UNIVERSIDADE FEDERAL DE ALAGOAS
INSTITUTO DE MATEMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA
DISSERTAÇÃO DE MESTRADO

Demonstração do Teorema de van der Waerden Via Sistemas Dinâmicos

Vanessa Lúcia da Silva

Maceió, Brasil
Fevereiro de 2016

VANESSA LÚCIA DA SILVA

DEMONSTRAÇÃO DO TEOREMA DE VAN DER WAERDEN VIA SISTEMAS DINÂMICOS

Dissertação de Mestrado, na área de concentração de Sistemas Dinêmicos, submetida em 12 de Fevereiro de 2016 à banca examinadora, designada pelo Programa de Mestrado em Matemática da Universidade Federal de Alagoas, como parte dos requisitos necessários à obtenção do grau de mestre em Matemática.

Orientador: Prof. Dr. Krerley Irraciel Martins Oliveira.

Maceió, Brasil
Fevereiro de 2016

Catalogação na fonte
Universidade Federal de Alagoas
Biblioteca Central
Divisão de Tratamento Técnico
Bibliotecário Responsável: Valter dos Santos Andrade

S586 Silva, Vanessa Lúcia da.
Demonstração do teorema de van Waerden via Sistemas dinâmicos / Vanessa Lúcia da Silva. - 2016.
32 f. : il.

Orientador: Krerley Irraciel Martins Oliveira.
Dissertação (Mestrado em Matemática) – Universidade Federal de Alagoas. Instituto de Matemática. Programa de Pós-Graduação em Matemática. Maceió, 2016.

Bibliografia: f. 32.

1. Partição finita. 2. Progressão aritmética. 3. Teorema de van der Waerden. I. Título.

CDU: 517.93

VANESSA LÚCIA DA SILVA

Demonstração do Teorema de van der Waerden via Sistemas Dinâmicos

Dissertação submetida ao curso de Mestrado do Programa de Pós-Graduação em Matemática da Universidade Federal de Alagoas e aprovada em 12 de fevereiro de 2016.

Banca Examinadora:

Prof. Dr. Krerley Irraciel Martins Oliveira – UFAL
(Orientador)

Prof. Dr. Ali Golmakani – UFAL

Prof. Dr. Mohammad Fanaee – UFF

"Até aqui nos ajudou o Senhor"
(Ebenézer)

Agradecimentos

Agradeço primeiramente a minha família, que sempre me apoiou ao longo desta jornada em especial para minha mãe Vera Lúcia, minha irmã Vânia Lúcia e meu irmão Ivan José.

Agradeço também aos meus amigos Eduardo Santana, Isnaldo Isaac, Dayane Dalysse, Joselma Lins e Robson Silva pelos incentivos e auxílios ao trabalho.

Por fim, quero agradecer ao meu orientador, o professor Krerley Oliveira, especialmente por sua atenção e grande paciência, mesmo diante das minhas faltas e de seus múltiplos compromissos. Aproveito ainda para agradecer aos professores Ali Golmakani e Mohammad Fanaee, membros da banca examinadora, pelas sugestões e apoio a respeito desta dissertação.

Resumo

Vamos apresentar uma demonstração alternativa ao âmbito da Aritmética Combinatória, do Teorema van der Waerden. Este resultado, consiste na busca por progressões aritméticas de tamanhos arbitrários dentro do conjunto dos números inteiros. Mais precisamente, mostraremos que ao particionarmos finitamente o conjunto dos números inteiros, uma dessas partes finitas de $\mathbb{Z}$, possui progressões aritméticas de qualquer comprimento. Para tal, faremos uso das ideias de Fustemberg com elementos da Dinâmica Simbólica.

Palavras-chave: Partição finita. Progressão Aritmética. Teorema de van der Waerden.

Abstract

We will present an alternative demonstration the scope of Arithmetic Combinatorics, Theorem van der Waerden. This result is the search for arithmetic progressions of arbitrary size within the set of integers. More precisely, we show the particionarmos the finitely set of integers, one of these parts of finite $\mathbb{Z}$, has arithmetic progressions of any length. To this end, we will make use of Fustemberg ideas with elements of symbolic dynamics.

Keywords: Finite Partition. Arithmetic Progression. Van der Waerden Theorem.

Sumário

Introdução

A Teoria dos Números vem ganhando destaque ao longo dos tempos com seus diversos problemas que desafiam gerações de matemáticos. Essa área da matemática estuda as propriedades dos números, dando maior atenção aos inteiros, com ênfase aos primos.
As engrenagens dessa ciência são questionamentos tais como o surgido em 300 a.C., cuja conjectura fora atribuída a Euclides de Alexandria , ainda em aberto, conhecida como conjectura dos primos gêmeos, segundo a qual existem infinitos pares de primos tais que p é primo e $p + 2$ também o é. Outra conjectura ,também em aberto, é a de Goldbach, o qual afirma que todo número inteiro par maior ou igual a 4 é a soma de dois primos.
Mediante estes questionamentos, daremos ênfase a uma conjectura respondida por Van der Waerden, atualmente classificada como Teorema de Van der Waerden que trata da existência de progressões aritméticas de tamanhos arbitrários em subconjuntos de $\mathbb{Z}$ e é a esse resultado que atentaremos neste trabalho. Tal problema fora obtido originalmente numa versão segundo elementos da teoria combinatória e, posteriormente, fora apresentado e resolvido por Fustemberg fazendo uso de elementos da Teoria Ergódica.
O teorema de Van der Waerden foi obtido em 1927 pelo matemático Bartel Leendert Van der Waerden (1903- 1996). Van der Waerden é oriundo dos Países Baixos, popularmente conhecido como Holanda, estudou na Universidade de Amsterdam e de Gottingen no período de 1919 a 1925. Apesar de ser mais conhecido no ramo da álgebra,também publicou trabalhos em geometria algébrica, topologia, teoria dos números, geometria, análise combinatória, análise matemática, e teoria da probabilidade.
Seguindo a mesma linha do problema proposto e resolvido por Van der Waerden, encontramos resultados posteriores mais gerais tais como: o teorema de Szemerédi (1975) o qual afirma que todo conjunto de inteiros positivos com densidade superior positiva possui progressões aritméticas de tamanhos arbitrários. Em 2004 fora demonstrada, por Ben Green e Terence Tao, uma antiga conjectura sobre os primos. Fora mostrado que existem progressões aritméticas arbitrariamente longas formada apenas por primos. Esse resultado, juntamente com outros, levaram Tao a ganhar medalha Fields em 2006.
Desta forma, observa-se a relevante busca por números primos através das progressões aritméticas.
Então, iniciaremos este trabalho identificando conceitos e teoremas necessários para demonstrar o resultado devido a van der Waerden, sobre progressões aritméticas. Posteriormente, serão abordadas outras formas do Teorema de van der Waerden e em sequência resultados mais gerais ao mesmo.

Capítulo 1

Preliminares

1.1 Lema de Zorn

Para concluirmos a demonstração da versão simples do Teorema de Recorrência de Birkhoff (Seção 2.1), faremos uso de um importante lema de caráter existencial, a qual trata de famílias de infinitos conjuntos, é o denominado Lema de Zorn, devido a Max Zorn.

Definição 1.1 *Uma ordenação parcial num conjunto M é uma relação binária $\prec$ em M que é reflexiva ($x \prec x, \forall x \in M$), transitiva ($x \prec y$ e $y \prec z \Rightarrow x \prec z$) e anti simétrica ($x \prec y$ e $y \prec x \Rightarrow x = y$). Neste caso, dizemos que $(M, \prec)$ é um conjunto parcialmente ordenado.*

O termo "parcial"na definição dada aparece pois pode haver elementos que não são comparáveis de acordo com a ordenação $\prec$ dada.

Exemplo 1 *Denotando $\mathcal{P}(X)$ o conjunto das partes de X. A inclusão de conjuntos, isto é, $A \prec B$ se, e somente se, $A \subset B$, define uma ordenação parcial em $\mathcal{P}(X)$.*

Definição 1.2 *Um conjunto parcialmente ordenado é dito totalmente ordenado quando quaisquer dois elementos são comparáveis de acordo com a ordenação parcial dada.*

Exemplo 2 *O conjunto $\mathbb{R}$ dos números reais com a ordenação usual $\leq$ é totalmente ordenado.*

Definição 1.3 *Seja $(M, \prec)$ um conjunto parcialmente ordenado. $x \in M$ é um elemento maximal (minimal) em M se para todo $y \in M$ com $x \prec y$ ($y \prec x$) segue-se $x = y$. Um elemento $x \in X$ é majorante (minorante) de $Y \subset X$ se $y \prec x$ ($x \prec y$) para todo $y \in Y$.*

Exemplo 3 *O conjunto $\mathcal{P}(X)$, do exemplo 1.1, possui como elemento maximal o próprio X. Enquanto no Exemplo 1.2, $\mathbb{R}$ não possui elemento maximal, nem minimal.*

Lema 1.1 (Lema de Zorn) *Um conjunto não vazio parcialmente ordenado, no qual todo subconjunto totalmente ordenado possui um majorante (minorante), possui elemento maximal (minimal).*

Este lema é muito útil, pois substitui cálculos técnicos envolvendo indução transfinita, que trata de técnica matemática que permite provar propriedades para todos números ordinais (ou, de forma mais geral, para qualquer conjunto (ou classe) bem ordenado) a partir de etapas finitas. Trata de uma generalização da indução finita. Além disso, o mesmo dá suporte a demonstração de importantes resultados tais como, para concluir que todo espaço vetorial não trivial possui uma base de Hamel, que todo conjunto infinito M é equipotente ao conjunto $M \times M$, dentre outros, cujas demonstrações não serão expostas pois fogem ao objetivo do trabalho.

1.2 Espaço de Sequências

Em matemática, dinâmica simbólica é a prática de modelagem de um sistema dinâmico topológico por um espaço discreto que consiste em sequências infinitas de símbolos abstratos, cada um dos quais corresponde a um estado do sistema, com a dinâmica dadas pelo operador de deslocamento a esquerda, o shift.

Considere, para cada $l \geq 2$ o espaço

$$\Sigma = \{1, 2, ..., l\}^{\mathbb{Z}} = \{\underline{\alpha} = (..., \alpha_{-1}, \alpha_0, \alpha_1, ...) : \alpha_i \in \{1, 2, ..., l\}, i \in \mathbb{Z}\}$$

o espaço das sequências bilaterais com l símbolos.

Definição 1.4 *Sejam $(X, \mathcal{T}_X)$ e $(Y, \mathcal{T}_Y)$ espaços topológicos. A topologia $\mathcal{T}$ em $X \times Y$ gerada pela base*

$$\mathcal{B} = \{U \times V; U \in \mathcal{T}_X, V \in \mathcal{T}_Y\}$$

chama-se topologia produto de $\mathcal{T}_X$ e $\mathcal{T}_Y$. O espaço topológico $(X \times Y, \mathcal{T})$ chama-se espaço produto.

Teorema 1.1 (Tychonof) *Seja $\{X_\alpha\}$ uma família de espaços topológicos compactos. Então, $X = \prod X_\alpha$ é compacto.*

Definiremos uma topologia, observando que Σ é o produto direto, de uma quantidade enumerável de cópias do conjunto finito $\{1, 2, ..., l\}$, com a topologia discreta, e usando a topologia produto. Desta forma, segue-se pelo Teorema de Tychonoff que Σ é compacto.

Fixando inteiros $n_1 < n_2 < ... < n_k$ e números $\alpha_1, ..., \alpha_k \in \{1, 2, ..., l\}$ chamaremos o subconjunto de Σ de cilindro o conjunto dado por:

$$C^{n_1,...,n_k}_{\alpha_1,...,\alpha_k} = \left\{\underline{\beta} \in \Sigma : \beta_{n_i} = \alpha_i \forall i = 1, ..., k\right\}.$$

O número k de dígitos fixados é chamado de categoria do cilindro.

Uma forma alternativa de definirmos uma topologia no espaço Σ e observando que todos os cilindros são conjuntos abertos e que formam uma base da topologia. Então, todo cilindro é fechado, pois o complementar de um cilindro é a união finita de cilindros.

Podemos também introduzir, em Σ, a métrica $d(\underline{\beta}, \underline{\gamma}) = 2^{-N(\underline{\beta},\underline{\gamma})}$ em Σ, onde

$$N(\underline{\beta}, \underline{\gamma}) = \max\{N \geq 0 : \beta_n = \gamma_n, \forall n \in \mathbb{Z}, |n| < N\}$$

. De fato,

- $d(\underline{\beta}, \underline{\beta}) = 0$, pois $N(\underline{\beta}, \underline{\beta}) = \infty$;
- $d(\underline{\beta}, \underline{\gamma}) = d(\underline{\gamma}, \underline{\beta})$ visto que $N(\underline{\beta}, \underline{\gamma}) = N(\underline{\gamma}, \underline{\beta})$;
- Dados $\underline{\alpha}, \underline{\beta}, \underline{\gamma} \in \Sigma$, $N_1 = N(\underline{\beta}, \underline{\alpha})$, $N_2 = N(\underline{\beta}, \underline{\gamma})$, $N_3 = N(\underline{\alpha}, \underline{\gamma})$ e sem perda de generalidade $N_2 \leq N_3$, temos que $\alpha_n = \beta_n$ para todo $|n| \leq N_2$. Logo, $N_2 \leq N_1$ e

$$\tfrac{1}{2^{N_1}} \leq \tfrac{1}{2^{N_2}} \leq \tfrac{1}{2^{N_2}} + \tfrac{1}{2^{N_3}} \iff d(\underline{\alpha}, \underline{\beta}) \leq d(\underline{\alpha}, \underline{\gamma}) + d(\underline{\gamma}, \underline{\beta}).$$

Logo, d é distância.

Observe que a topologia gerada pela métrica é mais fina que a topologia produto, visto que toda bola é um cilindro. De fato, se $\underline{\alpha}$ e $\underline{\beta}$ pertencem a uma bola de centro em $\underline{\alpha}$ e raio r, então $d(\underline{\alpha}, \underline{\beta}) < r$. Desta forma, $\beta_n = \alpha_n$ para todo $|n| < r$ e se algumas coordenadas coincidem temos um cilindro.

Por outro lado, a topologia produto é mais fina que a topologia gerada pela métrica. De fato, todo cilindro é interseção de imagens de bolas pelo operador deslocamento, que é um homeomorfismo como será demonstrado a seguir. Então, todo cilindro é a interseção de abertos da métrica. Desta forma, temos que a topologia gerada pela métrica coincide com a topologia produto.

1.3 A Transformação Shift

Nesta seção vamos introduzir uma dinâmica $\sigma : \Sigma \to \Sigma$ no espaço Σ, chamada deslocamento (ou shift) de Bernoulli, onde $\sigma((x_n)_n) = (x_{n+1})_n$. Os teoremas a seguir asseguram que a transformação shift é um homeomorfismo.

Teorema 1.2 *O deslocamento,definido no espaço de sequências bilaterais, é contínuo.*

Demonstração. Dado $\varepsilon > 0$, existe $N > 0$ suficientemente grande tal que $\varepsilon > \frac{1}{N}$. Seja n tal que

$$\frac{1}{2^i} < \frac{1}{N}.$$

Tome $0 < \delta < \frac{1}{2^i}$. Assim, para quaisquer $x, y \in \Sigma$, se $d(x, y) < \delta$, então $x_i = y_i$ para $-n - 1 \leq i \leq n + 1$, e portanto $\sigma(x)$ e $\sigma(y)$ coincidem nas n-ésimas coordenadas centrais, implicando que:

$$d(\sigma(x), \sigma(y)) < \frac{1}{2^i} < \frac{1}{N} < \varepsilon$$

□

O deslocamento a esquerda, definido no espaço de sequências bilaterais, é uma bijeção e o deslocamento a direita $\tilde{\sigma}((x_n)_n) = (x_{n-1})$ é a sua inversa. Note também que

$$\sigma(C^{n_1,...,n_k}_{\alpha_1,...,\alpha_k}) = \left\{\omega \in \Sigma : \omega_{n_{i+1}} = \alpha_i \forall i = 1, ..., n\right\}$$

é um cilindro, isto é, σ leva cilindros em cilindros. De forma análoga, vemos que a pré imagem de um cilindro é ainda um cilindro. Logo, σ é um homeomorfismo.

Observe que o operador deslocamento não é injetivo quando definido no espaço de sequências unilaterais com l símbolos, isto é, no espaço $\tilde{\Sigma} = \{\underline{\alpha} = (\alpha_0, \alpha_1, ...) : \alpha_i \in \{1, 2, ..., l\}, i \in \mathbb{Z}\}$,como o mostra o exemplo a seguir.

Exemplo 4 *Sendo* $\underline{\alpha} = (1, 2, 2, ...)$ *e* $\underline{\beta} = (2, 2, 2, ...)$*, temos que*

$$\sigma(\underline{\alpha}) = \sigma(\underline{\beta}) = (2, 2, 2, ...)$$

A restrição de σ a algum subconjunto fechado e invariante de Σ, é chamado de Sistema Dinâmico Simbólico.

1.4 Medidas Invariantes

A teoria da medida foi desenvolvida no final do século XIX e no início do século XX por Emile Borel, Henri Lebesgue, Johan Radon e Maurice Fréchet. As principais aplicações são:

- na fundamentação da integral de Lebesgue, que generaliza (com vantagens) a integral de Riemann;
- Na axiomatização da teoria de probabilidade feita por Andrey Kolmogorov;
- na definição mais geral de espaços não euclidianos.

Abordaremos nesta seção alguns dos conceitos e resultados relevantes da teoria da medida necessários para a discussão do Teorema de Szemerédi.

Uma medida é uma função dos subconjuntos (ou partes) de algum conjunto X, que associa "massas"a cada uma dessas partes. Uma medida deve satisfazer certas propriedades intuitivas que idealizamos para uma função "massa".

Definição 1.5 *Seja* X *um conjunto qualquer.* $\mathcal{B} \subset \mathcal{P}(X)$ *é uma* $\sigma-$*álgebra se*

i. $\varnothing \in \mathcal{B}$

ii. $A \in \mathcal{B}$ *implica* $A^c \in \mathcal{B}$

iii. $A_i \in \mathcal{B}$*, para todo* $i = 1, 2, ...,$ *implica* $\cup_{i=1}^{\infty} A_i \in \mathcal{B}$*.*

O par $(X, \mathcal{B})$ é chamado de **espaço mensurável**, enquanto um elemento $B \in \mathcal{B}$ é referido como um **conjunto mensurável**.

Exemplo 5 *Seja X um conjunto qualquer. Então, $\{\varnothing, X\}$ e $\mathcal{P}(X)$ são exemplos triviais de $\sigma-$álgebra. São a menor e maior $\sigma-$álgebra, respectivamente, de X.*

Exemplo 6 *Se X é um conjunto não enumerável, então*

$$\{E \subset X : E \text{ é enumerável ou } E^c \text{ é enumerável}\}$$

é uma $\sigma-$álgebra, chamada $\sigma-$álgebra dos conjuntos enumeráveis ou coenumeráveis.

Exemplo 7 *A interseção de uma família de $\sigma-$álgebras é uma $\sigma-$álgebra. Segue-se que, se $\mathcal{C} \subset \mathcal{P}(X)$, então existe uma menor $\sigma-$álgebra $\mathcal{M}(\mathcal{C})$ que contém $\mathcal{C}$; existe pelo menos uma $\sigma-$álgebra que contém $\mathcal{C}$, a $\sigma-$álgebra $\mathcal{P}(X)$. $\mathcal{M}(\mathcal{C})$ é chamada $\sigma-$álgebra gerada por $\mathcal{C}$.*

Definição 1.6 *Uma **medida** em $(X, \mathcal{B})$ é uma função $\mu : \mathcal{B} \to [0, \infty]$ que é contavelmente aditiva, isto é, se $\{B_i\}_{i\in\mathbb{N}}$ for uma coleção dois a dois disjunta de elementos de $\mathcal{B}$ então*

$$\mu(\overset{\infty}{\underset{n=1}{\cup}} B_i) = \sum_{i=1}^{\infty} \mu(B_i).$$

Exemplo 8 *Se X é qualquer conjunto não vazio, $\mathcal{M} = \mathcal{P}(X)$ e $f : X \to [0, \infty)$ é uma função qualquer, então f induz uma medida μ em $\mathcal{M}$ definindo-se:*

$$\mu(E) = \textstyle\sum_{x\in E} f(x) := \sup\left\{ \sum_{x\in F} f(x) : F \text{ é finito} \right\}$$

Em particular, se $f(x) = 1$ para todo x então μ é chamada a medida de contagem; se $f(x_0) = 1$ e $f(x) = 0$ para todo $x \neq x_0$, então μ é chamada a medida de Dirac ou medida de ponto de massa.

Exemplo 9 *Se X é um conjunto não enumerável, e $\mathcal{M}$ é a $\sigma-$álgebra dos conjuntos enumeráveis ou coenumeráveis, então a função μ definida por*

$$\mu(E) = \begin{cases} 0, \text{ se } E \text{ é enumerável,} \\ 1, \text{ se } E \text{ é coenumerável} \end{cases}$$

é uma medida.

Particularmente, se houver um elemento B de medida finita, então

$$\mu(B) = \mu(B \cup \varnothing) = \mu(B) \cup \mu(\varnothing) \Rightarrow \mu(\varnothing) = 0.$$

Se a medida tiver imagem em $[0, \infty)$ então diz-se que ela é finita. Se for finita, pode ser normalizada para que $\mu(X) = 1$, e neste caso é chamada de medida de probabilidade, ou simplesmente de probabilidade. A tripla $(X, \mathcal{B}, \mu)$ é chamada de espaço de medida, ou de espaço de probabilidade se $\mu(X) = 1$.

Definição 1.7 *Seja $(X, \mathcal{B}, \mu)$ um espaço mensurável. Diz-se que $f : X \to \mathbb{R}$ é mensurável se para todo boreleano B de $\mathbb{R}$ valer $f^{-1}(B) \in \mathcal{B}$.*

A mensurabilidade pode ser testada com conjuntos (c, ∞), pois eles geram a $\sigma-$álgebra dos boreleanos de $\mathbb{R}$. Se X for espaço topológico e $\mathcal{B}$ for a $\sigma-$álgebra de Borel de X então qualquer função contínua será mensurável.

Com o intuito de definirmos integração de uma função simples, que será definida logo abaixo, denotaremos por $\mathcal{X}_A$ a função característica de A:

$$\mathcal{X}_A = \begin{cases} 0, \text{se } x \notin A \\ 1, \text{ se } x \in A \end{cases}$$

Definição 1.8 *Seja $(X, \mathcal{B}, \mu)$ espaço de probabilidade. Diz-se que $f : X \to \mathbb{R}$ é simples se*

$$f = \sum_{i=1}^{n} a_i \mathcal{X}_{A_i},$$

onde os $A_i's$ são elementos dois a dois disjuntos de $\mathcal{B}$.

Definição 1.9 *A integral de uma função simples $f = \sum_{i=1}^{n} a_i \mathcal{X}_{A_i}$ é dada por*

$$\int f d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$$

Nesse caso também dizemos que f preserva μ. O conceito acima faz sentido, pois a pré-imagem de um conjunto mensurável por uma transformação mensurável ainda é um conjunto mensurável.

É fácil ver que qualquer função mensurável não negativa pode ser aproximada, por baixo e monotonamente, por uma sequência de funções simples. A integral de f é definida como sendo o limite das integrais das funções simples, provando-se que independe da sequência escolhida. Diz-se que f é integrável se $\int f d\mu < \infty$.

Para funções reais, basta decompor $f = f_+ - f_-$ e dizer que f é integrável se f_+ e f_- forem integráveis. A função f será integrável se, e somente se, $|f|$ for integrável.

Teorema 1.3 (da Convergência Monótona) *Considere a sequência $\{f_1 \leq f_2 \leq ...\}$ de funções reais integráveis em $(X, \mathcal{B}, \mu)$. Se $\left\{ \int f_n d\mu \right\}$ for sequência limitada, então $\lim_{n\to\infty} f_n$ existe em quase todo ponto e $\int (\lim f_n) d\mu = \lim \int f_n d\mu$. Se a sequência não for limitada, então ou $\lim f_n$ não existe num conjunto de medida positiva ou $\lim f_n$ existe num conjunto de medida total mas, não é integrável.*

Teorema 1.4 (Lema de Fatou) *Seja* $\{f_n\}_n$ *sequência de funções mensuráveis, limitadas inferiormente por uma função integrável. Se* $\liminf_n \int f_n d\mu < \infty$*, então* $\liminf_n f_n$ *é integrável e*

$$\int (\liminf_n f_n) d\mu = \liminf_n \int f_n d\mu$$

Um exemplo que ilustra um pouco esse teorema e mostra que não precisa haver igualdade é dado pelas funções $f_n : [0,1] \to \mathbb{R}$ com

$$f_n(x) = \begin{cases} n^2 x & , \text{ se } 0 \leq x \leq 1/n \\ 2n - nx^2 & , \text{ se } 1/n \leq x \leq 2/n \\ 0 & , \text{ se } 2/n \leq x \leq 1 \end{cases}$$

onde as integrais são sempre iguais a 1, as funções são limitadas inferiormente por qualquer função integrável não positiva, e a função limite é a função nula (não precisa que exista o limite das funções para que o Teorema seja verdadeiro).

Segue-se o seguinte corolário deste último :

Teorema 1.5 (da Convergência Dominada) *Se* g *é integrável e* $\{f_n\}_n$ *é sequência de funções mensuráveis com* $|f_n| \leq g$ *q.t.p. e* $\lim f_n = f$ *q.t.p., então* f *é integrável e*

$$\int f_n d\mu = \int f d\mu.$$

Definição 1.10 *Seja* $(M, \mathcal{B}, \mu)$ *um espaço de medida e seja* $f : M \to M$ *uma transformação mensurável. Dizemos que a medida* μ *é invariante por* f *se*

$$\mu(E) = \mu(f^{-1}(E)) \text{ para todo conjunto mensurável } E \subset M.$$

Teorema 1.6 *Sejam* $f : M \to M$ *uma transformação mensurável e* μ *uma medida em* M*. Então* f *preserva* μ *se, e somente se,*

$$\int \phi d\mu = \int \phi \circ f d\mu$$

para toda função $\mu-$*integrável* $\phi : M \to \mathbb{R}$*.*

Demonstração. Suponhamos que a medida μ é invariante, isto é, $\mu(B) = \mu(f^{-1}(B))$ para todo conjunto mensurável B.

$$\int \mathcal{X}_B d\mu = \mu(B) \text{ e } \mu(f^{-1}(B)) = \int (\mathcal{X}_B \circ f) d\mu$$

O que mostra a igualdade para as funções características e por linearidade da integral é válida para funções simples. Vamos concluir agora que vale para toda função integrável.Dada qualquer função integrável $\phi : M \to \mathbb{R}$, considere uma sequência $(s_n)_n$ de funções simples convergindo para ϕ e tal que $|s_n| \leq |\phi|$ para todo n. Então,

$$\int \phi d\mu = \lim_n \int s_n d\mu = \lim_n \int (s_n \circ f) d\mu = \int (\phi \circ f) d\mu.$$

Reciprocamente, se $\int \phi d\mu = \int \phi \circ f d\mu$ é suficiente mostrar que μ é invariante para um conjunto fechado C qualquer. Seja $\{U_n\}_n$ uma sequência decrescente de abertos tal que $\cap_{n=1}^{\infty} U_n = C$. Isso implica que $\cap_{n=1}^{\infty} f^{-1}(U_n) = f^{-1}(C)$. Tome ϕ_n contínua que valha 1 em C e zero fora de U_n. Então, $\phi_n \circ f$ vale 1 em $f^{-1}(C)$ e zero fora de $f^{-1}(U_n)$. Como μ é probabilidade, $\mu(U_n)$ converge a $\mu(C)$ e $\mu(f^{-1}(U_n))$ converge a $\mu(f^{-1}(C))$.

Além disso,

$$\mu(C) \leq \int \phi_n d\mu \leq \mu(U_n).$$

Segue que, $\int \phi d\mu$ converge para $\mu(C)$. Analogamente, $\int \phi_n \circ f d\mu$ converge para $\mu(f^{-1}(C))$.

Por hipótese

$$\int \phi_n d\mu = \int \phi_n \circ f d\mu \Rightarrow \mu(C) = \mu(f^{-1}(C))$$

□

Capítulo 2

Teoremas de Recorrência e Extensões Naturais

2.1 Teorema de Recorrência de Birkhoff

Nesta Seção demonstraremos o Teorema de Recorrência de Birkhoff e decorrente desse a sua versão múltipla, crucial para a conclusão do Teorema de Van der Waerden. Além desses, também faremos uma breve discussão sobre o Teorema de Recorrência Múltipla de Poincaré , a fim de dar uma ideia da demonstração do Teorema de Szemerédi.

Definição 2.1 *Considere M um espaço topológico. Dizemos que um ponto $x \in M$ é recorrente para uma transformação $f : M \to M$ se existe uma sequência de naturais $n_j \to \infty$ tal que $f^{n_j}(x) \to x$.*

A versão simples do Teorema de Birkhoff investiga a existência de pontos recorrentes, mais precisamente:

Teorema 2.1 (Birkhoff) *Seja $f : M \to M$ uma transformação contínua num espaço métrico compacto M, então existe algum ponto $x \in M$ que é recorrente para f.*

Para fazermos a demonstração deste resultado , consideremos o lema a seguir.

Lema 2.1 *Seja M espaço métrico compacto, $f : M \to M$ uma transformação contínua e $\mathcal{I}$ a família de conjuntos fechados não vazios de M invariantes por f, isto é, $\mathcal{I} = \left\{X \subset M | X = \overline{X}, f(X) \subset X\right.$ Então, um elemento $X \in \mathcal{I}$ é minimal para a relação de inclusão se, e somente, a órbita de qualquer elemento $x \in X$ é densa em X.*

Demonstração Observe que $\mathcal{I}$ é uma família não vazia, pois pelo menos $M \in \mathcal{I}$. Considere $X \subset M$ elemento minimal de $\mathcal{I}$ e $x \in X$. Note que o fecho da órbita de x, $\mathcal{O}(x) = \{f^n : n \in \mathbb{Z}\}$ é um elemento de $\mathcal{I}$. De fato,$\overline{\mathcal{O}(x)}$ é não vazio, já que M é compacto, fechado e invariante, pois dado

$$y \in \overline{\mathcal{O}}(x) \Rightarrow y = \lim f^{n_k}(x) \Rightarrow f(y) = \lim f(f^{n_k}(x))$$

devido a continuidade da função f. Esta última igualdade pode ser reescrita da seguinte forma $f(y) = \lim f^{n_k+1}(x)$ e desta forma concluímos que $f(y) \in \overline{\mathcal{O}}(x)$, isto é, $f(\overline{\mathcal{O}}(x)) \subset \overline{\mathcal{O}}(x)$. Além disso,

$$\mathcal{O}(x) \subset X \Rightarrow \overline{\mathcal{O}}(x) \subset X, \text{ pois } X \text{ é fechado.}$$

Pela minimalidade de X concluímos que $\overline{\mathcal{O}}(x) = X$. Reciprocamente, suponha que a órbita de qualquer elemento $x \in X$ é densa em X, isto é, que $\overline{\mathcal{O}}(x) = X$ para todo $x \in X$ e considere $Y \in \mathcal{I}$ tal que $Y \subset X$. Temos que, $y \in Y \Rightarrow f^n(y) \in Y$ para todo n natural. Logo,

$$\mathcal{O}(y) \subset Y \Rightarrow X = \overline{\mathcal{O}(y)} \subset Y \Rightarrow X = Y.$$

Portanto, X é minimal.

□

Desta forma, de acordo com o lema, é suficiente verificarmos que, mediante as hipóteses dadas, a família $\mathcal{I}$ de conjuntos fechados, não vazios de M e invariantes por f possui elemento minimal. Pois, teríamos desta forma que todo elemento $x \in X$ é recorrente para f, já que ele escreve-se como limite de uma sequência de $f^n(x)$.

Demonstração do Teorema de Birkhoff. Considere a família $\mathcal{I}$ de conjuntos fechados não vazios de M invariantes por f, isto é $f(X) \subset X$ e $\{X_\alpha\} \in \mathcal{I}$ um subconjunto qualquer de $\mathcal{I}$ totalmente ordenado. Vamos mostrar que $\{X_\alpha\}$ admite algum minorante. De fato, tome $X = \cap_\alpha X_\alpha$ e observe que $X \subset X_\alpha$ para todo α, é fechado, não vazio pois os X_α são compactos e constituem uma família encaixada. Além disso, X também é invariante por f pois, dado

$$x \in X \Rightarrow x \in X_\alpha, \forall \alpha \Rightarrow f(x) \in X_\alpha, \forall \alpha \Rightarrow f(x) \in X.$$

Logo, X é um minorante de $\{X_\alpha\}$ e pelo Lema de Zorn (lema 4.1, apêndice) $\mathcal{I}$ admite elemento minimal.

□

2.2 Teorema de Recorrência Múltipla de Birkhoff

Trataremos agora de uma versão mais geral do teorema de Birkhoff, onde são consideradas várias transformações. Para nossos fins, é suficiente que tais transformações sejam homeomorfismos, todavia o teorema que segue também é válido se as transformações forem apenas contínuas (Seção 2.4) .

Teorema 2.2 *Seja M um espaço métrico compacto e $f_1, ..., f_q : M \to M$ homeomorfismos que comutam entre si, isto é, $f_i \circ f_j = f_j \circ f_i$, então existe $a \in M$ e uma sequência $(n_k)_k \to \infty$ tal que*

$$\lim_k f_i^{n_k}(a) = a$$

para todo $i = 1, ..., q$.

Para demonstrarmos este resultado, faremos inicialmente a seguinte construção: $F : M^q \to M^q$ definida no espaço produto $M^q = M$ com $F(x_1, ..., x_n) = (f_1(x_1), ..., f_n(x_n))$, onde as f_i's são homeomorfismos. Defina a diagonal Δ_q de M^q como o conjunto formado pelos pontos $\tilde{x} = (x, ..., x) \in M^q$. Assim, a conclusão do Teorema de Recorrência Múltipla de Birkhoff pode ser equivalentemente interpretado da seguinte forma: existe $\tilde{a} \in \Delta_q$ e $(n_k)_k \to \infty$ tal que

$$\lim_k F^{n_k}(\tilde{a}) = \tilde{a}.$$

Usaremos indução no número q de transformações. Para $q = 1$ o teorema resume-se ao Teorema de Recorrência de Birkhoff. Suponha então que o teorema vale para um número $q - 1$ de homeomorfismos que comutam entre si, com $q \geq 2$. Vamos provar que vale para a família $f_1, ..., f_q$.
Denote por $\mathcal{G}$ o grupo abeliano gerado pelos homeomorfismos $f_1, ..., f_q$, com a operação de composição de funções. Dizemos que um conjunto X é $\mathcal{G}-$invariante quando $g(X) \subset X$, para toda função $g \in \mathcal{G}$.

Lema 2.2 *A família de subconjuntos não vazios, fechados, $\mathcal{G}$-invariantes de um compacto M admite elemento minimal.*

Demonstração. Considere $\mathcal{I}$ a família de subconjuntos não-vazios, fechados, $\mathcal{G}-$invariantes do compacto M e $\in \mathcal{I}$ totalmente ordenado. Defina $X = \cap_\alpha X_\alpha$. Temos que, $X \subset X_\alpha \forall \alpha$, é fechado e $\{X_\alpha\} \in \mathcal{I}$ para todo α, visto que os X_α são compactos e constituem uma família encaixada. Além disso, X é $\mathcal{G}-$invariante pois para todo homeomorfismo $g \in \mathcal{G}$ e $x \in X$ temos que $g(x) \in X_\alpha, \forall \alpha$.

Logo $g(\cap_\alpha X_\alpha) \subset \cap_\alpha X_\alpha$ para todo α, donde concluímos que X é um minorante de $\{X_\alpha\}$ e, pelo Lema de Zorn, $\mathcal{I}$ admite elemento minimal.

□

Desta forma, não perdemos a generalidade se considerarmos o espaço M como sendo minimal. Essa informação será de suma importância, para a sequência de lemas auxiliares a conclusão do teorema de Recorrência Múltipla de Birkhoff.

Para tal, façamos a construção da seguinte função:

$$\phi : \Delta_q \to [0, \infty), \phi(\tilde{x}) = \inf \{d(F^n(\tilde{x}), \tilde{x}) : n \geq 1)\}.$$

Note que ϕ é semicontínua, isto é, dado qualquer $\varepsilon > 0$, todo ponto $\tilde{x}$ admite alguma vizinhança V tal que $\phi(\tilde{y}) < \phi(\tilde{x}) + \varepsilon$. De fato, basta notar que dado $\tilde{x} \in \Delta_q$ e $\tilde{y} \in V$ de tal forma que $d(\tilde{x}, \tilde{y}) < \varepsilon/2$ temos que

$$d(F^n(\tilde{x}), \tilde{x}) \leq d(F^n(\tilde{x}), F^n(\tilde{y})) + d(F^n(\tilde{y}), \tilde{y}) + d(\tilde{x}, \tilde{y}) < \varepsilon + d(F^n(\tilde{y}), \tilde{y}).$$

Logo, ϕ admite algum ponto $\tilde{a}$ de continuidade. Construiremos agora uma sequência de resultados que levarão a conclusão de que esse ponto $\tilde{a}$ de continuidade satisfaz a conclusão do Teorema.

Lema 2.3 *Se M é minimal, então para todo aberto não vazio $U \subset M$ existe um subconjunto finito $\mathcal{H} \subset \mathcal{G}$ tal que*

$$M = \bigcup_{h \in \mathcal{H}} h^{-1}(U)$$

Demonstração. Dado $x \in M$ temos que o fecho da sua órbita, $\mathcal{O}(x) = \{g(x) : g \in \mathcal{G}\}$, é um subconjunto não vazio de M, fechado e $\mathcal{G}$-invariante. Então, da minimalidade de M, concluí-se que $\mathcal{O}(x)$ é densa em M. Em particular, existe $g \in \mathcal{G}$ tal que $g(x) \in U$. Assim, provamos que $\{g^{-1}(U) : g \in \mathcal{G}\}$ é uma cobertura aberta de M. Da compacidade de M , segue-se que existe uma subcobertura finita.

□

Note que a aplicação $M \to \Delta_q$, $x \mapsto \tilde{x} = (x, ..., x)$ é um homeomorfismo, então todo aberto $U \subset M$ corresponde a um aberto $\tilde{U} \subset \Delta_q$ via esse homeomorfismo. Dado qualquer $g \in \mathcal{G}$, representaremos por $\tilde{g} : M^q \to M^q$ o homeomorfismo definido por $\tilde{g}(x_1, ..., x_q) = (g(x_1), ..., g(x_n))$. Sendo $\mathcal{G}$ abeliano, então F comuta com $\tilde{g}$. Além disso, $\tilde{g}$ preserva a diagonal Δ_q. Aplicando o lema anterior no espaço Δ_q, temos que

$$\Delta_q = \bigcup_{h \in \mathcal{H}} h^{-1}(U)$$

Lema 2.4 *Dado $\varepsilon > 0$ existem $\tilde{x}, \tilde{y} \in \Delta_q$ e $n \geq 1$ tais que $d(F^n(\tilde{x}), \tilde{y}) < \varepsilon$.*

Demonstração. Defina $g_i = f_i \circ f_q^{-1}$ para cada $i = 1, ..., q-1$. Como as f_i comutam entre si, segue-se o mesmo para as g_i. Então, pela hipótese de indução existe $y \in M$ e $(n_k)_k \to \infty$ tal que

$$\lim_k g_i^{n_k}(y) = y$$

para todo $i = 1, ..., q-1$. Denote $x_k = f_q^{-n_k}(y)$ e considere $\tilde{x}_k = (x_k, ..., x_k) \in \Delta_q$. Então,

$$F^{n_k}(\tilde{x}_k) = (f_1^{n_k}(x_k), ..., f_{q-1}^{n_k}(x_k), f_q^{n_k}(x_k))$$

$$= (f_1^{n_k} f_q^{-n_k}(y), ..., f_{q-1}^{n_k} f_q^{-n_k}(y), y) = (g_1^{n_k}(y), ..., g_{q-1}^{n_k}(y), y)$$

que converge para $(y, ..., y)$ quando $k \to \infty$. O que conclui o lema para $\tilde{x} = \tilde{x}_k$, $\tilde{y} = (y, ..., y)$ e $n = n_k$, para k suficientemente grande.

□

Mostraremos agora que o ponto $\tilde{y}$ do lema anterior é arbitrário.

Lema 2.5 *Dado $\varepsilon > 0$ e $\tilde{z} \in \Delta_q$ existem $\tilde{w} \in \Delta_q$ e $m \geq 1$ satisfazendo $d(F^m(\tilde{w}), \tilde{z}) < \varepsilon$.*

Demonstração. Dado $\varepsilon > 0$ e $\tilde{w} \in \Delta_q$ considere $\tilde{U}$ a bola de centro em $\tilde{z}$ e raio $\varepsilon/2$. Então, pela observação do lema 7 existe um subconjunto finito $\mathcal{H} \subset \mathcal{G}$ tal que $\Delta_q = \bigcup_{h \in \mathcal{H}} h^{-1}(\tilde{U})$. Como os elementos de $\mathcal{G}$ são uniformemente contínuos, pois tratam-se de funções contínuas em um domínio compacto, então existe $\delta > 0$ tal que

$$d(\tilde{x}_1, \tilde{x}_2) < \delta \Rightarrow d(\tilde{h}(\tilde{x}_1), \tilde{h}(\tilde{x}_2)) < \varepsilon/2$$

para todo $h \in \mathcal{H}$. Pelo lema anterior existem $\tilde{x}, \tilde{y} \in \Delta_q$ e $n \geq 1$ tais que $d(F^n(\tilde{x}), \tilde{y}) < \varepsilon$. Fixe $h \in \mathcal{H}$ tal que $\tilde{y} \in \tilde{h}^{-1}(\tilde{U})$, isto é, $\tilde{h}(\tilde{y}) \in \tilde{U}$. Então,

$$d(\tilde{h}(F^n(\tilde{x})), \tilde{z}) \leq d(\tilde{h}(F^n(\tilde{x})), \tilde{h}(\tilde{y})) + d(\tilde{h}(\tilde{y}), \tilde{z}) < \varepsilon/2 + \varepsilon/2.$$

Tomando $\tilde{w} = \tilde{h}(\tilde{y})$ e notando que F^n comuta com $\tilde{h}$ segue-se o resultado $d(F^n(\tilde{w}), \tilde{z}) < \varepsilon$.

□

Mostraremos agora que é possível tomar $\tilde{x} = \tilde{y}$ no lema anterior.

Lema 2.6 *(Bowen) Dado $\varepsilon > 0$ existem $\tilde{v} \in \Delta_q$ e $k \geq 1$ tal que $d(F^k(\tilde{v}), \tilde{v}) < \varepsilon$.*

Demonstração. Dado $\varepsilon > 0$ e $\tilde{z}_0 \in \Delta_q$, considere as sequências ε_j, m_j e $\tilde{z}_j$, $j \geq 1$, definidas da seguinte forma: dado $\varepsilon_1 = \varepsilon/2$. Então, pelo lema anterior existem $\tilde{z}_1 \in \Delta_q$ e $m_1 \geq 1$ tais que $d(F^{m_1}(\tilde{z}_1), \tilde{z}_0) < \varepsilon_1$. Da continidade de F^{m_1}, existe um $\varepsilon_2 < \varepsilon_1$ tal que $d(\tilde{z}, \tilde{z}_1) < \varepsilon_2$ implica $d(F^{m_1}(\tilde{z}), \tilde{z}_0) < \varepsilon_1$.
Em geral, dado qualquer $j \geq 2$ e $\tilde{z}_{j-1} \in \Delta_q$, temos pelo lema anterior que: existem $m_j \geq 1$ e $\tilde{z}_j \in \Delta_q$ tal que $d(F^{m_j}(\tilde{z}_j), \tilde{z}_{j-1}) < \varepsilon_j$. Pela continuidade de F^{m_j}, existe $\varepsilon_{j+1} < \varepsilon_j$ tal que $d(\tilde{z}, \tilde{z}_j) < \varepsilon_{j+1}$ implica $d(F^{m_j}(\tilde{z}), \tilde{z}_{j-1}) < \varepsilon_j$. Em particular, para todo $i < j$ temos

$$d(F^{m_{i+1}+\ldots+m_j}(\tilde{z}_j), \tilde{z}_i) < \varepsilon_{i+1} \leq \varepsilon/2.$$

Sendo Δ_q compacto podemos encontrar, i, j com $i < j$ tais que $d(\tilde{z}_i, \tilde{z}_j) < \varepsilon/2$. Tomando $k = m_{i+1} + \ldots + m_j$, temos

$$d(F^k(\tilde{z}_j), \tilde{z}_j) \leq d(F^k(\tilde{z}_j), \tilde{z}_i) + d(\tilde{z}_j, \tilde{z}_i) < \varepsilon.$$

□

Estes resultados são suficientes para concluirmos que, de fato o ponto de continuidade $\tilde{a}$ de ϕ, definida inicialmente satisfaz a conclusão do Teorema. Para isso, observe que $\phi(\tilde{a}) = 0$. De fato, suponha que $\phi(\tilde{a}) > 0$. Então, por continuidade, existem $\beta > 0$ e uma vizinhança V de $\tilde{a}$ tais que $\phi(\tilde{y}) \geq \beta > 0$ para todo $\tilde{y} \in V$. Logo,

$$d(F^n(\tilde{y}), \tilde{y}) \geq \beta \text{ para todo } y \in V \text{ e todo } n \geq 1$$

Por outro lado, já vimos que para todo $\tilde{x} \in \Delta_q$ existe $h \in \mathcal{H}$ tal que $h(\tilde{x}) \in V$. Como as transformações h são uniformemente contínuas, podemos fixar $\alpha > 0$ tal que:

$$d(\tilde{w}, \tilde{z}) < \alpha \Rightarrow d(\tilde{h}(\tilde{w}), \tilde{h}(\tilde{z})) < \beta, \text{ para toda } h \in \mathcal{H}.$$

Então, pelo Lema de Bowen e lembrando que F^n comuta com $\tilde{h}$, existe $n \geq 1$ tal que

$$d(\tilde{x}, F^n(\tilde{x})) < \alpha \Rightarrow d(\tilde{h}(\tilde{x}), F^n(\tilde{h}(\tilde{x}))) < \beta.$$

Chegamos a uma contradição, concluíndo assim que $\phi(0) = 0$. Logo, existe uma sequência $(n_k)_k \to \infty$ tal que $d(F^n(\tilde{a}), \tilde{a}) \to 0$. □

2.3 Teorema de Recorrência múltipla de Poincaré

O Teorema de Recorrência Múltipla de Poincaré afirma, basicamente, que existe algum tempo n tal que os iterados de um subconjunto com medida positiva de pontos de E retornam para E, simultaneamente para todas as transformações f_i, nesse momento n. Mais precisamente:

Teorema 2.3 *Seja $(M, \mathcal{B}, \mu)$ um espaço de probabilidade e sejam $f_i : M \to M$, $i = 1, ..., q$ transformações mensuráveis que preservam μ e que comutam entre si. Então, para qualquer conjunto $E \subset M$ com medida positiva, existe $n \geq 1$ tal que*

$$\mu(E \cap f_1^{-n}(E) \cap ... \cap f_q^{-n}(E)) > 0.$$

A demonstração deste teorema pode ser encontrada no livro de Fustenberg [5], e sua apresentação não será feita neste trabalho. Vamos apenas mencioná-lo a fim de seu na prova do teorema de Szemerédi sobre existência de progressões aritméticas em subconjuntos densos dos números inteiros.

2.4 Extensões Naturais

Com o intuito de generalizar o Teorema de Recorrência Múltipla de Birkhoff, faremos uma breve análise ao conceito de extensões naturais, que nos permite reduzir a demostração de casos mais gerais de aplicações simplesmente contínuas, ao caso de funções inversíveis.

Dada uma transformação sobrejetiva $f : M \to M$ é sempre possível encontrar uma extensão $\tilde{f} : \tilde{M} \to \tilde{M}$ que é invertível, isto é, existe uma aplicação sobrejetiva $\pi : \tilde{M} \to \tilde{M}$ tal que $\pi \circ \tilde{f} = f \circ \pi$.

De fato, iniciaremos tomando $\tilde{M}$ como o conjunto de todas as pré órbitas de f , isto é, o conjunto de todas as sequências $(x_n)_{n\leq 0}$ indexadas pelos números inteiros não positivos e satisfazendo $f(x_n) = x_{n+1}$ para todo $n < 0$. Considere a aplicação $\pi : \tilde{M} \to M$ que associa a cada sequência $(x_n)_{n\leq 0}$ o seu termo x_0. Claramente, $\pi(\tilde{M}) = M$, pois $\pi(\tilde{M}) \subset M$ e todo $x \in M$ é o termo x_0 de alguma sequência em $\tilde{M}$. Por fim, definimos $\tilde{f} : \tilde{M} \to \tilde{M}$ como sendo o deslocamento a esquerda :

$$\tilde{f}(..., x_n, ..., x_0) = (..., x_n, ..., x_0, f(x_0)).$$

De acordo com a construção acima, temos que

$$\pi \circ \tilde{f}(..., x_n, ..., x_0) = f(x_0) = f \circ \pi(..., x_n, ..., x_0).$$

Além disso, $\tilde{f}$ é invertível. Sua inversa é o deslocamento a direita:

$$(..., y_n, ..., y_{-1}, y_0) \mapsto (..., y_n, ..., y_{-2}, y_{-1}).$$

A seguir serão demonstrados resultados com o intuito de discutir a noção de extensões naturais para transformações que comutam entre si. Para os três próximos lemas consideraremos M um espaço compacto, $f_1, ..., f_q : M \to M$ transformações contínuas, sobrejetivas que comutam entre si e $\hat{M}$ o conjunto das sequências $(x_{n_1,...,n_q})_{n_1,...,n_q\leq 0}$, indexadas pelas q-uplas de inteiros não positivos, tais que

$$f_i(x_{n_1,...,n_i,...,n_q}) = x_{n_1,...,n_{i+1},...,n_q} \text{ para todo } i \text{ e todo } (n_1, ..., n_q).$$

Lema 2.7 *$\hat{M}$ é um espaço métrico compacto. Além disso, $\hat{M}$ é metrizável se M é metrizável.*

Demonstração. Observe que $\hat{M}$ é um subconjunto fechado do compacto $M^{\mathbb{Z}^q_-}$, portanto compacto. Além disso, considere d uma distância em M. Então,

$$\hat{d}((x_{\overline{n}})_{\overline{n}}, (y_{\overline{n}})_{\overline{n}}) = \sum_{n_1,...,n_q\leq 0} 2^{n_1+...+n_q} \min\{d((x_{\overline{n}})_{\overline{n}}, (y_{\overline{n}})_{\overline{n}}), 1\}$$

define uma distância em $\hat{M}$, onde $\overline{n} = (n_1, ..., n_q)$. De fato,

- $\hat{d}((x_{\overline{n}})_{\overline{n}}, (x_{\overline{n}})_{\overline{n}}) = 0$, pois $\min\{d((x_{\overline{n}})_{\overline{n}}, (x_{\overline{n}})_{\overline{n}}), 1)\} = 0$ já que d é métrica.
- Claramente, $\hat{d}((x_{\overline{n}})_{\overline{n}}, (y_{\overline{n}})_{\overline{n}}) = \hat{d}((y_{\overline{n}})_{\overline{n}}, (x_{\overline{n}})_{\overline{n}})$.
- Sendo d distância temos que $d((x_{\overline{n}})_{\overline{n}}, (y_{\overline{n}})_{\overline{n}}) \leq d((x_{\overline{n}})_{\overline{n}}, (z_{\overline{n}})_{\overline{n}}) + d((z_{\overline{n}})_{\overline{n}}, (y_{\overline{n}})_{\overline{n}})$ e assim

$$\hat{d}((x_{\overline{n}})_{\overline{n}}, (y_{\overline{n}})_{\overline{n}}) = \sum_{n_1,...,n_q\leq 0} 2^{n_1+...+n_q} \min\{d((x_{\overline{n}})_{\overline{n}}, (y_{\overline{n}})_{\overline{n}}), 1\}$$

$$\leq \sum_{n_1,...,n_q\leq 0} 2^{n_1+...+n_q} \min\{d((x_{\overline{n}})_{\overline{n}}, (z_{\overline{n}})_{\overline{n}}) + d((z_{\overline{n}})_{\overline{n}}, (z_{\overline{n}})_{\overline{n}}), 1\}$$

$$\leq \sum_{n_1,...,n_q\leq 0} 2^{n_1+...+n_q} \min\{d((x_{\overline{n}})_{\overline{n}}, (z_{\overline{n}})_{\overline{n}}), 1\}$$

$$+ \sum_{n_1,...,n_q\leq 0} 2^{n_1+...+n_q} \min\{d((z_{\overline{n}})_{\overline{n}}, (y_{\overline{n}})_{\overline{n}}), 1\} = \hat{d}((x_{\overline{n}})_{\overline{n}}, (z_{\overline{n}})_{\overline{n}}) + \hat{d}((z_{\overline{n}})_{\overline{n}}, (y_{\overline{n}})_{\overline{n}})$$

□

Lema 2.8 *Seja $\pi : \hat{M} \to M$ a aplicação que envia $(x_{n_1,...,n_q})_{n_1,...,n_q \leq 0}$ no ponto $x_{0,...,0}$ e, para cada i, $\hat{f}_i : \hat{M} \to \hat{M}$ a aplicação que envia $(x_{n_1,...,n_i,...,n_q})_{n_1,...,n_q \leq 0}$ em $(x_{n_1,...,n_{i+1},...,n_q})_{n_1,...,n_q \leq 0}$. Cada $\hat{f}_i$ é um homeomorfismo com $\pi \circ \hat{f}_i = \hat{f}_i \circ \pi$. Além disso, esses homeomorfismos comutam entre si.*

Demonstração. A inversa de $\hat{f}_i$ envia $(x_{n_1,...,n_i,...,n_q})_{n_1,...,n_q \leq 0}$ no ponto $(x_{n_1,...,n_{i-1},...,n_q})_{n_1,...,n_q \leq 0}$. Além disso tanto $\hat{f}_i$ quanto a sua inversa são contínuas, pois cada coordenada da imagem é função contínua de um número finito de coordenadas da variável. Temos também, por definição:

$$\pi \circ \hat{f}_i((x_{n_1,...,n_i,...,n_q})_{n_1,...,n_q \leq 0}) = x_{0,...,1,...,0} = \hat{f}_i \circ \pi((x_{n_1,...,n_i,...,n_q})_{n_1,...,n_q \leq 0}).$$

Por fim,

$$\hat{f}_i \circ \hat{f}_j((x_{n_1,...,n_i,...,...,n_j,...,n_q})_{n_1,...,n_q \leq 0}) = (x_{n_1,...,n_{i+1},...,...,n_{j+1},...,n_q})_{n_1,...,n_q \leq 0}$$

$$= \hat{f}_j \circ \hat{f}_i((x_{n_1,...,n_i,...,...,n_j,...,n_q})_{n_1,...,n_q \leq 0})$$

□

Lema 2.9 *A aplicação $\pi : \hat{M} \to M$ a aplicação que envia $(x_{n_1,...,n_q})_{n_1,...,n_q \leq 0}$ no ponto $x_{0,...,0}$ é contínua e sobrejetiva. Em particular, $\hat{M}$ é não vazio.*

Demonstração. A aplicação π é contínua para a topologia herdada do espaço produto. Para mostrarmos a sobrejetividade, considere $x \in M$. Como supomos anteriormente as $f_1, ..., f_q$ são sobrejetivas,contínuas e comutam entre si, então podemos encontrar uma sequência $(x_n)_{n \leq 0}$ em M tal que $x_0 = 0$ e $f_1...f_n(x_n) = (x_{n+1})$ para cada $n < 0$. Defina

$$x_{n_1,...,n_q} = f^{n_1 - n}...f^{n_q - n}(x_n) \text{ para qualquer } n \leq \min\{n_1, ..., n_q\}.$$

Note que esta definição não depende do número n. De fato, seja $m \leq \min\{n_1, ..., n_q\}$ temos que

$$f^{n_1 - m}...f^{n_q - m}(x_m) = f^{n_1 - m}...f^{n_{q-1} - m}(x_{m,m,...,n_q}) = f^{n_1 - m}(x_{m,...,n_q}) = x_{n_1,...,n_q}.$$

Além disso, $\pi(x_{n_1,...,n_q}) = \pi(f^{n_1 - n}...f^{n_q - n}(x_n)) = x$.

□

Teorema 2.4 *Seja M espaço métrico compacto e $g_1, ..., g_q : M \to M$ transformações contínuas que comutam entre si. Defina $M_g = \cap_{n=1}^{\infty} g_1^n...g_q^n(M)$. Então, $M_g = \cap_{n_1,...,n_q} g_1^{n_1}...g_q^{n_q}(M)$, onde a interseção é sobre todas as q-uplas $(n_1, ..., n_q)$ com $n_i \geq 1$ para todo i. Além disso, $g_i(M_g) \subset M_g$ e a restrição $f_i = g_i | M_g$ é sobrejetiva para todo i.*

Demonstração. Claramente $\cap_{n_1,...,n_q} g_1^{n_1}...g_q^{n_q}(M) \subset \cap_{n=1}^{\infty} g_1^n...g_q^n(M)$, isto é,

$$\cap_{n_1,...,n_q} g_1^{n_1}...g_q^{n_q}(M) \subset M_g.$$

Por outro lado, note que se tomarmos $n = \max\{n_1, ..., n_q\}$, então

$$\cap_{n_1,...,n_q} g_1^{n}...g_q^{n}(M) \subset \cap_{n_1,...,n_q} g_1^{n_1}...g_q^{n_q}(M).$$

De fato,sejam $n = \max\{n_1, ..., n_q\}$ e $y \in g_1^n...g_q^n(M)$. Então, existe $x \in M$ tal que $y = g_1^n...g_q^n(x)$ e podemos escrever $g_i^n = g_i^{m_i} g_i^{n_i}$,onde $m_i + n_i = n$ para todo $i = 1, ..., q$. Daí,

$$y = g_1^n...g_q^n(x) = y = g_1^{m_1} g_1^{n_1} g_2^{m_2} g_2^{n_2}...g_q^{m_q} g_q^{n_q}(x)$$

Usando o fato que as transformações g_i comutam entre si, podemos reescrever a última igualdade da seguinte forma:

$$y = (g_1^{n_1} g_2^{n_2}...g_q^{n_q})(g_1^{m_1} g_2^{m_2}...g_q^{m_q})(x) = g_1^{n_1} g_2^{n_2}...g_q^{n_q} g_1^{n_1} g_2^{n_2}...g_q^{n_q}(\tilde{x})$$

onde $\tilde{x} = g_1^{m_1} g_2^{m_2}...g_q^{m_q}(x) \in M$.
Assim,

$$y \in g_1^{n_1}...g_q^{n_q}(M) \Rightarrow \cap_{n=1}^{\infty} g_1^n...g_q^n(M) \subset \cap_{n_1,...,n_q} g_1^{n_1}...g_q^{n_q}(M)$$

$$\Rightarrow M_g \subset \cap_{n_1,...,n_q} g_1^{n_1}...g_q^{n_q}(M) \Rightarrow M_g = \cap_{n_1,...,n_q} g_1^{n_1}...g_q^{n_q}(M).$$

Desta forma, dado $y \in M_g$ existe um $x_n \in M$ tal que $y = g_1^{n_1} g_2^{n_2}...g_q^{n_q}(x_n)$, para cada $q-$upla $(n_1, ..., n_q)$. Desta forma, $g_i(M_g) \subset M_g$. Por fim, para concluirmos a sobrejetividade de $f_i = g_i|M_g$ basta mostrar que $M_g \subset g_i(M_g)$. Para tal, considere $y \in M_g$. Então, de acordo com o que fora construído, para cada $n \geq 1$ existem $z_n \in M$ tal que $y = g_i(g_1^n g_2^n...g_q^n(z_n))$. Seja z o ponto de acumulação da sequência $(g_1^n g_2^n...g_q^n(z_n))_n$, temos que $z \in M_g$ e $z = g_i^{-1}(y)$. Logo, $y \in g_i(M_g)$.□

Mediante estes resultados, podemos agora estender a demonstração do Teorema de Recorrência Múltipla de Birkhoff para o caso em que as transformações f_i não são necessariamente invertíveis. De fato, pelo teorema anterior não é restrição supor que as transformações $f_1, ..., f_q$ são sobrejetivas. Sejam $\hat{f}_1, ..., \hat{f}_q : \hat{M} \to \hat{M}$ as extensões naturais, no sentido do lema 2.7. Pelo caso invertível do Teorema de Recorrência Múltipla de Birkhoff, existe algum $\hat{x} \in \hat{M}$ que é simultaneamente recorrente para $\hat{f}_1, ..., \hat{f}_q$, isto é, existe $(n_k)_k \to \infty$ e $\hat{x} \in \hat{M}$ tal que $\lim_k f_i^{n_k}(\hat{x}) = \hat{x}$. Daí,

$$\lim_k f_i^{n_k}(\pi(\hat{x})) = \lim_k \pi f_i^{n_k}(\hat{x}) = \pi(\hat{x}).$$

Portanto, $x = \pi(\hat{x})$ é simultaneamente recorrente para $f_1, ..., f_q$.

Capítulo 3

Teorema de van der Waerden

Vamos provar, nesta seção, um importante resultado da Aritmética Combinatória obtido originalmente pelo matemático holandês Bartel van der Waerden. Para tal, faremos uso de uma importante ferramenta de Sistemas Dinâmicos, a dinâmica simbólica.

Definição 3.1 *Uma* ***progressão aritmética finita*** *é uma sequência da forma* $m+n, m+2n, ..., m+qn$*, com* $m \in \mathbb{Z}$ *e* $n, q \geq 1$*. O número* q *é chamado* ***comprimento da progressão****.*

Definição 3.2 *Chamamos de* ***partição finita*** *do conjunto dos números inteiros* $\mathbb{Z}$*, a qualquer família finita de conjuntos* $S_1, ..., S_k \subset \mathbb{Z}$*, dois a dois disjuntos, cuja união é todo o* $\mathbb{Z}$*.*

Exemplo 10 $\mathbb{Z} = S_1 \cup S_2$*, onde* S_1 *é o conjunto dos números inteiros pares e* S_2 *os ímpares.*

Observe que toda sequência $\underline{\alpha} = (\alpha_n)_{n\in\mathbb{Z}} \in \Sigma$ determina uma partição finita de $\mathbb{Z}$ da seguinte forma $S_i = \{n \in \mathbb{Z} : \alpha_n = i\}$. Bem como, toda partição de $\mathbb{Z}$ faz corresponder um elemento $\underline{\alpha} \in \Sigma$ da forma:

$$\alpha_n = i \Rightarrow n \in S_i,$$

onde $\Sigma = \{1, ..., l\}^{\mathbb{Z}}$.

Teorema 3.1 (Van der Waerden) *Dada qualquer partição finita* $\{S_1, S_2, ..., S_l\}$ *de* $\mathbb{Z}$ *existe algum* $j \in 1, ..., l$ *tal que* S_j *contém progressões aritméticas de todos os comprimentos,isto é, para todo* $q \geq 1$ *existem* $m \in \mathbb{Z}$ *e* $n \geq 1$ *tais que* $m + in \in S_j$ *para todo* $1 \leq i \leq q$*.*

Demonstração. Note que, com o dicionário exposto acima é suficiente mostrarmos que para todo $\underline{\alpha} \in \Sigma$ e todo $q \geq 1$ existem $n \geq 1$ e $m \in \mathbb{Z}$ tais que $\alpha_{m+n} = ... = \alpha_{m+qn} = j$, pois desta forma teríamos que $m + n, ..., m + qn \in S_j$.

Já vimos que , $d(\underline{\beta}, \underline{\gamma}) = 2^{-N(\underline{\beta},\underline{\gamma})}$, onde $N(\underline{\beta}, \underline{\gamma}) = \max\{N \geq 0 : \beta_n = \gamma_n, \forall n \in \mathbb{Z}, |n| < N\}$ define uma métrica em Σ e, além disso, observe que $d(\underline{\beta}, \underline{\gamma}) < 1$ se e somente se $\beta_0 = \gamma_0$.

Consideremos agora o fecho da órbita de $\underline{\alpha} \in \Sigma$: $Z = \overline{\{\sigma^n(\underline{\alpha}) : n \in \mathbb{Z}\}}$. Temos então que Z é compacto, pois é um subconjunto fechado do compacto Σ. Além disso, Z também

é invariante pelo deslocamento de Bernoulli. De fato, se $\underline{\beta} \in Z$ então $\underline{\beta} = \sigma^n(\underline{\alpha})$ para algum $n \in \mathbb{Z}$. Assim,

$$\sigma(\underline{\beta}) = \sigma(\sigma^n(\alpha_n)) = \sigma^{n+1}(\alpha_n) \Rightarrow \sigma(\underline{\beta}) \in Z.$$

Dado um número q arbitrário, construiremos as transformações $f_1 = \sigma$, $f_2 = \sigma^2$,...,$f_q = \sigma^q$ definidas de Z em Z. Claramente estas funções comutam entre si, pois $f_i \circ f_j = \sigma^i\sigma^j = \sigma^{i+j} = \sigma^{j+i} = \sigma^j\sigma^i = f_j \circ f_i$. Logo, pelo Teorema de Recorrência Múltipla de Birkhof, existe um $\underline{\theta} \in Z$ e uma sequência $(n_k)_k \to \infty$ tal que

$$\lim_k f_i^{n_k}(\underline{\theta}) = \underline{\theta} \text{ para todo } i = 1, ..., q$$

$$\Rightarrow \lim_k \sigma^{in_k}(\underline{\theta}) = \underline{\theta} \text{ para todo } i = 1, ..., q$$

Em particular fixando um $n = n_j$ tal que $\sigma^n(\underline{\theta}), \sigma^{2n}(\underline{\theta})$,...,$\sigma^{qn}(\underline{\theta})$ estão a uma distância menor que $\frac{1}{2}$ de $\underline{\theta}$, temos que:

$$d(\sigma^{in}(\underline{\theta}), \sigma^{jn}(\underline{\theta})) \leq d(\sigma^{in}(\underline{\theta}), \underline{\theta}) + d(\sigma^{jn}(\underline{\theta}), \underline{\theta}) = d(f_i^n(\underline{\theta}), \underline{\theta}) + d(f_j^n(\underline{\theta}), \underline{\theta}) < \tfrac{1}{2} + \tfrac{1}{2} = 1$$

para todo $1 \leq i, j \leq q$. Sendo $\underline{\theta} \in Z$ podemos encontrar um $m \in \mathbb{Z}$ tal que $\sigma^m(\underline{\alpha})$ está muito próximo de $\underline{\theta}$ e por continuidade da aplicação σ temos que

$$d(\sigma^{in}(\sigma^m(\underline{\alpha})), \sigma^{jn}(\sigma^m(\underline{\alpha}))) < 1 \Rightarrow d(\sigma^{m+in}(\underline{\alpha}), \sigma^{m+jn}(\underline{\alpha})) < 1$$

para todo $1 \leq i, j \leq q$.

Segundo a observação feita no início desta demonstração, com respeito a distância, a última desigualdade só ocorrerá se os elementos que ocupam a posição 0, das sequências envolvidas, forem coincidentes.

Logo,

$$\alpha_{m+n} = ... = \alpha_{m+qn} = k \Rightarrow m + n, ..., m + qn \in S_k, \text{ para algum } k = 1, ..., l.$$

□

Os teoremas do tipo Van der Waerden tem sido extensivamente estudados nos últimos anos. Eles enquadram-se na classe de problemas conhecidos como extremais. Nestes tipos de problemas, normalmente procura-se funções limiares para o tamanho de certas estruturas de modo que sempre seja possível encontrar nestas estruturas uma certa subestrutura. Uma teoria mais geral desses tipos de problemas e encontrada nas literaturas que abordam a teoria de Ramsey ergódica que, por sua vez, aborda questões bem mais gerais, adentrando no campo da teoria dos grafos.

Posteriormente, os matemáticos húngaros Pal Erdos e Pal Turan formularam um resultado mais forte que o Teorema de van der Waerden: todo conjunto S com densidade superior positiva contém sequências aritméticas de tamanho arbitrário. Essa conjectura foi demonstrada pelo matemático húngaro, Endre Szemerédi, há quase quatro décadas. Tal demonstração fora de natureza essencialmente combinatória. No entanto, em 1977, Fustenberg proporcionou um desenvolvimento muito importante para ramos da matemática, com a apresentação de uma demonstração do Teorema de Szemerédi usando argumentos de Teoria Ergódica.

3.1 Outras Formas do Teorema de van der Waerden

Nesta seção, faremos breves abordagens a outras versões do teorema de van der Waerden observando as analogias entre estas formas e a versão já comentada via sistemas dinâmicos. Iniciaremos com a sua forma original, isto é, a sua primeira versão via combinatória através do método de colorir, a começar explanando em que consiste tal método.

Imagine que dispomos de duas cores, azul(A) e vermelho(V), para colorir os números do conjunto $\{1, ..., 9\}$. Temos, por exemplo:

1	2	3	4	5	6	7	8	9
V	V	A	A	V	V	A	A	V

Note que, nesta 2-coloração, há a aparição de uma sequência de três números com a mesma cor e equidistantes, estes são: $1, 5, 9$, todos vermelhos.

Outra 2-coloração de $\{1, ..., 9\}$ pode ser dada por:

1	2	3	4	5	6	7	8	9
V	A	V	A	A	V	A	A	V

Novamente ressaltamos a aparição de uma sequência de 3 números com a mesma cor e equidistantes: $2, 5, 8$, todos azuis.

Observe que não conseguimos o mesmo feito de uma 2-coloração do conjunto $\{1, ..., 8\}$, como por exemplo:

1	2	3	4	5	6	7	8
V	A	V	A	A	V	A	A

Generalizaremos a seguir estas observações.

Definição 3.3 *Dada uma coloração de um subconjunto de* $\mathbb{N}$*, chamaremos de progressão aritmética monocromática,* $k - PA$*, a progressão aritmética de comprimento* k *tal que todos os elementos tem a mesma cor.*

Fazendo analogia as terminologias utilizadas na versão via sistemas dinâmicos, é fácil ver que o método de colorir um subconjunto de $\mathbb{Z}$ em Combinatória é equivalente a construir uma partição deste conjunto, isto é, consiste em considerar partes disjuntas do conjunto dado e cuja união dessas partes resulta no original.

Observe que nas duas primeiras $2-$coloração de $\{1, ..., 9\}$ encontramos progressões aritméticas de monocromáticas de comprimento $k = 3$, isto é, encontramos uma $3-$PA. Já na coloração de $\{1, ..., 8\}$ não conseguimos encontrar uma $3-$PA. Vamos formalizar estes resultados a seguir e, para tal, faremos as seguintes convenções:

- $\mathbb{N} = \{1, 2, 3, ...\}$, isto é, não incluiremos o 0 nesta lista.
- Se $m \in \mathbb{N}$, então $[m] = \{1, ..., m\}$.

Mediante estes conceitos segue-se o Teorema de van der Waerden, provado pelo próprio e cuja demonstração será omitida neste trabalho, pois foge do nosso objetivo.

Teorema 3.2 *Sejam $k \geq 1$ e $c \geq 1$. Então, existe W tal que para cada $c-$coloração $\mathcal{X} : [W] \to [c]$ existe uma $k - PA$ monocromática, isto é, existem $a, d \in \mathbb{N}$ tal que:*

$$\mathcal{X}(a) = \mathcal{X}(a + d) = ... = \mathcal{X}(a + (k - 1)d).$$

Em suma, de todas as formas que partimos o conjunto $[W]$ em subconjuntos disjuntos sempre existirá progressões aritméticas de comprimentos arbitrários em, pelo menos, um dos subconjuntos que formam a partição. Nesta versão Combinatória as partições serão obtidas por meio de colorações e a conclusão é que sempre existem progressões aritméticas de comprimento arbitrários com todos os seus termos pintados de uma única cor e esse resultado vale para partições ou colorações arbitrárias.

Obviamente o número W referido no teorema não é único, por exemplo se $k = 3$ e $c = 2$, então qualquer $W \geq 9$ satisfaz o teorema de van der Waerden. De fato, para obtermos uma progressão aritmética monocromática de comprimento 3 devemos ter inicialmente $W \geq 3$.

Note que para $[W] = [8]$ a seguinte coloração não apresenta $3 - PA$ monocromática:

1	2	3	4	5	6	7	8
V	V	A	A	V	V	A	A

Desta forma, não haverá uma $3 - PA$ monocromática na mesma coloração restrita a $[W]$, para $W \leq 7$.

Se $[W] = [9]$ e na predisposição das cores tivermos até 3 A's sempre aparecerá uma sequência de 3 V's seguidos. Então, é suficiente analisar apenas os seguintes casos:

1. ($3A$ e $6V$) Se os $3A$'s estiverem juntos, em qualquer posição, esta sequência já caracteriza a $3 - PA$ monocromática.

1	2	3	4	5	6	7	8	9
A	A	A	V	V	V	V	V	V

   Mantendo agora um bloco $2A$ sempre junto e em qualquer posição, há sempre a aparição de $3 - PA$'s de razão 1:

1	2	3	4	5	6	7	8	9
A	A	V	V	V	A	V	V	V

Se os $3A$'s estiverem todos separados, há a aparição de Se os $3PA$'s monocromáticas de razões 1 e 2:

1	2	3	4	5	6	7	8	9
A	V	V	V	V	A	V	A	V

2. ($4A$ e $5V$) Se tivermos $4A$'s ou $3A$'s juntos, esta predisposição já caracteriza a aparição de $3PA$'s da cor A. Em qualquer outro caso contrário ao sitado, há a aparição de $3PA$'s monocromáticas de razão 1 ou 2.

   Todos os A's separados:

1	2	3	4	5	6	7	8	9
A	V	A	V	V	A	V	A	V

   Dois blocos de 2 A's:

1	2	3	4	5	6	7	8	9
A	A	V	V	V	A	A	V	V

   Apenas um bloco de 2 A's:

1	2	3	4	5	6	7	8	9
A	A	V	V	V	A	V	A	V

Desta forma, concluímos que 9 é o menor número que satisfaz o teorema de van der Waerden. Este menor número W que satisfaz o teorema recebe nomenclatura especial, como segue na definição:

Definição 3.4 *Sejam $k, c \in \mathbb{N}$. O número $W(k,c)$ é o menor valor W que satisfaz o Teorema de van der Waerden e é denominado o número de van der Waerden.*

Exemplo 11 *Se $c = 1$, então $W(k,c) = k$, visto que com uma única cor a própria sequência $[k]$ forma uma $k-PA$.*

Exemplo 12 *Se $k = 1$, então $W(k,c) = 1$, visto que uma $1-PA$ é formada por qualquer termo simples.*

Exemplo 13 *Se $k = 2$, então $W(2,c) = c + 1$ pelo princípio da casa dos pombos.*

Inspecionar o número de van der Waerden consiste em um processo exaustivo, apesar de simples, todavia o resultado a seguir estima limites inferiores para $W(k,c)$.

Teorema 3.3 *Para todo k e c, tem-se $W(k,c) \geq \sqrt{k-1}c^{(k-1)/2}$.*

Demonstração. Seja W o número em questão. Tentaremos encontrar uma $c-$coloração de $[W]$ tal que não possua uma $k-PA$ monocromática.

Considere o seguinte experimento: para cada $i \in [W]$ escolha aleatoriamente uma cor de $[c]$ para i. A distribuição é uniforme. Primeiro escolhemos a cor da $k-PA$ formada. Seja c a opção. Em seguida, escolhemos o primeiro valor da $k-PA$. Seja a o valor. Existem, no máximo, W opções. Agora, escolha um valor que difere d do ponto a. Desta forma, temos no máximo $W/(k-1)$ opções. Uma vez que estes estão determinados, k números distintos de mesma cor serão determinados. A saber:

$$\{a, a+d, a+2d, ..., a+(k-1)d\}.$$

Temos $W-k$ valores a esquerda. Então, o número de colorações é limitado por $cW^2c^{W-k}/k-1$. Desta forma, a probabilidade de uma $c-$coloração ser uma $k-PA$ monocromática é limitada por

$$\frac{cW^2c^{W-k}}{(k-1)c^W} = \frac{W^2}{(k-1)c^{k-1}}.$$

Este número é menor que 1, então a probabilidade de uma $k-PA$ não monocromática deve ser positiva, de modo que tal coloração deve existir. Temos então:

$$W^2 < (k-1)c^{k-1} \Rightarrow W < c^{(k-1)/2}\sqrt{k-1}.$$

Portanto, existe uma $c-$coloração de $(c^{(k-1)/2}-1)\sqrt{k-1}$ sem uma $k-PA$ monocromática. Então, $W(k,c) \geq c^{(k-1)/2}\sqrt{k-1}$. Note que a prova deste teorema foi não construtiva, uma vez que não foi construída uma coloração.

□

Uma vez clara a ideia central do Teorema de van der Waerden que consiste em colorir o conjunto dos números naturais, poderíamos tentar imaginar uma coloração de $\mathbb{N} \times \mathbb{N}$ ou de $\mathbb{N} \times \mathbb{N} \times \mathbb{N}$.

Definição 3.5 *Sejam $k, d \geq 1$ e $(a_0, a_1) \in \mathbb{Z} \times \mathbb{Z}$. Denotaremos por grade com canto (a_0, a_1) e diferença d o conjunto de pontos:*

$$k \times k = \{(a_0, a_1) + (id, jd) | 0 \leq i, j \leq k-1\}$$

Quando o canto não for especificado a grade será apenas referida como $k \times k$.

O seguinte resultado trata da forma multidimensional do Teorema de van der Waerden:

Teorema 3.4 *Para todo $c \in \mathbb{N}$ existe um inteiro $G = G(c)$ tal que para toda $c-$coloração de $G(c) \times G(c)$ existe um quadrado que possui os quatro lados com a mesma cor.*

3.2 Polinômio de van der Waerden

Atentando novamente para a essência do teorema de van der Waerden, que consiste na busca de progressões aritméticas de comprimentos k, arbitrários, é possível visualizar uma progressão aritmética com estrutura polinomial da seguinte forma:

$$a, a + p_1(d), ..., a + p_k(d)$$

onde $p_i(x) = ix$.

Denotaremos por $\mathbb{Z}$ o conjunto dos números inteiros e por $\mathbb{Z}[x]$ o conjunto dos polinômios com coeficientes inteiros e mediante este modo de ver a progressão aritmética, na forma polinomial, poderíamos idealizar que a conclusão do teorema de van der Waerden poderia ser equivalente a seguinte afirmação:

Afirmação: Para cada polinômio $p_1(x), ..., p_k(x) \in \mathbb{Z}[x]$ e para cada número natural c, existe um W tal que, para alguma $c-$coloração $\mathcal{X} : [W] \to [c]$ existem $a, d \in \mathbb{N}$ tais que:

$$\mathcal{X}(a) = \mathcal{X}(a + p_1(d)) = \mathcal{X}(a + p_2(d)) = ... = \mathcal{X}(a + p_k(d)).$$

Todavia, esta colocação é falsa, pois se considerarmos por exemplo $k = 1$, $p_1(x) = 1$ e $c = 2$ em qualquer $2-$coloração de $[W]$ há a aparição de pelo menos dois números consecutivos com cores distintas. O fato é que a mesma só será válida se os termos constantes de todos os polinômios forem nulos. Bergelson e Leibman foram os primeiros a provar esta afirmação. Logo, teorema polinomial equivalente ao teorema de van der Waerden seria:

Teorema 3.5 (Teorema Polinomial de van der Waerden) *Para cada polinômio*

$$p_1(x), ..., p_k(x) \in \mathbb{Z}[x], \text{ com } p_i(0) = 0 \text{ para todo } i = 1, ..., k$$

e para cada número natural c, existe um W tal que, para cada $c-$coloração $\mathcal{X} : [W] \to [c]$ existem $a, d \in \mathbb{N}$ tais que:

$$\mathcal{X}(a) = \mathcal{X}(a + p_1(d)) = \mathcal{X}(a + p_2(d)) = ... = \mathcal{X}(a + p_k(d)).$$

De forma análoga a versão combinatória do teorema de van der Waerden também define-se o número de van der Waerden nesta versão, porém estes números agora dependem dos polinômios $p_1, ..., p_k \in \mathbb{Z}[x]$, como segue a definição:

Definição 3.6 *Sejam $p_1, ..., p_k \in \mathbb{Z}[x]$ e $c \in \mathbb{N}$. $W(p_1, ..., p_k; c)$ é o menor W que satisfaz o Teorema Polinomial de van der Waerden, denominado: o número do polinômio de van der Waerden.*

O Teorema Polinomial de van der Waerden foi provado para $k = 1$ por Furstenberg (Ergodic behavior of diagonal measures and a theorem of Szemerédi's on arithmétic progressions) e Sarkozy (On difference sets of sequences of integers), independentes. A prova original deste resultado foi feita por Bergelson e Leibman usando métodos ergódicos. Também fora dada uma prova usando técnicas de combinatória, feita por Walters.

A existência de infinitas progressões polinomiais formadas por primos foi demonstrada por Tao e Ziegler, com a seguinte versão:

Teorema 3.6 *Sejam* $p_1, ..., p_k \in \mathbb{Z}[x]$, *onde* $p_i(m) = (i-1)m$ *são polinômios com valores inteiros. Dado* $\varepsilon > 0$ *existem infinitos inteiros* x *e* m *tais que* $1 \leq m \leq x(\varepsilon)$ *e* $x+p_1(m), ..., x+p_k(m)$ *são primos.*

Capítulo 4

Demonstração do Teorema de Szemerédi

Nesta seção daremos uma prova do Teorema de Szemerédi, o qual afirma que todo subconjunto de $\mathbb{Z}$ com densidade superior positiva contém progressões aritméticas de tamanhos arbitrários. Para tal, serão utilizados elementos da Teoria Ergódica, que segue a linha de raciocínio da prova dinâmica do teorema de van der Waerden, inclusive faz uso do mesmo dicionário entre partições de $\mathbb{Z}$ e sequências de inteiros.

Definição 4.1 *Chamaremos de intervalo do conjunto dos números inteiros $\mathbb{Z}$ qualquer subconjunto I da forma $\{n \in \mathbb{Z} : a \leq n \leq b\}$ com $a \leq b$. A cardinalidade do conjunto é o número $\#I = b - a$.*

Para enunciarmos o teorema de Szemerédi daremos uma definição de densidade superior de um subconjunto de $\mathbb{Z}$, equivalente a dada em 3.5.

Definição 4.2 *A densidade superior $D_s(S)$ de um subconjunto S de $\mathbb{Z}$ é o número*

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

onde I representa qualquer intervalo de $\mathbb{Z}$.

Em outras palavras, $D_s(S)$ é o maior número D tal que existe uma sequência de intervalos $I_j \subset \mathbb{Z}$ satisfazendo:

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

Exemplo 14 *Seja $S = \{n^2 : n \in \mathbb{N}\} = \{1, 4, 9, 16, 25, ...\}$ e I um intervalo de $\mathbb{Z}$ com cardinalidade n. Temos que*

$$D_s(S) = \limsup_{n \to \infty} \frac{\#\left\{j^2 : 1 \leq j^2 \leq n\right\}}{n} \leq \tfrac{\sqrt{n}}{n} \to 0, \text{ quando } n \to \infty.$$

Logo, $D_s(S) = 0$.

Exemplo 15 *Considere* $S = \{n \in \mathbb{N}$*: n é primo*$\}$ *e I um intervalo de* $\mathbb{Z}$ *com cardinalidade n. Temos que*

$$D_s(S) = \limsup_{n\to\infty} \frac{\#\{p-primo: 1 \leq p \leq n\}}{n} = \limsup_{n\to\infty} \frac{\pi(n)}{n} \to 0$$

onde, pelo Teorema Fundamental dos Números Primos, $\pi(n) \cong \frac{n}{\log n}$.

Exemplo 16 *Seja S o conjunto dos números pares. Dado qualquer intervalo I de* $\mathbb{Z}$*, temos que* $\#(S \cap I) = \#I/2$ *se o cardinal de I for par e* $\#(S \cap I) = (\#I \pm 1)/2$ *se o cardinal de I for ímpar, onde o sinal* $\pm$ *é positivo se o menor elemento de I for um número par e negativo caso contrário. Temos então* $D_s(S) = 1/2$.

Exemplo 17 *Seja S o seguinte subconjunto de* $\mathbb{Z}$*:*

$$S = \{1, 3, 4, 7, 8, 9, 13, 14, 15, 16, 21, 22, 23, 24, 25, 31, 32, 33, 34, 35, 26, 43...\}.$$

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 comprimentos arbitrariamente grandes . Logo,* $D_s(S) = 1$.

Note que nos dois últimos exemplos o conjunto S contém progressões aritméticas de tamanhos arbitrários , inclusive no exemplo 4.3 S contém progressões aritméticas de comprimento infinito o que não ocorre com S no exemplo 4.4.

Teorema 4.1 (Teorema de Szemerédi) *Se S é um subconjunto de* $\mathbb{Z}$ *com densidade superior positiva, então ele contém progressões aritmética*
s de tamanho arbitrário.

Demonstração Considere S um conjunto com densidade superior positiva, isto é, tal que, existe um $c > 0$ e intervalos $I_j = [a_j, b_j)$ de $\mathbb{Z}$ tais que

$$\lim_j \#I_j = \infty \text{ e } \lim_j \frac{\#S \cap I_j}{\#I_j} \geq c.$$

Associaremos a a S a sequência $\underline{\alpha} = (\alpha_j)_{j\in\mathbb{Z}} \in \Sigma = \{0, 1\}^{\mathbb{Z}}$ definida por:

$$\alpha_j = 1 \Leftrightarrow j \in S.$$

Considere o deslocamento $\sigma : \Sigma \to \Sigma$ e o subconjunto $A = \{\underline{\alpha} \in \Sigma : \alpha_0 = 1\}$ de Σ. Note que A é aberto e fechado pois trata-se de um cilindro de Σ. Observe também que , para qualquer $j \in \mathbb{Z}$,

$$\sigma(\underline{\alpha}) \in A \Leftrightarrow \alpha_j = 1 \Leftrightarrow j \in S.$$

Logo, para mostrar o Teorema de Szemerédi é suficiente mostrar que para todo $k \in \mathbb{N}$ existem $m \in \mathbb{Z}$ e $n \leq 1$ tais que

$$\sigma^{m+n}(\underline{\alpha}), \sigma^{m+2n}(\underline{\alpha}), ..., \sigma^{m+kn}(\underline{\alpha}) \in A$$

pois desta forma teríamos

$$\alpha_{m+n}, \alpha_{m+2n} = ... = \alpha_{m+kn} = 1$$

e assim

$$m + n, m + 2n, ..., m + kn \in S.$$

Para tal, considere a sequência

$$\mu_j = \frac{1}{\#I_j} \sum_{i \in I_j} \delta_{\sigma^i}(\underline{\alpha})$$

Como o conjunto $\mathcal{M}_1(\Sigma)$ das probabilidades em Σ é compacto, a menos de substituir $(\mu_j)_j$ por uma subsequência, podemos supor que ela converge na topologia fraca* para alguma probabilidade μ de Σ. Note que μ é uma probabilidade σ-invariante pois, para toda função contínua $\varphi : \Sigma \to \mathbb{R}$, vale

$$\int (\varphi \circ \sigma) d\mu_j = \frac{1}{\#I_j} \sum_{i \in I_j} \varphi(\sigma^i(\underline{\alpha})) + \frac{1}{\#I_j}[\varphi(\sigma^{b_j}(\underline{\alpha})) - \varphi(\sigma^{a_j}(\underline{\alpha}))]$$

$$= \int \varphi d\mu_j + \frac{1}{\#I_j}[\varphi(\sigma^{b_j}(\underline{\alpha})) - \varphi(\sigma^{a_j}(\underline{\alpha}))]$$

e, passando o limite quando $j \to \infty$, temos $\int (\varphi \circ \sigma) d\mu = \int \varphi d\mu$. Observe também que $\mu(A) > 0$. De fato, como A é fechado temos que:

$$\mu(A) \geq \limsup_j \mu_j(A) = \limsup_j \frac{\#(S \cap I_j)}{\#I_j} \geq c$$

. Dado qualquer $k \geq 1$, considere as transformações $f_i = \sigma^i$ para $i = 1, ..., k$. Claramente estas transformações comutam entre si. Então, pelo Teorema de Recorrência Múltipla de Poincaré, existe algum $n \geq 1$ tal que

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

Sendo A aberto temos que:

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

para qualquer l suficientemente grande. Pela definição de μ_l significa que existe algum $m \in I_l$ tal que

$$\sigma^m(\underline{\alpha}) \in A \cap \sigma^{-n}(A) \cap ... \cap \sigma^{-kn}(A).$$

Em particular, $\sigma^{m+in}(\underline{\alpha}) \in (A)$ para todo $i = 1, ..., k$ como queríamos provar.

□

O teorema de van der Waerden é um caso particular do teorema de Szemerédi. De fato, dada uma partição $S_1, ..., S_l$ de $\mathbb{Z}$ e sequências (x_{n_i}) em S_i, com $i = 1, ..., l$ temos:

$$\limsup (x_{n_1} + ... + x_{n_l}) \leq \limsup x_{n_1} + ... + \limsup x_{n_l}.$$

Assim,

$$D_s(\mathbb{Z}) \leq D_s(S_1) + ... + D_s(S_l).$$

Como,

$$D_s(\mathbb{Z}) = \limsup_{\#I \to \infty} \frac{\#(\mathbb{Z} \cap I)}{\#I} = \limsup_{\#I \to \infty} \frac{\#I}{\#I} = 1$$

onde $I \subset \mathbb{Z}$ representa qualquer intervalo dos inteiros, segue -se que:

$$1 \leq D_s(S_1) + ... + D_s(S_l).$$

Sabendo que a densidade superior é sempre um número maior ou igual a zero, segue-se desta desigualdade que existe um $j \in 1, ..., l$ tal que $D_s(S_j) > 0$, o que implica pelo teorema de Szemerédi que este S_j possui progressões aritméticas de tamanhos arbitrários.

Capítulo 5

Teorema de Green-Tao

O matemático inglês Ben Green e o matemático australiano Terence Tao formularam o espetacular resultado: *existem progressões aritméticas arbitrariamente longas formadas por números primos* [4]. O conjunto dos números primos não tem densidade superior positiva, portanto o teorema de Green-Tao não é consequência do teorema de Szemerédi. Por outro lado, o teorema de Green-Tao é um caso particular de outra conjectura devido a Erdos: se $S \subset \mathbb{N}$ é tal que a soma dos inversos diverge, ou seja, tal que:

$$\sum_{n \in S} \frac{1}{n} = \infty,$$

então S contém progressões aritméticas de qualquer comprimento. Este era um dos "prize money problems"de Paul Erdos pelo qual ele oferecia uma quantia em dinheiro para quem conseguisse resolver. Esta afirmação mais geral ainda permanece em aberto.

Salientaremos a seguir os passos fundamentais que Green e Tao desenvolveram para demonstrar o Teorema citado acerca da existência de progressões aritméticas nos primos, evitando definições formais e questões técnicas não triviais.

Trata-se de uma interessante demonstração resultante da fusão de de métodos e resultados de teoria dos números, teoria ergódica, análise harmônica, combinatória e geometria discreta. O ponto de partida é a demonstração do teorema de Szemerédi comentado no capítulo anterior. Todavia, o conjunto dos números primos não tem densidade superior positiva, portanto o teorema de Szemerédi não pode ser diretamente aplicado. Donde surge o segundo passo da demonstração de Green e Tao, usar um princípio da transferência o qual permite o uso do teorema de Szemerédi num contexto mais geral.

Assim, será possível deduzir do teorema de Szemerédi que qualquer subconjunto de um conjunto suficientemente $k-$pseudo-aleatório (que satisfazem uma condição de formas lineares e uma condição de correlação em termos do parâmetro k), de densidade relativa positiva contém progressões aritméticas de comprimento arbitrariamente longo.

O último ponto, consiste no uso de propriedades específicas dos números primos e da sua distribuição, baseados em resultados recentes de Goldston e Yildirim sobre o tamanho das lacunas nos primos [9], onde é mostrado que o Teorema de Szemerédi generalizado pode ser aplicado ao conjunto dos números primos.

Referências Bibliográficas

[1] B. Hasselblatt and A. Katok , *Introduction to the Modern Theory of Dynamical Systems*, volume 54 of Encyclopedia of Mathematics and its Applications, Cambridge University Press (1995),Whit a supplementary chapter by Katok and Leonardo Mendonza.

[2] K. Oliveira and M. Viana, *Foundations of Ergodic Theory*, Cambridge studies in advanced (2015).

[3] W. Gasarch Kruskel and P. Andy, *Van der Waerden's Theorem: Variants and Aplicati-ons*,210 (2012), no 1.

[4] B. Green and T. Tao, *The primos contain arbitrarily long arithmetic progressions*, Ann. of Math., 167:481-547 (2008).

[5] H. Fustemberg, *Ergodic behavior and a theorem of Szemerédi on arithmetic Progressions* . J. d' Analyse Math, 31:204-256 (1977).

[6] A. Arbieto, C. Matheus, C. Gustavo Moreira *Aspectos Ergódicos da Teoria dos Números*, IMPA, 26° Colóquio Brasileiro de Mtemática, 2007.

[7] P. Halmos, *Measure Theory*. D. Van Nostrand Company (1950).

[8] D. A. Goldston and C. Y. Yildirim, *Higher Correlations of Divisor Sums Related to primes, III*. Small gaps between primes (2003).