PICME/IME/USP COMBINATÓRIA

NOTAS - 2011 (SEMESTRE 2)

YOSHIHARU KOHAYAKAWA, MIGUEL ABADI E GUILHERME MOTA (IME/USP)

1. Esquemas de votação / Base de Fourier

09/08/2011

1.1. **Esquemas de votação.** Se existem n pessoas que desejam decidir sobre uma questão e cada uma tem sua opinião, então como é possível que estas pessoas decidam sobre tal questão? A maneira como isto é feito é denominada *esquema de votação*. Suponha que cada pessoa i (com i pertencente ao conjunto $[n]$ de todas as pessoas envolvidas) tem uma opinião ($x_i = 1$ ou $x_i = 0$) sobre o assunto. Desta forma, definimos formalmente um esquema de votação como sendo uma função $f\colon \{0,1\}^n \to \{0,1\}$. Abaixo definimos alguns tipos de esquemas de votação, onde, dado $x \in \{0,1\}^n$, a quantidade de 1's em x é denotada por $|x|$.

- **Democracia** $f(x) = \begin{cases} 1, & \text{se } |x| > n/2; \\ 0, & \text{caso contrário.} \end{cases}$

  A maioria dos votos determina o resultado da votação.

- **Ditadura** $f(x) = x_1$.

  A pessoa 1 escolhe o resultado da votação, sem levar em consideração o voto de qualquer outra pessoa.

- **Maioria das maiorias** $f(x) = \text{maj}(\text{maj}(x_1, \ldots, x_i), \ldots, \text{maj}(x_j, \ldots, x_n))$, onde temos que, para $x \in \{0,1\}^k$, $\text{maj}(x) = \begin{cases} 1, & \text{se } |x| > k/2; \\ 0, & \text{caso contrário.} \end{cases}$.

  As pessoas são divididas em grupos de votação e a decisão é tomada baseada no que a maioria dos grupos decidiu (cada grupo vota em esquema de democracia).

- **Paridade** $f(x) = |x|(\mod 2)$.

  A escolha é feita de acordo com a paridade da quantidade de pessoas que votaram da mesma forma.

Notas por Guilherme Mota

- **Junta** $f(x) = x_1 \wedge x_2$.

  A escolha é feita baseada no acordo entre duas pessoas.
- **Tribos** $f(x) = (x_1 \wedge \ldots \wedge x_a) \vee \ldots \vee (x_{n-a+1} \wedge \ldots \wedge x_n)$.

  Basta que todas as pessoas de pelo menos uma tribo tenham voto 1 para esta ser a escolha.

Estamos interessados em estudar a *influência* de uma pessoa i (da i-ésima variável) na votação (na saída da função f em questão). Definimos a influência da i-ésima coordenada como a probabilidade de que, ao escolher o vetor x uniformemente ao acaso, o valor de $f(x)$ é alterado caso mudemos o valor de x. Formalmente, dada uma função $f\colon \{0,1\}^n \to \{0,1\}$, definimos a influência da i-ésima variável como

$$I_i(f) = \Pr_x(f(x) \neq f(x + e_i)), \text{ onde a soma é } \bmod 2.$$

Para facilitar o entendimento, vamos analisar quais os valores da influência para algumas das funções definidas acima.

- **Democracia.** Uma coordenada pode mudar a votação somente se temos um empate em todas as outras $n-1$ coordenadas. Desta forma (considerando n ímpar), temos

$$I_i(f) = 2\frac{\binom{n-1}{\frac{n-1}{2}}}{2^{n-1}} = \frac{\binom{n-1}{\frac{n-1}{2}}}{2^n} \approx \frac{1}{\sqrt{2\pi n}}.$$

- **Ditatorial.** Neste caso, claramente $I_1(f) = 1$ e $I_i(f) = 0$ para $i \in \{2, \ldots, n\}$.
- **Paridade.** Como qualquer mudança em uma coordenada implica mudança na paridade, temos $I_i(f) = 1$ para $i \in \{1, \ldots, n\}$.
- **Junta.** Este é um caso simples. $I_1(f) = I_2(f) = 1/2$ e $I_i(f) = 0$ para $i \in \{3, \ldots, n\}$.

Gostaríamos de encontrar uma função balanceada (função em que obtemos 0's e 1's a mesma quantidade de vezes) com a menor influência máxima. Isto é, não desejamos que nenhuma pessoa tenha uma influência muito grande na votação.

Vamos ver que, com os parâmetros corretos, a função "tribos" é uma função (quase) balanceada com uma máxima influência "pequena" (na verdade, sua máxima influência é assintoticamente a menor possível, como veremos adiante). Seja $n = k \ln 2 \log_2 k$. Assim, considere que as coordenadas estão divididas em $k \ln 2$ tribos com $\log_2 k$ pessoas cada.

Dado que cada tribo tem voto 1 com probabilidade $(1/2)^{\log_2 k}$, temos que

$$\begin{aligned}\Pr_x(f(x)=0) &= \left(1-\frac{1}{2^{\log_2 k}}\right)^{\frac{k\ln 2\log_2 k}{\log_2 k}} \\ &= \left(\left(1-\frac{1}{k}\right)^k\right)^{\ln 2} \to \frac{1}{2}.\end{aligned}$$

Portanto, já vimos que a função é balanceada. Vamos calcular agora a influência $I_i(f)$ de uma coordenada. Observe que a mudança na coordenada só altera a decisão se todas as outras coordenadas de sua tribo tiverem valor 1 e todas as outras tribos votaram 0. Assim,

$$\begin{aligned}I_i(f) &= \left(\frac{1}{2}\right)^{\log_2 k-1}\left(1-\frac{1}{k}\right)^{k\ln 2-1} \\ &= \frac{2}{k}\left(1+\frac{1}{k}\right)\left(1-\frac{1}{k}\right)^{k\ln 2} \\ &= \Theta\left(\frac{1}{k}\right) \\ &= \Theta\left(\frac{\log_2 n}{n}\right).\end{aligned}$$

Como comentado anteriormente, esta é a menor máxima influência que uma função balanceada pode ter (será provado mais adiante).

1.2. **Base de Fourier.** Considere o espaço das funções $f\colon \{0,1\}^n \to \mathbb{R}$ e seja o produto interno entre duas funções f e g dado por

$$\langle f,g\rangle = \mathbb{E}[fg] = \frac{\sum_{x\in\{0,1\}^n} f(x)g(x)}{2^n}.$$

Vamos construir uma base para tal espaço de funções. Para isto, considere as funções ditatoriais $\chi_i\colon \{0,1\}^n \to \{-1,1\}$ tal que $\chi_i(x) = (-1)^{x_i}$. Equivalentemente, podemos utilizar o conjunto $S = \{i\colon x_i = 1\}$ para representar o vetor x:

$$\chi_i(S) = (-1)^{|\{i\}\cap S|} = \begin{cases} 1, & \text{se } i\notin S; \\ -1, & \text{se } i\in S.\end{cases}$$

Ademais, defina $\chi_S = \prod_{i\in S}\chi_i$. Para $T\subset[n]$, temos $\chi_S(T) = (-1)^{|T\cap S|}$. Para $x\in\{0,1\}^n$, dizemos que $\chi_S(x) = (-1)^{|X\cap S|}$, onde $X = \{i\colon x_i = 1\}$. Vejamos algumas propriedades destas funções.

- $\chi_S\chi_T = \chi_{S\triangle T}$, onde $S\triangle T = (S\cup T) - (S\cap T)$.

Demonstração.

$$\begin{aligned}\chi_S\chi_T(R) &= (-1)^{|S\cap R|}(-1)^{|T\cap R|}\\ &= (-1)^{|S\cap R|+|T\cap R|}\\ &= (-1)^{2|T\cap S\cap R|}(-1)^{|(T\triangle S)\cap R|}\\ &= (-1)^{|(T\triangle S)\cap R|}\\ &= \chi_{S\triangle T}.\end{aligned}$$

□

- $\chi_S(x+y) = \chi_S(x)\chi_S(y)$.

  Demonstração.

$$\begin{aligned}\chi_S(x+y) &= \chi_S(X\triangle Y)\\ &= (-1)^{|(X\triangle Y)\cap S|}\\ &= (-1)^{|X\cap S|}(-1)^{|Y\cap S|}\\ &= \chi_S(x)\chi_T(y).\end{aligned}$$

  □

- $\langle \chi_S, \chi_T\rangle = \begin{cases} 1, & \text{se } S = T;\\ 0, & \text{se } S \neq T.\end{cases}$

  Demonstração. Observe que a coleção de funções $\{\chi_1, \ldots, \chi_n\}$ é independente, desde que cada χ_i só depende da coordenada x_i, que por sua vez é independente das outras coordenadas. Desta forma, temos que

$$\begin{aligned}\mathbb{E}[\chi_{S\triangle T}] &= \mathbb{E}\left[\prod_{i\in S\triangle T}\chi_i\right]\\ &= \prod_{i\in S\triangle T}\mathbb{E}[\chi_i]\\ &= \begin{cases} 1, & \text{se } S\triangle T = \emptyset, \text{ isto é, } S = T;\\ 0, & \text{se } S\triangle T \neq \emptyset, \text{ isto é, } S \neq T.\end{cases}\end{aligned}$$

  □

Desde que o conjunto das 2^n funções em $\{\chi_S\}_{S\subset[n]}$ é ortonormal, tal conjunto forma uma base ortonormal para o espaço das funções $f\colon \{0,1\}^n \to \mathbb{R}$. A esta base damos o nome de *base de Fourier*.

Como $\{\chi_S\}_{S\subset[n]}$ é uma base, toda função $f\colon \{0,1\}^n \to \mathbb{R}$ pode ser escrita como $f = \sum_{S\subset[n]} \hat{f}(S)\chi_S$. Os termos $\hat{f}(S)$ são chamados *coeficientes de Fourier* e esta soma é conhecida como *expansão de Fourier* de f.

Por simplicidade, vamos limitar nosso estudo às funções $f\colon \{0,1\}^n \to \{-1,1\}$.

As seguintes duas igualdades seguem facilmente das propriedades das funções χ_S.

$$\hat{f}(S) = \langle f, \chi_S \rangle. \tag{1}$$

$$\sum_S \hat{f}(S)\hat{g}(S) = \langle f, g \rangle. \tag{2}$$

Como corolário destas propriedades, obtemos

$$\mathbb{E}[f^2] = \langle f, f \rangle = \sum_S \hat{f}^2(S). \tag{3}$$

Ademais, se a imagem de f está em $\{-1,1\}$, então $\mathbb{E}[f^2] = 1$.

Os valores de esperança e variância de f podem ser escritos de forma simples.

$$\mathbb{E}[f] = \langle f, 1 \rangle = \langle f, \chi_\emptyset \rangle = \hat{f}(\emptyset). \tag{4}$$

$$\mathrm{Var}[f] = \mathbb{E}[f^2] - \mathbb{E}[f]^2 = \left(\sum_S \hat{f}^2(S)\right) - \hat{f}^2(\emptyset) = \sum_{S\neq\emptyset} \hat{f}^2(S). \tag{5}$$

Vamos estudar agora a influência utilizando os coeficientes de Fourier. Definimos a *influência total* de uma função f como sendo $I(f) = \sum_i I_i(f)$.

No que segue, considere a função $f_{e_i}(x) = f(x + e_i) = \sum_S \hat{f}(S)\chi_S(x + e_i)$ e também a função $\chi_{S,e_i}(x) = \chi_S(x + e_i)$.

Afirmativa 1. $\hat{f}_{e_i}(S) = \begin{cases} +\hat{f}(S), & \textit{se } i \notin S; \\ -\hat{f}(S), & \textit{se } i \in S. \end{cases}$

Demonstração.

$$\hat{f}_{e_i}(S) = \langle f_{e_i}, \chi_S \rangle = \langle g, \chi_S \rangle = \hat{f}(S)\langle \chi_{S,e_i}, \chi_S \rangle$$

$$= \begin{cases} +\hat{f}(S), & \text{se } i \notin S; \\ -\hat{f}(S), & \text{se } i \in S. \end{cases}$$

□

Agora considere a função $f_i(x) = f(x) - f_{e_i}(x)$, que tem imagem em $\{-1, 0, 1\}$.

Afirmativa 2. $\hat{f}_i(S) = \begin{cases} 2\hat{f}(S), & \text{se } i \in S; \\ 0, & \text{se } i \notin S. \end{cases}$

Demonstração. Sabemos que, pela definição da função f_i, temos $\hat{f}_i(S) = \hat{f}(S) - \hat{f}_{e_i}(S)$. Pela Afirmativa 1, $\hat{f}_i(S) = 2\hat{f}(S)$ sempre que $i \in S$ e $\hat{f}_i(S) = 0$ sempre que $i \notin S$. □

Dada a afirmativa acima, temos

$$f_i = 2 \sum_{S\colon i \in S} \hat{f}(S)\chi_S. \tag{6}$$

Pela definição de influência, veja que $I_i(f)$ é igual à fração de vetores x tais que $f_i(x) \neq 0$, isto é, $f_i(x) \in \{-2, 2\}$. Desta forma, temos que $\mathbb{E}[f_i^2| = 4\Pr_x(f_i(x) \neq f_{e_i}(x))$. Portanto,

$$I_i(f) = \frac{\mathbb{E}[f_i^2|}{4}. \tag{7}$$

Teorema 3. $I_i(f) = \sum_{S\colon i \in S} \hat{f}^2(S)$.

Demonstração. Por (7), temos $I_i(f) = \mathbb{E}[f_i^2|/4 = \langle f_i^2, f_i^2\rangle/4$. Mas, por (3), $\langle f_i^2, f_i^2\rangle = \sum_S \hat{f}_i^2(S)$. Usando a Afirmativa 2,

$$\begin{aligned} I_i(f) &= \frac{\sum_{S\colon i \in S} \left(2\hat{f}(S)\right)^2}{4} \\ &= \sum_{S\colon i \in S} \hat{f}^2(S). \end{aligned}$$

□

O seguinte corolário é facilmente verificado.

Corolário 4. $I(f) = \sum_S |S|\hat{f}^2(S)$.

O seguinte resultado diz que em qualquer esquema de votação sempre existe uma pessoa com influência $\Omega(\log_2 n/n)$.

Teorema 5 (Teorema KKL [3])**.** *Para qualquer função $f\colon \{0,1\}^n \to \{-1,1\}$, existe $i \in [n]$ tal que $I_i(f) = \Omega(\mathrm{Var}(f)\frac{\log_2 n}{n})$.*

No que segue, dado $J \subset [n]$, considere $I_J(f) = \Pr_x(f \text{ não é constante em } F_J(x))$, onde temos $F_J(x) = \{y \in \{0,1\}^n\colon y_i = x_i \text{ para todo } i \notin J\}$. O seguinte teorema (que apresentamos sem

prova), devido a Kalai, Kahn e Linial, diz que, não importa qual seja o esquema de votação, sempre existe um "pequeno" grupo de pessoas que consegue determinar o resultado.

Teorema 6. *Para todo $\varepsilon > 0$ e qualquer função $f\colon \{0,1\}^n \to \{-1,1\}$, existe um subconjunto $J \subset [n]$ tal que $|J| \leq c_v(\log(1/\varepsilon))\frac{n}{\log_2 n}$ tal que $I_J(f) \geq 1-\varepsilon$, onde c_v depende da variância de f.*

Provaremos uma a seguinte versão mais simples do Teorema 5.

Teorema 7 (Teorema KKL simplificado)**.** *Para qualquer função $f\colon \{0,1\}^n \to \{-1,1\}$, existe i tal que $I_i(f) = \Omega(\mathrm{Var}(f)\frac{1}{n})$.*

Demonstração. Sabemos, pelo Corolário 4, que

$$\sum_{i=1}^{n} I_i(f) = I(f) = \sum_{S} \hat{|}S|f^2(S) \geq \mathrm{Var}[f].$$

Assim,

$$\frac{\sum_{i=1}^{n} I_i(f)}{n} \geq \frac{\mathrm{Var}[f]}{n}.$$

Logo, existe i tal que

$$I_i(f) = \Omega(\mathrm{Var}[f]\frac{1}{n}).$$

□

O seguinte resultado também é de simples verificação.

Teorema 8. *Para qualquer função balanceada $f\colon \{0,1\}^n \to \{-1,1\}$, temos $I(f) \geq 1$.*

Demonstração. Notando que $\mathbb{E}[f]^2 = 0$, pois a função é balanceada, temos

$$\begin{aligned} I(f) = \sum_{S} \hat{|}S|f^2(S) &\geq \left(\sum_{S\colon S\neq\emptyset} \hat{f}^2(S)\right) + \hat{f}^2(\emptyset) - \hat{f}^2(\emptyset) \\ &= \langle f, f\rangle - \hat{f}^2(\emptyset) \\ &= 1 - \hat{f}^2(\emptyset) \\ &= 1 - \mathbb{E}[f]^2 \\ &= 1. \end{aligned}$$

□

O seguinte teorema mostra que se a influência de uma função balanceada é igual a 1, então tal função certamente é uma função ditatorial.

Teorema 9. *Se f é uma função booleana balanceada tal que $I(f) = 1$, então $f(x) = x_i$ para algum $i \in [n]$.*

Agora veremos que toda função periódica em $[0, 2\pi]$ pode ser aproximada por uma combinação linear de senos e cossenos. Usaremos o fato de que $\{1, \operatorname{sen} mx, \cos mx\}$ é uma base ortogonal, onde estamos considerando o produto interno $\langle F, F\rangle = \int_0^{2\pi} f(x)g(x)dx$.

Considere a função

$$F(x) = c + \sum_{i=1}^{m} (a_i \cos ix + b_i \operatorname{sen} ix).$$

Calculemos agora algumas propriedades de tal função F.

- $\langle F, F\rangle = c^2 2\pi + \sum_{i=1}^{m} (\pi a_i^2 + \pi b_i^2)$;
- $\mathbb{E}[F] = \frac{\langle F,1\rangle}{2\pi} = c$;
- $\mathbb{E}[F^2] = \frac{\langle F,F\rangle}{2\pi} = c^2 + \sum_{i=1}^{m} \left(\frac{a_i^2}{2} + \frac{b_i^2}{2}\right)$;
- $\operatorname{Var}[F] = \sum_{i=1}^{m} \frac{a_i^2+b_i^2}{2}$.

Como exemplo, tome a função periódica $f(x) = x$ em $[0, 2\pi[$. Considere $F(x) = \pi + \sum_{i=1}^{m} (-2/i) \operatorname{sen} ix$. Temos $c = (1/2\pi) \int_0^{2\pi} f(x)dx$, $a_k = (1/\pi) \int_0^{2\pi} f(x) \cos kxdx$ e $b_k = (1/\pi) \int_0^{2\pi} f(x) \operatorname{sen} kxdx$. Pelo que vimos anteriormente, obtemos $\mathbb{E}[F] = \pi$ e $\operatorname{Var}[F] = \sum_{i=1}^{m} 2/i^2 = \pi^2/3$.

Diversas referências sobre o assunto podem ser encontradas na página do projeto PICME.

2. Problema de Waring

23/08/2011

Em 1770, Edward Waring levantou a seguinte questão: existe k tal que todo inteiro positivo pode ser escrito como a soma de no máximo k n-ésimas potências de números naturais?

Primeiramente, daremos algumas definições e propriedades. Feito isso, vamos analisar a questão levantada por Waring.

Se $A = (0, a_1, a_2, \ldots)$ é uma sequência estritamente crescente de inteiros, denotamos por $A(n)$ a quantidade de termos positivos na sequência que são menores ou iguais a n.

Dado $n > 0$, defina $\phi(n) = A(n)/n$ e veja que $0 \leq \phi(n) \leq 1$. A *densidade* de A é dada por $d(A) = \inf_{n>0} \phi(n)$.

Vamos agora provar algumas propriedades relacionadas com a densidade.

Propriedade 10. *Seja $A = (0, a_1, a_2, \ldots)$ uma sequência estritamente crescente de inteiros tal que $a_1 > 1$. Temos que $d(A) = 0$.*

Demonstração. O resultado segue do fato de $\phi(1) = 0$. □

Propriedade 11. *Se $A = (0, 1, r+1, 2r+1, \ldots)$, então $d(A) = 1/r$.*

Demonstração. Note que

$$\phi(n) = \frac{1 + \left\lfloor \frac{n-1}{r} \right\rfloor}{n}.$$

Com isso, basta realizar alguns cálculos.

$$\begin{aligned} \phi(n) &\leq \frac{1 + \frac{n-1}{r}}{n} \\ &= \left(\frac{r-1}{n} + 1 \right) \frac{1}{r} \\ &= \frac{1}{r} + \frac{r-1}{rn}. \end{aligned}$$

Mas $\phi(n) \geq 1/r - 1/nr$. Portanto, $\phi(n)$ tende a $1/r$ quando n tende a infinito. □

Propriedade 12. *Se $A = (0, 1, 2^2, 3^2, \ldots)$, então $d(A) = 0$.*

Demonstração. $0 \leq \phi(n) = \lfloor \sqrt{n} \rfloor / n \leq 1/\sqrt{n}$. Assim, $\lim_{n\to\infty} \phi(n) = 0$. □

Propriedade 13. *$d(A) = 1$ se e somente se $A = (0, 1, 2, 3, \ldots)$.*

Demonstração.

$$\begin{aligned} d(A) = 1 &\Leftrightarrow \inf_{n>0} \phi(n) = 1 \\ &\Leftrightarrow \phi(n) = 1 \text{ para todo } n > 0. \\ &\Leftrightarrow A(n) = n \text{ para todo } n > 0. \\ &\Leftrightarrow A = (0, 1, 2, 3, \ldots). \end{aligned}$$

□

Propriedade 14. *Seja $A = (0, a_1, a_2, \ldots)$ uma sequência estritamente crescente de inteiros. Se $d(A) = 0$ e $a_1 = 1$, então, para todos $\varepsilon > 0$ e M inteiro positivo, existe $m = m(\varepsilon, M) > M$ tal que $\phi(m) < \varepsilon$.*

Demonstração. Fixe $\varepsilon > 0$ e um inteiro positivo M. Como $a_1 = 1$, temos que $\phi(n) \geq 1/n$ para todo $n \geq 1$. Por outro lado, como $\inf_{n>0} \phi(n) = 0$, existe um m tal que $\phi(m) < \min\{\varepsilon, 1/M\}$. Portanto, $m > M$ e $\phi(m) < \varepsilon$. □

Propriedade 15. *Sejam A e B sequências estritamente crescente de inteiros com o primeiro elemento igual a 0. Temos que $1 - d(A + B) \leq (1 - d(A))(1 - d(B))$, onde $A + B$ é a sequência estritamente crescente formada pela ordenação dos elementos do conjunto $\{a + b \colon a \in A, b \in B\}$.*

Demonstração. Suponha que $d(A) = 0$ ou $d(B) = 0$ (s.p.g. suponha que $d(B) = 0$). Assim, como sabemos que $0 \in A$, temos $d(A + B) \geq d(A)$ e estamos feitos.

Suponha agora que $d(A) > 0$ e $d(B) > 0$. Por simplicidade, faça $C + A + B$, $\alpha = d(A)$, $\beta = d(B)$ e $\gamma = d(C)$. Primeiramente, note que existem $A(n)$ elementos de $[n]$ em A e, portanto, tais elementos, que denotamos por $a_1, a_2, \ldots, a_{A(n)}$, estão em C. Para cada par $\{a_k, a_{k+1}\}$ da sequência $(a_0 = 0, a_1, a_2, \ldots, a_{A(n)}, a_{A(n)+1} = n)$, considere os números $a_k, a_k + 1, a_k + 2, \ldots, a_k + \ell_k = a_{k+1} - 1, a_{k+1}$. Assim, vamos considerar, para $k = \{0, 1, \ldots, A(n)\}$, a sequência $(a_k + i)_{i=1}^{\ell_k}$ de tamanho ℓ_k , onde, claramente, definimos $\ell_k = a_{k+1} - a_k - 1$ quando $a_{k+1} > a_k + 1$, e $\ell_k = 0$ quando

$a_{k+1} = a_k + 1$. Com isso, tal sequência possui pelo menos $B(\ell_k)$ elementos em C. Portanto,

$$\begin{aligned} C(n) &\geq A(n) + \left(\sum_{k=0}^{A(n)-1} B(\ell_k) \right) + B(\ell_{A(n)} + 1) \\ &\geq A(n) + \beta \left(\left(\sum_{k=0}^{A(n)-1} \ell_k \right) + \ell_{A(n)} + 1 \right) \\ &= A(n) + \beta(n - A(n)) \\ &\geq \beta n + \alpha n - \alpha\beta n. \end{aligned}$$

Com isso, temos $\gamma \geq \alpha + \beta - \alpha\beta$. □

Observação. Com uma simples indução, é possível mostrar que a Propriedade 15 vale no caso geral, isto é, $1 - d(\sum_{i=1}^{k} A_i) \leq \prod_{i=1}^{k}(1 - d(A_i))$.

Propriedade 16. *Toda sequência A estritamente crescente com densidade positiva é uma base dos números naturais, i.e., existe k tal que $\sum_{i=1}^{k} A$ resulta no conjunto dos naturais.*

O seguinte fato é necessário para a prova da Propriedade 16.

Fato 17. *Dadas duas sequências estritamente crescentes A e B, temos que, se $A(n) + B(n) \geq n$, então n ocorre em $A + B$.*

Demonstração. Se n ocorre em A ou B, então estamos feitos, pois $0 \in A$ e $0 \in B$. Suponha que $n \notin A$ e $n \notin B$. Assim, $A(n) = A(n-1)$ e $B(n) = B(n-1)$. Desta forma, a hipótese pode ser reescrita como $A(n-1) + B(n-1) \geq n$. Sejam $(a_i)_{i=1}^{r}$ e $(b_i)_{i=1}^{s}$, respectivamente, as sequências dos inteiros em $[n-1]$ que estão em A e B, onde $r + s \geq n$. Assim, pelo princípio da casa dos pombos, existem p e q tais que $a_p = n - b_q$, dado que $(n - b_j)_{j=1}^{s} \in [n-1]$. Isto é, $n = a_p + b_q$ e o resultado segue. □

Demonstração da Propriedade 16. Considere uma sequência A estritamente crescente com densidade positiva. Seja $A_k = \sum_{i=1}^{k} A$. Tome k tal que $d(A_k) > 1/2$ e seja $2k = 2k$. Sabemos que $A_k(n) \geq d(A_k)n > n/2$ para todo $n > 0$. Ademais, $2A_k(n) > n$. Portanto, pelo Fato 17, $n \in A_{2k}$ para todo $n > 0$. Assim, $d(A_{2k}) = 1$. O resultado segue da Propriedade 13. □

3. Problema de Waring - Continuação

30/08/2011, 06/09/2011 e 13/09/2011

Em 1909, Hilbert deu uma resposta afirmativa ao questionamento feito por Waring.

Teorema 18 (Teorema de Hilbert–Waring [2])**.** *Dado um inteiro positivo n, existe k tal que todo inteiro positivo pode ser escrito como a soma de no máximo k n-ésimas potências de números naturais.*

Para provar o Teorema de Hilbert–Waring, vamos precisar do seguinte lema.

Lema 19 (Lema fundamental)**.** *Dado n natural, existem k, c tais que o número $r_k(m)$ de soluções inteiras não negativas da igualdade $\sum_{i=1}^{k} x_i^n = m$, com $x_i \geq 0$, satisfaz $r_k(m) < cN^{(k/n)-1}$ para todos m, N tais que $1 \leq m \leq N$.*

Demonstração do Teorema 18. Seja n inteiro positivo e sejam k, c obtidos através do Lema 19 aplicado com n. Fixe $0 < \varepsilon < 1/(2ck^{k/n})$ e $N > 1/(\varepsilon c)^{n/k}$.

Considere a sequência $A_n = (0^n, 1^n, 2^n, \ldots)$ e a sequência $A_{n,k} = kA_n$ formada pelos elementos que são soma de k elementos de A_n. Suponha, por contradição, que $d(A_{n,k}) = 0$. Aplicando a Propriedade 14 com ε e N, temos, pelo fato de $A_{n,k}$ conter 1 e $d(A_{n,k}) = 0$, que

$$A_{n,k}(N) < \varepsilon N. \tag{8}$$

Considere $R_k(N) = \sum_{m=0}^{N} r_k(m)$, isto é, o número de soluções inteiras não negativas da desigualdade $\sum_{i=1}^{k} x_i^n \leq N$. Observe que, se $x_i \leq (N/k)^{1/n}$ para $1 \leq i \leq k$, então temos uma solução para esta desigualdade. Portanto,

$$\begin{aligned} R_k(N) &\geq \left(\left\lfloor \left(\frac{N}{k} \right)^{1/n} \right\rfloor + 1 \right)^k \\ &> \left(\frac{N}{k} \right)^{k/n}. \end{aligned} \tag{9}$$

Mas note que

$$
\begin{aligned}
R_k(N) &= r_k(0) + \sum_{m=1}^{N} r_k(m) \\
&< 1 + A_{n,k}(N) c N^{k/n-1} \\
&< 1 + \varepsilon N c N^{k/n-1} \\
&= 1 + \varepsilon c N^{k/n} \\
&< 2\varepsilon c N^{k/n} \\
&< \left(\frac{N}{k}\right)^{k/n},
\end{aligned} \tag{10}
$$

onde a primeira desigualdade segue da desigualdade $r_k(m) < cN^{k/n-1}$, que é válida pelo Lema 19; a segunda desigualdade segue da desigualdade (8); a penúltima desigualdade segue da escolha de N e a última desigualdade segue da escolha de ε. Desta forma, temos um absurdo, desde que as desigualdades (9) e (10) são contraditórias.

Concluímos que $d(A_{n,k}) > 0$. Mas, pela Propriedade 16, tal sequência é uma base dos números naturais. Assim, o resultado está provado. □

No que segue, provaremos diversos lemas que serão úteis na prova do Lema 19 (Lema Fundamental).

Lema 20. *Considere a equação* $a_1 z_1 + a_2 z_2 = m$, *com* $|a_2| \leq |a_1| \leq A$ *e* $\mathrm{mdc}(a_1, a_2) = 1$. *O número de soluções desta equação satisfazendo* $\max\{z_1, z_2\} \leq A$ *é no máximo* $3A/|a_1|$.

Demonstração. Desde que $|a_1| \leq A$, se existem menos que duas soluções para a equação, então não há o que fazer. Desta forma, sejam (z_1, z_2) e (z_1', z_2') duas soluções distintas da equação. Portanto, $a_1 z_1 + a_2 z_2 = a_1 z_1' + a_2 z_2'$, isto é, $a_1(z_1 - z_1') = a_2(z_2' - z_2)$. Como $\mathrm{mdc}(a_1, a_2) = 1$, temos que a_1 divide $(z_2' - z_2)$. Como consequência disto, sabemos que $|a_1| < |z_2' - z_2|$. Deste modo, existem no máximo $2A/|a_1| + 1 \leq 3A/|a_1|$ soluções. □

Lema 21. *Considere a equação*

$$
\sum_{i=1}^{\ell} a_i z_i = m, \tag{11}
$$

com $M = \max_{1\le i\le \ell}\{|a_i|\} \le A$ *e* $\mathrm{mdc}(a_1,\ldots,a_\ell) = 1$. *O número de soluções da equação (11) satisfazendo* $\max\{z_1,\ldots,z_\ell\} \le A$ *é no máximo* $c(\ell)A^{\ell-1}/M$, *onde* $c(\ell)$ *é uma constante que só depende de* ℓ.

Demonstração. Provamos por indução em ℓ. Observe que para $\ell = 2$, o Lema 20 nos fornece o resultado. Suponha s.p.g. que $M = |a_\ell|$.

Se $a_i = 0$ para $1 \le i \le \ell - 1$), então temos que $M = a_\ell = 1$ e a equação (11) é equivalente à equação $z_l = m$, que contém no máximo $(2A+1)^{\ell-1} \le (3A)^{\ell-1} = c(\ell)A^{\ell-1}/M$ soluções, onde fizemos $c(\ell) = 3^{\ell-1}$.

Consideremos agora o caso onde existe um a_i com $1 \le i \le \ell - 1$ que é diferente de zero. Considere a equação

$$\sum_{i=1}^{\ell-1}\left(\frac{a_i}{\delta}\right) z_i = m', \tag{12}$$

onde $\delta = \mathrm{mdc}(a_1,\ldots,a_{\ell-1})$ e tomamos $M' = \max_{1\le i\le \ell-1}\{|a_i|/\delta\}$. Observe que a equação (11) é equivalente a

$$\delta m' + a_\ell z_\ell = m. \tag{13}$$

Veja que $|m'| \le \sum_{i=1}^{\ell-1}\left|\frac{a_i}{\delta}\right| |z_i| \le (\ell-1)M'A$ e, claramente, $|z_\ell| \le (\ell-1)M'A$. Portanto, podemos aplicar o Lema 20 com a equação (13), obtendo que existem no máximo $3(\ell-1)M'A/M$ soluções para (13).

Por hipótese indutiva, observamos que a equação (12) possui no máximo $c(\ell-1)A^{\ell-2}/M'$ soluções, para alguma constante $c(\ell-1)$ que depende somente de $\ell - 1$. Portanto, e equação (11) possui no máximo $(3(\ell-1)M'A)c(\ell-1)A^{\ell-2}/(M'M) = c(\ell)A^{\ell-1}/M$, onde $c(\ell) = 3(\ell-1)c(\ell-1)$. □

Lema 22. *Para todo* $\ell \ge 3$, *existem* $c'(\ell)$ *e* $c''(\ell)$ *tais que, para todos inteiros positivos* A, B *com* $1 \le A \le B \le c'(\ell)A^{\ell-1}$, *o somatório da quantidade de soluções de todas as equações na forma* $\sum_{i=1}^{\ell} a_i z_i = 0$, *com* $|a_i| \le A$ *e* $|z_i| \le B$ *para* $1 \le i \le \ell$ *é no máximo* $c''(\ell)(AB)^{\ell-1}$.

Demonstração. Fixe $\ell \ge 3$. Se $a_i = 0$ para $1 \le i \le \ell$, então $(2B+1)^\ell < 3^\ell B^\ell = 3^\ell c'(\ell)(AB)^{\ell-1}$ é o número máximo de soluções que podem existir.

Considere agora o caso onde existe $i \in \{1,\ldots,\ell\}$ tal que $a_i \ne 0$ e $\mathrm{mdc}(a_1,\ldots,a_\ell) = 1$. Vamos dividir as equações em tipos, onde consideramos $M = \max_{1\le i\le \ell}\{|a_i|\}$. Dizemos que uma equação é do tipo m se $A/2^{m+1} < M \le A/2^m$. Desta forma, existem no máximo $(2A/2^m + 1)^\ell$ equações do

tipo m. Isto é, no máximo $3^\ell A^\ell/2^{m\ell}$ equações do tipo m. Mas, pelo Lema 21, cada equação tem no máximo $c(\ell)B^{\ell-1}2^m/A$ soluções. Portanto, cada tipo m fornece no máximo $c_m(\ell)(AB)^{\ell-1}(2^m)^{1-\ell}$ soluções, onde $c_m(\ell)$ é uma constante que depende de ℓ. Desta forma, existem no máximo

$$\begin{aligned}\sum_{m=0}^{\infty} c_m(\ell)(AB)^{\ell-1}(2^{1-\ell})^m &= (AB)^{\ell-1}\sum_{m=0}^{\infty} c_m(\ell)(2^{1-\ell})^m \\ &= c''(\ell)(AB)^{\ell-1},\end{aligned}$$

onde $c''(\ell)$ é uma constante que depende apenas de ℓ.

Finalmente, considere o caso onde existe $i \in \{1,\ldots,\ell\}$ tal que $a_i \neq 0$ e $\mathrm{mdc}(a_1,\ldots,a_\ell) = \delta \neq 1$. As equações que estamos analisando são equivalentes às equações na forma $\sum_{i=1}^{\ell} a_i/\delta z_i = 0$. Como $\mathrm{mdc}(a_1/\delta,\ldots,a_\ell/\delta) = 1$, temos, do caso anterior, que o número de soluções de equações nesta forma é no máximo $c(\ell)(AB)^{\ell-1}/\delta^{\ell-1}$. Portanto, o máximo de soluções para equações na forma $\sum_{i=1}^{\ell} a_i z_i = 0$ é dado por

$$\begin{aligned}\sum_{\delta=2}^{A} \frac{c(\ell)(AB)^{\ell-1}}{\delta^{\ell-1}} &= c(\ell)(AB)^{\ell-1}\sum_{\delta=2}^{A}\frac{1}{\delta^{\ell-1}} \\ &< \frac{1}{\ell-2}c(\ell)(AB)^{\ell-1} \\ &= c''(\ell)(AB)^{\ell-1},\end{aligned}$$

onde $c''(\ell) = c(\ell)/(\ell-2)$. □

Lema 23. *Seja c um número real e considere multiconjuntos finitos A e B. O número de soluções da equação $x+y=c$ com $x \in A$ e $y \in B$ é no máximo metade do número de soluções da equação $x-y=0$ com $x \in A$, $y \in A$ ou $x \in B$, $y \in B$.*

Demonstração. Para cada par (x_i, y_j), com $x_i \in A$ e $y_j \in B$, que soluciona a equação $x+y=c$, existem $\mu_i\lambda_j \leq (\mu_i^2+\lambda_j^2)/2$ soluções, onde μ_i é a multiplicidade de x_i e λ_j é a multiplicidade de y_j. Portanto, existem no máximo

$$\frac{1}{2}\left(\sum_{x_i\in A}\mu_i^2 + \sum_{y_j\in B}\lambda_j^2\right)$$

soluções para $x+y=c$. Mas observe que, para cada par (x_i,x_i), existem μ_i^2 soluções para $x-y=0$ com $x \in A$, $y \in A$. Ademais, para cada par (y_j,y_j), existem λ_i^2 soluções para $x-y=0$ com $x \in B$,

$y \in B$. Isto é, existem

$$\left(\sum_{x_i \in A} \mu_i^2 + \sum_{y_j \in B} \lambda_j^2 \right)$$

soluções da equação $x - y = 0$ com $x \in A$, $y \in A$ ou $x \in B$, $y \in B$. Portanto, o resultado segue. □

Lema 24. *Seja c um número real, s, k inteiros positivos, $l = k2^s$ e considere multiconjuntos finitos A_i $(1 \leq i \leq l)$. O número de soluções de $\sum_{i=1}^{l} x_i = c$ com $x_i \in A_i$ é no máximo o número de soluções de*

$$\sum_{j=1}^{2^{s-1}} y^{(j)} - \sum_{j=2^{s-1}-1}^{2^s} y^{(j)},$$

com $y^{(j)} = \sum_{i=1}^{k} y_i^{(j)}$, onde $y_i^{(j)} \in A_{\omega k+i}$ $(0 \leq \omega \leq 2^s - 1)$.

Exemplo. O seguinte exemplo dá uma ideia de como funciona a prova para o lema acima.

Fixe c real, tome $s = 2$ e $k = 3$ (assim, temos que $l = 12$). Pomos $x = x_1 + \ldots + x_6$ e $y = x_7 + \ldots + x_{12}$. Pelo Lema 23, a equação $x + y = c$ não tem mais soluções que o número de soluções de

$$(x_1 + \ldots + x_6) - (x_7 + \ldots + x_{12}) = 0.$$

Vamos reescrever a última igualdade como $(x_1^{(1)} + \ldots + x_6^{(1)}) - (x_1^{(2)} + \ldots + x_7^{(2)}) = 0$. Desta forma, temos que $x_i^{(j)} \in A_i$ se $j = 1$ e $x_i^{(j)} \in A_{i+6}$ se $j = 2$.

Aplicando mais uma vez o Lema 23, mas desta vez com $x = ((x_1^{(1)} - x_1^{(2)}) + \ldots + (x_3^{(1)} - x_3^{(2)}))$ e $y = ((x_4^{(1)} - x_4^{(2)}) + \ldots + (x_6^{(1)} - x_6^{(2)}))$, temos não mais soluções que na equação

$$\left(\left(x_1^{(1)} - x_1^{(2)}\right) + \ldots + \left(x_3^{(1)} - x_3^{(2)}\right)\right) - \left(\left(x_4^{(1)} - x_4^{(2)}\right) + \ldots + \left(x_6^{(1)} - x_6^{(2)}\right)\right) = 0,$$

que é o mesmo que

$$\left(x_1^{(1)} + x_4^{(2)} - x_1^{(2)} - x_4^{(1)}\right) + \left(x_2^{(1)} + x_5^{(2)} - x_2^{(2)} - x_5^{(1)}\right) + \left(x_3^{(1)} + x_6^{(2)} - x_3^{(2)} - x_6^{(1)}\right) = 0.$$

Vamos reescrever a última igualdade como

$$\left(x_1^{(1)} + x_1^{(2)} - x_1^{(3)} - x_1^{(4)}\right) + \left(x_2^{(1)} + x_2^{(2)} - x_2^{(3)} - x_2^{(4)}\right) + \left(x_3^{(1)} + x_3^{(2)} - x_3^{(2)} - x_3^{(1)}\right) = 0.$$

Desta forma, temos que $x_i^{(j)} \in A_i$ ou A_{i+3} ou A_{i+6} ou A_{i+9}, isto é, $x_i^{(j)} \in A_{3\omega+i}$ $(0 \leq \omega \leq 2^s - 1)$. Pondo $x^{(j)} = x_1^{(j)} + x_2^{(j)} + x_3^{(j)}$, temos não mais soluções que a equação $x^{(1)} + x^{(1)} - x^{(3)} - x^{(4)} = 0$, o que confirma o lema para este exemplo específico.

A prova do Lema 24 é baseada essencialmente na ideia utilizada acima. Basta generalizar o procedimento que fizemos no exemplo.

Vamos trabalhar agora para mostrar o Lema 19 (Lema fundamental). A idéia é utilizar indução em n. Por conta disso, como acontece muitas vezes em uma prova por indução, para que possamos utilizar a hipótese indutiva, provamos um resultado mais forte que o desejado inicialmente. Segue abaixo a versão mais forte do Lema 19 que vamos provar.

Lema 25. *Dado n natural, existem $c_1 = c_1(n)$, $c_2 = c_2(n)$ e $k = k(n)$ tais que, para todo polinômio inteiro de grau n dado por $f(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$ com $|a_i| < c_2N^{i/n}$, o número $r_k(m)$ de soluções de*

$$\sum_{i=1}^{k} f(x_i) = m, \tag{14}$$

com $|x_i| \leq N^{1/n}$, satisfaz $r_k(m) \leq c_1N^{(k/n)-1}$ para todos m, N tais que $1 \leq m \leq N$.

Demonstração. A prova segue por indução em n. Considere $n = 1$. Faça $k = 1$ e considere o polinômio $f(x) = a_0x + a_1$. Note que temos (14) se e somente se $a_0(x_1 + x_2) = m - 2a_1$. Mas temos no máximo $2N^{1/n} + 1$ escolhas para x_1, dado que $|x_1| \leq N^{1/n}$. Uma vez que x_1 é fixado, só existe um valor possível para x_2. Portanto, $r_2(m) \leq 2N^{1/n} + 1 \leq 3N^{1/n} = 3N^{(k/n)-1}$.

Suponha agora que $n > 1$ e que o resultado é válido para $n - 1$. Por simplicidade, faça $k(1) = 2$ e $k(n-1) = k'$ e tome $s = \lfloor 4 \log_2 k' \rfloor - 1$. Seja $k = k(n) = (2n)2^{s+1}$. Desta forma, a igualdade 14 é equivalente à seguinte desigualdade.

$$\sum_{i=1}^{k/2} (f(x_i) + f(y_i)) = m, \tag{15}$$

onde $y_i = x_{k/2+i}$.

Aplicando o Lema 23 com $x = \sum_{i=1}^{k/2} f(x_i)$, $y = \sum_{i=1}^{k/2} f(y_i)$ e $c = m$, obtemos que (14) não tem mais soluções que $\sum_{i=1}^{k/2}(f(x_i) - f(y_i)) = 0$ com $|x_i|, |y_i| < N^{1/n}$.

Pondo $h_i = x_i - y_i$ $(1 \leq i \leq k/2)$, temos $|h_i| \leq 2N^{1/n}$. Vamos supor agora que $|y_i| \leq 2N^{1/n}$ (podemos fazer isso, pois tal suposição somente aumenta o número de soluções da equação). Vamos considerar agora equações sobre as variáveis y_i e h_i, em vez de x_i e y_i.

Fixe uma coleção h_i $(1 \leq i \leq k/2)$. Observe que

$$\begin{aligned} f(x_i) - f(y_i) &= f(y_i + h_i) - f(y_i) \\ &= \sum_{v=0}^{n-1} a_v \left((y_i + h_i)^{n-v} - y_i^{n-v} \right) \\ &= \sum_{v=0}^{n-1} a_v \left(\sum_{t=1}^{n-v} y_i^{n-v-t} h_i^t \right). \end{aligned}$$

Pondo $u = v + t$, temos que

$$\begin{aligned} f(x_i) - f(y_i) &= \sum_{v=0}^{n-1} a_v \left(\sum_{u=v+1}^{n} \binom{n-v}{u-v} y_i^{n-u} h_i^{u-v} \right) \\ &= h_i \sum_{v=0}^{n-1} a_v \left(\sum_{u=v+1}^{n} \binom{n-v}{u-v} y_i^{n-u} h_i^{u-v-1} \right) \\ &= h_i \sum_{u=1}^{n} a_{iu} y_i^{n-u}, \end{aligned}$$

onde $a_{iu} = \sum_{v=0}^{n-1} a_v \binom{n-v}{u-v} h_i^{u-v-1}$. Por simplicidade, faça $\varphi_i(y_i) = \sum_{u=1}^{n} a_{iu} y_i^{n-u}$. Assim, temos que $\sum_{i=1}^{k/2} (f(x_i) - f(y_i)) = 0$ se e somente se

$$\sum_{i=1}^{k_0 2^s} h_i \varphi_i(y_i) = 0, \tag{16}$$

onde fizemos $k_0 = 2n$, de modo que, pela escolha de k, temos $k/2 = k_0 2^s$.

Aplicando o Lema 24 na igualdade (16) com $l = k_0 2^s$, $c = 0$ e $x_i = h_i \varphi_i(y_i)$, temos que o número de soluções de (16) não é maior que o número de soluções de

$$\sum_{j=1}^{2^{s-1}} x^{(j)} - \sum_{j=2^{s-1}+1}^{2^s} x^{(j)} = 0, \tag{17}$$

onde, para $1 \leq j \leq 2^s$, temos $x^{(j)} = \sum_{i=1}^{k_0} x_i^{(j)} = \sum_{i=1}^{k_0} h_i \varphi_i(y_i^{(j)})$, com $|y_i^{(j)}| \leq 2N^{1/n}$ e, ademais, $x_i^{(j)} \in A_{\omega k_0 + i}$ $(1 \leq i \leq k_0)$. Note que A_r $(1 \leq r \leq 2^s)$ é o conjunto dos números na forma $h_r \varphi_r(y_r^{(j)})$ para um dado h_r e y_r arbitrário, com $|y_r| \leq 2N^{1/n}$. Mas note que, para o caso $\omega = 0$, a equação

(17) é equivalente a

$$\begin{aligned}
&h_1\varphi_1\left(y_1^{(1)}\right)+\ldots+h_{k_0}\varphi_{k_0}\left(y_{k_0}^{(1)}\right)\\
&+h_1\varphi_1\left(y_1^{(2)}\right)+\ldots+h_{k_0}\varphi_{k_0}\left(y_{k_0}^{(2)}\right)\\
&+\ldots\\
&+h_1\varphi_1\left(y_1^{(2^{s-1})}\right)+\ldots+h_{k_0}\varphi_{k_0}\left(y_{k_0}^{(2^{s-1})}\right)\\
&-h_1\varphi_1\left(y_1^{(2^{s-1}+1)}\right)+\ldots+h_{k_0}\varphi_{k_0}\left(y_{k_0}^{(2^{s-1}+1)}\right)\\
&-\ldots\\
&-h_1\varphi_1\left(y_1^{(2^s)}\right)+\ldots+h_{k_0}\varphi_{k_0}\left(y_{k_0}^{(2^s)}\right)=0.
\end{aligned}$$

Podemos reescrever a igualdade acima como

$$\sum_{i=1}^{k_0} h_i\left(\varphi_i\left(y_i^{(1)}\right)+\ldots+\varphi_i\left(y_i^{(2^{s-1})}\right)-\varphi_i\left(y_i^{(2^{s-1}+1)}\right)-\ldots-\varphi_i\left(y_i^{(2^s)}\right)\right).$$

Faça

$$z_i=\varphi_i\left(y_i^{(1)}\right)+\ldots+\varphi_i\left(y_i^{(2^{s-1})}\right).$$

Assim, nossa equação fica com a seguinte forma.

$$\sum_{i=1}^{k_0} h_i z_i=0. \tag{18}$$

Para estimar a quantidade de soluções possíveis da equação (18), precisamos primeiramente saber os limites sobre os quais as variáveis z_i podem variar. Mas z_i é definido em função de variáveis do tipo $\varphi_i(y_i^{(j)})$, onde $\varphi_i(y_i)=\sum_{u=1}^{n} a_{iu}y_i^{n-u}$, com $a_{iu}=\sum_{v=0}^{n-1} a_v\binom{n-v}{u-v}h_i^{u-v-1}$. Por nossas hipóteses, temos que

$$|a_v|\leq c_2N^{v/n},\ |h_i|\leq 2N^{1/n}.$$

Desta forma,

$$|a_v||h_i|^{u-v-1}\leq c_3N^{(u-1)/n},$$

onde $c_3=2^{u-v-1}c_2$. Com isso, temos que

$$|a_{iu}|\leq c_4N^{(u-1)/n},$$

onde $c_4 = c_3 \sum_{v=0}^{n-1} \binom{n-v}{u-v}$. Portanto,

$$\begin{aligned} |\varphi_i(y_i)| &= \sum_{u=1}^{n} a_{iu} y_i^{n-u} \\ &\leq c_4 \sum_{u=1}^{n} N^{(u-1)/n} \left(2N^{1/n}\right)^{n-u} \\ &= c_5 N^{(n-1)/n}, \end{aligned}$$

onde $c_5 = c_4 \sum_{u=1}^{n} 2^{n-u}$. Portanto, pela definição de z_i, concluímos que $|z_i| \leq c_6 N^{(n-1)/n}$ para alguma função c_6 que depende somente de n.

Considere agora $|z_i| = \overline{m}$, onde $\overline{m} \leq c_6 N^{(n-1)/n}$. Vamos estimar o número de soluções de $|z_i| = \overline{m}$, i.e., o número de soluções de

$$\sum_{j=1}^{2^{s-1}} \varphi_i(y_i^{(j)}) - \sum_{j=2^{s-1}+1}^{2^s} \varphi_i(y_i^{(j)}) = \overline{m}.$$

Como $k' < 2^{s-1}$, a equação acima pode ser reescrita como

$$\sum_{j=1}^{k'} \varphi_i(y_i^{(j)}) = m',$$

onde $m' = \overline{m} - \sum_{j=k'+1}^{2^{s-1}} \varphi_i(y_i^{(j)}) + \sum_{j=2^{s-1}+1}^{2^s} \varphi_i(y_i^{(j)})$. Reescrevendo novamente, temos

$$\sum_{j=1}^{k'} \left(\sum_{u=1}^{n} a_{iu} y_i^{(j)} \right) = m'. \tag{19}$$

Vamos aplicar a hipótese indutiva na igualdade (19). Mas para isso, precisamos verificar que as hipóteses necessárias são satisfeitas. Basta lembrar que $|a_{iu}| \leq c_4 N^{(u-1)/n} = c_4 (N^{(n-1)/n})^{(u-1)/(n-1)}$ e $|y_i^{(j)}| \leq 2N^{1/n} = 2(N^{(n-1)/n})^{1/(n-1)}$ e notar que $|m'| \leq c_7 N^{(n-1)/n}$ para alguma variável c_7 que depende somente de n. Portanto, pela hipótese indutiva, o número de equações da igualdade (19) não é maior que $c_8 N^{(k'-n+1)/n}$ para alguma variável c_8 que depende somente de n. Mas esta estimativa foi feita considerando valores fixos de $y_i^{(k'+1)}, \ldots, y_i^{(2^s)}$. Como vimos que $|\varphi_i(y_i)| \leq c_5 N^{(n-1)/n}$, a quantidade de soluções não ultrapassa $c_9 N^{(k'-n+1)/n} N^{(2^s-k')/n} = c_9 N^{(2^s-n+1)/n}$, para alguma variável c_9 que depende somente de n.

Lembre que tudo que fizemos até agora foi considerando que os números h_i estavam fixados. Precisamos saber a quantidade de soluções de

$$\sum_{i=1}^{k_0} h_{\omega k_0+i} z_{\omega k_0+i} = 0, \tag{20}$$

onde $0 \leq \omega \leq 2^s - 1$, os números h_i, com $1 \leq i \leq k_0 2^s$ são tais que $|h_i| \leq 2N^{1/n}$ e $|z_i| \leq c_6 N^{(n-1)/n}$.

Considere $K = \{h_i \colon 1 \leq i \leq k_0 2^s\}$ e $U_\omega(K)$ a quantidade de soluções de (20) dados por K considerando ω. Temos que o número que procuramos desde o início é dado por

$$r_k(m) \leq \sum_K \sum_\omega U_\omega(K) \leq 2^s \sum_K U_0(k).$$

Aplicando o Lema 22 com $l = 2n$, $A = 2N^{1/n}$ e $B = c_6 N^{(n-1)/n}$, temos que a equação (20) tem no máximo $c_{10} N^{2n-1}$ soluções, onde c_{10} depende somente de n. Assim, para algum c_{11} que depende somente de n, temos

$$\begin{aligned}
\sum_K U_0(K) &\leq c_{11} \left(N^{1/n}\right)^{(k/2)-2n} N^{2n-1} \left(N^{2s-n+1}\right) \\
&= c_{11} \left(N^{\frac{k}{2n}-2}\right) N^{2n-1} \left(N^{2s-n+1}\right) \\
&= c_{11} \left(N^{2^{s+1}-2}\right) N^{2n-1} \left(N^{2s-n+1}\right) \\
&= c_{11} N^{2^{s+2}-1}.
\end{aligned}$$

Portanto, concluímos que $r_k(m) \leq c_{12} N^{(k/n)-1}$, provando o resultado. □

4. Teorema de Paris–Harrington

20/09/2011, 27/09/2011, 04/10/2011, 11/20/2011, 18/10/2011

Comecemos com algumas definições. Dizemos que um conjunto $S \subset \mathbb{N}$ é dito *relativamente grande* se $|S| > \min\{x\colon x \in S\}$. Sejam n, m, k, $r \in \mathbb{N}$ com $m \geq n$. O conjunto $\{n, n+1, \ldots, m\}$ é denotado por $[n, m]$. Dizemos que $m \to (n)^k_r$ se, para toda coloração $P\colon [n,m]^k \to [r]$, existe subconjunto $H \subset [n, m]$ monocromático em P tal que H é relativamente grande.

Futuramente, provaremos o seguinte teorema.

Teorema 26. *Para todos n, k, $r \in \mathbb{N}$, existe $m \in \mathbb{N}$ tal que $m \to (n)^k_r$.*

O Teorema de Paris–Harrington diz que o resultado acima não pode ser provado na aritmética de Peano. Tal teorema foi o primeiro exemplo natural do Teorema da Incompletude de Gödel.

Teorema 27 (Teorema de Paris–Harrington [5])**.** *O Teorema 26 não pode ser provado na aritmética de Peano.*

Antes de discutirmos o teorema acima, vamos apresentar algumas ferramentas e contar um pouco de história.

4.1. **Números ordinais.** Em 1822, no livro "Théorie analytique de la chaleur" [1], de Joseph Fourier (1768–1830), apareceu pela primeira vez o conceito de séries de Fourier, que foi estudado em mais detalhes por Georg Cantor (1845–1918), que estudou, em 1870, o problema da unicidade da série de Fourier. Nesse estudo, dado um conjunto $X \subset \mathbb{R}$, Cantor definiu o que vem a ser sua derivada: $D^0(X) = X$ e $D^{n+1}(X)$ é o conjunto dos pontos de acumulação de $D^n(X)$. A unicidade foi provada nos casos em que $D^n(X) = \emptyset$ para algum $n \in \mathbb{N}$, onde X é o conjunto dos pontos em que a série não converge. Naturalmente, temos $D^\infty(X) = \bigcap_{n\geq 1} D^n(X)$. Assim, temos que $D^{\infty+1}(X) = D(D^\infty(X))$ e também $D^{\infty+\infty}(X) = \bigcap_{n\geq 1} D^{\infty+n}(X)$. Isto motivou Cantor a criar o que chamou de *números transfinitos*. Entre 1895 e 1897, Cantor desenvolveu a teoria dos números ordinais e cardinais, e demonstrou que $\mathbb{R}$ não é um conjunto enumerável.

Deixe-nos falar um pouco sobre a Teoria dos Conjuntos. Em 1908, Ernst Zermelo fez a primeira axiomatização da Teoria dos Conjuntos, que foi complementada em 1922/1923 com um novo axioma, proposto, independentemente, por Abraham Fraenkel e Thoralf Skolem. Tal axioma é chamado de *axioma da substituição*. A teoria criada por Zermello, juntamente com o axioma da substituição, formam a chamada *Teoria Axiomática de Zermello–Fraenkel* (por brevidade, usamos ZF quando queremos nos referir a tal teoria). Um dos axiomas de ZF é chamado de *axioma do infinito*, que

diz o seguinte: $\exists x(\emptyset \in x \wedge \forall y(y \in x \Rightarrow y \cup \{y\} \in x))$. Por tal axioma, observe que

$$\emptyset \in x.$$

$$\emptyset \cup \{\emptyset\} = \{\emptyset\} \in x.$$

$$\{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\} \in x.$$

$$\vdots$$

Definimos os números naturais da seguinte forma.

- $0 = \emptyset$;
- $n + 1 = n \cup \{n\}$.

Assim, fazemos

$$1 = \emptyset \cup \{\emptyset\} = \{\emptyset\} = \{0\};$$

$$2 = 1 \cup \{1\} = \{0\} \cup \{1\} = \{0, 1\};$$

$$3 = 2 \cup \{2\} = \{0, 1, 2\};$$

$$\vdots$$

$$n = \{0, 1, \ldots, n-1\}.$$

Vamos estudar algumas propriedades do conjunto $\omega = \{0, 1, \ldots\}$. Dado um conjunto A e $r \subset A \times A$. Dizemos que

- r é *reflexiva* se $\forall x(x \in A \Rightarrow (x, x) \in r)$;
- r é *simétrica* se $\forall x \forall y((x, y) \in r \Rightarrow (y, x) \in r)$;
- r é *transitiva* se $\forall x \forall y \forall z((x, y) \in r \wedge (y, z) \in r \Rightarrow (x, z) \in r)$;
- r é uma *ordem parcial* se é transitiva, $\forall x(x \in A \Rightarrow (x, x) \notin r)$ e $\forall x \forall y((x, y) \in r \Rightarrow (y, x) \notin r)$;
- r é uma *ordem total* (ou *linear*) se r é uma ordem parcial tal que $\forall x \forall y((x \in A \wedge y \in A) \Rightarrow ((x, y) \in r \vee (y, x) \in r \vee x = y))$ (esta última condição é chamada de *tricotomia*).

Como exemplo, se consideramos o conjunto $A = \mathcal{P}(n)$ e dizemos que $(x, y) \in r$ quando $x \subset y$ propriamente, então note que r é uma ordem parcial que não é total. Fazendo $A = \mathbb{R}$ e considerando que $(x, y) \in r$ se $x < y$, então r é uma ordem total.

Proposição 28. *A relação de pertinência "$\in$" em $\omega \times \omega$ é uma ordem linear.*

Demonstração. Dados n, m, $k \in \omega$, temos que $n \notin n$, pois $n = \{0, 1, \ldots, n-1\}$. Temos também que, se $n \in m$, então $m \notin n$. Por fim, não é difícil ver que a transitividade e a tricotomia também são válidas. □

Note que, se $x \in y \in \omega$, então $x \in \omega$. Daqui por diante, dizemos que $x < y$ se $x \in y$ em ω.

Uma relação $r \in A \times A$ é dita uma *boa ordem* se r é uma ordem linear tal que temos a seguinte propriedade: $\forall X(X \subset A \wedge X \neq \emptyset \Rightarrow \exists \min_r X)$, onde $x_0 = \min_r X$ se $\forall x \in X((x_0, x) \in r \vee x_0 = x)$. Dizemos que um conjunto A é *bem ordenado* se existe uma boa ordem $r \in A \times A$. Zermello provou o seguinte teorema em 1904.

Teorema 29. *Todo conjunto pode ser bem ordenado.*

Dizemos que um conjunto x é *transitivo* se $\forall y \forall z(z \in y \wedge y \in x \Rightarrow z \in x)$. Note que isto é o mesmo que $\forall y(y \in x \Rightarrow y \subset x)$. Não é difícil ver que ω é transitivo. Vejamos mais um exemplo.

Fato 30. *Todo $n \in \omega$ é transitivo.*

Demonstração. Lembre que $n+1 = n \cup \{n\}$. Vamos utilizar o princípio da indução. Note que $0 = \emptyset$ é transitivo e suponha que $n \in \omega$ é transitivo. Considere $k \in m \in n \cup \{n\}$. Se $m \in \{n\}$, então $m = n$, de onde concluímos que $k \in n$. Mas $n \subset n \cup \{n\}$. Logo, $k \in n + 1$. Agora considere o caso onde $k \in n$. Sabemos que $k \in m$. Utilizando a hipótese de que n é transitivo, temos, pelos mesmos argumentos utilizados anteriormente, que $k \in n + 1$. □

Dizemos que um conjunto α é *ordinal* se for transitivo e bem ordenado pela relação "$\in$". A classe dos conjuntos ordinais é denotada por Ord. É fácil ver que $\emptyset$, $n \in \omega$ e ω são conjuntos ordinais.

Lema 31. *Cada conjunto ordinal α é o conjunto de todos os ordinais menores que α*

Demonstração. Seja α um ordinal. Vamos mostrar que $\alpha = \{\beta \colon \beta \in \alpha \wedge \beta \in \text{Ord}\}$. Uma inclusão é óbvia. Para a outra, seja $\beta < \alpha$. Com α é transitivo, temos que $\beta \subset \alpha$. Logo, β é bem ordenado pela relação "$\in$". Pela transitividade de α, temos que, se $\gamma \in \eta \in \beta \in \alpha$, então $\gamma \in \alpha$ e $\eta \in \alpha$. Pela transitividade de "$\in$" em α, temos que $\gamma \in \beta$. Portanto, β é ordinal. □

Lema 32. *Se α for ordinal, então $\alpha \cup \{\alpha\}$ é ordinal.*

Lema 33. *Se α e β são ordinais, então $\alpha \in \beta$, $\alpha = \beta$ ou $\beta \in \alpha$.*

Lema 34. *Se X é ordinal, então $\bigcup X$ é um ordinal.*

Seja α um ordinal. Dizemos que $\alpha + 1 = \alpha \cup \{\alpha\}$ é um *ordinal sucessor*. Se α não é um ordinal sucessor, então é dito *ordinal limite*. Temos que $\emptyset$ e ω são ordinais limites.

Lema 35. *Seja X um conjunto de ordinais tal que $\forall(\alpha \in X)\exists(\beta \in X)(\alpha \in \beta)$. Então $\bigcup X$ é um ordinal limite.*

É um bom exercício provar as seguintes duas proposições.

Proposição 36. *λ é um ordinal limite se e somente se $\lambda = \bigcup \lambda$.*

Proposição 37 (Indução transfinita em ordinais)**.** *Seja $A(x)$ uma propriedade definida para todos os ordinais. Então*

$$(A(0) \wedge \forall\alpha \in \mathrm{Ord}(\forall\beta \in \mathrm{Ord}(\beta < \alpha \Rightarrow A(\beta)))) \Rightarrow (\forall\alpha(\alpha \in \mathrm{Ord} \Rightarrow A(\alpha)))$$

4.1.1. *Operações sobre ordinais.* Definimos agora algumas operações sobre os ordinais e suas propriedades.

Definição 38 (Adição de ordinais)**.** *Seja $+\colon \mathrm{Ord} \times \mathrm{Ord} \to \mathrm{Ord}$, onde $(\alpha, \beta) \mapsto \alpha + \beta$.*

- $\alpha + 0 = \alpha$.
- $\alpha + (\beta + 1) = (\alpha + \beta) + 1$, *para todos* α, $\beta \in \mathrm{Ord}$.
- $\alpha + \lambda = \bigcup_{\beta<\lambda}(\alpha + \beta)$, *se λ é um ordinal limite.*

Lema 39. *A operação de adição em ordinais é associativa. Isto é, Para todos α, β, $\gamma \in \mathrm{Ord}$, temos que $(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$.*

Demonstração. Provaremos por indução em γ. Se $\gamma = 0$, então $(\alpha + \beta) + 0 = \alpha + \beta = \alpha + (\beta + 0)$.

Considere, em primeiro lugar, o caso onde γ é ordinal sucessor. Neste caso, sabemos que $\gamma = \delta+1$ para algum ordinal δ. Suponha que $(\alpha+\beta)+\delta = \alpha+(\beta+\delta)$. Agora, usando este fato e as propriedades da operação de adição, temos que

$$\begin{aligned}
(\alpha + \beta) + \gamma &= (\alpha + \beta) + (\delta + 1) \\
&= ((\alpha + \beta) + \delta) + 1 \\
&= (\alpha + (\beta + \delta)) + 1 \\
&= \alpha + ((\beta + \delta) + 1) \\
&= \alpha + (\beta + (\delta + 1)) \\
&= \alpha + (\beta + \gamma).
\end{aligned}$$

Seja agora γ um ordinal limite. Assim, suponha que $(\alpha+\beta)+\theta = \alpha+(\beta+\theta)$, para todo $\theta < \gamma$. Temos

$$\begin{aligned}(\alpha+\beta)+\gamma &= \bigcup_{\theta<\gamma}((\alpha+\beta)+\theta)\\ &= \bigcup_{\theta<\gamma}(\alpha+(\beta+\theta))\\ &= \alpha(\beta+\gamma).\end{aligned}$$

□

Observação. A operação de adição em ordinais não é comutativa. Note que $1+\omega = \omega \neq \omega+1$.

Definição 40 (Produto de ordinais)**.** *Seja* $.: \mathrm{Ord}\times\mathrm{Ord}\to\mathrm{Ord}$*, onde* $(\alpha,\beta)\mapsto\alpha.\beta$ *(representamos* $\alpha.\beta$ *simplesmente por* $\alpha\beta$*).*

- $\alpha.0 = 0$.
- $\alpha(\beta+1) = \alpha\beta+\alpha$*, para todos* α*,* $\beta \in \mathrm{Ord}$.
- $\alpha\lambda = \bigcup_{\beta<\lambda}\alpha\beta$*, se* λ *é um ordinal limite.*

Lema 41. *A operação de produto em ordinais é associativa. Isto é, Para todos* α*,* β*,* $\gamma \in \mathrm{Ord}$*, temos que* $(\alpha\beta)\gamma = \alpha(\beta\gamma)$.

Lema 42. *A operação de produto em ordinais é distributiva. Isto é, Para todos* α*,* β*,* $\gamma \in \mathrm{Ord}$*, temos que* $\alpha(\beta+\gamma) = \alpha\beta+\alpha\gamma$.

Observação. A operação de produto em ordinais não é comutativa. Por exemplo, podemos notar que $2\omega = \bigcup_{n<\omega}2n = \bigcup_{n<\omega}\{0,1,\ldots,2n-1\} = \omega \neq \omega+\omega = \omega(1+1) = \omega.2$.

Definição 43 (Exponenciação de ordinais)**.** *Sejam* α*,* $\beta \in \mathrm{Ord}$*. Definimos* α^β *da seguinte forma.*

- $\alpha^0 = 1$.
- $\alpha^{\beta+1} = \alpha^\beta.\alpha$*, para todos* α*,* $\beta \in \mathrm{Ord}$.
- $\alpha^\lambda = \bigcup_{\beta<\lambda}\alpha^\beta$*, se* λ *é um ordinal limite.*

Lema 44. *Para todos* α*,* β*,* $\gamma \in \mathrm{Ord}$*, temos que* $\alpha^{\beta+\gamma} = \alpha^\beta.\alpha^\gamma$ *e* $\alpha^{(\beta\gamma)} = (\alpha^\beta)^\gamma$.

Teorema 45 (Forma normal de Cantor)**.** *Para todo ordinal* $\alpha > 0$*, existem* $t \in \omega\setminus\{0\}$*, com* $t > 0$*, ordinais* $\alpha_1 > \alpha_2 > \ldots\alpha_t \geq 0$ *e* $n_1,\ldots,n_t \in \omega\setminus\{0\}$ *tais que*

$$\alpha = \omega^{\alpha_1}n_1 + \omega^{\alpha_2}n_2 + \ldots + \omega^{\alpha_t}n_t.$$

Lema 46 (Princípio da compacidade)**.** *Seja k um inteiro positivo. Seja $\mathcal{A}$ uma família de subconjuntos finitos de $\mathbb{N}$. Suponha que, para toda coloração finita de $\mathbb{N}^k$, existe $A \in \mathcal{A}$ tal que A^k é monocromático. Então, para todo inteiro positivo r, existe $n_0 = n_0(r)$ tal que, se $n \geq n_0$, então, para toda r-coloração de $[n]^k$, existe $A \in \mathcal{A}$, com $A \subset [n]$ tal que A^k é monocromático.*

Teorema 47 (Teorema de Ramsey (versão infinita))**.** *Para todos inteiros positivos k, r, se fixamos uma r-coloração de $[\mathbb{N}]^k$, então existe um subconjunto de $\mathbb{N}$ que é infinito e monocromático.*

Provaremos agora o Teorema 26, isto é, para todos n, k, $r \in \mathbb{N}$, existe $m \in \mathbb{N}$ tal que $m \to (n)^k_r$.

Demonstração do Teorema 26. Fixe n, k, $r \in \mathbb{N}$. Considere uma r-coloração de $[n, +\infty[^k$. Pelo Teorema 47 (Ramsey infinito), existe $S \subset [n, +\infty[$ monocromático infinito. Seja T denotado pelos primeiros $\min(S)+1$ elementos de S. Logo, T é monocromático e $|T| = \min(S) + 1 > \min(S) = \min(T)$. Portanto, T é relativamente grande. Pelo Lema 46, existe o $m \in \mathbb{N}$ desejado. □

4.1.2. *Funções grandes.* Antes de tratar de teoremas que não podem ser provados na aritmética de Peano, vamos estudar um pouco alguns números muito grandes. Consideremos a seguinte brincadeira: descreva em um cartão retangular de 3 x 5 centímetros o maior inteiro positivo que você conseguir. O que você faria? Você pode encher o cartão com o número formado pela maior quantidade de 9's que você conseguir, mas certamente caberá um número maior no cartão.

Pensemos em como descrever uma função que cresce muito rapidamente. Nossa primeira tentativa é escrever no cartão as seguintes três linhas:

$$f_9(9), \text{ onde}$$

$$f_1(x) = 2x \text{ e}$$

$$f_{n+1}(x) = f_n^{(x)}(1), \text{ com } f^{(x)} = f(f^{(n-1)}).$$

Assim, notamos que, por exemplo, $f_3(x)$ representa um número que é uma torre formada por x algarismos 2. Mas vamos construir umas funções com maior crescimento. Dado que definimos essas funções $f_n(x)$, considere a função $f_\omega(n) = f_n(n)$, conhecida como *função de Ackermann.* Tal função nos permite definir novas funções que crescem ainda mais rapidamente. Definimos

$$f_{\omega+1}(n) = f_\omega^{(n)}(1)$$

$$f_{\omega+2}(n) = f_{\omega+1}^{(n)}(1)$$

$$\vdots$$

Com isso, podemos definir a função $f_{\omega.2}(n) = f_{\omega+n}(n)$.

Dado $\alpha = \omega^{\alpha_1}n_1 + \ldots + \omega^{\alpha_t}n_t$ ordinal, definimos a sequência $\alpha(n)$ tal que, se α_t é ordinal sucessor, então $\alpha(n) = \omega^{\alpha_1}n_1 + \ldots + \omega^{\alpha_t}(n_t - 1) + \omega^{\alpha_t - 1}n$ e, se α_t é ordinal limite, então temos $\alpha(n) = \omega^{\alpha_1}n_1 + \ldots + \omega^{\alpha_t}(n_t - 1) + \omega^{\alpha_t(n)}$. Podemos assim, definir a seguinte função, para todo ordinal α: $f_\alpha(n) = f^{(n)}_{\alpha-1}(1)$ se α é ordinal sucessor, e $f_\alpha(n) = f_{\alpha(n)}(n)$ se α é ordinal limite. Vamos definir agora

$$\omega_1 = \omega$$

$$\omega_2 = \omega^{\omega_1}$$

$$\omega_3 = \omega^{\omega_2}$$

$$\vdots$$

Fazendo $\varepsilon_0 = \lim_{n\to\infty} \omega_n$, vamos definir um último conjunto de funções.

$$f_{\varepsilon_0}(n) = f_{\omega_n}(n)$$

$$f_{\varepsilon_0+1}(n) = f^{(n)}_{\varepsilon_0}(n)$$

$$\vdots$$

Por enquanto, encerramos a discussão sobre funções com grande crescimento. Voltaremos a elas no momento oportuno.

4.2. **Paris–Harrington.** O Teorema 26 diz que, para todos n, k, $r \in \mathbb{N}$, existe $m \in \mathbb{N}$ tal que $m \to (n)^k_r$. Denote por $\mathrm{LR}(n, k, r)$ o menor m que torna tal afirmação verdadeira.

Considere as funções $f_1(n) = 2n$ e $f_r(n) = f^{(n)}_{r-1}(n)$. Se fixamos $k = 2$, obtemos o seguinte resultado.

Proposição 48. $\mathrm{LR}(n, 2, r) \geq f_r(n)$.

Demonstração. Exibiremos uma r-coloração em $[n, f_r(n)-1]^2$. Para x, y tais que $n \leq x < y < f_r(n)$, definimos $\overline{xy}$ pelo menor i tal que, para algum j, temos x, $y \in [f^{(j)}_i(n), f^{(j+1)}_i(n)[$. Note que

$$f^{(0)}_1(n) = n,\ f^{(1)}_1(n) = 2n,\ f^{(2)}_1(n) = 4n,\ \ldots,\ f^{(k)}_1(n) = 2^k n;$$
$$f^{(0)}_2(n) = n,\ f^{(1)}_2(n) = 2^n n,\ f^{(2)}_2(n) = 2^{2^n n}2^n n,\ \ldots$$

Se $i = r$ e $j = 0$, então x, $y \in [n, f_r(n)[$. Logo, $\overline{xy} \leq r$. Assim, vamos considerar a coloração onde $\{x, y\}$ recebe a cor $\overline{xy}$. Seja $H = \{x_1, \ldots, x_m\}$ um conjunto monocromático com a cor i. Desta forma, para $1 \leq u < v \leq m$, temos que x_u, $x_v \in [f^{(j)}_i(n), f^{(j+1)}_i(n)[$.

Se $i = 1$, então, fazendo $s = f_i^{(j)}(n)$, temos que $\{x_1, \ldots, x_m\} \subset [s, f_i(s)[$. Portanto,

$$\begin{aligned} m &\leq f_1(s) - s \\ &= f_1^{(j+1)}(n) - f_1^{(j)}(n) \\ &= 2^{j+1}n - 2^j n \\ &= f_1^{(j)}(n) \\ &= s. \end{aligned}$$

Mas sabemos que $s \leq \min\{x \colon x \in H\}$ e $m = |H|$. Portanto, $|H| \leq \min\{x \colon x \in H\}$, isto é, H não é relativamente grande.

Se $i > 1$, então, fazendo $s = f_{i-1}^{(k)}(n)$ para algum k, temos que

$$\begin{aligned} [s, f_i(s)[&= [f_{i-1}^{(k)}(n), f_{i-1}^{(s)}(s)[\\ &= [f_{i-1}^{(k)}(n), f_{i-1}^{(s+k)}(n)[\\ &= \bigcup_{l=k}^{s+k-1} [f_{i-1}^{(l)}(n), f_{i-1}^{(l+1)}(n)[. \end{aligned}$$

Assim, $\{x_1, \ldots, x_m\} \subset [s, f_i(s)[$. Como H é monocromático de cor i, se $x_u \in [f_{i-1}^{(l)}(n), f_{i-1}^{(l+1)}(n)[$ para algum $k \leq l \leq s + k - 1$, então, para todo $v \neq u$, temos que $x_v \notin [f_{i-1}^{(l)}(n), f_{i-1}^{(l+1)}(n)[$. Logo, $|H| = m \leq (s + k - 1) - k + 1 = s \leq \min\{x \colon x \in H\}$, isto é, H não é relativamente grande. □

Seja α um ordinal em sua forma normal de Cantor $\alpha = \omega^{\alpha_1} n_1 + \ldots + \omega^{\alpha_t} n_t$, com $\alpha_1 > \ldots > \alpha_t$. Definimos $T(\alpha)$ como o número de termos que aparece na forma normal de Cantor de α (temos $T(\alpha) = t$ para o α dado) e $N(\alpha) = \max\{n_1 + 1, \ldots, n_t + 1, N(\alpha_1), \ldots, N(\alpha_t)\}$, isto é, $N(\alpha)$ é igual ao sucessor do maior número inteiro que aparece na forma normal de Cantor de α. Por exemplo, se considerarmos $\alpha = \omega^{\omega^{\omega^7}}.5 + \omega^{\omega^4}.3 + 1$, temos que $T(\alpha) = 3$ e $N(\alpha) = 1 + 7 = 8$.

Dado um ordinal limite α e $n \in \omega$, defina

$$\begin{aligned}
\alpha_0 &= \alpha \\
\alpha_1 &= \alpha_0(n) \\
\alpha_2 &= \alpha_1(n) \\
&\vdots \\
\alpha_s &= \alpha_{s-1}(n),
\end{aligned}$$

onde $\alpha_{s-1}(n)$ é o único ordinal sucessor do conjunto $\{\alpha, \alpha_0(n), \alpha_1(n), \ldots, \alpha_{s-1}(n)\}$ (Escrevemos $\alpha((n))$ para representar $\alpha_{s-1}(n)$). Por exemplo, se $\alpha = \omega^{\omega+3} + \omega^2$, temos $\alpha_1 = \omega^{\omega+3} + \omega n$ e $\omega_2 = \omega^{\omega+3} + \omega(n-1) + n$. Para tornar esta definição mais clara, vejamos mais um exemplo: seja $\alpha = \omega^\omega$. Assim, temos

$$\begin{aligned}
\alpha_1 &= \omega^n \\
\alpha_2 &= \omega^{n-1} n \\
&\vdots \\
\alpha_{n+1} &= \omega^{n-1}(n-1) + \omega^{n-2}(n-1) + \ldots + \omega(n-1) + n.
\end{aligned}$$

Vamos definir agora uma hierarquia de funções conhecida como *Hierarquia de Wainer*, que serve como uma extensão das funções de Ackermann. Dado $\alpha \in \mathrm{Ord}$, definimos $f_\alpha \colon \mathbb{N} \to \mathbb{N}$ da seguinte forma:

$$\begin{aligned}
&f_1(n) = 2n; \\
&f_{\alpha+1}(n) = f_\alpha^{(n)}(n), \text{ se } \alpha + 1 \text{ é ordinal sucessor;} \\
&f_\alpha(n) = f_{\alpha((n))}(n), \text{ se } \alpha \text{ é ordinal limite.}
\end{aligned}$$

Definição 49. *Sejam f, $g \colon \mathbb{N} \to \mathbb{N}$. A função f* domina *g se existe $C \in \mathbb{N}$ tal que $f(n) > g(n)$ para todo $n \geq C$.*

Definição 50. *Uma afirmação $P(s,t)$ é dita* comprovadamente recursiva (CR) *se existe um algoritmo que decide se $P(s,t)$ é verdadeira e a prova desse algoritmo, em aritmética de Peano, sempre termina.*

Seja $\widetilde{P}$: $(\forall s \in \mathbb{N}\ \exists t \in \mathbb{N}(P(s,t)))$ uma afirmação em aritmética de Peano que é verdadeira nos naturais. Definimos $f_P(s)$ como o menor t que satisfaz $\widetilde{P}$.

Teorema 51 (Teorema de Kreisel [4])**.** *Se $\widetilde{P}$ é demonstrável na aritmética de Peano e $P(s,t)$ é CR, então existe $\alpha < \varepsilon_0$ tal que a função f_P é dominada por f_α.*

Definimos $\varepsilon_0(n) = \omega_n$ e $f_{\varepsilon_0}(n) = f_{\omega_n}(n)$. Discutiremos agora algumas propriedades importantes.

Propriedade 52. *f_{ε_0} domina f_α para todo $\alpha < \varepsilon_0$.*

Demonstração. Observe que se $\alpha < \beta < \varepsilon_0$ e $2 \leq m \leq N(\alpha)$ para algum m, onde β é ordinal limite, então $\alpha < \beta(m)$.

Mostremos primeiramente, por indução transfinita em β, que se $\alpha < \beta < \varepsilon_0$ e $2 \leq m \leq N(\alpha)$, então $f_\alpha(m) < f_\beta(m)$. Se $\beta = \alpha + 1$, então temos $f_\beta(m) = f_\alpha^{(m)}(m) = f_\alpha^{(m-1)}(f_\alpha(m)) > f_\alpha(m)$. Suponha agora que a afirmação é verdadeira para todo γ tal que $\alpha < \gamma < \beta$. Mostremos que, desta forma, o resultado é válido para β. Se $\beta = \gamma + 1$ para algum γ, então

$$\begin{aligned} f_\beta(m) &= f_\gamma^m(m) \\ &= f_\gamma^{(m-1)}(f_\gamma(m)) \\ &> f_\gamma(m) \\ &> f_\alpha(m), \end{aligned}$$

onde a última desigualdade segue da hipótese indutiva. Por outro lado, se β é ordinal limite, então temos $f_\beta(m) = f_{\beta(m)}(m) > f_\alpha(m)$.

Com isso, fixe $\alpha < \varepsilon_0$. Assim, existe s tal que $\alpha < \omega_s$. Tomando $m = \max\{s, N(\alpha)\}$, obtemos $f_{\varepsilon_0}(m) = f_{\omega_m}(m)$. Mas pelo que foi dito no primeiro parágrafo desta prova, sabemos que $f_{\omega_m}(m) > f_\alpha(m)$. □

Se α um ordinal em sua forma normal de Cantor $\alpha = \omega^{\alpha_1} n_1 + \ldots + \omega^{\alpha_t} n_t$, com $\alpha_1 > \ldots > \alpha_t$, então denotamos por $\nu_{\alpha_i}(\alpha)$ o i-ésimo coeficiente (n_i) de α.

Definição 53. *Dados $\beta < \alpha < \varepsilon_0$, definimos $\overline{\alpha\beta} = \max\{\delta \colon \nu_\delta(\alpha) \neq \nu_\delta(\beta)\}$.*

Propriedade 54. *Se $\alpha_1 > \ldots > \alpha_n$, onde $\alpha_i < \varepsilon_0$ para $1 \leq i \leq n$, então a seguinte igualdade é verdadeira: $\overline{\alpha_1\alpha_n} = \max\{\overline{\alpha_t\alpha_{t+1}} \colon 1 \leq t \leq n-1\}$.*

Considere para as próximas propriedades a 3-coloração χ^* sobre $[\varepsilon_0]^3$ dada por

$$\chi^*(\{\alpha,\beta,\gamma\}) = \begin{cases} 0, & \text{se } \overline{\alpha\beta} > \overline{\beta\gamma}; \\ 1, & \text{se } \overline{\alpha\beta} = \overline{\beta\gamma}; \\ 2, & \text{se } \overline{\alpha\beta} < \overline{\beta\gamma}. \end{cases}$$

Definição 55. *Seja* $S = \{\alpha_1, \ldots, \alpha_r\} \subset \mathrm{Ord}$ *tal que* $\max\{x\colon x \in S\} = \alpha_1$ *(Observe que* $|S| = r$*). Definimos também a função* $e_S(n)$ *como sendo uma torre de* n*'s com tamanho* r*.*

Propriedade 56. *Se* $\chi^*(S) = 1$*, então* $r \leq N(\alpha_1)$*.*

Propriedade 57. *Se* $\chi^*(S) = 2$*, então* $r \leq T(\alpha_1) + 1$*. Ademais, se* $\alpha < \omega_{S+1}$*, então temos que* $r \leq e_S(N(\alpha_1)) + 1$*.*

Propriedade 58. *Se* $\chi^*(S) = 0$ *e* $\alpha_1 < \omega^\omega$*, então* $r \leq N(\alpha_1) + 1$*.*

Seja α um ordinal e $n \in \mathbb{N}$. Definimos $T = T^{\alpha,n}$ a (n,α)-*função de translação* dada por

(i) $T(n) = \alpha$;

(ii) Suponha $T(m) = \beta$.

  (a) Se β é ordinal sucessor, então $T(m+1) = \beta - 1$;

  (b) Se β é ordinal limite, então $T(m+1) = \beta((m)) - 1$;

  (c) Se $T(m) = 0$, então o processo termina.

Definimos $U(\alpha, n) = u$ como sendo o menor natural tal que $T(u) = 0$. Observe que o domínio de T é dado por $\mathrm{dom}(T) = [n, U(\alpha,n)] \cap \mathbb{Z}$. Considere o seguinte exemplo para entender melhor a definição. Seja $\alpha = \omega^2$ e $n = 5$. Temos $T(5) = \omega^2$, $T(6) = \omega.4+4$, $T(7) = \omega.4+3, \ldots, T(10) = \omega.4$. Continuando, obtemos

$$T(20) = \omega.3$$

$$T(40) = \omega.2$$

$$T(80) = \omega$$

$$T(160) = 0,$$

de onde concluímos que $U(\omega^2, 5) = 160$.

Propriedade 59. $N(T^{\alpha,n}(m)) \leq \max\{N(\alpha), m\}$ *para todos* α*,* m*,* n*, desde que* $T^{\alpha,n}(m)$ *esteja definido.*

Propriedade 60. *Para todos* $1 \leq \alpha < \varepsilon_0$ *e* $n \in \mathbb{N}$*, temos* $U(n, \omega^\alpha) = f_\alpha(n)$*, onde* f_α *é como na hierarquia de Wainer.*

Defina a propriedade $P(S, m, n)$: $n > m$ e, para toda $(2s-1)$-coloração de $[m, n]^{s+1}$, existe um conjunto X tal que $|X| \geq h_s(\min\{x \colon x \in X\})$. Ademais, considere $P(s, t) = P(s+1, s, t)$.

Propriedade 61. *A afirmação* $P(s, m, n)$ *é expressível em aritmética de Peano.*

Propriedade 62. $(\forall s \forall m \exists n(P(s, m, n)))$.

Definiremos agora, para $x \geq 2$, a $(2x-1)$-coloração $\chi_x \colon [\omega_x]^{x+1} \to \{0, 1, \ldots, 2x-2\}$ e uma função monótona $h_x \colon \mathbb{N} \to \mathbb{N}$ tal que, se S (definido como na Definição 55) é monocromático em χ_x, então $|S| \leq h_x(N(\alpha_1))$.

Considere $x = 2$. Seja $\chi_2 \colon [\omega^\omega]^3 \to \{0, 1, 2\}$ tal que χ_2 é uma restrição de χ^* e $h_2(n) = n + 1$. Considere $x = 3$. Seja $\chi_3 \colon [\omega^{\omega^\omega}]^3 \to \{0, 1, 2, 3, 4\}$ tal que χ_3 é definida como segue, onde vamos considerar $S = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ e $\alpha'_i = \overline{\alpha_i \alpha_{i+1}}$, para $1 \leq i \leq 3$:

$$\chi_3(S) = \begin{cases} 3, & \text{se } \chi^*(\{\alpha_1, \alpha_2, \alpha_3\}) = 1; \\ 4, & \text{se } \chi^*(\{\alpha_1, \alpha_2, \alpha_3\}) = 2; \\ \chi^*(\{\alpha'_1, \alpha'_2, \alpha'_3\}), & \text{se } \alpha'_1 > \alpha'_2 > \alpha'_3; \\ 0, & \text{caso contrário.} \end{cases}$$

No que segue, vamos construir a função h_3. Seja S como na Definição 55 um conjunto monocromático em χ_3.

(a) Se $\chi_3(S) = 3$, então sabemos que $\chi^*(S \setminus \{\alpha_r\}) = 1$. Pela Propriedade 56, concluímos que $r \leq N(\alpha_1) + 1$.

(b) Se $\chi_3(S) = 4$, então sabemos que $\chi^*(S \setminus \{\alpha_r\}) = 2$. Pela Propriedade 57, concluímos que $r \leq e_2(N(\alpha_1)) + 2$.

(c) Se $\chi_3(S) \in \{0, 1, 2\}$, então sabemos que $\chi^*(S \setminus \{\alpha_r\}) = 0$. Pela definição de χ^*, temos $\alpha'_i > \alpha'_{i+1}$ para $1 \leq i \leq r-2$. Com um pouco de trabalho, é possível mostrar que o conjunto $\{\alpha'_1, \ldots, \alpha'_{r-2}\}$ é monocromático em χ_2. Mas, pela definição de χ_2, temos $r \leq h_2(N(\alpha'_1)) + 2 \leq h_2(N(\alpha_1)) + 2$ (pois h_2 é crescente).

Desta forma, façamos $h_3(n) = \max\{n + 1, e_2(n) + 2, h_2(n) + 2\}$.

Propriedade 63. $P(s, m, n)$ *é falsa para* $n \leq f_{\omega_{s-1}}(m)$.

Demonstração. Seja T a $(m, \omega_s(m))$-função de translação. Claramente, T está definida no intervalo $[m, U(m, \omega_s(m))]$. Note que $U(m, \omega_s(m)) = (m, \omega_s) = f_{\omega_{s-1}}(m)$. Assim, T está definida em $[m, n]$.

Seja a coloração $\chi'_s = [m, n]^{s+1} \to \{0, 1, \ldots, 2s-2\}$ dada por

$$\chi'_s = (\{x_1, \ldots, x_{s+1}\}_<) = \chi_s(\{T(x_1), T(x_2), \ldots, T(x_{s+1})\}_>).$$

Mas, sabemos que, para todo X monocromático em χ_S, temos $|X| \leq h'_s(\min X)$.

Considere $X = \{x_1, \ldots, x_{s+1}\}_<$ monocromático em χ'_s. Assim, $T(X) = \{T(x_1), \ldots, T(x_{s+1})\}_>$ é monocromático em χ_S. Portanto, $|X| = |T(X)| \leq h_s(N(T(x_1)))$. Pela Propriedade 59, temos $|X| \leq h_s(x_1) = h_s(\min X)$. □

Para demonstrar o Teorema 27 (Teorema de Paris–Harrington), basta utilizar a contra positiva do Teorema 51 (Teorema de Kreisel). Porém, para isso, precisamos que as seguintes condições sejam satisfeitas.

(PH0) $P(s, t)$ é expressível na aritmética de Peano;

(PH1) $P(s, t)$ é comprovadamente recursiva;

(PH2) $P(s, t)$ é falsa para todo $t < f_{\varepsilon_0}(s)$;

(PH3) Para todo s, existe t tal que $P(s, t)$ é verdadeira.

Mas das propriedades que discutimos, é fácil obter tais condições. A saber, utilizamos a Propriedade 61 para (PH0), Propriedade 63 para (PH2) e Propriedade 62 para (PH3).

Referências

[1] J. Fourier, *Théorie analytique de la chaleur*, Éditions Jacques Gabay, Paris, 1988, Reprint of the 1822 original.

[2] D. Hilbert, *Beweis für die Darstellbarkeit der ganzen Zahlen durch eine feste Anzahl n^{ter} Potenzen (Waringsches Problem)*, Math. Ann. **67** (1909), no. 3, 281–300.

[3] J. Kahn, G. Kalai, and N. Linial, *The influence of variables on boolean functions (extended abstract)*, 1988, pp. 68–80.

[4] G. Kreisel, *On the interpretation of non-finitist proofs. II. Interpretation of number theory. Applications*, J. Symbolic Logic **17** (1952), 43–58.

[5] J. Paris and L. Harrington, *A mathematical incompleteness in peano arithmetic*, Handbook for Mathematical Logic (Ed. J. Barwise) (1977).