

NOTAS DE AULA DO  
**PICME**  
PROGRAMA DE INICIAÇÃO CIENTÍFICA E MESTRADO  
EM  
COMBINATÓRIA

<http://www.ime.usp.br/~tcco/picme>

*Anotado por: Henrique Stagni*

*1º semestre de 2016*

## Conteúdo

<b>1</b>	<b>Paradoxo de Banach-Tarski</b>	<b>2</b>
1.1	Teorema de Cantor-Bernstein-Schröder-Banach . . . . .	2
<b>2</b>	<b>Polígonos equidecomponíveis geometricamente</b>	<b>4</b>
2.1	Teorema de Wallace-Bolyai-Gerwien . . . . .	4
<b>3</b>	<b>Paradoxo de Banach-Tarski (cont)</b>	<b>6</b>
3.1	Medidas Invariantes . . . . .	6
3.2	Rotações em $R_3$ . . . . .	8
3.3	Prova da Versão Fraca do Paradoxo de Hausdorff (Teorema 3.5) . . . . .	8
3.4	Prova do Paradoxo de Banach-Tarski . . . . .	9
<b>4</b>	<b>Teorema de Dehn</b>	<b>11</b>
<b>5</b>	<b>Ladrilhamento de Retângulos</b>	<b>13</b>
<b>6</b>	<b>Valuações não-arquimedianas</b>	<b>14</b>
6.1	Teorema de Mosky . . . . .	15
6.2	Valuações em grupos abelianos . . . . .	17
6.3	Valuações e Grupos Livres . . . . .	19
6.3.1	Quatérnios . . . . .	19
6.3.2	Grupos livres . . . . .	21
<b>7</b>	<b>Happy Ending Theorem</b>	<b>22</b>
7.1	Preliminares . . . . .	22
7.2	Demonstração do Teorema 7.1 . . . . .	26

<b>8</b>	<b>Flag Algebra</b>	<b>28</b>
8.1	Exemplo: Teorema de Mantel	28
8.2	Definições iniciais	30
8.3	Flag Álgebra	31
8.4	Método semidefinido	35
8.5	Exemplo método semidefinido: Teorema de Mantel	37
<b>9</b>	<b>Teorema de Erdős, Ginzburg e Ziv</b>	<b>38</b>
9.1	Primeira prova do Lema 9.2	38
9.2	Segunda prova do Lema 9.2	39

## 1 Paradoxo de Banach-Tarski

◇ ◇ ◇ *Aula 1 (01 de Março) — Yoshiharu Kohayakawa* ◇ ◇ ◇

Uma *isometria* entre dois conjuntos  $X, Y \subseteq \mathbb{R}^n$  é uma função  $\varphi : X \rightarrow Y$  bijetora que preserva distâncias, isto é, tal que  $d(x_1, x_2) = d(\varphi(x_1), \varphi(x_2))$  para todo  $x, y \in X$ . Dizemos que  $X$  e  $Y$  são *congruentes* se existe uma tal função e denotamos esse fato por  $X \cong Y$ .

Sejam  $X, Y \subseteq \mathbb{R}^n$ . Dizemos que  $X$  e  $Y$  são *equidecomponíveis* se se existem conjuntos  $X_1, \dots, X_k \subseteq X$ , dois a dois disjuntos e conjuntos  $Y_1, \dots, Y_k \subseteq Y$  dois a dois disjuntos tais que

- $X = \cup_{i=1}^k X_i, Y = \cup_{i=1}^k Y_i$
- $X_i \cong Y_i$  para todo  $1 \leq i \leq k$ .

Denotamos o fato de que  $X$  e  $Y$  é equidecomponível por  $X \equiv Y$  e também escrevemos  $X \equiv_k Y$  para denotar que  $X$  e  $Y$  são equidecomponíveis em no máximo  $k$  partes.

Note que  $\equiv$  é uma relação de equivalência. De fato, não é difícil mostrar que se  $X \equiv_k Y$  e  $Y \equiv_m Z$ , então  $X \equiv_{km} Z$ .

**Exercício 1.1.** Sejam  $X, Y \subseteq \mathbb{R}^2$  polígonos. Então  $X \equiv^8 Y$  se, e somente se,  $X \equiv Y$ .

**Teorema 1.2** (Banach-Tarski). *Suponha que  $B_1, B_2 \subseteq \mathbb{R}^3$  são bolas disjuntas de mesmo raio em  $\mathbb{R}^3$ . Então  $B_1 \cup B_2 \equiv_{10} B_1$ .*

### Pendente

◇ ◇ ◇ *Aula 2 (08 de Março) — Yoshiharu Kohayakawa* ◇ ◇ ◇

### 1.1 Teorema de Cantor-Bernstein-Schröder-Banach

O resultado demonstrado nesta aula permitirá resolver o seguinte exercício.

**Exercício 1.3.** Seja  $A = \{p \in \mathbb{R}^2 : |p| \leq \frac{1}{2}\}$  um disco de diâmetro unitário no plano e  $B = [0, 1] \times [0, 1]$  um quadrado de lado unitário. Prove que existem partições  $A = A_1 \cup A_2$  e  $B = B_1 \cup B_2$  (i.e., com  $A_1 \cap A_2 = \emptyset$  e  $B_1 \cap B_2 = \emptyset$ ) tais que  $A_1 \simeq B_1$  e  $A_2 \simeq B_2$ , onde  $X \simeq Y$  significa que  $X$  e  $Y$  são *semelhantes*.

**Teorema 1.4** (Cantor-Bernstein-Schröder-Banach). *Se  $A$  e  $B$  são conjuntos tais que  $|A| \leq |B|$  e  $|B| \leq |A|$ , então  $A \sim B$ .*

Provaremos a seguinte versão mais forte desse Teorema.

**Teorema 1.5** (Cantor-Bernstein-Schröder-Banach). *Suponha que  $f : A \rightarrow B$  e  $g : B \rightarrow A$  são funções injetoras. Então existem partições  $A = A_1 \cup A_2$  e  $B = B_1 \cup B_2$  tais que  $f(A_1) = B_1$  e  $g(B_2) = A_1$ .*

*Demonstração.* Dado  $x \in A$ , consideramos a seguinte sequência

$$x, g^{-1}(x), f^{-1}(g^{-1}(x)), g^{-1}(f^{-1}(g^{-1}(x))), \dots \quad (*)$$

Pomos  $A_1 = \{x \in A : (*) \text{ tem número finito e ímpar de elementos}\}$ . Por exemplo, se  $x \notin \mathcal{I}(g)$ , então  $(*)$  tem um único elemento e  $x \in A_1$ . Pomos  $A_2 = A \setminus A_1$  e  $B_1 = f(A_1)$ .

Pomos também  $B_2 = g^{-1}(A_2) = \{b \in B : g(b) \in A_2\}$ .

Afirmamos agora que valem os seguinte fatos.

- $B = B_1 \cup B_2$ .

*Demonstração.* Fixe  $y \in B$  e considere  $x = g(y)$ . Se  $x \in A_2$ , então  $y \in B_2$ . Suponha então que  $x \in A_1$ . Pela definição de  $A_1$ , existe  $x' = f^{-1}(y) = f^{-1}(g^{-1}(x))$ . Novamente pela definição de  $A_1$  devemos ter  $x' \in A_1$ . Mas então  $y = f(x')$  com  $x' \in A_1$  e, portanto,  $y \in f(A_1) = B_1$ .  $\square$

- $B_1 \cap B_2 = \emptyset$ .

*Demonstração.* Suponha que exista  $y \in B_1 \cap B_2$ . Então existe  $x' \in A_1$  tal que  $f(x') = y$  e existe  $x \in A_2$  tal que  $g(y) = x$ . Considere agora a sequência  $(*)$  para o elemento  $x$ . Vemos que  $x \in A_1$ , contradição.  $\square$

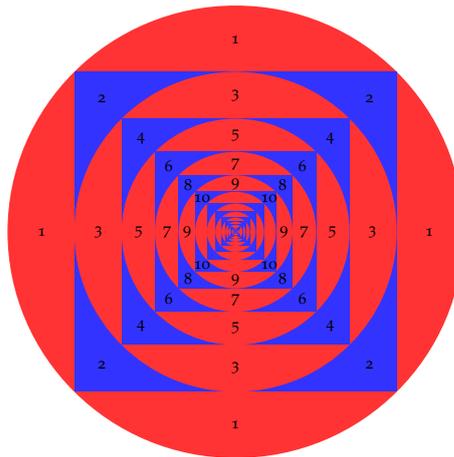
- $g(B_2) = A_2$ .

*Demonstração.* Fixe  $x \in A_2$ . Por definição de  $A_1$ , existe  $y$  tal que  $g(y) = x$ . Claramente, como  $B_2 = g^{-1}(A_2)$  e  $x \in A_2$  temos que  $y \in B_2$ . Segue que  $g(B_2) = A_2$ .  $\square$

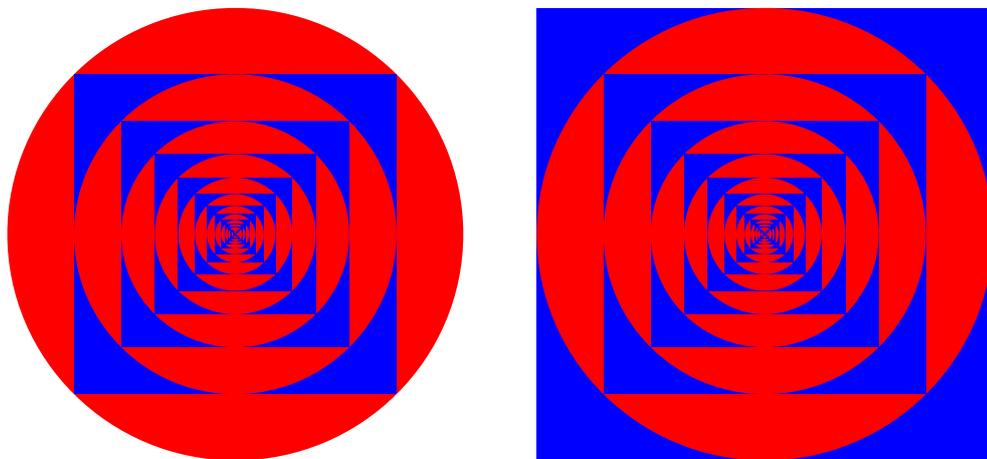
Segue dos três fatos acima que  $A_1, A_2, B_1, B_2$  são conjuntos como no enunciado do Teorema.  $\square$

*Solução Exercício 1.3.* Sejam  $A$  e  $B$ , respectivamente, o disco e o quadrado do enunciado. Defina  $f : A \rightarrow B$  como a função identidade e  $g : B \rightarrow A$ ,  $(x, y) \mapsto (\frac{x}{\sqrt{2}}, \frac{y}{\sqrt{2}})$ . Como  $A \cong f(A)$  e  $B \cong g(B)$ , ao aplicar o Teorema 1.5 obtemos partições  $A = A_1 \cup A_2$  e  $B = B_1 \cup B_2$  como desejado.

Na figura do disco  $A$  abaixo, a cor vermelha corresponde ao conjunto  $A_1$  e a azul ao conjunto  $A_2$ . Os números indicam o tamanho da sequência  $(*)$  para os pontos daquela região do conjunto  $A$ .



A figura abaixo mostra ambos os conjuntos  $A$  e  $B$  e suas respectivas partições.



□

## 2 Polígonos equidecomponíveis geometricamente

◇ ◇ ◇ *Aula 3 (08 de Março) — Victor Luiz Martins de Sousa* ◇ ◇ ◇

Dizemos que dois polígonos  $A, B \subset \mathbb{R}^2$  são *equidecomponíveis no sentido geométrico* se

1. Existem polígonos  $A_1, \dots, A_n$ , com interiores dois-a-dois disjuntos, satisfazendo  $A = \bigcup_{i=1}^n A_i$ .
2. Existem polígonos  $B_1, \dots, B_n$ , com interiores dois-a-dois disjuntos, satisfazendo  $B = \bigcup_{i=1}^n B_i$ .
3. Para todo  $1 \leq i \leq n$ , temos  $A_i \cong B_i$ , isto é, existe uma isometria entre  $A_i$  e  $B_i$ .

Escrevemos  $A \stackrel{\cong}{\equiv} B$  para denotar que  $A$  e  $B$  são equidecomponíveis (no sentido geométrico).

### 2.1 Teorema de Wallace-Bolyai-Gerwien

Sejam  $A$  e  $B$  dois polígonos. Uma condição trivialmente necessária para que  $A \stackrel{\cong}{\equiv} B$  é que  $A$  e  $B$  tenham a mesma área. Provaremos nesta seção o resultado a seguir, que mostra que esta condição é também suficiente.

**Teorema 2.1** (Wallace-Bolyai-Gerwien). *Se  $A$  e  $B$  são polígonos de mesma área, então  $A \stackrel{\cong}{\equiv} B$ .*

**Lema 2.2.** *Se  $A \stackrel{\cong}{\equiv} B$  e  $B \stackrel{\cong}{\equiv} C$ , então  $A \stackrel{\cong}{\equiv} C$ .*

*Demonstração.* A idéia da prova consiste em sobrepor a decomposição de  $B$  que mostra que  $A$  e  $B$  são equidecomponíveis com a decomposição de  $B$  que mostra que  $B$  e  $C$  são equidecomponíveis.

Como  $A \stackrel{\cong}{\equiv} B$ , então existem decomposições (i.e. partições em polígonos com interiores dois-a-dois disjuntos)  $A = \cup_{i=1}^n A_i$  e  $B = \cup_{i=1}^n B_i$  e isometrias  $f_i : A_i \rightarrow B_i$  (para todo  $1 \leq i \leq n$ ). Como  $B \stackrel{\cong}{\equiv} C$ , também existem decomposições  $B = \cup_{j=1}^m B'_j$  e  $C = \cup_{j=1}^m C_j$  e isometrias  $g_j : B'_j \rightarrow C_j$  (para todo  $1 \leq j \leq m$ ).

Como  $\cup_{i=1}^n \cup_{j=1}^m (B_i \cap B'_j)$  é uma decomposição de  $B$ , podemos definir decomposições  $A = \cup A_{i,j}$  e  $B = \cup B_{i,j}$  onde

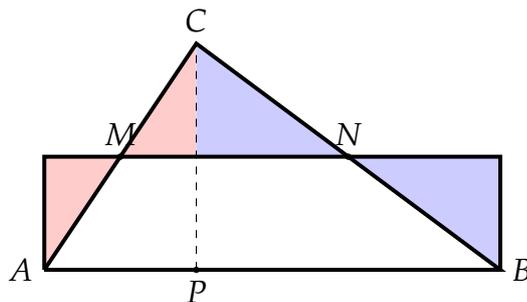
$$A_{i,j} = f_i^{-1}(B_i \cap B'_j) \quad \text{e} \quad C_{i,j} = g_j(B_i \cap B'_j).$$

Ademais a função  $h_{i,j} = (f_i \circ g_j)$  atesta que  $A_{i,j}$  é isométrico a  $C_{i,j}$  para todo  $1 \leq i \leq n$ ,  $1 \leq j \leq m$ .

Por fim, notamos que podemos supor que todos os  $A_i$  e  $B_i$  são polígonos *convexos*. Daí segue que  $A_{i,j}$  e  $B_{i,j}$  também são polígonos.  $\square$

**Lema 2.3.** *Para todo triângulo  $T$ , existe um retângulo  $R$  tal que  $T \stackrel{\cong}{\equiv} R$ .*

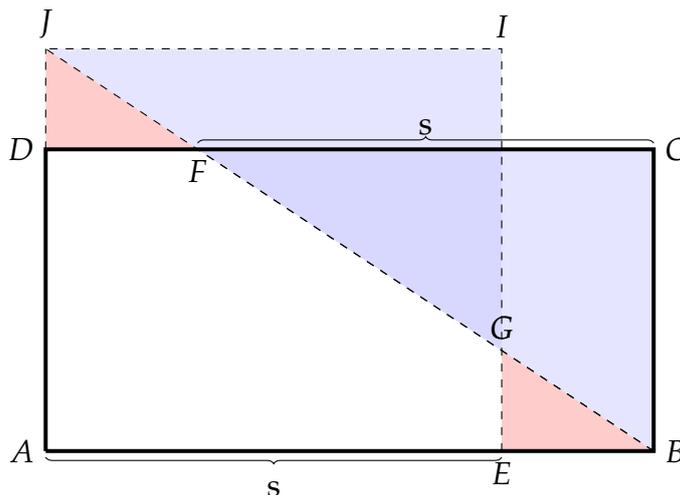
*Demonstração.* Seja  $T = ABC$  e suponha sem perda de generalidade que  $\hat{C}AB$  e  $\hat{A}BC$  são ângulos agudos. Sejam  $M, N$  os pontos médios dos segmentos  $CA, BC$ , respectivamente e seja  $P$  a projeção do ponto  $C$  sobre o segmento  $AB$ . A figura abaixo mostra que  $ABC$  é equidecomponível a um retângulo.



$\square$

**Lema 2.4.** *Se  $ABCD$  é um retângulo e  $S$  é um segmento de reta, então  $ABCD \stackrel{\cong}{\equiv} A'B'C'D'$ , onde  $A'B'$  tem o mesmo comprimento que  $S$ .*

*Demonstração.* Seja  $s = |S|$  o comprimento do segmento  $S$ . Podemos supor sem perda de generalidade que  $s < |AB| < 2s$ .



Como  $\triangle JDF \cong \triangle EBG$  e  $\triangle JIG \cong \triangle FCB$ , temos  $ABCD \stackrel{\cong}{\cong} AEIJ$ . □

*Prova do Teorema 2.1.* Seja  $S$  um segmento de comprimento  $s$  arbitrário. Fixe uma triangulação  $A = \cup_{i=1}^n T_i$ . Pelos Lemas 2.2 e 2.3, cada triângulo  $T_i$  é equidecomponível a um retângulo de base  $s$  e, portanto,  $A$  é equidecomponível a um retângulo  $R_A$  de base  $s$ .

Analogamente,  $B$  também é equidecomponível a um retângulo  $R_B$  de base  $s$ . Mas como  $A$  e  $B$  têm a mesma área devemos ter  $R_A \cong R_B$ . Segue que  $A \stackrel{\cong}{\cong} B$  (Lema 2.2). □

**Exercício 2.5.** Sejam  $X, Y \subseteq \mathbb{R}^2$  polígonos. Então  $X \stackrel{\cong}{\cong} Y$  se, e somente se,  $X \equiv Y$ .

### 3 Paradoxo de Banach-Tarski (cont)

◇ ◇ ◇ Aula 4 (15 de Março) — Yoshiharu Kohayakawa ◇ ◇ ◇

#### 3.1 Medidas Invariantes

Uma *medida* é uma função  $\mathcal{C} \rightarrow \mathbb{R}$ , com  $\mathcal{C} \subseteq X$ , que satisfaz as seguintes propriedades:

1. *não-negatividade:*  $m(A) \geq 0$ , para todo  $A \in \mathcal{C}$ .
2. *aditividade:*  $m(A \cup B) = m(A) + m(B)$ , para todo  $A, B \in \mathcal{C}$  com  $A \cap B = \emptyset$ .

Aqui, consideraremos sempre o caso em que  $X = \mathbb{R}^n$  para algum  $n > 0$ .

Uma medida  $m$  é *invariante por movimento rígido* se  $m(A) = m(\alpha(A))$  para toda função  $\alpha : \mathbb{R}^n \rightarrow \mathbb{R}^n$  que preserva distância e para todo  $A \in \mathcal{C}$  tal que  $\alpha(A) \in \mathcal{C}$ . Uma medida  $m$  é *normalizada* se  $m(Q_n) = 1$ , onde  $Q_n = [0, 1]^n \subset \mathbb{R}^n$  e  $Q_n \in \mathcal{C}$ .

**Fato 3.1.** Seja  $X = \mathbb{R}^2$  e  $\mathcal{C}$  o conjunto das regiões poligonais. Então a função  $m : \mathcal{C} \rightarrow \mathbb{R}$  tal que  $m(A) = \text{área de } A$  é uma medida normalizada e invariante por movimento rígido.

A medida de Borel-Lebesgue é uma medida  $m : \mathcal{C} \rightarrow \mathbb{R}$  sobre um conjunto  $\mathcal{C}$  que estende o conjunto de regiões poligonais, composto pelos chamados “conjuntos borelianos”. A medida  $m$  assim obtida é  $\sigma$ -aditiva isto é para qualquer sequência

$A_1, A_2, \dots \in \mathcal{C}$  de conjuntos dois-a-dois disjuntos vale que

$$m\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} m(A_i).$$

Uma pergunta natural é se não existe uma medida  $\sigma$ -aditiva, normalizada e invariante por movimento rígido com  $\mathcal{C} = \mathcal{P}(\mathbb{R}^2)$  (ou, mais genericamente com  $\mathcal{C} = \mathcal{P}(\mathbb{R}^n)$ ). O seguinte resultado responde negativamente tal pergunta.

**Teorema 3.2** (Vitali). *Seja  $m : \mathcal{P}(\mathbb{R}^n) \rightarrow \mathbb{R}$  uma medida definida sobre todos os subconjuntos de  $\mathbb{R}^n$ . Se  $m$  é invariante por translações e  $m(Q_n) = 1$ , então  $m$  não pode ser  $\sigma$ -aditiva.*

Para provar o resultado acima, precisaremos do seguinte Lema.

**Lema 3.3.** *Existe um conjunto  $A \subset \mathbb{R}^n$  satisfazendo as seguintes propriedades:*

1.  $Q_n$  contém uma infinidade de cópias transladadas de  $A$ , duas-a-duas disjuntas.
2.  $\mathbb{R}^n$  pode ser coberto com uma quantidade enumerável de cópias transladadas de  $A$ .

*Demonstração.* Faremos o caso  $n = 1$ . Se  $x, y \in \mathbb{R}$  são tais que  $x - y \in \mathcal{Q}$ , dizemos que  $x \sim y$ . Não é difícil verificar que  $\sim$  é uma relação de equivalência sobre  $\mathbb{R}$ . Para cada classe de equivalência de  $\sim$  escolha<sup>1</sup> um representante no intervalo  $[0, \frac{1}{2}]$ . Definimos  $A$  como sendo o conjunto desses representantes.

Para mostrar a propriedade (1) tomamos

$$A_i = A + \frac{1}{i+1} \quad i = 1, 2, \dots$$

Esses conjuntos  $A_i$  são dois-a-dois disjuntos pois se  $x, y \in \mathcal{Q}$  e  $x \neq y$ , então  $(A+x) \cap (A+y) = \emptyset$ . De fato, se  $z \in A+x$  e  $z \in A+y$ , então temos, respectivamente,  $z-x \in A$  e  $z-y \in A$ . Entretanto  $(z-x) - (z-y) = y-x \in \mathcal{Q}$  e portanto  $z-y$  e  $z-x$  pertencem à mesma classe de equivalência. Segue da definição de  $A$  que  $z-y = z-x$ , isto é,  $y = x$ .

Para mostrar a propriedade (2), basta notar que  $\mathbb{R} \subseteq \bigcup_{r \in \mathcal{Q}} A+r$ . De fato, fixe  $x \in \mathbb{R}$  e seja  $x_0 \in A$ , o representante da classe  $x$  em  $A$ , isto é, o elemento tal que  $x \sim x_0$ . Logo devemos ter  $x \in A+r$ , com  $r = x - x_0 \in \mathcal{Q}$ .  $\square$

*Prova do Teorema 3.2.* Suponha que exista uma medida  $m : \mathcal{P}(\mathbb{R}^n) \rightarrow \mathbb{R}$  como no enunciado. Note que toda medida  $m$  é monótona, isto é, se  $X \subseteq Y$ , então  $m(X) \leq m(Y)$  (pois  $m(Y) = m(X) + m(Y \setminus X)$ ). Tome  $A$  como no Lema anterior e sejam  $A_1, A_2, \dots \subseteq Q_n$  conjuntos transladados de  $A$ , dois-a-dois disjuntos e seja  $B = \bigcup_{i=1}^{\infty} A_i \subseteq Q_n$ . Temos

$$\sum_{i=1}^{\infty} m(A) = \sum_{i=1}^{\infty} m(A_i) = m(B) \leq m(Q_n) = 1,$$

da onde segue que  $m(A) = 0$ .

Por outro lado, sejam  $A_1, A_2, \dots \subset Q_n$  conjuntos transladados de  $A$ , dois-a-dois disjuntos tais que  $\mathbb{R} \subseteq \bigcup_{i=1}^{\infty} A_i$ . Então

$$1 = m(Q_n) \leq m(\mathbb{R}^n) \leq m\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} m(A_i) = 0,$$

<sup>1</sup>note que esse passo faz uso do Axioma de Escolha

um absurdo. □

*Pergunta:* Existe medida  $m : \mathcal{P}(\mathbb{R}^n) \rightarrow \mathbb{R}$  invariante por translação e normalizada no  $\mathbb{R}^n$ ? Note que o Teorema 3.2 mostra que não podemos exigir  $\sigma$ -aditividade.

Banach mostrou que existe uma tal medida para os casos  $n = 1$  e  $n = 2$ . Por outro lado, Hausdorff provou que tal medida não existe se  $n \geq 3$ . Seja  $S = S^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$ .

**Teorema 3.4** (Paradoxo de Hausdorff). *Existem decomposições  $S = A_1 \cup A_2 \cup C_1$  e  $S = A_3 \cup A_4 \cup A_5 \cup C_2$  tais que os conjuntos  $A_i$  são todos congruentes e  $C_1$  e  $C_2$  são enumeráveis.*

O resultado acima atesta que não existe medida  $m : \mathcal{P}(S) \rightarrow \mathbb{R}$  invariante por translação tal que  $m(S) = 1$ . De fato, como  $C_1$  e  $C_2$  são enumeráveis, uma tal medida  $m$  deve satisfazer  $m(C_1) = m(C_2) = 0$ . Portanto, as decomposições do Paradoxo de Hausdorff implicam, respectivamente, que  $m(A_1) = \frac{1}{2}$  e  $m(A_1) = \frac{1}{3}$ , absurdo.

Vamos provar o resultado a seguir que, apesar de ser mais fraco que o Paradoxo de Hausdorff, é suficiente para mostrar que não existe medida  $m : \mathcal{P}(S) \rightarrow \mathbb{R}$  invariante por translação tal que  $m(S) = 1$ .

**Teorema 3.5.** *Existe  $A \subseteq S = S^2 \subset \mathbb{R}^3$  tal que*

1. *Existe uma infinidade de conjuntos  $D_1, D_2, \dots \subset S$ , todos congruentes a  $A$  e dois-a-dois disjuntos.*
2. *Existem conjuntos  $A_1, A_2, A_3, A_4$ , todos congruentes a  $A$ , tais que  $S = A_1 \cup A_2 \cup A_3 \cup A_4$ .*

◇ ◇ ◇

Aula 5 (29 de Março) — Yoshiharu Kohayakawa

◇ ◇ ◇

### 3.2 Rotações em $R_3$

Seja  $SO(3)$  o conjunto das rotações sobre a origem em  $\mathbb{R}^3$ . É um fato de que  $SO(3)$  é um grupo. Para verificar esse fato, note que uma rotação pode ser representada por uma matriz ortogonal  $R$  que preserva a orientação dos eixos, isto é, que satisfaz a propriedade especial<sup>2</sup>  $\det(R) = 1$ .

Fixados dois elementos  $\phi, \psi \in SO(3)$ , consideramos sequências (finitas) da forma:

$$\phi^{a_1} \psi^{a_2} \phi^{a_3} \dots \in SO(3), \quad (1)$$

com  $a_2, a_3, \dots \neq 0$ .

**Lema 3.6.** *Suponha que  $\phi$  e  $\psi$  são rotações em torno dos eixos  $x$  e  $y$  (respectivamente) por um ângulo  $\alpha$  tal que  $\cos \alpha$  é transcendental. Então  $\phi$  e  $\psi$  geram um grupo de dois geradores, isto é,  $\phi^0$  é a única expressão da forma **1** que é igual à identidade.*

### 3.3 Prova da Versão Fraca do Paradoxo de Hausdorff (Teorema 3.5)

No que se segue, fixamos um par  $\phi, \psi$  como no Lema acima. Seja  $G$  o grupo formado por elementos da forma **1**.

<sup>2</sup>A notação  $SO$  vem de *special orthogonal group*

Dados  $x, y \in S$ , dizemos que  $x \sim y$  se existe  $g \in G$  tal que  $y = g(x)$ . Então  $\sim$  é uma relação de equivalência cujas classes de equivalência são denominadas *órbitas*. Defina

$$C = \{x \in S : \text{existe } \chi \in G, \chi \neq e, \text{ tal que } \chi(x) = x\}$$

o conjunto dos pontos em  $S$  que são pontos fixos de algum elemento de  $G$  diferente da identidade. Note que se  $x \sim y$  e  $x \in C$ , então  $y \in C$ . De fato, se existem  $\chi, g \in G, \chi \neq e$  tal que  $\chi(x) = x$  e  $y = g(x)$ , então  $y = (g\chi g^{-1})(x)$  e  $(g\chi g^{-1}) \neq e$ . Logo  $C$  é uma união de órbitas.

Fixe um conjunto  $H \subset S$  composto por um representante de cada órbita em  $S \setminus C$ . Seja  $U$  o conjunto de elementos da forma  $\mathbf{1}$  com  $a_1 \neq 0$ . Definimos

$$A = \{\chi(x) : x \in H, \chi \in U\}.$$

Vamos provar as asserções do Teorema.

1. Tomamos  $D_i = \psi^{(i)}(A)$  para  $i = 1, 2, \dots$

**Afirmção 3.7.** Se  $n \neq m$ , então  $\psi^{(n)}(A) \cap \psi^{(m)}(A) = \emptyset$

*Demonstração.* Suponha que existe  $y \in \psi^{(n)}(A) \cap \psi^{(m)}(A)$ . Então devem existir  $\chi_1, \chi_2 \in U$  e  $x_1, x_2 \in H$  tais que  $y = \psi^{(n)}(\chi_1(x_1))$  e  $y = \psi^{(m)}(\chi_2(x_2))$ . Então temos  $\chi_1(x_1) = \psi^{(m-n)}(\chi_2(x_2))$ , o que implica  $x_1 = (\chi_1^{-1}\psi^{(m-n)}\chi_2)(x_2)$ . Mas como  $H$  possui um única representante de cada órbita, devemos ter  $\chi_1^{-1}\psi^{(m-n)}\chi_2 = e$ , isto é,  $\chi_2 = \psi^{(n-m)}\chi_1$ . Segue da definição de  $U$  que  $n - m = 0$ , ou seja,  $n = m$ .  $\square$

2. Mostraremos que existe  $\rho \in SO(3)$  tal que  $S = A \cup \phi(A) \cup \rho^{-1}(A) \cup \rho^{-1}(\phi(A))$ .

**Afirmção 3.8.**  $S \setminus C \subseteq A \cup \phi(A)$ .

*Demonstração.* Seja  $y \in S \setminus C$  e seja  $x \in H$  tal que  $y \sim x$ . Logo existe  $\chi \in G$  tal que  $y = \chi(x)$ . Se  $\phi \in U$ , então  $y \in A$ . Por outro lado, se  $\chi = \psi^{a_2}\phi^{a_3}\dots$ , então  $\phi^{-1}\chi \in U$  e, portanto,  $(\phi^{-1}\chi)(x) \in A$ . Mas como  $y = \phi(\phi^{-1}(\chi(x)))$ , segue que  $y \in \phi(A)$ .  $\square$

Como  $C$  é enumerável, existe  $\rho \in SO(3)$  tal que  $C \cap \rho(C) = \emptyset$ . De fato, existe apenas um número enumerável de rotações que *não* satisfazem tal igualdade, a saber no máximo duas para cada par de elementos em  $C$ . Usando a afirmação acima, concluímos que

$$\rho(C) \subseteq S \setminus C \subseteq A \cup \phi(A)$$

e, portanto, que

$$C = \rho^{-1}(A) \cup \rho^{-1}(\phi(A)).$$

◇ ◇ ◇

Aula 6 (05 de Abril) — Yoshiharu Kohayakawa

◇ ◇ ◇

### 3.4 Prova do Paradoxo de Banach-Tarski

Nesta seção provaremos o Paradoxo de Banach-Tarski (Teorema 1.2)

**Lema 3.9.** Sejam  $X, Y \subseteq \mathbb{R}^n$ . Suponha que existam  $V \subseteq X$  e  $U \subseteq Y$  tais que  $X \equiv_k U$  e  $V \equiv_m Y$ . Então  $X \equiv_{k+m} Y$ .

*Demonstração.* Como  $X \equiv_k U$ , existem conjuntos  $X_1, \dots, X_k$  e  $U_1, \dots, U_k$  tais que  $X = \bigcup_{i=1}^k X_i$ ,  $U = \bigcup_{i=1}^k U_i$  e para todo  $1 \leq i \leq k$  temos  $X_i \cong U_i$ , isto é, existem isometrias  $\phi_i : X_i \rightarrow U_i$ .

Como  $Y \equiv_k V$ , existem conjuntos  $Y_1, \dots, Y_m$  e  $V_1, \dots, V_m$  tais que  $Y = \bigcup_{j=1}^m Y_j$ ,  $V = \bigcup_{j=1}^m V_j$  e para todo  $1 \leq j \leq m$  temos  $Y_j \cong V_j$ , isto é, existem isometrias  $\psi_j : Y_j \rightarrow V_j$ .

Defina funções  $f : X \rightarrow U$  e  $g : Y \rightarrow V$  pondo  $f(x) = \phi_i(x)$  para todo  $x \in X_i$  e  $g(y) = \psi_j(y)$  para todo  $y \in Y_j$ . Note que  $f$  e  $g$  são funções injetoras. Pelo Teorema 1.5, existem  $A_1, A_2 \subseteq X$  e  $B_1, B_2 \subseteq Y$  tais que  $X = A_1 \cup A_2$  e  $Y = B_1 \cup B_2$  e satisfazendo  $f(A_1) = B_1$  e  $g(B_2) = A_2$ . Então podemos escrever

$$X = A_1 \cup A_2 = A_1 \cup g(B_2) = \bigcup_{i=1}^k (X_i \cap A_1) \cup \bigcup_{j=1}^m \psi_j(Y_j \cap B_2).$$

De maneira análoga podemos escrever

$$Y = \bigcup_{i=1}^k \phi_i(X_i \cap A_1) \cup \bigcup_{j=1}^m Y_j \cap B_2.$$

Essas duas decomposições de  $X$  e  $Y$  atestam que  $X \equiv_{k+m} Y$ . □

*Prova do Teorema 1.2 (Paradoxo de Banach-Tarski).* Dado  $x \in S$ , definimos  $r_x = (0, x] = \{\lambda x : 0 < \lambda \leq 1\}$  como o segmento de 0 a  $x$ . Dado  $C \subseteq S$ , também definimos

$$C^* = \bigcup_{x \in C} r_x.$$

Sejam  $A_1, A_2, A_3, A_4 \subset S$  e  $D_1, D_2, \dots \subset S$  como no Teorema 3.5. Definimos os seguintes conjuntos disjuntos:  $C_1 = A_1$ ,  $C_2 = A_2 \setminus A_1$ ,  $C_3 = A_3 \setminus (A_1 \cup A_2)$  e  $C_4 = A_4 \setminus (A_1 \cup A_2 \cup A_3)$ . Temos  $S = C_1 \cup \dots \cup C_4$ . Assim, supondo sem perda de generalidade que  $S$  é a casca esférica de  $B_1$ , temos a decomposição

$$B_1 \setminus \{0\} = \bigcup_{i=1}^4 C_i^*.$$

Seja  $\gamma : B_1 \rightarrow B_2$  uma isometria. Então podemos decompor  $B_1 \cup B_2$  em 9 partes como a seguir:

$$B_1 \cup B_2 = (C_1^* \cup \{0\}) \cup \bigcup_{i=2}^4 C_i^* \cup \bigcup_{i=1}^4 \gamma(C_i^*) \cup \{\gamma(0)\}.$$

Note que  $C_1^* \cup \{0\} \cong D_1^* \cup \{0\}$ . Ademais,  $C_i^*$  é congruente a uma parte de  $D_i^*$  ( $2 \leq i \leq 4$ ) Finalmente,  $\gamma(C_i^*)$  é congruente a uma parte de  $D_{i+4}^*$  ( $1 \leq i \leq 4$ ) e  $\{\gamma(0)\}$  é congruente a qualquer subconjunto unitário em  $D_9$ . Consequentemente existe  $V \subseteq B_1$  tal que  $B_1 \cup B_2 \equiv_9 V$ . Temos  $B_1 \equiv_1 B_1 \subset B_1 \cup B_2$ . Segue do Lema 3.9 que  $B_1 \equiv_{10} B_1 \cup B_2$ . □

É possível generalizar o Teorema 1.2 para dimensões maiores e para bolas com raios distintos.

**Teorema 3.10.** *Seja  $n \geq 3$  e  $X, Y \subset \mathbb{R}^n$  tais que  $X \supset B_1$  e  $Y \supset B_2$ , onde  $B_1$  e  $B_2$  são bolas de raio positivo. Então  $X \equiv Y$ .* □

## 4 Teorema de Dehn

◇ ◇ ◇

Aula 7 (12 de Abril) — Marcelo Soares Campos

◇ ◇ ◇

Dizemos que dois poliedros  $A, B \subseteq \mathbb{R}^3$  são equidecomponíveis (no sentido geométrico) se podemos escrever  $A = A_1 \cup \dots \cup A_n$ , para poliedros  $A_1, \dots, A_n$  de interiores dois-a-dois disjuntos, e  $B = B_1 \cup \dots \cup B_n$ , para poliedros  $B_1, \dots, B_n$  de interiores dois-a-dois disjuntos, de tal forma que  $A_i \cong B_i$  para todo  $1 \leq i \leq n$ . Neste caso escrevemos  $A \stackrel{g}{\cong} B$ .

Defina uma transformação linear  $f: \mathbb{R} \rightarrow \mathbb{R}$ , onde  $\mathbb{R}$  é visto como espaço vetorial sobre  $\mathcal{Q}$ , tal que  $f(\pi) = 0$  e  $f(\arccos(\frac{1}{3})) = 1$ . Para que a função  $f$  esteja bem definida, precisamos do lema a seguir, que mostra que  $\pi$  e  $\arccos(\frac{1}{3})$  são elementos linearmente independentes nesse espaço vetorial.

**Lema 4.1.**  $\frac{\arccos(\frac{1}{3})}{\pi}$  é irracional.

*Demonstração.* Suponha que  $\arccos(\frac{1}{3}) = \frac{m}{n}\pi$ , com  $m, n \in \mathbb{Z}$ . Aplicando a função  $\cos$  em ambos os lados obtemos

$$\cos(n \arccos(\frac{1}{3})) = \pm 1.$$

Usando a identidade

$$\cos(nx) = n \sum_{k=0}^n (-2)^k \frac{(n+k+1)!}{(n-k)!(2k)!} (1 - \cos(x))^k,$$

temos

$$\cos(n \arccos(\frac{1}{3})) = n \sum_{k=0}^n (-2)^k \frac{(n+k+1)!}{(n-k)!(2k)!} (1 - \frac{1}{3})^k.$$

Multiplicando ambos os lados por  $3^n$ :

$$\pm 3^n = 3^n \cos(n \arccos(\frac{1}{3})) = n \sum_{k=0}^n (-2)^k (n+k+1) \binom{n+k}{2k} 2^k 3^{n-k}.$$

Note que as parcelas  $1 \leq k \leq n-1$  do somatório acima são divisíveis por 3. Logo para que valha a igualdade acima, a última parcela ( $k = n$ ) também deve ser divisível por 3, o que implica 3 dividir uma potência de 2, um absurdo.  $\square$

Seja  $\mathcal{P}$  o conjunto de todos os poliedros em  $\mathbb{R}^3$ . Para todo poliedro  $P \in \mathcal{P}$  definimos  $E(P)$  como o conjunto das arestas de  $P$ . Dada uma aresta  $e$  de um poliedro, denotamos por  $\alpha_e$  o seu ângulo *diedral*, isto é, o ângulo entre as duas faces que definem  $e$ . Também denotamos o comprimento de uma aresta  $e$  por  $|e|$ .

Definimos o *invariante de Dehn* como a seguinte função:

$$\begin{aligned} \phi: \mathcal{P} &\rightarrow \mathbb{R} \\ P &\mapsto \sum_{e \in E(P)} |e| f(\alpha_e). \end{aligned}$$

**Lema 4.2.** Sejam  $P_1$  e  $P_2$  poliedros com interiores disjuntos tais que  $P = P_1 \cup P_2$  é também um poliedro. Então

$$\phi(P) = \phi(P_1) + \phi(P_2).$$

*Demonstração.* Considere o plano  $h$  que corta  $P$  e dá origem aos poliedros  $P_1$  e  $P_2$ . Só precisamos analisar os casos de arestas  $e_1 \in P_1$  e  $e_2 \in P_2$  cuja interseção com  $h$  é não-nula (as demais correspondem a arestas em  $P$  de mesmo tamanho e mesmo ângulo diedral). Tais arestas podem ser geradas de três maneiras:

1. O plano  $h$  corta o interior de uma face de  $P$ , gerando novas arestas  $e_1 \in P_1$  e  $e_2 \in P_2$  (cujos interiores estão contidos no interior de uma face de  $P$ ). [FIGURA]

Temos  $|e_1| = |e_2|$  e  $\alpha_{e_1} + \alpha_{e_2} = \pi$ . Logo

$$|e_1|f(\alpha_{e_1}) + |e_2|f(\alpha_{e_2}) = |e_1|(f(\alpha_{e_1}) + f(\alpha_{e_2})) = |e_1|f(\alpha_{e_1} + \alpha_{e_2}) = 0.$$

2. O plano  $h$  intersecta uma aresta  $e \in P$  em exatamente um ponto, gerando novas arestas  $e_1 \in P_1$  e  $e_2 \in P_2$  de tamanhos diferentes mas com o mesmo ângulo diedral [FIGURA]

Temos  $|e_1| + |e_2| = |e|$  e  $\alpha_{e_1} = \alpha_{e_2} = \alpha_e$ . Logo

$$|e_1|f(\alpha_{e_1}) + |e_2|f(\alpha_{e_2}) = (|e_1| + |e_2|)f(\alpha_e) = |e|f(\alpha_e).$$

3. O plano  $h$  é tal que  $h \cap e = e$  para uma aresta  $e \in P$ , gerando arestas  $e_1 \in P_1$  e  $e_2 \in P_2$  de mesmo tamanho mas com ângulos diedrais distintos. [FIGURA]

Temos  $|e_1| = |e_2| = |e|$  e  $\alpha_{e_1} + \alpha_{e_2} = \alpha_e$ . Logo

$$|e_1|f(\alpha_{e_1}) + |e_2|f(\alpha_{e_2}) = |e|f(\alpha_{e_1} + \alpha_{e_2}) = |e|f(\alpha_e).$$

□

**Teorema 4.3** (Teorema de Dehn). *Seja  $C$  um cubo e  $T$  um tetraedro, ambos de volumes unitários. Então  $C$  e  $T$  não são equidecomponíveis.*

*Demonstração.* Primeiro note que

$$\phi(C) = \sum_{e \in E(C)} |e|f(\pi/2) = 0,$$

pois  $f(\pi/2) = f(\pi)/2 = 0$  e

$$\phi(T) = \sum_{e \in E(T)} |e|f(\arccos \frac{1}{3}) \neq 0,$$

uma vez que  $\arccos \frac{1}{3} = 1$ . Logo  $\phi(C) \neq \phi(T)$ .

Suponha que existam poliedros  $A_1, \dots, A_n$  com interiores dois-a-dois disjuntos e  $B_1, \dots, B_n$  com interiores dois-a-dois disjuntos tais que  $A_i \cong B_i$  ( $1 \leq i \leq n$ ) e  $C = \bigcup_{i=1}^n A_i$

e  $T = \bigcup_{i=1}^n B_i$ . Então

$$\begin{aligned}
 \phi(C) &= \phi\left(\bigcup_{i=1}^n A_i\right) \\
 &= \sum_{i=1}^n \phi(A_i) && \text{(Lemma 4.2 + indução)} \\
 &= \sum_{i=1}^n \phi(B_i) && \text{(pois } A_i \cong B_i) \\
 &= \phi\left(\bigcup_{i=1}^n B_i\right) && \text{(Lemma 4.2 + indução)} \\
 &= \phi(T),
 \end{aligned}$$

contradizendo o fato de que  $\phi(C) \neq \phi(T)$ . □

## 5 Ladrilhamento de Retângulos

◇ ◇ ◇

Aula 8 (12 de Abril) — Bruno Cavalari

◇ ◇ ◇

**Teorema 5.1.** *Seja  $x$  um número irracional positivo e  $R \subset \mathbb{R}^2$  um retângulo de lados 1 e  $x$ . Então é impossível ladrilhar  $R$  com um número finito de quadrados.*

*Demonstração.* Consideramos  $\mathbb{R}$  como um espaço vetorial  $V$  sobre  $\mathcal{Q}$ . Temos portanto que 1 e  $x$  são linearmente independentes em  $V$ .

Considere uma transformação linear  $f : \mathbb{R} \rightarrow \mathbb{R}$  tal que  $f(1) = 1$  e  $f(x) = -1$ . Se  $S$  é um retângulo de lados  $a$  e  $b$ , definimos  $v(S) = f(a)f(b)$ . Disto segue que  $v(R) = -1$  e que se  $Q$  é um quadrado, então  $v(Q) \geq 0$ .

Suponha que existe um ladrilhamento de  $R$  em um número finito de quadrados  $Q_1, Q_2, \dots, Q_\ell$ . Estenda os segmentos de reta que definem cada um dos quadrados  $Q_i$  de forma a obter um ladrilhamento de  $R$  em retângulos  $R_{1,1}, \dots, R_{n,m}$  tais que  $R_{i,j}$  tem base  $a_i$  e altura  $b_j$  ( $1 \leq i \leq n$ ,  $1 \leq j \leq m$ ). **[FIGURA]**

Temos

$$\begin{aligned}
 v(R) &= f(1)f(x) \\
 &= f\left(\sum_{i=1}^n a_i\right) f\left(\sum_{j=1}^m b_j\right) \\
 &= \sum_{i=1}^n f(a_i) \sum_{j=1}^m f(b_j) \\
 &= \sum_{i=1}^n \sum_{j=1}^m f(a_i)f(b_j) \\
 &= \sum_{i=1}^n \sum_{j=1}^m f(R_{i,j}).
 \end{aligned}$$

Pelo mesmo argumento, todo quadrado  $Q_k$  tem sua valuação igual à soma das valuações dos retângulos  $R_{i,j}$  contidos em  $Q_k$ , isto é,  $v(Q_k) = \sum_{R_{i,j} \subseteq Q_k} v(R_{i,j})$  (para todo  $1 \leq k \leq \ell$ ).

Mas como cada retângulo  $R_{i,j}$  está contido em exatamente um quadrado  $Q_k$ , segue que

$$-1 = v(R) = \sum_{i=1}^n \sum_{j=1}^m f(R_{i,j}) = \sum_{k=1}^{\ell} v(Q_k) \geq 0,$$

uma contradição. □

## 6 Valuações não-arquimedianas

◇ ◇ ◇

*Aula 9 (26 de Abril) — Bruno Cavalari*

◇ ◇ ◇

Seja  $K$  um corpo (geralmente tomaremos  $K = \mathbb{Q}$ ). Uma *valuação* é uma função  $v : K \rightarrow \mathbb{R}_{\geq 0}$  que admite as seguintes propriedades:

- i)  $v(0) = 0$ ;
- ii)  $v(xy) = v(x)v(y)$  para todo  $x, y \in K$ ;
- iii)  $v(x) + v(y) \leq \max\{v(x), v(y)\}$  para todo  $x, y \in K$ .

Note que para toda valuação  $v : K \rightarrow \mathbb{R}_{\geq 0}$  devemos ter  $v(1) = 1$  (pois  $v(1) = v(1)v(1)$ ),  $v(-1) = 1$  (pois  $v(1) = v(-1)v(-1)$ ) e, portanto,  $v(-a) = v(a)$ . Além disso, para todo  $0 \neq a \in K$ , temos  $v(a)v(a^{-1}) = v(aa^{-1}) = v(1) = 1$  e, portanto,  $v(a) = 1/v(a^{-1})$ .

**Proposição 6.1.** *Toda valuação  $v : K \rightarrow \mathbb{R}_{\geq 0}$  satisfaz a seguinte propriedade:*

$$iv) \text{ Se } v(x) \neq v(y), \text{ então } v(x+y) = \max\{v(x), v(y)\}.$$

*Demonstração.* Suponha sem perda de generalidade que  $v(y) > v(x)$ . Temos

$$v(y) = v(x+y-x) \leq \max\{v(x+y), v(x)\}.$$

Se  $\max\{v(x+y), v(x)\} = v(x)$ , teríamos  $v(y) \leq v(x)$  um absurdo. Logo  $\max\{v(x+y), v(x)\} = v(x+y)$  e, portanto

$$v(y) \leq v(x+y) \leq v(y),$$

da onde segue que  $v(y) = v(x+y)$ . □

**Proposição 6.2.** *Sejam  $a, b_1, \dots, b_n \in K$  tais que  $v(a) > v(b_i)$  para todo  $1 \leq i \leq n$ . Então  $v(a \pm b_1 \pm \dots \pm b_n) = v(a)$*

*Demonstração.* Segue da Proposição 6.2, por indução em  $n$ . □

Seja  $p$  um número primo. Note que qualquer racional  $0 \neq r \in \mathbb{Q}$  pode ser escrito da forma

$$r = p^k \frac{a}{b},$$

com  $a, b, k \in \mathbb{Z}$ ,  $b > 0$  e  $\text{mdc}(a, p) = \text{mdc}(b, p) = 1$ .

Definimos a *norma  $p$ -ádica*  $|\cdot|_p : \mathbb{Q} \rightarrow \mathbb{R}_{\geq 0}$  da seguinte forma:

1. Se  $0 \neq r = p^k \frac{a}{b}$ , então  $|r|_p = p^{-k}$ .

2. Se  $r = 0$ , então  $|r|_p = 0$ .

**Proposição 6.3.** *A norma  $p$ -ádica  $|\cdot|_p : \mathcal{Q} \rightarrow \mathbb{R}_{\geq 0}$  é uma valuação arquimediana.*

*Demonstração.* Basta provar o item iii) da definição de valuação arquimediana. Seja  $r = p^k \frac{a}{b}$  e  $s = p^l \frac{c}{d}$  e suponha que  $k \geq l$ . Note que  $|r|_p = p^{-k} \geq p^{-l} = |s|_p$ . Temos

$$|r + s|_p = \left| p^k \frac{a}{b} + p^l \frac{c}{d} \right|_p = \left| p^l \left( p^{k-l} \frac{a}{b} + \frac{c}{d} \right) \right|_p = \left| p^l \frac{p^{k-l} ad + bc}{bd} \right|_p.$$

Como  $\text{mdc}(bd, p) = 1$ , então  $\left| \frac{p^{k-l} ad + bc}{bd} \right|_p \leq 1$ . Logo  $|r + s|_p \leq p^{-l} = |s|_p = \max\{|r|_p, |s|_p\}$ .  $\square$

Para o caso  $p = 2$ , temos  $|\frac{1}{2}|_2 = 2 > 1$ . É um fato que a valuação  $|\cdot|_2 : \mathcal{Q} \rightarrow \mathbb{R}_{\geq 0}$  pode ser estendida para o conjunto dos números reais, de forma a obter uma valuação  $v : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  satisfazendo  $v(\frac{1}{2}) > 1$ . Consideraremos essa tal valuação no restante na próxima seção. Posteriormente, mostraremos uma outra forma de se obter uma valuação dos reais que satisfaz  $v(\frac{1}{2}) > 1$  (ver Teorema 6.8).

## 6.1 Teorema de Mosky

O objetivo desta seção é provar o seguinte resultado.

**Teorema 6.4.** *Toda dissecção do quadrado  $S = [0, 1]^2$  em triângulos tem um número ímpar de triângulos.*

A seguir consideramos uma valuação  $v : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}$  satisfazendo  $v(\frac{1}{2}) > 1$  (ver fim da seção anterior) e definimos a seguinte coloração do plano:

$$c : \mathbb{R}^2 \rightarrow \{\text{azul, verde, vermelho}\}$$

$$(x, y) \mapsto \begin{cases} \text{azul} & , \text{ se } v(x) \geq v(y) \text{ e } v(x) \geq v(1); \\ \text{verde} & , \text{ se } v(y) > v(x) \text{ e } v(y) \geq 1; \\ \text{vermelho} & , \text{ se } v(1) > v(y) \text{ e } v(1) > v(x). \end{cases}$$

**Lema 6.5.** *Sejam  $p_b = (x_b, y_b)$ ,  $p_g = (x_g, y_g)$  e  $p_r = (x_r, y_r)$  pontos azul, verde e vermelho respectivamente. Então*

$$\det \begin{pmatrix} x_b & y_b & 1 \\ x_g & y_g & 1 \\ x_r & y_r & 1 \end{pmatrix}$$

*tem valuação maior ou igual a 1.*

*Demonstração.* Seja  $z = x_b y_g 1$ . Temos

$$v(z) = v(x_b) v(y_g) v(1) \geq 1.$$

Qualquer outra parcela do determinante pode ser escrita da forma  $z' = \pm ABC$ , com  $A \in \{x_b, y_b, 1\}$ ,  $B \in \{x_g, y_g, 1\}$  e  $C \in \{x_r, y_r, 1\}$ . Segue que  $v(A) \leq x_b$ ,  $v(B) \leq y_g$  e  $v(C) \leq 1$ . Note também que alguma dessas igualdade deve ser estrita, caso contrário as entradas na matriz correspondentes a  $A, B, C$  estariam, todas, acima da diagonal. Logo

$v(z') = v(A)v(B)v(C) < v(z)$ . Segue da Proposição 6.2 que a valuação do determinante é maior ou igual a  $v(z)$ .  $\square$

Chamaremos de *triângulo arcoíris* um triângulo cujos vértices têm cores duas-a-duas distintas.

**Lema 6.6.** *Toda reta no plano contém pontos de no máximo duas cores e a área de um triângulo arcoíris não pode ser  $\frac{1}{n}$ , com  $n$  ímpar.*

*Demonstração.* Quaisquer pontos  $(x_b, y_b)$ ,  $(x_g, y_g)$  e  $(x_r, y_r)$  colineares são tais que o determinante do Lema 6.5 é nulo e portanto não podem ter cores duas-a-duas distintas.

Sejam  $(x_b, y_b)$ ,  $(x_g, y_g)$  e  $(x_r, y_r)$  vértices (com cores azul, verde e vermelho, respectivamente) de um triângulo de área  $a$  e  $d$  o determinante do Lema 6.5. Temos

$$v(a) = v\left(\frac{1}{2}d\right) = v\left(\frac{1}{2}\right)v(d) > v(d) \geq 1.$$

Por outro lado, mostramos que  $v\left(\frac{1}{n}\right) = 1$  se  $n$  é ímpar.  $\square$

**Teorema 6.7.** *Toda dissecção do quadrado  $S = [0, 1]^2$  tem um número ímpar de triângulos arcoíris.*

*Demonstração.* Considere os segmentos formados pela dissecção na borda do quadrado. Afirmamos que há um número *ímpar* de segmentos azul-vermelho (isto é, cujas extremidades são coloridas, respectivamente, com as cores azul e vermelho). De fato, como o ponto  $(0, 1)$  é da cor verde, então as retas  $(0, 0)(0, 1)$  e  $(0, 1)(1, 1)$  não podem ter segmentos azul-vermelho, caso contrário teriam pontos das três cores. Além disso, a reta  $(1, 0)(1, 1)$  não pode conter pontos vermelhos (pois  $v(x) = v(1) = 1$ ). Logo os únicos segmentos azul-vermelho na borda de  $S$  são os da reta  $(0, 0)(1, 0)$ . Como o ponto  $(0, 0)$  tem cor vermelha e o ponto  $(1, 0)$ , deve haver um número ímpar de segmentos azul-vermelho nessa reta, caso contrário ela teria pontos das três cores.

Agora, note que todo triângulo arcoíris tem um número *ímpar* de segmentos azul-vermelho. De fato, tais segmentos só podem aparecer no lado do triângulo cujos vértices têm cores azul e vermelho, respectivamente. Novamente, deve haver um número ímpar de segmentos nesse lado do triângulo, caso contrário ele conteria pontos das três cores.

Note também que todo triângulo que não é arcoíris tem um número *par* de segmentos azul-vermelho. De fato, se um triângulo não é arcoíris, então ele deve ter um número par (ou zero ou dois) de lados cujos vértices têm cores azul e vermelho, respectivamente. Novamente, cada um desses lados têm um número ímpar de segmentos azul-vermelho, totalizando um número par de tais segmentos em todo o triângulo.

Por fim, consideramos um grafo  $G$  definido a seguir. Cada vértice de  $G$  está associado a um triângulo da dissecção de  $S$  e há um vértice extra  $x$  associado à região externa a  $S$ . Dois vértices de  $G$  são adjacentes se há um segmento azul-vermelho na fronteira de suas respectivas regiões. Note que pelo o que foi discutido anteriormente, o vértice  $x$  deve ter grau ímpar e os demais vértices de  $G$  de grau ímpar devem ser exatamente aqueles associados a triângulos arcoíris. O resultado segue do fato que todo grafo tem um número par de vértices de grau ímpar.  $\square$

◇ ◇ ◇

Aula 10 (03 de Maio) — Bruno Cavalari

◇ ◇ ◇

## 6.2 Valuações em grupos abelianos

Uma tripla  $(G, \cdot, \leq)$  é um grupo abeliano ordenado se

1.  $(G, \cdot)$  é um grupo.
2.  $(G, \leq)$  é uma ordem total, isto é, valem as seguintes propriedades:
  - (a) se  $x \leq y$  e  $y \leq x$ , então  $y = x$  ( $x, y \in G$ );
  - (b) se  $x \leq y$  e  $y \leq z$ , então  $x \leq z$  ( $x, y, z \in G$ );
  - (c) para todo  $x, y \in G$ , ou  $x \leq y$  ou  $y \leq x$ .
3. se  $x, y \in G$  são tais que  $x \leq y$ , então  $xz \leq yz$  para todo  $z \in G$ .

Um exemplo de grupo abeliano ordenado é a tripla  $(\mathbb{R}_{\geq 0}, \cdot, \leq)$ . Diremos apenas que  $G$  é um grupo abeliano ordenado, quando não há possibilidade de confusão quanto a operação  $\cdot$  e a ordenação  $\leq$  que definem a tripla  $(G, \cdot, \leq)$ .

Dado um grupo abeliano ordenado  $G$  consideraremos o conjunto  $G \cup \{0\}$ , com  $0 \in G$ , e estendemos a operação  $\cdot$  e a relação  $\leq$  da seguinte forma:

1.  $0 \cdot a = a \cdot 0 = 0$ , para todo  $a \in G$ ;
2.  $0 \leq a$ , para todo  $a \in G$ .

Seja  $\mathbb{K}$  um corpo. Uma *valoração não-arquimediana* em um grupo abeliano ordenado  $G$  é uma função  $v: \mathbb{K} \rightarrow G \cup \{0\}$  satisfazendo

1.  $v(x) = 0$  se, e somente se,  $x = 0$ ;
2.  $v(xy) = v(x)v(y)$  para todo  $x, y \in \mathbb{K}$ ;
3.  $v(x + y) \leq \max\{v(x), v(y)\}$  para todo  $x, y \in \mathbb{K}$ .

Queremos mostrar o seguinte resultado.

**Teorema 6.8.** *Existe uma valoração arquimediana  $v: \mathbb{R} \rightarrow G \cup \{0\}$  tal que  $v(\frac{1}{2}) > 1$ .*

Dada uma valoração  $v: \mathbb{K} \rightarrow G \cup \{0\}$ , definimos  $R_v = \{x \in \mathbb{K} : v(x) \leq 1\}$  como o *anel de valoração* de  $v$ . Note que  $R_v$  é, de fato, um anel.

**Lema 6.9.** *Um subanel próprio  $R$  de  $\mathbb{K}$  é um anel de valoração para algum  $v$  (em algum grupo ordenado  $G$ ) se, e somente se,  $\mathbb{K} = R \cup R^{-1}$  ( $R^{-1} = \{x^{-1} : x \in U(R)\}$ ).*

*Demonstração.*  $(\Rightarrow)$  Basta notar que se  $x \notin R$ , então  $v(x) > 1$ . Logo  $v(x^{-1}) < 1$ , isto é,  $x^{-1} \in R$  e, portanto,  $x = (x^{-1})^{-1} \in R^{-1}$ .

$(\Leftarrow)$  Suponha que  $\mathbb{K} = R \cup R^{-1}$ . Vamos definir um grupo ordenado  $G$  e uma valoração  $v$  de forma que  $R = R_v$ . Seja  $U$  o grupo das unidades de  $R$ , isto é, o conjunto dos elementos que têm inverso em  $R$ . Note que  $U$  é um subgrupo de  $\mathbb{K}^* = \mathbb{K} \setminus \{0\}$ . Defina  $G := \mathbb{K}^*/U$ . Dados dois elementos  $xU, yU \in G$  fazemos  $xU \leq yU$  se, e somente se,  $xy^{-1} \in R$ . Vamos mostrar que  $\leq$  é de fato uma ordem em  $G$ :

1. Se  $xU \leq yU$  e  $yU \leq xU$ , então  $xy^{-1} \in R$  e  $yx^{-1} \in R$ . Logo  $(xy^{-1})^{-1} \in R$ , o que implica  $xy^{-1} \in U$ . Segue da definição de grupo quociente que  $xU = yU$ .
2. Se  $xU \leq yU$  e  $yU \leq zU$ , então  $xy^{-1} \in R$  e  $yz^{-1} \in R$ . Logo  $xy^{-1}yz^{-1} = xz^{-1} \in R$ , isto é,  $xU \leq zU$ .
3. Se não vale  $xU \leq yU$ , então  $xy^{-1} \notin R$ , da onde segue que  $(xy^{-1})^{-1} \in R$  e, portanto,  $yx^{-1} \in R$ , ou seja,  $yU \leq xU$ .
4. Se  $xU \leq yU$ , então  $xy^{-1} \in R$ . Logo  $(xz)(yz)^{-1} = xzz^{-1}y^{-1} \in R$ , isto é,  $xz \leq yz$ .

Definimos  $v : \mathbb{K} \rightarrow G \cup \{0\}$  fazendo  $v(x) = xU$  para todo  $x \in K \setminus \{0\}$  e  $v(0) = 0$ . Temos que provar que  $v$  é de fato uma valuação. A primeira propriedade segue imediatamente da definição. Também temos  $v(xy) = (xy)U = (xU)(yU) = v(x)v(y)$ , o que implica a segunda propriedade. Para a terceira propriedade, note primeiro que  $1 \in R$ . De fato, se  $1 \notin R$ , então teríamos  $1 \in R^{-1}$  e portanto  $1 = 1^{-1} \in R$ . Agora, suponha que  $xU$  e  $yU$  sejam tais que  $xU \leq yU$ , isto é,  $xy^{-1} \in R$ . Então como

$$(x+y)y^{-1} = \underbrace{xy^{-1}}_{\in R} + \underbrace{1}_{\in R} \in R,$$

temos  $v(x+y) = (x+y)U \leq yU = v(y) = \max\{v(x), v(y)\}$ .

Resta mostrar que  $R$  é o anel de valuação de  $v$ . Note que

$$x \in R \Leftrightarrow x \cdot 1 \in R \Leftrightarrow xU \leq U \Leftrightarrow v(x) \leq 1.$$

Portanto, de fato temos  $R = \{x \in K : v(x) \leq 1\}$ .

□

**Lema 6.10.** *Se  $B \subseteq \mathbb{R}$  é um subanel maximal (com respeito a inclusão) satisfazendo  $\frac{1}{2} \notin B$ , então  $B$  é um anel de valuação.*

*Demonstração.* Suponha por absurdo que  $B$  satisfaça as hipóteses do enunciado mas não seja um anel de valuação. Pelo Lemma 6.9 deve existir  $\alpha \in \mathbb{R} \setminus (B \cup B^{-1})$ . Considere o conjunto  $B[\alpha]$  dos polinômios em  $\alpha$  com coeficientes em  $B$ . Temos  $B \subset B[\alpha]$  e, portanto,  $\frac{1}{2} \in B[\alpha]$ , caso contrário  $B$  não seria maximal. Logo, como  $1 \in 2B[\alpha]$ , devem existir  $u_0, \dots, u_m \in B$  tais que

$$1 = 2u_0 + 2u_1\alpha + \dots + 2u_m\alpha^m. \quad (2)$$

Analogamente, também temos  $1 \in 2B[\alpha^{-1}]$  e portanto existem  $v_0, \dots, v_n \in B$  tais que

$$1 = 2v_0 + 2v_1\alpha^{-1} + \dots + 2v_n\alpha^{-n}. \quad (3)$$

Suponha sem perda de generalidade que  $m$  e  $n$  são minimais nas equações (2) e (3) e que  $m \geq n$ . Multiplicando a equação (3) por  $\alpha^m$ , obtemos

$$(1 - 2v_0)\alpha^m = 2v_1\alpha^{m-1} + \dots + 2v_n\alpha^{m-n}. \quad (4)$$

Multiplicando a equação (2) por  $(1 - 2v_0)$  obtemos

$$1 = 2(u_0(1 - 2v_0) + v_0) + 2u_1(1 - 2v_0)\alpha + \dots + 2u_m(1 - 2v_0)\alpha^m \quad (5)$$

Substituindo o termo  $(1 - 2v_0)\alpha^m$  pelo lado direito da equação (4), obtemos uma equação que contradiz a minimalidade de  $m$ .  $\square$

**Lema 6.11** (Lema de Zorn). *Seja  $(\emptyset \neq P, \leq)$  um conjunto parcialmente ordenado. Suponha que toda cadeia  $(A_i)_{i=1}^{\infty}$  de  $P$  admita um limite superior, isto é, um elemento  $U \in P$  tal que  $A_i \leq U$  para todo  $i \geq 1$ . Então  $P$  tem um elemento maximal.*  $\square$

*Demonstração do Teorema 6.8.* O conjunto  $P$  dos subanelis de  $\mathbb{R}$  que não contém  $\frac{1}{2}$  é parcialmente ordenado (por inclusão). Note que  $P \neq \emptyset$  pois  $\mathbb{Z} \in P$ .

Além disso,  $P$  satisfaz as hipóteses do Lema de Zorn. De fato, se  $B_1 \subseteq B_2 \subseteq \dots$  é uma cadeia, então  $B := \bigcup_{i \geq 1} B_i$  contém  $\frac{1}{2}$  e é um subanel de  $R$  (uma vez que se  $x \in B_i$  e  $y \in B_j$ , então  $x \in B_j$  e, portanto,  $x + y \in B_j \subseteq B$  e  $xy \in B_j \subseteq B$ ).

Portanto existe subanel  $B \subseteq \mathbb{R}$  maximal com  $\frac{1}{2} \notin B$ . O resultado segue do Lemma 6.10.  $\square$

◇ ◇ ◇

Aula 11 (10 de Maio) — Arnaldo Mandel

◇ ◇ ◇

## 6.3 Valuações e Grupos Livres

### 6.3.1 Quatérnios

Seja  $R$  um domínio de integridade, isto é, um anel comunitativo em que o produto de quaisquer elementos não-nulos é não-nulo. Suponha que  $\text{car } R \neq 2$  e seja  $Q$  o corpo de frações de  $R$ . Por exemplo, poderíamos tomar  $R = \mathbb{Z}$  e  $Q = \mathbb{Q}$ .

Dados  $0 \neq a, b \in R$  definimos o conjunto dos quatérnios

$$T = \begin{pmatrix} a, b \\ R \end{pmatrix}$$

como um espaço vetorial de dimensão quatro sobre  $\mathbb{R}$  dotado de uma operação associativa e distributiva e multiplicação definida como a seguir. Escrevemos os elementos de  $T$  na forma

$$\alpha + \beta i + \gamma j \delta k,$$

onde os elementos  $1, i, j, k$  formam uma base de  $T$  e satisfazem as seguintes identidades:

$$i^2 = a, j^2 = b \text{ e} \\ ij = k = -ji.$$

Temos, por exemplo,

$$jk = j(-ji) = -j^2i = -bi, \\ k^2 = ij(-ji) = -ij^2i = -bi^2 = -ab, \\ kj = ijj = bi = -jk, \\ ik = i(ij) = (i^2)j = ja = j(i^2) = (ji)i = -ki.$$

Os quatérnios de Hamilton são  $\mathcal{H} = \begin{pmatrix} -1, -1 \\ \mathbb{R} \end{pmatrix}$ .

Definimos também o conjunto dos *quaternios puros* como

$$P = \{\beta i + \gamma j + \delta k : \beta, \gamma, \delta \in R\}.$$

**Proposição 6.12.** Se  $q \in P$ , então  $q^2 \in R$ .

*Demonstração.* Seja  $q = \beta i + \gamma j + \delta k$ . Temos

$$\begin{aligned} q^2 &= (\beta i + \gamma j + \delta k)^2 \\ &= \beta^2 a + \gamma^2 b - \delta^2 ab + \underbrace{(\beta i \gamma j + \gamma j \beta i)}_{\rightarrow 0} + \underbrace{(\beta i \delta k + \delta j \beta i)}_{\rightarrow 0} + \underbrace{(\gamma j \delta k + \delta k \gamma j)}_{\rightarrow 0} \\ &= \beta^2 a + \gamma^2 b - \delta^2 ab \in R. \end{aligned}$$

□

**Proposição 6.13.** Seja  $q = (\alpha + p)$ . Então  $q^2 \in R$  se, e somente se,  $p = 0$  ou  $\alpha = 0$ .

*Demonstração.* Temos  $q^2 = (\alpha + p)^2 = (\alpha^2 + p^2) + 2\alpha p$ . Pela proposição anterior, temos que  $q^2 \in R$  se e somente se  $2\alpha p = 0$ . O resultado segue da premissa de  $R$  ser um domínio de integridade. □

Considere um elemento  $q = \alpha + p \in T$ , com  $\alpha \in R$  e  $p \in P$ . Definimos a *norma* de  $q$  como

$$N(q) = (\alpha + p)(\alpha - p) = \alpha^2 - p^2 = \alpha^2 - \beta^2 a - \gamma^2 b + \delta^2 ab.$$

**Proposição 6.14.** Seja  $q \in T$ . Se  $N(q)$  é inversível, então  $q$  é inversível.

*Demonstração.* Seja  $q = \alpha + p$ , com  $\alpha \in R$  e  $p \in P$ . Se  $N(q)$  é inversível, então temos  $q^{-1} = (\alpha - p)N(q)^{-1}$ . □

**Corolário 6.15.**  $\mathcal{H}$  é um anel de divisão.

*Demonstração.* Em  $\mathcal{H}$ ,  $N(\alpha + \beta i + \gamma j + \delta k) = \alpha^2 + \beta^2 + \gamma^2 + \delta^2$ . Logo se  $q \neq 0$ , então  $N(q) > 0$  e, portanto,  $q$  é inversível. □

O *centro* de um anel  $A$  é definido como  $Z(A) = \{q \in A : qx = xq \text{ para todo } x \in A\}$ , isto é, como o conjunto dos elementos que comutam com qualquer  $x \in A$ .

**Exercício 6.16.**  $Z(T) = R$

Seja  $\mathcal{U}(T)$  o conjunto dos elementos inversíveis de  $T$ . Denote por  $\text{aut } T$  o conjunto de automorfismos de  $T$ . Definimos a função

$$\begin{aligned} \rho : \mathcal{U}(T) &\rightarrow \text{aut } T \\ w &\mapsto (x \mapsto w^{-1}xw). \end{aligned}$$

Note que temos  $\rho(wu) = \rho(w)\rho(u)$  e  $\rho(w^{-1}) = \rho(w)^{-1}$ . Além disso, segue da definição de centro que  $\text{Ker } \rho = Z(\mathcal{U}(T)) = \mathcal{U}(R)$ .

**Exercício 6.17.** Para todo  $w \in \mathcal{U}(T)$ , a função  $\rho(w)$  é invariante em  $P$ , isto é,  $\rho(w)(P) \subseteq P$ .

Como podemos identificar  $P$  com  $\mathbb{R}^3$ , segue que para  $T = \mathcal{H}$ ,  $\rho$  induz elementos de  $GL(3)$ , isto é, que  $\rho(w)$  é uma transformação linear em  $P \simeq \mathbb{R}^3$  para todo  $w \in T$ . Mais do que isso, temos que  $\rho$  induz elementos em  $SO(3)$ . De fato, note que

$$\|\rho(w)(x)\|^2 = N(\rho(w)(x)) = -(\rho(w)(x))^2 = \rho(w)(-x^2) = \rho(w)(N(x)) = N(x) = \|x\|^2.$$

### 6.3.2 Grupos livres

Em um grupo, dizemos que um par  $(u, v)$  é *livre* se  $\langle u, v \rangle$  é livre gerado por  $u$  e  $v$ . Dizemos que  $(u, v)$  é *livre módulo centro* se  $\bar{u}, \bar{v}$  é livre em  $G/Z(G)$ .

Dizemos que  $(g, h)$  é *semilivre módulo centro* se

1.  $g^n \notin Z(G)$  para todo  $n \in \mathbb{Z} \setminus \{0\}$ ;
2.  $h \notin Z(G)$ ,  $h^2 \in Z(G)$ ; e
3. não existem inteiros não-nulos  $n_1, \dots, n_k$  tais que

$$g^{n_1} h g^{n_2} h \dots g^{n_k} h \in Z(G).$$

**Proposição 6.18.** *Se  $(g, h)$  é semilivre módulo centro, então  $(g, hgh)$  é livre.* □

**Lema 6.19** (Lema do Ping Pong). *Seja  $G$  um grupo agindo sobre um conjunto  $X$ . Sejam  $g, h \in G$ , com  $h^2 \in Z$ . Suponha que existam  $X_1, X_2 \subseteq X$  tais que*

1.  $X_1, X_2 \neq \emptyset$ ;
2.  $X_1 \cap X_2 = \emptyset$ ;
3.  $Z(G) \cdot X_1 \subseteq X_1$  e  $Z(G) \cdot X_2 \subseteq X_2$ .

*Se  $h(X_1) \subseteq X_2$  e  $g^n(X_2) \subseteq X_1$  para todo inteiro  $n \neq 0$ , então  $(g, h)$  é semilivre módulo centro.*

*Demonstração.* Suponha por absurdo que existam inteiros  $n_1, \dots, n_k$  não-nulos tais que

$$w := g^{n_1} h g^{n_2} h \dots h g^{n_k} h \in Z(G).$$

Seja  $x$  um elemento arbitrário em  $X_2$ . Como  $w \in Z(G)$ , temos por hipótese que  $w \cdot x \in X_2$ . Por outro lado, como  $w = gwg^{-1}$ , segue das hipóteses  $h(X_1) \subseteq X_2$  e  $g^n(X_2) \subseteq X_1$  que  $w \cdot x \in X_1$ , um absurdo. □

**Teorema 6.20.** *Seja  $R$  um domínio de integridade com  $\text{car } R \neq 2$  e corpo de frações  $Q$ . Sejam  $0 \neq a, b, \alpha, \beta \in R$ . Suponha que exista uma valuação não-arquimediana  $v$  em  $R[i] (\simeq R[\sqrt{a}])$  tal que  $v(a) = v(b) = v(\beta) = 1$  e  $v(1 + \alpha i) \neq v(1 - \alpha i)$ . Então  $(1 + \alpha i, 1 + \beta j)$  é semilivre em  $\left(\frac{a, b}{Q}\right)$ .*

*Demonstração.* Seja  $F = R + Ri$  e  $T = \left(\frac{a, b}{Q}\right)$ . Temos  $T = F + Fj$ , pois  $(\alpha + \beta i) + (\gamma + \delta i)j = \alpha + \beta i + \gamma j + \delta k$ . Seja  $G = U(T)$  um grupo agindo sobre  $X = T$  com  $g \cdot x = xg$ . Tome

$$X_1 = \{x + yj : v(x) = v(y)\} \quad \text{e} \quad X_2 = \{x + yj : v(x) \neq v(y)\}.$$

Considere  $x + yj \in X_1$ . Temos

$$(x + yj)(1 + \alpha i)^n = x(1 + \alpha i)^n + y(1 - \alpha i)^n j.$$

Como  $v(x(1 + \alpha i)^n) = v(x)v(1 + \alpha i)^n$  e  $v(y(1 - \alpha i)^n) = v(y)v(1 - \alpha i)^n$ , segue que  $(x + yj)(1 + \alpha i)^n \in X_2$ . Para podermos aplicar o Lema do Ping Pong, falta mostrar que  $(x + yj)(i + \beta j) \in X_1$ . De fato, temos

$$(x + yj)(i + \beta j) = (ix + b\beta y) + (\beta x - iy)j$$

Como  $v(ix) = v(x) \neq v(y) = v(\beta y)$ , segue que  $v(ix + b\beta y) = \max\{v(x), v(y)\}$ . Analogamente, também temos  $v(\beta x - iy) = \max\{v(x), v(y)\}$ , o que implica  $(x + yj)(i + \beta j) \in X_1$ .  $\square$

## 7 Happy Ending Theorem

◇ ◇ ◇ *Aula 12 (17 de Maio) — Marcelo Tadeu Sales* ◇ ◇ ◇

Seja  $ES(n)$  o número mínimo de pontos no plano em posição geral (isto é, três-a-três não colineares) tal que é garantida a existência de um subconjunto de  $n$  pontos que formam um  $n$ -ágono convexo.

Erdős e Szekeres provaram (1935) que

$$ES(n) \leq \binom{2n-4}{n-2} + 1 \approx \frac{4^n}{\sqrt{n}}$$

e conjecturaram que  $ES(n) = 2^{n-2} + 1$ . Em 1961, Erdős e Szekeres (1961) construíram exemplos que mostram que

$$ES(n) \geq 2^{n-2} + 1.$$

Veremos aqui uma prova recente que mostra que  $ES(n) \leq 2^{n+o(n)}$

**Teorema 7.1** (Andrew Suk, 2016).

$$ES(n) \geq 2^{n+2n^{4/5}}.$$

### 7.1 Preliminares

Sejam  $p_1 = (x_1, y_1), \dots, p_k = (x_k, y_k)$  pontos no plano e suponha que  $x_1 < x_2 < \dots < x_n$  (note que para qualquer conjunto de pontos em posição geral sempre podemos fixar um sistema de coordenadas tal que suas abscissas sejam todas distintas). Dizemos que  $\{p_1, p_2, \dots, p_k\}$  formam um  $k$ -cap se a linha poligonal  $p_1 p_2 \dots p_k$  for côncava, isto é, se  $\text{tg}(p_1, p_2) > \text{tg}(p_2, p_3) > \dots > \text{tg}(p_{k-1}, p_k)$  onde  $\text{tg}(p_i, p_j) = \frac{y_j - y_i}{x_j - x_i}$ .

De maneira análoga, dizemos que  $\{p_1, p_2, \dots, p_k\}$  formam um  $k$ -cup se a linha poligonal  $p_1 p_2 \dots p_k$  for convexa, isto é, se  $\text{tg}(p_1, p_2) < \text{tg}(p_2, p_3) < \dots < \text{tg}(p_{k-1}, p_k)$ .

**Teorema 7.2** (Caps e Cups). *Seja  $f(k, l)$  o número mínimo de pontos do plano em posição geral tal que é garantido a existência ou de um  $k$ -cup ou de um  $l$ -cap. Então*

$$f(k, l) = \binom{k+l-4}{k-2} + 1.$$

*Demonstração.* Vamos provar apenas a desigualdade  $f(k, l) \leq \binom{k+l-4}{k-2} + 1$ , por indução em  $k+l$ . Se  $k = 2$ , então quaisquer dois pontos determinam um  $k$ -cap. Logo  $f(k, 2) =$

$2 \leq \binom{l-2}{0} + 1$ . Da mesma forma, temos  $f(2, l) = 2 \leq \binom{k-2}{0} + 1$ .

Para o passo da indução, vamos mostrar que

$$f(k, l) \leq f(k-1, l) + f(k, l-1) - 1.$$

Seja  $n = f(k-1, l) + f(k, l-1) - 1$  e sejam  $p_1, \dots, p_n$  pontos ordenados pela abscissa. Coloque  $m = f(k-1, l)$  e considere os seguintes conjuntos de  $m$  pontos

$$\begin{aligned} U_1 &= \{p_1, \dots, p_{m-1}, p_m\} \\ U_2 &= \{p_1, \dots, p_{m-1}, p_{m+1}\} \\ U_3 &= \{p_1, \dots, p_{m-1}, p_{m+2}\} \\ &\vdots \\ U_{f(k, l-1)} &= \{p_1, \dots, p_{m-1}, p_n\} \end{aligned}$$

Se algum deles contiver um  $l$ -cap, não há nada a fazer. Suponha então que cada  $U_i$  contém um conjunto de pontos  $U_i = \{u_i^1, u_i^2, \dots, u_i^{k-1}\}$  (ordenados pela abscissa) que formam um  $(l-1)$ -cup. Seja  $Y = \{u_1^{k-1}, \dots, u_{f(k, l-1)}^{k-1}\}$  o conjunto formado pelos últimos pontos de cada um desses  $(k-1)$ -cups. Se  $Y$  contiver um  $k$ -cup não há nada a fazer. Suponha que  $Y$  contém um  $(l-1)$ -cap  $\Lambda$ . Sejam  $u_r^{k-1}$  e  $u_s^{k-1}$  o primeiro e o segundo ponto de  $\Lambda$  com menor abscissa, respectivamente.

Se  $\text{tg}(u_r^{k-1}, u_s^{k-1}) > \text{tg}(u_r^{k-2}, u_r^{k-1})$ , então  $U_r \cup u_s^{k-1}$  forma um  $k$ -cup. Caso contrário  $\Lambda \cup u_r^{k-2}$  forma um  $l$ -cup. Segue, usando a hipótese de indução, que

$$f(k, l) \leq f(k-1, l) + f(l, k-1) + 1 \leq \binom{k+l-5}{k-3} + \binom{k+l-5}{k-2} + 1 \leq \binom{k+l-4}{k-2},$$

como desejado. □

Seja  $H = ([n], \binom{[n]}{3})$  o hipergrafo 3-uniforme completo cujo conjunto de vértices é  $[n]$ . Dizemos que uma 2-coloração  $c : \binom{[n]}{3} \rightarrow \{0, 1\}$  das arestas de  $H$  é *transitiva* se para todo  $i_1 < i_2 < i_3 < i_4 \in [n]$  vale que

$$\text{se } c(\{i_1, i_2, i_3\}) = c(\{i_2, i_3, i_4\}), \text{ então } \{i_1, \dots, i_4\} \text{ é monocromático,}$$

isto é, todas as triplas em  $\{i_1, \dots, i_4\}$  têm a mesma cor.

**Teorema 7.3.** *Seja  $g(k, l)$  o número mínimo de vértices em um hipergrafo 3-uniforme completo com uma 2-coloração transitiva de forma a garantir a existência de um  $k$ -clique da cor 0 ou um  $l$ -clique da cor 1. Então  $g(k, l) = f(k, l)$ .*

*Demonstração.* Para mostrar que  $f(k, l) \leq g(k, l)$  considere um conjunto de  $n = g(k, l)$  pontos no plano e construa um hipergrafo completo onde cada tripla é colorida com a cor 0 ou 1 se formar um cup ou um cap, respectivamente. Agora basta notar que essa coloração é transitiva.

Mostraremos, então, que  $g(k, l) \leq f(k, l)$ . Novamente, mostramos por indução que  $g(k, l) \leq \binom{k+l-4}{k-2} + 1$ . A afirmação é vacuamente verdadeira para  $k = 2$  ou  $l = 2$ . Para o passo de indução, mostraremos que  $g(k, l) \leq g(k-1, l) + g(k, l-1) - 1$ . Ponha  $m = g(k-1, l)$  e considere os conjuntos  $X_i = \{1, 2, \dots, m-1, m+i-1\}$ , para  $i = 1, \dots, g(k, l-1)$

1). Suponha que nenhum deles contenha um  $l$  clique da cor 1. Então cada  $X_i$  deve conter um subconjunto  $U_i = \{u_i^1, \dots, u_i^{k-1}\}$  que forma um clique da cor 0. Seja  $Y = \{u_i^{k-1}\}_{i=1}^{g(k,l-1)}$ . Se  $Y$  não contiver um  $k$ -clique da cor 0, então  $Y$  deve conter um conjunto  $\Lambda$  que forma um  $(l-1)$ -clique da cor 0. Sejam  $u_r^{k-1} < u_s^{k-1}$  os dois menores elementos de  $\Lambda$ .

Suponha que  $c(u_r^{k-2}, u_r^{k-1}, u_s^{k-1}) = 0$ . Fixe  $i < j < k-2$ . Como  $c(u_r^i, u_r^{k-2}, u_r^{k-1}) = 0$ , devemos ter  $c(u_r^i, u_r^{k-1}, u_s^{k-1}) = 0$ , por transitividade. Como  $c(u_r^i, u_r^j, u_r^{k-1}) = 0$ , também segue da transitividade que  $c(u_r^i, u_r^j, u_s^{k-1}) = 0$ . Logo  $U_r \cup u_s^{k-1}$  é um clique da cor 0. Por outro lado, se  $c(u_r^{k-2}, u_r^{k-1}, u_s^{k-1}) = 1$  podemos usar a transitividade de maneira análoga para mostrar que  $\Lambda \cup u_r^{k-1}$  é um clique da cor 1.  $\square$

Dado um  $n$ -ágono  $X = \{x_1, \dots, x_n\}$  no plano definimos o *suporte* de  $X$  como a coleção de regiões  $\mathcal{C} = \{T_1, \dots, T_n\}$ , onde  $T_i$  é a região externa ao  $n$ -ágono definida pelas retas  $(x_{i-1}, x_i)$  e  $(x_{i+1}, x_{i+2})$ .

**figura: num pentágono regular as regiões são triângulos. Num triângulo, são regiões ilimitadas**

Dado um conjunto  $X$  de pontos no plano, dizemos que  $X$  é um  $\text{cup}^{\text{cap}}$  quando os pontos de  $X$  formam ou um cup ou um cap. Se  $|X| = k$ , dizemos que  $X$  é um  $k\text{-cup}^{\text{cap}}$ .

A seguir mostraremos que todo conjunto suficientemente grande de pontos do plano contém um  $k\text{-cup}^{\text{cap}}$  tal que todas as regiões de seu suporte contêm uma fração significativa do conjunto original de pontos.

**Lema 7.4.** *Sejam  $k \geq 3$  e  $P$  um conjunto finito de pontos no plano em posição geral tal que  $|P| \geq 4^k$ . Então existe um conjunto  $X \subseteq P$  tal que  $X$  é um  $k\text{-cup}^{\text{cap}}$  e as regiões  $T_1, \dots, T_k$  do suporte de  $X$  satisfazem*

$$|T_i \cap P| \geq \frac{|P|}{2^{40k}}.$$

*Demonstração.* Vamos estimar o número  $y$  de  $4k\text{-cup}^{\text{cap}}$ s em  $P$ . Para isso contamos o número de pares da forma  $(Z, S)$ , onde  $Z$  é um  $4k\text{-cup}^{\text{cap}}$  e  $S$  é um conjunto de  $4^k$  pontos de  $P$  que contém  $Z$ . Por um lado, o número de tais pares é igual a

$$\binom{n-4k}{4^{4k}-4k} y,$$

uma vez existem exatamente  $\binom{n-4k}{4^{4k}}$  conjuntos de  $4^k$  pontos que contêm um dado  $k\text{-cup}^{\text{cap}}$   $Z$ . Por outro lado, como qualquer conjunto de  $f(4k, 4k)$  pontos contém um  $4k\text{-cup}^{\text{cap}}$  e

$$f(4k, 4k) \leq \binom{8k-4}{4k-2} + 1 \leq 2^{8k} = 4^{4k},$$

o número de pares da forma  $(Z, S)$  é, no mínimo,  $\binom{n}{4^{4k}}$ . Concluimos que

$$y \leq \frac{\binom{n}{4^{4k}}}{\binom{n-4k}{4^{4k}-4k}} = \frac{\binom{n}{4k}}{\binom{n}{4k}}.$$

Dizemos que um conjunto  $W$  de  $2k$ -pontos *intercala* com um  $4k\text{-cup}^{\text{cap}}$   $Z = \{x_1, x_2, \dots, x_{4k}\}$  se  $W = \{x_1, x_3, \dots, x_{4k-1}\}$  ou se  $W = \{x_2, x_4, \dots, x_{4k}\}$ . Vamos contar o número de pares

$(W, Z)$ , onde  $Z$  é um  $4k$ - $\text{cap}_{\text{cup}}^{\text{cap}}$  e  $W$  é um conjunto de  $2k$ -pontos que intercala  $Z$ . Temos

$$\sum_W t_W = \#(W, Z) \geq 2y \geq 2 \frac{\binom{n}{4k}}{\binom{4k}{4k}},$$

onde  $t_W$  é o número de  $4k$ - $\text{cap}_{\text{cup}}^{\text{cap}}$  que  $W$  intercala. Logo deve existir  $W$  satisfazendo

$$\begin{aligned} t_W &\geq \frac{2 \binom{n}{4k}}{\binom{4k}{4k} \binom{n}{2k}} \\ &\geq \frac{2n(n-1)\cdots(n-4k+1)(2k)!}{\binom{4k}{4k} n(n-1)\cdots(n-2k+1)(4k)!} > \frac{2(2k)!(n-4k)^{2k}}{\binom{4k}{4k}(4k)!} \\ &> \frac{(n-4k)^{2k}}{(4k)^{4k}} > \frac{n^{2k}}{2^{2k} 2^{32k^2}} > \frac{n^{2k}}{2^{33k^2}}. \end{aligned}$$

Suponha sem perda de generalidade que  $W$  é um cap. Sejam  $T_1, \dots, T_{2k}$  as regiões do suporte de  $W$  e defina  $b_i = |T_i \cap P|$  para todo  $i = 1 \dots 2k$ . Temos

$$\frac{n^{2k}}{2^{33k^2}} \leq \prod_{i=1}^{2k} b_i.$$

Sejam  $i_1, \dots, i_{2k}$  índices tais que  $b_{i_1} < b_{i_2} < \dots < b_{i_{2k}}$ . Defina  $\alpha = b_{i_k}$ . Como

$$\prod_{i=1}^{2k} b_i = (b_{i_1} \dots b_{i_k})(b_{i_{k+1}} \dots b_{i_{2k}}) \leq \alpha^k n^k,$$

obtemos

$$\alpha n > \frac{n^2}{2^{33k}}, \text{ isto é, } \alpha > \frac{n}{2^{33k}} > \frac{n}{2^{40k}}.$$

Concluimos que cada uma das regiões  $T_{i_{k+1}}, \dots, T_{i_{2k}}$ , que estão no suporte de  $W$ , contém pelo menos  $n/(2^{40k})$  pontos de  $P$ . Para concluir a demonstração, basta considerar o  $k$ -cap  $X = \{w_{i_{k+1}}, \dots, w_{i_{2k}}\}$ , uma vez que cada região desse  $k$ -cap contém alguma das regiões  $T_{i_{k+1}}, \dots, T_{i_{2k}}$ .  $\square$

Seja  $(P, \leq)$  um conjunto parcialmente ordenado e finito. Dizemos que uma sequência  $(a_1, \dots, a_k)$  de elementos de  $P$  é uma *cadeia* se  $a_1 \leq a_2 \leq \dots \leq a_k$ . Dizemos que um conjunto  $A = \{a_1, \dots, a_k\} \in P$  é uma *anticadeia* se para todo  $a_i \not\leq a_j$  todo  $1 \leq i, j \leq k$ .

Dada uma anticadeia  $A$  e uma partição  $\mathcal{C} = \{C_1, \dots, C_l\}$  de  $P$  em cadeias, então devemos ter  $|A| \leq |\mathcal{C}|$ , uma vez que cada  $C_i$  pode conter no máximo um elemento de  $A$ . Segue, em particular, que o tamanho de uma anticadeia em  $P$  de tamanho máximo é menor ou igual ao tamanho de uma partição em cadeias de  $P$  de tamanho mínimo. Usaremos o resultado a seguir, que afirma que, nesse caso, vale a igualdade.

**Teorema 7.5** (Teorema de Dilworth). *Seja  $(P, \leq)$  um conjunto parcialmente ordenado e finito. O número máximo de elementos em uma anticadeia de  $P$  é igual ao número mínimo de cadeias em uma partição de  $P$  em cadeias.*  $\square$

**Corolário 7.6.** *Sejam  $a, b > 0$  tais que  $|P| = ab$ . Então  $P$  contém ou uma anticadeia de tamanho pelo menos  $a$ , ou uma cadeia de tamanho pelo menos  $b$ .*

*Demonstração.* Suponha que o número máximo de elementos em uma anticadeia seja  $s < a$ . Pelo Teorema de Dilworth, existe uma partição de  $P$  em  $s$  cadeias. Logo, alguma delas deve conter pelo menos  $|P|/s > b$  elementos.  $\square$

## 7.2 Demonstração do Teorema 7.1

Considere um conjunto  $P$  de pontos em posição geral no plano de cardinalidade  $|P| = N = 2^{n+2n^{4/5}}$ . Aplicamos o Lema 7.4 com  $k = t + 3$ , onde  $t = \sqrt{n}$  para obter um conjunto  $(t + 3)$ -cup  $X$  com suporte  $\{T_1, \dots, T_{t+3}\}$  satisfazendo

$$|T_i \cap P| \geq \frac{N}{2^{40(t+3)}} \geq \frac{N}{2^{50t}}.$$

A seguir, assumiremos que  $X$  é um  $(t + 3)$ -cup. O prova para o caso em que  $X$  é um  $(t + 3)$ -cup será análoga.

Para todo  $1 \leq i \leq (t + 3)$ , seja  $P_i = T_i \cap P$ . Definimos uma ordem parcial em  $P_i$  da seguinte forma: dados dois pontos  $p \neq q \in P_i$ , fazemos

$$p < q \text{ se, e somente se, } q \in \text{conv}(p, x_{i-1}, x_{i+2}),$$

isto é,  $p < q$  se  $q$  pertence ao triângulo cujos vértices são  $p$ ,  $x_{i-1}$  e  $x_{i+2}$ . Note que  $p$  e  $q$  são comparáveis se, e somente se, a reta  $(p, q)$  intersecta a reta  $r_i = (x_{i-1}, x_{i+2})$ .

Seja  $\alpha = n^{-\frac{1}{5}}$ . Dizemos que uma região  $P_i$  é

- do tipo A se  $P_i$  contém uma anticadeia de tamanho pelo menos  $|P_i|^\alpha$ ; ou
- do tipo B se  $P_i$  contém uma cadeia de tamanho pelo menos  $|P_i|^{1-\alpha}$ .

Pelo Corolário 7.6 toda região é ou do tipo A ou do tipo B. Note que existem ou  $\sqrt{t}$  regiões do tipo A ou  $\sqrt{t}$  regiões consecutivas do tipo B (de fato, considere  $\sqrt{t}$  conjuntos disjuntos de  $\sqrt{t}$  regiões consecutivas: se nenhum deles for composto apenas de regiões do tipo B, devemos ter pelo menos uma região do tipo A para cada um desses conjuntos). Analisaremos esses dois casos separadamente.

- *Caso 1: Existem pelo menos  $\sqrt{t}$  regiões do tipo A.* Como há pelo menos  $\sqrt{t}$  regiões do tipo A, podemos escolher um conjunto de  $s = \frac{1}{2}\sqrt{t} = \frac{1}{2}n^{1/4}$  regiões  $P_{i_1}, \dots, P_{i_s}$  duas-a-duas não consecutivas do tipo A. Note que  $P_{i_r}$  contém uma anticadeia  $A_{i_k}$  de tamanho

$$\begin{aligned} |A_{i_r}| &\geq |P_{i_r}|^\alpha \geq \left(\frac{N}{2^{50t}}\right)^{n^{-1/5}} = 2^{n^{4/5} + 2n^{3/5} - 50n^{3/10}} \\ &\geq 2^{2n^{3/4} + 2n^{3/4} \lg n} \geq (2n)^{2n^{3/4}} \geq (n + 2n^{\frac{3}{4}} - 4)2n^{3/4} \\ &\geq \left(\frac{n + 2n^{\frac{3}{4}} - 4}{2n^{\frac{3}{4}} - 2}\right) + 1 = f(n, 2n^{\frac{3}{4}}). \end{aligned}$$

Logo, em cada  $P_{i_r}$  deve haver ou um  $n$ -cup ou um  $2n^{\frac{3}{4}}$ -cup formado por pontos em uma anticadeia. Se algum  $P_{i_r}$  contiver um  $n$ -cup, não há nada a ser feito. Suponha então que cada  $P_{i_r}$  contém um  $2n^{\frac{3}{4}}$ -cup compostos por pontos em uma anticadeia. Afirmamos que  $\cup_{r=1}^s P_{i_r}$  é um  $n$ -cup. De fato,

- *Caso 2: Existe pelo menos  $\sqrt{t}$  regiões consecutivas do tipo B.* Suponha sem perda de generalidade que  $P_1, \dots, P_{\sqrt{t}}$  são regiões consecutivas do tipo B. Para cada inteiro  $1 \leq i \leq \sqrt{t}$ , fixamos uma cadeia de pontos  $Q_i$  de tamanho pelo menos  $|P_i|^{1-\alpha}$ . Dizemos que um conjunto de pontos  $A \subseteq Q_i$  é *convexo à direita*, se  $A \cup \{x_i\}$  forma um polígono convexo; ou *convexo à esquerda*, se  $A \cup \{x_{i+1}\}$  forma um polígono convexo.

**Afirmção 7.7.** *Para todo  $1 \leq i \leq \sqrt{t}$ , a cadeia  $Q_i$  contém ou um  $(n - (i-1)n^{\frac{3}{4}})$ -conjunto de pontos convexo à direita ou um  $(in^{\frac{3}{4}})$ -conjunto de pontos convexo à esquerda.*

*Demonstração.* Associe cada tripla  $\{p_{i_1}, p_{i_2}, p_{i_3}\} \subseteq Q_i$  à cor 0, se ela for uma tripla à esquerda ou à cor 1, se for uma tripla à direita. Note que toda tripla é associada a alguma cor e que a coloração assim definida é transitiva. Segue do Teorema 7.3, que se  $|Q_i| \geq g(k, l)$ , então  $Q_i$  contém um  $k$ -conjunto de pontos convexo à direita ou um  $l$ -conjunto convexo à esquerda. Como para quaisquer inteiros  $k, l > 0$  temos

$$g(k, l) = \binom{k+l-4}{k-2} + 1 \leq 2^{k+l-4} + 1 \leq 2^{k+l},$$

é suficiente mostrar que  $|Q_i| \geq 2^{n+n^{3/4}}$ . Mas como  $P_i$  é uma região do tipo B temos

$$\begin{aligned} |Q_i| &= |P_i|^{1-\alpha} \\ &\geq \frac{N^{1-n^{1/5}}}{2^{50t}} \\ &= (2^{n+2n^{4/5}-50n^{1/2}})1^{-n^{-1/5}} = \\ &= 2^{n+2n^{4/5}-50n^{1/2}-n^{4/5}-2n^{3/5}+50n^{3/10}} \\ &= 2^{n+n^{4/5}+o(n^{4/5})} \\ &\geq 2^{n+n^{3/4}}, \end{aligned}$$

como desejado. □

**Afirmção 7.8.** *Seja  $E$  um conjunto de pontos convexo à esquerda em uma região  $P_i$  do tipo B e  $D$  um conjunto de pontos convexo à direita em uma região  $P_{i+1}$  do tipo B. Então  $E \cup D$  forma um polígono convexo.*

*Demonstração.* Seja  $U = E \cup D$ . Para mostrar que  $U$  forma um polígono convexo, basta mostrar que quaisquer quatro pontos formam um quadrilátero convexo (de fato, se  $U$  não for convexo, então deve haver algum ponto no interior de um triângulo de uma triângulação do casco convexo de  $U$ ).

Sejam  $p_1, p_2, p_3, p_4 \in U$ . Suponha que  $p_1, p_2 \in E$  e  $p_3, p_4 \in D$ . Como  $p_1$  e  $p_2$  estão em uma cadeia, a reta  $(p_1, p_2)$  não pode cruzar a região  $P_{i+1}$ . Analogamente  $(p_3, p_4)$  não cruza a região  $P_i$ . Logo  $p_1, p_2, p_3, p_4$  formam um quadrilátero convexo. Agora suponha que  $p_1, p_2, p_3 \in E$  e  $p_4 \in D$ . Como  $p_1, p_2, p_3$  estão em uma cadeia, os pontos  $x_{i+1}$  e  $p_4$  devem estar na mesma região definida pelas retas  $(p_1, p_2), (p_2, p_3), (p_1, p_3)$ . Portanto, como  $p_1, p_2, p_3, x_{i+1}$  formam um quadrilátero

convexo, o mesmo deve ser verdade para os pontos  $p_1, p_2, p_3, p_4$ . Os demais casos seguem por simetria.  $\square$

Suponha que  $Q_1$  não possua um  $n$ -conjunto convexo à direita e que  $Q_{\sqrt{i}}$  não possui um  $\sqrt{i}p = n$ -conjunto de pontos convexo à esquerda (nesses casos não há nada a provar). Segue da Afirmação 7.7 que existe  $i$  tal que  $Q_i$  possui um  $ip$ -conjunto convexo à esquerda e  $Q_{i+1}$  possui um  $n - ip$  conjunto convexo à direita. Esses dois conjuntos de pontos formam um  $n$ -ágono convexo.

## 8 Flag Algebra

◇ ◇ ◇

Aula 13 (31 de Maio) — Fernando Mario de Oliveira Filho

◇ ◇ ◇

### 8.1 Exemplo: Teorema de Mantel

**Teorema 8.1** (Mantel 1910). *Todo grafo livre de triângulos de tamanho  $n$  tem no máximo  $\lfloor n^2/4 \rfloor$  arestas.*

Assintoticamente, o Teorema de Mantel afirma que todo grafo livre de triângulos tem densidade de arestas no máximo  $\frac{1}{2}$ . Nesta seção, veremos uma demonstração dessa versão assintótica do Teorema de Mantel que usa técnicas que estão por trás da teoria de Flag Algebras para grafos.

Seja  $p(F; G)$  a probabilidade de um conjunto  $U \subseteq V(G)$ , de cardinalidade  $|U| = |F|$ , escolhido uniformemente ao acaso ser tal que  $G[U] \simeq F$ , isto é, tal que o grafo induzido por  $U$  é isomorfo a  $F$ . Em outras palavras temos

$$p(F; G) = \left( \frac{|G|}{|F|} \right)^{-1} c(F; G),$$

onde  $c(F; G)$  é o número de cópias homomórficas (não-rotuladas) de  $F$  em  $G$ . Temos, por exemplo,  $c(\text{---}, \text{---}) = \binom{3}{2} = 3$ .

Seja  $\mathcal{H}$  uma coleção de grafos proibidos. Dizemos que  $G$  é  $\mathcal{H}$ -livre se  $G$  não possui subgrafo induzido isomorfo a  $\mathcal{H}$ , isto é, se  $p(F, G) = 0$  para todo  $F \in \mathcal{H}$ .

Fixe um grafo  $C$ . Definimos

$$\text{ex}(C, \mathcal{H}) = \sup_{(G_k)_{k \geq 0}} \lim_{k \rightarrow \infty} p(C; G_k),$$

onde o supremo é tomado sobre todas as sequências  $(G_k)_{k \geq 0}$  crescentes (dizemos que uma sequências  $(G_k)_{k \geq 0}$  é crescente se  $(|G_k|)_{k \geq 0}$  é crescente) de grafos  $\mathcal{H}$ -livres. Nessa linguagem, podemos enunciar o Teorema de Mantel da seguinte forma.

**Teorema 8.2** (Teorema de Mantel, versão assintótica).

$$\text{ex}(\text{---}, \{\text{---}\}) \leq \frac{1}{2}.$$

Fixe uma família  $\mathcal{H}$  de grafos e seja  $\mathcal{G}$  o conjunto de todos os grafos (a menos de isomorfismos)  $\mathcal{H}$ -livres. Dizemos que uma sequência  $(G_k)_{k \geq 0}$  é *convergente* se para

todo  $F \in \mathcal{G}$  existe  $\lim_{k \rightarrow \infty} p(F; G_k)$ .

**Proposição 8.3.** *Toda sequência crescente  $(G_k)_{k \geq 0}$  possui uma subsequência convergente.*

*Demonstração.* Como  $p(F; G) \in [0, 1]$ , podemos identificar a função  $\zeta_k : F \mapsto p(F, G_k)$  com um ponto do espaço  $[0, 1]^{\mathcal{G}}$ . Mas como  $[0, 1]$  é um espaço compacto, segue do Teorema de Tychonoff que  $[0, 1]^{\mathcal{G}}$  é compacto. Portanto, toda sequência  $(\zeta_k)_{k \geq 0}$  nesse espaço contém uma subsequência convergente.  $\square$

Dizemos que  $\phi : \mathcal{G} \rightarrow \mathbb{R}$  é um *funcional limite* se existe uma sequência convergente  $(G_k)_{k \geq 0}$  tal que  $\phi(F) = \lim_{k \rightarrow \infty} p(F; G_k)$  para todo  $F \in \mathcal{G}$ . Defina  $\Phi = \{\phi : \phi \text{ é funcional limite}\}$ . Note que podemos rescrever  $\text{ex}(C, \mathcal{H})$  como

$$\text{ex}(C, \mathcal{H}) = \sup\{\phi(C) : \phi \in \Phi\}.$$

Logo, uma maneira de se obter uma cota superior para  $\text{ex}(C, \mathcal{H})$  consiste em computar  $\sup\{\phi(C) : \phi \in \Phi'\}$  para algum  $\Phi' \supseteq \Phi$ .

*Demonstração do Teorema 8.2 (Razborov, Bondy).* Fixe  $\mathcal{H} = \{\text{triângulo}\}$ . Seja  $G$  um grafo  $\mathcal{H}$ -livre. Toda aresta de  $G$  pertence a  $|G| - 2$  subgrafos de  $G$  de três vértices. Portanto, temos

$$c(\text{triângulo}; G) + 2c(\text{triângulo}; G) = (|G| - 2)c(\text{caminho de 2}; G).$$

Dividindo a equação acima por  $\binom{|G|}{3}$ , obtemos

$$p(\text{triângulo}; G) + 2p(\text{triângulo}; G) = 3p(\text{caminho de 2}; G).$$

Logo, todo  $\phi \in \Phi$  satisfaz

$$\phi(\text{triângulo}) + 2\phi(\text{triângulo}) = 3\phi(\text{caminho de 2}). \tag{6}$$

A seguir obteremos outra condição que deve ser satisfeita por qualquer  $\phi \in \Phi$ . Para isso, precisaremos estender a definição de  $p(F; G)$  para grafos que possuem alguns vértices rotulados.

Seja  $F$  e  $G$  grafos contendo, respectivamente, vértices  $v_1 \in V(F)$  e  $w_1 \in V(G)$  com rótulo 1 (os demais vértices de  $F$  e  $G$  não são rotulados). Definimos  $p(F; G)$  como a probabilidade de um conjunto  $U \subseteq V(G) \setminus \{w_1\}$ , de cardinalidade  $|U| = |F| - 1$ , escolhido uniformemente ao acaso, ser tal que  $G[U \cup \{w_1\}] \simeq F$  (isto é, se há um isomorfismo de  $\phi$  de grafos entre  $G[U \cup \{w_1\}]$  e  $F$  satisfazendo  $\phi(w_1) = v_1$ ).

Seja  $G$  é um grafo  $\mathcal{H}$ -livre e  $v \in V(G)$ . Definimos  $G^v$  como o grafo isomorfo a  $G$  com rótulo 1 em  $v$ . Por exemplo, se  $G = \text{triângulo}$  e  $v$  é o vértice de grau dois de  $G$ , então  $G^v = \text{triângulo rotulado}$ .

Usaremos tais definições para estimar  $p(\text{triângulo}; G)$ . Note que para todo  $v \in V(G)$ , cada par de vértices na vizinhança de  $v$  forma (junto com  $v$ ) uma cópia de  $\text{triângulo}$  em  $G$ . Logo, temos

$$p(\text{triângulo}; G) = \binom{|G|}{3}^{-1} \sum_{v \in G} \binom{d(v)}{2} = \binom{|G|}{3}^{-1} \sum_{v \in G} p(\text{triângulo rotulado}; G^v) \binom{|G| - 1}{2} = \frac{3}{|G|} \sum_{v \in G} p(\text{triângulo rotulado}; G^v).$$

Agora afirmamos que se  $|G| \rightarrow \infty$ , então  $p(\text{triângulo}; G^v) \rightarrow p(\text{caminho}; G^v)^2$ . De fato, basta notar que  $\sum_{v \in G} p(\text{triângulo}; G^v)$  é a probabilidade de dois vértices distintos (escolhidos ao acaso) serem vizinhos de  $v$ , enquanto  $p(\text{caminho}; G^v)^2$  é a probabilidade de dois vértices escolhidos de forma independente serem vizinhos de  $v$ .

Seja  $\phi$  um funcional limite associado a uma sequência  $(G_k)_{k \geq 0}$  de grafos. Temos

$$\phi(\text{triângulo}) = \lim_{k \rightarrow \infty} p(\text{triângulo}; G_k) = \lim_{k \rightarrow \infty} \frac{3}{|G_k|} \sum_{v \in G_k} p(\text{triângulo}; G_k^v) = \lim_{k \rightarrow \infty} \frac{3}{|G_k|} \sum_{v \in G_k} p(\text{caminho}; G_k^v)^2.$$

Aplicando a desigualdade de Cauchy-Swartz na última soma da equação acima, obtemos

$$\phi(\text{triângulo}) \geq \lim_{k \rightarrow \infty} \frac{3}{|G_k|^2} \left( \sum_{v \in G_k} p(\text{caminho}; G_k^v) \right)^2.$$

Note que  $\sum_{v \in G} p(\text{caminho}; G_k^v) = p(\text{caminho}; G_k) |G_k|$ , uma vez que  $p(\text{caminho}; G_k^v) |G_k^v - 1| = \sum_{v \in G_k^v - 1} d(v) = 2p(\text{caminho}; G_k) \binom{|G_k|}{2}$ . Substituindo essa igualdade na inequação anterior, obtemos  $\phi(\text{triângulo}) \geq \lim_{k \rightarrow \infty} 3p(\text{caminho}; G_k)^2$ , ou seja,

$$\phi(\text{triângulo}) \geq 3\phi(\text{caminho})^2. \tag{7}$$

Como todo  $\phi \in \Phi$  deve satisfazer as condições (6) e (7), o valor ótimo do seguinte programa semidefinido é um limitante superior para  $\text{ex}(\text{caminho}, \{\text{triângulo}\}) = \sup\{\phi(\text{caminho}) : \phi \in \Phi\}$ :

$$\begin{aligned} \max \quad & \phi(\text{caminho}) \\ \text{s.a.} \quad & \phi(\text{triângulo}) + 2\phi(\text{caminho}) = 3\phi(\text{caminho}); \tag{6} \\ & \phi(\text{triângulo}) \geq 3\phi(\text{caminho})^2; \tag{7} \\ & \phi \in [0, 1]^{\mathcal{G}} \end{aligned}$$

Note que qualquer solução dual do programa semidefinido acima fornece cotas superiores para  $\text{ex}(\text{caminho}, \{\text{triângulo}\})$ . Em particular, fazendo  $2 \cdot (7) - (6)$ , obtemos a desigualdade

$$0 \leq \phi(\text{triângulo}) \leq 3\phi(\text{caminho}) - 6\phi(\text{caminho})^2,$$

da onde segue que  $2\phi(\text{caminho})^2 \leq \phi(\text{caminho})$ , isto é, que  $\phi(\text{caminho}) \leq \frac{1}{2}$ , como desejado.  $\square$

◇ ◇ ◇

Aula 14 (07 de Junho) — Fernando Mario de Oliveira Filho

◇ ◇ ◇

## 8.2 Definições iniciais

Fixe uma família  $\mathcal{H}$  de grafos proibidos e seja  $\mathcal{G}$  o conjunto de grafos  $\mathcal{H}$ -livres.

Um *tipo* de tamanho  $k$  é um grafo em  $\mathcal{G}$  com vértices em  $[k]$ . Dado um tipo  $\sigma$  de tamanho  $k$  e  $F \in \mathcal{G}$ , definimos uma *imersão* de  $\sigma$  em  $F$  como uma função injetora  $\theta: [k] \rightarrow V(F)$  que define um isomorfismo entre os grafos  $\sigma$  e  $F[\text{Im} \theta]$ .

Uma  $\sigma$ -*flag* é um par  $(F, \theta)$  onde  $F \in \mathcal{G}$  é um grafo e  $\theta$  é uma imersão de  $\sigma$  em  $F$ . Por simplicidade, às vezes omitimos a imersão  $\theta$  e dizemos apenas que  $F$  é uma  $\sigma$ -flag. Duas  $\sigma$ -flags  $(F_1, \theta_1)$  e  $(F_2, \theta_2)$  são *isomorfas* se existe um isomorfismo  $\xi$  entre  $F_1$  e  $F_2$  tal que  $\xi(\theta_1(i)) = \xi(i)$  para todo  $1 \leq i \leq |\sigma|$ . No que se segue, faremos j

O conjunto de todas as  $\sigma$ -flags de tamanho  $n$  será denotado por  $\mathcal{F}_n^\sigma$ . Também definimos  $\mathcal{F}^\sigma = \bigcup_{n \geq |\sigma|} \mathcal{F}_n^\sigma$  como o conjunto de todas as  $\sigma$ -flags.

Dadas  $\sigma$ -flags  $F_1, \dots, F_t$  e um inteiro  $n$ , dizemos que  $F_1, \dots, F_t$  cabem em  $n$  se

$$n - |\sigma| \geq \sum_{i=1}^t (|F_i| - |\sigma|).$$

Também dizemos que  $F_1, \dots, F_t$  cabem em uma  $\sigma$ -flag  $G$  se  $F_1, \dots, F_t$  cabem em  $|G|$ .

Sejam  $F_1, \dots, F_t, G \in \mathcal{F}^\delta$  tais que  $F_1, \dots, F_t$  cabem em  $G$ , definimos

$$p(F_1, \dots, F_t; G) = \mathbb{P}(F_i \simeq G[U_i \cup \text{Img} \theta]),$$

onde  $\theta$  é a imersão de  $\sigma$  e onde  $U_1, \dots, U_t \subseteq V(G) \setminus \text{Img} \theta$  são conjuntos dois-a-dois disjuntos, de cardinalidade  $|U_i| = |F_i| - |\sigma|$ , escolhidos uniformemente ao acaso.

**Teorema 8.4** (Regra da Cadeia). *Sejam  $F_1, \dots, F_t, G \in \mathcal{F}^\sigma$  tais que  $F_1, \dots, F_t$  cabem em  $G$ . Seja  $n$  um inteiro tal que para todo  $1 \leq s \leq t$ , vale que  $F_1, \dots, F_s$  cabem em  $n$  e que  $F_{s+1}, \dots, F_t, H$  cabem em  $G$ , onde  $H$  é uma  $\sigma$ -flag de tamanho  $n$ . Então*

$$p(F_1, \dots, F_t; G) = \sum_{F' \in \mathcal{F}_n^\sigma} p(F_1, \dots, F_s; F') p(F', F_{s+1}, \dots, F_t; G). \quad \square$$

O próximo resultado mostra que  $p(F_1, F_2; G)$  está arbitrariamente próximo de  $p(F_1; G)p(F_2; G)$  para  $\sigma$ -flags  $G$  de tamanho suficientemente grande.

**Teorema 8.5.** *Para todo  $F_1, F_2 \in \mathcal{F}^\sigma$ , existe  $f(n) = O(\frac{1}{n})$  tal que se  $F_1, F_2$  cabem em  $G$  então*

$$|p(F_1, F_2; G) - p(F_1; G)p(F_2; G)| \leq f(|G|). \quad \square$$

### 8.3 Flag Álgebra

Uma sequência de  $\sigma$ -flags  $(A_k)_{k \geq 0}$  é convergente se para todo  $F \in \mathcal{F}^\sigma$  existe  $\lim_{k \rightarrow \infty} p(F; A_k)$ .

Seja

$$\mathbb{R}\mathcal{F}^\sigma = \{f : \mathcal{F}^\sigma \rightarrow \mathbb{R} : |\sup f| < \infty\}.$$

o espaço vetorial sobre  $\mathbb{R}$  das funções de  $\mathcal{F}^\sigma \rightarrow \mathbb{R}$  com suporte finito. Representaremos um elemento  $f \in \mathbb{R}\mathcal{F}^\sigma$  como uma soma formal (finita) da forma  $\sum_{F \in \mathcal{F}^\sigma} f(F)F$ .

Um *funcional linear* é qualquer função linear  $\phi : \mathbb{R}\mathcal{F}^\sigma \rightarrow \mathbb{R}$ . Um funcional linear  $\phi$  é um *funcional limite* se para toda sequência convergente de  $\sigma$ -flags  $(A_k)_{k \geq 0}$ , temos

$$\phi(F) = \lim_{k \rightarrow \infty} p(F; A_k),$$

para todo  $F \in \mathcal{F}^\sigma$ . Note que como  $\mathcal{F}^\sigma$  é uma base de  $\mathbb{R}\mathcal{F}^\sigma$  (como espaço linear sobre  $\mathbb{R}$ ), um funcional linear é completamente determinado pela sequência convergente  $(A_k)_{k \geq 0}$ .

Como consequência da Regra da Cadeia, certos elementos de  $\mathcal{F}^\sigma$  devem ter a mesma imagem em  $\mathbb{R}$  para qualquer funcional limite  $\phi$ . Mais especificamente, temos o seguinte resultado.

**Proposição 8.6.** *Sejam  $F \in \mathcal{F}^\sigma$  uma flag e  $\phi : \mathbb{R}\mathcal{F}^\sigma \rightarrow \mathbb{R}$  um funcional limite. Para todo  $n \geq |F|$  temos:*

$$\phi(F) = \phi\left(\sum_{F' \in \mathcal{F}_n^\sigma} p(F; F')F'\right).$$

*Demonstração.* Seja  $(A_k)_{k \geq 0}$  uma seqüência convergente que determina o funcional limite  $\phi$ . Temos

$$\begin{aligned} \phi\left(\sum_{F' \in \mathcal{F}_n^\sigma} p(F; F')F'\right) &= \sum_{F' \in \mathcal{F}_n^\sigma} p(F; F')\phi(F') \\ &= \sum_{F' \in \mathcal{F}_n^\sigma} p(F; F') \lim_{k \rightarrow \infty} p(F'; A_k) \\ &= \lim_{k \rightarrow \infty} \sum_{F' \in \mathcal{F}_n^\sigma} p(F; F')p(F'; A_k) \\ &= \lim_{k \rightarrow \infty} p(F; A_k) \\ &= \phi(F). \end{aligned}$$

□

A seguir, iremos considerar o espaço vetorial  $\mathcal{A}^\sigma$  obtido (a partir de  $\mathbb{R}\mathcal{F}^\sigma$ ) ao identificarmos quaisquer elementos que, por consequência da Proposição 8.6, devem ter a mesma imagem para *qualquer* funcional limite.

Mais formalmente, definimos o espaço linear

$$\mathcal{K}^\sigma = \left\langle \left\{ F - \sum_{F' \in \mathcal{F}_n^\sigma} p(F; F')F' : F \in \mathcal{F}^\sigma, n \geq |F| \right\} \right\rangle.$$

Note que para todo funcional linear  $\phi$ , a Proposição 8.6 implica  $\mathcal{K}^\sigma \subseteq \text{Ker } \phi$ . Agora consideramos o espaço  $\mathcal{A}^\sigma := \mathbb{R}\mathcal{F}^\sigma / \mathcal{K}^\sigma$ , obtido fazendo o quociente de  $\mathbb{R}\mathcal{F}^\sigma$  por  $\mathcal{K}^\sigma$ . Dado  $F \in \mathbb{R}\mathcal{F}^\sigma$  denotaremos o elemento  $(F + \mathcal{K}^\sigma) \in \mathbb{R}\mathcal{F}^\sigma / \mathcal{K}^\sigma$  apenas por  $F$ , de forma a simplificar a notação.

Seja  $\phi : \mathbb{R}\mathcal{F}^\sigma \rightarrow \mathbb{R}$  um funcional limite. Como  $\mathcal{K}^\sigma \subseteq \text{Ker } \phi$ , então  $\phi$  é [pode ser estendido a] um funcional limite em  $\mathcal{A}^\sigma$ .

Definiremos, a seguir, uma operação bilinear de multiplicação  $\cdot : \mathcal{A}^\sigma \times \mathcal{A}^\sigma \rightarrow \mathcal{A}^\sigma$  de forma a obter uma álgebra  $(\mathcal{A}^\sigma, \cdot)$ . Dadas duas flags  $F_1, F_2 \in \mathcal{F}^\sigma$ , definimos

$$F_1 \cdot F_2 = \sum_{F \in \mathcal{F}_n^\sigma} p(F_1, F_2; F)F \in \mathcal{A}^\sigma.$$

onde  $n$  é um inteiro arbitrário tal que  $F_1$  e  $F_2$  cabem em  $n$ . Não é difícil mostrar que esse produto está bem definido em  $\mathcal{A}^\sigma$ , isto é, independe da escolha de  $n$ . De fato,

dados inteiros  $n > m$  tais que  $F_1$  e  $F_2$  cabem em, ambos,  $n$  e  $m$ , temos

$$\begin{aligned} \sum_{H \in \mathcal{F}_m^\sigma} p(F_1, F_2; H)H &= \sum_{H \in \mathcal{F}_m^\sigma} p(F_1, F_2; H) \left( \sum_{F \in \mathcal{F}_n^\sigma} p(H; F)F \right) \\ &= \sum_{F \in \mathcal{F}_n^\sigma} F \sum_{H \in \mathcal{F}_m^\sigma} p(F_1, F_2; H)p(H; F) \\ &= \sum_{F \in \mathcal{F}_n^\sigma} p(F_1, F_2; F)F. \end{aligned}$$

Uma vez definida a operação  $\cdot$  para flags, podemos estendê-la bilinearmente para uma operação  $\cdot : \mathbb{R}\mathcal{F}^\sigma \times \mathbb{R}\mathcal{F}^\sigma \rightarrow \mathcal{A}^\sigma$ . O resultado a seguir mostra que a extensão natural de  $\cdot$  para uma operação em  $\mathcal{A}^\sigma \times \mathcal{A}^\sigma \rightarrow \mathcal{A}^\sigma$  também está bem definida.

**Proposição 8.7.** *Seja  $f \in \mathcal{K}^\sigma$  e  $g \in \mathbb{R}\mathcal{F}^\sigma$ . Então  $f \cdot g = 0 \in \mathcal{A}^\sigma$ .*

*Demonstração.* Vamos considerar o caso particular em que existem  $F, G \in \mathcal{F}^\sigma$  e inteiro  $m$  tais que  $f = F - \sum_{F' \in \mathcal{F}_m^\sigma} p(F; F')F'$  e  $g = G$ . Temos

$$\begin{aligned} f \cdot g &= F \cdot G - \sum_{F' \in \mathcal{F}_m^\sigma} p(F; F')(F' \cdot G) \\ &= F \cdot G - \sum_{F' \in \mathcal{F}_m^\sigma} p(F; F') \sum_{H \in \mathcal{F}_n^\sigma} p(F', G; H)H \\ &= F \cdot G - \sum_{H \in \mathcal{F}_m^\sigma} p(F, G; H)H = 0. \end{aligned}$$

O caso geral segue por linearidade. □

Logo  $\cdot$  é um produto comutativo em  $\mathcal{A}^\sigma$ . A álgebra  $(\mathcal{A}^\sigma, \cdot)$  será denotada apenas por  $\mathcal{A}^\sigma$ .

**Proposição 8.8.** *O elemento neutro de  $\mathcal{A}^\sigma$  é  $\sigma \in \mathcal{A}$ .*

*Demonstração.* Basta provar que para toda flag  $F \in \mathcal{F}_n^\sigma$  temos  $F \cdot \sigma = F$ . De fato,

$$F \cdot \sigma = \sum_{H \in \mathcal{F}_{n+1}^\sigma} p(F, \sigma; H)H = \sum_{H \in \mathcal{F}_{n+1}^\sigma} p(F; H)H = F.$$

□

Seja  $\text{Hom}(A^\sigma, \mathbb{R})$  o conjunto de homomorfismos  $\phi : A^\sigma \rightarrow \mathbb{R}$  e  $\text{Hom}^+(A^\sigma, \mathbb{R})$  o conjunto de homomorfismos  $\phi : A^\sigma \rightarrow \mathbb{R}$  tais que  $\phi(F) \geq 0$  para todo  $F \in \mathcal{F}^\sigma$ .

**Proposição 8.9.** *Se  $\phi : A^\sigma \rightarrow \mathbb{R}$  é um funcional limite, então  $\phi \in \text{Hom}^+(A^\sigma, \mathbb{R})$ .*

*Demonstração.* Seja  $(A_k)_{k \geq 0}$  uma sequência convergente tal que  $\phi(F) = \lim_{k \rightarrow \infty} p(F; A_k) \geq 0$  para todo  $F \in \mathcal{F}^\sigma$ . Basta mostrar que  $\phi$  é homomorfismo.

- $\phi(\sigma) = \lim_{k \rightarrow \infty} p(\sigma, A_k) = 1$ .



Logo, todo conjunto  $\mathcal{C} \subseteq \mathcal{S}^\emptyset$  fornece a seguinte cota superior para  $\text{ex}(C, \mathcal{H})$ :

$$\text{ex}(C, \mathcal{H}) \leq \min\{\lambda : \lambda \emptyset - C \in \mathcal{C}\}.$$

## 8.4 Método semidefinido

Um método para obtermos um conjunto (não-trivial)  $\mathcal{C} \subseteq \mathcal{S}^\emptyset$  consiste em considerar elementos  $f \in \mathcal{A}^\emptyset$  que podem ser expressos como soma de quadrados, isto é, elementos da forma  $f = g_1^2 + \dots + g_t^2$ . De fato,  $\phi(f) = \sum_{i=1}^t \phi(g_i)^2 \geq 0$  para qualquer  $\phi \in (\mathcal{A}^\emptyset)^* \supseteq \text{Hom}^+(\mathcal{A}^\emptyset)$ . Infelizmente, conjuntos  $\mathcal{C}$  assim obtidos não fornecem, em geral, uma boa cota superior para  $\text{ex}(C, \mathcal{H})$ . Ao invés disso, fixaremos um tipo  $\sigma$  e introduziremos um operador  $\llbracket \cdot \rrbracket_\sigma$  que projeta elementos de  $\mathcal{A}^\sigma$  em elementos de  $\mathcal{A}^\emptyset$  com a propriedade que  $\llbracket \mathcal{S}^\sigma \rrbracket_\sigma \subseteq \mathcal{S}^\emptyset$ . Dessa forma, poderemos considerar conjuntos  $\mathcal{C} = \llbracket \mathcal{C}^\sigma \rrbracket_\sigma \subseteq \mathcal{S}^\emptyset$ , onde  $\mathcal{C}^\sigma \subseteq \mathcal{S}^\sigma$  é o conjuntos das somas de quadrados em  $\mathcal{A}^\sigma$ .

Defina  $\llbracket \cdot \rrbracket_\sigma : \mathbb{R}\mathcal{F}^\sigma \rightarrow \mathbb{R}\mathcal{F}^\emptyset$  como o operador linear que satisfaz, para toda flag  $F \in \mathcal{F}^\sigma$ ,

$$\llbracket F \rrbracket_\sigma = q_\sigma(F)A \downarrow F,$$

onde  $\downarrow F \in \mathcal{F}^\emptyset$  é a flag obtida a partir de  $F$  desconsiderando o tipo  $\sigma$  e  $q_\sigma(F)$  é a probabilidade de uma injeção  $\theta : [k] \rightarrow V(F)$  escolhida uniformemente ao acaso ser tal que  $(\downarrow F, \theta) \simeq F$ .

**Proposição 8.11.**  $\llbracket \mathcal{K}^\sigma \rrbracket_\sigma \subseteq \mathcal{K}^\emptyset$

*Demonstração.* Seja  $F - \sum_{H \in \mathcal{F}_n^\sigma} p(F; H)H \in \mathcal{K}^\sigma$ . Temos

$$\begin{aligned} \left\llbracket F - \sum_{H \in \mathcal{F}_n^\sigma} p(F; H)H \right\rrbracket_\sigma &= q_\sigma(F) \downarrow F - \sum_{H \in \mathcal{F}_n^\sigma} p(F; H) q_\sigma(H) \downarrow H \\ &= q_\sigma(F) \downarrow F - \sum_{H^\emptyset \in \mathcal{F}_n^\emptyset} \left( \sum_{\substack{H \in \mathcal{F}_n^\sigma \\ \downarrow H = H^\emptyset}} p(F; H) q_\sigma(H) \right) H^\emptyset \\ &= q_\sigma(F) \downarrow F - q_\sigma(F) \sum_{H^\emptyset \in \mathcal{F}_n^\emptyset} p(\downarrow F; H^\emptyset) H^\emptyset \in \mathcal{K}^\emptyset. \end{aligned}$$

□

Segue da proposição acima que  $\llbracket \cdot \rrbracket_\sigma$  está bem definido como um operador linear  $\llbracket \cdot \rrbracket_\sigma : \mathcal{A}^\sigma \rightarrow \mathcal{A}$ .

**Teorema 8.12.**  $\llbracket \mathcal{S}^\sigma \rrbracket_\sigma \subseteq \mathcal{S}^\emptyset$

□

A seguir veremos como elementos que são somas de quadrados podem ser codificados como matrizes positivas semidefinidas (p.s.d.)  $Q \in \mathcal{F}_n^\sigma \times \mathcal{F}_n^{\sigma_n} \rightarrow \mathbb{R}$ . Seja  $0 \neq g \in \mathbb{R}\mathcal{F}^\sigma$ . Definimos o grau de  $g$  como  $\text{deg } g = \max\{|F| : g(F) \neq 0\}$ . Para elementos  $f \in \mathcal{A}^\sigma$  definimos o grau de  $f$  como o menor grau dentre todos os elementos

equivalentes a  $f$ , isto é,  $\deg f = \min\{\deg g : f = g + K^\sigma, g \in \mathbb{R}\mathcal{F}^\sigma\}$ . Defina também

$$\begin{aligned} v_{\sigma,n} : \mathcal{F}^\sigma &\rightarrow \mathcal{A}^\sigma \\ F &\mapsto F. \end{aligned}$$

**Teorema 8.13.** *Sejam  $f \in \mathcal{A}^\sigma$  e  $n \geq |\sigma|$ . Para algum  $t \geq 1$  existem elementos  $g_1, \dots, g_t \in \mathcal{A}^\sigma$ , todos de grau no máximo  $n$ , tais que  $f = g_1^2 + \dots + g_t^2$  se, e somente se existe uma matriz p.s.d.  $Q : \mathcal{F}_n^\sigma \times \mathcal{F}_n^\sigma \rightarrow \mathbb{R}$  tal que  $f = v_{\sigma,n} Q v_{\sigma,n}$ .*

*Demonstração.* Suponha que  $f = g_1^2 + \dots + g_t^2$  com  $\deg g_i \leq n$ . Usando a regra da cadeia, podemos expressar cada  $g_i$  como  $g_i = \sum_{F' \in \mathcal{F}_n^\sigma} \alpha_{i,F'} F'$ . Seja  $c_i = (\alpha_{i,F'})_{F' \in \mathcal{F}_n^\sigma}$  o vetor de coeficientes tal que  $g_i = c_i^t v_{\sigma,n}$ . Temos

$$f = \sum_{i=1}^t g_i^2 = \sum_{i=1}^t (c_i^t v_{\sigma,n})^2 = \sum_{i=1}^t (v_{\sigma,n}^t c_i) (c_i^t v_{\sigma,n}) = \sum_{i=1}^t v_{\sigma,n} Q v_{\sigma,n},$$

onde  $Q = \sum_{i=1}^t (c_i c_i^t)$  e, portanto, p.s.d.

Por outro lado, suponha que  $Q$  é uma matriz p.s.d. de posto  $t$  tal que  $f = v_{\sigma,n}^t Q v_{\sigma,n}$ . Então  $Q = \sum_{i=1}^t c_i c_i^t$  para certos vetores de coeficientes  $c_i$ . Defina  $g_i = c_i^t v_{\sigma,n}$ . De modo análogo ao feito anteriormente, devemos ter  $f = \sum_{i=1}^t g_i^2$ .  $\square$

Fixe um tipo  $\sigma$  um inteiro  $n \geq |\sigma|$  e seja  $v = v_{\sigma,n}$ . Considere o conjunto

$$\mathcal{C} := \{r + \llbracket v^t Q v \rrbracket_\sigma : Q \in \mathcal{F}_n^\sigma \times \mathcal{F}_n^\sigma \rightarrow \mathbb{R} \text{ p.s.d. e } r \in \mathcal{S}_n^\sigma\}.$$

Como o conjunto  $\{v^t Q v : Q \in \mathcal{F}_n^\sigma \times \mathcal{F}_n^\sigma \rightarrow \mathbb{R} \text{ p.s.d.}\}$  é composto por elementos que são somas de quadrados (Teorema 8.13) e, portanto, pertencentes a  $\mathcal{S}^\sigma$ , segue do Teorema 8.12 que  $\mathcal{C} \subseteq \mathcal{S}^\sigma$ . Logo, podemos obter um limitante superior para  $\text{ex}(\mathcal{C}, \mathcal{H})$  resolvendo o seguinte problema de otimização

$$\begin{aligned} \min \lambda \\ \text{s.a. } \lambda \emptyset - C &= r + \llbracket v^t Q v \rrbracket_\sigma \\ r &\in \mathcal{S}^\sigma \\ Q &\in \mathcal{F}_n^\sigma \times \mathcal{F}_n^\sigma \rightarrow \mathbb{R} \text{ p.s.d.} \end{aligned}$$

Dadas matrizes  $A, B$  definimos o produto inteiro  $\langle A, B \rangle = A^t B$ . Em particular, para todo vetor  $v$  vale que  $v^t Q v = \langle v v^t, Q \rangle$ . Logo, podemos reescrever  $\llbracket v^t Q v \rrbracket_\sigma$  como

$$\llbracket v^t Q v \rrbracket_\sigma = \llbracket \langle v v^t, Q \rangle \rrbracket_\sigma = \langle \llbracket v v^t \rrbracket_\sigma, Q \rangle.$$

Fixe  $N \geq n$  e seja  $(\lambda, Q, r)$  uma solução do problema de otimização acima. Toda flag  $G \in \mathcal{F}_N^\sigma$  deve satisfazer

$$p(\lambda \emptyset - C) = p(r; G) + \langle p(\llbracket v v^t \rrbracket_\sigma; G), Q \rangle,$$

onde  $p(\llbracket v v^t \rrbracket_\sigma; G)$  é a matriz  $(p(\llbracket v v^t \rrbracket_\sigma(F, F'); G))_{F, F' \in \mathcal{F}_N^\sigma}$ . Como  $p(r; G) \geq 0$ , o seguinte

programa semidefinido fornece uma cota superior para  $\text{ex}(C; \mathcal{H})$ .

$$\begin{aligned} & \min \lambda \\ & \text{s.a. } \lambda - p(C; G) \geq \langle p(\llbracket vv^t \rrbracket_\sigma; G), Q \rangle \quad \forall G \in \mathcal{F}_N^\sigma \\ & \quad Q \in \mathcal{F}_n^\sigma \times \mathcal{F}_n^\sigma \rightarrow \mathbb{R} \text{ p.s.d.} \end{aligned} \quad (8)$$

Note que fixados  $\sigma, n, N$ , podemos obter computacionalmente cada matriz  $p(\llbracket vv^t \rrbracket_\sigma; G)$  (para todo  $G \in \mathcal{F}_N^\sigma$ ) e então computar o valor do programa acima. Desta forma, obtemos cotas superiores para  $\text{ex}(C, \mathcal{H})$  de uma forma completamente mecânica.

## 8.5 Exemplo método semidefinido: Teorema de Mantel

Seja  $\mathcal{H} = \{\triangle\}$  e  $C = \circ - \circ$ . Usaremos o programa 8 com  $n = 2, N = 3$  e  $\sigma = \bullet$  para obter uma cota superior para  $\text{ex}(\circ - \circ, \{\triangle\})$ .

Temos  $v = v_{\sigma, n} = \{\bullet \cdots \circ, \bullet \circ\}$  e

$$\begin{aligned} \llbracket vv^t \rrbracket_\sigma &= \left\| \left( \begin{array}{cc} \bullet \cdots \circ & \bullet \circ \\ \bullet \circ & \bullet \cdots \circ \end{array} \right) \right\|_\sigma = \left\| \left( \begin{array}{cc} \left( \triangle + \frac{1}{3} \bullet \triangle \right) & \frac{1}{3} \triangle + \frac{1}{3} \bullet \triangle \\ \frac{1}{3} \triangle + \frac{1}{3} \bullet \triangle & \frac{1}{3} \triangle \end{array} \right) \right\|_\sigma \\ &= \left( \begin{array}{cc} \triangle + \frac{1}{3} \bullet \triangle & \frac{1}{3} \triangle + \frac{1}{3} \bullet \triangle \\ \frac{1}{3} \triangle + \frac{1}{3} \bullet \triangle & \frac{1}{3} \triangle \end{array} \right). \end{aligned}$$

Agora podemos calcular as matrizes  $p(\llbracket vv^t \rrbracket_\sigma; G)$  para todo  $G \in \mathcal{F}_3^\sigma = \{\triangle, \triangle, \triangle\}$ :

$$\begin{aligned} p(\llbracket vv^t \rrbracket_\sigma; \triangle) &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \\ p(\llbracket vv^t \rrbracket_\sigma; \triangle) &= \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 \end{pmatrix}, \\ p(\llbracket vv^t \rrbracket_\sigma; \triangle) &= \begin{pmatrix} 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}. \end{aligned}$$

Também temos  $p(\circ - \circ; \triangle) = 0$ ,  $p(\circ - \circ; \triangle) = \frac{1}{3}$  e  $p(\circ - \circ; \triangle) = \frac{2}{3}$ . Logo, o programa 8 neste caso particular é:

$$\begin{aligned} & \min \lambda \\ & \text{s.a. } \lambda \geq \left\langle \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, Q \right\rangle \\ & \quad \lambda - \frac{1}{3} \geq \left\langle \begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 \end{pmatrix}, Q \right\rangle \\ & \quad \lambda - \frac{2}{3} \geq \left\langle \begin{pmatrix} 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}, Q \right\rangle \\ & \quad Q \in \mathcal{F}_2^\sigma \times \mathcal{F}_2^\sigma \rightarrow \mathbb{R} \text{ p.s.d.} \end{aligned}$$

Uma solução ótima para esse problema é  $\lambda = \frac{1}{2}, Q = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ . Neste caso, obtemos

uma solução que fornece a melhor cota possível para  $\text{ex}(\text{---}, \text{---}) = \frac{1}{2}$ .

## 9 Teorema de Erdős, Ginzburg e Ziv

◇ ◇ ◇

Aula 16 (21 de Junho) — Yoshiharu Kohayakawa

◇ ◇ ◇

**Teorema 9.1** (Erdős, Ginzburg e Ziv 1964). *Toda sequência de  $2n - 1$  inteiros contém uma subsequência de  $n$  elementos cuja soma é divisível por  $n$ .*

Veremos a seguir que o Teorema 9.1 segue da versão abaixo que se restringe ao caso em que  $n$  é primo.

**Lema 9.2.** *Seja  $p$  um inteiro primo. Toda sequência de  $2p - 1$  inteiros contém uma subsequência de  $p$  elementos cuja soma é divisível por  $p$ .*

*Demonstração do Teorema 9.1.* Usamos indução em  $n$ . Se  $n = 1$ , então o resultado é trivial. Suponha que  $n > 1$  e que o teorema vale para inteiros menores que  $n$ . Seja  $a_1, \dots, a_{2n-1}$  uma sequência dada e seja  $p$  um primo que divide  $n$ . Assumimos que  $p < n$ , caso contrário o resultado segue do Lema 9.2. Seja  $m = n/p$ .

A seguir, afirmamos que é possível encontrar conjuntos  $I_1, \dots, I_{2m-1} \subseteq [2n - 1]$  dois-a-dois disjuntos e todos de cardinalidade  $p$  tais que

$$b_j := \sum_{i \in I_j} a_i \equiv 0 \pmod{p}$$

para todo  $j = 1, \dots, 2m - 1$ . De fato, suponha já definidos  $I_1, \dots, I_l$ , com  $l < 2m - 1$ . Temos

$$\left| [2n - 1] \setminus \bigcup_{j \leq l} I_j \right| = 2n - 1 - lp \geq 2n - 1 - (2m - 2)p = 2p - 1.$$

Logo, pelo Lema 9.2 existe  $I_{l+1} \subseteq [2n - 1] \setminus \bigcup_{j \leq l} I_j$  tal que  $b_{l+1} = \sum_{i \in I_{l+1}} a_i \equiv 0 \pmod{p}$ , como requerido.

Considere agora a sequência  $(c_j)_{1 \leq j \leq 2m-1}$ , onde  $c_j := b_j/p$ . Por hipótese de indução (em  $m$ ) existe conjunto  $J \subseteq [2m - 1]$  de cardinalidade  $|J| = m$  tal que  $\sum_{j \in J} c_j \equiv 0 \pmod{m}$ . Mas então basta considerar o conjunto  $I = \bigcup_{j \in J} I_j$ . Note que  $|I| = |J|p = mp = n$ . Ademais, temos

$$\sum_{i \in I} a_i = \sum_{j \in J} \sum_{i \in I_j} a_i = p \sum_{j \in J} \frac{1}{p} \sum_{i \in I_j} a_i = p \sum_{j \in J} c_j \equiv 0 \pmod{n},$$

uma vez que  $\sum_{j \in J} c_j \equiv 0 \pmod{m}$ . □

A seguir veremos três provas para o Lema 9.2.

### 9.1 Primeira prova do Lema 9.2

Seja dada uma sequência  $a_1, \dots, a_{2p-1}$ . No que se segue supomos que

$$0 \leq a_1 \leq \dots \leq a_{2p-1} < p,$$

e trabalhamos sobre  $\mathbb{Z}_p = \mathbb{Z}/p\mathbb{Z}$ .

Se tivermos  $a_i = a_{i+p-1}$ , para algum  $1 \leq i \leq p$ , então a subsequência  $a_i, \dots, a_{i+p-1}$  é tal que

$$a_i + \dots + a_{i+p-1} = pa_i \equiv 0 \pmod{p}.$$

Assim, supomos  $a_i \neq a_{i+p-1}$  para todo  $1 \leq i \leq p$ . Para todo  $i = 1, 2, \dots, p-1$ , defina  $A_i = \{a_i, a_{i+p-1}\}$  e

$$B_i = A_1 + \dots + A_i,$$

onde definimos  $A + B := \{a + b : a \in A, b \in B\}$  (soma de Minkowski).

Afirmamos que  $|B_i| \geq i + 1$  para todo  $i = 1, 2, \dots, p-1$ . De fato, como  $|B_{i+1}| \geq |B_i|$  para todo  $i$ , se a afirmação não for válida deve existir algum  $k < p-1$  tal que  $|B_k| = |B_{k+1}|$ . Por simplicidade, ponha  $B = B_k$  e  $A_{k+1} = \{c, d\}$ . Como  $|B_{k+1}| = |B|$  e  $B_{k+1} = (B + c) \cup (B + d)$ , devemos ter  $B + c = B + d$ . Seja  $C = B + c$ . Note que  $C$  é invariante por somar  $e := d - c$ . Em geral temos

$$C = C + e = C + 2e = \dots C + (p-1)e.$$

Mas como  $e \neq 0$ , segue que  $C = \mathbb{Z}_p$ .

Agora, como  $|B_i| \geq i + 1$ , então  $B_{p-1} = \mathbb{Z}_p$ . Em particular,  $-a_{2p-1} \in B_{p-1}$ . Logo existem índices  $k_i \in \{i, i+p-1\}$  (para  $i = 1, 2, \dots, p-1$ ) tais que

$$-a_{2p-1} \equiv a_{k_1} + \dots + a_{k_{p-1}} \pmod{p},$$

isto é,

$$a_{k_1} + \dots + a_{k_{p-1}} + a_{2p-1} = 0 \pmod{p},$$

como desejado. □

## 9.2 Segunda prova do Lema 9.2

Seja

$$S = \sum_{\substack{I \subseteq [2p-1] \\ |I|=p}} \left( \sum_{i \in I} a_i \right)^{p-1}.$$

Podemos expressar  $S$  como uma soma de monômios (considerando que cada  $a_i$  é uma indeterminada) da forma

$$c_{\mathbf{k}} \prod_{i \in [2p-1]} a_i^{k_i}$$

onde  $\mathbf{k} = (k_1, \dots, k_{2p-1})$  é tal que  $\sum_{i=1}^{2p-1} k_i = p-1$  e  $1 \leq |\mathbf{k}| \leq p-1$ , onde  $|\mathbf{k}| := \#\{i : k_i \neq 0\}$  denota o número de entradas não-nulas de  $\mathbf{k}$ .

Fixe  $\mathbf{k}$  e seja  $\ell = |\mathbf{k}|$ . Observe que os conjuntos  $I$  que contribuem para  $c_{\mathbf{k}}$  são aqueles que satisfazem  $I \supseteq \text{supp}(\mathbf{k})$ , onde  $\text{supp}(\mathbf{k}) = \{i : k_i > 0\}$  denota o conjunto das entradas não-nulas de  $\mathbf{k}$ . É fácil ver que o número de tais conjuntos  $I$  satisfaz

$$\#\{I : I \supseteq \text{supp}(\mathbf{k})\} = \binom{2p-1-\ell}{p-\ell}$$

e que tais conjuntos  $I$  contribuem, todos, com a mesma quantidade para cada  $c_{\mathbf{k}}$ . Logo

todo  $c_k$  é múltiplo de  $\binom{2p-1-\ell}{p-\ell}$ . Mas como

$$\binom{2p-1-\ell}{p-\ell} = \frac{(2p-1)(2p-2)\cdots p}{(p-1)\cdots 1},$$

segue que  $p|c_k$  e, portanto, que  $S \equiv 0 \pmod{p}$ .

Por outro lado, suponha por absurdo que a conclusão do Lema 9.2 seja falsa. Então para todo  $I \subseteq [2p-1]$  com  $|I| = p$  temos  $\sum_{i \in I} a_i \not\equiv 0 \pmod{p}$ . Logo, pelo Pequeno Teorema de Fermat devemos ter

$$\left( \sum_{i \in I} a_i \right)^{p-1} \equiv 1 \pmod{p}$$

o que implica

$$S = \sum_{\substack{I \subseteq [2p-1] \\ |I|=p}} 1 = \binom{2p-1}{p} \equiv \binom{1}{1} \binom{p-1}{0} \equiv 1 \pmod{p},$$

contradizendo o fato que  $S = 0$ . □