RMU
REVISTA MATEMÁTICA UNIVERSITÁRIA

ISSN: 2675-5254

Ano 2022, Volume 1

Índice

Humberto Bortolossi
Editorial: Sobre a logomarca da RMU 1

Jean Cerqueira Berni, Hugo Luiz Mariano
A Geometria Diferencial Sintética e os mundos onde podemos interpretá-la: um convite ao estudo dos anéis $\mathcal{C}^\infty$ 5

Jaqueline Rojas, Sally Andria
Klein nos ajuda contar retas em superfífices projetivas 31

Jorge A. J. Avila, Elielsom D. Moreira, Bianca F. Guimarães
Padrões especiais de distribuição dos números primos: o $n-$*quadrado Zeta* 55

Thaís Fernanda Mendes Monis, Bruno Aparecido Pim
As muitas facetas da teoria dos jogos: da economia à biologia evolutiva 79

Marcio Colombo Fenille
Comparação de topologias sobre as partes de um espaço métrico . . . 97

Ronaldo Freire de Lima
O Teorema Espectral e suas misteriosas relações com o Teorema da Decomposição Primária 107

SBM
SOCIEDADE BRASILEIRA DE MATEMÁTICA

SOBRE A RMU

A Revista de Matemática Universitária (RMU) é uma publicação digital da Sociedade Brasileira de Matemática voltada para alunos e professores de cursos de nível superior de MatemÃ¡tica em suas várias modalidades. Seus objetivos são promover o intercâmbio entre os membros da comunidade matemática brasileira e apresentar a beleza, a vitalidade e os aspectos interdisciplinares, históricos, humanos e culturais da Rainha das Ciências.

Comitê Editorial

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

ISSN
2675–5254l

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`
Página web: `https://rmu.sbm.org.br/`

RMU
REVISTA MATEMÁTICA UNIVERSITÁRIA

SBM
SOCIEDADE BRASILEIRA DE MATEMÁTICA

RMU
REVISTA MATEMÁTICA UNIVERSITÁRIA

SBM
SOCIEDADE BRASILEIRA DE MATEMÁTICA

Revista Matemática Universitária, vol. 1, 2019
ISSN: 2675-5254 — **DOI:** 10.21711/26755254/rmu20221

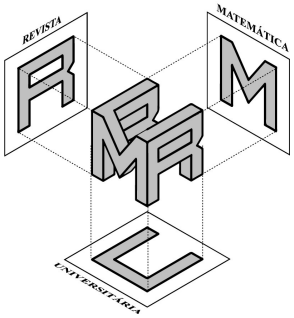
EDITORIAL: SOBRE A LOGOMARCA DA RMU

HUMBERTO BORTOLOSSI,

Resumo. Neste artigo discutimos a Matemática por trás da logomarca da Revista Matemática Universitária da SBM.

1. Introdução

A logomarca da Revista Matemática Universitária, que pode ser apreciada no canto superior esquerdo dessa página, é constituída por um objeto tridimensional que possui uma propriedade especial: ao ser observado de três direções diferentes vêem-se as iniciais R, M e U ou, dito de outra forma, as projeções ortogonais do objeto tridimensional nos planos coordenados são essas letras.

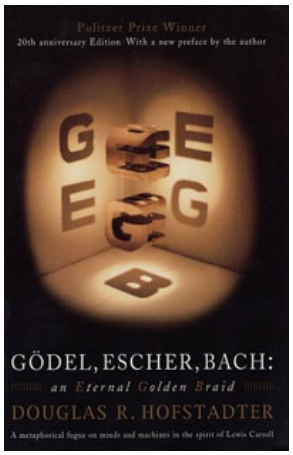


https://www.geogebra.org/m/gc54mm9r

Uma versão interativa do logo no Geogebra aparece em https://www.geogebra.org/m/gc54mm9r.

Esse tipo de sólido, denominado *trip-let* ou ambigrama 3D, foi idealizado por Douglas R. Hofstadter para ilustrar a capa de seu livro *Gödel, Escher, Bach: An Eternal Golden Braid* ([2]).

Data de aceitação: 24 de agosto de 2022.
Palavras chave. geometria, *triplet*, ambigrama superdícies cilíndricas.

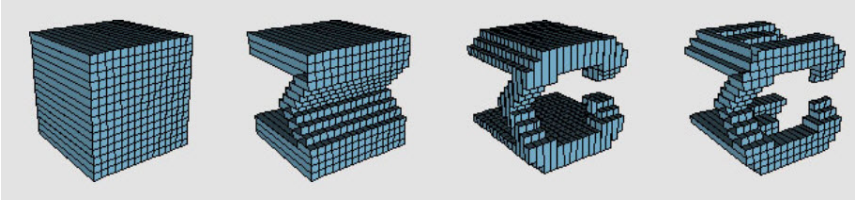


2. TRIPLETS: O JOGO

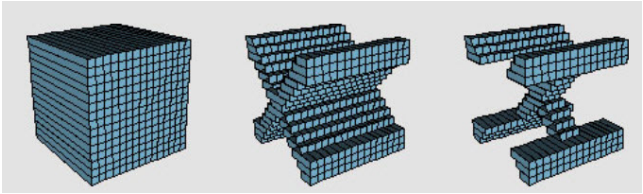
Inspirados pela ideia de Hofstadter, desenvolvemos um jogo interativo para exercitar visualização tridimensional em português, inglês, espanhol e francês. O jogador deve manipular uma *trip-let* aleatória e identificar três letras do alfabeto, com as quais ele deve compor uma palavra ou sigla sem acentos. Para auxiliar na visualização, os eixos e os planos coordenados podem ser exibidos. Além do exercício de vocabulário e do exercício cognitivo de percepção de formas tridimensionais, o software também pode ser usado para tratar de questões de simetria (dos formatos das letras) e de contagem (das possíveis palavras para um mesmo conjunto de letras).

http://www.cdme.im-uff.mat.br/html5/triplets/triplets-html/triplets-br.html

Como o aplicativo gera os objetos 3D? Tomemos, por exemplo, a sigla MEC (ou a palavra CEM). O software tenta montar a *trip-let* da seguinte maneira: a partir de um par de faces opostas de um cubo, ele corta o formato da letra M. Em seguida, a partir de um outro par de faces, o programa corta a letra C. Por fim, no par de faces restantes, ele corta a letra E, como indica a figura a seguir.

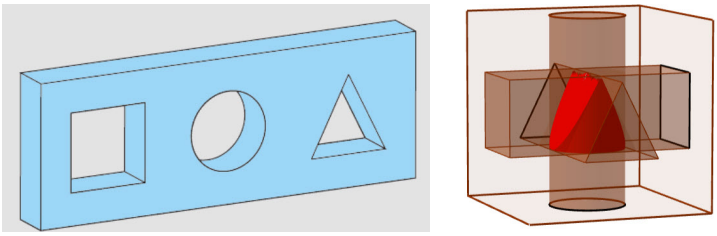


Dependendo da escolha das três letras, não é possível construir uma *triplet* seguindo este algoritmo. Por exemplo, não é possível construir uma *trip-let* para a palavra XIS. Ao se cortar as letras X e I em dois pares de faces opostas, o par restante fica automaticamente com a letra I, a partir da qual não é possível se obter a letra S.



3. Considerações Finais.

Um *trip-let* é definido pela interseção de três superfícies cilíndricas cujas bases são os formatos das letras. Mais geralmente, a interseção de superfícies cilíndricas faz parte da cultura matemática. O divulgador matemático americano Martin Gardner, por exemplo, propôs o seguinte quebra-cabeça: encontrar uma rolha que possa ser usada para tapar três buracos nas formas de um quadrado, de um círculo e de um triângulo. onde o diâmetro do círculo é igual ao lado do quadrado e o triângulo é isósceles, com base e altura com medidas também iguais ao lado do quadrado. A solução é obtida pela interseção de três superfícies cilíndricas. Uma versão interativa desse quebra-cabeça usando o Geogebra aparece em: https://www.geogebra.org/m/Q7eXY36j



Outro exemplo clássico, presente em quase todos os livros de Cálculo a Várias Variáveis, é o de determinar o volume do sólido resultante da interseção de três cilindros circulares $x^2 + y^2 = 1$, $x^2 + z^2 = 1$ e $y^2 + z^2 = 1$.

É natural perguntar se existem sólidos cujas projeções consistam de letras em mais que três planos. A resposta é sim, e não só para letras. Um lindo resultado da teoria de fractais ([3], teorema 6.9) garante que dada uma família de figuras, uma para cada subespaço 2-dimensional de $\mathbb{R}^3$, existe um conjunto em $\mathbb{R}^3$ tal que quase todas suas projeções nesses subespaços são as figuras dadas, a menos de conjuntos

de medida zero. Um vídeo exemplificando esse teorema através da construção de *trip-lets* em cristais por meio de raios laser pode ser encontrado em https://www.youtube.com/watch?v=EKsk8J1gGAA ([4]).

Agradecimentos

O autor agradece a Luis Edmundo Carlos Pinto Dantas, Lhaylla Crissaff, André Rapozo e Michel Spira pelos comentários, críticas, sugestões e correções.

Referências

[1] Gardner, Martin: *The Cork Plug.* Chapter 5 in *The Second Scientific American Book of Puzzles & Diversions: a New Selection. New York*, Simon and Schuster, pp. 52-59, 1961.

[2] Hofstadter, Douglas Richard: *Gödel, Escher, Bach: Um Entrelaçamento de Gênios Brilhantes*, Brasília, UnB, 2001.

[3] Falconer, Kenneth: *Fractal Geometry: Mathematical Foundations and Applications*, John Willey & Sons, 1990.

[4] Nakayama, Hirotaka *et al*: Three-dimensional volume containing multiple two-dimensional information patterns, *Nature*, Scientific Reports v. 3, 2013. doi: DOI:10.1038/srep01931

Departamento de Matemática Aplicada
Estrada Instituto de Matemática e Estatística
Universidade Federal Fluminense
Niterói – RJ
Orcid: https://orcid.org/0000-0003-1212-6252
Email address: humbertobortolossi@id.uff.br

RMU
REVISTA MATEMÁTICA UNIVERSITÁRIA

SBM
SOCIEDADE BRASILEIRA DE MATEMÁTICA

Revista Matemática Universitária, vol. 1, 2022
ISSN: 2675-5254 — **DOI:** 10.21711/26755254/rmu20222

A GEOMETRIA DIFERENCIAL SINTÉTICA E OS MUNDOS ONDE PODEMOS INTERPRETÁ-LA: UM CONVITE AO ESTUDO DOS ANÉIS $\mathcal{C}^\infty$

JEAN CERQUEIRA BERNI E HUGO LUIZ MARIANO

Sumário

Introdução 5
1. Um axioma da Geometria Diferencial Sintética 11
2. Mundos (modelos) para a Geometria Diferencial Sintética 15
   2.1. Incompatibilidade com a Lógica Clássica e com o "mundo" dos Conjuntos 18
   2.2. Um topos que valida o Axioma de Kock–Lawvere 19
   2.3. Um "Toy Model" para a Geometria Diferencial Sintética 23
   2.4. Há um mundo bem melhor: o topos de Dubuc 26
3. Considerações Finais 28
Referências 29

Introdução

A Geometria Diferencial tem como um de seus principais objetos de estudo a noção de variedade (uma "quantidade $n-$vezes estendida"). Este conceito remonta à palestra de habilitação[1] de B. Riemann, em 1854, e as teorias das variedades diferenciáveis e das variedades Riemannianas surgiram como um esforço de seus sucessores para esclarecer e formalizar as ideias ali apresentadas. Originalmente essa noção abrangia, inclusive, variedades de dimensão infinita, como por exemplo,

Data de aceitação: Outubro de 2020.
Palavras chave. Geometria Diferencial Sintética, Topoi, Anéis $\mathcal{C}^\infty$.
Projeto financiado pela CAPES, Projeto PNPD 88882.463201/2019-01.

[1] "Über die Hypothesen welche die Geometrie zu Grunde liegen" — ou, "Dos pressupostos que subjazem aos axiomas da Geometria".

a dos "valores possíveis de uma função em uma dada região", a "das formas possíveis de uma figura sólida" e assim por diante.

Ao longo do desenvolvimento desses esforços, a noção de quantidades "infinitamente pequenas" (ou "infinitésimos"), tão úteis nos processos de descobertas de novos resultados na Geometria — sobretudo para geômetras como S. Lie e E. Cartan — foi sendo reduzida a uma ideia vaga de, como expressou o bispo Berkeley, *"fantasmas de quantidades que já partiram"*, sem qualquer legitimidade no âmbito do rigor introduzido por K. Weierstrass. Com o passar do tempo, apesar das críticas, o rigor de Weierstrass prevaleceu, de modo que os infinitésimos - que tantas vezes conduziam a inconsistências lógicas - acabaram se tornando meramente uma ferramenta heurística que, embora útil, deveria ser "evitada" na demonstração de novos resultados.

Em 1913 surge explicitamente o conceito de variedade diferenciável como conhecemos hoje, em um trabalho de H. Weyl[2] tratando também do conceito de superfícies de Riemann. Como se sabe, na Geometria Diferencial das variedades diferenciáveis como as conhecemos hoje não há espaço para "infinitésimos", ou seja, para entes "infinitamente pequenos", o que acaba por obscurecer a natureza de diversas definições e a essência de vários teoremas dessa área do conhecimento matemático. Mas e se, de algum modo, pudéssemos conciliar os infinitésimos com o rigor?

Ao longo da História, a Geometria foi estudada essencialmente por dois pontos de vista principais: o ponto de vista analítico e o ponto de vista sintético. Enquanto a abordagem analítica é calcada na introdução de sistemas de coordenadas em seus objetos de estudo (ou seja, na atribuição a cada ponto de uma variedade diferenciável n-dimensional uma $n-$upla de números), a abordagem sintética pretende formular a Geometria em termos de axiomas capazes de "capturar" a intuição que fazemos dos entes geométricos e das relações entre eles, de modo que os resultados sobre Geometria sejam deduzidos "formalmente", mediante deduções lógicas.

O leitor possivelmente estará familiarizado com alguns exemplos de geometrias sintéticas: a Geometria contida em "Os Elementos", de Euclides, as Geometrias Finitas e a Geometria Projetiva Sintética, introduzida (canhestramente) por Desargues no século XVII, são geometrias nas quais o *modus operandi* é deduzir, a partir de um corpo inicial de axiomas, diversos resultados acerca de seus entes geométricos e das relações entre eles.

O poder da abordagem analítica (clássica) à Geometria Diferencial procede, em grande parte, da possibilidade de se empregar métodos do Cálculo Diferencial e, mais geralmente, da Análise Matemática, à Geometria. Podemos estudar "vetores tangentes", "ângulos entre curvas" e medir áreas com o auxílio do aparato do Cálculo Diferencial, ainda que apenas "localmente", por meio da introdução de sistemas de coordenadas. Posto desta forma, a ideia de uma Geometria Diferencial que prescinda desse aparato e se utilize somente de axiomas pode parecer, à primeira vista, "improdutiva", ou mesmo "estéril".

[2] "Die Idee der Riemannschen Fläche", ou "O conceito de superfícies de Riemann".

FIGURA 1. Anders Kock e Francis William Lawvere no *Café Odeon* em Zürich, 1966. Extraído de [**Pic**].

De acordo com J. Bell (cf. [**Bel**]), uma primeira tentativa de desenvolver uma Geometria Diferencial Sintética[3] foi feita por Herbert Busemann, nos anos 1940, baseada em um trabalho de Paul Finsler. A ideia era construir uma Geometria Diferencial que, nas palavras do autor, dispensasse o uso de "derivadas". Os objetos básicos de sua abordagem não eram, como hoje, as superfícies e variedades diferenciáveis, mas espaços métricos de um certo tipo nos quais a noção de geodésica era definida de modo intrínseco.

Em 1967, quando Francis W. Lawvere (homem à direita na Figura 1) era professor assistente na Universidade de Chicago e S. MacLane ministrava um curso sobre Mecânica, Lawvere passou a refletir sobre possíveis justificativas para os velhos métodos "intuitivos" (à maneira de S. Lie e de E. Cartan por exemplo, hoje muito bem conhecidos por físicos e engenheiros) empregados por MacLane para apresentar tópicos de Geometria Diferencial pertinentes àquele curso. Destas reflexões surgiu o programa da "Dinâmica Categorial" e da "Geometria Diferencial Sintética". O então estudante de doutorado de Lawvere, Anders Kock (à esquerda na Figura 1), prosseguiu com a pesquisa em Geometria Diferencial Sintética, assim como outros matemáticos, como Eduardo J. Dubuc - orientando de S. MacLane naquela época. Desde então tem havido esforços para compor um quadro axiomático capaz de "capturar" a noção de continuidade do espaço físico (muito utilizada em Mecânica, por exemplo) e do qual se possa deduzir a maior quantidade possível de resultados da Geometria Diferencial.

[3]na qual a noção de "curva geodésica" em variedades Riemannianas e de Finsler é descrita mediante axiomas, cf. [**Bus**].

A Geometria Diferencial Sintética proposta por Lawvere é formulada no contexto da Teoria das Categorias, um ramo da Matemática que trata dos "mundos" e estruturas matemáticas de um modo mais geral. Uma categoria corresponde, a grosso modo, a um "mundo matemático", composto por objetos e por transformações entre esses objetos. Temos, por exemplo, a categoria **Sets**, dos conjuntos e das funções entre conjuntos, **Grp**, dos grupos e homomorfismos de grupos, **Top**, a categoria dos espaços topológicos e funções contínuas e assim por diante.

Sendo a Geometria Diferencial um estudo sobre os objetos da categoria **Man**, das variedades diferenciáveis (espaços topológicos de Hausdorff, com base enumerável e munidos de um atlas maximal $\mathfrak{A}$ de dimensão n, para algum $n \in \mathbb{N}$ e de classe $\mathcal{C}^\infty$) e cujas transformações são as funções infinitamente diferenciáveis entre essas, pode-se supor que um dos objetivos da Geometria Diferencial Sintética seja dar uma descrição axiomática para essa categoria. Vejamos quais características de **Man** inviabilizam uma tal descrição.

A categoria **Man** padece de certas "deficiências" que dificultam uma axiomatização, a começar por sua "escassez" de certos tipos de construções (denominados na Teoria de Categorias por "limites inversos"), que não podem ser levadas a cabo em seu interior. Por exemplo, dadas duas variedades diferenciáveis M e N, uma subvariedade $Z \subseteq N$ e uma função suave (*i.e.*, de classe $\mathcal{C}^\infty$) qualquer $f : M \to N$, seu pullback:

$$\begin{array}{ccc} f^{-1}[Z] & \longrightarrow & M \\ \downarrow & & \downarrow f \\ Z & \hookrightarrow & N, \end{array} \tag{1}$$

em geral, *não* é uma variedade diferenciável [4]. De fato, a interseção não transversal de subvariedades suaves, na maioria das vezes, não é uma subvariedade suave[5]. Tem se tornado relativamente comum em alguns círculos referir-se a estas "lacunas" em **Man** como uma "crise nos fundamentos da Geometria Diferencial", embora talvez a palavra "crise" não seja a mais adequada. A percepção dessas lacunas varia de acordo com a subárea da Geometria Diferencial em que se trabalha: naquelas em que os problemas fundamentais são de caráter analítico, pouco se sente sobre a estrutura categorial, enquanto que naquelas em que se estuda quocientes, moduli spaces, folheações e ações de grupos (dentre outros temas), as deficiências categoriais implicam na introdução de uma profusão de outros objetos que tentam de alguma forma saná-las, como grupóides, stacks, orbifolds, e espaços de Chen, Frölicher e

[4] salvo mediante a condição de transversalidade: se $M \xrightarrow{f} N$ é uma função suave transversal à subvariedade $Z \subseteq N$, então (1) é um pullback em **Man**.

[5] Em virtude de um famoso teorema de H. Whitney, todo subconjunto fechado de $\mathbb{R}^n$, $n \in \mathbb{N}$, é o conjunto dos zeros de alguma função suave. Tomando, por exemplo, o conjunto de Cantor, C [que é fechado] em $\mathbb{R}$ (e que, por diversos motivos *não* é uma variedade suave), existe $f : \mathbb{R} \to \mathbb{R}$ suave tal que $C = f^{-1}[\{0\}]$. Desta forma, a interseção de Graf (f) (variedade suave de dimensão 1) com Graf (0) (variedade suave de dimensão 1), o gráfico da função identicamente nula, é C, que evidentemente não é uma variedade suave.

difeológicos, que são mais gerais que variedades suaves (e portanto não tão bem comportados) mas que compõem categorias que possuem melhores propriedades.

Frequentemente desejamos considerar, na categoria **Man**, e de um modo natural, um mesmo conceito sob pontos de vista distintos - conceitos importantes para a aplicação da Geometria Diferencial, por exemplo, à Mecânica dos Sólidos. Lawvere, por exemplo, ponderou o seguinte: o movimento de um certo corpo B (por exemplo, um sistema $0-$dimensional de partículas, uma corda elástica $1-$dimensional, uma membrana $2-$dimensional ou um sólido $3-$dimensional) é frequentemente representado por uma função:

$$q : B \times T \to E,$$

onde T é o espaço $1-$dimensional que representa os instantes (tempo) e E é o espaço $3-$dimensional onde o movimento se dá. Neste contexto, o movimento pode ser pensado como uma função que associa ao par:

$$(\text{partícula de } B, \text{instante}) \in B \times T$$

uma certa posição em E durante o movimento.

Para outros fins, porém, convém interpretar o movimento como uma função:

$$\bar{q} : B \to E^T,$$

que associa a cada partícula de B o seu trajeto através de E. Uma categoria em que existe uma correspondência natural biunívoca:

$$\frac{q : B \times T \to E}{\bar{q} : B \to E^T}$$

recebe a denominação de categoria "cartesiano-fechada" — o que, feliz ou infelizmente, não é o caso de **Man** (na verdade, dadas duas variedades suaves M e N, $M^N = \mathcal{C}^\infty(N, M)$, salvo em casos patológicos, não é uma variedade diferenciável e portanto, não é objeto de **Man**).

Finalmente, em **Man** não dispomos de uma linguagem adequada para lidar direta e explicitamente com estruturas "infinitesimais". Dentre essas estruturas, seria particularmente interessante dispor de certo objeto D, capaz de codificar o fibrado tangente de uma variedade diferenciável, M — ou seja, tal que $TM = M^D$. Este objeto D poderia ser considerado a "pedra filosofal" da Geometria Diferencial em virtude das inúmeras simplificações de conceitos que sua introdução proporcionaria, como os de vetores tangentes, fluxos, formas diferenciáveis, conexões etc. No entanto, em **Man**, nada se fala - por falta de vernáculo - acerca de infinitésimos.

Podemos agrupar essas características que desabonam **Man** para uma boa descrição axiomática em dois grupos principais:

(i) diversas construções envolvendo variedades diferenciáveis desbordam da categoria das variedades diferenciáveis - ou seja, não são, em si, variedades diferenciáveis;

(ii) não há, em **Man**, um objeto "infinitesimal" D, no sentido de ser o fibrado tangente de uma variedade diferenciável M, TM, identificado com uma "variedade diferenciável de todos os caminhos infinitesimais em M, M^D".

A ideia de Lawvere para superar esses obstáculos foi, então, "enriquecer" a categoria **Man**[6] a fim de obter uma categoria **S**, cujos objetos são denominados "espaços suaves", fechada sob essas construções e que admita uma descrição axiomática simples.

Passemos às características essenciais desejadas para uma tal categoria **S**, dos espaços suaves. São elas:

(1) **S** deve conter um "modelo algébrico" da reta geométrica (assim como **Man** contém $\mathbb{R}$), que denotaremos por R. O objeto R deve admitir (pelo menos) a estrutura de anel comutativo, munido de dois "elementos" especiais, 0 e e, $0 \neq e$, correspondendo respectivamente aos elementos neutros da soma e do produto[7]. Ademais, convém que R admita um subobjeto (não trivial) D, consistindo dos "elementos" infinitésimos quadrado-nilpotentes de R e capaz de representar o fibrado tangente de qualquer espaço suave;

(2) ao expandir **Man** para **S**, toda variedade diferenciável deve ser "mergulhada" em **S** como um espaço suave, e novas transformações entre essas variedades mergulhadas não devem ser admitidas - ou seja, os únicos mapas entre variedades mergulhadas devem ser os mapas suaves (o que não ocorre quando mergulhamos **Man** em **Sets**);

(3) **S** deverá ser uma categoria cartesiano-fechada, ou seja, dados espaços suaves M, N e P, $M \times N$ e P^N devem ser espaços suaves e, além disso, deverá haver uma bijeção natural relacionando estas construções:

$$\frac{M \times N \to P}{M \to P^N}$$

(4) **S** deve ter um vernáculo capaz de codificar a noção de nilpotência em R, ou seja, de "elementos" de R tais que alguma de suas potências seja 0. Note que obter nilpotentes diferentes de 0 é impossível em um corpo como $\mathbb{R}$. Consequentemente, em **S**, deve-se exigir de R que tenha uma estrutura um tanto mais fraca que a de corpo (no sentido clássico). Convém exigirmos que R seja uma $\mathbb{Q}-$álgebra, ou seja, que $e, 2 \cdot e, \cdots, n \cdot e, \cdots$ sejam todos elementos invertíveis em R. Dentre os subobjetos de R, deve haver um particularmente especial, D, consistindo desses "infinitésimos", incluindo os quadrados nilpotentes, e de representar[8] o funtor fibrado tangente.

Note, também, que sendo a reta geométrica modelada em **S** por R, o plano geométrico será modelado por $R \times R$, o espaço geométrico tridimensional por $R \times R \times R$ e assim por diante.

---

[6] tecnicamente isto corresponde a construir um mergulho pleno e fiel de **Man** em **S**.

[7] em qualquer categoria $\mathcal{C}$ com limites finitos, um objeto R tem estrutura de anel comutativo com unidade quando estão dados morfismos: $1 \underset{0}{\overset{e}{\rightrightarrows}} R \underset{\cdot}{\overset{+}{\leftleftarrows}} R \times R$ satisfazendo as identidades usuais que definem essa estrutura, expressas por diagramas comutativos.

[8] ou seja, tal que $T : \mathbf{S} \to \mathbf{S}$ seja naturalmente isomorfo ao funtor $\mathrm{Hom}_{\mathbf{S}}(D, \cdot) : \mathbf{S} \to \mathbf{S}$.

1. Um axioma da Geometria Diferencial Sintética

Vamos supor, para fins de discussão, que uma tal categoria **S**, detentora de todas as propriedades descritas anteriormente de fato exista. Denotaremos por $(R, +, \cdot, -, 0, e)$ seu objeto anel comutativo com unidade e, a fim enfatizar a importância deste objeto, escreveremos $(\mathbf{S}, (R, +, \cdot, -, 0, e))$ para denotar a categoria dos espaços suaves.

Sendo **S** uma categoria com limites finitos, podemos exprimir o subobjeto de "elementos quadrado-nilpotentes", D, do objeto anel R como o equalizador dos mapas:

$$\begin{array}{rccc} (\,\cdot\,)^2: & R & \to & R \\ & r & \mapsto & r^2 = r \cdot r \end{array}$$

e:

$$\begin{array}{rccc} 0: & R & \to & R \\ & r & \mapsto & 0 \end{array}$$

ou seja,

$$D = [\![d.R \mid d^2 = 0]\!] \rightarrowtail R \underset{0}{\overset{(\,\cdot\,)^2}{\rightrightarrows}} R$$

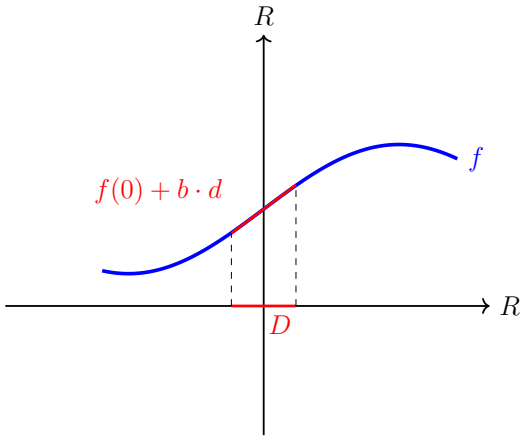
é um equalizador.

Por motivos que serão expostos posteriormente, podemos trabalhar em **S** "como se" estivéssemos na categoria dos conjuntos, fazendo menção à noção de pertencimento - embora, na realidade, em **S** haja uma noção mais flexível para o termo "elemento".

Um dos principais axiomas que dão fulcro à Geometria Diferencial Sintética é o:

Axioma 1 (**axioma de Kock–Lawvere**). *Em* $(\mathbf{S}, (R, +, \cdot, 0, e))$, *para qualquer* $f : D \to R$ *existe um único* $b \in R$ *tal que:*

$$(\forall d \in D)(f(d) = f(0) + b \cdot d)$$



O **axioma de Kock–Lawvere** afirma, portanto, que todo morfismo $f : D \to R$ é essencialmente uma "função afim", admitindo um único coeficiente linear, $f(0)$, e um

único coeficiente angular, b. Deste modo, é comum nos referirmos coloquialmente ao **axioma de Kock–Lawvere** como o **"princípio da microafinidade"**.

Anders Kock descreve este axioma de um modo particularmente pitoresco: D é tão "pequeno" que não se pode distinguir o gráfico de uma função de D em R de um segmento de reta e, ao mesmo tempo, é "grande" o suficiente para que o coeficiente angular desta reta seja unicamente determinado.

Sabemos, da Álgebra, que em um anel R não é possível qualquer elemento ser simultaneamente nilpotente e invertível, de modo que se $a, b \in R$ e $d \in D$ forem tais que $a \cdot d = b \cdot d$, não podemos concluir daí que $a = b$. O **axioma de Kock–Lawvere** implica, não obstante, uma espécie de "lei do cancelamento" para infinitésimos universalmente quantificados. Veja a seguinte:

Proposição 2. *Em* $(\mathbf{S}, (R, +, \cdot, -, 0, e))$ *vale a seguinte dedução, para quaisquer* $a, b \in R$*:*

$$\frac{(\forall d \in D)(a \cdot d = b \cdot d)}{a = b}$$

Demonstração. Consideremos a função $f(d) = a \cdot d$, que é, por hipótese, $f(d) = b \cdot d$. Do **axioma de Kock–Lawvere**, existe um único $c \in R$ tal que $f(d) = f(0) + c \cdot d = c \cdot d$. Como c é o *único* com esta propriedade e tanto a quanto b a satisfazem, segue que $a = c = b$. □

Quanto à estrutura algébrica de D, note que D *não* é fechado pela soma, uma vez que dados $d_1, d_2 \in D$, não é necessário que $(d_1 + d_2) \in D$ (de fato, $(d_1 + d_2)^2 = d_1^2 + 2 \cdot d_1 \cdot d_2 + d_2^2 = 2d_1 \cdot d_2$ e não há nenhum motivo para afirmarmos que $d_1 \cdot d_2 = 0$ [note que convencionamos munir R da estrutura de $\mathbb{Q}-$álgebra]). No que diz respeito à relação entre os "elementos" de D e os de R, note que D "absorve" produtos, ou seja, dados $b \in R, d \in D$, tem-se $b \cdot d \in D$, uma vez que $(b \cdot d)^2 = b^2 \cdot d^2 = b^2 \cdot 0 = 0$.

O **axioma de Kock–Lawvere** nos permite definir facilmente o que entendemos pela derivada de um mapa $f : R \to R$ em um ponto $a \in R$ sem fazer qualquer menção à noção de limite:

Definição 3 (derivada). *Em* $(\mathbf{S}, (R, +, \cdot, -, 0, e))$*, sejam* $f : R \to R$ *uma função,* $a \in R$*, e considere:*

$$\begin{array}{rccc} g_a : & D & \to & R \\ & d & \mapsto & f(a + d). \end{array}$$

Pelo ***axioma de Kock–Lawvere****, existe um único* $b \in R$ *tal que:*

$$(\forall d \in D)(g_a(d) = g_a(0) + b \cdot d),$$

ou seja, tal que:

$$(\forall d \in D)(f(a + d) = f(a) + b \cdot d).$$

Denotamos o b *assim obtido por* $f'(a)$ *e o denominamos* ***a derivada de*** f ***em*** a*. Podemos, portanto, escrever:*

$$(\forall d \in D)(f(a + d) = f(a) + f'(a) \cdot d).$$

A ideia básica por detrás da definição acima é tomar, para cada $a \in R$, $f'(a)$ como a parte linear de g_a. Por exemplo, para $f(x) = x^3$, tem-se $g_a(d) = f(a+d) = (a+d)^3 = a^3 + 3 \cdot a^2 \cdot d + 3 \cdot a \cdot d^2 + d^3 = a^3 + 3 \cdot a^2 \cdot d = f(a) + 3 \cdot a^2 \cdot d$, e como:

$$(\forall d \in D)(f(a+d) = a^3 + 3 \cdot a^2 \cdot d),$$

segue que $f'(a) = 3 \cdot a^2$, coincidindo com a derivada "clássica" de $f(x) = x^3$ em $x = a$.

Sabendo calcular a derivada de uma função $f : R \to R$ em qualquer ponto $a \in R$, podemos definir naturalmente sua "*função* derivada", como segue.

Definição 4. *Em* $(\mathbf{S}, (R, +, \cdot, -, 0, e))$, *dada qualquer* $f : R \to R$, *a* ***função derivada de*** f *é a função:*

$$\begin{array}{rccc} f' : & R & \to & R \\ & x & \mapsto & f'(x). \end{array}$$

Assim, retomando o exemplo dado anteriormente, para $f(x) = x^3$ temos $f'(x) = 3 \cdot x^2$.

O leitor poderá verificar que, utilizando o **axioma de Kock–Lawvere**, as demonstrações das propriedades básicas das derivadas (regras da soma, do produto, da cadeia e outras) se reduzem a simples manipulações algébricas.

Observe que ao assumirmos o **axioma de Kock–Lawvere**, *toda função* $f : R \to R$ *se torna infinitamente diferenciável*, uma vez que o processo dado na **Definição 3** acima pode ser repetido tantas vezes quanto quisermos.

Vejamos, na sequência, como alguns conceitos da Geometria Diferencial se tornam simples no contexto sintético:

Definição 5 (**vetor tangente**). *Em* $(\mathbf{S}, (R, +, \cdot, -, 0, e))$, *sejam* M *um espaço suave e* $x \in M$. *Um* ***vetor tangente a*** M ***em*** x *é um morfismo* $t : D \to M$ *tal que* $t(0) = x$.

Sendo assim, um vetor tangente a M em $x \in M$ pode ser concebido como um "caminho infinitesimal" em M tal que $t(0) = x$, dispensando-nos de todo o trabalho de tomar classes de equivalências de certas curvas contendo o ponto x. Dito de outro modo, os vetores tangentes a M em x são simplesmente *micro-caminhos em* M *que contêm o ponto-base* x. Mais geralmente, definimos o **morfismo ponto-base** por:

$$\begin{array}{rccc} \pi : & TM = M^D & \to & M \\ & D \xrightarrow{t} R & \mapsto & t(0) = x, \end{array}$$

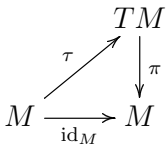
o que nos permite enunciar a seguinte:

Definição 6 (**espaço tangente**). *Sejam* M *um espaço suave,* $x \in M$ *e* $\pi : TM \to M$ *o morfismo ponto-base* x. *O espaço tangente a* M *em* x *é* $T_xM = \pi^{-1}[\{x\}]$.

Na Geometria Diferencial Sintética temos a seguinte:

Definição 7 (**campo vetorial tangente**). *Um* ***campo vetorial tangente a*** M *é um morfismo* $\tau : M \to TM = M^D$ *tal que para todo* $x \in M$, $\tau(x)(0) = x$, *ou seja, tal que* $\pi \circ \tau = \mathrm{id}_M$ *(uma seção global do morfismo ponto-base). Em outros termos,*

um campo vetorial tangente ao espaço suave M é um morfismo $\tau : M \to TM$ tal que o seguinte diagrama comuta:



Munidos destes conceitos de simples formulação,podemos definir também a versão sintética de $k-$forma diferencial, com ajuda da seguinte:

Definição 8 ($k-$**tangente**)**.** *Para $k \in \mathbb{N}$, um morfismo $\tau : D^k \to M$, onde $D^k = D \times \cdots \times D$, é um $k-$**tangente em** M. O espaço dos $k-$tangentes é denotado, portanto, por M^{D^k}.*

No contexto sintético, uma $k-$**forma diferencial em** M com valores em um $R-$módulo euclidiano E é simplesmente um morfismo:

$$\omega : M^{D^k} \to E,$$

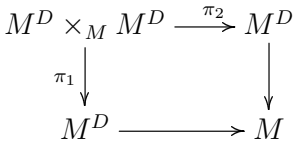
satisfazendo certos axiomas.

Outro conceito fundamental em Geometria Diferencial é o de "conexão afim", que é, a grosso modo, um modo de identificar localmente vetores tangentes a uma variedade em pontos diferentes (porém próximos) de uma curva. Formalmente, no contexto sintético, temos a seguinte:

Definição 9 (**conexão afim**)**.** *Seja M um espaço suave. Uma **conexão afim** no fibrado tangente $TM = M^D$ é um morfismo:*

$$\nabla : M^D \times_M M^D \to M^{D \times D} \tag{2}$$

onde $M^D \times_M M^D$ é tal que o diagrama abaixo é um pullback:



tal que para todo $(t_1, t_2) \in M^D \times_M M^D$ e para quaisquer $d_1, d_2 \in D$ e $\lambda \in R$ tem-se:

(CA1) $\nabla(t_1, t_2)(d_1, 0) = t_1(d_1)$;
(CA2) $\nabla(t_1, t_2)(0, d_2) = t_2(d_2)$;
(CA3) $\nabla(\lambda \cdot t_1, t_2)(d_1, d_2) = \nabla(t_1, t_2)(\lambda \cdot d_1, d_2)$;
(CA4) $\nabla(t_1, \lambda \cdot t_2)(d_1, d_2) = \nabla(t_1, t_2)(d_1, \lambda \cdot d_2)$.

Note que esta definição de conexão é uma corporificação bastante fiel à ideia original de Élie Cartan:

"Uma variedade com uma conexão afim é uma variedade que, na vizinhança imediata de qualquer um de seus pontos, manifesta todas as propriedades de um espaço afim e na qual se tem uma lei que relaciona duas vizinhanças infinitesimalmente próximas".([Car] *apud* [Kos]*.)*

De fato, no contexto sintético podemos formalizar de modo coerente os termos utilizados por Cartan, sendo fiéis à sua ideia original. A presença de objetos infinitesimais em **S** nos permite formalizar o que entendemos por pontos "infinitesimalmente próximos", mediante uma relação reflexiva e simétrica, $\sim$. Por exemplo, em R podemos dizer que dois pontos $x, y \in R$ estão infinitesimalmente próximos, $x \sim y$, se, e somente se, $x - y \in D$. Desta forma, dado um espaço suave qualquer M, munido de uma relação de "proximidade infinitesimal" $\sim$, uma conexão no fibrado tangente $TM = M^D$ é simplesmente um morfismo que associa ao par (x, y) com $x \sim y$, um morfismo $\nabla(x, y) : T_xM \to T_yM$ tal que:

$$\nabla(x, x) = \nabla(y, x) \circ \nabla(x, y) = \mathrm{id}_{T_xM}.$$

A partir deste conceito seguem-se muitos outros, como o de "transporte paralelo", em termos igualmente claros.

Ainda outros conceitos da Geometria Diferencial se exprimem de modo particularmente simples em **S**. Veja a seguinte:

Definição 10 (**microfluxo**). *Seja M um espaço suave. Um* ***microfluxo em*** *M é um morfismo:*

$$\phi : M \times D \to M$$

Além da evidente clareza desses conceitos expressos em termos de construções na categoria dos espaços suaves, sendo **S** uma categoria cartesiano-fechada, há uma correspondência natural entre campos vetoriais e microfluxos, dada por:

$$\frac{\tau : M \to M^D}{\widehat{\tau} : M \times D \to M} : \quad \widehat{\tau}(x, d) = \tau(x)(d)$$

Por sua vez, temos também uma correspondência natural entre microfluxos e micro-caminhos em M^M com ponto-base id_M, dada por :

$$\frac{\widehat{\tau} : M \times D \to M}{\tau^* : D \to M^M} : \quad \tau^*(d)(x) = \widehat{\tau}(x, d) = \tau(x)(d)$$

Desta maneira, em **S**, *campos vetoriais, micro-fluxos e micro-caminhos são noções (naturalmente) equivalentes*, o que em Geometria Diferencial Clássica pode ser dito, na melhor das hipóteses, em sentido metafórico.

As noções de geodésica e todo o aparato da Geometria Riemanniana também podem ser desenvolvidas no contexto da Geometria Diferencial Sintética. Remetemos o leitor interessado a [Bel] e a [Lav].

2. Mundos (modelos) para a Geometria Diferencial Sintética

Conforme discutido anteriormente, a Geometria Diferencial Sintética é formulada no contexto da Teoria das Categorias. Também vimos que, para simplificar a formulação de vários conceitos, faz-se necessário, dentre outras coisas, um "fechamento" da categoria que nos servirá de modelo, **S**, sob diversas construções, como

limites finitos[9] (produtos, espaços de funções, pullbacks etc.), bem como uma adjunção entre produtos e exponenciais que nos permita, por exemplo, "unificar" conceitos como campos vetoriais, microfluxos e micro-caminhos.

Como não se obtém resultados "sofisticados" em Geometria Diferencial sem se operar com variáveis e elementos, é natural desejarmos que **S** se comporte essencialmente como **Sets**, no sentido de que em **S**, para cada objeto, tenhamos uma noção de "elemento" ou "pertinência".

Uma classe importante de categorias que possuem todas essas características é a classe dos *toposes* (ou "topoi"). Estas categorias são, a grosso modo, categorias bastante parecidas com **Sets** sob diversos aspectos. Como os toposes têm objeto terminal (ou seja, um objeto - usualmente denotado por $\mathbf{1}$ - tal que para qualquer objeto A existe um único morfismo $A \xrightarrow{!} \mathbf{1}$), neles podemos formular o que entendemos por "elemento global". Em um topos $\mathcal{E}$, um elemento global b de um objeto $B \in \mathbf{Obj}(\mathcal{E})$ é simplesmente um morfismo $\mathbf{1} \xrightarrow{b} B$, o que corresponde com a identificação mental que fazemos, em **Sets**, entre funções de um conjunto unitário genérico, $\{*\}$, em um conjunto A e seus elementos:

$$\frac{\{*\} \xrightarrow{f} A}{f(*) \in A}$$

Passemos, portanto, à definição de topos:

Definição 11 (**topos**). *Um **topos** é uma categoria $\mathcal{E}$ que satisfaz as seguintes propriedades:*

(I) *$\mathcal{E}$ tem objeto terminal, $\mathbf{1}$, e pullbacks; equivalentemente, $\mathcal{E}$ tem todos os limites finitos;*

(II) *$\mathcal{E}$ é cartesiano-fechada, ou seja, para todo objeto X temos um funtor* exponencial, *$(\cdot)^X : \mathcal{E} \to \mathcal{E}$ que é adjunto à direita do funtor $(\cdot) \times X : \mathcal{E} \to \mathcal{E}$;*

(III) *$\mathcal{E}$ tem classificador de subobjetos, ou seja, um objeto Ω e um morfismo $\mathbf{1} \xrightarrow{\top} \Omega$ tal que para todo monomorfismo $Y \overset{\sigma}{\rightarrowtail} X$ existe um único morfismo $\chi_\sigma : X \to \Omega$ tal que o seguinte diagrama é um pullback:*

$$\begin{array}{ccc} Y & \overset{\sigma}{\rightarrowtail} & X \\ \downarrow{\scriptstyle !} & & \downarrow{\scriptstyle \chi_\sigma} \\ \mathbf{1} & \xrightarrow{\top} & \Omega \end{array}$$

O exemplo mais familiar de topos é **Sets**. Uma classe de grande importância de toposes pode ser obtida considerando-se $\mathbf{Sets}^{\mathcal{C}^{\mathrm{op}}}$, a categoria de todos os funtores de $\mathcal{C}^{\mathrm{op}}$ em **Sets**, onde $\mathcal{C}$ é uma categoria pequena. Outra classe especial de toposes, denominada por "toposes de Grothendieck", é obtida tomando-se feixes sobre um espaço topológico (X, τ) ou, mais geralmente, tomando-se feixes sobre locales ou sítios pequenos.

[9] que, inclusive, nos permita definir estruturas como a de "anel comutativo com unidade" em seu bojo.

Conforme já esperávamos, as simplificações proporcionadas pela robustez estrutural de **S** não vêm gratuitamente. Vimos, por exemplo, que **S** deve conter um objeto R, munido de uma estrutura de anel comutativo com unidade, que satisfaça o **axioma de Kock–Lawvere**. Abaixo enunciamos o sentido exato do que entendemos por isso:

Definição 12. *Sejam $\mathcal{C}$ uma categoria com todos os limites finitos e $(R, +, \cdot, -, 0, e)$ um objeto anel comutativo com unidade em $\mathcal{C}$, munido do subobjeto $D = [\![d.R \mid d^2 = 0]\!] \rightarrowtail R$. Dizemos que R **satisfaz o axioma de Kock–Lawvere** se, e somente se, para qualquer $f : D \to R$ existir um único $b \in R$ tal que:*

$$(\forall d \in D)(f(d) = f(0) + b \cdot d).$$

A proposição a seguir nos dá uma formulação do **axioma de Kock–Lawvere** unicamente em termos de isomorfismos, o que torna sua verificação particularmente mais simples em categorias de pré-feixes.

Proposição 13 (axioma de Kock–Lawvere, 2ª versão). *Seja $\mathcal{C}$ uma categoria com todos os limites finitos e suponha que $(R, +, \cdot, -, 0, e)$ seja um objeto anel comutativo com unidade. Seja, também, $D = [\![d.R \mid d^2 = 0]\!] \rightarrowtail R$ o seu subobjeto de quadrados nilpotentes. Então R satisfaz o **axioma de Kock–Lawvere** se, e somente se:*

$$\begin{array}{rccc} \alpha: & R \times R & \to & R^D \\ & (x, y) & \mapsto & \alpha(x,y): \begin{array}{ccc} D & \to & R \\ d & \mapsto & x + y \cdot d \end{array} \end{array}$$

é um isomorfismo, ou seja, se, e somente se $R \times R \cong R^D$.

Objetos anéis que satisfazem o axioma de Kock–Lawvere são tão especiais que recebem uma denominação própria: **anéis do tipo linha**. Com essas definições em mãos, podemos enunciar o que, exatamente, entendemos por um "mundo" (modelo) onde podemos interpretar a Geometria Diferencial Sintética, ou seja, que tipo de categoria **S** deve ser.

Definição 14. *Um **modelo** para a Geometria Diferencial Sintética*[10] *é um par $(\mathcal{E}, R)$, onde:*

(I) *$\mathcal{E}$ é um topos, de modo que em $\mathcal{E}$ podemos trabalhar "como se" estivéssemos em **Sets**, ao menos no que diz respeito à noção de "elemento";*

(II) *R é um objeto anel comutativo do tipo linha em $\mathcal{E}$;*

(III) *O subobjeto $D = [\![d.R \mid d^2 = 0]\!] \rightarrowtail R$ (que pode ser visto como o equalizador dos morfismos $0 : D \to R$ e $(\cdot)^2 : D \to R$) é exponenciável - ou seja, para todo objeto $A \in \mathrm{Obj}(\mathcal{E})$ tem-se $A^D \in \mathrm{Obj}(\mathcal{E})$;*

(IV) *O funtor $\mathrm{Hom}(D, \cdot) : \mathcal{E} \to \mathcal{E}$ preserva* todos *os limites indutivos.*

Uma questão natural que nos ocorre é: quais são os toposes munidos de objeto anel do tipo linha que satisfazem o **axioma de Kock–Lawvere**? Na sequência analisaremos este tema, exibindo primeiramente um "mundo" onde a admissão do **axioma de Kock–Lawvere** nos conduz à trivialização - e portanto inexistência, pois requeremos $e \neq 0$ - do anel do tipo linha.

[10] cf. p. 231,232 de [**Dub**].

2.1. **Incompatibilidade com a Lógica Clássica e com o "mundo" dos Conjuntos.** Comentamos, de passagem, que **Sets** é um topos, de modo que é legítimo que nos perguntemos se a categoria dos conjuntos, onde podemos definir anel do tipo linha, serve de modelo para a Geometria Diferencial Sintética. Veremos, a seguir, um resultado que exclui **Sets** de nosso rol de possíveis modelos.

Proposição 15 (**Sets não é modelo para a Geometria Diferencial Sintética**)**.** *Ao considerarmos o topos dos conjuntos,* ***Sets****, se* $(R, +, \cdot, -, 0, e)$ *satisfaz o* ***axioma de Kock–Lawvere****, então* $R = \{0\}$.

Demonstração. Consideremos a seguinte função:

$$\begin{array}{rccc} f: & D & \to & R \\ & d & \mapsto & \begin{cases} 0, & \text{se } d = 0 \\ e, & \text{se } d \neq 0 \end{cases} \end{array}$$

Como R satisfaz o **axioma de Kock–Lawvere**, existe um único $b \in R$ tal que:

$$(\forall d \in D)(f(d) = f(0) + b \cdot d).$$

Se existisse $\delta \in D$ tal que $\delta \neq 0$, teríamos:

$$f(\delta) = f(0) + \delta \cdot b$$

$$e = \delta \cdot b.$$

Elevando os dois membros da igualdade acima ao quadrado, teríamos:

$$e = e^2 = (\delta \cdot b)^2 = \delta^2 \cdot b^2 = 0 \cdot b^2 = 0,$$

de modo que $R = \{0\}$. □

Na demonstração da **Proposição 15**, fizemos uso do fato que, em **Sets**, podemos decompor D como reunião de dois subconjuntos, um consistindo dos elementos que são 0, e seu complementar em D, consistindo dos elementos que não são zero, cuja reunião é todo D (não há uma terceira possibilidade). Desta forma, sempre que pudermos decompor nosso objeto de quadrados nilpotentes, $D \rightarrowtail R$, desta forma[11], ao satisfazer o **axioma de Kock–Lawvere** teremos, fatalmente, $R = \{0\}$.

Uma vez que em **Sets**, para qualquer conjunto D, $\mathrm{Sub}(D)$ é uma álgebra de Boole, a categoria dos conjuntos não nos fornecerá um "bom" modelo para o axioma de Kock–Lawvere. Devemos, assim, buscar outra categoria da qual D seja um objeto em que $\mathrm{Sub}(D)$ *não* seja uma álgebra de Boole. Desejamos assim que em **S**, para todo objeto D, tenha-se em $\mathrm{Sub}(D)$ a estrutura de Álgebra de Heyting,[12] que é uma estrutura mais geral do que a de álgebra de Boole no que diz respeito a "complementos". Mas que tipo de categoria podemos tomar com esta propriedade? Veremos a seguir um exemplo.

[11] ou seja, sempre que $\mathrm{Sub}(D)$ for uma álgebra de Boole.

[12] Um bom exemplo de álgebra de Heyting é a álgebra de abertos de um espaço topológico (X, τ). Intuitivamente, podemos "pensar" nos subobjetos de X não como a totalidade de seus subconjuntos, mas como a totalidade de seus subconjuntos abertos, ou seja, $\mathrm{Sub}(X) = \{U \in \tau \mid U \subseteq X\}$. Assim como "álgebras de Boole"são a contraparte algébrica da Lógica Clássica, as "álgebras de Heyting" são a contraparte algébrica da Lógica Intuicionista (ou "construtivista").

Consideremos, dada uma categoria pequena $\mathcal{C}$, a categoria $\mathbf{Sets}^{\mathcal{C}^{\mathrm{op}}}$ dos funtores (também denominados "pré-feixes") de $\mathcal{C}^{\mathrm{op}}$ - a categoria oposta de $\mathcal{C}$ - em **Sets**. Dado qualquer $D : \mathcal{C}^{\mathrm{op}} \to \mathbf{Sets}$, $\mathrm{Sub}\,(D) = \{S \in \mathbf{Sets}^{\mathcal{C}^{\mathrm{op}}} \mid S \text{ é subfuntor de } D\}$ tem a estrutura de álgebra de Heyting dada por:

- o "maior" subfuntor de D é o próprio D;
- o "menor" subfuntor de D é $0 : \mathcal{C}^{\mathrm{op}} \to \mathbf{Sets}$, dado por $0(C) = \varnothing$ para todo $C \in \mathbf{Obj}\,(\mathcal{C})$;
- dados S e T, subfuntores de D, tem-se para qualquer $C \in \mathbf{Obj}\,(\mathcal{C})$, $(S \vee T)(C) = S(C) \cup T(C)$ e $(S \wedge T)(C) = S(C) \cap T(C)$;
- se S é um subfuntor de D, então para qualquer $C \in \mathbf{Obj}\,(\mathcal{C})$, tem-se:

$$(\neg S)(C) = \{x \in D(C) \mid (\forall f : B \to C)(f(x) \notin S(B))\}$$

- dados S e T, subfuntores de D, para qualquer $C \in \mathbf{Obj}\,(\mathcal{C})$, tem-se:

$$(S \Rightarrow T)(C) = \\ = \{x \in D(C) \mid (\forall f : B \to C)(f(x) \in S(B) \Rightarrow f(x) \in T(C))\}.$$

Neste nosso caso, verifica-se que, em geral:

$$(S \vee \neg S)(C) \neq D(C)$$

(tome, por exemplo, $\mathcal{C} = \mathbf{2}$ como a categoria discreta de dois objetos, ou seja, tal que $\mathbf{Obj}\,(\mathbf{2}) = \{0, 1\}$ e $\mathbf{Mor}\,(\mathbf{2}) = \{\mathrm{id}_0 : 0 \to 0, \mathrm{id}_1 : 1 \to 1\}$). Esta condição nos desapossaria do princípio do terceiro excluído usado na demonstração da **Proposição 15** , nos permitindo admitir o **Axioma de Kock–Lawvere** sem que, fatalmente, se tenha $R = \{0\}$.

Vimos acima que, em categorias da forma $\mathbf{Sets}^{\mathcal{C}^{\mathrm{op}}}$, onde $\mathcal{C}$ é uma categoria pequena, não é obrigatório que $\mathrm{Sub}\,(D)$ seja uma álgebra de Boole, o que sugere que categorias de pré-feixes sejam bons candidatos a "mundos" onde podemos interpretar o **axioma de Kock–Lawvere** para R e D sem nos reduzirmos a $R = \{0\}$. Categorias de pré-feixes também possuem a (grande) vantagem de serem toposes, de modo a serem cartesiano-fechadas e possuírem uma noção interna de "elemento" e "pertencimento". Por terem todos os limites finitos, nessas categorias podemos definir estruturas algébricas como anéis, objeto de quadrado-nilpotentes dentre muitas outras.

Uma boa questão a se colocar neste momento é: por qual(quais) categoria(s) podemos substituir $\mathcal{C}$ de modo a poder interpretar, do melhor modo possível, a Geometria Diferencial Sintética em $\mathbf{Sets}^{\mathcal{C}^{\mathrm{op}}}$? Veremos, na sequência, duas categorias de pré-feixes que validam o axioma de Kock–Lawvere. Posteriormente, faremos algumas restrições aos pré-feixes aceitos como objetos de nossa categoria a fim de obter um modelo com as características adequadas ao nosso propósito.

2.2. **Um topos que valida o Axioma de Kock–Lawvere.** O primeiro exemplo de um topos com um anel do tipo linha, R, é um bem simples - e na verdade simples demais para ser realmente útil em Geometria Diferencial.

Seja $\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}$ a categoria cujos objetos são as $\mathbb{R}-$álgebras[13] comutativas finitamente geradas, ou seja, quocientes de $\mathbb{R}-$álgebras livres finitamente geradas por algum de seus ideais [14]:

$$\mathbf{Obj}\left(\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}\right) = \\ = \left\{ \frac{\mathbb{R}[X_1,\cdots,X_n]}{I} \mid (n\in\mathbb{N})\&(I\subseteq\mathbb{R}[X_1,\cdots,X_n] \text{ ideal}) \right\}.$$

Considere, agora, o topos de pré-feixes $\mathbf{Set}^{\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}}$[15] e tome, como objeto anel do tipo linha, o funtor esquecimento:

$$\begin{array}{rccc} R: & \mathbb{R}-\mathbf{Alg}_{\text{f.g.}} & \to & \mathbf{Set} \\ & A & \mapsto & |A| \\ & A\overset{h}{\to}B & \mapsto & |A|\overset{|h|}{\to}|B| \end{array}$$

que a cada $\mathbb{R}-$álgebra finitamente gerada associa seu conjunto subjacente, e a cada homomorfismo de $\mathbb{R}-$álgebras associa a função subjacente.

Proposição 16. *Em* $\mathbf{Set}^{\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}}$, $R:\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}\to\mathbf{Set}$ *é um anel comutativo com unidade.*

Demonstração. Demonstrar isto significa mostrar que existem transformações naturais $+,\cdot : R\times R\Rightarrow R$ e "elementos" $0,e:\mathbf{1}\Rightarrow R$, onde $\mathbf{1}=\{0\}$, satisfazendo os diagramas comutativos correspondentes à estrutura de anel comutativo com unidade.

Definimos, em cada estágio ($\mathbb{R}-$álgebra) $(A,+,\cdot,-,0,e)$:

$$\begin{array}{rccc} +_A: & R(A)\times R(A)=A\times A & \to & R(A)=A \\ & (a,b) & \mapsto & a+b \end{array}$$

$$\begin{array}{rccc} \cdot_A: & R(A)\times R(A)=A\times A & \to & R(A)=A \\ & (a,b) & \mapsto & a\cdot b \end{array}$$

$$\begin{array}{rccc} -_A: & R(A)=A & \to & R(A)=A \\ & a & \mapsto & -a \end{array}$$

$$\begin{array}{rccc} 0_A: & R(\mathbf{1})=\{0\} & \to & R(A)=A \\ & 0 & \mapsto & 0 \end{array}$$

$$\begin{array}{rccc} e_A: & R(\mathbf{1})=\{0\} & \to & R(A)=A \\ & 0 & \mapsto & 1 \end{array}$$

A naturalidade destas transformações segue da própria definição de homomorfismo de $\mathbb{R}-$álgebras.

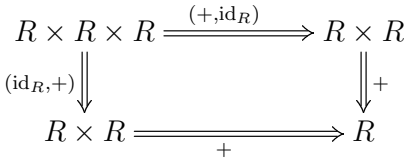
A comutatividade dos diagramas a seguir expressam:

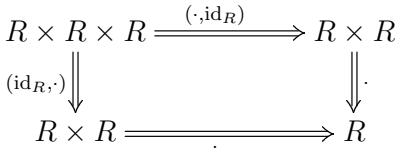
[13] uma $\mathbb{R}-$álgebra é, a grosso modo, um $\mathbb{R}-$espaço vetorial A munido de um produto $\cdot : A\times A\to A$ bilinear.

[14] observe que esta categoria é *pequena*.

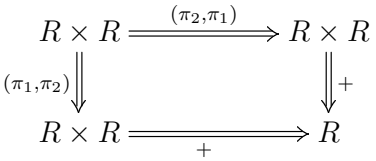
[15] os objetos desta categoria são todos os funtores de $\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}$ em **Set**, e os morfismos são as transformações naturais entre esses funtores.

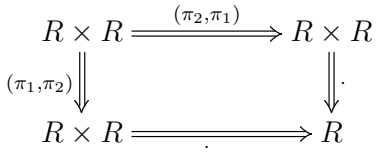
- **Associatividade da soma e do produto, respectivamente:**



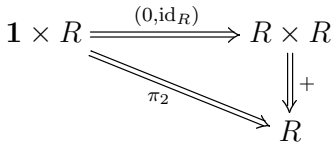


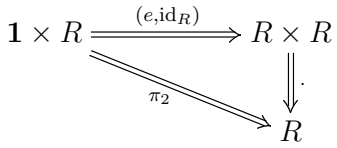
- **Comutatividade da soma e do produto, respectivamente:**



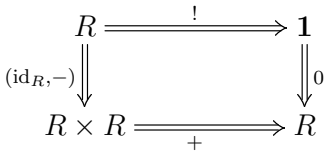


- **Elemento neutro da soma e do produto, respectivamente:**

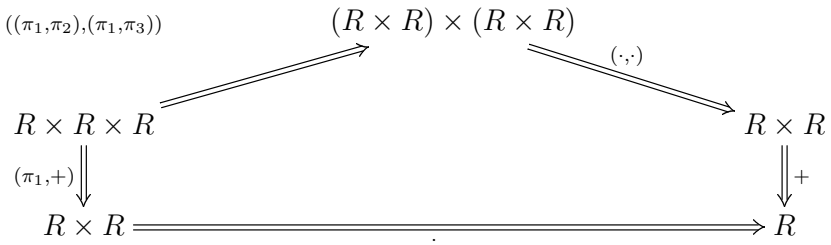




- **Elemento simétrico da soma:**



- **Distributividade do produto sobre a soma:**



onde $\pi_1, \pi_2, \pi_3 : R \times R \times R \to R$ são as projeções na primeira, segunda e terceira coordenada, respectivamente. □

Proposição 17. *No contexto acima, o funtor R é representável.*

Demonstração. Dada qualquer $\mathbb{R}-$álgebra finitamente gerada, A, temos um isomorfismo natural:

$$R(A) \cong \mathrm{Hom}_{\mathbb{R}-\mathbf{Alg}_{\mathrm{f.g.}}}(\mathbb{R}[X], A)$$

que associa a cada $a \in R(A)$ o único $\mathbb{R}-$homomorfismo:

$$\begin{array}{rccc} \phi_a: & \mathbb{R}[X] & \to & A \quad \text{tal que} \\ & X & \mapsto & a \end{array}$$

□

Desta forma, neste topos, R é um funtor representável por $\mathbb{R}[X]$, *i.e.*,

$$R \cong \mathrm{Hom}_{\mathbb{R}-\mathbf{Alg}_{\mathrm{f.g.}}}(\mathbb{R}[X], \bullet).$$

Nosso objeto de quadrados nilpotentes nesta categoria é o subfuntor de R dado por:

$$\begin{array}{rccc} D: & \mathbb{R}-\mathbf{Alg}_{\mathrm{f.g.}} & \to & \mathbf{Set} \\ & A & \mapsto & \{a \in A \mid a \cdot a = 0\} \\ & A \xrightarrow{f} B & \mapsto & \{a \in A \mid a^2 = 0\} \overset{|f|\restriction_{\{a\in A|a^2=0\}}}{\to} \{b \in B \mid b^2 = 0\} \end{array}$$

que é representado pela $\mathbb{R}-$álgebra dos números duais, $\mathbb{R}[\varepsilon] = (\mathbb{R}\times\mathbb{R}, +, \cdot)$, onde:

$$\begin{array}{rccc} +: & (\mathbb{R}\times\mathbb{R})\times(\mathbb{R}\times\mathbb{R}) & \to & \mathbb{R}\times\mathbb{R} \\ & ((a,b),(c,d)) & \mapsto & (a+c, b+d) \end{array}$$

$$\begin{array}{rccc} \cdot: & (\mathbb{R}\times\mathbb{R})\times(\mathbb{R}\times\mathbb{R}) & \to & \mathbb{R}\times\mathbb{R} \\ & ((a,b),(c,d)) & \mapsto & (a\cdot c, a\cdot d + b\cdot c) \end{array}$$

Isto significa que:

$$D \cong \mathrm{Hom}_{\mathbb{R}-\mathbf{Alg}_{\mathrm{f.g.}}}(\mathbb{R}[\varepsilon], \bullet)\,.$$

Recorde que a segunda versão do **axioma de Kock–Lawvere** nos diz que $R\times R \cong R^D$, de modo que para interpretá-lo em nosso topos precisamos explicitar o que entendemos por R^D e por $R \times R$. Definimos o produto de dois funtores R e S por:

$$\begin{array}{rccc} R\times S: & \mathbb{R}-\mathbf{Alg}_{\mathrm{f.g.}} & \to & \mathbf{Set} \\ & A & \mapsto & R(A)\times S(A) \\ & A \xrightarrow{f} B & \mapsto & R(A)\times S(A) \overset{(R(f),S(f))}{\to} R(B)\times S(B) \end{array}$$

Tendo a definição de produto de dois funtores, em qualquer topos de pré-feixes podemos definir:

$$\begin{array}{rccc} R^D: & \mathbb{R}-\mathbf{Alg}_{\mathrm{f.g.}} & \to & \mathbf{Set} \\ & A & \mapsto & \mathrm{Nat}\left(\mathrm{Hom}_{\mathbb{R}-\mathbf{Alg}_{\mathrm{f.g.}}}(A,\bullet)\times D, R\right). \end{array}$$

Teorema 18. *R satisfaz o **axioma de Kock–Lawvere**.*

Demonstração. De fato, para qualquer $\mathbb{R}-$álgebra finitamente gerada A, temos:

$$\begin{aligned} R^D(A) &\overset{\text{def.}}{=} \mathrm{Nat}(\mathrm{Hom}_{\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}}(A,\bullet)\times D, R) \overset{\textbf{Prp. 17}}{=} \\ &= \mathrm{Nat}\left(\mathrm{Hom}\left(A,\bullet\right)\times \mathrm{Hom}\left(\mathbb{R}[\varepsilon],\bullet\right), \mathrm{Hom}\left(\mathbb{R}[X],\bullet\right)\right) \cong \\ &\cong \mathrm{Nat}\left(\mathrm{Hom}_{\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}}(A\otimes\mathbb{R}[\varepsilon],\bullet), \mathrm{Hom}_{\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}}(\mathbb{R}[X],\bullet)\right) \overset{\text{Yoneda}}{\cong} \\ &\cong \left|\mathrm{Hom}_{\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}}(\mathbb{R}[X], A\otimes\mathbb{R}[\varepsilon])\right| = |A\otimes\mathbb{R}[\varepsilon]| \overset{\text{como conjunto}}{=} A\times A = \\ &= (R\times R)(A), \end{aligned}$$

ou seja, $R^D \overset{\alpha}{\cong} R\times R$. □

Assim, $(\mathbf{Set}^{\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}}, R)$, onde R é o funtor esquecimento, é um topos que satisfaz o **axioma de Kock–Lawvere** e, portanto, neste sentido, um "modelo" para a Geometria Diferencial Sintética. No entanto, pode-se verificar que nosso "modelo da reta geométrica", R, neste topos não é sequer um anel local. Não temos disponível, tampouco, um objeto neste topos que represente o intervalo $[0,1]$, imprescindível para desenvolver a teoria de integração. Além disto, este topos claramente não contém as variedades diferenciáveis (apenas as variedades algébricas, representadas por seus anéis de coordenadas), de modo que não é possível se "fazer" Geometria Diferencial nele. Convém, portanto, buscar outra categoria na qual além de valer o **Axioma de Kock–Lawvere**, possamos codificar as variedades diferenciáveis suaves.

2.3. **Um "Toy Model" para a Geometria Diferencial Sintética.** Conforme prometido no título deste trabalho, a busca por modelos da Geometria Diferencial Sintética nos motiva a estudar os anéis $\mathcal{C}^\infty$ [16]. Nesta seção definiremos, funtorialmente, o que é um anel $\mathcal{C}^\infty$ e daremos, em seguida, uma definição da categoria dos anéis $\mathcal{C}^\infty$ finitamente gerados, cujos objetos talvez sejam mais familiares para o leitor. Citaremos alguns exemplos destes anéis e exibiremos um topos munido de um objeto anel do tipo linha que, além de satisfazer o **axioma de Kock–Lawvere**, contém uma "cópia" de da categoria das variedades diferenciáveis suaves.

Definição 19 (**anel $\mathcal{C}^\infty$**). *Seja $\mathcal{C}^\infty$ a categoria cujos objetos são as potências cartesianas de $\mathbb{R}$, ou seja,* $\mathbf{Obj}\left(\mathcal{C}^\infty\right) = \{\{*\}, \mathbb{R}, \mathbb{R}^2, \cdots, \mathbb{R}^n, \cdots\}$*, e cujos morfismos são as funções suaves (isto é, de classe $\mathcal{C}^\infty$) entre esses objetos:*

$$\mathrm{Hom}_{\mathcal{C}^\infty}\left(\mathbb{R}^m, \mathbb{R}^n\right) = \{\mathbb{R}^m \xrightarrow{f} \mathbb{R}^n \mid f \text{ é uma função suave}\} = \mathcal{C}^\infty(\mathbb{R}^m, \mathbb{R}^n)$$

Um anel $\mathcal{C}^\infty$ é um funtor $A : \mathcal{C}^\infty \to \mathbf{Set}$ que preserva produtos finitos. Desta forma, pode-se conceber um anel $\mathcal{C}^\infty$ como uma interpretação, em **Set**, *da teoria das funções suaves.*

Uma vez que todas as funções polinomiais são funções suaves, podemos considerar um anel $\mathcal{C}^\infty$ como uma extensão natural do conceito de $\mathbb{R}-$álgebra, no sentido de interpretar não apenas funções polinomiais mas *todas* as funções suaves.

[16] o leitor interessado encontrará um estudo mais aprofundado sobre esta classe de objetos matemáticos em [Ber].

Definição 20 (**anel $\mathcal{C}^\infty$ finitamente gerado**). *Um anel comutativo com unidade A é um* ***anel $\mathcal{C}^\infty$ finitamente gerado*** *se, e somente se, existirem $n \in \mathbb{N}$ e um ideal $I \subseteq \mathcal{C}^\infty(\mathbb{R}^n)$ tais que:*

$$A \cong \frac{\mathcal{C}^\infty(\mathbb{R}^n)}{I}$$

Os morfismos de anéis $\mathcal{C}^\infty$ finitamente gerados são exatamente os morfismos entre os anéis comutativos unitários subjacentes. Denotamos a categoria de todos os anéis $\mathcal{C}^\infty$ finitamente gerados e seus morfismos por $\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}$.

Dado qualquer aberto $U \subseteq \mathbb{R}^n$, $\mathcal{C}^\infty(U)$ é um anel $\mathcal{C}^\infty$ finitamente gerado. De fato, consideramos uma função característica suave para U, digamos $\chi_U : \mathbb{R}^n \to \mathbb{R}$ (uma função que é diferente de zero exatamente em U) e tomamos $\widehat{U} = \{(x, y) \in \mathbb{R}^n \times \mathbb{R} \mid y \cdot \chi_U(x) = 1\} \subseteq \mathbb{R}^{n+1}$. Verifica-se, então, que $\mathcal{C}^\infty(U) \cong \mathcal{C}^\infty(\mathbb{R}^{n+1})/\langle y \cdot \chi_{\widehat{U}}(x) - 1\rangle$.

Uma outra classe importante de exemplos de anéis $\mathcal{C}^\infty$ finitamente gerados é obtida utilizando-se de variedades diferenciáveis suaves. Dada qualquer variedade suave M, por exemplo, $\mathcal{C}^\infty(M) = \mathcal{C}^\infty(M, \mathbb{R})$ é um anel $\mathcal{C}^\infty$ finitamente gerado. De fato, pelo **Teorema do Mergulho de Whitney** existe $n \in \mathbb{N}$ tal que $M \subseteq \mathbb{R}^n$. Pelo **Teorema da $\varepsilon-$vizinhança** existe um aberto $U \subseteq \mathbb{R}^n$ tal que $M \subseteq U$ e existe uma retração $r : U \twoheadrightarrow M$. Assim, como $\mathcal{C}^\infty(M)$ é um retrato do anel $\mathcal{C}^\infty$ finitamente gerado $\mathcal{C}^\infty(U)$, segue que $\mathcal{C}^\infty(M)$ é finitamente gerado. Um outro importante exemplo de anel $\mathcal{C}^\infty$ finitamente gerado é o dos números duais, $\mathbb{R}[\varepsilon]$, uma vez que $\mathbb{R}[\varepsilon] \cong \mathcal{C}^\infty(\mathbb{R})/\langle x^2\rangle$.

Observa-se que a categoria dos anéis $\mathcal{C}^\infty$ finitamente gerados é capaz de codificar todas as variedades suaves de um modo análogo ao que a categoria das $\mathbb{R}-$álgebras finitamente apresentadas é capaz de codificar variedades algébricas, além de codificar o objeto "infinitesimal" $\mathbb{R}[\varepsilon]$. Deste modo, é justo esperamos obter um modelo melhor para a Geometria Diferencial Sintética tomando o topos de pré-feixes de $\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}$ em **Sets** ao invés de $\mathbf{Sets}^{\mathbb{R}-\mathbf{Alg}_{\text{f.g.}}}$.

O teorema a seguir nos garante que há uma réplica suficientemente bem comportada da categoria das variedades diferenciáveis suaves na categoria dos anéis $\mathcal{C}^\infty$ finitamente gerados. O mergulho de **Man** em $\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}$ é descrito a seguir.

Teorema 21. *O funtor contravariante:*

$$\begin{array}{rccc} m: & \mathbf{Man} & \to & (\mathcal{C}^\infty\mathbf{Rng})^{\text{op}}_{\text{f.g.}} \\ & M & \mapsto & \mathcal{C}^\infty(M) \\ & M \xrightarrow{g} N & \mapsto & \mathcal{C}^\infty(M) \overset{(-\circ g)^{\text{op}}}{\to} \mathcal{C}^\infty(N) \end{array}$$

é pleno, fiel e preserva pullbacks transversais.

Demonstração. A fim de verificar que m aplica variedades suaves sobre anéis $\mathcal{C}^\infty$ finitamente gerados, veja o **Teorema 2.3** da p. 25 de [MR]. Para ver que m aplica pullbacks transversais sobre pushouts, veja a **Proposição 2.6** da p. 27 de [MR]. Para uma demonstração de que m é pleno e fiel, veja o **Teorema 2.8** da p. 30 de [MR]. □

Consideremos, finalmente, o seguinte topos:

$$\mathbf{S} = \mathbf{Set}^{(\mathcal{C}^\infty\mathbf{Rng})_{\text{f.g.}}}$$

e vejamos como mergulhar $\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}$ nele. Ao considerarmos o funtor de Yoneda contravariante:

$$\begin{array}{rccc} Y: & (\mathcal{C}^\infty\mathbf{Rng})^{\text{op}}_{\text{f.g.}} & \to & \mathbf{Set}^{(\mathcal{C}^\infty\mathbf{Rng})_{\text{f.g.}}} \\ & A & \mapsto & \text{Hom}_{(\mathcal{C}^\infty\mathbf{Rng})_{\text{f.g.}}}(A, \bullet) \end{array}$$

obtemos um mergulho contravariante pleno, fiel e que preserva pullbacks transversais. Por conseguinte, a composição dos dois funtores plenos, fiéis e contravariantes:

$$\begin{array}{ccc} & & \mathbf{Sets}^{(\mathcal{C}^\infty\mathbf{Rng})_{\text{f.g.}}} \\ & \overset{s}{\nearrow} & \uparrow{\scriptstyle Y} \\ \mathbf{Man} & \xrightarrow[m]{} & \mathcal{C}^\infty\mathbf{Rng}^{\text{op}}_{\text{f.g.}} \end{array}$$

nos dá um mergulho pleno, fiel e covariante de **Man** em $\mathbf{Sets}^{\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}}$.

Pode-se demonstrar que:

• Em **S** o funtor esquecimento:

$$\begin{array}{rccc} R: & (\mathcal{C}^\infty\mathbf{Rng})_{\text{f.g.}} & \to & \mathbf{Set} \\ & A & \mapsto & |A| \end{array}$$

é representado por $\mathcal{C}^\infty(\mathbb{R})$, ou seja, $R \cong \text{Hom}_{(\mathcal{C}^\infty\mathbf{Rng})_{\text{f.g.}}}(\mathcal{C}^\infty(\mathbb{R}), \bullet)$.

• O funtor:

$$\begin{array}{rccc} D: & (\mathcal{C}^\infty\mathbf{Rng})_{\text{f.g.}} & \to & \mathbf{Set}^{\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}} \\ & A & \mapsto & \text{Hom}_{(\mathcal{C}^\infty\mathbf{Rng})_{\text{f.g.}}}(\mathbb{R}[\varepsilon], A) \end{array}$$

é, em **S**, o objeto de quadrados nilpotentes. Ademais, tem-se:

Teorema 22. *Em* $(\mathbf{Sets}^{\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}}, R)$ *vale o* ***axioma de Kock–Lawvere****.*

Demonstração. Dado qualquer anel $\mathcal{C}^\infty$ finitamente gerado A, tem-se:

$$\begin{aligned} R^D(A) &= \text{Nat}\left(\text{Hom}_{\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}}(A, \bullet) \times D, R\right) = \\ &= \text{Nat}\left(\text{Hom}(A, \bullet) \times \text{Hom}\left(\mathbb{R}[\varepsilon], \bullet\right), \text{Hom}(\mathcal{C}^\infty(\mathbb{R}), \bullet)\right) = \\ &= \left|\text{Hom}_{\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}}\left(\mathcal{C}^\infty(\mathbb{R}), A \otimes_\infty \mathbb{R}[\varepsilon]\right)\right| = \left|A \otimes_\infty \mathbb{R}[\varepsilon]\right| = A \times A = \\ &= (R \times R)(A). \end{aligned}$$

(cf. **Teorema 2.4** da p. 87 de [MR]). □

O topos $(\mathbf{Sets}^{\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}}, R)$, no entanto, é muito "grande" - e tem uma estrutura muito "grosseira", no sentido em que:

- O mergulho $s = Y \circ m : \mathbf{Man} \to \mathbf{S}$ não preserva alguns colimites interessantes, como coberturas abertas;
- O objeto anel do tipo linha, R, não é um anel local;
- O objeto anel do tipo linha (nosso modelo da reta geométrica), R, não é arquimediano.
- Em **S** não temos um objeto que nos valha pelo intervalo unitário $[0, 1]$.

Veremos, na sequência, que ao considerar somente certos objetos de $\mathbf{Sets}^{\mathcal{C}^\infty\mathbf{Rng}_{\text{f.g.}}}$ (feixes sobre um sítio de anéis $\mathcal{C}^\infty$ de um certo tipo especial), podemos obter um topos que preenche todas essas lacunas.

2.4. **Há um mundo bem melhor: o topos de Dubuc.** Consideramos, agora, uma subcategoria da dos pré-feixes de certa classe de anéis $\mathcal{C}^\infty$ sobre **Set**. Este topos é mais sofisticado, por se tratar de um topos de feixes. Usa-se aqui a noção de topologia de Grothendieck (um conceito que generaliza a noção de cobertura aberta de um espaço topológico), que nos permite interpretar fórmulas (geométricas) e conceitos no topos. Referimos, em benefício do leitor, o **Capítulo 8** de [Lav] para um bom estudo sucinto das noções de semântica categorial.

Para construir o topos de Dubuc, precisamos nos restringir a certos anéis $\mathcal{C}^\infty$ finitamente gerados que satisfazem uma propriedade especial, a saber, os anéis $\mathcal{C}^\infty$ germe-determinados (também denominados "de caráter local").

Definição 23 (**anel $\mathcal{C}^\infty$ germe-determinado**). *Seja I um ideal de $\mathcal{C}^\infty(\mathbb{R}^n)$. Dizemos que I é um* ***ideal germe-determinado*** *se, e somente se para todo $f \in \mathcal{C}^\infty(\mathbb{R})$ tivermos:*

$$\left(\forall x \in Z(I) = \bigcap_{g\in I} g^{-1}[\{0\}]\right)(f \restriction_x \in I \restriction_x \to f \in I).$$

A categoria dos anéis $\mathcal{C}^\infty$ (finitamente gerados) da forma $\frac{\mathcal{C}^\infty(\mathbb{R}^n)}{I}$, onde I é um ideal germe-determinado, juntamente com seus morfismos, será denotada por $\mathbb{G}$.

Uma classe importante de exemplos de anéis $\mathcal{C}^\infty$ germe-determinados é a classe dos anéis $\mathcal{C}^\infty$ que são quocientes de $\mathcal{C}^\infty(\mathbb{R}^n)$, para algum $n \in \mathbb{N}$, por ideais finitamente gerados - ou seja, os anéis $\mathcal{C}^\infty$ finitamente apresentados (cf. **Proposição 12** de [Dub1]). Assim, tanto $\mathcal{C}^\infty(\mathbb{R})$ quanto $\mathcal{C}^\infty(\mathbb{R})/\langle x^2\rangle$ são anéis $\mathcal{C}^\infty$ germe-determinados, e como para qualquer variedade diferenciável suave M tem-se $\mathcal{C}^\infty(M)$ finitamente apresentado, segue que $\mathcal{C}^\infty(M)$ é germe-determinado.

No entanto, nem todo ideal germe-determinado é finitamente gerado. De fato, dado qualquer $X \subsetneq \mathbb{R}^n$ fechado, em $\mathcal{C}^\infty(\mathbb{R}^n)$ o ideal:

$$\mathfrak{m}_X^g = \{f \in \mathcal{C}^\infty(\mathbb{R}^n) \mid (\exists U \subseteq \mathbb{R}^n)\&(U \text{ aberto})\&(f \restriction_U \equiv 0)\}$$

embora germe-determinado, não é sequer enumeravelmente gerado. Mais concretamente, $\frac{\mathcal{C}^\infty(\mathbb{R})}{\mathfrak{m}^g_{[0,1]}}$ é um exemplo de anel $\mathcal{C}^\infty$ germe-determinado que *não* é finitamente apresentado.

O ideal $\mathfrak{m}^\infty_{[0,1]} = \{f \in \mathcal{C}^\infty(\mathbb{R}) \mid D^\alpha f \restriction_{[0,1]} = 0 \text{ para todo } \alpha\}$ é, também, germe-determinado, de modo que $\mathcal{C}^\infty(\mathbb{R})/\mathfrak{m}^\infty_{[0,1]}$ é um anel $\mathcal{C}^\infty$ germe-determinado. Este fato nos permite, por assim dizer, conceder "cidadania" ao intervalo $[0,1]$ em nosso topos. Munidos deste objeto em nosso "universo", podemos interpretar o **Axioma da Integração**:

$$(\forall f \in R^{[0,1]})(\exists! g \in R^{[0,1]})((g(0) = 0)\&(g' \cong f)),$$

com o qual se desenvolve uma grande parte da Geometria Diferencial, que engloba, por exemplo, o teorema de Stokes - cuja demonstração no caso sintético se reduz a poucas linhas.

Eduardo J. Dubuc demonstrou, dentre muitas outras coisas, que a classe dos ideais germe-determinados de $\mathcal{C}^\infty(\mathbb{R}^n)$ (que ele chama de "ideais de caráter local") é a maior classe de ideais que satisfaz o $\mathcal{C}^\infty-$**Nullstellensatz** (fraco):

$$Z(I) = \varnothing \iff 1 \in I,$$

o que nos sugere uma forte conexão entre a variedades diferenciáveis e os anéis $\mathcal{C}^\infty$ germe-determinados, semelhante àquela estabelecida pelo **Nullstellensatz fraco**, entre variedades algébricas e $\mathbb{K}-$álgebras reduzidas de tipo finito, onde $\mathbb{K}$ denota um corpo algebricamente fechado.

Deve-se ressaltar, ainda, que se I é um ideal germe-determinado, $I\restriction_U = \{f\restriction_U : U \to \mathbb{R} \mid f \in I\}$ *não é, necessariamente,* germe-determinado: de fato, embora o ideal $\mathfrak{m}^g_{\{0\}}$ seja germe-determinado, considerando $U = \{t \in \mathbb{R} \mid t > 0\}$ obtemos o ideal $\mathfrak{m}^g_{\{0\}}\restriction_U$, que não é germe-determinado (cf. p. 98 de [MR]).

A definição a seguir é importante para a construção do topos de Dubuc.

Definição 24. *Uma família de funções $(f_\alpha : M \to \mathbb{R})_\alpha$ é* ***localmente finita*** *se para todo $x \in M$ existir uma vizinhança U_x tal que $f_\alpha\restriction_{U_x} = 0$ exceto por uma quantidade finita de índices.*

Se I é um ideal de $\mathcal{C}^\infty(\mathbb{R}^n)$, existe um menor ideal germe-determinado, $\widetilde{I}$, denominado o **germe-radical de** I, que contém I. Basta tomarmos:

$$\widetilde{I} = \left\{ \sum_{\alpha\in\Lambda} f_\alpha \mid (f_\alpha)_{\alpha\in\Lambda} \text{ é uma família localmente finita de } I \right\}$$

Observação 25. *Nós precisamos da "germe determinação" dos ideais a fim de que a topologia de Grothendieck do sítio a ser construído seja subcanônica - ou seja, a fim de que os funtores representáveis sejam feixes (e em particular, o objeto anel do tipo linha, que é representável, esteja na categoria, cf.* ***Lema 1.3*** *da p. 101 de* [MR]*), e portanto objetos da categoria de feixes. Veja, por exemplo, p. 102 de* [MR].

Para construir o sítio e, posteriormente o topos de Dubuc, vamos munir $\mathbb{G}^{\text{op}}$ de uma pré-topologia de Grothendieck, *i.e.*, vamos definir, para cada objeto de $\mathbb{G}$ o que entendemos por famílias recobridoras:

Dado um anel $\mathcal{C}^\infty$ germe-determinado, $A = \mathcal{C}^\infty(\mathbb{R}^n)/I$, definimos:

$$\text{Cov}\left(\frac{\mathcal{C}^\infty(\mathbb{R}^n)}{I}\right) = \left\{ \frac{\mathcal{C}^\infty(\mathbb{R}^n)}{I} \xrightarrow{u_\alpha} \frac{\mathcal{C}^\infty(U_\alpha)}{\widetilde{I\restriction_{U_\alpha}}} \mid \{U_\alpha \mid \alpha \in \Lambda\} \in \mathcal{O}(Z(I)) \right\},$$

onde $\mathcal{O}(Z(I))$ denota a classe de todas as coberturas abertas do conjunto $Z(I) \subseteq \mathbb{R}^n$.

Denotando por $J_{\text{Cov}} : \mathbf{Obj}\left(\mathbb{G}^{\text{op}}\right) \to \wp(\wp(\mathbf{Mor}\left(\mathbb{G}^{\text{op}}\right)))$ a topologia de Grothendieck gerada por Cov, obtemos um sítio pequeno, $(\mathbb{G}^{\text{op}}, J_{\text{Cov}})$, onde J_{Cov} é subcanônica - ou seja, para todo anel $\mathcal{C}^\infty$ germe determinado A, o funtor $\text{Hom}_{\mathbb{G}^{\text{op}}}(A, \bullet)$ é feixe.

O **topos de Dubuc** é o topos[17] de feixes sobre o sítio dos anéis $\mathcal{C}^\infty$ germe-determinados:

$$\mathcal{G} = \text{Sh}\left(\mathbb{G}^{\text{op}}, J_{\text{Cov}}\right)$$

[17] para uma prova de que $\mathcal{G}$ é um topos, ver **Proposição 1.5**, da p.102 de [MR].

Sendo J_{Cov} subcanônica, tem-se que $R = \text{Hom}_{\mathbb{G}^{\text{op}}}(\mathcal{C}^\infty(\mathbb{R}), \bullet)$ e $D = \text{Hom}_{\mathbb{G}^{\text{op}}}(\mathbb{R}[\varepsilon], \bullet)$ são feixes e, portanto, estão em $\mathcal{G}$.

Demonstra-se que $\mathcal{G}$ tem as seguintes características:

- existe um mergulho covariante pleno e fiel de **Man** em $\mathcal{G}$, que preserva pullbacks transversais e leva coberturas abertas em famílias recobridoras: (cf. **Corolário 1.4** de [MR])
- $\mathcal{G}$ valida o axioma de Kock–Lawvere e o objeto anel do tipo linha, R, que é representável, é um feixe;
- O objeto anel do tipo linha de $\mathcal{G}$, é um objeto anel local (cf. **Proposição 1.7**, da p. 108 de [MR]) e até mesmo um corpo internalizado, num sentido específico (cf. **Proposição 1.9**, p. 109 de [MR]);
- o objeto anel do tipo linha, R, munido da relação de ordem conveniente em $\mathcal{G}$ é arquimediano.
- Em $\mathcal{G}$, temos disponível o correspondente ao intervalo $[0, 1]$, que é representado por $\mathcal{C}^\infty(\mathbb{R})/\mathfrak{m}^\infty_{[0,1]}$.

3. Considerações Finais

Apesar de terem sido extensivamente utilizados para descobertas e raciocínios sobre diversos conceitos da Geometria, os infinitésimos permaneceram durante muito tempo com o status de "ficções úteis". Cronologicamente, a primeira tentativa bem sucedida de formalizar uma noção de infinitésimo (invertível) foi a de A. Robinson, nos anos 1960, com sua Análise Não-Standard. Foi somente ao final da mesma década que surgiu uma abordagem capaz de "conceder cidadania" aos entes infinitésimos nilpotentes no mundo das "variedades" estudadas pela Geometria Diferencial, capazes de explicitar a intuição por detrás de conceitos intrincados, como conexões, fibrados vetoriais e muitos outros. Os argumentos e intuições de Lie, Cartan, Darboux e muitos outros tornam-se, com a Geometria Diferencial Sintética, finalmente vindicados - o que tem implicações óbvias no ensino da Geometria Diferencial, por exemplo.

Da busca de modelos para a Geometria Diferencial Sintética surge uma motivação para o estudo da teoria algébrica dos anéis $\mathcal{C}^\infty$. A partir deste estudo surgem novos ramos da Matemática, como o estudo de uma Geometria Algébrica sobre anéis $\mathcal{C}^\infty$ (cf. [Joy]), bem como o de sua contraparte algébrica, a Álgebra Comutativa Suave (cf. [BM]). Estes anéis vêm assumindo um protagonismo cada vez maior em novas áreas da Matemática, como a Geometria Diferencial Derivada, bem como em outras partes do conhecimento. Esperamos, com este trabalho, ter contribuído com a familiarização do leitor com a Geometria Diferencial Sintética e com a utilidade dos anéis $\mathcal{C}^\infty$.

Agradecimentos: *Os autores gostariam registrar seus sinceros agradecimentos ao parecerista anônimo pelas excelentes sugestões e pelos comentários os quais este, gentilmente, nos permitiu, através do editor, o professor Daniel Gonçalves, incorporá-los ao presente texto enriquecendo-o grandemente, sobretudo na contextualização do tema apresentada na introdução.*

Referências

[**Bel**] Bell, J. *Two approaches to Modelling the Universe: Synthetic Differential Geometry and Frame-Valued Sets*, 2008. Disponível em http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.114.1930

[**Ber**] Berni, J. C., *Some Algebraic and Logical Aspects of $\mathcal{C}^\infty-$Rings*, tese de doutorado, Instituto de Matemática e Estatística - IME-USP, 2018.

[**BM**] Berni, J. C., Mariano, H.L. *Topics on Smooth Commutative Algebra (preprint)*, 2019, disponível em https://arxiv.org/pdf/1904.02725.pdf.

[**Bus**] Busemann, H. *Metric Methods in Finsler Spaces in the Foundations of Geometry*, Princeton University Press,1942.

[**Car**] Cartan, E. *On Manifolds with Affine Connection and the Theory of General Relativity*, Bibliopolis, Edizioni di Filosofia e Science, Napoli, 1986.

[**Dub1**] Dubuc, E. *$\mathcal{C}^\infty-$Schemes*, American Journal of Mathematics, vol. 103, n. 4, pp. 683-690, 1981.

[**Dub**] Dubuc, E. *Sur les modèles de la Géométrie Différentielle Synthétique*, Cahiers de Topologie et Géométrie Différentielle Catégorique, tome 20, n. 3 (1979), p. 231-279.

[**Gol**] Goldblatt, R. *Topoi - the categorial analysis of logic*, Dover Publications, Inc., Mineola, 2006.

[**End**] Enderton, H. *A Mathematical Introduction to Logic*, Academic Press, Inc., 1972.

[**Joy**] Joyce, D. *Algebraic Geometry over $\mathcal{C}^\infty-$Rings*, 2015, disponível em https://arxiv.org/pdf/1001.0023.pdf.

[**Koc**] Kock, A. *Synthetic Differential Geometry*, London Mathematical Society Lecture Note Series 333, Cambridge University Press, 2006.

[**Kos**] Kostecki, R. *Differential Geometry in Toposes*, disponível em https://www.fuw.edu.pl/~kostecki/sdg.pdf.

[**Law**] Lawvere, F. W. *Outline of Synthetic Differential Geometry*, disponível em http://www.acsu.buffalo.edu/~wlawvere/SDG_Outline.pdf

[**Lam**] Lamers, J. *Smooth Infinitesimal Analysis*, Bachelor Thesis, Utrecht University, 2009.

[**Lav**] Lavendhomme, R. *Basic Concepts of Synthetic Differential Geometry*, Kluwer Texts in Mathematical Sciences, 1996.

[**Mac**] MacLane, S. *Saunders MacLane: a mathematical autobiography*, A. K. Peters Ltd., 2005.

[**MMo**] MacLane, S., Moerdijk, I. *Sheaves in Geometry and Logic - A First Introduction to Topos Theory*, Springer Verlag, 1992.

[**MMR**] Moerdijk, I., Reyes, G. *A Smooth Version of the Zariski Topos.* Advances in Mathematics 65, 229-253 (1987).

[**MR**] Moerdijk, I., Reyes, G. *Models for Smooth Infinitesimal Analysis*, Springer Verlag New York Inc, 1991.

[**Pic**] Picado, J. *An Interview with F. William Lawvere*, disponível em http://www.mat.uc.pt/~picado/lawvere/interview.pdf.

Instituto de Matemática e Estatística, Rua do Matão, 1010, Cidade Universitária, São Paulo - SP

Email address: jeancb@ime.usp.br, hugomar@ime.usp.br

RMU
REVISTA MATEMÁTICA UNIVERSITÁRIA

SBM
SOCIEDADE BRASILEIRA DE MATEMÁTICA

Revista Matemática Universitária, vol. 1, 2022
ISSN: 2675-5254 — **DOI:** 10.21711/26755254/rmu20223

KLEIN NOS AJUDA CONTAR RETAS EM SUPERFÍCIES PROJETIVAS

JACQUELINE ROJAS E SALLY ANDRIA

Resumo. Os grupos de simetrias por rotações dos Sólidos Platônicos reaparecem no *Teorema de classificação dos subgrupos finitos de* $\mathrm{PGL}_2(\mathbb{C})$, dado por *F. Klein* em 1870. Neste artigo veremos como este teorema pode ser utilizado na contagem de retas em certas superfícies projetivas não singulares, seguindo as linhas do artigo *Counting lines on surfaces* de *Boissière & Sarti* publicado em 2007 (cf. Ann. Scuola Norm. Sup. Pisa Cl. Sci. 5).

1. Introdução

Desde a Grécia Antiga até os dias atuais, cinco sólidos capturaram a imaginação da humanidade, desempenhando um papel importante em várias áreas do conhecimento. Há milênios, os geômetras estudaram sua beleza matemática e simetrias únicas. No início da cosmologia, eles também tiveram um papel importante, sendo utilizados pelo astrónomo alemão *Johannes Kepler* (1571-1630) para explorar seu primeiro modelo do sistema solar. No modelo de Kepler, as órbitas dos planetas estariam circunscritas pelos cinco sólidos platônicos, que embora errada, levaria mais adiante a descoberta das órbitas elípticas do sistema solar (cf. p. 76 em [23]). Além disso, na Biologia e na Química também existem reflexos das propriedades e simetrias destes sólidos, por exemplo ao estudar a morfologia de um vírus ou a geometria molecular de diversos elementos (veja [12]).

Em 1884, o matemático alemão *Felix Christian Klein* (1849-1925) explicou em seu livro [11], como expressar as raízes de qualquer polinômio de grau 5 em termos de funções modulares elípticas. A ideia dele foi relacionar o grupo das rotações do icosaedro com o grupo de automorfismos dos pontos de torção de ordem 5 de uma determinada curva elíptica. De fato, ele criou uma teoria que relacionava o grupo das rotações de cada um dos cinco sólidos platônicos com o grupo de automorfismos de certos subconjuntos de pontos de torção.

Data de aceitação: 16 de março de 2021.
Palavras chave. Retas em superfícies. Teorema de classificação de Klein.

Um dos resultados obtidos por *Klein* (em [11]) é o *Teorema de classificação dos subgrupos finitos de* $\mathrm{PGL}_2(\mathbb{C})$, o qual foi usado em 2007 por *Boissière & Sarti* (em [5]) na contagem de retas em superfícies projetivas não singulares.

A referida contagem, surge como uma tentativa de responder, ou mais precisamente, achar cotas inferiores para a clássica questão da área de geometria enumerativa (ainda em aberto para $d \geqslant 5$):

Qual é o número máximo de retas que uma superfície não singular de grau d em $\mathbb{P}^3$ contém?

No qual $\mathbb{P}^3$ é o espaço projetivo sobre o corpo dos números complexos.

Este texto tem como objetivo principal mostrar para o leitor (seguindo as linhas da demonstração da Proposição 3.3 em [5]), como a partir de um subconjunto C da reta projetiva, constituído por d pontos distintos (com $d \geqslant 3$), obtemos um subgrupo finito Γ_C de $\mathrm{PGL}_2(\mathbb{C})$ e também uma superfície não singular de grau d em $\mathbb{P}^3$ contendo exatamente $d^2+d\cdot|\Gamma_C|$ retas. Veremos que a noção de ação de um grupo sobre um conjunto, o estudo das famílias de retas de certas superfícies quádricas não singular em $\mathbb{P}^3$, e principalmente o *Teorema de classificação dos subgrupos finitos de* $\mathrm{PGL}_2(\mathbb{C})$ são elementos fundamentais para a determinação das cotas inferiores N_d em (12), sendo N_d o valor máximo para $d^2 + d \cdot |\Gamma_C|$.

Nas próximas três seções, distribuímos os assuntos da seguinte forma:

Na Seção 2 apresentaremos os sólidos Platônicos e uma descrição de seus grupos de simetrias. Além disso, veremos uma breve descrição da relação que existe entre os subgrupos finitos do grupo das rotações em $\mathbb{R}^3$ e os subgrupos finitos de $\mathrm{PGL}_2(\mathbb{C})$. Encerramos a seção com o enunciado do *Teorema de Classificação dos subgrupos finitos de* $\mathrm{PGL}_2(\mathbb{C})$ estabelecido por Klein em [11].

Na Seção 3, em um primeiro momento, introduziremos as noções mínimas necessárias de geometria algébrica (tais como: espaço projetivo e superfície não singular) para tornar este texto compreensível para aqueles leitores que não tem conhecimentos prévios nessa área. A seguir, serão apresentadas a noção de ação de um grupo sobre um conjunto e o grupo dos automorfismos da reta projetiva complexa $\mathrm{Aut}(\mathbb{P}^1)$. O foco desta seção é o estudo da ação de subgrupos finitos de $\mathrm{Aut}(\mathbb{P}^1)$ sobre o conjunto formado pelos seus pontos fixos. Uma consequência muito importante desse estudo é a classificação dos subgrupos de $\mathrm{Aut}(\mathbb{P}^1)$ que deixam invariantes d pontos distintos da reta projetiva, obtida no Teorema 3.1 (que é utilizado para obter as cotas inferiores N_d).

A Seção 4 começa com uma breve descrição dos resultados conhecidos sobre a contagem de retas em superfícies não singulares de grau d no espaço projetivo complexo $\mathbb{P}^3$. Em seguida, para cada subconjunto C (da reta projetiva $\mathbb{P}^1$) formado por d pontos distintos, associaremos uma superfície não singular em $\mathbb{P}^3$. Prosseguindo, descreveremos o esquema que será utilizado para contar as retas nessas superfícies e as linhas principais para demonstrar que tais superfícies contém $d^2 + d \cdot |\Gamma_C|$ retas. Concluímos com a estimativa do valor máximo:

$$N_d = \max\left\{ d^2 + d \cdot |\Gamma_C| \;\middle|\; C \text{ consiste de } d \text{ pontos distintos} \right\}.$$

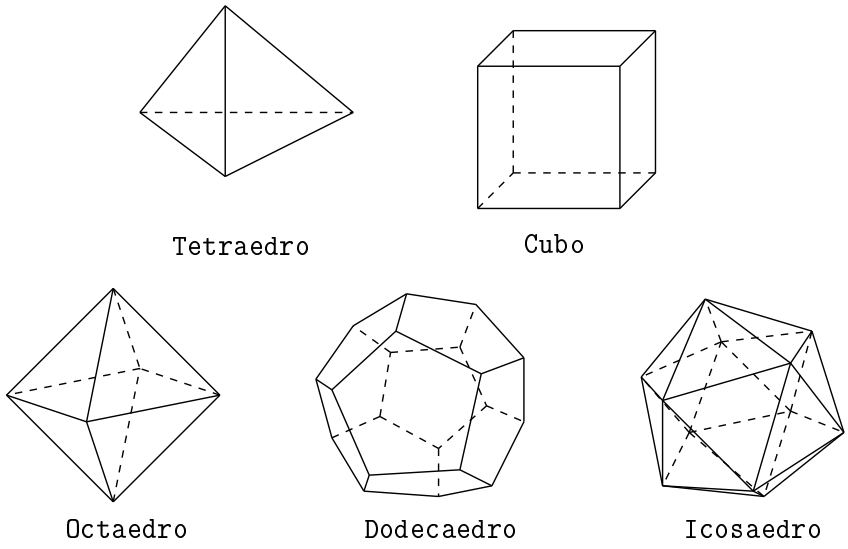
2. Sólidos Platônicos

Um Sólido Platônico é um *poliedro convexo*[1], no qual:

- todas as suas faces são polígonos regulares iguais;
- em todos os seus vértices encontram-se o mesmo número de arestas[2].

Estes sólidos também são denominados *poliedros regulares*.

Um dos fatos que mais surpreende, como também o fez na Grécia antiga a mais de 2000 anos, é a existência de exatamente 5 desses poliedros regulares, apresentados na figura a seguir.



Por certo, a primeira demonstração conhecida de que existem exatamente cinco é atribuída a Theaetetus (c. 417-369 a.C.). Entretanto, foi o filósofo grego Platão (c. 427-347 a.C.) quem primeiro postulou a existência desses poliedros em seu diálogo *Timaeus*[3], a posteriori nomeados *Sólidos Platônicos* em sua homenagem. Posteriormente, Euclides deu uma descrição matemática completa dos sólidos platônicos no livro XIII final dos Elementos de Euclides.

[1] Poliedro é um sólido limitado por superfícies planas que chamamos polígonos. Um poliedro é convexo quando qualquer reta secante possui exatamente dois pontos em comum com o poliedro.

[2] As arestas são as linhas resultantes do encontro de duas faces (se a interseção for não vazia), e os Vértices são os pontos de encontro das arestas.

[3] Nesse texto o narrador de Platão, o astrônomo Timeu de Locri, usa triângulos - como figuras fundamentais - para criar quatro dos cinco poliedros regulares (tetraedro, hexaedro, octaedro, icosaedro). Além disso, em Timeu, Platão associa cada um dos elementos clássicos (terra, ar, água e fogo) a um poliedro regular. A terra é associada ao cubo, o ar ao octaedro, a água ao icosaedro e o fogo ao tetraedro. Com relação ao quinto sólido platônico, o dodecaedro, Platão escreve: "Faltava ainda uma quinta construção que o deus utilizou para organizar todas as constelações do céu.". Aristóteles introduziu um quinto elemento, o éter, e postulou que os céus eram feitos deste elemento, mas ele não teve interesse em associá-lo com o quinto sólido de Platão.

Dualidade. Um olhar atencioso em cada sólido platônico nos permite determinar os seguintes números.

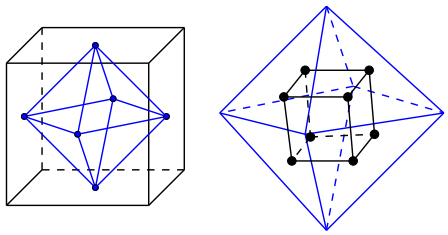
Sólido Platônico	Número de faces	Número de vértices	Número de arestas	Número de arestas por face
Tetraedro	4	4	6	3
Cubo	6	8	12	4
Octaedro	8	6	12	3
Dodecaedro	12	20	30	5
Icosaedro	20	12	30	3

Ao analisar a Tabela acima com atenção, percebemos uma dualidade entre o número de arestas, faces e vértices, para os dois pares cubo-octaedro e dodecaedro-icosaedro: em cada par, os dois sólidos têm o mesmo número de arestas e o número de faces de cada um deles é igual ao número de vértices do outro.

De fato, se S for um sólido platônico, o *dual* de S é o poliedro determinado da seguinte forma:

1º. considere o centro de cada face;
2º. conecte por segmentos os pontos que se encontram em faces vizinhas.

Note que, o dual do tetraedro é o próprio tetraedro. O dual do cubo é o octaedro. E o dual do octaedro é o cubo. Tal como, o icosaedro é o dual do dodecaedro e vice-versa.



Essa dualidade permite concluir que um dado sólido platônico e seu dual tem o mesmo grupo de simetria.

Grupos de simetria dos Sólidos Platônicos. Lembremos que uma *isometria* em $\mathbb{R}^n$ é uma função bijetora $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ que preserva distâncias, ou seja, $||f(\mathrm{x}) - f(\mathrm{y})|| = ||\mathrm{x} - \mathrm{y}||$ para todo $\mathrm{x}, \mathrm{y} \in \mathbb{R}^n$ (sendo "$||\ ||$" a norma Euclidiana). E o conjunto $\mathrm{Isom}(\mathbb{R}^n)$ constituído por tais aplicações é um grupo com a composição de funções.

Uma descrição mais precisa destas isometrias é a seguinte[4]:

$$f \in \mathrm{Isom}(\mathbb{R}^n) \Longleftrightarrow \begin{cases} \exists\, A \text{ matriz ortogonal de ordem } n \times n \text{ e} \\ \mathrm{v} \in \mathbb{R}^n \text{ tal que } f(\mathrm{x}) = A\mathrm{x} + \mathrm{v}. \end{cases}$$

[4]De fato, se $f \in \mathrm{Isom}(\mathbb{R}^n)$ e $0 \in \mathbb{R}^n$ for a origem, então considere $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ dada por $\varphi(\mathrm{x}) = f(\mathrm{x}) - f(0)$. A partir da identidade $2\langle \mathrm{x}, \mathrm{y}\rangle = ||\mathrm{x}||^2 + ||\mathrm{y}||^2 - ||\mathrm{x} - \mathrm{y}||^2$ e o fato de f ser uma isometria, segue-se que $2\langle \varphi(\mathrm{p}), \varphi(\mathrm{q})\rangle = 2\langle \mathrm{p}, \mathrm{q}\rangle$ para quaisquer p, q em $\mathbb{R}^n$. Agora, considere a base canônica $\{e_1, ..., e_n\}$ de $\mathbb{R}^n$ e note que $\langle \varphi(\mathrm{p} + \mathrm{q}), \varphi(e_i)\rangle = \langle \varphi(\mathrm{p}) + \varphi(\mathrm{q}), \varphi(e_i)\rangle$ e $\langle \varphi(\lambda \mathrm{p}), \varphi(e_i)\rangle = \langle \lambda\varphi(\mathrm{p}), \varphi(e_i)\rangle$ para todo $i \in \{1, \ldots, n\}$, $\lambda \in \mathbb{R}$ e p, q $\in \mathbb{R}^n$. Assim, $\varphi(\mathrm{p} + \mathrm{q}) = \varphi(\mathrm{p}) + \varphi(\mathrm{q})$ e $\varphi(\lambda \mathrm{p}) = \lambda\varphi(\mathrm{p})$, ou seja, φ é uma transformação linear ortogonal.

Por exemplo, para $n = 1$ tem-se

$$\text{Isom}(\mathbb{R}) = \Big\{ f : \mathbb{R} \to \mathbb{R} \mid f(\text{x}) = a\text{x} + b,\ a \in \{-1, 1\}, b \in \mathbb{R} \Big\}.$$

Observe que subgrupos de $\text{Isom}(\mathbb{R}^n)$ aparecem sob a condição de invariância para subconjuntos de $\mathbb{R}^n$.

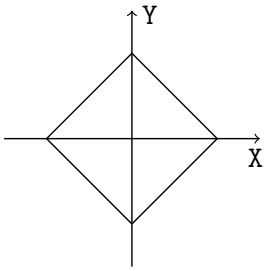
Concretamente, para cada subconjunto X de $\mathbb{R}^n$, considere

$$S(X) = \Big\{ f \in \text{Isom}(\mathbb{R}^n) \mid f(X) = X \Big\}$$

denominado de agora em diante, *grupo de simetrias* de X.

A título de exemplo, se $P_m \subset \mathbb{R}^2$ for um polígono regular de m-lados ($m \geqslant 3$) com seu centro na origem, o grupo de simetrias $S(P_m)$ é isomorfo ao grupo diedral[5] D_m de ordem $2m$.

Observe que para $m = 4$, ao colocarmos o quadrado com seu centro na origem e seus vértices sobre os eixos coordenados, tem-se que $S(P_4)$ é gerado pela rotação em um ângulo $\frac{\pi}{2}$ (dada por $(x, y) \mapsto (-y, x)$, que por sua vez gera um grupo cíclico de ordem 4) e a reflexão em torno do eixo `X` (dada por $(x, y) \mapsto (x, -y)$ de ordem 2).



Antes de apresentar a descrição dos grupos de simetria dos sólidos platônicos, vamos colocar uma pergunta que possui o mesmo teor do Teorema de classificação de Klein, que apresentaremos em breve, a saber:

Quais são os subgrupos finitos de $\text{Isom}(\mathbb{R}^n)$*, a menos de isomorfismo?*

No caso $n = 1$, observe que a ordem dos elementos $x \mapsto -x + b$ é 2 para todo b e eles não comutam entre si, já a ordem dos elementos $x \mapsto x + b$ é infinita para $b \neq 0$. A partir disso podemos concluir que os subgrupos finitos de $\text{Isom}(\mathbb{R})$ são os cíclicos de ordem 1 e 2.

No caso $n = 2$, é atribuído a *Leonardo da Vinci* (1452-1519) ter mostrado que os únicos subgrupos finitos de $\text{Isom}(\mathbb{R}^2)$ são o grupo cíclico C_m de ordem m e o grupo diedral D_m de ordem $2m$ (cf. p. 66 em [23]). E como já comentamos, esses grupos aparecem no estudo dos grupos de simetrias de um polígono regular de m lados.

O caso $n \geqslant 3$ deixamos como uma pesquisa para o leitor, por não ser o foco deste artigo.

Notação: Seja Y um conjunto. De agora em diante usaremos a notação $|Y|$ para indicar a cardinalidade de Y.

[5]Se $I_m = \{1, ..., m\}$ é o conjunto dos vértices de um polígono regular de m lados, então D_m é o subgrupo de $S_m = \{f : I_m \longrightarrow I_m \mid f \text{ é uma bijeção}\}$ (S_m é um grupo com a composição de funções, usualmente denominado *grupo das permutações de m elementos*) gerado pelas funções $\theta, r \in S_m$ dadas por $\theta(i) = i + 1$ para $i \neq m$ e $\theta(m) = 1$; $r(i) = m + 2 - i$ para $i \neq 1$ e $r(1) = 1$. De fato, $D_m = \langle \theta, r \rangle$ é um grupo de ordem $2m$ tal que $ord(\theta) = m$, $ord(r) = 2$ e $\theta \circ r = r \circ \theta^{m-1}$ (cf. p. 196 em [10]).

O subgrupo das rotações $\mathrm{Rot}(X)$ de $S(X)$. Seja X um subconjunto de $\mathbb{R}^n$ contendo a origem. O grupo de *simetria por rotação* de X é o subgrupo $\mathrm{Rot}(X)$ de $S(X)$ constituído pelas simetrias $f(\mathrm{x}) = A\mathrm{x}$, com A ortogonal e determinante de A igual a um.

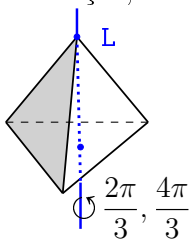
Doravante, S_n denota o grupo das permutações de n-elementos e A_n o subgrupo de S_n formado pelas permutações pares (também denominado grupo alternado)[6].

A seguir faremos uma breve descrição do grupo de rotações Rot(X), no caso do tetraedro e do cubo (este segundo também vale para seu dual, o octaedro).

Tetraedro. No caso deste sólido temos duas escolhas para eixo de rotação, a saber:

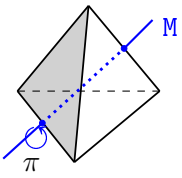
- Eixo de rotação L passando por um vértice e pelo centro da face oposta. Com ângulo de rotação $\frac{2\pi}{3}$ e $\frac{4\pi}{3}$.

  Levando em conta as quatro possíveis escolhas de vértice, obtemos 8 rotações.



- Eixo de rotação M passando pelo ponto médio de uma das arestas e pelo ponto de médio da aresta oposta (aquela que resulta da interseção das faces, que não contém a aresta onde fixamos o primeiro ponto médio). Com ângulo de rotação π.

  Considerando as três possíveis escolhas de pares de arestas, obtemos 3 rotações.



Ao incluirmos a identidade, obtemos exatamente 12 rotações. Assim, $|\mathrm{Rot(X)}| = 12$. Observe que essas rotações permutam os vértices do tetraedro, e portanto, a cada uma dessas rotações podemos associar uma permutação em S_4. Para isto, podemos enumerar os vértices do tetraedro de 1 a 4, e denotarmos por $p_{i,j}$ o ponto médio da aresta com extremos nos vértices i e j. Deste modo, podemos representar as rotações da seguinte forma:

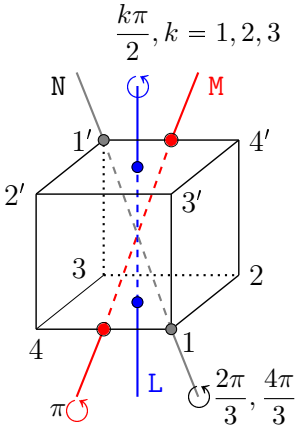
$$\begin{array}{lcl} \text{Eixo L passando pelo vértice 1 e ângulo } \frac{2\pi}{3} & \longmapsto & (2\,3\,4). \\ \text{Eixo L passando pelo vértice 1 e ângulo } \frac{4\pi}{3} & \longmapsto & (2\,4\,3). \\ \text{Eixo M passando por } p_{2,3},\ p_{1,4} \text{ e ângulo } \pi & \longmapsto & (2\,3)(1\,4). \end{array}$$

A partir da descrição acima o leitor pode deduzir que $\mathrm{Rot(X)} \cong A_4$

Cubo. Pare este sólido temos três escolhas para eixo de rotação, a saber:

[6]Lembramos que $S_n = \{f : I_n \longrightarrow I_n \mid f \text{ é uma bijeção}\}$, sendo $I_n = \{1, ..., n\}$, é um grupo com a composição de funções. De fato, $|S_n| = n!$ e $|A_n| = \frac{n!}{2}$ se, $n \geqslant 2$ (cf. p. 218-224 em [10]).

- Eixo de rotação L passando pelo ponto médio de faces opostas, com ângulo de rotação $\frac{k\pi}{2}$, $k = 1, 2, 3$. Obtemos ao todo $3 \cdot 3 = 9$ rotações.
- Eixo de rotação N passando por vértices opostos (isto é, uma de suas diagonais), com ângulo de rotação $\frac{2k\pi}{3}$ $k = 1, 2$. Obtemos ao todo $2 \cdot 4 = 8$ rotações.
- Eixo de rotação M passando pelo ponto médio de uma aresta e da sua oposta superior (conforme ilustra a figura), com ângulo de rotação π. Neste caso obtemos, 6 rotações.



Ao incluirmos a identidade, obtemos exatamente 24 rotações. Logo, $|\mathrm{Rot(X)}| = 24$. Note que essas rotações permutam as diagonais ($\overline{11'}$, $\overline{22'}$, $\overline{33'}$ e $\overline{44'}$). Podemos associar a cada uma dessas rotações uma permutação em S_4, via sua ação nas diagonais (ao fazermos a identificação $1 \equiv \overline{11'}$, $2 \equiv \overline{22'}$, $3 \equiv \overline{33'}$ e $4 \equiv \overline{44'}$). De fato, conforme a figura acima, temos:

$$\begin{array}{lcl} \text{Eixo L com ângulo } \frac{\pi}{2} & \longmapsto & (1\,2\,3\,4). \\ \text{Eixo N com ângulo } \frac{2\pi}{3} & \longmapsto & (2\,3\,4). \\ \text{Eixo M (com ângulo } \pi) & \longmapsto & (1\,4). \end{array}$$

A descrição acima permite concluir que $\mathrm{Rot(X)} \cong S_4$.

De fato, a classificação dos grupos de simetrias dos sólidos platônicos (cf. p. 52–56 em [15]) é dada por :

Grupos de simetria dos Sólidos Platônicos. Seja X um sólido platônico e C_2 um grupo cíclico de ordem 2, então

(i) $S(\mathrm{X}) \cong S_4$ e $\mathrm{Rot(X)} \cong A_4$, se X for um tetraedro.
(ii) $S(\mathrm{X}) \cong S_4 \times C_2$ e $\mathrm{Rot(X)} \cong S_4$, se X for um cubo ou um octaedro.
(iii) $S(\mathrm{X}) \cong A_5 \times C_2$ e $\mathrm{Rot(X)} \cong A_5$, se X for um dodecaedro ou um icosaedro.

Subgrupos finitos de $\mathrm{Rot}(\mathbb{R}^3)$ e o Teorema de classificação de Klein. Na subseção anterior constatamos que os grupos A_4 e S_4 são obtidos a partir dos grupos de simetrias por rotações do tetraedro e cubo, respectivamente. Cabe assinalar que esses grupos reaparecem na classificação dos subgrupos finitos de $\mathrm{Rot}(\mathbb{R}^3)$ (ou seja, das rotações no espaço 3-dimensional) e no *Teorema de classificação de Klein* (que será enunciado em breve). De fato, tem-se (cf. p. 105 [2], p. 149 [23]) que:

Subgrupos finitos de $\mathrm{Rot}(\mathbb{R}^3)$. *Um subgrupo finito de* $\mathrm{Rot}(\mathbb{R}^3)$ *é isomorfo a exatamente um dos seguintes grupos:*

- *cíclico;*
- *diedral de ordem* $2m$, $m \geqslant 2$;
- *grupo de simetrias por rotações de um dos sólidos platônicos.*

O leitor poderá apreciar em breve a similitude que existe entre o resultado acima (classificação dos subgrupos finitos de $\mathrm{Rot}(\mathbb{R}^3)$) e o Teorema de classificação de

Klein. De fato, veremos que essa semelhança tem como base a conexão que existe entre o grupo $\mathrm{Rot}(\mathbb{R}^3)$ e o grupo projetivo linear $\mathrm{PGL}_2(\mathbb{C})$ via o grupo unitário especial $\mathrm{SU}_2(\mathbb{C})$, definidos a seguir.

Os grupos $\mathrm{PGL}_2(\mathbb{C})$, $\mathrm{SU}(2)$ e rotações em $\mathbb{R}^3$

Seja $\mathrm{GL}_2(\mathbb{C})$ o grupo[7] formado pelas matrizes 2×2 complexas invertíveis e $\mathrm{E}_2(\mathbb{C})$ o subgrupo normal de $\mathrm{GL}_2(\mathbb{C})$ formado por todas as matrizes escalares. O grupo quociente $\mathrm{GL}_2(\mathbb{C})/\mathrm{E}_2(\mathbb{C})$ será denotado por $\mathrm{PGL}_2(\mathbb{C})$ e denominado *grupo projetivo linear* associado ao espaço vetorial complexo $\mathbb{C}^2$. Considere o subgrupo (denominado *grupo unitário especial*)

$$\mathrm{SU}_2(\mathbb{C}) = \left\{ \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \in \mathrm{GL}_2(\mathbb{C}) \;\middle|\; |z|^2 + |w|^2 = 1 \right\},$$

sendo $\overline{z} = a - bi$ o conjugado de $z = a + bi$ e $|z|^2 = \overline{z}z = a^2 + b^2$. De fato, se $A = \begin{bmatrix} z & w \\ -\overline{w} & \overline{z} \end{bmatrix} \in \mathrm{SU}_2(\mathbb{C})$, com $z = a + bi$ e $w = c + di$, ao considerarmos o vetor $\mathrm{q} = (b, c, d) \in \mathbb{R}^3$ de norma $||\mathrm{q}|| = \sqrt{b^2 + c^2 + d^2}$, vamos definir a transformação linear $R_A : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ dada por

$$\mathrm{v} \longmapsto (a^2 - ||\mathrm{q}||^2)\mathrm{v} + 2(\mathrm{q}\cdot\mathrm{v})\mathrm{q} + (2a)\mathrm{q}\times\mathrm{v} \tag{1}$$

sendo "$\cdot$" o produto interno usual e "$\times$" o produto vetorial em $\mathbb{R}^3$.

Observe que, se q é o vetor nulo (logo $|z|^2 + |w|^2 = |a| = 1$ e $A = \pm\mathrm{I}$, sendo I a matriz identidade), então R_A é a função identidade. Entretanto, se q for não nulo e considerarmos o ângulo $\theta \in [0, 2\pi]$ tal que $a = \cos\frac{\theta}{2}$ e o vetor unitário $\mathrm{u} = \frac{\mathrm{q}}{||\mathrm{q}||}$, obtemos $\mathrm{q} = \operatorname{sen}\frac{\theta}{2}\mathrm{u}$ e após substituição em (1), segue que

$$R_A(\mathrm{v}) = (\cos\theta)\mathrm{v} + (\operatorname{sen}\theta)\mathrm{u}\times\mathrm{v} + (1-\cos\theta)(\mathrm{u}\cdot\mathrm{v})\mathrm{u}$$

sendo o lado direito exatamente fórmula de *Euler-Rodrigues*[8]. Assim, R_A é uma rotação com eixo contendo a origem, determinado pelo vetor diretor u, no ângulo θ, e portanto, $R_A \in \mathrm{Rot}(\mathbb{R}^3)$. Deixamos a cargo do leitor mostrar que $R : \mathrm{SU}_2(\mathbb{C}) \longrightarrow \mathrm{Rot}(\mathbb{R}^3)$ dada por $A \longmapsto R_A$ é um homomorfismo sobrejetivo de grupos tendo por núcleo $\{\pm\mathrm{I}\}$[9]. Assim, o grupo quociente $\mathrm{PSU}_2(\mathbb{C}) = \mathrm{SU}_2(\mathbb{C})/\{\pm\mathrm{I}\}$ é isomorfo ao grupo $\mathrm{Rot}(\mathbb{R}^3)$.

Além disso, observe que $\mathrm{PSU}_2(\mathbb{C})$ é um subgrupo de $\mathrm{PGL}_2(\mathbb{C})$ (visto que $\mathrm{SU}_2(\mathbb{C})$ é um subgrupo de $\mathrm{GL}_2(\mathbb{C})$ e tem-se que $\{\pm\mathrm{I}\} \leqslant \mathrm{E}_2(\mathbb{C})$). Pode-se mostrar que todo subgrupo finito de $\mathrm{PGL}_2(\mathbb{C})$ é conjugado[10] a um subgrupo de $\mathrm{PSU}_2(\mathbb{C})$ (cf. p. 9 [21], p. 84 de [13]). Assim, subgrupos finitos de $\mathrm{PGL}_2(\mathbb{C})$ irão se identificar com conjugados de subgrupos de $\mathrm{Rot}(\mathbb{R}^3)$ (uma vez que $\mathrm{Rot}(\mathbb{R}^3)$ é isomorfo ao grupo $\mathrm{PSU}_2(\mathbb{C})$).

[7] Com o produto usual de matrizes.

[8] Tal fórmula foi descoberta inicialmente por *Leonhard Paul Euler* (1707-1783) em torno de 1770 (cf. [9]) e posteriormente de forma independente por *Benjamin Olinde Rodrigues* (1795-1851) por volta de 1840 (cf. [17]).

[9] Uma abordagem via quatérnios unitários encontra-se na p. 2 das notas de *Dolgachev* [8].

[10] Os subgrupos H_1 e H_2 do grupo $(G, *)$ são ditos *conjugados*, se existe $g \in G$ tal que $H_2 = gH_1g^{-1}$ sendo $gH_1g^{-1} = \{g * h * g^{-1} \in G \mid h \in H_1\}$.

O leitor poderá apreciar a importância do Teorema de classificação de Klein, para o desenvolvimento deste texto, no momento que formos estudar os conjuntos invariantes de pontos na reta projetiva[11].

Teorema de classificação de Klein (**TCK**). *Um subgrupo finito de* $\mathrm{PGL}_2(\mathbb{C})$ *é isomorfo a exatamente um dos seguintes grupos:*

(i) C_m *o grupo cíclico de ordem* m*;*
(ii) D_m *o grupo diedral de ordem* $2m$*,* $m \geqslant 2$*;*
(iii) A_4 *o grupo alternado de ordem 12;*
(iv) S_4 *o grupo das permutações de ordem 24;*
(v) A_5 *o grupo alternado de ordem 60.*

Se dois subgrupos de $\mathrm{PGL}_2(\mathbb{C})$ *são isomorfos, então eles são conjugados.*

3. Sobre Geometria algébrica e ações de grupos

Iniciamos esta seção revisando algumas noções de geometria algébrica (em consideração aos leitores que não tem conhecimentos prévios dessa área).

3.1. **Preliminares de Geometria Algébrica.** O n–espaço projetivo complexo $\mathbb{P}^n$ é constituído de todos os subespaços de dimensão 1 de $\mathbb{C}^{n+1}$ (considere $\mathbb{C}^{n+1}$ como espaço vetorial complexo). Ou seja, $\mathbb{P}^n$ é formado pelas retas que passam pela origem em $\mathbb{C}^{n+1}$. Assim, cada elemento em $\mathbb{P}^n$ é da forma $[\mathrm{v}]$, com v vetor não nulo em $\mathbb{C}^{n+1}$. Se $\mathrm{v} = (v_0, ..., v_n)$, usaremos a notação $[\mathrm{v}] = [v_0 : ... : v_n]$ na qual v_0,...,v_n são denominadas de coordenadas homogêneas do ponto $[\mathrm{v}]$. Observe que, os subespaços $[a_0 : ... : a_n]$ e $[b_0 : ... : b_n]$ são iguais (ou seja, definem o mesmo ponto em $\mathbb{P}^n$) se, e somente se, $(a_0, ..., a_n) = \lambda(b_0, ..., b_n)$ para algum $\lambda \in \mathbb{C}$ não nulo. Logo, pontos distintos são determinados por vetores linearmente independentes.

Neste contexto, $\mathbb{P}^1$, $\mathbb{P}^2$ e $\mathbb{P}^3$ são usualmente denominados de *reta projetiva*, *plano projetivo* e *espaço projetivo*, respectivamente.

Assim como $x+y=0$ define uma reta no plano cartesiano $\mathbb{R}^2$ e $x^2+2y^2+4z^2-4=0$ define um elipsoide em $\mathbb{R}^3$, retas e superfícies em $\mathbb{P}^n$ também são definidos a partir dos zeros de polinômios, só que tais polinômios precisam ser homogêneos.

Polinômios homogêneos e seu conjunto de zeros. Considere $f \in \mathbb{C}[x_0, ..., x_n]$ não nulo e $d \geqslant 0$ inteiro. Dizemos que f é *homogêneo de grau* d se, para todo $\lambda \in \mathbb{C}$,

$$f(\lambda x_0, ..., \lambda x_n) = \lambda^d f(x_0, ..., x_n).$$

Se $f \in \mathbb{C}[x_0, ..., x_n]$ for homogêneo e não constante[12], definimos o *conjunto dos zeros* [13] de f por:

$$\mathcal{Z}(f) = \left\{ [\mathrm{v}] \in \mathbb{P}^n \;\middle|\; f(\mathrm{v}) = 0 \right\}.$$

Usualmente, $\mathcal{Z}(f)$ é denominada *hipersuperfície* definida por f em $\mathbb{P}^n$.

[11] Visto que o grupo $\mathrm{PGL}_2(\mathbb{C})$ é isomorfo ao grupo dos automorfismos da reta projetiva (veja Proposição 3.2).

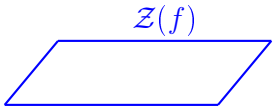
[12] No caso dos polinômios constantes, definimos $\mathcal{Z}(0) = \mathbb{P}^n$ e $\mathcal{Z}(a) = \varnothing$, se $a \in \mathbb{C}$ não nulo.

[13] Ressaltamos que a homogeneidade de f nos garante a boa definição de $\mathcal{Z}(f)$, uma vez que $f(\lambda \mathrm{v}) = \lambda^d f(\mathrm{v})$, $f(\mathrm{v}) = 0$ implica em $f(\lambda \mathrm{v}) = 0$.

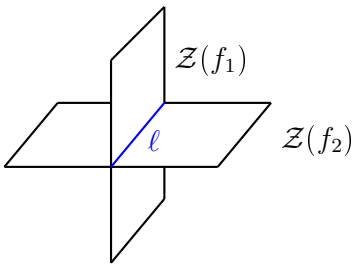
Notação: Se $f_1, f_2, ..., f_k$ forem polinômios homogêneos não constantes (não necessariamente do mesmo grau) em $\mathbb{C}[x_0, ..., x_n]$, então $\mathcal{Z}(f_1, f_2, ..., f_k)$ denotará a interseção $\bigcap_{i=1}^{k} \mathcal{Z}(f_i)$.

Por exemplo, para

- $n = 1$: $\mathcal{Z}(f)$ é um conjunto finito de pontos na reta projetiva[14].
- $n = 2$: $\mathcal{Z}(f)$ é denominada *curva projetiva plana*. Se o grau de f for 1, então $\mathcal{Z}(f)$ é uma *reta* no plano projetivo $\mathbb{P}^2$.
- $n = 3$: $\mathcal{Z}(f)$ é denominada *superfície projetiva*. $\mathcal{Z}(f)$ é um *plano* em $\mathbb{P}^3$ se f for homogêneo de grau 1. Uma *reta* em $\mathbb{P}^3$ é determinada pela interseção de planos distintos[15], conforme ilustra a próxima figura.



Plano $\mathcal{Z}(f)$ em $\mathbb{P}^3$



Reta $\ell = \mathcal{Z}(f_1) \cap \mathcal{Z}(f_2)$ em $\mathbb{P}^3$

Outra descrição de retas e planos em $\mathbb{P}^3$.

(i) H é um plano em $\mathbb{P}^3 \iff$ Existe W subespaço vetorial de dimensão 3 de $\mathbb{C}^4$ tal que $H = \left\{ [\mathrm{w}] \in \mathbb{P}^3 \;\middle|\; \mathrm{w} \in W, \text{ não nulo} \right\}$.

De forma análoga,

(ii) ℓ é uma reta em $\mathbb{P}^3 \iff$ Existe U subespaço vetorial de dimensão 2 de $\mathbb{C}^4$ tal que $\ell = \left\{ [\mathrm{u}] \in \mathbb{P}^3 \;\middle|\; \mathrm{u} \in U, \text{ não nulo} \right\}$.

Neste caso, usamos a notação $H = \mathbb{P}(W)$ e $\ell = \mathbb{P}(U)$. Ao trocarmos 3 por n, nesta abordagem obtemos uma definição para reta e plano em $\mathbb{P}^n$.

(iii) Se $\mathrm{p} = [\mathrm{u}_1]$ e $\mathrm{q} = [\mathrm{u}_2]$ são pontos distintos em $\mathbb{P}^n$, então existe uma única reta passando por p e q. De fato, tal reta é determinada pelo subespaço $U = [\mathrm{u}_1, \mathrm{u}_2]$ e

$$\mathbb{P}(U) = \left\{ [a_0\mathrm{u}_1 + a_1\mathrm{u}_2] \in \mathbb{P}^n \;\middle|\; [a_0 : a_1] \in \mathbb{P}^1 \right\}.$$

A noção de ponto singular, apresentada a seguir, nos permitirá introduzir o conceito de superfície não singular.

Ponto singular de $\mathcal{Z}(f)$. Considere $f \in \mathbb{C}[x_0, ..., x_n]$ homogêneo de grau $d \geqslant 1$.

[14] Se $f \in \mathbb{C}[x_0, x_1]$ é homogêneo de grau d então verifica-se que $f = (b_0x_0 - a_0x_1)^{n_0} \cdots (b_kx_0 - a_kx_1)^{n_k}$, com $n_i \in \mathbb{N}$ e $d = n_0 + \cdots + n_k$, $a_i, b_i \in \mathbb{C}$ não ambos nulos. Logo, $\mathcal{Z}(f) = \{[a_0 : b_0], ..., [a_k : b_k]\}$.

[15] $\mathcal{Z}(f_1)$ e $\mathcal{Z}(f_2)$ são planos distintos em $\mathbb{P}^3$ se, e somente se, f_1 e f_2 são linearmente independentes. Um fato interessante da geometria em $\mathbb{P}^3$ é que dois planos sempre se intersectam.

Um ponto $[\mathrm{v}] \in \mathbb{P}^n$ é dito *ponto singular de* $\mathcal{Z}(f)$ se, $\partial_i f(\mathrm{v}) = 0$ com $i \in \{0, ..., n\}$, onde ∂_i denota a derivada parcial de f em relação à variável x_i.

Os pontos singulares são zeros de f. Ao derivarmos $f(\lambda x_0, ..., \lambda x_n) = \lambda^d f(x_0, ..., x_n)$ em relação a λ, obtemos

$$\sum_{i=0}^{n} \frac{\partial f}{\partial(\lambda x_i)} \cdot x_i = d\lambda^{d-1} f(x_0, ..., x_n).$$

Fazendo $\lambda = 1$ nessa última igualdade chegamos na *Identidade de Euler*:

$$\partial_0 f \cdot x_0 + \cdots + \partial_n f \cdot x_n = d \cdot f.$$

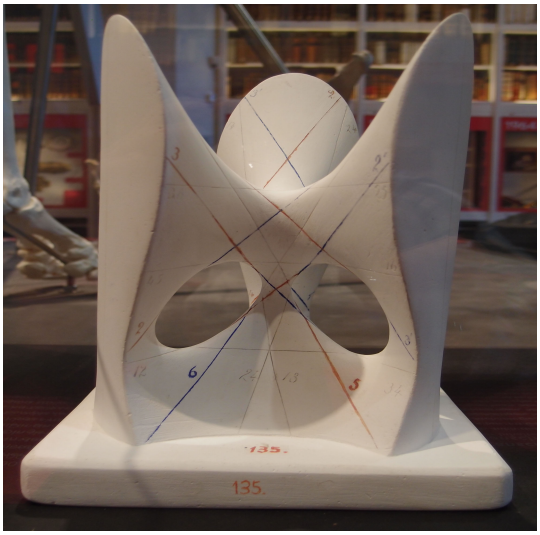
De onde concluímos que, se $[\mathrm{v}]$ é ponto singular de $\mathcal{Z}(f)$, então $[\mathrm{v}] \in \mathcal{Z}(f)$.

Dizemos que $\mathcal{Z}(f)$ é *não singular* se $\mathcal{Z}(f)$ não possui pontos singulares. Por exemplo, todo plano em $\mathbb{P}^3$ é não singular.

Superfícies não singulares de grau d em $\mathbb{P}^3$. Se $f \in \mathbb{C}[x_0, x_1, x_2, x_3]$ for homogêneo de grau $d \geqslant 1$ com $\mathcal{Z}(f) \subset \mathbb{P}^3$ não singular, então $\mathcal{Z}(f)$ será denominada *superfície não singular de grau d*. Também são usados os epítetos superfície quádrica, cúbica, quártica para $d = 2, 3, 4$, respectivamente.

Na Georg-August Universität Göttingen encontra-se um modelo local, da assim denominada cúbica de Clebsch definida por

$$x_0^3 + x_1^3 + x_2^3 + x_3^3 = (x_0 + x_1 + x_2 + x_3)^3.$$



Observe que a superfície cúbica definida acima é não singular[16]. Vale ressaltar que as "linhas" destacadas nessa figura representam retas contidas nessa superfície.

[16]De fato, ao fazer $L = x_0 + x_1 + x_2 + x_3$ e $f = x_0^3 + x_1^3 + x_2^3 + x_3^3 - L^3$ tem-se que $\partial_i f = 3x_i^2 - 3L^2 = 3(x_i - L)(x_i + L)$ para todo i. Assim, se $[\mathrm{v}] \in \mathbb{P}^3$ com $\mathrm{v} = (v_0, v_1, v_2, v_3)$ for um ponto singular de $\mathcal{Z}(f)$, segue que $L(\mathrm{v}) = \pm v_i$ para todo i (o que é absurdo). Portanto, $\mathcal{Z}(f)$ é uma superfície não singular em $\mathbb{P}^3$.

3.2. **Ação de um grupo sobre um conjunto.** Se $(G, *)$ for um grupo com elemento neutro e, considere $G^\star = G - \{e\}$.

Lembremos que o grupo G *age pela esquerda* num conjunto C, se existe uma função de $G \times C$ em C que associa a cada par (g, x) um único elemento em C, que denotaremos por $g{\cdot}\mathrm{x}$, satisfazendo às condições:

(i) $e{\cdot}\mathrm{x} = \mathrm{x}$, para todo $\mathrm{x} \in C$,

(ii) $h{\cdot}(g{\cdot}\mathrm{x}) = (h * g){\cdot}\mathrm{x}$ para todo $\mathrm{x} \in C$, $h, g \in G$.

Para cada $\mathrm{x} \in C$, $G_\mathrm{x} = \{g \in G \mid g{\cdot}\mathrm{x} = \mathrm{x}\}$ é denominado de *estabilizador* de x em G e $O_\mathrm{x} = \{g{\cdot}\mathrm{x} \in C \mid g \in G\}$ é denominado de *órbita* de x (relativa a ação definida por "$\cdot$").

Um resultado importante neste contexto é a seguinte Proposição.

Proposição 3.1. *Com as notações acima, verifica-se que:*

(i) $O_\mathrm{x} = O_\mathrm{y}$ *se, e somente se, existe* $g \in G$, *tal que* $\mathrm{y} = g{\cdot}\mathrm{x}$.

(ii) $O_\mathrm{x} \cap O_\mathrm{y} = \varnothing$ *ou* $O_\mathrm{x} = O_\mathrm{y}$.

(iii) *Para todo* $\mathrm{y} \in O_\mathrm{x}$ *tem-se que* G_y *e* G_x *são grupos conjugados* (*logo isomorfos*).

(iv) $|O_\mathrm{x}| = (G : G_\mathrm{x})$. *Portanto, se* G *for um grupo finito, o número de elementos de cada órbita, é um divisor da ordem do grupo* G.

Se C for um conjunto finito, a Proposição 3.1 nos garante que podemos escolher uma quantidade finita de elementos em C, digamos $\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_k$, tais que:

$$C = O_{\mathrm{x}_1} \cup O_{\mathrm{x}_2} \cup \cdots \cup O_{\mathrm{x}_k} \text{ sendo que } O_{\mathrm{x}_i} \cap O_{\mathrm{x}_j} = \varnothing, \quad \forall\, i \neq j. \tag{2}$$

Neste caso, $\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_k$ denomina-se *sistema de representantes* determinado pela ação de G em C. A partição em (2) é chamada de *decomposição de* C *em órbitas* (valendo a unicidade a menos de uma permutação dos índices).

A seguir, iniciamos o estudo dos automorfismos de $\mathbb{P}^1$ e seus subgrupos finitos, visto que desempenham um papel muito importante na contagem de retas que faremos mais adiante.

Automorfismos da reta projetiva. Seja $\mathrm{Aut}(\mathbb{C}^2)$ o grupo constituído pelos isomorfismos lineares de $\mathbb{C}^2$ em $\mathbb{C}^2$ sob a operação de composição de funções. Como cada aplicação $T \in \mathrm{Aut}(\mathbb{C}^2)$ leva retas pela origem em retas pela origem, T induz a função

$$\mathbf{T} : \mathbb{P}^1 \longrightarrow \mathbb{P}^1, \quad \text{dada por} \quad \mathbf{T}([\mathrm{v}]) = [T(\mathrm{v})].$$

O conjunto formado por tais funções será denotado por $\mathrm{Aut}(\mathbb{P}^1)$ e seus elementos denominados *automorfismos* de $\mathbb{P}^1$.

Lembre que a função de $\mathrm{Aut}(\mathbb{C}^2)$ em $\mathrm{GL}_2(\mathbb{C})$ dada por $T \longmapsto [T]$, sendo $[T] \in \mathrm{GL}_2(\mathbb{C})$ a matriz associada a T da base canônica na base canônica, é um isomorfismo de grupos. Por exemplo, se $T \in \mathrm{Aut}(\mathbb{C}^2)$ é tal que $[T] \in \mathrm{E}_2(\mathbb{C})$, então $\mathbf{T} = \mathrm{id}_{\mathbb{P}^1} \in \mathrm{Aut}(\mathbb{P}^1)$.

Listamos a seguir algumas das propriedades essenciais dos elementos de $\mathrm{Aut}(\mathbb{P}^1)$ que usaremos neste texto.

Proposição 3.2. *Sejam* $\mathbf{T}, \mathbf{S} \in \mathrm{Aut}(\mathbb{P}^1)$ *determinadas por* $T, S \in \mathrm{Aut}(\mathbb{C}^2)$, *respectivamente. Então verifica-se que:*

(i) $\mathbf{T} = \mathbf{S}$ *se, e somente se,* $T = \lambda S$ *para algum* $\lambda \in \mathbb{C}$ *não nulo.*
(ii) *Dados os pontos* $\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3, \mathrm{q}_1, \mathrm{q}_2$ *e* q_3 *em* $\mathbb{P}^1$ *tais que* $\mathrm{p}_i \neq \mathrm{p}_j$ *e* $\mathrm{q}_i \neq \mathrm{q}_j$ *para todo* $i \neq j$*, existe um único* $\mathbf{T} \in \mathrm{Aut}(\mathbb{P}^1)$*, tal que* $\mathbf{T}(\mathrm{p}_i) = \mathrm{q}_i$ *para* $i = 1, 2, 3$*.*
(iii) $\mathrm{Aut}(\mathbb{P}^1)$ *com a composição de funções é um grupo, tendo* $\mathrm{id}_{\mathbb{P}^1}$ *por elemento neutro.*
(iv) *A função* $\Psi : \mathrm{Aut}(\mathbb{P}^1) \longrightarrow \mathrm{PGL}_2(\mathbb{C}) \stackrel{def}{=} \mathrm{GL}_2(\mathbb{C})/\mathrm{E}_2(\mathbb{C})$ *dada por* $\mathbf{T} \mapsto [T] \cdot \mathrm{E}_2(\mathbb{C})$ *é um isomorfismo de grupos.*

Salientamos que a próxima subseção começa com a noção de ponto fixo, além de introduzir notações para certos automorfismos da reta projetiva complexa (que poderão ser utilizados pelo leitor na determinação do valor de $|\Gamma_C|$ para um dado subconjunto C de $\mathbb{P}^1$ na Seção 4). A Proposição 3.4 é o resultado que nos permitirá classificar os subgrupos finitos de $\mathrm{Aut}(\mathbb{P}^1)$ a partir da decomposição em órbitas de seus pontos fixos.

Ação de subgrupos de $\mathrm{Aut}(\mathbb{P}^1)$ **sobre pontos fixos.** Lembremos que $\mathrm{p} \in \mathbb{P}^1$ é um ponto fixo de $\mathbf{T} \in \mathrm{Aut}(\mathbb{P}^1)$ se $\mathbf{T}(\mathrm{p}) = \mathrm{p}$.

Notações: No que segue considere $\mathbf{e}_1 = [1 : 0]$ e $\mathbf{e}_2 = [0 : 1]$. Se $\kappa \in \mathbb{C}$ for não nulo, então $\mathbf{C}_\kappa$, $\mathbf{P}_\kappa$ e $\mathbf{R}_\kappa$ denotarão os automorfismos de $\mathbb{P}^1$ dados por: $\mathbf{C}_\kappa([x : y]) = [x + \kappa y : y]$, $\mathbf{P}_\kappa([x : y]) = [\kappa y : x]$ e $\mathbf{R}_\kappa([x : y]) = [\kappa x : y]$. Em particular, temos que $\mathbf{R}_1 = \mathrm{id}_{\mathbb{P}^1}$ (o qual possui infinitos pontos fixos). Observe que:

(3)

Automorfismo	Pontos fixos	Valor em $\mathbf{e}_1$	Valor em $\mathbf{e}_2$	Ordem
$\mathbf{C}_\kappa$	$\mathbf{e}_1$	$\mathbf{e}_1$	$[\kappa : 1]$	infinita
$\mathbf{P}_\kappa$	$[\sqrt{\kappa} : 1], [-\sqrt{\kappa} : 1]$	$\mathbf{e}_2$	$\mathbf{e}_1$	2
$\mathbf{R}_\kappa$, $\kappa \neq 1$	$\mathbf{e}_1, \mathbf{e}_2$	$\mathbf{e}_1$	$\mathbf{e}_2$	$ord(\kappa)$ em $(\mathbb{C}^\star, \cdot)$

A seguir para cada subgrupo G de $\mathrm{Aut}(\mathbb{P}^1)$, considere

$$\mathrm{Fix}(G) = \left\{ \mathrm{p} \in \mathbb{P}^1 \;\middle|\; \mathrm{p} \text{ é ponto fixo de algum elemento em } G^\star \right\}.$$

Observe que:

(i) Se $\mathrm{x} \in \mathrm{Fix}(G)$ então $\mathbf{T}(\mathrm{x}) \in \mathrm{Fix}(G)$ para todo $\mathbf{T} \in G$.[17]
(ii) $(\mathbf{T}, \mathrm{x}) \longmapsto \mathbf{T}(\mathrm{x})$ define uma ação pela esquerda de G em $\mathrm{Fix}(G)$.

Na próxima proposição considere a ação pela esquerda $(\mathbf{T}, \mathrm{x}) \longmapsto \mathbf{T}(\mathrm{x})$, tanto para G quanto para H.

Proposição 3.3. *Seja* G *um subgrupo finito de* $\mathrm{Aut}(\mathbb{P}^1)$ *com* $|G| > 1$ *tal que* $\mathrm{Fix}(G) = O_1 \cup \cdots \cup O_k$ *é a decomposição em órbitas de* $\mathrm{Fix}(G)$*. Se* H *for um subgrupo de* $\mathrm{Aut}(\mathbb{P}^1)$ *isomorfo ao grupo* G*, então existe* $\mathbf{L} \in \mathrm{Aut}(\mathbb{P}^1)$ *tal que* $\mathrm{Fix}(H) = \mathbf{L}(O_1) \cup \cdots \cup \mathbf{L}(O_k)$ *é a decomposição em órbitas de* $\mathrm{Fix}(H)$*.*

Demonstração. Como $\mathrm{Fix}(G)$ (resp. $\mathrm{Fix}(H)$) é finito, podemos considerar o sistema finito de representantes $\mathrm{x}_1, ..., \mathrm{x}_k$ em $\mathrm{Fix}(G)$ (resp. $\mathrm{y}_1, ..., \mathrm{y}_l$ em $\mathrm{Fix}(H)$). Por

[17] De fato, se $\mathrm{x} \in \mathrm{Fix}(G)$, então existe $\mathbf{S} \in G^\star$ tal que $\mathbf{S}(\mathrm{x}) = \mathrm{x}$. Agora, basta observar que $\mathbf{T} \circ \mathbf{S} \circ \mathbf{T}^{-1} \in G^\star$ e fixa $\mathbf{T}(\mathrm{x})$ se $\mathbf{T} \in G$.

simplicidade, seja $O_i = O_{\mathrm{x}_i}$ e $\mathcal{O}_j = \mathcal{O}_{\mathrm{y}_j}$. Assim temos:

$$\begin{aligned} \mathrm{Fix}(G) &= O_1 \cup \cdots \cup O_k \text{ com } O_i \cap O_j = \varnothing, \qquad \forall\, i \neq j. \\ \mathrm{Fix}(H) &= \mathcal{O}_1 \cup \cdots \cup \mathcal{O}_l \text{ com } \mathcal{O}_i \cap \mathcal{O}_j = \varnothing, \qquad \forall\, i \neq j. \end{aligned}$$

Como G e H são isomorfos e finitos, o Teorema de classificação de Klein (**TCK**) nos garante que eles são conjugados, logo existe $\mathbf{L} \in \mathrm{Aut}(\mathbb{P}^1)$, tal que o automorfismo interno $I_{\mathbf{L}}$ (note que $I_{\mathbf{L}}(\mathbf{T}) = \mathbf{T} \circ \mathbf{L} \circ \mathbf{T}^{-1} \in H$, para $\mathbf{T} \in G$) induz um isomorfismo de G em H. Deixamos a cargo do leitor verificar que:

Afirmação 1: $\mathrm{Fix}(H) = \mathbf{L}(\mathrm{Fix}(G))$.

Afirmação 2: $\mathbf{L}(O_{\mathrm{x}}) = \mathcal{O}_{\mathbf{L}(\mathrm{x})}$, para todo $\mathrm{x} \in \mathrm{Fix}(G)$.

Portanto,

$$\mathrm{Fix}(H) = \mathcal{O}_1 \cup \cdots \cup \mathcal{O}_l = \mathbf{L}(O_1) \cup \cdots \cup \mathbf{L}(O_k) = \mathcal{O}_{\mathbf{L}(\mathrm{x}_1)} \cup \cdots \cup \mathcal{O}_{\mathbf{L}(\mathrm{x}_k)}.$$

Segue do item (ii) na Proposição 3.1 que para cada $i \in \{1, ..., k\}$, existe um único $j \in \{1, ..., l\}$, tal que $\mathcal{O}_j = \mathcal{O}_{\mathbf{L}(\mathrm{x}_i)} = \mathbf{L}(O_i)$ e necessariamente $k = l$.

■

A demonstração da Proposição 3.4, que enunciaremos a seguir, encontra-se na seção 6.12 do texto de *Artin* [3] (também na p. 3 de [8]).

Proposição 3.4. *Seja G é um subgrupo finito de $\mathrm{Aut}(\mathbb{P}^1)$ de ordem $N \geqslant 2$ tal que*

$$\mathrm{Fix}(G) = O_1 \cup \cdots \cup O_k, \text{ na qual } O_i = O_{\mathrm{x}_i}$$

é a decomposição de $\mathrm{Fix}(G)$ em órbitas induzida pela ação $(\mathbf{T}, \mathrm{x}) \longmapsto \mathbf{T}(\mathrm{x})$ (da Proposição 3.3). Então

(i) *Se $\mathrm{x} \in O_i$ então $|G_{\mathrm{x}}| = |G_{\mathrm{x}_i}| := E_i \geqslant 2$, para cada i.*

(ii) *$k \in \{2, 3\}$. Além disso,*

- *$k = 2$ se, e somente se, G é cíclico.*
- *Se $k = 3$ e $E_1 \leqslant E_2 \leqslant E_3$, então só uma das seguintes possibilidades ocorre:*

E_1	E_2	E_3	G *é isomorfo ao grupo*
2	2	$m \geqslant 2$	D_m
2	3	3	A_4
2	3	4	S_4
2	3	5	A_5

E consequentemente, se $n_i = \frac{|G|}{E_i}$ (é a cardinalidade das órbitas) para $i = 1, 2, 3$, temos:

(4)

n_1	n_2	n_3	G *é isomorfo ao grupo*
m	m	2	D_m
6	4	4	A_4
12	8	6	S_4
30	20	12	A_5

Na próxima subseção será utilizada a proposição acima (junto com o Teorema de Classificação de Klein) para classificar os subgrupos finitos de $\mathrm{Aut}(\mathbb{P}^1)$ que deixam invariantes subconjuntos finitos da reta projetiva (cf. Teorema 3.1). Sendo esse Teorema um dos resultados que nos permitirão determinar a cota N_d em (12).

Classificando subgrupos de $\mathrm{Aut}(\mathbb{P}^1)$ que fixam um conjunto. Para cada subconjunto C de $\mathbb{P}^1$, considere

$$\Gamma_C = \left\{ \mathbf{T} \in \mathrm{Aut}(\mathbb{P}^1) \mid \mathbf{T}(C) = C \right\}.$$

Proposição 3.5. *Γ_C é um subgrupo de $\mathrm{Aut}(\mathbb{P}^1)$, cuja ordem é finita se $|C| \geqslant 3$.*

Demonstração. A seguir mostraremos que $|\Gamma_C| < \infty$ se $|C| \geqslant 3$.

Se $|C| \geqslant 3$ então considere $\{\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3\} \subseteq C$. Segue do item (ii) da Proposição 3.2 que cada elemento $\mathbf{T} \in \Gamma_C$ é determinado pela escolha de

$$\{\mathbf{T}(\mathrm{p}_1), \mathbf{T}(\mathrm{p}_2), \mathbf{T}(\mathrm{p}_3)\} \subseteq C,$$

que é finita e dada pelo binômio $\binom{d}{3}$, sendo $d = |C|$. Agora, levando em conta as permutações (6 ao todo), concluímos que $|\Gamma_C| \leqslant \binom{d}{3}$. ∎

Exemplo 3.1. *Se $C = \{\mathrm{p}_1, \mathrm{p}_2, \mathrm{p}_3\}$, então cada $\sigma \in D_3$ determina $\mathbf{T}_\sigma \in \Gamma_C$ dado por $\mathbf{T}_\sigma(\mathrm{p}_i) = p_{\sigma(i)}$ ((ii) na Proposição 3.2). Deixamos a cargo do leitor a verificação de que $\Psi : D_3 \to \Gamma_C$ dada por $\sigma \mapsto \mathbf{T}_\sigma$ é um isomorfismo de grupos.*

A seguir, assuma que O_{x} denota a órbita de x relativa à ação $(\mathbf{T}, \mathrm{p}) \mapsto \mathbf{T}(\mathrm{p})$ de Γ_C em C, e que $C = O_1 \cup \cdots \cup O_k$ é a decomposição em órbitas relativa a essa ação. Lembre que $\mathrm{Fix}(\Gamma_C)$ é o conjunto formado pelos pontos fixos dos elementos de $(\Gamma_C)^\star$.

Proposição 3.6. *Se $\mathrm{x} \in C$ e $|C| = d \geqslant 3$, temos:*

(i) *$O_{\mathrm{x}} \subseteq \mathrm{Fix}(\Gamma_C) \cap C$ se $\mathrm{x} \in \mathrm{Fix}(\Gamma_C)$, caso contrário $O_{\mathrm{x}} \cap \mathrm{Fix}(\Gamma_C) \cap C = \varnothing$ e $|O_{\mathrm{x}}| = |\Gamma_C|$.*

(ii) *Se Γ_C não for cíclico e $\mathrm{Fix}(\Gamma_C) = \mathcal{O}_1 \cup \mathcal{O}_2 \cup \mathcal{O}_3$ for a decomposição em órbitas de $\mathrm{Fix}(\Gamma_C)$ sob a ação de Γ_C (dada por $(\mathbf{T}, \mathrm{p}) \mapsto \mathbf{T}(\mathrm{p})$), então:*

$$d = \alpha n_1 + \beta n_2 + \gamma n_3 + \delta |\Gamma_C|, \tag{5}$$

na qual $\alpha, \beta, \gamma \in \{0, 1\}$, $\delta \geqslant 0$ e $n_i = |\mathcal{O}_i|$.

Demonstração. (i) Deixamos a cargo do leitor.

(ii) Sabemos que $O_i \cap \mathrm{Fix}(C) \cap C = O_i$ ou $O_i \cap \mathrm{Fix}(\Gamma_C) \cap C = \varnothing$ para cada $i \in \{1, ..., k\}$. Se acontecer que $O_i \cap \mathrm{Fix}(\Gamma_C) \cap C = O_i$ então $O_i = \mathcal{O}_j$ para algum $j \in \{1, 2, 3\}$. Assim, no máximo as três órbitas em $\mathrm{Fix}(\Gamma_C) = \mathcal{O}_1 \cup \mathcal{O}_2 \cup \mathcal{O}_3$ podem aparecer na decomposição em órbitas de C (isto, vai depender da quantidade de pontos em $\mathrm{Fix}(\Gamma_C) \cap C$). Seja $\delta = \left|\{i \mid O_i \cap \mathrm{Fix}(\Gamma_C) \cap C = \varnothing\}\right|$, note que para cada um desses índices $|O_i| = |\Gamma_C|$. Agora, como $|C| = d = |O_1| + \cdots + |O_k|$, segue que:

$$d = \alpha|\mathcal{O}_1| + \beta|\mathcal{O}_2| + \gamma|\mathcal{O}_3| + \delta|\Gamma_C| \quad \text{com} \quad \alpha, \beta, \gamma \in \{0, 1\} \ \text{ e } \ \delta \geqslant 0.$$ ∎

O próximo Teorema contém os resultados mais importantes para a determinação de Γ_C (a menos de isomorfismo). Sua demonstração tem como base o Teorema de classificação de Klein (**TCK**), a Proposição 3.4 e a equação (5) acima.

Teorema 3.1. *Seja C um subconjunto de $\mathbb{P}^1$ com d pontos ($d \geqslant 3$). Verifica-se que:*

(i) *Se Γ_C for cíclico então $|\Gamma_C| \leqslant d$.*

(ii) *Se* $\Gamma_C \cong D_k$ *o grupo diedral, sendo* $k \geqslant 2$, *então* $k \,|\, d$ *ou* $k \,|\, (d-2)$. *Em particular,* $k \leqslant d$.

(iii) *Se* $C = \{[1 : \xi^j]\}_{j=1}^{d}$, *sendo* ξ *uma raiz primitiva* d*-ésima da unidade, então* $\Gamma_C \cong D_d$.

(iv) *Se* d *é ímpar e* Γ_C *não é cíclico, então* $\Gamma_C \cong D_k$ *para algum* k, $3 \leqslant k \leqslant d$.

Demonstração. De acordo com a tabela (3), considere:

$$[u : v] \overset{\mathbf{R}_a}{\longmapsto} [au : v] \quad \text{e} \quad [u : v] \overset{\mathbf{P}_a}{\longmapsto} [av : u]$$

com $a \in \mathbb{C}$ não nulo. Por simplicidade, vamos escrever $\mathbf{P}$ em lugar de $\mathbf{P}_1$ (ou seja, $\mathbf{P} = \mathbf{P_1}$).

(i) A menos de conjugação podemos assumir que $\Gamma_C = \langle \mathbf{R}_\kappa \rangle$ (sendo κ uma raiz primitiva da unidade). Como $d \geqslant 3$ existe $\mathrm{x} \in C$ que não é ponto fixo de $\mathbf{R}_\kappa$ (logo de nenhum elemento de $(\Gamma_C)^\star$), assim $|O_\mathrm{x}| = |\Gamma_C|$. Como $O_\mathrm{x} \subseteq C$, concluímos que $|\Gamma_C| \leqslant d = |C|$.

(ii) Se $\Gamma_C \cong D_k$, então segue de (4) e (5) que: $d = \alpha \cdot k + \beta \cdot k + \gamma \cdot 2 + \delta \cdot 2k$ com $\alpha, \beta, \gamma \in \{0, 1\}$ e $\delta \geqslant 0$ inteiro. Assim, se $\gamma = 0$ então k divide d, caso contrário ($\gamma = 1$), concluímos que k divide $d - 2$.

(iii) Se $d = 3$ então $\Gamma_C \cong D_3$ (veja exemplo 3.1). A seguir assuma que $d \geqslant 4$. Temos $C = \{[1 : \xi^j]\}_{j=0}^{d-1}$, sendo ξ uma raiz primitiva d-ésima da unidade. Observe que $\mathbf{R}_\xi, \mathbf{P} \in \Gamma_C$ e $\langle \mathbf{R}_\xi, \mathbf{P} \rangle$ é isomorfo a D_d (desde que $ord(\mathbf{R}_\xi) = d$, $ord(\mathbf{P}) = 2$ e $\mathbf{P} \circ \mathbf{R}_\xi \circ \mathbf{P} = \mathbf{R}_\xi^{d-1}$).

Assim, Γ_C não é cíclico e $|\Gamma_C| = 2dm$ para algum m natural.

Caso 1: $d \in \{4, 5\}$. Neste caso, segue do Teorema de Classificação de Klein (**TCK**) e do item (ii) (nesta Proposição) que:

d	Γ_C é isomorfo ao grupo	$\lvert\Gamma_C\rvert$
4	S_4 ou D_4	24 ou 8
5	A_5 ou D_5	60 ou 10

Se $\Gamma_C \cong A_4$ ou $\Gamma_C \cong S_5$ então a partir de (4) e (5), concluímos que $4 = \alpha 6 + \beta 8 + \gamma 12 + \delta 24$ ou $5 = \alpha 12 + \beta 20 + \gamma 30 + \delta 60$, sendo que essas equações não admitem soluções com $\alpha, \beta, \gamma \in \{0, 1\}$ e $\delta \geqslant 0$ inteiro. Portanto, $\Gamma_C \cong D_d$ para $d = 4, 5$.

Caso 2: $d \geqslant 6$. Observe que $\mathbf{R}_\xi \in \Gamma_C$ é um elemento de ordem $d \geqslant 6$ (e Γ_C não é cíclico). Assim, segue do Teorema de classificação de Klein (**TCK**) que $\Gamma_C \cong D_k$, com $k \leqslant d$. Portanto, $\Gamma_C \cong D_d$ (pois $|\Gamma_C| = 2dm = 2k \leqslant 2d$).

(iv) Se Γ_C não for cíclico, segue do Teorema de classificação de Klein (**TCK**) que $\Gamma_C \cong G$ com $G \in \{D_k, A_4, S_4, A_5\}$ e $k \geqslant 3$. Assim a ordem de Γ_C é par e caso $\Gamma_C \cong G$ com $G \in \{A_4, S_4, A_5\}$ segue de (4) que n_i é par para todo i. Agora, como $d = \alpha n_1 + \beta n_2 + \gamma n_3 + \delta |\Gamma_C|$ com $\alpha, \beta, \gamma \in \{0, 1\}$ e $\delta \geqslant 0$ inteiro (veja (5)), concluímos que d teria que ser par. Assim $\Gamma_C \cong D_k$ para algum $k \geqslant 3$. De fato, segue do item (ii) desta Proposição que $k \leqslant d$. ∎

Finalmente temos a base para determinar a ordem dos grupos Γ_C, se C for um subconjunto finito da reta projetiva. Assim, na próxima e última seção retornamos ao mundo geométrico (revisando inicialmente a contagem de retas em superfícies

projetivas não singulares no espaço projetivo $\mathbb{P}^3$, para logo associar a cada subconjunto C de $\mathbb{P}^1$ constituído por d pontos, uma superfície não singular de grau d em $\mathbb{P}^3$, continuando com a contagem das retas contidas nessas superfícies). Concluímos com a estimativa de N_d em (12), referente à quantidade máxima de retas que tais superfícies podem conter.

4. Contagem de retas em superfícies projetivas

Como já comentamos na introdução, uma pergunta clássica em Geometria Algébrica é:

Qual é a quantidade máxima de retas que uma superfície projetiva não singular de grau d em $\mathbb{P}^3$ contém?

Agora, já estamos prontos para saber a resposta.

Como vai essa contagem?

- $d \in \{1, 2\}$: Planos e superfícies quádricas contêm infinitas retas.
- $d = 3$: As superfícies cúbicas não singulares contêm exatamente 27 retas. A solução neste caso foi encontrada por *Cayley* & *Salmon* em 1849 ([6], [18]). O cálculo de tal número (de forma bastante elementar, para quem deseja se aventurar pelas trilhas da geometria enumerativa) pode ser encontrado nos textos [16] e [4].
- $d = 4$: Tais superfícies contêm no máximo 64 retas. Neste caso, é importante destacar que em 1882 *Friedrich Schur* exibiu (em [19]) uma superfície não singular de grau 4 que contém exatamente 64 retas, hoje denominada *Quártica de Schur* em sua homenagem. E somente em 1943, *Beniamino Segre*, mostrou que tais superfícies contêm no máximo 64 retas ([20]). Recentemente (2015), foi constatado um erro na abordagem utilizada por *Segre* em [20], por *Rams* & *Schütt*, segundo os quais, após fazer a correção nas premissas usadas por *Segre*, ele só poderia concluir que tais superfícies contêm no máximo 72 retas ([14]). De fato (em [14]), os autores corrigem o erro de *Segre* e além disso, mostram que uma superfície não singular de grau 4, definida por um polinômio com coeficientes num corpo $\mathbb{K}$, com característica diferente de 2 e 3, ainda contém no máximo 64 retas.
- $d \geqslant 5$: Ainda é um problema em aberto.

Uma das formas de abordar essa questão (no caso $d \geqslant 5$) é através do cálculo de cotas inferiores, ou seja, explicitar superfícies não singulares de grau d e contar as retas que tais superfícies contêm.

Por exemplo, a superfície de Fermat $x_0^d + x_1^d + x_2^d + x_3^d = 0$ contém exatamente $3d^2$ retas para todo $d \geqslant 3$. Logo o número procurado é maior ou igual a $3d^2$.

As superfícies de Fermat fazem parte de uma família de superfícies que podem ser construídas a partir d pontos distintos na reta projetiva, e as descrevemos a seguir.

Superfícies definidas a partir de d pontos distintos em $\mathbb{P}^1$. O leitor deve ter percebido que ao considerarmos um polinômio homogêneo f no qual, por exemplo, só comparecem duas variáveis, ao escrevermos $\mathcal{Z}(f)$ podemos de fato querer determinar os zeros de f em $\mathbb{P}^1$; ou em algum outro $\mathbb{P}^n$ para $n \geqslant 2$. Por este motivo, no que

segue do texto usaremos a notação $\mathcal{Z}_{\mathbb{P}^1}(f)$ para indicar o cálculo dos zeros de um polinômio homogêneo f, no qual comparecem duas variáveis (não necessariamente, x_0 e x_1). E deixamos a notação $\mathcal{Z}(f)$ para indicar os zeros em $\mathbb{P}^3$ de polinômios homogêneos em $\mathbb{C}[x,y,z,t]$.

Construção de uma superfícies a partir d pontos distintos em $\mathbb{P}^1$.

A essência desta construção repousa no seguinte fato: se C for um subconjunto da reta projetiva que consiste de d pontos distintos, então existe $\phi \in \mathbb{C}[u,v]$ polinômio homogêneo[18] de grau d, cujo conjunto de zeros (ou raízes) em $\mathbb{P}^1$ é igual ao conjunto C, ou seja, $\mathcal{Z}_{\mathbb{P}^1}(\phi) = C$. Por exemplo, se $C = \{[1 : \xi^j]\}_{j=1}^{d}$ sendo ξ uma raiz primitiva d-ésima da unidade, então $\mathcal{Z}_{\mathbb{P}^1}(u^d - v^d) = C$.

A tais $\phi \in \mathbb{C}[u,v]$ faremos corresponder a superfície de grau d em $\mathbb{P}^3$

$$\mathcal{S}_\phi = \mathcal{Z}\left(\phi(x,y) - \phi(z,t)\right).$$

$\mathcal{S}_\phi$ é não singular. Se $d = 1$, então $\mathcal{S}_\phi$ é um plano em $\mathbb{P}^3$, logo não singular[19].

Assuma que $d \geqslant 2$ e, suponha (por absurdo) que $\mathrm{p} = [\alpha_0 : \alpha_1 : \beta_0 : \beta_1]$ é um ponto singular da superfície $\mathcal{S}_\phi$. Assim, ao considerar $f = \phi(x,y) - \phi(z,t) \in \mathbb{C}[x,y,z,t]$ tem-se que

$$\begin{aligned} \partial_x f(\mathrm{p}) = \partial_u \phi(\alpha_0,\alpha_1) = 0, \qquad \partial_y f(\mathrm{p}) = \partial_v \phi(\alpha_0,\alpha_1) = 0, \\ \partial_z f(\mathrm{p}) = \partial_u \phi(\beta_0,\beta_1) = 0, \qquad \partial_t f(\mathrm{p}) = \partial_v \phi(\beta_0,\beta_1) = 0. \end{aligned}$$

Agora, segue da identidade de Euler $\partial_u\phi \cdot u + \partial_v\phi \cdot v = d \cdot \phi$ que $\phi(\alpha_0,\alpha_1) = 0$ e $\phi(\beta_0,\beta_1) = 0$. Portanto, como estamos assumindo que $\mathrm{p} = [\alpha_0 : \alpha_1 : \beta_0 : \beta_1]$ é um ponto singular da superfície $\mathcal{S}_\phi$ então alguma das coordenadas de p é não nula, ou seja, $[\alpha_0 : \alpha_1] \in C = \mathcal{Z}_{\mathbb{P}^1}(\phi)$ ou $[\beta_0 : \beta_1] \in C = \mathcal{Z}_{\mathbb{P}^1}(\phi)$.

Suponha que $[\alpha_0 : \alpha_1] \in C = \mathcal{Z}_{\mathbb{P}^1}(\phi)$. Assim, a menos de uma reordenação no conjunto dos pontos em C, podemos escrever $\phi = L_1 \cdot L_2 \cdots L_d$ sendo $L_1 = \alpha_1 u - \alpha_0 v$ e $L_i = b_i u - a_i v$ para $i = 2, ..., d$. Assim, tendo em consideração que $L_1(\alpha_0,\alpha_1) = 0$, concluímos que

$$\begin{aligned} \partial_u\phi(\alpha_0,\alpha_1) = 0 = \alpha_1 L_2 \cdots L_d(\alpha_0,\alpha_1), \\ \partial_u\phi(\alpha_0,\alpha_1) = 0 = -\alpha_0 L_2 \cdots L_d(\alpha_0,\alpha_1). \end{aligned} \tag{6}$$

Como os pontos do conjunto C são distintos e $\mathcal{Z}_{\mathbb{P}^1}(L_i) = \{\mathrm{p}_i\}$, com $\mathrm{p}_i \neq \mathrm{p}_j$ para todo $i \neq j$, segue que $L_2 \cdots L_d(\alpha_0,\alpha_1) \neq 0$ (visto que $\mathrm{p}_1 = [\alpha_0 : \alpha_1]$). Assim, segue de (6) que $\alpha_0 = 0 = \alpha_1$ (o que é um absurdo). Analogamente, chegamos num absurdo ao supor que $[\beta_0 : \beta_1] \in C = \mathcal{Z}_{\mathbb{P}^1}(\phi)$. Portanto, $\mathcal{S}_\phi$ não possui pontos singulares.

Esquema da contagem das retas na superfície $\mathcal{S}_\phi$. Para calcularmos

$$n_\phi = \left| \left\{ \text{ retas contidas em } \mathcal{S}_\phi \right\} \right|,$$

[18] Note que tais polinômios são unicamente determinados a menos de um múltiplo escalar não nulo. Ou seja, $\mathcal{Z}_{\mathbb{P}^1}(\phi) = \mathcal{Z}_{\mathbb{P}^1}(\phi_1) \Longleftrightarrow \phi = \lambda\phi_1$, com $\lambda \in \mathbb{C}$ não nulo.

[19] Se $d = 1$ então $\phi = b_1 u - a_1 v \in \mathbb{C}[u,v]$ com $\mathrm{p}_1 = [a_1 : b_1] \in \mathbb{P}^1$ e $\mathcal{S}_\phi = \mathcal{Z}(b_1 x - a_1 y - b_1 z + a_1 t)$.

considere a seguinte partição do conjunto das retas contidas na superfície $\mathcal{S}_\phi$.

$$\left\{\ell \mid \ell \overset{reta}{\subset} \mathcal{S}_\phi\right\} = \underbrace{\left\{\ell \mid \ell \cap L \neq \varnothing\right\}}_{\mathcal{L}_\phi} \bigcup \underbrace{\left\{\ell \mid \ell \cap L = \varnothing\right\}}_{\mathcal{L}'_\phi} \tag{7}$$

sendo $L = \mathcal{Z}(z,t) \subset \mathbb{P}^3$.

A seguir, assuma que $\mathcal{Z}_{\mathbb{P}^1}(\phi) = \left\{[a_i : b_i]\right\}_{i=1}^{d}$ e considere os pontos p_1,..., p_d e q_1,..., q_d em $\mathbb{P}^3$ dados por $\mathrm{p}_i = [a_i : b_i : 0 : 0]$ e $\mathrm{q}_i = [0 : 0 : a_i : b_i]$.

Pontos chave para determinar n_ϕ.

Denotemos por $\ell_{i,j}$ a reta que passa pelos pontos p_i e q_j.

(i) A reta $\ell_{i,j}$ está contida na superfície $\mathcal{S}_\phi$ para todo i, j.

(ii) $\mathcal{L}_\phi = \left\{\ell_{i,j} \mid i,j \in \{1,...,d\}\right\}$. Logo, $|\mathcal{L}_\phi| = d^2$.

(iii) $|\mathcal{L}'_\phi| = d|\Gamma_C|$, sendo $C = \mathcal{Z}_{\mathbb{P}^1}(\phi)$ e $\Gamma_C = \left\{\mathbf{T} \in \mathrm{Aut}(\mathbb{P}^1) \mid \mathbf{T}(C) = C\right\}$.

Portanto,

$$n_\phi = d^2 + d|\Gamma_C|.$$

Linhas para demonstrar que $n_\phi = d^2 + d|\Gamma_C|$**.** A seguir demonstraremos as afirmações (i) e (ii) que fazem parte do esquema de contagem das retas na superfície $\mathcal{S}_\phi$. Entretanto, para o item (iii) só faremos um esboço com as principais ideias que são usadas para chegar nesse resultado. Para mais detalhes, consultar a dissertação [22].

Considere $\mathcal{S}_\phi = \mathcal{Z}(\phi(x,y) - \phi(z,t))$ não singular de grau d tal que $\mathcal{Z}_{\mathbb{P}^1}(\phi) = \left\{[a_i : b_i]\right\}_{i=1}^{d}$. Seja $\ell_{i,j} \subset \mathbb{P}^3$ a reta determinada por $\mathrm{p}_i = [a_i : b_i : 0 : 0]$ e $\mathrm{q}_j = [0 : 0 : a_j : b_j]$, com $i, j \in \{1,...,d\}$.

Afirmação 1: $\ell_{i,j} \subset \mathcal{S}_\phi$.

Se $\mathrm{p} \in \ell_{i,j}$, então $\mathrm{p} = [u_0 a_i : u_0 b_i : u_1 a_j : u_1 b_j]$ com $u_0, u_1 \in \mathbb{C}$ não ambos nulos. Assim, se $f = \phi(x,y) - \phi(z,t)$, tem-se que

$$f(\mathrm{p}) = u_0^d \phi(a_i, b_i) - u_1^d \phi(a_j, b_j) = 0, \text{ pois } \mathcal{Z}_{\mathbb{P}^1}(\phi) = \left\{[a_i : b_i]\right\}_{i=1}^{d} \Longrightarrow \mathrm{p} \in \mathcal{S}_\phi.$$

Afirmação 2: $\mathcal{L}_\phi := \left\{\ell \subset \mathcal{S}_\phi \mid \ell \cap L \neq \varnothing\right\} = \left\{\ell_{i,j} \mid i,j \in \{1,...,d\}\right\}$. Logo, $|\mathcal{L}_\phi| = d^2$.

Observe que $\mathrm{p}_i \in \ell_{i,j} \cap L$, logo $\ell_{i,j} \in \mathcal{L}_\phi$ (visto que $\ell_{i,j} \subset \mathcal{S}_\phi$). A outra inclusão é consequência da Proposição que mostraremos a seguir.

Proposição 4.1. *Sejam $L = \mathcal{Z}(z,t)$ e $M = \mathcal{Z}(x,y)$ retas em $\mathbb{P}^3$. Se ℓ é uma reta contida na superfície $\mathcal{S}_\phi$ com $d \geqslant 2$, então*

$$\ell \cap M \neq \varnothing \iff \ell \cap L \neq \varnothing.$$

Demonstração. Considere $f = \phi(x,y) - \phi(z,t)$ de grau $d \geqslant 2$ tal que $\mathcal{S}_\phi = \mathcal{Z}(f) \subset \mathbb{P}^3$. A seguir mostraremos a implicação direta, sendo a recíproca análoga.

Assuma que $\ell = \mathbb{P}(W)$. Como M não está contida em $\mathcal{S}_\phi$ e $\ell \cap M \neq \varnothing$, então $\ell \cap M = \{[\mathrm{w_M}]\}$ para algum $\mathrm{w_M} \in W$. A seguir, escolha $\mathrm{v} \in \mathbb{C}^4$ tal que $\{\mathrm{w_M}, \mathrm{v}\}$ seja uma base de W. Note que podemos escrever de forma única $\mathrm{v} = \mathrm{v_L} + \mathrm{v_M}$, com $[\mathrm{v_M}] \in M$ e $[\mathrm{v_L}] \in L$.

Suponha que $\{\mathrm{w_M}, \mathrm{v_M}\}$ é L.I. Assim, $\mathrm{w_M} = (0,0,\alpha,\beta)$ e $\mathrm{v_M} = (0,0,\gamma,\delta)$ com $\alpha\delta - \gamma\beta \neq 0$. A seguir, considere $T \in \mathrm{Aut}(\mathbb{C}^2)$ dada por $T(x,y) = (\alpha x + \gamma y, \beta x + \delta y)$. Agora, como $\ell \subset \mathcal{S}_\phi$ e todo ponto da reta ℓ é da forma $[u_0\mathrm{w_M} + u_1(\mathrm{v_L} + \mathrm{v_M})]$ com $\mathrm{u} = [u_0 : u_1] \in \mathbb{P}^1$, temos que:

$$\begin{aligned} f(u_0\mathrm{w_M} + u_1(\mathrm{v_L} + \mathrm{v_M})) = 0 &\iff {u_1}^d\phi(\mathrm{v_L}) = \phi(u_0\mathrm{w_M} + u_1\mathrm{v_M}) \\ &\iff {u_1}^d\phi(\mathrm{v_L}) = \phi(T(u_0,u_1)) \\ &\iff \lambda(\alpha y - \beta x)^d\phi(\mathrm{v_L}) = \phi(x,y) \\ &\qquad \text{com } \lambda = (\alpha\delta - \gamma\beta)^{-d} \text{ e } (x,y) = T(u_0,u_1). \end{aligned}$$

Analisando a última igualdade chegamos a um absurdo. De fato, se $\phi(\mathrm{v_L}) = 0$ então $\phi = 0$, do contrário temos que ϕ não é livre de quadrados (o que não ocorre pois $\mathcal{S}_\phi$ é não singular de grau $d \geqslant 2$).

Portanto, $\{\mathrm{w_M}, \mathrm{v_M}\}$ é L.D., logo $\mathrm{v_M} = \lambda\mathrm{w_M}$, para algum $\lambda \neq 0$. Agora, tendo em consideração que $\{\mathrm{w_M}, \mathrm{v} = \mathrm{v_L} + \mathrm{v_M}\}$ é base de W, concluímos que $\mathrm{v_L}$ é não nulo. Note que $\mathrm{v_L} = \mathrm{v} - \lambda\mathrm{w_M} \in W$ é não nulo, logo determina o ponto $[\mathrm{v_L}] \in \ell \cap L$. ∎

Corolário 4.1. *Se ℓ for uma reta contida na superfície $\mathcal{S}_\phi$ tal que $\ell \cap L \neq \varnothing$, então $\ell = \ell_{i,j}$ para algum $i, j \in \{1, ..., d\}$.*

Demonstração. Como a reta L não está contida na superfície $\mathcal{S}_\phi$, então $L \neq \ell$. Assim, $\ell \cap L$ consiste de um único ponto $\mathrm{p} = [a : b : 0 : 0]$ com a e b complexos não ambos nulos. Entretanto, a condição $\ell \subset \mathcal{S}_\phi$ nos leva à conclusão $\phi(a,b) = 0$. Assim, $\mathrm{p} = \mathrm{p}_i$ para algum i.

Por outro lado, a Proposição 4.1 implica em $\ell \cap M \neq \varnothing$. Como a reta M também não esta contida em $\mathcal{S}_\phi$, o raciocínio anterior nos permite concluir que $\ell \cap M = \{\mathrm{q}_j\}$ para algum j. Portanto, $\ell = \ell_{i,j}$ (ou seja, ℓ é determinada pelos pontos p_i e q_j). ∎

Corolário 4.2. *Com as notações em (7). Tem-se que $|\mathcal{L}_\phi| = d^2$ e $n_\phi = d^2 + |\mathcal{L}'_\phi|$.*

Afirmação 3: $|\mathcal{L}'_\phi| = d|\Gamma_C|$, sendo $C = \mathcal{Z}_{\mathbb{P}^1}(\phi)$ e

$$\Gamma_C = \Big\{ \mathbf{T} \in \mathrm{Aut}(\mathbb{P}^1) \mid \mathbf{T}(C) = C \Big\}.$$

A ideia chave para demonstrar a Afirmação 3 é associar a cada reta $\ell \in \mathcal{L}'_\phi$ a um automorfismo em Γ_C, como faremos a seguir.

Proposição 4.2. *Se $\ell \in \mathcal{L}'_\phi$, então ℓ induz um automorfismo $\mathbf{T}_\ell \in \Gamma_C$, no qual $C = \mathcal{Z}_{\mathbb{P}^1}(\phi)$.*

Demonstração. Sendo ℓ uma reta em $\mathbb{P}^3$, existem f_1 e f_2 homogêneos de grau 1 e L.I. tais que $\ell = \mathcal{Z}(f_1, f_2)$. Se $f_i = a_ix + b_iy + c_iz + d_it$, então $\ell \cap L = \mathcal{Z}(f_1, f_2, z, t) = \mathcal{Z}(a_1x + b_1y, a_2x + b_2y, z, t)$. Agora, a condição $\ell \cap L = \varnothing$ assegura que o sistema $a_1x + b_1y = 0, a_2x + b_2y = 0$, possui solução única $x = y = 0$, e portanto $a_1b_2 - a_2b_1 \neq 0$. Assim, podemos escolher para ℓ equações da forma:

$x = \alpha z + \beta t$ e $y = \gamma z + \delta t$, com α, β, γ e δ em $\mathbb{C}$. Além disso, a condição $\ell \cap M = \varnothing$ (Proposição 4.1) nos garante que $\alpha\delta - \beta\gamma \neq 0$.

Assim, obtemos $\mathbf{T}_\ell \in \mathrm{Aut}(\mathbb{P}^1)$ dada por $[z : t] \mapsto [\alpha z + \beta t : \gamma z + \delta t]$. Observe que $[\alpha c + \beta d : \gamma c + \delta c : c : d] \in \ell$ para todo $[c : d] \in \mathbb{P}^1$. Entretanto, se considerarmos $[c : d] \in \mathcal{Z}_{\mathbb{P}^1}(\phi)$ (neste caso $\phi(c, d) = 0$) concluímos que $\phi(\alpha c + \beta d, \gamma c + \delta c) = 0$, pois $\ell \subset \mathcal{S}_\phi$. Portanto, $\mathbf{T}_\ell([c : d]) \in \mathcal{Z}_{\mathbb{P}^1}(\phi)$. Finalmente, como $\mathbf{T}_\ell$ é uma bijeção e $|\mathcal{Z}_{\mathbb{P}^1}(\phi)| = d$, o resultado segue. ∎

Proposição 4.3. *Notações como na Proposição 4.2. A função $\Omega : \mathcal{L}'_\phi \longrightarrow \Gamma_C$ dada por $\ell \mapsto \mathbf{T}_\ell$ é sobrejetora e $|\Omega^{-1}(\mathbf{T})| = d$ para todo $\mathbf{T} \in \Gamma_C$.*

Demonstração. Considere $\mathbf{T} \in \Gamma_C$ dado por $\mathbf{T}([z : t]) = [\alpha z + \beta t : \gamma z + \delta t]$.

O primeiro passo é associar ao automorfismo $\mathbf{T}$ a superfície quádrica

$$Q_{\mathbf{T}} = \mathcal{Z}(x(\gamma z + \delta t) - y(\alpha z + \beta t)) \subset \mathbb{P}^3.$$

A seguir, citamos (sem demonstrar) os fatos essenciais que relacionam a superfície quádrica $Q_{\mathbf{T}}$ com a contagem das retas na superfície $\mathcal{S}_\phi$.

Fato 1: *$Q_{\mathbf{T}}$ é uma superfície quádrica não singular.*[20]

Sabe-se que toda superfície quádrica não singular em $\mathbb{P}^3$ possui exatamente duas famílias de retas (p. 66 e 67 em [1], p. 406 em [7]). No caso da superfície quádrica $Q_{\mathbf{T}}$ essas famílias são dadas por:

Família $\mathcal{L}$: Para cada $\mathrm{y} = [c : d] \in \mathbb{P}^1$, $\mathcal{L}_\mathrm{y}$ é definida pelas equações:

$$(\gamma c + \delta d)x - (\alpha c + \beta d)y = 0 \quad \text{e} \quad dz - ct = 0. \tag{8}$$

Família $\mathcal{M}$: Para cada $\mathrm{x} = [a : b] \in \mathbb{P}^1$, $\mathcal{M}_\mathrm{x}$ é definida pelas equações:

$$ax = b(\alpha z + \beta t) \quad \text{e} \quad ay = b(\gamma z + \delta t). \tag{9}$$

Note que $\mathcal{M}_{[0:1]} = L = \mathcal{Z}(z, t)$ e $\mathcal{M}_{[1:0]} = M = \mathcal{Z}(x, y)$.

Fato 2: *Seja $\mathcal{L}$ a família de retas em (8). Temos que:*
(i) *$\mathcal{L}_\mathrm{y} \cap M \neq \varnothing$ para todo $\mathrm{y} \in \mathbb{P}^1$.*
(ii) *$\mathcal{L}_\mathrm{y} \subset \mathcal{S}_\phi$ se, e somente se, $\mathrm{y} \in \mathcal{Z}_{\mathbb{P}^1}(\phi)$. Assim, existem d retas (distintas) da família $\mathcal{L}$ contidas em $Q_{\mathbf{T}} \cap \mathcal{S}_\phi$ (visto que $\mathcal{L}_\mathrm{y} \in \mathcal{L}_\phi$).*

Fato 3: *Existem exatamente d retas da família $\mathcal{M}$ (em (9)) contidas em $Q_{\mathbf{T}} \cap \mathcal{S}_\phi$ e tais retas pertencem a $\Omega^{-1}(\mathbf{T})$.*

Fato 4: *Sejam $L_0, ..., L_{d-1}$ e $M_0, ..., M_{d-1}$ as únicas retas da família $\mathcal{L}$ e $\mathcal{M}$, respectivamente, contidas em $\mathcal{S}_\phi \cap Q_{\mathbf{T}}$. Então*

$$\mathcal{S}_\phi \cap Q_{\mathbf{T}} = L_0 \cup \cdots \cup L_{d-1} \cup M_0 \cup \cdots \cup M_{d-1}.$$

Observe que o Fato 3 nos garante que Ω é sobrejetora e $|\Omega^{-1}(\mathbf{T})| \geqslant d$.

[20] Note que o sistema $\partial_x Q_{\mathbf{T}} = \gamma z + \delta t = 0$, $\partial_y Q_{\mathbf{T}} = -(\alpha z + \beta t) = 0$, $\partial_z Q_{\mathbf{T}} = \gamma x - \alpha y = 0$, $\partial_t Q_{\mathbf{T}} = \delta x - \beta y = 0$, só admite a solução trivial $x = y = z = t = 0$, pois $\alpha\delta - \beta\gamma \neq 0$.

Se $\ell \in \Omega^{-1}(\mathbf{T})$ então $\mathbf{T}_\ell = \mathbf{T}$, e a partir de (i) na Proposição 3.2 concluímos que $Q_{\mathbf{T}_\ell} = Q_{\mathbf{T}}$, que implica em $\ell \subset Q_{\mathbf{T}} \cap \mathcal{S}_\phi$ com $\ell \in \mathcal{L}'_\phi$. Como $\{M_0, ..., M_{d-1}\}$ são as únicas retas que pertencem a $\mathcal{L}'_\phi$ e estão contidas em $Q_{\mathbf{T}} \cap \mathcal{S}_\phi$, a partir do Fato 4 concluímos que $\ell \in \{M_0, ..., M_{d-1}\}$. ∎

Corolário 4.3. *Verifica-se que* $|\mathcal{L}'_\phi| = d|\Gamma_C|$ *com* $C = \mathcal{Z}_{\mathbb{P}^1}(\phi)$.
Demonstração. A sobrejetividade de Ω nos garante que $\mathcal{L}'_\phi = \bigcup_{\mathbf{T}\in\Gamma_C} \Omega^{-1}(\mathbf{T})$, sendo esta união disjunta (e finita), conclui-se que $|\mathcal{L}'_\phi| = \sum_{\mathbf{T}\in\Gamma_C} |\Omega^{-1}(\mathbf{T})| = d|\Gamma_C|$. ∎

Sobre a estimativa do valor máximo de n_ϕ. Tendo em consideração que $n_\phi = d^2 + d|\Gamma_C|$, segue-se que o valor máximo de n_ϕ depende da ordem do subgrupo Γ_C, ou de forma mais explicita, das escolhas de d pontos distintos em $\mathbb{P}^1$.

Sobre o valor máximo de $|\Gamma_C|$. Lembre que $|C| = d \geqslant 3$. De acordo com a paridade de d, temos:

- d ímpar: A Proposição 3.1 nos garante que: $|\Gamma_C| \leqslant d$ se Γ_C for cíclico, caso contrário $\Gamma_C \cong D_k$ com $k \leqslant d$. Portanto, para d ímpar o valor máximo de $|\Gamma_C|$ é $2d$ (basta escolher $C = \{[1 : \xi^j]\}_{j=1}^d$, sendo ξ uma raiz primitiva d-ésima da unidade).
- d par: Segue da Proposição 3.1 e do Teorema de Classificação de Klein que o valor máximo de $|\Gamma_C|$ pertence ao conjunto $\{2d, 12, 24, 60\}$.
  Assim, para $d \geqslant 30$ o valor máximo de $|\Gamma_C|$ é $2d$.
  Agora, vamos analisar o caso d par, para $4 \leqslant d \leqslant 28$.
  Note que, o Teorema de Classificação de Klein (**TCK**) nos garante que $\Gamma_C \cong G$ com $G \in \{D_k, A_4, S_4, A_5\}$[21]. Neste caso, segue de (4) e (5) que:

(10)

G	$d = \alpha n_1 + \beta n_2 + \gamma n_3 + \delta\lvert G\rvert$
D_k	$d = \alpha k + \beta k + \gamma 2 + \delta 2k$
A_4	$d = \alpha 6 + \beta 4 + \gamma 4 + \delta 12$
S_4	$d = \alpha 12 + \beta 8 + \gamma 6 + \delta 24$
A_5	$d = \alpha 30 + \beta 20 + \gamma 12 + \delta 60$

com $\alpha, \beta, \gamma \in \{0, 1\}$ e $\delta \geqslant 0$ inteiro.

Se considerarmos somente o grupo diedral de ordem máxima, ao qual Γ_C pode ser isomorfo, segue do item (ii) na Proposição 3.1 e da tabela em (10) que:

(11)

$4 \leqslant d \leqslant 28$ par	Γ_C pode ser isomorfo a	ordem máxima de Γ_C
$d = 4$	A_4, D_4	12
$d \in \{6, 8\}$	D_d, A_4, S_4	24
$d = 10$	D_{10}, A_4	20
$d = 12$	D_{12}, A_4, S_4, A_5	60
$d \neq 20, 14 \leqslant d \leqslant 28$	D_d, A_4, S_4	2d
$d = 20$	D_{20}, A_5	60

[21] Ao procurarmos o valor máximo de $|\Gamma_C|$, descartamos o caso em que Γ_C é cíclico.

Por simplicidade, no que segue do texto usaremos a notação: $\infty = [0:1]$ e $a = [1:a]$ com $a \in \mathbb{C}$.

$\underline{d=4}$ A partir de (11) temos que $\Gamma_C \cong D_4$ ou $\Gamma_C \cong A_4$. Se escolhermos $C = \{0, 1, \omega, \omega^2\}$, sendo ω raiz cúbica primitiva da unidade. Tem-se que $[u:v] \longmapsto [\omega u : v]$ é um elemento de ordem 3 de Γ_C, logo $\Gamma_C \cong A_4$ e a ordem máxima é 12.

Para os outros valores de d, nos remetemos a citar a escolha do conjunto C tal que $|\Gamma_C|$ atinge o valor máximo listado em (11).

$\underline{d=6}$ Considere $C = \{\infty, 0, 1, i, -1, -i\}$ sendo $i \in \mathbb{C}$ raiz quarta primitiva da unidade.

$\underline{d=8}$ Considere $C = \{\kappa, i\kappa, -\kappa, -i\kappa, \kappa^{-1}, i\kappa^{-1}, -\kappa^{-1}, -i\kappa^{-1}\}$, sendo $\kappa \in \mathbb{C}$ uma raiz da equação $x^2 - (i+1)x - i = 0$.

$\underline{d=10}$ Considere C sendo o conjunto das raízes 10-ésimas da unidade.

$\underline{d=12}$ Seja ω raiz quinta primitiva da unidade, $\theta = \omega^3 + \omega^2$ e considere

$$C = \{\infty, 0, \theta, \theta\omega, \theta\omega^2, \theta\omega^3, \theta\omega^4, -\theta^{-1}, -\theta^{-1}\omega, -\theta^{-1}\omega^2, -\theta^{-1}\omega^3, -\theta^{-1}\omega^4\}.$$

$\underline{d \neq 20, 14 \leqslant d \leqslant 28 \text{ par.}}$ Considere C sendo o conjunto das raízes d-ésimas da unidade.

$\underline{d=20}$ Veja [22].

Assim, se $N_d = n_\phi$ com ϕ de grau d escolhido de modo que $|\Gamma_C|$ tenha a maior ordem possível, então tem-se que:

$$\begin{cases} N_d = 3d^2, & \text{se } d \notin \{4, 6, 8, 12, 20\}, \\ N_4 = 64,\ N_6 = 180,\ N_8 = 256,\ N_{12} = 864 \text{ e } N_{20} = 1600. \end{cases} \tag{12}$$

Agradecimentos. Agradecemos a *Samuel Boissière* (Laboratoire de Mathématiques et Applications - Université de Poitiers) e a *Alessandra Sarti* (LMA - Université de Poitiers) pela disponibilidade em nos ajudar na compreensão deste belo assunto. E ao professor *Marivaldo Pereira Matos* (DM-UFPB), pela leitura do texto original. E de forma muito especial ao revisor pelas sugestões e comentários que foram de fundamental importância para o aprimoramento desse texto.

Referências

[1] Ramón Orestes Mendoza Ahumada e Jacqueline Fabiola Rojas Arancibia, *Álgebra Linear e o Problema das quatro retas do Cálculo de Schubert*, Revista Matemática Universitária, **45**, pp. 55–69, 2009.

[2] Mark Anthony Armstrong, *Groups and symmetry*, Undergraduate Texts in Mathematics. Springer-Verlag, New York, 1988.

[3] Michael Artin, *Algebra*, Pearson, 2010.

[4] Cibelle de Fátima Castro de Assis, *As 27 Retas Sobre uma Superfície Cúbica não Singular*, Dissertação de Mestrado, DM–UFPB, 2004.

[5] Samuel Boissière e Alessandra Sarti, *Counting lines on surfaces*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. 5, pp. 39–52, 2007.

[6] Arthur Cayley, *On the triple tangent planes of surfaces of the third order*, Cambridge and Dublin Math. J., **4**, pp. 118–138, 1849.

[7] David Vassall Cox, John Little e Donal O'Shea, *Ideals, Varieties and Algorithms*, Springer-Verlag, New York, 1996.
[8] Igor V. Dolgachev , *McKay correspondence*, http://www.math.lsa.umich.edu/~idolga/McKaybook.pdf, 2009. Acessado em 03/08/2020.
[9] Leonhard Paul Euler, *Problema algebraicum ob affectiones prorsus singulares memorabile*, Novi Commentarii academiae scientiarum Petropolitanae, Volume 15, pp. 75–106, 1771, reimpresso em L. Euleri Opera Omnia, 1st series, Vol. 6, pp. 287–315.
[10] Arnaldo Garcia e Yves Lequain, *Elementos de Álgebra*, Projeto Euclides, IMPA, 2003.
[11] Christian Felix Klein, *Lectures on the Icosahedron and the Solution of Equations of Fifth Degree*, Dover, New York, 1956.
[12] Eugene Klotz, *Os Sólidos Platônicos*, https://youtu.be/bZvVwXb0Ejw. Acessado em 29/03/2021.
[13] Graham J. Leuschk e Roger Wiegand, *Cohen-Macaulay representations* Mathematical Surveys and Monographs , Vol. 181. American Mathematical Society. Providence, RI, 2012.
[14] Sławomir Rams e Matthias Schütt, *64 Lines on smooth quartic surfaces*, Math. Ann. 362, no. 1-2, pp. 679–698, 2015.
[15] Eduardo Francisco Rêgo, *Tópicos de Geometria*, http://cmup.fc.up.pt/cmup/eerego/TG.pdf, 2004. Acessado em 29/03/2021.
[16] Miles Anthony Reid, *Undergraduate Algebraic Geometry*, Cambridge University Press, London Mathematical Society Student Texts 12, 1988.
[17] Benjamin Olinde Rodrigues, *Des lois géométriques qui régissent les déplacements d'un système solide dans l'espace, et de la variation des coordonnées provenant de ces déplacements considérés indépendamment des causes qui peuvent les produire*, Journal de Mathématiques 5, pp. 380–440, 1840.
[18] George Salmon, *On the triple tangent planes of surfaces of the third order*, Cambridge and Dublin Math. J., **4**, pp. 252–260, 1849.
[19] Friedrich Heinrich Schur, *Ueber eine besondere Classe von Flächen vierter Ordnung*, Math. Ann. 20, pp. 254-296, 1882.
[20] Beniamino Segre, *The maximum number of lines lying on a quartic surface*, Quart. J. Math. Soc. (2) **49**, pp. 86–96, 1943.
[21] Javier Carrasco Serrano, *Finite Subgroups of* $\mathrm{SL}(2,\mathbb{C})$ *and* $\mathrm{SL}(3,\mathbb{C})$, http://homepages.warwick.ac.uk/~masda/McKay/Carrasco_Project.pdf, Warwick, 2014. Acessado em 29/03/2021.
[22] Sally Andria Vieira da Silva, *Sobre o número máximo de retas em superfícies no espaço projetivo* $\mathbb{P}^3$, Dissertação de Mestrado, DM–UFPB, 2016. https://repositorio.ufpb.br/jspui/bitstream/tede/9272/2/arquivototal.pdf Acessado em 29/03/2021.
[23] Hermann Klaus Hugo Weyl, *Symmetry*, Princeton University Press, 1952.

CCEN - Departamento de Matemática - UFPB
Cidade Universitária - João Pessoa - PB
58051-900
Email address: jfra@academico.ufpb.br

Instituto de Matemática e Estatística -UFF
São Domingos - Niterói - RJ
24210-200
Email address: sally_andria@id.uff.br

RMU
REVISTA MATEMÁTICA UNIVERSITÁRIA

SBM
SOCIEDADE BRASILEIRA DE MATEMÁTICA

Revista Matemática Universitária, vol. 1, 2022
ISSN: 2675-5254 — **DOI:** 10.21711/26755254/rmu20224

PADRÕES ESPECIAIS DE DISTRIBUIÇÃO DOS NÚMEROS PRIMOS: O $n-$QUADRADO ZETA

JORGE A. J. AVILA, ELIELSOM D. MOREIRA, BIANCA F. GUIMARÃES

Resumo. A busca por padrões de números primos, isto é, saber como os números primos estão distribuídos entre os números naturais, sempre foi e será uns dos desafios mais fascinantes da matemática. Este trabalho apresenta um novo arranjo dos números naturais, sobre um quadrado, para identificar padrões especiais de distribuição dos números primos. Por ter características específicas é chamado de $n-$quadrado Zeta. Nele, é possível encontrar sequências "longas" de números primos (caminhos). Foram estudados dois tipos de caminhos: diagonal e vertical, e modelados por polinômios de 2° grau. Uma atenção especial foi dada ao caminho vertical $\mathcal{V}_{41}$ por possuir grande quantidade de números primos. Surpreendentemente, o polinômio v_{41}, o qual modela o caminho vertical $\mathcal{V}_{41}$, resultou ser uma translação do polinômio de Euler. Finalmente, fizemos uma comparação do perfil do polinômio de Euler, tanto na espiral de Ulam, quanto no $n-$quadrado Zeta.

1. Introdução

A seguir, apresentamos alguns resultados preliminares.

Definição 1.1 (Número Primo)**.** Seja $p \in \mathbb{N}$, $p > 1$. Dizemos que p é um número primo se ele tem como únicos divisores positivos o 1 e ele mesmo.

Conjunto dos Números Primos: $\mathbb{P} = \{2, 3, 5, 7, 11, 13, 17, 19, 23, \dots\}$

Teorema 1.1. *A série Harmônica generalizada*

$$\sum_{n=1}^{\infty} \frac{1}{n^x} = 1 + \frac{1}{2^x} + \frac{1}{3^x} + \dots, \quad x \in \mathbb{R} \tag{1}$$

é convergente se $x > 1$ e divergente se $x \leqslant 1$.

Data de aceitação: 2 de maio de 2022.
Palavras chave. Números Primos, $\boldsymbol{n}-$quadrado Zeta, Espiral de Ulam, Polinômio de Euler, Polinômio $\boldsymbol{v_{41}}$.

Definição 1.2 (Igualdade Assintótica)**.** Sejam f e g funções reais de variável real.

$$f(x) \sim g(x) \iff \lim_{x\to\infty} f(x)/g(x) = 1 \tag{2}$$

onde, $f(x) \sim g(x)$ lê-se: $f(x)$ é assintoticamente igual a $g(x)$.

Definição 1.3 ("grande O")**.** Sejam f e g funções reais de variável real. $f(x)$ é "grande O" de $h(x)$, denotamos por $f(x) = O(h(x))$, se $\exists C > 0$ e um $x_0 \in \mathbb{R}$ tal que

$$|f(x)| \leqslant C|h(x)|, \quad \forall x \geqslant x_0$$

Definição 1.4 (Função Geradora de Números Primos)**.** Dizemos que $f : \mathbb{Z} \longrightarrow \mathbb{Z}$ é uma função geradora de números primos se $f(n) \in \mathbb{P}$, para algum $n \in \mathbb{Z}$.

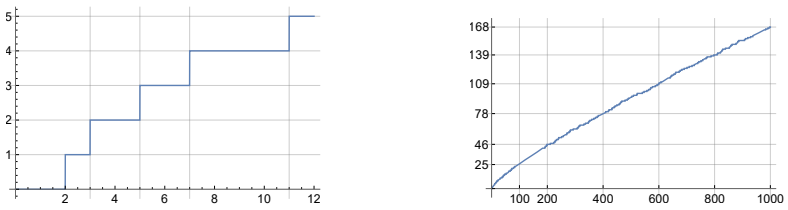
Definição 1.5 (Função Contagem de Primos)**.**

$$\pi(x) = \#\{p \in \mathbb{P} : p \leqslant x\}, \quad x > 0$$

onde, # denota o cardinal de um conjunto.

Na Figura 1 apresentamos o gráfico de $\pi(x)$.

FIGURA 1. Gráfico da função $\pi(x)$.



(a) $0 < x < 12$ (b) $0 < x < 1000$

Fonte: Os autores (2022).

Definição 1.6 (Quantidade de Números Primos de uma Função f até n)**.**

$$\pi_f(n) = \#\{p \in \mathbb{P} : f(k) = p,\ 0 < k \leqslant n\} \tag{3}$$

Em Teoria dos números, um dos conceitos essenciais para alavancar a matemática são os números primos. Eles são os responsáveis por muitos problemas em aberto, ou não resolvidos, nas diferentes áreas da matemática.

1.1. Breve História dos Números Primos.

◦ *Euclides de Alexandria (Alexandria, 330 - 277 a.C.).*

Teorema 1.2 (Teorema de Euclides)**.** *O conjunto dos números primos é infinito.*

Demonstração. Uma prova é encontrada nos Elementos (Livro IX, Proposição 20).

◦ *Eratóstenes de Cirene (Cirene, 276 - Alexandria, 194 a.C.).* Obteve os números primos até um valor de n usando tabelas que continham alguns números naturais. O processo consistia em eliminar múltiplos de números: primeiro do 2, depois do 3, e assim sucessivamente. Os primos eram aqueles números que restaram após a eliminação. Essa tabela ficou conhecida como "Crivo de Eratóstenes".

◦ *Pierre de Fermat (França, 1601 - 1665).*

Definição 1.7 (Primos de Fermat)**.**

$$F_n = 2^{2^n} + 1 \in \mathbb{P}, \quad n \in \mathbb{N}_0$$

Observação 1.

(a) Os quatro primeiros primos de Fermat são:

$$F_0 = 3, \quad F_1 = 5, \quad F_2 = 17, \quad F_3 = 257, \quad F_4 = 65.537$$

(b) Em 1732, foi provado que F_5 é composto:

$$F_5 = 2^{2^5} + 1 \ = \ 4.294.967.297 = 641 \times 6.700.417$$

◦ *Marin Mersenne (França, 1588 - 1648).*

Definição 1.8 (Primos de Mersenne)**.**

$$M_p = 2^p - 1 \in \mathbb{P}, \quad p \in \mathbb{P}$$

Observação 2.

(a) Os cinco primeiros primos de Mersenne são:

$$M_2 = 3, \quad M_3 = 7, \quad M_5 = 31, \quad M_7 = 127, \quad M_{13} = 8.191$$

(b) O maior primo de Mersenne até 03/03/2022 é: o 51º primo de Mersenne:

$$M_{82.589.933} = 2^{82.589.933} - 1$$

com 24.862.048 dígitos. (21 de dezembro de 2018, GIMPS).

(c) A Infinitude dos primos de Mersenne é um problema em aberto.

◦ *Leonhard Paul Euler (Suíça, 1707 - Russia, 1783).* Em 1732, provou que F_5 é um número composto, e em 1737, apresentou uma fórmula que relaciona os números primos com a série Harmônica generalizada.

Teorema 1.3 (Fórmula do Produto de Euler)**.**

$$\sum_{n=1}^{\infty} \frac{1}{n^x} \ = \ \prod_{p \in \mathbb{P}} \frac{1}{1 - p^{-x}}, \quad (x > 1)$$

◦ *Johann Peter Gustav Lejeune Dirichlet (Alemanha, 1805 - 1859).* O teorema de Dirichlet, sobre Progressão Aritmética (PA), conjeturado por Adrien-Marie Legendre (França, 1752 - 1833), foi provado por Dirichlet em 1837.

Teorema 1.4 (Teorema de Dirichlet)**.** *Sejam a, b inteiros positivos. Se a e b são coprimos, então existem infinitos números primos da forma $a + nb$, onde n é um inteiro positivo.*

O teorema de Dirichlet pode expressar-se, também, da seguinte forma:

(a) Sejam a, b inteiros positivos. Se a e b são coprimos, então existem infinitos números primos p tal que $p \equiv a \mod b$.

(b) Sejam a, b inteiros positivos. Se a e b são coprimos, então a PA: $a+nb$ possui infinitos números primos.

(c) Sejam a, b inteiros positivos. Se a e b são coprimos, então o conjunto dos números primos p, com $p \equiv a \mod b$, tem densidade de Dirichlet $1/\varphi(b)$.

No item (c), φ é a função totiente de Euler, e essa formulação é as vezes conhecida como: "O teorema da densidade de Dirichlet". A *densidade de Dirichlet ou analítica* de um conjunto S é definida pelo limite

$$\lim_{x\to 1^+}\left(\sum_{p\in S}(1/p^x)\,/\,\sum_{p\in\mathbb{P}}(1/p^x)\right)$$

e a *densidade natural* de um conjunto S é definida pelo limite

$$\lim_{x\to\infty}\frac{\#\{p\leqslant x:\ p\in\mathbb{P}\ \text{e}\ p\in S\}}{\#\{p\leqslant x:\ p\in\mathbb{P}\}}$$

O seguinte resultado relaciona estas duas densidades: "Se S tem densidade natural k, então a densidade analítica existe e é igual a k", (veja [15], pg. 76). Segundo [17], o teorema da densidade de Dirichlet, apresenta um padrão bastante intuitivo. Por exemplo, ao classificarmos os números primos de 1 até 100 de acordo com seu último dígito, não par maior que 2, teremos os seguintes grupos de primos:

1	11,21,31,41,61,71	5	5
2	2	7	7,17,37,47,67,97
3	3,13,23,43,53,73,83	9	19,29,59,79,89

Os grupos que terminam em $1, 3, 7$ e 9 têm, aproximadamente, a mesma quantidade de primos. Nesse sentido, podemos pensar que a quantidade de primos, em cada grupo, poderia ser infinita, se considerarmos uma quantidade infinita de números. Na verdade, esses grupos de primos são subconjuntos das classes residuais $[1], [3], [7]$ e $[9]$ módulo 10, respectivamente. Por outro lado, os números naturais menores e coprimos de 10 são: $1, 3, 7$ e 9, e a quantidade deles é determinada por $\varphi(10) = 4$. Agora, para $x = 100$, considere o seguinte quociente:

$$\frac{\#\{p\leqslant x:\ p\in\mathbb{P}\ \text{e}\ p\in[1]\}}{\#\{p\leqslant x:\ p\in\mathbb{P}\}} \quad = \quad \frac{6}{25} = 0,24 \approx \frac{1}{4} = \frac{1}{\varphi(10)} \tag{4}$$

Procedendo de forma análoga, o quociente $1/\varphi(10)$ se repete nos outros três grupos de primos. Se x tende-se ao infinito, (4) seria a densidade natural do conjunto $[1]$ com valor $1/\varphi(10)$. No limite, o denominador é infinito (Teorema 1.2). Assim, o numerador (a quantidade de primos do grupo que termina em 1) é infinito. Note que a densidade analítica de $[1]$ também é $1/\varphi(10)$. Portanto, esse resultado representa o teorema da densidade de Dirichlet, para o caso $b = 10$.

A prova analítica do teorema de Dirichlet, utilizada pelo mesmo, baseou-se na sua própria teoria de caráteres e $L-$funções, [15].

Ainda que, os números primos possuam uma definição bem simples, o grande problema é sua distribuição dentre os números naturais.

1.2. **Teorema do Número Primo (TNP).** Com a finalidade de encontrar resultados que determinem o comportamento dos números primos dentre o conjunto dos números naturais, o comportamento assintótico da quantidade de números primos despertou interesse na comunidade matemática. Uma formulação assintótica para a quantidade dos números primos é conhecida como Teorema do Número Primo (TNP). A seguir descreveremos, brevemente, alguns fatos históricos do TNP.

◦ *Carl Friedrich Gauss (Alemanha, 1777 - 1855).* Entre os 15 e 16 anos de idade, Gauss conjeturou a primeira versão do TNP.

Conjetura 1.1 (TNP, 1792/1793)**.**

$$\pi(x) \sim x/\ln x \tag{5}$$

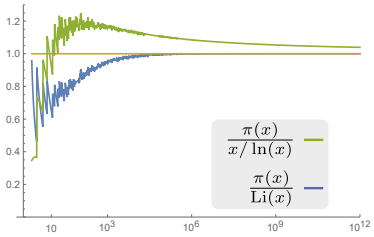
Em 1849, Gauss escreveu uma carta a seu amigo, o astrônomo Johann Encke, onde explica a segunda versão de seu TNP. Ele utiliza a função Integral Logarítmica, definida por $\mathrm{Li}(x) := \int_2^x \frac{dt}{\ln t} = \mathrm{li}(x) - \mathrm{li}(2)$, onde $\mathrm{li}(x) := \int_0^x \frac{dt}{\ln t}$.

Conjetura 1.2 (TNP, 1849)**.**

$$\pi(x) \sim \mathrm{Li}(x) \tag{6}$$

Na Figura 2 apresentamos o gráfico das funções: $\frac{\pi(x)}{x/\ln(x)}$ e $\frac{\pi(x)}{\mathrm{Li}(x)}$, $1 < x \leqslant 10^{12}$, com o domínio em escala logarítmica. Note que o limite, quando x tende ao infinito de ambas funções, é 1. Confirmando que, nesse intervalo, as conjeturas de Gauss são verdadeiras.

FIGURA 2. Confirmação computacional das conjeturas de Gauss.



Fonte: Os autores (2022).

◦ *Adrien-Marie Legendre (França, 1752 - 1833).*

Conjetura 1.3 (TNP, 1798)**.**

$$\pi(x) \sim x/(A\ln x - B) \tag{7}$$

Em 1808, Legendre, apresentou os valores $A = 1$ e $B = 1,08366$ para sua conjetura.

◦ *Pafnuti Tchebychev (Russia, 1821 - 1894).* Tchebychev propus $B = 1$ na conjetura de Legendre. Apresentou uma prova do seguinte teorema como uma tentativa de se aproximar à demonstração do TNP.

Teorema 1.5 (Tchebychev, 1852)**.**

$$0,92129 < \pi(x)/(x/\ln(x)) < 1,10555 \tag{8}$$

◦ *Jacques Hadamard (França, 1865 - 1963) e Charles-Jean de La Vallée Poussin (Bélgica, 1866 - 1962).* Eles, independentemente, demonstraram o TNP em 1896.

Teorema 1.6 (TNP, 1896)**.**

$$\lim_{x\to\infty} \pi(x)/(x/\ln(x)) \quad = \quad 1 \tag{9}$$

Outra forma de se obter relações assintóticas para a quantidade de números primos é através da Função Riemann de Contagem de Primos.

Definição 1.9 (Função Riemann de Contagem de Primos)**.** Seja μ a função de Möbius.

$$\mathrm{Ri}(x) \quad := \quad \sum_{n=1}^{\infty} \frac{\mu(n)}{n} \mathrm{li}(x^{1/n}) \tag{10}$$

Riemann, também, estudou a aproximação assintótica de $\pi(x)$. Conjeturou que

$$\pi(x) \quad \sim \quad \mathrm{Ri}(x) - \sum_{\rho}^{\infty} \mathrm{Ri}(x^{\rho}) \tag{11}$$

onde, ρ são os zeros não triviais da função Zeta de Riemann. Por outro lado, existe uma fórmula mais tratável, em questão de cálculos numéricos, para trabalhar no lugar da função de Riemann. Em 1893, Jørgen Gram, apresentou a série:

$$G(x) \quad := \quad 1 + \sum_{k=1}^{\infty} \frac{1}{kk!\zeta(k+1)} (\ln x)^k \quad \approx \quad \mathrm{Ri}(x) \tag{12}$$

Encontrar padrões, nos números primos, que descrevam seu comportamento tem sido tema de pesquisa e assunto de problemas em aberto.

1.3. **Problemas em Aberto em Teoria dos Números.** Descreveremos três problemas em aberto da Teoria dos Números: A Hipótese de Riemann, a Conjetura de Erdős sobre PA e a Conjetura de Goldbach.

1.3.1. *A Hipótese de Riemann (HR).* Foi proposta pelo matemático alemão Bernhard Riemann (1826 - 1866) em 1859. No 2º Congresso Internacional de Matemáticos (París, 1900) David Hilbert propôs 23 problemas de matemática. Sendo o 8º, "A hipótese de Riemann e a Conjetura de Goldbach". O Instituto Clay de Matemática fundado em Oxford (Reino Unido, 1998) promove o programa: "Problemas do Prémio Millenium", que no ano 2000 lançou os 7 problemas do milênio. Sendo o 4º, "A hipótese de Riemann". Quem resolva qualquer um desses problemas recebe o

prêmio de um milhão de dólares.

A convergência da série Harmônica generalizada, (1), é a base para a definição da função Zeta de Riemann.

Definição 1.10 (Função Zeta de Riemann)**.** A função $\zeta : \{s \in \mathbb{C} : \mathrm{Re}(s) > 1\} \longrightarrow \mathbb{C}$ é definida por

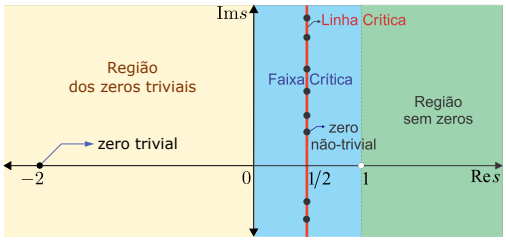
$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \tag{13}$$

Observação 3. A função Zeta, definida no semiplano $\mathrm{Re}(s) > 1$, estende-se por Continuação Analítica para todo o plano complexo e, cuja única singularidade é o polo simples $s = 1$. De aqui em diante, consideraremos a função Zeta de Riemann definida em todo o plano complexo, exceto em $s = 1$.

Conjetura 1.4 (Hipótese de Riemann, 1859)**.** *Todos os zeros não-triviais da função Zeta de Riemann estão na linha crítica* $\mathrm{Re}(s) = 1/2$.

Na Figura 3 apresentamos o domínio da função Zeta de Riemann.

FIGURA 3. Domínio da função Zeta de Riemann: $\mathbb{C} - \{1\}$.



Fonte: Os autores (2022).

◦ **Se a HR fosse verdadeira.** Existem muitos resultados supondo que a HR fosse verdadeira. Consideraremos duas aplicações:

(a) *Estimativas de erro.* No TNP, a melhor estimativa do erro foi sob a HR: Se a HR é verdadeira, então

$$\pi(x) - \mathrm{Li}(x) = O(\sqrt{x}\ln x) \tag{14}$$

(b) *Distância entre primos.* Em 1919, Harald Cramér, provou o seguinte resultado: Seja p_n o $n-$ésimo primo. Se a HR é verdadeira, então

$$p_{n+1} - p_n = O(\sqrt{p_n}\ln p_n)$$

◦ **A Hipótese de Riemann Generalizada (HRG).** A generalização da HR envolve as $L-$funções de Dirichlet $L(s,\chi)$, associados com o caráter de Dirichlet primitivo χ, módulo $q > 1$, [2].

Conjetura 1.5 (Hipótese de Riemann Generalizada)**.** *Todos os zeros não-triviais da função* $L(s,\chi)$ *estão na linha crítica* $\mathrm{Re}(s) = 1/2$.

Observação 4. Da mesma forma que se obtêm resultados assumindo que a HR é verdadeira, também ocorre com a HRG, obtendo-se resultados importantes na Teoria analítica dos números.

◦ **Verificação Computacional da HR.** Na Tabela 1, mostramos a quantidade de zeros da função Zeta de Riemann que se encontram na linha crítica.

Tabela 1. Os $n-$primeiros zeros não triviais da função Zeta de Riemann.

Ano	$n-$ primeiros zeros	Autor
1903	15	J. P. Gram
1956	25×10^3	D. H. Lehmer
2003	25×10^{10}	S. Wedeniwski
2004	10^{13}	Xavier Gourdon

Fonte: [5].

1.3.2. *A Conjetura de Erdős sobre PA.* Esta conjetura foi enunciada entre os anos 1940 e próximo aos anos 1955, [16]. O matemático húngaro Paul Erdős (1913 - 1996) escreveu em 1996 (publicado postumamente em 1997) que aumentou o prêmio de \$3.000 para \$5.000 pela prova ou refutação dessa conjetura.

Conjetura 1.6 (Erdős sobre PA)**.** *Seja* $A = \{a_1, a_2, \ldots, a_n, \ldots\} \subset \mathbb{Z}^+$, $a_i < a_{i+1}, \forall i$. *Se* $\sum_{n\in A} 1/a_n = \infty$, *então* A *contém PA arbitrariamente longas.*

A conjetura de Erdős ([6], [7] e [16]) implica dois grandes teoremas: de Szemerédi, [16], e de Green-Tao, [7].

Teorema 1.7 (Szemerédi, 1972)**.** *Qualquer subconjunto dos números naturais de densidade superior positiva contém PA arbitrariamente longas.*

Endre Szemerédi recebeu \$1.000, de Paul Erdős, pela prova desse teorema.

Teorema 1.8 (Green-Tao, 2008)**.** *O conjunto dos números primos contém infinitas PA de comprimento* k *para todo* k.

O teorema de Green-Tao pode, também, expressar-se do seguintes modo:

$$\forall k \in \mathbb{Z}^+,\ \exists\, a, b \in \mathbb{Z},\ b \neq 0;\quad \{a + nb\}_{n=0}^{k-1} \subset \mathbb{P}$$

Por exemplo, para $k = 3$: $\{3 + 2n\}_{n=0}^{2} = \{3, 5, 7\}$ é uma PA de 3 números primos. Por outro lado, para alguns valores de k podem existir várias PA de tamanho k (denotadas por $k-$PA). Na Tabela 2, apresentamos algumas das $k-$PA Minimal.

Tabela 2. Algumas $k-$PA Minimal (sendo minimizado o último termo).

k	Primos para $n = 0$ até $k-1$	k	Primos para $n = 0$ até $k-1$
3	$3 + 2n$	15	$115453391 + 4144140n$
5	$5 + 6n$	20	$214861583621 + 18846497670n$
10	$199 + 210n$	21	$5749146449311 + 26004868890n$

Fonte: [13].

Observação 5. Pelo TNP, temos que $\lim_{x\to\infty} \pi(x)/x = 0$, ou seja, os números primos tem densidade natural nula. Desse modo o teorema de Szemerédi não poderia implicar o teorema de Green-Tao.

1.3.3. *A Conjetura de Goldbach.* Em 1742, o matemático alemão Christian Goldbach (1690 - 1764) escreveu uma carta para Euler, onde explicava suas duas conjeturas: a "forte ou binária" e a "fraca ou ternária". A conjetura de Goldbach, a forte, é, sem dúvida, um dos problemas mais antigos que tem fascinado e inspirado a muitos matemáticos. No 2º Congresso Internacional de Matemáticos (Paris, 1990), Hilbert propôs 23 problemas, sendo a conjetura de Goldbach parte do 8º problema, [18].

Conjetura 1.7 (Forte ou binária, Goldbach, 1742)**.** *Todo número par maior que* 2 *pode ser expressado como a soma de dois números primos*

Exemplos desta conjetura: $4 = 2 + 2$, $6 = 3 + 3$, $8 = 5 + 3$ e $10 = 7 + 3 = 5 + 5$.

Observação 6.

(a) Olivier Ramaré (França, 1965) provou em 1995 que: "Todo número par $n > 2$ é a soma de no máximo seis primos".
(b) Tomás Oliveira e Silva (Portugal) verificou em 2012 que a conjetura é válida para: $2 < n \leqslant 4 \times 10^{18}$.
(c) O matemático australo-americano Terence Tao (Austrália, 1975) provou em 2012 que: "Todo número ímpar $n > 1$ é a soma de no máximo cinco primos".

Maiores informações da conjetura forte pode ser encontrada em [8].

Conjetura 1.8 (Fraca ou ternária, 1742)**.** *Todo número ímpar maior que* 5 *pode ser expressado como a soma de três números primos*

Exemplos: $7 = 3 + 2 + 2$, $9 = 3 + 3 + 3 = 5 + 2 + 2$ e $11 = 7 + 2 + 2 = 5 + 3 + 3$.

Observação 7.

(a) Em 2013, o matemático Harald Andrés Helfgott (Peru, 1977) submeteu o artigo: "The ternary Goldbach conjecture is true". Até o momento é o único artigo que prova que a conjetura fraca de Goldbach é verdadeira, [9].
(b) Em 2013, H. Helfgott e D. Platt verificaram que a conjetura é válida para: $5 < n \leqslant 8.875 \times 10^{30}$.
(c) Algumas Teorias abordadas para esta conjetura são: Método do círculo (Series de Fourier, Transformada de Fourier, Transformada de Laplace).

Sabe-se até o presente momento que não existe nenhuma fórmula que descreva todos os números primos, para todo n natural. O Teorema de Wilson (atribuído a Jhon Wilson em 1770, onde, ainda, era uma conjetura) nos fornece as condições necessária e suficiente para que um número seja primo. A condição suficiente: "Seja $n > 1$ um inteiro. Se $(n-1)! \equiv -1 \mod n$, então n é primo" nos propicia um Teste de Primalidade, mas infelizmente não é eficiente a medida que n cresce. Porém, indo em uma direção menos abrangente é possível encontrar certos "*Padrões especiais de distribuição dos números primos*". É importante para a teoria matemática saber se esses padrões especiais descrevem, ou não, infinitos números primos, e qual é sua densidade. Encontramos, pelo menos, duas formas de padrões especiais: (a) *Forma Analítica*: consiste em encontrar diretamente funções geradoras de números primos (pode ser explicitamente ou através de problemas de existência). (b) *Forma Geométrica*: consiste em usar arranjos dos números naturais sobre geometrias planas,

com a finalidade de identificar padrões especiais de distribuição dos números primos; identificado o padrão, procede-se à obtenção da função geradora de números primos.

1.4. **Padrões Especiais de Números Primos.**

1.4.1. *Com relação à forma analítica.* Citaremos alguns exemplos:

∘ FUNÇÃO EXPONENCIAL. Os primos de Mersenne $M_p = 2^p - 1$, onde p é primo, representam uma função exponencial, na variável p, geradora de número primos.

∘ FUNÇÃO POLINOMIAL DE 1º GRAU. A PA $a + nb$, onde $\text{mdc}(a, b) = 1$, é uma função polinomial, na variável n, geradora de números primos. Devido ao Teorema de Dirichlet, essa função possui infinitos números primos.

∘ FUNÇÃO POLINOMIAL DE 2º GRAU. O polinômio de Euler $n^2 - n + 41$, onde $n = 1, 2, 3, \ldots$ é uma função geradora de números primos, [11].

Observação 8.

(a) Polinômios não geram, totalmente, números primos.

Teorema 1.9 (Goldbach, 1752)**.** *Seja $f(x)$ um polinômio de $n-$ésimo grau com coeficientes inteiros. Se $f(k) \in \mathbb{P}$, $\forall\, k \in \mathbb{Z}_0^+$, então $f(x)$ é constante.*

Prova. Para $k = 0$, é evidente. Seja $k_0 > 0$ e $f(k_0) = p,\ p \in \mathbb{P}$.

$$\begin{aligned} a_0 &\equiv a_0 \mod p \\ a_1 k_0 &\equiv a_1(k_0 + tp) \mod p \\ &\vdots \\ a_n k_0^n &\equiv a_n(k_0 + tp)^n \mod p \\ \hline f(k_0) &\equiv f(k_0 + tp) \mod p \end{aligned} \tag{15}$$

De (15), temos que $p \mid f(k_0 + tp) - f(k_0), \forall t \in \mathbb{Z}$. Como $p \mid f(k_0)$, então $p \mid f(k_0 + tp)$. Logo, $f(k_0 + tp) = 1$ ou $f(k_0 + tp) = p$. Assim, $\forall t \in \mathbb{Z}$, $f(k_0 + tp) = p$, pois $f(k_0 + tp)$ é primo por hipótese. Agora, definamos a função: $g(k) = f(k) - p, \forall k \in \mathbb{Z}^+$. Note que, $\forall t \in \mathbb{Z}$, $k_0 + tp$ são zeros da função $g(k)$. Logo, $g(k) = 0, \forall k \in \mathbb{Z}^+$. Portanto, $f(k) = p, \forall k \in \mathbb{Z}^+$.

(b) Infinitude dos Números Primos.

∘ *Função polinomial de 2º grau.* A infinitude dos números primos gerados por polinômios de 2º grau é um problema em aberto.

Conjetura 1.9 (Conjetura F de Hardy-Littlewood, 1922)**.** *Sejam $a, b, c \in \mathbb{Z}, a > 0$, $mdc(a, b, c) = 1$, $\Delta = b^2 - 4ac$ não é um quadrado e, $a + b$ e c não são ambos pares. Então, existem infinitos números primos da forma $f(n) = an^2 + bn + c$. Além disso,*

$$\pi_f(n) \sim C\varepsilon\sqrt{n}/(\sqrt{a}\ln n) \tag{16}$$

onde,

$$C = \prod_{\substack{p>2 \\ p|a \wedge p|b}} \left(\frac{p}{p-1}\right) \prod_{\substack{p>2 \\ p\nmid a}} \left(1 - \frac{1}{p-1}\left(\frac{\Delta}{p}\right)\right) \tag{17}$$

$\left(\frac{\Delta}{p}\right)$: *denota o símbolo de Legrendre*

$$\varepsilon = \begin{cases} 1/2 & se \quad 2 \nmid a+b \\ 1 & de\ outro\ modo \end{cases}$$

∘ *Função polinomial de $n-$ésimo grau.* Não existe uma prova sobre a infinitude de números primos gerados por polinômios de $n-$ésimo grau de uma variável. Veja a Conjetura de Buniakovsky, ([12], p. 145).

Conjetura 1.10 (Conjetura de Buniakovsky, 1857)**.** *Se $f(x)$ é um polinômio de $n-$ésimo grau com coeficientes inteiros, irreduzível e $N = mdc(f(m) : m \in \mathbb{N})$, então existem infinitos $m \in \mathbb{N}$ tal que $\frac{1}{N}|f(m)|$ é primo.*

1.4.2. *Com relação à forma geométrica.* Consideraremos três exemplos:

∘ A ESPIRAL DE ULAM. Em 1963, Stanislaw Ulam elaborou um arranjo dos números naturais, sobre um quadrado, em uma ordem que segue o formato de uma espiral "quadrada", conhecida como a Espiral de Ulam. Ulam observou, nesse arranjo, que existem certos padrões de números primos determinados por polinômios quadráticos da forma: $4n^2 + bn + c$, [1].

∘ ORLOWSKI E CHMIELEWSKY, [14] usaram processamento de imagens na espiral de Ulam para detectar linhas com muitos números primos. Eles registraram que o polinômio $4n^2 - 1260n + 98827$ gera 613 primos, para $n = 0, 1, \ldots, 1000$, resultando ser mais denso que o polinômio de Euler.

∘ JARA-VERA E SÁNCHEZ-ÁVILA, [10] estudaram padrões da adjacência de números primos em linhas da espiral de Ulam. Eles mostraram que as adjacências diminuem rapidamente, mas não desaparecem, conforme a espiral cresce.

O presente trabalho utiliza a forma geométrica para encontrar padrões especiais de distribuição de números primos. Criamos um novo arranjo dos números naturais, sobre um quadrado, o qual chamamos de $n-$quadrado Zeta. Da infinidade de "caminhos" que este quadrado pode fornecer-nos, temos considerado apenas dois. Eles são modelados por funções geradoras de números primos que são polinômios de 2° grau da forma: $k^2 + ak + b$, sendo a ímpar e b um quadrado perfeito.

2. O $n-$QUADRADO

Nessa Seção, definiremos um $n-$quadrado e os conceitos que se relacionam com a geometria desse quadrado. Assim, como também, seu arranjo de formação.

Definição 2.1 (**Casa de um Quadrado**)**.** Seja Q um quadrado. Uma casa de Q é um quadrado contido em Q, cujo valor é um número inteiro positivo.

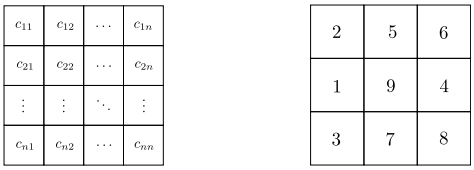
Definição 2.2 (**O $n-$quadrado**)**.** Dado $n \in \mathbb{Z}^+$. Um $n-$quadrado é um quadrado formado por n^2 casas que se interceptam em um lado. Cada casa c_{ij}, $i, j = 1, 2, \ldots, n$, segue uma ordem, como descrita na Figura 4(a), e seu valor não se repete.

Notação 1. O $n-$quadrado é denotado por

$$\{c_{11}, c_{12}, \ldots, c_{1n};\ c_{21}, c_{22}, \ldots, c_{2n};\ \ldots\ ; c_{n1}, c_{n2}, \ldots, c_{nn}\}$$

Exemplo 1. Na Figura 4(b) temos o $3-$quadrado definido por $\{2, 5, 6; 1, 9, 4; 3, 7, 8\}$.

FIGURA 4. (a) O $n-$quadrado. (b) Um 3$-$quadrado.



(a) (b)

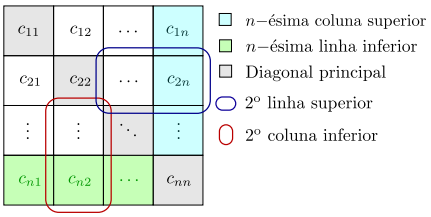
Fonte: Os autores (2022).

2.1. **Geometria de um $n-$quadrado.** Alguns elementos geométricos são:

$i-$ésima linha: $(c_{i1}, c_{i2}, \ldots, c_{in})$
Diagonal principal: $(c_{11}, c_{22}, \ldots, c_{nn})$
$i-$ésima linha superior: $(c_{i(i+1)}, c_{i(i+2)}, \ldots, c_{in})$
$j-$ésima coluna inferior: $(c_{(j+1)j}, c_{(j+2)j}, \ldots, c_{nj})$
$j-$ésima coluna: $(c_{1j}, c_{2j}, \ldots, c_{nj})$
$i-$ésima linha inferior: $(c_{i1}, c_{i2}, \ldots, c_{i(i-1)})$
$j-$ésima coluna superior: $(c_{1j}, c_{2j}, \ldots, c_{(j-1)j})$

Na Figura 5 esboçamos no $n-$quadrado alguns desses elementos.

FIGURA 5. Geometria de um $n-$quadrado.



Fonte: Os autores (2022).

2.2. **Arranjo de Formação.**

Definição 2.3 (**Arranjo de Formação**)**.** Chama-se um arranjo de formação de um $n-$quadrado a qualquer preenchimento não-aleatório dos valores de suas casas.

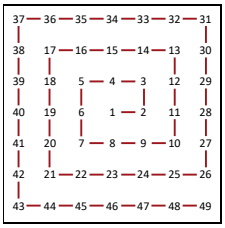
Observação 9. O arranjo de formação de um $n-$quadrado pode vir acompanhado de linhas cuja única finalidade é expressar de forma clara sua formação.

2.2.1. *A Espiral de Ulam e o $n-$quadrado Zeta.* A espiral de Ulam e o $n-$qZeta são dois arranjos diferentes dos números naturais que servem para identificar padrões de distribuição de números primos sobre $n-$quadrados.

Definição 2.4 (**A Espiral de Ulam**)**.** A espiral de Ulam é um arranjo dos números naturais sobre um $n-$quadrado, cujo arranjo de formação consiste em distribuir os números naturais a partir do 1, sendo ele colocado no centro do quadrado e os outros números seguem segundo a forma de uma "espiral quadrada".

Exemplo 2. A Figura 6 apresenta a Espiral de Ulam em um 7$-$quadrado com arranjo de formação através de linhas.

FIGURA 6. A espiral de Ulam, em um 7−quadrado.



Fonte: Os autores (2022).

Definição 2.5 (O n−quadrado Zeta). O n−quadrado Zeta, denotado por n−qZeta, é um n−quadrado, cujo arranjo de formação segue os 3 passos seguintes:

(1º) Casas da diagonal principal. São preenchidas, nessa ordem, pelos quadrados perfeitos: $1, 4, 9, \ldots, n^2$.

(2º) Casas da j−ésima coluna superior. Para cada j, $2 \leqslant j \leqslant n$, temos a sequência:

$$\Big((j-1)^2+1, \quad (j-1)^2+2, \quad (j-1)^2+3, \quad \ldots \quad , \quad j(j-1)\Big) \tag{18}$$

que possui $(j-1)$ casas. As casas com números primos são pintadas de cor verde.

(3º) Casas da i−ésima linha inferior. Para cada i, $2 \leqslant i \leqslant n$, temos a sequência:

$$\Big(i(i-1)+1, \quad i(i-1)+2, \quad i(i-1)+3, \quad \ldots \quad , \quad (i-1)(i+1)\Big) \tag{19}$$

que possui $(i-1)$ casas. As casas com números primos são pintadas de cor amarela.

Exemplo 3. Na Figura 7(a) apresentamos o 7−quadrado Zeta que mostra seu arranjo de formação através de linhas (note que os quatro primeiros números formam a letra Z, que no alfabeto grego seria o Zeta, por isso o nome Zeta no arranjo). Já, a Figura 7(b) representa o 7−quadrado Zeta obtido pela Definição 2.5, que a seguir explicamos:

(i) *Diagonal principal.* A diagonal principal $(c_{11}, c_{22}, \ldots, c_{77})$ é preenchida pela sequência $(1, 4, 9, 16, 25, 36, 49)$.

(ii) *Colunas superiores e linhas inferiores.* Na Tabela 3 observamos que há uma ordem de preenchimento, conforme indicamos nas colunas 1 e 4. Nas colunas 2 e 5, indicamos a j−ésima coluna superior e a i−ésima linha inferior do 7−qZeta, respectivamente. Já, nas colunas 3 e 6 expressamos as sequências que resultam ao utilizar-se as fórmulas (18) e (19), respectivamente.

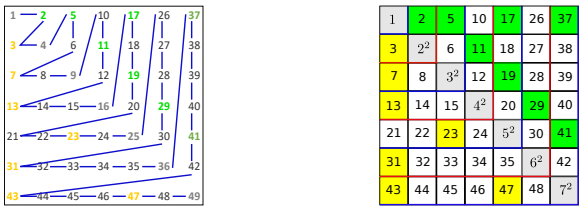
Exemplo 4. Na Figura 8 apresentamos quatro n−quadrados Zeta: (a) O 10-qZeta. (b) O 20-qZeta. (c) O 50-qZeta. (d) O 100-qZeta, e na Figura 15 (em ANEXO): (e) O 200-qZeta. (f) O 300-qZeta. Para a elaboração de um n−qZeta foi escrito um código em Linguagem Visual Basic do Excel.

TABELA 3. Preenchimento das casas do $7-$quadrado Zeta.

Ord.	Coluna superior	Sequências	Ord.	Linha inferior	Sequências
1°	2	(2)	2°	2	(3)
3°	3	(5, 6)	4°	3	(7, 8)
5°	4	(10, 11, 12)	6°	4	(13, 14, 15)
7°	5	(17, 18, 19, 20)	8°	5	(21, 22, 23, 24)
9°	6	(26, 27, 28, 29, 30)	10°	6	(31, 32, 33, 34, 35)
11°	7	(37, 38, 39, 40, 41, 42)	12°	7	(43, 44, 45, 46, 47, 48)

Fonte: Os autores (2022).

FIGURA 7. (a) O $7-$qZeta mostrando o arranjo de formação através de linhas. (b) O $7-$qZeta, obtido pela Definição 2.5.



(a) (b)

Fonte: Os autores (2022).

3. CAMINHOS EM UM $n-$QUADRADO

Nessa Seção, definiremos os conceitos relacionados com caminhos no $n-$quadrado.

Definição 3.1 (Sequência de casas)**.** Uma sequência de casas no $n-$quadrado, denotado por (c_{ij}), $i, j = 1, 2, 3, \ldots, n$, é qualquer agrupamento "contínuo" de casas.

Definição 3.2 (Caminho)**.** Um caminho no $n-$quadrado, denotado por $\mathbf{C}(n) = (c_{ij})$, $i, j = 1, 2, 3, \ldots, n$, é qualquer sequência de casas com pelo menos um primo.

Definição 3.3 (Caminho Retilíneo)**.** Um caminho retilíneo é um caminho no $n-$quadrado que, geometricamente, descreve uma linha reta.

Exemplo 5.

Considere o $3-$quadrado $\{1, 4, 5; 7, 2, 6; 8, 9, 3\}$ (veja Figura 9):

(a) O caminho $\mathbf{C}_1(3) = (c_{22}, c_{23}, c_{32}, c_{21}, c_{11}) = (2, 6, 9, 7, 1)$.
(b) O caminho retilíneo $\mathbf{C}_2(3) = (c_{21}, c_{32}) = (7, 9)$.

Considere o $7-$qZeta (veja Figura 9):

(c) $\mathbf{C}_1(7) = (8, 15, 23, 34, 45)$. $\mathbf{C}_2(7) = (11, 17)$.
$\mathbf{C}_3(7) = (5, 11, 19, 29, 41)$. $\mathbf{C}_4(7) = (3, 7, 13, 21, 31, 43)$.
O único caminho que não é retilíneo é o $\mathbf{C}_1(7)$.

Considere o $41-$qZeta (veja Figura 9):

(d) Temos o seguinte caminho retilíneo:

$$\mathbf{C}(41) = (c_{20\,22}, c_{19\,23}, c_{18\,24}, \ldots, c_{3\,39}, c_{2\,40}, c_{1\,41}) = (461, 503, 547, \ldots, 1447, 1523, 1601)$$

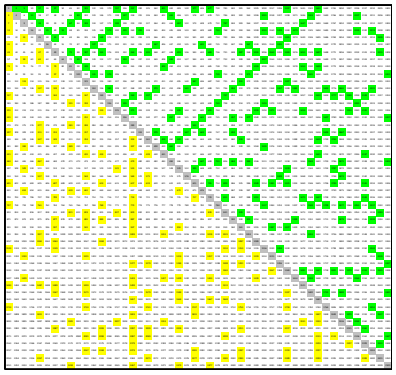
FIGURA 8. O $n-$quadrado Zeta.

1	2	5	10	17	26	37	50	65	82
3	4	6	11	18	27	38	51	66	83
7	8	9	12	19	28	39	52	67	84
13	14	15	16	20	29	40	53	68	85
21	22	23	24	25	30	41	54	69	86
31	32	33	34	35	36	42	55	70	87
43	44	45	46	47	48	49	56	71	88
57	58	59	60	61	62	63	64	72	89
73	74	75	76	77	78	79	80	81	90
91	92	93	94	95	96	97	98	99	100

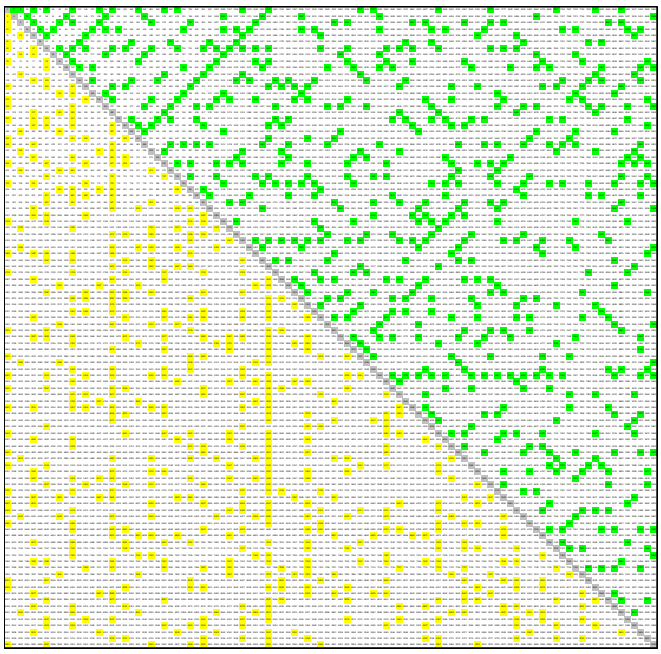
(a) O 10−quadrado Zeta.

1	2	5	10	17	26	37	50	65	82	101	122	145	170	197	226	257	290	325	362
3	4	6	11	18	27	38	51	66	83	102	123	146	171	198	227	258	291	326	363
7	8	9	12	19	28	39	52	67	84	103	124	147	172	199	228	259	292	327	364
13	14	15	16	20	29	40	53	68	85	104	125	148	173	200	229	260	293	328	365
21	22	23	24	25	30	41	54	69	86	105	126	149	174	201	230	261	294	329	366
31	32	33	34	35	36	42	55	70	87	106	127	150	175	202	231	262	295	330	367
43	44	45	46	47	48	49	56	71	88	107	128	151	176	203	232	263	296	331	368
57	58	59	60	61	62	63	64	72	89	108	129	152	177	204	233	264	297	332	369
73	74	75	76	77	78	79	80	81	90	109	130	153	178	205	234	265	298	333	370
91	92	93	94	95	96	97	98	99	100	110	131	154	179	206	235	266	299	334	371
111	112	113	114	115	116	117	118	119	120	121	132	155	180	207	236	267	300	335	372
133	134	135	136	137	138	139	140	141	142	143	144	156	181	208	237	268	301	336	373
157	158	159	160	161	162	163	164	165	166	167	168	169	182	209	238	269	302	337	374
183	184	185	186	187	188	189	190	191	192	193	194	195	196	210	239	270	303	338	375
211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	240	271	304	339	376
241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	272	305	340	377
273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	306	341	378
307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	342	379
343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	380
381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400

(b) O 20−quadrado Zeta.



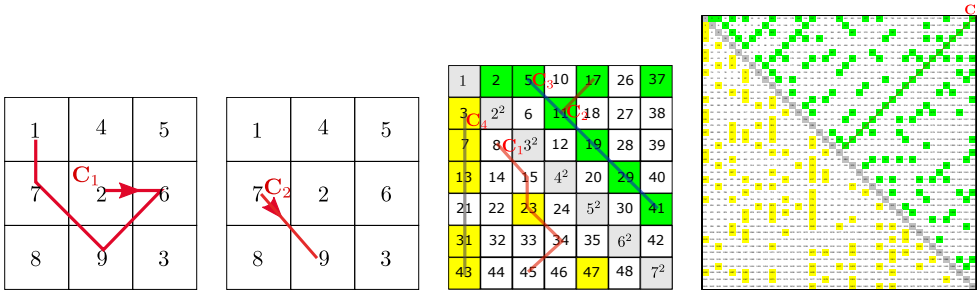
(c) O 50−quadrado Zeta.



(d) O 100−quadrado Zeta.

Fonte: Os autores (2021).

FIGURA 9. Caminhos e caminhos retilíneos.



(a) 3−quadrado. (b) 3−quadrado. (c) 7−qZeta. (d) 41−qZeta.

Fonte: Os autores (2022).

Definição 3.4 (Comprimento de um Caminho)**.** No $n-$quadrado, o comprimento de um caminho $\mathbf{C}(n)$, denotado por $|\mathbf{C}(n)|$, é o número total de casas que ele possui.

Definição 3.5 (Densidade de um Caminho)**.** No $n-$quadrado, a densidade de um caminho $\mathbf{C}(n)$, denotado por $\rho(n)$, define-se por

$$\rho(n) \quad = \quad |p| \,/\, |\mathbf{C}(n)| \tag{20}$$

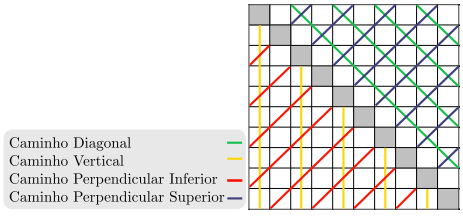
onde, $|p|$ denota a quantidade de números primos que contem o caminho $\mathbf{C}(n)$.

Exemplo 6. Considere alguns caminhos do Exemplo 5. Então,

(a) $|\mathbf{C}_1(3)| = 5$, $|p| = 2$ e $\rho(3) = \frac{2}{5} = 0,4$. Logo, $\mathbf{C}_1(3)$ é um caminho 40%.
(b) $|\mathbf{C}(41)| = 20$, $|p| = 20$ e $\rho(41) = 1$. Assim, $\mathbf{C}(41)$ é um caminho 100%.

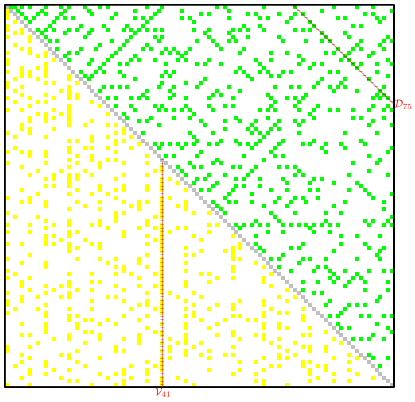
3.1. **Tipos de Caminhos Retilíneos.** Existem vários tipos de caminhos retilíneos em um $n-$quadrado. Escolheremos quatro caminhos retilíneos que partem da diagonal principal: Diagonal, vertical, perpendicular inferior e perpendicular superior. É importante destacar que a primeira casa desses quatro caminhos retilíneos podem ou não iniciar-se na casa da diagonal principal. Veja na Figura 10 cada um deles.

FIGURA 10. Quatro tipos de caminhos retilíneos em um $n-$quadrado.



Fonte: Os autores (2022).

FIGURA 11. Os caminhos $\mathcal{D}_{75}(100)$ e $\mathcal{V}_{41}(100)$, no $100-$qZeta.



Fonte: Os autores (2022).

4. Caminhos Diagonais e Verticais no $n-$qZeta

Nessa Seção, definiremos, no $n-$qZeta, os caminhos retilíneos: diagonal e vertical que foram escolhidos dentre uma infinidade de caminhos retilíneos.

Definição 4.1 (Caminho Diagonal)**.** Seja o $n-$qZeta. Um caminho diagonal, denotado por $\mathcal{D}_j(n), 3 \leqslant j \leqslant r$, é definido por

$$\mathcal{D}_j(n) = (c_{1\,j}, c_{2\,(j+1)}, \ldots, c_{(n-j+1)\,n}) \tag{21}$$

onde j é ímpar, e $r = n$ se n é ímpar e $r = n-1$ se n é par.
O subíndice j indica que o caminho diagonal se inicia na coluna j.

Definição 4.2 (Comprimento de um Caminho Diagonal)**.**

$$|\mathcal{D}_j(n)| = n - (j-1)$$

Observação 10. Os caminhos diagonais são "paralelos" à diagonal principal, e possuem um comprimento "infinito", pois crescem a medida que n aumenta.

Exemplo 7. No $100-$qZeta (veja Figura 11) temos o caminho diagonal $61,54\%$:

$$\mathcal{D}_{75}(100) = (c_{1\,75}, c_{2\,76}, \ldots, c_{25\,99}, c_{26\,100}) = (5.477, 5.627, \ldots, 9.629, 9.827)$$

Definição 4.3 (Caminho Vertical)**.** Seja o $n-$qZeta. Um caminho vertical, denotado por $\mathcal{V}_j(n), 1 \leqslant j \leqslant r$, é definido por

$$\mathcal{V}_j(n) = (c_{(j+1)\,j}, c_{(j+2)\,j}, \ldots, c_{n\,j}) \tag{22}$$

onde j é ímpar, e $r = n-2$ se n é ímpar e $r = n-1$ se n é par.
O subíndice j indica que o caminho se encontra na coluna j.

Definição 4.4 (Comprimento de um Caminho Vertical)**.**

$$|\mathcal{V}_j(n)| = n - j$$

Observação 11. Os caminhos verticais estão formados por casas da coluna inferior, e possuem um comprimento "infinito", pois crescem conforme n aumenta.

Exemplo 8 (Caminho vertical $\mathcal{V}_{41}(100)$)**.** No 100−qZeta (veja a Figura 11),

$$\begin{aligned} \mathcal{V}_{41}(100) &= (c_{42\,41}, c_{43\,41}, c_{44\,41}, \ldots, c_{98\,41}, c_{99\,41}, c_{100\,41}) \\ &= (1.763, 1.847, 1.933, \ldots, 9.547, 9.743, 9.941) \end{aligned} \tag{23}$$

é um caminho vertical $77,97\%$.

5. Formulação Matemática de Caminhos Retilíneos no $n-$qZeta

Nessa Seção, vamos determinar uma formulação matemática para os caminhos retilíneos: diagonal e vertical, que se encontram no $n-$qZeta. Veremos que dado um caminho retilíneo existe um polinômio quadrático associado a ele.

Proposição 5.1 (Polinômio associado a um Caminho Diagonal)**.** *Seja o $n-$qZeta. O polinômio que descreve os valores das casas do caminho diagonal $\mathcal{D}_j(n)$, $3 \leqslant j \leqslant r$, e comprimento s é dado por*

$$d_j(k) = k^2 + (2j-3)k + (j-2)^2, \quad \forall k = 1, 2, 3, \ldots, s \tag{24}$$

Prova. De (21), temos a seguinte sequência:

$$\begin{aligned} a_1 &= c_{1\,j} = j^2 - 2j + 2, & a_2 &= c_{2\,(j+1)} = a_1 + 2j, \\ a_3 &= c_{3\,(j+2)} = a_2 + 2j + 1 \cdot 2, & a_4 &= c_{4\,(j+3)} = a_3 + 2j + 2 \cdot 2, \quad \ldots \\ a_k &= c_{k\,(j+k-1)} = a_{k-1} + 2j + (k-2) \cdot 2 \end{aligned}$$

Observamos que a sequência é, na verdade, uma recorrência linear não-homogênea:

$$a_1 = j^2 - 2j + 2 \quad \text{e} \quad a_k = a_{k-1} + 2k + 2(j-2),\ k \geqslant 2 \tag{25}$$

A solução de (25) é dada por

$$a_k = a_1 + \sum_{l=2}^{k} \Big(2l + 2(j-2)\Big) = k^2 + (2j-3)k + (j-2)^2, \quad \forall k \geqslant 1$$

Essa recorrência define o seguinte polinômio:

$$d_j(k) = k^2 + (2j-3)k + (j-2)^2, \quad 1 \leqslant k \leqslant s$$

Exemplo 9. O polinômio associado a $\mathcal{D}_{75}(100)$, comprimento $26 = 100 - 75 + 1$, é

$$d_{75}(k) = k^2 + 147k + 5329,\ \forall k = 1, 2, \ldots, 26$$

Proposição 5.2 (Polinômio associado a um Caminho Vertical)**.** *Seja o $n-$qZeta. O polinômio que descreve os valores das casas do caminho vertical $\mathcal{V}_j(n)$, $1 \leqslant j \leqslant r$, e comprimento s é dado por*

$$v_j(k) = k^2 + (2j-1)k + j^2, \quad \forall k = 1, 2, 3, \ldots, s \tag{26}$$

Prova. A prova é análoga à Proposição 5.1, considerando a seguinte recorrência:

$$a_1 = j(j+2) \quad \text{e} \quad a_k = a_{k-1} + 2k + 2(j-1),\ k \geqslant 2 \tag{27}$$

Definição 5.1 (O Polinômio v_{41})**.** Seja o $n-$qZeta, $n \geq 42$. O polinômio associado ao caminho vertical $\mathcal{V}_{41}(n)$ é chamado Polinômio v_{41}, e é dado por

$$v_{41}(k) = k^2 + 81k + 1681,\ \forall k = 1, 2, \ldots, (n-41) \tag{28}$$

Exemplo 10. O polinômio, associado a $\mathcal{V}_{41}(100)$, comprimento $59 = 100 - 41$, é

$$v_{41}(k) = k^2 + 81k + 1681,\ \forall k = 1, 2, \ldots, 59 \tag{29}$$

6. Função Geradora de Números Primos: O Polinômio v_{41}

Nessa Seção, estudaremos o polinômio v_{41}, que é um polinômio de 2º grau, gerador de números primos.

No $n-$ qZeta, os polinômios quadráticos (24) e (26) associados aos caminhos diagonais e verticais, respectivamente, são funções geradoras de números primos. Outras funções geradoras de números primos foram estudadas na Subseção 1.4.1. Também, podemos considerar [19] para o caso de funções polinomiais. No entanto, o mais admirável é o Polinômio de Euler, o qual é definido por:

$$p_E(k) = k^2 - k + 41, \quad k = 1, 2, 3, \ldots \tag{30}$$

Este polinômio gera quarenta números primos para todo $k = 1, 2, 3, \ldots, 40$, eles são:

$$\{41, 43, 47, 53, 61, \ldots, 1.447, 1.523, 1.601\}$$

A sequência não possui números primos consecutivos, pois o 59 não está presente. Mas, ele continua fornecendo mais primos para outros valores de k. Por outro lado, o Polinômio v_{41}, definido em (28), é, também, uma função geradora de números primos. Infelizmente ele está limitado a $(n-41)$ valores para k, por causa do caminho vertical $\mathcal{V}_{41}(n)$ possui $(n-41)$ casas no $n-$qZeta. Devido a essa restrição consideramos necessário extrair o caminho vertical $\mathcal{V}_{41}(n)$ do $n-$qZeta para que o polinômio v_{41} possa fornecer-nos mais números primos para outros valores de k.

6.1. **O Caminho Vertical $\mathcal{V}_{41}(n)$ fora do $n-$qZeta.** Para cada $n \geqslant 42$, existe um único caminho vertical $\mathcal{V}_{41}(n)$ de comprimento $(n-41)$, no $n-$qZeta. Se pensamos, por um momento, em um $\infty-$qZeta teremos um único caminho vertical $\mathcal{V}_{41}$ independente de n e de comprimento "infinito". Assim, o caminho vertical $\mathcal{V}_{41}$ seria:

$$\mathcal{V}_{41} = (c_{42}, c_{43}, c_{44}, \ldots) = (1.763, 1.847, 1.933, \ldots) \tag{31}$$

Sendo assim, o polinômio v_{41}, associado ao caminho vertical $\mathcal{V}_{41}$, é dado por

$$v_{41}(k) = k^2 + 81k + 1681, \ \forall k = 1, 2, 3, \ldots \tag{32}$$

Este polinômio de 2º grau é uma função geradora de números primos.

6.2. **O Polinômio de Euler como Translação do Polinômio v_{41}.** Por inspeção,

$$v_{41}(1) = p_E(42) = 1763, \ v_{41}(2) = p_E(43) = 1847, \quad \ldots \quad , v_{41}(k) = p_E(k+41), \ k \geqslant 1$$

Assim, o polinômio v_{41} resulta ser uma translação de 41 unidades, para trás, do polinômio de Euler. Ou, ainda,

$$p_E(k) = v_{41}(k-41) \tag{33}$$

Nesse caso o polinômio v_{41} pode, também, ser avaliado nos inteiros não positivos.

7. O Perfil do Polinômio de Euler no $n-$quadrado

Nessa Seção, estudaremos o perfil do polinômio de Euler nos $n-$quadrados: Espiral de Ulam e o $n-$qZeta.

Definição 7.1 (Perfil de um Polinômio)**.** O perfil de um polinômio $p(x)$ em um $n-$quadrado está formado por sequências de casas, cujos valores das casas coincidem com os valores do polinômio $p(k)$ para todo $k = 1, 2, 3, \ldots$

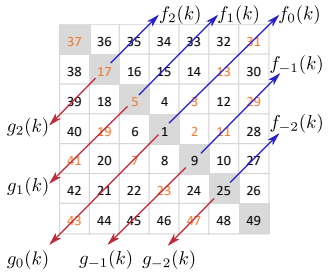
Exemplo 11. Seja o polinômio $p(x) = x^2 + x + 1$. Na Figura 12 mostramos o perfil do polinômio $p(x)$ no $5-$quadrado $\{2, 3, 5, \ldots\}$. Esse perfil está formado pelo caminho $\mathbf{C}(5) = (3, 7, 13)$ e pela sequência de casa única: (21).

FIGURA 12. Perfil de $p(x) = x^2 + x + 1$, no $5-$quadrado.

2	3	5	6	1
7	8	10	4	11
12	13	9	14	15
17	16	18	19	20
25	21	22	23	24

Fonte: Os autores (2022).

FIGURA 13. Funções Diagonais na Espiral de Ulam, no $7-$quadrado.



Fonte: Os autores (2022).

7.0.1. *Perfil do Polinômio de Euler.* Estudaremos o perfil do polinômio de Euler $p_E(x) = x^2 - x + 41$, tanto na espiral de Ulam, quanto no $n-$quadrado Zeta. Estes perfis foram obtidos com um código escrito em Linguagem Visual Basic do Excel.

Na Espiral de Ulam. Na Figura 14(a) mostramos o perfil do polinômio de Euler na Espiral de Ulam (com o número 1 no centro) em um 99-quadrado. A 1ª casa está pintada de cor azul escuro. É importante salientar que só aparecem impressos os números primos que compõem o perfil. Este perfil apresenta, geometricamente, as seguintes características: (a) As 12 primeiras casas são primos e não seguem nenhum padrão específico. (b) A partir da 13ª casa, encontramos 6 caminhos retilíneos:

$S_1 = (197, 251, 313, 383)$ $S_2 = (223, 281, 347, 421)$
$S_3 = (461, 547, 641, \ldots, 1523)$ $S_4 = (503, 593, 691, \ldots, 1447)$
$S_5 = (1601, 1763, 1933, \ldots, 9547, \ldots)$ $S_6 = (1681, 1847, 2021, \ldots, 9743, \ldots)$

onde, $S_1 \parallel S_2$, $S_3 \parallel S_4$, e $S_5 \parallel S_6$ e $|S_1| = |S_2| = 4, |S_3| = 10, |S_4| = 9$. Observamos que os dois primeiros pares de paralelas: $\{S_1, S_2\}$ e $\{S_3, S_4\}$ seguem uma relação de "reflexão" considerando como espelho o centro da espiral, ou seja, um primo aparece em uma paralela e, o seguinte, na outra. Os caminhos S_1, S_2, S_3 e S_4 têm densidade 1. O último par de paralelas $\{S_5, S_6\}$ correspondem a valores de $k \geqslant 40$ no polinômio de Euler, sendo 1601 o valor da 1ª casa de S_5 um número primo para $k = 40$, e o número composto 1681 o valor da 1ª casa de S_6 para $k = 41$. Estas paralelas apresentam, também, uma relação de reflexão, porém "descontínua", ou seja, as vezes ocorrem dois ou mais primos seguidos em uma mesma paralela e, o seguinte, na outra paralela.

Os caminhos retilíneos S_5 ***e*** S_6**.**
Como descrito acima, os caminhos perpendiculares, superior S_5 e inferior S_6, pertencem à espiral de Ulam. Steven Bolde,[3], apresentou as **funções diagonais**: $f_\alpha(n)$ e $g_\alpha(n)$, $\alpha \in \mathbb{Z}$, que caracterizam todos os caminhos perpendiculares da espiral de Ulam, veja a Figura 13. Observamos que o caminho perpendicular que passa pela

casa central da diagonal principal da espiral de Ulam é descrita pela função $f_0(k)$, no caso do caminho perpendicular superior, e pela função $g_0(k)$, no caso do caminho perpendicular inferior. O autor define cada um dessas funções diagonais, por

$$f_\alpha(k) = \begin{cases} 4k^2 + (8\alpha - 10)k + (4\alpha^2 - 8\alpha + 7) & \text{se} \quad \alpha \geqslant 0 \\ 4k^2 - (8\alpha + 10)k + (4\alpha^2 + 12\alpha + 7) & \text{se} \quad \alpha < 0 \end{cases}$$

e

$$g_\alpha(k) = \begin{cases} 4k^2 + (8\alpha - 6)k + (4\alpha^2 - 8\alpha + 3) & \text{se} \quad \alpha \geqslant 0 \\ 4k^2 - (8\alpha + 6)k + (4\alpha^2 + 4\alpha + 3) & \text{se} \quad \alpha < 0 \end{cases}$$

E, verifica-se, computacionalmente, que as funções diagonais $f_{20}(k)$ e $g_{-20}(k)$ são as "maiores" funções geradores de números primos, na espiral de Ulam. Isto é condicente com os resultados obtidos, neste trabalho, pois os caminhos S_5 e S_6 tem associados os polinômios $p_1(k) = 4k^2+150k+1447 = f_{20}(k)$ e $p_2(k) = 4k^2+154k+1523 = g_{-20}(k)$, para todo $k = 1, 2, 3, \ldots$, respectivamente.

No $n-$quadrado Zeta. Na Figura 14(b) mostramos o perfil do polinômio de Euler em um $100-$qZeta. A 1ª casa do perfil está pintada de cor verde escuro. É importante salientar que somente foram impressos números primos das casas do perfil. Geometricamente, o perfil do polinômio de Euler apresenta as seguintes características: (a) As 9 primeiras casas só contem números primos e estão distribuídos aleatoriamente. (b) A partir da décima casa, seguem 4 caminhos retilíneos:

$\mathbf{C}_1 = (131, 151, 173, 197)$ $\quad$ $\mathbf{C}_2 = (223, 251, 281, 313, 347, 383, 421)$
$\mathbf{C}_3 = (461, 503, 547, \ldots, 1523, 1601)$ $\quad$ $\mathcal{V}_{41} = (1763, 1847, 1933, \ldots, 9941, \ldots)$

Os três primeiros caminhos: $\mathbf{C}_1, \mathbf{C}_2$ e $\mathbf{C}_3$ possuem 31 números primos, e somando com os 9 primeiros primos aleatórios, nos dão os 40 primeiros números primos do polinômio de Euler. Note que, a primeira casa do caminho vertical $\mathcal{V}_{41}$ é um número composto.

Foi observado, também, que as casas dos caminhos S_3 e S_4 coincidem com as casas do caminho $\mathbf{C}_3$, e as casas dos caminhos S_5 e S_6 são as mesmas que as do caminho vertical $\mathcal{V}_{41}$, exceto pela primeira casa de S_5, pois corresponde a uma casa da diagonal principal do $n-$qZeta, que é um quadrado perfeito.

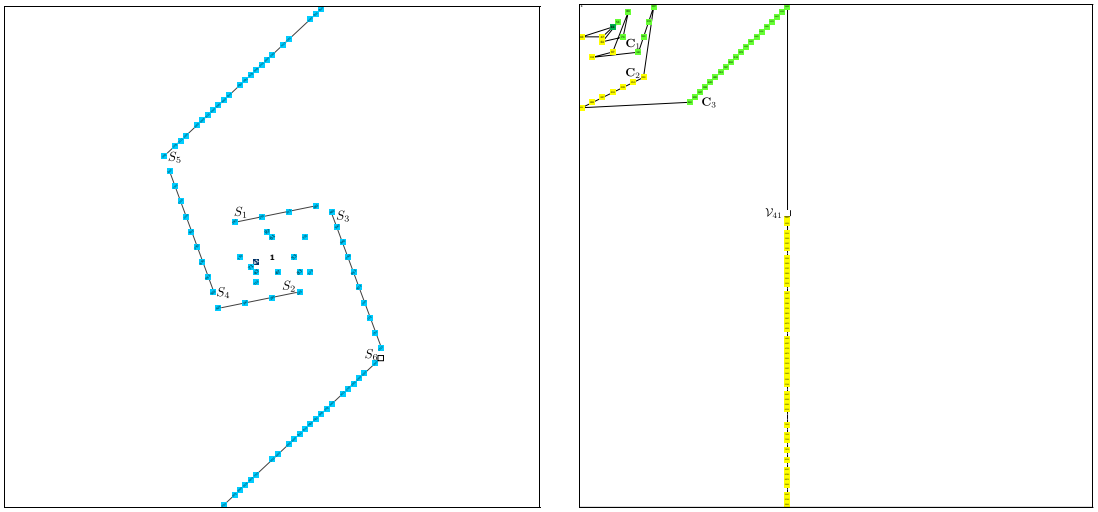
8. Conclusões

Com a elaboração do $n-$quadrado Zeta conseguiu-se: (a) Um novo arranjo dos números naturais para a identificação de padrões especiais de distribuição dos números primos. (b) Caminhos diagonais e verticais para encontrar sequências "longas" de números primos. (c) Polinômios geradores de números primos: $d_j(k)$ e $v_j(k)$. (d) O polinômio gerador de números primos, $v_{41}(k)$, $k = 1, 2, 3, \ldots$ que resultou ser uma Translação do polinômio de Euler. (e) Informações detalhadas do perfil do polinômio de Euler, tanto na espiral de Ulam, quanto no n-quadrado Zeta.

Agradecimento

À FAPEMIG por ter financiado o Projeto que deu origem a esse artigo.

FIGURA 14. O perfil do polinômio de Euler.



(a) Espiral de Ulam (em um 99−quadrado). (b) O 100−quadrado Zeta.

Fonte: Os autores (2022).

REFERÊNCIAS

[1] Aletheia-Zomlefer, S. L.; Fukshansky, L.; Garcia, S. R. The Bateman-Horn conjecture: Heuristics, history, and applications. **arXiv:1807.08899v4**, 1-47. 2019. https://arxiv.org/pdf/1807.08899.pdf

[2] Borwein, P. et al. **The Riemann Hypothesis**: A Resource for the Afficionado and Virtuoso Alike. 1. ed. Springer-Verlag, NY, 2008. https://doi.org/10.1007/978-0-387-72126-2

[3] Bode, S. Quadratic Polynomials With High Asymptotic Prime Density in The Ulam Spiral. **Research Gate**. Thesis, December 2018. https://www.researchgate.net/publication/341540065

[4] Clay Mathematics Institute. **Millennium Problems**: Riemann Hypothesis. 2021. Disponível em: https://www.claymath.org/millennium-problems/riemann-hypothesis. Acesso em: 03 set. 2021.

[5] Gourdon, X. The 10^{13} first zeros of the Riemann Zeta function, and zeros computation at very large height, October, 2004. http://numbers.computation.free.fr/Constants/Miscellaneous/zetazeros1e13-1e24.pdf

[6] Gowers, W. T. Erdős and Arithmetic Progressions. **ArXiv:1509.03421v1 [math.CO]**, 1-21. 2015. https://arxiv.org/pdf/1509.03421v1.pdf

[7] Green, B.; Tao, T. The primes contain arbitrarily long arithmetic progressions. **Annals of mathematics**, v. 167, p. 481-547, 2008. https://doi.org/10.4007/annals.2008.167.481

[8] Guy, R. K. **Unsolved Problems in Number Theory**. 3. ed. Springer-Verlag New York, LLC, 2004. https://doi.org/10.1007/978-0-387-26677-0

[9] Helfgott, H. A. The ternary Goldbach conjecture is true. **arXiv:1312.7748v2[math.NT]**, p. 1-79, 2014. https://arxiv.org/abs/1312.7748

[10] Jara-Vera, V.; Sánchez-Ávila, C. Distribution of adjacent prime numbers in the Ulam spiral. **Engineering Computations**, 1633-1651. 2020. https://doi.org/10.1108/EC-09-2019-0402

[11] Mollin, R. A. Prime-Producing Quadratics. **The American Mathematical Monthly**, 529-544. 1997. https://doi.org/10.2307/2975080

[12] Murty, M. R.; Srinivas, K.; Subramani, M. Admissible primes and Euclidean quadratic fields. **J. Ramanujan Math. Soc.**, v. 33, n. 2, p. 135-147. 2018. http://www.mathjournals.org/jrms/2018-033-002/2018-033-002-003.html

[13] OEIS. **Primes in arithmetic progression**. 2022. Disponível em: https://oeis.org/wiki/Primes_in_arithmetic_progression. Acesso em: 23 fev. 2022.

[14] Orlowski, A; Chmielewsky, L. J. Ulam Spiral and Prime-Rich Polynomials. **Computer Vision and Graphics. ICCVG 2018. Lecture Notes in Computer Science**, v. 10114. 2018. https://doi.org/10.1007/978-3-030-00692-1_45

[15] Serre, J-P. **A Course in Arithmetic**. 1 ed. Springer-Verlag New York, p 119, 1973. https://doi.org/10.1007/978-1-4684-9884-4

[16] Soifer, A. **The Mathematical Coloring Book**. 1. ed. Springer ggScience+Business Media, p 637, 2009. https://doi.org/10.1365/s13291-014-0101-y

[17] Stevenhagen, P.; Lenstra, H. W. Chebotarëv and his density theorem. **The Mathematical Intelligencer**, 18, 26-37. 1996. https://doi.org/10.1007/BF03027290

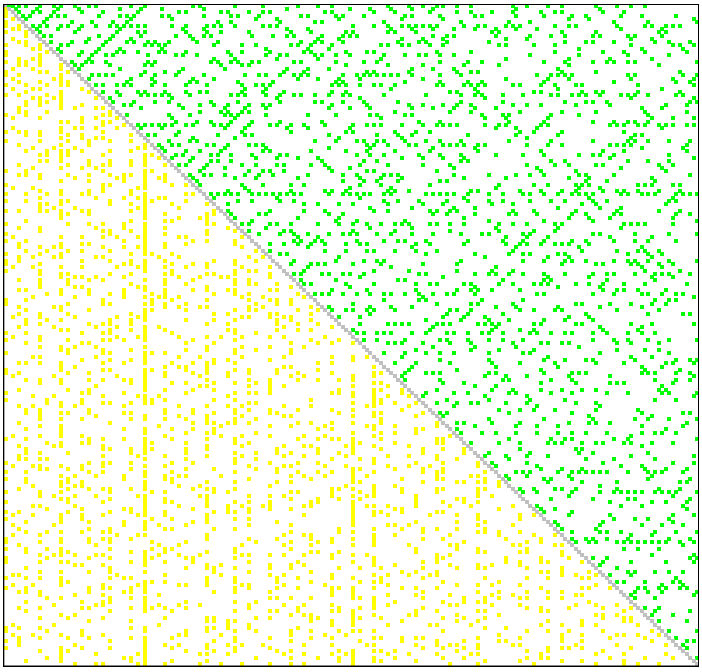
[18] Wang, Y. **The Goldbach Conjecture**. 2nd ed. World Scientific Co. Pte. Ltd, p 344, 2002. https://doi.org/10.1142/5096

[19] Weisstein, E. W. **Prime-Generating Polynomial**. Fonte: MathWorld - A Wolfram Web Resource. 2021. Disponível em: https://mathworld.wolfram.com/Prime-GeneratingPolynomial.html. Acesso em: 03 set. 2021.

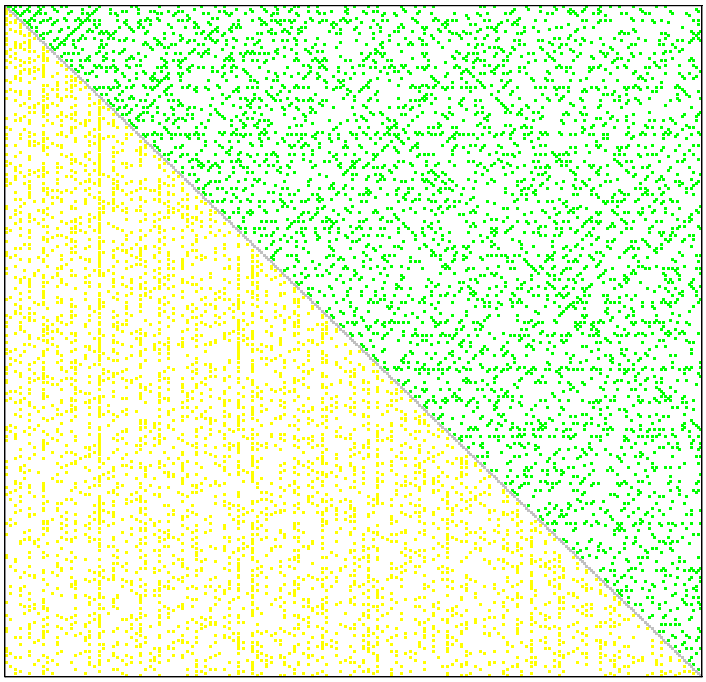
Departamento de Matemática e Estatística, Universidade Federal de São João del-Rei, MG, Brasil.

Email address: avila_jaj@ufsj.edu.br

FIGURA 15. O $n-$quadrado Zeta.



(a) O 200$-$quadrado Zeta.



(b) O 300$-$quadrado Zeta.

Fonte: Os autores (2022).

RMU
REVISTA MATEMÁTICA UNIVERSITÁRIA

SBM
SOCIEDADE BRASILEIRA DE MATEMÁTICA

Revista Matemática Universitária, vol. 1, 2022
ISSN: 2675-5254 — **DOI:** 10.21711/26755254/rmu20225

AS MUITAS FACETAS DA TEORIA DOS JOGOS: DA ECONOMIA À BIOLOGIA EVOLUTIVA

THAÍS FERNANDA MENDES MONIS, BRUNO APARECIDO PIM

Resumo. Nesse artigo, pretendemos divulgar ideias introdutórias da Teoria dos Jogos tais como os conceitos de jogos não-cooperativos, de solução para um jogo não-cooperativo - o assim chamado equilíbrio de Nash - também o conceito de equilíbrio social, bem como modelos aplicados ao estudo de questões em economia e em biologia evolutiva. Recentemente, publicamos o artigo [15] - um texto igualmente de divulgação da área de Teoria dos Jogos - onde já abordamos as noções de jogos não-cooperativos e o equilíbrio de Nash. A novidade aqui é: o desenvolvimento da noção de equilíbrio social, também chamado de ponto ótimo-Pareto, e uma análise crítica, por meio de exemplos, onde pomos em contra-ponto essas duas noções de solução para um jogo; a apresentação de uma demonstração para o Teorema de existência de equilíbrio de Nash fazendo uso do Teorema de ponto fixo de Brouwer; e uma discussão sobre como o conceito de equilíbrio de Nash é utilizado em Biologia a fim de se lançarem novas luzes sobre a teoria de evolução de Darwin.

Apesar de não ser objeto habitual presente nos cursos de Matemática, acreditamos que a riqueza de ideias apresentada na Teoria dos Jogos, bem como a sua história e vocação natural de produzir diálogo entre a Matemática e outras áreas do saber tais como a Economia, as Ciências Sociais e a Biologia, podem contribuir para uma formação mais humanizada e plena do estudante.

1. Introdução

Os seres humanos não seríamos capazes de sobreviver se não interagíssemos uns com os outros. E esse mesmo mecanismo de sobrevivência nos leva muitas vezes a conflitos, motivados pela busca de algum ganho ou bem estar pessoal, mas condicionados a uma interdependência recíproca. O estudo de economia, ciências políticas e, de modo geral, das ciências sociais, aparece como uma tentativa de entender o modo como os seres humanos interagem entre si e tomam suas decisões, tanto individuais quanto em grupo. O objetivo, geralmente, é o de aplicar o conhecimento científico

Data de aceitação: 12 de julho de 2022.
Palavras chave. Jogos não-cooperativos, equilíbrio de Nash, pontos ótimo-Pareto, equilíbrio evolutivamente estável.

à análise das questões pertinentes ao funcionamento das sociedades e de suas instituições tais como mercados financeiros, governos, instituições legais, etc. Cientistas intentam desenvolver modelos rigorosos, abstraindo a realidade que observam, não apenas para explicitá-la de modo sistematizado, mas também para poder predizer resultados e obter critérios para tomadas de decisão. Nessa direção, a Teoria dos Jogos se encarrega da construção de modelos matemáticos que descrevem situações de conflito e de cooperação entre tomadores de decisão racionais. O axioma de racionalidade colocado aqui significa que cada envolvido toma suas decisões com o objetivo de maximizar algum tipo de ganho pessoal.

Observamos que a Teoria dos Jogos não é objeto habitual dos programas de disciplinas de cursos de Matemática. No entanto, acreditamos que a riqueza de ideias nela apresentada, bem como a sua história e vocação natural de produzir diálogo entre a Matemática e outras áreas do saber tais como a Economia, as Ciências Sociais e a Biologia, podem contribuir para uma formação mais humanizada e plena do estudante. E é com o espírito de divulgar essas ideias que escrevemos esse artigo. Observamos ainda que, recentemente, publicamos o artigo [15] - um texto igualmente de divulgação da área de Teoria dos Jogos - onde já abordamos as noções de jogos não-cooperativos e o equilíbrio de Nash. Algumas novidades trazidas no presente texto são as seguintes: o desenvolvimento da noção de equilíbrio social, também chamado de ponto ótimo-Pareto, e uma análise crítica, por meio de exemplos, onde pomos em contra-ponto essas duas noções de solução para um jogo; a apresentação de uma demonstração para o Teorema de existência de equilíbrio de Nash fazendo uso do Teorema de ponto fixo de Brouwer; e uma discussão sobre como o conceito de equilíbrio de Nash é utilizado em Biologia a fim de se lançarem novas luzes sobre a teoria de evolução de Darwin.

2. O que é um jogo?

Na Teoria dos Jogos, uma linha de interesse é o estudo do comportamento estratégico dos envolvidos em uma situação de conflito ou de competição, situações que podem ir desde uma ingênua brincadeira de "pedra-papel-tesoura" até modelos de leilão, disputas eleitorais, conflitos entre nações para diminuição de emissão de CO_2 na atmosfera, etc. Tais situações podem ser modeladas pelo que chamamos de **jogo na forma normal**: explicita-se um conjunto finito de jogadores, $\{1, 2, \ldots, n\}$; para cada jogador $i \in \{1, \ldots, n\}$, explicita-se o seu conjunto de estratégias (ou movimentos possíveis), que denotaremos por S_i; e, dado que cada jogador escolheu o seu movimento, cada um terá um retorno, o qual não só dependerá da sua escolha de estratégia mas também das estratégias adotadas pelos demais jogadores. Esse retorno é modelado por uma função pagamento (*"payoff"*), que denotaremos por p_i, que quantifica quão vantajoso foi o resultado final para o jogador i. Como esse retorno p_i do jogador i depende das estratégias assumidas por todos os jogadores, p_i é uma função definida no produto cartesiano dos espaços de estratégias, $S = S_1 \times S_2 \times \cdots \times S_n$, e com valores reais. Esse conceito pode ser visto como um modelo para qualquer situação de competição na qual o resultado de cada competidor não depende apenas de suas escolhas mas também das escolhas dos demais envolvidos. O axioma de racionalidade que se coloca é que cada jogador i busca

fazer uma escolha s_i de modo a maximizar sua payoff $p_i(s_1, \ldots, s_n)$, entendido que, para cada $j \neq i$, o j-ésimo jogador está simultaneamente escolhendo s_j, também com o objetivo de maximizar o valor da sua payoff p_j.

Assim, sempre que um conjunto de indivíduos, empresas, partidos políticos etc., estiver em uma situação de interdependência recíproca, em que as decisões tomadas por cada agente influenciam não somente o seu próprio resultado como também os resultados dos demais envolvidos, diremos que esses agentes encontram-se em um **JOGO**.

Explicitamente, define-se um jogo não-cooperativo na forma normal por:

Definição 2.1. *Um **jogo não-cooperativo na forma normal** é constituído de três elementos:*

- *um conjunto finito de **jogadores,** $N = \{1, 2, ..., n\}$;*
- *a cada jogador $i \in N$ associa-se um conjunto S_i, chamado de seu **espaço de estratégias**;*
- *e cada jogador $i \in N$ possui uma função $p_i : S_1 \times \cdots \times S_n \to \mathbb{R}$, chamada de sua função **payoff** (função de retorno ou de pagamento).*

Assim, um jogo na forma normal será representado por uma terna $(N, (S_i)_{i\in N}, (p_i)_{i\in N})$.

O produto cartesiano dos espaços de estratégias, $S_1 \times \cdots \times S_n$, será denotado por S. Os elementos de S são chamados de **perfis de estratégias puras**.

Como ilustração da Definição 2.1, apresentamos o jogo *pedra-papel-tesoura* nessa linguagem. Nesse jogo, há sempre dois competidores, os quais nomearemos Jogador 1 e Jogador 2. O conjunto de estratégias do Jogador 1 é $S_1 = \{\text{pedra, papel, tesoura}\}$, três possíveis movimentos. O mesmo para o Jogador 2, $S_2 = \{\text{pedra, papel, tesoura}\}$. Como a regra do jogo é pedra ganha de tesoura, tesoura ganha de papel e papel ganha de pedra, podemos quantificar o retorno desses jogadores com a seguinte tabela:

Tabela 1. Jogo pedra-papel-tesoura

J.1 \ J.2	pedra	papel	tesoura
pedra	(0,0)	(0,1)	(1, 0)
papel	(1,0)	(0,0)	(0,1)
tesoura	(0,1)	(1,0)	(0,0)

Nessa tabela, a primeira coluna está representando as estratégias do Jogador 1, a primeira linha, as estratégias do Jogador 2 e, para as entradas com pares ordenados de números, a primeira coordenada do par representa o retorno do Jogador 1 e a segunda coordenada representa o retorno do Jogador 2 no respectivo perfil de estratégias, sendo que o valor 1 está sendo associado à situação de vitória e o valor 0 à situação de derrota do jogador.

Dada a definição de jogo na forma normal, informamos que a Teoria dos Jogos se encarrega da construção de modelos que descrevem situações de interação estratégica

entre tomadores de decisão racionais. Para cada uma dessas situações, é preciso explicitar os agentes envolvidos, os conjuntos de estratégias e as funções de retorno. A força motriz da Teoria dos Jogos é a busca por uma abordagem unificada no tratamento de qualquer situação de interação estratégica, por meio de um princípio universal que nos leve a prever os possíveis desdobramentos em cada uma dessas situações. A resposta para esse busca por um princípio fundamental foi dada pelo brilhante matemático John Forbes Nash Jr. em sua tese de doutorado, com o seu conceito de equilíbrio para jogos não-cooperativos. No artigo [12], John Forbes Nash Jr. apresenta os principais resultados de sua tese de doutorado. Dentre eles, o seu reconhecido conceito de soluçao para jogos não cooperativos - o chamado **equilíbrio de Nash** - e um teorema de existência de equilíbrio cuja demonstração envolve o Teorema de ponto fixo de Brouwer. Abaixo, apresentamos a definição.

Definição 2.2 (**Equilíbrio de Nash**). *Seja $(N, (S_i)_{i\in N}, (p_i)_{i\in N})$ um jogo na forma normal, $N = \{1, \ldots, n\}$. O perfil de estratégia pura $s^* = (s_1^*, s_2^*, ..., s_n^*) \in S$ é um **equilíbrio de Nash** para esse jogo se*

$$p_i(s^*) = p_i(s_1^*, \ldots, s_{i-1}^*, s_i^*, s_{i+1}^*, \ldots, s_n^*) \geq p_i(s_1^*, \ldots, s_{i-1}^*, s_i, s_{i+1}^*, \ldots, s_n^*)$$

para todo $s_i \in S_i$, $i = 1, \ldots, n$, ou seja, para cada $i \in \{1, \ldots, n\}$, se cada jogador $j \neq i$ assumir a estratégia s_j^*, o jogador i não consegue um resultado melhor do que aquele obtido assumindo a estratégia s_i^*.

A interpretação para o equilíbrio de Nash é que, em um tal perfil de estratégias, não há motivação para mudanças unilaterais por parte dos jogadores, ou seja, nenhum jogador possui incentivo para mudar a sua estratégia se os demais não o fizerem.

Note que o jogo *pedra-papel-tesoura* apresentado acima não possui equilíbrio de Nash. Apesar de nem todo jogo na forma normal possuir equilíbrio de Nash, John Nash demonstrou que, considerando cada jogador com um conjunto finito de estratégias, se considerarmos distribuições de probabilidade sobre as estratégias, então o jogo resultante **sempre** possui equilíbrio de Nash (2.5), o qual é chamado de **equilíbrio de Nash em estratégias mistas**. No jogo *pedra-papel-tesoura*, por exemplo, uma estratégia mista para o Jogador 1 é uma função $\sigma : \{\text{pedra, papel, tesoura}\} \to [0, 1]$ com $\sigma(\text{pedra}) + \sigma(\text{papel}) + \sigma(\text{tesoura}) = 1$, sendo que a interpretação é que σ representa a estratégia mista onde o jogador assume a probabilidade $\sigma(\text{pedra})$ de jogar pedra, $\sigma(\text{papel})$ de jogar papel e $\sigma(\text{tesoura})$ de jogar tesoura. Outro exemplo onde podemos trabalhar com estratégias mistas é em uma disputa de penaltis, onde jogador e goleiro escolhem os lados com probabilidades de acordo com as suas facilidades e características do oponente.

Dado um conjunto finito A, denotamos por $\Delta(A)$ o conjunto de todas as distribuições de probabilidade sobre A. Ou seja,

$$\Delta(A) := \left\{ \sigma : A \to [0, 1] \mid \sum_{a\in A} \sigma(a) = 1 \right\}.$$

Definição 2.3 (**Estratégias mistas**). *Seja $G = (N, (S_i)_{i\in N}, (p_i)_{i\in N})$ um jogo no qual cada conjunto de estratégias S_i é não-vazio e finito. Uma* ***estratégia mista*** *do jogador i é uma distribuição de probabilidade sobre o seu conjunto de estratégias S_i. Assim,*

$$\Delta(S_i) = \{\sigma_i : S_i \to [0,1] \mid \sum_{s_i \in S_i} \sigma_i(s_i) = 1\}$$

é o conjunto das estratégias mistas do jogador i.

Note que cada estratégia pura do jogador i, $s_i \in S_i$, pode ser vista como uma estratégia mista identificando s_i com a distribuição de probabilidade $\sigma_i : S_i \to [0,1]$ sobre S_i dada por: $\sigma_i(s_i) = 1$ e $\sigma_i(t) = 0$ para todo $t \neq s_i$.

Definição 2.4. *Seja $G = (N, (S_i)_{i\in N}, (p_i)_{i\in N})$ um jogo no qual cada conjunto de estratégias S_i é não-vazio e finito, $N = \{1, \ldots, n\}$. A extensão de G para um jogo de estratégias mistas é o jogo*

$$\Gamma = (N, (\Delta(S_i))_{i\in N}, (P_i)_{i\in N}),$$

onde a função payoff $P_i : \Delta = \Delta(S_1) \times \cdots \times \Delta(S_n) \to \mathbb{R}$ é dada pela extensão multilinear da payoff p_i. Isso significa que, dado $\sigma = (\sigma_1, \ldots, \sigma_n) \in \Delta(S_1) \times \cdots \times \Delta(S_n)$,

$$P_i(\sigma) = \sum_{(s_1,\ldots,s_n)\in S} \sigma_1(s_1)\sigma_2(s_2)\cdots\sigma_n(s_n)\, p_i(s_1, s_2, \ldots, s_n),$$

o qual é o retorno estimado do jogador i quando os jogadores assumem o perfil de estratégias $\sigma = (\sigma_1, \ldots, \sigma_n) \in \Delta$.

Nesse caso, um perfil de estratégias mistas $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$ é um ***equilíbrio de Nash em estratégias mistas*** *quando*

$$P_i(\sigma^*) = P_i(\sigma_1^*, \ldots, \sigma_{i-1}^*, \sigma_i^*, \sigma_{i+1}^*, \ldots, \sigma_n^*) \geq P_i(\sigma_1^*, \ldots, \sigma_{i-1}^*, \sigma_i, \sigma_{i+1}^*, \ldots, \sigma_n^*),$$

para todo $\sigma_i \in \Delta(S_i)$, $i = 1, \ldots, n$.

Note que se o jogo $G = (N, (S_i)_{i\in N}, (p_i)_{i\in N})$ admitir um equilíbrio de Nash em estratégias puras, digamos $s^* = (s_1^*, s_2^*, ..., s_n^*) \in S$, então $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*)$, onde

$$\sigma_i^*(t) = \begin{cases} 0, & \text{se } t \neq s_i^* \\ 1, & \text{se } t = s_i^* \end{cases},$$

é um equilíbrio de Nash em estratégias mistas. De fato, nesse caso, por um lado,

$$\begin{aligned} P_i(\sigma^*) &= \sum_{(s_1,\ldots,s_n)\in S} \sigma_1^*(s_1)\cdots\sigma_n^*(s_n)\, p_i(s_1, \ldots, s_n) \\ &= p_i(s_1^*, \ldots, s_n^*) = p_i(s^*). \end{aligned}$$

Por outro,

$$\begin{aligned} P_i(\sigma_1^*, \ldots, \sigma_{i-1}^*, \sigma_i, \sigma_{i+1}^*, \ldots, \sigma_n^*) &= \sum_{(s_1,\ldots,s_n)\in S} \sigma_1^*(s_1)\cdots\sigma_i(s_i)\cdots\sigma_n^*(s_n)\, p_i(s_1,\ldots,s_n) \\ &= \sum_{s_i\in S_i} \sigma_i(s_i)\, p_i(s_1^*,\ldots,s_{i-1}^*, s_i, s_{i+1}^*,\ldots,s_n^*) \\ &\leq \sum_{s_i\in S_i} \sigma_i(s_i)\, p_i(s^*) = p_i(s^*). \end{aligned}$$

Portanto,

$$P_i(\sigma^*) = P_i(\sigma_1^*, \ldots, \sigma_{i-1}^*, \sigma_i^*, \sigma_{i+1}^*, \ldots, \sigma_n^*) \geq P_i(\sigma_1^*, \ldots, \sigma_{i-1}^*, \sigma_i, \sigma_{i+1}^*, \ldots, \sigma_n^*),$$

para todo $\sigma_i \in \Delta(S_i)$, $i = 1, \ldots, n$.

A recíproca desse fato não é verdadeira, ou seja, um jogo pode admitir equilíbrio de Nash em estratégias mistas sem que haja um equilíbrio em estratégias puras, como é o caso do jogo *pedra-papel-tesoura* comentado anteriormente.

Teorema 2.5 (Teorema de Nash)**.** *Qualquer jogo de n jogadores na forma normal com conjuntos finitos de estratégias S_i para todos os jogadores possui um* ***equilíbrio de Nash*** *em estratégias mistas.*

Como veremos, a demonstração do teorema de equilíbrio de Nash é uma aplicação do Teorema de ponto fixo de Brouwer. Na verdade, esses dois resultados são equivalentes, como é demonstrado em [19] e em [21].

Finalizamos essa seção com um conceito de solução de natureza distinta do conceito de equilíbrio de Nash. Trata-se do conceito de ponto **ótimo-Pareto** de um jogo, também conhecido por **equilíbrio social**.

Definição 2.6 (**Ótimo-Pareto**)**.** *Um ponto $s \in S$ é dito ótimo-Pareto (ou eficiente no sentido de Pareto) para $p_1, \ldots, p_n$ quando satisfaz a seguinte condição:*

Se $p_i(s') \geq p_i(s)$ para todo $i \in \{1, \ldots, n\}$ então $p_i(s') = p_i(s)$ para todo $i \in \{1, \ldots, n\}$.

Essa definição traduz a ideia de Pareto de que uma situação é eficiente quando não for possível aumentar o lucro de um agente sem diminuir o lucro de outro.

Veja que $s \in S$ não ser ótimo-Pareto para $p_1, \ldots, p_n$ significa que existe $s' \in S$ tal que $p_i(s') \geq p_i(s)$ para todo i e $p_j(s') > p_j(s)$ para algum j. Assim, quando $S = S_1 \times \cdots \times S_n$ e $p_1, \ldots, p_n : S \to \mathbb{R}$ são as funções payoff de uma competição em n jogadores, com espaços de estratégia $S_1, \ldots, S_n$, se $s = (s_1, \ldots, s_n) \in S$ não é ótimo-Pareto, significa que existe uma solução $s^* = (s_1^*, \ldots, s_n^*) \in S$ tão boa quanto s para todos os jogadores e mais lucrativa para pelo menos um dos jogadores.

Como dissemos, a interpretação de um ponto de equilíbrio de Nash é que, nele, nenhum jogador possui motivação para mudar sua estratégia se os demais não o

fizerem. Isso não significa que um equilíbrio de Nash seja um resultado particularmente desejável para um jogo. Na verdade, devemos pensar num ponto de equilíbrio de Nash simplesmente como uma descrição do que é provável que aconteça em uma situação na qual os jogadores perseguem seus objetivos individuais sem qualquer tipo de cooperação, quer por falta de comunicação ou simplesmente por não existir o interesse em cooperar. Um exemplo clássico é o bem conhecido e estudado Dilema do Prisioneiro no qual existem dois jogadores (prisioneiros), I e II, que são chamados em salas isoladas para confessarem um crime que cometeram. A matriz de retorno (número de anos de prisão) é a seguinte:

Tabela 2. Dilema do Prisioneiro

I \ II	não confessa	confessa
não confessa	(1,1)	(5,0)
confessa	(0,5)	(2,2)

onde o par (a, b) indica a anos de prisão para I e b anos de prisão para II. O objetivo de cada prisioneiro, é claro, é passar o menor número de anos possível na prisão. O único equilíbrio de Nash para esse jogo é a situação em que ambos confessam o crime. Porém, solução mais interessante para ambos é aquela na qual ambos não confessam, que pode ser obtida caso os prisioneiros consigam agir de forma cooperativa entre si. Assim, o equilíbrio de Nash desse jogo não é eficiente no sentido de Pareto. Na verdade, no artigo [4], P. Dubey apresenta uma prova matemática de que os equilíbrios de Nash são, "em geral", ineficientes no sentido de Pareto. Vemos, então, que essas duas importantes definições dificilmente são equivalentes dentro de um jogo, ou seja, raramente correspondem à mesma combinação de estratégias. Para além de uma demonstração matemática, podemos fazer esta observação ao notarmos que, na Definição (2.2), o equilíbrio é obtido quando os jogadores utilizam as estratégias que mais os beneficiam, dadas as estratégias dos oponentes, se importando apenas com o *seu próprio ganho*, enquanto que na Definição (2.6), os jogadores também levam em consideração se o *outro jogador* será prejudicado ou não. Por isso, comumente um ponto ótimo-Pareto é chamado um *equilíbrio social*. O equilíbrio social no Dilema do Prisioneiro acima é a configuração estratégica em que ambos os prisioneiros não confessam.

3. A Demonstração do Teorema de Equilíbrio de Nash: uma aplicação do Teorema de Ponto Fixo de Brouwer

A demonstração de existência de equilíbrio dada por John Nash é uma bonita aplicação do Teorema de ponto fixo de Brouwer. Para o seu entendimento, alguns pontos preliminares são os seguintes.

Definição 3.1. *O simplexo canônico de dimensão n (ou n-simplexo canônico) é o subconjunto de $\mathbb{R}^{n+1}$ dado por*

$$\Delta^n = \{(t_0, \ldots, t_n) \in \mathbb{R}^{n+1} \mid \sum_{i=0}^{n} t_i = 1 \quad e \quad t_i \geq 0, \ \textit{para } i = 0, \ldots, n\}.$$

Abaixo, enunciamos o Teorema de ponto fixo de Brouwer. Uma demonstração para ele utilizando-se de métodos da Topologia Algébrica pode ser encontrada em [20]. Para outra abordagem ver [10].

Teorema 3.2 (Teorema de ponto fixo de Brouwer [2])**.** *Seja $K \subset \mathbb{R}^N$ um subconjunto compacto e convexo de um espaço euclidiano. Então, para toda função contínua $f : K \to K$, existe $x \in K$ tal que $f(x) = x$.*

Para uma demonstração do lema abaixo, ver [[6], Theorem 1, pág. 69].

Lema 3.3. *Sejam H um espaço de Hilbert e C um subconjunto fechado e convexo de H. Para cada $u_0 \in H$ existe um único $v_0 \in C$ tal que*

$$\|u_0 - v_0\| = \inf_{v \in C} \|u_0 - v\|.$$

A aplicação

$$u_0 \to v_0$$

é uma função contínua, a qual chamaremos de retração natural de H sobre C, e que denotaremos por $r : H \to C$. Se H é um espaço vetorial real então a retração natural r satisfaz a seguinte desigualdade variacional:

$$\langle u_0 - r(u_0), w - r(u_0) \rangle \leq 0, \ \textit{para todo } w \in C,$$

onde $\langle \ , \ \rangle$ denota o produto interno em H.

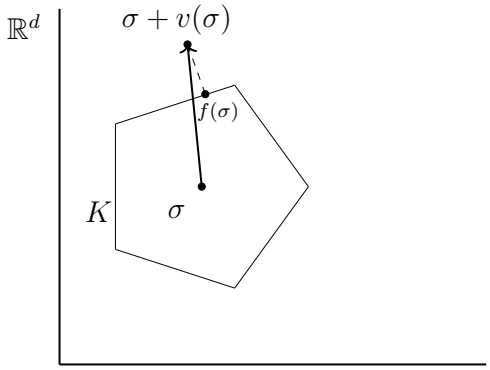
Demonstração do Teorema de Equilíbrio de Nash. Se $S_1, \ldots, S_n$ são os espaços de estratégia dos jogadores $1, 2, \ldots, n$, respectivamente, todos finitos e não vazios, o conjunto de estratégias mistas do jogador i, $\Delta(S_i)$, pode ser visto, sem nenhum prejuízo, como sendo o $(d_i - 1)$-simplexo canônico $\Delta^{d_i - 1} \subset \mathbb{R}^{d_i}$, onde d_i é a cardinalidade de S_i. Sob essa identificação, as funções payoff $P_i : \Delta^{d_1 - 1} \times \cdots \times \Delta^{d_n - 1} \to \mathbb{R}$, devido a sua multilinearidade, são da forma

$$P_i(\sigma_1, \ldots, \sigma_n) = \langle v_i(\sigma_1, \ldots, \sigma_n), \sigma_i \rangle + u_i(\sigma_1, \ldots, \sigma_n),$$

onde $v_i : \Delta^{d_1 - 1} \times \cdots \times \Delta^{d_n - 1} \to \mathbb{R}^{d_i}$ e $u_i : \Delta^{d_1 - 1} \times \cdots \times \Delta^{d_n - 1} \to \mathbb{R}$ são funções contínuas que não dependem da coordenada $\sigma_i \in \Delta^{d_i - 1}$, $1 \leq i \leq n$. Aqui, $\langle \ , \ \rangle$ denota o produto interno canônico de $\mathbb{R}^{d_i}$. Sejam $K = \Delta^{d_1 - 1} \times \cdots \times \Delta^{d_n - 1} \subset \mathbb{R}^d$, $d = d_1 + \cdots + d_n$, e $r : \mathbb{R}^d \to K$ a retração natural dada pelo Lema 3.3. Ainda, seja $f : K \to K$ a função definida por

$$f(\sigma) = r(\sigma + v(\sigma))$$

para todo $\sigma = (\sigma_1, \ldots, \sigma_n) \in K$, onde $v(\sigma) = (v_1(\sigma), \ldots, v_n(\sigma))$.



Pelo Teorema de Ponto Fixo de Brouwer, existe $\sigma^* = (\sigma_1^*, \ldots, \sigma_n^*) \in K$ tal que $f(\sigma^*) = \sigma^*$.

Afirmação: σ^* é um equilíbrio de Nash. De fato, como vimos, a retração natural $r : \mathbb{R}^d \to K$ é caracterizada pela desigualdade variacional:

$$\langle y - r(y), w - r(y) \rangle \leq 0, \text{ para todo } w \in K \text{ e todo } y \in \mathbb{R}^d. \tag{1}$$

Sejam $y = \sigma^* + v(\sigma^*)$ e $w = \sigma \in K$ qualquer, $\sigma = (\sigma_1, \ldots, \sigma_n)$. Por hipótese,

$$r(y) = r(\sigma^* + v(\sigma^*)) = f(\sigma^*) = \sigma^*.$$

Então, de (1), concluímos que

$$\langle v(\sigma^*), \sigma - \sigma^* \rangle \leq 0, \text{ para todo } \sigma \in K. \tag{2}$$

De (2), segue que

$$\langle v_i(\sigma^*), \sigma_i - \sigma_i^* \rangle \leq 0, \text{ para todo } \sigma_i \in \Delta^{d_i - 1}. \tag{3}$$

De fato, basta tomar o ponto $\sigma = (\sigma_i, \sigma_{-i}^*) \in K$ e aplicar (2).

Por (3), e desde que v_i e u_i não dependem de σ_i, temos

$$\begin{aligned} P_i(\sigma_i, \sigma_{-i}^*) &= \langle v_i(\sigma_i, \sigma_{-i}^*), \sigma_i \rangle + u_i(\sigma_i, \sigma_{-i}^*) \\ &= \langle v_i(\sigma^*), \sigma_i \rangle + u_i(\sigma^*) \\ &\leq \langle v_i(\sigma^*), \sigma_i^* \rangle + u_i(\sigma^*) \\ &= P_i(\sigma^*). \end{aligned}$$

Portanto, σ^* é um equilíbrio de Nash em estratégias mistas para o jogo dado.

□

No exemplo do Jogo "pedra-papel-tesoura" que descrevemos anteriormente, vimos que esse não admite equilíbrio de Nash. No entanto, quando consideramos estratégias mistas (distribuições de probabilidade sobre as estratégias puras), o Teorema de Nash nos garante a existência de pelo menos um equilíbrio. Pode-se verificar que o único equilíbrio em estratégias mistas do referido jogo é a situação em que cada

jogador assume cada uma das possibilidades de movimento (pedra, papel, tesoura) com probabilidade 1/3.

4. O clássico modelo de oligopólio devido a Cournot

Suponha n ($n \geq 2$) empresas, $\{1, 2, \ldots, n\}$, produtoras de um mesmo produto homogêneo. Denote por s_i a quantidade desse produto produzida pela empresa i e seja f_i sua função custo de produção, ou seja, para produzir a quantidade s_i, a empresa i tem um gasto de $f_i(s_i)$ reais. Suponha ainda que o preço de venda do produto no mercado, denotado por q, depende da quantidade total de produção, ou seja, $q = q(\sum_{j=1}^{n} s_j)$ (função demanda inversa). Supondo que o mercado consome a quantidade total produzida pelas empresas, $\sum_{j=1}^{n} s_j$, o lucro da empresa i é expresso por

$$p_i(s_1, \ldots, s_n) = q\left(\sum_{j=1}^{n} s_j\right) \cdot s_i - f_i(s_i).$$

Aqui, o dilema da empresa i é decidir qual a quantidade s_i de produção é mais interessante, lembrando que as empresas concorrentes estão simutaneamente com o mesmo dilema.

Por exemplo, se a situação é de duopólio (apenas duas empresas no mercado, $i \in \{1, 2\}$), com preço de produto no mercado dado por $q(s_1 + s_2) := 100 - (s_1 + s_2)$ e funções custo de produção $f_i(s_i) := 10s_i$, $i = 1, 2$, qual deve ser a previsão de quantidades produzidas por cada empresa? Com o princípio de que a empresa toma a sua decisão visando o lucro máximo, se a empresa 1 supõe que a empresa 2 decidirá pela quantidade s_2 de produção, ela buscará a quantidade s_1 que maximiza sua função lucro,

$$p_1(s_1, s_2) = (100 - s_1 - s_2)s_1 - 10s_1,$$

e se a empresa 2 supõe que a empresa 1 decidirá pela quantidade s_1, ela buscará a quantidade s_2 que maximiza sua função lucro,

$$p_2(s_1, s_2) = (100 - s_1 - s_2)s_2 - 10s_2.$$

O equilíbrio ocorre quando é alcançada uma configuração estratética $(\tilde{s}_1, \tilde{s}_2)$ tal que

$$\begin{cases} p_1(s_1, \tilde{s}_2) \leq p_1(\tilde{s}_1, \tilde{s}_2), & \text{para todo } s_1 \\ p_2(\tilde{s}_1, s_2) \leq p_2(\tilde{s}_1, \tilde{s}_2), & \text{para todo } s_2. \end{cases} \tag{4}$$

Dos métodos do Cálculo Diferencial, os candidatos a satisfazerem o Sistema 4 de desigualdades são os que satisfazem o sistema

$$\begin{cases} \dfrac{\partial p_1}{\partial s_1}(\tilde{s}_1, \tilde{s}_2) = 0 \\ \dfrac{\partial p_2}{\partial s_2}(\tilde{s}_1, \tilde{s}_2) = 0 \end{cases} \tag{5}$$

que possui única solução $\tilde{s}_1 = \tilde{s}_2 = 30$. Pode-se verificar que, de fato, $(30, 30)$ é o único equilíbrio de Nash para a situação acima. Uma vez alcançada tal configuração, nenhuma das empresas possuirá motivação para alterar unilateralmente a sua quantidade de produção.

Veja que no equilíbrio de Nash da situação hipotética acima os lucros serão de R$ 900,00 reais para cada empresa. Mas e se as empresas tivessem o interesse de cooperar entre si? Existiria um equilíbrio social entre elas? Ou seja, existiria um ponto ótimo-Pareto para esse jogo? Observamos que todo par (s_1^*, s_2^*) que maximiza $p_1(s_1, s_2) + p_2(s_1, s_2)$ é um ótimo-Pareto, não sendo necessariamente verdadeira a recíproca. De qualquer modo, quando analisamos a função

$$p_1(s_1, s_2) + p_2(s_1, s_2) = (90 - s_1 - s_2)(s_1 + s_2) = 2025 - (s_1 + s_2 - 45)^2 \quad (6)$$

vemos que seu valor máximo é atingido quando $s_1 + s_2 = 45$. Assim, toda configuração estratégica do tipo $(s_1, 45 - s_1)$ é um ótimo-Pareto do jogo. Dentre essas configurações, a que agradará igualmente ambas as empresas é aquela em que $s_1 = s_2 = 22,50$, e verifica-se que, aqui, os lucros serão de R$ 1012,50 reais para cada empresa.

5. A Teoria dos Jogos aplicada à Biologia Evolutiva

"Desde que Darwin leu o trabalho de Malthus, a Teoria de Evolução se beneficiou da interação entre a Ecologia e a Economia. A Teoria dos Jogos Evolutivos pertence a esta tradição: ela funde a ecologia populacional com a teoria dos jogos" (SIGMUND; NOWAK, 1999, p. 503, tradução nossa)[1]. A Teoria de Evolução de Darwin baseia-se no princípio da sobrevivência do mais apto, segundo o qual as novas gerações da flora e da fauna sofrem mutações. Um indivíduo que sofreu uma mutação irá passá-la para seus descendentes. Apenas aqueles indivíduos mais adaptados ao seu ambiente alcançam sucesso na luta pela sobrevivência. Decorre desse princípio que, no contexto do processo de evolução, todo organismo age como se fosse uma criatura racional, o que significa uma criatura cujo comportamento é direcionado para um objetivo: maximizar o número esperado de seus descendentes. Dizemos que o organismo age "como se fosse racional" para enfatizar que ele, por si só, não é uma criatura estrategicamente planejadora.

A Teoria dos Jogos tratava originalmente de problemas enfrentados por tomadores de decisão com interesses divergentes (por exemplo, empresas competindo por um mercado, como no Modelo de Oligopólio de Cournot). No contexto da biologia evolutiva, as duas noções básicas da teoria dos jogos, a saber, estratégia e recompensa, precisam ser reinterpretadas. Aqui, uma estratégia não é um curso de ação deliberada, mas um traço hereditário; a recompensa é a aptidão darwiniana (sucesso reprodutivo médio). Os "jogadores" são membros de uma população, todos competindo por uma parcela maior de descendentes. Se várias variantes de uma característica ocorrem em uma população, a seleção natural leva a um aumento na frequência daquelas variantes com maior aptidão. Se o sucesso de uma característica não depende de sua frequência, isso acabará por levar à fixação da variante ótima. Mas se o sucesso de um traço depende da frequência, então podemos modelar uma situação de competição, e isso pode ser analisado por meio da teoria dos jogos. A situação é semelhante ao que acontece na ecologia de populações. Se a presa for

[1] "Ever since Darwin read Malthus, the theory of evolution has benefited from the interaction of ecology with economics. Evolutionary game theory belongs to this tradition: it merges population ecology with game theory."

abundante, os predadores aumentarão por um tempo. Mas esse aumento reduz a abundância de presas e, eventualmente, leva a uma diminuição dos predadores. Se entendermos o número de descendentes de um organismo como uma recompensa, descrevemos um processo que é impulsionado pela maximização das recompensas. Uma vez que o conceito de equilíbrio em um jogo também se baseia na ideia de que apenas estratégias que maximizam os retornos esperados (contra as estratégias utilizadas pelos outros jogadores) serão escolhidas, temos uma motivação para usar ideias da teoria dos jogos para explicar fenômenos evolutivos. Em [17], Maynard Smith e Price mostraram que, de fato, é possível usar o conceito de equilíbrio de Nash para lançar uma nova luz sobre a teoria de Darwin. O exemplo que apresentaremos a seguir foi tirado e adaptado de [17] e contém as principais ideias e a abordagem geral dessa área de pesquisa.

5.1. **Exemplo: Tipos pacíficos e tipos agressivos.** Lutas intraespecíficas fornecem um primeiro exemplo de mudanças em uma população que dependem da frequência de uma característica.

Na discussão que se segue, suponha que um determinado animal possa exibir um de dois tipos de comportamento: o agressivo ou o pacífico. Os tipos diferentes de comportamento são expressos quando um animal invade o território de outro animal da mesma espécie. Um tipo agressivo escala numa luta para repelir o invasor até que a lesão de um dos competidores resolva a questão. Um tipo pacífico apresenta alguma exibição convencional (uma luta de empurrar, por exemplo, onde as lesões são praticamente nulas) e cede o território se o adversário aumentar o embate. Se um dos animais for do tipo agressivo e o outro do tipo pacífico, o resultado dessa luta é que o agressivo ficará no território, enquanto que o pacífico será expulso, ficando exposto a predadores e outros perigos. Se os dois animais forem pacíficos, um deles acabará saindo do território. Suponha que cada um deles saia com probabilidade $\frac{1}{2}$. Se os dois animais forem agressivos, inicia-se uma luta, durante a qual os dois são feridos, talvez fatalmente, e no máximo um deles permanecerá no território e produzirá descendentes. A tabela abaixo apresenta um exemplo de matriz que descreve o número esperado de descendentes de cada tipo de animal considerando os possíveis encontros.

TABELA 3. Número esperado de descendentes

		Invasor	
		Pacífico	Agressivo
Defensor	Pacífico	4, 4	2, 8
	Agressivo	8, 2	1, 1

Observe que o jogo na figura acima é simétrico; isto é, ambos os jogadores têm o mesmo conjunto de estratégias $S_1 = S_2$, e suas funções payoff satisfazem $p_1(s_1, s_2) = p_2(s_2, s_1)$, para cada $s = (s_1, s_2) \in S = S_1 \times S_2$.

Esse é um exemplo de uma população de espécie única, ou seja, uma população composta por apenas uma espécie animal, onde cada indivíduo pode exibir um de vários comportamentos possíveis.

O foco, agora, é entender o processo dinâmico que se desenvolve a partir de muitos encontros aleatórios entre indivíduos na população, juntamente com o aparecimento de mutações aleatórias.

O número esperado de descendentes de um indivíduo que encontra aleatoriamente outro indivíduo na população depende tanto de seu tipo quanto do tipo de indivíduo que encontrou; para ser mais preciso, o número esperado depende da probabilidade y de que o indivíduo encontrado seja do tipo pacífico (ou da probabilidade $1-y$ de que seja do tipo agressivo).

Cada composição populacional determina um único número real y $(0 \leq y \leq 1)$ que é a probabilidade de que um indivíduo escolhido aleatoriamente na população seja do tipo pacífico. Suponha agora que uma mutação, antes de nascer, possa "decidir" que tipo será (apresentar o tipo pacífico ou o tipo agressivo). Esta "decisão" e as subsequentes interações que a mutação terá com a população podem ser descritas pela matriz

Tabela 4. Matriz da dinâmica mutação x população

		População	
		Pacífico	Agressivo
		y	1-y
Mutação	Pacífico	4, 4	2, 8
	Agressivo	8, 2	1, 1

Na matriz acima, as colunas e as linhas representam os tipos de comportamento. Se tratarmos a matriz como um jogo, o jogador da coluna é a "população", que está implementando uma estratégia mista fixa $[y(\text{Pacífico}), (1-y)(\text{Agressivo})]$; ou seja, com probabilidade y, um indivíduo escolhido aleatoriamente na população será do tipo pacífico e com probabilidade $1-y$ será do tipo agressivo. O jogador da linha, que é a mutação, em contraste, escolhe seu tipo.

Se, por exemplo, a população é composta de 80% de animais do tipo pacífico $(y = 0,8)$ e 20% do tipo agressivo, e uma nova mutação é chamada a escolher o seu tipo quando nasce. O que a mutação "deveria escolher"? Se a mutação escolhe nascer do tipo pacífico, seu número esperado de descendentes é

$$(0,8) \times 4 + (0,2) \times 2 = 3,6 \text{ descendentes},$$

enquanto que se escolher nascer do tipo agressivo, seu número esperado de descendentes é

$$(0,8) \times 8 + (0,2) \times 1 = 6,6 \text{ descendentes}.$$

É, portanto, nessa situação, mais vantajoso para a mutação escolher nascer do tipo agressivo. Nenhuma mutação, é claro, tem a capacidade de decidir se será do tipo agressivo ou do tipo pacífico, porque essas características são ou herdadas ou o resultado de uma mudança aleatória na composição genética. O que acontece na prática é que indivíduos que possuem as características de um tipo agressivo irão se reproduzir mais que indivíduos que possuem as características de um tipo pacífico. Ao longo das gerações, o número dos do tipo agressivo irá aumentar, e a proporção de pacíficos para agressivos não será mais 80% para 20%. Uma população na qual

a proporção de pacíficos para agressivos é 80% : 20% é, portanto, evolutivamente instável.

Agora, se assumirmos que a mutação implementa uma estratégia mista fixa $[x(\text{Pacífico}), (1-x)(\text{Agressivo})]$, as payoffs são dadas por:

$$\textbf{Mutação: } p_1(x,y) = x(4y + 2(1-y)) + (1-x)(8y + (1-y))$$

$$\textbf{População: } p_2(x,y) = y(4x + 2(1-x)) + (1-y)(8x + (1-x))$$

Nesse jogo em estratégias mistas, pode-se mostrar que os equilíbrios de Nash são $(1/5, 1/5)$, $(0,1)$ e $(1,0)$. O primeiro corresponde à situação de uma população composta por 20% de seus elementos sendo do tipo pacífico e, das mutações, 20% "escolhe nascer" do tipo pacífico. Para o segundo, a população é toda composta por tipos pacíficos e uma mutação sempre "escolhe nascer" do tipo agressivo. E o terceiro é o oposto: população toda composta por tipos agressivos, com as mutações "escolhendo nascer" do tipo pacífico. Vemos que os equilíbrios de Nash $(0,1)$ e $(1,0)$ não são evolutivamente estáveis. Na verdade, a definição de estratégia evolutivamente estável é dada pelo seguinte:

Definição 5.1. *Uma estratégia mista x^* em um jogo simétrico de 2 jogadores é uma estratégia evolutivamente estável (EEE) se, para toda estratégia mista x diferente de x^*, existe $\varepsilon_0 = \varepsilon_0(x) > 0$ tal que, para todo $\varepsilon \in (0, \varepsilon_0)$,*

$$(1-\varepsilon)p_1(x,x^*) + \varepsilon p_1(x,x) < (1-\varepsilon)p_1(x^*,x^*) + \varepsilon p_1(x^*,x). \tag{7}$$

A interpretação biológica para essa definição é a seguinte. Uma vez que mutações ocorrem na natureza em uma base regular, estamos lidando com populações majoritariamente compostas de indivíduos "normais", com uma minoria de mutações. Interpretaremos x^* como a distribuição dos tipos entre os indivíduos normais. Considere uma mutação fazendo uso da estratégia x, e assuma que a proporção dessa mutação na população é ε. Todo indivíduo do tipo x irá encontrar um indivíduo do tipo x^* com probabilidade $1-\varepsilon$, recebendo, nesse caso, o retorno $p_1(x,x^*)$, e encontrará uma mutação do tipo x com probabilidade ε, recebendo, nesse caso, o retorno $p_1(x,x)$. A desigualdade dada em (7) nos diz, então, que em uma população na qual a proporção de mutações é ε, o retorno esperado de uma mutação (o lado esquerdo da desigualdade) é menor do que o retorno esperado de um indivíduo normal (o lado direito da desigualdade) e, assim, a proporção de mutações irá decrescer e, eventualmente, desaparecer com o tempo, com a composição da população retornando a ser majoritariamente do tipo x^*. Um "equilíbrio evolutivamente estável" é, portanto, uma estratégia mista do jogador coluna que corresponde à população que é imune a ser superada por mutações.

Uma propriedade das estratégias evolutivamente estáveis é a seguinte.

Teorema 5.2. *Se x^* é uma estratégia evolutivamente estável em um jogo simétrico de 2 jogadores, então (x^*, x^*) é um equilíbrio de Nash simétrico no jogo.*

Demonstração. Suponha x^* uma estratégia evolutivamente estável (EEE). Para mostrarmos que (x^*, x^*) é um equilíbrio de Nash, devemos provar que, para todo $x \neq x^*$,

$p_1(x, x^*) \leq p_1(x^*, x^*)$ e $p_2(x^*, x) \leq p_2(x^*, x^*)$. Como o jogo é simétrico, a desigualdade $p_2(x^*, x) \leq p_2(x^*, x^*)$ coincide com a primeira, $p_1(x, x^*) \leq p_1(x^*, x^*)$. Assim, provemos que, dado $x \neq x^*$, $p_1(x, x^*) \leq p_1(x^*, x^*)$.

Seja $x \neq x^*$ qualquer. Como x^* é EEE, existe $\varepsilon_0 > 0$ tal que, para todo $\varepsilon \in (0, \varepsilon_0)$,

$$(1-\varepsilon)p_1(x, x^*) + \varepsilon p_1(x, x) < (1-\varepsilon)p_1(x^*, x^*) + \varepsilon p_1(x^*, x).$$

Fazendo $\varepsilon \to 0$, obtemos

$$p_1(x, x^*) \leq p_1(x^*, x^*).$$

□

É importante salientarmos que a recíproca do teorema acima não é verdadeira, ou seja, existem jogos simétricos em 2 jogadores, com payoffs não negativas, admitindo equilíbrio de Nash simétrico (x^*, x^*) sem que x^* seja uma estratégia evolutivamente estável.

O exemplo que discutimos nessa seção é comumente encontrado na literatura como *Jogo Falcão-Pomba*, em analogia à situação da dinâmica populacional presa-predador. O tipo agressivo é identificado por falcão e o tipo pacífico por pomba. Em geral, a matriz de payoff que se coloca é a seguinte:

Tabela 5. Jogo Falcão-Pomba

	Falcão	Pomba
Falcão	(G-C)/2, (G-C)/2	G, 0
Pomba	0, G	G/2, G/2

onde G é o ganho em aptidão darwiniana resultante de vencer a competição, e C é o custo em aptidão darwiniana devido a uma lesão.

6. Discussões finais

Vimos que a Teoria dos Jogos é uma área de pesquisa que se encarrega da construção de modelos matemáticos que descrevem situações de competição e de cooperação entre tomadores de decisão racionais. Uma força motivadora da Teoria dos Jogos é a busca por uma abordagem unificada no tratamento de qualquer situação de interação estratégica, e de um princípio universal que nos leve a prever os possíveis desdobramentos em cada uma dessas situações, às vezes fornecendo até mesmo indicações aos tomadores de decisão sobre maneiras pelas quais eles podem alcançar seus objetivos. Inicialmente desenvolvida para dar conta de questões da Economia, a Teoria dos Jogos possui hoje aplicações nas mais diversas áreas do conhecimento humano tais como Ciência Política (e.g. [8]), Psicologia e Ciências Sociais (e.g. [16]), Ciência da Computação (e.g. [13]), Biologia (e.g. [14]), Engenharia (e.g. [5]), Relações Internacionais (e.g. [1]), Saúde (e.g. [18]), Filosofia (e.g. [3]), etc.

Com essas notas, esperamos ter instigado o leitor para as potencialidades dessa área de pesquisa cujo desenvolvimento firma a sua vocação de produzir diálogo entre a Matemática e diversas outras áreas do saber humano.

7. Agradecimentos

Gostaríamos de agradecer o parecerista anônimo pelos apontamentos de várias correções e sugestões que certamente contribuíram para melhorar o presente texto.

Referências

[1] Peter G. Bennett, *Modelling decisions in international relations: game theory and beyond*, Mershon International Studies Review (1995), vol. 39, Issue supplement 1, 19–52.

[2] Luitzen Egbertus Jan Brouwer, *Über Abbildung von Mannigfaltigkeiten*, Math. Ann. **71**, 97–115 (1911).

[3] B. Bruin, *Game Theory in Philosophy*, Topoi **24**, 197–208 (2005).

[4] Pradeep Dubey, *Inefficiency of Nash Equilibria*, Mathematics of Operations Research **11** (1986), no. 1, 1–8.

[5] Oleg Kapliński and Jolanta Tamošaitiene, *Game theory applications in construction engineering and management*, Ukio Technologinis ir Ekonominis Vystymas, (2010), vol. 16, n. 2, 348–363.

[6] David G. Luenberger, *Optimization by vector space methods*, John Wiley & Sons, Inc., 1969.

[7] Michael Maschler, Eilon Solan, Shmuel Zamir, *Game Theory*. Cambridge University Press, 2013.

[8] B. de Mesquita, *Game Theory, Political Economy, and the Evolving Study of War and Peace*, American Political Science Review, 100(4) (2006), 637–642.

[9] John Milnor, *A nobel prize for John Nash*, Math. Intelligencer **17** (1995), no. 3, 11–17.

[10] John Milnor, *Analytic Proofs of the "Hairy Ball Theorem" and the Brouwer Fixed Point Theorem*, The American Mathematical Monthly, vol. 85, no. 7 (1978), 521–524.

[11] Thaís Fernanda Mendes Monis, *Sobre Teoremas de Equilíbiro de Nash*, Tese de doutorado, ICMC/USP, 2010.

[12] John Forbes Nash, *Non-cooperative games*, Ann. of Math., **54** (1951), no. 2, 286–295.

[13] Yoav Shoham, *Computer science and game theory*, Communications of the ACM, vol. 51, Issue 8 (2008), 74–79.

[14] Karl Sigmund, Martin A. Nowak, *Evolutionary game theory*, Current Biology (1999), Vol. 9, no. 14, 503–505.

[15] Pim, B. A.; Souza, L. L. G.; Monis, T. F. M., *Uma breve introdução à teoria dos jogos*, C.Q.D. - Revista Eletrônica Paulista de Matemática, Bauru, v. 21, p. 69–80, dez. 2021.

[16] H. M. Samieh and K. Wahba, *Knowledge Sharing Behavior from Game Theory and Socio-Psychology Perspectives*, 2007 40th Annual Hawaii International Conference on System Sciences (HICSS'07), 2007, pp. 187c–187c.

[17] J. Maynard Smith, G. R. Price, *The logic of animal conflict*, Nature **246** (1973), 15–18.

[18] Kateřina Staňková; Joel S. Brown; William S. Dalton; et al, *Optimizing Cancer Treatment Using Game Theory - A Review*, JAMA Oncol. (2019), vol. 5, n. 1, 96–103.

[19] Juan Pablo Torres-Martínez, *Fixed points as Nash equilibria*, Fixed Point Theory Appl. 2006, Art. ID 36135, 4 pp.

[20] James W. Vick, *Homology Theory - an introduction to Algebraic Topology*, Second Edition, Springer-Verlag, 1994.

[21] Jingang Zhao, *The equivalence between four economic theorems and Brouwer's fixed point theorem,* working paper, Dept. Economics, Iowa State Univ., 2002; available at econ.iastate.edu/faculty/jingang/wp/fpt.pdf

Universidade Estadual Paulista (UNESP)
Instituto de Geociências e Ciências Exatas (IGCE)
Avenida 24 A, 1515, Bela Vista
Rio Claro, SP
13506-900
Email address: thais.monis@unesp.br, b.pim@unesp.br

RMU
REVISTA MATEMÁTICA UNIVERSITÁRIA

SBM
SOCIEDADE BRASILEIRA DE MATEMÁTICA

Revista Matemática Universitária, vol. 1, 2022
ISSN: 2675-5254 — **DOI:** 10.21711/26755254/rmu20226

COMPARAÇÃO DE TOPOLOGIAS SOBRE AS PARTES DE UM ESPAÇO MÉTRICO

MARCIO COLOMBO FENILLE

Resumo. No reticulado das topologias sobre as partes fechadas e limitadas de um espaço métrico X com mais que um ponto há duas particularmente importantes para as quais não há comparação imediata, a saber, a topologia métrica de Hausdorff e a topologia semifinita superiormente. Mostramos que estas topologias são comparáveis se, e somente se, não existem em X dois fechados não vazios e disjuntos, um dos quais limitado, com distância nula entre si, caso em que a topologia métrica de Hausdorff é estritamente mais fina que a topologia semifinita superiormente. Apresentamos alguns elementos típicos da diferença das duas topologias, considerando os casos em que o espaço contém ou não triângulos escalenos.

1. Introdução e teorema principal

A coleção de todas as topologias sobre um conjunto não vazio constitui um *reticulado* sob a ordem parcial de inclusão; ver [9]. Duas topologias que se relacionam segundo esta ordem são ditas *comparáveis* e, neste caso, aquela que contém a outra é dita *mais fina* que a outra, ou *estritamente mais fina*, caso a inclusão seja estrita. Decidir se duas dadas topologias sobre um mesmo conjunto são comparáveis e, no caso positivo, estabelecer a comparação, é um típico exercício inicial de qualquer curso de Topologia. No entanto, isto nem sempre é tarefa fácil. Mais que simples (ou não tão simples) exercício, estabelecer comparação entre topologias é um engenho por vezes decisivo para a resolução de problemas complexos, tendo em conta que algumas propriedades de um espaço podem ou não se manter quando trocamos sua topologia por uma mais fina ou por uma menos fina, e lidar com certos problemas assumindo uma ou outra topologia sobre o espaço pode ser decisivo para torná-lo mais simples ou mais complexo.

Este tipo de empreitada aparece com frequência em situações em que duas ou mais topologias sobre um dado conjunto possam ser consideradas, digamos, *naturais*, isto

Data de aceitação: 15 de julho de 2022.
Palavras chave. Comparação de topologias, partes fechadas e limitadas, topologia métrica de Hausdorff, topologia semifinita superiormente, conjuntos isósceles.

é, construídas naturalmente a partir de algo conhecido ou condição dada *a priori*. É um cenário assim que passamos a descrever.

Iniciamos com um espaço topológico (X,τ) e consideramos o conjunto $\mathfrak{p}(X)$, que chamamos as *partes* de X, constituído de todos os subconjuntos não vazios de X. A coleção das partes dos abertos não vazios de X, isto é, $\mathcal{B}=\{\mathfrak{p}(A):\varnothing\neq A\in\tau\}$ é base para uma topologia sobre $\mathfrak{p}(X)$ que chamamos a *topologia semifinita superiormente* e denotamos por $\tau_{\mathfrak{p}}$. Impossível dizer que esta não seja uma maneira natural de se construir uma topologia sobre $\mathfrak{p}(X)$ a partir de τ; aliás, esta naturalidade carrega o fato de que a função $\mu:(X,\tau)\to(\mathfrak{p}(X),\tau_{\mathfrak{p}})$ dada por $\mu(x)=\{x\}$ é contínua. Não só isso, esta topologia desempenha um papel importante na teoria de *funções a multivalores* (funções que a cada elemento do domínio associam um subconjunto não vazio do contradomínio). Explicamos: dada uma função a multivalores $f:Y\multimap X$ entre espaços topológicos, podemos considerá-la como uma função (no sentido usual) $f:Y\to\mathfrak{p}(X)$. Se consideramos sobre $\mathfrak{p}(X)$ a topologia semifinita superiormente, então $f:Y\multimap X$ é semicontínua superiormente (como função a multivalores) se, e somente se, $f:Y\to\mathfrak{p}(X)$ é contínua (como função no sentido usual). Para detalhes sobre este assunto recomendamos [7, 8].

Ocorre que, exceto no caso em que o espaço X é unitário (possui um único ponto), o espaço topológico $(\mathfrak{p}(X),\tau_{\mathfrak{p}})$ não possui boas propriedades de separação; ele não satisfaz sequer o axioma T_1 de separação (Lema 2.1) e, portanto, não é metrizável, ainda que (X,τ) o seja. Por óbvio, é certo dizer que $\tau_{\mathfrak{p}}$ depende apenas da topologia τ, sendo pois independente de qualquer métrica que acaso exista em X a induzir τ.

Suponhamos então que seja este o caso, isto é, que (X,τ) seja metrizável e que seja d uma métrica em X que induz a topologia τ. Não há sobre $\mathfrak{p}(X)$ uma métrica natural induzida por d, mas há sobre $\mathfrak{q}(X)$, o subconjunto de $\mathfrak{p}(X)$ formado pelos fechados não vazios e limitados de X, a *métrica de Haudorff* d_H (que não pode ser estendida para $\mathfrak{p}(X)$; ver Seção 3). A topologia induzida pela métrica d_H sobre $\mathfrak{q}(X)$ é chamada a *topologia métrica de Hausdorff* e denotada por τ_H. A naturalidade a que nos referimos sobre a métrica d_H contém a informação de que a função $\mu:(X,d)\to(\mathfrak{q}(X),d_H)$ dada por $\mu(x)=\{x\}$ é uma isometria.

Portanto, dado um espaço métrico X, acabamos de definir duas topologias sobre $\mathfrak{q}(X)$, a saber, a topologia métrica de Hausdorff τ_H, induzida pela métrica d de X, e a topologia $\tau_{\mathfrak{q}}$ que faz de $(\mathfrak{q}(X),\tau_{\mathfrak{q}})$ subespaço topológico de $(\mathfrak{p}(X),\tau_{\mathfrak{p}})$, sendo pois induzida pela topologia τ de X que, por sua vez, é induzida pela métrica d. Para que não reste dúvida, $\tau_{\mathfrak{q}}=\{V\cap\mathfrak{q}(X):V\in\tau_{\mathfrak{p}}\}$, e esta, assim como $\tau_{\mathfrak{p}}$, também será chamada topologia semifinita superiormente. Portanto, ao cabo, tanto τ_H quanto $\tau_{\mathfrak{q}}$ são induzidas pela métrica de X (embora $\tau_{\mathfrak{q}}$ dependa apenas da topologia), porém de maneiras distintas e com resultado distinto, em geral, como logo se vê.

Se X é um espaço métrico unitário, então $\mathfrak{q}(X)$ também o é, e, por conseguinte, $\tau_{\mathfrak{q}}=\tau_H$. Caso contrário, as topologias $\tau_{\mathfrak{q}}$ e τ_H são necessariamente distintas; de fato, $\tau_{\mathfrak{q}}$ não pode conter τ_H, uma vez que o espaço $(\mathfrak{q}(X),\tau_H)$ satisfaz o axioma T_1 de separação (por ser metrizável), mas o espaço $(\mathfrak{q}(X),\tau_{\mathfrak{q}})$ não satisfaz (Lema 2.2). Mais do que isso, há casos em que as topologias $\tau_{\mathfrak{q}}$ e τ_H não são sequer comparáveis. O resultado principal deste artigo, que anunciamos a seguir, fornece uma condição

necessária e suficiente, sobre o espaço X, para que as topologias $\tau_{\mathfrak{q}}$ e τ_H sejam comparáveis e, sob tal condição, estabelece a comparação.

Teorema 1.1. *Sejam X um espaço métrico com mais que um ponto e $\mathfrak{q}(X)$ a coleção dos fechados não vazios e limitados de X. A topologia semifinita superiormente $\tau_{\mathfrak{q}}$ e a topologia métrica de Hausdorff τ_H sobre $\mathfrak{q}(X)$ são comparáveis se, e somente se, não existem em X dois fechados não vazios e disjuntos, um dos quais limitado, com distância nula entre si; neste caso, τ_H é estritamente mais fina que $\tau_{\mathfrak{q}}$.*

Importa observar que a condição sobre o espaço X enunciada no Teorema 1.1 não depende apenas da topologia de X, mas também, como se haveria de esperar, de sua métrica. Isto ficará mais claro com os exemplos da Seção 4.

A prova do Teorema 1.1 é apresentada na Seção 5. Antes, nas Seções 2 e 3, relembramos alguns fatos sobre as topologias $\tau_{\mathfrak{q}}$ e τ_H. Na Seção 4 introduzimos uma terminologia para elucidar a hipótese sobre o espaço X expressa no Teorema 1.1.

Como já dito, sejam as topologias $\tau_{\mathfrak{q}}$ e τ_H comparáveis ou não, $\tau_{\mathfrak{q}}$ jamais contém τ_H (exceto se X é unitário). Aliás, o Lema 2.2 fornece elementos típicos da diferença $\tau_H \setminus \tau_{\mathfrak{q}}$, a saber, todos os conjuntos $\mathfrak{q}(X) \setminus \{\{x\}\}$ com $x \in X$. No caso em que $\tau_{\mathfrak{q}}$ e τ_H não são comparáveis, tem-se ainda que τ_H não contém $\tau_{\mathfrak{q}}$ e, conforme se vê na prova do Teorema 1.1, um elemento típico da diferença $\tau_{\mathfrak{q}} \setminus \tau_H$ é da forma $\mathfrak{p}(A) \cap \mathfrak{q}(X)$, onde $A \subset X$ é um aberto que contém um fechado não vazio e limitado C tal que $d(C, X \setminus A) = 0$. Este conjunto $\mathfrak{p}(A) \cap \mathfrak{q}(X)$ é um aberto básico do espaço $(\mathfrak{q}(X), \tau_{\mathfrak{q}})$, ao passo que o anteriormente considerado $\mathfrak{q}(X) \setminus \{\{x\}\}$, embora aberto do espaço $(\mathfrak{q}(X), \tau_H)$, pode não ser um aberto básico, ou seja, não ser uma bola aberta em $\mathfrak{q}(X)$ segundo a métrica d_H. Por outro lado, é óbvio que se a diferença $\tau_H \setminus \tau_{\mathfrak{q}}$ é não vazia, então existe uma bola aberta no espaço métrico $(\mathfrak{q}(X), d_H)$ que não pertence a $\tau_{\mathfrak{q}}$. O Teorema 6.1 e sua prova mostram como obter uma tal bola. Curiosamente, a prova deste teorema traz separadamente os casos em que o espaço X é discreto ou não discreto e, no segundo caso, considera que o espaço X pode ou não conter *triângulos* (subconjuntos de três pontos) escalenos.

Toda a terminologia de topologia geral e espaços métricos segue a referência [6].

Para finalizar esta introdução, enfatizamos que, além da topologia semifinita superiormente e da topologia métrica de Hausdorff, há outras topologias importantes sobre $\mathfrak{q}(X)$ ou, mais geralmente, sobre o conjunto dos fechados não vazios de X, conhecido na literatura como *hiperespaço* de X. Deixamos aqui duas referências: o artigo [2], um dos primeiros a reunir conteúdo substancial sobre o tema, e o livro [3], com muita informação consolidada.

2. A topologia semifinita superiormente

Sejam (X, τ) um espaço topológico e $\mathfrak{p}(X)$ o conjunto dos subconjuntos não vazios de X, que doravante chamamos as *partes* de X. A coleção das partes dos abertos de X, isto é, $\mathcal{B} = \{\mathfrak{p}(A) : \varnothing \neq A \in \tau\}$ é base para uma topologia sobre $\mathfrak{p}(X)$ chamada a *topologia semifinita superiormente* e denotada por $\tau_{\mathfrak{p}}$.

Se X possui mais que um ponto, a topologia $\tau_{\mathfrak{p}}$ não é discreta (ainda que τ seja discreta). De fato, dados $a, b \in X$, o conjunto $\{\{a\}, \{a, b\}\}$ é um subconjunto de $\mathfrak{p}(X)$ que não pertence a $\tau_{\mathfrak{p}}$ (verificação por conta do leitor ou da leitora). Mais que

isso, exceto no caso em que X é unitário, o espaço $(\mathfrak{p}(X), \tau_{\mathfrak{p}})$ não satisfaz sequer o axioma T_1 de separação. Esta é uma consequência do seguinte lema:

Lema 2.1. *Seja (X, τ) um espaço topológico com mais que um ponto. Para cada $x \in X$, o conjunto unitário $\{\{x\}\}$ não é fechado no espaço $(\mathfrak{p}(X), \tau_{\mathfrak{p}})$.*

Demonstração. Dado um ponto $x \in X$, como X tem mais que um ponto, temos que $X \in \mathfrak{p}(X) \setminus \{\{x\}\}$. Seja U uma vizinhança básica de X em $\mathfrak{p}(X)$. Então $X \in U = \mathfrak{p}(A)$ para algum A aberto em X. Segue que $A = X$ e, assim, $\{x\} \in U$, o que mostra que $\mathfrak{p}(X) \setminus \{\{x\}\}$ não é aberto. □

Suponhamos que o espaço (X, τ) seja metrizável e que seja d uma métrica em X que induz a topologia τ. Consideramos o subconjunto $\mathfrak{q}(X) \subset \mathfrak{p}(X)$ cujos elementos são as *partes fechadas e limitadas* de X, isto é, os subconjuntos de X que são não vazios, fechados e limitados. Observamos que ser fechado é uma propriedade topológica (depende apenas de τ) ao passo que ser limitado é uma propriedade métrica (depende da métrica fixada d induzindo τ).

A topologia $\tau_{\mathfrak{p}}$ induz sobre $\mathfrak{q}(X)$ uma topologia que denotamos por $\tau_{\mathfrak{q}}$ e continuamos a chamar *topologia semifinita superiormente*. Os abertos básicos do subespaço topológico $(\mathfrak{q}(X), \tau_{\mathfrak{q}})$ são da forma $\mathfrak{p}(A) \cap \mathfrak{q}(X)$ com $\varnothing \neq A \in \tau$.

Observamos que, em geral, $\mathfrak{p}(A) \cap \mathfrak{q}(X) \neq \mathfrak{q}(A)$. Por exemplo, se $X = [0, 3]$ com a métrica usual herdada da reta, então $A = (1, 3]$ é aberto em X e o intervalo $(1, 2]$ pertence a $\mathfrak{q}(A)$, por ser fechado em A, mas não pertence a $\mathfrak{p}(A) \cap \mathfrak{q}(X)$, por não ser fechado em X.

O resultado expresso no Lema 2.1 permanece válido para o subespaço $(\mathfrak{q}(X), \tau_{\mathfrak{q}})$.

Lema 2.2. *Seja X um espaço métrico com mais que um ponto. Para cada $x \in X$, o conjunto unitário $\{\{x\}\}$ não é fechado no espaço topológico $(\mathfrak{q}(X), \tau_{\mathfrak{q}})$.*

Demonstração. Dado $x \in X$, seja $C \in \mathfrak{q}(X)$ o conjunto formado por x e por um segundo ponto de X. Seja U uma vizinhança básica de C no espaço $(\mathfrak{q}(X), \tau_{\mathfrak{q}})$. Então $U = \mathfrak{p}(A) \cap \mathfrak{q}(X)$ para algum aberto A de X contendo C. Segue que $\{x\} \in U$. Portanto, toda vizinhança de C em $\mathfrak{q}(X)$ contém $\{x\}$, o que implica que o conjunto unitário $\{\{x\}\}$ não é fechado. □

Segue do Lema 2.2 que o espaço $(\mathfrak{q}(X), \tau_{\mathfrak{q}})$ não satisfaz o axioma T_1 de separação e, por conseguinte, não é metrizável.

3. A topologia métrica de Hausdorff

Seja (X, d) um espaço métrico e, como antes, seja $\mathfrak{q}(X)$ o conjunto das partes fechadas e limitadas de X. Para cada subconjunto não vazio $C \subset X$ e cada número real $\varepsilon > 0$, chamamos de *ε-vizinhança* de C em X ao aberto $\mathcal{V}_\varepsilon(C) = \{x \in X : d(x, C) < \varepsilon\}$. Claro que para um subconjunto unitário $\{x\} \subset X$, sua ε-vizinhança $\mathcal{V}_\varepsilon(\{x\})$ é simplesmente a bola aberta $B_\varepsilon(x)$ de centro x e raio ε em X.

A *métrica de Hausdorff* d_H sobre $\mathfrak{q}(X)$ é definida como segue: dados $A, B \in \mathfrak{q}(X)$,

$$d_H(A, B) = \inf\{\varepsilon > 0 : A \subset \mathcal{V}_\varepsilon(B) \text{ e } B \subset \mathcal{V}_\varepsilon(A)\}.$$

Uma prova de que d_H é de fato uma métrica sobre $\mathfrak{q}(X)$ pode ser vista em [7, Proposição 4.3]. Em [1, Exercício 53 do Capítulo 1] há a seguinte definição alternativa para a métrica de Hausdorff: dados $A, B \in \mathfrak{q}(X)$,

$$d_H(A,B) = \max\Big\{ \sup_{a\in A} d(a,B),\ \sup_{b\in B} d(b,A)\Big\}.$$

Observações simples mas que justificam a atenção voltada apenas às partes *fechadas e limitadas* do espaço X são as seguintes: (i) a métrica de Haudorff não pode ser estendida para o conjunto das partes (apenas) limitadas, uma vez que se um subconjunto $A \subset X$ é limitado mas não é fechado, então A está estritamente contido em seu fecho $\mathrm{cl}(A)$, que é ainda limitado, e tem-se $d_H(A, \mathrm{cl}(A)) = 0$; (ii) do mesmo modo, d_H não pode ser estendida para as partes (apenas) fechadas, pois se o espaço X não é limitado, dados um ponto $x \in X$ e um subconjunto fechado e ilimitado $A \subset X$ (por exemplo, $A = X$), não existe $\varepsilon > 0$ tal que $A \subset \mathcal{V}_\varepsilon(\{x\})$, donde $d_H(\{x\}, A)$ não está definido.

Os seguintes são alguns exemplos simples para auxiliar o leitor ou a leitora que pela primeira vez se depara com a métrica de Hausdorff:

(1) Dados $x, y \in X$, tem-se $d_H(\{x\},\{y\}) = d(x,y)$; assim, a função $\mu : (X,d) \to (\mathfrak{q}(X), d_H)$ dada por $\mu(x) = \{x\}$ é uma isometria.

(2) Se $x \in X$ e $B \in \mathfrak{q}(X)$, então $d_H(\{x\}, B) = \sup_{b\in B} d(x,b)$.

(3) Para os subconjuntos $A = \{(x,0) : 0 \le x \le a\}$ e $B = \{(x,y) : x^2 + y^2 \le b^2\}$ do plano euclidiano, com $a \ge b > 0$ dados, tem-se $d_H(A,B) = \max\{a-b, b\}$.

A topologia métrica induzida sobre $\mathfrak{q}(X)$ pela métrica d_H é chamada a *topologia métrica de Hausdorff* e denotada por τ_H.

Dados $C \in \mathfrak{q}(X)$ e um número real $r > 0$, a bola aberta $B_r^H(C)$ em $\mathfrak{q}(X)$, de centro C e raio r, pode ser caracterizada como segue: para cada $A \in \mathfrak{q}(X)$,

$$A \in B_r^H(C) \ \Leftrightarrow\ \exists\ \varepsilon \in (0,r)\ :\ A \subset \mathcal{V}_\varepsilon(C) \text{ e } C \subset \mathcal{V}_\varepsilon(A).$$

Exemplo simples mas elucidativo é o da bola aberta cujo centro é um subconjunto unitário $\{x\}$ de X; temos:

$$B_r^H(\{x\}) = \bigcup_{\varepsilon\in(0,r)} \mathfrak{p}(B_\varepsilon(x)) \cap \mathfrak{q}(X).$$

4. Fechados limitadamente assintóticos

Nesta breve seção introduzimos um conceito para simplificar a referência e elucidar a hipótese expressa no Teorema 1.1 como sendo equivalente a comparabilidade das topologias métrica de Hausdorff e semifinita superiormente.

Definição 4.1. Diz-se que um espaço métrico (X,d) tem a *propriedade dos fechados limitadamente assintóticos* (a *propriedade* F.L.A.) se existem fechados não vazios e disjuntos $C, D \subset X$, um dos quais limitado, tais que $d(C,D) = 0$.

Lembramos que se $C, D \subset X$ são fechados não vazios e disjuntos, um dos quais é compacto, então tem-se necessariamente $d(C,D) > 0$. Por conseguinte:

(1) Os espaços euclidianos não tem a propriedade F.L.A., já que seus subconjuntos fechados e limitados são compactos.

(2) Se X é compacto, então X não tem a propriedade F.L.A., uma vez que seus subconjuntos fechados são compactos.

Para ilustrar, apresentamos exemplo de um espaço métrico que tem a propriedade F.L.A. Seja X o plano cartesiano munido da métrica euclidiana d. Considere sobre X a métrica limitada padrão $\bar{d}$ induzida por d, isto é, $\bar{d}(x,y) = \min\{d(x,y), 1\}$. As métricas d e $\bar{d}$ induzem a mesma topologia sobre X, já que as bolas abertas de raio menor que 1 coincidem sob ambas as métricas. Considere os subconjuntos $C = \{(x,0) : x \in \mathbb{R}\}$ e $D = \{(x,1/x) : x \in (0,\infty)\}$. Temos que C e D são fechados não vazios e disjuntos e, sob a métrica $\bar{d}$, ambos são limitados. Além disso, $\bar{d}(C,D) = 0$. Portanto, o espaço métrico $(X,\bar{d})$ tem a propriedade F.L.A.

Uma forma alternativa que utilizamos na introdução do artigo para expressar a propriedade F.L.A. é a seguinte: *(X,d) tem a propriedade* F.L.A. *se existe um aberto A em X contendo um fechado não vazio e limitado C tal que $d(X \setminus A, C) = 0$.*

5. Prova do teorema principal

Com a terminologia introduzida na Seção 4, o Teorema 1.1 pode ser enunciado alternativamente como segue: *para um espaço métrico X com mais que um ponto, as topologias $\tau_{\mathfrak{q}}$ e τ_H sobre $\mathfrak{q}(X)$ são comparáveis se, e somente se, X não tem a propriedade* F.L.A.; *neste caso, τ_H é estritamente mais fina que $\tau_{\mathfrak{q}}$.*

Prova do Teorema 1.1. Seja (X,d) um espaço métrico com mais que um ponto.

Seja $x \in X$ um ponto. Pelo Lema 2.2, $\mathfrak{q}(X) \setminus \{\{x\}\}$ não pertence a $\tau_{\mathfrak{q}}$. Por outro lado, como τ_H é uma topologia métrica, o conjunto unitário $\{\{x\}\}$ é fechado em $\mathfrak{q}(X)$ segundo τ_H, donde $\mathfrak{q}(X) \setminus \{\{x\}\}$ pertence a τ_H. Isto mostra que *τ_H não está contida em $\tau_{\mathfrak{q}}$*. Observamos que esta conclusão é independente do espaço X ter ou não ter a propriedade F.L.A.

Suponha que X tenha a propriedade F.L.A. e sejam $C, D \subset X$ fechados não vazios e disjuntos, com C limitado, tais que $d(C,D) = 0$. Então $C \in \mathfrak{q}(X)$ e $A = X \setminus D$ é um aberto em X contendo C, o que implica que $U = \mathfrak{p}(A) \cap \mathfrak{q}(X)$ é uma vizinhança de C em $\mathfrak{q}(X)$ segundo a topologia $\tau_{\mathfrak{q}}$. Afirmamos que U não pertence a τ_H. De fato: dado um número real $r > 0$, como $d(C,D) = 0$, existe $y \in D$ tal que $d(y,C) < r/2$. Seja $C' = C \cup \{y\}$. Então $C' \in \mathfrak{q}(X)$ e temos $C' \subset \mathcal{V}_{r/2}(C)$ e $C \subset \mathcal{V}_{r/2}(C')$, donde $C' \in B_r^H(C)$. Por outro lado, $C' \notin U$, pois $y \notin A$. Isto mostra que não existe $r > 0$ tal que $B_r^H(C) \subset U$ e, assim, U não é aberto em $\mathfrak{q}(X)$ segundo τ_H. Portanto, *$\tau_{\mathfrak{q}}$ não está contida em τ_H*. Como antes concluímos que τ_H não está contida em $\tau_{\mathfrak{q}}$, segue que as topologias $\tau_{\mathfrak{q}}$ e τ_H não são comparáveis.

Suponha que X não tenha a propriedade F.L.A. Vamos provar que $\tau_{\mathfrak{q}}$ está contida em τ_H. Seja U um aberto básico de $\mathfrak{q}(X)$ segundo $\tau_{\mathfrak{q}}$, isto é, $U = \mathfrak{p}(A) \cap \mathfrak{q}(X)$, onde A é um aberto não vazio de X. Seja C um elemento de U. Então C é um fechado não vazio e limitado de X, com $C \subset A$. Segue que $D = X \setminus A$ é um fechado de X disjunto de C. Como X não tem a propriedade F.L.A., a distância $r = d(C,D)$ é positiva. Afirmamos que $B_r^H(C) \subset U$. De fato: se $K \in B_r^H(C)$, então existe $\varepsilon \in (0,r)$ tal que $K \subset \mathcal{V}_\varepsilon(C) \subset A$, o que implica que $K \in U$. Portanto, *$\tau_{\mathfrak{q}}$ está*

contida em τ_H. Como antes concluímos que τ_H não está contida em $\tau_{\mathfrak{q}}$, segue que τ_H é estritamente mais fina que $\tau_{\mathfrak{q}}$. □

6. Elementos típicos da diferença das topologias

Nesta seção final cumprimos nosso compromisso, assumido na introdução do artigo, de descrever bolas abertas em $(\mathfrak{q}(X), d_H)$ que não pertencem à topologia $\tau_{\mathfrak{q}}$.

Teorema 6.1. *Seja X um espaço métrico com mais que um ponto. Então existem um real $r > 0$ e um subconjunto $F \subset X$, com dois ou três pontos (logo $F \in \mathfrak{q}(X)$), tais que $B_r^H(F) \notin \tau_{\mathfrak{q}}$ e, portanto, $B_r^H(F)$ é um elemento da diferença $\tau_H \setminus \tau_{\mathfrak{q}}$.*

Para provar o Teorema 6.1 consideramos separadamente o caso em que o espaço X é discreto e o caso em que X não é discreto. No segundo caso, utilizamos dois resultados preliminares envolvendo a (in)existência de triângulos escalenos em X.

Por um *triângulo* em um espaço métrico (X, d) entendemos um subconjunto $\{x_1, x_2, x_3\} \subset X$ composto de três pontos distintos. Diz-se que um tal triângulo é *escaleno* se as três distâncias $d(x_1, x_2)$, $d(x_1, x_3)$ e $d(x_2, x_3)$ são distintas; caso contrário, diz-se que o triângulo é *isósceles*. Um número real $r > 0$ é chamado uma *distância factível* em (X, d) se existem $x, y \in X$ tais que $r = d(x, y)$.

Lema 6.2. *Seja X um espaço métrico com quatro pontos no qual todos os triângulos são isósceles. Então há no máximo três distâncias factíveis em X.*

Demonstração. Seja $X = \{x_1, x_2, x_3, x_4\}$. As distâncias factíveis em X são $d(x_1, x_2)$, $d(x_1, x_3)$, $d(x_1, x_4)$, $d(x_2, x_3)$, $d(x_2, x_4)$ e $d(x_3, x_4)$, e queremos provar que no máximo três delas são distintas. Suponhamos, ao contrário disso, que (pelo menos) quatro destas distâncias sejam distintas. Então os quatro pontos de X comparecem na realização destas distâncias, pois se apenas três comparecessem teríamos apenas três distâncias. Além disso, ou todos os pontos comparecem duas vezes ou há um ponto que comparece três vezes. No primeiro caso, as quatro distâncias distintas seriam $d(x_i, x_j), d(x_i, x_k), d(x_j, x_l)$ e $d(x_k, x_l)$ para alguma permutação (i, j, k, l) de $(1, 2, 3, 4)$. Como o triângulo $\{x_i, x_j, x_k\}$ deve ser isósceles, a distância $d(x_j, x_k)$ deve ser igual a $d(x_i, x_j)$ ou $d(x_i, x_k)$. Por outro lado, também o triângulo $\{x_j, x_k, x_l\}$ deve ser isósceles, o que força que $d(x_j, x_k)$ seja igual a $d(x_j, x_l)$ ou $d(x_k, x_l)$. Assim chegamos a uma contradição. No segundo caso, isto é, se há um ponto de X que comparece três vezes na realização das quatro distâncias distintas, então estas distâncias devem ser $d(x_a, x_b), d(x_a, x_c), d(x_a, x_d)$ e $d(x_b, x_c)$ para alguma permutação (a, b, c, d) de $(1, 2, 3, 4)$. Mas disto resulta que o triângulo $\{x_a, x_b, x_c\}$ é escaleno, o que está em contradição com a hipótese sobre X. Portanto, há no máximo três distâncias factíveis em X, conforme enunciado. □

Os seguintes são modelos simples para um espaço métrico $X = \{x_1, x_2, x_3, x_4\}$ que tem três distâncias factíveis e cujos triângulos são todos isósceles:

(1) No plano cartesiano, sejam x_1, x_2 e x_3 os vértices de um triângulo isósceles cujos ângulos internos, medidos em graus, não sejam $15, 15, 150$ ou $30, 75, 75$ ou $60, 60, 60$ ou $120, 30, 30$, e seja x_4 o circuncentro do triângulo.

(2) Para um modelo no espaço tridimensional, iniciamos com um triângulo isósceles $\{x_1, x_2, x_3\}$ que não seja equilátero, e consideramos a semirreta L que inicia no circuncentro e é perpendicular ao plano do triângulo. Para qualquer ponto x_4 escolhido sobre L, o espaço $X = \{x_1, x_2, x_3, x_4\}$ terá todos os seus triângulos isósceles. Além disso, há apenas duas maneiras de escolher o ponto x_4 em L de modo que haja apenas duas distâncias factíveis em X, havendo três para todas as outras escolhas.

O leitor ou a leitora terá divertimento ao verificar os detalhes destes exemplos.

Lema 6.3. *Seja X um espaço métrico não discreto em que todos os triângulos são isósceles. Então existe um subespaço $X' = \{x_1, x_2, x_3, x_4\}$ de X no qual as distâncias factíveis satisfazem*

$$0 < d(x_1, x_4) < d(x_1, x_3) = d(x_3, x_4) < d(x_1, x_2) = d(x_2, x_3) = d(x_2, x_4).$$

Demonstração. Seja $x_1 \in X$ um ponto de acumulação de X e seja $x_2 \in X$ um ponto qualquer diferente de x_1. Então existem $x_3, x_4 \in X$ tais que

$$0 < 4d(x_1, x_4) < 2d(x_1, x_3) < d(x_1, x_2).$$

Como todos os triângulos em X são isósceles, o mesmo ocorre com os triângulos em $X' = \{x_1, x_2, x_3, x_4\}$. Pelo Lema 6.2, existem (no máximo, e, portanto) exatamente três distâncias factíveis em X', a saber, $d(x_1, x_2)$, $d(x_1, x_3)$ e $d(x_1, x_4)$. Além disso, temos: (i) como o triângulo $\{x_1, x_3, x_4\}$ é isósceles e $d(x_1, x_4) < d(x_1, x_3)/2$, segue que $d(x_3, x_4) = d(x_1, x_3)$; (ii) como o triângulo $\{x_1, x_2, x_4\}$ é isósceles e $d(x_1, x_4) < d(x_1, x_2)/2$, segue que $d(x_2, x_4) = d(x_1, x_2)$; (iii) como o triângulo $\{x_2, x_3, x_4\}$ é isósceles e $d(x_3, x_4) = d(x_1, x_3) < d(x_1, x_2)/2$, segue que $d(x_2, x_3) = d(x_2, x_4)$. □

Provados estes lemas, passamos à prova do Teorema 6.1.

Demonstração do Teorema 6.1. Conforme antecipado, consideramos dois casos:

• X DISCRETO: provaremos que *para cada conjunto finito $F \subset X$, com mais que um ponto, existe um número real $r > 0$ tal que $B_r^H(F) = \{F\} \notin \tau_{\mathfrak{q}}$.*

Seja $F = \{x_1, \ldots, x_n\} \subset X$, com $n \geq 2$. Então F é aberto, fechado e limitado.

Suponha $F \neq X$. Como F é finito e $X \setminus F$ é fechado, a distância $\delta = d(F, X \setminus F)$ é positiva. Considere o real $\varepsilon = \min\{d(x_i, x_j) : x_i, x_j \in F,\ x_i \neq x_j\}$ e seja $r = \min\{\varepsilon, \delta\}$. Então $r > 0$ e temos $B_r^H(F) = \{F\}$. De fato: dado $K \in \mathfrak{q}(X)$, temos:

(i) se K não está contido em F, então existe $x \in K \setminus F$ e $d(x, F) \geq \delta$, o que implica que $x \notin \mathcal{V}_\delta(F) \supset \mathcal{V}_r(F)$, donde $K \notin B_r^H(F)$;

(ii) se K está contido propriamente em F, então existe $x_i \in F \setminus K$ e $d(x_i, K) \geq \varepsilon$, o que implica que $F \nsubseteq \mathcal{V}_\varepsilon(K) \supset \mathcal{V}_r(K)$, donde $K \notin B_r^H(F)$.

Segue de (i) e (ii) que: se $K \in B_r^H(F)$, então $K = F$, ou seja, $B_r^H(F) = \{F\}$. Vamos provar que $\{F\} \notin \tau_{\mathfrak{q}}$. Suponha o contrário; então, como $\{F\}$ é um conjunto unitário, deve existir um aberto $A \subset X$ tal que $\{F\} = \mathfrak{p}(A) \cap \mathfrak{q}(X)$, o que é impossível, já que F tem mais que um ponto.

Suponha agora $F = X$ (consequentemente, X finito). Considere o real positivo $r = \min\{d(x_i, x_j) : x_i, x_j \in X,\ x_i \neq x_j\}$. Dado um subconjunto próprio $K \subset X$,

para cada $x_i \in X \setminus K$ temos $d(x_i, K) \geq r$, o que implica que $X \notin \mathcal{V}_r(K)$, donde $K \notin B_r^H(X)$. Segue que $B_r^H(X) = \{X\}$, ao passo que, pela razão de antes, $\{X\} \notin \tau_{\mathfrak{q}}$.

• X NÃO DISCRETO: provaremos que *existem um real* $r > 0$ *e um subconjunto* $F \subset X$, *com dois ou três pontos, tais que* $B_r^H(F) \notin \tau_{\mathfrak{q}}$.

Consideramos primeiro o caso em que X possui um triângulo escaleno e, em seguida, o caso em que todos os triângulos em X são isósceles. A prova é similar em ambos os casos.

Suponhamos que exista um triângulo escaleno $\{x_1, x_2, x_3\}$ em X e sejam $r_1 = d(x_1, x_2)$, $r_2 = d(x_2, x_3)$ e $r_3 = d(x_1, x_3)$. Sem perda de generalidade, podemos assumir $r_1 < r_2 < r_3$. Ponhamos $F = \{x_1, x_2\}$ e $G = \{x_2, x_3\}$. Seja $\delta > 0$ um número real tal que $r_2 + 2\delta < r_3$, e sejam $r = r_2 + 2\delta$ e $\varepsilon = r_2 + \delta$. Então $r_2 < \varepsilon < r < r_3$. Segue que $G \subset \mathcal{V}_\varepsilon(F)$ e $F \subset \mathcal{V}_\varepsilon(G)$, o que mostra que $G \in B_r^H(F)$. Vamos provar que $B_r^H(F) \notin \tau_{\mathfrak{q}}$. Suponhamos o contrário; então existe um aberto $A \subset X$ tal que $G \in \mathfrak{p}(A) \cap \mathfrak{q}(X) \subset B_r^H(F)$. Logo, $G \subset A$ e, assim, $\{\{x_3\}\} \in \mathfrak{p}(A) \cap \mathfrak{q}(X) \subset B_r^H(F)$, o que implica que $\{x_3\} \in B_r^H(F)$. Por outro lado, como $x_1 \in F$ e $r < d(x_1, x_3)$, temos que $F \notin \mathcal{V}_r(\{x_3\})$. Isto perfaz uma contradição. Portanto, $B_r^H(F) \notin \tau_{\mathfrak{q}}$.

Suponhamos agora que todos os triângulo em X sejam isósceles. Pelo Lema 6.3, existe um subespaço $X' = \{x_1, x_2, x_3, x_4\}$ de X tal que $0 < d(x_1, x_4) < d(x_1, x_3) = d(x_3, x_4) < d(x_1, x_2) = d(x_2, x_3) = d(x_2, x_4)$. Ponhamos $F = \{x_2, x_3, x_4\}$ e $G = \{x_1, x_2, x_4\}$. Sejam ε e r número reais tais que $d(x_1, x_4) < \varepsilon < r < d(x_1, x_3)$. Segue que $G \subset \mathcal{V}_\varepsilon(F)$ e $F \subset \mathcal{V}_\varepsilon(G)$, o que mostra que $G \in B_r^H(F)$. Como antes, vamos provar que $B_r^H(F) \notin \tau_{\mathfrak{q}}$. Suponhamos o contrário; então existe um aberto $A \subset X$ tal que $G \in \mathfrak{p}(A) \cap \mathfrak{q}(X) \subset B_r^H(F)$. Logo $G \subset A$ e, assim, $\{\{x_2\}\} \in \mathfrak{p}(A) \cap \mathfrak{q}(X) \subset B_r^H(F)$, o que implica que $\{x_2\} \in B_r^H(F)$. Por outro lado, como $x_3 \in F$ e $r < d(x_2, x_3)$, temos que $F \notin \mathcal{V}_r(\{x_2\})$. Isto perfaz uma contradição. Portanto, $B_r^H(F) \notin \tau_{\mathfrak{q}}$. □

À primeira vista, a possibilidade de que todos os triângulos em um espaço métrico não discreto possam ser isósceles pode parecer estranha. Para evitar que este estranhamento persista (e provocar no leitor ou na leitora questionamentos para além da existência), apresentamos um exemplo: consideramos um conjunto infinito enumerável $X = \{x_0, x_1, x_2, \ldots\}$ e definimos uma métrica d sobre X estabelecendo que $d(x_m, x_n) = d(x_n, x_m)$ para quaisquer $m, n \geq 0$ e, para $m \leq n$,

$$d(x_m, x_n) = \begin{cases} 0 & \text{if} \quad m = n \\ 1/n & \text{se} \quad 0 = m < n \\ 1/m & \text{se} \quad 0 < m < n \end{cases} .$$

Deixamos ao leitor ou à leitora o trabalho de verificar que d é de fato uma métrica sobre X, que o espaço métrico (X, d) tem x_0 como (seu único) ponto de acumulação e que todos os triângulos neste espaço são isósceles.

Subespaços de espaços euclidianos nos quais todos os triângulos são isósceles são necessariamente finitos; eles são chamados *conjuntos isósceles*. Vários artigos sobre tais conjuntos foram publicados na primeira década do século corrente. Indicamos as referências [4, 5, 10] para quem se interessar pelo tema.

Referências

[1] Elon Lages Lima, *Espaços Métricos*, Rio de Janeiro: IMPA, CNPq, 1977.

[2] Ernest Michael, *Topologies on spaces of subsets*, Transactions of the American Mathematical Society, v. 71, (1951), 152–182.

[3] Gerald Beer, *Topologies on closed and closed convex sets*, Mathematics and its applications 266, Kluwer Academic Publishers, 1993.

[4] Hiroaki Kido, *Classification of isosceles eight-point sets in three-dimensional euclidean space*, European Journal of Combinatorics, v. 27, no. 3, (2006), 329–341.

[5] Hiroaki Kido, *On isosceles sets in the 4-dimensional Euclidean space*, International Journal of Combinatorics, v. 2010, Art. ID 803210, (2010), 30 pp.

[6] James Raymond Munkres, *Topology*, Second Edition, Prentice Hall, Inc., 2000.

[7] Lech Górniewicz, *Topological Fixed Point Theory of Multivalued Mappings*, Second Edition, Springer, Netherlands, 2006.

[8] Marcio Colombo Fenille, *A reverse viewpoint on the upper semicontinuity of multivalued maps*, Mathematica Bohemica, v. 138, no. 4, (2013), 415–423.

[9] Roland Edwin Larson and Susan Jane Andima, *The Lattice of Topologies: A Survey*, Rocky Mountain Journal of Mathematics, v. 5, No. 2, 1975, 177–198.

[10] Yury Ionin, *Isosceles Sets*, The Electronic Journal of Combinatorics, v 16, no. 1, (2009), Research Paper 141, 24 pp.

Universidade Federal de Uberlândia – Faculdade de Matemática
Av. João Naves de Ávila, 2121, Santa Mônica.
38400-902 – Uberlândia MG, Brasil.
Email address: mcfenille@gmail.com

RMU
REVISTA MATEMÁTICA UNIVERSITÁRIA

SBM
SOCIEDADE BRASILEIRA DE MATEMÁTICA

Revista Matemática Universitária,
ISSN: 2675-5254 — **DOI:** 10.21711/26755254/rmu20227

O TEOREMA ESPECTRAL E SUAS MISTERIOSAS RELAÇÕES COM O TEOREMA DA DECOMPOSIÇÃO PRIMÁRIA

RONALDO FREIRE DE LIMA

Resumo. Neste artigo, apresentamos demonstrações alternativas do Teorema Espectral e do Teorema da Decomposição Primária, mostrando que o primeiro decorre diretamente do segundo.

1. Introdução

É notório que uma parte considerável da Álgebra Linear, enquanto teoria, dedica-se ao estudo dos operadores lineares, isto é, das transformações lineares $T : V \to V$, em que V é um espaço vetorial arbitrário. (Aqui, nos restringiremos aos de dimensão finita.) É também estabelecido que a abordagem mais natural aos operadores lineares dá-se por meio de decomposições desses em "suboperadores", o que corresponde a um processo de diagonalização em blocos das matrizes que os representam.

Mais especificamente, a cada operador linear T definido num espaço vetorial real V de dimensão $n \in \mathbb{N}$, como se sabe, está associada uma família de matrizes reais $n \times n$, em que cada uma delas representa T com respeito a alguma base de V. Uma decomposição de T, então, é um processo que visa determinar, nessa família, a matriz mais simples possível, sendo o caso ótimo aquele em que essa matriz é diagonal. Nessa ocorrência, diz-se que o operador T é *diagonalizável*.

A simplicidade das matrizes diagonais, inclusive do ponto de vista operacional, concede aos operadores diagonalizáveis um status de excelência, fato que torna o resultado seguinte um dos mais notáveis da Álgebra Linear.

Teorema Espectral. *Seja V um espaço vetorial real de dimensão finita munido de um produto interno. Então, todo operador linear autoadjunto $T\colon V \to V$ é diagonalizável. Consequentemente, existe uma base ortonormal de V formada por autovetores de T.*

Data de aceitação: 21 de fevereiro de 2022.
Palavras chave. teorema espectral – teorema da decomposição primária – polinômio mínimo.

Relembremos que, dado um espaço vetorial V munido de um produto interno $\langle\ ,\ \rangle$, um operador linear $T\colon V \to V$ é dito *autoadjunto* quando cumpre a condição:

$$\langle Tv, w\rangle = \langle v, Tw\rangle \ \ \forall v, w \in V.$$

Aplicando-se a igualdade acima aos vetores de uma base ortonormal $\mathcal{B}$, de V, conclui-se facilmente que um operador $T : V \to V$ é autoadjunto se, e somente se, a matriz de T com respeito à base $\mathcal{B}$ é simétrica. Dessa forma, os operadores autoadjuntos não só são facilmente identificáveis, mas também existem em abundância. Mais que isso, o conjunto $\mathcal{A}$ formado por todos os operadores autoadjuntos $T : V \to V$ (munido das operações usuais de adição e multiplicação por escalar) constitui um espaço vetorial real, o qual é isomorfo àquele formado pelas matrizes reais simétricas de ordem $n = \dim V$. Esses fatos, dentre outros, atestam a referida notabilidade do Teorema Espectral.

Neste artigo, temos como primeiro propósito o de mostrar que o Teorema Espectral decorre diretamente de um interessante resultado sobre decomposição de operadores, conhecido como Teorema da Decomposição Primária, o qual é raramente considerado em textos introdutórios de Álgebra Linear. Assim, o nosso segundo propósito é o de resgatar o Teorema da Decomposição Primária, apresentando, inclusive, uma demonstração alternativa do mesmo, mais simples que as usuais.

Um dos grandes atributos do Teorema da Decomposição Primária, o qual justifica plenamente o seu resgate, é a aplicabilidade. Com efeito, além do Teorema Espectral, muitos resultados fundamentais em Álgebra Linear podem ser obtidos do Teorema da Decomposição Primária. O mais imediato é a caracterização de operadores lineares diagonalizáveis como aqueles cujos polinômios mínimos são produtos de polinômios mônicos de grau 1. (A prova desse fato está contida no argumento que conclui a nossa demonstração do Teorema Espectral.)

Decorre também do Teorema da Decomposição Primária que as raízes do polinômio mínimo de um operador linear são precisamente seus autovalores, e que, para todo operador linear $T : V \to V$, existe um subespaço de V o qual é invariante por T e tem dimensão 1 ou 2. Por fim, outros importantes teoremas de decomposição podem ser estabelecidos a partir do Teorema da Decomposição Primária, tais como o Teorema da Decomposição Cíclica (cf. [2]) e o célebre Teorema (da forma canônica) de Jordan (cf. [1]).

2. Preliminares

No que se segue, denotaremos por V um espaço vetorial real arbitrário de dimensão $n \in \mathbb{N}$ e munido de um produto interno $\langle\ ,\ \rangle$. O conjunto dos polinômios de variável t e coeficientes no corpo dos reais $\mathbb{R}$ será denotado por $\mathscr{P}[t, \mathbb{R}]$.

Dados um operador linear $T : V \to V$ e um polinômio

$$p(t) := a_k t^k + a_{k-1} t^{k-1} + \cdots + a_1 t + a_0 \in \mathscr{P}[t, \mathbb{R}],$$

define-se o operador $p(T) : V \to V$ por

$$p(T) := a_k T^k + a_{k-1} T^{k-1} + \cdots + a_1 T + a_0 I,$$

em que I denota o operador identidade de V. Quando $p(T)$ é o operador nulo, o qual indicaremos por $\mathbf{0}$, diz-se que p *anula* T ou, equivalentemente, que p é um *anulador* de T.

Por exemplo, o polinômio $p(t) = t^2 + 1$ anula o operador $T : \mathbb{R}^2 \to \mathbb{R}^2$ cuja matriz com respeito à base canônica é:

$$[T] = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

De fato, identificando-se T com sua matriz $[T]$, temos que

$$p([T]) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^2 + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix},$$

donde $p(T) = \mathbf{0}$, isto é, p anula T.

Para todo operador linear $T : V \to V$, existe um polinômio $p \in \mathscr{P}[t, \mathbb{R}]$ que o anula. Para constatarmos isso, lembremos que o espaço vetorial formado pelos operadores lineares de V, o qual denotamos por $\mathscr{L}(V)$, tem dimensão n^2 (por ser isomorfo ao espaço das matrizes $n \times n$). Logo, o conjunto

$$\{I, T, T^2, \ldots, T^{n^2}\},$$

por conter $n^2 + 1$ operadores, é necessariamente LD em $\mathscr{L}(V)$. Assim, existem escalares $a_0, \ldots, a_{n^2} \in \mathbb{R}$, não todos nulos, para os quais a igualdade

$$a_0 I + a_1 T + \cdots + a_{n^2} T^{n^2} = \mathbf{0}$$

se cumpre, donde se conclui que o polinômio

$$p(t) = a_0 + a_1 t + \cdots + a_{n^2} t^{n^2}$$

anula o operador linear $T \in \mathscr{L}(V)$.

Definição 1. O *polinômio mínimo* de um operador linear $T \in \mathscr{L}(V)$ é o polinômio mônico de menor grau que anula T. (Lembre que um polinômio é dito *mônico* quando o coeficiente do seu termo de maior grau é 1.)

A existência de anuladores de um operador $T : V \to V$ assegura a existência do polinômio mínimo de T. Verifiquemos, através do algoritmo da divisão, a sua unicidade. Para tanto, suponhamos que $m_1, m_2 \in \mathscr{P}[t, \mathbb{R}]$ sejam polinômios mínimos de T. Pelo algoritmo da divisão, existem $q, r \in \mathscr{P}[t, \mathbb{R}]$, com $r = 0$ ou $\text{grau}\,(r) < \text{grau}\,(m_1)$, tais que $m_2 = m_1 q + r$, donde $r = m_2 - m_1 q$. Se r fosse não nulo, multiplicando-o por $c = 1/a_k$, em que a_k é o coeficiente de seu termo de maior grau, teríamos que cr seria um polinômio mônico, e também um anulador de T, já que m_1 e m_2 o são. Isso, porém, violaria a minimalidade de m_1. Assim, devemos ter $r = 0$. Daí, e do fato de m_1 e m_2 serem ambos mônicos e de mesmo grau, segue-se que $q = 1$ e, portanto, que $m_1 = m_2$.

De modo inteiramente análogo, isto é, através do algoritmo da divisão, estabelece-se a propriedade fundamental dos polinômios mínimos, a saber: *Para todo operador $T \in \mathscr{L}(\mathbb{R}^n)$, e todo anulador p de T, o polinômio mínimo m de T divide p.* Com efeito, pelo algoritmo da divisão, existem polinômios $q, r \in \mathscr{P}[t, \mathbb{R}]$, com $r = 0$ ou $\text{grau}(r) < \text{grau}(m)$, tais que $p = qm + r$. Argumentando como acima, concluiríamos

que se r fosse não nulo, então cr seria um polinômio mônico e anulador de T, o que contradiria a minimalidade de m.

Dado $v \in V$, diz-se também que $p \in \mathscr{P}[t, \mathbb{R}]$ *anula* T *em* v, quando $p(T)v = 0$. O *polinômio mínimo de* T *em* v é então definido como o polinômio mônico de menor grau que anula T em v. Analogamente, verifica-se que o polinômio mínimo de T em v divide qualquer polinômio que anule T em v.

Denotaremos o núcleo de um operador linear $T : V \to V$ por $\ker T$, e o seu conjunto imagem por $T(V)$, isto é,

$$\ker T = \{v \in V \,;\, Tv = 0\} \quad \text{e} \quad T(V) = \{Tv \,;\, v \in V\}.$$

Verifica-se facilmente que $\ker T$ e $T(V)$ são ambos subespaços vetoriais de V, bem como que um operador linear $T : V \to V$ é injetivo se, e somente se, $\ker T = \{0\}$. Além disso, pelo Teorema do Núcleo e da Imagem,

$$\dim V = \dim \ker T + \dim T(V),$$

de modo que T é um isomorfismo (isto é, é invertível) se, e somente se, é injetivo.

Dado um operador linear $T : V \to V$, diz-se que um subespaço vetorial W de V é T-*invariante*, se $T(W) \subset W$. Note que, quando W é T-invariante, o operador linear $T|_W : W \to W$ fica bem definido.

O *autoespaço* de um operador T associado a um número real λ,

$$W_\lambda := \ker(T - \lambda I),$$

é um importante exemplo de subespaço T-invariante. Observe que todo vetor não nulo $v \in W_\lambda$ é um autovetor de T cujo autovalor associado é λ, isto é, $Tv = \lambda v$.

3. Demonstrações dos Teoremas

Inicialmente, relembremos que, pelo Teorema Fundamental da Álgebra, todo polinômio $p \in \mathscr{P}[t, \mathbb{R}]$ se decompõe de forma única como

$$p(t) = (p_1(t))^{s_1} \dots (p_k(t))^{s_k}, \quad s_1, \dots, s_k \in \mathbb{N},$$

em que $p_1, \dots, p_k$ são polinômios irredutíveis (i.e., que não admitem divisores), de grau 1 ou 2, e distintos entre si. Nesse contexto, os polinômios $p_1, \dots, p_k$ são ditos os *fatores primos* de p.

O Teorema da Decomposição Primária estabelece que, para todo operador linear $T : V \to V$, existe uma decomposição de V em subespaços T-invariantes, os quais são determinados pelos fatores primos do polinômio mínimo de T. (Note que os fatores primos de um polinômio mínimo são todos mônicos, uma vez que m o é.) Mais precisamente:

Teorema da Decomposição Primária. *Considere um operador linear* $T : V \to V$, *bem como a decomposição de seu polinômio mínimo* m *por fatores primos:*

$$m(t) = (p_1(t))^{s_1} \dots (p_k(t))^{s_k}, \quad s_1, \dots, s_k \in \mathbb{N}.$$

Nessas condições, V *admite a seguinte decomposição:*

$$V = W_1 \oplus \dots \oplus W_k \,,$$

em que W_i denota o núcleo de $p_i(T)^{s_i}$. Além disso, para todo índice $i = 1, \ldots, k$, são válidas as seguintes afirmações:

- *W_i é T-invariante;*
- *O polinômio mínimo de $T|_{W_i}$ é $p_i^{s_i}$.*

Vejamos, primeiramente, que o Teorema da Decomposição Primária implica diretamente o Teorema Espectral. Antes, convém observarmos que toda potência de um operador autoadjunto T é também um operador autoadjunto, como pode ser facilmente verificado. Consequentemente, vale o mesmo para $p(T)$, qualquer que seja o polinômio $p \in \mathscr{P}[t, \mathbb{R}]$. Lembramos ainda que autovetores de um operador autoadjunto associados a autovalores distintos são, necessariamente, ortogonais.

No lema seguinte, estabelecemos algumas propriedades adicionais dos operadores autoadjuntos.

Lema 1. *São verdadeiras as seguintes afirmações a respeito de um operador autoadjunto $T : V \to V$:*

i) *Dado $k \in \mathbb{N}$, se $T^k = \mathbf{0}$, então $T = \mathbf{0}$;*
ii) *Se $p(t) = t^2 + bt + c$ é irredutível, então $p(T)$ é invertível.*

Demonstração. (i) O resultado é óbvio para $k = 1$. Suponhamo-lo verdadeiro, então, para $k \in \mathbb{N}$. Assim, supondo-se $T^{k+1} = \mathbf{0}$, para todo $v \in V$, tem-se

$$\langle T^k v, T^k v\rangle = \langle T^{k+1} v, T^{k-1} v\rangle = 0,$$

donde se obtém $T^k v = 0$ e, pela hipótese de indução, que $T = \mathbf{0}$.

(ii) Suponhamos, por absurdo, que exista um vetor $u \in V$, tal que

$$p(T)u = 0, \ \ \|u\| = 1.$$

Nesse caso, uma vez que $|\langle Tu, u\rangle| \leq \|Tu\|$ (pela desigualdade de Cauchy-Schwarz) e $\langle T^2 u, u\rangle = \|Tu\|^2$ (por T ser autoajunto), para $b \geq 0$, tem-se

$$0 = \langle p(T)u, u\rangle = \|Tu\|^2 + b\langle Tu, u\rangle + c \geq \|Tu\|^2 - b\|Tu\| + c = p(-\|Tu\|).$$

Analogamente, se $b < 0$, tem-se $p(\|Tu\|) \leq 0$. Porém, sendo p um polinômio mônico quadrático e irredutível, devemos ter $p(t) > 0\, \forall t \in \mathbb{R}$. Dessa contradição, segue-se que $\ker p(T) = \{0\}$, donde se infere que $p(T)$ é um operador injetivo e, portanto, invertível. □

Demonstração do Teorema Espectral. Consideremos a decomposição do polinômio mínimo m de T por fatores primos:

$$m(t) = (p_1(t))^{s_1}(p_2(t))^{s_2} \ldots (p_k(t))^{s_k}\,, \ \ s_i \in \mathbb{N}, \ \ i = 1, \ldots k,$$

lembrando que cada p_i é um polinômio mônico irredutível de grau 1 ou 2. Designando-se por W_i o núcleo de $(p_i(T))^{s_i}$, segue-se do Teorema da Decomposição Primária que W_i é T-invariante e que vale a igualdade

$$W = W_1 \oplus \cdots \oplus W_k.$$

Além disso, $m_i := (p_i)^{s_i}$ é o polinômio mínimo do operador (autoadjunto)

$$T_i = T|_{W_i} : W_i \to W_i.$$

Em particular, $p_i(T_i)$ é autoadjunto e

$$m_i(T_i) = (p_i(T_i))^{s_i} = \mathbf{0} \ \ \forall i = 1, \dots k.$$

Segue-se dessa última igualdade e do Lema 1-(i) que $p_i(T_i) = \mathbf{0}$, donde se infere que p_i é o polinômio mínimo de T_i, isto é, $s_i = 1$. Além disso, pelo Lema 1-(ii), p_i tem grau 1, uma vez que $p_i(T_i)$ não é invertível. Dessa forma, o polinômio mínimo de T se escreve como

$$m(t) = (t - \lambda_1) \dots (t - \lambda_k), \ \ \lambda_i \in \mathbb{R}, \ i = 1, \dots, k,$$

donde $W_i = \ker(T - \lambda_i I) = W_{\lambda_i}$ e, portanto,

$$V = W_{\lambda_1} \oplus \dots \oplus W_{\lambda_k}.$$

Conclui-se facilmente dessa última igualdade e da T-invariância dos autoespaços W_{λ_i} que o operador T é diagonalizável, bem como que $\lambda_1, \dots, \lambda_k$ são seus autovalores. Além disso, uma vez que autovetores de T em diferentes autoespaços são ortogonais, tomando-se uma base ortonormal em cada um dos autoespaços W_{λ_i}, obtém-se uma base ortonormal de V formada por autovetores de T. □

Passemos à demonstração do Teorema da Decomposição Primária, na qual faremos uso do seguinte

Lema 2. *Para quaisquer operadores lineares $T_1, \dots, T_k \colon V \to V$, tem-se*

$$\dim \ker(T_1 T_2 \dots T_k) \leq \dim \ker T_1 + \dim \ker T_2 + \dots + \dim \ker T_k.$$

Demonstração. Pelo Princípio da Indução, basta considerarmos o caso $k = 2$. Fazendo-se $K := \ker(T_1 T_2)$, é imediato que $T_2(K) \subset \ker T_1$. Daí, e do Teorema do Núcleo e da Imagem, tem-se

$$\dim K = \dim T_2(K) + \dim \ker T_2|_K \leq \dim \ker T_1 + \dim \ker T_2,$$

como queríamos demonstrar. □

Demonstração do Teorema da Decomposição Primária. Vejamos, inicialmente, que

$$W_i := \ker p^{s_i}(T) \neq \{0\} \ \ \forall i = 1, \dots, k.$$

Supondo, por absurdo, que $W_i = \{0\}$ para algum i, temos que $p_i(T)^{s_i} : V \to V$ é um isomorfismo. Claramente, os operadores $p_j(T)^{s_j}$ comutam, donde podemos supor, sem perda de generalidade, que $i = k$. Agora, existe $w \in V$, tal que $p_1(T)^{s_1} \dots p_{k-1}(T)^{s_{k-1}} w \neq 0$. Caso contrário, $p_1^{s_1} \dots p_{k-1}^{s_{k-1}}$ seria o polinômio mínimo de T. Porém, sendo $p_k(T)^{s_k}$ um isomorfismo, existe $v \in V$, tal que $p_k(T)^{s_k} v = w$. Logo,

$$m(T)v = p_1(T)^{s_1} \dots p_{k-1}(T)^{s_{k-1}} p_k(T)^{s_k} v = p_1(T)^{s_1} \dots p_{k-1}(T)^{s_{k-1}} w \neq 0,$$

o que, obviamente, é uma contradição. Assim, $W_i \neq \{0\} \, \forall i = 1, \dots, k$.

Suponhamos, agora, que exista um vetor não nulo $w \in W_i \cap W_j$, $i \neq j$. Nesse caso, $p_i^{s_i}$ e $p_j^{s_j}$ anulam T em w e, portanto, o polinômio mínimo de T em w divide ambos, o que é absurdo, visto que $p_i^{s_i}$ e $p_j^{s_j}$ são primos entre si. Logo,

$$W_i \cap W_j = \{0\} \ \forall i \neq j \in \{1, \dots, k\}.$$

Além disso, temos

$$V = \ker m(T) = \ker(p_1(T)^{s_1} \dots p_k(T)^{s_k}),$$

o que, juntamente com o Lema 2, nos dá

$$\dim V = \dim \ker m(T) \leq \dim W_1 + \cdots + \dim W_k = \dim (W_1 \oplus \cdots \oplus W_k),$$

donde $V = W_1 \oplus \cdots \oplus W_k$.

Segue-se facilmente da comutatividade de T com $p_i(T)^{s_i}$ que cada subespaço W_i é T-invariante. De fato, dado $v \in W_i$, tem-se

$$p_i(T)^{s_i} Tv = Tp_i(T)^{s_i} v = 0,$$

o que implica $Tv \in W_i$.

Provemos, finalmente, que o polinômio mínimo de $T|_{W_i}$, o qual denotaremos por m_i, é $p_i^{s_i}$. Supondo-se, sem perda de generalidade, que $i = 1$, definamos o polinômio

$$p = m_1 p_2^{s_2} \dots p_k^{s_k}$$

e vejamos que o mesmo anula T. De fato, tomando-se $v = w_1 + \cdots + w_k \in V$, $w_i \in W_i$, e lembrando-se que $p_j^{s_j}(T) w_j = 0$, tem-se

$$\begin{aligned} p(T)v &= m_1(T)p_2^{s_2}(T) \dots p_k^{s_k}(T)(w_1 + \cdots + w_k) \\ &= m_1(T)p_2^{s_2}(T) \dots p_k^{s_k}(T)(w_1 + \cdots + w_{k-1}) + \\ &\quad m_1(T)p_2^{s_2}(T) \dots p_k^{s_k}(T) w_k \\ &= m_1(T)p_2^{s_2}(T) \dots p_k^{s_k}(T)(w_1 + \cdots + w_{k-1}). \end{aligned}$$

Repentindo-se esse procedimento $k-1$ vezes, chega-se facilmente a:

$$\begin{aligned} p(T)v &= m_1(T)p_2^{s_2}(T) \dots p_k^{s_k}(T) w_1 = p_2^{s_2}(T) \dots p_k^{s_k}(T) m_1(T) w_1 \\ &= p_2^{s_2}(T) \dots p_k^{s_k}(T) m_1(T|_{W_1}) w_1 = 0. \end{aligned}$$

Desta forma, p anula T e, portanto, m divide p. Em particular, $p_1^{s_1}$ divide m_1. Porém, uma vez que W_1 é o núcleo de $p_1^{s_1}(T)$, temos que

$$p_1^{s_1}(T|_{W_1}) = p_1^{s_1}(T)|_{W_1} = \mathbf{0},$$

isto é, $p_1^{s_1}$ anula $T|_{W_1}$. Logo, pela minimalidade de m_1, devemos ter $p_1^{s_1} = m_1$. □

Referências

[1] K. Hoffman, R. Kunze: Linear Algebra, Prentice-Hall (1971).

[2] H. G. Jacob: Another proof of the rational decomposition theorem. Amer. Math. Monthly. 1131–1134 (1973).

UFRN-CCET-Departamento de Matemática
Lagoa Nova, Natal, RN
59072-970
Email address: ronaldo.freire@ufrn.br, zabumbe@gmail.com