



# RMU

**REVISTA MATEMÁTICA UNIVERSITÁRIA**

ISSN: 2675-5254

Ano 2021, Volume 1

## Índice

Ivo Terek	
<i>Mergulhos clássicos de variedades Grassmannianas: uma visão geral</i> .....	1
Ágda Talita Galvão e Jocemar de Quadros Chagas	
<i>Uma introdução a somatórios fracionários: aprendendo a somar uma quantidade não inteiras de parcelas</i> .....	15
José Claudinei Ferreira	
<i>Quando os métodos de Euler e de Newton coincidem</i> .....	34
Ton Marar	
<i>Modelos do plano projetivo no espaço tridimensional</i> .....	47
Marcelo Bongarti e German Lozada-Cruz	
<i>Alguns teoremas do tipo valor médio: De Lagrange a Malesevich</i> ....	56



# SBM

**SOCIEDADE BRASILEIRA DE MATEMÁTICA**

## **SOBRE A RMU**

A Matemática Universitária é uma publicação da Sociedade Brasileira de Matemática semestral de divulgação de ideias e estímulos ao estudo e à curiosidade intelectual, dirigida a todos que se interessam pelo ensino e estudo da Matemática em nível Superior. É direcionada a professores, pesquisadores, alunos de graduação e pós graduação, promovendo e fortalecendo o intercâmbio entre os membros dessa comunidade.

### **Comitê Editorial**

- Editor-chefe:
  - Paolo Piccione (USP)
- Editores:
  - Humberto Bortolossi (UFF)
  - Daniel Gonçalves (UFSC)
  - Fernando Manfio (USP)
  - Michel Spira (UFMG)

### **Periodicidade**

Semestral

### **Contato**

*Endereço:*

Sociedade Brasileira de Matemática  
Estrada Dona Castorina, 110 sala 109  
Jardim Botânico  
22460-320 Rio de Janeiro – RJ

*Email:* [rmu@sbm.org.br](mailto:rmu@sbm.org.br)

*Página web:* <https://rmu.sbm.org.br/>



## MERGULHOS CLÁSSICOS DE VARIEDADES GRASSMANNIANAS: UMA VISÃO GERAL

IVO TEREK COUTO

RESUMO. Neste artigo discutiremos conceitos básicos sobre variedades Grassmannianas — em particular a motivação e o contexto histórico por trás do seu surgimento, como tais objetos aparecem em Geometria (Diferencial e Algébrica) hoje em dia, e alguns mergulhos isométricos de Grassmannianos em esferas e espaços projetivos.

### 1. INTRODUÇÃO E UM POUCO DE HISTÓRIA

Em Matemática, Geometria Diferencial é o estudo das chamadas variedades diferenciáveis, objetos que “localmente se parecem” com o espaço Euclidiano usual  $\mathbb{R}^n$  com o qual estamos acostumados, desde cursos básicos de Álgebra Linear. Trocando  $\mathbb{R}^n$  por  $\mathbb{C}^n$  e “diferenciabilidade” por “holomorfia”, inicia-se a discussão na categoria complexa. Em todo caso, sendo uma área ativa de pesquisa, é natural que cursos de Geometria Diferencial sejam frequentes e bem estabelecidos em programas de pós-graduação não apenas no Brasil, mas em universidades no mundo inteiro afora.

Neste texto, vamos brevemente explorar uma classe de variedades rica de simetrias e propriedades interessantes, injustamente negligenciada na maioria dos cursos introdutórios de variedades diferenciáveis e Geometria Riemanniana. Começamos com uma ideia simples: fixado de uma vez por todas um espaço vetorial  $V$  sobre um corpo  $\mathbb{K}$ , pode-se considerar o conjunto de todas as  $\mathbb{K}$ -retas em  $V$  passando pela origem. Explicitamente:

$$PV \doteq \{L \subseteq V \mid L \text{ é um subespaço vetorial de } V \text{ e } \dim L = 1\}.$$

Dizemos que  $PV$  é o *espaço projetivo associado a  $V$* . Quando  $V = \mathbb{K}^{n+1}$ , escrevemos simplesmente  $P(\mathbb{K}^{n+1}) = \mathbb{K}P^n$ . Uma generalização natural consiste em passar de retas (i.e., subespaços 1-dimensionais) para subespaços  $k$ -dimensionais, e assim passamos a considerar

$$\text{Gr}_k V \doteq \{W \subseteq V \mid W \text{ é um subespaço vetorial de } V \text{ e } \dim W = k\},$$

---

Data de aceitação: Setembro de 2020.

*Palavras chave.* Geometria Diferencial, Grassmannianos.

de modo que claramente temos  $PV = \text{Gr}_1V$ . Este conjunto  $\text{Gr}_kV$  é chamado o *Grassmanniano de  $k$ -planos de  $V$* , e será o protagonista desta história.

Agora, há dois fatos históricos particularmente surpreendentes e que merecem algum destaque. O primeiro deles, é que a ideia de um “espaço Grassmanniano” precede a formalização do conceito de variedade diferenciável! Com efeito, a primeira edição da Teoria da Extensão (*Ausdehnungslehre*) de Hermann Grassmann (1809–1877) foi publicada em 1844, enquanto a grande aula inaugural de Bernhard Riemann (1826–1866) na Universidade de Göttingen ocorreu apenas em 1854! O segundo deles, é que Grassmann nunca teve uma educação formal em Matemática e nunca foi professor universitário, tendo iniciado sua carreira como matemático relativamente “tarde” e daí em diante trabalhando como professor no que hoje mais ou menos equivale ao nosso Ensino Médio (o *gymnasium* alemão). Por outro lado, é seguro dizer que Grassmann foi um exemplo típico da versatilidade dos cientistas de antigamente, tendo investido muito de seu tempo também em Teologia, Filologia, Filosofia e Física.

Fica então a pergunta: se a ideia original de Grassmann não começou como sugerimos acima, qual foi a motivação? Surpreendentemente simples, foi representar uma partícula com massa  $m > 0$  localizada em coordenadas  $(x, y, z)$  em um espaço Euclidiano tridimensional como uma quádrupla  $(mx, my, mz, m)$ . Deste modo, o centro de massa de um sistema formado por duas tais partículas  $(m_i x_i, m_i y_i, m_i z_i, m_i)$ ,  $i = 1, 2$ , é descrito apenas pela soma

$$(m_1 x_1 + m_2 x_2, m_1 y_1 + m_2 y_2, m_1 z_1 + m_2 z_2, m_1 + m_2).$$

Quando o “peso” de tal centro de massa é zero, descrevemos a direção dada pelo vetor  $(x_1 - x_2, y_1 - y_2, z_1 - z_2)$  — neste momento, o leitor familiarizado com Geometria Projetiva talvez reconheça algum semblante de coordenadas homogêneas de um plano projetivo. Uma massa não-nula pode ser assumida como 1, e isto levou Grassmann a analisar matrizes das formas

$$\begin{pmatrix} x_1 & y_1 & z_1 & 1 \\ x_2 & y_2 & z_2 & 1 \end{pmatrix} \quad \text{e} \quad \begin{pmatrix} x_1 & y_1 & z_1 & 0 \\ x_2 & y_2 & z_2 & 0 \end{pmatrix},$$

bem como análogos de ordem mais alta. Por outro lado, mediante escolhas de bases, note que elementos de  $\text{Gr}_k(\mathbb{K}^n)$  são identificados com classes de equivalência de matrizes de ordem  $k \times n$  com entradas em  $\mathbb{K}$  e posto máximo, onde diremos que  $A \sim B$  se os espaços-linha de  $A$  e  $B$  coincidem (isto é, as linhas de  $A$  geram o mesmo subespaço de  $\mathbb{K}^n$  que as linhas de  $B$ ).

Vale observarmos que mesmo tal linguagem simplificada também não foi utilizada no *Ausdehnungslehre* original, que dispensou completamente fórmulas matemáticas e tratou apenas de argumentos filosóficos extremamente gerais, sendo (por falta de um termo melhor) intragável. Mas de qualquer modo, daí em diante o resto foi história.

Infelizmente, o potencial das descobertas de Grassmann não foi devidamente reconhecido na época. Mas hoje, vemos claramente como estes objetos são ubíquos em Geometria, não apenas pura mas também aplicada. A título de ilustração, alguns exemplos:

1. Em Geometria Riemanniana, Grassmannianos complexos são exemplos particulares de variedades de Kähler-Einstein homogêneas e localmente simétricas (veremos isto na Seção 3). Recordamos que uma variedade Riemanniana  $(M, \mathbf{g})$  é Einstein se o seu tensor de Ricci é um múltiplo constante de  $\mathbf{g}$ ; Kähler se estiver equipada com uma estrutura quase-complexa paralela; localmente simétrica se o seu tensor de curvatura é paralelo. Aqui, “paralelismo” é sempre relativo à conexão de Levi-Civita de  $\mathbf{g}$ .
2. Em Geometria Algébrica, Grassmannianos são exemplos clássicos de variedades projetivas (entenderemos isto melhor na Seção 4). Por variedades projetivas queremos dizer subvariedades de espaços projetivos descritas como o conjunto de zeros de alguma coleção de polinômios homogêneos.
3. Em Geometria Simplética, o Grassmanniano de subespaços Lagrangeanos de um espaço vetorial simplético  $(\mathbf{V}, \Omega)$  é uma ferramenta importante para estabelecer propriedades do chamado índice de Maslov: um isomorfismo explícito entre  $\pi_1 \text{Sp}(\mathbf{V}, \Omega)$  e  $\mathbb{Z}$ , onde  $\text{Sp}(\mathbf{V}, \Omega)$  é o grupo de simplectomorfismos de  $(\mathbf{V}, \Omega)$  e  $\pi_1$  é o funtor grupo fundamental. Para mais detalhes, veja [8] e [10].
4. Em Otimização, há o chamado algoritmo de Riemann-Newton: uma generalização do método de Newton conhecido dos cursos básicos de Cálculo, agora para aproximar zeros de campos de vetores em variedades Riemannianas. Tal algoritmo é descrito explicitamente para variedades Grassmannianas em [1]. Claro que esta não é a única situação em que Grassmannianos aparecem em problemas mais computacionais — veremos na Seção 4 um mergulho motivado por problemas desta natureza.

Para mais contexto histórico e detalhes sobre a vida de Grassmann, recomendamos o Capítulo 3 de [6]. Neste ponto, o leitor deve estar razoavelmente convencido da relevância de entender o mínimo sobre Grassmannianos. Mas com o que  $\text{Gr}_k \mathbf{V}$  se parece? Nosso objetivo final aqui é tentar responder, mesmo que de maneira possivelmente vaga, tal pergunta. E para isto, vamos ter que sujar nossas mãos um pouco.

Gostaria de agradecer a Andrzej Derdzinski pelos vários comentários úteis e ao parecerista anônimo por apontar algumas correções a serem feitas na versão inicial deste texto.

## 2. CÁLCULO EM GRASSMANNIANOS

Como nosso objetivo inicial é realizar Cálculo, comecemos assumindo que o corpo com o qual trabalharemos é  $\mathbb{K} = \mathbb{R}$  ou  $\mathbb{K} = \mathbb{C}$ , e também que  $n = \dim_{\mathbb{K}} \mathbf{V} > 1$  (a não ser que digamos o contrário). Denotando uma  $k$ -upla ordenada  $[x_1 \cdots x_k] \in \mathbf{V}^{\times k}$  de elementos de  $\mathbf{V}$  apenas pela letra  $\mathfrak{x}$  (digamos, em fonte gótica), consideramos a variedade de Stiefel de  $k$ -referenciais de  $\mathbf{V}$ :

$$\text{St}_k \mathbf{V} \doteq \{ \mathfrak{x} \in \mathbf{V}^{\times k} \mid \mathfrak{x} \text{ é linearmente independente} \}.$$

Vemos que  $\text{St}_k \mathbf{V}$  é um subconjunto aberto do espaço vetorial  $\mathbf{V}^{\times k}$ , e portanto é uma variedade diferenciável. Ainda, temos uma projeção canônica  $\Pi: \text{St}_k \mathbf{V} \rightarrow \text{Gr}_k \mathbf{V}$ , que leva  $\mathfrak{x}$  no subespaço de  $\mathbf{V}$  gerado por  $\mathfrak{x}$ . Com isto, equipamos  $\text{Gr}_k \mathbf{V}$  com a topologia quociente induzida por  $\Pi$ , que herda (direta ou indiretamente) todas as propriedades

boas de  $\text{St}_k\mathbf{V}$ , e mais. Isto é,  $\text{Gr}_k\mathbf{V}$  torna-se um espaço topológico compacto, conexo, Hausdorff e com base enumerável. A construção de um atlas para  $\text{Gr}_k\mathbf{V}$  parte de uma ideia simples: dada qualquer aplicação linear *sobrejetora*  $f: \mathbf{V} \rightarrow \mathbb{K}^k$  e um ponto  $\mathbf{W} \in \text{Gr}_k\mathbf{V}$ , vale de duas uma: ou  $f|_{\mathbf{W}}$  é um isomorfismo, ou não é. Então considere o subconjunto aberto

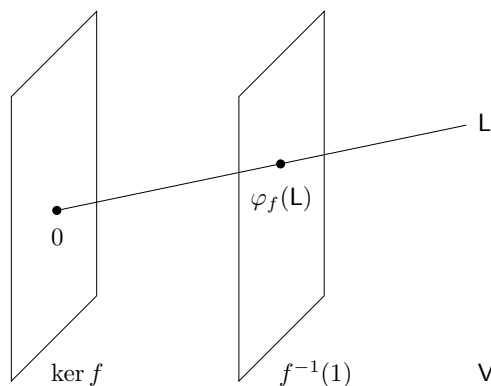
$$U_f \doteq \{\mathbf{W} \in \text{Gr}_k\mathbf{V} \mid f|_{\mathbf{W}} \text{ é um isomorfismo}\}$$

de  $\text{Gr}_k\mathbf{V}$ , e defina  $\varphi_f: U_f \rightarrow f^{-1}(e_1) \times \cdots \times f^{-1}(e_k)$  declarando  $\varphi_f(\mathbf{W})$  como a única base (ordenada) de  $\mathbf{W}$  que é levada por  $f$  na base canônica  $(e_1, \dots, e_k)$  de  $\mathbb{K}^k$ . Um cálculo direto mostra que a função de transição entre duas tais cartas é uma aplicação diferenciável (resp., holomorfa) entre abertos de  $\mathbb{R}^k$  (resp.,  $\mathbb{C}^k$ ) — observando que as entradas da inversa de uma matriz não-singular são funções racionais das entradas da matriz original). Assim, concluímos a:

**Proposição 1.**  *$\text{Gr}_k\mathbf{V}$  é uma variedade (real ou complexa) de dimensão (real ou complexa) igual a  $k(n - k)$ .*

Antes de prosseguirmos, há três comentários instrutivos a serem feitos.

1. Como recuperar, neste contexto, a descrição de  $\text{Gr}_k(\mathbb{K}^n)$  em termos de classes de equivalências de matrizes, dada na Seção 1? Basta notar que temos uma ação (livre) de  $\text{GL}_k(\mathbb{K})$  em  $\text{St}_k\mathbf{V}$  pela direita, dada por mudança de bases:  $(\mathfrak{X}, A) \mapsto \mathfrak{X}A$ . Assim,  $\text{Gr}_k\mathbf{V}$  é o quociente  $\text{St}_k(\mathbf{V})/\text{GL}_k(\mathbb{K})$ .
2. Há outra ação em  $\text{St}_k\mathbf{V}$ , desta vez pela esquerda, pelo grupo de automorfismos  $\text{GL}(\mathbf{V})$ : nada mais é do que uma ação diagonal dada por avaliação. Em outras palavras, se  $\mathfrak{X} = [x_1 \cdots x_k]$  e  $T \in \text{GL}(\mathbf{V})$ , então colocamos  $T\mathfrak{X} \doteq [Tx_1 \cdots Tx_k]$ . E o ponto chave para descrevermos concretamente os espaços tangentes a  $\text{Gr}_k\mathbf{V}$  na Proposição 2 adiante é que esta ação é compatível com a ação descrita no ponto acima, isto é, vale que  $T(\mathfrak{X}A) = (T\mathfrak{X})A$ , em vista da linearidade de  $T$ .
3. Quando  $k = 1$  e lidamos com  $\text{PV}$ , é simples descrever as cartas  $(U_f, \varphi_f)$  explicitamente: dado  $\mathbf{L} \in U_f$ , tome um vetor  $x \in \mathbf{L}$  não nulo (ou seja, tal que  $\mathbf{L} = \mathbb{K}x$ ), de modo que  $\varphi_f(\mathbf{L}) = x/f(x)$ . Veja a ilustração a seguir:



Em particular, quando  $\mathbf{V} = \mathbb{K}^{n+1}$ , as cartas  $(U_i, \varphi_i) \doteq (U_{\pi_i}, \varphi_{\pi_i})$  já cobrem  $\text{KP}^n$ , onde para cada  $i$  de 0 até  $n$ ,  $\pi_i: \mathbb{K}^{n+1} \rightarrow \mathbb{K}$  é a projeção na  $i$ -ésima componente. Cada pré-imagem  $\pi_i^{-1}(1)$  é isomorfa à  $\mathbb{K}^n$  mediante o deletamento da  $i$ -ésima

componente. Isto nos dá

$$\varphi_i([x_0 : \cdots : x_n]) = \left( \frac{x_1}{x_i}, \dots, \frac{x_{i-1}}{x_i}, \frac{x_{i+1}}{x_i}, \dots, \frac{x_n}{x_i} \right),$$

onde  $[x_0 : \cdots : x_n]$  denota a reta gerada pelo vetor  $(x_0, \dots, x_n)$ , como talvez seja mais familiar para o leitor.

Grassmannianos são suficientemente abstratos para que não seja imediato reconhecer seus espaços tangentes. Afinal, o que seria o vetor velocidade de uma curva de subespaços de  $V$ ?

**Proposição 2.**  $T_W \text{Gr}_k V \cong \text{Hom}(W, V/W)$ .

**Demonstração:** Recordando que temos a projeção  $\Pi: \text{St}_k V \rightarrow \text{Gr}_k V$  e, portanto, a sua derivada  $d\Pi_{\mathfrak{x}}: V^{\times k} \rightarrow T_{\Pi(\mathfrak{x})} \text{Gr}_k V$ , a descrição do isomorfismo é simples:

$$\text{Hom}(W, V/W) \ni H \mapsto d\Pi_{\mathfrak{x}}(\tilde{H}\mathfrak{x}) \in T_W \text{Gr}_k V,$$

onde  $\mathfrak{x} \in \text{St}_k V$  é qualquer base de  $W$  e  $\tilde{H}: W \rightarrow V$  é qualquer levantamento linear de  $V$ . Obviamente a aplicação acima será linear e sobrejetora (pois cada  $d\Pi_{\mathfrak{x}}$  o é), de modo que ser um isomorfismo segue de ambos  $\text{Hom}(W, V/W)$  e  $T_W \text{Gr}_k V$  possuírem a mesma dimensão (finita). Assim, o verdadeiro obstáculo aqui consiste em mostrar que nada depende das escolhas de  $\mathfrak{x}$  e  $\tilde{H}$ .

- Se  $\tilde{H}_1$  e  $\tilde{H}_2$  são dois levantamentos lineares do mapa  $H$  fixado, então a diferença  $\tilde{H}_2 - \tilde{H}_1: W \rightarrow W$  toma valores em  $W$ . Para  $t$  real pequeno o suficiente,  $\mathfrak{x} + t(\tilde{H}_2 - \tilde{H}_1)\mathfrak{x}$  é uma base de  $W$ , de modo que aplicar  $\Pi$  e derivar em  $t = 0$  nos dá que  $d\Pi_{\mathfrak{x}}(\tilde{H}_1\mathfrak{x}) = d\Pi_{\mathfrak{x}}(\tilde{H}_2\mathfrak{x})$ .
- Se  $\mathfrak{x}_1$  e  $\mathfrak{x}_2$  são duas bases de  $W$ , podemos escrever  $\mathfrak{x}_2 = \mathfrak{x}_1 A$ , onde  $A \in \text{GL}_k(\mathbb{K})$  é a matriz de mudança de base. Denotando também por  $A$  a sua ação (linear!) pela direita  $A: \text{St}_k V \rightarrow \text{St}_k V$ , vale que  $\Pi \circ A = \Pi$ . Tomando a derivada total dos dois lados desta igualdade no ponto  $\mathfrak{x}_1$ , a regra da cadeia nos dá que  $d\Pi_{\mathfrak{x}_2} \circ A = d\Pi_{\mathfrak{x}_1}$ . E avaliando no vetor  $\tilde{H}\mathfrak{x}_1$ , finalmente obtemos que  $d\Pi_{\mathfrak{x}_1}(\tilde{H}\mathfrak{x}_1) = d\Pi_{\mathfrak{x}_2}((\tilde{H}\mathfrak{x}_1)A) = d\Pi_{\mathfrak{x}_2}(\tilde{H}(\mathfrak{x}_1 A)) = d\Pi_{\mathfrak{x}_2}(\tilde{H}\mathfrak{x}_2)$ .

□

E como usar tal isomorfismo na prática? Se temos uma curva  $t \mapsto W(t)$  em  $\text{Gr}_k V$ , usamos  $W$  também para denotar  $W(0)$ , e queremos entender o que o vetor  $H = W'(0): W \rightarrow V/W$  faz, devemos descrever  $Hx$ , para  $x \in W$  arbitrário. Para tanto, escreva  $W(t) = \ker T(t)$ , com  $T(t): V \rightarrow V$  para cada  $t$ . Tome uma curva  $t \mapsto x(t)$  com  $x(0) = x$  e  $x(t) \in W(t)$  para cada  $t$ , de modo que  $T(t)x(t) = 0$ . Derivando tal igualdade em  $t = 0$ , obtemos que  $T'(0)x + T(0)x'(0) = 0$ . Moralmente, o que queremos é dizer que  $Hx = x'(0)$ , mas a relação anterior nos diz que  $x'(0)$  só está determinado módulo  $\ker T(0) = W$ . Então que seja  $Hx = x'(0) + W$ , e vamos em frente.

Próximo passo: se  $N$  é qualquer variedade (real ou complexa), uma função diferenciável (ou holomorfa)  $f: \text{Gr}_k V \rightarrow N$  é identificada unicamente com uma função  $\tilde{f}: \text{St}_k V \rightarrow N$  do mesmo tipo, com a propriedade adicional de que  $\tilde{f}(\mathfrak{x}) = \tilde{f}(\mathfrak{x}A)$



para toda  $A \in \text{GL}_k(\mathbb{K})$  — via a relação  $f \circ \Pi = \tilde{f}$ . Então a regra da cadeia nos diz que  $df_{\mathbb{W}}(H) = d\tilde{f}_{\tilde{\mathbf{x}}}(\tilde{H}\mathbf{x})$ , onde  $\mathbf{x}$  é qualquer base de  $\mathbb{W}$  e  $\tilde{H}$  é qualquer levantamento linear de  $H$ .

Note também que uma vez escolhido um produto interno em  $\mathbb{V}$ , o quociente  $\mathbb{V}/\mathbb{W}$  é identificado com o complemento ortogonal  $\mathbb{W}^\perp$ . Isto geralmente facilita alguns cálculos, como veremos no restante do texto. Vejamos alguns exemplos bem concretos de tudo o que foi descrito até agora:

**Exemplo 3** (Projetivizações). Cada aplicação linear *injetora*  $T: \mathbb{V}_1 \rightarrow \mathbb{V}_2$  entre dois espaços vetoriais induz uma aplicação  $\text{Gr}_k(T): \text{Gr}_k\mathbb{V}_1 \rightarrow \text{Gr}_k\mathbb{V}_2$  via imagens diretas, ou seja,  $\text{Gr}_k(T)(\mathbb{W}) \doteq T[\mathbb{W}]$ . Com esta definição, vemos que  $\text{Gr}_k$  de fato tem um comportamento functorial (ou seja, valem  $\text{Gr}_k(T \circ S) = \text{Gr}_k(T) \circ \text{Gr}_k(S)$  e também  $\text{Gr}_k(\text{Id}_{\mathbb{V}}) = \text{Id}_{\text{Gr}_k\mathbb{V}}$ ). Suponha agora, para facilitar, que  $\mathbb{V} = \mathbb{V}_1 = \mathbb{V}_2$  e que  $T \in \text{GL}(\mathbb{V})$ . Dado um vetor tangente  $H \in T_{\mathbb{W}}\text{Gr}_k\mathbb{V}$  qualquer, realizado por uma curva suave  $t \rightarrow \mathbb{W}(t)$ , tomamos uma curva com  $t \mapsto y(t) \in T[\mathbb{W}(t)]$ , de modo que  $d(\text{Gr}_k(T))_{\mathbb{W}}(H) = y'(0) + T[\mathbb{W}]$ . Mas derivando em  $t = 0$  a curva invertida  $t \mapsto T^{-1}y(t) \in \mathbb{W}(t)$ , vemos que  $HT^{-1}y = T^{-1}(y'(0)) + \mathbb{W}$ , e assim concluímos que  $d(\text{Gr}_k(T))_{\mathbb{W}}(H) = \tilde{T}HT^{-1}y + T[\mathbb{W}]$ , onde  $\tilde{T}$  é o isomorfismo  $\mathbb{V}/\mathbb{W} \cong \mathbb{V}/T[\mathbb{W}]$  induzido por  $T$ . O grupo de transformações  $\{\text{Gr}_k(T) \mid T \in \text{GL}(\mathbb{V})\}$  terá um papel importante na Seção 3. Quando  $k = 1$ , tal grupo é denotado por  $\text{PGL}(\mathbb{V})$ . Restringindo de  $\text{GL}(\mathbb{V})$  para  $\text{SL}(\mathbb{V})$ , ou  $\text{O}(\mathbb{V})$  ou  $\text{U}(\mathbb{V})$  (quando  $\mathbb{V}$  tiver estrutura adicional), também faz sentido falarmos em  $\text{PSL}(\mathbb{V})$ ,  $\text{PO}(\mathbb{V})$  ou  $\text{PU}(\mathbb{V})$ .

**Exemplo 4** (Projeções ortogonais). Equipe  $\mathbb{V}$  com um produto interno  $\langle \cdot, \cdot \rangle$ , real ou complexo, fixe  $u \in \mathbb{V}$  e defina  $F: \text{Gr}_k\mathbb{V} \rightarrow \mathbb{V}$  por  $F(\mathbb{W}) = \text{pr}_{\mathbb{W}}u$ . Se  $H: \mathbb{W} \rightarrow \mathbb{W}^\perp$  é um vetor tangente a  $\text{Gr}_k\mathbb{V}$  em  $\mathbb{W}$ , tome uma base ortonormal  $\mathbf{x}$  para  $\mathbb{W}$  e calcule

$$F(\mathbb{W}) = \sum_{i=1}^k \langle u, x_i \rangle x_i \implies dF_{\mathbb{W}}(H) = \sum_{i=1}^k \langle u, Hx_i \rangle x_i + \sum_{i=1}^k \langle u, x_i \rangle Hx_i.$$

Como  $H$  toma valores em  $\mathbb{W}^\perp$ , segue que  $dF_{\mathbb{W}}(H) = H^* \text{pr}_{\mathbb{W}^\perp}u + H \text{pr}_{\mathbb{W}}u$ . No cálculo acima, há talvez um passo não-trivial: a saber, que de fato podemos tomar as curvas  $t \mapsto x_i(t) \in \mathbb{W}(t)$  *sempre* ortonormais (para todo instante  $t$  e não apenas para  $t = 0$ ) e mantendo a propriedade  $x'_i(t) = Hx_i(t)$ . Caso seja necessário, substituímos  $H$  por  $H + T$ , para algum operador anti-simétrico  $T: \mathbb{W} \rightarrow \mathbb{W}$ , afinal valerá que

$$\sum_{i=1}^k \langle u, Tx_i \rangle x_i + \sum_{i=1}^k \langle u, x_i \rangle Tx_i = -T \text{pr}_{\mathbb{W}}u + T \text{pr}_{\mathbb{W}}u = 0$$

e, como  $T$  toma valores em  $\mathbb{W}$ ,  $H + T$  determina o mesmo vetor que  $H$  no espaço tangente  $T_{\mathbb{W}}\text{Gr}_k\mathbb{V}$ .

**Exemplo 5** (Dualidade). Considere desta vez a *aplicação polar*, denotada aqui por  $\text{Pol}: \text{Gr}_k\mathbb{V} \rightarrow \text{Gr}_{n-k}\mathbb{V}^*$ , que leva um subespaço  $\mathbb{W}$  no seu espaço aniquilador  $\mathbb{W}^0 \doteq \{f \in \mathbb{V}^* \mid f|_{\mathbb{W}} = 0\}$ . Sabemos da Álgebra Linear que  $\text{Pol}$  é bijetiva, e também que  $\mathbb{V}^*/\mathbb{W}^0 \cong \mathbb{W}^*$ . Se  $t \mapsto x(t) \in \mathbb{W}(t)$  e  $t \mapsto f(t) \in \mathbb{W}(t)^0$  são curvas com  $\mathbb{W}(0) = \mathbb{W}$ ,  $\mathbb{W}'(0) = H$ ,  $x(0) = x$  e  $f(0) = f$ , derivar os dois lados de  $f(t)x(t) = 0$  imediatamente implica que  $f'(0)x = -f(Hx)$ . Isto quer dizer que  $d(\text{Pol})_{\mathbb{W}}(H) = -H^*$ , onde  $H^*$  é

a aplicação dual de  $H$ . Logo  $\text{Pol}$  é um difeomorfismo local pelo Teorema da Função Inversa (e portanto um difeomorfismo global por já ser bijetora).

Como uma última curiosidade para esta seção, note que se  $E \rightarrow M$  é um fibrado vetorial sobre uma variedade  $M$ , tudo o que fizemos até agora pode ser aplicado de forma suave ponto a ponto, nos dando fibrações localmente triviais  $\text{Gr}_k E \rightarrow M$ , para cada  $k$ . Note que o espaço total  $\text{Gr}_k E$  é compacto sempre que a base  $M$  o for. Esta construção pode ser utilizada para expressar de forma elegante algumas curvaturas de uma variedade Riemanniana  $(M, \mathbf{g})$ . Por exemplo, a curvatura seccional é uma função  $K: \text{Gr}_2(TM) \rightarrow \mathbb{R}$ , o tensor de Ricci avaliado quadraticamente ao longo de vetores unitários nos dá uma função  $\text{Ric}: PTM \rightarrow \mathbb{R}$  que é constante precisamente quando  $(M, \mathbf{g})$  é Einstein, e por aí vai. Se  $M^k$  é uma subvariedade de algum  $(V, \langle \cdot, \cdot \rangle)$ , a aplicação de Gauss generalizada é  $G: M \rightarrow \text{Gr}_k V$  que leva  $x$  em  $T_x M$ , e sua derivada (identificada com uma aplicação bilinear  $T_x M \times T_x M \rightarrow (T_x M)^\perp$ ) é precisamente a Segunda Forma Fundamental de  $M$  em  $V$ .

### 3. A MÉTRICA NATURAL EM $\text{Gr}_k V$

Suponha que  $V$  está equipado com um produto interno (real ou complexo)  $h$ , e seja  $g = \text{Re } h$  (no caso real temos  $g = h$ , e no caso complexo  $h$  pode ser recuperado de  $g$  via a relação  $h = g - ig(i \cdot, \cdot)$ ). Por  $W^\perp$  sempre denotaremos o complemento  $h$ -ortogonal de  $W$ . A métrica natural em  $\text{Gr}_k V$  é definida do seguinte modo: dados  $W \in \text{Gr}_k V$  e  $H_1, H_2 \in T_W \text{Gr}_k V$ , escolhemos uma base  $h$ -ortonormal  $\mathfrak{X}$  para  $V$ , e (ligeiramente abusando da notação) decretamos que

$$\mathbf{g}_W(H_1, H_2) \doteq \sum_{i=1}^k g(H_1 x_i, H_2 x_i).$$

Para ver que a escolha de  $\mathfrak{X}$  não importa na definição acima, observe que vale a igualdade  $\mathbf{g}_W(H_1, H_2) = \text{Re tr}(H_1^* H_2)$ . Com isto, temos uma métrica Riemanniana  $\mathbf{g}$  em  $\text{Gr}_k V$  (que no caso  $\mathbb{K} = \mathbb{C}$  é compatível com a estrutura complexa de  $\text{Gr}_k V$ , ou seja,  $\mathbf{g}_W(iH_1, iH_2) = \mathbf{g}_W(H_1, H_2)$  para todos os vetores tangentes  $H_1, H_2 \in T_W \text{Gr}_k V$ ). A significância geométrica deste produto é que se consideramos a variedade de Stiefel de  $k$ -referenciais  $h$ -ortonormais de  $V$ ,

$$\text{St}_k(V, h) \doteq \{\mathfrak{X} \in V^{\times k} \mid h(x_i, x_j) = \delta_{ij} \text{ para todos os } 1 \leq i, j \leq n\},$$

a restrição  $\Pi: \text{St}_k(V, h) \rightarrow \text{Gr}_k V$  torna-se uma *submersão Riemanniana* — um conceito não apenas amplamente estudado por si só em Geometria, mas também muito útil para o estudo de variedades homogêneas, por exemplo. O artigo original [9] onde tais submersões aparecerem pela primeira vez é acessível e apresenta várias fórmulas relacionando as geometrias do espaço total e da base da submersão. Por outro lado, expressar a geometria de  $\text{St}_k(V, h)$  (i.e., a conexão  $\nabla$ , a curvatura  $R$ , etc.) de forma limpa quando  $k > 1$  é surpreendentemente trabalhoso (veja [11] para cálculos matriciais da curvatura seccional de  $\text{St}_k(\mathbb{R}^n)$ , e [13] e [14] para  $\text{Gr}_k(\mathbb{K}^{n+k})$ ). Quando  $k = 1$ ,  $\text{St}_1(V, h)$  é apenas uma esfera e a métrica definida acima em  $PV$  é chamada a *métrica de Fubini-Study*. Mesmo este caso particular é de extrema relevância pois,

quando  $\mathbb{K} = \mathbb{C}$ ,  $PV$  com esta métrica acaba por ser o modelo de forma espacial complexa com curvatura seccional holomorfa constante e positiva (veja [7]). De qualquer modo, mencionamos na Seção 1 que  $\text{Gr}_k V$  é rico em simetrias. A seguinte classe de isometrias é grande o suficiente para os nossos propósitos:

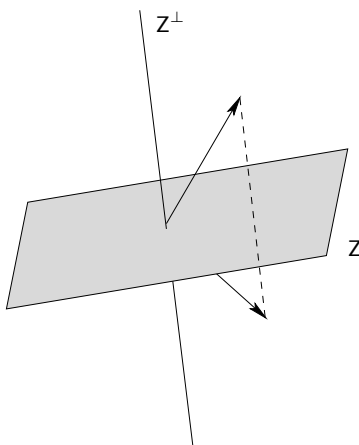
**Exemplo 6.** O grupo de transformações

$$\{\text{Gr}_k(T) \mid T \in \text{GL}(V) \text{ e } T^*h = h\}$$

age em  $\text{Gr}_k V$  por isometrias. De fato, como estamos usando  $h$  para identificar quocientes de  $V$  com seus subespaços, a derivada de  $\text{Gr}_k(T)$  é dada simplesmente por  $d(\text{Gr}_k(T))_W(H) = TH_1T^{-1}$ . Se  $\mathfrak{z}$  é uma base  $h$ -ortonormal para  $W$ ,  $T\mathfrak{z}$  é uma base  $h$ -ortonormal para  $T[W]$ , e assim segue que

$$\mathfrak{g}_{T[W]}(TH_1T^{-1}, TH_2T^{-1}) = \mathfrak{g}_W(H_1, H_2)$$

para quaisquer  $H_1, H_2 \in T_W \text{Gr}_k V$ , usando a definição da métrica  $\mathfrak{g}$  em  $\text{Gr}_k V$ . Como caso particular, fixado  $W$ , podemos considerar  $T$  como a reflexão ortogonal sobre  $W$ , ou seja,  $T = \text{Id}_V - 2 \text{pr}_{W^\perp}$ . Daí  $\varphi = \text{Gr}_k(T)$  satisfaz  $\varphi(W) = W$  e  $d\varphi_W = -\text{Id}_{T_W \text{Gr}_k V}$  — e isto tem inúmeras consequências que exploraremos em breve. Note que se  $T: V \rightarrow V$  é a reflexão ortogonal sobre algum subespaço  $Z$  fixado, um subespaço de  $V$  é  $T$ -invariante se e somente se for a soma direta de um subespaço de  $Z$  com um subespaço de  $Z^\perp$ :



A primeira consequência é que  $\text{Gr}_k Z$  é uma subvariedade totalmente geodésica de  $\text{Gr}_k V$  para todo subespaço  $Z \subseteq V$  com dimensão pelo menos  $k$ , pelo seguinte argumento: a discussão acima nos leva a considerar o subconjunto  $Q$  de  $\text{Gr}_k V$  formado pelos pontos que são fixos por todas as isometrias de  $\text{Gr}_k V$  induzidas por isometrias lineares de  $V$  que restritas a  $Z$  são  $\text{Id}_Z$ ; Sendo precisamente o conjunto de pontos fixos de uma família de isometrias de  $\text{Gr}_k V$ , vale que  $Q$  é a união disjunta de subvariedades totalmente geodésicas de  $\text{Gr}_k V$  (possivelmente de dimensões diferentes). Tais subvariedades acabam por ser  $\text{Gr}_k Z$  e  $\{W \in \text{Gr}_k V \mid Z^\perp \subseteq W\}$ .

Neste ponto, podemos nos perguntar se estas são todas as subvariedades totalmente geodésicas de  $\text{Gr}_k V$ . Em geral, a resposta é não, e há vários estudos sobre tais subvariedades com propriedades geométricas adicionais, como por exemplo em [2] e [12]. Quando  $k = 1$ , a situação é bem mais simples: por exemplo, quando  $\mathbb{K} = \mathbb{C}$

vale que toda subvariedade (complexa, conexa, completa) e totalmente geodésica  $N$  de  $PV$  é  $N = PZ$  para algum subespaço  $Z$  de  $V$ : se  $r = \dim N$ , tomamos  $Z = L \oplus \bigoplus_{i=1}^r H_i[L]$ , onde  $L \in N$  é qualquer reta e  $\{H_1, \dots, H_r\}$  é uma base para  $T_L N$  (afinal, subvariedades totalmente geodésicas nas condições dadas são determinadas por um ponto e seu espaço tangente).

**Exemplo 7.** Continuando o Exemplo 6, nos concentremos por um instante no caso  $k = 1$ . Mostramos que  $PO(V) \subseteq \text{Iso}(PV, \mathfrak{g})$  no caso real e  $PU(V) \subseteq \text{Iso}(PV, \mathfrak{g})$  no caso complexo. Na verdade, vale que estas são *todas* as isometrias (holomorfas, quando  $\mathbb{K} = \mathbb{C}$ ), i.e., valem igualdades acima. Mais precisamente, é possível mostrar que dada uma isometria linear  $\psi: T_{L_1}(PV) \rightarrow T_{L_2}(PV)$ , existe uma única isometria linear  $T: V \rightarrow V$  tal que  $d(PV)_{L_1} = \psi$  (construída explicitamente a partir de um levantamento de  $\psi$ ). Para  $\text{Gr}_k V$  com  $k \neq 1$  e  $n - 1$ , isto já não é mais verdade em geral: usando o produto  $h$  para identificar  $V$  com  $V^*$ , a aplicação  $\text{Gr}_k V \rightarrow \text{Gr}_{n-k} V$  que leva  $W$  em  $W^\perp$  nada mais é do que  $\text{Pol}$ , vista no Exemplo 5. Vejamos que  $\text{Pol}$  é uma isometria. Com efeito,  $d(\text{Pol})_W(H) = -H^*$ , onde  $H^*: W^\perp \rightarrow W$  agora denota o adjunto de  $H$ , donde

$$\mathfrak{g}_{W^\perp}(-H^*, -H^*) = \text{Re tr}((H^*)^* H^*) = \text{Re tr}(H^* H) = \mathfrak{g}_W(H, H).$$

Mas quando  $n = 2k \geq 4$ ,  $\text{Pol}$  não é induzida por nenhuma isometria linear de  $V$ .

O Exemplo 6 nos permite aplicar para  $(\text{Gr}_k V, \mathfrak{g})$  uma observação geral da Álgebra Linear: se  $B$  é um tensor  $r$  vezes contravariante e  $s$  vezes covariante em um espaço vetorial  $Z$ , com  $r + s$  ímpar, que é preservado por  $-\text{Id}_Z$ , então  $B = 0$ . Com efeito,  $(-\text{Id}_Z)^* B$  é simultaneamente igual a  $B$  (por hipótese) e também igual a  $-B$  (pois  $r + s$  é ímpar). Como a aplicação antípoda de cada espaço tangente a  $\text{Gr}_k V$  é realizada como a derivada de uma isometria, podemos aplicar isto ponto a ponto para  $\nabla R$  e  $\nabla J$  (onde  $R$  é o tensor de curvatura e, no caso  $\mathbb{K} = \mathbb{C}$ ,  $J$  é a estrutura quase-complexa dada por  $JH \doteq iH$ ), que possuem valência total 5 e 3, respectivamente. Obtemos o:

**Corolário 8.**  $(\text{Gr}_k V, \mathfrak{g})$  é uma variedade localmente simétrica e, quando  $\mathbb{K} = \mathbb{C}$ , também é de Kähler.

A última propriedade mencionada na Seção 1 que ainda não estabelecemos é que  $(\text{Gr}_k V, \mathfrak{g})$  é uma variedade de Einstein. Isto acaba sendo um caso particular de um argumento mais geral sobre variedades homogêneas, entre as quais Grassmannianos estão inclusos. Digamos, no caso  $\mathbb{K} = \mathbb{C}$ , fixado  $W \in \text{Gr}_k V$ , o estabilizador de  $W$  sob a ação  $U(V) \curvearrowright \text{Gr}_k V$  é o produto  $U(W) \times U(W^\perp)$ , de modo que pelo Teorema da Órbita e do Estabilizador, temos

$$\text{Gr}_k V \cong \frac{U(V)}{U(W) \times U(W^\perp)}.$$

No caso real, trocamos  $U$  por  $O$  acima. Dito isto, se  $G$  é um grupo agindo por isometrias em uma variedade Riemanniana  $(M, \mathfrak{g})$  e fixamos  $x \in M$ , o estabilizador  $G_x$  age em  $T_x M$  via derivadas (mais precisamente, define-se  $g \cdot v \doteq dg_x(v)$ , onde  $g: M \rightarrow M$ ), o que nos permite considerar as chamadas *representações de isotropia* da ação. E vemos que se em algum  $x \in M$ , esta representação é *irredutível*, ou seja, o

espaço tangente  $T_x M$  não possui subespaços  $G_x$ -invariantes, então  $(M, \mathbf{g})$  é Einstein, pelo seguinte motivo: se usamos  $\mathbf{g}_x$  para considerar  $\text{Ric}_x$  como um operador linear auto-adjunto em  $T_x M$ , o fato de  $G_x$  agir por isometrias nos diz que os autoespaços de  $\text{Ric}_x$  são  $G_x$ -invariantes, e logo a irreducibilidade implica que  $\text{Ric}_x = \lambda \mathbf{g}_x$  para alguma constante  $\lambda \in \mathbb{R}$ . Mas novamente,  $G$  age por isometrias em  $M$ , então se  $\text{Ric}_x = \lambda \mathbf{g}_x$  vale para algum ponto  $x \in M$ , vale para *todos* os pontos. Um argumento ligeiramente técnico mostra que a representação de isotropia da ação de  $U(\mathbf{V})$  em  $\text{Gr}_k \mathbf{V}$  é irreduzível, e assim obtemos o:

**Corolário 9.**  $(\text{Gr}_k \mathbf{V}, \mathbf{g})$  é uma variedade de Einstein.

#### 4. ALGUNS MERGULHOS CLÁSSICOS

**4.1. O mergulho de Plücker.** Os produtos  $h$  e  $g$  em  $\mathbf{V}$  induzem produtos (que também denotaremos por  $h$  e  $g$ ) na  $k$ -ésima potência exterior  $\mathbf{V}^{\wedge k}$ , via determinantes de Gram. Mais precisamente, definimos

$$h(x_1 \wedge \cdots \wedge x_k, y_1 \wedge \cdots \wedge y_k) = \det \left( (h(x_i, y_j))_{i,j=1}^k \right),$$

e similarmente para  $g$ . Deste modo, o espaço projetivo  $\mathbb{P}(\mathbf{V}^{\wedge k})$  também ganha uma métrica de Fubini-Study. Se  $W \in \text{Gr}_k \mathbf{V}$  e escolhemos uma base  $\mathbf{x}$  para  $W$ , podemos considerar o  $k$ -vetor (não-nulo)  $x_1 \wedge \cdots \wedge x_k$ , e mudando a base para  $\mathbf{x}' = \mathbf{x}A$ , com  $A \in \text{GL}_k(\mathbb{K})$ , vale que

$$x'_1 \wedge \cdots \wedge x'_k = \det(A) x_1 \wedge \cdots \wedge x_k.$$

Isto quer dizer que  $W$  determina uma certa reta em  $\mathbf{V}^{\wedge k}$ . O *mergulho de Plücker*  $\mathfrak{P}: \text{Gr}_k \mathbf{V} \rightarrow \mathbb{P}(\mathbf{V}^{\wedge k})$  leva  $W$  nesta reta  $\mathfrak{P}(W)$ . Verifiquemos que  $\mathfrak{P}$  é, de fato, um mergulho isométrico. Olhando para expressões em coordenadas não é difícil ver que  $\mathfrak{P}$  é diferenciável (ou holomorfa se  $\mathbb{K} = \mathbb{C}$ ). Em seguida, vejamos que  $\mathfrak{P}$  é injetor: se  $\mathfrak{P}(W) = \mathfrak{P}(W')$  e escolhemos bases  $\mathbf{x}$  e  $\mathbf{x}'$  para  $W$  e  $W'$ , escrevemos  $x'_1 \wedge \cdots \wedge x'_k = x_1 \wedge \cdots \wedge x_k$ . Por absurdo, se  $x'_1 \notin W$ , então temos que

$$0 \neq x'_1 \wedge x_1 \wedge \cdots \wedge x_k = x'_1 \wedge x'_1 \wedge \cdots \wedge x'_k = 0.$$

Repetir o argumento com qualquer  $x'_j$  no lugar de  $x'_1$  nos dá a inclusão  $W' \subseteq W$ , e logo  $W' = W$  pois ambos tem dimensão  $k$ . Assim, como  $\text{Gr}_k \mathbf{V}$  é compacto, uma vez que também mostrarmos que  $\mathfrak{P}$  é uma imersão,  $\mathfrak{P}$  será automaticamente um mergulho. E para mostrar que cada derivada  $d\mathfrak{P}_W$  é uma aplicação linear injetora, vamos colocar a carroça na frente dos bois e mostrar logo que é uma isometria linear. A regra do produto nos dá que  $d\mathfrak{P}_W: \text{Hom}(W, W^\perp) \rightarrow \text{Hom}(\mathfrak{P}(W), \mathfrak{P}(W)^\perp)$  é dada por

$$d\mathfrak{P}_W(H)(x_1 \wedge \cdots \wedge x_k) = \sum_{i=1}^k x_1 \wedge \cdots \wedge x_{i-1} \wedge Hx_i \wedge x_{i+1} \wedge \cdots \wedge x_k.$$

Com efeito, note que o lado direito da relação acima está em  $\mathfrak{P}(W)^\perp$ , visto que o  $i$ -ésimo termo da soma

$$\sum_{i=1}^k h(x_1 \wedge \cdots \wedge x_{i-1} \wedge Hx_i \wedge x_{i+1} \wedge \cdots \wedge x_k, x_1 \wedge \cdots \wedge x_k)$$

é o determinante da matriz obtida substituindo a  $i$ -ésima linha da matriz de Gram  $(h(x_i, x_j))_{i,j=1}^k$  por  $(h(Hx_i, x_1) \cdots h(Hx_i, x_k)) = (0 \cdots 0)$ .

E finalmente, para mostrar que  $d\mathfrak{P}_W$  preserva as métricas naturais, suponha agora que a base  $\mathfrak{X}$  escolhida para  $W$  é  $h$ -ortonormal. Com isto, é suficiente mostrar que

$$\|d\mathfrak{P}_W(H)(x_1 \wedge \cdots \wedge x_k)\|^2 = \sum_{i=1}^k \|Hx_i\|^2,$$

visto que o lado esquerdo é igual a  $\mathfrak{g}_{\mathfrak{P}(W)}(d\mathfrak{P}_W(H), d\mathfrak{P}_W(H))$  e o direito é  $\mathfrak{g}_W(H, H)$ . Isto talvez requeira um esforço de visualização do leitor, mas o lado esquerdo é a soma sobre  $i$  e  $j$  dos determinantes das matrizes descritas do seguinte modo: comece com a identidade de ordem  $k$ , troque a  $i$ -ésima linha e a  $j$ -ésima coluna por zeros, e insira na posição  $(i, j)$  o produto  $g(Hx_i, Hx_j)$ . Os únicos termos não-nulos ocorrem para  $i = j$  e assim obtemos a soma dos  $\|Hx_i\|^2$ , como desejado.

Isto conclui a demonstração, mas ainda há dois comentários instrutivos a serem feitos, agora que o trabalho duro está feito.

- Primeiro, para cada elemento  $\alpha \in V^{\wedge k}$ , podemos considerar o subespaço  $W_\alpha \doteq \{x \in V \mid \alpha \wedge x = 0\}$  de  $V$ . Quando  $\alpha = x_1 \wedge \cdots \wedge x_k$ , ou seja,  $\alpha$  é decomponível (e portanto na imagem de  $\mathfrak{P}$ ), temos que  $\alpha \mapsto W_\alpha$  é a inversa de  $\mathfrak{P}$ .
- Segundo, podemos reconciliar o que foi feito acima com o a situação em que  $V = \mathbb{K}^n$  e temos a descrição de  $\text{Gr}_k(\mathbb{K}^n)$  dada na Seção 1. A aplicação  $\mathfrak{P}$  deve levar a classe de equivalência de uma matriz  $A$  de ordem  $k \times n$  com posto máximo em uma certa reta em um espaço de dimensão  $n!/k!(n-k)!$ . Mas este é exatamente o número de menores principais de ordem  $k$  que uma matriz  $k \times n$  tem! Ou seja, teremos que  $\mathfrak{P}([A]) = [p_{i_1 \dots i_k} \mid 1 \leq i_1 < \cdots < i_k \leq n]$ , onde  $p_{i_1 \dots i_k}$  é o subdeterminante de  $A$  formado pelas colunas  $i_1, \dots, i_k$ . O fato de  $A$  ter posto máximo garante que realmente temos um elemento bem definido em  $\mathbb{K}P^{\binom{n}{k}-1}$ . A letra  $p$ , naturalmente, vem de Plücker. E este ponto de vista é o que evidencia que Grassmannianos são variedades projetivas. Por exemplo, temos coordenadas  $[p_{11} : p_{12} : p_{13} : p_{14} : p_{23} : p_{24} : p_{34}]$  para  $\text{Gr}_2(\mathbb{K}^4)$  mergulhado em  $\mathbb{K}P^5$ . Como  $\text{Gr}_2(\mathbb{K}^4)$  tem codimensão 1, esperamos poder descrevê-lo como o conjunto de zeros de um único polinômio. Dito e feito, temos  $p_{12}p_{34} - p_{13}p_{24} + p_{14}p_{23} = 0$ . Isto pode ser repetido em dimensões maiores, com mais equações, e descrito de forma intrínseca (isto é, sem apelar diretamente para matrizes) — veja [4], por exemplo.

**4.2. O mergulho de Conway-Hardin-Sloane.** Dado  $W \in \text{Gr}_k V$ , sempre podemos escrever a soma direta  $V = W \oplus W^\perp$ , e tal decomposição nos fornece a projeção ortogonal  $\text{pr}_W : V \rightarrow V$ . Claramente  $\text{pr}_W$  é um operador auto-adjunto com traço  $k/n$  que determina  $W$  completamente. Por outro lado, recorde que dado qualquer operador linear  $T \in \mathfrak{gl}(V)$ , definimos a *parte sem traço* de  $T$  como  $T_0 \doteq T - (\text{tr}(T)/n)\text{Id}_V$ . E  $T_0$  juntamente com o número  $\text{tr}(T)$  determina  $T$  completamente, no sentido de que se denotamos por  $\mathfrak{gl}_0(V)$  o espaço dos operadores lineares em  $V$  com traço zero, a aplicação

$$\mathfrak{gl}(V) \ni T \mapsto (T_0, \text{tr}(T)) \in \mathfrak{gl}_0(V) \times \mathbb{K}$$

é um isomorfismo de álgebras de Lie. Então, denotando por  $\mathfrak{gl}_0^{\text{sym}}(\mathbf{V})$  o espaço dos operadores auto-adjuntos em  $\mathbf{V}$  com traço zero, é natural considerar a aplicação  $\mathfrak{C}: \text{Gr}_k \mathbf{V} \rightarrow \mathfrak{gl}_0^{\text{sym}}(\mathbf{V})$  dada por  $\mathfrak{C}(W) \doteq (\text{pr}_W)_0$ . Como estamos interessados em Geometria Diferencial, precisamos de uma métrica no espaço  $\mathfrak{gl}_0^{\text{sym}}(\mathbf{V})$  para continuar a discussão. A boa e velha métrica Euclideana servirá para os nossos propósitos:  $\langle T, S \rangle \doteq \text{Re tr}(TS)/2$ .

Com isto em mãos, a afirmação é que não apenas  $\mathfrak{C}$  é um mergulho isométrico, mas também a imagem de  $\text{Gr}_k \mathbf{V}$  sob  $\mathfrak{C}$  é um subconjunto da esfera em  $\mathfrak{gl}_0^{\text{sym}}(\mathbf{V})$  de raio  $\sqrt{k(n-k)}/2n$ .

A discussão acima já mostra que  $\mathfrak{C}$  é injetora. Escrevendo

$$(\text{pr}_W)_0 = \frac{1}{n} ((n-k)\text{pr}_W - k\text{pr}_{W^\perp}),$$

vemos que

$$\langle (\text{pr}_W)_0, (\text{pr}_W)_0 \rangle = \frac{1}{2n^2} ((n-k)^2 k + k^2(n-k)) = \frac{k(n-k)}{2n}.$$

Finalmente, vejamos que a derivada  $d\mathfrak{C}_W$  preserva as métricas dadas. Dado qualquer  $H \in T_W \text{Gr}_k \mathbf{V}$ , como  $\text{pr}_W$  e  $(\text{pr}_W)_0$  diferem por algo (constante) que não depende de  $W$ , o Exemplo 4 nos dá que  $d\mathfrak{C}_W(H) = H^* \text{pr}_{W^\perp} + H \text{pr}_W$ . Um cálculo direto implica que vale a identidade  $d\mathfrak{C}_W(H)^2 = H^* H \text{pr}_W + H H^* \text{pr}_{W^\perp}$ , e portanto

$$\langle d\mathfrak{C}_W(H), d\mathfrak{C}_W(H) \rangle = \frac{1}{2} (2 \text{Re tr}(H^* H)) = \mathfrak{g}_W(H, H),$$

concluindo o argumento (pois novamente, uma imersão isométrica injetora com domínio compacto é automaticamente um mergulho).

Historicamente não é claro se tal mergulho de fato é devido a Conway e seus colaboradores, mas a atribuição não deixa de ser válida em vista de suas surpreendentes aplicações. Em [3], tal mergulho é utilizado no estudo de empacotamentos de retas em  $\mathbb{R}^3$  (e mais geralmente, no empacotamento de subespaços  $k$ -dimensionais de  $\mathbb{R}^n$ , que nos traz de volta a  $\text{Gr}_k \mathbf{V}$ ), em uma pesquisa interdisciplinar relacionada com o tratamento de tumores por meio de lasers de alta energia.

**4.3. Os mergulhos de Segre e Veronese.** Se  $(\mathbf{V}_1, h_1)$  e  $(\mathbf{V}_2, h_2)$  são espaços com produtos internos (reais ou complexos) e escrevemos  $g_i = \text{Re } h_i$ , para  $i = 1, 2$ , como na Seção 3, o produto tensorial  $\mathbf{V}_1 \otimes \mathbf{V}_2$  também ganha um produto  $h$  caracterizado pela relação

$$h(x_1 \otimes x_2, y_1 \otimes y_2) = h_1(x_1, y_1) h_2(x_2, y_2),$$

e similarmente definimos um produto real  $g$  a partir de  $g_1$  e  $g_2$ . Em particular, cada Grassmanniano de  $\mathbf{V}_1 \otimes \mathbf{V}_2$  ganha uma métrica natural. E a operação de produto tensorial pode ser vista como uma aplicação

$$\mathfrak{T}: \text{Gr}_{k_1} \mathbf{V}_1 \times \text{Gr}_{k_2} \mathbf{V}_2 \rightarrow \text{Gr}_{k_1 k_2}(\mathbf{V}_1 \otimes \mathbf{V}_2),$$

explicitamente dada por  $\mathfrak{T}(W_1, W_2) \doteq W_1 \otimes W_2$ . Vejamos que  $\mathfrak{T}$  é injetora: suponha que  $W_1 \otimes W_2 = W'_1 \otimes W'_2$ , onde assumimos que  $W_1, W'_1 \in \text{Gr}_{k_1} \mathbf{V}_1$  e  $W_2, W'_2 \in \text{Gr}_{k_2} \mathbf{V}_2$ . Sem perda de generalidade, é suficiente mostrar que  $W_1 \subseteq W'_1$ . Tomamos  $x \in W_1$  e fixamos  $y \in W_2$ , não-nulo. Se não tivemos  $x \in W'_1$ , existirá  $\phi \in V_1^*$  com  $\phi(x) = 1$  e

$\phi|_{W'_1} = 0$ , donde segue o absurdo  $y = (\phi \otimes \text{Id}_{V_2})(x \otimes y) = 0$ , visto que  $x \otimes y$  é uma soma de tensores puros em  $W'_1 \otimes W'_2$ . Portanto devemos ter  $x \in W'_1$ .

Prosseguindo, temos que a derivada desta aplicação  $\mathfrak{T}$  é a aplicação linear

$$d\mathfrak{T}_{(W_1, W_2)}: T_{W_1} \text{Gr}_{k_1} V_1 \times T_{W_2} \text{Gr}_{k_2} V_2 \rightarrow T_{W_1 \otimes W_2} \text{Gr}_{k_1 k_2} (V_1 \otimes V_2)$$

dada por

$$d\mathfrak{T}_{(W_1, W_2)}(H_1, H_2) = H_1 \otimes \text{Id}_{W_2} + \text{Id}_{W_1} \otimes H_2.$$

Mas se  $\mathfrak{x}$  é uma base  $h_1$ -ortonormal para  $W_1$  e  $\mathfrak{y}$  é uma base  $h_2$ -ortonormal para  $W_2$ , então  $\{x_i \otimes y_j \mid 1 \leq i \leq k_1, 1 \leq j \leq k_2\}$  é uma base  $g$ -ortonormal para  $W_1 \otimes W_2$ . Usando a caracterização de  $g$  em  $V_1 \otimes V_2$ , obtemos

$$g_{(W_1, W_2)}((H_1, H_2), (H_1, H_2)) = k_2 g_{W_1}(H_1, H_1) + k_1 g_{W_2}(H_2, H_2),$$

donde pelo menos segue que  $\mathfrak{T}$  é um mergulho suave, apesar de não ser isométrico. Mas quando temos que  $k_1 = k_2 = 1$  (que incidentalmente é equivalente à relação  $k_1 = k_2 = k_1 k_2$  entre as dimensões dos subespaços envolvidos), obtemos o chamado mergulho de Segre de um produto de espaços projetivos:  $PV_1 \times PV_2 \rightarrow P(V_1 \otimes V_2)$ . Isto imediatamente se generaliza para um mergulho isométrico

$$\prod_{i=1}^r PV_i \hookrightarrow P\left(\bigotimes_{i=1}^r V_i\right)$$

quando temos  $r$  espaços  $V_1, \dots, V_r$ . Quando tomamos  $V = V_1 = \dots = V_r$ , compor  $\mathfrak{T}$  com o mergulho diagonal  $\Delta$  nos dá ainda outro mergulho isométrico

$$PV \hookrightarrow (PV)^{\times r} \hookrightarrow P(V^{\odot r}),$$

onde  $V^{\odot r}$  denota a  $r$ -ésima potência simétrica de  $V$ . Observe que a homotetia causada pelo mergulho diagonal  $\Delta$  é desfeita pela normalização da métrica natural em  $V^{\odot r}$  (ou seja, considerar  $P(V^{\otimes r})$  ao invés de  $P(V^{\odot r})$  nos dá apenas um mergulho homotético). Este é chamado o mergulho de Veronese de grau  $r$  de  $V$ .

Estes dois mergulhos combinados nos dão uma gama de exemplos de subvariedades de espaços projetivos, e são frequentes tanto em Geometria Diferencial e Geometria Algébrica. Por exemplo, quando temos  $V = \mathbb{K}^3$  e  $r = 2$ , o mergulho de Veronese  $\nu: \mathbb{K}P^2 \rightarrow \mathbb{K}P^5$  resulta nas coordenadas homogêneas consistindo de todos os produtos possíveis (dois a dois) entre as variáveis  $x, y, z$ , isto é,  $\nu([x : y : z]) = [x^2 : y^2 : z^2 : yz : xz : xy]$ . Uma exposição bastante detalhada sobre variedades de Segre e Veronese, do ponto de vista da Geometria Algébrica, pode ser consultada em [5].

#### REFERÊNCIAS

- [1] Absil, P. A.; Mahony, R.; Sepulchre, R.; *Riemannian geometry of Grassmann manifolds with a view on algorithmic computation*, Acta Applicandae Mathematicae (Volume 80, Issue 2, pp. 199-220), 2004.
- [2] Chen, B. Y.; Nagano, T.; *Totally geodesic submanifolds of symmetric spaces, I*, Duke Math Journal (Volume 44, pp. 745-755), 1977.
- [3] Conway, J. H.; Hardin, R. H.; Sloane, N. J. A.; *Packing lines, planes, etc.: packings in Grassmannian spaces*, Experiment. Math. (Volume 5, Issue 2, pp. 139-159), 1996.
- [4] Griffiths, P.; Harris, J.; *Principles of Algebraic Geometry*, John Wiley & Sons, Inc., 1994.



- [5] Harris, J., *Algebraic Geometry: A First Course*, Springer-Verlag, 1995.
- [6] Klein, F., *Development of Mathematics in the 19<sup>th</sup> Century*, tradução por M. Ackerman, apêndices por R. Hermann, *Math Sci Press* (Lie Groups: History, Frontiers and Applications — volume IX), 1979.
- [7] Kobayashi, N.; Nomizu, K.; *Foundations of Differential Geometry, Volume 2*, Interscience Publishers, 1969.
- [8] McDuff, D.; Salamon, D.; *Introduction to Symplectic Topology*, 3<sup>a</sup> edição, Oxford Graduate Texts in Mathematics (Volume 27), 2017.
- [9] O’Neill, B., *The Fundamental Equations of a Submersion*, Michigan Math J. (volume 13, pp. 459-469), 1966.
- [10] Piccione, P.; Tausk, D. V.; *On the Geometry of Grassmannians and the Symplectic Group: the Maslov Index and Its Applications*, 2007.
- [11] Rapcsák, T., *Sectional curvatures in non linear optimization*, J. Glob. Optim. (Volume 40, pp. 375-388), 2008.
- [12] Wolf, J., *Elliptic Spaces in Grassmann Manifolds*, Illinois J. Math (Volume 7, Issue 3, pp. 447-462), 1963.
- [13] Wong, Y. C., *Differential Geometry of Grassmann Manifolds*, PNAS March 1 (Volume 57, Issue 3, pp. 589-594), 1967.
- [14] Wong, Y. C., *Sectional Curvatures of Grassmann Manifolds*, Proc. Natl. Aca. Sci. USA (Volume 60, Issue 1, pp. 75-79), 1968.

THE OHIO STATE UNIVERSITY  
DEPARTMENT OF MATHEMATICS  
*Email address:* terekcoutho.1@osu.edu

**UMA INTRODUÇÃO A SOMATÓRIOS FRACIONÁRIOS:  
APRENDENDO A SOMAR UMA QUANTIDADE NÃO INTEIRA DE  
PARCELAS**

ÁGDA TALITA GALVÃO E JOCEMAR DE QUADROS CHAGAS

RESUMO. Neste artigo apresentamos uma introdução à teoria de *somatórios fracionários*, que permite considerar somas finitas com quantidades não inteiras de parcelas. Este tipo de somatório tem seu primeiro registro em um exemplo devido a Euler, em uma publicação de 1755 [7], mas só voltou a ter vez na literatura em 2005 [18]. Expomos os axiomas, a definição de *funções fracionário-somáveis* e a fórmula fundamental para somatórios fracionários propostos por M. Müller e D. Schleicher, além de alguns exemplos, incluindo o exemplo de Euler. Como este artigo é, possivelmente, o primeiro registro em língua portuguesa sobre somatórios fracionários, excluída a dissertação [11], para facilitar a leitura do texto e a ambientação do leitor com os detalhes introduzidos nesta teoria, mostramos o passo-a-passo de muitos dos procedimentos realizados.

## 1. INTRODUÇÃO

Ao longo da história o conceito de *número* foi desenvolvido gradualmente. Inicialmente eram conhecidos apenas números naturais, depois números racionais, negativos, reais e complexos, introduzidos em alguma ordem conforme o decorrer da evolução da humanidade, quando surgiam necessidades, por exemplo, de contar animais, particionar objetos, registrar dívidas, etc.

É provável que a primeira operação matemática a surgir tenha sido a *adição*, surgida a partir da sistematização do processo natural de contagem. O modo como a contagem foi abstraída, junto com alguns dos sistemas usados para representar quantidades por meio de símbolos, os números, permitiram que algumas propriedades das quantidades fossem observadas, como por exemplo a possibilidade de efetuar de uma vez só a adição de dois ou mais objetos a um grupo, em vez de somar um objeto de cada vez. O entendimento

---

Data de aceitação: Setembro de 2020.

*Palavras chave.* Somas, Somatório, Somatório Fracionário, Parcelas Não Inteiras.

Autora 1, Ágda T. Galvão, agradece a SBM, IMPA, UEPG e UAB/CAPES pela oferta e manutenção do PROFMAT.

da operação aditiva foi sendo constantemente aprimorado, por exemplo, com o estabelecimento das propriedades comutativa e associativa:

- i) A adição é comutativa:  $a + b = b + a$  ;
- ii) A adição é associativa:  $a + (b + c) = (a + b) + c$  ,

entre outras. A notação que usamos para escrever sinteticamente uma soma longa (com uma quantidade finita ou infinita de parcelas) utiliza o símbolo  $\sum$ . Nos referimos a esta forma de denotar um *somatório* como *notação sigma*, pois “Euler estabeleceu a notação de somatório que usamos hoje, usando sigma, uma letra grega maiúscula, para o simbolizar a soma.” (ROSA, [22]). A origem e difusão desta notação são melhores pontuadas por Cajori ([3]):

O sinal de somatório é devido a Euler (1755), que diz: “*summam indicabimus signo  $\Sigma$* ”, ou seja,  $\Sigma$  indica o sinal de soma. Este símbolo foi utilizado por Lagrange, mas, de resto, recebeu pouca atenção durante o Séc. XVIII. . . . O símbolo  $\Sigma$  para expressar “soma” aparece novamente em *La Théorie Analytique de la chaleur* de Fourier, publicada em 1822, e nas funções elípticas de Jacobi em 1829. Cauchy utilizou três índices  $m, n, r$ , como em  $\sum_m^n r f r$ .

Segundo Graham ([12]), a expressão  $\sum_{k=1}^n f(k)$ , onde  $n$  é um número inteiro positivo, nos diz para incluir na soma precisamente aqueles termos  $f(k)$  cujo índice  $k$  é um inteiro que se encontra entre o menor e o maior limites 1 e  $n$ , inclusive. Por ser uma notação compacta, elegante, e por ter propriedades que facilitam operações algébricas, a notação sigma para somatórios é muito importante em vários campos da matemática, bem como em outras áreas como estatística, física, química, etc.

Quando são considerados somatórios com um número infinito de parcelas, conhecidos por *séries numéricas* (cuja notação traz o símbolo  $\infty$  no limite superior do somatório), inicialmente deve-se decidir se a série é *convergente*, e se for, é interessante tentar calcular seu valor numérico, sendo importante conhecer o comportamento das sequências envolvidas (ver, e.g., [17]).

Estes conceitos (somatório clássico e séries numéricas) estão bem claros para os matemáticos da atualidade, e encontram-se inseridos em livros didáticos dos mais variados níveis de estudo, na maioria das vezes usando a notação sigma para somatórios diretamente (ou seja, sem uma seção prévia introduzindo a notação), indicando que ela é considerada pelos autores de tais livros como sendo de conhecimento pleno por todos (e é muito provável que seja).

Mas que significado podemos atribuir a um somatório do tipo  $\sum_{\nu=1}^x f(\nu)$ , quando  $x \in \mathbb{R}$  ou  $\mathbb{C}$ ?

Passamos agora a olhar para somatórios onde os índices são de ordens não-inteiras, chamados índices de ordens arbitrárias, que abrangem índices racionais, reais e até mesmo complexos, aos quais convencionou-se chamar *índices fracionários* (da mesma forma que na nomenclatura do *cálculo fracionário* - ver, e.g., [4], [14], [21]). Por exemplo, quais seriam as somas dos primeiros  $-7$  números de uma sequência, ou dos  $\pi$  primeiros termos

da série harmônica? Este tipo de somatório vem sendo tratado, recentemente, com o nome de *somatório de índices fracionários*, ou apenas *somatório fracionário* ([1], [18]).

Müller e Schleicher ([20]) comentam que poderíamos pensar que um método para calcular este tipo de somas deveria ter sido descoberto há pelo menos duzentos anos, mas para surpresa desses autores (e nossa) uma tal teoria não parece ter sido investigada na literatura ou ser conhecida pelos especialistas antes de 2005 ([18]), além de breves observações esporádicas, nem mesmo no trabalho de Euler ([6], [7]), que nos brinda com um belo exemplo de uma soma com um número racional de termos, ao utilizar uma soma com um número não inteiro de termos como um dos métodos para introduzir funções. De fato, Euler apresentou o primeiro exemplo de somatório fracionário:

$$(1) \quad \sum_{\nu=1}^{-1/2} \frac{1}{\nu} = -2 \ln(2).$$

No entanto, uma formulação adequada do problema e sua solução foram apresentadas apenas a partir de 2005 por Müller e Schleicher ([18], [19], [20]), que entre vários exemplos encontraram o interessante caso onde somatórios fracionários podem ser utilizados para obter a função gama, denotada por  $\Gamma(x)$ , e que em certo sentido estende a noção de fatorial de um número para os números reais (e complexos).

Como afirmam Tenreiro Machado et al. ([16]) os conceitos de “fracionário” embutidos nos conceitos de diferencial e de integral de funções permitem uma notável e frutífera generalização do sistema de operadores do cálculo clássico, e acreditamos que o mesmo ocorrerá com o conceito de somatório. A extensão do significado da inserção adequada de ordens fracionárias nos operadores clássicos da matemática pode ser comparada ao avanço que ocorreu na antiguidade quando surgiu a necessidade do conceito de números racionais, e na sequência e em momentos posteriores, dos irracionais e dos reais. O sucesso destas “novas” ferramentas em aplicações às ciências de alguma forma podem superar outras possíveis generalizações matemáticas e ajustar-se melhor à modelagem de fenômenos físicos, contribuindo assim para o desenvolvimento da fronteira da ciência, mas ainda é cedo para fazer a mensuração desta contribuição no caso dos somatórios fracionários.

Pretendemos, com base nos trabalhos de Müller e Schleicher ([18], [19], [20]) e de Euler ([6], [7]), apresentar em língua portuguesa os conceitos básicos da teoria para somatórios fracionários. Partiremos de um pequeno número de axiomas bem motivados, e apresentaremos uma definição única para somatórios com um número não inteiro de termos. Utilizando a definição, mostraremos exemplos de identidades dos somatórios clássicos que permanecem válidas no contexto dos somatórios fracionários, e além disso, apresentaremos dois exemplos de somas fracionárias que não encontram expressão equivalente para somatórios clássicos.

## 2. SOMATÓRIOS FRACIONÁRIOS - OS AXIOMAS

Consideraremos o contexto de funções complexas definidas em  $\mathbb{C}$ , isto é,

$$\begin{aligned} f : \mathbb{C} &\rightarrow \mathbb{C} \\ z &\mapsto f(z), \end{aligned}$$

onde se deseja somar quantidades não inteiras de termos deste tipo. Neste texto,  $x, y, z$  e  $s$  são usados para representar números complexos, enquanto  $f$  e  $g$  representam funções complexas definidas em  $\mathbb{C}$  ou seus subconjuntos, sujeitos a condições que serão especificadas na próxima seção.

A primeira necessidade é apresentar condições naturais para somas com um número complexo arbitrário de termos: os axiomas, que listamos a seguir.

**Axioma (S1) - Continuidade da soma:**

$$\sum_{\nu=x}^y f(\nu) + \sum_{\nu=y+1}^z f(\nu) = \sum_{\nu=x}^z f(\nu).$$

**Axioma (S2) - Invariância por translação:**

$$\sum_{\nu=x+s}^{y+s} f(\nu) = \sum_{\nu=x}^y f(\nu + s).$$

**Axioma (S3) - Linearidade para constantes arbitrárias  $\lambda, \mu \in \mathbb{R}$ :**

$$\sum_{\nu=x}^y (\lambda f(\nu) + \mu g(\nu)) = \lambda \sum_{\nu=x}^y f(\nu) + \mu \sum_{\nu=x}^y g(\nu).$$

**Axioma (S4) - Consistência com a definição clássica:**

$$\sum_{\nu=1}^1 f(\nu) = f(1).$$

**Axioma (S5) - Somas de monômios:**

Para cada  $d \in \mathbb{N}$ , a aplicação

$$z \mapsto \sum_{\nu=1}^z \nu^d$$

é holomórfica em  $\mathbb{C}$  (funções holomorfas são funções definidas sobre um subconjunto aberto do plano complexo  $\mathbb{C}$  com valores em  $\mathbb{C}$  e que são diferenciáveis em cada ponto; ver e.g. [9]).

**Axioma (S6) - Continuidade no deslocamento à direita:**

Se  $\lim_{n \rightarrow \infty} f(z + n) = 0$  para cada ponto  $z \in \mathbb{C}$ , então

$$(2) \quad \lim_{n \rightarrow \infty} \sum_{\nu=x}^y f(\nu + n) = 0.$$

Mais geralmente, se existe uma sequência de polinômios  $(p_n)_{n \in \mathbb{N}}$  de grau fixo, tal que  $|f(z + n) - p_n(z + n)| \rightarrow 0$  quando  $n \rightarrow \infty$  para todo  $z \in \mathbb{C}$ , então vale

$$(3) \quad \left| \sum_{\nu=x}^y f(\nu + n) - \sum_{\nu=x}^y p_n(\nu + n) \right| \rightarrow 0.$$

Os primeiros quatro axiomas (S1) - (S4) são tão óbvios e necessários, que é difícil imaginar construir uma teoria para somatórios que os violem. Com esses quatro axiomas,

facilmente pode ser verificado que vale

$$\sum_{\nu=1}^n f(\nu) = f(1) + f(2) + \cdots + f(n)$$

para cada  $n \in \mathbb{N}$ , de forma que, ao considerar esses quatro axiomas, fica estabelecida a consistência com a definição clássica de soma.

O axioma (S5) é motivado pelas conhecidas fórmulas

$$\sum_{\nu=1}^n \nu = \frac{n(n+1)}{2}; \quad \sum_{\nu=1}^n \nu^2 = \frac{n(n+1)(2n+1)}{6}; \quad \sum_{\nu=1}^n \nu^3 = \left(\frac{n(n+1)}{2}\right)^2;$$

e fórmulas similares para potências superiores. Se respeitamos os axiomas (S1) - (S5), todas essas fórmulas permanecem válidas ao considerarmos valores  $n \in \mathbb{C}$  arbitrários.

Finalmente, o axioma (S6) também pode ser apresentado como uma condição natural. No caso mais simples (2) está expresso o fato que, se  $f$  tende a zero quando seu argumento tende ao infinito, então a soma sobre "o domínio limitado"  $[x, y]$  deve fazer o mesmo. Em (3) o mesmo se aplica, apenas é acrescentada uma aproximação polinomial para a função  $f$ .

Para uma grande classe de funções  $f$ , há uma maneira única de definir um somatório  $\sum_{\nu=1}^z f(\nu)$  com  $z \in \mathbb{C}$  que respeita a todos os axiomas (S1) - (S6). Na próxima seção iremos apresentar tal definição, e denotaremos tais somas, os *somatórios fracionários*, com o símbolo  $\sum_{\nu=1}^z f(\nu)$ .

### 3. SOMATÓRIOS FRACIONÁRIOS - UMA DEFINIÇÃO ÚNICA

Para ganhar experiência no uso dos axiomas (S1) - (S6) para o manejo de somas e começar a compreender como tais axiomas podem ser usados para determinar de forma única um método para somar uma quantidade não inteira de termos, começaremos somando polinômios. O caso mais simples é uma soma do tipo  $\sum_{\nu=1}^x c$ , onde  $c \in \mathbb{C}$  é constante e  $x \in \mathbb{R}$ .

Como um primeiro exemplo, vamos somar  $\sum_{\nu=1}^{1/2} c$ . Se consideramos o axioma (S1) válido, podemos escrever

$$\sum_{\nu=1}^{1/2} c + \sum_{\nu=3/2}^1 c = \sum_{\nu=1}^1 c.$$

Ao aplicarmos o axioma (S2) na segunda parcela do lado esquerdo desta equação, e o axioma (S4) no lado direito, obtemos

$$\sum_{\nu=1}^{1/2} c + \sum_{\nu=1}^{1/2} c = c,$$

e assim

$$\sum_{\nu=1}^{1/2} c = \frac{c}{2}.$$

Por linearidade, este procedimento simples pode ser estendido para definir qualquer soma, com um número racional de termos, de polinômios.

O lema 3.1, a seguir, será necessário na proposição 3.2 para a definição de soma fracionária de uma função polinomial (soma com quantidade de parcelas racional, real, ou mesmo complexa).

**Lema 3.1:** *Para qualquer polinômio  $p : \mathbb{C} \rightarrow \mathbb{C}$ , de grau  $m$ , existe um único polinômio  $P : \mathbb{C} \rightarrow \mathbb{C}$  tal que  $P(0) = 0$  e  $P(z) - P(z - 1) = p(z)$  para todo  $z \in \mathbb{C}$ .*

**Demonstração:** Fixado qualquer polinômio  $p : \mathbb{C} \rightarrow \mathbb{C}$  de grau  $m$ , para  $n = 1, 2, \dots, m+1$  definimos  $\alpha_n = \sum_{i=1}^n p(i)$  e consideramos os  $m+2$  pontos  $(0, 0)$  e  $(n, \alpha_n)$ . Pelo teorema de interpolação de Lagrange (ver, e.g., [5], [13]), existe um único polinômio  $P$  de grau máximo  $m+1$  que passa por esses  $m+2$  pontos, isto é, com  $P(0) = 0$  e  $P(n) = \alpha_n$  para  $1 \leq n \leq m+1$ , tal que  $P(n) - P(n-1) = p(n)$ .

Tomamos agora  $q : \mathbb{C} \rightarrow \mathbb{C}$  definido por  $q(z) = P(z) - P(z-1)$ ; o grau de  $q$  é no máximo  $m$ , e vale  $q(n) = P(n) - P(n-1) = p(n)$  para  $1 \leq n \leq m+1$ . Segue portanto que  $q = p$ , ou seja,  $P(z) - P(z-1) = p(z)$ , para todo  $z \in \mathbb{C}$  (em particular, o grau do polinômio  $P$  é  $m+1$ ).  $\square$

Apresentamos, a seguir, a definição para a soma fracionária de uma função polinomial.

**Proposição 3.2:** *Para qualquer polinômio  $p : \mathbb{C} \rightarrow \mathbb{C}$ , seja  $P : \mathbb{C} \rightarrow \mathbb{C}$  o único polinômio tal que  $P(0) = 0$  e  $P(z) - P(z-1) = p(z)$  para todo  $z \in \mathbb{C}$ . Então:*

(I) *A definição de soma fracionária de polinômios dada por*

$$(4) \quad \sum_{\nu=x}^y p(\nu) := P(y) - P(x-1)$$

*satisfaz os axiomas (S1) a (S6).*

(II) *Inversamente, toda teoria de soma que satisfaça aos axiomas (S1), (S2), (S3), e (S4) também deve satisfazer a definição (4) para todo polinômio  $p$  e para todo  $x, y \in \mathbb{C}$  com diferença racional, i.e., com  $y - x \in \mathbb{Q}$ .*

(III) *Toda teoria de soma que satisfaça aos axiomas (S1), (S2), (S3), (S4) e (S5) também deve satisfazer a definição (4) para todo polinômio  $p$  e para todo  $x, y \in \mathbb{C}$ .*

**Demonstração:** Dado um polinômio  $p : \mathbb{C} \rightarrow \mathbb{C}$ , para provar a afirmação (I), tomamos o único polinômio  $P$  como no lema 3.1.

Para verificar que o axioma (S1) é satisfeito pela eq. (4), calculamos:

$$\begin{aligned} \sum_{\nu=x}^z p(\nu) + \sum_{\nu=z+1}^y p(\nu) &= P(z) - P(x-1) + P(y) - P((z+1)-1) \\ &= P(y) - P(x-1) \\ &= \sum_{\nu=x}^y p(\nu). \end{aligned}$$

Para verificar que o axioma (S4) é satisfeito pela eq. (4), calculamos:

$$\sum_{\nu=1}^1 p(\nu) = P(1) - P(1-1) = P(1) - P(0) = p(1).$$

Para verificar que o axioma (S5) é satisfeito pela eq. (4), basta notar que para cada  $d \in \mathbb{N}$ , existe um único polinômio  $P : \mathbb{C} \rightarrow \mathbb{C}$  tal que  $P(0) = 0$  e  $P(z) - P(z-1) = z^d$ , de forma que podemos escrever

$$\sum_{\nu=1}^z \nu^d = P(z) - P(1-1) = P(z) - P(0) = P(z)$$

para qualquer  $z \in \mathbb{Z}$ , e polinômios são holomorfos em  $\mathbb{C}$  (ver, e.g., [9]).

Para verificar que o axioma (S2) é satisfeito pela eq. (4), consideramos o polinômio  $p$ , e para cada  $s \in \mathbb{C}$ , definimos os polinômios  $\tilde{p}, \tilde{P} : \mathbb{C} \rightarrow \mathbb{C}$  respectivamente por  $\tilde{p}(z) := p(z+s)$  e  $\tilde{P}(z) := P(z+s) - P(s)$ , de forma que o polinômio  $\tilde{P}$  satisfaz  $\tilde{P}(0) = 0$  e  $\tilde{P}(z) - \tilde{P}(z-1) = \tilde{p}(z)$ . De fato:

$$\begin{aligned} \tilde{P}(0) &= P(0+s) - P(s) = 0; \text{ e} \\ \tilde{P}(z) - \tilde{P}(z-1) &= P(z+s) - P(s) - (P(z-1+s) - P(s)) \\ &= P(z+s) - P(z+s-1) \\ &= p(z+s) \\ &= \tilde{p}(z). \end{aligned}$$

Segue portanto

$$\begin{aligned} \sum_{\nu=x}^y p(\nu+s) &= \sum_{\nu=x}^y \tilde{p}(\nu) \\ &= \tilde{P}(y) - \tilde{P}(x-1) \\ &= (P(y+s) - P(s)) - (P(x-1+s) - P(s)) \\ &= P(y+s) - P(x+s-1) \\ &= \sum_{\nu=x+s}^{y+s} p(\nu). \end{aligned}$$

Para verificar que o axioma (S3) é satisfeito pela eq. (4), consideramos  $\lambda, \mu \in \mathbb{R}$ , o polinômio  $p$ , um outro polinômio  $q : \mathbb{C} \rightarrow \mathbb{C}$  e o único polinômio  $Q : \mathbb{C} \rightarrow \mathbb{C}$  que satisfaz  $Q(0) = 0$  e  $Q(z) - Q(z-1) = q(z)$  para todo  $z \in \mathbb{C}$ , conforme o lema 3.1.



Definindo o polinômio  $r : \mathbb{C} \rightarrow \mathbb{C}$  por  $r(z) := \lambda p(z) + \mu q(z)$  para todo  $z \in \mathbb{C}$ , o polinômio  $R : \mathbb{C} \rightarrow \mathbb{C}$  definido por  $R(z) := \lambda P(z) + \mu Q(z)$  para todo  $z \in \mathbb{C}$  satisfará  $R(z) - R(z - 1) = r(z)$  e  $R(0) = 0$ . De fato:

$$R(0) = \lambda P(0) - \lambda Q(0) = \lambda \cdot 0 - \mu \cdot 0 = 0 - 0 = 0; \text{ e}$$

$$\begin{aligned} R(z) - R(z - 1) &= (\lambda P(z) + \mu Q(z)) - (\lambda P(z - 1) + \mu Q(z - 1)) \\ &= (\lambda P(z) - \lambda P(z - 1)) + (\mu Q(z) - \mu Q(z - 1)) \\ &= \lambda(P(z) - P(z - 1)) + \mu(Q(z) - Q(z - 1)) \\ &= \lambda p(z) + \mu q(z) = r(z). \end{aligned}$$

Então, calculamos:

$$\begin{aligned} \sum_{\nu=x}^y (\lambda p(\nu) + \mu q(\nu)) &= \sum_{\nu=x}^y r(\nu) \\ &= R(y) - R(x - 1) \\ &= (\lambda P(y) + \mu Q(y)) - (\lambda P(x - 1) + \mu Q(x - 1)) \\ &= (\lambda P(y) - \lambda P(x - 1)) + (\mu Q(y) - \mu Q(x - 1)) \\ &= \lambda(P(y) - P(x - 1)) + \mu(Q(y) - Q(x - 1)) \\ &= \lambda \sum_{\nu=x}^y p(\nu) + \mu \sum_{\nu=x}^y q(\nu). \end{aligned}$$

Para verificar que o axioma (S6) é satisfeito pela eq. (4), consideramos o espaço linear  $V_m$  de polinômios complexos de grau menor ou igual a  $m \in \mathbb{N}$ , e ao definirmos  $\|p\| := \sum_{i=0}^m |p(i)|$  para  $p \in V_m$ , introduzimos uma norma em  $V_m$ . Definimos então o operador

linear  $\sum_x^y : V_m \rightarrow \mathbb{C}$  através de  $\sum_x^y p := \sum_{\nu=x}^y p(\nu)$ . Como  $\dim(V_m) = m + 1 < \infty$ , o operador  $\sum_x^y$  é limitado (e, portanto, contínuo).

Assim, se  $(q_n)_{n \in \mathbb{N}} \subset V_m$  é uma sequência de polinômios que converge para zero na norma  $\|\cdot\|$  ao  $n \rightarrow \infty$ , isto é, se  $\lim_{n \rightarrow \infty} \|q_n\| = 0$ , vale a convergência  $\lim_{n \rightarrow \infty} \sum_x^y q_n = 0$  em  $\mathbb{C}$ .

Nesse contexto, a convergência constante na afirmação (3) do axioma (S6) é obtida ao considerarmos a sequência de polinômios  $(q_n)_{n \in \mathbb{N}}$  dada por  $q_n(z) := p(z + n) - p_n(z + n)$  e observarmos que a convergência pontual de  $|q_n(z)|$  para zero, dada como hipótese no axioma (S6), implica na convergência para zero da sequência  $q_n$  na norma  $\|\cdot\|$  e, por seguinte, na convergência para zero de  $\sum_x^y q_n$  em  $\mathbb{C}$ , isto é,

$$\lim_{n \rightarrow \infty} \left( \sum_{\nu=x}^y p(z + n) - \sum_{\nu=x}^y p_n(z + n) \right) = 0.$$

**Obs.:** Para a verificação do axioma (S6) foram necessários conceitos de análise funcional, que podem ser consultados em [2].

Para provarmos a afirmação (II), supomos que existe alguma definição que permite obter a soma fracionária  $\sum_{\nu=x}^y p(\nu)$  e que respeite os axiomas (S1) - (S4), e estendemos a ideia que usamos para mostrar que  $\sum_{\nu=1}^{1/2} c = \frac{c}{2}$ . Tomamos inteiros  $r, s \geq 1$  e, usando o axioma (S1), podemos escrever

$$\sum_{\nu=1}^r \nu^d = \sum_{\nu=1}^{\frac{r}{s}} \nu^d + \sum_{\nu=\frac{r}{s}+1}^{\frac{2r}{s}} \nu^d + \dots + \sum_{\nu=\frac{(s-2)r}{s}+1}^{\frac{(s-1)r}{s}} \nu^d + \sum_{\nu=\frac{(s-1)r}{s}+1}^r \nu^d,$$

onde o lado esquerdo é um somatório clássico, que sabemos determinar o valor. Usando os axiomas (S2) e (S3), podemos rearranjar o lado direito como segue:

$$\begin{aligned} \sum_{\nu=1}^r \nu^d &= \sum_{\nu=1}^{\frac{r}{s}} \nu^d + \sum_{\nu=\frac{r}{s}+1}^{\frac{r}{s}+\frac{r}{s}} \nu^d + \sum_{\nu=\frac{2r}{s}+1}^{\frac{r}{s}+\frac{2r}{s}} \nu^d + \dots + \sum_{\nu=\frac{(s-1)r}{s}+1}^{\frac{r}{s}+\frac{(s-1)r}{s}} \nu^d \\ &= \sum_{\nu=1}^{\frac{r}{s}} (\nu+0)^d + \sum_{\nu=1}^{\frac{r}{s}} \left(\nu+\frac{r}{s}\right)^d + \sum_{\nu=1}^{\frac{r}{s}} \left(\nu+\frac{2r}{s}\right)^d + \dots + \sum_{\nu=1}^{\frac{r}{s}} \left(\nu+\frac{(s-1)r}{s}\right)^d \\ &= \sum_{k=1}^{s-1} \sum_{\nu=1}^{\frac{r}{s}} \left[\left(\nu+\frac{kr}{s}\right)^d\right] \\ &= \sum_{k=1}^{s-1} \sum_{\nu=1}^{\frac{r}{s}} \left[\left(\nu+\frac{kr}{s}\right)^d + \nu^d - \nu^d\right] \\ &= \sum_{k=1}^{s-1} \sum_{\nu=1}^{\frac{r}{s}} \nu^d + \sum_{k=1}^{s-1} \sum_{\nu=1}^{\frac{r}{s}} \left[\left(\nu+\frac{kr}{s}\right)^d - \nu^d\right] \\ &= s \sum_{\nu=1}^{\frac{r}{s}} \nu^d + \sum_{k=1}^{s-1} \sum_{\nu=1}^{\frac{r}{s}} q_{d-1,k}(\nu) \end{aligned}$$

onde os polinômios  $q_{d-1,k}(\nu) = \left(\nu+\frac{kr}{s}\right)^d - \nu^d$  tem grau  $d-1$  (e observamos que  $q_{-1,k} \equiv 0$ ).

Agora argumentamos por indução. Se  $d=0$ , a equação anterior determina claramente o valor de  $\sum_{\nu=1}^{r/s} 1$ . Por linearidade, fica também determinada a soma de constantes arbitrárias  $c \in \mathbb{C}$  de 1 até  $r/s$ . Ao supormos que o valor da soma de qualquer polinômio de grau  $d-1$  é bem determinado, a igualdade anterior também permite encontrar o valor de  $\sum_{\nu=1}^{r/s} \nu^d$  e, por linearidade, concluímos ser possível determinar a soma de cada polinômio de grau  $d$ .

Então, usamos o axioma (S2) para escrever

$$(5) \quad \sum_{\nu=x}^y p(\nu) = \sum_{\nu=x-(x-1)+(x-1)}^{y-(x-1)+(x-1)} p(\nu) = \sum_{\nu=1+(x-1)}^{y-x+1+(x-1)} p(\nu) = \sum_{\nu=1}^{(y-x+1)} p(\nu+x-1)$$

e concluímos que os axiomas (S1) - (S4) nos permitem determinar o valor da soma  $\sum_{\nu=x}^y p(\nu)$  para todo  $x, y \in \mathbb{C}$  tais que  $y-x \in \mathbb{Q}$ .

Como visto na demonstração da afirmação (I), a equação (4) é uma possível definição que satisfaz os axiomas (S1) - (S4); então, usamos (5) e (4) para escrever

$$\sum_{\nu=1}^{(y-x+1)} p(\nu+x-1) = \sum_{\nu=x}^y p(\nu) = P(y) - P(x-1),$$

e concluir que (4) é a única definição possível para  $y-x \in \mathbb{Q}$ .

Finalmente, para concluir a demonstração da afirmação (III), basta notar que a restrição  $y-x \in \mathbb{Q}$  pode ser retirada ao supormos que a definição para obter a soma fracionária  $\sum_{\nu=x}^y p(\nu)$  respeite os axiomas (S1) - (S4) e também o axioma (S5). A holomorfia é necessária apenas para  $y-x \in \mathbb{C}$ , pois para  $y-x \in \mathbb{R}$  a equação (4) já será satisfeita ao exigirmos apenas continuidade no axioma (S5).  $\square$

Para estabelecer uma definição de soma fracionária para uma classe mais ampla de funções  $f : \mathbb{C} \rightarrow \mathbb{C}$ , tudo o que necessitamos exigir sobre a função  $f$  é que os valores de  $f(n+z)$  possam ser aproximados por alguma sequência de polinômios  $p_n(n+z)$ , de grau fixo, quando  $n \rightarrow +\infty$ , além de um pequeno cuidado com os domínios de definição  $U$  para o somatório fracionário: é necessária a propriedade

$$z \in U \quad \Rightarrow \quad z+1 \in U.$$

Convencionando que o polinômio nulo é o único polinômio de grau  $-\infty$ , apresentamos a seguinte definição:

**Definição 3.3:** Dados  $U \subset \mathbb{C}$  e  $m \in \mathbb{N} \cup \{-\infty\}$ , uma função  $f : U \rightarrow \mathbb{C}$  será chamada **fracionário-somável** de grau  $m$  se forem satisfeitas as seguintes condições:

- (i)  $x+1 \in U$  para todo  $x \in U$ ;
- (ii) Existe uma sequência de polinômios  $(p_n)_{n \in \mathbb{N}}$  de grau  $m$  fixo tal que, para todo  $x \in U$ , ocorre

$$|f(n+x) - p_n(n+x)| \rightarrow 0 \quad \text{quando } n \rightarrow +\infty;$$

- (iii) Para todo  $x, y \in U$ , existe o limite

$$\lim_{n \rightarrow \infty} \left( \sum_{\nu=n+x}^{n+y} p_n(\nu) + \sum_{\nu=1}^n (f(\nu+x-1) - f(\nu+y)) \right),$$

onde a soma  $\sum p_n$  é definida como em (4).

Se uma função  $f$  for fracionário-somável, usaremos a notação

$$\sum_{\nu=x}^y f(\nu) \quad \text{ou, brevemente,} \quad \sum_x^y f,$$

e vale a **fórmula fundamental do somatório fracionário**:

$$(6) \quad \sum_{\nu=x}^y f(\nu) = \sum_{\nu=1}^{\infty} (f(\nu+x-1) - f(\nu+y)).$$

Além disso, podemos definir **produtório fracionário** por

$$(7) \quad \prod_{\nu=x}^y f(\nu) := \exp\left(\sum_{\nu=x}^y \ln(f(\nu))\right),$$

sempre que  $\ln(f)$  for fracionário-somável.

A definição 3.3 não depende da escolha dos polinômios  $(p_n)_{n \in \mathbb{N}}$  que aproximam  $f$ . De fato, se  $(\tilde{p}_n)_{n \in \mathbb{N}}$  é outra escolha de polinômios que aproximam  $f$ , então vale

$$\lim_{n \rightarrow \infty} (p_n(n+x) - \tilde{p}_n(n+x)) = 0, \quad \forall x \in U$$

(e portanto para todo  $x \in \mathbb{C}$ ), pois o conjunto de polinômios de grau máximo  $m$  é um espaço linear de dimensão finita. Como visto na proposição 2.2, somas de polinômios satisfazem o axioma (S6), então, substituindo  $f$  por 0 e  $p_n$  por  $(\tilde{p}_n - p_n)$  na expressão (3), conclui-se facilmente que

$$\lim_{n \rightarrow \infty} \left( \sum_{\nu=n+x}^{n+y} p_n(\nu) - \sum_{\nu=n+x}^{n+y} \tilde{p}_n(\nu) \right) = 0.$$

A soma fracionária apresentada na definição 3.3 é a única que satisfaz os axiomas (S1) - (S6), como mostra o teorema a seguir:

**Teorema 3.4:** *A definição 3.3, de funções fracionário-somáveis, atende todos os axiomas (S1) - (S6) para domínios de definição adequados, e a fórmula fundamental do somatório fracionário (6) é a única que satisfaz aos axiomas (S1) - (S6) para a classe de funções considerada.*

**Demonstração:** Para mostrar que a fórmula fundamental do somatório fracionário (6) é a única possível para somas fracionárias satisfazendo a todos os axiomas (S1) - (S6),

consideramos uma soma do tipo  $\sum_{\nu=x}^y f(\nu)$  com  $x, y \in \mathbb{C}$ , e, com o uso do axioma (S1), escrevemos

$$\sum_{\nu=x}^y f(\nu) + \sum_{\nu=y+1}^{y+n} f(\nu) = \sum_{\nu=x}^{y+n} f(\nu) = \sum_{\nu=x}^{x+n-1} f(\nu) + \sum_{\nu=x+n}^{y+n} f(\nu),$$

onde  $n \in \mathbb{N}$  é arbitrário, e reorganizamos os termos como segue:

$$\sum_{\nu=x}^y f(\nu) = \sum_{\nu=x}^{x+n-1} f(\nu) - \sum_{\nu=y+1}^{y+n} f(\nu) + \sum_{\nu=x+n}^{y+n} f(\nu)$$

$$\begin{aligned}
&= \sum_{\nu=1+(x-1)}^{n+(x-1)} f(\nu) - \sum_{\nu=1+(y)}^{n+(y)} f(\nu) + \sum_{\nu=x+n}^{y+n} f(\nu) \\
(8) \quad &= \sum_{\nu=1}^n (f(\nu+x-1) - f(\nu+y)) + \sum_{\nu=x}^y f(\nu+n).
\end{aligned}$$

Na última linha da eq. (8) o primeiro somatório envolve um número inteiro de termos e pode ser avaliado classicamente. O problema a resolver fica então restrito à última soma do lado direito, onde o domínio da soma aparece deslocado  $n$  unidades para a direita. Como a eq. (8) vale para cada inteiro  $n$ , podemos passar a eq. (8) ao limite com  $n \rightarrow \infty$ , e então usar o axioma (S6): se  $f(n+z) \rightarrow 0$  quando  $n \rightarrow \infty$  para todo  $z$ , então o axioma (S6) implica que o limite quando  $n \rightarrow \infty$  na última soma da eq. (8) desaparece, e ficamos apenas com

$$\sum_{\nu=x}^y f(\nu) = \sum_{\nu=1}^{\infty} (f(\nu+x-1) - f(\nu+y)),$$

que apresentamos em (6) como a fórmula fundamental do somatório fracionário.

Resta provar que a fórmula fundamental do somatório fracionário (6) realmente satisfaz todos os axiomas (S1) - (S6).

Claramente, os axiomas (S3) e (S5) são automaticamente satisfeitos.

Para provar que os axiomas (S1), (S2) e (S4) são satisfeitos, basta substituir a soma  $\sum_x^y f(\nu)$  pela fórmula fundamental do somatório fracionário (6) no respectivo axioma, e fazer alguns poucos cálculos diretos. Por exemplo, substituindo a fórmula fundamental do somatório fracionário no lado esquerdo do axioma (S1), obtemos o lado direito:

$$\begin{aligned}
\sum_{\nu=x}^y f(\nu) + \sum_{\nu=y+1}^z f(\nu) &= \\
&= \sum_{\nu=1}^{\infty} (f(\nu+x-1) - f(\nu+y)) + \sum_{\nu=1}^{\infty} (f(\nu+y+1-1) - f(\nu+z)) \\
&= \sum_{\nu=1}^{\infty} ((f(\nu+x-1) - f(\nu+y)) + (f(\nu+y) - f(\nu+z))) \\
&= \sum_{\nu=1}^{\infty} (f(\nu+x-1) - f(\nu+z)) \\
&= \sum_{\nu=x}^z f(\nu).
\end{aligned}$$

Para provar que o axioma (S6) é satisfeito, utilizamos a fórmula fundamental do somatório fracionário (6) junto com os outros axiomas, e buscamos avaliar

$$\lim_{n \rightarrow \infty} \left( \sum_{\nu=x}^y f(\nu+n) - \sum_{\nu=x}^y p_n(\nu+n) \right)$$

da seguinte forma:

$$\begin{aligned}
& \lim_{n \rightarrow \infty} \left( \sum_{\nu=x}^y f(\nu+n) - \sum_{\nu=x}^y p_n(\nu+n) \right) = \\
& = \lim_{n \rightarrow \infty} \left( \sum_{\nu=x}^y f(\nu+n) - \sum_{\nu=x}^y p_n(\nu+n) \right. \\
& \quad \left. - \sum_{\nu=1}^n (f(\nu+x-1) - f(\nu+y)) + \sum_{\nu=1}^n (f(\nu+x-1) - f(\nu+y)) \right) \\
& = \lim_{n \rightarrow \infty} \left( \sum_{\nu=x}^y f(\nu+n) - \sum_{\nu=x}^y f(\nu+n) \right. \\
& \quad \left. - \sum_{\nu=1}^n (f(\nu+x-1) - f(\nu+y)) + \sum_{\nu=1}^n (f(\nu+x-1) - f(\nu+y)) \right) \\
& = \lim_{n \rightarrow \infty} \left( \sum_{\nu=x}^y f(\nu+n) - \sum_{\nu=x}^y f(\nu) + \sum_{\nu=1}^n (f(\nu+x-1) - f(\nu+y)) \right) \\
& = \lim_{n \rightarrow \infty} \left( \sum_{\nu=x}^y f(\nu+n) - \sum_{\nu=x}^y f(\nu) + \sum_{\nu=1}^n f(\nu+x-1) - \sum_{\nu=1}^n f(\nu+y) \right) \\
& = \lim_{n \rightarrow \infty} \left( \sum_{\nu=x+n}^{y+n} f(\nu) - \sum_{\nu=x}^y f(\nu) + \sum_{\nu=x}^{x+n-1} f(\nu) - \sum_{\nu=y+1}^{y+n} f(\nu) \right) \\
& = \lim_{n \rightarrow \infty} \left( \sum_{\nu=x}^{x+n-1} f(\nu) + \sum_{\nu=x+n}^{y+n} f(\nu) - \sum_{\nu=x}^y f(\nu) - \sum_{\nu=y+1}^{y+n} f(\nu) \right) \\
& = \lim_{n \rightarrow \infty} \left( \sum_{\nu=x}^{y+n} f(\nu) - \sum_{\nu=x}^{y+n} f(\nu) \right) = 0.
\end{aligned}$$

Isso completa a demonstração de que a fórmula fundamental do somatório fracionário (6), para uma função  $f : \mathbb{C} \rightarrow \mathbb{C}$  atendendo às condições dadas na definição 3.3, satisfaz a todos os axiomas (S1) - (S6).  $\square$

Na próxima seção apresentamos alguns exemplos de somas fracionárias obtidas com a aplicação da definição 3.3 e da fórmula fundamental do somatório fracionário (6).

#### 4. ALGUNS EXEMPLOS DE SOMAS FRACIONÁRIAS

Agora que conhecemos uma definição adequada para funções fracionário-somáveis e uma fórmula que permite obter o valor da soma de uma quantidade não inteira de termos (definição 3.3 e fórmula (6)), é natural a curiosidade por conhecer algumas das propriedades satisfeitas pelos somatórios clássicos (finitos) que permanecem válidas com essa abordagem mais geral, e também por conhecer algumas outras identidades que não encontram equivalência no caso clássico.

Uma das mais básicas propriedades para somas finitas é a soma da série geométrica, que permanece válida para somatórios fracionários. De fato, considerando  $0 \leq q < 1$  por simplicidade, como a função  $f(z) = q^z$  é “aproximadamente zero” para valores grandes de  $z$ , e vale  $\lim_{n \rightarrow \infty} q^{z+n} = 0$  para todo  $z \in \mathbb{C}$ , concluímos que  $q^z$  é fracionário-somável, e, usando a fórmula fundamental do somatório fracionário (6), obtemos:

$$\sum_{\nu=0}^z q^{\nu} = \sum_{\nu=1}^{\infty} (q^{\nu-1} - q^{\nu+z}) = (1 - q^{z+1}) \sum_{\nu=1}^{\infty} q^{\nu-1} = \frac{1 - q^{z+1}}{1 - q},$$

ou seja, a conhecida fórmula para a soma da série geométrica permanece válida para todo  $z \in \mathbb{C}$ .

No próximo exemplo, consideramos uma função cujo comportamento assintótico não se aproxima de zero, e usamos somatórios e produtórios fracionários para extrapolar a função fatorial clássica para todo  $z \in \mathbb{C}$ , exceto  $z \in \{0, -1, -2, -3, \dots\}$ . Inicialmente, reescrevemos a eq. (8) para a função  $f(z) = \ln(z)$ :

$$(9) \quad \sum_{\nu=x}^y \ln(\nu) = \sum_{\nu=1}^n (\ln(\nu+x-1) - \ln(\nu+y)) + \sum_{\nu=x}^y \ln(\nu+n)$$

onde, para o último termo da eq.(9), notamos que para  $z \in [x, y] \subset \mathbb{R}^+$ , os valores  $\ln(z+n)$  são relativamente bem aproximados pela função constante  $\ln(n)$ , com um erro que tende a zero quando  $n \rightarrow \infty$  (dizemos que  $f = \ln$  é “aproximadamente constante” quando seu argumento tende ao infinito). Então, usando o axioma (S3), podemos escrever

$$(10) \quad \sum_{\nu=x}^y \ln(\nu+n) = \sum_{\nu=x}^y \ln n + \sum_{\nu=x}^y (\ln(\nu+x) - \ln(n))$$

para cada  $n \in \mathbb{N}$ . Na eq. (10), a primeira soma do lado direito tem somando constante e pode ser avaliada através da eq. (4), e, pelo axioma (S6), a última soma se anula quando  $n \rightarrow \infty$ . Assim, considerando  $\ln(n)$  constante antes de passar a eq. (10) ao limite, podemos concluir que

$$\lim_{n \rightarrow \infty} \sum_{\nu=x}^y \ln(\nu+n) = \ln(n) \sum_{\nu=x}^y 1 = (y-x+1) \ln(n).$$

Portanto, ao passar a eq. (9) ao limite com  $n \rightarrow \infty$ , obtemos

$$\sum_{\nu=x}^y \ln(\nu) = \lim_{n \rightarrow \infty} \left( \sum_{\nu=1}^n (\ln(\nu+x-1) - \ln(\nu+y)) + (y-x+1) \ln n \right),$$

de onde segue que

$$\exp\left(\sum_{\nu=1}^z \ln \nu\right) = \lim_{n \rightarrow \infty} \left( \exp\left(\sum_{\nu=1}^n (\ln(\nu) - \ln(\nu+z)) + z \ln n\right) \right).$$

Então, tomando a liberdade de escrever  $z!$  para  $z \in \mathbb{R}$  ou  $z \in \mathbb{C}$ , e usando a definição de produtório fracionário dada na eq. (7), com  $\nu$  no lugar de  $f(\nu)$ , calculamos

$$\begin{aligned}
 z! &\equiv \prod_{\nu=1}^z \nu = \exp\left(\sum_{\nu=1}^z \ln(\nu)\right) \\
 &= \lim_{n \rightarrow \infty} \left( n^z \exp \sum_{\nu=1}^n \ln\left(\frac{\nu}{\nu+z}\right) \right) \\
 &= \lim_{n \rightarrow \infty} \left( n^z \cdot \prod_{\nu=1}^n \left(\frac{\nu}{\nu+z}\right) \right) \\
 (11) \quad &= \Gamma(z+1);
 \end{aligned}$$

obtendo assim a *função gama*, que, para  $z \in \mathbb{C} \setminus \{0, -1, -2, -3, \dots\}$ , é definida por

$$\Gamma(z) := \int_0^{\infty} e^{-t} t^{z-1} dt,$$

e tem a seguinte representação na forma de limite (ver, e.g., [14] ou [21]):

$$\Gamma(z) = \lim_{n \rightarrow \infty} \frac{n! n^z}{z(z+1) \cdots (z+n)},$$

utilizada na última passagem para obter a eq. (11), junto com a propriedade  $z \Gamma(z) = \Gamma(z+1)$ .

Outra propriedade que permanece válida para somas fracionárias é a expansão da série binomial. Para recuperá-la, inicialmente lembramos que se  $\alpha$  e  $j$  são naturais, vale

$$\binom{\alpha}{j} = \frac{\alpha!}{j!(\alpha-j)!},$$

mas se  $\alpha$  e  $j \in \mathbb{C}$ , então vale (ver [14] ou [21]):

$$\binom{\alpha}{j} = \frac{\Gamma(\alpha+1)}{\Gamma(j+1)\Gamma(\alpha-j+1)}.$$

Consideramos então  $x \in \mathbb{C}$ , com  $|x| < 1$ , e  $\alpha \in \mathbb{C} \setminus \{-1, -2, -3, \dots\}$  e, usando a fórmula fundamental do somatório fracionário (6) para a função

$$f(\nu) = \binom{\alpha}{\nu} z^{\nu},$$

obtemos

$$(12) \quad \sum_{\nu=0}^x \binom{\alpha}{\nu} z^{\nu} = \sum_{\nu=1}^{\infty} \left( \binom{\alpha}{\nu-1} z^{\nu-1} - \binom{\alpha}{\nu+x} z^{\nu+x} \right).$$

Então fazemos  $x = \alpha$  na eq. (12). Em seu último termo ficamos com

$$\binom{\alpha}{\nu+\alpha} = \frac{\Gamma(\alpha+1)}{\Gamma(\nu+\alpha+1)\Gamma(-\nu+1)},$$



e como para todo  $\nu \in \mathbb{N}^*$  vale  $\Gamma(1 - \nu) = \pm\infty$  e isso ocorre no último termo do somatório no lado direito da eq. (12), sempre teremos “algo” do tipo  $\frac{\Gamma(\alpha + 1)}{\pm\infty}$ , que deve ser considerado nulo. Assim, a eq. (12) pode ser reescrita como

$$\begin{aligned} \sum_{\nu=0}^{\alpha} \binom{\alpha}{\nu} z^{\nu} &= \sum_{\nu=1}^{\infty} \binom{\alpha}{\nu-1} z^{\nu-1} \\ &= (1+z)^{\alpha}, \end{aligned}$$

onde aparece a expressão expandida da série binomial (ver [8], [15]).

Apresentamos a seguir duas novas identidades obtidas com o uso de somatórios e produtórios fracionários, e que não encontram alguma expressão equivalente com somatórios clássicos.

Com auxílio das seguintes identidades da função gama

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\operatorname{sen}(\pi z)} = \frac{i\pi}{\sinh(\pi iz)}$$

(ver [21]) e da expressão da função gama em termos de um produtório fracionário obtida na eq. (11), podemos obter a primeira nova identidade:

$$(13) \quad \tanh(\pi) = \prod_{\nu=1}^{-1/2} (\nu^2 + 1).$$

A prova de (13), por ser um pouco mais trabalhosa, não será apresentada neste artigo, mas o leitor interessado pode acompanhar o desenvolvimento passo-a-passo em [11].

Para a segunda nova identidade, que apresentamos como exemplo final neste artigo, começamos por considerar a série harmônica. A função  $f(z) = z^{-1}$ , que pode ser considerada “aproximadamente zero” para  $z$  grande, é fracionária-somável e podemos, portanto, aplicar a fórmula fundamental do somatório fracionário (6), obtendo:

$$\sum_{\nu=1}^x \frac{1}{\nu} = \sum_{\nu=1}^{\infty} \left( \frac{1}{\nu+1-1} - \frac{1}{\nu+x} \right) = \sum_{\nu=1}^{\infty} \left( \frac{1}{\nu} - \frac{1}{\nu+x} \right)$$

e, em particular:

$$\begin{aligned} \sum_{\nu=1}^{-1/2} \frac{1}{\nu} &= \sum_{\nu=1}^{\infty} \left( \frac{1}{\nu} - \frac{1}{\nu - \frac{1}{2}} \right) = \sum_{\nu=1}^{\infty} \left( \frac{1}{\nu} - \frac{2}{2\nu-1} \right) \\ &= (1-2) + \left( \frac{1}{2} - \frac{2}{3} \right) + \left( \frac{1}{3} - \frac{2}{5} \right) + \left( \frac{1}{4} - \frac{2}{7} \right) + \dots \\ &= (-2+1) + \left( -\frac{2}{3} + \frac{1}{2} \right) + \left( -\frac{2}{5} + \frac{1}{3} \right) + \left( -\frac{2}{7} + \frac{1}{4} \right) + \dots \\ &= (-2+1) + \left( -\frac{2}{3} + \frac{2}{4} \right) + \left( -\frac{2}{5} + \frac{2}{6} \right) + \left( -\frac{2}{7} + \frac{2}{8} \right) + \dots \\ &= -2 \left( 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10} + \dots \right) \\ &= -2 \ln(2). \end{aligned}$$

Recuperamos, assim, o resultado (1), isto é,

$$\sum_{\nu=1}^{-1/2} \frac{1}{\nu} = -2 \ln 2,$$

o primeiro exemplo de um somatório fracionário, obtido e anunciado por Euler em 1755 [7]. Como uma curiosidade final, destacamos que a soma

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10} + \dots$$

é a expansão em série de MacLaurin da função  $\ln(x)$ , avaliada em  $x = 2$ .

## 5. CONSIDERAÇÕES FINAIS

Apresentamos, neste texto, uma introdução em língua portuguesa à expansão da teoria concernente à mais básica das operações matemáticas: a adição. Somatórios clássicos, usados para somar uma quantidade finita de termos, cuja notação sigma foi introduzida por Euler (e.g., [6]) e difundida a partir de seu uso por Fourier [10], são considerados conhecidos por todos, em maior ou menor profundidade. Somatórios infinitos de termos reais, conhecidos por séries numéricas, são ensinados em nossos cursos de graduação e alguns casos até no ensino médio. A grande “novidade” introduzida por M. Müller e D. Schleicher ([18], [19], [20]) mas já apresentada, de alguma forma, por Euler [7] é a possibilidade de efetuar somas finitas contendo uma quantidade não inteira de parcelas. Com o advento da teoria para somatórios fracionários (consistente com a teoria para somatórios clássicos) passa a ser permitido somar sobre uma quantidade racional, real e mesmo complexa de termos.

Em um primeiro contato com esta teoria, é comum pensar que se trata de uma espécie de “soma contínua”, talvez um pouco semelhante com o somatório infinitesimal costumadamente apresentado como definição da integral definida; alertamos que isso é uma falácia, como o axioma (S1) induz a compreender. Quando consideramos um somatório do tipo

$$\sum_{\nu=x}^y f(\nu)$$

onde  $f$  é uma função fracionário-somável e  $x, y \in U \subset \mathbb{R}$  são dois números bem determinados, tal soma tem uma quantidade finita de parcelas, onde (possivelmente) apenas uma parcela é fracionária, sendo as demais parcelas inteiras. Por exemplo, com o uso do axioma (S1) a expressão

$$\sum_{\nu=1}^{3,5} f(\nu)$$

pode ser reescrita em uma das seguintes formas:

$$\sum_{\nu=1}^2 f(\nu) + \sum_{\nu=3}^{3,5} f(\nu) \quad \text{ou} \quad \sum_{\nu=1}^3 f(\nu) + \sum_{\nu=4}^{3,5} f(\nu);$$

ou seja,

$$f(1) + f(2) + \sum_3^{3,5} f(v) \quad \text{ou} \quad f(1) + f(2) + f(3) + \sum_4^{3,5} f(v),$$

em ambas temos uma soma de uma quantidade inteira de parcelas do tipo  $f(v)$  com uma parcela fracionária, ou seja, de “comprimento” menor que uma unidade. Quando  $x, y \in U \subset \mathbb{C}$ , ainda não está totalmente claro para os autores qual seria a natureza da soma  $\sum_x^y f(v)$ ; a única certeza é que não se trata de uma soma de infinitas parcelas. Além dessa certeza, por enquanto apenas conseguimos conjecturar que existe algum tipo de “independência do caminho” da soma, pois a fórmula fundamental do somatório fracionário (6) considera explicitamente valores de  $f$  com argumentos dependendo dos limites inferior e superior  $x$  e  $y$  do somatório fracionário (fato também percebido em outros cálculos rabiscados e ainda não publicados).

Acreditamos que esta teoria para somatórios fracionários, desenvolvida recentemente, guarda ainda belas propriedades e identidades surpreendentes a serem descobertas e talvez até encontre utilidade na modelagem de situações reais, talvez com um belo futuro no desenvolvimento tecnológico. Esperamos que este artigo desperte o interesse no leitor em buscar aprender mais sobre o tema, e fazer suas próprias especulações. Recomendamos aos interessados a leitura da dissertação [11], escrita em português, dos artigos [18], [19], e [20], escritos pelos autores que deram o passo inicial desta teoria, e do livro [1].

#### REFERÊNCIAS

- [1] Alabdulmohsin, I. M. - *Summability Calculus* A comprehensive theory of fractional finite sums. Springer, Switzerland (2018).
- [2] Botelho, G. M. A.; Pellegrino, D. M.; Teixeira, E. V. - *Fundamentos de Análise Funcional*. Textos Universitários, 2ª ed., Editora SBM, Rio de Janeiro (2015).
- [3] Cajori, F. - *A History of Mathematical Notations*. Dover Publications, New York (1993).
- [4] Camargo, R. F.; Oliveira, E. C. - *Cálculo Fracionário*. Editora Livraria da Física, São Paulo (2015).
- [5] Caminha, A. - *Tópicos de Matemática Elementar* Volume 6 - Polinômios. 2ª ed., Editora SBM, Rio de Janeiro (2016).
- [6] Euler, L. - *Institutiones Calculi Differentialis* cum eius usu in Analysi Finitorum ac Doctrina Serierum [E212]. Petropolitanae Academiae Sumptibus, (1755); Republicado em Opera Omnia, series 1, vol 10.
- [7] Euler, L. - *Dilucitationes in capita postrema calculi mei differentialis de funcionibus inexplicabilibus* [E613]. Supplementum in [E212] (1755); republicado em Memoires de l'academie des sciences de St.-Petersbourg 4 (1813), 88-119; republicado em Opera Omnia, series 1, vol 16, 1-33.
- [8] Euler, L. - *Nova demonstratio quod potestatem Binomii Newtoniana etiam pro exponentibus fractis valeat* [E637]. Nova Acta Academiae Scientiarvm Imperialis Petropolitanae, Tomvs V (1787), 52-58 (Mathematica).
- [9] Fernandez, C. S.; Bernardes Jr., N. C. - *Introdução às Funções de uma Variável Complexa*. Textos Universitários, 5ª ed., Editora SBM, Rio de Janeiro (2019).
- [10] Fourier, J.-B. J. - *La théorie analytique de la chaleur*. Chez Firmin Didot, Père ed Fils, Paris (1822).
- [11] Galvão, A. T. - *Somas, somatórios e termos não inteiros*. Dissertação de Mestrado (PROFMAT), UEPG, Ponta Grossa (2020).
- [12] Graham, R. L.; Knuth, D. E.; Patashnik, O. - *Concrete mathematics: a foundation for computer science*. 2ª ed., Addison-Wesley Publishing, New York (1994).
- [13] Hefez, A.; Villela, M. L. T. - *Polinômios e Equações Algébricas*. Coleção PROFMAT, Editora SBM, Rio de Janeiro (2012).

- [14] Kobuchei, A.; Luchko, Y. (Eds.) - *Handbook of Fractional Calculus with Applications*, vol. 1: Basic Theory. De Gruyter, Berlin (2019).
- [15] Leachenski, A. A. - *Binômio de Newton com expoente negativo e fracionário*. Dissertação de Mestrado (PROFMAT), UEPG, Ponta Grossa (2017).
- [16] Tenreiro Machado, J.; Lopes, A. M.; Duarte, F. B.; Ortigueira, M. D.; Rato, R. T. - *Rhapsody in Fractional*. *Fractional Calculus & Applied Analysis*, Vol. 17, No. 4 (2014), 1188-1214.
- [17] Matos, M. P. - *Séries e Equações Diferenciais*. Ed. Ciência Moderna, Rio de Janeiro (2017).
- [18] Müller, M.; Schleicher, D. - *How to Add a Non-Integer Number of Terms, and How to Produce Unusual Infinite Summations*. *Journal of Computational and Applied Mathematics* 178 1–2 (2005), 347–360.
- [19] Müller, M.; Schleicher, D. - *Fractional sums and Euler-like identities*. *Ramanujan J.*, 21 (2010). 347-360.
- [20] Müller, M.; Schleicher, D. - *How to add a noninteger number of Terms: From axioms to new identities*. *The American Mathematical Monthly*. Vol. 118, nº 2 (2011). 136-152.
- [21] Podlubny, I. - *Fractional differential equations: an introduction to fractional derivatives, fractional differential equations, to methods of their solution and some of their applications*. Vol. 198, Academic Press, London (1999).
- [22] Rosa, C. A. P. - *História da ciência: a ciência moderna*. 2ª ed., FUNAG, Brasília (2012).

DEPARTAMENTO DE MATEMÁTICA E ESTATÍSTICA, UEPG, 84030-900, PONTA GROSSA, PR  
Email address: agdatg@gmail.com, jochemarchagas@uepg.br

## QUANDO OS MÉTODOS DE EULER E DE NEWTON COINCIDEM

JOSÉ CLAUDINEI FERREIRA

**Resumo:** Este texto é dedicado ao estudo do uso de problemas de otimização e equações diferenciais para a dedução de métodos numéricos para a determinação de zeros de funções. Em particular, tratamos de estimativas sobre convergência exponencial para o método de Euler, em duas abordagens distintas, sendo que em uma delas, o método de Newton surge como uma das possibilidades. Acreditamos que a estrutura do texto não é usual, em textos introdutórios ou avançados sobre o tema. Aproveitamos a oportunidade para estimular a repensarmos as formas de apresentar conteúdos ou conceitos em matemática, porque acreditamos que a organização de resultados, ou a performance em um texto, sobre matemática pode dar maior ou menor impacto na compreensão e no interesse do leitor.

### 1. INTRODUÇÃO

Um raciocínio comum para aqueles que lidam com matemática, no meio das chamadas ciências exatas, envolve lidar com variáveis e, para alguns, é isso que amplia o conceito do que é matemática para além do conceito de lidar com números, ou de fazer aquilo que a calculadora do caixa de supermercado ou de alguma loja faz.

Um primeiro contato com variáveis ou incógnitas aparece quando os jovens alunos do Ensino Fundamental se deparam com equações como  $x + 2 = 4$  ou  $x^2 + 2x - 1 = 0$ . Mais adiante, talvez em cursos de graduação, alguns se deparam com equações como  $3 \cos(x) = 2$ ,  $x^5 + x - 1 = 0$  ou uma mais interessante (ou complexa)  $x^2 + 1 = 0$ . Mas será que sabemos resolver tais equações? Bem, sem calculadora, muito poucas. E com calculadoras ou computadores, só algumas. Isso talvez motive alguns professores de matemática a dizerem que matemático não é calculadora, e de fato não é, e achamos que nem deveria tentar ser. Mas, certamente, alguns dos argumentos que contribuíram para o

---

Data de aceitação: Janeiro de 2021.

*Palavras chave.* Resolução numérica de equações, Método de Euler, Método de Newton, Otimização.

Agradeço ao parecerista pela leitura cuidadosa e pelos ótimos comentários e sugestões, que certamente melhoraram versão anterior deste texto. Agradeço aos professores José P. C. dos Santos e Nelson H. T. Lemes pela oportunidade de discutir e tentar pensar em questões envolvendo Análise Numérica e Equações Diferenciais. Agradeço minha esposa Rejane Siqueira Julio pelas conversas, pela companhia, pelo incentivo para tentar escrever e pelos nossos filhos Alice Julio Ferreira e Miguel Julio Ferreira que, de uma forma ou de outra, também contribuem com sua companhia entre pausas, brincadeiras e muitas outras boas coisas!

funcionamento das calculadoras e dos computadores surgiram com a matemática, ou com algum uso dela.

Como as equações estão presentes em muitos problemas que surgem em aplicações e em teoria matemática é comum uma breve apresentação desse tema, em cursos de formação, para tratar de métodos para resolver equações, ou para aproximação de soluções, quando soluções existem.

Neste texto tratamos de duas formas de aproximação de soluções de equações, por meio do uso do método de Euler para a resolução numérica de equações diferenciais ordinárias que surgem em problemas de minimização. Em uma dessas formas o método de Newton surge como uma das possibilidades.

Os métodos de Euler e de Newton costumam ser apresentadas a discentes de cursos de graduação em matemática, computação e engenharia, por exemplo. Mas nem sempre são apresentados da forma que pretendemos aqui. Aproveitamos isso como uma oportunidade de mostrar que a matemática, seja lá o que isso for, não parece ser algo dividido em partes disjuntas, mas sim possuir intersecções importantes entre os mais diversos temas.

O problema que tentamos introduzir nesses breves comentários e o caminho que pretendemos seguir neste texto, pode, de certa forma, ser formalizado neste final de seção.

Antes de qualquer formalismo, alertamos o leitor que vamos usar símbolo em negrito para denotar vetores, como o vetor  $\mathbf{u}$ , e itálico para denotar números reais, como o número  $t$ .

Vamos denotar o módulo ou a norma euclidiana de um vetor por  $\|\cdot\|$ , que depende do produto interno usual (ou escalar) aqui denotado por  $\langle \cdot, \cdot \rangle$ . Quanto a matrizes  $A_{m \times m}$ , denotamos

$$\|A\| = \max_{\|\mathbf{u}\|=1} \|A\mathbf{u}\|, \quad \mathbf{u} \in \mathbb{R}^m.$$

Consideramos uma função  $\mathbf{f} : U \rightarrow \mathbb{R}^m$ , que por simplicidade supomos que possua derivadas contínuas até a ordem 2, no conjunto aberto  $U \subset \mathbb{R}^m$ . Supomos ainda que a equação

$$\mathbf{f}(\mathbf{u}) = (f_1(\mathbf{u}), f_2(\mathbf{u}), \dots, f_m(\mathbf{u})) = \mathbf{0}$$

possua solução  $\mathbf{a} \in U$ . Nesse caso, cada  $f_i : U \rightarrow \mathbb{R}$  é uma função que possui derivada de segunda ordem contínua e tal que  $f_i(\mathbf{a}) = 0$ . Argumento semelhante pode ser aplicado à função

$$g(\mathbf{u}) = \frac{1}{2} \|\mathbf{f}(\mathbf{u})\|^2, \quad \mathbf{u} \in U,$$

em que

$$0 = g(\mathbf{a}) = \min_{\mathbf{u} \in U} g(\mathbf{u}).$$

Isso mostra que a determinação de raízes de equações está ligado a problemas de otimização. O que, de certa forma, é sabido por quem já cursou disciplinas de Cálculo Diferencial, porque nessas disciplinas problemas de otimização estão relacionados ao problema de determinar as raízes da função gradiente de  $g(\mathbf{u})$ , por exemplo. É um tema antigo e que está ligado a muitos artigos recentes ([1, 2, 3, 4]); o tema vai e volta em meio a outros interesses que vão surgindo ([5, 6]).

Nesse ponto, cabe mencionar que a ideia de escrever este texto surgiu após leitura do artigo [2] e de perceber que a compreensão do tema tratado lá parecera-nos muito mais intuitiva se o texto fosse conduzido numa ordem diferente, e é isso que buscamos aqui.

## 2. A DERIVADA E A DERIVADA DIRECIONAL

Lembramos rapidamente que a derivada da função  $f$ , no ponto  $\mathbf{u}$ , é uma transformação linear que denotamos como  $f'(\mathbf{u}) : \mathbb{R}^m \rightarrow \mathbb{R}^m$ , que melhor aproxima a variação da função  $f$ , quando aplicada a pontos próximos de  $\mathbf{u}$ . Podemos também visualizar  $f'(\mathbf{u}) = Jf(\mathbf{u})$  como a matriz

$$Jf(\mathbf{u}) = \begin{bmatrix} \frac{\partial f_1(\mathbf{u})}{\partial u_1} & \frac{\partial f_1(\mathbf{u})}{\partial u_2} & \cdots & \frac{\partial f_1(\mathbf{u})}{\partial u_m} \\ \frac{\partial f_2(\mathbf{u})}{\partial u_1} & \frac{\partial f_2(\mathbf{u})}{\partial u_2} & \cdots & \frac{\partial f_2(\mathbf{u})}{\partial u_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{u})}{\partial u_1} & \frac{\partial f_m(\mathbf{u})}{\partial u_2} & \cdots & \frac{\partial f_m(\mathbf{u})}{\partial u_m} \end{bmatrix},$$

cujas entradas são derivadas parciais de  $f_i$ , e que também é conhecida como matriz Jacobiana de  $f(\mathbf{u})$ . A fim de simplificar a notação vamos considerar o vetor  $\mathbf{u}$  como uma matriz coluna e  $f'(\mathbf{u})\mathbf{h}$  denotará a multiplicação da matriz  $Jf(\mathbf{u})$  por  $\mathbf{h}$ , ou a aplicação da transformação linear  $f'(\mathbf{u})$  no vetor  $\mathbf{h}$ , que é a derivada direcional de  $f$ , no ponto  $\mathbf{u}$  e na direção  $\mathbf{h}$  ([7, 8, 9]).

Em resumo, vale a relação

$$(1) \quad \mathbf{f}(\mathbf{u} + \mathbf{h}) = \mathbf{f}(\mathbf{u}) + f'(\mathbf{u})\mathbf{h} + \mathbf{r}(\mathbf{h}),$$

com

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|\mathbf{r}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0,$$

que é equivalente à definição da derivada de  $f$  no ponto  $\mathbf{u}$ .

Isso (ou a regra da cadeia) nos ajuda a calcular a derivada de  $g$ , no ponto  $\mathbf{u}$ . Antes, observe que

$$(2) \quad g(\mathbf{u}) = \frac{1}{2}\|\mathbf{f}(\mathbf{u})\|^2 = \frac{1}{2}\langle \mathbf{f}(\mathbf{u}), \mathbf{f}(\mathbf{u}) \rangle,$$

denotando o produto escalar usual de  $\mathbb{R}^m$  como  $\langle \cdot, \cdot \rangle$  ([9]). Segue então que

$$\begin{aligned} g(\mathbf{u} + \mathbf{h}) &= \frac{1}{2}\langle \mathbf{f}(\mathbf{u} + \mathbf{h}), \mathbf{f}(\mathbf{u} + \mathbf{h}) \rangle \\ &= \frac{1}{2}\langle \mathbf{f}(\mathbf{u}) + f'(\mathbf{u})\mathbf{h} + \mathbf{r}(\mathbf{h}), \mathbf{f}(\mathbf{u}) + f'(\mathbf{u})\mathbf{h} + \mathbf{r}(\mathbf{h}) \rangle \\ &= \frac{1}{2}\langle \mathbf{f}(\mathbf{u}), \mathbf{f}(\mathbf{u}) \rangle + \langle f'(\mathbf{u})\mathbf{h}, \mathbf{f}(\mathbf{u}) \rangle + \mathbf{R}(\mathbf{h}) \end{aligned}$$

com

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\|\mathbf{R}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0.$$

O que nos diz que

$$g'(\mathbf{u})\mathbf{h} = \langle f'(\mathbf{u})\mathbf{h}, \mathbf{f}(\mathbf{u}) \rangle = \langle \mathbf{h}, [f'(\mathbf{u})]^* \mathbf{f}(\mathbf{u}) \rangle,$$

em que  $[f'(\mathbf{u})]^*$  denota a matriz adjunta (no caso, a transposta) da matriz  $Jf(\mathbf{u})$ , considerando  $\mathbf{f}(\mathbf{u})$  também como matriz coluna.

Como  $g(\mathbf{u})$  é um número real, é comum denotarmos

$$g'(\mathbf{u}) = [f'(\mathbf{u})]^* \mathbf{f}(\mathbf{u}) = \nabla g(\mathbf{u}),$$

em que  $\nabla$  representa o vetor gradiente, o que implica na igualdade

$$(3) \quad g'(\mathbf{u})\mathbf{h} = \langle \mathbf{h}, \nabla g(\mathbf{u}) \rangle,$$

que é a derivada direcional de  $g$ , no ponto  $\mathbf{u}$  e na direção  $\mathbf{h}$ .

Para uso seguinte, observamos que a Expressão (1) pode também ser escrita como

$$(4) \quad \mathbf{f}(\mathbf{u} + \mathbf{h}) = \mathbf{f}(\mathbf{u}) + \int_0^1 \mathbf{f}'(\mathbf{u} + t\mathbf{h})\mathbf{h} dt.$$

Isso produz de imediato a desigualdade do valor médio

$$(5) \quad \|\mathbf{f}(\mathbf{u} + \mathbf{h}) - \mathbf{f}(\mathbf{u})\| \leq \max_{0 \leq t \leq 1} \|\mathbf{f}'(\mathbf{u} + t\mathbf{h})\| \|\mathbf{h}\|.$$

**2.1. A direção de maior decrescimento.** Lembramos que a [direção de maior decrescimento](#) de  $g(\mathbf{u})$ , partindo de um ponto  $\mathbf{u}_0 \in U$ , é dada pela direção oposta a do vetor gradiente  $\nabla g(\mathbf{u}_0)$ .

Isso define uma curva (ou caminho)  $\mathbf{u}(t) \in U$ , para  $0 \leq t \leq T$ , cuja derivada coincide com o oposto do vetor gradiente de  $g$ , no ponto  $\mathbf{u}(t)$ . Ou seja,  $\mathbf{u}(t)$  é solução da equação diferencial ordinária ([10, 11]), ou [problema de valor inicial](#),

$$(6) \quad \begin{cases} \mathbf{u}'(t) = -\nabla g(\mathbf{u}(t)) \\ \mathbf{u}(0) = \mathbf{u}_0 \end{cases},$$

o que quer dizer que  $g(\mathbf{u}(t))$  é decrescente, desde que  $\nabla g(\mathbf{u}(t)) \neq \mathbf{0}$ .

No caso em que  $\mathbf{u}_0$  é suficientemente próximo de  $\mathbf{a}$ , isso sugere que  $\mathbf{u}(t)$  está ainda mais próximo de  $\mathbf{a}$ , para  $t > 0$ , o que melhora quando  $t$  cresce ([1]). Um olhar mais atento para questões de unicidade da solução da Equação (6) sugere também que  $\mathbf{u}(t)$  nunca atinge  $\mathbf{a}$ , a menos que  $\mathbf{u}_0$  seja  $\mathbf{a}$ .

Nesse ponto, podemos pensar, bem, eu tinha uma equação  $\mathbf{f}(\mathbf{u}) = \mathbf{0}$  para resolver e agora tenho uma [equação diferencial](#) para resolver. Parece que a situação piorou um pouco! E isso tem como agravante o fato de não sabermos resolver a maioria das equações diferenciais. Ou ainda o fato de equações diferenciais não ser um tema tratado de forma explícita em muitos cursos de graduação em áreas exatas. Mas acreditamos que isso mudaria se prestássemos mais atenção às [leis de Newton](#) e se permitíssemos e incentivássemos, em salas aula, mais conversas direcionadas aos motivos de estudarmos certos conceitos, como os relacionados com as disciplinas de Cálculo e de Física.

### 3. O MÉTODO DE EULER

Nem sempre é possível determinar a função  $\mathbf{u}(t)$  que é solução da Equação (6) e, por isso, vamos discutir como aproximar  $\mathbf{u}(h)$ ,  $\mathbf{u}(2h)$ ,  $\dots$ ,  $\mathbf{u}(ph)$ , para algum número real  $0 < T$  e  $ph = T$ , para algum número natural  $p$ .

Como supomos que  $\mathbf{f}(\mathbf{u})$  tem derivadas contínuas até a ordem 2, então podemos usar  $\mathbf{u}(t)$  e  $\mathbf{u}'(t)$  no lugar de  $\mathbf{f}(\mathbf{u})$ , nas expressões (1) e (4), para escrever a igualdade

$$(7) \quad \mathbf{u}((j+1)h) = \mathbf{u}(jh) - \nabla g(\mathbf{u}(jh))h + \tau_j(h),$$

em que  $\tau_j(h)$  é o erro de truncamento ([7, 11, 12]), que representa o erro cometido na aproximação

$$\mathbf{u}((j+1)h) \approx \mathbf{u}(jh) - \nabla g(\mathbf{u}(jh))h,$$



sendo

$$\|\tau_j(h)\| \leq Mh^2.$$

Esse argumento produz o método de Euler para aproximar  $\mathbf{u}(jh)$  como

$$\begin{aligned} \mathbf{u}(h) &\approx \mathbf{u}_0 - h\nabla g(\mathbf{u}_0) = \mathbf{u}_1 \\ \mathbf{u}(2h) &\approx \mathbf{u}_1 - h\nabla g(\mathbf{u}_1) = \mathbf{u}_2 \\ &\vdots \\ \mathbf{u}(ph) &\approx \mathbf{u}_{p-1} - h\nabla g(\mathbf{u}_{p-1}) = \mathbf{u}_p \end{aligned}$$

Cabe aqui observar que, para obtermos uma aproximação de  $\mathbf{u}(T)$ , fazemos  $p$  aproximações e que em cada uma delas pode haver um erro

$$\mathbf{E}_j = \mathbf{u}(t_j) - \mathbf{u}_j, \quad 1 < j \leq p,$$

que certamente interfere na aproximação final.

Vamos estimar até que ponto os erros  $\mathbf{E}_j$ , em cada estimativa  $\mathbf{u}_j$ , podem interferir na aproximação  $\mathbf{u}_p \approx \mathbf{u}(ph)$ .

Da nossa hipótese sobre  $f(\mathbf{u})$  e da Expressão (5), diminuindo um pouco o conjunto  $U$  e deixando-o **convexo**, se necessário, segue que

$$(8) \quad \|\nabla g(\mathbf{u}) - \nabla g(\mathbf{v})\| \leq L\|\mathbf{u} - \mathbf{v}\|, \quad \mathbf{u}, \mathbf{v} \in U,$$

para  $L > 0$ , que depende da segunda derivada de  $g$ .

A Expressão (7), o método de Euler e a desigualdade (8) implicam que

$$\begin{aligned} \|\mathbf{E}_{j+1}\| &= \|\mathbf{u}(t_j) - \nabla g(\mathbf{u}(t_j))h + \tau_j(h) - (\mathbf{u}_j - \nabla g(\mathbf{u}_j)h)\| \\ &\leq \|\mathbf{E}_j\| + hL\|\mathbf{E}_j\| + Mh^2, \end{aligned}$$

quando  $h$  for suficientemente pequeno.

Repetindo o raciocínio, para  $0 < j < p - 1$ , obtemos

$$\begin{aligned} \|\mathbf{E}_{j+1}\| &\leq (1 + hL)\|\mathbf{E}_j\| + Mh^2 \\ &\leq (1 + hL)((1 + hL)\|\mathbf{E}_{j-1}\| + Mh^2) + Mh^2 \\ &\vdots \\ &\leq ((1 + hL)^j + \cdots + (1 + hL) + 1) Mh^2 \\ &= \frac{(1 + hL)^{j+1} - 1}{hL} Mh^2 \\ &< (e^{TL} - 1) \frac{M}{L} h \end{aligned}$$

As duas últimas desigualdades seguem da soma dos termos de uma progressão geométrica e do fato de que  $1 + hL < e^{hL}$ , para  $0 < h$ .

Esse raciocínio para estimar  $|\mathbf{E}_j|$  mostra que este depende linearmente de  $h$  e demonstra o lema a seguir ([10, 11, 12]). Em particular, segue que, se  $T$  for suficientemente grande,  $\mathbf{u}_0$  for suficientemente próximo de  $\mathbf{a}$  e  $h$  for suficientemente pequeno, então (a menos de limitação computacional)  $\mathbf{u}_p$  será uma boa aproximação para  $\mathbf{a}$ .

**Lema 3.1.** *Considere a equação diferencial*

$$(9) \quad \begin{cases} \mathbf{u}'(t) = \mathbf{F}(t, \mathbf{u}(t)) \\ \mathbf{u}(0) = \mathbf{u}_0 \end{cases}, \quad 0 \leq t \leq T,$$

com  $\mathbf{F}$  contínua e tal que

$$\|\mathbf{F}(t, \mathbf{u}) - \mathbf{F}(t, \mathbf{v})\| \leq L\|\mathbf{u} - \mathbf{v}\|.$$

Existe única solução  $\mathbf{u}(t)$  para a Equação (9), para  $0 \leq t \leq T$ .

Ainda, se

$$\left\| \frac{1}{2} \frac{d\mathbf{F}(t, \mathbf{u}(t))}{dt} \right\| \leq M$$

e

$$\mathbf{u}_{j+1} = \mathbf{u}_j + h\mathbf{F}(jh, \mathbf{u}_j), \quad 0 \leq j \leq p-1,$$

então

$$\|\mathbf{u}(jh) - \mathbf{u}_j\| \leq ((1 + hL)^{j+1} - 1) \frac{M}{L} h < (e^{TL} - 1) \frac{M}{L} h.$$

O erro que pode ocorrer em cálculos aproximados, causados pela precisão da calculadora que usarmos, é importante, mas não vamos nos ater a isso aqui. Esse erro tende a diminuir cada vez mais, com o passar do tempo ou com o avanço da tecnologia (se puder veja [12, 13]).

O Lema 3.1 é bastante importante em análise de métodos numéricos. Mas como podemos ver, é resultado geral, e não diz algo sobre a estimativa de  $\|\mathbf{u}_p - \mathbf{a}\|$ , que é o que nos interessa nessa discussão. Será que a forma da Equação (6) pode nos mostrar algo nessa linha?

#### 4. CONTROLANDO A VELOCIDADE DE CONVERGÊNCIA

A forma de determinar uma aproximação para  $\mathbf{a}$ , tal que  $\mathbf{f}(\mathbf{a}) = \mathbf{0}$ , por meio da direção de maior decrescimento de  $g(\mathbf{u})$  e do método de Euler, parece que pode ser bastante lenta, porque o erro que estimamos é proporcional a  $h$ , e não sabemos o quão perto  $\mathbf{u}(T)$  está de  $\mathbf{a}$ .

Sendo assim, procuramos determinar agora se é possível, a partir de certo ponto, estimar a velocidade de convergência de  $\mathbf{u}_j$  para  $\mathbf{a}$ , ou mesmo escolher outra curva  $\mathbf{u}(t)$  em que  $g(\mathbf{u}(t))$  tenha um decrescimento que conseguimos controlar ou estimar.

Para termos uma ideia de possíveis caminhos a escolher observe o argumento que segue. Este raciocínio está relacionado com teoria de controle e de [funcionais de Lyapunov](#) (se puder veja [1, 2, 3, 6]).

A dedução da Expressão (3) e a regra da cadeia nos ajudam a concluir que

$$(10) \quad \begin{aligned} \frac{dg(\mathbf{u}(t))}{dt} &= g'(\mathbf{u}(t))\mathbf{u}'(t) \\ &= \langle \mathbf{u}'(t), \nabla g(\mathbf{u}(t)) \rangle \\ &= \langle \mathbf{u}'(t), [J\mathbf{f}(\mathbf{u}(t))]^* \mathbf{f}(\mathbf{u}(t)) \rangle \\ &= \langle J\mathbf{f}(\mathbf{u}(t))\mathbf{u}'(t), \mathbf{f}(\mathbf{u}(t)) \rangle \end{aligned}$$

Supondo que  $\mathbf{u}(t)$  é solução da Equação (6) e usando a desigualdade de [Cauchy-Schwarz](#) concluímos então que

$$\begin{aligned} \frac{dg(\mathbf{u}(t))}{dt} &= -\langle \nabla g(\mathbf{u}(t)), \nabla g(\mathbf{u}(t)) \rangle \\ &= -\langle [J\mathbf{f}(\mathbf{u}(t))]^* \mathbf{f}(\mathbf{u}(t)), [J\mathbf{f}(\mathbf{u}(t))]^* \mathbf{f}(\mathbf{u}(t)) \rangle \\ &= -\|[J\mathbf{f}(\mathbf{u}(t))]^* \mathbf{f}(\mathbf{u}(t))\|^2 \\ &\leq -\alpha(t) \|\mathbf{f}(\mathbf{u}(t))\|^2 \\ &= -2\alpha(t)g(\mathbf{u}(t)) \end{aligned}$$

sendo  $0 \leq \alpha(t)$  o menor autovalor de matriz  $[J\mathbf{f}(\mathbf{u}(t))]^* J\mathbf{f}(\mathbf{u}(t))$ , ou seja  $\sqrt{\alpha(t)}$  é o menor [valor singular](#) de  $J\mathbf{f}(\mathbf{u}(t))$  ([9]).

Nos argumentos a seguir aproveitamos para apresentar uma equação diferencial que pode ser resolvida. Por isso usamos  $g(t)$  no lugar de  $g(\mathbf{u}(t))$ , e podemos supor que  $\alpha(t)$  é uma função contínua qualquer. Definimos

$$\lambda(t) = \int_0^t \alpha(s) ds.$$

Seja  $r : [0, T] \rightarrow [0, +\infty)$  uma função contínua e tal que

$$g'(t) + 2\alpha(t)g(t) + r(t) = 0.$$

Então, pelas regras do produto e da cadeia,

$$\begin{aligned} \frac{d(e^{2\lambda(t)}g(t))}{dt} &= e^{2\lambda(t)}g'(t) + 2\alpha(t)e^{2\lambda(t)}g(t) \\ &= -e^{2\lambda(t)}r(t) \end{aligned}$$

Integrando ambos os lados da igualdade e usando o [Teorema Fundamental do Cálculo](#) temos que

$$e^{2\lambda(t)}g(t) - g(0) = -\int_0^t e^{2\lambda(s)}r(s) ds.$$

Como  $0 \leq r(t)$ , concluímos que

$$0 \leq \int_0^t e^{2\lambda(s)}r(s) ds$$

e que

$$g(t) \leq g(0)e^{-2\lambda(t)}.$$

Isso demonstra o lema que segue.

**Lema 4.1.** *Seja  $\mathbf{u}(t)$  uma solução da Equação (6) e  $0 \leq \alpha(t)$  o menor autovalor de matriz  $[J\mathbf{f}(\mathbf{u}(t))]^* J\mathbf{f}(\mathbf{u}(t))$ . Então*

$$g(\mathbf{u}(t)) \leq g(\mathbf{u}_0)e^{-2\int_0^t \alpha(s) ds}.$$

No caso em que a matriz  $J\mathbf{f}(\mathbf{a})$  for invertível, existe um número real  $0 < \alpha < \alpha(t)$  e o lema anterior pode ser reformulado como segue.

**Lema 4.2.** *Nas condições Lema 4.1, se  $Jf(\mathbf{a})$  for invertível e  $\mathbf{u}_0$  for suficientemente próximo de  $\mathbf{a}$ . Então existe  $0 < \alpha$  tal que*

$$(11) \quad g(\mathbf{u}(t)) \leq g(\mathbf{u}_0)e^{-2\alpha t}.$$

Isso nos diz que a escolha da direção de maior decréscimo nos dá uma curva  $\mathbf{u}(t)$  que, em tese, nos permite aproximar  $\mathbf{a}$  de tal forma  $g(\mathbf{u}(t))$  tenha decaimento exponencial, desde que  $Jf(\mathbf{a})$  seja invertível. Em tempos de pandemia principalmente, é conhecido que esse decaimento é rápido. Mas determinar o valor de  $\alpha$  pode ser uma tarefa complicada, afinal queremos descobrir o valor de  $\mathbf{a}$ , que supomos existir, mas que não conhecemos.

Por outro lado, como geralmente precisamos aproximar  $\mathbf{u}(t)$  numericamente, uma outra pergunta interessante é a seguinte: Será que, nas condições do Lema 4.1, é possível que o método de Euler para resolver a Equação (6), com  $\mathbf{u}_0$  suficientemente próximo de  $\mathbf{a}$ , seja tal que  $\|\mathbf{u}_j - \mathbf{a}\|$  tenha decaimento exponencial?

A resposta dessa pergunta é sim, e o curioso é que  $Jf(\mathbf{a})$  precisa apenas ser uma matriz não nula. Se for esse o caso, então  $Jf(\mathbf{u})$  será também não nula para todo  $\mathbf{u} \in U$  e suficientemente próximo de  $\mathbf{a}$ , digamos, para todo  $\mathbf{u}$  tal que  $\|\mathbf{u} - \mathbf{a}\| < r$ , para algum  $0 < r$ .

Para nos auxiliar, definimos a função  $\phi$  como

$$\phi(\mathbf{u}) = \mathbf{u} - h [Jf(\mathbf{u})]^* f(\mathbf{u}), \quad \|\mathbf{u} - \mathbf{a}\| < r.$$

Sendo assim, se  $\|\mathbf{u}_0 - \mathbf{a}\| < r$  e  $\mathbf{u}_j$  for obtido pelo método de Euler, lembrando que  $\nabla g(\mathbf{u}) = [Jf(\mathbf{u})]^* f(\mathbf{u})$ , então

$$\mathbf{u}_{j+1} = \phi(\mathbf{u}_j).$$

Como  $f(\mathbf{a}) = \mathbf{0}$ , temos ainda que  $\phi(\mathbf{a}) = \mathbf{a}$ .

Isso nos leva a observar que o método de Euler toma a forma de um [método de ponto fixo](#) (se puder veja [12, 14]).

Usando a [regra da cadeia](#) e a regra do produto obtemos a igualdade

$$\phi'(\mathbf{u}) = I - h [(Jf(\mathbf{u}))']^* f(\mathbf{u}) - h [Jf(\mathbf{u})]^* Jf(\mathbf{u})$$

em que  $I$  denota a matriz identidade.

Seja

$$S^*S = I, \quad D = \begin{bmatrix} s_1^2 & 0 & \cdots & 0 \\ 0 & s_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & s_m^2 \end{bmatrix}, \quad S^*DS = [Jf(\mathbf{a})]^* Jf(\mathbf{a})$$

a decomposição em autovalores e autovetores da matriz  $[Jf(\mathbf{a})]^* Jf(\mathbf{a})$  ([9]). Segue que  $\|S\| = \|S^*\| = 1$  e que

$$\begin{aligned} \|\phi'(\mathbf{a})\| &= \|I - h [Jf(\mathbf{a})]^* Jf(\mathbf{a})\| \\ &= \|S^*(I - hD)S\| \\ &\leq \|I - hD\| \\ &= \max_{1 \leq i \leq m} |1 - hs_i^2| \end{aligned}$$

Supondo que  $s_m \leq \dots \leq s_2 \leq s_1$ , podemos escolher  $0 < h < 2s_1^{-2}$  e obtermos

$$\|\phi'(\mathbf{a})\| = N_0 < N < 1.$$

Diminuindo  $r$ , se necessário, a desigualdade do valor médio implica que

$$\|\phi(\mathbf{u}) - \phi(\mathbf{v})\| \leq N\|\mathbf{u} - \mathbf{v}\|, \quad \|\mathbf{u} - \mathbf{a}\| < r, \quad \|\mathbf{v} - \mathbf{a}\| < r.$$

Em particular, se  $\|\mathbf{u}_0 - \mathbf{a}\| < r$ , então

$$\begin{aligned} \|\mathbf{u}_{j+1} - \mathbf{a}\| &= \|\phi(\mathbf{u}_j) - \phi(\mathbf{a})\| \\ &\leq N\|\mathbf{u}_j - \mathbf{a}\| \\ &= N\|\phi(\mathbf{u}_{j-1}) - \phi(\mathbf{a})\| \\ &\leq N^2\|\mathbf{u}_{j-1} - \mathbf{a}\| \\ &\vdots \\ &\leq N^{j+1}\|\mathbf{u}_0 - \mathbf{a}\| \\ &\leq e^{-(j+1)\beta}\|\mathbf{u}_0 - \mathbf{a}\| \end{aligned}$$

em que  $e^{-\beta} = N$ , o que nos dá o decaimento exponencial para a  $\|\mathbf{u}_j - \mathbf{a}\|$ , e demonstra o teorema que segue. Observe que nos argumentos anteriores poderíamos ter tomado  $h$  variável em cada iteração, mantendo a restrição  $|1 - hs_1^2| < N < 1$ .

**Teorema 4.3.** *Se a matriz  $J\mathbf{f}(\mathbf{a})$  for não nula e  $|1 - hs_1^2| < N < 1$ , então existe  $0 < r$ , tal que*

$$\|\mathbf{u}_0 - \mathbf{a}\| < r \implies \|\mathbf{u}_j - \mathbf{a}\| \leq e^{-j\beta}\|\mathbf{u}_0 - \mathbf{a}\|,$$

em que  $e^{-\beta} = N$ , quando  $\mathbf{u}_j$  for dado pelo método de Euler para resolver a Equação (6).

Porém, como não conhecemos  $\mathbf{a}$ , não temos como conhecer  $s_1$  e escolher  $h$  que nos dê o decaimento estimado para  $\|\mathbf{u}_j - \mathbf{a}\|$ , o que nos leva a seção seguinte.

**4.1. O método de Newton.** Com a resposta que obtemos nos lemas 4.1 e 4.2, nos perguntamos se é possível escolher antes um valor para  $\alpha$  e encontrar um caminho  $\mathbf{u}(t)$  que satisfaça a Desigualdade (11)?

No ponto em que estamos, a resposta vem rápido. Segue da Expressão (10) que

$$\frac{dg(\mathbf{u}(t))}{dt} = \langle J\mathbf{f}(\mathbf{u}(t))\mathbf{u}'(t), \mathbf{f}(\mathbf{u}(t)) \rangle,$$

o que demonstra o resultado que queremos e apresentamos agora.

**Lema 4.4.** *Seja  $\alpha > 0$ . Se  $\mathbf{u}(t)$  for solução da equação*

$$(12) \quad \begin{cases} J\mathbf{f}(\mathbf{u}(t))\mathbf{u}'(t) = -\alpha\mathbf{f}(\mathbf{u}(t)) \\ \mathbf{u}(0) = \mathbf{u}_0 \end{cases},$$

então

$$\frac{dg(\mathbf{u}(t))}{dt} = -2\alpha g(\mathbf{u}(t)).$$

Em particular,

$$g(\mathbf{u}(t)) = g(\mathbf{u}_0)e^{-2\alpha t}.$$

Em tese, obtemos uma nova equação diferencial, dada pela Expressão (12), que nos permite controlar a velocidade em que  $g(\mathbf{u}(t))$  converge para 0. Mas como ocorre com a Equação (6) pode ser necessário o uso de um método numérico para a aproximação da solução dessa nova equação.

Supondo que a matriz  $J\mathbf{f}(\mathbf{a})$  seja invertível e que  $\mathbf{u}_0$  esteja suficientemente próximo de  $\mathbf{a}$ , podemos mais uma vez aplicar o método de Euler para aproximar a (única) solução da Equação (12), que pode agora ser escrita como

$$\begin{cases} \mathbf{u}'(t) = -\alpha [J\mathbf{f}(\mathbf{u}(t))]^{-1} \mathbf{f}(\mathbf{u}(t)) \\ \mathbf{u}(0) = \mathbf{u}_0 \end{cases}.$$

Para evitarmos o cálculo da matriz inversa de  $J\mathbf{f}(\mathbf{u}(t))$ , escrevemos o método de Euler como

$$(13) \quad \begin{cases} J\mathbf{f}(\mathbf{u}_j)\mathbf{w}_j = -h\alpha\mathbf{f}(\mathbf{u}_j) \\ \mathbf{u}_{j+1} = \mathbf{u}_j + \mathbf{w}_j \end{cases},$$

em que  $\mathbf{u}(t_{j+1}) \approx \mathbf{u}_{j+1}$ ,  $t_{j+1} = t_j + h$ , sendo  $0 < h$ .

O restante da seção nos mostra que essa nova forma de determinar  $\mathbf{u}_j$  traz alguma melhora em relação a escolha de  $h$  e quanto a convergência de  $\mathbf{u}_j$ , quando comparamos com o Teorema 4.3. Os argumentos que usaremos são quase idênticos aos que usamos antes, por isso apresentamos aqui apenas o que achamos que difere um pouco.

Se  $J\mathbf{f}(\mathbf{a})$  for invertível, então  $J\mathbf{f}(\mathbf{u})$  será também invertível, quando  $\|\mathbf{u} - \mathbf{a}\| < r$ , para algum  $0 < r$ , e podemos definir a função

$$\psi(\mathbf{u}) = \mathbf{u} - h\alpha [J\mathbf{f}(\mathbf{u})]^{-1} \mathbf{f}(\mathbf{u}), \quad \|\mathbf{u} - \mathbf{a}\| < r.$$

A Expressão (13) implica que

$$\mathbf{u}_{j+1} = \psi(\mathbf{u}_j),$$

caso  $\|\mathbf{u}_0 - \mathbf{a}\| < r$ , e, mais uma vez, vemos um método de ponto fixo relacionado ao método de Euler, agora para outra equação diferencial, pois

$$\mathbf{f}(\mathbf{a}) = \mathbf{0}, \quad \psi(\mathbf{a}) = \mathbf{a}.$$

Ainda,

$$\begin{aligned} \psi'(\mathbf{u}) &= I - h\alpha [(J\mathbf{f}(\mathbf{u}))']^{-1} \mathbf{f}(\mathbf{u}) - h\alpha [J\mathbf{f}(\mathbf{u})]^{-1} J\mathbf{f}(\mathbf{u}) \\ &= (1 - h\alpha)I - h\alpha [(J\mathbf{f}(\mathbf{u}))']^{-1} \mathbf{f}(\mathbf{u}) \end{aligned}$$

e

$$\|\psi'_j(\mathbf{a})\| = |1 - h\alpha|.$$

Escolhendo  $0 < h$ , tal que  $0 \leq |1 - h\alpha| \leq N < 1$ , e repetindo argumentos anteriores, obtemos a demonstração do teorema que segue, o que nos dá novamente um decaimento exponencial para  $\|\mathbf{u}_j - \mathbf{a}\|$ , mas agora podemos escolher  $h$  de forma clara. Em particular, a restrição imposta a  $h$  não impede que o mesmo varie, em cada iteração.

**Teorema 4.5.** *Se a matriz  $J\mathbf{f}(\mathbf{a})$  for invertível e  $|1 - \alpha h| \leq N < 1$ , então existe  $0 < r$ , tal que*

$$\|\mathbf{u}_0 - \mathbf{a}\| < r \implies \|\mathbf{u}_j - \mathbf{a}\| \leq e^{-j\beta} \|\mathbf{u}_0 - \mathbf{a}\|,$$

em que  $e^{-\beta} = N$ , quando  $\mathbf{u}_j$  for dado pela Expressão (13).

Quanto melhor for a escolha de  $\mathbf{u}_0$ , menor pode ser  $N$  no teorema anterior. A convergência tende a ser cada vez mais rápida a cada iteração, porque  $\mathbf{u}_j$  pode ser visto como condição inicial para a próxima iteração.

Em tese, podemos ainda tomar  $N$  tão pequeno quanto quisermos, quando escolhermos  $h\alpha = 1$ , o que produz o método de Newton para resolver a equação  $\mathbf{f}(\mathbf{u}) = \mathbf{0}$ . Sob o ponto de vista prático, de certa forma, isso desmotiva a escolha de  $\alpha \neq 1$  na Equação (12), caso tenha a intenção de usar o método de Euler. Entretanto, tendo em mente o Lema 3.1, a escolha de  $h$  menor pode aumentar a região de escolha de  $\mathbf{u}_0$  em que  $\mathbf{u}_j$  converge para  $\mathbf{a}$ .

O Lema 3.1 sugere que o método de Euler pode ser usado inicialmente para aproximar o valor de  $\mathbf{u}_p \approx \mathbf{u}(T) \approx \mathbf{a}$ , usando a Equação (6), que é computacionalmente mais simples que a Equação (12). Depois disso, se desconfiarmos que  $J\mathbf{f}(\mathbf{a})$  possui inversa, podemos usar a Expressão (13) para calcular  $\mathbf{u}_j$ , para  $j > p$ , para tentar convergência mais rápido para  $\mathbf{a}$ . Mas se desconfiarmos que  $J\mathbf{f}(\mathbf{a})$  não possui inversa, ou que está próxima disso, o Teorema 4.3 pode ser reconfortante.

**4.2. Outras possibilidades.** Uma outra possibilidade para tentar aumentar a velocidade de convergência de  $\mathbf{u}_j$ , dado pelo método de Euler, para a solução da equação  $\mathbf{f}(\mathbf{u}) = \mathbf{0}$ , é não se preocupar muito com a curva  $\mathbf{u}(t)$ , que é solução de alguma equação diferencial, e deixar  $h$  variar em cada iteração, com algum critério para que  $g(\mathbf{u}_{j+1}) < g(\mathbf{u}_j)$  (veja [2, 3]).

Podemos também usar outros métodos numéricos para resolver as equações diferenciais que aparecem neste texto e, com isso, diminuir o erro que cometemos nas aproximações ([11, 12]). Por simplicidade, vamos descrever o caso do método de Euler Modificado, que é dado por

$$\mathbf{u}((j+1)h) \approx \mathbf{u}_{j+1} = \mathbf{u}_j + h\mathbf{u}'\left(\mathbf{u}_j + \frac{h}{2}\mathbf{u}'(\mathbf{u}_j)\right).$$

No caso da Equação (12), por exemplo, esse método pode ser aplicado em duas etapas. Obtemos primeiro uma estimativa  $\tilde{\mathbf{u}}_j$  como

$$\tilde{\mathbf{u}}_{j+1} = \mathbf{u}_j + \frac{h}{2}\mathbf{u}'(\mathbf{u}_j) \implies \begin{cases} J\mathbf{f}(\mathbf{u}_j)\tilde{\mathbf{w}}_j = -\frac{h}{2}\alpha\mathbf{f}(\mathbf{u}_j) \\ \tilde{\mathbf{u}}_{j+1} = \mathbf{u}_j + \tilde{\mathbf{w}}_j \end{cases},$$

sendo  $\tilde{\mathbf{u}}_0 = \mathbf{u}_0$ , e depois obtemos  $\mathbf{u}_j$  como

$$\mathbf{u}_{j+1} = \mathbf{u}_j + h\mathbf{u}'(\tilde{\mathbf{u}}_j) \implies \begin{cases} J\mathbf{f}(\tilde{\mathbf{u}}_j)\mathbf{w}_j = -h\alpha\mathbf{f}(\tilde{\mathbf{u}}_j) \\ \mathbf{u}_{j+1} = \mathbf{u}_j + \mathbf{w}_j \end{cases}.$$

Entretanto, esse método é computacionalmente mais caro que o método de Euler. Mesmo assim, podemos usar argumentos da demonstração do Teorema 4.5 para mostrar que a convergência desse último método é mais rápida que a do método de lá, mas deixamos isso para o leitor.

É possível manipular a Expressão (4) para obter métodos numéricos que têm semelhanças com o que discutimos neste texto ([6]). Se o leitor tiver interesse em outros tipos de métodos, ou em estimativas de custo computacional, ou desejar evitar o cálculo de derivadas em suas implementações computacionais, sugerimos, por exemplo, a leitura do texto [4] e referências.

## 5. COMENTÁRIOS FINAIS

Para finalizar este texto, observamos que os teoremas 4.3 e 4.5 sugerem que a equação  $\mathbf{f}(\mathbf{u}) = \mathbf{0}$  pode ainda ser resolvida por meio da equação diferencial

$$(14) \quad \begin{cases} \mathbf{u}'(t) = -A(\mathbf{f}(\mathbf{u}(t)))\mathbf{f}(\mathbf{u}(t)) \\ \mathbf{u}(0) = \mathbf{u}_0 \end{cases},$$

em que  $A(\mathbf{f}(\mathbf{u}(t)))$  é uma matriz que faz com que  $g(\mathbf{u}(t))$  tenha decaimento controlado, ou estimado, de alguma forma. O uso de método numéricos para aproximar  $\mathbf{u}(t)$ , dado pela Equação (14), nos dá vislumbres para criar vários métodos de ponto fixo para aproximar a solução da equação  $\mathbf{f}(\mathbf{u}) = \mathbf{0}$ .

Como exemplo particular, seja  $\mathbf{f}(\mathbf{u}) = B\mathbf{u} - \mathbf{v}$ , para alguma matriz  $B_{n \times n}$  e algum vetor  $\mathbf{v}$  conhecido. Escolha  $A(\mathbf{f}(\mathbf{u}(t))) = [\text{diag}(B)]^{-1}$ , sendo  $\text{diag}(B)$  a matriz cuja diagonal principal é igual a de  $B$  e com zeros nas demais entradas. Então o método de Euler, para resolver numericamente a versão da Equação (14), nesse caso, inclui o método de ponto fixo de Jacobi-Richardson, para resolver a equação  $Bu = v$ , quando  $h = 1$ . Se, em vez disso, tomarmos  $A(\mathbf{f}(\mathbf{u}(t))) = [\text{diag}(B) + L]^{-1}$ , sendo  $\text{diag}(B) + L$  a matriz cujos termos acima da diagonal principal são nulos e os demais são iguais aos de  $B$ . Então o método de Euler, com  $h = 1$ , é agora o método de ponto fixo de Gauss-Seidel.

## REFERÊNCIAS

- [1] Y. C. Huang, S. H. Sun, J. Q. Han, Neural networks based on Lagrange multiplier for nonlinear programming, *Acta Electron. Sin.* 26, 24–28, 1998.
- [2] L. Zhou, Y. Wu, L. Zhang, G. Zhang, Convergence analysis of a differential equation approach for solving nonlinear programming problems, *Applied Mathematics and Computation*, Volume 184, Issue 2, 789-797, 2007. <https://doi.org/10.1016/j.amc.2006.05.190>
- [3] I. M. Ross, An optimal control theory for nonlinear optimization, *Journal of Computational and Applied Mathematics*, Volume 354, 39-51, 2019. <https://doi.org/10.1016/j.cam.2018.12.044>
- [4] J. R. Sharma, D. Kumar, On a reduced cost derivative-free higher-order numerical algorithm for nonlinear systems. *Comp. Appl. Math.* 39, 202, 2020. <https://doi.org/10.1007/s40314-020-01218-3>
- [5] J. C. Ferreira, M. C. Baquião, A least square point of view to reproducing kernel methods to solve functional equations, *Applied Mathematics and Computation*, 2019. <https://doi.org/10.1016/j.amc.2019.04.008>.
- [6] C. A. Tavares, T. M. R. Santos, N. H.T. Lemes, J. P.C. Santos, J. C. Ferreira, J. P. Braga, Solving ill-posed problems faster using fractional-order Hopfield neural network, *Journal of Computational and Applied Mathematics*, Volume 381, 2021. <https://doi.org/10.1016/j.cam.2020.112984>
- [7] E. L. Lima, *Análise Real*, Vol. 2, Rio de Janeiro, Instituto Nacional de Matemática Pura e Aplicada, 2004.
- [8] G. B. Thomas, *Cálculo*, Vol. 2, Pearson, Ed. 11, 2009.
- [9] H. P. Bueno, *Álgebra Linear: Um segundo curso*, Rio de Janeiro, Sociedade Brasileira de Matemática, 2006.
- [10] D. G. Zill, *Equações diferenciais com aplicações em modelagem*. 3. ed., São Paulo: Cengage Learning, 2016.
- [11] D. J. Lambert, *Numerical Methods for Ordinary Differential Systems: The Initial Value Problem*, John Wiley & Sons, Inc., USA, 1991.
- [12] R. L. Burden, J. D. Faires, *Análise Numérica*, São Paulo, Cengage Learning, 2008.



- [13] D. S. Clark, Short proof of a discrete Gronwall inequality, *Discrete Applied Mathematics*, Volume 16, Issue 3, 279-281, ISSN 0166-218X, 1987. [https://doi.org/10.1016/0166-218X\(87\)90064-3](https://doi.org/10.1016/0166-218X(87)90064-3)
- [14] A. B. M. Brunner, P. R. S. Malta, Teoremas de ponto fixo de Banach e Knaster–Tarski, *revistos*, *Revista Matemática Universitária*, vol. 2, 2020. <https://doi.org/10.21711/26755254/rmu202021>

UNIVERSIDADE FEDERAL DE ALFENAS

*Email address:* jose.ferreira@unifal-mg.edu.br

## MODELOS DO PLANO PROJETIVO NO ESPAÇO TRIDIMENSIONAL

TON MARAR

RESUMO. Os bordos de um 2-disco e de uma faixa de Möbius são curvas fechadas simples. Quando identificamos ponto-a-ponto essas duas curvas, a superfície sem bordo resultante é um modelo do plano projetivo real  $\mathbb{P}^2$ . Vamos apresentar três modos de fazer esta identificação e obter três modelos de  $\mathbb{P}^2$  em  $\mathbb{R}^3$ , topologicamente distintos. Tais modelos aparecem na literatura com a denominação: esfera com cross-cap, superfície romana de Steiner e superfície de Boy, porém sem os detalhes geométricos que apresentamos aqui.

### 1. INTRODUÇÃO

Hilbert & Cohen-Vossen no clássico *Geometry and the Imagination* ([7] p.116), introduzem o plano projetivo como o plano euclidiano aumentado pela adição dos *pontos no infinito*. Cada ponto no infinito é a direção de uma reta. Assim, retas paralelas têm um ponto em comum, no infinito. Uma reta com seu ponto no infinito é chamada *reta projetiva*. *Movendo-se* ao longo de uma reta projetiva de um lado para o outro ou vice-versa somos levados ao mesmo ponto no infinito. Portanto, um círculo é um modelo geométrico da reta projetiva.

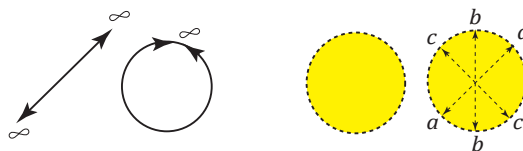


FIGURA 1

Analogamente, um modelo geométrico do plano projetivo é um 2-disco aberto, representando o plano euclidiano, aumentado com todos os seus pontos no infinito, isto é, as direções de todas as retas no plano. As direções das retas que passam pelo centro do disco correspondem a todos os pontos no infinito do plano; de fato,

qualquer reta no plano tem uma paralela passando pelo centro do disco. Cada uma dessas retas passando pelo centro do disco tem um par de pontos antipodais no bordo do disco, e eles devem ser identificados. Portanto, o disco com esses pares de pontos antipodais identificados é um modelo geométrico do plano projetivo (Fig.1).

Alternativamente, o plano projetivo é obtido geometricamente de uma faixa de Möbius e um 2-disco identificados ao longo do bordo.

Para ver isso, considere o disco e os seus pontos de bordo antipodais que serão identificados. Dividimos o disco em três partes, a saber, uma parte central e os setores à esquerda e direita (Fig.2 (a)). É fácil ver que depois da identificação dos antipodais, a parte central dará origem a uma faixa de Möbius. O restante do disco, os dois setores, depois de identificados ao longo do arco tracejado, desde  $a$  até  $c$ , é homeomorfo a um novo disco cujo bordo é exatamente o bordo da faixa de Möbius.

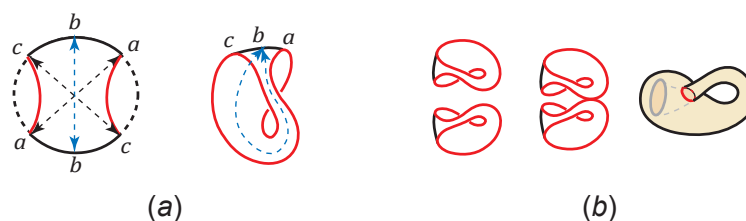


FIGURA 2

De modo mais geral, o espaço projetivo  $n$ -dimensional  $P^n(E)$  associado ao espaço vetorial  $E$  de dimensão  $n + 1$  sobre um corpo  $K$  é o conjunto dos subespaços unidimensionais de  $E$ . Existe uma aplicação natural  $p : E - \{0\} \rightarrow P^n(E)$  que associa cada vetor não nulo  $\mathbf{u}$  ao subespaço gerado por  $\mathbf{u}$ . Uma topologia pode ser dada a  $P^n(E)$  como a topologia quociente da aplicação natural  $p$ . Em particular, o plano projetivo real  $P^2$  é o conjunto de todas as retas que passam pela origem do  $\mathbb{R}^3$ .

Outra superfície obtida de faixas de Möbius é a *garrafa de Klein*. Um modelo da garrafa de Klein é construído identificando, ponto-a-ponto, os bordos de duas faixas (Fig.2(b)).

Garrafa de Klein e plano projetivo são exemplos de superfícies finitas em tamanho e sem bordo. Tais superfícies são chamadas *superfícies fechadas*.

O resultado da classificação topológica das superfícies fechadas foi publicado na Princeton Ph.D. Thesis de Henry Brahana em 1920 ([3]). Existe uma certa disputa de quem foi o primeiro a obter este resultado ([10] p. 53), fato não incomum em matemática.

A lista das superfícies fechadas se divide em duas, a saber, as superfícies orientáveis e as não-orientáveis.



FIGURA 3

Modelos no  $\mathbb{R}^3$  de superfícies orientáveis, como a esfera, o toro, bitoro, tritoro etc, dividem o espaço tridimensional em interior e exterior (Fig.3).

Modelos de superfícies não-orientáveis no  $\mathbb{R}^3$  são mais complicados. Enquanto superfícies orientáveis podem ser modeladas no  $\mathbb{R}^3$  sem auto-interseções, qualquer modelo de superfície não-orientável fechada no  $\mathbb{R}^3$  apresenta auto-interseções. Este é um corolário de um teorema de L. Brouwer, que pode ser lido assim:

*Nenhum subconjunto fechado do espaço euclidiano  $n$ -dimensional é homeomorfo a uma  $(n - 1)$ -variedade não-orientável ([5]).*

No caso do plano projetivo, o processo para a obtenção do modelo geométrico identificando os bordos da faixa de Möbius e do 2-disco não é trivial. De fato, enquanto o bordo da faixa dá duas voltas até se fechar, o bordo do disco fecha depois de apenas uma volta. Assim, para que a identificação dos bordos ocorra no  $\mathbb{R}^3$ , é necessário permitir auto-interseções adequadas na faixa, como veremos na próxima seção.

A curva central da faixa de Möbius, conhecida como *alma* da faixa, é chamada *caminho desorientador*. Se percorrermos a faixa ao longo da sua alma, retornaremos ao ponto inicial de cabeça para baixo (ignorando a gravidade!) (Fig. 4(a)).

Diferente de um cilindro que possui dois lados, isto é, interior e exterior podem ser definidos, a faixa de Möbius possui apenas um lado.

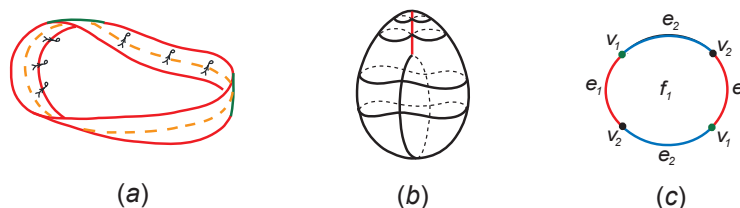


FIGURA 4

A presença de um caminho desorientador caracteriza a não-orientabilidade, isto é, uma superfície é não-orientável se, e somente se, contém uma faixa de Möbius.

Enquanto as superfícies orientáveis fechadas dividem o  $\mathbb{R}^3$  em interior e exterior, as não-orientáveis fechadas, como o plano projetivo e a garrafa de Klein, não permitem a definição de lado de dentro e lado de fora. De fato, qualquer superfície mergulha no  $\mathbb{R}^4$ , isto é, podem ser representadas sem auto-interseção. No caso das superfícies não-orientáveis fechadas, a presença de uma faixa de Möbius, que possui apenas um lado, impossibilita a definição de interior e exterior.

É importante distinguir a superfície abstrata de seu modelo no  $\mathbb{R}^3$ . Auto-interseções presentes nos modelos de superfícies não-orientáveis no  $\mathbb{R}^3$  permitem definir um interior e um exterior, lado de dentro e lado de fora, para os modelos de superfícies não-orientáveis. Além disso, as possíveis auto-interseções de um modelo podem afetar o tipo topológico do subconjunto do  $\mathbb{R}^3$  que o representa. Esta diferença é percebida no cálculo da *característica de Euler*, isto é, a soma alternada do número de 0-células (vértices), 1-células (arestas) e 2-células (faces) de uma decomposição celular do modelo.

Hilbert & Cohn-Vossen ([7], p.315) descrevem o modelo de  $\mathbb{P}^2$  no  $\mathbb{R}^3$  representado aqui na Fig.4(b), e afirmam que este modelo é topologicamente equivalente ao plano projetivo. Contudo, é fácil ver que a característica de Euler do modelo é igual a 2,

enquanto tomando uma decomposição celular do plano projetivo (Fig. 4(c)), com dois vértices  $v_1$  e  $v_2$ , duas arestas  $e_1$  e  $e_2$  e uma face  $f_1$ , obtém-se o valor 1 para a característica de Euler de  $\mathbb{P}^2$ . Portanto, o modelo do plano projetivo no  $\mathbb{R}^3$  (Fig. 4(b)) e a superfície abstrata  $\mathbb{P}^2$  **não** são topologicamente equivalentes.

## 2. TRÊS MODELOS DE $\mathbb{P}^2$ NO $\mathbb{R}^3$

A esfera com cross-cap, a superfície romana de Steiner e a superfície de Boy, representadas respectivamente pelas Fig.5(a), Fig.5(b) e Fig.5(c), são os três modelos do plano projetivo no espaço tridimensional mais conhecidos na literatura.

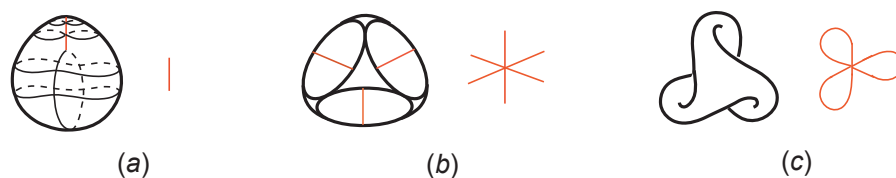


FIGURA 5

São obtidos da identificação do bordo de uma faixa de Möbius com o bordo de um 2-disco, porém com diferentes auto-interseções que permitem tal identificação no  $\mathbb{R}^3$ . A esfera com cross-cap tem um segmento de auto-interseção com dois pontos singulares nas extremidades. A superfície romana de Steiner tem três segmentos não coplanares de auto-interseção e cada um deles possui dois pontos singulares nas extremidades; além disso, os três segmentos se interceptam transversalmente num ponto triplo. A auto-interseção da superfície de Boy é uma rosácea de três pétalas não coplanares, que se interceptam transversalmente num ponto triplo.

Se considerarmos os modelos como imagem de aplicações  $f : \mathbb{P}^2 \rightarrow \mathbb{R}^3$  diferenciáveis, então a esfera com cross-cap e a superfície romana terão pontos onde a matriz jacobiana de  $f$  tem posto igual a 1, pontos singulares. A superfície de Boy é a imagem de uma imersão, portanto posto constante e igual a 2.

H. Whitney ([12]) demonstrou que uma parametrização local dos pontos singulares que ocorrem nesses modelos é dada por  $(x, y) \mapsto (x, y^2, xy)$ . A imagem desta aplicação é conhecida como *guarda-chuva de Whitney*. A razão deste nome é que a superfície imagem, vista como solução da equação algébrica  $Z^2 - YX^2 = 0$ , inclui todo o eixo  $Y$  (Fig.6), como o cabo de um guarda-chuva!

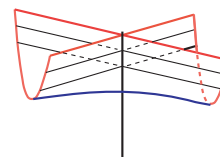


FIGURA 6

**2.1. A esfera com cross-cap.** Os pontos de  $\mathbb{P}^2$  são as retas pela origem de  $\mathbb{R}^3$ , ou equivalentemente, são os pares de pontos antipodais de uma esfera com centro na origem de  $\mathbb{R}^3$ , isto é, cada reta pela origem do  $\mathbb{R}^3$  determina um par de pontos naquela esfera e esses devem ser identificados.

Faremos a identificação dos antipodais de uma esfera por partes. Inicialmente dividimos a esfera em equador e dois hemisférios, um superior e outro inferior. Identificamos todos os pontos do hemisfério superior com os pontos do inferior após uma rotação de  $180^\circ$ . Em seguida identificamos os pontos antipodais do equador (Fig.7).

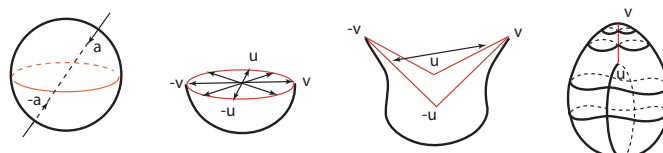


FIGURA 7

Esta construção encontra-se em Hilbert & Cohn-Vossen ([7], p. 314).

Uma construção mais explícita da esfera com cross-cap pode ser feita usando uma adequada projeção no  $\mathbb{R}^3$  de um mergulho da faixa de Möbius no  $\mathbb{R}^4$ .

Considere uma faixa de Möbius mergulhada no  $\mathbb{R}^4$  e sua alma obtida da interseção com um plano transversal à faixa. Tomemos a projeção  $\mathbb{R}^4 \rightarrow \mathbb{R}^3$  que envia todo o plano contendo a alma sobre uma reta no  $\mathbb{R}^3$ . Esta projeção identifica dois-a-dois os pontos da alma, enviando-a sobre um intervalo de reta com dois pontos singulares como extremos. Esses pontos singulares são precisamente a imagem dos pontos da alma onde a direção de projeção é tangente à faixa.

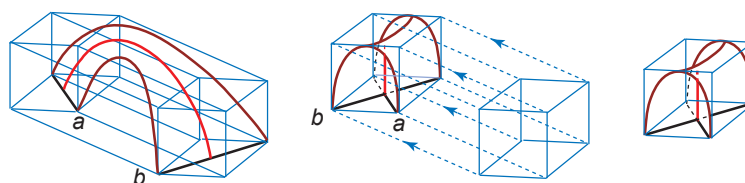


FIGURA 8

A figura 8 descreve a projeção de um hipercubo, contendo meia faixa de Möbius, sobre um cubo que contém um guarda-chuva de Whitney, imagem da projeção da meia faixa.

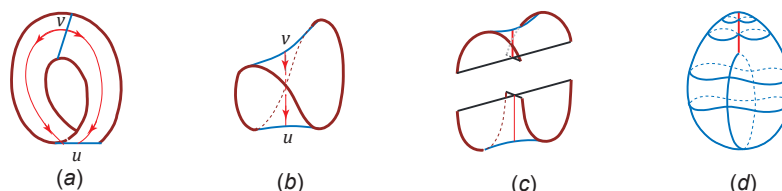


FIGURA 9

A imagem de toda a faixa de Möbius (Fig.9(a)) por esta projeção é chamada conóide de Plücker (Fig.9(b)). Enquanto o bordo da faixa dá duas voltas até se fechar, o bordo do conóide é um bordo circular, fecha depois de apenas uma volta.

Assim, a identificação no  $\mathbb{R}^3$  dos pontos do bordo do conóide com os pontos do bordo do 2-disco é direta. O resultado é o modelo esfera com cross-cap do plano projetivo (Fig.9(d)).

É fácil ver que o conóide é obtido da identificação de dois guarda-chuvas de Whitney (Fig.9(c)), ambos projeções, do  $\mathbb{R}^4$  no  $\mathbb{R}^3$ , de meia faixa de Möbius, sendo um rotacionado de  $\pi/2$  em relação ao outro. Isso simplifica o entendimento dos dois pontos singulares do modelo esfera com cross-cap (Fig.9(d)).

Também é fácil ver que o conóide (Fig.9(b)) é contrátil a um ponto. Assim, o modelo esfera com cross-cap (Fig.9(d)) tem característica de Euler igual a 2.

**2.2. Superfície romana de Steiner.** A superfície romana é um modelo de  $\mathbb{P}^2$  no  $\mathbb{R}^3$  cuja auto-interseção se dá ao longo de três segmentos não coplanares que se cruzam em um ponto triplo. Nas extremidades dos três segmentos, o modelo possui seis pontos singulares do tipo guarda-chuva de Whitney. Supostamente Jakob Steiner (1796-1863) encontrou este modelo quando esteve em férias em Roma em 1844.

Hilbert & Cohn-Vossen ([7] p.302) fazem uma descrição poliedral deste modelo que denominam *heptaedro*. Também se sabe que esta superfície é uma superfície algébrica dada pela equação quártica  $x^2y^2 + x^2z^2 + y^2z^2 - xyz = 0$ .

Assim como fizemos com o cross-cap, usaremos uma projeção adequada da faixa de Möbius do  $\mathbb{R}^4$  no  $\mathbb{R}^3$  para facilitar a visualização da superfície romana.

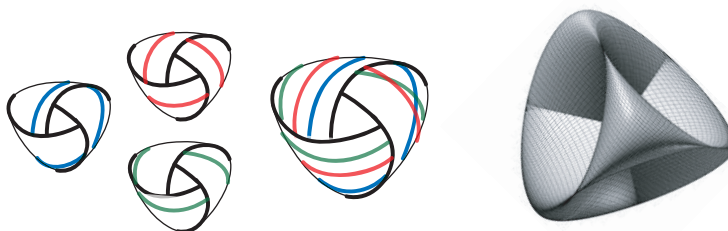


FIGURA 10

Desta vez porém consideramos três curvas em três planos distintos, representando três caminhos desorientadores da faixa, e de modo que esses caminhos se cruzam aos pares (Fig.10). Com esta escolha, após a projeção do  $\mathbb{R}^4$  no  $\mathbb{R}^3$ , cada um dos três planos é projetado em uma reta e essas três retas são não coplanares.

Esta projeção do  $\mathbb{R}^4$  no  $\mathbb{R}^3$  é uma generalização da projeção do  $\mathbb{R}^3$  no  $\mathbb{R}^2$  que leva dois planos em duas retas (Fig.11). O  $\mathbb{R}^4$  é amplo o suficiente para conter os três planos que são projetados em três retas não coplanares no  $\mathbb{R}^3$ .

Além disso, cada um dos três caminhos desorientadores é projetado em um segmento que tem dois pontos singulares nas extremidades e os três possuem um ponto em comum. De fato, uma vizinhança de cada um dos três caminhos

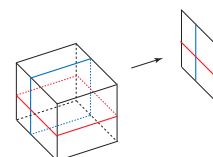


FIGURA 11

desorientadores, isto é, uma faixa de Möbius, é projetada no  $\mathbb{R}^3$  num conóide de Plücker (Fig.9(b)).

Na Fig.12(a) os três caminhos são desenhados num modelo plano da faixa de Möbius. Três pequenos discos são removidos para facilitar o entendimento da projeção. A projeção terá quatro componentes de bordo, três correspondentes aos discos removidos e um que é o bordo da faixa.

A Fig.12(b), composição de três conóides foi desenhada por Scott Carter ([6] p.202), baseado no que ele denomina *Marar's models of singular surfaces*.

Observe que podemos retrair as quatro componentes de bordo até os segmentos de auto-interseção e esses a um ponto. Portanto, a superfície Fig.12(b) é contrátil e o modelo da superfície romana é obtido colando-se os quatro discos ao longo das suas componentes de bordo. Portanto sua característica de Euler é igual a 5.

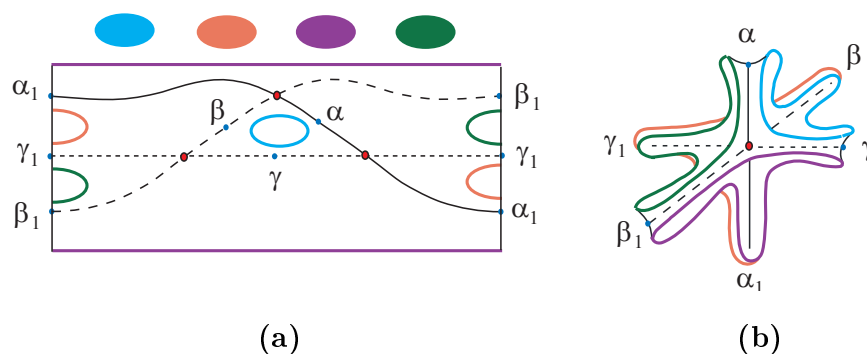


FIGURA 12

**2.3. A superfície de Boy.** Esta superfície é um modelo de  $\mathbb{P}^2$  no  $\mathbb{R}^3$  com um único ponto triplo. Foi descoberto em 1901 por Werner Boy (1879-1914). O modelo não possui pontos singulares, é uma imersão de  $\mathbb{P}^2$  no  $\mathbb{R}^3$ . François Apéry ([1]) encontrou uma parametrização dessa imersão, assim como uma equação algébrica de grau seis definindo a superfície de Boy.

Um corolário de um teorema de Thomas Banchoff ([2]) assegura que toda imersão de  $\mathbb{P}^2$  no  $\mathbb{R}^3$  possui um número ímpar de pontos triplos. De fato, Banchoff mostra que a característica de Euler  $\chi(S)$ , de uma superfície fechada  $S$ , tem a mesma paridade que o número de pontos triplos de qualquer imersão de  $S$  no  $\mathbb{R}^3$ . No caso de  $\mathbb{P}^2$ , a característica de Euler  $\chi(\mathbb{P}^2) = 1$ , e assim, um número ímpar de pontos triplos ocorre em qualquer imersão de  $\mathbb{P}^2$  no  $\mathbb{R}^3$ .

Werner Boy foi aluno de doutorado de David Hilbert. Alguns dizem que Hilbert pediu para Boy demonstrar que qualquer modelo de  $\mathbb{P}^2$  no  $\mathbb{R}^3$  possui no mínimo dois pontos singulares. Contudo, a superfície de Boy é um contra-exemplo a esta suposta conjectura de Hilbert. No prefácio do livro de Apéry ([1]), Egbert Brieskorn escreve: ... *the famous mathematician David Hilbert was mistaken.*



FIGURA 13



A projeção sobre  $\mathbb{R}^3$  do mergulho da faixa de Möbius no  $\mathbb{R}^4$  que mostramos na Fig.8, exibe *inevitáveis* pontos de tangência da direção de projeção com a faixa. De fato, uma conjectura de Whitney de 1940, demonstrada por William Massey ([11]), tem como corolário que toda projeção sobre  $\mathbb{R}^3$ , de uma mergulho de  $\mathbb{P}^2$  no  $\mathbb{R}^4$ , possui no mínimo dois pontos singulares.

Portanto, Boy mergulhou a faixa de Möbius no  $\mathbb{R}^5$  e projetou diretamente no  $\mathbb{R}^3$ , evitando qualquer ponto de tangência com a direção de projeção.

A tese de doutorado de Boy foi publicada em 1903 ([4]) e contém belíssimos desenhos e fotos de modelos de sua imersão de  $\mathbb{P}^2$  no  $\mathbb{R}^3$ . Este é o único registro acadêmico deixado por Boy que, como soldado na Primeira Guerra Mundial, foi abatido logo no seu início, em 6 de setembro de 1914.

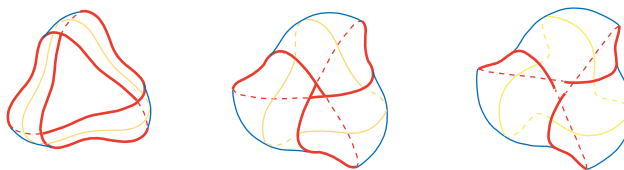


FIGURA 14

Para ver a superfície de Boy como uma faixa de Möbius com um disco colado ao longo de seu bordo, deformamos a faixa (Fig.14) criando um ponto triplo (Fig.15). Este é um processo análogo ao que fizemos na deformação da faixa de Möbius quando criamos o conóide de Plücker (Fig.9(b)).

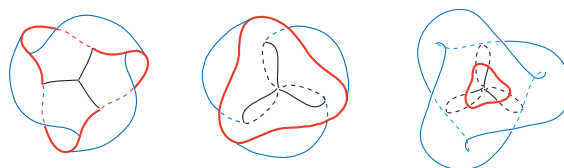


FIGURA 15

A característica de Euler da superfície de Boy é igual a 2, mas isso não é fácil de ver. Neste caso, uma fórmula do tipo Riemann-Hurwitz, como explicitaremos abaixo, pode ajudar.

O conceito de *aplicação semiregular*  $f : N^n \rightarrow \mathbb{R}^{2n-1}$ , onde  $N^n$  é uma variedade  $n$ -dimensional,  $n \geq 2$ , foi introduzido por Whitney no artigo de 1944 publicado no *Annals of Mathematics* ([12]). No caso  $n = 2$ , uma aplicação semiregular de uma superfície com imagem no  $\mathbb{R}^3$  é uma imersão, a menos de um número finito de pontos singulares do tipo guarda-chuva de Whitney.

Se no  $\mathbb{R}^3$  considerarmos o modelo  $M$  de uma superfície  $S$  como a imagem de uma aplicação semiregular então as características de Euler  $\chi(M)$  de  $M$  e  $\chi(S)$  de  $S$  são relacionadas por uma fórmula, obtida por S. Izumiyia e o autor ([8]).

**Teorema.** *Seja  $f : S \rightarrow \mathbb{R}^3$  uma aplicação semiregular de uma superfície fechada  $S$  e cuja imagem  $f(S) = M$  é um modelo de  $S$  no  $\mathbb{R}^3$ . Sejam  $T$  e  $C$  respectivamente*

o número de pontos triplos e guarda-chuvas de Whitney do modelo  $M$ . Então,

$$\chi(M) = \chi(S) + T + C/2.$$

Como a superfície  $S$  é fechada, o número  $C$  é par. De fato, os pontos singulares (guarda-chuvas de Whitney) são os pontos iniciais (ou finais) da curva de auto-interseção do modelo  $M$ . Assim, como a superfície é fechada, a presença de um guarda-chuva (ponto inicial) requer a presença de outro (ponto final).

Esta fórmula facilita o cálculo da diferença das características de Euler da superfície abstrata e de seu modelo no  $\mathbb{R}^3$ . Por exemplo, como  $\chi(\mathbb{P}^2) = 1$ , a fórmula nos dá imediatamente a característica de Euler da esfera com cross-cap, superfície romana e superfície de Boy, a saber, 2, 5 e 2. De fato, a esfera com cross-cap tem  $(C, T) = (2, 0)$ , a superfície romana tem  $(C, T) = (6, 1)$  e a superfície de Boy tem  $(C, T) = (0, 1)$ .

**Agradecimentos.** Agradeço ao referee anônimo e aos meus colegas François Apéry, J. Scott Carter e Thomas Banchoff por sugestões que melhoraram o texto.

#### REFERÊNCIAS

- [1] Apéry, F.; *Models of the real projective plane*, Vieweg 1987.
- [2] Banchoff, T. F.; Triple points and surgery of immersed surfaces; Proc. Amer. Math. Soc. 46 (1974) 407-413.
- [3] Brahana, H. R. ; Systems of circuits on two-dimensional manifolds; Annals of Mathematics 23 (1921) 144-168.
- [4] Boy, W.; Über die Curvatura integra und die Topologie geschlossener Flächen, Mathematische Annalen 57 (1903) 151-184.
- [5] Brouwer, L.E.J.; Sur le théorème de M. Jordan dans l'espace à  $n$  dimensions, Comptes Rendus Paris 153 (1911) 542-544.
- [6] Carter, J. Scott; *How surfaces intersect in space*, Second edition, World Scientific 1995.
- [7] Hilbert, D. & Cohn-Vossen, S. ; *Geometry and the imagination*, Chesea 1952 (a translation of *Anschauliche geometrie*, Springer 1932).
- [8] Izumyia, S. & Marar, W. L.; On topologically stable singular surfaces in a 3-manifold, Journal of Geometry 52 (1995) 108-119.
- [9] Marar, Ton; *Topologia geométrica para inquietos*, Edusp 2019.
- [10] Massey, W. S.; *Algebraic Topology: An Introduction*, Springer 1967.
- [11] Massey, W. S.; Proof of a conjecture of Whitney, Pacific J. Math. 31 (1969) 143-156.
- [12] Whitney, H.; The singularities of a smooth  $n$ -manifold in  $(2n - 1)$ -space, Annals of Mathematics 45 (1944) 247-293.

DEPARTAMENTO DE MATEMÁTICA APLICADA E ESTATÍSTICA  
 INSTITUTO DE CIÊNCIAS MATEMÁTICAS E DE COMPUTAÇÃO  
 UNIVERSIDADE DE SÃO PAULO  
 EMAIL: TON@ICMC.USP.BR

## ALGUNS TEOREMAS DO TIPO VALOR MÉDIO: DE LAGRANGE A MALESEVIC

MARCELO BONGARTI E GERMAN LOZADA-CRUZ

RESUMO. Nosso objetivo neste trabalho é apresentar alguns teoremas do tipo valor médio que não são estudados em disciplinas clássicas de cálculo e análise matemática. Trata-se de teoremas simples e de grande aplicabilidade na análise matemática (por exemplo no estudo de equações funcionais, operadores integrais, etc), matemática computacional, economia e outras áreas.

### 1. INTRODUÇÃO

Dos teoremas clássicos da análise matemática, o *teorema do valor médio* se destaca por sua simplicidade e vasta aplicabilidade. É, sem dúvida, um dos resultados mais conhecidos pela comunidade matemática, e, sem favor algum, um dos tijolos que constituem os alicerces do Cálculo e da Análise Matemática como um todo.

O objetivo deste artigo é apresentar uma gama de outros teoremas do tipo valor médio que somam-se ao clássico e que também são de altíssima aplicabilidade, simplicidade e contribuem para o avanço da Matemática.

Atualmente existem diversas maneiras para abordar o teorema do valor médio, cada qual ligada a alguma meta específica a qual se pretende chegar. Neste caso, como queremos um panorama acerca dos diferentes teoremas do tipo valor médio, utilizaremos a abordagem clássica.

Nossa história começa em 1691 quando Rolle<sup>1</sup> usou técnicas do Cálculo diferencial e integral para provar o seguinte resultado, nosso primeiro teorema do tipo valor médio:

**Teorema 1.1** (Teorema de Rolle). *Seja  $f : [a, b] \rightarrow \mathbb{R}$  uma função contínua em  $[a, b]$  e diferenciável em  $(a, b)$ . Se  $f(a) = f(b)$ , então, existe  $c \in (a, b)$  tal que  $f'(c) = 0$ , isto é, a reta tangente ao gráfico de  $f$  no ponto  $(c, f(c))$  é horizontal.*

---

Data de aceitação: Janeiro de 2021.

*Palavras chave.* Teorema de Lagrange, Teorema de Flett, Condição de Tong, Condição de Malesevic.

<sup>1</sup>Michel Rolle (1652-1719), matemático francês.

Ainda que Rolle tenha sido merecidamente homenageado pela formalização do resultado, é interessante comentar que Bhaskara II<sup>2</sup>, demonstrou um caso particular do teorema de Rolle muito tempo antes, embora sem pouca ou nenhuma formalidade.

O teorema de Rolle ficou mais conhecido depois que Drobisch<sup>3</sup> usou o termo pela primeira vez em 1834, seguido por Bellavitis<sup>4</sup> em 1846. Maiores detalhes podem ser encontrados em [2].

O teorema do tipo valor médio mais famoso da história é o *teorema do valor médio de Lagrange*:

**Teorema 1.2** (Teorema do valor médio de Lagrange). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua em  $[a, b]$  e derivável em  $(a, b)$ , então existe  $c \in (a, b)$  tal que*

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Este resultado foi inicialmente descoberto por Lagrange<sup>5</sup>, que o demonstrou sem, inicialmente, fazer menção ao teorema de Rolle. Entretanto, a dedução mais conhecida tem como ideia principal a aplicação do teorema de Rolle à função auxiliar

$$\varphi(x) = f(x) - \left[ \frac{f(b) - f(a)}{b - a}(x - a) + f(a) \right],$$

e esta foi feita por Bonnet<sup>6</sup>. Publicamente, o teorema do valor médio de Lagrange foi citado pela primeira vez em um trabalho do renomado físico Ampère<sup>7</sup>. Para maiores detalhes, a referência [16] pode ser consultada.

Geometricamente, o teorema do valor médio de Lagrange diz que existe um ponto  $c$  dentro do intervalo  $(a, b)$ , tal que a reta tangente ao gráfico de  $f$  no ponto  $(c, f(c))$  é paralela à reta secante que passa pelos pontos  $(a, f(a))$  e  $(b, f(b))$ .

Fisicamente, o teorema do valor médio de Lagrange garante que se uma partícula possui uma trajetória suave  $(t, f(t))$  no intervalo de tempo  $[a, b]$ , então existirá um instante  $t_c \in (a, b)$  tal que a velocidade instantânea da partícula (em  $t = t_c$ ) coincide com a velocidade média de todo o percurso.

Motivado por esta aplicação, Cauchy<sup>8</sup> se perguntou o que poderia ser dito a respeito de uma partícula de trajetória suave  $(f(t), g(t))$  no intervalo de tempo  $[a, b]$ . Então, Cauchy aplicou o teorema de Rolle à função

$$\varphi(x) = [g(b) - g(a)]f(x) - [f(b) - f(a)]g(x)$$

e o resultado foi o que conhecemos hoje por *Teorema do valor médio de Cauchy*.

**Teorema 1.3** (Teorema do valor médio de Cauchy). *Se  $f, g : [a, b] \rightarrow \mathbb{R}$  são funções contínuas em  $[a, b]$  e deriváveis em  $(a, b)$ , então existe  $c \in (a, b)$  tal que*

$$f'(c)[g(b) - g(a)] = g'(c)[f(b) - f(a)].$$

<sup>2</sup>Bhaskara Akaria (1114-1185), matemático indiano.

<sup>3</sup>Moritz Wilhelm Drobisch (1802-1896), matemático alemão.

<sup>4</sup>Giusto Bellavitis (1803-1880), matemático italiano

<sup>5</sup>Joseph Louis Lagrange (1736-1813), matemático italiano.

<sup>6</sup>Pierre Ossian Bonnet (1819-1892), matemático francês.

<sup>7</sup>André-Marie Ampère (1775-1836), físico francês.

<sup>8</sup>Augustine-Louis Cauchy (1789-1857), matemático francês.

Geometricamente, o teorema do valor médio de Cauchy garante que dada uma trajetória suave  $(f(t), g(t))$  com  $t \in [a, b]$ , existirá  $c \in (a, b)$  de tal forma que a reta tangente à trajetória no ponto  $(f(c), g(c))$  é paralela à reta que passa pelos pontos  $(f(a), g(a))$  e  $(f(b), g(b))$ .

As demonstrações detalhadas dos teoremas do tipo valor médio discutidos nesta seção podem ser encontradas em [1, Teorema 6.2.4 e Teorema 6.3.2] ou em [16, Teorema 2.2 e Teorema 2.17]. O leitor interessado em variações do teorema de Lagrange pode consultar [6]. Para variações e aplicações do teorema de Cauchy, recomendamos [7] e [8]

## 2. TEOREMA DE FLETT E SUAS VARIAÇÕES

O teorema a seguir é uma versão do clássico teorema do valor médio para integrais e a partir de observações acerca dele que nasce a motivação para o teorema do tipo valor médio que abordaremos a seguir.

**Teorema 2.1.** *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua, então existe  $\eta \in [a, b]$  tal que*

$$\int_a^b f(x)dx = f(\eta)(b - a).$$

A demonstração do Teorema 2.1 pode ser encontrada em [16, Teorema 7.1].

Considere, agora, uma função  $g : [a, b] \rightarrow \mathbb{R}$  como no Teorema 2.1, então existe  $\xi \in (a, b)$  tal que

$$g(\xi) = \frac{1}{b - a} \int_a^b g(t)dt.$$

Além disso, se considerarmos uma tal função  $g$ , contínua, e tal que

$$g(a) = 0, \quad \int_a^b g(t)dt = 0,$$

e definirmos a função

$$\varphi(x) = \begin{cases} \frac{1}{x - a} \int_a^x g(t)dt, & x \in (a, b] \\ 0, & x = a. \end{cases}$$

A função  $\varphi$  é contínua em  $[a, b]$ , derivável em  $(a, b)$  e para  $x \in (a, b)$ ,

$$\varphi'(x) = -\frac{1}{(x - a)^2} \int_a^x g(t)dt + \frac{g(x)}{x - a}.$$

Observemos que,  $\varphi(a) = 0 = \varphi(b)$ , logo segue do teorema de Rolle que existe  $\xi \in (a, b)$  tal que  $\varphi'(\xi) = 0$ , i.e., existe  $\xi \in (a, b)$  tal que

$$(1) \quad g(\xi) = \frac{1}{\xi - a} \int_a^{\xi} g(t) dt.$$

Notemos que (1) é exatamente o teorema do valor médio para integrais com a substituição de  $b$  por  $\xi$ . Olhando rapidamente para o teorema fundamental do cálculo, podemos escrever a equação (1) da seguinte forma:

$$(2) \quad G'(\xi) = \frac{G(\xi) - G(a)}{\xi - a}$$

onde  $G$  é uma primitiva de  $g$ , ou seja,  $G' = g$ .

Neste sentido, dada uma função  $g$ , é natural nos perguntarmos se podemos trocar a condição de que  $\int_a^b g(t) dt = 0$  simplesmente por  $g(b) = 0$ . A seguir, como uma consequência do teorema do valor intermediário, vemos que é possível. Estas observações foram feitas por Flett<sup>9</sup> e o resultado (de 1958) leva o seu nome como homenagem.

O teorema de Flett (veja [4]) é uma variação do teorema de Rolle onde a condição  $f(a) = f(b)$  foi substituída por  $f'(a) = f'(b)$ . Por este motivo, dizemos que o teorema de Flett é um teorema do tipo de Lagrange com uma condição do tipo Rolle, ou simplesmente, teorema do tipo valor médio com uma condição do tipo Rolle.

**Teorema 2.2** (Teorema do valor médio de Flett [4]). *Seja  $f : [a, b] \rightarrow \mathbb{R}$  uma função diferenciável em  $[a, b]$  com  $f'(a) = f'(b)$ . Então, existe  $\xi \in (a, b)$  tal que*

$$(3) \quad f'(\xi) = \frac{f(\xi) - f(a)}{\xi - a}.$$

**Demonstração.** Sem perda de generalidade podemos supor  $f'(a) = f'(b) = 0$ , pois caso contrário fazemos  $\psi(x) = f(x) - xf'(a)$  e daí teremos  $\psi'(a) = \psi'(b) = 0$ . Definamos a função  $\varphi : [a, b] \rightarrow \mathbb{R}$  dada por

$$\varphi(x) = \begin{cases} \frac{f(x) - f(a)}{x - a}, & x \in (a, b] \\ f'(a), & x = a. \end{cases}$$

A função  $\varphi$  é contínua em  $[a, b]$ , derivável em  $(a, b)$  e para  $x \in (a, b)$ ,

$$\varphi'(x) = \frac{f'(x)}{x - a} - \frac{\varphi(x)}{x - a}.$$

Observemos que,  $\varphi(a) = 0$ . Se  $\varphi(b) = 0$ , pelo teorema de Rolle existe  $\xi \in (a, b)$  tal que  $\varphi'(\xi) = 0$  e o teorema está provado.

<sup>9</sup>Thomas Muirhead Flett (1923-1976), matemático britânico.

Suponhamos  $\varphi(b) \neq 0$ . Se  $\varphi(b) > 0$ , segue que

$$\varphi'(b) = \frac{f'(b) - \varphi(b)}{b - a} = -\frac{\varphi(b)}{b - a} < 0.$$

Logo, para  $\epsilon > 0$  suficientemente pequeno existe  $x_1 \in (b - \epsilon, b)$  tal que  $\varphi(b) < \varphi(x_1)$ . Como  $\varphi$  é contínua em  $(a, x_1)$  e  $0 = \varphi(a) < \varphi(b) < \varphi(x_1)$ , segue do teorema do valor intermediário, que existe  $\eta \in (a, x_1)$  tal que  $\varphi(\eta) = \varphi(b)$ . E então, do teorema de Rolle aplicado ao intervalo  $[\eta, b]$ , existe  $\xi \in (\eta, b) \subset (a, b)$  tal que  $\varphi'(\xi) = 0$ , isto é,  $f'(\xi) = \frac{f(\xi) - f(a)}{\xi - a}$ .

O caso  $\varphi(b) < 0$  é análogo. □

Geometricamente, o teorema de Flett diz que se uma curva  $(t, f(t))$  é suave no intervalo  $[a, b]$  e as retas tangentes nos extremos  $(a, f(a))$  e  $(b, f(b))$  são paralelas, então, existe um ponto  $\xi \in (a, b)$  de modo que a reta tangente ao gráfico de  $f$  que passa por  $(\xi, f(\xi))$  também passa por  $(a, f(a))$ , como podemos ver na Figura 1.

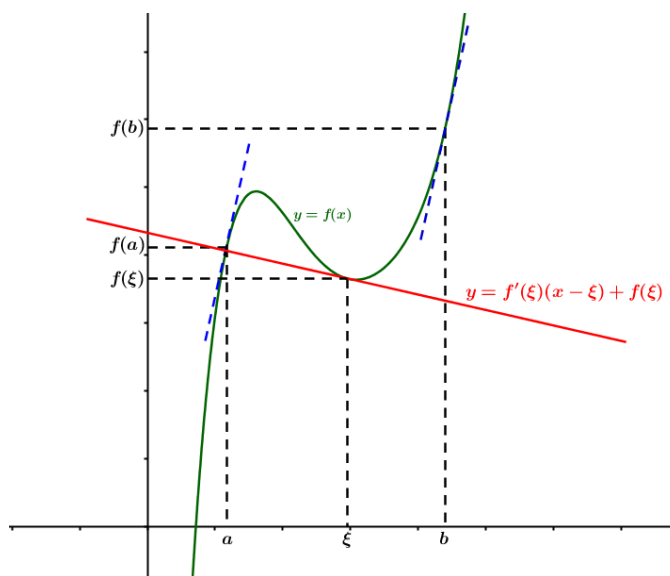


FIGURA 1. Interpretação geométrica do teorema de Flett

Por outro lado, do ponto de vista cinemático, Flett concluiu que, se as velocidades inicial e final de uma partícula com trajetória  $(t, f(t))$  suave no intervalo de tempo  $[a, b]$  forem iguais, então, existe um momento  $t_j \in (a, b)$  tal que a velocidade instantânea da partícula neste instante, é exatamente a velocidade média do percurso até o instante  $t_j$ .

Por comodidade, dada uma função  $f : [a, b] \rightarrow \mathbb{R}$ , chamaremos o ponto  $\xi \in (a, b)$  tal que satisfaz a conclusão do teorema de Flett simplesmente de *ponto de Flett*.

Um exemplo básico para exemplificar o uso do teorema de Flett pode ser dado quando consideramos a função  $f : [-2, 2] \rightarrow \mathbb{R}$  dada por  $f(x) = x^3 + 2x - 1$ . Como

$f$  é um polinômio segue que  $f$  é derivável em  $[-2, 2]$  e usando (3) vemos facilmente que  $\xi = 1 \in (-2, 2)$  é um ponto de Flett de  $f$ .

A seguir trataremos brevemente dos resultados apresentados por R. Meyers em 1977. Estes apresentam variações do teorema de Flett. As interpretações geométricas e físicas são análogas às interpretações do teorema de Flett, portanto, não serão verbalmente discutidas nesta seção.

**Teorema 2.3** ([13, Teorema 1']). *Seja  $f : [a, b] \rightarrow \mathbb{R}$  uma função diferenciável em  $[a, b]$  com  $f'(a) = f'(b)$ . Então, existe  $\xi \in (a, b)$  tal que*

$$f'(\xi) = \frac{f(b) - f(\xi)}{b - \xi}.$$

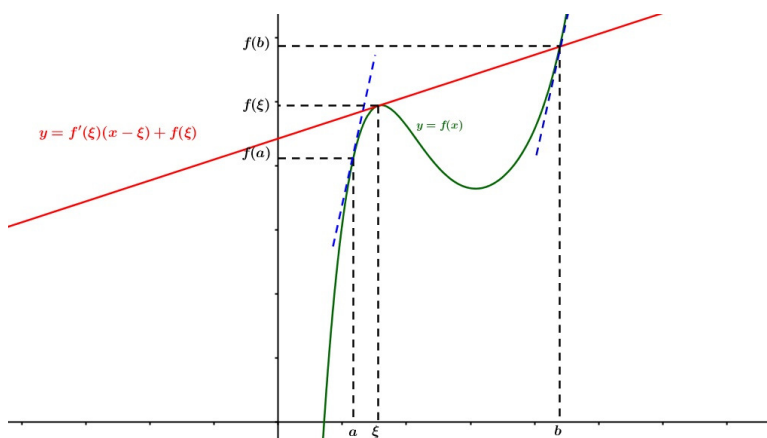


FIGURA 2. Interpretação geométrica Teorema 2.3

**Teorema 2.4** ([13, Teorema 2]). *Seja  $f : [a, b] \rightarrow \mathbb{R}$  uma função diferenciável em  $[a, b]$  com  $f'(a) = f'(b)$ . Então, existe  $\xi \in (a, b)$  tal que*

$$f'(\xi) = \frac{f(b) - f(\xi)}{\xi - a}.$$

**Teorema 2.5** ([13, Teorema 2']). *Seja  $f : [a, b] \rightarrow \mathbb{R}$  uma função diferenciável em  $[a, b]$  com  $f'(a) = f'(b)$ . Então, existe  $\xi \in (a, b)$  tal que*

$$f'(\xi) = \frac{f(\xi) - f(a)}{b - \xi}.$$

**Teorema 2.6** ([13, Teorema 3]). *Se  $f$  é diferenciável e  $f'$  é contínua em  $[a, b]$  e  $[f(b) - f(a)][f(b) - f(a) - (b - a)f'(b)] < 0$ . Então, existe  $\xi \in (a, b)$  tal que*

$$f'(\xi) = \frac{f(b) - f(a)}{\xi - a}.$$



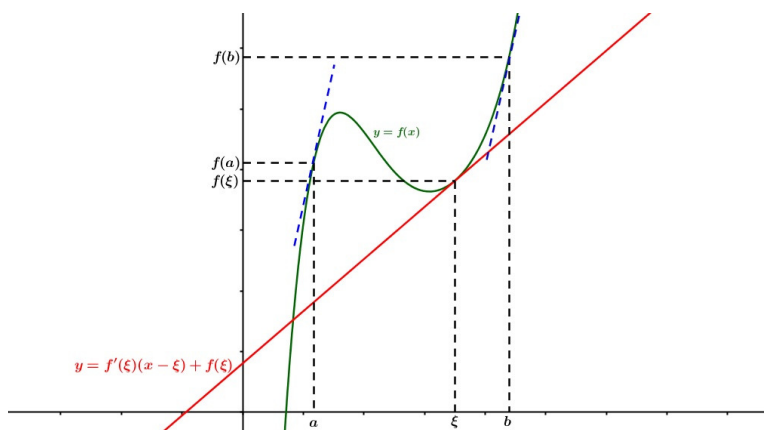


FIGURA 3. Interpretação geométrica Teorema 2.4

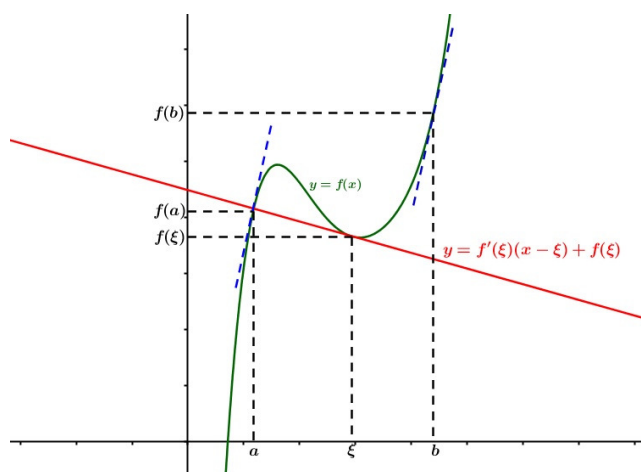


FIGURA 4. Interpretação geométrica Teorema 2.5

**Teorema 2.7** ([13, Teorema 3']). *Se  $f$  é diferenciável e  $f'$  é contínua em  $[a, b]$  e  $[f(b) - f(a)][f(b) - f(a) - (b - a)f'(a)] < 0$ . Então, existe  $\xi \in (a, b)$  tal que*

$$f'(\xi) = \frac{f(b) - f(a)}{b - \xi}.$$

**Teorema 2.8** ([13, Teorema 4]). *Se  $f$  é diferenciável e  $f'$  é contínua em  $[a, b]$  e  $f'(a)[f(b) - f(a) - (b - a)f'(b)] > 0$ . Então, existe  $\xi \in (a, b)$  tal que*

$$f'(\xi) = \frac{f(\xi) - f(a)}{b - a}.$$

**Teorema 2.9** ([13, Teorema 4']). *Se  $f$  é diferenciável e  $f'$  é contínua em  $[a, b]$  e  $f'(b)[f(b) - f(a) - (b - a)f'(a)] > 0$ . Então, existe  $\xi \in (a, b)$  tal que*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

### 3. EXISTÊNCIA DE PONTOS DE FLETT

O assunto discutido nessa seção tem motivação no exemplo a seguir.

Seja  $[a, b]$  um intervalo fechado que contém 0 no seu interior e considere a função  $f : [a, b] \rightarrow \mathbb{R}$  dada por  $f(x) = |x|$ , que não é diferenciável em  $x = 0$ , entretanto, supondo que  $a < x < 0$  temos,

$$\frac{f(x) - f(a)}{x - a} = \frac{|x| - |a|}{x - a} = \frac{-x + a}{x - a} = -1 = f'(x), \quad \forall x \in (a, 0).$$

Portanto, existem infinitos pontos de Flett em  $(a, 0) \subset (a, b)$ .

Este exemplo mostra que o conjunto das funções que satisfazem as hipóteses do teorema de Flett está estritamente contido no conjunto das funções que têm um ponto de Flett. Portanto, é natural perguntarmos que outras condições suficientes existem que garantam a existência de pontos de Flett.

Originalmente, em 1958, T.M. Flett demonstrou que pontos de Flett existem sob as hipóteses de que  $f$  seja diferenciável no intervalo fechado  $[a, b]$  e que  $f'(a) = f'(b)$ . Mas esta não é a única.

Os primeiros estudos sobre os resultados de T.M Flett e suas generalizações foram feitos em 1966 pelo matemático Donald. H. Trahan em 1966 (ver [18]). Ele deu uma nova condição para a existência de um ponto de Flett através de algumas desigualdades, usando uma comparação entre a inclinação da reta secante ao gráfico da função  $f : [a, b] \rightarrow \mathbb{R}$  passando pelos extremos  $(a, f(a))$  e  $(b, f(b))$  e a inclinação das retas tangentes ao gráfico passando pelos mesmos.

Os seguintes resultados são necessários para a compreensão da condição de Trahan.

**Lema 3.1** ([18, Lema 1]). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua, diferenciável em  $(a, b)$  e  $f'(b)[f(b) - f(a)] \leq 0$ , então existe  $c \in (a, b)$  tal que  $f'(c) = 0$ .*

**Lema 3.2** ([18, Lema 2]). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é uma função contínua, diferenciável em  $(a, b)$  e  $f'(b)[f(b) - f(a)] < 0$ , existe  $c \in (a, b)$  tal que  $f'(c) = 0$ .*

Observemos que os Lemas 3.1 e 3.2 são generalizações do Teorema de Rolle.

**Teorema 3.1 (Condição de Trahan [18]).** *Seja  $f : [a, b] \rightarrow \mathbb{R}$  uma função diferenciável e tal que*

$$(4) \quad \left(f'(b) - \frac{f(b)-f(a)}{b-a}\right) \left(f'(a) - \frac{f(b)-f(a)}{b-a}\right) \geq 0.$$

*Então, existe um ponto de Flett em  $(a, b)$ .*

**Demonstração.** Considere a função  $\varphi : [a, b] \rightarrow \mathbb{R}$  definida por

$$\varphi(x) = \begin{cases} \frac{f(x) - f(a)}{x - a}, & x \in (a, b] \\ f'(a), & x = a \end{cases}$$

Observemos que  $\varphi$  é contínua em  $[a, b]$ , é diferenciável em  $(a, b)$  e que  $\varphi'(b)[\varphi(b) - \varphi(a)] \leq 0$ .

Logo, pelo Lema 3.1, existe  $\xi \in (a, b)$  tal que  $\varphi'(\xi) = 0$ , o que significa que

$$f'(\xi) = \frac{f(\xi) - f(a)}{\xi - a},$$

ou seja,  $\xi$  é um ponto de Flett em  $(a, b)$ . □

No exemplo a seguir, temos uma função que não satisfaz a condição de Flett, mas satisfaz a de Trahan e possui, portanto, um ponto de Flett.

**Exemplo 3.1.** Considere a função  $f : [-\frac{1}{2}, 1] \rightarrow \mathbb{R}$  dada por  $f(x) = x^3$ .

Notemos que  $f$  é diferenciável e que  $f'(-\frac{1}{2}) \neq f'(1)$ , logo,  $f$  não satisfaz a condição de Flett.

No entanto,  $f$  satisfaz a condição de Trahan e, portanto, possui um ponto de Flett, a saber  $\xi = \frac{1}{4} \in (-\frac{1}{2}, 1]$ .

Outra condição suficiente para a existência de um ponto de Flett foi provada por J. Tong em [17]. Um ponto interessante desta condição é que Tong só exige a diferenciabilidade de  $f$  em  $(a, b)$ , mas usa os conceitos de média aritmética de  $f$ ,

$$\mathcal{M}(f) := \frac{f(a)+f(b)}{2} \text{ e média de } f, \mathcal{I}(f) := \frac{1}{b-a} \int_a^b f(t)dt.$$

**Teorema 3.2 (Condição de Tong [17, Teorema 2]).** Seja  $f : [a, b] \rightarrow \mathbb{R}$  uma função contínua em  $[a, b]$  e diferenciável em  $(a, b)$ . Se  $\mathcal{M}(f) = \mathcal{I}(f)$  então,  $f$  admite um ponto de Flett em  $(a, b)$ .

**Demonstração.** Basta observar que a função  $h$  dada por

$$\begin{aligned} h : [a, b] &\rightarrow [a, b] \\ x &\mapsto h(x) = \frac{f(x)+f(a)}{2}(x-a) - \int_a^x f(t)dt. \end{aligned}$$

é contínua em  $[a, b]$  e diferenciável em  $(a, b)$  com derivada

$$h'(x) = \frac{1}{2}f'(x)(x-a) + \frac{1}{2}(f(x) + f(a)) - f(x).$$

Como  $h(a) = 0$  e  $\mathcal{M}(f) = \mathcal{I}(f)$ , segue que  $h(b) = 0$ . Agora, usando o teorema de Rolle obtemos a conclusão do eorema. □

O exemplo a seguir traz uma função que não satisfaz a condição de Flett, nem de Trahan, mas satisfaz a de Tong.

**Exemplo 3.2.** Considere a função  $f(x) = \arcsin x$  sobre o intervalo  $[-1, 1]$ .

Notemos que  $f$  não satisfaz a condição de Flett, nem a de Trahan, pois não é diferenciável nos extremos. No entanto, um cálculo simples usando integração por partes garante que  $\mathcal{M}(f) = \mathcal{I}(f)$ , o que prova que a função  $\arcsin x$  satisfaz a condição de Tong no intervalo  $[-1, 1]$  e que, portanto, possui um ponto de Flett.

A terceira condição se deve ao matemático B. Malesevic e é feita em termos de uma função infinitesimal.

Para tanto, seja  $f : [a, b] \in \mathbb{R}$  uma função diferenciável em  $[a, b]$  e diferenciável um número arbitrário de vezes numa vizinhança à direita do ponto  $x = a$ .

Considere a expansão de Taylor de ordem um, com resto, dado por

$$f(x) = f(a) + f'(a)(x - a) + \varphi(x)(x - a),$$

onde  $\lim_{x \rightarrow a^+} \varphi(x) = 0$ . Então, definimos a função  $\varphi_1 : [a, b] \in \mathbb{R}$  por

$$(5) \quad \varphi_1(x) = \begin{cases} \frac{f(x) - f(a)}{x - a} - f'(a), & x \in (a, b] \\ 0, & x = a. \end{cases}$$

A partir desta função  $\varphi_1$  Malesevic demonstrou o seguinte resultado:

**Teorema 3.3** (Condição de Malesevic [11]). *Seja  $f : [a, b] \rightarrow \mathbb{R}$  é uma função diferenciável e  $\varphi_1$  como em (5). Se uma das seguintes condições*

$$T_1 : \varphi_1'(b) \varphi_1(b) < 0 \text{ e}$$

$$M_1 : \varphi_1'(a) \varphi_1(b) < 0$$

é satisfeita, então  $f$  possui um ponto de Flett.

**Demonstração.** Se a condição  $T_1$  é satisfeita então  $\varphi_1'(b)[\varphi_1(b) - \varphi_1(a)] < 0$ . Logo, do Lema 3.2 existe  $\xi_1 \in (a, b)$  tal que  $\varphi_1'(\xi_1) = 0$ , i.e.,

$$\frac{1}{\xi_1 - a} (f'(\xi_1) - \frac{f(\xi_1) - f(a)}{\xi_1 - a}) = 0 \Leftrightarrow f'(\xi_1) = \frac{f(\xi_1) - f(a)}{\xi_1 - a},$$

Agora se a condição  $M_1$  é satisfeita então  $\varphi_1'(a)[\varphi_1(b) - \varphi_1(a)] < 0$ . Logo, do Corolário 3 do Teorema 2 em [12] existe  $\xi_2 \in (a, b)$  tal que  $\varphi_1'(\xi_2) = 0$ , i.e.,

$$\frac{1}{\xi_2 - a} (f'(\xi_2) - \frac{f(\xi_2) - f(a)}{\xi_2 - a}) = 0 \Leftrightarrow f'(\xi_2) = \frac{f(\xi_2) - f(a)}{\xi_2 - a}.$$

□

**Observação 3.1.** *Se ambas as condições  $T_1$  e  $M_1$  do Teorema 3.3 são satisfeitas, então existem dois pontos (distintos) de Flett. (veja [12]).*

A relação entre funções que satisfazem as condições de Flett, Trahan, Tong e Malesevic é representada na Figura 5.

Observemos que  $\Delta_{12} \neq \emptyset$ , uma vez que  $f(x) = \text{sgn}(x)$  é uma função que não satisfaz nenhuma das condições, pois não é diferenciável em  $(a, b)$  para qualquer intervalo  $[a, b]$  da reta, que contenha o zero. Entretanto, possui infinitos pontos de Flett.

Analogamente,

(i)  $f(x) = x^3$ ,  $x \in [-1, 1]$  está em  $\Delta_1$ .

(ii)  $f(x) = \sin(x)$ ,  $x \in [-\frac{\pi}{2}, \frac{5\pi}{2}]$  está em  $\Delta_2$ .

(iii)  $f(x) = x^3$ ,  $x \in [-\frac{2}{3}, 1]$  está em  $\Delta_3$ .

(iv)  $f(x) = \arcsin(x)$ ,  $x \in [-1, 1]$  está em  $\Delta_{12}$ .

Prova-se que todos os conjuntos  $\Delta_i$ ,  $i = 1, 2, \dots, 12$  são não vazios, mas a confecção de exemplos para  $i = 4, 5, 6, 7, 8, 10$  e  $11$  fogem ao objetivo deste trabalho e por isso não serão discutidos aqui.

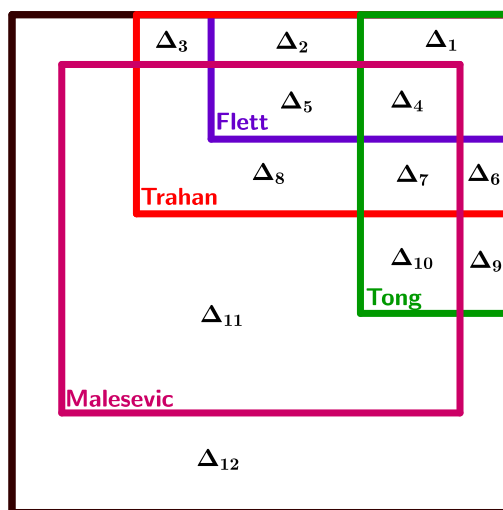


FIGURA 5. Relação entre as condições de Flett, Tong, Trahan e Malesevic

Maiores detalhes das demonstrações e exemplos vistos nesta seção, sugerimos a referência [5].

#### 4. GENERALIZAÇÕES E APLICAÇÕES

**4.1. Algumas Generalizações.** A seguir trataremos de algumas generalizações e consequências do teorema de Flett. O teorema abaixo foi demonstrado em 1998. Este resultado é uma generalização do teorema de Flett que não exige a condição tipo Rolle, nesse caso o resultado é mais geral. Notemos que nos dois teoremas a seguir, o caso em que  $f'(a) = f'(b)$  é exatamente o teorema de Flett.

**Teorema 4.1 (Riedel-Sahoo [16, Teorema 5.2]).** *Se  $f : [a, b] \rightarrow \mathbb{R}$  é diferenciável em  $[a, b]$ , então, existe  $\xi \in (a, b)$  tal que*

$$(6) \quad f(\xi) - f(a) = (\xi - a)f'(\xi) - \frac{1}{2} \frac{f'(b) - f'(a)}{b - a} (\xi - a)^2.$$

**Demonstração.** Definamos a função  $\varphi : [a, b] \rightarrow \mathbb{R}$  por

$$\varphi(x) = f(x) - \frac{1}{2} \frac{f'(b) - f'(a)}{b - a} (x - a)^2.$$

Fácilmente vemos que  $f$  é diferenciável em  $[a, b]$  e

$$\varphi'(x) = f'(x) - \frac{f'(b) - f'(a)}{b - a} (x - a).$$

Como  $\varphi'(a) = f'(a) = \varphi'(b)$ , segue do teorema de Flett que existe  $\xi \in (a, b)$  tal que

$$\varphi'(\xi) = \frac{\varphi(\xi) - \varphi(a)}{\xi - a}.$$

Portanto, existe  $\xi \in (a, b)$  tal que a equação (6) é satisfeita, o que prova o teorema.  $\square$

Inspirados, então, pela afirmação do Teorema 2.3, demonstra-se o resultado abaixo.

**Teorema 4.2** ([3, Teorema 2.1]). *Se  $f : [a, b] \rightarrow \mathbb{R}$  é diferenciável em  $[a, b]$ , então, existe  $\xi \in (a, b)$  tal que*

$$f(b) - f(\xi) = (b - \xi)f'(\xi) + \frac{1}{2} \frac{f'(b) - f'(a)}{b - a} (b - \xi)^2.$$

O resultado abaixo é também uma generalização do teorema de Flett com outra condição do tipo Rolle  $f''(a) = f''(b)$ .

**Teorema 4.3** ([15, Exercício 5.3.11(b)]). *Seja  $f : [a, b] \rightarrow \mathbb{R}$  duas vezes diferenciável e tal que  $f''(a) = f''(b)$ . Então, existe  $\xi \in (a, b)$  tal que*

$$f(\xi) - f(a) = (\xi - a)f'(\xi) - \frac{(\xi - a)^2}{2} f''(\xi).$$

O resultado abaixo é análogo ao Teorema anterior, também com a condição do tipo Rolle  $f''(a) = f''(b)$ .

**Teorema 4.4.** *Seja  $f : [a, b] \rightarrow \mathbb{R}$  duas vezes diferenciável e tal que  $f''(a) = f''(b)$ . Então, existe  $\xi \in (a, b)$  tal que*

$$f(b) - f(\xi) = (b - \xi)f'(\xi) - \frac{(b - \xi)^2}{2} f''(\xi).$$

A demonstração do Teorema 4.4 é análoga à demonstração do Teorema 4.3 e portanto deixamos como exercício para o leitor.

Os Teoremas 2.2 e 4.3 foram generalizados por I. Pawlikowska em [14] para funções  $n$  vezes diferenciáveis, com a condição do tipo Rolle  $f^{(n)}(a) = f^{(n)}(b)$ .

**Teorema 4.5** ([14, Lema 2.2]). *Seja  $f : [a, b] \rightarrow \mathbb{R}$   $n$  vezes diferenciável e tal que  $f^{(n)}(a) = f^{(n)}(b)$ . Então, existe  $\xi \in (a, b)$  tal que*

$$f(\xi) - f(a) = \sum_{i=1}^n \frac{(-1)^{i+1}}{i!} (\xi - a)^i f^{(i)}(\xi).$$

**4.2. Algumas Aplicações.** A seguir vamos apresentar algumas aplicações do teorema de Flett. Vamos tratar, principalmente, dos trabalhos feitos por C. Lupu e T. Lupu em [9] e por C. Lupu em [10].

O objetivo desta seção é apresentar algumas propriedades importantes sobre alguns operadores integrais, como o de Volterra.

Lembremos que  $C([0, 1])$  denota o conjunto das funções contínuas reais definidas em  $[0, 1]$  e  $C^1([0, 1])$  é o conjunto de todas as funções reais continuamente diferenciáveis definidas no mesmo intervalo.

Definamos operadores  $T, S : C([0, 1]) \rightarrow C([0, 1])$  por

$$(T\varphi)(t) = \varphi(t) - \int_0^t \varphi(x) dx$$

$$(S\psi)(t) = t\psi(t) - \int_0^t x\psi(x) dx.$$

As seguintes propriedades valem para os operadores  $T$  e  $S$ .

**Teorema 4.6** ([9, Teorema 2.11]). *Se  $f, g : [0, 1] \rightarrow \mathbb{R}$  são funções contínuas, então existem  $\xi_1, \xi_2, \xi_3 \in (0, 1)$  tal que*

$$\int_0^1 f(x) dx (Tg)(\xi_1) = \int_0^1 g(x) dx (Tf)(\xi_1)$$

$$(Tf)(\xi_2) = (Sf)(\xi_2)$$

$$\int_0^1 f(x) dx (Sg)(\xi_3) = \int_0^1 g(x) dx (Sf)(\xi_3).$$

**Teorema 4.7** ([9, Teorema 2.12]). *Se  $f, g : [0, 1] \rightarrow \mathbb{R}$  são funções contínuas, então existem  $\xi_1, \xi_2 \in (0, 1)$  tal que*

$$\int_0^1 (1-x)f(x) dx (Tg)(\xi_1) = \int_0^1 (1-x)g(x) dx (Tf)(\xi_1)$$

$$\int_0^1 (1-x)f(x) dx (Sg)(\xi_2) = \int_0^1 (1-x)g(x) dx (Sf)(\xi_2).$$

As demonstrações dos Teoremas 4.6 e 4.7 encontram-se com detalhes em [9, Teorema 2.11 e Teorema 2.12].

O seguinte resultado é a versão equivalente do teorema do valor médio de Flett para o teorema do valor médio de Cauchy e é criticamente utilizado para obter os resultados seguintes.

**Teorema 4.8** ([10, Lema 2.1]). *Sejam  $f, g : [a, b] \rightarrow \mathbb{R}$  funções diferenciáveis em  $[a, b]$  com  $g'(x) \neq 0$  para todo  $x \in [a, b]$  e  $\frac{f'(a)}{g'(a)} = \frac{f'(b)}{g'(b)}$ . Então, existe  $\xi \in (a, b)$  tal que*

$$\frac{f(\xi) - f(a)}{g(\xi) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Denotemos por  $L^2((0, 1))$  o espaço vetorial das funções reais quadrado integráveis a Lebesgue sobre  $(0, 1)$ , i.e.,

$$L^2((0, 1)) = \left\{ f : (0, 1) \rightarrow \mathbb{R} : f \text{ é Lebesgue mensurável e } \int_a^b f^2(x) dx < \infty \right\}.$$

Observemos que para funções contínuas no intervalo  $(0, 1)$  a integral de Lebesgue e de Riemann coincidem, logo podemos pensar num primeiro momento no espaço  $L^2((0, 1))$  como sendo o espaço das funções contínuas em  $(0, 1)$  cujo quadrado é Riemann integrável em  $(0, 1)$ . Mais ainda, o espaço  $L^2((0, 1))$  é um espaço vetorial normado com a norma

$$\|f\|_{L^2((0,1))} = \left( \int_a^b f^2(x) dx \right)^{1/2}, \quad f \in L^2((0, 1)).$$

**Definição 4.1** (Operador de Volterra). *Sejam  $f \in L^2((0, 1))$  e  $x \in (0, 1)$ . Definimos o operador de Volterra,  $V$ , por*

$$\begin{aligned} V : L^2((0, 1)) &\rightarrow L^2((0, 1)) \\ f &\rightarrow V(f)(x) = \int_0^x f(t) dt. \end{aligned}$$

Sejam  $\Psi, \phi : [0, 1] \rightarrow \mathbb{R}$  funções sendo  $\Psi$  contínua e  $\phi$  diferenciável com  $\phi'(x) \neq 0$  para todo  $x \in (0, 1)$ . Definimos o operador do *tipo Volterra com peso*

$$V_\phi \Psi(t) = \int_0^t \phi(x) \Psi(x) dx.$$

No que segue, definimos os espaços

$$\mathfrak{C}([a, b]) := \{ \phi \in C^1([a, b]), \phi'(x) \neq 0, x \in [a, b], \phi(a) = 0 \} \text{ e}$$

$$C_{nula}([a, b]) := \left\{ f \in C([a, b]) : \int_a^b f(x) dx = 0 \right\}.$$

**Teorema 4.9.** *Seja  $f \in C_{nula}([a, b])$  e  $g \in C^1([a, b])$ , com  $g'(x) \neq 0$  para todo  $x \in [a, b]$ . Então, existe  $\xi \in (a, b)$  tal que*

$$V_g f(\xi) = g(a) \cdot V f(\xi).$$

**Demonstração.** Consideremos as funções  $\varphi, \eta : [a, b] \rightarrow \mathbb{R}$  dadas por

$$\begin{aligned} \varphi(t) &= \int_a^t f(x) g(x) dx - g(t) \int_a^t f(x) dx \text{ e} \\ \eta(t) &= g(t). \end{aligned}$$



Como  $\varphi$  é diferenciável segue que

$$\varphi'(t) = f(t)g(t) - \left( g'(t) \int_a^t f(x)dx + g(t)f(t) \right) = -g'(t) \int_a^t f(x)dx.$$

Observemos que  $\varphi'(a) = 0$ , assim  $\frac{\varphi'(a)}{\eta'(a)} = 0$ .

Por outro lado,  $\varphi'(b) = -g'(b) \int_a^b f(x)dx = 0$  pois  $f \in C_{nula}([a, b])$ , e daí,

$$\frac{\varphi'(b)}{\eta'(b)} = 0.$$

Logo, do Teorema 4.8 segue que existe  $\xi \in (a, b)$  tal que

$$\frac{\varphi(\xi) - \varphi(a)}{\eta(\xi) - \eta(a)} = \frac{\varphi'(\xi)}{\eta'(\xi)}.$$

Equivalentemente

$$\frac{\int_a^\xi f(x)g(x)dx - g(\xi) \int_a^\xi f(x)dx}{g(\xi) - g(a)} = \frac{-g'(\xi) \int_a^\xi f(x)dx}{g'(\xi)}.$$

Disso segue que

$$\int_a^\xi f(x)g(x)dx = g(a) \int_a^\xi f(x)dx.$$

□

**Teorema 4.10.** *Se  $f, g$  são funções reais contínuas em  $[0, 1]$  e  $\phi \in \mathfrak{C}([0, 1])$ , então existe  $\xi \in (0, 1)$  tal que*

$$\begin{aligned} V_\phi f(\xi) \int_0^1 g(x)dx - V_\phi g(\xi) \int_0^1 f(x)dx \\ = \phi(0) \left( Vf(\xi) \int_0^1 g(x)dx - Vg(\xi) \int_0^1 f(x)dx \right). \end{aligned}$$

Os detalhes da demonstração do Teorema 4.10 podem ser encontrados em [10, Teorema 2.4]. Além disso, algumas observações a respeito deste teorema são pertinentes:

(i) Se  $\phi(0) = 0$  então existe  $\xi \in (0, 1)$  tal que

$$V_\phi f(\xi) \int_0^1 g(x)dx = V_\phi g(\xi) \int_0^1 f(x)dx;$$

Em particular se  $\phi(x) = x$ , existe  $\xi \in (0, 1)$  tal que

$$\int_0^1 f(x)dx \int_0^\xi xg(x)dx = \int_0^1 g(x)dx \int_0^\xi xf(x)dx.$$

(ii) Considere o espaço  $L^2$  com peso dado por

$$L_\phi^2(0, \xi) = \left\{ u : (0, \xi) \rightarrow \mathbb{R} : \int_0^\xi u^2(x)\phi(x) dx < \infty \right\}$$

e equipado com a norma

$$\|u\|_{L_\phi^2(0, \xi)} = \left( \int_0^\xi u^2(x)\phi(x) dx \right)^{1/2}, \quad u \in L_\phi^2(0, \xi).$$

Substituindo  $f$  e  $g$  por  $f^2$  e  $g^2$  temos que existe  $\xi \in (0, 1)$  tal que

$$\begin{aligned} V_\phi f^2(\xi) \int_0^1 g^2(x)dx - V_\phi g^2(\xi) \int_0^1 f^2(x)dx \\ = \phi(0) \left( V f^2(\xi) \int_0^1 g^2(x)dx - V g^2(\xi) \int_0^1 f^2(x)dx \right), \end{aligned}$$

ou seja,

$$\begin{aligned} \|f\|_{L_\phi^2(0, \xi)}^2 \|g\|_{L^2(0, 1)}^2 - \|g\|_{L_\phi^2(0, \xi)}^2 \|f\|_{L^2(0, 1)}^2 \\ = \phi(0) (\|f\|_{L^2(0, \xi)}^2 \|g\|_{L^2(0, 1)}^2 - \|g\|_{L^2(0, \xi)}^2 \|f\|_{L^2(0, 1)}^2). \end{aligned}$$

Desta última igualdade, se  $\phi(0) = 0$ , obtemos

$$(7) \quad \|f\|_{L_\phi^2(0, \xi)} \|g\|_{L^2(0, 1)} = \|g\|_{L_\phi^2(0, \xi)} \|f\|_{L^2(0, 1)}.$$

Escrevendo a equação (7) da seguinte maneira

$$(8) \quad \frac{\|f\|_{L_\phi^2(0, \xi)}}{\|g\|_{L_\phi^2(0, \xi)}} = \frac{\|f\|_{L^2(0, 1)}}{\|g\|_{L^2(0, 1)}},$$

concluimos a seguinte propriedade interessante: dadas duas funções  $f, g$  que tem normas iguais (ou proporcionais) em  $L^2(0, 1)$ , e se for dada uma função peso não constante  $\phi$ , então existe um número  $\xi \in (0, 1)$  onde as normas das funções serão iguais (ou proporcionais) em  $L_\phi^2(0, \xi)$ .

## 5. PROBLEMAS EM ABERTO

O estudo de condições necessárias e suficientes para a existência de pontos de Flett (veja seção 2) não é, até onde sabemos, completo. Recorde que a função  $f(x) = \text{sgn}(x)$  pertence ao conjunto  $\Delta_{12}$  (veja a Figura 5), portanto, não satisfaz nenhuma das condições discutidas anteriormente. No entanto, possui infinitos pontos de Flett. Esta observação, sozinha, torna natural três perguntas, ainda não respondidas na literatura:

**Pergunta 5.1.** *Além das apresentadas neste trabalho, existem outras condições suficientes para a existência de pontos de Flett?*

**Pergunta 5.2.** *Existe uma condição necessária para a existência de pontos de Flett?*

**Pergunta 5.3.** *Assumindo que uma função possua pelo menos um ponto de Flett. Sob quais condições este seria único?*

**Agradecimentos:** Os autores agradecem ao parecerista pelos comentários e observações, que ajudaram a melhorar de maneira significativa a apresentação deste trabalho. Este trabalho é fruto da Iniciação Científica realizada pelo primeiro autor, quem agradece o apoio da FAPESP através do Processo 2013/03866-9.

#### REFERÊNCIAS

- [1] R.G. Bartle e Donald R. Sherbert. *Introduction to real analysis*. Four edition. John Wiley & Sons, Inc., New York, 2011.
- [2] F. Cajori. *A history of mathematics*. Fifth Edition. AMS, 1991.
- [3] D. Cakmak, A. Tiryaki. *Mean value theorem for holomorphic functions*, Vol. 2012 (2012), No. 34, pp. 1–6.
- [4] T.M. Flett. *A mean value problem*, Math. Gaz. 42(1958), pp. 38-39.
- [5] O. Hutník and J. Molnárová. *On Flett's mean value theorem*. Aequat. Math., Published online: 22 October 2014.
- [6] G. Lozada-Cruz. *Some variants of Lagrange's mean value theorem*. Sel. Mat. 7 (1)(2020), pp. 144–150.
- [7] G. Lozada-Cruz. *Some variants of Cauchy's mean value theorem*. Internat. J. Math. Ed. Sci. Tech. 51 (2020), no. 7, pp. 1155–1163.
- [8] G. Lozada-Cruz. *Some applications of Cauchy's mean value theorem*. Asia Pac. J. Math. 7 (30) (2020), pp. 1–9.
- [9] C. Lupu e T. Lupu. *Mean value theorems for some linear integral operators*, Electron. J. Diff. Eqn., 2009, no. 117, pp. 1–15.
- [10] C. Lupu. *Mean value problems of Flett type for a Volterra operator*, Electron. J. Diff. Eqn., 2013, No. 53, pp. 1–7.
- [11] B.J. Malešević. *Some mean value theorem in terms of an infinitesimal function*. Mat. Vesnik 51 (1999), pp. 9–13.
- [12] J. Malešević. *O jednoj teoremi G. Darboux-a i teoremi D. Trahan-a*. Mat. Vesnik 4 (13) (32) (1980), pp. 304–312.
- [13] R.E. Meyers. *Some elementary results related to the mean value theorem*, The two-year college Mathematics Journal. Vol. 8, No. 1 (1977), pp. 51-53.
- [14] I. Pawlikowska. *An extension of a theorem of Flett*. Demonstratio Math. 32 (1999), no. 2, pp. 281–286.
- [15] T.L. Radulescu, V.D. Radulescu, T. Andreescu. *Problems in Real Analysis: advanced calculus on the real axis*, Springer Verlag, 2009.
- [16] P.K. Sahoo and T. Riedel. *Mean Value Theorems and Functional Equations*, World Scientific Publishing Company, River Edge, NJ, 1998.
- [17] J. Tong. *On Flett's mean value theorem*. Internat. J. Math. Ed. Sci. Tech. 35 (2004), 936-941.
- [18] D.H. Trahan. *A new type of mean value theorem*. Math. Mag. 39, pp. 264–268, (1966).

DEPARTMENT OF MATHEMATICAL SCIENCES, THE UNIVERSITY OF MEMPHIS, MEMPHIS, TN 38152, USA. *Email address:* msngrti@memphis.edu

DEPARTAMENTO DE MATEMÁTICA, INSTITUTO DE BIOCÊNCIAS, LETRAS E CIÊNCIAS EXATAS (IBILCE)–UNIVERSIDADE ESTADUAL PAULISTA (UNESP), SÃO JOSÉ DO RIO PRETO, BRASIL. *Email address:* german.lozada@unesp.br