satisfazendo *(iii)*. De fato, pela propriedade *(ii)* temos que L é uniformemente contínua no subconjunto denso $\mathcal{A}$ de $\mathscr{C}(X)$. Estendendo L por continuidade, obtemos linearidade para coeficientes racionais diretamente da propriedade *(i)* e subsequentemente para todos os coeficientes reais.

Se $f \geq 0$ e $f_n \in \mathcal{A}$ com

$$\left| f_n - \left(f + \frac{1}{n} \right) \right| < \frac{1}{n},$$

então $f_n \geq 0$ e $f_n \to f$ uniformemente. Como $L(f_n) \geq 0$ teremos que $L(f) \geq 0$. Agora a cada funcional linear em $\mathscr{C}(X)$ satisfazendo $L(1) = 1$ e $f \geq 0 \Rightarrow L(f) \geq 0$ corresponde a uma medida em $\mathcal{M}(X)$. Concluímos que uma medida em $\mathcal{M}(X)$ é determinada por um funcional linear em $\mathcal{A}$ satisfazendo as propriedades de *(i)* a *(iv)*.

Tomemos para cada $f \in \mathcal{A}$, $E(f|\mathbf{Y})$. Se fixarmos $L_y(f) = E(f|\mathbf{Y})(y)$, então vemos que quase todo L_y satisfaz as propriedades *(i)* a *(iv)*. Além disso , $\{y : L_y(f) > 0\}$ é um conjunto mensurável para $(Y, \mathcal{D}, \nu)$, então a aplicação $y \mapsto Ly$ é mensurável . Indentificando L_y com uma medida μ_y por meio de $L_y(f) = \int f d\mu_y$, definimos a aplicação $y \mapsto \mu_y$ e vemos que a propriedade *(i)* do teorema é automaticamente satisfeita por $f \in \mathcal{A}$. Usando convergência uniforme vemos que a propriedade é válida para $f \in \mathscr{C}(X)$.

Para $f \in L^1(X, \mathcal{B}, \mu)$ temos que quase sempre $f = \sum g_n$, $g_n \in \mathscr{C}(X)$, onde $\sum \|g_n\|_1 < \infty$. Então $\sum \|E(|g_n||\mathbf{Y})\|_1 < \infty$ e ainda $\sum E(|g_n||\mathbf{Y}) < \infty$ em quase todo ponto. Em algum ponto y que esta série convirja, $\sum |g_n| \in L^1(X, \mathcal{B}, \mu_y)$, e assim

$$\int (\sum g_n) d\mu_y = \sum \int g_n d\mu_y.$$

Temos ainda que para quase todo ponto, $\int g_n d\mu_y = E(g_n|\mathbf{Y})(y)$, em que

$$\int (\sum g_n) d\mu_y = \sum E(g_n|\mathbf{Y}(y), \text{ quase sempre.}$$

Como $f \to E(f|\mathbf{Y})$ é uma contração em L^1, temos

$$\sum E(g_n|Y) \to E(\sum g_n|\mathbf{Y}),$$

na norma L^1.

Então necessáriamente

$$\int (\sum g_n) d\mu_y = E(\sum g_n|\mathbf{Y}), \text{ quase sempre.}$$

Para completar a prova, precisamos mostrar qua a função nula $f - \sum g_n$ permanece nula para quase todo μ_y. Seja $\mu(A) = 0$ e seja A_n aberto com $A_n \supset A, \mu(A_n) \to 0$. Seja h_n uma função contínua com $h_n = 0$ fora de A_n e $0 < h_n \leq 1$ em A. Para cada $\epsilon > 0$, a potência h_n^ϵ é contínua e

$$\int \left\{ \int h_n^\epsilon d\mu_y \right\} d\nu(y) = \int h_n^\epsilon d\mu \leq \mu(A_n).$$

Fazendo $\epsilon \to 0$, encontramos

$$\int \mu_y(A_n) d\nu(y) \leq \mu(A_n) \to 0.$$

Isto completa a prova da propriedade *(i)*. Por outro lado sempre que a propriedade *(i)* é válida para uma função f, então a propriedade *(ii)* também é válida por virtude da propriedade *(v)* da Proposição 2.3. Para completar a prova do teorema, lembremos que a medida em $(X, \mathcal{B})$ é determinada pela integral de funções contínuas com respeito a este, assim a propriedade *(i)* determina claramente μ_y em quase todo ponto. ■

Finalmente consideremos a desintegração de medida para sistemas preservando medida.

Proposição 2.6. *Seja $(Y, \mathcal{D}, \nu, S)$ um fator de $(X, \mathcal{B}, \mu, T)$, e seja $\mu = \int \mu_y d\nu(y)$. Para quase todo $y \in Y$,*

$$\mu_{Sy} = T\mu_y \tag{2.7}$$

Demonstração: Pela propriedade *(i)* do Teorema 2.3 e a Proposição 2.5 temos

$$\int Tf d\mu_y = \int f d\mu_{Sy}.$$

Substitua f por $T^{-1}f$:

$$\int f d\mu_y = \int T^{-1} f d\mu_{Sy} = \int f d(T^{-1}\mu_{Sy})$$

em virtude de (2.3). Assim $\mu_y = T^{-1}\mu_{Sy}$ em quase todo ponto. ■

2.5 Produto relativo de espaços de medida

Nesta seção nós definiremos o produto de dois espaços de medida relativos a um fator comum. Isto terá o mesmo papel na teoria de medida que o produto fibrado tem na teoria de conjuntos.

Seja $(X_1, \mathcal{B}_1, \mu_1)$ e $(X_2, \mathcal{B}_2, \mu_2)$ dois espaços regulares de medida que são extensões de um mesmo espaço $(Y, \mathcal{D}, \nu)$:

$$\alpha_1 : (X_1, \mathcal{B}_1, \mu_1) \to (Y, \mathcal{D}, \nu), \ \ \alpha_2 : (X_2, \mathcal{B}_2, \mu_2) \to (Y, \mathcal{D}, \nu).$$

Definição 2.7. *Denote por $\mu_1 \times_{\boldsymbol{Y}} \mu_2$ a medida em $X_1 \times X_2$ definida por*

$$\mu_1 \times_{\boldsymbol{Y}} \mu_2(A) = \int \mu_{1,y} \times \mu_{2,y}(A) d\nu(y) \tag{2.8}$$

para $A \in \mathcal{B}_1 \times \mathcal{B}_2$, onde $\mu_i = \int \mu_{i,y} d\nu(y)$ é a desintegração de μ_i com respeito ao fator $(Y, \mathcal{D}, \nu)$. O espaço de medida $(X_1, \times X_2, \mathcal{B}_1 \times \mathcal{B}_2, \mu_1 \times_{\boldsymbol{Y}} \mu_2)$ é chamado o produto de $(X_1, \mathcal{B}_1, \mu_1)$ e $(X_2, \mathcal{B}_2, \mu_2)$ relativo a $(Y, \mathcal{D}, \nu)$ e é denotado $\boldsymbol{X}_1 \times_{\boldsymbol{Y}} \boldsymbol{X}_2$.

Para justificar a mensurabilidade do integrando em (2.8), nós notamos que sempre que $\mu = \int \mu_y d\nu(y)$ é a desintegração de μ com respeito a $(Y, \mathcal{D}, \nu)$, sendo μ a medida em um espaço métrico compacto X, então $\int f d\mu_y$ é mensurável para cada função contínua $f \in \mathscr{C}(X)$, pela definição da mensurabilidade de μ_y. Como o conjunto de f para o qual $\int f d\mu_y$ é mensurável é fechado para limites monótonos, o mesmo vale para f mensurável sobre Borel. Concluímos agora que $\mu_{1,y} \times \mu_{2,y}(A)$ é mensurável para $A = A_1 \times A_2$, $A_i \in \mathcal{B}_i$, $i = 1, 2$. Então por passagem de limite, o resultado é válido para todo $A \in \mathcal{B}_1 \times \mathcal{B}_2$. Uma vez que a expressão (2.8) é vista com significado, é fácil verificar que se definimos uma probabilidade σ-aditiva em $(X_1 \times X_2, \mathcal{B}_1 \times \mathcal{B}_2)$.

Se f_1 é uma função no espaço de medida X_1 e f_2 é uma função no espaço de medida X_2, denotamos $f_1 \otimes f_2$ a função

$$f_1 \otimes f_2(x_1, x_2) = f_1(x_1) f_2(x_2).$$

Temos o seguinte resultado.

Proposição 2.7. *A medida $\mu_1 \times_{\mathbf{Y}} \mu_2$ é caracterizada pela igualdade*

$$\int f_1 \otimes f_2 d\mu_1 \times_{\mathbf{Y}} \mu_2 = \int E(f_1 | \boldsymbol{Y}) E(f_2 | \boldsymbol{Y}) d\nu \tag{2.9}$$

satisfeita sempre que $f_1 \in L^2(\boldsymbol{X}_1), f_2 \in L^2(\boldsymbol{X}_2)$.

Demonstração: Temos

$$\int f_1(x_1) f_2(x_2) d\mu_1 \times_{\mathbf{Y}} \mu_2(x_1, x_2)$$

$$\begin{aligned}
&= \int \{ \int f_1(x_1) f_2(x_2) d\mu_{1,y} \times \mu_{2,y}(x_1, x_2) \} d\nu(y) \\
&= \int \left\{ \int f_1(x_1) d\mu_{1,y}(x_1) \right\} \left\{ \int f_2(x_2) d\mu_{2,y}(x_2) \right\} d\nu(y) \\
&= \int E(f_1 | \mathbf{Y})(y) E(f_2 | \mathbf{Y})(y) d\nu(y). \quad \blacksquare
\end{aligned}$$

Sendo, respectivamente, π_1, π_2 as projeções de $X_1 \times X_2$ em suas componentes X_1 e X_2, podemos verificar que são aplicações preservando medida de

$$(X_1 \times X_2, \mathcal{B}_1 \times \mathcal{B}_2, \mu_1 \times_{\mathbf{Y}} \mu_2) \to (X_i, \mathcal{B}_i, \mu_i).$$

Por exemplo,

$$
\begin{aligned}
\mu_1 \times_{\mathbf{Y}} \mu_2(\pi_1^{-1}(A_1)) &= \mu_1 \times_{\mathbf{Y}} \mu_2(A_1 \times X_2) \\
&= \int \mu_{1,y} \times \mu_{2,y}(A_1 \times X_2) d\nu(y) \\
&= \int \mu_{1,y}(A_1) d\nu(y) = \mu_1(A_1).
\end{aligned}
$$

Deste modo $(X_1 \times X_2, \mathcal{B}_1 \times \mathcal{B}_2, \mu_1 \times_{\mathbf{Y}} \mu_2)$ é uma extensão de ambos $(X_i, \mathcal{B}_i, \mu_i)$ e estes são ambos extensões de $(Y, \mathcal{D}, \nu)$. Podemos dizer então que $(X_1 \times X_2, \mathcal{B}_1 \times \mathcal{B}_2, \mu_1 \times_{\mathbf{Y}} \mu_2)$ é uma extensão de $(Y, \mathcal{D}, \nu)$, e podemos desintegrar $\mu_1 \times_{\mathbf{Y}} \mu_2$ com respeito a $(Y, \mathcal{D}, \nu)$.

Proposição 2.8. *Para $f_1 \in L^2(\boldsymbol{X}_1), f_2 \in L^2(\boldsymbol{X}_2)$,*

$$
E(f_1 \otimes f_2 | \boldsymbol{Y}) = E(f_1 | \boldsymbol{Y}) E(f_2 | \boldsymbol{Y}), \tag{2.10}
$$

e a desintegração de $\mu_1 \times_{\mathbf{Y}} \mu_2$ com respeito a $(Y, \mathcal{D}, \nu)$ é

$$
(\mu_1 \times_{\mathbf{Y}} \mu_2)_y = \mu_{1,y} \times \mu_{2,y}.
$$

Demonstração: É fácil ver da definição de esperança condicional que $E(f|\mathbf{Y}) = E(E(f|\mathbf{X}_1)|\mathbf{Y}) = E(E(f|\mathbf{X}_2)|\mathbf{Y})$. Fazemos $E(f_1 \otimes f_2|\mathbf{X}_1) = f_1 E(f_2|\mathbf{Y})^{\alpha_1}$, e para provar isto usaremos o corolário da Proposição 2.3. Precisamos mostrar que

$$
\int f_1(x_1) f_2(x_2) h(x_1) d\mu_1 \times_{\mathbf{Y}} \mu_2 = \int f_1(x_1) E(f_2|\mathbf{Y})^{\alpha_1} h(x_1) d\mu_1(x_1).
$$

Aplicando a propriedade *(v)* da Proposição 2.3 para o homomorfismo α_1, o lado direito torna-se

$$
\int E(f_1 h|\mathbf{Y}) E(f_2|\mathbf{Y}) d\nu = \int f_1(x_1) h(x_1) f_2(x_2) d\mu_1 \times_{\mathbf{Y}} \mu_2(x_1, x_2),
$$

de acordo com a Proposição 2.7. Isto estabelece nossa afirmação. Agora usando a propriedade *(iv)* da Proposição 2.3, obtemos

$$
\begin{aligned}
E(f_1 \otimes f_2|\mathbf{Y}) &= E(E(f_1 \otimes f_2|\mathbf{X}_1)|\mathbf{Y}) = E(f_1 E(f_2|\mathbf{Y})^{\alpha_1}|\mathbf{Y}) \\
&= E(f_2|\mathbf{Y}) E(f_1|\mathbf{Y}) = \int f_1 \otimes f_2 d\mu_{1,y} \times \mu_{2,y}.
\end{aligned}
$$

Isto prova ambas as afirmações da proposição. ■

A próxima proposição mostra que as propriedades definidas nesta seção podem ser definidas se o sistema em questão não é regular.

Proposição 2.9. *Se $\boldsymbol{X}_1, \boldsymbol{X}_2$ são extensões de $\boldsymbol{Y}$ e $\boldsymbol{X}_1'$ é equivalente a $\boldsymbol{X}_1$ e $\boldsymbol{X}_2'$ é equivalente a $\boldsymbol{X}_2$, então $\boldsymbol{X}_1'$ e $\boldsymbol{X}_2'$ são extensões de $\boldsymbol{Y}$ e $\boldsymbol{X}_1' \times_{\boldsymbol{Y}} \boldsymbol{X}_2'$ é equivalente a $\boldsymbol{X}_1 \times_{\boldsymbol{Y}} \boldsymbol{X}_2$.*

Demonstração: Devemos usar os critério da Proposição 2.1 para equivalência de espaços de medida. Estabelecemos um isomorfismo entre as álgebras geradas por conjuntos produtos em $\mathbf{X}_1 \times \mathbf{X}_2$ e em $\mathbf{X}_1' \times \mathbf{X}_2'$. Temos por hipótese uma correspondência entre $L^2(\mathbf{X}_i)$ e $L^2(\mathbf{X}_i')$ para $i = 1, 2$, então é suficiente mostrarmos que se $f_1' \in L^2(X_1', \mathcal{B}_1', \mu_1')$ corresponde a $f_1 \in L^2(X_1, \mathcal{B}_1, \mu_1)$ e $f_2' \in L^2(X_2', \mathcal{B}_2', \mu_2')$ corresponde a $f_2 \in L^2(X_2, \mathcal{B}_2, \mu_2)$, então

$$\int f_1' \otimes f_2' d\mu_1' \times_{\mathbf{Y}} \mu_2' = \int f_1 \otimes f_2 d\mu_1 \times_{\mathbf{Y}} \mu_2.$$

Mas isto segue da Proposição 2.7 e do fato de que $E(f_i'|\mathbf{Y}) = E(f_i|\mathbf{Y})$. ■

Passando a tratar de transformações que preservam medidas, temos um importante resultado.

Proposição 2.10. *Suponha que $(X_1, \mathcal{B}_i, \mu_i, T_i)$, $i = 1, 2$, são extensões do sistema preservando medida $(Y, \mathcal{D}, \nu, S)$. Então a medida $\mu_1 \times_{\mathbf{Y}} \mu_2$ é T-invariante, onde a ação de $T = T_1 \times T_2$ em $X_1 \times X_2$ é dada por $T_1 \times T_2(x_1, x_2) = (Tx_1, Tx_2)$. Assim $(X_1 \times X_2, \mathcal{B}_1 \times \mathcal{B}_2, \mu_1 \times_{\mathbf{Y}} \mu_2, T_1 \times T_2)$ é um s.p.m. que é o produto relativo $(X_1, \mathcal{B}_1, \mu_1, T_1) \times_{\mathbf{Y}} (X_2, \mathcal{B}_2, \mu_2, T_2)$.*

Demonstração. Temos que verificar que

$$\mu_1 \times_{\mathbf{Y}} \mu_2((T_1 \times T_2)^{-1} A) = \mu_1 \times_{\mathbf{Y}} \mu_2(A)$$

para $A \in \mathcal{B}_1 \times \mathcal{B}_2$. Mas de acordo com a equação (2.8),

$$\begin{aligned} \mu_1 \times_{\mathbf{Y}} \mu_2(T^{-1}A) &= \int \mu_{1,y} \times \mu_{2,y}(T^{-1}A) d\nu(y) \\ &= \int T_1\mu_{1,y} \times T_2\mu_{2,y}(A) d\nu(y). \end{aligned}$$

A última identidade é estabelecida para todo $A \in \mathcal{B}_1 \times \mathcal{B}_2$, verificando $A = A_1 \times A_2, A_i \in \mathcal{B}_i$. Agora, usando a Proposição 2.6,

$$\begin{aligned}
\mu_1 \times_{\mathbf{Y}} \mu_2(T^{-1}A) &= \int \mu_{1,Sy} \times \mu_{2,Sy}(A) d\nu(y) \\
&= \int \mu_{1,y} \times \mu_{2,y}(A) dS\nu(y),
\end{aligned}$$

Como $S\nu = \nu$ temos a proposição. ∎

Capítulo 3

Teorema de Szemerédi

Em 1936 Erdös e Turán [2] conjecturaram que todo subconjunto de $\mathbb{Z}$ com densidade superior positiva contém progressões aritméticas finitas com comprimento arbitrariamente grande. Esta conjectura foi demosntrada por Szemerédi [14] em 1975.

Teorema 3.1 (Teorema de Szemerédi)**.** *Se S é um subconjunto de $\mathbb{Z}$ com densidade superior positiva, então S contém progressões aritméticas arbitrariamente grande.*

Em 1976, Furstenberg [4] notou que o argumento deste teorema é equivalente a um argumento de "recorrência múltipla" de transformações preservando medida.

Neste capítulo daremos uma prova do Teorema de Szemerédi com argumentos de Teoria Ergódica, usando o Teorema de Recorrência Múltipla de Furstenberg. Ao fim do capítulo veremos que estes teoremas são equivalentes.

3.1 Densidade superior

Chamamos de *intervalo* do conjunto $\mathbb{Z}$ dos números inteiros qualquer subconjunto I da forma $\{n \in \mathbb{Z} : a \leq n < b\}$, para quaisquer $a \leq b$ em $\mathbb{Z}$. O seu *cardinal* é $\#I = b - a$. Do mesmo modo definimos intervalos em $\mathbb{N}$.

Definição 3.1. *A densidade superior $D_S(S)$ de um subconjunto S em $\mathbb{Z}$ é*

$$D_S(S) = \limsup_{\#I \to \infty} \frac{\#(S \cap I)}{\#I}$$

onde I representa qualquer intervalo em $\mathbb{Z}$. De maneira análoga definimos a densidade inferior $D_I(S)$ trocando limite superior por limite inferior. Dizemos que um conjunto possui densidade $D(S)$ se $D_I(S) = D_S(S) = D(S)$

Em outras palavras, $D_S(S)$ é o maior número D tal que existe uma seqüência de intervalos $I_j \subset \mathbb{Z}$ tais que

$$\#I_j \to \infty \text{ e } \frac{\#(S \cap I_j)}{\#I_j} \to D.$$

e $D_I(S)$ é o menor número nessas condições.

Exemplo 3.1. *Seja S o seguinte subconjunto de $\mathbb{Z}$:*

$$\{1, 3, 4, 7, 8, 9, 13, 14, 15, 16, 21, 22, 23, 24, 25, 31, 32, 33, 34, 35, 36, 42, \ldots\}$$

Isto é, para cada $k \geq 1$ incluímos em S um bloco de k inteiros consecutivos e omitimos os k inteiros seguintes. Este conjunto contém intervalos com comprimento arbitrariamente grande. Portanto $D_S(S) = 1$. Por outro lado, o complementar de S também contém intervalos arbitrariamente grandes. Portanto, $D_I(S) = 0$.

Daremos a seguir uma proposição que caracteriza uma seqüência convergir em densidade:

Proposição 3.1. *Se $\{a_n\}$ é uma seqüência limitada de números reais então as seguintes afirmações são equivalentes:*

(i) $\displaystyle\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} |a_i| = 0.$

(ii) *Existe um subconjunto J de $\mathbb{Z}^+$ de densidade zero tal que $\lim_n a_n = 0$ para $n \notin J$.*

(iii) $\displaystyle\lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} |a_i|^2 = 0.$

Demonstração: Se $M \subset \mathbb{Z}^+$ seja $\alpha_M(n)$ a cardinalidade de $\{0, 1, \ldots, n-1\} \cap M$.

(i)⇒ (ii). Seja $J_k = \{n \in \mathbb{Z}^+ : |a_n| \geq 1/k\}(k \geq 0)$. Então $J_1 \subset J_2 \subset \ldots$. Cada J_k possui densidade zero com

$$\frac{1}{n}\sum_{i=0}^{n-1}|a_i| \geq \frac{1}{n}\frac{1}{k}\alpha_{J_k}(n).$$

Deste modo, existem inteiros $0 = l_0 < l_1 < l_2 < \ldots$ tal que para $n \geq l_k$,

$$\frac{1}{n}\alpha_{J_{k+1}}(n) < \frac{1}{k+1}.$$

Seja $J = \bigcup_{k=0}^{\infty}\{J_{k+1} \cap [l_k, l_{k+1})\}$. Mostraremos agora que J possui densidade zero. Como $J_1 \subset J_2 \subset \ldots$, se $l_k \leq n < l_{k+1}$, temos

$$J \cap [0, n) = \{J \cap [0, l_k)\} \cup \{J \cap [l_k, n)\} \subset \{J_k \cap [0, l_k)\} \cup \{J_{k+1} \cap [0, n)\},$$

e deste modo,

$$\frac{1}{n}\alpha_J(n) \leq \frac{1}{n}\left[\alpha_{J_k}(l_k) + \alpha_{J_{k+1}}(n)\right] \leq \frac{1}{n}\left[\alpha_{J_k}(n) + \alpha_{J_{k+1}}(n)\right] < \frac{1}{k} + \frac{1}{k+1}.$$

Como $(1/n)\alpha_J(n) \to 0$ com $n \to \infty$, então J possui densidade zero. Se $n > l_k$ e $n \notin J$, então $n \notin J_{k+1}$, e deste modo $|a_n| < 1/(k+1)$. Com

$$\lim_{n \notin J \to \infty} |a_n| = 0.$$

(ii)⇒ (i) Suponha que $|a_n| \leq K$ $\forall n$. Seja $\epsilon > 0$. Existe N_ϵ tal que $n \geq N_\epsilon$, $n \notin J$ implica $|a_n| < \epsilon$ e tal que $N \geq N_\epsilon$ implica $(\alpha_J(n)/n) < \epsilon$. Então $n \geq N_\epsilon$ implica

$$\begin{aligned}
\frac{1}{n}\sum_{i=0}^{n-1}|a_i| &= \frac{1}{n}\left[\sum_{i \in J \cap \{0,1,\ldots,n-1\}}|a_i| + \sum_{i \notin J \cap \{0,1,\ldots,n-1\}}|a_i|\right] \\
&< \frac{K}{n}\alpha_J(n) + \epsilon < (K+1)\epsilon.
\end{aligned}$$

(i)⇔ (iii) Pelo que foi feito acima podemos notar que $\lim_{n \notin J \to \infty}|a_n| = 0$ se, e somente se, $\lim_{n \notin J \to \infty}|a_n|^2 = 0$ ■

3.2 Teorema de Szemerédi

Em 1975, E. Szemerédi provou o seguinte teorema conjecturado por Erdös e Turan:

Teorema 3.2 (Teorema de Szemerédi). *Se S é um subconjunto de $\mathbb{Z}$ com densidade superior positiva, então S contém progressões aritméticas arbitrariamente grandes.*

Em 1976, Furstenberg possibilitou uma prova embasada na Teoria Ergódica com o seguinte teorema:

Teorema 3.3 (Teorema de Recorrência Multipla de Fusrstenberg.). *Seja $(X, \mathcal{B}, \mu)$ um espaço de probabilidade, seja T uma transformação inversível preservando medida em $(X, \mathcal{B}, \mu)$, e seja $A \in \mathcal{B}$ um conjunto de medida positiva. Então para qualquer inteiro k, existe um subconjunto $B \subset A$ com $\mu(B) > 0$ e um inteiro $n \geq 1$ com*

$$T^n B \subset A, \quad T^{2n} B \subset A, \quad \ldots, \quad T^{kn} B \subset A,$$

ou de forma equivalente,

$$\mu\left(\bigcap_{j=0}^{k} T^{-jn} A\right) > 0.$$

Assumindo este resultado provaremos o Teorema de Szemerédi.

Demonstração do Teorema de Szemerédi: Seja $X = \{0,1\}^{\mathbb{Z}}$ e seja $f : X \to X$ a aplicação deslocamento. Dado $S \subset \mathbb{Z}$ podemos definir uma sequência $\alpha = (\alpha_n)_{n \in \mathbb{Z}} \in X$ dada por

$$\alpha_n = \begin{cases} 1 & \text{se } n \in S, \\ 0 & \text{se } n \notin S. \end{cases}$$

Se S possui densidade superior positiva, existe $c > 0$ e uma sequência de intervalos $I_n = [a_n, b_n)$ de $\mathbb{Z}$ com $\lim \# I_n = \infty$ e tais que

$$\limsup_{\# I_n \to \infty} \frac{\#(S \cap I_n)}{\# I_n} > c > 0.$$

Defina o subconjunto $A \subset X$ por $A = \{y \in X; y_0 = 1\}$. Assim

$$f^j(\alpha) \in A \Leftrightarrow j \in S \tag{3.1}$$

Deste modo, mostrar que $m + in \in S$ equivale a mostrar que $f^{m+in}(\alpha) \in A$.

O conjunto A é um aberto e ao mesmo tempo um fechado de X, considerando a topologia dada onde os cilindros são abertos, pois A é um cilindro de comprimento 1 em X e seu complementar é uma união de cilindros.

Definimos a sequência μ_n de probabilidades em X por:

$$\mu_n = \frac{1}{\#I_n} \sum_{i=a_n}^{b_n-1} \delta_{f^i(\alpha)} \tag{3.2}$$

Sabemos que o conjunto das probabilidades $\mathcal{M}(X)$ munido com a topologia fraca* é compacto. Assim, podemos garantir que alguma subsequência μ_{n_i} converge para uma probabilidade μ de X. Para não carregar a notação, supomos que a própria μ_n converge para μ na topologia fraca*. Observe que μ é uma probabilidade f-invariante, pois para toda função limitada continua $\phi : X \to \mathbb{R}$ vale

$$\int \phi \circ f d\mu = \lim_{n\to\infty} \int \phi \circ f d\mu_n = \lim_{n\to\infty} \frac{1}{\#I_n} \sum_{i=a_n}^{b_n-1} \phi(f^i(\alpha))$$

$$+ \lim_{n\to\infty} \frac{\phi(f^{b_n}(\alpha)) - \phi(f^{a_n}(\alpha))}{\#I_n} = \lim_{n\to\infty} \int \phi d\mu_n = \int \phi d\mu.$$

Temos que A é um conjunto fechado e aberto de X. Logo, pela Proposição 1.5 temos que

$$\mu(A) = \lim_{n\to\infty} \mu_n(A) = \lim_{\#I_n\to\infty} \frac{\#(S \cap I_n)}{\#I_n} > c > 0.$$

Dado $k \in \mathbb{N}$, o Teorema de Recorrência Multipla de Poincaré nos garante que existe algum $n \geq 1$ tal que

$$\mu(A \cap f^{-n}(A) \cap f^{-2n}(A) \cap ... \cap f^{-kn}(A)) > 0.$$

Em particular, existe algum l tal que

$$\mu_l(A \cap f^{-n}(A) \cap f^{-2n}(A) \cap ... \cap f^{-kn}(A)) > 0.$$

Como $\mu_l = (1/\#I_l) \sum_{i=a_l}^{b_l-1} \delta_{f^i(\alpha)}$, podemos garantir que pelo menos para algum $a_l \leq m \leq b_l - 1$, o ponto$f^m(\alpha)$ pertence a $A \cap f^{-n}(A) \cap f^{-2n}(A) \cap ... \cap f^{-kn}(A)$. Assim, $f^{m+in}(\alpha) \in A$, para $i = 0, 1, ..., k$, como queríamos provar. ■

3.3 Equivalência

Verificaremos agora a equivalência dos teoremas, demonstrando que o Teorema de Szemerédi implica o Teorema de Recorrência Multipla de Furstenberg. Para isso daremos primeiro uma versão finita do Teorema de Szemerédi.

Teorema 3.4. *Para todo $\epsilon > 0$ e k inteiro positivo, existe $N = N(\epsilon, k)$ tal que se S é um conjunto de inteiros contido em algum intervalo $[a, b]$ tal que $b - a > N$ e $\#S \geq \epsilon(b - a)$, então S contém uma progressão aritmética de comprimento k.*

Demonstração: Afirmamos que 3.2 $\Rightarrow$ 3.4. De fato, se 3.4 fosse falso nós poderiamos ter um $\epsilon > 0$ e um inteiro positivo k tal que para todo N existe uma sequência S_N carregada de um intervalo $[a_N, b_N]$, $b_N - a_N > N$, $\#S_N > \epsilon(b_N - a_N)$ e S_N não contendo uma progressão aritmética de comprimento k. A propriedade listada para S_N é invariante por translação, assim podemos assumir que os S_N's estão bem separados, digamos $a_{N+1} > b_N + (b_N - a_1)$. Escrevendo $S = \cup S_N$ vemos que S possui densidade superior positiva ($> \epsilon$) e não contém progressões aritméticas de comprimento k que, por causa da separação, tal progressão teria de estar contida em um dos S_N's. Temos assim uma contradição do Teorema 3.2. ■

Deduzimos do Teorema 3.4 a seguinte consequência.

Teorema 3.5. *Sejam $\epsilon > 0$ e k dados, escreva $N_1 = N(\epsilon/2, k)$ (de 3.4). Sejam $(X, \mathcal{B}, \mu)$ um espaço de probabilidade e $B_l \in \mathcal{B}$, $\mu(B_l) \geq \epsilon$ para $l = 1, 2, \ldots, N_1$. Então existe uma progressão aritmética de comprimento k em $\{1, \ldots, N_1\}$, digamos $\{a + mb\}_{m=0}^{k-1}$, tal que*

$$\mu\left(\bigcap_{m=0}^{k-1} B_{a+mb}\right) > \frac{\epsilon}{2} N_1^{-2}. \tag{3.3}$$

Demosntração: Para $x \in X$ escreva $S(x) = \{l; 1 \leq l \leq N_1, x \in B_l\}$. Temos que $\#S(x) = \sum_{l=1}^{N_1} 1_{B_l}(x)$ e assim

$$\int \#S(x) d\mu = \sum_{l=1}^{N_1} \mu(B_l) \geq N_1 \epsilon,$$

e consequentemente

$$\mu(\{x; \#S(x) \geq \epsilon N_1/2\}) \geq \epsilon/2. \tag{3.4}$$

De fato, pelo contrário, sendo $A = \{x; \#S(x) \geq \epsilon N_1/2\}$, teríamos

$$\begin{aligned} N_1\epsilon &\leq \int \#S(x)d\mu(x) = \int_A \#S(x)d\mu(x) + \int_{X-A} \#S(x)d\mu(x) \\ &< \frac{\epsilon N_1}{2} + \frac{\epsilon N_1}{2} = \epsilon N_1. \end{aligned}$$

Pelo Teorema 3.4 e pela escolha de N_1, para cada ponto $x \in A$, $S(x)$ contém uma progressão aritmética de comprimento k, digamos $\{a(x) + mb(x)\}_{m=0}^{k-1}$. Existem menos do que N_1 escolhas para $a(x)$ ou $b(x)$ que implica, em vista de (3.4), que para algum par (a, b) temos (3.3). ■

A prova do Teorema de Recorrência Multipla de Furstenberg assumindo que vale o Teorema de Szemerédi (e consequentemente 3.5) é feita escrevendo $B_l = T^{-l}A$ e notando que

$$\mu\left(\bigcap_{m=0}^{k-1} T^{-(a+mb)}A\right) = \mu\left(T^{-a}\bigcap_{m=0}^{k-1} T^{-bm}A\right) = \mu\left(\bigcap_{m=0}^{k-1} T^{-bm}A\right).$$

Capítulo 4

Teorema de Recorrência Múltipla de Furstenberg

O objetivo deste capítulo é demonstrar o Teorema de Recorrência mencionado no capítulo anterior.

Teorema 4.1 (Teorema de Recorrência Multipla de Furstenberg.)**.** *Seja $(X, \mathcal{B}, \mu)$ um espaço de probabilidade, seja T uma transformação preservando medida inversível em $(X, \mathcal{B}, \mu)$, e seja $A \in \mathcal{B}$ um conujnto de medida positiva. Então para qualquer inteiro k, existe um subconjunto $B \subset A$ com $\mu(B) > 0$ e um inteiro $n \geq 1$ com*

$$T^n B \subset A, \quad T^{2n} B \subset A, \quad \ldots, \quad T^{kn} B \subset A.$$

ou que resulta no mesmo,

$$\mu\left(\bigcap_{j=0}^{k-1} T^{-jn} A\right) > 0.$$

O resultado desejado será conseqüência do teorema seguinte.

Teorema 4.2. *Para qualquer sistema preservando medida $(X, \mathcal{B}, \mu, T)$ e $A \in \mathcal{B}$ com $\mu(A) > 0$, e para qualquer $k = 1, 2, \ldots$*

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n} A \cap \ldots \cap T^{-kn} A) > 0. \tag{4.1}$$

Nas duas primeiras seções deste capítulo demonstraremos que este teorema é válido para dois casos especiais de sistemas preservando medida $(X, \mathcal{B}, \mu, T)$, *sistemas fracamente mixing* e *sistemas compactos*. Porém estes dois sistemas não exaurem todas as possibilidades. Contudo, na Seção 4.3 mostraremos que se $(X, \mathcal{B}, \mu, T)$ não é fracamente mixing, então existe uma σ-álgebra T-invariante $\mathcal{B}_l \subset \mathcal{B}$ tal que T restrito à $\mathcal{B}_l$ cumpre a afirmação do Teorema 4.2.

Para concluir o Teorema 4.2, na Seção 4.4 mostraremos que existe uma sub-σ-álgebra $\mathcal{B}_l \subset \mathcal{B}$ que é maximal, com respeito à inclusão, na classe das subálgebras T-invariantes de $\mathcal{B}$ em que o argumento do Teorema é válido. Assumindo $\mathcal{B}_l \neq \mathcal{B}$, estudaremos nas últimas seções a ação de conjuntos $A \in \mathcal{B}$ por T "relativamente à $\mathcal{B}_l$" de maneira resumida e poderemos concluir que ou a ação é "relativamente fracamente mixing" ou existe $\mathcal{B}_2 \subset \mathcal{B}_l$ para o qual a ação de T é "relativamente compacto". Em ambos os casos, teremos que existe uma subálgebra maior para o qual é válido o Teorema, contradizendo a maximalidade de $\mathcal{B}_l$. Isto implica que $\mathcal{B}_l = \mathcal{B}$ completando a prova.

4.1 Sistemas fracamente mixing

Começamos a demonstrar o Terorema de Recorrencia Múltipla de Furstenberg para uma classe especial de sistemas.

Definição 4.1. *Dizemos que um sistema preservando medida $(X, \mathcal{B}, \mu, T)$ é mixing se para quaisquer dois conjuntos $A, B \in \mathcal{B}$, $\lim_{n\to\infty} \mu(A \cap T^{-n}B) = \mu(A)\mu(B)$. O sistema é fracamente mixing se para quaisquer $A, B \in \mathcal{B}$ tivermos que*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left[\mu(A \cap T^{-n}B) - \mu(A)\mu(B)\right]^2 = 0. \tag{4.2}$$

Exemplo 4.1. ***Sistema de Bernoulli.***

Seja M o conjunto das seqüências $\{\alpha_n\}_{n\in\mathbb{Z}}$, com α_n assumindo valores em um conjunto finito $\Gamma = \{1, 2, \ldots, r\}$. Tomemos a σ-álgebra gerada pelos cilindros da forma

$$[k, l; a_k, \ldots, a_l] = \{\alpha \in M : \alpha_k = a_k, \ldots, \alpha_l = a_l\}$$

onde $k, l \in \mathbb{Z}$, com $k \leq l$ e cada $a_j \in \{1, 2, \ldots, r\}$. Definimos

$$\mu([k, l; a_k, \ldots, a_l]) = p(a_k) \ldots p(a_l)$$

onde $p(a_j) \geq 0$ e $\sum_{j=1}^{r} p(a_j) = 1$. Estendemos μ à σ-álgebra dos cilindros e a chamamos de medida de Bernoulli. Desta forma a este sistema chamamos Sistema de Bernoulli. Consideremos a transformação deslocamento

$$T : M \to M; \quad T(\{\alpha_n\}_{n \in \mathbb{Z}}) = \{\alpha_{n+1}\}_{n \in \mathbb{Z}}.$$

Afirmamos que esta transformação preserva a medida de Bernoulli. De fato se $A = [k, l; a_k, \ldots, a_l]$, então $T^{-1}(A) = [k+1, l+1; a_k, \ldots, a_l]$ e assim

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

O sistema de Bernoulli com o deslocamento definido acima, formam um sistema mixing, ou ainda, fracamente mixing. De fato, sendo $A = [k, l; a_k, \ldots, a_l]$ e $B = [p, q; b_p, \ldots, b_q]$. Para cada n tem-se $T^{-n}(B) = [p+n, q+n; b_p, \ldots, b_q]$ e assim tomando n de modo que $p + n > l$,

$$\begin{aligned} A \cap T^{-n}(B) &= \{\alpha \in M : \alpha_k = a_k, \ldots, \alpha_l = a_l, \alpha_{p+n} = b_p, \ldots, \alpha_{q+n} = b_q\} \\ &= \bigcup [k, q+n; a_k, \ldots, a_l, c_{l+1}, \ldots, c_{p+n-1}, b_p, \ldots, b_q], \end{aligned}$$

onde a união é sobre os possíveis valores de $c_{l+1}, \ldots, c_{p+n-1}$. Concluímos que $\mu(A \cap T^{-n}(B)) = \mu(A)\mu(B)$. Esta conclusão foi feita para quando os conjuntos são cilindros, para o caso geral usamos o fato de μ ser finitamente aditiva.

Em seguida daremos dois lemas a respeito dos sistemas fracamente mixing que serão muito úteis no que tange trabalhar com tais sistemas. O primeiro nos dá uma caracterização equivalente, enquanto o segundo mostra que fracamente mixing é uma propriedade mais forte que a ergodicidade.

Lema 4.1. *São equivalentes:*

(i) *O sistema $(X, \mathcal{B}, \mu, T)$ é fracamente mixing;*

(ii) *Para $f, g \in L^2(X, \mathcal{B}, \mu)$ temos que*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \left(\int f T^n g d\mu - \int f d\mu \int g d\mu \right)^2 = 0. \tag{4.3}$$

Demosntração: Tomando $f = 1_A$ e $g = 1_B$ na da equação acima, observamos que

$$\int 1_A \, T^n 1_B d\mu = \int 1_A 1_{T^{-n}B} d\mu = \int 1_{A\cap T^{-n}B} d\mu = \mu(A \cap T^{-n}B).$$

E que por outro lado,

$$\int 1_A d\mu \int 1_B d\mu = \mu(A)\mu(B)$$

Assim supondo (4.3) válido, temos que o sistema é fracamente mixing.

A recíproca se faz para funções simples, já que vale para características. E depois tomamos funções mensuráveis próximas a estas funções simples [15]. ∎

Lema 4.2. *Um sistema fracamente mixing é necessariamente ergódico.*

Demonstração: Assumindo por contradição que o sistema $(X, \mathcal{B}, \mu, T)$ é fracamente mixing mas não ergódico, pode existir um conjunto T-invariante $E \in \mathcal{B}$ onde $T^{-1}E = E$ com $0 < \mu(E) < 1$. Se tomarmos $A = E$ e $B = X\backslash E$ em (4.2), então $\mu(T^{-n}E \cap (X\backslash E)) = \mu(E \cap (X\backslash E)) = 0$ para todo $n \geq 0$, deduzimos que $\mu(E)\mu(X\backslash E) = 0$ dando a desejada contradição. ∎

Temos ainda o fato de que se o sistema $(X, \mathcal{B}, \mu, T)$ é fracamente mixing, o sistema produto $(X \times X, \mathcal{B} \times \mathcal{B}, \mu \times \mu, T \times T)$ também será. Para tanto, é suficiente que dadas duas funções em $L^2(X \times X)$, estas cumpram (4.3). Tais funções são da forma $f \otimes g(x_1, x_2)$ onde $f \otimes g(x_1, x_2) = f(x_1)g(x_2)$. Bem, (4.3) é equivalente à afirmação que para qualquer $\epsilon > 0$

$$\left| \int f T^n g d\mu - \int f d\mu \int g d\mu \right| < \epsilon, \tag{4.4}$$

porém para um conjunto de densidade zero. A mesma afirmação agora segue para o produto tensorial. Com efeito, sabendo que

$$\begin{aligned} \int f_1 \otimes f_2 (T \times T)^n g_1 \otimes g_2 d(\mu \times \mu) &= \int f_1 T^n g_1 d\mu \int f_2 T^n g_2 d\mu, \\ \int f_1 \otimes f_2 d(\mu \times \mu) &= \int f_1 d\mu \int f_2 d\mu, \\ \int g_1 \otimes g_2 d(\mu \times \mu) &= \int g_1 d\mu \int g_2 d\mu, \end{aligned} \tag{4.5}$$

temos que

$$\left|\int f_1 \otimes f_2 (T\times T)^n g_1 \otimes g_2 d(\mu\times\mu) - \int f_1\otimes f_2 d(\mu\times\mu)\int g_1\otimes g_2 d(\mu\times\mu)\right|$$

$$\begin{aligned}
&= \left|\int f_1 T^n g_1 d\mu \int f_2 T^n g_2 d\mu - \int f_1 d\mu \int g_1 d\mu \int f_2 d\mu \int g_2 d\mu\right| \\
&\leq \left|\int f_1 T^n g_1 d\mu \int f_2 T^n g_2 d\mu - \int f_1 T^n g_1 d\mu \int f_2 d\mu \int g_2 d\mu\right| \\
&+ \left|\int f_1 T^n g_1 d\mu \int f_2 d\mu \int g_2 d\mu - \int f_1 d\mu \int g_1 d\mu \int f_2 d\mu \int g_2 d\mu\right| \\
&\leq \left|\int f_1 T^n g_1 d\mu\right| \left|\int f_2 T^n g_2 d\mu - \int f_2 d\mu \int g_2 d\mu\right| \\
&+ \left|\int f_1 T^n g_1 d\mu - \int f_1 d\mu \int g_1 d\mu\right| \left|\int f_2 d\mu \int g_2 d\mu\right|.
\end{aligned}$$

e concluimos usando o fato de que f_1, f_2, g_1 e g_2 são limitadas.

Daremos agora uma recíproca fraca para o Lema 4.2.

Proposição 4.1. *Se o sistema* $(X\times X, \mathcal{B}\times\mathcal{B}, \mu\times\mu, T\times T)$ *é ergódico então* $(X, \mathcal{B}, \mu, T)$ *é fracamente mixing.*

Demonstração: Sabemos do Corolário da Teorema 1.6, que se um sistema $(X, \mathcal{B}, \mu, T)$ é ergódico, então para quaisquer $f, g \in L^2(X, \mathcal{B}, \mu)$,

$$\frac{1}{N+1}\sum_{n=0}^{N}\int fT^n g d\mu \to \int f d\mu \int g d\mu. \tag{4.6}$$

Se assumirmos $T\times T$ ergódico, então ainda o será T e teremos a equação acima bem como a afirmação correspondente para $f\otimes f$ e $g\otimes g$. Como em (4.6) teremos

$$\frac{1}{N+1}\sum_{n=0}^{N}\left(\int fT^n g d\mu\right)^2 \to \left(\int f d\mu\right)^2\left(\int g d\mu\right)^2. \tag{4.7}$$

Usaremos agora o fato de que se $(1/N)\sum_{n=1}^{N} a_n \to \alpha$ e também $(1/N)\sum_{n=1}^{N} a_n^2 \to \alpha^2$, então $(1/N)\sum_{n=1}^{N}(a_n-\alpha)^2 \to 0$, temos que (4.6) e (4.7) implicam que

$$\frac{1}{N+1}\sum_{n=0}^{N}\left(\int fT^n g d\mu - \int f d\mu \int g d\mu\right)^2 \to 0 \tag{4.8}$$

o que prova a proposição. ∎

A prova do Teorema de Recorrência no caso fracamente mixing segue da próxima proposição.

Proposição 4.2. *Se $(X, \mathcal{B}, \mu, T)$ é um sistema fracamente mixing e $A_0, A_1, \ldots, A_k$ são conjuntos em $\mathcal{B}$, então*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \left[\mu(A_0 \cap T^{-n}A_1 \cap \ldots \cap T^{-kn}A_k) - \mu(A_0)\mu(A_1)\ldots\mu(A_k)\right]^2 = 0. \tag{4.9}$$

Demonstração: Para provar a proposição, daremos a prova de duas variantes da equação (4.9) por indução em k. Sejam $f_0, f_1, \ldots, f_k \in L^\infty(X, \mathcal{B}, \mu)$ funções não negativas.

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \left[\int \prod_{l=0}^{k} T^{ln} f_l d\mu - \prod_{l=0}^{k} \int f_l d\mu\right]^2 = 0, \tag{4.10}$$

$$\lim_{N\to\infty} \left\| \frac{1}{N}\sum_{n=1}^{N} \prod_{l=0}^{k} T^{ln} f_l - \prod_{l=0}^{k} \int f_l d\mu \right\|_{L^2(X)} = 0. \tag{4.11}$$

A equação (4.9) é um caso especial da equação de (4.10) aplicada para funções característica, como feito no Lema 4.1. A equação (4.11) se referi à convergência em $L^2(X, \mathcal{B}, \mu)$ e esta convergência implica uma convergência mais fraca, assim implica que

$$\frac{1}{N}\sum_{n=1}^{N} \int f_0 \prod_{l=1}^{k} T^{ln} f_l d\mu \to \prod_{l=0}^{k} \int f_l d\mu. \tag{4.12}$$

Como $X \times X$ é também fracamente mixing, nós obtemos o análogo de (4.12) substituindo f por $f \otimes f$ e T por $T \times T$. As integrais em $X \times X$ tornam-se produtos de integrais em X e como feito em (4.6), obtemos

$$\frac{1}{N}\sum_{n=1}^{N} \left[\int f_0 \prod_{l=1}^{k} T^{ln} f_l d\mu\right]^2 \to \prod_{l=0}^{k} \left[\int f_l d\mu\right]^2. \tag{4.13}$$

Usando o mesmo argumento que em (4.8) temos que (4.12) e (4.13) implicam (4.10) e, deste modo, supondo (4.11) válido, também vale (4.10).

Por outro lado, (4.10) com $k = 1$ é simplismente a equação (4.3). Portanto, para concluir a hipótese de indução e obter a proposição precisamos mostrar que se (4.10) é válida para $k-1$, então (4.11) vale para k.

Agora, para obter (4.11) é suficiente considerar em geral o caso onde algum $\int f_0 d\mu = 0$ e obter assim

$$\frac{1}{N}\sum_{n=1}^{N}\prod_{l=1}^{k} T^{ln} f_l \to 0 \tag{4.14}$$

em $L^2(X, \mathcal{B}, \mu)$.

De fato, considerando a válida identidade

$$\prod_{l=1}^{k} a_l - \prod_{l=1}^{k} b_l = \sum_{j=1}^{k}\left(\prod_{l=1}^{j-1} a_l\right)(a_j - b_j)\left(\prod_{l=j+1}^{k} b_l\right), \tag{4.15}$$

e tomando

$$g_1 = f_1,\ g_2 = f_2,\ \ldots,\ g_{j-1} = f_{j-1},\ g_j = \left(f_j - \int f_j d\mu\right),\ g_{j+1} = \int f_{j+1} d\mu,\ \ldots, g_k = \int f_k d\mu,$$

teremos que

$$\begin{aligned}\prod_{l=1}^{k} T^{ln} g_l &= \left(\prod_{l=1}^{j-1} T^{ln} f_l\right)\left(T^{jn}\left(f_j - \int f_j d\mu\right)\right)\left(\prod_{l=j+1}^{k} T^{ln}\int f_l d\mu\right)\\ &= \left(\prod_{l=1}^{j-1} T^{ln} f_l\right)\left(T^{jn} f_j - \int f_j d\mu\right)\left(\prod_{l=j+1}^{k}\int f_l d\mu\right).\end{aligned}$$

Assim, podemos substituir $\prod_{l=1}^{k} T^{ln} f_l - \prod_{l=1}^{k}\int f_l d\mu$ por uma soma de produtos com a propriedade da equação (4.14).

Agora tomamos

$$\psi_N = \frac{1}{N}\sum_{n=1}^{N}\prod_{l=1}^{k} T^{ln} f_l.$$

e seja $N \to \infty$. Fixamos um número H muito grande. Mostraremos a convergência exigida a ψ_N em (4.14) dividindo-o em duas partes. Chamando $G_n = \prod_{l=1}^{k} T^{ln} f_l$ temos que

$$\begin{aligned}\psi_N &= \frac{1}{N}\left[G_1 + G_2 + \ldots + G_N\right]\\ &= \frac{1}{N}\left[\frac{1}{H}G_1 + \frac{2}{H}G_2 + \ldots + \frac{H-1}{H}G_{H-1} + G_H + G_{H+1} + \ldots + G_N\right] +\\ &\quad \frac{1}{N}\left[\frac{H-1}{H}G_1 + \frac{H-2}{H}G_2 + \ldots + \frac{1}{H}G_{H-1}\right].\end{aligned}$$

Reescrevemos $\psi_N = \psi_N^{'} + \psi_N^{''}$, onde

$$\begin{aligned}
\psi_N^{'} &= \frac{1}{N}\left[\frac{1}{H}G_1 + \frac{2}{H}G_2 + \ldots + \frac{H-1}{H}G_{H-1} + G_H + G_{H+1} + \ldots + G_N\right] \\
&= \frac{1}{N}\sum_{j=1}^{N}\left[\frac{1}{H}\sum_{n=j}^{\min\{j+H-1,N\}} G_n\right], \\
\psi_N^{''} &= \frac{1}{N}\left[\frac{H-1}{H}G_1 + \frac{H-2}{H}G_2 + \ldots + \frac{1}{H}G_{H-1}\right].
\end{aligned}$$

ou seja,

$$\psi_N = \frac{1}{N}\sum_{j=1}^{N}\left[\frac{1}{H}\sum_{n=j}^{\min\{j+H-1,N\}}\prod_{l=1}^{k} T^{ln} f_l\right] + \psi_N^{''} = \psi_N^{'} + \psi_N^{''}. \tag{4.16}$$

Observamos que como f_l são limitadas, $\psi_N^{''} \to 0$ com $N \to \infty$ por $\psi_N^{''}$ se tratar da divisão, por N, de uma soma finita de funções todas cotadas superiormente por um mesmo valor. Disto, vemos que será suficiente mostrar, para algum H apropriado, que $\limsup_{N\to\infty} \left\|\psi_N^{'}\right\|_{L^2(X)} < \epsilon$. Sabendo que a função quadrática é convexa, temos que o quadrado da média é menor ou igual à média dos quadrados, logo

$$\psi_N^{'\,2} \leq \frac{1}{N}\sum_{j=1}^{N}\left[\frac{1}{H}\sum_{n=j}^{\min\{j+H-1,N\}}\prod_{l=1}^{k} T^{ln} f_l\right]^2,$$

e assim

$$\begin{aligned}
\left\|\psi_N^{'}\right\|_{L^2(X)}^2 &\leq \frac{1}{N}\sum_{j=1}^{N}\frac{1}{H^2}\sum_{n,m=j}^{\min\{j+H-1,N\}}\int\prod_{l=1}^{k} T^{ln} f_l T^{lm} f_l d\mu \\
&= \frac{1}{NH^2}\sum_{j=1}^{N}\sum_{n,m=j}^{\min\{j+H-1,N\}}\int\prod_{l=1}^{k} T^{ln}\left(f_l T^{l(m-n)} f_l\right) d\mu.
\end{aligned}$$

Como T, e deste modo T^n, é uma transformação preservando medida, nós podemos substituir cada T^{ln} por $T^{(l-1)n}$ na expressão anterior, obtendo assim,

$$\left\|\psi_N^{'}\right\|_{L^2(X)}^2 \leq \frac{1}{NH^2}\sum_{j=1}^{N}\sum_{n,m=j}^{\min\{j+H-1,N\}}\int\prod_{l=0}^{k-1} T^{ln}\left(f_{l+1} T^{(l+1)(m-n)} f_{l+1}\right) d\mu. \tag{4.17}$$

Observamos que as integrais acima são as que ocorrem em (4.10) para $k-1$ com as funções f_l substituidas por $g_{l,m-n} = f_{l+1}T^{(l+1)(m-n)}f_{l+1}$.

Sendo $r = m - n$, notamos que o par (n, m) aparece em (4.17) somente se $|r| = |m - n| < H$, e então para $H - |r|$ valores de j, assim reescrevemos (4.17) como

$$\left\|\psi_N'\right\|_{L^2(X)}^2 \leq \frac{1}{H} \sum_{r=1-H}^{H-1} \left(1 - \frac{|r|}{H}\right) \left(\frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=0}^{k-1} T^{ln} g_{l,r} d\mu\right). \tag{4.18}$$

O fato de tal substituição está melhor demonstrado no Apêndice deste texto.

Finalmente, pela hipótese de indução, podemos usar (4.10) para substituir o lado direito da desigualdade acima e obetermos

$$\left\|\psi_N'\right\|_{L^2(X)}^2 \leq \frac{1}{H} \sum_{r=1-H}^{H-1} \left(1 - \frac{|r|}{H}\right) \left(\prod_{l=0}^{k-1} \int g_{l,r} d\mu\right) + \frac{\epsilon}{2}, \tag{4.19}$$

para N grande. Por outro lado, sendo algum $\int f_{l_0} d\mu = 0$ e por (4.3) a maior parte dos termos $\int g_{l_0-1,r} d\mu$ é pequeno para H grande, pois

$$\int g_{l_0-1,r} d\mu = \int f_{l_0} T^{l_0 r} f_{l_0} d\mu$$

está, em média, próximo de $\left(\int f_{l_0} d\mu\right)^2 = 0$.

Por outro lado, como todos os integrandos da equação (4.19) são limitados, nós podemos tomar H tão grande que

$$\left|\frac{1}{H} \sum_{r=1-H}^{H-1} \left(1 - \frac{|r|}{H}\right) \left(\prod_{l=0}^{k-1} \int g_{l,r} d\mu\right)\right| < \frac{\epsilon}{2}$$

e assim

$$\limsup_{N\to\infty} \left\|\psi_N'\right\|_{L^2(X)}^2 < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

completando a indução e a prova da proposição. ■

4.2 Sistemas compactos

Nesta seção demonstraremos o Teorema de Recorrência Múltipla de Furstenberg para um segundo caso de sistemas.

Definição 4.2. *Dizemos que o sistema preservando medida $(X, \mathcal{B}, \mu, T)$ é compacto se para qualquer função $f \in L^2(X, \mathcal{B}, \mu)$ o fecho da órbita $\{f, Tf, T^2 f, \ldots, T^n f, \ldots\}$ é compacta em $L^2(X, \mathcal{B}, \mu)$.*

Exemplo 4.2. *Um exemplo bem simples ocorre para transformações periódicas, isto é,* $T^p =$ *identidade para algum* p*. Isto ocorre por exemplo com a rotação racional do círculo. De uma maneira menos trivial é ainda possível mostrar que a rotação irracional no círculo também forma um sistema compacto.*

A seguir veremos que esta classe de sistemas também cumpre as condições do Teorema 4.2.

Proposição 4.3. *Se o sistema inversível preservando medida* $(X, \mathcal{B}, \mu, T)$ *é compacto, então para qualquer* $f \in L^\infty(X, \mathcal{B}, \mu), f \geq 0$ *com* f *não quase sempre nula e para qualquer inteiro* k*,*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int f T^n f T^{2n} f \dots T^{kn} f d\mu > 0. \tag{4.20}$$

Tomando $f = 1_A$ a função característica de um conjunto $A \in \mathcal{B}$ com $\mu(A) > 0$ na equação (4.20), obtemos

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int 1_A T^n 1_A T^{2n} 1_A \dots T^{kn} 1_A d\mu$$

$$\begin{aligned}
&= \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int 1_A \ 1_{(T^{-n}A)} \ 1_{(T^{-2n}A)} \ \dots \ 1_{(T^{-kn}A)} d\mu \\
&= \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int 1_{(A\cap T^{-n}A\cap T^{-2n}A\dots\cap T^{-kn}A)} d\mu \\
&= \liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n}A \cap T^{-2n}A \dots \cap T^{-kn}A),
\end{aligned}$$

portanto válida a proposição, teremos que o Teorema de Recorrência Múltipla de Furstenberg é válido para sistemas compactos.

Demonstração da proposição 4.3: Seja $a = \int f^{k+1} d\mu$, temos então que $a > 0$. Podemos assumir sem perda da generalidade que $0 \leq f \leq 1$. Tomemos $\epsilon < a/(k+1)$. Sejam $g_0, \dots, g_k$ funções mensuráveis com $0 \leq g_i \leq 1$ e com $\|f - g_i\| < \epsilon$, $i = 0, 1, \dots, k$. Com base na equação (4.15), usada na demonstração da Proposição 4.2, temos

$$\begin{aligned}
\left|\int \prod_{l=0}^{k} g_l d\mu - \int f^{k+1} d\mu\right| &= \left|\int \prod_{l=0}^{k} g_l d\mu - \int \prod_{l=0}^{k} f d\mu\right| \\
&= \left|\int \left[\prod_{l=0}^{k} g_l - \prod_{l=0}^{k} f\right] d\mu\right| \\
&= \left|\int \sum_{j=0}^{k} \prod_{l=1}^{j-1} g_l (g_j - f) f^{k-j} d\mu\right| \\
&\leq \sum_{j=0}^{k} \int \prod_{l=1}^{j-1} g_l \left|g_j - f\right| f^{k-j} d\mu \\
&\leq (k+1)\epsilon < a.
\end{aligned}$$

Se tomarmos $a' = a - (k+1)\epsilon$ então devemos ter $\int \prod_{l=0}^{k} g_l d\mu \geq a'$.

Usando esta observação, nosso trabalho então será mostrar que para um conjunto de n com densidade inferior positiva, $\left\|T^{ln} f - f\right\| < \epsilon$. Ainda, se mostrarmos que $\|T^n f - f\| < \epsilon/k$ para um conjunto de n com densidade inferior positiva, então como T preserva medida teremos que $\left\|T^{(j+1)n} f - T^{jn} f\right\| < \epsilon/k$ para este conjunto de n, e pela desigualdade triangular, $\left\|T^{ln} f - f\right\| < \epsilon$ para $l = 0, 1, \ldots, k$.

A propriedde desejada segue da compacidade do fecho da órbita $\{T^n f\} \subset L^2(X, \mathcal{B}, \mu)$, pois teremos que o conjunto $\{T^n f, n = 0, 1, 2, \ldots\}$ é totalmente limitado. Por este fato podemos encontrar um subconjunto $\{T^{n_1} f, T^{n_2} f, \ldots, T^{n_r} f\}$ de modo $\|T^{n_i} f - T^{n_j} f\| \geq \epsilon/k$, tal que qualquer elemento da órbita diste menos que ϵ/k de alguém deste subconjunto. Agora, como T preserva medida, para qualquer n, o subconjunto $\{T^{n+n_1} f, T^{n+n_2} f, \ldots, T^{n+n_r} f\}$ possui a mesma propriedade. Deste modo, para cada n existe $1 \leq i(n) \leq r$ tal que $\|T^{n+n_{i(n)}} f - f\| < \epsilon$. Em particular, a seqüência $\{n + n_{i(n)}\}_{n=0}^{\infty}$ é uma seqüência de densidade inferior positiva, e temos o resultado desejado. ∎

4.3 Sistemas fracamente mixing e fatores compactos

Nesta seção começaremos a nos aproximar de fato da demonstração do Teorema 4.2. Diremos que um fator $(X, \mathcal{B}_l, \mu, T)$ será não trivial se $\mathcal{B}_l$ contiver conjuntos de medida estritamente entre 0 e 1.

O propósito central desta seção é assegurar que se um sistema não é fracamente mixing, algum fator não trivial será compacto. Deste modo, com os resultados obtidos nas duas seções anterios o Teorema 4.2 é válido para algum fator não trivial de um sistema arbitrário.

Proposição 4.4. *Um sistema preservando medida $(X, \mathcal{B}, \mu, T)$ é fracamente mixing se, e somente se, este não possui um fator compacto não trivial.*

Definimos anteriormente que um sistema seria compacto desde que para toda $f \in L^2(\mathbf{X})$ o fecho de sua órbita por T fosse compacto. Agora olharemos para quando isto acontece para alguma f e não todas.

Definição 4.3. *Dizemos que $f \in L^2(\boldsymbol{X})$ é AP (quase periódica) se o fecho de sua órbita for compacto.*

Proposição 4.5. *Se para um sistema preservando medida $(X, \mathcal{B}, \mu, T)$ o quadrado $T \times T$ não é ergódico, então existirá um função não constante $f \in L^2(\boldsymbol{X})$ que é AP.*

Demosntração: Seja $H(x, x')$ uma função não constante $T \times T$-invariante em $L^2(\mathbf{X} \times \mathbf{X})$.

Sendo T não ergódica, existirá $A \in \mathcal{B}$ com $T^{-1}A = A$ onde $0 < \mu(A) < 1$, logo a função 1_A será uma função não constante quase periódica e o Lema estará demonstrado. Suponhamos então que T seja ergódica.

A função $\int H(x, x')d\mu(x)$ é T-invariante, e portanto constante. Subtraindo essa constante a H nós podemos supor que esta desaparece. Como H não é nula, existe necessariamente uma função $\phi \in L^2(X, \mathcal{B}, \mu)$ com $\int H(x, x')\phi(x')d\mu(x') \neq 0$ para um conjunto, de x, de medida positiva. Disto segue que

$$f(x) = \int H(x, x')\phi(x')d\mu(x') \tag{4.21}$$

é ainda não constante, como $\int f(x)d\mu(x) = \int \phi(x') \int H(x, x')d\mu(x)d\mu(x') = 0$.

Agora a função em (4.21) é quase periódica. Por

$$T^n f(x) = \int H(T^n x, x')\phi(x')d\mu(x') = \int H(T^n x, T^n x')\phi(T^n x')d\mu(x'),$$

pela invariancia de μ, ou

$$T^n f(x) = \int H(x, x')T^n\phi(x')d\mu(x').$$

Se $\tilde{H} : L^2(X, \mathcal{B}, \mu) \to L^2(X, \mathcal{B}, \mu)$ denota o operador integral

$$\tilde{H}\psi(x) = \int H(x, x')\psi(x')d\mu(x'),$$

então $\overline{\{T^n f\}} = \overline{\{\tilde{H}(T^n\phi)\}}$. Entretanto é sabido que o operador $\tilde{H}$ é compacto e como a norma de $T^n\phi$ é compacta, temos que $\overline{\{\tilde{H}(T^n\phi)\}}$ é compacto. $\blacksquare$

Para uma leitura sobre operadores compactos [12].

Para provar a Proposição 4.4 provaremos primeiro os seguintes Lemas:

Lema 4.3. *O conjunto das funções $\phi \in L^2(X, \mathcal{B}, \mu)$ que são AP é um subespaço linear fechado de $L^2(X, \mathcal{B}, \mu)$. Ainda, este subespaço é fechado para as operações $(\phi_1, \phi_2) \mapsto \max\{\phi_1, \phi_2\}$ e $(\phi_1, \phi_2) \mapsto \min\{\phi_1, \phi_2\}$.*

Demonstração: Primeiro, lembremos que um subconjunto de um espaço métrico completo possui fecho compacto se, e somente se, para todo $\epsilon > 0$, o subconjunto pode ser coberto por uma quantidade finita de bolas de raio $\leq \epsilon$. Usando isto, pode-se verificar que o conjunto das funções AP em $L^2(X, \mathcal{B}, \mu)$ é um subespaço linear fechado de $L^2(X, \mathcal{B}, \mu)$. $\blacksquare$

Lema 4.4. *Se f é AP e $\mathcal{B}_0$ é a menor σ-álgebra de conjuntos com respeito a que f é mensurável, então cada $1_A, A \in \mathcal{B}_0$ é AP.*

Demonstração: Temos que $\mathcal{B}_0$ é gerada pelas imagens inversas de intervalos abertos. Se $A_1 = f^{-1}((a, b))$, então 1_{A_1} é AP. De fato, pelo Lema 4.3, são AP as funções:

$h_1 = \min\{f, b\}$

$h_2 = \max\{a, h_1\}$

$h_3 - \frac{h_2 - a}{b - a}$

$h_4 = 1 - h_3$

$h = \min\{h_3, h_4\}$

$g_n = n \min\{h, 1/n\}$.

Notando que $g_n \to 1_{A_1}$, usando novamente o Lema 4.3 temos que 1_{A_1} é AP. Por fim, podemos estender para funções características dos conjuntos na σ-álgebra $\mathcal{B}_0$. $\blacksquare$

Retornaremos à Proposição 4.4.

Demonstração da proposição 4.4: Assuma que $(X, \mathcal{B}, \mu, T)$ é um sistema fracamente mixing. Devemos mostrar que não existe função não constante $f \in L^2(X, \mathcal{B}, \mu)$ com o fecho de sua órbita $\overline{\{T^n f\}} \subset L^2(X, \mathcal{B}, \mu)$ compacto.

A prova da Proposição 4.3 mostra que se f possui o fecho da órbita compacto, para qualquer $\epsilon > 0$, existe um subconjunto $S \subset \mathbb{N}$ de densidade inferior positiva de modo que para $n \in S$, $\|f - T^n f\|_{L^2(X)} < \epsilon$. Por outro lado (4.4) implica que para qualquer $\delta > 0$, $f, g \in L^2(X, \mathcal{B}, \mu)$

$$\left| \int f T^n g d\mu - \int f d\mu \int g d\mu \right| < \delta$$

para um conjunto n de densidade 0. Em particular,

$$\left| \int f T^n f d\mu - \left(\int f d\mu \right)^2 \right| < \delta$$

mas para um conjunto de n de densidade 0. Para algum $n \in S$ teremos

$$\left| \int f^2 d\mu - \left(\int f d\mu \right)^2 \right| < \delta + \epsilon \|f\|_{L^2(X)},$$

e como ϵ, δ são arbitrários $\int f^2 d\mu = \left(\int f d\mu\right)^2$, que para f real implica $f =$ constante em quase todo ponto. De fato, $\int \left(f - \int f d\mu\right)^2 d\mu = \int f^2 d\mu - 2(\int f d\mu)^2 + (\int f d\mu)^2 = 0$, logo $(f - \int f d\mu) = 0$ para quase todo ponto.

Suponha que $(X, \mathcal{B}, \mu, T)$ não é mixing. Então, pela Proposição 4.1, existe uma função não trivial invariante em $X \times X$. Pela Proposição 4.5 existe uma função não constante $f \in L^2(X, \mathcal{B}, \mu)$ que é AP. Construiremos então uma σ-álgebra não trivial $\mathcal{B}_l$, invariante com respeito a T, de modo que o fator $(X, \mathcal{B}_l, \mu, T)$ é compacto.

Usamos o Lema 4.4, como ϕ é AP se, e somente se, $T\phi$ é AP, o mesmo é verdade para $\mathcal{B}_l$ como sendo a menor σ-álgebra de conjuntos com respeito a que $f, Tf, T^2 f, \ldots$ é mensurável. Finalmente, se cada 1_A, $A \in \mathcal{B}_l$ é AP, então será cada $\phi \in L^2(X, \mathcal{B}, \mu)$. Disto segue que este fator $(X, \mathcal{B}_l, \mu, T)$ é compacto. ■

Definição 4.4. *Diremos que um sistema* $(X, \mathcal{B}, \mu, T)$ *é um sistema SZ quando para todo* $A \in \mathcal{B}$ *com* $\mu(A) > 0$ *e para todo* $k \geq 1$ *tivermos que*

$$\liminf_{N\geq 1} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n}A \cap T^{-2n}A \cap \ldots \cap T^{-kn}A) > 0. \tag{4.22}$$

4.4 Fatores maximais SZ

Seja $(X, \mathcal{B}, \mu, T)$ um sistema inversível preservando medida.

Proposição 4.6. *Seja* $\{\mathcal{B}_t\}$ *uma família totalmente ordenada de sub-*σ*-álgebras de* $\mathcal{B}$*. Seja* $\tilde{\mathcal{B}}$ *a* σ*-álgebra gerada pela união* $\bigcup \mathcal{B}_t$*. Se cada* $(X, \mathcal{B}_t, \mu, T)$ *é um sistema SZ, então* $(X, \tilde{\mathcal{B}}, \mu, T)$ *também é um sistema SZ.*

Demonstração: Seja $A \in \tilde{\mathcal{B}}$, com $\mu(A) > 0$, seja k um dado inteiro e seja $\epsilon > 0$. Como $\bigcup \mathcal{B}_t$ é uma álgebra que gera $\tilde{\mathcal{B}}$ como σ-álgebra, podemos encontrar $A' \in \mathcal{B}_{t'}$ para algum t' com $\mu(A \Delta A') < \epsilon$ (Teorema 1.3). Então, pelo fato de $\|1_{A\Delta A'}\|_{L^2} = \|1_A - 1'_A\|_{L^2}$, temos $\|1_A - 1_{A'}\|_{L^2} < \sqrt{\epsilon}$ em $L^2(\mathbf{X})$. Se $\mathbf{X}_{t'} = (X, \mathcal{B}_{t'}, \mu, T)$, então $E(1_A|\mathbf{X}_{t'})$ é a projeção ortogonal de 1_A no subespaço das funções $L^2(\mathbf{X}_{t'})$. Como $1_{A'}$ está neste subespaço, nós devemos ter

$$\|1_A - E(1_A|\mathbf{X}_{t'})\|_{L^2} < \sqrt{\epsilon}.$$

Disto podemos deduzir que se ϵ for suficientemente pequeno, $E(1_A|\mathbf{X}_{t'}) > 1 - (1/2k)$ em um conjunto de medida positiva. Do contrário, caso $E(1_A|\mathbf{X}_{t'}) \leq 1 - (1/2k)$ quase sempre temos $1_A - E(1_A|\mathbf{X}_{t'}) \geq (1/2k)$ no conjunto de medida $\mu(A)$ onde

$$\|1_A - E(1_A|\mathbf{X}_{t'})\|_{L^2} \geq \frac{\sqrt{\mu(A)}}{2k}.$$

Assim tomando $\epsilon < \mu(A)/4k^2$, podemos supor que $E(1_A|\mathbf{X}_{t'}) > 1 - (1/2k)$ em B onde $B \in \mathcal{B}_{t'}$, com $\mu(B) > 0$. Agora seja $\mu = \int {\mu'}_x d\mu(x)$ a desintegração de μ com respeito ao fator $\mathbf{X}_{t'}$. Então para $x \in B$, ${\mu'}_x(A) > 1 - (1/2k)$.

Usando agora a hipótese de que $\mathbf{X}_{t'}$ é um sistema SZ, temos para o conjunto B com medida positiva

$$\liminf_{N\geq 1} \frac{1}{N}\sum_{n=1}^{N} \mu(B\cap T^{-n}B\cap T^{-2n}B\cap \ldots \cap T^{-kn}B) = \eta > 0. \tag{4.23}$$

Sendo $x \in B\cap T^{-n}B\cap T^{-2n}B\cap \ldots \cap T^{-kn}B$, temos

$$\mu'_x(A\cap T^{-n}A\cap T^{-2n}A\cap \ldots \cap T^{-ln}A)$$

$$\begin{aligned}
&= 1-\mu'_x(X-(A\cap T^{-n}A\cap T^{-2n}A\cap \ldots \cap T^{-kn}A))\\
&= 1-\mu'_x\left(\bigcup_{i=1}^{k}(X-T^{-in}A)\right) \geq 1-\sum_{i=1}^{k}(1-\mu'_x(T^{-n}A))\\
&= 1-\sum_{i=1}^{k}(1-\mu'_{T^{in}x}(A)).
\end{aligned}$$

Usamos a Proposição 2.6 para igualar $T^{in}\mu'x$ com $\mu'_{T^{in}x}$. Como $x \in B\cap T^{-n}B\cap T^{-2n}B\cap \ldots \cap T^{-kn}B$, $T^{in}x \in B$ e $\mu'_{T^{in}x}(A) > 1-(1/2k)$, temos

$$\mu'_x(A\cap T^{-n}A\cap T^{-2n}A\cap \ldots \cap T^{-kn}A) > \frac{1}{2}$$

e pelo Teorema da Desintegração

$$\mu(A\cap T^{-n}A\cap T^{-2n}A\cap \ldots \cap T^{-kn}A) \geq \frac{1}{2}\mu(B\cap T^{-n}B\cap T^{-2n}B\cap \ldots \cap T^{-kn}B)$$

Finalmente este resultado combinado com (4.23) nos fornece

$$\liminf_{N\geq 1} \frac{1}{N}\sum_{n=1}^{N} \mu(A\cap T^{-n}A\cap T^{-2n}A\cap \ldots \cap T^{-kn}A) \geq \frac{\eta}{2} > 0. \quad \blacksquare$$

Mostramos então que no conjunto dos fatores SZ do sistema $(X, \mathcal{B}, \mu, T)$, toda cadeia ordenada possui cota superior. Pelo Lema de Zorn esta cota deve ser um sistema SZ maximal.

4.5 Extensões fracamente mixing

A partir de agora sejam $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ um sistema preservando medida e $\mathbf{Y} = (Y, \mathcal{D}, \nu, S) = (X, \mathcal{B}_l, \mu, T)$ um fator de $\mathbf{X}$. Por conveniência assumimos que $\mathbf{X}$ é um sistema ergódico.

Definição 4.5. *Dizemos que o sistema $\boldsymbol{X}$ é uma extensão fracamente mixing relativa a $\boldsymbol{Y}$ se $\boldsymbol{X}\times_{\boldsymbol{Y}}\boldsymbol{X}$ é um sistema ergódico.*

Mostraremos, sem muitos detalhes, nesta seção, que se uma extensão de um fator SZ é fracamente mixing, então esta extensão também será SZ.

Por simplicidade de notação, escreveremos $\tilde{\mathbf{X}} = \mathbf{X} \times_{\mathbf{Y}} \mathbf{X}$ e da mesma forma $(\tilde{X}, \tilde{\mathcal{B}}, \tilde{\mu}, \tilde{T}) = (X \times X, \mathcal{B} \times \mathcal{B}, \mu \times_{\mathbf{Y}} \mu, T \times T)$.

Lema 4.5. *Seja $(X, \mathcal{B}, \mu, T)$ uma extensão fracamente mixing com relação a $(Y, \mathcal{D}, \nu, S)$ e sejam $f, g \in L^{\infty}(\boldsymbol{X})$. Então*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int [E(fT^n g|\,\boldsymbol{Y})) - E(f|\,\boldsymbol{Y})S^n E(g|\,\boldsymbol{Y})]^2 d\nu = 0. \tag{4.24}$$

Demonstração: Tome $f, g \in L^{\infty}(\mathbf{X})$, temos que $f \otimes f, g \otimes g \in L^2(\mathbf{X} \times_{\mathbf{Y}} \mathbf{X})$. Pela Proposição 2.8,

$$E((g \otimes g)T^n(f \otimes f)|\mathbf{Y}) = E(gT^n f \otimes gT^n f|\mathbf{Y}) = E(gT^n f|\mathbf{Y})^2.$$

Assumimos que $E(f|\mathbf{Y}) = 0$. Precisamos mostrar que

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \int [E(fT^n g|\mathbf{Y})]^2 d\nu = 0 \tag{4.25}$$

Temos que

$$\begin{aligned} \frac{1}{N} \sum_{n=0}^{N} \int [E(fT^n g|\mathbf{Y})]^2 d\nu &= \frac{1}{N} \sum_{n=0}^{N} \int E(f \otimes f\tilde{T}^n g \otimes g|\mathbf{Y}) d\nu \\ &= \frac{1}{N} \sum_{n=0}^{N} \int \left[f \otimes f\tilde{T}^n g \otimes g\right] d\tilde{\mu} \\ &= \int f \otimes f \left(\frac{1}{N} \sum_{n=0}^{N} \tilde{T}^n (g \otimes g) \right) d\tilde{\mu}. \end{aligned}$$

Mas como $\tilde{T}$ é ergódico, temos pelo Teorema Ergódico

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N} \tilde{T}^n (g \otimes g) = \int g \otimes g d\tilde{\mu} = const.$$

e por outro lado

$$\int f \otimes f d\tilde{\mu} = \int E(f|\mathbf{Y})^2 d\nu = 0,$$

e assim temos (4.25).

Agora substituimos f por $f - E(f|\mathbf{Y})$. Teremos $E(f - E(f|\mathbf{Y})) = 0$ e obtemos pela Proposição 2.3

$$E((f - E(f|Y))T^n g|\mathbf{Y}) = E(fT^n g|\mathbf{Y}) - E(f|\mathbf{Y})S^n E(g|\mathbf{Y}). \quad \blacksquare$$

Lema 4.6. *Seja* $(X, \mathcal{B}, \mu, T)$ *uma extensão fracamente mixing relativa a* $\boldsymbol{Y} = (Y, \mathcal{D}, \nu, S)$*. Então o sistema* $(X \times X, \mathcal{B} \times \mathcal{B}, \mu \times_{\mathbf{Y}} \mu, T \times T)$ *também é uma extensão fracamente mixing relativa a* $\boldsymbol{Y}$*.*

Demonstração: Denotaremos $\mathbf{X} \times_{\mathbf{Y}} \mathbf{X}$ por $\tilde{\mathbf{X}}$ e $\tilde{\mathbf{X}} \times_{\mathbf{Y}} \tilde{\mathbf{X}}$ por $\hat{\mathbf{X}}$. Queremos então mostrar que $(\hat{X}, \hat{\mathcal{B}}, \hat{\mu}, \hat{T})$ é ergódico. Para isso é suficiente mostrar que para uma sequência de conjuntos de funções $F, G \in L^2(\hat{X}, \hat{\mathcal{B}}, \hat{\mu})$

$$\frac{1}{N}\sum_{n=1}^{N} \int F\hat{T}^n G d\hat{\mu} \to \int F d\hat{\mu} \int G d\hat{\mu}. \tag{4.26}$$

Por conseguinte é suficiente mostrar (4.26) para F e G da forma

$$F = f_1 f_2 f_3 f_4$$

$$G = g_1 g_2 g_3 g_4.$$

Deste modo, temos então

$$\begin{aligned} \int F\hat{T}^n G d\hat{\mu} &= \int \left[\int f_1 T^n g_1 d\mu_y \int f_2 T^n g_2 d\mu_y \int f_3 T^n g_3 d\mu_y \int f_3 T^n g_3 d\mu_y\right] d\nu \\ &= \int E(f_1 T^n g_1|\mathbf{Y}) E(f_2 T^n g_2|\mathbf{Y}) E(f_3 T^n g_3|\mathbf{Y}) E(f_4 T^n g_4|\mathbf{Y}) d\nu. \end{aligned}$$

Mas agora pelo Lema 4.5, cada expressão $E(f_i T^n g_i|\mathbf{Y})$ pode ser substituida por $E(f_i|\mathbf{Y})S^n E(g_i|\mathbf{Y})$ e assim o lado esquerdo de (4.26) pode ser substituido por

$$\frac{1}{N}\sum_{n=1}^{N} \int E(f_1|\mathbf{Y})E(f_2|\mathbf{Y})E(f_3|\mathbf{Y})E(f_4|\mathbf{Y}) \cdot S^n \left[E(g_1|\mathbf{Y})E(g_2|\mathbf{Y})E(g_3|\mathbf{Y})E(g_4|\mathbf{Y})\right] d\nu.$$

Como $(Y, \mathcal{D}, \nu, S)$ é ergódico, temos

$$\frac{1}{N}\sum_{n=1}^{N} S^n \left[E(g_1|\mathbf{Y})E(g_2|\mathbf{Y})E(g_3|\mathbf{Y})E(g_4|\mathbf{Y})\right] \to \int \left[\int g_1 d\mu_y \int g_2 d\mu_y \int g_3 d\mu_y \int g_4 d\mu_y\right] d\nu = \int G d\hat{\mu},$$

e segue (4.26). ■

Proposição 4.7. *Seja $(X,\mathcal{B},\mu,T)$ uma extensão fracamente mixing de $(Y,\mathcal{D},\nu,S)$. Então se $f_l \in L^\infty(X,\mathcal{B},\mu)$, $l=0,1,\ldots,k$, temos*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\int\left[E\left(\prod_{l=0}^{k}T^{ln}f_l|\,\mathbf{Y}\right)-\prod_{l=0}^{k}S^{ln}E(f_l|\,\mathbf{Y})\right]^2 d\nu=0, \tag{4.27}$$

e

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=1}^{N}\left(\prod_{l=0}^{k}T^{ln}f_i-\prod_{l=0}^{k}T^{ln}E(f_l|\,\mathbf{Y})\right)\right\|_{L^2(\mu)}=0. \tag{4.28}$$

Note que $E(f_l|Y)$ é uma função em $L^\infty(X,\mathcal{B},\mu)$ e em $L^\infty(Y,\mathcal{D},\nu)$.

Demonstração: A demonstração desta proposição é semelhante à demonstração da Proposição 4.2. Procederemos por indução em k. Para $k=1$, o Lema 4.5 nos dá (4.27) enquanto (4.28) segue do Teorema Ergódico. Assumimos que (4.27) é válido para $k-1$ em total generalidade, que é, para todas extensões fracamente mixing de $\mathbf{Y}$ (em paticular, para $X\times_{\mathbf{Y}}X$) e provaremos

(i) (4.27) válido para $k-1 \;\Rightarrow\;$ (4.28) válido para k,

(ii) (4.28) válido para k (para $X\times_{\mathbf{Y}}X$) $\;\Rightarrow\;$ (4.27) válido para k (para X).

Começaremos por *(ii)*. Se f_0 é mensurável em $\mathcal{B}_1=\alpha^{-1}(\mathcal{D})$, a integral em (4.27) tem a forma

$$\int {f_0}^2\left[E\left(\prod_{l=1}^{k}T^{ln}f_l|\mathbf{Y}\right)-\prod_{l=1}^{k}S^{ln}E(f_l|\mathbf{Y})\right]^2 d\nu$$

$$\leq \sup|{f_0}^2|\int\left[E\left(\prod_{l=0}^{k-1}T^{ln}f_{l+1}|\mathbf{Y}\right)-\prod_{l=0}^{k-1}S^{ln}E(f_{l+1}|\mathbf{Y})\right]^2 d\nu,$$

usando a propriedade *(iv)* da Proposição 2.3 e a Proposição 2.5 combinado com o fato de S^n preservar medida. Temos que (4.27) é reduzida ao caso $k-1$. Isto nos permite assumir, como fizemos no Lema 4.5, que $E(f_0|\mathbf{Y})=0$. Com isto, (4.27) para o caso $k=1$ tem a forma

$$\lim_{N\to\infty}\int f_0\otimes f_0\left(\frac{1}{N}\sum_{n=1}^{N}\prod_{l=1}^{k}\tilde{T}^{ln}f_l\otimes f_l\right)d\tilde{\mu}. \tag{4.29}$$

Por (4.28) aplicado a $\mathbf{X} \times_{\mathbf{Y}} \mathbf{X}$, o limite em (4.29) é o memos que

$$\lim_{N\to\infty} \int f_0 \otimes f_0 \left(\frac{1}{N} \sum_{n=1}^{N} \prod_{l=1}^{k} \tilde{T}^{ln} E(f_l \otimes f_l | \mathbf{Y}) \right) d\tilde{\mu}.$$

que é 0 (para todo N) pois a soma acima é constante pelo Teorema Ergódico e $E(f_0|\mathbf{Y}) = 0$ (Como feito no Lema 4.5).

Para provar *(i)*, lembramos da equação (4.15)

Então podemos escrever

$$\prod_{l=1}^{k} T^{ln} f_l - \prod_{l=1}^{k} T^{ln} E(f_l|\mathbf{Y}) = \sum_{j=1}^{k} \left(\prod_{l=1}^{j-1} T^{ln} f_l \right) T^{jn} \left((f_j - E(f_j|\mathbf{Y})) \prod_{l=j+1}^{k} T^{ln} E(f_l|\mathbf{Y}) \right)$$

e note que isto nos ajuda a provar (4.28) sobre a condição adicional que par algum l_0, $1 \le l_0 \le k$, $E(f_{l_0}|\mathbf{Y}) = 0$. agora temos que mostrar que sobre esta condição, $\lim \|\psi_N\|_2 = 0$ quando $\psi_N = (1/N) \sum_{n=1}^{N} \prod_{l=1}^{K} T^{ln} f_l$.

Reescreva

$$\psi_N = \frac{1}{N} \sum_{j=1}^{N} \left(\frac{1}{H} \sum_{n=j}^{j+H-1} \prod_{l=1}^{k} T^{ln} f_l \right) + O(H/N),$$

onde H será escolhido grande mas necessariamente menor do que N. Pela convexidade da função $\phi(x) = x^2$, temos

$${\psi_N}^2 \le \frac{1}{N} \sum_{j=1}^{N} \left(\frac{1}{H} \sum_{n=j}^{j+H-1} \prod_{l=1}^{K} T^{ln} f_l \right)^2.$$

Por integração e do fato de T preservar medida,

$$\begin{aligned} \|\psi_N\|_2^2 &\le \frac{1}{N} \sum_{j=1}^{N} \frac{1}{H^2} \sum_{n,m=1}^{j+H-1} \int \prod_{l=1}^{k} T^{ln} f_l T^{lm} f_l d\mu \qquad (4.30) \\ &= \frac{1}{(N)H^2} \sum_{j=1}^{N} \sum_{n,m=j}^{j+H-1} \int \prod_{l=1}^{k} T^{(l-1)n} (f_l T^{l(m-n)} f_l) d\mu. \end{aligned}$$

Como feito na demonstração da Proposição 4.2, reescrevemos (4.31) como

$$\|\psi_N\|_2^2 \le \frac{1}{H} \sum_{r=1-H}^{H-1} \left(1 - \frac{|r|}{H}\right) \left[\frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=1}^{k} T^{(l-1)n} f_l T^{lr} f_l d\mu \right] + O(H/N). \qquad (4.31)$$

Por (4.27) para $k-1$ e para um H fixo e todo r tal que $|r| < H$ nós podemos substituir o termo integrável em (4.31), pondo N suficientemente grande, por $\prod_{l=1}^{k} T^{(l-1)n} E(f_l T^{lr} f_l | \mathbf{Y})$ e obtemos

$$\|\psi_N\|_2^2 \leq \frac{1}{H} \sum_{r=1-H}^{H-1} \left(1 - \frac{|r|}{H}\right) \left[\frac{1}{N} \sum_{n=1}^{N} \int \prod_{l=1}^{k} T^{(l-1)n} E(f_l T^{lr} f_l | \mathbf{Y}) d\mu\right] + O(H/N). \tag{4.32}$$

Agora estimando a integral dada, aparecerá em (4.32)

$$\left\|E(f_{l_0} T^{l_0 r} f_{l_0} | \mathbf{Y})\right\|_2 \prod_{l \neq l_0} \|f_l\|_\infty^2$$

e lembrando que $E(f_{l_0}|\mathbf{Y}) = 0$ temos pelo Lema 4.5 que a maior parte dos termos em (4.32) são pequenos contanto que H seja suficientemente grande. Como todos os termos são limitados por $\prod \|f_l\|_\infty^2$ e que a maior parte dos termo são pequenos, sua média dada por (4.32) é pequena, arbitrariamente para H e N grandes. ∎

Proposição 4.8. *Seja $(X, \mathcal{B}, \mu, T)$ uma extensão fracamente mixing relativa a $(Y, \mathcal{D}, \nu, S)$. Se $(Y, \mathcal{D}, \nu, S)$ é um sistema SZ, então $(X, \mathcal{B}, \mu, T)$ também é um sistema SZ.*

Demonstração: Seja $A \in \mathcal{B}$, $\mu(A) > 0$. Seja $a > 0$ o menor número tal que para $A_1 = \{y : E(1_A|\mathbf{Y}) \geq a\}$, tenhamos $\nu(A_1) > 0$. Pelo Teorema da desintegração e pela Proposição 4.7 (e $E(1_A|\mathbf{Y}) \geq a 1_{A_1}$) temos para todo k

$$\frac{1}{N} \sum_{n=1}^{N} \mu\left(\bigcap_{l=0}^{k} T^{-ln} A\right) > \frac{1}{2} a^{k+1} \cdot \frac{1}{N} \sum_{n=1}^{N} \nu\left(\bigcap_{l=0}^{k} S^{-ln} A_1\right),$$

contanto que N seja suficientemente grande. ∎

Notamos que se tomarmos para o sistema $(X, \mathcal{B}, \mu, T)$ o fator $(Y, \mathcal{D}, \nu, S) \approx (X, \{X, \emptyset\}, \mu, T)$ teremos a igualdade $\mu \times_{\mathbf{Y}} \mu = \mu \times \mu$. Obtemos então o seguinte resultado.

Corolário 4.1. *Se $(X, \mathcal{B}, \mu, T)$ é um sistema fracamente mixing, então este sistema é SZ.*

4.6 Extensões compactas

Podemos também definir uma extensão compacta e assim provar que o Terorema de Recorrência Multipla de Furstenberg também é valido para sistemas "relativamente" compactos.

Definição 4.6. *Uma função* $f \in L^2(X, \mathcal{B}, \mu)$ *é quase periódica (AP) relativa a um fator* $\mathbf{Y}$ *se para todo* $\delta > 0$ *existem funções* $g_1, g_2, \ldots, g_k \in L^2(X, \mathcal{B}, \mu)$ *tais que para todo* $n \in \mathbb{Z}$, $\min_{1 \leq j \leq k} \|T^n f - g_j\|_{L^2(\mu_y)} < \delta$ *p.q.t.* $y \in Y$.

Quando não houver confusão denotaremos por AP ao conjunto das funções AP em $L^2(X, \mathcal{B}, \mu)$.

Definição 4.7. *Dizemos que a extensão* $(X, \mathcal{B}, \mu, T)$ *é uma extensão compacta de* $(Y, \mathcal{D}, \nu, S)$ *se o conjunto AP é denso em* $L^2(X, \mathcal{B}, \mu)$.

Proposição 4.9. *Se* $(X, \mathcal{B}, \mu, T)$ *é uma extensão compacta de* $(Y, \mathcal{D}, \nu, S)$ *que por vez é um sistema SZ, então* $(X, \mathcal{B}, \mu, T)$ *é um sistema SZ.*

Deja-se provar para algum $A \in \mathcal{B}$, com $\mu(A) > 0$ e k inteiro, que

$$\liminf_{N \to \infty} \frac{1}{N} \sum_{j=1}^{N} \mu \left(\bigcap_{l=1}^{k} T^{-jl} A \right) > 0 \tag{4.33}$$

que claramente continua válido para um subconjunto de A. Assumindo que $\mu_y(A) \geq \lambda = \frac{1}{2}\mu(A)$ para $y \in A_1$, $\nu(A) > \frac{1}{2}\mu(A)$, e $\mu_y(A) = 0$ para $y \notin A_1$.

É possível remover de A menos que a metade de sua medida e obter um conjunto cuja função característica é AP.

Considerando vetores da forma $(f, T^n f, T^{2n} f, \ldots, T^{kn} f)$ nas fibras de $y \in Y$, e denotando por $\oplus_{l=0}^{k} L^2(X, \mathcal{B}, \mu_y)$ a soma direta de $k+1$ copias de $L^2(X, \mathcal{B}, \mu_y)$ dotada da norma $\|(f_0, f_1, \ldots, f_k)\|_y = \max \|f_i\|_{L^2(X, \mathcal{B}, \mu_y)}$, se $f \in$ AP, o conjunto $\mathfrak{L}(k, f) = \{(f, T^n f, \ldots, T^{kn} f)\}_{n \in \mathbb{Z}}$ será totalmente limitado em $\oplus_{l=0}^{k} L^2(X, \mathcal{B}, \mu_y)$ para quase todo $y \in Y$, uniformemente em $y \in Y$.

Assumindo que $f = 1_A$ é AP e sustentand a notação $A_1 = \{y \in Y; \mu_y(A) > \frac{1}{2}\mu(A)\}$, seja $k > 0$ dado, escreve-se

$$\mathfrak{L}(k, f, y) = \left\{(f, T^n f, \ldots, T^{kn} f)_y\right\}_{n \in \mathbb{Z}} \subset \oplus_{l=0}^{k} L^2(X, \mathcal{B}, \mu_y).$$

$\mathfrak{L}(k, f, y)$ é uniformente totalmente limitado em y. Para $y \in A_1$, nos elementos de $\mathfrak{L}(k, f, y)$ para os quais todas as componentes são não nulas, denota-se este subconjunto de $\mathfrak{L}(k, f, y)$ por $\mathfrak{L}^*(k, f, y)$. Para $y \in A_1$ e $\epsilon > 0$ seja $M(\epsilon, y)$ denotando a cardinalidade máxima de um subconjunto ϵ-separado de $\mathfrak{L}^*(k,f,y)$. $M(\epsilon,y)$ é limitado em A_1. Como uma função de y, $M(\epsilon,y)$ é mensurável e pode-se encontrar algum $\epsilon_0 < \mu(A)/10k, \eta > 0$, e $A_2 \subset A_1$, $\nu(A_2) > 0$, em que $M(\epsilon,y)$ é constante, digamos M, para $\epsilon_0 - \eta \leq \epsilon \leq \epsilon_0$ e $y \in A_2$.

Por fim é possível tomar A_3 o subconjunto de A_2 dos pontos y para os quais, para cada i,j,l com $1 \leq i < j \leq M, 0 \leq l \leq k$,

$$\left\|T^{lm_i}f - T^{lm_j}f\right\|_{L^2(X,\mathcal{B},\mu_y)} > \left\|T^{lm_i}f - T^{lm_j}f\right\|_{L^2(X,\mathcal{B},\mu_{y_0})} - \eta. \tag{4.34}$$

E este será um subconjunto com $\nu(A_3) > 0$.

Usando a hipótese de que o sistema $(Y,\mathcal{D},\nu,S)$ é SZ e aplicando isto a A_3, têm-se a Proposição.

4.7 Existência de extensões compactas

Nesta seção enunciamos uma proposição que garante a existencia de extensões compactas intermediárias de um fator de uma extensão não relativamente mixing. Daremos apenas um roteiro da prova feita por Furstenberg em [5].

Proposição 4.10. *Se $\boldsymbol{X} = (X,\mathcal{B},\mu,T)$ é uma extensão de $\boldsymbol{Y} = (Y,\mathcal{D},\nu,S)$ que não é fracamente mixing, então existe um fator intermediário $\boldsymbol{Z}$ entre $\boldsymbol{Y}$ e $\boldsymbol{X}$ onde $\boldsymbol{Z}$ é uma extensão compacta de $\boldsymbol{Y}$.*

A prova dá-se supondo $\mathbf{X}$ ergódico. Como $\tilde{T}$ em $\mathbf{X} \times_{\mathbf{Y}} \mathbf{X}$ não é ergódico, existe uma função $H(x,x')$ em $\mathbf{X} \times_{\mathbf{Y}} \mathbf{X}$ que é invariante sobre $\tilde{T}$, mas não é uma função somente de x nem somente de x'. Existe uma função $\phi \in L^2(X,\mathcal{B},\mu)$, em que a convolução $H * \phi$ definida por

$$H * \phi(x) = \int H(x,x')\phi(x')d\mu_{\alpha(x)}(x)$$

não é uma função totalmente em $L^2(\mathbf{Z})^\alpha$. Encontra-se funções $\phi \in L^2(\mathbf{X})$, $f \notin L^2(\mathbf{Y})^\alpha$ com a propriedade de que para todo $\delta > 0$ existe um conjunto finito de funções $g_1, g_2, \ldots, g_k \in L^2(\mathbf{X})$ tal que para cada n,

$$\min_{1<j<k} \left\|T^n\phi - g_j\right\|_y < \delta,$$

para cada $y \in \mathbf{Y}$. Dizendo que $\phi \in L^2(\mathbf{X})$ é *compacta* se para cada $\epsilon, \delta > 0$ existir $g_1, g_2, \ldots, g_k \in L^2(\mathbf{X})$ tal que para cada n

$$\min_{1 \leq j \leq k} \|T^n\phi - g_j\|_y < \delta,$$

para cada $y \in \mathbf{Y}$ exceto para um conjunto de y com medida $< \epsilon$. Têm-se que, se $\mathcal{C} = \{A \in \mathcal{B} : 1_A$ é compacto $\}$, então $\mathcal{C}$ é uma σ-álgebra e o espaço das funções compactas de $L^2(X, \mathcal{B}, \mu)$ consiste do espaço de funções mensuráveis com respeito a $\mathcal{C}$ que é T-invariante. Daí, o conjunto $\mathbf{Z} = (X, \mathcal{C}, \mu, T)$ é um fator de $\mathbf{X}$ que é compacto. Ainda $\mathcal{C} \subset \alpha^{-1}(\mathcal{D})$ e $\mathbf{Y}$ é fator de $\mathbf{Z}$. E prova-se que $\mathbf{Y} \to \mathbf{Z}$ é relativamente compacto.

4.8 Conclusão

Seja $(X, \mathcal{B}, \mu, T)$ um espaço de probabilidade. Se este é um sistema fracamente mixing, então pela Proposição 4.1 ele é um sistema SZ. Por outro lado se o sistema $(X, \mathcal{B}, \mu, T)$ não é fracamente mixing a Proposição 4.6 assegura que existe uma σ-álgebra maximal em que o sistema com esta σ-álgebra é um sistema SZ. Se esta, por sua vez, é igual a $\mathcal{B}$ a prova está completa. Caso isto não aconteça, a Proposição 4.4 garante que esta é uma sub-σ-álgebra não trivial de $\mathcal{B}$, e podemos assumir que o sistema não é uma extensão fracamente mixing deste fator, pois do contrário pela Proposição 4.8 teríamos que o sistema é SZ. Mas, supondo que o sistema não é uma extensão fracamente mixing do fator maximal SZ a Proposição 4.10 garante a existência de uma extensão intermediária não trivial que é relativamente compacta, que por sua vez, pela Proposição 4.9 é um sistema SZ contrariando a maximalidade do fator.

Proposição 4.2

Conhecida a desigualdade

$$\left\| \psi_N' \right\|^2_{L^2(X)} \leq \frac{1}{NH^2} \sum_{j=1}^{N} \sum_{n,m=j}^{\min\{j+H-1,N\}} \int \prod_{l=0}^{k-1} T^{ln} \left(f_{l+1} T^{(l+1)(m-n)} f_{l+1} \right) d\mu, \tag{35}$$

olharemos para os índices das funções $F_{n,r} = \int \prod_{l=0}^{k-1} T^{ln} g_{l,r}$, onde $r = m - n$, para estudar a presença de tais funções no somatório acima. Temos que $m, n \in \{1, 2, \ldots, N\}$ e observamos as seguintes matrizes de índices, para cada $j \in \{1, 2, \ldots, N\}$:

$j = 1$;

...	1	2	3	$\cdots$	H
1	(1,0)	(1,1)	(1,2)	$\cdots$	(1,H-1)
2	(2,-1)	(2,0)	(2,1)	$\cdots$	(2,H-2)
3	(3,-2)	(3,-1)	(3,0)	$\cdots$	(3,H-3)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
H-1	(H-1,2-H)	(H-1,3-H)	(H-1,4-H)	$\cdots$	(H-1,1)
H	(H,1-H)	(H,2-H)	(H,3-H)	$\cdots$	(H,0)

$j = 2$;

...	2	3	3	$\cdots$	H+1
2	(2,0)	(2,1)	(2,2)	$\cdots$	(2,H-1)
3	(3,-1)	(3,0)	(3,1)	$\cdots$	(3,H-2)
4	(4,-2)	(4,-1)	(4,0)	$\cdots$	(4,H-3)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
H	(H,2-H)	(H,3-H)	(H,4-H)	$\cdots$	(H,1)
H+1	(H+1,1-H)	(H+1,2-H)	(H+1,3-H)	$\cdots$	(H+1,0)

$\vdots \quad \vdots$

$j = N;$

...	N
N	(N,0)

Fazendo então a soma no índice r, encontramos:

$$
\begin{aligned}
r = 1 - H \quad &\rightarrow \quad F_{H,1-H} + F_{H+1,1-H} + F_{H+2,1-H} + \ldots + F_{N,1-H} \\
&\vdots \\
r = -1 \quad &\rightarrow \quad F_{1,-1} + 2F_{2,-1} + 3F_{3,-1} + \ldots (H-1)F_{H-1,-1} + \ldots + (H-1)F_{N,-1} \\
r = 0 \quad &\rightarrow \quad F_{1,0} + 2F_{2,0} + HF_{H,0} + \ldots + HF_{N,0} \\
r = 1 \quad &\rightarrow \quad F_{1,1} + 2F_{2,1} + 3F_{3,1} + (H-1)F_{H-1,1} + \ldots + (H-1)F_{N-1,1} \\
&\vdots \\
r = H - 1 \quad &\rightarrow \quad F_{1,H-1} + F_{2,H-1} + F_{3,H-1} + \ldots + F_{4,H-1}
\end{aligned}
$$

Assim, lembrando que as funções $f_0, f_1, ldots, f_k \in L^{\infty}(X, \mathcal{B}, \mu)$ são não negativas, temos que,

$$
\begin{aligned}
&\frac{1}{NH^2} \sum_{j=1}^{N} \sum_{n,m=j}^{\min\{j+H-1,N\}} \int \prod_{l=0}^{k-1} T^{ln} \left(f_{l+1} T^{(l+1)(m-n)} f_{l+1} \right) d\mu \\
&\leq \frac{1}{NH^2} \sum_{r=1-H}^{H-1} (1 - |r|) \left(\sum_{n=1}^{N} \int \prod_{l=0}^{k-1} T^{ln} g_{l,r} d\mu \right).
\end{aligned}
$$

Referências Bibliográficas

[1] A. A. Castro. *Curso da Teoria da Medida.* Projeto Euclides, IMPA, Rio de Janeiro, 2004.

[2] P. Erdös and P. Turán. On some sequences of integers. *J. London Math. Soc.*, 11:261–264, 1936.

[3] R. Fernandes. *Introdução à Teoria da Medida.* Projeto Euclides, IMPA, Rio de Janeiro, 2002.

[4] H. Furstenberg. Ergodic bhavior of diagonal measures and a theorem of szemerédi on arithmetic progressions. *J. d'Analyse Math.*, 31:204–256, 1977.

[5] H. Furstenberg. *Ergodic behavior and a theorem and combinatorial number theory.* Princeton University Press, New Jersey, 1981.

[6] Y. Katznelson H. Furstenberg and D. Ornstein. The ergodic theoretical proof of Szemerédi's theorem. *Bull. Amer. Math. Soc.*, 7(3):527–552, 1982.

[7] M. Viana K. Oliveira. *Introdução à Teoria Ergódica.* Preprint, IMPA, Rio de Janeiro, 2000.

[8] E. L. Lima. *Espaços Métricos.* Projeto Euclides, IMPA, Rio de Janeiro, 2003.

[9] M. Yuri M. Pollicott. *Dynamical Systems and Ergodic Theory.* Cambridge University Press, UK, 1998.

[10] R. Mañé. *Teoria Ergódica.* Projeto Euclides, IMPA, Rio de Janeiro, 1983.

[11] K. Oliveira. *Um primeiro curso sobre Teoria Ergódica com aplicações.* Publicaçoes Matemáticas, IMPA, Rio de Janeiro, 2005.

[12] F. Riesz and B. Nagy. *Function analysis*. Ungar, New York, 1955.

[13] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 3 edition, 1987.

[14] S. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.

[15] P. Walters. *An Introduction to Ergodic Theory*. Springer, Berlin, 1983.

Baixar livros de Literatura
Baixar livros de Literatura de Cordel
Baixar livros de Literatura Infantil
Baixar livros de Matemática
Baixar livros de Medicina
Baixar livros de Medicina Veterinária
Baixar livros de Meio Ambiente
Baixar livros de Meteorologia
Baixar Monografias e TCC
Baixar livros Multidisciplinar
Baixar livros de Música
Baixar livros de Psicologia
Baixar livros de Química
Baixar livros de Saúde Coletiva
Baixar livros de Serviço Social
Baixar livros de Sociologia
Baixar livros de Teologia
Baixar livros de Trabalho
Baixar livros de Turismo