O Teorema de Green-Tao: Progressões Aritméticas de Tamanho Arbitrariamente Grande Formadas por Primos

Matheus Gonçalves Cassiano da Cunha

Dissertação apresentada
ao
Instituto de Matemática e Estatística
da
Universidade de São Paulo
para
obtenção do título
de
Mestre em Ciências

Programa: Matemática

Orientador: Prof. Dr. Manuel Valentim de Pera Garcia

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES

São Paulo, 23 de agosto de 2019

O Teorema de Green-Tao:
Progressões Aritméticas de Tamanho Arbitrariamente Grande Formadas por Primos

Esta versão da dissertação contém as correções e alterações sugeridas pela Comissão Julgadora durante a defesa da versão original do trabalho, realizada em 27/06/2019. Uma cópia da versão original está disponível no Instituto de Matemática e Estatística da Universidade de São Paulo.

Comissão Julgadora:

- Prof. Dr. Manuel Valentim de Pera Garcia - IME-USP
- Prof. Dr. Fábio Armando Tal - IME-USP
- Prof. Dr. Nicolau Corção Saldanha - PUC-Rio

Ao povo da cruz.

Agradecimentos

Agradeço primeiramente a Deus, pelas inúmeras bençãos dadas de maneira bondosa — e entre as quais estão incluídas todos os nomes que seguirão. Ao Rei eterno, imortal, invisível, Deus único, sejam honra e glória pelos séculos dos séculos. Amém!

Este não é um trabalho individual. Pode até ter sido que eu fui quem investi meu tempo e esforço no texto, mas cada pessoa aqui mencionada, e muitas outras que não o puderam ser, investiram tempo e esforço em mim. Eu tinha, ao menos, a vantagem de saber de antemão que os resultados com os quais trabalhava indubitavelmente dariam certo.

Obviamente, gostaria de registrar meu apreço por toda minha família, primeiramente a meus pais Enoc e Cristiane, que sempre me amaram e incentivaram, bem como para meus irmãos Leonardo, Marcos Paulo e Lucas, companheiros durante a vida, sempre tornando-a mais prazerosa e significativa. Mesmo que eles não entendam os teoremas, corolários, proposições e conjecturas enunciados aqui, tenho certeza que entendem o quanto foram, são e serão importantes para mim, de uma maneira que nem mesmo cardinais inacessíveis poderiam mensurar.

Uma outra família, unida por outro sangue, merece meus agradecimentos. A todos os meus irmãos da amada Igreja Batista Redenção, que me ensinaram verdades ainda mais belas e mais profundas do que as descritas nesta dissertação. Em especial, a cada um dos jovens do Darash e dos homens do REMAR, do RELP e, claro, do "Vidas...". Agradeço de coração toda a comunhão, instrução e amizade que recebo continuamente.

Agradeço ao meu orientador Prof. Dr. Manuel Valentim de Pera Garcia, o famosíssimo Mané, por todo o ensino, paciência, humor e amizade em todos estes anos (e, claro, pelas ótimas pizzas nas noites de sexta-feira). Adicionalmente, não poderia me esquecer de reconhecer o quanto a Prof. Dra. Sônia Regina Leite Garcia, minha "co-orientadora", me ajudou durante meus estudos. Espero ter tornado todos estes momentos especiais para ambos, assim como eles os tornaram para mim.

Não poderia esquecer de meus amigos e colegas do IME-USP, que durante minha graduação e mestrado tornaram o processo todo mais divertido. Nomes como Luciana, Marisa, Karina e Felipe se sobressaem durante todos estes anos em que lutamos (e vencemos!) juntos.

Gostaria de reconhecer também o trabalho dos funcionários do IME-USP, sempre resolvendo os problemas com os quais eu surgia para eles (matemáticos por vezes esquecem que nem todos *gostam* de problemas).

Agradeço aos membros da banca, pela disponibilidade e prontidão em participar deste momento singular para mim.

Muitas das idéias deste texto — certo, eu confesso: a *maioria* delas — foram realizadas durante meu trajeto diário. Devo dizer obrigado às companhias de ônibus e metrô por proporcionarem um ambiente tão propício ao pensamento acadêmico (e pessoal também).

Por fim, agradeço a CAPES pelo apoio financeiro. Realmente foi de grande ajuda.

A man has to live with himself, and he should see to it that he always has good company.

Charles Evans Hughes
(estadista americano)

Resumo

CUNHA, M. G. C. da **O Teorema de Green-Tao: Progressões Aritméticas de Tamanho Arbitrariamente Grande Formadas por Primos**. 2019. 61 f. Dissertação (Mestrado) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2019.

Encontrar subestruturas aditivas que revelam um certo grau de organização em certos conjuntos contidos nos números naturais é o foco do estudo da combinatória aditiva. Desta área, resultados como os famosos Teorema de Van der Waerden e o Teorema de Szemerédi se destacam, revelando através de métodos combinatoriais que certas propriedades referentes ao "tamanho" de subconjuntos de inteiros implicam a existência de progressões aritméticas de tamanho arbitrariamente grande.

Em meados de 1970, Furstenberg causou certa comoção no meio matemático ao publicar provas para ambos os teoremas usando métodos e ferramentas da teoria ergódica.

Apesar de tal abordagem ter apresentado uma nova e profunda ligação entre as áreas, houve certa crítica pelo fato de não gerar resultados originais e por suas limitações (por exemplo, seus resultados costumam ser de caráter assintótico, sem lidar com limitantes e cotas, amplamente conhecidos pelos métodos combinatórios).

Tais críticas foram silenciadas quando Ben Green e Terence Tao, usando tais métodos de teoria ergódica, demonstraram a incrível e bela afirmação de que os primos possuem progressões aritméticas de tamanho arbitrariamente grande, dando uma resposta definitiva para um enunciado conjecturado há muito tempo. Certamente, este foi um grande passo na matemática do século XXI.

Deste então, novas abordagens foram amplamente estudadas e analisadas, de modo a aumentar ainda mais nossa compreensão sobre estes impressionantes conceitos.

Palavras-chave: Teorema de Green-Tao, progressões aritméticas, primos.

Abstract

CUNHA, M. G. C. da **The Green-Tao Theorem: arbitrarily long arithmetic progressions on primes**. 2019. 61 p. Thesis (Master's Degree) - Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2019.

Finding additive substructures that reveal a certain degree of organization in certain sets contained in the set of the natural numbers is the focus of the study of additive combinatorics. From this area, results such as the famous Van der Waerden's Theorem and Szemerédi's Theorem stand out, revealing through combinatorial methods that certain properties concerning the "size" of subsets of integers imply the existence of arbitrarily long arithmetic progressions.

In the mid-1970s Furstenberg caused some commotion in the mathematical world by publishing proofs for both theorems using methods and tools of ergodic theory rather than combinatorial methods.

Although this approach had presented a new and deep link between those areas, there was some criticism for the lack of original results and some limitations of this technique (for instance, its results usually have an asymptotic flavour without dealing with bounds widely known by combinatorial methods).

Such criticisms were silenced when Ben Green and Terence Tao, using such methods of ergodic theory, demonstrated the incredible and beautiful theorem that the primes have arithmetic progressions of arbitrarily large size, giving a definitive answer to a statement conjectured a long time ago. Certainly, this was a major step for the mathematics of the 21st century.

Hence, new approaches have been extensively studied and analyzed in order to further increase our understanding of these impressive concepts.

Keywords: Green-Tao Theorem, arithmetic progressions, primes.

Sumário

Definições, Notações e Propriedades Básicas

Neste capítulo, definiremos alguns objetos básicos e veremos algumas de suas propriedades que serão usados durante o resto do texto.

Definições e Notações

Como comumente ocorre, consideramos $\mathbb{Z}_N := \mathbb{Z}/N\mathbb{Z}$ como o *anel dos inteiros módulo* N. Durante os capítulos iniciais, N assumirá o papel de qualquer inteiro estritamente positivo, mas nos capítulos específicos sobre o Teorema de Green-Tao, N será sempre um número primo (suficientemente) grande — e, portanto, poderemos lidar com inversos multiplicativos de elementos não-nulos sem maiores preocupações.

Dado $A \subseteq X$, a *função característica de* A, denotada por $\chi_A : X \to \{0,1\}$, é definida por $\chi_A(x) = 1$, se $x \in A$, e $\chi_A(x) = 0$, se $x \notin A$. Naturalmente, escreveremos $\chi_{\{a\}}$ como χ_a.

Dado um conjunto A finito, $|A|$ é a *quantidade de elementos de* A, ou, mais explicitamente, $|A| := \sum_{x \in A} 1$.

Sendo $X = \mathbb{R}$ ou $X = \mathbb{Z}_N$, consideremos $\emptyset \neq A \subseteq X$ um subconjunto finito, assim como uma função $f : A \to \mathbb{R}$. Podemos então definir o funcional *esperança de* f *em relação a* A ou *valor esperado de* f *em relação a* A, dado por

$$\mathbb{E}(f(x) : x \in A) = \mathbb{E}(f : A) := \frac{\sum_{x \in A} f(x)}{|A|}.$$

Quando não for causa de confusão, escreveremos simplesmente $\mathbb{E}(f)$.

Usaremos constantemente a abreviação *PA* para nos referirmos a progressões aritméticas no geral, e *k-PA* para progressões aritméticas de tamanho k (isto é, formadas por k elementos).

Se $r \neq 0$, para nos referirmos à k-PA formada por $\{a, a+r, a+2r, \ldots, a+(k-1)r\}$ usaremos $\{a + [0,k)r\}$ (note que esta PA possui, de fato, k elementos).

Estaremos muitas vezes interessados em mensurar a quantidade de k-PAs "bem-distribuídas" em relação aos suportes de uma k-úpla $(f_0, f_1, \ldots, f_{k-1})$ de funções de $\mathbb{Z}_N$ em $\mathbb{R}$. Para realizarmos esta análise, consideramos o operador *contagem normalizada de k-PAs em* $f_0, f_1, f_2, \ldots, f_{k-1}$, definido por

$$\Upsilon_k(f_0, f_1, \ldots, f_{k-1}) := \mathbb{E}(f_0(x) f_1(x+r) \cdots f_{k-1}(x+(k-1)r) : x, r \in \mathbb{Z}_N).$$

No caso de $f_0 = f_1 = \ldots = f_{k-1} = f$, escreveremos $\Upsilon_k(f)$ para denotar $\Upsilon_k(f, f, \ldots, f)$.

Dado A um subconjunto não-vazio de $\mathbb{Z}_N$ e $f_0, f_1, \ldots, f_{k-1}$ funções reais com suporte contido em A, note que o elemento $x + jr$ da k-PA $\{x + [0,k)r\}$ não pertence ao suporte de alguma f_j se, e somente se, o produto $f_0(x) \cdots f_{k-1}(x+(k-1)r)$ é zero. Isso justifica o nome dado a $\Upsilon_k(f_0, \ldots, f_{k-1})$. Também vale a pena ressaltar que $\Upsilon_k(\chi_A)$ resulta na probabilidade de, dados $x, r \in \mathbb{Z}_{\mathbb{N}}$, o conjunto $\{x + [0,k)r\}$ definir uma k-PA formada exclusivamente por elementos de A.

Além disso, uma função $\phi : A \to B$, onde $A \subseteq \mathbb{Z}_N^n, B \subseteq \mathbb{Z}_N^m$, é dita uma *cobertura uniforme de* B *por* A se ϕ é sobrejetora e $|\phi^{-1}(b)| = |A|/|B|$, para todo $b \in B$.

Ainda mais, usaremos as consagradas notações $o(1)$ para um termo que tende a 0, quando N tende a $+\infty$, e $O(1)$ para um termo que permanece limitado sob as mesmas condições. Por fim, usaremos as abreviações $o(X) = o(1)X$ e $O(X) = O(1)X$ e, quando a convergência ou o fator limitante dependerem de certos parâmetros (digamos, m e q), tais parâmetros serão indicados nos subscritos, como $o_m(1)$ (dependência em apenas um dos parâmetros) ou $O_{m,q}(X)$ (dependência em ambos os parâmetros).

Por fim, uma última convenção que usaremos: sempre escreveremos as expressões de forma que os expoentes nas funções seguirão o seguinte padrão: $f^n(x) := \underbrace{(f \circ \ldots \circ f)}_{n \text{ vezes}}(x)$, enquanto que $f(x)^n := \underbrace{(f \cdot \ldots \cdot f)}_{n \text{ vezes}}(x)$.

Propriedades

Eis algumas propriedades básicas que usaremos durante o texto:

Proposição 0.0.1 (Propriedades da Esperança)**.** *Dados $A \subseteq \mathbb{Z}_N^m, B \subseteq \mathbb{Z}_N^n$ dois conjuntos não-vazios, consideremos $f : A \to \mathbb{R}$ e $g : B \to \mathbb{R}$. Temos que:*

1. *Se $m = n$, $A = B$, e dados $a, b \in \mathbb{R}$, então*

$$\mathbb{E}(af(x) + bg(x) : x \in A) = a\mathbb{E}(f(x) : x \in A) + b\mathbb{E}(g(x) : x \in A);$$

2. *Vale que*

$$\begin{aligned}\mathbb{E}(f(x)g(y) : x \in A, y \in B) &= \mathbb{E}(f(x)\mathbb{E}(g(y) : y \in B) : x \in A)\\ &= \mathbb{E}(f(x) : x \in A) \cdot \mathbb{E}(g(y) : y \in B);\end{aligned}$$

   Em particular, obtemos

$$\mathbb{E}(f(x) : x \in A) = \mathbb{E}(f(x) : x \in A, y \in B);$$

3. *Se $\phi, \psi : A \times B \to \mathbb{R}$, então*

$$\mathbb{E}(\phi(x,y)\psi(x,y) : x \in A, y \in B) = \mathbb{E}(\mathbb{E}(\phi(x,y)\psi(x,y) : y \in B) : x \in A);$$

4. *Se I for um conjunto finito de índices e tivermos $h_i : A \to \mathbb{R}$, para todo $i \in I$, temos que*

$$\mathbb{E}\left(\prod_{i\in I} h_i(x) : x \in A\right)^2 = \mathbb{E}\left(\prod_{i\in I} h_i(x)h_i(y) : x, y \in A\right);$$

5. *Se $\phi : A \to B$ é uma cobertura uniforme de B por A e $f : B \to \mathbb{R}$, temos que*

$$\mathbb{E}(f(\phi(a)) : a \in A) = \mathbb{E}(f(b) : b \in B).$$

Capítulo 1

Introdução

É inegável que os números primos despertaram, e ainda despertam, uma espécie de fascínio e desafio à humanidade de maneira singular. Desde o começo da Matemática Dedutiva, em *Os Elementos*, de Euclides, por volta de 300 a. C., até o século XXI, eles foram objetos de especial estudo e peças fundamentais em áreas de estudo como Aritmética, Álgebra, Análise Real e Complexa, Análise Funcional, Teoria Ergódica, Topologia, Lógica, Ciências da Computação e Criptografia, para citar as mais proeminentes.

Entre tantas propriedades dos números primos, duas notadamente foram alvo de um interesse particular entre matemáticos das mais diversas eras: sua *distribuição* e suas *propriedades aditivas*.

O primeiro questionamento surge de maneira natural, dada a importância que os números primos possuem, sendo os "blocos fundamentais" na construção dos números naturais do ponto de vista multiplicativo.

Sobre a distribuição dos primos, desde os primórdios da Matemática temos resultados importantíssimos, como o Teorema de Euclides[1] (em que ele mostra que o conjunto dos primos é infinito), até resultados mais modernos, como o Teorema do Número Primo.

Já quando tratamos das propriedades aditivas, é surpreendente a facilidade com que podemos apresentar problemas com enunciados simples e que, no entanto, resistem a séculos de esforços matemáticos; para corroborarmos isso, basta lembrarmo-nos de conjecturas como a dos Primos Gêmeos ou a de Goldbach.

Apesar de não terem sido completamente respondidas, tais perguntas impulsionaram vários desbravamentos em diversas áreas da Matemática, e tais esforços geraram uma enormidade de resultados e conhecimentos valiosos.

De um modo especial, um assunto une ambos os questionamentos sobre a distribuição e as propriedades aditivas dos números primos de um modo bastante interessante: o que podemos dizer sobre *progressões aritméticas e sua relação com os números primos*?

Com não pouca aclamação, a comunidade matemática do século XXI presenciou como resposta o seguinte

Teorema 1.0.1 (Teorema de Green-Tao)**.** *Existem progressões aritméticas de tamanho arbitrariamente grande formadas exclusivamente por primos.*

Este teorema é o resultado de uma seqüência de resultados que tratam de Combinatória Aditiva, analisando estruturas que levam em conta as propriedades aditivas de seus elementos.

Agora, procuraremos entender o papel de cada uma das peças deste grandioso resultado.

1.1 Progressões Aritméticas

A relativa simplicidade da definição de progressões aritméticas esconde a profundidade de conceitos relacionados. Não por acaso, tais relações já foram amplamente investigadas, sobre vários

[1] É de se admirar que um resultado incrível como este já esteja presente tão cedo na História da Matemática.

pontos de vista diferentes, por matemáticos como Dirichlet, Lagrange, Waring, Vinogradov, entre vários outros.

De todos os resultados sobre progressões aritméticas, porém, um dos mais fundamentais é o

Teorema 1.1.1 (Szemerédi). *Seja $A \subseteq \mathbb{N}$, de modo que $\limsup_{n\to\infty} \frac{|A \cap \{1,2,3,\ldots,n\}|}{n} > 0$.*
Então existem infinitas k-PAs em A, para qualquer $k \in \mathbb{N}$.

De fato, este teorema é central para o nosso estudo, dado que a estratégia é provar uma generalização dele que se aplica aos números primos.

Especificando um pouco mais no assunto do nosso interesse, dada a estrutura fortemente organizada das PAs e a aparente aleatoriedade dos números primos, ligações referentes às propriedades de primalidade são um tanto quanto surpreendentes.

A título de curiosidade, exemplificamos dois resultados clássicos sobre a relação de primos e PAs:

Teorema 1.1.2 (Dirichlet). *Sejam $a, b \in \mathbb{N}$ tal que $mdc(a, b) = 1$.*
Então a progressão aritmética $a + nb$ possui infinitos números primos.

Teorema 1.1.3 (Balog). *Dado $m \in \mathbb{N}$, existem $p_1, p_2, \ldots, p_m$ primos distintos tais que todas as médias $(p_i + p_j)/2$ são números primos.*

1.2 Uma Abordagem Probabilística Sobre a Distribuição os Primos

Um dos principais resultados sobre a distribuição dos primos é o seguinte

Teorema 1.2.1 (Teorema do Número Primo). *Seja $\pi(x)$ a função que determina a quantidade de números primos p com $2 \leq p \leq x$.*
Então $\lim_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1$.

Tal resultado proporcionou uma abordagem distinta da comumente usada no estudo dos números primos até então quando, além de considerar apenas a distribuição exata dos primos no conjunto dos inteiros, ele também permitiu expressar esta distribuição como um modelo probabilístico de pertinência, com probabilidade $1/\log(x)$. Deste modo, de um ponto de vista probabilístico, dado o conjunto $\{1, 2, 3, \ldots, n\}$, ao escolhermos ao acaso um elemento, ele possui $1/\log(n)$ de chance de ser primo (em caráter assintótico).

Apesar de tal ponto de vista ser uma grande simplificação, podemos perceber várias propriedades sobre os primos, como o fato de que eles vão se tornando cada vez mais raros, além de serem um subconjunto relativamente pequeno dos naturais ($\pi(x)/x$ tende a 0, por exemplo). Além disso, conseguimos uma ajuda valiosa na hora de fazer conjecturas e heurísticas.

Olhemos para o Teorema do Número Primo em Progressão Aritmética (Teorema 2.4.3), por exemplo: ele nos diz que $\sum_{p\leq x,\ p\equiv a \mod b} \log p = \frac{x}{\phi(b)}(1 + o(1))$, o que pode ser visto como um resultado claro de modelar os primos com uma distribuição uniforme com valor $1/\log(n)$ dada pelo Teorema do Número Primo juntamente com o fato elementar de que as classes dos resíduos se dispersam com uma concentração de fator $1/\phi(b)$.

Apesar disso, esta é uma abordagem que deve ser realizada com cuidado, já que despreza várias sutilezas da distribuição dos primos.

Caso contrário, poderíamos supor que p e $p+2$ possuem, individualmente e independentemente, uma chance de $1/\log(n)$ de serem primos, logo a probabilidade de ambos serem *simultaneamente* primos seria de $1/\log(n)^2$. Fazendo a contagem até n, obteríamos $\sum_{k=1}^{n} \frac{1}{\log(n)^2} = \frac{n}{\log(n)^2}$, que diverge quando n tende a infinito e, portanto, parece que ficou demonstrada a existência de infinitos pares de primos gêmeos (nada mal, hein?).

Porém, se olharmos mais atentamente, veremos que poderíamos deduzir igualmente que existem infinitos pares de primos gêmeos *pares*, já que bastaria adicionarmos um fator multiplicativo de $1/2$, o que não afetaria a divergência da expressão. Mas isso claramente é um absurdo!

O erro está em considerar que p e $p+2$ serem primos são eventos independentes. Se p for par, por exemplo, isto afeta diretamente o evento "$p+2$ é primo".

O Teorema do Número Primo é, indubitavelmente, um grande resultado, porém ele nos dá informações apenas sobre questões de *densidade* dos primos, mas não oferece informação direta sobre questões aritméticas, e é por isso que a "demonstração" acima está errada, e que temos que lidar com mais cuidado quando nos deparamos com questões como estas.

1.3 Uma Idéia Surge...

Se estivermos atentos e utilizarmos um pouco de imaginação (qualidades indispensáveis aos matemáticos), podemos notar que o Teorema de Szemerédi (Teorema 1.1.1) lida com a noção de densidade de um conjunto, juntamente com a existência de certa estrutura aritmética em tal conjunto.

Além disso, a noção de densidade costuma lidar bastante com conceitos de Probabilidade e de Teoria da Medida mesma maneira, a seção anterior tratou diretamente de uma abordagem probabilística dos primos inicialmente proporcionada pelo Teorema do Número Primo (Teorema 1.2.1).

Será que, então, poderíamos usar ambas tais idéias de uma maneira criativa (mais uma qualidade necessária aos matemáticos) para aumentarmos nosso conhecimento sobre progressões aritmética em primos?

1.4 Uma (Rápida) Jornada Histórica

A resposta é **sim**, e um pouco de História da Matemática nos ajudará a seguir a linha de pensamento dos conceitos que serão explicados no restante do texto.

A questão de lidar com progressões aritmética nos primos toca profundamente num ramo da Matemática conhecido como *Combinatória Aditiva*, que que procura estruturas e seus relacionamentos com as propriedades aditivas de objetos combinatórios.

Um dos primeiros e principais resultados da área foi o chamado Teorema de Van der Waerden, enunciado pela primeira vez em [vdW27], relacionando a existência de PAs arbitrariamente grandes nos naturais com uma partição finita destes.

Algum tempo depois, respondendo a uma conjectura de vários anos, o Teorema de Szemerédi é publicado em [Sze75]. Tal teorema acaba por generalizar o resultado de Van der Waerden, provendo uma relação entre a existência de k-PAs, para qualquer k, de um subconjunto dos naturais com uma noção adequada de densidade (que é satisfeita quando é realizada uma partição finita).

Alguns anos se passaram, até que Furstenberg realizou um incrível feito ao publicar, em [Fur81], provas de natureza ergódica a ambos os teoremas, apontando uma ligação entre as áreas.

Esta e outras ligações continuaram a ser estudadas. Havia de certa parte um ceticismo, dado que nenhum resultado novo tinha sido produzido, e também pelo fato de que a abordagem ergódica costumava não fornecer informações conhecidas por outros métodos, tais como cotas superiores para as quantidades envolvidas.

O estudo, no entanto, mostrou-se frutífero de um modo extraordinário, respondendo a uma pergunta que por séculos já estava despertando o interesse dos matemáticos.

Após o resultado, Gowers publicou mais uma prova em [Gow01], esta usando elementos de Análise de Fourier, mostrando que o resultado de Szemerédi possui uma característica de conexão entre distintas áreas da Matemática.

Obviamente, no ínterim entre cada um destes resultados notáveis, diversos outros teoremas eram publicados e aumentavam o entendimento sobre tais questões. A título de exemplo, dadas as mesmas hipóteses usadas por Szemerédi, um teorema de Roth ([Rot53]) já enunciava a existência

de progressões aritméticas compostas por três elementos, e um resultado de Van der Corput já assegurava a existência de progressões aritméticas formadas por três elementos primos ([VdC39]).

Por fim, em [GT04], Been Green e Terence Tao, usando a abordagem ergódica, responderam de maneira afirmativa o questionamento sobre a existência de progressões aritméticas arbitrariamente longas formadas exclusivamente por primos.

Logo em seguida, generalizações do teorema (em versões multidimensionais ou com progressões polinomiais, por exemplo), aumentando ainda mais o conhecimento matemático sobre tais estruturas.[2]

1.5 Objetivo e Estrutura do Texto

O objetivo deste texto é apresentar uma explicação de um dos maiores avanços que a área de Combinatória Aditiva experimentou recentemente: o Teorema de Green-Tao.

A estrutura do texto começa com uma introdução seguindo uma abordagem histórica, apresentando os primeiros teoremas citados anteriormente (Van der Waerden e Szemerédi), e demonstrando-os com as ferramentas desenvolvidas por Furstenberg. Após isso, adentramos de fato nas definições e resultados necessários para o teorema central do texto. Apesar de começarmos seguindo de maneira bem próximo o artigo de Green e Tao, vale ressaltar que aproximadamente na metade da exposição nos guiaremos por um ponto de vista mais próximo ao utilizado por Gowers (explicitamente, toda a parte que utiliza Análise Funcional). Por fim, os resultados sobre primos estão também presentes no artigo original de Green e Tao.

Apesar do intuito do texto de prover um desenvolvimento claro e sistemático sobre o assunto, procurando construir metodicamente os conceitos, definições e teorias necessários, e priorizando a didática e compreensão, estes mesmos paradigmas forçaram certos limites em seu desenvolvimento.

Assim sendo, a fim de termos um guia compreensível e relativamente simples ao entendimento do Teorema de Green-Tao, certas diretrizes foram adotadas: assumiremos resultados conhecidos e de fácil acesso, somente dando uma explicação heurística sobre seus papéis, e indicando outras fontes ao leitor interessado; e também, durante certos momentos em que a explanação de certos conceitos partindo de suas bases causaria uma digressão tão grande que destruiria o fluxo de idéias e se chocaria com um dos propósitos principais do texto.

Além disso, algumas discussões de caráter técnico são essenciais para o desenvolvimento das idéias, e enquanto que inseri-los diretamente causaria uma ruptura na sucessão de idéias, sua omissão seria negligenciosa. Tais discussões foram adicionadas como apêndices.

O texto, esperamos, não deve apresentar muitas dificuldades para alguém com conhecimentos comumente obtidos numa graduação em Matemática, aliado a um possível primeiro curso em Sistemas Dinâmicos e familiaridade com resultados amplamente conhecidos.

[2] Tais generalizações são rapidamente discutidas no Capítulo 5

Capítulo 2

Ferramentas

Neste capítulo, apresentaremos as ferramentas matemáticas necessárias para o desenvolvimento do restante do texto.

2.1 Análise Funcional

A Análise Funcional lida com espaços normados e a análise de suas funções. Nosso foco será principalmente estudar os espaços $L^q(\mathbb{Z}_N)$, para $p \geq 1$, isto é, o conjunto das funções $f : \mathbb{Z}_N \to \mathbb{R}$ com as normas q dadas por

$$\|f\|_q := \mathbb{E}(|f|^q)^{1/q},$$

além do caso $L^\infty(\mathbb{Z}_N)$, com a norma

$$\|f\|_\infty := \sup_{x \in \mathbb{Z}_N} |f(x)|.$$

Além disso, vale lembrar que podemos tomar $L^2(\mathbb{Z}_N)$ como um espaço de Hilbert, se considerarmos no conjunto das funções $f : \mathbb{Z}_N \to \mathbb{R}$ o produto interno $\langle f, g \rangle \mapsto \mathbb{E}(fg)$.

Esta visão nos possibilita utilizar a desigualdade triangular, a desigualdade de Jensen, bem como as desigualdades de Hölder e de Cauchy-Schwarz para médias:

Teorema 2.1.1 (Desigualdade de Cauchy-Schwarz para médias)**.** *Sejam $f, g : A \to \mathbb{R}$ funções a valores reais, com $A \subseteq \mathbb{Z}_N$. Então*

$$\mathbb{E}(|f(x)g(x)| : x \in A)^2 \ \leq \ \mathbb{E}(f(x)^2 : x \in A) \cdot \mathbb{E}(g(x)^2 : x \in A).$$

Teorema 2.1.2 (Desigualdade de Hölder para médias)**.** *Sejam $f_j : A \to \mathbb{R}, j = 1, 2, \ldots, m$, com $A \subseteq \mathbb{Z}_N$, e sejam $q_1, q_2, \ldots, q_m > 0$ e $q \geq 1$ tais que $\dfrac{1}{q_1} + \dfrac{1}{q_2} + \cdots + \dfrac{1}{q_m} = \dfrac{1}{q}$.*

Então

$$\mathbb{E}\left(\prod_{j=1}^{m} |f_j(x)|^q : x \in A\right)^{1/q} \leq \prod_{j=1}^{m} \mathbb{E}(|f_j(x)|^{q_j} : x \in A)^{1/q}.$$

Teorema 2.1.3 (Desigualdade de Jensen)**.** *Seja $f : A \to \mathbb{R}$, com $A \subseteq \mathbb{Z}_N$. Se $\phi : \mathbb{R} \to \mathbb{R}$ é uma função convexa*[1]*, então*

$$\phi(\mathbb{E}(f(x) : x \in A)) \leq \mathbb{E}(\phi(f(x)) : x \in A).$$

Um dos principais teoremas da análise funcional é o chamado Teorema de Hahn-Banach, que, entre as várias maneiras em que pode ser enunciado, possui a seguinte formulação, com um toque bastante geométrico:

[1] Dado U um conjunto convexo, uma função $f : U \to \mathbb{R}$ é dita *convexa* se, para todos $x, y \in U$ e para qualquer $t \in [0, 1]$, temos que a desigualdade $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$ é satisfeita.

Teorema 2.1.4 (Hahn-Banach). *Seja K um corpo convexo em $\mathbb{R}^n$ (isto é, um subconjunto compacto convexo de $\mathbb{R}^n$ com interior não-vazio), e seja $x \notin K$.*[2]

Então existem uma constante β e um funcional linear não-nulo ϕ tais que $\phi(x) \geq \beta$ e $\phi(y) \leq \beta$, para todo $y \in K$.

Em especial, uma versão específica do Teorema de Hahn-Banach, que apresentaremos a seguir, será nossa ferramenta principal, originalmente usada em [Gow10]. Nesse artigo, Gowers expõe decomposições como um modo de realizar certas transferências de propriedades entre objetos, já que usualmente uma decomposição é realizada de modo que um termo contenha o atributo requerido e o outro seja, de certo ponto de vista, negligenciável.

Para o próximo resultado, é interessante relembrarmos o *Teorema da Representação de Riesz*, que nos garante que, dado um funcional linear ϕ em $\mathbb{R}^n$ (equipado com um produto interno $\langle \cdot, \cdot \rangle$), temos que existe um vetor $\tilde{\phi} \in \mathbb{R}^n$ tal que $\phi(x) = \left\langle \tilde{\phi}, x \right\rangle$, para todo $x \in \mathbb{R}^n$.

Teorema 2.1.5 (Hahn-Banach). *Sejam K_1, K_2 subconjuntos fechados e convexos de $\mathbb{R}^N$, ambos contendo a origem, e suponha que $x \notin K_1 + K_2$ (isto é, $x \notin \{k_1 + k_2 : k_1 \in K_1, k_2 \in K_2\}$).*

Então existe $\phi \in \mathbb{R}^N$ tal que $\langle \phi, x \rangle > 1$ e $\langle \phi, y \rangle \leq 1$, para todo $y \in K_1 \cup K_2$.

Demonstração. Seja K o corpo convexo dado por $K_1 + K_2$. Como K é fechado e, por hipótese, $x \notin K$, temos que existe um $\varepsilon > 0$ tal que $(1+\varepsilon)^{-1}x \notin K$.

Logo, pelo Teorema de Hahn-Banach (Teorema 2.1.4), identificando o funcional obtido com o vetor dado pelo Teorema da Representação de Riesz, temos que existe um vetor ϕ e uma constante β tal que $(1+\epsilon)^{-1} \langle \phi, x \rangle \geq \beta$ e $\langle \phi, y \rangle \leq \beta$, para qualquer $y \in K$, e em particular, sempre que $y \in K_i, i = 1, 2$.

Como $0 \in K$, temos que $\beta > 0$, e dividindo ϕ por β, vemos que podemos tomar a constante como sendo 1, com o resultado de que $\langle x, \phi \rangle \geq (1+\varepsilon)\beta > 1$.

Por fim, como 0 pertence a cada um dos K_i's, podemos concluir que $\langle y, \phi \rangle \leq 1$, para todo $y \in K_1 + K_2$.

□

A partir deste ponto, sempre que for feita uma referência ao Teorema de Hahn-Banach, estaremos nos refererindo à última versão, o Teorema 2.1.5, salvo menção contrária.

2.2 Sistemas Dinâmicos

A área de Sistemas Dinâmicos estuda as propriedades de um espaço e uma transformação dele em si mesmo, bem como sua evolução pela suas iteradas (no caso discreto) ou pelo fluxo (no caso contínuo).

Mais especificamente no caso discreto, que é o que usaremos, é o estudo de um conjunto X, geralmente munido com alguma propriedade métrica ou topológica, relacionada à medida (ou alguma outra propriedade que permita o estudo de resultados mais contextuais), juntamente com uma transformação $T : X \to X$ que possui ligação com a propriedade do conjunto X (contrações em espaços métricos, continuidade em espaços topológicos e preservação de medida em espaços com medida, etc), para assim obtermos resultados sobre a *órbita de um elemento x por T*, isto é, o conjunto $\mathcal{O}_T(x) := \{T^n(x) : n \in \mathbb{N}\}$ (convencionamos que $T^0 = Id_X$), bem como sobre o comportamento assintótico de T.

O exemplo mais relevante, em nosso caso, é o da *dinâmica simbólica*, onde dado um *alfabeto* X (um conjunto não-vazio qualquer) consideramos as *palavras infinitas de X* (seqüências de símbolos de X indexadas em $\mathbb{N}$), e estudamos o espaço $\Omega(X) := \{(x_n)_{n\in\mathbb{N}} : x_n \in X\}$ (também conhecido como *dicionário de X*), juntamente com a aplicação *shift* (comumente denotada por σ), definida por $\sigma(x_1, x_2, x_3, \ldots) = (x_2, x_3, x_4, \ldots)$.

[2] Não confundir com o conceito algébrico de *corpo*. No inglês, tal confusão não ocorre, dado que corpo convexo é denominado por "*convex body*", enquanto a noção mais comum de corpo é denominada por "*field*".

Podemos tornar $\Omega(X)$ um espaço métrico, já que dados $x = (x_n)_{n\in\mathbb{N}}$ e $y = (y_n)_{n\in\mathbb{N}}$ elementos de X, temos que a função $d : X \times X \to \mathbb{R}^+$ definida por $d(x, y) = 0$, se $x = y$, e $d(x, y) := 1/l$, se $x \neq y$, onde l é o menor inteiro tal que $x_l \neq y_l$, satisfaz os axiomas de métrica.

Em nosso caso, tomaremos um conjunto $A = \{a_1, a_2, a_3, \ldots, a_n\}$ finito como alfabeto. Nestas condições, temos que $\Omega(A)$ com a métrica d é um conjunto compacto, e, além disso, nesta métrica σ é uma função contínua.

O seguinte teorema será usado repetidamente no texto. A demonstração deles pode ser encontrada em [Fur81]:

Teorema 2.2.1 (Recorrência Múltipla Topológica de Furstenberg-Weiss)**.** *Sejam X um espaço métrico compacto e $T : X \to X$ uma função contínua.*

Então para todos $k \in \mathbb{N}$ e $\varepsilon > 0$, existem $x \in X$ e $n \in \mathbb{N}$ tais que $d(T^{jn}(x), x) < \varepsilon$, para $j = 1, 2, \ldots, k$.

Além disso, dado $Y \subseteq X$ denso, podemos tomar $x \in Y$.

2.3 Teoria Ergódica

A Teoria Ergódica une as áreas de Sistemas Dinâmicos e Teoria da Medida, ao estudar sistemas dinâmicos munidos de uma medida que é invariante pela transformação que caracteriza o sistema.

Mais detalhadamente, é o estudo da estrutura formada pela quádrupla $(X, \mathcal{B}, \mu, T)$, sendo que $(X, \mathcal{B}, \mu)$ é um espaço de medida, isto é, X é um espaço-fase, $\mathcal{B}$ é uma σ-álgebra sobre X, e μ é uma medida sobre $(X, \mathcal{B})$, e onde $T : X \to X$ é uma transformação que preserva μ, ou seja, $\mu(T^{-1}(A)) = \mu(A)$, para todo $A \subseteq X$ mensurável (nesse caso também dizemos que T é $(\mu, \mathcal{B})$-mensurável ou, quando não for caso de confusão, μ-mensurável).

Quando a medida de X é finita, dizemos que X é um *espaço de probabilidade*.

A estrutura de um sistema dinâmico munido com uma medida de probabilidade proporciona muita informação de caráter probabilístico sobre a transformação, principalmente no que se refere a conceitos de recorrência. Não é surpreendente, portanto, que vários dos teoremas tratados aqui são justamente deste tipo.

Um dos teoremas que usaremos é o seguinte:

Teorema 2.3.1 (Recorrência Múltipla Ergódica de Furstenberg)**.** *Seja (X, μ) um espaço de probabilidade, $T : X \to X$ uma função μ-invariante, $k \geq 3$ um inteiro e $A \subseteq X$ tal que $\mu(A) > 0$.*

Então existe $N > 0$ tal que $\mu(\cap_{j=0}^{k-1} T^{-jN}(A)) > 0$.

A prova deste teorema também pode ser encontrada em [Fur81].

2.4 Teoria Elementar e Analítica dos Números

Durante nossos estudos sobre o conjunto dos primos e suas propriedades, necessitaremos de alguns conceitos e resultados básicos da Teoria dos Números.

As definições que usaremos são

Definição (Funções Piso e Teto)**.** Dado $x \in \mathbb{R}$, a função *piso de x* é determinada por

$$\lfloor x \rfloor := \max \{m \in \mathbb{Z} : m \leq x\},$$

e a função *teto de x* é dada por

$$\lceil x \rceil := \min \{m \in \mathbb{Z} : m \geq x\}.$$

Definição (Função Totiente de Euler)**.** A *função totiente de Euler* $\phi : \mathbb{N} \to \mathbb{N}$ é definida por

$$\phi(n) = |\{1 \leq k \leq n : mdc(k, n) = 1\}|.$$

Definição (Função de Möbius)**.** A *função de Möbius* $\mu : \mathbb{N} \to \{-1, 0, 1\}$ é dada por

$$\mu(n) := \begin{cases} 1, & \text{se } n \text{ é livre de quadrados e possui um número par de fatores primos;} \\ -1, & \text{se } n \text{ é livre de quadrados e possui um número ímpar de fatores primos;} \\ 0, & \text{se } n \text{ não for livre de quadrados.} \end{cases}$$

Precisaremos também dos seguintes resultados:

Teorema 2.4.1 (Fórmula da Inversão de Möbius)**.** *Seja $f : \mathbb{N} \to \mathbb{R}$ e consideremos $g : \mathbb{N} \to \mathbb{R}$ dada por $g(n) = \sum_{d|n} f(d)$.*

Então a seguinte identidade

$$f(n) = \sum_{d|n} \mu(d) g\left(\frac{n}{d}\right)$$

é válida.

A validade desta fórmula é provada em [MMST13].

Também faremos uso do seguinte

Teorema 2.4.2 (Ordem máxima de $d(n)$)**.** *Seja $n \geq 2$ um natural e consideremos $d(n)$ a função que a cada natural n associa o número de divisores (positivos) de n.*

Então, para todo $\varepsilon > 0$, existe $n_0 \in \mathbb{N}$ tal que, se $n > n_0$, temos que $d(n) < 2^{(1+\varepsilon)\log n/\log\log n}$.

Demonstração. Seja $n = \prod_{j=1}^{k} p_j^{\alpha_j}$, com $p_i < p_j$ se $i < j$ a fatoração em números primos de n.

Temos então que

$$d(n) = \prod_{j=1}^{k}(1+\alpha_j) = \prod_{p_j \leq (\log n)^{1-\delta}} (1+\alpha_j) \cdot \prod_{p_j > (\log n)^{1-\delta}} (1+\alpha_j),$$

onde $\delta = \varepsilon/2(1+\varepsilon)$.

Analisando o primeiro produtório, dado que $n \geq p_j^{\alpha_j} \geq 2^{\alpha_j}$, para todo j, e conseqüentemente $\log n \geq \alpha_j \log 2$, temos que para todo j segue que $1 + \alpha_j \leq 1 + \dfrac{\log n}{\log 2}$.

Assim, vemos que

$$\prod_{p_j \leq (\log n)^{1-\delta}} (1+\alpha_j) \leq \prod_{j=1}^{(\log n)^{1-\delta}} (1+\alpha_j) = (1+\alpha_j)^{(\log n)^{1-\delta}} \leq \left(1 + \frac{\log n}{\log 2}\right)^{(\log n)^{1-\delta}},$$

donde concluímos que $\prod_{p_j \leq (\log n)^{1-\delta}}(1+\alpha_j) \leq 2^{O(\log\log n \cdot (\log n)^{1-\delta})}$.

Por outro lado, para todo j, obtemos $1 + \alpha_j \leq 2^{\alpha_j}$, e daí,

$$\prod_{p_j > (\log n)^{1-\delta}} (1+\alpha_j) \leq 2^{\sum_{p_j > (\log n)^{1-\delta}} \alpha_j} \leq 2^{\log n/\log((\log n)^{1-\delta})} = 2^{\log n/(1-\delta)\log\log n}$$

(de fato, temos que $n \geq \prod_{p_j > (\log n)^{1-\delta}} p_j^{\alpha_j} \geq ((\log n)^{1-\delta})^{\sum_{p_j > (\log n)^{1-\delta}} \alpha_j}$ implica diretamente que $\log n \geq (\log((\log n)^{1-\delta})) \cdot \sum_{p_j > (\log n)^{1-\delta}} \alpha_j$, e daí $\dfrac{\log n}{\log((\log n)^{1-\delta})} \geq \sum_{p_j > (\log n)^{1-\delta}} \alpha_j$).

Concluímos então que

$$d(n) = \prod_{j=1}^{k}(1+\alpha_j) \leq 2^{\log n/(1-\delta)\log\log n + O(\log n/\log\log n)} < 2^{(1+\varepsilon)\log n/\log\log n},$$

se n é suficientemente grande, já que $\frac{1}{1-\delta} < \frac{1}{1-2\delta} = 1 + \varepsilon$. □

Por fim, um último resultado também será do nosso interesse:

Teorema 2.4.3 (Teorema do Número Primo para Progressões Aritméticas)**.** *A seguinte identidade é verdadeira:*

$$\sum_{\substack{p \leq x \\ p \equiv a \pmod{b}}} \log p = \frac{x}{\phi(b)}(1 + o(1)).$$

A prova deste teorema pode ser encontrada em [Sel50].

Capítulo 3

Alguns Resultados Preliminares

Neste capítulo, teremos contato e provaremos os resultados que levam ao teorema principal deste texto. Destacamos que a maior parte destes teoremas possuem várias demonstrações, quer usando métodos combinatórios, métodos da análise de Fourier, ou ainda ferramentas de sistemas dinâmicos e teoria ergódica. Usaremos, seguindo o caminho de Furstenberg, as duas últimas abordagens mencionadas.

3.1 O Teorema de Van der Waerden

O Teorema de Van der Waerden foi obtido em 1927, em [vdW27], e foi de uma enorme importância para o desenvolvimento de áreas como combinatória extremal e combinatória aditiva. Em [Fur81], Furstenberg deu uma prova usando noções de teoria ergódica, que é a que apresentaremos aqui.

Antes de enunciar o teorema, precisamos definir um conceito simples. Uma *m-coloração de um conjunto X* é uma função $c : X \to \{1, 2, 3, \ldots, m\}$.

Caso $A \subseteq c^{-1}(j)$, para algum j, dizemos que A é um *conjunto monocromático (de cor j pela coloração c).*

Por fim, notemos que realizar uma m-coloração de X é equivalente a realizar uma partição de X em m subconjuntos distintos.

Teorema 3.1.1 (Van der Waerden)**.** *Seja $c : \mathbb{N} \to \{1, 2, 3, \ldots, m\}$ uma m-coloração dos naturais.*

Então existe $j \in \{1, 2, 3, \ldots, m\}$ tal que o conjunto $c^{-1}(j)$ possui progressões aritméticas de tamanho arbitrariamente grande.

Em outras palavras, se $\mathbb{N}$ é colorido com m cores, existe um conjunto monocromático A com k-PAs, para qualquer $k \in \mathbb{N}$.

Demonstração. A prova será feita através dos seguintes passos:

1. Associaremos, de maneira apropriada, uma m-coloração dos naturais a uma palavra infinita e a noção de PA para uma recorrência da aplicação *shift*, podendo então usar as ferramentos da dinâmica simbólica;

2. Com um elemento apropriado, construiremos um conjunto compacto no qual a órbita deste elemento é densa e a função *shift* é contínua;

3. Podemos então usar o Teorema da Recorrência Múltipla Topológica para escolhermos $k \in \mathbb{N}$ e $\varepsilon > 0$ de modo a obtermos uma "relação monocromática" entre as letras da nossa palavra;

4. Por fim, aplicaremos o resultado ao problema original.

Seja $c : \mathbb{N} \to \{1, 2, 3, \ldots, m\}$ uma m-coloração dos naturais.

Primeiramente lembremos que $\Omega(\{1, 2, 3, \ldots, m\})$ com a métrica d definida anteriormente é compacto.

A partir daí, uma escolha natural para um elemento de $\Omega(\{1, 2, 3, \ldots, m\})$ que mantém as informações da coloração é $z = (z_n)_{n\in\mathbb{N}} = (c(n))_{n\in\mathbb{N}}$. Observe que, se $d(\sigma^r(z), \sigma^s(z)) < 1$, então $z_{r+1} = z_{s+1}$.

Considere $Z = \mathcal{O}_\sigma(z)$, e tome $X = \overline{Z} = \overline{\mathcal{O}_\sigma(z)}$. Temos então que X é compacto, σ é contínua e Z é denso em X.

Nestas condições, podemos aplicar o Teorema da Recorrência Múltipla Topológica de Furstenberg e Weiss (Teorema 2.2.1), e sendo $k \in \mathbb{N}$ o tamanho da PA desejado e $\varepsilon = \frac{1}{2}$, temos que existem $\tilde{z} \in Z$, isto é, $\tilde{z} = \sigma^r(z) = (z_{r+1}, z_{r+2}, \ldots)$ e $n \in \mathbb{N}$ tal que $d(\sigma^{jn}(\tilde{z}), (\tilde{z})) < \frac{1}{2}$, para todo $j = 1, 2, \ldots, k$.

Pela observação acima, isto significa que $z_{r+1} = z_{n+(r+1)} = z_{2n+(r+1)} = \ldots = z_{kn+(r+1)}$, ou seja, pela definição de z, temos que $c(r + 1) = c(n + (r + 1)) = c(2n + (r + 1)) = \ldots = c(kn + (r + 1))$, e portanto $\{(r + 1) + [0, k + 1)n\}$ contém, em particular, uma k-PA monocromática.

Assim, para cada k, existe uma k-PA monocromática. Como o número de cores é finito, pelo Princípio da Casa de Pombos, uma das cores contém k-PAs, para infinitos valores de k e, portanto, tal conjunto contém PAs de tamanho arbitrariamente grande. □

Vale a pena fazermos uma consideração sobre este resultado: ele não diz que existe um conjunto com uma PA *infinita* monocromática.

De fato, é possível particionarmos $\mathbb{N}$ em dois subconjuntos A, B de maneira que nenhum deles possua uma PA de tamanho infinito. Um exemplo simples em que isso ocorre é tomando $A = \{0\} \cup \left\{2^l + 1 \leq n \leq 2^{l+1} : l \text{ é par}\right\}$ e $B = \mathbb{N} \setminus A = \left\{2^l + 1 \leq n \leq 2^{l+1} : l \text{ é ímpar}\right\}$. Como cada um deles possui "buracos" de tamanho 2^l, para qualquer l, é impossível que contenham uma PA infinita, já que os saltos que um conjunto pode ter para que possua uma PA são limitados pela sua razão. Apesar disto, ambos possuem PAs de tamanho arbitrariamente grande.

3.2 O Teorema de Szemerédi

Dificilmente alguém poderia exagerar a importância do Teorema de Szemerédi. Conjecturado por Erdős e Turán em 1936 [ET36], foi respondido afirmativamente por Szemerédi em 1975 [Sze75], numa prova foi considerada por Erdős "uma obra-prima do pensamento combinatório" ([Erd97]).

Hoje são conhecidas várias provas, usando ferramentas como teoria ergódica [Fur81] e análise de Fourier [Gow01], entre outras; de fato, o próprio Terence Tao revelou que percebia as diversas provas do Teorema de Szemerédi como uma espécie de "Pedra de Rosetta" que conecta distintas áreas da Matemática ([Tao05b]).

Ainda seguindo a nossa proposta, veremos a demonstração ergódica de Furstenberg.

Primeiramente, precisamos apresentar a seguinte

Definição (Densidade Superior)**.** Seja $A \subseteq \mathbb{N}$.

Definimos a *densidade superior de* A, denotada por $d(A)$, como sendo o valor

$$d(A) = \limsup_{n\to\infty} \frac{|A \cap \{1, 2, 3, \ldots, n\}|}{n}.$$

Podemos notar a estreita ligação do conceito de densidade superior com o de medida: propriedades como monotonicidade em relação à inclusão e positividade são claras.

Porém, podemos ver que a propriedade de σ-aditividade não é satisfeita, já que, caso fosse válida, teríamos $1 = d(\mathbb{N}) = d(\cup_{n\in\mathbb{N}} \{n\}) = \sum_{n\in\mathbb{N}} d(\{n\}) = \sum_{n\in\mathbb{N}} 0 = 0$, um absurdo.

Mesmo assim, ainda podemos perceber como tais conceitos estão próximos, e é tal proximidade que possibilita termos noções probabilísticas sobre subconjuntos dos números naturais e suas relações com progressões aritméticas (que, conforme veremos, podem ser encaradas sob o ponto de vista do conceito de recorrência, como enfatizamos na seção 2.3).

Para provar o Teorema de Szemerédi, usaremos, juntamente com o Teorema da Recorrência Múltipla Ergódica de Furstenberg (Teorema 2.3.1), o seguinte lema:

Lema 3.2.1. *Sejam A um alfabeto finito e $\xi \in \Omega(A)$. Para $a \in A$, consideremos o conjunto* $\Lambda(a) = \left\{\omega = (\omega_0, \omega_1, \omega_2, \ldots) \in \overline{\mathcal{O}_\sigma(\xi)} : \omega_0 = a\right\}$.

Então o símbolo a ocorre em ξ com densidade superior positiva se, e somente se, existe uma medida μ em $\overline{\mathcal{O}_\sigma(\xi)}$ que é σ-invariante e com $\mu(\Lambda(a)) > 0$.

Como esperado, este lema é desenvolvido no livro de Furstenberg [Fur81].

Teorema 3.2.1 (Szemerédi)**.** *Seja $A \subseteq \mathbb{N}$ com densidade superior positiva.*

Então A possui infinitas progressões aritméticas de tamanho arbitrariamente grande.

Demonstração. Sejam $A \subseteq \mathbb{N}$ o conjunto com densidade positiva e ξ a palavra $(\chi_A(j))_{j\in\mathbb{N}}$, (isto é, a palavra formada é aquela constituída por 0s nos naturais que não pertencem a A e por 1s nos naturais que pertencem a A), $X = \overline{\mathcal{O}_\sigma(\xi)}$ e $\Lambda(1) = \{\omega \in X : \omega_0 = 1\}$.

Pelo Lema 3.2.1, como o símbolo 1 aparece em ξ com densidade superior positiva, temos que existe uma medida σ-invariante μ em X tal que $\mu(\Lambda(1)) > 0$.

Além disso, pelo Teorema da Recorrência Múltipla Ergódica (Teorema 2.3.1), dado $k \in \mathbb{N}$ temos que existe $n \in \mathbb{N}$ tal que $\mu(\cap_{j=0}^{k-1}\sigma^{jn}(\Lambda(1))) > 0$, isto é, existe $z \in X$ com $z(0) = \sigma^n(z(0)) = \sigma^{2n}(z(0)) = \ldots = \sigma^{(k-1)n}(z(0)) = 1$, mas como $\sigma^j(z(0)) = z(j)$, temos que $z(0) = z(n) = z(2n) = \ldots = z((k-1)n) = 1$.

Obtemos também, dado que $z = \sigma^l(c)$, que $c(l) = c(l+n) = c(l+2n) = \ldots = c(l+(k-1)n) = 1$, e portanto A possui uma k-PA monocromática.

Além disso, como o conjunto $A \setminus \{l + [0, k)n\}$ possui a mesma densidade de A, reaplicando a demonstração acima acharemos infinitas k-PAs. □

Ressaltamos que o Teorema de Szemerédi implica o Teorema de Van der Waerden.

De fato, dada $\{A_1, A_2, \ldots, A_m\}$ uma partição finita de $\mathbb{N}$, temos que

$$\begin{aligned}
1 &= \lim_{n\to\infty} \frac{|\{1,2,3,\ldots,n\}|}{n} = \lim_{n\to\infty} \frac{|\mathbb{N} \cap \{1,2,3,\ldots,n\}|}{n} \\
&= \lim_{n\to\infty} \frac{|(\cup_{j=1}^m A_j) \cap \{1,2,3,\ldots,n\}|}{n} = \lim_{n\to\infty} \frac{|\cup_{j=1}^m (A_j \cap \{1,2,3,\ldots,n\})|}{n} \\
&= \lim_{n\to\infty} \sum_{j=1}^m \frac{|A_j \cap \{1,2,3,\ldots,n\}|}{n} = \sum_{j=1}^m \lim_{n\to\infty} \frac{|A_j \cap \{1,2,3,\ldots,n\}|}{n} \\
&= \sum_{j=1}^m d(A_j),
\end{aligned}$$

e portanto, como a soma é positiva, ao menos um dos fatores $d(A_j)$ é positivo, e então, por Szemerédi, A possui k-PAs, para qualquer $k \in \mathbb{N}$.

3.3 Versões Finitárias

Nesta seção, enunciaremos resultados equivalentes aos mostrados anteriormente, sob um ponto de vista *finitário*, no sentido de que agora os conjuntos em questão são da forma $\{1, 2, 3, \ldots, N\}$ ou da forma $\mathbb{Z}_N$, ao invés do conjunto infinito $\mathbb{N}$.

A razão desta abordagem se dá pelo fato de que tais enunciados (especificamente, o resultado equivalente ao Teorema de Szemerédi enunciado na seção anterior) tornam a generalização que buscamos mais transparente e fácil de trabalhar.

Teorema 3.3.1 (Teorema de Van der Waerden Finitário)**.** *Para todos os inteiros positivos r e k existe um inteiro $N_W = N_W(r,k)$ tal que, para todo $N \geq N_W$, se $\{1,2,3,\ldots,N\}$ é colorido com r cores, existe uma k-PA monocromática.*

Teorema 3.3.2 (Teorema de Szemerédi Finitário)**.** *Para todo $k \geq 1$ e $0 < \delta \leq 1$, existe um inteiro $N_S = N_S(k,\delta) \geq 1$ tal que, para qualquer $N \geq N_S$, temos que todo conjunto $A \subseteq \{1,2,3,\ldots,N\}$ com cardinalidade $|A| \geq \delta N$ contém pelo menos uma k-PA.*

Notemos que a expressão relacionando a cardinalidade de A com δ e N pode ser encarada como uma propriedade sobre a densidade do conjunto A — de fato, é bem perceptível a semelhança de $|A|/N \geq \delta$ com a propriedade de densidade superior positiva —, no mesmo sentido que o Teorema de Szemerédi se baseia.

Assim, podemos reescrevê-lo da seguinte maneira, que será útil para nossa busca por generalizações:

Teorema 3.3.3 (Szemerédi Finitário Alternativo)**.** *Para cada $k > 0$ e $\delta > 0$, existem $N_S = N_S(k,\delta)$ e $c_{k,\delta} > 0$ tais que, para todo $N \geq N_S$ e toda $f : \mathbb{Z}_N \to \{0,1\}$ com $\mathbb{E}(f) \geq \delta$, vale que*

$$\Upsilon_k(f) \geq c_{k,\delta}.$$

Além disso, o valor de $c_{k,\delta}$ é independente de f.

Uma generalização deste teorema com uma sutil, porém profunda diferença, ao permitir que a função analisada possuísse mais liberdade em seu contradomínio:

Teorema 3.3.4 (Szemerédi-Gowers)**.** *Para cada $k > 0$ e $\delta > 0$, existem $N_S = N_S(k,\delta)$ e $c_{k,\delta} > 0$ tais que, para todo $N \geq N_S$ e toda $f : \mathbb{Z}_N \to [0,1]$ com $\mathbb{E}(f) \geq \delta$, vale que*

$$\Upsilon_k(f) \geq c_{k,\delta}.$$

Além disso, o valor de $c_{k,\delta}$ é independente de f.

Como dito, a única diferença é a mudança do contradomínio da f, de $\{0,1\}$ para $[0,1]$, cuja forma equivalente $0 \leq f \leq 1$ tornará a transição para a forma geral que queremos mais natural.

Observação. Uma última, porém importantíssima, observação: teremos que ter cuidado com o fato de que existem PAs em $\mathbb{Z}_N$ que não são PAs em $\mathbb{N}$; como exemplo, basta notarmos que $\{1,4,2\}$ é uma 3-PA em $\mathbb{Z}_5$, mas não é uma PA em $\mathbb{N}$.

Este problema é facilmente solucionado limitanto o intervalo nos quais os elementos da PA podem ocupar dentro de $\mathbb{Z}_N$ (por exemplo, restringindo o suporte das progressões no "terço central" de $\mathbb{Z}_N$). Resolveremos esta questão em detalhes na seção 4.4.

Capítulo 4

O Teorema de Green-Tao

Neste capítulo, definiremos os objetos e provar os resultados usados para enunciar e provar o Teorema de Green-Tao.

Primeiramente, para auxiliar nosso entendimento, daremos uma visão geral de cada tópico que será abordado, provendo também uma pequena explicação sobre sua importância para o esquema completo. Em seguida, mostraremos a estratégia em que cada um desses tópicos contribui, a fim de alcançarmos nossos objetivos. Então, finalmente, começaremos a "sujar as mãos" e estudarmos os conceitos com rigor e profundidade.

Vale ressaltar que, por fins didáticos (ao menos do ponto de vista do autor desta monografia), esta abordagem difere da usada originalmente por Green e Tao, baseando-se fortemente nos estudos de Gowers [Gow10].

4.1 Visão Geral

1. Analisar as várias versões do Teorema de Szemerédi e compreender o tipo de generalização que procuramos para lidar com o objetivo de aplicá-lo aos primos;
2. Iniciaremos os estudos do Produto Interno e das Normas de Gowers, um modo engenhoso de medir propriedades relacionadas com a existência e a quantidade de PAs num conjunto;
3. Com as ferramentas estudadas anteriormente, podemos entender o Teorema de von Neumann Generalizado, que nos dá certos limitantes sobre como a contagem de k-PAs é governada pela norma de Gowers das funções envolvidas;
4. Iniciaremos a análise dos efeitos da dualidade nas normas de Gowers, e como estas podem nos ajudar a conseguir limitantes interessantes para nossos fins;
5. Em posse de tais análises, provaremos o Teorema da Decomposição, que nos permitirá decompor nossa função em uma parte em que poderemos aplicar o Teorema de Szemerédi original e em uma parte "negligenciável";
6. Obtendo tais resultados, provaremos o Teorema de Szemerédi Relativo, uma generalização do resultado original que permitirá sua aplicação a um número muito maior de casos;
7. Em seguida, estudaremos a distribuição dos números primos e veremos que, de fato, eles são passíveis de aplicação do Teorema de Szemerédi Relativo;
8. Por fim, em posse de cada uma dessas ferramentas, provaremos o Teorema de Green-Tao.

4.2 Estratégia

A estratégia para provarmos o Teorema de Green-Tao pode ser entendida como um processo com dois passos principais, seguindo o seguinte esquema:

1. Provamos o que chamaremos de *Teorema de Szemerédi Relativo*, usando o maquinário das chamadas *normas de Gowers*, que induzem a uma noção de distância entre subconjuntos de $\mathbb{N}$ de modo que:
   (a) Tais normas preservam a quantidade de k-PAs, isto é, subconjuntos próximos nestas normas contêm, aproximadamente, o mesmo número de k-PAs;
   (b) Há um certo *princípio de transferência*, no sentido de que um subconjunto denso em relação a um subconjunto pseudoaleatório fica próximo, do ponto de vista de tais normas, de um subconjunto denso nos naturais. Isto será feito usando o chamado *Teorema da Decomposição*;
   (c) Em posse de ambos os resultados e do Teorema de Szemerédi original, mostramos que este subconjunto maior possui as procuradas k-PAs, e então, usando o *Teorema de von Neumann Generalizado*, provamos que o subconjunto original também possui k-PAs.
2. Verificamos que, efetivamente, os primos possuem densidade positiva num subconjunto pseudoaleatório.

A principal idéia é a de que não precisamos que o conjunto seja necessariamente $\mathbb{N}$, mas que seja "regular" o suficiente para que se comporte, do ponto de vista aritmético, de maneira similar ao conjunto dos naturais, e deste modo possamos usar o Teorema de Szemerédi tendo tal conjunto como base, e usando a noção de densidade relativa.

A definição de *k-pseudoaleatoriedade* fornecerá a formalização do conceito de regularidade que buscamos, e a proximidade de sua semelhança em relação aos aspectos aritméticos dos números naturais é desenvolvida pela estudo das *normas de Gowers*.

Em seguida, a aplicação aos números primos é dada por uma medida criada sobre uma variação da *função de von Mangoldt*, que lida de maneira especial com os primos, e pela verificação que tais alterações satisfazem a versão mais geral do Teorema de Szemerédi e também o possibilita a dar informações sobre distribuições no conjunto dos primos.

4.3 O Teorema de Szemerédi Relativo

4.3.1 A Evolução do Teorema de Szemerédi

Durante nossa incursão sobre as propriedades que relacionam a densidade e a cardinalidade de subconjuntos de $\mathbb{N}$ com progressões aritméticas, lidaremos com vários enunciados que retomam o constante tema do Teorema de Szemerédi, e uma visão *a priori* desta evolução pode ser vantajosa em nos guiar e orientar para nossos objetivos.

Primeiramente, temos o enunciado inicial simples, porém um tanto quanto difícil de utilizar em nosso caso (dada nossa discussão sobre a densidade superior do conjuntos dos números primos ser nula), relacionando a densidade superior de um subconjunto de $\mathbb{N}$ com o fato deste possuir progressões aritméticas de tamanho arbitrário.

Em seguida, olhamos para a versão que traduz (de maneira equivalente) a noção de densidade como a esperança de uma "função característica" do conjunto, como mostrado no Teorema 3.3.2.

A generalização provida pelo Teorema de Szemerédi-Gowers (Teorema 3.3.4) se dá com a pequena, mas extremamente profunda, mudança de não necessitar que a "função característica" do conjunto possua como contradomínio o conjunto $\{0, 1\}$, possibilitando uma espécie de ponderamento sobre os elementos do conjunto, ao permitir que a função varie por todo o intervalo $[0, 1]$, isto é, temos a hipótese mais geral de que $0 \leq f \leq 1$.

Por fim, após definirmos o conceito de *pseudoaleatoriedade*, conseguiremos formular um enunciado para o teorema que permite uma generalização ainda maior, ao substituir a hipótese de $0 \leq f \leq 1$ por $0 \leq f \leq \nu$, onde ν é uma medida k-pseudoaleatória *qualquer*. Conforme veremos, este resultado de fato generaliza o anterior, dado que a medida associada à funcão constante $\nu \equiv 1$ é uma medida k-pseudoaleatória.

Conforme veremos, o conceito de pseudoaleatoriedade está fortemente relacionado com o fato de certas propriedades aritméticas serem "bem distribuidas", e por isso não é de todo surpreendente que a função constante 1 é, de fato, facilmente associada a uma medida pseudoaleatória, bem como o fato de que todas as medidas pseudoaleatórias são, num sentido que analisaremos, próximas à medida associada à função constante 1.

Com este comentário em mente, percebemos que esta última versão é realmente uma generalização do enunciado anterior, de modo que poderemos analisar conjuntos muito mais diversos que $\mathbb{N}$, e será a forma final que utilizaremos para atingir nosso objetivo de entender melhor o comportamento aritmético dos primos[1].

Para introduzirmos o conceito, primeiramente precisamos considerar a seguinte

Definição (Medida[2])**.** Seja $\nu = (\nu_1, \nu_2, \nu_3, \ldots) = (\nu_n)_{n\geq 1}$, de modo que $\nu_n : \mathbb{Z}_n \to \mathbb{R}^+$. Dizemos que ν é uma *medida* se

$$\mathbb{E}(\nu_n) = 1 + o(1).$$

Daqui em diante, ao nos referirmos a uma medida, permitiremo-nos um abuso de notação e escreveremos $\mathbb{E}(\nu) = 1 + o(1)$.

E, após definirmos o conceito de pseudoaleatoriedade, poderemos entender com clareza a forma final do teorema que buscamos:

Teorema 4.3.1 (Szemerédi Relativo)**.** *Seja ν uma medida, e sejam dados $k > 0$ e $\delta > 0$.*

Então existem $N_{GT} = N_{GT}(k, \delta)$ e $c'_{k,\delta} > 0$ tais que, para todo $N \geq N_{GT}$, para toda ν medida k-pseudoaleatória, e para toda $f_N : \mathbb{Z}_N \to \mathbb{R}^+$ com $0 \leq f_N \leq \nu_N$ (dominância ponto a ponto) e $\mathbb{E}(f_N) \geq \delta$, existe uma constante $c'_{k,\delta}$ tal que

$$\Upsilon_k(f_N) \geq c'_{k,\delta}.$$

Além disso, os valores de N_{GT} e de $c'_{k,\delta}$ são independentes de f_N.[3]

Provaremos esta forma final, com um detalhamento ainda maior (a saber, a relação entre a constantes $c_{k,\delta}$ do Teorema 3.3.4 e a constante $c'_{k,\delta}$ do Teorema 4.3.5), sob o nome de *Teorema de Szemerédi Relativo* (Teorema 4.3.5), na subseção 4.3.7 (O Enunciado e a Prova do Teorema de Szeremédi Relativo).

4.3.2 A Noção de k-Pseudoaleatoriedade

O conceito de k-pseudoaleatoriedade, que será fundamental daqui para a frente, se baseia na idéia de tentar transportar a "uniformidade" das propriedades aritméticas presentes em $\mathbb{N}$ para contextos mais amplos. Se tomarmos, por exemplo, como "conjunto-base" um subconjunto A de $\mathbb{N}$ de modo que $\mathbb{N} \setminus A$ seja finito, não é difícil ver que o Teorema de Szemerédi pode ser usado para nos garantir a existência de progressões aritméticas de tamanho arbitrariamente grande em A.

Conforme veremos, isto é facilmente transportado para o contexto de um conjunto particular de famílias de funções pseudoaleatórias. O surpreendente, no entanto, é que situações *muito mais gerais* possam ser estudadas ao se aplicar tais conceitos. Usando ainda o mesmo exemplo, não é

[1] Na verdade, ficará cada vez mais claro que iremos deixar de focar nossa atenção em conjuntos para olharmos mais atentamente para *funções*, o que só nos trará vantagens, já que ganharemos todo um arsenal de ferramentas como normas, produtos internos e operações que não possuímos ao lidar diretamente com conjuntos. Por outro lado, não é difícil perceber as relações que podem ser associadas entre uma função e os conjuntos nos quais ela se baseia.

[2] Como o próprio artigo de Green e Tao ressalta, usa-se somente a nomenclatura de *medida*, apesar de ser destacado que tal escolha foi usada considerando sua brevidade, e que o termo *densidade de probabilidade normalizada* seria mais acurado. O uso do termo "medida" neste contexto não deve ser confudido com o conceito clássico de medida, que ocupou certo espaço na seção 2.3, ainda que não tão expressivo na parte principal deste texto.

[3] Claramente, estamos mais preocupados com o comportamento assintótico das funções consideradas. Por isso, de agora em diante, nos permitiremos mais alguns abusos de notação e será comum nos referirmos diretamente a f e ν, sem nos preocuparmos com os índices. Seguindo esta filosofia, poderíamos escrever o enunciado com $f : \mathbb{Z}_N \to \mathbb{R}^+$, $0 \leq f \leq \nu$, $\mathbb{E}(f)$, entre outras substituições similares.

necessário o maquinário da pseudoaleatóriedade para lidarmos com o conjunto A acima, dada sua densidade superior positiva, mas no caso dos primos, que como sabemos possuem densidade superior nula, é este maquinário que permitirá a generalização que buscamos.

Começaremos as definições introduzindo o conceito de *peso* de uma fração (irredutível) a/b, que é igual ao $\max\{|a|, |b|\}$, e dado $q = (q_1, q_2, \ldots, q_p) \in \mathbb{Q}^p$, diremos que o *peso da p-úpla q* é o máximo dos pesos entre as coordenadas $q_1, q_2, \ldots, q_p$.

Agora definiremos algumas propriedades importantes:

Definição (Condição Das Formas Lineares)**.** Se m_0, t_0 e L_0 são naturais não-nulos, dizemos que uma medida ν satisfaz a (m_0, t_0, L_0)*-condição das formas lineares* se, dados $m \leq m_0, t \leq t_0$ e uma família $\left\{L_i : L_i \in \mathbb{Q}^t\right\}_{1\leq i\leq m}$ de t-úplas racionais não nulas e que não são múltiplas racionais entre si, tais que cada elemento possui seu *peso* limitado por L_0, então as formas lineares

$$\psi_i(x) := \sum_{j=1}^{t} L_{ij}x_j + b_i,$$

satisfazem a identidade

$$\mathbb{E}(\nu(\psi_1(x)) \cdot \nu(\psi_2(x)) \cdot \ldots \cdot \nu(\psi_m(x)) : x \in \mathbb{Z}_N^t) = 1 + o_k(1)$$

Se uma medida satisfaz a $(k2^k, 3k-4, k)$-condição das formas lineares, dizemos que tal medida é *linearmente k-pseudoaleatória*[4].

Primeiramente, notemos que se tomarmos $m = 1$, a condição das formas lineares se reduz a dizer que $\mathbb{E}(\nu) = 1 + o(1)$, que é a definição de uma medida. Analisando mais profundamente, é possível ver que podemos encarar a condição das formas lineares como uma espécie de independência entre $\nu(\psi_1), \ldots, \nu(\psi_m)$, o que está de acordo com nossos objetivos de encontrar um paralelo com uma distribuição uniforme sobre o suporte das medidas.

Além disso, vale a pena destacar que, se ν satisfaz a (m_0, t_0, L_0)-condição das formas lineares, então para todos $m \leq m_0, t \leq t_0, L \leq L_0$, a (m, t, L)-condição das formas lineares também é satisfeita por ν.

Definição (Condição De Correlação)**.** Dizemos que uma medida ν satisfaz a m_0*-condição de correlação* se, para todo $1 < m \leq m_0$, existe uma função-peso $\tau = \tau_m : \mathbb{Z}_N \to \mathbb{R}^+$ que satisfaz a chamada *condição de momento*

$$\mathbb{E}(\tau(x)^q) = O_{m,q}(1),$$

para todo $1 \leq q < \infty$ e, além disso,

$$\mathbb{E}(\nu(x+h_1) \cdot \nu(x+h_2) \cdot \ldots \cdot \nu(x+h_m) : x \in \mathbb{Z}_N) \leq \sum_{1\leq i<j\leq m} \tau(h_i - h_j),$$

para todos $h_1, h_2, \ldots, h_m \in \mathbb{Z}_N$ (não necessariamente distintos).

Se uma medida satisfaz a 2^{k-1}-condição de correlação, dizemos que tal medida é *simplesmente k-pseudoaleatória.*

A condição de correlação será utilizada principalmente para estimar certos produtos envolvendo funções que estarão intimamente ligadas à contagem de k-PAs. Ela provê certos limitantes para o quanto as "variáveis aleatórias" $\nu(\psi_j)$ são correlacionas entre si, quando as ψ_j são funções lineares com coeficientes lineares iguais.

Definição (k-Pseudoaleatoriedade)**.** Uma medida ν é dita *k-pseudoaleatória* se satisfaz, simultaneamente, a $(k2^{k-1}, 3k-4, k)$-condição das formas lineares e a 2^{k-1}-condição de correlação, ou, em outras palavras, se ν é tanto linearmente k-pseudoaleatória quanto simplesmente k-pseudoaleatória.

[4]Como recurso mnemônico, basta lembrar que uma medida ser *linearmente* k-pseudoaleatória possui relação com a condição das formas *lineares.*

O conceito de k-pseudoaleatoriedade pode parecer, à primeira vista, um pouco estranho (principalmente a escolha dos coeficientes). Sua definição vem de resultados de Goldston e Yıldırım[5], que procuram majorantes para funções de von Mangoldt modificadas, que possuem ligação com os números primos (NÃO ENTRE EM PÂNICO!: estudaremos mais detalhadamente estes resultados na seção 4.4).

De maneira intuitiva, tais condições nos dizem que o conjunto de inteiros no suporte de ν possui propriedades aritméticas aleatórias, isto é, se comportam aritmeticamente de maneira parecida com o que se esperaria de uma distribuição uniforme sobre os inteiros (se tal objeto existisse). Por outro lado, são suficientemente fracas para que possam ser aplicadas nos primos e permitir o uso da versão mais geral do Teorema de Szemerédi (o já citado Teorema de Szemerédi Relativo (Teorema 4.3.1)) no suporte de medidas focadas em primos.

A seguir, veremos algumas propriedades básicas que se mostrarão importantes para nossas considerações.

Para entendermos a primeira propriedade, usaremos o conceito de *conjunto estrelado ao redor de a* em um dado um espaço vetorial X, que introduziremos a seguir: um subconjunto não-vazio $A \subseteq X$ é dito *estrelado ao redor de a* se, para todo $x \in A$, temos que o ponto médio entre x e a também pertence a A, isto é, $(a+x)/2 \in A$.

Lema 4.3.1 (Conjunto estrelado ao redor de 1)**.** *Seja ν uma medida k-pseudoaleatória. Então $\nu_{1/2} := (\nu+1)/2$ também é uma medida k-pseudoaleatória.*

Demonstração. Claramente temos que $\mathbb{E}(\nu_{1/2}) = \frac{1}{2}(\mathbb{E}(\nu)+1) = 1+o(1)$, e portanto satisfaz as condições de uma medida.

Para provar que $\nu_{1/2}$ satisfaz a $(k2^{k-1}, 3k-4, k)$-condição das formas lineares, consideremos $m \leq k2^{k-1}, t \leq 3k-4$ e $\psi_1, \ldots, \psi_m$ conforme as exigências requeridas na definição das formas lineares.

Então, temos que

$$\begin{aligned}
&\mathbb{E}(\nu_{1/2}(\psi_1(x)) \cdots \nu_{1/2}(\psi_m(x)) : x \in \mathbb{Z}_N^t) \\
&= \mathbb{E}\left(\frac{\nu(\psi_1(x)+1)}{2} \cdots \frac{\nu(\psi_m(x)+1)}{2} : x \in \mathbb{Z}_N^t\right) \\
&= \frac{1}{2^m}\mathbb{E}(\nu(\psi_1(x)+1) \cdots (\psi_m(x)+1) : x \in \mathbb{Z}_N^t) \\
&= \frac{1}{2^m} \sum_{A \subseteq \{1,2,\ldots,m\}} \mathbb{E}\left(\prod_{i \in A} \nu(\psi_i(x)) : x \in \mathbb{Z}_N^t\right).
\end{aligned}$$

Agora mostraremos que cada parcela desta soma é da forma $1+o(1)$.

De fato, em cada um dos termos $\mathbb{E}\left(\prod_{i\in A}\nu(\psi_i(x)) : x \in \mathbb{Z}_N^t\right)$ do somatório podemos aplicar a propriedade descrita na definição da condição das formas lineares, e dado que ν é, por hipótese, pseudoaleatória, temos que cada termo da soma é da forma $1+o(1)$.

Como a soma possui 2^m termos — um para cada subconjunto de $\{1, 2, \ldots, m\}$ —, o resultado da expressão completa é $1+o(1)$.

Resta provar agora que $\nu_{1/2}$ satisfaz a 2^{k-1}-condição de correlação. Para isso, consideremos $m \leq 2^{k-1}$ e $h_1, h_2, \ldots, h_m \in \mathbb{Z}_N$.

Temos, por cálculos semelhantes aos detalhados anteriormente, que

$$\mathbb{E}(\nu_{1/2}(x+h_1) \cdots \nu_{1/2}(x+h_m)) : x \in \mathbb{Z}_N^t) = \frac{1}{2^m} \sum_{A \subseteq \{1,2,\ldots,m\}} \mathbb{E}\left(\prod_{i \in A} \nu(x+h_i) : x \in \mathbb{Z}_N^t\right).$$

[5] Um dos fatos mais notáveis descrito nesta dissertação é que o nome dele realmente não possui os pingos nos i's.

Sabemos que existe τ com a qual ν satisfaz a condição de correlação, e que certamente satisfaz a identidade $\mathbb{E}(\tau(x)^q) = O_{m,q}(1)$, para todo $q \geq 1$.

Consideremos um dos termos da soma. Para cada $A \subseteq \{1, 2, \ldots, m\}$, temos que

$$\mathbb{E}\left(\prod_{i\in A} \nu(x + h_i) : x \in \mathtt{Z}_N\right) \leq \sum_{1\leq i<j\leq m} \tau(h_i - h_j),$$

dado que ν satisfaz a condição de correlação.

Novamente, dado que a soma apresenta 2^m termos, o resultado final nos dá a desigualdade

$$\mathbb{E}(\nu_{1/2}(x + h_1) \cdots \nu_{1/2}(x + h_m)) : x \in \mathbb{Z}_N^t) \leq \sum_{1\leq i<j\leq m} \tau(h_i - h_j),$$

e isto prova o resultado. □

Observação. Claramente, podemos iterar o resultado acima, e assim obtemos que $(\nu_{1/2} + 1)/2$ também é uma medida k-pseudoaleatória.

De fato, definindo ν_n como sendo a n-ésima medida obtida a partir de ν por este processo — e, portanto, temos que $\nu_0 = \nu$ e $\nu_1 = \nu_{1/2} = (\nu + 1)/2$ —, obtemos a fórmula $\nu_n = (\nu + 2^n - 1)/2^n$ e decorre que tal medida é k-pseudoaleatória, resultado este que será extremamente útil nos capítulos que seguirão.

4.3.3 Produto Interno e Norma de Gowers

De um modo incrivelmente engenhoso, Gowers introduziu em seu artigo sobre o Teorema de Szemerédi [Gow01] uma norma que leva em conta conceitos relacionados com a noção de progressões aritméticas.

Assim, um modo de medir e comparar "distâncias" entre conjuntos, através de funções com suporte em progressões aritméticas, foi investigado com todas as ferramentas que um espaço normado proporciona.

Estudar e compreender este caminho será nosso objetivo nesta subseção.

Definição (Produto Interno e Norma de Gowers)**.** Seja $d \geq 1$ um inteiro, e consideremos o conjunto $\{0,1\}^d$, bem como um elemento $\omega = (\omega_1, \ldots, \omega_d) \in \{0,1\}^d$.

Seja $h = (h_1, \ldots, h_d) \in \mathbb{Z}_N^d$, e então definimos a operação $\omega \cdot h := \omega_1 h_1 + \ldots + \omega_d h_d \in \mathbb{Z}_N$ (com a mulplicação entre as coordenadas de ω e de h, bem como as subseqüentes somas, interpretadas da maneira natural em $\mathbb{Z}_N$).

Dada $(f_\omega)_{\omega\in\{0,1\}^d}$ uma $\{0,1\}^d$-úpla de funções $f_\omega : \mathbb{Z}_N \to \mathbb{R}$, temos que o *d-produto interno de Gowers* é definido por

$$\left\langle (f_w)_{\omega\in\{0,1\}^d} \right\rangle_{U^d} := \mathbb{E}\left(\prod_{\omega\in\{0,1\}^d} f_w(x + \omega \cdot h) : x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^d\right),$$

e, se $f_\omega = f$, para todo $\omega \in \{0,1\}^d$, definimos a *d-norma de Gowers* como (usando $f_\omega = f$)

$$\|f\|_{U^d} = \left\|(f)_{\omega\in\{0,1\}^d}\right\|_{U^d} := \left\langle (f)_{\omega\in\{0,1\}^d} \right\rangle_{U^d}^{1/2^d}.$$

Tal quantidade também será denotada por *U^d-norma*.

Definição (η-Uniformidade)**.** Fixada uma dimensão d, dizemos que f é *η-uniforme* se $\|f\|_{U^d} \leq \eta$.

Proposição 4.3.1 (Propriedades das Normas de Gowers)**.** *Seja $(f_\omega)_{\omega\in\{0,1\}^d}$ uma $\{0,1\}^d$-úpla de funções $f_\omega : \mathbb{Z}_N \to \mathbb{R}$, bem como $f, g : \mathbb{Z}_N \to \mathbb{R}$.*

Então as seguintes afirmação são válidas:

1. *A norma é positiva:*
$$\|f\|_{U^d} \geq 0;$$

2. *Desigualdade de Cauchy-Schwarz-Gowers:*
$$\left|\left\langle (f_\omega)_{\omega\in\{0,1\}^d}\right\rangle_{U^d}\right| \leq \prod_{\omega\in\{0,1\}^d} \|f_\omega\|_{U^d};$$

3. *Desigualdade Triangular:*
$$\|f+g\|_{U^d} \leq \|f\|_{U^d} + \|g\|_{U^d};$$

4. *Multiplicação por escalar:*
$$\|\lambda f\|_{U^d} = |\lambda|\,\|f\|_{U^d};$$

5. *Monotonicidade em relação à dimensão:*
$$\|f\|_{U^{d-1}} \leq \|f\|_{U^d};$$

6. *Para $d=1$, a norma de Gowers é o valor absoluto da esperança*[6]*:*
$$\|f\|_{U^1} = |\mathbb{E}(f)|;$$

7. *Para $d\geq 2$, a norma de Gowers é uma norma (de verdade!).*

Todas estas propriedades são provadas no artigo [GT04].

Lema 4.3.2 (Uniformidade de Funções k-Pseudoaleatórias)**.** *Seja ν uma medida k-pseudoaleatória. Então temos que*
$$\|\nu - 1\|_{U^d} = o(1),$$
para todo $1\leq d\leq k-1$.

Demonstração. Dado que a U^d-norma é crescente em relação à d, precisamos provar o resultado apenas para $d=k-1$.

Aplicando a definição da U^d-norma, obtemos
$$\|\nu-1\|_{U^d} = \mathbb{E}\left(\prod_{\omega\in\{0,1\}^d} (\nu(x+\omega\cdot h)-1) : x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^d\right)^{1/2^d},$$
que pode ser expandida para
$$\sum_{A\subseteq\{0,1\}^{k-1}} (-1)^{|A|}\mathbb{E}\left(\prod_{\omega\in\{0,1\}^{k-1}} \nu(x+\omega\cdot h) : x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\right)^{1/2^{k-1}}.$$

Analisemos cada parcela $\mathbb{E}\left(\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h) : x\in\mathbb{Z}_N, h\in\mathbb{Z}_N^{k-1}\right)$ da soma.

Como ν é k-pseudoaleatória, cada uma dessas parcelas (ignorando o sinal) satisfaz a $(2^{k-1},k,1)$-condição das formas lineares, a qual podemos aplicar, já que as formas lineares $x+\omega\cdot h$ não são múltiplas racionais entre si, e portanto cada uma delas vale $1+o(1)$.

Como temos que $|\{A\subseteq\{0,1\}^n : |A| \text{ é par}\}| = |\{A\subseteq\{0,1\}^n : |A| \text{ é ímpar}\}|$, para qualquer $n\in\mathbb{N}$, os sinais se pareiam, e portanto temos que na soma $\sum_{A\subseteq\{0,1\}^{k-1}}(-1)^{|A|}(1+o(1))$ as parcelas se cancelam e o resultado final é da forma $o(1)$. □

[6] Vale ressaltarmos que, neste caso, é possível termos $\|f\|_{U^1}=0$ sem que $f\equiv 0$ e, portanto, $\|\cdot\|_{U^1}$ não é uma norma, sendo apenas o que é chamado de *seminorma*.

Vale a pena olharmos com um pouco mais de atenção para entender o significado deste resultado mais claramente.

Basicamente, ele nos diz que as medidas k-pseudoaleatórias são uniformemente distribuídas sob o ponto de vista de conterem $(k-1)$-progressões aritméticas em seu suporte, já que, ao estarem próximas à função constante 1 numa norma que está essencialmente nos contando a quantidade de k-PAs, garantimos uma espécie de proximidade na contagem de k-PAs para qualquer ν pseudoaleatória.

4.3.4 O Teorema de von Neumann Generalizado

Nosso próximo ingrediente será o Teorema de von Neumann Generalizado, cuja importância reside no fato de que fornece uma relação direta de como a quantidade de PAs contidas no suporte de uma função é governada pela sua norma de Gowers.

Aqui enunciaremos o teorema e veremos suas conseqüências e relações com o Teorema de Green-Tao. A fim de não interromper o fluxo de idéias para o objetivo principal, a apresentação da prova é realizada no Apêndice A.

O enunciado é o seguinte:

Teorema 4.3.2 (von Neumann Generalizado)**.** *Sejam ν uma medida k-pseudoaleatória, e consideremos $f_0, f_1, \ldots, f_{k-1} \in L^1(\mathbb{Z}_N)$ satisfazendo*

$$|f_j(x)| \leq \nu(x), \text{ para todo } x \in \mathbb{Z}_N, 0 \leq j \leq k-1.$$

Então temos que

$$\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x+jr) : x, r \in \mathbb{Z}_N\right) = \Upsilon_k(f_0, f_1, \ldots, f_{k-1}) = O_k\left(\inf_{0\leq j\leq k-1} \|f_j\|_{U^{k-1}}\right) + o_k(1).$$

A importância deste teorema se dá, especialmente, pelo corolário a seguir:

Corolário 4.3.1. *Se ν é linearmente k-pseudoaleatória e $f = g+h$, onde $|g|, |h|$ são ambas limitadas (pontualmente) por ν, e h é η-uniforme, então*

$$\Upsilon_k(f) = \Upsilon_k(g) + O_k(\eta) + o_k(1).$$

Demonstração. Expandindo os termos da esperança, obtemos que

$$\Upsilon_k(f) = \Upsilon_k(g+h) = \Upsilon_k(g) + \sum_{\emptyset \neq I \subseteq \{0,1,\ldots,k-1\}} \Upsilon_k(f_0^I, \ldots, f_{k-1}^I),$$

onde $f_i^I = h$ se $i \in I$, e $f_i^I = g$ se $i \notin I$.

Basta então aplicar o Teorema de von Neumann Generalizado (Teorema 4.3.2) a cada uma das parcelas $\Upsilon_k(f_0^I, \ldots, f_{k-1}^I)$, para $I \neq \emptyset$, juntamente com o fato de que h é η-uniforme (isto é, $\|h\|_{U^{k-1}} \leq \eta$), e obtemos que o somatório é limitado superiormente por $O_k(\eta) + o_k(1)$. □

Observação. Notemos que podemos trabalhar com a hipótese mais geral de que $|f_j(x)| \leq \nu(x) + 1$, já que definindo $\tilde{f}_j := f_j/2$ e $\tilde{\nu} := (\nu+1)/2$, têm-se que, como $|f_j| \leq \nu$, então $|f_j| \leq \nu + 1$ e, daí, concluímos que $|f_j/2| \leq (\nu+1)/2$, isto é, $\tilde{f}_j \leq \tilde{\nu}$.

Supondo que o teorema é válido, podemos aplicá-lo a esta última desigualdade, já que as funções k-pseudoaleatórias formam um conjunto estrelado ao redor de 1 e, portanto, $\tilde{\nu}$ também é k-pseudoaleatória.

Assim, obtemos que

$$\mathbb{E}\left(\prod_{j=0}^{k-1}\tilde{f}_j(x+jr):x,r\in\mathbb{Z}_N\right)=\Upsilon_k(\tilde{f}_0,\tilde{f}_1,\ldots,\tilde{f}_{k-1})=O_k\left(\inf_{0\leq j\leq k-1}\left\|\tilde{f}_j\right\|_{U^{k-1}}\right)+o_k(1).$$

Porém, a expressão acima pode ser convertida para

$$\frac{1}{2^k}\mathbb{E}\left(\prod_{j=0}^{k-1}f_j(x+jr):x,r\in\mathbb{Z}_N\right)=\frac{1}{2^k}\Upsilon_k(f_0,f_1,\ldots,f_{k-1})=O_k\left(\inf_{0\leq j\leq k-1}\frac{1}{2}\left\|f_j\right\|_{U^{k-1}}\right)+o_k(1),$$

que por sua vez é equivalente a

$$\mathbb{E}\left(\prod_{j=0}^{k-1}f_j(x+jr):x,r\in\mathbb{Z}_N\right)=\Upsilon_k(f_0,f_1,\ldots,f_{k-1})=O_k\left(\inf_{0\leq j\leq k-1}\left\|f_j\right\|_{U^{k-1}}\right)+o_k(1),$$

que é a forma original do teorema.

Na verdade, reiterando o processo da demonstração acima e aplicando a observação em 4.3.4, vemos que pode-se usar o mesmo argumento realizado para usar a hipótese $|f_j|\leq\nu+K$, para algum $K>0$, de modo que ainda garantimos o resultado desejado.

Claramente, as constantes "escondidas" nas notações O_k e o_k se alteram em cada caso, e é óbvio que, apesar de para cada N as funções $f_{j,N}:\mathbb{Z}_N\to\mathbb{R}$ serem limitadas, temos que lembrar que estamos lidando com comportamentos assintóticos e com seqüências de funções, de modo que mesmo que cada elemento $f_{j,N}$ seja limitado por uma expressão da forma $\nu+K_N$, pode não existir um único K que limite simultaneamente todos os elementos de f_j, de modo que o teorema de fato possui certas restrições.

4.3.5 Funções Duais e Anti-Uniformes

De maneira geral[7], sempre que temos uma norma $\|\cdot\|$ em $\mathbb{R}^n$, podemos definir sua norma dual da seguinte maneira:

$$\|f\|^*:=\sup\left\{|\langle f,g\rangle|:\|g\|\leq 1\right\}.$$

Além disso, em espaços de dimensão finita, temos que a norma dual é realmente uma norma.

Aplicando este conceito à norma de Gowers, obtemos a seguinte

Definição (Norma de Gowers Dual)**.** Seja $f:\mathbb{Z}_N\to\mathbb{R}$. Definimos a *norma de Gowers dual de* f como sendo

$$\|f\|^*_{(U^d)}:=\sup_{\|g\|_{U^d}\leq 1}\left\{|\mathbb{E}(fg)|\right\}.$$

Nomearemos tal norma como $(U^d)^*$-norma.

Para entendermos a utilidade desta definição, precisamos do seguinte

Lema 4.3.3. *Dadas uma norma $\|\cdot\|$ e um produto interno $\langle\cdot,\cdot\rangle$ (que não precisam estar relacionados entre si), a seguinte desigualdade é válida:*

$$|\langle f,g\rangle|\leq\|f\|\,\|g\|^*.$$

[7] Quando falamos "de maneira geral", não estamos gastando palavras: as normas em questão não precisam ter relação alguma com o produto interno!

Demonstração. Se $f \equiv 0$, o resultado é trivial.

Suponhamos, então, que $f \not\equiv 0$. Neste caso, temos que $\frac{f}{\|f\|}$ possui norma $\|\cdot\|$ de valor unitário, e portanto sabemos que

$$\|g\|^* \geq \left| \left\langle \frac{f}{\|f\|}, g \right\rangle \right| = \frac{1}{\|f\|} \left| \langle f, g \rangle \right|,$$

e portanto $\|f\| \, \|g\|^* \geq |\, \langle f, g \rangle \,|$. □

A ênfase na independência entre a norma $\|\cdot\|$ e o produto interno $\langle \cdot, \cdot \rangle$ é justificada na demonstração anterior, pelo fato de que em nenhum momento foi usado alguma espécie de relação entre estes objetos.

Esta desigualdade nos permite concluir que sempre que $\langle f, g \rangle$ for grande e $\|g\|^*$ for pequena, somos forçados a ter $\|f\|$ também grande. Em outras palavras, quando duas funções se correlacionam fortemente, "pequenez" da norma dual de um dos fatores evita "pequenez" da norma original do outro fator.

Agora definiremos um tipo específico de funções serão de grande valia na obtenção de limitantes superiores:

Definição (Função Dual)**.** Seja $f : \mathbb{Z}_N \to \mathbb{R}$. Definimos a *função dual de* f (pontualmente) por

$$\mathcal{D}f(x) := \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} f(x + \omega \cdot h) : h \in \mathbb{Z}_N^{k-1} \right).$$

A primeira propriedade interessante destas funções é sua relacão com o produto interno da função original:

Lema 4.3.4. *A seguinte identidade é verdadeira:*

$$\langle f, \mathcal{D}f \rangle = \|f\|_{U^{k-1}}^{2^{k-1}}.$$

Demonstração. Usando as propriedades da esperança e a definição de norma de Gowers dual, temos a seguinte cadeia de igualdades:

$$\langle f, \mathcal{D}f \rangle = \mathbb{E}(f \mathcal{D}f) = \mathbb{E} \left(f(x) \mathbb{E} \left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} f(x + \omega \cdot h) : h \in \mathbb{Z}_N^{k-1} \right) : x \in \mathbb{Z}_N \right),$$

Nesta última expressão, consideramos $f(x)$ justamente como um elemento do produtório com $\omega = 0$ e aplicamos o item 2 da proposição 0.0.1 pra obter $\mathbb{E}\left(\prod_{\omega \in \{0,1\}^{k-1}} f(x + \omega \cdot h)\right)$, que é, por definição, $\|f\|_{U^{k-1}}^{2^{k-1}}$. □

Por fim, para que nossos estudos prossigam, precisamos que tais funções estejam relacionadas, de alguma maneira, com as hipóteses de limitação em relação às medidas. Isso é realizado a partir da seguinte

Definição (Funções Anti-Uniformes)**.** Seja ν uma medida k-pseudoaleatória. Uma função $g : \mathbb{Z}_N \to \mathbb{R}$ é chamada uma *função anti-uniforme (com respeito a* ν*)* se existe uma função $f : \mathbb{Z}_N \to \mathbb{R}$ tal que $g = \mathcal{D}f$ e $|f(x)| \leq \nu(x)$, para todo $x \in \mathbb{Z}_N$.

Logo mais, veremos o uso das funções anti-uniformes.

4.3.6 O Teorema da Decomposição

É especificamente nesta parte que nossa abordagem toma um rumo diferente da abordagem original de Green e Tao. No artigo, eles usam certas construções de teoria ergódica para também provar uma espécie de decomposição e aplicar o mesmo tipo de idéia que faremos. Da nossa parte, trilharemos um caminho que utiliza análise funcional para chegar ao mesmo objetivo, seguindo o caminho mostrado por Gowers em [Gow10]. A vantagem (novamente, na opinião do autor) que obtemos é uma enorme simplificação da demonstração. Por outro lado, como é comum ao lidarmos com resultados de análise funcional, a prova é não-construtível, enquanto que a abordagem ergódica possui um aspecto mais algorítmico (ainda que não tão explicitamente construtível).

Nosso objetivo nesta seção será provar o *Teorema da Decomposição*, que afirma que se uma função f satisfaz a condição $0 \leq f \leq \nu$, onde ν é pseudoaleatória, então ela pode ser decomposta em duas componentes $f = g+h$, de modo que $0 \leq g \leq 2$ — que é limitada e portanto podemos, com algumas adaptações simples, utilizar o Teorema de Szemerédi — e h é η-uniforme (com $0 < \eta < 1$) — que, pelo Teorema de von Neumann Generalizado, é negligenciável do ponto de vista das normas de Gowers.

Com isto em mente, poderemos manter praticamente intacta a contagem de k-PAs ao passar de g para f.

Utilizaremos a seguinte estratégia para mostrar que existe a decomposição $f = g + h$ com as propriedades mencionadas:

Suporemos, por absurdo, que tal decomposição não exista e, graças ao Teorema de Hahn-Banach (que podemos aplicar, dado que os conjuntos das funções g com $0 \leq g \leq 2$ e o conjunto das funções η-uniformes são convexos e fechados), chegaremos a uma contradição.

Em termos bem gerais, esta contradição será alcançada com o seguinte raciocínio: se a decomposição não existisse, seria possível provar a existência de uma função ϕ anti-uniforme com $\langle f, \phi\rangle$ grande e que, por outro lado, teria $\langle \nu, \phi\rangle$ pequeno, já que é a soma de $\langle 1, \phi\rangle$, que é limitada, com $\langle \nu - 1, \phi\rangle$, que é $o(1)$, dado que $\nu - 1$ é uniforme e ϕ é anti-uniforme; isto contrariaria, portanto, o fato de que $f \leq \nu$.

Os próximos dois lemas nos darão limitantes para a anti-uniformidade, de modo a nos auxiliar na construção da função anti-uniforme ψ que irá aproximar ϕ (na verdade, ϕ_+) no lema 4.3.10.

Lema 4.3.5 (Funções Anti-Uniformes São Limitadas)**.** *Seja $0 \leq f \leq \nu$, para alguma ν pseudoaleatória.*

Então $\|\mathcal{D}f\|_\infty = O_k(1)$.

Demonstração. Se $0 \leq f \leq \nu$, então $0 \leq f \leq \nu + 1$ e, portanto, $0 \leq f \leq 2\nu_{1/2}$.

Daí, segue que

$$\|\mathcal{D}f\|_\infty = \sup_{x\in\mathbb{Z}_N} \{|\mathcal{D}f(x)|\} = \sup_{x\in\mathbb{Z}_N} \left\{ \left| \mathbb{E}\left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} f(x+\omega\cdot h) : h \in \mathbb{Z}_N^{k-1} \right) \right| \right\}$$

$$\leq \sup_{x\in\mathbb{Z}_N} \left\{ \left| \mathbb{E}\left(\prod_{\substack{\omega \in \{0,1\}^{k-1} \\ \omega \neq 0}} 2\nu_{1/2}(x+\omega\cdot h) : h \in \mathbb{Z}_N^{k-1} \right) \right| \right\}.$$

Como $\{0,1\}^{k-1}\setminus\{0\}$ possui $2^{k-1}-1$ elementos, temos, por linearidade, que o fator 2 possui um expoente igual a $2^{k-1}-1$, isto é, temos que a expressão anterior é igual a

$$2^{2^{k-1}-1}\sup_{x\in\mathbb{Z}_N}\left\{\left|\mathbb{E}\left(\prod_{\substack{\omega\in\{0,1\}^{k-1}\\ \omega\neq 0}}\nu_{1/2}(x+\omega\cdot h):h\in\mathbb{Z}_N^{k-1}\right)\right|\right\}.$$

Dado que $\nu_{1/2}$ também é uma medida k-pseudoaleatória (vide Lema 4.3.1), podemos aplicar a $(2^{k-1}-1,k-1,1)$-condição das formas lineares com as $2^{k-1}-1$ formas lineares $x+\omega\cdot h$, para $\omega\neq 0$, que, claramente, não são múltiplas racionais entre si, e então teremos que cada um dos valores de $\mathbb{E}\left(\prod_{\substack{\omega\in\{0,1\}^{k-1}\\ \omega\neq 0}}\nu_{1/2}(x+\omega\cdot h):h\in\mathbb{Z}_N^{k-1}\right)$ é da forma $1+o(1)$, seja qual for $x\in\mathbb{Z}_N$.

Logo, a expressão é da forma $2^{2^{k-1}-1}(1+o(1))$ e podemos obter a majoração $\|\mathcal{D}f\|_\infty\leq 2^{2^{k-1}-1}+o(1)$ e, enfim, $\|\mathcal{D}f\|_\infty=O(1)$. □

Lema 4.3.6 (Produtos Anti-Uniformes São Anti-Uniformes[8])**.** *Sejam ν uma medida pseudoaleatória e $0\leq f_j\leq\nu$, para $j=1,\dots,m$.*
Então $\|\mathcal{D}f_1\cdots\mathcal{D}f_m\|_{U^{k-1}}^*=O_m(1)$.

A prova deste lema, apesar de não ser necessariamente difícil ou necessitar de alguma dose extra de criatividade, é de um caráter razoavelmente técnico. Desta maneira, nos abstemos de realizá-la aqui e indicamos a prova dada no artigo original [GT04].

Nosso objetivo, agora, é construir uma função ψ, de tal maneira que seja uma aproximação anti-uniforme para ϕ_+. Mostraremos que ϕ é uma "pequena" combinação linear de funções uniformes básicas, e usaremos este fato para construirmos um polinômio em ϕ que, além de ser anti-uniforme, aproxima ϕ_+.

Para isso, precisaremos da seguinte norma:

Definição (Norma Básica)**.** Dada uma função $\phi:\mathbb{Z}_N\to\mathbb{R}$, bem como uma medida pseudoaleatória $\nu>0$, definimos a *norma básica de ϕ (em relação a ν)* como sendo

$$\|\phi\|_B:=\inf\left\{\sum_{i=1}^n|\lambda_i|\ :\phi=\sum_{i=1}^n\lambda_i\mathcal{D}f_i,\ 0\leq f_i\leq\nu,i=1,\dots,n\right\}.$$

Além disso, se $\|\phi\|_B\leq\eta$, dizemos que ϕ é *η-básica.*

Pode-se dizer que a norma $\|\cdot\|_B$ mede o quão bem podemos aproximar seu argumento por uma combinação linear de funções anti-uniformes (que, pelos lemas anteriores, são bem-comportadas).

É um simples exercício mostrar que a definição realmente satisfaz as condições de uma norma - *caso esteja bem-definida.* Para conseguirmos isso, ainda precisamos do fato de que exista um conjunto $\{\mathcal{D}f_i\}_{1\leq i\leq n}$ que seja gerador. Isso será provado no seguinte

Lema 4.3.7. *Existe um conjunto de funções $\{f_j\}_{1\leq j\leq N}$ tal que $\{\mathcal{D}f_j\}_{1\leq j\leq N}$ gera o espaço das funções de $\mathbb{Z}_N$ em $\mathbb{R}$.*

Demonstração. Tomemos $f_j:\mathbb{Z}_N\to\mathbb{R}$ como a função $f_j=\chi_j$ (a função característica de j). Claramente, este é um conjunto linearmente independente.

[8] Aqui, grande parte da condição de pseudoaleatoriedade simples é requerida. Vale notar que, se um modo de evitar este lema fosse encontrado, o Teorema da Decomposição poderia ser provado somente com a hipótese de que $\|\nu-1\|_{U^d}$ é suficientemente pequeno, para algum d suficientemente grande.

Seja $d \geq 2$. Primeiramente, notemos que para $h = (0, 0, ..., 0)$, temos que a expressão se torna $\prod_{\omega \in \{0,1\}^d,\, \omega \neq 0} f_j(x)$.

Caso $x = j$, esta expressão retorna 1, e como $f_j \geq 0$, temos que as outras parcelas da esperança não afetam a positividade de $f_j(j)$ e, portanto, $\mathcal{D}f_j(j) > 0$.

Caso $x \neq j$, para a parcela com $h = (0, 0, ..., 0)$ temos que $\prod_{\omega \in \{0,1\}^d, \omega \neq 0} f_j(x) = 0$.

Analisemos agora as outras parcelas da esperança.

Como as outras parcelas envolvem $h \neq (0, 0, ..., 0)$, existe um k tal que $h = (h_1, h_2, ..., h_k, ..., h_d)$ com $h_k \neq 0$.

Consideremos então $\omega = (1, 1, ..., 1)$ e $\tilde{\omega} = (1, 1, ..., 1, 0, 1, ..., 1)$, onde o 0 está na posição k.

Isso nos dá que $\omega \cdot h = 1h_1 + 1h_2 + ... + 1h_d$ e $\tilde{\omega} \cdot h = 1h_1 + 1h_2 + ... + 1h_d - h_k$. Como $h_k \neq 0$, estas expressão são claramente diferentes, e portanto, por f_j ser uma função característica de um conjunto unitário, podemos concluir que $f_j(x + \omega \cdot h)$ ou $f_j(x + \tilde{\omega} \cdot h)$ é igual a 0, e daí teremos que a expressão $\prod_{\omega \in \{0,1\}^d, \omega \neq 0} f_j(x + \omega \cdot h)$ é nula.

Com isso, concluímos que $\mathcal{D}f_j(x) > 0$, se $x = j$, e $\mathcal{D}f_j(x) = 0$, se $x \neq j$.

Mais explicitamente, temos que $\mathcal{D}f_j(x) = \lambda(x) f_j(x)$, onde λ é uma função estritamente positiva.

Com estes resultados, claramente temos que $\mathcal{D}f_1, \mathcal{D}f_2, ..., \mathcal{D}f_N$ formam um conjunto linearmente independente, e portanto a norma básica está bem-definida. [9] □

Além disso, notemos que dado que $\|\cdot\|_B$ é uma norma em um espaço de dimensão finita, é uma norma reflexiva, isto é, temos que $\|\cdot\|_B = \|\cdot\|_B^{**}$. Uma demonstração deste fato pode ser encontrada em [Hö90].

É fácil deduzir os seguintes resultados análogos aos lemas 4.3.5 e 4.3.6:

Lema 4.3.8 (Funções Básicas São Limitadas)**.** *Se ϕ é η-básica, então $\|\phi\|_\infty = O(\eta)$.*

Lema 4.3.9 (Potências Básicas São Anti-Uniformes)**.** *Se ϕ é η-básica, então $\|\phi^m\|_{U^{k-1}}^* = O_m(\eta^m)$, para todo m inteiro positivo.*

O seguinte lema nos dá a construção de uma aproximação anti-uniforme para ϕ, desde que ela seja "suficientemente" básica.

Lema 4.3.10 (Aproximação Por Polinômio Anti-Uniforme)**.** *Existe um polinômio P tal que, se $0 < \eta < 1$, então para toda função ϕ que for η-básica, temos que*

1. $\|P\phi - \phi_+\|_\infty \leq \frac{1}{8}$;
2. $\|P\phi\|_{U^{k-1}}^* \leq A$, *para alguma constante A dependente apenas de η.*

Demonstração. Pelo lema 4.3.8, sabemos que para toda ϕ que seja η-básica, vale que $\|\phi\|_\infty \leq C\eta < C$, para alguma constante C independente de ϕ e η. Pelo Teorema de Weierstrass, obtemos a existência de um polinômio P tal que valha $|P(x) - x_+| \leq \frac{1}{8}$ em $[-C, C]$, e portanto $\|P\phi - \phi_+\|_\infty \leq \frac{1}{8}$. Notemos que P não depende de ϕ ou η.

Além disso, pelo lema 4.3.9, temos que $\|\phi^m\|_{U^{k-1}}^* = O_m(\eta^m)$. Escrevendo P como $a_m x^m + \cdots + a_1 x + a_0$, e aplicando a desigualdade triangular, chegamos a

$$\|P\phi\|_{U^{k-1}}^* \leq |a_m| \, \|\phi^m\|_{U^{k-1}} + \cdots + |a_1| \, \|\phi\|_{U^{k-1}} + |a_0| = O(a_m \eta^m + \cdots + a_1 \eta + a_0),$$

provando o resultado desejado.

□

Finalmente, mostramos que a a condição de $\langle h, \phi \rangle \leq 1$, para toda h η-básica, implica que ϕ é suficientemente básica para que possamos aplicar a construção da aproximação anti-uniforme anterior.

[9] O leitor atento[10] perceberá que ainda falta lidarmos com o fato de que $f \leq \nu$, logo não podemos utilizar as funções características se nossa medida possuir pontos em que se anula. Logo mais, utilizaremos este resultado e veremos que na aplicação necessária, tal questão não afetará as considerações com as quais nos importaremos.

[10] E, agora, também o leitor desatento, já que a situação foi descrita explicitamente...

Lema 4.3.11 (Limitante Básico)**.** *Se $\langle h, \phi\rangle \leq 1$ para toda h η-uniforme e que satisfaz $0 \leq h \leq \nu$, então ϕ é $\eta^{-2^{k-1}}$-básica.*

Demonstração. Primeiramente, notemos que se $\|h\|_B^* \leq \eta^{2^{k-1}}$, então temos que $\|h\|_{U^{k-1}} \leq \eta$.

Para perceber isto, basta ver que

$$\|h\|_B^* = \sup_{\|g\|_B \leq 1} \{|\langle h, g\rangle|\} \leq \eta^{2^{k-1}}.$$

Porém, como por definição $\|\mathcal{D}h\|_B \leq 1$, temos que $|\langle h, \mathcal{D}h\rangle| \leq \eta^{2^{k-1}}$, e pelo Lema 4.3.4 sabemos que esta expressão é equivalente a $\|h\|_{U^{k-1}}^{2^{k-1}} \leq \eta^{2^{k-1}}$, de onde concluímos que h é η-uniforme.

Assim sendo, dado que $\|\cdot\|_B$ é reflexiva (isto é, $\|\cdot\|_B^{**} = \|\cdot\|_B$), temos que

$$\|\phi\|_B = \|\phi\|_B^{**} = \sup_{\|g\|_B^* \leq 1} \{|\langle g, \phi\rangle|\} = \sup_{\|h\|_B^* \leq \eta^{2^{k-1}}} \left\{\left|\left\langle \frac{h}{\|h\|_B^*}, \phi \right\rangle\right|\right\} = (\|h\|_B^*)^{-1} \sup_{\|h\|_B^* \leq \eta^{2^{k-1}}} \{|\langle h, \phi\rangle|\}.$$

Sabemos pela observação imediatamente anterior que $\|h\|_B^* \leq \eta^{2^{k-1}}$ implica em h ser η-uniforme; além disso, pela hipótese da η-uniformidade de h resulta que $\langle h, \phi\rangle \leq 1$, nos permitindo concluir que

$$\|\phi\|_B = \|\phi\|_B^{**} = (\|h\|_B^*)^{-1} \cdot \sup_{\|h\|_B^* \leq \eta^{2^{k-1}}} \{|\langle h, \phi\rangle|\} \leq (\eta^{2^{k-1}})^{-1} \cdot 1 \leq \eta^{-2^{k-1}},$$

que é o resultado que queríamos demonstrar. □

Agora basta concatenar os lemas 4.3.10 e 4.3.11 para obter o seguinte

Teorema 4.3.3. *Se $\langle h, \phi\rangle \leq 1$ para toda h η-uniforme, então existe um polinômio P que satisfaz $\|P\phi - \phi_+\|_\infty \leq \frac{1}{8}$ e $\|P\phi\|_{U^{k-1}}^* \leq A$, para alguma constante A dependente apenas de η.*

Demonstração. Imediata. □

Podemos, então, finalmente completar nossa estratégia e provar o Teorema da Decomposição:

Teorema 4.3.4 (Decomposição)**.** *Seja ν simplesmente pseudoaleatória, e η um parâmetro que satisfaz $0 < \eta < 1$, e suponha que N é suficientemente grande (dependendo de η).*

Então, para toda função f com $0 \leq f \leq \nu$, podemos realizar a decomposição $f = g + h$, onde $0 \leq g \leq 2$ e h é η-uniforme.

Demonstração. Comecemos verificando que os conjuntos $K_1 = \{g : \mathbb{Z}_N \to \mathbb{R} : 0 \leq g \leq 2\}$ e $K_2 = \{h : \mathbb{Z}_N \to \mathbb{R} : \|h\|_{U^{k-1}} \leq \eta\}$ são convexos. Para isso, tomemos $t \in [0, 1]$.

Dadas $g_1, g_2 \in K_1$, temos que $0 \leq (1-t)g_1 + tg_2 \leq (1-t) \cdot 2 + t \cdot 2 = 2$, o que prova que o conjunto K_1 é convexo.

Por outro lado, dadas $h_1, h_2 \in K_2$, temos que $\|(1-t)h_1 + th_2\|_{U^{k-1}} \leq \|(1-t)h_1\|_{U^{k-1}} + \|th_2\|_{U^{k-1}} \leq (1-t)\|h_1\|_{U^{k-1}} + t\|h_2\|_{U^{k-1}} \leq (1-t)\eta + t\eta = \eta$, e portanto K_2 também é convexo.

Assim, como K_1 e K_2 são convexos, fechados e contêm a origem, podemos aplicar a eles o Teorema de Hahn-Banach (Teorema 2.1.5).

Suponhamos, por absurdo, que tal decomposição não exista. Então, aplicando o Teorema de Hahn-Banach, obtemos a existência de uma ϕ tal que

1. $\langle f, \phi\rangle > 1$;
2. $\langle g, \phi\rangle \leq 1$, para toda g com $0 \leq g \leq 2$;

3. $\langle h, \phi \rangle \leq 1$, para toda h η-uniforme.[11]

Notemos agora que, se definirmos uma função $\tilde{g}$, com $\tilde{g}(x) = 2$, se $\phi(x) \geq 0$, e $\tilde{g}(x) = 0$, se $\phi(x) < 0$, podemos aplicar a condição 2 acima a $\tilde{g}$, de modo a deduzir que $2\langle 1, \phi_+ \rangle = \langle 2, \phi_+ \rangle = \langle \tilde{g}, \phi \rangle \leq 1$, e portanto concluímos que $\langle 1, \phi_+ \rangle \leq 1/2$.

Pelo Teorema 4.3.3 deduzimos, pela condição 3 obtida pelo Teorema de Hahn-Banach, que existe um polinômio P tal que $\|P\phi - \phi_+\|_\infty \leq \frac{1}{8}$ e $\|P\phi\|^*_{U^{k-1}} \leq A$, para alguma constante A dependente apenas de η. Assim, obtemos o limitante

$$\begin{aligned}\langle \nu, \phi_+ \rangle &= \langle 1, \phi_+ \rangle + \langle 1, P\phi - \phi_+ \rangle + \langle \nu - 1, P\phi \rangle + \langle \nu, \phi_+ - P\phi \rangle \\ &\leq \frac{1}{2} + \frac{1}{8} + A\|\nu - 1\|_{U^{k-1}} + \frac{1}{8}(1 + o(1)) = \frac{3}{4} + o(1),\end{aligned}$$

já que $\|\nu - 1\|_{U^{k-1}} = o(1)$.

Como $f \leq \nu$, e tanto f quanto ϕ_+ são positivas, podemos deduzir que $\langle f, \phi_+ \rangle \leq \langle \nu, \phi_+ \rangle$.

Tendo em vista a condição 1 obtida anteriormente pelo Teorema de Hahn-Banach, juntamente com as desigualdades imediatamente anteriores, chegamos a

$$1 < \langle f, \phi \rangle \leq \langle f, \phi_+ \rangle \leq \langle \nu, \phi_+ \rangle \leq \frac{3}{4} + o(1),$$

que é, obviamente, uma contradição, com o que podemos de fato concluir que a decomposição desejada existe. □

4.3.7 O Enunciado e a Prova do Teorema de Szemerédi Relativo

Agora, iremos finalmente provar o teorema principal desta seção, que permitirá que usemos o Teorema de Szemerédi original em contextos mais amplos (e possibilitará seu posterior uso no caso dos primos).

Teorema 4.3.5 (Teorema de Szemerédi Relativo[12])**.** *Sejam $k \geq 3$ e $\delta > 0$, e seja $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ uma medida k-pseudoaleatória. Suponha que $f : \mathbb{Z}_N \to \mathbb{R}$ satisfaz $0 \leq f(x) \leq \nu(x)$, para todo $x \in \mathbb{Z}_N$, e $\mathbb{E}(f) \geq \delta$.*

Então, para todo N suficientemente grande, temos que

$$\Upsilon_k(f) \geq \frac{c_{k,\delta/3}}{2},$$

onde $c_{k,\delta} > 0$ é a constante do Teorema de Szemerédi-Gowers (Teorema 3.3.4).

Demonstração. Seja η um parâmetro que definiremos posteriormente e que satisfaz $0 < \eta < 1$, e seja $f = g + h$ a decomposição dada pelo Teorema da Decomposição, isto é, $0 \leq g \leq 2$ e h é η-uniforme.

Uma primeira idéia seria aplicar o Teorema de Szemerédi diretamente em g, mas esta é limitada superiormente por 2 ao invés de 1, e sua densidade (esperança), dado que a decomposição é dependente de η, é limitada inferiormente por uma função de η, o que impossibilita a aplicação direta do teorema.

Para contornar este obstáculo, consideramos a função $\dfrac{g + \eta}{2 + \eta}$.

[11] Em relação ao problema citado na nota de rodapé do Lema 4.3.7, eis o motivo: garantimos a existência de uma ϕ com as propriedades requeridas. Construímos então $\tilde{\phi}(x) := \begin{cases} \phi(x), & \text{caso } \nu(x) > 0 \\ 0, & \text{caso contrário.} \end{cases}$. Neste caso, temos que os produtos internos $\left\langle f, \tilde{\phi} \right\rangle, \left\langle g, \tilde{\phi} \right\rangle$ e $\left\langle h, \tilde{\phi} \right\rangle$ ainda satisfazem as propriedades 1, 2 e 3 usadas na demonstração do Teorema 4.3.4, e é de tal forma que $supp(\tilde{\phi}) \subseteq supp(\nu)$.

[12] Compare com a versão anteriormente enunciada do Teorema de Szemerédi Relativo (Teorema 4.3.1), na página 19.

Agora temos que

$$\mathbb{E}\left(\frac{g+\eta}{2+\eta}\right) = \frac{\mathbb{E}(f)-\mathbb{E}(h)+\eta}{2+\eta} \geq \frac{\delta}{2+\eta} > \frac{\delta}{3},$$

dado que $|\mathbb{E}(h)| \leq \mathbb{E}(|h|) = \|h\|_{U^1} \leq \|h\|_{U^{k-1}} \leq \eta$. Além disso, temos também que $0 \leq \dfrac{g+\eta}{2+\eta} \leq 1$.

Portanto, as condições do Teorema de Szemerédi-Gowers são satisfeitas para a função $\dfrac{g+\eta}{2+\eta}$, e deduzimos que existe uma constante $c_{k,\delta/3} > 0$, dependente somente de k e δ que satisfaz

$$\Upsilon_k(g+\eta) \geq \Upsilon_k\left(\frac{g+\eta}{2+\eta}\right) \geq c_{k,\delta/3}.$$

Como $\eta < 2$, substituindo os majorantes na definição de Υ_k, temos que

$$\Upsilon_k(f_0, \ldots, f_{k-1}) \leq 2^{k-1}\eta = O_k(\eta)$$

sempre que $f_j = \eta$ ou $f_j = g$, para $0 \leq j \leq k-1$, e pelo menos um dos f_i é igual a η. Logo, $\Upsilon_k(g) \geq c_{k,\delta/3} - O_k(\eta)$.

Agora, percebamos que se $f = g + h$, então $h = f - g$, e a partir daí podemos concluir que $|h| = |f-g| \leq |f| + |g| = f + g$, já que f e g são funções positivas. Além disso, como $0 \leq f \leq \nu$ e $0 \leq g \leq 2$, temos que $f + g \leq \nu + 2$, resultando em $|g|, |h| \leq \nu + 2$, e assim, aplicando a observação resultante do corolário do Teorema de von Neumann Generalizado (página 24), deduzimos que

$$\Upsilon_k(f) = \Upsilon_k(g) + O_k(\eta) + o_k(1),$$

ou, mais explicitamente,

$$\Upsilon_k(f) = c_{k,\delta/3} + O_k(\eta) + o_k(1).$$

Tomando η pequeno o suficiente e N suficientemente grande, podemos garantir que as quantidades $O_k(\eta)$ e $o_k(1)$, são, ambas, menores do que $\dfrac{c_{k,\delta/3}}{4}$ em valor absoluto.

Concluímos, portanto, que

$$\Upsilon_k(f) \geq \frac{c_{k,\delta/3}}{2},$$

como desejado. □

4.4 Uma Medida Para Os Primos

Na seção anterior, mostramos que há uma classe maior de conjuntos em que podemos deduzir resultados semelhantes aos que obtemos ao utilizar o Teorema de Szemerédi original, se usarmos o Teorema de Szemerédi Relativo.

Claramente, como já discutimos na introdução, não podemos aplicar o Teorema de Szemerédi diretamente na função χ_P, já que o fato de que o conjunto dos primos possuir densidade nula é equivalente a $\mathbb{E}(\chi_P) = O\left(\frac{1}{\log N}\right) = o(1)$.

Assim, precisamos compreender melhor algumas propriedades dos números primos[13], a fim de podermos usar o Teorema de Szemerédi Relativo para concluirmos o resultado que desejamos.

Uma primeira idéia seria então modificar tal função, o que é feito por um ponderamento, obtido a partir da *função de von Mangoldt*[14]:

[13] Um fato realmente incrível é a conclusão de Tao de que a única informação necessária sobre os primos para o desenvolvimento do Teorema de Green-Tao é a de que a função ζ possui um pólo simples em $z = 1$. Apesar disso, seguir por este caminho, apesar de elementar (no sentido matemático), não é nada fácil ou simples, de modo que nosso estudo resolveu seguir um caminho diferente, que apesar de usar resultados preliminares (ver o Apêndice B), é mais compreensível para o leitor. O texto em que tal abordagem é explicitada pode ser encontrado em [Tao].

[14] Na verdade, apesar de estarmos inicialmente definindo a função de Mangoldt da maneira usual, com suporte sobre as *potências dos primos*, para nossas considerações isto adicionaria pouca vantagem, de maneira que logo mais

Definição (Função de von Mangoldt)**.** A seguinte função é conhecida como *função de von Mangoldt*:

$$\Lambda(n) := \begin{cases} \log p, & \text{se } n = p^k \text{ para algum } p \text{ primo;} \\ 0, & \text{caso contrário.} \end{cases}$$

.

Esta função possui a propriedade de que $\mathbb{E}(\Lambda) = 1 + o(1)$, porém não é limitada superiormente, impossibilitando o uso direto do Teorema de Szemerédi, de modo que iremos usar o Teorema de Szemerédi *Relativo*.

Uma das maiores dificuldades ao lidar com o conjunto dos primos é o fato de que *os primos não possuem comportamento aritmético aleatório*. Para perceber isto, basta notar que, num conjunto pseudoaleatório, o fato de possuir semelhança aditiva com os naturais o forçaria a ter uma quantidade próxima de números pares e ímpares, o que é evidentemente falso no caso dos primos (já que é possível provar que a quantidade de primos pares é *muito* menor que a quantidade de primos ímpares).

Tal dificuldade será superada graças à observação de que não precisamos que o conjunto tenha um comportamento *totalmente* pseudoaleatório, basta que seu comportamente seja *suficientemente* aleatório, num sentido que irá depender do tamanho da progressão aritmética considerada.

Para isso usaremos o chamado W-truque, onde olhamos para classes de congruências específicas, de modo a evitar certos obstáculos como, por exemplo, a falta de uniformidade na distribuição dos primos nas classes de congruências de certos resíduos.

4.4.1 O W-Truque

Definimos $w = w(N) := \log\log N$ e

$$W := \prod_{p \leq w} p,$$

isto é, o produto de todos os primos menores que w.

Desta forma, nos restringiremos a analisar a classe de congruência $n \equiv 1 \pmod{W}$, e então olharemos para Λ não no intervalo $\{1, 2, \dots N\}$, mas ao invés disso a modificaremos para a analisarmos tendo como seu suporte o conjunto $\{W + 1, 2W + 1, \dots, NW + 1\}$.

4.4.2 Construindo as Funções

Aplicando o W-truque e modificando um pouco a função de von Mangoldt anteriormente definida:

Definição (Função de von Mangoldt Modificada)**.** Seja ϕ a função totiente de Euler. Definimos então a função

$$\tilde{\Lambda}(n) := \begin{cases} \frac{\phi(W)}{W} \log(Wn + 1), & \text{se } Wn + 1 \text{ é primo;} \\ 0, & \text{caso contrário.} \end{cases}$$

Uma propriedade importante desta função é o seguinte resultado:

Proposição 4.4.1. *Para N suficientemente grande, vale que*

$$\mathbb{E}(\tilde{\Lambda}(x)\chi_{[\varepsilon N, 2\varepsilon N]}(x) : x \in \mathbb{Z}_N) = \varepsilon(1 + o(1)).$$

Demonstração. Para provar esta afirmação, basta expandirmos a notação do funcional esperança e usarmos o Teorema do Número Primo para Progressões Aritméticas (Teorema 2.4.3).

a modificaremos para versões mais simples da função.

$$\begin{aligned}\mathbb{E}(\tilde{\Lambda}(x)\chi_{[\varepsilon N,2\varepsilon N]}(x) : x \in \mathbb{Z}_N) &= \frac{\phi(W)}{W}\frac{1}{N}\sum_{\substack{\varepsilon N \leq n \leq 2\varepsilon N \\ Wn+1 \text{ é primo}}} \log(Wn+1) \\ &= \frac{1}{N}\frac{\phi(W)}{W}\sum_{\substack{\varepsilon WN \leq p \leq 2\varepsilon WN \\ p \equiv 1 \pmod W}} \log(p) + o(1) \\ &= \frac{1}{N}\frac{\phi(W)}{W}\left(\frac{2\varepsilon WN}{\phi(W)} - \frac{\varepsilon WN}{\phi(W)}\right)(1+o(1)) \\ &= \varepsilon(1+o(1)).\end{aligned}$$

□

Observemos agora que, se $n = p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_k^{\alpha_k}$, com $p_1 < p_2 < \ldots < p_k$, é a decomposição em primos de n, temos que

$$\sum_{d|n}\Lambda(d) = \sum_{j=1}^{k}\sum_{r=1}^{\alpha_j}\Lambda(p_j^r) = \sum_{j=1}^{k}\sum_{r=1}^{\alpha_j}\log p_j = \sum_{j=1}^{k}\alpha_j \log p_j = \log n,$$

e portanto, aplicando a Foŕmula da Inversão de Möbius (Teorema 2.4.1), obtemos a identidade $\Lambda(n) = \sum_{d|n}\mu(d)\log_+(n/d)$.

Isto motiva a seguinte definição, usada por Goldston e Yıldırımem seus resultados:

Definição (Função de von Mangoldt-Goldston-Yıldırım)**.** Seja $R > 0$ um parâmetro. Definimos, então, a seguinte função

$$\Lambda_R(n) := \sum_{\substack{d|n \\ d \leq R}} \mu(d)\log\left(\frac{R}{d}\right) = \sum_{d|n}\mu(d)\log_+\left(\frac{R}{d}\right)$$

Vale a pena notar que a única razão para considerarmos as potências de primos no suporte da função de von Mangoldt (Λ) foi como motivação para definirmos a função de von Mangoldt-Goldston-Yıldırım. Para nossos objetivos, é muito melhor trabalhar desconsiderando potências de primos, que é o que faremos deste ponto em diante.

E, por fim, podemos definir a medida ν:

Definição (Medida de von Mangoldt-Goldston-Yıldırım)**.** Sejam $R = N^{k^{-1}2^{-(k+4)}}$ e $\varepsilon = 2^{-k}/(k+4)!$. Definimos $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ por

$$\nu(n) := \begin{cases}\dfrac{\phi(W)}{W}\dfrac{\Lambda_R(Wn+1)^2}{\log R}, & \text{se } n \in [\varepsilon N, 2\varepsilon N]; \\ 1, & \text{caso contrário.}\end{cases}$$

Definida tal função, resta-nos apenas provar que ν, de fato, possui as propriedades desejadas, isto é, majora $\dfrac{1}{k2^{k+5}}\tilde{\Lambda}$ e, mais ainda, é uma medida k-pseudoaleatória.

Proposição 4.4.2. *Seja N um primo suficientemente grande.*

Temos então que

$$\nu(x) \geq \frac{1}{k2^{k+5}}\tilde{\Lambda}(x),$$

para todo $\varepsilon N \leq x \leq 2\varepsilon N$. Além disso, ν é uma medida k-pseudoaleatória.

Para provar a proposição anterior, precisamos provar os quatro seguintes tópicos:

1. Temos que $\nu(x) \geq 0$, para todo $x \in \mathbb{Z}_N$, e $\nu(x) \geq \frac{1}{k2^{k+5}}\tilde{\Lambda}(x)$, para todo $\varepsilon N \leq x \leq 2\varepsilon N$;
2. ν é uma medida, isto é, $\mathbb{E}(\nu) = 1 + o(1)$;
3. ν satisfaz a $(k2^{k-1}, 3k-4, k)$-condição das formas lineares (ν é linearmente k-pseudoaleatória);
4. ν satisfaz a 2^{k-1}-condição de correlação (ν é simplesmente k-pseudoaleatória).

Por hora, veremos somente que ν majora a função desejada. A prova completa de que ν, além disso, é de fato uma medida k-pseudoaleatória é dada no Apêndice B.

ν majora $\tilde{\Lambda}$

Demonstração. Claramente temos que $\nu(x) \geq 0$, já que cada um dos fatores é positivo, bem como o quadrado em $\tilde{\Lambda}_R$ para todo $x \in \mathbb{Z}_N$, portanto só nos resta provar a parte final.

Suponhamos então que $\varepsilon N \leq x \leq 2\varepsilon N$.

Se $Wn + 1$ não é um primo, então o lado direito da desigualdade é igual a 0, e não nos resta mais nada a provar neste caso.

Se $Wn + 1$ é um primo, temos que provar que

$$\frac{\phi(W)}{W}\frac{\Lambda_R(Wx+1)^2}{\log R} \geq \frac{1}{k2^{k+5}}\frac{\phi(W)}{W}\log(Wx+1),$$

que é equivalente a provar que

$$\frac{\Lambda_R(Wx+1)^2}{\log R} \geq \frac{1}{k2^{k+5}}\log(Wx+1).$$

Como $\varepsilon N \leq x \leq 2\varepsilon N$, temos que $Wx + 1 = O(N\log(N))$, que é maior que R quando N é suficientemente grande, dado que $R = N^a$, com $a < 1$.

Logo, quando $Wx + 1$ é primo, só há um divisor d de $Wx + 1$ que satisfaz $d \leq R$, que é $d = 1$, e portanto

$$\Lambda_R(Wx+1) = \log R = k^{-1}2^{-k-4}\log N,$$

e dividindo por $k^{-1}2^{-k-4}$, vemos que temos que provar que

$$\log N \geq \frac{1}{2}\log(Wx+1),$$

o que é válido para N suficientemente grande, já que $W = O(\log N)$. □

4.5 A Prova do Teorema de Green-Tao

Temos, afinal, todas as ferramentas que precisamos para conseguirmos o resultado desejado, já que possuímos:

- O Teorema de Szemerédi Relativo (Teorema 4.3.5);
- Uma função de contagem com suporte nos primos;
- Uma medida majorante que satisfaz os requisitos que necessitamos.

Assim sendo, vamos provar o teorema principal deste texto:

Teorema 4.5.1 (Green-Tao)**.** *Existem k-PAs formadas exclusivamente por primos, para todo $k \in \mathbb{N}$.*

Demonstração. Tomando

$$f(n) = \frac{1}{k2^{k+5}}\tilde{\Lambda}(n)\chi_{[\varepsilon N, 2\varepsilon N]},$$

temos então que

$$\mathbb{E}(f) = \frac{1}{k2^{k+5}} \sum_{\varepsilon N \leq n \leq 2\varepsilon N} \frac{\tilde{\Lambda}(n)}{N} = \frac{1}{k2^{k+5}}\varepsilon(1 + o(1)) > 0.$$

Além disso, a medida ν de von Mangoldt-Goldston-Yıldırımé k-pseudoaleatória e majora f (conforme visto na subseção 4.4.2).

Aplicando o Teorema de Szemerédi Relativo (Teorema 4.3.5), concluímos que

$$\Upsilon_k(f) \geq c'_{k,\delta},$$

onde $\delta = \varepsilon/k2^{k+5}$.

Além disso, sabemos que o termo $f(x)f(x+r)\dots f(x+(k-1)r)$, com $r = 0$, contribui com $o(1)$ para $\Upsilon_k(f)$.

Como o suporte de f está contido no conjunto dos primos e k é um inteiro arbitrário, temos que existem progressões aritméticas de tamanho k formadas exclusivamente por primos, para qualquer $k \in \mathbb{N}$. □

Capítulo 5

Ao Arbitrariamente Grande...E Além!

Precisamos encarar a realidade: ser matemático é ser insatisfeito. Afinal, depois de ver todo o texto sobre o Teorema de Green-Tao, é normal esperar que a curiosidade inerente aos matemáticos logo tome forma e lance perguntas como "Podemos generalizar tal resultado?", "O que podemos saber sobre as constantes envolvidas no teorema?", ou ainda "Este foi difícil. Qual problema *ainda mais difícil* será meu alvo agora?".

Inúmeros resultados seguiram o artigo de Green e Tao: primos de determinadas formas com progressões aritméticas arbitrariamente longas, majorantes para os primos em k-PAs, etc.

Para ajudarmos o leitor nesta jornada, reunimos alguns resultados e conjecturas para animá-lo:

5.1 Resultados Posteriores

Podemos tentar generalizar o Teorema de Green-Tao olhando de um modo especial para cada uma de suas "componentes principais". Veremos um resultado cujo foco está na generalização do conceito de primos para o plano complexo, e outro que expande a noção de progressões aritméticas geradas com base em polinômios.

5.1.1 Constelações de Primos

A noção de primalidade estende-se naturalmente para o domínio dos *inteiros gaussianos* $\mathbb{Z}[i] := \{a + ib : a, b \in \mathbb{Z}\}$, onde p é primo se seus únicos divisores são $\pm 1, \pm i, \pm p$ e $\pm ip$.

Seja J um conjunto finito de índices e $(v_j)_{j\in J} \in (\mathbb{Z}[i])^J$. Chamamos $(v_j)_{j\in J}$ de uma *forma* em $\mathbb{Z}[i]$. Uma *constelação* em $\mathbb{Z}[i]$ com relação a esta forma é qualquer J-upla $(a + rv_j)_{j\in J} \in (\mathbb{Z}[i])^J$ (com $a, r \in \mathbb{Z}[i]$ inteiros gaussianos distintos).

A existência de muitas constelações formadas por primos gaussianos foi mostrada por Tao em [Tao05c] com o seguinte

Teorema 5.1.1. *Seja $(v_j)_{j\in J}$ uma forma qualquer de inteiros gaussianos.*
Então os primos gaussianos contêm infinitas constelações desta forma.

5.1.2 Progressões Polinomiais

Podemos também generalizar a noção de progressão aritmética da seguinte maneira: consideramos a k-PA constituída por $x, x + y, x + 2y, \ldots, x + (k - 1)y$ como sendo da forma $x, x + P_1(y), x + P_2(y), \ldots, x + P_k(y)$ (ou, resumidamente, $x + [P_1, P_k]$), onde $P_i(y) = (i - 1)y$, para $i = 1, 2, \ldots, k$.

Assim, podemos estender a noção de progressões permitindo que $P_i \in \mathbb{Z}[y]$ sejam quaisquer polinômios com coeficientes inteiros tais que $P_i(0) = 0$.

Um teorema de existência de tais progressões polinomiais formadas exclusivamente por primos foi demonstrada por [TZ08], com o seguinte

Teorema 5.1.2. *Sejam $P_1, P_2, \ldots, P_k$ polinômios satisfazendo as condições acima.*

Então dado $\varepsilon > 0$, existem infinitos inteiros x e m tais que $1 \leq m \leq x^{\varepsilon}$ e $x + P_i(m)$ são primos para $i = 1, 2, \ldots, k$.

5.2 Perguntas

Por fim, há também algumas perguntas que se seguem diretamente do resultado de Green-Tao:

Pergunta 5.2.1. *Existe uma k-PA formada por primos* consecutivos*, para todo $k \in \mathbb{N}$?*

Pergunta 5.2.2. *Como estimar $p_{GT}(k)$, o menor primo tal que, para algum inteiro $r > 0$, temos que $p_{GT}(k) + [0, k)r$ é uma k-PA formada exclusivamente por primos?*

Pergunta 5.2.3. *Como caracterizar os primos que fazem parte de alguma k-PA exclusiva de primos?*

Pergunta 5.2.4. *Como as séries $\sum_{p \in P(k)} \frac{1}{p}$ se comportam, onde $P(k)$ são os primos que pertencem a alguma k-PA exclusiva de primos?*

Entre tantas que poderiam ser feitas, estas duas são algumas das mais diretas e interessantes, do ponto de vista do autor, que encerra este texto na pretensão de, um dia, conhecer as respostas.

Apêndice A

O Teorema de von Neumann Generalizado

Como vimos, o Teorema de von Neumann Generalizado (Teorema A.2.1), apresentado na Subseção 4.3.4, foi crucial para a prova do Teorema de Szemerédi Relativo (Teorema 4.3.5).

Neste apêndice realizaremos a apresentação dos conceitos preliminares necessários, para então enunciarmos e provarmos o resultado.

Este capítulo foi fortemente baseado em [Jen09].

A.1 Preliminares

Primeiramente, precisaremos introduzir uma notação e provar mais uma versão da Desigualdade de Cauchy-Schwarz:

Definição (Vetor Indicador de S)**.** Sejam $k \geq 2$ e $0 \leq d \leq k-1$, e consideremos os vetores $y = (y_1, \ldots, y_{k-1}) \in \mathbb{Z}_N^{k-1}$ e $y' = (y'_{k-d}, \ldots, y'_{k-1}) \in \mathbb{Z}_N^d$ e $S \subseteq \{k-d, \ldots, k-1\}$.

Então definimos o vetor $y^{(S)} = (y_1^{(S)}, \ldots, y_{k-1}^{(S)}) \in \mathbb{Z}_N^{k-1}$ por

$$y_i^{(S)} = \begin{cases} y'_i, & \text{se } i \in S; \\ y_i, & \text{caso contrário.} \end{cases}$$

Antes de enunciarmos e provarmos esta nova versão da Desigualdade de Cauchy-Schwarz, lidaremos com a seguinte proposição, que terá sua utilidade.

Proposição A.1.1. *Seja $y \in \mathbb{Z}_N^{k-1}$ e consideremos $\phi_0(y) = \sum_{i=1}^{k-1} y_i$.*

A seguinte igualdade é válida:

$$\left\{\phi_0(y^{(S)}) : S \subseteq \{1, \ldots, k-1\}, y' \in \mathbb{Z}_N^{k-1}\right\} = \left\{\phi_0(y) + \omega \cdot h : \omega \in \{0,1\}^{k-1}, h \in \mathbb{Z}_N^{k-1}\right\}.$$

Demonstração. Consideremos $\phi_0(y^{(S)})$ um elemento do lado esquerdo da igualdade e consideremos $\omega \in \{0,1\}^{k-1}$ definido por $\omega_i = \chi_S(i)$, para $1 \leq i \leq k-1$.

Então, se $h = (h_1, \ldots, h_{k-1})$ é dado por $h_i = y'_i - y_i$, vem que

$$\phi_0(y^{(S)}) = \sum_{i \in S} y'_i + \sum_{i \notin S} y_i = y_1 + \ldots + y_{k-1} + \sum_{i=1}^{k-1} \chi_S(i)(y'_i - y_i) = \phi_0(y) + \omega \cdot h,$$

logo $\phi_0(y^{(S)})$ é também um elemento do lado direito da igualdade, com os ω e h especificados.

Agora seja $x + \omega \cdot h$ um elemento do lado direito da igualdade. Seja $S \subseteq \{1, \ldots, k-1\}$ definido por $i \in S \Leftrightarrow \omega_i = 1$, e seja $y' \in \mathbb{Z}_N^{k-1}$ definida por $y'_i = h_i + y_i$. Os mesmos cálculos acima nos garantem que $x + \omega \cdot h = \phi_0(y^{(S)})$, o que prova o A.1.1. □

Um dos principais resultados que usaremos partindo da definição anterior é (mais) uma versão da Desigualdade de Cauchy-Schwarz:

Lema A.1.1 (Desigualdade de Cauchy-Schwarz Alternativa)**.** *Sejam $k \geq 2$, $\nu : \mathbb{Z}_N \to \mathbb{R}^+$ uma medida, e $\phi_0, \phi_1, \ldots, \phi_{k-1} : \mathbb{Z}_N^{k-1} \to \mathbb{Z}_N$ tais que, para todo i, temos que ϕ_i não depende da i-ésima coordenada. Além disso, sejam $f_0, f_1, \ldots, f_{k-1} : \mathbb{Z}_N \to \mathbb{R}$ tais que*

$$|f_i(x)| \leq \nu(x),$$

para todo $x \in \mathbb{Z}_N$ e $0 \leq i \leq k-1$.

Consideremos, para $0 \leq d \leq k-1$, as seguintes quantidades:

$$J_d = \mathbb{E}\left(\prod_{S \subseteq \{k-d,\ldots,k-1\}} \left(\left(\prod_{i=0}^{k-d-1} f_i(\phi_i(y^{(S)}))\right)\left(\prod_{i=k-d}^{k-1} \nu^{1/2}(\phi_i(y^{(S)}))\right)\right) : y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d\right)$$

e

$$P_d = \mathbb{E}\left(\prod_{S \subseteq \{k-d,\ldots,k-1\}} \nu(\phi_{k-d-1}(y^{(S)})) : y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d\right).$$

Então, para todo $0 \leq d \leq k-2$, temos que

$$|J_d|^2 \leq P_d J_{d+1}.$$

Demonstração. Considere J_d. Como sabemos que ϕ_{k-d-1} não depende da $(k-d-1)$-ésima coordenada, podemos, usando o item 2 da Proposição 0.0.1, expandir a média em duas partes, uma dependente de y_{k-d-1} e outra independente de y_{k-d-1}.

Então, escrevemos

$$J_d = \mathbb{E}(G(y, y')H(y, y') : y_1, y_2, \ldots, y_{k-d-2}, y_{k-d}, \ldots, y_{k-1} \in \mathbb{Z}_N, y' \in \mathbb{Z}_N^d),$$

onde

$$G(y, y') = \prod_{S \subseteq \{k-d,\ldots,k-1\}} f_{k-d-1}(\phi_{k-d-1}(y^{(S)}))\nu^{-1/2}(\phi_{k-d-1}(y^{(S)}))$$

e

$$H(y, y') = \mathbb{E}\left(\prod_{S \subseteq \{k-d,\ldots,k-1\}} \left(\left(\prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)}))\right)\left(\prod_{i=k-d-1}^{k-1} \nu^{1/2}(\phi_i(y^{(S)}))\right)\right) : y_{k-d-1} \in \mathbb{Z}_N\right).$$

Notemos que foi adicionado um fator $\nu^{1/2}(\phi_{k-d-1}(y^{(S)}))$ em H e este mesmo fator foi dividido em G.

Aplicando Cauchy-Schwarz (mais especificamente, a versão apresentada no Teorema 2.1.1), obtemos

$$\begin{aligned} |J_d|^2 \quad \leq \quad & \mathbb{E}(G(y, y')^2 : y_1, y_2, \ldots, y_{k-d-2}, y_{k-d}, \ldots, y_{k-1} \in \mathbb{Z}_N, y' \in \mathbb{Z}_N^d) && \text{(A.1)} \\ & \cdot \mathbb{E}(H(y, y')^2 : y_1, y_2, \ldots, y_{k-d-2}, y_{k-d}, \ldots, y_{k-1} \in \mathbb{Z}_N, y' \in \mathbb{Z}_N^d). && \text{(A.2)} \end{aligned}$$

No primeiro fator, dado que $|f_{k-d-1}(x)| \leq |\nu(x)|$, para todo $x \in \mathbb{Z}_N$, podemos concluir que

$$
\begin{aligned}
G(y,y')^2 &= \left(\prod_{S\subseteq\{k-d,\dots,k-1\}} f_{k-d-1}(\phi_{k-d-1}(y^{(S)}))\nu^{-1/2}(\phi_{k-d-1}(y^{(S)})) \right)^2 \\
&\leq \left(\prod_{S\subseteq\{k-d,\dots,k-1\}} \nu(\phi_{k-d-1}(y^{(S)}))\nu^{-1/2}(\phi_{k-d-1}(y^{(S)})) \right)^2 \\
&= \left(\prod_{S\subseteq\{k-d,\dots,k-1\}} \nu^{1/2}(\phi_{k-d-1}(y^{(S)})) \right)^2 \\
&= \prod_{S\subseteq\{k-d,\dots,k-1\}} \nu(\phi_{k-d-1}(y^{(S)})),
\end{aligned}
$$

que são exatamente os fatores dentro do funcional esperança em P_d.

Agora, pelo caso particular notado no item 2 da Proposição 0.0.1, não há diferença se tomamos a média no primeiro fator do último membro da equação A.1 sobre todas as variáveis ou se omitirmos y_{k-d-1}, já que $G(y,y')$ não depende de y_{k-d-1}. Assim, temos que

$$
\begin{aligned}
&\mathbb{E}(G(y,y')^2 : y_1, y_2, \dots, y_{k-d-2}, y_{k-d}, \dots, y_{k-1} \in \mathbb{Z}_N, y' \in \mathbb{Z}_N^d) \\
&= \mathbb{E}\left(G(y,y')^2 : y \in \mathbb{Z}_N, y' \in \mathbb{Z}_N^d \right) \\
&\leq \mathbb{E}\left(\prod_{S\subseteq\{k-d,\dots,k-1\}} \nu(\phi_{k-d-1}(y^{(S)})) : y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d \right) \\
&= P_d.
\end{aligned}
$$

Consideremos agora o termo A.2. Queremos provar que ele é menor ou igual a J_{d+1}. Com d incrementado por 1, o vetor y' terá uma dimensão a mais, portanto precisaremos de uma variável extra y'_{k-d-1}. Também teremos que tomar produtos sobre mais subconjuntos de S e alterar os índices dos produtórios internos.

Nosso objetivo é mostrar que

$$
\begin{aligned}
&H(y,y')^2 \\
&\leq \mathbb{E}\left(\prod_{S\subseteq\{k-d-1,\dots,k-1\}} \left(\left(\prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)})) \right) \left(\prod_{i=k-d-1}^{k-1} \nu^{1/2}(\phi_i(y^{(S)})) \right) \right) : y_{k-d-1}, y'_{k-d-1} \in \mathbb{Z}_N \right)^2 .
\end{aligned} \tag{A.3}
$$

Fazemos então:

$$
\begin{aligned}
H(y,y')^2 &= \left(\frac{1}{N}\sum_{y_{k-d-1}\in\mathbb{Z}_N}\prod_{S\subseteq\{k-d,\ldots,k-1\}}\prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)}))\prod_{i=k-d-1}^{k-1}\nu^{1/2}(\phi_i(y^{(S)}))\right)^2 \\
&= \frac{1}{N^2}\sum_{y_{k-d-1},y'_{k-d-1}\in\mathbb{Z}_N}\prod_{S\subseteq\{k-d,\ldots,k-1\}}\prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)}))\prod_{i=k-d-1}^{k-1}\nu^{1/2}(\phi_i(y^{(S)})) \\
&\qquad \prod_{S\subseteq\{k-d,\ldots,k-1\}}\prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S\cup\{k-d-1\})}))\prod_{i=k-d-1}^{k-1}\nu^{1/2}(\phi_i(y^{(S\cup\{k-d-1\})}) \\
&= \mathbb{E}\left(\prod_{S\subseteq\{k-d,\ldots,k-1\}}\prod_{i=0}^{k-d-2} f_i(\phi_i(y^{(S)}))\prod_{i=k-d-1}^{k-1}\nu^{1/2}(\phi_i(y^{(S)})) : y_{k-d-1}, y'_{k-d-1}\in\mathbb{Z}_N\right).
\end{aligned}
$$

A segunda e a terceira igualdades são um pouco complicadas.

Considerando a segunda igualdade, como estamos tomando o quadrado da soma sobre $y_{k-d-1} \in \mathbb{Z}_N$, obtemos N^2 pares, representados pelas duas variáveis, denotadas y_{k-d-1} e y'_{k-d-1}, ambas percorrendo $\mathbb{Z}_N$. Então na parte y'_{k-d-1} do par, devemos trocar y_{k-d-1} por y'_{k-d-1}, o que pode ser feito adicionando o elemento $k-d-1$ ao conjunto S.

Consideremos agora a terceira igualdade, e notemos que a expressão pode ser simplificada, já que na primeira componente estamos tomando os produtos sobre todos os $S \subseteq \{k-d,\ldots,k-1\}$, e na segunda componente estamos tomando os produtos sobre os mesmos subconjuntos unidos com $\{k-d-1\}$.

Assim sendo, estamos considerando os subconjuntos $S \subseteq \{k-d-1, k-d, \ldots, k-1\}$ em sua totalidade, e então podemos reescrever a expressão toda na forma descrita pela identidade A.3.

Concluímos, então, que

$$
\begin{aligned}
|J_d|^2 \leq\ & \mathbb{E}(G(y,y')^2 : y_1, y_2, \ldots, y_{k-d-2}, y_{k-d}, \ldots, y_{k-1}\in\mathbb{Z}_N, y'\in\mathbb{Z}_N^d) \\
& \cdot\mathbb{E}(H(y,y')^2 : y_1, y_2, \ldots, y_{k-d-2}, y_{k-d}, \ldots, y_{k-1}\in\mathbb{Z}_N, y'\in\mathbb{Z}_N^d) \\
\leq\ & P_d J_{d+1}
\end{aligned}
$$

□

Corolário A.1.1. *Sejam J_d e P_d como anteriormente definidos.*

Temos, para cada $k \geq 2$,

$$
|J_0|^{2^{k-1}} \leq J_{k-1}\prod_{d=0}^{k-2}|P_d|^{2^{k-2-d}}.
$$

Demonstração. Faremos a prova por indução em k.

Para o caso base $k=2$, temos $|J_0|^2 \leq J_1 P_0$, verdadeiro pelo Lema A.1.1.

Suponhamos agora o resultado válido para k, isto é, que $|J_0|^{2^{k-1}} \leq J_{k-1}\prod_{d=0}^{k-2}|P_d|^{2^{k-2-d}}$ e mostraremos que isso implica diretamente que o mesmo resultado é válido para $k+1$, ou seja, $|J_0|^{2^k} \leq J_k\prod_{d=0}^{k-1}|P_d|^{2^{k-1-d}}$.

Consideremos, para isso, a seguinte cadeia de desigualdades:

$$|J_0|^{2^k} = \left(|J_0|^{2^{k-1}}\right)^2 \leq \left(J_{k-1}\prod_{d=0}^{k-2}|P_d|^{2^{k-2-d}}\right)^2 \leq |J_{k-1}|^2\left(\prod_{d=0}^{k-2}|P_d|^{2^{k-2-d}}\right)^2$$
$$\leq J_k P_{k-1}\prod_{d=0}^{k-2}\left(|P_d|^{2^{k-2-d}}\right)^2 = J_k P_{k-1}\prod_{d=0}^{k-2}|P_d|^{2^{k-1-d}} = J_k\prod_{d=0}^{k-1}|P_d|^{2^{k-1-d}},$$

onde usamos a hipótese de indução na passagem entre a segunda e terceira expressões, e a Desigualdade de Cauchy-Schwarz (Proposição A.1.1) na passagem da quarta para quinta expressão.

E provamos, assim, o resultado desejado. □

Agora entraremos no contexto do Teorema de von Neumann Generalizado mais especificamente.

Portanto, a partir deste ponto, consideraremos ν uma medida k-pseudoaleatória, $c_0, \ldots, c_{k-1}$ uma permutação de k elementos consecutivos de $\{-(k-1), \ldots, -1, 0, 1, \ldots, k-1\}$, e as funções $f_0, f_1, \ldots, f_{k-1} \in L^1(\mathbb{Z}_N)$ satisfazendo

$$|f_j(x)| \leq \nu(x), \text{ para todo } x \in \mathbb{Z}_N, 0 \leq j \leq k-1.$$

Primeiramente, podemos permutar os f_j's e os c_j's, de modo que podemos também assumir que $\inf_{0\leq j\leq k-1}\|f_j\|_{U^{k-1}}$ é alcançado quando $j = 0$, e por fim, transladando x por $c_0 r$, podemos assumir que $c_0 = 0$ ou, mais geralmente, $c_j = j$.

Para usar o lema que provamos anteriormente, precisamos definir certas ϕ_i's de maneira que cada ϕ_i seja independente da i-ésima coordenada, para $0 \leq i \leq k-1$.

Definimos $\phi_0, \phi_1, \ldots, \phi_{k-1} : \mathbb{Z}_N^{k-1} \to \mathbb{Z}_N$ por

$$\phi_i(y) := \sum_{j=1}^{k-1}\left(1 - \frac{c_i}{c_j}\right)y_j,$$

para $i = 0, 1, \ldots, k-1$, onde $y = (y_1, \ldots, y_{k-1})$.

Notemos agora que, para $i \geq 0$,

$$\phi_i(y) = x(y) + c_i r(y),$$

sendo que $x(y) = y_1 + \ldots + y_{k-1}$ e

$$r(y) = -\sum_{j=1}^{k-1}\frac{y_j}{c_j}.$$

Notemos, também, que $\phi_i(y)$ não depende de y_i, já que para $i = 0$ temos que $\phi_0(y) = x(y) = y_1 + \ldots + y_{k-1}$ e, para $1 \leq i \leq k-1$, o termo envolvendo y_i possui coeficiente $(1 - c_i/c_i) = 0$, e portanto estamos numa situação em que podemos usar o lema anterior (Lema A.1.1) e, assim, podemos construir os elementos J_d's e P_d's da maneira previamente definida.

Para provar o resultado principal, precisamos antes de mais quatro lemas:

Lema A.1.2. *A seguinte identidade é verdadeira:*

$$J_0 = \mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x + jr) : x, r \in \mathbb{Z}_N\right).$$

Demonstração. Definamos $\Phi : \mathbb{Z}_N^{k-1} \to \mathbb{Z}_N^2$ por

$$\Phi(y) := (x(y), -r(y)).$$

Queremos agora provar que Φ é uma cobertura uniforme.

Assim, resta-nos provar apenas que $|\Phi^{-1}(z_1, z_2)| = |\mathbb{Z}_N^{k-1}|/|\mathbb{Z}_N^2| = N^{k-3}$, para todos $z_1, z_2 \in \mathbb{Z}_N$.

Isto pode ser feito usando Álgebra Linear, já que $\Phi^{-1}(z_1, z_2)$ é a solução $(t_1, \ldots, t_{k-1})$ do sistema de equações

$$\begin{aligned} z_1 &= t_1 + \ldots + t_{k-1} \\ z_2 &= c_1^{-1} t_1 + \ldots + c_{k-1}^{k-1} t_{k-1}, \end{aligned}$$

cuja matriz de coeficientes assume a forma

$$\begin{pmatrix} 1 & 1 & \ldots & 1 \\ c_1^{-1} & c_2^{-1} & \ldots c_{k-1}^{-1} & \end{pmatrix}$$

Como todos os c_i's são não-nulos e distintos entre si, o sistema possui posto máximo, e daí podemos concluir que a função é sobrejetora e o espaço de soluções terá dimensão $(k-1)-2 = k-3$, e portanto terá N^{k-3} elementos.

Definamos, para $j = 0, \ldots, k-1$, as funções $g_j : \mathbb{Z}_N^2 \to \mathbb{Z}_N$ por $g_j(x, r) := f(x + c_j r)$. Como Φ é uma cobertura uniforme, usando o item 5 da Proposição 0.0.1 temos que

$$\mathbb{E}\left(\prod_{j=0}^{k-1} g_j(x, r) : x, r \in \mathbb{Z}_N\right) = \mathbb{E}\left(\prod_{j=0}^{k-1} g_j(\Phi(y)) : y \in \mathbb{Z}_N^{k-1}\right),$$

e usando as definições dos g_j's e de Φ, obtemos

$$\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x + c_j r) : x, r \in \mathbb{Z}_N\right) = \mathbb{E}\left(\prod_{j=0}^{k-1} f_j(\phi_j(y)) : y \in \mathbb{Z}_N^{k-1}\right).$$

Agora, observando que pela definição de J_d temos

$$J_0 = \mathbb{E}\left(\prod_{j=0}^{k-1} f_j(\phi_j(y)) : y \in \mathbb{Z}_N^{k-1}\right),$$

que combinando com a igualdade anterior, nos garante que

$$J_0 = \mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x + c_j r) : x, r \in \mathbb{Z}_N\right),$$

e portanto concluímos o lema.

□

Lema A.1.3. *A seguinte desigualdade é válida:*

$$|J_0|^{2^{k-1}} \leq (1 + o(1)) J_{k-1}.$$

Demonstração. Recordemos a seguinte definição:

$$P_d = \mathbb{E}\left(\prod_{S \subseteq \{k-d, \ldots, k-1\}} \nu(\phi_{k-d-1}(y^{(S)})) : y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d\right),$$

para $d \leq k-1$.

Por hipótese, ν é k-pseudoaleatória, logo satisfaz a $(k2^k, 3k-4, k)$-condição das formas lineares,

e pela observação feita na página 20, podemos deduzir que também satisfaz a $(2^d, k-1+d, k)$-condição das formas lineares, para todo $d \leq k-1$, já que cada um dos parâmetros é menor que seu correspondente.

Agora, dado $d \geq 0$, consideremos os subconjuntos $S_i \subseteq \{k-d, \ldots, k-1\}$ e definimos os funcionais $\psi_{S_i} : \mathbb{Z}_N^{k-d-1} \to \mathbb{Z}_N$ por

$$\psi_{S_i}(y, y') := \phi_{k-d-1}(y^{(S_i)}),$$

para $S_i \subseteq \{k-d, \ldots, k-1\}$ (lembremos que há 2^d subconjuntos contidos num conjunto com d elementos).

Usando a $(2^d, k-1+d, k)$-condição das formas lineares em cada ψ_i, obtemos

$$\begin{aligned} P_d &= \mathbb{E}\left(\prod_{S_i \subseteq \{k-d,\ldots,k-1\}} \nu(\phi_{k-d-1}(y^{(S_i)})) : y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d \right) \\ &= \mathbb{E}\left(\prod_{S \subseteq \{k-d,\ldots,k-1\}} \nu(\psi_{S_i}(y, y')) : y \in \mathbb{Z}_N^{k-1}, y' \in \mathbb{Z}_N^d \right) = 1 + o(1), \end{aligned}$$

para todo $0 \leq d \leq k-1$. Aplicando o Corolário A.1.1, garantimos que

$$|J_{k-1}| \prod_{d=0}^{k-2} |P_d|^{2^{k-2-d}} = J_{k-1}(1 + o_k(1)),$$

que é o resultado desejado, considerando que é permitido o termo $o(1)$ depender de k. □

Lema A.1.4. *Definamos a função* $W : \mathbb{Z}_N \times \mathbb{Z}_N^{k-1} \to \mathbb{R}$ *por*

$$W(x, h) = \mathbb{E}\left(\prod_{\omega \in \{0,1\}^{k-1}} \prod_{i=1}^{k-1} \nu^{1/2}(\phi_i(\widehat{y_x} + \widehat{\omega h})) : y_1, \ldots, y_{k-2} \in \mathbb{Z}_N \right),$$

onde $\widehat{\omega h}$ *é o vetor tal que* $(\widehat{\omega h})_i := \omega_i h_i$, *e* $\widehat{y_x} = (y_1, \ldots, y_{k-2}, x - y_1 - \ldots - y_{k-2})$.

Então temos que

$$J_{k-1} = \mathbb{E}\left(W(x, h) \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) : x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right).$$

Demonstração. Consideremos a expressão

$$\begin{aligned} &\mathbb{E}\left(W(x, h) \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) : x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right) \\ &= \mathbb{E}\left(\mathbb{E}\left(\prod_{\omega \in \{0,1\}^{k-1}} \prod_{i=1}^{k-1} \nu^{1/2}(\phi_i(\widehat{y_x} + \widehat{\omega h})) : y_1, \ldots, y_{k-2} \in \mathbb{Z}_N \right) \cdot \prod_{\omega \in \{0,1\}^{k-1}} f_0(x + \omega \cdot h) : x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1} \right). \end{aligned}$$

Pelo item 2 na Proposição 0.0.1, sabemos que podemos juntar tudo numa única esperança. Reescrevendo de forma que o produto sobre $\omega \in \{0,1\}^{k-1}$ seja tomado num único produtório,

obtemos

$$\mathbb{E}\left(\prod_{\omega\in\{0,1\}^{k-1}}\left(f_0(x+\omega\cdot h)\prod_{i=1}^{k-1}\nu^{1/2}(\phi_i(\widehat{y_x}+\widehat{\omega h}))\right):x,y_1,\ldots,y_{k-2}\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right),$$

e que, graças à correspondência biunívoca deduzida na Proposição A.1.1, podemos reescrever como

$$\mathbb{E}\left(\prod_{S\subseteq\{1,\ldots,k-1\}}\left(f_0(\phi_0(\widehat{y_x}^{(S)})\prod_{i=1}^{k-1}\nu^{1/2}(\phi_i(\widehat{y_x}^{(S)}))\right):x,y_1,\ldots,y_{k-2}\in\mathbb{Z}_N,y'\in\mathbb{Z}_N^{k-1}\right),$$

que é igual a J_{k-1}. □

E, por fim:

Lema A.1.5.

$$\mathbb{E}\left((W(x,h)-1)\prod_{\omega\in\{0,1\}^{k-1}}f_0(x+\omega\cdot h):x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right)=o(1).$$

Demonstração. Dado que assumimos que $|f_0(x)|\leq\nu(x)$, para todo $x\in\mathbb{Z}_N$, basta provarmos que

$$\mathbb{E}\left((W(x,h)-1)\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h):x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right)=o(1).$$

Se considerarmos o lado esquerdo da expressão ao quadrado, temos, por Cauchy-Schwarz (Lema A.1.1), que seu valor é menor ou igual a

$$\mathbb{E}\left(|W(x,h)-1|^2\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h):x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right)$$
$$\cdot\mathbb{E}\left(\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h):x\in\mathbb{Z}_N,h\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right),$$

e portanto é suficiente provar que esta expressão é $o(1)$. Logo, se provarmos que

$$\mathbb{E}\left(|W(x,h)-1|^2\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h):x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right)=o(1)\tag{A.4}$$

e

$$\mathbb{E}\left(\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h):x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right)=1+o(1),\tag{A.5}$$

completamos a prova, dado que $(1+o(1))\cdot o(1)=o(1)$.

Para isso, consideremos o lado esquerdo da identidade A.4. Expandindo o quadrado, podemos

reescrever a expressão como

$$\mathbb{E}\left(\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h):x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right) \tag{A.6}$$

$$-2\mathbb{E}\left(W(x,h)\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h):x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right) \tag{A.7}$$

$$+\mathbb{E}\left(W(x,h)^2\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h):x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right), \tag{A.8}$$

então é suficiente provar que cada uma das expressões de valor esperado são iguais a $1+o(1)$, já que a expressão inteira seria $(1+o(1))-2(1+o(1))+(1+o(1))=o(1)$. Como o último termo é igual ao lado esquerdo da expressão A.5, é o mesmo que prová-la.

A.6 é da forma $1+o(1)$*:*
Consideremos o item A.6. Dado que ν é k-pseudoaleatória, sabemos que ela satisfaz a $(k2^{k-1},3k-4,k)$-condição das formas lineares, e portanto a $(2^{k-1},k,1)$-condição das formas lineares. Definamos agora 2^{k-1} formas lineares $\mathbb{Z}_N^k\to\mathbb{Z}_N$ (uma para cada $\omega\in\{0,1\}^{k-1}$) por

$$(x,h_1,\ldots,h_{k-1})\mapsto x+\omega\cdot h,$$

onde $h=(h_1,\ldots,h_{k-1})$.

Então, a condição das formas lineares nos garante que o termo em A.6 é $1+o(1)$, como desejado. □

A.7 é da forma $-2(1+o(1))$*:* Para realizarmos esta dedução, mostraremos o resultado equivalente de que o termo da esperança é igual a $1+o(1)$.

Primeiramente, notemos que $W(x,h)$ pode ser escrito de outra maneira, explicitamente

$$W(x,h)=\mathbb{E}\left(\prod_{i=1}^{k-1}\prod_{\omega\in\{0,1\}^{k-1},\omega_i=0}\nu(\phi_i(\widehat{y_x}+\widehat{\omega h})):y_1,\ldots,y_{k-2}\in\mathbb{Z}_N\right), \tag{A.9}$$

porque se trocarmos a ordem dos produtórios e então relembrarmos que cada ϕ_i é independente na i-ésima variável, então é indiferente o fato de $\omega_i=0$ ou $\omega_i=1$, logo obtemos o mesmo resultado se calcularmos sobre os ω com $\omega_i=0$ e em seguida tomarmos o quadrado em cada um dos fatores $\nu(\phi_i(\widehat{y_x}+\widehat{\omega h}))^{1/2}$.

Mais uma vez, usando o fato de que ν é k-pseudoaleatória, então em particular satisfaz a $(2^{k-2}(k+1),2k-2,k)$-condição das formas lineares. Queremos definir $2^{k-1}(k+1)$ formas lineares para usar a condição, e a usaremos sobre as variáveis $x,h_1,\ldots,h_{k-1},y_1,\ldots,y_{k-2}$.

Definamos as 2^{k-1} primeiras formas lineares (uma para cada $\omega\in\{0,1\}^{k-1}$) por

$$(x,h_1,\ldots,h_{k-1},y_1,\ldots,y_{k-2})\mapsto x+\omega\cdot h,$$

e definamos $(k-1)2^{k-2}$ formas lineares (uma para cada $1\le i\le k-1$ e cada $\omega\in\{0,1\}^{k-1}$ com $\omega_i=0$) por

$$(x,h_1,\ldots,h_{k-1},y_1,\ldots,y_{k-2})\mapsto\phi_i(\widehat{y_x}+\widehat{\omega h}),$$

onde $h=(h_1,\ldots,h_{k-1})$, onde os ϕ_i's são definidos como no começo da prova e $\widehat{y_x}=(y_1,\ldots,y_{k-1})$, onde $y_{k-1}=x-y_1-\ldots-y_{k-2}$. Note que são formas lineares, já que cada ϕ_i's são formas lineares. Isto nos dá $2^{k-1}+(k-1)2^{k-2}=(k+1)2^{k-1}$ formas lineares no total, então resta-nos somente verificar que estas são realmente as formas lineares que procuramos.

Usando a identidade A.9 e o ponto 2 da Proposição 0.0.1, obtemos que

$$\mathbb{E}\left(W(x,h)\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h):x\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right)=$$

$$\mathbb{E}\left(\prod_{i=1}^{k-1}\prod_{\omega\in\{0,1\}^{k-1},\omega_i=0}\nu(\phi_i(\widehat{y_x}+\widehat{\omega h}))\prod_{\omega\in\{0,1\}^{k-1}}\nu(x+\omega\cdot h):x,y_1,\ldots,y_{k-2}\in\mathbb{Z}_N,h\in\mathbb{Z}_N^{k-1}\right),$$

logo com as formas lineares escolhidas e usando as condições das formas lineares, vemos que é valorada em $1+o(1)$. □

A.8 é da forma $1+o(1)$*:* Por fim, consideremos o termo A.8. Agora necessitaremos de toda a força dos parâmetros da condição das formas lineares, especificamente a $(k2^{k-1},3k-4,k)$-condição.

Definamos $(k-1)2^{k-2}$ formas lineares (para cada $1\leq i\leq k-1$ e para cada $\omega\in\{0,1\}^{k-1}$, com $\omega_i=0$) por

$$(x,h_1,\ldots,h_{k-1},y_1,\ldots,y_{k-2},y_1',\ldots,y_{k-2}')\mapsto\phi(\widehat{y_x}+\widehat{\omega h}),$$

sendo $h=(h_1,\ldots,h_{k-1})$, os ϕ_i's são definidos como anteriormente, e $\widehat{y_x}=(y_1,\ldots,y_{k-1})$, de modo que $y_{k-1}=x-y_1-\ldots-y_{k-2}$.

De modo semelhante, definamos outras $(k-1)2^{k-2}$ formas lineares por

$$(x,h_1,\ldots,h_{k-1},y_1,\ldots,y_{k-2},y_1',\ldots,y_{k-2}')\mapsto\phi(\widehat{y_x}'+\widehat{\omega h}),$$

onde $\widehat{y_x}'=(y_1',\ldots,y_{k-1}')$, de modo que $\widehat{y_x}'_{k-1}=x-y_1'-\ldots-y_{k-2}'$.

Por fim, definamos 2^{k-1} formas lineares (uma para cada $\omega\in\{0,1\}^{k-1}$) da seguinte maneira:

$$(x,h_1,\ldots,h_{k-1},y_1,\ldots,y_{k-2},y_1',\ldots,y_{k-2}')\mapsto\widehat{x}+\widehat{\omega h}.$$

Portanto, temos $(k-1)2^{k-2}+(k-1)2^{k-2}+2^{k-1}=k2^{k-1}$ formas lineares, que são exatamente as que necessitamos.

Agora, analisemos o termo $W(x,h)^2$. Pela expressão A.9 e o resultado 4 de 0.0.1, obtemos que

$$W(x,h)^2=\mathbb{E}\left(\prod_{i=1}^{k-1}\prod_{\substack{\omega\in\{0,1\}^{k-1}\\ \omega_i=0}}\nu(\phi_i(\widehat{y_x}+\widehat{\omega h}))\prod_{\substack{\omega'\in\{0,1\}^{k-1}\\ \omega_i'=0}}\nu(\phi_i(\widehat{y_x}'+\omega' h))\right),$$

com $y_1,\ldots,y_{k-1},y_1',\ldots,y_{k-1}'\in\mathbb{Z}_N$.

Logo, usando o 2 de 0.0.1, temos que podemos colocar tudo sobre um único funcional esperança, podemos deduzir que o termo A.9 é o valor esperado sobre o produto de todas as formas lineares que definimos previamente, e daí, usando a condição das formas lineares, concluímos que A.9 é da forma $1+o(1)$. □

Agora que cada uma das expressões está devidamente estimada, podemos finalmente concluir a prova de que $A.6+A.7+A.8=(1+o(1))+(-2\cdot(1+o(1)))+(1+o(1))=o(1)$, como desejávamos. □

Agora que provamos estes lemas, faremos o enunciado de uma versão um pouco mais forte do Teorema de von Neumann do que a utilizada na Subseção 4.3.4, da qual a conclusão daquela é imediata, e mostraremos como os lemas anteriores implicam em sua validade.

A.2 Enunciado e Prova

Teorema A.2.1 (von Neumann Generalizado)**.** *Consideremos ν uma medida k-pseudoaleatória, $c_0, \dots, c_{k-1}$ uma permutação de k elementos consecutivos de $\{-(k-1), \dots, -1, 0, 1, \dots, k-1\}$, e $f_0, f_1, \dots, f_{k-1} \in L^1(\mathbb{Z}_N)$ satisfazendo*

$$|f_j(x)| \leq \nu(x), \text{ para todo } x \in \mathbb{Z}_N, 0 \leq j \leq k-1.$$

Então temos que

$$\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x + c_j r) : x, r \in \mathbb{Z}_N\right) = O\left(\inf_{0\leq j\leq k-1} \|f_j\|_{U^{k-1}}\right) + o(1).$$

Demostração do Teorema de von Neumann Generalizado (Teorema A.2.1). Pelo Lema A.1.2, temos que o lado esquerdo da igualdade é igual a J_0, o que pelo Lema A.1.3 nos garante que

$$\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x + c_j r) : x, r \in \mathbb{Z}_N\right)^{2^{k-1}} = J_0^{2^{k-1}} \leq J_{k-1} + o(1). \tag{A.10}$$

Agora, pela definição da U^{k-1}-norma, sabemos que

$$\|f_0\|_{U^{k-1}}^{2^{k-1}} = \mathbb{E}\left(\prod_{\omega\in\{0,1\}^{k-1}} f_0(x + \omega \cdot h) : x, h \in \mathbb{Z}_N\right),$$

e daí, usando os Lemas A.1.4 e A.1.5, temos que

$$\begin{aligned} o(1) &= \mathbb{E}\left((W(x,h) - 1) \prod_{\omega\in\{0,1\}^{k-1}} f_0(x + \omega \cdot h) : x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\right) \\ &= \mathbb{E}\left((W(x,h)) \prod_{\omega\in\{0,1\}^{k-1}} f_0(x + \omega \cdot h) : x \in \mathbb{Z}_N, h \in \mathbb{Z}_N^{k-1}\right) - \|f_0\|_{U^{k-1}}^{2^{k-1}} \\ &= J_{k-1} - \|f_0\|_{U^{k-1}}^{2^{k-1}}. \end{aligned}$$

Aplicando isto na desigualdade A.10, temos que

$$\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x + c_j r) : x, r \in \mathbb{Z}_N\right)^{2^{k-1}} \leq \|f_0\|_{U^{k-1}}^{2^{k-1}} + o(1), \tag{A.11}$$

mas relembremos que no começo da prova permutamos os c_j's e os f_j's, de modo que

$$\|f_0\|_{U^{k-1}} = \inf_{0\leq j\leq k-1} \|f_j\|_{U^{k-1}},$$

logo, ao tomar a 2^{k-1}-ésima raiz de ambos os lados de A.11, obtemos o resultado que desejamos, dado que $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, e portanto

$$\left|\mathbb{E}\left(\prod_{j=0}^{k-1} f_j(x + c_j r) : x, r \in \mathbb{Z}_N\right)\right| \leq \left(\|f_0\|_{U^{k-1}}^{2^{k-1}} + o(1)\right)^{1/2^{k-1}} \leq O(\|f_0\|_{U^{k-1}}) + o(1),$$

e, assim, terminamos a prova. □

Apêndice B

ν é k-Pseudoaleatória

Neste apêndice, verificaremos que a medida ν definida na Seção 4.4 é, de fato, uma medida pseudoaleatória.

ν é uma medida

Para provar que ν é uma medida, precisaremos da seguinte proposição:

Proposição B.0.1. *Sejam $m, t \in \mathbb{N}$, e para cada $1 \leq i \leq m$, definamos as formas lineares $\psi_i : \mathbb{R}^t \to \mathbb{R}$ por*

$$\psi_i(x_1, \ldots, x_t) := \sum_{j=1}^{t} L_{ij} x_j + b_i,$$

onde $(L_{ij})_{m \times t}$ é uma matriz com valores inteiros, de modo que não possua duas linhas que sejam múltiplas racionais entre si, e tal que satisfaça a desigualdade $|L_{ij}| \leq \sqrt{w(N)}/2$, para todo $i = 1, \ldots, m$ e $j = 1, \ldots, t$. Daí, definamos $\theta_i = W\psi_i + 1$, para $i = 1, \ldots, m$.

Assumamos que $B \subseteq \mathbb{R}^t$ é um produto de intervalos $I_j \subseteq \mathbb{R}$ de modo que cada intervalo tenha comprimento no mínimo R^{10m}.

Então vale que

$$\mathbb{E}(\Lambda_R(\theta_1(x))^2 \cdots \Lambda_R(\theta_m(x))^2) : x \in B) = (1 + o_{m,t}(1)) \left(\frac{W \log R}{\phi(W)} \right)^m.$$

Uma prova para esta proposição pode ser encontrada no apêndice do artigo original de Green e Tao, [GT04].

Agora estamos em condições de provar o seguinte resultado:

Teorema B.0.1. *A função ν definida na Seção 4.4 satisfaz $\mathbb{E}(\nu(x) : x \in \mathbb{Z}_N) = 1 + o(1)$.*

Demonstração. Aplicaremos a proposição anterior com $m = t = 1$ e $B = [\varepsilon N, 2\varepsilon N]$, para N suficientemente grande. Temos, portanto, que

$$\mathbb{E}(\Lambda_R(\theta_1(x))^2 : x \in [\varepsilon N, 2\varepsilon N]) = (1 + o(1)) \left(\frac{W \log R}{\phi(W)} \right),$$

e recordando a definição de ν, temos que isto é equivalente a

$$\mathbb{E}(\nu(x) : x \in [\varepsilon N, 2\varepsilon N]) = 1 + o(1).$$

Mas como $\nu(x) = 1$, para $x \notin [\varepsilon N, 2\varepsilon N]$, temos que

$$\mathbb{E}(\nu(x) : x \notin [\varepsilon N, 2\varepsilon N]) = 1,$$

e combinando ambas, segue que

$$\mathbb{E}(\nu(x) : x \in \mathbb{Z}_N) = 1 + o(1),$$

que é o resultado desejado. □

ν satisfaz a condição das formas lineares

Demonstração. Queremos provar que ν satisfaz a $(k2^{k-1}, 3k-4, k)$-condição das formas lineares.

Consideremos então as formas lineares

$$\psi_i(x_1, \ldots, x_t) = \sum_{j=1}^{t} L_{ij}x_j + b_i,$$

para $i = 1, \ldots, m$, com $m \leq k2^{k-1}$, $t \leq 3k-4$ e tais que todos os coeficientes L_{ij} são números racionais com peso no máximo k. Além disso, nenhuma linha da $(m \times t)$-matriz (L_{ij}) é múltipla racional de alguma outra.

Pela definição da condição das formas lineares, precisamos provar que

$$\mathbb{E}(\nu(\psi_1(x)) \cdots \nu(\psi_m(x)) : x \in \mathbb{Z}_N^t) = 1 + o(1). \tag{B.1}$$

Os denominadores de L_{ij} são menores do que k em valor absoluto, portanto se multiplicarmos todos os coeficientes L_{ij} por $k!$, obtemos o limitante $|L_{ij}| \leq k \cdot k! \leq (k+1)!$, e tomando N suficientemente grande garantimos que

$$|L_{ij}| \leq (k+1)! < \frac{\sqrt{w(N)}}{2},$$

satisfaz as hipóteses da proposição demonstrada no tópico anterior.

Tomemos agora $Q = Q(N)$ uma função com $Q(N) \to \infty$ quando $N \to \infty$, e que $Q(N) < N$. Seja também $(u_1, \ldots, u_t) \in \mathbb{Z}_Q^t$, e definamos

$$B_{u_1,\ldots,u_t} = \left\{(x_1, \ldots, x_t) \in \mathbb{Z}_N^t : \left\lfloor u_j \frac{N}{Q} \right\rfloor \leq x_j < \left\lfloor (u_j+1)\frac{N}{Q} \right\rfloor, j = 1, 2, \ldots, t\right\}.$$

Chamaremos cada um desses elementos de *caixa*.

Precisaremos agora do seguinte lema, que nos diz que se quisermos tomar a média em um conjunto, podemos particioná-lo em subconjuntos de tamanhos quase iguais, tomar a média em cada um destes subconjuntos, e então tomar a média sobre estes valores, e obtemos o resultado desejado com um fator multiplicativo de erro da forma $1 + o(1)$.

Lema B.0.1. *Seja $f : \mathbb{Z}_N^n \to \mathbb{R}$ e seja $\left\{B_i^N\right\}_{i\in I}$ uma partição de $\mathbb{Z}_N^n$ tal que $|B_i^N| \to \infty$, para $N \to \infty$, para todo $i \in I$, e*

$$|B_i^N| - |B_j^N| = O(1),$$

para todo $i, j \in I$.

Então temos que

$$\mathbb{E}(\mathbb{E}(f(x) : x \in B_i) : i \in I) = (1 + o(1))\mathbb{E}(f). \tag{B.2}$$

Demonstração. Escreveremos $B_i = B_i^N$. Seja $b = \mathbb{E}(|B_i| : i \in I)$. Então $|B_i| = b + O(1)$, para todo $i \in I$, e

$$\left|\frac{1}{B_i} - \frac{1}{b}\right| = \frac{O(1)}{b(b + O(1))} = \frac{1}{b}o(1),$$

porque $b \to \infty$ quando $N \to \infty$, e daí

$$\frac{1}{B_i} = (1 + o(1))\frac{1}{b}.$$

Usando isto no lado esquerdo da igualdade do lema, obtemos que

$$\frac{1}{|I|}\sum_{i\in I}\frac{1}{|B_i|}\sum_{x\in B_i} f(x) = \frac{1}{|I|}\sum_{i\in I}\frac{1}{b}(1+o(1))\sum_{x\in B_i} f(x) = (1+o(1))\mathbb{E}(f).$$

□

Notemos agora que $\{B_{u_1,\dots,u_t} : u_1,\dots,u_t \in \mathbb{Z}_Q\}$ é uma partição de $\mathbb{Z}_N^t$, e portanto este lema nos garante que

$$\mathbb{E}(\mathbb{E}(\nu(\psi_1(x))\nu(\psi_2(x))\dots\nu(\psi_t(x)) : x \in B_{u_1,\dots,u_t}) : u_1,\dots,u_t \in \mathbb{Z}_Q) \tag{B.3}$$

é igual ao lado direito da identidade B.1 vezes um fator multiplicativo de $(1+o(1))$, dado que duas caixas diferem por no máximo 2 em tamanho, e como podemos escolher Q de maneira que $Q \to \infty$ quando $N \to \infty$ suficientemente lento, tal que o tamanho das caixas também tenda a ∞ conforme $N \to \infty$.

Portanto, resta provarmos que a expressão acima é $1 + o_{m,t}(1)$. Para isso, precisaremos da seguinte definição e do seguinte lema:

Definição (t-úplas comportadas)**.** Uma t-úpla $(u_1,\dots,u_t) \in \mathbb{Z}_Q^t$ é dita *comportada* se, para todo $1 \le i \le m$, ocorre que

$$\psi_i(B_{u_1,\dots,u_t}) \subseteq [\varepsilon N, 2\varepsilon N] \text{ ou } \psi_i(B_{u_1,\dots,u_t}) \cap [\varepsilon N, 2\varepsilon N] = \emptyset.$$

Lema B.0.2. *Se $(u_1,\dots,u_t) \in \mathbb{Z}_Q^t$ é comportada, temos que*

$$\mathbb{E}(\nu(\psi_1(x))\nu(\psi_2(x))\dots\nu(\psi_t(x)) : x \in B_{u_1,\dots,u_t}) = 1 + o_{m,t}(1).$$

Se $(u_1,\dots,u_t) \in \mathbb{Z}_Q^t$ não é comportada, temos que

$$\mathbb{E}(\nu(\psi_1(x))\nu(\psi_2(x))\dots\nu(\psi_t(x)) : x \in B_{u_1,\dots,u_t}) = O_{m,t}(1) + o_{m,t}(1).$$

Por fim, a proporção de t-úplas não-comportadas é de $O_{m,t}(1/Q)$.

Demonstração. Suponha que $(u_1,\dots,u_t)$ é comportada. Pela relação B.2, temos que

$$\mathbb{E}(\nu(\psi_1(x))\cdots\nu(\psi_t(x)) : x \in B_{u_1,\dots,u_t}) = 1 + o_{m,t}(1),$$

porque podemos substituir cada $\nu(\psi_i(x))$ por ou 1 ou $\frac{\phi(W)}{W\log(R)}\Lambda_R(\omega_i(x))^2$, e porque se Q cresce lento o suficiente em N, então $N/Q \ge R^{10m}$ para N suficientemente grande, e portanto as condições da identidade B.2 são satisfeitas.

Assuma agora que $(u_1,\dots,u_t)$ não é comportada. Usamos então o limitante

$$\|\nu\|_\infty \le 1 + \frac{\phi(W)}{W\log(R)}\Lambda_R(\omega_i(x))^2$$

para obter

$$\mathbb{E}(\nu(\psi_1(x))\cdots\nu(\psi_t(x)) : x \in B_{u_1,\dots,u_t}) \le \sum_{A\subseteq\{1,\dots,m\}} \mathbb{E}\left(\prod_{i\in A}\frac{\phi(W)}{W\log(R)}\Lambda_R(\omega_i(x))^2 : x \in B_{u_1,\dots,u_t}\right).$$

Usando a identidade B.2 novamente, obtemos

$$\mathbb{E}\left(\prod_{i\in A}\frac{\phi(W)}{W\log(R)}\Lambda_R(\omega_i(x))^2 : x\in B_{u_1,\dots,u_t}\right) = 1+o_{m,t}(1),$$

e portanto a soma sobre todos os subconjuntos $A\subseteq\{1,\dots,m\}$ resulta em

$$\mathbb{E}(\nu(\psi_1(x))\cdots\nu(\psi_t(x)) : x\in B_{u_1,\dots,u_t}) = O_{m,t}(1)+o_{m,t}(1),$$

já que há $O_m(1)$ de tais A's.

Resta-nos somente provar que a proporção de t-úplas não-comportada é da forma $O_{m,t}(1/Q)$.

Suponha que $(u_1,\dots,u_t)$ é não-comportada. Então um dos $\psi_i(B_{u_1,\dots,u_t})$ possui intersecção não-vazia com o intervalo $[\varepsilon N, 2\varepsilon N]$, mas não está completamente contido no intervalo, logo existem $x,y\in B_{u_1,\dots,u_t}$ tais que

$$\psi(x)\in[\varepsilon N,2\varepsilon N] \text{ e } \psi(y)\in[\varepsilon N,2\varepsilon N].$$

Mas como $L_{ij}=O(1)$, ambos $\psi(x)$ e $\psi(y)$ podem ser estimados por $\psi(u_1,\dots,u_t)$, e daí conseguimos

$$\psi(x)=\sum_{j=1}^{t}L_{ij}\left\lfloor u_j\frac{N}{Q}\right\rfloor+b_i+O_{m,t}(N/Q)$$

e

$$\psi(x)=\sum_{j=1}^{t}L_{ij}\left\lfloor u_j\frac{N}{Q}\right\rfloor+b_i+O_{m,t}(N/Q),$$

onde o termo de erro é de magnitude N/Q, já que este é o tamanho da caixa e portanto é a distância maximal que a i-ésima coordenada x pode possuir de $\left\lfloor u_j\frac{N}{Q}\right\rfloor$, e a dependência em m e t ser em razão do tamanho da soma.

Se $\psi(y)<\varepsilon N$, então temos que $\psi(y)<\varepsilon N<\psi(x)$, e portanto

$$\varepsilon N=\sum_{j=1}^{t}L_{ij}\left\lfloor u_j\frac{N}{Q}\right\rfloor+b_i+O_{m,t}(N/Q),$$

pelas expressões para $\psi(x)$ e $\psi(y)$. Se $\psi(y)>2\varepsilon N$, obtemos uma expressão similar, com $2\varepsilon N$ no lado esquerdo da expressão. Em resumo, temos que

$$a\varepsilon N=\sum_{j=1}^{t}L_{ij}\left\lfloor u_j\frac{N}{Q}\right\rfloor+b_i+O_{m,t}(N/Q),$$

para $a=1$ ou $a=2$. Dividindo por N/Q, obtemos

$$a\varepsilon Q=\sum_{j=1}^{t}L_{ij}+\frac{b_iQ}{N}+O_{m,t}(1).$$

Nenhuma das t-úplas $(L_{ij})_{j=1}^{t}$ é nula, e portanto as t-úplas do tipo $(u_1,\dots,u_t)$ que satisfazem a equação são, no máximo, da ordem de $O_{m,t}(Q^{t-1})$, ou ainda $O_{m,t}(1/Q)$ das possíveis t-úplas. □

Voltemos agora para a prova de que a expressão B.3 é igual a $1+o_{m,t}(1)$.

Temos que

$$(1-O_{m,t}(1/Q))(1+o_{m,t}(1))+O_{m,t}(1/Q)(O_{m,t}(1)+o_{m,t}(1))=1+o_{m,t}(1),$$

porque escolhemos Q de maneira que $Q(N)\to\infty$, temos que $O_{m,t}(1/Q)=o_{m,t}(1)$, e pelo lema

podemos deduzir que

$$\mathbb{E}(\mathbb{E}(\nu(\psi_1(x))\cdots\nu(\psi_t(x)) : x \in B_{u_1,\ldots,u_t}) : u_1,\ldots,u_t \in \mathbb{Z}_Q) = 1 + o_{m,t}(1),$$

e portanto obtemos o resultado desejado. □

ν satisfaz a condição de correlação

Demonstração. Provaremos que ν satisfaz a 2^{k-1}-condição de correlação, e para isto usaremos dois lemas.

Lema B.0.3. *Seja $m \geq 1$ um inteiro e B um intervalo de comprimento $\geq R^{10m}$. Além disso, sejam $h_1,\ldots,h_m \in \mathbb{Z}$ distintos e tais que $|h_i| \leq N^2$, para todo i. Por fim, defina*

$$\Delta = \prod_{1\leq i<j\leq m} |h_i - h_j|$$

.

Então, para N suficientemente grande, temos que

$$\begin{aligned}&\mathbb{E}(\Lambda_R(W(x+h_1)+1)^2\cdots\Lambda_R(W(x+h_m)+1)^2 : x \in B)\\ &\leq (1+o_m(1))\left(\frac{W\log(R)}{\phi(W)}\right)^m \prod_{p|\Delta}(1+O_m(p^{-1/2})),\end{aligned}$$

onde o produto é sobre todos os primos p que dividem Δ.

A prova deste primeiro lema pode ser encontrada no apêndice do artigo original de Green e Tao, [GT04].

Lema B.0.4. *Seja $m \leq 1$ um inteiro. Existe uma função $\tau_m : \mathbb{Z} \to \mathbb{R}^+$ tal que $\tau_m(n) \geq 1$, para todo $n \neq 0$, e tal que*

$$\prod_{p|\Delta}(1+O_m(p^{-1/2})) \leq \sum_{1\leq i<j\leq m} \tau_m(h_i - h_j),$$

para todos os $h_1,\ldots,h_m \in [\varepsilon N, 2\varepsilon N]$ distintos, onde Δ é como no lema anterior.

Além disso, temos que

$$\mathbb{E}(\tau_m(n)^q : 0 < |n| \leq N) = O_{m,q}(1),$$

para todo $q \geq 1$.

Demonstração. Pela definição de Δ, podemos facilmente concluir que

$$\prod_{p|\Delta}(1+O_m(p^{-1/2})) \leq \prod_{1\leq i<j\leq m} \prod_{p\left|\right.|h_i-h_j|} (1+O_m(p^{-1/2})),$$

onde a desigualdade ocorre porque um mesmo primo pode ocorrer várias vezes no lado direito e, portanto, possui pelo menos os fatores que aparecem no lado esquerdo da desigualdade.

Temos também em nosso poder o seguinte limitante:

$$\prod_{p\left|\right.|h_i-h_j|} (1+O_m(p^{-1/2})) \leq \prod_{p\left|\right.|h_i-h_j|} \left((1+p^{-1/2})\right)^{O_m(1)}.$$

Agora, para todo $n \in \mathbb{Z}$, definamos $\tau_m : \mathbb{Z} \to \mathbb{R}^+$ da seguinte maneira:

$$\tau_m(n) = O_m(1)\cdot\prod_{p|n}\left((1+p^{-1/2})\right)^{O_m(1)}.$$

Obtemos, então, a seguinte cadeia de desigualdades:

$$\begin{aligned}
&\prod_{p|\Delta}(1+O_m(p^{-1/2}))\\
&\leq \prod_{1\leq i<i\leq m}\prod_{p\left||h_i-h_j|\right.}(1+O_m(p^{-1/2}))\\
&\leq \prod_{1\leq i<i\leq m}\prod_{p\left||h_i-h_j|\right.}(1+p^{-1/2})^{O_m(1)}\\
&\leq \prod_{1\leq i<i\leq m}\tau_m(h_i-h_j)\\
&\leq \prod_{1\leq i<i\leq m}\tau_m(h_i-h_j)^{\binom{m}{2}} \leq \frac{1}{\binom{m}{2}}\sum_{1\leq i<i\leq m}\tau_m(h_i-h_j),
\end{aligned}$$

onde a última desigualdade é válida pela desigualdade entre a média geométrica e a média aritmética. Isso prova a primeira parte do lema.

Para a segunda afirmação precisamos mostrar que

$$\mathbb{E}\left(\prod_{p|n}(1+p^{-1/2})^{O_m(q)} : 0<|n|\leq N\right) = O_{m,q}(1),$$

para todo $q \geq 1$.

Para todos os primos suficientemente grandes (de fato, para todos os primos exceto os da forma $O_{m,q}(1)$), temos que

$$(1+p^{-1/2})^{O_m(q)} \leq 1+p^{-1/4},$$

e portanto

$$\mathbb{E}\left(\prod_{p|n}(1+p^{-1/2})^{O_m(q)} : 0<|n|\leq N\right) \leq O_{m,q}(1)\cdot\mathbb{E}\left(\prod_{p|n}(1+p^{-1/4}) : 0<n\leq N\right).$$

Porém, realizando as multiplicações, obtemos a desigualdade

$$\prod_{p|n}(1+p^{-1/4}) \leq \sum_{d|n} d^{-1/4},$$

onde no lado direito incluímos somente os divisores positivos. Note que não é uma igualdade pois o lado direito pode incluir divisores que são divisíveis por potências de primos.

Usando este resultado, deduzimos que

$$\mathbb{E}\left(\prod_{p|n}(1+p^{-1/2})^{O_m(q)} : 0<|n|\leq N\right) \leq O_{m,q}(1)\frac{1}{2N}\sum_{1\leq|n|\leq N}\sum_{d|n} d^{-1/4}.$$

Vemos que $d^{-1/4}$ aparece na soma dupla um número $2N/d$ de vezes, para cada $1 \leq d \leq N$, porque aparece N/d vezes para cada n positivo e N/d vezes para cada n negativo, logo a expressão pode ser resumida em

$$O_{m,q}(1)\frac{1}{2N}\sum_{1\leq|n|\leq N}\sum_{d|n} d^{-1/4} = O_{m,q}(1)\sum_{d=1}^{N}\frac{N}{d}d^{-1/4} = O_{m,q}(1),$$

donde concluímos a prova do lema. □

Relembrando que para provarmos que ν satisfaz a 2^{k-1}-condição de correlação, precisamos provar que para qualquer $1 \leq m \leq 2^{k-1}$, existe uma função $\tau = \tau_m : \mathbb{Z}_N \to \mathbb{R}^+$ tal que $\mathbb{E}(\tau(x)^q) = O_{m,q}(1)$, para todo $q \geq 1$, e que também satisfaça

$$\mathbb{E}(\nu(x+h_1)\cdots\nu(x+h_m) : x \in \mathbb{Z}_N) \leq \sum_{1\leq i<j\leq m} \tau(h_i - h_j),$$

para todos $h_1, \ldots, h_m$.

Fixemos m e definamos $\tau : \mathbb{Z}_N \to \mathbb{R}^+$ por

$$\tau = \tau_m(n) = \begin{cases} n - \lfloor N/2 \rfloor, & \text{se } n \neq 0; \\ \exp(Cm\log(N)/\log(\log(N))), & \text{se } n = 0. \end{cases}$$

onde C é uma constante que será determinada posteriormente, e consideramos $n - \lfloor N/2 \rfloor \in \mathbb{Z}_N$ da maneira óbvia, identificando $\mathbb{Z}_N$ com os inteiros entre $-N/2$ e $N/2$. Pelo Lema B.0.4, podemos ver que

$$\mathbb{E}(\tau(x)^q : x \neq 0) = O_{m,q}(1),$$

e dado que

$$\frac{\exp(Cm\log(N)/\log\log N)}{N} = N^{Cm/\log\log N - 1} = o_{m,q}(1),$$

o caso $x = 0$ contribui apenas com $o_{m,q}(1)$, e portanto temos que $\mathbb{E}(\tau(x)^q) = O_{m,q}(1)$.

Vamos considerar o caso onde ao menos dois dos h_i's são iguais. É suficiente mostrar que

$$\mathbb{E}(\nu(x+h_1)\cdots\nu(x+h_m) : x \in \mathbb{Z}_N) \leq \exp(Cm\log N/\log\log N),$$

pela definição de $\tau(0)$. Temos que

$$\mathbb{E}(\nu(x+h_1)\cdots\nu(x+h_m) : x \in \mathbb{Z}_N) \leq \|\nu\|_\infty^m.$$

Precisamos provar que

$$\|\nu\|_\infty \leq \exp(C\log N/\log\log N),$$

e pela definição de ν só temos

$$\|\nu\|_\infty \leq O(\Lambda_R(Wx+1)^2/\log R),$$

já que $\phi(n)/n \leq 1$, para todo n. Agora, como $\log R = O(\log N)$ e

$$|\Lambda_R(Wx+1)| \leq d(Wx+1)\log R,$$

onde $d(\cdot)$ é a função que conta a quantidade de divisores positivos de um inteiro. Usando o que sabemos sobre a ordem máxima de $d(\cdot)$ (Proposição 2.4.2) (e tomando e ao invés de 2 como base da função exponencial), temos que

$$d(Wx+1) \leq \exp(C'\log(Wx+1)/\log\log(Wx+1)),$$

para alguma constante C', e dado que $Wx+1 = O(N\log N)$, isto é menos do que

$$\exp(C''\log(N\log N)/\log\log(N\log N)) \leq \exp(C\log N/\log\log N),$$

para certas constantes C'' e C.

Agora suponhamos que todos os h_is são distintos entre si. Definamos $g : \mathbb{Z}_N \to \mathbb{R}$ por

$$g(x) = \frac{\phi(W)}{W}\frac{\Lambda_R(Wx+1)^2}{\log R}\chi_{[\varepsilon N, 2\varepsilon N]}(x).$$

Então temos que $\nu(x) \leq 1 + (g(x))$, para todo $x \in \mathbb{Z}_N$, pela definição de ν, e portanto

$$\mathbb{E}(\nu(x+h_1)\cdots\nu(x+h_m) : x \in \mathbb{Z}_N) \leq \mathbb{E}((1+g(x+h_1))\cdots(1+g(x+h_m)) : x \in \mathbb{Z}_N).$$

O lado direito da desigualdade pode ser escrito como

$$\sum_{A\subseteq\{1,\ldots,m\}} \mathbb{E}\left(\prod_{i\in A} g(x+h_i) : x \in \mathbb{Z}_N\right).$$

Notemos agora que se $i, j \in A$ com $|h_i - h_j| > \varepsilon N$, então ou $g(x+h_i) = 0$ ou $g(x+h_j) = 0$, portanto para a expressão acima, podemos assumir que $|h_i - h_j| \leq \varepsilon N$, para todos $i, j \in A$. Pelo Lema B.0.3, onde talvez tenhamos que aumentar o valor de N de maneira que $N \geq R^{10m}$, temos que

$$\mathbb{E}\left(\prod_{i\in A} g(x+h_i) : x \in \mathbb{Z}_N\right) \leq (1 + o_{m,q}(1))\prod_{p|\Delta}(1 + O_m(p^{-1/2})),$$

e como $|h_i - h_j| \leq \varepsilon N$, obtemos

$$o_m(1)\prod_{p|\Delta}\left(1 + O_m(p^{-1/2})\right) = o_m(1).$$

Usando o Lema B.0.4, temos que

$$\mathbb{E}\left(\prod_{i\in A} g(x+h_i) : x \in \mathbb{Z}_N\right) \leq \sum_{1\leq i<j\leq m} \tau(h_i - h_j) + o_m(1).$$

Por fim, queremos somar sobre todo o conjunto A, mas para conseguirmos o limitante desejado, precisamos "escalar" τ por um fator limitado, no entanto este fator depende de m, o que é permitido graças à construção de τ na prova do lema B.0.4 como

$$O_m(1)\prod_{p|n}(1 + p^{-1/2})^{O_m(1)}.$$

Isto conclui a prova de que ν satisfaz a condição de correlação. □

Assim, com cada um dos tópicos provados, conseguimos o resultado de que ν é, de fato, uma medida k-pseudoaleatória.

Referências Bibliográficas

[AMM07] Alexander Arbieto, Carlos Matheus e Carlos Gustavo Moreira. *Aspectos Ergódicos Da Teoria Dos Números*. IMPA, 2007.

[Bal92] Antal Balog. Linear equations in primes. *Mathematika*, 39(2):367–378, 1992.

[Blo10] Thomas Bloom. The Green-Tao Theorem On Arithmetic Progressions Within The Primes. Dissertação de Mestrado, University of Bristol, 2010.

[CFZ14] David Conlon, Jacob Fox e Yufei Zhao. The Green-Tao Theorem: An Exposition. *arXiv preprint arXiv:1403.2957*, 2014.

[Erd49] Paul Erdös. On A New Method In Elementary Number Theory Which Leads To An Elementary Proof Of The Prime Number Theorem. *Proceedings Of The National Academy Of Sciences*, 35(7):374–384, 1949.

[Erd97] Paul Erdős. Some Of My Favorite Problems And Results. Em *The Mathematics of Paul Erdös I*, páginas 47–67. Springer, 1997. 14

[ET36] Paul Erdös e Paul Turán. On Some Sequences Of Integers. *Journal Of The London Mathematical Society*, 1(4):261–264, 1936. 14

[Fur81] Hiller Harry Furstenberg. *Recurrence In Ergodic Theory And Combinatorial Number Theory*. Princeton University Press, 1981. 5, 9, 13, 14, 15

[Gow01] William T. Gowers. A New Proof Of Szemerédi's Theorem. *Geometric and Functional Analysis*, 11(3):465–588, 2001. 5, 14, 22

[Gow10] William T. Gowers. Decompositions, Approximate Structure, Transference, And The Hahn–Banach Theorem. *Bulletin Of The London Mathematical Society*, 42(4):573–606, 2010. 8, 17, 27

[GT04] Ben Green e Terence Tao. The Primes Contain Arbitrarily Long Arithmetic Progressions. *arXiv preprint math/0404188*, 2004. 6, 23, 28, 51, 55

[Hö90] Chaim Samuel Hönig. *Análise Funcional E Aplicações*. Publicações Do Instituto De Matemática E Estatística Da Universidade De São Paulo, 1990. 29

[Jen09] Jonas Lindstrøm Jensen. Master Thesis - On The Existence Of Long Arithmetic Progressions In The Primes. Dissertação de Mestrado, University of Aarhus, 2009. 39

[MMST13] Fabio Brochero Martinez, Carlos Gustavo Moreira, Nicolau Saldanha e Eduardo Tengan. Teoria dos Números: Um Passeio Com Primos E Outros Números Familiares Pelo Mundo Inteiro. *Coleção Projeto Euclides*, 2013. 10

[Pea89] Giuseppe Peano. *Arithmetices Principia: Nova Methodo*. Fratres Bocca, 1889.

[Rot53] Klaus F Roth. On certain sets of integers. *Journal of the London Mathematical Society*, 1(1):104–109, 1953. 5

[Sel50] Atle Selberg. An Elementary Proof Of The Prime Number Theorem For Arithmetic Progressions. *Canadian J. Math*, 2:66–78, 1950. 11

[Sze75] Endre Szemerédi. On Sets Of Integers Containing No k Elements In Arithmetic Progression. *Acta Arith*, 27:299–345, 1975. 5, 14

[Tao] Terence Tao. A Remark On Goldston-Yıldırım Correlation Estimates. 32

[Tao05a] Terence Tao. The dichotomy between structure and randomness, arithmetic progressions, and the primes. *arXiv preprint math/0512114*, 2005.

[Tao05b] Terence Tao. The Dichotomy Between Structure And Randomness, Arithmetic Progressions, And The Primes. *arXiv preprint math/0512114*, 2005. 14

[Tao05c] Terence Tao. The Gaussian Primes Contain Arbitrarily Shaped Constellations. *arXiv preprint math/0501314*, 2005. 37

[TZ08] Terence Tao e Tamar Ziegler. The Primes Contain Arbitrarily Long Polynomial Progressions. *Acta Mathematica*, 201(2):213–305, 2008. 37

[VdC39] JG Van der Corput. Über summen von primzahlen und primzahlquadraten. *Mathematische Annalen*, 116(1):1–50, 1939. 6

[vdW27] Bartel L. van der Waerden. Beweis Einer Baudetschen Vermutung, Nieiw. Arch. Wisk. 15 (1927), 212–216. *Nieuw Arch. Wisk. 15, 212-216*, 1927. 5, 13

Índice Remissivo