

Um passeio  
matemático  
singular

Étienne Ghys

Um passeio matemático singular

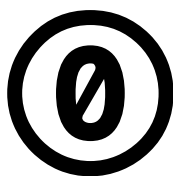


Caspar David Friedrich  
"Caminhante sobre o  
mar de neblina" (1818) †

ÉTIENNE GHYS

UM PASSEIO  
MATEMÁTICO  
SINGULAR

## Um passeio matemático singular



O texto e as ilustrações sem o símbolo  $\dagger$  foram produzidos por Étienne Ghys, que renunciou a todos os direitos de autor e direitos conexos. Étienne Ghys dedicou o trabalho ao domínio público, renunciando a todos os seus direitos sob as leis de direito de autor e/ou de direitos conexos referentes ao trabalho, em todo o mundo, na medida permitida por lei. Você pode copiar, modificar, distribuir e executar o trabalho, mesmo para fins comerciais, tudo sem pedir autorização. [Image Credits page](#).

Veja outras Informações CC0 [Creative Commons](#).

Este livro pdf está disponível gratuitamente na página [do autor](#).

Endereço:

Étienne Ghys  
CNRS-ENS Lyon, UMPA  
46 Allée d'Italie  
69364 Lyon  
France  
[etienne.ghys@ens-lyon.fr](mailto:etienne.ghys@ens-lyon.fr)  
[perso.ens-lyon.fr/ghys](http://perso.ens-lyon.fr/ghys)

*Pour Martine, qui m'a toujours  
encouragé à écrire "un petit livre".*



## *Um passeio tropical?*

Este livro deve muito ao Brasil. Minha primeira viagem ao Rio de Janeiro foi há quarenta anos, durante um estágio de pós-doutorado. Você pode imaginar o choque que isso representou para um jovem francês dos pontos de vista cultural, linguístico, climático, mas também... matemático.

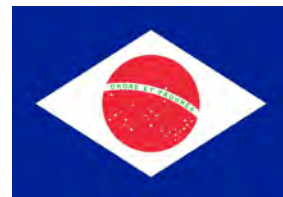
Estava acostumado com a matemática à francesa, imbuído da abordagem Bourbakista, quando descobri a matemática praticada nos trópicos.

Percebi que a teoria dos sistemas dinâmicos poderia ser qualitativa, que seria possível raciocinarmos do particular para o geral, que exemplos concretos frequentemente precedem os teoremas abstratos. Compreendi especialmente que poderíamos abordar a pesquisa matemática com alegria e descontração: duas características brasileiras, presentes, assim espero, neste passeio-texto singular.

Os alunos que assistiram meu primeiro curso eram apenas um pouco mais novos do que eu e conheciam topologia diferencial quase igual a mim. Paulo Ney de Souza participou desse curso e foi o primeiro a escrever sua dissertação de mestrado em topologia das folheações sob minha orientação. Quarenta anos depois, ele mora em Berkeley, mas retorna frequentemente ao seu país de origem, onde é responsável pelas publicações do IMPA (Instituto de Matemática Pura e Aplicada). Tornou-se um dos melhores  $\text{\TeX}$ perts do mundo e escreveu o maravilhoso livro “Berkeley Problems in Mathematics”, em parceria com Jorge-Nuno Silva.

Ele gentilmente se ofereceu para finalizar a versão brasileira deste livro. Eu sou infinitamente grato ao Paulo.

Nestes quarenta anos, tive muitas oportunidades de visitar o Brasil. Quando vim em 2017, lecionei um curso de verão no IMPA enquanto escrevia este livro. Os alunos eram de toda a América do Sul, e o ambiente excelente, uma mistura de seriedade e relaxamento (lembro-me, por exemplo, que os alunos tinham algumas dificuldades para chegar no horário).



Bandeira franco-brasileira



Paulo Ney de Souza



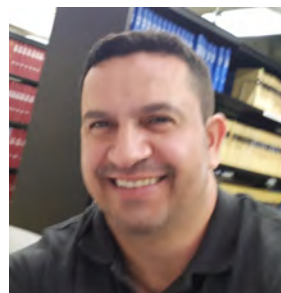
Para minha alegria, durante o curso, em relação ao livro, fui surpreendido duas vezes. Primeiro, Filipe Bellio da Nóbrega e Gregory Cosac Daher iniciaram a tradução de alguns dos capítulos. Agradeço muito a ambos. Segundo, como as aulas foram filmadas (fato habitual no IMPA), recebi um e-mail de um estudante de doutorado brasileiro, o Carlos Alexandre Gomes da Silva: ele participava on-line do curso enquanto finalizava um período de pesquisa na Universidade de Lyon (cidade onde eu trabalho!) e propôs traduzir o livro para o português. Hoje, Carlos é professor da Universidade Federal do Rio Grande do Norte (UFRN). Agradeço-lhe sinceramente pelo tremendo trabalho feito com tanto entusiasmo.

Finalmente, Daniel Weller leu o texto e fez sugestões que melhoraram muito o estilo e a qualidade da língua. Quero agradecer a ele também.

Sempre achei que a linguagem em que expressamos a matemática é parte integrante da própria matemática. Por isso, estou particularmente feliz com esta bela tradução para o português, uma forma de agradecer ao Brasil por tudo o que devo aos brasileiros e a este país tropical, abençoado e bonito por natureza.

Étienne Ghys

Paris, 7 de Abril 2020



Carlos Alexandre Gomes da Silva

# Conteúdo

<i>Prefácio</i>	1
<i>Roteiro</i>	5
<i>Polinômios que se intersectam</i>	11
<i>Padrões e permutações: Donald Knuth</i>	17
<i>Permutações separáveis</i>	29
<i>Hiparco e Schroeder</i>	39
<i>De methodis serierum et fluxionum: Método de Newton</i>	47
<i>De methodis serierum et fluxionum: Séries de Newton</i>	55
<i>Um pouco de álgebra formal</i>	65
<i>Gauss e as curvas algébricas</i>	75
<i>Prova da afirmação de Gauss sobre singularidades</i>	87
<i>De seriebus divergentibus: Euler, Cauchy e Poincaré</i>	97
<i>Convergência: Cauchy</i>	107

<i>Möbius e a sua faixa</i>	125
<i>Colares de Möbius</i>	137
<i>Resolução de singularidades</i>	151
<i>A esfera de dimensão 3 e a fibração de Hopf</i>	163
<i>A cúspide e o trevo</i>	179
<i>Victor Puiseux, finalmente!</i>	193
<i>Jack Milnor e a sua fibração</i>	203
<i>O associaedro de Hiparco–Schroeder–Tamari–Stasheff</i>	219
<i>Jim Stasheff e os espaços de laços</i>	229
<i>Operads</i>	241
<i>Operads singulares</i>	253
<i>Gauss está de volta: curvas no plano</i>	263
<i>Diagramas de cordas analíticos: um algoritmo</i>	279
<i>Diagramas de cordas analíticos: grafos de entrelaçamento</i>	293
<i>Gauss e o número de entrelace</i>	313
<i>Kontsevich está de volta: um invariante universal</i>	323
<i>Posfácio</i>	331
<i>Agradecimentos</i>	333
<i>Crédito das imagens</i>	335

# Prefácio

EM MARÇO DE 2009, estava numa reunião administrativa ao lado do meu colega, o Maxim Kontsevich, ainda mais entediado do que eu. Obviamente, ele tinha outra coisa em mente. De repente, ele me passou um bilhete de metrô parisiense contendo um rabisco e uma única palavra: “impossível”. Esse era um novo teorema que ele queria compartilhar comigo!

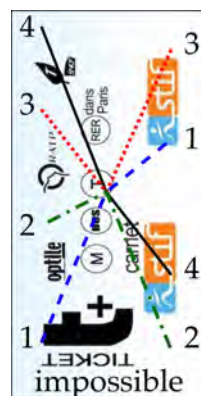
Um breve tempo passou e alguns sussurros entre nós antes que eu pudesse adivinhar a afirmação do teorema e mais alguns minutos para encontrar a prova. Aqui está o enunciado.

**Teorema.** Quatro polinômios  $P_1, P_2, P_3, P_4$  de variável real  $x$  não podem satisfazer

- $P_1(x) < P_2(x) < P_3(x) < P_4(x)$  para pequenos valores de  $x < 0$ ,
- $P_2(x) < P_4(x) < P_1(x) < P_3(x)$  para pequenos valores de  $x > 0$ .

A posição relativa dos gráficos de quatro polinômios reais está sujeita a algumas restrições. Eu fiquei fascinado: um novo resultado elementar sobre quatro polinômios em 2009!

Mais tarde, eu tentei colocar isso num contexto mais geral, para estudar a situação com mais de quatro polinômios. O resultado foi uma prazerosa jornada pela matemática, com bastante desvios, em diversas áreas surpreendentemente diferentes, em muitos períodos da história. Como de costume, isso levou a problemas que eu só consegui resolver parcialmente. Compartilho essa viagem neste pequeno livro que tem o propósito de convidar o leitor para esse passeio matemático. Não escolhi a maneira mais eficiente de atingir uma meta específica porque, na verdade, não há uma meta para esse texto. Quase todos os capítulos são basicamente independentes entre si, e você pode pular quantos



Maxim Kontsevich .

deles quiser. Se você achar uma seção muito árdua ou muito simples, você pode facilmente ignorá-la. Faremos uma visita a Hiparco, Newton e Gauss, mas também a muitos matemáticos contemporâneos. Brincaremos com um pouco de álgebra, topologia, geometria, análise complexa, combinatória e ciência da computação. Um verdadeiro passeio no mundo matemático.

No entanto, para alcançar algum objetivo e não transformar esse passeio em um percurso completamente aleatório, deixe-me citar um resultado que será provado em um dos últimos capítulos. Este é provavelmente o único novo resultado neste trabalho.

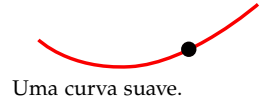
Vamos considerar um ponto  $p$  em uma curva plana  $C$ .

Se  $C$  é suave, a sua representação gráfica local não é tão interessante.

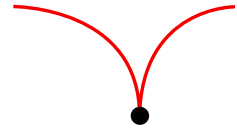
Se  $C$  é singular em  $p$ , a sua representação gráfica pode ser bem mais complicada, como, por exemplo, a origem na cúspide  $x^2 = y^3$ . Vamos restringir nosso estudo a curvas algébricas definidas por alguma equação  $F(x, y) = 0$  em que  $F$  é um polinômio em  $x, y$  com coeficientes reais. Acontece que, na vizinhança de um de seus pontos, tal curva é a união de um número finito de pedaços irredutíveis, geralmente chamados de ramos. A natureza desses ramos foi um tema de muitos debates no passado, e nós discutiremos esse tópico em detalhe. O principal resultado é que topologicamente os ramos são suaves! Mais precisamente, para cada um dos ramos, existe um homeomorfismo local do plano que o envia para uma linha reta. De forma que cada ramo intersecta um pequeno círculo centrado em  $p$  em exatamente dois pontos.

A posição relativa de muitos ramos de uma curva é muito mais sutil. Na vizinhança de um ponto singular, a topologia é descrita por um número par de pontos em um círculo, pareados dois a dois: o pareamento é dado pelos ramos. Assim, obtemos  $2n$  pontos em um círculo agrupados em  $n$  pares, cada um deles com uma cor ou uma letra.

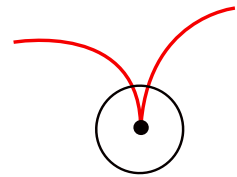
Agora eu posso enunciar um teorema que será mais ou menos o nosso destino final, algum tipo de farol mostrando alguma direção.



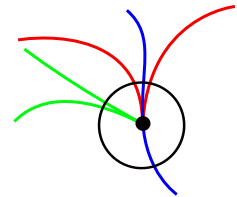
Uma curva suave.



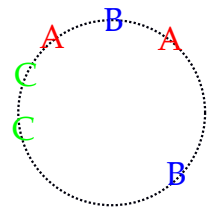
Um ponto cuspidal.



Um pequeno círculo intersecta a curva em dois pontos.



Uma curva com três ramos.



A palavra cíclica associada ABACCB.

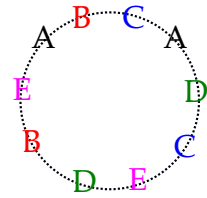
**Teorema.** *Não há singularidade de uma curva algébrica real no plano que consista em cinco ramos A, B, C, D, E cruzando um pequeno círculo como na imagem ao lado.*

Realmente, eu irei provar um teorema muito mais preciso, dando uma descrição completa de todas as possíveis configurações topológicas dos ramos de uma curva analítica.

Eu escrevi esse “petit livre” para um leitor específico na minha mente: eu mesmo, quando eu era um estudante de graduação... Para ser mais específico, eu limitei os pré-requisitos para meu conhecimento quando eu passei no exame da “agrégation”, exatamente há quarenta anos ☺! Eu lembro perfeitamente que eu tinha (e ainda tenho) grandes dificuldades de ler longos tratados matemáticos, cheios de detalhes técnicos, porque eu sempre preferi olhar para as figuras. Ao longo desses anos, eu aprendi que a precisão e os detalhes são frequentemente necessários em matemática, mas eu ainda gosto muito de passeios. Eu tentei imaginar quais teriam sido minhas próprias reações diante deste livro, como se eu fosse um iniciante. Essa “conversa” entre duas “versões de mim mesmo” tem sido interessante e lembrou-me do conto “El Otro” por Jorge Luis Borges. Foi um sonho? Uma reconstrução do passado?

Uma palavra de cautela: este não é um livro escrito na tradicional estrutura definição-teorema-prova. Você tem que estar preparado para se perder de vez em quando, como em muitos passeios. Eu sei que você vai reclamar por causa da falta de definições precisas, mas, de fato, você terá que aceitar definições incompletas... Claro, livros didáticos são necessários e eu forneço muitas referências nas margens. No entanto, estou convencido de que ideias e exemplos matemáticos precedem a provas e a definições formais. Como d’Alembert disse uma vez: “*Apenas vá em frente... e a fé irá alcançá-lo!*”. Você pode ver de vez em quando um belo panorama emergindo da neblina, como aquele no frontispício deste ensaio, de Caspar David Friedrich

Espero que alguns graduandos motivados desfrutem de alguns desses panoramas. Agora podemos embarcar para a nossa viagem.



Cinco ramos impossíveis.



“*Allez en avant, et la foi vous viendra.*”

Der Wanderer über dem Nebelmeer (1818): uma sugestão do mundo matemático?



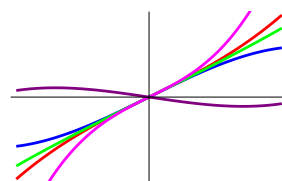
# Roteiro

DEFINITIVAMENTE NÃO SEGUIREMOS UMA ROTA DIRETA.

Como você já deve estar preparado para alguns desvios opcionais, apresento um esboço provavelmente útil do nosso itinerário. Como se fosse uma apresentação promocional de um pacote turístico, ofertado por um funcionário de uma agência de turismo.

Os primeiros quatro capítulos tratam das *posições relativas* entre os gráficos de uma família de polinômios reais  $P_1, \dots, P_n$ , no espírito do **teorema de Kontsevich** que mencionei no prefácio. Comparando os valores de  $P_i(x)$  para pequenos valores positivos e negativos de  $x$  produz uma permutação de  $\{1, \dots, n\}$ , que descreve a figura local numa vizinhança do 0. Eu darei uma caracterização bastante precisa dessas permutações, que já foram consideradas em um disfarce diferente pelos especialistas da área da combinatória, sob o nome “*permutações separáveis*”. Em seguida, examinaremos *push e pop stacks* na forma apresentada por **Donald Knuth** em “A arte da programação de computadores”. Também *determinaremos* o número de permutações **separáveis**, oportunidade para descobrir que esses números já foram considerados por **Hiparco** há mais de dois milênios.

Em seguida, tentaremos generalizar o problema dos gráficos de polinômios para *curvas planas* definidas implicitamente por alguma equação polinomial real  $F(x, y) = 0$ , que requer a compreensão da topologia de uma curva algébrica (ou analítica) na vizinhança de um ponto singular. Os primeiros resultados importantes relacionados a esse problema são devidos à Newton em 1669, em um artigo extraordinário intitulado ***Tractatus de methodis serierum et fluxionum***, que estudaremos em dois capí-



Uma permutação definida por 5 polinômios.

specimen aliqua proponitur, para

$x^4$	$x^3y$	$x^2y^2$	$xy^3$	$y^4$
$x^3$	$x^2y$	$xy^2$	$x^2y^3$	$x^3y^4$
$x^2$	$x^2y$	$x^2y^2$	$xy^3$	$x^2y^4$
$x$	$xy$	$xy^2$	$xy^3$	$xy^4$
0	y	y <sup>2</sup>	y <sup>3</sup>	y <sup>4</sup>

A

correspondencia inordinis nota al  
fonti plura in usquibus para

De Newton *De methodis*. +



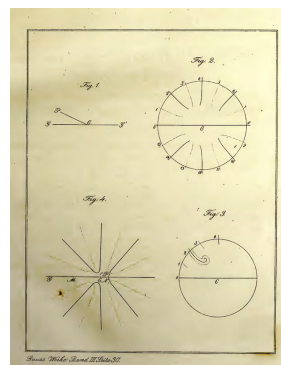
tulos. Nesse artigo, há uma apresentação detalhada do famoso *método de Newton* para encontrar aproximações das raízes de polinômios, introduzindo a ideia dos *polígonos de Newton*. Estrictamente falando, Newton não forneceu provas, mas entendeu que, localmente, uma curva analítica consiste em um número finito de ramos, que são “gráficos” de *séries de potências formais com expoentes racionais*. O capítulo adicional — **álgebra formal** — explicará os resultados de Newton na terminologia moderna e oferecerá provas.

Até aqui, até esse ponto do passeio, a discussão será puramente algébrica. Então, revisaremos a **primeira prova de Gauss do teorema fundamental da álgebra** — sua tese de doutorado em 1799 — usando argumentos de natureza topológica, revolucionários para a época, baseados na afirmação não comprovada de que *uma curva algébrica que entra em um disco deve sair*. A **prova desta afirmação** é mais sutil do que se pode imaginar. No século XIX, dois matemáticos que compartilham o mesmo nome não puderam provar.

**Euler, Cauchy e Poincaré** foram grandes mestres na manipulação de séries. Dois capítulos tratam sobre suas descobertas, usando o **Cálculo dos limites de Cauchy** de forma a obtermos a prova da convergência *da série de Newton*, que nos permite mostrar que um pequeno círculo em torno de uma singularidade de uma curva analítica real do plano cruza a curva em um número par de pontos, definindo um *diagrama de cordas*, ou seja,  $2n$  pontos ordenados ciclicamente e agrupados em pares.

Os três capítulos seguintes tratam da topologia das singularidades das curvas analíticas planas. Nós explicaremos o método *blow-up*, um tipo de microscópio que nos permite examinar profundamente a singularidade. Topologicamente, isso introduz uma **faixa de Möbius**, ou **colares de Möbius** se o microscópio for usado várias vezes. Como veremos, o blow-up será uma ferramenta poderosa na **resolução de singularidades**.

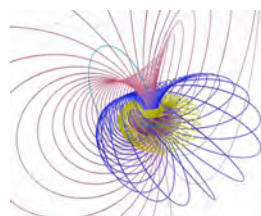
As imagens locais das curvas planas *complexas* são lindas e merecem uma visita. Como  $\mathbb{C}^2$  têm dimensão real 4, nós intersectaremos a curva com pequenas esferas tridimensionais em torno da singularidade. Desse ponto de vista, até mesmo linhas retas



Da Tese de doutorado de Gauss



Um disco sofrendo três blow-ups.



As fibras de Hopf.

produzem objetos notáveis, como a **Fibração de Hopf**.

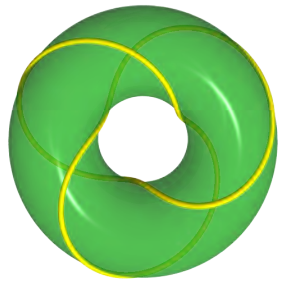
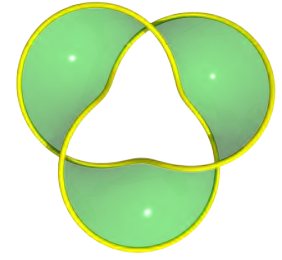
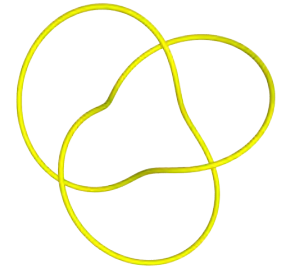
As singularidades mais complexas, por exemplo, o **cúspide**  $x^3 - y^2 = 0$  são descritas por *nós* e *links*. Para entender o caso geral, nós visitaremos **Victor Puiseux**, que propôs em 1850 uma abordagem completamente nova para a série de Newton. Em 1968, **Jack Milnor** usou essas ideias para resultar numa figura topológica completa, mas ainda considerando os *números complexos*.

Curiosamente, descobriremos uma conexão entre permutações separáveis e o **associaedro**, família de polítopos convexos introduzidos por Tamari e Stasheff para entender o significado de “associatividade a menos de homotopia”. Usando seus polítopos, **Jim Stasheff** foi capaz de fornecer uma caracterização dos espaços com o mesmo tipo de homotopia que os grupos topológicos. Descobriu-se que esse era o ponto de partida da **teoria dos operads**, que desempenha um papel fundamental na teoria da homotopia moderna e na topologia algébrica. Ressalto que os Operads são estruturas algébricas muito gerais e estão perfeitamente adaptadas à nossa situação, sendo exemplos típicos as árvores, tranças, espaços de configuração etc. Veremos que a coleção de todas as singularidades, à exceção dos homeomorfismos, pode ser vista como um **operad singular**, ajudando a entender o quadro global das ideias.

Apenas por diversão, examinaremos uma pequena **nota de Gauss**, referente a curvas fechadas no plano com pontos duplos simples. Indo ao redor da curva, cada ponto duplo é visitado duas vezes, de maneira que se defina um diagrama de cordas. Como podemos caracterizar este tipo de diagrama?

Ao final, alcançaremos nosso objetivo: a caracterização completa, em dois capítulos, dos **diagramas de cordas associados a singularidades** de curvas planas analíticas reais.

Dois capítulos adicionais concluem nosso passeio. Um sobre a abordagem dos números de entrelace de **Gauss**, e o último, sem provas, sobre o **invariante universal de Kontsevich para nós**, encorajando o leitor a continuar a exploração.



O nó de trevo. +

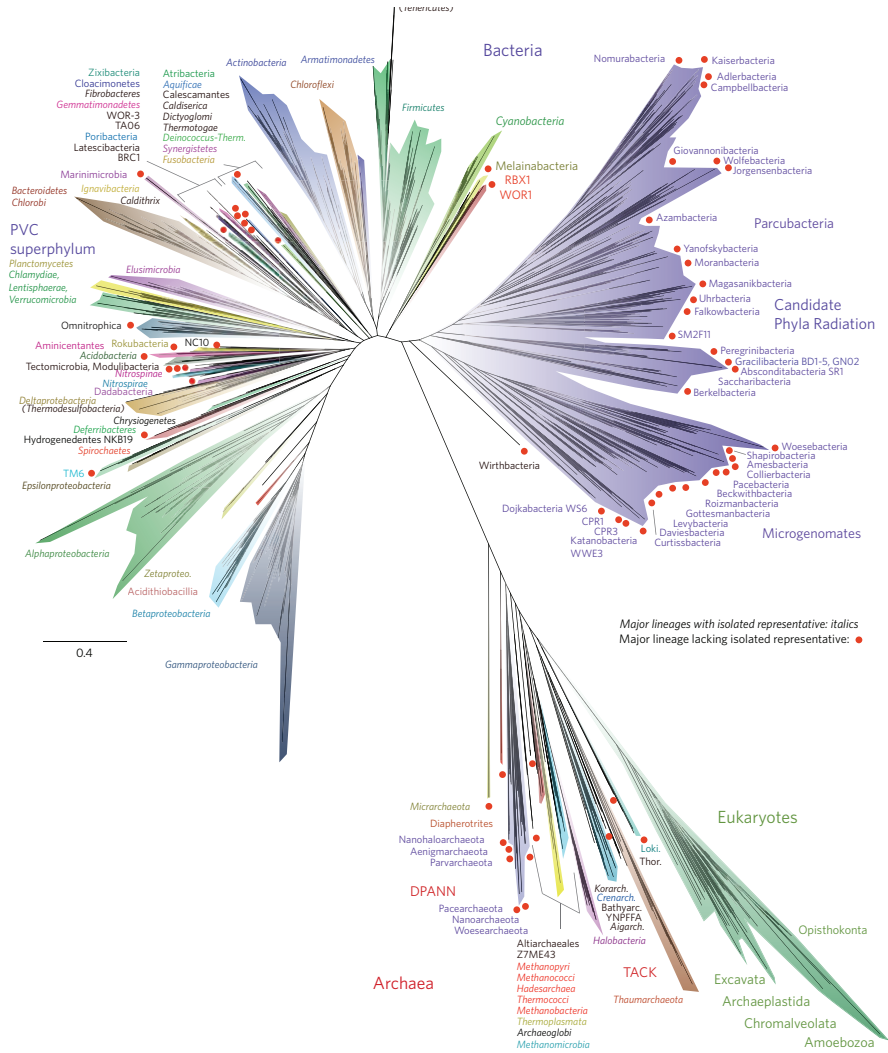


O associaedro. +



As cores (verde, azul, vermelho e preto) dão uma ideia muito subjetiva da dificuldade.





De “Uma nova visão da árvore da vida” *Nature Microbiology* 1, (2016).

Essas ramificações podem ser realizadas por grafos de polinômios?

# Polinômios que se intersectam

## Maxim Kontsevich

### Intercâmbios polinomiais

ANTES DE PROVAR O TEOREMA DE KONTSEVICH, deixe-me começar com uma consideração muito elementar. Considere a posição do gráfico de um polinômio real não nulo  $P(x)$  com respeito ao eixo  $x$ , numa vizinhança do 0.

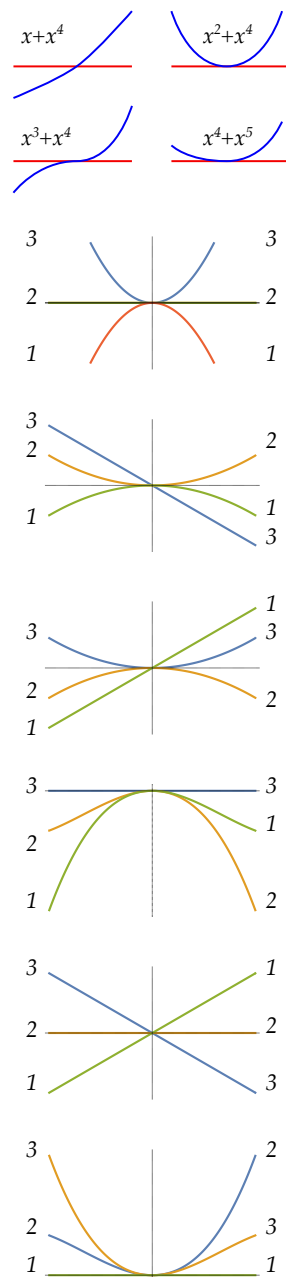
Existem duas possibilidades, a saber: ou o gráfico de  $P$  cruza o eixo  $x$  no 0, ou ele fica do mesmo lado. Para distinguir esses dois casos, introduzimos a seguinte definição.

**Definição.** Seja  $P(x) = a_0 + a_1x + a_2x^2 + \dots$  um polinômio (ou uma série formal). A *avaliação*  $v(P)$  de  $P$  (em 0) é o menor inteiro  $k$  tal que  $a_k \neq 0$ . Por definição, o valor do polinômio nulo é  $\infty$ .

Claramente, o gráfico de  $P$  cruza o eixo  $x$  sendo o valor  $v(P)$  um inteiro ímpar.

Dados dois polinômios distintos  $P_1, P_2$ , o sinal de  $P_1(x) - P_2(x)$  muda em 0  $v(P_1 - P_2)$  é ímpar.

Suponha agora que há três polinômios  $P_1, P_2, P_3$  e analisemos as possíveis configurações numa vizinhança da origem. As seis figuras na margem mostram que todas as seis permutações de  $\{1, 2, 3\}$  podem ser realizadas caso escolhermos os polinômios adequadamente.



Por exemplo:

$P_1(x) = x^2$	$P_2(x) = 0$	$P_3(x) = -x^2$
$P_1(x) = -x$	$P_2(x) = x^2$	$P_3(x) = -x^2$
$P_1(x) = x^2$	$P_2(x) = -x^2$	$P_3(x) = x$
$P_1(x) = 0$	$P_2(x) = -x^2 - x^3$	$P_3(x) = -x^2 + x^3$
$P_1(x) = -x$	$P_2(x) = 0$	$P_3(x) = x$
$P_1(x) = x^2 - x^3$	$P_2(x) = x^2 + x^3$	$P_3(x) = 0$

Portanto, o fenômeno de Kontsevich começa com *quatro* polinômios.

Note que as figuras sugerem que assumi  $P_i(0) = 0$ , mas essa suposição não é necessária. Considerarei apenas porque nesse livro discutiremos, principalmente, propriedades locais na vizinhança de um único ponto  $(0,0)$ .

Podemos, agora, provar o teorema do bilhete de metrô mencionado no prefácio.

Por contradição, assumimos que existam quatro polinômios  $P_1, P_2, P_3, P_4$  tais que:

- $P_1(x) < P_2(x) < P_3(x) < P_4(x)$  para  $x < 0$  pequeno,
- $P_2(x) < P_4(x) < P_1(x) < P_3(x)$  para  $x > 0$  pequeno.

Substituindo  $P_i$  por  $P_i - P_1$ , podemos considerar que  $P_1 = 0$ .

Como  $P_2$  e  $P_4$  mudam de sinal na origem, seus valores  $v(P_2), v(P_4)$  são ímpares.

Como  $P_3$  não muda de sinal, será  $v(P_3)$  é par.

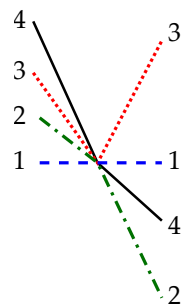
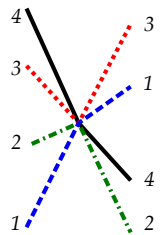
De  $0 < P_2(x) < P_3(x) < P_4(x)$  para  $x$  negativo pequeno, deduzimos que  $v(P_2) \geq v(P_3) \geq v(P_4)$ .

Analogamente, de  $|P_4(x)| < |P_2(x)|$  para  $x$  positivo pequeno, deduzimos que  $v(P_4) \geq v(P_2)$ .

Isso forçaria as três valuações a serem iguais, mas duas delas são ímpares e a outra é par!

Contradição. □

Note que a mesma prova se aplica a funções reais analíticas, mas não a funções suaves. De fato, o leitor pode facilmente encontrar quatro funções  $P_i$  ( $i = 1, 2, 3, 4$ ) de classe  $C^\infty$  cruzando na origem de acordo com a permutação “proibida”.



Eu uso o símbolo □ ao final de uma prova. Meus leitores astutos podem adivinhar por que eu ponho um · em um □? Dica: Pense em Francês.

Por quê?

Mudando a orientação ao longo do eixo  $x$ , podemos ver que a permutação inversa é também proibida. Como exercício, recomendo mostrar que as 22 permutações restantes  $\{1, 2, 3, 4\}$  ocorrem para uma escolha apropriada dos  $P_i$  ( $i = 1, 2, 3, 4$ ).

Tentemos, agora, analisar a situação com um número qualquer de polinômios.

**Definição.** Seja  $n \geq 2$  um inteiro e  $\pi$  alguma permutação de  $\{1, 2, \dots, n\}$ . Dizemos que  $\pi$  é um *intercâmbio polinomial* se existirem  $n$  polinômios  $P_1, \dots, P_n$  tais que:

- $P_1(x) < P_2(x) < \dots < P_n(x)$  para  $x$  negativo pequeno.
- $P_{\pi(1)}(x) < P_{\pi(2)}(x) < \dots < P_{\pi(n)}(x)$  para  $x$  positivo pequeno.

Nosso objetivo é dar uma descrição razoavelmente precisa dos intercâmbios polinomiais.

### Permutações Separáveis

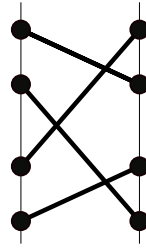
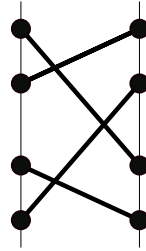
**Definição.** Seja  $n \geq 2$  um inteiro e  $\pi$  uma permutação de  $\{1, 2, \dots, n\}$ . Dizemos que  $\pi$  é *separável* se ela não “contém” nenhuma das duas permutações proibidas, i.e. se não existem quatro índices  $1 \leq i_1 < i_2 < i_3 < i_4 \leq n$  tais que  $\pi(i_2) < \pi(i_4) < \pi(i_1) < \pi(i_3)$  ou  $\pi(i_3) < \pi(i_1) < \pi(i_4) < \pi(i_2)$ .

Em outras palavras, uma permutação é separável se ela não contém uma das duas permutações de Kontsevich em quatro letras. Como já ficou claro, um intercâmbio polinomial é necessariamente separável. Nesta seção, proveremos a recíproca.

Começemos com um lema que parece ser “folclore” na literatura de Combinatória<sup>1</sup>.

**Lema.** *Seja  $\pi$  um permutação separável de  $\{1, 2, \dots, n\}$  (para  $n \geq 3$ ). Então, existe um intervalo próprio  $I = \{k, k + 1, \dots, k + l\}$  (onde  $1 \leq k \leq k + l \leq n$ ) de comprimento  $l + 1 \geq 2$  cuja imagem para  $\pi$  é um intervalo.*

Podemos assumir que  $\pi(1) < \pi(2)$  pois, caso contrário, basta substituir  $\pi$  pela permutação “reversa”  $\bar{\pi}(k) = n + 1 - \pi(k)$ .



As permutações proibidas.

A razão para a terminologia “separável” ficará clara no próximo capítulo.

<sup>1</sup> P. Bose, J. F. Buss, and A. Lubiw. Pattern matching for permutations. *Inform. Process. Lett.*, 65(5):277–283, 1998.

Observe que se  $\pi$  é um intercâmbio polinomial,  $\bar{\pi}$  também será (multiplique todos os polinômios por  $-x$ ). De modo similar, pela própria definição de permutação separável,  $\pi$  e  $\bar{\pi}$  são simultaneamente separáveis.



Se  $\pi(2) = \pi(1) + 1$  não há o que provar, pois a imagem de  $\{1, 2\}$  será o intervalo  $\{\pi(1), \pi(2)\}$ . Portanto, podemos assumir que  $\pi(2) > \pi(1) + 1$ . Considere o menor inteiro  $k$  tal que  $\pi(\{2, \dots, k\})$  contém o intervalo  $J$  entre  $\pi(1) + 1$  e  $\pi(2)$ . observe que  $\pi(k)$  está em  $J$  de modo que  $\pi(1) < \pi(k) < \pi(2)$ .

Se a imagem  $\pi(\{2, \dots, k\})$  for exatamente igual ao intervalo  $J$ , encontramos um intervalo não trivial cuja imagem por  $\pi$  é um intervalo.

Do contrário, podemos tomar um elemento  $l$  entre  $2$  e  $k$  cuja imagem por  $\pi$  está “fora” de  $J$ . Temos  $\pi(l) < \pi(1)$  ou  $\pi(l) > \pi(2)$ .

Se  $\pi(l) < \pi(1)$ , os quatro elementos  $1, 2, l, k$  satisfazem  $1 < 2 < l < k$  e  $\pi(l) < \pi(1) < \pi(k) < \pi(2)$ . São portanto ordenados de acordo com a “permutação proibida”, sendo impossível pela definição de permutação separável.

Podemos então assumir que todos os elementos de  $\pi(\{2, \dots, k\})$  são maiores ou iguais a  $\pi(1)$ .

Podemos também assumir que  $\pi(\{2, \dots, k\})$  é não trivial, pois, caso contrário, acabaríamos. Existe; portanto, pelo menos uma “lacuna” em  $\pi(\{2, \dots, k\})$  que deve estar à direita de  $\pi(2)$ . Logo, existe  $m$  tal que  $k < m$  e  $l$  tal que  $2 < l < k$  com  $\pi(m) < \pi(l)$ . Os quatro elementos  $2, l, k, m$  são tais que  $2 < l < k < m$  e  $\pi(k) < \pi(2) < \pi(m) < \pi(l)$  estando ordenados de acordo com a “permutação proibida” torna-se impossível para uma permutação separável.

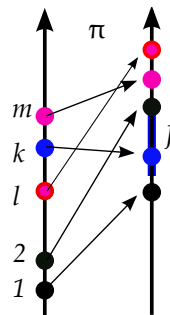
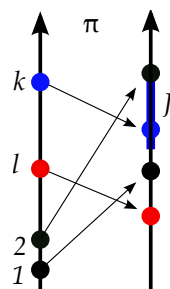
O lema está provado. □

É fácil melhorar o lema:

**Lema.** *Seja  $\pi$  uma permutação separável de  $\{1, 2, \dots, n\}$ . Podemos então encontrar dois inteiros consecutivos cujas imagens são consecutivas.*

A prova por indução é imediata, uma vez que qualquer intervalo próprio cuja imagem define outra permutação separável para um valor menor de  $n$  contém dois elementos consecutivos com imagens consecutivas. □

Podemos agora melhorar o resultado principal desse capítulo. O contra-exemplo de Kontsevich é, de alguma forma, o único contra-exemplo.



**Teorema.** *Uma permutação é um intercâmbio polinomial separável.*

Já havíamos observado que intercâmbios polinomiais são separáveis, pois essa é a observação de Kontsevich.

Novamente, por indução em  $n$ , mostraremos que toda permutação separável é um intercâmbio polinomial. Seja  $\pi$  uma permutação separável de  $\{1, 2, \dots, n\}$ . Sabemos que existem dois inteiros consecutivos  $i, i + 1$  com imagens consecutivas  $\pi(i), \pi(i + 1)$ .

Se “colapsarmos”  $\{i, i + 1\}$  e  $\{\pi(i), \pi(i + 1)\}$  em um único ponto, produziremos uma permutação  $\pi'$  em  $n - 1$  objetos, que é obviamente separável, sendo; portanto, um intercâmbio polinomial pela hipótese de indução. Segue que podemos encontrar  $n - 1$  polinômios

$$P_1, \dots, P_{n-1}$$

que se intersectam na origem de acordo com  $\pi'$ . O único que resta a fazer é dividir o  $i$ -ésimo polinômio  $P_i$  de modo a produzir  $n$  polinômios

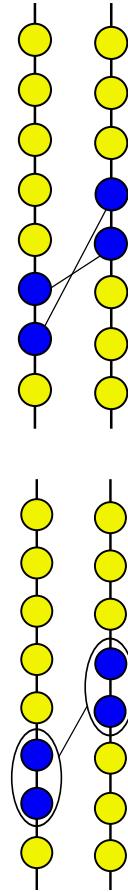
$$P_1, \dots, P_{i-1}, P'_i, P''_i, P_{i+1}, \dots, P_{n-1}$$

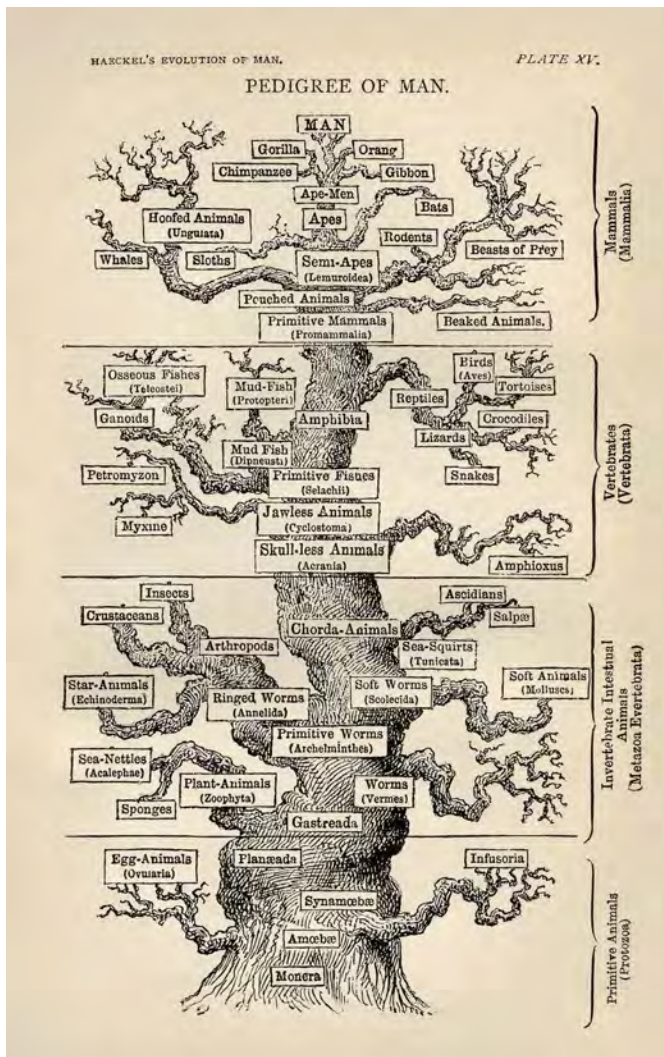
que se intersectam segundo  $\pi$ . É suficiente definir

$$P'_i(x) = P_i(x) \quad ; \quad P''_i(x) = P_i(x) + (-x)^N$$

para um valor de  $N$  suficientemente grande, par ou ímpar, de acordo com  $\pi(i + 1) > \pi(i)$  ou  $\pi(i + 1) < \pi(i)$ . □

Agora que identificamos os intercâmbios polinomiais, nossa próxima tarefa será entender a estrutura dessas permutações separáveis.





“Árvore da vida” (1879) de Ernst Haeckel.  
 O homem no topo da árvore da vida? +

# Padrões e permutações

## Donald Knuth

### Permutações

SE VOCÊS, LEITORES, FORAM TREINADOS MATEMATICAMENTE COMO EU FUI, podem ter sentido algum desconforto no capítulo anterior. Afinal, permutações costumam ser definidas como bijeções de um conjunto para si próprio e sua *raison d'être* é que elas constituem um grupo. Mas, não fizemos isso, manipulamos as permutações de uma maneira estranha quando usamos a expressão “A permutação  $\pi$  contém uma das duas permutações proibidas do Kontsevich” para significar que existem quatro índices  $1 \leq i_1 < i_2 < i_3 < i_4 \leq n$  tais que

$$\pi(i_2) < \pi(i_4) < \pi(i_1) < \pi(i_3), \text{ ou}$$
$$\pi(i_3) < \pi(i_1) < \pi(i_4) < \pi(i_2).$$

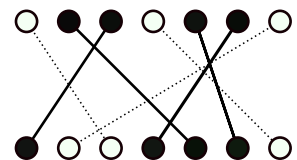
Isso certamente não significa que o conjunto com quatro elementos  $\{i_1, i_2, i_3, i_4\}$  é invariante por  $\pi$ . Nós *não* estamos tomando a restrição a um subconjunto invariante.

Usaremos a palavra “permutação” de um outro ponto de vista, mais próximo da ciência da computação. Essa abordagem se deve, em boa parte, à Donald Knuth em seu notável livro “The art of computer programming”<sup>2</sup>. O livro mais recente<sup>3</sup> é uma boa fonte de informação e mostra que essa área está florescendo.

Considere um conjunto finito  $E$  equipado com duas relações de ordem totais  $\ll$  e  $\lll$ . Pode-se ordenar os elementos usando a



Donald Knuth. +



<sup>2</sup> D. E. Knuth. *The art of computer programming. Vol. 1: Fundamental algorithms.* Second printing. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont, 1969.

<sup>3</sup> S. Kitaev. *Patterns in permutations and words.* Monographs in Theoretical Computer Science. An EATCS Series. Springer, Heidelberg, 2011. Com um prólogo por Jeffrey B. Remmel.

primeira relação

$$x_1 \ll x_2 \ll \dots \ll x_n$$

e olhar para a ordenação sob  $\ll$ , definindo uma permutação  $\pi$  de  $\{1, 2, \dots, n\}$  tal que

$$x_{\pi(1)} \ll\ll x_{\pi(2)} \ll\ll \dots \ll\ll x_{\pi(n)}$$

Nós vamos adotar esse ponto de vista: uma permutação é uma comparação entre duas ordens totais.

Por exemplo, o conjunto  $\{1, 2, 3, 4\}$  pode ser equipado com as ordens  $1 \ll 2 \ll 3 \ll 4$  e  $2 \ll\ll 4 \ll\ll 1 \ll\ll 3$ . Vamos denotar a permutação associada por  $(2, 4, 1, 3)$ .

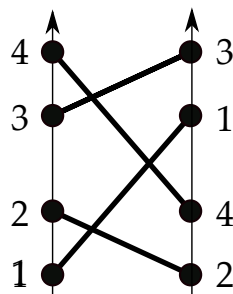
Qualquer conjunto finito de polinômios reais  $\{P_i(x)\}$  pode ser ordenado pelo menos de duas maneiras: comparando os valores de  $P_i(x)$  para pequenos valores negativos ou para pequenos valores positivos de  $x$ . Tais ordens nos conduzem aos intercâmbios polinomiais.

É sempre possível *restringir* relações de ordem a subconjuntos, o que define o conceito de inclusão para permutações.

**Definição.** Seja  $\pi$  uma permutação de  $\{1, \dots, n\}$  associada a duas relações de ordem totais  $\ll$  e  $\ll\ll$  em um conjunto  $E$  com  $n$  elementos. Seja  $F \subset E$  um subconjunto com  $p$  elementos. A restrição de  $\ll$  e  $\ll\ll$  a  $F$  define a permutação  $\sigma : \{1, \dots, p\} \rightarrow \{1, \dots, p\}$ . Em tal caso, dizemos que  $\sigma$  é *contida* em  $\pi$  e escrevemos  $\sigma \leq \pi$ .

Denote por  $\Sigma_n$  o conjunto (não considerado como um grupo) de permutações de  $\{1, \dots, n\}$  e  $\Sigma_\infty$  a união disjunta dos  $\Sigma_n$ , para  $n \geq 1$ . Assim, definimos uma relação de ordem parcial  $\leq$  sobre  $\Sigma_\infty$ . A compreensão dessa ordem é chamada de “reconhecimento de padrões”, pois também se diz que  $\sigma$  é *um padrão em  $\pi$*  quando  $\sigma \leq \pi$ .

Um subconjunto  $\mathcal{C} \subset \Sigma_\infty$  é chamado de *classe de permutação* se  $\pi \in \mathcal{C}$  e  $\sigma \leq \pi$  implicam  $\sigma \in \mathcal{C}$ . Dada uma classe de permutação, pode-se considerar sua *base*  $\mathcal{B}$ , que consiste nas permutações  $\pi$  que *não* estão em  $\mathcal{C}$ , mas tais que quaisquer  $\sigma \leq \pi$ , diferente de  $\pi$ , pertence a  $\mathcal{C}$ . Portanto, uma permutação  $\pi$  pertence a  $\mathcal{C}$  e não contém elementos de  $\mathcal{B}$ . Denota-se  $\mathcal{C} = Ev(\mathcal{B})$  e diz-se que  $\mathcal{C}$  consiste em permutações que *evitam*  $\mathcal{B}$ .



Cuidado. Nessa figura,  $1 \ll 2 \ll 3 \ll 4$  and  $2 \ll\ll 4 \ll\ll 1 \ll\ll 3$  e  $(\pi(1), \pi(2), \pi(3), \pi(4)) = (2, 4, 1, 3)$ , e, então,  $\pi$  é na realidade a *inversa* da permutação que se obtém acompanhando os segmentos dos pontos da esquerda para os da direita.

Por exemplo, toda permutação diferente da identidade contém  $(2, 1)$ .

O site [Database of Permutation Pattern Avoidance](#) contém uma vasta coleção de exemplos.

$Av((1, 2, 3))$  consiste nas permutações que podem ser escritas como união de duas seqüências decrescentes.

Por exemplo, o conjunto  $Inter \subset \Sigma_\infty$  dos intercâmbios polinomiais é obviamente uma classe de permutação. Já vimos que sua base consiste em dois elementos  $(2, 4, 1, 3), (3, 1, 4, 2)$ .

A seguir estão algumas questões típicas da teoria. Dada uma classe de permutação  $\mathcal{C}$ :

- é possível determinar sua base? Quando ela é finita?
- é possível contar o número de elementos em  $\mathcal{C} \cap \Sigma_n$ ? Ou, pelo menos, é possível estimá-lo?
- é possível determinar algoritmicamente se uma dada permutação  $\pi$  pertence à  $\mathcal{C}$ ? Qual é a complexidade de tal algoritmo?

Responderemos a todas essas perguntas no tempo devido para a classe dos intercâmbios polinomiais/permutações separáveis.

### Ordenações por pilha

A teoria de padrões de permutações recebeu um grande impulso a partir de um exercício do volume 1 de “The Art of computer programming”. Donald Knuth teve a ideia de atribuir um grau de dificuldade aos exercícios de seu livro.

Um “0” significa que o leitor deve ser capaz de resolvê-lo imediatamente.

Um “10” requer um minuto.

Um “20” pode demandar várias horas etc.

A escala é de fato logarítmica e parece conter um polo na vizinhança de 50...

O exercício que queremos discutir é rotulado como [M28]. O M significa que ele é destinado a leitores com enfoque matemático, e o 28 é uma indicação do tempo necessário para resolvê-lo (na escala logarítmica previamente explicada). Atualmente, essa questão não é tão difícil, mas ela teve uma influência duradoura sobre a área da combinatória.

Vamos descrever a classe de permutação que é definida por uma pilha.

Imagine  $n$  objetos rotulados  $1, 2, \dots, n$  alinhados sobre uma reta horizontal, nessa ordem da esquerda para a direita.  $\hat{A}$

Tente provar o [teorema de Erdős–Szekeres](#): toda permutação  $\pi \in \Sigma_n$  com  $n > (p - 1)(q - 1)$  contém  $(1, 2, 3, \dots, p)$  ou  $(q, q - 1, \dots, 2, 1)$ .

Encontre algum exemplo de classe de permutação com base infinita.



The art of computer programming. +

direita de  $n$ , há um “buraco”, uma espécie de poço onde os objetos podem ser empilhados uns sobre os outros.

Inicialmente, o buraco está vazio. Toma-se o objeto  $n$  e empurra-o para dentro do buraco. Então, há duas opções. Ou *empurrar* o último elemento da linha para o topo da pilha, originalmente denominado *push*. Ou *tirar* o elemento do topo da pilha para a direita, originalmente denominado *pop*.

Observe as figuras na margem e acompanhe a evolução dos objetos sob a sequência Push, Push, Pop, Push, Pop, Pop, Push, Push, Push, Pop, Pop, Pop.

No final do processo, a sequência  $(1, 2, 3, 4, 5, 6)$  se tornou  $(3, 2, 1, 6, 4, 5)$ . Pode ser pensado como uma permutação ou como duas ordens  $1 \ll 2 \ll 3 \ll 4 \ll 5 \ll 6$  (do lado esquerdo) e  $3 \lll 2 \lll 1 \lll 6 \lll 4 \lll 5$  (do lado direito).

**Definição.** Uma permutação  $\pi$  é uma *ordenação por pilha* se ela é o resultado de uma sequência de Push’s e Pop’s aplicados a  $\{1, 2, \dots, n\}$ .

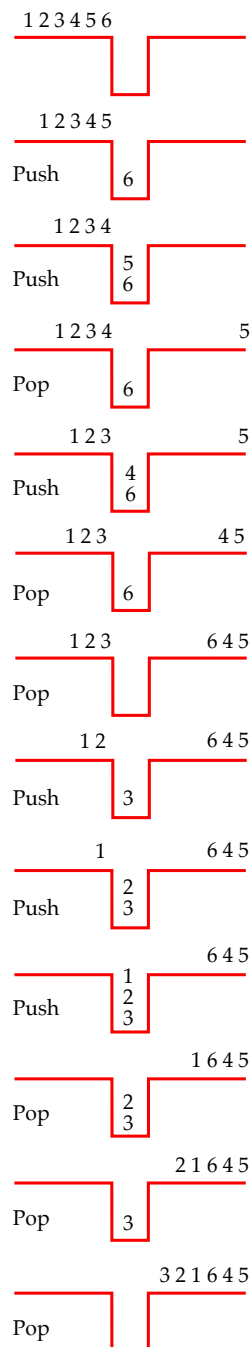
O exercício 5 do livro de Knuth, do capítulo 2, avaliado como [M28], é o seguinte:

**Teorema.** Uma permutação é um ordenação por pilha se, e somente se, ela não contém  $(2, 3, 1)$ .

Vamos resolver esse exercício:

Comece com uma permutação, por exemplo  $(3, 2, 1, 6, 4, 5)$ . O último elemento é 5. Se queremos formar essa permutação como uma pilha, não há escolhas: devemos empurrar todos os elementos da lista  $(1, 2, 3, 4, 5, 6)$  até que o 5 esteja disponível no topo da pilha, então, o tiramos e colocamos na posição devida, isto é, no final da lista resultante. Em seguida, procuramos pelo penúltimo elemento da permutação desejada, no caso 4, e seguimos empurrando até que o 4 esteja no topo da pilha etc. Portanto, se queremos formar uma permutação, só há uma maneira de fazê-lo.

Precisamos apenas entender quando o processo pode dar errado. No caso, ocorrerá exatamente quando for a hora de tirar algum objeto  $a$  que infelizmente já está na pilha e não no topo,

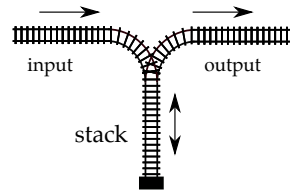


mas abaixo de algum objeto  $b \ll a$ . Se  $b$  já foi empurrado para a pilha, é porque tivemos que tirar anteriormente algum outro objeto  $c \ll b \ll a$ .

Temos  $a \lll c$ , pois  $c$  já foi tirado e estamos tentando tirar  $a$ . Similarmente, temos  $b \lll a$ , pois não queremos tirar  $b$  antes de  $a$ . Portanto, o subconjunto  $\{c, b, a\}$  em  $\{1, 2, \dots, n\}$  provoca a inclusão  $(2, 3, 1) \leq \pi$  que é o que tínhamos que provar.  $\square$

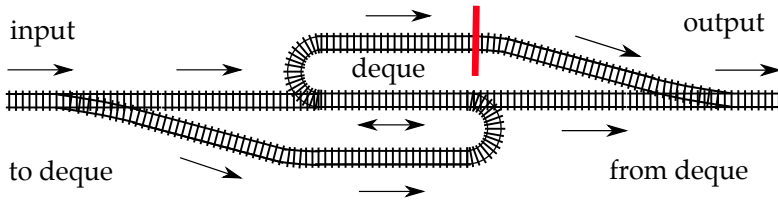
Eu o encorajo fortemente, leitor, a fazer *todos* os exercícios do livro do Knuth.

O traçado dos trilhos na margem produz ordenações por pilha. Um trem com vagões rotulados  $(1, 2, \dots, n)$  chega pela esquerda. Os vagões podem se separar e cada um deve seguir os trilhos nas direções dadas pelas setas. O trem é remontado na saída, na direita da figura.



Essa simples ferrovia produz ordenações por pilha.

Knuth também definiu “deques” (uma aglutinação de “deck” e “queue”), produzidos pelo traçado mais complicado mostrado abaixo.



Isso é um deck-queue = um deque.

Quais são as ordenações por deques?

Se fecharmos o portão vermelho, temos o chamado “output-restricted deque”. As permutações associadas são exatamente aquelas que não contêm  $(4, 2, 3, 1), (4, 1, 3, 2)$ . Estamos nos aproximando da caracterização dos intercâmbios polinomiais, que evitam  $(2, 4, 1, 3), (3, 1, 4, 2)$ .

Cuidado! a dificuldade dessa questão pode ser em torno de  $60!$

Para conhecer muito mais sobre essa área fascinante, eu recomendo o já mencionado livro do Kitaev.



*Catalan onipresente*

O exercício 4 do mesmo capítulo do livro do Knuth é avaliado [M34]. Entretanto, é mais fácil resolvê-lo se já foi resolvido o exercício 5.

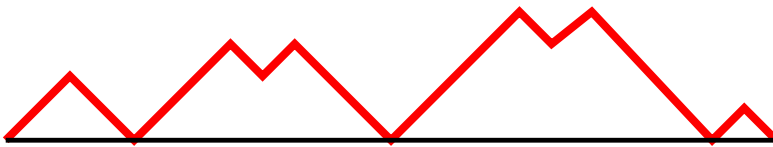
O problema é contar o número de ordenações por pilha de comprimento  $n$ .

Esse é o famoso  $n$ -ésimo número de Catalan  $C_n$ , que aparece em quase toda parte na matemática.

Os primeiros valores são:

1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, 208012, 742900, 2674440, 9694845, 35357670 etc. (sequência [A000108 na OEIS](#)).

Vimos que as ordenações por pilha são descritas *unicamente* por uma sequência de  $2n$  Push's e Pop's. Reciprocamente, se começarmos com uma sequência de Push's e Pop's, obtemos uma permutação desde que não tiremos algo de uma pilha vazia. Dito de outra forma, todo segmento inicial da sequência deve conter pelo menos tantos Push's quanto Pop's. Também é possível descrever a sequência considerando a evolução do número de elementos na pilha. A pilha deve estar vazia nos tempos 0 e  $2n$ , ganha +1 ou perde -1 nível para cada Push e Pop respectivamente e sempre é não negativa. Isso é chamada uma *palavra de Dyck* de comprimento  $2n$ . Aqui está um exemplo de uma palavra de Dyck de comprimento 24.

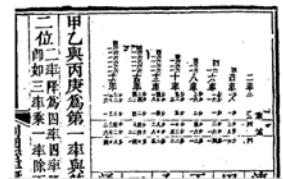


O número de palavras de Dyck é uma das muitas definições para o  $n$ -ésimo número de Catalan. Alternativamente, pode-se encarar um Push como um parêntese aberto "(" e um Pop como um fechado ")". A condição de não aplicar um Pop a uma pilha vazia é agora equivalente à sequência de parênteses estar bem balanceada: cada "(" aberto é pareado com um ")" fechado e eles



Eugène Catalan nasceu em 1814, em Bruges, na época pertencente ao Império Napoleônico.

Walther von Dyck (1856–1934) foi "o primeiro a definir matematicamente um grupo, no sentido moderno". Pelo menos é isso que se lê na *Wikipedia*. A questão é muito mais sutil e muitos outros nomes deveriam ser citados. Entretanto, ele foi realmente um dos primeiros a manipular apresentações de grupos, com geradores e relações.



A primeira publicação dos números de Catalan no [livro de Mingantu](#) "The Quick Method for Obtaining the Precise Ratio of Division of a Circle" por volta de 1730.

estão nessa ordem.

Por exemplo, para  $n = 3$ , temos as seguinte  $C_3 = 5$  sequências

$$((\ ))(\ ) ; (\ )((\ )) ; (((\ ))) ; (\ )(\ )(\ ) ; ((\ ))(\ )).$$

Há também uma interpretação em termos de “árvores planares enraizadas”. Uma imagem vale mais que mil palavras. Por alguma razão, matemáticos e cientistas da computação costumam desenhar árvores de cabeça para baixo: a raiz fica no topo, e as folhas na base.

A imagem na margem é um exemplo desse tipo de árvore. Ela tem uma raiz, três vértices internos e quatro folhas. A árvore é “planar” porque os filhos de seus vértices são ordenados da esquerda para a direita, digamos do mais velho para o mais novo.

Pode-se associar uma palavra de Dyck a uma dessas árvores. Basta começar pela raiz e seguir a árvore externamente, indo no sentido anti-horário. A cada etapa, você ou se afasta ou se aproxima da raiz, produzindo a sequência de +1 e -1, ou “Push” e “Pop”.

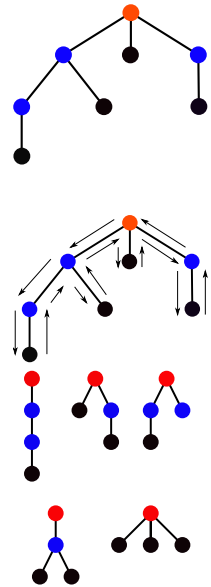
Nesse exemplo, obtemos a sequência  $+++--+-+--+-+---$ . Reciprocamente, pode-se transformar uma palavra de Dyck em uma árvore planar enraizada.

Juntando tudo, temos bijeções entre:

- Ordenações por pilha com  $n$  objetos.
- Palavras de Dyck de comprimento  $2n$ .
- Sequências balanceadas com  $2n$  parênteses ( $n$  abertos e  $n$  fechados).
- Árvores planares enraizadas com  $n$  arestas.

A cardinalidade de qualquer um desses conjuntos é o  $n$ -ésimo número de Catalan  $C_n$ .

Dada uma ordenação por pilha  $\pi$ , considere a posição do último elemento  $n$  na lista de objetos (à esquerda, em preto na imagem) e a sua posição  $k$  à direita. A permutação envia o intervalo (vermelho)  $\{k, \dots, n - 1\}$  em  $\{k + 1, \dots, n\}$ , e o intervalo



Cinco árvores planares enraizadas com três arestas.

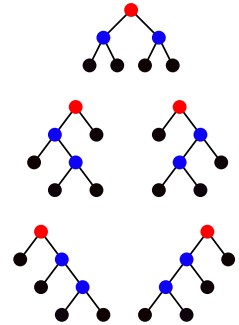
amarelo  $\{1, 2, \dots, k-1\}$  é invariante, induzindo portanto uma ordenação por pilha em cada um dos intervalos. Assim, chegamos à relação de recorrência:

$$C_n = \sum_{k=1}^n C_{k-1}C_{n-k}.$$

Essa é a marca característica de Catalan, encontrada em muitos contextos diferentes.

Por exemplo, considere árvores planares *binárias* enraizadas. Sua definição depende do autor, mas irei defini-las como árvores planares enraizadas tais que todo vértice não tem filho ou tem dois filhos, um sendo “à esquerda” e o outro “à direita”. Essas árvores com  $n$  vértices internos têm  $n + 2$  folhas e  $2n + 2$  arestas. Se removemos sua raiz, obtemos duas árvores planares binárias enraizadas. Reciprocamente, pode-se adicionar uma raiz comum a duas árvores planares binárias enraizadas para produzir uma outra ainda maior. Após um momento de reflexão, você, leitor, pode constatar que o número de árvores planares binárias enraizadas com  $n + 1$  folhas satisfaz a relação de recorrência de Catalan. Também pode-se verificar que há 1, 1, 2, 5 árvores planares binárias enraizadas com 1, 2, 3, 4 folhas, e, então, por indução, temos que  $C_n$  é também o número de árvores planares binárias enraizadas com  $n + 1$  folhas.

Olhe para essa relação de recorrência em termos de palavras de Dyck, seqüências de parênteses e árvores planares enraizadas.



Cinco árvores planares binárias enraizadas com quatro folhas.

Isso sugere que deve haver alguma correspondência entre “árvores planares enraizadas” e “árvores planares binárias enraizadas”. E de fato esse é o caso. Vou apresentar uma leve variação na chamada “transformada de Knuth” ou “representação primeiro filho-próximo irmão”: Começando com uma árvore planar enraizada  $T$  com  $n$  arestas (primeira figura), queremos construir uma árvore planar binária enraizada  $T_{bin}$  com  $n + 1$  folhas (última figura). Inicialmente constrói-se uma árvore auxiliar  $T'$  (segunda e terceira imagens). O conjunto de vértices de  $T'$  é o mesmo conjunto de vértices de  $T$ . A raiz é a mesma. Todo vértice  $v$  de  $T'$  tem no máximo dois filhos. O primeiro é o filho mais velho de  $v$  em  $T$  se ele existir. O segundo é o próximo irmão de  $v$  em  $T$ , o mais velho entre os irmãos mais novos de  $v$ , se houver. Então, é possível transformar  $T'$  em uma árvore planar binária enraizada  $T_{bin}$  da seguinte maneira: Primeiro apague a raiz e a aresta

partindo dela. A nova raiz de  $T_{bin}$  é o filho mais velho da raiz de  $T$ . Para cada vértice de  $T'$ , adicione um filho à esquerda se esse vértice não tiver filhos em  $T$  e adicione um filho à direita se esse vértice não tiver irmãos mais novos em  $T$ . Assim, se o vértice é uma folha de  $T'$  (i.e. ele não tem filhos nem irmãos mais novos em  $T$ ), adicione dois filhos em  $T_{bin}$  (observe os pontos verdes na quarta figura). Dessa forma, verifique que isso define uma bijeção entre árvores planares enraizadas com  $n$  arestas e árvores planares binárias enraizadas com  $n + 1$  folhas ([M15]).

Como de costume em combinatória, codificamos essa sequência  $C_n$  pela “série formal geradora”

$$C(t) = \sum_{n \geq 0} C_n t^n$$

e a recorrência se torna:

$$C(t) = tC^2(t) + 1.$$

Recordando equações quadráticas do ensino fundamental, obtemos

$$C(t) = \frac{1 - \sqrt{1 - 4t}}{2t}.$$

Segue dessa fórmula que o raio de convergência de  $C(t)$  é  $1/4$ . Então, pelo teorema de Cauchy–Hadamard, conseguimos uma estimativa para o crescimento de  $C_n$ :

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \ln C_n = \ln 4.$$

Também podemos usar essa fórmula para chegar a uma expressão bonitinha de  $C_n$ : basta usar a série binomial de Newton.

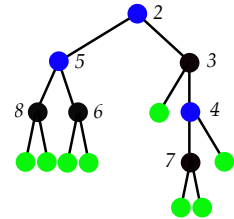
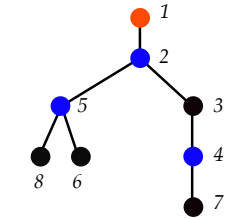
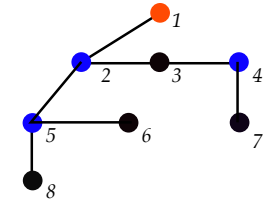
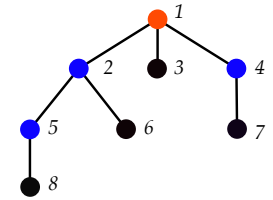
$$\sqrt{1 - 4t} = \sum_{n=0}^{\infty} \binom{1/2}{n} (-4t)^n.$$

Comparando os coeficientes de  $t^n$ , temos:

$$C_n = -\frac{1}{2} \binom{1/2}{n+1} (-4)^{n+1}.$$

Agora limpamos um pouco:

$$\begin{aligned} C_n &= -\frac{1}{2} \frac{1}{(n+1)!} \left(\frac{1}{2}\right) \left(\frac{1}{2} - 1\right) \dots \left(\frac{1}{2} - n\right) (-1)^{n+1} 2^{2(n+1)} = \dots \\ &= \frac{1}{n+1} \binom{2n}{n}. \end{aligned}$$



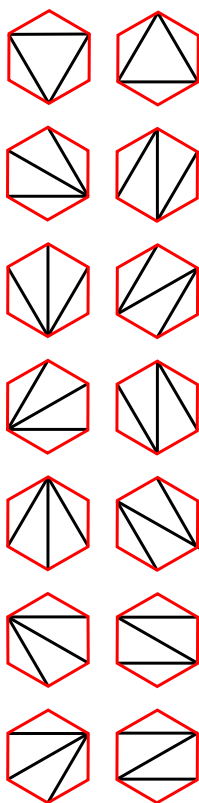
Por que eu escolhi o sinal – em frente a raiz quadrada?

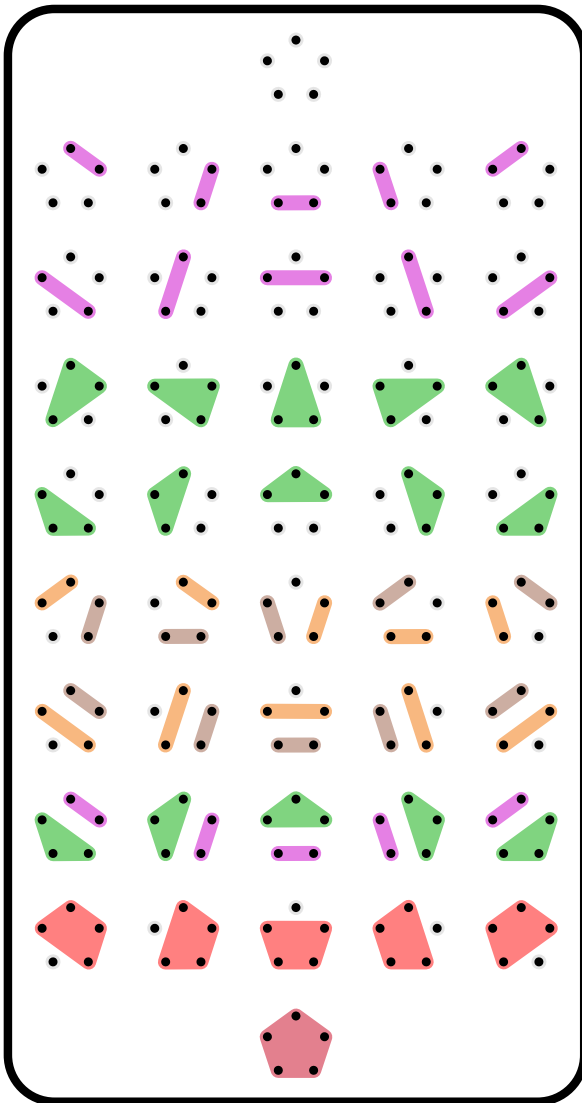
O segundo “...” significa que você é encorajado a calcular você mesmo!

Antes de finalizar, reforço que as bíblias do assunto sobre números de Catalan são os livros do Stanley <sup>4,5</sup>. Também, como fonte de informação notável, indico o site de Catalan “[Catalan Numbers](#)”, mantido por Igor Pak, e para obter uma perspectiva mais ampla (veja em particular o capítulo 6 sobre árvores): o livro escrito por Flajolet e Sedgewick <sup>6</sup>.

Há  $C_5 = 14$  maneiras de subdividir um hexágono em triângulos cujos vértices são vértices do hexágono e cujos interiores são disjuntos.

<sup>6</sup> R. S. P. Flajolet. *An introduction to the analysis of algorithms*. Addison-Wesley, 2nd ed edition, 2013.





42 partições sem interseções de um conjunto com 5 elementos. †



Uma árvore plana em Guadeloupe, comumente conhecida como a palmeira do viajante.

# Permutações separáveis

## Dos polinômios às árvores

O ANEL DE POLINÔMIOS  $K[x]$  COM COEFICIENTES EM UM CORPO  $K$  de característica 0 é equipado com uma *avaliação*  $v$ , dada pelo grau do primeiro coeficiente não nulo, e com uma *distância ultramétrica* natural, definida em termos de  $v$ :

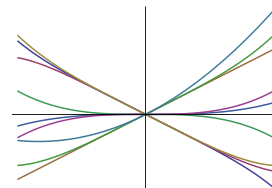
$$\text{dist}(P, Q) = \exp(-v(P - Q)).$$

Em linguagem clara, dois polinômios estão próximos se suas  $k$  primeiras derivadas em 0 coincidem para um valor grande de  $k$ . A propriedade ultramétrica para uma distância significa precisamente que para todo  $\varepsilon > 0$  a relação

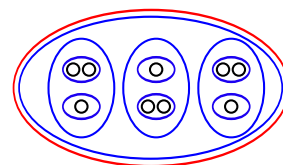
$$\text{dist}(P, Q) < \varepsilon$$

é uma relação de equivalência. À medida que  $\varepsilon$  diminui, essas relações de equivalência ficam mais finas e suas interseções tornam-se triviais.

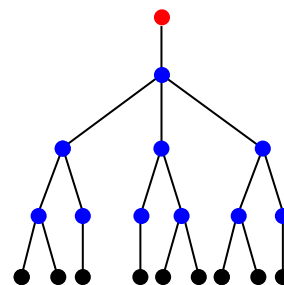
Considere um conjunto finito  $E$  de polinômios. Pode-se definir uma árvore da seguinte maneira: a raiz é rotulada pelo conjunto  $E$ . Os filhos da raiz são rotulados pelas classes de equivalência da relação  $v(P - Q) \geq 1$ . Os netos da raiz são rotulados pelas classes de equivalência da relação  $v(P - Q) \geq 2$ . E, em geral, a  $k$ -ésima geração corresponde às classes de equivalência da relação  $v(P - Q) \geq k$ . Assim, essa árvore é infinita, mas as classes de equivalência estabilizam em conjuntos unitários quando  $k$  é grande, de modo que podemos podá-la e produzir uma árvore finita



Nove polinômios cujas árvores associadas se encontram abaixo. Você consegue propor nove equações?



Relações de equivalência encaixadas.



Árvore associada.



cujas folhas são rotuladas pelos elementos de  $E$ . Reciprocamente, podemos recuperar a estrutura de valuação pela árvore. Dados dois elementos  $P, Q$  de  $E$ , vistos como folhas, procuramos na árvore seu ancestral comum mais próximo. A valuação de  $P - Q$  é o nível desse ancestral, ou seja, o comprimento do caminho ligando esse ancestral à raiz.

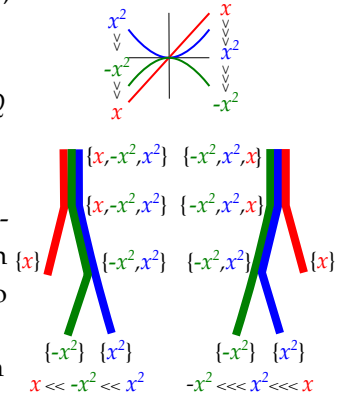
Suponha agora que  $K$  é o corpo dos números reais. Como já foi observado, qualquer conjunto finito de polinômios é equipado com duas ordens totais, comparando valores para  $x$  pequeno negativo e positivo, fornecendo uma ordem no conjunto das folhas da árvore dada pela valuação. Observe que os descendentes de um vértice, i.e. uma classe de equivalência a um certo nível, definem um intervalo nas duas ordenações. Portanto, nossa árvore é uma árvore planar de dois modos distintos.

Por convenção, escolhemos a primeira ordem (i.e. para  $x < 0$  pequeno) e associamos ao nosso conjunto  $E$  de polinômios a árvore planar correspondente. A comparação das duas ordens define uma permutação  $\pi$  que havíamos chamado de um intercâmbio polinomial.

O conjunto de folhas de qualquer árvore planar é equipado com duas ordens canônicas. Sendo que a primeira, denotada por  $\ll$ , é simplesmente a ordem dada pela definição de planaridade, e a segunda, denotada por  $\lll$ , é definida da seguinte maneira: Dadas duas folhas  $a$  e  $b$ , denote por  $a \vee b$  seu ancestral comum mais próximo. Então,  $a \ll b$  e  $a \lll b$  valem simultaneamente o nível de  $a \vee b$  é par. Deve-se checar se isso define de fato uma ordem  $\lll$ , em outras palavras, que  $a \lll b \lll c \lll a$  não é possível. Podemos assumir que  $a \ll b \ll c$  ou que  $c \ll b \ll a$  e, assim, o segundo caso se reduz ao primeiro por simetria. Se tivermos  $a \lll b \lll c \lll a$ , o nível do ancestral comum mais próximo de  $(a, b)$  e  $(b, c)$  deve ser ímpar, e o de  $(c, a)$  deve ser par. As figuras na margem mostram que isso não é possível.

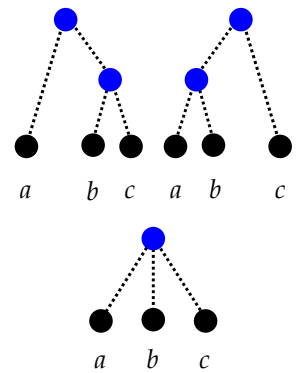
Vamos recapitular:

- Um conjunto finito de polinômios reais produz uma árvore planar com raiz.
- Uma árvore planar com raiz define duas ordens no seu conjunto de folhas e, portanto, uma permutação das folhas.



As duas árvores planas associadas aos polinômios  $\{x, -x^2, x^2\}$ , para valores pequenos de  $x$  negativos (left  $\ll$ ) e positivos (right  $\lll$ ).

Se  $P(x) < R(x) < Q(x)$  para valores positivos pequenos  $x$  (ou para  $x$  negativo pequeno), então  $v(P - Q) \geq v(R - Q)$ . Segue que todas as classes de equivalência  $val(P - Q) \geq k$  são intervalos em ambas as ordens: nossas árvores são de fato planas.



- A permutação associada à árvore planar que é relacionada a um conjunto finito de polinômios é simplesmente o intercâmbio polinomial correspondente.

Nossas árvores contêm informação demais e por isso vamos podar suas arestas.

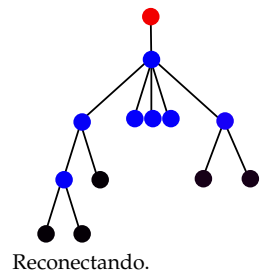
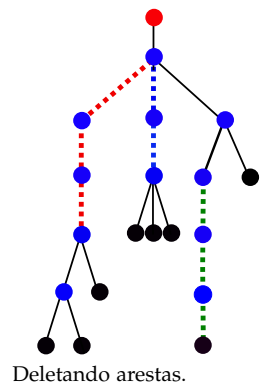
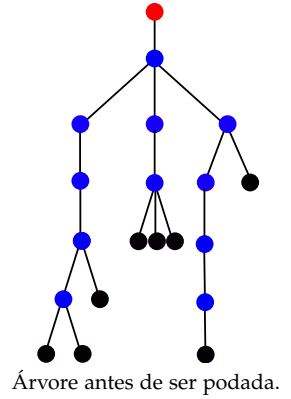
Primeiro, se um vértice interno tem somente uma folha como descendente, podemos remover todos os descendentes sem modificar a permutação associada (caminho verde na margem).

Suponha que dois vértices  $P, Q$  são conectados na árvore por algum caminho tal que todos os vértices entre  $P$  e  $Q$  são não ramificados, i.e., têm apenas um filho (caminhos vermelho e azul). Se o número de arestas nesse caminho for par, apenas as delete e identifique os dois pontos terminais em um único vértice. Se o número de arestas nesse caminho for ímpar, apenas as delete e conecte os dois pontos terminais por uma única aresta. Isso produz uma nova árvore. Nesse processo, o nível de alguns vértices mudou, mas apenas por um número par. Portanto, se computarmos a valuação na nova árvore, a paridade não mudará, constituindo-se na única informação que importa para construir o intercâmbio polinomial. Note que a árvore podada compartilham da propriedade de que todos os seus vértices internos têm ao menos dois filhos.

Resumindo, dados  $n$  polinômios, pode-se construir uma *árvore planar* tal que:

- A raiz pode ter qualquer número de filhos.
- Cada vértice interno tem ao menos dois filhos.
- Existem exatamente  $n$  folhas, rotuladas pelos  $n$  polinômios.

Dizemos que uma árvore planar é *podada* se ela satisfaz essas propriedades. Deve ser claro que, para qualquer árvore podada, pode-se encontrar  $n$  polinômios tais que a árvore podada associada seja a árvore dada. Vimos que, para qualquer conjunto finito de polinômios, o intercâmbio polinomial associado pode ser lido da árvore. Em particular, o número de intercâmbios polinomiais é menor ou igual ao número de árvores podadas.



*De uma permutação a uma árvore*

Mostraremos agora que o número de intercâmbios polinomiais é igual ao número de árvores podadas. O problema é mostrar que duas árvores podadas diferentes com  $n$  folhas produzem permutações diferentes.

Rotulamos as folhas da nossa árvore planar  $T$  por  $1 \ll 2 \ll \dots \ll n$ , da esquerda para a direita. Seja  $\pi$  a permutação associada ( $n \geq 2$ ).

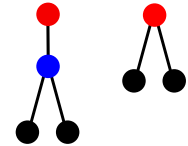
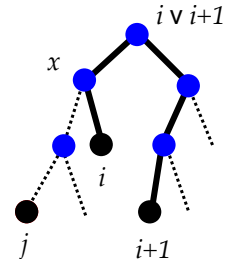
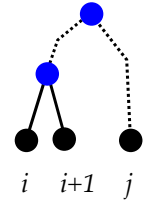
**Lema.** *As imagens por  $\pi$  de dois inteiros consecutivos  $i, i + 1$  são consecutivos as folhas  $i, i + 1$  são irmãos, i.e. têm o mesmo pai.*

De fato, se  $i$  e  $i + 1$  têm o mesmo pai, temos  $i \vee j = (i + 1) \vee j$  por todo  $j \neq i, i + 1$ . Segue que  $i \lll j \lll i + 1$  e  $i + 1 \lll j \lll i$  é impossível. Ou seja,  $i, i + 1$  são consecutivos por  $\lll$ .

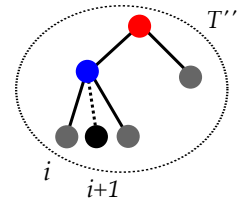
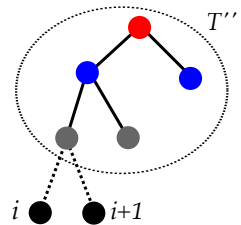
Reciprocamente, suponha que  $i \vee (i + 1)$  não seja o pai. Nesse caso, o caminho que conecta  $i$  e  $i + 1$  na árvore tem comprimento pelo menos 3 e contém, portanto, algum vértice interno  $x$  tal que os níveis de  $x$  e  $i \vee (i + 1)$  sejam diferentes. Escolha uma folha  $j$  que é um descendente de  $x$  diferente de  $i, i + 1$  (que existe desde que  $T$  está podada). Segue-se que  $j$  está entre  $i$  e  $i + 1$  para a ordem  $\lll$ . □

Assim temos outra prova de que qualquer intercâmbio polinomial contém pelo menos dois inteiros consecutivos com imagens consecutivas. Para isso, basta considerar um vértice interno com o nível mais alto em  $T$ : ele tem pelo menos dois filhos, que devem ser folhas e, portanto, irmãos.

Agora podemos provar, por indução em  $n \geq 2$ , que há no máximo uma árvore podada produzindo uma dada permutação  $\pi$ . Isso é trivial para  $n = 2$ . Se  $T_1$  e  $T_2$  tiverem  $n$  folhas e definirem a mesma permutação  $\pi$ , o lema anterior mostra que há alguns pares ou folhas consecutivas  $i, i + 1$  que são irmãos, para ambos  $T_1$  e  $T_2$ . Delete a folha  $i + 1$  de  $T_1$  e  $T_2$ , produzindo árvores  $T'_1$  e  $T'_2$  com  $n - 1$  folhas. Elas claramente  $T'_1$  e  $T'_2$  definem a mesma permutação  $\pi'$  com  $n - 1$  folhas. Entretanto, devemos ter cuidado, pois  $T'_1$  ou  $T'_2$  podem não ser podadas. Isso acontece precisamente quando  $i$  e  $i + 1$  são os únicos filhos de algum vértice interno em  $T_1$  ou  $T_2$ . A hipótese de indução mostra que as árvores podadas  $T''_1$  e  $T''_2$



As duas árvores podadas com duas folhas, definindo a transposição e a identidade.



são iguais a algumas  $T''$ . Nossas árvores  $T_1$  e  $T_2$  são obtidas de  $T''$  por meio de uma das duas operações a seguir: adicionando dois filhos à  $i$ -ésima folha de  $T''$ , ou adicionando uma irmã à  $i$ -ésima folha de  $T''$ . Assumindo que  $T_1$  e  $T_2$  produzem a mesma permutação, os pais de  $i$  e  $i + 1$  em  $T_1$  e  $T_2$  têm níveis com a mesma paridade. Segue-se que  $T_1$  e  $T_2$  são obtidos de  $T''$ , executando a mesma operação. Então,  $T_1 = T_2$ , conforme desejado.  $\square$

Portanto, o número de intercâmbios polinomiais de tamanho  $n$  é igual ao número de árvores podadas com  $n$  de folhas.

*De uma árvore podada a um intercâmbio polinomial e a uma permutação separável*

Isso nos leva à definição original de *permutação separável*: Dadas duas permutações  $\pi_1$  e  $\pi_2$  de  $n_1$  e  $n_2$  objetos ordenados, podemos pensar em duas maneiras para produzir uma permutação de  $n_1 + n_2$  objetos, a saber: enumere os primeiros  $n_1$  objetos como  $\{1, 2, \dots, n_1\}$ , e os  $n_2$  seguintes como  $\{n_1 + 1, n_1 + 2, \dots, n_1 + n_2\}$ . Denote por  $\pi_1 \oplus \pi_2$  a permutação definida por

$$\begin{aligned} \pi_1 \oplus \pi_2(k) &= \pi_1(k) && \text{se } 1 \leq k \leq n_1 \\ &= n_1 + \pi_2(k - n_1) && \text{se } n_1 + 1 \leq k \leq n_1 + n_2. \end{aligned}$$

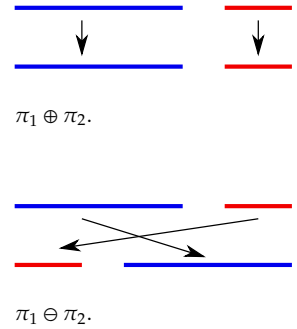
Defina também

$$\begin{aligned} \pi_1 \ominus \pi_2(k) &= \pi_2(k) + n_1 && \text{se } 1 \leq k \leq n_2 \\ &= \pi_1(k - n_2) && \text{se } n_2 + 1 \leq k \leq n_1 + n_2. \end{aligned}$$

Segundo a definição<sup>7</sup> de 1998, uma permutação é separável se ela é obtida a partir de diversas cópias da permutação trivial em um objeto, por meio de sucessivas operações  $\oplus$  e  $\ominus$ . Temos, de fato, um entendimento razoavelmente bom dessas permutações<sup>8</sup>.

**Teorema.** *Seja  $\pi$  uma permutação de  $\{1, \dots, n\}$ . As seguintes condições são equivalentes:*

1.  $\pi$  é o intercâmbio polinomial associado a  $n$  polinômios distintos  $P_1, \dots, P_n$ .



<sup>7</sup>P. Bose, J. F. Buss, and A. Lubiw. Pattern matching for permutations. *Inform. Process. Lett.*, 65(5):277–283, 1998.

Note que, se uma permutação  $\pi$  é separável, sua reversa  $\bar{\pi}(k) = n + 1 - \pi(k)$  também será.

<sup>8</sup>E. Ghys. Intersecting curves (variation on an observation of Maxim Kontsevich). *Amer. Math. Monthly*, 120(3):232–242, 2013.

2.  $\pi$  não contém  $(2,4,1,3)$  ou  $(3,1,4,2)$ .
3.  $\pi$  é a permutação definida por uma árvore podada.
4.  $\pi$  é obtida a partir de diferentes cópias da permutação trivial em um objeto por sucessivas operações  $\oplus$  e  $\ominus$ .

Essas permutações já foram definidas como separáveis.

Até aqui já provamos tudo exceto a equivalência entre 3 e 4. Provemos essa equivalência por indução:

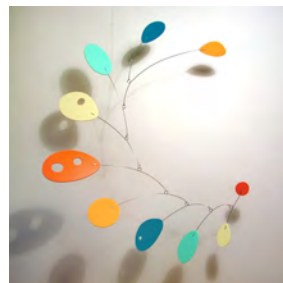
Seja  $\pi$  descrita por uma árvore podada  $T$ . A raiz dessa árvore pode ter um único filho. Se esse for o caso, os descendentes desse único filho definem uma outra árvore podada  $\bar{T}$  na qual a nova raiz tem vários filhos. A permutação associada à  $\bar{T}$  é a reversa  $\bar{\pi}$ . Podemos, portanto, assumir que a raiz tem vários filhos, que definem árvores podadas e intercâmbios polinomiais  $\pi_1, \pi_2, \dots, \pi_k$  cuja soma  $\oplus$  é  $\pi$ . Por indução, as permutações  $\pi_1, \pi_2, \dots, \pi_k$  são obtidas da permutação trivial pelas operações  $\oplus$  e  $\ominus$  de modo que o mesmo vale para a permutação  $\pi$ .

A recíproca é igualmente fácil: Deve-se mostrar que as somas  $\oplus$  e  $\ominus$  de duas permutações associadas às árvores podadas também estão relacionadas a uma árvore podada. Para isso, basta juntar as raízes das duas árvores podadas com um pai comum, ou adicionar um avô, dependendo do sinal. □

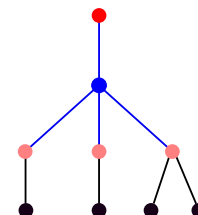
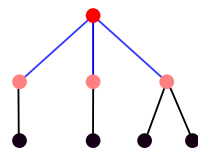
### Trilhos de trem, pilhas, plantas e permutons

Imagine um trem (matemático) constituído por  $n$  vagões. Insira uma *plataforma rotatória* matemática, um daqueles dispositivos que vemos às vezes em ferrovias, que permite rotações de 180 graus. A plataforma rotatória é matemática, pois assumimos que ele pode comportar um número qualquer de vagões consecutivos. Consideramos também que revertido um segmento de vagões, eles ficarão enganchados de modo permanente, sem possibilidade de desconexão. Permita que o trem se mova várias vezes na plataforma rotatória. A permutação final dos carros é separável quase que por definição.

Pode-se expressar a mesma ideia fazendo “pop de pilhas em série”, como mostrado no capítulo anterior. Imagine um número



Alexander Calder, *MóBILE*. Evidentemente, os móveis de Calder não são feitos para serem planares, mas quaisquer das suas duas configurações planares diferem por uma permutação separável de suas folhas. †



Adicionando um novo pai ou avô para juntar várias árvores.

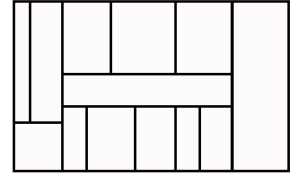


Uma plataforma rotatória. †

indefinido de pilhas alinhadas à direita da sequência  $1, 2, \dots, n$ . As regras do jogo de ordenação são diferentes do caso de uma só pilha: A cada passo é permitido empurrar um elemento da lista (“push”) no topo da primeira pilha, ou empilhar (“pop”) *todo o conteúdo* de alguma pilha sobre a pilha seguinte.

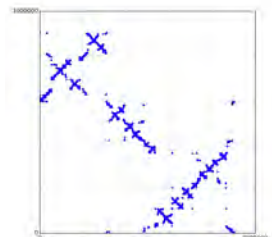
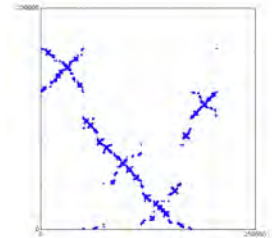
Aqui temos uma outra definição de permutação separável: Comece com um retângulo e o decomponha em diversos “quartos retangulares” por meio de sucessivas divisões horizontais ou verticais. Obtém-se uma planta, como na margem. Encontre uma boa definição para plantas equivalentes e para uma bijeção com permutações separáveis <sup>9</sup>.

Para concluir este capítulo, deixe-me mencionar um artigo recente <sup>10</sup>, que descreve o comportamento probabilístico de uma permutação separável quando  $n$  tende a infinito. Dada uma permutação  $\pi$  de  $\{1, \dots, n\}$ , pode-se considerar seu gráfico: o subconjunto  $\{(i, \pi(i))\} \subset \{1, \dots, n\}^2$ . Podemos mudar a escala dessa figura de modo a desenhá-la no quadrado unitário  $[0, 1]^2$ : para cada permutação  $\pi$ , associamos uma medida de probabilidade  $\mu_\pi$  no quadrado que é a soma das  $n$  medidas de Dirac com peso  $1/n$  no ponto  $(i/n, \pi(i)/n)$ . O espaço  $Prob([0, 1]^2)$  de medidas de probabilidade no quadrado é compacto (na topologia fraca) de forma que podemos estudar os pontos de acumulação das  $\mu_\pi$  em  $Prob([0, 1]^2)$ . É fácil ver que qualquer ponto de acumulação  $\mu$  é o que se chama de *permuton*: uma medida de probabilidade no quadrado cujas distribuições marginais (suas duas projeções nos eixos) são a medida de Lebesgue em  $[0, 1]$ . O artigo de Bassino et al. discute o limite de permutações separáveis: Para cada  $n$ , escolha uma permutação separável aleatoriamente (uniformemente distribuída dentre todas as permutações separáveis), produzindo uma probabilidade no espaço  $Prob([0, 1]^2)$  para cada  $n$ . Os autores provam que essa sequência de probabilidades converge para uma probabilidade bem definida no espaço  $Prob([0, 1]^2)$ . Para esse limite, uma probabilidade concentrada no espaço de permutons, denomina-se: o “permuton separável aleatório”. As duas figuras na margem, extraídas desse estudo de Bassino et al. mostram gráficos típicos de permutações separáveis para valores grandes de  $n$ .



<sup>9</sup> E. Ackerman, G. Barequet, and R. Y. Pinter. [A bijection between permutations and floorplans, and its applications](#). *Discrete Appl. Math.*, 154(12):1674–1684, 2006.

<sup>10</sup> F. Bassino, M. Bouvel, V. Féray, L. Gerin, and A. Pierrot. [the brownian limit of separable permutations](#). *Ann. Probab.*, 46(4):2134–2189, 07 2018.



†

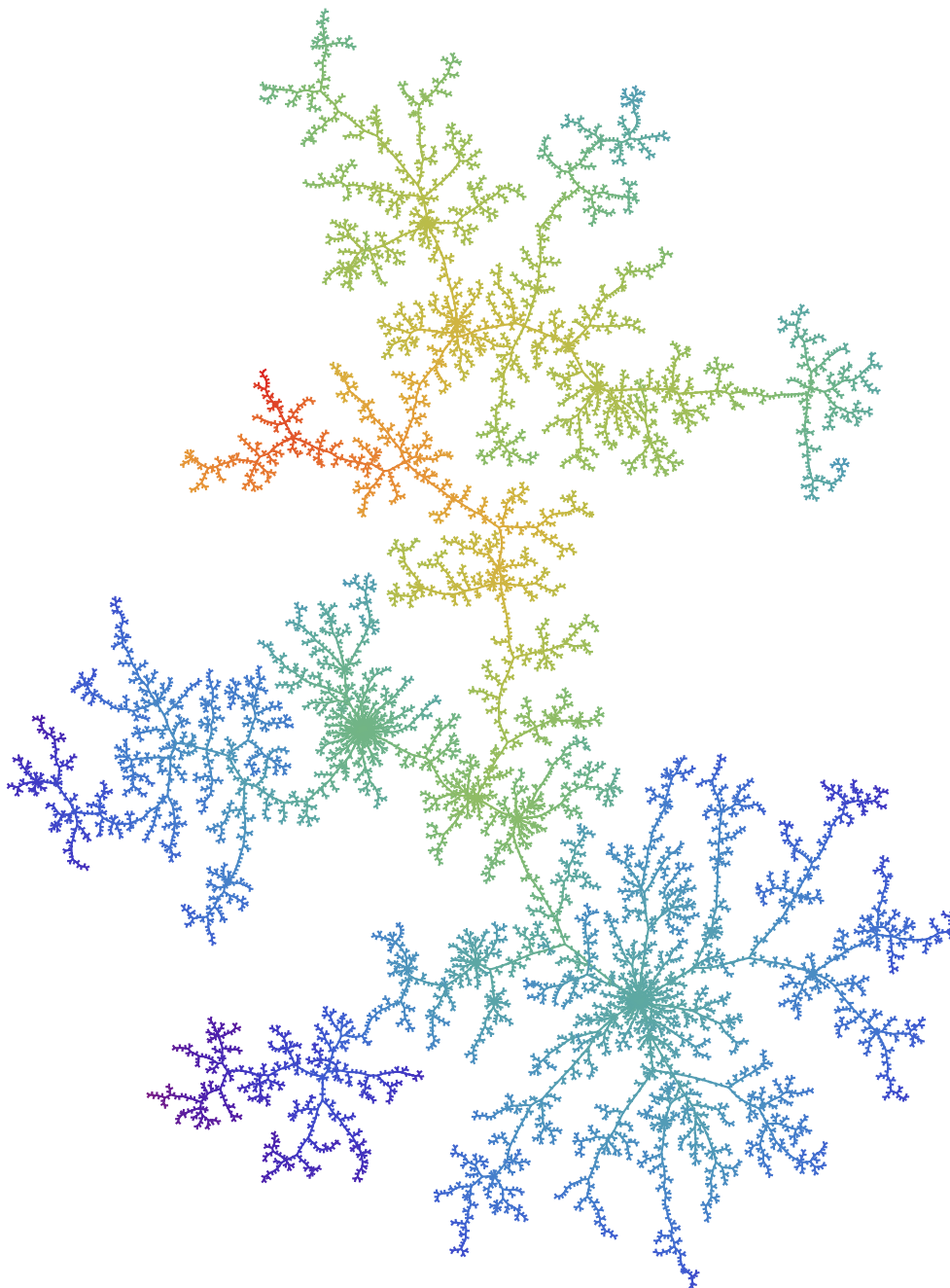
Para aprofundamentos em combinatória de permutações separáveis, recomendo dois livros <sup>11</sup> ou <sup>12</sup>. Antes de encerrar, deixe-me, no entanto, propor um exercício:

*Mostre que existe um algoritmo que decide se uma permutação de  $\{1, 2, \dots, n\}$  dada é ou não separável em tempo linear (em  $n$ ).*

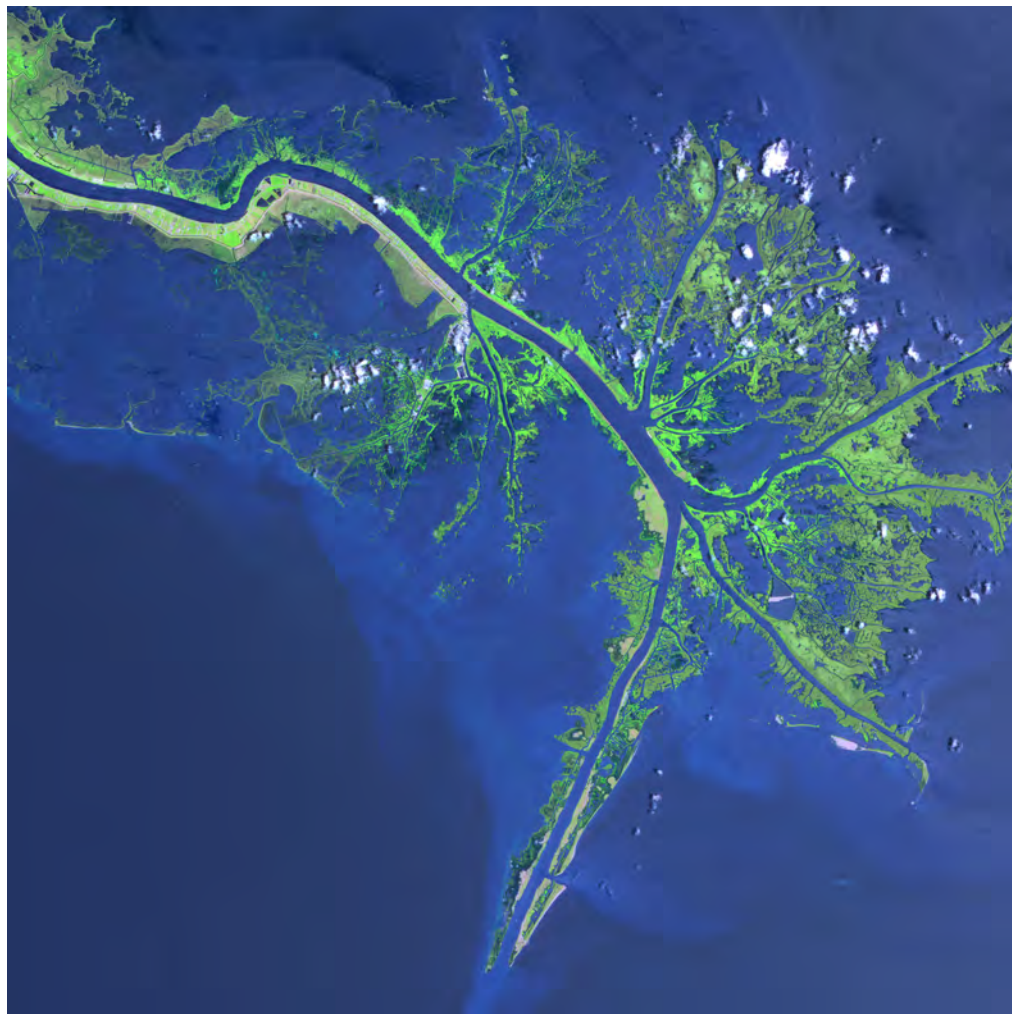
Note que existe um algoritmo óbvio em *tempo polinomial*: para cada quádrupla  $1 \leq i_1 \leq i_2 \leq i_3 \leq i_4 \leq n$  verifique se suas imagens estão ordenadas como uma das permutações proibidas. Ir de tempo polinomial para tempo linear pode ser importante visto que  $\log n!$  cresce mais rápido que uma função linear em  $n$ , porém, mais devagar que uma função quadrática pela fórmula de Stirling. Portanto, se você conseguir encontrar um algoritmo com tempo linear, você estará provando, em particular, que o número de permutações separáveis é pequeno quando comparado a  $n!$ . Como dica para o exercício, sugestão: leia novamente a prova da bijeção entre as permutações separáveis e as árvores podadas com raiz.

<sup>11</sup> S. Kitaev. *Patterns in permutations and words*. Monographs in Theoretical Computer Science. An EATCS Series. Springer, Heidelberg, 2011. Com um prólogo por Jeffrey B. Remmel.

<sup>12</sup> M. Bóna. *Combinatorics of permutations*. Discrete Mathematics and its Applications. CRC Press, Boca Raton, FL, second edition, 2012. Com um prólogo por Richard Stanley.







Delta do rio Mississippi. +

# Hiparco e Schroeder

Vamos contar árvores

NÓS VAMOS CONTAR o número  $a(n)$  de intercâmbios polinomiais (ou permutações separáveis) com  $n$  objetos.

Seja  $b(n)$  o número de árvores podadas com  $n$  folhas tais que a raiz tem *pelo menos dois filhos* se  $n \geq 2$  (e não tem filhos, se  $n = 1$ ). Para cada uma dessas árvores, pode-se criar uma nova raiz que tem a raiz original como seu único filho. Segue que  $a(n) = 2b(n)$  para  $n \geq 2$ . Os primeiros valores de  $b$  são:

$b(1) = 1$ : uma pequena árvore cuja raiz é sua única folha.

$b(2) = 1$ : uma pequena árvore com dois galhos e duas folhas.

$b(3) = 3$ .

É muito tentador estabelecer uma relação de recorrência para  $b(n)$ .

Comece com uma árvore podada com  $n$  folhas tal que a raiz tenha pelo menos dois filhos. Apagando a raiz e suas arestas, obtém-se um certo número de árvores, havendo um total de  $n$  folhas. Reciprocamente, começando com um conjunto ordenado contendo pelo menos duas árvores podadas e com  $n$  folhas no total, pode-se adicionar uma nova raiz e conectá-la às raízes anteriores, para assim construir uma árvore podada com  $n$  folhas.

Portanto, temos a seguinte relação:

$$b(n) = \sum_{i_1, i_2, \dots, i_k; i_1 + \dots + i_k = n} b(i_1)b(i_2)\dots b(i_k).$$

Agora usamos o método clássico de *série geradora*. Defina a série

Os primeiros valores de  $a(n)$ :

1 1  
2 2  
3 6  
4 22  
5 90  
6 394  
7 1806  
8 8558  
9 41586  
10 206098  
11 1037718  
12 5293446  
13 27297738  
14 142078746  
15 745387038  
16 3937603038  
17 20927156706  
18 111818026018  
19 600318853926  
20 3236724317174  
21 17518619320890  
22 95149655201962  
23 518431875418926  
24 2832923350929742  
25 15521467648875090  
26 85249942588971314  
27 469286147871837366  
28 2588758890960637798  
29 14308406109097843626  
30 79228031819993134650  
31 439442782615614361662

de potências formal  $H$  por:

$$H(t) = \sum_{n=1}^{\infty} b(n)t^n = t + t^2 + 3t^3 + \dots$$

Vamos tomar o quadrado de  $H$ :

$$H(t)^2 = t^2 + 2t^3 + 7t^4 + \dots$$

O coeficiente de  $t^n$  nessa nova série é  $\sum_{i_1+i_2=n} b(i_1)b(i_2)$ , sendo igual ao número de árvores podadas com  $n$  folhas tais que a raiz tenha exatamente dois filhos. Usando  $H(t)^3$ , contamos o número de árvores cujas raízes têm três filhos etc.

A série infinita

$$H(t)^2 + H(t)^3 + \dots$$

conta todas as árvores, exceto a única que possui apenas uma folha. Consequentemente, essa soma infinita é  $H(t) - t$ .

Provamos que:

$$H(t) - t = H(t)^2 + H(t)^3 + \dots$$

Somando a série geométrica, obtemos:

$$H(t) - t = \frac{H(t)^2}{1 - H(t)}$$

ou

$$2H(t)^2 - (1 + t)H(t) + t = 0,$$

que fornece:

$$H(t) = \sum_{n=1}^{\infty} b(n)t^n = (1 + t - \sqrt{1 - 6t + t^2})/4.$$

Como uma função de uma variável complexa,  $(1 + t + \sqrt{1 - 6t + t^2})/4$  é bem definida e a holomorfa no disco de centro 0 cujo raio é a menor das duas raízes de  $1 - 6t + t^2 = 0$ , que é  $t = 3 - 2\sqrt{2}$ . O raio de convergência de  $H(t)$  é, portanto,  $3 - 2\sqrt{2}$ . Em outras palavras

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \log b(n) = \log(3 + 2\sqrt{2}).$$

O leitor vai facilmente mostrar que o  $\limsup$  pode ser substituído por um  $\lim$ . Os  $a(n)$  são os *grandes números de Schroeder*, e



E. SCHRÖDER.

Friedrich Wilhelm Karl Ernst Schroeder (1841–1902) tinha muitos passatempos esportivos: andar de bicicleta, caminhar, nadar, patinar no gelo, cavalgar e cuidar do jardim. Como ele era sempre visto andando com sua bicicleta por Karlsruhe, ele era conhecido localmente como o 'Professor bicicleta' (veja [MacTutor History of Mathematics archive](#)). +



Hiparco  
(c. 190 – c. 120 ac.) +

os  $b(n)$  são os *pequenos números de Schroeder*. Não se esqueça que  $a(n) = 2b(n)$  para  $n \geq 2$ .

A incrível *Online Encyclopedia of Integer Sequences* tem várias [páginas](#) dedicadas a essas duas sequências (dentre muitas outras páginas) e contém provavelmente informações demais! Por exemplo, pode-se encontrar a estimativa assintótica:

$$a(n) \sim \frac{(3 + 2\sqrt{2})^n}{2n\sqrt{2\pi n}\sqrt{3\sqrt{2} - 4}\left(1 - \frac{9\sqrt{2}+24}{32n} + \dots\right)}.$$

### *Hiparco e Schroeder*

O *Princípio de Arnold* afirma que

Se uma noção tem um nome próprio, não é o do seu criador

e seu complementar, o *Princípio de Berry*:

O princípio de Arnold aplica-se a ele próprio<sup>13</sup>.

Isso se aplica em particular à descoberta dos números de Schroeder. Ernst Schroeder foi um lógico alemão importante. Ele explicava que seu objetivo era:<sup>14</sup>:

Transformar a lógica em uma disciplina do cálculo, especialmente possibilitar o trabalho exato com conceitos relativos, para, daí em diante, libertar-se dos marcos rotineiros da linguagem falada, libertar-se também dos clichês filosóficos, preparando as bases de uma língua internacional com fundamento científico, que se pareça mais com uma linguagem de sinais do que com uma linguagem sonora.

Considerando seu ponto de vista sobre a lógica, para ele, era uma questão muito natural contar o número de maneiras apropriadas de se pôr parênteses em uma palavra de tamanho  $n$ . Esse é o propósito de seu artigo de 1870<sup>15</sup>.

Para uma palavra de tamanho 2, temos duas possibilidades:

$$ab \text{ e } (ab).$$

Seis possibilidades para uma palavra de tamanho 3:

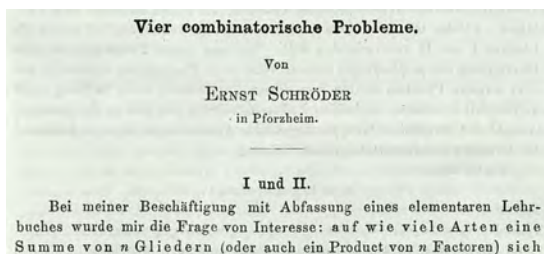
$$\begin{array}{ccc} abc & (ab)c & a(bc) \\ (abc) & ((ab)c) & (a(bc)). \end{array}$$



<sup>13</sup> V. I. Arnold. [On teaching mathematics](#).

<sup>14</sup> V. Peckhaus. 19th century logic between philosophy and mathematics. *Bull. Symbolic Logic*, 5(4):433–450, 1999.

<sup>15</sup> E. Schröder. [Vier combinatorische probleme](#). *Bull. Symbolic Logic*, 15:361–376, 1870.



As regras do jogo são que uma letra não pode ficar individualmente dentro de parênteses como  $(a)$  e não se pode duplicar parênteses como  $((ab))$ . A palavra completa pode estar dentro de parênteses ou não (e é por isso que os grandes números de Schroeder são pares). Note que um par de parênteses pode cercar mais de duas letras. Há 22 possibilidades para uma palavra de tamanho 4.

Deve estar claro para o leitor que essas 22 expressões não são nada além da lista de 22 árvores podadas com quatro folhas. De fato, palavras com parênteses podem ser descritas por árvores podadas, como mostrado na margem. Schroeder estava simplesmente contando árvores podadas, ou intercâmbios polinomiais, ou permutações separáveis. Seu artigo contém a relação de recorrência e a função geradora descritas acima.

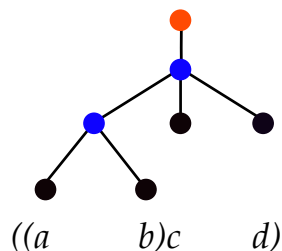
Em 1994, David Hough, um aluno de pós-graduação da George Washington University (USA), estava lendo o exercício 1.45 do livro do Stanley <sup>16</sup>:

A seguinte citação é da obra de Plutarco *Conversa de Mesa* VIII.9.732 'Crisipo diz que o número de proposições compostas que podem ser feitas a partir de somente dez proposições simples excede um milhão'. (Hiparco, para ter certeza, refutou essa afirmação, mostrando que do lado afirmativo há 103 049 proposições compostas, e do lado negativo 310,952.

Segundo Y. Heath, *A History of Greek Mathematics*, vol 2, p. 245; "parece impossível entender o significado desses números." [Heath também nota que uma leitura alternativa de 103 049 é 101 049.]

Será possível dar algum sentido à afirmação de Plutarco?

- $abcd$
- $(ab)cd$
- $a(bc)d$
- $(ab)(cd)$
- $(ab)(cd)$
- $(abc)d$
- $a(bcd)$
- $((ab)c)d$
- $(a(bc))d$
- $a((bc)d)$
- $(a(b(cd)))$
- $(abcd)$
- $((ab)cd)$
- $(a(bc)d)$
- $((ab)(cd))$
- $((ab)(cd))$
- $((ab)c)d$
- $((ab)c)d$
- $((ab)c)d$
- $((a(bc))d)$
- $((a(bc))d)$
- $(a((bc)d))$
- $(a((bc)d))$



<sup>16</sup> R. P. Stanley. *Enumerative combinatorics. Vol. I. The Wadsworth & Brooks/Cole Mathematics Series.* Wadsworth & Brooks/Cole Advanced Books & Software, Monterey, CA, 1986. Com um prólogo por Gian-Carlo Rota.

Hough observou que 103 049 é o décimo pequeno número de Schroeder  $b(10)$ , fato que não poderia ser mera coincidência.

Plutarco foi um historiador e biógrafo Grego, cujo papel (na nossa história) se limita à transcrição de uma citação de Hiparco feita duzentos anos mais cedo. É difícil imaginar que esse número 103 049 pôde ser lembrado por tanto tempo sem estar preservado em um livro na posse de Plutarco.

Hiparco é provavelmente o astrônomo mais importante da Antiguidade. Ele é conhecido por ter descoberto a precessão dos equinócios e, acima de tudo, por ter construído uma descrição científica coerente do movimento dos planetas. Seu sucessor, Ptolomeu, trezentos anos mais tarde, é famoso pelo *Modelo geocêntrico Ptolomaico* que se tornou o dogma astronômico até Copérnico introduzir o modelo heliocêntrico, muitos séculos mais tarde. Ptolomeu deve muito a Hiparco e nem sempre reconheceu essa sua dívida. Mas isso não tem a ver com a nossa história.

Então, segundo Hough, Hiparco, sob a transcrição de Plutarco, estava contando palavras com parênteses de tamanho 10. Sobre essa descoberta dos números de Schroeder por Hiparco, diversos artigos históricos foram escritos <sup>17 18</sup>.

Um [artigo](#) <sup>19</sup> no [MathPages](#) dá uma explicação um pouco mais elaborada em termos de lógica estoica (uma lógica pré-Aristotélica ensinada em particular por Crisipo, e criticada por Hiparco).

Dado um certo número de asserções lógicas  $a_1, a_2, \dots, a_k$  há pelo menos duas maneiras de combiná-las por conjunção ou disjunção:

- $a_1$  OU  $a_2$  OU  $\dots$  OU  $a_n$ , que é uma função  $n$ -ária  $OU(a_1, \dots, a_n)$ ,
- $a_1$  E  $a_2$  E  $\dots$  E  $a_n$ , que é uma função  $n$ -ária  $E(a_1, \dots, a_n)$ .

Na notação booleana moderna, usa-se + para OU e um ponto ou concatenação para E. Agora, considere uma palavra de tamanho  $n$  (por exemplo  $abcd$  para  $n = 4$ ). Para cada um dos  $n - 1$  espaços entre letras, escolha um "+" ou um ".". Temos  $2^{n-1}$  possibilidades (8 no nosso exemplo).

Nós estamos acostumados a dar prioridade à multiplicação



Ta sage instruction fert de riche couronne  
A Trajan, élevé par dessus tous humains.  
Si les grands te porcioient au cœur & dans leurs mains,  
Verru viuroit au lieu de Venus & Bellone

Plutarco  
(c. AD 46 – AD 120).

†

<sup>17</sup> R. P. Stanley. [Hipparchus, Plutarch, Schröder, and Hough](#). *Amer. Math. Monthly*, 104(4):344–350, 1997.

<sup>18</sup> F. Acerbi. On the shoulders of Hipparchus: a reappraisal of ancient Greek combinatorics. *Arch. Hist. Exact Sci.*, 57(6):465–502, 2003.

<sup>19</sup> K. Brown. Hipparchus on compound statements.

$abcd$	$abc + d$
$ab + cd$	$ab + c + d$
$a + bcd$	$a + bc + d$
$a + b + cd$	$a + b + c + d$

sobre a soma, mas se quisermos especificar uma ordem para avaliar essa função lógica, devemos colocar parênteses.

Essa construção na margem pode ser descrita por uma árvore podada, associando o símbolo OU a um vértice de nível ímpar e E a um vértice de nível par. Dessa forma, cada vértice atua no conjunto de seus filhos, ficando esses filhos entre parênteses. Se a raiz tiver um único filho, não é necessário rotulá-la com E, pois ela age sobre um conjunto unitário.

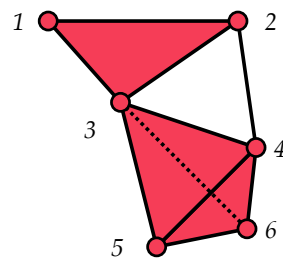
Como um exercício, o leitor é encorajado a mostrar que duas expressões diferentes, i.e., duas árvores podadas diferentes, definem duas funções booleanas diferentes  $\{0, 1\}^n \rightarrow \{0, 1\}$  quando avaliadas em  $a_i = 0$  ou 1 (falso ou verdadeiro).

Portanto, Hiparco estava certo: há  $a(10) = 2 \times 103049$  maneiras de se combinar 10 asserções usando OU ou E, no sentido recém-descrito. Alguém poderia se perguntar por que ele mencionou  $b(10)$  e não  $a(10)$ . Será que ele percebeu a involução natural entre *proposições compostas* dada pela negação, que basicamente permuta E com OU?

Existe uma questão em aberto relacionada, o [Problema de Dedekind](#). Há  $2^{2^n}$  funções booleanas, ou seja, funções de  $\{0, 1\}^n$  para  $\{0, 1\}$ . É fácil constatar que qualquer função booleana pode ser escrita como uma fórmula usando OU, E e NÃO. Tais funções descritas por fórmulas que não envolvem NÃO são chamadas de *funções booleanas monótonas* (mas observe que nós não impomos que cada variável apareça apenas uma vez na fórmula, como no caso de Hiparco). *A questão de calcular o número de funções booleanas monótonas está em aberto para  $n > 8$ .* Esse número também tem uma interpretação topológica interessante: ele é o número de complexos simpliciais cujos vértices são  $\{1, 2, \dots, n\}$ .

A maioria dos matemáticos, eu inclusive, tem uma ideia ingênua da matemática grega. Nós acreditamos que ela consiste apenas de Geometria, no espírito de Euclides. O exemplo do cálculo feito por Hiparco do décimo número de Schroeder pode ser uma pista de que os gregos antigos desenvolveram uma compreensão elaborada da combinatória, justamente o tema do artigo do Acerbi anteriormente citado.

$abcd$	$abc + d$
$ab(c + d)$	$a(bc + d)$
$ab + cd$	$a(b + c)d$
$a(b + cd)$	$(ab + c)d$
$ab + c + d$	$a(b + c) + d$
$a(b + c + d)$	$a + bcd$
$(a + b)cd$	$(a + bc)d$
$a + bc + d$	$(a + b)(c + d)$
$(a + b)c + d$	$a + b(c + d)$
$a + (b + c)d$	$a + b + cd$
$(a + b + c)d$	$a + b + c + d$



Um complexo simplicial com conjunto de vértices  $V = \{1, \dots, 6\}$ . A função associando 0 a um subconjunto  $X \subset V$  se  $X$  for um simplexo, e 1 caso contrário, é uma função booleana monótona.

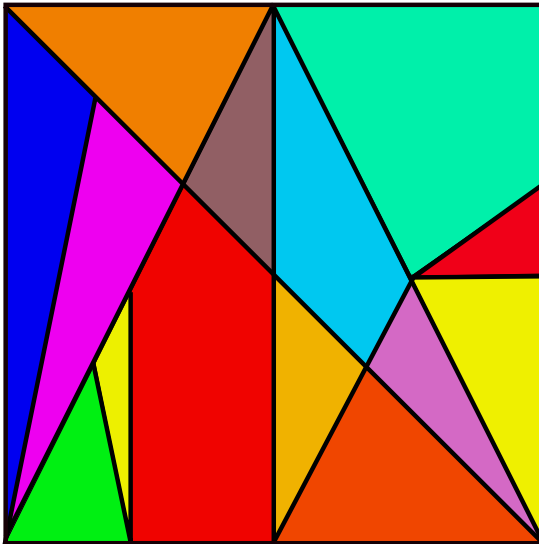
O livro escrito por Netz <sup>20</sup> traz novas perspectivas a essa história. O primeiro capítulo discute combinatória grega e em particular o número 103 049. Ele também contém a descrição de um outro enigma combinatório, que foi encontrado no famoso palimpsesto de Arquimedes. O leitor é instigado a ler e decifrar <sup>21</sup> como uma história de detetive. O jogo é composto de 14 peças poligonais e é parecido com o *Tangram*.

Netz “pediu ajuda a seu colega em Stanford, Persi [Diaconis], um notável combinatorialista, para resolver o que ele acreditava ser uma questão simples: de quantas maneiras é possível reconstruir o quadrado? [...] Diaconis precisou de alguns meses de trabalho colaborativo com três colegas para obter o número de soluções: 17 142.”

Será que Arquimedes sabia a resposta?

<sup>20</sup> R. Netz. *Ludic proof, reek mathematics and the Alexandrian aesthetic*. Cambridge UP, 2009.

<sup>21</sup> R. Netz and W. Noel. *The Archimedes codex, Revealing the secrets of the world's greatest palimpsest*. Phoenix, London, 2008.



Stomachion de Arquimedes.



THE  
METHOD of FLUXIONS  
AND  
INFINITE SERIES;

WITH ITS  
Application to the Geometry of CURVE-LINES.

---

By the INVENTOR  
*Sir ISAAC NEWTON, K<sup>t</sup>.*  
Late President of the Royal Society.

---

*Translated from the AUTHOR'S LATIN ORIGINAL  
not yet made publick.*

---

To which is subjoin'd,  
A PERPETUAL COMMENT upon the whole Work,  
Consisting of  
ANNOTATIONS, ILLUSTRATIONS, and SUPPLEMENTS,  
In order to make this Treatise  
*A compleat Institution for the use of LEARNERS.*

---

By *JOHN COLSON*, M. A. and F. R. S.  
Master of *Sir Joseph Williamson's* free Mathematical-School at *Rockester*.

---

L O N D O N :  
Printed by HENRY WOODFALL ;  
And Sold by JOHN NOURSE, at the *Lamb* without *Temple-Bar*.  
M.DCC.XXXVI.

*Adam*  
*83.12*

*275*

# *De methodis serierum et fluxionum*

## *O método de Newton*

### *Curvas algébricas*

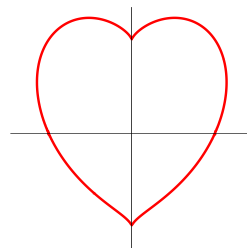
DESDE A INTRODUÇÃO DE COORDENADAS POR RENÉ DESCARTES, o estudo de curvas planas e, especialmente, de *curvas planas algébricas* (definidas por alguma equação polinomial  $P(x, y) = 0$ ), tornou-se e continua a ser um tema central na matemática. Evidentemente, equações de grau 1 e 2 (retas e cônicas) eram bem conhecidas. Quando matemáticos do século XVIII olharam para curvas de grau maior, encontraram uma selva de diversas formas, as quais eles tentaram domar. Isaac Newton, por exemplo, escreveu um longo tratado sobre curvas de grau 3, decompondo-as num grande número de “espécies”. Veja por exemplo as discussões <sup>22</sup> <sup>23</sup>.

Rapidamente, tornou-se claro que os pontos singulares desempenhavam um papel central na geometria dessas curvas. Um ponto  $(x_0, y_0)$  é *singular* se ele está na curva, i.e. se  $P(x_0, y_0) = 0$ , e se as derivadas parciais  $\partial P/\partial x$  e  $\partial P/\partial y$  se anulam em  $(x_0, y_0)$ . Numa vizinhança de um ponto *regular* (i.e. não-singular), um matemático moderno não teria problemas em aplicar o teorema da função implícita: em coordenadas suaves apropriadas no entorno de tal ponto, a curva é uma linha reta. Pontos singulares, no entanto, podem ser muito mais complicados e decifrar sua natureza levou um longo tempo.

Neste capítulo, vou descrever um dos passos mais importan-

<sup>22</sup> J. Stillwell. *Mathematics and its history*. Undergraduate Texts in Mathematics. Springer, New York, third edition, 2010.

<sup>23</sup> W. W. R. Ball. On Newton’s Classification of Cubic Curves. *Proc. London Math. Soc.*, S1-22(1):104–143, 1890.



Uma curva algébrica.

tes na direção de entender essa natureza, seguindo o livro de Newton *De methodis serierum et fluxionum*.

Não quero entrar em nenhum detalhe histórico sobre a rivalidade entre Newton e Leibniz em relação à invenção do cálculo. Deixe-me recomendar, especificamente para nossos propósitos, a excelente biografia de Newton por Westfall <sup>24</sup>.

Essa é a descrição do próprio Newton sobre seu *Annus mirabilis* <sup>25</sup>:

*No início do ano de 1665, descobri o método de aproximação a uma série desse tipo e a regra para reduzir qualquer potência de qualquer binômio a tal série. No mesmo ano, em maio, descobri o método das tangentes de Gregory e Slusius; em novembro, obtive o método direto dos fluxões; no ano seguinte, em janeiro, a teoria das cores; no maio posterior desvendi o método inverso dos fluxões. No mesmo ano, comecei a pensar na gravidade como se estendendo até a órbita da Lua [...]. Tudo isso foi nos dois anos da peste, 1665–1666. Pois, nessa época, eu estava no auge de minha fase de invenção, interessava-me mais pela matemática e pela filosofia do que em qualquer ocasião posterior.*

Em julho de 1669, baseado nas suas ideias de 1665, Newton escreveu *De Analysi per aequationes numero terminorum infinitas*.

Em 1671, escreveu *De methodis serierum et fluxionum*, mas não o publicou.

Em 1676, escreveu duas famosas cartas à Leibniz (por meio de Oldenburg como intermediário): *Epistola prior* e *Epistola posterior*.

A [tradução](#) inglesa (por Colson) do *De methodis* apareceu em 1736 (nove anos após a morte de Newton). Uma [tradução](#) francesa da tradução inglesa, por Buffon, apareceu em 1740.

Todos esse artigos contêm uma descrição um tanto precisa de pontos singulares em curvas algébricas, em termos do que hoje é chamado de “séries de Puiseux”, seguindo novamente o princípio de Arnold.

Queremos estudar a curva  $P(x, y) = 0$ , onde  $P$  é um polinômio com *coeficientes complexos*. Devemos entender primeiramente que Newton não pensa nisso como uma “curva”, mas sim como uma “função”: dado  $x$ , ele quer resolver a equação  $P(x, y) = 0$  e encontra  $y$  como uma função  $y(x)$ . Seu resultado principal é que esse fato é possível, contanto que se esteja disposto a considerar

<sup>24</sup> R. S. Westfall. *Never at rest, a biography of Isaac Newton*. Cambridge UP, 1980.

<sup>25</sup> R. S. Westfall. [Newton’s marvelous years of discovery and their aftermath: myth versus manuscript](#). *Isis*, 71(256):109–121, 1980.

In the beginning of the year 1665 I found the Method of approximating series and the Rule for reducing any dignity of any Binomial into such a series. The same year in May I found the method of Tangents of Gregory and Slusius, and in November had the direct method of fluxions and the next year in January had the theory of Colours and in May following I had entrance into ye inverse method of fluxions. And the same year I began to think of gravity extending to ye orb of the Moon [...]. All this was in the two plague years of 1665–1666. For in those days I was in the prime of my age for invention and minded Mathematicks and Philosophy more then at any time since.

Aparentemente Colson não aceitou mostrar os manuscritos em latim para Buffon!

Posteriormente, estudaremos o caso com coeficientes reais assim como funções  $P$  que são apenas analíticas.

$y(x)$  como sendo uma série infinita de potências racionais de  $x$ . Deixe-me enunciar um teorema que tornaremos preciso posteriormente, e que Newton “quase” provou.

**Teorema.** Qualquer equação polinomial  $P(x, y) = 0$  (onde  $P$  não é divisível por  $x$ ) tal que  $P(0, 0) = 0$  é equivalente, numa vizinhança de  $(0, 0)$ , a um número finito de igualdades da forma  $y = f_i(x)$  (com  $i = 1, \dots, n$ ) onde  $f_i$  é uma “série de Puiseux” da forma:

$$f_i(x) = \sum_{k=1}^{\infty} a_{i,k} x^{\frac{k}{m_i}}$$

para determinados coeficientes complexos  $a_k$  e inteiros positivos  $m_i$ .

Em outras palavras,  $\{P = 0\}$  é a união de um número finito de “gráficos” de séries  $f_i$ . Estamos numa posição similar a da pergunta original de Kontsevich, sendo natural nos perguntarmos qual é a natureza topológica desses gráficos. No entanto, antes de estudarmos essa questão, temos muitos detalhes para consertar, dado que, em particular, essas  $f_i$  não são “funções” de fato. Pense, por exemplo, no “gráfico” da raiz quadrada.

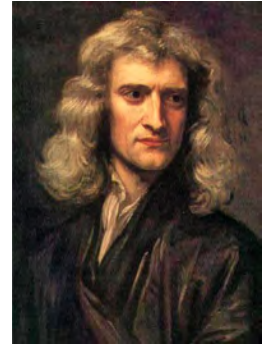
Convido o meu leitor a olhar mais de perto para a primeira parte do *De methodis serierum et fluxionum*. A folha de rosto desse importante livro está na primeira página deste capítulo. Para simplificar a tarefa para você, leitor, seguirei a tradução inglesa.

## O método de Newton

Começemos a ler Newton.

*Visto que há uma enorme conformidade entre as Operações em Espécies, e as mesmas Operações em Números comuns; e que tampouco elas parecem diferir, exceto nos símbolos pelos quais são representadas, o primeiro sendo geral e indefinido, e o outro definido e particular: não posso deixar de perguntar-me como ninguém pensou em acomodar a recentemente descoberta Doutrina de Frações Decimais de maneira análoga às Espécies [...] especialmente pela possibilidade de abrir um caminho para Descobertas mais obscuras.*

*Explicação:* Por “número comum”, Newton quer dizer... um número comum, isto é, o que hoje chamamos de um número



Newton em 1689, por Godfrey Kneller.



Uma página do *De methodis*.

Since there is a great conformity between the Operations in Species, and the same Operations in common Numbers; nor do they seem to differ, except in the Characters by which they are represented, the first being general and indefinite, and the other definite and particular: I cannot but wonder that no body has thought of accommodating the lately discover'd Doctrine of Decimal Fractions in like manner to Species [...] especially since it might have open'd a way to more abstruse Discoveries.

*complexo*. Note que pouquíssimos matemáticos naquela época considerariam esses números “comuns”. Por “espécies”, ele quer dizer um polinômio em  $x$ , ou uma série inteira, ou o que chamamos hoje de uma “série de Laurent”, ou talvez uma “série de Puiseux”, i.e. uma série de potências racionais de  $x$ . Em qualquer caso, nas palavras de Newton uma “espécie” é algum tipo de função.

*Mas dado que essa Doutrina de Espécies tem a mesma relação com a Álgebra que a Doutrina de Números Decimais tem com a Aritmética comum; as Operações de Adição, Subtração, Multiplicação, Divisão, e Extração de Raízes, podem ser facilmente deduzidas dessa, caso o Estudante seja ao menos habilitado na Aritmética Decimal, e na Álgebra Vulgar, e observe a correspondência que se obtém entre Frações Decimais e Termos Algébricos indefinidamente contínuos.*

*Explicação:* Newton observa que se pode manipular séries do mesmo modo que se faz com números, para os quais se têm quatro operações (+, −, ×, ÷). Na terminologia moderna, ele observa que “números comuns” e “séries de Laurent” são ambos corpos.

*Pois como em Números, os Lugares à direita decrescem continuamente em Proporção Decimal ou Subdécupla; e o mesmo acontece respectivamente em Espécies, nos quais os Termos são dispostos (como imposto no que se segue), numa Progressão uniforme indefinidamente continuada, de acordo com a Ordem das Dimensões de qualquer Numerador ou Denominador.*

*Explicação:* Novamente na terminologia anacrônica moderna, Newton nos fala sobre a topologia desses dois corpos. Dois números reais estão próximos se suas expansões decimais concordam até uma ordem grande, analogamente, dois polinômios em  $x$ , ou duas séries, estão próximos na vizinhança de 0 se a valuação da sua diferença é grande.

A essa altura, podemos adivinhar a estratégia de Newton. Ele nos ensinará um jeito de resolver equações polinomiais  $P(x) = 0$  sendo  $P$  um polinômio em algum corpo, que pode consistir em “números comuns” ou em “espécies”. Isso se aplicará, portanto, a equações da forma  $P(x, y) = 0$  sendo  $P$  um polinômio em duas variáveis, visto como um polinômio  $P(x)(y)$  em uma variável

But since this Doctrine of Species, has the same relation to Algebra, as the Doctrine of Decimal Numbers has to common Arithmetick; the Operations of Addition, Subtraction, Multiplication, Division, and Extraction of Roots, may easily be learned from thence, if the Learner be but skill'd in Decimal Arithmetick, and the Vulgar Algebra, and observes the correspondence that obtains between Decimal Fractions and Algebraick Terms infinitely continued.

For as in Numbers, the Places towards the right-hand continually decrease in a Decimal or Subdecuple Proportion; so it is in Species respectively, when the Terms are disposed, (as is often enjoyn'd in what follows) in an uniform Progression on infinitely continued, according to the Order of the Dimensions of any Numerator or Denominator.



Capa da tradução francesa por Buffon.

$y$  com coeficientes no corpo das funções racionais  $\mathbb{C}(x)$ , ou no corpo  $\mathbb{C}((x))$  das séries de Laurent.

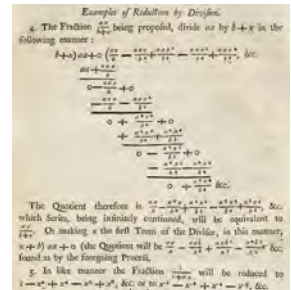
Numa apresentação bastante pedagógica, Newton apresenta diversos exemplos, mostrando a analogia entre “espécies” e “números comuns”. Primeiro, ele mostra como expandir  $a^2/(b-x)$  em uma série em  $x$ . Exatamente do mesmo modo que ele explicaria numa escola primária que dividir 1 por  $0.9 = 1 - 0.1$  dá  $1.11111111\dots$ . Isso é fácil e deve também ter sido fácil para seus leitores.

Então, ele explica o significado de expoentes racionais  $x^{p/q}$ , que também deveria ser familiar para a maioria de seus leitores. Ele pode, então, apresentar sua famosa fórmula binomial para uma potência racional de  $(x+a)$ , como uma série infinita em  $x$ .

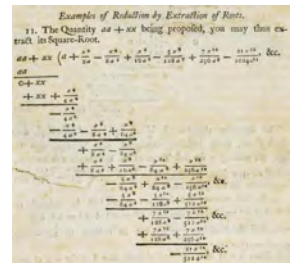
$$(a+x)^{\frac{p}{q}} = a^{\frac{p}{q}} + (p/q)a^{\frac{p}{q}-1}x + \frac{1}{2}(p/q)(p/q-1)a^{\frac{p}{q}-2}x^2 + \dots$$

Chegamos agora na parte mais interessante para nós.

Ele gostaria de resolver o que ele chama de “equações afetadas”, isto é, equações polinomiais cujos coeficientes são Espécies, ou seja, equações  $P(x,y) = P(x)(y) = 0$ . Novamente, de modo bastante pedagógico, ele declara que começará por resolver equações *ordinárias* em números comuns da forma  $P(y) = 0$ , sendo  $P$  um polinômio em  $\mathbb{C}[y]$ .



+



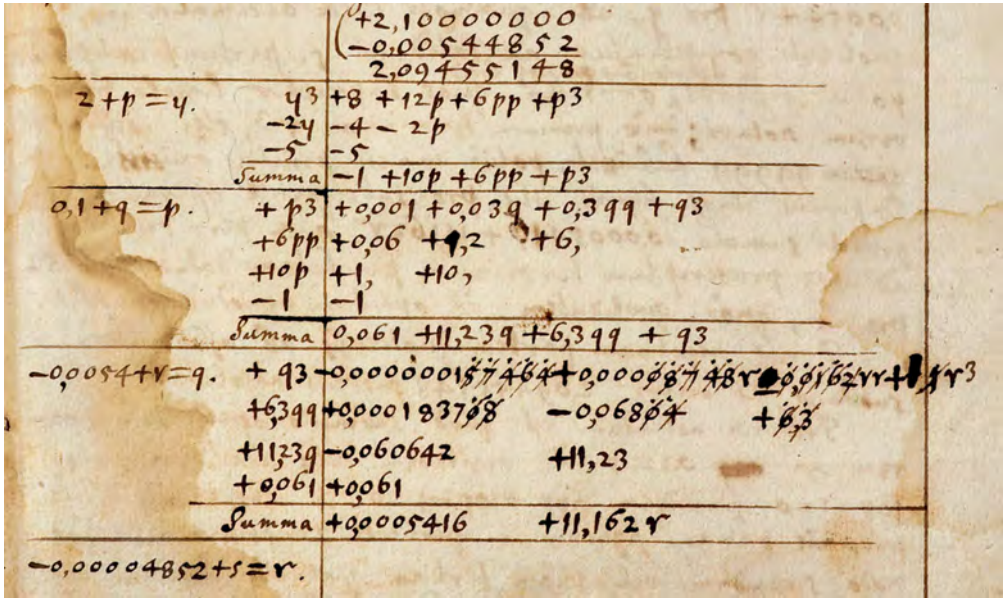
+

*Of the Reduction of affected Equations.*

19. As to affected Equations, we must be something more particular in explaining how their Roots are to be reduced to such Series as these; because their Doctrine in Numbers, as hitherto deliver'd by Mathematicians, is very perplexed, and incumber'd with superfluous Operations, so as not to afford proper Specimens for performing the Work in Species. I shall therefore first shew how the Resolution of affected Equations may be compendiously perform'd in Numbers, and then I shall apply the same to Species.

+

Esse é o famoso *método de Newton*, uma das ferramentas mais fundamentais em análise.



Observe para o modo como ele apresenta os cálculos. Seu exemplo é a equação cúbica

$$y^3 - 2y - 5 = 0.$$

Pode-se ver, por tentativa e erro, que existe uma raiz que não difere muito de 2. Por esse motivo, ele olha para  $y$  como  $y = 2 + p$ , com um  $p$  pequeno. Substituindo na equação original, ele encontra

$$p^3 + 6p^2 + 10p - 1 = 0.$$

Ele pode então “rejeitar  $p^3 + 6p^2$  pelo fato dele ser muito pequeno” e obter

$$10p - 1 \approx 0$$

de modo que  $p$  está perto de  $1/10$ . Ele, então, põe  $p = 0.1 + q$  e substitui na equação

$$q^3 + 6.3q^2 + 11.23q + 0.061 = 0$$

e “como  $11.23q + 0.061 = 0$  está próximo da verdade” ele sabe que  $q$  está próximo de  $-0.061/11,23 \approx -0.0054$ . Escrevendo  $q =$

$-0.0054 + r$ , ele pode substituir como antes e continuar com a operação “tanto quanto eu queira”. Ele finalmente encontra a solução perto de 2.09455148.

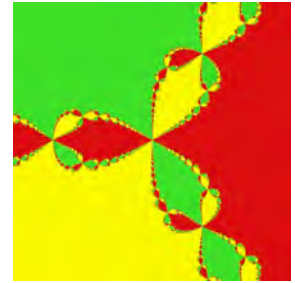
O próximo parágrafo mostra, se necessário, que Newton era incrivelmente dotado para cálculos. “O trabalho pode ser bastante abreviado” de fato. Ele explica o que todo estudante universitário sabe (ou deveria saber): que a cada passo o número de casas decimais corretas é essencialmente dobrado e que portanto não é necessário computar exatamente os  $p, q, r, s$  etc. É por isso que, em sua tabela, alguns dígitos estão cortados: isso não é um erro; é uma simplificação inteligente!

Em 1690, Joseph Raphson (1648–1715) (membro da Royal Society e, portanto, conhecendo Newton muito bem) publicou um método para resolver equações em *Analysis æquationum universalis*. Comece com uma solução aproximada  $y_0$  de  $P(y) = 0$  e considere a sequência definida por

$$y_{k+1} = y_k - \frac{P(y_k)}{P'(y_k)}$$

que, se tudo der certo, convergirá para a solução. Raphson não menciona Newton. Alguns historiadores alegam que os dois métodos são *bem* diferentes. No caso de Raphson, mantém-se a mesma equação e computa-se a sequência  $y_k$ . No caso de Newton, a cada passo computa-se uma nova equação. Os dois métodos dão exatamente o mesmo resultado e são formalmente idênticos mas, claramente, se os cálculos forem feitos à mão, a apresentação de Newton é muito mais eficiente. Poderíamos dizer que Raphson usa uma *iteração* e Newton uma *recursão* <sup>26</sup>. Alguns matemáticos alegam que Raphson entendeu o papel da derivada de  $P$  e que Newton estava apenas linearizando a equação. Bem, quem pode dizer que Newton, o inventor da derivada, não notou que a parte linear era a derivada? Como de costume, eu continuarei falando do *método de Newton* e não do método de Newton–Raphson.

Como um comentário final, é desnecessário dizer que Newton não discute nenhuma questão sobre convergência. Note também, leitor, que seu exemplo envolve apenas raízes reais de polinômios reais.



O método de Newton pode ser usado para encontrar raízes de polinômios  $P(z)$  com coeficientes complexos. Começando com  $z_{init}$ , esperamos que a iteração do algoritmo de Newton convergirá para uma raiz. Pode-se, então, decompor o plano (ou pelo menos o conjunto dos  $z_{init}$  para os quais o método funciona) em vários domínios, de acordo com a raiz limite. Em 1880, Cayley indagou sobre a descrição dessa decomposição, escrevendo que a questão é fácil para o grau 2 (exercício para o leitor) e que para grau 3 isso é “qualquer coisa exceto óbvio”. De fato, sabemos hoje que essa decomposição é de natureza fractal. Isso é conhecido como “o coelho de Newton”.

<sup>26</sup> C. Christensen. [Newton's method for resolving affected equations](#). *Collegè Math. J.*, 27(5):330–340, 1996.





Árvore da maçã de Newton no Trinity College. Às vezes circula um mito de que essa era a árvore da qual a maçã caiu sobre Isaac Newton. De fato, ele não estava em Cambridge durante o seu *Annus mirabilis*.

# *De methodis serierum et fluxionum*

## Séries de Newton

NEWTON AGORA PODE RESOLVER AS “EQUAÇÕES IMPLÍCITAS”, cujos coeficientes são *funções* de  $x$ . O exemplo sugerido por Newton é

$$y^6 - 5xy^5 + (x^3/a)y^4 - 7a^2x^2y^2 + 6a^3x^3 + b^2x^4 = 0.$$

Nessa equação,  $a, b$  são parâmetros. Note que Newton toma um grande cuidado para escrever equações *homogêneas*.

Por simplicidade, eu serei menos cuidadoso e escolherei  $a = b = 1$ . Newton enxerga inicialmente uma “solução aproximada” da forma  $y = ux^\alpha$  sendo  $u$  uma constante não nula e desconhecida e  $\alpha$  um número racional desconhecido.. Substituindo, obtemos

$$u^6x^{6\alpha} - 5u^5x^{1+5\alpha} + u^4x^{3+4\alpha} - 7u^2x^{2+2\alpha} + 6x^3 + x^4 = 0.$$

Essa é uma expressão envolvendo “monômios” com potências racionais de  $x$ . Os expoentes são  $6\alpha, 1 + 5\alpha, 3 + 4\alpha, 2 + 2\alpha, 3, 4$ . Se estudarmos essa situação numa vizinhança do 0, o maior termo corresponde ao menor desses expoentes. Para uma escolha genérica de  $\alpha$ , os seis expoentes correspondentes são diferentes. Caso desejarmos expressar o fato de que o termo dominante zera, nós poderíamos forçar e escolher  $u = 0$ , mas certamente não é o que queremos fazer. Assim sendo, temos que selecionar  $\alpha$  tal que pelo menos dois dos seis expoentes sejam iguais e, além disso, os menores possíveis. Newton expressa essa condição usando o seu famoso *polígono*, desenhando uma espécie de

Não esqueça que Newton foi também um grande físico!



Uma página do *De methodis*.

+

tabuleiro de damas subdividido em quadrados (que ele chama “paralelogramos”). Na equação original, para cada monômio não nulo  $a_{ij}x^jy^i$  ( $i, j \geq 0$ ), ele assinala uma estrela no paralelogramo  $(i, j)$ . Nesse exemplo, são marcadas seis estrelas.

Escolhendo  $\alpha$  e comparando os expoentes  $j + i\alpha$ , podemos dar uma interpretação geométrica, claramente explicada por Newton: Coloque uma régua no tabuleiro de xadrez e mova-a até que toque as estrelas marcadas.

*Então, quando qualquer equação é proposta, marca-se tal paralelogramo assim como todos os correspondentes termos, e usa-se a régua duas ou talvez mais vezes nos paralelogramos assim marcados, dos quais um é o mais baixo na coluna da esquerda em AB, e o outro tocando a régua pela direita; e todo o resto, não tocando a régua, situa-se acima dela. Em seguida, selecione os termos da equação que são representados pelos paralelogramos que tocam na régua.*

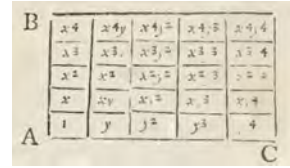
### Equações implícitas



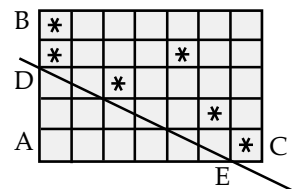
Então, nesse caso, o coeficiente  $\alpha$  é escolhido como sendo  $1/2$  (inclinação da reta  $DE$ ) e os três monômios dominantes  $x^3, x^2y^2$  and  $y^6$  são escolhidos. De fato, para  $\alpha = 1/2$ , a equação torna-se ordenada nas potências crescentes de  $x$ :

$$(u^6 - 7u^2 + 6)x^3 - 5u^5x^{7/2} + b^2x^4 + u^4x^5 = 0.$$

Por algum motivo, Newton marca os monômios  $x^j$  no eixo vertical e os  $y^i$  no eixo horizontal.



Then, when any Equation is proposed, mark such of the Parallelograms as correspond to all its Terms, and let a Ruler be apply'd to two, or perhaps more, of the Parallelograms so mark'd, of which let one be the lowest in the left-hand Column at AB, the other touching the right-hand; and let all the rest, not touching the Ruler, lie above it. Then select those Terms of the Equation which are represented by the Parallelograms that touch the Ruler.





Nós estamos, portanto, escolhendo  $u$  como uma solução da equação

$$u^6 - 7u^2 + 6 = 0$$

que contém três monômios, visto que a régua toca três estrelas. Existem seis soluções

$$u = \pm 1 \quad ; \quad \pm\sqrt{2} \quad ; \quad \pm\sqrt{-3}.$$

Newton parece ignorar as duas últimas soluções imaginárias. Talvez por estar interessado apenas pelas soluções reais, mas mesmo assim, se esse for o caso, é um engano, como veremos mais adiante nesta seção.

Então, ele escolhe a primeira solução e escreve  $y = \sqrt{x} + p$ , como no seu método com os “números comuns”. Pode-se, então, “continuar o processo com prazer”.

Entretanto, sem nenhuma explicação, Newton abandona seu primeiro exemplo e muda para outros casos numéricos, mostrando como o seu método funciona na prática.

Deixe-me mostrar como continuar o primeiro exemplo de Newton. Novamente, em nome simplicidade, altero ligeiramente sua apresentação.

Em vez de melhorar a primeira solução aproximada  $y \approx \sqrt{x}$  pela adição de uma variável  $p$ , vamos definir

$$x = x_1^2 \quad ; \quad y = x_1(1 + y_1).$$

Nós substituímos esses valores na equação original e simplificando por  $x_1^6$ . Obtemos:

$$\begin{aligned} & -5x_1 + x_1^2 + x_1^4 - 8y_1 - 25x_1y_1 + 4x_1^4y_1 \\ & + 8y_1^2 - 50x_1y_1^2 + 6x_1^4y_1^2 + 20y_1^3 \\ & - 50x_1y_1^3 + 4x_1^4y_1^3 + 15y_1^4 - 25x_1y_1^4 \\ & + x_1^4y_1^4 + 6y_1^5 - 5x_1y_1^5 + y_1^6 = 0. \end{aligned}$$

Nessa nova equação, os coeficientes de  $x_1$  e  $y_1$  são não nulos ambos. Então, a “régua de Newton” passa através do pontos  $(0,1)$  e  $(1,0)$ . Ressalto que essa é uma outra maneira de dizer que a nova equação não é singular na origem. Assim, os termos dominantes são lineares

$$-5x_1 - 8y_1$$

Em 24 de Outubro de 1676, Newton enviou uma [carta](#) para Leibniz, “descrevendo” a sua contribuição ao *calculus*. No final dessa carta, ele escreve “[...] problemas inversos sobre tangentes estão sobre o nosso poder, e outros mais difíceis que esses e, para resolvê-los, eu usei um método duplo de que uma parte é mais pura, a outra mais geral. Neste momento, julguei oportuno registrá-las em cartas transpostas...”, e depois ele esconde o seu método num anagrama:

```
5accdaei1oeffh11i4l3m9n6oqq
r8s1t9y3x:11ab3cdd1oeæg1
oill4m7n6o3p3q6r5s11t8vx,
3accæ4egh5i4l4m5n8oq4r3s6
t4v,aaddæceceiijmmnnoopr
rrssssttuu
```

Pobre Leibniz! Ele deve ter lutado para encontrar o significado do anagrama. Pobre leitor! Mesmo se eu desse a solução, ele terá que traduzi-la do latim para o seu próprio idioma e, então, entenderá que o conteúdo não é tão claro!

“Una methodus consistit in extractione fluentis quantitatis ex aequatione simul involvente fluxionem ejus: altera tantum in assumptione seriei pro quantitate qualibet incognita ex qua caetera commodè derivari possunt, et in collatione terminorum homologorum aequationis resultantis, as eruendis terminis assumptae seriei”.

Assim, segue que

$$y_1 \simeq -\frac{5}{8}x_1.$$

Continuando o processo, obtemos:

$$x_1 = x_2 \quad ; \quad y_1 = -\frac{5}{8}x_1(1 + y_2)$$

e assim por diante, nós poderíamos determinar uma expansão de  $y$  como o produto de  $\sqrt{x}$  e uma série de potências inteiras de  $x$ .

Com a ajuda do *Mathematica*, obtemos:

$$\begin{aligned} y(x) = & x^{1/2} - 5 \cdot 2^{-3}x + 79 \cdot 2^{-5}x^{3/2} - 14185 \cdot 2^{-10}x^2 \\ & + 3118083 \cdot 2^{-15}x^{5/2} - 189696965 \cdot 2^{-18}x^3 \\ & + 24625187405 \cdot 2^{-22}x^{7/2} - 1670815928565 \cdot 2^{-25}x^4 + \dots \end{aligned}$$

Para outra solução  $u = \sqrt{2}$ , obtemos:

$$\begin{aligned} y(x) = & \sqrt{2}x^{1/2} + 2x - 13\sqrt{2} \cdot 2^{-2}5^{-4}x^{3/2} + 3825 \cdot 2^{-2}x^2 \\ & - 267229\sqrt{2} \cdot 2^{-5}5^{-3}x^{5/2} + 903813 \cdot 2^{-1}5^{-4}x^3 \\ & - 1661176381\sqrt{2} \cdot 2^{-7}5^{-5}x^{7/2} + 777992628 \cdot 5^{-6}x^4 + \dots \end{aligned}$$

Agora um comentário final sobre a motivação de Newton: ele “provou” que qualquer “função”  $y(x)$  definida por alguma relação implícita  $P(x, y) = 0$  pode ser expandida como séries de potências de  $x$  (ao custo de usar expoentes racionais) desde que ela, é claro, conhecia muito bem derivadas e primitivas de qualquer potência  $x^\alpha$ , ele usa essa técnica para calcular derivadas e primitivas de quaisquer séries. Em outras palavras, ele permite determinar as derivadas e primitivas de “qualquer” função. O restante do *De methodis serierum et fluxionum* é dedicado a muitas aplicações desse método.

### *Um engano de Newton?*

É surpreendente perceber que Newton esqueceu uma raiz da equation

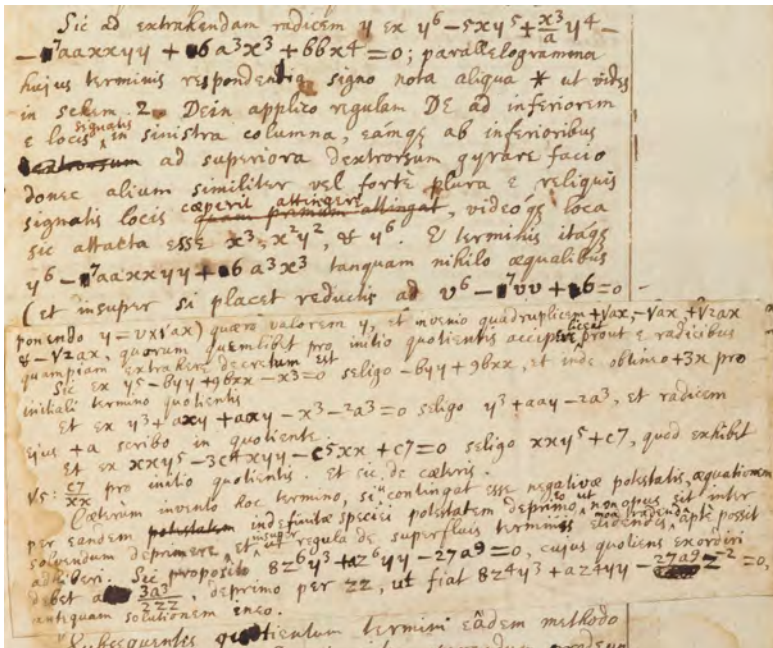
$$u^6 - 7u^2 + 6 = 0.$$

Pode-se acreditar que as raízes imaginárias  $\pm\sqrt{-3}$  conduzem a soluções imaginárias para  $y(x)$ . Mas isso não é assim, e eu acredito que foi realmente um engano.



Newton no momento preciso do erro? Como visto por Gotlib. (Homenagem a Gotlib). †

Em 2016, descobrir um engano em um trabalho importante escrito por Newton por volta de 1669 é uma experiência interessante! Olhando para o manuscrito original, descobre-se que Newton teve de corrigir um equívoco e colar um pedaço de papel acima da página original. Suspeito que a biblioteca do Trinity College não concordaria em retirar o precioso manuscrito para ver o que está por baixo. Deve ser necessário usar raios-X.



Um engano? +

<sup>27</sup> I. Newton. *The correspondence of Isaac Newton*, Vol. II: 1676–1687. Published for the Royal Society. Cambridge UP, New York, 1960.

Na verdade, poderia ser dada uma outra interpretação. No seu comentário da *Epistola posterior*, Turnbull <sup>27</sup> (nota 68, página 159) menciona outro “engano”: a “cobertura” do sinal da raiz quadrada não é suficientemente longo, e Newton escreveu equivocadamente  $\sqrt{2x}$  no lugar de  $\sqrt{2}x$ . Então, ele comenta que “Newton rejeita a raiz imaginária no caso dado por  $v^2 + 3 = 0$ ”. De fato, é possível que Newton tenha cometido um erro com  $\sqrt{2x}$ , o que o levou a pensar em  $\sqrt{-3x}$  como imaginário, para depois rejeitá-lo. Se ele tivesse escrito  $\sqrt{-3x}$ , ele teria visto que essa solução não é imaginária quando  $x < 0$ . Bem, mas isso nós nunca saberemos!

Na realidade,  $\pm\sqrt{-3}$  é imaginário, mas a função aproximada  $y \simeq \pm\sqrt{-3}\sqrt{x}$  é real se  $x$  é um *número real negativo*, então, ele não deveria ter sido descartado.

Para a raiz real  $u = 1$ , nós temos

$$x = x_1^2 \quad ; \quad y = x_1(1 + y_1).$$

Para a raiz imaginária  $\sqrt{-3}$ , obtém-se

$$x = -x_1^2 \quad ; \quad y = 3x_1(1 + y_1)$$

e nós procedemos como antes. Finalmente, consideramos a terceira raiz *real*.

$$\begin{aligned} y(x) = & -3^{1/2}(-x)^{1/2} - 9 \cdot 2^{-3}(-x) - 721 \cdot 2^{-6}5^{-1}3^{-1/2}(-x)^{3/2} - 36543 \cdot 2^{-10}5^{-2}(-x)^2 \\ & - 27986569 \cdot 2^{-15}3^{-3/2}5^{-3}(-x)^{5/2} - 96025589 \cdot 2^{-18}5^{-4}(-x)^3 \\ & + 169264391911 \cdot 2^{-22}3^{-5/2}5^{-5}(-x)^{7/2} + 1398151100829 \cdot 2^{-25}5^{-6}(-x)^4 + \dots \end{aligned}$$

Pode-se perguntar por que encontramos três soluções e não seis. Afinal, a equação  $u^6 - 7u^2 + 6 = 0$  possui, de fato, seis soluções. Entretanto, isso ocorre simplesmente porque as raízes opostas dão origem à mesma solução.

Não esqueça que Newton considera  $\sqrt{x}$  como uma função, tal que  $\sqrt{x}$  e  $-\sqrt{x}$  produzem “o mesmo”. Concordo que escrever  $\sqrt{x} = -\sqrt{x}$  pode levar a contradições, mas não na mão do grande Newton. Nós somos sábios para ensinar nossos alunos que  $\sqrt{x}$  é a raiz positiva de  $x$ , que é um número real e positivo, e para escolher alguma “determinação principal” para  $x \in \mathbb{C} \setminus \mathbb{R}_-$ .

Usando a terminologia moderna, poderíamos dizer que as curvas parametrizadas  $(t, t^2)$  and  $(-t, t^2)$  são as mesmas curvas com diferentes parametrizações.



Monumento a Newton na Abadia de Westminster.



Mão direita nos seus quatro livros principais.



Mão esquerda apontando para os anjos mostrando... a série binomial.

### O que Newton não provou

A definição de “convergência” não estava à disposição de Newton. No entanto, seus cálculos numéricos sugerem que as suas séries são convergentes e de fato ele as usa, até mesmo o vocábulo “convergente”. Para ser honesto, poderíamos dizer que ele somente mostra que suas séries fornecem *expansões assintóticas*. Na prática, as séries

$$a_1x^{\alpha_1} + a_2x^{\alpha_2} + \dots$$

(sendo  $\alpha_1 < \alpha_2 < \dots$  expoentes racionais) são *assintóticas* para a função  $f(x)$  se para todo  $k \geq 1$ :

$$f(x) - \sum_{k=1}^n a_k x^{\alpha_k} = \mathcal{O}(x^{\alpha_n}).$$

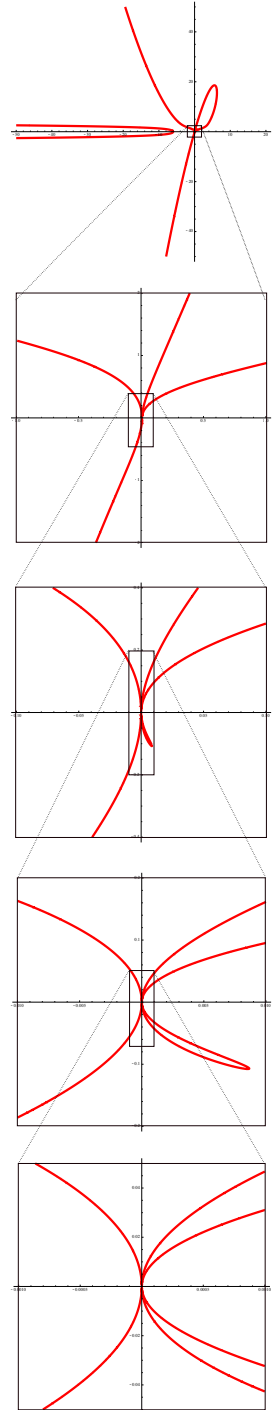
Isso não implica que  $f_n$  converge para  $f$ , mas é ainda frequentemente utilizada, porque às vezes é mais útil do que uma convergência habitual.

Outro aspecto que não é discutido corresponde à natureza dos expoentes racionais que aparece nessas séries. Em cada passo, um novo número racional se apresenta, não sendo claro se essa sequência dos expoentes converge para o infinito. Mas parece ser claro o fato de que todos os denominadores sejam *limitados*. Entretanto, Newton observa que o seu método não está restrito a equações polinomiais  $P(x, y) = 0$ , mas trabalha perfeitamente para “*æquationes numero terminorum infinitas*” da forma  $\sum_{i,j \geq 0} a_{ij} x^i y^j = 0$  (com  $a_{00} = 0$ ), envolvendo o que hoje chamamos de *funções analíticas*.

Para concluir esse capítulo, vamos olhar com atenção para a curva original de Newton.

$$P(x, y) = y^6 - 5xy^5 + x^3y^4 - 7x^2y^2 + 6x^3 + x^4 = 0.$$

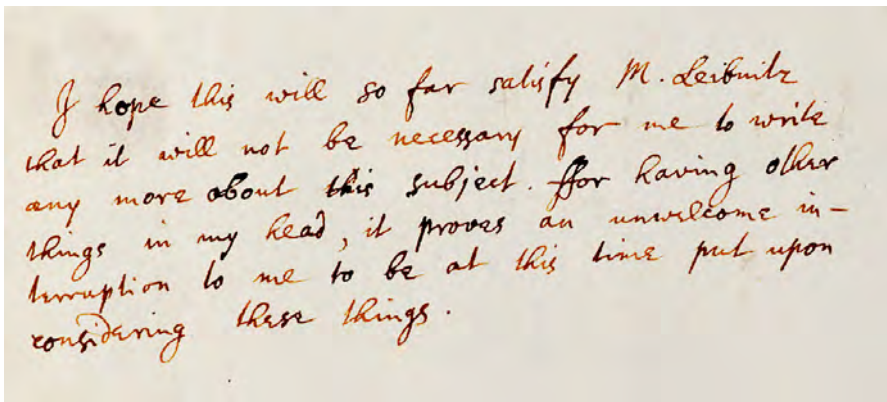
Se eu pedir para um computador plotar essa curva na caixa  $[-50, +20] \times [-50, +50]$ , ele fornecerá uma primeira curva na margem. Isto pode parecer surpreendente, pois vemos apenas dois ramos na vizinhança da origem. Se olharmos mais de perto, isto é, numa caixa menor, por exemplo,  $[-1, 1] \times [-2, 2]$  (segunda curva), podemos ver um segundo ramo. Ampliando ainda





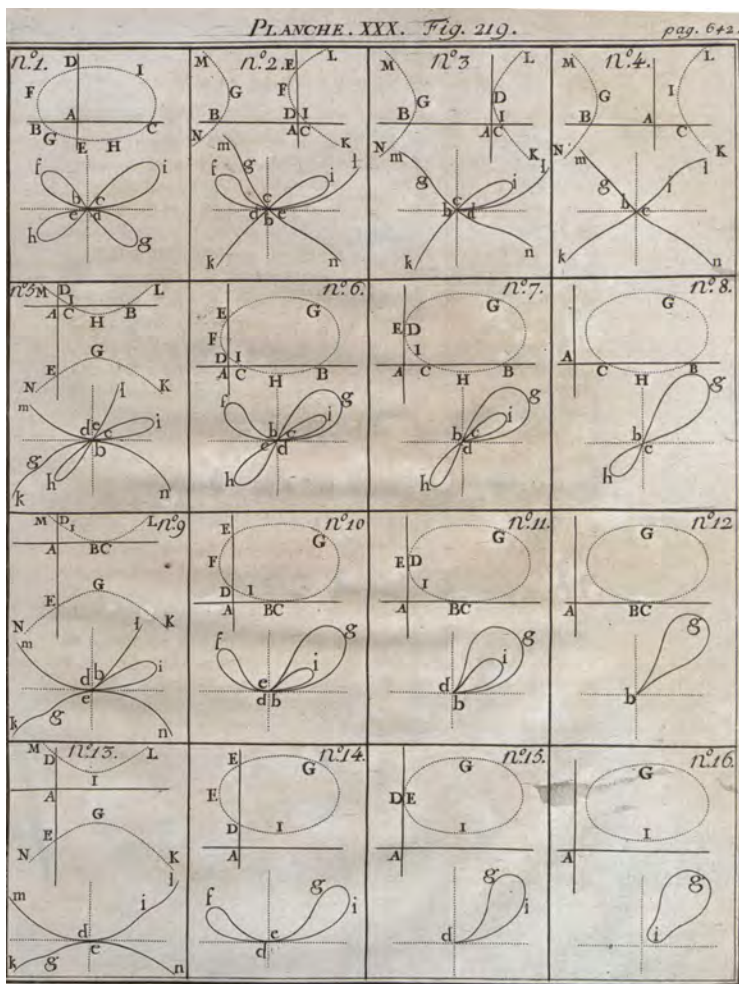
mais, olhando em  $[-.1, .1] \times [-.4, .4]$  (terceira curva), fica mais fácil de observar. A situação local é completamente clara em  $[-.01, .01] \times [-.2, .2]$  (quarta curva). Finalmente, em  $[-.001, .001] \times [-.05, .05]$ , podemos identificar três ramos assintóticos para  $\pm\sqrt{x}, \pm\sqrt{2x}, \pm\sqrt{-3x}$  como foi previsto por Newton (à exceção do terceiro, para  $x < 0$ , que ele se esqueceu).

Como esse polinômio  $P(x, y)$  é primo, ele não pode ser escrito, de modo não trivial, por meio do produto de dois polinômios. Entretanto, como série de potência convergente em  $x, y$ , na vizinhança da origem, pode ser escrito como um produto de três fatores, correspondendo aos três ramos.



I hope this will so far satisfy M. Leibnitz  
that it will not be necessary for me to write  
any more about this subject. For having other  
things in my head, it proves an unwelcome in-  
terruption to me to be at this time put upon  
considering these things.

Quando Newton questiona Oldenbourg para encaminhar sua Epistola Posterior para Leibniz, ele adiciona os seguinte P.S. Sim, ele tinha de fato outras coisas em sua cabeça.



Uma figura do livro de Cramer sobre curvas. †

# Um pouco de álgebra formal

A ÁLGEBRA NESTE CAPÍTULO SERÁ “FORMAL” uma vez que consideraremos séries formais.

## Encontrando uma solução

Repetirei os argumentos de Newton, expressando-os numa terminologia algébrica mais moderna. Atribuir tudo que segue à Newton exige imaginação e extrapolação infinitas. Enfatizaremos, no entanto, uma importante contribuição de Cramer. Geralmente, uma boa parte também é atribuída a Puiseux, mas isso exigiria, pelo menos, o mesmo tanto de extrapolação. Descreveremos a contribuição de Puiseux no seu devido tempo.

Existem livros excelentes sobre esse assunto. Eu recomendo, em particular, Walker <sup>28</sup>, Brieskorn and Knörrer <sup>29</sup>, Wall <sup>30</sup> e Casas-Alvero <sup>31</sup>.

Seja  $K$  um corpo algebricamente fechado de característica zero. O principal exemplo que tenho em mente é o corpo  $\mathbb{C}$  dos números complexos.

Algumas notações:

- $K[x]$  é o anel de polinômios em  $x$  com coeficientes em  $K$ .
- $K(x)$  é o corpo de funções racionais em  $x$  com coeficientes em  $K$  (corpo de frações de  $K[x]$ ).
- $K[[x]]$  é o anel de séries formais em  $x$ : expressões da forma  $\sum_{i=0}^{\infty} a_i x^i$  tais que os  $a_i$  estão em  $K$ , sem nenhuma referência à questão de convergência.

<sup>28</sup> R. J. Walker. *Algebraic curves*. Springer-Verlag, New York-Heidelberg, 1978. Reimpressão da edição de 1950.

<sup>29</sup> E. Brieskorn and H. Knörrer. *Plane algebraic curves*. Birkhäuser/Springer Basel AG, Basel, 1986.

<sup>30</sup> C. T. C. Wall. *Singular points of plane curves*, volume 63 of *London Mathematical Society Student Texts*. Cambridge UP, 2004.

<sup>31</sup> E. Casas-Alvero. *Singularities of plane curves*. London Mathematical Society Lecture Note Series. Cambridge UP, 1 edition, 2000.

– De modo similar, em duas variáveis, definimos  $K[x, y]$  (polinômios),  $K(x, y)$  (funções racionais) e  $K[[x, y]]$  (séries formais).

Podemos agora enunciar o teorema de Newton de forma precisa.

**Teorema (Newton–Cramer).** *Seja  $F(x, y)$  uma série formal em  $K[[x, y]]$  que se anula na origem e não é divisível por  $x$ . Então, existe um inteiro  $m \geq 1$  e uma série de potências formal  $f(t) \in K[[t]]$  se anulando em 0 e tal que  $F(t^m, f(t))$  seja identicamente nula. Em outras palavras, a equação  $F(x, y) = 0$  tem ao menos uma “solução” da forma  $y = f(x^{1/m})$ .*

Já vimos a estrutura geral da prova.

Deixe-me escrever  $F(x_0, y_0)$  em vez de  $F(x, y)$  uma vez que descreveremos uma construção iterativa, envolvendo alguns  $x_k, y_k$ . Então, seja  $F_0(x_0, y_0) = \sum_{i,j} a_{ij} x_0^i y_0^j$  (com  $a_{00} = 0$ ) uma série formal. Para cada  $(i, j)$  tal que  $a_{ij} \neq 0$ , consideramos o quarto de plano  $x \geq i; y \geq j$ . O *polígono de Newton* é o fecho convexo da união desses quartos de plano. A figura na margem mostra o polígono para

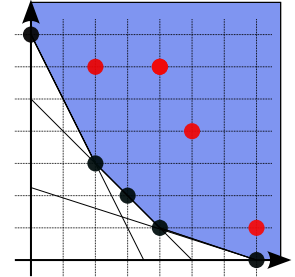
$$F = y_0^7 - x_0^2 y_0^3 + x_0^2 y_0^6 + x_0^3 y_0^2 + x_0^4 y_0 + x_0^4 y_0^6 + x_0^5 y_0^4 + x_0^7 + x_0^7 y_0.$$

Observe que sempre podemos assumir que  $F_0$  não é divisível por  $y_0$ , pois, caso contrário, dividiríamos por algum monômio  $y_0^j$  sem alterar o problema. Em outras palavras, cada um dos eixos intersecta o polígono de Newton. A fronteira desse polígono, longe dos eixos, consiste num número finito de segmentos, contidos em *retas suportes* cujas equações têm a forma  $\alpha i + \beta j = \gamma$  sendo  $\alpha, \beta$  inteiros positivos que podemos assumir serem primos entre si. Escolha uma dessas retas  $\alpha_0 i + \beta_0 j = \gamma_0$  e selecione os coeficientes  $a_{ij}$  tais que  $(i, j)$  estejam nessa reta (existe uma quantidade finita deles). Isso define um polinômio “dominante”

$$F_{dom}(x_0, y_0) = \sum_{\alpha_0 i + \beta_0 j = \gamma_0} a_{ij} x_0^i y_0^j.$$

Então, procuramos por uma solução “aproximada” parametrizada por  $t$ , da forma

$$x = t^{\alpha_0} \quad ; \quad y = ut^{\beta_0}.$$



Um polígono de Newton: Termos na série com  $a_{ij}$  não nulo são representados por pontos. Existem três retas suporte. Note que eu não estou seguindo a ideia estranha de Newton de escrever os  $i$  no eixo vertical. Para o benefício do leitor, siga a tradição e uso  $x$  e  $y$  para os eixos horizontais e verticais respectivamente.

Exemplo: Escolha a reta suporte

$$2i + j = 7,$$

de modo que  $(\alpha, \beta, \gamma) = (2, 1, 7)$ .  $F_{dom}(x_0, y_0) = y^7 - x^2 y^3$  e o polinômio  $p$  seja

$$p(u) = u^7 - u^3$$

Escolha  $u_0 = 1$  de modo que a solução aproximada seja

$$x_0 = t^2; y_0 = t.$$

Por “aproximada”, queremos dizer que ela resolve a parte dominante da equação

$$F_{dom}(t^{\alpha_0}, ut^{\beta_0}) = 0.$$

Obtemos uma equação polinomial em  $u$ :

$$p(u) = \sum_{\alpha_0 i + \beta_0 j = \gamma_0} a_{ij} u^j = 0.$$

Como  $K$  é algebricamente fechado, existe ao menos uma solução não nula  $u_0$ . Voltamos, então, à equação original  $F_0(x_0, y_0) = 0$  e substituímos  $x_0$  por  $x_1^{\alpha_0}$  e  $y_0$  por  $u_0 x_1^{\beta_0} (1 + y_1)$ . Obtém-se uma nova série formal em  $(x_1, y_1)$  que, por construção, é divisível por  $x_1^{\gamma_0}$ . Dividindo por  $x_1^{\gamma_0}$ , obtém-se outra equação equivalente  $F_1(x_1, y_1) = 0$

... e o processo pode ser continuado “tanto quanto se queira”, produzindo uma sequência de equações  $F_k(x_k, y_k)$  ( $k \geq 1$ ) e de inteiros  $\alpha_k, \beta_k, \gamma_k$ .

Está faltando uma propriedade importante que *não* foi discutida por Newton. Temos que mostrar que, após um número finito de passos, os coeficientes  $\alpha_k$  são sempre iguais a 1, o que significa que a inclinação de todas as retas suportes são inversos de inteiros e não apenas números racionais. Essa questão é importante porque cada passo implica na introdução de uma raiz  $x_{k+1} = x_k^{1/\alpha_k}$  e, então, teríamos um problema caso tivéssemos que realizar um número infinito de vezes.

Isso é analisado em detalhes no Capítulo VII do excelente livro de Cramer *Introduction à l'analyse des lignes courbes algébriques*<sup>32</sup>, publicado em 1750. O autor dá créditos a Newton, porém:

*O verdadeiro Método das Séries é baseado no Paralelogramo do Sr. Newton, excelente invenção, mas o Autor não deu a Demonstração, aparentando nem mesmo ter percebido todo seu valor.*

**Definição.** Se  $F$  é uma série formal em  $K[[x, y]]$  não divisível por  $x$ , a *multiplicidade*, denotada  $\text{mult}(F)$ , é a valoração de  $F(0, y)$  como uma série em  $y$ . É também a menor altura de um ponto do polígono de Newton de  $F$  no eixo vertical.

Primeiro pass: Em  $F_0$  substitua

$$x_0 \rightarrow x_1^2, y_0 \rightarrow x_1(1 + y_1).$$

e divide por  $x_1^7$ . Obtemos

$$F_1(x_1, y_1) = x_1 + 4y_1 + 2x_1^2 + 2x_1y_1 + 18y_1^2 + 6x_1^2y_1 + x_1y_1^2 + 34y_1^3 + 10x_1^2y_1^2 + 35y_1^4 + 10x_1^2y_1^3 + 21y_1^5 + 5x_1^2y_1^4 + 7y_1^6 + 3x_1^7 + x_1^2y_1^5 + y_1^7 + x_1^8 + 10x_1^7y_1 + x_1^8y_1 + 21x_1^7y_1^2 + 24x_1^7y_1^3 + 16x_1^7y_1^4 + 6x_1^7y_1^5 + x_1^7y_1^6,$$

que tem um termo linear não trivial em  $y_1$  de modo que  $y_1$  pode ser expresso como uma série de potências em  $x_1$ .

<sup>32</sup> G. Cramer. *Introduction à l'Analyse des lignes courbes algébriques*. Frères Cramer et Cl. Philibert, 1750.

La vraie Méthode des Séries est fondée sur le Parallélogramme de Mr. Newton, invention excellente, mais dont l'Auteur n'a pas donné la Démonstration, dont il semble même n'avoir pas senti tout le prix.

Note que, pela convexidade, qualquer reta suporte intersecta o eixo  $j$  abaixo de  $\text{mult}(F)$ . Em particular, o grau do polinômio  $p(u)$  é no máximo  $\text{mult}(F)$ .

**Lema.**  $\text{mult}(F_1) \leq \text{mult}(F)$ .

Por definição,

$$F_1(x_1, y_1) = x_1^{-\gamma} \sum_{i,j} a_{ij} x_1^{\alpha_0 i + \beta_0 j} u_0^j (1 + y_1)^j.$$

Para encontrar  $\text{mult}(F_1)$ , fazemos  $x_1 = 0$  e olhamos para a valoração de  $p(u_0(1 + y_1))$  como polinômio em  $y_1$ .

$$\text{mult}(F_1) \leq \text{degree}(p) \leq \text{mult}(F).$$

□

Então, pelo algoritmo de Newton, a sequência de multiplicidades  $\text{mult}(F_k)$  é não crescente. Veremos agora que essa desigualdade é estrita exceto quando  $F_0$  tem uma forma muito especial.

Tem-se a igualdade  $p(u_0(1 + y_1))$  contém apenas um monômio de grau  $\text{mult}(F)$ . Isso implica, em particular, que o grau de  $p$  é igual a  $\text{mult}(F)$ . Dito de outra forma, a raiz  $y_1 = 0$  de  $p(u_0(1 + y_1)) = 0$  deve ter multiplicidade  $\text{mult}(F)$ . Isso significa que  $p$  tem a forma

$$p(u) = C(u - u_0)^{\text{mult}(F)}.$$

Esse polinômio tem coeficientes não nulos em todos os graus de 0 a  $\text{mult}(F)$ . Por conseguinte, o segmento da fronteira do polígono de Newton que escolhemos contém pontos para cada valor de  $j$  desde  $j = 0$  até  $j = \text{mult}(F)$ . Isso implica que o polígono de Newton tem apenas um lado (além dos segmentos nos eixos) e que  $\alpha_0 = 1$ .

Vamos recapitular: Ao longo do algoritmo, as multiplicidades  $\text{mult}(F_k)$  são não crescentes e, portanto, têm que ser constantes após um determinado tempo. Nesse ponto, todos os polígonos de Newton têm  $\alpha_k = 1$  (e, mais ainda, têm a estrutura muito especial que acabamos de descrever).

Temos

$$x_0 = x_1^{\alpha_0} = x_2^{\alpha_0 \alpha_1} = \dots = x_k^{\alpha_0 \alpha_1 \dots \alpha_{k-1}} = \dots$$



Gabriel Cramer (1704–1752). Seu livro sobre curvas contém, além de uma séria discussão sobre séries de Newton, uma teoria de equações lineares com  $n$  incógnitas (a famosa Regra de Cramer) e elementos da teoria de eliminação. Gosto do título do apêndice 1: “De l'évanouissement des inconnues” (N.T. em Português: “Sobre o desmaio das incógnitas”) o que soa mais atraente que “eliminação”.

+



Retire a hipótese que  $K$  tem característica zero e considere o fecho algébrico do corpo  $\mathbb{F}_p$  com  $p$  elementos. Tente o algoritmo com  $F = y^p + y^{p+1} - x$  e mostre que algo de estranho ocorre.

Et dès-lors la Série devient régulière, parceque toutes les déterminatrices suivantes partant du point T, on ne tombe plus dans des équations qui ayent plusieurs racines. Tous les termes suivants de la Série peuvent même se calculer avec plus de facilité par la Méthode qu'on va expliquer.

Trecho da prova de Cramer (página 200).

e

$$y_0 = u_0 x_1^{\beta_1} (1 + y_1) = u_0 x_1^{\beta_1} (1 + u_1 x_2^{\beta_2} (1 + y_2)) = \dots$$

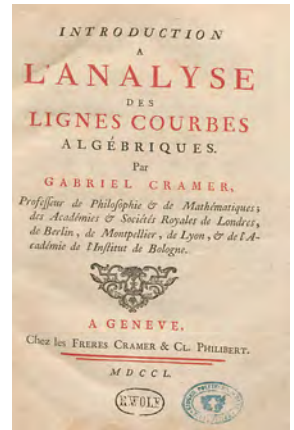
Uma vez que sabemos que os  $\alpha_k$  são iguais a 1 para valores grandes de  $k$ , podemos tomar  $m$  como sendo o produto de todos os  $\alpha_k$  e chamar de  $t$  os valores de  $x_k$  para  $k$  grande. Temos  $x_0 = t^m$  e, mais geralmente, cada  $x_k$  é uma potência de  $t$  com algum expoente inteiro. A fórmula indutiva

$$y_{k+1} = u_{k+1} x_k^{\beta_k} (1 + y_k)$$

define a sequência dos polinômios  $y_k(t)$  na variável  $t$ . Essa sequência “converge” para uma série limite  $f(t) \in K[[t]]$ . Isso quer dizer que a valoração de  $f(t) - y_k(t)$  vai para infinito quando  $k$  converge para infinito.

Para ser completo, deve-se checar se encontramos de fato uma solução para nosso problema, i.e. que  $F(t^m, f(t))$  se anula identicamente. Encorajo você, meu leitor, a fazer essa verificação. Afinal, todo o algoritmo tinha apenas um objetivo: encontrar uma solução!

Essa é a prova do teorema de Newton: toda equação da forma  $F(x, y) = 0$  tem alguma solução, num sentido apropriado. □



### Fecho Algébrico

É hora de construir uma definição precisa de séries com expoentes racionais.

Denotamos por  $K[[x]][x^{-1}]$  o corpo das *séries formais de Laurent*, isto é, expressões formais da forma  $\sum_{i \geq i_0} a_i x^i$  ( $i_0$  pode ser um inteiro negativo). Esse é o o corpo de frações do anel  $K[[x]]$ .

“Un spectacle dont on ne se lasse jamais”.



Geralmente, se  $n$  é um inteiro, denotamos por  $K[[x^{1/n}]][[x^{-1/n}]]$  o corpo das séries de Laurent formais na variável  $x^{1/n}$ : expressões formais da forma  $\sum_{i \geq i_0} a_i x^{i/n}$  ( $i_0$  pode ser um inteiro negativo). O subcorpo das séries tais que  $a_i = 0$  sempre que  $i$  não for divisível por  $n$  é canonicamente isomorfo a  $K[[x]][[x^{-1}]]$  de forma que podemos ver  $K[[x^{1/n}]][[x^{-1/n}]]$  como uma extensão de  $K[[x]][[x^{-1}]]$ . O grupo de Galois dessa extensão é fácil de se descrever, pois consiste nas raízes  $n$ -ésimas da unidade. A ação de uma tal raiz  $\omega$  em

$$\sum_{i \geq i_0} a_i x^{i/n}$$

produz

$$\sum_{i \geq i_0} \omega^i a_i x^{i/n}.$$

Essa extensão é de Galois: os elementos de  $K[[x^{1/n}]][[x^{-1/n}]]$  que são invariantes pela ação do grupo de Galois estão em  $K[[x]][[x^{-1}]]$ .

Do mesmo modo, se  $n_1$  divide  $n_2$ , o corpo  $K[[x^{1/n_1}]][[x^{-1/n_1}]]$  é um subcorpo de  $K[[x^{1/n_2}]][[x^{-1/n_2}]]$ . O limite direto de todas essas extensões de  $K[[x]][[x^{-1}]]$  é denotado por  $K[[x^*]][[x^{*-1}]]$ .

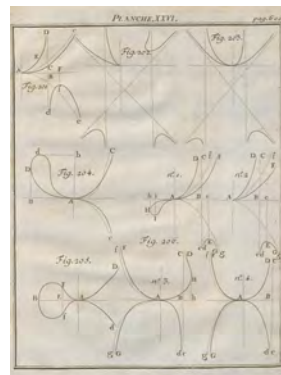
Esse é o corpo de séries de Puiseux: séries com expoentes racionais, tendo um denominador comum. Em termos concretos, uma série de Puiseux é uma expressão formal da forma  $\sum_{i \geq i_0} a_i x^{i/n}$  para algum inteiro positivo  $n$ . Séries de Puiseux com  $i_0 \geq 0$  constituem um anel, que denotamos por  $K[[x^*]]$ .

**Teorema** (Newton–Cramer). *O corpo de séries de Puiseux  $K[[x^*]][[x^{*-1}]]$  é algebricamente fechado, constituindo-se no fecho algébrico das séries de Laurent  $K[[x]][[x^{-1}]]$ .*

Esse teorema não é nada além de uma reformulação do teorema principal desse capítulo.

O fato de  $K[[x^*]][[x^{*-1}]]$  ser uma extensão algébrica de  $K[[x]][[x^{-1}]]$  é claro. De fato, qualquer série de Puiseux está em algum  $K[[x^{1/n}]][[x^{-1/n}]]$  sendo, portanto, algébrica sobre  $K[[x]][[x^{-1}]]$ .

Considere uma equação polinomial com coeficientes em  $K[[x^*]][[x^{*-1}]]$  e variável  $y$ . Introduzindo uma nova variável  $\bar{x} = x^{1/n}$  para algum  $n$  e multiplicando todos os coeficientes por uma



†

potência alta de  $\bar{x}$  podemos assumir que os coeficientes do nosso polinômio estão em  $K[[\bar{x}]]$ . Nossa equação é, por conseguinte, da forma  $F(\bar{x}, y) = 0$  sendo  $F$  uma série formal. Sabemos que tal equação possui como solução uma série em  $\bar{x}^{1/m}$  para algum  $m$  que, em particular, está em  $K[[x^*]][x^{*-1}]$ . Então, para qualquer polinômio não constante com coeficientes em  $K[[x^*]][x^{*-1}]$ , encontramos uma raiz em  $K[[x^*]][x^{*-1}]$ .  $\square$

### Encontrando todas as Soluções

Se pensarmos em  $F(x, y) = 0$  como uma equação na qual a incógnita é uma série  $y(x) \in K[[x]][x^{-1}]$ , podemos tentar, como faríamos com uma equação polinomial qualquer, fatorar  $F$  como um produto de fatores lineares no fecho algébrico.

$$F = A(x, y)(y - f_1(x))(y - f_2(x)) \cdots (y - f_n(x))$$

sendo  $A(0, 0) \neq 0$  e as  $n$  soluções  $f_i(x)$  em  $K[[x^*]][x^{*-1}]$ . Se  $F$  fosse um polinômio na variável  $y$ , seria óbvio, mas é apenas uma série formal. No caso, não é claro que nossa equação tenha uma quantidade finita de soluções.

Na verdade, Newton estava certo, e nossa equação está, de fato, bem próxima de ser uma equação polinomial “padrão”, como explicarei agora.

Deixe-me começar com algumas observações elementares.

**Lema.** *Suponha que uma série formal  $y = f(x) \in K[[x]]$  é solução de uma equação  $F(x, y) = 0$  com  $F \in K[[x, y]]$ . Então,  $F$  é divisível por  $y - f(x)$  em  $K[[x, y]]$ .*

Isso é óbvio se  $f(x) = 0$ . Agora, a transformação formal  $(x, y) \mapsto (x, y - f(x))$  induz um automorfismo de  $K[[x, y]]$  que envia  $y$  em  $y - f(x)$ .  $\square$

**Lema.** *Para  $f \in K[[x^*]]$ , define*

$$\bar{f}(x, y) = (y - f_1(x))(y - f_2(x)) \cdots (y - f_n(x))$$

sendo  $f_1, \dots, f_n$  os conjugados de Galois de  $f$ . Então,  $\bar{f}(x, y)$  está em  $K[[x, y]]$ .



+

Uma vez que  $\bar{f}(x, y)$  é um polinômio em  $y$  cujos coeficientes são invariantes pelo grupo de Galois. Observe que esse polinômio é o polinômio minimal do elemento  $f$  de  $K[[x^*]][x^{*-1}]$  como uma extensão algébrica de  $K[[x]][x^{-1}]$ .  $\square$

**Lema.** *Suponha que uma série de Puiseux formal  $y = f(x) \in K[[x^*]]$  seja a solução de uma equação  $F(x, y) = 0$  com  $F \in K[[x, y]]$ . Então, a série associada  $\bar{f}(x, y) \in K[[x, y]]$  divide  $F(x, y)$  no anel  $K[[x, y]]$ .*

Uma vez que  $f$  é uma solução e a equação é invariante pelo grupo de Galois todos os conjugados são também soluções. Mostra-se, então, usando o primeiro lema  $n$  vezes, que  $F$  é divisível por  $\bar{f}(x, y)$  em  $K[[x^{1/n}, y]]$ . Agora, o quociente  $F/\bar{f}$  é invariante pelo grupo de Galois de onde ele estiver, de fato, em  $K[[x, y]]$ .  $\square$

Podemos provar agora o chamado *teorema de preparação de Weierstrass* para séries formais.

**Teorema.** *Seja  $F(x, y) \in K[[x, y]]$ . Assuma que  $F$  não seja divisível por  $x$  e denote por  $\text{mult}(F)$  sua multiplicidade. Pode-se, então, escrever  $F$  como um produto  $A(x, y)P(x, y)$  considerando  $A, P$  em  $K[[x, y]]$  e*

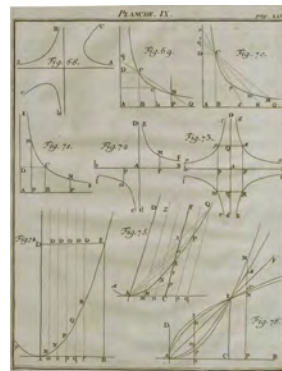
- $A(0, 0) \neq 0$  de forma que  $A$  é um elemento invertível.
- $P(x, y)$  é um polinômio em  $y$  de grau  $\text{mult}(F)$ .

A prova é por indução em  $\text{mult}(F)$ . Note que a valoração de um produto é a soma das valorações e que  $\text{mult}(F) = 0$  significa precisamente que  $F(0, 0) \neq 0$ . Se  $\text{mult}(F) \geq 1$ , sabemos que  $F$  tem ao menos uma solução em  $y = f(x) \in K[[x^*]]$  e que  $F$  é divisível por  $\bar{f}(x, y) \in K[[x, y]]$ . O quociente tem uma multiplicidade menor.  $\square$

Agora, podemos colher os frutos do nosso trabalho e enunciar dois corolários que seguem facilmente do teorema anterior. As provas são as mesmas do caso clássico de anéis de polinômios sobre corpos.

**Teorema.** *Qualquer elemento não nulo  $F$  de  $K[[x, y]]$  pode ser decomposto como*

$$F = A(x, y)x^r(y - f_1(x))(y - f_2(x)) \cdots (y - f_k(x))$$



†

sendo  $A \in K[[x, y]]$  tal que  $A(0, 0) \neq 0$ , as  $k$  soluções  $f_i(x)$  estão em  $K[[x^*]]$  e  $r \geq 0$ .

**Teorema.** O anel  $K[[x, y]]$  é um domínio de fatoração única.

Os fatores irredutíveis são os  $\bar{f}_i \in K[[x, y]]$ .

Para concluirmos, proponho dois exercícios.

*Exercício.* Vimos que o algoritmo de Newton produz soluções  $y(x)$ . A cada passo, temos que escolher uma das retas suportes na fronteira do polígono e uma raiz da equação polinomial correspondente. Mostre que esse algoritmo produz *todas* as soluções  $f_i(x)$  de  $F(x, y) = 0$ .

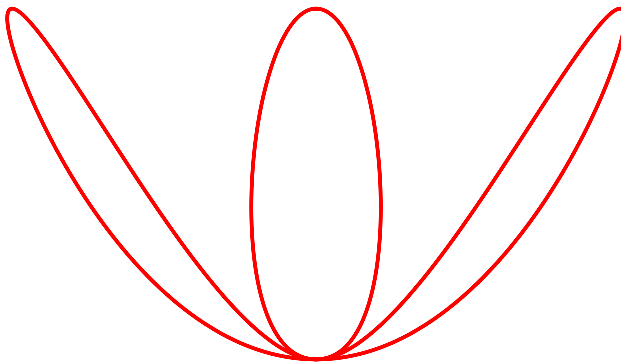
*Exercício.* Suponha que sigamos o algoritmo de Newton com determinadas escolhas de segmentos, levando finalmente a uma solução  $y = f(x)$ . Durante esse processo, obtemos uma sequência de séries formais  $F_k(x_k, y_k)$  e provamos que as multiplicidades  $m_k$  de  $F_k$  são eventualmente constantes, iguais a um inteiro  $m \geq 1$ . Mostre que essa “multiplicidade final”  $m$  é apenas a multiplicidade da raiz, ou seja, o número de fatores iguais a  $(y - f(x))$  na decomposição acima  $F = A(x, y)x^r(y - f_1(x))(y - f_2(x)) \cdots (y - f_k(x))$ .

Basta de álgebra por hora!

É lamentável que o Sr. Newton tenha se contentado em espalhar suas descobertas sem apresentar demonstrações e que ele tenha preferido o prazer de ser admirado ao de instruir.

*Il est fâcheux que Mr. Newton se soit contenté d'étaler ses découvertes sans y joindre de démonstrations et qu'il ait préféré le plaisir de se faire admirer à celui d'instruire.*

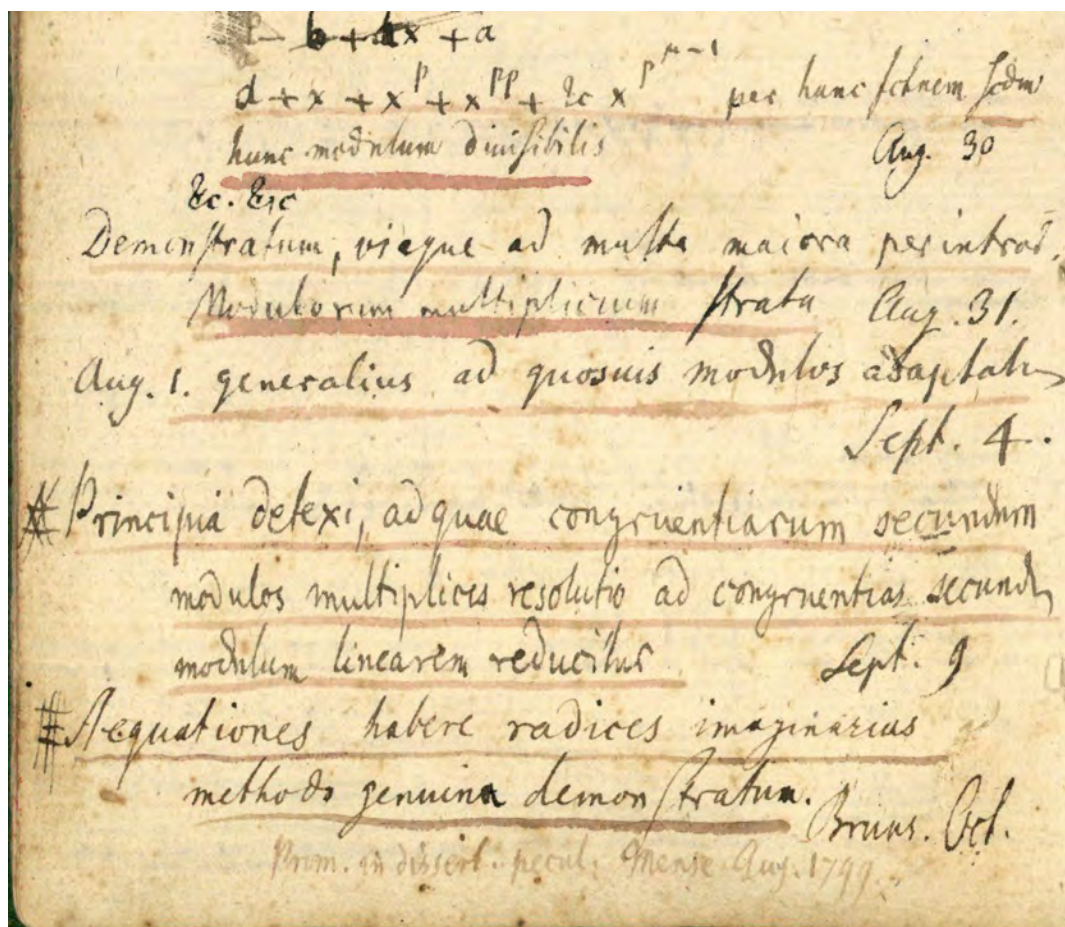
Do prefácio de Cramer.



Uma das curvas de grau 6 estudadas por Cramer. A equação é :

$$y^6 - (y - x^2)(y - 4x^2)^2 = 0.$$

O ponto singular tem 3 ramos. Você pode desenhar esses ramos?



A partir de 1796 (quando ele tinha 19 anos) Gauss registrou suas descobertas matemáticas no seu famoso [Tagebuch](#). Uma lista impressionante de resultados. Veja os comentários de Klein <sup>33</sup> and <sup>34</sup> para uma tradução em inglês. Esta página acima diz respeito a agosto a outubro de 1797. O último item “*Aequationes habere radices imaginarias methodo genuina demonstratum*” enuncia sua prova do teorema fundamental da álgebra. Abaixo desta linha, com uma tinta diferente, uma adição posterior menciona que este era o tema de sua tese: “*Prom[ulgatum] in dissert[atione] peculiari] mense Aug. 1799*”

<sup>33</sup> F. Klein. [Gauß' wissenschaftliches Tagebuch 1796–1814](#). *Math. Ann.*, 57(1):1–34, 1903.

<sup>34</sup> J. J. Gray. A commentary on Gauss's mathematical diary, 1796–1814, with an English translation. *Exposition. Math.*, 2(2):97–130, 1984.

# *Curvam algebraicam neque alicubi subito abrumpi posse Gauss e as curvas algébricas*

## *O teorema fundamental da álgebra*

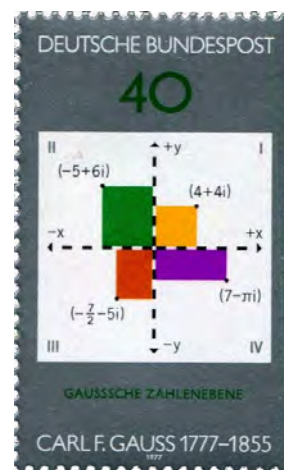
CARL FRIEDRICH GAUSS TINHA 22 ANOS QUANDO DEFENDEU A SUA TESE DE DOUTORADO EM 1799. Isso é [um trabalho extraordinário](#)<sup>35</sup> contendo o que pode ser considerado a primeira “prova” do teorema fundamental da álgebra.

*Qualquer polinômio com coeficientes complexos e não constante possui pelo menos uma raiz complexa.*

Usando uma terminologia ligeiramente diferente e sem a presença das palavras “complexo” ou “imaginário”, que eram suspeitas naquela época, Gauss provou que qualquer polinômio real é um produto de fatores de primeiro ou de segundo graus. Numa linguagem diferente, o título da sua tese de doutorado é:

Demonstratio nova theorematis omnem functionem algebraicam rationalem integram unius variabilis in factores reales primi vel secundi gradus resolvi posse

Essa não é uma prova pelos padrões atuais, mas apresentarei uma variação ligeira no mesmo tema, perfeitamente aceitável pelos matemáticos do século XXI. Não foi a primeira tentativa de uma prova. Entre os predecessores de Gauss, pode-se mencionar d’Alembert, Euler e Lagrange. Nenhuma dessas “provas”



Um selo comemorativo do plano complexo de Gauss. +

<sup>35</sup> C. F. Gauß. *Werke*. Band III. Georg Olms Verlag, Hildesheim, 1973. Reimpressão do original de 1866.

Gauss obteve o título de doutor na Universidade de Helmstedt, sendo orientado formalmente por Johann Friedrich Pfaff, que leu cuidadosamente a sua tese. Entretanto, esse doutorado foi *in absentia*, i.e., não houve apresentação oral. O manuscrito menciona que o principal resultado foi obtido em outubro de 1797. Uma tradução para o inglês da tese de Gauss, feita por Ernest Fandreyer, é [disponível online](#).

anteriores eram sólidas, mesmo naquela época, mas tentarei reconstruir a prova de d’Alembert uma vez que ele usou o polígono de Newton.

A primeira metade da tese de Gauss trata sobre uma crítica às tentativas de provas de seus três predecessores, explicando cuidadosamente porque as provas são falhas. Hoje é difícil imaginar uma situação semelhante: um homem muito jovem defendendo seu doutorado, começando pela destruição sistemática de grandes e respeitados mestres que haviam morrido há apenas quinze anos ou, no caso de Lagrange, que ainda estava vivo. Depois dessa primeira parte, Gauss, então, na segunda metade da tese, apresenta a sua prova. Bela prova, de fato, mas não totalmente isenta de “fatos não provados”. Num momento crucial, a ser descrito a seguir, ele necessita de uma descrição bastante precisa da estrutura local de uma curva algébrica real. Ele então afirma, sem prova, que

*Mas, de acordo com a matemática superior, qualquer curva algébrica (ou as partes individuais de uma tal curva algébrica, talvez consista de várias partes) ou volta para si mesma ou se estende até o infinito. Consequentemente, um ramo de qualquer curva algébrica que entra em um espaço limitado, deve necessariamente sair desse espaço em algum lugar.*

Em outras palavras, Gauss está afirmando que uma curva algébrica não para simplesmente num ponto. A “prova” é dada numa nota de rodapé, um exemplo típico de uma prova por intimidação:

*Parece ter sido provado com suficiente certeza que uma curva algébrica não pode ser quebrada de repente em qualquer lugar (como acontece, por exemplo, com a curva transcendental cuja equação é  $y = 1/\log x$ ). Nem se perde, por assim dizer, em algum ponto depois de infinitas voltas (como a espiral logarítmica). Tanto quanto sei, ninguém levantou dúvidas sobre isso. No entanto, se alguém a exigir, então, comprometo-me a dar uma prova que não está sujeita a qualquer dúvida, em alguma outra ocasião.*

Ninguém levantou dúvidas, e ele vai dar uma prova em alguma outra ocasião ☺! Acontece que ele nunca provou esse fato (ainda que publicou mais três provas do teorema Fundamental

Os “fatos não provados” de Gauss não têm nada a ver com “fake news”: afinal todos eles são verdadeiros.

*Iam ex geometria sublimiori constat, quamvis curvam algebraicam, (siue singulas cuiusvis curvae algebraicae partes, si forte e pluribus composita sit) aut in se redientem aut utrimque in infinitum excurrentem esse, adeoque si ramus aliquis curvae algebraicae in spatium definitum intret, eundem necessario ex hoc spatio rursus alicubi exire debere.*

Satis bene certe demonstratum esse videtur, curvam algebraicam neque alicubi subito abrumpi posse (uti e.g. evenit in curva transcendente, cuius aequatio  $y = 1/\log x$ ), neque post spiras infinitas in aliquo puncto se quasi perdere (ut spiralis logarithmica), quantumque scio nemo dubium contra hanc rem movit. Attamen si quis postulat, demonstrationem nullis dubiis obnoxiam alia occasione tradere suscipiam.

da álgebra, como se ele próprio não estivesse convencido). Que jovem arrogante (e brilhante)!

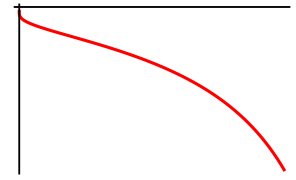
Gauss forneceu dois exemplos de curvas: O primeiro é o gráfico de  $1/\log(x)$ , e o segundo, a espiral logarítmica ( $r = \exp(\theta)$  em coordenadas polares). Ambos podem ser definidos por uma equação  $F(x, y) = 0$  e têm algum tipo de *ponto de parada*. Se desenharmos um pequeno disco em torno desse ponto, a curva entra nesse disco, mas não sai dele. A afirmação (correta) de Gauss é que esse fato se deve à natureza transcendental dessas curvas, mas não acontece para curvas algébricas para as quais  $P(x, y)$  é um polinômio.

### Uma reconstrução da prova de Gauss

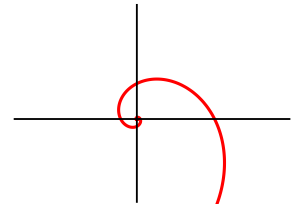
Minha intenção certamente não é discutir essa prova de um ponto de vista histórico. Há muito a ser discutido: o conceito de continuidade, de curva, os argumentos topológicos e acima de tudo o uso geométrico de números complexos como pontos em um plano. Recomendo fortemente os livros de Dhombres e Alvarez <sup>36</sup> e van der Waerden <sup>37</sup>. Permita-me apenas mencionar um ponto de vista lúcido expresso por Gauss, mais de vinte anos antes de Abel e Galois. Detalhe: o que ele chama de *equação pura*, a seguir, é uma equação da forma  $x^n = a$ .

[...] Depois de tanto trabalho de tão grandes matemáticos há muito pouca esperança deixada para chegar a uma solução geral de equações algébricas. Parece cada vez mais provável que tal solução seja inteiramente impossível e contraditória. Isso não deve ser considerado paradoxal, como o que é geralmente chamado de solução de uma equação é de fato nada além de sua redução para equações puras. Pois a solução de equações puras não é aqui ensinada mas pressuposta; e se você expressa as raízes de uma equação  $x^m = H$  por  $\sqrt[m]{H}$ , você não a tem resolvido, e você não fez mais do que se tivesse inventado algum símbolo para denotar a raiz de uma equação  $x^h + Ax^{h-1} + \text{etc.} = 0$  e definir a raiz igual a essa.

Aqui meu modesto propósito é fazer uma moderna reconstrução da prova, mostrando o quanto Gauss precisava de alguma compreensão da natureza local das curvas algébricas.



A curva  $x \exp(-1/y) = 1$  (i.e.  $y = 1/\log x$ ) tem um "ponto final" em  $(0,0)$ .



A espiral logarítmica de equação  $y - x \log \text{tg}(x^2 + y^2) = 0$  ( $\rho = \exp(-\theta)$  em coordenadas polares) tem uma infinidade de voltas e converge para a origem.

<sup>36</sup> J. Dhombres and C. Alvarez. *Une histoire de l'invention mathématique : les démonstrations du théorème fondamental de l'algèbre dans le cadre de l'analyse réelle et de l'analyse complexe de Gauss à Liouville*. Hermann, 2013.

<sup>37</sup> P. D. B. L. van der Waerden. *A History of Algebra: From al-Khwarizmi to Emmy Noether*. Springer-Verlag Berlin Heidelberg, 1 edition, 1985.



Seja  $P(z)$  um polinômio mônico de grau  $n \geq 1$  com coeficientes complexos. A ideia principal consiste em pensar no número complexo  $z = x + iy$  como um ponto do plano e  $P(x + iy)$  como  $p(x, y) + iq(x, y)$ , definindo dois polinômios reais em  $(x, y)$ . Provar a existência de uma raiz complexa de  $P$  é equivalente a mostrar que as duas curvas algébricas  $p(x, y) = 0$  e  $q(x, y) = 0$  se intersectam não trivialmente. Vamos analisar o comportamento qualitativo dessas duas curvas na vizinhança do infinito.

Quando o módulo de  $z$  é grande,  $P(z)$  e  $z^n$  são equivalentes, de modo que se considera como primeira aproximação as curvas

$$\Re(x + iy)^n = 0 \quad \Im(x + iy)^n = 0.$$

Essas equações podem ser facilmente resolvidas: elas definem linhas radiais

$$\arg z = \frac{(2k+1)\pi}{2n} \quad (0 \leq k \leq 2n-1) \quad \arg z = \frac{2k\pi}{2n} \quad (0 \leq k \leq 2n-1)$$

as quais se intersectam na origem. Essas  $2n$  linhas intersectam cada círculo  $|z| = R$  em  $4n$  pontos. A primeira afirmação feita por Gauss foi a seguinte:

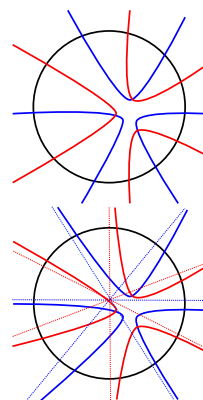
**Lema.** Quando  $R$  é suficientemente grande, cada uma das duas curvas algébricas  $p(x, y) = 0$  e  $q(x, y) = 0$  intersecta o círculo  $|z| = R$  em  $2n$  pontos que são próximos aos pontos anteriores.

As partes real e imaginária de  $\frac{1}{R^n} P(R \exp(i\theta))$  são polinômios trigonométricos de grau  $n$  na variável  $\theta$ , aparecendo  $\cos(n\theta)$  e  $\sin(n\theta)$ . Portanto, cada um pode zerar no máximo  $2n$  vezes, e eles zeram  $2n$  vezes pelo teorema do valor intermediário. Os detalhes elementares são deixados para você, meu leitor. A prova desse ponto por Gauss é perfeita.  $\square$

Agora vem a parte topológica da prova:

Inicialmente suponha que as curvas algébricas  $p(x, y) = 0$  e  $q(x, y) = 0$ , que vamos chamar de curva azul e vermelha, respectivamente, são suaves. No interior do disco  $|z| \leq R$ , elas consistem em um número finito de arcos, cada um deles difeomorfo ao intervalo  $[0, 1]$ , e em um certo número de curvas fechadas, difeomorfas a um círculo. Isso segue da classificação das variedades diferenciáveis compactas unidimensionais (veja, por exemplo, <sup>38</sup>

Note que a “ideia simples” de pensar um polinômio como uma aplicação de um plano em outro plano era um ideia inovadora em 1797.



<sup>38</sup> J. W. Milnor. *Topology from the differentiable viewpoint*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997.

ou <sup>39)</sup>. Então, temos  $4n$  pontos num círculo, azuis e verdes, onde essas cores se alternam. Vamos dizer que dois desses pontos (com a mesma cor) são *pareados* se eles são as duas extremidades de um desses arcos azuis ou vermelhos *no interior do disco*. Então, o conjunto dos nossos  $4n$  pontos consistem em  $2n$  pares de pontos.

Considere quatro pontos distintos no círculo, dois deles coloridos de vermelho e os outros dois de azul. Do ponto de vista topológico, existem duas possibilidades: Podem ser *entrelaçados* ou *não entrelaçados*. Contornando o círculo, lê-se cores alternadas, como “azul, vermelho, azul, vermelho” (caso entrelaçado), e “azul, azul, vermelho, vermelho” (caso não entrelaçado). O lema topológico crucial, bastante intuitivo, é o seguinte:

**Lema.** *Sejam  $b_0, b_1, r_0, r_1$  quatro pontos de um círculo tais que  $\{b_0, b_1\}$  e  $\{r_0, r_1\}$  são entrelaçados. Sejam  $b$  ( $r$ , respectivamente) um arco suave no disco, conectando  $b_0$  e  $b_1$  ( $r_0$  and  $r_1$ , respectivamente). Então,  $b$  e  $r$  se intersectam não trivialmente.*

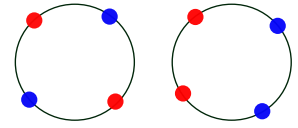
Isso decorre de um dos primeiros teoremas em topologia algébrica (portanto, não estava formalmente à disposição de Gauss). *Duas curvas fechadas no plano que se intersectam transversalmente têm um número par de pontos de interseção (veja por exemplo, o livro do Milnor).* Se existirem arcos *disjuntos* conectando os pontos azul e vermelho no *interior do disco*, poderíamos construir duas curvas fechadas no plano que se cruzam exatamente em um ponto (veja a figura na margem). □

Pelo teorema de Jordan, cada vez que  $c_2$  entra em uma componente conexa do complemento de  $c_1$  (modificado), ele tem que sair de modo que haja realmente um número par de interseções. Agora, leitor, tente provar o teorema de Jordan a partir da paridade de interseção.

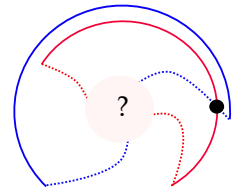
**Lema.** *Suponha que  $2k$  pessoas sentam em torno de uma mesa e elas apertam as mãos dois a dois, sem cruzar os braços! Então, pelo menos dois vizinhos apertam as mãos.*

Para  $k = 2$ , é o lema anterior. Considere duas pessoas apertando as mãos. Eles decompõem a borda da mesa em dois

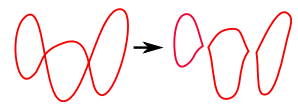
<sup>39)</sup> V. Guillemin and A. Pollack. *Differential topology*. AMS Chelsea Publishing, Providence, RI, 2010. Reimpressão do original de 1974.



Entrelaçados ou não.



O fato de duas curvas fechadas transversais no plano se cruzarem em um número par de pontos é mais ou menos equivalente ao teorema de Jordan: “o complemento de uma curva fechada mergulhada no plano tem exatamente duas componentes conexas”. De fato, se  $c_1$  e  $c_2$  são fechadas e transversais, pode-se primeiro modificar  $c_1$  ligeiramente, sem alterar sua interseção com  $c_2$ , de tal forma que  $c_1$  se torne uma imersão com autointerseções transversais. Então, é possível modificar  $c_1$ , como na figura abaixo, novamente sem alterar a interseção com  $c_2$ , para substituí-la por uma união disjunta de curvas fechadas mergulhadas.



intervalos. Se um está vazio, o resultado está demonstrado. Caso contrário, prossiga por indução.  $\square$

Ainda assumindo que as duas curvas  $p(x, y) = 0$  e  $q(x, y) = 0$ , azul e vermelha, são suaves, nós podemos agora provar o teorema fundamental da álgebra, seguindo Gauss.

Por contradição, suponha que os arcos azul e vermelho não se cruzam. Pelo lema anterior, dois vizinhos no círculo são conectados. Mas isso é impossível uma vez que pontos consecutivos não têm a mesma cor.  $\square$

Agora nós entendemos a dificuldade e o porquê “ninguém levantou dúvidas”. Se houvesse uma curva algébrica com um beco sem saída, um arco poderia penetrar *no interior do disco* e parar lá, sem sair: fatal para a prova.

Agora vamos enunciar precisamente a afirmação de Gauss.

**Teorema.** *Seja  $(x_0, y_0)$  um ponto não isolado em alguma curva algébrica real  $C$  definida por  $F(x, y) = 0$  sendo  $F$  um polinômio real em  $\mathbb{R}[x, y]$ . Então, há um homeomorfismo de algum pequeno disco centrado em  $(x_0, y_0)$  enviando  $C$  para a união de um número par de raios distintos.*

Essa afirmação é de fato verdadeira e “Eu me comprometo a dar uma prova que não está sujeita a qualquer dúvida, em alguma outra ocasião.” ☺

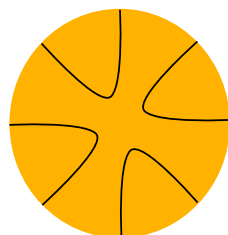
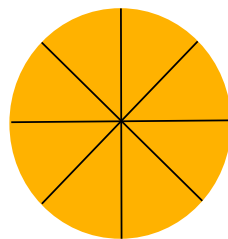
Assumindo que isso é verdade, é fácil finalizar a prova. Se as curvas azul e vermelha  $p(x, y) = 0$ ,  $q(x, y) = 0$  são singulares e disjuntas, podemos ligeiramente modificá-las como na margem, conectando os raios em pares, para que elas se tornem *arcos disjuntos não singulares*. Mas nós tínhamos visto que isso não é possível.  $\square$

### Comentários sobre essa prova

Steve Smale discutiu essa prova em um artigo que trata de versões efetivas do teorema Fundamental da álgebra <sup>40</sup>. Ele enfatizou a afirmação de Gauss:

Mas, por enquanto, gostaria de salientar a imensa lacuna que a prova de Gauss continha. É um ponto sutil até hoje que uma

Considere um número par de pontos em um círculo e conte o número de maneiras de parear esses pontos de modo que os pares não se entrelacem. Mostre que é um número de Catalan.



<sup>40</sup> S. Smale. [The fundamental theorem of algebra and complexity theory](#). *Bull. Amer. Math. Soc. (N.S.)*, 4(1):1–36, 1981.

curva plana algébrica real não pode entrar em um disco sem sair.

Ele também comenta a interminável discussão sobre quem deu a “primeira” prova aceita.

Pode-se entender melhor a situação histórica talvez do ponto de vista de Imre Lakatos <sup>41</sup>. Lakatos na tradição de Hegel, por um lado, e Popper, por outro, veem a Matemática como um desenvolvimento que procede como uma série de ‘provas e refutações’.

Existem muitas maneiras de “corrigir” a prova e preencher a “imensa lacuna”. Primeiro, deve-se mencionar o longo artigo detalhado de Ostrowski, datado de 1920, totalmente dedicado à prova da afirmação de Gauss <sup>42</sup>. As curvas  $p(x, y) = 0$  e  $q(x, y) = 0$  usadas por Gauss são realmente curvas algébricas, mas são curvas algébricas muito especiais. Na terminologia moderna, esses polinômios são partes reais e imaginárias de uma função holomorfa  $P(z)$  e, portanto, são polinômios *harmônicos*. A prova detalhada de Ostrowski trata realmente de polinômios harmônicos, suficiente para o nosso problema atual. Com noções elementares sobre a análise complexa, é realmente fácil preencher os detalhes, como mostro agora.

Considere  $P(z) = P(x + iy) = p(x, y) + iq(x, y)$  como uma aplicação de  $\mathbb{C} \simeq \mathbb{R}^2$  para outra cópia de si mesmo. A diferencial dessa aplicação  $P$  no ponto  $z_0 = x_0 + iy_0$  pode ser vista como uma matriz real  $2 \times 2$  ou como o número complexo  $P'(z_0)$ . Portanto, os pontos críticos são simplesmente os zeros finitos da derivada  $P'$ , e as curvas azuis e vermelhas são as imagens inversas dos dois eixos. Vamos analisar a imagem inversa por  $P$  de alguma reta:

No caso de emergência, o livro <sup>43</sup> pode ser muito útil para entender essas figuras.

Na vizinhança de algum ponto  $z_0$ , temos:

$$P(z) - P(z_0) = c_k(z - z_0)^k + c_{k+1}(z - z_0)^{k+1} + \dots + c_n(z - z_0)^n$$

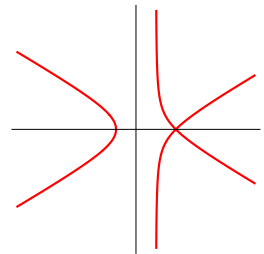
para algum  $k \geq 1$  (a valuação de  $P(z) - P(z_0)$  em  $z_0$ ). Então,

$$P(z) - P(z_0) = \left( (z - z_0) \sqrt[k]{c_k} \sqrt{1 + \frac{c_{k+1}}{c_k}(z - z_0) + \dots} \right)^k = \varphi(z)^k.$$

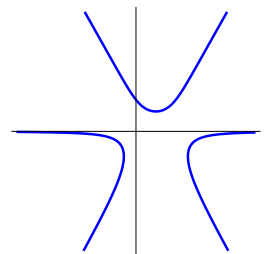
<sup>41</sup> I. Lakatos. *Proofs and refutations*. Cambridge Philosophy Classics. Cambridge UP, paperback edition, 2015. Publicado originalmente em 1976.

<sup>42</sup> A. Ostrowski. *Über den ersten und vierten Gauss'schen Beweis des Fundamental satzes der Algebra*. *Nachrichten der Gesellschaft der Wissenschaften Göttingen*, 1920.

<sup>43</sup> T. Needham. *Visual complex analysis*. The Clarendon Press, Oxford University Press, New York, 1997.



A imagem inversa do eixo vertical pelo polinômio  $2z^3 - 3z^2 + 1 + i$ . Os pontos críticos são  $z = 0, 1$  e os valores críticos são  $1 + i$  e  $i$ : um deles, então, está sobre o eixo vertical.



A imagem inversa do eixo horizontal por  $2z^3 - 3z^2 + 1 + i$ . Não existem valores críticos nesse eixo.

Aqui  $\sqrt[k]{c_k}$  é qualquer escolha da raiz  $k$ -ésima e a segunda raiz  $k$ -ésima é uma série de potências convergente pelo teorema binomial de Newton. A diferencial em  $z_0$  de  $\varphi$  não é zero de modo que  $\varphi$  é um difeomorfismo local. Em suma,  $P(z)$  é a composição local de um difeomorfismo e da aplicação  $z \mapsto P(z_0) + (z - z_0)^k$ . Portanto, é óbvio que a imagem inversa por  $P$  de uma curva suave passando por  $P(z_0)$  é a união de  $k$  curvas suaves através de  $z_0$ . Em particular, temos localmente  $2k$  meias-retas, e isso prova a afirmação de Gauss no caso especial que ele precisava. Esse caso é realmente muito especial, já que temos  $k$  curvas suaves fazendo ângulos iguais.  $\square$

*Mas não se esqueça que ainda não provamos a afirmação de Gauss em sua generalidade.*

Existe outra maneira de preencher a “imensa lacuna” de Gauss. Girando o eixo por um ângulo  $\theta$ , podemos substituir  $P(z)$  por  $\exp(i\theta)P(z)$ . A curva  $p(x, y) = 0$  (ou  $q(x, y) = 0$ ) é singular se, e somente se, um dos valores críticos de  $P$  estiver no eixo real (ou imaginário). Portanto, basta rotacionar com um  $\theta$  adequado para evitar essa situação, de modo que Gauss também poderia ter começado com a afirmação de que se pode sempre supor que as curvas azuis e vermelhas são suaves. Esse argumento é fácil, mas não era fácil em 1797.

Hoje há muitas provas do teorema fundamental da álgebra. Eu recomendo o trabalho de Eisermann <sup>44</sup> para uma visão geral lúcida. Essa prova de Gauss certamente não é a mais direta, nem a mais fácil. Limpá-la requer alguns argumentos topológicos sutis, porém, seguindo esse caminho, ganhamos grandes recompensas, por exemplo: entendemos muito melhor sobre polinômios complexos como aplicações.

Deixe-me apresentar a minha prova favorita, no espírito da prova topológica de Gauss, que encontramos no artigo de Smale já mencionado. Para mim, essa é a mais simples <sup>45</sup>. Escolha um ponto  $z_0$  de tal forma que o segmento  $\gamma$  conectando 0 a  $P(z_0)$  não contenha um dos valores críticos finitos de  $P$ . Isso é possível se 0 não é um valor crítico, mas se esse é o caso 0 é um *valor* de modo que  $P$  tenha uma raiz! Considere a imagem inversa de  $\gamma$  por  $P$ , uma variedade suave e compacta de dimensão 1, com

<sup>44</sup> M. Eisermann. [The fundamental theorem of algebra made effective: an elementary real-algebraic proof via Sturm chains](#). *Amer. Math. Monthly*, 119(9):715–752, 2012.

O que eu quero dizer com “a mais simples”? Provavelmente não a mais curta, pois essa prova contém muitos fatos implícitos que devem ser provados. A simplicidade é um conceito sutil e muito pessoal na matemática. Nesse caso especial, eu diria que ela é simples porque eu acho que não poderia esquecê-la.

<sup>45</sup> E. Ghys. Inner simplicity vs. outer simplicity. In R. Kossak and P. Ording, editors, *Simplicity: Ideals of Practice in Mathematics and the Arts Conferences (CUNY New York, 2013)*. Springer Verlag, 2017.

borda. O ponto  $z_0$  está na borda de uma componente. O outro ponto da borda dessa componente é claramente uma raiz de  $P$ . Voilà!  $\square$

Essa prova simples realmente dá muito mais. Longe dos valores críticos, pode-se fazer um pull-back do campo de vetores radiais  $-x\partial/\partial x - y\partial/\partial y$  pela diferencial de  $P$ . Temos um campo vetorial no plano, longe dos pontos críticos. As trajetórias desse campo vetorial são precisamente as imagens inversas das linhas radiais. Portanto, partindo de um ponto e resolvendo essa equação diferencial, devemos chegar às raízes de  $P$ . Uma maneira de aproximar as soluções de uma EDO é usar o método padrão de Euler. Acontece que o esquema iterativo de Euler coincide com o método de Newton. E assim, Newton, Gauss e Euler juntos!

### A prova de d'Alembert

Vou descrever também uma prova de d'Alembert <sup>46</sup> por duas razões. A primeira é que, na França, o teorema Fundamental da álgebra é frequentemente chamado de teorema de d'Alembert  $\odot$ . A segunda é que isso está intimamente relacionado com os polígonos de Newton que discutimos anteriormente. Veja <sup>47</sup> e <sup>48</sup> para muito mais sobre essa prova. Um fato: D'Alembert não menciona Newton. Como poderia um francês reconhecer a contribuição de um inglês?

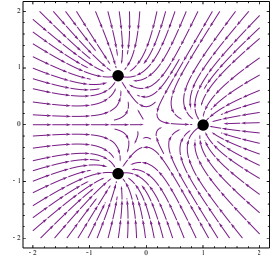
Surpreendentemente d'Alembert publicou uma segunda versão dessa prova num artigo discutindo da *causa dos ventos*<sup>49</sup>. Suponhamos que se queira resolver

$$z^n + a_{n-1}z^{n-1} + \dots + a_0 = 0.$$

seja  $z = y/\varepsilon$ . obtemos uma equação estritamente equivalente:

$$y^n + a_{n-1}\varepsilon y^{n-1} + \dots + \varepsilon^n a_0 = 0.$$

é claro que  $y = 0$  é uma solução para  $\varepsilon = 0$  e queremos uma solução para  $\varepsilon \neq 0$ . Considere a equação acima como uma equação da forma  $F(\varepsilon, y) = 0$ . Sabemos, por Newton e Cramer, que existem soluções não triviais  $y(\varepsilon)$  expressas como séries de Puiseux em  $\varepsilon$ . Portanto, encontramos uma raiz de nossa equação.  $\square$



O retrato de fase do campo vetorial  $-P(z)/P'(z)$  for  $P(z) = z^3 - 1$ . As trajetórias são mapeadas por  $P$  em linhas radiais.

<sup>46</sup> J. D'Alembert. [Recherches sur le calcul intégral](#). *Histoire de l'Acad. Royale Berlin*, pages 182–224, 1748.

<sup>47</sup> C. Baltus. [D'Alembert's proof of the fundamental theorem of algebra](#). *Historia Math.*, 31(4):414–428, 2004.

<sup>48</sup> C. Gilain. Sur l'histoire du théorème fondamental de l'algèbre: théorie des équations et calcul intégral. *Arch. Hist. Exact Sci.*, 42(2):91–136, 1991.

<sup>49</sup> J. D'Alembert. [Réflexions sur la cause générale des vents](#). David l'ainé, Paris, 1747.

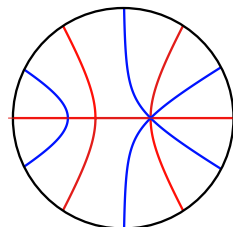


Jean Le Rond d'Alembert (1717–1783).  $\dagger$

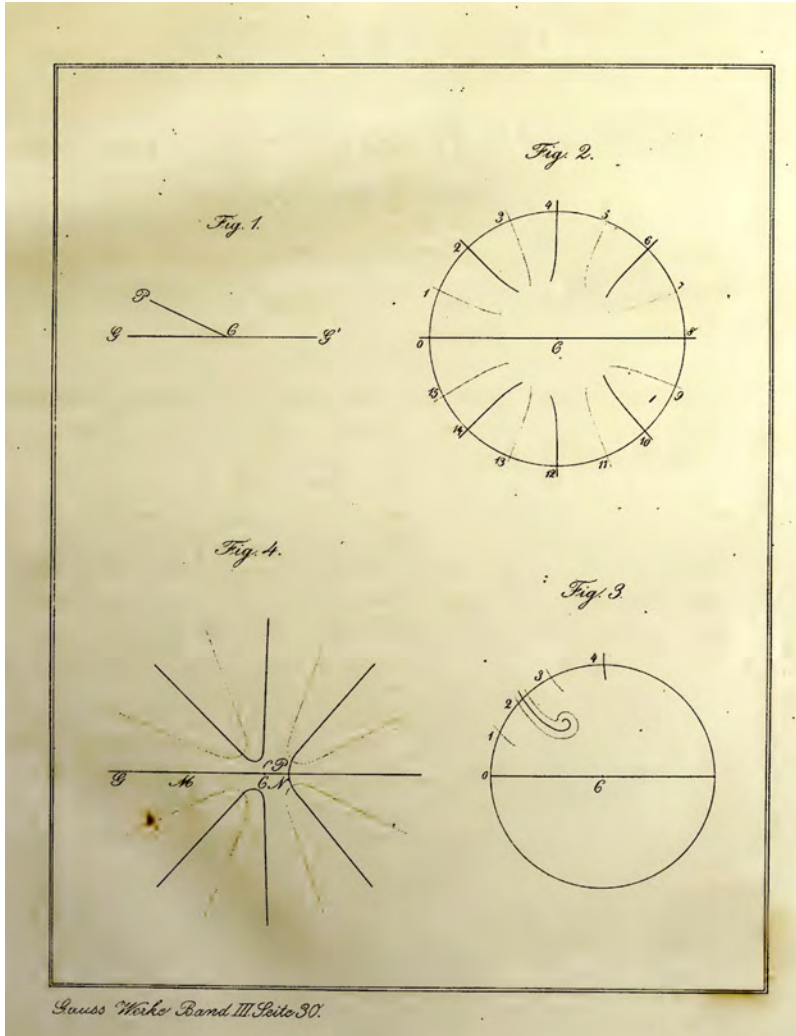
Claro que o “portanto” anterior está sujeito a muita discussão. Uma das principais dificuldades é que nem Newton, nem Cramer, nem d’Alembert provaram a convergência da série. Pior ainda, o algoritmo de Newton, que constrói a série de Puiseux, usa o teorema fundamental da álgebra.

Deixe-me concluir esse capítulo com um exercício, sugerido pelo meu ex-aluno Victor Kleptsyn. Considere as imagens inversas do eixo real (digamos: vermelho) e do eixo imaginário (azul) por algum polinômio complexo  $P(z)$ , produzindo dois grafos (azul e vermelho) em um disco de raio suficientemente grande. Assim, os vértices do grafo azul (ou vermelho) são pontos críticos de  $P$ , mapeados para o eixo real (ou imaginário), apresentando um número par de arestas azuis (respectivamente vermelhas) saindo de um vértice. Geralmente, não existe tal ponto singular. A imagem local em torno da interseção dos dois grafos foi descrita acima:  $4k$  bordas saindo do vértice, alternando ciclicamente azul e vermelho. Tais interseções existem pelo teorema fundamental da álgebra. Além disso, sabemos que em grandes círculos, temos alternância entre vermelho e azul.

A questão, leitor, diz respeito ao inverso. Suponha que temos dois grafos, azul e vermelho, em um disco apresentando todas as características qualitativas anteriores. Sob que condições existe algum polinômio  $P(z)$  tal que seus grafos coloridos associados sejam homeomorfos aos grafos dados, sob algum homeomorfismo do disco?



As imagens inversas dos eixos real e imaginário por  $z^3 - 3z + 2$ . Existem dois valores críticos: 0 and 4.



Essa é a única figura da Tese de Gauss.

+





Joseph Alfred Serret  
(1819–1885).

†

# Prova da afirmação de Gauss sobre singularidades de curvas algébricas: Dois artigos por dois Serret

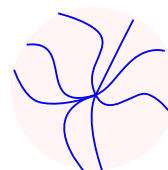
É HORA DE PROVAR A AFIRMAÇÃO DE GAUSS: “a vizinhança de um ponto de uma curva algébrica plana é homeomorfa a um número *par* de raios em um disco”.

## Provas Insuficientes

Joseph Alfred Serret (1819–1885) não deve ser confundido com Paul Joseph Serret (1827–1898).

O primeiro deles, Joseph Alfred, teve uma carreira brilhante na matemática. Ele assinava seus livros como “Membre de l’Institut et du bureau des longitudes, Professeur au Collège de France et à la faculté des sciences de Paris”. Em 1849, publicou o bastante influente *Cours d’algèbre supérieure* em dois volumes contendo uma das primeiras exposições sistemáticas da Teoria de Galois. Também se encontra na origem do triedro de Frenet-Serret para curvas em um espaço tridimensional.

Já o segundo, Paul Joseph, era mais jovem e teve uma carreira bem mais modesta, assinando seus livros como “Docteur ès sciences, membre de la société philomatique” e lecionando no Collège Sainte-Barbe in Paris. Não sei se é um indicador válido sobre sua carreira modesta, mas não consegui encontrar uma fotografia dele dessa época!



Figuras locais de uma curva algébrica.



Em 1847, Joseph Alfred escreveu um artigo <sup>50</sup> no qual ele “prova” uma afirmação de Newton:

Se uma curva é assintótica a um ramo de uma curva algébrica, então, ela é também assintótica a outro ramo.

Note que o que Newton chama aqui de “ramo” é metade do que nós chamamos de um ramo... Como um exemplo simples, olhe para o folium de Descartes  $x^3 + y^3 = 3xy$ . Sua assíntota é aproximada pela curva quando  $x$  tende a  $+\infty$  e a  $-\infty$ , correspondendo a dois ramos na terminologia de Newton e a um ramo no infinito na nossa nomenclatura.

A prova de Joseph Alfred Serret consiste no seguinte: Seja  $F(x, y) = 0$  a equação da curva em um sistema de coordenadas de modo que  $y = 0$  seja a assíntota. Mudemos  $x$  por  $1/x$ , que produz uma segunda curva algébrica  $F_1(x, y) = 0$ . Agora se a curva original tivesse apenas um ramo assintótico a  $y = 0$ , então, a curva algébrica  $F_1(x, y) = 0$  teria um “point d’arrêt”, i.e. um *ponto de parada*, o que é impossível. Surpreendentemente, Joseph Alfred toma como certo que um tal ponto de parada é impossível. Claramente, isso não é uma prova em nenhum sentido da palavra.

Ironicamente, ele critica Euler por sua falta de rigor. Ao final do seu curto artigo, ele cita Euler *Introductio in analysin infinitorum* (volume 2, chapter 7, section 174):

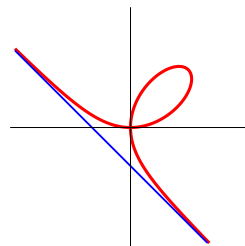
**Quam ob** rem Linea curva duos habebit ramos in infinitum excurrentes inter se oppositos. . .

A última frase do artigo de Joseph Alfred é: “Isso *quam ob rem* precisava de uma prova”. Será que ele acreditava mesmo que Euler, ou Newton, não teriam pensado na mudança de variáveis  $x \mapsto 1/x$ ?

Dezoito anos mais tarde, o desconhecido Paul Joseph escreveu outro artigo curto <sup>51</sup>, no mesmo jornal científico, criticando o artigo de seu quase xará e prestigioso colega, dizendo que Joseph Alfred reduz o problema de assíntotas ao problema de pontos de parada de curvas algébricas, mas isso era “a priori óbvio”. Agora — Paul Joseph insiste — a questão principal continua em aberto: ainda é preciso provar que uma curva algébrica não pode ter

<sup>50</sup> J. A. Serret. *Théorème sur les courbes algébriques asymptotiques*. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 6:217–218, 1847.

De acordo com Joseph Alfred, “*Ce théorème est dû à Newton, et est énoncé, si je ne me trompe, dans son Enumeratio Linearum tertio ordains*”. “Esse teorema é devido à Newton e está enunciado, se não me engano, em seu *Enumeratio Linearum tertio ordains*.”



A assíntota do folium de Descartes  $x^3 + y^3 - 3xy = 0$ .

“*ce qui ne peut arriver pour une courbe algébrique*”.

Por essa razão, a curva tem dois ramos no infinito que serão opostos um ao outro. . .

<sup>51</sup> P. J. Serret. *Note sur les courbes algébriques*. *Nouvelles annales de mathématiques, journal des candidats aux écoles polytechnique et normale*, 4:311–313, 1865.

um ponto de parada. Para isso, ele finalmente propõe a seguinte prova.

Seja  $(0,0)$  um ponto em uma curva algébrica  $F(x,y) = 0$ . Intersectemos a curva com um pequeno círculo centrado na origem  $x^2 + y^2 = r^2$ . Obtemos a seguinte parametrização.

$$x = \frac{2rt}{1+t^2} \quad y = \frac{(1-t^2)r}{1+t^2}.$$

Substituindo em  $F(x,y) = 0$  e multiplicando por  $(1+t^2)^d$  sendo  $d$  o grau de  $F$ , obtemos uma equação  $\varphi_{2d}(t) = 0$ , sendo  $\varphi_{2d}$  um polinômio de grau  $2d$ . Agora se o ponto  $(0,0)$  fosse um ponto de parada, a curva intersectaria um pequeno círculo em um único ponto, de modo que uma equação em  $t$  de grau  $2d$  teria uma única raiz, “*ce qui serait absurde*”.

Incrível. Como pode Paul Joseph não saber que  $t^2$  tem grau dois e tem uma única raiz? Essa raiz é “dupla”, mas esse é exatamente nosso problema. Poderia-se imaginar uma curva algébrica indo a um ponto e voltando, seguindo o *mesmo* caminho.

Neste ponto, há algo a ser provado.

### Dois fatos importantes em álgebra comutativa

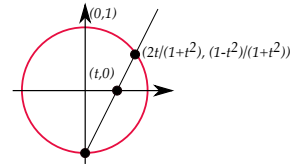
Coletei aqui dois teoremas básicos sobre polinômios que nos permitirá consertar a prova de Paul Joseph. Para aprofundar, leitor, na álgebra, recomendo esses livros <sup>52</sup> <sup>53</sup>, ou <sup>54</sup>. Todos os anéis serão assumidos comutativos. Algumas definições úteis estão na margem.

**Teorema.** *Seja  $\mathcal{R}$  um domínio de fatoração única. Então o anel polinomial  $\mathcal{R}[x]$  também é um domínio de fatoração única.*

Dizemos que um polinômio em  $\mathcal{R}[x]$  é *primitivo* se seus coeficientes são primos entre si. O ponto principal é o chamado... “lemma de Gauss”.

**Lema.** *O produto de dois polinômios primitivos em  $\mathcal{R}[x]$  é primitivo.*

A prova (moderna) é fácil (mas, de certo modo, indireta). Se  $p$  é primo em  $\mathcal{R}$ , o anel  $\mathcal{R}/p$  é um domínio de integridade Se



Por grau de  $F$ , quero dizer o máximo  $i + j$  com  $x^i y^j$  variando entre todos os monômios com coeficientes não triviais.

Como veremos, Paul Joseph está certo: o grau é exatamente  $2d$  e não apenas  $\leq 2d$ .



<sup>52</sup> S. Lang. *Algebra*, volume 211 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, third edition, 2002.

<sup>53</sup> M. Artin. *Algebra*. Prentice Hall, Inc., Englewood Cliffs, NJ, 1991.

<sup>54</sup> N. Bourbaki. *Commutative algebra. Chapters 1–7*. Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 1998.

Provado por Gauss no Artigo 42 de seu *Disquisitiones Arithmeticae* em 1801, três anos após seu PhD.

$P_1(x)$  e  $P_2(x)$  são dois polinômios cujo produto não é primitivo, todos os coeficientes de  $P_1P_2$  são divisíveis por algum primo  $p$  e, então, pode-se reduzir todos os coeficientes módulo  $p$  e obter a seguinte igualdade em  $(\mathcal{R}/p)[x]$ :

$$\overline{P}_1(x)\overline{P}_2(x) = 0.$$

Como um anel de polinômios sobre um domínio de integridade é um domínio de integridade, concluímos que  $\overline{P}_1(x)$  ou  $\overline{P}_2(x)$  é zero em  $(\mathcal{R}/p)[x]$ .  $\square$

Definimos o *conteúdo*  $\text{cont}(P)$  de um polinômio  $P(x) \in \mathcal{R}[x]$  como o máximo divisor comum de seus coeficientes. Claramente, todo polinômio  $P(x)$  pode ser escrito como o produto  $\text{cont}(P)\tilde{P}(x)$  sendo  $\tilde{P}(x)$  primitivo. O lema de Gauss significa simplesmente que o conteúdo do produto é o produto dos conteúdos.

Podemos agora provar o teorema. Vamos mostrar que elementos primos em  $\mathcal{R}[x]$  são:

1. Elementos primos de  $\mathcal{R}$ , vistos como polinômios constantes,
2. Polinômios primitivos em  $\mathcal{R}[x]$  que são primos quando vistos como polinômios sobre o corpo de frações  $\text{Frac}(\mathcal{R})$  de  $\mathcal{R}$ .

O anel de polinômios sobre um corpo é euclidiano. Isso se aplica ao  $\text{Frac}(\mathcal{R})[x]$  de modo que qualquer elemento  $P(x)$  de  $\mathcal{R}[x]$  pode ser escrito como um produto de polinômios primos em  $\text{Frac}(\mathcal{R})[x]$ . Eliminando denominadores, podemos escrever  $P$  como um produto de elementos dos tipos 1 e 2 :

$$P(x) = u \cdot r_1 \cdots r_k \cdot P_1(x) \cdots P_l(x).$$

Aqui  $u$  é uma unidade em  $\mathcal{R}$ , os  $r_i$  são primos em  $\mathcal{R}$  e os  $P_i$  são primitivos e irredutíveis em  $\text{Frac}(\mathcal{R})[x]$ . Por Gauss, o produto  $r_1 \cdots r_k$  é o conteúdo de  $P$  e, portanto, fica unicamente determinado por  $P$ .

Como  $\mathcal{R}$  é um domínio de fatoração única, os  $r_i$  são unicamente determinados por  $P$  a menos de unidades e permutações.

Como  $\text{Frac}(\mathcal{R})[x]$  é também um domínio de fatoração única, os fatores  $P_i(x)$  são também unicamente determinados a menos

Um domínio de integridade é um anel no qual o produto de dois elementos não nulos é não nulo.

Uma *unidade* em um anel é um elemento que admite inverso.

Dois elementos  $a, b$  em um anel são chamados *associados*, denotado por  $a \equiv b$ , se existe uma unidade  $u$  tal que  $b = ua$ .

Um *elemento primo*  $p$  em um domínio de integridade  $\mathcal{R}$  é um elemento tal que o anel quociente  $\mathcal{R}/p$  é um domínio de integridade.

Um elemento  $a$  em um domínio de integridade é *irredutível* se não é o produto de dois elementos que não sejam unidades. Elementos primos são irredutíveis, mas a recíproca não vale em geral.

Um *domínio de fatoração única* (às vezes chamado de domínio fatorial) é um anel no qual todo elemento é um produto de elementos primos, único a menos de ordenação e de multiplicação por unidades. Anéis euclidianos e principais — por exemplo o anel de polinômios sobre um corpo — são anéis de fatoração única. Nesse caso, os conceitos de primos e irredutíveis coincidem e pode-se definir máximo divisor comum etc.

de unidades e permutações, em  $\text{Frac}(\mathcal{R})[x]$ . Agora uma igualdade  $Q(x) = aP(x)$  sendo  $P(x), Q(x)$  primitivos em  $\mathcal{R}[x]$  e  $a$  está em  $\text{Frac}(\mathcal{R})$  implicando que  $a$  é uma unidade em  $\mathcal{R}$ .  $\square$

O corolário imediato que usaremos é que para qualquer corpo  $K$ , os anéis de polinômios  $K[x_1, \dots, x_n]$  são domínios de fatoração única. Nesse caso especial, o teorema significa que qualquer polinômio não constante em  $K[x_1, \dots, x_n]$  pode ser escrito como produto de fatores irredutíveis, de modo único, a menos de permutações e de multiplicações por fatores constantes (em  $K$ ).

O segundo fato algébrico que usaremos diz respeito ao *resultante*. Sejam  $P_1(x), P_2(x)$  dois polinômios no anel de polinômios  $\mathcal{R}[x]$  sobre algum domínio de integridade  $\mathcal{R}$ , de graus  $d_1, d_2 \geq 1$ . Denotemos por  $\mathcal{R}_d[x]$  o  $\mathcal{R}$ -módulo dos polinômios de grau menor ou igual a  $d$ , isomorfo a  $\mathcal{R}^{d+1}$ . Considere a aplicação

$$\Phi : (A_1, A_2) \in \mathcal{R}[x]_{d_2-1} \times \mathcal{R}[x]_{d_1-1} \mapsto A_1P_1 - A_2P_2 \in \mathcal{R}_{d_1+d_2-1}[x].$$

Isso pode ser vista como uma aplicação linear do espaço  $\mathcal{R}^{d_1+d_2}$  nele mesmo. Seu determinante é chamado *resultante* de  $P_1$  e  $P_2$ , sendo denotado por  $\text{Res}(P_1, P_2)$ . Esse elemento de  $\mathcal{R}$  é uma expressão polinomial universal, com coeficientes em  $\mathbb{Z}$ , nos coeficientes de  $P_1$  e  $P_2$ .

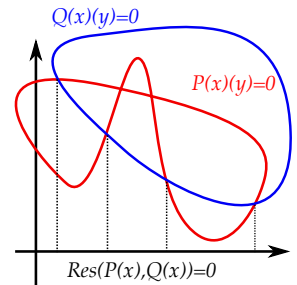
**Teorema.** *Suponha que  $\mathcal{R}$  seja um domínio de fatoração única. O resultante  $\text{Res}(P_1, P_2)$  é igual a zero  $P_1$  e  $P_2$  têm um divisor não trivial comum em  $\mathcal{R}[x]$ .*

De fato, se  $P_1 = QQ_1$  e  $P_2 = QQ_2$ , o elemento  $(Q_2, Q_1)$  está no núcleo de  $\Phi$  de modo que o resultante se anula.

Reciprocamente, se o resultante se anula, o núcleo de  $\Phi$  é não trivial, existindo elementos  $A_2, A_1$  em  $\mathcal{R}[x]_{d_1-1}$  e  $\mathcal{R}[x]_{d_2-1}$  tais que  $A_1P_1 = A_2P_2$ . A conclusão segue do fato de que  $\mathcal{R}[x]$  é um domínio de fatoração única: se  $P_1$  e  $P_2$  fossem primos entre si,  $P_1$  dividiria  $A_2$  o que é impossível uma vez que o grau de  $A_2$  é menor que o grau de  $P_1$ .  $\square$

### Prova da afirmação de Gauss

Podemos agora provar que a vizinhança de um ponto numa curva algébrica real consiste num número *par* de arcos se intersec-



A projeção da interseção de duas curvas  $P(x, y) = 0$  e  $Q(x, y) = 0$  no eixo  $x$  é dada pelos zeros do resultante.

tando apenas na origem.

Seja  $F(x, y) = 0$  a equação da nossa curva algébrica real passando pela origem  $(0, 0)$ . Sabemos que  $F$  pode ser escrito como um produto de fatores irredutíveis:

$$F(x, y) = F_1(x, y) \cdot \cdots \cdot F_n(x, y).$$

Sem mudar o lugar geométrico dos zeros de  $F$  em uma vizinhança de  $(0, 0)$ , podemos deletar os fatores duplicados na decomposição anterior assim como os fatores que não se anulam em  $(0, 0)$ . Podemos, portanto, assumir que todos os  $F_i$  se anulam em  $(0, 0)$  e que são fatores irredutíveis não associados.

O conjunto dos zeros de  $F$  em uma vizinhança de  $(0, 0)$  é a união dos conjuntos dos zeros das  $F_i$ .

Para provar a afirmação de Gauss, provamos dois lemas.

**Lema.** *Seja  $P(x, y)$  um polinômio irredutível e  $Q(x, y)$  um polinômio qualquer. Suponha que as curvas  $P(x, y) = 0$  e  $Q(x, y) = 0$  tenham um número infinito de pontos de interseção em alguma vizinhança da origem. Então,  $P$  divide  $Q$  em  $\mathbb{R}[x, y]$ . Em particular, se  $P$  e  $Q$  são ambos irredutíveis e não-associados, então, numa vizinhança suficientemente pequena de um ponto dado, as correspondentes curvas podem se intersectar apenas naquele ponto.*

**Lema.** *Se  $P(x, y)$  for irredutível, seu conjunto de zeros numa vizinhança da origem consiste em um número par de arcos convergindo para  $(0, 0)$ .*

Começemos com o primeiro lema:

Se  $P(x, y) = 0$  contém um número infinito de pontos no mesmo eixo vertical  $x = x_0$ , o polinômio  $P(x, y)$  deve ser divisível por  $(x - x_0)$  e, uma vez que estamos assumindo que esse polinômio é irredutível e se anula na origem, isso implica que  $P(x, y)$  é um múltiplo constante de  $x$  para o qual o lema é óbvio. Sem perda de generalidade, podemos assumir que  $P(x, y) = 0$  intersecta cada linha vertical num número finito de pontos.

Se  $P(x_0, y_0) = Q(x_0, y_0) = 0$ , os dois polinômios  $P(x_0, y)$ ,  $Q(x_0, y)$ , vistos como elementos de  $\mathbb{R}[y]$ , têm uma raiz comum  $y_0$  e, portanto, seu resultante é nulo como elemento de  $\mathbb{R}$ .

Os polinômios  $F_i, F_j$  são ditos não associados quando  $F_i \neq F_j$  para  $i \neq j$ : não existe constante  $c$  tal que  $F_j = cF_i$ .

Assuma que as curvas  $P(x, y) = 0$  e  $Q(x, y) = 0$  tenham um número infinito de pontos de interseção em uma vizinhança pequena da origem. Olhemos para o resultante de  $P, Q \in \mathbb{R}[x][y]$  como um elemento de  $\mathbb{R}[x]$ . Esse resultante se anula para uma quantidade infinita de valores  $x_0$  e, portanto, se anula identicamente. Vimos que isso implica que  $P, Q$  têm um fator comum em  $\mathbb{R}[x, y]$ . Como  $P$  é irredutível, isso mostra que  $P$  divide  $Q$ .  $\square$

Provaremos agora o segundo lema seguindo a ideia de Paul Joseph:

Coloquemos

$$F(x, y) = \sum_{i,j} a_{ij} x^i y^j.$$

Denotemos por  $d$  o grau de  $F$  (que é por definição o valor máximo de  $i + j$  para o qual  $a_{ij} \neq 0$ ). Note que, uma vez que  $F$  é irredutível, ele não é divisível por  $x$  (a menos que seja um múltiplo constante de  $x$ ) de modo que um dos coeficientes  $a_{0j}$  não é 0. Fixando  $r$ , obtemos uma parametrização  $t$  do círculo de raio  $r$  (menos o ponto  $(0, -r)$ ):

$$x = r \frac{2t}{1+t^2} \quad y = r \frac{1-t^2}{1+t^2}.$$

Substituindo em  $F(x, y)$  e multiplicando o resultado por  $(1+t^2)^d$ :

$$\varphi_{2d,r}(t) = \sum_{i,j} a_{ij} r^{i+j} (2t)^i (1-t^2)^j (1+t^2)^{d-i-j}.$$

Isso é um polinômio em  $t$  cujo monômio de grau mais alto é

$$\left( \sum_j (-1)^j a_{0j} r^j \right) t^{2d}$$

que certamente não é zero para  $r \neq 0$  pequeno. Portanto, Paul Joseph estava certo e o grau de  $\varphi_{2d,r}(t)$  é igual a  $2d$ .

Para concluir a prova, ainda precisamos mostrar que as raízes de  $\varphi_{2d,r}(t) = 0$  são simples para  $r \neq 0$ , o que implicará que existem um número par de raízes. Numa raiz dupla  $t_0$ , o polinômio  $\varphi_{2d,r}(t)$  e sua derivada se anulam simultaneamente. Geometricamente, isso significa que a tangente ao círculo nesse ponto é também tangente a curva  $F(x, y) = 0$ . Dito de outra forma, os pontos duplos que queremos excluir correspondem à interseção

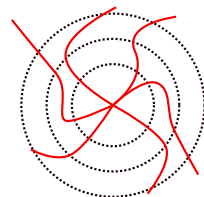
Observe que quando  $F(x, y) = x$ , o lema é óbvio.



de  $F(x, y) = 0$  com a curva  $y\partial F/\partial x - x\partial F/\partial y = 0$ . Dado que assumimos que  $F$  é irredutível, o primeiro lema implica que essas curvas intersectam num número finito de pontos, a menos que  $y\partial F/\partial x - x\partial F/\partial y$  divida  $F$ . Por questões de grau, isso somente pode acontecer se  $y\partial F/\partial x - x\partial F/\partial y$  é um múltiplo constante de  $F$ , significando, por sua vez, que  $F$  é constante no círculo. Isso implica que  $F$  é um polinômio em  $x^2 + y^2$  e, uma vez que se anula na origem, deve ser divisível por  $x^2 + y^2$ .  $F$  é, por conseguinte, um múltiplo constante de  $x^2 + y^2$ , cujo conjunto de zeros se reduz à origem.  $\square$

A prova da afirmação de Gauss está essencialmente concluída. A restrição de  $F$  a cada círculo  $x^2 + y^2 = r^2$  para  $r$  pequeno e não nulo tem uma quantidade par de zeros que são simples. Usando o teorema da função implícita nessa situação bastante elementar, concluímos que esse zeros definem um número par de curvas disjuntas convergindo para a origem. Mas essa conclusão não diz nada sobre a direção limite dessas curvas: elas podem a priori convergir para a origem sem ter uma tangente limite  $\square$

Realmente: Carl Friedrich Gauss e Paul Joseph Serret estavam certos.



Veremos que qualquer ramo de uma curva algébrica tem de fato uma tangente.

**COURS**  
DE  
**CALCUL DIFFÉRENTIEL**  
**ET INTÉGRAL,**

PAR J.-A. SERRET,  
MEMBRE DE L'INSTITUT ET DU BUREAU DES LONGITUDES.

TROISIÈME ÉDITION.

TOME PREMIER.  
CALCUL DIFFÉRENTIEL.

PARIS,  
GAUTHIER-VILLARS, IMPRIMEUR-LIBRAIRE  
DE L'ÉCOLE POLYTECHNIQUE, DU BUREAU DES LONGITUDES,  
SUCCESSEUR DE MALLET-BACHELIER,  
Quai des Augustins, 55. }

1886  
(Tous droits réservés.)



Uma figura do livro de J. Lamouroux, datado de 1821, representando *Oculina Hirtella*. Este livro estava na biblioteca do *HMS Beagle*, que também foi a cabine de Darwin por cinco anos. Ramis sparsis divergentibus!

# De seriebus divergentibus: Euler, Cauchy e Poincaré

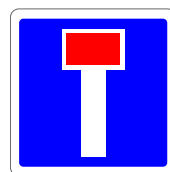
## Seriebus divergentibus por Euler

NEWTON NÃO FEZ O USO DE SÉRIES INFINITAS PARA AS EQUAÇÕES DA FORMA  $F(x, y) = 0$ . Também utilizou um caminho usual para resolver *equações diferenciais*, em uma abordagem essencialmente prática: Olha-se para uma solução como uma série formal e, então, pode-se calcular indutivamente um grande número de termos da série, a fim de obter alguma “precisão”. Na época, não havia uma compreensão sistemática em relação ao conceito de convergência, mas as séries eram de fato convergentes em todos os casos tratados por Newton.

Mais tarde, Euler tornou-se o grande mestre das séries. É uma opinião comum entre matemáticos contemporâneos que Euler foi descuidado com séries e manipulou algumas que “não faziam sentido”. Por exemplo, a seguinte fórmula <sup>55</sup>

$$1 - 2^3 + 3^3 - 4^3 + 5^3 - 6^3 + \dots = -\frac{1}{8}$$

É chocante para os estudantes de graduação que foram ensinados, desde cedo, sobre a definição de uma série convergente. Consequentemente, eles se recusam a considerar esses horrores. Não! Mas Euler sabe o que faz. Ele discute vários procedimentos para atribuir uma soma a uma série, mesmo que ela seja divergente, tentando comparar os diversos métodos. As suas séries não são as mais gerais, são implicitamente definidas por algum



Fim da linha. Este capítulo é completamente independente do resto do livro.



Leonhard Euler (1707–1783)  
†

<sup>55</sup> L. Euler. Remarques sur un beau rapport entre les séries des puissances tant directes que réciproques. *Mémoires de l'académie des sciences de Berlin*, 17:83–106, 1768. Veja o [Euler Archive](#) para traduções para o Inglês e comentários.

tipo de algoritmo, usando um anacronismo. Ele estava convencido de que séries divergentes representam “algo” inerentemente ligado à natureza da série.

O seu trabalho *De seriebus divergentibus*<sup>56</sup> é uma joia e eu o recomendo fortemente a qualquer matemático.

Um dos seus exemplos é famoso:

$$S = 1 - 1! + 2! - 3! + 4! - 5! + \dots$$

Usando cinco diferentes métodos para somar a série, Euler obtém que o valor indicado por  $S$  poderia ser 0.5963473621237. Um dos mais convincentes métodos considera o fato de que a série formal

$$\widehat{f}(x) = x - 1!x^2 + 2!x^3 - \dots$$

é uma solução da equação diferencial *linear*

$$x^2y' + y = x.$$

Uma equação bastante elementar que pode determinar uma solução explícita, que é 0 para  $x = 0$ :

$$f(x) = \exp\left(\frac{1}{x}\right) \int_0^x \frac{1}{t} \exp\left(-\frac{1}{t}\right) dt.$$

De alguma forma, pode-se dizer que  $f$  representa o valor da série formal  $\widehat{f}(x)$ . O valor numérico encontrado por Euler 0.596347362123 é o valor de  $f(1)$ .

Basta digitar o seguinte em seu computador

$$N[Exp[1] * Integrate[Exp[-1/t]/t, t, 0, 1], 100]$$

e obter aproximadamente

0.5963473623231940743410784993692793760741778601525487815734849104823272191148744174

Não totalmente de acordo com o resultado numérico de Euler.

### Cauchy

Então, chegou o período de desgraça para as séries divergentes. O novo mestre foi Augustin Cauchy, que definiu claramente o



Cum series convergentes ita definiantur, ut constant terminis continuo decrecentibus, qui tandem, si series in infinitum procederet, penitus evanescant; sic intelligitur, quomodo seriem terminis infinitis non in silibum abeat, sed vel finit maneat, vel in infinitum excreseat, eas, quae non sunt convergentes, ad classis serierum divergentium referri oportere. Propterea igitur termini seriei vltimi, ad quos progressionem in infinitum evanescens pertinet, fuerint vel magnitudo finita, vel infinita, duo habebuntur serierum divergentium genera, quorum utraque porro in duas species subdividitur, prout vel omnes termini eodem sint affecti signo, vel signa + et - alternatim se excipiant. Omnino ergo habebimus quatuor serierum divergentium species, ex quibus materiae perspicuitatis gratia aliquae exempla subiungam.



Ser cuidadoso! Eu certamente não gostaria de dar ao meu jovem leitor um sentimento errado de que qualquer série divergente tem uma soma bem definida. Algumas séries divergentes são de fato somáveis usando diferentes métodos que produzem somas diferentes. Um exemplo fácil, consiste em pensar na série  $1 - 1 + 1 - 1 + \dots$ .



Augustin Cauchy (1789–1857).

conceito de *convergência*, geralmente associado ao rigor matemático. Esse ponto de vista não é completamente errado, mas sem dúvida representa uma simplificação exagerada. Mas, por um lado, o rigor existia antes de Cauchy e, por outro lado, Cauchy não rejeitava completamente as séries divergentes <sup>57</sup>.

Infelizmente, mesmo hoje, muitos estudantes ainda estão convencidos de que as séries divergentes são coisas do satanás. . .

Em 1821, no prefácio de *Cours d'analyse*, Cauchy escreveu que foi *forçado* a abandonar as séries divergentes!

Tenho sido forçado a admitir algumas proposições que parecerão, talvez, difíceis de aceitar. Por exemplo, que uma série divergente não tem soma.

Em uma famosa carta à Holmboe, Abel escreveu em 1826:

As séries divergentes são, em geral, algo fatal, e é uma vergonha basear qualquer prova nelas.

## Poincaré

O grande mestre seguinte foi Poincaré que compreendeu claramente que as séries divergentes não são apenas úteis, mas também necessárias para resolver questões naturais da mecânica celeste. Vou abster-me de discutir esses aspectos dinâmicos, embora sejam fascinantes e ligados à atividade atual de pesquisa.

Permita-me citar o segundo volume do tratado *Les méthodes nouvelles de la mécanique céleste*.

*Existe uma espécie de mal-entendido entre os geômetras e os astrônomos sobre o significado da palavra convergência. Os geômetras, preocupados com o rigor absoluto e não incomodados pelo comprimento dos cálculos inextricáveis concebidos, mas sem tentar realizá-los explicitamente, diriam que uma série é convergente quando a soma dos termos tende a um limite definido, mesmo se os primeiros termos diminuam muito lentamente. Pelo contrário, os astrônomos adquiriram o hábito de dizer que uma série converge quando, por exemplo, os primeiros 20 termos diminuem muito rapidamente, mesmo se os termos restantes crescerem para sempre. Assim, vamos dar um exemplo simples e considerar duas séries abaixo que têm como termo geral*

<sup>57</sup> A. Cauchy. *Sur un emploi légitime des séries divergentes*. *Comptes Rendus Académie des Sciences*, XVII:18–25, 1843.

*“J’ai été forcé d’admettre diverses propositions qui paraîtront peut-être un peu dures. Par exemple qu’une série divergente n’a pas de somme. . .”*

*Les séries divergentes sont en général quelque chose de bien fatal et c’est une honte qu’on ose y fonder aucune démonstration.* Volume 2 das Obras completas de Abel.

$$\frac{1000^n}{n!} \quad \text{and} \quad \frac{n!}{1000^n}.$$

Os geômetras vão dizer que a primeira série converge, e seus termos convergem rápido [...], mas a segunda série diverge [...]. Ao contrário, os astrônomos irão considerar a primeira série como divergente [...] e a segunda como convergente. E as duas regras são legítimas: a primeira no campo das pesquisas teóricas e a segunda nas aplicações numéricas.

Il y a entre les géomètres et les astronomes une sorte de mal-entendu au sujet de la signification du mot convergence. Les géomètres, préoccupés de la parfaite rigueur et souvent trop indifférents à la longueur de calculs inextricables dont ils conçoivent la possibilité, sans songer à les entreprendre effectivement, disent qu'une série est convergente quand la somme des termes tend vers une limite déterminée, quand même les premiers termes diminueraient très lentement. Les astronomes, au contraire, ont coutume de dire qu'une série converge quand les vingt premiers termes, par exemple, diminuent très rapidement, quand même les termes suivants devraient croître indéfiniment. Ainsi, pour prendre un exemple simple, considérons les deux séries qui ont pour terme général  $\frac{1000^n}{n!}$  et  $\frac{n!}{1000^n}$ . Les géomètres diront que la première série converge, et même qu'elle converge rapidement, [...] mais ils regarderont la seconde comme divergente [...]. Les astronomes, au contraire, regarderont la première série comme divergente, [...] et la seconde comme convergente [...]. Les deux règles sont légitimes : la première, dans les recherches théoriques ; la seconde, dans les applications numériques.

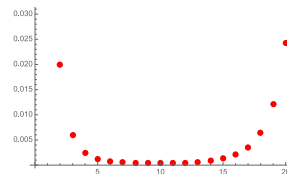
O exemplo de Poincaré é perfeito: olhe para os valores de  $\frac{n!}{1000^n} x^n$  e  $\frac{1000^n}{n!} x^n$  para  $x = .01$  and  $n = 1, \dots, 20$ .

### O sela-nó e a equação de Euler

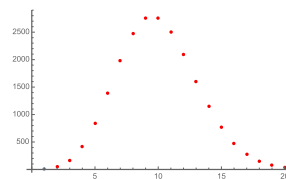
Vejamos um exemplo muito simples que mostra que não há escolhas: se queremos entender equações diferenciais ordinárias, mesmo com coeficientes polinomiais, nós *temos* que lidar com séries divergentes.

Considere o simples sistema a seguir:

$$\frac{dx}{dt} = x^2 \quad \frac{dy}{dt} = -y + x.$$



$\frac{n!}{10^n}$  for  $n = 1, \dots, 20$ .



$\frac{10^n}{n!}$  for  $n = 1, \dots, 20$ .

Ele é chamado de *sela-nó* porque parece realmente uma sela para  $x > 0$  e um nó quando  $x < 0$ . Pode-se pensar que essa é uma situação muito degenerada, mas aparece em *codimensão 1*: na origem, a parte linear do campo vetorial tem um autovalor que é nulo. Portanto, devemos esperar encontrar nós de sela semelhantes em um parâmetro famílias de campos vetoriais no plano.

A imagem na margem mostra o retrato de fase desse campo vetorial. Claramente, vemos uma curva invariante suave que passa pela origem. Isso é chamado de *variedade central*.

Considerando essa curva como um gráfico  $y(x)$ , obtemos imediatamente a equação de Euler  $x^2 \frac{dy}{dx} + y = x$ . Então, a equação da principal variedade é uma função  $C^\infty$  definida por

$$y(x) = \exp\left(\frac{1}{x}\right) \int_0^x \frac{1}{t} \exp\left(-\frac{1}{t}\right) dt,$$

E temos que entender como essa função está relacionada com a série formal divergente  $\hat{f}$ .

*Função de Euler, fenômeno de Stokes etc.*

Aqui sigo a apresentação feita por Hardy <sup>58</sup>.

Fazendo a mudança de variáveis  $t = x/(1+xw)$ , segue que

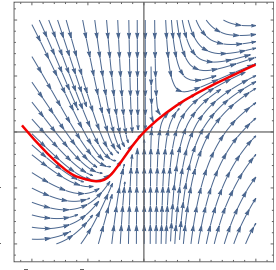
$$f(x) = x \int_0^\infty \frac{\exp(-w)}{1+xw} dw.$$

Assim,

$$\begin{aligned} f(x) &= \int_0^\infty \exp(-w) (x - x^2w + x^3w^2 - \dots + (-1)^{n-1}x^n w^{n-1}) dw \\ &\quad + (-1)^n x^{n+1} \int_0^\infty \frac{\exp(-w)w^n}{1+xw} dw \\ &= x - 1!x^2 + 2!x^3 - \dots + (-1)^{n-1}(n-1)!x^n + R_n(x). \end{aligned}$$

O termo  $R_n$  é facilmente majorado. Se  $x, w > 0$ , temos que  $1+xw > 1$ , e, então, obtemos

$$|f(x) - (x - 1!x^2 + 2!x^3 - \dots + (-1)^{n-1}(n-1)!x^n)| \leq n!x^{n+1}.$$



Retrato de fase do nó da sela.



Cuidado: mudanças de variáveis perigosas!!

<sup>58</sup> G. H. Hardy. *Divergent series*. Éditions Jacques Gabay, Sceaux, 1992. Com um prefácio por J. E. Littlewood e uma nota por L. S. Bosanquet, Reimpressão da edição revisada de 1963.



“Os jovens devem provar teoremas, os velhos devem escrever livros”. (Hardy 1877–1947)



Em outras palavras, a série formal  $\widehat{f}$  é assintótica à função  $f$  que é  $C^\infty$ .

Na verdade, podemos ser muito mais precisos. Suponha agora que  $x$  é um número complexo que não é um número real negativo. Então, a fórmula que define  $f$  faz todo o sentido porque  $f$  é uma função holomorfa em  $\mathbb{C} \setminus \mathbb{R}_-$ . Suponhamos agora restringir  $x$  a uma região onde seu argumento está em  $[-\pi + \delta, \pi - \delta]$  para algum  $\delta > 0$ . Nessa região, podemos minorar  $|1 + xw|$  ( $w$  ainda é um número real positivo) para que possamos obter alguma desigualdade

$$|f(x) - (x - 1)!x^2 + 2!x^3 - \dots + (-1)^{n-1}(n-1)!x^n| \leq C(\delta)n!|x|^{n+1}$$

Para  $x$  nessa região. Dito de forma diferente, a série formal  $\widehat{f}$  é assintótica à função holomorfa  $f$  em qualquer região sem conter a parte negativa do eixo real.

Mais ainda pode ser dito. Continuemos com a apresentação de Hardy das manipulações de Euler.

$$\begin{aligned} f(x) &= \exp\left(\frac{1}{x}\right) \int_0^x \frac{\exp\left(\frac{1}{t}\right)}{t} dt = \exp\left(-\frac{1}{x}\right) \int_{\frac{1}{x}}^\infty \frac{\exp(-u)}{u} du \\ &= -\exp\left(\frac{1}{x}\right) li\left(\exp\left(-\frac{1}{x}\right)\right) \end{aligned}$$

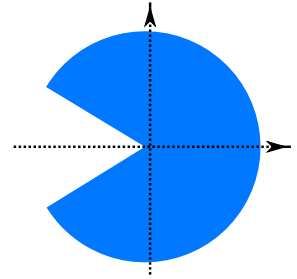
sendo  $li$  a integral logarítmica definida para  $0 < v < 1$  por

$$\begin{aligned} li(v) &= \int_0^v \frac{dt}{\log t} = - \int_{\log \frac{1}{v}}^\infty \frac{\exp(-u)}{u} du. \\ -li(\exp(-y)) &= \int_y^\infty \frac{\exp(-u)}{u} du \\ &= \int_1^\infty \frac{\exp(-u)}{u} du - \int_0^1 \frac{1 - \exp(-u)}{u} du \\ &\quad - \int_1^y \frac{du}{u} + \int_0^y \frac{1 - \exp(-u)}{u} du \\ &= -\gamma - \log y + y - \frac{1}{2 \cdot 2!} y^2 + \frac{1}{3 \cdot 3!} y^3 - \dots \end{aligned}$$

sendo  $\gamma \dots$  a constante de Euler. Assim, segue que

$$f(x) = \exp\left(\frac{1}{x}\right) \log x + S\left(\frac{1}{x}\right)$$

Lembre que uma série  $\sum_k a_k x^k$  é assintótica a uma função  $f(x)$  se para todo  $n$ , temos que  $f(x) - \sum_{k=1}^n a_k x^k = O(x^n)$ .



Pacman.

Note que  $li(x) = \int_0^\infty \frac{dx}{\log x}$  também é definida para  $x > 1$  como a integral imprópria:

$$\lim_{\epsilon \rightarrow 0} \left( \int_0^{1-\epsilon} + \int_{1+\epsilon}^x \right) \frac{dt}{\log t}$$

Ela é famosa na Teoria dos Números por fornecer uma estimativa bem precisa para a função  $\pi(x)$ , que representa a quantidade de números primos  $\leq x$  quando  $x$  tende ao infinito. Em particular, sabe-se que

$$|\pi(x) - li(x)| = o\left(\frac{x}{(\log x)^N}\right)$$

para todo  $N \geq 1$ .

com

$$S(y) = -\exp(y) \left( \gamma - y + \frac{1}{2 \cdot 2!} y^2 - \frac{1}{3 \cdot 3!} y^3 + \dots \right).$$

Note que  $S(y)$  é uma função inteira, i.e. holomorfa e uniforme em todo o plano complexo.

Isso fornece alguma extensão holomorfa de  $f$  no recobrimento universal de  $\mathbb{C} \setminus \{0\}$ . À medida que dá uma volta em torno da origem, a função muda de  $2i\pi \exp(\frac{1}{x})$ .

Vamos resumir as propriedades de  $f$ .

- É uma função holomorfa multivalente que é definida em todo o plano, ou mais precisamente uma função holomorfa na superfície de Riemann do logaritmo.
- Num setor de ângulo  $< 2\pi$ , a função  $f$  é assintótica à série formal  $\widehat{f}(x) = x - 1!x^2 + 2!x^3 - \dots$ .
- A monodromia, ou seja, a mudança no valor de  $f(x)$  como  $x$  circunda a origem, é  $2i\pi \exp(\frac{1}{x})$ , sendo plana em qualquer setor de ângulo  $< 2\pi$ , significando que quaisquer duas determinações de um único valor de  $f$  em um setor têm a mesma expansão assintótica  $\widehat{f}$ .

A divergência da série formal  $\widehat{f}$  corresponde ao fato de que a função  $f$  é multivalorada a menos de uma função plana. Essa não é a primeira vez que se explica um fenômeno no domínio real por outro no domínio complexo. Isso é chamado de *fenômeno Stokes* (descoberto em 19 de março de 1857 às 3 da manhã).

É claro que o exemplo de Euler é apenas um exemplo. O fato notável é que o caso foi tão significativo que implicou em uma bela teoria desenvolvida. Equações diferenciais algébricas ou analíticas podem ter soluções que são séries divergentes, mas, mesmo divergentes, pode-se dar um significado perfeitamente bem definido à sua soma, como funções holomorfas *multivalentes*.

Eu me abstenho de continuar nessa direção, já que o nosso passeio não iria nos levar para onde eu planejo ir. Mesmo em passeios é bom navegar para algum tipo de lugar bem definido.

Para uma descrição fascinante do desenvolvimento histórico da teoria, eu recomendo dois artigos <sup>59</sup>. Para uma descrição mais

Como esse capítulo discute séries divergentes, pode ser uma boa ideia mencionar a expansão assintótica divergente

$$li(x) = \frac{x}{\log x} \left( \sum_{k=0}^{\infty} \frac{k!}{(\log x)^k} \right)$$

, que você querido leitor, deve ser capaz de provar por si mesmo, integrando por partes.

Uma “Uma função holomorfa na superfície de Riemann do logaritmo” é uma maneira antiquada de falar de  $\varphi(\log z)$  sendo  $\varphi$  a função holomorfa no plano complexo. Ela é multivalente, pois  $\log z$  é definida “a menos de  $2i\pi$ ”.

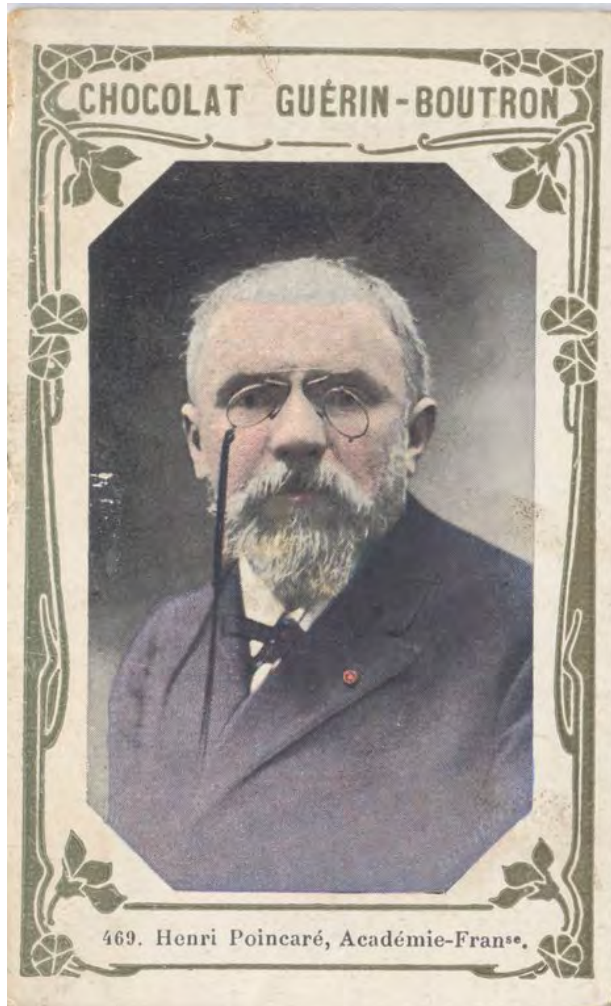
<sup>59</sup> J.-P. Ramis. [Poincaré et les développements asymptotiques \(première partie\)](#). *Gaz. Math.*, 133:33–72, 2012.; and J.-P. Ramis. [Les développements asymptotiques après Poincaré: continuité et divergences](#). *Gaz. Math.*, 134:17–36, 2012.

sistemática e acessível, sugiro as notas <sup>60</sup>, que poderão ser úteis.

Além disso, destaco o trabalho <sup>61</sup>, uma apresentação moderna das séries divergentes publicadas em comemoração ao aniversário de 300 anos de Leonhard Euler.

<sup>60</sup> M. Loday-Richaud. *Divergent series, summability and resurgence. II*, volume 2154 of *Lecture Notes in Mathematics*. Springer, [Cham], 2016. Simple and multiple summability, With prefaces by Jean-Pierre Ramis, Éric Delabaere, Claude Mitschi and David Sauzin.

<sup>61</sup> V. S. Varadarajan. *Euler and his work on infinite series*. *Bull. Amer. Math. Soc. (N.S.)*, 44(4):515–539, 2007.



Henri Poincaré (1854–1912).  
Quantos matemáticos  
chegaram a ser tão famosos  
em vida que sua fotografia  
foi impressa em barras de  
chocolate?



Augustin Cauchy.  
(1789 –1857).

+

# Convergência

## *Le calcul des limites de Cauchy*

EU PROVAREI AGORA O CHAMADO “TEOREMA DE PUISEUX” por meio de uma parametrização local de uma curva algébrica complexa numa vizinhança de um ponto singular, em termos de séries de potências *convergentes*. Não seguirei a abordagem original de Puiseux. Em vez disso, usarei um método introduzido por Cauchy sob o nome de “Calcul des limites”<sup>62</sup>.

### *O teorema da função implícita*

Começarei com uma prova *à la Cauchy* do teorema clássico da função implícita. Essa costumava ser a prova padrão nos livros-textos antigos, mas é frequentemente ignorada atualmente e substituída por métodos mais poderosos, baseados em teoremas de ponto fixo. No entanto, tem algumas vantagens: é elementar e quase inteiramente combinatória. Recomendo esse livro<sup>63</sup> para uma abordagem histórica interessante.

Denotemos por  $K$  um corpo de característica 0, equipado com uma norma, aplicação  $x \in K \mapsto |x| \in \mathbb{R}_+$  tal que  $|1| = 1$ ,  $|xy| = |x||y|$  e  $|x + y| \leq |x| + |y|$ . Assumimos que  $|x| = 0 \iff x = 0$  e que  $K$  equipado com  $|\cdot|$  é completo: seqüências de Cauchy convergem.

Eu tenho em mente basicamente os casos  $\mathbb{C}$  e  $\mathbb{R}$ , mas existem muitos outros exemplos (corpos  $p$ -ádicos em particular).

<sup>62</sup> A. Cauchy. *Œuvres complètes*, volume 12. Gauthier Villars, 1882.

A palavra “limite” não deve ser entendida como “limite” mas como “cota”.

<sup>63</sup> S. G. Krantz and H. R. Parks. *The implicit function theorem, history, theory, and applications*. Birkhäuser Boston, Inc., Boston, MA, 2002.

Assumiremos também que a norma não é a norma trivial, com  $|x| = 1$  para todos elementos não nulos  $x$ .

Denotemos por  $K\{x\}$  o anel das séries

$$f(x) = \sum_{k \geq 0} u_k x^k$$

para as quais os  $u_k$  em  $K$  e satisfazem alguma desigualdade da forma

$$|u_k| \leq Cr^k$$

para determinados  $C, r > 0$ . Uma vez que  $K$  é completo, isso corresponde a séries que são absolutamente convergentes em alguma vizinhança do 0 (germes de *funções analíticas* para  $\mathbb{C}, \mathbb{R}$ ). Por simplicidade, diremos que os elementos de  $K\{x\}$  são *séries convergentes*.

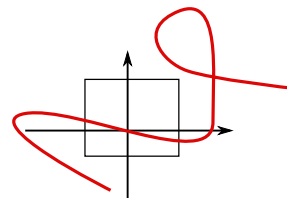
Similarmente, denotamos por  $K\{x, y\}$  o anel das séries convergentes

$$F(x, y) = \sum_{i, j \geq 0} a_{ij} x^i y^j$$

para as quais os  $a_{ij}$  estão em  $K$  e existem  $C, r > 0$  tais que para todos  $i, j$ :

$$|a_{ij}| \leq Cr^{i+j}.$$

**Teorema** (da função implícita). *Seja  $F \in K\{x, y\}$  tal que  $F(0, 0) = 0$  e  $\partial F / \partial y(0, 0) \neq 0$ . Então, existe uma série convergente  $f(x) \in K\{x\}$  tal que  $f(0) = 0$  e  $F(x, f(x)) = 0$ . As soluções  $(x, y)$  da equação  $F(x, y) = 0$  numa vizinhança de  $(0, 0)$  em  $K^2$  são precisamente os pares  $(x, f(x))$ .*



A prova é a seguinte: Se substituirmos a série formal  $y = \sum_{k \geq 1} u_k x^k$  na série formal  $\sum_{i, j \geq 0} a_{ij} x^i y^j$  (com  $a_{00} = 0$ ), obtemos uma série formal  $\sum_{l \geq 1} v_l x^l$  cujos coeficientes dependem dos  $u_k$  e dos  $a_{ij}$ . Vamos computar os primeiros termos

$$\sum_{i, j} a_{ij} x^i \left( \sum_{k \geq 1} u_k x^k \right)^j = \sum_{l \geq 1} v_l x^l.$$

Encontramos

$$v_1 = a_{10} + a_{01}u_1$$

$$v_2 = a_{20} + a_{11}u_1 + a_{02}u_1^2 + a_{01}u_2$$

$$v_3 = a_{30} + a_{21}u_1 + a_{12}u_1^2 + a_{03}u_1^3 + a_{11}u_2 + 2a_{02}u_1u_2 + a_{01}u_3$$

$$v_4 = a_{40} + a_{31}u_1 + a_{13}u_1^3 + a_{04}u_1^4 + a_{21}u_2 + 2a_{12}u_1u_2 + 3a_{03}u_1^2u_2 + a_{02}u_2^2 + a_{11}u_3 + 2a_{02}u_1u_3 + a_{01}u_4$$

etc.

Mesmo sendo complicado, prova-se imediatamente por indução que  $v_l$  se escreve como

$$v_l = G_l((a_{ij})_{i+j \leq l}, (u_k)_{k \leq l-1}) + a_{01}u_l$$

sendo  $G_l$  uma expressão polinomial com coeficientes inteiros positivos envolvendo os  $a_{ij}$  para  $i + j \leq l$  e os  $u_k$  para  $k \leq l - 1$ . Nosso problema é mostrar que dada uma série *convergente*  $F$ , existe uma única  $f(x)$  *convergente* tal que  $F(x, f(x)) = 0$ . Em outras palavras, nos são dados  $a_{ij}$  tais que  $|a_{ij}| \leq Cr^{i+j}$  e queremos mostrar que as equações  $v_l = 0$  com incógnitas  $u_l$  têm uma única solução convergente.

Pelas nossas hipóteses,  $a_{01} \neq 0$ , de modo que, multiplicando  $F$  por  $-1/a_{01}$ , podemos supor que  $a_{01} = -1$ . Do mesmo modo, mudando  $x, y$  por múltiplos constantes, podemos assumir que  $C = 1$  e  $r = 1$ . Em outras palavras, assumimos que  $|a_{ij}| \leq 1$  para todos  $i, j \geq 0$ .

Como  $G_l$  depende apenas dos  $u_k$  para  $k \leq l - 1$  (e dos  $a_{ij}$ ), as fórmulas anteriores definem indutivamente uma única série  $u_l$  (dependendo dos  $a_{ij}$ ):

$$\begin{aligned} u_1 &= a_{10} \\ u_2 &= a_{20} + a_{11}u_1 + a_{02}u_1^2 = a_{20} + a_{11}a_{10} + a_{02}a_{10}^2 \\ &\dots \\ u_l &= G_l((a_{ij})_{i+j \leq l}, (u_k)_{k \leq l-1}). \end{aligned}$$

Nossa tarefa é mostrar que essa série  $\sum_l u_l x^l$  é convergente.

Agora vem a ideia simples e bela de Cauchy<sup>64</sup>. Verificaremos o teorema em *um* exemplo específico e, então, mostraremos que isso implica no caso geral.

Para esse exemplo, escolhemos  $\bar{F}$  tal que  $\bar{a}_{01} = -1$  e todos os outros  $\bar{a}_{ij} = 1$ :

$$\bar{F}(x, y) = \frac{1}{(1-x)(1-y)} - 1 - 2y = -y + x^2 + xy + y^2 + x^3 + x^2y + xy^2 + y^3 + \dots$$

Seja  $\bar{u}_l$  a sequência correspondente associada a essa escolha de  $\bar{F}$  definida por:

$$\bar{u}_l = G_l((1), (\bar{u}_k)_{k \leq l-1}) \quad k = 1, 2, \dots$$

<sup>64</sup> U. Bottazzini and J. Gray. *Hidden harmony—geometric fantasies. Sources and Studies in the History of Mathematics and Physical Sciences*. Springer, New York, 2013.



Sabemos resolver  $\bar{F}(x, y) = 0$  uma vez que a equação

$$\frac{1}{(1-x)(1-y)} - 1 - 2y = 0$$

é equivalente a  $y = \frac{1}{4} \left( 1 \pm \sqrt{\frac{1-9x}{1-x}} \right)$ . Na vizinhança do 0, deve-se escolher o sinal -- e obtemos uma única solução analítica

$$\begin{aligned} y &= f(x) = \frac{1}{4} \left( 1 - \sqrt{\frac{1-9x}{1-x}} \right) \\ &= \bar{u}_1 x + \bar{u}_2 x^2 + \dots \\ &= x + 3x^2 + 13x^3 + 71x^4 + 441x^5 + 2955x^6 + \dots \end{aligned}$$

Os coeficientes  $\bar{u}_k$  obviamente satisfazem alguma desigualdade  $\bar{u}_k \leq c\rho^k$  uma vez que  $f$  é analítica em alguma vizinhança de 0.

Agora estudamos o caso de uma  $F$  geral para a qual assumimos  $|a_{ij}| \leq 1$ . Dado que os polinômios  $G_l$  têm *coeficientes inteiros positivos*, obtemos por indução que  $|u_l| \leq \bar{u}_l$ . De fato:

$$\begin{aligned} |u_{l+1}| &= |G_l((a_{ij})_{i+j \leq l+1}, (u_k)_{k \leq l})| \\ &\leq |G_l((1), (|u_k|)_{k \leq l})| \\ &\leq |G_l((1), (\bar{u}_k)_{k \leq l})| \\ &\leq \bar{u}_{l+1} \end{aligned}$$

Em particular,  $|u_l| \leq c\rho^l$  e a série  $f(x) = \sum_k u_k x^k$  é convergente.

A prova do teorema está quase concluída. Encontramos uma solução convergente  $y = f(x)$  e ainda temos que mostrar que todas as soluções de  $F(x, y) = 0$  na vizinhança da origem são da forma  $(x, f(x))$ .

No anel  $K\{x, y\}$ , é claro que um elemento  $F(x, y)$  é divisível por  $y$  se, e somente se, ele se anula quando substituimos  $y$  por 0. A transformação  $(x, y) \mapsto (x, y - f(x))$  induz um automorfismo de  $K\{x, y\}$ , levando  $y$  em  $(y - f(x))$ . Sabemos que  $y = f(x)$  é solução de  $F(x, y) = 0$  de modo que a observação anterior implica que  $F$  é divisível por  $y - f(x)$  in  $K\{x, y\}$ . O quociente é não nulo em  $(0, 0)$  uma vez que

$$F(x, y) = -y + a_{10}x + \dots \quad \text{e} \quad f(x) = a_{10}x + \dots$$

Temos, portanto, que

$$F(x, y) = U(x, y)(y - f(x))$$

sendo  $U \in K\{x, y\}$  tal que  $U(0, 0) \neq 0$ . Em particular, na vizinhança de  $(0, 0)$ , a equação  $F(x, y) = 0$  é de fato equivalente a  $y = f(x)$ . O teorema está provado.  $\square$

Para uma boa descrição de *la méthode des limites*, recomendo o livro de Hille <sup>65</sup>.

<sup>65</sup> E. Hille. *Ordinary differential equations in the complex domain*. Dover Publications, Inc., Mineola, NY, 1997. Reimpressão do original de 1976.

### Teorema de Puiseux

Lembre que já resolvemos equações implícitas da forma  $F(x, y) = 0$  quando  $F$  é uma série *formal* não trivial em  $K[[x, y]]$  onde  $K$  é um corpo algebricamente fechado.

Mostramos (com ajuda de Newton e Cramer) que qualquer elemento não nulo  $F$  em  $K[[x, y]]$  pode ser decomposto como um produto de séries de Puiseux formais:

$$F = A(x, y)x^r(y - f_1(x))(y - f_2(x)) \cdots (y - f_n(x))$$

sendo  $A(0, 0) \neq 0$  e as  $n$  soluções  $f_i(x)$  *séries de Puiseux* formais em  $K[[x^*]]$ .

Nosso objetivo agora é mostrar que se  $F$  é uma série convergente, também serão as  $f_i(x)$ .

*Assumimos agora que  $K$  seja um corpo algebricamente fechado equipado com uma norma completa*. Uma vez que todo elemento  $f(x)$  em  $K[[x^*]]$  está num anel  $K[[x^{1/N}]]$  para algum  $N$ , i.e. é uma série na variável  $x^{1/N}$ , não há dificuldade em definir séries de Puiseux convergentes.

Denotamos por  $K\{x^*\}$  e  $K\{x^*, y^*\}$  os anéis das séries de Puiseux convergentes em uma e duas variáveis.

Mesmo que séries em  $K\{x^*\}$  convirjam, devemos ser cuidadosos: não são funções de fato definidas na vizinhança de 0. São funções “multivalentes” de  $x$ .

**Teorema (de Puiseux).** *Qualquer elemento não nulo  $F$  em  $K\{x, y\}$  pode ser decomposto como*

$$F = A(x, y)x^r(y - f_1(x))(y - f_2(x)) \cdots (y - f_n(x))$$

sendo  $A(0, 0) \neq 0$  e as  $n$  soluções  $f_i(x)$  séries de Puiseux convergentes em  $K\{x^*\}$ .

O completamento do fecho algébrico do corpo dos números  $p$ -ádicos é um bom exemplo.

Espero que meu leitor tenha adivinhado a definição do anel das séries de Puiseux convergentes em duas variáveis: basta considerar séries de potências convergentes em  $(x, y)$  e substituir formalmente  $x$  e  $y$  por  $x^{1/m}$  e  $y^{1/n}$ .

A prova pode parecer um pouco complicada, mas o leitor dever ter em mente que esse teorema é uma leve generalização do teorema da função implícita.

Lembremos dos seguintes fatos:

1. Se a série de Puiseux formal  $y = f(x)$  é uma raiz de  $F(x, y) = 0$ , ela é obtida a partir de uma aplicação do algoritmo de Newton, fazendo-se uma escolha do lado do polígono de Newton a cada passo.
2. A cada passo do algoritmo, define-se  $x_k = x_{k+1}^{\alpha_k}$  e  $y_k = u_k x_k^{\beta_k} (1 + y_{k+1})$  para determinados inteiros positivos  $\alpha_k, \beta_k$ , e substitui-se  $F_k(x_k, y_k)$  por  $F_{k+1}(x_{k+1}, y_{k+1}) = x_{k+1}^{-\gamma_k} F_k(x_k, y_k)$  (para um determinado inteiro positivo  $\gamma_k$ ). Portanto, a série  $f(x)$  pode ser também descrita como a série  $y_k(x_k)$  ( $k \geq 1$ ). Claramente, é equivalente provar a convergência de  $f(x)$  ou de qualquer uma das  $y_k(x_k)$ .
3. Após um certo número de passos, as multiplicidades de  $F_k(x_k, y_k)$  (i.e. as valorações de  $F_k(0, y_k)$ ) se mantêm igual a uma "multiplicidade final"  $m \geq 1$  (teorema de Cramer).
4. Essa "constante final"  $m$  associada à raiz  $y = f(x)$  de  $F(x, y) = 0$  é também a multiplicidade da raiz, em outras palavras, é o número de fatores iguais a  $(y - f(x))$  aparecendo na decomposição de  $F$ .

Podemos terminar a demonstração do teorema de Puiseux:

Seja  $F$  em  $K\{x, y\}$  e seja

$$F = A(x, y)x^r(y - f_1(x))(y - f_2(x)) \cdots (y - f_n(x))$$

sua decomposição em séries de Puiseux formais. Escolha  $N$  tal que todas as  $f_i(x)$  pertençam a  $K[[x^{1/N}]]$  e tome  $\bar{x} = x^{1/N}$  de modo que  $F$  também possa ser visto como um elemento de  $K\{\bar{x}, y\}$  e todas as  $f_i$  como elementos de  $K[[\bar{x}]]$ .

Reduzimos ao caso em que  $F$  está em  $K\{x, y\}$  e todas as  $f_i(x)$  são séries formais em  $K[[x]]$  e temos que provar que essas  $f_i$  são na verdade séries convergentes, i.e. pertencem a  $K\{x\}$ .

Se a "multiplicidade final"  $m$  de uma raiz  $y = f(x)$  é igual a 1, sabemos que o caminho seguido pelo algoritmo de Newton

levando à solução  $f(x)$  trará finalmente uma  $F_k(x_k, y_k)$  com multiplicidade 1. O teorema da função implícita aplicado a  $F_k$  convergente mostra que a solução  $f(x)$  é também convergente.

Se um polinômio tem raiz múltipla, essa raiz é também raiz de sua derivada. Em nosso contexto, isso significa que se  $f(x)$  é uma série formal em  $K[[x]]$ , solução de  $F(x, y) = 0$  com multiplicidade  $m \geq 2$ , então, a mesma série é uma solução de  $\partial F / \partial y(x, y) = 0$  com uma multiplicidade menor. Evidentemente, se  $F(x, y)$  é convergente, sua derivada parcial com respeito a  $y$  também será. Uma simples indução termina a prova.  $\square$

### Corolários

Fizemos a maior parte do trabalho. É hora da sobremesa!

Primeiro, obtemos os mesmos corolários que tínhamos para séries formais com as mesmas provas.

*Continuamos a assumir que  $K$  é um corpo algebricamente fechado, com característica 0 e equipado com uma norma completa.*

**Teorema** (Teorema de preparação de Weierstrass). *Seja  $F(x, y)$  uma série convergente no anel  $K\{x, y\}$  com multiplicidade  $\text{mult}(F)$  e que não é divisível por  $x$ . Então, podemos escrever  $F$  como um produto  $A(x, y)P(x, y)$  estando  $A, P$  em  $K\{x, y\}$  e*

- $A(0, 0) \neq 0$  de modo que  $A$  é um elemento invertível.
- $P(x, y)$  é um polinômio em  $y$  de grau  $\text{mult}(F)$ .  $\square$

**Teorema.** *O anel  $K\{x, y\}$  é um domínio de fatoração única.*  $\square$

Uma formulação bastante útil do teorema de Puiseux é dada em termos de parametrizações.

**Teorema** (Parametrizações de Puiseux). *Seja  $F(x, y)$  uma série convergente não nula no anel  $K\{x, y\}$ , anulando-se na origem e não divisível por  $x$ . Então, existem*

1. inteiros  $n_i \geq 1$
2. conjuntos abertos  $U_i \subset K$ , contendo 0 (na topologia definida pela norma)

3. séries  $g_i \in K\{x\}$  convergindo em  $U_i$

tais que a interseção da curva  $F(x, y) = 0$  com uma vizinhança pequena de  $(0, 0) \in K^2$  é a união das imagens de aplicações

$$\varphi_i : t \in U_i \mapsto (t^{n_i}, g_i(t)) \in K^2.$$

Mais ainda, essas aplicações  $\varphi_i$  são injetivas e suas imagens intersectam apenas na origem.

Se  $f(x) \in K\{x^*\}$  é uma solução de  $F(x, y) = 0$ , denotamos por  $n$  o menor inteiro tal que  $f(x) \in K\{x^{1/n}\}$ . Sabemos que isso define  $n$  conjugados distintos pelo grupo de Galois  $f_1(x), \dots, f_n(x)$  sob a ação das raízes  $n$ -ésimas da unidade. Nenhuma dessas  $f_i(x)$  é uma “função” de  $x$  no sentido usual. No entanto, a escolha de alguma raiz  $n$ -ésima de  $x$  define  $n$  valores para  $f_i(x)$ . Mudando a raiz de  $x$  há apenas a permutação dos valores de  $f_i(x)$ . Dito de outra forma, existe uma  $g(t) \in K\{t\}$  convergente tal que esses  $n$  valores são os  $n$  valores de  $g(\sqrt[n]{x})$  para as  $n$  possibilidades de  $\sqrt[n]{x}$ . Todos esses pontos estão parametrizados numa vizinhança de 0 por:

$$\varphi : t \in U \mapsto (t^n, g(t)) \in K^2.$$

Obtemos, desta forma, um número finito de  $\varphi_i$  como no teorema cujas imagens recobrem o lugar geométrico dos zeros  $F$  (sempre numa vizinhança da origem).

Agora falta mostrar que as  $\varphi_i$  são injetivas e que elas se intersectam apenas na origem.

É conhecido que os zeros de uma função analítica são isolados. O teorema seguinte diz simplesmente que isso é verdade em  $K\{x\}$  para um  $K$  geral. Deixamos a prova como exercício para você, prezado leitor.

**Lema.** *Seja  $h$  uma série convergente em  $K\{x\}$ . Se existir uma sequência  $(x_n)_{n \geq 0} \in K \setminus \{0\}$  convergindo para 0 tal que  $h(x_n) = 0$ , então,  $h = 0$ .  $\square$*

Suponha agora que  $\varphi$  não seja injetiva na vizinhança de 0. Isso implicaria que existe alguma raiz  $n$ -ésima da unidade  $\omega$  tal que as soluções de  $g(\omega t) = g(t)$  se acumulam em 0. De acordo com o

lema, teríamos que  $g(\omega t) = g(t)$  identicamente, contradizendo o fato de  $n$  ser o menor inteiro tal que  $f(x) \in K\{x^{1/n}\}$ .

O mesmo argumento mostra que as imagens de

$$\varphi_1 : t \in U \mapsto (t^{n_1}, g_1(t)) \in K^2 \quad \varphi_2 : t \in U_2 \mapsto (t^{n_2}, g_2(t)) \in K^2$$

se intersectam não trivialmente (i.e. suas interseções se acumulam na origem)  $n_1 = n_2 = n$  e  $g_2, g_1$  são conjugados pelo grupo de Galois, de modo que  $g_2(t) = g_1(\omega t)$ , identicamente, para alguma raiz  $n$ -ésima da unidade. Nesse caso, as duas imagens de fato coincidem na vizinhança do 0. □

As imagens das  $\varphi_i$  são normalmente chamadas de *ramos* da curva  $F(x, y) = 0$ . As parametrizações tipo Puiseux são únicas a menos da ação pelo grupo de Galois.

Em particular, uma vizinhança da origem em  $\{F(x, y) = 0\}$  é homeomorfa à união de um número finito de bolas em  $K$  se intersectando num número finito de pontos. Note que “uma bola” é um intervalo em  $\mathbb{R}$ , um disco em  $\mathbb{C}$  e um conjunto de Cantor para os números  $p$ -ádicos.

### Números reais

Até agora, assumimos que o corpo  $K$  é algebricamente fechado. Vamos estudar o caso dos números reais que, afinal, está na origem da nossa caminhada.

Seja  $F(x, y) \in \mathbb{R}\{x, y\}$  uma série convergente não-nula que se anula na origem. Podemos olhar para seu conjunto de zeros  $\{(x, y) \mid F(x, y) = 0\}$  como uma curva complexa em  $\mathbb{C}^2$  ou como uma curva real em  $\mathbb{R}^2$ , numa vizinhança de  $(0, 0)$ . Aqui, estamos principalmente interessados na descrição de uma curva real.

Sobre os números complexos, sabemos que esse conjunto de zeros é a união de alguns ramos parametrizados por:

$$\varphi_i : t \in U_i \mapsto (t^{n_i}, g_i(t)) \in \mathbb{C}^2.$$

Como  $F(x, y)$  tem coeficientes reais, o lugar geométrico de seus zeros em  $\mathbb{C}^2$  é globalmente invariante por conjugação complexa. Uma vez que ramos são disjuntos longe da origem, um ponto real diferente da origem deve pertencer a uma ramo que coincide

Para alguma  $F(x, y) \in \mathbb{R}\{x, y\}$ , pode acontecer que a parte real de seus zeros se reduzem à origem. O exemplo mais óbvio é  $x^2 + y^2 = 0$ . Sobre os números complexos, essa curva consiste em dois ramos imaginários  $y = ix$  and  $y = -ix$ , que se intersectam somente em  $(0, 0)$ . Naturalmente, como estamos apenas interessados na parte real dos conjuntos de zeros de  $F$ , poderíamos simplesmente descartar os fatores irredutíveis de  $F$  cujos zeros se reduzem à origem (sobre os reais).

com seu conjugado. O conjugado complexo da imagem de  $\varphi_i$  é a imagem de

$$\overline{\varphi_i}: t \in U_i \mapsto (t^{n_i}, \overline{g_i(\overline{t})}) \in \mathbb{C}^2.$$

Portanto, ramos contendo pontos reais diferentes da origem são tais que

$$\overline{g_i(\overline{t})} = g_i(\omega t)$$

para alguma raiz  $n_i$ -ésima da unidade  $\omega$ . Escrevendo

$$g_i(t) = \sum_{k \geq 1} a_k t^k,$$

essa condição significa que

$$\overline{a_k} = a_k \omega^k.$$

Seja  $\mu$  uma das duas raízes quadradas de  $\omega$  e seja  $t = \mu s$ . Então

$$t^{n_i} = \mu^{n_i} s^{n_i} = \pm s^{n_i}$$

e

$$g_i(t) = \sum_{k \geq 1} a_k t^k = \sum_{k \geq 1} a_k \mu^k s^k = \sum_{k \geq 1} b_k s^k.$$

Agora os coeficientes  $b_k$  são reais uma vez que

$$\overline{b_k} = \overline{a_k} \overline{\mu}^k = a_k \omega^k \mu^{-k} = a_k \mu^k = b_k.$$

Vamos resumir a discussão:

**Teorema.** *Seja  $F(x, y) \in \mathbb{R}\{x, y\}$  uma série convergente não nula com coeficientes reais, anulando-se na origem e não divisível por  $x$ . Assuma que o lugar geométrico dos zeros de  $F$  na vizinhança de  $(0, 0) \in \mathbb{R}^2$  não se reduz à origem. Então, esse conjunto de zeros é a união de um número finito de curvas da forma*

$$\varphi_i: t \in ]-\varepsilon_i, +\varepsilon_i[ \mapsto (\pm t^{n_i}, g_i(t)) \in \mathbb{R}^2.$$

sendo  $g_i$  uma série convergente com coeficientes reais. As  $\varphi_i$  são injetivas e suas imagens se intersectam apenas na origem.  $\square$

É fácil ver que essas curvas  $\varphi_i$  são transversais a um pequeno círculo centrado na origem. De fato, pontos de tangência correspondem aos zeros de

$$\frac{d}{dt}(t^{2n_i} + g_i^2(t)) = 2n_i t^{n_i-1} + 2g_i(t)g_i'(t)$$

que são isolados. Note que essa expressão não pode ser identicamente 0, caso contrário, a curva seria um círculo!

Provamos mais do que a afirmação de Gauss. Mostramos que, localmente, uma curva real analítica é feita de um número finito de ramos.

- Cada ramo é homeomorfo a  $] - \varepsilon, +\varepsilon[$ .
- Cada ramo é transversal a círculos pequenos.
- Cada ramo intersecta círculos pequenos em dois pontos (um para  $t > 0$  e outro para  $t < 0$ ).
- Dois ramos distintos se intersectam apenas na origem.
- Ao longo de um ramo  $y/x$  converge a um limite em  $\mathbb{R} \cup \{\infty\}$  quando  $t$  tende a zero. Isso significa que cada ramo tem uma tangente bem definida na origem.

Aqui temos um simples corolário, análogo ao fato de que todo polinômio real de grau ímpar tem uma raiz real.

*Seja  $F(x, y)$  uma série convergente não nula com coeficientes reais, se anulando na origem, não divisível por  $x$ , e com multiplicidade ímpar. Então, a curva real  $F(x, y) = 0$  não se reduz à origem. Para valores reais pequenos de  $x$ , existe ao menos uma solução real de  $F(x, y) = 0$ .*

Esse fato simples foi transformado numa ferramenta poderosa por Poincaré que a usou em inúmeras situações, como, por exemplo, para provar a existência de órbitas periódicas no problema de três corpos (vide <sup>66</sup> página 70). Esse é o seu “método de continuidade”.

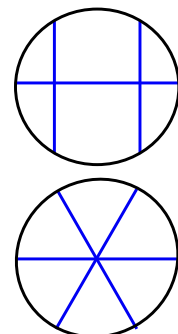
### Diagramas de Cordas

A topologia local de uma curva analítica na vizinhança de um ponto singular sugere a definição a seguir, muito importante para o restante deste nosso livro-passeio.

**Definição.** 1. Um *diagrama de cordas* é um conjunto de  $2n$  pontos no círculo equipado com alguma involução sem pontos fixos.

Em particular, uma curva algébrica não pode chegar a origem como uma espiral infinita.

<sup>66</sup> H. Poincaré. *Les méthodes nouvelles de la mécanique céleste*. Tome I. Les Grands Classiques Gauthier-Villars. Librairie Scientifique et Technique Albert Blanchard, Paris, 1987. Reimpressão do original de 1892.



Dois diagramas com três cordas.



Em outras palavras, uma coleção de  $2n$  pontos pareados dois a dois.

2. Dois diagramas de cordas são considerados *equivalentes* se existe um homeomorfismo do círculo que preserva orientação e leva o primeiro no segundo, comutando com a involução. Em outras palavras, consideramos uma palavra cíclica de  $2n$  letras na qual cada letra aparece exatamente duas vezes. Imaginamos cordas conectando os pares. Isso às vezes é também chamado de *palavra de Gauss* ou um *pareamento*, dependendo do contexto. Deve-se fazer uma escolha, e eu decido por “diagrama de cordas”.
3. O diagrama de cordas associado a uma curva analítica em algum ponto (singular) é o diagrama de cordas que se obtém ao intersectar-se a curva com um círculo pequeno ao redor do ponto, onde os pontos pareados correspondem aos ramos. Diremos que um tal diagrama de cordas é *analítico*.

Gostaria de entender diagramas de cordas analíticos e a topologia das curvas reais analíticas?

Seja paciente! Chegaremos lá.

### *Uma controvérsia que diz respeito à forma dos bicos de pássaros?*

Em 1751, Euler escreveu um artigo muito interessante (em francês) sobre a geometria das curvas algébricas. Na introdução, ele menciona que

*Mesmo a Geometria não está isenta de controvérsias e aparentes contradições, embora nós frequentemente sustentemos o contrário.*

A controvérsia que Euler queria discutir diz respeito ao formato dos pontos cuspidais<sup>67</sup>. Houve uma discordância entre o Sr. le Marquis de l’Hôpital e o Sr. Guà de Malves. Euler agiu como juiz e dissipou a aparente contradição de modo brilhante.

Até o momento, discutimos apenas a *topologia* de ramos na vizinhança de um ponto singular e não dissemos muito sobre

Même la géométrie n’est pas exemte [sic] de controverses, & des contradictions apparentes, quoi qu’on soutienne souvent le contraire.

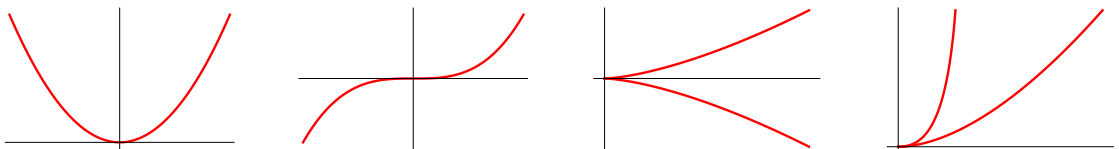
Será que deveria dar exemplos?

<sup>67</sup> L. Euler. Sur le point de rebroussement de la seconde espèce de M. le Marquis de l’Hôpital. *Mémoires de l’académie des sciences de Berlin*, 5:203–221, 1751. Veja o [Euler archive](#) para traduções para o Inglês e comentários.

sua *geometria*. Mencionamos apenas que um ramo tem uma tangente no ponto singular.

O livro de l'Hôpital é intitulado “*Analyse des infiniment petits pour l'intelligence des lignes courbes*” e foi publicado em 1696. É o primeiro livro-texto em cálculo diferencial. Contém uma classificação de ramos singulares de curvas algébricas em quatro categorias. Deixe-me expor em terminologia moderna: Escolha coordenadas de modo que a tangente seja  $y = 0$ . Localmente, nosso ramo é a união de dois “semirramos” que são gráficos de duas funções  $f_1(x), f_2(x)$ , definidas num pequeno intervalo da forma  $] - \varepsilon, 0]$  ou  $[0, \varepsilon[$ . Essas funções são suaves, longe da origem. Os quatro casos são:

1.  $f_1$  é definida em  $] - \varepsilon, 0]$  e  $f_2$  em  $[0, \varepsilon[$  e suas segundas derivadas têm mesmo sinal. Nesse caso, a curva é convexa (ou côncava) e está em um lado da sua tangente.
2. Um *ponto de inflexão*. O mesmo que antes exceto que as segundas derivadas têm sinais diferentes.
3. Um *ponto cuspidal* padrão. Aqui  $f_1$  e  $f_2$  estão definidas no mesmo lado da origem e suas segundas derivadas têm sinais opostos. Então, ambos os semirramos têm concavidades opostas.
4. Um *bico de pássaro* (“point de rebroussement à bec d’oiseau”) no qual as segundas derivadas têm o mesmo sinal nas duas metades de ramo.



É fácil encontrar exemplos das três primeiras categorias. Para a quarta categoria, l'Hôpital deu o seguinte exemplo: Enrole um fio em alguma curva com um ponto de inflexão e com uma extremidade presa em algum outro ponto. Quando você o

A fronteira entre geometria e topologia não é clara. Deixe-me dizer que a topologia lida com propriedades invariantes por homeomorfismos e a geometria, invariantes por... grupos menores, como, por exemplo, isometrias euclidianas, automorfismos projetivos ou simplesmente difeomorfismos.

desenrolar, mantendo-o apertado, o ponto terminal descreverá uma curva (chamada de *envolvente*) que apresenta o tal bico de pássaro.

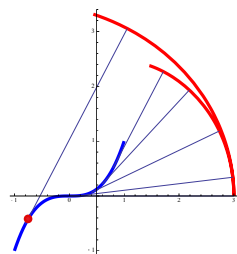
Escolhi simplesmente  $y = x^3$  como uma curva com inflexão e pedi ao meu computador que desenhasse a curva de l'Hôpital. O resultado está na margem. De fato, conforme afirmado, o ponto terminal do fio descreve a curva vermelha que apresenta um bico de pássaro quando o fio é tangente ao ponto de inflexão. Os semirramos têm a mesma concavidade. Para l'Hôpital, isso é "mecanicamente" óbvio.

Em 1740, o Sr. Guà de Malves publicou um [livro incrível](#)<sup>68</sup> cujo propósito era evitar as técnicas de Newton e usar apenas as de Descartes! Deve-se lembrar das controvérsias entre os ingleses e os franceses durante o século XVIII em torno de Descartes e Newton. Como uma ilustração dessa guerra anglo-francesa, eu recomendo as "[lettres sur Descartes et Newton](#)", por Voltaire.

*Um francês chegando a Londres encontrará a filosofia, assim como todo o resto, muito diferente lá. Ele deixou o mundo pleno, e o encontra vazio. Em Paris, vê-se o universo composto por vórtices de matéria sutil; em Londres, não se vê nada disso. Na França, é a pressão da Lua que causa as marés; mas na Inglaterra é o mar que gravita em direção à Lua; de modo que, quando vós acreditais que a lua deveria fazer a maré baixa, esses senhores acreditam que deveria ser maré alta, o que infelizmente não pode ser provado, pois seria necessário, para esclarecer isso, investigar a Lua e as marés no primeiro instante da criação.*

De qualquer forma, o livro de Guà é sobre um debate, ainda ativo: a geometria algébrica deveria usar ferramentas transcendentais da geometria diferencial? Entre os "teoremas" nesse livro, encontra-se a afirmação que l'Hôpital estava errado e que bicos de pássaro não existem.

Guà está ciente do exemplo de l'Hôpital, mas ele o critica do seguinte modo: Suponha que se tenha duas parábolas  $y = x^2$  e  $y = 2x^2$ , mas somente para  $x \geq 0$ . Obtêm-se duas metades de parábolas convexas cuja união se parece com um bico de pássaro. Portanto, de acordo com Guà, o bico que se observa no exemplo de l'Hôpital é artificial: a curva algébrica completa contém dois ramos suaves, como no exemplo das parábolas e na construção mecânica usando o fio, está faltando uma metade da



O comprimento do fio mais o comprimento curvilíneo ao longo da curva é constante.

<sup>68</sup> J.-P. Guà de Malves. *Usages de l'analyse de Descartes pour découvrir, sans le secours du Calcul Différentiel, les Propriétés, ou affectations principales des lignes géométriques de tous les ordres*. Briasson, 1740.

Un Français qui arrive à Londres trouve les choses bien changées en philosophie comme dans tout le reste. Il a laissé le monde plein; il le trouve vide. À Paris, on voit l'univers composé de tourbillons de matière subtile; à Londres, on ne voit rien de cela. Chez nous, c'est la pression de la lune qui cause le flux de la mer chez les Anglais, c'est la mer qui gravite vers la lune, de façon que, quand vous croyez que la lune devrait nous donner marée haute, ces Messieurs croient qu'on doit avoir marée basse; ce qui malheureusement ne peut se vérifier, car il aurait fallu, pour s'en éclaircir, examiner la lune et les marées au premier instant de la création.

curva algébrica. Convincente? Guà continua e “prova” que bicos de pássaro são impossíveis para uma curva algébrica.

A “prova” segue mais ou menos a seguinte linha: Um ramo tem a forma  $y = ax^{p/q} + o(x^{p/q})$  para algum par de inteiros  $p, q$  primos entre si com  $p > q$  se  $y = 0$  é a tangente em 0. Se existe um bico,  $q$  deve ser par, pois, do contrário,  $y$  estaria definido para todo  $x$ , positivo ou negativo. A concavidade é dada pelo sinal da segunda derivada, que é proporcional a

$$x^{(\frac{p}{q}-1)(\frac{p}{q}-2)} = x^{\frac{(p-q)(p-2q)}{q^2}}$$

Como  $q^2$  é par e  $(p - q)(p - 2q)$  é ímpar, as duas determinações da segunda derivada têm sinais diferentes, de modo que seus dois semirramos têm concavidades opostas: isso não é um bico de pássaro.

Agora vem o grande Euler: Seu artigo é muito claro e inquestionável. Inicialmente, ele foi convencido pelo argumento de Guà, mas encontrou um erro em 1744. Numa [coluna](#) intitulada “Did Euler prove Cramer’s rule”, Rob Bradley menciona uma carta entre Euler e Cramer, discutindo esse tópico.

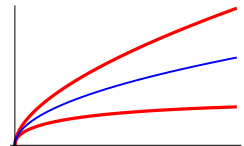
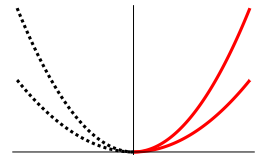
Euler, Guà e l’Hôpital usam livremente séries de Puiseux e não levantam dúvidas quanto a sua convergência. O que é notável no artigo de Euler é a descrição do papel dos números *complexos* quando se quer entender curvas algébricas *reais* (em 1751). Aqui está um dos exemplos de Euler:

$$y = x^{1/2} \pm x^{3/4}.$$

O gráfico na margem de fato se parece com um bico de águia! Como saber se esses gráficos, com sinal  $\pm$ , pertencem ao mesmo ramo e não podem ser completados como em nosso exemplo com duas parábolas? Euler dá um argumento convincente usando números complexos. Eu encorajo fortemente, meu leitor, a encontrar o erro na “prova” de Guà.

Pode-se também eliminar os radicais e escrever explicitamente uma equação polinomial. Euler encontra

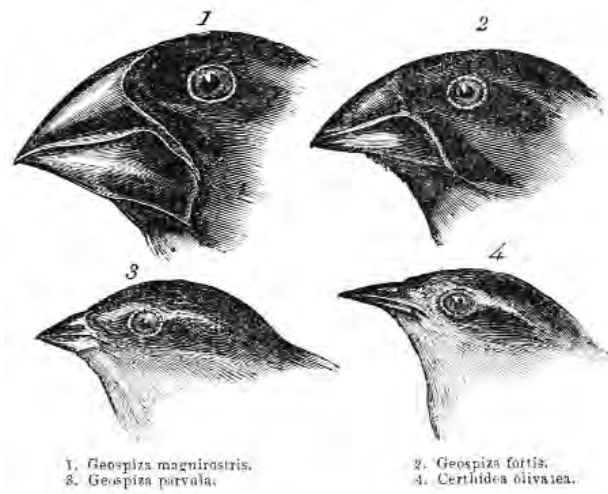
$$y^4 - 2xy^2 + x^2 - x^3 - 4yx^2 = 0$$



e eu lhe instigo a desenhar o polígono de Newton e checar que existe de fato um único ramo na origem.

Hoje não se menciona mais bicos de pássaros. Esses pontos são agora chamados de “cúspides de segunda ordem” de um jeito mais neutro. Às vezes, ainda se encontra o nome *curva ranfóide*, do grego “ραμφος” associado ao bico curvo de aves de rapina.

Para concluir este capítulo, deixo apenas um conselho: pare de ler este livro e vá ler os artigos de Euler. Agora!



O formato do bico de pardais das Ilhas Galápagos foi importante na descoberta de Darwin da evolução. †



Uma faixa de Möbius no saguão principal do IMPA, onde uma primeira versão desse livro foi escrita.

# Möbius e a sua faixa

ESSE É O TÍTULO DE UM LIVRO <sup>69</sup> dedicado à Matemática Alemã do século XIX. Neste capítulo, nós vamos discutir a topologia associada ao processo de *desingularização* de curvas analíticas, levando a alguns belos colares feitos de faixas Möbius.

## Coordenadas polares

Olhe para a seguinte aplicação familiar: um cilindro para um plano.

$$\Phi : (\rho, \theta) \in \mathbb{R} \times \mathbb{R}/2\pi\mathbb{Z} \mapsto (\rho \cos \theta, \rho \sin \theta) \in \mathbb{R}^2.$$

Ela tem as seguintes propriedades.

1.  $\Phi$  restrita a  $\mathbb{R}_+^* \times \mathbb{R}/2\pi\mathbb{Z}$  é um difeomorfismo sobrejetivo na região  $\mathbb{R}^2 \setminus \{(0,0)\}$ .
2.  $\Phi$  “colapsa” o círculo  $\{0\} \times \mathbb{R}/2\pi\mathbb{Z}$  para a origem.
3. A imagem inversa por  $\Phi$  de um ponto que é diferente da origem contém precisamente dois pontos, da forma  $(\rho, \theta)$  e  $(-\rho, \theta + \pi)$ .

A propriedade 3 não é muito conveniente para um sistema de coordenadas e essa é a principal razão pela qual vamos modificar ligeiramente  $\Phi$  em um dado momento. Às vezes, usa-se a restrição de  $\Phi$  à  $\mathbb{R}_+ \times \mathbb{R}/2\pi\mathbb{Z}$ , mas isso introduz uma fronteira artificial.



August Ferdinand Möbius  
(1825–1884).

<sup>69</sup> R. F. John Fauvel, Robin Wilson. *Moebius and his Band: Mathematics and Astronomy in Nineteenth-Century Germany*. OUP, 1993.





A propriedade 2 é interessante no contexto da *desingularização*. Numa pequena vizinhança da origem,  $\Phi^{-1}$  está se comportando como um microscópio: pequenos círculos  $x^2 + y^2 = \varepsilon^2$ , de perímetro  $2\pi\varepsilon$ , são mapeados por  $\Phi^{-1}$  em dois grandes círculos  $\{\pm\varepsilon\} \times \mathbb{R}/2\pi\mathbb{Z}$ , de perímetro  $2\pi$ .

Como um primeiro exemplo ingênuo, considere uma reta  $D$  passando pela origem. A imagem inversa  $\Phi^{-1}(D)$  consiste de duas “linhas”  $\theta = \alpha$  e  $\theta = \alpha + \pi$  mais o círculo  $\{0\} \times \mathbb{R}/2\pi\mathbb{Z}$ . Portanto, se duas linhas distintas  $D_1$  and  $D_2$  se intersectam na origem, suas imagens inversas tornam-se de algum modo disjuntas. A operação  $\Phi^{-1}$  remove o ponto de interseção. O de “alguma forma” é devido ao fato de que  $\Phi^{-1}(D)$  contém  $\Phi^{-1}(0,0)$  de modo que as imagens inversas de duas linhas que se intersectam não podem ser disjuntas.

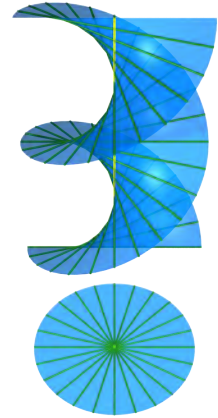
Um procedimento melhor é o seguinte: Dado um subconjunto  $X$  do plano, denota-se por  $\widehat{\Phi^{-1}}(X)$  o fecho em  $\mathbb{R} \times \mathbb{R}/2\pi\mathbb{Z}$  de  $\Phi^{-1}(X \setminus \{0,0\})$ . Com essa definição  $\widehat{\Phi^{-1}}(D_1)$  e  $\widehat{\Phi^{-1}}(D_2)$  são de fato disjuntos. Dizemos que  $\widehat{\Phi^{-1}}$  é a *transformação estrita*.

A fim de visualizar  $\Phi$ , podemos considerar a superfície  $S$  mergulhada em  $\mathbb{R}^2 \times \mathbb{R}/2\pi\mathbb{Z}$ , definida por  $x \sin \theta = y \cos \theta$ . Isso é análogo a uma escada espiral dupla. A imagem na margem representa uma escada simples em  $\mathbb{R}^2 \times [0, 2\pi[$ . Note que  $S$  é uma superfície *suave*. Nossa aplicação  $\Phi$  corresponde à projeção sobre o plano horizontal  $\mathbb{R}^2$ , e  $\Phi^{-1}(0,0)$  é a vertical  $\mathbb{R}/2\pi\mathbb{Z} \times \{(0,0)\}$ .

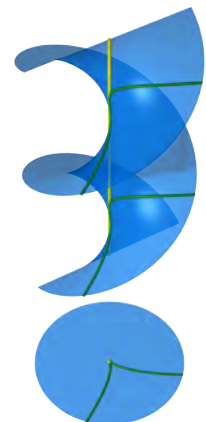
Como um segundo exemplo simples, considere a curva plana  $x^3 = y^2$ , tendo um ponto singular cuspidal na origem. Sua transformação estrita tem a equação  $\rho = \sin^2 \theta / \cos^3 \theta$  (com duas componentes, como deve ser) e não é mais singular, sendo agora suave e tangente ao círculo  $\mathbb{R}/2\pi\mathbb{Z} \times \{(0,0)\}$ .

A ideia geral é que a transformada estrita de uma curva seja “menos singular” do que a curva original na origem. Repetindo a operação, um certo número de vezes, esperamos transformá-la em uma curva suave.

Porém, antes de continuar, temos de corrigir o problema de que pré-imagens por  $\Phi$  contêm dois pontos. Iterando o processo  $n$  vezes, obtemos  $2^n$  pontos e isso é muito difícil de manusear.



Chapa 27 de [Instruction en la science de perspective](#), H. Hondius (1625).



### A faixa de Möbius

Note que a involução enviando  $(\rho, \theta)$  para  $(-\rho, \theta + \pi)$  não tem pontos fixos. Uma maneira fácil de se livrar das pré-imagens duplas de  $\Phi$  é identificar os pontos  $(\rho, \theta)$  e  $(-\rho, \theta + \pi)$  em  $\mathbb{R} \times \mathbb{R}/2\pi\mathbb{Z}$ . O quociente é uma superfície suave. A involução inverte orientação desde que o seu determinante Jacobiano seja  $-1$ . Portanto, segue que a superfície quociente não é orientável, constituindo-se na famosa *faixa de Möbius*<sup>70</sup>.

Poderíamos expressar o mesmo da seguinte maneira: O conjunto das retas que passam pela origem é um círculo, que pode ser parametrizado

– por sua *inclinação*  $t$  a qual é um elemento do plano projetivo real  $P^1_{\mathbb{R}} \simeq \mathbb{R} \cup \{\infty\}$ .

– ou por seu ângulo  $\theta$ , agora *definido módulo*  $\pi$ .

Considere o espaço  $\mathcal{M}$ , consistindo dos pares  $(p, D)$  sendo  $p$  um ponto do plano e  $D$  uma reta passando através da origem e de  $p$ . Pode-se ver isso como

$$\mathcal{M} = \{((x, y), t) \in \mathbb{R}^2 \times (\mathbb{R} \cup \{\infty\}) \mid y = tx\}$$

ou

$$\mathcal{M} = \{((x, y), \theta) \in \mathbb{R}^2 \times \mathbb{R}/\pi\mathbb{Z} \mid x \sin \theta = y \cos \theta\}.$$

A primeira apresentação tem a vantagem de ter uma equação muito simples e a desvantagem de não mostrar imediatamente que  $\mathcal{M}$  é uma superfície suave na vizinhança de  $t = \infty$ . No entanto, em um segundo pensamento, substituindo  $t$  por  $t' = 1/t$ , a equação torna-se  $x = t'y$  e a desvantagem desaparece. A segunda apresentação mostra que  $\mathcal{M}$  é de fato o quociente de  $\mathbb{R} \times \mathbb{R}/2\pi\mathbb{Z}$  pela involução mencionada acima.

Note que  $x = y = 0$  define um círculo  $E$  mergulhado em  $\mathcal{M}$ : chamado de o *divisor excepcional*.

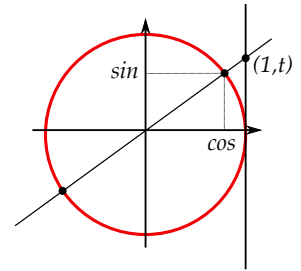
Agora consideremos a aplicação

$$\Psi : ((x, y), t) \in \mathcal{M} \mapsto (x, y) \in \mathbb{R}^2.$$

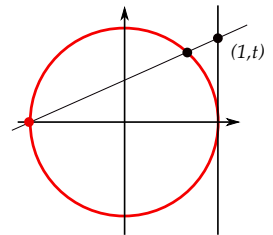
Ela tem exatamente as propriedades que queríamos:

<sup>70</sup> P. Popescu-Pampu. *La bande que tout le monde connaît. Images des Mathématiques*, 2010.

Em 1895, nos seus artigos *Analysis Situs*, Henri Poincaré não fez menção sobre a faixa de Möbius, mas da “surface unilatère que tout le monde connaît” (a superfície unilateral que todo mundo conhece).



O círculo de ângulos módulo  $\pi$  é uma linha projetiva real.



O círculo de ângulos módulo  $2\pi$  também é uma linha projetiva real.

1.  $\Psi$  restrita ao complemento do divisor excepcional é um difeomorfismo sobrejetivo no complemento da origem no plano.
2.  $\Psi$  “colapsa” o divisor excepcional na origem.

Por essa razão, dizemos que  $\mathcal{M}$  foi obtida a partir do plano por um *blow-up da origem*. Reciprocamente, usualmente, diz-se que  $\Psi$  é uma aplicação *blow-down*.

Como queremos trabalhar localmente, é frequentemente útil restringir a uma superfície compacta com borda.

$$\overline{\mathcal{M}} = \{(x, y, \theta) \mid x^2 + y^2 \leq 1\} \subset \mathcal{M}.$$

Claramente  $\overline{\mathcal{M}}$  é obtida de  $[-1, +1] \times [0, \pi]$  após colar  $(t, 0)$  e  $(-t, \pi)$ .

Temos a imagem familiar da faixa de Möbius: um retângulo onde dois lados opostos são identificados após uma torção.

Deve ficar claro que a borda de  $\overline{\mathcal{M}}$  é conexa, visto que é levada homeomorficamente da borda de um disco pela aplicação  $\Psi$ .

Deve ficar igualmente claro que o complemento do divisor excepcional em  $\overline{\mathcal{M}}$  é conexo. Essa é uma maneira complicada de dizer que se cortarmos a faixa ao longo do “círculo do meio”, a espinha dorsal, obtém-se um anel *bona fide*. Na verdade, o complemento de  $E$  é homeomorfo a um disco perfurado.

Finalmente, a imagem inversa de um círculo, digamos

$$x^2 + y^2 = 1/2$$

é um círculo mergulhado em  $\overline{\mathcal{M}}$  que o desconecta em duas partes. O primeiro é um anel não torcido, mapeado para  $x^2 + y^2 \geq 1/2$  por  $\Psi$ , e o segundo é uma faixa de Möbius menor, mapeada para  $x^2 + y^2 \leq 1/2$  por  $\Psi$ .

### Algumas imagens

A faixa de Möbius é, sem dúvida, um dos poucos objetos matemáticos que são famosos fora do mundo matemático. Em ficção científica, em arte, filosofia etc.

A terminologia vem da geometria algébrica e é confusa, pois o divisor excepcional é um círculo mergulhado em  $\mathcal{M}$  que *não* divide a superfície em duas componentes, ao contrário da espinha dorsal de um anel.

A terminologia *excepcional* também tem procedência da geometria algébrica, mas é bem mais difícil de explicar. Duas curvas *complexas* e suaves numa superfície *complexa* têm um número positivo de interseção. Um “blow-up” de um ponto em uma superfície *complexa* produz um divisor, que agora é uma reta projetiva *complexa* que não pode ser deformada holomorficamente. Para calcular o número de autointerseções, temos que usar deformações não holomorfas. Acontece que sua autointerseção é  $-1$ , considerada surpreendente pelos geométricos-algébricos do passado.



O site [Impact Earth](#) permite que você exploda nosso planeta em qualquer ponto.

†

Apenas por diversão, deixe-me citar algumas frases do famoso psicanalista Jacques Lacan <sup>71</sup> no seu seminário “l'Étourdit”, em 1972:

Le non-enseignable, je l'ai fait mathème de l'assurer de la fixation de l'opinion vraie, fixation écrite avec un x mais non sans ressource d'équivoque. Ainsi un objet aussi facile à fabriquer que la bande de Möbius en tant qu'elle s'imagine, met à portée de toutes mains ce qui est unimaginable dès que son dire à s'oublier, fait le dit s'endurer. D'où a procédé ma fixation de ce point doxa que je n'ai pas dit, je ne le sais pas et ne peux donc - pas plus que FREUD - en rendre compte de ce que j'enseigne, sinon à suivre ses effets dans le discours analytique, effet de sa mathématisation qui ne vient pas d'une machine, mais qui s'avère tenir du machin une fois qu'il l'a produite.

Para grandes (e sérios) comentários sobre a faixa de Möbius, convido o leitor a *olhar para* o livro de J. Scott Carter <sup>72</sup>.

A faixa tem o nome de Möbius, que a publicou em 1865, mas — como de costume — ele não foi o primeiro. Listing tinha descrito o mesmo objeto em 1862.

Poder-se-ia facilmente produzir um livro cheio de faixas de Möbius, de diferentes formas, cores etc. Deixe-me apresentar aqui apenas uma pequena amostra:

Faça um nó simples com uma faixa de papel e aperte-o. Você obtém algo como na figura abaixo. Quando você fecha seu pentágono regular, você produz uma faixa de Möbius.

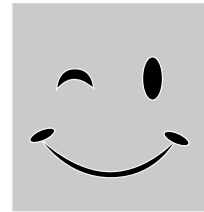


O capítulo 4 do belo *Topological picturebook* <sup>73</sup> é dedicado ao impossível *tribar*.

Considere um disco ou o interior de uma elipse no plano. Seu exterior tem a topologia de um anel. Agora pense nessa elipse no

<sup>71</sup> J. Lacan. *L'Étourdit*. Seuil, 1973.

Não sou capaz de traduzir para o Português (ou mesmo para o francês compreensível).



“olhe para” é mais apropriado que “leia” nesse caso.

<sup>72</sup> J. S. Carter. *How surfaces intersect in space: an introduction to topology*. K & E series on knots and everything. World Scientific, 2nd ed edition, 1993.

<sup>73</sup> G. K. Francis. *A topological picturebook*. Springer-Verlag, 2006.



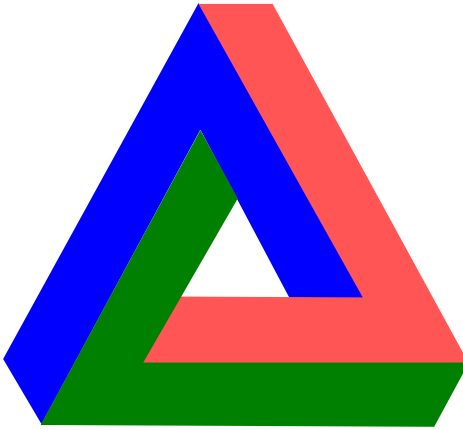
Quadrinhos de Möbius, por Étienne [Lecroart](#). Você pode ler a história e voltar ao início de cabeça para baixo e a história começa de novo! †



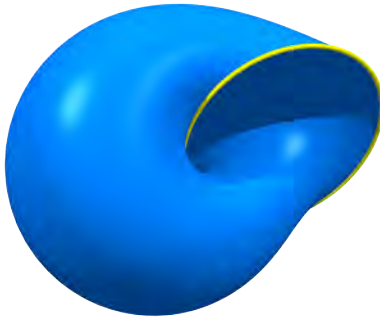
Esta é uma homenagem de J. Leys a J. S. Bach [Bach crab cannon](#) em uma faixa de Möbius (1747), por Jos Leys. †



Isso pode parecer enganoso. É possível realizá-lo de tal forma que ele é feito de três *trapézios planares* como parece na imagem? Se as três peças forem torcidas, esse objeto descreve uma faixa de Möbius no espaço. Sua borda é um círculo, como deve ser, mas ele é um nó não trivial no espaço: é um nó de trevo. Isso é diferente do quadro usual em que a borda é um nó trivial.



Um *objeto impossível*. †



Partindo de uma faixa de Möbius usual, e Deformando-se até que o limite de torne-se um círculo redondo, obtemos o *Caracol de Möbius*. †



Você já notou que o símbolo internacional de reciclagem é uma faixa de Möbius? †

*plano projetivo real* onde se tem que adicionar a linha no infinito, que é um círculo: um ponto para cada direção. Cada linha no plano projetivo intersecta o infinito em exatamente um ponto. Mostre que isso implica que o complemento de um disco no plano projetivo é uma faixa de Möbius.



Cada dobradiça tem quatro braços. Portanto, essa é a união de muitos segmentos. Verifique se isso representa duas faixas de Möbius que se cruzam ao longo da sua espinha dorsal comum, como na margem.

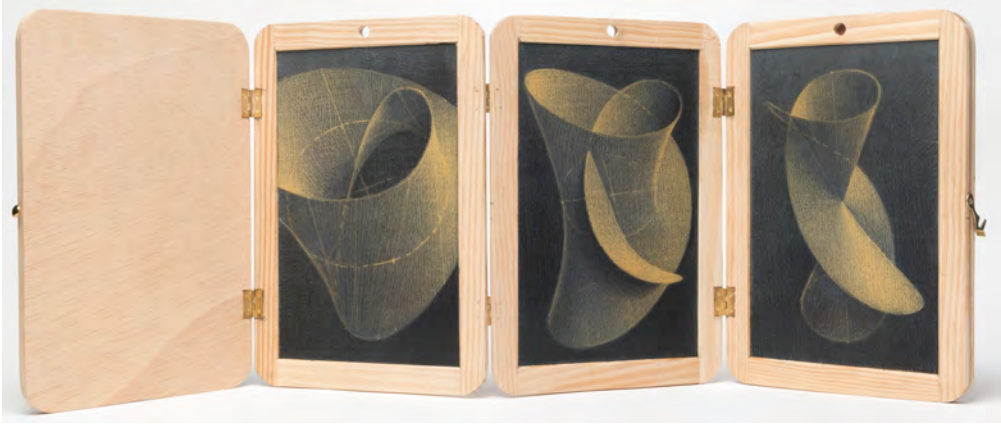
Os seguintes desenhos mostram o colapso do divisor excepcional numa faixa de Möbius.

A imagem final é um cone sobre algum laço imerso na esfera com dois pontos duplos. Um cone sobre um círculo é um disco, como esperamos de um mapa de blow-down. Cortando o cone com um plano, encontra-se um “folio de Descartes”, e a equação do cone poderia ser  $x^3 + y^3 - 3xyz = 0$ , como nos seguintes modelos de fios.

Não se deve confundir o plano projetivo real, obtido do plano  $\mathbb{R}^2$  pela adição de uma reta real projetiva (i.e. um círculo) no infinito, e a reta projetiva complexa, obtida de  $\mathbb{C}$  por adição de um ponto no infinito.



Uma escultura cinética por Anthony Howe.



Reducing the Möbius,  
tríptico em lápis de cor em  
ardósia, madeira e metal, de  
Sylvie Pic. †



Um cone de terceira ordem.  
†



### Testando nosso microscópio

Vamos testar a eficiência do nosso microscópio  $\Psi^{-1}$ :

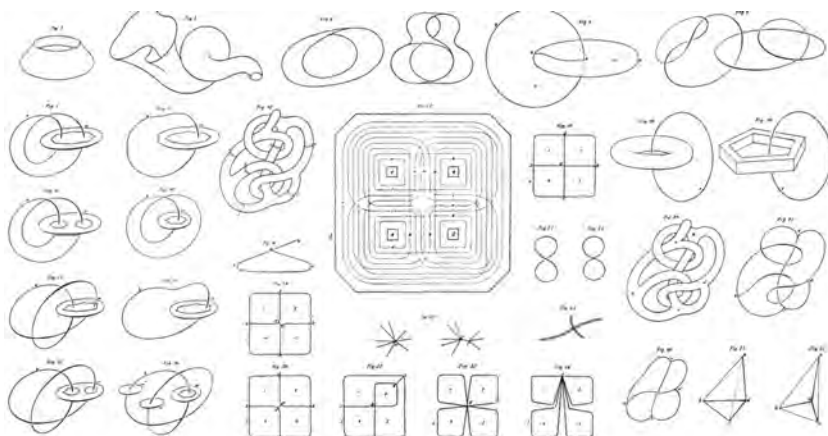
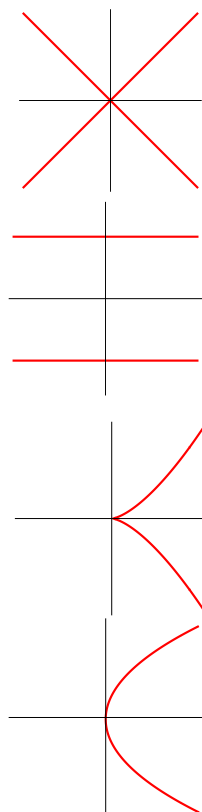
Se  $X \subset \mathbb{R}^2$  é um conjunto qualquer, ele é *estritamente transformado* no fecho de  $\Psi^{-1}(X \setminus \{(0,0)\})$  em  $\mathcal{M}$ .

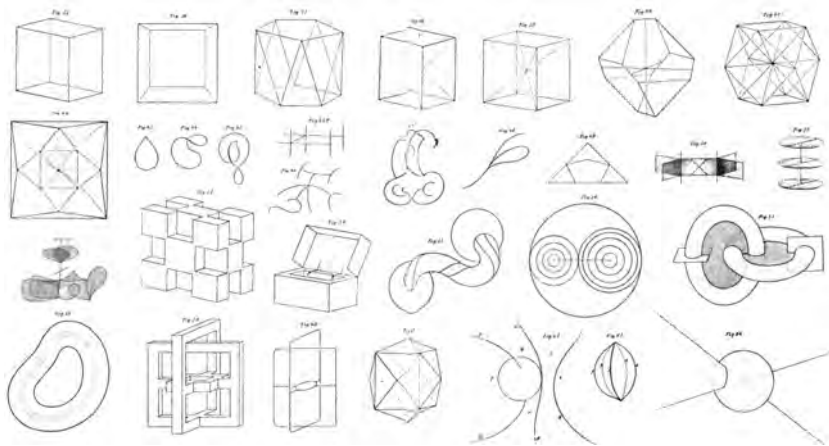
Vamos tentar primeiro com duas retas  $x^2 - y^2 = 0$ . Portanto, definimos  $y/x = t$ , que juntamente com  $y = \pm x$  dá  $tx = \pm x$ . Como calculamos a transformada estrita, trabalhamos fora da origem e obtemos  $t = \pm 1$ . Note que  $t = \infty$  não está na transformada estrita, como se vê facilmente na coordenada  $t' = 1/t$ . Agora o fecho de  $t = \pm 1$  em  $\mathcal{M}$  consiste de duas curvas disjuntas. Portanto, a transformada estrita de duas curvas suaves que se cruzam transversalmente na origem produz *curvas suaves disjuntas*.

O que acontece com um ponto cuspidal  $y^2 - x^3 = 0$ ? Esse dá  $t^2x^2 - x^3$  e nós podemos simplificar por  $x^2$  e obter  $x = t^2$ . Portanto, nas coordenadas  $(x, t)$  de  $\mathcal{M}$ , a transformação estrita é uma parábola suave, tangente ao divisor excepcional ( $x = 0$ ).

Agora vamos considerar  $y^2 - x^5 = 0$ . Portanto, a transformada estrita é  $t^2 = x^3$ , um ponto cuspidal.

Percebemos que um único blow-up será insuficiente e precisaremos fazer um blow-up novamente. Da mesma forma que o algoritmo de Newton: nem sempre terminaremos no primeiro passo.





Figuras provenientes do livro de Listing "Der Census räumlicher Complexe, der Verallgemeinerung des Euler'schen Satzes von den Polyedern" onde apareceu a faixa de Möbius pela primeira vez, em 1862.



Max Bill,  
"Unidade tripartida", 1948–  
49, escultura, MAC/USP,  
São Paulo, Brasil. O "tripar-  
tido" se relaciona ao fato de  
que esta é a soma conexa de  
três planos projetivos menos  
um disco? †

# Colares de Möbius

## Blowing up várias vezes

NÓS VIMOS COMO FAZER UM BLOW-UP DE UM PONTO DO PLANO. A construção pode ser generalizada: Dado um ponto  $p$  em uma superfície suave  $S$ , um blow-up  $S$  em  $p$  produz uma outra superfície suave  $S_p$  e uma aplicação blowing down  $\Psi_p : S_p \rightarrow S$  como antes. A imagem inversa de  $p$  é o divisor excepcional  $E_p$ : seus elementos são as linhas tangentes em  $p$  e é identificado com a linha projetiva  $P^1(T_p(S))$ , construída a partir do plano tangente  $T_p(S)$  de  $S$  em  $p$ . Fora do divisor excepcional,  $\Psi_p$  é um difeomorfismo em  $S \setminus \{p\}$ .

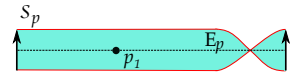
Vamos iterar o processo. Escolha um ponto  $p_1$  no divisor excepcional  $E_p = \Psi_p^{-1}(p)$  e blow-up  $S_p$  em  $p_1$ . O resultado é uma superfície suave  $S_{p,p_1}$  e uma aplicação blowing down  $\Psi_{p_1}$  de  $S_{p,p_1}$  para  $S_p$  com um divisor excepcional  $E_{p_1} \subset S_{p,p_1}$ . A imagem inversa  $(\Psi_p \circ \Psi_{p_1})^{-1}(p)$  consiste na união de  $E_{p_1}$  e da transformação estrita de  $E_p$  em  $\Psi_{p_1}$ . Essa união é chamada de divisor excepcional da composição  $\Psi_p \circ \Psi_{p_1} : S_{p,p_1} \rightarrow S$ . Fora deste divisor, a aplicação  $\Psi_p \circ \Psi_{p_1}$  é um difeomorfismo em  $S \setminus \{p\}$ .

Escolha um ponto  $p_2$  em  $(\Psi_p \circ \Psi_{p_1})^{-1}(p)$  e continue o processo em qualquer número finito de vezes, tantas vezes quanto for necessário. O resultado final é:

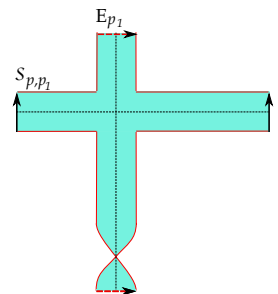
- uma superfície suave  $\bar{S}$ ,
- uma aplicação suave  $\bar{\Psi} : \bar{S} \rightarrow S$ , que envia difeomorficamente  $\bar{\Psi}^{-1}(S \setminus \{p\})$  para  $S \setminus \{p\}$ .



A fim de fazer um blow-up em um ponto de uma superfície, exclua um disco em torno desse ponto e faça um blow-up desse disco, colando a borda da faixa de Möbius à borda do complemento do disco.



Nessas figuras, deve-se colar as flechas correspondentes. Esta é uma faixa de Möbius.



Blowing-up duplo.

Observe, em particular, que a borda de  $\bar{S}$  é conexa. A imagem inversa  $\bar{\Psi}^{-1}(p)$  é o *divisor excepcional*. É uma união finita de círculos suaves mergulhados em  $\bar{S}$ . Quaisquer dois desses círculos são disjuntos ou se cruzam transversalmente em um único ponto. Três círculos diferentes não se intersectam. A imagem é uma reminiscência do logotipo dos Jogos Olímpicos. A diferença é que os anéis olímpicos são disjuntos, diferentemente de nossa situação em que alguns círculos se intersectam.

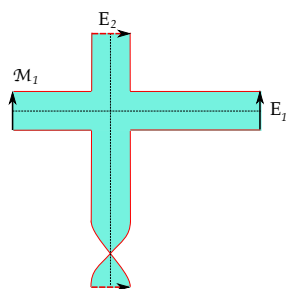


Esta composição de blowing ups é o microscópio multilentes que usaremos e que nos permitirá analisar todos os pontos singulares.

### *O microscópio*

Antes de usar nosso microscópio, vamos examiná-lo. Se começarmos com um disco  $S$ , o único blow-up  $\bar{S}$  é uma faixa de Möbius. Ilustraremos a topologia de  $\bar{S}$  no caso geral de um número finito de blowing ups.

Vamos começar com os dois blow-ups: O resultado de um blow-up de um ponto é uma faixa de Möbius. Considere agora uma faixa de Möbius  $\mathcal{M}$  contendo o divisor excepcional  $E$  como sua espinha dorsal. Como antes, escolha um ponto  $p_1$  em  $E$  e blow-up  $\mathcal{M}$  em  $p_1$ . O resultado será uma superfície  $\mathcal{M}_1$  contendo dois círculos  $E_1$  e  $E_2$  intersectando em um ponto. Aqui  $E_1$  é a transformação estrita de  $E$  e  $E_2$  é o divisor excepcional do



segundo blow-up.

Seja  $\gamma$  um laço em alguma superfície  $S$ . Comece com alguma orientação do espaço tangente de  $S$  em  $\gamma(0)$  e siga-o ao longo de  $\gamma$ . Quando o laço retorna à sua origem, a orientação é a original ou foi invertida. Por conseguinte, direi que  $\gamma$  é *orientante* ou *desorientante*. Formalmente, isso define um homomorfismo do grupo fundamental de  $S$  (ou sua primeira homologia) para  $\mathbb{Z}/2\mathbb{Z}$ .

Dê uma olhada, leitor, no logotipo dos Jogos Olímpicos do Rio de Janeiro e nos laços que são orientantes/desorientantes.

Voltando ao nosso blow-up da faixa de Möbius  $\mathcal{M}_1$ , veremos que  $E_1$  é orientante e  $E_2$  desorientante.

O fato de que  $E_2$  é desorientante deve ser claro. Quando fizemos um blow-up  $\mathcal{M}$  em  $p_1$ , introduzimos uma faixa de Möbius cuja espinha dorsal é  $E_2$ . Quanto a  $E_1$ , é a transformação estrita da espinha dorsal  $E$  de  $\mathcal{M}$ . Claramente  $E$  é desorientante em  $\mathcal{M}$ , mas isso não implica que sua transformação estrita seja desorientante também. Muito pelo contrário, como veremos. Para construir  $\mathcal{M}_1$ , cavamos um buraco no  $\mathcal{M}$  original e colamos outra faixa de Möbius em sua borda. Como estamos lidando com topologia, podemos cavar um “buraco quadrado”.

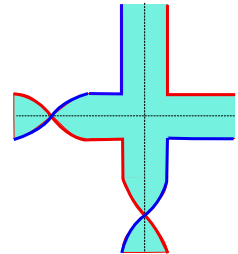
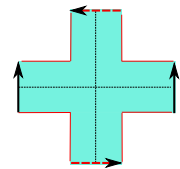
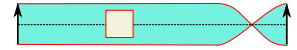
Excluindo uma pequena vizinhança anular do bordo de  $\mathcal{M}$ , podemos até imaginar que o buraco quadrado cruza  $\mathcal{M}$  “de lado a lado” (sem esquecer que uma faixa de Möbius tem apenas uma borda circular). Nesse caso, o complemento do quadrado em  $\mathcal{M}$  é outro quadrado. Portanto, a construção de  $\mathcal{M}_1$  pode ser feita de outra maneira. Comece com uma faixa de Möbius, escolha dois intervalos separados em sua borda e cole os dois lados opostos de um quadrado nesses dois intervalos. Mas ainda resta uma pergunta. A colagem dos dois lados pode ser feita de duas maneiras: com ou sem uma torção.

A construção anterior pode ser visualizada da seguinte maneira: Considere uma cruz. Cole os lados superior e inferior com uma torção para que a parte vertical da cruz se torne uma faixa de Möbius. O eixo vertical é a curva desorientante  $E_2$ .

Agora temos que colar os lados esquerdo e direito da cruz e decidir se eles devem ser torcidos ou não.

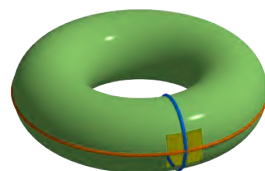
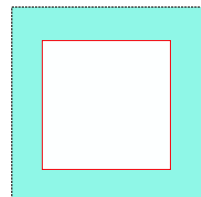
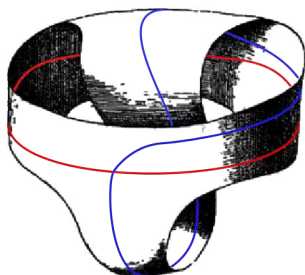


+



Vamos tentar primeiro com uma torção. O bordo da superfície resultante não é conexa: não pode ser nossa superfície  $\mathcal{M}_1$ . Portanto, os dois lados devem ser colados sem torcer: a curva  $E_1$  é de fato orientante.

Nós obtivemos uma boa imagem para  $\mathcal{M}_1$ . Um amigo recomendou que eu mostrasse a melhor imagem de  $\mathcal{M}_1$  ☺! Eu desenhei em vermelho e azul o divisor excepcional.

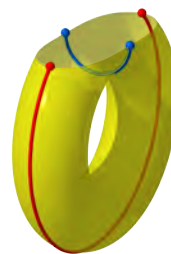


### Corações entrelaçados

Encorajo o leitor a praticar os seguintes truques topológicos:

Comece com uma cruz, cole os lados opostos sem torcer. Corte a superfície resultante ao longo da cruz central, em outras palavras, ao longo dos dois círculos. O resultado é um quadro quadrado. Você teria adivinhado? Imagine um toro no espaço tridimensional e cave um quadrado nele. Em seguida, corte e abra-o ao longo de um meridiano e um paralelo. Claramente, o que resta é um quadrado com um buraco quadrado: um quadro quadrado.

Surpreendentemente, este exemplo de uma cruz com lados opostos identificados já foi considerado por Gauss sob o nome *Doppelring*. Em seu notável artigo *Gauss als Geometer*, Stäckel <sup>74</sup> relata uma conversa entre Gauss e Möbius. Gauss observa que o *Doppelring* tem uma borda conexa. Mais interessante, ele observa que existem dois arcos *disjuntos* conectando dois pares de pontos *entrelaçados* na fronteira. Lembro que a impossibilidade de tal configuração em um disco era o ponto crucial em sua



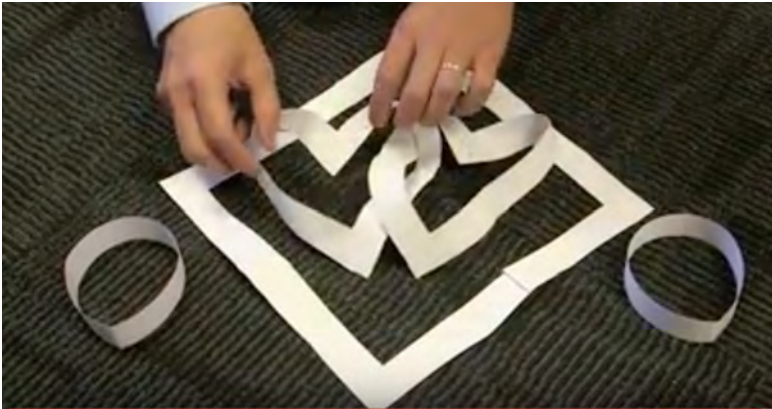
<sup>74</sup> P. Stäckel. *Literaturberichte: Materialien für eine wissenschaftliche Biographie von Gauß. Monatsh. Math. Phys.*, 32(1):A5, 1922. Gesammelt von F. Klein, M. Brendel und L. Schlesinger. Heft V: C. F. Gauss als Geometer. In Kommission bei B. G. Teubner in Leipzig, 1918.

prova do Teorema Fundamental da Álgebra.

Pegue novamente a mesma cruz e cole os lados opostos, agora com uma torção. Corte a superfície resultante ao longo dos dois círculos. O resultado é . . .

Bem, depende de como você torceu. Como uma superfície abstrata é bem definida: consiste de duas componentes conexas, cada uma homeomorfa a um anel. No entanto, o modo como é mergulhado no espaço depende da torção. Experimente!

O resultado mais impressionante ocorre quando ambos os lados opostos da cruz são torcidos, mas de uma maneira diferente: à esquerda e à direita, por assim dizer. Isso produz dois corações entrelaçados.



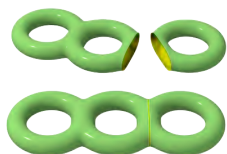
Sobre este tópico, o leitor deve olhar para Tadashi Tokieda apresentado em [Unexpected shapes](#) no Youtube (em duas partes). +

Finalmente, colando dois dos lados opostos com uma torção e não os outros dois, nós pegamos nossa faixa de Möbius blown-up, com uma borda conexa. O que acontece quando o abrimos ao longo dos dois círculos excepcionais? Isso é fácil, pois o blow-up é um homeomorfismo fora do divisor excepcional. Nós temos algo como um homeomorfo em um disco perfurado. De fato, também obtemos um quadro quadrado. Além disso, a maneira como esse quadro é mergulhado no espaço depende da torção. Pratique esses truques topomágicos!



### Blowing up mais pontos

Eu agora descrevo a situação quando mais pontos sofrem blow-ups. É fácil descrever a topologia da superfície resultante. Fazer um blow-up em um ponto equivale a cavar um buraco na superfície e a colar uma faixa de Möbius na borda. Dito de maneira diferente, a operação de fazer um blow-up é equivalente à soma conexa com um plano projetivo. Dadas duas superfícies conexas



$M_1$  e  $M_2$ , sua soma conexa  $M_1 \# M_2$  é obtida excluindo um disco de cada uma delas e colando-os ao longo das bordas recém-criadas. Como já observamos que, quando deletamos um disco em um plano projetivo, obtemos uma faixa de Möbius, o efeito topológico de um blowing up é a soma conexa com um plano projetivo. Portanto, se fizermos um blow-up de um disco  $k$  vezes em uma linha, a superfície resultante será a soma conexa de  $k$  planos projetivos menos um disco. Lembre-se de que qualquer superfície não orientável e compacta com uma borda conexa e homeomorfa a uma superfície desta forma, e também que o número  $1 - k$  é conhecido como a característica de Euler–Poincaré da superfície. Veja por exemplo os livros <sup>75</sup> e <sup>76</sup>. No entanto, essa é apenas uma descrição parcial do resultado, uma vez que ainda temos que descrever a posição e a natureza do divisor excepcional. O último não depende apenas de  $k$ , mas também das escolhas dos  $k$  pontos sucessivos que sofreram blow-up.

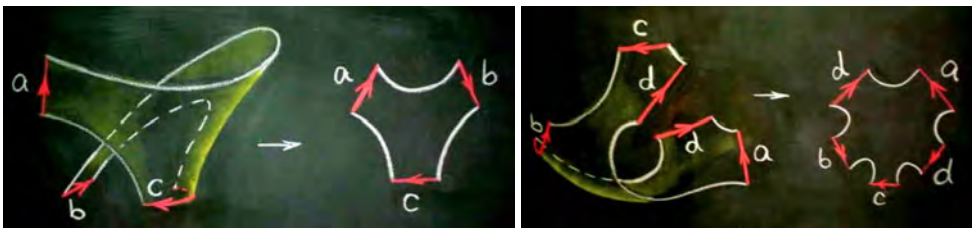
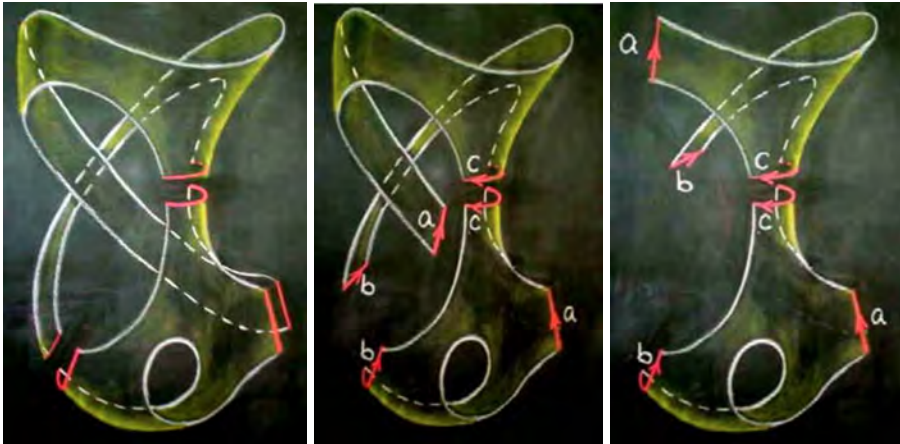
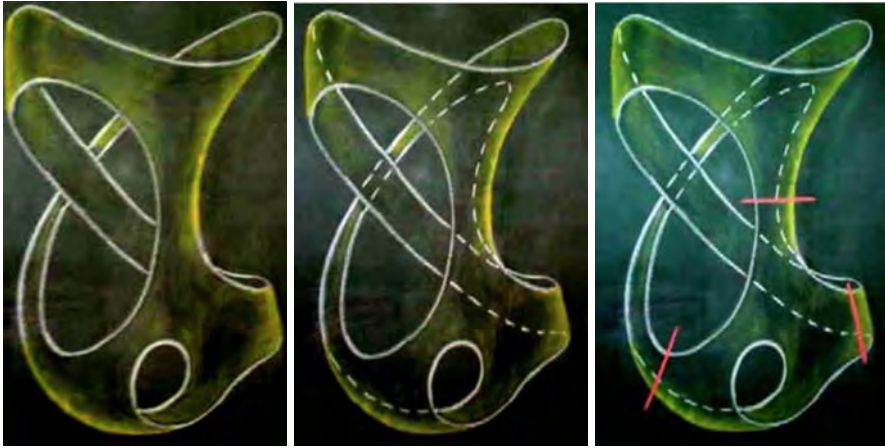
Veja a bela escultura de Max Bill ilustrando este capítulo. Um parágrafo do livro de Ton Marar <sup>77</sup> é dedicado a mostrar que essa escultura representa uma soma conexa de *três planos projetivos* (menos um disco). Isso é explicado nas imagens a seguir, extraídas deste artigo. O mesmo artigo contém outra versão da mesma superfície, inspirada no já mencionado livro de Francis (página 101).

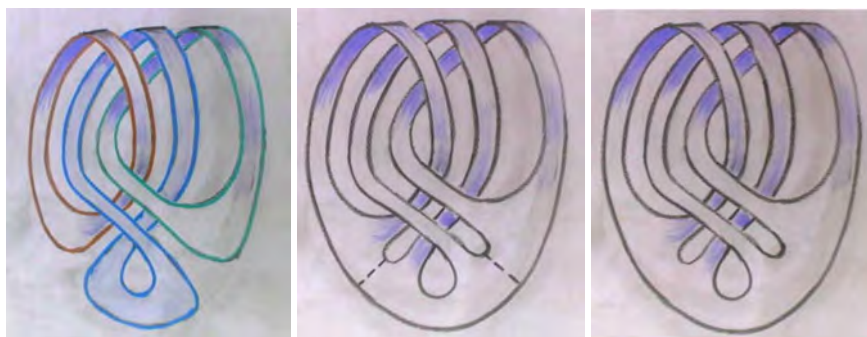
Existe uma questão de orientação sutil aqui. As duas superfícies perfuradas podem ser coladas de duas maneiras diferentes, uma vez que o círculo do bordo tem duas orientações. No entanto, superfícies orientáveis têm homeomorfismos revertendo a orientação. Verifique se isso implica em que a soma conexa é de fato bem definida entre superfícies conexas, orientáveis ou não.

<sup>75</sup> S. Barr. *Experiments in topology*. Dover Publications, Inc., Mineola, NY, 1989. Reimpressão do original de 1964.

<sup>76</sup> V. G. Boltyanskiĭ and V. A. Efremovich. *Intuitive combinatorial topology*. Universitext. Springer-Verlag, New York, 2001.

<sup>77</sup> T. Marar. [Aspectos topológicos na arte concreta](#), 2004. Notas de um minicurso da II Bienal da Sociedade Brasileira de Matemática, Salvador, UFBA.





### Colares de divisores

Ainda temos que descrever a topologia do divisor excepcional *dentro* da soma conexa dos planos projetivos.

No primeiro passo, não há surpresa: o divisor é a espinha dorsal da faixa de Möbius.

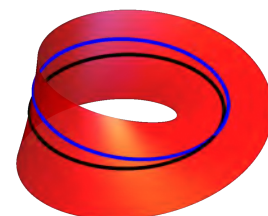
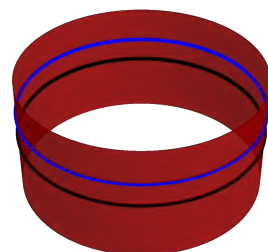
Na segunda etapa, fazemos um blow-up de um ponto da faixa de Möbius. O caso de interesse é quando fazemos um blow-up em um ponto de  $E_1$ , conforme discutido anteriormente. Os geométricos algébricos pensam na linha projetiva como uma linha ... e a desenham como uma linha, embora ela seja homeomorfa a um círculo. ... Quando chegamos ao terceiro blow-up, podemos escolher o ponto em  $E_1$ , ou em  $E_2$ , ou no cruzamento de  $E_1$  e  $E_2$ . Em todos os casos, a superfície blow-up é uma soma conexa de *três planos projetivos* (menos um disco), isto é, a superfície de Max Bill. No entanto, a localização do divisor excepcional nesta superfície não é a mesma. Como exercício, o leitor deve tentar (mentalmente) desenhar os três possíveis divisores excepcionais, diretamente na escultura.

A situação geral é agora fácil de descrever. Combinatoriamente, as várias componentes do divisor excepcional são organizadas como uma árvore.

A cada blow-up, uma nova faixa de Möbius é anexada ao colar anterior.

No entanto, isso altera a orientabilidade da faixa na qual a nova faixa está conectada.

Para perdôá-los, devo lembrar que a linha projetiva sobre os números complexos é homeomorfa a uma esfera de dimensão 2 e a um conjunto de Cantor no caso dos números  $p$ -ádicos.



Para provar isso, considere uma curva fechada  $\gamma$  em uma superfície  $S$ . Deforme levemente  $\gamma$  para algum  $\gamma'$ , transversal a  $\gamma$ , e conte o número de pontos de interseção em  $\gamma \cap \gamma'$  modulo 2. Isso é chamado de *autointerseção* de  $\gamma$ . É igual a 0 ou 1 de acordo com o fato de  $\gamma$  ser orientante ou desorientante.

Agora vamos fazer um blow-up em algum ponto  $p$  de  $\gamma \subset S$ . Escolha  $\gamma'$  que também passa por  $p$  e deixe-nos aumentar a imagem em  $p$ . As duas transformações estritas  $\bar{\gamma}$  e  $\bar{\gamma}'$  têm precisamente um ponto a menos de interseção que  $\gamma$  e  $\gamma'$ , já que as tangentes em  $p$  são diferentes. Segue-se que a autointerseção de  $\bar{\gamma}$  é igual à autointerseção de  $\gamma$  menos (ou mais, já que contamos módulo 2!) um.

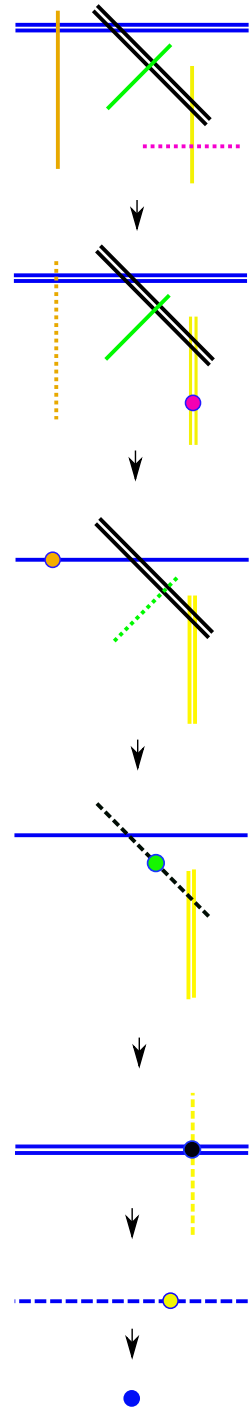
Vamos elaborar um exemplo: A imagem na margem ilustra uma sucessão de seis blow-ups. Os pontos grossos representam os centros dos blow-ups. As retas representam as retas projetivas (não esqueça que elas são na verdade círculos). As retas tracejadas representam os novos divisores que aparecem em cada etapa. Assim, as aplicações blow-down, representadas por setas para baixo, estão colapsando essas retas pontilhadas em pontos da mesma cor. Retas duplas representam as componentes orientantes. No final do processo, o divisor excepcional consiste em seis círculos.

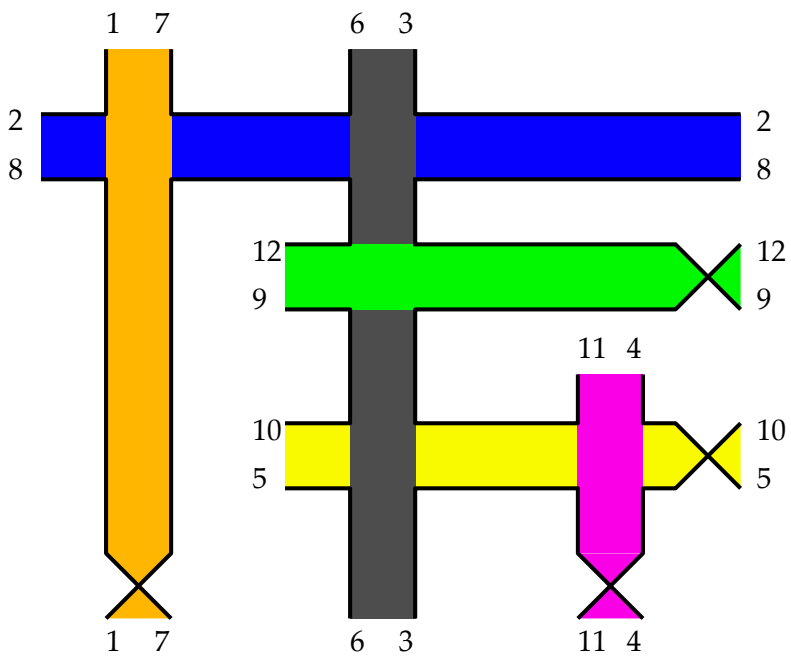
Agora podemos desenhar o colar correspondente, feito de quatro faixas de Möbius e dois anéis. Os lados opostos das seis faixas devem ser colados como sugere a figura. Você deve verificar se a borda é de fato conexa, como deveria ser. Vá ao redor da borda, seguindo os números de 1 a 12 e depois de volta para 1.

## Colagem

Aqui está outra visão do exemplo anterior:

Um topólogo diria que esta superfície é obtida por meio da *colagem* de várias faixas de Möbius e anéis. Esta operação é muito simples. Suponha que você tenha duas superfícies  $S_1, S_2$  com borda não vazia. Escolha dois mergulhos  $i_1, i_2$  do quadrado  $[-1, 1]^2$  em  $S_1$  e  $S_2$  de tal forma que as imagens  $i_1(\{\pm 1\} \times [-1, 1])$





e  $i_2([-1, 1] \times \{\pm 1\})$  encontram-se no bordo de  $S_1$  e  $S_2$  respectivamente.

Agora, para cada  $(x, y) \in [-1, 1]^2$ , identifique  $i_1(x, y)$  e  $i_2(x, y)$ . O resultado é a colagem de  $S_1$  e  $S_2$  ao longo de  $i_1, i_2$ . Esta é uma superfície com borda (e cantos que podem ser suavizados facilmente). Veja este artigo <sup>78</sup> para uma apresentação de algumas variações em torno dessa construção.

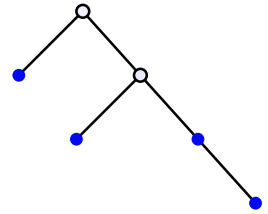
Vamos começar agora com uma árvore planar enraizada. Para cada vértice, pegue um anel ou uma faixa de Möbius. Agora junte todas essas faixas de acordo com o diagrama dado pela árvore. Cada faixa é colada para todas as faixas associadas aos seus filhos na árvore, como na figura na margem. Note que o anel e a faixa de Möbius admitem quatro homeomorfismos, permutando lados opostos do quadrado, de modo que a operação de colagem de tal faixa é bem definida. O resultado final desta colagem é uma superfície  $S$  com borda.

Cada faixa (anel ou faixa de Möbius) contém um círculo como sua espinha dorsal. A união desses círculos define um grafo  $E \subset S$  que podemos chamar de divisor, mesmo que nosso  $S$  não tenha sido necessariamente construído por uma sequência de blowing ups. Existe uma projeção  $\pi$  de  $S$  em  $E$  tal que a imagem inversa  $\pi^{-1}(x)$  consiste em um arco se  $x$  for um ponto regular de  $E$  e dois arcos intersectados de outra forma. Vamos denotar por  $S/E$  o espaço topológico obtido colapsando  $E$  para um único ponto. Se  $S$  é o resultado de uma sequência de blow-ups, sabemos que  $S/E$  é um disco fechado e a projeção de  $S$  para  $S/E$  é a aplicação blowing down.

*Exercício:* Mostre que o espaço  $S/E$  é homeomorfo a um cone cuja base é a união disjunta de  $k$  círculos, sendo  $k$  o número de componentes conexas da borda de  $S$ .

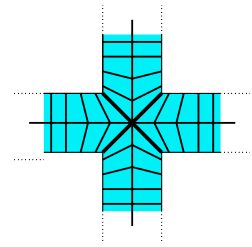
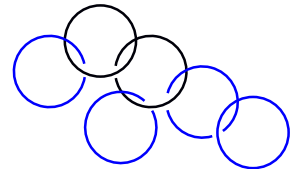
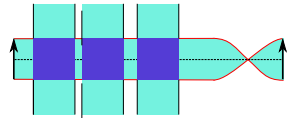
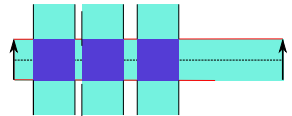
Em particular, o espaço do quociente  $S/E$  é homeomorfo a um disco se, e somente se, o bordo de  $S$  é conexo.

O exercício a seguir fornece um critério simples que nos permite verificar diretamente a partir do diagrama se a borda de  $S$  é conexa. Isso é mais fácil do que desenhar a figura em uma folha de papel e seguir cuidadosamente a borda. A solução desse exercício requer alguma compreensão da homologia das



Pontos azuis correspondem a faixas de Möbius e brancas a anéis.

<sup>78</sup> B. Ozbagci and P. Popescu-Pampu. [Generalized plumbings and Murasugi sums.](#) *Arnold Math. J.*, 2(1):69–119, 2016.



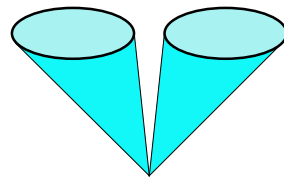
As fibras da projeção  $\pi : S \rightarrow E$ . Não confunda  $\pi$  de  $S$  no divisor  $E$  com a aplicação blowing down  $\Psi$  de  $S$  para  $S/E$ .

superfícies. Suponha que a árvore tenha  $n$  vértices. Considere a matriz simétrica  $A$  de ordem  $n$ , com os coeficientes em  $\mathbb{Z}/2\mathbb{Z}$ , definidos da seguinte maneira: Defina  $a_{ii} = 0$  se o vértice  $i$  for um anel e  $a_{ii} = 1$  se for uma faixa de Möbius. Se  $i \neq j$ , defina  $a_{ij} = 1$  se os vértices  $i, j$  forem adjacentes na árvore e 0 caso contrário.

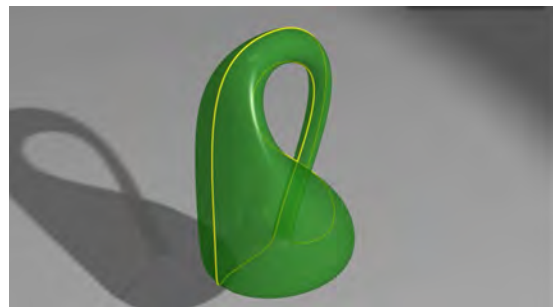
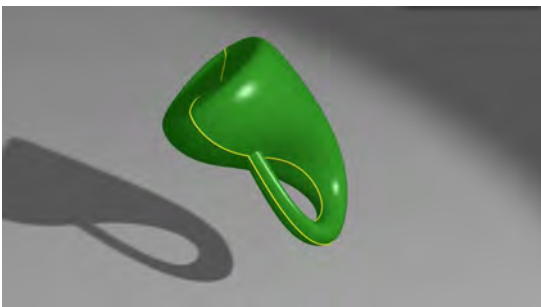
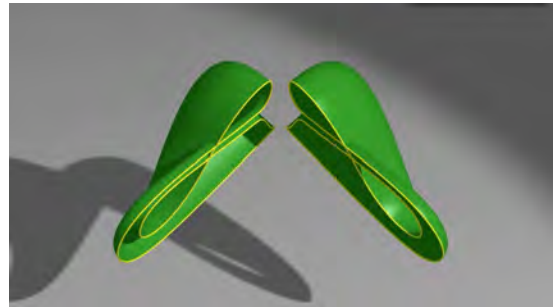
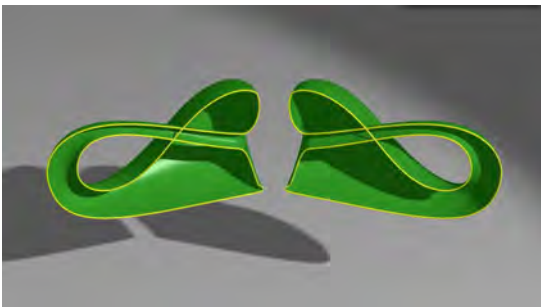
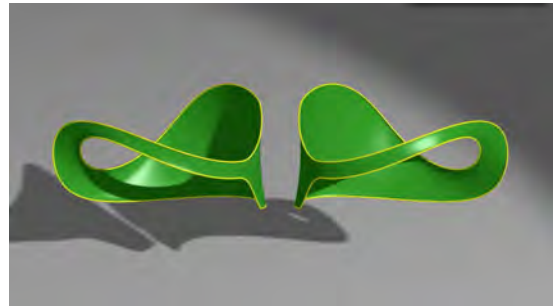
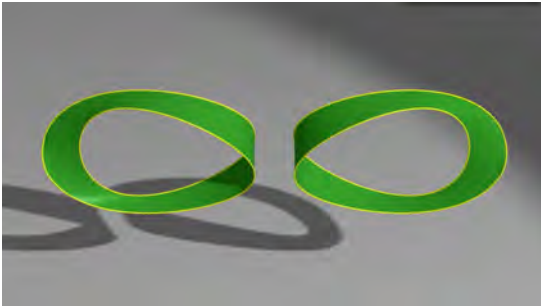
*Exercício:* Mostre que a borda de  $S$  é conexa se, e somente se, a matriz  $A$  for invertível (sobre  $\mathbb{Z}/2\mathbb{Z}$ ).

*Sugestão:* Verifique o seguinte:

- A injeção  $E \subset S$  e a projeção  $\pi : S \rightarrow E$  induz isomorfismos inversos entre  $H_1(S, \mathbb{Z}/2\mathbb{Z})$  e  $H_1(E, \mathbb{Z}/2\mathbb{Z})$ .
- Uma base de  $H_1(E, \mathbb{Z}/2\mathbb{Z})$  é dada pelas espinhas dorsais das  $n$  faixas.
- A forma de interseção simétrica em  $H_1(E, \mathbb{Z}/2\mathbb{Z})$  é dada pela matriz  $A$ .
- O núcleo da forma de interseção é a imagem de  $H_1(\partial S, \mathbb{Z}/2\mathbb{Z})$  em  $H_1(S, \mathbb{Z}/2\mathbb{Z})$ .



O cone em dois círculos.



Colando duas faixas de Möbius:  
uma garrafa de Klein.





Um microscópio de 1882 . †

# Resolução de singularidades

NÓS VAMOS UTILIZAR UM MICROSCÓPIO PARA ANALISAR A NATUREZA DAS SINGULARIDADES. Provaremos um teorema que é essencialmente devido à Max Noether <sup>79</sup>.

## Fazendo blowing up em um ramo

Considere algum ponto singular de uma curva analítica real plana definida por uma equação do tipo  $F(x, y) = 0$ .

Suponhamos que encontramos um ramo real dessa curva, ou seja, uma solução da forma

$$x = \pm t^m \quad y = \sum_{k \geq 1} a_k t^k.$$

Vamos olhar para o conjunto  $I \subset \mathbb{N} \subset \mathbb{Z}$  dos inteiros  $k$  tais que  $a_k \neq 0$ . Nós podemos sempre assumir que o máximo divisor comum dos elementos de  $I$  é igual a 1. Em outras palavras, o subgrupo de  $\mathbb{Z}$  gerado por  $I$  é  $\mathbb{Z}$ .

Seja  $\mu \geq 1$  o menor inteiro tal que  $a_\mu \neq 0$ .

Se  $\mu < m$ , a série  $y/x$  “tende ao infinito” quando  $t$  tende para 0, o que geometricamente significa que o eixo vertical  $x = 0$  é tangente ao ramo na origem.

Se  $\mu > m$ , a série  $y/x$  “tende para 0” quando  $t$  tende para 0, o que geometricamente significa que o eixo horizontal  $y = 0$  é tangente ao ramo na origem.

Se  $\mu = m$ , a reta tangente ao ramo na origem é  $y = a_m x$ .

Agora, seguindo Newton, nós olhamos para  $y$  como uma

<sup>79</sup> M. Noether. [Rationale Ausführungen der Operationen in der Theorie der algebraischen Funktionen.](#) *Math. Ann.*, 23:311–358, 1883.



[Max Noether](#) (1844–1921). †

“função” de  $x$ . Nós estamos mais interessados na curva  $F(x, y) = 0$  e, assim, podemos permutar os papéis de  $x$  e  $y$ .

Então, sempre podemos assumir que  $\mu \geq m$ .

De fato, podemos definir  $\tau$  como alguma  $\mu$ -ésima raiz de  $y = \sum_{k \geq \mu} a_k t^k$  de modo que  $\tau$  é uma série de potência em  $t$ . Nós invertemos os papéis de  $x$  e  $y$  e temos que agora  $y = \tau^\mu$  e  $x$  é uma série de potências inteiras de  $\tau$ .

Se o nosso ramo for singular, i.e. se  $m > 1$ , nós procedemos como se segue.

1. Seja  $\beta_1$  o menor inteiro em  $I$  que não é múltiplo de  $m$ .
2. Seja  $\beta_2$  o menor inteiro em  $I$  que não pertence ao grupo gerado por  $m$  e  $\beta_1$ .

Continuando nesse caminho, você obtém uma família de inteiros em  $I$  gerando  $\mathbb{Z}$ . Isso define uma sequência de inteiros  $m < \beta_1 < \beta_2 < \dots < \beta_g$ . Essa lista é a característica de Puiseux do ramo. Mais uma vez, Puiseux não é responsável por essa definição, que foi introduzida por Halphen e Smith<sup>80</sup>. Pobre Puiseux!

Nós agora vamos analisar o efeito de um blow-up sobre o nosso ramo. Perceba que na prática olhamos para as coordenadas  $(x, y_1)$  onde  $y_1 = y/x$  que é a inclinação da reta passando pela origem e pelo ponto  $(x, y)$ . Nessas coordenadas  $(x, y_1)$ , temos:

$$x = \pm t^m \quad y_1 = \sum_{k \geq \mu} a_k t^{k-m}.$$

A divisão euclideana de  $\beta_1$  por  $m$  revela que

$$\beta_1 = mq + m_1 \quad \text{com} \quad 0 < m_1 < m$$

Assim,

$$x = \pm t^m \quad y_1 = a_m + a_{2m}t^m + \dots + a_{qm}t^{(q-1)m} + \sum_{k \geq \beta_1} a_k t^{k-m}$$

Agora, se definirmos  $y_2 = y_1 - a_m$ , segue que

$$x = \pm t^m \quad y_2 = a_{2m}t^m + \dots + a_{qm}t^{(q-1)m} + \sum_{k \geq \beta_1} a_k t^{k-m}$$



Esse procedimento pode parecer complicado. É muito semelhante ao algoritmo de Euclides... Dados dois inteiros positivos  $0 < a \leq b$ , subtraindo  $a$  de  $b$  obtemos  $a, b - a$ . Se  $0 < a \leq b - a$ , continuamos o processo e obtemos  $a, b - 2a$ . Nós continuamos enquanto o primeiro inteiro for menor que o segundo. Isso nada mais é do que a divisão euclideana de  $b$  por  $a$ . Então, nós permutamos os dois inteiros e continuamos o processo. O algoritmo termina após um número finito de passos, quando o segundo inteiro for igual a 0. Nesse passo final, o primeiro número é o m.d.c de  $a$  e  $b$ . Por exemplo  $(6, 9) \rightarrow (6, 3) \rightarrow (3, 6) \rightarrow (3, 3) \rightarrow (3, 0)$ . No caso da nossa situação, que é mais complicada, podemos proceder no mesmo caminho, fazendo blowing up quantas vezes forem necessárias e ainda podemos permutar os papéis de  $x$  e  $y$  e nós continuamos...

<sup>80</sup> E. R. García Barroso, P. D. González Pérez, and P. Popescu-Pampu. Variations on inversion theorems for Newton–Puiseux series. *Math. Ann.*, 368(3-4):1359–1397, 2017.

e nós podemos fazer um blow-up novamente se  $q \geq 2$ . Fazendo um blow-up  $q$  vezes, finalmente obtemos:

$$x = \pm t^m \quad y_q = \sum_{k \geq \beta_1} a_k t^{k-qm}.$$

Como  $\beta_1 - qm = m_1 < m$ , o eixo vertical é tangente a essa curva na origem. Como antes, nós permutamos papéis das duas coordenadas e, então,

$$y_q = \pm \tau^{m_1} \quad x = \sum_{k \geq 1} b_k \tau^k.$$

Em outras palavras, após um certo número de blow-ups, obtemos uma nova curva com multiplicidade  $m_1 < m$ . Continuando nesse caminho, após um número finito de passos, obtemos uma curva suave.

Provamos, portanto, o seguinte:

**Teorema.** *Seja  $C$  um ramo de alguma curva analítica  $F(x, y) = 0$  na vizinhança da origem. Então, a transformação estrita de  $C$  por uma apropriada sucessão de blow-ups é uma curva suave.*

### *Blowing up em todos os ramos*

Em uma vizinhança da origem, a curva  $F(x, y) = 0$  consiste de vários ramos. Nós já aprendemos a desingularizar cada um desses ramos, mas as muitas curvas suaves que obtemos podem estar em uma situação relativa bastante complicada. Nós ainda temos que fazer mais blow-ups para desvendar as cordas.

Usando o teorema anterior, podemos desingularizar todos os ramos, um por um. Temos

- uma aplicação blow-down  $\Psi$  de alguma superfície  $S$  para a vizinhança da origem,
- um divisor excepcional  $E \subset S$  mapeado na origem por  $\Psi$ , tais que a transformada estrita de nossa curva é a união de um certo número de curvas suaves. Cada uma delas intersecta o divisor excepcional em um único ponto.

Se todos esses pontos são distintos, nosso trabalho é finalizado: nós teremos desingularizado a nossa curva singular  $F = 0$  em uma união de curvas suaves disjuntas.

Estritamente falando, nós apenas definimos o blowing up para superfícies reais. Também poderíamos definir blowing ups semelhantes sobre os números complexos. Nesse caso, o resultado de uma sucessão de blow-ups seria uma superfície holomorfa (isto é, de dimensão complexa 2), que contém um divisor excepcional, agora uma união de linhas projetivas complexas que se cruzam transversalmente. Poderíamos até mesmo definir essas operações de blowing up no contexto de superfícies algébricas em qualquer corpo. A maior parte deste capítulo se adaptaria literalmente a esse caso geral.

Em dimensão 3 ou mais, a desingularização é muito mais sutil e os blow-ups não são suficientes. Hiro-naka provou, em 1964, que qualquer variedade algébrica sobre um corpo de característica zero pode ser desingularizada. Sempre foi dito que essa prova é um *tour de force* que é muito difícil de digerir. No entanto, em suas palestras de 2007 sobre a resolução de singularidades (Princeton University Press), J. Kollár escreve que “A persistente percepção de que a prova é muito difícil divergiu gradualmente da realidade. ... é possível provar a resolução nas duas últimas semanas de um curso de geometria algébrica inicial”. Então, por que não tentar ler Kollár?

A única tarefa que ainda temos que fazer é lidar com um certo número de *curvas suaves*, passando pelo mesmo ponto  $p$  no divisor excepcional. Observe que algumas dessas curvas podem ser tangentes ao divisor.

O ponto  $p$  pode pertencer a uma ou duas componentes de  $E$ . Adicione esta ou estas componentes à lista de nossas curvas suaves que passam por  $p$ . Escolha as coordenadas locais  $(x, y)$  na vizinhança de  $p$  tais que as equações das curvas suaves são  $y = f_i(x)$  onde os  $f_i$  são séries de potências convergentes distintas.

Então, fazemos um blow-up novamente, introduzindo uma nova linha projetiva. As transformadas estritas de todas as curvas permaneceram suaves e cruzarão a nova componente do divisor num ponto correspondente às derivadas dos  $f_i$  na origem.

Duas curvas podem ser tangentes na origem, isto é, dois  $f_i$  podem ter a mesma derivada em 0, mas podemos fazer um blow-up novamente. Nesse processo, separaremos os  $f_i$  por alguns de seus polinômios de Taylor. Eventualmente, o resultado é uma coleção de curvas suaves que são disjuntas e transversais. Podemos até assumir que a coleção final de curvas só intersectam o divisor em pontos regulares, ou seja, não na interseção de dois círculos.

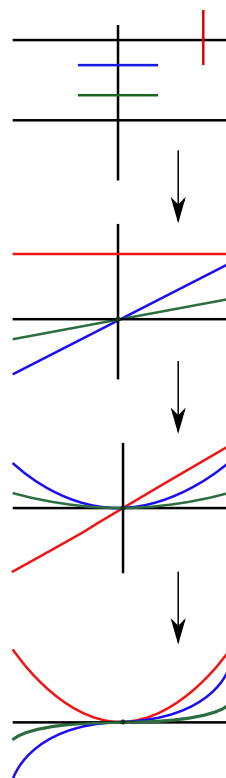
Vamos provar o teorema de Noether:

**Teorema.** *Seja  $C$  uma curva analítica na vizinhança da origem. Então, a transformada estrita de  $C$  sobre uma sucessão apropriada de blow-ups é uma união disjuntas de curvas suaves de curvas transversais ao divisor excepcional.*

### Transformações quadráticas

Max Noether estava trabalhando no contexto global de *curvas algébricas* e não de singularidades locais de curvas analíticas, como discutimos. Seu microscópio era ligeiramente diferente e é chamado de *Transformação quadrática*.

Introduzimos primeiro o grupo *Cremona* do plano projetivo  $P^2(K)$  sobre algum corpo  $K$ <sup>81</sup>. É o grupo de *automorfismos*



As curvas vermelha, azul e verde são tangentes ao divisor em preto. Três blow-ups produzindo um (novo) divisor.



Isto é um desvio.

<sup>81</sup> S. Cantat. [The Cremona group in two variables](#). In *European Congress of Mathematics*, pages 211–225. Eur. Math. Soc., Zürich, 2013.

do corpo  $K(x, y)$  de funções racionais em duas variáveis. Tal automorfismo é completamente definido por duas funções racionais  $f(x, y), g(x, y)$  que são as imagens de  $x$  e  $y$ . Duas funções  $f, g$  definem um elemento do grupo de Cremona se a transformação  $(x, y) \rightarrow (f(x, y), g(x, y))$  é um *isomorfismo biracional*.

As transformações projetivas em  $PGL(3, K)$  são isomorfismos biracionais do plano, mas o grupo de Cremona é muito grande. Um exemplo típico é a involução quadrática

$$\sigma : (x, y) \rightarrow (1/x, 1/y)$$

Que se pode ver também em coordenadas homogêneas  $[x : y : z]$

$$\sigma : [x : y : z] \rightarrow [yz : zx : xy].$$

Esta involução não é definida nos três pontos  $[1 : 0 : 0]$ ,  $[0 : 1 : 0]$  e  $[0 : 0 : 1]$ . Além disso, ela colapsa a reta que contém dois desses pontos para o terceiro ponto. Longe das três retas,  $\sigma$  é uma bijeção e até mesmo uma involução. Note também que  $[1 : 1 : 1]$  é um ponto fixo de  $\sigma$ .

Se  $A, B, C, M$  é uma base projetiva de  $P^2(K)$  (i.e. nenhum de três são colineares), existe uma transformação projetiva  $\varphi$  enviando-os para  $[1 : 0 : 0]$ ,  $[0 : 1 : 0]$ ,  $[0 : 0 : 1]$  e  $[1 : 1 : 1]$ . O conjugado  $\varphi^{-1} \circ \sigma \circ \varphi$  é a *transformação quadrática* associada ao triângulo  $A, B, C$  (e ponto fixo  $M$ ).

Max Noether usou esses mapas em vez dos blow-ups que utilizamos.

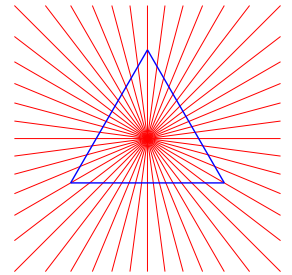
A vantagem é que toda a discussão é feita no plano projetivo sem ter que introduzir uma nova superfície. A desvantagem é que  $\sigma$  é um blow-up e um blow-down ao mesmo tempo. Ele colapsa retas e blow-up pontos de modo que ao resolver algumas singularidades, cria novas.

Noether mostrou que  $\sigma$  e  $PGL(3, K)$  geram todo o grupo Cremona.

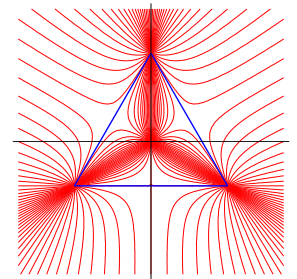
Comece com uma curva algébrica definida por alguma equação polinomial  $P(x, y) = 0$ . Escolha um ponto singular  $A$  e selecione dois pontos  $B, C$  que não estejam na curva e tal que as linhas  $AB, BC, CA$  intersectem a curva algébrica transversalmente (exceto em  $A$ ). Então, considere a imagem da curva



Luigi Cremona  
(1830–1903). +



A involução quadrática mapeia o pincel de linhas através de  $M$  para o pincel de cônicas passando por  $A, B, C, M$ .



por uma transformação quadrática associada a  $A, B, C$  (e algum ponto fixo  $M$  que não é relevante). No ponto  $A$  fizemos um blow-up. As outras interseções da curva com  $AB, BC, CA$  produzem curvas suaves que se cruzam transversalmente. Então, podemos fazer um blow-up quantas vezes quisermos, ao custo de introduzir vários pontos onde as curvas suaves se cruzam transversalmente.

Essa é a maneira como Noether expressou seu teorema:

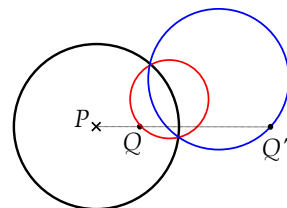
**Teorema.** *Qualquer curva algébrica pode ser transformada sob um automorfismo de Cremona adequado em outra curva cujas únicas singularidades são ordinárias, isto é, consistem em alguns ramos suaves que se cruzam transversalmente.*

Há outra involução famosa no plano: a *inversão*. Há 60 anos, todos os alunos do ensino médio estavam familiarizados com ela. Os livros-textos estavam cheios de exercícios do seguinte estilo: pegue seu teorema favorito em geometria plana, transforme-o por inversão e produza um novo teorema. A definição é muito simples.

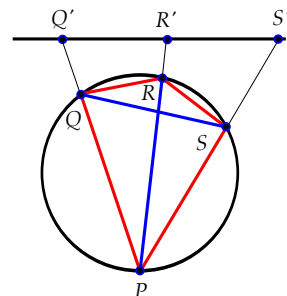
Escolha um ponto  $P$  no plano euclidiano, chamado de *polo* de inversão. Todo ponto  $Q$  é enviado por inversão num ponto  $Q'$  tal que  $P, Q, Q'$  estão numa mesma reta e tal que o produto  $\overline{PQ} \cdot \overline{PQ'} = 1$ . Essa involução não é definida para o polo  $P$ , transforma círculos não contendo  $P$  em círculos, e círculos contendo  $P$  em retas não contendo  $P$ . Se escolhermos  $P$  como sendo a origem do plano complexo, essa transformação é dada explicitamente por  $z \in \mathbb{C}^* \mapsto 1/\bar{z} \in \mathbb{C}^*$ .

Por exemplo, considere o teorema que todas as crianças francesas chamam de relação de Chasles: Por 3 pontos  $Q', R', S'$  numa reta orientada, tem-se que  $\overline{Q'R'} + \overline{R'S'} = \overline{Q'S'}$ . Aplicando a inversão a esse teorema, você pode obter o teorema de Ptolomeu: "Seja  $PQRS$  um quadrilátero convexo inscrito num círculo. Então, a soma dos produtos dos dois pares de lados opostos é igual ao produto das suas duas diagonais."

Acontece que a inversão é um caso especial de transformação quadrática. Para o primeiro vértice  $A$  do nosso triângulo, escolhemos o ponto  $[0 : 0 : 1]$  em  $P^2(\mathbb{R})$ , que é a origem do plano  $\mathbb{R}^2$ , ponto  $0 \in \mathbb{C}$ . Para o segundo e o terceiro vértices,  $B, C$ ,



$$\overline{PQ} \cdot \overline{PQ'} = 1.$$



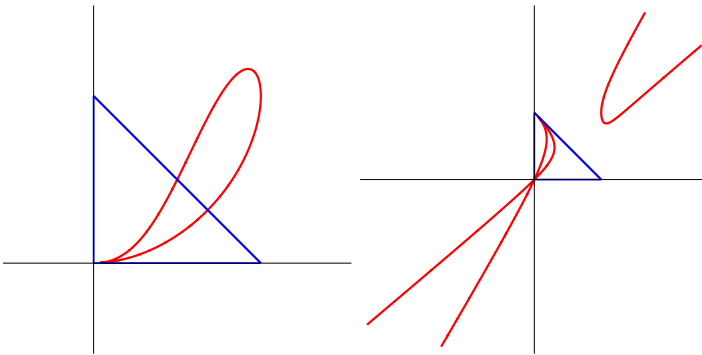
$$QR \cdot SP + PQ \cdot RS = PR \cdot QS.$$

vamos escolher os chamados *pontos cíclicos*: aqueles pontos que costumavam ser famosos entre os alunos, que são ao mesmo tempo *no infinito* e *imaginário*. Mais precisamente, eles são os pontos  $[1 : i : 0]$  e  $[1 : -i : 0]$  ( $i$  é  $\sqrt{-1}$ ). Eles são chamados *cíclicos* visto que todos os círculos no plano euclideo passam por esses pontos. Para um ponto fixo  $M$ , escolha, por exemplo, o ponto  $[1 : 0 : 1]$ , i.e., o ponto  $1 \in \mathbb{C}$ . Eu lhe encorajo, meu leitor, a mostrar que a transformação quadrática nesse caso é justamente a inversão. Isso pode ser verificado diretamente ou usando geometria projetiva clássica. Observe que a imagem de uma reta por uma transformação quadrática é uma cônica que passa pelos três vértices do triângulo. Note também que qualquer cônica que passa pelos pontos cíclicos é um círculo. Aprecie a prova!

Mais precisamente, eu deveria falar de *complexificação* de círculos, mas essa complexificação era sempre implícita no passado. O conjunto dos três pontos  $[0 : 0 : 1]$ ,  $[1 : i : 0]$  e  $[1 : -i : 0]$  está contido em  $P^2(\mathbb{C})$  e não em  $P^2(\mathbb{R})$ , mas a involução correspondente em  $P^2(\mathbb{C})$  preserva  $P^2(\mathbb{R})$  e induz a inversão no plano real euclideo, considerado como o complemento da reta no infinito em  $P^2(\mathbb{R})$ . Verifique!

### Vamos elaborar um exemplo

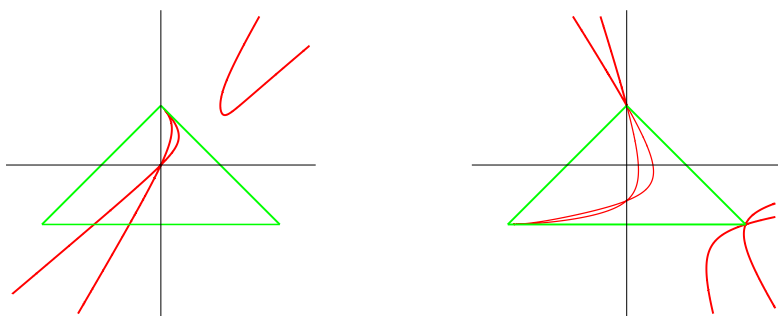
Considere a curva cuja equação é  $x^4 + x^2y^2 - 2x^2y - xy^2 + y^2 = 0$ . Essa é uma curva de Euler, que faz parte das *curvas ranfoides*, que serão discutidas mais tarde, com uma cúspide de segunda ordem. Vamos escolher um triângulo (azul) com um vértice no ponto singular e transversal à curva em qualquer outro lugar. Vamos realizar uma transformação quadrática.



O resultado é mostrado na imagem da direita acima (que eu ampliei). O ponto singular da curva ranfoide sendo localizado em um vértice, a transformação se comporta como um blow-up na vizinhança deste ponto. Nesse vértice é feito um blow-up

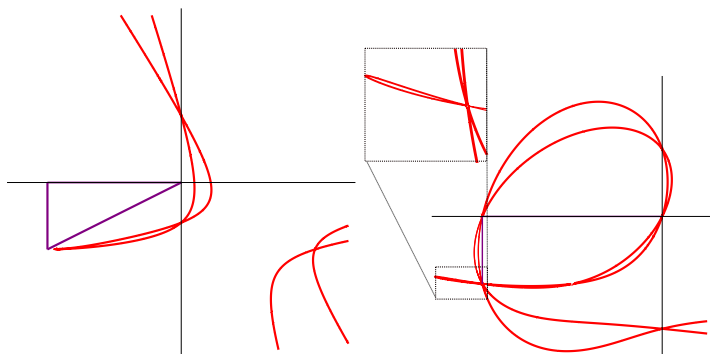


para o lado oposto do triângulo. No entanto, a singularidade é muito profunda para ser resolvida no primeiro passo. A nova curva ainda tem um ponto singular em um vértice (algum outro). Cada lado no triângulo é colapsado para o vértice oposto e isso cria um ponto duplo na origem.



Portanto, escolhamos outro triângulo (maior e verde) com um vértice no ponto singular, como mostrado na próxima imagem. Aplicamos uma vez mais a correspondente transformada quadrática. O resultado é mostrado à direita. A nova curva ainda é singular no canto inferior esquerdo, enquanto os outros vértices são pontos duplos comuns.

Nós escolhemos algum outro triângulo (roxo). Uma outra transformada quadrática leva finalmente a uma curva cujas únicas singularidades são interseções transversais de curvas suaves.



O teorema de Noether é indiscutivelmente belo, mas essas

singularidades *ordinárias* não são tão simples afinal. O exercício seguinte mostra que  $n$  curvas suaves que se cruzam transversalmente ainda contêm muita informação.

*Exercício (não tão fácil):* Suponha que você tenha um número finito  $n$  de curvas analíticas *suaves* que se cruzam transversalmente em um ponto. Mostre que para  $n = 1, 2, 3, 4$  pode-se encontrar um difeomorfismo analítico local do plano enviando-os para  $n$  retas no plano. Mostre que isso não é necessariamente verdadeiro quando  $n \geq 5$ . Você pode descrever o *espaço de moduli* de  $n$  curvas suaves transversais, ou seja, o espaço quociente sob difeomorfismos locais?

Há uma outra abordagem: Considere o espaço tangente ao espaço projetivo  $P^d(K)$  de dimensão  $d$  sobre um corpo  $K$ . Pode-se projetivizar esse espaço tangente para produzir uma variedade algébrica de dimensão  $2d - 1$ , que pode, portanto, ser mergulhada em algum espaço projetivo de maior dimensão  $P^{2(2d-1)+1}$ .

Dada uma curva algébrica  $C$  em  $P^d(K)$ , pode-se tomar o fecho (Zariski) do conjunto de suas linhas tangentes em pontos regulares. Isso produz uma outra curva algébrica  $C_1$  em algum outro espaço projetivo de dimensão  $d_1$ . Repetindo o processo, obtém-se finalmente uma curva mergulhada suave  $C_n$  em um espaço projetivo de alguma dimensão alta  $d_n$ . Agora escolha uma projeção *genérica* para uma curva em  $P^2(K)$ . Terminamos com uma curva  $\bar{C}$  que é suave com um número finito de *pontos duplos ordinários*.

Uma variedade algébrica projetiva suave de dimensão  $k$  pode ser, por definição, mergulhada em algum espaço projetivo de alguma dimensão. Pode-se então projetá-lo genericamente em um subespaço projetivo com dimensão  $2k + 1$  para produzir esse mergulho.

**Teorema.** *Qualquer curva algébrica é birracionalmente equivalente a outra curva cujas únicas singularidades são pontos duplos comuns onde dois ramos lisos se cruzam transversalmente.*

Podemos ser otimistas e esperar que qualquer curva algébrica planar seja birracionalmente equivalente a uma curva planar *suave*, mas isso está longe de ser verdade. O gênero de uma curva suave planar de grau  $d$  é  $(d - 2)/2$  de modo que se uma curva algébrica tiver um gênero que não seja um inteiro dessa forma, os pontos duplos são obrigatórios.

Poderíamos ser menos otimistas e esperar que qualquer curva algébrica possa ser transformada em alguma curva cujas singularidades sejam pontos duplos comuns usando alguma transformação de Cremona. Infelizmente, isso também não é

verdade. A equivalência birracional fornecida pelo teorema anterior pode não ser induzida por alguma transformação de Cremona <sup>82</sup> (página 42).

Para uma apresentação moderna de todos esses conceitos, eu recomendo os livros de Wall e Dolgachev <sup>83</sup> <sup>84</sup> e, para uma versão tradicional, o livro de Semple e Roth <sup>85</sup>.

<sup>83</sup> I. V. Dolgachev. *Classical algebraic geometry. A modern view*. Cambridge UP, 2012.

<sup>84</sup> C. T. C. Wall. *Singular points of plane curves*, volume 63 of *London Mathematical Society Student Texts*. Cambridge UP, 2004.

<sup>85</sup> J. G. Semple and L. Roth. *Introduction to algebraic geometry*. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1985. Reimpressão do original de 1949.



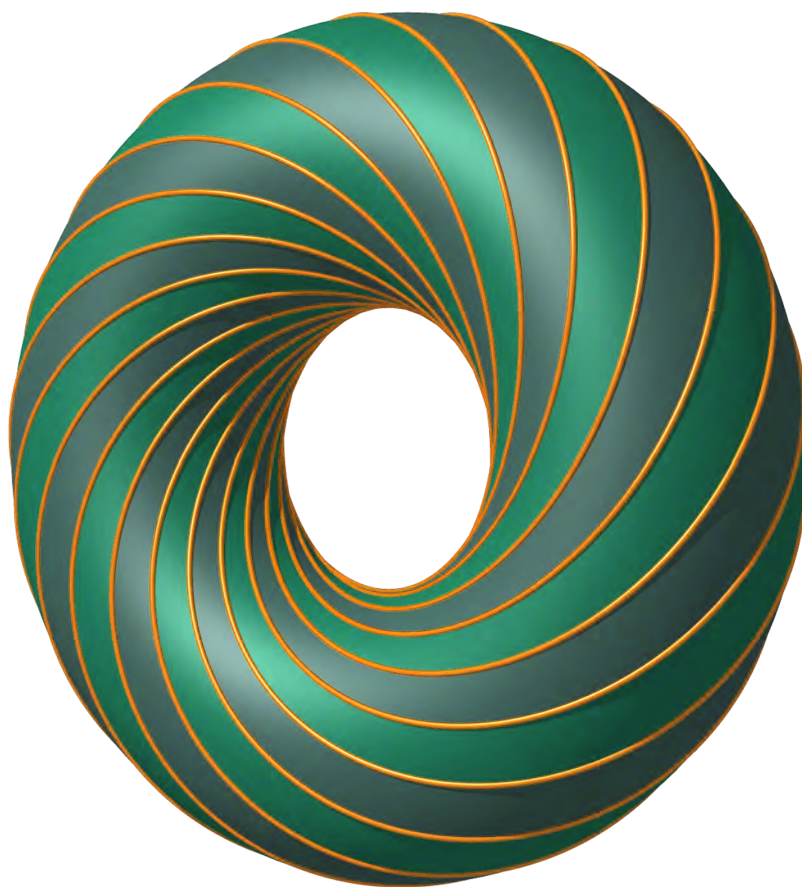


Projeções de uma curva do espaço nos três planos de coordenadas.

*Acht Draht-Modelle über die Abhängigkeit der Rückkehrelemente der Projectionen einer unebenen Curve von denen der Curve selbst (1884).*  
Geometric Models collection,  
Karazin Kharkiv University.

†





†

Um toro de Clifford preenchido com os chamados círculos de Villarceau. Cada um destes círculos é a interseção de alguma reta em  $\mathbb{C}^2$  (passando pela origem) com a esfera unitária  $S^3$  (e projetado estereograficamente no espaço tridimensional).

# A esfera de dimensão 3 e a fibração de Hopf

*Um mundo complexo?*

DEMOROU MUITO TEMPO ATÉ QUE NÚMEROS COMPLEXOS PUDESSEM SER ACEITOS PELOS MATEMÁTICOS COMO NÚMEROS GENUÍNOS.

Com o surgimento da álgebra, as raízes complexas das equações reais clamaram mais e mais insistentemente por reconhecimento.

Estas são as palavras de Coolidge em seu maravilhoso livro<sup>86</sup> descrevendo o lento surgimento da geometria complexa na matemática. Como vimos, Gauss foi um dos mais importantes pioneiros, pensando em um número complexo como um ponto no plano. Visualizar  $\mathbb{C}^2$  foi muito mais difícil, pois tem dimensão 4 em relação aos números reais e apenas visionários poderiam imaginar a quarta dimensão durante o século XIX. Muitas tentativas malsucedidas são explicadas no livro de Coolidge.

With the rise of algebra, the complex roots of real equations clamoured more and more insistently for recognition.

<sup>86</sup> J. L. Coolidge. *Geometry of the complex domain*. Clarendon Press, Oxford, 1924.



Um modelo à la Riemann de [Göttinger Sammlung mathematischer Modelle und Instrumente](#).

Mesmo Riemann, com seu conceito revolucionário agora chamado de “*superfície de Riemann*”, teve que “mergulhar” algumas superfícies no espaço de dimensão real 3, exibindo algumas

estranhas linhas onde a superfície se autointersecta, de alguma forma virtual. O mínimo que se pode dizer é que a geometria sobre os números complexos carregava algum ar de mistério.

No entanto, ficou progressivamente claro que a geometria complexa não é nada complexa, e que é de grande ajuda para a compreensão do domínio real. A seguinte citação de Paul Painlevé, em 1900<sup>87</sup>, é um bom exemplo.

Parece que, entre duas verdades do domínio real, o caminho mais fácil e mais curto passa muitas vezes pelo domínio complexo.

Atualmente a geometria complexa é melhor compreendida. A grosso modo, existem dois tipos de abordagens.

A primeira consiste em usar formalmente números complexos, como elementos de algum corpo algebricamente fechado, sem qualquer tentativa de visualizá-los. Isso tem sido muito eficiente na geometria algébrica moderna e, de fato, as propriedades algébricas de  $\mathbb{C}$  são incrivelmente poderosas. A desvantagem é que as perguntas originais, vindas de números reais, são geralmente esquecidas. A história de um famoso geômetra algébrico em uma palestra sobre variedades Abelianas (complexas) mostra bem a realidade (ou a não realidade). Depois de sua apresentação, no final de sua palestra, surgiu uma pergunta sobre as variedades abelianas *reais*. O palestrante ficou surpreso e demorou algum tempo até responder com seriedade:

*Desculpe, nunca pensei em realidade!*

A segunda abordagem consiste em *construir figuras*, projeções, seções etc. Mais importante, tenta-se desenvolver alguma intuição sobre espaços de dimensões altas, baseada em alguma *analogia*. Topólogos e geômetras modernos não têm mais medo de objetos em  $\mathbb{C}^2$  e até os consideram muito concretos. Neste capítulo, tentamos desenvolver algumas dessas intuições.

De acordo com uma estória circulando na internet, Sophus Lie teria dito:

A vida é complexa porque tem uma parte real e uma parte imaginária.

<sup>87</sup> P. Painlevé. *Ceuvres de Paul Painlevé. Tome I*. Éditions du Centre National de la Recherche Scientifique, Paris, 1973. Analyse des travaux scientifiques, pages 72–73.

*Il apparut que, entre deux vérités du domaine réel, le chemin le plus facile et le plus court passe bien souvent par le domaine complexe.*

é difícil acreditar que um matemático tão sério tenha dito tal coisa.

### *A esfera redonda de dimensão 3*

Na maioria das vezes, os geômetras desenham uma reta no quadro-negro quando eles querem significar  $P^1(\mathbb{C})$  em  $P^2(\mathbb{C})$ , mesmo sabendo que  $P^1(\mathbb{C})$  é uma esfera de dimensão 2 (Riemann) e que  $P^2(\mathbb{C})$  é uma variedade de dimensão 4 não-contrátil, que não tem muito em comum com um quadro-negro. Eles frequentemente desenham um círculo no plano quando eles querem se referir a uma esfera de dimensão 3 em  $\mathbb{C}^2$ . Eles desenham um ramo real de uma curva  $P(x, y) = 0$ , mesmo sabendo que a topologia real sobre os números complexos é muito mais rica.

Usamos essas “figuras erradas”, pois elas são muitas vezes a única aproximação possível da “realidade” no mundo complexo.

Nosso objetivo é fornecer uma descrição, tão visual quanto possível, da vizinhança de um ponto em uma curva analítica  $F(x, y) = 0$  em  $\mathbb{C}^2$ . Aqui  $x$  e  $y$  são números complexos  $x_1 + ix_2$  e  $y_1 + iy_2$  e  $x_1, x_2, y_1, y_2$  são números reais. A curva é realmente dada por duas equações

$$\Re(F(x_1 + ix_2, y_1 + iy_2)) = \Im(F(x_1 + ix_2, y_1 + iy_2)) = 0$$

em  $\mathbb{R}^4$ , de modo que, do ponto de vista dos números reais, nossa curva é uma superfície. A ideia muito natural é cruzar nossa curva/superfície com uma pequena esfera de dimensão 3, de raio  $\varepsilon$  e esperamos ver algo de dimensão 1 (sobre os reais).

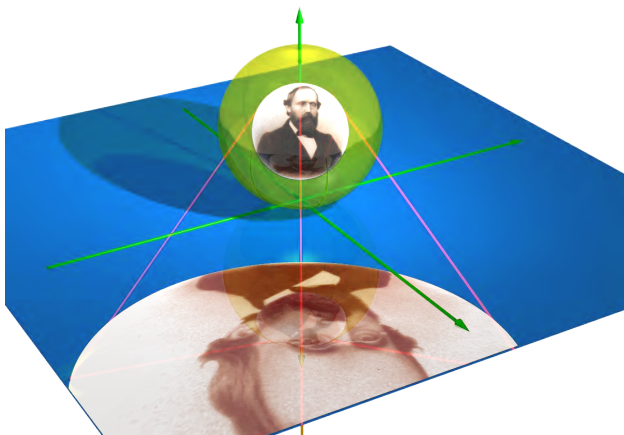
Portanto, começamos com uma descrição da esfera de dimensão 3. A interseção de nossa curva com a esfera será exibida posteriormente.

Existem várias maneiras de visualizar a esfera unitária de dimensão 3

$$\begin{aligned} \mathbb{S}^3 &= \{(x, y) \in \mathbb{C}^2 \mid |x|^2 + |y|^2 = 1\} \\ &= \{(x_1, x_2, y_1, y_2) \in \mathbb{R}^4 \mid x_1^2 + x_2^2 + y_1^2 + y_2^2 = 1\}. \end{aligned}$$

Poderíamos primeiro usar a *projeção estereográfica*. Escolha





Projetando Bernhard Riemann estereograficamente.

†

por exemplo o ponto  $N = (0, 0, 1, 0) \in \mathbb{R}^4$  como o *polo norte* de  $\mathbb{S}^3$  e projete a partir de  $N$  para o plano tangente no polo sul  $(0, 0, -1, 0) \in \mathbb{R}^4$ . O ponto  $(x_1, x_2, y_1, y_2) \in \mathbb{S}^3$  é mapeado para  $(u, v, -1, w)$  de modo que os pontos  $N$ ,  $(x_1, x_2, y_1, y_2)$  e  $(u, v, -1, w)$  estão alinhados. Na fórmula,

$$\Pi : (x_1, x_2, y_1, y_2) \in \mathbb{S}^3 \setminus \{N\} \mapsto \left( \frac{2x_1}{1-y_1}, \frac{2x_2}{1-y_1}, \frac{2y_2}{1-y_1} \right) \in \mathbb{R}^3.$$

A esfera *menos um ponto* pode, portanto, ser representada como o espaço comum de dimensão 3. Algumas simetrias são perdidas, no entanto, desde que o polo norte seja completamente arbitrário. Essa é uma boa oportunidade para recomendar o famoso livro<sup>88</sup> sobre a imaginação na geometria. Eu também recomendo o vídeo [Dimensions](#).

As seguintes propriedades da projeção estereográfica são bem conhecidas:

- a projeção é *conforme* e sua diferencial em qualquer ponto é uma semelhança.
- a imagem de um círculo na esfera de dimensão 3 é um círculo no 3-espaço (ou uma reta se o círculo original passar pelo polo norte).

<sup>88</sup> D. Hilbert and S. Cohn-Vossen. *Geometry and the imagination*. Chelsea Publishing Company, New York, N. Y., 1952. Traduzido por P. Neményi.

De acordo com alguns historiadores, estas propriedades foram estabelecidas (no caso bidimensional) por Hiparco, a quem já conhecemos neste livro.

O grupo  $SO(4)$  de rotações positivas da esfera  $S^3$  pode, portanto, ser visto como um grupo de difeomorfismos conformes  $\mathbb{R}^3 \cup \{\infty\}$ .

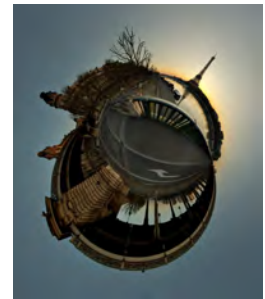
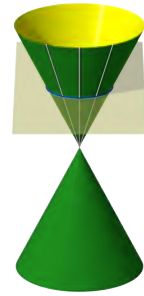
O grupo de difeomorfismos conformes da esfera de dimensão  $n$  é realmente muito maior que  $SO(n+1)$ , pois é não compacto. Por exemplo,  $\mathbb{R}^{n+2}$  pode ser equipado com a forma quadrática de assinatura  $(n+1, 1)$  dada por  $q = x_1^2 + \dots + x_{n+1}^2 - x_{n+2}^2$ , de forma que a esfera de dimensão  $n$  possa ser interpretada como a interseção do cone isotrópico  $q = 0$  com o hiperplano  $x_{n+2} = 1$ . Equivalentemente, a esfera de dimensão  $n$  pode ser pensada como o espaço das retas isotrópicas. O grupo não compacto  $SO(n+1, 1)$  induz uma ação conforme na  $n$ -esfera.

A geometria conforme das esferas é muito rica.

Deixe-me mencionar apenas duas propriedades: Qualquer difeomorfismo conforme entre dois abertos conexos em uma esfera de dimensão de pelo menos 3 é a restrição de um difeomorfismo conforme global (o teorema de Liouville). Isto está em forte contraste com o caso da dimensão 2 no qual os difeomorfismos conformes coincidem com difeomorfismos holomorfos ou anti-holomorfos, e a paisagem matemática seria muito menos bela se as aplicações holomorfas se reduzissem aos automorfismos de Möbius  $(az + b)/(cz + d)$  da esfera de Riemann.

Se o grupo conforme de uma variedade Riemanniana é não compacto, essa variedade é conforme com a esfera ou com o espaço euclidiano. Este é o teorema de Obata e Lelong-Ferrand.

Eu me abstenho de continuar nessa direção, pois nós poderíamos facilmente nos perder e nunca mais voltar do nosso passeio matemático. Eu recomendo o livro de Berger <sup>89</sup> assim como o seu vasto panorama <sup>90</sup>. Para o leitor interessado em apresentações antigas, o livro de Coolidge <sup>91</sup> é lindo.



Uma vista conforme de Paris.

†

<sup>89</sup> M. Berger. *Geometry I*. Universitext. Springer-Verlag, Berlin, 2009.

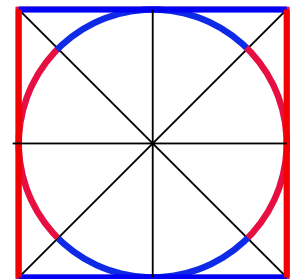
<sup>90</sup> M. Berger. *Geometry revealed. A Jacob's ladder to modern higher geometry*. Springer, Heidelberg, 2010.

<sup>91</sup> J. L. Coolidge. *A treatise on the circle and the sphere*. Clarendon Press, Oxford, 1916.

### A esfera de dimensão 3 “quadrada”

Como  $|x|^2 + |y|^2 = 1$ , a esfera de dimensão 3 pode ser dividida em duas partes  $T_1, T_2$  definidas por

$$T_1 = \{(x, y) \in S^3 \mid |x|^2 \leq 1/2\} \quad \text{e} \quad T_2 = \{(x, y) \in S^3 \mid |y|^2 \leq 1/2\}.$$



Quadrando o círculo?

A interseção de  $T_1$  e  $T_2$  é um *toro de Clifford* parametrizado por

$$(\theta, \varphi) \in (\mathbb{R}/2\pi\mathbb{Z})^2 \mapsto \left( \frac{\sqrt{2}}{2} \exp(i\theta), \frac{\sqrt{2}}{2} \exp(i\varphi) \right) \in \mathbb{S}^3.$$

Quanto à  $T_1$  e  $T_2$ , eles são *toros sólidos*, parametrizados por um produto de um disco unitário  $D^2$  em  $\mathbb{C}$  e um círculo.

$$(z, \varphi) \in D^2 \times (\mathbb{R}/2\pi\mathbb{Z}) \mapsto \left( \frac{\sqrt{2}}{2} z, \sqrt{1 - \frac{|z|^2}{2}} \exp(i\varphi) \right) \in T_1 \subset \mathbb{S}^3$$

$$(\theta, z) \in (\mathbb{R}/2\pi\mathbb{Z}) \times D^2 \mapsto \left( \sqrt{1 - \frac{|z|^2}{2}} \exp(i\theta), \frac{\sqrt{2}}{2} z \right) \in T_2 \subset \mathbb{S}^3.$$

A esfera de dimensão 3 é, portanto, a união de dois sólidos toros, colados ao longo de suas bordas. Os meridianos de  $\partial T_1$ , ou seja, os círculos que enlacam um disco em  $T_1$  são colados a paralelos de  $\partial T_2$ , que não são bordas de um disco em  $T_2$ , e reciprocamente. Nós também poderíamos usar a “bola quadrada”

$$D^2 \times D^2 = \{(x, y) \in \mathbb{C}^2 \mid |x| \leq 1 \text{ e } |y| \leq 1\}.$$

Seu bordo consiste em dois toros sólidos  $T'_1 = \{|x| \leq 1 \text{ e } |y| = 1\}$  e  $T'_2 = \{|x| = 1 \text{ e } |y| \leq 1\}$ . Usando a projeção radial, os dois toros sólidos  $T_1, T_2$  são identificados com  $T'_1, T'_2$ . Frequentemente é mais conveniente usar a esfera quadrada, já que podemos desenhar figuras no toro sólido sem ter que usar a projeção estereográfica. Esta ideia simples, mas muito útil, é devida à Kähler <sup>92</sup>.

### A esfera de dimensão 3 é muito redonda

William Thurston <sup>93</sup>, um dos mestres do aspecto visual da matemática, costumava dizer que a esfera de dimensão 3 é “mais redonda” do que as outras esferas. Ele tinha em mente o fato importante de que o grupo  $SO(n+1)$  não é um grupo simples se, e somente se,  $n = 3$  (e, claro,  $n = 0, 1$ ). Isso está relacionado ao que foi chamado na antiga literatura de *paralelismo de Clifford*.

Lembre-se de que os quatérnios são expressões formais da forma  $q = x_1 + ix_2 + jy_1 + ky_2$  em que  $x_1, x_2, y_1, y_2$  são números

<sup>92</sup> E. Kähler. über die Verzweigung einer algebraischen Funktion zweier Veränderlichen in der Umgebung einer singulären Stelle. *Math. Z.*, 30(1):188–204, 1929.

<sup>93</sup> W. P. Thurston. How to see 3-manifolds. *Classical Quantum Gravity*, 15(9):2545–2571, 1998. Topology of the Universe Conference (Cleveland, OH, 1997).



Uma captura de tela de *Knots to Narnia*. William Thurston (1946–2012) mostra que quando ele dá um nó, ele chega “em outro lugar”. Este vídeo vintage do YouTube é obrigatório. †

reais e os símbolos formais  $i, j, k$  satisfazem  $ij = -ji = k; jk = -kj = i; ki = -ik = j$  e  $i^2 = j^2 = k^2 = -1$ . Isso define uma álgebra de divisão não comutativa  $\mathbb{H}$ . O conjugado  $\bar{q}$  de  $q$  é definido como  $x_1 - ix_2 - jy_1 - ky_2$  e a norma  $N(q)$  é o produto  $q\bar{q} = x_1^2 + x_2^2 + y_1^2 + y_2^2$ . Essa norma é multiplicativa, ou seja,  $N(q_1q_2) = N(q_1)N(q_2)$  e o inverso de um quaternião diferente de zero é  $q^{-1} = \bar{q}/N(q)$ .

Segue-se que a esfera de dimensão 3 é identificada com o grupo dos quaterniões unitários  $\{q \in \mathbb{H} \mid N(q) = 1\}$ . É um dos grandes sucessos do século XX matemático provar que as únicas esferas que podem ser equipadas com a estrutura de um grupo topológico são  $S^0 \simeq \mathbb{Z}/2\mathbb{Z}$ ,  $S^1 \simeq SO(2)$  e  $S^3$ . Um bom ponto de partida para este tópico está em *Numbers* <sup>94</sup>.

Mas esta não é a única razão pela qual a esfera de dimensão 3 é redonda. Assim como qualquer grupo, ele pode ser visto como homogêneo de duas maneiras usando as translações à direita e à esquerda. Dados dois quaterniões unitários  $q_1, q_2$ , a aplicação  $q \in \mathbb{H} \mapsto q_1q_2q^{-1} \in \mathbb{H}$  é uma isometria e define um elemento de  $SO(4)$ . Acontece que esse homomorfismo de  $S^3 \times S^3$  a  $SO(4)$  é sobrejetivo e seu núcleo contém apenas  $\pm(1, 1)$ . Noutras palavras, cada rotação da esfera de dimensão 3 é a composição de uma rotação à esquerda e uma rotação à direita que comutam. Essa situação é exclusiva da dimensão 3, pois todos os outros grupos de rotação são simples (com a exceção óbvia de  $SO(2)$ ).

### A fibração de Hopf

Podemos agora iniciar nossa descrição da topologia das curvas algébricas em  $\mathbb{C}^2$  e começar com a curva mais simples possível: uma reta.

Vamos ver a interseção das retas  $x = 0$  e  $y = 0$  com a esfera unitária.

Sob a projeção estereográfica, como a reta  $x = 0$  passa pelo polo norte, sua imagem é simplesmente uma reta vertical. A outra reta  $y = 0$  é projetada em um círculo que “gira em torno da reta vertical  $x = 0$ ”.

Em nossa decomposição em dois toros sólidos,  $x = 0$  torna-se o círculo  $\{0\} \times (\mathbb{R}/2\pi\mathbb{Z})$ , que é a espinha dorsal de  $T_1$ . Recipro-

<sup>94</sup> H.-D. Ebbinghaus, H. Hermes, F. Hirzebruch, M. Koecher, K. Mainzer, J. Neukirch, A. Prestel, and R. Remmert. *Numbers*, volume 123 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1990.

Um grupo (não trivial) é *simples* se não contiver um subgrupo normal próprio. É comum dizer que um grupo de Lie é simples se todo subgrupo normal de Lie é discreto ou aberto. Isso equivale a dizer que sua álgebra de Lie é simples, ou seja, não contém um ideal próprio. O único subgrupo normal não próprio de  $SO(n)$  (para  $n \neq 1$ ) é  $\{\pm Id\}$  para  $n$  par, portanto,  $SO(n)$  não é simples como um grupo para  $n$  par, mas é simples como um grupo de Lie.



Heinz Hopf (1894–1971) não deve ser confundido com Eberhard Hopf (1902–1983) (o único matemático que se mudou dos EUA para a Alemanha em 1936?).

camente,  $y = 0$  torna-se o círculo  $(\mathbb{R}/2\pi\mathbb{Z}) \times \{0\}$ , que é a espinha dorsal de  $T_2$ . Observe que esses dois círculos estão entrelaçados.

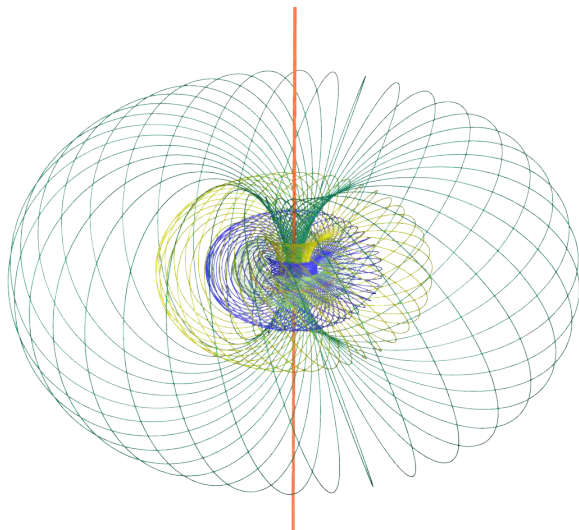
A reta  $y = x$  intersecta a esfera em um círculo que está em  $T_1$  e  $T_2$ : não é nem um meridiano, nem um paralelo, mas sua classe de homotopia é  $(1,1)$  em ambos  $T_1$  e  $T_2$ .

Toda essa estrutura é descrita globalmente pela chamada *fibração de Hopf*. Cada ponto  $(x, y)$  do plano perfurado  $\mathbb{C}^2 \setminus \{(0,0)\}$  pertence a uma única reta complexa que passa pela origem, ou seja, define um elemento de  $P^1(\mathbb{C})$ . Em outras palavras, uma reta através da origem tem uma equação  $y = \lambda x$  com  $\lambda$  pertencendo a  $\mathbb{C} \cup \{\infty\}$ , identificado com a esfera de Riemann ou com uma esfera bidimensional  $S^2$ . Isso define uma aplicação

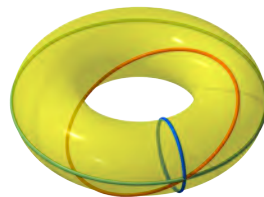
$$\pi : S^3 \rightarrow S^2$$

cujas fibras são círculos, intersecções de retas complexas com a esfera. Quaisquer duas fibras estão entrelaçadas.

Aqui estão algumas figuras da fibração de Hopf, sob a projeção estereográfica, extraídas do filme [Dimensions](#).



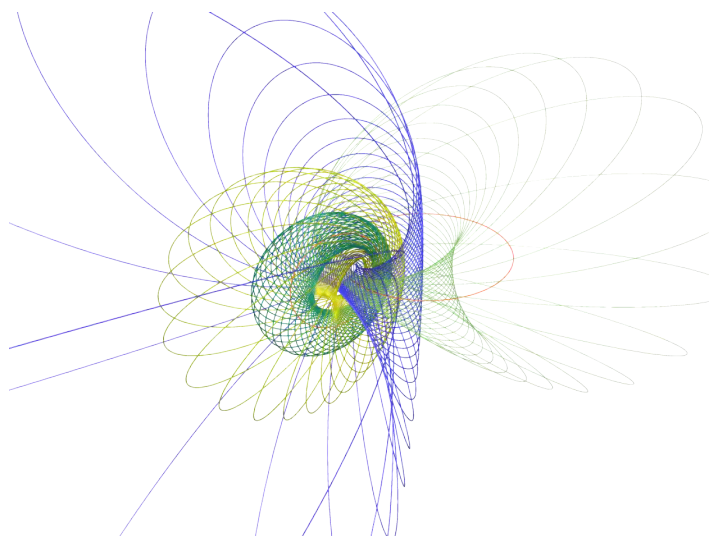
Qual é a versão real da fibração de Hopf? Ela existe, mas é um pouco decepcionante. Cada ponto  $(x, y)$  de  $\mathbb{R}^2 \setminus \{(0,0)\}$  pertence a uma reta única que passa pela origem e define um elemento de



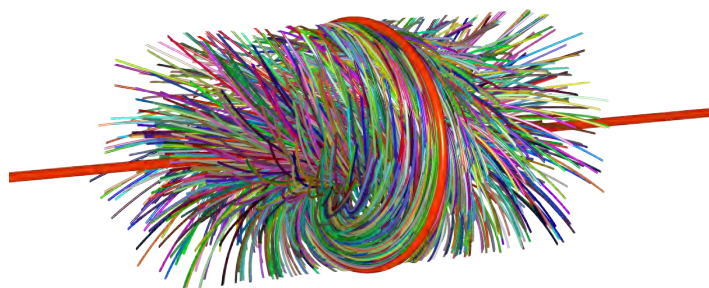
Um meridiano, um paralelo e um  $(1,1)$ -círculo em um toro. †

Claro, esta aplicação não foi inventada por Hopf! Sua contribuição mostrou que não é homotópico para uma aplicação constante, mas isso é outra história.

A fibração de Hopf: cada círculo é uma fibra de  $\pi$ . A imagem inversa de um círculo por  $\pi$  é um toro de Clifford, que é uma união de fibras (da mesma cor na imagem). †



A mesma imagem acima, depois de uma rotação da esfera de dimensão 3, que corresponde a uma aplicação conforme em  $\mathbb{R}^3 \cup \{\infty\}$ . +



Círculos de Hopf, na vizinhança de um deles, projetado como uma reta no espaço (em vermelho). Uma vez, não me esqueci, um dos meus leitores me disse que ele vê essa figura como "assustadora". Você concorda com ele? +

$P^1(\mathbb{R})$ . Essa reta tem a forma  $y = \lambda x$  tal que  $\lambda$  pertence a  $\mathbb{R} \cup \{\infty\}$ , uma esfera unidimensional, i.e., um círculo  $S^1$ . Isso define uma aplicação

$$\pi : S^1 \rightarrow S^1$$

cujas fibras são  $S^0$ , interseções de retas reais com o círculo unitário. Essa é apenas a multiplicação por 2 em  $\mathbb{R}/\mathbb{Z}$ . Não esqueça que uma esfera de dimensão zero é um par de pontos.

### *Links de Hopf*

Uma fibra Hopf é apenas um círculo redondo na esfera, então, não há muito a dizer sobre isso. (Isso não é bem verdade: a geometria do espaço de círculos no espaço de dimensão 3 é maravilhosa. Veja o livro moderno de Cecil <sup>95</sup> ou o clássico Vorlesungen <sup>96</sup> de Blaschke).

*Dois círculos Hopf* são mais interessantes, pois definem o link não trivial mais simples. Observe que, mesmo estando entrelaçados, os dois círculos limitam um anel. De fato, olhe para a pré-imagem pela fibração de Hopf de algum arco conectando dois pontos: é um anel.

*Três círculos Hopf* (ou mais) dão um *link de Hopf*. Cada componente é um círculo e quaisquer duas componentes são entrelaçadas uma vez. É fácil encontrar uma superfície orientável com tal link como borda. De fato, vamos considerar  $n$  números complexos  $\lambda_1, \dots, \lambda_n$  e o polinômio

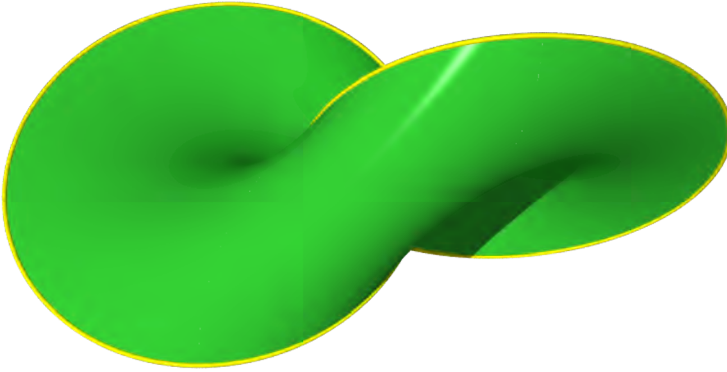
$$F(x, y) = (y - \lambda_1 x)(y - \lambda_2 x) \cdots (y - \lambda_n x).$$

A interseção da esfera de dimensão 3 com o conjunto de  $(x, y)$  tal que  $F(x, y)$  é um número real positivo é uma superfície cuja borda consiste em  $n$  círculos de Hopf. Tudo isso será bastante generalizado nos capítulos seguintes.

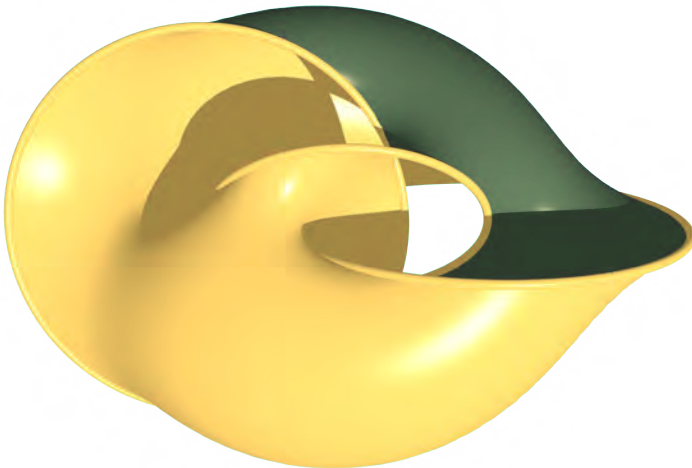
<sup>95</sup> T. E. Cecil. *Lie sphere geometry. With applications to submanifolds*. Universitext. Springer-Verlag, New York, 1992.

<sup>96</sup> W. Blaschke. *Vorlesungen über Differentialgeometrie und geometrische Grundlagen von Einsteins Relativitätstheorie. Band I. Elementare Differentialgeometrie*. Dover Publications, New York, N. Y., 3 edition, 1945.

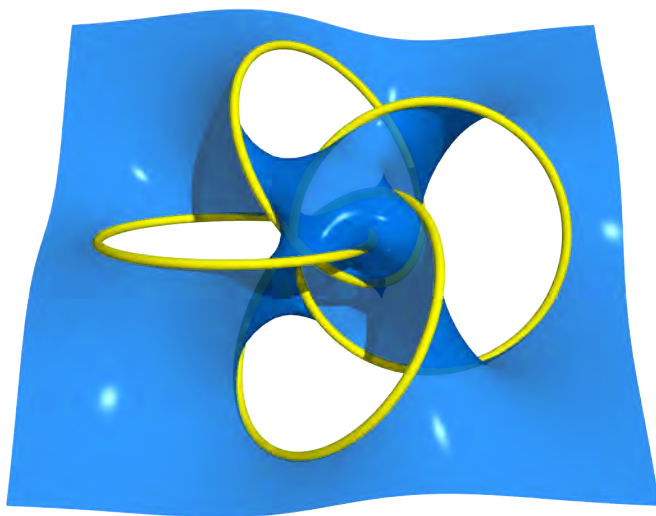
Dois círculos de Hopf,  
limitando um anel. †



Três círculos de Hopf,  
limitando uma superfície. †







Quatro círculos de Hopf, limitando uma superfície. †

### *Dante, La Divina Commedia e a esfera de dimensão 3*

Foi argumentado por Mark Peterson <sup>97</sup> que o universo de Dante, como descrito na *Divina Comédia*, é homeomorfo a uma esfera de dimensão 3. Mesmo que eu não esteja totalmente convencido de que “seja claro que Dante inventou a noção de variedade <sup>98</sup>”, eu gosto desta visão cosmológica. Pelo menos, responde a uma pergunta que todos nós perguntamos a nossos pais quando éramos crianças: o que acontece quando chegamos ao limite do universo? Bem, o universo de Dante é uma variedade de dimensão 3 compacta sem borda.

Deixe-me lembrar que o mundo herdado dos gregos antigos é finito <sup>99</sup>. Basicamente, no sistema de Aristóteles–Ptolomeu, a Terra está fixada no centro do Universo e é cercada por sete esferas celestes, cada uma carregando um “planeta”: Lua, Mercúrio, Vênus, Sol, Marte, Júpiter e Saturno. Uma oitava esfera carrega as *estrelas fixas*. Finalmente, uma nona esfera, chamada *Primum Mobile*, serve como um recipiente para o sistema completo e gera o movimento das outras esferas.

Esta seção, obviamente, não é necessária para o restante do livro.

<sup>97</sup> M. A. Peterson. [Dante and the 3-sphere](#). *American Journal of Physics*, 47:1031–1035, 1979.

<sup>98</sup> M. A. Peterson. The geometry of *paradise*. *Math. Intelligencer*, 30(4):14–19, 2008.

<sup>99</sup> A. Koyré. [From the Closed World to the Infinite Universe](#). Johns Hopkins University Press, 1957.



Gravura de Flammarion (1888). †

Sistema geocêntrico de Aristóteles e Ptolomeu. †

O mundo é considerado, portanto, uma bola de dimensão 3, cuja borda é o *Primum Mobile*. Além desse limite, começa o reino do chamado *Empíreo* cuja natureza não é muito clara. De acordo com Aristóteles<sup>100</sup>: “é claro que [não contém] nem espaço, nem vazio, nem tempo”.

A *Divina Comédia* é um longo poema escrito entre 1308 e 1320 que oferece uma descrição fascinante do mundo cristão no século XIV<sup>101 102</sup>. Dante nos conta sobre sua jornada através do Inferno, do Purgatório e do Paraíso, como uma alegoria para a salvação das almas. Na parte final, no Paraíso, sua musa Beatrice o ajuda a visitar sucessivamente as nove esferas celestes. Quando ele chega ao *Primum Mobile*, ele pode contemplar o mundo a partir de sua fronteira, com a minúscula Terra no centro. De repente, ele volta e descobre que o *Empíreo* tem exatamente a mesma estrutura que o mundo sensível. Consiste no mesmo número de esferas que estão agora centradas em Deus. Estas esferas angélicas, simétricas às esferas celestes, têm os seguintes nomes (indo de Deus ao *Primum Mobile*): Serafins, Querubins, Tronos, Dominações, Virtudes, Poderes, Principados, Arcanjos, Anjos.

*O universo de Dante é, portanto, a união de duas bolas de dimensão*

É claro que a Lua e o Sol não são planetas, mas giram em torno da Terra ... no sistema geocêntrico.

<sup>100</sup> Aristotle. [On the Heavens](#). Traduzido por J. L. Stocks.

<sup>101</sup> D. Alighieri. [Divina Commedia](#). Recomendo essa bela tradução comentada para o Inglês de 2017.

<sup>102</sup> D. Alighieri. [Divina Commedia](#). Nova Fronteira, 2003. Tradução para Português Brasileiro por Xavier Pinheiro (1822–1882).

3, coladas ao longo do *Primum Mobile*. Portanto, é uma esfera de dimensão 3. □

Veja o livro<sup>103</sup> para uma discussão muito mais profunda sobre a visão medieval do Universo e o pequeno livro<sup>104</sup> para fruir mais poesia e matemática.

<sup>103</sup> J. Grzybowski. *Cosmological and Philosophical World of Dante Alighieri: "The Divine Comedy"*. Peter Lang GmbH, 1st new edition, 2015.

<sup>104</sup> R. Osserman. *Poetry of the universe, From the Divine Comedy to Riemann and Einstein*. Anchor, 1995.

Os últimos versos da Comédia:

*“A l’alta fantasia qui mancò possa;  
ma già volgeva il mio disio e ’l velle,  
sì come rota ch’igualmente è mossa,  
l’amor che move il sole e l’altre stelle.”*

*Para a alta fantasia aqui o poder falhou;  
mas já meu desejo e minha vontade foram rolados -  
mesmo como uma roda que se move igualmente -  
pelo Amor que move o sol e as outras estrelas.*



Rosa Celeste: Dante e Beatrice olhando o céu, a Empyrean (figura por Gustave Doré).

†



Uma página do livro aberto  
de Milnor associado à curva  
 $y^2 - x^3 = 0$ . †

# A cúspide e o trevo

O PROPÓSITO DO NOSSO PASSEIO é descrever a topologia das singularidades das curvas analíticas *reais*. Conforme explicado anteriormente, um atalho através do domínio *complexo* pode possivelmente lançar alguma luz sobre a nossa discussão “real”. Em qualquer caso, neste livro, estamos mais interessados nos desvios do que nos atalhos.

Para uma descrição completa da topologia das singularidades das curvas algébricas *complexas*, eu recomendo fortemente o excelente livro <sup>105</sup> por Brieskorn and Knörrer, de 721 páginas. Entretanto:

Un petit livre est rassurant.

Como escreveu Jules Tannery no prefácio da sua bela e concisa introdução <sup>106</sup> à teoria de Galois. Seguindo esse conselho, irei me limitar a características básicas da teoria com o objetivo único de convencer o leitor da riqueza inacreditável da topologia local de uma singularidade no domínio complexo.

## O link de uma singularidade

A ideia de intersectar uma curva analítica complexa  $F(x, y) = 0$  por uma pequena esfera  $S_\varepsilon^3 = \{(x, y) \in \mathbb{C}^2 \mid |x|^2 + |y|^2 = \varepsilon^2\}$  é provavelmente bastante antiga. O primeiro trabalho explicando essa construção foi feito por Brauner <sup>107</sup>, publicado em 1928, que seguiu a ideia proposta pelo seu orientador de doutorado, Wirtinger, em 1905. Veja <sup>108</sup> para uma apresentação inspiradora

<sup>105</sup> E. Brieskorn and H. Knörrer. *Plane algebraic curves*. Birkhäuser/Springer Basel AG, Basel, 1986.

<sup>106</sup> H. Vogt. *Leçons sur la résolution algébrique des équations*. Librairie Nony, 1895.

<sup>107</sup> K. Brauner. Zur Geometrie der Funktionen zweier komplexer Veränderliche. *Abh. Math. Sem. Univ. Hamburg*, 6(1):1–55, 1928.

<sup>108</sup> M. Epple. [Branch points of algebraic functions and the beginnings of modern knot theory](#). *Historia Math.*, 22(4):371–401, 1995.

É importante lembrar que toda curva sobre os números complexos tem dimensão 1 sobre  $\mathbb{C}$  e dimensão 2 sobre  $\mathbb{R}$ ; então uma curva complexa é uma superfície real. Esse equilíbrio constante entre curvas e superfícies é um dos encantos da teoria.

do desenvolvimento histórico dessas ideias.

Anteriormente nós tínhamos olhado o caso simples  $F(x, y) = y - \lambda x$ , considerando a fibração de Hopf. Agora nós vamos ver um segundo e significativo exemplo: a *singularidade cuspidal*, definida por  $F(x, y) = y^2 - x^3$ .

Como escolher um raio pequeno  $\varepsilon$ ? Poderíamos usar também alguma outra hiperfície, como por exemplo um elipsoide?

A palavra *hiperfície* foi introduzida por Elon Lages Lima que achava que “*hypersurface*” tem pouco sentido em inglês sendo que “*hyper*” e “*sur*” significam a mesma coisa em grego e latim respectivamente!

A resposta é que sob suposições muito moderadas, todas essas interseções definem o *mesmo* objeto topológico, a menos de homeomorfismos. O caso de  $y^2 - x^3 = 0$  é particularmente simples. Considere o seguinte fluxo linear em  $\mathbb{C}^2$ :

$$\varphi^t(x, y) = (e^{2t}x, e^{3t}y) \quad t \in \mathbb{R}.$$

O espaço  $\mathcal{O}$  das órbitas de  $\varphi^t$  em  $\mathbb{C}^2 \setminus \{(0, 0)\}$  é homeomorfo a  $\mathbb{S}^3$ . De fato, ao longo de cada órbita, a norma  $|x|^2 + |y|^2$  é estritamente crescente e cada órbita intersecta cada órbita a esfera exatamente uma vez. O mesmo argumento poderia ser usado com um elipsoide centrado na origem, ou com nossa “esfera quadrada”  $\max(|x|, |y|) = \varepsilon$ , ou com muitas outras hiperfícies.

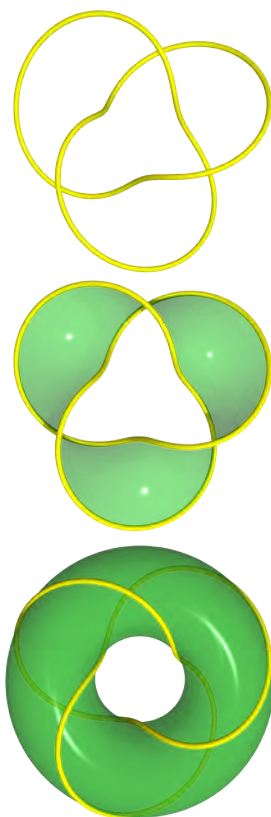
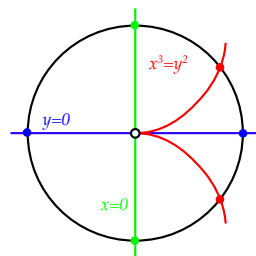
Agora observe que o fluxo  $\varphi^t$  preserva nossa curva, cuja equação é  $y^2 - x^3 = 0$ , que define *canonicamente* um subconjunto  $K$  de  $\mathcal{O}$ . Se identificarmos  $\mathcal{O}$  com  $\mathbb{S}^3_\varepsilon$ , nós realizamos que, a menos de homeomorfismos, a interseção de uma curva com  $\mathbb{S}^3_\varepsilon$  é de fato independente de  $\varepsilon$  e que poderíamos também usar a esfera quadrada.

Então, vamos intersectar  $y^2 - x^3 = 0$  com  $\max(|x|, |y|) = \varepsilon$ . Se  $\varepsilon < 1$ , é parametrizada por  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$

$$x = \varepsilon \exp(2i\theta) \quad y = \varepsilon^{3/2} \exp(3i\theta)$$

no toro sólido (sendo  $|x| = \varepsilon$  e  $|y| \leq \varepsilon$ ). Esse é o *nó de trevo*, visto como um *toro nó* de tipo  $(3, 2)$ : construído e é desenhado em um toro padrão de revolução no espaço 3-dimensional, que dá três voltas ao redor do meridiano e duas ao longo do paralelo.

Existem muitos livros excelentes sobre a topologia dos nós. Recomendo o “pequeno livro” de Sossinsky <sup>109</sup> e o livro de



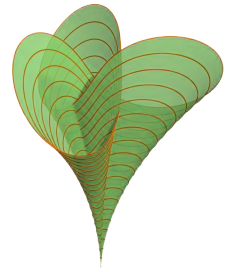
O nó de trevo.

<sup>109</sup> A. Sossinsky. *Knots. Mathematics with a twist.* Harvard University Press, Cambridge, MA, 2002.

Kauffman <sup>110</sup>, bastante visual.

A fim de entender a topologia da curva cuspidal  $y^2 - x^3 = 0$ , na pequena bola  $|x|^2 + |y|^2 \leq \varepsilon^2$ , é suficiente notar que todas as esferas concêntricas intersectam a curva em um trevo. Segue que em uma bola pequena, nossa curva é homeomorfa a um cone topológico sobre o nó de trevo. O trevo é um círculo, mergulhado de uma maneira não trivial. Então, o cone é um disco mergulhado de uma maneira complicada no espaço de dimensão 4. Sobre os números complexos, a curva é topologicamente suave: é localmente homeomorfa a um disco, mas ele está mergulhado em  $\mathbb{C}^2$  de uma forma não é trivial. Esse é um exemplo típico de um fenômeno que pode ser identificado sobre os números complexos, podendo ser invisível sobre os números reais, visto que temos para a curva real  $y^2 - x^3 = 0$  e todos os ramos de uma curva analítica real são localmente homeomorfos a uma reta do plano: essa foi o que chamamos anteriormente de afirmação de Gauss.

<sup>110</sup> L. H. Kauffman. *On knots*, volume 115 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, NJ, 1987.



A figura local de um ramo (projetado no espaço tridimensional), onde o ramo não está mergulhado.

### Fibração de Milnor

Descrevemos agora a curva cúspide e mostramos um caso muito especial de um teorema geral de Milnor que será apresentado posteriormente com mais detalhes.

Considere a aplicação:

$$\mu : (x, y) \in \mathbb{S}^3 \mapsto y^2 - x^3 \in \mathbb{C}.$$

A imagem inversa do 0 é um nó de trevo. Queremos ver a imagem inversa  $\Sigma_\theta$  da semirreta saindo da origem, cuja equação é  $\arg(z) = \theta \in \mathbb{R}/2\pi\mathbb{Z}$ . Em outras palavras, desejamos ver as fibras da aplicação composta  $\arg \circ \mu$  definidas no complemento do nó de trevo com valores no círculo  $\mathbb{R}/2\pi\mathbb{Z}$ . É fácil ver que  $\arg \circ \mu$  é uma submersão. De fato, o fluxo

$$\psi^s(x, y) = (e^{2is}x, e^{3is}y)$$

preserva as esferas e satisfaz  $\arg(\mu \circ \psi^s) = \arg(\mu) + 6s$ . Segue que o campo de vetores associado a  $\psi^s$  não está no núcleo da diferencial de  $\arg \circ \mu$ . Observe que  $\psi^s$  permuta as fibras  $\Sigma_\theta$ .

Na vizinhança do nó de trevo, a situação é muito fácil de analisar. Nós ainda podemos utilizar a esfera quadrada, visto





que  $\arg \circ \mu$  é invariante por  $\varphi^t$  e nós ainda estamos de fato no espaço das órbitas  $\mathcal{O}$ . Parametrizar uma vizinhança de um trevo por pares  $(\theta, \zeta)$  com  $\theta \in \mathbb{R}/2\pi\mathbb{Z}$  e  $\zeta$  um pequeno número complexo:

$$x = \varepsilon \exp(2i\alpha) \quad y = \varepsilon^{3/2} (\exp(3i\alpha) + \exp(-3i\alpha)\zeta).$$

Nessas coordenadas,  $\arg \circ \mu$  é igual a  $\arg(\zeta)$  na primeira ordem. Segue que na vizinhança do trevo, os  $\Sigma_\theta$  são superfícies que se comportam como as páginas de um livro em torno da sua lombada.

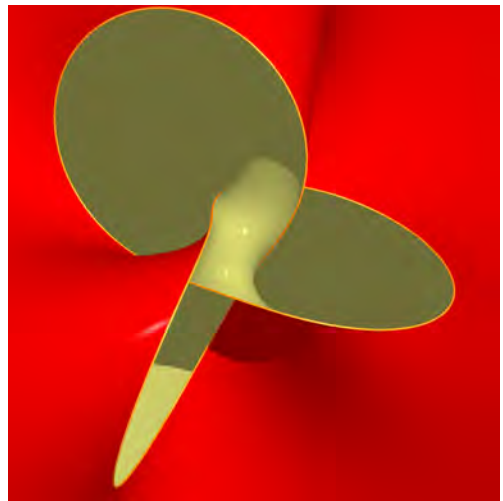
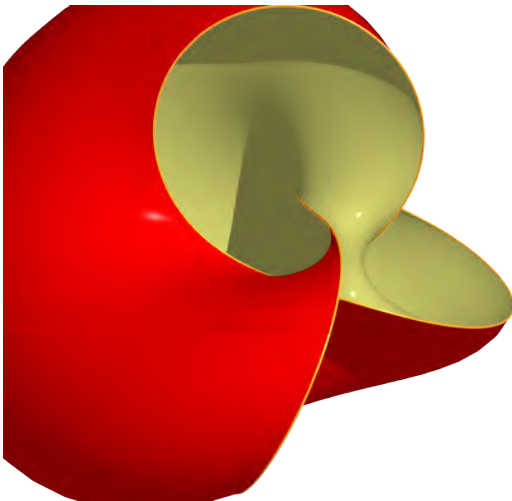
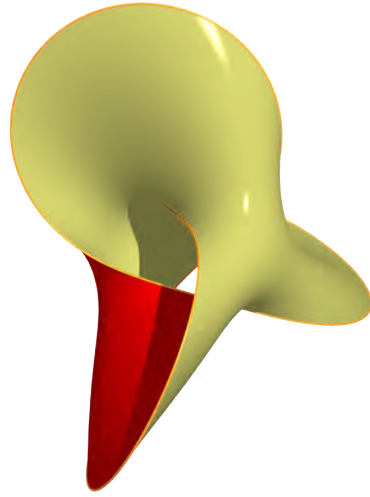
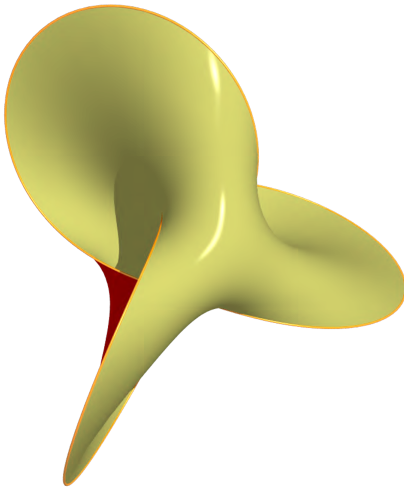


Um livro estranho em que as páginas são ordenadas ciclicamente e sem primeira página. Um livro de sonhos que você lê para sempre. Pelo menos as páginas são orientáveis.

Afirmarei, então, que o nó de trevo é *fibrado* ou que seu complemento fibra sobre o círculo. As fibras são *páginas* disjuntas cujos fechos na esfera de dimensão 3 têm todos o nó como fronteira comum. Observe que uma página passa pelo polo norte na esfera de dimensão 3, o centro da projeção estereográfica. Essa página, quando projetada no espaço euclidiano tridimensional, não é compacta.

### *Monodromia*

O fluxo  $\psi^s(x, y) = (e^{2is}x, e^{3is}y)$  permuta as páginas do nosso livro. Mais precisamente,  $\psi^s$  leva a página  $\Sigma_\theta$  na página  $\Sigma_{\theta+6s}$ . Note que  $\psi^{2\pi}$  é a identidade e que  $\psi^{\pi/3}$  fixa globalmente cada página, induzindo um homeomorfismo de ordem 6, para cada página, denominado de *monodromia* da cúspide.



Algumas páginas.

†

Agora nosso objetivo é descrever a topologia das páginas e a ação da monodromia.

Por definição, uma página  $\Sigma$  é o conjunto dos pares  $(x, y)$  em  $S^3$  tal que o número complexo  $y^2 - x^3$  está em alguma semirreta, por exemplo, o eixo real positivo.  $\mathbb{R}_+^* \subset \mathbb{C}$ . Seja  $\mathcal{C}$  a curva algébrica definida por  $y^2 - x^3 = 1$  em  $\mathbb{C}^2$ . Pense na esfera de dimensão 3 como o espaço de órbitas do fluxo  $\varphi^t(x, y) = (e^{2t}x, e^{3t}y)$  agindo sobre  $\mathbb{C}^2 \setminus \{(0, 0)\}$ . As duas superfícies reais  $\Sigma$  e  $\mathcal{C}$  definem um mesmo objeto nesse espaço de órbitas, então, vamos trabalhar com  $\mathcal{C}$ . A ação da monodromia corresponde a

$$(x, y) \in \mathcal{C} \mapsto (\omega^2 x, \omega^3 y) \in \mathcal{C}$$

sendo  $\omega = \exp(2i\pi/6)$  uma raiz primitiva 6-ésima da unidade.

A topologia de  $\mathcal{C}$  é fácil de descrever se você conhece algo sobre o gênero da superfície de Riemann ou das curvas algébricas. Em  $P^2(\mathbb{C})$ , a curva cúbica homogeneizada  $y^2z - x^3 = z^3$  é uma curva elíptica suave intersectando triplamente a linha do infinito no ponto  $[0 : 1 : 0]$ . “Então”, a curva afim  $\mathcal{C}$  é homeomorfa a um toro perfurado uma vez.

De uma forma mais elementar, procedemos da seguinte maneira. Seja  $Y = y^2$ , de modo que a aplicação  $(x, y) \in \mathcal{C} \mapsto Y$  é um recobrimento com seis folhas de  $\mathbb{C}$ , ramificado no 0 e 1, com ordens 2 e 3.

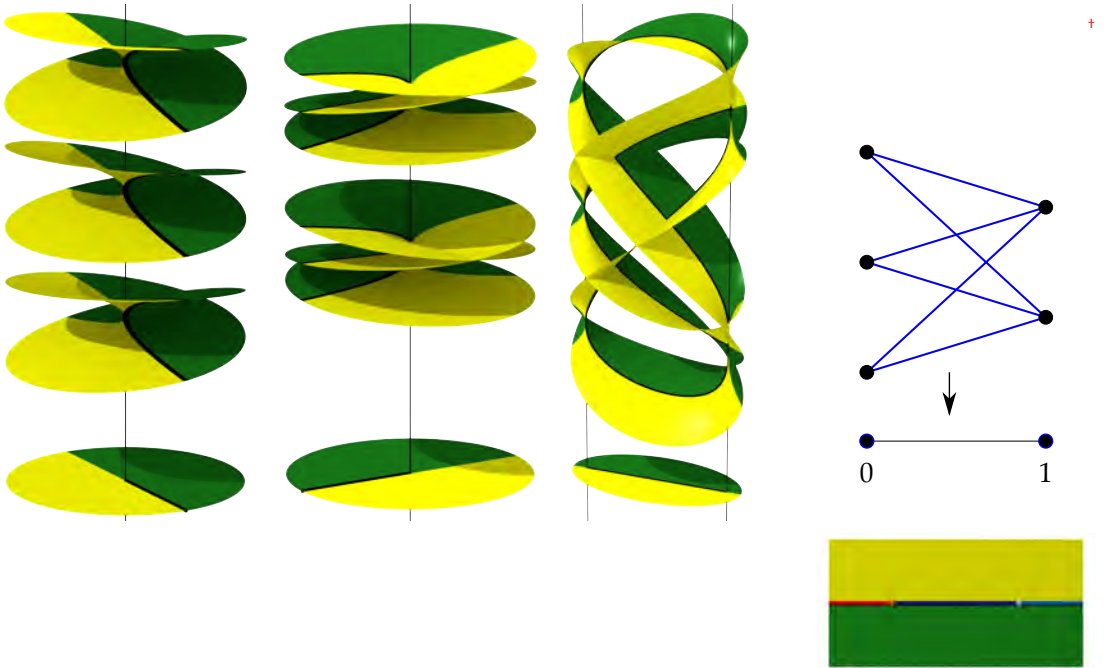
Construa um arco no plano complexo conectando  $Y = 0$  e  $Y = 1$ . Ele pode ser levantado em seis arcos em  $\mathcal{C}$

Acima de  $Y = 0$ , existem 3 pontos, onde os 6 arcos se reúnem em 3 grupos de 2. Acima de um pequeno círculo centrado em 0, existem 3 *placas duplas*, como na parte esquerda da imagem a seguir. Acima de  $Y = 1$ , como na segunda figura, existem dois pontos, onde os arcos se reúnem em dois grupos de três.

Como é impossível desenhar no espaço 4-dimensional  $\mathbb{C}^2$ , essas imagens representam o gráfico de alguma combinação apropriada das partes real e imaginária de  $\sqrt{Y} + \sqrt[3]{1-Y}$ .

A combinatória dos seis arcos é representada na margem direita.

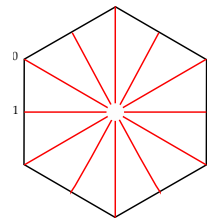




Cortando  $C$  ao longo de três arcos, de  $-\infty$  a  $0$ , de  $0$  a  $1$ , e de  $1$  a  $\infty$ , decompos  $C$  em dois *triângulos*, sendo a parte imaginária positiva ou negativa (amarelo e verde na figura). Esses são, de fato, os triângulos cujos vértices são  $0, 1$  and  $\infty$ . Em  $C$ , produz-se no total 18 arcos, e  $12 = 6 \times 2$  triângulos.

Há um outro caminho possível para vermos os nós na mesma figura. Para isso, identifique os lados opostos de um hexágono regular por translações. Dessa forma, obteremos um toro que ao ser apagado o centro, teremos um toro perfurado.

A partir do centro do hexágono, construa os 6 segmentos indo para os vértices e as 6 alturas dos lados. Nosso toro está decomposto em 12 triângulos, tendo no total 18 lados. As seis raízes da unidade agem por rotações no hexágono (furado), permutando os triângulos exatamente como no caso de  $C$ .



A apresentação mais econômica e comum de um toro é a partir de um quadrado, mas a nossa apresentação com o hexágono é a mais bela. Observe que os seis lados definem três arcos no toro e os seis vértices definem dois pontos nesse toro.

Resumindo, *cada página do livro associado a cúspide é um toro furado como o toro acima, e a monodromia é simplesmente uma rotação de  $1/6$  de um volta completa.*

### Nós do tipo toro

A maior parte do que foi observado para a cúspide  $y^2 = x^3$  estende-se a uma curva geral  $F(x, y) = 0$ . Isso exigirá algum trabalho, mas há pelo menos uma família de exemplos na qual não há trabalho extra. Sejam  $p, q$  dois inteiros positivos relativamente primos e vejamos a curva  $y^p - x^q = 0$ . Podemos assumir  $q > p$ .

Assim como antes, olhamos para a interseção com a esfera quadrada.

$$x = \varepsilon \exp(ip\theta) \quad y = \varepsilon^{q/p} \exp(iq\theta).$$

Isso é um nó toral  $K_{p,q}$  traçado num toro padrão no espaço 3-dimensional, indo  $p$  vezes em torno de um paralelo e  $q$  vezes de um meridiano.

Exatamente pela mesma razão, há uma decomposição como um livro aberto e uma fibração sobre o círculo. Qualquer página é homeomorfa a curva algébrica afim  $y^p - x^q = 1$ , cuja topologia pode ser descrita da mesma maneira. Defina  $Y = y^p$  e observe esta curva como espalhada em cima de  $Y$  e ramificada em 0 e 1. Sobre um ponto  $Y$  diferente de 0, 1, existem  $pq$  pontos. Acima de 0 (resp. 1), existem apenas  $q$  (resp.  $p$ ) pontos, mas cada um com multiplicidade  $p$  (resp.  $q$ ). Substitua o hexágono por um  $pq$ -gono e a situação é a mesma. Existem  $2pq$  triângulos ( $pq$  de cada cor) e  $3pq$  arestas. Existem  $p$  vértices acima de 0,  $q$  acima de 1 e um acima do infinito.

Isso dá um número de Euler–Poincaré igual a

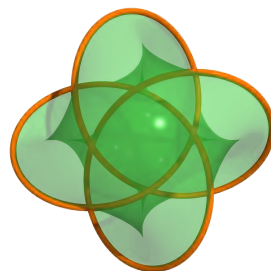
$$p + q + 1 - 3pq + 2pq = 2 - 2 \frac{(p-1)(q-1)}{2}.$$

Portanto, cada página é uma superfície de gênero  $(p-1)(q-1)/2$  furada num ponto.

Esboço, agora, uma demonstração de um fato fundamental: a topologia recupera boa parte da curva algébrica  $y^p = x^q$ .

**Teorema.** *Se algum homeomorfismo da esfera de dimensão 3 envia o nó toral  $(p_1, q_1)$  em  $(p_2, q_2)$ , então, os conjuntos  $\{p_1, q_1\}$  e  $\{p_2, q_2\}$  são iguais.*

A demonstração requer um pouco de topologia algébrica. Da



Um nó toral (3,4).

+

No prefácio, afirmei que tentaria escrever um livro que até um estudante de graduação entenderia. Para o final deste capítulo, temo que não seja esse o caso. Se for o caso, basta ignorá-lo!



topologia do nó toral, nós construímos um truque algébrico que nos permitirá reconstruir  $p, q$ .

O complemento do nó toral  $(p, q)$  na esfera de dimensão 3 é uma 3-variedade. O seu invariante mais primitivo é o seu grupo fundamental, representado por  $\Gamma_{p,q}$ . O ponto chave é extrair algebricamente  $p$  e  $q$  desse grupo. Nós vamos provar que  $\Gamma_{p_1,q_1}$  é isomorfo a  $\Gamma_{p_2,q_2}$  somente se  $\{p_1, q_1\} = \{p_2, q_2\}$ .

Primeiro, observe que a aplicação  $\arg \mu = \arg(y^p - x^q)$  do complemento do nó toral  $(p, q)$  na esfera de dimensão 3 para  $\mathbb{R}/2\pi\mathbb{Z}$  pode induzir um homomorfismo sobrejetivo.

$$\lambda : \Gamma_{p,q} \rightarrow \pi_1(S^1) \simeq \mathbb{Z}$$

entre os grupos fundamentais. De fato, considere o seguinte laço na esfera unitária  $t \in [0, 1]$  definido por:

$$x(t) = \frac{\sqrt{2}(1 + \zeta(t))}{2|1 + \zeta(t)|} \exp(2ip\pi t) \quad y = \frac{\sqrt{2}}{2} \exp(2iq\pi t).$$

Para valores pequenos de  $\zeta$ , esse é um pequeno laço que dá uma volta ao redor do nó toral  $(p, q)$ . O argumento de  $\mu$  desse laço é próximo a

$$\pi + \arg \zeta(t) + 2\pi pqt$$

de modo que podemos escolher  $\zeta(t) = \varepsilon \exp(i(1 - 2pqt)\pi)$  e garantir que a imagem deste laço por  $\lambda$  é 1.

No primeiro passo, nós mostramos que esse homeomorfismo, a menos de sinal, é a única sobrejeção de  $\Gamma_{p,q}$  em  $\mathbb{Z}$ . Segue que o núcleo de  $\lambda$  depende apenas da topologia do nó.

No segundo passo, nós vamos analisar a abelianização do núcleo de  $\lambda$ , mostrando que é um grupo abeliano livre finitamente gerado, e encontrar seu posto. Isso nos permitirá recuperar  $\{p, q\}$  a partir do grupo  $\Gamma_{p,q}$  conforme necessário.

O primeiro passo pode ser explicado de várias maneiras, mais ou menos sofisticadas, a maioria baseada na dualidade de Lefschetz. Suponha que alguma variedade orientável e fechada  $X$ , por exemplo, o círculo, seja mergulhada em uma esfera de dimensão 3. Então, a homologia do complementar de  $X$  não depende do modo como a variedade  $X$  é mergulhada na esfera. Em particular, a homologia do complementar de um nó na esfera

Os iniciantes são encorajados a olhar para o notável site de Henri Paul de Saint Gervais dedicado a [Análise Situs](#).

de dimensão 3 é a mesma que no caso de um nó trivial, então, essa homologia é simplesmente isomorfa a  $\mathbb{Z}$  no grau 1.

Eu poderia apresentar o mesmo fato da seguinte maneira: Seja  $\gamma$  um laço suave em  $S^3 \setminus K_{p,q}$ . Desde que a esfera seja simplesmente conexa,  $\gamma$  é a borda de alguma aplicação suave  $D \rightarrow S^3$  que poderia não ser um mergulho. Ponha esse disco numa posição genérica com  $K_{p,q}$ , tal que as interseções entre  $K_{p,q}$  e  $D$  sejam transversais. Conte o número de interseções entre os discos  $D$  e  $K_{p,q}$ , sendo a contagem algébrica, tendo em conta as orientações. Esse número é o *número de entrelace*  $lk(\gamma)$ . Acontece que só depende da classe de homologia de  $\gamma$  em  $S^3 \setminus K_{p,q}$ . Isso segue do fato de que uma superfície sem borda em  $S^3$  tem uma interseção algébrica trivial com qualquer curva fechada.

Portanto, define-se um homomorfismo sobrejetivo:

$$lk : H_1(S^3 \setminus K(p, q), \mathbb{Z}) \rightarrow \mathbb{Z}$$

Agora, se  $\gamma$  está no núcleo de  $lk$ , isso significa que sinais  $+$  e  $-$  podem ser acoplados na interseção. Cave buracos em  $D$ , em torno dos pontos de interseção, e conecte suas bordas em pares com tubos, a fim de construir uma superfície cuja borda ainda é  $\gamma$  e que não intersecta mais  $K(p, q)$ . Desse modo, mostramos que os elementos do núcleo de  $lk$  são homólogos a zero. Em outras palavras,  $lk$  é um isomorfismo. Finalmente, lembre-se que o primeiro grupo de homologia é a abelianização do grupo fundamental, de modo que qualquer homomorfismo  $\pi_1(S^3 \setminus K) \rightarrow \mathbb{Z}$  fatoriza por meio de  $lk$ . Segue-se em particular que  $lk$  coincide (a menos do sinal) com o  $\lambda$  definido previamente. Esse é o primeiro passo.

Vamos agora proceder o segundo passo. Denote o núcleo de  $lk$  por  $G(p, q)$ . É, portanto, o grupo fundamental de algum recobrimento de Galois de  $S^3 \setminus K(p, q)$  cujo grupo de automorfismos é infinito cíclico  $\mathbb{Z} = \Gamma_{p,q}/\text{Ker } lk$ . Esse recobrimento é claramente o produto  $\Sigma \times \mathbb{R}$  de uma página com  $\mathbb{R}$  e o grupo de transformações do recobrimento sendo simplesmente gerado por

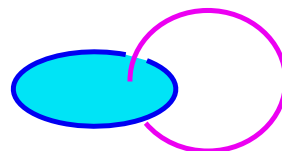
$$(p, t) \in \Sigma \times \mathbb{R} \rightarrow (M(p), t + 2\pi) \in \Sigma \times \mathbb{R}$$

seja  $M$  representando a aplicação de monodromia. Segue que  $G(p, q)$  é o grupo fundamental de uma página  $\Sigma$ . Como vocês

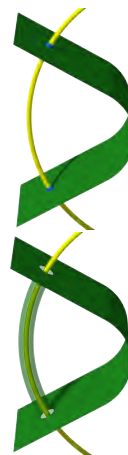
Note que o complemento de um nó trivial na esfera de dimensão 3 é homeomorfo a  $\mathbb{R}^2 \times S^1$ .

Qualquer aplicação de  $S^1$  para  $S^3$  estende-se ao disco unitário  $D^2$ . Se essa extensão for um mergulho, então,  $\gamma$  é um nó trivial. Entretanto, qualquer nó é borda de uma superfície orientável (de gênero maior): essa é a chamada superfície de Seifert.

Mais precisamente esse é o número de entrelace de  $\gamma$  e  $K_{p,q}$ . Voltaremos a esse conceito mais tarde neste passeio, digo, livro.



número de entrelace.



Removendo dois pontos de interseção.

devem lembrar, já escrevemos anteriormente a topologia de uma página. Caso  $G(p, q)$  não seja abeliano, pode ser mais fácil torná-lo abeliano. Denote por  $H(p, q)$  essa abelianização, que não é nada mais do que a primeira homologia de uma página.

Vamos descrever esse grupo abeliano  $H(p, q)$  e a ação de  $M$ . Em  $\Sigma$  existe um grafo contendo  $pq$  arcos, obtido por meio do levantamento do arco 0 e 1. Ele contém  $q$  vértices acima de 0 e  $p$  vértices acima de 1. Relembre que  $\Sigma$  é obtido a partir de uma superfície triangulada fechada por supressão de um vértice que é comum a todos os triângulos. Portanto, a superfície perfurada  $\Sigma$  pode ser deformada para a união de todos as bordas opostas a esse vértice, que corresponde ao nosso grafo com  $pq$  arestas, geralmente denominado de *grafo bipartido completo*. Isso produz um 1-complexo muito simples que calcula  $H(p, q)$ .

O grupo abeliano de 1-cadeias é livremente gerado pelos arcos  $c_{i,j}$  sendo  $i \in \mathbb{Z}/p\mathbb{Z}$  e  $j \in \mathbb{Z}/q\mathbb{Z}$ . O grupo abeliano de 0 cadeias é gerado por  $p$  pontos  $a_i$  e  $q$  pontos  $b_j$  com  $i \in \mathbb{Z}/p\mathbb{Z}$  e  $j \in \mathbb{Z}/q\mathbb{Z}$ . O operador borda mapeia  $c_{i,j}$  em  $b_j - a_i$ . Finalmente, a monodromia  $\mathbb{Z}/pq\mathbb{Z} \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$  age de uma maneira óbvia sobre os índices  $i, j$ .

A homologia  $H(p, q)$  se encaixa em uma sequência exata

$$0 \longrightarrow H(p, q) \longrightarrow \mathbb{Z}^p \otimes \mathbb{Z}^q \xrightarrow{\partial} \mathbb{Z}^p \oplus \mathbb{Z}^q \longrightarrow \mathbb{Z} \longrightarrow 0$$

equivariante em relação às ações de  $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$  para cada nível. O gerador  $M$  de monodromia é associado com a ação de  $(1, 1)$ . Fazendo o produto tensorial por  $\mathbb{R}$ , para obter espaços vetoriais e aplicações lineares, segue a dimensão de  $H(p, q) \otimes \mathbb{R}$  is  $pq - (p + q) + 1$ , i.e.  $(p - 1)(q - 1)$ . O polinômio característico da ação de  $M$  sobre  $H(p, q)$  pode ser calculado usando a seguinte sequência exata:

$$P(X) = \frac{(X^{pq} - 1)(X - 1)}{(X^p - 1)(X^q - 1)}.$$

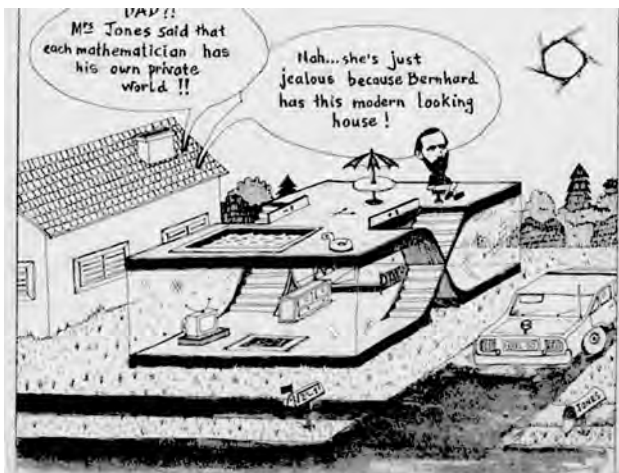
Observe que as raízes de  $P$ , autovalores de  $M_*$ , são as  $pq$ -ésimas raízes da unidade, menos as  $p$ -ésimas e as  $q$ -ésimas raízes, mais 1. A partir deste espectro, os valores de  $p, q$  podem ser extraídos.

A demonstração do teorema está finalizada.



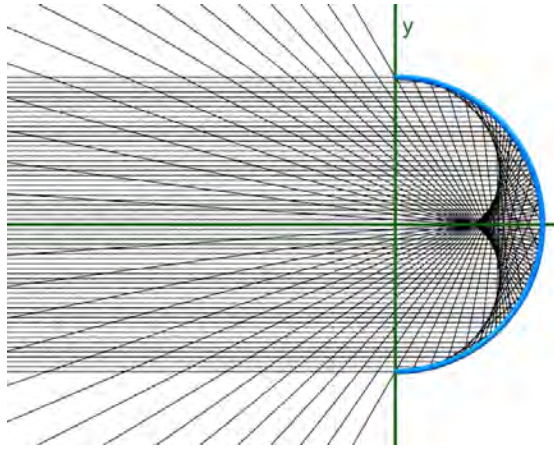
Do grupo fundamental  $\Gamma$  do (complemento do) nó  $K(p, q)$ , construa seu primeiro grupo derivado  $\Gamma_1 = [\Gamma, \Gamma]$ , que é também, como temos visto, o núcleo de  $lk$ . Em seguida, faça  $\Gamma_1$  abeliano e defina o grupo  $\Gamma_1/[\Gamma_1, \Gamma_1]$ . Agora considere algum elemento  $g$  em  $\Gamma$  com  $lk(g) = \pm 1$  e a conjugação por  $g$  em  $(\Gamma_1/[\Gamma_1, \Gamma_1]) \otimes \mathbb{R}$ . Examinando os autovalores desse mapa linear, recuperamos os valores de  $p, q$ .  $\square$ .

Esse truque algébrico é realmente uma técnica muito geral e poderosa e não está restrito aos nós. Dado qualquer grupo  $\Gamma$ , veja a ação da abelianização  $\Gamma_{ab} = \Gamma/\Gamma_1 = \Gamma/[\Gamma, \Gamma]$  por conjugação na abelianização  $(\Gamma_1/[\Gamma_1, \Gamma_1]) \otimes \mathbb{R}$ . Isso define uma família comutativa de automorfismos cujas classes de conjugação são invariantes do grupo  $\Gamma$ . Fala-se do *módulo de Alexander* de  $\Gamma$ , um dos invariantes mais primitivos de um grupo.



Por André Nachbin.

†



Uma cúspide no café. †



Pico Puiseux.



## *Victor Puiseux, finalmente!*

O NOME DE PUISEUX JÁ APARECEU VÁRIAS VEZES NESTE LIVRO.

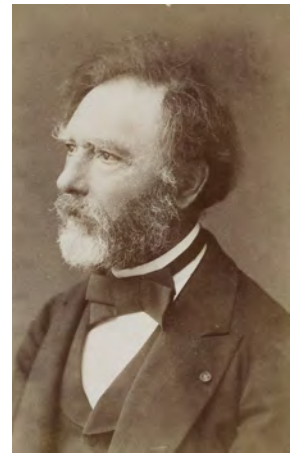
O leitor pode estar ansioso para saber o que ele realmente fez. Infelizmente, o “famoso teorema de Puiseux” não se deve a ele, mas, como vimos, à Newton, com alguma ajuda posterior de Cramer. Pode-se argumentar que nem Newton, nem Cramer provaram a convergência das séries associadas, mas ela pode ser facilmente provada, por exemplo, usando o *calcul des limites* de Cauchy.

No entanto, Puiseux abordou o problema da estrutura local das singularidades de uma maneira totalmente diferente e sua contribuição é fundamental. Neste capítulo, gostaria de explicar ponto de vista dele, mas, infelizmente, seria inútil manter sua apresentação original.

Destino estranho para um matemático: ele é “famoso” por um teorema que era conhecido muito antes dele, e que entendemos hoje muito melhor do que ele, usando técnicas que vieram muito depois dele.

Felizmente, o Puiseux é ainda mais famoso entre os alpinistas, já que o pico mais alto do Monte Pelvoux (3.946 m), no Maciço dos Écrins, é chamado *pico Puiseux*.

Ele chegou a este pico em 9 de agosto de 1848. Infelizmente, novamente, a mesma peculiaridade: nem sequer é possível ter certeza de que ele foi “o primeiro”, já que o capitão Durand afirmou ter alcançado o cume 18 anos antes. Um eterno segundo?



Victor Puiseux (1820–1883) †

## A abordagem topológica de Puiseux

Vamos nos lembrar do que geralmente é chamado de teorema de Puiseux.

**Teorema.** *Seja  $F(x, y)$  uma função holomorfa diferente de zero definida na vizinhança da origem em  $\mathbb{C}^2$  e tal que  $F(0, 0) = 0$ . Então, existe um número finito de funções holomorfas  $g_1, \dots, g_n$  definidas na vizinhança de  $0 \in \mathbb{C}$  e inteiros positivos  $m_1, \dots, m_n$  tais que a curva  $F(x, y) = 0$ , novamente na vizinhança de  $(0, 0)$ , é a união de  $n$  ramos  $t \mapsto (t^{m_i}, g_i(t))$  (para  $i = 1, \dots, n$ ) (mais, possivelmente, o eixo  $y$ ). Além disso, esses ramos são injetivos e só se cruzam na origem.*

Como vocês estão lembrados, já discutimos uma prova pré-Puiseux, de espírito muito algébrico, quando encontramos primeiro as série formais  $g_i$  antes de provar que elas convergem. Puiseux propôs uma abordagem topológica<sup>111</sup> em 1850, pouco antes dos grandes trabalhos de Riemann, apresentando ideias topológicas em geometria algébrica. Devemos, portanto, “perdoá-lo”, pois, é claro, ele não poderia se expressar em termos de superfícies de Riemann.

Deixe-me esboçar essa prova topológica:

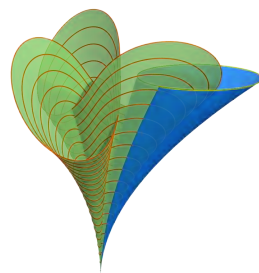
Considere primeiro  $F(0, y)$ . Se for igual a zero,  $F$  pode ser dividido por uma potência de  $x$  sem alterar o problema. Podemos, portanto, assumir que a valoração de  $F(0, y)$  (também chamada de multiplicidade) é um inteiro positivo  $m > 0$ . Em particular  $F(0, y)$  tem um zero isolado na origem (de multiplicidade  $m$ ). Escolha algum  $\varepsilon > 0$  tal que  $0$  é a única raiz de  $F(0, y) = 0$  em  $|y| \leq \varepsilon$ . Por um simples argumento de continuidade, existe algum  $\eta > 0$  tal que não existe raiz de  $F(x, y) = 0$  no toro sólido  $\{(x, y) \mid |x| \leq \eta; |y| = \varepsilon\}$ . Dividindo  $x$  e  $y$  por  $\varepsilon$  e  $\eta$ , assumimos que  $\varepsilon = \eta = 1$ .

Vamos agora fazer algumas suposições que serão analisadas em detalhes mais adiante.

*Suponha que a derivada parcial  $\partial F / \partial y$  não se anula na curva  $F(x, y) = 0$ , exceto na origem.*

Denote por  $\mathcal{C}^*$  a curva perfurada

$$\{(x, y) \in \mathbb{C}^2 \mid (x, y) \neq (0, 0); F(x, y) = 0; |x| \leq 1; |y| \leq 1\}.$$



Dois ramos podem ser entrelaçados.



Dois modelos de [Coleção de modelos e instrumentos matemáticos de Göttingen](#) que o leitor deve definitivamente visitar.

<sup>111</sup> V. Puiseux. [Recherches sur les fonctions algébriques](#). *Journal de Mathématiques Pures et Appliquées*, 15:365–480, 1850.

Não se esqueça que toda essa discussão é local, então, quando escrevo “não se anula”, quero dizer “não se anula em alguma vizinhança da origem”.

A principal afirmação é que a projeção de  $C^*$  sobre o disco perfurado  $D^* = \{x \mid |x| \leq 1\} \setminus \{0\}$  é uma aplicação de recobrimento.

Deixe-me lembrar rapidamente a definição das aplicações de recobrimento e como elas diferem dos homeomorfismos locais.

Uma aplicação contínua  $p : X \rightarrow Y$  é um *homeomorfismo local* (às vezes chamado de aplicação étale) se cada ponto em  $X$  tiver uma vizinhança aberta  $U$  tal que  $p(U)$  é aberto e a restrição de  $p$  a  $U$  é um homeomorfismo em  $p(U)$ .

Uma aplicação contínua  $p : X \rightarrow Y$  é uma *aplicação de recobrimento* se cada ponto em  $Y$  tiver uma vizinhança aberta  $V$  tal que  $p^{-1}(V)$  é uma união disjunta de conjuntos abertos  $U_i$  e tal que a restrição de  $p$  a cada  $U_i$  é um homeomorfismo em  $V$ .

Claramente, uma aplicação de recobrimento é um homeomorfismo local, mas exemplos simples mostram que o inverso não é verdadeiro. É fácil mostrar que um homeomorfismo local é um espaço de recobrimento se for *própria*.

Vamos agora mostrar nossa afirmação de que  $C^*$  é um recobrimento do disco perfurado. O fato de a projeção ser um homeomorfismo local decorre imediatamente de nossa suposição de que  $\partial F/\partial y$  não zera em  $C^*$  e do teorema da função implícita. A propriedade da projeção também é clara, já que uma sequência de pontos em  $C^*$  escapa de um conjunto compacto se, e somente se, convergir para a origem.

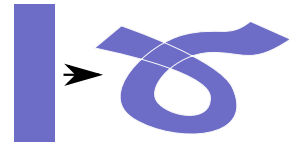
O principal teorema da teoria dos espaços de recobrimento está associado com os espaços de recobrimento conexos (e localmente simplesmente conexo) de um espaço conexo descritos, a menos de isomorfismos, pelos subgrupos do grupo fundamental. Por exemplo, os espaços de recobrimento conexos de  $D^*$  são equivalentes a alguma aplicação potência  $x \in D^* \mapsto x^m \in D^*$ , por algum inteiro  $m \geq 1$ , ou para a exponencial complexa restrita ao semi plano  $\Re(x) \leq 0$ .

Escolha alguma componente conexa  $C_0^*$  de  $C^*$ . Como um recobrimento de  $C_0^* \rightarrow D^*$  tem fibras finitas, ele é isomorfo a algum recobrimento  $x \in D^* \mapsto x^m \in D^*$ . Dito de forma diferente, há alguns *homeomorfismos*

$$\varphi : x \in D^* \mapsto (x^m, g(x)) \in C_0^*.$$

Mais uma vez, não perca [Analysis Situs](#), de Henri Paul de Saint Gervais, disponível online.

A palavra *étale* foi introduzida por geômetras algébricos franceses. Significa “estacionário”, sendo frequentemente usado para descrever a superfície do mar quando em repouso.



Uma aplicação étale que não é um recobrimento.

Uma aplicação é *própria* se a imagem inversa de um conjunto compacto for compacta. Claro que o inverso não é verdadeiro e uma aplicação de recobrimento pode não ser própria (como por exemplo  $t \in \mathbb{R} \mapsto \exp(it) \in S^1$ ).

Na verdade, Puiseux usava implicitamente recobrimentos, quando ele descrevia alguns laços seguidos por  $x$  em torno da origem, e a permutação associada dos valores de  $y$  satisfazendo  $F(x, y) = 0$ .

Esse  $\varphi$  é claramente holomorfo no disco perfurado e ainda temos que mostrar que ele se estende como uma função holomorfa no disco. Isto segue do teorema de extensão de Riemann: uma função holomorfa limitada em um disco perfurado é holomorfa no disco inteiro.

O teorema está provado, supondo que a derivada parcial  $\partial F/\partial y$  não se anula na curva perfurada  $F(x, y) = 0$ , que será discutida no próximo parágrafo.

Mais um anacronismo!

### *Raízes simples*

Uma função holomorfa de uma variável complexa  $y$  e sua derivada se anulam simultaneamente em algum  $y_0$  se, e somente se, esse zero for múltiplo. Portanto, temos que mostrar que no teorema de Puiseux, pode-se sempre supor que  $F$  tem a propriedade que, para  $x_0$  pequeno e diferente de zero, não existem raízes múltiplas de  $F(x_0, y) = 0$ .

Acontece que Puiseux não considerou funções holomorfas gerais  $F(x, y)$ , mas polinômios em  $x, y$ . Nesse caso, é fácil lidar com várias raízes e, na verdade, Puiseux descarta o problema em uma frase (exatamente na página 135 do seu artigo). Deixe-nos ser um pouco mais cuidadosos do que ele.

Considere o polinômio  $F$  como um elemento de  $\mathbb{C}[x][y]$ . Esse polinômio  $F$  pode ser visto como um polinômio em uma variável  $y$  com coeficientes em um anel fatorial. Podemos escrever  $F$  como um produto de fatores irredutíveis para que a curva  $F(x, y) = 0$  seja a união das curvas associadas a esses fatores irredutíveis. Portanto, podemos supor que  $F$  é irredutível.

Suponha agora que há uma sequência  $(x_k, y_k)$  convergindo para  $(0, 0)$  com  $x_k \neq 0$  tal que  $y_k$  é uma raiz múltipla de  $F(x_k, y) = 0$ . Então, o discriminante do polinômio  $F(x_k, y)$  é igual a 0. Portanto, o discriminante de  $F$ , como um elemento de  $\mathbb{C}[x]$ , anula-se de forma idêntica, uma vez que possui um número infinito de raízes. Se o discriminante de algum polinômio  $P$  se anula, o polinômio e sua derivada têm um fator comum. Isso é impossível se  $P$  for irredutível.

O discriminante de um polinômio é o resultante deste polinômio e do seu polinômio derivado.

Portanto, se  $F$  é irredutível em  $\mathbb{C}[x][y]$  e se  $x$  é pequeno e

diferente de zero, então,  $F(x, y) = 0$  não tem nenhuma raiz múltipla como um equação em  $y$ . Esse é o ingrediente que faltava para a prova do teorema de Puiseux, *para uma equação polinomial*  $F(x, y) = 0$ .

Para uma função holomorfa geral  $F(x, y) = 0$  (mais uma vez, Puiseux não considerou), ainda há algum trabalho a ser feito.

### O teorema de preparação de Weierstrass

Anteriormente já encontramos o teorema de preparação de Weierstrass, que provamos primeiro no contexto de séries formais antes de estabelecer a convergência. O objetivo agora é provar o mesmo teorema usando análise complexa.

Vamos lembrar o enunciado:

**Teorema.** *Seja  $F(x, y)$  uma função holomorfa diferente de zero definida em alguma vizinhança da origem em  $\mathbb{C}^2$ . Então, existem  $m$  funções holomorfas  $a_0(x), \dots, a_{m-1}(x)$  definidas em alguma vizinhança de  $0 \in \mathbb{C}$ , uma função holomorfa  $U(x, y)$  que não zera na origem e um inteiro  $r \geq 0$ , tal que*

$$F(x, y) = x^r U(x, y) (y^m + a_{m-1}(x)y^{m-1} + \dots + a_1(x)y + a_0(x)).$$

Esse teorema é exatamente o que é necessário. Ele afirma, a menos de funções que não zeram, podemos sempre assumir que a função  $F$  em estudo é um polinômio na variável  $y$ , com coeficientes no anel  $\mathbb{C}\{x\}$  de séries convergentes em  $x$ . A prova anterior de Puiseux (sempre é possível supor que  $\partial F/\partial y$  não se anula identicamente na curva, exceto na origem), pode, portanto, ser reproduzida palavra por palavra, substituindo o anel de polinômios em  $x$  pelo anel de séries convergentes. Portanto, o teorema de Puiseux é provado, usando o teorema de Weierstrass.

Agora apresento a prova analítica de Weierstrass.

Suponha, depois de dividir  $F$  por alguns  $x^r$ , que  $F(x, y)$  não se anula para  $|x| \leq 1$  e  $|y| = 1$ . Fixando  $x$  com  $|x| \leq 1$ , a função  $y \mapsto F(x, y)$  tem um número finito de zeros  $y_1(x), y_2(x), \dots, y_m(x)$  no disco da unitário, contando as multiplicidades. A principal dificuldade é a impossibilidade de escolher essas funções  $y_i(x)$  como funções holomorfas de  $x$  ou contínuas, precisamente pelo



Karl Weierstrass  
(1815–1897).

†



caráter multivalorado de  $y(x)$ , implícito em  $F(x, y) = 0$ . No entanto, mostraremos que todas as funções simétricas dos  $y_i(x)$  são, de fato, funções holomorfas de  $x$ .

A prova mais simples usa a fórmula de Cauchy. Vamos avaliar

$$s_k(x) = \frac{1}{2i\pi} \int_{|y|=1} \frac{y^k F'_y(x, y)}{F(x, y)} dy.$$

O resíduo de  $y^k F'_y(x, y)/F(x, y)$  em função de  $y$ , em uma das raízes  $y_i(x)$ , é a  $k$ -ésima potência  $y_i(x)^k$ , de modo que  $s_k(x)$  é a soma das  $k$ -ésimas potências das raízes. A integral mostra claramente que  $s_k(x)$  é uma função holomorfa de  $x$ .

Como os  $s_k$  geram as funções simétricas, todas as funções simétricas do  $y_i(x)$  são funções holomorfas de  $x$ , em particular as funções simétricas elementares  $a_i(x)$ . Pelo teorema de Viète, o polinômio

$$y^m - a_{m-1}(x)y^{m-1} + \dots + (-1)^{m-1}a_1(x)y + (-1)^m a_0(x)$$

anula-se exatamente nos mesmos pontos que  $F$  com as mesmas multiplicidades, de modo que o quociente  $U(x, y)$  não se anula.

O teorema da preparação de Weierstrass e o teorema de Puiseux estão provados.  $\square$

### Quem provou o teorema de preparação de Weierstrass?

Meu leitor já deve ter adivinhado: a resposta simples para essa pergunta não é certamente Weierstrass. Os historiadores da matemática sabem muito bem que perguntas como “quem provou isso primeiro?” são muito ingênuas e frequentemente não têm uma resposta muito precisa. No entanto, é interessante notar que dois importantes matemáticos do século XX, Henri Cartan <sup>112</sup> e Carl Siegel <sup>113</sup> escreveram artigos detalhados tentando desdobrar o desenvolvimento de ideias em torno desse teorema. Seus artigos, no entanto, não são completamente convergentes. Deixe-me mencionar apenas alguns passos.

– O fato de que as funções simétricas das raízes de alguma equação holomorfa  $F(x, y) = 0$ , sendo  $y$  a incógnita e  $x$  um

$s_0$  é o número de raízes. Sendo um inteiro e uma função holomorfa de  $x$  constante. Foi usado implicitamente algumas linhas acima!

Outro teorema bem conhecido de Newton.

<sup>112</sup> H. Cartan. Sur le théorème de préparation de Weierstrass. In *Festschr. Gedächtnisfeier K. Weierstrass*, pages 155–168. Westdeutscher Verlag, Cologne, 1966.

<sup>113</sup> C. L. Siegel. Zu den Beweisen des Vorbereitungssatzes von Weierstrass. In *Number Theory and Analysis (Papers in Honor of Edmund Landau)*, pages 297–306. Plenum, New York, 1969.

parâmetro, dependem holomorficamente em  $x$  já era conhecido por Cauchy em 1831, estando presente na prova que apresentei.

– Weierstrass publicou sua prova em 1886, mas menciona em uma nota de rodapé que estava dando palestras sobre esse teorema desde 1860. Não é surpreendente que ele evitou, tanto quanto possível, o uso de resíduos de Cauchy, mas não completamente, e trabalhou com séries. Sua prova é apenas parcialmente algébrica.

– O teorema é provado por Poincaré em sua tese, em 1879, sem menção a Cauchy. Como de costume, a palavra “prova” deve ser tomada com grande cuidado na escrita de Poincaré, sendo especialmente verdadeiro neste artigo. Muito mais tarde, por exemplo, em seus *Méthodes Nouvelles*, ele se referiu à sua tese, sem fornecer uma prova melhor e sem mencionar Weierstrass. Curiosamente, Henri Cartan, um dos pais fundadores de Bourbaki, não menciona Poincaré em seu artigo.

– Em 1905, Lasker <sup>114</sup> forneceu uma prova totalmente algébrica e deduziu as consequências algébricas para os anéis das séries formais e convergentes.

– Siegel também enfatiza que, segundo ele, a prova mais curta é devida à Stickelberger <sup>115</sup> em 1887.

Para uma apresentação moderna e elementar do teorema, veja o livro de Ebeling <sup>116</sup>. Para uma descrição cuidadosa das muitas variantes do teorema e dos comentários históricos adicionais, consulte Grauert e Remmert <sup>117</sup>.

Não posso terminar este capítulo sem mencionar que há uma versão desse teorema para funções  $C^\infty$ , conjecturada por Thom e provada por Malgrange <sup>118</sup>. Mas essa é uma outra história <sup>119</sup>...

<sup>114</sup> E. Lasker. *Zur Theorie der Moduln und Ideale*. *Math. Ann.*, 60(1):20–116, 1905.

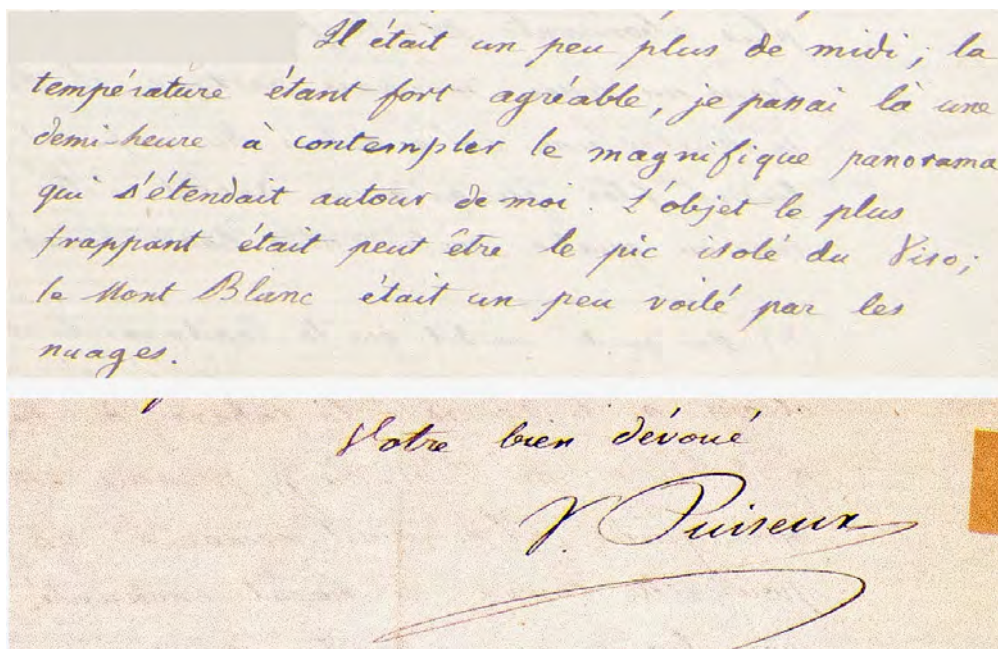
<sup>115</sup> L. Stickelberger. *Ueber einen Satz des Herrn Noether*. *Math. Ann.*, 30(3):401–409, 1887.

<sup>116</sup> W. Ebeling. *Functions of several complex variables and their singularities*. GSM083. AMS, 2007.

<sup>117</sup> R. R. H. Grauert. *Analytische Stellenalgebren*. Die Grundlehren der mathematischen Wissenschaften 176. Springer-Verlag Berlin Heidelberg, 1971.

<sup>118</sup> B. Malgrange. *Ideals of differentiable functions*. Tata Institute of Fundamental Research Studies in Mathematics, No. 3. Tata Institute of Fundamental Research, Bombay; Oxford University Press, London, 1967.

<sup>119</sup> V. I. Arnold, S. M. Gusein-Zade, and A. N. Varchenko. *Singularities of differentiable maps. Classification of critical points, caustics and wave fronts*. Modern Birkhäuser Classics. Birkhäuser/Springer, New York, 2012. Reimpressão da edição de 1985.



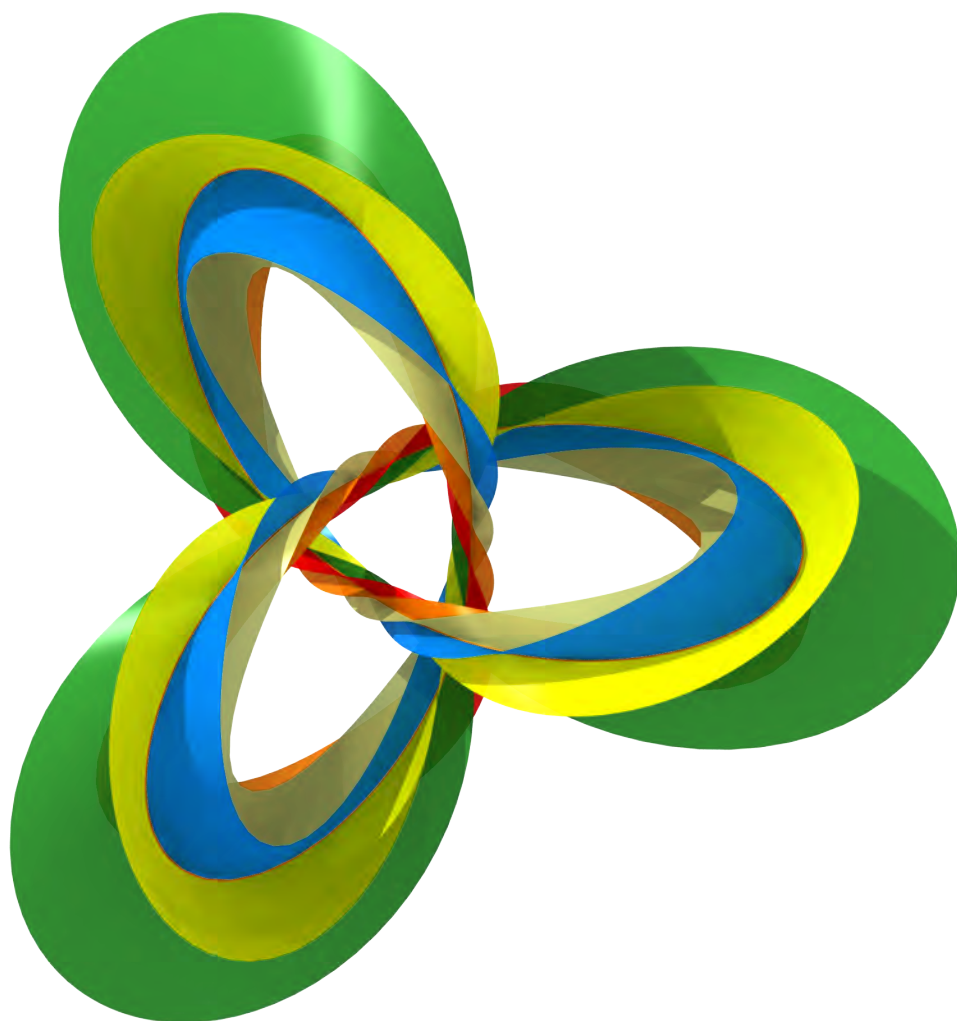
De uma carta de Victor Puiseux descrevendo sua expedição ao Monte Pelvoux: “Contemplando o magnífico panorama ao meu redor”. Victor Puiseux poderia ter sido um modelo para Caspar David Friedrich quando pintou o *Caminhante sobre o mar de neblina*, mas a expedição Pelvoux aconteceu 30 anos depois! +

“Il était un peu plus de midi, la température étant fort agréable, je passai là une demi-heure à contempler le magnifique panorama qui s’étendait autour de moi. L’objet le plus frappant était peut-être le pic isolé du Viso; le Mont Blanc était un peu voilé par les nuages.”



Victor Puiseux, na época da expedição ao Monte Pelvoux.

†



As fibras de Milnor de  
 $x^3 - y^2$ .

†

## Jack Milnor e a sua fibração

QUANDO EU ENTRO NUMA BIBLIOTECA DE MATEMÁTICA, ou quando eu navego através dos *Mathematical Reviews*, ou simplesmente no *Google*, sou frequentemente dominado pela vastidão do mundo matemático. Mesmo tópicos que podem parecer microscópicos para o leigo, como, por exemplo, a topologia de curvas algébricas, são realmente territórios enormes cuja exploração poderia facilmente exigir várias vidas. Esse sentimento pode ser deprimente ou intoxicante, dependendo do meu humor ☺. Neste pequeno livro, o melhor que posso fazer é descrever um exemplo significativo, mencionar alguns dos principais resultados e referir alguns dos livros (longos) que propõem uma discussão completa do estado da arte.

Em qualquer caso, um único livro deve ser enfatizado como uma joia e *tem que ser lido por todos os alunos* interessados por este tópico: *Pontos singulares de hipersuperfícies complexas*<sup>120</sup> por Milnor : um grande mestre na arte de escrever matemática.

### Um exemplo

Considere a curva

$$F(x, y) = x^9 - x^{10} + 6x^8y - 3x^6y^2 + 2x^5y^3 + 3x^3y^4 - y^6 = 0.$$

Esse  $F$  não foi escolhido aleatoriamente. Qualquer equação  $F = 0$  pode ser resolvida usando séries de Puiseux. Nesse exemplo, eu trapaceei e comecei da solução

$$y = x^{3/2} + x^{5/3}.$$



John Milnor.



<sup>120</sup> J. Milnor. *Singular points of complex hypersurfaces*. Annals of Mathematics Studies, No. 61. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1968.

E eu olhei para a equação! Nesse caso,

$$(y - x^{3/2})^3 = x^5.$$

Expandindo e elevando a uma potência adequada para eliminar os expoentes racionais, encontramos realmente  $F(x, y) = 0$ . Na verdade, definindo  $x = x_1^2$  e  $y = x_1^3(1 + y_1)$ , como se deve fazer usando o algoritmo de Newton, o resultado fatoriza conforme esperado:

$$F(x_1^2, x_1^3(1 + y_1)) = -x_1^{18} (x_1 - y_1^3) (-8 + x_1 - 12y_1 - 6y_1^2 - y_1^3).$$

Segue-se que o conjunto dos zeros de  $F$  na vizinhança da origem contém exatamente um ramo  $x_1 = y_1^3$  ou

$$x = t^6 \quad y = t^9 + t^{10}$$

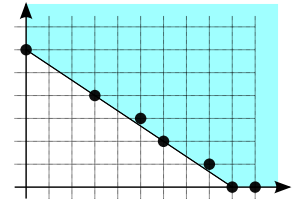
o que revela que  $y = x^{3/2} + x^{5/3}$  como previsto. Para um  $F$  qualquer, mesmo para um polinômio, devemos esperar uma série de Puiseux *infinita*, mas vamos primeiro olhar para esse exemplo específico.

Examinemos o *link da singularidade*, interseção da curva  $F = 0$  com uma pequena esfera  $S_\varepsilon^3$ . A transversalidade da interseção da curva com pequenas esferas é fácil de ver. Na verdade, o quadrado da norma

$$\varphi : t \in \mathbb{C} \mapsto |t^6|^2 + |t^9 + t^{10}|^2 \in \mathbb{R}_+$$

é equivalente a  $|t|^{12}$  para  $t$  pequeno, e a equação  $\varphi(t) = \varepsilon^2$  define um laço em  $\mathbb{C}$ , perto de  $|t| = \varepsilon^{1/6}$ , transversal às linhas radiais. Em outras palavras, a interseção de  $F = 0$  com uma pequena esfera  $S_\varepsilon^3$  é um círculo mergulhado, i.e., um *nó*, a imagem por  $\varphi$  desse laço.

A menos do homeomorfismos da esfera, esse nó é independente de  $\varepsilon$ . Poderíamos até usar elipsoides em vez de esferas ou mesmo nossa esfera quadrada  $\max(|x|, |y|) = \varepsilon$ . A prova detalhada é técnica e chata, mas a ideia-chave é bastante simples. Dadas duas normas euclidianas  $N_0, N_1$  em  $\mathbb{R}^4$ , podemos intersectar a curva  $F(x, y) = 0$  com esferas pequenas  $N_0 = \varepsilon$  e  $N_1 = \varepsilon$ , resultando em dois nós de duas variedades homeomorfas a uma esfera. Nós agora construímos um caminho de normas



O polígono Newton de  $F$ .

Um *nó* é um mergulho do círculo na 3-esfera. Um *link* é uma união disjunta de uma quantidade finita de muitos nós. Dois nós ou links são considerados equivalentes se existe um homeomorfismo, que preserve a orientação da esfera enviando o primeiro ao segundo.

A não prova a seguir oculta um fato fundamental da topologia diferencial: Se tivermos uma família de mergulhos  $i_\lambda : X \rightarrow Y$  (for  $0 \leq \lambda \leq 1$ ) de uma variedade compacta  $X$  em alguma outra variedade  $Y$ , então, existe uma *isotopia*, isto é, uma família de difeomorfismos  $\Phi_\lambda$ , de  $Y$  tal que  $i_\lambda = \Phi_\lambda \circ i_0$ .

$\lambda N_1 + (1 - \lambda)N_0$  (for  $0 \leq \lambda \leq 1$ ) de forma que realmente temos uma família contínua de círculos mergulhados em esferas que, portanto, definem “o mesmo nó”. Um argumento semelhante poderia ser usado para a “esfera quadrada”. Para uma ilustração de possíveis erros que um principiante ingênuo poderia cometer, veja este artigo<sup>121</sup>.

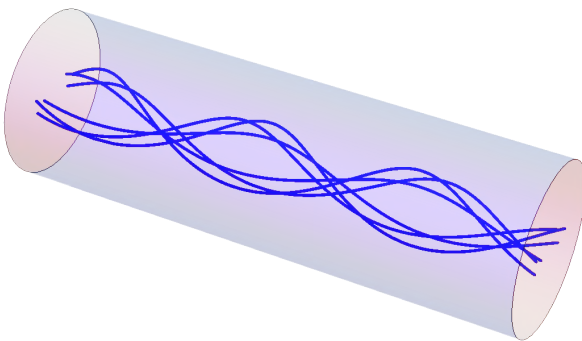
Denote esse nó por  $K_F$ . Vamos usar uma esfera quadrada conveniente:  $\max(|x|, |y|) = \varepsilon$ . A interseção com a curva está localizada no toro sólido  $|x| = \varepsilon$  e  $|y| \leq \varepsilon$  tal que  $|t| = \varepsilon^{1/6}$ . Vamos trocar de escala, pondo  $X = x/\varepsilon$  e  $Y = y/\varepsilon$ . Em particular,  $X$  está no círculo unitário e  $Y$  no disco unitário. Se  $t = \varepsilon^{1/6}\tau$ , segue que

$$X = \tau^6 \quad Y = \varepsilon^{1/2}\tau^9 + \varepsilon^{2/3}\tau^{10}$$

com  $\tau$  descrevendo o círculo unitário.

*Em todas as figuras seguintes, o toro sólido  $S^1 \times D^2$  é desenhado com o cilindro  $[0, 2\pi] \times D^2$  e as duas faces  $\{0\} \times D^2$  e  $\{2\pi\} \times D^2$  deveriam ser coladas.*

Para todo  $X$  no círculo unitário, existem exatamente seis valores de  $Y$ , diferindo pela multiplicação de  $\tau$  por algumas raízes sextas da unidade. Dizemos que o nó  $K_F$  está em *formato de trança*, pois ele intersecta transversalmente todos os discos  $\{*\} \times D^2$ , em torno do círculo, sendo esses seis pontos permutados em um caminho que não está descrito.



Observe que  $\varepsilon^{2/3}$  é pequeno comparado à  $\varepsilon^{1/2}$  para todo  $\varepsilon$

<sup>121</sup> F. Deloup. [The fundamental group of the circle is trivial.](#) *Amer. Math. Monthly*, 112(5):417–425, 2005.



Não tenha pressa! Olhe para as imagens neste capítulo com muito cuidado mesmo que não seja fácil.

+

+



pequeno. Observe também que  $\varepsilon^{1/2}\tau^9$  recebe apenas dois valores quando um deles multiplica  $\tau$  por um raiz sexta da unidade. O nó associado para

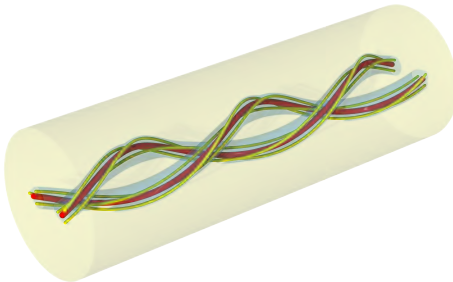
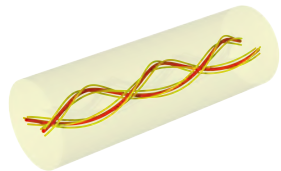
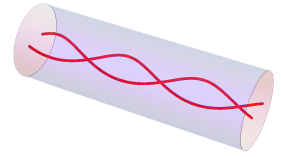
$$X = \tau^6 \quad Y_0 = \varepsilon^{1/2}\tau^9$$

é simplesmente nosso velho amigo trevo  $x^3 = y^2$ . Quando  $X$  varia em torno do círculo, os correspondentes dois pontos em  $\{X\} \times D^2$  giram por três meia voltas, produzindo o nó de trevo.

Considere  $Y = \varepsilon^{1/2}\tau^9 + \varepsilon^{2/3}\tau^{10} = Y_0 + Y_1$  como uma pequena perturbação de  $Y_0$ . Note que

$$X = \tau^6 \quad Y_1 = \varepsilon^{2/3}\tau^{10}$$

dados três (muito pequenos) valores de  $Y_1$  para cada valor de  $X$  no círculo unitário. Então, os seis pontos  $Y$ , em cada disco  $\{*\} \times D^2$ , vêm em dois grupos de três pontos. Em outras palavras, o nó  $K_F$  vive numa estreita vizinhança tubular do nó de trevo e intersecta pequenos discos transversais ao nó de trevo em três pontos.



Essa vizinhança tubular do nó de trevo pode ser parametrizada da seguinte maneira:

$$(\mu, \zeta) \in S^1 \times D^2 \mapsto (X = \mu^2, Y = \varepsilon^{1/2}\mu^3 - 2\varepsilon^{2/3}\mu^{-3}\zeta).$$

O círculo  $S^1 \times \{0\}$ , espinha dorsal desse toro sólido, é levado ao nó de trevo. Poderíamos perguntar porque escolher  $\varepsilon^{1/2}\mu^3 - 2\varepsilon^{2/3}\mu^{-3}\zeta$  e não simplesmente  $\varepsilon^{1/2}\mu^3 + 2\varepsilon^{2/3}\zeta$ , que também poderia ser uma parametrização. O ponto é que, com essa escolha de

O coeficiente 2 diante do  $\varepsilon^{2/3}\mu^{-3}$  não é importante: o seu único papel é dar um diâmetro suficiente para o tubo que contém o nosso nó.

coordenadas, nessa vizinhança tubular,  $x^3 - y^2$  é igual a

$$\varepsilon^3(\mu^2)^3 - \varepsilon^2(\varepsilon^{1/2}\mu^3 - 2\varepsilon^{2/3}\mu^{-3}\zeta)^2$$

que é da ordem de  $4\varepsilon^{19/6}\zeta$  para um pequeno  $\zeta$ , então, o argumento de  $x^3 - y^2$  é próximo ao argumento de  $\zeta$ . Entretanto, as fibras de Milnor de  $x^3 - y^2 = 0$  numa vizinhança tubular estão próximas às páginas  $\arg \zeta = \text{constante}$ .

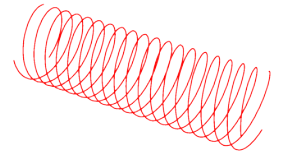
Nessas coordenadas, podemos comparar as expressões  $\varepsilon^{1/2}\mu^3 - 2\varepsilon^{2/3}\mu^{-3}$  e  $\varepsilon^{1/2}\tau^9 + \varepsilon^{2/3}\tau^{10}$  para  $Y$ . Usando  $\mu = \tau^3$ , nosso nó  $K_F$  é a imagem de

$$\tau \in S^1 \mapsto (\mu, \zeta) = (\tau^3, -\frac{1}{2}\tau^{19}) \in S^1 \times D^2.$$

Isto é, um nó toral  $(19, 3)$ .

Esse 19 pode parecer estranho. Note que existe um homeomorfismo do toro sólido  $S^1 \times D^2$  que leva o nó toral  $(19, 3)$  no nó toral  $(19 - 3k, 3)$  para qualquer  $k$  (por exemplo, o nó toral  $(1, 3)$ , que é muito mais simples), mas tal homeomorfismo não pode ser tanto homotópico à identidade na fronteira quanto estendido a toda a esfera.

Segue que  $K_F$  é obtido intersectando um nó toral  $(19, 3)$  em uma vizinhança de um nó toral  $(3, 2)$ . Este é um exemplo típico de um nó toral *iterado*. Às vezes, eles são chamados de *nó cabo* porque se assemelham à construção de cabos feitos de fios torcidos trançados em conjunto, ou *satélites* girando em torno de um planeta que, por sua vez, gira em torno do sol.



De fato, a aplicação  $(\tau, \zeta) \mapsto (\tau, \tau^k \zeta)$  é um homeomorfismo que torce o toro sólido.

### Fibração de Milnor

No caso do nó de trevo, as superfícies de nível de  $\arg(x^3 - y^2)$  preenchem o complemento do nó como as páginas de um livro cuja lombada é o nó de trevo. Descrevemos a topologia dessas páginas como um toro furado.

Em 1968, no seu livro seminal, Milnor mostrou que esse é um fato geral. Ele realmente provou um teorema em todas as dimensões, mas nos limitaremos ao caso de dimensão 2.

**Teorema.** *Seja  $F(x, y)$  um função holomorfa reduzida definida numa vizinhança da origem de  $\mathbb{C}^2$  tal que  $F(0, 0) = 0$ .*

Uma série  $F(x, y)$  é *reduzida* se ela não possui fatores múltiplos na sua decomposição em fatores irredutíveis.



- A curva  $F(x, y) = 0$  intersecta transversalmente pequenas esferas  $S_\epsilon^3$  ao longo de algum link  $L_\epsilon \subset S_\epsilon^3$  cuja topologia é independente de  $\epsilon$ .
- Além disso, a aplicação

$$(x, y) \in S_\epsilon^3 \setminus L_\epsilon \mapsto \arg(F(x, y)) = \frac{F(x, y)}{|F(x, y)|} \in S^1$$

é uma fibração localmente trivial. Os fechos das fibras são superfícies compactas cujas bordas coincidem com  $L_\epsilon$ . Numa vizinhança tubular desse link, eles são parecidos com um livro aberto: as fibras são localmente produtos de um segmento radial em  $D^2$  e um segmento.

Esse teorema é a ferramenta fundamental para o estudo local de singularidades. Entretanto, preciso confessar: por um longo tempo, não tinha olhado a prova e estava convencido de que ela teria que ser elementar e direta. Temos uma aplicação muito natural para o círculo; por que não seria uma fibração? Descobri que eu estava errado, e a prova é bastante sutil. Surpreendentemente, os livros que lidam com essa questão são de dois tipos: O primeiro, ao chegar ao ponto-chave da prova, de uma forma muito discreta, escrevem “Ver Milnor, capítulo 2”. O segundo, também no mesmo ponto, apenas copiam praticamente palavra por palavra do conteúdo apresentado em “Milnor, capítulo 2”. Na verdade, ambos estão certos, pois é difícil escrever melhor

do que “Milnor, capítulo 2”. Assim, meu objetivo aqui não é inovar, mas dar alguma intuição sobre esse teorema.

Inicialmente, observe que o teorema é verdadeiro e elementar em dimensão 1. Seja  $f(x)$  alguma função holomorfa não nula definida numa vizinhança da origem de  $\mathbb{C}$  e tal que  $f(0) = 0$ . Escreva  $f(x) = ax^n + \dots$  with  $a \neq 0$  e veja a aplicação

$$x \in S_\varepsilon^1 \mapsto \arg(f(x)) = \frac{f(x)}{|f(x)|} \in S^1.$$

Essa é obviamente um recobrimento para  $\varepsilon$  suficientemente pequeno. De fato, isso é próximo na topologia  $C^1$  da aplicação de recobrimento  $x \mapsto \arg(a) \arg(x)^n$ .

Após esse caso trivial, vamos considerar a curva  $F(x, y) = 0$  : Assuma que  $F$  é reduzida e, então,  $\partial F/\partial y$  não zera em  $F = 0$  (exceto na origem). Em vez de usar a esfera redonda, usaremos a nossa esfera quadrada  $\max(|x|, |y|) = \varepsilon$ . Por simplicidade, assuma que a curva intersecta apenas o toro sólido  $T_1$  definido por  $|x| = \varepsilon, |y| < \varepsilon$ .

Nosso objetivo é mostrar que o argumento de  $F(x, y)$ , restrito a esse toro sólido, é uma submersão (fora de  $F = 0$ ). Em outras palavras, dado um ponto  $(x, y)$  em  $T_1$ , nós olhamos para alguma direção tangente em  $T_1$  ao longo da qual a derivada de  $\arg F$  é não nula. Como uma primeira tentativa, podemos olhar para alguma direção vertical, fixando  $x$ . Então,  $F(x, y)$  muda de acordo com a derivada parcial  $\partial F/\partial y$  de modo que  $\arg F$  seja de fato uma submersão, *pelo menos fora do conjunto dos zeros de  $\partial F/\partial y$* .

Para esse conjunto dos zeros de  $\partial F/\partial y(x, y) = 0$ , agora, é alguma outra curva, apenas intersectando  $F = 0$  na origem, podendo ser parametrizada à la Puiseux por  $x = t^n$  e  $y = f(t)$ . Usando o nosso caso (trivial) de dimensão 1,

$$t \in S_{\varepsilon^{1/n}}^1 \mapsto \arg F(t^n, f(t)) \in S^1$$

é uma aplicação de recobrimento. Portanto, nos pontos onde  $\partial F/\partial y(x, y) = 0$  pode ser encontrada outra direção na qual a derivada do argumento não se anula.

Mas esse argumento não é uma prova completa do teorema de Milnor por várias razões.

O argumento de um número complexo não nulo  $z$  pode ser definido de várias maneiras. Ele poderia ser um elemento do intervalo  $[0, 2\pi[$ , ou  $\mathbb{R}/2\pi\mathbb{Z}$ , ou  $z/|z|$  no círculo unitário. Para o que segue, sempre escolherei a maneira mais conveniente: acredito que isso não causará dificuldades.

A primeira: não podemos usar uma esfera quadrada em vez de uma esfera redonda. Embora isso não seja tão sério assim e o argumento possa facilmente ser adaptado para uma esfera redonda.

A segunda: uma submersão não precisa ser uma fibração sem assumirmos a compacidade das fibras. Nós precisamos estudar a estrutura local da nossa submersão próximo do link  $L_\varepsilon \subset S_\varepsilon^3$ . E isso não representa nenhuma dificuldade. O ponto-chave consiste em se  $F(x, y) = 0$  e se consideramos alguma reta complexa em  $\mathbb{C}^2$ , passando por  $(x, y)$  e transversal à curva, então, podemos aplicar o caso (trivial) dimensão 1 para analisar o argumento  $\arg F$ , numa vizinhança de  $(x, y)$ , para obter uma imagem local em torno do link.

Note que nossa apresentação é simples e limitada para o caso de dimensão 2, diferente do teorema de Milnor: verdadeiro para qualquer dimensão.

Para uma excelente apresentação, sugiro Milnor, capítulo 2 ☺.

### *Fibração de Milnor no nosso exemplo*

Vamos voltar ao nosso exemplo:

$$F(x, y) = x^9 - x^{10} + 6x^8y - 3x^6y^2 + 2x^5y^3 + 3x^3y^4 - y^6 = 0.$$

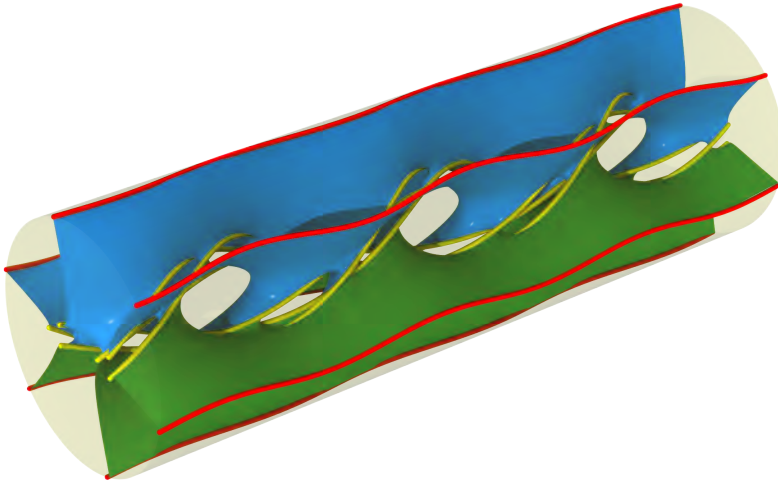
A curva  $F(x, y) = 0$  intersecta uma pequena esfera ao longo de um nó, um satélite do nó de trevo. Agora desejemos descrever a topologia das fibras de Milnor  $\arg F(x, y) = \text{const}$ . Se eu pedir ao meu computador para desenhar uma dessas fibras, o resultado é a figura abaixo.

Mas isso é complicado e requer uma análise cuidadosa. Note que pelo menos uma dessas superfícies intersecta a fronteira de um toro sólido ao longo de seis curvas (em vermelho). O nó, representado em amarelo, é também uma componente da fronteira. As cores azul e verde das duas faces destacadas mostram que essa superfície é orientável. Mantenha em mente que existem sete componentes limitadas.

A fim de entender essa figura, vamos inicialmente olhar para as fibras de Milnor de uma  $p, q$  curva  $x^p - y^q = 0$ , sendo  $p$

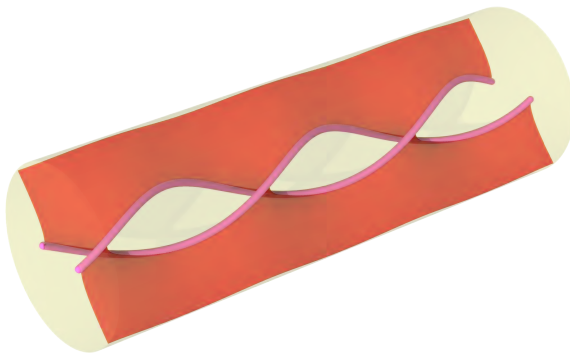


O final deste capítulo requer uma grande atenção, mesmo que não seja necessário para o restante do livro. †



+

e  $q$  relativamente primos (com  $p > q$ ). Sabemos que elas são superfícies de gênero  $(p - 1)(q - 1)/2$  com um disco removido. Vamos olhar para as suas posições relativas na nossa esfera quadrada  $\max(|x|, |y|) = \varepsilon$ . Desde que  $p > q$ , a interseção com  $F = 0$  vive no toro sólido  $T_1$  definido por  $|x| = \varepsilon$ . Na fronteira  $|x| = \varepsilon$  e  $|y| = \varepsilon$ , o valor de  $x^p - y^q$  é muito próximo de  $y^q$  e o argumento de  $x^p - y^q$  é quase igual a  $q$  vezes o argumento de  $y$ . Segue que a fibra de Milnor de  $x^p - y^q$  intersecta a fronteira de  $T_1$  ao longo de  $q$  curvas próximas de  $q$  paralelas.

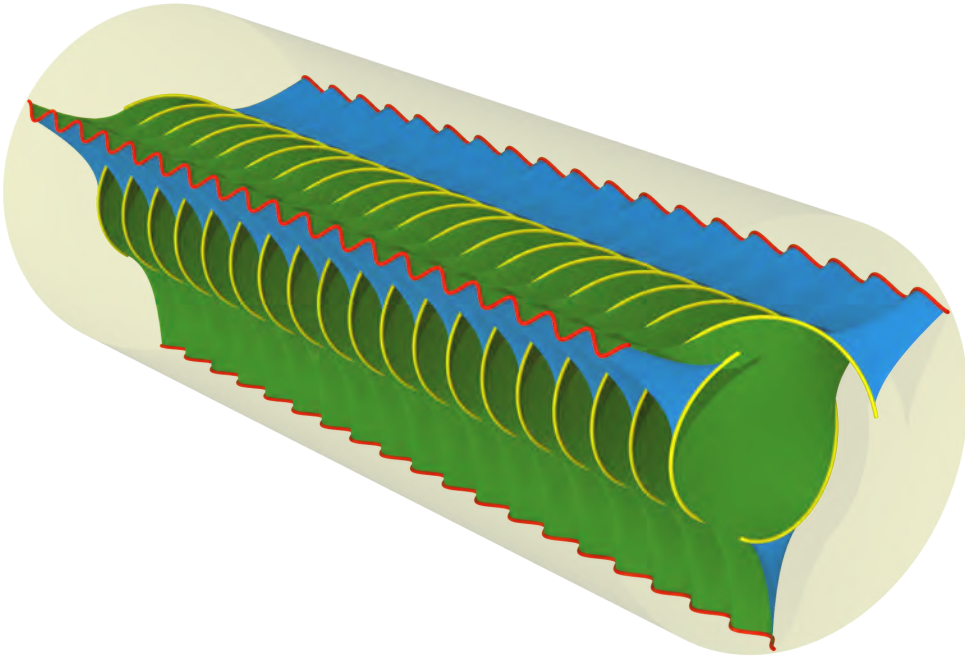


Uma fibra de Milnor do nó de trevo vista como um nó toral  $(3, 2)$ . A fronteira da superfície é um nó e sua interseção com a fronteira de um cilindro consiste de duas paralelas.

+

Pela mesma razão, em um outro toro sólido  $T_2$  definido

por  $|y| = \varepsilon$ , um fibra de Milnor quase coincide com  $q$  discos, com o argumento de  $y$  tomando  $q$  valores e  $x$  descrevendo o disco de raio  $\varepsilon$ . Em outras palavras, a interseção de alguma fibra de Milnor de  $x^p - y^q$  com  $T_1$  é uma superfície de gênero  $(p-1)(q-1)/2$ , onde  $1+q$  discos foram removidos. A fronteira do primeiro disco removido é um nó toral, posicionado no interior de  $T_1$  e os outros  $q$  discos têm como fronteiras  $q$  círculos fronteiros de  $T_1$ .



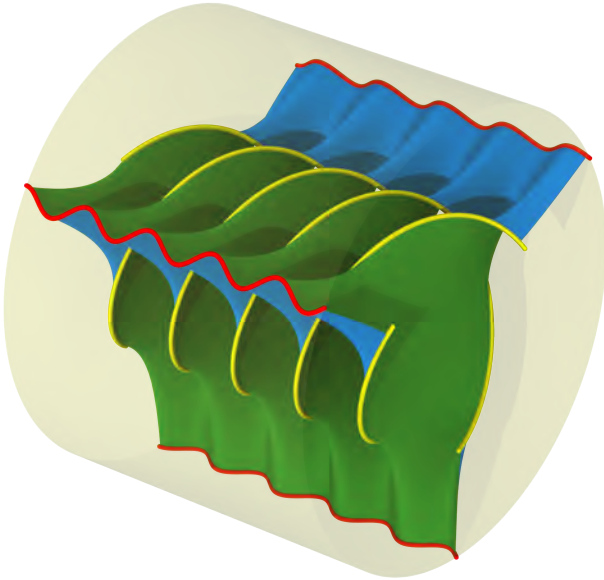
Vamos voltar ao nosso *exemplo mais complicado definido por*  
 $F(x, y) = 0$ .

Extraindo os termos dominantes do polígono de Newton, obtemos:

$$F(x, y) = (x^3 - y^2)^3 - x^{10} + 6x^8y + 2x^5y^3.$$

Relembre que construímos uma vizinhança tubular  $\mathcal{T}_\varepsilon$  do nó de trevo, parametrizada por  $(\mu, \zeta) \in \mathbb{S}^1 \times D^2$ , na qual  $x^3 - y^2$  é de ordem  $4\varepsilon^{19/6}\zeta$ . Na fronteira desse toro sólido, onde  $|\zeta| = 1$ , nós

Uma fibra de Milnor de um nó toral de tipo  $(19, 3)$ . A fronteira da superfície é o nó (em amarelo) e sua interseção com a fronteira do cilindro consiste de três paralelas (em vermelho). +



Aqui está uma pequena fatia para entender melhor a figura anterior. As cores azul e verde mostram que tal superfície é, de fato, orientável. †

temos  $|(x^3 - y^2)^3| \simeq 64\varepsilon^{57/6}$  e  $|x|, |y| \leq \varepsilon$ , em que  $F(x, y)/(x^3 - y^2)^3$  é muito próximo de 1, e o argumento de  $F(x, y)$  é próximo de  $\arg(\zeta)^3$ . Em particular, cada fibra de Milnor de  $F$  na fronteira de  $\mathcal{T}_\varepsilon$  é muito próxima de três paralelas tais que  $\arg \zeta = \text{constant}$ .

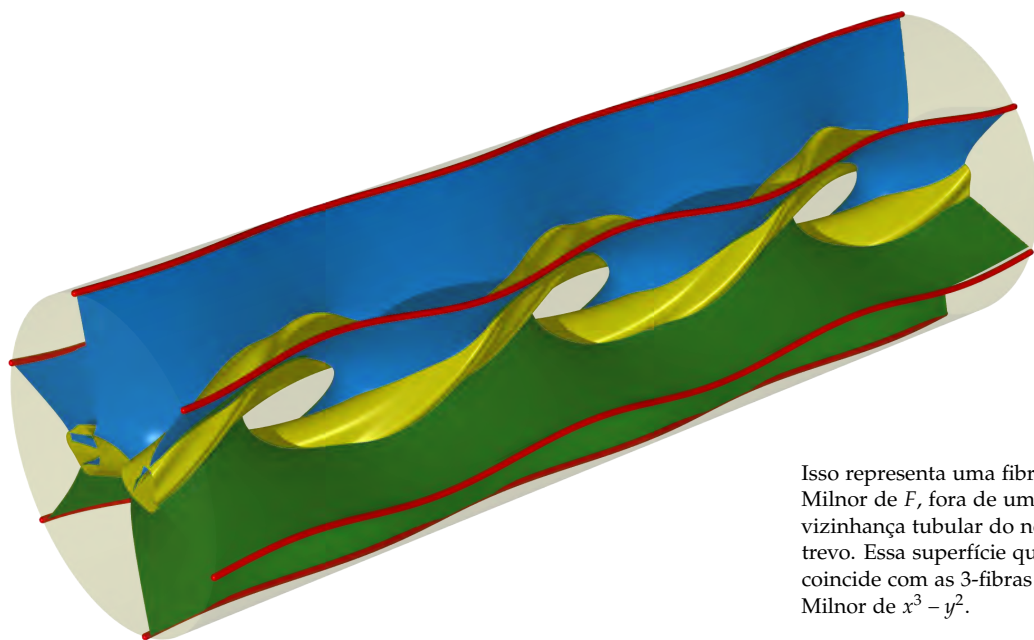
Isso também ocorre fora de  $\mathcal{T}_\varepsilon$ : cada fibra de Milnor de  $F$  fora de  $\mathcal{T}_\varepsilon$  é muito próxima das três fibras de Milnor de  $x^3 - y^2 = 0$ . Não esqueça que as fibras de Milnor de  $x^3 - y^2 = 0$  são toros perfurados. Suas interseções com  $T_2$  consistem de dois discos e sua interseção com  $T_1$  tem três componentes fronteiras: duas na fronteira de  $T_1$  e a terceira corresponde ao próprio nó. Isso é, de fato, o que pode ser visto escondido no interior de  $\mathcal{T}_\varepsilon$ .

O interior de  $\mathcal{T}_\varepsilon$  é o domínio do nó  $(19, 3)$  que é inserido no tubo. Vamos avaliar  $F(x, y) = F(\varepsilon X, \varepsilon Y)$  no interior do tubo, em coordenadas  $(\mu, \zeta)$ :

$$F(\varepsilon\mu^2, \varepsilon^{3/2}\mu^3 - 2\varepsilon^{5/3}\mu^{-3}\zeta).$$

Isso se anula exatamente em  $32\zeta^3 = -3\mu^{19}$ , para a primeira ordem. Portanto, uma fibra de Milnor de  $F$ , interior a  $\mathcal{T}_\varepsilon$ , é próxima a uma fibra de Milnor de  $y^3 = x^{19}$ . É uma superfície





Isso representa uma fibra de Milnor de  $F$ , fora de uma vizinhança tubular do nó de trevo. Essa superfície quase coincide com as 3-fibras de Milnor de  $x^3 - y^2$ . †

de gênero  $(3-1)(19-1)/2 = 18$  com 4 discos removidos. Uma das fronteiras é o nosso nó, com deveria ser, e as três outras são paralelas à fronteira de  $\mathcal{T}_\varepsilon$ .

*Em resumo, uma fibra de Milnor de  $F(x, y) = 0$  é homeomorfa a uma superfície fechada e orientável de gênero 18, com três somas conexas com um toro, onde finalmente eliminamos um disco. É uma superfície de gênero 21. Bastante complicada.*

O resto desse capítulo será bastante vago: não posso dar mais do que um vislumbre da teoria.

Dividindo o complemento de  $K_F$  ao longo de uma fibra de Milnor  $\Sigma$ , obtemos um produto  $\Sigma \times [0, 1]$ . Para reconstruir o complemento do nó, deve-se colar  $\Sigma \times \{0\}$  a  $\Sigma \times \{1\}$  usando algum difeomorfismo de  $\Sigma$ . Este difeomorfismo, bem definido a menos de isotopia, é chamado de *monodromia* do nó. A ação da monodromia na primeira homologia possui um polinômio característico denominado de *polinômio de Alexander* do nó. No nosso exemplo, tudo pode ser descrito de forma bastante concreta.

Na verdade, a fibração de Milnor no complemento de *uma* fibra é uma fibração em  $[0, 1]$ , daí uma fibração trivial, uma vez que  $[0, 1]$  é contrátil.

Deixe-me dar apenas um resultado. Nossa superfície  $\Sigma$  contém três curvas fechadas  $\gamma_i$  ( $i = 1, 2, 3$ ) ao longo das quais nós realizamos uma soma conexa. Cortando ao longo das  $\gamma_i$ , obtemos quatro componentes:  $S, \Sigma_i$ , sendo  $S$  uma superfície de gênero 18 menos 4 discos e cada  $\Sigma_i$  é um toro perfurado. Alguma aplicação de monodromia  $\psi$  preserva as curvas  $\gamma_i$  sendo um *Dehn twist* em algum anel em torno dessas curvas. Isso significa que  $\psi$ , numa vizinhança dessas curvas, é parecido com a figura na margem, mostrando que a ação da monodromia na homologia é periódica. Entretanto, poderia não ser verdadeiro na homotopia: as curvas  $\gamma_i$  são homólogas a zero, mas não são homotópicas a zero.

Se cortarmos o aberto  $\Sigma$  ao longo de  $\gamma_i$ , nós determinaremos, três vezes, as monodromias de  $x^3 - y^2$  e, uma vez, de  $x^{19} - y^3$ . Segue que o polinômio de Alexander é o produto do cubo do polinômio para  $x^3 - y^2$  e do polinômio para  $x^{19} - y^3$ . Portanto, temos:

$$\frac{(X^6 - 1)^3(X - 1)^3(X^{57} - 1)(X - 1)}{(X^2 - 1)^3(X^3 - 1)^3(X^{19} - 1)(X^3 - 1)}$$

que é igual a

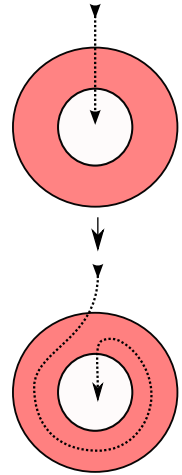
$$(1 - X + X^2)^3(1 - X + X^3 - X^4 + X^6 - X^7 + X^9 - X^{10} + X^{12} - X^{13} + X^{15} - X^{16} + X^{18} - X^{20} + X^{21} - X^{23} + X^{24} - X^{26} + X^{27} - X^{29} + X^{30} - X^{32} + X^{33} - X^{35} + X^{36}).$$

### O caso geral

Deixe-me apenas mencionar alguns resultados mais expressivos:

Os nós associados a um ramo de uma curva são sempre nós do tipo toro, iterados.

Os nós associados a duas curvas *irredutíveis*  $F_1(x, y) = 0$  e  $F_2(x, y) = 0$  são topologicamente equivalentes por meio de um homeomorfismo da 3-esfera se, e somente se, os dois ramos associados tenham o mesmo invariante característico de Puiseux. Na verdade, distingue-se esses nós usando os polinômios de Alexander. Isso foi provado há muito tempo para o caso de nós e também para o caso de curvas com dois ramos. O fato análogo, para as curvas não irredutíveis, produzindo *links* consistindo de vários nós disjuntos, foi estabelecido muito mais recentemente.



Um Dehn twist: esse homeomorfismo, como mostrado, é a identidade na fronteira do anel, torcendo os círculos concêntricos, mas os preservando.

As *monodromias* associadas às curvas gerais têm sido descritas lindamente por A'Campo <sup>122</sup>.

Eu poderia dizer que a situação está agora muito melhor entendida.

Neste ponto, é aconselhável parar nossa excursão aqui se quisermos continuar nosso passeio: há outros lugares para visitar. No entanto, eu entendo perfeitamente alguma frustração do leitor obrigado a retroceder e não continuar nesse caminho que parece ser (e realmente é) bastante bonito. Para os leitores frustrados e os demais, se quiserem saber mais sobre o tema, com uma perspectiva histórica, minha sugestão é olhar na maravilhosa pesquisa de Weber <sup>123</sup> e no já mencionado <sup>124</sup>.

<sup>122</sup> N. A'Campo. [Sur la monodromie des singularités isolées d'hypersurfaces complexes](#). *Invent. Math.*, 20:147–169, 1973.

<sup>123</sup> C. Weber. On the topology of singularities. In *Singularities II*, volume 475 of *Contemp. Math.*, pages 217–251. Amer. Math. Soc., Providence, RI, 2008.

<sup>124</sup> E. Brieskorn and H. Knörrer. *Plane algebraic curves*. Birkhäuser/Springer Basel AG, Basel, 1986.

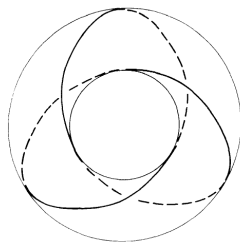


Figure 1. The torus knot of type (2, 3).

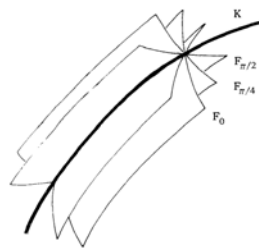


Figure 3-A. The curve  $y = \pm x\sqrt{1-x^2}$



Figure 3-C. The curve  $x^2 = y(1 + \sqrt{1+y})$

Figure 3-B. The curve  $y = \pm x\sqrt{x-1}$

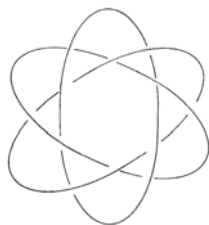
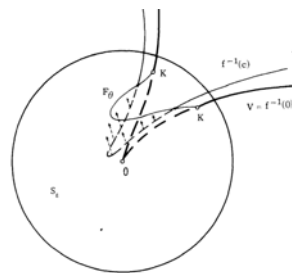


Figure 5. The link K associated with  $x_1^3 + x_2^6 = 0$ .

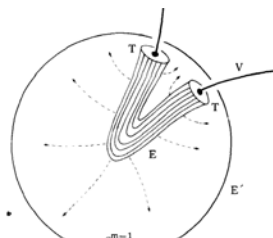


Figure 6.

As seis figuras do livro de J. Milnor "Singularities of complex hypersurfaces" (1968).



A famosa figura *Melencolia* de Dürer (1514). O politopo não é  $K_5$ ! Eu recomendo um artigo do Günter Ziegler no *The Guardian* [Dürer's polyhedron: five theories that explain Melencolia's crazy cube.](#)

+

# O associaedro de Hiparco–Schroeder–Tamari–Stasheff

NÓS VAMOS ESQUECER AS CURVAS ANALÍTICAS por enquanto e vamos voltar para as árvores, as palavras e a combinatória. Temos um dicionário natural entre três tipos de objetos.

- *Árvores planas enraizadas binárias* com  $n$  folhas.
- *Parênteses binários* em uma palavra de tamanho  $n$ .
- *Partições de um polígono convexo* com  $(n + 1)$  arestas (uma delas sendo chamada “a raiz”) em  $n$  triângulos.

Isso está ilustrado pelas figuras na margem.

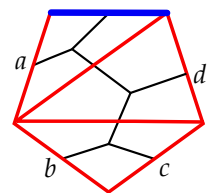
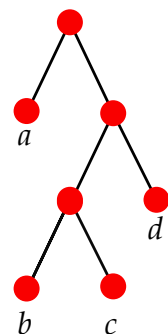
Também temos discutido árvores enraizadas com  $n$  folhas tais que todo vértice interno tenha *pelo menos* dois filhos. Essas árvores estão associadas a parênteses de Schroeder em uma palavra de comprimento  $n$  que não sejam necessariamente binárias. Em termos das diagonais, em um polígono convexo de  $(n + 1)$  lados, elas correspondem a coleções de  $k$  diagonais sem interseção, com  $0 \leq k \leq n$ . O número desses objetos é o  $n$ -ésimo (pequeno) número de Hiparco–Schroeder.

## Um politopo abstrato

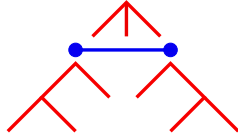
Estamos indo para a construção de uma sequência de politopos  $K_n$ , de dimensão  $n - 2$ , chamados de *associaedros*.



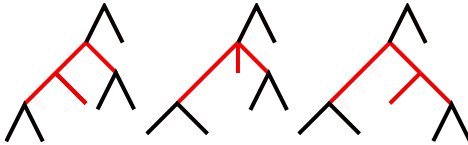
Um modelo de  $K_5$ .



Desenhe um intervalo e rotule-o com a única árvore de raiz que tenha três filhos e nenhum outro vértice. Rotule seus pontos finais com as duas árvores binárias planas com três folhas. Nós temos a seguinte figura. Isso é  $K_3$ : apenas um intervalo.



Como sugerido na figura anterior, se podemos passar de uma para outra por uma transição local, é muito tentador conectar duas árvores binárias por uma aresta. Se você detectar dentro de uma árvore binária alguma subárvore com três folhas, você a apaga e substitui pela outra árvore com três folhas, definindo uma aresta no associaedro.

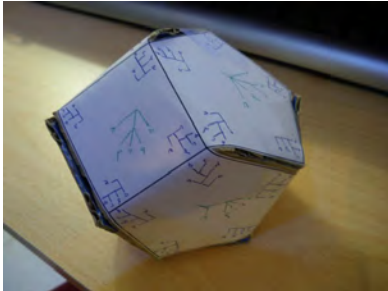
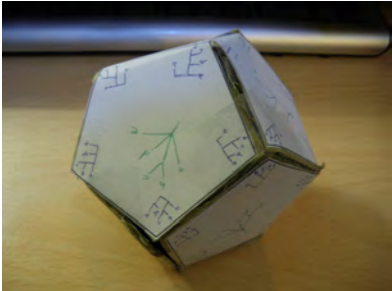
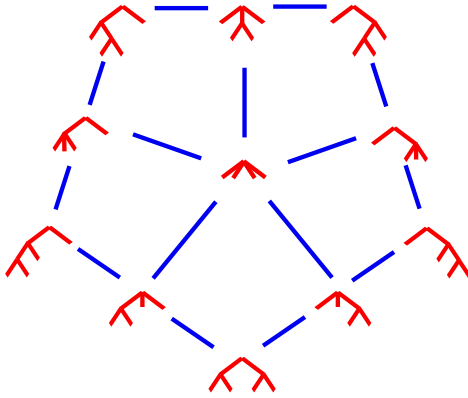


Vamos construir uma figura para  $n = 4$ . Existem cinco árvores com quatro folhas. Nós as colocamos nos vértices de um pentágono. As cinco folhas são rotuladas por cinco árvores planares de quatro folhas e exatamente um vértice com 3-filhos. Temos mais uma árvore planar enraizada com quatro filhos. Nós a colocamos no centro do pentágono, como um rótulo para a face de dimensão 2 do Pentágono. Isso é o  $K_4$ .

Isso sugere que se poderia definir algum “politopo” de dimensão  $n - 2$  cujos vértices são rotulados por árvores binárias com  $n$  folhas, cujas arestas são rotuladas por árvores com um único vértice com três filhos etc. cuja única face de dimensão máxima  $n - 2$  é rotulada pela única árvore enraizada com  $n$  filhos (geralmente chamada de *corola*).

Indo para  $n = 5$ , ainda podemos construir uma figura.

Acontece que é realmente possível construir um tal politopo





para todos os valores de  $n$ . Mas o primeiro problema é dar uma definição precisa da palavra “politopo” em um *contexto combinatório*. Gostaríamos de uma definição inspirada em nossa intuição geométrica de um politopo *no espaço euclidiano*, mas que não deveria levar em conta um mergulho em algum espaço.

Existe um conceito bem definido de *poliedro combinatório* cujas faces são segmentos, triângulos e simplexos em geral. Comece com um conjunto  $V$  de pontos chamados “vértices” e selecione alguns subconjuntos de  $V$ , denominados “faces”, respeitando apenas uma condição: um subconjunto de uma face deve ser uma face. Se uma face contém  $k + 1$  elementos, diz-se que é um simplexo de dimensão  $k$ . Essa é uma definição bastante fácil, mas isso não é adequado em nossa situação. Por exemplo, o politopo dimensão 3  $K_5$  acima tem faces bidimensionais que são quadrados ou pentágonos, não triângulos.

Há, de fato, várias definições combinatórias (não equivalentes) de *politopos abstratos*, mas não as usaremos, uma vez que eventualmente perceberemos nosso politopo como um objeto geométrico no espaço euclidiano. Não obstante, um “politopo abstrato” deveria pelo menos ser feito de “faces” tendo alguma dimensão e haver alguma ordenação parcial entre faces, correspondendo a ideia intuitiva de adjacência. Então, vamos nos contentar com a definição de um conjunto parcialmente ordenado  $K_n$  de dimensão  $n - 2$ . Mas isso é muito fácil.

Para simplificar, escolheremos alguns polígonos convexos no plano  $\Pi_{n+1}$  com  $(n + 1)$  vértices, mas a construção a seguir é independente da escolha desse polígono. Vamos também selecionar um lado do polígono, chamado de *raiz*.

Uma face de dimensão  $d$  de  $K_n$  é, por definição, um conjunto de  $F$  de  $n - 2 - d$  diagonais sem interseção em  $\Pi_{n+1}$ . A relação de adjacência é definida usando a inclusão *reversa*: dizemos que uma face associada a um subconjunto  $F_1$  é uma *subface* de  $F_2$  se  $F_2 \subset F_1$ . Por exemplo, vértices de  $K_n$ , de dimensão 0, correspondem a partições de  $\Pi_{n+1}$  em  $(n - 1)$  triângulos por  $n - 2$  diagonais. Usando a raiz, esses vértices estão associados a árvores binárias planares, como queríamos.

Uma face de codimensão  $q$  de  $K_n$  está associada a uma árvore

A dimensão de um conjunto parcialmente ordenado é a cardinalidade de um subconjunto máximo totalmente ordenado menos 1.

planar com raiz com  $n$  folhas e exatamente  $q$  vértices internos (diferentes da raiz e das folhas) ou, equivalentemente, com  $q$  arestas internas. Visto do “ponto de vista da árvore”, pode-se dizer que a face associada a alguma árvore  $T_1$  é uma subface daquela associada a  $T_2$  se obtiver  $T_2$  de  $T_1$  colapsando algumas arestas.

Por enquanto, só defini alguns conjuntos parcialmente ordenados. Não seria difícil verificar se isso satisfaz os axiomas que definem os politopos abstratos. . . que escolhi não explicitar.

Este é o *associaedro de Hiparco–Schroeder–Tamari–Stasheff*.

### Algumas histórias

Como de costume, dar um único nome a um objeto matemático é quase impossível.

Como sabemos, Catalan contou o número de vértices de  $K_n$  e Hiparco e Schroeder contaram suas faces.

Dov Tamari (anteriormente Bernhard Teitler) definiu o objeto combinatório em 1951 em sua dissertação.

Sugiro que você leia o primeiro capítulo do Festschrift de Tamari <sup>125</sup> para entender a motivação dele e conhecer a sua biografia (um “passeio” na Alemanha, Palestina, França, Israel, EUA, Brasil e Holanda ao longo do século XX). Aprende-se, por exemplo, que

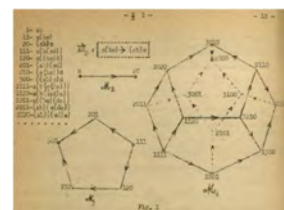
Pelo menos a partir de 1948, Tamari se opôs às injustiças que os israelenses fizeram aos palestinos, bem como a discriminação dirigida contra os imigrantes judeus dos países do Oriente Médio (essas visões não eram amplamente aceitas naqueles dias).

Em 1963, Jim Stasheff definiu o mesmo objeto, também em sua dissertação, mas em um contexto topológico muito diferente, que discutiremos com algum detalhe no próximo capítulo. Destaco que ele não estava ciente do trabalho anterior de Tamari. A imagem na margem mostra o “politopo curvo” de seu artigo original.

A construção de um politopo convexo em algum *espaço euclidiano* era uma questão muito natural. De acordo com uma anedota,

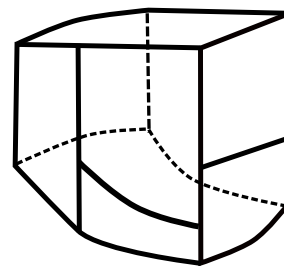


Dov Tamari (1911–2006).



Uma figura da dissertação de Tamari.

<sup>125</sup> F. Müller-Hoissen, J. M. Pallo, and J. Stasheff, editors. *Associahedra, Tamari lattices and related structures. Tamari memorial Festschrift*, volume 299 of *Progress in Mathematical Physics*. Birkhäuser/Springer, Basel, 2012.



A figura de Stasheff.

Milnor compareceu a defesa do PhD de Stasheff com um modelo de papelão de  $K_5$

O nome “associaedro” foi cunhado por Gil Kalai que perguntou a Mark Haiman se existiria um politopo convexo geométrico (não abstrato) em  $\mathbb{R}^n$  que realiza  $K_{n-2}$ . Haiman forneceu alguma construção em 1984, mas não a publicou. Uma construção foi publicada por Carl W. Lee em 1989. Vários autores forneceram outras construções. Para maiores detalhes, veja o capítulo correspondente de Cesar Ceballos e Günter Ziegler no Festschrift de Tamari.

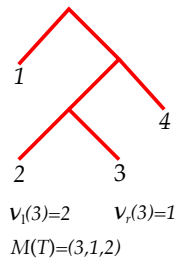
*A construção de Loday*

Para encerrar este capítulo, descrevo uma bela construção de Jean-Louis Loday <sup>126</sup> em 2004, de um politopo convexo *no espaço euclidiano* cujas faces (no sentido geométrico) realizam precisamente a combinatória do associaedro de Hiparco e outros.

Considere uma árvore binária planar enraizada  $T$  com  $n$  folhas, pensada como um vértice de  $K_n$ . Rotule as folhas de 1 a  $n$  da esquerda para a direita. Para cada par de folhas  $i, j$ , denotamos por  $i \vee j$  o vértice de  $T$  que é o mais jovem ancestral comum de  $i$  e  $j$ . Para cada inteiro  $i$  tal que  $1 \leq i \leq n - 1$ , considere o vértice  $i \vee (i + 1)$  e denote por  $v_l(i)$  (resp.  $v_r(i)$ ) o número de seus descendentes ao longo do ramo esquerdo (resp. direito). Associamos a árvore  $T$  ao ponto

$$M(T) = (v_l(1)v_r(1), v_l(2)v_r(2), \dots, v_l(n - 1)v_r(n - 1)) \in \mathbb{R}^{n-1}.$$

<sup>126</sup> J.-L. Loday. Realization of the Stasheff polytope. *Arch. Math. (Basel)*, 83(3):267–278, 2004.



**Teorema.** *O fecho convexo do conjunto de pontos  $M(T) \in \mathbb{R}^{n-1}$  que  $T$  descreve todas as árvores binárias planares enraizadas é um politopo convexo cuja combinatória é precisamente a do associaedro de Hiparco–Schroeder–Tamari–Stasheff.*

Primeiro mostraremos que todos os pontos  $M(T)$  estão no hiperplano de  $\mathbb{R}^{n-1}$  cuja equação é

$$x_1 + x_2 + \dots + x_{n-1} = \frac{n(n - 1)}{2}.$$

Uma maneira de provar é contar o número de triplas  $(a, b, v)$  em que  $a < b$  são duas folhas com  $v = a \vee b$ . Como  $v$  é determinado por  $a, b$ , esse número é igual ao número de pares  $a < b$ , igual a  $n(n-1)/2$ . Se contarmos este mesmo número de acordo com o vértice  $v$ , obteremos a soma do  $v_l(i)v_r(i)$  de  $i = 1$  a  $n-1$ . Isso prova a afirmação.  $\square$

Para provar o teorema de Loday, primeiro identificamos as faces  $F$  de codimensão 1 de  $K_n$ . Elas são rotuladas por árvores (não necessariamente binárias) com um único vértice interno. Elas são definidas por dois inteiros  $1 \leq p < p+q-1 \leq n$  e obtidas enxertando uma  $q$ -corola no  $p$ -ésimo vértice da  $n-q+1$ -corola. O conjunto  $F_{p,q}$  dos vértices dessa face é o conjunto das árvores (planares enraizadas) binárias que são obtidas enxertando qualquer árvore (binária, planar e com raiz) com  $q$  folhas na folha rotulada  $p$  em alguma árvore (binária, planar e com raiz) com  $(n-q+1)$  folhas.

Pode-se expressar o mesmo ainda de uma maneira diferente. Uma árvore binária  $T$  pertence a  $F_{p,q}$  se, e somente se, a menor subárvore binária  $T_{p,q}$  de  $T$  contendo as folhas  $\{p, \dots, p+q-1\}$  não contém nenhuma outra folha.

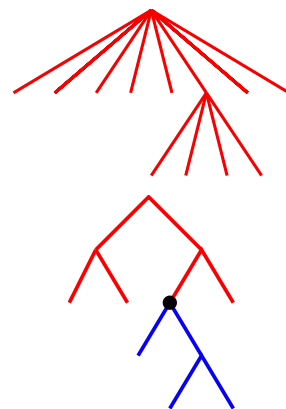
Defina uma função linear  $l_{p,q}$  em  $\mathbb{R}^{n-1}$  por:

$$l_{p,q}(x_1, x_2, \dots, x_{n-1}) = x_p + x_{p+1} + \dots + x_{p+q-1}.$$

Vamos avaliar  $l_{p,q}(M(T))$  quando  $T$  é um vértice da face  $F_{p,q}$  de codimensão 1. Sabemos que  $T$  é um enxerto de alguma árvore (binária, planar e enraizada) com  $q$  folhas na folha rotulada  $p$  em alguma árvore (binária, planar e enraizada) com  $(n-q+1)$  folhas. Claramente, a  $i$ -ésima coordenada de  $M(T)$  para  $p \leq i \leq p+q-1$ , é a  $(ip)$ -ésima coordenada de  $M(T_1) \in \mathbb{R}^{q-1}$  de modo que  $l_{p,q}(T) = q(q-1)/2$  como vimos.

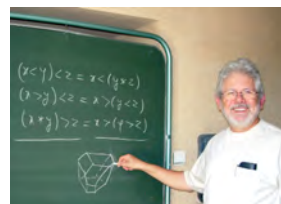
Suponha agora que  $T$  não esteja em  $F_{p,q}$ , então, existe pelo menos alguma folha  $i$  no intervalo  $\{p, \dots, p+q-1\}$  tal que  $i \vee (i+1)$  tem algum descendente fora dele. Ao computar  $l_{p,q}(M(T))$ , sabemos que obtemos  $q(q-1)/2$  se contarmos apenas os descendentes de  $i \vee (i+1)$  que estão dentro do intervalo  $\{p, p+q-1\}$ . Qualquer descendente que esteja fora do intervalo  $\{p, p+q-1\}$  produz uma soma maior.

A prova do teorema está terminada. De fato, para cada face



$F_{p,q}$  de codimensão 1 de  $K_n$ , mostramos que a função afim  $l_{p,q} - q(q-1)/2$  é zero em todos os  $M(T)$  para todos os vértices  $T$  de  $F_{p,q}$  e positivo em todos os  $M(T)$  para vértices  $T$  de  $K_n$  que não estejam em  $F_{p,q}$ . Em outras palavras, encontramos funções afins de apoio que mostram explicitamente que o fecho convexo dos pontos  $M(T)$  tem, de fato, a combinatoria de nosso politopo abstrato.

Jean-Louis Loday explicou a descoberta desse mergulho em um belo artigo online <sup>127</sup>. Ele frequentemente usou "Guillaume William Zinbiel" como pseudônimo, por sua admiração à Leibniz.



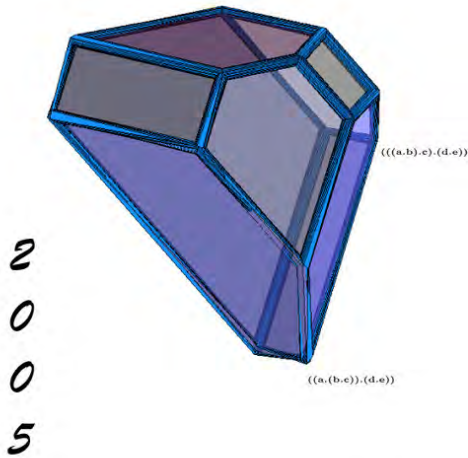
Jean-Louis Loday (1946–2012) falando sobre o associaedro. +

<sup>127</sup> J.-L. Loday. [Comment j'ai trouvé l'associaèdre.](#)

# C.I.R.M.

*Centre International de Rencontres Mathématiques*

Este é o cartão da temporada CIRM para 2005, representando  $K_3$ . Eles até produziram uma camiseta. †



Associahédre  $K_3$



“Cherry Tree” de Hasegawa Kyūzō, biombo pintado com papel folheado a ouro (1592).  
†

## Jim Stasheff e os espaços de laços



Jim Stasheff. +

NÃO HÁ NECESSIDADE DE RECORDAR A IMPORTÂNCIA DOS GRUPOS na matemática em geral e em particular na topologia. Ocorre que um dos problemas é o caráter bastante sutil desse conceito na *teoria da homotopia* conforme explicado neste capítulo.

Relembre que duas aplicações contínuas  $f_0, f_1 : X \rightarrow Y$  são *homotópicas* quando existe uma aplicação contínua  $F : X \times [0, 1] \rightarrow Y$  (chamada homotopia) tal que  $F(x, 0) = f_0(x)$  and  $F(x, 1) = f_1(x)$ . Dois espaços topológicos  $X, Y$  têm o mesmo *tipo de homotopia* se houver *equivalência homotópica*  $f : X \rightarrow Y$  e  $g : Y \rightarrow X$ , i.e., uma aplicação tal que  $f \circ g$  and  $g \circ f$  são homotópicas à identidade. A teoria da homotopia estuda a *categoria homotópica* cujos objetos são espaços topológicos e cujas morfismos são aplicações a menos de homotopia.

Como um exemplo trivial, um intervalo fechado não pode ser *homeomorfo* a um grupo topológico, pois o conjunto de seus dois pontos finais é invariante sob qualquer homeomorfismo, enquanto qualquer grupo age transitivamente em si mesmo por meio de translações. No entanto, o intervalo é *contrátil*: tem o mesmo tipo de homotopia que um ponto, que é um grupo (trivial). Em sua dissertação de 1961 (publicada em 1963 <sup>128</sup>), Jim Stasheff abordou a questão de determinar quais espaços têm o tipo de homotopia de um grupo topológico.



Atenção: desvio! Estritamente falando, este capítulo não é necessário para o restante do livro, mas servirá como uma motivação para o conceito de operad que, por sua vez, também não é necessário, apesar de lançar luz sobre a figura geral. Provavelmente, leitor, é a parte mais difícil do nosso passeio. +

<sup>128</sup> J. D. Stasheff. [Homotopy associativity of H-spaces. I, II.](#) *Trans. Amer. Math. Soc.* 108 (1963), 275-292; *ibid.*, 108:293-312, 1963.



## Grupos Topológicos e fibrados principais

Nesta seção, dou uma breve visão geral do papel dos grupos topológicos na teoria da homotopia. Meu único objetivo é introduzir a terminologia e fatos básicos suficientes para explicar a contribuição de Stasheff. Mais uma vez, este é um território enorme e eu tenho que me referir a livros excelentes, como por exemplo, o de ... Milnor e Stasheff <sup>129</sup>, lidando com as chamadas classes características.

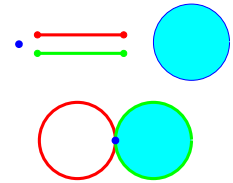
Seja  $G$  um grupo topológico, ou seja, um grupo equipado com uma topologia de tal forma que a composição e as aplicações inversas sejam contínuas. Um  $G$ -fibrado principal é uma ação livre de  $G$  em algum espaço  $X$  com um “bom quociente”  $B$ . Cada ponto em  $X$  deve ter uma vizinhança  $G$ -invariante homeomorfa a um produto  $U \times G$  em que a  $G$ -ação é apenas a ação por translações no segundo fator. Frequentemente, é mais conveniente pensar no fibrado como uma aplicação de projeção  $p : X \rightarrow B$  no espaço  $B$  das  $G$ -órbitas. Um  $G$ -fibrado  $p' : X' \rightarrow B$  é isomorfo a  $p$  se houver um homeomorfismo  $G$ -equivariante entre  $X$  e  $X'$ , induzindo a identidade em  $B$ . Diremos que o espaço total  $X$  está sobre a base  $B$  e que a imagem inversa de um ponto por  $p$ , uma  $G$ -órbita, é uma fibra.

Neste ponto, devemos ser muito cautelosos sobre o tipo de espaços topológicos que serão usados. Eles devem ser Hausdorff e não muito patológicos. Normalmente, restringe-se o estudo a CW-complexos. Como não é minha intenção dar uma descrição precisa desses espaços, apenas mencionarei que tal espaço  $X$  é, por definição, uma união crescente de subespaços  $Sk_n(X)$ , chamados de  $n$ -ésimos esqueletos. De forma que o  $(n+1)$ -ésimo esqueleto  $Sk_{n+1}(X)$  é obtido a partir de  $Sk_n(X)$ , colando algumas bolas  $(n+1)$ -dimensionais  $B^{n+1}$  ao longo de algumas “aplicações de colagem”  $u : \partial B^{n+1} \rightarrow Sk_n(X)$ . O livro de Hatcher <sup>130</sup> (disponível gratuitamente na Internet) é uma excelente referência.

Os fibrados principais são objetos fundamentais na topologia (diferencial). Por exemplo, dada uma variedade diferenciável  $M$  de dimensão  $m$ , seja  $Fr(M)$  o espaço de pares  $(x, f)$  sendo  $x$  um ponto de  $M$  e  $f$  um referencial em  $x$ , em outras palavras,

<sup>129</sup> J. W. Milnor and J. D. Stasheff. *Characteristic classes*, volume 76 of *Annals of Mathematics Studies*. Princeton University Press, Princeton, N.J., 1974.

Tradicionalmente, no contexto dos fibrados principais, os grupos atuam à direita.



Tome um ponto, dois intervalos e um disco. Cole os pontos finais dos intervalos num ponto, produzindo uma figura oito. Cole a borda do disco para uma componente dos oito. Você obtém um exemplo muito simples de um CW-complexo. É claro que, em geral, as aplicações de colagem podem ser muito mais complicadas.

<sup>130</sup> A. Hatcher. *Algebraic topology*. Cambridge UP, 2002.

uma base do espaço tangente  $T_x(M)$ . Há uma ação livre óbvia do grupo linear  $GL(m, \mathbb{R})$  em  $Fr(M)$  e o mapa  $p$ , enviando  $(x, f)$  em  $Fr(M)$  para  $x \in M$  um fibrado principal.

Dado um  $G$ -fibrado principal  $p : X \rightarrow B$  e uma aplicação  $i : B_1 \rightarrow B$ , o pull-back  $p$  produz um fibrado principal  $p_1 : X_1 \rightarrow B_1$ . Formalmente,  $X_1$  é o subespaço de  $B_1 \times X$  que consiste de pares  $(b_1, x)$  tal que  $i(b_1) = p(x)$  e  $p_1(b_1, x) = b_1$ . Por exemplo, se  $i$  é uma inclusão,  $p_1$  é a restrição para “o que está acima de  $i(B_1)$  em  $X$ ”.

Aqui está um exemplo importante: Seja  $Gr_{k,n}$  o espaço de subespaços lineares de dimensão  $k$  em  $\mathbb{R}^n$ . Essa é uma variedade compacta, chamada *Grassmanniana*. Existe um fibrado tautológico  $GL(k, \mathbb{R})$  sobre  $Gr_{k,n}$  cuja fibra sobre algum subespaço consiste nas bases daquele subespaço. Se uma variedade  $k$  dimensional  $M$  está imersa em  $\mathbb{R}^n$ , então, a diferencial dessa imersão dá uma aplicação de  $M$  para  $Gr_{k,n}$ . O pull-back do  $GL(k, \mathbb{R})$  é um fibrado tautológico sobre  $Gr_{k,n}$  que é isomorfo ao fibrado dos referenciais de  $M$ .

Deixe-me mencionar dois fatos importantes sobre os fibrados principais:

– Qualquer fibrado com uma base contrátil  $B$  é trivial(izável), ou seja, isomorfo a  $B \times G$ . Veja <sup>131</sup> para a história desse teorema.

– Se  $i, i' : B_1 \rightarrow B$  são homotópicos, então o pull-back dos fibrados principais  $p_1, p'_1$  de  $p$  por  $i, i'$  são isomorfos.

Essas duas propriedades mostram que o conjunto de classes de isomorfismos dos  $G$ -fibrados principais sobre algum espaço  $B$  depende apenas do tipo de homotopia de  $B$  e define algum functor contravariante na categoria de homotopia.

### *Classificando os espaços*

Um fibrado  $G$ -principal  $p_G : E(G) \rightarrow B(G)$  é chamado *universal* se todo  $G$ -fibrado principal  $p : X \rightarrow B$  é isomorfo ao pull-back de  $p_G$  para alguma aplicação  $i : B \rightarrow B(G)$  que é única, a menos de homologia. Mais tarde, iremos esboçar uma prova do seguinte teorema.

**Teorema.** *Para todo grupo topológico  $G$ , existe um fibrado universal*

<sup>131</sup> M. Audin. *Publier sous l'Occupation. I. Autour du cas de Jacques Feldbau et de l'Académie des sciences. Rev. Histoire Math.*, 15(1):7–57, 2009.



Sequências espectrais e órbitas da ação de grupos, por A. Fomenko. +

$$p_G : E(G) \rightarrow B(G).$$

Em outras palavras, existe uma bijeção natural entre:

– (classes de isomorfismos de) fibrados  $G$ -principais sobre algum espaço  $B$ .

– Classes de homotopia  $[B, B(G)]$  de aplicações de  $B$  em  $B(G)$ .

Diremos que  $B(G)$  é o *espaço classificante* de  $G$ .

Deixe-me descrever dois exemplos importantes: Como primeiro exemplo, suponha que  $G$  seja um *grupo discreto*. Em tal situação, um fibrado  $G$ -principal nada mais é do que uma aplicação de recobrimento galoisiano com grupo Galois  $G$ . Os espaços de recobrimento de um espaço  $B$  são descritos por subgrupos de seu grupo fundamental. Nesse caso,  $B(G)$  é um espaço de *Eilenberg–MacLane*  $K(G, 1)$ : seu grupo fundamental é  $G$  e seu recobrimento universal  $E(G)$  é contrátil (equivalentemente todos grupos de homotopia mais altos de  $K(G, 1)$  são triviais). Um recobrimento de algum espaço  $B$  e de grupo de Galois  $G$  é equivalente a uma classe de homotopia de aplicações  $B \rightarrow K(G, 1)$ .

*Uma observação importante:*

Para definir o grupo fundamental de um espaço  $X$ , precisamos de um ponto base  $x \in X$ . Uma notação como  $f : (X, x) \rightarrow (Y, y)$  significa que  $f(x) = y$ . Para discutir homotopia de aplicações, devo mencionar explicitamente se essa homotopia preserva os pontos de base ou não. Eu deveria ... mas não vou! Se fizesse, implicaria frases longas e técnicas e, como o leitor já percebeu, este livro não é uma enciclopédia completa. Espero que meu querido leitor perdoe essa falta de precisão.

Como segundo exemplo, considere o grupo  $U(1)$  dos números complexos de módulo 1. Para cada  $n$ , há uma ação de  $U(1)$  na esfera unitária  $S^{2n-1}$  em  $\mathbb{C}^n$ . O elemento  $\omega$  age em  $(z_1, \dots, z_n)$  para produzir  $(\omega z_1, \dots, \omega z_n)$ . Isso define um  $U(1)$ -fibrado principal

$$p_n : S^{2n-1} \rightarrow \mathbb{C}P^n$$

sobre o espaço projetivo complexo. Todas essas esferas e espaços projetivos são naturalmente encaixadas, mergulhando  $(z_1, \dots, z_n)$  em  $(z_1, \dots, z_n, 0, \dots, 0)$ , para que possamos definir um  $U(1)$ -fibrado principal cujo espaço total é a esfera dimensional

Para a teoria de espaços de recobrimento, veja o livro de Hatcher ou [Analysis Situs](#).



Poliedros e cadeias simpliciais 1973, A. Fomenko.

†

infinita e cuja base é o espaço projetivo dimensional de dimensão infinita.

$$p_\infty : S^\infty \rightarrow CP^\infty.$$

Veremos que esse é o fibrado universal para  $G = U(1)$ . O ponto-chave é o seguinte:

**Proposição.** *Um  $G$ -fibrado é universal e o seu espaço total é contrátil.*

A prova desse fato fundamental é um exemplo típico da *teoria da obstrução*. Comece com algum  $G$ -fibrado  $p_G : E(G) \rightarrow B(G)$  tal que  $E(G)$  seja contrátil e deixe-nos mostrar que é universal. Considere algum outro  $G$ -fibrado  $p : E \rightarrow B$  e queremos mostrar que é o pull-back de  $p_G$  por alguma aplicação  $i : B \rightarrow B(G)$ . Construímos  $i$  nos esqueletos de  $B$  (que é, como sempre, um CW-complexo), por indução em suas dimensões. A cada passo, alguma aplicação contínua tem que ser estendida e a contratibilidade de  $E(G)$  é precisamente o necessário para esta construção. Veja o livro de Milnor e Stasheff para os detalhes e para a prova da recíproca.

No exemplo de  $U(1)$ , o espaço classificante  $B(U(1))$  é  $CP^\infty$  já que a esfera dimensional infinita é de fato contrátil.

### A construção de Milnor

A construção de  $B(G)$  por Milnor é linda e fácil <sup>132</sup>. Qualquer grupo topológico  $G$  age livremente sobre si mesmo, mas é claro que o grupo não precisa ser contrátil. Portanto, temos que forçar a contratibilidade, preservando uma ação livre de um grupo. Seja  $E(G)$  o *simplexo sobre  $G$* . Um elemento de  $E(G)$  é, por definição, uma combinação baricêntrica formal finita de elementos de  $G$ , isto é, uma soma formal

$$\lambda_0 g_0 + \lambda_1 g_1 + \dots + \lambda_k g_k$$

sendo  $\lambda_i \geq 0$  and  $\sum_i \lambda_i = 1$ . Este espaço é convexo, portanto, contrátil, estando equipado com uma ação livre de  $G$ . A projeção de  $E(G)$  no seu quociente  $B(G)$  é, portanto, um espaço classificante. Et voilà!

Sugestão: verifique tudo sozinho, sem abrir muito os livros recomendados! Porque a esfera  $S^\infty$  é contrátil?



Espaços celulares, por A. Fomenko.

<sup>132</sup> J. Milnor. [Construction of universal bundles. II.](#) *Ann. of Math.* (2), 63:430–436, 1956.

Poderíamos ser mais cuidadosos na definição de  $E(G)$ . Começando com a união disjunta dos produtos  $G^{n+1} \times \Delta_n$  sendo

$$\Delta_n = \{(\lambda_0, \dots, \lambda_n) \mid \lambda_i \geq 0 \text{ and } \sum_i \lambda_i = 1\}$$

é um simplexo padrão. Então, introduzimos uma relação de equivalência “óbvia” gerada por

$$\begin{aligned} & ((g_0, \dots, g_i, g_{i+1}, g_{i+2}, \dots, g_n), (\lambda_0, \dots, \lambda_i, 0, \lambda_{i+2}, \dots, \lambda_n)) \in G^{n+1} \times \Delta_n \\ & \equiv ((g_0, \dots, g_i, g_{i+2}, \dots, g_n), (\lambda_0, \dots, \lambda_i, \lambda_{i+2}, \dots, \lambda_n)) \in G^n \times \Delta_{n-1} \end{aligned}$$

definindo  $E(G)$  como o espaço quociente. Isso é a *construção “join” de Milnor*. O nome vem do fato de que “conexões virtuais” foram criadas, juntando pontos em  $G$ .

### Laços e suas composições

Dado um espaço  $B$  com um ponto base  $\star \in B$ , o *espaço dos laços baseados*  $\Omega(B, \star)$  é o espaço de ... laços baseados  $\odot$ , isto é, das aplicações contínuas  $\gamma : [0, 1] \rightarrow B$  tais que  $\gamma(0) = \gamma(1) = \star$  equipado com a topologia aberta compacta. Dados dois laços baseados  $\gamma_1, \gamma_2$ , eles podem ser *concatenados*. Uma definição possível é definir  $\gamma_1 \bullet \gamma_2(t)$  como  $\gamma_1(2t)$  para  $0 \leq t \leq 1/2$  e  $\gamma_2(2t - 1)$  para  $1/2 \leq t \leq 1$ . Essa aplicação de composição

$$\Omega(B, \star) \times \Omega(B, \star) \rightarrow \Omega(B, \star)$$

certamente não é associativa. Na composição  $(\gamma_1 \bullet \gamma_2) \bullet \gamma_3$  ao longo de  $\gamma_1$  quando  $t \in [0, 1/4]$ , então, ao longo de  $\gamma_2$  quando  $t \in [1/4, 1/2]$  e, finalmente, ao longo de  $\gamma_3$  quando  $t \in [1/2, 1]$ . Esse não é o mesmo caminho que  $\gamma_1 \bullet (\gamma_2 \bullet \gamma_3)$ , mesmo que esses dois laços sejam homotópicos.

Alterando um pouco as definições, obtemos um espaço de laços que é *estritamente associativo*, não apenas a menos de homotopia. Vamos usar os chamados *laços de Moore*, que consiste em algum número  $l \geq 0$  (pensado com algum comprimento) e alguma aplicação contínua  $\gamma : \mathbb{R}_+ \rightarrow B$  tal que  $\gamma(0) = \star$  e  $\gamma(t) = \star$  para  $t \geq l$ . Há uma topologia natural no espaço desses laços estranhos, denotada por  $\Omega_M(B, \star)$ , que tem o mesmo tipo de



Complexos simpliciais em 1973, por A. Fomenko. †

homotopia de  $\Omega(B, \star)$ . Dados  $(l_1, \gamma_1)$  e  $(l_2, \gamma_2)$ , a composição é definida como  $(l_1 + l_2, \tilde{\gamma})$  onde  $\tilde{\gamma}(t) = \gamma_1(t)$  para  $t \leq l_1$  e  $\gamma_2(t - l_1)$  para  $t \geq l_1$ . Isso é claramente associativo.

Isso é devido à Milnor (encore lui!). É descrito em um livro <sup>133</sup> de Stasheff. No caso, fazemos uma suposição muito leve:  $B$  é a realização geométrica de um complexo simplicial com um número enumerável de faces. Defina um grupo  $G(B)$  da seguinte maneira: Comece com a união disjunta dos  $B^n$  com  $n \geq 0$ . Pense em um elemento de  $B^n$  como um caminho discreto  $b_1, \dots, b_n$  com  $n$  passos, de forma que um ponto, em vez de seguir um caminho contínuo, salta de ponto a ponto. Considere a relação de equivalência:

$$(b_1, \dots, b_{i-1}, b_i, b_{i+1}, \dots, b_n) \equiv (b_1, \dots, b_{i-1}, b_{i+1}, \dots, b_n)$$

se  $b_i = b_{i+1}$  ou  $b_{i-1} = b_i$ .

Existe outro truque, menos conhecido, para obter outro espaço, ainda com o mesmo tipo de homotopia  $\Omega(B, \star)$ , que é um grupo topológico.

No espaço do quociente, defina  $G(B)$  como o subespaço de classes que tem um representante  $(b_1, \dots, b_n)$  tal que  $b_1 = b_n = \star$  e tal que quaisquer dois elementos consecutivos  $b_i, b_{i+1}$  estão no mesmo simplexo (então, mentalmente, podemos conectá-los por um segmento). A estrutura do grupo é apenas concatenação. É um exercício simples verificar que é realmente um grupo topológico, com o mesmo tipo de homotopia que o espaço de laços  $\Omega(B, \star)$ .

Bem, essa construção não é tão complicada, mas é preciso ter em mente que o grupo produzido é bastante extenso, mesmo que  $B$  seja muito simples. Esse grupo é muito raramente usado “na prática”.

De qualquer forma, devemos lembrar que um espaço  $(B, \star)$  define um usual  $\Omega(B, \star)$  equipado com alguma aplicação de concatenação, não associativa, mas que pode ser transformada em uma lei associativa ou mesmo em um grupo topológico ao custo de algumas contorções topológicas.

Uma observação final nesta seção:

**Proposição.** *Qualquer grupo topológico  $G$  tem o mesmo tipo de*

<sup>133</sup> J. Stasheff. *H-spaces from a homotopy point of view*. Lecture Notes in Mathematics, Vol. 161. Springer-Verlag, Berlin-New York, 1970.

Um *complexo simplicial* é um conceito combinatório que consiste de um conjunto  $V$  cujos elementos são chamados *vértices*, e uma família de subconjuntos finitos de  $V$ , chamados de *faces*. Temos um único axioma: um subconjunto não vazio de uma face é uma face. Dado um complexo simplicial, existe um espaço topológico associado chamado *realização geométrica*. Consiste de funções  $t : V \rightarrow [0, 1]$  tal que  $\sum_{x \in V} t(x) = 1$  e tal que  $\{x | t(x) \neq 0\}$  é uma face.

homotopia que o espaço de laços de seu espaço classificante.

Apenas listo as palavras-chave na prova para ilustrar o tipo de ginástica requerida nesta parte da topologia. Se  $(X, x)$  for um espaço com um ponto base, sua *suspensão*  $S(X, x)$  é obtida de  $X \times [0, 1]$  pelo colapso de  $X \times \{0\}$ ,  $X \times \{1\}$  e  $\{x\} \times [0, 1]$ , para um único ponto. Uma aplicação de  $S(X, x)$  para algum outro espaço  $(Y, y)$  é equivalente a uma aplicação de  $(X, x)$  para o espaço de laços  $\Omega(Y, y)$ . Os seguintes são equivalentes:

- uma classe de homotopia de aplicações de  $(X, x)$  para  $\Omega(B(G), e)$  ( $e$  é a unidade em  $G$ );
- uma classe de homotopia de aplicações de  $S(X, x)$  para  $(B(G), e)$ ;
- uma classe de isomorfismo de um  $G$ -fibrado sobre  $X \times [0, 1]$  trivializado sobre  $X \times \{0\}$ ,  $X \times \{1\}$  e  $\{x\} \times [0, 1]$ ;
- um caminho de  $G$ -fibrados  $p_t$  sobre  $X$  e isomorfismos entre  $p_0$  e  $p_1$  com o fibrado trivial  $X \times G$ .

Observe que, dado um fibrado trivializado  $X \times G \rightarrow X$ , as outras trivializações são simplesmente dadas pelas aplicações  $X \rightarrow G$ . De fato, um isomorfismo de  $X \times G \rightarrow X$  para si mesmo envia  $(b, g)$  para  $(b, u(g)g)$  para algumas aplicações  $u : B \rightarrow G$ . Portanto, as classes de homotopia de mapas de  $(X, x)$  a  $\Omega(B(G), e)$  estão em bijeções canônicas com as classes de homotopia de  $(X, x)$  a  $(G, e)$ . □

### O teorema de Stasheff em $H$ -espaços

Um espaço  $(X, *)$  é chamado de  $H$ -espaço se for equipado com uma “multiplicação”

$$m_2 : X \times X \rightarrow X$$

tal que  $m_2(x, *) = m_2(*, x) = x$ .

A questão estudada por Stasheff é denominada de *problema de reconhecimento*. É possível decidir a partir de  $X$  e  $m_2$  se existe algum espaço  $Y$  e alguma *equivalência de homotopia* de  $(X, x)$  para o espaço de laço  $\Omega(Y, y)$  que transforma  $m_2$  na concatenação de laços em  $\Omega(Y, y)$ ? Isso equivale ao reconhecimento de grupos topológicos entre  $H$ -espaços, a menos de homotopia.



Um espaço fibrado por A. Fomenko.

Note que o fato de que as translações à direita e à esquerda comutarem não é nada mais do que a associatividade.

$H$  é em homenagem a Heinz Hopf e não de Homotopia.

Esta é realmente uma questão fundamental:

*Qual é o conceito correto de grupo na categoria homotópica?*

Como o leitor certamente adivinhou, a resposta dada por Stasheff envolverá o associaedro introduzido no capítulo anterior.

Suponha que  $(X, x)$  tenha de fato o tipo de homotopia de algum espaço de laços  $\Omega(Y, y)$  e que  $m_2$  seja homotópico à concatenação. É conveniente usar o espaço de laços de Moore  $\Omega_M(Y, y)$  com sua concatenação associativa  $\mu$ . As duas aplicações

$$\begin{aligned} (x_1, x_2, x_3) \in X^3 &\mapsto m_2(m_2(x_1, x_2), x_3) \in X \\ &\mapsto m_2(x_1, m_2(x_2, x_3)) \in X \end{aligned}$$

devem ser homotópicas, uma vez que  $\mu$  é associativo. Essa condição não necessariamente ocorre, mas se acontecer, o  $H$ -espaço é associativo a menos de homotopia e dizemos que o  $H$ -espaço  $(X, m_2)$  é um  $A_1$ -espaço. A homotopia, neste caso, é uma aplicação de  $X^3 \times [0, 1]$  para  $X$  e o fator  $[0, 1]$  deve ser visto como o associaedro de  $K_3$ .

Quatro termos definem cinco aplicações de  $X^4 \rightarrow X$  associados às cinco árvores binárias planares com quatro folhas. Estas cinco árvores podem ser vistas como os vértices de um pentágono  $K_4$ . Na etapa anterior, consideramos cinco aplicações  $[0, 1] \times X^4 \rightarrow X$ , que concordam nas suas bordas e definem uma aplicação  $\partial K_4 \times X^4 \rightarrow X$ . Se  $X$  tem o tipo de homotopia de um espaço de laços  $\Omega_M(Y)$ , com multiplicação estritamente associativa  $\mu$ , esta aplicação tem que se estender ao pentágono completo  $K_4 \times X^4 \rightarrow X$ . Se isso acontecer, dizemos que  $X$  é um  $A_2$ -espaço.

Deve ficar claro que essa imagem continua em todas as dimensões.

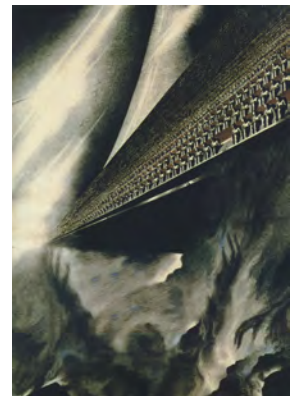
Podemos finalmente enunciar o teorema de Stasheff:

**Teorema.** *Um  $H$ -espaço  $m_2 : X \times X \rightarrow X$  é homotopicamente equivalente a algum espaço de laços se, e somente se, for um espaço  $A_\infty$ -espaço, ou seja, se existirem aplicações coerentes  $m_n : K_n \times X^n \rightarrow X$  compatíveis com as faces do associaedro  $K_n$  ( $n \geq 1$ ).*

A condição necessária é clara agora. A parte mais interessante do teorema é, evidentemente, a condição suficiente, que eu não irei provar na próxima seção.



Anti-Dürer - Do ciclo : Diálogo com autores do século XVI, 1975, por A. Fomenko.



Matar os espaços na topologia homotópica, por A. Fomenko.



Comece com algum espaço  $X$  de tipo  $A_\infty$ , com aplicações compatíveis  $m_n : K_n \times X^n \rightarrow X$ , nosso objetivo é produzir um espaço  $Y$  cujo espaço de laços tenha o mesmo tipo de homotopia como  $X$ . Sabemos que  $\Omega(Y, y)$  tem o mesmo tipo de homotopia que algum grupo topológico  $G$  que, por sua vez, tem o tipo de homotopia de  $\Omega(B(G), *)$ . Portanto, é tentador escolher  $Y = B(G)$ , mas não conhecemos  $G$ .

Afinal, só sabemos a coleção de mapas  $m_n : K_n \times X^n \rightarrow X$  que é algum tipo de substituto para uma estrutura de grupo. Portanto, nossa estratégia é clara: temos que adaptar a construção de Milnor de  $B(G)$  para essas  $A_\infty$ -estruturas mais gerais. Este projeto foi realizado por Stasheff.

Em vez de começar com a união disjunta de  $G^{n+1} \times \Delta_n$  e identificar pontos de acordo com alguma relação de equivalência “óbvia”, comece com a união disjunta de  $K_n \times X^n$  e defina alguma relação de equivalência “óbvia” nessa união disjunta. Isso produzirá um espaço  $B(X)$ , o *espaço classificante* do  $A_\infty$ -espaço  $X$ .

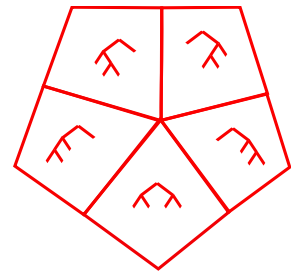
A única coisa que ainda temos que “mostrar” é, como esperado, o espaço de laços de  $B(X)$  ser uma solução para o nosso problema: isto é, um *desenlaçamento* de  $X$ . Isso não é fácil e Stasheff provou isso com alguma hipótese adicional (menor) sobre a topologia de  $X$ .

### Cerejeiras

A fim de obter alguma intuição por trás de  $B(X)$ , deixe-me descrever brevemente as *Cerejeiras* introduzidas por Boardman e Vogt<sup>134</sup>. Considere uma árvore binária planar enraizada com  $n$  folhas. Se as  $n - 2$  arestas internas estiverem equipadas com algum comprimento em  $[0, 1]$ , obteremos uma árvore *métrica*. O conjunto dessas métricas define um cubo  $[0, 1]^{n-2}$  para cada árvore binária.

Se uma ou mais arestas internas tiverem comprimento 0, elas poderão ser colapsadas e o resultado será uma árvore planar com raiz, que deixa de ser binária, mas cujas arestas internas ainda têm comprimento. Isso produz algumas identificações ao longo das bordas desses cubos. O conjunto dessas árvores

<sup>134</sup> J. M. Boardman and R. M. Vogt. *Homotopy invariant algebraic structures on topological spaces*. Lecture Notes in Mathematics, Vol. 347. Springer-Verlag, Berlin-New York, 1973.



métricas define uma decomposição cúbica de  $K_n$ . Portanto,  $K_n$  pode ser visto como um espaço de árvores métricas. Por exemplo, o pentágono é decomposto em cinco quadrados. Esta apresentação de  $K_n$  como um espaço de árvores métricas nos permite definir *aplicações de enxerto*  $\iota_{k_1, \dots, k_n}$ :

$$K_n \times (K_{k_1} \times K_{k_2} \times \dots \times K_{k_n}) \rightarrow K_{k_1 + \dots + k_n}.$$

Assim, basta colar árvores métricas nas folhas de uma árvore métrica.

Agora, queremos visualizar  $K_n \times X^n$ . Simplesmente imagine que cada folha de nossas árvores métricas carrega algum elemento de  $X$ , pensado como uma cereja. Esta é uma *cerejeira*: uma árvore métrica com cerejas nas folhas.

O espaço classificante  $B(X)$  pode ser descrito usando essa terminologia. Suponha que uma cerejeira  $T$  contenha uma aresta  $e$  que seja *totalmente crescida* de comprimento 1. Cortando  $e$ , a árvore  $T$  é decomposta em duas árvores métricas. Deixe-nos denotar por  $T_1$  a parte que consiste em descendentes do ponto final de  $e$ : esta é uma cerejeira com  $k \leq n$  folhas. A outra árvore,  $T_0$ , contendo a raiz de  $T$ , não é bem uma cerejeira, pois a folha recém-criada, na origem de  $e$ , não está equipada com uma cereja. Agora podemos avaliar  $m_k$  na árvore de cereja  $T_1$  e depositar o resultado como uma nova cereja na folha de  $T_0$ , que estava esperando por sua cereja. Isso produz uma nova cerejeira.

Por definição, o espaço classificante  $B(X)$  é o quociente do espaço das cerejeiras por meio dessa operação de corte de arestas totalmente crescidas e aplicação de  $m_k$  conforme explicado.

Eu *não* vou mostrar que o espaço de laços de  $B(X)$  tem o mesmo tipo de homotopia de  $X$ . Aqui tenho que admitir: as referências que eu dei contêm provas, mas definitivamente não são fáceis de ler ☹.

A referência mais acessível que eu conheço é este artigo <sup>135</sup>.



<sup>135</sup> E. Hoefel, M. Livernet, and J. Stasheff. [A<sub>∞</sub>-actions and recognition of relative loop spaces](#). *Topology Appl.*, 206:126–147, 2016.



Uma árvore trançada,  
ilustrando um operad?

# Operads

DEIXE-ME COMEÇAR ESTE CAPÍTULO COM UMA CITAÇÃO de Peter May <sup>136</sup>.

O nome ‘operad’ é uma palavra que eu me concentrei, passando uma semana pensando em nada mais. Além de soar bem, o nome deve trazer à mente tanto as operações quanto as mônadas. [...] O que eu não previ era o quão flexível seria a noção, quantos contextos matemáticos essencialmente diferentes existem em que desempenharia um papel natural, quantos modos filosoficamente diferentes poderiam ser explorados.

De acordo com [Wikipedia](#), outra razão para este nome é que a mãe de May era uma cantora de ópera. Como quase todos os conceitos que discutimos até agora, os operads “existiam” muito antes do seu nascimento <sup>137</sup>, melhor dizendo, antes do momento quando eles foram “batizados”... A definição de May é destinada a encapsular muitos tipos de “operações”, a maioria lembrando os enxertos de árvores que já encontramos neste passeio.

Sabemos que um grupo é muito mais do que um conjunto equipado com alguma aplicação de composição satisfazendo alguns axiomas. Grupos só existem através de suas representações como automorfismos de “algo”. Da mesma forma, os operads só existem por meio de suas representações e não vamos gastar muito tempo com as definições abstratas.

Um *operad* consiste de

- conjuntos  $\mathcal{O}_n$  para  $n \geq 0$  (pensados como operações  $n$ -árias).

<sup>136</sup> J. P. May. Operads, algebras and modules. In *Operads: Proceedings of Renaissance Conferences (Hartford, CT/Luminy, 1995)*, volume 202 of *Contemp. Math.*, pages 15–31. Amer. Math. Soc., Providence, RI, 1997.

<sup>137</sup> J. Stasheff. The pre-history of operads. In *Operads: Proceedings of Renaissance Conferences (Hartford, CT/Luminy, 1995)*, volume 202 of *Contemp. Math.*, pages 9–14. Amer. Math. Soc., Providence, RI, 1997.

Nesta palestra “[On teaching mathematics](#)”, V. Arnold critica fortemente a abordagem axiomática da teoria dos grupos, como é geralmente ensinado na França. E eu concordo plenamente com ele.

- algum elemento 1 em  $\mathcal{O}_1$  chamado de unidade,
- para todo  $n, k_1, \dots, k_n$ , uma *operação*, i.e. uma aplicação

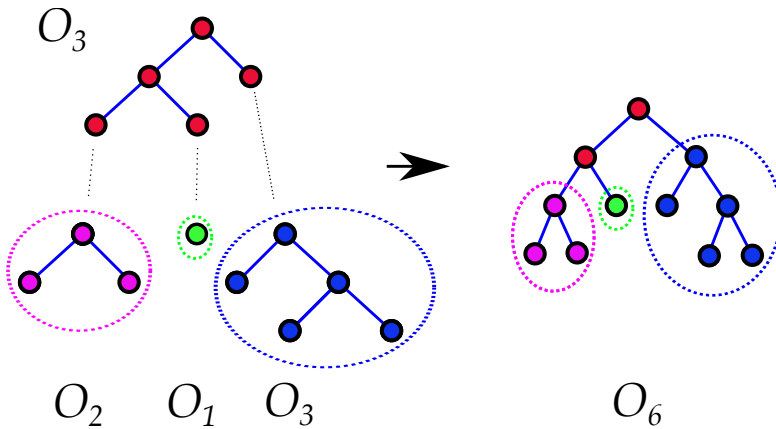
$$\mathcal{O}_n \times (\mathcal{O}_{k_1} \times \mathcal{O}_{k_2} \times \dots \times \mathcal{O}_{k_n}) \rightarrow \mathcal{O}_{k_1 + \dots + k_n}$$

satisfazendo ... alguns axiomas. Não quero escrever as fórmulas que expressam esses axiomas, pois eu mesmo não conseguiria ler as fórmulas que escreveria.

Eu prefiro dar primeiro um exemplo (que o leitor provavelmente já tenha adivinhado) antes de usar palavras para descrever os axiomas.

O exemplo é dado por árvores planares binárias: Denotemos por  $\mathcal{O}_n$  o conjunto das árvores planares binárias enraizadas com  $n$  folhas e chamamos de 1 a árvore com uma folha que é ao mesmo tempo a raiz. A operação de enxerto que já usamos várias vezes define o exemplo mais fácil de um operad.

Os cientistas da computação nos ensinaram que algumas bijeções computáveis  $\mathbb{N} \rightarrow \mathbb{N}$  são fáceis de avaliar, mas possuem inversos muito complicados. Escrever uma fórmula geralmente é fácil, mas as entender pode ser terrivelmente complicado.



Agora, quais são os axiomas de um operad geral? Enxertar 1 em alguma árvore não muda a árvore. Se selecionarmos duas folhas em alguma árvore e enxertarmos algo na primeira folha e outro algo na segunda folha, poderíamos também começar enxertando o outro algo na segunda folha e depois o algo na primeira. Na realidade, os axiomas operacionais são nada mais do que isso, substituindo a palavra "enxerto" pela operação operad.

Naturalmente, poderíamos também enxertar árvores planas enraizadas que não são necessariamente binárias. Por exemplo, poderíamos usar o que chamamos anteriormente de “árvores podadas”, ou seja, árvores planas enraizadas tais que cada vértice interior tem pelo menos dois filhos. Isso produz o operad de *Hiparco–Schroeder–Tamari–Stasheff*.

Para uma introdução de duas páginas sobre os operads, veja este artigo <sup>138</sup>. Para uma apresentação de 27 páginas, consulte este livro <sup>139</sup>. Para um livro de 634 páginas mais recente sobre o mesmo tópico, consulte esse enciclopédia <sup>140</sup>. Qual o meu querido leitor escolherá? Diga-me o que escolhes e sabereis quem és? ☺

Aqui está outro exemplo de uma operação simples: Escolha algum conjunto  $E$  e defina  $\mathcal{O}_n$  como o conjunto das aplicações  $E^n \rightarrow E$ . Neste exemplo,  $1$  é a identidade e as operações são simplesmente dadas por *substituição*. Se tivermos uma aplicação  $f : E^n \rightarrow E$  e  $n$  aplicações  $f_i : E^{k_i} \rightarrow E$ , pode-se substituir  $x_i$  por  $f_i$  em  $f(x_1, \dots, x_n)$  para obter aplicações  $E^{k_1 + \dots + k_n} \rightarrow E$ . Isso satisfaz os axiomas que não escrevemos. . . Denotamos este operad por  $\text{End}(E)$ .

Uma *álgebra* sobre um operad  $\mathcal{O}$  (também chamada de representação) é um conjunto  $E$  e uma homomorfismo de operad de  $\mathcal{O}$  para  $\text{End}(E)$ . Em outras palavras, cada elemento de  $\mathcal{O}_n$  define uma operação  $n$ -ária  $E^n \rightarrow E$  de uma “maneira compatível”.

Pode-se trabalhar em muitas categorias diferentes. Em vez de conjuntos, pode-se usar espaços topológicos, tipos de homotopia, espaços vetoriais etc.

Agora descrevo exemplos mais interessantes de operads. Outros virão nos capítulos seguintes.

## Permutações

Lembre-se que no início deste livro estudamos algumas questões combinatórias relativas ao “reconhecimento de padrões”. Isso se encaixa muito bem com o seguinte operad.

$\mathcal{O}_n$  é o conjunto de permutações de  $\{1, \dots, n\}$ . Eu escrevi “conjunto” e não “grupo” porque não vamos compor estas permu-

<sup>138</sup> J. Stasheff. [What is ... an operad?](#) *Notices Amer. Math. Soc.*, 51(6):630–631, 2004.

<sup>139</sup> F. Chapoton. [Operads and algebraic combinatorics of trees.](#) *Sém. Lothar. Combin.*, 58:Art. B58c, 27, 2007/08.

<sup>140</sup> J.-L. Loday and B. Vallette. *Algebraic operads*, volume 346 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2012.

tações. Em vez disso, pensamos em uma permutação como duas relações de ordem total em  $\{1, \dots, n\}$ .

Suponha que tenhamos algumas ordenações totais  $\sigma$  em  $\{1, \dots, n\}$  e  $\sigma_1, \dots, \sigma_n$  em  $\{1, \dots, k_1\}, \dots, \{1, \dots, k_n\}$ . Escreva  $\{1, \dots, k_1 + \dots + k_n\}$  como uma união disjunta de  $n$  intervalos consecutivos de tamanhos  $k_1, k_2, \dots, k_n$ . Ordene estes intervalos de acordo com  $\sigma$  e, dentro do  $i$ -ésimo intervalo, ordene os elementos de acordo com a ordem  $\sigma_i$ . Isso produz uma ordenação total natural em  $\{1, \dots, k_1 + \dots + k_n\}$ . Claramente essas “operações de enxerto” em ordenações satisfazem os axiomas de um operad.

No capítulo que discutimos permutações separáveis, notamos que qualquer árvore podada enraizada planar define uma permutação das folhas.

Em outras palavras, construímos um homomorfismo do operad de Hiparco–Schroeder–Tamari–Stasheff para o operad de Permutações. Nós também provamos que isso é uma injeção: pode-se reconstruir a árvore a partir da permutação.

### *O operad livre: Hiparco–Schroeder novamente*

Considere uma sequência de conjuntos  $(E_n)_{n \geq 1}$  e a definição do operad livre gerado por  $E_n$ . Nós temos que criar os conjuntos  $\mathcal{O}_n$  cujos elementos são produzidos sob as operações a partir dos elementos dos  $E_n$ . Uma vez que queremos um “operad livre”, todos esses novos elementos devem ser assumidos como diferentes, a menos que algum uso dos axiomas implique que eles sejam iguais. Não é difícil construir este operad livre.

Um elemento de  $\mathcal{O}_n$  é uma árvore planar enraizada com  $n$  folhas, de tal forma que cada vértice com  $i$  filhos está equipado com algum rótulo pertencente a  $E_i$ . As operações deste operad são novamente definidas através de enxertos.

Como exemplo, vamos considerar o caso em que todos os  $E_n$  são vazios, exceto  $E_2$  contendo um elemento. Então, o operad livre construído a partir um elemento “de grau 2” é o operad das árvores binárias enraizadas. Uma álgebra sobre este operad é apenas um conjunto com uma operação binária.

Como outro exemplo, vamos considerar o caso em que cada

Estou consciente do fato de que não é uma definição precisa, mas estou relutante em definir usando objetos iniciais em categorias.

$E_n$  contém um único elemento para  $n \geq 2$ . Obtemos árvores planas enraizadas, não necessariamente binárias, de modo que aos parênteses de Hiparco–Schroeder. Uma álgebra sobre este operad é apenas um conjunto com uma operação  $n$ -ária para cada  $n$ .

### *Simétrica e não simétrica*

Estritamente falando, lidamos até agora com operads que são chamados *não simétricos*. Em muitos casos, há ações dos grupos simétricos  $\Sigma_n$  em  $\mathcal{O}_n$  que são compatíveis com as operações da operad

$$\mathcal{O}_n \times (\mathcal{O}_{k_1} \times \mathcal{O}_{k_2} \times \cdots \times \mathcal{O}_{k_n}) \rightarrow \mathcal{O}_{k_1 + \cdots + k_n}.$$

Nesse caso, o operad é chamado *simétrico*. Frequentemente, o adjetivo “simétrico” é omitido. Aqui convido o leitor a determinar quais dos exemplos anteriores são simétricos.

Encontre a definição correta!

### *Pequenos cubos e Stasheff novamente*

Lembre-se de que interpretamos o associaedro  $K_n$  como espaços de árvores métricas, onde cada aresta interior tem um comprimento em  $[0, 1]$ . Enxertar essas árvores produz aplicações

$$K_n \times (K_{k_1} \times K_{k_2} \times \cdots \times K_{k_n}) \rightarrow K_{k_1 + \cdots + k_n}.$$

Em outras palavras, a verdadeira natureza da sequência de politopos  $K_n$  é a de um operad: o *operad de Stasheff*. Este é um operad topológico, pois  $\mathcal{O}_n$  agora é visto como um espaço topológico.

Uma álgebra sobre  $K_n$  é, por definição, uma família de aplicações de  $K_n \times X^n$  a  $X$ , que define um homomorfismo de operad. Claramente, todas as definições foram preparadas de tal maneira que a condição “homomorfismo de operad” coincide com a definição dos espaços de tipo  $A_\infty$ . Podemos, portanto, reafirmar o teorema de Stasheff como

**Teorema.** *Um  $H$ -espaço  $m_2 : X \times X \rightarrow X$  é homotopicamente equivalente a um espaço de laços se, e somente se, ele se estender como uma álgebra sobre o operad de Stasheff.*



Boardman e Vogt transformaram essa afirmação, escrevendo ainda de outra maneira e introduziram o “operad dos pequenos cubos”. Vamos escolher alguma dimensão  $d \geq 1$  e definir um operad topológico  $Cub_d$  da seguinte maneira:

- $Cub_d(n)$  é o espaço de  $n$ -uplos  $(c_1, c_2, \dots, c_n)$  dos mergulhos  $[0, 1]^d \rightarrow [0, 1]^d$  de tal forma que os interiores das imagens são disjuntos. O  $c_i$  deve ser uma aplicação afim e, mais precisamente, da forma  $c_i(x_1, \dots, x_d) = (\alpha_{1i}x_1 + \beta_{1i}, \dots, \alpha_{di}x_d + \beta_{di})$  ( $\alpha_{ij} > 0$ ).
- As operações deste operad

$$Cub_d(n) \times (Cub_d(k_1) \times \dots \times Cub_d(k_n)) \rightarrow Cub_d(k_1 + \dots + k_n)$$

são “óbvias”. Basta inserir os cubos como na figura.

Se  $(Y, \star)$  é um espaço topológico (na verdade, como de costume, um CW-complexo), o espaço dos  $d$ -laços  $\Omega^d(Y, \star)$  é o espaço das aplicações contínuas (com ponto de base) da esfera de dimensão  $d$  para  $Y$ . Pode-se também definir  $\Omega^d(Y, \star)$  como o espaço das aplicações  $[0, 1]^d \rightarrow Y$  que enviam a borda do cubo para o ponto base.

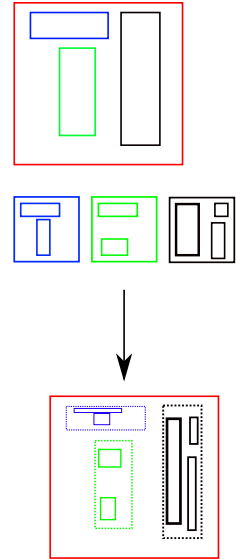
O espaço dos  $d$ -laços de  $Y$  aparece naturalmente como uma álgebra sobre o operad dos pequenos cubos. As operações do operad

$$Cub_d(n) \times \Omega^d(Y, \star)^n \rightarrow \Omega^d(Y, \star)$$

são “óbvias”. Dados  $n$  cubos pequenos  $(c_1, c_2, \dots, c_n)$  e  $n$  elementos  $\gamma_i : [0, 1]^d \rightarrow Y$  de  $\Omega^d(Y, \star)$ , pode-se definir uma aplicação de  $[0, 1]^d \rightarrow Y$ . Dentro da imagem de  $c_i$ , use a composição de  $c_i$  e  $\gamma_i$  e, de fora, use a função constante enviando tudo para o ponto base.

Em seu parecer do livro de Markl, Shinder e Stasheff sobre operads <sup>141</sup>, John Baez explica uma das motivações para os operads.

A maioria dos especialistas da teoria da homotopia alegremente venderiam suas almas pela capacidade de computar os grupos de homotopia de um espaço arbitrário.



<sup>141</sup> M. Markl, S. Shnider, and J. Stasheff. *Operads in algebra, topology and physics*, volume 96 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2002.

De fato, Boardman, Vogt<sup>142</sup> e May<sup>143</sup> generalizaram o teorema de Stasheff:

**Teorema.** *Se um espaço conexo  $X$  é uma álgebra sobre o operad dos pequenos cubos  $Cub_d$ , então, é homotopicamente equivalente ao espaço dos  $d$ -laços  $\Omega^d(Y)$  de algum espaço  $Y$ .*

### Mais operads

Acredito que meu jovem leitor tenha entendido que os operads ocorrem em quase toda parte na matemática, em um nível fundamental. Talvez (quem sabe) essa grande generalidade torne a teoria um pouco abstrata demais? Para terminar este capítulo conceitual, deixe-me dar mais alguns exemplos.

Insira uma fórmula em um programa matemático, por exemplo, o seguinte no *Mathematica*,

$$\text{Sqrt}[\text{Sin}[a + b + c]^2 + b^2 + c^2 + a/b]$$

$$\sqrt{\frac{a}{b} + b^2 + c^2 + \text{Sin}[a + b + c]^2}$$

Se você quiser saber como o seu computador “entende” essa fórmula, basta digitar o seguinte.

**TreeForm[Sqrt[Sin[a + b + c]^2 + b^2 + c^2 + a/b]]**

Você tem ... uma árvore.

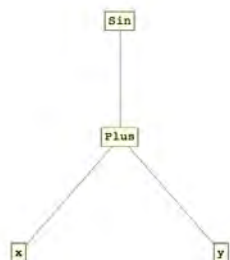
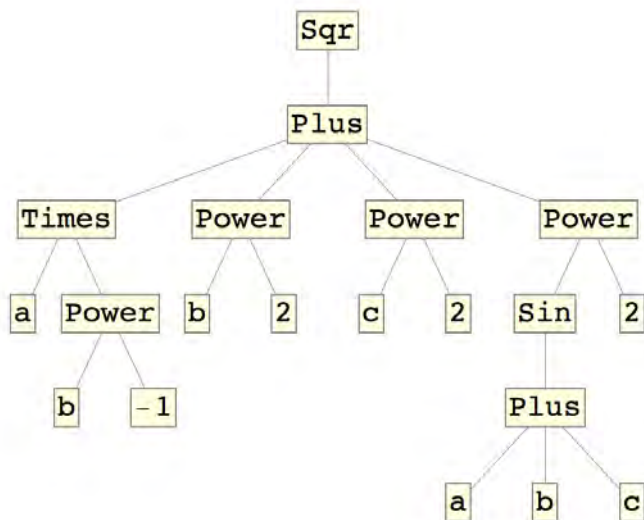
Os vértices são rotulados por operadores, o que poderiam ser  $n$ -ários para cada  $n$ . As folhas são “átomos”. Portanto, a linguagem de um software como o *Mathematica* é na verdade um operad.

Note, no entanto, que este operad não é livre. Por exemplo, o *TreeForm* de  $\text{sen}(x + y)$  é a árvore na margem. Se eu substituir  $a$  por  $x$  e  $\pi - a$  por  $y$  e se eu perguntar novamente pela árvore correspondente:

<sup>142</sup> J. M. Boardman and R. M. Vogt. [Homotopy-everything H-spaces](#). *Bull. Amer. Math. Soc.*, 74:1117–1122, 1968.

<sup>143</sup> J. P. May. *The geometry of iterated loop spaces*, volume 271 of *Lectures Notes in Mathematics*. Springer-Verlag, Berlin-New York, 1972.

Alguns matemáticos reclamam que o grupo livre é muito abstrato para merecer o apelido de grupo, sendo apenas um monte de palavras.

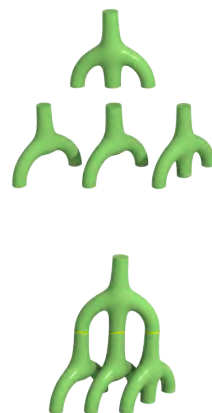


```
TreeForm[Sin[x + y] /. x -> a /. y -> Pi - a]
```

nós obtemos a árvore trivial com apenas um vértice rotulado com 0. O *Mathematica* “sabe” que  $\sin(\pi) = 0$ , então, a árvore entrou em colapso. Em outras palavras, o operad *Mathematica* é definido por geradores e relações, que são incorporados. O usuário tem permissão para adicionar regras locais e trabalha no operad quociente correspondente.

*Superfícies de Riemann* fornecem outra boa fonte de operads. Suponha que tenhamos duas superfícies Riemann  $\Sigma_1, \Sigma_2$  com borda. Note que uma superfície de Riemann é canonicamente orientada e isso induz uma orientação na borda. Suponha que tenhamos algum difeomorfismo  $f$  invertendo a orientação entre um círculo  $S_1$  contido na borda  $\partial\Sigma_1$  e um círculo  $S_2$  contido em  $\partial\Sigma_2$ . Pode-se colar as duas superfícies ao longo de  $f$  para produzir uma nova superfície (orientada!).

Acontece que essa nova superfície é canonicamente uma



Colando superfícies de Riemann.

superfície de Riemann, isto é, está equipada com uma estrutura de uma variedade holomorfa de dimensão 1. Isto é fácil de ver se  $f$  é analítica real, já que neste caso é possível ser estendido a um difeomorfismo holomorfo entre pequenos anéis, que podem ser usados para definir uma estrutura holomorfa na superfície colada. Pode-se também colar superfícies de Riemann ao longo de difeomorfismos não analíticos, mas isso não é importante em nosso contexto.

Usando esta operação de colagem, pode-se construir um operad. Um elemento de  $\mathcal{O}_n$  é uma classe de isomorfismo de uma superfície Riemann compacta com  $(n + 1)$  bordas rotuladas, sendo um deles chamado de entrada e os outros  $n$  saindo. Além disso, cada componente da borda é equipada com um difeomorfismo com o círculo. Colando superfícies ao longo de suas bordas, como no jogo de *Lego*, dá um exemplo de um operad.

Pode-se também aplicar vários *functores* a operads para produzir mais operads. Por exemplo, vamos olhar para o operad  $Cub_2$  de pequenos quadrados.  $Cub_2(n)$  tem o tipo de homotopia do espaço de  $n$  pontos distintos em um quadrado. Vamos considerar seu grupo fundamental  $PB_n$ . Isso é chamado de *grupo de trança pura*.

Um elemento de  $PB_n$  consiste em  $n$  pequenos quadrados, numerados em  $1, 2, \dots, n$  em um quadrado (ou um disco) que se movem ao longo de  $n$  laços sem se cruzarem. No final dos laços, os quadrados voltariam às suas posições iniciais e essa última propriedade é o que se entende por “puro”.

Usando grupos fundamentais, obtemos aplicações

$$PB_n \times (PB_{k_1} \times \dots \times PB_{k_n}) \rightarrow PB_{k_1 + \dots + k_n}$$

e nós temos um operad de grupo. Isso não é muito complicado. Um elemento de  $PB_n$  dá origem a  $n$  tubos em um cilindro  $[0, 1]^2 \times [0, 1]$ . A estrutura de operad consiste em inserir tubos em tubos.

Não há razão para nos limitarmos ao caso de dimensão 2 e ao grupo fundamental. Dado um espaço topológico  $X$  e um inteiro positivo  $n$ , pode-se considerar o *espaço de configuração*  $X^{[n]}$  definido como o espaço de  $n$ -uplas de *pontos distintos* em  $X$ . Se  $X$  é um disco de dimensão 2, pode-se mostrar que o recobrimento universal desse espaço é contrátil para que, do ponto de



Três pequenos quadrados seguem um laço.

†

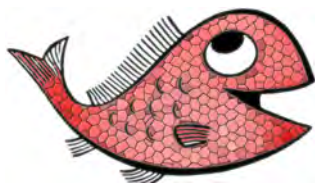
vista da homotopia, apenas seu grupo fundamental seja interessante: esse é o grupo de tranças puras que acabamos de discutir. No entanto, se  $X$  for uma bola de dimensão mais alta,  $X^{[n]}$  é simplesmente conexo e tentaremos descrever sua topologia.

Por exemplo, se  $X = \mathbb{R}^3$  e  $n = 2$ , a resposta é fácil: dois pontos distintos  $x, y$  no espaço são completamente definidos pelo seu ponto médio  $(x + y)/2$  e o vetor diferente de zero  $xy$ . Portanto,  $(\mathbb{R}^3)^{[2]}$  tem o mesmo tipo de homotopia que uma esfera de dimensão 2. A situação se torna mais complicada para  $n = 3$ : três corpos no espaço. . .

Uma boa abordagem é estudar a homologia ou cohomologia desses espaços não individualmente, para cada valor de  $n$ , mas globalmente: a cohomologia do operad dos pequenos cubos.

Peço ao meu leitor que leia o artigo (muito acessível) <sup>144</sup>, que poderá servir como porta de entrada para a teoria dos operads. Você aprenderá, por exemplo, que “A homologia do operad dos  $d$ -cubos é o grau  $d$  do *Poisson operad*” (seja lá o que isso signifique).

<sup>144</sup> D. P. Sinha. [The \(non-equivariant\) homology of the little disks operad](#). In *OPERADS 2009*, volume 26 of *Sémin. Congr.*, pages 253–279. Soc. Math. France, Paris, 2013.



O operad de Poisson (Peixe em francês).

September 16, 2011

# The Chicago Maroon

$\mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \Pi_n) \quad \Pi_n = \Sigma_n$

$\mathcal{O}_{\mathcal{G}}(\mathcal{Q}) = \mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \mathcal{O}(\mathcal{Q})) = \mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \tilde{\Sigma}_j)$

$\mathcal{O}_{\mathcal{G}}(\mathcal{Q}) \times \mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \mathcal{Q}^j) = \mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \tilde{\Sigma}_j) \times \mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \mathcal{Q}^j)$   
 $= \mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \tilde{\Sigma}_j \times \mathcal{Q}^j) \rightarrow \mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \mathcal{Q})$

$\mathbb{B}(\mathcal{G}, \Pi_n) \rightsquigarrow \mathbb{B}Cat_{\mathcal{G}}(\tilde{\mathcal{G}}, \Pi_n) ?$

$\mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \Pi \amalg \Pi') \leftarrow \mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \Pi) \amalg \mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \Pi')$

$\Pi$  and  $\Pi'$  groups considered as 1-object cate

$F: \tilde{\mathcal{G}} \rightarrow \Pi \amalg \Pi' \quad F(g) = *_1, F(h) = *_2$

the unique map  $g \rightarrow h$  has no nodes to go, so this is iso of Cate.

May 1985 Some remarks on equiv. bilie & chas, &  $Map(EG, B\Gamma) = E_{\Gamma}B\Gamma / \Gamma = B_{\mathcal{G}}\Gamma$

classifying space for  $\Gamma$ -bundles

$\Gamma = G \times \Pi \Rightarrow$

See  $(EG, B\Gamma) = Map(EG, B\Gamma)$

$\alpha: B(\Pi; \Gamma) \xrightarrow{\cong} Sec(EG, B\Gamma)$

if  $\Gamma$  is discrete. If also  $\Gamma = G \times \Pi$   $\mathbb{C}at_{\mathcal{G}}(\tilde{\mathcal{G}}, \Pi)$  ( $\Pi$  is

$\alpha: B(\Pi; \Gamma) \xrightarrow{\cong} Map(EG, B\Gamma)$

$Map(EG, E\Pi) \simeq E_{\mathcal{G}}(\Pi)$

$\downarrow$

$Map(EG, E\Pi) / \Pi \simeq B_{\mathcal{G}}(\Pi)$

$\downarrow$

$Map(EG, B\Pi) \quad \Pi$  discrete.

$B_{\mathcal{G}}\Pi = B(\Pi, G \times \Pi) \xrightarrow{\cong} Map(EG, B\Pi)$

$(\mathbb{B}Cat_{\mathcal{G}}(\tilde{\mathcal{G}}, \tilde{\Pi})) / \Pi \rightarrow \mathbb{B}Cat_{\mathcal{G}}(\tilde{\mathcal{G}}, \tilde{\Pi} / \Pi)$

$J(\mathcal{O}(\mathcal{Q}, X)) =$

maps of sets  $\mathcal{G} \rightarrow \Pi$

$z: \mathcal{F} \rightarrow \mathcal{F}' \quad f, f': \tilde{\mathcal{G}} \rightarrow$

$z_g: f(g) \rightarrow f'(g)$

$\forall g \rightarrow h \quad f(g) \rightarrow f(h)$

$\downarrow z_g \quad \downarrow z_h$

$\tilde{f}(g) \rightarrow \tilde{f}(h)$

$Hom(\mathcal{Q}, \tilde{\Pi}) = J(\mathcal{O}(\mathcal{Q}, \tilde{\Pi}))$

$E(\mathcal{G}, \tilde{\Sigma}_j) \leftarrow E(\mathcal{Q})$  auto

$Cat_{\mathcal{G}}(\tilde{\mathcal{G}}, \tilde{\Pi}) / \Pi \rightarrow$

$B(\mathcal{G}, \Pi_n)$

$E(\mathcal{G}, \Pi_n) / \Pi_n$

$(\mathbb{B}Cat_{\mathcal{G}}(\tilde{\mathcal{G}}, \tilde{\Pi})) / \Pi_n$

$Cat_{\mathcal{G}}(\tilde{\mathcal{G}}, \tilde{\Pi} / \Pi)$

$hg': g \rightarrow h \quad h'ih \rightarrow i$

$F: \tilde{\mathcal{G}} \rightarrow \Pi$

$J(\mathcal{O}(\mathcal{Q}, \tilde{X}))$

functors

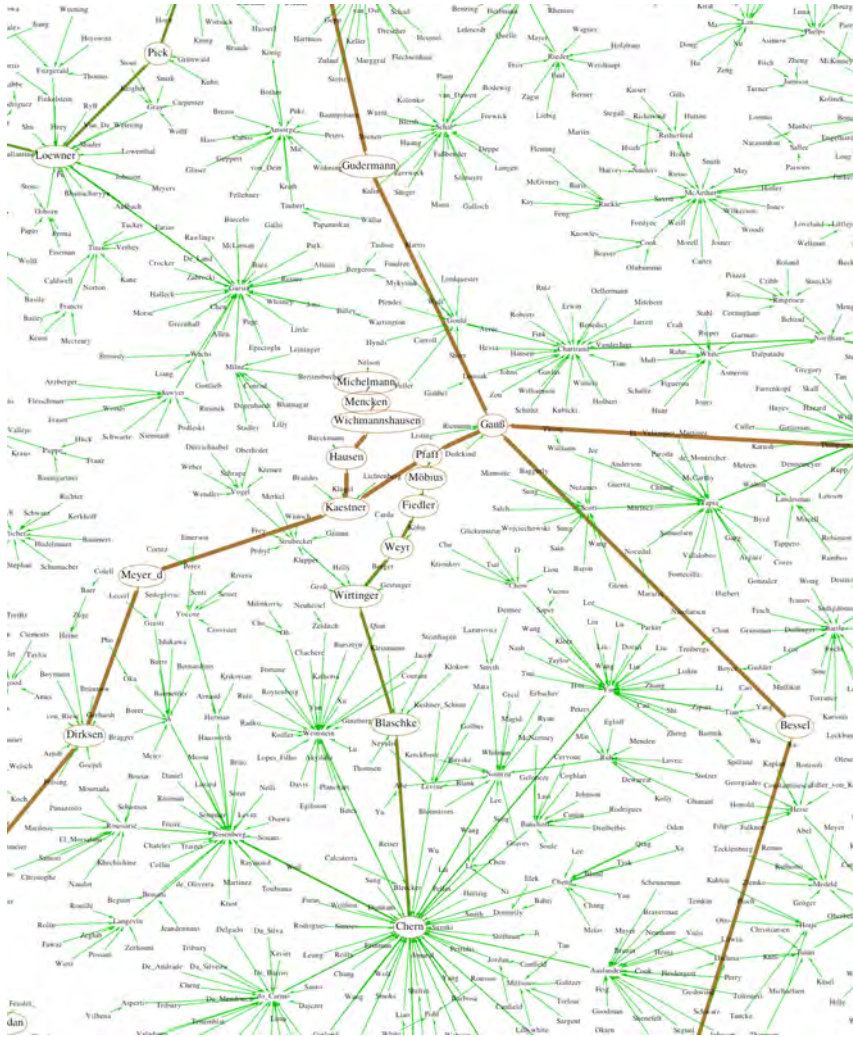
discrete case;

sets

one object

$Cat_{\mathcal{G}}(\tilde{\mathcal{G}})$

Peter May. +



Esta é uma visão local, centrada em Gauss, da “árvore dos matemáticos”, onde dois matemáticos estão conectados, se um é o orientador do outro. †

# Operads singulares

VOLTAMOS À NOSSA DISCUSSÃO INICIAL SOBRE A POSIÇÃO RELATIVA DOS GRÁFICOS DE UMA FAMÍLIA DE POLINÔMIOS.

## Os operads de polinômios reais

Sejam  $(P_1, P_2, \dots, P_n)$  e  $(Q_1, Q_2, \dots, Q_n)$  duas  $n$ -uplas de polinômios distintos em  $\mathbb{R}[x]$  que zeram na origem. Direi que são *topologicamente equivalentes* se para valores pequenos não nulos de  $x$ , as duas  $n$ -uplas de números reais  $(P_1(x), \dots, P_n(x))$  e  $(Q_1(x), \dots, Q_n(x))$  são ordenadas da mesma maneira. Seja  $Pol_{\mathbb{R}}(n)$  o conjunto (finito) de classes de equivalência de tais  $n$ -uplas. Vamos construir uma estrutura de operad muito simples em  $Pol_{\mathbb{R}}(n)$ .

Suponha que sejam dados

- $(P_1, P_2, \dots, P_n)$  (um representante de) um elemento de  $Pol_{\mathbb{R}}(n)$ ,
- para cada  $i = 1, \dots, n$ , um elemento de  $Pol_{\mathbb{R}}(k_i)$  dado por (a classe de)  $(P_{i,1}, \dots, P_{i,k_i})$ .

Nós queremos *enxertar* os  $P_j$  nos  $P_i$ . Apenas considere os polinômios  $k_1 + \dots + k_n$

$$P_i(x) + x^{2N} P_{i,j}(x) \quad (1 \leq i \leq n \text{ and } 1 \leq j \leq k_i)$$

na ordem lexicográfica de  $(i, j)$ , sendo  $N$  um inteiro suficientemente grande.

Algumas explicações podem ser úteis: O papel de  $x^{2N}$  é ter certeza de que os termos que são adicionados aos  $P_i$  são muito



Raízes ou ramos?  
(Rio Preguiças, Brasil).



Raízes e ramos.  
(Hong Kong).



menores que as diferenças  $P_i - P_j$  ( $i \neq j$ ). Para garantir essa propriedade, basta escolher  $2N$  maior que todos os  $v(P_i - P_j)$  ( $i \neq j$ ). Fixando  $i$ , os gráficos dos  $k_i$  polinômios  $P_i(x) + x^{2N}P_j(x)$  estão muito próximos do gráfico de  $P_i$ . O expoente *par*  $2N$  implica que, fixando  $i$ , a ordem entre  $P_i(x) + x^{2N}P_j(x)$  é a mesma que a ordem entre o  $P_{i;j}(x)$ . Topologicamente, os gráficos dos  $P_i$  foram transformados em algumas fatias finas nas quais os  $P_j$  foram inseridos.

Deve ficar claro que isso está bem definido e fornece uma estrutura de operad em  $Pol_{\mathbb{R}}(n)$ . Um operad simétrico, pois os polinômios  $(P_1, P_2, \dots, P_n)$  podem ser permutados.

Deve ficar igualmente claro, desde os primeiros capítulos deste livro, que esse operad está muito próximo do operad (não simétrico) das permutações separáveis.

### O operad dos polinômios complexos

Vamos brincar de um jogo similar com polinômios *complexos*. Se  $(P_1, P_2, \dots, P_n)$  é uma  $n$ -upla de polinômios distintos em  $\mathbb{C}[x]$ , que zeram na origem, olhamos para os seguintes laços em  $\mathbb{C}$  (por  $1 \leq i \leq n$ ):

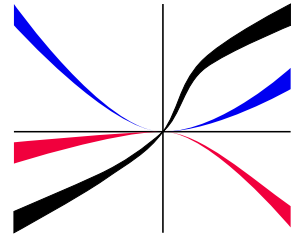
$$\gamma_i : \theta \in \mathbb{R}/2\pi\mathbb{Z} \mapsto P_i(\varepsilon \exp(\sqrt{-1}\theta)) \in \mathbb{C}.$$

Escolha  $\varepsilon$  muito pequeno. Então, para cada  $\theta$ , os  $n$  pontos  $\gamma_1(\theta), \gamma_2(\theta), \dots$  são distintos. Isso define um laço no espaço das  $n$ -uplas de pontos distintos no plano, ou seja, um elemento do grupo de tranças puras  $PB_n$ . Para ser exato, eu deveria falar de uma classe de conjugação de uma trança pura, uma vez que os pontos iniciais  $\gamma_i(0)$  poderiam estar em qualquer lugar e a definição do grupo de trança pura requer algum ponto de base. Essa classe de conjugação é independente da escolha do pequeno  $\varepsilon$ .

Dizemos que  $(P_1, P_2, \dots, P_n)$  e  $(Q_1, Q_2, \dots, Q_n)$  são *topologicamente equivalentes* se as tranças correspondentes forem conjugadas em  $PB_n$ . Denote por  $Pol_{\mathbb{C}}(n)$  o conjunto de classes de equivalência.

Lembre-se de que os grupos de tranças puras definem um dos nossos exemplos de operads, em que as operações consistem em

Lembre-se que  $v(P)$  indica a valoração de  $P$ .



Um desafio: você, leitor, consegue encontrar uma relação explícita entre os dois operads?

É equivalente à existência de um homeomorfismo local de  $\mathbb{C}^2$  na vizinhança da origem, enviando cada curva complexa  $P_i = 0$  para a curva complexa  $Q_i = 0$ ?

inserir tranças em vizinhanças tubulares de uma determinada trança. Isso sugere que  $Pol_{\mathbb{C}}(n)$  poderia ser um suboperad do  $PB_n$ . Este é realmente o caso como eu mostrarei agora:

Sabemos que  $\exp(-v(P_i - P_j))$  define uma distância ultramétrica em  $\{P_1, \dots, P_n\}$ , que pode ser codificada por uma árvore enraizada. Por sua vez, a raiz corresponde ao conjunto completo  $\{P_1, P_2, \dots, P_n\}$  e seus filhos são as classes de equivalência da relação  $v(P_i - P_j) \geq 2$  etc. Até chegarmos aos singletons  $\{P_i\}$  que são as folhas, rotuladas por  $1, 2, \dots, n$ .

Ressalto que existem duas diferenças principais em relação ao caso dos polinômios reais:

- Não há estrutura de ordem natural nos nós, de modo que nossa árvore não seja planar (afinal, a maioria das árvores na natureza não são planares).

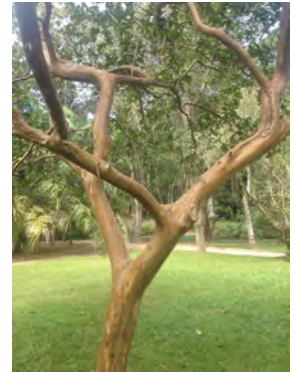
- No caso real, fizemos algumas podas na árvore. Fizemos isso porque, por exemplo, os pares  $(0, x)$  e  $(0, x^3)$  são topologicamente equivalentes *sobre os reais*:  $x$  e  $x^3$  têm os mesmos sinais. Mas isso não é mais verdade no domínio complexo: a trança associada a  $(0, x^3)$  gira 3 vezes ao contrário de  $(0, x)$  que gira apenas uma vez, pois  $x$  descreve a borda de um pequeno disco centrado em 0.

O seguinte resultado (elementar) fornece uma descrição precisa de  $Pol_{\mathbb{C}}(n)$ .

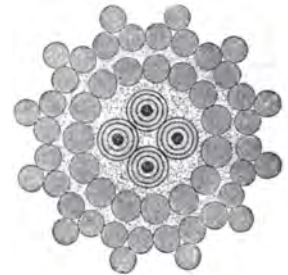
**Teorema.** *Duas  $n$ -uplas de polinômios complexos  $(P_1, P_2, \dots, P_n)$  e  $(Q_1, Q_2, \dots, Q_n)$  são topologicamente equivalentes se, e somente se, eles definirem a mesma árvore enraizada ou equivalentemente se  $v(P_i - P_j) = v(Q_i - Q_j)$  para todos os  $i, j$ .*

Inicialmente vamos mostrar que a árvore determina a trança: Comece pela raiz e desça  $p$  arestas até chegar ao primeiro nó com  $q \geq 2$  filhos. Isso significa que os  $p$ -ésimos polinômios de Taylor de todos os  $P_i$  são todos iguais e que existem  $q$  polinômios  $(p + 1)$ -ésimos polinômios de Taylor distintos.

Se  $t_p$  denota este  $p$ -ésimo polinômio de Taylor comum, substituímos  $t_p$  de todos os  $P_i$  sem alterar a trança correspondente. Portanto,  $P_1, P_2, \dots, P_n$  vêm em  $q$  grupos, onde colocamos no mesmo grupo dois  $P_i$  com o mesmo polinômio de Taylor de



Uma árvore não planar no Jardim Botânico, Rio de Janeiro.



Um cabo telegráfico submarino de 1858 ligando a Inglaterra e a Holanda e sua seção transversal.

ordem  $(p + 1)$ , da forma  $a_{\nu}x^{p+1}$ , para  $\nu = 1, \dots, q$ , sendo todos os  $a_{\nu}$  números complexos distintos. Quando  $x$  descreve o círculo de raio  $\varepsilon$ , esses  $q$  pontos  $a_{\nu}x^{p+1}$  descrevem círculos pequenos sob a mesma rotação rígida e giram em  $p + 1$  voltas completas. Essa (classe de conjugação dessa) trança em  $PB_q$  depende apenas de  $q$  e  $p$  e não da posição dos  $a_{\nu}$ : temos  $q$  pontos que simplesmente giram em bloco e descrevem  $p + 1$  ciclos completos.

Em torno de cada  $a_{\nu}\varepsilon^{p+1} \exp(\sqrt{-1}(p + 1)\theta)$  construa um pequeno disco que contenha todos os  $P_i(\varepsilon \exp(\sqrt{-1}\theta))$  com  $(p + 1)$ -ésimo Polinômio de Taylor  $a_{\nu}x^{p+1}$ . Continue o processo dentro de cada um desses discos, dividindo novamente cada grupo de acordo com os polinômios de Taylor de ordem superior.

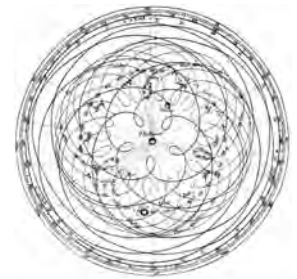
A conclusão é que a trança é como um sistema solar, movendo-se ao longo de epiciclos, à la Hiparco . Ela consiste de um grupo de  $q$  discos pequenos que giram rigidamente em  $p + 1$  voltas . Dentro de cada disco, a imagem é semelhante. E assim por diante, até chegarmos a  $\gamma_i(\theta)$ . Todos esses números  $p, q$ , para todos esses discos, são dados pela combinatória da árvore, de modo que a trança (sempre a menos de uma conjugação) é de fato determinada pela árvore, de uma maneira muito concreta.

Explico agora como construir a árvore a partir de uma trança.

Escolha dois inteiros  $1 \leq i < j \leq n$ . Uma trança pura em  $PB_n$  é uma classe de homotopia de um laço de  $n$  pontos distintos  $(x_1, \dots, x_n)$  no plano. Esqueça todos os pontos, exceto  $x_i$  e  $x_j$ : isso define algum homomorfismo de  $PB_n$  a  $PB_2$ , para cada  $i, j$ . A estrutura de  $PB_2$  é muito simples: é isomorfo a  $\mathbb{Z}$ . Quando dois pontos distintos se movem e retornam à sua posição original, o vetor  $x_i - x_j$  é um laço no plano perfurado e tem um índice relativo a 0. Este é o isomorfismo entre  $PB_2$  e  $\mathbb{Z}$ . Isso define homomorfismos  $n(n - 1)/2 lk_{i,j} : PB_n \rightarrow \mathbb{Z}$ , que simplesmente expressam o número de voltas de  $x_i - x_j$ . Duas tranças conjugadas têm as mesmas imagens por  $lk_{i,j}$ .

Vamos voltar à nossa trança definida pelos  $\gamma_i$ . Avaliando  $lk_{i,j}$  nesta trança, estamos contando o número de voltas de  $P_i(x) - P_j(x)$  quando  $x$  ultrapassa a borda de um pequeno disco centrado na origem. Essa é obviamente a valoração  $v(P_i - P_j)$ . Portanto, as valorações  $v(P_i - P_j)$  podem ser lidas a partir da

Encore lui !



O movimento aparente do Sol, Mercúrio e Vênus visto da Terra.

classe de conjugação da trança, conforme desejado. □

Reciprocamente, dada uma árvore enraizada com  $n$  folhas, é fácil construir  $n$  polinômios cuja árvore associada é aquela dada.

O fato é que nossas árvores não são planares. No entanto, as árvores enraizadas não planares podem ser enxertadas assim que suas folhas forem rotuladas de 1 a  $n$ . Dada uma árvore enraizada  $T$  com  $n$  folhas, rotuladas  $1, \dots, n$  e  $n$  árvores  $T_1, \dots, T_n$ , podemos enxertar  $T_i$  na folha numerada  $i$  de  $T$ . Isso define a operação de *árvores enraizadas não planares rotuladas*.

Portanto, essa operação de enxerto define uma estrutura natural de operad em  $Pol_{\mathbb{C}}(n)$ . Além disso, a discussão anterior mostra que  $Pol_{\mathbb{C}}(n)$  é um suboperad do operad de tranças puras.

Podemos descrever essa estrutura em fórmulas, diretamente em termos de polinômios. Dado algumas  $n$ -tuplas  $(P_1, \dots, P_n)$ , defina

$$\delta_i = \max_{j \neq i} v(P_i - P_j).$$

A operação de  $(P_1, \dots, P_n)$  em uma família  $(P_{i,j})$  ( $1 \leq i \leq n$  and  $1 \leq j \leq k_i$ ) é definida como a  $(k_1 + \dots + k_n)$ -upla de polinômios (em ordem lexicográfica)

$$P_i(x) + x^{\delta_i} P_{i,j}(x) \quad (1 \leq i \leq n \text{ and } 1 \leq j \leq k_i).$$

Em resumo  $Pol_{\mathbb{C}}(n)$  é isomorfo a um operad de árvores enraizadas não-planares rotuladas e aparece como um suboperad do operad  $PB_n$ .

### Um operad associado às curvas complexas singulares

Para dois polinômios complexos  $P_i(x)$  e  $P_j(x)$  anulando-se em 0, a valuação  $v(P_i - P_j)$  também é chamada de *multiplicidade* de interseção das duas curvas suaves  $y = P_i(x)$  e  $y = P_j(x)$  na origem. Não há surpresa nessa terminologia, já que essa é, de fato, a multiplicidade, no sentido usual da palavra, da raiz 0 de  $(P_j - P_i)(x) = 0$ . O parágrafo anterior também pode ser reinterpretado da seguinte maneira:

As curvas  $y - P_i(x) = 0$  são suaves em  $\mathbb{C}^2$ . Elas se cruzam transversalmente uma pequena esfera  $S^3_\epsilon$  em um nó trivial. Nós



Uma trança pura com três fios.

Este é um operad *simétrico*, uma vez que os polinômios podem ser permutados.

$\delta_i$  é o nível no qual a folha  $P_i$  está anexada à árvore.

Verifique se esta definição realiza o enxerto das árvores marcadas associadas, conforme desejado.

olhamos para estes nós na esfera quadrada  $\max(|x|, |y|) = \varepsilon$  e nós os denotamos por  $\gamma_i$ . Acontece que o número de entrelace de  $\gamma_i$  e  $\gamma_j$  nada mais é do que sua multiplicidade de interseção. Esta é uma interpretação topológica da multiplicidade.

O número de entrelace de dois nós orientados  $k_1, k_2$  em  $S^3$  é definido da seguinte maneira: Escolha uma superfície orientada mergulhada cuja borda orientada seja  $k_1$  e conte o número de interseção (algébrico) dessa superfície com  $k_2$ .

Em nosso caso simples, considere o link da curva lisa  $y - P_i(x) = 0$  na 3-esfera como a borda de uma de suas fibras Milnor, sendo  $y - P_i(x)$  real e positivo. Para calcular o número de entrelace das duas curvas  $y - P_i(x) = 0$  e  $y - P_j(x) = 0$  na 3-esfera, temos que contar a interseção (algébrica) do nó  $y - P_j(x) = 0$  com a fibra de Milnor  $y - P_i(x) \in \mathbb{R}_+$ . Agora  $(P_j - P_i)(x) = ax^{v(P_i - P_j)} \dots$ , de forma que o segundo nó cruze a fibra Milnor exatamente  $v(P_i - P_j)$  vezes (deve-se checar se as interseções são positivas). Segue que o número de entrelace entre  $\gamma_i$  e  $\gamma_j$  é de fato a multiplicidade  $v(P_i - P_j)$ .

Acontece que a maior parte do que foi dito é verdadeiro para ramos de curvas não suaves. Seja  $F(x, y) = 0$  uma curva analítica complexa singular (assumida reduzida), admitindo  $n$  ramos. Escreva  $F = F_1 \dots F_n$ , uma decomposição em fatores irredutíveis. Cada ramo  $F_i = 0$  admite uma parametrização de Puiseux:

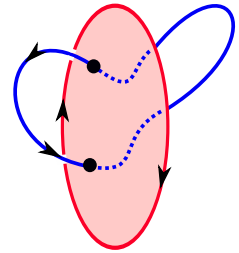
$$t \in \mathbb{C} \mapsto (t^{m_i}, g_i(t)).$$

O inteiro  $m_i$  é a ordem do ramo. A interseção de cada ramo com uma pequena esfera é um nó  $k_i$ , o que não é trivial se o ramo não for suave. O número de entrelace  $m_{ij}$  de  $k_i$  e  $k_j$  é a multiplicidade de interseção dos dois ramos. O leitor pode ver isso como uma definição, se sua mente for topologicamente orientada. Se ela preferir álgebra, ela poderia proceder da seguinte maneira: Insira a parametrização de um ramo na equação do outro e observe a multiplicidade do zero  $t = 0$  de

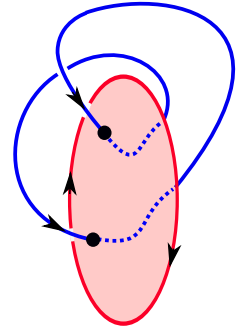
$$t \in \mathbb{C} \mapsto F_j(t^{m_i}, g_i(t)).$$

Observe a analogia das duas definições: “multiplicidade algébrica de interseção de dois ramos” e “número de entrelace”

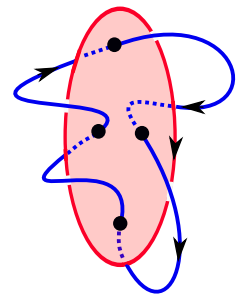
Os dois últimos capítulos deste livro tratam do número de entrelace em mais detalhes.



número de entrelace =  $1 - 1 = 0$ .



número de entrelace =  $1 + 1 = 2$ .



número de entrelace = 4.

Você pode provar que um ramo define um nó trivial se, e somente se, a ramificação for suave?

de dois nós. Ambos parecem assimétricos, mas na verdade são simétricos.

Acontece que as multiplicidades da interseção  $m_{ij}$  entre os ramos sempre têm as propriedades importantes que notamos para curvas suaves.

- $m_{ij}$  é um inteiro positivo. Isso é fácil.
- Os  $m_{ij}/m_i m_j$  satisfazem algum tipo de *desigualdade ultramétrica*. Em outras palavras, para cada  $\varepsilon > 0$ , a relação  $m_{ij}/m_i m_j \geq \varepsilon$  é uma relação de equivalência em  $\{1, \dots, n\}$  para que possamos construir um árvore, como antes (exceto quando os comprimentos das arestas não são inteiros, mas números racionais).

O segundo item é devido a Płoski <sup>145</sup> em 1985. Veja o artigo <sup>146</sup> para uma apresentação e aplicação modernas e este outro <sup>147</sup> para uma ampla generalização.

Podemos agora construir um *operad de curvas singulares complexas*. Na verdade, como antes, precisamos rotular os ramos de nossa curva  $y = P_1(x), \dots, y = P_n(x)$  (sendo os  $P_i$  agora as séries de Puiseux  $g_i(x^{\frac{1}{m_i}})$ ). Em outras palavras, consideramos o conjunto de  $n$ -tuplas de séries distintas do Puiseux, denotadas como  $Curv_n$ , como *curvas complexas com ramificações rotuladas*. Nós determinamos uma estrutura de operad nos  $Curv_n$  exatamente como fizemos com os polinômios. Defina números positivos racionais  $\delta_1, \dots, \delta_n$ , associados a  $(P_1, \dots, P_n) \in Curv_n$  por

$$\delta_i = \max_{j \neq i} v(P_i - P_j)$$

sendo  $v$  agora o menor expoente racional de um termo não trivial de uma série de Puiseux.

Dados os elementos de  $Curv_{k_1}, \dots, Curv_{k_n}$ , definidos pela série de Puiseux  $P_{i,j}$  para  $1 \leq i \leq n$  e  $1 \leq j \leq k_i$ , podemos construir o elemento de  $Curv_{k_1 + \dots + k_n}$  como a  $(k_1 + \dots + k_n)$ -tupla

$$P_i(x) + x^{\delta_i} P_{i,j}(x) \quad (1 \leq i \leq n \text{ and } 1 \leq j \leq k_i)$$

em ordem lexicográfica. Isso define uma operação (simétrica).

*Exercício:*

<sup>145</sup> A. Płoski. Remarque sur la multiplicité d'intersection des branches planes. *Bull. Polish Acad. Sci. Math.*, 33(11–12):601–605 (1985), 1985.

<sup>146</sup> I. Abío, M. Alberich-Carramiñana, and V. González-Alonso. [The ultrametric space of plane branches](#). *Comm. Algebra*, 39(11):4206–4220, 2011.

<sup>147</sup> E. García Barroso, P. González Pérez, and P. Popescu-Pampu. [Ultrametric spaces of branches on arborescent singularities](#). Preprint no arXiv, 2016.

A interseção de  $F(x, y) = 0$  com  $|x| = \varepsilon$  define uma trança, que não é mais uma trança pura. Na verdade, cada ramo intersecta  $x = \varepsilon$  um número de vezes igual ao seu próprio  $m_i$ . Ao contrário do caso anterior, o nó definido por cada ramo não é trivial.

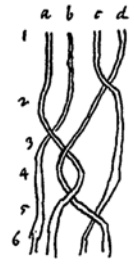
Desafio: encontre uma versão topológica do operad anterior. Em outras palavras, encontre uma relação de equivalência natural entre as curvas  $F(x, y) = 0$  (com ramificações rotuladas) em termos das tranças associadas que elas definem. Essa relação de equivalência deve ser tal que o operad age sobre classes de equivalência.



Uma trança “não pura” com 3 fios .

+

*A*




Veränderung der Ordnung

a	1	1	2+i	3+i	2+2i	2+2i
b	2	2	1	4	1	3+i
c	3	4	2	1	4	1
d	4	3+i	3+i	2+2i	3+2i	4+3i

Es kommt dazu den Index der Verkettung  
 so als Äquivalent von Werten anzustellen daß  
 man nicht welche Teile einander destruiert.

Mehrschleifen sind gegeben die halben Umwicklungen  
 einer Linie um die andere nach einer bestimmten Richtung.  
 Dies anzugeben.  
 In jedem Beispiel  
 entsprechend  
 ab, ab

Man braucht nur in jeder Linie zu zählen wie oft + und - wechselt



*A*

Um manuscrito inédito de Gauss, no qual ele inicia o estudo topológico de tranças.





Gauss em 1840.

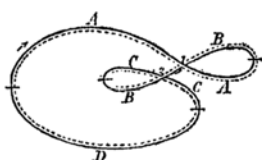
†

# Gauss está de volta: curvas no plano

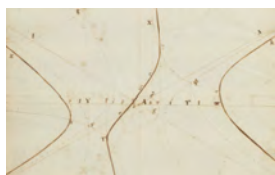
MUITOS GRANDES MATEMÁTICOS, PASSADOS OU PRESENTES, DESFRUTARAM OU GOSTARAM DE DESENHAR CURVAS. Como teste, eu coloco algumas imagens, e meu leitor deve adivinhar seus autores.



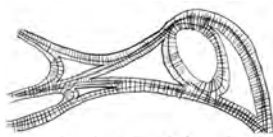
Desvio! Eu gosto deste capítulo, mas é completamente independente do resto do livro.



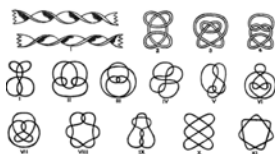
1



2



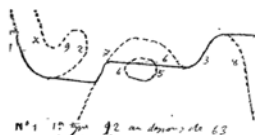
3



4



5



6

Solução: próxima página.

A **Figura 1** foi desenhada por *Gauss* em seu *Nachlass*, em torno do problema discutido neste capítulo.

A **Figura 2** é uma curva cúbica desenhada por *Newton* em seu *Enumeratio Linearum tertii Ordinis* (1667). Na terminologia moderna, *Newton* tenta classificar essas curvas algébricas sob transformações projetivas. Ele fala da “geração de curvas pelas sombras”, que é uma maravilhosa definição de geometria projetiva.

A **Figura 3** é um desenho de *Thurston* em suas *Notes on the topology and geometry of 3-manifolds* (1980), que têm sido muito influentes. Essa figura representa um *train-track*, um conceito importante que é útil na descrição da borda do espaço de Teichmüller.

A **Figura 4** é uma projeção de um nó feita por *Tait* (1884) em uma série de artigos que representam a primeira tentativa seria de classificar os nós.

A **Figura 5** é de uma nota suíça em homenagem a *Euler*. Em sua *Lettres à une princesse d’Allemagne*, ele explicou a lógica *booleana* a uma jovem princesa alemã, usando diagramas limitados por elipses. Ele descobriu todas as possíveis configurações de interseções de quatro elipses e os muitos casos possíveis estão no fundo da nota.

A **Figura 6** é de *Poincaré* em uma tentativa de provar seu “último teorema”, sobre o número de pontos fixos de um difeomorfismo do anel que preserva a área. Na verdade, ele não provou isso e faleceu alguns meses depois. O teorema foi provado logo após a morte de *Poincaré* por *Birkhoff*.



Assinatura de  
Gauss em 1794  
(ele tinha 17 anos).

+

## Palavras de Gauss

Neste capítulo, revisaremos uma bela pergunta levantada por Gauss sobre curvas no plano. Lembre-se de que sua primeira prova do teorema fundamental da álgebra se baseia no comportamento qualitativo das curvas dentro de um disco.

O volume 8 de *Gauss Werke* contém algumas páginas <sup>148</sup>(página 272 e 282–286) sobre curvas imersas no plano. Devemos ter cuidado: estas são as chamadas *Nachlasse*, notas que permaneceram inéditas durante a vida de Gauss. Não esqueça o seu lema *Pauca sed matura* (Poucos, mas maduros). Essas páginas não devem ser consideradas como uma publicação real, mas como um rascunho privado.

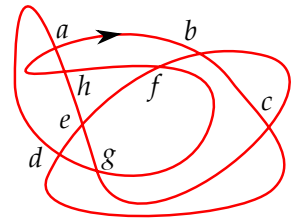
Desenhe uma curva imersa fechada e orientada genérica  $i : S^1 \rightarrow \mathbb{R}^2$  no plano. Genérico significa que pontos múltiplos são apenas pontos duplos com duas tangentes diferentes. Rotule os pontos duplos por  $n$  símbolos  $a_1, \dots, a_n$  (que Gauss chamou de *Knoten*). Indo ao redor da curva, nós passamos duas vezes em cada um desses  $a_i$ , definindo uma palavra cíclica de comprimento  $2n$  na qual cada  $a_i$  aparece duas vezes. *A curva fechada, portanto, define um diagrama de cordas com uma corda para cada  $a_i$ .*

A questão levantada por Gauss é reconhecer quais diagramas de cordas provém de alguma curva planar. Ele primeiro listou todas as possibilidades para  $n \leq 8$  à mão! Então, ele encontrou uma condição *necessária*. Escrevendo a palavra em um círculo, entre duas ocorrências de um  $a_i$ , deve haver um número *par* de letras. No exemplo na margem, a corda  $d$  decompõe o círculo em dois intervalos contendo respectivamente 6 e 8 letras.

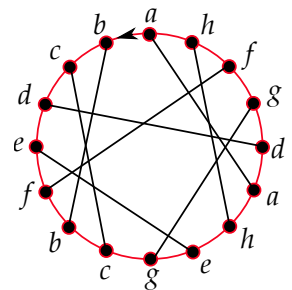
Alguns autores modernos afirmam que esta era uma conjectura de Gauss e que ele não podia provar isso. Que falta de respeito! Parece claro, para mim, que Gauss podia provar isso e não quis perder tempo a escrevê-lo em seu caderno particular.

Um dos primeiros teoremas de topologia, conhecido por Gauss em seu doutorado, como já citamos, é que duas curvas fechadas no plano que se cruzam transversalmente têm um *número par de pontos de interseção*. Uma das provas possíveis é mover a primeira curva por um caminho genérico de translações,

<sup>148</sup> C. F. Gauß. *Werke*. Band VIII. Georg Olms Verlag, Hildesheim, 1973. Reimpressão do original de 1900.



*abcdefghijklgh*



de modo que, no final do movimento, não haja mais ponto de interseção. Examinando como os pontos de interseção aparecem ou desaparecem em situações genéricas, não é difícil ver que, genericamente, os pontos aparecem ou desaparecem em pares.

Um fato intimamente relacionado foi notado por todos os alunos que desenhavam rabiscos, durante as aulas chatas de matemática. Se você desenhar uma curva fechada genérica no plano, as componentes conexas do complemento podem ser coloridas em preto e branco, como em um tabuleiro de xadrez. Apenas use branco para a componente no infinito e para alguma outra componente a conecte ao infinito por algum arco genérico, sendo cor branca se este arco cruzar a curva fechada em um número par de pontos, e preta de outra forma. Isso é coerente por causa da observação anterior.

Agora, a condição necessária de Gauss é um corolário fácil: Qualquer corda decompõe o círculo em dois intervalos  $I_1, I_2$  que definem duas curvas fechadas no plano, digamos  $\gamma_1$  e  $\gamma_2$ , começando do mesmo ponto. Mova levemente essas duas curvas para torná-las transversais. O número de letras em  $I_1$  é igual ao número de pontos de interseção entre  $\gamma_1$  e  $\gamma_2$ , mais o dobro do número de autointerseções de  $\gamma_1$ . Portanto, este número é par. Na imagem na margem, um dos dois laços de  $b$  para  $a$  é ligeiramente desviado e mostrado como um laço azul pontilhado.

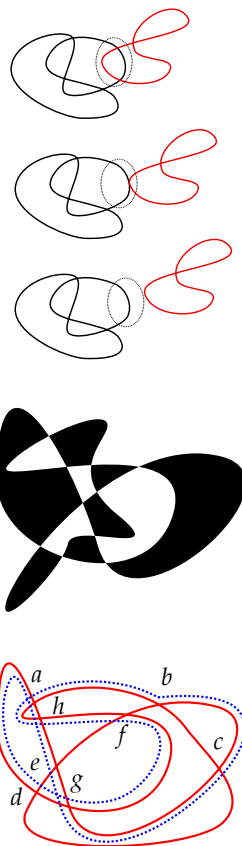
Essa condição necessária não é suficiente, como Gauss sabia muito bem.

### *Palavras de Gauss com sinais*

O problema de Gauss foi resolvido muitas vezes, de muitas maneiras diferentes, em diferentes comunidades matemáticas, basicamente topológicas ou combinatórias. Isto está em sintonia com a citação de Poincaré:

[...] il n'y a plus des problèmes résolus et d'autres qui ne le sont pas, il y a seulement des problèmes plus ou moins résolus.

Eu descreverei apenas duas soluções.



Deixe-me apresentar primeiro uma solução de um problema mais simples, usando um argumento topológico, misturando <sup>149</sup>, <sup>150</sup> e <sup>151</sup>. Se o plano é orientado, cada ponto duplo da nossa curva genérica define dois vetores tangentes, de modo que um deles é “o primeiro” e o outro é “o segundo”. Indo ao redor da curva, quando encontramos um ponto duplo, marcamos um sinal de + se passarmos pelo primeiro e um sinal de – caso contrário. Assim, a palavra cíclica é agora decorada com sinais, ou expoentes,  $\pm 1$ . Cada letra ocorre duas vezes, com sinais diferentes. Em termos equivalentes, poderíamos pensar em um *diagrama de cordas orientado* em que cada corda vai, digamos, de seu início + para seu final –. O *problema de Gauss com sinais* (que Gauss *não* estudou) é o seguinte. Dada uma *palavra com sinais*, pode-se decidir se ela está associada a alguma curva genérica no plano?

Para cada símbolo  $a_1, a_2, \dots, a_n$ , desenhe um cruzamento como na margem. Cada cruzamento tem dois lados entrando e dois lados saindo. Cada cruzamento tem dois braços, rotulados + e –.

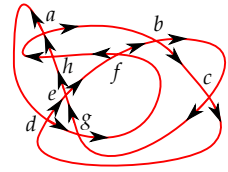
Uma palavra cíclica com sinais  $w$  define unicamente uma maneira de colar cada lado de saída de cada cruzamento para algum lado de entrada de algum outro cruzamento. O resultado dessa operação de colagem é alguma superfície orientada  $S$  com borda contendo uma curva orientada imersa  $\gamma$ . Indo em torno desta curva, lemos precisamente a palavra com sinais  $w$ . Se  $S$  for planar, pode ser mergulhado no plano, nosso problema seria resolvido desde que construamos uma curva no plano. Inversamente, se a palavra vem de uma curva imersa no plano, alguma vizinhança de sua imagem é claramente uma união de cruzamentos, montada como em  $S$ .

*Portanto  $w$  é realizável por alguma curva planar imersa se, e somente se,  $S$  for planar.*

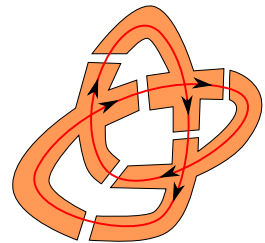
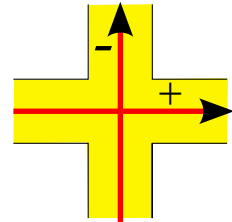
Para o restante deste capítulo, é necessária alguma familiaridade com a teoria básica da homologia das superfícies. Esta é uma boa oportunidade para recomendar o livro visual de Fomenko <sup>152</sup>. Ao abrir este livro, meu leitor entenderá imediatamente por que gosto disso. O livro mais padrão e Massey é muito acessível <sup>153</sup>.

<sup>150</sup> G. Cairns and D. M. Elton. The planarity problem for signed Gauss words. *J. Knot Theory Ramifications*, 2(4):359–367, 1993.

<sup>151</sup> G. Cairns and D. M. Elton. The planarity problem. II. *J. Knot Theory Ramifications*, 5(2):137–144, 1996.



$$a^+ b^+ c^+ d^+ e^+ f^+ b^- c^- g^- e^- h^+ a^- d^+ g^+ f^- h^-$$



<sup>152</sup> A. Fomenko. *Visual geometry and topology*. Springer-Verlag, Berlin, 1994. Traduzido do Russo por Marianna V. Tsaplina.

<sup>153</sup> W. S. Massey. *A basic course in algebraic topology*, volume 127 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1991.

Seja  $k$  o número de componentes da borda de  $S$ . Seja  $\hat{S}$  a superfície obtida de  $S$  colando um disco em cada um de seus componentes de fronteira. A superfície  $S$  tem o tipo de homotopia de um grafo com  $n$  vértices e  $2n$  arestas. Assim, a característica de Euler–Poincaré de  $\hat{S}$  é  $n - 2n + k$ . Superfícies orientadas compactas sem borda são classificadas pela característica de Euler–Poincaré. Como um corolário,  $S$  é planar se, e somente se,  $k - n = 2$ .

Isso dá um algoritmo muito simples para decidir se  $w$  é realizável: Comece com  $w$ , cole os cruzamentos e conte o número de componentes da borda, que deve ser  $n + 2$ . Isso é essencialmente devido à Carter <sup>154</sup>.

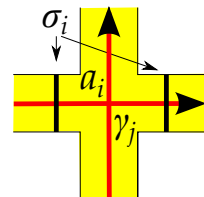
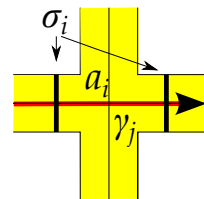
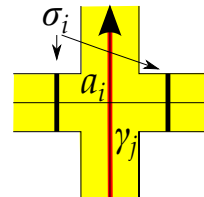
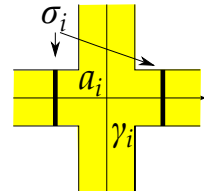
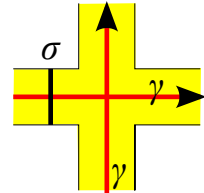
Agora apresento outro ponto de vista: Segue-se da classificação de superfícies orientadas compactas com borda - uma superfície é planar se, e somente se, quaisquer duas curvas transversais fechadas se cruzarem em um número par de pontos. De fato, se o gênero de uma superfície for  $\geq 1$ , ela contém um toro perfurado que contém duas curvas que se intersectam exatamente em um ponto. Portanto, para verificar se o gênero de  $S$  é 0, é suficiente encontrar uma base de sua homologia  $H_1(S; \mathbb{Z}/2\mathbb{Z})$  módulo 2, e calcular a forma de interseção.

Existe uma maneira fácil de encontrar uma base de  $H_1(S; \mathbb{Z}/2\mathbb{Z})$ . Como observação preliminar, note que um par ordenado de pontos  $(a^+, a^-)$ , no círculo orientado, define um intervalo aberto no círculo, enquanto viajamos na direção positiva de  $a^+$  para  $a^-$ . Eu direi que os elementos desse intervalo são *entre*  $a^+$  e  $a^-$ . Tenha cuidado, no entanto: o intervalo é alterado para o seu complemento se os dois pontos forem permutados.

A curva original  $\gamma$  é desenhada em  $S$  e, portanto, define uma classe de homologia  $[\gamma]$ . Além disso, para cada  $i = 1, \dots, n$ , o intervalo de  $a_i^+$  a  $a_i^-$  no círculo define um laço  $\gamma_i$  em  $S$  e uma classe de homologia  $[\gamma_i]$  em  $H_1(S; \mathbb{Z}/2\mathbb{Z})$ . Note que quando  $\gamma_i$  entra em um cruzamento, ele não muda de direção com a exceção da cruz marcada com  $a_i$ , onde ela se modifica. Dito de forma diferente, a interseção de  $\gamma_i$  com um cruzamento diferente daquele marcado com  $a_i$  é vazia, ou um segmento reto, ou dois segmentos perpendiculares.

**Lema.** As classes  $[\gamma], [\gamma_i]_{1 \leq i \leq n}$  definem uma base para  $H_1(S; \mathbb{Z}/2\mathbb{Z})$ .

<sup>154</sup> J. S. Carter. *Classifying immersed curves*. *Proc. Amer. Math. Soc.*, 111(1):281–287, 1991.



A superfície  $S$  tem o tipo de homotopia de um grafo conexo com  $n$  vértices e  $2n$  arestas. A característica de Euler–Poincaré é  $-n$  e é igual a 1 menos o posto de  $H_1(S; \mathbb{Z}/2\mathbb{Z})$ . Portanto, esse posto é  $(n + 1)$ . Para provar o lema, basta mostrar que os  $[\gamma], [\gamma_i]_{1 \leq i \leq n}$  são linearmente independentes em  $H_1(S; \mathbb{Z}/2\mathbb{Z})$ .

Qualquer arco  $\sigma$  em  $S$  com pontos finais na borda de  $S$  define uma forma linear em  $H_1(S; \mathbb{Z}/2\mathbb{Z})$ : apenas conta pontos de intersecção com  $\sigma$  (sempre módulo 2). Por exemplo, escolha  $\sigma$  em algum cruzamento, como na margem, na página anterior. Claramente  $\gamma$  intercepta  $\sigma$  em apenas um ponto para que  $[\gamma]$  não seja trivial. Na cruz associada à letra  $a_i$ , deixe  $\sigma_i$  ser a soma das duas curvas mostradas na margem. A intersecção de  $\sigma_i$  com  $\gamma_j$  é 0 se  $i \neq j$  e 1 se  $i = j$ . A intersecção de  $\gamma$  e  $\sigma_i$  é 0. Portanto, é possível usar as  $n + 1$  formas lineares  $\sigma_i$  e  $\sigma$  para mostrar que  $[\gamma], [\gamma_i]_{1 \leq i \leq n}$  são linearmente independentes e o lema está provado. □

Agora, a superfície  $S$  tem o gênero zero se, e somente se, os números de intersecção dos laços  $\gamma, \gamma_i$  forem todos 0, módulo 2.

Observe que a autointersecção de qualquer curva fechada em qualquer superfície orientável é 0 (sobre  $\mathbb{Z}$  e, portanto, sobre  $\mathbb{Z}/2\mathbb{Z}$ ).

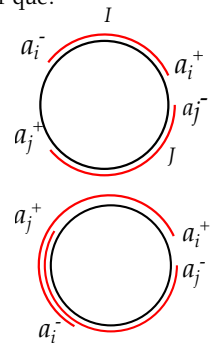
Vamos calcular o número de intersecção de  $\gamma$  e  $\gamma_i$ , denotado por  $\gamma \cdot \gamma_i$ . Para torná-los transversais, mova  $\gamma$  ligeiramente para a direita para obter uma curva  $\gamma$  que é transversal a  $\gamma_i$ . Não esqueça que a superfície  $S$  e os laços  $\gamma_i$  estão orientados. Segue-se que o número de intersecção de  $\gamma$  e  $\gamma_i$  é o número de letras entre  $a_i^+$  e  $a_i^-$ . Assim, recuperamos a condição necessária de Gauss.

*Suponha, a partir de agora, que essa condição esteja satisfeita.*

Vamos calcular o número de intersecção módulo 2 de  $\gamma_i$  e  $\gamma_j$ , denotado por  $\gamma_i \cdot \gamma_j$ .

Se as letras  $a_i^\pm, a_j^\pm$  estiverem não entrelaçadas, existem dois intervalos disjuntos  $I, J$  no círculo (ou em a palavra cíclica) cujos pontos finais são  $a_i^+, a_i^-$  e  $a_j^+, a_j^-$  respectivamente. Como  $\gamma_i$  pode ser substituído por  $\gamma_i - \gamma$  no cálculo do número de intersecção, segue-se que, neste caso não entrelaçado,  $\gamma_i \cdot \gamma_j$  é o número de letras na palavra com um ocorrência em  $I$  e a segunda em  $J$ . Esta

Por quê?





é, portanto, uma segunda condição de paridade, necessária para a realização de  $w$ .

Se as letras  $a_i^\pm, a_j^\pm$  estiverem *entrelaçadas*, os laços  $\gamma_i$  e  $\gamma_j$  não são transversais, uma vez que coincidem em algum intervalo não trivial. Mova  $\gamma_i$  levemente para a direita, para produzir  $\gamma'_i$ , e mova  $\gamma_j$  para a esquerda, para obter  $\gamma'_j$ . As curvas  $\gamma'_i$  e  $\gamma'_j$  estão agora paralelas nesta parte comum. Agora contamos o número de interseção de  $\gamma'_i$  e  $\gamma'_j$ . Veja a figura.

Há uma interseção no cruzamento  $a_j$  e nenhuma na cruzamento  $a_i$ . As outras interseções correspondem a letras entre  $a_i^+$  e  $a_i^-$  cuja segunda ocorrência é entre  $a_j^+$  e  $a_j^-$ .

Portanto, quando  $a_i^\pm, a_j^\pm$  estão *entrelaçados*, o número de interseções  $\gamma_i \cdot \gamma_j$  é igual a 1 mais o número de letras entre  $a_i^+$  e  $a_i^-$  cuja outra ocorrência é entre  $a_j^+$  e  $a_j^-$ .

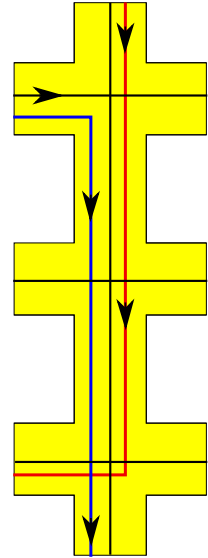
Portanto, obtemos uma resposta muito simples para o problema de Gauss com sinais.

**Teorema.** *Uma palavra de Gauss com sinais é realizável por uma curva imersa planar se, e somente se, as seguintes condições forem satisfeitas.*

1. Para cada letra  $a_i$ , existe um número par de letras entre  $a_i^+$  e  $a_i^-$  (condição de paridade de Gauss).
2. Para cada  $i, j$  tal que as letras  $a_i^\pm$  e  $a_j^\pm$  não estão *entrelaçadas*, sejam  $I, J$  os intervalos disjuntos cujos pontos finais são  $a_i^+, a_i^-$  e  $a_j^+, a_j^-$  (excluindo estes pontos). O número de letras na palavra com uma ocorrência em  $I$  e a outra em  $J$  é par.
3. Para cada  $i, j$  tal que as letras  $a_i^\pm$  e  $a_j^\pm$  estão *entrelaçadas*, o número de letras entre  $a_i^+$  e  $a_i^-$  cuja outra ocorrência é entre  $a_j^+$  e  $a_j^-$  é ímpar.

### O problema de Gauss

Vamos voltar ao problema original: palavras *sem sinais*. É claro que poderíamos trapacear e tentar todas as formas de escolher sinais na palavra. Isso pode demorar muito tempo. Até mesmo a força computacional de Gauss poderia ter sido derrotada por  $2^n$ . Além disso, isso não seria muito esclarecedor.



Equivalentemente, as duas primeiras condições necessárias podem ser expressas em termos do diagrama de cordas.

- 1 / Toda corda é intersectada por um número par de cordas.
- 2 / Dadas duas cordas sem interseção, o número de cordas que cruzam ambas é par.

Note que o critério de paridade de Gauss é independente dos sinais. A segunda condição, no caso em que  $a_i^\pm$  e  $a_j^\pm$  não estão entrelaçados, também é independente dos sinais.

Portanto, assumimos que ambos estão satisfeitos com uma palavra  $w$  sem sinais.

Deixe-me apresentar o grafo de entrelaçamento  $G(w)$ . Seus vértices são os inteiros  $1, \dots, n$  ou as cordas  $a_i$ , e há uma aresta entre duas cordas se elas se intersectarem.

Aqui está um exemplo do artigo de Cairns e Elton:

$$w = a_1 a_2 a_3 a_4 a_5 a_1 a_6 a_3 a_2 a_5 a_4 a_6.$$

Verifica-se facilmente que a condição de paridade de Gauss é satisfeita, bem como a segunda condição no caso não entrelaçado.

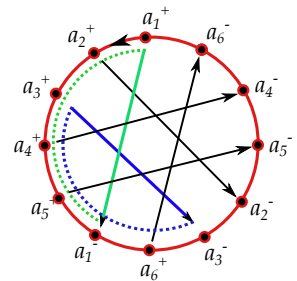
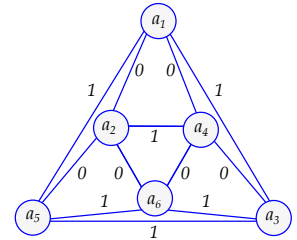
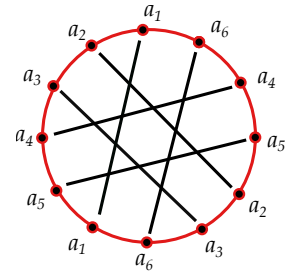
O grafo de entrelaçamento está representado na margem.

Escolha alguma palavra com sinais  $\bar{w}$  cuja palavra associada sem sinais seja  $w$ , por exemplo:

$$\bar{w} = a_1^+ a_2^+ a_3^+ a_4^+ a_5^+ a_1^- a_6^+ a_3^- a_2^- a_5^- a_4^- a_6^-.$$

Como estamos de volta às palavras com sinais, sabemos como calcular os números de interseção  $\gamma_i \cdot \Gamma_j$  dos laços associados. Isso define um elemento  $f(e) \in \mathbb{Z}/2\mathbb{Z}$  para cada aresta  $e$  de  $G(w)$ . Por exemplo, as duas cordas 1 e 3 se cruzam. Na figura da margem, não há letras no intervalo de  $a_1^+$  a  $a_1^-$  (em verde pontilhado) cuja outra ocorrência é entre  $a_3^+$  e  $a_3^-$  (em azul pontilhado). Portanto,  $\gamma_1 \cdot \gamma_3$  é igual a  $0 + 1$  módulo 2 e nós escrevemos um 1 na aresta conectando  $a_1$  e  $a_3$  no grafo. Como vemos 1 em algumas arestas em nosso exemplo, essa palavra com sinal não é realizada por uma curva imersa no plano. Agora temos que decidir se poderia haver uma mudança inteligente nos sinais nas letras para que todas as arestas pudessem ser rotuladas por 0.

Essa rotulação  $f(e)$  pode ser considerada como 1-cocadeia  $f$  (com valores em  $\mathbb{Z}/2\mathbb{Z}$ ) no grafo, portanto, um 1-cociclo (já que não há 2-faces em um grafo). Vamos examinar como esse cociclo muda quando os sinais em  $\bar{w}$  são modificados. Uma mudança de sinais é definida por alguma função  $u$  de  $\{1, \dots, n\}$  to  $\mathbb{Z}/2\mathbb{Z}$ , que podemos considerar como uma 0-cocadeia no nosso grafo. Afirimo que o novo 1-cociclo  $f'$  em  $G(w)$ , após a mudança de



sinal associado a  $u$ , é simplesmente igual a  $f + du$ , em que  $du$  é o *coborda* de  $u$ . Esse  $du$ , avaliado em alguma aresta  $e$  é, por definição, a diferença (ou soma desde que trabalhamos o módulo 2) dos valores de  $u$  nas duas extremidades de  $e$ .

Vamos começar alterando apenas o sinal de *uma letra*, digamos  $a_k$ . Para cada  $i, j$  tal que as letras  $a_i^\pm$  e  $a_j^\pm$  estão entrelaçadas, em outras palavras, para cada aresta do grafo  $G(w)$ , temos que comparar dois números de interseção, para as duas palavras com sinais  $\bar{w}, \bar{w}'$  cujos sinais diferem apenas na letra  $k$ . Claramente, esses números de interseção são iguais se  $k$  for diferente de  $i$  e  $j$ . Acontece que eles diferem em 1 (módulo 2), quando  $k = i$  ou  $k = j$ . Por exemplo, se  $k = i$ , temos que comparar:

1. o número de letras entre  $a_i^+$  e  $a_i^-$  cuja outra ocorrência é entre  $a_j^+$  e  $a_j^-$ , e
2. o número de letras entre  $a_i^-$  e  $a_i^+$  cuja outra ocorrência é entre  $a_j^+$  e  $a_j^-$ .

Assim, o módulo 2 é a diferença entre o número de letras  $a_j^+$  e  $a_j^-$ , diferente de  $a_i$ . Esse número é ímpar, pois assumimos que a condição de paridade de Gauss vale e  $a_i^\pm$  e  $a_j^\pm$  estão entrelaçados. Portanto, o efeito de alterar os sinais em uma única letra  $a_k$  é alterar em 1 os rótulos nas arestas de  $G(w)$  com  $a_k$  como um ponto final, e não modificar os outros rótulos. Portanto, para esse caso simples, a fórmula  $f' = f + du$  é válida. O caso geral segue desde que possamos mudar sinais um por um.

Como consequência, o objeto que é bem definido, independentemente dos sinais, é a *classe de cohomologia* de  $f$  em  $H^1(G(w); \mathbb{Z}/2\mathbb{Z})$ . Essa classe é zero se, e somente se, o cociclo é zero para alguma escolha dos sinais, ou seja, se, e somente se, a palavra sem sinais  $w$  é realizável por uma curva genérica imersa no plano.

Finalmente, uma classe de cohomologia em um grafo é trivial se, e somente se, for zero quando avaliada em ciclos. Isso dá um algoritmo muito eficiente: Escolha qualquer palavra com sinal  $\bar{w}$ , calcule o 1-cociclo e some seus valores em ciclos no grafo  $G(w)$ .

Em nosso exemplo, o ciclo  $a_1 \rightarrow a_3 \rightarrow a_5 = a_1$  em  $G(w)$  dá um total de 1 de tal modo que a palavra *sem sinal*  $w$  não seja associada a uma curva imersa no plano.

Veja <sup>155</sup> para uma história do problema até 1972 e <sup>156</sup> para um livro mais recente.

### O gênero de um diagrama de cordas

Seja  $w$  um diagrama com  $n$  cordas. Pegue um anel e cole as bordas de  $n$  bandas de acordo com  $w$ , como na figura da próxima página. Você obtém uma superfície  $S(w)$  com borda, que possui algum gênero (que é por definição o gênero da superfície fechada obtida após a colagem de discos em cada componente de fronteira): isto é, o gênero  $g(w)$  do diagrama de cordas.

Existe uma boa maneira de calcular esse gênero, devido à Moran <sup>157</sup>. Considere a matriz  $n \times n$  cujos coeficientes  $a_{ij} \in \mathbb{Z}/2\mathbb{Z}$  são iguais a 1 se as duas ocorrências de  $a_i$  estiverem entrelaçadas às duas ocorrências de  $a_j$  e 0 caso contrário. Essa é a *matriz de incidência* do grafo de entrelaçamento, módulo 2.

**Teorema.** O gênero de um diagrama de cordas é metade do posto da matriz de incidência módulo 2 do grafo  $G(w)$ .

A prova dada por Moran é bastante complicada, mas pode ser apresentada de uma maneira mais simples.

As cordas podem ser vistas como arcos em  $S(w)$ .

Nós também podemos ver essas cordas como arcos no disco de dimensão 2.

Cole o disco e  $S(w)$  ao longo do círculo externo para obter uma superfície  $S'(w)$ . As duas cópias de cada corda definem laços  $b_1, \dots, b_n$  gerando a homologia de  $S'(w)$ . Nessa base, a intersecção de  $b_i$  e  $b_j$  é 0 se  $i, j$  não se entrelacem e 1 caso contrário.

Nós agora colamos  $k$  discos ao longo dos componentes do bordo de  $S'(w)$  para produzir uma superfície orientada fechada  $\hat{S}(w)$ . O mergulho de  $S'(w)$  em  $\hat{S}(w)$  induz uma sobrejeção (módulo 2) na primeira homologia. Na verdade, qualquer laço em  $\hat{S}(w)$  pode ser deformado homotopicamente fora dos discos que foram adicionados. No entanto, esse mergulho não induz uma injeção na homologia. De fato, quando colamos discos ao longo da fronteira, os componentes de contorno morrem em homologia, uma vez que agora são bordas de discos. No

<sup>155</sup> B. Grünbaum. *Arrangements and spreads*, volume 10 of *Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics*. American Mathematical Society Providence, R.I., 1972.

<sup>156</sup> C. Godsil and G. Royle. *Algebraic graph theory*, volume 207 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2001.



Desvio no desvio.

<sup>157</sup> G. Moran. [Chords in a circle and linear algebra over GF\(2\)](#). *J. Combin. Theory Ser. A*, 37(3):239–247, 1984.



entanto, qualquer elemento no núcleo está no núcleo da forma de interseção de  $S(w)$ , uma vez que é homóloga a uma coleção de curvas paralelas à borda. Segue-se que a forma de interseção em  $S'(w)$ , módulo seu núcleo, é isomorfo à forma de interseção de  $\hat{S}(w)$ . Para uma superfície orientada fechada do gênero  $g$ , a forma de interseção é não degenerada de posto  $2g$ .  $\square$

*Um teorema de Lovász e Marx*

Existe uma solução diferente para o problema de Gauss. Ela se encaixa muito bem com a nossa descrição de permutações separáveis como aqueles que evitam os dois padrões 3142 e 2413. Curiosamente, este teorema é publicado sem provas <sup>158</sup>. Tenho certeza que o meu leitor gostará de encontrar uma prova pessoal.

Dada uma curva planar imersa genérica, há duas maneiras de excluir um determinado ponto duplo, ilustradas na margem. Na primeira, a curva é dividida em dois componentes, de tal forma que podemos escolher um deles.

Do ponto de vista combinatório, essas duas operações podem ser expressas da seguinte maneira.

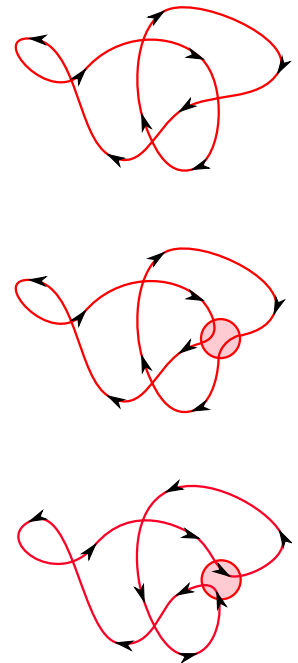
- Partindo de uma palavra  $w = aUaV$ , exclua  $a$  e considere a palavra  $UV^{-1}$  (todas as palavras são escritas ciclicamente).
- Começando com uma palavra  $w = aUaV$ , exclua  $a$  e todas as letras que aparecem em  $V$ .

Portanto, cada palavra  $w$  produz algumas outras palavras com menos letras. As imagens na margem mostram que as novas palavras são realizáveis se a primeira foi. Continue e produza novas palavras mais curtas. Dizem que estas palavras mais curtas são *contidas* na palavra inicial  $w$ .

**Teorema.** *Uma palavra de Gauss é realizável se ela não contiver a palavra  $a_1a_2 \cdots a_n a_1a_2 \cdots a_n$  com  $n$  par.*

Observe que o grafo de entrelaçamento de  $a_1a_2 \cdots a_n a_1a_2 \cdots a_n$  é o grafo completo em  $n$  vértices.

<sup>158</sup> L. Lovász and M. L. Marx. [A forbidden substructure characterization of Gauss codes.](#) *Acta Sci. Math. (Szeged)*, 38(1-2):115-119, 1976.



*Um operad Gaussiano*

Defino uma estrutura de operad no conjunto de palavras de Gauss. Mais precisamente, vou lidar com curvas imersas orientadas *marcadas* genéricas no plano  $\mathbb{R}^2 \simeq \mathbb{C}$ . *Genérico* significa, como antes, que vários pontos são apenas pontos duplos com duas tangentes diferentes. *Marcado* significa que escolhemos um dos pontos duplos como o ponto inicial e que os pontos duplos restantes são rotulados de 1 a  $n$ .

Como o plano e nossa curva estão orientados, qualquer ponto duplo define uma rosa bússola (rosa dos ventos): os quatro pontos de interseção com um pequeno círculo ao redor do ponto podem ser rotulados pelas direções cardeais.

Se o ponto de partida marcado for enviado no infinito por inversão, obtemos uma imagem composta de duas curvas *longas*,  $\gamma_b, \gamma_r : \mathbb{R} \rightarrow \mathbb{R}^2$  (azul e vermelho) com as seguintes propriedades.

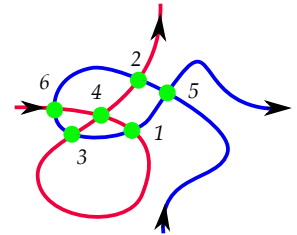
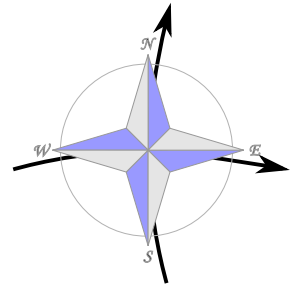
- A curva azul (resp. vermelho)  $\gamma_b$  (resp.  $\gamma_r$ ) vai do sul para o leste (respectivamente de oeste para norte). Mais precisamente  $\gamma_b(t)$  é igual a  $it$  para grandes valores negativos de  $t$  e  $t$  para grandes valores positivos de  $t$  (resp.  $T$  e  $it$ ).
- $\gamma_b$  e  $\gamma_r$  estão imersos e os únicos pontos múltiplos de sua união são pontos duplos transversais chamados  $1, \dots, n$ .

Denote por  $\Gamma_n$  o conjunto de pares de curvas satisfazendo essas propriedades e tendo  $n$  pontos duplos, a menos de difeomorfismos preservando a orientação do plano.

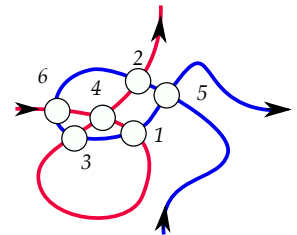
Existe uma estrutura natural de operad na união dos  $\Gamma_n$ .

Pegue um par  $(\gamma_b, \gamma_r)$  de curvas azul-vermelho como acima, tendo  $n$  pontos rotulados duplos.

Escolha  $n$  pares de curvas azul-vermelho  $(\gamma_{b,i}, \gamma_{r,i})$  (para  $i = 1, \dots, n$ ), tendo  $k_1, k_2, \dots, k_n$  como pontos duplos. Cavar pequenos discos em torno dos pontos duplos de  $(\gamma_b, \gamma_r)$ . Gostariamos de inserir os  $(\gamma_{b,i}, \gamma_{r,i})$  no disco com o rótulo  $i$ , respeitando as direções cardeais. No entanto, isso não é possível. Quando cavamos um buraco, as curvas azuis vão do sul para o norte e as curvas vermelhas vão do oeste para o leste, de modo que isso não é coerente com o comportamento do sudeste e oeste-norte



Um elemento de  $\Gamma_6$ .

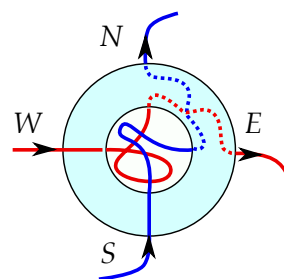


das curvas azul e vermelha  $(\gamma_{b,i}, \gamma_{r,i})$  que queremos inserir. É fácil contornar esse problema. Antes de inserir os  $(\gamma_{b,i}, \gamma_{r,i})$ , é suficiente inserir primeiro um anel padrão contendo arcos orientados, alternando norte e leste.

O resultado dessa operação de recortar e colar é um par de curvas com  $k_1 + \dots + k_n$  pontos duplos.

Este é o operad *Gaussiano (simétrico)*. Leitor, você consegue encontrar um sistema de geradores? E as relações entre geradores?

Para concluir este capítulo sobre uma ampla abertura, recomendando o livro <sup>159</sup>, uma introdução notável e compreensível sobre os invariantes de nós de Vassiliev, nos quais os diagramas de cordas desempenham um papel crucial.



<sup>159</sup> S. Chmutov, S. Duzhin, and J. Mostovoy. *Introduction to Vassiliev knot invariants*. Cambridge UP, 2012.

**Ein vollständiger Knoten.**

1.  $aa$

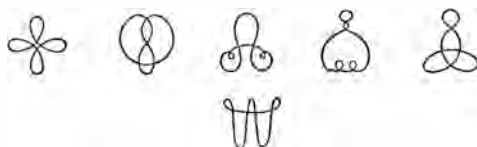
**Zwei vollständige Knoten.**

1.  $aabb$
2.  $abab$  \*
3.  $abba$

**Drei vollständige Knoten.**

- |                |                 |
|----------------|-----------------|
| 1. $aabbcc$    | 9. $abbcca$     |
| 2. $abcbbc$ *  | 10. $abcabc$    |
| 3. $aabccb$    | 11. $abcacb$ *  |
| 4. $ababcc$ *  | 12. $abcbac$ *  |
| 5. $abacbc$ ** | 13. $abcbca$ ** |
| 6. $abaccb$ *  | 14. $abccab$ *  |
| 7. $abbacc$    | 15. $abccba$    |
| 8. $abbcac$ *  |                 |

A lista de Gauss, com 1, 2, 3 letras.



## Vier vollständige Knoten.

- |                        |                        |                         |
|------------------------|------------------------|-------------------------|
| 1. <i>aabbccdd</i>     | 36. <i>abbcaddc</i> *  | 71. <i>abccadbd</i> **  |
| 2. <i>aabbcdcd</i> *   | 37. <i>abbccadd</i>    | 72. <i>abccaddd</i> *   |
| 3. <i>aabbcdcc</i>     | 38. <i>abbccdada</i> * | 73. <i>abccbadd</i>     |
| 4. <i>aabcbcdd</i> *   | 39. <i>abbcddda</i>    | 74. <i>abccbada</i> *   |
| 5. <i>aabcdbcd</i> **  | 40. <i>abbcdacd</i>    | 75. <i>abccbdda</i>     |
| 6. <i>aabcdbdc</i> *   | 41. <i>abbcdadc</i> *  | 76. <i>abccdabd</i>     |
| 7. <i>aabccbdd</i>     | 42. <i>abbcdcad</i> *  | 77. <i>abccdadbb</i> *  |
| 8. <i>aabccdbd</i> *   | 43. <i>abbcddca</i> ** | 78. <i>abccdbad</i> *   |
| 9. <i>aabccddb</i>     | 44. <i>abbcddac</i> *  | 79. <i>abccdbda</i> **  |
| 10. <i>aabcbcbd</i>    | 45. <i>abbcddca</i>    | 80. <i>abccddab</i> *   |
| 11. <i>aabcbddc</i> *  | 46. <i>abcabcdd</i>    | 81. <i>abccddba</i>     |
| 12. <i>aabcbdbd</i> *  | 47. <i>abcabdcd</i> ** | 82. <i>abcdabcd</i> *   |
| 13. <i>aabcbddb</i> ** | 48. <i>abcabddc</i>    | 83. <i>abcdabdc</i> *   |
| 14. <i>aabcbddc</i> *  | 49. <i>abcacbdd</i> *  | 84. <i>abcdacbd</i> **  |
| 15. <i>aabcbddc</i>    | 50. <i>abcacdbd</i> ** | 85. <i>abcdacdb</i> **  |
| 16. <i>ababccdd</i> *  | 51. <i>abcacddb</i> *  | 86. <i>abcdadbc</i> *   |
| 17. <i>ababdcdd</i> *  | 52. <i>abcadbcd</i> *  | 87. <i>abcdadcb</i> *   |
| 18. <i>ababdcdd</i> *  | 53. <i>abcadbdc</i> ** | 88. <i>abcdbacd</i> *   |
| 19. <i>ababcadd</i> ** | 54. <i>abcadcdb</i>    | 89. <i>abcdbadc</i>     |
| 20. <i>ababcadd</i> ** | 55. <i>abcadcdb</i> ** | 90. <i>abcdbcad</i> **  |
| 21. <i>ababcadd</i> ** | 56. <i>abcadddc</i>    | 91. <i>abcdbcda</i>     |
| 22. <i>abaccbdd</i> ** | 57. <i>abcadddc</i> *  | 92. <i>abcbddac</i> **  |
| 23. <i>abaccbdd</i> ** | 58. <i>abcbacdd</i> *  | 93. <i>abcbddca</i> **  |
| 24. <i>abaccddb</i> ** | 59. <i>abcbadcd</i> ** | 94. <i>abcbdcad</i> **  |
| 25. <i>abacdbcd</i> ** | 60. <i>abcbaddd</i> *  | 95. <i>abcbdcad</i> **  |
| 26. <i>abacdbdc</i> ** | 61. <i>abcbacdd</i> ** | 96. <i>abcbdcad</i> *   |
| 27. <i>abacdbdd</i> ** | 62. <i>abcbcdad</i> ** | 97. <i>abcbdcda</i> **  |
| 28. <i>abacdbdd</i> ** | 63. <i>abcbcdad</i> ** | 98. <i>abcbdcad</i> **  |
| 29. <i>abacdbdd</i> ** | 64. <i>abcbdcad</i> ** | 99. <i>abcbdcba</i> **  |
| 30. <i>abacdbdd</i> ** | 65. <i>abcbdcad</i> ** | 100. <i>abcbdcad</i>    |
| 31. <i>abbaacdd</i>    | 66. <i>abcbdcad</i> ** | 101. <i>abcbdcad</i> *  |
| 32. <i>abbaacdd</i> *  | 67. <i>abcbdcad</i> ** | 102. <i>abcbddbc</i> *  |
| 33. <i>abbaacdd</i>    | 68. <i>abcbddac</i> *  | 103. <i>abcbddbc</i> ** |
| 34. <i>abbaacdd</i> *  | 69. <i>abcbddca</i> ** | 104. <i>abcbddcb</i> *  |
| 35. <i>abbaacdd</i> ** | 70. <i>abcbddca</i> *  | 105. <i>abcbddcb</i>    |





O carácter Kanji da palavra  
"árvore".

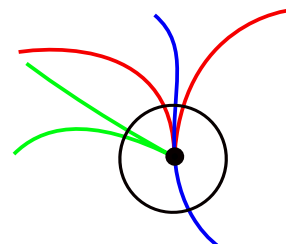


## *Diagramas de cordas analíticas: um algoritmo*

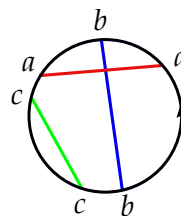
NESTE CAPÍTULO, ALCANÇAMOS UM DE NOSSOS OBJETIVOS: a descrição algorítmica dos diagramas de cordas que ocorrem na vizinhança de um ponto singular de uma curva analítica real plana. Lembre-se de que essa curva intersecta um pequeno círculo em torno do ponto singular em um número par de pontos que vêm em pares, cada par sendo associado a um ramo real.

Pode-se pensar nessa estrutura como uma palavra cíclica de comprimento  $2n$  na qual cada letra ocorre exatamente duas vezes (os nomes das letras são irrelevantes). Para ser mais pedante (e preciso), estamos discutindo pontos fixos de involuções livres de  $\mathbb{Z}/2n\mathbb{Z}$  a menos de conjugações por permutações cíclicas. Também podemos desenhar  $n$  cordas em um círculo.

O número total desses *diagramas de cordas* de comprimento  $2n$  foi estudado em muitos artigos. Veja por exemplo <sup>160</sup> com fortes motivações da teoria dos nós. O problema seria fácil se, em vez de uma palavra cíclica, procurássemos palavras padrão (não cíclicas) de comprimento  $2n$  nas quais cada letra ocorre exatamente duas vezes e na qual os nomes das letras são irrelevantes. De fato, escreva a primeira letra da palavra e, em seguida, escolha qualquer um dos  $2n - 1$  locais restantes para a outra letra que seja idêntica à primeira, escreva a segunda letra no primeiro lugar livre disponível e escolha outra letra idêntica em qualquer um dos  $2n - 3$  locais restantes etc. Portanto,



Uma curva com três ramos.



*abaccb*

Às vezes eu pulo a palavra “corda” para “diagrama de corda”.

<sup>160</sup> A. Stoimenow. [On the number of chord diagrams.](#) *Discrete Math.*, 218(1-3):209–233, 2000.

o número total dessas palavras é  $(2n - 1) \cdot (2n - 3) \cdots 3 \cdot 1$ . Às vezes esses números são chamados de *fatoriais duplos* e denotados por  $(2n - 1)!!$ . Veja <sup>161</sup> para uma apresentação de suas propriedades combinatórias.

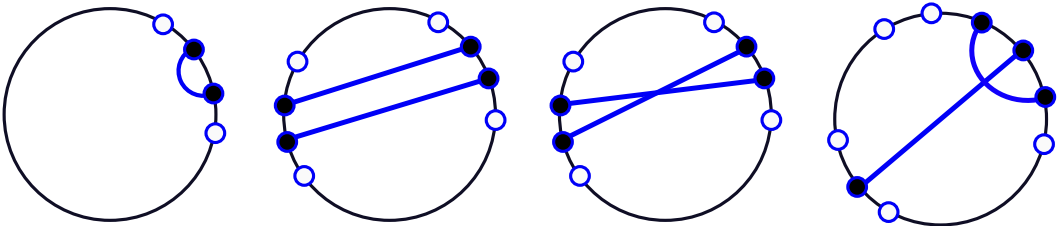
Seria tentador dividir  $(2n - 1)!!$  por  $2n$  para levar em conta as permutações cíclicas, mas algumas palavras admitem simetrias, por isso, que a combinatória é mais sutil. Em qualquer caso, destas considerações, segue-se que o número de diagramas de cordas de comprimento  $2n$  cresce exponencialmente com  $n$ .

Veremos que uma proporção muito pequena de diagramas de cordas são *analíticos*, no sentido de que eles estão associados a alguma singularidade de uma curva analítica plana.

### Uma condição necessária

Lembre-se de que no primeiro capítulo deste livro mostramos que, para qualquer permutação separável, existem dois inteiros consecutivos com imagens consecutivas. Esse foi o ponto-chave que nos permitiu produzir um algoritmo que decide se uma permutação é separável. Agora irei provar uma propriedade similar para diagramas de cordas analíticos.

Direi que uma corda em um diagrama é *solitária* se ela conecta dois pontos consecutivos do diagrama como na imagem à esquerda abaixo. Duas cordas são *paralelas* (resp. *antiparalelas*) se estiverem na segunda (terceira) imagem, ou seja, se as letras correspondentes  $a, b$  ocorrerem na palavra cíclica como  $\cdots ab \cdots ba \cdots$  (respectivamente  $\cdots ab \cdots ab \cdots$ ). Finalmente, duas cordas, como na quarta imagem, constituem uma *forquilha* ( $\cdots a \cdots bab \cdots$ ).



<sup>161</sup> D. Callan. [A combinatorial survey of identities for the double factorial](#). Preprint no arXiv, 2009.

Cuidado! O fatorial duplo  $(2n - 1)!!$  não é nem o fatorial do fatorial nem um ponto de exclamação!

Use a fórmula de Stirling para mostrar que  $(2n - 1)!!$  é equivalente a  $\sqrt{2} \left(\frac{2}{e}\right)^n e^{n \log n}$  quando  $n$  tende ao infinito. Então,  $(2n - 1)!!$  está de fato crescendo superexponencialmente, mas não muito mais rápido. Por exemplo, é pequeno quando comparado com  $C\lambda^{n^{1+\epsilon}}$  para qualquer  $\epsilon > 0$  e  $\lambda > 1$ .

**Lema fundamental** ☺. *Qualquer diagrama de cordas analítico contém uma corda solitária, um par de cordas paralelas ou antiparalelas ou uma forquilha.*

Observemos que isso implica imediatamente o teorema que foi enunciado no prefácio.

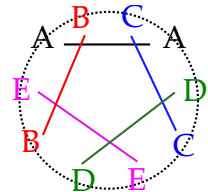
**Teorema.** *Não há uma curva analítica singular no plano que consiste em cinco ramos cruzando um pequeno círculo como na imagem na margem.* □

*Vamos ao blow-up*

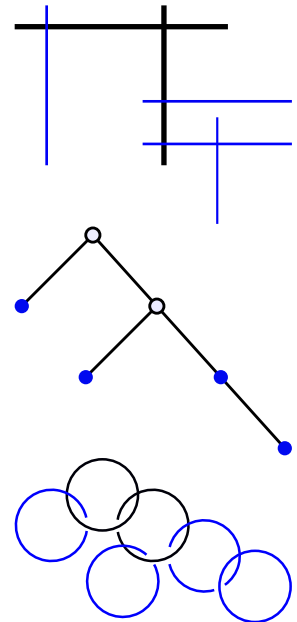
Comece com algum ponto singular de alguma curva analítica no plano (real). Faça um blow-up pela primeira vez. O resultado é uma curva em alguma faixa de Möbius, cujos pontos singulares estão no excepcional divisor, no meio da faixa. Se as coisas forem bem, o ponto singular se divide em vários pontos singulares, presumivelmente mais simples. Vamos fazer um blow-up de todos eles. Pode acontecer que, após um blow-up, ainda exista um único ponto singular no divisor. Então, vamos fazer um blow-up pela segunda vez. Vamos continuar o processo de fazer blow-up quantas vezes forem necessárias. Sabemos que depois de algum tempo, a singularidade será *resolvida*. Isso significa que a transformação estrita da curva inicial é agora uma coleção de  $n$  curvas analíticas suaves disjuntas, que intersectam transversalmente o divisor excepcional.

Esse divisor excepcional é uma união de retas projetivas reais que são círculos que se cruzam transversalmente. Considere o grafo cujos vértices são essas retas projetivas e uma aresta conecta dois vértices se as retas projetivas se cruzarem. Esse grafo é uma árvore, como pode ser facilmente visto por indução. De fato, no processo indutivo de desingularização, a cada passo nós fazemos um blow-up num ponto que pode ser um ponto suave do divisor excepcional ou uma interseção de duas retas projetivas. No primeiro caso, uma nova folha é enxertada em uma árvore e, no segundo caso, uma aresta é dividida em duas arestas. A primeira reta projetiva, vinda do primeiro blow-up, pode ser escolhida como a raiz dessa árvore.

© Alguns teoremas ou lemas são tão famosos que parece proibido usar palavras como “Teorema A, B” (Cartan) ou “Lema Fundamental” (Ngô) .



Cinco ramos impossíveis.

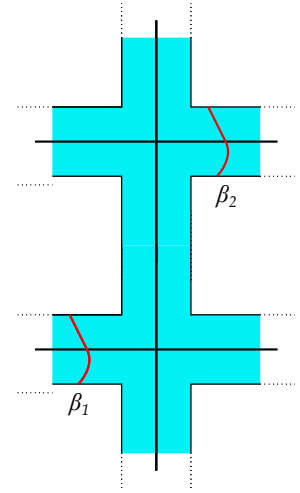


Será conveniente aumentar e fazer mais um blow-up de cada um dos  $n$  pontos no divisor excepcional, se necessário, introduzindo novas retas projetivas, a fim de garantir que no final do processo cada reta projetiva contenha *no máximo* um ponto da transformação estrita. Podemos até supor que aqueles componentes do divisor que tocam a transformação estrita são folhas da árvore de desingularização.

Vamos resumir. Dada alguma curva analítica  $\mathcal{C}$  definida em uma vizinhança de  $(0,0)$  em  $\mathbb{R}^2$  por alguma equação  $F(x,y) = 0$ , podemos construir os seguinte objetos:

- Uma superfície  $S$  com borda conexa e orientada.
- Um divisor excepcional  $E \subset S$ , consistindo em um certo número de círculos se intersectando transversalmente, cada par se encontrando no máximo uma vez. O grafo de interseção associado é uma árvore com raiz. O mergulho  $E \subset S$  é uma equivalência de homotopia.
- Uma união disjunta finita  $\hat{\mathcal{C}}$  de arcos analíticos *suaves*  $\beta_1, \dots, \beta_n$  em  $S$  intersectando transversalmente  $E$ . A interseção de  $\hat{\mathcal{C}}$  com uma componente de  $E$  será vazia se este componente não for uma folha da árvore e contiver no máximo um ponto se for uma folha. Podemos supor, além disso, que  $\hat{\mathcal{C}}$  é transversal à borda de  $S$  e que cada arco  $\beta_i$  cruza a borda em dois pontos.
- Uma aplicação de blow-down analítica  $\Psi : S \rightarrow \mathbb{R}^2$ , colapsando  $E$  para a origem, um difeomorfismo de  $S \setminus E$  em algum pequeno disco perfurado, que colapsa  $\hat{\mathcal{C}}$  para nossa curva singular  $\mathcal{C}$ .

Lembre-se de que cada laço em  $S$  pode ser orientante ou desorientante. Seja  $\gamma$  uma curva imersa fechada em uma superfície, passando uma vez por um ponto  $x$ . Quando é realizado um blow-up da superfície em  $x$ , a autointerseção (*módulo 2*) da transformação estrita de  $\gamma$  (no blow-up da superfície) é igual à autointerseção de  $\gamma$  (na superfície original) mais 1. Na construção indutiva, quando uma reta projetiva aparece pela primeira vez no divisor excepcional, é a espinha dorsal de uma faixa de Möbius, de autointerseção 1. Mais tarde, podemos fazer um



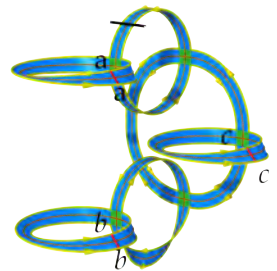
blow-up em alguns de seus pontos. Cada um desses blow-ups permuta o status de orientante/desorientante de uma componente do divisor. As imagens anteriores na margem (seis retas, uma árvore com seis vértices e seis círculos) correspondem ao mesmo exemplo que foi analisado no capítulo sobre colares, sendo obtido após seis blow-ups. Eu indiquei em azul as componentes que correspondem às faixas de Möbius.

Algumas componentes de  $E$  cruzam a curva desingularizada  $\hat{C}$ : elas definem algumas das folhas na árvore de desingularização. Chame essas folhas de *coloridas*. Observe que algumas folhas podem estar sem cor.

Observe que, se escolhermos alguma orientação de cada componente de  $E$ , a árvore correspondente é planar de modo que os filhos de qualquer vértice são ordenados linearmente. Alterar a orientação inverte essa ordem.

### *Um exemplo*

Olhe o colar na margem. Este ainda é o mesmo objeto que já foi discutido no capítulo sobre colares de Möbius. Seis blow-ups produziram seis faixas, duas orientáveis e quatro não orientáveis. O divisor excepcional consiste nas seis espinhas dorsais das seis faixas. A curva desingularizada consiste em três arcos vermelhos, denominados  $a, b, c$ , cada um cruzando a borda de  $S$  em dois pontos. No topo, vemos em preto a transformação estrita do eixo  $y$ . Indo ao redor da borda de  $S$ , podemos ler o diagrama de cordas analítico correspondente. Basta seguir a seta e ler  $ababc$ . Pois é, aqui, devo admitir que não é tão fácil seguir as flechas!



### *A prova do lema fundamental*

Uma primeira observação trivial é que a operação de deletar uma corda em um diagrama analítico o transforma em algum outro diagrama analítico. Isso corresponde a excluir um ramo.

Comece com algum diagrama de cordas  $w$  associado a alguma singularidade de uma curva analítica planar. Considere uma árvore de desingularização como antes. Há uma projeção  $\pi$  da

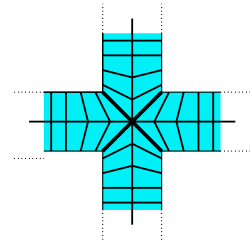
superfície  $S$  no divisor excepcional  $E$  que é uma equivalência homotópica. Para cada ponto  $x \in E$ , a fibra  $\pi^{-1}(x)$  é um arco conectando dois pontos da borda se  $x$  for um ponto regular, e uma cruz se  $x$  for um ponto a intersecção de dois círculos.

Seja  $L$  algum vértice da árvore, ou seja, uma das retas projetivas que constituem o divisor excepcional  $E$ . Existe um único caminho de vértices indo de  $L$  até a raiz. Corte dois arcos disjuntos em  $S$ , como na figura, para desconectar  $L$  da raiz em  $S$ . Os quatro pontos finais desses arcos decompõem o bordo do círculo de  $S$  em quatro intervalos. Dois deles (de cor vermelha) correspondem a “o que está em  $L$  ou abaixo de  $L$ ” na árvore. Indo ao redor da borda de  $S$  e lendo o diagrama de cordas, encontramos dois intervalos de letras, abaixo de  $L$ , cuja união é estável sob a involução, enviando cada ocorrência de uma letra para a outra ocorrência. Observe que esses intervalos podem estar vazios se não houver uma folha colorida abaixo de  $L$ . Se houver uma folha colorida abaixo de  $L$ , pelo menos um dos dois intervalos não está vazio, mas pode ser que apenas um deles não seja vazio.

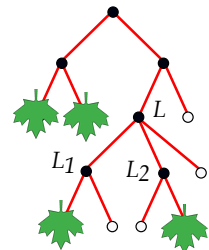
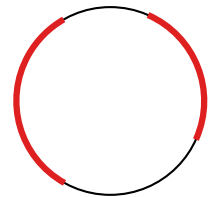
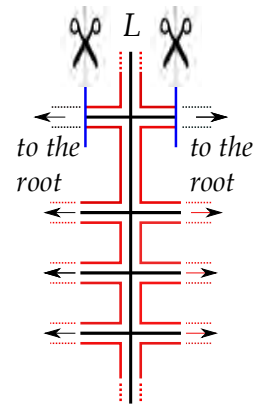
Em resumo, *todo vértice  $L$  na árvore define um diagrama de cordas  $w(L)$  que é um subdiagrama do diagrama original  $w$  e que é “conexo” no sentido de que suas letras formam um ou dois intervalos em  $w$ .*

Pense em uma árvore enraizada como uma árvore genealógica, sendo a raiz o membro fundador da família. Cada vértice tem um certo número de descendentes, alguns deles sendo folhas coloridas. Seja  $L$  um dos membros mais jovens da família com pelo menos duas folhas coloridas como descendentes. Entre os filhos de  $L$ , seja  $L_1, \dots, L_k$  a lista daqueles com pelo menos um descendente colorido (assim ordenado ao longo de  $L$ ). Temos  $k \geq 2$  já que de outra forma um dos filhos de  $L$  teria pelo menos dois descendentes coloridos. Pelo mesmo motivo, cada  $L_i$  tem um descendente colorido.

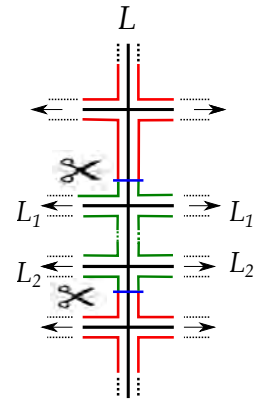
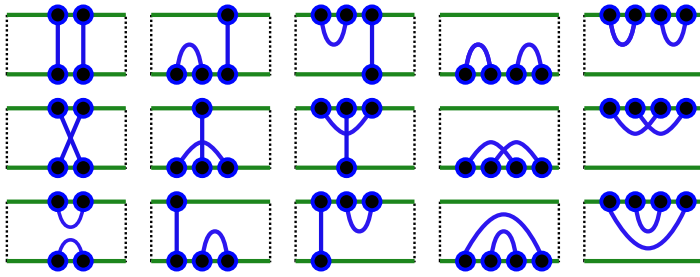
Agora corte a superfície  $S$  como na figura da margem, desconectando  $L_1$  e  $L_2$  da raiz e de todas as outras folhas coloridas. Como antes, isso define dois intervalos (verdes) no limite de  $S$  cuja união contém exatamente quatro pontos do nosso diagrama inicial, associados a duas cordas, que em dois intervalos podem ser organizados das quinze maneiras a seguir:



As fibras de  $\pi$ . Não confunda  $\pi$  (de  $S$  para o divisor  $E$ ) com  $\Psi$  (de  $S$  para o disco) o qual é um blowing down  $E$  para um ponto.



As folhas coloridas são verdes!



Em cada caso, há uma corda solitária, uma forquilha ou um par de cordas paralelas ou antiparalelas.

Este é o fim da prova do lema fundamental. □

### Mais diagramas não analíticos

Observamos que a exclusão de algumas letras em um diagrama analítico de cordas produz outro diagrama analítico. Digamos que um diagrama seja *básico não analítico* se não for analítico, mas se torna analítico quando uma única corda é excluída. Claramente, um diagrama de cordas analítico não contém um *diagrama básico de cordas não analítico*. Lembre-se que uma permutação é separável se, e somente se, não contém a permutação de Kontsevich  $(2, 4, 1, 3)$  ou sua permutação reversa  $(3, 1, 4, 2)$ , de modo que, neste caso, existem apenas duas permutações básicas não separáveis. A situação é mais complicada no caso de diagramas de cordas.

**Teorema.** *Existe um número infinito de diagramas básicos não analíticos de cordas.*

Aqui é um exemplo que será denotado por  $C_n$  ( $n \geq 5$ ). Considere os  $2n$  pontos de  $\mathbb{Z}/2n\mathbb{Z}$  ordenados de forma natural no círculo. O diagrama de cordas combina  $2k$  e  $2k + 3$  para  $k = 1, \dots, n$ . Para  $n = 5$ , é o nosso exemplo anterior de diagrama não analítico com cinco cordas. Esse diagrama  $C_n$  ( $n \geq 5$ ) não é analítico pelo mesmo motivo que no caso  $n = 5$ . Ainda temos que mostrar que, se uma letra for excluída, o diagrama restante é analítico. Para isso, precisamos de uma *condição suficiente de analiticidade*, que



será fornecido por meio de um algoritmo muito simples que determina se um diagrama de cordas é analítico.

**Teorema.** *O algoritmo a seguir decide se um diagrama de corda é analítico:*

1. *Se não houver uma corda solitária, nenhuma forquilha e nenhum par de cordas paralelas ou antiparalelas, o diagrama não é analítico.*
2. *Se houver uma corda solitária, apague-o e continue. Se houver uma forquilha, exclua a “pequena corda” e mantenha a alça. Se houver um par de cordas paralelas ou antiparalelas, exclua uma delas e continue.*
3. *Se você acabar com o diagrama vazio, o original era analítico.*

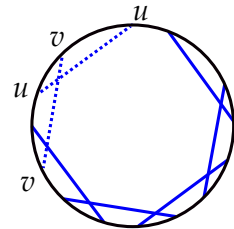
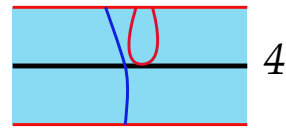
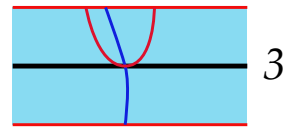
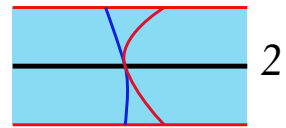
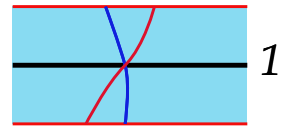
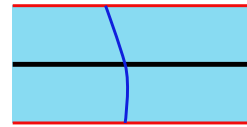
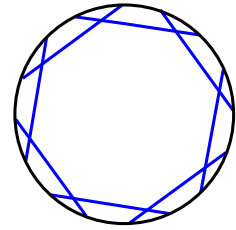
A prova é fácil. Nós temos que mostrar que se  $w$  é um diagrama e se  $\bar{w}$  é o novo diagrama com uma corda a menos obtido após uma etapa do algoritmo, então,  $w$  é analítico se  $\bar{w}$  é analítico. O inverso é fácil: se  $w$  é analítico,  $\bar{w}$  também é analítico, pois corresponde à exclusão de um ramo. Agora temos que *adicionar* um ramo adicional.

Escolha alguma desingularização de algum ponto singular em alguma superfície  $S$ . Uma corda corresponde a um arco suave  $\gamma$  (em azul na margem) conectando dois pontos no bordo, transversal ao divisor em algum ponto  $x$ . Adicione uma nova curva analítica (vermelha)  $\gamma'$  em  $S$ , também transversal ao divisor, muito próxima de  $\gamma$  e transversal a  $\gamma$ , como em 1. O blow-down de  $S$  produz um novo ponto singular com mais um ramo. Claramente, o novo diagrama tem mais uma corda paralela ou antiparalela à corda inicial, dependendo das orientações da borda. Agora escolha  $\gamma'$  como em 2 com uma tangência quadrática com  $\gamma$  em  $x$  e você obtém a outra situação paralela ou antiparalela.

Caso você queira criar uma forquilha com uma dada alça, basta adicionar uma curva suave  $\gamma'$  próxima a  $\gamma$  com uma tangência quadrática com o divisor como em 3.

Finalmente, se você quiser adicionar uma corda solitária logo após alguma carta, prossiga como em 4. □

Vamos testar nosso algoritmo no diagrama anterior em  $\mathcal{C}_n$  ( $n \geq 5$ ). Excluindo uma corda, obtemos uma cadeia não cíclica



de  $n - 1$  cordas. As duas primeiras cordas  $u, v$  definem uma forquilha, então, podemos excluir a primeira corda e continuar até que haja apenas uma corda. Este diagrama é, portanto, analítico. Existe, de fato, um número infinito de diagramas básicos não analíticos.

### Usando o computador

Para contar os diagramas analíticos de cordas, vamos usar um computador para testar valores pequenos de  $n$ . Isso é fácil. Primeiro listamos todas as palavras possíveis de comprimento  $2n$  em que cada letra ocorre duas vezes. A única sutileza é levar em conta o caráter cíclico da palavra em consideração. Aqui está o resultado para  $n \leq 7$ .

Na tabela a seguir:

$n$	2	3	4	5	6	7
Palavras	3	15	105	945	10395	135135
Diagramas de cordas	2	5	18	105	902	9749
A menos de simetria	2	5	17	79	554	5283

– *Palavras* significa “palavras de comprimento  $2n$  em que cada letra ocorre duas vezes”. O número dessas palavras é o fatorial duplo de  $2n - 1$ .

– *diagramas de cordas*, como os definimos, são palavras a menos de permutações cíclicas.

– A imagem de um diagrama de cordas por uma simetria em relação a alguma reta é outro diagrama, que pode ser o mesmo diagrama ou não. O item “a menos de simetria” conta o número de palavras a menos dessas simetrias diedrais.

Em seguida, contamos o número de diagramas *analíticos*. Em princípio, isso não é difícil, usando o algoritmo descrito anteriormente.

O resultado é:

$n$	2	3	4	5	6	7
Diagramas analíticos	2	5	18	102	817	7641
A menos de simetrias	2	5	17	76	499	4132

The On-Line Encyclopedia of Integer Sequences, [A007769](#), [A054499](#), [A279207](#), [A279208](#). Neil Sloane adicionou as duas últimas sequências baseadas em uma versão preliminar deste livro.

Segue que para  $n \leq 4$ , todos os diagramas são analíticos.

Entre os 105 diagramas com 5 cordas, apenas os 3 exemplos na margem não são analíticos.

O primeiro diagrama já é familiar, sob o nome  $C_5$ . Deixei-me denotar os outros por Não foi difícil para mim adivinhar o primeiro, mas devo admitir que não encontrei os outros dois à mão, mas com um computador.

Entre os 902 diagramas de com 6 cordas, 85 não são analíticos.

No entanto, a maioria deles não é analítico porque um de seus subdiagramas não é analítico. Existem apenas dois diagramas com 6 cordas que são básicos: não são analíticos e todos os seus subdiagramas são analíticos.

Observe que o primeiro é o membro  $C_6$  da família infinita dos diagramas básicos  $C_n$ . Corresponde a  $\mathbb{Z}/12\mathbb{Z}$ , sendo cada número par  $k \pmod{12}$  conectado a  $k + 3 \pmod{12}$ . O segundo será denotado por

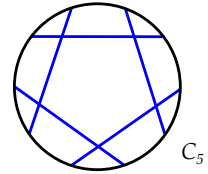
Entre os 9749 diagramas de cordas com 7 cordas, 2108 não são analíticos. O único exemplo básico não analítico é  $C_7$ .

No próximo capítulo, mostrarei que meu computador encontrou todos os diagramas não analíticos básicos.

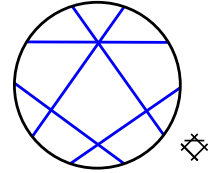
### Diagramas de cordas marcados

Para concluir este capítulo e colocá-lo em um contexto mais geral, gostaria de descrever uma estrutura que está por trás da cortina. Será conveniente primeiro introduzir um leve fortalecimento da noção de diagramas de corda analíticos.

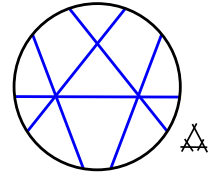
Quando provamos que nosso algoritmo decide se um diagrama é analítico, o ponto principal é a possibilidade de inserir um novo ramo. Acontece que singularidades mais complicadas também podem ser inseridas, como explicarei agora: Considere uma desingularização de alguma curva  $C$  como antes, para que tenhamos uma superfície  $S$ , um divisor  $E$  e uma coleção de curvas suaves  $\beta_1, \dots, \beta_n$  intersectando  $E$  transversalmente em um número finito de pontos,  $p_1, \dots, p_n$ , sendo  $n$  é o número de ramos reais. Escolha um desses pontos, digamos,  $p_1$ . Escolha agora algumas outras curvas singulares  $C_1$ , com  $n_1$  ramos reais e



$C_5$

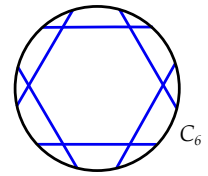


✖

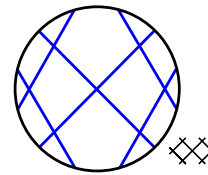


△

Os 3 diagramas de cordas básicos não analíticos com 5 cordas.



$C_6$



✖

Os 2 diagramas de cordas não analíticos com 6 cordas

suponha que não contém o eixo  $y$ . Exclua  $\beta_1$  e substitua por uma cópia de  $C_1$  na superfície  $S$ , de modo que o eixo  $y$  para  $C_1$  seja mapeado para o divisor  $E$  e o ponto singular de  $C_1$  é mapeado para  $p_1$ . Agora podemos fazer um blow-up a união dessa cópia de  $C_1$  e  $\beta_2, \dots, \beta_n$ . O resultado é um novo ponto singular, com  $n + n_1 - 1$  ramos: um deles  $C$  foi substituído por uma cópia de  $C_1$ .

Vamos examinar o efeito desse tipo de operação nos diagramas de cordas associados. Observando o diagrama associado a  $C_1$ , vemos que o eixo  $y$  decompõe a palavra de comprimento  $2n_1$  em dois componentes, *Esquerda* e *Direita*. No novo diagrama de cordas com  $2(n + n_1 - 1)$  letras, um par de letras idênticas do diagrama antigo com  $2n$  letras foi substituído por dois intervalos, que são *Esquerda* e *Direita*. Devemos ter cuidado, no entanto, que neste processo, as ordens das letras *Esquerda* e *Direita* poderiam ter sido revertidas. De fato, as duas interseções da borda orientada de  $S$  com  $\beta_1$  podem ser de sinais diferentes. Além disso, a inserção de  $C_1$  em  $S$  pode ser feita de quatro maneiras, uma vez que  $S$  não é orientável e  $E$  não é orientado.

É claro que, podemos proceder da mesma maneira com todos os outros ramos do  $C$ , usando outras curvas singulares  $C_2, \dots, C_n$ .

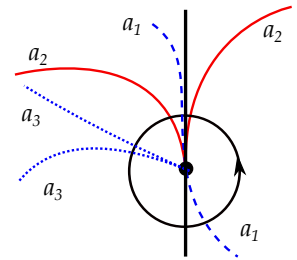
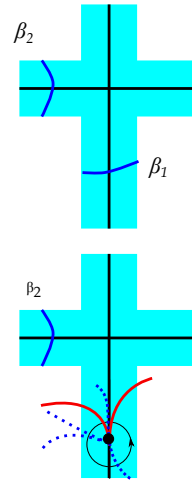
Todas essas observações sugerem a seguinte definição:

**Definição.** Um diagrama de cordas marcado é uma coleção de  $2n$  pontos distintos  $a_1^{\pm 1}, \dots, a_n^{\pm 1}$  na união de dois lados opostos de um quadrado  $\{-1, 1\} \times [-1, 1]$  (a menos de homeomorfismos, preservando a orientação de cada lado).

Observe os recursos adicionais se compararmos com diagramas padrão. Os diagramas marcados têm uma parte direita e uma esquerda. Além disso, cada corda  $a_i$  é agora rotulada com um número  $i$  de 1 a  $n$  e é orientada de  $a_i^{-1}$  a  $a_i^{+1}$ .

Deixe-me denotar por  $\mathcal{AMC}$  o conjunto dos diagramas de cordas marcados que são *analíticos*, i.e., que surgem de alguma curva analítica  $F(x, y) = 0$  que não contém o eixo  $y$ . Observe que a analiticidade de um diagrama marcado não depende nem da orientação das cordas nem da rotulagem.

Agora espero que o leitor tenha adivinhado a estrutura de operad em  $\mathcal{AMC}$ . Seja  $w$  algum diagrama analítico de cordas marcado com  $n$  cordas. Dados  $n$  diagramas analíticos de cordas



*Right* =  $a_1 a_2$ , *Left* =  $a_1 a_2 a_3 a_3$   
 Note que *Esquerda* ou *Direita* podem estar vazios.

Em uma equação  $F(x, y) = 0$ , podemos substituir  $(x, y)$  por  $(-x, y)$  ou  $(x, -y)$  ou  $(x, xy)$ . A transformação

$$(x, y) \mapsto (x, xy)$$

preserva cada reta vertical, colapsa o eixo  $x = 0$  para a origem e inverte a orientação no  $x < 0$ . Mas isso não é uma surpresa: exatamente uma aplicação de blow-down. O quadrado dessa transformação preserva a orientação em cada reta vertical (por  $x \neq 0$ ).

(marcados)  $w_1, \dots, w_n$ , com  $k_1, \dots, k_n$  cordas, defina a ação de  $w$  em  $(w_1, \dots, w_n)$  da seguinte maneira: Desenhe  $w$  e engrosse cada corda  $a_i$  de  $w$ , criando retângulos. Use os  $a_i^{-1}$  e  $a_i^{+1}$  como os lados esquerdo e direito desses novos retângulos. Agora insira  $w_1, \dots, w_n$  nesses retângulos respeitando os rótulos e as orientações. Renomeie as cordas, de 1 para  $k_1 + k_2 + \dots + k_n$ , usando a ordenação lexicográfica. O resultado é outro diagrama de cordas analítico marcado, pois essa operação corresponde à inserção de curvas analíticas anteriormente descrita.

Portanto,  $\mathcal{AMC}$  tem, de fato, uma estrutura natural de operad. Para completar, vamos observar que a ação natural dos grupos simétricos, permutando os rótulos dos cordas, mostra que é realmente uma operação simétrica.

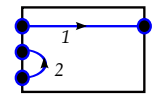
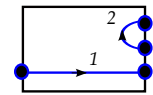
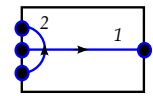
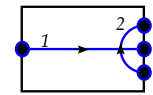
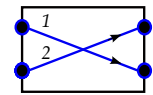
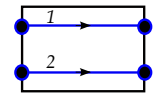
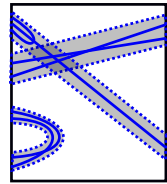
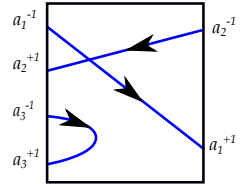
*Vamos limitar o número de diagramas de cordas*

Seria ótimo ter algumas informações precisas sobre o número  $a_n$  de diagramas analíticos com  $n$  cordas. Por exemplo, uma fórmula explícita para a série geradora  $\sum a_n t^n$  daria a exata taxa de crescimento exponencial de  $A_n$ . Infelizmente, não consegui calcular esta série ☹. Nesta seção, mostro que o lema fundamental fornece pelo menos uma cota razoável.

Considere uma árvore planar binária enraizada finita. Equipe cada um de seus vértices interiores (incluindo a raiz) com um dos seis exemplos de diagramas marcados com duas cordas representadas na margem. Por inserções recursivas dos diagramas dos irmãos no diagrama de seus pais, produz-se um diagrama analítico marcado e, portanto, um diagrama analítico, esquecendo os rótulos, as orientações e os dois lados do quadrado.

Afirmo que todos os diagramas analíticos com  $n$  cordas são produzidos com esta receita, que vale para  $n = 2$ , pois os dois diagramas com duas cordas (entrelaçadas ou não) aparecem quando alguém esquece a marcação nos seis exemplos. Agora seja  $w$  algum diagrama analítico com  $n + 1$  cordas e aplique o lema fundamental. Portanto, encontramos  $\dots aa \dots$ ,  $\dots ab \dots ba \dots$ ,  $\dots ab \dots ab \dots$  ou  $\dots b \dots aba \dots$  no diagrama. No caso de  $aa$ , chame  $b$  a letra que vem antes de  $a$  na ordem cíclica. Nosso algoritmo

O papel das rotulagens e das orientações é simplesmente fornecer as informações relevantes sobre qual diagrama marcado está inserido em cada corda e de que maneira é essa inserção.



exclui  $a$  e produz um diagrama analítico  $\bar{w}$  com  $n$  cordas, para o qual podemos aplicar a indução. Isto significa que  $\bar{w}$  pode ser “vestido” com rótulos e orientações de tal forma que é produzido por uma árvore binária, como acima. Nosso diagrama  $w$  é obtido de  $\bar{w}$ , substituindo uma corda por duas delas. É fácil verificar se nossos seis exemplos são suficientes para realizar essa duplicação usando uma inserção no operad. Portanto,  $w$  é construído a partir de uma árvore binária com  $n + 1$  folhas com a mesma receita.

Uma árvore binária enraizada com  $n$  folhas tem  $n - 1$  vértices interiores (incluindo a raiz) de modo que temos  $6^{n-1}$  etiquetas possíveis nos vértices interiores. O número de árvores binárias planares com  $n$  folhas é dado pelo  $(n - 1)$ -ésimo número de Catalan. Portanto, obtemos a seguinte estimativa:

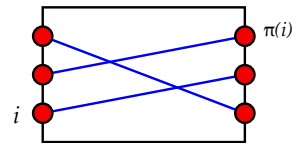
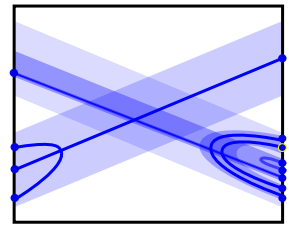
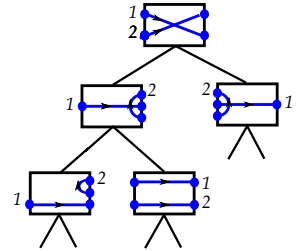
**Teorema.** O número  $A_n$  de diagramas de cordas analíticos com  $n$  cordas é menor que  $6^{n-1}$  vezes o número  $(n - 1)$ -ésimo número de Catalan  $C_{n-1}$ .

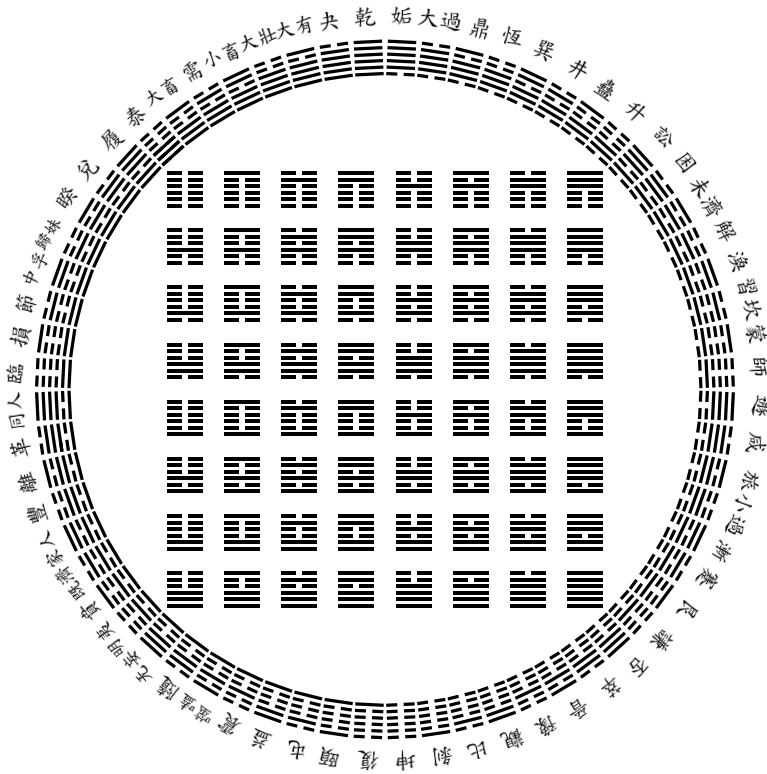
Lembre-se que  $\frac{1}{n} \log C_n$  converge para  $\log 4$  quando  $n$  tende para o infinito. Portanto,

$$\limsup \frac{1}{n} \log A_n \leq \log(24).$$

Note que qualquer permutação de letras  $n$  pode ser vista como um diagrama com  $n$  cordas, de forma que todos as suas cordas conectem pontos em ambos os lados de um quadrado. Em particular, as permutações separáveis produzem diagramas analíticos marcados. Essa situação dá um limite inferior para o crescimento de  $A_n$ , uma vez que já contamos permutações separáveis.

$$\liminf \frac{1}{n} \log A_n \geq \log(3 + 2\sqrt{2}).$$

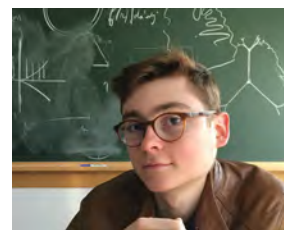




Os  $2^6 = 64$  hexagramas chineses consistem em seis barras horizontais que podem ser conectadas ou desconectadas. Eles apareceram no *I Ching* — o livro das Mudanças — escrito há mais de 2500 anos e são comumente usados como uma ferramenta de adivinhação. Originalmente, eles foram ordenados de uma maneira misteriosa, geralmente atribuída ao rei Wen, que os estudiosos ainda estão tentando decifrar. Mil anos atrás, Shao Yong ordenou-os como mostrado na foto, em um círculo e em um quadrado. Em 1701, o jesuíta Joachim Bouvet enviou uma cópia dessa configuração para Leibniz, que a explicou em termos de expansões binárias e escreveu uma das primeiras exposições sistemáticas de aritmética na base 2. Este é um exemplo interessante de interação entre a filosofia oriental e a ciência ocidental. Vou discutir esses I Chings um pouco mais na seção final deste capítulo. †

## *Diagramas de cordas analíticos: grafos de entrelaçamento*

OS ASPECTOS AGRADÁVEIS DOS PASSEIOS ALEATÓRIOS são as muitas surpresas que eles nos oferecem. Christopher-Lloyd Simon é estudante de graduação na *École Normale Supérieure de Lyon* e gentilmente concordou em ler o primeiro rascunho deste livro. Enquanto ele estava lendo uma versão preliminar do capítulo anterior, ele teve a brilhante ideia de transferir a discussão dos diagramas de cordas para seus *grafos de entrelaçamento* associados. Já encontramos esse conceito em nosso estudo das palavras de Gauss associadas a curvas imersas genéricas no plano. Dado um diagrama de cordas, o conjunto de vértices de seu grafo de entrelaçamento é simplesmente o conjunto de cordas, e as arestas conectam cordas entrelaçadas (isto é, intersectando-as). Nem todo grafo vem de um diagrama, podendo vir de vários diagramas. No entanto, os grafos de entrelaçamento provenientes de diagramas analíticos revelaram-se fáceis de analisar. A cereja do bolo é que esses grafos foram introduzidos há quarenta anos em um contexto totalmente diferente, sendo muito bem compreendidos. Graças a essa nova perspectiva, obteremos a lista completa de diagramas básicos de cordas não analíticos.



Christopher-Lloyd Simon.

### *Voltando às permutações separáveis*

A fim de motivar o que se segue, vamos revisitar rapidamente a situação muito mais simples de intercâmbios polinomiais



(também denominadas de permutações separáveis), que já examinamos nos primeiros capítulos.

Seja  $\pi$  uma permutação de  $\{1, \dots, n\}$ . O grafo de permutação  $G(\pi)$  associado a  $\pi$  tem  $\{1, \dots, n\}$  como vértices, e uma aresta conecta  $i$  e  $j$  se  $\pi$  inverte a ordem de  $(i, j)$ . Esse é também o grafo de entrelaçamento do diagrama de cordas marcado associado, com letras  $n$  em cada lado.

Sabemos que se  $\pi$  é um intercâmbio polinomial, pelo menos dois inteiros consecutivos possuem imagens consecutivas. Portanto, as cordas correspondentes têm a seguinte propriedade: qualquer corda que intersecta uma delas intersecta a outra.

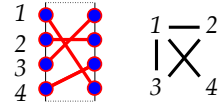
considerando o grafo  $G(\pi)$ , há a seguinte definição: Dois vértices  $x, y$  em um grafo são chamados *gêmeos* se tiverem os mesmos vizinhos (diferentes de  $x$  ou  $y$ ). Eles são chamados *gêmeos verdadeiros ou falsos* dependendo da existência de uma aresta que os conecta. Dois gêmeos em um grafo podem ser mesclados em um único vértice, produzindo um grafo menor com um vértice a menos.

Portanto, o grafo  $G(\pi)$  vindo de um intercâmbio polinomial  $\pi$  contém pelo menos dois gêmeos, correspondendo a dois inteiros consecutivos  $i, i + 1$  tal que  $\pi(i + 1) = \pi(i) \pm 1$ . Mesclar os gêmeos no gráfico equivale a mesclar os dois elementos  $i, i + 1$ . Os intercâmbios polinomiais são caracterizados pelo fato de que a iteração desse procedimento de fusão leva à permutação trivial com  $n = 1$ .

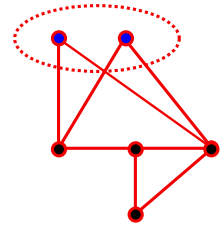
**Definição:** Um grafo finito é chamado de *cografo* se puder ser reduzido a um grafo de apenas 1 vértice trivial, fundindo gêmeos sucessivamente.

**Proposição.** *Uma permutação é um intercâmbio polinomial se, e somente se, seu grafo de permutação for um cografo.*

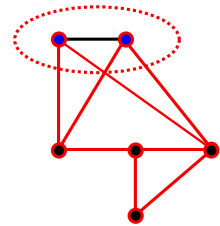
Acabei de explicar por que o grafo de permutação de um intercâmbio polinomial é um cografo. Para provar o contrário, basta mostrar que, se  $G(\pi)$  é um cografo, existem dois inteiros consecutivos com imagens consecutivas. A prova é por indução sobre  $n$ . Se  $i < j$  são falsos (resp. verdadeiro) gêmeos, a imagem por  $\pi$  do intervalo  $\{i, i + 1, \dots, j\}$  é  $\{\pi(i), \pi(i + 1), \dots, \pi(j)\}$  (resp.  $\{\pi(j), \pi(j + 1), \dots, \pi(i)\}$ ). Se  $j \geq i + 2$ , a imagem  $\pi(\{i, \dots, j - 1\})$



Uma permutação e seu grafo.



Gêmeos falsos. Observe que dois vértices solitários são gêmeos falsos. Não sou responsável pela terminologia clássica, mas parece-me mais conveniente do que o termo gêmeos dizigóticos (ou fraternos), aproximando-se mais do francês “faux jumeaux”.



Gêmeos verdadeiros, (idênticos ou monozigóticos), vrais jumeaux.

A terminologia “cografo” surge do fato de que o complemento de um cografo é também um cografo. Um grafo  $G$  e seu complemento  $\bar{G}$  têm os mesmos vértices e dois vértices são adjacentes em  $\bar{G}$  se não forem adjacentes em  $G$ .

também é um intervalo e aplicamos a hipótese de indução à restrição de  $\pi$  para  $\{i, \dots, j - 1\}$  para que se encontrem dois inteiros consecutivos com imagens consecutivas.  $\square$

Os cografos foram introduzidos na década de 1970 com diferentes nomes: ( $D^*$ -grafos, hereditários de Dacey e grafos de 2-paridade <sup>162</sup>). Eles não são muito difíceis de descrever. Deixe-me listar algumas de suas propriedades e as provas (elementares) ficam para você, leitor.

A seguir, todos os grafos são finitos, sem laços e sem arestas múltiplas. Um grafo conexo define um espaço métrico em seu conjunto de vértices. A distância entre dois vértices é, por definição, o comprimento do caminho mais curto que os conecta.

Os especialistas em teoria dos grafos dizem que um subgrafo  $H$  de um grafo  $G$  é *induzido* se qualquer aresta de  $G$  conectando dois vértices de  $H$  for também uma aresta de  $H$ .

**Teorema.** *As seguintes propriedades de um grafo finito  $G$  são equivalentes.*

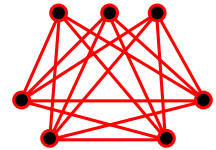
1.  $G$  é um cografo.
2. Qualquer subgrafo induzido e conexo de  $G$  tem um diâmetro menor ou igual que 2.
3.  $G$  é o grafo de permutação de algum intercâmbio polinomial.
4. Não há subgrafo  $P_4$  induzido com quatro vértices como na margem.

Note que os grafos de permutação associados às duas permutações proibidas de Kontsevich  $(2, 4, 1, 3)$  e  $(3, 1, 4, 2)$  são isomorfos a  $P_4$ .

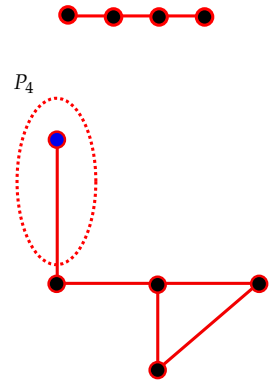
### Grafos colapsáveis

A partir de uma árvore, você pode tirar as folhas e fazê-lo novamente até que a árvore tenha sido completamente despida. Digamos que um vértice em um grafo seja *pendente* se for adjacente a um único vértice. Qualquer árvore pode ser construída por adições sucessivas de vértices pendentes, começando a partir da árvore com apenas um vértice.

<sup>162</sup> A. Brandstädt, V. B. Le, and J. P. Spinrad. *Graph classes: a survey*. SIAM Monographs on Discrete Mathematics and Applications. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999.



Um cografo.



Vértice pendente.

**Definição.** Um grafo finito é *colapsável* se puder ser reduzido a um grafo de um vértice, aplicando dois tipos de operações elementares: excluindo um vértice pendente e mesclando gêmeos.

Se você é mais construtivo do que destrutivo, pode expressar a mesma coisa de outra maneira. Comece com o grafo trivial com um vértice e aplique dois tipos de operações: adicionando um vértice pendente ou criando um par de gêmeos. A segunda operação consiste simplesmente em duplicar um vértice e conectar o gêmeo recém-nascido ao resto do grafo como o vértice original. Então, decida se você quer gêmeos verdadeiros ou falsos.

O ponto chave é o seguinte:

**Proposição.** Um diagrama de cordas é analítico se, e somente se, o seu grafo de entrelaçamento for colapsável.

Isso seguirá a partir da descrição algorítmica dos diagramas analíticos apresentados no capítulo anterior.

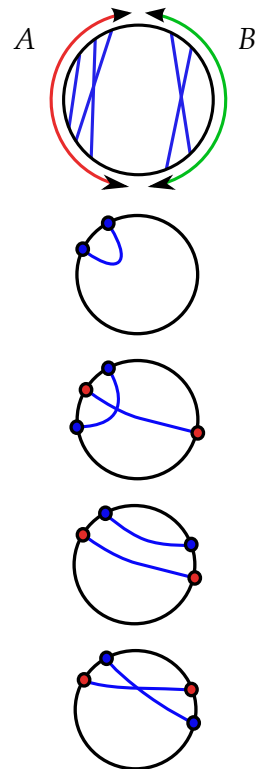
Antes da prova, deixe-me fazer uma observação elementar, como aperitivo.

Seja  $w$  um diagrama e  $A$  seja um subconjunto de suas  $2n$  letras num círculo. Eu direi que  $A$  é *estável* sob  $w$  se qualquer corda com uma extremidade em  $A$  tiver sua outra extremidade em  $A$ . Dito de forma diferente,  $A$  é um diagrama de subcordas  $w_A$  de  $w$ .

Suponha que haja um intervalo  $A$  que seja estável sob  $w$  e que seja  $B$  seu complemento. Claramente, o grafo de entrelaçamento  $G(w)$  de  $w$  é a união disjunta dos grafos  $G(w_A)$  e  $G(w_B)$  de  $w_A$  e  $w_B$ . Segue-se que  $G(w)$  é colapsável se, e somente se,  $G(w_A)$  e  $G(w_B)$  forem colapsáveis. Nosso algoritmo mostra que se  $w_A, w_B$  são analíticos, então, é  $w$ . Por outro lado, se  $w$  é analítico, o mesmo acontece com os subdiagramas  $w_A$  e  $w_B$ .

Vamos provar agora a proposição:

Comece com um diagrama analítico  $w$ . Se duas cordas de  $w$  são paralelas ou antiparalelas, os vértices associados no grafo de entrelaçamento são gêmeos e nosso algoritmo os reúne. Uma forquilha fornece um vértice pendente no grafo e o algoritmo exclui a corda curta e mantém a alça. Uma corda solitária define



um vértice pendente, removido pelo algoritmo. Segue-se que o grafo de entrelaçamento associado a um diagrama analítico é colapsável.

*Para a recíproca, mostramos que todo diagrama  $w$  cujo grafo de entrelaçamento  $G(w)$  é colapsável contém uma corda solitária, uma forquilha ou um par de cordas paralelas ou antiparalelas.*

Uma corda solitária em um diagrama corresponde a duas letras idênticas consecutivas  $\dots aa \dots$  na palavra cíclica.

Uma forquilha corresponde a uma subpalavra da forma  $\dots aba \dots$ .

Um par de cordas paralelas (resp. Antiparalelas) corresponde a  $\dots ab \dots ba \dots$  (respectivamente  $\dots ab \dots ab \dots$ ).

Nossa prova será por contradição: Considere um possível contra-exemplo  $w$  para a asserção anterior *com um número mínimo de cordas*. Então,  $G(w)$  é colapsável e  $w$  não contém cordas solitárias, nenhum par de cordas paralelas ou antiparalelas e nenhuma forquilha.

Como  $G(w)$  é colapsável, existe um vértice  $\alpha$  que é solitário, ou pendente, ou faz parte de um par de gêmeos. Seja  $\bar{w}$  o diagrama obtido excluindo  $\alpha$  de  $w$ .

É claro que  $G(\bar{w})$  é colapsável para que, por minimalidade,  $\bar{w}$  contenha uma subpalavra  $\dots aa \dots$  ou  $\dots aba \dots$  ou  $\dots ab \dots ba \dots$  ou  $\dots ab \dots ab \dots$ . O problema é que estas são subpalavras de  $\bar{w}$  e não de  $w$ , que também contém duas cópias da letra  $\alpha$ , que poderiam entrar nas subpalavras acima.

Observe que, por minimalidade, qualquer intervalo estável sob  $w$  é vazio ou tudo.

A priori:

- 0, 1 ou 2 letras  $\alpha$  poderiam infiltrar-se na subpalavra;
- A subpalavra de  $\bar{w}$  poderia corresponder a uma corda solitária, a uma forquilha ou a um par de cordas paralelos ou antiparalelos;
- $\alpha$  pode ser solitário, pendente ou gêmeo, verdadeiro ou falso, em  $G(w)$ .

Isso faz com que tenhamos  $3 \times 4 \times 4$  casos para examinar! Felizmente, muitos casos podem ser estudados simultaneamente.

Esta prova caso a caso não é particularmente agradável. Você pode pulá-la se quiser, mas se fizer isso, você deve simpatizar-se comigo, que teve que listar todos os casos um por um.

1. Se nenhuma letra  $\alpha$  chegar nas subpalavras acima, não há problema: nossa corda solitária ou forquilha, ou um par de cordas paralelas, ou antiparalelas em  $\bar{w}$  tem a mesma propriedade para  $w \not\equiv$ .

2. Se  $\alpha$  é solitário em  $G(w)$ , isso significa que nenhuma corda intersecta  $\alpha$ . Portanto,  $\alpha$  decompõe o círculo em dois intervalos estáveis, que devem ser vazios  $\not\equiv$ .

3. Se duas letras  $\alpha$  entrarem, elas não podem ocorrer como letras consecutivas, pois isso forçaria à corda  $\alpha$  a ser solitária em  $w \not\equiv$ .

Então, ainda neste caso, temos que olhar

$$\dots\alpha b \alpha a \dots, \text{ or } \dots\alpha \alpha b \dots b \alpha a \dots, \text{ or } \dots\alpha \alpha b \dots \alpha \alpha b \dots$$

Isso produz, respectivamente, uma forquilha  $(\alpha, b)$ , um par de cordas paralelas  $(\alpha, b)$  ou cordas antiparalelas  $(\alpha, b)$  em  $w \not\equiv$ .

4. Inserindo um  $\alpha$  em uma corda solitária, rende  $\dots\alpha \alpha a \dots$  que produz uma forquilha em  $w$  com alça  $\alpha \not\equiv$ .

Até agora, não usamos o fato de que  $\alpha$  é pendente ou gêmeo. Isso será usado nos casos restantes, quando um único  $\alpha$  entra numa forquilha ou um par de cordas paralelas ou antiparalelas de  $\bar{w}$ .

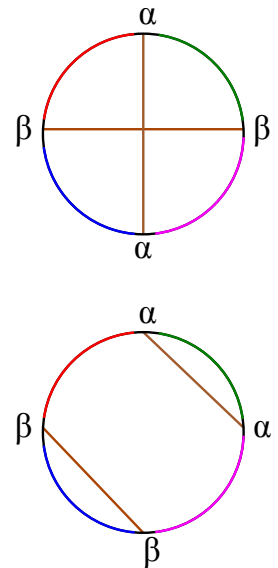
Se  $\alpha$  é pendente, seja  $\beta$  a única corda em  $w$  que intersecta  $\alpha$ . Se  $\alpha$  tiver alguns irmãos gêmeos, vamos denotar um deles por  $\beta$ . No caso, as duas cordas  $\alpha, \beta$  determinam quatro intervalos no círculo, excluindo  $\alpha, \beta$ , que chamarei de setores. Se  $\alpha$  é pendente, a união de dois setores que estão no mesmo lado de  $\alpha$  é estável. Se  $\alpha, \beta$  são gêmeos, a união de setores opostos é estável.

5. Suponha agora que uma única letra  $\alpha$  insira  $\dots\alpha b a \dots$ , ou  $\dots\alpha b \dots b a \dots$ , ou  $\dots\alpha b \dots \alpha b \dots$  e uma dessas duas letras  $a, b$  é igual a  $\beta$ . Então, as duas letras  $\alpha$  e  $\beta$  são consecutivas em  $w$ . Isto implica que um dos setores está vazio.

5-1. No caso pendente, isso implica que o outro setor, do mesmo lado de  $\alpha$ , é um intervalo estável e, portanto, também vazio. Logo,  $\alpha, \beta$  é uma forquilha em  $w$  com o alça  $\beta \not\equiv$ .

5-2. No caso gêmeo, isto implica que o setor oposto é um intervalo estável e, portanto, vazio. Então,  $\alpha, \beta$  é um par de cordas paralelas ou antiparalelas em  $w \not\equiv$ .

Uso o símbolo  $\not\equiv$  para significar “contradição”.



6. Finalmente, suponha que uma única letra  $\alpha$  entre  $\dots aba\dots$ , ou  $\dots ab\dots ba\dots$ , ou  $\dots ab\dots ab\dots$  e que nenhuma das duas letras  $a, b$  é igual a  $\beta$ .

6-1. Suponha que  $(a, b)$  sejam cordas paralelas ou antiparalelas em  $\overline{w}$ . Inserindo um  $\alpha$  em  $ab$ , produz-se  $\dots a\alpha b\dots ba\dots$  ou  $\dots a\alpha b\dots ab\dots$ . Das letras consecutivas  $ba$  (ou  $ab$ ), segue-se que essas duas ocorrências estão no mesmo lado da corda  $\alpha$ . Das letras consecutivas  $a\alpha b$  em  $w$ , segue-se que essas outras ocorrências de  $a$  e  $b$  estão em lados diferentes de  $\alpha$ . Assim, a corda  $\alpha$  intersecta apenas uma das cordas  $a, b$ .

6-1-1. Se  $\alpha$  é pendente, isso força  $a$  ou  $b$  a ser igual a  $\beta$   $\cancel{!}$ .

6-1-2. Se  $\alpha, \beta$  são gêmeos, isso não é possível  $\cancel{!}$ .

6-2. Suponha que  $(a, b)$  seja uma forquilha em  $\overline{w}$ . Inserindo uma letra  $\alpha$  em  $\dots aba\dots$  tem-se  $\dots a\alpha ba$  or  $\dots ab\alpha a$ , então, a corda  $\alpha$  deve intersectar a corda  $a$ .

6-2-1. Se  $\alpha$  é pendente, isso força  $a = \beta$   $\cancel{!}$ .

6-2-2. Se  $(\alpha, \beta)$  forem gêmeos, as letras consecutivas  $a\alpha ba$  mostram que as duas letras  $a$  estão em lados opostos de  $\alpha$ , portanto, em setores opostos. Segue-se que  $b = \beta$   $\cancel{!}$ .

Isso finaliza a prova.

Ufa!  $\square$

Agora, precisamos entender a natureza dos grafos colapsáveis.

### Grafos colapsáveis, distância hereditária

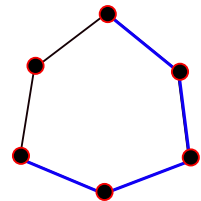
Os grafos colapsáveis foram definidos por vários autores há quarenta anos, sob diferentes nomes, com motivações muito diferentes. Veremos que esses grafos estão muito próximos de serem árvores.

Howorka <sup>163</sup> definiu grafos de distância hereditária em 1977.

**Definição.** Um grafo finito  $G$  é de *distância hereditária* se para cada subgrafo induzido conexo  $H \subset G$ , a distância entre dois vértices de  $H$  em  $H$  é igual à distância entre os mesmos vértices em  $G$ .

Por exemplo, uma árvore é de *distância hereditária* e um ciclo de comprimento de pelo menos cinco não é. Basta escolher  $H$  como o subgrafo induzido definido por um caminho dentro do

<sup>163</sup> E. Howorka. A characterization of distance-hereditary graphs. *Quart. J. Math. Oxford Ser. (2)*, 28(112):417-420, 1977.



Um ciclo de comprimento  $\geq 5$  que não é de distância hereditária.

ciclo cujo comprimento é maior que a metade do comprimento do ciclo (em azul na figura).

Considere um grafo finito e escolha um comprimento para cada aresta, que pode ser qualquer número real positivo. Defina o comprimento de um caminho como a soma dos comprimentos de suas arestas e a distância entre dois vértices como o menor comprimento de um caminho conectando-os. Estamos diante de um *grafo métrico*.

Estamos procurando uma caracterização de espaços métricos (geralmente chamadas de *árvores métricas*) que surgem dessa maneira a partir de *árvores*. Aqui está a resposta: Seja  $(E, d)$  um espaço métrico finito. Escolha quatro pontos  $x_1, x_2, x_3, x_4$  em  $E$  e calcule as somas dos comprimentos dos três pares de diagonais:

$$d(x_1, x_2) + d(x_3, x_4); d(x_1, x_3) + d(x_2, x_4); d(x_1, x_4) + d(x_2, x_3).$$

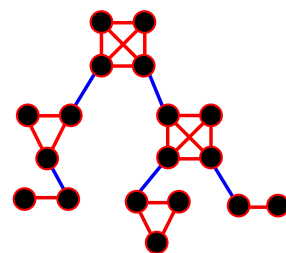
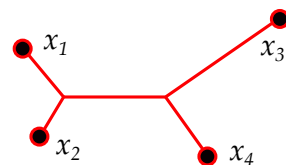
Sejam  $s$  (resp.  $M, l$ ) o menor (resp. médio, maior) desses três números:  $s \leq m \leq l$ . Acontece que um espaço métrico finito é isométrico a um subconjunto de uma árvore métrica  $m = l$  para cada quádrupla de pontos. Isso não é difícil de provar, e eu deixo como um exercício M25. O leitor preguiçoso pode ver a prova neste breve artigo <sup>164</sup>.

Nós devemos ter cuidado. Um grafo, onde todas as arestas têm comprimento 1, pode ser isométrico a um subconjunto de uma árvore métrica sem ser em si uma árvore. Veja o exemplo na margem.

Na teoria dos grafos, esses grafos são chamados *grafos de blocos*. Para construí-los, comece com uma árvore, delete alguns dos seus vértices e substitua-os por *cliques*, ou seja, grafos finitos com todos os pares de vértices são adjacentes, como na figura. Sugiro que meu leitor prove que isso é de fato uma caracterização de grafos de blocos (M15 e, em caso de emergência, ver este livro <sup>165</sup>).

Na década de 1980, Gromov desenvolveu uma teoria geométrica para *espaços hiperbólicos* que teve uma influência muito forte sobre a teoria combinatória e geométrica dos grupos (infelizmente, essa parada não faz parte do nosso passeio).

A definição é a seguinte: Um espaço métrico  $(E, d)$  é chamado

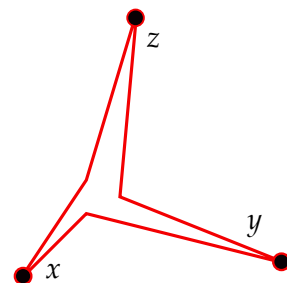


Um grafo bloco.

<sup>164</sup> P. Buneman. *A note on the metric properties of trees*. *J. Combinatorial Theory Ser. B*, 17:48–50, 1974.

<sup>165</sup> F. Harary. *A characterization of block-graphs*. *Canad. Math. Bull.*, 6:1–6, 1963.

Um espaço métrico  $(E, d)$  é *geodésico* se para cada par de pontos  $(x, y)$  existir um mergulho *isométrico*  $i : [0, d(x, y)] \rightarrow E$  tal que  $i(0) = x$  and  $i(d(x, y)) = y$ .



hiperbólico se existir algum  $\delta \geq 0$  tal que para cada quádrupla de pontos como acima,  $m$  e  $l$  são “quase iguais”, i. e.,  $l - m \leq \delta$ . Observe que qualquer espaço métrico finito é trivialmente hiperbólico (para  $\delta$  suficientemente grande), de modo que esse conceito é relevante somente para geometria em grande escala.

Existem muitas formulações equivalentes dessa propriedade: a mais popular (para espaços métricos geodésicos) é aquela que todos os triângulos geodésicos são *magros*. Considere três pontos  $x, y, z$  e escolha três geodésicas  $[x, y], [x, z], [y, z]$  conectando-os. Cada ponto em  $[x, y]$  deve estar a uma distância uniformemente limitada da união  $[x, z] \cup [y, z]$ , independentemente da escolha de  $x, y, z$  (veja a figura).

Esse conceito é notavelmente robusto. Por exemplo, o recobrimento universal de uma variedade Riemanniana compacta de curvatura negativa é hiperbólica. Esses espaços métricos são bem aproximados por árvores, de forma quantitativa. Para mais sobre esta teoria, o leitor é encorajado a ler este livro <sup>166</sup>.

Em 1986, Bandelt e Mulder publicaram um artigo <sup>167</sup> propondo caracterizações puramente métricas para grafos hereditários, próximas das condições de hiperbolicidade de Gromov.

**Definição.** Um grafo finito  $G$  é tipo árvore se para cada 4-tupla de vértices  $x_1, x_2, x_3, x_4$  dois dos três números seguintes são iguais:

$$d(x_1, x_2) + d(x_3, x_4); \quad d(x_1, x_3) + d(x_2, x_4); \quad d(x_1, x_4) + d(x_2, x_3)$$

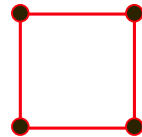
Meu leitor provavelmente adivinhou que todas essas definições acabam sendo equivalentes.

**Teorema.** Seja  $G$  um grafo finito. As propriedades a seguir são equivalentes:

1.  $G$  é o grafo de entrelaçamento de algum diagrama analítico de cordas;
2.  $G$  é colapsável;
3.  $G$  é de distância hereditária;
4.  $G$  é tipo árvore;

<sup>166</sup> E. Ghys and P. de la Harpe, editors. *Sur les groupes hyperboliques d'après Mikhael Gromov*, volume 83 of *Progress in Mathematics*. Birkhäuser Boston, Inc., Boston, MA, 1990.

<sup>167</sup> H.-J. Bandelt and H. M. Mulder. [Distance-hereditary graphs](#). *J. Combin. Theory Ser. B*, 41(2):182–208, 1986.



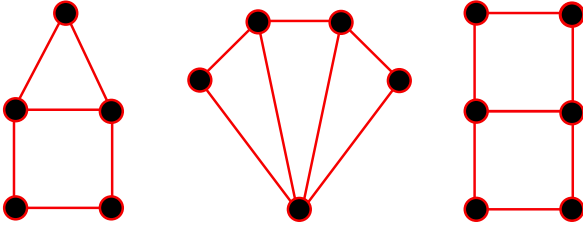
Um ciclo de comprimento quatro é tipo árvore, mas não é uma árvore.

*Exercício:* Mostre que um grafo tipo árvore é hiperbólico no sentido de Gromov com  $\delta=2$ .



5.  $G$  não contém um ciclo de comprimento de pelo menos cinco, ou uma casa, ou uma gema, ou um dominó, como um subgrafo induzido.

Uma casa, uma gema e um dominó são mostrados abaixo.



Todas as equivalências no teorema anterior (exceto, é claro, o primeiro item) são provadas nos artigos acima mencionados. No entanto, vou propor em breve algumas provas elementares.

Agora é hora de colher os frutos do nosso trabalho e obter uma descrição bastante simples dos diagramas analíticos de cordas.

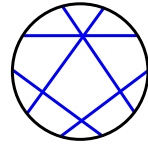
Você não deve se surpreender que os grafos de entrelaçamento de  $\diamond$ ,  $\triangle$ ,  $\diamond\diamond$  são a casa, a gema e o dominó.

*Exercício:* Mostre que  $\diamond\diamond$ ,  $\diamond$ ,  $\triangle$  são os únicos diagramas de cordas cujos grafos de entrelaçamento são a casa, o dominó e a gema.

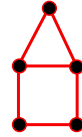
Da mesma forma, já descrevemos o diagrama de cordas não analítico  $C_n$  definido por  $\mathbb{Z}/2n\mathbb{Z}$  ( $n \geq 5$ ), existindo uma corda conectando  $2k$  e  $2k + 3$  (para  $k = 1, \dots, n$ ). De forma que seu grafo de entrelaçamento é um ciclo de comprimento  $n$ .

*Exercício:* Mostre que  $C_n$  é o único diagrama de cordas cujo grafo de entrelaçamento é um ciclo de comprimento  $n$ .

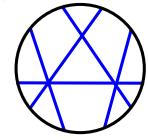
Por fim, observe que um diagrama de subcordas define um subgrafo induzido no grafo de entrelaçamento. Portanto, obtenemos uma descrição muito satisfatória dos diagramas analíticos de cordas. Uma vez que é um destaque em nosso passeio, imprimimos o seguinte teorema em azul:



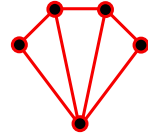
$\diamond$



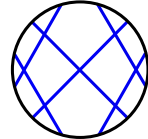
A casa.



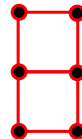
$\triangle$



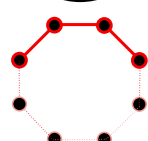
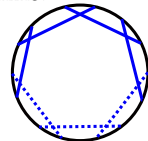
A gema.



$\diamond\diamond$

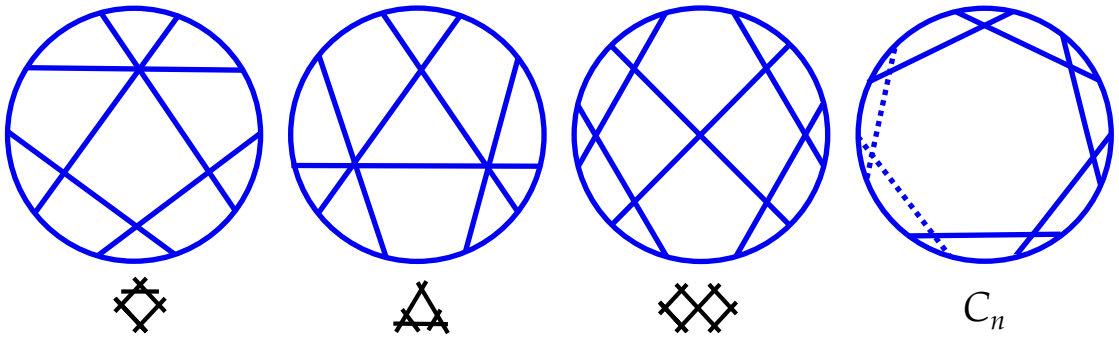


O dominó.



$C_n$  e seu grafo de entrelaçamento: o  $n$ -ciclo.

**Teorema.** *Um diagrama de cordas é analítico se, e somente se, não contiver  $\diamond\diamond$ ,  $\diamond$ ,  $\triangle$  ou  $C_n$  ( $n \geq 5$ ) com um subdiagrama de cordas*



Note a analogia completa com a nossa caracterização de intercâmbios polinomiais como as *permutações separáveis*, que são, por sua vez, precisamente aquelas que não contêm os exemplos de Kontsevich  $(2, 4, 1, 3)$  e  $(3, 1, 4, 2)$ .

### Algumas demonstrações

Apresento agora as provas das equivalências das definições da seção anterior. Eles são principalmente elementares, e eu sugiro que o leitor interessado tente prová-las sozinho. E nesse processo, é importante desenhar figuras porque, nesse caso específico, provavelmente foi mais desafiador encontrar as definições significativas do que provar sua equivalência.

No  $\triangle \diamond \square \circ$  é um subgrafo induzido  $\implies$  distância hereditária.

Seja  $H$  um subgrafo induzido de um grafo  $G$ . Conecte dois vértices  $p, q$  de  $H$  a uma distância  $n$  em  $H$  por um caminho  $c = (x_0, x_1, \dots, x_n)$  (com  $p = x_0$  e  $q = x_n$ ) em  $H$ . Dois vértices  $x_i, x_j$  são adjacentes se, e somente se,  $i, j$  forem consecutivos, caso contrário, haverá um atalho. Em outras palavras, o caminho  $c$  é induzido em  $G$ . Segue-se que, para mostrar que um grafo é de distância hereditária, devemos provar que a distância entre os pontos finais de qualquer caminho induzido em  $G$  é igual ao

comprimento do caminho.

Suponha que  $\triangle \nabla \square \boxtimes$  não é induzido em  $G$ . Tome o caminho induzido  $c_1 = (x_0, x_1, \dots, x_n)$  e mostremos, por indução em  $n$ , que a distância em  $G$  entre  $x_0$  e  $x_n$  é exatamente  $n$ .

[2cm]

Conecte  $x_0$  a  $x_n$  por um caminho mais curto  $c_2 = (y_0, y_1, \dots, y_l)$  em  $G$  (com  $y_0 = x_0$  e  $y_l = x_n$ ). Naturalmente,  $c_2$  também é induzido e  $d(y_0, y_i) = i$  para  $0 \leq i \leq l$ . Por indução,  $d(x_0, x_i) = i$  para  $0 \leq i \leq n - 1$ . Segue-se que  $l$  é igual a  $n - 2$ ,  $n - 1$  ou  $n$ . Mostramos que os dois primeiros casos não são possíveis. Suponha que  $l = n - 2$  ou  $n - 1$ . Por indução, podemos supor que os dois caminhos  $c_1, c_2$  somente se intersectam em seus pontos de extremos: qualquer outro ponto de interseção poderia ser usado como ponto de partida dos caminhos mais curtos  $c'_1$  e  $c'_2$ .

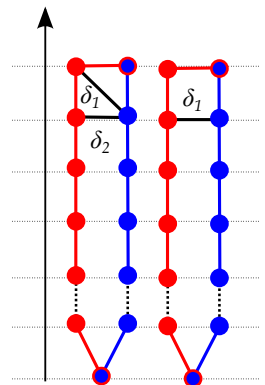
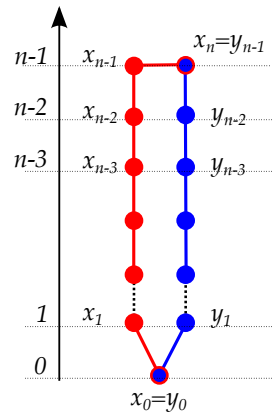
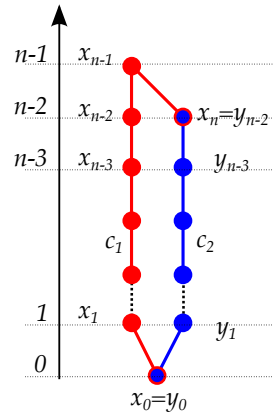
Desenhe uma imagem no plano de tal forma que a altura de um vértice de  $c_1$  ou  $c_2$  seja a distância de  $x_0$ . Os casos  $l = n - 2$  e  $l = n - 1$  são mostrados na margem. Vértices de  $c_1$  são vermelhos e vértices de  $c_2$  são azuis. A união de  $c_1$  e  $c_2$  define um ciclo  $c$  em  $G$ . O comprimento de  $c$  é pelo menos cinco. O ciclo  $c$  não pode ser induzido, pois não há ciclo induzido de comprimento  $\geq 5$ . Portanto, deve haver diagonais conectando vértices de  $c_1$  com vértices de  $c_2$ .

Pela desigualdade do triangular, a diferença de altura dos dois pontos finais de uma diagonal pode ser apenas  $-1, 0, 1$ . Além disso, as diagonais conectam pontos de cores diferentes.

Vamos ordenar as diagonais  $(x_i, y_j)$  de cima para baixo, ou seja,  $(x_i, y_j)$  é antes de  $(x_{i'}, y_{j'})$  se  $j > j'$  ou  $j = j'$  e  $i > i'$ .

Agora, tente construir a escada: uma diagonal de cada vez. A regra do jogo é a seguinte: Você tem que desenhar uma sequência ordenada de diagonais  $\delta_1, \delta_2, \dots$ , respeitando as condições acima e sem criar nenhuma  $\triangle \nabla \square \boxtimes$  induzida. Note que a diagonal  $\delta_k$  junto com a parte de  $c$  que está acima dela define um ciclo. Qualquer corda neste ciclo tem que ser uma das cordas previamente escolhidas  $\delta_1, \dots, \delta_{k-1}$ .

No caso  $l = n - 1$ , existem apenas duas possibilidades para a primeira diagonal  $\delta_1$ . Pode ser  $(x_{n-1}, y_{n-2})$  ou  $(x_{n-2}, y_{n-2})$ . No caso de  $l = n - 2$ , existe apenas uma possibilidade para  $\delta_1$ .



Em seguida, tente selecionar a segunda diagonal  $\delta_2$ , evitando  $\triangle \nabla \square \circ$ . Apenas uma das três opções de  $\delta_1$  permite que você faça isso.

Finalmente, tente desenhar a terceira diagonal, no único caso em que você poderia desenhar  $\delta_1, \delta_2$ . Não é possível continuar sem criar um dos grafos proibidos.  $\square$

*Distância hereditária  $\implies$  Colapsável.*

Deixe-me primeiro fazer uma observação elementar.

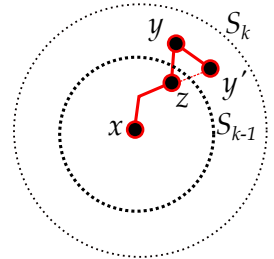
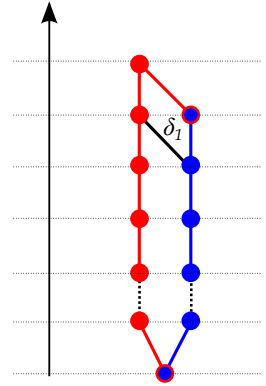
Escolha um vértice  $x$  em um grafo conexo de distância hereditária  $G$  e observe o maior  $k$  tal que a esfera  $S_k$  em  $G$  de raio  $k$  e centrada em  $x$  não está vazia. Seja  $C$  um componente conexo de  $S_k$ . Se  $C$  contém apenas um elemento, então, é um vértice pendente em  $G$ .

Escolha dois vértices  $y, y'$  em  $C$  que são adjacentes em  $C$ . Escolha algum ponto  $z$  que seja adjacente a  $y$ , à distância  $k - 1$  de  $x$ . Escolha uma cadeia  $c$  de comprimento  $k - 1$  de  $x$  a  $z$  e chame  $c'$  a cadeia de comprimento  $k + 1$  obtida adicionando a aresta entre  $z$  e  $y$  e de  $y$  a  $y'$ . Como a distância entre  $x$  e  $y'$  é exatamente  $k$ , esta cadeia não pode ser induzida e  $y'$  tem que ser adjacente a  $z$ . Isto implica que dois pontos em  $C$  são simultaneamente adjacentes ou simultaneamente não adjacentes a qualquer ponto  $z$  à distância  $k - 1$  de  $x$ . Segue-se que  $C$  não pode conter um caminho induzido  $P_4$  de comprimento 3, já que junto com  $z$  produziria uma gema em  $G$ , que não é de distância hereditária. Portanto,  $C$  é um cografo e, em particular, contém um par de gêmeos em  $G$ .  $\square$

*Colapsável  $\implies$  Tipo árvore.*

Fácil por indução: Pegue quatro pontos em um grafo, exclua um vértice pingente ou junte dois gêmeos. Um dos quatro pontos pode ser o vértice que foi removido. Se este for o caso, substitua-o pela outra extremidade da aresta removida. Olhe para os pontos correspondentes no grafo diminuído (tendo em conta, por exemplo, o fato de que dois dos nossos quatro pontos podem ser os dois gêmeos que foram fundidos). Aplique a hipótese de indução.  $\square$

*Tipo árvore  $\implies$  Não induzido  $\triangle \nabla \square \circ$ .*



Óbvio, pois verifica-se facilmente que nenhum desses exemplos de grafos são grafos tipo árvores. □

*Grafos completamente decomponíveis*

Grafos que podem ser levados para um ponto, excluindo somente vértices pendentes, são árvores. Grafos que podem ser levados para um ponto, unindo apenas pares de gêmeos, são cografos.

Meu leitor provavelmente adivinhou que os grafos colapsáveis não devem estar longe de serem árvores. Isso é verdade, como explico agora:

Seja  $G$  um grafo conexo finito. Suponha que seus vértices tenham sido particionados em duas partes  $A_1$  e  $A_2$ . Seja  $B_1 \subset A_1$  (resp.  $B_2 \subset A_2$ ) o conjunto de vértices de  $A_1$  (resp.  $A_2$ ) adjacentes a algum vértice em  $A_2$  (resp.  $A_1$ ). Suponha que todo elemento de  $B_1$  seja adjacente a todos os elementos de  $B_2$ . Esta condição é trivialmente satisfeita se  $A_1$  ou  $A_2$  contiverem apenas um (ou zero!) elemento, por isso, assumimos que  $A_1$  e  $A_2$  contenham pelo menos dois elementos cada. Nesta situação, o grafo  $G$  é chamado *decomponível* e a partição  $A_1, A_2$  é uma *decomposição*. Para acompanhar essa decomposição, vamos criar dois grafos,  $G_1, G_2$ , da seguinte maneira. O conjunto de vértices de  $G_1$  (resp.  $G_2$ ) é  $A_1$  mais um vértice extra  $x_1$  (resp.  $x_2$ ) chamado de *vértice de controle*. Quanto às arestas de  $G_1$  (resp.  $G_2$ ), escolha aquelas de  $G$  mais as arestas extras conectando  $x_1$  (resp.  $x_2$ ) a todos os elementos de  $B_1$  (resp.  $B_2$ ).

O grafo  $G$  pode ser reconstruído a partir de  $(G_1, x_1)$  e  $(G_2, x_2)$  por uma construção elementar. Observe que os pontos de controle  $x_1, x_2$  não são vértices de  $G$ : eles são úteis apenas para definir as arestas que conectam as duas partes.

Note que quando  $A_2$  contém dois elementos, como na margem, o grafo  $G$  tem um vértice pendente ou um par de gêmeos.

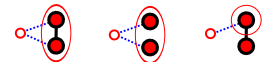
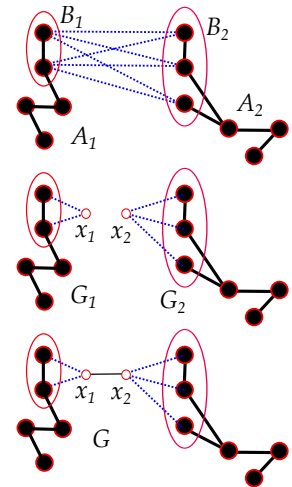
Hammer e Maffray <sup>168</sup> introduziram outra definição em 1987.

**Definição.** Um grafo finito é *completamente decomponível* se todo subgrafo conexo induzido com pelo menos quatro vértices for decomponível.



Aqui há outro desvio porque o único propósito desta seção é descrever a estrutura dos grafos colapsáveis.

Prefiro a terminologia “decomponível” em vez de “separável”, bastante comum na área, uma vez que já usamos a palavra “separável” para permutações.



<sup>168</sup> P. L. Hammer and F. Maffray. Completely separable graphs. *Discrete Appl. Math.*, 27(1-2):85-99, 1990. Computational algorithms, operations research and computer science (Burnaby, BC, 1987).

Não é difícil provar que grafos completamente decompostos (conexos) são precisamente os grafos colapsáveis (conexos).

De fato, na construção da junção, se  $G_1$  e  $G_2$  forem colapsáveis, o mesmo vale para  $G$ , de modo que os grafos completamente decomponíveis são colapsáveis por indução.

Reciprocamente, vimos que um vértice pendente ou um par de gêmeos dá origem a uma decomposição. Portanto, os grafos colapsáveis são decomponíveis e até mesmo completamente decomponíveis, uma vez que um subgrafo induzido de um grafo colapsável é colapsável.

A fim de fornecer uma descrição precisa dos grafos completamente decomponíveis, deixe-me mostrar primeiro o importante teorema da *decomposição* para grafos (conexos e finitos) *gerais*.

Se um grafo conexo finito  $G$  for decomponível, considere-o como a junção de  $G_1$  e  $G_2$  como antes. Em seguida, tente decompor  $G_1$  e  $G_2$  etc. até que os grafos resultantes não sejam decompostos. O resultado final dessa decomposição em “peças primárias” pode ser convenientemente descrito por um *grafo rotulado árvore*, conforme explicado abaixo:

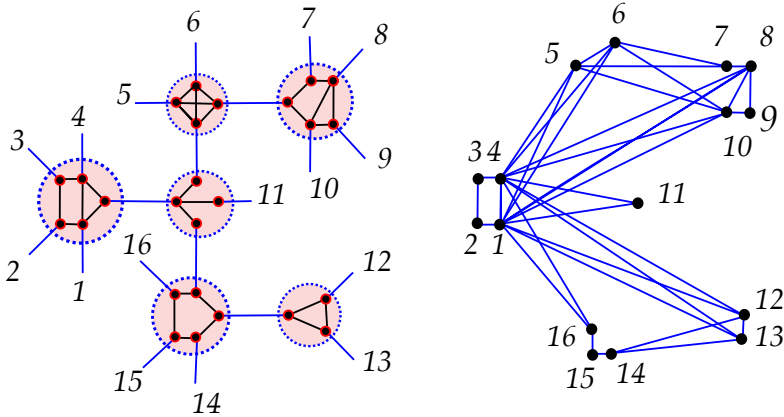
Consiste em uma árvore  $T$  onde cada vértice interno  $x$  é equipado com um grafo finito conexo  $G_x$ . Além disso, foi escolhida alguma bijeção entre os vértices de  $G_x$  e as arestas saindo do vértice  $x$  em  $T$ . Suponha que a valência de cada vértice seja pelo menos 3. Dada tal estrutura, construímos um grafo  $G(T)$  que é uma “composição dos  $G_x$ ’s controlados por  $T$ ”. A definição é a seguinte:

Os vértices de  $G(T)$  são as folhas de  $T$ . A fim de entender as arestas de  $G(T)$ , deixe-me desenhar uma figura, inspirada no artigo de Gioan e Paul <sup>169</sup>, que introduziu esse conceito de *grafos de árvore rotulada*. Nós vemos uma árvore com 16 folhas e 6 vértices internos em rosa. O grafo associado, com 16 vértices, é desenhado à direita.

Escolha duas folhas de  $T$  e conecte-as pelo caminho mais curto da árvore. Para cada vértice  $x$  visitado por esse caminho, há uma aresta de entrada e uma aresta de saída. Por sua vez, essas duas arestas definem dois vértices de  $G_x$ . Dois vértices

Observe a analogia com as árvores: Um grafo conexo finito é uma árvore se, e somente se, cada subgrafo conexo induzido contiver uma *aresta de corte*, ou seja, uma aresta que a desconecte.

<sup>169</sup> E. Gioan and C. Paul. [Split decomposition and graph-labelled trees: characterizations and fully dynamic algorithms for totally decomposable graphs](#). *Discrete Appl. Math.*, 160(6):708–733, 2012.



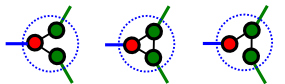
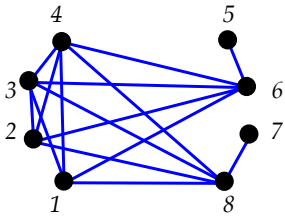
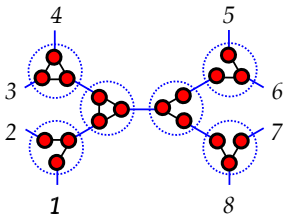
de  $G(T)$ , isto é, duas folhas de  $T$ , são adjacentes em  $G(T)$  se para cada vértice  $x$  visitado por este caminho, os dois vértices correspondentes de  $G_x$  são adjacentes em  $G_x$ . Os vértices dos  $G_x$  generalizam os dois vértices de controle, como no caso simples em que  $T$  contém apenas uma aresta.

O principal resultado, provado por Cunningham e Edmonds <sup>170</sup> em 1980 (e reformulado por Gioan e Paul): *qualquer grafo conexo finito é obtido por tal construção na qual o  $G_x$ 's são indecomponíveis, de uma maneira essencialmente única. A existência desta divisão é fácil. A parte difícil é a "unicidade essencial" que não defino, pois não precisarei dela.*

Agora vamos voltar aos grafos completamente decomponíveis. Neste caso especial, os  $G_x$  devem ter no máximo três vértices. De fato, são subgrafos indecomponíveis e induzidos de  $G$ , de modo que a afirmação decorre da definição de decomposibilidade completa. Isto dá uma descrição geométrica razoavelmente precisa de grafos completamente decomponíveis: Tome uma árvore de modo que cada vértice tenha uma valência três. Para cada vértice, escolha um grafo conexo com três vértices (não há muitas opções!) E construa uma árvore marcada com grafo de árvore rotulada como na margem. Todos os grafos completamente decomponíveis são produzidos desta maneira.

Mas isso não é absolutamente uma surpresa. De fato, olhe para a terceira figura na margem: três pequenos grafos com três

<sup>170</sup> W. H. Cunningham and J. Edmonds. *A combinatorial decomposition theory*. *Canad. J. Math.*, 32(3):734-765, 1980.



folhas. Quando você conecta um deles a uma folha (azul) de uma árvore rotulada com um grafo  $T$ , você obtém outro grafo de árvore rotulada com o nome  $T'$  com mais uma folha. Se você examinar o efeito no grafo associado  $G(T)$ , verá que dividiu um vértice em um par de gêmeos (verdadeiro ou falso) ou criou um vértice pendente, dependendo dos casos. Estamos de volta à definição original de grafos colapsáveis como grafos que podem ser construídos a partir de um ponto por sucessivas introduções de gêmeos ou de vértices pendentes. Nós também estamos de volta aos operads e este foi realmente um ótimo desvio!

### *Computabilidade*

Existe um algoritmo que decide em tempo quadrático (em  $n$ ) se um grafo de tamanho  $n$  é um grafo de entrelaçamento. Isto foi provado em <sup>171</sup> após um longo período de sucessivas melhorias a partir de um algoritmo  $n^9$ , de 1987.

Dado um diagrama com  $n$  cordas, construir seu grafo de entrelaçamento requer um tempo que é quadrático em  $n$ . Em seguida, procure vértices, gêmeos e pendentes, repetindo o processo  $n$  vezes para que você decida em tempo quadrático se é analítico.

<sup>171</sup> J. Spinrad. Recognition of circle graphs. *J. Algorithms*, 16(2):264–282, 1994.

### *Um exercício esotérico*

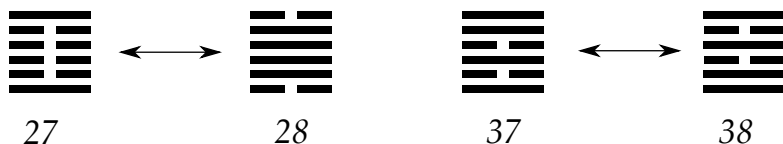
Os 64 hexagramas retratados na primeira página deste capítulo são tradicionalmente agrupados em 32 pares de *hexagramas complementares*. Pense no *Yin e Yang*. Para obter o *dual de um hexagrama*, basta virar de cabeça para baixo. No caso em que o hexagrama é simétrico, substitua cada linha conexa por uma desconexa e vice-versa.

Aqui estão dois exemplos de pares duais:

Os números (27–28) e (37–38) são relativos à ordenação do rei Wen. Muitos especialistas gostam de desenhar um segmento entre hexagramas duplos. Trabalhando com o arranjo circular de Shao Yong, produz-se um diagrama com 32 cordas.

Meu leitor terá paciência para desenhar essas cordas e *decidir*





*quando ou não esse diagrama I Ching é analítico?*

Em 14 de novembro de 1701, Leibniz recebeu uma cópia do arranjo circular de Shao Yong do jesuíta francês Joachim Bouvet, que vivia na China. Dois anos depois, publicou um notável artigo sobre aritmética binária <sup>172</sup> em *Mémoires de l'Académie des Sciences*. De acordo com ele:

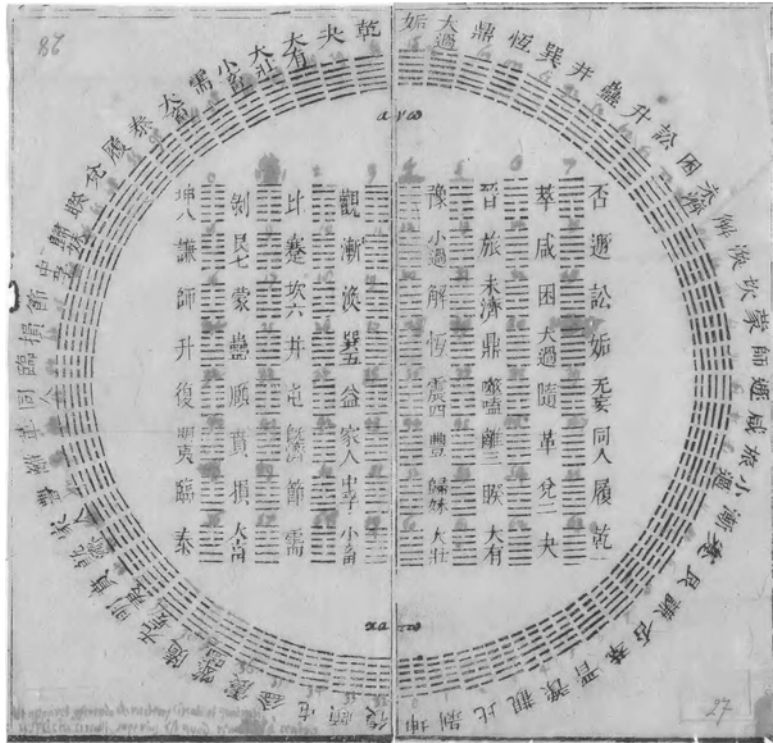
Essas figuras talvez sejam o monumento mais antigo da ciência existente no mundo.

“Leibniz esperava que sua análise astuta dos trigramas do *I Ching* despertasse na China um profundo apreço pela ciência ocidental e, por fim, pelo cristianismo <sup>173</sup>.”

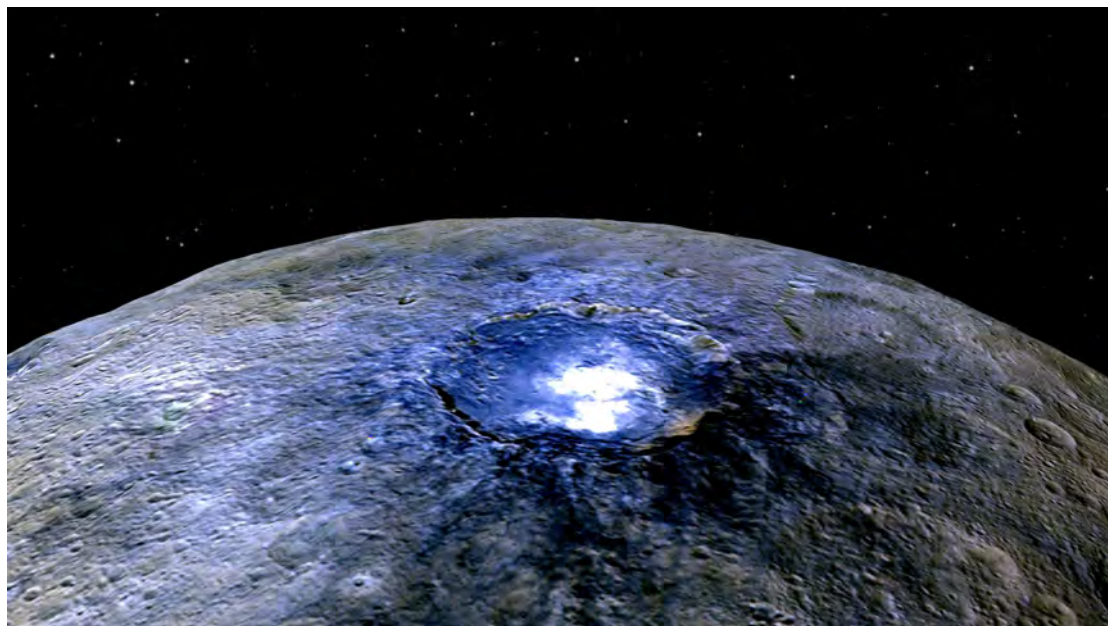
<sup>172</sup> G.-G. Leibniz. *Explication de l'arithmétique binaire, qui se sert des seuls caractères 0 et 1 avec des remarques sur son utilité et sur ce qu'elle donne le sens des anciennes figures chinoises de Fohy. Mémoires de mathématique et de physique de l'Académie royale des sciences*, pages 85–89, 1703.

Ces figures sont peut-être le plus ancien monument de science qui soit au monde.

<sup>173</sup> D. Lach. *Leibniz and China. Journal of the History of Ideas*, 6(4):436–455, 1945.



O arranjo circular que Leibniz recebeu.

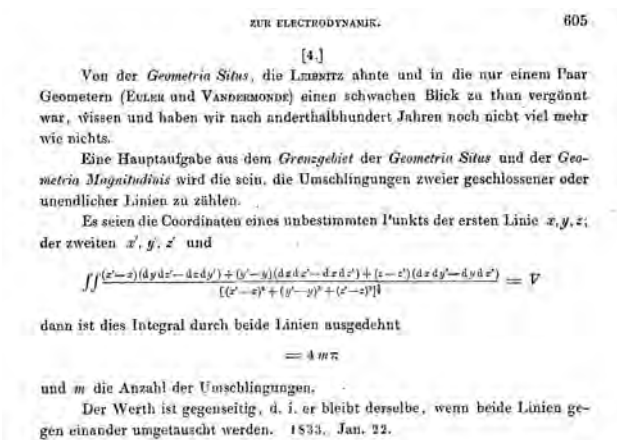


O planeta anão Ceres, como visto pela missão Dawn, em julho de 2016. A determinação de sua órbita foi um sucesso espetacular de Gauss. †

# Gauss novamente: ligando o magnetismo e a astronomia

Gauss e os números de entrelace

EM 22 DE JANEIRO DE 1833, GAUSS ESCREVEU UMA FÓRMULA ENIGMÁTICA NO SEU CADERNO <sup>174</sup>.



Uma moeda alemã, produzida em 1977, celebrando o 200º aniversário de Gauss. +

<sup>174</sup> C. F. Gauß. *Werke. Ergänzungsreihe. Band V.* Georg Olms Verlag, Hildesheim, 1975. Briefwechsel: C. F. Gauss–H. C. Schumacher. Vol. 3. Editado por C. A. F. Peters, Reimpressão dos originais de 1863 e 1865.

Sua finalidade é “contar o número de entrelace de duas curvas fechadas: um inteiro associado a duas curvas fechadas disjuntas no espaço de dimensão três que é invariante sob deformação.

Em 1833, a Topologia ainda não existia. . . Até mesmo a palavra só apareceria impressa doze anos depois em um [livro](#) por Listing. Leibniz já havia cunhado a palavra *Analysis Situs* e es-

“Die Umschlingungen zweier geschlossener Lieben zu zählen.”

tava apenas sonhando com alguma ciência manipulando formas, assim como a álgebra manipula símbolos. Gauss usa a terminologia *Geometria Situs* e menciona Euler e Vandermonde como precursores.

Não esqueça que estes *Nachlasse* não foram destinados para publicação. O que ele teria pensado se soubesse que seus rascunhos privados se tornariam disponíveis publicamente? Essa nota de 1833 foi de fato publicada em 1867, após a morte de Gauss, e o editor a incluiu em um volume dedicado ao eletromagnetismo. Essa foi uma escolha razoável, e um recente artigo <sup>175</sup> propõe uma boa interpretação eletromagnética. Outro trabalho <sup>176</sup> afirma, pelo contrário, que a fórmula tem uma origem astronômica e o artigo parece bastante crível. Quem está certo? Ambos, claro! Gauss estava convencido pela profunda unidade da matemática, não construindo fronteiras entre matemática, astronomia, física etc. Eu certamente não farei uma escolha e apresentarei três pontos de vista paralelos: três definições do número de entrelace de duas curvas fechadas disjuntas no espaço de dimensão três.

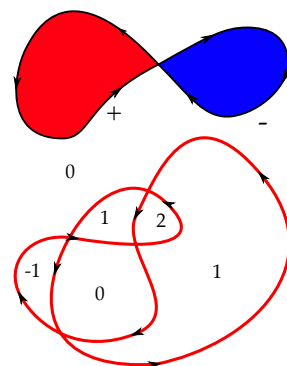
### *Geometria*

Uma curva fechada suave, orientada e *mergulhada* no plano limita um domínio, que possui alguma área. Digamos que essa área seja positiva se a curva for orientada no sentido anti-horário e negativa no outro caso. Isso é fácil. Agora, se a curva não está mergulhada, a situação é um pouco mais complicada, como, por exemplo, na figura oito, limitando o domínio. O laço esquerdo é orientado no sentido anti-horário e o direito no sentido horário, de modo que somos levados a definir a *área com sinal* como a soma algébrica das duas áreas.

No caso geral de uma curva imersa com muitos pontos duplos finitos, procedemos de maneira semelhante. A curva decompõe o plano em componentes conexas. Vamos equipar cada componente ilimitado com o coeficiente 0. Agora, equipamos cada componente com algum inteiro a partir da convenção de que quando cruzarmos a curva positivamente, esse inteiro salta em +1. Em outras palavras, um ponto se movendo na curva na

<sup>175</sup> R. L. Ricca and B. Nipoti. Gauss' linking number revisited. *J. Knot Theory Ramifications*, 20(10):1325–1343, 2011.

<sup>176</sup> M. Epple. Orbits of asteroids, a braid, and the first link invariant. *Math. Intelligencer*, 20(1):45–52, 1998.



Tente e prove a existência de tal rotulagem usando o fato de que o número de interseções (algébricas) de duas curvas orientadas transversais fechadas no plano é 0. Sejam  $c_1$  e  $c_2$  duas curvas orientadas transversalmente em uma superfície orientada. Qualquer ponto de interseção de  $c_1$  e  $c_2$  tem um sinal  $\pm 1$  dependendo da orientação dada pelo par de vetores tangentes de  $c_1$  e  $c_2$  neste ponto. A soma de todos esses sinais para todos os pontos de interseção é o *número de interseção algébrica* de  $c_1$  e  $c_2$ , que caso sejam curvas no plano, o número de interseção é 0. Em termos modernos, isso decorre do fato de que a homologia do plano é trivial. Gauss sabia disso. Você pode produzir uma prova de que ele poderia ter aceitado?

direção positiva vê um coeficiente à esquerda igual ao coeficiente à direita +1. Acontece que essa rotulagem existe e é única.

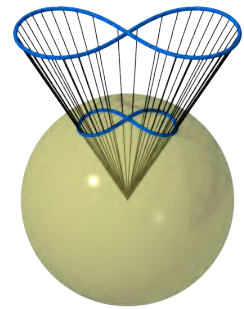
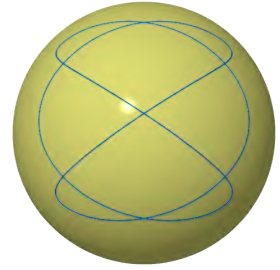
Em seguida, definimos a área sinalada da curva como a combinação linear das áreas geométricas dos componentes com esses coeficientes inteiros. Esta definição é natural e é devida à ... Gauss.

Outra definição vem do fato de que, afinal, a área *abaixo de uma curva*  $y(x)$  é a integral de  $y dx$ . Considere a 1-forma diferencial  $\omega = -y dx$  no plano e integre-o ao longo da curva. Encorajo ao meu leitor, parceiro desse passeio, a verificar se essas duas definições fornecem o mesmo número. En passant, note que a diferencial de  $\omega$  é a 2-forma  $dx \wedge dy$ , que é a forma da área. Isto não é uma surpresa para um matemático do século XXI, mas estava longe de ser óbvio no início do século XIX.

Considere agora uma curva orientada fechada na esfera unitária de dimensão dois. Podemos definir a área fechada? Se a curva estiver mergulhada, não há problema. A curva decompõe a esfera em dois domínios, um deles tendo a curva orientada (borda) no sentido anti-horário, e podemos definir a área como a área deste domínio. Agora, se a curva é complicada, o que podemos fazer? Ainda podemos atribuir números a cada componente conexa do complemento com a mesma propriedade de antes, mas eles não podem ser normalizados pedindo que alguma componente no infinito tenha o peso 0, desde que não haja o infinito. Portanto, todos esses inteiros são bem definidos a menos da adição do mesmo inteiro a cada componente. A área com sinal delimitada pela curva é definida apenas a menos de uma adição de um múltiplo integral da área da esfera, ou seja, módulo  $4\pi\mathbb{Z}$ .

Uma curva suave fechada na esfera define um cone em espaço de dimensão três, com o vértice na origem. A área da curva é, por definição, o *ângulo sólido* do cone, portanto, definida como módulo  $4\pi\mathbb{Z}$ . Assim como um ângulo orientado no plano, é definido como módulo  $2\pi\mathbb{Z}$ .

Suponha agora que temos uma curva orientada fechada  $\gamma$  no espaço, não necessariamente mergulhada. Para cada ponto  $x$  fora de  $\gamma$ , veja o ângulo sólido  $A_\gamma(x)$  do cone com o vértice  $x$  e com



Uma curva fechada de dimensão três é a imagem de uma aplicação de  $S^1$  para  $\mathbb{R}^3$ . Em dimensão dois, podemos considerar uma aplicação de  $S^0$  para  $\mathbb{R}^2$ , ou seja, dois pontos  $P, Q$  no plano! O cone agora é substituído por um triângulo e a função ângulo  $A : \mathbb{R}^2 \setminus \{P, Q\} \rightarrow \mathbb{R}/2\pi\mathbb{Z}$  associado para cada  $x$  o ângulo  $\widehat{PxQ}$ . Ao olhar para as fibras de  $A$ , pode-se lembrar da escola secundária...

base em  $\gamma$ . Seu ângulo sólido define uma função

$$A_\gamma : \mathbb{R}^3 \setminus \gamma \rightarrow \mathbb{R}/4\pi\mathbb{Z}.$$

Note que se  $\gamma$  é um nó, ou seja, se é um círculo mergulhado, a pré-imagem  $A_\gamma^{-1}(\theta)$  é uma superfície orientável com  $\gamma$  como sua borda para cada valor regular  $\theta$  de  $A_\gamma$ . Já encontramos essas superfícies com o nome de Seifert.

Vamos calcular a diferencial  $dA_\gamma$ . Sejam  $x, x'$  dois pontos próximos no espaço, ambos longe de  $\gamma$ . Para calcular  $A_\gamma(x)$ , devemos transladar  $\gamma$  por  $-x$ , projetar radialmente o resultado na esfera unitária e calcular sua área com sinal. A diferença de áreas  $A_\gamma(x) - A_\gamma(x')$  é a área com sinal da projeção na esfera unitária do anel delimitando as bordas de  $\gamma$  por  $-x$  e  $-x'$ . Aproximando  $\gamma$  por alguma curva poligonal de modo que  $A_\gamma(x) - A_\gamma(x')$  é aproximada pela soma de áreas sinaladas das projeções de alguns paralelogramos. Note que se  $\delta x, \delta' x$  são dois vetores no espaço, o volume da pirâmide com ápice 0 e base

$$x, x + \delta x, x + \delta' x, x + \delta x + \delta' x$$

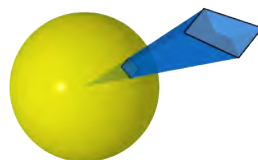
é

$$\frac{1}{3} \det(x, \delta x, \delta' x).$$

Se  $\delta x$  e  $\delta' x$  são muito pequenos, o ângulo sólido correspondente é obtido aproximadamente dividindo este valor pela norma de  $x$  ao cubo. Colocando tudo junto, passando o limite, obtemos uma fórmula para  $dA$  no ponto  $x$ , no vetor  $v$ :

$$dA_\gamma(x, v) = \int_\gamma \frac{1}{\|\gamma(t) - x\|^3} \det\left(\gamma(t), \frac{d\gamma}{dt}(t), v\right) dt.$$

Conhecemos o conceito de *índice de Cauchy* de uma curva fechada  $c$  no plano perfurado  $\mathbb{C} \setminus \{0\}$ : o número de “voltas” ao redor da origem quando se caminha ao longo de  $c$ . Siga o argumento por continuidade enquanto você gira em torno de  $c$  e depois compute o incremento do argumento quando você está de volta ao ponto de partida. Pode-se também usar a forma diferencial  $\frac{1}{2i\pi} dz/z$  e integrá-la em  $c$ . Tudo isso parece fácil para os estudantes de hoje, mas não era óbvio para os pais fundadores Gauss–Cauchy etc.



ângulo sólido.

†

Agora, faça exatamente o mesmo procedimento no espaço de dimensão três, substituindo o argumento pelo ângulo sólido criado por alguma curva fechada  $\gamma$ . Se uma curva  $\gamma'$  não intersectar  $\gamma$ , contorne  $\gamma'$  e observe o aumento do ângulo sólido ao fazer uma volta completa (dividido por  $4\pi$ ). Este índice é chamado de *número de entrelace* de  $\gamma$  e  $\gamma'$ : esse é um inteiro.

Usando a fórmula para  $dA_\gamma$ , obtemos a fórmula de Gauss para o número de entrelace  $link(\gamma, \gamma')$ :

$$\frac{1}{4\pi} \iint \frac{1}{\|\gamma(t) - \gamma'(t')\|^3} \det \left( \gamma(t) - \gamma'(t'), \frac{d\gamma}{dt}(t), \frac{d\gamma'}{dt'}(t') \right) dt dt'.$$

Isso é exatamente o que Gauss escreveu em seu caderno em 22 de janeiro de 1833.

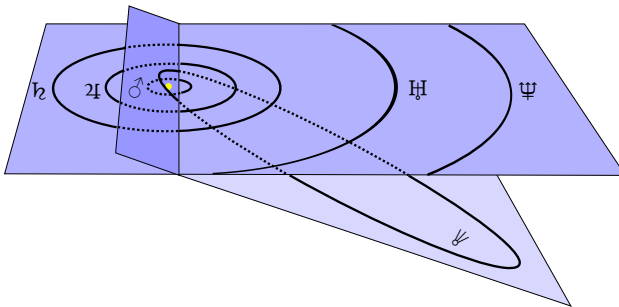
Note que a fórmula acima mostra que o número de entrelace é simétrico  $link(\gamma, \gamma') = link(\gamma', \gamma)$ , que não era óbvia a partir da definição. Isso é o que Gauss escreveu:

O valor é simétrico i.e. permanece o mesmo quando se troca uma curva pela outra.

Note também que se  $\gamma$  e  $\gamma'$  são deformadas continuamente de tal forma que elas não se cruzam durante a deformação, o número de entrelace tem que ser constante: um inteiro não pode mudar continuamente. Essa é a característica mais importante do número de entrelace: *invariante sob deformação*.

*Der Werth ist gegenseitig, d. i. er bleibt derselbe, wenn beide Linien gegen einander umgetauscht werden.*

### Astronomia



A órbita do cometa de Halley, juntamente com as órbitas de Marte, Júpiter, Saturno, Urano e Netuno.



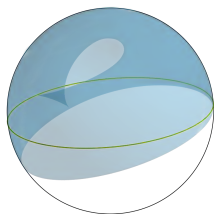
O artigo de Epple mencionado acima propõe uma possível abordagem para números de entrelace. Uma das primeiras realizações de Gauss que o tornou famoso foi sua determinação em 1801 da órbita do planeta anão Ceres, que acabara de ser descoberto. Suponha que observemos um planeta a partir de uma posição fixa no nosso planeta Terra. Onde devemos olhar no céu? Mais precisamente, seja  $\gamma$  a trajetória da Terra em espaço fixo (fixado em relação ao Sol) e deixe que  $\gamma'$  seja a trajetória do planeta que queremos observar. Por simplicidade, eu assumo que  $\gamma$  e  $\gamma'$  são disjuntos ☺. Se os períodos de rotação são racionalmente independentes, as posições da Terra e do planeta *em suas órbitas* são variáveis aleatórias independentes. Gauss chama *zodíaco* do planeta (em relação à Terra) a imagem da aplicação:

$$\omega : (t, t') \in (\mathbb{R}/\mathbb{Z})^2 \mapsto \frac{\gamma(t) - \gamma'(t')}{\|\gamma(t) - \gamma'(t')\|} \in \mathbb{S}^2.$$

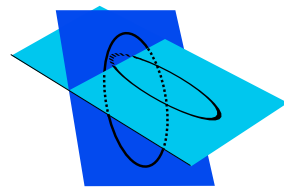
Essa é a zona na esfera celeste onde o observador deve procurar pelo planeta.

O integrando na fórmula de Gauss para o número de entrelace é simplesmente o determinante Jacobiano desta aplicação, de modo que o número de entrelace é  $1/4\pi$  vezes a área com sinal do zodíaco. Um matemático moderno sabe que a integral do determinante jacobiano de um mapa entre duas variedades orientadas da mesma dimensão é o grau topológico desta aplicação. Portanto, o *número de entrelace também pode ser definido como o grau da aplicação zodíaco*  $\omega$ .

Naturalmente, Gauss estudou em detalhes o caso de duas elipses no espaço.



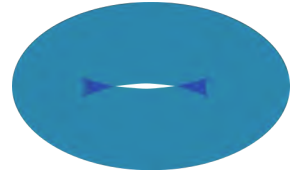
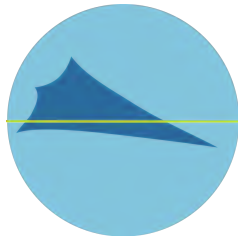
Quando as duas elipses não estão entrelaçadas, como na imagem acima, o zodíaco não cobre toda a esfera celeste.



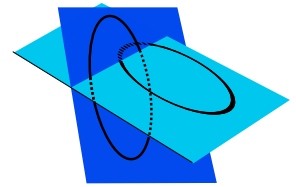
Duas elipses não entrelaçadas. O zodíaco está à esquerda.

A zona azul clara no zodíaco corresponde a pontos que são cobertos duas vezes por  $\omega$ . A zona mais escura, com dois pontos singulares, é coberta quatro vezes. Compare com a imagem usual (na margem) da perspectiva de um toro de revolução e as singularidades que aparecem em seu contorno.

Quando as elipses estão entrelaçadas, como na imagem abaixo, o zodíaco é a esfera completa. A zona azul clara no zodíaco corresponde a pontos que são cobertos apenas uma vez por  $\omega$ . A zona mais escura, com quatro pontos singulares, é coberta três vezes.



Projeção de um toro em um plano.



Duas elipses entrelaçadas. O zodíaco está à esquerda.

Dada uma aplicação suave  $f : M \rightarrow N$  entre duas variedades compactas e orientadas, sem borda, existem várias definições possíveis para o seu grau topológico. A primeira consiste em escolher alguma forma volume  $vol$  em  $N$ , do volume total 1, e integrar o pull-back  $f^*vol$  em  $M$ . Não é difícil ver que isso é independente de  $vol$ . De fato, se  $vol'$  é outra opção de forma de volume,  $vol' - vol$  é uma forma exata e, portanto, a integral de  $f^*vol' - f^*vol$  é zero. A partir dessa definição, é fácil ver que isso é invariante sob deformação. De fato, se duas aplicações  $f_0, f_1$  são homotópicas,  $f_0^*vol - f_1^*vol$  é uma forma exata. É bem menos fácil ver que esse grau é um inteiro.

Uma segunda definição consiste em escolher um valor regular (de  $f$ )  $y \in N$  e observar as pré-imagens  $x_1, \dots, x_n$  de  $y$  em  $M$

Em cada uma dessas pré-imagens, o diferencial de  $f$  preserva ou inverte a orientação, e atribuímos a elas um sinal + ou - de acordo. O grau de  $f$  é a soma desses sinais. É preciso mostrar que isso não depende da escolha do valor regular e de que é um invariante por homotopia. Isto está provado brilhantemente no livro de Milnor <sup>177</sup>. Deve-se também provar que as duas

Veja novamente as figuras anteriores do zodíaco de duas elipses e descubra os sinais + ou -.

<sup>177</sup> J. W. Milnor. *Topology from the differentiable viewpoint*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997.

definições são equivalentes . . . Uma possibilidade é usar uma sequência de formas de volume em  $N$  que converge para a massa de Dirac no valor regular  $y$ .

Vamos usar o ponto de vista do valor regular para calcular o grau da aplicação zodiaco  $\varphi : \mathbb{R}^2/\mathbb{Z}^2 \rightarrow S^2$ . Escolha o polo sul de  $S^2$  como o ponto  $y$ . As pré-imagens de  $y$  consistem dos pares de pontos  $\gamma(t)$  e  $\gamma'(t')$  tal que  $\gamma(t)$  está acima de  $\gamma'(t')$ : eles têm as mesmas coordenadas  $x, y$  e a coordenada  $z$  de  $\gamma(t)$  é maior que  $\gamma'(t')$ . A diferencial de  $\varphi$  em tal ponto é fácil de calcular. Não é degenerada se as projeções de  $\gamma$  e  $\gamma'$  no plano  $(x, y)$  cruzarem-se transversalmente no ponto correspondente. Seu determinante Jacobiano é *positivo* (resp. *negativo*) se a intersecção das projeções for positiva (resp. negativa).

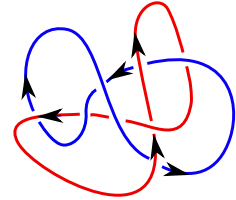
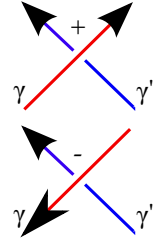
Isso leva à definição combinatória do número de entrelace, que Gauss obviamente sabia. Projete as duas curvas  $\gamma, \gamma'$  em um plano genérico para que as duas projeções  $\tilde{\gamma}, \tilde{\gamma}'$  intersectam-se transversalmente. Marque um índice  $+1$  ou  $-1$  para cada ponto de intersecção, conforme os vetores tangentes em  $\gamma$  e  $\gamma'$  definam uma base positiva ou negativa. Entre esses pontos de intersecção, selecione apenas aqueles em que  $\gamma$  está *sobre*  $\gamma'$ . A soma dos sinais correspondentes é o número de entrelace de  $\gamma, \gamma'$ . Na imagem, a curva vermelha passa três vezes sobre a azul com os sinais  $+1, +1, -1$ . O número de entrelace é de 1.

O chamado *link de Whitehead* na margem tem o número de entrelace 0, mas isso não significa que as duas componentes possam ser separadas por alguma deformação <sup>178</sup>. Mostre que não há curva algébrica complexa com dois ramos, de forma que o link associado seja esse link de Whitehead.

*Eletromagnetismo*

A fórmula de Gauss é uma reminiscência da lei Biot–Savart da Física. Uma corrente elétrica gera um campo magnético. Suponha que um fio fechado  $\gamma$  é atravessado por uma corrente constante, com intensidade de  $i$  e seja  $x$  um ponto fora do fio. Então, o campo magnético criado em  $x$  é

$$B(x) = \frac{\mu_0 i}{4\pi} \int_{\gamma} \frac{1}{\|(\gamma(t) - x)\|^3} \left( (\gamma(t) - x) \wedge \frac{d\gamma(t)}{dt} \right) dt$$

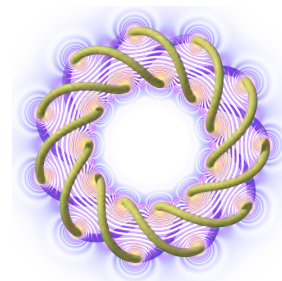


Link de Whitehead.

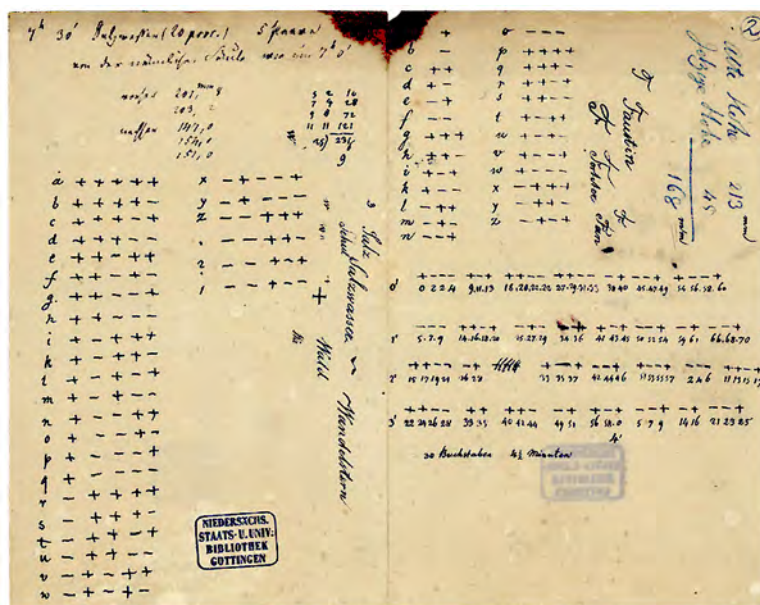
<sup>178</sup> D. Rolfsen. *Knots and links*, volume 7 of *Mathematics Lecture Series*. Publish or Perish, Inc., Houston, TX, 1990. Reimpressão corrigida do original de 1976.

sendo  $\mu_0$  a constante magnética. Esse campo vetorial é o dual da 1-forma  $dA_\gamma$  (com relação à métrica euclideana no espaço físico de dimensão 3. Ele também pode ser interpretado como o campo de gradiente de uma primitiva local da 1-forma  $A_\gamma$ . Segue-se que a circulação do campo magnético em algum laço  $\gamma'$  é a mesma que a integral de  $dA_\gamma$  em  $\gamma'$ , ou seja, o número de entrelace. Portanto, o número de entrelace  $link(\gamma, \gamma')$  é a circulação do campo magnético criado por uma corrente.

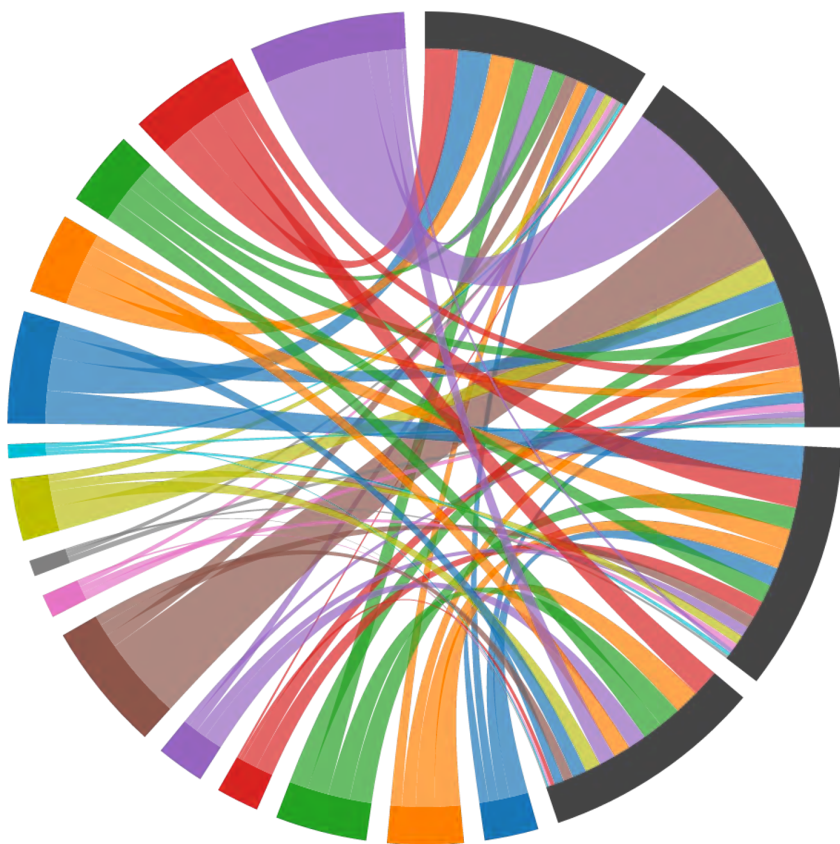
O artigo acima mencionado de Ricca e Nipoti dá uma interessante reconstrução do que poderia ter sido a interpretação magnética na mente de Gauss. Não se esqueça que, junto com Weber, Gauss estabeleceu o primeiro telégrafo transmitindo mensagens através de Goettingen.



O campo magnético gerado por um toro.



O código de Gauss-Weber para o seu telégrafo: combinatoria de novo!



Um **diagrama da corda** é também um método gráfico de exibir as inter-relações entre os dados. Dada uma matriz estocástica  $a_{ij}$  (i.e.  $a_{ij} > 0$  e  $\sum_j a_{ij} = 1$ ), pode-se pensar em  $a_{ij}$  como a proporção da entidade  $i$  interagindo com  $j$ . Desenhamos  $n$  intervalos  $I_1, \dots, I_n$  ao redor do círculo, cujos comprimentos  $l_1, \dots, l_n$  precisam ser determinados, e bandas conectando  $I_i$  e  $I_j$  com larguras  $a_{ij}l_i$ . A condição de compatibilidade pode ser expressa como  $l_j = \sum_i a_{ij}l_i$ . A existência de uma solução é garantida pelo teorema de Perron–Frobenius.

# Kontsevich está de volta: Um invariante universal

## ESTE PASSEIO É UM LAÇO HOMOTÓPICO A UM PONTO?

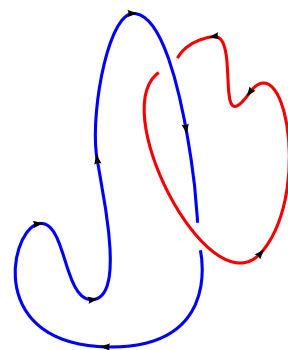
Estamos de volta ao nosso ponto de partida: Maxim Kontsevich. Este capítulo não é uma conclusão, mas uma abertura para um vasto domínio que mostra: na matemática, é possível voltar a ideias muito antigas com uma perspectiva completamente nova. Para isso, quero apresentar uma pequena introdução de um desenvolvimento maravilhoso, na teoria dos nós, envolvendo diagramas de cordas, de um artigo de 1993 de Kontsevich <sup>179</sup>.

<sup>179</sup> M. Kontsevich. Vassiliev's knot invariants. In *I. M. Gelfand Seminar*, volume 16 of *Adv. Soviet Math.*, pages 137–150. Amer. Math. Soc., Providence, RI, 1993.

## Um novo ponto de vista sobre o número de entrelace

Sejam  $\gamma_1, \gamma_2 : \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{R}^3$  duas curvas fechadas orientadas e disjuntas. Sabemos que o número de entrelace de  $\gamma_1$  e  $\gamma_2$  é o grau topológico da *aplicação zodíaco* de Gauss, do produto das duas curvas para a esfera unitária. A fórmula de Kontsevich irá expressar o mesmo número que um grau topológico de uma aplicação *de uma variedade dimensão 1 orientada, portanto, uma união de círculos para um círculo*. A grande vantagem é que esse novo ponto de vista nos permite definir muitos mais invariantes.

Vamos pensar no espaço  $\mathbb{R}^3$  com as coordenadas  $(x, y, t)$  como o produto da linha complexa  $\mathbb{C}$  (com coordenada  $z = x + iy$ ) e  $\mathbb{R}$  (com coordenada  $t$ ). Suponha que nossas curvas  $\gamma_1, \gamma_2$  sejam de tipo *Morse*. Isso significa simplesmente que a projeção para a coordenada  $t$  tem um número finito de pontos críticos e que a



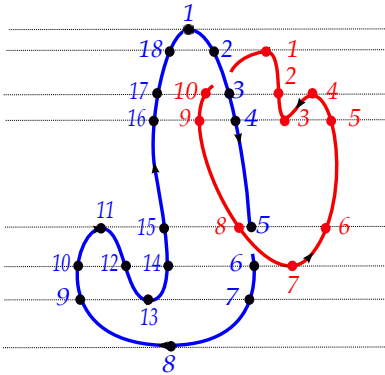
segunda derivada não é zero nesses pontos críticos. Suponha que os valores críticos das  $t$ -coordenadas sejam todos distintos.

Considere agora o conjunto de pares de pontos em  $\gamma_1, \gamma_2$  que têm a mesma coordenada  $t$ . Formalmente, este é o conjunto

$$X = \{(s_1, s_2) \in (\mathbb{R}/\mathbb{Z})^2 \mid t(\gamma_1(s_1)) = t(\gamma_2(s_2))\}.$$

Essa é uma curva suave do toro de dimensão 2. O único fato (fácil) que se deve verificar é que este é realmente o caso na vizinhança de pontos críticos.

Veja este exemplo: Existem oito valores críticos, decompondo a primeira curva em 18 fios e a segunda em 10 fios.

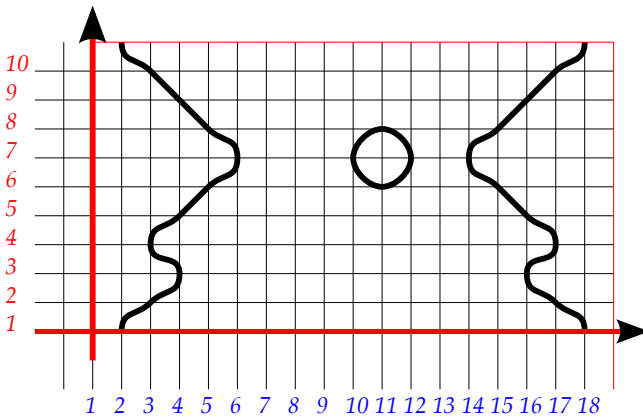


A subvariedade  $X$  está representada na figura a seguir:

Existe uma orientação canônica em  $X$ . Escolha um pequeno intervalo  $I$  em  $X$ , longe dos valores críticos. Esse intervalo é mapeado de forma difeomorfa em algum intervalo  $I_1$  em  $\gamma_1$  e em algum outro intervalo  $I_2$  em  $\gamma_2$ . Um intervalo não crítico em  $\gamma_1$  (ou em  $\gamma_2$ ) está equipado com duas orientações, vindas da orientação  $\gamma_1$  (ou  $\gamma_2$ ) por um lado, e das  $t$ -coordenadas, por outro lado. Então, direi que tal intervalo é *positivo* se essas duas orientações concordarem e *negativo* caso contrário. Oriente  $I$  usando o aumento de  $t$  se  $I_1$  e  $I_2$  forem ambos positivos ou negativos, e usando menos  $t$  da outra forma.

Verifique se isso define uma orientação em  $X$ .

Agora, um ponto em  $X$  define dois pontos  $\gamma_1(s_1)$  e  $\gamma_2(s_2)$ , que projetam para pontos *distintos* no plano complexo  $x + iy$ . O



argumento da diferença define um mapa  $\pi : X \rightarrow S^1$ . Esta é uma aplicação entre variedades orientadas de dimensão 1.

*Eu afirmo que o grau de  $\pi$  é o número de entrelace de  $\gamma_1$  e  $\gamma_2$ .*

Vamos provar essa afirmação: O número de entrelace é o grau topológico da aplicação

$$\omega : (s_1, s_2) \in (\mathbb{R}/\mathbb{Z})^2 \mapsto \frac{\gamma_1(s_1) - \gamma_2(s_2)}{\|\gamma_1(s_1) - \gamma_2(s_2)\|} \in S^2$$

entre superfícies orientadas. A esfera unitária  $S^2$  contém o equador horizontal  $S^1$  (sendo  $t = 0$ ). A suposição de que  $\gamma_1$  e  $\gamma_2$  são do tipo Morse com valores críticos distintos implica que  $\omega$  é transversal a  $S^1 \subset S^2$ . A imagem inversa  $\pi^{-1}(S^1)$  é  $X$ , por definição. A diferencial de  $\omega$  identifica o fibrado normal de  $X$ , no 2-toro, com o fibrado normal do equador, na esfera. Nossa convenção de orientação em  $X$  é tal que essa identificação é positiva.

Queremos comparar os dois graus topológicos de  $\omega$  e  $\pi$ . Tome um valor regular  $v \in S^1 \subset S^2$  de  $\pi$  e deixe  $u$  ser um ponto na sua pré-imagem. Observe que  $v$  também é um valor normal de  $\omega$ . O sinal do jacobiano da diferencial de  $\pi$  em  $u$  é o mesmo que o sinal do jacobiano de  $\omega$  em  $u$ . Segue que os graus de  $\omega$  e  $\pi$  são iguais. □

Agora, obtemos uma nova fórmula para o número de entrelace, usando índices do tipo Cauchy. Este é um caso especial do teorema de Kontsevich.



**Teorema.** *Fatie  $\gamma_1$  e  $\gamma_2$  por planos horizontais que passam pelos pontos críticos da  $t$ -coordenada de  $\gamma_1$  ou  $\gamma_2$ . Entre dois planos consecutivos,  $\gamma_1, \gamma_2$  definem um certo número de filamentos que são positivos ou negativos. Escolha um dos fios correspondentes a  $\gamma_1$ , definida por algum grafo  $(\zeta_1(t), t)$  (para  $t_- \leq t \leq t_+$ ). Escolha um dos fios  $(\zeta_2(t), t)$  para a curva  $\gamma_2$  (também para  $t_- \leq t \leq t_+$ ). Calcule a quantidade de rotação*

$$\varepsilon \frac{1}{2i\pi} \int_{t_-}^{t_+} \frac{d(\zeta_1(t) - \zeta_2(t))}{\zeta_1(t) - \zeta_2(t)}$$

sendo  $\varepsilon +1$  se os dois fios escolhidos tiverem o mesmo sinal e  $-1$  caso contrário.

Some todos esses números para todos os possíveis pares de planos horizontais consecutivos, e para todos os pares de um fio para  $\gamma_1$  e um fio para  $\gamma_2$ . O resultado é o número de entrelace  $lk(\gamma_1, \gamma_2)$ .

No exemplo anterior, os 8 valores singulares definem 7 intervalos contendo 2, 2, 2, 2, 4, 4, 2 fios azuis e 0, 2, 4, 2, 2, 0 fios vermelhos. Portanto, existem

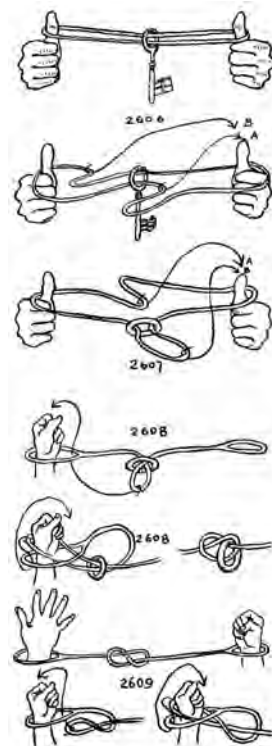
$$2 \times 0 + 2 \times 2 + 2 \times 4 + 2 \times 2 + 4 \times 2 + 4 \times 0 + 4 \times 0 = 24$$

emparelhamentos entre fios, correspondendo ao número de intervalos em  $X$ .

### Invariante Kontsevich universal de um nó com valores na álgebra de cordas

Para concluir, eu esboço a definição de um invariante associado a um nó com valores em séries formais com coeficientes em diagramas de cordas. Esta é uma ideia brilhante de Kontsevich, do seu famoso artigo de 1993.

Seja  $Chord(n)$  o conjunto de diagramas de cordas com  $n$  cordas. Como vimos muitas vezes, são conjuntos de  $2n$  pontos em um círculo orientado, agrupados em pares, até a orientação de preservar os homeomorfismos do círculo. Denotar por  $\mathbb{C}[Chord]$  o espaço vetorial com a união  $Chord$  de todos os  $Chord(n)$  como base. Seus elementos são, portanto, somas finitas  $\sum_{w \in Chord} \lambda_w \cdot w$  em que  $\lambda_w = 0$  para todos, exceto um número finito de  $w$ . Con-

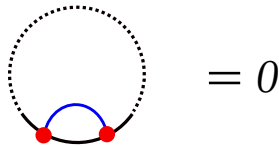


Alguns truques de mágica tentando convencê-lo de que o número de entrelace não existe. Do maravilhoso "Ashley's book of knots" (1944), contendo 7000 figuras representando 3800 nós. †

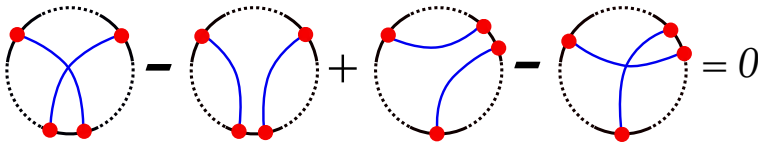
sidere  $\mathbb{C}[Chord]$  como um espaço vetorial *graduado*, sendo a graduação dada por  $n$ .

Denote por  $\mathcal{A}$  o quociente de  $\mathbb{C}[Chord]$  pelo subespaço gerado por duas relações, que podem parecer artificiais à primeira vista:

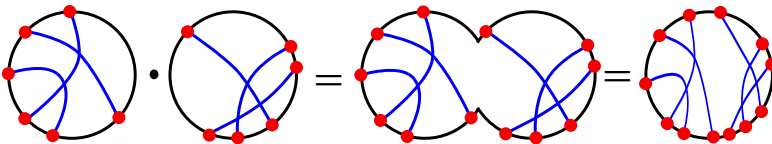
— uma *relação de um termo*. Isso significa que qualquer diagrama de cordas obtido a partir da imagem abaixo, completado em *qualquer caminho* na parte pontilhada do círculo, é declarado como 0 em  $\mathcal{A}$ . Dito de forma diferente, cada diagrama contendo uma corda solitária é igual a 0 em  $\mathcal{A}$ .



— uma *relação de quatro termos*. Analogamente, a parte pontilhada do círculo pode ser completada de qualquer maneira, mas é claro que da mesma maneira nos quatro constituintes da relação.



Esse espaço vetorial  $\mathcal{A}$  é realmente uma *álgebra graduada*  $\oplus_{n \geq 0} \mathcal{A}_n$ . Dois diagramas de cordas podem ser multiplicados da seguinte forma.



A relação de quatro termos é exatamente o que é necessário para garantir que essa operação seja bem definida, independentemente do lugar da *soma conexa*.

Considere o complemento  $\hat{\mathcal{A}}$ , no qual adicionamos somas formais infinitas  $\sum_{w \in Chord_n} \lambda_w \cdot w$  sem condição nos números  $\lambda_w$ .

Vamos chamar  $\hat{\mathcal{A}}$  the *álgebra de cordas*.

Agora posso definir a invariante universal *Kontsevich de um nó*, com valores em  $\hat{\mathcal{A}}$ .

Seja  $\gamma$  um nó no espaço, assumido do tipo Morse.

Corte-o por planos horizontais passando pelos pontos críticos das coordenadas  $t$ . Isso decompõe o nó em um número finito de fios, que podem ser positivos ou negativos, em relação à orientação do nó.

Escolha algum inteiro  $n$ . Considere o espaço de  $2n$ -tuplos de pontos distintos  $(p_1, q_1, \dots, p_n, q_n)$  no nó de modo que

$$t(p_1) = t(q_1) < t(p_2) = t(q_2) < \dots < t(p_n) = t(q_n)$$

Esta é uma subvariedade  $n$ -dimensional  $X_n$  com borda sendo o toro de dimensão 2, canonicamente orientado pela orientação do círculo.

Note que qualquer elemento de  $X_n$  define um diagrama de corda com  $n$  cordas.

Existe uma aplicação natural  $\omega$  de  $X_n$  para  $(\mathbb{C}^*)^n$ . De fato, se  $p$  e  $q$  são dois pontos distintos em  $\gamma$ , com a mesma coordenada  $t$ , sua diferença é um número complexo diferente de zero. Portanto, associamos o  $n$ -upla  $(q_1 - p_1, \dots, q_n - p_n) \in (\mathbb{C}^*)^n$  para  $(p_1, q_1, \dots, p_n, q_n)$ .

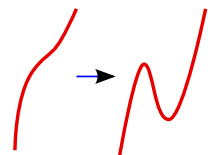
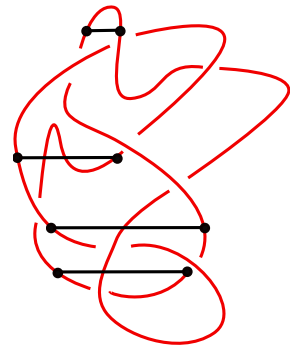
Considere, agora, a  $n$ -forma diferencial (complexa)

$$\frac{1}{(2i\pi)^n} \omega^* \left( \frac{d\zeta_1}{\zeta_1} \wedge \dots \wedge \frac{d\zeta_n}{\zeta_n} \right)$$

em  $X_n$ . Integrando-a em cada componente conexa de  $X_n$ , multiplicando com o elemento correspondente de  $Chord(n)$  e somando todos os componentes de  $X_n$ , obtemos um elemento de  $\mathcal{A}_n$ . A soma formal de todos esses elementos, por todos os valores de  $n$ , define um elemento  $\hat{\mathcal{A}}$ : esta é o invariante *Kontsevich de*  $\gamma$ , denotado por  $Z(\gamma)$ , que é um elemento de  $\hat{\mathcal{A}}$ .

Estritamente falando, isso ainda não é um invariante! Acontece que será apenas um invariante se o nó  $\gamma$  for deformado entre os nós de tipo Morse, preservando o número de pontos críticos. Isso já é um fato não trivial.

Uma deformação geral de  $\gamma$  poderia introduzir uma *corcunda*.



No entanto, a mudança nesta introdução de uma corcunda pode ser completamente descrita. Seja  $Z(H)$  a invariante da corcunda na margem.

Pode-se mostrar que se a  $t$ -coordenada de um nó  $\gamma$  tiver  $2c$  pontos críticos, o quociente

$$I(K) = Z(K)/Z(K)^{c/2} \in \hat{A}$$

é um invariante verdadeiro do nó  $\gamma$ , para qualquer isotopia, que é qualquer deformação do nó, evitando a criação de pontos duplos.

Mas ainda tenho que justificar a divisão por  $Z(K)^{c/2}$  na álgebra  $\hat{A}$ . Fato que não é difícil, pois é fácil ver que  $Z(K)^{c/2}$  tem o formato  $1 + a$  com  $a$  de grau  $> 1$ , de modo que o inverso de  $1 + a$  é  $1 - a + a^2 - a^3 - \dots$ .

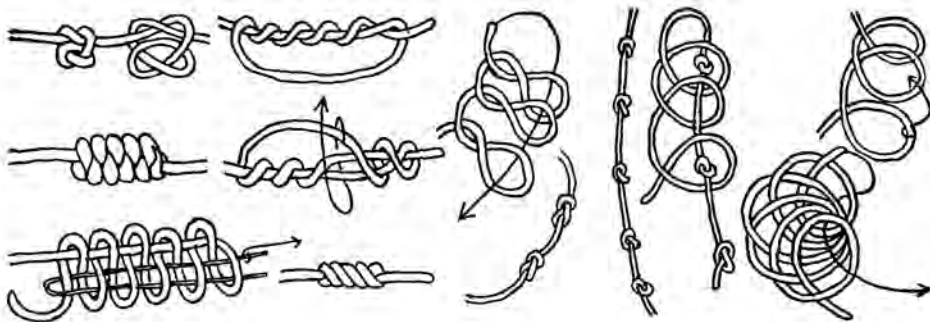
Eu não provei absolutamente nada.

Eu não expliquei em que sentido esse invariante é universal. De fato, não se sabe se dois nós são equivalentes se, e somente se, eles tiverem um mesmo invariante: isso seria fantástico.

Para uma apresentação detalhada, recomendo fortemente o artigo <sup>180</sup> e este livro <sup>181</sup>.

<sup>180</sup> S. Chmutov and S. Duzhin. [The Kontsevich integral](#). *Acta Appl. Math.*, 66(2):155–190, 2001.

<sup>181</sup> S. Chmutov, S. Duzhin, and J. Mostovoy. [Introduction to Vassiliev knot invariants](#). Cambridge UP, 2012.



Mais nós do livro dos nós de Ashley. †



Caspar David Friedrich:  
Árvore dos corvos.

+

## Posfácio

Nosso PASSEIO ACABOU. Vagamos por muitas florestas matemáticas, vimos realmente muitas árvores, mas nossa viagem definitivamente não era um caminho geodésico, e espero que meu leitor queira viajar mais e explorar novos territórios com muito mais detalhes e talvez mais a sério.

Já que nosso passeio foi uma espécie de circuito fechado que começou com o *Caminhante sobre o mar de neblina*, de Caspar David Friedrich, acredito ser apropriado, neste final, admirar *Árvore dos corvos* do mesmo artista, de 1822. Naquele período, Gauss sonhava com a geometria não euclideana.

A pintura foi escolhida como frontispício para um dos meus livros matemáticos favoritos que também lida com árvores <sup>182</sup>, embora muito diferente dos caminhos e das vistas que experimentamos durante o nosso passeio. O próximo destino para o meu leitor?

<sup>182</sup> J.-P. Serre. *Trees*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003. Traduzido do original em Francês por John Stillwell, Segunda impressão corrigida da tradução para o Inglês de 1980.



Caspar David Friedrich:  
Mann und Frau in Betrachtung des Mondes (Homem e mulher contemplando a Lua) (1818–1824). +

# Agradecimentos

A matemática seria muito triste se não pudesse ser compartilhada com os outros. Bons amigos e colegas se ofereceram para ler rascunhos preliminares deste livro.

Thierry Barbot foi um dos meus primeiros alunos de doutorado e agora é um renomado especialista em sistemas dinâmicos e geometria lorentziana. Ele propôs espontaneamente comentar sobre uma versão inicial do manuscrito.

Eu conheci Grant Cairns há muito tempo quando éramos ambos alunos, e escrevemos dois artigos juntos. Ele não é especificamente um especialista em singularidades, mas tem um amplo interesse em todos os aspectos da matemática. E também, como eu, gosta de passeios.

Hesitei em pedir a opinião de Pierre de la Harpe, já que estou bem ciente de seus padrões de qualidade. Suas sinceras observações foram totalmente apreciadas.

Jos Leys é artista e geômetra e colaboramos muito, principalmente na produção de filmes matemáticos. Muitas das imagens mais sofisticadas são devidas a ele.

Patrick Popescu-Pampu é um dos melhores especialistas da teoria da singularidade e ama a história da matemática. Somos ambos colaboradores de Henri-Paul de Saint Gervais.

Meu colega Bruno Sévenec sabe tudo em matemática. Ele detectou um grande número de erros.

Christopher-Lloyd Simon concordou em fazer o papel hipotético do “graduando motivado”, tentando decifrar essas anotações. O cobaia se transformou em colaborador. Suas sugestões melhoraram os principais resultados do manuscrito de



Thierry Barbot.



Grant Cairns.



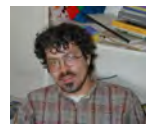
Pierre de la Harpe.



Jos Leys.



Patrick Popescu Pampu.



Bruno Sévenec.



Christopher-Lloyd Simon.



maneira significativa.

Uma versão deste texto estava disponível em [Arxiv](#) e em [Notas de matemática aberta do AMS](#) em novembro de 2016. Desde então, dois outros ex-alunos examinaram o livro em detalhes e enviaram comentários muito interessantes.

Serge Cantat geralmente lida com geometria e dinâmica algébrica no domínio complexo.

Michele Triestino está mais interessado em probabilidade, teoria ergódica e física teórica, basicamente ausente deste passeio!

Tive o privilégio de ministrar um curso de pós-graduação sobre esse tema no IMPA, no Rio de Janeiro, em janeiro e fevereiro de 2017. Entre os alunos que participaram das aulas, gostaria de destacar sete, que iniciaram uma tradução para o português e o espanhol: Aguiar Alina Sotolongo, Azul Fatalini, Carlos A. Gomes, Esteban Arreaga, Filipe Bellio da Nóbrega, Gregory Cosac e Thiago Dourado Dali Shen, pós-doutorando no IMPA, também me enviou uma longa lista de comentários. Também recebi comentários de Jean Barge, Brian Davey, Maxim Kontsevich, Manuel Ojanguren e Jean-Pierre Serre.

É um grande prazer agradecer a todos.

Não vou esquecer o leitor anônimo que me enviou chocolates porque gostou do livro. Identificou-se como Paul Ynomial, rue du nœud de trèfle, Ker Lann.

Eu provavelmente não deveria acrescentar o comentário óbvio de que sou responsável por todos os erros remanescentes.

É também um prazer agradecer a três instituições importantes.

O [Centre National de la Recherche Scientifique](#) tem me apoiado por muitos anos.

O [Instituto de Matemática Pura e Aplicada](#), Rio de Janeiro, proporcionou uma excelente atmosfera científica e paz suficiente para escrever o primeiro rascunho deste livro.

A [Unité de Mathématiques Pures et Appliquées da l'École Normale Supérieure de Lyon](#) tem sido a minha instituição de origem há muito tempo.

Étienne Ghys

Paris, 7 de Abril 2019



Serge Cantat.



Michele Triestino.



## Crédito das imagens



O texto e as ilustrações sem o símbolo † foram produzidos por Étienne Ghys, que renunciou a todos os direitos autorais e direitos relacionados ou vizinhos. Você pode copiar, modificar e distribuir, mesmo para fins comerciais, tudo sem pedir permissão.

A mesma licença CC0 aplica-se as imagens de [Jos Leys](#) nas seguintes páginas:

[7](#), [7](#), [53](#), [126](#), [126](#), [130](#), [131](#), [140](#), [140](#), [162](#), [170](#), [173](#), [173](#), [174](#), [178](#), [180](#), [181](#), [202](#), [182](#), [183](#), [185](#), [186](#), [188](#), [194](#), [195](#), [205](#), [206](#), [206](#), [206](#), [207](#), [211](#), [211](#), [212](#), [213](#), [214](#), [230](#), [248](#), [249](#), [257](#), [260](#), [316](#), [321](#),

e os seguintes documentos de [archive.org](#):

[Gauss's dissertation](#): [6](#) — [Colson's translation](#): [46](#) — [Buffon's translation](#): [50](#) — [Cramer's book](#): [69](#) — [Cramer's book](#) : [64](#) — [Gauss figure](#): [85](#) — [Serret's book](#): [95](#) — [Serret's book](#): [87](#) — [Divergentibus](#): [96](#) — [Architecture book](#): [126](#) — [Cable](#): [255](#) — [Ashley](#): [326](#),

bem como as seguindo [Wikipedia](#):

[Wanderer](#): [ii](#) — [Haeckel](#): [16](#) — [Catalan](#): [22](#) — [Catalan](#): [22](#) — [Catalan2](#): [27](#) — [Schroeder](#): [40](#) — [Hiparco](#): [40](#) — [US signs](#): [41](#), [205](#) — [Plutarch](#): [43](#) — [Newton](#): [49](#) — [Cramer](#): [68](#), — [d'Alembert](#): [83](#) — [Euler](#): [97](#), — [Hardy](#): [101](#) — [Cauchy](#): [106](#) — [Darwin finches](#): [123](#) — [Möbius](#): [125](#) — [Recycling](#): [131](#) — [Olympic](#): [138](#), — [Noether](#): [151](#) — [Cremona](#): [155](#) — [Paris](#): [167](#) — [G. Doré](#): [177](#) — [Puisseux](#): [193](#) — [Weierstrass](#): [197](#) — [Dürer](#): [218](#) — [Cerejeira](#): [228](#) — [Epicycles](#): [256](#) — [Gauss](#): [262](#) — [Gauss's signature](#): [264](#) — [Árvore dos corvos](#): [330](#) — [Mann und Frau](#): [332](#),

e

[Four seasons](#): [9](#) — [Árvore da vida](#): [10](#) — [Mississippi](#): [38](#) — [Circular I Ching](#) : [311](#) — [Chord](#): [322](#).

As outras ilustrações são publicadas sob a [Creative Commons](#). As licenças são descritas a seguir:



Signs: 3, 152, 210, 229 — Knuth: 17, 19 — Mobile: 34 — : 34 — : 37 — Cauldron: ?? — Cereja : 239 — IconPark : 278 — Nasa: 312.



Earth: 128 — Projeções de uma curva : 161 — Logo: 139 — Dimensions: 166 — Hopf: 169 — Flammarion: 175 — Geocentric: 175 — Pointe: 192 — Stasheff: 229.



Agradeço às seguintes pessoas e instituições por permitirem que eu use alguns dos seus documentos.

Maxim Kontsevich: 1 — Bib. Diderot ENS Lyon: 4 — Cambridge University library: 5, 55, 51, 52, 56, 56, 59, 63 — Museu de Arte da USP: 6, 136 — Frédérique Bassino: 35 — Pierre Gallais: 58 — U. Göttingen: 74, 321 — Scot Kolodziejcki: 75, 313 — Institut de France: 86 — Gallica: 201 — Étienne Lecroart: 130 — Sylvie Pic: 133 — Laurent Bartholdi: 163, 133, 194 — Tadashi Tokieda: 141 — Ton Marar: 142 — Barry Sobel: 150 — Tony Philips: 168 — 191, André Nachbin: Riemann cartoon. — Olivier Joseph: 200 — John Milnor: 203 — Rémy Oudompheng: 221 — Robert Coquereaux: 227 — Eliane Loday: 226 — Anatoly Fomenko: 231, 232, 233, 234, 236, 237, 237 — Yifan Hu: 252, —

Este livro foi elaborado usando o “[tufte-latex Document class](#)” inspirado pelo trabalho de [Edward Tufte](#).

Limitei os hiperlinks da bibliografia a artigos de acesso livre. Disseram-me que quase todas as referências são facilmente acessíveis em sites cuja legalidade é questionável.