

# Dicionário de Matemática

Praciano-Pereira, T

Sobral Matemática

10 de outubro de 2015

[tarcisio@member.ams.org](mailto:tarcisio@member.ams.org)

Textos da Sobral Matemática

Editor Tarcisio Praciano-Pereira,

[tarcisio@member.ams.org](mailto:tarcisio@member.ams.org)



Este é um projeto em andamento que talvez se torne permanente, um dicionário de Matemática em Português. Há diversas obras semelhantes em outras línguas, e a própria *wikipedia* representa um exemplo de enciclopédia.

Coloquei “dicionário de Matemática” mas isto não representa algo definitivo, pode, em algum momento passar a ser “dicionário Brasileiro das Ciências Exatas” que estou evitando, no momento, porque eu não teria capacidade para carregar sozinho este projeto, mas se a idéia vingar e mais gente se aproximar, mudamos o nome, afinal isto aqui é apenas um arquivo eletrônico...qualquer coisa pode ser rapidamente alterada, e  $\text{\LaTeX}$  permite que isto seja feito com grande rapidez. Portanto, não tenha receio de criticar, corrija os erros que encontrar, por exemplo. Como trabalho como  $\text{\LaTeX}$ , em um ambiente `Linux`, está todo o sistema de produção do dicionário automatizado: depois de redigido, a compilação e o envio para o site está tudo automaizado, basta executar um `make livro` e `pluft` - vai para o site. Pode criticar que é fácil corrigir, ou sugerir verbetes, com o corpo do mesmo.

O número de verbetes desta edição ainda é pequeno, não fiz nenhuma contagem mas acredito que seja algo no entorno de 400, mas começou com 50 e o objetivo é que o trabalho apareça e aos poucos cresça, oxalá com a colaboração de outros autores.

Mesmo assim vou me aventurar a construir aos poucos este projeto, tentando não competir com a *wikipedia*, que seria um desastre, mas tentando oferecer uma informação resumida, com indicações de onde se possa encontrar informações mais completas sobre os termos que aparecem na literatura científica. Há outro objetivo que tenho em mente, *criar curiosidade!* Como seria impossível dizer tudo, e o projeto Bourbaki nos lembra desta impossibilidade, ao criar um verbete tento ir o mais longe que possa sem ultrapassar os limites do bom senso, e o resultado não pode ser outro, estimular a curiosidade do leitor para ir em busca de mais informações e para isto existe a *wikipedia*.

Ao mesmo tempo convido colegas que desejem contribuir para este projeto, e aqueles que o fizerem lhes será garantido o lugar de co-autor na capa do dicionário, não importa qual o tamanho da contribuição. Uma forma de contribuir pode ser corrigindo algum erro, ou incompletitude que for observada no que aqui for publicado, tendo sempre em mente que o objetivo não pode ser enciclopédico.

Entretanto eu não vou estabelecer objetivos! Este é um projeto aberto, no sentido de código aberto, os que desejarem participar são bemvidos e juntos refaremos as regras do projeto.

Para participar, basta enviar-me o seu verbete, por e-mail, eu incluo a sua contribuição e o seu nome na lista de autores. Sempre me lembro do projeto de `Linux`, que começou com um sistema que podia abrir a gaveta de disquetes, em 1991, e quando chegou a ser capaz de criar uma rede de computadores, acendeu o interesse do mundo inteiro! Ousadia?

A estrutura dos verbetes pode ser vista neste exemplo:

**redundante\*\*** Que contem repetições de informações.  
Uma base ortogonal de vetores é um exemplo de sistema não

redundante. Uma base `\index{redundante}` pode ser ainda redundante.

- O verbete dentro do “ambiente”

```
\Underline{\bf }
```

Observe que

```
\underline{\bf }
```

é um comando do  $\text{\LaTeX}$ , e eu criei o comando

```
\Underline{\bf }
```

para marcar melhor onde termina um verbete e onde começa outro.

- inclua

```
\index{ }
```

contendo as palavras-chave do seu verbete, serve para criar o índice remissivo ao final.

Basta copiar este esqueleto

```
\Underline{\bf seu verbete} \index{seu verbete}
\index{verbeta, o seu}...
```

e preencher com sua contribuição. O índice deve ser, sempre que possível no singular, em letra minúscula quando não for um nome próprio. É uma regra de normatização para evitar entradas duplas na lista de índices:

```
Em vez de
raízes, funções polinomiais
use
raíz, função polinomial.
```

Inclua fórmulas, figuras (por favor, me envie no formato eps - “encapsulated postscript” - eu não sei ainda trabalhar com outros formatos - ou me ensine como!

Ao contribuir, por favor, tente se adequar às regras da publicação que objetivam apenas criar uma organização, mas como qualquer outro aspecto do projeto, se encontra livre para discussão e alteração. Redija o seu verbete em  $\text{\LaTeX}$  e analise o que já está feito para tentar criar alguma coisa seja parecida. Se quiser propor modificações, não tenha dúvidas em fazê-lo, mas pense que sejam exequíveis uma vez que somos nós, os autores, os que administram o projeto. Por exemplo inclua a indexação dentro do verbete, se tiver dúvidas como isto é feito, pergunte-me, mas eu logo vou criar um arquivo de FAQ para responder perguntas frequentes.

A estrutura inicial é muito simples, afinal, neste momento há um pouco mais de 400 verbetes... quando o trabalho atingir um nível adequado eu vou fazer uma rodada de discussões com os envolvidos para encontrar uma estrutura mais adequada, possivelmente dividindo em capítulos que reúnem os verbetes pela letra inicial, como é costume em dicionários, ou qualquer outra forma de sistematização o que é extremamente simples de fazer com apoio de  $\text{\LaTeX}$ . Por exemplo, logo terei que ter um corpo editorial para ter cuidado com áreas específicas porque o meu conhecimento não é, e nem pretende ser, enciclopédico.

Como todo bom livro em  $\text{\LaTeX}$ , este tem um índice remissivo alfabético que se encontra no final. Construção de índices é uma das facilidades que o  $\text{\LaTeX}$  nos oferece e eu o estou usando aqui. Chamo sua atenção para uma palavra-chave deste índice, *erro*. Sempre que eu observar, ou alguém observar um erro, vou corrigí-lo deixando no local uma indexação para o erro observado.

O trabalho promete ser divertido, lutei um pouco para redigir “aproximação” e qualquer crítico deve encontrar no texto incompletitudes ou imprecisões, e sou um “especialista” da área de *aproximação*... não tenha pudor, critique!

Este dicionário está sendo compilado com  $\text{\LaTeX}$  e o trabalho está sendo automatizado com `make` um programa de domínio público produzido e distribuído pela fundação FSF. A saída de dados é um arquivo de tipo `pdf` produzido com `pdfTEX` e que pode ser lido com `xpdf` que também é de domínio público. Tudo isto rodando dentro dum ambiente Debian/GNU/Linux.

Tarcisio Praciano-Pereira

Sobral, 23 de Janeiro de 2014

---

- **Abel, Lema de** Considere a série de potências  $S_n = \sum_{k=0}^n a_n z^n$  por comparação com séries geométricas se pode deduzir que se  $\overline{\lim} \sqrt[n]{|a_n|} = \frac{1}{r}$  então  $S_n$  converge absolutamente e uniformemente no disco  $B(0, \rho); \rho < r$ . Nada se pode dizer sobre o que acontece na fronteira deste disco. O número  $r$  é o raio de convergência da série de potências. As séries de potências definem funções de classe  $C^\infty$  no interior do disco de convergência e tais funções satisfazem às equações de Cauchy-Riemann são as funções analíticas, ou holomorfas.

---

- **absolutamente somável**

Uma série  $S_n = (\sum_{k=0}^n a_k)_n$  se diz **absolutamente somável** ou **absolutamente convergente** se a série obtida com a substituição  $a_k := |a_k|$  for convergente.

**Theorem 1 (comutatividade)** Séries e comutatividade Se  $S_n$  for **absolutamente convergente** e se  $\alpha$  for uma bijeção de  $\mathbf{N}$  então, [3, página 39]

$$\left(\sum_{k=0}^n a_k\right)_n = \left(\sum_{k=0}^n a_{\alpha(k)}\right)_n$$

A série harmônica,  $S_n = (\sum_{k=1}^n (-1)^{k+1}/k)_n$  converge mas a permutação

$$\left(\sum_{j=1}^{2^{n-1}} \frac{1}{2^n + (2j+1)} - \sum_{j=1}^{2^{n-1}} \frac{1}{2^n - 2j}\right)_n$$

não converge.

---

- **algébrico, número** são os números definidos por uma equação algébrica, por exemplo, como  $\sqrt{2}$  é uma solução da equação polinomial  $P(x) = 0$  com

$$P(x) = x^2 - 2 \quad (1)$$

então o número irracional  $\sqrt{2}$  é um número algébrico o que mostra que existem números algébricos são irracionais e então o conjunto dos números tem interseção não vazia com o conjunto dos números irracionais. Mas como há números irracionais que não são algébricos,  $\pi$  então estes dois conjuntos são distintos.

---

- **algoritmo** é um método descrevendo a execução de uma tarefa. Um programa, escrito em uma linguagem de computação, é um algoritmo. Algumas equações podem representar um algoritmo, como

$$p = dq + r; p, d, q, r \in \mathbf{N}; r < d; q > 0 \quad (2)$$

é o algoritmo da divisão euclidiana de  $p$  por  $d$ , porque, dados  $p, d$  podemos encontrar dois únicos números  $q, r$  de modo a definir a divisão de  $p$  por  $d$ . Embora este “algoritmo” seja passivo, ele é um antigo exemplo de expressão algorítmica em Matemática. Como exemplo de algoritmo, fere um pouco a concepção atual desta palavra uma vez que ele não produz os números  $q, r$ , apenas serve para testar uma quantidade finita de pares  $(q, r)$  com objetivo de encontrar um que sirva. Mas, como esta expressão podemos construir um *método*, com divisões sucessivas, e expressar esta sucessão de divisões com uma linguagem de programação que seria um algoritmo na concepção atual.

---

- analítica, geometria confira *geometria analítica*.

---

- ângulo é um conceito típico da geometria euclidiana ainda que usualmente deixado pouco claro. Ângulo é um conceito plano, quer dizer de um espaço de dimensão dois porque representa uma medida associada a dois segmentos de reta. Isto não nos impede de falar de ângulo num espaço de dimensão maior apenas tudo vai se passar como se acontece dentro de um plano deste espaço.

Vou aqui definir ângulo como um número, é um número que mede um arco do círculo trigonométrico  $\mathbf{S}^1$ . Na figura (fig 1), página 6, você pode ver o ângulo  $\alpha$  que a reta  $r$  faz com o eixo  $OX$ . Mas se considerarmos o *círculo trigonométrico*,  $\mathbf{S}^1$  como um padrão, podemos simplificar a linguagem dizendo apenas o *ângulo da reta  $r$* , querendo com isto dizer que colocamos o centro de  $\mathbf{S}^1$  sobre a reta e encontramos o arco  $\alpha$  determinado a partir da origem de  $\mathbf{S}^1$  até o ponto em que  $r$  corta  $\mathbf{S}^1$ .

Vou abstrair o plano cartesiano e redefinir ângulo entre duas retas quaisquer usando uma representação de  $\mathbf{S}^1$  no plano que estas duas retas determinam.

- Se os dois segmentos de reta forem paralelos, o ângulo entre eles é zero.
- Se os dois segmentos de reta não forem paralelos, então eles determinam um plano e suas retas suporte são concorrentes num ponto  $P$  e podemos traçar um círculo com centro em  $P$  sobre o qual as retas suporte determinam quatro segmentos de círculo iguais dois a dois. Considere  $P$  como a origem comum destas retas e o raio como sendo a unidade então o círculo considerado é unitário,  $\mathbf{S}^1$ . O menor segmento de círculo determinado pelas retas é o ângulo entre elas, determinando também a origem de  $\mathbf{S}^1$ .

Precisei criar a convenção de que o círculo traçado seria  $\mathbf{S}^1$ , o círculo unitário e este círculo tem um ponto inicial que é o ângulo zero. Se trata de uma *convenção*, ou uma *codificação*, escolhermos o *ponto inicial*, onde  $\mathbf{S}^1$  corta o semi-eixo positivo horizontal, no caso do plano cartesiano, ou o ponto que vai determinar o menor segmento do círculo unitário quando estivermos determinando o ângulo entre duas retas quaisquer.

Como  $\mathbf{S}^1$  tem raio 1, as coordenadas de qualquer ponto  $p \in \mathbf{S}^1$  serão os números  $\cos(\alpha), \sin(\alpha)$  do arco marcado a partir do *ponto inicial*. E isto funciona perfeitamente em qualquer plano do espaço em que as duas retas se encontrarem, um conceito plano, portanto.

$\mathbf{S}^1$  funciona como um *transferidor universal* para determinar ângulos: a medida dum arco determinado neste *transferidor*.

Ângulo é um número que fica no intervalo  $[0, \pi)$  e diremos que duas retas são perpendiculares se determinarem o ângulo  $\frac{\pi}{2}$ , paralelas se se determinarem o ângulo 0. Podemos estender o conceito de ângulo ao perímetro do círculo trigonométrico e então seria um número no intervalo  $[0, 2\pi)$ .

Algumas vezes precisamos de estender ainda mais o conceito de ângulo então é um número real qualquer... medindo o perímetro de uma curva que envolva o círculo trigonométrico. Conceitos associados: *grau*, *número de voltas de uma curva*.

Se atribuí a Euler a fórmula

$$P = (\cos(\alpha), \sin(\alpha)); P \in \mathbf{S}^1; \quad (3)$$

$$e^{i\alpha} = (\cos(\alpha) + i \sin(\alpha)) \equiv (\cos(\alpha), \sin(\alpha)) \quad (4)$$

que permite facilmente encontrarmos coseno, seno e tangente dos arcos soma.

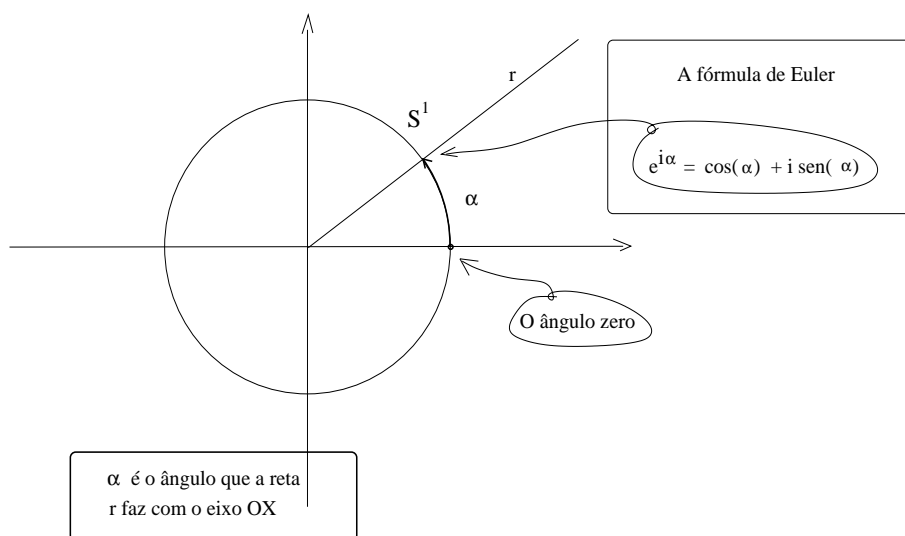


Figura 1: O ângulo da reta  $r$

Confira trigonometria.

Ao longo do tempo a Humanidade escolheu diversas formas como medição para arcos. Duas chegaram a até nós, o grau hexadecimal baseado numa divisão de  $\mathbf{S}^1$  em 360 partes chamadas *grau* e a centesimal baseada numa divisão de  $\mathbf{S}^1$  em 400 partes chamadas grau centesimal. Aqui estou falando da medida natural do ângulo, ou ainda chamada de  $\pi$ -radiano.

A *medida natural* é que melhor se adapta ao trabalho científico uma vez que é definida por uma medida tomada em cima de um padrão que é o *círculo trigonométrico*. Confira a tabela de equivalências entre estas três medidas para algumas medidas naturais bem conhecidas.



medida natural	grau hexadecimal	grau centesimal
0	0°	0°
$\frac{\pi}{4}$	45°	50°
$\frac{\pi}{2} = \frac{2\pi}{4}$	90°	100°
$\frac{3\pi}{4}$	135°	150°
$\pi = \frac{4\pi}{4}$	180°	200°
$\frac{5\pi}{4}$	225°	250°
$\frac{3\pi}{2} = \frac{6\pi}{4}$	270°	300°
$\frac{7\pi}{4}$	315°	350°
$2\pi = \frac{8\pi}{4}$	360°	400°

- **ângulo central**, na geometria euclidiana é o angulo determinado por dois vetores de mesmo módulo,  $\vec{u}, \vec{v}$ . Confira a figura (fig 2), página 7, e sua me-

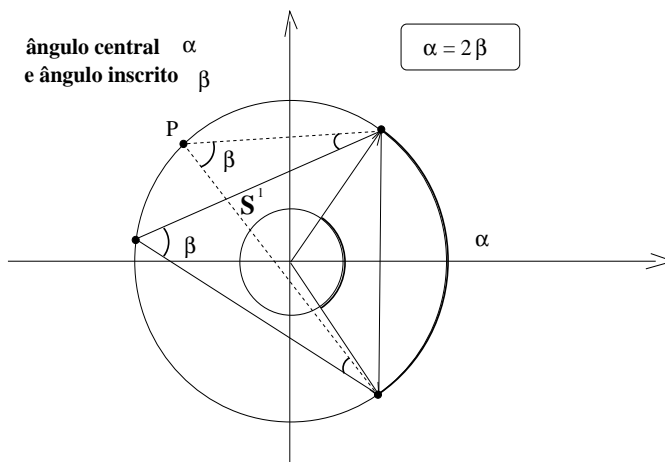


Figura 2: ângulo central

didada corresponde à medida do segmento equivalente do círculo trigonométrico que as retas suporte destes vetores determinam. Na figura (fig 2), o círculo trigonométrico  $S^1$  está representando como um círculo de raio menor que o círculo determinado pelos dois vetores  $\vec{u}, \vec{v}$ . Confira também *ângulo inscrito*.

O ângulo inscrito mede a metade do ângulo central se ambos determinarem o mesmo arco de círculo.

- **ângulo inscrito**, na *geometria euclidiana* é o angulo determinado por dois segmentos de reta de mesma origem  $P$ , em que  $P$  é um ponto sobre um círculo.

A figura (fig 2), página 7 mostra um exemplo de *ângulo inscrito*  $\beta$  comparado com o correspondente *ângulo central*,  $\alpha$ .

O *ângulo inscrito* mede a metade do *ângulo central* relativamente ao mesmo arco de círculo, confira a figura (fig 2).

**Dem**:

Dado um segmento dum círculo, considere a figura (fig 2), é possível ter-se-lhe associado um ângulo central  $\alpha$  e uma infinidade de ângulos inscritos.

Para obter um ângulo inscrito, selecione um ponto no complemento do arco  $\alpha$  e ligue-o aos extremos do segmento  $\alpha$  do círculo.

Vou mostrar que

- todos os ângulos assim obtidos são iguais ao ângulo  $\beta$ , confira a figura (fig. 2),
- $\beta = \frac{\alpha}{2}$ .

Tire do ponto  $P$  uma perpendicular à mediatriz do ângulo inscrito  $\gamma$  determinando com a mediatriz um triângulo retângulo tendo um dos catetos com origem em  $P, \overline{PQ}$ , chame este triângulo *primeiro*.

A mediatriz de  $\beta$  sendo perpendicular à corda associado ao arco  $\alpha$  determina dois triângulos retângulos um dos quais com o vértice  $P$ , chame este triângulo de *segundo* e considere o cateto  $\overline{PM}$ . As etiquetas "1" e "2" que identificam estes triângulos se encontram próximas das respectivas hipotenusas na figura (fig. 2).

Os triângulos retângulos *primeiro* e *segundo* são semelhantes porque

- O ângulo entre o cateto  $\overline{PQ}$  sendo perpendicular à mediatriz de  $\gamma$  determina com a reta suporte do cateto  $\overline{PM}$ , do *segundo* triângulo, o mesmo ângulo entre as mediatrizes de  $\beta$  e  $\gamma$ .
- 

Para demonstrá-lo vou colocar-me numa situação particular que em nada restringe a geral como depois vou deixar claro ao final.

Vou supor que o círculo na figura (fig. 3), página 8, é o círculo trigonométrico  $\mathbf{S}^1$ . Se

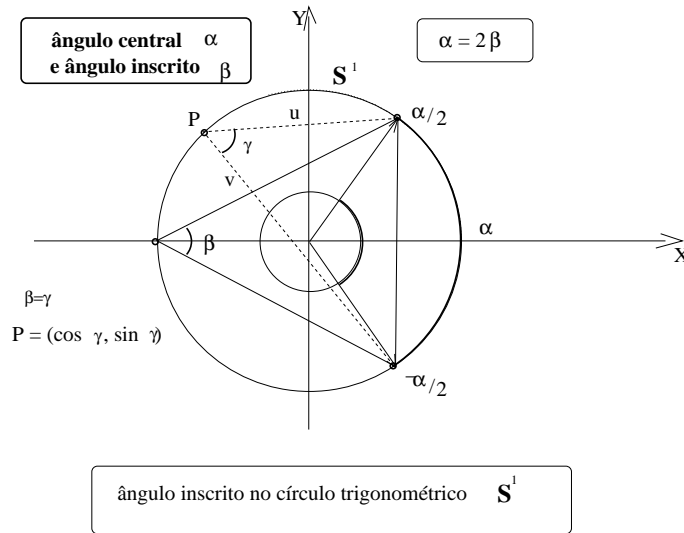


Figura 3: ângulo inscrito, círculo trigonométrico  $\mathbf{S}^1$

considerar o arco  $\alpha$  centrado em volta da origem de  $\mathbf{S}^1$  então os extremos deste arco serão  $\pm \frac{\alpha}{2}$ .

O ponto  $P$  tem coordenadas  $P = (\cos(\gamma), \sin(\gamma))$ ;  $|\gamma| > \frac{\alpha}{2}$  e os vetores que determinam o ângulo inscrito no ponto  $P$  são, respectivamente

$$\begin{cases} u = (\cos(\gamma), \sin(\gamma)) - (\cos(\frac{\alpha}{2}), \sin(\frac{\alpha}{2})); \\ v = (\cos(\gamma), \sin(\gamma)) - (\cos(-\frac{\alpha}{2}), \sin(-\frac{\alpha}{2})); \end{cases} \quad (5)$$

$$\begin{cases} u = (\cos(\gamma) - \cos(\frac{\alpha}{2}), \sin(\gamma) - \sin(\frac{\alpha}{2})); \\ v = (\cos(\gamma) - \cos(-\frac{\alpha}{2}), \sin(\gamma) - \sin(-\frac{\alpha}{2})); \end{cases} \quad (6)$$

$$\begin{cases} u = (\cos(\gamma) - \cos(\frac{\alpha}{2}), \sin(\gamma) - \sin(\frac{\alpha}{2})); \\ v = (\cos(\gamma) - \cos(\frac{\alpha}{2}), \sin(\gamma) + \sin(\frac{\alpha}{2})); \end{cases} \quad (7)$$

O produto escalar entre estes vetores é

$$\langle u, v \rangle = (\cos(\gamma) - \cos(\frac{\alpha}{2}))(\cos(\gamma) - \cos(\frac{\alpha}{2})) + (\sin(\gamma) - \sin(\frac{\alpha}{2}))(\sin(\gamma) + \sin(\frac{\alpha}{2})) \quad (8)$$

$$\langle u, v \rangle = \cos^2(\gamma) + \cos^2(\frac{\alpha}{2}) - 2\cos(\gamma)\cos(\frac{\alpha}{2}) + \sin^2(\gamma) - \sin^2(\frac{\alpha}{2}); \quad (9)$$

$$\langle u, v \rangle = 1 + \cos^2(\frac{\alpha}{2}) - \sin^2(\frac{\alpha}{2}) - 2\cos(\gamma)\cos(\frac{\alpha}{2}); \quad (10)$$

$$\langle u, v \rangle = 1 + \cos(\alpha) - 2\cos(\gamma)\cos(\frac{\alpha}{2}); \quad (11)$$

$$\langle u, v \rangle^2 = 1 + \cos^2(\alpha) + 4\cos^2(\gamma)\cos^2(\frac{\alpha}{2}) + 2\cos(\alpha) - 4\cos(\gamma)\cos(\frac{\alpha}{2}) - 4\cos(\alpha)\cos(\gamma)\cos(\frac{\alpha}{2}) \quad (12)$$

$$\langle u, v \rangle^2 = \quad (13)$$

$$|u|^2 = (\cos(\gamma) - \cos(\frac{\alpha}{2}))^2 + (\sin(\gamma) - \sin(\frac{\alpha}{2}))^2; \quad (14)$$

$$|u|^2 = \cos^2(\gamma) + \cos^2(\frac{\alpha}{2}) - 2\cos(\gamma)\cos(\frac{\alpha}{2}) + \sin^2(\gamma) + \sin^2(\frac{\alpha}{2}) - 2\sin(\gamma)\sin(\frac{\alpha}{2}) \quad (15)$$

$$|u|^2 = 1 + 1 - 2\cos(\gamma + \frac{\alpha}{2}); \quad (16)$$

$$|u|^2 = 2 - 2\cos(\gamma + \frac{\alpha}{2}); \quad (17)$$

$$|v|^2 = (\cos(\gamma) - \cos(\frac{\alpha}{2}))^2 + (\sin(\gamma) + \sin(\frac{\alpha}{2}))^2 \quad (18)$$

$$|v|^2 = \cos^2(\gamma) + \cos^2(\frac{\alpha}{2}) - 2\cos(\gamma)\cos(\frac{\alpha}{2}) + \sin^2(\gamma) + \sin^2(\frac{\alpha}{2}) + 2\sin(\gamma)\sin(\frac{\alpha}{2}); \quad (19)$$

$$|v|^2 = 1 + 1 + 2\cos(\gamma + \frac{\alpha}{2}) = 2 + 2\cos(\gamma + \frac{\alpha}{2}); \quad (20)$$

$$\frac{\langle u, v \rangle^2}{|u|^2|v|^2} = \frac{(2 - 2\cos(\gamma + \frac{\alpha}{2}))^2}{(2 - 2\cos(\gamma + \frac{\alpha}{2}))(2 + 2\cos(\gamma + \frac{\alpha}{2}))} = \quad (21)$$

$$\frac{\langle u, v \rangle^2}{|u|^2|v|^2} = \frac{4 - 4\cos^2(\gamma + \frac{\alpha}{2})}{4 - 4\cos^2(\gamma + \frac{\alpha}{2})} = \quad (22)$$

$$\frac{\langle u, v \rangle^2}{|u|^2|v|^2} = \frac{4\sin^2(\gamma + \frac{\alpha}{2})}{4\sin^2(\gamma + \frac{\alpha}{2})} \quad (23)$$

**q.e.d .**

- **antecedente** numa implicação  $A \implies B$   $B$  é o conseqüente e  $A$  é o antecedente.

- **aproximação** É um método pelo qual construímos objetos, dentro de um conjunto (ou espaço), que representam um outro objeto com um erro aceitável. Por exemplo a imagem transmitida à distância não corresponde a uma cópia exata do objeto captado, há um erro que é conseqüência da necessidade que temos de considerar apenas uma quantidade finita pixels (no caso da imagem) para compor a representação do objeto no espaço de chegada. Outro exemplo é o número  $\pi$ , que é obtido, aproximadamente, quando consideramos o quociente entre o perímetro de um polígono regular convexo inscrito no círculo unitário dividido pelo diâmetro 2. O resultado deste quociente é uma aproximação de  $\pi$ . Maior o número  $n$ , melhor a aproximação obtida de  $\pi$ .

- **arranjo** é um item da Análise Combinatória que é a parte elementar da

*combinatória*. Confira também *arranjo com repetição*, *combinação*, *permutação*.

A *análise combinatória* estuda *arranjos*, *permutações*, *combinações* e os *arranjos* podem ser *simples* ou *com repetição*. Os *arranjos simples* também são chamados de *arranjos sem repetição*.

As *combinações* são *exatamente os subconjuntos dum conjunto dado* e seria interessante que você começasse a leitura pelo verbete dedicado à *combinação*, que eu vou começar analisando as combinações, quando a maioria dos textos começa pelos arranjos, mas não é exatamente para ser diferente e sim porque desta forma vou conseguir chegar rapidamente a uma fórmula para calcular as combinações da qual vou deduzir a fórmula dos arranjos. E você logo vai ver que é um método mais concreto.

Se você tiver um conjunto  $A$  com  $n$  objetos, por exemplo,

$$A = \{1, 2, 3, \dots, n\} \quad (24)$$

então você pode produzir *arranjos* com estes elementos.

Considere um conjunto com 3 elementos:  $A = \{1, 2, 3\}$ .

Posso agora extrair vários subconjuntos de  $A$ , e vou começar pelo subconjunto vazio, todo conjunto tem o vazio como subconjunto.

- $C_3^0 = 1$  é o número de subconjuntos com zero elementos tirados de  $A$ , o vazio é único!
- $C_3^1 = 3$  é o número de subconjuntos unitários que podem ser tirados de  $A$ :  $\{1\}$ ,  $\{2\}$ ,  $\{3\}$ .
- $C_3^2 = 3$  é o número de subconjuntos 2-a-2 que podem ser tirados de  $A$ :  $\{1, 2\}$ ,  $\{1, 3\}$ ,  $\{2, 3\}$ . Agora você pode ver a diferença entre *arranjos* e *combinações*. Os *arranjos* 2-a-2 serão

$$(1, 2), (2, 1), (1, 3), (3, 1), (2, 3), (3, 2)$$

totalizando 6 arranjos, 2-a-2, que posso fazer com os elementos do conjunto  $A$ . Cada *subconjunto* “gerou” dois arranjos porque seus elementos foram *permutados*:

$$A_3^2 = 2C_3^2 = 2 * 3;$$

- $C_3^3 = 1$  é o número de subconjuntos 3-a-3 que podem ser tirados de  $A$ :  $\{1, 2, 3\} = A$ , somente o subconjunto  $A$ . Mas os arranjos serão 6:

$$(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1); \quad (25)$$

$$(1, 2, 3) \neq (1, 3, 2); \quad (26)$$

$$A_3^3 = 6C_3^3 = 6 * 1; \quad (27)$$

Pense nos arranjos de flores, um assunto comum. Tudo muda se você trocar as flores de lugar: são *arranjos de flores*, não são *subconjuntos de flores*.

Se 1 representar o branco, dois o verde e três o vermelho:

- $(1, 2, 3)$  é a flor branca seguida da verde e depois da vermelha,
- $(1, 2, 3)$  é a flor branca seguida da vermelha e depois da verde.

Os arranjos são deduzidos dos subconjuntos e a quantidade deles é obtida por uma multiplicação, observe como.

- $C_3^0 = 1 = A_3^0$  é o número de subconjuntos 0-a-0 tirados de  $A$  que é apenas o vazio. A quantidade de arranjos 0-a-0 é ainda 1. E aqui se trata de uma abstração, precisamos em Matemática de arranjos 0-a-0, não tente usar as flores agora!  $C_3^0 = 1 = A_3^0$ ;
- $C_3^1 = 3$  é o número de subconjuntos 1-a-1 que podem ser tirados de  $A$ :  $\{1\}, \{2\}, \{3\}$  e aqui não tem como alterar a ordem uma vez que cada subconjunto tem apenas um elemento:  $C_3^1 = 3 = A_3^1$ ;
- $C_3^2 = 3$  é o número de subconjuntos 2-a-2 que podem ser tirados de  $A$ :  $\{1, 2\}, \{1, 3\}, \{2, 3\}$  e já vimos que cada subconjunto podia ser repetido então  $A_3^2 = 2 * 3 = 2 * C_3^2 = 6$ ;
- $C_3^3 = 1$  é o número de subconjuntos 3-a-3 que podem ser tirados de  $A$ :  $\{1, 2, 3\} = A$ , apenas o próprio  $A$ ,  $A_3^3 = 1 * 6 = 6$ .

E tudo que precisamos saber é *que fator é este* que nos fornece o número de arranjos a partir do número de combinações:

$$A_n^p = K * C_n^p; K \text{ depende de } p; \quad (28)$$

logo acima você leu que dos subconjuntos foram feitas as *permutações dos seus elementos* para obter todos os arranjos.

Quantas são as  $P_p$ , permutações de  $p$  elementos? É este o número  $K$  que vai nos dar o valor de  $A_n^p$  multiplicando  $C_n^p$ .

As permutações são um caso particular dos arranjos e a notação é a seguinte:

$$P_p = A_p^p; \quad (29)$$

E você já viu nas equações acima:

- $P_0 = 1 = 0!$ ;
- $P_1 = 1 = 1!$ ;
- $P_2 = 2 = 2!$ ;
- $P_3 = 6 = 3!$ ;

Se você tiver  $n$  elementos,

- o primeiro pode ser escolhido de  $n = A_n^1$  maneiras diferentes;

- mas para escolher o segundo, sem repetição, você já tem apenas  $n - 1$  possibilidades,

$$A_n^2 = n(n - 1);$$

- para escolher o terceiro, sem repetição, você já tem apenas  $n - 2$  possibilidades:

$$A_n^3 = n(n - 1)(n - 2) = \frac{n!}{(n - 3)!};$$

- e sucessivamente, para escolher o  $n$ -ésimo, só lhe resta uma escolha:

$$A_n^n = P_n = n(n - 1)(n - 2) \dots 1 = n! = \frac{n!}{0!}$$

- $P_n = n!$

E chegamos à fórmula dos arranjos a partir das combinações:

$$A_n^p = p!C_n^p = p! \frac{n!}{(n-p)!p!} = \frac{n!}{(n-p)!} = \frac{P_n}{P_{(n-p)}}; \quad (30)$$

$$A_n^p = \frac{n!}{(n-p)!}; \quad (31)$$

porque será possível construir  $p!$  arranjos  $p - a - p$  com cada subconjunto com  $p$  elementos.

A fórmula que você vai encontrar em qualquer livro é esta da equação (eq. 31).

- **arranjo com repetição** é um item da Análise Combinatória que é a parte elementar da *combinatória*. Confira também *arranjo simples*, *combinação*, *permutação*.

Dado um conjunto  $A = \{1, 2, \dots, n\}$  o conjunto dos arranjos,  $p - a - p$ , com repetição destes  $n$  elementos é o *produto cartesiano*

$$A^p = \underbrace{A \times A \times \dots \times A}_p \text{ fatores} \quad (32)$$

e o número dos elementos de um produto cartesiano é o produto dos números de elementos de cada conjunto fator então

$$\mathcal{A}_n^p = n^p \quad (33)$$

O símbolo  $\mathcal{A}_n^p$  não é universalmente adotado.

Os arranjos com repetição ocorrem quando houver sentido em se repetirem os elementos dum conjunto, como é o caso com os números, *número de telefone*, *placa de carro*, palavras de uma linguagem que são *arranjos com repetição* das letras de um determinado alfabeto.

- **atlas** um conceito de geometria diferencial e também da topologia de variedades.

A ideia intuitiva vem do *atlas mundi*, como a Terra pertence a classe topológica da esfera, é impossível definirmos uma única função de uma região do plano na superfície da Terra, então precisamos de vários mapas formando um *atlas* para descrever com peças planas a superfície da Terra. É esta a ideia de um atlas em topologia ou em geometria diferencial.

**Definição 1 (atlas)** *Atlas* Se  $\mathcal{V}$  for uma variedade de dimensão  $n$ , dizemos que  $((U_i, \phi_i)_{i \in I})$  é um atlas para  $\mathcal{V}$  se  $(R_i)_{i \in I}$  for uma coleção de abertos do  $\mathbf{R}^n$  e  $(R_i \xrightarrow{\phi_i} U_i)_{i \in I}$  for uma coleção de aplicações injetivas contínuas (homeomorfismos) de abertos do  $\mathbf{R}^n$  em abertos de  $\mathcal{V}$  satisfazendo às condições seguintes:

- $\bigcup_{i \in I} U_i = \mathcal{V}$ , que garante que os mapas cobrem a variedade;
- para qualquer par  $(i, j)$  de índices,  $R_i \cap R_j \neq \emptyset$ , que garante que não haja regiões com mapas incompatíveis junto com a próxima condição;
- para qualquer par  $(i, j)$  de índices, as aplicações  $\phi_i, \phi_j$  coincidem na interseção  $R_i \cap R_j$  que garante que não haja regiões com mapas incompatíveis;

Se a variedade  $\mathcal{V}$  for diferenciável (em geometria diferencial e topologia diferencial) se precisa que  $\phi_i$  sejam difeomorfismos

Cada aplicação

$$R_i \xrightarrow{\phi_i} U_i$$

se chama um mapa do atlas.

Na figura (4), página 13, você dois mapas para o círculo, mas é preciso de quatro mapas para formar um atlas para o círculo, para isto para selecionar dois pares de *pontos antípodos*, diferentes, e a cada um destes pares corresponde um par de mapas.

---

- **autovalor** da Álgebra Linear, associado ao conceito de *autovetor*. Confira o sinônimo *valor próprio*.

---

- **autovetor** da Álgebra Linear, associado ao conceito de *autovalor* ou *valor próprio*. Confira o sinônimo *vetor próprio*.

---

- **Baire, teorema de** ainda chamado de teorema das categorias de Baire é um dos teoremas básicos da topologia dos espaços vetoriais, análise funcional, classificando os espaços pela densidade dos seus abertos em duas grandes classes, *magros*, *gordos* que ainda assim são duas classes relativas. Iremos dizer que  $X$  é *magro* em  $Y$ .

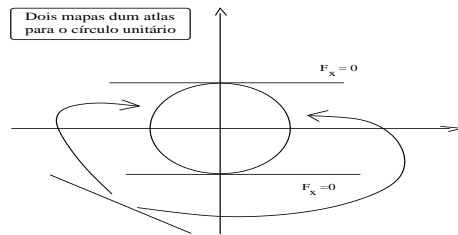


Figura 4: Um atlas para o círculo

O teorema de das categorias de Baire foi estabelecido na tese de doutoramento de René-Louis Baire em 1899.

---

- **Banach, espaço de** é um tipo de espaço que generaliza os espaços vetoriais normados de dimensão finita como  $\mathbf{R}^n$  ou  $\mathbf{C}^n$  em que os vetores são “médidos” com a *norma euclidiana*. Há duas teorias elementares que fazem esta generalização, a dos espaços de Hilbert e a dos espaços de Banach, [9, segunda parte].

A formulação do que é um espaço de Banach pode se expressar de forma absolutamente simples, com a linguagem do Cálculo no  $\mathbf{R}^n$ , apenas com a substituição dos vetores  $x = (x_1, \dots, x_n)$  pelo símbolo  $f$  representando os elementos de um certo espaço de funções.

Por exemplo, se considerarmos o conjunto de todas as funções contínuas definidas, definidas num intervalo fechado da reta,  $\mathcal{C}([a, b])$ , podemos provar que a equação

$$\|f\|_\infty = \sup_{x \in [a, b]} |f(x)| \quad (34)$$

tem as mesmas propriedades que

$$|x| = \sqrt{\sum_{k=1}^n x_k^2}; x \in \mathbf{R}^n \quad (35)$$

portanto  $(\mathcal{C}([a, b]), \|f\|_\infty)$  é um exemplo de uma estrutura semelhante a  $(\mathbf{R}^n, | \cdot |)$ .

A equação (34) recebe a denominação de *norma* caracterizando que obtivemos uma generalização do conceito tradicional “módulo”, ou seja, o módulo é um exemplo de *norma*. Os espaços vetoriais em que for possível definir uma norma se chamam espaços vetoriais normados.

O mesmo se poderia fazer substituindo o “espaço”  $[a, b]$  por um espaço topológico  $X$  e o resultado desta generalização conduziu à descoberta de propriedades topológicas finas sobre os espaços topológicos que tornam a família dos espaços de funções  $\mathcal{C}(X)$  em que  $X$  é um espaço genérico, uma teoria bastante complexa inclusive ainda com alguns resultados abertos associados aos tipos de medida que é possível definir em  $X$  associadas à dimensão<sup>1</sup> do resultante espaço  $\mathcal{C}(X)$ .

Um exemplo, dentre muitos que podemos dar, seria o espaço vetorial das *séries trigonométricas absolutamente convergentes*.

A soma (ponto a ponto) de duas tais séries é outra do mesmo tipo das anteriores, assim como o produto por um escalar (real ou complexo, e neste caso falaríamos de espaços vetoriais reais ou complexos) temos um espaço vetorial sobre um destes corpos. É interessante este caso porque ele estabelece ligação com outro tipo espaço associado aos coeficientes das séries de Fourier.

---

- **Banach-Steinhaus**, é um teorema da teoria dos espaços de Banach que

---

<sup>1</sup>Resultados muito recentes sugerem que esta seria uma forma de entender as dimensões fracionárias.



estabelece uma condição para que um conjunto de operadores lineares seja limitado. É comumente referido como “*condição de limitação uniforme*” porque foi inicialmente formulado para o espaço de Banach  $\mathcal{C}([a, b])$ , das funções contínuas, cuja norma usual e a *norma do supremo* que corresponde à “*convergência uniforme*”. Mas é apenas uma condição, usada com grande frequência, quando se deseja demonstrar ou verificar que um conjunto seja limitado num espaço de Banach.

**Teorema 1** *Banach-Steinhaus*

Considere  $X$ , um espaço de Banach,  $Y$  um espaço vetorial normado, e uma família  $(\Lambda)_{\alpha \in A \neq \emptyset}$  de transformações lineares  $\Lambda_\alpha : X \rightarrow Y$ . Então uma das condições ocorre

- (A)  $\exists M > 0; \forall \alpha \|\Lambda_\alpha\| \leq M$ , ou equivalentemente,
 
$$\exists M > 0; (\Lambda)_{\alpha \in A \neq \emptyset} \subset \text{cal}B(0, M);$$
- (B)  $\sup_{\alpha \in A} \|\Lambda_\alpha(f)\| = \infty$  para  $f \in G_\delta$ ,  $G_\delta$  é uma interseção enumerável de abertos densos em  $X$ .

No contexto  $X = \mathcal{C}(K, \mathbf{R})$  em que  $(K, \sigma, \mu)$  é um espaço topológico medido, vale a pena observar que  $\text{não}(B) \implies (A)$  e portanto o teorema de Banach-Steinhaus pode ser expresso com esta implicação tendo por redação “*a família  $(\Lambda)_{\alpha \in A \neq \emptyset}$  de operadores lineares ser essencialmente limitada implica que  $(\Lambda)_{\alpha \in A \neq \emptyset}$  é uma família de operadores lineares contínuos contida na bola  $\mathcal{B}(0, M) \subset \mathcal{C}(K, \mathbf{R})$  para algum número real  $M > 0$ .*”

A expressão “*essencialmente*” é sinônima de *quase sempre*, ou ainda que o domínio em que ela não vale tem medida zero. Como  $X$  um espaço de Banach é então, canonicamente, um espaço medido com os seus borelianos.

Em geral se omite “*essencialmente*” porque sendo uma relação de equivalência se considera o espaço quociente módulo conjuntos de medida zero tornando (A) e (B) equivalentes. Portanto o contexto natural do teorema de Banach-Steinhaus é  $\mathcal{C}(K, \mathbf{R})$  ou  $\mathcal{C}(K, \mathbf{C})$  e  $K$  é um espaço topológico compacto e de Hausdorff como é o caso de  $[a, b] \subset \mathbf{R}$ .

- **binomial, coeficiente** Procure *coeficiente binomial*

- **Bourbaki, N** Nicolas Bourbaki, o nome de um grupo de matemáticos franceses, criado por volta de 1940, idealizando reescrever toda a Matemática de forma rigorosa e axiomática.

O projeto Bourbaki *nasceu morto* porque na mesma época de sua criação Gödel havia demonstrado a impossibilidade de completção axiomática dos naturais, aliás, o que se tornou na razão de uma grande inimizade entre ele e Hilbert. Hilbert talvez tenha sido a grande influência dos membros do grupo Bourbaki uma vez que ele já vinha alimentando a ideia duma axiomática geral para Matemática já se tendo intrigado com matemático holandês Luitzen E. J.

Brouwer com a exclusão deste último do *Mathematische Annalen*. Brouwer é o fundador do *intuicionismo* como fundamentação da Matemática.

Mesmo que o seu objetivo fosse impossível, a contribuição do grupo Bourbaki para a Matemática foi muito grande tendo influenciado profundamente a discussão sobre os fundamentos. Serve de exemplo, por um lado, para mostrar que um grupo de matemáticos excelentes pode errar em sua visão geral da Matemática, como todos os seres humanos e o erro faz parte do processo de construção do conhecimento, e por outro lado, para nós alertar sobre as burocracias governamentais que podem, e costumam, parar projetos porque os burocratas não conseguem entender os objetivos difusos dos germens duma construção científica que mesmo errada, ou inválida, tem os seus méritos.

---

- **Cálculo\*\*** É uma disciplina da Matemática que estuda o comportamento das funções com o objetivo de descrever a continuidade, diferenciabilidade e integrabilidade das mesmas. Ver continuidade, diferenciabilidade e integrabilidade.

---

- **campo escalar** é uma função, em geral multivariada, e tomando valores em  $\mathbf{R}$  ou  $\mathbf{C}$ . É uma antiga denominação para funções. O adjetivo *escalar* caracteriza que o conjunto de chegada é de dimensão 1 (real ou complexa).

---

- **campo vetorial** A palavra *campo* é uma antiga denominação para funções multivariadas e que permaneceu na Física sendo usada também na literatura matemática. É mais frequente o uso de

- *campo vetorial* para funções multivariadas cuja imagem é um vetor,
- e de *campo escalar* para funções multivariadas cuja imagem é um número real ou complexo. variáveis.

Embora não haja uma uniformidade de notação é possível pensar que *campo escalar* está associado um potencial, uma função que mede a intensidade dum fenômeno, um escalar, e cuja derivada é um *campo vetorial*.

Há várias formas de apresentar este conceito. Se  $F$  for um *campo escalar*,

$$F : \mathbf{R}^n \rightarrow \mathbf{R}; J(F) : \mathbf{R}^n \rightarrow \mathbf{R}^n \quad (36)$$

então a sua derivada,  $J(F)$ , a *jacobiana de  $F$* , é um *campo vetorial*, uma função definida no mesmo domínio de  $F$  mas agora tendo tantas funções-coordenadas quantas sejam as variáveis: cada função-coordenada é uma das derivadas parciais de  $F$ . A derivada de um *campo escalar* é um *campo vetorial* que recebeu o nome jacobiana.

No caso da jacobiana, as funções coordenadas possuem uma notação particular:

$$\frac{\partial F}{\partial x_1} \cdots \frac{\partial F}{\partial x_n} \quad (37)$$

Se  $\mathbf{R}^n \xrightarrow{F} \mathbf{R}^m$ , um campo vetorial, então o valor da derivada de  $F$  em cada ponto é uma matriz (funcional) de dimensão  $m \times n$  formada com as  $mn$

*derivadas parciais* das  $m$  coordenadas de  $F$  relativas a cada uma das  $n$  variáveis de  $F$ .

A derivada de funções multivariadas recebeu o nome de *jacobiana* numa época em que não era reconhecida como *a derivada*, e algumas vezes se usa o símbolo  $J(F)$  para representá-la.

- **característica, equação** é uma equação polinomial associada às matrizes quadradas. Se  $T$  for uma matriz  $n \times n$  então a equação polinomial

$$\det(T - \lambda I) = 0; I, T \in \mathcal{L}(\mathbf{K}^n); I \text{ a identidade} \quad (38)$$

é a equação característica de  $T$ . Aqui o corpo  $\mathbf{K}$  é  $\mathbf{R}, \mathbf{C}$ . As raízes da equação característica são os *valores próprios*, ou *autovalores* de  $T$ .

A equação característica é *invariante* sob a *transformação de semelhança de matriz*, ou em outras palavras, se  $A$  for uma matriz semelhante à matriz  $T$  no sentido de que as duas representam o mesmo sistema de equações lineares, então elas terão a mesma equação característica:

$$\det(AB) = \det(A) \det(B); \det(S^{-1}) = \det(S)^{-1}; \det(SCS^{-1}) = \det(C); \quad (39)$$

$$A = STS^{-1}; \quad (40)$$

$$\det(A - \lambda I) = \det(STS^{-1} - S\lambda I S^{-1}) = \det(S(T - \lambda I)S^{-1}); \quad (41)$$

$$\det(A - \lambda I) = \det(S) \det(T - \lambda I) \det(S^{-1}) = \det(T - \lambda I); \quad (42)$$

$$\det(STS^{-1} - \lambda I) = \det(T - \lambda I); \quad (43)$$

- **cardinalidade** é a generalização do conceito de número de elementos de conjuntos finitos. O conceito “conjunto finito” precisa ser definido e uma forma de o fazer consiste da dicotomia seguinte:

1. conjunto infinito existe uma bijeção entre  $A$  e uma parte própria de  $A$ ;
2. conjunto finito não pode haver uma bijeção entre  $A$  e uma parte própria de  $A$ ;

Como um exemplo da primeira sentença, considere o conjunto  $\mathbf{N}$  dos números naturais, é possível estabelecer uma bijeção entre  $\mathbf{N}$  e  $2\mathbf{N}$ , o conjunto dos números naturais pares, que é dada pela equação  $n \mapsto 2n$ , que é uma bijeção. A primeira relação caracteriza os conjunto infinitos, enquanto que a segunda caracteriza os conjuntos finitos.

Há ainda uma subclassificação dos *conjuntos infinitos* em

- conjuntos enumeráveis que são aqueles que tiverem a mesma cardinalidade de  $\mathbf{N}$ , como  $\mathbf{Z}, \mathbf{Q}$ .
- conjuntos não enumeráveis que são aqueles que estiverem fora da classe de  $\mathbf{N}$ , como  $\mathbf{R}$ .

Cantor estabeleceu a sua *hipótese de cardinalidade*, chamada de *conjectura de Cantor* ou ainda *hipótese do contínuo*, confirmada na década de 60 do século 20, por Paul Cohen, que demonstrou que a *hipótese do contínuo* era um axioma da teoria dos conjuntos, estabelecendo que havia “classes disjuntas de cardinalidade”:

- a classe dos conjuntos finitos, cujas cardinalidades são elementos do conjunto  $\mathbf{N}$ ,
- depois a classe  $\text{card}(\mathbf{N})$  a que pertencem  $\mathbf{N}, \mathbf{Z}, \mathbf{Q}, \dots$ ,
- depois a classe  $\text{card}(\mathbf{P}(\mathbf{N}))$  a que pertencem  $\mathbf{R}, \mathbf{C}, \dots$ .

E pensou Cantor, se  $A$  for um conjunto infinito, então  $\text{card}(\mathbf{P}(A))$  é uma nova classe a qual  $A$  não pertence: o salto de cardinalidade.

Se um conjunto  $A$  for finito, o operador  $\mathbf{P}$  produz um novo conjunto  $\mathbf{P}(A)$  cuja classe é diferente da classe de  $A$  porém sem alterar sua classificação como *conjunto finito*, a classe de  $\mathbf{P}(A)$  é  $2^{\text{card}(A)} \in \mathbf{N}$ , se  $A$  for um conjunto finito. Os símbolos “ $\leq$ ” e “ $\geq$ ” continuam sendo usados porém perdendo os seus significados da *aritmética*, a nova definição destes símbolos de desigualdade é:

- $\text{card}(A) = \text{card}(B)$  Se dois conjuntos tiverem a mesma cardinalidade dizemos que existe uma bijeção entre eles.
- $\text{card}(A) \leq \text{card}(B)$  dizemos que  $\text{card}(A) \leq \text{card}(B)$  se houver uma função injetiva  $A \xrightarrow{f} B$  ou equivalentemente
- $\text{card}(A) \geq \text{card}(B)$  dizemos que  $\text{card}(A) \geq \text{card}(B)$  se houver uma função sobrejetiva  $A \xrightarrow{g} B$ .

A figura (5), página 19, mostra uma sobrejeção do conjunto  $A$  sobre o conjunto  $B$  e a forma de construir uma função sobrejetiva  $g$  usando as classes de equivalência módulo  $g$  que é simples de construir quando  $A$  for finito. Se  $A$  não for finito é “simples” aplicação do *axioma da escolha* ... selecionar, para cada classe em  $A/\approx$  um elemento de  $A$  que lhe corresponda, e como  $A/\approx$  é equipotente com  $B$  então existe uma bijeção entre estes dois conjuntos cuja composição com a anterior “seleção” dos elementos de  $A$  é a sobrejeção procurada.

---

- **categoria** É uma classe de estruturas algébricas, por exemplo, *Grupo* é a categoria de todos os grupos e *Abel* é a sub-categoria dos grupos comutativos. Entre os membros de uma mesma categoria se estabelecem *funtores* que colocam em correspondência as respectivas operações e os elementos privilegiados destas estruturas, como, por exemplo, elementos neutros. Podemos estabelecer um funtor especial entre uma categoria mais complexa, como a *EV*, dos espaços vetoriais, e outra mais simples, como *Grupo*, que é um *funtor esquecido* que esquece aspectos da estrutura para produzir uma estrutura mais simples, no caso do funtor esquecido de *EV* para *Grupo* *esqueceríamos* o produto por um escalar para obter a categoria dos grupo aditivos subjacentes ao espaço vetorial.

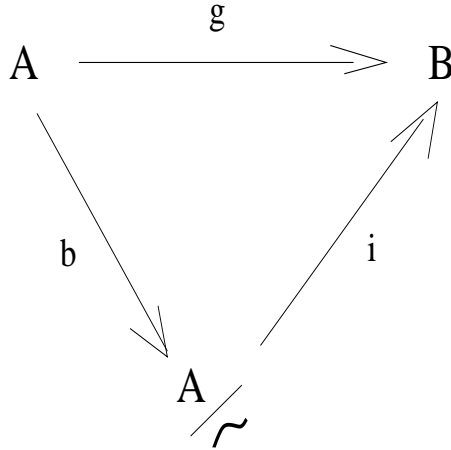


Figura 5:  $\text{card}(A) \geq \text{card}(B)$

A Teoria das Categorias foi construída por Cartan e Eilenberg que, se diz, num primeiro momento, a chamaram de *general abstract nonsense* porque, embora não houvesse dúvida do poder de abstração que esta teoria nos traria, eles sabiam que ela dificilmente se tornaria um instrumento popular dentro da Matemática. Mesmo assim uma consequência prática da Teoria das Categorias é a orientação à objeto em Computação que é uma aplicação das *categorias concretas*.

- **Cauchy, problema de.** É um problema de condição inicial, PCI, de equações diferenciais, portanto o problema

$$\begin{cases} y'(t) = f(t) \\ y(a) = b; t \in [a, b]; b \in \mathbf{R} \end{cases} \quad (44)$$

é um problema de Cauchy com condição inicial  $(a, b)$ . Observe que  $f$  pode ser uma função com valores num espaço vetorial de dimensão  $n$  e neste caso  $y, y'$  serão também funções vetoriais no mesmo espaço,  $b \in \mathbf{R}^n$ , o que tornaria a equação (44) uma equação diferencial de ordem  $n$ , a dimensão do espaço. Um problema com *condição inicial*, ou um *problema de Cauchy* tem como solução uma curva que se origina no ponto  $(a, b) \in [a, b] \times \mathbf{R}^n$ .

Um exemplo um pouco diferente e que abre possibilidades para apresentarmos uma versão mais avançada do *problema de Cauchy*, seria o *teorema da função implícita* que pode ser visto como o inverso dum problema de Cauchy de equações diferenciais parciais.

A versão no  $\mathbf{R}^3$  do *teorema da função implícita* pode ser expressa assim:

1. Temos uma curva  $\alpha$  contida em  $\mathbf{R}^3$  parametrizada no intervalo  $[a, b]$  cujo gráfico se encontra num plano paralelo ao plano  $XOY$ ;

2. um campo de vetores  $T(\alpha(t))_{t \in [a,b]}$  do  $\mathbf{R}^3$  indexado nos pontos da curva  $\alpha$ , quer dizer, uma função de  $\alpha$  em  $\mathbf{R}^3$ ;
3. Então existe uma única superfície  $z = F(x, y)$  e um número  $c$  tal que  $\alpha$  é a curva de nível  $c$  de  $F$ , ou  $\alpha = \{(x, y); F(x, y) = c\}$

A condição (1) faz de  $\alpha$  uma curva de nível de uma função  $z = F(x, y)$  e o *teorema da função implícita* (leia este teorema para ver mais detalhes) nos garante então que existe uma única função  $f$  tal que nas vizinhanças de um ponto  $(a, b, c)$  que pertence à curva  $\alpha$

1. O campo  $T$  é o gradiente de  $F$  restrito à curva  $\alpha$ ;
2.  $y = f(x); y' = -\frac{F_x}{F_y}$  numa região em que  $F_y \neq 0$ ; ou
3.  $x = f(y); y' = -\frac{F_y}{F_x}$  numa região em que  $F_x \neq 0$ .

Isto, resumidamente, é o que nos diz o *teorema da função implícita* e o *problema de Cauchy* é uma espécie de recíproca do *teorema da função implícita*.

No teorema da função implícita, existe um ponto  $(a, b, c); c = F(a, b)$  em que uma das derivadas parciais de  $F$  é diferente de zero e em geral é tudo que se conhece mas apenas com isto é possível obterem-se boas aproximações para a função  $f$  que explicita a equação  $F(x, y) = c$ , entretanto, o nosso objetivo aqui é de partir deste teorema do Cálculo para obter uma formulação do *teorema de Cauchy-Kovalevskaya*, e precisamos de uma notação adequada que vamos agora desenvolver ampliando a descrição feita do *teorema da função implícita*.

Considerando o problema recíproco, o que temos é a curva  $\alpha$  e um campo de vetores definidos sobre ela, para o *problema de Cauchy*, a curva  $\alpha$  e o campo de vetores  $T(\alpha(t))_{t \in [a,b]}$  do  $\mathbf{R}^3$ , formado pelo gradiente de  $F$  ao longo da curva  $\alpha$  se chamam de dados iniciais do problema, ou algumas vezes se os chama de *dados de Cauchy* do problema. A solução é a superfície  $graf(F)$ .

O campo vetorial dos gradientes pode ser substituído por um campo escalar, das derivadas na direção da normal à curva  $\alpha$ , e a curva  $\alpha$  é uma variedade de dimensão 1 que pode ser substituída por uma variedade de dimensão  $n$  quando considerarmos o problema em  $n$  variáveis.

O problema de Cauchy é uma generalização da recíproca do *teorema da função implícita* que pode ser expressa assim:

O sistema de equações  $\frac{\partial^k u_j}{\partial x^k} = f_j(x, y_1, \dots, y_n, (\frac{\partial^i u_j}{\partial x^i}), (\frac{\partial^i u_j}{\partial y_j^i})_{j,i})$

- temos uma variedade de dimensão  $n$ ,  $\alpha$ ;
- temos um campo vetorial definido sobre  $\alpha$

e queremos encontrar uma família de variedades de dimensão  $n + 1$  com a restrição de  $f_i$  seja analítica em todas as variáveis.

Sonja Kovalevskaya, uma matemática russa, demonstrou em sua tese de doutorado a versão que hoje se conhece, chamada Cauchy-Kovalevskaya generalizando a expressão que Cauchy havia formulado.

- **Cauchy, teorema de** O teorema de Cauchy fornece a expressão integral de uma função analítica a partir dos seus valores na fronteira do domínio sendo portanto a solução de um PVF, problema de valor na fronteira, ou em inglês, BVP-boundary value problem. Se  $\Omega$  for um domínio do plano e  $\gamma = \partial\Omega$  então

$$f(a)Ind_{\gamma}(a) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(t)dt}{t-a}; \quad (45)$$

em que podemos identificar um *produto por convolução* de  $f$  pelo núcleo de Cauchy.

Esta integral é chamada de *integral de Cauchy*.

O número  $Ind_{\gamma}(a)$ , é o *índice de  $\underline{a}$ , relativamente à curva  $\gamma$* , conta o número de vezes que  $\gamma$  circula à volta do ponto  $a$ , por, exemplo, se  $a$  for um ponto no exterior da curva  $\gamma$  a fórmula de Cauchy retorna zero. Não podemos calcular os valores de  $f$  em cima de  $\gamma$ . Das duas uma, ou conhecemos estes valores, o que significa que  $f$  é uma *condição de fronteira* e então não precisamos aplicar a fórmula de Cauchy para calculá-los ou  $f$  é uma função ou mesmo uma *medida* definida na fronteira mas, possivelmente, com valores não explicitos pelo menos em alguns pontos, pontos de salto, por exemplo, ou melhor dizendo não é uma função definida ponto a ponto e sim uma medida definida em  $\gamma$  para a qual não tem sentido buscar valores nos pontos: medidas são funções aditivas de conjuntos.

Se considerarmos  $a \in K \subset \Omega$ ,  $K$  um conjunto compacto, então a integral de Cauchy é *absolutamente convergente* podendo ser derivada. Usando a propriedade da derivada das convoluções

$$(f * \eta)' = f' * \eta \Rightarrow f'(a) \in \mathbf{C} \quad (46)$$

vemos que uma função definida pela integral de Cauchy é derivável e sua derivada novamente é definida pela *integral de Cauchy*. Isto significa que a derivada  $f'(a)$  é um número complexo ou seja uma função linear complexa.

Assim a integral de Cauchy define uma função complexa no interior do domínio  $\Omega$  limitado pela curva  $\gamma$ . A notação usual para  $f$  como função de  $\mathbf{R}^2$  em  $\mathbf{R}^2$  é  $f = u + iv$  em que  $u, v$  são funções reais bivariadas com

$$f'(a) = \begin{bmatrix} u_x(a) & u_y(a) \\ v_x(a) & v_y(a) \end{bmatrix} = J(f); \quad (47)$$

uma função linear complexa.

As funções lineares complexas formam um subconjunto próprio das funções lineares do  $\mathbf{R}^2, \mathcal{L}(\mathbf{R}^2)$ , elas são representadas pelas matrizes da forma

$$\mathcal{L}(\mathbf{R}^2) \ni \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \approx (\alpha - i\beta) \in \mathbf{C} \quad (48)$$

$$\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \equiv (\alpha - i\beta)(x + iy); \quad (49)$$

são as funções lineares do  $\mathbf{R}^2$  que se identificam com a multiplicação por um número complexo, quando vistas como funções lineares complexas, como está indicado na equação (eq. 49).

Se aplicarmos a equação (eq. 49) à jacobiana de  $f$  no ponto  $\underline{a}$ ,  $f'(a)$  vamos obter as chamadas *equações de Cauchy-Riemann*:

$$J(f)(a) = \begin{bmatrix} u_x(a) & u_y(a) \\ v_x(a) & v_y(a) \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix}; \quad (50)$$

$$\begin{cases} u_x(a) = v_y(a); \\ u_y(a) = -v_x(a); \end{cases} \quad (51)$$

O sistema de equações diferenciais parciais de primeira ordem na (eq. 51) são as *equações de Cauchy-Riemann*.

Usando a equação (eq. 48) podemos escrever facilmente a derivada de  $f$  no ponto  $a$

$$f'(a) = u_x(a) - iu_y(a)$$

Como  $f$  pode ser uma medida sobre  $\gamma$  então sua derivada pode não ser mais uma medida e sim uma distribuição, entretanto a regularização pela convolução resulta numa função de classe  $C^\infty(\Omega)$  no interior do domínio  $\Omega$ .

Uma consequência simples da regularização por convolução é que  $f$ , definida pela integral de Cauchy é infinitamente derivável e sempre suas derivadas definidas pela integral de Cauchy agora pela derivada correspondente de  $f$  na fronteira.

*Uma pequena observação, em passant, foi resolvido, assim, o sistema de equações diferenciais parciais de primeira ordem, as equações de Cauchy-Riemann. As contas que acabei de fazer mostram que a integral de Cauchy define uma solução para estas equações. A recíproca é mais fácil, qualquer solução das equações de Cauchy-Riemann define uma função complexa com derivada complexa, e sua restrição à curva  $\gamma$  é o valor de fronteira da integral de Cauchy pela unicidade da equação (eq. 45). Os matemáticos levaram século e meio para compreender que a chamada teoria das funções complexas era a construção da solução desta equação diferencial. Resolver esta equação de forma completa precisa dumas 450 páginas, é o número de páginas do livro de Titchmarsh, *Theory of Functions*<sup>2</sup>.*

Há vários detalhes importantes que foram ignorados nesta curta teoria, por exemplo, a definição de função analítica.

As funções dadas pela integral de Cauchy são as funções analíticas. Como elas são infinitamente diferenciáveis, então elas tem uma *série de Taylor* em qualquer ponto do domínio  $\Omega$  e as séries de Taylor estão associadas a um raio de convergência que definem um disco dentro do qual elas convergem uniforme. O raio de convergência será o daquele disco contido em  $\Omega$ , entretanto existe uma fórmula para o cálculo do raio de convergência que pode produzir um raio maior do que o disco que se encontra contido em  $\Omega$ , se assim for, há um domínio que contem  $\Omega$  onde  $f$  é analítica este tópico se chama de *extensão analítica* e

<sup>2</sup>que pode ser livremente baixado da página <https://archive.org/details/TheTheoryOfFunctions>



trata da busca do maior domínio de analiticidade de uma função. O domínio máximo é o plano complexo e as funções analíticas com este domínio se chamam de *inteiras*, é o caso dos polinômios, mas também da função exponencial, as funções trigonométrica.

As funções racionais são uma outra classe interessante de funções que incluem um novo tópico, os polos. As funções racionais são os quocientes de funções polinômiais e aqui temos as raízes do numerador e as raízes do denominador: os polos. As séries de Laurent entram neste contexto.

Seriam necessárias as 450 páginas para cobrir todos os detalhes, leia o livro de Titchmarsh, mas se prepare para encontrar um ponto de vista diferente. Os livros clássicos começam com a teoria geral das séries definindo funções analíticas como sendo aquelas que têm uma série de potências convergente num disco, a partir de onde chegam no domínio de analiticidade. Depois partem dum ponto aparentemente inteiramente diferente para chegar na integral de Cauchy e finalmente mostrar que as funções definidas pela integral de Cauchy tem séries de potências. Optei em parte pelo método de Rudin, *Real and Complex analysis*, que consegue fazer esta teoria num capítulo de 40 páginas mas apresenta a a integral de Cauchy no meio do capítulo.

- **Cauchy-Riemann, equações de** Considere a função complexa

$$w = f(z) = u(z) + iv(z); u, v : \Omega \rightarrow \mathbf{R} \quad (52)$$

Se  $f$  for diferenciável, como função complexa, então

$$u_x = v_y; u_y = -v_x; \text{(Cauchy-Riemann)} \quad (53)$$

Estas equações são necessárias e suficientes para que o  $(u_x, u_y)$ , ou  $(-v_x, v_y)$ , sejam *diferenciais exatos* com a integral de linha se anulando sobre qualquer curva fechada dentro do domínio de validade das equações de Cauchy-Riemann,  $\Omega$ , portanto, pelo teorema de Green,

$$f' = u_x + iu_y = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} v_y & -v_x \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} u_x & u_y \\ -u_y & u_x \end{pmatrix} \quad (54)$$

em que a matriz na equação (54) é a matriz da transformação linear tangente de  $f$  vista como função vetorial de variável vetorial. As funções de  $\mathbf{R}^2$  em  $\mathbf{R}^2$  que satisfizerem às equações de Cauchy-Riemann, são um sub-espço vetorial do espaço das funções deriváveis de  $\mathbf{R}^2$  em  $\mathbf{R}^2$ , as funções analíticas ou deriváveis no sentido complexo.

- **ciclo** é um conceito difuso, e muito importante em diversos aspectos da ciência. Poderíamos defini-lo como *comprimento de onda*, entretanto nem sempre é possível entender claramente o que seria uma onda.

Em computação, nos computadores, o ciclo é o *comprimento de onda* de uma célula de *quartz* excitada por um pulso de energia elétrica que é o substituto do “*cabelo*” nos relógios mecânicos existentes até a década de 60 do século 20.

O pêndulo é uma outra forma estabelecer o ciclo, e na verdade o *cabelo* é um tipo de pêndulo, ambos representam a oscilação que hoje é em geral usada na célula de *quartz*. Possivelmente nós iremos retornar ao uso destes *osciladores* mecânicos dentro do quadro de economia de energia que temos pela frente.

Entre os seres vivos poderia ser o *comprimento médio de vida*, os seres vivos são estruturas que podem ser muito complexas, agregados de outros seres vivos, as células, e estas tem distintos ciclos. Seria possível definir um *espectro* para seres vivos que identificasse cada espécie?

- **círculo**. Confira *equação do círculo*

- **coeficiente binomial** Dados dois números  $a, b$  podemos expressar  $(a + b)^n$  como uma soma em que aparecem  $a, b$  e os coeficientes  $C_n^k$ , os números combinatorios:

$$(a + b)^n = \sum_{k=0}^n C_n^k a^k b^{n-k} = \sum_{k=0}^n C_n^k a^{n-k} b^k \quad (55)$$

Esta afirmação é conhecida como *teorema do binômio de Newton*.

Estes coeficientes, conhecidos também como *números binomiais* se dispostos em linhas crescentes pelo índice  $n$  da potência, formam o *triângulo de Pascal* também, possivelmente, já conhecido por matemáticos chineses há oito mil anos.

**programa em Python: 11 linhas do Triângulo de Pascal**

```

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
1 8 28 56 70 56 28 8 1
1 9 36 84 126 126 84 36 9 1
1 10 45 120 210 252 210 120 45 10 1
1 11 55 165 330 462 462 330 165 55 11 1

```

A figura (27) página 71, mostra as  $n = 15$  primeiras linhas do *triângulo*, construídas com um programa em *python* que pode imprimir um número arbitrário de linhas do *triângulo de Pascal* em alguns segundos, mesmo que  $n$  seja grande.

É interessante observar que se  $a = b = 1$  se obtém esquematicamente o *triângulo de Pascal* e se pode ler em cada linha uma nova potência de 11, apenas temos que adaptar à base de numeração. Por exemplo, até a quarta potência se podem ler as potências de 11 diretamente do *triângulo*, dentro do nosso sistema de base 10, mas para  $11^5$  temos que incluir “10” como um algarismo:

1 5 10 10 5 1

afim de ainda ler uma potência de 11, portanto, passar à uma base de numeração em que “10” (ou outro símbolo) seja o dígito para representar o “10” da nossa base decimal. Por exemplo, na base hexadecimal, o símbolo A representa o “algarismo 10”:

1 5 A A 5 1

é o resultado que você vai ver no visor de uma máquina de calcular que tenha a habilidade com números na base hexadecimal, figura (6), página 25, efetuando o cálculo  $11^5$ .

Mas podemos simplesmente considerar “10” como um novo algarismo... e expandir a base de numeração progressivamente a medida que lemos as linhas do *triângulo*.



Figura 6:  $11^5$  na base hexadecimal

Mas simples, aplique a *regra de passagem para a casa seguinte* usada nas adições multiplicações e você pode ler, diretamente do *triângulo de Pascal* qualquer potência de 11. Por exemplo, na linha de ordem 8 podemos ler

1 8 28 56 70 56 28 8 1

à qual aplicando a *regra de passagem para a casa seguinte*, a partir da direita, como é habitual, temos

1. 1, 8 são algarismos, ficam: 81;
2. 28 deixa 8 e passa 2 para a casa seguinte: 56 fica 58;
3.  $56+2=58$ , deixa 8, passa 5 para a casa seguinte: 5-881;
4.  $70+5=75$ , deixa 5, passa 7 para a casa seguinte: 7-58881;

5.  $56+7=63$ , deixa 3, passa 6 para a casa seguinte: 6-358881;
6.  $28+6=34$ , deixa 4, passa 3 para a casa seguinte: 3-4358881;
7.  $8+3=11$ , deixa 1, passa 1 para a casa seguinte: 1-14358881;
8.  $1+1=2$ ,  $214358881 = 11^8$ ;

Você pode conferir, usando `python` ou `calc` como calculadora, em `calc`; `power(11,8)`; 214358881.

Repetindo estas operações, você pode *recuperar* para ler na base 10, qualquer potência de 11 que aparece nas linhas do *triângulo de Pascal*.

---

- **combinação** é um item da Análise Combinatória que é a parte elementar da *combinatória*. Confira também *arranjos*, *permutações*.

A *análise combinatória* estuda *arranjos*, *permutações*, *combinações* e os *arranjos* podem ser simples ou com repetição. Os *arranjos simples* também são chamados de *arranjos sem repetição*.

As combinações são *exatamente os subconjuntos dum conjunto dado* e é por elas que vou começar, quando a maioria dos textos começa pelos arranjos, mas não é exatamente para ser diferente e sim porque desta forma vou conseguir chegar rapidamente a uma fórmula para calcular as combinações da qual vou deduzir a fórmula dos arranjos. E você logo vai ver que é um método mais concreto.

Considere um conjunto com 3 elementos:  $A = \{1, 2, 3\}$ .

Posso agora extrair vários subconjuntos de  $A$ , e vou começar pelo conjunto vazio, todo conjunto tem o vazio como subconjunto.

- $C_3^0 = 1$  é o número de subconjuntos com zero elementos tirados de  $A$ . O vazio é único!
- $C_3^1 = 3$  é o número de subconjuntos unitários que podem ser tirados de  $A$ :  $\{1\}, \{2\}, \{3\}$ .
- $C_3^2 = 3$  é o número de subconjuntos 2-a-2 que podem ser tirados de  $A$ :  $\{1, 2\}, \{1, 3\}, \{2, 3\}$ .
- $C_3^3 = 1$  é o número de subconjuntos 3-a-3 que podem ser tirados de  $A$ :  $\{1, 2, 3\} = A$ .

Observe a mudança na linguagem, “subconjuntos 2-a-2” esta é forma de falar na *análise combinatória* para significar que estou selecionando os subconjuntos com dois elementos. Então refazendo as outras afirmações fica:

- $C_3^0 = 1$  é o número de subconjuntos 0-a-0 tirados de  $A$  que é apenas o vazio.
- $C_3^1 = 3$  é o número de subconjuntos 1-a-1 que podem ser tirados de  $A$ :  $\{1\}, \{2\}, \{3\}$

- $C_3^2 = 3$  é o número de subconjuntos 2-a-2 que podem ser tirados de  $A$ :  $\{1, 2\}, \{1, 3\}, \{2, 3\}$
- $C_3^3 = 1$  é o número de subconjuntos 3-a-3 que podem ser tirados de  $A$ :  $\{1, 2, 3\} = A$ , apenas o próprio  $A$

Deixe-me escrever esta informação resumidamente, em quatro colunas, a coluna do 0-a-0, coluna do 1-a-1, coluna do 2-a-2 e coluna do 3-a-3:

1 3 3 1

Uma observação sobre a notação,  $C_n^p$  é o número de subconjuntos com  $p$  elementos que podemos extrair dum conjunto que tenha  $n$  elementos. Neste momento ainda não conheço uma fórmula para este número, e o meu objetivo é construir esta fórmula.

Vou agora extrair os subconjuntos do conjunto  $A = \{1, 2\}$  e assim obter os números  $C_2^0, C_2^1, C_2^2$ .

- $C_2^0 = 1$  apenas o vazio como subconjunto 0-a-0 que pode ser retirado de  $A$ .
- $C_2^1 = 2$  a quantidade de subconjuntos 1-a-1 que podem ser retirados de  $A$ :  $\{1\}, \{2\}$
- $C_2^2 = 1$  a quantidade de subconjuntos 2-a-2 que podem ser retirados de  $A$ :  $\{1, 2\} = A$

Deixe-me ampliar o resumo anterior agora com estas novas informações, sempre separando por colunas a quantidade de subconjuntos p-a-p:

1 2 1  
1 3 3 1

Se agora eu fizer o mesmo com o conjunto  $A = \{1\}$  eu vou ter

- $C_1^0 = 1$  apenas o vazio como subconjunto 0-a-0 que pode ser extraído de  $A$ .
- $C_1^1 = 1$  apenas um subconjunto 1-a-1 que pode ser extraído de  $A$  que o próprio  $A$ .

Completando o resumo:

1 1  
1 2 1  
1 3 3 1

Vou agora extrair os subconjuntos do conjunto  $A = \{\}$  que é o conjunto vazio, eu vou ter

$C_0^0 = 1$  apenas o próprio vazio é subconjunto de  $A = \{\}$ . Completando o esquema:

1  
1 1  
1 2 1  
1 3 3 1

que são as quatro primeiras linhas do *triângulo de Pascal*, um *esquema algorítmico* muito antigo e muito cheio de surpresas e vou logo lhe mostrar algumas dessas surpresas.

Observe que cada linha pode ser obtida da anterior com três regras bem simples:

1. O primeiro e último número é sempre 1. Porque num conjunto com  $n$  elementos tem apenas 1 subconjunto que é vazio, e corresponde às combinações 0-a-0 dum subconjunto com  $n$  elementos,  $C_n^0 = 1$ , e as combinações n-a-n de  $n$  elementos,  $C_n^n = 1$ .

Na linha de ordem  $n$  eles vão representar a quantidade de subconjuntos 0-a-0 que o é vazio, e a quantidade de subconjuntos n-a-n que podem ser extraídos dum conjunto com  $n$  elementos que é apenas o próprio conjunto.

$$C_n^0 = 1 = C_n^n; \quad (56)$$

2. a linha de ordem zero, a primeira linha, tem apenas um elemento que sendo o primeiro e último, tem que ser 1. A segunda linha, a de ordem 1, tem dois elementos que têm que ser 1.
3. Lei de Stifel A partir da linha de ordem 2, cada novo elemento numa linha é obtido pela soma de dois elementos da linha anterior: o que lhe estiver acima, e o da esquerda. Confira o esquema acima e depois o próximo em que estou acrescentando as linhas de ordem 4 e 5 que corresponde às combinações de 4 elementos p-a-p e as combinações de 5 elementos p-a-p. Em símbolos, a *Lei de Stifel* fica:

$$C_n^p = C_{n-1}^{p-1} + C_{n-1}^p \quad (57)$$

```

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1

```

4. Um programa como equação O meu objetivo era obter uma fórmula para o cálculo de  $C_n^p$  e as propriedades listadas acima já permitem escrever um programa de computador para calcular este número para valores arbitrários de  $n, p \in \mathbf{N}$ . Eis uma função recursiva, escrita em `calc`, para calcular qualquer *número combinatório*

```

define c(n,p) {
  if (p == 0) return 1;
  else if (p == 1) return n;
  else if (p == n) return 1;
  else return c(n-1,p-1) + c(n-1,p);
}

```

```

}

define EscreveTriangulo(n) {
    local k = 0, p = 0;
    while (k <= n) {
        p = 0;
        while (p <= k) {
            printf("%f ", c(k,p));
            p++;
        }
        printf("\n");
        k++;
    }
}

```

```

EscreveTriangulo(15);
quit;

```

Se você raspar e colar o texto acima num terminal do `calc` você verá o triângulo ser calculado até a linha de ordem 15. Se você melhorar o programa para obter uma saída de dados com formatação bonita eu agradeceria que me enviasse... como está, vai ficar feio. Um programa é uma equação então parte do projeto inicial está atingido, mas vou obter uma expressão algébrica para  $C_n^p$ .

Ao final apresento de forma resumida a demonstração por *indução finita* da fórmula  $C_n^p$ .

5. As potências de 11 Como eu falei, este é um algoritmo cheio de propriedades, a segunda que lhe vou apresentar agora é que podemos identificar as potências de 11 nas linhas do triângulo de Pascal. Experimente:

$$\left\{ \begin{array}{ll} 11^0 & = 1 \\ 11^1 & = 11 \\ 11^2 & = 121 \\ 11^3 & = 1331 \\ 11^4 & = 14641 \\ 11^5 & = 161051 \end{array} \right. \quad (58)$$

e parece que falha no caso do  $11^5 = 161051!$

Mas vale também neste caso como vale em todas as potências inteiras de 11, apenas temos que aplicar a *regra da passagem do algarismo da dezena para casa seguinte* quando o algarismo for maior do que 9. Apliquei esta regra nos dois “10” da linha de ordem 5 para encontrar  $161051 = 11^5$ .

Para valores maiores  $n$  da ordem da linha, será preciso passar para frente os algarismos das dezenas, das centenas...

E aqui você tem as 15 primeiras linhas do triângulo de Pascal, na figura (fig 7), página 30, calculado com um programa em Python.



```

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
1 8 28 56 70 56 28 8 1
1 9 36 84 126 126 84 36 9 1
1 10 45 120 210 252 210 120 45 10 1
1 11 55 165 330 462 462 330 165 55 11 1
1 12 66 220 495 792 924 792 495 220 66 12 1
1 13 78 286 715 1287 1716 1716 1287 715 286 78 13 1
1 14 91 364 1001 2002 3003 3432 3003 2002 1001 364 91 14 1
1 15 105 455 1365 3003 5005 6435 6435 5005 3003 1365 455 105 15 1

```

Figura 7: 15 primeiras linhas do Triângulo de Pascal

Experimente ler

$$11^{15} = 4177248169415651; 4177248169415651 \quad (59)$$

Na expansão de  $11^{15}$ , eu obtive o valor aplicando a regra de passagem para as casas seguintes, escrevendo o resultado logo depois da igualdade, e confirmando o valor usando o comando `power(11,15)` num terminal do `calc`.

Vou aproveitar as potências de 11 para mostrar-lhe mais uma propriedade que podemos tirar do triângulo de Pascal. Na verdade vou começar mostrando-lhe porque as potências de 11 aparecem nas linhas do triângulo.

A justificativa vem de duas propriedades da aritmética, *distributividade* e *associatividade*. É que

$$11 = (10 + 1); 11^2 = (10 + 1)^2 = 10 * (10 + 1) + 1 * (10 + 1); \quad (60)$$



a *associatividade* é difícil de ver-se neste exemplo, ficaria mais visível em  $11^3 = 11^2 * 11$ .

A *distributividade* é a que permite que a multiplicação seja distribuída com os termos da soma. Finalmente entra a multiplicação por 10 que faz correr todas as casas uma unidade para a esquerda:

$$\begin{aligned} 11^3 &= 11^2 * 11 \\ 11^3 &= 11^2 * (10 + 1) \\ 11^3 &= 11^2 * 10 + 11^2 * 1 \\ 11^3 &= 1210 + 121 \\ 11^3 &= 1331 \end{aligned}$$

$$\begin{array}{r} \phantom{11^3} \phantom{=} \phantom{=} \phantom{=} \phantom{=} \\ \phantom{11^3} \phantom{=} \phantom{=} \phantom{=} \phantom{=} \\ \phantom{11^3} \phantom{=} \phantom{=} \phantom{=} \phantom{=} \\ \hline 11^3 \phantom{=} \phantom{=} \phantom{=} \phantom{=} \phantom{=} \end{array} \begin{array}{r} 1 \phantom{0} \phantom{0} \phantom{0} \\ \phantom{1} 2 \phantom{0} \phantom{0} \\ \phantom{1} \phantom{2} 1 \phantom{0} \\ \hline 1 \phantom{0} 3 \phantom{0} 3 \phantom{0} 1 \end{array}$$

$$10^3 + 3 * 10^2 + 3 * 10^1 + 10^0$$

e esta análise nos mostra que  $11^3$  é uma soma de potências de 10 multiplicadas por coeficientes retirados da linha de ordem 3 do triângulo. Isto é verdade para todas as linhas, porque qualquer linha traz as potências de 11. Este resultado se aplica diretamente para qualquer potência do formato  $(a + 1)^n$  com  $a$  em lugar de 10.

$$\left\{ \begin{array}{l} (a + 1)^n = C_n^0 a^n + \dots + C_n^p a^{n-p} + \dots + C_n^n a^{n-n} \\ (a + 1)^n = \sum_{p=0}^n C_n^p a^{n-p} \end{array} \right. \quad (61)$$

e como  $(a + 1)^n$  é a expressão usada para calcular juros acumulados, então o triângulo de Pascal serve como calculadora avançada para cálculo de juros. Se em vez de usar  $(a + 1)^n$  tivesse usado  $(a + b)^n$  eu teria obtido a expressão do Binômio de Newton a respeito de que você pode ler em outro lugar. Não se esqueça, quando for calcular juros que pode ser necessário aplicar a *regra da passagem para a casa seguinte* para obter a expressão correta quando você “traduzir”  $(a + 1)^n$  numa expressão decimal.

6. simetria Os elementos em cada linha do triângulo se distribuem simetricamente em relação ao centro da linha. Assim  $C_n^p = C_n^{n-p}$ .
7. Colunas E agora falando das colunas, todas são dadas por um polinômio cujo grau é a ordem da coluna. Assim a primeira de ordem zero é  $n^0 = 1$ , a coluna de ordem 1 é  $n$ . A justificação é simples: as diferenças de termos sucessivos em qualquer coluna é dada pela coluna anterior o que implica que o grau do polinômio que descrevem as colunas diferem de 1. A demonstração:

**Dem.**:

Aplicação direta da lei de Stifel.

**q.e.d.**

A segunda coluna é dada pela equação  $P(x) = \frac{x^2-x}{2}$  e você pode obter a equação de qualquer coluna, usando a identidade

$$P(x+1) - P(x) = Q(x); \quad (62)$$

em que  $P(x)$  é o polinômio que descreve a coluna de ordem  $n+1$ , *desconhecido*, enquanto que  $Q(x)$  é o polinômio que descreve a coluna de ordem  $n$ , *conhecido*!

Eu usei esta identidade para encontrar a expressão da coluna de ordem 2.

8. A linha de ordem  $n$  do triângulo de Pascal é a listagem dos números combinatórios:

$$C_n^0, C_n^1, \dots, C_n^n \quad (63)$$

Existe, também, a pirâmide de Pascal...

O que diferencia um arranjo numa combinação é a ordem dos elementos. Os dois conjuntos  $\{1, 2\}$  e  $\{2, 1\}$  são o mesmo conjunto. A palavra *arranjo* é usada na linguagem coloquial de forma bem parecida como a entendemos em Matemática. Um *arranjo de flores* num bouquet, altera significativamente o bouquet quando trocamos o lugar de algumas flores, com jeito se pode obter um belo bouquet de flores. Com jeito e com arte! Mas o conjunto das flores é o mesmo.

Então podemos obter os arranjos a partir dos subconjuntos, como vou fazer agora.

Voltando ao conjunto  $A = \{1, 2, 3\}$  vou agora extrair do mesmo os arranjos simples 0-a-0, 1-a-1, 2-a-2, 3-a-3.

Qualquer destes arranjos pode ser obtido diretamente do correspondente conjunto p-a-p que já obtivemos anteriormente:

- $C_3^0 = 1$  é o número de subconjuntos 0-a-0 tirados de  $A$  que apenas o vazio. Este é também o número de arranjos 0-a-0 tirados de  $A$ .
- $C_3^1 = 3$  é o número de subconjuntos 1-a-1 que podem ser tirados de  $A$ :  $\{1\}$ ,  $\{2\}$ ,  $\{3\}$ , também aqui não há alteração, são 3 os arranjos 1-a-1 tirados de  $A$  um para cada um dos 3 subconjuntos unitários.
- $C_3^2 = 3$  é o número de subconjuntos 2-a-2 que podem ser tirados de  $A$ :  $\{1, 2\}$ ,  $\{1, 3\}$ ,  $\{2, 3\}$  e aqui começa a haver alguma novidade,

$$\{(12), (13), (23), (21), (31), (32)\};$$

são os arranjos, tomados 2-a-2, que podemos fazer com os elementos de  $A$ . O que diferencia um arranjo numa combinação é a permutação dos seus elementos, então os arranjos  $(12)$ ,  $(21)$  são diferentes. Cada subconjunto produz dois arranjos, então

$$A_3^2 = 2 * C_3^2;$$

- $C_3^3 = 1$  é o número de subconjuntos 3-a-3 que podem ser tirados de  $A$ :  $\{1, 2, 3\} = A$  e os os arranjos serão:

$$\{1, 2, 3\}, \{1, 3, 2\}, \{3, 1, 2\}, \{3, 2, 1\}, \{2, 3, 1\}, \{2, 1, 3\} \quad (64)$$

$$A_3^3 = 6C_3^3; \quad (65)$$

A relação entre  $A_n^p$  e  $C_n^p$  pode ser deduzida destes exemplos. Os arranjos p-a-p são obtidos quando alteramos a ordem dos elementos de um subconjunto p-a-p tirado dum universo com  $n$  elementos. É isto que se chama “permutar” os  $p$  elementos do subconjunto. Deixe-me apenas introduzir o símbolo das permutações e logo vamos descobrir a fórmula para o seu cálculo:  $P_p$  é o número de arranjos diferentes que podemos obter dum subconjunto com  $p$  elementos:

$$A_n^p = P_p C_n^p \quad (66)$$

e nos exemplos acima temos

$$P_0 = 1; P_1 = 1; P_2 = 2; P_3 = 6; \quad (67)$$

ou seja,

- permutar os elementos do vazio corresponde apenas ao vazio,
- e permutar os elementos dum subconjunto unitário corresponde ao mesmo subconjunto unitário.

São duas abstrações necessárias à teoria.

Neste momento você já pode ver que a metodologia que estou usando é mais precisa. Nos textos que começam tratando dos arranjos ou se eliminam os arranjos 0-a-0 da *conversa* ou se fica com *uma conversa muito difícil*.

Afinal, como *arranjar* zero elementos?

Como defini os arranjos a partir dos conjuntos, simplesmente pude *arranjar* os zero elementos do vazio dizendo que não havia diferença possível entre os arranjos: tem um “único *arranjo*”.

O vazio é uma abstração, apenas é uma abstração bem conhecida, há muito que falamos de *conjunto vazio*. E nesta forma de apresentar, os arranjos 0-a-0 se encaixam naturalmente.

### O número de arranjos

Vou passar à construção da fórmula para o cálculo  $A_n^p$  e depois deduzir a fórmula de cálculo de  $C_n^p$  para a qual conhecemos apenas um programa e agora quero obter uma fórmula algébrica. Mas sem preconceitos, programas são fórmulas também.

Já chegamos a uma primeira informação:  $A_n^p = P_p C_n^p$  Vou construir a fórmula dos arranjos por *indução finita*. Acompanhe as contas, e leia sobre *indução finita* em outro lugar.

Observe que os cálculos vão ser feitos considerando os elementos dum conjunto que tem  $n$  elementos dos quais vou descobrir quantos arranjos podemos fazer p-a-p.

$$A_n^0 = 1; \text{ porque o vazio é único.} \quad (68)$$

$$A_n^1 = n; \text{ porque há } n \text{ subconjuntos unitários;} \quad (69)$$

$$A_n^2 = n(n-1); \quad (70)$$

$$A_n^p = \underbrace{n(n-1) \cdots (n-(p-1))}_{p \text{ fatores}}; \quad (71)$$

Deixe-me justificar as equações (eq. 68) (eq. 71).

1.  $A_n^1 = C_n^1$  porque os subconjuntos 1-a-1 coincidem com os arranjos.
2. Para obter  $A_n^2$  selecionamos todos os subconjuntos unitários, em número de  $n$  e queremos saber quantos arranjos podemos fazer 2-a-2. Como, já temos um elemento selecionado em cada conjunto unitário, o segundo elemento somente pode ser escolhido entre  $n-1$  elementos, donde o número:  $n(n-1)$ .
3. Para obter os arranjos 3-a-3 consideramos todos os arranjos 2-a-2 e neles vou inserir um novo elemento, mas isto pode ser feito de  $(n-2)$  diferentes maneiras para não repetir elemento, então  $A_n^3 = n(n-1)(n-2)$  e observe que há 3 fatores, que eles vão decrescendo e que o último fator é  $n-(3-1)$  o que me permite formular a *hipótese de indução*

$$A_n^p = \underbrace{n(n-1) \cdots (n-(p-1))}_{p \text{ fatores}}; \quad (72)$$

4. Para obter os arranjos (p+1)-a-(p+1) vou inserir em todos os arranjos p-a-p um novo elemento, mas isto pode ser feito de  $(n-(p-2))$  diferentes maneiras que é o número que resta de elementos diferentes para cada um dos subconjuntos (p+1)-a-(p+1) para não repetir elemento, então

$$A_n^{p+1} = \underbrace{n(n-1) \cdots (n-(p-2))}_{p+1 \text{ fatores}} \quad (73)$$

confirmando a *hipótese de indução* e demonstrando assim o teorema:

**Teorema 2 (do número)** *de arranjos simples*

*Dados  $n$  elementos diferentes, um conjunto com  $n$  elementos, é possível obter*

$$A_n^p = n(n-1) \cdots (n-(p-1)) = \frac{n!}{(n-p)!} \quad (74)$$

arranjos simples  $p$ -a- $p$ . **Dem**:

Falta demonstrar a expressão depois da igualdade uma vez que a primeira parte foi obtida pela confirmação da indução finita. Acompanhe os cálculos:

$$n(n-1) \cdots (n-(p-1)) = \frac{n(n-1) \cdots (n-(p-1)) \cdots (n-p) \cdots 1}{(n-p) \cdots 1}; \quad (75)$$

$$n(n-1) \cdots (n-(p-1)) = \frac{n!}{(n-p)!} \quad (76)$$

**q.e.d.**

Observe que no denominador se encontra o fatorial de  $n-p$  que é a permutação de  $n-p$  elementos dados. Ou seja

$$P_n = A_n^n = n \cdots (n - (n-1)) = n \cdots 1 = n! \quad (77)$$

As permutações de  $n$  elementos são os arranjos de  $n$  elementos tomados  $n$ -a- $n$ .

Resumindo as fórmulas:

$$\begin{cases} A_n^p = \frac{n!}{(n-p)!}; 0 \leq p \leq n \\ C_n^p = \frac{A_n^p}{p!} = \frac{n!}{(n-p)!p!}; 0 \leq p \leq n; \\ P_p = p!; p \geq 0; \end{cases} \quad (78)$$

Demonstração por indução finita de  $C_n^p$

A verificação inicial será feita sobre os elementos da linha de ordem 2, vou ignorar as linhas de ordem zero e um.

Nesta linha de ordem 2 há três elementos: 1,2,1 que correspondem  $C_2^0, C_2^1, C_2^2$  e ela foi obtida aplicando a *Lei de Stifel* à linha anterior se tem:

$$C_2^0 = 1 = \frac{2!}{(2-0)!0!}; C_2^1 = 2 = \frac{2!}{(2-1)!1!}; C_2^2 = 1 = \frac{2!}{(2-2)!2!} \quad (79)$$

$$C_n^p = \frac{n!}{(n-p)!p!} \quad (80)$$

$$C_{n+1}^0 = 1 = \frac{(n+1)!}{(n+1)!0!} = 1 = C_{n+1}^{n+1} = \frac{(n+1)!}{(n+1-(n+1))!(n+1)!}; \quad (81)$$

$$0 < p < n+1 \Rightarrow C_{n+1}^p = C_n^p + C_n^{p-1}; \quad (82)$$

$$C_n^p = \frac{n!}{(n-p)!p!}; \quad (83)$$

$$C_n^{p-1} = \frac{n!}{(n-p+1)!(p-1)!}; \quad (84)$$

$$\frac{n!}{(n-p)!p!} + \frac{n!}{(n-p+1)!(p-1)!} = \frac{(n-p+1)n!+pn!}{(n-p+1)!p!} = \quad (85)$$

$$= \frac{(n-p+1+p)n!}{(n-p+1)!p!} = \frac{(n+1)n!}{(n-p+1)!p!} = \frac{(n+1)!}{(n+1-p)!p!} = \quad (86)$$

$$= C_m^p; m = n+1 \quad (87)$$

- Na equação (79) apenas constatei que a fórmula vale, como estabelece o *princípio da indução finita*.
- Na equação (80), está enunciada a *hipótese de indução* num forma em que ela significa que a expressão vale para todos os valores de  $p \leq n$  para um valor genérico  $n$ ;
- Na equação (81) e seguintes, estou usando a *hipótese de indução* para calcular um número combinatório  $C_{n+1}^p$  qualquer, da linha de ordem  $n+1$ .

---

- **cometa Ison**

O caso do Ison, um belo exemplo de equações diferenciais.

Ison era um aglomerado de poeira espacial navegando em conjunto, aliás, como o é também a nossa Terra. Ison foi oficialmente declarado extinto depois de sua passagem extremamente próxima do sol ai pelo dia 28 de novembro. Extinto é uma forma de falar errada, segundo Lavoisier, nada na Natureza se cria (ou se extingue) tudo se transforma, e foi o caso do Ison, parte aderiu ao Sol e outra parte adquiriu mais energia e partiu em outra órbita, apenas esta parte virou o que se chama poeira cósmica.

Esta foto foi tirada pelo telescópio TRAPPIST da ESO, na manhã de 15 de novembro, TRAPPIST fica em La Silla e é um telescópio operado remotamente a partir de uma central situada em Liège, Bélgica.

Na figura (fig 8), página 36, você pode ver o cometa com sua cauda apon-



Figura 8: O cometa e sua cauda

tando na direção contrário do Sol.

E nesta página você pode ver mais fotografias: <http://www.space.com/19973-comet-ison.html>

A Terra é um aglomerado de “poeira espacial” apenas muito maior do que era o Ison, envolta com uma capa de gases que também estão presos pela gravidade terrestre que ajudam a manter colada a poeira de que a Terra é feita, é aquilo que nós chamamos de pressão atmosférica. Tem também a a somatória da gravidade de toda esta massa que atua sobre cada partícula e inclusive sobre a massa de gases criando a pressão atmosférica. Mas se a Terra passasse a uma distância pequena do Sol, este milhão de quilômetros a que Ison passou, toda a nossa poeira seria chupada pela gravidade do Sol como aconteceu com Ison. Na verdade, ao se aproximar demasiado de uma grande fonte de gravidade como é o Sol, os pedaços de que era formado antes o cometa, deixaram de se comportar de forma unificada passando a agir sob a influência mais forte da gravidade do Sol.

Cada uma das “partículas” de que era formado o Ison adquiriu a sua própria órbita, algumas foram ao encontro do Sol, e outras adquiriram diferentes órbitas se espalhando pelo espaço.

A passagem por “perto” de uma fonte de gravidade tão grande como o Sol é utilizada com frequência na navegação espacial, é o uso da segunda lei de Newton, uma equação diferencial. Uma passagem devidamente calculada, com acerto, pode ser utilizada para que uma nave adquira velocidade, melhor, altere sua velocidade o que pode ser uma simples mudança de direção, mas também pode ser mudança de direção com aumento de velocidade, e corrija seu rumo em uma direção desejada. A gravitação é a grande fonte de energia que existe no espaço, a força de gravidade. No caso do Ison apenas ele chegou muito perto o que fez com que suas partes perdessem a coesão se transformando num coleção solta de pedaços que partiram cada um deles em órbita própria ocasionando a declaração de que Ison ficou extinto, em 28 de Novembro. Mas esta é uma declaração errada, nada se extingue, tudo se transforma.

Aqui mais outra foto também obtida pelo TRAPPIST na figura (fig 9), página 37, você vê o fenômeno da cauda do cometa. O cometa está iluminado



Figura 9:

essencialmente pela luz solar e puxa atrás de si uma poeira cósmica que também está iluminada pela luz solar criando a “cauda do cometa”, esta é a justificção de por que a cauda do cometa “aponta” na direção contrária do Sol.

Mas eu queria agora explicar porque a “extinção”, do Ison, ou melhor, sua transformação em poeira astral, é um exemplo bonito de equações diferenciais, um tipo de equações que começa a ser estudada nos cursos de Cálculo por volta do terceiro semestre da Universidade, usualmente nos cursos de Cálculo III. É a segunda lei de Newton que se expressa sob forma da equação diferencial

$$\vec{F} = m\dot{\vec{s}}(t) = m\frac{d\vec{v}(t)}{dt} = m\vec{a}(t) \quad (88)$$

em que a última expressão na sucessão de equações acima é “massa vezes aceleração” que é comumente chamada de equação do movimento acelerado. Ao se aproximar excessivamente do sol, Ison perdeu a sua aparente coerência ficando cada partícula que o compunha entregue à ação da segunda lei de Newton. Também temos que considerar uma perturbação desta equação representada pela velocidade que o conjunto tinha relativamente ao Sol que passa a representar energia cinética, portanto uma força que se contrapõe à aceleração da gravidade solar, então, partículas com maior massa adquiriram mais velocidade e partiram em nova órbita relativamente semelhante à órbita que Ison poderia ter seguido se você formado de uma massa mais coesa que não se tivesse espatifado frente à aceleração da gravidade solar.

Uma das fotos que pode ser vista na página <http://www.space.com/19973-comet-ison.html> mostra a trajetória do Ison antes de chegar perto do Sol assim como se podem ver algumas partículas que se afastam do Sol, as mais pesadas que adquiriram mais velocidade com a gravitação solar e seguiram seus caminhos em outra direção. Abaixo você vê uma simulação feita manualmente do que pode ter acontecido com algumas partículas e é possível fazer esta simulação com boa precisão tendo-se a massa e a velocidade do conjunto relativamente ao Sol, basta aplicar-se na segunda lei de Newton e rodar um programinha com as distintas massas para fazer o gráfico das trajetórias.

Como não tenho estas informações estou aqui fazendo uma simulação artística na figura (fig 10), página 39, em que você pode ver o caminho do cometa se não houvesse se quebrado assim como as trajetórias possíveis de algumas das partículas. Algumas caindo mais rápido sobre o So, outras um pouco depois e algumas seguindo por uma nova trajetória, quase certamente diferente da trajetória original do cometa porque a trajetória original correspondia à massa total e as partículas com massa menor mas com a a energia cinética herdada do conjunto, prosseguiram em uma órbita se afastando mais do sol, numa elipse mais aberta do a que teria sido a trajetória do conjunto original. As partículas mais leves foram as que caíram contra o Sol, elas tinham menos inércia e puderam ser mais violentamente aspiradas pela gravitação solar. As mais pesadas, com mais inércia saíram mais para fora da curva elíptica que o conjunto perseguia. Há uma grande quantidade de curvas possíveis, o número corresponde à quantidade de partículas independentes em que se partiu Ison, cada uma com uma curva determinada por sua massa.

O telescópio Hubble, que também é controlado remotamente, foi apontado para o ponto em que deveria sair o cometa Ison logo depois do dia 28 de novembro não tendo podido detectar nada possivelmente porque estaria programado para observar um corpo do tamanho do que seria Ison quando na verdade o que havia era poeira cósmica formada de pedaços de tamanho muito inferior ao esperado.

Assim se acabou o cometa do ano, embora sua matéria persista, o conjunto que formou o cometa resplandecente como seria visto se não houvesse se desintegrado. Você pode ver na Wikipedia

[http://en.wikipedia.org/wiki/Comet\\_Ison](http://en.wikipedia.org/wiki/Comet_Ison) um filme da NASA (National Agency for Space Administration), não confundir com a pernóstica NSA. A



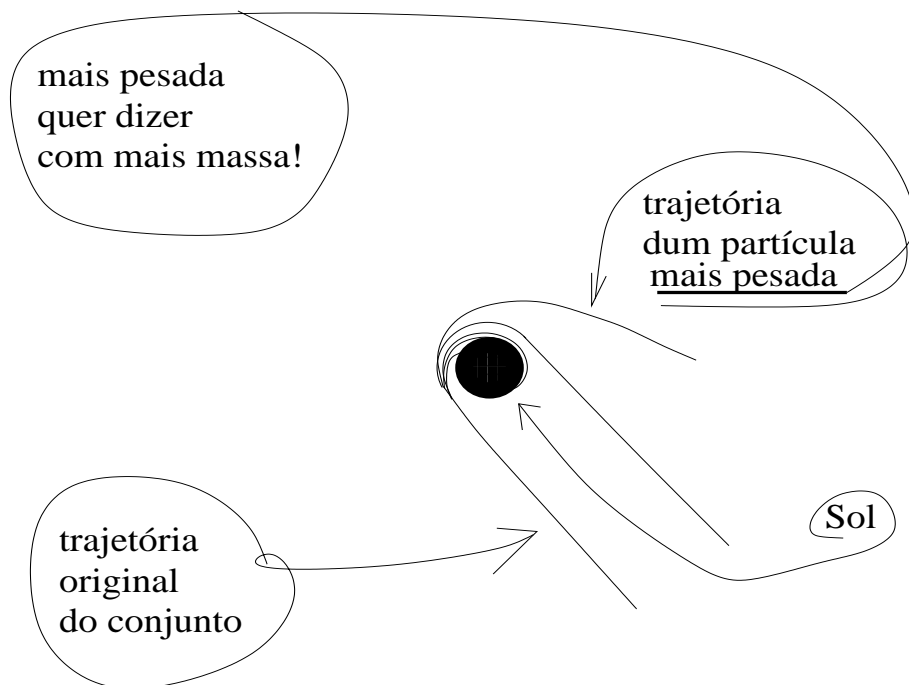


Figura 10:

NASA estabeleceu um programa para acompanhar o passeio do Ison em volta do Sol, você pode ver o que ela coletou na página <http://www.isoncampaign.org/>

- **complexidade** pode ser descrita usando como modelo a *teoria dos conjuntos*.

O tipo mais simples (menos complexo) de conjunto é o conjunto finito, aqui estou me referindo a uma classe de conjuntos. O representante da classe dos conjuntos com  $n$  elementos é

$$A = \{1, 2, \dots, n\}; n \geq 1; \quad (89)$$

Claro que logo temos uma discussão a fazer? onde fica o conjunto  $\{\}$  = nesta classificação? Vou deixar de lado esta querela, por enquanto! Neste mesmo grau de complexidade podemos encontrar uma variante  $\mathbf{P}(A)$ , conjunto das partes de  $A$ . Ele contém parte dos modelos  $A^k$ , os produtos cartesianos de  $A$ , porém sem repetições.  $A^k$  seriam os arranjos com repetições de  $n$  elementos tomados  $k$  a  $k$  e em  $\mathbf{P}(A)$  as repetições estão eliminadas. Mas podemos dizer que se trata da mesma *complexidade*. O próximo grau de complexidade seria o conjunto  $\mathbf{N}$  dos números naturais que é o conjunto infinito mais simples (menos complexo), aqui não cabe mais falar em “números de elementos” que é um conceito generalizado pela *cardinalidade*. Há diversos conjuntos com a mesma complexidade de  $\mathbf{N}$ ,  $\mathbf{Z}$ ,  $\mathbf{Q}$ ,  $\mathbf{Q}^n$ , se  $n$  for um inteiro... e se não for, como esta

expressão representa um conjunto de funções, pode haver uma complexidade mais elevada nesta expressão!

$\mathbf{P}(\mathbf{N})$  é uma nova classe de complexidade.

Aqui vem a descoberta fundamental de Cantor o operador  $\mathbf{P}(\mathbf{N})$  produz um novo grau de complexidade. A cardinalidade de  $\mathbf{P}(\mathbf{N})$  é diferente da  $card(\mathbf{N})$  com a sua hipótese de que não há complexidade intermediária entre  $card(\mathbf{N})$  e  $card(\mathbf{P}(\mathbf{N})) = card(\mathbf{R})$  a cardinalidade do conjunto dos números reais, também chamada de *cardinalidade do contínuo*. Assim

$$A, \mathbf{N}, \mathbf{P}(\mathbf{N}), \mathbf{P}(\mathbf{P}(\mathbf{N})), \dots \quad (90)$$

seria uma lista de complexidades sucessivas, em que  $A$  é um subconjunto finito de  $\mathbf{N}$ .

A hipótese de Cantor estabelece que não há cardinalidades intermediárias entre as que se obtiver com o operador  $\mathbf{P}$ , são os chamados “saltos de cardinalidade”. Em 1900 Hilbert enunciou 23 classes de problemas em sua fala no Congresso Internacional de Matemática, entre estes problemas se encontrava a hipótese de Cantor que foi estudada por K. Gödel in 1940 e finalmente Paul Cohen in 1963 completou o trabalho de Gödel com a conclusão de que a hipótese de Cantor teria que ser considerada um axioma da Teoria dos Conjuntos de Zermelo–Fraenkel que é considerada um dos fundamentos da Matemática, [12, Cantor].

- **complexo, número** é um número da forma

$$a + bi; a, b \in \mathbf{R} \quad (91)$$

Estes número surgem quando se tenta resolver uma equação do segundo grau usando a *fórmula de Bhaskara*.

$$ax^2 + bx + c = 0 \Rightarrow x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} \quad (92)$$

$$\Delta = b^2 - 4ac < 0 \Rightarrow x \notin \mathbf{R} \quad (93)$$

$$\Delta < 0 \Rightarrow \sqrt{\Delta} = \pm i \sqrt{|\Delta|} = \pm id; d = \sqrt{|\Delta|}; i = \sqrt{-1} \quad (94)$$

$$\Delta < 0 \Rightarrow x = \frac{-b \pm id}{2a} \quad (95)$$

Foi feita uma invenção:  $\sqrt{-1} = i$ . Até então, antes desta invenção, tínhamos uma regra com *uma exceção*:

$$\sqrt{ab} = \sqrt{a}\sqrt{b} \Leftrightarrow a > 0; b > 0;$$

A exceção sendo que “a regra deixava de valer se algum dos números,  $a$  ou  $b$  fosse negativo”. Agora a regra é, simplesmente,

$$\sqrt{ab} = \sqrt{a}\sqrt{b}$$

para quaisquer que sejam os os números reais. Por exemplo,

$$\sqrt{-3} = \pm \sqrt{-1}\sqrt{3} = \pm i\sqrt{3}; \sqrt{-4} = \pm \sqrt{-1}\sqrt{4} = \pm 2i;$$

Não existe mais exceção, a regra vale sempre.

Esta invenção,  $i = \sqrt{-1}$  foi mal aceita e até recentemente os números complexos eram considerados *imaginários*. . .

Na verdade, *imaginário* é o nome que se deu ao  $i = \sqrt{-1}$ , a *unidade imaginária*, desta forma o preconceito fica sendo repassado de geração em geração.

Podemos resolver a equação abaixo usando a fórmula de Baskara, mas também podemos fazê-lo diretamente:

$$x^2 + 1 = 0 \Rightarrow x^2 = -1 \Rightarrow x = \pm i \quad (96)$$

a solução é um *número imaginário puro*

Ainda um outro exemplo

$$x^2 - 6x + 10 = 0 \Rightarrow x = \frac{6 \pm \sqrt{36 - 40}}{2} = \frac{6 \pm \sqrt{-4}}{2} = \frac{6 \pm 2i}{2} = 3 \pm i \quad (97)$$

$$x \in \{3 + i; 3 - i\} \quad (98)$$

em que vemos os números  $a \pm bi$ ;  $a = 3$ ;  $b = 1$  aparecendo como soluções de uma equação do segundo grau. Você pode criar uma infinidade de exemplos deste tipo partindo do final da questão:

- escreva

$$(x - a - bi)(x - a + bi) = (x - a)^2 - (bi)^2 = (x - a)^2 + b = 0$$

- selecione:  $a, b$ ;

monte de volta uma equação do segundo grau que terá os números complexos  $a + bi$ ;  $a - bi$  como solução, para os dois números  $a, b$  que você tiver selecionado.

Com a criação dos números complexos as *equações do segundo grau* passam a ter sempre solução apesar de que, cuidadosamente, se acrescente a observação, “*raízes imaginárias*” quando  $\Delta < 0$ .

Isto mostra que a invenção do  $i$  tem sentido e que nada têm de imaginários os números complexos que, além do mais, aparecem em fórmulas de eletricidade.

E para caracterizar esta nova forma de ver as coisas vou dar um nome ao novo conjunto, é o conjunto dos “números complexos”,  $\mathbf{C}$ .

Sem querer manifestei o meu preconceito colocando *aspas* em torno da expressão, *números complexos*, traduzindo um sentimento de que não são “números” como os outros, *naturais*, *racionais* ou *reais*.

É preciso agora mostrar que podemos fazer operações aritméticas com estes números para que os possamos aceitar “como números”.

Dados  $u = a + bi$ ;  $v = c + di$  podemos somá-los usando as regras da álgebra de polinômios como faríamos com os polinômios

$$u(x) = a + bx; v(x) = c + dx$$

resultando em

$$u(i) + v(i) = (a + c) + (b + d)i;$$

Está no momento de dar um nome adequado aos componentes do “número complexo”  $u(i) = a + bi$ . Observe que somamos os “termos independentes” de cada um deles, e depois somamos os coeficientes de  $i$ . A definição é a seguinte:

**Definição 2** *Parte real e parte imaginária* Dado um número complexo  $u = a + bi = (a, b)$  designamos

- parte imaginária  $Im(u) = b \in \mathbf{R}$
- parte real  $Re(u) = a \in \mathbf{R}$

Observe que  $Re, Im$  são duas funções definidas em  $\mathbf{C}$  e tomando valores em  $\mathbf{R}$ .

De forma semelhante, mas agora usando a regra da multiplicação de polinômios, podemos efetuar:

$$(a + bi)(c + di) = ac + adi + bci + bdi^2 = \quad (99)$$

$$(a + bi)(c + di) = ac + adi + bci - bd = ac - bd + (ad + bc)i \quad (100)$$

que você pode ver, esquematicamente, na figura figura (fig 11), página 42,

## Multiplicação de números complexos

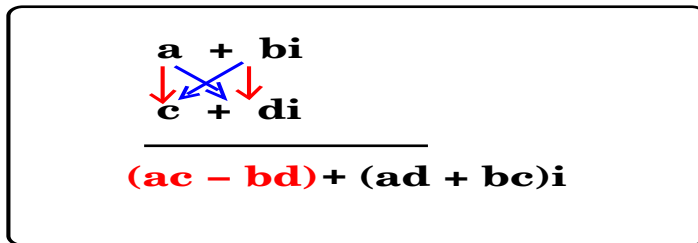


Figura 11: produto  $(a + bi)(c + di)$

O interessante é que podemos fazer *interpretação geométrica* dos números complexos mostrando que eles nada tem de imaginário e, muito pelo contrário, até são geométricos.

Os números complexos se infiltraram em nosso sistema cultural com duas apresentações:

$$\text{expressão algébrica } \mathbf{C} \ni a + bi \equiv (a, b) \in \mathbf{R}^2 \text{ entidade geométrica.} \quad (101)$$

eles podem ser um número,  $u = a + bi$  ou um ponto do plano  $(a, b) \in \mathbf{R}^2$ .

A última parte na equação (eq. 101),  $(a, b) \in \mathbf{R}^2$ , é uma *representação geométrica* para os números complexos, uma vez que estamos dizendo que existe um ponto do plano,

$$(a, b) \in \mathbf{R}^2 \quad (102)$$

que é equivalente ao número complexo

$$a + bi \in \mathbf{C}. \quad (103)$$

A descoberta da *representação geométrica* para os números complexos, representa um *salto qualitativo*. Como eles têm uma *representação geométrica*, não podem ser tão estranhos como no começo pareciam. Observe a figura (fig. 12), página 43, nela há alguns números complexos representados no plano.

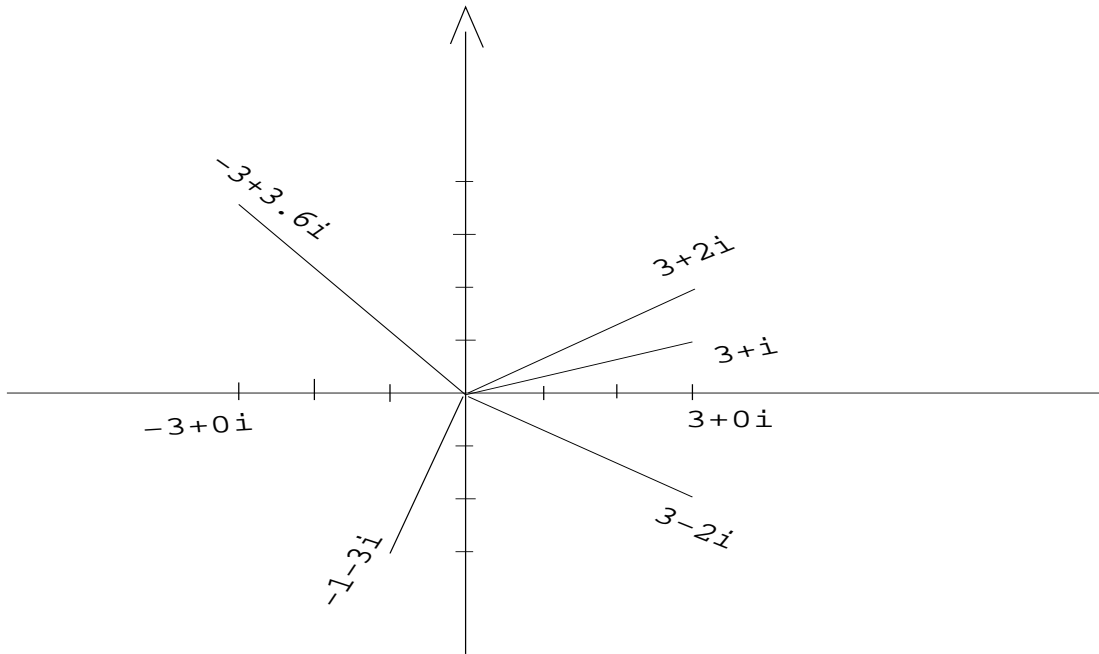


Figura 12: Representação geométrica dos complexos

Vou avançar mais a fundo na representação geométrica dos números complexos para *descobrir a fórmula de Euler* e mostrar outro método para calcular o produto de números complexos que é mais simples do que a definição apresentada acima com a interpretação polinomial.

Mais simples e mais computacional muito fácil de ser enfiada num programa de computador para construir uma calculadora para números complexos que também entenderia os números naturais, racionais e reais. Uma grande unificação dos números.

Você deve ter achado estranho que para a *adição* apresentei uma definição formal não o fazendo para o *produto de números complexos*, a razão disto é que, avançando na interpretação geométrica mais um pouco, vou poder apresentar uma fórmula para o produto muito simples que vai merecer o destaque de uma definição.

O primeiro passo nesta direção é que os números tem módulo, como os reais,

apenas muito mais significativo. Como é um ponto do plano, o módulo de um número complexo sai direto como aplicação do *teorema de Pitágoras*:

Na (eq. 101) você viu a equivalência entre a forma algébrica e a geométrica dum número complexo.

$$\mathbf{C} \ni v = c + di \equiv (c, d) \in \mathbf{R}^2, \quad (104)$$

o par  $(c, d)$  era um ponto do plano e assim estamos *representando* um número complexo com uma entidade geométrica, um ponto.

Desta forma os números complexos trouxeram, para o reino dos números, os conceitos da geometria: ângulo, módulo, direção e sentido, e a Física, desde cedo, lançou mão deles, com muito sucesso, por exemplo, na eletricidade.

A figura (fig. 13), página 44 descreve vários dos aspectos geométricos dos números complexos.

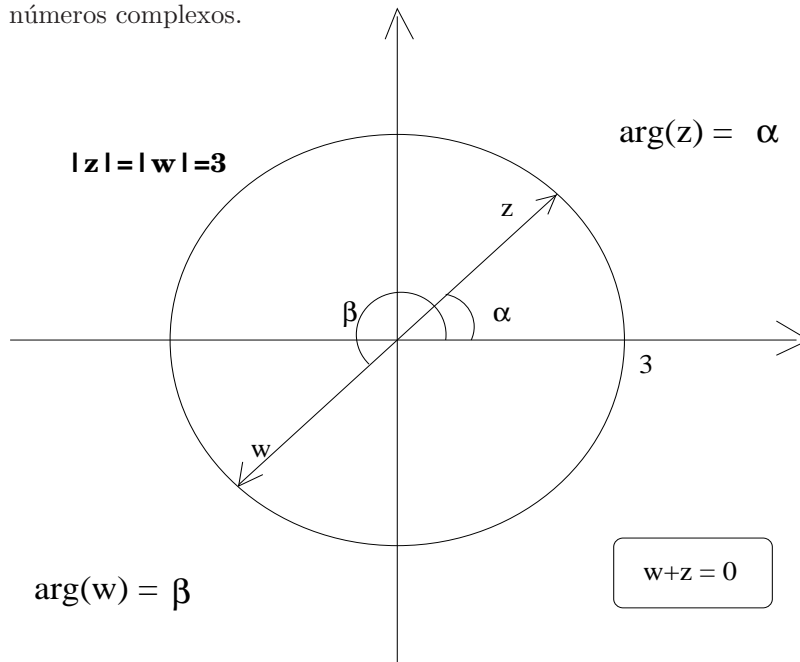


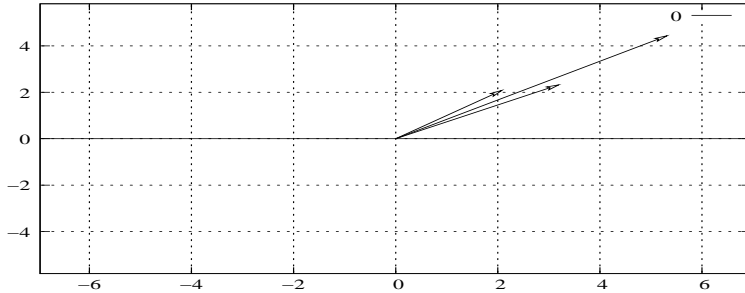
Figura 13:

- Os dois números complexos,  $z, w$  têm mesmo módulo, no plano complexo isto significa que eles se encontram num mesmo círculo de raio 3, na figura (fig. 13). Para os números reais (ou racionais, ou inteiros) isto se resume à troca de sinais. Os números complexos oferecem mais opções na expressão  $|z| = 3 \dots$
- Os dois números complexos,  $z, w$ , tem sinais contrários e como nos reais, um é o inverso aditivo do outro. Nos complexos isto significa estarem diametralmente opostos.

- Um número complexo tem um ângulo, relativamente ao eixo  $OX$ , na figura (fig. 13), o ângulo de  $z$  é  $\alpha$  e o ângulo de  $w$  é  $\beta$ . Não chamamos “ângulo”, a palavra que usamos é *argumento* e a notação é a que aparece na figura (fig. 13),  $\arg(z) = \alpha, \arg(w) = \beta$
- Se trocarmos de círculo, suponha que na figura (fig. 13) tenhamos um círculo de raio 1, faça um esforço de abstração, suponha que está vendo o círculo trigonométrico, então:

$$z = (\cos(\alpha), \sin(\alpha)); w = (\cos(\beta), \sin(\beta)); \beta = \alpha + \frac{\pi}{2} \quad (105)$$

Na figura (fig 14), página 45, você pode ver o cálculo feito com uma



$$\mathbf{u} = (3 * \cos(\pi/4), 3 * \sin(\pi/4))$$

$$\mathbf{v} = (4 * \cos(\pi/5), 4 * \sin(\pi/5))$$

$$\text{A soma de } \mathbf{u} = 2.12132 + 2.12132 \mathbf{i}$$

$$\text{com } \mathbf{v} = 3.23606 + 2.35114 \mathbf{i}$$

$$\text{é } 5.35738 + 4.47246 \mathbf{i}$$

Figura 14: A regra do paralelograma em ação

*calculadora* escrita em C++ com saída de dados para *gnuplot* que é possível editar para incluir num texto como foi feito aqui.

Retornando à equação (eq. 105) deixe-me convidá-la para mais um exercícios de abstração, esqueça-se que  $\beta = \alpha + \frac{\pi}{2}$ . Considere dois *argumentos* quaisquer e deixe-me escrever esta equação assim:

$$z = (\cos(\alpha), \sin(\alpha)); w = (\cos(\beta), \sin(\beta)); \quad (106)$$

$$z \cos(\alpha) + i \sin(\alpha) = e(i\alpha); w = \cos(\beta) + i \sin(\beta) = e(i\beta); \quad (107)$$

em que estou exercendo o meu direito de entender um número complexo ora como um número, ora como um ponto do plano. Estou também incluindo uma nova notação:  $z = e(i\alpha); w = e(i\beta)$ .

Se eu multiplicar os dois números:

$$zw = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) + i(\cos(\alpha)\sin(\beta) + \cos(\beta)\sin(\alpha)) \quad (108)$$

$$zw = \cos(\alpha + \beta) + i\sin(\alpha + \beta) = e^{i(\alpha + \beta)} \quad (109)$$

de acordo com as fórmulas de soma de arcos do cosseno e do seno.

Euler viu isto e identificou com a exponencial escrevendo sua famosa fórmula:

$$e^{i\alpha} = \cos(\alpha) + i\sin(\alpha); \quad (110)$$

$$e^{i\beta} = \cos(\beta) + i\sin(\beta); \quad (111)$$

$$e^{i(\alpha+\beta)} = e^{i\alpha}e^{i\beta} = \cos(\alpha + \beta) + i\sin(\alpha + \beta) \quad (112)$$

que permitiu-lhe escrever a famosa fórmula que é também identificada como fórmula de Euler:

$$e^{i\pi} + 1 = 0 \quad (113)$$

envolvendo os números,  $e, i, \pi, 0, 1$  numa única fórmula, que é realmente uma obra de arte.

Aqui encontramos uma aplicação prática dos números complexos, na memorização ou determinação das fórmulas de somas de arcos. Se você quiser lembrar-se das fórmulas  $\cos(\alpha + \beta), \sin(\alpha + \beta), \tan(\alpha + \beta)$ , use o produto de  $e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)}$ .

Se multiplicarmos por um número positivo  $\rho$  um número complexo que se encontre no círculo trigonométrico podemos assim obter qualquer número outro número que se encontre no plano complexo, basta que ambas tenham o mesmo argumento. Isto dá nascimento a expressão chamada polar de um número complexo:

$$u = \rho e^{i\alpha} = a + bi; a = \rho \cos(\alpha); b = \rho \sin(\alpha); \quad (114)$$

$$\rho = \|u\| = \sqrt{a^2 + b^2}; \cos(\alpha) = \frac{a}{\rho}; \sin(\alpha) = \frac{b}{\rho} \quad (115)$$

E posso agora escrever a definição do produto de números complexos:

**Definição 3 (Produto de números complexos)** *Produto de números complexos*

Dados dois números complexos  $u = a + bi = \|u\|e^{i\alpha}; v = c + di = \|v\|e^{i\beta}$ ; então  $uv = \|u\|\|v\|e^{i(\alpha+\beta)}$  O produto de  $u$  por  $v$  é obtido com o produto dos módulos e a soma dos argumentos.

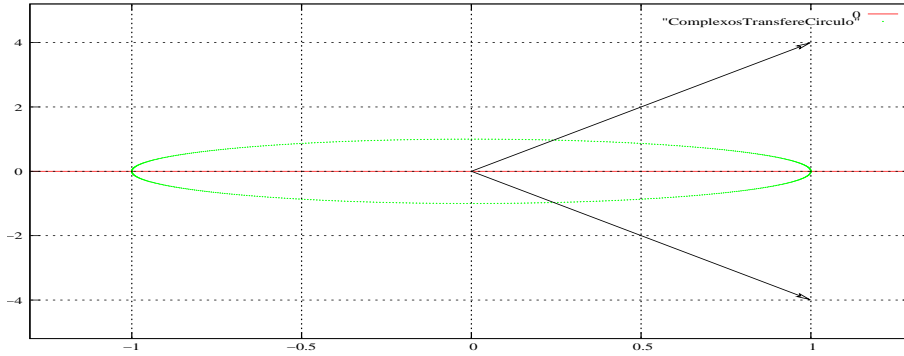
que pode ser expresso numa frase simples, como fizemos para a adição,

Há livros com centenas de páginas sobre números complexos e um dicionário não pode concorrer com esta extensão de informação, portanto aqui falta muita coisa que pode ser dita sobre estes números. Mas seria uma grande falta não mencionar uma importante operação com números complexos que é o *conjugado*:

**Definição 4 (Conjugado de  $z$ )** *Conjugado de  $z$*

Se  $z = a + bi = \rho e^{i\alpha}$  então  $\bar{z} = a - bi = \rho e^{-i\alpha}$





### O conjugado de $(1 + 4i)$ é $(1 - 4i)$

Figura 15: Um número complexo e seu conjugado

Na figura (fig 15), página 47, você encontra o círculo unitário e um número complexo de módulo maior do que 1 assim como seu conjugado.

A importância do conjugado se vê neste cálculo:

$$z\bar{z} = \rho e^{i\alpha} \rho e^{-i\alpha} = \rho^2 \quad (116)$$

$$\|z\| = \sqrt{z\bar{z}} \quad (117)$$

o produto de  $z$  pelo seu conjugado é o quadrado do módulo de  $z$  permitindo uma fórmula prática para módulo de  $z$ .

$$\|z\| = \sqrt{z\bar{z}}$$

Para finalizar, como todo número real diferente de zero, todo número complexo diferente de zero tem um inverso multiplicativo, os cálculos seguintes mostram como obter a fórmula para o inverso de  $z$ .

$$z = a + bi = \rho e^{i\alpha}; w = c + di = \lambda e^{i\beta}; \quad (118)$$

$$1 = e^{i0}; \quad (119)$$

$$zw = \rho e^{i\alpha} \lambda e^{i\beta} = 1 \Rightarrow \lambda = \frac{1}{\rho}; \beta = -\alpha \quad (120)$$

$$w = \frac{1}{\rho} e^{-i\alpha} \quad (121)$$

$$\frac{1}{z} = \frac{e^{-i\alpha}}{\rho} = \frac{\rho e^{-i\alpha}}{\rho^2} = \frac{\bar{z}}{\|z\|^2} \quad (122)$$

Dado  $z = a + bi$  posso calcular  $\rho e^{i\alpha}$  e conseqüentemente também  $-\alpha, \frac{1}{\rho}$  com o que escrevo o inverso

$$\frac{1}{z} = \frac{\bar{z}}{\|z\|^2}$$

Para os números reais diferentes de zero é verdade que se  $\|z\| < 1$  então  $\|\frac{1}{z}\| > 1$  e facilmente deduzimos da fórmula do inverso que esta relação continua

verdadeira para números complexos, apenas agora a frase fica mais imponente: “se um número complexo  $z$ , diferente de zero, estiver dentro do círculo trigonométrico então o seu inverso  $\frac{1}{z}$  estará fora do círculo trigonométrico”.

**condução** É um dos três métodos de transmissão de calor: condução, convecção e radiação.

A linguagem da termodinâmica é imprecisa, onde menciona *superfície*, que seria uma variedade de dimensão dois, na verdade quer se referir a uma variedade de dimensão 3, cuja espessura seria “despresível” e que serve de invólucro para um fluido que se encontra separado de outro pela “*superfície*”. Aqui consideramos então dois fluidos com temperaturas diferentes e o calor que irá se transmitir através da *superfície* para o outro fluido atendendo à lei do equilíbrio.

Um exemplo concreto de transmissão de calor em que existe uma “*superfície*” separando dois meios, seria um computador e sua caixa. A caixa seria a “*superfície*” que encerra componentes eletrônicos que se aquecem, fonte de calor, e este calor deve ser transmitido para fora da caixa onde se presume que haja menos calor. O manual de qualquer aparelho indica qual é a temperatura ambiente na qual o aparelho deve funcionar, se o ambiente estiver mais quente do que o desejado, a transmissão do calor produzido pelos componentes não se dará na velocidade adequada e haverá superaquecimento.

O exemplo anterior pode ser considerado “muito complexo” porque os componentes internos do computador novamente representam uma situação semelhante a do computador. Cada componente é uma “pastilha” em que o fluxo da corrente produz aquecimento e este deve ser transmitido para fora da pastilha senão a mesma se danifica.

Enfim, para simplificar vou me manter no exemplo do computador como se fosse uma unidade homogênea e isto já é uma *aproximação*.

Na *condução* se considera o processo de transmissão do calor quando dois meios se encontram em contacto direto um dos quais representa uma fonte de calor (está mais quente). Na condução as moléculas do meio mais quente cedem movimento para as moléculas do meio mais frio por impacto elástico.

---

- **cone** é uma superfície gerada por uma reta chamada *geratriz* que roda no espaço tridimensional em torno de outra reta a ela oblíqua e chamada de eixo do cone. Confira a figura (fig 16), página 49.

---

- **congruência** é um tipo de *relação de equivalência* entre os inteiros positivos, pelo resto que eles deixam na divisão por outro número inteiro. Fixe o número inteiro  $n$  e considere dois outros números inteiros,  $p, q$ . Dizemos que  $p$  é congruente a  $q$  se eles deixarem o mesmo resto na divisão por  $n$ . **Notação:**  $p \equiv q \pmod{n}$ . Como a quantidade de restos possíveis é  $n$  então esta *relação de equivalência* determina  $n$  classes no conjunto dos números inteiros positivos. As duas operações, soma e produto, podem ser definidas sobre estas classes criando uma aritmética semelhante a dos números inteiros sobre este conjunto finito.

Por exemplo, o caso  $n = 2$  corresponde aos números binários da computação, a aritmética dos computadores. Quando  $n$  for um número primo esta aritmética

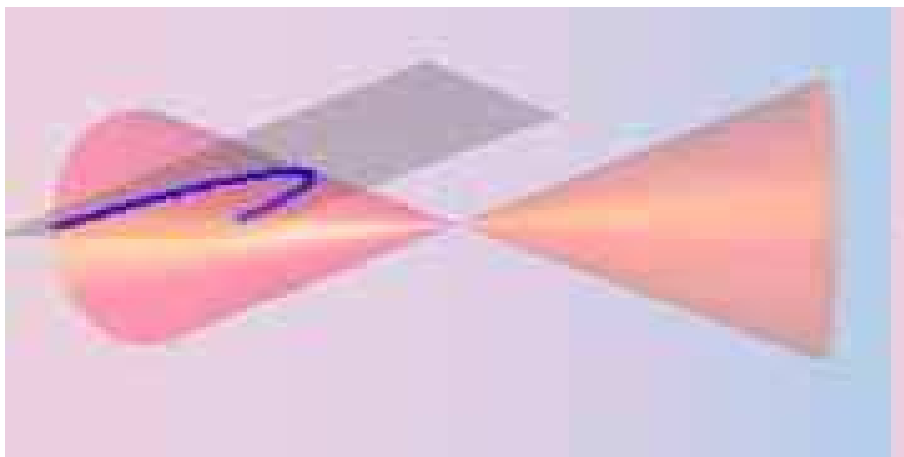


Figura 16: Cone de duas folhas gerado pela geratriz

é semelhante a dos números reais tendo como base um conjunto finito, que é o caso de  $n = 2$ , os números binários da computação. A estrutura aritmética facilmente se estende para *enúplas* de elementos da classe o que nos permite trabalhar com “números” congruentes com quantidade de dígitos arbitrariamente grandes (ainda um conjunto finito) como é a aritmética dos computadores.

Na teoria das congruências, os restos são chamados *resíduos*, e o conjunto das classes na congruência *mod n* de conjunto dos resíduos *mod n*.

Como exemplo confira a tabela de adição módulo 13

+	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0	1	2	3	4	5	6	7	8	9	10	11	12
1	1	2	3	4	5	6	7	8	9	10	11	12	0
2	2	3	4	5	6	7	8	9	10	11	12	0	1
3	3	4	5	6	7	8	9	10	11	12	0	1	2
4	4	5	6	7	8	9	10	11	12	0	1	2	3
5	5	6	7	8	9	10	11	12	0	1	2	3	4
6	6	7	8	9	10	11	12	0	1	2	3	4	5
7	7	8	9	10	11	12	0	1	2	3	4	5	6
8	8	9	10	11	12	0	1	2	3	4	5	6	7
9	9	10	11	12	0	1	2	3	4	5	6	7	8
10	10	11	12	0	1	2	3	4	5	6	7	8	9
11	11	12	0	1	2	3	4	5	6	7	8	9	10
12	12	0	1	2	3	4	5	6	7	8	9	10	11

que junto com a tabela da multiplicação módulo 13 se define o corpo finito módulo 13

*	1	2	3	4	5	6	7	8	9	10	11	12
1	1	2	3	4	5	6	7	8	9	10	11	12
2	2	4	6	8	10	12	1	3	5	7	9	11
3	3	6	9	12	2	5	8	11	1	4	7	10
4	4	8	12	3	7	11	2	6	10	1	5	9
5	5	10	2	7	12	4	9	1	6	11	3	8
6	6	12	5	11	4	10	3	9	2	8	1	7
7	7	1	8	2	9	3	10	4	11	5	12	6
8	8	3	11	6	1	9	4	12	7	2	10	5
9	9	5	1	10	6	2	11	7	3	12	8	4
10	10	7	4	1	11	8	5	2	12	9	6	3
11	11	9	7	5	3	1	12	10	8	6	4	2
12	12	11	10	9	8	7	6	5	4	3	2	1

$$(\mathbf{Z}_{13}, +, \cdot) \quad (123)$$

é um corpo comutativo finito.

- **conjectura** É uma afirmação que se considera verdadeira mas da qual não se conseguiu ainda estabelecer uma prova. Nos últimos 20 anos duas conjecturas importantes foram provadas, em Matemática, o *último teorema de Fermat*, por Wiles e outros, e a *conjectura de Poincaré*, por Perelman. Um pouco antes, na década de 60 foi encerrada a discussão sobre a *conjectura de Cantor* sobre a cardinalidade, confira *cardinalidade*.

- **conjectura de Poincaré** Não é mais uma conjectura, embora ainda seja conhecida assim, agora é um teorema demonstrado pelo russo Grigori Perelman, em 2002, que por tal ganhou a medalha Fields (mas a recusou). É extremamente difícil enunciar este teorema e mais ainda entender a sua demonstração, um livro publicado pela American Mathematical Society, com 520 páginas é a metade do projeto para explicar a teoria que conduz à demonstração (há outro livro que o completa com igual número de páginas), mas intuitivamente pode ser expressa dizendo-se que, se um elástico muito tenso, preso em volta de uma esfera, for deslizado para fora da mesma, se vai reduzir a um ponto, e que isto vale em qualquer esfera de qualquer dimensão. Para dimensões maiores de que 3 a conjectura já havia sido provada, mas não para a comum e corrente dimensão 3. Na linguagem da topologia isto se expressa dizendo-se que uma esfera é *homeotópica* a um ponto: pode ser deformada continuamente para se transformar um ponto. Uma busca na Internet com a palavra chave *Poincaré* vai levá-l@ a uma quantidade grande publicações, inclusive filmes, descrevendo esta conjectura.

- **conjunto** é a mais simples das estruturas de que trata a Matemática o que não significa que seja um conceito *simples*. Pelo contrário, exatamente por ser a mais simples das estruturas, o conceito de conjunto invade praticamente todos os outros conceitos da Matemática e faz parte da própria fundamentação desta

área do conhecimento. Confira por exemplo a *conjectura de Cantor* para ter uma visão das dificuldades que temos na *teoria dos conjuntos*. Mas vou adotar um ponto de vista intuitivo, pelo menos inicialmente.

Na década de 60 se iniciou uma renovação de linguagem em matemática colocando o conceito de conjunto como módulo central de toda a construção matemática.

A razão bem simples para isto se encontra nos seguintes fatos:

1. As operações fundamentais com conjuntos servem de *modelo concreto* para as operações fundamentais da lógica. Em suma, estudar Teoria dos Conjuntos equivale a estudar uma *realização do modelo da lógica formal*. Esta afirmação esta intimamente ligada com a *conjectura de Cantor*, leia a respeito.
2. Todas as estruturas matemáticas tem como objeto inicial uma família de conjuntos à qual se associam relações típicas da estrutura. Existem algumas exceções a esta regra, *teoria dos grafos* por exemplo, mas se tratam de autênticas exceções confirmando a regra geral . . .

Quer dizer que, estudando conjuntos se desenvolve a ferramenta básica para produzir matemática, a *lógica formal*, e também se produzem os blocos básicos desta construção.

Mas é difícil a construção desta fundamentação, e você deve observar que a redação contém defeitos lógicos, alguns dos quais eu vou registrar logo no início, e se devem ao fato de que estou tratando dum conceito que se encontra mesmo na fronteira da Matemática, então fica fácil sair da fronteira . . . fique alerta!

Existe um conceito que separa a teoria dos conjuntos entre *avançada* e *elementar*: conjunto finito. Vou deixar os conjuntos infinitos para um segundo momento, inclusive a definição do que seja *conjunto infinito*

A grande dificuldade de se iniciar qualquer *conversação* ou explanação teórica reside na *definição das ideias básicas, nas convenções iniciais* que vão servir de alicerce para o resto da construção. No início do século 20 este sentimento se concretizou vindo das dificuldades sentidas pelos nossos predecessores no século 19 e se criou o conceito de *noções básicas* que, junto com os *postulados* formariam, o background da teoria e seria aceita sem discussão, a menos que outra teoria seja desejada.

*Conjunto* é, para a Teoria dos Conjuntos, esta *noção primeira*. Os que nos precederam no início do século 20 e escreveram sobre esta teoria, ficaram circulando entre palavras como *agregado, lista* ou *conjunto*, tentando com uma, justificar a outra. Depois de algum tempo a frase “conjunto é uma ideia básica, que não iremos definir”, começou a prevalecer nos textos.

Não vou definir *conjunto* como ninguém definiu para você as primeiras palavras da língua que você fala. Diziam-lhe, no começo, que um determinado objeto era uma *cadeira* e que outro era uma *mesa* sem lhe apresentar nenhuma lógica porque uma cadeira não seria uma mesa, ou vice-versa. Somente depois, quando você já havia adquirido algum vocabulário básico é que lhe foi dado o direito de fazer perguntas. Para não agir de forma tão autoritária, vou dar alguns

exemplos de conjuntos, escrever algumas frases iniciais de forma semelhante ao modo como você aprendeu a falar...

Escrevo:

$\{a, e, i, o, u\}$  é um conjunto,

“a” é um elemento deste conjunto,

$e, i, o, u$  também o são.

Uma das forma de caracterizar o conceito de conjunto é o símbolo  $\{\}$  e logo aqui uma *complicação*, o conjunto vazio. Teria mesmo sentido em falarmos de conjunto vazio? Um exemplo bem concreto mostra que sim. Na figura (fig 17), página 52, você dois subconjuntos<sup>3</sup> do plano, duas retas  $r, s$  que se “cortam” no

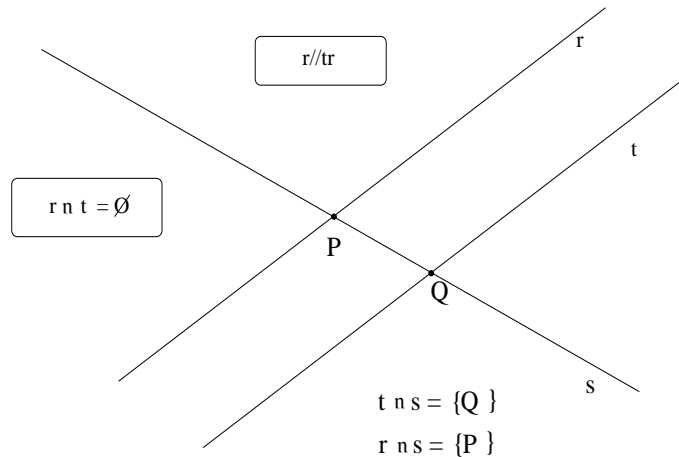


Figura 17: Interseção de dois conjuntos do plano

ponto  $P$ . Deixe-me ler a figura (fig 17) para incluir sob forma de texto o seu conteúdo. Nela há duas retas, que são dois subconjuntos do plano, e o plano é um conjunto de pontos. Os pontos são os objetos mais elementares da geometria e os objetos geométricos são conjuntos de pontos, como as três que aparecem na figura (fig 17). A reta  $t$  corta também a reta  $s$  determinando o ponto  $Q$ . Com a linguagem da teoria dos conjuntos eu vou dizer:

- a interseção dos conjuntos  $r$  e  $s$  é o conjunto  $\{Q\}$ .
- a interseção dos conjuntos  $t$  e  $s$  é o conjunto  $\{P\}$ .
- a interseção dos conjuntos  $t$  e  $r$  é o conjunto  $\{\}$  o que caracteriza que estas duas são paralelas, como você pode encontrar em alguns livros de geometria mais antigos “duas retas paralelas se encontram no infinito” para traduzir a impossibilidade de que duas retas paralelas tenham por interseção algum ponto do plano.

<sup>3</sup>E o que seria um subconjunto? deixe-me seguir sem definir, mas a particular “sub” indica algo que é parte de uma outra coisa.

Com estes exemplos geométricos encontramos alguns conceitos que poderiam ser considerados absurdos: *conjunto vazio*, que é o conjunto que caracteriza os pontos comuns de duas retas paralelas, *conjunto unitário*, um conjunto que tem um único elemento. Nós precisamos do conjunto vazio e de conjuntos unitários.

Vou voltar mais a frente a esta figura quando falar de conjunto infinito. . .

Se deixarmos a geometria e cairmos na Idade da Pedra, vamos encontrar os pastores que possivelmente deixavam as ovelhas saírem do cercado associando uma pedra a cada ovelha para no final do dia poderem conferir se todas haviam voltado. Então usavam dois conjuntos, o conjunto das ovelhas e um conjunto de pedras e até mesmo o conceito de função entre estes dois conjuntos. É certo que nesta época não pensavam em conjuntos e possivelmente nem mesmo tinham claro o que seriam números para contar as ovelhas. Mas se avançarmos um pouco mais no tempo um certo pastor poderia ter o conjunto

$$O = \{1, 2, 3, 4, 5, 6, 7, 8, 9, \dots, 30\} \quad (124)$$

para representar as sua posse de 30 ovelhas e tinha avançado na teoria porque agora bastava-lhe “contar” as ovelhas para saber se todas estavam dentro do cercado.

Novamente aqui incluí um novo conceito representado pelas reticências indicando que há uma lei para descrever o conjunto  $O$ , das ovelhas. E se aqui observarmos em detalhe este novo exemplo, como fiz a figura (fig 17), posso dizer

- Caracterizei que tinha um conjunto com os símbolos  $\{, \}$ .
- Associei ao conjunto  $\{1, 2, 3, 4, 5, 6, 7, 8, 9, \dots, 30\}$  uma variável  $O$  para denominá-lo.
- Posso dizer que 1 pertence ao conjunto  $O$  ou expressar isto com um símbolo,  $1 \in O$ .
- A sentença  $31 \in O$  é falsa.
- Posso selecionar alguns dos elementos de  $O$  para produzir um outro conjunto, agora um subconjunto de  $O$ ,  $P = \{1, 2, 3, 4, 5\}$ , e expressar isto com um novo símbolo  $P \subset O$ .
- Na medida em que a linguagem cresce, com a construção de novos conceitos e símbolos que os representam, precisamos de regras de sintaxe que regem as relações entre os conceitos. Então é sintaticamente errado escrever  $1 \subset O$  e o correto é dizer
  - $\{1\} \subset O$ , caracterizando um *conjunto* e um seu *subconjunto*.
  - $1 \in O$ , caracterizando um *conjunto* e um seu *elemento*.

---

- **consequente** numa implicação  $A \implies B$   $B$  é o consequente e  $A$  é o antecedente.

- **Continuidade** é uma *propriedade* de uma *classe* de funções. Se diz que “*f* é uma função contínua se o seu gráfico puder ser obtido sem que você tire o lápis do papel”. Mais a frente você vai encontrar um exemplo mostrando que este critério do lápis para testar continuidade, pode ser inútil.

A *continuidade* é um conceito do Cálculo Diferencial e Integral, da Topologia que é uma teoria que foi desenvolvida em torno deste conceito e da Análise Matemática, que é a grande área da Matemática à qual estes conceitos pertencem.

Continuidade é um conceito muito intuitivo, como está expresso na frase inicial, e conseqüentemente muito difícil de ser formalizado. Está muito ligado ao conceito de *limite* do qual Richard Courant dizia, “*é o vestibulo da Matemática Superior*”...

Esta forma intuitiva de continuidade, contida na frase inicial sobre “lápis e gráfico de funções”, aparece na obra de Gottfried Leibniz numa frase que ficou famosa, “*Natura non facit saltum*”, a *Natureza não dá saltos*, contraditada por Darwin ao estabelecer que, sim, a *Natureza dá saltos, porém pequenos*.

Uma das primeiras definições formais de continuidade, que se deve à Bolzano, tem a seguinte formulação

$$(\forall \epsilon) (\exists \delta) (|x - a| < \delta \implies |f(x) - f(a)| < \epsilon) \quad (125)$$

significando que *f* é contínua no ponto  $x = a$ . A figura (fig 18), página 54, é

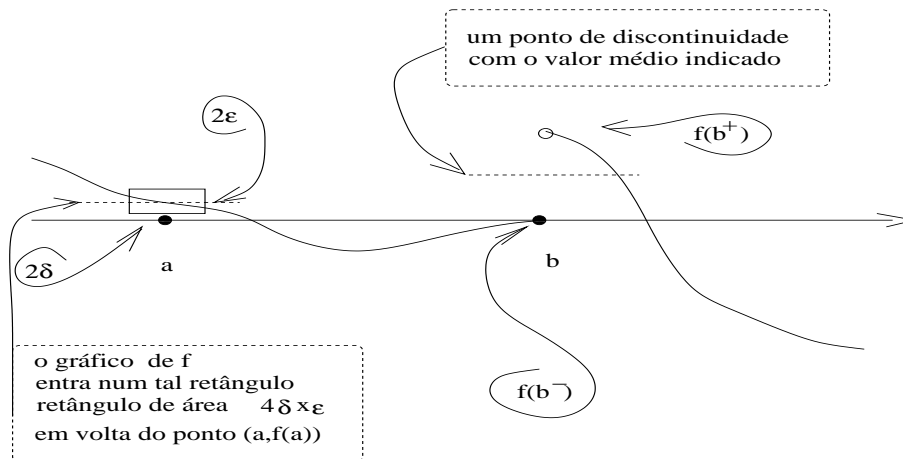


Figura 18: Interpretação geométrica da continuidade

uma representação geométrica da equação (eq. 125). Nela você vê um retângulo de área  $4\delta\epsilon$  e qualquer que seja a altura  $2\epsilon$  existe uma base  $2\delta$  de modo que o gráfico entra num tal retângulo, em volta do ponto  $(a, f(a))$  onde *f* é contínua.

Isto deixa de ser verdadeiro em volta do ponto  $(b, f(b))$  que pode até mesmo não estar definido. Dizemos que o ponto  $x = b$  é um ponto de salto da função *f*, ou que  $x = b$  é um ponto de descontinuidade de *f*. Algumas vezes é possível



calcular-se o valor médio num ponto de descontinuidade de  $f$  como está indicado na figura (fig 18) e neste caso esse valor seria o valor mais conveniente para  $f$  e é possível então verificar que “*existe um  $\epsilon$  para o qual não é possível encontrar  $\delta$  para satisfazer a (eq. 125), relativamente ao ponto  $x = b$ .*”

$x = b$  é um ponto de salto de  $f$ .

Num ponto de *salto finito* é possível determinar o valor à esquerda,  $f(b^-)$  e o valor à direita,  $f(b^+)$  que aparecem indicados no gráfico na figura (fig 18) com duas bolinhas, uma preenchida com preto, o valor à esquerda,  $f(b^-)$ , e a outra na cor do papel, o valor à direita,  $f(b^+)$ . O valor médio é

$$\frac{f(b^-) + f(b^+)}{2} \quad (126)$$

se existir, porque o salto pode ser *infinito*.

A definição contida na equação (eq. 125) é *definição local* de continuidade. Nesta definição se diz que  $f$  é contínua no ponto  $x = a$ . Existe uma definição de continuidade global, em que dizemos apenas,  $f$  é *contínua*, o que significa que é contínua em todos os pontos do domínio de definição.

Usando a notação de espaços métricos posso traduzir a equação (eq. 125) assim

$$(\forall \epsilon) (\exists \delta) d(x, a) < \delta \implies d(f(x), f(a)) < \epsilon \quad (127)$$

e como  $d(x, a) < \delta$  significa que  $x$  pertence a uma *bola aberta* centrada em  $\underline{a}$  e com raio  $\underline{\delta}$ ,

$$x \in \mathcal{B}(a, \delta) \implies f(x) \in \mathcal{B}(a, \epsilon) \quad (128)$$

se pode chegar assim à formulação de continuidade global da topologia:

$$(\forall \mathcal{O} \text{ aberto de } Y) (\exists \mathcal{B} \text{ aberto de } X) f(\mathcal{B}) \subset \mathcal{O} \quad (129)$$

Mas esta não é definição *usual* de continuidade da topologia e sim uma consequência, ou um corolário. A definição *usual* é

$$\left\{ \begin{array}{l} f \text{ é uma função contínua } X \xrightarrow{f} Y; \\ (\forall \mathcal{O} \text{ aberto de } Y) (f^{-1}(\mathcal{O}) \text{ é um aberto de } X); \end{array} \right. \quad (130)$$

de onde se pode deduzir a equação (eq. 129) uma vez que sendo  $O$  um aberto de  $Y$  então  $f^{-1}(O)$  é um aberto de  $X$  portanto existe um aberto  $\mathcal{B}$  de  $X$  tal que

$$f(\mathcal{B}) \subset \mathcal{O}; \quad (131)$$

No Cálculo existe uma outra forma de definir continuidade, chamada *continuidade sequencial* que é muito mais simples apesar de ser considerada *mais avançada*. Considere uma sucessão  $(s_k)_{k \in \mathbf{N}}$  convergindo para o ponto  $x = a$ :

$$s_k \longrightarrow a; \quad (132)$$

$$\text{Se } f \text{ for contínua no ponto } x = a \text{ então} \quad (133)$$

$$y_k = f(s_k) \longrightarrow f(a); \quad (134)$$

Quer dizer que as funções contínuas preservam a convergência. Observe analisando a figura (fig 18) que se uma sucessão  $(t_k)_{k \in \mathbf{N}}$  convergir para  $x = b$  a sucessão imagem  $(f(t_k))_{k \in \mathbf{N}}$  pode ser divergente, pode ter como limite  $f(b^-)$  ou  $f(b^+)$ . Por exemplo, se  $(t_k)_{k \in \mathbf{N}}$  oscilar a volta de  $x = b$  então a sucessão imagem será divergente. Num ponto de salto  $f$  pode não preservar a convergência de alguma sucessão o que implica na falha da *continuidade sequencial*.

A continuidade sequencial permite o enunciado dum critério muito útil para testar a continuidade que é “comutatividade” entre  $f$ ,  $\lim$  estabelecendo quando  $f$  é contínua no ponto  $x = a$

$$\lim_k f(t_k) = f(\lim_k t_k); \quad (135)$$

Esta troca de símbolos permite a criação de funções contínuas a partir de expressões complicadas como  $f(x) = x \sin(\frac{1}{x})$  que não está definida para  $x = 0$ . Mas, como podemos provar que o limite desta expressão no zero é zero, então podemos redefinir  $f$  como  $f(0) = 0$  obtendo a função contínua

$$\begin{cases} \text{Se } x \neq 0 & f(x) = x \sin(\frac{1}{x}); \\ \text{Se } x = 0 & f(x) = 0; \end{cases} \quad (136)$$

e aqui você tem um *exemplo duma função contínua cujo gráfico não pode ser feito sem que se retire a ponta do lápis do papel...*

A respeito de funções do tipo mostrada na equação (eq. 136), posso mencionar alguns conceitos que você deverá encontrar em outro lugar, como *oscilação*, *limite superior* e *limite inferior*. A função  $f(x) = x \sin(\frac{1}{x})$  tem oscilação  $\omega_f(0) = 0$  e isto implica que é possível calcular-se o seu limite no ponto zero.

Do texto se deduz que o conceito de limite é um prerequisite para discutir a continuidade. Isto não é verdade e aqui voltamos a ver que a continuidade é mais difícil do que parece... nos espaços topológicos o conceito de limite fica difuso e algumas vezes simplesmente deixa de ter sentido quando topologia é muito rica. É quando aparece o conceito de *ponto de acumulação* que é uma das formas de generalizar o *limite*. É neste momento que a equação (eq. 130) é usada como definição de continuidade. A equação (eq. 130) se verifica extremamente flexível para ser usada em demonstrações e junto com continuidade sequencial são as duas formas mais simples de caracterizar a continuidade.

O defeito da continuidade sequencial é que ela está associada às topologias semelhantes a que a reta real tem como *topologia usual* associadas *riqueza* de abertos da topologia.

A leitora pode ver que onde se encontra a razão da dificuldade inicial do Cálculo: *limite e continuidade*. Outra razão desta dificuldade se encontra no fato de que a grande maioria das funções que utilizamos são contínuas: polinômios, funções trigonométricas, desconsiderando a tangente e a cotangente, exponencial, módulo e logaritmo. Mais do que isto, se duas funções  $f, g$  forem contínuas, então as funções  $f + g, -f, -g, fg, fog, gof$  também são contínuas valendo a afirmação tanto para a continuidade local como global. E se o limite de  $g$  no ponto  $x = a$  for diferente de zero, então  $f/g$  será contínua no ponto  $x = a$ , continuidade local.

Estas propriedades estabelecem um estoque imenso de funções contínuas que cobre a grande maioria das funções que usamos no cotidiano. Quem está se iniciando no Cálculo tem a impressão que *descontinuidade* é um conceito inventado para usar em provas... Na verdade o próprio exemplo da divisão mostra que é fácil construir exemplos de funções que deixam de ser contínuas n'algum ponto como é o caso das funções trigonométricas *tan*, *cotan* ou de qualquer função racional cujo denominador tenha zeros.

Na verdade é preciso avançar um pouquinho mais para compreender que a continuidade é essencial e aparece de forma natural. A derivação oferece esta oportunidade. Considere a função módulo,

$$f(x) = |x|; \quad (137)$$

uma função contínua cujo gráfico é formado, à esquerda do zero pela semireta de equação  $y = -x$  e à direita do zero pela semireta de equação  $y = x$ . Confira o gráfico na figura (fig 19), página 57,

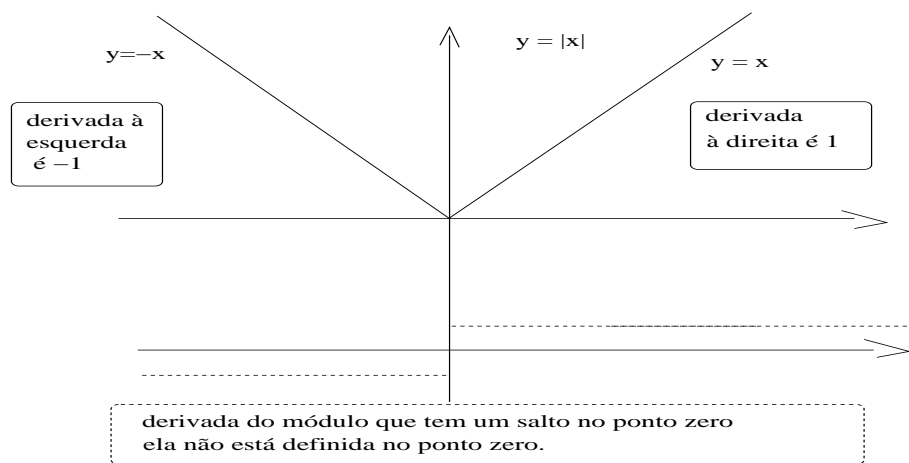


Figura 19: A derivada de  $f(x) = |x|$  é descontínua no zero

Como a derivada duma função num ponto onde o gráfico seja um segmento de reta é exatamente o coeficiente angular da reta, então derivada à esquerda é -1 e a derivada à direita é 1. Na figura (fig 19) você pode ver dois gráficos sincronizados, o segundo é o gráfico da derivada do módulo que tem um salto no ponto zero onde ela não está definida.

Então a função módulo é contínua e é *quase sempre derivável*, não sendo derivável apenas no zero onde a derivada tem um ponto de salto. Os valores laterais da derivada do módulo são

$$|0^-|' = -1; |0^+|' = 1; \quad (138)$$

consequentemente a derivada do módulo não existe no zero.

Qualquer função cujo gráfico seja uma poligonal, como as que aparecem com frequência ilustrando estatísticas, podem apresentar este fenômeno em que suas derivadas sejam descontínuas.

Mas, ao estudar a equação do movimento dum carro você verá outro exemplo simples e cotidiano de descontinuidade: da aceleração. Cada vez que o motorista troca a marcha há um salto na aceleração.

- a segunda derivada do movimento, a aceleração tem saltos nas trocas de marcha, em geral será uma função descontínua.
- a derivada do movimento, a velocidade em geral será contínua, mesmo nas freiadas bruscas ou nas trocas de marcha;
- a função do movimento, o gráfico do movimento será contínuo, e derivável, com derivada contínua.

Este exemplo mostra que existe uma descontinuidade escondida na segunda derivada de uma função, no caso da função do movimento. Este fato dá aso à criação de *classes de continuidade* que ficam marcadas pela continuidade até uma certa ordem de derivada. O movimento dum corpo seria de classe  $\mathcal{C}^1$  porque até a derivada de primeira ordem se tem continuidade que é perdida na derivada de segunda ordem, quando a aceleração tiver pontos de salto.

O estudo da continuidade e do limite pode se alongar por dezenas de páginas, sendo um verbete de dicionário apenas um resumo rápido não pode se alongar na análise de todas as nuances do conceito. Vou terminar apresentando a demonstração da afirmação de que soma e produto de funções contínuas, continuidade global, é também uma função contínua. Vou fazer a demonstração no caso real para usar a continuidade sequencial o que cobre as necessidades do Cálculo. Vale, obviamente, para o caso local permitindo escrever-se um teorema semelhante para a divisão de funções que inclua a restrição sobre os zeros da função no quociente.

- Então considero duas funções,  $f, g$  que sejam globalmente contínuas e uma sequência qualquer  $t_k \rightarrow x \in \mathbf{R}$ .
- então as sequencias imagens  $f(t_k), g(t_k)$  convergem

$$f(t_k) \rightarrow f(x); g(t_k) \rightarrow g(x)$$

- e como *o limite da soma é a soma dos limites* então,

$$f, g(t_k) \rightarrow f, g(x)$$

Esta demonstração tem uma lacuna contida na frase “*o limite da soma é a soma dos limites*” que uma propriedade do “limite” que está sendo usada aqui. Também estou usando implicitamente que a função “*soma*” definida no espaço  $\mathbf{R} \times \mathbf{R}$  é uma função contínua e no texto anterior não há nenhuma menção à definição de continuidade de funções duas variáveis.

Esta observação serve para que a leitora sinta como é difícil manter um sistema lógico completo e sem bolhas. Não pode ser a pretensão do autor dum verbete de dicionário atingir esta perfeição, assim como também tem que ficar claro que melhoras do texto podem ser alcançadas pela crítica colaborativa. Mas seria desonesto não apontar as falhas. Seria preciso alertar para o fato de que  $\mathbf{R}$  é apenas um conjunto de sucessões convergentes de números racionais e que portanto o conceito de limite pertence à construção dos números reais o que torna o teorema sobre limite mencionado acima apenas um propriedade dos números reais, coisa que fica invisível dentro dos cursos de Cálculo.

Um resultado semelhante pode ser proposto para o produto de funções que é uma função contínua se cada fator o for. Novamente aqui está envolvida a continuidade do produto como função bilinear de  $\mathbf{R} \times \mathbf{R}$  e dum teorema sobre o produto de limites, que, novamente, é apenas um propriedade dos números reais.

Para terminar um teorema que ficará completamente contido no texto. Se

$$\mathbf{R} \xrightarrow{f} \mathbf{R}, \mathbf{R} \xrightarrow{g} \mathbf{R}$$

forem funções contínuas então as composições de função  $f \circ g, g \circ f$  são contínuas.

Porque, como  $f$  é contínua, então para toda sequência convergente  $t_k \rightarrow x$  se produz uma sucessão convergente  $y_k = f(t_k) \rightarrow f(x)$  e pela continuidade de  $g$  se transforma a sucessão convergente  $(y_k)_{k \in \mathbf{N}}$  n'outra sucessão convergente

$$z_k = g(y_k) = g(f(t_k)) \rightarrow g(f(x)) = g \circ f(x)$$

provando a convergência da função composta  $g \circ f$ . É semelhante a prova de que  $f \circ g$  é contínua.

Compare esta demonstração *crystalina* com a demonstração tocada a es e  $\delta s$  para a continuidade da função composta.

- **continuidade sequencial** É um tipo restrito de continuidade: se uma função for *sequencialmente contínua* então será também *contínua*.

**Observação 1** *continuidade é um item da topologia*

*A continuidade sequencial está associada a um tipo de topologia que pode ser descrita por famílias enumeráveis de abertos e portanto uma classe de topologia fraca.*

*O conjunto dos números reais ou mesmo os espaços reais de dimensão finita,  $\mathbf{R}^n$ , tem uma topologia chamada usual cuja topologia pode ser definida por famílias enumeráveis de abertos e neste topologia usual os dois conceitos, continuidade sequencial e continuidade são equivalentes. A continuidade do Cálculo poderia ser simplesmente a continuidade sequencial que é mais ágil e serve inteiramente às necessidades do Cálculo tornando as demonstrações mais simples e mesmo mais elegantes.*

Para uma função

$$\mathbf{R} \xrightarrow{f} \mathbf{R}$$

ou

$$(a, b) \xrightarrow{f} \mathbf{R}$$

uma função definida num subintervalo da reta, a continuidade sequencial significa *preservar convergência de sucessões* para os pontos do intervalo  $(a, b)$

**Definição 5 (continuidade sequencial)** de funções reais

Se  $f$  estiver definida no intervalo  $(a, b)$

$$(a, b) \xrightarrow{f} \mathbf{R}; c \in (a, b); \quad (139)$$

$$\lim_n x_n = c \Rightarrow \lim_n f(x_n) = f(c) \quad (140)$$

Esta definição é dita *continuidade sequencial* que é uma forma particular que vale para as funções reais de variável real, quando considerarmos o conjunto dos números reais munido da sua *topologia usual*.

Em outras palavras,  $f$  é contínua se  $(x_n)_n$  definir o número  $c$  então  $(f(x_n))_n$  define o número  $f(c)$ . Isto é verdade para a grande maioria das funções que usamos no Cálculo o que torna este conceito difícil pela prática ausência de contra-exemplos, os contra-exemplos em geral parecem “fabricados”, e a grande quantidade de funções não contínuas é pouco intuitiva. Um exemplo “não fabricado” pode ser o seguinte: a função  $y = f(x) = \|x\|$  é contínua, mas sua *derivada*<sup>4</sup>

$$f'(x) = \begin{cases} x < 0 & -1 \\ x > 0 & 1 \end{cases}$$

é descontínua no ponto  $c = 0$ , aliás, não é definida para  $c = 0$ , observe que a definição exige que  $f(c)$  esteja definido. Podemos facilmente verificar que esta função contradiz a definição, apenas para ficar mais fácil, vamos fazer uma pequena modificação em  $f'$  chamando esta modificação de  $g$

$$g(x) = \begin{cases} x \leq 0 & -1 \\ x > 0 & 1 \end{cases}$$

que está definida no ponto  $c = 0$  onde vale  $-1$ . A seleção de uma sucessão que convirja para zero, por exemplo,  $(\frac{1}{n})_{n \in \mathbf{N}}$  tem uma imagem constante, igual a 1 que não converge para  $g(0) = -1$ .

Fizemos esta alteração porque a função  $f'$  pode ser dita “contínua” se alterarmos o domínio para  $(-2, 2) - \{0\}$  “retirando o ponto de descontinuidade”. Definindo assim diríamos que  $\|x\|$  é uma função derivável no domínio  $(-2, 2) - \{0\}$ . Esta é uma forma de dar exemplos artificiais, retirando algum ponto do conjunto de definição de uma função descontínua.

A teoria não está errada e nem tem engodos, algumas vezes precisamos de fazer restrições deste tipo na solução de problemas, e isto mostra como “continuidade” é difícil, embora, de certa forma seja um conceito relativamente simples, expresso num axioma da natureza - *a natureza não dá saltos*<sup>5</sup>. Se  $\|x\|$  estiver definida na reta inteira, então sua derivada não estará definida no ponto zero o que a torna imediatamente descontínua neste ponto uma vez que não lhe podemos aplicar a definição. Logo aqui podemos ver uma dificuldade na “*pedagogia do*

<sup>4</sup>Aqui havia um erro, em edição anterior do dicionário.

<sup>5</sup>Natura non facit salta - Gottfried Leibniz, que depois de Charles Darwin deve ser **natura saltus facit, sed parvos** a natureza dá saltos, porém pequenos...

ensino da continuidade”, no desenvolvimento da disciplina, do Cálculo, a continuidade é um dos primeiros assuntos e *derivada* vem depois, tornando difícil, mas não impossível, apresentar este exemplo.

O conceito de *continuidade* é nato dos espaços topológicos, é nesta estrutura que se define continuidade de forma natural para isto se usa o conceito de *vizinhança aberta* que aparece na maioria dos livros de Cálculo:

**Definição 6 (continuidade)** *Definição topológica de continuidade*

Se  $(a, b) \xrightarrow{f} \mathbf{R}$  e  $c \in (a, b)$  e se para qualquer seja  $\epsilon > 0$  existir  $\delta > 0$  tal que

$$\|x - c\| < \delta \Rightarrow \|f(x) - f(c)\| < \epsilon \quad (141)$$

então  $f$  é contínua no ponto  $c$ . Se for contínua em qualquer ponto do intervalo  $(a, b)$  dizemos que é contínua em  $(a, b)$ .

Modificando a linguagem podemos escrever, primeiro:

$$(\forall \mathcal{V}_{f(c)}) (\exists \mathcal{V}_c); (x \in \mathcal{V}_c) \Rightarrow f(x) \in \mathcal{V}_{f(c)}; \quad (142)$$

em que você pode reconhecer “ $\epsilon, \delta$ ” substituídos por vizinhança aberta de um ponto, que são os elementos básicos dos espaços topológicos. Como a implicação define inclusão de conjuntos, podemos finalmente escrever:

$$(\forall \mathcal{V}_{f(c)}) (\exists \mathcal{V}_c); (f(\mathcal{V}_c) \subset \mathcal{V}_{f(c)}); \quad (143)$$

Conclusão, os livros de Cálculo, habitualmente, usam a definição mais difícil e abstrata para continuidade que não é necessário.

- **convecção** É um dos três métodos de transmissão de calor: condução, convecção e radiação. Ver *condução*

- **convergência** É um conceito associado à aproximação. Há várias formas de considerar-se “convergência” e elas dependem da *densidade* do espaço em que se estiver trabalhando, ou sua categoria que é uma forma de classificar conjuntos considerando “quantidade” (cardinalidade) de abertos de sua topologia. Por exemplo não tem sentido em falar-se de convergência dentro do conjunto dos números naturais<sup>6</sup>, mas cabe falar em convergência dentro de  $\mathbf{Q}$  e neste caso se podem descobrir as “falhas” do conjunto dos números racionais. Uma forma de abordar convergência em  $\mathbf{Q}$  pode ser descrita sumariamente assim:

Sumário da construção dos números reais. Este é um resumo da forma como Cauchy fez a construção dos números reais.

1. considerar o conjunto das sucessões de de números racionais, este conjunto é muito amplo<sup>7</sup>, por exemplo, a sucessão dos números naturais é um dos elementos deste conjunto, e no contexto deste verbete, é um elemento

<sup>6</sup>Tem sentido, ver topologia...

<sup>7</sup>Havia um erro, neste ponto, em edição anterior.

inútil... É possível encontrar uma “lei” que elimine elementos deste tipo que posteriormente serão classificados como “sucessões não convergentes”, porém isto tornaria a teoria mais difícil, e vamos usar uma alternativa.

Temos que classificar os seus elementos em duas classes, das *sucessões convergentes* e a classe das *não convergentes*. Uma forma bonita e efetiva passa pela estrutura de grupo:

- (a) O conjunto de todas as sucessões é um grupo aditivo;
- (b) o subgrupo nulo O subconjunto das sucessões que satisfazem á condição

$$(\forall \epsilon > 0)(\exists N \in \mathbf{N})(\forall n > N)(\|x_n\| < \epsilon) \quad (144)$$

é um sub-grupo aditivo do grupo de todas as sucessões que vou designar por  $c_0$  e a razão é porque ele vai ser o *zero* do novo conjunto. É o subgrupo formado de todas as sucessões que vou definir como “convergentes para zero”, observe que é uma definição, ou apenas uma etiqueta. Vou chamar esta “classe” de “*classe do zero*” e logo você verá a razão.

Como estamos num grupo comutativo é desnecessário mencionar que este grupo é normal, propriedade necessária no próximo passo.

- (c) Quando consideramos o quociente de um grupo por um seu grupo (normal) o resultado é um grupo também - das classes quociente. Foi criada uma classificação, quando translataremos a classe do zero, vamos obter todas as outras classes. Mas esta linguagem é insuficiente uma vez que estamos criando novos objetos que não conhecemos e ao falar em “translatar” queremos dizer  $c_0 + a$  em que  $a$  é um elemento do grande grupo, entretanto o conjunto de todas as classes quociente formam o total destas “translações”. Por exemplo, um elemento do grande grupo é uma sucessão constante de números racionais e o resultado da translação é o conjunto de todas as sucessões equivalentes a esta - quer dizer as sucessões que vamos “etiquetar” com a expressão  $\lim_n x_n = a$  em que “ $a$ ” é o valor constante da sucessão  $x$ . Mas há sucessões de números racionais para as quais não sabemos escrever uma etiqueta tão simples, por exemplo as que convergirem<sup>8</sup> para  $\pi$ , ou para  $\sqrt[3]{2}$ . Por isto usei a expressão “etiqueta” e vou escrever  $c_0 + \pi$  para me referir a esta classe. Se você conhecer algum método para construir  $\pi$  você pode descobrir elementos nesta classe, e se você descobrir um elemento os outros passam a ser sem importância (são equivalentes ao que você tiver descoberto...) - procure  $\pi$ .
- (d) sucessões convergentes Para eliminar os elementos inconvenientes, como a sucessão dos números naturais, e outras sucessões “não convergentes” Cauchy criou um teste, definindo as *sucessões de Cauchy*

---

<sup>8</sup>Havia um erro, em edição anterior, neste ponto.



**Definição 7 (Teste) de Cauchy**

$$(\forall \epsilon > 0) (\exists N \in \mathbf{N}) (n, m > N) (\|x_n - x_m\| < \epsilon) \quad (145)$$

Uma sucessão que satisfaça (145) se chama sucessão de Cauchy

É relativamente simples mostrar que a sucessão dos números naturais não é uma sucessão de Cauchy, a desigualdade que aparece na equação (145) falha para qualquer valor de  $\epsilon$  que seja inferior a 1, ... depois vou dizer que “sucessão convergente é equivalente a sucessão de Cauchy” então a sucessão dos números naturais é uma sucessão divergente.

Vou chamar o conjunto de todas as sucessões de Cauchy de  $c$ . As sucessões que satisfazem a este teste formam um grupo, é este o grande grupo,  $c$ , que nos interessa, e acabei de mostrar que  $\mathbf{N}$  não é deste tipo, nem sucessões como  $((-1)^n)_{n \in \mathbf{N}}$ , uma sucessão oscilante com oscilação constante. A classe  $c_0$  é um subconjunto de  $c$ , um subgrupo.

Sim, temos que mostrar que  $(c, +)$  é um grupo, mas tem apenas uma pequena dificuldade que fica como exercício: mostrar que é fechado para soma, e sai usando a *desigualdade triangular*.

2. O grupo quociente  $c/c_0$  é um novo conjunto que contém  $\mathbf{Q}$ , é o conjunto  $\mathbf{R}$  dos números reais.
3. Esta construção define a convergência de sucessões numéricas (podemos falar em sucessões de números reais agora - é o mesmo conjunto definido acima... das classes quocientes do grupo de todas as *sucessões de Cauchy* de números racionais pelo subgrupo  $c_0$ . O operador  $\lim$ , aqui, apenas identifica um número real, como um elemento de uma destas classes de equivalência.

---

- **convolução** O produto de convolução é uma operação definida em conjuntos de *funções*, de *distribuições* ou de sucessões (convolução discreta). Se  $f, g$  forem duas funções integráveis sobre o conjunto dos números reais, podemos definir uma terceira função com a equação

$$\phi(x) = \int_{-\infty}^{\infty} f(t)g(x-t)dt; \phi = f * g; \quad (146)$$

O físico Dirac, precisou deste produto para montar um espaço de Hilbert de “objetos” representando *ondas* na sua construção teórica da Mecânica Quântica e com isto criou o que durante muito tempo foi chamado de “*função de Dirac*” que é a unidade relativamente à convolução:

$$(\delta_0 * g)(x) = \int_{-\infty}^{\infty} \delta_0(t)g(x-t)dt = g(x); \delta_0 * g = g = g * \delta_0; \quad (147)$$

apenas esta expressão não pode ser vista como uma integral porque a *delta de Dirac* não é uma função. Com a teoria das distribuições esta situação incômoda foi resolvida. A *distribuição de Dirac* é a unidade relativamente ao produto de convolução, entretanto o anel das distribuições (com o produto por convolução) é uma extensão dum anel de funções integráveis (com o produto por convolução) que é um anel sem unidade, porque a unidade é a *delta de Dirac* que não é uma função.

A transformada de Fourier (ou a transformada de Gelfand) relaciona anéis de funções em que o produto é definido ponto a ponto, com anéis de distribuições em que o produto é o *produto por convolução*:  $\widehat{fg} = \widehat{f} * \widehat{g}$ . O símbolo  $\widehat{f}$  representa a *transformada de Fourier* de  $f$ .

Durante anos a fórmula na equação (147) representou apenas um aspecto teórico sem grande “interesse prático” porque calcular convoluções era muito difícil para o cálculo manual. Hoje até algoritmos não otimizados calculam convoluções com relativa rapidez o que re-acende o interesse por esta fórmula. Na figura (fig 20), página 64, você pode ver a convolução da função característica

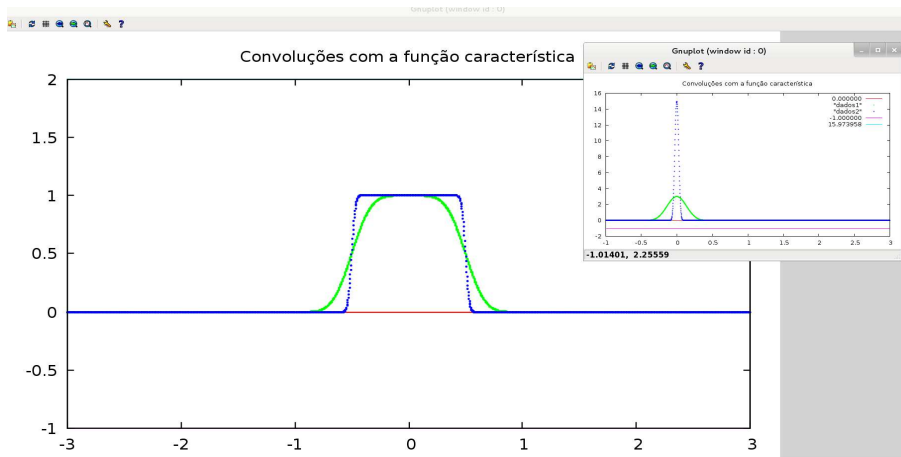


Figura 20: convolução de  $\chi_{[-0.5,0.5]}$  com dois núcleos

$\chi_{[-0.5,0.5]}$  com duas aproximações da medida de Dirac, no detalhe aparecem dois elementos de uma “unidade aproximada” utilizadas no cálculo com um programa em `python` e o processamento durou 18 segundos e como foram obtidos 2 mil pontos para cada um dos gráficos então foram calculadas 4 mil integrais. O termo *núcleo* é usado com frequência para fazer referência a um elemento duma *unidade aproximada* e uma *unidade aproximada* é uma sucessão de núcleos cujos suportes encolhem “convergindo” para um ponto.

As “unidades aproximadas” são as aproximações que Dirac usou e nas quais se inspirou para afirmar que havia uma função *nula em todos os pontos da reta, dando um salto para o infinito no ponto zero e cuja integral seria 1*.

Além do mais a convolução aparece junto com o *núcleo de Dirichlet* na

definição da transformada de Fourier que pode ser expressa como uma sucessão de produtos de convolução por unidades aproximadas convergindo para  $\delta_0$ . Em artigo escrito em 1953, D. Widder, afirma que “todas as transformações integrais podem ser colocadas na forma de um produto por convolução”, e hoje a teoria dos operadores confirma a importância do produto de convoluções na forma como Widder pensava. Na figura (fig. 20), página 64, você pode ver dois elementos de uma tal sucessão se aproximando da função  $\chi_{[-0.5,0.5]}$  e na figura (fig 21), página 65, você vê a convolução da função característica  $\chi_{[-0.5,0.5]}$  com

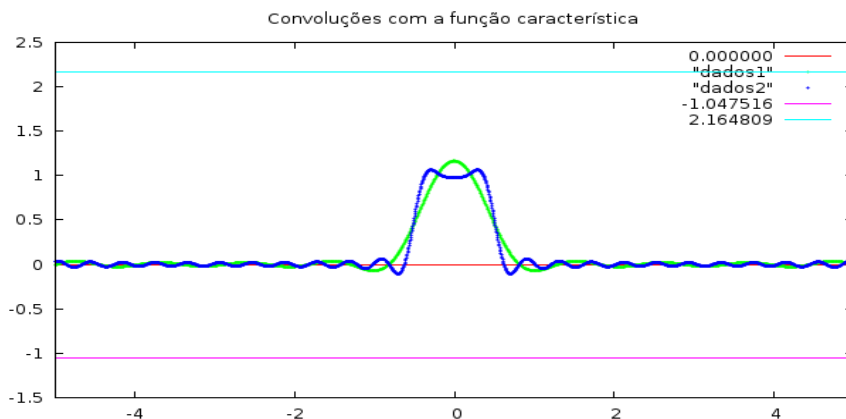


Figura 21: convolução de  $\chi_{[-0.5,0.5]}$  com núcleos de Dirichlet

dois valores do núcleo de Dirichlet,  $D_n$  cuja equação é

$$D_n = \frac{1}{2\pi} \frac{\sin(n + 1/2)x}{\sin(x/2)} \quad (148)$$

Isto é uma média Considere agora que na equação (146), uma das funções, a função  $g$  tenha integral 1 e seja positiva, como na figura (22), página 66. A integral na equação (146) é o valor médio de  $f$  no suporte de  $g$ . Se o suporte de  $g$  estiver contido numa vizinhança de raio  $\epsilon$  do ponto  $x = a$  então a equação (146) é o valor médio integral de  $f$  nesta vizinhança. Este raciocínio nos conduz ao valor de médio de  $f$  no ponto  $x = a$ , mesmo que neste ponto  $f$  não seja contínua, como mostra a figura (23), página 67.

Como a integral sobre um intervalo compacto converge uniformemente, podemos derivar dentro da integral, comutando os operadores *derivada* e *integral*, e se  $g$  for diferenciável então  $f * g$  será diferenciável com

$$(f * g)' = f * g' \quad (149)$$

É o que se chama de regularização por convolução, porque se obtém assim uma função diferenciável,  $f * g$ , a partir de uma função que não precisa ser diferenciável,  $f$ . Se o suporte de  $g$  tiver uma medida pequena então  $f * g \approx f$ .

Isto é uma média

$$\int_{-\infty}^{\infty} f(t)g(x-t) dt = f * g(x)$$

$$\int_{-\infty}^{\infty} g(t)dt = 1; \quad g > 0$$

Figura 22: Convolução com uma função cuja integral seja 1

É isto que se chama de aproximação por convolução. Neste caso a função  $g$  recebe um nome especial, se chama *núcleo*, ou *kernel*, e tem uma definição apropriada para lidar com aproximações, é uma função cuja integral é 1, é positiva, e se deseja que tenha um suporte cuja medida seja “pequena” (para produzir aproximações).

O *núcleo de Dirichlet* representa uma exceção uma vez que não é uma função positiva, mas é uma exceção perfeitamente aceitável, primeiro porque se encontra dentro da definição da *transformada de Fourier*, depois porque, embora não seja positivo, quando  $n$  cresce, ele se “concentra” em volta da origem com valores positivos, sua integral é 1. Os valores negativos se tornam desprezíveis a medida que  $n$  cresce. De qualquer forma é o fato de que  $D_n$  oscile em torno de zero que faz com que a transformada de Fourier de uma função oscile em volta da função. A figura (fig 24), página 68, mostra a diferença entre a série trigonométrica<sup>9</sup> e a convolução com o núcleo de Dirichlet usando a função  $f(x) = x^2$ . A convolução  $D_{30} * f$  é uma “aproximação” de  $f$  e como tal uma função ilimitada. A série de Fourier de  $f$  é uma função periódica e  $S_{30}$  coincide com  $D_{30} * f$  no intervalo de expansão da série de Fourier que na figura (fig 24) é o intervalo  $[-\pi, \pi]$ . As séries de Fourier, como este exemplo bem ilustra, não podem ser vistas como “aproximações” funções, o seu significado é bem outro e passam a ter sentido

<sup>9</sup>Na verdade um polinômio trigonométrico, uma reduzida da série.

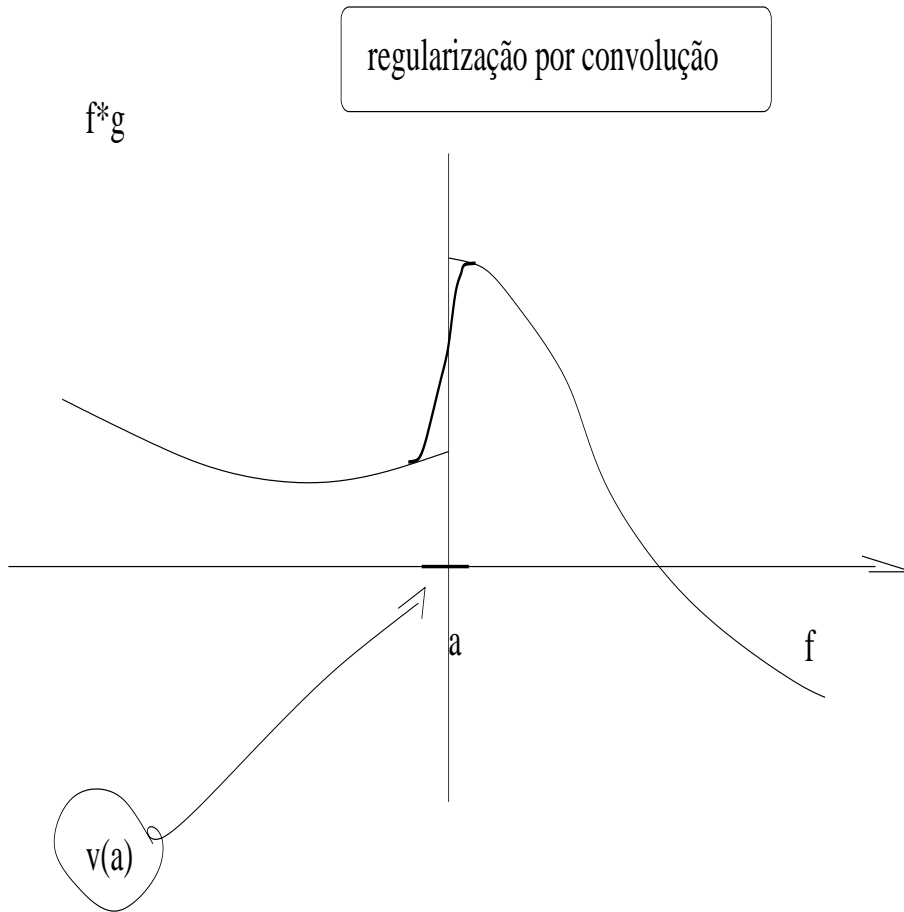


Figura 23: regularização por convolução

quando aplicadas a sinais para codificá-los e depois transmiti-los, foram muito importante até a década de 70 nas comunicações com este objetivo mas depois substituídas pelas wavelets.

A equação (149) estabelece a possibilidade de calcularmos a derivada de uma função, mesmo que ela não seja derivável usando a aproximação  $f * g$  quando  $g$  for um núcleo diferenciável porque então  $f * g$  é uma função diferenciável que é uma aproximação de  $f$  se o suporte de  $g$  tiver uma medida “pequena”. A figura (fig 25), página 69, mostra três gráficos, de  $f, f', Dif(f, \rho')$  em que  $f$  é o polinômio mônico com raízes nos inteiros  $\{-10, -7, -5, -1, 2, 5, 7, 10, 13\}$ , sua derivada, e a imagem da convolução de  $f$  com a derivada do núcleo  $\rho$ , consequentemente  $Dif(f, \rho')$  é uma aproximação de  $f'$ . Na figura os gráficos de  $f'$  e de  $Dif(f, \rho')$  se confundem dando a errônea impressão de que são idênticos. Há um erro que não seria visível ao olho mas fazendo uma varredura no intervalo  $[-5, 5]$

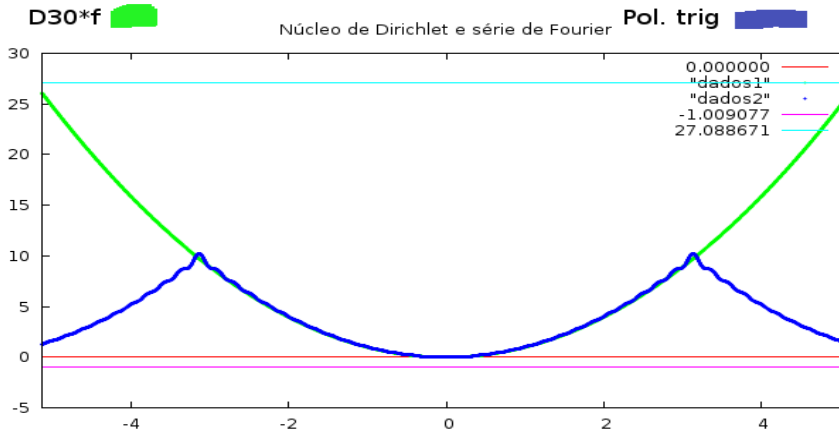


Figura 24:  $D_{30} * f$ ;  $f(x) = x^2$ ;  $S_{30}(x)$

o erro máximo encontrado é 225.3144 para a expressão  $\|f'(x) - Dif(x, f, \rho')\|$  que é um erro bastante grande, entretanto esta é uma função que assume valores muito grandes e a variação da derivada é ainda maior! Interessa usar a convolução, como aproximação para funções cujas imagens não assumam grandes valores, como os sinais de telecomunicação. Se multiplicarmos  $f$  por 0.0001 este erro máximo cai para 0.02253144 tornando a figura (fig 25) um exemplo confiável.

Os polinômios trigonométricos associados como uma determinada função  $f$  podem ser expressos como convoluções com o núcleo de Dirichlet o que permite então ver um polinômio trigonométrico como uma aproximação da função  $f$  que serviu de base para encontrar os coeficientes de Fourier deste polinômio. Estes coeficientes ainda são chamados de *coeficientes de Fourier de  $f$* . Mas o núcleo de Dirichlet não é positivo o que implica em conceber um polinômio trigonométrico como uma expansão do conceito de média e justifica a oscilação dos polinômios trigonométricos em volta da função (ou onda) dos quais eles são uma aproximação. Como os núcleos de Dirichlet se concentram na vizinhança de zero a medida que  $n$  cresce, então podemos usar a convolução para entender porque os polinômios trigonométricos “passam” no ponto médio do gráfico de  $f$  nos pontos de salto produzindo o fenômeno de Gibbs.

Uma outra aplicação da convolução pode ser vista na construção de *splines*. Se  $f = \chi_{[-\epsilon, \epsilon]}$ , a função característica de um intervalo centrado na origem de medida  $2\epsilon$ , então  $f * f$  é uma função linear por pedaços cujo gráfico será um triângulo com suporte medindo  $4\epsilon$  que pode ser visto na figura (26) página 70, portanto um 1-splines com suporte compacto. Calculando outra potência por convolução, a terceira potência por convolução de  $f$ , vamos encontrar pedaços de parábola formando agora uma função diferenciável com derivada contínua, de classe  $\mathcal{C}^1$ , portanto um 2-splines a suporte compacto. O suporte desta nova função mede  $6\epsilon$ , pois os suportes vão sendo somados (soma de conjuntos).

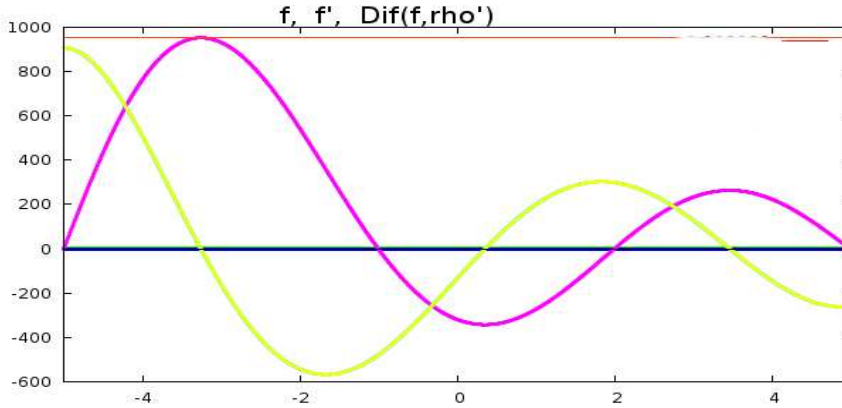


Figura 25: aproximação por convolução da derivada

Podemos agora estabelecer a hipótese de indução de que a  $k$ -ésima potência por convolução seja um  $(k-1)$ -splines a suporte compacto, quer dizer uma função de classe  $\mathcal{C}^{k-1}$  formada de polinômios de grau menor ou igual a  $k-1$ , com suporte medindo  $2k\epsilon$ . Uma nova potência por convolução vai nos fornecer um grau a mais na classe de continuidade e nos pedaços de polinômios, quer dizer um  $k$ -splines, mas com suporte  $(2k+2)\epsilon$ . É possível manter o suporte sob tamanho controlado com uma mudança de variável adequada da função característica cujas potências se estiver calculando. Este método mostra que é possível obter-se funções altamente diferenciáveis,  $n$ -splines com suporte concentrado em volta da origem e com integral 1. Uma família de tais funções é chamada de *unidade aproximada*, e representam a medida de Dirac no sentido de que “convergem” para ela, apenas esta convergência precisa ser definida de forma adequada. É uma convergência chamada fraca, ou no sentido das *distribuições*. Observe que para garantir que a integral sempre seja 1, é preciso que a transformação seja do tipo  $K\eta(Kx)$ ;  $K \in \{1, 2, \dots\}$  fazendo com o suporte encolha em volta de um ponto e que o máximo cresça proporcionalmente a  $K$ . Dirac deve ter trabalhado experimentalmente com tais funções o que o levou a pensar numa “função” concentrada num ponto e neste ponto o seu valor fosse  $\infty$  e com integral 1.

- **convolução discreta** O produto de dois polinômios pode ser facilmente definido usando o *produto de convolução discreto*. Observe que o produto de dois polinômios, ou a soma de polinômios, é uma operação que envolve apenas os coeficientes, o uso das “variáveis” se reduz apenas ao mecanismo posicional dos coeficientes do resultado. Assim para multiplicar  $1 + x + x^2$  por  $1 + x$  preciso apenas das sucessões  $(1, 1, 1)$  e  $(1, 1)$  e a metodologia é semelhante a que usamos para multiplicar números, que, por sinal, podem ser “codificados” como polinômios em que a variável representa a base de numeração em que eles estiverem “codificados”. Experimente e se convença do que está dito a cima.

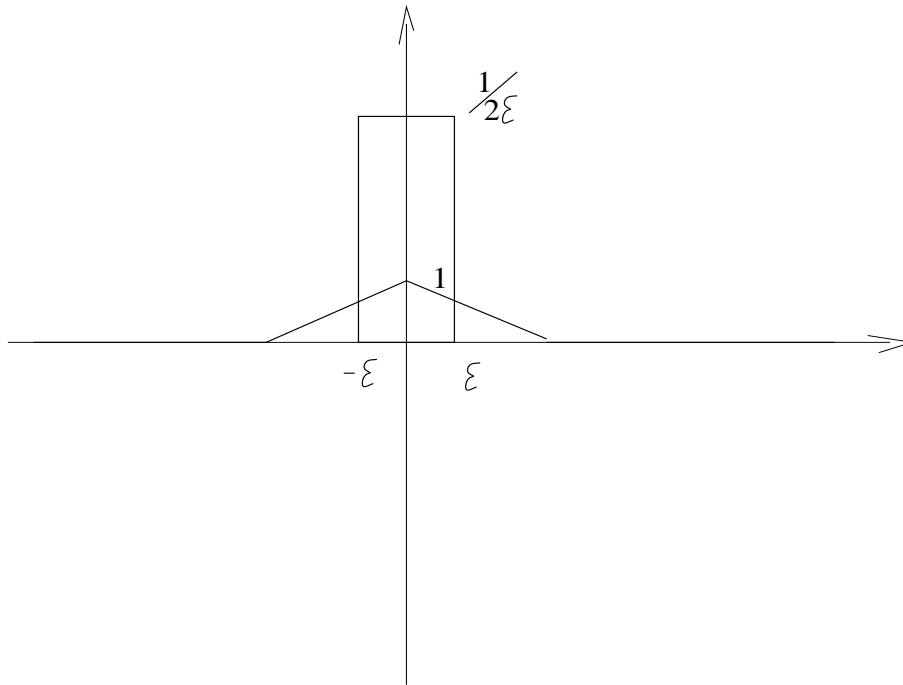


Figura 26: Segunda potência de convolução da função característica

Partindo da suposição de que você já se convenceu, passemos ao caso geral.

A figura (27) página 71, mostra as 15 primeiras linhas do *triângulo de Pascal* obtido com um programa em `python` para convolução discreta. É exatamente o mesmo que multiplicar o polinômio  $(1, 1)$  por ele mesmo 15 vezes.

Quero multiplicar os dois polinômios  $(a_0, \dots, a_n)$  e  $(b_0, \dots, b_m)$ . A maneira mais simples de fazê-lo consiste em criar uma *representação* para o conjunto dos polinômios. Um polinômio (mais geralmente, uma função racional) é uma sucessão a *suporte finito*, entenda tais sucessões como sendo sucessões não finitas, porém tendo apenas um número (finito) de valores não nulos. Podemos simplificar, sem perder a generalidade, translatando tais sucessões para que o primeiro elemento não nulo corresponda ao índice zero, para dois exemplos que estejamos multiplicando. Deixo como conjectura (não estou vendo uma demonstração) *o conjunto de todas as sucessões a suporte finito não é equipotente ao conjunto das sucessões a suporte finito tendo como primeiro índice um número inteiro positivo*. Possivelmente isto sai da comparação das funções polinômiais com as funções racionais. O conjunto das sucessões a suporte finito cujo primeiro índice seja um número natural representa (é equipotente a) o conjunto dos polinômios (a coeficientes reais ou complexos, importa pouco qual você escolher neste contexto). A retirada do predicado *suporte finito* criaria um problema de convergência que nos obrigaria a restringir esta questão aos espaços de sucessões





1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
1 8 28 56 70 56 28 8 1
1 9 36 84 126 126 84 36 9 1
1 10 45 120 210 252 210 120 45 10 1
1 11 55 165 330 462 462 330 165 55 11 1
1 12 66 220 495 792 924 792 495 220 66 12 1
1 13 78 286 715 1287 1716 1716 1287 715 286 78 13 1
1 14 91 364 1001 2002 3003 3432 3003 2002 1001 364 91 14 1
1 15 105 455 1365 3003 5005 6435 6435 5005 3003 1365 455 105 15 1

Figura 27: Triângulo de Pascal e convolução discreta

$l^p$ . Se  $(a_{k \in \mathbf{N}})$  e  $(b_{k \in \mathbf{N}})$  forem duas sucessões à suporte finito então elas representam polinômios. Partindo da experiência prática na multiplicação de dois polinômios, em que você usa apenas os coeficientes, os dados se agrupam em colunas nas quais a soma  $k + j = i$  é a constante  $i$  (que corresponde a soma de potências  $k + j$ , é apenas para isto que serve a variável). Esta soma pode ser reescrita como  $j = i - k$  e o resultado da multiplicação é a soma dupla

$$\sum_{j=0}^{n+m} \sum_{i-k=j} a_k b_j = \sum_{j=0}^{n+m} \sum_{k=0}^{k=m} a_k b_{i-k} \quad (150)$$

que resulta no polinômio de grau  $n+m$ . Como num polinômio não tem variáveis - polinômio é uma lista de coeficientes e as funções polinomiais transformam o valor fornecido via “ $x$ ” na soma habitual, então a “soma” externa é na verdade apenas a indexação da lista que resulta desta multiplicação portanto, na segunda

“soma dupla” podemos identificar o valor do coeficiente de ordem  $j$  como sendo

$$\sum_{k=0}^{k=m} a_k b_{i-k} = \sum_{k=0}^{k=m} b_{i-k} a_k = (b * a)(j) \quad (151)$$

uma expressão semelhante a que aparece na “convolução contínua” que identifiquei na equação (151) como “ $(b * a)(j)$ ”, o valor do coeficiente de ordem  $j$  do produto de polinômios.

- coordenadas cartesianas confira *plano cartesiano*.

- coordenadas polares é um sistema de referência para os pontos do plano determinando um endereço para cada ponto  $P$  do plano em função de sua distância à origem e do ângulo do vetor que este ponto determina com o eixo  $OX$ . O ângulo é o segmento do círculo trigonométrico,  $\mathbf{S}^1$ , determinado pela projeção do ponto  $P$  sobre  $\mathbf{S}^1$ , ou seja, na prática determinamos as coordenadas polares  $P = (\rho, \theta)$  por sua projeção sobre o círculo trigonométrico  $\mathbf{S}^1$  que determina o ângulo  $\theta$  e a distância à origem que determina  $\rho$ .

Para o espaço de dimensão 3 existem as coordenadas esféricas de  $P = (\rho, \alpha, \theta)$  em que a esfera  $\mathbf{S}^2$  substitui  $\mathbf{S}^1$  para determinar os ângulos  $\alpha, \theta$  e a distância do ponto à origem determina  $\rho$ .

As coordenadas polares essencialmente consistem em identificar um ponto  $P$  do plano a partir de sua projeção sobre o círculo unitário  $\mathbf{S}^1$ , determinando o ângulo  $\theta$  e depois à distância à origem,  $\rho$ , então  $P = (\rho, \theta)$ . Compare com a equação de Euler

$$P = (\rho, \theta); P \in \rho\mathbf{S}^1 = \rho(\cos(\theta), \sin(\theta)); \quad (152)$$

e confira a figura (fig 28), página 73, as coordenadas polares dum ponto  $P$  do plano mostram como o ponto  $P$  pode ser projetado sobre o círculo trigonométrico  $\mathbf{S}^1$  e vice-versa como podemos mapear o plano todo a partir de endereços sobre o círculo trigonométrico. Para isto se usam duas referências,

- a distância de  $P$  até a origem,  $\rho$ ,
- e o ângulo que a projeção de  $P$  determina sobre o círculo trigonométrico  $\mathbf{S}^1$ ,  $\theta$ .

As coordenadas polares de  $P$  são então  $(\rho, \theta)$  quando a projeção de  $P$  sobre  $\mathbf{S}^1$  for  $(\cos(\theta), \sin(\theta))$ .

As coordenadas polares e as coordenadas cartesianas se relacionam com as equações:

$$P = (\rho, \theta); \begin{cases} x = \rho \cos(\theta) \\ y = \rho \sin(\theta) \end{cases} ; \Rightarrow P = (x, y); \quad (153)$$

$$P = (x, y); \begin{cases} \rho = \sqrt{x^2 + y^2} \\ \theta = \text{Atan}(y/x) \end{cases} ; \Rightarrow P = (\rho, \theta); \quad (154)$$

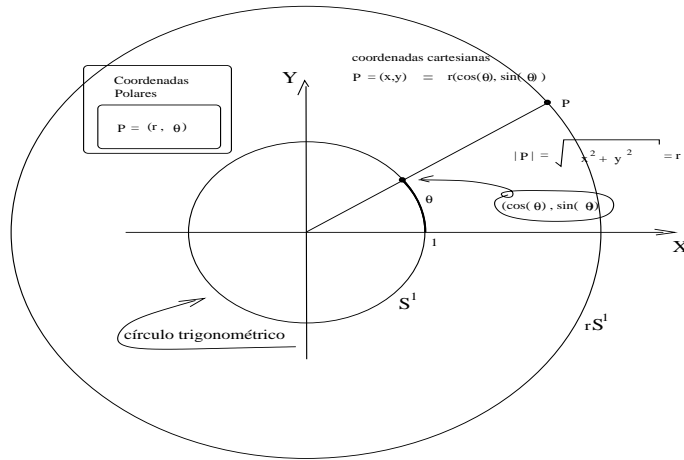


Figura 28: coordenadas polares dum ponto no plano

Somente vale a pena usar coordenadas polares para descrever um lugar geométrico quando este for de tipo radial, quer dizer, depender apenas do raio. Por exemplo, a equação do círculo de centro na origem e raio 4 em coordenada polares seria  $\rho = 4$ , significando que o valor de  $\rho$  é constante e que  $\theta$  é qualquer.

Experimente com `gnuplot`

```
set polar
rho = 4;
plot rho
```

**Observação 2** *Contra exemplos do uso das coordenadas polares:*

- O primeiro é a equação

$$x^2 + y^2 - 4x - 4y + 1 = 0; \quad (155)$$

pode ser facilmente transformada em coordenadas polares numa aplicação direta das equações (eq. 153) e (eq. 154).

$$\rho^2 - 4\rho \cos(\theta) - 4\rho \sin(\theta) + 1 = 0; \quad (156)$$

mas esta segunda equação não oferece nenhuma vantagem sobre a equação (eq.155) para identificarmos que curva é esta no plano, ao passo que, partindo da equação (eq. 155), completando os quadrados se encontra

$$x^2 - 4x + 4 - 4 + y^2 - 4y + 4 - 4 + 1 = 0; \quad (157)$$

$$x^2 - 4x + 4 + y^2 - 4y + 4 = 7; \quad (158)$$

$$(x - 2)^2 + (y - 2)^2 = (\sqrt{7})^2; \quad (159)$$

que é o círculo de raio  $\sqrt{7}$  com centro no ponto  $(2, 2)$ . Como esta equação não é radial, sua expressão em coordenadas polares não oferece nenhuma vantagem.

- Um outro exemplo, a equação

$$\rho = f(\rho, \theta) = \frac{2}{1 - \cos(\theta)} \quad (160)$$

que é difícil de interpretar com coordenadas polares, mas facilmente podemos transformar em uma equação cartesiana:

$$\rho = \frac{2}{1-\cos(\theta)} \equiv \rho - \rho \cos(\theta) = 2; \quad (161)$$

$$\sqrt{x^2 + y^2} - x = 2 \equiv x^2 + y^2 = (x+2)^2; \quad (162)$$

$$y^2 = 4x + 4 = 4(x+1) \equiv x = \frac{y^2}{4} - 1 = f(y); \quad (163)$$

$$y = \pm 4 \Rightarrow x = 3 \Rightarrow (3, \pm 4) \in \{(x, y); (f(y), x)\}; \quad (164)$$

$$2ydy = 4dx; A(x-a) = B(y-b); -B(x-a) = A(y-b); \quad (165)$$

$$(3, 4) \rightarrow 8 * (y-4) = 4(x-3); 4(y-4) = -8(x-3); \quad (166)$$

$$(3, -4) \rightarrow -8(y+4) = 4(x-3); 4(y+4) = 8(x-3); \quad (167)$$

Na equação (eq. 164) calculei dois pontos sobre a curva. Nas equações (eq. 165)- (eq. 166) calculei as diferenciais da equação da parábola em dois pontos, (3, 4), (3, -4), as equações das retas tangentes e de suas perpendiculares, nestes pontos. Por definição, na parábola, as retas perpendiculares em qualquer ponto, passam pelo foco, então duas tais retas identificam o o foco, a interseção das retas perpendiculares é o foco da parábola:

$$\begin{cases} 4(y+4) = 8(x-3); \\ 4(y-4) = -8(x-3); \end{cases} \quad (168)$$

$$\begin{cases} 32 = 16x - 48; \Rightarrow x = 80/16 = 5 \\ x = 5 \Rightarrow 4y + 16 = 16 \Rightarrow y = 0; \\ (5, 0) \text{ é o foco da parábola}; \end{cases} \quad (169)$$

- A equação da reta com coeficiente angular  $m$  e passando pelo ponto  $(a, b)$  é  $y = m(x - a) + b$  então

$$y = m(x - a) + b \Rightarrow \rho \sin(t) = m(\rho \cos(t) - a) + b; \quad (170)$$

$$\rho = \frac{b - ma}{\sin(t) - m \cos(t)}; \quad (171)$$

$$\text{para gnuplot: } \rho(t) = (b - m * a) / (\sin(t) - m * \cos(t)); \quad (172)$$

O código, em `gnuplot`

```
set polar
m = -1; a = 1; b = 1;
r(t) = (b - m*a)/(sin(t) - m*cos(t));
plot r(t)
```

produz o gráfico de  $y = m(x - a) + b$ .

Se uma reta passar pela origem fazendo o ângulo  $\alpha$  com o eixo  $OX$  então é um lugar geométrico radial, qualquer ponto sobre esta reta é função do raio mas o método descrito nas equações (eq. 170) -(eq. 172) não funciona porque

$$b = a = 0 \Rightarrow b - m * a = 0; \quad (173)$$

- Uma reta que passe pela origem é

$$\begin{cases} y = mx; \\ \{(r, \theta); \theta = \text{Atan}(m)\}; \end{cases} \quad (174)$$

Seria a equação em coordenadas polares da reta  $y = mx$ , mas não tem como traduzir esta equação para fazer o gráfico com `gnuplot` uma vez que este motor espera uma equação no formato  $r = f(\theta)$  que é impossível tirar do sistema de equações (eq. 174).

Gráficos bonitos podem ser obtidos com coordenadas polares, por exemplo  $r(t) = t$  em que o raio é igual ao ângulo, é uma espiral, experimente com `gnuplot`, mas não esqueça do comando crucial: `set polar`; . Depois tente os gráficos

$$\text{pow}(t, n) = t * n; \# \text{ para definir potência}; \quad (175)$$

$$r(t) = \text{pow}(t, 2); \quad (176)$$

$$r(t) = \text{pow}(t, 3); \quad (177)$$

$$r(t) = \sin(\text{pow}(t, 2)); \quad (178)$$

Estes exemplos mostram a fragilidade do sistema de coordenadas polares para ser usado *genericamente*. Se passarmos para derivadas então as coordenadas polares podem oferecer cálculos simplesmente estrambóticos um bom exemplo é a representação do Laplaciano em coordenadas polares. . . e o objetivo que temos com Matemática absolutamente não é este e sim, usar um sistema de representação que seja o mais adequado para atingir a simplicidade, quando isto for possível.

O exemplo do círculo de centro na origem e raio  $R$  cuja equação em coordenadas polares é  $r = R$  mostra que há casos em que vale a pena usar coordenadas polares, como é o caso do próximo exemplo no cálculo de integrais de sólidos de revolução.

Vou dar mais outro exemplo mostrando como algumas vezes é importante usar coordenadas polares, é o cálculo duma integral que é facilmente obtido com a chamada de *mudança de variáveis* na integral. A integral da gaussiana, sendo o exemplo clássico do cálculo da integral de  $f(x) = e^{-x^2}$ , passando pela integral de  $F(x, y) = e^{-x^2-y^2}$  que é uma função radial.

$$R = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2-y^2} dx dy = \int_{-\infty}^{\infty} e^{-x^2} dx \int_{-\infty}^{\infty} e^{-y^2} dy = \left( \int_{-\infty}^{\infty} e^{-t^2} dt \right)^2; \quad (179)$$

$$R = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2-y^2} dx dy = \int_0^{\infty} \int_0^{2\pi} e^{-r^2} \det(J(r, \theta)) dr d\theta; \quad (180)$$

$$\begin{pmatrix} dx & dy \end{pmatrix} = J(r, \theta) \begin{pmatrix} dr \\ d\theta \end{pmatrix}; \quad (181)$$

$$J(r, \theta) = \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & \rho \cos(\theta) \end{pmatrix}; \det(J(r, \theta)) = \rho; \quad (182)$$

$$R = \int_0^{\infty} \int_0^{2\pi} e^{-r^2} \det(J(r, \theta)) dr d\theta = \int_0^{\infty} \int_0^{2\pi} e^{-r^2} r dr d\theta = 2\pi \int_0^{\infty} e^{-r^2} \rho dr; \quad (183)$$

$$R = \pi \int_0^{\infty} e^{-r^2} 2r dr = -\pi e^{-r^2} \Big|_0^{\infty} = \pi; \quad (184)$$

$$\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{R} = \sqrt{\pi}; \quad (185)$$

Os sistema de equações na (eq. 182) representam transformações, não lineares, do plano para o plano que é diferenciável fora da origem que é um ponto crítico para as coordenadas polares. É interessante registrar a jacobiana destas transformadas que podem ser usadas para simplificar integrais.

Esta é a grande aplicação das coordenadas polares, no cálculo de volumes de sólidos de revolução quando o cálculo se processa de forma semelhante ao da sequência nas equações (eq. 179)-(eq. 185).

A justificativa da passagem em que aparece o determinante da jacobiana da transformada de coordenadas é que o determinante é o fator de correção local da área entre a região de saída e a de chegada pela transformação *mudança de coordenadas*.

---

- **cosenos diretores** confira *equação polar da reta*. Os *cosenos diretores* são os nomes dados às coordenadas dum ponto sobre a esfera de dimensão  $n - 1$ ,  $\mathbf{S}^{n-1}$ .

$$X \in \mathbf{S}^{n-1}; X = (\cos(\alpha_1), \cos(\alpha_2), \dots, \cos(\alpha_n)); \quad (186)$$

Eles satisfazem a relação fundamental da trigonometria,

$$\sum_{k=1}^n \cos(\alpha_k)^2 = 1; \quad (187)$$

---

- **curva** é uma variedade de dimensão 1. Uma reta é uma curva, assim com qualquer outras das cônicas de dimensão 1: parábola, círculo, elipse ou hipérbole. Um ponto é uma variedade de dimensão zero, e um plano é uma variedade de dimensão 2. Uma outra forma de definir curva se dá por parametrização quando se entende uma curva como uma função definida em um intervalo da reta e tomando valores em um espaço vetorial:

$$\alpha : [a, b] \rightarrow \mathbf{R}^n; \alpha(s) = (x_1(s), \dots, x_n(s)); \quad (188)$$

O traço de  $\alpha$  é uma variedade de dimensão 1 contida na variedade linear de dimensão  $n$ ,  $\mathbf{R}^n$ .

Entretanto as curvas podem ser objetos bastante complicados, como as curvas que preenchem o espaço (space filling curves), você pode encontrar um exemplo deste tipo de curva em [9], a curva de Peano que passa por todos os pontos de coordenadas racionais de um retângulo. Uma tal curva não é diferenciável. Para a grande maioria dos teoremas do Cálculo se impõe a condição de diferenciabilidade sobre as curvas para eliminar os problemas que uma curva genérica poderia criar.

Um exemplo simples de curva é o gráfico de uma função diferenciável.

$$[a, b] \xrightarrow{f} \mathbf{R}; x \mapsto f(x) \in \mathbf{R} \quad (189)$$

Retirando o adjetivo “diferenciável” o exemplo pode ficar bem complicado. Este exemplo pode facilmente se acomodar à notação da equação (188) se escrevermos

$$(s, f(s)) = (x_1(s), x_2(s)) \quad (190)$$

---

- **curva algébrica** é uma variedade de dimensão 1 definida como os zeros dum polinômio em duas variáveis com coeficientes sobre um corpo  $\mathbf{K}$ . Por exemplo, se o corpo  $\mathbf{K} = \mathbf{R}$  e sendo

$$F(x, y) = \sum_{k=0}^n c_k x^{n-k} y^k; c_k \in \mathbf{R}; \quad (191)$$

$$F(x, y) = 0 \text{ é uma curva algébrica;} \quad (192)$$

Todas as cônicas da *geometria analítica* são curvas algébricas sobre o corpo  $\mathbf{R}$  e a *geometria analítica* é o estudo de um tipo particular de curvas algébricas sobre o corpo dos reais.

Como  $F(x, y)$  definido pela equação (eq. 191) é derivável, então a derivada implícita

$$\frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy \quad (193)$$

e mais um ponto  $(a, b); F(a, b) = 0$ , um ponto sobre a curva algébrica  $F(x, y) = 0$ , produzem uma variedade tangente à curva algébrica no ponto  $(a, b)$  que é dada por

$$\frac{\partial F}{\partial x}|_{(a,b)}(x - a) + \frac{\partial F}{\partial y}|_{(a,b)}(y - b) \quad (194)$$

Deduzi a equação (eq. 194) da equação (eq. 191) sob a hipótese de que a solução  $(a, b); F(a, b) = 0$  exista e a equação (eq. 194) é um polinômio do primeiro grau em  $x, y$  o que define uma reta.

Este cálculo pode ser considerado uma demonstração de que as curvas algébricas, como definidas acima, são variedades de dimensão 1 porque suas variedades tangentes são retas. Isto porque a *tangência* é uma relação de equivalência, então se a tangente num ponto de  $F(x, y) = 0$  for uma variedade de dimensão 1, então  $F(x, y) = 0$  é uma variedade com a mesma dimensão 1.

- **curva de Jordan** é um tipo de curva fechada, diferenciável, que não se corta a si mesma. Se define o índice de um ponto relativamente a uma curva como o número de vezes que a curva circula em torno do ponto. Em variáveis complexas isto pode ser obtido como o valor da integral

$$Ind_{\gamma}(a) = \frac{1}{2\pi i} \int_{\gamma} \frac{1}{z - a}; \quad (195)$$

Se a curva  $\gamma$  for de Jordan, a integral na equação (195) somente poderá valer 0 ou  $n \neq 0$ . Se valer  $n \neq 0$  então o ponto  $a$  é um ponto interior de  $\gamma$  e valerá 0 se  $a$  for um ponto exterior, nada se podendo dizer se  $a$  estiver sobre a curva. Observe a lógica do nome da equação (195), ela expressa o número de voltas que a curva  $\gamma$  der em volta do ponto  $a$ , em inglês *winding number*, número de voltas. Esta integral é estudada em variáveis complexas e faz parte da fórmula *integral de Cauchy*.

A fronteira de um retângulo ainda é uma curva de Jordan porque se estende o conceito às curvas que não sejam diferenciáveis em um número finito de pontos. O cálculo da integral na equação (195) pode ser facilmente feito neste caso, do retângulo ou de um quadrado com lados paralelos aos eixos. No gráfico (29) há pontos cujos índices relativamente a  $\gamma$  é dois, portanto não se trata de uma curva de Jordan. As curvas de Jordan dividem o plano em duas regiões disjuntas chamadas de interior e exterior da curva. É um teorema, que levou anos para ser demonstrado, que leva o nome de Jordan, estabelece que se ligarmos um

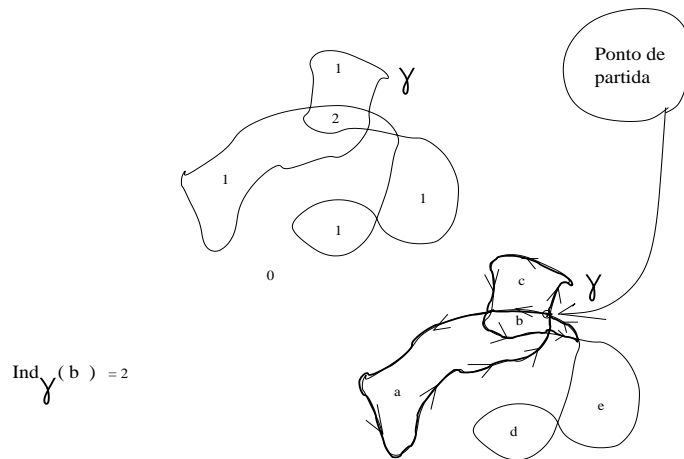


Figura 29: Um curva que não é de Jordan

ponto do interior de uma curva de Jordan,  $\gamma$  com um ponto do exterior da mesma curva, por meio de uma curva  $\beta$  então haverá um ponto do plano que será comum às duas curvas  $\gamma$  e  $\beta$ , elas se interceptam. Uma reta é uma curva de Jordan! resta apenas a decisão da escolha da região interior, é um ponto de vista!

- **derivada** Se uma função real  $y = f(x)$  de variável real tiver retas tangentes ao seu gráfico numa vizinhança do ponto  $a$ , então dizemos que  $f$  é diferenciável em  $a$  e o coeficiente angular da reta tangente ao ponto  $(a, f(a))$ , designado com o símbolo  $f'(a)$ , é a *derivada* de  $f$  no ponto  $a$ . A função  $y = f'(x)$  que fornece o coeficiente angular da reta tangente ao gráfico de  $f$  no ponto  $(x, f(x))$  se chama derivada de  $f$ .

**gnuplot** é um programa para fazer gráficos e ele entende um pouco de Matemática, podemos fazer algumas operações algébricas com ele. Os gráficos na figura (fig 30), página 79, foram feitos com **gnuplot** para ilustrar a frase anterior. Foi este o programa escrito em **gnuplot** que eu rodei:

```
pow(x,n) = x**n; ## uma expressão melhor para potencia
f(x) = pow(x,2) - x - 6; ## a equação da função
df(x) = 2*x - 1; ## a derivada atualize se alterar f
a1 = -3; A1 = df(a1); B1 = f(a1); ## o valor de a, A, B - troque apenas "a"
a2 = 3; A2 = df(a2); B2 = f(a2);
r1(x) = A1*(x-a1) + B1 ## uma reta tangente
r2(x) = A2*(x-a2) + B2 ## outra reta tangente
plot f(x), r1(x), 0, r2(x);
```

Se você souber calcular a derivada de  $f$  você pode fazer experiências com este programinha do **gnuplot**. O objetivo lhe mostrando logo este programa é o de traduzir o poder da derivada. Se você souber calcular a derivada de uma



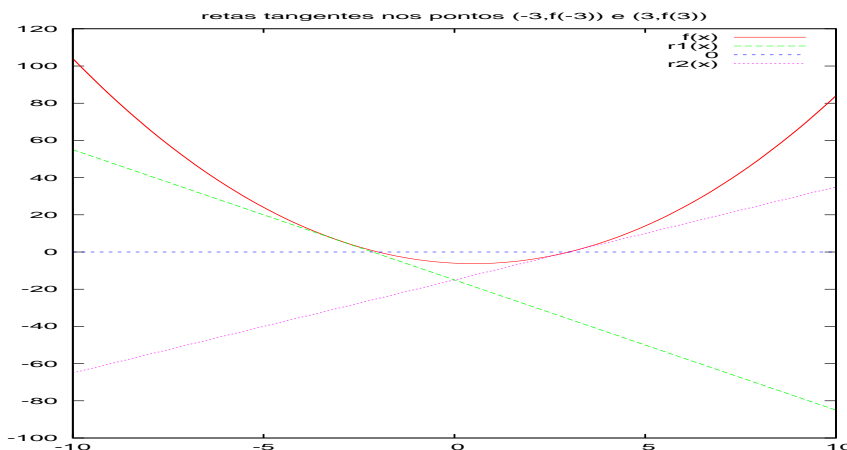


Figura 30: tangentes ao gráfico de  $f(x) = x^2 - x - 6$

função então você tem o poder de colocar uma reta tangente ao gráfico de  $f$  em qualquer ponto que você desejar.

Se você souber calcular a derivada de uma função pode imediatamente usar este programa para fazer gráficos de funções. Raspe, cole, e se divirta!

Observe que  $r_1(x)$  é a equação de uma reta apenas os coeficientes  $A_1, B_1$  foram calculados usando  $f$ , o coeficiente angular da reta foi calculado usando a derivada de  $f$ . Este é o significado da derivada:  $f'(a)$  é coeficiente angular instantâneo do gráfico de  $f$  no ponto  $(a, f(a))$ .

A equação da reta tangente ao gráfico de  $f$  no ponto  $(a, f(a))$  se deduz direto da equação da reta que passa no ponto  $(a, f(a))$  e tem coeficiente angular  $m$

$$y = P(x) = b + m(x - a); (a, b) = (a, f(a)); m = f'(a); \quad (196)$$

$$P(x) = f(a) + f'(a)(x - a); \quad (197)$$

que é a expressão do *polinômio de Taylor* do primeiro grau desenvolvido no ponto  $x=a$ .

Para que você entenda melhor esta questão do *coeficiente angular instantâneo* um exemplo da Física é bem instrutivo. Considere a seguinte situação, representada no gráfico da figura (fig 31), página 80. Neste gráfico estou simulando o que aconteceria se eu estivesse rodando uma pedra presa a um cordão que, num determinado instante, se quebrasse. A pedra, enquanto presa à minha mão pelo cordão, em movimento, muda de direção a cada instante. Mas quando o cordão se quebra ela se mantém numa direção fixa, a da tangente. Obviamente que em seguida ela vai mudar de direção porque passa a ficar presa à Terra pela força de gravidade. Mas, se não houvesse a *força de gravidade*, então, sim, ela seguiria por uma reta - *movimento uniforme não acelerado* como diz a Física.

A verdade é que, em todo o Universo, não existe um só corpo em *movimento uniforme não acelerado*, todos os corpos sofrem a ação da gravidade dos outros

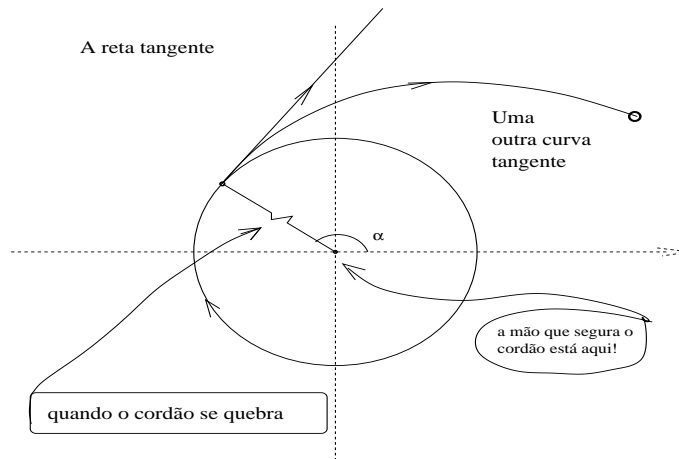


Figura 31: o cordão se quebra, a reta sai pela tangente

corpos no Universo e assim em todo o Universo não existe uma única reta... mas na Geometria Euclidiana existem retas e disso que estamos falando para discutir a *derivada*.

Observe que neste simples exemplo, estou mostrando que Geometria Euclidiana, que parece tão concreta, na verdade é uma grande *abstração* da mente humana.

Retornando à figura (fig 31), o coeficiente angular da reta que ali aparece, é a derivada do círculo naquele ponto. Você pode ver logo aqui um método geométrico para o cálculo de derivadas:

- coloque uma régua tangente ao gráfico de  $f$  no ponto  $(a, f(a))$ ;
- calcule o coeficiente angular,  $m$ , desta reta;
- $f'(a) = m$ .

Outra vez usando `gnuplot`, com os comandos abaixo, que você pode raspar e colar num terminal do `gnuplot`,

```
f(x) = (x+3)*(x-5)*sin(x/5.0);
df(x) = (x-5)*sin(x/5.0) + \
(x+3)*sin(x/5.0)+0.2*(x+3)*(x-5)*cos(x/5.0);
P(x) = f(a) + df(a)*(x-a)
a = 4;
plot f(x),P(x),0
```

você pode repetir a figura (32) página 81, em que está representada a reta tangente ao gráfico de  $f(x) = (x + 3)(x - 5) \sin(x/5.0)$  no ponto  $(4, f(4))$ .

Apenas trocando valor de a você pode obter gráficos de outras retas tangentes ao gráfico desta mesma função ou outra de sua escolha (redefina  $f$  e  $df$

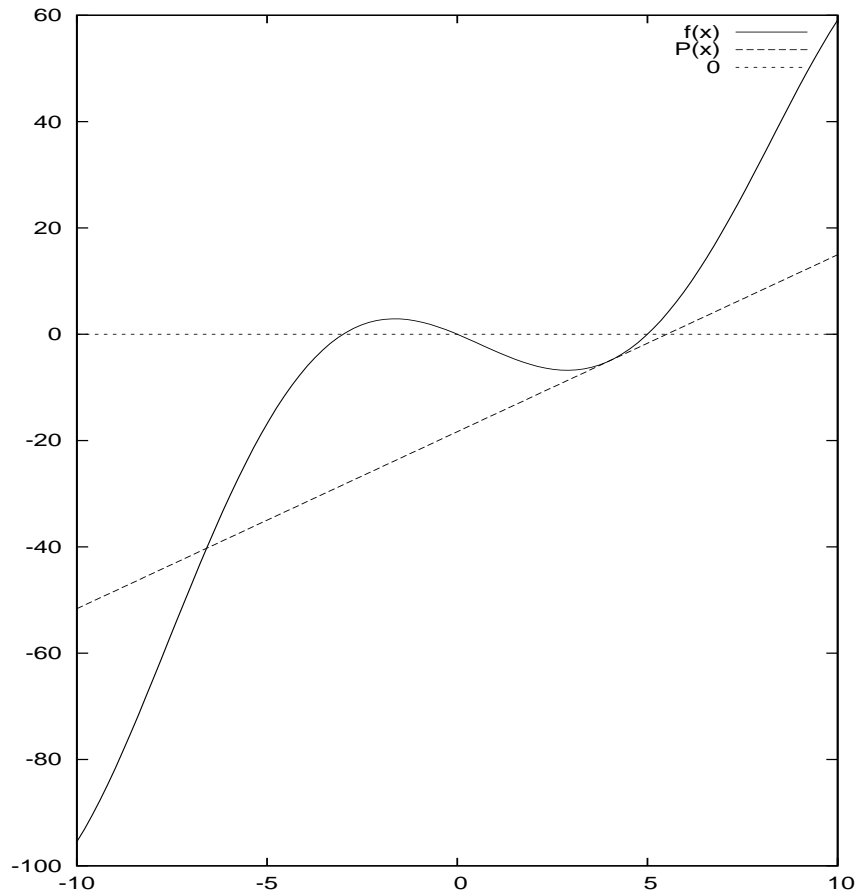


Figura 32: Reta tangente ao gráfico de uma função - derivada

no programa), escolha distintos valores para  $a$  e repita o comando `plot` para ver tangentes em diversos pontos do gráfico. Por exemplo, raspe e cole num terminal do gnuplot este programa:

```
f(x) = (x+3)*(x-5)*sin(x/5.0);
df(x) = (x-5)*sin(x/5.0) + \
(x+3)*sin(x/5.0)+0.2*(x+3)*(x-5)*cos(x/5.0);
P(x) = f(a) + df(a)*(x-a)
a = -4;
plot f(x),P(x),0 ;
pause -2 "Aperte enter para continuar"
a = -2;
plot f(x),P(x),0;
pause -2 "Aperte enter para continuar"
```

```

a= 0;
plot f(x),P(x),0 ;
pause -2 "Aperte enter para continuar"
a= 1;
plot f(x),P(x),0 ;
pause -2 "Aperte enter para continuar"
a= 2;
plot f(x),P(x),0 ;
pause -2 "Aperte enter para continuar"
a= 3;
plot f(x),P(x),0 ;
pause -2 "Aperte enter para continuar"
a= 5;
plot f(x),P(x),0 ;
pause -2 "Aperte enter para continuar"

```

e você vai ver sucessivos gráficos de retas tangentes ao gráfico da função  $y = f(x)$

Até este momento adotei um ponto de vista que pode ser constrangedor para você, leitora. Estou usando a derivada, e supondo que você já sabe derivar, mas eu lhe pedi que aceitasse esta forma de comunicação com paciência, e agora vou mostrar-lhe o caminho para aprender a derivar. Vou ser resumido do contrário eu iria escrever um capítulo do livro de Cálculo, e não é este o objetivo aqui. Você pode ler o livro de Cálculo na biblioteca, aqui estou apenas tentando estimulá-la a fazer isto.

#### A definição da derivada para funções reais

A derivada mede uma taxa de variação instantânea, portanto um *cálculo de limite*. Mas deixe-me começar do começo, taxa de variação quer dizer um quociente de diferenças que você pode ver indicado no gráfico na figura (fig 33), página 83. O quociente entre duas diferenças:

$$\Delta_a(f) = \frac{f(a + \Delta x) - f(a)}{\Delta x} \quad (198)$$

é a tangente do ângulo  $\alpha$  no triângulo retângulo que você pode ver na figura (fig 33).

Mas eu quero o coeficiente angular da reta tangente no ponto  $(a, f(a))$  que é  $f'(a)$ . Se a curva  $graf(f)$  fosse a trajetória de uma nave que fosse liberar no ponto  $(a, f(a))$  um foguete, este foguete seguiria pela reta tangente com coeficiente angular  $f'(a)$ . É uma *situação semelhante* a da pedra presa ao cordão que se quebra na figura (fig 31).

Eu vou agora mostrar-lhe o cálculo deste limite.

Dada uma função  $y = f(x)$  podemos aplicar-lhe o operador diferença caracterizado pelo símbolo  $\Delta$ , seguido do operador *quociente* para o qual não há um símbolo padrão. A seqüência de operações é esta:

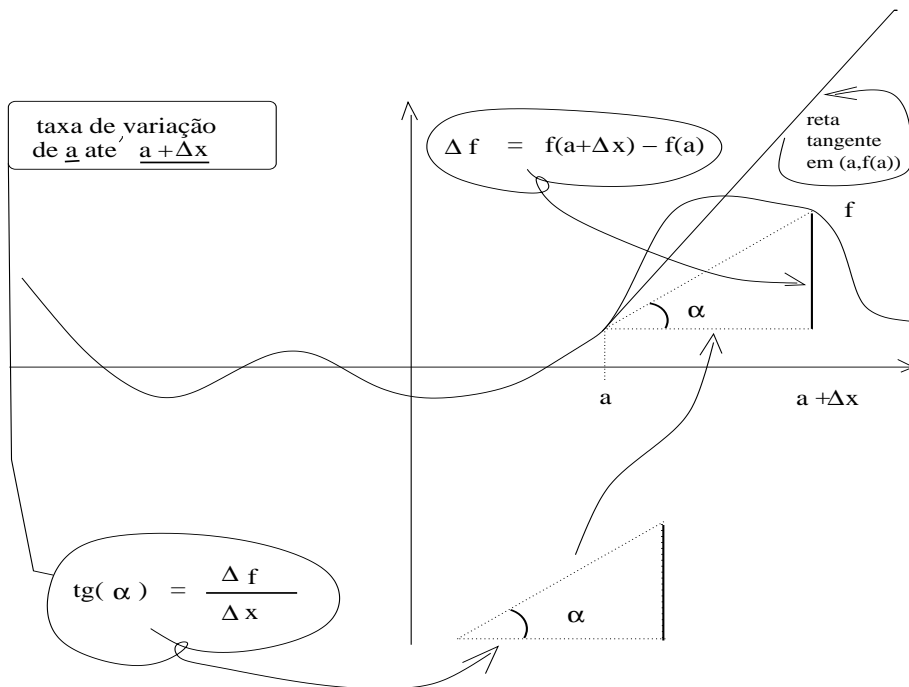


Figura 33: taxa de variação de  $f$  no intervalo  $[a, a + \Delta x]$

$$\Delta_{a,h}(f) = f(a + h) - f(a) \quad (199)$$

$$Q_{a,h}(f) = \frac{\Delta_a(f)}{h} \quad (200)$$

$$\lim_{h \rightarrow 0} Q_{a,h}(f) = f'(a) \quad (201)$$

Se soubermos o significado da última linha, calculamos a derivada. Para “decifrar” a última linha algumas vezes é preciso muito prática. Entretanto, salvo alguns casos muito especiais que inclusive são chamados de *limites notáveis* a grande maioria dos casos simples do dia-a-dia são resolvidos com as regras de derivação descritas abaixo e mais um banco de derivadas conhecidas, e *alguma prática de cálculo...*

### Um exemplo simples

A derivada da função  $f(x) = x^3$  é  $f'(x) = 3x^2$  e vou calcular esta derivada usando quociente de diferenças seguido do limite. Acompanhe a sequência de operações para a qual vou fazer uma legenda em seguida, e você pode saltar da legenda para a sequência de operações para procurar entender as passagens.

$$f(x) = x^3; \Delta_{a,h}(f) = f(a+h) - f(a) = 3a^2h + 3ah^2 + h^3; \quad (202)$$

$$Q_{a,h}(f) = \frac{\Delta_{a,h}(f)}{h} = 3a^2 + 3ah + h^2; \quad (203)$$

$$\lim_{h=0} Q_{a,h}(f) = 3a^2 = f'(a) \quad (204)$$

- Na equação (eq. 202) calculei a diferença com auxílio do triângulo de Pascal, confira no *triângulo* a linha de ordem 3, cancelei o termo em  $a^3$  quando subtraí  $f(a)$ ;
- Na equação (eq. 203) calculei o quociente de  $f$  no ponto  $a$  com a diferença  $h$ . No quociente  $Q_{a,h}(f)$  as potências de  $h$  caíram de uma unidade, ficando um termo livre de  $h$ .
- Na equação (eq. 204) calculei o limite de  $Q_{a,h}(f)$  quando  $h = 0$  que neste caso é simples, os termos que tem  $h$  se anulam. Resta o termo que não tem  $h$  com a conclusão  $f'(a) = 3a^2$ .

Para todas as funções polinomiais é assim simples, o cálculo vai ter mais termos e ficar um pouco mais complicado de fazer a redação, mas tem alguns truques redacionais que vou mostrar-lhe.

Considere apenas  $f(x) = x^n$  e confira que é semelhante ao caso acima. A diferença é calculada com a linha de ordem  $n$  do triângulo de Pascal que começa com  $a^n$  seguido de  $\binom{n}{n-1}a^{n-1}h$  e termina com  $h^n$ . Coloque reticências no meio que é tudo “irrelevante”. Se convença!

Quando calcular  $\Delta_{a,h}(f)$  vai se cancelar o termo de maior grau, em  $a$  ficando de  $\binom{n}{n-1}a^{n-1}h$  e terminando com  $h^n$ . Coloque reticências no meio que é tudo “irrelevante”. Se convença!

$$\binom{n}{n-1}a^{n-1}h = na^{n-1}h$$

Quando calcularmos o quociente  $Q_{a,h}(f)$  caem de uma unidade todas as potências de  $h$  ficando de  $na^{n-1}$  e terminando com  $h^{n-1}$ . Coloque reticências no meio que é tudo “irrelevante”. Se convença!

Todos os termos, exceto o primeiro,  $na^{n-1}$  se anulam. Conclusão:

$$f(a) = a^n \implies f'(a) = na^{n-1}$$

que é a regra de derivação da enésima potência.

Logo você vai ver as regras de derivação, abaixo, e como calcular a derivada de um polinômio qualquer.

### O cálculo da derivada

A derivada não é uma operação aritmética, ela é o resultado da aplicação do *operador limite* a uma sucessão de *quocientes de diferenças* isto torna pouco provável que se consiga implementar a derivação em *Computação Algébrica* ou seja “calcular a derivada automaticamente com um programa de computador”.

Você pode ver facilmente que procede o que foi dito no parágrafo anterior analisando as contas que fiz para determinar a derivada de  $f(x) = x^n$ . Observe a passagem da equação (eq. 205) para a equação (eq. 206). Não foi uma passagem algébrica, houve um salto lógico que se traduziu na frase “todos termos que contém  $h$  se anulam exceto o primeiro”. Este raciocínio não é algébrico e nós apenas sabemos dizer que aplicamos o operador limite  $\lim_{h=0}$ . Sabemos fazer este cálculo, mas não sabemos traduzí-lo com um algoritmo o que torna impossível, no momento pelo menos, traduzir esta passagem para um programa de computador.

Ainda assim os programas de *Computação Algébrica* conseguem calcular derivadas de forma mais efetiva que o humano ao aplicar as regras do Cálculo para diferenciação que se podem resumir nas seguintes:

- A derivada de uma função constante é zero.
- Se uma função for linear ela é a sua própria função linear tangente portanto a derivada de uma função linear é ela mesma;
- derivada da soma Se  $f$  e  $g$  forem deriváveis, então  $f + g$  é derivável e

$$(f + g)' = f' + g'.$$

- derivada do produto que se  $f, g$  forem duas funções diferenciáveis, então

$$(fg)' = f'g + fg'$$

- derivada de uma função multiplicada por um número está contida na primeira regra e na regra da derivada do produto, mas merece destaque. Se  $r$  for um número, e se  $f$  for derivável, então  $(rf)' = rf'$ . Aplique a derivada do produto e você verá que obtém esta regra.
- derivada de funções polinômiais a derivada de  $f(x) = x^n$  é  $f'(x) = n * x^n$  para funções reais de variável real (ou complexa);. É uma aplicação direta do operador diferença ao monômio  $f(x) = x^n$  seguido do cálculo do limite.
- Para um polinômio qualquer se aplica a regra da soma de derivadas, e para cada um dos monômios da soma,  $a_k x^k$  se aplica a regras derivada da multiplicação pelo número real  $a_k$ .
- derivada da  $\frac{1}{f}$

$$\frac{\frac{1}{f(a+\Delta x)} - \frac{1}{f(a)}}{\Delta x} = \frac{f(a) - f(a + \Delta x)}{\Delta x f(a + \Delta x) f(a)} = \frac{\frac{f(a) - f(a + \Delta x)}{\Delta x}}{f(a + \Delta x) f(a)} \quad (205)$$

na última equação tanto o numerador como denominador têm limite e o limite do denominador é diferente de zero então podemos aplicar o operador limite  $\lim \Delta x = 0$  tendo por resultado

$$\left(\frac{1}{f}\right)' = \frac{f'(a)}{f(a)^2} \quad (206)$$

- derivada do quociente quando  $g(x)$  não se anular no ponto  $a$ , então numa vizinhança de  $a$

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$

- a regra da cadeia, a derivada da função composta, que se  $f, g$  forem duas funções diferenciáveis e se a composta  $f(g(x))$  existir então

$$(f(g(x)))' = f'(g(x))g'(x)$$

A regra da cadeia se aplica *ipsis literis* em qualquer dimensão em que as compostas estejam definidas.

Estas regras junto com *um banco de derivadas conhecidas* permitem que os programas de Computação Algébrica calculem derivadas de forma muito mais efetiva que o ser humano sugerindo a existência de *inteligência artificial*.

Uma alternativa à Computação Algébrica é a diferenciação algorítmica que tem conseguido alguns avanços, mais ainda não se pode comparar com as possibilidades da Computação Algébrica, e como esta, esbarra no salto lógico entre operações aritméticas e operador limite.

### Dois exemplos, um deles difícil

Mostro-lhe um exemplo, aliás, dois:

- Uma derivada difícil:  $f(x) = \sin(x)$ ;  $f'(x) = \cos(x)$ ; Esta derivada é obtida uma uma desigualdade geométrica e mais algumas propriedades do limite.
- Mas  $\cos(x) = \sin(x - \frac{\pi}{2})$ ;
- $g(x) = \cos(x)$ ; então podemos calcular  $g'(x)$  usando uma propriedade da trigonometria, junto com a derivada do sin.

$$g'(x) = \lim_{h \rightarrow 0} Q_{x,h}(g); \quad (207)$$

$$Q_{x,h}(g) = \frac{\cos(x+h) - \cos(x)}{h} = \frac{\sin(x+h - \frac{\pi}{2}) - \sin(x - \frac{\pi}{2})}{h}; \quad (208)$$

$$Q_{x,h}(g) = Q_{x - \frac{\pi}{2}, h}(f); \quad (209)$$

$$g'(x) = f'(x - \frac{\pi}{2}) = \cos(x - \frac{\pi}{2}) = \sin(x); \quad (210)$$

$$\cos(x)' = \sin(x); \quad (211)$$

Conclusão, há algumas derivadas que são difíceis de serem calculadas, como é o caso da derivada do  $\sin(x)$ , mas na medida em que aumentarmos o banco de dados de derivadas conhecidas, e ampliarmos a lista de regras com novas propriedades, conseguimos ampliar muito rapidamente o banco de dados. Esta é uma breve descrição da metodologia com que os programas de *Computação Algébrica* calculam derivadas e conseqüentemente também integrais, ou seja apenas eles representam uma automatização dos métodos que nós, os humanos, usamos para calcular derivadas, e conseguem fazer o trabalho com muito mais rapidez e sem os erros que os humanos frequentemente cometem...



### Que é mesmo a derivada?

A comparação da derivada de funções univariadas com a jacobiana das funções multivariadas levou a uma generalização do conceito de derivada. Observe como isto foi feito, obviamente, olhando depois que tudo aconteceu... Quem nos dá o fio da meada é a derivada implícita. Se  $z = F(x, y)$  for diferenciável então podemos obter, derivando implicitamente:

$$dz = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy = \quad (212)$$

$$dz = \begin{pmatrix} \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} \quad (213)$$

Na equação (eq. 213) deduzi da equação (eq. 212) um produto de matrizes em que aparecem as estranhas variáveis  $dx, dy, dz$  nas quais não há nem  $x$  nem  $y$  e nem  $z$ . Isto é uma outra história que criou um tremendo drama na Matemática numa tentativa de explicar o que seriam os *infinitesimos* com que os matemáticos até o início do século 20 identificavam os símbolos  $dx, dy, dz$ .

Para entender como o drama foi resolvido, vou derivar implicitamente, coisa que ninguém faz, uma função univariada:  $y - f(x) = 0$ . Para fazê-lo escrevi a função univariada de forma implícita...

$$y - f(x) = 0 \Rightarrow dy - f'(x)dx \Rightarrow dy = f'(x)dx; \quad (214)$$

Se considerarmos o ponto  $(a, f(a))$  do gráfico da função diferenciável  $y = f(x)$  podemos obter da equação (eq. 214) a equação da reta tangente ao gráfico da função no ponto  $(a, f(a)) = (a, b)$

$$(a, f(a)) = (a, b); \quad (215)$$

$$dx := (x - a); dy := (y - b); \quad (216)$$

$$dy = f'(x)dx \Rightarrow y - b = f'(a)(x - a); \quad (217)$$

um novo produto de matrizes na equação (eq. 217). Apenas agora uma matriz  $1 \times 1$  que se identifica com um número. Matrizes são identificadas com funções lineares e  $f'(a)$  é o coeficiente angular da reta tangente ao gráfico de  $y = f(x)$  no ponto  $(a, f(a))$ .

Com isto resolvemos o significado das variáveis estranhas, agora  $dx, dy$  no caso univariado. Quando derivamos encontramos o modelo para o objeto linear tangente, não o próprio objeto linear tangente. Então a derivada é uma função linear que serve como modelo para o objeto linear tangente. Observe o gráfico na figura (fig 34), página 88.

Podíamos ter feito dois gráficos, um num espaço identificado pelas variáveis  $dx, dy$  e outro onde se encontra o gráfico de  $y = f(x)$ . Simplesmente não precisamos de infinitesimais,  $dx, dy$  são os nomes de duas variáveis em relação às quais representamos a função linear tangente em parte devido a um defeito de linguagem e comunicação porque é complicado falar da função linear  $f'(a)$  e mais natural falar da função linear que associa à variável  $dx$  a sua imagem  $f'(a)dx = dy$ . Então, como

$$f : \mathbf{R} \longrightarrow \mathbf{R}; x \xrightarrow{f} f(x) \in \mathbf{R}; f' : \mathbf{R} \longrightarrow \mathcal{L}(\mathbf{R}, \mathbf{R}); x \xrightarrow{f'} f'(x)\mathcal{L}(\mathbf{R}, \mathbf{R}) \quad (218)$$

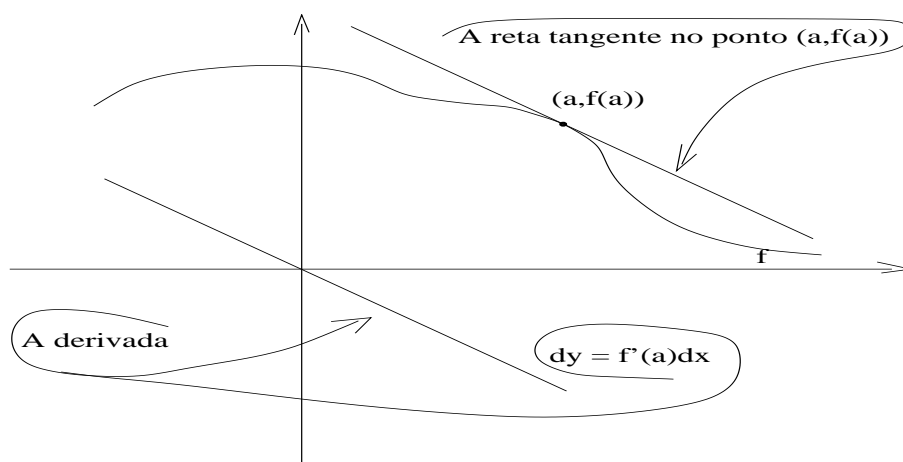


Figura 34: a reta tangente e o gráfico da derivada

a derivada é uma função que associa a cada ponto do domínio uma função linear cujo gráfico é paralelo ao gráfico da reta tangente, no caso univariado ou em qualquer outra dimensão finita.

Esta formulação está pronta para a generalização para espaços de dimensão não finita. O Cálculo que fazemos com espaços de dimensão finita pode ser generalizado para os espaços abstratos, no caso dos espaços vetoriais normados praticamente podemos repetir as relações existentes nos livros de Cálculo evitando as questões dimensionais, e Henri Cartan escreveu, na década de 60, um livrinho intitulado *Calcul différentiel dans les espaces de Banach*.

É isto que é jacobiana de uma função multivariada, a função linear que serve de modelo para a variedade linear tangente ao gráfico de uma função diferenciável em cada um dos pontos de seu gráfico. O nome jacobiana surgiu pela falta de claro entendimento do que seria o conjunto de derivadas parciais que aparecem quando derivamos uma função multivariada e que a derivada implícita, claramente, mostra que se trata de uma função linear sendo a jacobiana a matriz da mesma.

E os infinitésimos? como muita coisa em ciência, foi trabalho perdido e muita pesquisa feita para justificá-los até mesmo sendo criada uma teoria complicada de expansão dos números reais onde eles aparecem.

- **derivada complexa** O conjunto dos números complexos tem as mesmas propriedades que o conjunto dos números reais (exceto a ordem) e assim é um *corpo*. Desta forma podemos aplicar a definição de derivada usual das *funções reais de variável real* às *funções complexas de variável complexa* que é o que se costuma chamar de *derivada complexa*, e neste momento surge um dos resultados mais intrigantes da análise: *se uma função complexa de variável complexa tiver derivada complexa ela será infinitamente diferenciável*. São as *funções analíticas*, as *funções complexas que têm derivada complexa*.

Uma forma simples de chegarmos a este resultado pode ser esquematizada na seguinte sequência em que estamos usando derivação implícita para fazer aparecer as equações de *Cauchy-Riemann*, também estamos usando a dualidade de interpretação  $\mathbf{C} \xrightarrow{f} \mathbf{C}$ ,  $\mathbf{R}^2 \xrightarrow{f} \mathbf{R}^2$ , conforme for conveniente:

$$f \text{ uma função complexa de variável complexa;} \quad (219)$$

$$f = u + iv; u, v \text{ funções reais de variável complexa;} \quad (220)$$

$$J(f) = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \alpha + i\beta = f'(z) \in \mathbf{C}; \quad (221)$$

$$df = J(f) \begin{pmatrix} dx \\ dy \end{pmatrix} = (\alpha + i\beta)(dx + idy) = (\alpha dx - \beta dy) + i(\alpha dy + \beta dx) \quad (222)$$

$$df = f'(z)dz = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} \quad (223)$$

$$u_x = v_y; u_y = -v_x \quad (224)$$

A igualdade na equação (221) vem da afirmação inicial,  $\mathbf{C}$  é um corpo, como  $\mathbf{R}$ , a derivação das funções reais de variável real, se aplica verbatim ao caso complexo, portanto, como no caso real,  $f'(z) \in \mathbf{C}$ , a derivada complexa é o número complexo  $\alpha + i\beta$ .

Este fato volta a ser usado na equação (222) para identificar um tipo particular de matriz jacobiana, a derivada de  $f$ , agora vista como função de  $\mathbf{R}^2 \rightarrow \mathbf{R}^2$ , na equação (223). Vamos poder assim destacar, entre as funções  $\mathbf{R}^2 \rightarrow \mathbf{R}^2$ , uma classe particular de funções cuja matriz jacobiana tem o formato apresentado na equação (223), as funções analíticas.

A equação (ou sistema de equações diferenciais parciais de primeira ordem), equação (224), obtida quando igualamos as matrizes nas equações (221) e (223), é conhecida como *equações de Cauchy-Riemann*, e elas caracterizam quando uma função  $f = u + iv$  é analítica e são usadas com frequência como definição de função analítica.

A derivada complexa de  $f$ , se existir, é uma nova função complexa de variável complexa e se calcularmos sua derivada veremos aparecer novamente as *equações de Cauchy-Riemann*. Por indução se conclue que se  $f$  for uma função complexa, de variável complexa, então será infinitamente diferenciável se for derivável no sentido complexo.

Quer dizer que se voltarmos a olhar para as funções vetoriais de variável vetorial de dimensão dois haverá duas classes disjuntas de funções, aquelas que satisfazem às *equações de Cauchy-Riemann*, as funções analíticas, e as outras que podem ser de classe  $\mathcal{C}^\infty$  mas que não são analíticas. Por exemplo

$$g(x, y) = (x, -y) = (u(x, y), v(x, y)); g(z) = \bar{z} \quad u_x = 1 \neq v_y = -1 \quad (225)$$

não é uma função analítica mas é de classe  $\mathcal{C}^\infty$ .

Uma das implicações mais fortes da analiticidade é que se  $f$  for analítica irá transformar *abertos do plano complexo* em *abertos do plano complexo* mas não é uma propriedade fácil de ser demonstrada. Esta propriedade fundamental caracteriza as *funções analíticas* como *aplicações abertas*.

A derivada complexa de  $f$  pode ser escrita numa das formas alternativas seguintes, usando as *equações de Cauchy-Riemann*:

$$u_x + iv_x = u_x - iu_y = v_y + iv_x = v_y - iu_y \quad (226)$$

em outras palavras, o número  $f'(a + ib) = \alpha + i\beta$  pode ser obtido com um qualquer das expressões da equação (226). Se usarmos o conceito de operador diferencial, podemos deduzir das expressões na equação (226) as expressões, usando sempre a mesma notação:

$$\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(u) = \left(\frac{\partial}{\partial y} + i\frac{\partial}{\partial x}\right)(v) = (-i\frac{\partial}{\partial y} + \frac{\partial}{\partial x})(iv) = \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(iv) \quad (227)$$

$$\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(u) = \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(iv) = \alpha + i\beta = f'(a + ib) \quad (228)$$

$$\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(u) + \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(iv) = 2(\alpha + i\beta) \quad (229)$$

$$\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(u + iv) = 2(\alpha + i\beta) \quad (230)$$

$$\frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(f) = (\alpha + i\beta) = f'(a + ib) \quad (231)$$

$$\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(u) - \left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(iv) = 0 \quad (232)$$

$$\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(u - iv) = 0 \quad (233)$$

$$\overline{\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)}(u + iv) = 0 \quad (234)$$

$$\overline{\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)}(f) = 0; \left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)(f) = 0 \quad (235)$$

Destes cálculos surgiram duas expressões mais simples que se tornaram dois operadores diferenciais clássicos permitindo uma forma concisa de expressar tanto as *equações de Cauchy-Riemann* como a definição da derivada de uma função analítica:

$$\partial = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) = \frac{1}{2}\left(\frac{\partial}{\partial x} + \frac{1}{i}\frac{\partial}{\partial y}\right) \quad (236)$$

$$\bar{\partial} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right) = \frac{1}{2}\left(\frac{\partial}{\partial x} - \frac{1}{i}\frac{\partial}{\partial y}\right) \quad (237)$$

$$\partial(f) = \alpha + i\beta = f'(a + ib) \quad (238)$$

$$\bar{\partial}(f) = 0 \iff f \text{ satisfaz às equações de Cauchy-Riemann} \quad (239)$$

Embora a formulação à direita, nas equações (236) e (237) sejam mais didáticas (ligadas à definição de conjugado), a expressão que parece ser a mais comum são as que ficam à esquerda, para definir os operadores  $\partial, \bar{\partial}$ .

É interessante observar que as equações de Cauchy-Riemann são um exemplo de equação diferencial parcial que foi resolvida ao longo de mais de um século, resultando na construção do que se chamava de *teoria das funções* que se pode dizer, com alguma dose de exagero, que é *a solução das equações de Cauchy-Riemann*, ou, a solução destas equações é uma função analítica e vice-versa. As funções analíticas são também chamadas de *funções holomorfas*.

É interessante observar que se  $f = u + iv$  for analítica, então as duas funções reais  $u, v$  são harmônicas, quer dizer, satisfazem à equação homogênea de Laplace  $\Delta(u) = \Delta(v) = 0$ , isto é consequência direta das equações de Cauchy-Riemann e do teorema de Schwarz-Clairaut das derivadas mistas. As funções  $u, v$  chamam-se *conjugados harmônicos*. A recíproca é verdadeira e passa pela solução da equação diferencial de Cauchy-Riemann (as equações de Cauchy-Riemann) em que uma das duas funções,  $u$  ou  $v$ , é um dado do problema. A solução é única a menos de uma constante. Para resolver a *equação diferencial parcial*  $\Delta(F) = 0$  foi preciso montar toda a teoria das funções analíticas.

### - derivada parcial

Quando uma função,  $F$  for multivariada, há derivadas “parciais” que podem ser calculadas relativamente a cada uma de suas variáveis considerando então as demais variáveis como constantes. Notações:

$$\frac{\partial F}{\partial x} = F_x$$

indica que a derivada foi calculada relativamente a variável  $x$  considerando as demais variáveis “constantes”. É uma contradição difícil de resolver (a não ser com uma notação mais complicada e pouco usada, de múltiplos índices), porque  $F_x$  é uma função das mesmas variáveis que  $F$ .

Quando se fala, “considerando as demais variáveis constantes”, isto vale apenas para efeito do cálculo da derivada. Por exemplo, se

$$F(x, y, z) = x^2 + 2xyz + y^2 + z^3$$

então

$$F_x(x, y, z) = 2x + 2yz; F_y(x, y, z) = 2xz + 2y; F_z(x, y, z) = 2xy + 3z^2;$$

que são, respectivamente as derivadas

$$\frac{\partial F}{\partial x}(x, y, z); \frac{\partial F}{\partial y}(x, y, z); \frac{\partial F}{\partial z}(x, y, z)$$

A jacobiana é a matriz (funcional) das derivadas parciais.

A letra “ $x$ ” que aparece no símbolo do operador derivada, é apenas um índice indicando relativamente a que variável a derivada foi calculada. Se eu quiser calcular o valor da derivada, por exemplo, no ponto  $(-1, 2, 3)$  eu vou escrever:

$$\frac{\partial F}{\partial x}(-1, 2, 3) = 2x + 2yz|_{x=-1, y=2, z=3} \quad (240)$$

$$\frac{\partial F}{\partial x}(-1, 2, 3) = -2 + 12 = 10 \quad (241)$$

$$F_x(-1, 2, 3) = 10 \quad (242)$$

- desigualdade de Hardy Há duas formas desta desigualdade que foi enunciada em 1920 por Hardy. Na forma discreta ele estabelece que se  $a = (a_i)_{i=1}^{\infty}$  for

uma sucessão de números positivos então para todo  $p > 1$

$$\sum_{i=1}^{\infty} \left( \frac{a_1 + \dots + a_i}{i} \right)^p \leq \left( \frac{p}{p-1} \right)^p \sum a^p = \| \| a \| \|_p^p \quad (243)$$

Há alguns casos triviais desta desigualdade que mostram mostram como ela era um resultado difícil em 1920:

- se  $a$  for uma sucessão convergente a série á esquerda converge para o limite de  $a$ , que é a convergência à Cesaro, convergência em média aritmética que nada acrescenta a uma sucessão convergente (séries são sucessões).
- se  $a$  for limitada, a série á direita tem limite que é menor do que o  $\sup(a)$ , mas a série à esquerda não precisa ter limite, e em geral não tem, (uma condição necessária (não suficiente) para que uma série seja convergente é que a sucessão do “termo geral” convirja para zero).
- se  $a \in l(\mathbf{N}^p)$ , à direita se tem  $\| \| a \| \|_p^p$ , como valor da série, e como a convergência à Cesaro coincide com a convergência no caso de series convergentes, então também à esquerda se tem  $\| \| a \| \|_p^p$ . Porém  $\frac{p}{p-1} > 1$  valendo portanto a desigualdade com a observação de que para grandes valores de  $p$  decai a diferença do segundo membro da desigualdade, mas com isto nos aproximamos do segundo caso ...

A forma integral estabelece que se  $f$  for uma função integrável e positiva, então

$$\int_0^{\infty} \left( \frac{1}{x} \int_0^x f(t) dt \right)^p dx \leq \frac{p}{p-1} \int_0^{\infty} f(x)^p dx = \| \| f \| \|_p^p \quad (244)$$

Por alto valem as considerações feitas para o caso discreto com as devidas diferenças valendo portanto a desigualdade sempre exceto no caso trivial, quando  $f$  for quase sempre zero, quando as integrais valem zero nos dois lados da desigualdade.

- **determinante** , Álgebra Linear e Multilinear. Determinante é um número que aparece associado a uma matriz retangular  $n \times n$ . A solução dum *sistema de equações lineares* é muito semelhante, formalmente, à solução da equação do primeiro grau  $ax + b = c$ . Ignore a comutatividade, e você tem dois tipos de equação do primeiro grau:

$$ax + b = c \Rightarrow x = a^{-1}(c - b); \text{ inverso à esquerda de } a \quad (245)$$

$$xa + b = c \Rightarrow x = (c - b)a^{-1}; \text{ inverso à direita de } a \quad (246)$$

$$(247)$$

num *anel não comutativo* em que  $a$  tenha por inverso  $a^{-1}$ . Se o inverso existir, ele é único, mas na equação (245) ou (246) eu usei o inverso de  $a$  multiplicando à direita ou à esquerda porque assumi que estrutura não era comutativa. Isto

significa que numa estrutura *não comutativa* há duas soluções, possivelmente, para uma equação como esta da equação (245). É o caso das equações matriciais, confira *matrizes*.

Em  $\mathbf{N}, \mathbf{Z}, \mathbf{Q}, \mathbf{R}, \mathbf{C}$  estas equações são idênticas e se tem o hábito de escrever

$$x = \frac{c - b}{a}$$

como solução, devido à comutatividade<sup>10</sup> da multiplicação.

Mas  $a, b, c$  forem matrizes e estas equações forem possíveis (as dimensões envolvidas forem compatíveis) a solução do sistema seria expressa por uma das equações (245) ou (246). A razão é que a multiplicação de matrizes não é comutativa.

Para que se possa escrever as soluções que aparecem na equação (eq. 245) ou (246) é *suficiente*<sup>11</sup> que  $\det(a) \neq 0$  e neste caso podemos aplicar a fórmula para o inverso de uma matriz para calcular  $a^{-1}$ , confira *matriz*. Porém esta forma de resolver um sistema linear é muito pouco prática e apenas serve para introduzir o conceito de *determinante* no cálculo da matriz inversa.

O determinante duma matriz  $n \times n$  é uma forma  $n$ -linear alternada aplicada aos vetores-coluna (ou vetores-linha) da matriz. Este assunto é estudado na disciplina *álgebra multilinear* e o determinante é uma *forma multilinear* alternada. O adjetivo “*alternada*” vem de uma propriedade dos determinantes: se trocarmos duas colunas (ou duas linhas) o determinante muda de sinal. O adjetivo “*multilinear*” vem da propriedade de que se uma coluna (ou linha) for substituída por uma combinação linear de vetores se terá a combinação linear dos determinantes obtidos usando, em cada caso, um dos vetores desta combinação e usando os mesmo coeficientes para combinar os determinantes. Como isto vale para qualquer coluna (ou linha) então o determinante é “multi”-linear.

Os determinantes de matrizes  $2 \times 2$  e  $3 \times 3$  são fáceis de calcular, mas a expressão geral para o cálculo de um determinante envolve o conceito de *permutação* como é relativamente fácil de mostrar, mas não de calcular. Suponha que

$$\mathcal{A} = (a_{ij}); i, j = 1, \dots, n; \mathcal{A} = \begin{pmatrix} a_{11} & \dots & a_{1j} \dots & a_{1n} \\ & & \dots & \\ a_{k1} & \dots & a_{ij} \dots & a_{in} \\ & & \dots & \\ a_{n1} & \dots & a_{nj} \dots & a_{nn} \end{pmatrix} \quad (248)$$

e que  $\sigma \in \text{sim}(n)$ , um elemento,  $\sigma$ , do grupo das permutações de  $n$  elementos, então

$$\det(\mathcal{A}) = \sum_{\sigma \in \text{sim}(n)} \text{sinal}(\sigma) \prod_{i=1 \dots n} a_{i\sigma(i)} \quad (249)$$

ou seja a soma sobre todos as possíveis permutações  $\sigma$  de  $n$  elementos dos produtos das entradas da matriz  $\mathcal{A}$  obtidos quando se permutam os elementos

<sup>10</sup>A notação  $\frac{a}{b}$  somente pode ser usada em estruturas comutativas.

<sup>11</sup>Mas não é necessário. . .

que aparecem no produto. Um exemplo de termo da soma seria

$$a_{1\sigma(1)}a_{2\sigma(2)} \dots a_{n\sigma(n)}; \sigma \in \text{sim}(n); \quad (250)$$

ou para visualizar melhor, considere a matriz

$$\mathcal{A}_\sigma \begin{pmatrix} a_{11} & \dots & a_{1\sigma(j)} \dots & a_{1j} \\ & & \dots & \\ a_{\sigma(j)1} & \dots & a_{i\sigma(j)} \dots & a_{\sigma(i)j} \\ & & \dots & \\ a_{n1} & \dots & a_{n\sigma(j)} \dots & a_{nn} \end{pmatrix} \quad (251)$$

que foi obtida da matriz  $\mathcal{A}$  aplicando a permutação  $\sigma$  em sua coluna de ordem  $j$  então o número que aparece na equação (eq. 250) é o produto dos termos da diagonal de  $\mathcal{A}_\sigma$ , então, para obter o  $\det(\mathcal{A})$  se calculam todas as imagens de  $\mathcal{A}$  pela permutações  $\sigma \in \text{sim}(n)$  e se somam estes produtos usando *sin*al(*sigma*) como coeficiente. Este é o método usado para o cálculo do determinante numa matriz  $3 \times 3$ , observe que  $\text{sim}(3)$  tem  $3! = 6$  elementos portanto haveria 6 termos na soma da equação (eq. 249) que podem ser obtidos com esquema

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{pmatrix} \quad (252)$$

em que as duas primeiras colunas foram repetidas. Esta expressão não é *computacional*, entretanto. O método para resolver sistema de  $n$  equações lineares com  $n$  incógnitas passa por triangularizar as matrizes quando o cálculo do determinante se transforma no produto dos termos da diagonal. Este assunto pertence à *álgebra linear computacional*, ao *cálculo numérico* ou à *análise numérica*.

---

- **diagrama** confira:

- Hasse, diagrama. Teoria dos conjuntos.
- Venn, diagrama. Teoria dos conjuntos.

---

- **diferenciabilidade** É uma propriedade das funções, estudadas no Cálculo Diferencial e Integral que diz respeito à possibilidade do gráfico de uma função poder ter retas tangentes nas vizinhanças dum ponto. O coeficiente angular de qualquer uma tal reta tangente ao gráfico de  $f$  no ponto  $(a, f(a))$  é a *derivada* de  $f$  no ponto  $a$  designada com o símbolo  $f'(a)$ . É um conceito importante da *geometria diferencial* que estuda as propriedades dos objetos geométricos suficientemente macios a ponto de poderem ter tangentes. A *geometria diferencial* é uma divisão da *topologia diferencial*.

A possibilidade de ter variedades lineares afins tangentes (objetos lineares tangentes) permite que se calcule o coeficiente angular instantâneo (local) de uma variedade não linear. Observe que o coeficiente angular pode ser uma matriz de coeficientes angulares.



É a propriedade de uma variedade ter derivada.

- **diferenciável** Uma propriedade de certas funções. Diz-se que uma função  $f$  é diferenciável em um ponto  $a$ , se for possível encontrar uma reta tangente ao gráfico de  $f$  no ponto  $(a, f(a))$ , e neste caso a equação desta reta tangente é

$$y = f(a) + f'(a)(x - a)$$

Por exemplo, a função  $y = f(x) = \|x\|$  não é diferenciável no ponto  $a = 0$  porque neste ponto há duas retas tangente ao gráfico de  $f$ , as retas

$$y = -x; y = x;$$

como há duas retas tangentes, então dizemos que  $f$  não é diferenciável neste ponto. Observe que esta função  $y = f(x) = \|x\|$  é diferenciável em qualquer ponto diferente de  $a = 0$ . Dizemos que a diferenciabilidade é uma característica avançada de *continuidade*, porque é preciso que a função seja contínua para ser diferenciável (a recíproca é falsa,  $y = f(x) = \|x\|$  é contínua na reta inteira). De forma mais ampla ser *diferenciável* significa ter uma função linear tangente. Para espaços de dimensão diferente de 1 é preciso contornar o quociente na definição de *tangência* com uma relação de equivalência.

$$\exists A_a; y = f(a) + A_a(x - a) \quad (253)$$

$$y + R = f(a) + rA_a(x - a) \quad (254)$$

$$(255)$$

$A_a$  é a derivada de  $f$  no ponto  $a$  e esta forma de falar conduz à definição de derivada em espaços de dimensão diferente de 1:  $f$  é *diferenciável* se  $f(x) - f(a)$  for tangente a uma função linear  $A_a$ .

- A função identicamente nula é diferenciável e sua derivada é ela mesma.
- Mais geral, se  $f$  for linear então é a tangente a si própria, porque a tangência é uma relação de equivalência, logo é a sua própria derivada.
- Como a translação de uma função linear, uma função linear afim, tem a função linear tangente, então as funções constantes podem ser vistas como *lineares afins* tendo por derivada a função identicamente nula. Isto põe em evidência uma propriedade de *inversão da derivação*: há uma infinidade de funções com a mesma derivada, aquelas cuja diferença seja uma constante, todas as constantes têm a mesma derivada, a função identicamente nula, que é linear.
- Se  $f - g = r$ , uma constante, então é derivável. Se  $f$  for derivável então  $g$  é derivável e  $g' = f'$ .

- **dimensão** É um conceito da *estrutura de espaços vetoriais*. Intuitivamente falando é o

- número exato de informações,
- independentes,
- para gerar um espaço vetorial.

Para concretizar estes conceitos precisamos de um pouco de Álgebra Linear. Vou começar pela *independência linear* porque é fácil criar um caminho intuitivo, baseado na geometria, para defini-la, num espaço bidimensional, o plano, apenas você precisa de fazer algum esforço para entender.

Com dois vetores podemos “gerar um plano”, que é um espaço vetorial de dimensão dois, mas estes vetores não podem ser *colineares*. Esta afirmação vem direto da Geometria estudada no Ensino Médio porque dois vetores são determinados por três pontos. Se não forem *colineares* determinam um triângulo que é uma figura plana, assim dois vetores *não colineares* determinam um plano.

Aqui estou usando um conceito que é preciso explicitar: os vetores tem uma origem comum, é por esta razão que fiz a afirmação: dois vetores, correspondem a três pontos, a figura(fig 35) o mostra , página 96,

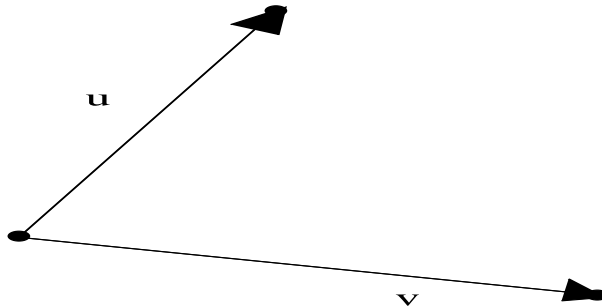


Figura 35: Dois vetores correspondem a três pontos

Se os dois vetores forem colineares, quaisquer operações entre eles dois resultaria noutro vetor em cima da mesma reta em que eles se encontram. Isto significa que eles “apenas geram um espaço de dimensão 1” - uma reta. Ou ainda, eles estão sobre uma mesma reta. Dois vetores *colineares* são vetores que são, *um, múltiplo do outro*. Vou colocar isto em símbolos:

$$\text{dois vetores } u, v \quad (256)$$

$$\text{são colineares: } u = \lambda v; \lambda \in \mathbf{R}; \lambda \neq 0; \quad (257)$$

$$\text{deduzimos: } u - \lambda v = 0; \quad (258)$$

Na equação (257) eu impuz a condição,  $\lambda \neq 0$  porque, caso contrário,  $u = 0$  e apenas teríamos um vetor não nulo gerando uma reta.

Da equação (258) podemos concluir que existem dois escalares,  $\alpha = 1, \beta = -\lambda$  tal que

$$\alpha u + \beta v = 0; \alpha \neq 0; \beta \neq 0; \quad (259)$$

e é possível reescrever (258) assim

$$\gamma u + v = 0; \gamma = \frac{-1}{\lambda}; \quad (260)$$

porque já vimos que não teria sentido que  $\lambda$  fosse zero.

Isto nos permite completar a equação (259) dizendo que, se os vetores  $u, v$  forem *colineares*, é possível encontrar um par de escalares  $\alpha, \beta$ , ambos diferentes de zero de modo a obter uma combinação linear nula com  $u, v$ , que é o conteúdo da equação (259).

Definimos, formalmente, *colinearidade* que é sinônimo de *dependência linear* no plano, num espaço de dimensão dois.

O contrário disto é *independência linear* de dois vetores que na linguagem da geometria é *não colinearidade*.

Resumindo, dois vetores,  $u, v$ , não são colineares se não for possível encontrar dois números reais  $\alpha \neq 0, \beta \neq 0$  tal que

$$\alpha u + \beta v = 0 \quad (261)$$

isto é a definição de vetores *linearmente independentes*.

Vou agora dar o salto para a dimensão três que será definitivo para qualquer outra dimensão. Se três vetores forem *linearmente independentes* então eles geram um espaço de dimensão três, se não forem, no máximo geram um espaço de dimensão dois.

Aqui vale lembrar uma questão de Física do Ensino Médio: três forças estão em equilíbrio estático se os vetores que as representarem formarem um triângulo o que significa: a soma delas é zero, consequência da Lei de Chasles.

Vamos considerar três vetores:  $u, v, w$  elas podem não representar um sistema em equilíbrio, mas se pudermos encontrar os escalares  $\alpha, \beta, \gamma$  tal que

$$\alpha u + \beta v + \gamma w = 0; \alpha \neq 0, \beta \neq 0, \gamma \neq 0 \quad (262)$$

então elas formam um triângulo que é uma figura plana.

Repetimos a expressão contida na equação (259) agora para três vetores que são *linearmente dependentes*. Descobrimos assim a definição geral da *dependência linear*, e *independência linear* é a sua negação.

### Definição 8 () *independência linear*

*Os vetores*

$$u_1, \dots, u_n$$

*são linearmente independentes se dados os escalares*

$$\lambda_1, \dots, \lambda_n$$

*tivermos*

$$\lambda_1 u_1 + \dots + \lambda_n u_n = 0$$

*então todos  $\lambda_k$  são nulos. Algumas vezes isto é resumido na frase: os vetores  $(u_{k;k=1\dots n})$  são l.i. se a única combinação linear nula é a trivial, com todos os escalares nulos.*

Parece ser muito forte a exigência de que *todos*  $\lambda_k$  sejam nulos, mas suponha que apenas um deles seja nulo, por exemplo o primeiro, então elimine  $\lambda_1, u_1$  e você conclui que o restante não pode ser *linearmente independente*, portanto repetindo, sucessivamente vamos chegar ao caso do plano.

Um exemplo que também faz parte do Ensino Médio ajudaria até mesmo para mostrar-lhe que os espaços de dimensão alta fazem parte da sua experiência.

Um polinômio do grau  $n$  que seja *identicamente nulo* tem todos os coeficientes nulos, reza um teorema sobre polinômios. Por que os monômios  $1, x, x^2, \dots, x^n$  são vetores linearmente independentes... Os polinômios são vetores.

O espaço dos polinômios de grau menor ou igual a  $n$  é um espaço vetorial de dimensão  $n + 1$ , porque os vetores

$$1, x, x^2, \dots, x^n \quad (263)$$

são linearmente independentes. Este conjunto de geradores se chama de *uma base* do espaço.

1. Eles fornecem o número exato de informações para obtermos qualquer polinômio de grau menor ou igual a  $n$ ;
2. eles são linearmente independentes;
3. eles geram um espaço de dimensão  $n + 1$ .

E assim obtivemos as três condições iniciais para compreender dimensão.

E qual seria a dimensão do conjunto de todos os polinômios? Uma resposta rápida seria: não pode ser finita, porque sempre haverá um polinômio de grau maior do que um que escolhermos. É isto que significa que há espaços de *dimensão infinita*.

Você poderia testar este conceito pedindo que eu escrevesse um (vetor) polinômio de *dimensão infinita* e isto é impossível! Esta impossibilidade, entretanto, não serve para negar que exista, pelo menos um espaço de dimensão infinita, o dos polinômios. O *infinito* na verdade é uma abstração Matemática e serve para fechar a teoria e neste ponto se criam discussões bizantinas negando a existência daquilo que não podemos exibir.

Há vários espaços de dimensão infinita à nossa volta, um deles é o espaço das ondas eletromagnéticas, aquelas que permitem as telecomunicações e que na prática garantem a poluição dos telefones celulares: *será sempre possível ter mais um celular* porque a cada celular corresponde, de uma certa forma, uma única *frequência* eletromagnética, a frequência específica em que o aparelho funciona. Os parâmetros que identificam os vetores, neste espaço, se chamam *frequências*

Os vetores básicos, linearmente independentes, para gerar este espaço são

$$\sin(kx), \cos(kx); k \in \mathbf{N} \quad (264)$$

e eles aparecem na teoria dos polinômios trigonométricos de Fourier associados às *séries de Fourier*. Os vetores na equação (264) formam uma base para

o espaço das ondas eletromagnéticas, ou para as séries de Fourier... É outro espaço de dimensão infinita, mas nunca será possível espalhar numa calçada um número infinito de aparelhos celulares, apenas, sempre será possível acrescentar mais um! Como no caso dos polinômios.

### Dimensão e cardinalidade

A *dimensão finita* corresponde à cardinalidade dum conjunto de vetores chamado de *base* do espaço vetorial. Para os espaços de dimensão infinita esta correspondência parece que “desaparece” porque *há diversas classes de infinito*. Os dois exemplos de espaços vetoriais de dimensão infinita, dos polinômios ou dos polinômios trigonométricos tem uma mesma classe de infinito para a “quantidade de vetores” da base. Aqui a palavra “quantidade” perde sentido, a forma correta de falar é *cardinalidade*. O que podemos dizer é que eles têm uma base *enumerável* que é a cardinalidade de  $\mathbf{N}$ .

Se a dimensão for finita  $n$  um exemplo de espaço é  $\mathbf{R}^n$ , o espaço das énuplas ordenadas de números reais. Esta notação de potência foi muito bem explorada para generalizações.

$\mathbf{R}^n$  é o espaço das funções

$$\{f; f \in R^n\} = \{f; \{1, 2, \dots, n\} \xrightarrow{f} \mathbf{R}\}$$

se convença disto. O conjunto de todas as funções definidas no conjunto  $X$  com valores no conjunto  $Y$  é representado com o símbolo  $Y^X$  porque, se os dois conjuntos forem finitos:

- Se  $\text{card}(X) = n$ ;
- se  $\text{card}(Y) = m$ ;
- então  $\text{card}(Y^X) = m^n$

Isto é um resultado de *análise combinatória simples*, este conjunto  $Y^X$  é dos arranjos com repetição dos elementos de  $Y$  tomados  $n$  a  $n$ . São as énuplas de elementos de  $Y$ . No caso de  $\mathbf{R}^n$  têm-se também arranjos apenas não podemos mais falar da quantidade deles que é um cardinal não finito, *infinito*, que pertence a uma classe diferente dos dois outros mencionados no caso dos polinômios algébricos ou dos polinômios trigonométricos (ou celulares ...). Então

$$\mathbf{R}^n = \mathbf{R}^{\{1,2,\dots,n\}} \quad (265)$$

Assim poderíamos nos referir a diversos espaços vetoriais:

$$\mathbf{R}^{[0,1]}; \text{ das funções reais definidas em } [0, 1] \quad (266)$$

$$\mathbf{R}^X; \text{ das funções reais definidas em } X \quad (267)$$

$$\mathbf{R}^{\mathbf{N}}; \text{ das sucessões de números reais} \quad (268)$$

O estudo da dimensão de tais espaços, ou para falar mais exato, de subespaços destes, é assunto de artigos de pesquisa... para determinar a classe a que pertencem!

- **Dirac, medida** Ver medida de Dirac.

- **Dirichlet, núcleo** ver núcleo de Dirichlet.

- **distância** Havia apenas uma distância, euclidiana, que vem do teorema de Pitágoras

$$d(P, Q) = \sqrt{(a - m)^2 + (b - m)^2}; P = (a, b); Q = (m, n) \quad (269)$$

e que serve para medir o comprimento da diagonal de um retângulo ou da hipotenusa no triângulo retângulo, como você pode ver na figura (fig 36), página 100,

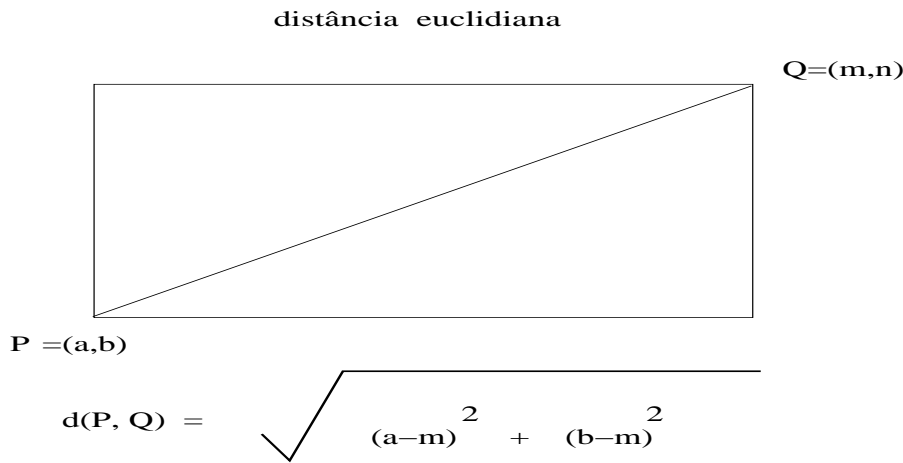


Figura 36: O comprimento da diagonal dum retângulo

Esta fórmula também serve para calcular a distância entre dois pontos, no plano ou no espaço:

$$P, Q \in \mathbf{R}^2; P = (a_1, a_2); Q = (b_1, b_2); \quad (270)$$

$$d(P, Q) = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2}; \quad (271)$$

$$P, Q \in \mathbf{R}^n; P = (a_1, \dots, a_n); Q = (b_1, \dots, b_n); \quad (272)$$

$$d(P, Q) = \sqrt{\sum_{k=1}^{k=n} (a_k - b_k)^2}; \quad (273)$$

O conceito de distância se generaliza para qualquer função de duas variáveis que tenham as mesmas propriedades que a *distância euclidiana*.

Quando pudermos definir uma distância  $d$  em um conjunto  $X$ , diremos que se trata de um *espaço métrico*  $(X, d)$

**Definição 9 (distância)** *Distância ou métrica*

Uma distância é uma função positiva,  $d$ , definida para todos os pares  $(x, y)$  de elementos de um conjunto  $X$  satisfazendo as propriedades

1. positiva  $d(x, y) \geq 0$  e  $d(x, y) = 0$  se e somente se  $x = y$ ;
2. reflexividade  $d(x, y) = d(y, x)$ ;
3. desigualdade triangular  $d(x, y) + d(y, z) \geq d(x, z)$ ;

Um exemplo bem simples, da *vida real*, pode explicar necessidade de diferentes formas de distância: uma firma que esteja instalando fios de telefone<sup>12</sup> precisa de uma forma diferente para o cálculo da distância. A distância da central telefônica para uma residência não pode ser medida com a distância geométrica, do compasso, que corresponde

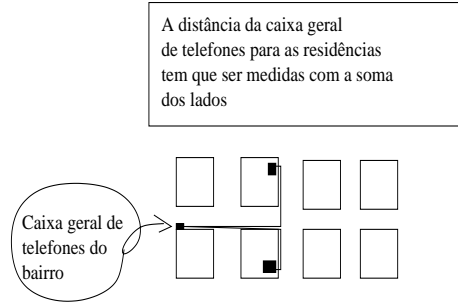


Figura 37: vizinhança, soma das medidas dos lados

à distância euclidiana, tem que ser usada a somas dos lados (ou fração dos lados) do retângulo, como pode ser visto na figura (??), página ??.

Se a medida fosse calculada com a métrica euclidiana

$$d(P, Q) = \sqrt{(a - m)^2 + (b - m)^2}; P(a, b); Q = (m, n) \tag{274}$$

iria faltar cabo para chegar da caixa telefônica até as residências. A distância euclidiana é menor do que a *distância-1*, da soma dos lados.

Podemos também medir a distância entre duas funções, você pode ver isto lendo a respeito de *vizinhanças*, há outras necessidades mais complexas que conduzem à distâncias mais envolvidas até porque as funções, algumas vezes tem uma quantidade muito grande de informações.

Foi feita referência à “distância-1” no exemplo do cabo telefônico, o sistema de equações (260), mostra algumas variações que generalizam a distância euclidiana:

$$d_1(P, Q) = |(a - m)| + |(b - m)|; P(a, b); Q = (m, n); \tag{275}$$

$$d_2(P, Q) = \sqrt{(a - m)^2 + (b - m)^2}; P(a, b); Q = (m, n); \tag{276}$$

$$d_3(P, Q) = \sqrt[3]{(a - m)^3 + (b - m)^3}; P(a, b); Q = (m, n); \tag{277}$$

$$d_p(P, Q) = \sqrt[p]{(a - m)^p + (b - m)^p}; P(a, b); Q = (m, n); \tag{278}$$

<sup>12</sup>isto ainda existe e seguirá existindo, dá mais segurança do que a ausência de fios.

em que a equação (278) mostra o modelo geral que vale para as equações anteriores. Se  $p \geq 1$  a expressão tem as três propriedades caracterizam a distância euclidiana, generalizando o conceito de distância.

Quando  $p = 1$  você tem a distância necessária para medir a quantidade de cabo telefônico.

Mas  $p$  pode ser  $\infty$ , apenas a equação tem que ser expressa de forma inteiramente diferente do modelo na equação (278)

$$d_{\infty}(P, Q) = \text{Max}\{|a - m|, |b - m|\}; P(a, b); Q = (m, n); \quad (279)$$

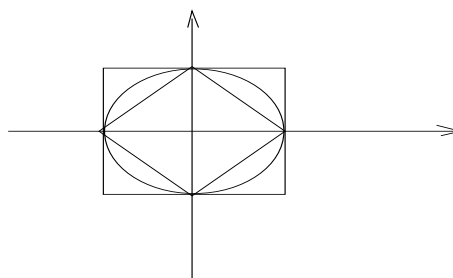
que é chamada de “distância do sup” porque algumas vezes temos que usar o “supremo” porque nem sempre existe o máximo. A medida que  $p$  cresce, fica menor a distância entre dois pontos, inversamente as bolas definidas com estas distâncias aumentam, geometricamente, a medida que  $p$  cresce. A figura (38), página 102,

mostra a bola calculada com as distâncias- $p$ ,  $p \in \{1, 2, \infty\}$  sendo a maior delas a que foi obtida com a distância do sup,  $p = \infty$ , e a que “parece” uma bola, é a bola euclidiana.

A generalização de um conceito perturba os nossos conceitos intuitivos, o conceito de distância generalizado altera a nossa concepção do que pode ser uma bola como na figura (38). Compare com as bolas que podemos definir num espaço de funções. . .

Se os técnicos que colocam os fios, usassem a *distância do supremo*, os fios ficariam se dobrando dentro dos conduites e fio em excesso seria gasto.

Tudo que foi dito acima para pontos do  $\mathbf{R}^2$  vale, com as mesmas palavras, para os pontos do  $\mathbf{R}^n$  e até mesmo para os espaços de sucessões que são espaços de dimensão infinita.



Bolas, centradas na origem, calculadas com as distâncias- $p$

Figura 38: Bolas com as distâncias  $d_p(P, Q)$

- **distância da reta à origem** A equação cartesiana da reta no plano é  $Ax + By + C = 0$  e não tem sentido que os números  $A, B$  sejam ambos zero.

Se  $A = 0$  então  $y = -\frac{C}{B}$  e a reta é paralela ao eixo  $OX$ . Se  $B = 0$  então  $x = -\frac{C}{A}$  e a reta é paralela ao eixo  $OY$ . O vetor  $(A, B)$  é perpendicular à reta  $Ax + By + C = 0$ .

A reta  $Ax + By + C = 0; C \neq 0$  não passa na origem, confira na figura (fig 39), página 103, a posição relativa da reta  $r$  de equação  $Ax + By + C = 0$  relativamente à origem dos eixos.

O vetor  $(A, B)$  é perpendicular à reta  $Ax + By + C = 0$  e isto é consequência do produto escalar  $\langle (A, B), (x, y) \rangle = 0$  significar que os vetores  $(A, B), (x, y)$



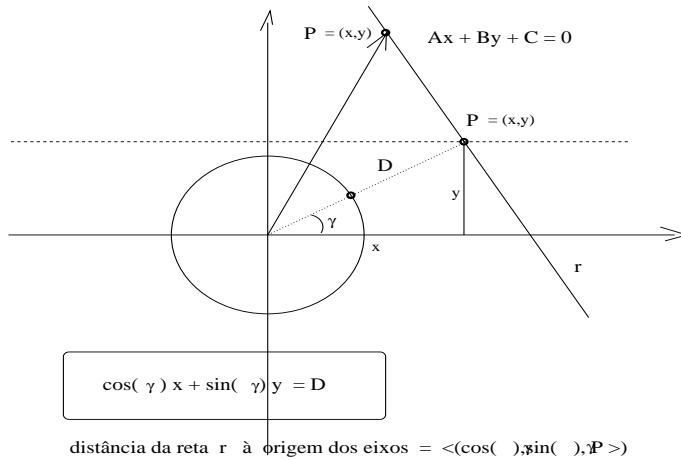


Figura 39: equação normal da reta e distância à origem

são perpendiculares, aplique no caso  $C = 0$  que corresponde a uma reta paralela passando na origem.

Observe que qualquer valor para  $C$  produz uma reta paralela à reta

$$Ax + By + C = 0;$$

e o caso  $C = 0$  corresponde à reta que passa na origem.

Como  $\sqrt{A^2 + B^2} = D \neq 0$ , podemos “normalizar” a equação da reta:

$$\begin{cases} \sqrt{A^2 + B^2} = D; \\ \frac{A}{D}x + \frac{B}{D}y + \frac{C}{D} = 0; \text{ equação normal da reta;} \end{cases} \quad (280)$$

então o vetor  $(\frac{A}{D}, \frac{B}{D})$  é um vetor unitário e o produto escalar dele por qualquer vetor descrevendo a reta produz a distância da reta  $r$  à origem:

$$\left\langle \left( \frac{A}{D}, \frac{B}{D} \right), (x, y) \right\rangle = -\frac{C}{D}$$

A figura (fig 40), página 104, mostra um vetor unitário na direção da perpendicular à reta  $r$  e que existe uma única reta perpendicular ao vetor  $(\cos(\gamma), \sin(\gamma))$  cuja distância à origem seja  $-\frac{C}{D}$  então a combinação linear  $x \cos(\gamma) + y \sin(\gamma) = -\frac{C}{D}$  identifica esta reta sendo sua equação em *coordenadas polares*. Mas as coordenadas polares são interessantes quando o lugar geométrico for radial. No caso de retas, daquelas que passam na origem.

Esta figura, (fig 40), página 104, é interessante porque permite generalizar o resultado da distância duma variedade linear para qualquer dimensão: existe uma única esfera, de centro na origem, tangente à variedade linear cujo raio mede a distância da variedade linear à origem. O vetor unitário da direção do raio no ponto de tangência é o que aparece nas contas dentro do produto escalar.

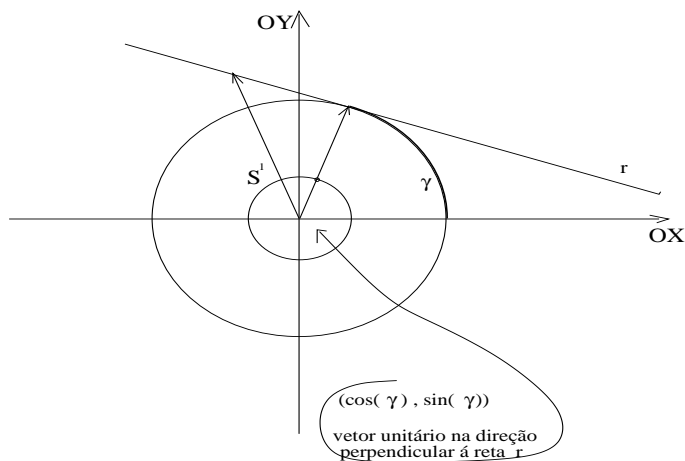


Figura 40: círculo unitário e distância à origem

Estes cálculos podem ser imediatamente generalizados para a equação do plano,  $Ax + By + Cz + D = 0$  no espaço  $\mathbf{R}^3$ :

$$\begin{cases} \Delta = \sqrt{A^2 + B^2 + D^2} \neq 0; \\ \frac{A}{\Delta}x + \frac{B}{\Delta}y + \frac{C}{\Delta}z + \frac{D}{\Delta} = 0 \text{ equação normal do plano;} \end{cases} \quad (281)$$

então a distância do plano  $Ax + By + Cz + D = 0$  à origem dos eixos é dada pelo produto escalar

$$\left\langle \left( \frac{A}{\Delta}, \frac{B}{\Delta}, \frac{C}{\Delta} \right), (x, y, z) \right\rangle = -\frac{D}{\Delta}; \quad (282)$$

- **distância do plano à origem** A equação cartesiana do plano no espaço  $\mathbf{R}^3$  é  $Ax + By + Cz + D = 0$  e não tem sentido que os números  $A, B, C$  sejam todos zero. Compare com a distância da reta à origem.

O plano  $Ax + By + Cz + D = 0$ ;  $D \neq 0$  não passa na origem, confira na figura (fig 39), página 103, por comparação, a posição relativa do plano  $\pi$  de equação  $Ax + By + Cz + D = 0$  relativamente à origem dos eixos.

O vetor  $(A, B, C)$  é perpendicular ao plano  $Ax + By + Cz + D = 0$  e isto é consequência do produto escalar  $\langle (A, B, C), (x, y, z) \rangle = 0$  significar que os vetores  $(A, B, C), (x, y, z)$  são perpendiculares.

Qualquer valor para  $D$  produz um plano paralelo ao plano

$$Ax + By + Cz + D = 0$$

e o caso  $D = 0$  corresponde ao plano que passa na origem.

Como  $\sqrt{A^2 + B^2 + C^2} = \Delta \neq 0$ , podemos “normalizar” a equação do plano:

$$\begin{cases} \sqrt{A^2 + B^2 + C^2} = \Delta; \\ \frac{A}{\Delta}x + \frac{B}{\Delta}y + \frac{C}{\Delta}z + \frac{D}{\Delta} = 0; \text{ equação normal do plano;} \end{cases} \quad (283)$$

então o vetor  $(\frac{A}{\Delta}, \frac{B}{\Delta}, \frac{C}{\Delta})$  é um vetor unitário e o produto escalar dele por qualquer vetor descrevendo o plano produz a distância do plano  $\pi$  à origem:

$$\langle (\frac{A}{\Delta}, \frac{B}{\Delta}, \frac{C}{\Delta}), (x, y, z) \rangle = -\frac{D}{\Delta}$$

Observe que estes cálculos podem ser imediatamente generalizados para qualquer dimensão, a equação da variedade linear

$$A_1x_1 + \dots + A_nx_n + D_n = 0$$

no espaço  $\mathbf{R}^n$ .

$$\begin{cases} \Delta = \sqrt{\sum_{k=1}^n A_k^2} \neq 0; \\ \sum_{k=1}^n \frac{A_k}{\Delta} x_k + \frac{D_n}{\Delta} = 0 \text{ equação normal da variedade;} \end{cases} \quad (284)$$

$\Delta$  é diferente de zero e pode ser calculado porque pelo menos um dos coeficientes  $A_k$  tem que ser diferente de zero. A distância da variedade à origem dos eixos,  $-\frac{D_n}{\Delta}$ , é dada pelo produto escalar

$$\langle (\frac{A_k}{\Delta})_k, (x_k)_k \rangle = -\frac{D_n}{\Delta}; \quad (285)$$

- **distância entre dois pontos** é conceito da Geometria Analítica generalizado na *teoria dos espaços métricos* que define *distância* abstraindo o sentido geométrico mas guardando as propriedades fundamentais da distância. Confira *distância*.

A Geometria Analítica *digitalizou* o espaço e a geometria euclidiana, tornando possível a algebrização da geometria. Assim  $(x, y)$  para dois números reais dados representam um ponto do espaço  $\mathbf{R}^2$ , que é um plano ou um espaço bidimensional, ou  $(x, y, z)$ , para três números reais dados representam um ponto do espaço  $\mathbf{R}^3$ , um espaço tridimensional, ou

$$P = (x_1, x_2, \dots, x_n); x_k \in \mathbf{R}; P \in \mathbf{R}^n \text{ um espaço de dimensão } n; \quad (286)$$

Dados dois pontos no  $P, Q \in \mathbf{R}^2$ , confira a figura (fig 79), página 197, podemos *calcular* a distância entre eles usando o teorema de Pitágoras,

$$d(P, Q) = \sqrt{(m-a)^2 + (n-b)^2} = \sqrt{(a-m)^2 + (b-n)^2}; \quad (287)$$

Observe que o módulo das diferenças que aparecem na equação (eq. 287) medem o comprimento dos lados do triângulo retângulo que o segmento  $\overline{PQ}$  determina com retas paralelas aos eixos coordenados.

As propriedades da distância são

1. reflexividade  $d(P, Q) \geq 0$  e  $d(P, Q) = 0 \iff P = Q$

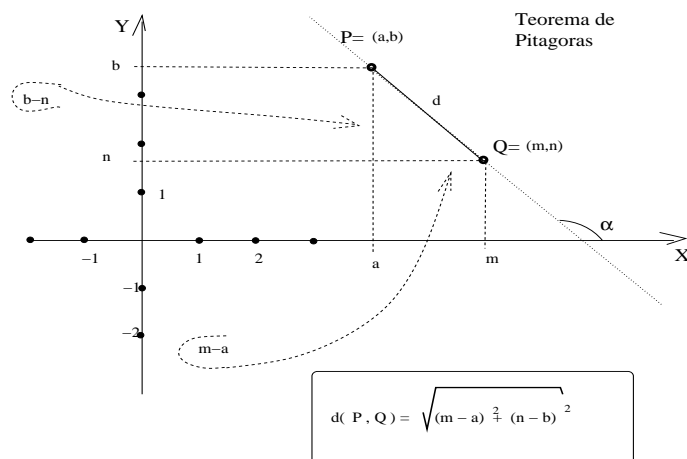


Figura 41: Distância entre dois pontos e o teorema de Pitágoras

2. simetria  $d(P, Q) = d(Q, P)$ ;
3. desigualdade triangular  $d(P, Q) \leq d(P, R) + d(R, Q)$

Qualquer função de duas variáveis que tenha estas propriedades define uma distância num conjunto  $M$  e é quando se diz que  $(M, d)$  é um espaço métrico.

A distância é uma propriedade bidimensional porque envolve em suas propriedades no máximo três elementos do espaço que determinam um plano deste espaço, o que facilita muito nas demonstrações das propriedades mais gerais da distância. Para o  $\mathbf{R}^n$  a fórmula na equação (eq. 287) fica

$$\begin{cases} d(P, Q) = \sqrt{\sum_{k=1}^n (p_k - q_k)^2}; \\ P = (p_1, \dots, p_n); Q = (q_1, \dots, q_n); \end{cases} \quad (288)$$

Uma forma de definir a distância passa pelo *produto interno* ou *produto escalar* que está definido em todos os espaços de dimensão finita e também em espaços de dimensão não finita. O *produto escalar*, como meio de definir distância também pode ser definido de várias maneiras. A definição usual no  $\mathbf{R}^n$  é

$$\langle P, Q \rangle = \sum_{k=1}^n p_k q_k = \|P\| \|Q\| \cos(\gamma); \quad (289)$$

em que  $\gamma$  é o ângulo entre  $P$  e  $Q$ .

A primeira expressão na (eq. 289) é uma *forma bilinear* definida em  $\mathbf{R}^n$  e a segunda é a expressão bidimensional da mesma no plano determinado pelos dois vetores.

Observe que a equação (eq. 289) nos oferece a oportunidade de definir ângulo entre dois vetores dum espaço qualquer em que esteja definido um produto

escalar, usando a função  $\text{acos}()$ . Tais espaços se chamam *espaço com produto interno*.

Confira *produto escalar* ou *produto interno*.

### Observação 3 *triedro positivo*

A ordenação de vetores no espaço é uma questão importante e apenas traduz as leis da Física ou da Natureza. Observe como isto é feito usando a regra do sacacolhas que também poderia ser chamada de regra do parafuso.

Como dois vetores no  $\mathbf{R}^n$  determinam um plano do  $\mathbf{R}^n$ , então o ângulo é um conceito geométrico claro, é o menor segmento do círculo trigonométrico que os dois vetores determinam. Na figura (fig 80), página 199, você vê dois vetores do  $\mathbf{R}^n$  que determinam um plano

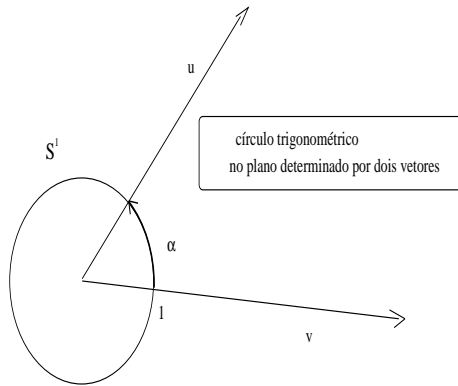


Figura 42: o ângulo entre dois vetores do  $\mathbf{R}^n$

e tomando a origem comum aos dois vetores como centro de um círculo de raio 1 um dos vetores corta o círculo trigonométrico na origem do círculo.

Aqui parece haver alguns conceitos difusos, mas não é bem assim. Existe uma orientação padrão para os vetores no espaço, dada pela regra do sacacolha. O sacacolha segue a orientação dos parafusos, para enfiá-lo na rolha a rotação é negativa, é a mesma orientação para enfiar um parafuso numa peça. Com esta regra a figura (fig 80), mostra no círculo trigonométrico o ponto inicial marcado pela reta suporte do vetor  $\vec{v}$ , estando os vetores  $\vec{v}, \vec{u}$  orientados positivamente pela regra do sacacolhas. Se houver um terceiro vetor ele vai ocupar a posição do sacacolhas pronto para “sacar uma rolha” mostrando a posição de tres vetores formando um triedro com ordenação positiva.

Para que você se convença de que esta regra está consistente, deixe-me sugerir-lhe um experimento: considere dois vetores perpendiculares, tome a figura (fig 80) como padrão, pensando nos dois vetores  $\vec{v}, \vec{u}$  cortando o círculo trigonométrico na quarta parte,  $\frac{\pi}{2}$ .

Mas agora considere que  $|\vec{v}| = |\vec{u}| = 1$ , por construção, você selecionou  $\vec{v}, \vec{u} \in \mathbf{S}^1$ , no círculo trigonométrico.

O sacacolhas na origem dos eixos, o ponto comum dos vetores  $\vec{v}, \vec{u}$ , representando um terceiro vetor “arrancando a rolha”, na direção positiva, é o vetor  $\vec{z}$  e também  $|\vec{z}| = 1$ , tudo por construção.

Os dois vetores  $\vec{u}, \vec{z}$ , determinam um plano e o vetor  $\vec{v}$  é o sacacolhas nesta nova situação. O produto vetorial de vetores executa, algebricamente, esta operação:

$$\vec{v} \times \vec{u} = \vec{z}; \vec{u} \times \vec{z} = \vec{v}; \quad (290)$$

ou ainda, como na Física:

$$\vec{i} \times \vec{j} = \vec{k}; \vec{j} \times \vec{k} = \vec{i}; \quad (291)$$

sendo que esta operação não é comutativa...

$$\vec{k} \times \vec{j} = -\vec{i}; \quad (292)$$

A eletricidade se comporta de acordo com esta regra para definir o sentido da corrente positiva gerada por gerador montado com ímãs eletromagnéticos. A regra do sacacolhas apenas traduz o que se passa no mundo eletromagnético. Se você inverter as conexões nos polos duma bateria, o eletroímã vai rodar no sentido inverso:

$$\vec{j} \times \vec{k} = \vec{i}; \vec{k} \times \vec{j} = -\vec{i}; \quad (293)$$

Uma propriedade que tem uso intenso é a desigualdade de Cauchy-Bunyakovsky-Schwarz

$$| \langle P, Q \rangle | \leq \|P\| \|Q\| \quad (294)$$

que tem uma demonstração simples no plano, mas eu já observei que, dados dois vetores, eles determinam um plano do espaço a que eles pertencem portanto demonstrá-la num plano não é uma restrição.

Para isto preciso mostrar que a forma trigonométrica da definição na equação (eq. 289) é idêntica à expressão da soma de produtos na mesma equação porque

$$| \langle P, Q \rangle | = \|P\| \|Q\| |\cos(\gamma)| \leq \|P\| \|Q\|;$$

**Dem**:

Se  $n = 2$  é consequência direta da fórmula de Euler porque, dividindo pelo módulo dos vetores teremos:

$$\langle \frac{P}{\|P\|}, \frac{Q}{\|Q\|} \rangle = \frac{|\langle P, Q \rangle|}{\|P\| \|Q\|} \leq 1; \quad (295)$$

$$\langle \frac{P}{\|P\|}, \frac{Q}{\|Q\|} \rangle = \langle (\cos(\alpha), \sin(\alpha)), (\cos(\beta), \sin(\beta)) \rangle; \quad (296)$$

$$\langle \frac{P}{\|P\|}, \frac{Q}{\|Q\|} \rangle = \cos(\alpha - \beta) = \cos(\gamma); \quad (297)$$

Na equação (eq. 295) usei a propriedade das formas bilineares que permite a distribuição dos coeficientes  $\|P\|, \|Q\|$  entre os dois fatores<sup>13</sup>.

Na equação (eq. 296) usei uma representação plana dos dois vetores que podem ser expressos como pontos do círculo trigonométrico sendo  $\alpha, \beta$  os seus ângulos e  $\gamma$  é o ângulo entre eles. Na passagem da equação (eq. 296) para a equação (eq. 297) usei a expressão do produto escalar

$$\langle (\cos(\alpha), \sin(\alpha)), (\cos(\beta), \sin(\beta)) \rangle = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) = \cos(\alpha - \beta); \quad (298)$$

Os cálculos mostram que posso me reduzir, na demonstração, ao caso dos vetores unitários. Então, na sequência vou supor que  $\|P\| = \|Q\| = 1$ .

A passagem crítica, na demonstração por indução que está na passagem de  $n - 1$  para  $n$  que é semelhante à passagem da dimensão 2 para 3, mas esta é mais intuitiva. Deixe-me agora considerar  $P, Q \in \mathbf{R}^3$ , vetores unitários, portanto pontos de  $\mathbf{S}^2$  a esfera unitária do  $\mathbf{R}^3$ . Mas estes dois vetores determinam um plano que corta  $\mathbf{S}^2$  segundo um círculo unitário, uma curva plana, e, pela demonstração anterior,  $\cos(\gamma)$  é o valor do produto escalar entre eles, em que  $\gamma$  é ângulo entre os vetores.

Mas esta demonstração não seria necessária dentro do escopo duma demonstração por indução, ela é simplesmente tomada como verdadeira para todo  $N \leq n - 1$  e temos que considerar agora  $P, Q \in \mathbf{R}^n$ , dois vetores unitários. Mas estes vetores definem um plano do  $\mathbf{R}^n$  e o valor do produto escalar para dois vetores, no plano, em dimensão  $N < n$ , por hipótese de indução, é  $\cos(\gamma)$ , em que  $\gamma$  é o ângulo entre os vetores, o que termina a demonstração.

**q.e.d.**

- **distribuição** Em 1945, aproximadamente, Laurent Schwartz, um matemático francês, conseguiu sintetizar, simultaneamente com um matemático português,

<sup>13</sup>as forma bilineares generalizam o produto de números que um exemplo de forma bilinear.

Sebastião e Silva, a Teoria das<sup>14</sup> Distribuições, resolvendo uma situação incômoda criada pelo físico Paul Adrien Maurice Dirac, que, ao formular a *mecânica quântica* precisou de criar uma unidade para o produto de convolução que tinha que ser uma função nula em todos os pontos da reta, exceto na origem onde seria infinita, e com integral igual a 1, que durante muito tempo se chamou de “função de Dirac”.

A *função de Dirac*, com esta propriedades, não pode ser uma função sem colocar em cheque todas as teorias de integração existentes, mas Dirac respondia que isto não era seu problema, que os matemáticos corrigissem as teorias de integração porque para ele tudo funcionava perfeitamente bem.

A descoberta de Schwartz e de Sebastião e Silva, ambos pesquisadores de equações diferenciais parciais, resolveu o problema criado por Dirac criando um novo objeto matemático, a *distribuição* mostrando que a chamada *função de Dirac*, é uma distribuição que é a derivada da função de Rademacher  $H_0$ , a função que é zero se  $x \leq 0$  e 1 se  $x > 0$

$$H_0(x) = \begin{cases} 0 & x \leq 0 \\ 1 & x > 0 \end{cases} \quad (299)$$

No seu livro intitulado *Téorie des Distributions*, Laurent Schwartz, construiu uma generalização do Cálculo Diferencial e Integral usando as distribuições como elemento em lugar das funções e é em função deste livro que o trabalho de Sebastião e Silva ficou na sombra durante muito tempo.

Embora Laurent Schwartz tenha durante algum tempo carregado sozinho os louros da construção da teoria, e possivelmente ele tenha sido o “inventor” do nome, a ideia já estava latente desde o século 19 e a teoria recupera nomes como de Cauchy que já havia pensado em um objeto que atuasse sobre classes de funções, uma distribuição é um funcional linear, criando uma distribuição chamada *valor principal* que era o nome que Cauchy dava um certo de integral que não teria sentido usual de uma integral e que, como a “função de Dirac”, encontrou um lugar preciso dentro da teoria das distribuições.

A teoria das distribuições, durante algum tempo eletrizou todas as atenções dos que estudavam equações diferenciais parciais porque parece que se havia descoberto a teoria final... não foi, mas marcou profundamente toda a linguagem com que se fala hoje as equações diferenciais parciais, criando uma forma nova de resolvê-las, as *soluções fracas* que é uma solução-distribuição.

- **divisão euclidiana** é a forma tradicional que temos para dividir um número natural,  $p$ , chamado *dividendo* por outro,  $d$ , chamado *divisor*

$$p = dq + r; p, d, q, r \in \mathbf{N}; r < d; q > 0 \quad (300)$$

Esta equação gera um *método* em que divisões sucessivas nos permitem encontrar  $q, r$  e algumas vezes é este método que é designado como *algoritmo da divisão euclidiana*.

<sup>14</sup>Uma distribuição é também chamada de *função generalizada* e esta denominação remonta a Euler, Cauchy e outros.

- **divisões da Matemática**. Não há um consenso sobre as *grandes divisões da Matemática* que alguns consideram como sendo, Álgebra, Análise, Estatística, Geometria, Lógica, Topologia

Tão pouco estas divisões conseguiriam descrever o escopo muito grande que a Matemática atingiu nos últimos 100 anos, e basta ver o grande projeto, falido, infelizmente, Bourbaki, que um grupo de matemáticos franceses encetou na década de 40, quando, prudentemente, criaram uma figura, Nicolas Bourbaki, sob a qual se esconderam para reescrever toda a Matemática. Dizer que o projeto é falido não significa que ele não tenha tido uma influência significativa no desenvolvimento da Matemática, mas significa que ele não conseguiu preencher seus objetivos. De certa forma o projeto nasceu morto, um poco antes, 1931, Gödel havia demonstrado que seria impossível “descrever” de forma completa os *números naturais* que foi a base do monumental tratado de Russel e Whitehead, *Principia Mathematica*, escrito em 1910.

A American Mathematical Society mantém uma tabela das áreas da Matemática, *1991 Mathematics Subject Classification*, [1] que é quase universalmente utilizada como descritiva de toda a Matemática contendo 100 grandes itens com os quais pretende descrever toda a atividade Matemática hoje conhecida.

- **domínio** é um conjunto aberto e conexo de um espaço topológico. Uma bola aberta  $\mathcal{B}(a, \rho)$ ;  $\rho > 0$  é um exemplo de domínio num espaço métrico. Uma reunião de bolas que se interceptem duas a duas, com interseção diferente do vazio, é um outro exemplo num espaço métrico. A figura (fig 43), página 110, mostra dois exemplos de domínio num espaço métrico.

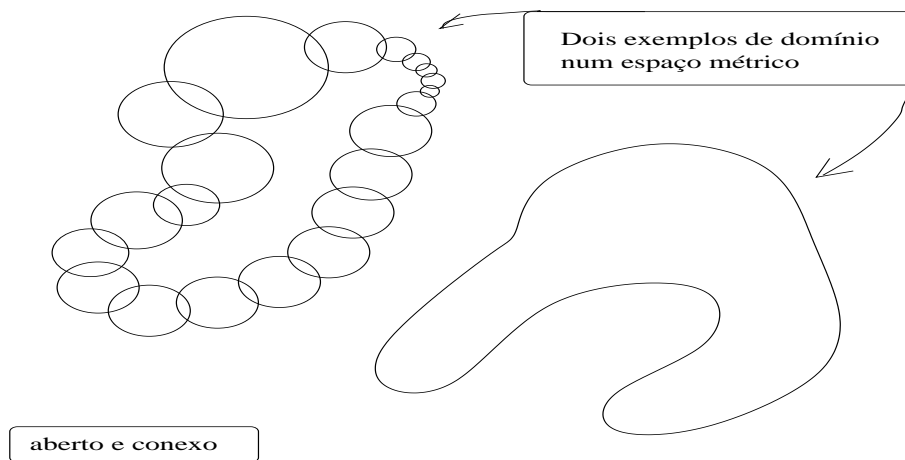


Figura 43: domínio num espaço métrico

-  $e$  o número “e” uma das constantes clássicas,  $e \approx 2.71828182845904523536$ ,



valor obtido com `calc` executando, num terminal, `exp(1)`. Este número é a base do *logaritmo neperiano* e as duas funções

$$y = \ln(x); y = e^x; \text{ um par de funções inversas ;} \quad (301)$$

$$\mathbf{R}^{++} \longrightarrow \mathbf{R}; x \mapsto \ln(x); \quad (302)$$

$$\mathbf{R} \longrightarrow \mathbf{R}^{++}; x \mapsto e^x = \exp(x) \quad (303)$$

Este número aparece na fórmula de Euler

$$e^{it} = \cos(t) + i \sin(t); t \in \mathbf{R}; \quad (304)$$

que é um caso particular da exponencial complexa,

$$z = x + iy; e^z = e^x e^{iy}; \quad (305)$$

confira *exponencial complexa*.

A definição deste número pode ser

$$\int_1^e \frac{dx}{x} = 1; \ln(t) = \int_1^t \frac{dx}{x}; \quad (306)$$

e o seu cálculo é um dos *limites notáveis* porque podemos demonstrar, usando o sistema de equações (eq. 306) que  $\ln$  é uma função indefinidamente diferenciável e portanto também a sua inversa o é. Como  $\exp$  tem por domínio  $\mathbf{R}$  então tem uma série de Taylor cujo raio de convergência é infinito então a série converge em qualquer ponto da reta e assim podemos escrever

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (307)$$

e qualquer reduzida desta série produz uma aproximação para  $e^x$ , quando  $x = 1$  podemos calcular  $\exp(1)$ . Fazendo esta soma com `calc`

```
define expo(x){
local k=0, soma =0, n=100;
while (k<=n){
    soma = soma + power(x,k)/fact(k);
    k +=1;
}
return soma;
}
```

```
expo(1);
```

```
define expon(x){
local n = 10000;
```

```
return power(1 + 1/n, n);
}
```

```
expon(1);
expo(x) defined
~2.71828182845904523536
expon(x) defined
2.71814592682522486404
```

Coloque este texto no arquivo `E.calc` e execute num terminal

```
calc < E.calc
```

supondo que a linguagem `calc` esteja instalada, o que é fácil de ser verdade uma vez que se trata duma linguagem livremente distribuída. Confira *programação*.

Em `calc` a função `fact` está definida e parando em  $k = 100$  obtive o valor aproximado 2.71828182845904523536 para  $e$  atingindo a precisão com que `calc` escreve `exp(1)`. Porém esta série converge muito lentamente uma vez que o cálculo do fatorial consome muito tempo de computação mesmo que o fatorial seja uma função natural da linguagem, como é o caso de linguagem `calc`.

O argumento do *raio de convergência* pode ser substituído pelo cálculo do limite notável

$$\lim_n (1 + 1/n)^n = e; \lim_n (x + x/n)^n = e^x \quad (308)$$

- **EDO** é uma sigla que representa *equações diferenciais ordinárias* um tipo de equação em que uma das operações é a derivada e portanto a “incógnita” é uma função. É uma equação funcional.

As equações diferenciais aparecem logo nas primeiras aulas de um curso de Cálculo Diferencial quando se procura saber qual é a função  $F$  que corresponde a  $f$  tal que  $F' = f$ . Neste contexto estas duas funções recebem os nomes

- $F$  é uma primitiva de  $f$ , ou ainda  $F' = f$ ;
- $f$  é a derivada de  $F$

Há várias formas de entender a “indefinição” da frase  $F$  é uma primitiva de  $f$ , significando que  $f$  tem muitas primitivas. A mais simples vem das expressões

$$\int_a^x f(t)dt = F_1(x); \quad (309)$$

$$\int_b^x f(t)dt = F_2(x); \quad (310)$$

$$a < b; \quad (311)$$

$$(312)$$

cuja interpretação gráfica aparece na figura (44) página 113. A equação (309)

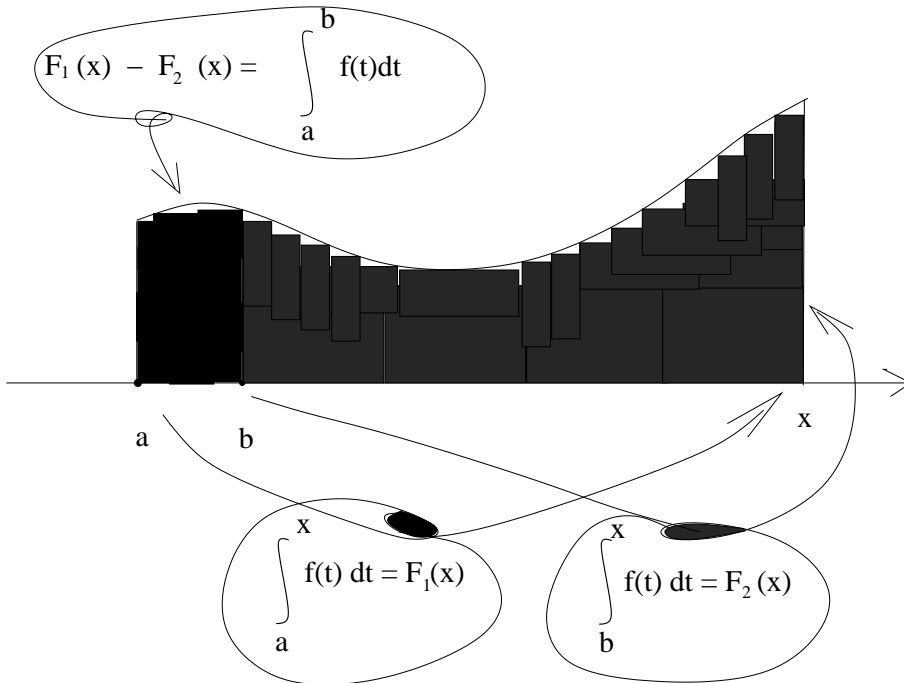


Figura 44: Diferença entre duas primitivas é uma constante

define a função  $F_1$  e a equação (310) define a função  $F_2$ . Estas duas áreas diferem pela *condição inicial*,  $F_1$  está definida com a *condição inicial*  $\underline{a}$  e  $F_2$  está definida com a *condição inicial*  $\underline{b}$  e a diferença entre as duas áreas é a constante

$$\int_a^b f(t) dt = C; F_1(x) = F_2(x) + C; \quad (313)$$

Partindo deste exemplo simples de *equação diferencial ordinária* podemos entender que as equações diferenciais tem uma infinidade de soluções como consequência da escolha da *condição inicial*.

Ainda um outro exemplo de equação diferencial ordinária pode ser a curva de crescimento biológico.

Cada espécie dos seres vivos tem uma curva específica de crescimento, é uma primitiva da “velocidade de crescimento” específica da espécie.

Dois seres da mesma espécie podem ter curvas diferentes, mas, se o “crescimento de ambos for normal”, ambos tiverem a *velocidade de crescimento normal de sua espécie*,  $f$ , a diferença entre as curvas de crescimento vai ser uma constante, a figura (45), página 114, mostra alguns exemplos de *curvas de crescimento* partindo de um ponto no eixo vértical que é o *tamanho ao nascer* do ser vivo a que corresponde a curva.

Curvas de crescimento

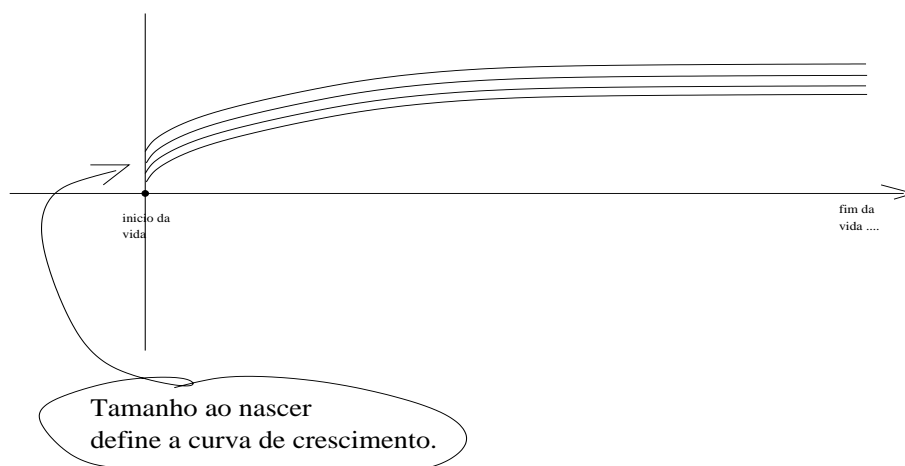


Figura 45: Curvas de crescimento diferem de uma constante

Este exemplo mostra a importância das equações diferenciais no estudo dos seres vivos e de certa forma *dramático* do ponto de vista social: dois seres humanos que ao nascer tenham tido condições diferentes de gestação, alimentação diferenciada da mãe, condições habitacionais da mãe, entre muitas outras condições, resultam em crianças que vão nascer com tamanhos diferentes: vão ter selecionada, ao nascerem, sua curva de crescimento, mas não apenas esta, a curva de conhecimento, as diversas curvas de desenvolvimento. As condições iniciais determinam de forma inexorável o resto da vida, porque a “derivada é a mesma”, uma característica da espécie, mas as curvas de desenvolvimento, são as primitivas que dependem *das condições iniciais*.

Este exemplo da biologia mostra que existem equações diferenciais cujas soluções dependem de várias variáveis, cada uma das variáveis com “velocidades” distintas, as *derivadas parciais* do crescimento.

A denominação, *equação diferencial ordinária*, se deve a que neste tipo de equação as soluções são funções univariadas, dependem de uma única variável. Existem também as *equações diferenciais parciais* que descrevem o comportamento de funções multivariadas.

Embora estas equações já apareçam muito cedo no estudo de Matemática, no primeiro ano de Cálculo, quando surgem no cálculo de primitivas, a referência às equações diferenciais em geral não é feita neste momento ficando adiada para um terceiro curso de Cálculo.

Uma forma aparentemente simples de definir uma equação diferencial or-

dinária é expressá-la sob a forma

$$P(x, y, y', \dots, y^{(n)}) = 0 \quad (314)$$

em que  $P$  é uma expressão envolvendo as operações da álgebra, tendo como “coeficientes” constantes (números reais ou complexos) ou *funções elementares* e as derivadas de uma função que é a incognita da equação representada na equação (314) pelo símbolo  $y$ .

Se os coeficientes forem constantes, se fala de uma equação diferencial *a coeficientes constantes*, e no outro caso se diz uma equação diferencial *a coeficientes variáveis*.

O *teorema da função implícita* estabelece as condições para que se possa escrever esta equação explicitando a maior derivada o que pode conduzir a uma expressão conhecida (solução conhecida).

A maior derivada caracteriza a ordem da equação. Podemos transformar e equação (314) num sistema de  $n$  equações de primeira ordem e desta forma uma equação diferencial ordinária de ordem  $n$  é equivalente a um sistema de equações de primeira ordem. Se nesta transformação se obtiver um *sistema linear* então a equação se chama *equação diferencial linear*. Se o sistema assim resultante não for linear, temos uma *equação diferencial não linear*.

Sabemos tudo sobre as *equações diferenciais lineares* exceto resolvê-las! A teoria das equações diferenciais lineares está praticamente pronta e a dificuldade na solução depende de que outras teorias não estejam prontas ou dificilmente algum dia estarão, como a teoria das equações algébricas, ou a Álgebra Linear. Obviamente, o defeito não se encontra na teoria das *equações diferenciais lineares*...

Como é possível transformar a equação (314) num sistema de equações de primeira ordem, torna-se importante saber resolver as equações de primeira ordem. Mesmo aqui o problema não é fácil e ainda está muito longe de ser resolvido, uma simples troca de “coeficiente variável” altera completamente o comportamento da equação. Este problema é bem descrito numa teoria que vem se desenvolvendo muito e que tem representação significativa dentro da Matemática brasileira, chamada de *sistemas dinâmicos*.

Se considerarmos os casos simples da teoria, exatamente os que são considerados nos cursos de Cálculo, podemos descrever de maneira muito elegante o que acontece com uma equação diferencial. Esta descrição terá que ser ajustada posteriormente, mas serve como descrição inicial.

Toda função *razoavelmente bem comportada*,  $f$ , é uma *derivada* e o cálculo de uma primitiva é feita com a fórmula

$$\int_a^x f(t)dt = F(x); \quad (315)$$

em que  $F$  é a primitiva de  $f$  associada á condição inicial  $\underline{a}$ . A figura 44 que pode ser vista na página 113, mostra que a diferença entre duas primitivas é uma constante neste caso. Mas, em geral, nem sempre é uma translação

como no caso do Cálculo, porque podem surgir pontos críticos que mascaram esta propriedade. Isto se consegue entender bem dentro da teoria dos sistemas dinâmicos.

Entretanto podemos dizer que as equações diferenciais ordinárias de primeira ordem tem *uma condição inicial* e a escolha desta condição inicial seleciona as distintas primitivas.

A denominação *ordinária* é um exemplo de preconceito, ou de vocabulário mal utilizado que domina a linguagem científica que esperamos que seja imparcial mas não pode ser uma vez que é produzida por seres que tem sentimento e individualidade. É possível que a palavra *ordinária* aqui tenha tido um significado de *mais fácil* ou *simples* com alguma razão porque a solução das equações diferenciais parciais com frequência depende da solução de alguma equação diferencial ordinária e elas podem ser consideradas mais difíceis.

Um exemplo,  $y - y' = 0$  esta equação é resolvida no cálculo por um longo circuito que passa pela função logaritmo. Uma outra forma de resolvê-la passa pela transformação

$$y - y' = y - \frac{dy}{dx} \Rightarrow y = \frac{dy}{dx} \quad (316)$$

$$dx = \frac{dy}{y} \quad (317)$$

O algebrismo contido na passagem da equação (316) para a equação (317) é muito controverso. No Cálculo com frequência as professores dizem que que o símbolo  $\frac{dy}{dx}$ , a notação de Leibniz é indivisível e não se deve nele *ver* uma fração e nem dois termos. Neste momento estou fazendo o contrário, e funciona! mas obtive uma nova expressão que se for colocado sob o sinal de integração, integrando ambos os membros na equação (317) nos conduz a

$$\int_a^t dx = \int_a^t \frac{dy}{y} = \int_a^t \frac{dx}{x} \quad (318)$$

que temos resolver de formas diferentes. No primeiro membro vemos a integral da função constante 1 e no segundo membro vemos a integral que resulta no logaritmo quando  $a = 1$ . Também reescrevi a integral no segundo membro usando a “letra x”, observe que ela representa a mesma integral. Usando a condição inicial  $a = 1$  temos (318)

$$\int_1^t dx = t - 1 = \int_1^t \frac{dx}{x} = \ln(t) - \ln(1) = \ln(t) \quad (319)$$

$$t = \exp(t - 1) = \frac{e^t}{e} \quad (320)$$

$$y = \frac{e^x}{e} \quad (321)$$

é uma solução da equação diferencial. Como  $(Ky)' = Ky' = Ky$  então vemos que a constante,  $e$  que surgiu na equação poderia ser qualquer uma, e na verdade esta que apareceu é consequência da *condição inicial* usada. Vemos que as soluções desta equação são da forma

$$y = K e^x \quad (322)$$

Apenas temos que analisar se realmente  $K$  pode ser qualquer.  $K = 0$ , por exemplo serve. A que encontramos,  $e$  é positiva, e qualquer constante positiva também serve. Se  $K < 0$  também funciona e a figura (fig 46), página 117,

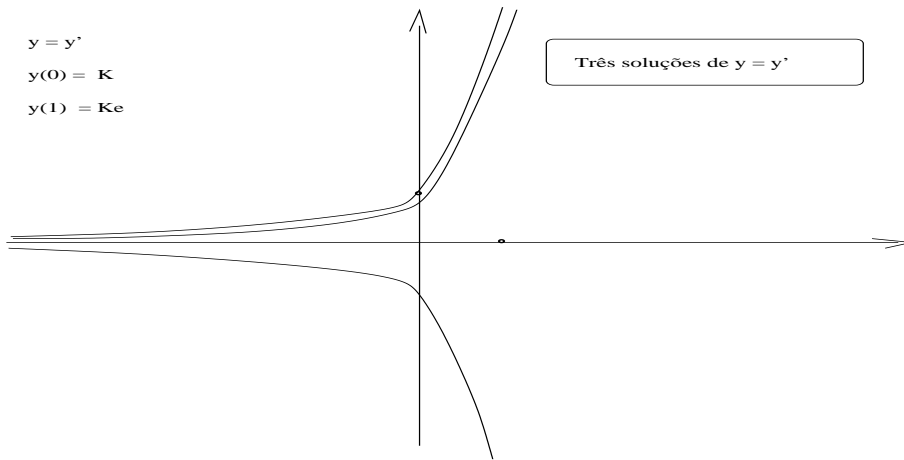


Figura 46: três soluções dd  $y - y' = 0$

Este exemplo mostra bem as dificuldades que envolveram as soluções das equações diferenciais no século 18 e que ainda faziam parte da dificuldade linguística dos textos de Matemática até a metade do século passado. Hoje nós temos vários métodos para enfrentar a contradição registrada na passagem da equação (316) para a equação (317) e que não está sendo considerada aqui. Ela será melhor entendida quando você passar pela *derivada implícita*, neste dicionário.

---

- **EDP** Equações à derivadas parciais. Uma grande quantidade de fenômenos que nos rodeiam podem ser descritos como soluções de equações à derivadas parciais, o que torna este ramo do conhecimento de grande importância. Um exemplo simples foi descrito em *equações diferenciais ordinárias* das *curvas de desenvolvimento de uma espécie*, aliás, mostrando a relação entre equações diferenciais ordinárias e parciais.

Um outro exemplo bem ligado à *vida real* é a difusão do calor, uma boa geladeira seria aquela em que apenas se perdesse calor ao se lhe abrir a porta, o que depende de uma boa isolamento. Uma geladeira é um tubo,  $\Omega$ , rodeado de calor por todos os lados com um material isolante protegendo sua *fronteira*,  $\partial\Omega$ , que impede que o calor se difunda para dentro do tubo.

Neste exemplo você encontra alguns dos termos mais comuns no estudo das equações diferenciais parciais: o domínio (o tubo),  $\Omega$ , a fronteira do domínio,  $\partial\Omega$ . A equação que nos interessa aqui (neste exemplo) é a equação de Laplace

ainda chamada *equação do calor*

$$(x, y, z) \in \Omega; \Delta(u(x, y, z)) = k; \quad (323)$$

$$(x, y, z) \in \partial\Omega; u(x, y, z) = f(x, y, z); \quad (324)$$

$$k \text{ é a temperatura constante dentro da geladeira}; \quad (325)$$

$$f \text{ é a temperatura variável em volta da geladeira}; \quad (326)$$

O sistema de equações (323)- (326) é chamado de *um problema*. A equação (326) se chama de condição de fronteira, e a função (obtida experimentalmente) que descreve a distribuição do calor na fronteira da geladeira,  $\Omega$ . A equação (323) é a distribuição do calor dentro da geladeira, que neste exemplo se espera que seja uma constante (ou tenha dois valores constantes, um deles no congelador).

Mas este exemplo precisa ser complicado um pouquinho mais para tornar-se mais real! Observe que o calor em volta da geladeira  $\Omega$ , não é constante, depende do tempo, e assim, para chegarmos a um modelo mais preciso para o problema da geladeira temos que considerar a variação da temperatura ao longo do tempo:  $\frac{\partial u}{\partial t}$  e conseqüentemente também a função que descreve a temperatura á volta da geladeira  $\Omega$  seria  $w = f(t, x, y, z)$  e desta forma chegamos a equação do calor,

$$(t, x, y, z) \in [a, b] \times \Omega; \frac{\partial u}{\partial t} = \Delta(u(x, y, z)); \quad (327)$$

$$(x, y, z) \in \partial\Omega; t \in [a, b]; u(t, x, y, z) = f(t, x, y, z); \quad (328)$$

Resolver esta equação significa encontrar uma descrição, uma equação, para a função  $u$ . Com frequência isto somente é possível aproximadamente, então  $u$  pode ser representada por um programa de computador que é obtido experimentalmente. O *problema* está agora mais bem colocado nas equações (327)-(328). Este é um *problema de valor na fronteira* representado nas equações (327)- (328), (PVF), usualmente referido pela sigla em inglês, BVP - *boundary value problem*. Na literatura se *complica*<sup>15</sup> um pouco mais a notação colocando  $u$  em evidência escrevendo-se

$$(t, x, y, z) \in [a, b] \times \Omega; F(u) = \left(\frac{\partial}{\partial t} - \Delta\right) (u) = 0; \quad (329)$$

$$(x, y, z) \in \partial\Omega, t \in [a, b]; u(t, x, y, z) = f(t, x, y, z); \quad (330)$$

em que nós entendemos que um operador  $F$  foi aplicado à função  $u$ . Esta abstração tem utilidade, existe uma *teoria de operadores* em que se conhecem diversas propriedades dos operadores tornando mais ágil o estudo das equações diferenciais, ou pelo menos criando uma linguagem mais padronizada em que elas possam ser expressas, ou, mesmo ainda, permitindo que os algoritmos sejam mais uniformes e portanto se consiga melhores programas de computador para representar as soluções coisa que de fato vem acontecendo: as equações diferenciais parciais, hoje, estão sendo resolvidas computacionalmente e na década de 80 havia apenas o sentimento da impossibilidade de resolvê-las.

<sup>15</sup>Complicar não é um defeito, consiste realmente em acrescentar mais dados para obter uma expressão mais geral, se houver uma teoria que abrigue esta formulação, e neste caso tem. É uma abstração, ou classificação.



No caso da equação do calor é bem simples mostrar como esta generalização representa uma linguagem adequada. Observe que na equação (329) ocorrem três tipos de operação, derivadas parciais de segunda ordem, derivada parcial de primeira ordem, e a adição. Estas três operações podem ser concatenadas na expressão  $P(t, x, y, z) = t - x^2 - y^2 - z^2$  e se “codificarmos” *derivada com potência* podemos “recodificar” o operador  $F$  usando o polinômio  $P$ , escrevendo  $F(D) = P(D)$  o que nos permite rescrever a equação (329) na forma

$$F(u) = P(D)(u) = 0; (t, x, y, z) \in [a, b] \times \Omega \quad (331)$$

$$u(t, x, y, z) = f(t, x, y, z); (x, y, z) \in \partial\Omega; t \in [a, b] \quad (332)$$

Você poderia argumentar que apenas houve uma *complicação inútil* com esta “codificação”, e, aparentemente, teria razão. Entretanto, observe agora que sabemos do Cálculo que a operação de derivação é *linear*, e com isto queremos dizer que  $(f + g)' = f' + g'$ ;  $(\alpha f)' = \alpha f'$ , propriedades fundamentais da Álgebra Linear aparecem aqui! Como o polinômio  $P$  é uma soma de operações linear, então o operador  $P(D)$  herda a linearidade se tornando um *operador linear*. Quando o operador que define uma equação diferencial for linear, diremos que se trata de uma *equação diferencial linear*. Com isto derrubei sua argumentação de que se tratava de uma *complicação inútil*, porque toda a Álgebra Linear vem em nossa ajuda para resolver as *equações parciais diferenciais lineares*, LPDE, EDPL tornando muito válida esta complicação. Este é um pequeno exemplo mostrando a linguagem com que estudamos as equações diferenciais parciais hoje, com auxílio da teoria dos operadores. Melhor, até no estudo das equações não lineares esta *codificação* nos traz bons resultados! Nós continuamos sem resolver as equações diferenciais, mas as entendemos muito melhor e avanços consideráveis foram feitos nas soluções aproximadas usando esta *complicada codificação*.

Um adendo, as equações que realmente nos interessam não são lineares, como as importantes equações de Navier-Stokes que são não lineares. Mas como no caso do Cálculo, uma equação não linear é aproximada por uma equação linear. Isto nem sempre ajuda, mas é um primeiro passo no processo de solução aproximada.

Tópicos importantes de EDP, são, equação do calor, equações de Maxwell, equação da onda, equação de Poisson, equações de Navier-Stokes, problema de Cauchy, problema de Dirichlet . . .

As equações de Navier-Stokes se encontram em grande evidência, hoje, porque elas descrevem o movimento dos flúidos e se encontram por trás do estudo das variações climáticas. Esta lista está longe de ser completa!

- **eixos coordenados** é um tópico da Geometria Analítica plana, também chamado de *Plano Coordenado* ou *Plano Cartesiano*. O plano é um ente geométrico de dimensão dois, dimensão é um conceito estudado em Álgebra Linear, mas é possível criar uma visão intuitiva do que pode significar “dimensão dois” com auxílio de exemplos.

1. **dimensão 0** O ponto é o mais elementar dos objetos geométricos e sua

dimensão é zero. Apoie um lápis numa folha de papel e você tem uma visão, falsa, dum ponto. . . é falso porque você imprimiu no papel um objeto tridimensional, tem altura, largura e comprimento! Mas estamos tratando intuitivamente, e a intuição passa por erros!

2. dimensão 1 Se agora você mover o lápis, “o ponto se movendo” vai traçar uma curva. Se você fizer isto com auxílio duma régua o resultado será um *segmento de reta*. Exemplos de curvas são círculos, parábolas, retas. São os objetos de dimensão um e apenas algumas curvas mais importantes é que tem nomes.
3. a reta numérica. Há uma diferença entre segmento de reta e uma reta. A reta é uma curva infinita, coisa impossível de obter com papel e lápis! Mas estamos trabalhando intuitivamente, imagine então que o segmento de reta que você traçou com lápis e régua se expande indefinidamente.
  - Nesta reta escolha um ponto, chame-o de 0, e estamos *formatando* a reta que agora ficou dividida em duas semiretas.
  - Numa das semiretas selecione outro ponto, chame-o de 1 e você definiu qual é a *semireta positiva*. Você também definiu um *segmento de reta com medida 1*. Usando este segmento de medida 1, você pode identificar na reta todos os inteiros positivos, na mesma semireta positiva em que você marcou 1, e os inteiros negativos, na semireta negativa. Esta reta é a reta numérica, confira a figura (fig 47), página 121.
  - Usando semelhança de triângulos é possível também encontrar a posição de qualquer número racional na reta numérica.
4. dimensão 2 Deixe-me supor que você usou uma régua e construiu a *reta numérica*, agora coloque a régua perpendicularmente ao segmento de reta traçado e trace outra reta. Você tem agora um sistema de duas retas perpendiculares como na figura (fig 47), página 121, Este conjunto deixa de ser simplesmente um plano da geometria, agora é uma *representação* do produto cartesiano  $\mathbf{R} \times \mathbf{R}$  ou como também é comum dizermos  $\mathbf{R}^2$ , é o *plano numerizado* ou seria o plano digitalizado. Na linguagem técnica dizemos *plano numérico*. Na figura (fig 47) você pode ver os pontos

$$(-2, -2), (-1, -1), (3, -3), (1, 1), (3, 3), (3, -3), (5, 1)$$

representados no *plano cartesiano*.

O plano é um objeto de dimensão dois que contém curvas e pontos.

5. dimensão 3 O espaço em que vivemos é de dimensão três e nele se podem conter planos, curvas e pontos. Podemos também *numerisar* também o espaço tridimensional se acrescentarmos uma reta numérica, passando pela origem  $(0, 0)$ , perpendicular ao plano numérico. Com isto associamos endereços no formato  $(x, y, z)$  para qualquer ponto do espaço.

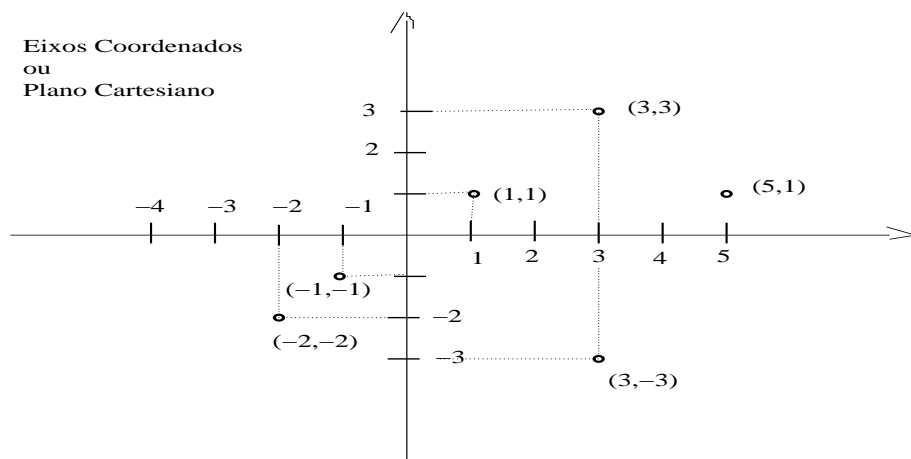


Figura 47: O plano cartesiano ou eixos coordenados

6. dimensão 4 e nós vivemos imersos n'algum espaço de dimensão 4 cujos objetos fogem a nossa compreensão ou capacidade física de perceber completamente. Poderíamos ter uma ideia distorcida d'alguma interseção destes objetos com o espaço em que vivemos. Por exemplo, uma *esfera* da dimensão 4 entrando em nosso espaço seria visível para nós como uma esfera, mas um parabolóide da dimensão 4, também poderia ter como interseção uma esfera, na nossa dimensão, dependendo da forma como se desse esta interseção. Não saberíamos se teria sido um esfera ou um parabolóide de dimensão 4 que teria entrado no nosso espaço. . . Poderíamos também *numerisar* o espaço se conseguíssemos descobrir a direção perpendicular ao  $(0,0,0)$  para neste ponto colocarmos uma reta perpendicular e assim associarmos endereços no formato  $(x, y, z, w)$  para qualquer ponto do espaço  $\mathbf{R}^4$ .

Na verdade somente podemos perceber os objetos tridimensionais, foi por esta razão que acima está dito que o ponto é uma falsidade. . . Um ponto, de verdade, não é visível para nós, entes tridimensionais, como também não podemos perceber nenhuma curva . . . Os pontos e curvas que representamos numa folha de papel são todos objetos tridimensionais que *imaginamos perceber* os objetos de dimensão *dois*, *um* ou de *dimensão zero*. Somente a Matemática é capaz de nos libertar da prisão tridimensional em que vivemos.

Se diz que René Descartes teria sido o idealizador da Geometria Analítica ao criar esta representação algébrica para os pontos do plano e muitas vezes chamamos este plano numerizado de *plano cartesiano*

No plano cartesiano podemos identificar objetos geométricos e descobrir a equação algébrica de alguns destes objetos. Confira

1. equação do círculo;

2. equação da elipse
3. equação da hipérbole
4. equação da parábola
5. equação do plano;

---

- **equação da elipse**, é uma expressão polinomial do segundo grau que descreve, no plano, uma elipse. É um item da *geometria analítica*.

A geometria analítica pode ser vista como uma introdução à *geometria algébrica* que é a parte da Matemática que estuda os objetos geométricos que pode ser obtidos por equações algébricas. A geometria analítica se especializa nas cônicas que são as curvas que podemos obter quando fazemos interseções dum plano com um cone.

A elipse é uma das cônicas e sua importância se encontra na Astronomia uma vez que é, aproximadamente, a forma das curvas descritas pelos astros em seus movimentos em torno dum polo de gravidade, como a Terra em volta do Sol ou da Lua em volta da Terra. Foi Kepler que corrigiu a teoria heliocêntrica de Copêrnico analisando os dados que herdou de Tycho Brae. Copêrnico afirmou que os astros percorriam uma órbita circular em volta do Sol.

Há várias formas de definir esta curva chamada elipse. Kepler, se beneficiando das medições que Tycho Brae fez ao longo de toda sua vida, conseguiu matar a charada que Copernico havia deixado quando criou a teoria *heliocêntrica*, que quase lhe custou a vida nas mãos da cúpula da igreja católica, que impunha a “teoria” *geocêntrica*. Copernico propoz que a rota dos planetas seriam círculos tendo o Sol como centro. Em sua teoria, entretanto, ficou o problema do movimento de Marte que parecia ter uma velocidade negativa relativamente a dos demais planetas, andava para trás. Analisando os dados coletados por Tycho Brae, Kepler concluiu que seriam elipses as rotas dos planetas em volta do Sol, e assim conseguiu resolver o problema deixado por Copêrnico com a rota de Marte que é uma elipse com *excentricidade* um pouco maior do que a da Terra<sup>16</sup>, e, em sua segunda lei, estabeleceu que “*se um planeta estivesse ligado ao Sol por uma linha, esta linha cobriria iguais áreas durante iguais intervalos de tempo*”. Observou também que o Sol estaria colocado num ponto fora do centro do movimento. Na figura (fig 48), página 123, você vê o caminho da Terra em volta do Sol, uma elipse em que o Sol ocupa um dos focos.

Um feito relativamente recente, pode motivá-la a compreender a equação da elipse. O módulo espacial Rosetta da agência espacial europeia, fez grande parte de sua viagem que durou 10 anos, para depositar um robot na superfície do cometa *67P/Churyumov-Gerasimenko*, usando a gravitação universal como força de propulsão. Este feito, que pode ser visto no link

---

<sup>16</sup>A excentricidade é a medida de como uma elipse distorce um círculo a ela relacionado, é a razão  $\frac{2c}{2a} = \frac{c}{a}$ , em que  $2a$  é soma das distâncias dum ponto sobre a elipse aos focos e  $2c$  é a distância entre os focos. No círculo  $c = 0$  zerando a excentricidade do círculo visto como elipse.

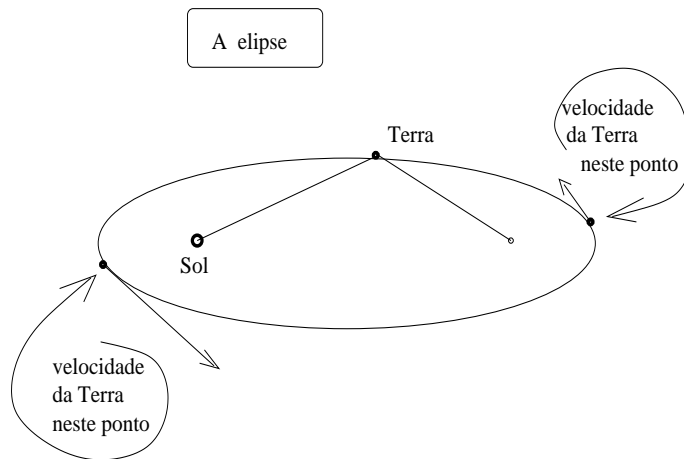


Figura 48: elipse é a rota da Terra em volta do Sol

<http://wp.me/psvm-2fu>

mostra como uma nave espacial pode obter propulsão da gravidade de um planeta para seguir sua viagem pelo Universo. Para fazer uso da gravidade é preciso compreender, entre outras coisas, a equação da trajetória dos astros e a elipse é um caso comum.

A definição geométrica da elipse é “o lugar geométrico dos pontos do plano cuja soma das distâncias a dois pontos fixos, chamados focos, é constante”. Então, você pode construir uma elipse com auxílio dum cordão com extremos fixos em dois pontos, os focos, fazendo a ponta dum lápis se movimentar mantendo o cordão esticado.

#### Cenário geométrico das cônicas

Vou mostrar que as cônicas se alteram na passagem entre círculos e parábolas. Por exemplo, uma elipse é apenas um círculo deformado e uma pequena deformação dada a um elipse faz surgir uma parábola. Inclusive isto tem aplicação na Astronomia ou melhor na exploração espacial em que uma nave espacial pode ser deixada hibernando numa série de elipses e repentinamente sair da elipse para uma parábola mudando a trajetória como foi o caso do projeto com a nave espacial Rosetta da agência espacial europeia. Uma pequena alteração num plano que produz uma parábola e o resultado é uma hipérbole. Você vai ver, também, que existem *hipérboles degeneradas*.

Vou começar discutindo a origem dos círculos e das parábolas.

Para estabelecer um *cenário* deixe-me fixar alguns “quadros” da geometria envolvendo um cone de duas folhas. Há duas classes de planos que delimitam os acontecimentos produzindo círculos e parábolas e depois você verá que elipses e hipérboles são variações destas duas curvas, na verdade *eventos estáveis* cujos limites são os *eventos instáveis círculos e das parábolas*. . . mas você não precisa dar muita importância a esta frase se ainda não souber o que são *sistemas dinâmicos*.

Vou também definir um *plano codificador* de onde vou deduzir um *círculo codificador* obtendo uma codificação mais efetiva e mais simples porque vai classificar melhor os eventos e descrever todas as possibilidades de cônicas.

- círculo qualquer plano perpendicular ao eixo do cone, eles cortam o cone segundo um círculo;
- parábola um plano tangente à superfície do cone, que contém uma imagem da *diretriz*, e qualquer plano paralelo a ele. Um tal plano corta a superfície do cone segundo uma parábola se não contiver a *diretriz*. O caso do plano tangente à superfície do cone que são aqueles que contém a *diretriz*, é uma *situação limite* que pertence ao “caso da hipérbole”. Eu vou discutir este caso sob o título *hipérbolas degeneradas*.

Estes dois pequenos exemplos mostram que preciso relacionar ângulos entre planos. Vou agora precisar de fazer referência aos ângulos que um plano faria com os planos das duas classes descritas com as etiquetas *círculo* e *parábola*. Determinamos o ângulo entre dois planos usando duas retas perpendiculares, cada uma delas, a um dos planos. É o ângulo entre estas retas que define o ângulo entre os planos.

Mas tem uma outra forma de codificar o cenário com uma linguagem que dominamos melhor, projetando num plano.

Existe *exatamente* um plano, deixe-me chamá-lo *plano codificador*, que é determinado pela geratriz e pelo eixo do cone. Posso codificar todo o cenário usando *este plano* como vou passar a fazer agora. O *plano codificador* é o que aparece na figura (fig 55), página 136, com a etiqueta 0 para a classe do *círculo* e a etiqueta 1 para a classe da *parábola*, contendo as imagens da geratriz e do eixo do cone. A geratriz é única, mas ela roda no espaço produzindo duas imagens no *plano codificador*.

Deixe-me descrever a figura (fig 55), página 136, e partir para uma especialização da codificação que vai agora ficar restrita apenas ao círculo que aparece na figura (fig 55), o *círculo codificador*. Este círculo tem o centro em um ponto  $Q$  do eixo do cone e raio qualquer. Apenas para obter uma figura menos poluída é bom selecionar o raio menor do que  $d(Q, O)$  em que  $O$  é ponto de concorrência da diretriz e o eixo do cone.

Nela você pode identificar seis pontos que foram etiquetados com os números  $\{0, 1, 3, 4, 5\}$  e que vou agora associar às cônicas que se podem obter.

- círculo a classe dos planos perpendiculares ao eixo marcada com os números 0, 3 e reduzida a uma reta perpendicular ao eixo do cone passando pelo centro do *círculo codificador* que aparece na figura (fig 55).

Qualquer reta perpendicular ao eixo do cone, na figura (fig 55) representa um plano desta classe que está agora reduzida as etiquetas 0, 3 do *círculo codificador*.

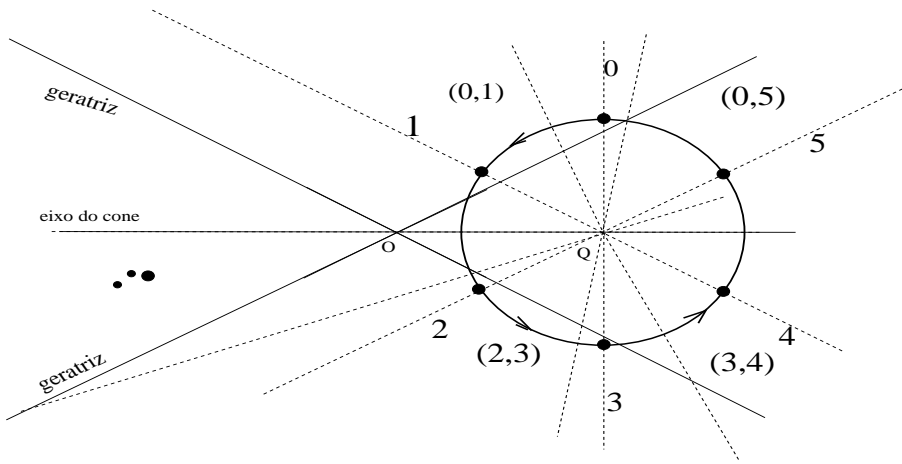


Figura 49: o círculo codificador

- parábola a classe dos planos cuja interseção produzem parábolas, marcada com os números 1, 2, 4, 5 e reduzida às duas retas paralelas à alguma das representações da geratriz, passando pelo centro do *círculo codificador* que aparece na figura (fig 55). A geratriz é única, mas ela roda no espaço e assim aparece no plano da figura (fig 55) duas vezes.

As imagens das diretrizes não pertencem a esta classe porque elas vão gerar *hipérbolas degeneradas* de que vou tratar mais a frente. Observe que círculos e parábolas são determinados por pontos isolados do *círculo codificador* em comparação com as elipses e hipérbolas que vão ficar determinadas por intervalos.

- o universo dos planos que nos interessam são os planos perpendiculares ao plano da (fig 55). Outra forma de fixar os planos interessantes é considerando a reta perpendicular ao *plano codificador*, (fig 55), passando pelo centro do *círculo codificador*: são os planos do espaço que contenham esta reta ou planos paralelos a estes.

O conjunto dos planos que contem um reta se chama *um feixe de planos*, portanto nos interessam todos os planos paralelos a um dos planos do feixe determinado pela reta perpendicular ao *círculo codificador* passando pelo seu centro.

Entenda os números 0, 1, 2, 3, 4, 5 como etiquetas no *círculo codificador* que se vê na figura (fig 55). E agora estou partindo para forma ainda mais especializada (e simples) de codificar usando apenas pontos ou setores do *círculo codificador* que aparece na (fig 55).

Estes números, 0, 1, 2, 3, 4, 5 não seriam as melhores escolhas para etiquetar, talvez melhor fosse usar o ângulo da diretriz com o eixo do cone, entretanto isto dependeria da escolha da diretriz e implicaria numa maior complexidade

da linguagem. A escolha que fiz usa a figura (fig 55) e portanto mostra, por semelhança de triângulos a dependência do ângulo que a diretriz faz com o eixo do cone. Enfim, fazer uma *codificação* implica em tomar uma decisão para escolher o código.

Todos os planos do universo que nos interessa, agora ficam classificados assim, usando apenas pontos ou segmentos de arco do *círculo codificador*:

- Os planos que produzem cônicas são os planos do feixe de planos determinado pela reta perpendicular ao *círculo codificador* passando pelo centro mesmo, ou planos paralelos a algum dos planos deste feixe.
- círculo os planos que determinam a reta etiquetada com 0,3 ou uma sua paralela.
- parábola os planos que determinam as retas etiquetadas com 1, 2, 4, 5, ou uma paralela a ela, excetuando as imagens da geratriz, confira o item hipérbole.
- elipse os planos que determinarem com o centro do *círculo codificador* uma reta nos intervalos (0, 1), (2, 3), (3, 4), (0, 5) Estas etiquetas significam mesmo um intervalo aberto porque os extremos do intervalo ficam para círculo ou parábola. Como observei acima, círculos e parábolas ficam determinadas por pontos do *círculo codificador*. Estas figuras são muito instáveis, por exemplo um objeto astronômico raramente se encontra sobre uma trajetória do tipo círculo ou parábola. . . se isto ocorrer será por uma instabilidade que deve ser corrigida.
- hipérbole os planos do feixe que pertencerem ao complementar das opções anteriores, ou ainda, quando um plano determinar uma reta que passando pelo centro do círculo que não esteja nas classes anteriores. Um plano da classe *parábola*, representa uma *situação limite* exatamente representado pelos planos tangentes ao cone, eles determina *hipérboles degeneradas* que se reduzem a duas semiretas representadas por qualquer das diretrizes. Eu vou voltar a esta questão na construção da equação da hipérbole.
- retas A interseção dum plano com o cone pode ser uma reta, se ele for tangente à superfície do cone. Tecnicamente, tais retas seriam hipérboles degeneradas. Observe que as hipérboles têm dois ramos disconexos, neste caso seriam duas semiretas conetadas na origem. . . os dois ramos duma hipérbole degenerada.

É um bom exercício corrigir as etiquetas deste modelo transformando-as em ângulos no forma  $\pi$ -radianos usando o ângulo  $\alpha$  entre uma das imagens da diretriz e o eixo do cone. O autor agradece se você lhe enviar esta a proposta e promete analisá-la para alteração do texto como uma contribuição sua.

O plano que corta o cone, na figura (fig 52), página 130, foi selecionado no feixe de planos de modo que a reta com que ele corta o *círculo codificador*



determine neste círculo um ângulo no intervalo  $(0, 1)$ , confira o *plano codificador* que aparece na figura (fig 55), página 136.

Para obter elipses basta alterar a posição do plano da posição 0.

Na figura (fig 52), página 130, você vê o corte do plano produzindo uma

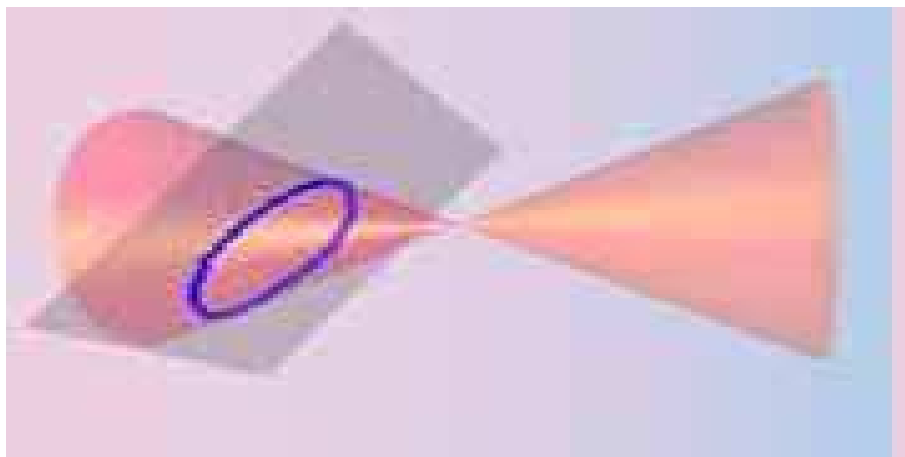


Figura 50: Plano corta cone segundo uma elipse

elipse.

Uma elipse é um círculo deformado, e vou mostrar como podemos desfazer a deformação para obter dados da elipse, como por exemplo a sua área. Então é importante compreender como se vai e volta duma elipse para um círculo, um exemplo de codificação e decodificação.

Vou considerar uma elipse desenhada com os dois focos sobre o eixo  $OX$  com a origem dos eixos no ponto médio dos pontos focais, como você pode ver na figura (fig 51), página 128. Na figura você também vê dois círculos, um deles inscrito na elipse, com raio  $b$  e a elipse está inscrita no outro com raio  $B$ .

Com esta definição vou obter os dados de medição da elipse.

- Chamando  $F_1 = (-c, 0)$ ,  $F_2 = (c, 0)$ ;  $c \geq 0$ , aos dois pontos focais, que se encontram sobre o eixo  $OX$  e considerando  $P = (x, y)$  um ponto genérico que descreve a elipse inicialmente localizado sobre o eixo  $OY$ .
- Nesta localização inicial escolhida, o ponto  $P = (x, y)$  que descreve a elipse, se encontra numa posição de simetria determinando dois triângulos retângulos semelhantes, com a origem e qualquer dos focos. A hipotenusa mede  $a$  e os dois catetos medem, respectivamente,  $b, c$ . O cateto que mede  $c$  é a metade da distância entre os focos, então a distância focal é convenientemente chamada de  $2c$ . Como o outro cateto mede  $b$ , pelo teorema de Pitágoras,

$$a^2 = b^2 + c^2;$$

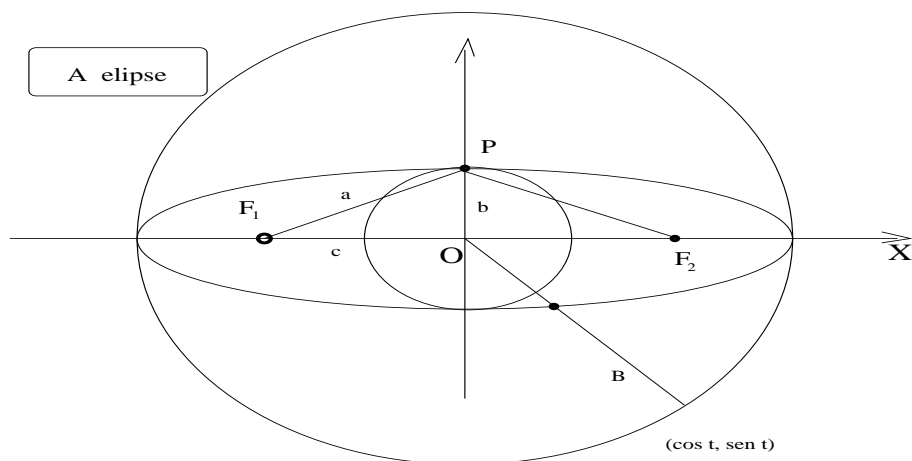


Figura 51: As medidas da elipse

- Se o ponto  $P$  estiver sobre o eixo  $OX$ , se conclui que a soma das distâncias de  $P$  aos focos é  $2B$ , o diâmetro do círculo de raio maior:

$$d(F_1, P) + d(P, F_2) = \overline{F_1O} + \overline{OP} + \overline{PF_2}; \quad (333)$$

$$c + B + |B - c| = \quad (334)$$

$$c + B + B - c = 2B; \quad (335)$$

$$d(F_1, P) + d(P, F_2) = 2B; \quad (336)$$

- Suponha agora que o ponto  $P = (x, y)$  se encontra numa posição arbitrária. Posso calcular a expressão da soma das distâncias aos focos:

$$d(P, F_1) + d(P, F_2) = 2B; \quad (337)$$

$$\sqrt{(x+c)^2 + y^2} + \sqrt{(x-c)^2 + y^2} = 2B; \quad (338)$$

$$\sqrt{(x+c)^2 + y^2} = 2B - \sqrt{(x-c)^2 + y^2}; \quad (339)$$

Elevando ao quadrado ambos os membros e simplificando os termos seme-

lhantes vem:

$$(x+c)^2 + y^2 = 4B^2 - 4B\sqrt{(x-c)^2 + y^2} + (x-c)^2 + y^2; \quad (340)$$

$$(x+c)^2 + y^2 - (x-c)^2 - y^2 - 4B^2 = -4B\sqrt{(x-c)^2 + y^2}; \quad (341)$$

$$4xc - 4B^2 = -4B\sqrt{(x-c)^2 + y^2}; \quad (342)$$

$$xc - B^2 = -B\sqrt{(x-c)^2 + y^2}; \quad (343)$$

$$(xc - B^2)^2 = B^2((x-c)^2 + y^2); \quad (344)$$

$$x^2c^2 - 2B^2xc + B^4 = B^2((x-c)^2 + y^2); \quad (345)$$

$$x^2c^2 - 2B^2xc + B^4 = B^2(c^2 + x^2 + y^2 - 2cx); \quad (346)$$

$$x^2c^2 + B^4 = B^2c^2 + B^2x^2 + B^2y^2; \quad (347)$$

$$x^2c^2 = -B^4 + B^2c^2 + B^2x^2 + B^2y^2; \quad (348)$$

$$x^2c^2 - B^2x^2 = -B^4 + B^2c^2 + B^2y^2; \quad (349)$$

$$x^2(c^2 - B^2) = B^2(c^2 + y^2 - B^2); \quad (350)$$

$$x^2(B^2 - c^2) = B^2(B^2 - c^2 - y^2); \quad (351)$$

Deixe-me chamar o número positivo  $B^2 - c^2$  de  $r^2$ :

$$r^2 = B^2 - c^2 = (B-c)(B+c) > 0$$

Posso simplificar a expressão na equação (eq. 351), repetida abaixo, com os seguintes cálculos:

$$x^2(B^2 - c^2) = B^2(B^2 - c^2 - y^2); \quad (352)$$

$$x^2r^2 = B^2(r^2 - y^2); \quad (353)$$

$$x^2 = B^2\frac{r^2 - y^2}{r^2}; \quad (354)$$

$$\frac{x^2}{B^2} = \frac{r^2 - y^2}{r^2}; \quad (355)$$

$$\frac{x^2}{B^2} = 1 - \frac{y^2}{r^2}; \quad (356)$$

$$\frac{x^2}{B^2} + \frac{y^2}{r^2} = 1; \quad (357)$$

$$\frac{x^2}{B^2} + \frac{y^2}{B^2 - c^2} = 1; \quad (358)$$

Esta não é a forma comum como aparece a equação da elipse na literatura, e sim com a seguinte expressão:

$$p = B^2; q = B^2 - c^2; \quad (359)$$

$$\frac{x^2}{p} + \frac{y^2}{q} = 1; \quad (360)$$

em que  $p, q$  são dois números positivos, aqui você está vendo de onde vêm estes dois números positivos associando-os aos dois círculos que elipse tangencia, por fora, o de menor raio, e por dentro o de maior raio.

Posso expressar  $y$  como função de  $x$  obtendo então duas equações de funções

$$f(x) = \sqrt{1 - \frac{x^2}{p}}; |x| \leq \sqrt{p}; \quad (361)$$

$$g(x) = -\sqrt{1 - \frac{x^2}{p}}; |x| \leq \sqrt{p}; \quad (362)$$

e usar `gnuplot` para obter o gráfico da elipse dando valores para  $B, p$  que pode ser vista na figura (fig 52), página 130,

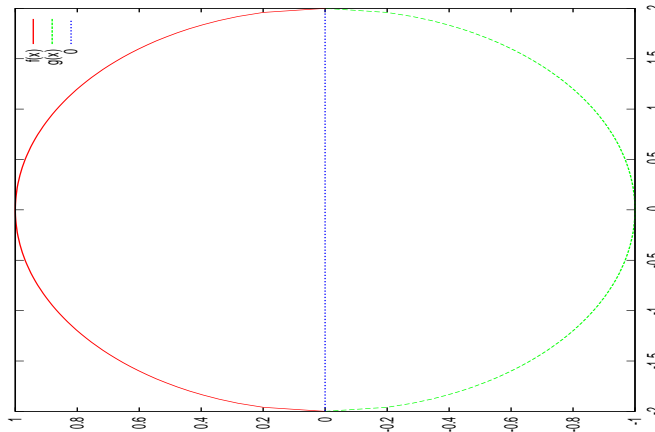


Figura 52: gráfico duma elipse com `gnuplot`

Se fosse possível associar botões controlando potenciômetros que alterassem o tamanho de  $b \in (0, B)$  poderíamos

- ver a elipse colapsar para um segmento de reta e coincidir com o diâmetro do círculo maior, quando  $b = 0$  e neste caso  $c = a$ ,
- ou, quando  $b = B$ , e neste caso  $c = 0$ , a elipse coincidiria com o círculo maior, e a equação (eq. 369) se transformaria na equação do círculo de raio  $B$

$$\frac{x^2}{B^2} + \frac{y^2}{B^2} = 1 \equiv x^2 + y^2 = B^2 \quad (363)$$

- excentricidade da elipse. Ainda observando a figura (fig 51), se  $c = 0$  então os dois círculos ficam coincidentes e a elipse se transforma num círculo. A excentricidade é definida pela razão  $\frac{c}{a}$ , usando a hipotenusa e um dos catetos do triângulo retângulo. É uma convenção a escolha do cateto que mede  $c$ , bem conveniente.  $c \in [0, a]$  e quando  $c = 0$ , os dois círculos na figura (fig 51) ficam coincidentes. O cateto que mede  $b$  não oferece esta

variação e certamente foi esta razão da escolha que se tem para a definição da excentricidade.

Um círculo é uma elipse com excentricidade nula. A excentricidade da rota da Terra em volta do Sol fica próximo de zero variando próximo do número 0.016710 ou seja é uma órbita quase circular.

Uma forma de perceber a excentricidade duma elipse se dá quando você desenhar um círculo numa folha de papel. Ao olhar para figura com o papel perpendicular ao eixo do seu *cone de visão* você verá um círculo, produzindo um ângulo diferente do ângulo reto, você terá a impressão de ver uma elipse. Esta experiência tem o que ver com a próxima.

Um cone, em geometria, é uma figura espacial que se obtém quando se fizer girar uma reta, *chamada geratriz do cone*, em torno de outra que lhe seja oblíqua, *chamada eixo do cone*. O resultado será um *cone de duas folhas* como aparece na figura (fig 55), página 136.

Uma terceira forma de obter uma elipse consiste em cortar um cone de forma que o plano de corte não seja perpendicular ao eixo do cone. Você pode ter uma visão disto se desenhar um círculo sobre uma folha de papel e olhar o desenho com o papel inclinado em relação ao *eixo do cone de sua visão*. Se a inclinação for de 90 graus, o papel estiver perpendicular ao eixo do cone de sua visão, você estará vendo um círculo, com qualquer outro ângulo, até zero, verá uma elipse. Esta definição ainda tem um outro grande interesse: o plano que corta o cone irá produzir todas as cônicas dependendo do ângulo do corte.

Não é fácil passar da forma geométrica para a forma algébrica na definição da elipse, as contas são muito complicadas.

Vou mostrar como se pode obter a equação duma elipse que esteja colocada numa posição arbitrária, mas com dados conhecidos, no plano, como *transformação* da equação obtida com os eixos da elipse sobre os eixos coordenados.

Primeiro deixe-me aplicar uma translação à elipse usando o vetor  $(r, s)$  transformo, assim, a equação

$$\frac{x^2}{B^2} + \frac{y^2}{B^2 - c^2} = 1; \quad (364)$$

na equação

$$\frac{(x - r)^2}{B^2} + \frac{(y - s)^2}{B^2 - c^2} = 1; \quad (365)$$

esta é a equação da elipse com polos  $F_1 = (-c, r)$  e  $F_2 = (c, s)$  obtidos com a soma do vetor de translação aos focos anteriores. Observe que a translação é  $(r, s)$ , mas aparece dentro da equação  $(x - r, y - s)$ .

---

- **equação do círculo**, é um item da *geometria analítica*. Usando uma *linguagem moderna*, podemos dizer que a *geometria analítica* digitaliza algumas figuras geométricas criando equações o que torna possível, por exemplo, usar estas figuras em programas de computação. A *geometria analítica* pode ser vista como uma introdução à *geometria algébrica* no sentido em que esta disciplina

da Matemática trata os entes geométricos que podem ser definidos através de equações algébricas, mais precisamente como expressões polinomiais.

Há um erro neste texto que será corrigido mais a frente, mas procure descobri-lo você mesma.

A definição geométrica de *círculo* é “o lugar geométrico dos pontos do plano que equidistam dum ponto chamado centro”. Deixe-me identificar o centro com as coordenadas  $C = (a, b)$  e um ponto genérico,  $P$ , do plano com as coordenadas  $(x, y)$ . A frase geométrica que define círculo pode então ser traduzida algébricamente como

$$d(P, C) = r; \sqrt{(x - a)^2 + (y - b)^2} = r \quad (366)$$

em que  $r$  é a distância constante entre  $P$  e  $C$ . A figura (fig 65), página 156, lhe

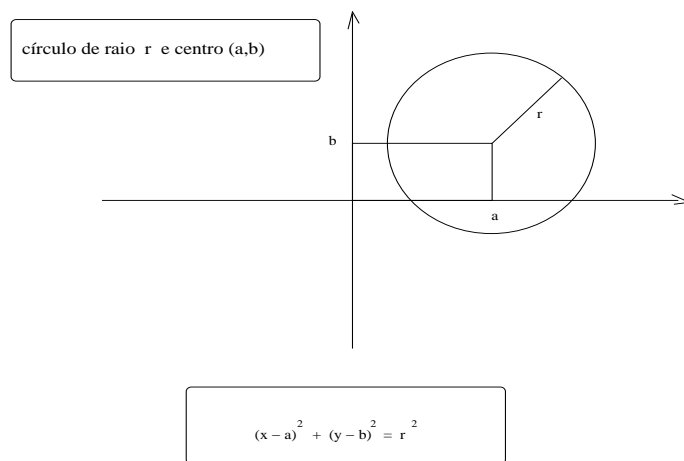


Figura 53: círculo de raio  $r$  e centro  $(a, b)$

apresenta a ideia geométrica contida na definição do círculo.

A expressão na equação (eq. 366) poderia ser considerada a equação do círculo, mas não é nesta forma que vamos encontrá-la, habitualmente, mas sim na forma

$$P(x, y) = (x - a)^2 + (y - b)^2 = r^2; \quad (367)$$

entretanto qualquer uma das duas equações é perfeita para representar o círculo apenas no primeiro caso podem surgir problemas se for usada num programa de computador pela presença da raiz que exige que o seu “parâmetro” seja positivo.

O objetivo da *geometria analítica* é duplo, por um lado, codificar em equações algumas figuras geométricas, as cônicas, e por outro lado, identificar numa expressão algébrica “geral” que tipo de figura algébrica se encontra codificada na expressão. Este segundo problema não é nada fácil e na verdade me-lhor faz parte do ambiente matemático mais amplo que é a *geometria algébrica*, mas é exatamente este segundo problema que aparece na maioria dos exercícios de

geometria analítica: “dado o polinômio

$$P(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (368)$$

em que  $A, B, C, D, E, F$  são constantes dadas, identifique qual é a cônica que a equação (eq. 368) representa”.

No caso do círculo posso encontrar as regras para fazer esta identificação se expandir a equação (eq. 367) quando vou encontrar a expressão

$$P(x, y) = x^2 + 2ax + a^2 + y^2 + 2by + b^2 - r^2 = 0 \quad (369)$$

e comparando as duas equações deduzo que  $A = C$  e se  $A \neq C$  a expressão pode representar uma elipse, uma hipérbole ou nenhuma cônica. Também

$$B = 0; D = 2a; E = 2b; F = a^2 + b^2 - r^2; \quad (370)$$

e desta forma, se a equação (eq. 368) representar um círculo é então possível encontrar o centro,  $C$ , e o raio,  $r$ . Analise um exemplo, considere a sequência de cálculos, que vou considerar “validados” porque  $A = C$ , então poderá ser um círculo.

$$P(x, y) = x^2 - 4x + y^2 + 6y - 3 = 0; \quad (371)$$

$$P(x, y) = x^2 - 4x + 4 - 4 + y^2 + 6y + 9 - 9 - 3 = 0; \quad (372)$$

$$P(x, y) = x^2 - 4x + 4 + y^2 + 6y + 9 - 9 - 3 - 4 = 0; \quad (373)$$

$$P(x, y) = x^2 - 4x + 4 + y^2 + 6y + 9 - 16 = 0; \quad (374)$$

$$P(x, y) = (x - 2)^2 + (y + 3)^2 - 4^2; \quad (375)$$

$$(x - 2)^2 + (y - (-3))^2 = 4^2; \quad (376)$$

que é a representação algébrica do círculo de centro  $C = (2, -3)$  e raio  $r = 4$ . Na sequência de contas das equações (eq. 371) -(eq. 376) usei uma técnica denominada completção de quadrados em vez de procurar identificar os coeficientes da equação geral (eq. 368). Quando identificamos  $A = C$ , é mais prático seguir por este método para verificar se a expressão representa um círculo. Para dominar as técnicas envolvidas no trabalho com *geometria analítica*, como em qualquer disciplina ou atividade humana, a regra é bem simples: fazer uns mil exercícios... e você os encontra aos milhares em textos livres distribuídos na Internet.

O erro mencionado inicialmente se encontra em que, ao escrever-se uma expressão  $P(x, y)$  não está implicitamente identificada a dimensão do espaço em que esta expressão “codifica” algum objeto. Por exemplo, no espaço de dimensão 3,  $x^2 + y^2 = 1$  representa o cilindro cuja base é o círculo unitário do plano  $XOY$  e a geratriz é o eixo  $OZ$ . Apenas mencionar esta equação sem identificar o universo em que ela estiver sendo considerada deixa uma indefinição. Este erro é muito comum nos textos, mas é difícil de ser corrigido sem criar uma redação muito complexa, em geral se admite implicitamente que se houver apenas duas variáveis, então a dimensão do universo é dois.

É interessante responder a uma pergunta que sempre é feita. Porque a  $P(x, y) = 0$  é que representa a equação do círculo e não apenas  $P(x, y)$ ? É que escrevendo “apenas  $P(x, y)$ ” você tem uma expressão livre que pode assumir qualquer valor em função dos valores dados às variáveis  $x, y$  e conseqüentemente  $P(x, y)$  se refere a qualquer ponto  $(x, y, z)$  do  $\mathbf{R}^3$  em que  $z = P(x, y)$ . A restrição  $P(x, y) = 0$  limita esta expressão a  $z = 0$ . Outra forma de dizê-lo, é que  $P(x, y) = 0$  corresponde ao sistema de equações

$$\begin{cases} z = P(x, y); \\ z = 0; \end{cases} \quad (377)$$

em que a primeira equação representa o gráfico da função bivariada  $z = P(x, y)$  e a segunda equação representa o gráfico da função bivariada  $z = 0$  cuja gráfico é o plano  $XOY$  ou seja, a equação (eq. 377) representa a interseção de duas superfícies o que pode ser uma curva. As curvas num espaço de dimensão maior, como aqui seria o  $\mathbf{R}^3$ , são determinadas por interseção de superfícies. Assim, o erro mencionado acima fica amenizado com um sistema de equações como a equação (eq. 377). Mas indefinição persiste... uma curva é uma variedade de dimensão 1 e a expressão  $P(x, y) = 0$  apenas sugere que existam duas variáveis o que daria para calcular a dimensão por uma regra prática que diz que a dimensão é uma unidade menor do que o número de variáveis. Mas escrever

$$x^2 - 4x + y^2 + 6y - 3 = 0$$

não contém nenhuma informação sobre a variável  $z$  ou significa que ela pode ter qualquer valor então esta expressão no  $\mathbf{R}^3$  é um cilindro uma vez que não está estabelecida nenhuma restrição sobre a mesma e existe restrições sobre os valores de  $x, y$ .

Entre escrever textos extremamente complicados termina valendo o bom senso que é determinado pelo contexto.

---

- **equação da hipérbole** é um item da *geometria analítica*. Usando uma *linguagem moderna*, podemos dizer que a *geometria analítica* “digitaliza” algumas figuras geométricas criando equações o que torna possível, por exemplo, usar estas figuras em programas de computação.

A *geometria analítica* pode ser vista como uma introdução à *geometria algébrica* no sentido em que esta disciplina da Matemática trata os entes geométricos que podem ser definidos através de equações algébricas, mais precisamente como expressões polinomiais. Porém a *geometria analítica* se especializa nas cônicas que são curvas, ou superfícies que podemos destacar dum *cone de duas folhas*.

Deixe-me introduzir a nomenclatura do cone que é uma superfície gerada no espaço quando uma reta, chamada *diretriz*, que é oblíqua a uma outra chamada *eixo do cone*, gira no espaço em torno do *eixo do cone*. O resultado é um cone de duas folhas que se conectam por um ponto.

Confira a figura (fig 54), página 135, em que se pode ver a interseção dum plano paralelo ao eixo, produzindo os ramos duma hipérbole.



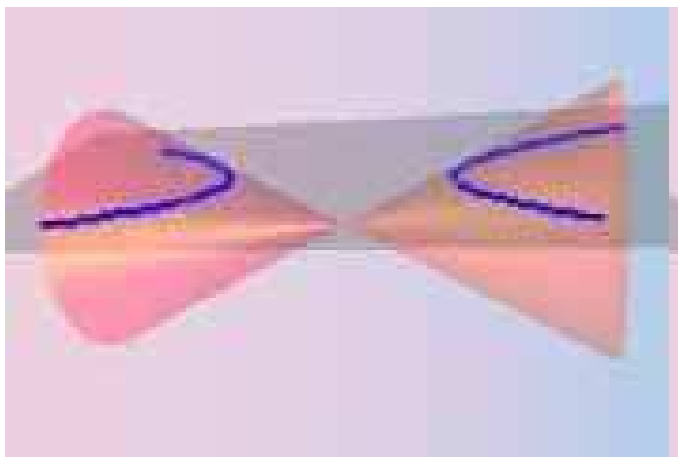


Figura 54: Plano paralelo ao eixo do cone: hipérbole

### Cenário geométrico das cônicas

Vou mostrar que as cônicas se alteram na passagem entre círculos e parábolas. Por exemplo, uma elipse é apenas um círculo deformado e uma pequena deformação dada a um elipse faz surgir uma parábola. Inclusive isto tem aplicação na Astronomia ou melhor na exploração espacial em que uma nave espacial pode ser deixada hibernando numa série de elipses e repentinamente sair da elipse para uma parábola mudando a trajetória como foi o caso do projeto com a nave espacial Rosetta da agência espacial europeia. Uma pequena alteração num plano que produz uma parábola e o resultado é uma hipérbole. Você vai ver, também, que existem *hipérboles degeneradas*.

Vou começar discutindo a origem dos círculos e das parábolas.

Para estabelecer um *cenário* deixe-me fixar alguns “quadros” da geometria envolvendo um cone de duas folhas. Há duas classes de planos que delimitam os acontecimentos produzindo círculos e parábolas e depois você verá que elipses e hipérboles são variações destas duas curvas, na verdade *eventos estáveis* cujos limites são os *eventos instáveis círculos e das parábolas*. . . mas você não precisa dar muita importância a esta frase se ainda não souber o que são *sistemas dinâmicos*.

Vou também definir um *plano codificador* de onde vou deduzir um *círculo codificador* obtendo uma codificação mais efetiva e mais simples porque vai classificar melhor os eventos e descrever todas as possibilidades de cônicas.

- círculo qualquer plano perpendicular ao eixo do cone, eles cortam o cone segundo um círculo;
- parábola um plano tangente à superfície do cone, que contém uma imagem da *diretriz*, e qualquer plano paralelo a ele. Um tal plano corta a superfície do cone segundo uma parábola se não contiver a *diretriz*. O caso do plano

tangente à superfície do cone que são aqueles que contêm a *diretriz*, é uma *situação limite* que pertence ao “caso da hipérbole”. Eu vou discutir este caso sob o título *hipérbolas degeneradas*.

Estes dois pequenos exemplos mostram que preciso relacionar ângulos entre planos. Vou agora precisar de fazer referência aos ângulos que um plano faria com os planos das duas classes descritas com as etiquetas *círculo* e *parábola*. Determinamos o ângulo entre dois planos usando duas retas perpendiculares, cada uma delas, a um dos planos. É o ângulo entre estas retas que define o ângulo entre os planos.

Mas tem uma outra forma de codificar o cenário com uma linguagem que dominamos melhor, projetando num plano.

Existe *exatamente* um plano, deixe-me chamá-lo *plano codificador*, que é determinado pela geratriz e pelo eixo do cone. Posso codificar todo o cenário usando *este plano* como vou passar a fazer agora. O *plano codificador* é o que aparece na figura (fig 55), página 136, com a etiqueta 0 para a classe do *círculo* e a etiqueta 1 para a classe da *parábola*, contendo as imagens da geratriz e do eixo do cone. A geratriz é única, mas ela roda no espaço produzindo duas imagens no *plano codificador*.

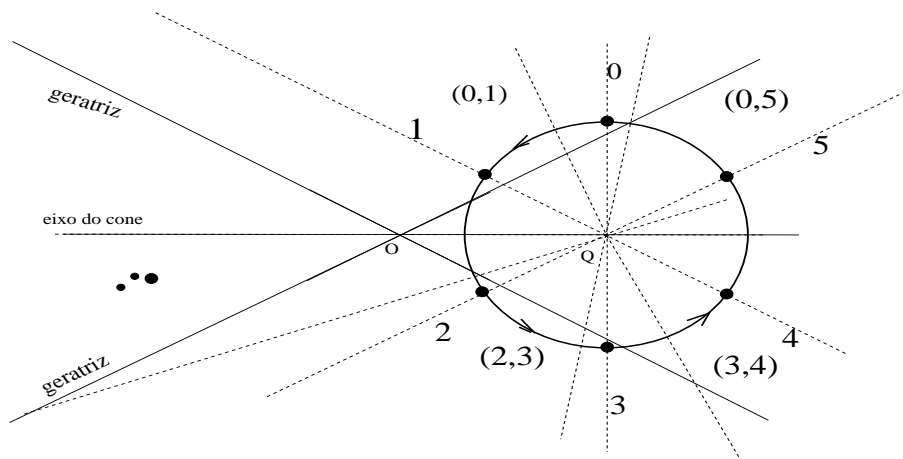


Figura 55: o círculo codificador

Deixe-me descrever a figura (fig 55), página 136, e partir para uma especialização da codificação que vai agora ficar restrita apenas ao círculo que aparece na figura (fig 55), o *círculo codificador*. Este círculo tem o centro em um ponto  $Q$  do eixo do cone e raio qualquer. Apenas para obter uma figura menos poluída é bom selecionar o raio menor do que  $d(Q, O)$  em que  $O$  é ponto de concorrência da diretriz e o eixo do cone.

Nela você pode identificar seis pontos que foram etiquetados com os números  $\{0, 1, 3, 4, 5\}$  e que vou agora associar às cônicas que se podem obter.

- círculo a classe dos planos perpendiculares ao eixo marcada com os números 0, 3 e reduzida a uma reta perpendicular ao eixo do cone passando pelo centro do *círculo codificador* que aparece na figura (fig 55).

Qualquer reta perpendicular ao eixo do cone, na figura (fig 55) representa um plano desta classe que está agora reduzida as etiquetas 0, 3 do *círculo codificador*.

- parábola a classe dos planos cuja interseção produzem parábolas, marcada com os números 1, 2, 4, 5 e reduzida às duas retas paralelas à alguma das representações da geratriz, passando pelo centro do *círculo codificador* que aparece na figura (fig 55). A geratriz é única, mas ela roda no espaço e assim aparece no plano da figura (fig 55) duas vezes.

As imagens das diretrizes não pertencem a esta classe porque elas vão gerar *hipérboles degeneradas* de que vou tratar mais a frente. Observe que círculos e parábolas são determinados por pontos isolados do *círculo codificador* em comparação com as elipses e hipérboles que vão ficar determinadas por intervalos.

- o universo dos planos que nos interessam são os planos perpendiculares ao plano da (fig 55). Outra forma de fixar os planos interessantes é considerando a reta perpendicular ao *plano codificador*, (fig 55), passando pelo centro do *círculo codificador*: são os planos do espaço que contenham esta reta ou planos paralelos a estes.

O conjunto dos planos que contem um reta se chama *um feixe de planos*, portanto nos interessam todos os planos paralelos a um dos planos do feixe determinado pela reta perpendicular ao *círculo codificador* passando pelo seu centro.

Entenda os números 0, 1, 2, 3, 4, 5 como etiquetas no *círculo codificador* que se vê na figura (fig 55). E agora estou partindo para forma ainda mais especializada (e simples) de codificar usando apenas pontos ou setores do *círculo codificador* que aparece na (fig 55).

Estes números, 0, 1, 2, 3, 4, 5 não seriam as melhores escolhas para etiquetar, talvez melhor fosse usar o ângulo da diretriz com o eixo do cone, entretanto isto dependeria da escolha da diretriz e implicaria numa maior complexidade da linguagem. A escolha que fiz usa a figura (fig 55) e portanto mostra, por semelhança de triângulos a dependência do ângulo que a diretriz faz com o eixo do cone. Enfim, fazer uma *codificação* implica em tomar uma decisão para escolher o código.

Todos os planos do universo que nos interessa, agora ficam classificados assim, usando apenas pontos ou segmentos de arco do *círculo codificador*:

- Os planos que produzem cônicas são os planos do feixe de planos determinado pela reta perpendicular ao *círculo codificador* passando pelo centro mesmo, ou planos paralelos a algum dos planos deste feixe.

- círculo os planos que determinam a reta etiquetada com 0,3 ou uma sua paralela.
- parábola os planos que determinam as retas etiquetadas com 1, 2, 4, 5, ou uma paralela a ela, excetuando as imagens da geratriz, confira o item hipérbole.
- elipse os planos que determinarem com o centro do *círculo codificador* uma reta nos intervalos (0, 1), (2, 3), (3, 4), (0, 5) Estas etiquetas significam mesmo um intervalo aberto porque os extremos do intervalo ficam para círculo ou parábola. Como observei acima, círculos e parábolas ficam determinadas por pontos do *círculo codificador*. Estas figuras são muito instáveis, por exemplo um objeto astronômico raramente se encontra sobre uma trajetória do tipo círculo ou parábola... se isto ocorrer será por uma instabilidade que deve ser corrigida.
- hipérbole os planos do feixe que pertencerem ao complementar das opções anteriores, ou ainda, quando um plano determinar uma reta que passando pelo centro do círculo que não esteja nas classes anteriores. Um plano da classe *parábola*, representa uma *situação limite* exatamente representado pelos planos tangentes ao cone, eles determina *hipérboles degeneradas* que se reduzem a duas semiretas representadas por qualquer das diretrizes. Eu vou voltar a esta questão na construção da equação da hipérbole.
- retas A interseção dum plano com o cone pode ser uma reta, se ele for tangente à superfície do cone. Tecnicamente, tais retas seriam hipérboles degeneradas. Observe que as hipérboles têm dois ramos desconexos, neste caso seriam duas semiretas conetadas na origem... os dois ramos duma hipérbole degenerada.

É um bom exercício corrigir as etiquetas deste modelo transformando-as em ângulos no forma  $\pi$ -radianos usando o ângulo  $\alpha$  entre uma das imagens da diretriz e o eixo do cone. O autor agradece se você lhe enviar esta a proposta e promete analisá-la para alteração do texto como uma contribuição sua.

Embora estas curvas sejam espaciais, uma vez são o resultado da interseção de duas superfícies, um cone e um plano, eu vou aqui estudar a hiperbole que é a projeção desta interseção no plano  $XOY$ , portanto vou me restringir ao estudo duma curva plana.

Confira a equação da parábola que é obtida quando o plano que intercepta o cone se encontra exatamente na direção de uma das imagens da *diretriz* no plano  $XOY$ .

A figura (fig 56), página 139, lhe mostra o resultado do corte que o plano produziu no cone quando estiver num ângulo, relativamente ao eixo do cone, que fique entre os ângulos determinados pelas imagens da *diretriz* no plano  $XOY$ , ou ainda nos intervalos [1, 2], [4, 5] do *círculo codificador*, na figura (fig 55).

As imagens da *diretriz* são duas retas concorrentes num ponto do plano, e aqui vou introduzir uma particularização que não irá prejudicar o desenvolvimento geral da teoria, mas irá simplificar as contas. Vou escolher a origem dos

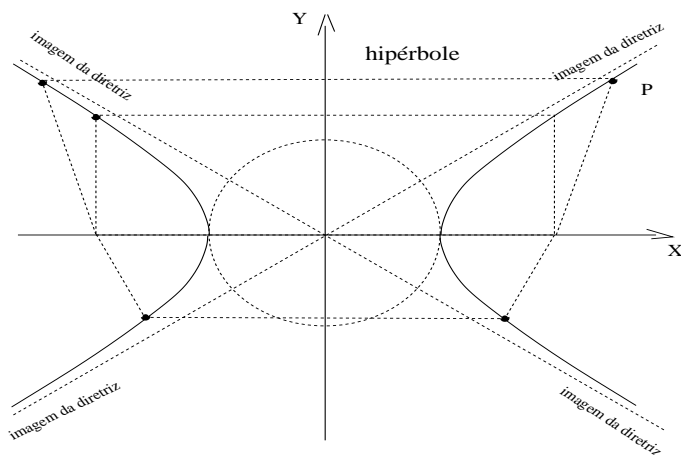


Figura 56: hipérbole com suas assíntotas

eixos no plano  $XOY$  como o ponto de concorrência das imagens da diretriz e do eixo do cone. As imagens da diretriz são as assíntotas ao gráfico da hipérbole: quer dizer, o gráfico da hipérbole se aproxima indefinidamente das assíntotas, ou ainda as assíntotas são uma *posição limite* para a qual tendem os ramos da hipérbole. Observe que as duas imagens da diretriz são também as *hipérboles degeneradas*.

As hipérboles têm uma definição da geometria, *é o lugar geométrico dos pontos  $P$  tal que o valor absoluto da diferença entre as distâncias do ponto  $P$  a dois pontos chamados focos, é constante*. Compare com a definição da elipse.

Posso formalizar esta definição para obter a equação da hipérbole. Vou usar uma propriedade da geometria que é a invariância das propriedades das figuras geométricas sob os movimentos rígidos: rotação, translação. Na verdade eu já usei esta propriedade de invariância quando selecionei a origem como ponto de concorrência das imagens da diretriz.

Isto me permite escolher a posição dos referenciais para construir a equação da hipérbole de forma mais simples. Então vou selecionar os dois pontos, os focos, como:  $F_1 = (-c, 0)$ ;  $F_2 = (c, 0)$ . Estou considerando os focos sobre o eixo  $OX$ .

#### Análise geométrica

Antes de calcular a equação, vou explorar os aspectos geométricos e com eles construir alguns pontos da hipérbole a guisa de motivação para os cálculos.

Observe, e ao final vou tirar uma conclusão a respeito, eu poderia ter selecionado os focos sobre o eixo  $OY$ , o resultado seria uma hipérbole com rotação de  $\frac{\pi}{2}$ .

É fácil deduzir que se um ponto  $P$  do plano satisfizer à definição, então um ponto simétrico a ele, em relação ao eixo  $OY$  também satisfaria à definição, porque, posso construir um triângulo semelhante com a mesma base  $\underline{F_1F_2}$ , usando

um ponto simétrico à  $P$  relativamente ao eixo  $OY$ .

Isto significa que as duas componentes da hipérbole são duas curvas simétricas em relação ao eixo  $OY$ . Também a mesma semelhança de triângulos nos permite concluir que cada ponto da hipérbole tem um simétrico em relação ao eixo  $OX$  e assim o gráfico da hipérbole é simétrico relativamente aos dois eixos  $OY$  e  $OX$ .

Se um ponto estiver sobre a hipérbole, e fora do eixo  $OY$ , a diferença  $d(P, F_1) - d(P, F_2)$  é necessariamente diferente de zero o que garante que nenhum ponto da hipérbole pode se encontrar sobre o eixo  $OY$ , mas vou mostrar que é possível que  $P$  esteja sobre o eixo  $OX$  então o eixo  $OY$  divide o plano em dois semiplanos cada um contendo um ramo da hipérbole.

Há dois pontos sobre o eixo  $OX$ , sejam eles  $P_1, P_2$ , simétricos relativamente ao eixo  $OY$ , que pertencem, cada um deles, a um dos ramos da hipérbole. São os dois pontos de menor distância entre os dois ramos da hipérbole, eles fazem parte a construção da hipérbole no sentido de que a cada seleção de  $P_1, P_2$ , corresponde uma hipérbole.

Então a hipérbole fica definida pela seleção de

$$P_1, P_2, F_1, F_2; \quad (378)$$

Usando estes dois pontos posso calcular a diferença constante

$$d(P, F_1) - d(P, F_2);$$

confira a figura (fig 57), página 141,

$$|d(P_1, F_1) - d(P_1, F_2)| = d(P_1, P_2) = r \quad (379)$$

porque é uma subtração de segmentos de reta sobre a mesma reta suporte  $OX$ .

Vou chamar estes dois pontos em que os ramos da hipérbole cortam o eixo  $OX$  de vértices, de cada um dos ramos, eles são, pela simetria, os dois pontos mais próximos dos dois ramos da hipérbole.

A figura (fig 57), página 141, lhe mostra um resumo da análise geometrica: a hipérbole é uma curva plana que se compõe de dois ramos, cada um deles se encontra num dos semiplanos em que  $OY$  divide o plano, é uma curva simétrica em relação aos dois eixos coordenados e os dois pontos mais próximos dos dois ramos são os pontos  $P_1, P_2$  que se encontram sobre o eixo  $OX$  e a hipérbole fica inteiramente determinada pela escolha dos pontos  $\{P_1, P_2, F_1, F_2\}$  de uma forma que ficará bem clara ao final das contas seguintes.

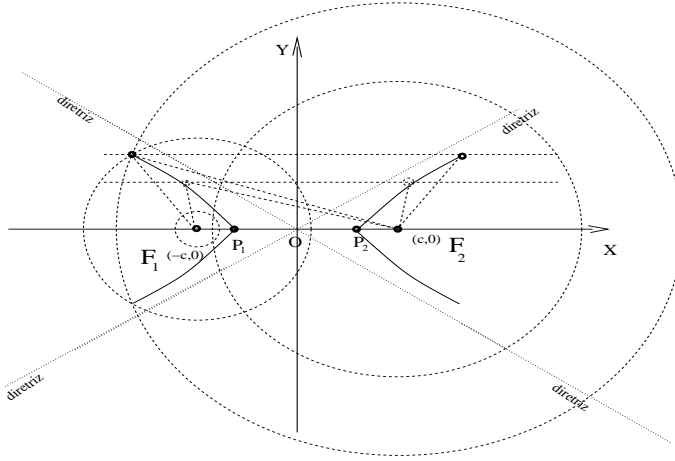


Figura 57: definição da hipérbole

$$d(P, F_1) - d(P, F_2) = r = d(P_1, P_2) < 2c = d(F_1, F_2); \quad (380)$$

$$0 < r < 2c; 4c^2 - r^2 = \delta^2 > 0; c^2 - \frac{r^2}{4} = \frac{\delta^2}{4} \quad (381)$$

$$d(P, F_1) - d(P, F_2) = \sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2}; \quad (382)$$

$$\sqrt{(x+c)^2 + y^2} - \sqrt{(x-c)^2 + y^2} = r; \quad (383)$$

$$\sqrt{(x+c)^2 + y^2} = \sqrt{(x-c)^2 + y^2} + r; \quad (384)$$

$$(x+c)^2 + y^2 = (\sqrt{(x-c)^2 + y^2} + r)^2; \quad (385)$$

$$4cx - r^2 = 2r\sqrt{(x-c)^2 + y^2}; \quad (386)$$

$$16c^2x^2 - 8cxr^2 + r^4 = 4r^2((x-c)^2 + y^2); \quad (387)$$

$$16c^2x^2 - 8cxr^2 + r^4 = 4r^2(x^2 - 2cx + c^2 + y^2); \quad (388)$$

$$16c^2x^2 - 4r^2x^2 = -r^4 + 4r^2c^2 + 4r^2y^2 = 0; \quad (389)$$

$$4(4c^2 - r^2)x^2 = r^2(4c^2 - r^2 + 4y^2); \quad (390)$$

$$4\delta^2x^2 = r^2(\delta^2 + 4y^2); 4\delta^2x^2 = r^2\delta^2 + 4r^2y^2 \quad (391)$$

$$\text{fatorando } r^2\delta^2; \frac{4}{r^2}x^2 - \frac{4}{\delta^2}y^2 = 1; \quad (392)$$

$$\frac{r^2}{4} = a^2; \frac{\delta^2}{4} = b^2; \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1; \quad (393)$$

O número positivo  $c$  é a metade da distância entre os focos,  $r$  é a distância

entre os vértices,

$$r < c; \quad (394)$$

$$\frac{\delta}{2} = \sqrt{c^2 - \frac{r^2}{4}}; \quad (395)$$

$$a = \frac{r}{2}; b = \frac{\delta}{2} = \sqrt{c^2 - \frac{r^2}{4}}; \quad (396)$$

$$a = \frac{r}{2}; b = \sqrt{c^2 - a^2}; \quad \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1; \quad (397)$$

Para construir a hipérbole, como já disse anteriormente, basta-nos selecionar  $\{P_1, P_2, F_1, F_2\}$  com os quais podemos calcular a distância entre os focos,  $2c$  e a distância  $r < 2c$  entre os ramos da hipérbole. Posso expressar qualquer uma das variáveis,  $x, y$  em função da outra para concluir que

$$x = \pm a \sqrt{1 + \frac{y^2}{b^2}} \quad (398)$$

onde podemos ver que para grandes valores de  $|y|$ , como  $a, b > 0$

$$|x| \approx \frac{a}{b}|y|; |x| < \frac{a}{b}|y|; |y| > \frac{b}{a}|x|; \quad (399)$$

que mostra que as retas  $y = \pm \frac{b}{a}x$  se encontram arbitrariamente próximas do gráfico da hipérbole, e que os ramos da hipérbole ficam situados acima ou abaixo dos gráficos destas retas. Estas retas se chamam de assíntotas e são as imagens das diretrizes da hipérbole e para obter um rápido borrão do gráfico da mesma basta traçar estas retas limitando os dois ramos da hipérbole como assíntotas. A figura (fig 58), página 142, mostra uma hipérbole desenhada usando `xfig`

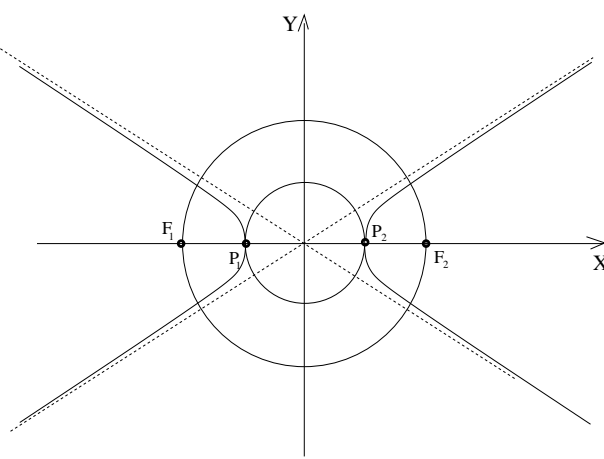


Figura 58: hipérbole desenhada a mão com `xfig`

a partir das quatro informações  $\{P_1, P_2, F_1, F_2\}$  selecionadas sobre o eixo  $OX$ ,



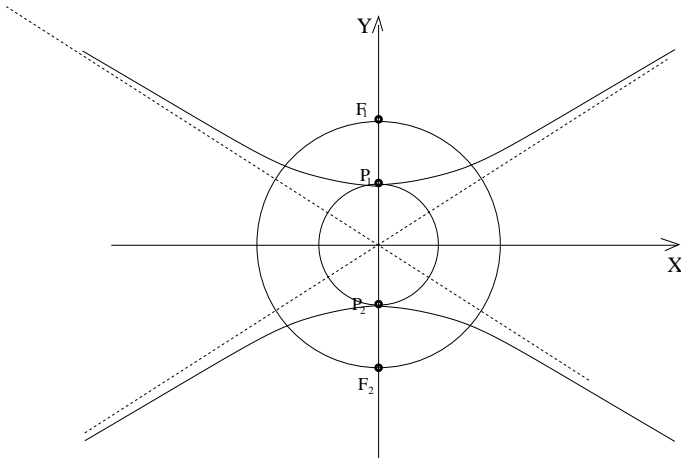


Figura 59: hipérbole com os polos sobre o eixo  $OY$

enquanto a figura (fig 59), página 143, mostra uma rotação de  $\pi$  dada à figura (fig 58) transformando os quatro pontos,  $\{P_1, P_2, F_1, F_2\}$ , para localizá-los sobre o eixo  $OY$ .

As equações (eq. 394)-(eq.399) podem ser reaproveitadas, apenas trocando  $x$  por  $y$  para obter:

$$r < c; \quad (400)$$

$$\frac{\delta}{2} = \sqrt{c^2 - \frac{r^2}{4}}; \quad (401)$$

$$a = \frac{r}{2}; b = \frac{\delta}{2} = \sqrt{c^2 - \frac{r^2}{4}}; \quad (402)$$

$$a = \frac{r}{2}; b = \sqrt{c^2 - a^2}; \quad \frac{y^2}{a^2} - \frac{x^2}{b^2} = 1; \quad (403)$$

$$y = \pm a \sqrt{1 + \frac{x^2}{b^2}} \quad (404)$$

Como  $a, b$  são obtidos em função dos pontos  $\{P_1, P_2, F_1, F_2\}$  posso agora definir duas funções

```
pow(x,n) = x**n;
r = 1.0; c = 2.0; a = r/2; delta = 2*sqrt(pow(c,2) - pow(r,2)/4);
b = sqrt(pow(c,2) - pow(a,2));
f(x) = a*sqrt(1 + pow(x,2)/pow(b,2));
g(x) = -a*sqrt(1 + pow(x,2)/pow(b,2));
r1(x) = (a/b)*x; r2(x) = -(a/b)*x;
plot f(x),g(x),0,r1(x),r2(x);
```

para obter o gráfico na figura (fig 60), página 144, com `gnuplot`.

- equação da reta, é um tópico da Geometria Analítica. Usando uma *lin-*

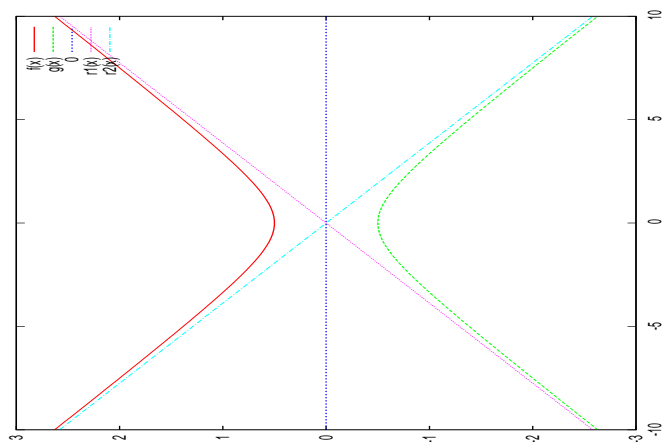


Figura 60: gráfico da hipérbole feito com gnuplot

guagem moderna, podemos dizer que *Geometria Analítica* digitaliza algumas figuras geométricas criando equações o que torna possível, por exemplo, usar estas figuras em programas de computação. A *Geometria Analítica* pode ser vista como uma introdução à *geometria algébrica* no sentido em que esta disciplina da Matemática trata os entes geométricos que podem ser definidos através de equações algébricas, mais precisamente, como expressões polinomiais.

A reta é uma figura básica da *geometria euclidiana* e não teria sentido dar-lhe uma definição. Vou estabelecer algumas relações que vão abrir caminho para construir a equação da reta repetindo alguns postulados da *geometria euclidiana*. Ao mesmo tempo é preciso que você se alerte para o fato de que é a *geometria euclidiana* que se encontra nos fundamentos da *Geometria Analítica* e que existem outras geometrias *não euclidianas*.

A *geometria euclidiana* estabelece como determinar retas:

**Teorema 3 (axioma) dos dois pontos**

Dados dois pontos,  $P, Q$ , eles determinam uma reta  $r$  e podemos dizer que  $P, Q \in r$  entendendo que retas são conjuntos de pontos.

Confira a figura (fig 62), página 148, e considere os pontos  $P, Q$ . A *geometria euclidiana* também afirma que as retas são conjuntos infinitos de pontos:

**Teorema 4 (axioma) infinidade de pontos na reta**

Dados dois pontos diferentes,  $P, Q$ , na reta  $r$ , existe um terceiro ponto entre os pontos  $P, Q$  e um quarto ponto  $X$  fora do segmento determinado pelos pontos  $P, Q$ .

Com esta afirmação a *geometria euclidiana* garante que numa reta existe uma infinidade de pontos *entre* os pontos  $P, Q$  assim como também fora do segmento de reta que eles determinam . . .

Finalmente a *geometria euclidiana*<sup>17</sup> garante existência de paralelas:

**Teorema 5 (axioma)** da paralela da geometria euclidiana

Dado um ponto  $Q$  e uma reta  $r$ , existe uma única reta paralela a esta reta passando pelo ponto  $Q$ ,

Na figura (fig 62), página 148 selecionei os pontos  $P, Q, X \in r$  e baseado no axioma 5, eu posso traçar retas paralelas ao eixo  $OX$  passando por  $X$  e por  $Q$  assim como posso traçar retas paralelas ao eixo  $OY$  passando por  $P$  e por  $Q$  construindo assim dois triângulos retângulos semelhantes, confira a figura (fig 62). Esta construção me permite escrever a proporção:

$$\frac{x-a}{y-b} = \frac{p-a}{q-b} \quad (405)$$

se a reta  $r$  não for paralela com o eixo  $OX$ , porque neste caso

$$y-b = q-b = 0$$

e logo vou discutir este caso.

Aplicando a lei das proporções *produto dos meios é igual ao produto dos extremos* posso escrever, sucessivamente:

$$(x-a)(q-b) = (y-b)(p-a); \quad (406)$$

$$(q-b)x - a(q-b) = (p-a)y - (p-a)b; \quad (407)$$

$$A = (q-b); B = -(p-a); C = (p-a)b - a(q-b); \quad (408)$$

$$Ax + By + C = 0; \quad (409)$$

Considerando agora o caso da reta paralela com o eixo  $OX$ , ela vai cortar o eixo  $OY$  num único ponto em que  $y = y_0$  numa aplicação reversa do axioma 5. E qualquer ponto sobre esta reta tem  $y_0$  por ordenada então a equação

$$y = y_0 \quad (410)$$

identifica esta reta no sentido que *permite encontrar qualquer ponto sobre a mesma* e ela tem o formato da equação (eq. 406) se fizermos  $A = 0; B = 1; C = -y_0$ .

A afirmação “ $Ax + By + C = 0$  é a equação da reta” é, infelizmente, falsa, poderia ser a equação dum plano ou dum hiperplano. Para evitar esta ambiguidade é preciso acrescentar “no plano”,

“ $Ax + By + C = 0$  é a equação da reta no plano” é a afirmação correta porque num espaço de dimensão três esta equação representa um plano.

A Geometria Analítica trata basicamente das cônicas, no plano, são *reta, círculo, elipse, parábola e hipérbole* e no espaço tridimensional, *plano, esfera, elipsóide, parabolóide e hiperbolóide*.

Vou restringir-me ao plano, aqui. Então uma equação como  $y = x$  que também pode ser escrita como  $y - x = 0$  representa uma reta. Se você rodar

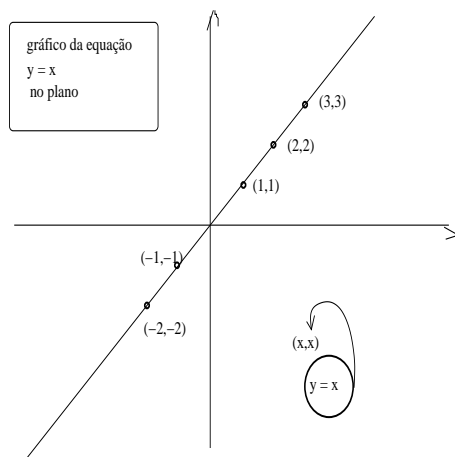


Figura 61: gráfico da equação  $y = x$  no plano

algum programa que produza gráficos, com esta equação você vai visualizar o que a figura (fig 61), página 146, lhe mostra. Esta figura foi desenhada com `xfig` que é um editor gráfico. Você pode obter um gráfico semelhante a este usando `gnuplot`. Num terminal do `gnuplot` execute os comandos:

```
f(x) = x
plot f(x), 0
pause -2 "Aperte para terminar"
```

para ver um gráfico semelhante ao da figura (fig 61).

Explicando, `gnuplot` “*não sabe*” fazer gráficos de equações, mas sabe gráficos de funções. Então defini  $f(x) = x$  ou ainda

$$y = f(x) = x \quad (411)$$

e depois o comando

```
plot f(x), 0
```

pede que `gnuplot` faça os gráficos das duas funções

$$y = f(x) = x; y = 0; \quad (412)$$

a segunda função é o eixo OX. Confira *eixos coordenados*.

Há várias formas de representar algébricamente uma reta, *equação vetorial*, *equação paramétrica* da reta. Confira *equações paramétricas*.

A equação vetorial, ou paramétrica da reta, é mais precisa. Uma reta é uma curva, confira *curva*, que é a imagem de uma função que depende de um parâmetro, o que caracteriza sua dimensão como sendo 1. Genericamente uma curva  $\alpha$  seria um objeto definido por

$$\alpha(t) = (x_1(t), \dots, x_n(t)); \quad n \text{ é a dimensão do espaço}; \quad (413)$$

<sup>17</sup>Esta é a “verdade” que separa a geometria euclidiana das geometrias não euclidianas...

Este é a equação duma curva definida no  $\mathbf{R}^n$ , é uma *variedade* de dimensão 1 imersa num espaço de dimensão  $n$ .

O que caracteriza uma reta é que todas as equações  $x_k(t)$  sejam do primeiro grau. Assim

$$\alpha(t) = (t, \dots, t) = (x_1(t), \dots, x_n(t)) \quad (414)$$

é a equação da primeira bissetriz no  $\mathbf{R}^n$ , assim como

$$\alpha(t) = (t, t) \quad (415)$$

é a equação da primeira bissetriz do plano. As equações que aparecem nas equações (eq. 414), (eq. 415) passam na origem. A equação paramétrica da reta na direção do vetor  $\vec{u}$  passando pelo ponto  $P$  é

$$P + t\vec{u}; \quad (416)$$

Escolha  $P$  e  $\vec{u}$  para obter a equação paramétrica da reta na dimensão que você desejar. Por exemplo, se você escolher

$$\vec{u} = (1, 2, 1) \in \mathbf{R}^3; P = (0, 0, 0)$$

o resultado da equação, (eq. 416), será o gráfico duma reta paralela ao vetor  $\vec{u} = (1, 2, 1)$  passando pelo ponto  $P = (0, 0, 0)$ , portanto uma reta paralela ao vetor  $\vec{u} = (1, 2, 1)$  passando na origem no espaço tridimensional  $\mathbf{R}^3$ .

Se você escolher

$$\vec{u} = (1, 2, 0, 1) \in \mathbf{R}^4; P = (-1, 2, 3, -2)$$

o resultado será uma reta do espaço de dimensão 4 passando pelo ponto  $P = (-1, 2, 3, -2)$ .

Esta equação é também chamada de *equação vetorial da reta*.

A equação cartesiana da reta, no plano, é uma expressão do primeiro grau envolvendo duas variáveis:

$$Ax + By + C = 0 \quad (417)$$

se aceitarmos que é a equação duma figura plana, o resultado é uma reta. Aqui você pode observar a indefinição das equações cartesianas, é preciso indicar em que dimensão estamos para saber o que uma equação cartesiana representa.

Podemos isolar  $y$ , escrevendo

$$y = -\frac{A}{B}x - \frac{C}{B}; \text{ quando } B \neq 0; \quad (418)$$

Se calcularmos, fazendo  $x = 0$ , ainda supondo que  $B \neq 0$ , vamos obter o ponto  $(0, -\frac{C}{B})$  na equação (eq. 418) e assim sabemos que esta reta passa neste ponto.

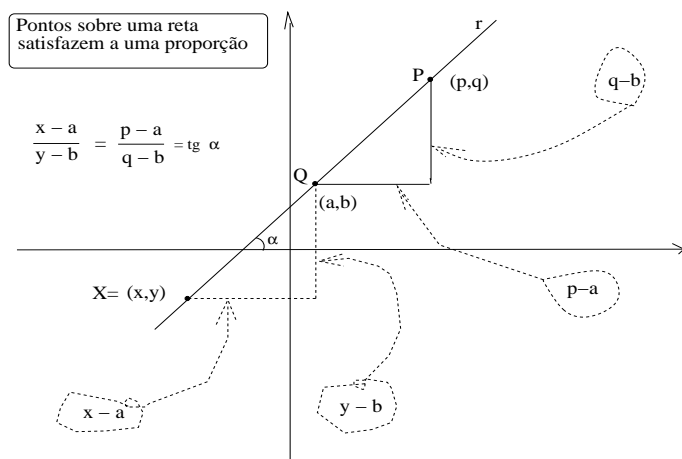


Figura 62: Coeficiente angular e coeficiente linear

O coeficiente de  $x$  na equação (eq. 418) é o *coeficiente angular da reta*,  $-\frac{A}{B}$ , é a tangente trigonométrica<sup>18</sup> do ângulo  $\alpha$  que a reta faz com o eixo  $OX$ . Confira o significado geométrico na figura (fig 62), página 148,

Os pontos sobre a reta satisfazem a uma proporção porque com eles montamos razões iguais entre os lados homólogos dos triângulos semelhantes que eles determinam com uma reta que lhe seja oblíqua. Na figura (fig 62) a reta  $r$  é oblíqua ao eixo  $OX$  e determina com ele o ângulo  $\alpha$  cuja tangente é a razão entre os catetos *oposto* e *adjacente* de qualquer um dos triângulos que você pode ver nesta figura.

Na figura (fig 62) marquei os pontos  $(a, b)$ ,  $(p, q)$  que determinam a reta e mais um ponto genérico,  $(x, y)$  que é a variável que vai permitir-me montar uma equação. Os lados homólogos são proporcionais:

$$\frac{y-b}{x-a} = \frac{q-b}{p-a} = \operatorname{tg}(\alpha) = m; \quad (419)$$

é a proporção obtida fazendo uso dos catetos do triângulo de vértices  $(a, b)$ ,  $(p, q)$  juntamente com triângulo formado pelos pontos  $(x, y)$  e  $(a, b)$ . Desta proporção

<sup>18</sup>Se  $B = 0$  então  $x = -\frac{C}{A}$  porque não tem sentido que também  $A = 0$ . Se  $A = 0$  não tem sentido que  $B = 0$ .

saem alguma das expressões que são bem conhecidas como “*equação da reta*”:

$$(q - b)(x - a) = (p - a)(y - b); \quad (420)$$

$$(q - b)x - (q - b)a = (p - a)y - (p - a)b; \quad (421)$$

$$p \neq a \Rightarrow y - b = \frac{q-b}{p-a}(x - a); \quad (422)$$

$$m = \frac{q-b}{p-a} \Rightarrow y = m(x - a) + b; \quad (423)$$

$$(q - b)x - (p - a)y - (q - b)a + (p - a)b; \quad (424)$$

$$\text{faça: } A = (q - b); B = -(p - a); C = -(q - b)a + (p - a)b; \quad (425)$$

$$Ax + By + C = 0; \quad (426)$$

$$B \neq 0; y = -\frac{A}{B}x - \frac{C}{B}; m = -\frac{A}{B}; n = -\frac{C}{B}; \quad (427)$$

$$y = mx + n; \quad (428)$$

- Na equação (eq. 423) obtive a equação na forma  $y = m(x - a) + b$  que evidencia que esta reta passa no ponto  $(a, b)$  e tem coeficiente angular  $m$ . Esta é uma das melhores formas de apresentar a equação da reta em situações em que você precise dum ponto onde a reta passe com uma inclinação adequada. Seria por exemplo o caso do módulo espacial Rosetta da Agência Espacial Européia que no momento certo despachou um robot para aterrisar na superfície do cometa *67P/Churyumov-Gerasimenko*. O módulo contendo o robot deve ter sido lançado num trajetória tangencial à trajetória da nave Rosetta para otimizar energia usando a direção em que trafegava a nave mãe que seguia numa rota sob atração gravitacional do cometa. Este é o formato requerido por muitas das aplicações do Cálculo Diferencial e Integral.

Por exemplo, a equação da *reta tangente ao gráfico da função diferenciável*  $y = f(x)$  no ponto  $(a, f(a))$  é

$$y = f'(a)(x - a) + f(a); m = f'(a); b = f(a); \quad (429)$$

- A equação (eq. 426) é comumente chamada de *equação geral da reta* porque apresenta a reta como consequência da combinação linear de dois números dados  $A, B$ . Observe o desprezo dado ao  $C$  que é apenas um coeficiente posicional: se  $C = 0$  a reta passa na origem ...
- Na equação (eq. 428) ficaram expressos dois coeficientes, o *coeficiente angular*,  $m$  e o *coeficiente linear*  $n$ . O *coeficiente angular* expressa a inclinação relativa da reta com o eixo  $OX$  e o *coeficiente linear* mostra em que ponto do eixo  $OY$  a reta passa.

O *coeficiente angular* é um conceito preso ao sistema de coordenadas cartesianas e de certo forma restritivo. Por exemplo, as retas paralelas ao eixo  $OY$  não tem *coeficiente angular*. Não teria sentido, também, falar no *coeficiente angular* com que a nave Rosetta expeliu o módulo em direção ao cometa *67P/Churyumov-Gerasimenko*.

Na verdade o caso da nave Rosetta e dos eventos espaciais caem no estudo vetorial das equações com o qual se resolve a deficiência da geométrica do *coeficiente angular*

Entretanto o *coeficiente angular*, pese sua restrição, ele tem usos e deixe-me mostrar algumas retas associadas aos seus *coeficientes angulares*

$$y = -x + 1; \text{ coeficiente angular } -1, \text{ passando em } (0, 1); \quad (430)$$

$$y = x + 1; \text{ coeficiente angular } 1, \text{ passando em } (0, 1); \quad (431)$$

$$y = 2x - 1; \text{ coeficiente angular } 2, \text{ passando em } (0, -1); \quad (432)$$

$$y = -2x - 1; \text{ coeficiente angular } -2, \text{ passando em } (0, -1); \quad (433)$$

$$y = 0.5x - 2; \text{ coeficiente angular } \frac{1}{2}, \text{ passando em } (0, -2); \quad (434)$$

Você pode ver os gráficos destas retas na figura (fig 63), página 150,

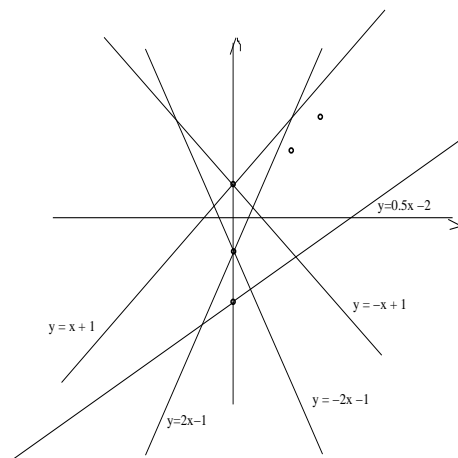


Figura 63: retas e suas equações no plano

O padrão para todas estas equações são as equações

$$y = mx + n \text{ coeficiente angular } m, \text{ passando em } (0, n); \quad (435)$$

$$y = m(x - a) + b; \text{ coeficiente angular } m, \text{ passando em } (a, b); \quad (436)$$

$$y - b = m(x - a); \text{ coeficiente angular } m, \text{ passando em } (a, b); \quad (437)$$

e os dois números  $m, n$  recebem os nomes *coeficiente angular* e *coeficiente linear*, respectivamente.

Se você deslizar estas figuras, paralelamente, usando um esquadro, para que passem no ponto  $(0, 0)$ , na origem dos eixos, a equação das retas assim obtidas,



serão:

$$y = -x; \text{ coeficiente angular } -1, \text{ passando em } (0,0); \quad (438)$$

$$y = x; \text{ coeficiente angular } 1, \text{ passando em } (0,0); \quad (439)$$

$$y = 2x; \text{ coeficiente angular } 2, \text{ passando em } (0,0); \quad (440)$$

$$y = -2x; \text{ coeficiente angular } -2, \text{ passando em } (0,0); \quad (441)$$

$$y = 0.5x; \text{ coeficiente angular } \frac{1}{2}, \text{ passando em } (0,0); \quad (442)$$

e você pode ver o resultado obtido com `xfig` na figura (fig 64), página 151,

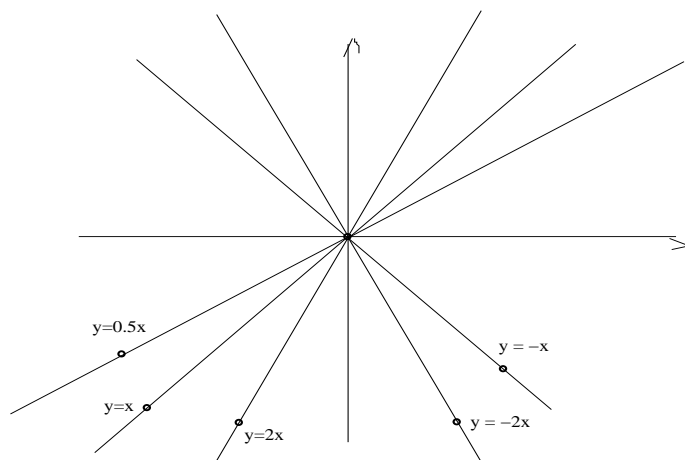


Figura 64: retas e coeficiente angular

Você pode repetir os gráficos que se encontram nas figuras (fig 63) e (fig 64) com este código do `gnuplot`:

```
m = 1; a=-2; b = -3; f(x) = m*(x-a);
set arrow from 0,-15 to 0,15 nohead;
plot f(x),0;
```

trocando os valores de  $a, b, m$  para ver distintas retas. Se você quiser ver os gráficos simultâneos, inclua índices, observe como:

```
m1 = 1; a1=-2; b1 = -3; f1(x) = m1*(x-a1);
set arrow from 0,-15 to 0,15 nohead;
plot f1(x),0;
m2 = 2; a2=-2; b2 = -3; f2(x) = m2*(x-a2);
plot f2(x), f1(x), 0;
m3 = 3; a3=-2; b3 = -3; f3(x) = m3*(x-a3);
plot f3(x),f2(x), f1(x), 0;
```

e você pode alterar o ponto onde passa algumas das retas, raspe o texto e cole num terminal do `gnuplot` ou num arquivo para depois chamar este arquivo com `gnuplot`.

- **equação diferencial** É uma equação em que a *principal operação* é a diferenciabilidade e conseqüentemente a variável é uma função, é portanto uma equação funcional. Estas equações, quando tiverem solução, em geral têm um *conjunto infinito de soluções*.

- Se as funções-solução destas equações forem univariadas, elas são chamadas *equações diferenciais ordinárias*, resumidamente EDO. Ver **EDO**.
- se as funções-solução forem multivariadas, elas são chamadas *equações diferenciais parciais*, resumidamente EDP. Ver **EDP**

Uma equação diferencial é uma expressão envolvendo uma função,  $y$ , a *incognita*

$$y = y(x_1, \dots, x_n); \quad (443)$$

$$F(x_1, \dots, x_n, y_{j,k}) = 0; \quad (444)$$

$$y_{j,k} = \frac{\partial^j y}{\partial x_k} \text{ derivadas parciais de } y; \quad (445)$$

as operações usuais da Álgebra, representada pela função  $F$ , no caso das equações diferenciais parciais, ou

$$y = y(x); \quad (446)$$

$$F(x, y_1, \dots, y_n) = 0; \quad (447)$$

$$y_j = \text{derivada de ordem } j \text{ de } y; \quad (448)$$

e  $F$  é uma expressão “algébrica” envolvendo estas variáveis.

Os coeficientes podem ser constantes ou outras funções e assim separamos em *equações a coeficiente constantes ou variáveis*. O mais alto índice de derivação é dito a ordem da equação diferencial.

Se  $y$  for uma função de várias variáveis, as derivadas envolvidas serão as parciais, e equação é chamada EDP, equação diferencial parcial, ver **EDP**,

Se  $y$  for uma função univariada, as derivadas envolvidas serão as *ordinárias*, e equação é chamada EDO, equação diferencial ordinária, ver **EDO**,

O adjetivo *parcial* tem o que ver com as *derivadas parciais* envolvidas na definição da equação e serve para colocar em *oposição* as *equações diferenciais ordinárias*, ver **EDO**.

Aplicando o *teorema da função implícita* à equação (444) ou (447) é, em princípio, possível explicitar a maior ordem de derivação.

Embora seja possível identificar *equações diferenciais ordinárias* dentro de um *problema de equações diferenciais parciais*, as metodologias para encontrar soluções diferem de forma fundamental. Um pouco desta ideia pode ser apresentada aqui.

As *equações diferenciais ordinárias* podem ser inicialmente descritas de uma forma muito simples, que embora não represente completamente estas equações

servem como uma forma inicial de descrição: toda função univariada *bem comportada* é uma derivada, ou dito em outras palavras, *tem uma primitiva* que pode ser obtida usando-se o *teorema fundamental do Cálculo* com uma *condição inicial*. Isto é absolutamente falso quando se passa ao caso multivariado.

A maneira mais fácil de entender esta diferença passa por dos mais bonitos teoremas da Análise, o *teorema de Green*, que tem duas versões: a trivial e a não trivial! Na versão trivial, aplicada a funções bivariadas, o teorema de Green descreve duas integrais nulas:

$$\int_{\gamma} \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy = \int_{\Omega} \int \left( \frac{\partial^2 F}{\partial y \partial x} - \frac{\partial^2 F}{\partial x \partial y} \right) dx dy; \gamma = \partial\Omega; \quad (449)$$

em que  $\gamma$  é uma curva fechada que é a fronteira da região  $\Omega$ ,  $\gamma = \partial(\Omega)$ , e desta forma separa as funções em duas classes:

- no caso trivial as funções cujas integrais (de linha) não dependem do caminho escolhido entre dois pontos do domínio, ou equivalentemente, suas integrais (de linha) se anulam sobre qualquer caminho fechado. Estas funções têm primitivas que podem ser calculadas com uma integral de linha. Quando uma integral de linha não depende do caminho  $\gamma$  que une dois pontos do plano,

$$\int_{\gamma} P(x, y) dx + Q(x, y) dy$$

dizemos que  $(P, Q) = \left( \frac{\partial F}{\partial x}, \frac{\partial F}{\partial y} \right)$  é um campo vetorial conservativo. Como a integral não depende do caminho, ela se anula sobre qualquer curva fechada e corresponde, na Física, à lei de conservação da energia: a circulação sobre um contorno fechado qualquer, é nula.

- no caso não trivial as funções cujas integrais (de linha) dependem do caminho escolhido entre dois pontos do domínio, ou equivalentemente, suas integrais (de linha) não se anulam sobre algum caminho fechado  $\gamma$

$$\int_{\gamma} \frac{\partial P}{\partial x} dx + \frac{\partial Q}{\partial y} dy = \int_{\Omega} \int \left( \frac{\partial^2 Q}{\partial y \partial x} - \frac{\partial^2 P}{\partial x \partial y} \right) dx dy; \gamma = \partial\Omega; \quad (450)$$

A curva  $\gamma$  é a fronteira,  $\partial\Omega$ , da região  $\Omega$ . Estas funções não têm primitivas, ou ainda, não é possível encontrar-se  $z = F(x, y)$  diferenciável duas vezes continuamente, tal que

$$\frac{\partial F}{\partial x} = P(x, y); \frac{\partial F}{\partial y} = Q(x, y); \quad (451)$$

Diz-se, neste caso, que  $(P, Q)$  não define um *potencial*, ou ainda que é um *campo vetorial não conservativo*.

As primeiras tem primitiva e a expressão da integral de linha do teorema de Green, na sua versão trivial é uma generalização direta do cálculo de primitivas de funções univariadas ou do teorema fundamental do Cálculo.

A segunda forma do teorema de Green diz que integral depende do contorno e neste caso a função  $(x, y) \mapsto (P(x, y), Q(x, y))$  não tem primitiva. Este é um primeiro exemplo de equação diferencial parcial

$$P(x, y)dx + Q(x, y)dy = 0; \quad (452)$$

$$\frac{\partial F}{\partial x} = P(x, y); \frac{\partial F}{\partial y} = Q(x, y); \quad (453)$$

que é chamada *exata* e que se puder ter solução, se for exata, tem como solução a família das curvas de nível de  $z = F(x, y)$ . Como a solução é uma família de curvas esta equação é considerada uma equação diferencial ordinária sendo estudada como tal.

O teorema de Green separa as funções vetoriais de variável vetorial,  $(P, Q)$  em duas classes: (1) as que são derivadas e então  $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$  e (2) as que não são derivadas,  $\frac{\partial P}{\partial y} \neq \frac{\partial Q}{\partial x}$ . É esta a diferença com o caso univariado mencionado acima, no caso univariado “qualquer função que não seja muito descontínua” tem primitiva. As funções vetoriais, por exemplo as que forem da forma  $(P, Q)$ , podem ser contínuas sem terem uma primitiva, elas podem não ser derivadas. O teste é feito pelo Teorema de Green.

Este exemplo de equação à derivadas parciais tem um significado especial de mostrar que a solução de uma equação diferencial parcial depende da fronteira de uma região, esta é outra das diferenças entre equações diferenciais ordinárias e parciais, a dependência da fronteiras ou ainda como se diz, *dependência das condições de contorno*.

Há algumas poucas equações diferenciais parciais de fácil solução e que tem importância prática porque sempre aparecem nas tentativas de resolver as mais difíceis:

$$u_y = \frac{\partial u}{\partial y} = 0; u(x, y) = f(x); \quad u_{xy} = 0; u(x, y) = f(x) + g(y) \quad (454)$$

em ambos os casos,  $f, g$  são funções arbitrárias, diferenciáveis, univariadas. Estas duas equações são do tipo *homogênea e linear* e podemos ver, neste caso, facilmente a influência da Álgebra Linear na solução das equações diferenciais. Para isto considere a equação mais difícil

$$u_{xy} = h(x, y) \quad (455)$$

Sua solução geral é a *solução geral da equação homogênea*,  $u_{xy} = 0$  mais *uma solução particular da geral* que é a expressão da solução geral de um sistema linear (da Álgebra Linear):

$$u(x, y) = \int_a^x \int_b^y h(s, t) ds dt + f(x) + g(y) \quad (456)$$

basta calcular  $u_{xy}$  na equação (456) para verificar que esta é de fato uma solução da equação (455).

As equações diferenciais parciais lineares (EDPL ou LPDE) se classificam com nomes tirados da Geometria Analítica, em *hiperbólicas*, *parabólicas*, *elípticas* porque podem ser expressas com *operadores lineares* que tem o aspecto das equações destas cônicas.

Este ramo da Matemática é possivelmente o mais profundamente estudado e por uma quantidade muito grande de matemáticos. Por um lado, porque apresenta dificuldades muito grandes, e por outro lado, por seu envolvimento com todos os ramos das ciências onde se possa verificar uma *dinâmica* ou variação, ou seja pela enorme influência junto as demais ciências.

Para desenvolver as diversas teorias dentro deste campo do conhecimento matemático, é necessário usar ferramentas avançadas da Álgebra, como Teoria dos grupos, dos semi-grupos, anel e módulos assim como da Geometria Diferencial.

Em 1945 Laurent Schwartz, um matemático francês, conseguiu sintetizar, simultaneamente com um matemático português, Sebastião e Silva, a Teoria das Distribuições, criando um novo objeto matemático, diferente de funções, e por algum tempo se teve a quase certeza de que se havia encontrado uma metodologia para resolver as equações diferenciais parciais. As *distribuições* são ainda chamadas de *funções generalizadas*. Ainda não sabemos resolver equações diferenciais parciais com as distribuições, mas este método abstrato tem se mostrado muito prático nas soluções aproximadas de equações diferenciais.

As equações à derivadas parciais se classificam em duas grandes classes, as *lineares* e as *não lineares*. Uma forma de determinar se a equação é *linear* ou *não linear* consiste em transformar a equação (444) ou (447) em um sistema de  $m$  equações em que  $m$  é a ordem da equação diferencial. Se a equação for linear o resultante sistema será um sistema linear nas incógnitas  $y_{j,k}$  ou respectivamente  $y_k$ .

Como no caso das *equações diferenciais ordinárias*, nós sabemos tudo sobre as lineares, *exceto resolvê-las*, porque dependemos de outros ramos da Matemática que ainda não conseguiram desenvolver os métodos necessários... Já as equações *não lineares* quase que formam, cada uma delas, a sua própria teoria.

Com a transformada de Fourier, que alterna *derivada* em *produto por convolução*, é possível identificar as equações diferenciais lineares parciais com expressões vindas da Geometria Analítica que definem as cônicas, *parábolas*, *hipérbolas* e *elipses* estabelecendo uma grande classificação das equações lineares ou não lineares.

Desde o advento da computação científica, na década de 50, avançou muito a pesquisa de soluções aproximadas de equações diferenciais com alguns resultados teóricos que surgiram com a possibilidade de compreender computacionalmente algumas soluções. O método dos elementos finitos é um dos métodos computacionais que mais tem se mostrado promissor quanto a construir aproximações de soluções e na construção de visualizações gráficas das mesmas.

---

- equação diferencial ordinária. Ver **EDO**.

- equação diferencial parcial Ver EDP

- equação do círculo, é um item da *geometria analítica*. Usando uma *linguagem moderna*, podemos dizer que a *geometria analítica* “digitaliza” algumas figuras geométricas criando equações o que torna possível, por exemplo, usar estas figuras em programas de computação. A *geometria analítica* pode ser vista como uma introdução à *geometria algébrica* no sentido em que esta disciplina da Matemática trata os entes geométricos que podem ser definidos através de equações algébricas, mais precisamente como expressões polinomiais.

O círculo é uma cônica, quer dizer uma curva determinada pela interseção dum plano com um cone de duas folhas, e aqui estou me referindo a uma curva espacial. Confira também a *equação da elipse* para uma análise destas inteseções que geral elipses, parábolas, hipérbolos ou retas. . .

Há um erro neste texto que será corrigido mais a frente, mas procure descobri-lo você mesma.

Equação geral do círculo

A definição , da geometria, de *círculo* é “o lugar geométrico dos pontos do plano que equidistam dum ponto chamado centro”. Deixe-me identificar o centro com as coordenadas  $C = (a, b)$  e um ponto genérico,  $P$ , do plano com as coordenadas  $(x, y)$ . A frase da geometria que define círculo pode então ser traduzida algébricamente como

$$d(P, C) = r; \sqrt{(x - a)^2 + (y - b)^2} = r \quad (457)$$

em que  $r$  é a *distância* constante entre  $P$  e  $C$ . A figura (fig 65), página 156, lhe apresenta a ideia , da geometria, contida na definição do círculo.

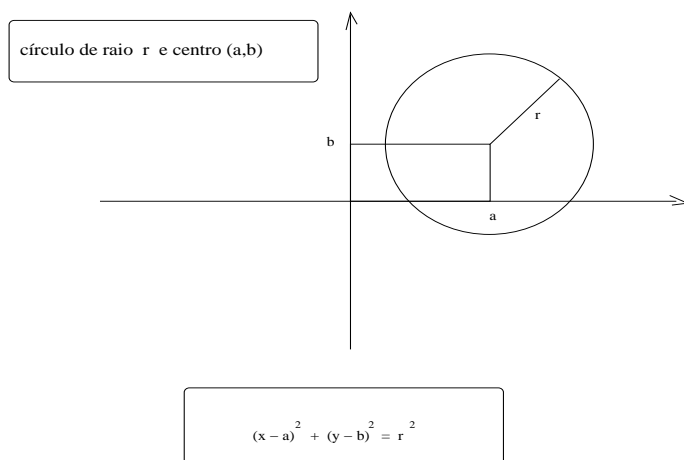


Figura 65: círculo de raio  $r$  e centro  $(a, b)$

A expressão na equação (eq. 457) poderia ser considerada a equação do círculo, mas não é nesta forma que vamos encontrá-la, habitualmente, mas sim

na forma

$$P(x, y) = (x - a)^2 + (y - b)^2 = r^2; \quad (458)$$

entretanto qualquer uma das duas equações é perfeita para representar o círculo apenas no primeiro caso podem surgir problemas se for usada num programa de computador pela presença da raiz que exige que o seu “parâmetro” seja positivo.

Um caso importante da equação do círculo é quando a origem é o centro do círculo quando a equação fica

$$P(x, y) = x^2 + y^2 = r^2; \quad (459)$$

Basta que você substitua  $a = 0, b = 0$  na equação (eq. 458).

### **Identificando um lugar geométrico**

O objetivo da *geometria analítica* é duplo, por um lado, codificar em equações algumas figuras geométricas, as cônicas, e por outro lado, identificar numa expressão algébrica “geral” que tipo de figura algébrica se encontra codificada na expressão. Este segundo problema não é nada fácil e na verdade melhor faz parte do ambiente matemático mais amplo que é a *geometria algébrica*, mas é exatamente este segundo problema que aparece na maioria dos exercícios de *geometria analítica*: “*dado o polinômio*

$$P(x, y) = Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \quad (460)$$

em que  $A, B, C, D, E, F$  são constantes dadas, identifique qual é a cônica que a equação (eq. 460) representa”.

No caso do círculo posso encontrar as regras para fazer esta identificação se expandir a equação (eq. 458) quando vou encontrar a expressão

$$P(x, y) = x^2 + 2ax + a^2 + y^2 + 2by + b^2 - r^2 = 0 \quad (461)$$

e comparando as duas equações deduzo que  $A = C$  e se  $A \neq C$  a expressão pode representar uma elipse, uma hipérbole ou nenhuma cônica. Também

$$B = 0; D = 2a; E = 2b; F = a^2 + b^2 - r^2; \quad (462)$$

e desta forma, se a equação (eq. 461) representar um círculo é então possível encontrar o centro,  $C$ , e o raio,  $r$ . Analise um exemplo, considere a sequência de cálculos, que vou considerar “validados” porque  $A = C$ , então poderá ser um círculo.

$$P(x, y) = x^2 - 4x + y^2 + 6y - 3 = 0; \quad (463)$$

$$P(x, y) = x^2 - 4x + 4 - 4 + y^2 + 6y + 9 - 9 - 3 = 0; \quad (464)$$

$$P(x, y) = x^2 - 4x + 4 + y^2 + 6y + 9 - 9 - 3 - 4 = 0; \quad (465)$$

$$P(x, y) = x^2 - 4x + 4 + y^2 + 6y + 9 - 16 = 0; \quad (466)$$

$$P(x, y) = (x - 2)^2 + (y + 3)^2 - 4^2; \quad (467)$$

$$(x - 2)^2 + (y - (-3))^2 = 4^2; \quad (468)$$

que é a representação algébrica do círculo de centro  $C = (2, -3)$  e raio  $r = 4$ . Na sequência de contas das equações (eq. 463) -(eq. 468) usei uma técnica denominada completção de quadrados em vez de procurar identificar os coeficientes da equação geral (eq. 460). Quando identificamos  $A = C$ , é mais prático seguir por este método para verificar se a expressão representa um círculo. Para dominar as técnicas envolvidas no trabalho com *geometria analítica*, como em qualquer disciplina ou atividade humana, a regra é bem simples: fazer uns mil exercícios... e você os encontra aos milhares em textos livres distribuídos na Internet.

O erro mencionado inicialmente se encontra em que, ao escrever-se uma expressão  $P(x, y)$  não está implicitamente identificada a dimensão do espaço em que esta expressão “codifica” algum objeto. Por exemplo, no espaço de dimensão 3,

$$x^2 + y^2 = 1;$$

representa o cilindro cuja base é o círculo unitário do plano  $XOY$  e a geratriz é o eixo  $OZ$ . Apenas mencionar esta equação sem identificar o universo em que ela estiver sendo considerada deixa uma indefinição. Este erro é muito comum nos textos, mas é difícil de ser corrigido sem criar uma redação muito complexa, em geral se admite implicitamente que se houver apenas duas variáveis, então a dimensão do universo é dois.

É interessante responder a uma pergunta que sempre é feita. Porque a  $P(x, y) = 0$  é que representa a equação do círculo e não apenas  $P(x, y)$ ? É que escrevendo “apenas  $P(x, y)$ ” você tem uma expressão livre que pode assumir qualquer valor em função dos valores dados às variáveis  $x, y$  e conseqüentemente  $P(x, y)$  se refere a qualquer ponto  $(x, y, z)$  do  $\mathbf{R}^3$  em que  $z = P(x, y)$ . A restrição  $P(x, y) = 0$  limita esta expressão a  $z = 0$ . Outra forma de dizê-lo, é que  $P(x, y) = 0$  corresponde ao sistema de equações

$$\begin{cases} z = P(x, y); \\ z = 0; \end{cases} \quad (469)$$

em que a primeira equação representa o gráfico da função bivariada  $z = P(x, y)$  e a segunda equação representa o gráfico da função bivariada  $z = 0$  cuja gráfico é o plano  $XOY$  ou seja, a equação (eq. 469) representa a interseção de duas superfícies o que pode ser uma curva. As curvas num espaço de dimensão maior, como aqui seria o  $\mathbf{R}^3$ , são determinadas por interseção de superfícies. Assim, o erro mencionado acima fica amenizado com um sistema de equações como a equação (eq. 469). Mas indefinição persiste... uma curva é uma variedade de dimensão 1 e a expressão  $P(x, y) = 0$  apenas sugere que existam duas variáveis o que daria para calcular a dimensão por uma regra prática que diz que a dimensão é uma unidade menor do que o número de variáveis. Mas escrever

$$x^2 - 4x + y^2 + 6y - 3 = 0$$

não contém nenhuma informação sobre a variável  $z$  ou significa que ela pode ter qualquer valor então esta expressão no  $\mathbf{R}^3$  é um cilindro uma vez que não



está estabelecida nenhuma restrição sobre a mesma e existe restrições sobre os valores de  $x, y$ .

Entre escrever textos extremamente complicados termina valendo o bom senso que é determinado pelo contexto.

### O círculo trigonométrico

O círculo trigonométrico, ou círculo unitário, é uma figura geométrica que aparece em muitas situações merecendo assim um tratamento especial. Posso partir da equação (eq. 459), basta dar o valor 1 ao  $r = 1$

$$P(x, y) = x^2 + y^2 = 1; \quad (470)$$

$$x = \cos(\theta); y = \sin(\theta); (x, y) = (\cos(\theta), \sin(\theta)); \quad (471)$$

A afirmação contida na equação (eq. 471) é a da relação fundamental da trigonometria e vale sempre que a soma de dois quadrados for 1. Explicando melhor, se a soma de dois quadrados for 1, como na equação (eq. 470) então os números  $x, y$  são, respectivamente coseno e seno de algum ângulo. Temos assim a forma polar do círculo.

- equação do plano é um tópico da Geometria Analítica.

É comum a afirmação de que  $Ax + By + Cz + D = 0$  é a equação do plano. Infelizmente isto é falso, poderia ser a equação dum *hiperplano*. Para evitar esta ambiguidade é preciso acrescentar “no espaço de dimensão 3” porque, num espaço de dimensão  $n; n > 3$ , por exemplo num espaço de dimensão quatro, esta equação representa um hiperplano.

Entenda porque,  $x = 4$  não é a equação duma reta! se for no  $\mathbf{R}^2$ , é verdade, é uma reta, mas se for no  $\mathbf{R}^3$  é um plano formado dos pontos  $\{(4, y, z); y, z \in \mathbf{R}\}$  uma vez que a única *restrição* imposta pela equação é sobre a variável  $x$ .

Uma equação deste tipo, combinação linear de coeficientes e variáveis, um polinômio do primeiro grau, representa uma *variedade linear*, retas, planos, hiperplanos são exemplos de *variedades lineares*. Representa a *variedade linear* maximal que as restrições permitirem.

A equação paramétrica do plano, é mais precisa por deixar claro a dimensão. Um plano é uma superfície, confira *superfície*, que é a imagem de uma função que depende de dois parâmetros, o que caracteriza sua dimensão como sendo 2. Genericamente, uma superfície  $\alpha$  seria um objeto definido por

$$\alpha(s, t) = (x_1(s, t), \dots, x_n(s, t)); n \text{ é a dimensão do espaço}; \quad (472)$$

na equação (eq. 472), uma superfície definida no  $\mathbf{R}^n$ , é uma *variedade* de dimensão 2 *imersa* no  $\mathbf{R}^n$ . Por exemplo

$$\alpha(s, t) = (s, t, \dots, x_n(s, t)); \quad (473)$$

$$x_1(s, t) = s; \quad (474)$$

$$x_2(s, t) = t; \quad (475)$$

$$k > 2; x_k(s, t) = s^k t^k; \quad (476)$$

*Plano* é a imagem de uma superfície cujas equações,  $x_k(s, t)$ , sejam todas do primeiro grau.

Assim

$$\alpha(s, t) = (a_1s + b_1t, \dots, a_ns + b_nt); a_k, b_k \in \mathbf{R}; \quad (477)$$

$$\beta(s, t) = (s, t, 3, 3s + 4t); \in \mathbf{R}^4; \quad (478)$$

$$\gamma(s, t) = (s, t, 3s + 4t); \in \mathbf{R}^3; \quad (479)$$

$$\eta(s, t) = (a_1s + b_1t + c_1, a_2s + b_2t + c_2, a_3s + b_3t + c_3) \in \mathbf{R}^3; \quad (480)$$

são equações dos planos, variedades lineares de dimensão 2

- $\alpha^*$  imerso no  $\mathbf{R}^n$ ;
- $\beta^*$  imerso no  $\mathbf{R}^4$ ;
- $\gamma^*, \eta^*$  imersos no  $\mathbf{R}^3$ ;

Observe que na equação (eq. 480) temos a expressão de todas as coordenadas com equações do primeiro grau, logo uma variedade linear, um lugar geométrico do  $\mathbf{R}^3$  que depende de dois parâmetros portanto uma variedade linear de dimensão 2 que é o mesmo que podemos ver nas duas equações anteriores.

Usamos o mesmo nome da função que define a variedade acrescentando-lhe asterísco para designar a imagem geométrica da variedade, em outras palavras

$$graf(\alpha) = \alpha^*; graf(\beta) = \beta^*; graf(\gamma) = \gamma^*;$$

O plano  $\gamma^*$  definido na equação (479) “passa” na origem, enquanto que o plano  $\beta^*$  definido na equação (478) passa no ponto  $(0, 0, 3, 0) \in \mathbf{R}^3$  e não passa na origem.

Vamos fazer uma comparação com a reta cuja visualização é fácil, para depois extrapolarmos para o caso do plano. Reta é uma *curva* que tem um único *coeficiente angular*, e sua direção é uma reta paralela a um vetor dado que define os *coeficientes diretores* desta reta. Um plano tem *coeficiente angular* em qualquer direção que escolhermos, uma infinidade de *coeficientes angulares*, mas tem *um vetor que lhe é perpendicular* e quaisquer outros planos que sejam perpendiculares a este vetor são paralelos entre si.

Portanto é importante determinar um vetor que seja perpendicular a um plano, ele determina uma classe de planos paralelos num determinado espaço e é a forma de construir sua equação.

A *Geometria Euclidiana* nos diz que dado uma direção  $r$ , e um ponto,  $P$ , existe um único plano,  $\pi$  passando por  $P$  e perpendicular a  $r$ . É difícil de ver-se isto graficamente, mas é fácil ver o caso da reta, no  $\mathbf{R}^2$ , a figura (fig 66), página 161, mostra a direção representada pelo vetor  $\vec{r}$  e um ponto do plano por onde passa uma única reta que é perpendicular à  $\vec{r}$ . Também há um único plano passando por  $P$  e perpendicular  $\vec{r}$ .

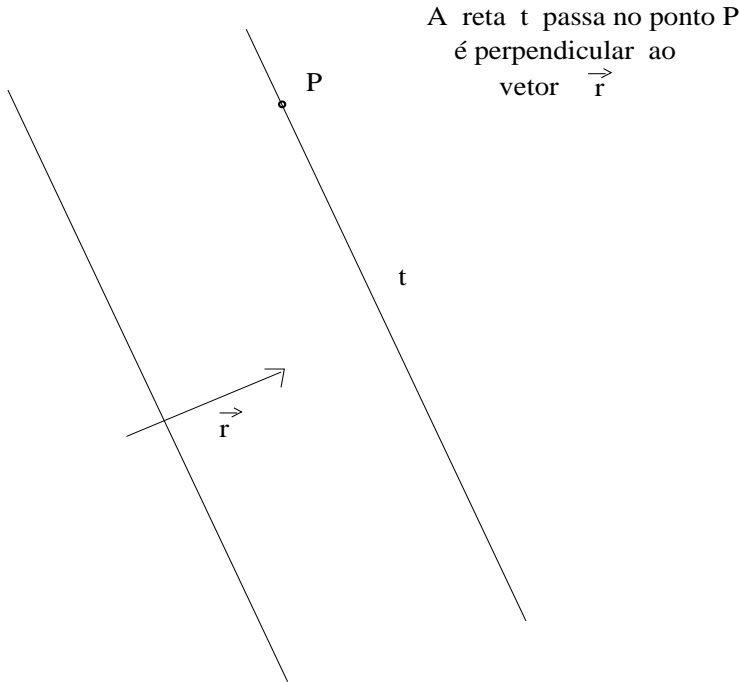


Figura 66: A reta  $t$  passa no ponto  $P$  é perpendicular  $\vec{r}$

É neste ponto em que as *geometrias não euclidianas* se separam da *geometria euclidiana*<sup>19</sup>. Por exemplo, na geometria esférica as retas são os círculos máximos, e as paralelas todas elas se encontram nos polos. Na *geometria euclidiana* elas se encontram no infinito. . .

Mas estamos falando da *Geometria Analítica Euclidiana!*

Retornando à equação do plano, precisamos identificar um vetor perpendicular ao plano para podermos construir sua equação, ou, vice-versa, identificar na equação o vetor perpendicular para poder imaginar o plano!

Na equação  $Ax + By + Cz = 0$  podemos identificar um *produto escalar*:

$$Ax + By + Cz = \langle (A, B, C), (x, y, z) \rangle = 0; \quad (481)$$

$$(A, B, C) \perp (x, y, z); \quad (482)$$

$$A(x - a) + B(y - b) + C(z - c) = 0; \quad (483)$$

$$\langle (A, B, C), (x - a, y - b, z - c) \rangle = 0; \quad (484)$$

$$(A, B, C) \perp (x - a, y - b, z - c); \quad (485)$$

Na equação (eq. 481) estamos traduzindo uma soma nula com o *produto*

<sup>19</sup>E se registra um famoso erro de um famoso e importante matemático, Lagrange, que tentou provar o axioma das paralelas.

*escalar* e concluindo que o vetor  $(A, B, C)$  é perpendicular a um vetor genérico do lugar geométrico estudado. Esta propriedade caracteriza as *variedades lineares* elas são o lugar geométrico dos pontos do espaço perpendiculares a um vetor dado. Na equação (eq. 483) aplicamos uma translação por  $(a, b, c)$  concluindo que o vetor  $(A, B, C)$  é perpendicular ao vetor  $(x - a, y - b, z - c)$ .

Vamos interpretar esta equação definindo um lugar geométrico do  $\mathbf{R}^3$ , então um plano.

- no plano isto significa *reta*;
- no espaço tridimensional, significa *plano*;
- num espaço de dimensão  $n \geq 4$  significa um hiperplano;

Na equação (eq. 482) aplicamos uma translação

$$Ax + By + Cz = 0 \xrightarrow{(a,b,c)} A(x - a) + B(y - b) + C(z - c) = 0$$

um movimento rígido do plano, sem rotação, ou com rotação zero. Esta equação representa um plano paralelo ao da equação anterior, (eq. 481). Se fizermos  $(x, y, z) = (a, b, c)$  vemos que o ponto  $(a, b, c)$  satisfaz à equação (eq. 482) portanto  $(a, b, c)$  pertence à variedade definida por esta equação que é um plano que passa no ponto  $(a, b, c)$ . Demonstramos o teorema

**Teorema 6 (Equação do plano)** *Equação do plano*

No  $\mathbf{R}^3$  a equação do plano perpendicular ao vetor  $(A, B, C)$  passando pelo ponto  $(a, b, c)$  é

$$A(x - a) + B(y - b) + C(z - c) = 0 \equiv Ax + By + Cz = D; D = Aa + Bb + Cc;$$

Para obter uma equação paramétrica do plano, partindo da equação (eq. 481), explicita  $z$  como função de  $(x, y)$ :

$$z = -\frac{Ax}{C}x - \frac{By}{C} = f(x, y) \Rightarrow (x, y, f(x, y)); \quad (486)$$

$$(s, t, f(s, t)) = (s, t, -\frac{As}{C} - \frac{Bt}{C}); \quad (487)$$

A expressão se diz *equação paramétrica* do plano, e habitualmente é escrita com os símbolos  $s, t$  representando os parâmetros e obtivemos por outro caminho a equação (eq. 479).

Para compreender a terminologia *hiperplano* observe as semelhanças:

- Uma reta, um espaço de dimensão 1, é dividido por um ponto, em dois semi-espaços chamados semi-retas, porque um ponto é um *objeto maximal* das retas;
- Num plano, as retas dividem o espaço em dois semi-espaços chamados semi-planos significando que as retas são os objetos maximais dos planos;

- num espaço tridimensional, os *planos* dividem o espaço em dois semi-espacos para os quais não temos terminologia geométrica devido ao fato de sermos, culturalmente, *prisioneiros da terceira dimensão*. Os planos são as variedades lineares maximais dos espaços tridimensionais.
- num espaço de dimensão  $n \leq 4$  chamamos de hiperplano as variedades lineares maximais que dividem estes espaços em dois semi-espacos para os quais a nossa linguagem tridimensional não tem vocábulos.

- equação linear, indexequação linear confira *sistema linear*.

- equação normal da reta A equação cartesiana da reta é a expressão

$$Ax + By + C = 0; \quad (488)$$

que também pode representar qualquer variedade linear portanto é preciso que se diga que é a equação duma figura plana. Associada a esta equação temos a equação da reta, *variedade linear*, que passa na origem

$$Ax + By = 0; \quad (489)$$

em que podemos claramente identificar o produto escalar

$$\langle (x, y), (A, B) \rangle = 0 \quad (490)$$

dum vetor genérico,  $(x, y)$ , do lugar geométrico que esta equação representa, pelo vetor dado  $(A, B)$ . Como o nenhum desses vetores deve ser nulo<sup>20</sup>, então quem é zero é o coseno do ângulo entre os vetores, conduzindo, sucessivamente, à conclusão de que eles são perpendiculares e que este lugar geométrico é uma reta,  $r$ , que passa na origem.

O *lugar geométrico* correspondente à equação (eq. 488) é uma reta paralela a esta da equação (eq. 489).

Como o vetor  $(A, B)$  é perpendicular à  $r$  podemos calcular o vetor unitário de sua direção

$$\vec{N} = \left( \frac{A}{\sqrt{A^2 + B^2}}, \frac{B}{\sqrt{A^2 + B^2}} \right); \quad (491)$$

produzindo assim um *vetor normal* à reta  $r$  ou, de forma mais geral, a uma reta qualquer como esta da equação (eq. 488).

*Vetor normal* é qualquer vetor unitário que é perpendicular a alguma variedade, linear ou não. Como eles se encontram dentro duma esfera unitária de algum espaço, então suas coordenadas são do tipo  $\sin, \cos$  e neste caso são

$$\vec{N} = (\cos(\alpha), \sin(\alpha)); \quad (492)$$

em que  $\alpha$  é ângulo entre a reta  $r$  e o eixo  $OX$ . A figura (fig 67), página 164, ilustra o significado do vetor normal para o cálculo da distância duma reta à

<sup>20</sup>Salvo algum ponto quando o zero pertencer ao lugar geométrico.

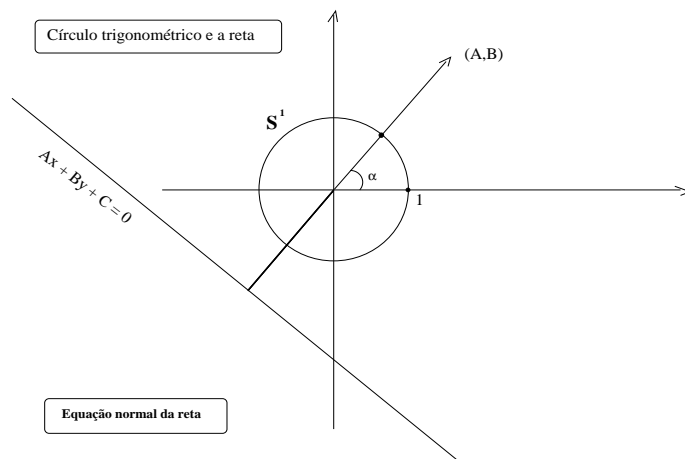


Figura 67: vetor normal e distância da reta à origem

origem dos eixos.

Uma tal equação é importante porque se dividirmos a equação (eq. 488) pelo módulo do vetor  $(\vec{A}, \vec{B})$  encontramos a fórmula da distância da reta

$$Ax + By + C = 0$$

à origem:

$$\frac{A}{\sqrt{A^2+B^2}} - \frac{B}{\sqrt{A^2+B^2}} - \frac{C}{\sqrt{A^2+B^2}} = 0; \quad (493)$$

$$\cos(\alpha)x + \sin(\alpha)y + \frac{C}{\sqrt{A^2+B^2}} = 0; \quad (494)$$

$$-\frac{C}{\sqrt{A^2+B^2}} = \cos(\alpha)x + \sin(\alpha)y; \quad (495)$$

Na equação (eq. 495) se encontra uma constante no primeiro membro e no segundo membro a projeção do vetor genérico da reta na direção do vetor normal, portanto esta constante, em módulo

$$\left| \frac{C}{\sqrt{A^2+B^2}} \right| \quad (496)$$

é a distância da reta  $Ax + By + C = 0$  à origem.

Considerando a observação inicial que  $Ax + By + C = 0$  é uma variedade linear portanto sua dimensão depende do espaço em que ela estiver considerada, estes cálculos servem *ipsis literis* como demonstração da fórmula da distância de qualquer variedade linear à origem dos eixos, apenas considerando o módulo

do vetor perpendicular:

$$\sum_{k=1}^n A_k x_k + A_{n+1} = 0; \quad (497)$$

$$|(A_1, \dots, A_n)| = \sqrt{\sum_{k=1}^n A_k^2} = R; \quad (498)$$

$$\vec{N} = \frac{1}{R}(A_1, \dots, A_n); \quad (499)$$

$$d(V, 0) = \left| \frac{A_{n+1}}{R} \right|; \quad (500)$$

e aparece na equação (eq. 497) a expressão que generaliza a distância da reta à origem agora representando a distância duma *variedade linear* de dimensão  $n$  à origem.

Podemos obter também a distância dum ponto  $P$  que se encontre sobre uma variedade  $P \in \mathcal{S}$ , por exemplo uma superfície, à origem usando cálculos semelhantes a estes. Para isto precisamos que o ponto  $P$  em questão admita uma *variedade linear tangente* e sendo assim basta calcularmos a distância desta *variedade linear tangente* à origem. Mas observe que esta distância somente pode ser considerada a distância da variedade  $\mathcal{S}$  à origem nas seguintes condições:

- $\mathcal{S}$  for convexa;
- O *variedade linear tangente* no ponto  $P$  divide o espaço em dois semi-espacos, isto sempre será verdadeiro, e a variedade  $\mathcal{S}$  esteja no semi-espaco diferente daquele em que estiver a origem.

Esta questão se situa no estudo da *convexidade* e de *variedade diferenciável*.

Embora na literatura se ponha algumas vezes ênfase no vetor unitário e que se chama esta expressão como *equação normal da reta* a real importância da “*equação normal da reta*” vem do fato de que ela oferece um substituto ao conceito de coeficiente angular que nem sempre existe oferecendo um vetor perpendicular, ou um vetor normal à reta. Esta equação está pronta para ser usada em qualquer dimensão e para oferecer um vetor tangente a uma variedade num espaço qualquer. Aplicações: *obter a direção que tangencia uma variedade no espaço, direção tangencial para expelir um módulo numa nave espacial*. Confira Rosetta, a nave espacial.

- **equação polar da reta** Uma reta paralela ao eixo  $OY$  não tem *coeficiente angular* quando discutimos sua *equação cartesiana* o que mostra que este conceito, *coeficiente angular*, é limitado. A solução são os chamados *cosenos diretores* que produzem a *equação polar da reta*.

Considere uma *reta*  $r$  que seja paralela ao vetor  $\vec{u} = (a, b)$  e passe pelo ponto  $P = (p, q)$ . Ela é também paralela ao vetor unitário da direção do vetor  $\vec{u}$  que é o vetor

$$(\cos(\alpha), \sin(\alpha)) = (\cos(\alpha), \frac{b}{R}); \alpha + \beta = \frac{\pi}{2} \quad (501)$$

$\alpha$  é o ângulo entre o vetor  $\vec{u}$  e o eixo  $OX$ .

$$\vec{n} = \begin{cases} \cos(\alpha) = \frac{a}{|\vec{u}|}; \\ \sin(\alpha) = \frac{b}{|\vec{u}|}; \end{cases} \quad (502)$$

$$X = \vec{P} + t\vec{n}; \quad (503)$$

$$X = \vec{P} + t(\cos(\alpha), \sin(\alpha)); t \in \mathbf{R}; \quad (504)$$

<sup>21</sup>  $X$  é um *vetor posição* sobre a reta: um ponto genérico sobre a reta. Confira a figura (fig 68), página 166, as coordenadas de  $X$  são determinadas como um

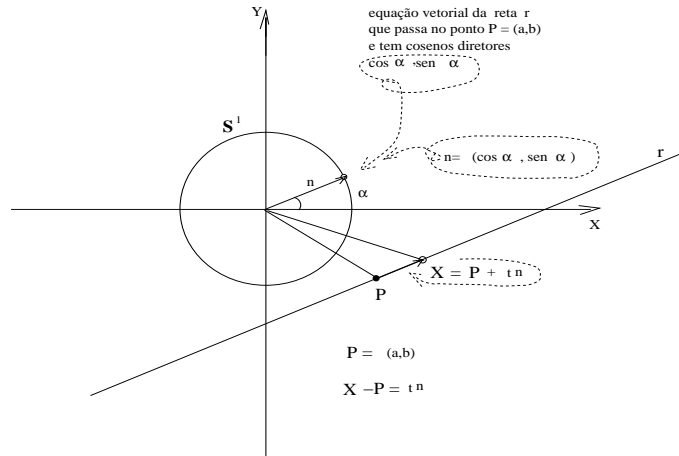


Figura 68: equação vetorial da reta: cossenos diretores

múltiplo dos *cossenos diretores* da reta somados ao vetor  $P$  por onde a reta passa, confira a equação (503), detalhada na equação (504).

A generalização desta equação para retas num espaço de dimensão  $n \geq 2$  é pura questão de aumentar o número de coordenadas e agora sim a expressão *cossenos diretores* aquirem o seu sentido mais completo.

Dado um vetor  $\vec{u} \in \mathbf{R}^n$ , a divisão pelo número  $|\vec{u}|$  produz um vetor da esfera unitária do  $\mathbf{R}^n$  que é designada com o símbolo,  $\mathbf{S}^{n-1}$  indicando que se trata duma variedade de dimensão  $n - 1$  do espaço  $\mathbf{R}^n$ , sendo uma generalização do *círculo trigonométrico*<sup>22</sup>  $\mathbf{S}^1$ , e um ponto qualquer  $X \in \mathbf{S}^{n-1}$  satisfaz à identidade fundamental da trigonometria

$$X = (x_1, \dots, x_n); \sum_{k=1}^n x_k^2 = 1; \quad (505)$$

<sup>21</sup>Aqui há um vício de notação contrariando um hábito estabelecido, o símbolo  $\vec{n}$  é reservado para o *vetor normal a uma variedade* sendo impróprio o uso deste símbolo aqui.

<sup>22</sup>Os índices usados como expoente caracterizam a dimensão destas variedades. O círculo trigonométrico é uma variedade de dimensão 1, uma curva.



e como temos apenas dois nomes para as funções trigonométricas fundamentais, fruto de nossa prisão tridimensional, a saída é chamá-las, todas as coordenadas, de “cos”, como algumas vezes fazemos no plano designando

$$\sin(\alpha) = \cos(\alpha_2); \alpha + \alpha_2 = \frac{\pi}{2}; \quad (506)$$

$$X = (\cos(\alpha_1), \cos(\alpha_2)) \in \mathbf{S}^1; \alpha_1 + \alpha_2 = \frac{\pi}{2}; \quad (507)$$

$$X = (\cos(\alpha_1), \cos(\alpha_2), \dots, \cos(\alpha_n)); \sum_{k=1}^n \cos(\alpha_k)^2 = 1; \quad (508)$$

O método consiste em

- sucessivamente cortando  $\mathbf{S}^{n-1}$  com planos perpendiculares;
- no primeiro selecionamos a direção  $OX_1$  e determinamos  $\alpha_1$  como o ângulo que a projeção de  $X$  fizer com  $OX$ ;
- se selecionarmos  $OX_1$  como a direção de  $X$  o processo está terminado!
- selecionamos a direção que for perpendicular a  $OX_1$  chamada  $OX_2$  e calculamos a projeção de  $X$  nesta direção encontrando  $\cos(\alpha_2)$  e assim sucessivamente até encontrarmos  $\cos(\alpha_n)$ .

Este processo está perfeitamente descrito no algoritmo de Gram–Schmidt para ortonormalização dum vetor cuja apresentação informal é a sucessão de passos acima.

---

- **equivalência** é uma generalização do conceito de igualdade. A igualdade é muito restrita, a equivalência é mais ampla: é possível ter uma infinidade de objetos equivalentes mas que de alguma forma nós tenhamos o interesse em *vê-los como* iguais. Por exemplo, os números racionais se agrupam em classes de equivalência, são equivalentes quando estiverem numa proporção:

$$\frac{p}{q} \equiv \frac{r}{s} \iff \frac{p}{q} = \frac{r}{s} \quad (509)$$

em que o segundo membro, na equivalência, é uma *proporção*, (*o produto dos extremos é igual ao produto dos meios*), e nós precisamos que eles representem o mesmo resultado.

A soma e o produto de números racionais respeitam as classes desta relação de equivalência (ou, são compatíveis com esta relação de equivalência) no sentido de que as duas ações,

- somar ou multiplicar dois números racionais e depois identificar a classe de equivalência do resultado;
- somar ou multiplicar os *representantes de classe* de dois números racionais e depois identificar a classe a que pertence este resultado.

produzem o mesmo resultado.

O que caracteriza uma *relação de equivalência* são as propriedades:

1.  $a \equiv a$ , *reflexividade*.
2.  $a \equiv b \Rightarrow b \equiv a$ , *simetria*.
3.  $a \equiv b$  e  $b \equiv c \Rightarrow a \equiv c$ , *transitividade*.

Estas três propriedades, *reflexividade*, *simetria*, *transitividade* são importantes por si próprias e aparecem em outras relações, parcialmente. Por exemplo, a *implicação lógica* não tem a propriedade 2 mas tem as propriedades 1,3. Consequentemente a *implicação lógica* não é uma *relação de equivalência* e nem seria desejável que o fosse.

A *congruência*, definida pelos restos na divisão por um número inteiro maior ou igual que 2, é uma *relação de equivalência*.

Dado um conjunto  $A$  e uma relação de equivalência  $R(x, y)$  definida entre os elementos de  $A$ , chamamos de *classes quociente* aos subconjuntos definidos em  $A$  formados por todos os elementos equivalentes entre si. Por exemplo, os números inteiros positivos que deixam o mesmo resto na divisão por um inteiro  $n \geq 2$  são as *classes quociente* na divisão por  $n$  em  $\mathbf{N}$ .

Nas escolas primárias, antigamente, as alunas aprendiam a *prova dos 9*, que é relativamente inútil mas já representava um começo de discussão de um assunto importante, a congruência. Ela se baseia na equivalência entre os números na divisão por 9 porque a soma e a multiplicação sendo compatíveis com as classes na divisão por 9, podemos verificar se a conta foi feita corretamente. Deixe-me introduzir a linguagem adequada, para simplificar o discurso. Dizemos que  $x \equiv y \pmod{9}$  se eles pertencerem a mesma classe dos restos na divisão por 9.

Assim  $1234343 \equiv 2 \pmod{9}$  e  $234244243 \equiv 1 \pmod{9}$  e o produto destes restos é 2.

Fazendo a multiplicação entres os dois números encontrei

$$289137741637349 \equiv 2 \pmod{9}$$

então sei que não errei na conta<sup>23</sup>, porque  $1 * 2 = 2$ , o produto dos restos. Então concluí que muito provavelmente não errei na conta. Como são iguais os resultados testados pela congruência, a *probabilidade* é grande de que a conta tenha sido feita corretamente.

A *prova dos nove* não é exata, mas a probabilidade de erro é muito pequena. Se ao calcular eu tivesse cometido o erro, encontrando  $289137741637439 = 289137741637349 + 90$  a *prova dos nove* não verificaria o meu erro, porque

$$289137741637439 \equiv 2 \pmod{9}$$

mas este “erro” seria, evidentemente muito raro. A probabilidade de erro na multiplicação é maior do que na adição, porque  $9 = 3 \times 3$ . Melhor seria a prova dos 11 que é um número primo, entretanto não temos meios simples para encontrar o resto na divisão por 11 como temos para com a divisão por 9. Este assunto supera muito a questão cultural importante de manter a *prova dos nove*

<sup>23</sup>Não errei, com certeza porque usei `calc` que uma linguagem de programação de domínio público que tenho instalada no computador.

na Escola, está ligada com encriptação de dados, mas neste caso o “módulo” na divisão tem que ser um número primo e bastante grande para que as *classes de equivalência* também possam ser grandes. Manter a *prova dos nove* na Escola significa preparar as estudantes para entender melhor a aritmética.

Um outro exemplo de relação de equivalência é *equivalência de triângulos*, confira *triângulos semelhantes*.

- **Eratóstenes** , o crivo de Eratóstenes é uma listagem em sequência dos números primos até um certo ponto, uma vez que o conjunto dos números primos é infinito. . .

#### crivo de Eratóstenes

1	2	3	5	7	11
13	17	19	23	31	37
41	43	47	53	59	61
67	71	73	79	83	89
97	101	103	107	109	113
127	131	137	139	149	151

Aqui o *crivo de Eratóstenes* construído com auxílio do **factor** até 151 em que você pode ver alguns dos fatos curiosos, os números primos gêmeos , que são os números primos seguidos pela soma de dois: (3, 5), (5, 7), (11, 13), (17, 19), (29, 31), (41, 43), (59, 61), (71, 73), (101, 103), (107, 109), (137, 139), (149, 151) . . .

Segundo a Wikipédia, existem cerca de mil números primos gêmeos abaixo de cem mil e oito mil abaixo de um milhão.

Os números primos foram e continuam sendo alvo de pesquisa, para aumentar a tabela dos primos conhecidos que hoje se encontra na casa dos bilhões havendo clusters de computadores que ficam dedicados nos finais de semana nesta busca. Por exemplo, a questão dos gêmeos merece atenção com problemas abertos envolvendo a sua distribuição.

Os números primos são importantes em criptografia por uma razão bem simples, que os restos na divisão por um número primo forma uma álgebra finita muito parecida com álgebra dos números inteiro positivos ou o corpo dos números racionais. Esta possibilidade algébrica permite um processo simples de encriptação que foi usado pelos alemães na segunda guerra com sua máquina de encriptação cujo segredo foi desvendado pelo matemático inglês Alan Turin. Na tabela operatória dos restos por um número primo, a multiplicação por qualquer número primo, provoca uma permutação dos elementos da linha representando assim um método algébrico que pode ser usado num programa de computação, para produzir permutações e portanto encriptar dados.

A tabela seguinte é da multiplicação dos restos na divisão por 7 em que cada linha é uma permutação da primeira obtida pela multiplicação pelo resto que se encontra na primeira coluna.

*	1	2	3	4	5	6
1	1	2	3	4	5	6
2	2	4	6	1	3	5
3	3	6	2	5	1	4
4	4	1	5	2	6	3
5	5	3	1	6	4	2
6	6	5	4	3	2	1

e a próxima é a tabela da adição dos restos na divisão por 7

+	0	1	2	3	4	5	6
0	0	1	2	3	4	5	6
1	1	2	3	4	5	6	0
2	2	3	4	5	6	0	1
3	3	4	5	6	0	1	2
4	4	5	6	0	1	2	3
5	5	6	0	1	2	3	4
6	6	0	1	2	3	4	5

Usando estas duas tabelas você pode resolver a equação apenas usando as propriedades fundamentais da aritmética dos números inteiros, como existência do neutro multiplicativo ou neutro aditivo, existência do inverso aditivo ou multiplicativo e a propriedade associativa,

$$4x + 3 = 2$$

$$4 * x + 3 = 2 \tag{510}$$

$$4 * x + 3 + 4 = 2 + 4 \Rightarrow 4 * x + 0 = 2 + 4 = 6 \tag{511}$$

$$4 * x = 6 \tag{512}$$

$$2 * (4 * x) = 2 * 6 = 5 \tag{513}$$

$$2 * (4 * x) = (2 * 4) * x = 5 \tag{514}$$

$$1 * x = x = 5 \text{ a solução da equação} \tag{515}$$

$$\text{testando } 4 * 5 + 3 = 6 + 3 = 2 \Leftarrow 4 * x + 3 = 2 \tag{516}$$

Considere um exercício identificar onde foi usada, cada uma das propriedades, para obter a linha seguinte no sistema de equações (eq. 510)-(eq. 516).

Assim se pode *encriptar* ou *desencriptar* qualquer informação baseada nesta tabela de multiplicação e adição dos restos na divisão por 7.

Aqui você a importância de descobrir números inteiros primos muito grandes: eles permitem uma álgebra mais generosa o que justifica a *generosidade* de certos financiadores que garantem computadores para matemáticos fazerem a pesquisa de números primos muito grandes. Não é exatamente porque eles sejam *mecenas* da ciência ...

As tabelas da adição e multiplicação módulo 7 foram construídas por um programa em C++ que vou voltar a usar para construir a maior tabela que for

possível colocar na página que é a tabela seguinte, da adição dos inteiros módulo 19, foi produzida por um programa em C++ que calculou os dados da tabela e construiu o layout da tabela em  $\text{\LaTeX}$ .

### A tabela da adição módulo 19

+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	0
2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	0	1
3	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	0	1	2
4	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	0	1	2	3
5	5	6	7	8	9	10	11	12	13	14	15	16	17	18	0	1	2	3	4
6	6	7	8	9	10	11	12	13	14	15	16	17	18	0	1	2	3	4	5
7	7	8	9	10	11	12	13	14	15	16	17	18	0	1	2	3	4	5	6
8	8	9	10	11	12	13	14	15	16	17	18	0	1	2	3	4	5	6	7
9	9	10	11	12	13	14	15	16	17	18	0	1	2	3	4	5	6	7	8
10	10	11	12	13	14	15	16	17	18	0	1	2	3	4	5	6	7	8	9
11	11	12	13	14	15	16	17	18	0	1	2	3	4	5	6	7	8	9	10
12	12	13	14	15	16	17	18	0	1	2	3	4	5	6	7	8	9	10	11
13	13	14	15	16	17	18	0	1	2	3	4	5	6	7	8	9	10	11	12
14	14	15	16	17	18	0	1	2	3	4	5	6	7	8	9	10	11	12	13
15	15	16	17	18	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
16	16	17	18	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
17	17	18	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
18	18	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17

### A tabela do produto módulo 19

*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
2	2	4	6	8	10	12	14	16	18	1	3	5	7	9	11	13	15	17
3	3	6	9	12	15	18	2	5	8	11	14	17	1	4	7	10	13	16
4	4	8	12	16	1	5	9	13	17	2	6	10	14	18	3	7	11	15
5	5	10	15	1	6	11	16	2	7	12	17	3	8	13	18	4	9	14
6	6	12	18	5	11	17	4	10	16	3	9	15	2	8	14	1	7	13
7	7	14	2	9	16	4	11	18	6	13	1	8	15	3	10	17	5	12
8	8	16	5	13	2	10	18	7	15	4	12	1	9	17	6	14	3	11
9	9	18	8	17	7	16	6	15	5	14	4	13	3	12	2	11	1	10
10	10	1	11	2	12	3	13	4	14	5	15	6	16	7	17	8	18	9
11	11	3	14	6	17	9	1	12	4	15	7	18	10	2	13	5	16	8
12	12	5	17	10	3	15	8	1	13	6	18	11	4	16	9	2	14	7
13	13	7	1	14	8	2	15	9	3	16	10	4	17	11	5	18	12	6
14	14	9	4	18	13	8	3	17	12	7	2	16	11	6	1	15	10	5
15	15	11	7	3	18	14	10	6	2	17	13	9	5	1	16	12	8	4
16	16	13	10	7	4	1	17	14	11	8	5	2	18	15	12	9	6	3
17	17	15	13	11	9	7	5	3	1	18	16	14	12	10	8	6	4	2
18	18	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1

Um programa como este pode ser usado para encriptar dados com valores grande para o número primo  $p$ , mas 19 foi o maior número primo que eu pude usar para que ainda fosse possível apresentar, de maneira decente, estas tabelas, numa página de texto. Estas são as tabelas da adição e do produto módulo 19. Testei a tabela para o número primo 23 mas não foi possível apresentá-la na página.

Vale a pena estudar o assunto!

- **esfera**. Confira *equação da esfera*.

- **espaço vetorial topológico** É uma generalização do conceito de *espaço vetorial* que se aplica naturalmente aos espaços de funções. O exemplo comum de espaço vetorial é o  $\mathbf{R}^n$  estudado nos cursos de Álgebra Linear, seus elementos são as *énuplas* de números reais  $P = (x_1, \dots, x_n)$ . Mas  $P$  pode ser vista como uma função definida no conjunto  $\{1, \dots, n\}$  com valores em  $\mathbf{R}$ , quer dizer que  $\mathbf{R}^n$  é o *espaço das funções definidas no conjunto  $\{1, \dots, n\}$*  portanto os

espaços vetoriais da Álgebra Linear são espaços de funções. O que torna distintos os diferentes espaços de funções é a topologia que estiver definida neles. A Álgebra Linear usa como primeiro exemplo de topologia para espaço vetorial aquela que é induzida pela norma euclidiana. Há possibilidade de definir diversas normas para  $\mathbf{R}^n$  mas se pode provar que todas são equivalentes ou seja, na prática os espaços vetoriais topológicos de dimensão finita são todos idênticos. Saindo da dimensão finita começam a acontecer eventos interessantes, uma multitude de topologias. Algumas destas topologias pode ser deduzida de alguma norma, mais geral um pouco, algumas podem ser induzidas por uma distância (ou métrica) e finalmente, o caso mais geral é quando se pode encontrar uma topologia que seja compatível com a estrutura de espaço vetorial fazendo contínuas as duas operações, adição e multiplicação por um escalar, neste caso se tem um *espaço vetorial topológico* que é o caso destes outros já citados. É uma generalização do conceito. As generalizações criam um ambiente em que, com menos propriedades, se consegue abranger mais objetos ou resolver problemas de outra que de outra forma seria muito difícil. Os *espaços vetoriais topológicos* surgem de forma natural no estudo das equações diferenciais parciais quando definições adequadas de topologias permitem criar condições para a construção de soluções aproximadas.

---

- **existência, teorema de**

Os *teoremas de existência* são instrumentos poderosos da Matemática e ao mesmo tempo muito disputados. Há *escolas de matemáticos* que não aceitam os teoremas de existência considerando-os algoritmicamente impossíveis, isto é, representam afirmações para as quais não seja possível construir um algoritmo com tempo de execução finito. Mas sem eles a Matemática iria se reduzir muito, e o raciocínio de algoritmicidade é fraco porque a Matemática não pretende ser algorítmica, esta é uma pretensão (bem sucedida) da *Matemática Computacional* que contém a *Computação Algébrica*.

A importância dos teoremas de existência consiste em provar, com alguma lógica, e aqui temos que considerar as limitações impostas pelo teorema de Gödel, a existência de um objeto matemático criando condições, então sim, para que produzem algoritmos que devem chegar a pelo menos uma aproximação do objeto.

Alguns teoremas de existência são enunciados em *equações diferenciais* garantindo a existência das soluções para uma certa classe de equações, o *teorema da função implícita* é um teorema de existência, o axioma da escolha teorema do valor médio da derivada, teorema de Rolle, são exemplos bem conhecidos de teoremas de existência.

---

- **fatoração** é a busca dos fatores primos dum inteiro positivo. Vale também para inteiros negativos, mas é um complicação inútil. Entre os inteiros positivos existe uma classe de números chamados primos. Confira *primo*, número.

Dado um inteiro qualquer ele pode se decompor de forma única como um

produto de números primos. Exemplo

$$12 = 2 * 2 * 3; 3600 = 2 * 2 * 2 * 2 * 3 * 3 * 5 * 5; \quad (517)$$

em que usei o símbolo “\*” para representar multiplicação e obtive os resultados digitando

```
factor 3600
3600: 2 2 2 2 3 3 5 5
```

num terminal da minha caixinha Linux, porque **factor** é um utilitário que vem junto com os programas da distribuição Linux que uso. Mas eu posso editar os resultados escrevendo

$$12 = 2^2 * 3; 3600 = 2^4 * 3^2 * 5^2; \quad (518)$$

e assim eu chego na notação habitual da fatoração: dado um número inteiro positivo qualquer,  $P$ , existem números primos  $p_1, \dots, p_k$  e números naturais  $\alpha_1, \dots, \alpha_k$  tal que

$$P = p_1^{\alpha_1} * \dots * p_k^{\alpha_k} \quad (519)$$

no caso de

$$3600; p_1 = 2, \alpha_1 = 4; p_2 = 3, \alpha_2 = 2; p_3 = 5, \alpha_3 = 2;$$

Claro,  $P$  pode ser um número primo, então  $p_1 = P, \alpha_1 = 1$  e por aí vai uma demonstração por indução finita que você pode encontrar em qualquer livro bom de Álgebra, do *Jacy Monteiro*, por exemplo.

Existe um algoritmo para determinação dos fatores primos que passa pela listagem dos primos e um teste de divisão de  $P$ :

	2	3	5	7	11	13	...
3600	2						
1800	2						
900	2						
450	2						
225	3						
75	3						
25	5						
5	5						
1							

Se baseia nas divisões sucessivas dos primos que aparecem no *crivo de Eratóstenes* que uma listagem dos números primos até um certo ponto... porque o conjunto dos números primos é infinito, outro teorema que você pode demonstrar por indução ou encontrar a demonstração num bom livro de Álgebra. Não é difícil construir a demonstração, é um bom exercício de *indução finita*.

Na tabela acima, em que aparece o algoritmo aplicado a 3600 você pode ver o *crivo de Eratóstenes* desenvolvido até o número primo 13 e de onde fui selecionando os fatores para testar e construir a fatoração de 3600.

Use **factor** na sua caixinha Linux e se divirta vendo a fatoração de grandes números:

```

factor 123343123
123343123: 4561 27043
factor 12334312334400
12334312334400: 2 2 2 2 2 2 3 5 5 7 367092629
factor 12334312334400334
12334312334400334: 2 89 113 51721 11856311
factor 11856311
11856311: 11856311

```

A Álgebra generaliza a definição de número primo para a classe dos ideais dum anel, definindo os ideais primos. Nesta forma de ver os números primos da Aritmética definem os ideais primos do anel dos inteiros. Dado um número inteiro  $p$ ,  $p\mathbf{Z}$ , o conjunto dos múltiplos de  $p$  é um ideal de  $\mathbf{Z}$  e se  $p$  for primo,  $p\mathbf{Z}$  é um ideal primo, e além disto maximal da classe de todos os ideais de  $\mathbf{Z}$ . Os restos na divisão por  $p$ , quando  $p$  for primo, representam as classes quocientes de  $\mathbf{Z}$  pelo ideal  $p\mathbf{Z}$  sendo então um corpo. Confira o crivo de Erastótenes e número primo.

---

- **Fermat, o último teorema de** O advogado, e matemático Pierre de Fermat, afirmou, nas margens de um seus livros que  $a^n + b^n = c^n$  não poderia ter soluções inteiras quando  $n > 2$  e que ele tinha uma simples demonstração para esta afirmação que infelizmente não cabia na margem do livro, por volta de 1632. Um caso bem conhecido, em que esta afirmação é verdadeira, é o Teorema de Pitágoras expresso para alguns inteiros quando  $n = 2$  o que define os chamados *números pitagóricos*, por exemplo, 3, 4, 5;  $3^2 + 4^2 = 5^2$ . Em 1993 Wiles anunciou a demonstração do último teorema de Fermat, mas um erro foi descoberto em sua demonstração que foi finalmente corrigida em 1995 num artigo publicado por ele e Taylor.

---

- **filtro crescente**, em topologia, se  $X$  for um espaço localmente compacto um *filtro crescente* é uma família não vazia  $\Phi$  de funções definidas em  $X$  e tomando valores na reta estendida,  $(-\infty, \infty]$ , com a propriedade

$$f, g \in \Phi \exists h \in \Phi; \sup(f, g) < h; \quad (520)$$

propriedade esta que pode ser iterada para garantir um majorante para uma enupla de elementos da família  $\Phi$ . Esta família se chama *filtro crescente* sobre o conjunto das funções reais definidas em  $X$  e definimos o *limite ao longo deste filtro* para um *operador*  $S$  definido em  $\Phi \subset \overline{\mathbf{R}}^X$  como

**Definição 10 (generalização do limite)** *limite ao longo do filtro*  $\Phi$

Se  $S : \Phi \rightarrow \mathbf{R} \lim_{\Phi} S(f) = c$ , se existir, significa que

$$(\forall \epsilon > 0) (\exists h \in \Phi) (f \in \Phi \ \& \ f \geq h) \implies |S(f) - c| < \epsilon$$

Vou apresentar três exemplos para atestar a efetividade desta *generalização do limite*. O exemplo mais comum, embora geralmente mal entendido no momento em que surge nas primeiras versões do Cálculo, é a integral no sentido



de Riemann que é usualmente apresentada, erroneamente, como um *limite*, um erro compreensível uma vez que não haveria maturidade da audiência para usar a linguagem *correta* que seria o *limite ao longo do filtro das funções escada*, definidas a seguir no primeiro exemplo. Depois vou generalizar a integral de Riemann definindo a integração no sentido de Lebesgue mostrando que ela é apenas uma seleção de um novo filtro com a adição de uma generalização da medida geométrica que se encontra à base da integral de Riemann. O terceiro exemplo será bem simples, é o próprio limite de sucessões ao qual vou aplicar *limite ao longo de um filtro* mostrando que a generalização se aplica ao *caso primitivo*.

**Exemplo 1 (integral de Riemann)** *A integral de Riemann*

*Considere uma função real  $\mathbf{f}$  positiva e um intervalo  $I$  fechado e limitado.*

*Associadas a cada partição  $(\Pi_n(I))_{n \in \mathbf{N}}$  existe uma infinidade de funções*

$$f = \sum_{k=1}^n \mathbf{c}_k \chi_{I_k}; \quad (521)$$

*resultado da combinação linear das funções características dos subintervalos de  $I$  com os coeficientes  $\mathbf{c}_k$ .*

*Como nos interessa associar as funções  $f$  à função  $\mathbf{f}$  consideramos*

$$\mathbf{c}_k = \mathbf{f}(x_k) \quad (522)$$

*para escolhas arbitrárias de  $x_k \in I_k$  que formam a classe  $\Phi$  deste exemplo, mas logo em seguida você verá que somos obrigados a fazer uma associação indireta com  $\mathbf{f}$  para obter a propriedade definidora na equação (520).*

*Observe que  $\mathbf{f}$  é uma constante dada, é a função cuja integral, no sentido de Riemann, deve ser calculada, se existir.*

*Uma função do tipo definido na equação (521) é usualmente chamada de função escada, embora seja uma escada que pode subir e descer ao sabor dos coeficientes tirados de  $\mathbf{f}$  e o seu conjunto forma o filtro  $\Phi$  deste exemplo.*

*Antes de caracterizarmos a família  $\Phi$  como um filtro teríamos que provar que a propriedade expressa na equação (520) se verifica. A existência de  $h > \sup(f, g); f, g \in \Phi$  que majore duas funções escadas é obtida considerando o refinamento comum às duas partições definidoras de  $f, g$  substituindo  $\mathbf{c}_k$  pelo supremo de  $\mathbf{f}$  no novo subintervalo que corresponder à combinação linear depois de calculado o refinamento das partições. Observe que estamos sempre tratando com um número  $n \in \mathbf{N}$  de subintervalos, apenas este supremo pode ser  $\infty$  o que pode forçar a busca de técnicas apropriadas para tratar um subintervalo em que  $\mathbf{f}$  tenha pontos críticos desta natureza, lembre-se que a função  $\log$  é integrável à Riemann num intervalo contendo o zero, ela tem uma primitiva num tal intervalo, por exemplo  $F(x) = x \log|x| - x$ .*

*O operador  $S$  definido na classe  $\Phi$  é obtido com as somas de Riemann calculadas a partir de cada função escada  $f \in \Phi$ ,*

$$S(f) = \sum_{k=1}^n \mathbf{c}_k m(I_k); \quad (523)$$

$$x_k \in I_k \text{ e } m(I_k) \text{ a medida do subintervalo } I_k; \quad (524)$$

O limite de  $S$  ao longo do filtro  $\Phi$ , se pudermos provar que existe, é a

$$S(f) = \int_I \mathbf{f}(x) dx \quad (525)$$

e a forma de provar a existência deste limite consiste em construirmos dois operadores,  $\overline{S(f)} \leq S(f)$  em que  $\mathbf{c}_k$  é tomado, respectivamente, como o supremo ou ínfimo de  $\mathbf{f}$  em cada subintervalo  $I_k$  e provar que todas as cadéias de  $\Phi$  são equivalentes porque definem o mesmo número real, ou ainda que para qualquer cadéia em  $\Phi$  a diferença

$$\overline{S(f)} - \underline{S(f)}$$

é arbitrariamente pequena.

Observe que neste momento estamos fazendo novamente associações indiretas dos operadores  $\overline{S(f)}$  e  $\underline{S(f)}$  com  $\mathbf{f}$  usando os operadores  $\sup$  e  $\inf$ .

Ou seja, uma função é integrável à Riemann sobre um intervalo limitado  $I$  se houver o limite para o operador  $S$  relativamente ao filtro  $\Phi$  e neste caso este limite é denotado com a expressão da equação (525), e este método de integração é chamado de integração à Riemann.

Você deve ter observado que definimos apenas a integral de *funções positivas* e portanto vem a pergunta: e a integral das outras funções. Como no Cálculo isto é feito trocando o sinal se  $\mathbf{f}$  for negativa e no caso geral considerando a identidade

$$\mathbf{f} = \mathbf{f}^+ - \mathbf{f}^- \quad (526)$$

em que  $\mathbf{f}^+$ ,  $\mathbf{f}^-$  são, respectivamente a *parte positiva* ou a *parte negativa* de  $\mathbf{f}$ , com a linearidade da integral aplicada à equação (526).

No próximo exemplo vou *alterar um pouco* o filtro  $\Phi$  tendo como resultado a integral no sentido de Lebesgue, na verdade a alteração não será pequena porque será considerada uma classe muito mais ampla de funções cujas integrais existem, e também outra alteração, *conceito de integrabilidade* muda. Se estabeleceu a ruptura entre a compreensão de integral que tinham os matemáticos até à época de Lebesgue, mas parece que o próprio Lebesgue não entendeu o salto que ele mesmo produziu e que somente ficou claro com o passar dos tempos em que a integral se afastou de sua interpretação geométrica como área (ou volume, hipervolume) para se tornar num *operador linear* definido num espaço de funções deixando a concepção anterior como um caso particular em que a integral podia ser interpretada como uma medida geométrica, *área, volume, hipervolume*.

Na verdade Lebesgue escreveu uma variante do que poderia ter sido a continuação da integral de Riemann–Stieltjes que é a real predecessora da integral como operador linear.

De forma idêntica, como limite ao longo de um filtro, a alteração se dá no filtro, que podemos identificar como uma classe elementar de funções integráveis. Para a integração à Riemann seriam as funções características de retângulos, e agora serão as *funções simples* que vou definir no próximo exemplo. Como no caso da construção da integral à Riemann vamos definir a integral de Lebesgue de uma função positiva  $f$  e posteriormente expandir a qualquer função com a identidade expressa na equação (526).

**Exemplo 2 (integral de Lebesgue)** *Função integrável no sentido de Lebesgue*

*As modificações introduzidas por Lebesgue foram:*

1. Medida *O primeiro aspecto da generalização é considerar  $I$  um conjunto mesurável, que pode ser um intervalo da reta ou um conjunto arbitrário. Aqui mesurável se refere à existência de uma família de subconjuntos de  $I$  que formam uma álgebra de conjuntos relativamente as operações de união, interseção ou diferença de conjuntos e neste ponto há diversas possibilidades tornando a teoria de Lebesgue muito rica, estou falando da teoria das medidas. Por exemplo,  $I$  pode ser uma região do plano ou do espaço 3D ou de qualquer outra dimensão e os conjuntos mesuráveis uma álgebra de conjuntos de  $I$ . A teoria das medidas define uma função que se chama “medida”, e podem ser múltiplas, criando uma grande variedade de integrais.*
2. funções simples *a família  $\Phi$  é formada das funções simples positivas, semelhantes à construção do exemplo 1 a diferença é que a partição de  $I$  é tomada entre os conjuntos mesuráveis e não subintervalos, outro aspecto da generalização. A definição de função simples é uma função mesurável que assume exatamente  $n$  valores, aqui estou usando o número  $n$  para criar a compatibilidade entre os dois exemplos de que vou fazer uso em seguida.*

*Estas dois novos aspectos são fundamentais e alteraram profundamente o curso da análise matemática, embora elas já viessem sendo fermentadas antes de Lebesgue, foi ele que, em sua tese de doutorado, montou a teoria das medidas.*

*Com estas duas alterações praticamente se aproveita o restante da construção do exemplo 1:*

1. *temos uma família que é um filtro crescente e a propriedade definidora expressa na equação (520) se verifica de forma similar porque a álgebra de conjuntos é fechada para interseção e assim podemos calcular um refinamento das partições que definem os dois elementos  $f, g \in \Phi$  para obter o majorante  $h \in \Phi$  exatamente como fizemos com os intervalos, apenas que agora estamos operando com os elementos da álgebra de conjuntos e uma família de subintervalos de um intervalo  $I$  também é uma álgebra de conjuntos, estamos numa generalização;*

2. a metodologia para garantir a existência do limite é semelhante a que usamos no exemplo 1, trabalhamos com os dois operadores  $\overline{S}, \underline{S}$  e nos dois casos se trabalha com o cone das funções positivas e assim calculamos a integral de uma função positiva, se o limite existir.

A integral de Lebesgue de função positiva  $\mathbf{f}$  é o limite, se existir, da integral  $S$  das funções simples alteradas com os coeficientes  $f(x_k); x_k \in I_k; \cup_k I_k = I$  sobre o filtro das funções simples.

Como era o meu objetivo, mostrei que a integração à Riemann ou à Lebesgue são construídas de forma semelhante como limite ao longo de um filtro, o que as distingue é o filtro crescente utilizado que por sua vez, na integral no sentido de Lebesgue agora se estende para uma variedade imensa de espaços  $X$  sendo por isto algumas vezes chamada de integração abstrata.

No terceiro exemplo vou mostrar que o *limite corriqueiro* do Cálculo é também uma aplicação do *limite ao longo de um filtro crescente* e portanto o método é uma generalização do limite. A metodologia que vou usar consiste da análise de um exemplo que depois vou mostrar que se aplica em qualquer caso.

### Exemplo 3 (o limite do Cálculo) Limite de sucessões

Suponha que a sucessão  $\mathbf{s} = (x_k)_{k \in \mathbf{N}}$  tenha limite:

$$\lim_k x_k = a; \quad (527)$$

Como  $\mathbf{s}$  é convergente, então existe uma sucessão  $\mathbf{t}$ , crescente, que é equivalente a  $\mathbf{s}$ , elas tem o mesmo limite. Então vou considerar  $\mathbf{s} := \mathbf{t}$ . O que estou fazendo aqui é semelhante à suposição feita nos dois exemplos de integração considerando  $\mathbf{f}$  positiva. O caso das sucessões decrescentes é construído com uma simples simetria. Como há sucessões convergentes que não são monótonas precisamos aqui de uma técnica que vou apenas citar: é possível construir a partir de uma sucessão não monótona uma sucessão crescente (ou decrescente) que lhe é equivalente, usando  $\liminf$  ou  $\limsup$ .

Vou continuar sob a suposição de que  $\mathbf{s}$  é crescente. e definir uma sucessão de sucessões

$$(X_k)_{k \in \mathbf{N}}; X_k = (x_0, \dots, x_k, x_k, \dots, x_k, \dots) \quad (528)$$

uma sucessão de sucessões construídas a partir da primitiva sucessão  $\mathbf{s}$  tornando constante os termos a partir do termo de ordem  $k$ . Esta é a família que vou usar como dos elementos da básico dos filtros  $\Phi$ .

Agora vem a condição definidora da equação (520). Como dadas duas sucessões,  $\hat{f}, \hat{g}$ , são crescentes, então  $\hat{h}$  é a que tiver mais elementos. Desta forma a família de sucessões a suporte compacto assim construída é um filtro crescente formado de sucessões convergentes cujos limites se aproximam arbitrariamente de  $a = \lim_k x_k$  satisfazendo a condição de limite do filtro crescente.

Não há dúvida de que o terceiro exemplo tem apenas o objetivo de mostrar que o cálculo do limite usando uma família filtrante também se aplica às sucessões mas é uma forma complicada de calcular um limite simples mas que serve para provar que o processo de convergência ao longo de um filtro generaliza o caso simples.

- **fórmula de Euler** é uma expressão envolvendo os números  $e, i, 0, 1, \pi$  numa única igualdade:

$$e^{i\pi} = -1 \Rightarrow e^{i\pi} + 1 = 0$$

Aceite como uma notação que um número complexo unitário seja chamado de  $e^{i\alpha}$  em que  $\alpha$  é o seu argumento.

$$e^{i\alpha} = \cos(\alpha) + i \sin(\alpha); \alpha \in \mathbf{R}; \quad (529)$$

Podemos demonstrar, usando um raciocínio geométrico e apenas com as técnicas da Geometria Analítica, que

$$e^{i\alpha} e^{i\beta} = e^{i(\alpha+\beta)} \quad (530)$$

Para começar é muito provável, porque como é o produto de dois números de módulo 1 o resultado teria de ser um número de módulo 1. Isto tem uma tradução interessante: o produto de dois números que estejam no círculo unitário é novamente um elemento do círculo unitário, e o conteúdo da *fórmula de Euler* diz que você pode obter um, somando o argumento do outro.

Se isto puder ser provado, e pode, então a *fórmula de Euler* é mais do que uma notação ou uma fórmula, é uma autêntica expressão de função exponencial.

Selecione  $\alpha = 0$  e o resultado é o número complexo 1 que também é o mesmo número real 1, um elemento do círculo unitário.

- **fórmula trigonométrica**

$$\cos(\alpha)^2 + \sin(\alpha)^2 = 1; \quad (531)$$

$$\cos(\alpha + \beta) = \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta); \quad (532)$$

$$\sin(\alpha + \beta) = \cos(\alpha) \sin(\beta) + \sin(\alpha) \cos(\beta); \quad (533)$$

$$\cos(\alpha - \beta) = \cos(\alpha) \cos(\beta) + \sin(\alpha) \sin(\beta); \quad (534)$$

$$\sin(\alpha - \beta) = \sin(\alpha) \cos(\beta) - \cos(\alpha) \sin(\beta); \quad (535)$$

$$\tan(\alpha + \beta) = \frac{\tan(\beta) + \tan(\alpha)}{1 - \tan(\alpha) \tan(\beta)} \quad (536)$$

$$\tan(\alpha - \beta) = \frac{\tan(\alpha) - \tan(\beta)}{1 + \tan(\alpha) \tan(\beta)} \quad (537)$$

$$\sin(-x) = -\sin(x); \cos(-x) = \cos(x); \quad (538)$$

$$\sin(x + 2k\pi) = \sin(x); \cos(x + 2k\pi) = \cos(x); \quad (539)$$

$$\sin(x) = \cos(x - \frac{\pi}{2}); \sin(x - \pi) = -\sin(x); \cos(x - \pi) = -\cos(x); \quad (540)$$

$$\frac{d \sin(x)}{dx} = \cos(x); \frac{d^2 \sin(x)}{dx^2} = -\sin(x); \frac{d^3 \sin(x)}{dx^3} = -\cos(x); \frac{d^4 \sin(x)}{dx^4} = \sin(x) \quad (541)$$

$$\frac{d \cos(x)}{dx} = -\sin(x); \frac{d^2 \cos(x)}{dx^2} = -\cos(x); \frac{d^3 \cos(x)}{dx^3} = \sin(x); \frac{d^4 \cos(x)}{dx^4} = \cos(x) \quad (542)$$

estas fórmulas podem ser deduzidas com auxílio da fórmula de Euler

$$e^{i\alpha} = \cos(\alpha) + i \sin(\alpha); \alpha \in \mathbf{R}; \quad (543)$$

que tem as propriedades<sup>24</sup> da *exponencial*

A (eq. 531) é a expressão do módulo do raio do círculo trigonométrico, porque a fórmula de Euler descreve o número complexo unitário que o ângulo  $\alpha$  determina no *círculo trigonométrico*  $\mathbf{S}^1$ , ela afirma que  $\|e^{i\alpha}\| = 1$ .

Para demonstrar as equações (eq. 532) e (eq. 533) vou mostrar que

1.  $e^{i\alpha}e^{i\beta} = e^{i\gamma} = (\cos(\gamma) + i \sin(\gamma)) \in \mathbf{S}^1$ , o produto de dois números complexos de módulo 1 é um número complexo de módulo 1,
2. e preciso provar é que  $\gamma = \alpha + \beta$ .

- inicialmente observe que dados dois arcos  $\alpha, \beta$  ficam determinados dois números complexos  $e^{i\alpha}, e^{i\beta} \in \mathbf{S}^1$ , dois pontos do círculo trigonométrico, dois números complexos de módulo 1. O produto de dois números de módulo 1 também tem módulo 1, é um elemento do círculo trigonométrico pela propriedade do *módulo do produto*:

$$1 = \|zw\| = \|z\|\|w\| \Leftrightarrow \|z\| = \|w\| = 1 \quad (544)$$

a demonstração se faz calculando o módulo de  $e^{i\alpha}e^{i\beta}$ , ao fatorar a expressão aplique a equação (eq. 531). Provei que  $e^{i\alpha}e^{i\beta} = e^{i\gamma} \in \mathbf{S}^1$  e quero provar que  $\gamma = \alpha + \beta$ .

- O produto dos números complexos  $e^{i\alpha}, e^{i\beta}$  é o número complexo  $e^{i\gamma}$  e vou mostrar que  $\gamma = \alpha + \beta$ .

$$(\cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta) + i(\cos(\alpha) \sin(\beta) + \sin(\alpha) \cos(\beta))) = \quad (545)$$

$$a + ib; \quad (546)$$

$$a = \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta) = \quad (547)$$

$$= \cos(\alpha) \cos(\beta) + \sin(\alpha) \sin(-\beta) = \quad (548)$$

$$= \cos(\alpha) \cos(-\beta) + \sin(\alpha) \sin(-\beta); \quad (549)$$

$$a = \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta) = \langle (\cos(\alpha), \sin(\alpha)), (\cos(-\beta), \sin(-\beta)) \rangle \quad (550)$$

$$a = \cos(\alpha) \cos(\beta) - \sin(\alpha) \sin(\beta) = \cos(\alpha + \beta); \quad (551)$$

$$b = \cos(\alpha) \sin(\beta) + \sin(\alpha) \cos(\beta) = \langle (\cos(\alpha), \sin(\alpha)), (\sin(\beta), \cos(\beta)) \rangle; \quad (552)$$

$$b = \langle (\cos(\alpha), \sin(\alpha)), (\cos(\delta), \sin(\delta)) \rangle; \quad (553)$$

$$b = \cos(\alpha - \delta) = \cos(\alpha + \beta - \frac{\pi}{2}) = \sin(\alpha + \beta); \quad (554)$$

$$a + bi = e^{i\gamma} = \cos(\alpha + \beta) + i \sin(\alpha + \beta) \quad (555)$$

– na equação (eq. 549) porque  $\cos$  é uma função par e  $\sin$  é uma função ímpar.

---

<sup>24</sup>Por que é realmente uma exponencial, a base é o número  $e \approx 2.71828182845904523536$ , um número irracional, mais exatamente, *transcendente*.

- na equação (eq. 551) estou usando a definição geométrica do *produto escalar* que também está sendo usada na equação (eq. 554)
- $\delta, \beta$  na equação (eq. 553) são simétricos em relação ao arco  $\frac{\pi}{4}$ , em outras palavras em relação a primeira bissetriz dos eixos. Logo

$$\beta - \frac{\pi}{4} = \frac{\pi}{4} - \delta \Rightarrow \delta = \frac{\pi}{2} - \beta$$

- Para demonstrar as equações (eq. 534) e (eq. 535) vou considerar o produto  $e^{i\alpha}e^{-i\beta}$  que já sabemos que resulta num novo número complexo de módulo 1 então  $e^{i\alpha}e^{-i\beta} = e^{i\gamma}$  e temos que mostrar que  $\gamma = \alpha - \beta$ . Basta analisar um caso particular  $e^{i\beta}e^{-i\beta} = 1 = e^{0i} = e^0$  que implica que  $e^{-i\beta} = \frac{1}{e^{i\beta}}$ . Os cálculos seguintes encerram a demonstração:

$$e^{i\beta}e^{-i\beta} = 1; e^{-i\beta} = \frac{1}{e^{i\beta}} \quad (556)$$

$$e^{i(\alpha+\beta)}e^{-i\beta} = (e^{i\alpha}e^{i\beta})e^{-i\beta} = e^{i\alpha}(e^{i\beta}e^{-i\beta}) = e^{i\alpha}; \quad (557)$$

$$e^{i(\alpha+\beta)}e^{-i\beta} = e^{i\alpha}e^{-i\beta}e^{i\beta} = e^{i\gamma}e^{i\beta} = e^{i\alpha}; \quad (558)$$

$$\gamma + \beta = \alpha \Rightarrow \gamma = \alpha - \beta; \quad (559)$$

1. Na equação (eq. 557) usei a equação (eq. 532) e a propriedade associativa do produto de números complexos.
  2. Na equação (eq. 558) usei as propriedades associativa e a comutativa do produto de números complexos.
  3. Na equação (eq. 559) usei a equação (eq. 532).
- Na demonstração da equação (eq. 536) use a definição de tan e as equações (eq. 532) e (eq. 533) e elimine os fatores comuns.

As demonstrações das equações (eq. 532) e (eq. 533) têm consequências mais profundas, elas mostram que o grupo aditivo dos números reais esta em correspondência com o grupo multiplicativo dos números complexos unitários,  $\mathbf{S}^1$ , e que esta correspondência é um *morfismo de grupos*. Isto é, para “localizar o produto de dois números complexos, some seus argumentos” o que leva à construção do *logaritmo complexo*.

A figura (fig 69), página 182, sugere um método para obter a *regra do coseno* usando o Cálculo da distância  $h = d(e^{i\alpha}, e^{-i\beta})$ .

$$h^2 = (\cos(\alpha) - \cos(-\beta))^2 + (\sin(\alpha) - \sin(-\beta))^2 = \quad (560)$$

$$1 + 1 - 2\cos(\alpha)\cos(-\beta) + 2\sin(\alpha)\sin(-\beta) = \quad (561)$$

$$\|u\|^2 + \|v\|^2 - 2\langle e^{i\alpha}, e^{-i\beta} \rangle = \quad (562)$$

$$\|u\|^2 + \|v\|^2 - 2\|u\|\|v\|\cos(\alpha - \beta); \quad (563)$$

$$u = e^{i\alpha}; v = e^{-i\beta}; \quad (564)$$

Uma transformação da (fig 69) pode ser vista na figura (fig 70), página 183, mostrando como obter a lei coseno num triângulo qualquer. Os dois lados do

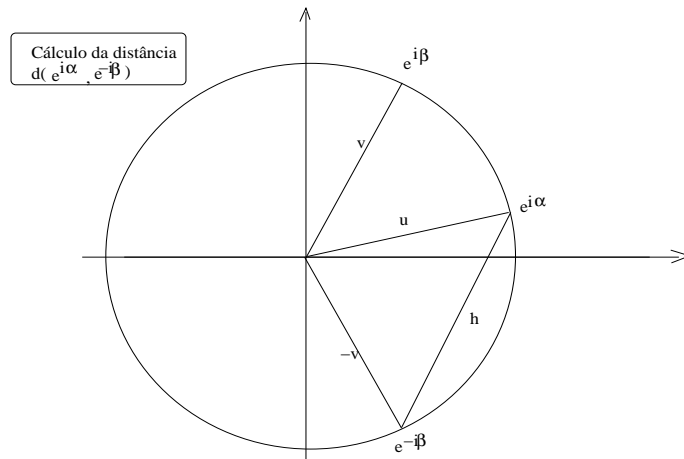


Figura 69: seno da soma de arcos

novo triângulo foram obtidos multiplicando  $u$  por  $\rho$ , para obter o segmento  $\overline{OP}$  e  $v$  por  $\sigma$ , para obter o segmento  $\overline{OQ}$  resultando na lei do cosseno

$$\|u\|^2 + \|v\|^2 - 2\|u\|\|v\|\cos(\alpha - \beta) \quad (565)$$

- **Fourier**, transformada de -. Ver transformada de Fourier.

- **fração** é a referência a número que significa uma parte dum inteiro, designada como *fração própria*<sup>25</sup>. É um exemplo de defeito do linguajar matemático oriundo de uma época em que as frações não eram consideradas números em todos os seus direitos.

Uma fração é um objeto da forma  $\frac{p}{q}$  em que  $p, q \in \mathbf{N}$ ;  $q \neq 0$  e que representa um número racional.

A história é uma arte muito difícil e um exemplo disto é o *triângulo de Pascal*, denominado em homenagem a Blaise Pascal que viveu no século 16, mas, aparentemente, os matemáticos chineses já conheciam este algoritmo dois mil anos antes de Pascal. Também parece que a geometria, dita dos gregos, era conhecida por outros povos anteriores aos helenos que, apenas teriam compilado de forma organizada aquilo que chegou até nós como *geometria euclidiana*.

Isto para afirmar, sem me sentir obrigado a grandes justificativas que, num certo momento da história humana, provavelmente na Idade Média, se começou a conceber que objetos como  $\frac{1}{3}$  *seriam números*. O uso do condicional tem sentido porque até recentemente as escolas ensinavam que  $\frac{5}{3}$  seria uma *fração imprópria* sugerindo que não seria bem uma fração! Os números negativos ainda

<sup>25</sup>em oposição às frações impróprias que são frações tão boas como as próprias apenas chamadas assim porque representam números racionais que não são inteiros mas que são em módulo maior do que 1.



**Lei do  
coseno**

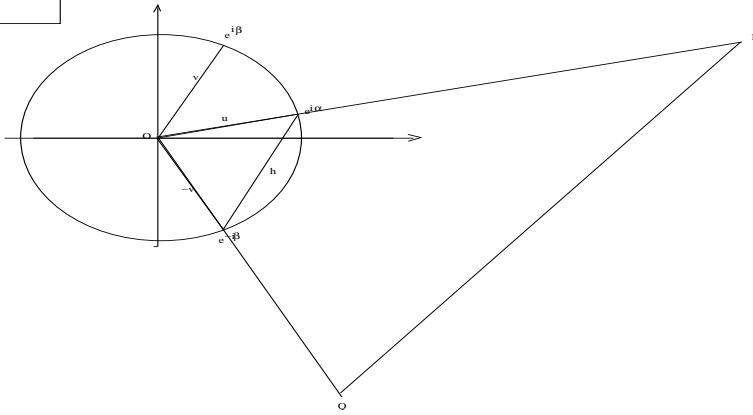


Figura 70: Lei do coseno

hoje são tratados *negativamente* e os números complexos nem sempre fazem parte do currículo do Ensino Médio ... para não mencionar os *quaternions* que, possivelmente, há gente que nem sabe que existem.

Então fazendo uma história romanceada, digamos que inventamos os objetos do tipo  $\frac{1}{n}$  quando  $n \in \mathbf{N}; n \neq 0$  para representar os inversos multiplicativos dos números naturais, com exceção do zero, e aos poucos construímos uma aritmética com estes novos objetos com as seguintes regras:

1. O símbolos  $\frac{p}{q}$  representam números sempre que  $q \neq 0$  em que os membros  $p, q \in \mathbf{N}$ . Temos o hábito de designar  $p$  como numerador e  $q$  como denominador. A razão destes nomes vem da ideia intuitiva da invenção das frações em que  $\frac{2}{3}$  significaria a quantidade 2 de uma coisa chamada “terço” donde o 2 é o “numerador” enquanto que 3 dá o nome, “denominador”.
2. O inverso aditivo do número  $\frac{p}{q}$  é o número  $\frac{-p}{q}$  que também pode ser escrito como  $-\frac{p}{q}$  e então entendemos que o sinal  $-$  é um modificador de tal modo que  $a + (-a) = 0$ .
3. Vale regra  $-(-a) = a$ ;
4. O zero Sempre que  $q \neq 0$ ,  $\frac{0}{q} = 0 \in \mathbf{N}$  e aqui guarde como observação para uso posterior, então existe uma quantidade imensa de representantes do zero e isto é um problema que preciso resolver!
5. Sempre podemos reduzir uma fração à sua expressão mais simples eliminando fatores comuns ao *numerador* e ao *denominador*. Assim

$$\frac{8}{80} = \frac{2^3}{2^4 5} = \frac{1}{2 \times 5} = \frac{1}{10} \quad (566)$$

ou seja, fatoramos numerador e denominador e eliminamos os fatores comuns para obter a forma irredutível duma fração. Esta operação cria um problema de que vou tratar ao final criticando todo o processo. Agora temos pelos menos dois objetos representando a mesma coisa, a fração simplificada e anterior que pode ser simplificada:

$$\frac{8}{80} = \frac{1}{10}$$

6. Esta regra de simplificação pode ser usada ao reverso para permitir a definição da soma de frações. Se quisermos somar

$$\frac{1}{p} + \frac{1}{q}$$

vamos acrescentar fatores comuns até obter denominadores iguais:

$$\frac{1}{p} + \frac{1}{q} = \frac{q}{qp} + \frac{p}{pq} \quad (567)$$

$$\frac{1}{p} + \frac{1}{q} = \frac{q}{pq} + \frac{p}{pq} = \frac{1}{pq}(q + p) \quad (568)$$

$$\frac{1}{p} + \frac{1}{q} = \frac{q+p}{pq} \quad (569)$$

porque temos  $q$  objetos do tipo  $\frac{1}{pq}$  e  $p$  objetos do tipo  $\frac{1}{pq}$ . Estou usando, a comutatividade da multiplicação de números naturais, na lista de operações acima junto a regra que nos permite eliminar ou incluir fatores comuns e finalmente, silenciosamente, estou usando *distributividade* da multiplicação em relação a adição que vou incluir como a próxima regra.

7. Vale a distributividade da multiplicação relativamente à soma.  
8. Podemos agora deduzir da regra anterior uma regra geral para soma de frações

$$\frac{m}{n} + \frac{p}{q} = \frac{qm}{nq} + \frac{np}{nq} \quad (570)$$

$$\frac{m}{n} + \frac{p}{q} = \frac{qm+np}{nq} \quad (571)$$

sendo a regra:

- multiplicamos os denominadores para formar o novo denominador;
- multiplicamos em cruz, os numeradores e denominadores e os somamos para formar o novo numerador.
- o resultado nem sempre será uma fração na forma mais simples, e este um problema de que tenho que tratar em seguida.

A figura (71), página 185, apresenta um algoritmo gráfico para ilustrar a regra de soma de frações.

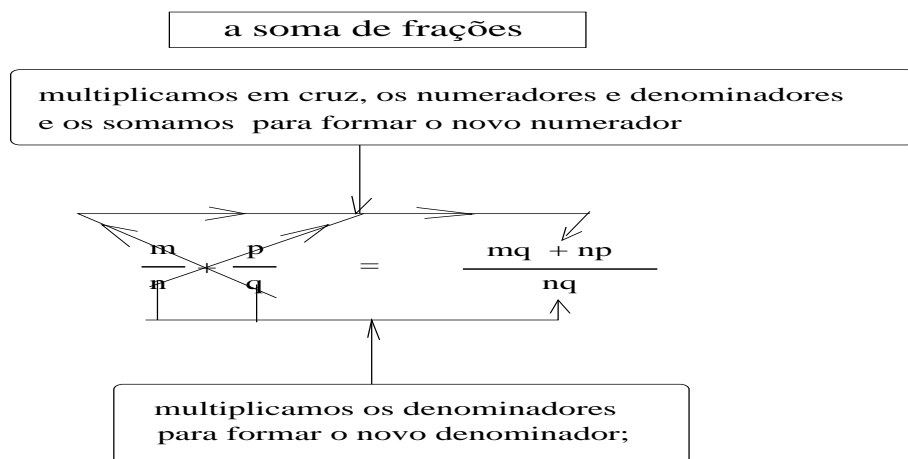


Figura 71: gráfico mostrando a soma de frações

O principal objetivo era conseguir que toda fração tivesse um inverso multiplicativo:

$$n \neq 0 \implies \frac{n}{m} \frac{m}{n} = \frac{mn}{mn} = 1 \quad (572)$$

não sendo necessário indicar que  $m \neq 0$  porque já excluímos a possibilidade de haver frações com denominador nulo.

O conjunto das frações é o conjunto dos números racionais e é designado com o símbolo  $\mathbf{Q}$ .

Resumindo, valem para este novo conjunto de objetos,  $\mathbf{Q}$ , todas as propriedades

1. A-1 No conjunto  $\mathbf{Q}$  existe um elemento neutro relativamente à **adição**.
2. A-2 Existe um inverso para todo número racional, relativamente à **adição** (inverso aditivo).
3. A-3 A *adição* é comutativa.
4. A-4 A *adição* é associativa.
5. M-1 Existe um elemento neutro relativamente à **multiplicação**.
6. M-2 Para todo número racional diferente de zero existe um inverso multiplicativo.
7. M-3 A *multiplicação* é comutativa.
8. M-4 A *multiplicação* é associativa.
9. AM-10 elemento neutro da adição, zero, multiplicado por qualquer número natural resulta em zero.

10. AM-2 a multiplicação é distributiva relativamente à adição.

Agrupei as propriedades em três classes, da adição (A), da multiplicação (M), e relativas à adição e multiplicação (AM). Observe que as propriedades da adição são as mesmas que da multiplicação com um pequeno detalhe sobre a exceção do zero na multiplicação. Estas quatro propriedades são as que definem a estrutura de grupo. Temos aqui um grupo aditivo, um grupo multiplicativo e as duas propriedades relativas a estes dois grupos. São estas propriedades que fazem de  $(\mathbf{Q}, +, \cdot)$  um corpo comutativo, porque a multiplicação é comutativa.

### Relação de equivalência

As relações de equivalência resolvem problemas formais de unicidade e outros problemas de identificação como é o caso de polígonos semelhantes que a geometria precisa de identificar. São objetos diferentes mas eles precisam ser “equivalentes”. A relação de equivalência é uma generalização da igualdade.

No caso das frações temos uma infinidade de frações que representam o mesmo número,  $\frac{n}{n} \equiv 1$ ;  $n \neq \text{zero}$ , por exemplo. Isto cria problemas para a Álgebra que precisa que o inverso dum número seja único, então como no caso dos triângulos semelhantes, precisamos de identificar as frações que representarem o mesmo número coloncando-as todas numa mesma *classe de equivalência*. Mas duas frações iguais forma o que chamamos de *proporção* então o *o produto dos meios é igual ao produto dos extremos* e assim chegamos à regra de equivalência de frações:

$$\frac{n}{m} \equiv \frac{p}{q} \iff nq = mp \quad (573)$$

é interessante observar que todas as frações equivalentes ficam sobre uma mesma reta determinada pela representação mais simples, pela fração irredutível, a figura (72) página 186, mostra as classes de equivalência das frações como

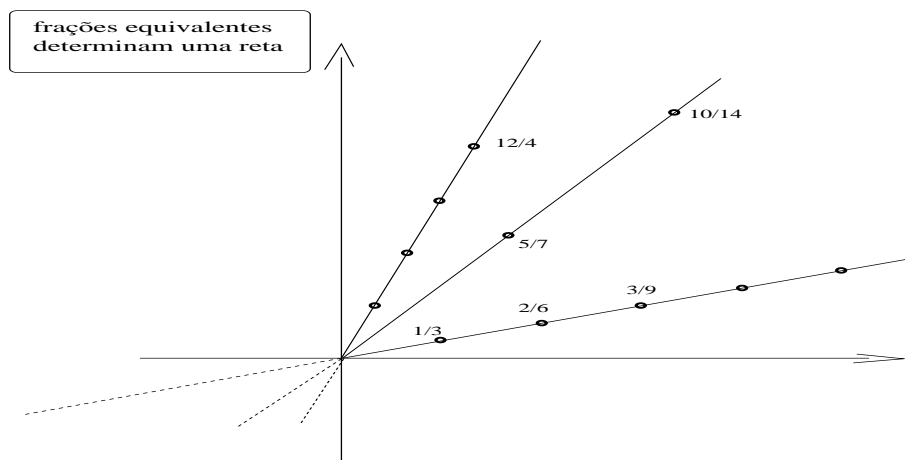


Figura 72: retas do plano de Gauss são classes de equivalência de frações

pontos das retas contidas no *plano de Gauss*, o produto cartesiano  $\mathbf{Z} \times \mathbf{Z}$ .

### Representação geométrica de $\mathbb{Q}$

Os números racionais tem uma propriedade de “densidade” que os inteiros não têm:

1. entre dois números racionais, sempre tem outro número racional;
2. dados dois número racionais, sempre tem outro à esquerda;
3. dados dois número racionais, sempre tem outro à direita;

A figura (73), página 187, compara com o que acontece na reta a propriedade

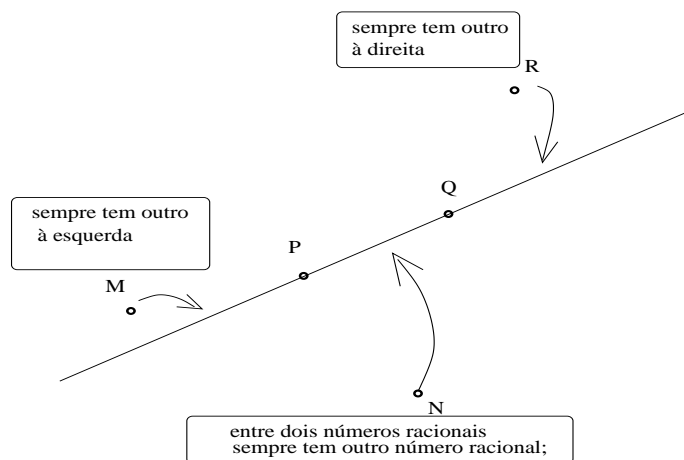


Figura 73: entre dois pontos, na reta ...

de densidade dos números racionais. Esta comparação sugere fazer uma interpretação geométrica dos números racionais.

1. Selecionamos um ponto na reta para a dividirmos em duas semiretas: a semireta positiva e a semireta negativa. É o representante do zero.
2. Depois, com um compasso selecionamos os inteiros à distâncias iguais, os inteiros positivos na semireta positiva, e os inteiros negativos na semireta negativa.
3. O espaço que sobra é para as frações que não são inteiros.
4. Depois veremos que que na reta ainda tem números que não são racionais, os números irracionais, que também encontram lugar na reta.

Aqui, novamente, estamos ante uma multitude de exemplares de retas representando  $\mathbb{Q}$  e simplesmente diremos que todas estas são equivalentes como representação de  $\mathbb{Q}$ , e isto vai nos permitir a definição geométrica das operações aritméticas de  $\mathbb{Q}$ . Por exemplo, a figura (74) página 188, ilustra o produto

A multiplicação geométrica

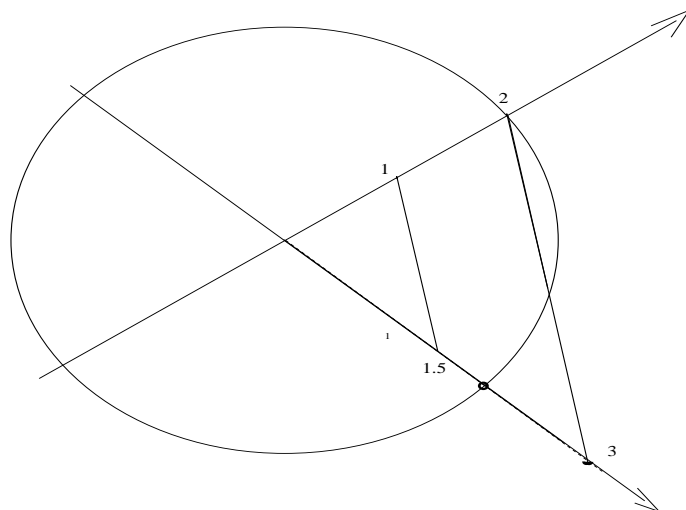


Figura 74: A multiplicação geométrica

$1.5 \times 2$  usando semelhança de triângulos. Escolhi duas retas para representar  $\mathbf{Q}$  e encontrei o resultado da multiplicação em uma das retas.

- **função** é um tipo de relação  $f$  entre dois objetos satisfazendo às propriedades

1. Existem dois conjuntos  $x \in A$  e  $y \in B$  designados, respectivamente, domínio de  $f$  e contradomínio de  $f$ .
2. Todo elemento  $x \in A$ , do domínio, tem um único elemento  $y \in B$ , do contradomínio, tal que  $y = f(x)$ , a notação para indicar que  $x, y$  estão relacionados, com frequência se usa a notação  $x \xrightarrow{f} y$ .

Notação:  $A \xrightarrow{f} B; A \ni x \mapsto y = f(x) \in B$ .

Observe que não há a condição de usar de todos os elementos do contradomínio  $B$ , em particular, a *função constante*, que associa todos os elementos do domínio  $A$  com um único elemento do contradomínio  $B$  é uma função de grande importância.

O subconjunto de  $B$  formado por todas as imagens  $f(x); x \in A$  é chamado de *conjuntos dos valores* de  $f$ .

Se todos os elementos do contradomínio forem utilizados, cabe pensar numa função inversa e ela existe de em dois casos:

1. se  $f$  for *injetiva*,

$$x \neq y \Rightarrow f(x) \neq f(y) \text{ ou } f(x) = f(y) \Rightarrow x = y$$

neste caso  $f$  é também bijetiva representando uma forma de identificar os dois conjuntos  $A, B$ .

Podemos usar isto para identificar duas estruturas algébricas, como  $(R, +)$  e  $(R^{++}, \cdot)$  porque a função  $\log$  é uma bijeção entre os dois conjuntos subjacentes a estas duas estruturas.

2. Se o conjunto  $B$  for todo utilizado, mas  $f$  não for injetiva confundindo elementos que tenha a mesma imagem, podemos definir uma relação de equivalência ( $\text{mod } f$ ) e definir uma função bijetiva do conjunto quociente ( $\text{mod } f$ ) em  $B$  que então será bijetiva. É isto que ocorre na congruência, nos restos da divisão por um número inteiro maior ou igual a 2. As classes de equivalência ( $\text{mod } f$ ) são definidas como os subconjuntos de  $A$ , do domínio, que tenham a mesma imagem em  $B$ .

As funções que utilizam todos os elementos do contradomínio, portanto quando *contradomínio* e *conjunto de valores* coincidem, se chamam *sobrejetivas*. Estas funções podem ser “*corrigidas*” quando definirmos uma função semelhante a  $f$  sobre as classes quociente ( $\text{mod } f$ ).

- **função aritmética** é uma função  $f$  definida de  $\mathbf{N}$  em  $\mathbf{C}$  tal que

$$f(1) = 1; \quad m, n \text{ primos entre si } f(mn) = f(m)f(n); \quad (574)$$

Uma função é dita ser *totalmente aritmética* se não houver a restrição de que  $m, n$  sejam primos entre si.

Este tipo de função é muito importante em *teoria dos números* e alguns exemplos são:

- $\tau(n)$  o número de divisores positivos de  $n$ ;
- $\sigma(n)$  a soma dos divisores de  $n$ ;
- $\omega(n)$  o número de fatores primos distintos de  $n$ ;
- $\Omega(n)$  o número de fatores primos de  $n$ ;
- A função de Euler,

$$\phi(n) \text{card}(\{m; m < n; (m, n) = 1\})$$

- **função generalizada** Ver *distribuição*.

- **funcional** Um *funcional* é uma função definida num conjunto de funções. A palavra foi inventada para evitar que falássemos de *uma função tomando valor numa função*, leia também sobre *funtor*. Mas a importância deste conceito vai muito além da semântica que pode estar por trás de sua invenção, os funcionais

são hoje um caso particular de *operadores* que também são *funções definidas em conjuntos de funções* e a *teoria dos operadores* é uma vasta área de investigação em Matemática que oferece boas possibilidades para resolver um tipo de equações, as *equações diferenciais*.

Podemos fazer uma classificação preliminar dos funcionais entre *funcionais lineares* e *funcionais não lineares*.

Alguns dos seguintes exemplos se encontram em [4, página 1].

- No conjunto das curvas do plano, podemos selecionar um subconjunto, daquelas que possamos calcular o comprimento, chamadas *retificáveis*.  $F$  pode ser o funcional que associa a cada curva retificável o seu comprimento. Pode-se ver que este funcional é uma medida semelhante à medida que fazemos, por exemplo, de figuras planas, o comprimento da união de duas curvas retificáveis é a soma dos comprimentos destas curvas, se elas forem disjuntas, como fariamos com área de figuras planas disjuntas, quando as áreas forem finitas. Isto mostra o nosso interesse em reagrupar conceitos com o objetivo de sistematizar a teoria, as medidas são uma classe de funcionais e elas podem ser redefinidas para ser vistas como *funcionais lineares*.
- Um segundo exemplo de [4, página 1] anexa um conceito auxiliar. Podemos ver as curvas como arames de um material, cobre por exemplo, ou fibra de vidro, com isto surge uma modificação do conceito de medida, porque tais arames deixam de ser *uniformes*, quer dizer que a medida de dois pedaços do mesmo “tamanho” não precisa ser a mesma, a densidade muda. Isto mostra que as medidas são um conceito que generaliza a medida geométrica com uma “densidade” multiplicativa. Para que você veja que o objetivo não é simples complicar, quando generalizamos, o próximo exemplo vai mostrar-lhe um caso *bem concreto* que é semelhante a este, tirado da *vida real*. As medidas são funcionais, aqui, vistas como funcionais definidos em “subconjuntos de curvas”, e uma curva é uma função, portanto a definição está preservada.
- Há duas formas de medirmos um terreno, um pedaço de terra. Uma delas é simplesmente geométrico, sua área. Outra consiste no preço que é uma deformação econômica da área. O preço usa a área como um coeficiente multiplicativo (ou vice-versa), este exemplo é semelhante ao anterior em que medimos arames considerando uma distorção produzida pela densidade não uniforme do material de que eles são feitos.
- O exemplo do comprimento de curva, ou do preço de terrenos, pode ser expresso como uma integral o que pode tornar mais claro para que servem *medidas*. Do Cálculo, o comprimento de uma curva é dado pela integral

$$\int_a^b \sqrt{1 + f'(x)^2} dx \quad (575)$$



este é o caso particular em que a curva é o  $\text{graf}(f); [a, b] \xrightarrow{f} \mathbf{R}$  e agora podemos recuperar o exemplo do arame escrevendo

$$\int_a^b \sqrt{1 + f'(x)^2} g'(x) dx \quad (576)$$

em que  $g'(x)$ , escrito como derivada de uma função  $g$  representa a variação (instantânea) da densidade do material de que é feito o arame. E como o objetivo de um dicionário é exatamente *complicar, para criar curiosidades* podemos dizer que esta integral não é exatamente mais uma *integral de Riemann*, mas uma integral de Riemann–Stieltjes em que você encontra o peso  $g'(x)$  alterando a integral e criando uma forma diferente de medir. A *teoria das medidas* é a generalização da integral de Riemann estudada no Cálculo.

- Quero terminar este verbete com uma bonita fórmula que também pode ser encontrada em [4, página 2], e vou partir da equação (575) em que posso interpretar a função, o integrando, como  $F(x, f(x), f'(x)) = \sqrt{1 + f'(x)^2}$  o que me leva a escrever

$$J(f) = \int_a^b F(x, f(x), f'(x)) dx \quad (577)$$

Observe a notação  $J(f)$ , se trata de um funcional, não me interessa por valores de  $x$  e até mesmo posso usar uma linguagem correta, mas “intrigante”, dizendo que nas equações (575), (576) e (577) não tem “x”! A equação (577), assim como as equações (575), (576), define um funcional em que a variável é  $f$ . A integral que se estuda no Cálculo é um *funcional*, é o primeiro exemplo, mas em geral não é apresentada assim!

Funcionais deste tipo foram estudados por alguns dos Bernouilli, Newton e L’Hôpital portanto no século 15, é o caso do célebre problema da braquistocrona em que se procurava o caminho mínimo que um corpo descreveria entre dois pontos do espaço em que se busca minimizar o funcional definido na equação (577), melhor, no caso da equação (576), porque ela mede o comprimento de curvas. Problemas deste tipo se chamam problemas variacionais

---

- **geometria analítica** foi criada para aplicar os métodos da álgebra e da aritmética às relações geométricas, comumente ligada ao nome de *René Descartes* que teria sido o idealizador do sistema de *coordenadas cartesianas*.

A Geometria Analítica define as equações de alguns lugares geométricos, como retas, planos, círculos, elipses, parábolas, hipérbolas, as chamadas cônicas. Confira

- equação do círculo, equação da elipse, equação da hipérbole, equação da parábola,
- equação da reta;
- plano coordenado, eixos coordenados;

O sistema de coordenadas cartesianas basicamente identifica os pontos da *reta numérica* com um número real e é isto que chamamos de *coordenada* dum ponto na reta.

Se identificarmos um ponto, numa reta qualquer, como sendo o *zero*, à direita do qual, por convenção se identifica um outro ponto como sendo o 1 como você pode ver na figura (fig 132), página 352, então criamos uma sistema para

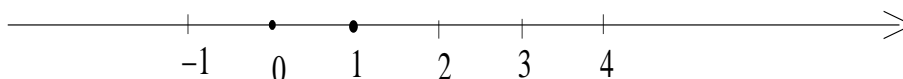


Figura 75: reta numérica, representação geométrica de  $\mathbf{R}$

“numerizar a reta” ou “digitalizar a reta”. Desta forma a reta deixa de ser apenas um ente geométrico e passa agora a ser uma representação do conjunto dos números reais  $\mathbf{R}$ .

- A escolha dum ponto para representar o zero dividiu a reta em duas semiretas.
- A escolha dum ponto para representar o número 1 selecionou uma das semiretas como a *semireta positiva*, mas fez mais do que isto definiu um segmento de reta com medida 1 o que nos permite propagar pela reta numérica todos os números inteiros usando, por exemplo, um compasso para marcar os inteiros positivos e negativos.
- Os números negativos são marcados na *semireta negativa*, naturalmente.

A semelhança de triângulos nos permite registrar a posição dos números racionais que não sejam inteiros, confira a figura (fig 76), página 193, onde você pode ver as frações da forma  $\frac{p}{7}; 0 \leq p \leq 7$ .

A *reta numérica* herda da *reta geométrica* uma propriedade importante dos segmentos de reta, confira a figura (fig 77), página 193,

- dois pontos diferentes,  $A, B$  determinam um segmento de reta,
- no segmento de reta  $\overline{AB}$  tem um ponto,  $C$  diferente dos extremos
- e um quarto ponto  $D$  fica fora do segmento.
- Como a reta real é orientada então o ponto que fica fora do segmento de reta, que agora chamamos de *intervalo*, pode ser maior do que  $B$  ou menor do que  $A$ , confira a figura (fig 77).

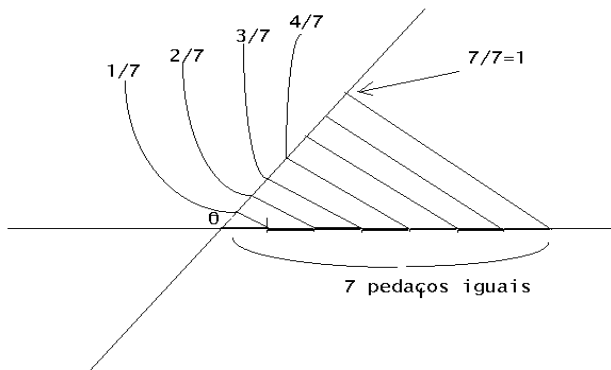


Figura 76: frações triângulos semelhantes

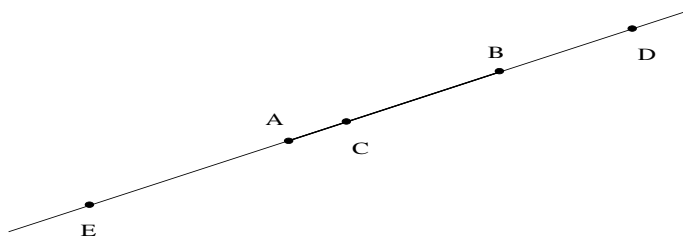


Figura 77: segmentos de reta na reta orientada

Para os números a propriedade adquire a redação “entre dois números diferentes quaisquer, sempre tem outro entre eles, um terceiro que é maior dos que estes dois e um quarto que é menor do que eles todos”.

Isto é consequência da escolha do 0 e do 1 criando uma *ordem na reta*.

Em símbolos ficaria assim, em que  $\mathbf{R}$  representa a reta numérica

$$a, b \in \mathbf{R}; a \leq b; \Rightarrow (\exists c \in \mathbf{R})(a \leq c \leq b); \quad (578)$$

$$a, b \in \mathbf{R}; a \leq b; \Rightarrow (\exists d \in \mathbf{R})(a \leq b \leq d); \quad (579)$$

$$a, b \in \mathbf{R}; a \leq b; \Rightarrow (\exists e \in \mathbf{R})(e \leq a \leq b \leq d); \quad (580)$$

e podemos iterar esta propriedade indefinidamente *porque a reta é infinita*.

Todas as propriedades da *reta geométrica* se aplicam a *reta numérica*, mas a *reta numérica* é uma *reta especializada* que tem mais propriedades o que a *transforma no conjunto dos números reais*.

Se além disto considerarmos um par de retas numéricas concorrentes no zero como mostra a figura (fig 78), página 194, podemos agora “numerisar” o plano. Com três retas podemos numerizar o espaço 3D e assim por diante. A figura (fig 78) é uma representação do produto cartesiano  $\mathbf{R} \times \mathbf{R} = \mathbf{R}^2$  criando o *plano*

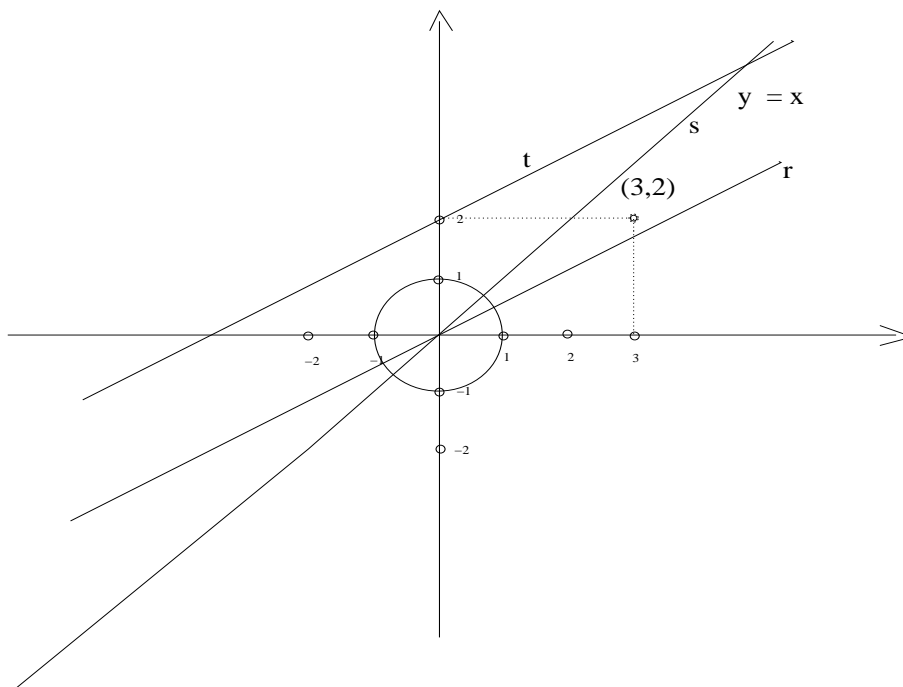


Figura 78: Sistema de coordenadas cartesianas

*numérico*. Mas esta denominação não é usada, dizemos *o plano coordenado*, a *reta numérica*.

Com três *retas numerizadas* se cortando perpendicularmente na origem, poderíamos representar geometricamente  $\mathbf{R}^3$ , ou numerizar o espaço  $\mathbf{R}^3$ .

Na figura (fig 78) estão representados vários pontos do plano, os pontos

$$\{(-2, 0), (-1, 0), (1, 0), (2, 0), (3, 0)\}$$

todos no que chamamos *eixo OX*, o eixo horizontal. Também estão marcados os pontos

$$\{(0, -2), (0, -1), (0, 1), (0, 2)\}$$

no eixo *OY*, e o ponto  $(3, 2)$ , que não pertence a nenhum dos eixos, também está marcado.

Agora é possível fazer referência a conjuntos geométricos com equações. Por exemplo, se estabelecermos a convenção de que a primeira coordenada ficará representada pela variável  $x$  e que a segunda coordenada será representada pela variável  $y$  então a equação da primeira bissetriz dos eixos, a reta que divide ao meio o ângulo determinado pelos eixos  $OX$  e  $OY$ , é  $y = x$ , porque todos os pontos sobre a primeira bissetriz tem as duas coordenadas iguais.

As duas coordenadas dum ponto no plano algumas vezes são chamadas de *abscissa*, a primeira coordenada, e *ordenada*, a segunda coordenada. Então sobre

a primeira bissetriz dos eixos qualquer ponto tem *abscissa* igual à *ordenada*.

Reciprocamente, se um ponto do plano tiver *abscissa* igual à *ordenada*, então é um ponto da forma  $(x, x)$ , quer dizer que satisfaz à equação  $y = x$ .

Com exemplo estou mostrando que o lugar geométrico do plano designado como primeira bissetriz fica inequivocamente identificado pela equação  $y = x$ . Você pode traduzir a equação  $y = x$  com a frase *é o lugar geométrico dos pontos  $P = (x, y)$  do plano tal que a abscissa de  $P$  é igual a ordenada de  $P$* .

#### **Observação 4** *As cônicas no Universo*

*Qualquer lugar geométrico do plano, ou do espaço, poderia, em princípio ser identificado com uma equação, mas nem sempre seria fácil encontrar uma equação para um conjunto arbitrário do plano. As cônicas, retas, planos, círculos, elipses, hipérbolas e parábolas são bem conhecidas e sabemos encontrar as equações destes lugares geométricos.*

*Também estes entes geométricos, as cônicas, respondem de forma muito boa às nossas necessidades de descrever a Natureza.*

*A forma esférica é um bom exemplo para começar. As circunferências ou esferas, são as formas geométricas naturais quando todas as forças estiverem equilibradas. No espaço, afastado de fortes influências gravitacionais, qualquer líquido toma a forma esférica.*

*A trajetória dos planetas no espaço tende a ser do tipo elipse que podemos entender como uma deformação do círculo. Se um corpo no espaço não estiver numa trajetória elíptica é quase certo que estará sobre uma parábola.*

*Retas é que praticamente não existem em nenhum ponto do Universo, devido força gravitacional entre os planetas da qual nem mesmo a luz escapa.*

Podemos identificar o *coeficiente angular* de algumas retas<sup>26</sup> estabelecendo uma proporção entre as coordenadas horizontais e verticais de cada ponto sobre ela. No caso da reta  $y = x$  esta proporção é 1 e então escrevemos  $y = 1x$  e neste caso se segue a convenção da álgebra não escrevendo 1.

A outra reta que aparece na figura (fig 78) teria um *coeficiente angular* positivo e menor do que 1 porque a coordenada vertical será sempre, em módulo, menor do que a coordenada horizontal, escrevemos  $y = mx$  em que o número  $m < 1$ .

Este coeficiente pode ser calculado, considerando a reta que aparece na figura, podemos medir o comprimento do segmento de reta que parte do ponto  $(1, 0)$  até encontrar a reta e este valor é  $m$ , no caso da reta  $r$  este valor é  $m = 0.56$  aproximadamente. Então a equação da reta  $r$  é  $y = 0.56x$ . Observe que o plano, agora, é uma região “métrica”, podemos medir as distâncias entre os pontos do plano, neste caso usamos uma régua para fazê-lo.

Isto vale para as retas que passem na origem. Para outra reta qualquer, que não passe na origem, confira na figura (fig 78), é o caso da reta  $t$ , podemos encontrar-lhe uma paralela passando na origem e aplicar o mesmo método para encontrar  $m$  porém a equação que podemos escrever facilmente seria a da reta paralela passando na origem:  $y = mx$ . É o caso da reta  $t$  que é paralela a reta  $r$ , na figura (fig 78).

Você pode ver que precisamos duma metodologia mais avançada para escrever a equação de qualquer reta. Podemos construir um método sem grandes dificuldades se aceitarmos algumas afirmações como óbvias, por exemplo, que a

<sup>26</sup>O coeficiente angular tem um defeito importante: as retas perpendiculares ao eixo  $OX$  não têm coeficiente angular.

equação de *quase todas*<sup>27</sup> as retas seria da forma  $y = mx + b$  então o número  $b$  é a distância ao longo do eixo  $OY$  entre as duas paralelas que no caso da reta  $t$  é 2 então a equação da reta  $t$  é  $y = 0.56x + 2$ .

Estes dois coeficientes recebem os nomes de *coeficiente angular* e *coeficiente linear* e na construção que acabei de fazer você viu as razões dos nomes escolhidos:  $m$ , *coeficiente angular*, é a razão de proporcionalidade entre as coordenadas horizontal e vertical, quando a reta passa na origem, mas ainda tem o mesmo sentido em qualquer reta paralela a elas.  $b$  é a distância de uma reta, ao longo do eixo  $OY$  à paralela que passa na origem, o *coeficiente linear*.

Agora, com uma expressão do tipo  $y = mx + b$ , em que você escolhe os valores de  $m$  e de  $b$  você pode facilmente fazer o gráfico da reta que tenha esta equação. Experimente com `gnuplot`, e aqui está o código

```
m1 = 0.5; m2 = 1; m3 = 1.5; m4 = -0.5; m5 = -1; m6 = -1.5;
b = 0.5; b = 1; b = 1.5; b = -0.5; b = -1; b = -1.5;
r1(x) = m1*x + b; r2(x) = m2*x + b; r3(x) = m3*x + b;
r4(x) = m4*x + b; r5(x) = m5*x + b; r6(x) = m6*x + b;
plot r1(x), r2(x), r3(x), r4(x), r5(x), r6(x), 0;
pause -2 "Aperte enter para terminar "
```

raspe e cole num terminal do `gnuplot` e você vai ver o gráfico das retas. `gnuplot` vai ler todos os valores da segunda linha mas vai usar o último apenas. Basta você apagar o último para ver novos gráficos. Ou digite, no terminal do `gnuplot`

```
b = 3;
replot;
```

e novos gráficos de retas, agora  $b = 3$  serão desenhados. Identifique os coeficientes angulares em cada gráfico para adquirir intuição a respeito do coeficiente angular.

Mais uma correção correção pode ser introduzida na equação da reta para encontrarmos sua equação usando duas informações:

- a reta pelo ponto  $(a, b)$ ;
- o seu coeficiente angular é  $m$

esta reta tem por equação

$$y = m(x - a) + b; \quad (581)$$

esta é a formulação ideal para trabalhar com equações de retas tangentes aos gráficos de funções no Cálculo.

O método que descrevi não vale para as retas que sejam perpendiculares ao eixo  $OX$ , mas a equação duma reta deste tipo é fácil de ser obtida, nelas a primeira coordenada será sempre constante, o valor, no eixo  $OX$  onde elas

<sup>27</sup>Este método não serve para as retas paralelas ao eixo  $OY$ , voltarei a seguir a este questão.

passam. Por exemplo  $x = 0$  é a equação do eixo  $OY$ , e  $x = 4$  é a equação de uma reta paralela ao eixo  $OY$  passando no ponto 4 do eixo  $OX$ . Mais geral, se a reta passa pelo ponto  $(a, b)$  e é paralela ao eixo  $OY$  então sua equação será  $x = a$ .

- **distância entre dois pontos** é conceito da Geometria Analítica generalizado na *teoria dos espaços métricos* que define *distância* abstraindo o sentido geométrico mas guardando as propriedades fundamentais da distância. Confira *distância*.

A Geometria Analítica *digitalizou* o espaço e a geometria euclidiana, tornando possível a algebrização da geometria. Assim  $(x, y)$  para dois números reais dados representam um ponto do espaço  $\mathbf{R}^2$ , que é um plano ou um espaço bidimensional, ou  $(x, y, z)$ , para três números reais dados representam um ponto do espaço  $\mathbf{R}^3$ , um espaço tridimensional, ou

$$P = (x_1, x_2, \dots, x_n); x_k \in \mathbf{R}; P \in \mathbf{R}^n \text{ um espaço de dimensão } n; \quad (582)$$

Dados dois pontos no  $P, Q \in \mathbf{R}^2$ , confira a figura (fig 79), página 197, podemos *calcular* a distância entre eles usando o teorema de Pitágoras,

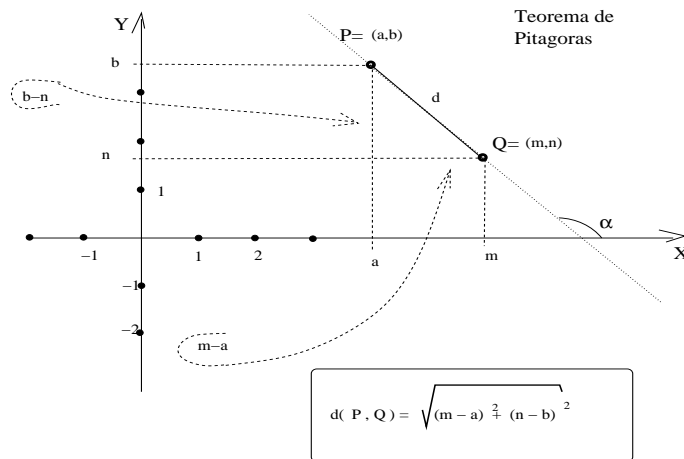


Figura 79: Distância entre dois pontos e o teorema de Pitágoras

$$d(P, Q) = \sqrt{(m-a)^2 + (n-b)^2} = \sqrt{(a-m)^2 + (b-n)^2}; \quad (583)$$

Observe que o módulo das diferenças que aparecem na equação (eq. 583) medem o comprimento dos lados do triângulo retângulo que o segmento  $\overline{PQ}$  determina com retas paralelas aos eixos coordenados.

As propriedades da distância são

1. reflexividade  $d(P, Q) \geq 0$  e  $d(P, Q) = 0 \iff P = Q$

2. simetria  $d(P, Q) = d(Q, P)$ ;
3. desigualdade triangular  $d(P, Q) \leq d(P, R) + d(R, Q)$

Qualquer função de duas variáveis que tenha estas propriedades define uma distância num conjunto  $M$  e é quando se diz que  $(M, d)$  é um espaço métrico.

A distância é uma propriedade bidimensional porque envolve em suas propriedades no máximo três elementos do espaço que determinam um plano deste espaço, o que facilita muito nas demonstrações das propriedades mais gerais da distância. Para o  $\mathbf{R}^n$  a fórmula na equação (eq. 583) fica

$$\begin{cases} d(P, Q) = \sqrt{\sum_{k=1}^n (p_k - q_k)^2}; \\ P = (p_1, \dots, p_n); Q = (q_1, \dots, q_n); \end{cases} \quad (584)$$

Uma forma de definir a distância passa pelo *produto interno* ou *produto escalar* que está definido em todos os espaços de dimensão finita e também em espaços de dimensão não finita. O *produto escalar*, como meio de definir distância também pode ser definido de várias maneiras. A definição usual no  $\mathbf{R}^n$  é

$$\langle P, Q \rangle = \sum_{k=1}^n p_k q_k = \|P\| \|Q\| \cos(\gamma); \quad (585)$$

em que  $\gamma$  é o ângulo entre  $P$  e  $Q$ .

A primeira expressão na (eq. 585) é uma *forma bilinear* definida em  $\mathbf{R}^n$  e a segunda é a expressão bidimensional da mesma no plano determinado pelos dois vetores.

Observe que a equação (eq. 585) nos oferece a oportunidade de definir ângulo entre dois vetores dum espaço qualquer em que esteja definido um produto escalar, usando a função  $\text{acos}()$ . Tais espaços se chamam *espaço com produto interno*.

Confira *produto escalar* ou *produto interno*.

### Observação 5 *triedro positivo*

A ordenação de vetores no espaço é uma questão importante e apenas traduz as leis da Física ou da Natureza. Observe como isto é feito usando a regra do saca-rolhas que também poderia ser chamada de regra do parafuso.

Como dois vetores no  $\mathbf{R}^n$  determinam um plano do  $\mathbf{R}^n$ , então o ângulo é um conceito geométrico claro, é o menor segmento do círculo trigonométrico que os dois vetores determinam. Na figura (fig 80), página 199, você vê dois vetores do  $\mathbf{R}^n$  que determinam um plano e tomando a origem comum aos dois vetores como centro de um círculo de raio 1 um dos vetores corta o círculo trigonométrico na origem do círculo.

Aqui parece haver alguns conceitos difusos, mas não é bem assim. Existe uma orientação padrão para os vetores no espaço, dada pela regra do saca-rolha. O saca-rolha segue a orientação dos parafusos, para enfiá-lo na rolha a rotação é negativa, é a mesma orientação para enfiar um parafuso numa peça. Com esta regra a figura (fig 80), mostra no círculo trigonométrico o ponto inicial marcado pela reta suporte do vetor  $\vec{v}$ , estando os vetores  $\vec{v}, \vec{u}$  orientados positivamente pela regra do saca-rolhas. Se houver um terceiro vetor ele vai ocupar a posição do saca-rolhas pronto para “sacar uma rolha” mostrando a posição de tres vetores formando um triedro com ordenação positiva.



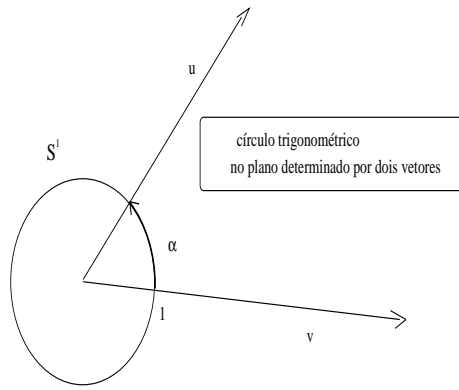


Figura 80: o ângulo entre dois vetores do  $\mathbf{R}^n$

Para que você se convença de que esta regra está consistente, deixe-me sugerir-lhe um experimento: considere dois vetores perpendiculares, tome a figura (fig 80) como padrão, pensando nos dois vetores  $\vec{v}, \vec{u}$  cortando o círculo trigonométrico na quarta parte,  $\frac{\pi}{2}$ .

Mas agora considere que  $|\vec{v}| = |\vec{u}| = 1$ , por construção, você selecionou  $\vec{v}, \vec{u} \in \mathbf{S}^1$ , no círculo trigonométrico.

O sacarolhas na origem dos eixos, o ponto comum dos vetores  $\vec{v}, \vec{u}$ , representando um terceiro vetor “arrancando a rolha”, na direção positiva, é o vetor  $\vec{z}$  e também  $|\vec{z}| = 1$ , tudo por construção.

Os dois vetores  $\vec{u}, \vec{z}$ , determinam um plano e o vetor  $\vec{v}$  é o sacarolhas nesta nova situação. O produto vetorial de vetores executa, algebricamente, esta operação:

$$\vec{v} \times \vec{u} = \vec{z}; \vec{u} \times \vec{z} = \vec{v}; \quad (586)$$

ou ainda, como na Física:

$$\vec{i} \times \vec{j} = \vec{k}; \vec{j} \times \vec{k} = \vec{i}; \quad (587)$$

sendo que esta operação não é comutativa...

$$\vec{k} \times \vec{j} = -\vec{i}; \quad (588)$$

A eletricidade se comporta de acordo com esta regra para definir o sentido da corrente positiva gerada por gerador montado com ímãs eletromagnéticos. A regra do sacarolhas apenas traduz o que se passa no mundo eletromagnético. Se você inverter as conexões nos polos duma bateria, o eletroímã vai rodar no sentido inverso:

$$\vec{j} \times \vec{k} = \vec{i}; \vec{k} \times \vec{j} = -\vec{i}; \quad (589)$$

Uma propriedade que tem uso intenso é a desigualdade de Cauchy-Bunyakovsky-Schwarz

$$| \langle P, Q \rangle | \leq \|P\| \|Q\| \quad (590)$$

que tem uma demonstração simples no plano, mas eu já observei que, dados dois vetores, eles determinam um plano do espaço a que eles pertencem portanto demonstrá-la num plano não é uma restrição.

Para isto preciso mostrar que a forma trigonométrica da definição na equação (eq. 585) é idêntica à expressão da soma de produtos na mesma equação porque

$$| \langle P, Q \rangle | = \|P\| \|Q\| | \cos(\gamma) | \leq \|P\| \|Q\|;$$

**Dem**:

Se  $n = 2$  é consequência direta da *fórmula de Euler* porque, dividindo pelo módulo dos vetores teremos:

$$\left\langle \frac{P}{\|P\|}, \frac{Q}{\|Q\|} \right\rangle = \frac{|\langle P, Q \rangle|}{\|P\|\|Q\|} \leq 1; \quad (591)$$

$$\left\langle \frac{P}{\|P\|}, \frac{Q}{\|Q\|} \right\rangle = \langle (\cos(\alpha), \sin(\alpha)), (\cos(\beta), \sin(\beta)) \rangle; \quad (592)$$

$$\left\langle \frac{P}{\|P\|}, \frac{Q}{\|Q\|} \right\rangle = \cos(\alpha - \beta) = \cos(\gamma); \quad (593)$$

Na equação (eq. 591) usei a propriedade das formas bilineares que permite a distribuição dos coeficientes  $\|P\|, \|Q\|$  entre os dois fatores<sup>28</sup>.

Na equação (eq. 592) usei uma representação plana dos dois vetores que podem ser expressos como pontos do círculo trigonométrico sendo  $\alpha, \beta$  os seus ângulos e  $\gamma$  é o ângulo entre eles. Na passagem da equação (eq. 592) para a equação (eq. 593) usei a expressão do produto escalar

$$\langle (\cos(\alpha), \sin(\alpha)), (\cos(\beta), \sin(\beta)) \rangle = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) = \cos(\alpha - \beta); \quad (594)$$

Os cálculos mostram que posso me reduzir, na demonstração, ao caso dos vetores unitários. Então, na sequência vou supor que  $\|P\| = \|Q\| = 1$ .

A passagem crítica, na demonstração por indução que está na passagem de  $n - 1$  para  $n$  que é semelhante à passagem da dimensão 2 para 3, mas esta é mais intuitiva. Deixe-me agora considerar  $P, Q \in \mathbf{R}^3$ , vetores unitários, portanto pontos de  $\mathbf{S}^2$  a esfera unitária do  $\mathbf{R}^3$ . Mas estes dois vetores determinam um plano que corta  $\mathbf{S}^2$  segundo um círculo unitário, uma curva plana, e, pela demonstração anterior,  $\cos(\gamma)$  é o valor do produto escalar entre eles, em que  $\gamma$  é ângulo entre os vetores.

Mas esta demonstração não seria necessária dentro do escopo duma demonstração por indução, ela é simplesmente tomada como verdadeira para todo  $N \leq n - 1$  e temos que considerar agora  $P, Q \in \mathbf{R}^n$ , dois vetores unitários. Mas estes vetores definem um plano do  $\mathbf{R}^n$  e o valor do produto escalar para dois vetores, no plano, em dimensão  $N < n$ , por hipótese de indução, é  $\cos(\gamma)$ , em que  $\gamma$  é o ângulo entre os vetores, o que termina a demonstração.

**q.e.d .**

Um outro exemplo simples de lugar geométrico é o círculo. A geometria o define como *é o lugar geométrico dos pontos P do plano cuja distância a um ponto fixo C é constante*. Esta distância constante de  $P$  a  $C$  é o raio  $r$ .

A equação do círculo que aparece na figura (fig. 78) é simples de ser obtida. Observe que é o círculo trigonométrico. Qualquer ponto em cima deste círculo fica à distância 1 da *origem*  $(0, 0)$  *dos eixos*. Chamando de  $(x, y)$  um ponto genérico sobre o círculo e aplicando o teorema de Pitágoras para representar a distância dele até a *origem* encontramos  $\sqrt{x^2 + y^2} = 1$  ou ainda:

$$x^2 + y^2 = 1 \text{ equação do círculo trigonométrico} \quad (595)$$

A equação de um círculo de raio  $R$  com centro no ponto  $(a, b)$  do plano seria

$$d((x, y), (a, b)) = R; d((x, y), (a, b))^2 = (x - a)^2 + (y - b)^2 = R^2 \quad (596)$$

que você pode obter usando novamente o teorema de Pitágoras e a definição de círculo.

Muito mais difícil seria, dada uma equação, como

$$x^3 + 3x^2y + xy^2 + y^3 = 4 \quad (597)$$

<sup>28</sup>as forma bilineares generalizam o produto de números que um exemplo de forma bilinear.

descobrir que figura seria representada por esta equação. Mas é possível com um programa de computador encontrar alguns pontos que satisfaçam esta equação e obter um gráfico de uma parte da tal figura, aproximadamente, se for possível, se a equação representar alguma figura “real”. A geometria algébrica é a parte da Matemática que tenta responder a esta questão para caso das expressões algébricas. Na figura (fig 81), página 201, você pode ver parte da curva algébrica

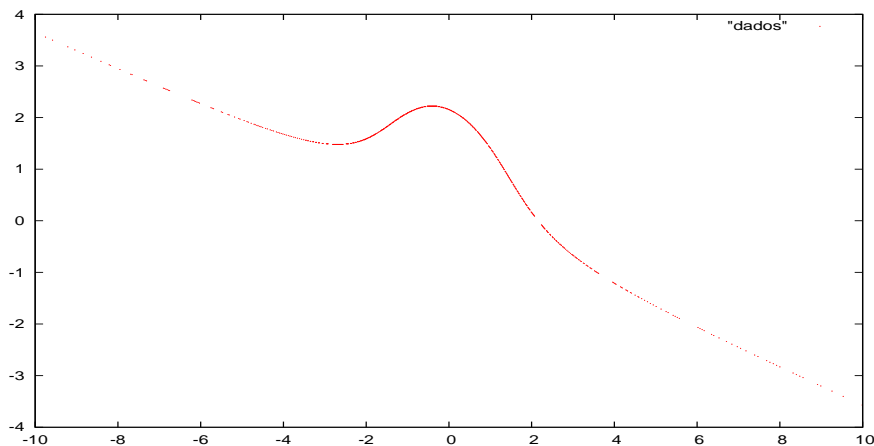


Figura 81: Parte da curva algébrica

cuja equação se encontra na (eq.597). Ela foi obtida com um programa escrito em Python que varreu um retângulo do plano procurando por pontos que estivessem próximos desta equação. Do gráfico se pode produzir a pergunta: “teria este gráfico uma reta assíntota?”. Uma tal pergunta se pode resolver com métodos avançados do Cálculo.

O programa pode ser visto na figura (fig 82), página 202, E você pode alterar a equação de  $F(x, y)$  para visualizar o gráfico de outra equação, se este gráfico puder ser feito.

Entenda como funciona o programa e modifique-o para obter outras visualizações desta curva ou de outras. Na figura (fig 82), há uma indicação sobre “terminal” o programa está preparado para produzir um arquivo “postscript”, se você quiser que gnuplot exiba o gráfico na tela, simplesmente comente as linhas:

```
# transfere.write( "set terminal postscript eps enhanced color \n");
# transfere.write( "set output \"GeometriaAnalitica_03.eps\" \n");
```

isto é feito colocando o sinal do “jogo da velha na primeira posição da linha, em python, respeitando a tabulação.

Na imagem do programa, há mais duas setas indicando os pontos onde selecionar a precisão do gráfico, **passo** é a precisão da malha, e **delta** é a precisão com a desigualdade vai ser resolvida. Experimente que não há riscos. Tente o valor grande, por exemplo, 0.5 para **delta** e você verá uma “faixa” no centro da

```

#! /usr/bin/python
# -*- coding: utf-8 -*-
import sys
from posix import popen
from math import *

def F(x,y):
    return pow(x,3) + 3*pow(x,2)*y + x*pow(y,2) + pow(y,3);

def cria_dados(inicio, fim, passo, delta,valor):
    dados = open('dados','w')
    x=inicio; y=fim; A=valor;
    while(x<fim):
        y = inicio;
        while(y<fim):
            if (abs(F(x,y) - A) < delta):
                dados.write(str(x)+" "+str(y)+"\n");
                y += passo;
            x += passo;
        dados.close();

def cria_transfere():
    transfere = open('transfere','w')
    transfere.write("set pointsize 0.1 \n");
    transfere.write("plot \"dados\" with points\n");
    transfere.write("set terminal postscript eps enhanced color \n");
    transfere.write("set output \"GeometriaAnalitica_03.eps\" \n");
    transfere.write("plot \"dados\" with points\n");
    transfere.write("pause -2 \"Aperte enter para terminar\" \n");
    transfere.close();

def principal():
    print "Lados do retângulo para o gráfico"
    inicio = input("na horizontal, inicio =");
    fim = input("na horizontal, fim =");
    passo = input("O passo da malha, passo = ");
    delta = input("Precisão para resolver a desigualdade, delta = ");
    print "Vou na verdade calcular uma curva de nível e preciso o \"n\". \n";
    "o valor de A na expressão F(x,y) = A ";
    A = input("O valor de A = ");
    cria_dados(inicio, fim, passo, delta,A);
    cria_transfere();
    popen("gnuplot transfere");

```

Figura 82: programa para traçar curvas de nível

qual se encontra a curva. Na verdade, para qualquer valor escolhido o resultado será uma “faixa”.

Manipule os valores da precisão onde está indicado na figura para conseguir melhor resultado mas comece com valores não muito pequenos, tente algo em torno de  $\delta = 0.1$  para começar, depois use um  $\delta$  menor. O tempo de processamento é consequência do incremento dado a  $x$  e a  $y$  guardado na variável `passo`, o valor dado a esta variável será o responsável pelo tempo de processamento. Se você escolher um valor muito pequeno prepare-se para aguardar algumas horas até que o gráfico apareça na tela.

Você pode baixar o programa da página

<http://calculo-numerico.sobralmatematica.org/programas//CurvaAlgebrica.py>

Algebricamente, e com algum trabalho, podemos escrever as equações de muitos dos subconjuntos do plano, em particular círculos, hipérbolas e parábolas são estudadas na Geometria Analítica, são as chamadas *cônicas*.

Esta curvas se originam geometricamente de cortes de um cone com duas folhas. Na figura (83) página 203, você vê um cone cortado por um plano paralelo ao eixo do cone, mas não precisava que o plano fosse paralelo ao eixo, apenas não podia chegar a ser paralelo à geratriz do cone. Este detalhe pode ser melhor discutido na construção da equação da hipérbole, confira esta equação em outro lugar.

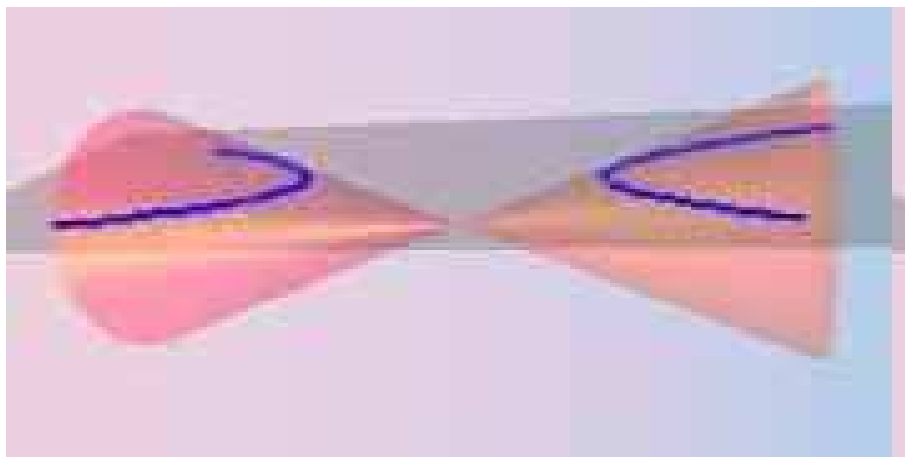


Figura 83: Plano paralelo ao eixo do cone: hipérboles

Se o plano for paralelo à geratriz o resultado é uma parábola que pode ser vista na figura (84) página 203, Mas também aqui se tem variantes, o plano

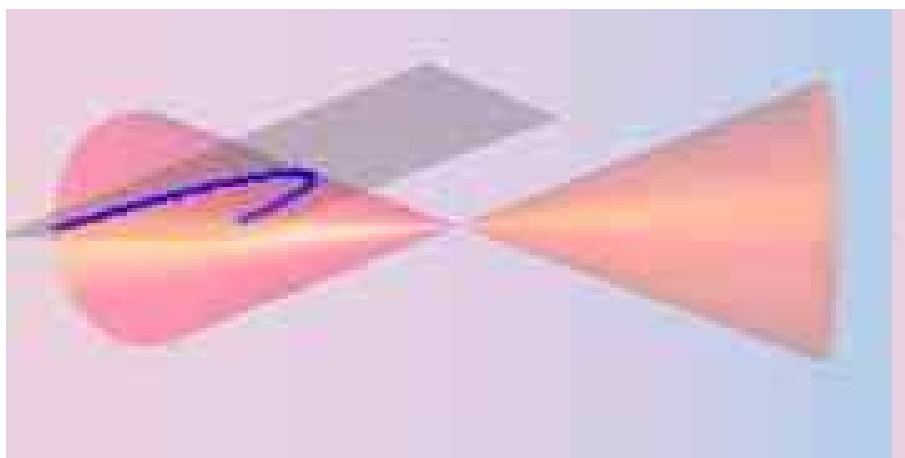


Figura 84: Quando o plano é paralelo à geratriz: parábola

pode cortar a folha do cone em vários pontos resultando numa parábola mais ou menos aberta, vou discutir isto quando construir a equação da parábola na última seção.

Se o plano tiver inclinação maior do que a da geratriz o resultado será uma curva fechada que pode ir de elipse até círculo, vou discutir isto na próxima seção com a equação do círculo quando vou mostrar que o círculo é uma elipse degenerada que é o contrário do que todo mundo diz. . . a figura (85) página 204,

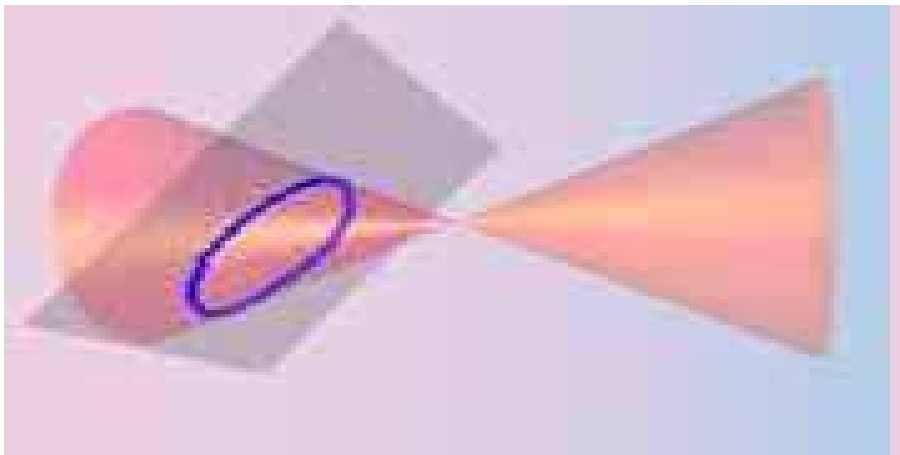


Figura 85: A elipse degenerada: círculo

As fotos das interseções do cone foram copiadas, com autorização, da página do Dr. Anthony Rynne, University of Limerick

[http://www3.ul.ie/~rynet/swconics/planes\\_cutting\\_coneA.htm](http://www3.ul.ie/~rynet/swconics/planes_cutting_coneA.htm)

Na figura (86) página 205, você vê plano  $XOY$  e o eixo do cone que aparece nas figuras 83-85 é perpendicular ao plano  $XOY$  passando pelo ponto  $(0, 0)$ . No plano  $ZOY$  se podem ver duas retas que se cruzam no ponto  $(0, 0, 0)$  a interseção do cone com este plano. Análise que corresponde às figuras 83-85 é relativa às possíveis forma como um plano intercepte o plano com o plano  $ZOY$  que estão representadas na figura (87) página 206,

A região hachuriada é a projeção do cone no plano  $ZOY$ . Se um plano cortar perpendicularmente  $ZOY$  com inclinação **entre** os limites da região hachuriada, corta o cone segundo uma hipérbole. Se o plano tiver **exatamente** a inclinação das retas que limitam a região hachuriada irá cortar uma das folhas do cone segundo uma parábola. Se a inclinação do plano estiver **fora** da região hachuriada irá cortar uma das folhas do cone segundo uma curva fechada, elipse ou círculo e o caso do círculo corresponde **exatamente** ao eixo  $OZ$ .

Há uma infinidade de variantes para as equações de cônicas que podem ser assim obtidas, mas basicamente elas dependem da inclinação da *geratriz do cone* e do ponto  $(0, a, b)$  por onde passa o eixo do cone e isto pode dar um livro de 400 páginas como o livro de Lehmann, ou qualquer outro que se tenha ao trabalho de descrever estas possibilidades. A minha opção é descrever o caso que vou chamar de “*padrão*” em que a interseção do cone que aparece na figura 87 são as bissetrizes dos eixos do plano  $YOZ$ . Qualquer outra variante pode ser obtida com uma mudança de coordenadas da forma

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} \quad (598)$$

em que  $x, y$  são as coordenadas usada pela equação padrão e  $x', y'$  são as novas

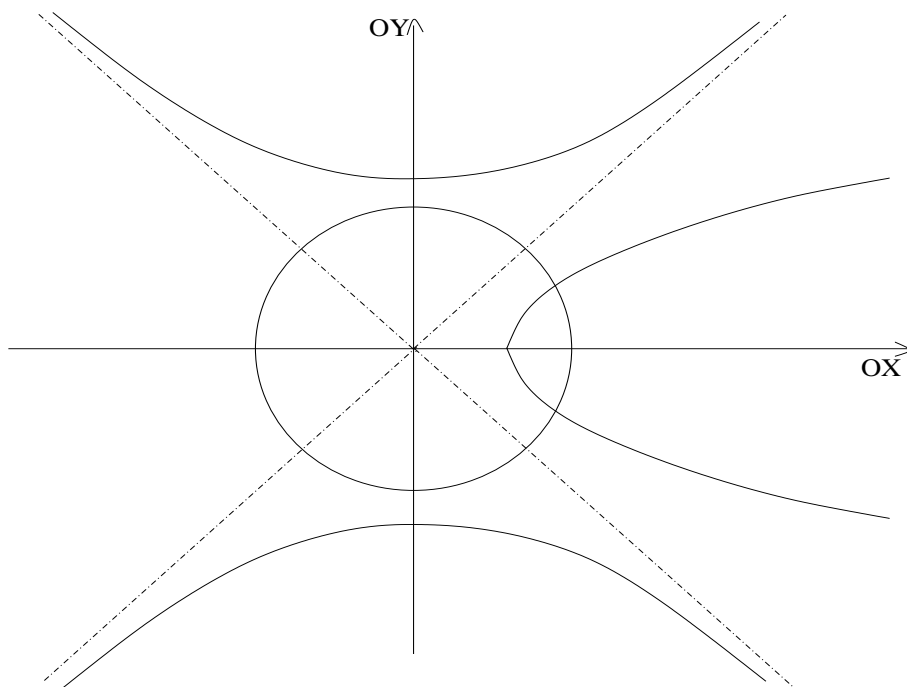


Figura 86: imagem das cônicas no plano  $XOY$

coordenadas depois da transformação. A matriz quadrada deve ter determinante diferente de zero para que seja possível desfazer a mudança. O resultado desta mudança de variáveis será uma forma quadrática

$$Ax'^2 + By'^2 + Cx'y' + Dx' + Ey' + F = 0 \quad (599)$$

e existe uma análise da relação dos coeficientes que expressa qual é a forma padrão que originou esta forma quadrática. É importante repetir esta análise rigorosamente porque a expressão na equação (599) serve também de padrão para classificação das equações diferenciais parciais entre parabólicas, elípticas ou hiperbólicas.

- **geometria algébrica** é o estudo das variedades geométricas que possam ser obtidas como os zeros de expressões algébricas. Por exemplo,

$$F(x, y) = x^2 + y^2 - 1; F(x, y) = 0 \equiv x^2 + y^2 = 1; \quad (600)$$

o círculo unitário é conjunto dos zeros do polinômio  $w = F(x, y)$ . Qualquer círculo é uma *variedade algébrica*, isto é uma *variedade* que corresponde aos zeros dum polinômio.

Todas as cônicas, da Geometria Analítica, são exemplos de variedades algébricas fazendo da Geometria Analítica um caso particular da *geometria algébrica*.

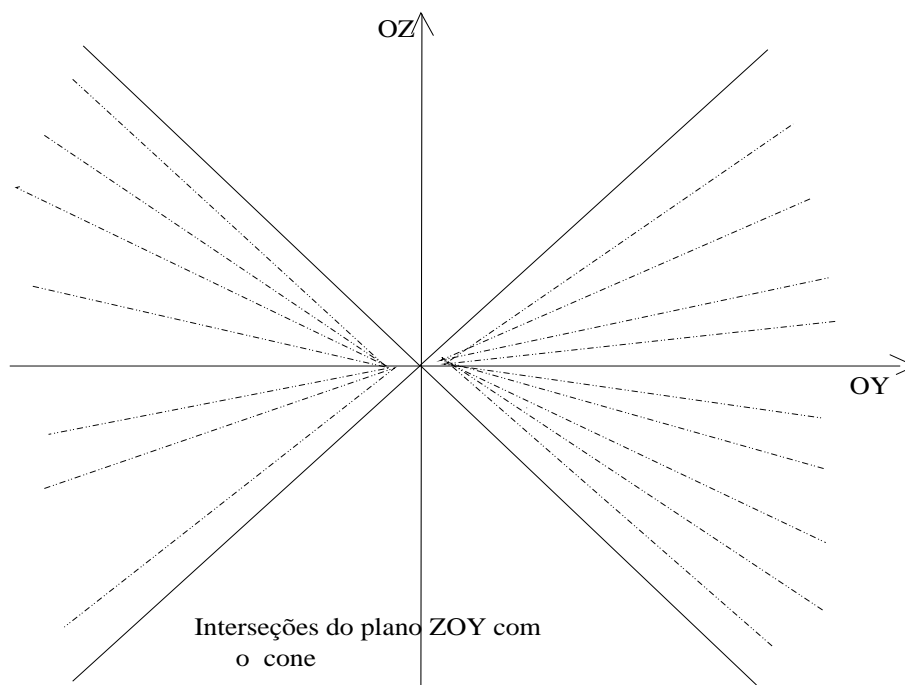


Figura 87: Interseções do plano ZOY com o cone

Como curiosidade, há números algébricos, são pontos da reta, variedades de dimensão zero, que são soluções de equações algébricas cujos coeficientes sejam racionais, e o número  $\pi$  não é algébrico pois não é solução de nenhuma equação algébrica com coeficientes racionais. É o caso do número de Euler,  $e$ , também. Números algébricos são assim um caso particular de variedades algébricas, as de dimensão zero.

A geometria algébrica pode ser vista como o estudo de um caso particular de anéis, o dos polinômios com coeficientes racionais. Parte dos métodos usados por Andrew Wiles para demonstrar o último teorema de Fermat vêm da geometria algébrica, mas isto já deu um livro de mil páginas e primeira demonstração de Andrew Wiles continha um erro que levou mais dois anos para ser corrigido.

O caso mais geral da geometria algébrica retira esta restrição sobre os coeficientes e poderia descrever qualquer tipo de variedade?

---

- **gnuplot** é um programa que tem por objetivo fazer gráficos de funções ou curvas definidas parametricamente. É distribuído livremente, mas, apesar do nome, não faz parte do projeto GNU. A página do **gnuplot** é

<http://www.gnuplot.info>

É interessante observar que **gnuplot** pode ser usado como uma linguagem de programação, pois admite funções recursivas e assim é possível fazerem-se



programas em `gnuplot`, [7].

- **gradiente** é o vetor funcional formado pelas derivadas parciais de uma função multivariada. Por exemplo, se  $z = F(x, y)$  então

$$\text{grad}(F) = \left( \frac{\partial F}{\partial x}, \frac{\partial F}{\partial y} \right); \quad (601)$$

Ao calcularmos a derivada implícita de  $z = F(x, y)$  o gradiente de  $F$  aparece naturalmente:

$$dz = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy \quad (602)$$

de onde podemos deduzir a equação do plano tangente com as substituições

$$dx := x - a; dy := y - b; dz := z - c; \quad (603)$$

$$z - c = \frac{\partial F}{\partial x} \Big|_{x=a, y=b} (x - a) + \frac{\partial F}{\partial y} \Big|_{x=a, y=b} (y - b); \quad (604)$$

$$dz = \text{grad}(F) \begin{pmatrix} dx \\ dy \end{pmatrix} \quad (605)$$

Na equação (eq. 603) as variáveis  $dx, dy, dz$  foram interpretada como diferenças permitindo-nos obter a equação dum polinômio do primeiro grau, observe que  $\frac{\partial F}{\partial x} \Big|_{x=a, y=b}$  é um número, o valor desta derivada calculada quando  $x = a, y = b$ , o mesmo significando  $\frac{\partial F}{\partial y} \Big|_{x=a, y=b}$  que é um número.

Na equação (eq. 604) foi feita uma nova interpretação da equação (eq. 603) em que  $\text{grad}(F)$  aparece como uma matriz aplicada a um vetor de dimensão dois. Comparada esta equação com a derivação de funções univariadas, vemos que o valor da derivada de funções multivariadas, em cada ponto onde forem calculadas, representam funções lineares que conduzem à *variedade linear tangente*. No caso univariado a *reta tangente*, no caso bivariado ao *plano tangente*.

Se calcularmos a derivada implícita da expressão  $F(x, y) = c$  podemos deduzir a seguinte sucessão de afirmações:

- $F(x, y) = c$  é uma curva de nível, se existir.
- da equação (eq. 604) que o gradiente é perpendicular à curva de nível no ponto  $(a, b)$  se  $F(a, b) = c$ ;
- $(a, b)$  for uma solução da equação  $F(x, y) = c$ , ou, em outras palavras, se o ponto  $(a, b, c)$  pertencer ao gráfico da função  $z = F(x, y)$ , então o  $\text{grad}(F)$  é perpendicular à curva de nível no ponto  $(a, b)$ .

- **Gram-Schmidt** é um algoritmo para completar a base dum espaço vetorial a partir de alguns vetores dados, ou de um vetor dado. A ideia do algoritmo pode ser assim descrita:

- Um espaço vetorial de dimensão  $n, E$ ;

- Você tem uma coleção de vetores linearmente independentes,  $u_1, \dots, u_m \in E; m \leq n$ ;
- quer uma base ortonormal para  $E$  constituída a partir deste vetores dados.

O processo de *Gram-Schmidt* constrói a base  $e_1, \dots, e_n$

1.  $e_1 = \frac{u_1}{\|u_1\|}$ ;
2.  $e_2 = u_2 - \langle u_2, e_1 \rangle e_1$ ;
3.  $e_2 := \frac{e_2}{\|e_2\|}$ ;
4.  $k > 2$ 
  - (a)  $e_k = u_k - \sum_{j=1}^{k-1} \langle u_k, e_j \rangle e_j$  ;
  - (b)  $e_k := \frac{e_k}{\|e_k\|}$ ;

Se  $m < n$  a partir de  $k = m + 1$  defina  $u_k$  sendo qualquer vetor linearmente independente dos vetores  $\{e_1, \dots, e_m\}$  e existe algum porque a  $\dim(E) = n > m$ .

Ou seja, o processo se inicia pela substituição de  $u_1$  por um vetor unitário da direção de  $u_1$ . Constrói  $e_2$  em dois passos, primeiro define  $e_2$  subtraindo de  $u_2$  a componente na direção de  $e_1$ , em seguida normaliza  $e_2$  dividindo-o pelo seu módulo.

Segue-se o processo sempre eliminando de  $u_k$  as componentes na direção dos  $e_j; j < k$ ; e depois normaliza o vetor  $u_k$  dividindo-o pelo seu módulo.

O resultado é uma lista de  $n$  vetores todos de módulo 1 perpendiculares a todos os anteriores já construídos. Como o resultado é uma base ortonormal para o espaço  $E$ , este processo também é chamado de *processo de ortonormalização de Gram-Schmidt*.

- **Green, teorema de** Este teorema é um dos resultados mais importantes do *Cálculo multivariado* junto com outros teoremas que podem ser considerados extensões ou complementações dele: teorema de Stokes e o teorema da divergência de Gauss.

O teorema de Green tem uma versão trivial pela qual vou começar e que serve para classificar os campos vetoriais que vou usar ao final na expressão do teorema.

Se  $F$  for uma campo vetorial, uma função de duas variáveis, por exemplo, continuamente diferenciável, então, pelo teorema de *Schwartz-Clairaut*, as derivadas mistas são iguais

$$\frac{\partial F^2}{\partial x \partial y} = F_{xy} = F_{yx} = \frac{\partial F^2}{\partial y \partial x}$$

o que torna a integral

$$\int_D \int (F_{yx} - F_{xy}) dx dy = 0$$

nula. Como podemos calcular as primitivas destas funções, é possível deduzir desta integral a integral de linha

$$\oint_{\partial\mathcal{D}} F_x dx + F_y dy$$

em que agora o símbolo  $\partial\mathcal{D}$  representa a fronteira do domínio  $\mathcal{D}$  e esta integral é também nula. Se eu alterar um pouquinho a notação vou obter a expressão comum nos livros de Cálculo.

$$P(x, y) = F_x(x, y); \quad Q(x, y) = F_y(x, y); \quad (606)$$

$$\oint_{\partial\mathcal{D}} P(x, y)dx + Q(x, y)dy = \int_{\mathcal{D}} (Q_x - P_y) dxdy \quad (607)$$

que é a expressão (trivial) do *teorema de Green* quando partimos de uma função diferenciável  $F$ , porque todas as integrais envolvidas são nulas. Se  $(P, Q)$  for um campo vetorial diferenciável continuamente, ainda vale o teorema de Green mas as integrais não precisam ser nulas sobre um domínio  $\mathcal{D}$  qualquer. A integral de linha, por exemplo, separa os campos vetoriais em duas classes:

- Campos conservativos, é o caso trivial, quando o campo vetorial é a derivada de um campo escalar. Então a integral de linha sobre qualquer curva fechada é zero, é uma aplicação direta do *teorema fundamental do Cálculo*.
- Campos não conservativos, quando houver uma curva fechada, fronteira de um domínio  $\mathcal{D}$  sobre a qual a integral de linha na equação (607) é diferente de zero.

o valor da integral de linha é então a perda (ou ganho) de energia que o campo escalar sofre ao longo da curva  $\partial\mathcal{D}$ . Neste caso o campo vetorial  $(P, Q)$  não tem primitiva. Esta formulação permite ainda explicar dois tipos de integrais,

- integrais independentes do caminho aquelas, da forma

$$\oint_{\partial\mathcal{D}} P(x, y)dx + Q(x, y)dy$$

que são nulas sobre qualquer curva fechada. O campo escalar é conservativo, tem primitiva (vem da derivada de um campo escalar diferenciável).

- integrais que dependem do caminho aquelas, da forma

$$\oint_{\partial\mathcal{D}} P(x, y)dx + Q(x, y)dy$$

que podem ser não nulas sobre uma curva fechada. O campo escalar é não conservativo e não tem primitiva (não vem da derivada de um campo

um domínio não convexo no plano

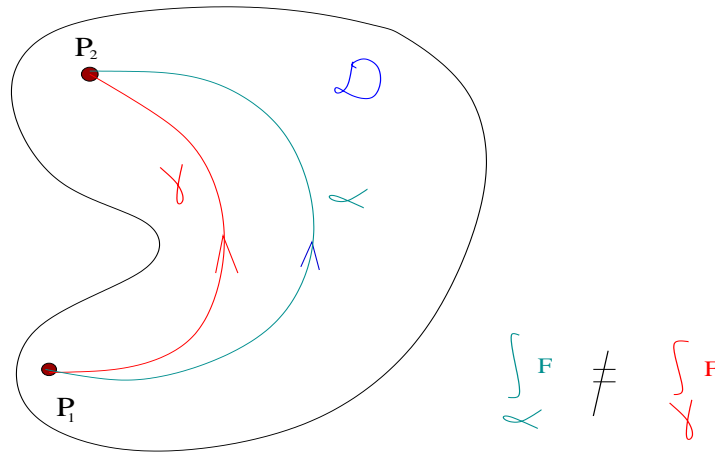


Figura 88: duas curvas ligando um ponto

escalar diferenciável). Dizemos que integral depende do caminho porque, escolhidos dois caminhos entre dois pontos dados  $P_1, P_2$ , como se pode ver na figura (88) página 210, se o valor da integral sobre um dos caminhos, de  $P_1$  até  $P_2$  for diferente do valor da integral sobre o outro caminho, também de  $P_1$  até  $P_2$ , podemos definir uma curva fechada, indo de  $P_1$  até  $P_2$ , então a integral será diferente zero sobre esta curva fechada. Isto equivale a dizer-se que o campo vetorial  $(P, Q)$  não tem primitiva, não é a derivada de um campo escalar.

- **grupo** É uma das estruturas algébricas da Matemática.

**Definição 11 (grupo)**  $(G, *)$

Considere  $G$  um conjunto no qual esteja definida uma relação binária  $*$  tal que

$$(\forall a, b \in G) a * b \in G \quad (608)$$

$$(\forall a, b, c \in G) a * (b * c) = (a * b) * c; \quad (609)$$

$$(\exists! e \in G) (\forall a \in G) a * e = e * a = a \quad (610)$$

$$(\forall a \in G) (\exists! b \in G) a * b = b * a = e; \quad (611)$$

Enão dizemos que  $(G, *)$  é um grupo.

Com frequência os autores usam uma frase mais simples dizendo apenas "G é um grupo" quando é possível deixar implícito qual é a operação binária que está definida em G.

O axioma expresso na equação (608) estabelece a existência de uma relação binária,  $*$ , e que G é fechado para a mesma. A equação (609) estabelece que  $*$  é associativa. A equação (610) estabelece que para a operação  $*$  existe um único elemento de G, chamado *identidade* e aqui designado com o símbolo  $e$ . A equação (611) estabelece a existência de um *inverso* para todo elemento de G relativamente à operação  $*$ .

A equação (611) também garante a existência de dois tipos de bijeção: para cada elemento  $x \in G$ , de uma solução única para a equação  $x * b = e$  cuja demonstração é consequência das equações:

$$x \mapsto b * x \text{ é uma bijeção de } G; \quad (612)$$

$$x \mapsto x * c \text{ é uma bijeção de } G; \quad (613)$$

1. A sentença na equação (612) é verdadeira porque, por absurdo, falhando a sobrejetividade,  $b * x = c$  não teria solução para algum elemento de G mas como  $b$  tem inverso chega-se ao absurdo:  $x = b^{-1} * c$ . Falhando a injetividade haveria duas soluções para  $b * x = c$  o que contradiz a unicidade do inverso.
2. A sentença na equação (613) tem uma demonstração simétrica.

Estas duas equações (612) (613), implicam na existência de bijeções obtidas na multiplicação por um elemento fixo, à direita ou à esquerda. Uma das formas de trabalhar com grupos finitos consiste em analisar as tabelas operatórias, estas bijeções tem um efeito importante: em cada coluna ou linha da tabela temos a presença de todos os elementos de G, (sobrejetividade) sem repetição (injetividade).

Uma outra forma de parafrasear a sentença acima é a seguinte: As equações (612) e (613), têm uma importância particular: elas criam permutações dos elementos de G indexadas pelo elemento multiplicador<sup>29</sup>. Se G for finito a quantidade de tais permutações é  $n = \text{card}(G) = \text{ordem}(G)$ . Se  $n > 2$  estas permutações formam um subconjunto próprio de todas as permutações dos elementos de G.

Interpretando G como um conjunto, como as permutações são bijeções de G, portanto funções inversíveis cuja composição é outra permutação, então o conjunto das permutações de G tem também a estrutura de grupo designado com o símbolo  $Sim(n)$  e a indexação mencionada acima identifica G como um subconjunto de  $Sim(n)$ .

Isto sugere pensar em *subgrupo*.

**Definição 12 (subgrupo)** Se  $H \subset G$  for fechado para a transformação

$$(x, y) \in H \times H \mapsto x * y^{-1} \quad (614)$$

<sup>29</sup>Os operadores multiplicação e translação aparecem em praticamente todas as estruturas em Matemática.

então  $(H, *)$  é também um grupo, e dizemos que  $H$  é subgrupo de  $G$ .

Em particular a equação (614) vale para  $G$  que é então um subgrupo de si próprio. Se  $H$  for subgrupo de  $G$ , diferente de  $G$ , observamos isto dizendo que  $H$  é um *subgrupo próprio* de  $G$ .

Também podemos usar a equação (614) considerando  $G$  como subconjunto de  $Sim(n)$ ;  $n = card(G)$  o que prova que  $G$  pode ser identificado como um subgrupo das permutações de  $G$ .

Consequentemente, encontraremos *todos os grupos finitos* dentro dos *grupos de permutações*.

Há diversos desenvolvimentos que podem seguir esta trilha, num deles se associa ao grupo das matrizes quadradas inversíveis de ordem  $n$  (entradas reais, complexas ...) com a operação de multiplicação de matrizes, este é um grupo multiplicativo de matrizes que não é comutativo. As permutações das colunas da matriz identidade produz uma imagem matricial de  $Sim(n)$ . O resultado é a *teoria da representação de grupos finitos*, com uso em codificação, por exemplo.

Existe interesse em considerar  $G$  como um subgrupo de si mesmo, em alguns problemas se consegue realçar propriedades significativas com auxílio da transformação  $h : x \mapsto a * x * a^{-1}$ , chamada *conjugação*, quando se considera um elemento fixo  $a \in G$ . Por exemplo, é possível descobrir pares de subgrupos que tem as mesmas propriedades, subgrupos conjugados e isomorfos. A conjugação é um endomorfismo de  $G$ .

Uma outra transformação importante é *inversão*

$$x \in G \mapsto x^{-1} \quad (615)$$

que é uma bijeção de  $G$  preservando todas as propriedades (608)-(611). Esta função é uma “reparametrização” de  $G$  sendo um *endomorfismo de  $G$* .

Vou usar na continuação a notação multiplicativa para a operação de  $G$  o que implica na suposição de que  $G$  não é comutativo, como se faz usualmente.

O conjunto de todos os *endomorfismo de  $G$*  tem naturalmente a estrutura de anel e podemos identificar  $G$  como subconjunto deste anel. Notação:  $End(G)$  é o anel de todos os endomorfismo de  $G$  em que a adição é definida ponto-a-ponto e a multiplicação é a composição de morfismos:  $(End(G), +, o)$ .

Existe um endomorfismo considerado *canônico* de  $G$  que é construído da seguinte forma.

- Considere a permutação dos elementos de  $G$  que equivale a multiplicação de  $G$  por um elemento dado  $g \in G$

$$\phi_g \in (End(G); \phi_g(x) = xg; \quad (616)$$

- e obviamente também teríamos

$${}_g\phi \in (End(G); {}_g\phi(x) = gx; \quad (617)$$

- cabe discutir quando  $g\phi = \phi_g$  o que leva à definição de *centro*( $G$ ), que é o subgrupo de  $G$  dos elementos que comutam com todos os elementos de  $G$ . Então as permutações de  $G$  são morfismo de grupo de  $G$  quando obtidas via um elemento do *centro*( $G$ ).

Mas se definirmos

$$\phi \in \text{End}(G); \phi(x) = g^{-1}xg \quad (618)$$

então temos o endomorfismo canônico de  $G$ , sendo um simples cálculo a verificação de que valem as propriedades dos morfismo de grupo para  $\phi$ .

Então, dado um elemento  $g \in G$  ele identifica um endomorfismo de  $G$ , o endomorfismo canônico de  $G$ , dado pela equação (eq.618) e assim vemos uma imersão de  $G$  no anel  $(\text{End}(G), +, \circ)$  o que produz o resultado surpreendente de que a estrutura de anel pode substituir em simplicidade a estrutura de grupo uma vez que todo grupo se identifica canonicamente com um anel.

Um questão que se pode levantar neste contexto é se a imagem de  $G$  seria o centro de  $\text{End}(G)$  e a resposta é não o que aliás mostra que esta imersão preserva uma propriedade importante de  $G$  porque a imagem do centro de  $G$  está estritamente contida no centro de  $\text{End}(G)$ .

- **grupo comutativo** Se  $(G, *)$  for um grupo e a operação binária for comutativa

$$(\forall a, b \in G) a * b = b * a; \quad (619)$$

então dizemos que  $(G, *)$  é um *grupo comutativo*. Não havendo dúvida dizemos apenas que  $G$  é um grupo comutativo. Há uma tradição de usar a notação aditiva quando se tratar de grupos comutativos:  $(G, +)$  é um grupo comutativo:

$$(\forall a, b \in G) a + b \in G; + \text{ é uma operação binária} \quad (620)$$

$$(\forall a, b, c \in G) a + (b + c) = (a + b) + c; \text{ propriedade associativa} \quad (621)$$

$$(\exists 0 \in G) (\forall a \in G) a + 0 = 0 + a = a; \text{ existência do elemento neutro} \quad (622)$$

$$(\forall a, b \in G) a + b = b + a; \text{ propriedade comutativa} \quad (623)$$

$$(\forall a \in G) (\exists ! b \in G) a + b = b + a = 0; \text{ existência do inverso} \quad (624)$$

e neste caso o elemento neutro é usualmente designado pelo símbolo 0.

Entretanto há exemplos de *grupos multiplicativos comutativos*, portanto esta forma simplista de ver *grupos multiplicativos* como *não comutativos* não é um padrão embora seja usada.

#### Exemplo 4 (Exemplos) de grupos

1. Os números inteiros com a adição usual é um grupo comutativo:  $(\mathbf{Z}, +)$ .
2. Os números racionais, os reais e os complexos eliminados do zero, com a multiplicação usual são grupos comutativos, e a notação para estes grupos é

$$(\mathbf{Q}^*, \cdot); (\mathbf{R}^*, \cdot); (\mathbf{C}^*, \cdot);$$

observe que nos dois primeiros casos, o cone formado pelos elementos estritamente positivos são subgrupos. Os logaritmos identificam estes subgrupos com os grupos aditivos dos reais (ou de  $\mathbf{Q}$ ). No caso dos complexos isto também é possível fazer quando se chega ao logaritmo complexo passando pela fórmula de Euler

$$e^{i\theta}$$

3.  $M_n(K)$  em que  $K \in \mathbf{Q}, \mathbf{R}, \mathbf{C}$  das matrizes quadradas de ordem  $n$  são grupos aditivos (comutativos). Os produtos cartesianos de qualquer quantidade dos grupos  $K$  também são grupos comutativos com a soma definida coordenada a coordenada como no caso  $M_n(K)$ , estes produtos cartesianos são apenas um caso particulares de matrizes não quadradas sobre  $K$ . O caso genérico sendo se considerarmos a dimensão fixa  $m \times n$  então  $M_{nm}(K)$  é também um grupo comutativo com a adição definida coordenada a coordenada.
4. Um grupo não comutativo é obtido quando se considera em  $M_n(K)$  a multiplicação de matrizes, então o conjunto dos elementos inversíveis, com determinante diferente de zero, em  $M_n(K)$ , relativamente à multiplicação de matrizes é um grupo. Este grupo designado por  $GL_n(K)$ ;  $K \in \mathbf{Q}, \mathbf{R}, \mathbf{C}$  ou o grupo geral linear.
5. Um caso bem simples mostra a importância de  $GL_n(\mathbf{R})$ , quando  $n = 2$ , as matrizes da forma

$$\begin{bmatrix} \cos(\theta) & -\text{sen}(\theta) \\ \text{sen}(\theta) & \cos(\theta) \end{bmatrix} \quad (625)$$

em que  $\theta$  é um número real qualquer (usualmente chamado de ângulo). Ao multiplicar matrizes deste tipo, se conclui, com auxílio da relação de soma de ângulos da trigonometria, que este conjunto é fechado para a multiplicação com o inverso obtido pela troca de sinal de  $\theta$  portanto um subgrupo<sup>30</sup> de  $GL_2(\mathbf{R})$ . Essas são as matrizes de rotação da Geometria Analítica plana. De forma semelhante podemos obter as matrizes de rotação da Geometria Analítica em  $3D$ .

6. O conjunto das permutações de um número finito,  $n$ , de objetos também pode ser visto como um grupo chamado  $Sim(n)$  que é um exemplo de grupo finito, não comutativo quando  $n > 2$ .

Trabalhar com permutações é muito complicado e inclusive os programas para construir permutações, diretamente, são difíceis de serem construídos. Como a multiplicação de um grupo finito, por um seu elemento fixo, produz uma permutação dos seus  $n$  então  $Sim(n)$  pode ser representado, e facilmente manipulado, ao ser identificado às permutações das colunas (ou linhas) da matriz identidade de  $GL_n(\mathbf{R})$  reduzindo as operações

<sup>30</sup>Se  $A, B$  forem duas matrizes de rotação, então  $AB^{-1}$  é uma matriz de rotação.



de  $Sim(n)$  à multiplicações de matrizes e já existem vários pacotes computacionais em que a multiplicação de matrizes está programada. Esta é uma breve introdução à representação de grupos finitos porque qualquer grupo finito se encontra representado n'algum  $Sim(n)$ .

7. Os restos na divisão por um inteiro fixo  $n$ , com a adição de restos, é um grupo comutativo designado por  $\mathbf{Z}_n$  em que cada elemento representa uma classe de equivalência (mod  $n$ ).

- **grupo normal** assunto da teoria dos grupos não comutativos. Se um grupo for comutativo, todos os seus subgrupos são normais. Os grupos não comutativos são todos isomorfos a algum subgrupo de permutações porque qualquer elemento do grupo induz uma permutação no conjunto dos elementos do grupo, os conceitos *elemento do grupo* e *permutação* se confundem, uma vez que todo elemento dum grupo pode ser considerado como uma permutação dos elementos do grupo.

Na teoria dos grupos existe o hábito de designar por multiplicativos os *grupos não comutativos*, estou aderindo a esta convenção.

Seja  $(G, *)$  um grupo e  $H$  um subgrupo de  $G$ ,  $H$  é dito normal se for invariante por conjugação:

$$(\forall a \in G) aHa^{-1} = c_a(H) = H; c_a(h) = a(ha^{-1}); \quad (626)$$

$$c_a : G \longrightarrow G; G \ni g \mapsto aga^{-1}; \quad (627)$$

a função, ou permutação, definida na equação (627) é chamada de *conjugação*.

A equação (626) pode ser escrita como,  $a*H = H*a$ , que são duas *translações* de  $H$  chamadas *classe lateral à esquerda* e *classe lateral à direita*, respectivamente. Um grupo é normal se estas classes coincidirem para qualquer elemento  $a \in G$  permitindo a definição de uma classe de equivalência módulo  $H$ . Se um grupo não for normal as classes à direita e à esquerda serão diferentes mas ainda é possível definir duas relações de equivalência módulo  $H$ , chamadas de à direita e à esquerda. A vantagem dos grupos normais é a unificação das classes que permite a definição do grupo quociente  $G/H$ .

**grupo finito normal** O símbolo  $|G|$  representa a *ordem de  $G$*  é o *card*( $G$ ). Se  $|G|$  for finita (grupo finito) então  $|H|$  divide  $|G|$  (porque todas as classes tem o mesmo número de elementos) e há  $m = \frac{|G|}{|H|}$  “translações” distintas de  $H$ . O conjunto destas translações<sup>31</sup> é o conjunto  $G/H$ , das classes *mod* $H$ . Em geral estes dois conjuntos são distintos, no caso dos subgrupos normais eles coincidem e representam uma partição de  $G$  definindo uma única relação de equivalência<sup>32</sup> em  $G$  dita *módulo  $H$*

Dadas duas *translações* de  $H$ ,  $h_i, h_j \in G$ , afirmo que

<sup>31</sup>Na literatura, as classes são com frequência chamadas de translações de  $H$  pensando no caso aditivo.

<sup>32</sup>Quando  $H$  não for normal, definem relações de equivalência, ou classes de equivalência, ditas à esquerda ou à direita.

- $h_i * H$  e  $h_j * H$  coincidem, ou
- $h_i * H$  e  $h_j * H$  são disjuntas,

e então posso representar

$$G/H = \{\overline{h_1}, \dots, \overline{h_m}\} = \{h_1 * H, \dots, h_m * H\} \quad (628)$$

usando um subconjunto de  $m$  elementos de  $G$  como “representantes” das possíveis classes.

**Teorema 7 (Grupo quociente)  $G/H$**

Se  $H$  for um subgrupo normal de  $G$ , notação:  $H \triangleleft G$ , as classes formadas pelas translações de  $H$  pelos elementos de  $G$  tem a estrutura de grupo, designado pelo símbolo  $G/H$ , é o grupo quociente de  $G$  pelo subgrupo normal  $H$ .

**Dem**:

**Lema 1 (a relação de subgrupo)** isomorfismo com um subgrupo de  $G$

Observando que  $G/H = \{\overline{h_1}, \dots, \overline{h_p}\}$  está representado como um subconjunto de  $G$  (isomorfismo de conjuntos) e como  $H$  é normal então é verdade que

$$\overline{a}, \overline{b} \in G/H \Rightarrow \overline{a * b^{-1}} \in G/H;$$

**Dem**:

Considere os seguintes cálculos em que estou usando distintos elementos

$$h, h', h'' \in H$$

para representar um elemento de uma classe porque nem sempre seria o mesmo elemento de  $H$  uma vez que o grupo  $G$  não é necessariamente comutativo.

$$\overline{a}, \overline{b} \in G/H \iff \quad (629)$$

$$\overline{a} = a * H; \overline{b^{-1}} = b^{-1} * H; \quad (630)$$

$$\overline{a * b^{-1}} = \{(a * h') * (b^{-1} * h'')\}_{h', h'' \in H} = \quad (631)$$

$$= \{(a * h') * (h * b^{-1})\}_{h', h \in H} = \{a * (h' * h) * b^{-1}\}_{h', h \in H} = \quad (632)$$

$$= \{a * (h * b^{-1})\}_{h \in H} = \{(a * b^{-1}) * h'\}_{h' \in H} \iff \quad (633)$$

$$\iff \overline{a * b^{-1}} \in G/H \quad (634)$$

Estou usando a associatividade, implicitamente, nas equações (632), (633). **q.e.d.**

Pelo lema, a relação para  $G/H$  seja um grupo é satisfeita o que demonstra o teorema.

**q.e.d.**

**Exemplo 5 (Grupo quociente) casos comutativo e não comutativo**

1.  $(\mathbf{Z}, +)$  sendo comutativo todos os seus subgrupos são normais e todos são da forma  $n\mathbf{Z}$  para algum inteiro positivo  $n$  (vale para zero). Então  $\mathbf{Z}/n\mathbf{Z}$  são as classes dos restos na divisão por  $n$  e no caso  $n = 0$  resulta no próprio  $\mathbf{Z}$ , mas usualmente não se considera este caso pela perturbação que ele causa na aritmética....

Ainda assim é interessante considerar estes dois casos:

- quando  $n = 0$ ,  $n\mathbf{Z} = 0$  portanto as classes são todas unitárias o que faz de  $\mathbf{Z}/n\mathbf{Z} \equiv \mathbf{Z}$  e o resultado fica coerente com a discussão feita anteriormente. Entretanto se fica impedido em falar de resto na divisão por  $n$ .
- quando  $n = 1$  existe uma única classe e  $\mathbf{Z}/n\mathbf{Z} \equiv \{0\}$  o que dá uma justificativa para que não se considere o número 1 como primo.

$$G = \mathbf{Z}_{10}; H = 2\mathbf{Z}_{10}; \quad (635)$$

$$\mathbf{Z}_{10}/2\mathbf{Z}_{10} = \mathbf{Z}_2 = 5\mathbf{Z}_{10} \subset \mathbf{Z}_{10}; \quad (636)$$

2.  $G = \text{Sim}(4)$  é o grupo das permutações de quatro elementos,

$$\{1, 2, 3, 4\}; |\text{Sim}(4)| = 4! = 24; \quad (637)$$

Usando a notação de permutações considere o subgrupo

$$H = \{e, (12), (34), (12)(34)\} \quad (638)$$

então a ordem de  $H$  é 4 e se for um grupo normal de  $\text{Sim}(4)$  haverá 6 classes à direita que serão idênticas às classes à esquerda, ou simplesmente classes quociente, que formarão o grupo quociente  $G/H$ .

### Verificando se $H$ é normal, ou não

Primeiro caberia mostrar que  $H$  é um subgrupo e a construção da tabela é simples pela quantidade de elementos:

	$e$	$(12)$	$(34)$	$(12)(34)$
$e$	$e$	$(12)$	$(34)$	$(12)(34)$
$(12)$	$(12)$	$e$	$(12)(34)$	$(34)$
$(34)$	$(34)$	$(12)(34)$	$e$	$(12)$
$(12)(34)$	$(12)(34)$	$(34)$	$(12)$	$e$

que é uma tabela de grupos. Como, em qualquer linha desta tabela tenho uma listagem dos elementos de  $H$ , vou copiar esta linha formando outra tabela multiplicando-a, sucessivamente por elementos de  $\text{sim}(4)$ . Após algumas multiplicações irão aparecer todos os elementos de  $\text{sim}(4)$  totalizando seis “translações”<sup>33</sup> e vou parar assim que eu obtiver seis conjuntos distintos: as seis classes possíveis.

Na primeira coluna vou deixar o multiplicador e vou usar a marca “✓” para indicar um conjunto diferente de  $H$ , uma nova classe.

<sup>33</sup>Continua-se a usar a linguagem dos grupos comutativos para fazer referência às classes laterais como “translações”.

	$e$	$(12)$	$(34)$	$(12)(34)$
✓	$ee$	$(12)$	$(34)$	$(12)(34)$
✓	$e(1234)$	$(234)$	$(124)$	$(24)$
✓	$e(1243)$	$(243)$	$(123)$	$(23)$
✓	$e(1324)$	$(13)(24)$	$(14)(23)$	$(1423)$
✓	$e(1342)$	$(134)$	$(142)$	$(14)$
repetida	$e(1423)$	$(14)(23)$	$(13)(24)$	$(1324)$
✓	$e(1432)$	$(143)$	$(132)$	$(13)$

As contas foram feitas com `scilab`, eis a operação feita para obter a ultima linha da tabela acima:

`I_12*I_1432, I_34*I_1432, (I_12*I_34)*I_1432`

basta raspar, colar no terminal do `scilab` e interpretar o resultado lendo as permutações das linhas da matriz identidade.

Possivelmente se pode automatizar todo este processo fazendo um programa na linguagem do `scilab` que reconheça o resultado e já identifique as permutações como um produto de ciclos.

Observe que usei os 4-círculos porque eles são os mais fáceis de serem escritos e agora tenho uma descrição completa de  $\text{sim}(4)$ :

$$\begin{aligned}
 \text{sim}(4) &= & (639) \\
 &= \{e, (12), (13), (14), (23), (24), (34)\} \cup & (640) \\
 &\quad \cup \{(12)(34), (13)(24), (14)(23)\} \cup & (641) \\
 &\quad \cup \{(123), (124), (132), (134), (142), (143), (234), (243)\} \cup & (642) \\
 &\quad \cup \{(1234), (1243), (1324), (1342), (1423), (1432)\} & (643)
 \end{aligned}$$

Vou agora multiplicar  $H$  à esquerda, para obter as classes à esquerda, se forem as mesmas classes então  $H$  é normal.

$H$	$e$	$(12)$	$(34)$	$(12)(34)$
✓	$ee$	$(12)$	$(34)$	$(12)(34)$
✓	$(1234)e$	$(134)$	$(123)$	$(13)$
✓	$(1243)e$	$(143)$	$(124)$	$(14)$
✓	$(1324)e$	$(14)(23)$	$(13)(24)$	$(1423)$
✓	$(1342)e$	$(234)$	$(132)$	$(23)$
repetida	$(1423)e$	$(13)(24)$	$(14)(23)$	$(1324)$
✓	$(1432)e$	$(243)$	$(142)$	$(24)$
✓	$(124)e$	$(14)$	$(1243)$	$(143)$

A operação para obter a última coluna:

`I_124*I_12, I_124*I_34, I_124*(I_12*I_34)`

Coletando os elementos tenho:

$$e, (12), (13), (14), (23), (24), (34); \leftarrow 7 \quad (644)$$

$$(12)(34), (13)(24), (14)(23); \leftarrow 3 \quad (645)$$

$$(123), (124), (132), (134), (142), (143), (234), (243); \leftarrow 8 \quad (646)$$

$$(1234), (1243), (1324), (1342), (1423), (1432); \leftarrow 6 \quad (647)$$

foram obtidos todos os elementos de  $\text{sim}(4)$  e resta verificar se as classes são idênticas: não são, a classe à direita de  $\overline{(1234)}$  é composta por

$$\{(1234), (234), (124), (24)\} \quad (648)$$

e a classe à esquerda de  $\overline{(1234)}$  é composta por

$$\{(1234), (134), (123), (13)\} \quad (649)$$

então  $H$  não é um subgrupo normal de  $\text{sim}(3)$ .

No quadro seguinte você tem a representação matricial, ou representação linear do grupo  $\text{sim}(4)$  com a qual eu fiz a verificação das contas aplicando as permutações do grupo à linhas da matriz identidade. Eu obtive a expressão dos ciclos dos elementos de  $\text{sim}(4)$  quando coletei os dados da tabela 2 colocados nas equações (eq. 639)-(eq. 642).

$$\begin{aligned} I &= [1,0,0,0; 0,1,0,0; 0,0,1,0; 0,0,0,1] \\ I_{12} &= [0,1,0,0; 1,0,0,0; 0,0,1,0; 0,0,0,1] \\ I_{13} &= [0,0,1,0; 0,1,0,0; 1,0,0,0; 0,0,0,1] \\ I_{14} &= [0,0,0,1; 0,1,0,0; 0,0,1,0; 1,0,0,0] \\ I_{23} &= [1,0,0,0; 0,0,1,0; 0,1,0,0; 0,0,0,1] \\ I_{24} &= [1,0,0,0; 0,0,0,1; 0,0,1,0; 0,1,0,0] \\ I_{34} &= [1,0,0,0; 0,1,0,0; 0,0,0,1; 0,0,1,0] \\ I_{12_{34}} &= [0,1,0,0; 1,0,0,0; 0,0,0,1; 0,0,1,0] \\ I_{13_{24}} &= [0,0,1,0; 0,0,0,1; 1,0,0,0; 0,1,0,0] \\ I_{14_{23}} &= [0,0,0,1; 0,0,1,0; 0,1,0,0; 1,0,0,0] \\ I_{123} &= [0,0,1,0; 1,0,0,0; 0,1,0,0; 0,0,0,1] \\ I_{124} &= [0,0,0,1; 1,0,0,0; 0,0,1,0; 0,1,0,0] \\ I_{132} &= [0,1,0,0; 0,0,1,0; 1,0,0,0; 0,0,0,1] \\ I_{134} &= [0,0,0,1; 0,1,0,0; 1,0,0,0; 0,0,1,0] \\ I_{142} &= [0,1,0,0; 0,0,0,1; 0,0,1,0; 1,0,0,0] \\ I_{143} &= [0,0,1,0; 0,1,0,0; 0,0,0,1; 1,0,0,0] \\ I_{234} &= [1,0,0,0; 0,0,0,1; 0,1,0,0; 0,0,1,0] \\ I_{243} &= [1,0,0,0; 0,0,1,0; 0,0,0,1; 0,1,0,0] \\ I_{1234} &= [0,0,0,1; 1,0,0,0; 0,1,0,0; 0,0,1,0] \\ I_{1243} &= [0,0,1,0; 1,0,0,0; 0,0,0,1; 0,1,0,0] \\ I_{1324} &= [0,0,0,1; 0,0,1,0; 1,0,0,0; 0,1,0,0] \\ I_{1342} &= [0,1,0,0; 0,0,0,1; 1,0,0,0; 0,0,1,0] \\ I_{1423} &= [0,0,1,0; 0,0,0,1; 0,1,0,0; 1,0,0,0] \\ I_{1432} &= [0,1,0,0; 0,0,1,0; 0,0,0,1; 1,0,0,0] \end{aligned}$$

Raspe e cole num terminal do *scilab* ou *octave* para obter a representação matricial de  $\text{sim}(4)$ .

3. Verificando se o subgrupo  $H$  gerado pelo elemento  $(12345) \in \text{sim}(5)$  é normal.

$$H = \{e, (12345), (13524), (14253), (15432)\} \quad (650)$$

	$I$	$(12345)$	$(13524)$	$(14253)$	$(15432)$
$I$	$I$	$(12345)$	$(13524)$	$(14253)$	$(15432)$
$(12345)$	$(12345)$	$(13524)$	$(14253)$	$(15432)$	$I$
$(13524)$	$(13524)$	$(14253)$	$(15432)$	$I$	$(12345)$
$(14253)$	$(14253)$	$(15432)$	$I$	$(12345)$	$(13524)$
$(15432)$	$(15432)$	$I$	$(12345)$	$(13524)$	$(14253)$

$H$	$I$	$(12345)$	$(13524)$	$(14253)$	$(15432)$
$I$	$I$	$(12345)$	$(13524)$	$(14253)$	$(15432)$
$(12345)$	$(12345)$	$(13524)$	$(14253)$	$(15432)$	$I$
$(13524)$	$(13524)$	$(14253)$	$(15432)$	$I$	$(12345)$
$(14253)$	$(14253)$	$(15432)$	$I$	$(12345)$	$(13524)$
$(15432)$	$(15432)$	$I$	$(12345)$	$(13524)$	$(14253)$

O  $\text{sim}(5)$  tem  $5!$  elementos e vai ser difícil colocá-los todos, de forma decente, numa página de texto. Como  $|H| = 5$  então há apenas  $4!$  classes laterais à direita. Calculando as classes laterais à direita de  $H$  em  $\text{sim}(5)$ :

✓	$I^*I$	(12345)	(13524)	(14253)	(15432)
✓	$I^*(21345)$	(14)(35)	(15423)	(243)	(125)
✓	$I^*(21354)$	(143)	(15)(23)	(245)	(12534)
repetida	$I^*(21354)$	(143)	(15)(23)	(245)	(12534)
✓	$I^*(21453)$	(154)	(234)	(12435)	(13)(25)
✓	$I^*(21534)$	(354)	(123)	(13245)	(14)(25)
✓	$I^*(12354)$	(13)(24)	(14325)	(152)	(345)
repetida	$I^*(13542)$	(143)	(15)(23)	(245)	(12534)
✓	$I^*(12)$	(1345)	(14)(352)	(153)(24)	(2543)
✓	$I^*(13)$	(145)(23)	(1524)	(2534)	(12)(354)
✓	$I^*(14)$	(15)(234)	(2435)	(1253)	(132)(45)
✓	$I^*(15)$	(2345)	(124)(35)	(13)(254)	(1432)
✓	$I^*(23)$	(1245)	(134)(25)	(142)(35)	(1543)
✓	$I^*(24)$	(125)(34)	(1352)	(1453)	(154)(23)
repetida	$I^*(25)$	(12)(345)	(1354)	(1423)	(15)(243)
repetida	$I^*(32)$	(1245)	(134)(25)	(142)(35)	(1543)
✓	$I^*(34)$	(1235)	(13)(245)	(14)(253)	(1542)
✓	$I^*(35)$	(123)(45)	(1324)	(1425)	(152)(34)
✓	$I^*(15423)$	(324)	(125)	(13452)	(14)(35)
✓	$I^*(12)(34)$	(135)	(14523)	(15324)	(2543)
	$I^*(13425)$	(14352)	(153)	(23)(45)	(124)

métodologia crie novo elemento, memorize no scilab troque acima pelo novo execute a linha no scilab e identifique as permutações copie no lixo e transforme numa coluna para organizar passe sort no lixo  $\xi$  lixo2 verifique lixo2 se há repetições, se não houver guarde no lixo coloque ✓ ou repetida, conforme apropriado. recomece o trabalho

$$e, 1 \quad (651)$$

$$, 0 \quad (652)$$

$$(13)(25), (13)(24), (14)(25), (14)(35), (15)(23), 5 \quad (653)$$

$$(12)(354), (14)(352), (145)(23), (153)(24), 4 \quad (654)$$

$$(123), (125), (143), (152), (154), (234), (243), (245), (345), (354), 10 \quad (655)$$

$$(1345), (2543), 2 \quad (656)$$

$$(12345), (12354), (12435), (12534), (13245), (13524), (14253), (14325), 6 \quad (657)$$

$$(15423), (15432), (21345), (21354), (21453), (21534), 6 \quad (658)$$

$$I = [1,0,0,0,0; 0,1,0,0,0; 0,0,1,0,0; 0,0,0,1,0; 0,0,0,0,1]$$

$$I_{12} = [0,1,0,0,0; 1,0,0,0,0; 0,0,1,0,0; 0,0,0,1,0; 0,0,0,0,1]$$

$$I_{13} = [0,0,1,0,0; 0,1,0,0,0; 1,0,0,0,0; 0,0,0,1,0; 0,0,0,0,1]$$

$$I_{14} = [0,0,0,1,0; 0,1,0,0,0; 0,0,1,0,0; 1,0,0,0,0; 0,0,0,0,1]$$

$$I_{15} = [0,0,0,0,1; 0,1,0,0,0; 0,0,1,0,0; 0,0,0,1,0; 1,0,0,0,0]$$

$I_{23} = [1,0,0,0,0;0,0,1,0,0;0,1,0,0,0;0,0,0,1,0; 0,0,0,0,1]$   
 $I_{24} = [1,0,0,0,0;0,0,0,1,0;0,0,1,0,0;0,1,0,0,0;0,0,0,0,1]$   
 $I_{25} = [1,0,0,0,0;0,0,0,0,1;0,0,1,0,0; 0,0,0,1,0;0,1,0,0,0]$   
 $I_{31} = [0,0,1,0,0; 0,1,0,0,0; 1,0,0,0,0; 0,0,0,1,0; 0,0,0,0,1]$   
 $I_{32} = [1,0,0,0,0; 0,0,1,0,0; 0,1,0,0,0; 0,0,0,1,0; 0,0,0,0,1]$   
 $I_{34} = [1,0,0,0,0; 0,1,0,0,0; 0,0,0,1,0; 0,0,1,0,0; 0,0,0,0,1]$   
 $I_{35} = [1,0,0,0,0; 0,1,0,0,0; 0,0,0,0,1; 0,0,0,1,0; 0,0,1,0,0]$   
 $I_{45} = [1,0,0,0,0; 0,1,0,0,0; 0,0,1,0,0; 0,0,0,0,1; 0,0,0,1,0]$   
 $I_{15423} = [0,0,1,0,0; 0,0,0,1,0; 0,1,0,0,0; 0,0,0,0,1; 1,0,0,0,0]$   
 $I_{324} = [1,0,0,0,0; 0,0,1,0,0; 0,0,0,1,0; 0,1,0,0,0; 0,0,0,0,1]$   
 $I_{12\_34} = [0,1,0,0,0;1,0,0,0,0;0,0,0,1,0;0,0,1,0,0;0,0,0,0,1]$   
 $I_{14\_35} = [0,0,0,1,0; 0,1,0,0,0; 0,0,0,0,1; 1,0,0,0,0; 0,0,1,0,0]$   
 $I_{243} = [1,0,0,0,0; 0,0,1,0,0; 0,0,0,1,0; 0,1,0,0,0; 0,0,0,0,1]$   
 $I_{125} = [0,0,0,0,1; 1,0,0,0,0; 0,0,1,0,0; 0,0,0,1,0; 0,1,0,0,0]$   
 $I_{153} = [0,0,1,0,0; 0,1,0,0,0; 0,0,0,0,1; 0,0,0,1,0; 1,0,0,0,0]$   
 $I_{23\_45}$

$I_{12345} = [0,0,0,0,1;1,0,0,0,0;0,1,0,0,0;0,0,1,0,0;0,0,0,1,0]$   
 $I_{12354} = [0,0,0,1,0;1,0,0,0,0;0,1,0,0,0;0,0,0,0,1;0,0,1,0,0]$   
 $I_{13425} = [0,0,0,0,1; 0,0,0,1,0; 1,0,0,0,0; 0,0,1,0,0; 0,1,0,0,0]$   
 $I_{13524} = [0,0,0,1,0;0,0,0,0,1;1,0,0,0,0;0,1,0,0,0;0,0,1,0,0]$   
 $I_{13542} = [0,1,0,0,0;0,0,0,1,0;1,0,0,0,0;0,0,0,0,1;0,0,1,0,0]$   
 $I_{14253} = [0,0,1,0,0;0,0,0,1,0;0,0,0,0,1;1,0,0,0,0;0,1,0,0,0]$   
 $I_{14352} = [0,1,0,0,0; 0,0,0,0,1; 0,0,0,1,0; 1,0,0,0,0; 0,0,1,0,0]$   
 $I_{15432} = [0,1,0,0,0;0,0,1,0,0;0,0,0,1,0;0,0,0,0,1;1,0,0,0,0]$   
 $I_{21345} = [0,1,0,0,0;0,0,0,0,1;1,0,0,0,0;0,0,1,0,0;0,0,0,1,0]$   
 $I_{21354} = [0,1,0,0,0;0,0,0,1,0;1,0,0,0,0;0,0,0,0,1;0,0,1,0,0]$   
 $I_{21435} = [0,1,0,0,0;0,0,0,0,1;0,0,0,1,0;1,0,0,0,0;0,0,1,0,0]$   
 $I_{21453} = [0,1,0,0,0;0,0,1,0,0;0,0,0,0,1;1,0,0,0,0;0,0,0,1,0]$   
 $I_{21534} = [0,1,0,0,0;0,0,0,1,0;0,0,0,0,1;0,0,1,0,0;1,0,0,0,0]$

Verificando se  $H$  é normal. Considerando  $g \in \text{sim}(5)$  vou calcular a classe à direita duma das classes à esquerda obtida na tabela 3 e tem que resultar numa das classes da referida tabela.

$$\overline{I_{13425}} = \{I_{13425}, I_{14352}, I_{153}, I_{23}I_{45}, I_{12}I_{35}\} \quad (659)$$

$$g = I_{14}I_{35}; \quad (660)$$

$$I_{14}I_{35} * \overline{I_{13425}} = \{I_{15423}, I_{245}, I_{134}, I_{14325}\} \quad (661)$$

$$I_{21345}, I_{14}I_{35}, I_{15423}, I_{243}, I_{125} \quad (662)$$

$$(663)$$

O código para *scilab* é

$I_{14\_35} * I_{13425}, \quad I_{14\_35} * I_{14352}, \quad I_{14\_35} * I_{153}, \quad I_{14\_35} * I_{23\_45}, I_{14\_35} * I_{124}$



e a classe à direita de  $I_15423$  sendo

$$\{I_{21345}, I_{14}I_{35}, I_{15423}, I_{243}, I_{125}\}$$

a conclusão é que  $H$  não é normal. Na verdade  $H$  é um grupo cíclico e a teoria dos grupos, que vale a pena estudar para evitar contas inúteis, nos ensina que se o conjugado do gerador não pertencer ao grupo então  $H$  não é normal:

$$H = [(12345)]; \quad (664)$$

$$I_{14}I_{35} * I_{12345} * I_{14}I_{35} = I_{13425} \quad (665)$$

$$(13425) \notin H; \quad (666)$$

que portanto gera um grupo diferente de  $H$  e aqui entra outra teorema da teoria dos grupos finitos, qualquer elemento dum grupo cíclico finito, é um gerador do grupo, portanto como  $(13425) \notin H$  então  $(13425)$  gera um grupo diferente de  $H$ .

4. Um exemplo de grupo normal Seria interessante que entre os exemplo visse um que não fosse um contra-exemplo, embora os contra-exemplos formem uma classe interessante de exemplos, inclusive ao mostrarem percalços que podem ser evitados com um estudo mais aprofundado da teoria...

**Teorema 8 (ordem de subgrupo)** grupo de ordem  $\frac{n}{2}$

Se ordem do grupo  $G$  for o número par  $n$ , qualquer subgrupo  $H$  de ordem  $\frac{n}{2}$  será normal.

**Dem**: Considere  $g_1, g_2 \notin H$  então as classes laterais à esquerda  $H, g_1H$  formam uma partição de  $G$ , assim como também as classes laterais à direita  $H, Hg_2$ . Mas

$$\text{card}(H) = \text{card}(g_1H) = \text{card}(Hg_2) \quad (667)$$

porque  $g_1, g_2$  são permutações dos elementos de  $G$  e qualquer duas dessas classes podem ser tomadas, duas a duas, como dois ciclos. Devido à cardinalidade, na equação (eq. 667) não resta outra saída que  $g_1H = Hg_2$  para quaisquer dois elementos de  $G$  que não pertençam à  $H$  logo

$$gH = Hg; g \notin H; \Rightarrow gHg^{-1} = H \quad (668)$$

Como a relação na equação (eq. 668) é trivial para  $g \in H$  então vale para todo elemento de  $G$  que é a definição de subgrupo normal.

**q.e.d.**

- **grupo quociente** Seja  $G, H$  um subgrupo de  $G$  elemento  $a \in G$ , notação:  $aG, Ga$  simbolizam as permutações de  $G$  obtidas pela multiplicação à esquerda ou à direita, respectivamente, pelo elemento  $\underline{a}$ .

- **grupo topológico** Se  $(G, *)$  for um grupo e  $(G, \tau)$  for um espaço topológico e se a operação binária  $*$  for contínua na topologia produto induzida por  $\tau$  em  $G \times G$  então diremos que  $(G, *, \tau)$  é um grupo topológico. Se não houver dúvida quanto a seleção da topologia e da operação binária, diremos simplesmente que

$G$  é um grupo topológico. Exemplos de grupos topológicos são os grupos lineares  $GL_n(K)$ ;  $K \in \{\mathbf{R}, \mathbf{C}\}$  a topologia natural sendo a induzida de  $K$  no grupo das matrizes. Como  $GL_n(K)$  é definido por uma desigualdade algébrica (determinante diferente de zero) então é um aberto do grupo das matrizes quadradas de ordem  $n$ . Como os grupos tem propriedades diferentes quando  $K = \mathbf{R}$  ou  $K = \mathbf{C}$ , vamos nos restringir ao caso  $K = \mathbf{R}$  na sequência, são duas teorias com diferenças importantes.

Um subgrupo de  $GL_n(\mathbf{R})$  é  $SL_n(\mathbf{R})$  definido como o conjunto das matrizes cujo determinante vale 1. Exemplos de subgrupos de  $SL_n(\mathbf{R})$  são obtidos considerando matrizes triangulares superiores ou inferiores cujos elementos da diagonal sejam todos 1.

- **HTML\*\*** É uma das 'markup languages'. Junto com TeX, Postscript, PDF e outras variantes destas, é uma das linguagens para automatizar a produção de textos. HTML está voltada para processar textos e imagens no vídeo com particular uso na Internet.

- **Hasse**, diagrama. É uma expressão gráfica da relação de inclusão entre conjuntos. Teoria dos conjuntos.

O conjunto  $A = \{a, e, i, o, u\}$  tem

$$C_5^0, C_5^1, C_5^2, C_5^3, C_5^4, C_5^5 \quad (669)$$

subconjuntos distribuídos pela quantidade de elementos e cada uma dessas quantidades indicadas com um número combinatório, porque *conjunto é combinação*.

Com um diagrama de Hasse posso exibir, neste caso, todos os subconjuntos de  $A = \{a, e, i, o, u\}$

$$\{e, i, o, u\}, \{a, i, o, u\}, \{a, e, o, u\}, \{a, e, i, u\}, \{a, e, i, o\} \quad (670)$$

$$\{i, o, u\}, \{e, o, u\}, \{e, i, u\}, \{e, i, o\}, \{a, o, u\} \quad (671)$$

$$\{a, i, u\}, \{a, i, o\}, \{a, e, u\}, \{a, e, o\}, \{a, e, i\} \quad \{a, e\}, \{a, i\}, \{a, o\}, \{a, u\}, \{a, o, u\} \quad (672)$$

$$\{e, o\}, \{e, u\}, \{i, o\}, \{i, u\}, \{o, u\} \quad (673)$$

$$\{a\}, \{e\}, \{i\}, \{o\}, \{u\} \quad (674)$$

$$\emptyset \quad (675)$$

que você pode ver na figura (fig 89), página 225,

- **Hausdorff, espaço de** Diz-se dum espaço topológico  $X$  em que dados dois pontos distintos,  $a, b \in X$  existem vizinhanças abertas  $\mathcal{V}(a)$  de  $a$ , e  $\mathcal{V}(b)$  de  $b$ , que são disjuntas. Ou ainda, podemos separar os pontos do espaço com conjuntos abertos disjuntos.

O conjunto dos números reais com sua topologia usual induzida pela função valor absoluto, é a topologia da norma, tem esta propriedade sendo assim um *espaço de Hausdorff*. O caso de  $\mathbf{R}$  se enquadra nos *espaços métricos*: se dois

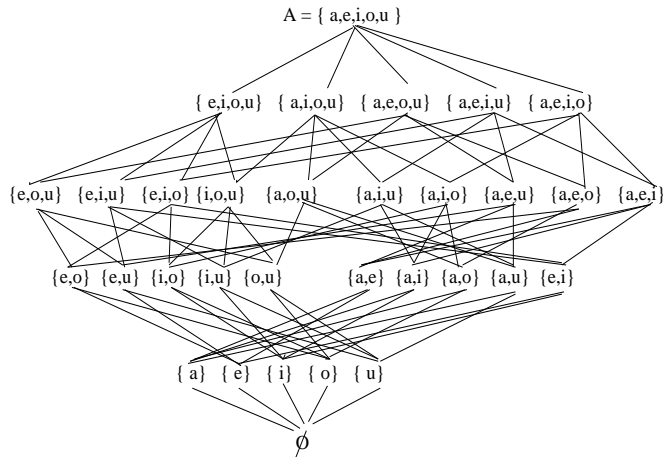


Figura 89: diagrama de Hasse

pontos,  $a, b$ , forem diferentes então  $r = d(a, b) \neq 0$  e  $\mathcal{B}(a, r/3), \mathcal{B}(b, r/3)$  são abertos que separam  $a, b$ . Se a métrica for discreta então

$$\mathcal{B}(a, r/3) = \{a\}; \mathcal{B}(b, r/3) = \{b\}$$

**Teorema 9 (Espaço métrico)** *espaço de Hausdorff* Todo espaço métrico é um espaço de Hausdorff. **Dem.**

Seja  $(E, d)$  um espaço métrico e tome dois pontos diferentes  $a, b \in E$  então  $d(a, b) = \epsilon \neq 0$  e as bolas  $\mathcal{B}(a, \epsilon/3)$  e  $\mathcal{B}(b, \epsilon/3)$  são disjuntas. **q.e.d.**

O espaço vetorial normado das funções contínuas definidas num intervalo fechado  $[a, b]$  da reta é um *espaço de Hausdorff*, confira *Stone-Weierstrass, teorema de*.

Esta propriedade garante que a topologia do espaço seja suficientemente rica em abertos e a demonstração, no caso do  $\mathcal{C}([a, b])$  é representante dum tipo particular de espaço, os espaços *vetoriais topológicos* em que as vizinhanças são obtidas por translação duma classe particular de vizinhanças, as vizinhanças da origem.

Este exemplo nos permite ver que os espaços  $\mathcal{C}([a, b])$  e  $\mathcal{C}((a, b))$  são profundamente diferentes e vai me permitir de mostrar a existência dum espaço vetorial topológico que não seja de Hausdorff.

Primeiro observe que  $\mathcal{C}([a, b]) \subset \mathcal{C}((a, b))$  e basta um exemplo para mostrá-lo, defina  $f$  identicamente nula exceto em nos pontos  $a, b$  onde  $f$  vale 1, é um exemplo de função contínua em  $(a, b)$  que não pertence a  $\mathcal{C}([a, b])$ .

Se você estiver registrando um defeito lógico no parágrafo acima, tem razão, mas ele pode ser sanado uma vez que é um erro de linguagem comum, considere  $\mathcal{C}([a, b]), \mathcal{C}((a, b))$  duas “etiquetas” representando dois subconjuntos de todas as funções reais e agora vale a inclusão  $\mathcal{C}([a, b]) \subset \mathcal{C}((a, b))$ .

Considere a *topologia usual* de  $\mathcal{C}([a, b])$  e defina os abertos de  $\mathcal{C}((a, b))$  como sendo as imagens dos abertos de  $\mathcal{C}([a, b])$ . Esta é uma forma de definir uma topologia num superconjunto: a chamada *topologia induzida*. O resultado, neste caso, será uma *topologia fraca no super espaço*

Por exemplo, as vizinhanças da origem em  $\mathcal{C}((a, b))$  serão as vizinhanças tubulares de raio  $\epsilon$  apenas nelas estão funções como  $\epsilon f$ :

$$\mathcal{C}((a, b)) \ni \epsilon f \notin \mathcal{C}([a, b]);$$

Tome  $g(x) \equiv 0; g \in \mathcal{C}([a, b])$  a função identicamente nula e temos:

$$\mathcal{C}((a, b)) \ni g \in \mathcal{C}([a, b]);$$

Qualquer vizinhança tubular de  $g$  em  $\mathcal{C}((a, b))$  com a topologia induzida do subespaço, contém  $f$ , não sendo possível separar  $f, g$  com a topologia induzida em  $\mathcal{C}((a, b))$  pela topologia do subespaço  $\mathcal{C}([a, b])$ .

Para conseguir separar  $f, g$  seria preciso acrescentar mais abertos na topologia de  $\mathcal{C}((a, b))$  e isto se faz considerando a topologia *tubular-compacta*, também designada como *compacto-aberta* em que os tubos em volta da origem correspondem aos compactos de  $(a, b)$  e agora será possível encontrar uma vizinhança de  $g$  e outra de  $f$  que sejam disjuntas. Basta considerar “tubos” de largura inferior a  $\frac{1}{2}$ .

Como  $\mathbf{R}$  é topologicamente equivalente a  $(a, b)$  então é possível compreender que este exemplo não é absurdo. A topologia da convergência uniforme seria inútil em  $\mathcal{C}(\mathbf{R})$ , por exemplo, neste espaço, a integral de Riemann não é contínua com a topologia da convergência uniforme, mas é contínua com a *topologia tubular-compacta* e assim definimos uma topologia em  $\mathcal{C}(\mathbf{R})$  que salva um instrumento que nos é muito caro, a integral de Riemann . . .

---

- **Hilbert, espaço de**  $\hat{E}$  é a generalização mais imediata dos espaços de dimensão finita,  $n$ . A redação que vou dar se aplica *ipsis literis* aos casos  $\mathbf{R}^n, \mathbf{C}^n$  com pequenas exceções que vou indicar entre parentesis no local.

Seja  $E$  é um espaço vetorial com *produto escalar*, conseqüentemente o módulo de um vetor é dado pela expressão  $\|x\| = \sqrt{\langle x, x \rangle}$  em que  $\langle, \rangle$  representa o produto escalar. No caso complexo a expressão deve ser  $\|x\| = \sqrt{\langle x, \bar{x} \rangle}$ .

**Definição 13** *espaço de Hilbert* O espaço  $E$  com produto escalar é um espaço Hilbert se for completo.

Num *espaço completo* os dois conceitos, *sucessão de Cauchy* e *sucessão convergente*, coincidem.

Observe que como definimos um módulo podemos imediatamente aplicar a definição de continuidade usada no Cálculo e o módulo é uma função contínua. Um teorema importante dos espaços de Hilbert é a desigualdade de Cauchy-Bunyakovsky-Schwarz

**Teorema** 10 *Cauchy-Bunyakovsky-Schwarz*

$$|\langle x, y \rangle| \leq \|x\| \|y\|$$

**Dem**:

Como esta desigualdade envolve apenas dois vetores, ela se situa num subespaço de dimensão 2, num plano do espaço considerado. Vou usar esta facilidade para construir uma demonstração simples. Vou usar um caso particular porque vou seleccionar um tipo de produto escalar na demonstração. Um produto escalar é uma forma bilinear que é positiva definida. As formas bilineares definidas em espaços vectoriais de dimensão finita podem ser representadas por um produto matricial:

$$\langle x, y \rangle = xAy^t;$$

em que  $A$  é uma matriz tal que  $xAx^t \geq 0$  sendo zero apenas quando  $x = 0$ . Como o problema se reduz à dimensão dois, uma tal matriz é da forma

$$A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}; \lambda_1, \lambda_2 > 0 \quad (676)$$

e eu vou usar  $\lambda_1 = \lambda_2 = 1$ .

Confira a demonstração geral da desigualdade de Cauchy-Bunyakovsky-Schwarz para uma demonstração geral.

Para isto preciso mostrar que a forma trigonométrica da definição do produto escalar é idêntica à expressão da soma de produtos na mesma equação porque

$$|\langle x, y \rangle| = \|x\|\|y\| \cos(\gamma) \leq \|x\|\|y\|;$$

Se  $n = 2$  é consequência direta da fórmula de Euler porque, dividindo pelo módulo dos vetores teremos:

$$\left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle = \frac{|\langle x, y \rangle|}{\|x\|\|y\|} \leq 1; \quad (677)$$

$$\left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle = \langle (\cos(\alpha), \sin(\alpha)), (\cos(\beta), \sin(\beta)) \rangle; \quad (678)$$

$$\left\langle \frac{x}{\|x\|}, \frac{y}{\|y\|} \right\rangle = \cos(\alpha - \beta) = \cos(\gamma); \quad (679)$$

Na equação (eq. 677) usei a propriedade das formas bilineares que permite a distribuição dos coeficientes  $\|x\|, \|y\|$  entre os dois fatores<sup>34</sup>.

Na equação (eq. 678) usei uma representação plana dos dois vetores que podem ser expressos como pontos do círculo trigonométrico sendo  $\alpha, \beta$  os seus ângulos e  $\gamma$  é o ângulo entre eles.

Os cálculos mostram que posso me reduzir, na demonstração, ao caso dos vetores unitários. Então, na sequência vou supor que  $\|x\| = \|y\| = 1$ .

A passagem crítica, na demonstração por indução que está na passagem de  $n - 1$  para  $n$  que é semelhante à passagem da dimensão 2 para 3, mas esta é mais intuitiva. Portanto deixeme agora considerar  $x, y \in \mathbf{R}^3$ , vetores unitários, portanto pontos de  $\mathbf{S}^2$  a esfera unitária do  $\mathbf{R}^3$ . Mas estes dois vetores determinam um plano que corta  $\mathbf{S}^2$  segundo um círculo unitário, uma curva plana, e, pela demonstração anterior,  $\cos(\gamma)$  é o valor do produto escalar entre eles, em que  $\gamma$  é ângulo entre os vetores.

Mas esta demonstração não seria necessária dentro do escopo duma demonstração por indução, ela é simplesmente tomada como verdadeira para todo  $N \leq n - 1$  e temos que considerar agora  $x, y \in \mathbf{R}^n$ , dois vetores unitários. Mas estes vetores definem um plano do  $\mathbf{R}^n$  e o valor do produto escalar para dois vetores, no plano, em dimensão  $N < n$ , por hipótese de indução, é  $\cos(\gamma)$ , em que  $\gamma$  é ângulo entre os vetores, o que termina a demonstração.

O uso dos espaços  $\mathbf{R}^n$  se justifica porque estou trabalhando em um subespaço de dimensão finita do espaço  $E$  e todos os espaços de dimensão finita são isomorfos a  $\mathbf{R}^n$  ou a  $K^n$  em que  $K$  é o corpo sobre o qual o espaço  $E$  estiver definido.

**q.e.d.**

No caso complexo é preciso colocar conjugado na expressão do produto escalar.

<sup>34</sup>as formas bilineares generalizam o produto de números que um exemplo de forma bilinear.

É um corolário que o produto escalar é contínuo. Todos os teoremas envolvendo módulo e produto escalar da Álgebra Linear ou da Geometria Euclidiana valem nos Espaços de Hilbert.

Vou dar um exemplo da dimensão não finita,  $E = \mathcal{C}([a, b])$ , o espaço vetorial das funções contínuas, no qual podemos definir o produto escalar (627)

$$\langle f, g \rangle = \int_a^b f(x)g(x)dx \quad (680)$$

sendo necessário usar conjugação em um dos termos dentro da integral, no caso complexo.

Neste ponto, se você for crítica, poderá registrar um defeito na exposição<sup>35</sup> que vou deixar claro mais a frente.

Com isto você pode repetir os teoremas da Geometria Euclidiana usando funções, em particular pode definir o ângulo entre duas funções:

$$\text{angulo}(f, g) = \text{Acos}\left(\frac{\langle f, g \rangle}{|f||g|}\right) \quad (681)$$

e aqui um exemplo de duas funções ortogonais:

$$[a, b] = [-1, 1] \quad (682)$$

$$\begin{cases} x \leq 0 & f(x) = 0; \\ x > 0 & f(x) = 1; \end{cases} \quad (683)$$

$$\begin{cases} x \leq 0 & g(x) = 1; \\ x > 0 & g(x) = 0; \end{cases} \quad (684)$$

dois sinais que contém informações independentes:  $f \perp g$ , ou, um exemplo do Cálculo,

$$\sin(x), \cos(x), \sin(2x), \cos(2x), \dots, \sin(nx), \cos(nx) \quad (685)$$

para quaisquer valores de  $n$  estes vetores são ortogonais entre si quando considerarmos como elementos do espaço  $E = \mathcal{C}([- \pi, \pi])$  definindo o produto escalar como

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx \quad (686)$$

No caso complexo é preciso usar complexo conjugado em dos elementos no integrando. Esta definição tem um defeito estético: os vetores não tem módulo 1 e a Álgebra Linear prefere que os vetores básicos tenham módulo 1, e estes vetores geram um espaço vetorial (dos polinômios trigonométricos, ou das séries de Fourier). Colocar uma constante positiva multiplicativa não altera as propriedades do produto escalar que é uma forma bilinear, então podemos multiplicá-lo

<sup>35</sup>E você sempre deve ser crítica e procurar defeitos nas exposições, mas ser também *decente*, e informar sobre os erros encontrados.

por uma constante que torne um desses vetores unitários:

$$f(x) = \sin(x); \quad \langle f, f \rangle = A \int_{-\pi}^{\pi} f(x)f(x)dx = A\pi \Rightarrow A = \frac{1}{\pi} \quad (687)$$

e você pode agora ler a teoria das séries de Fourier dentro dum espaço de Hilbert<sup>36</sup> (que contém o espaço das funções contínuas) sendo os coeficientes de Fourier apenas a projeção de uma função qualquer na direção dos vetores básicos.

O interessante é que podemos rapidamente colocar estes exemplos num quadro bem mais geral. Em todos os dois casos temos “integrais” definindo o produto escalar, porque uma soma finita é um tipo de integral, assim como as séries o são. O que caracteriza uma integral é uma forma de medir conjuntos, se tivermos uma função “medida” num certo conjunto  $X$ , podemos nele definir as integrais das funções reais (ou complexas) nele definidas e com isto ter um espaço com produto escalar. Mas esta teoria é muito rica para ficar contida num verbete de dicionário! Apenas para que você se divirta com com o que pode acontecer, observe que, por precaução, usei o conjunto  $E = \mathcal{C}([a, b])$  e não  $E = \mathcal{C}((a, b)) \dots$

O espaço de Hilbert mencionado é  $\mathcal{L}^2([a, b])$ .

#### **- Hilbert, espaço de - reproduzindo kernel**

Esta classe de espaços recentemente se tornou muito importante devido a certas aplicações que se puderam fazer com eles. *kernel* é o nome que se usa para as funções que ocupam o lugar das matrizes quando usamos integrais para definir funções lineares:

$$J(f)(y) = \int_E f(x)K(x, y)d\mu(x) = \phi(y); \phi = J(f) \quad (688)$$

define um operador linear entre dois espaços de funções definidas no (espaço de medida  $(E, \mu)$ ). As séries de Fourier usam o kernel  $K(n, x) = e^{2\pi nx}$  para transformar funções definidas em  $E = [-\pi, \pi]$  em funções definidas em  $\mathbf{N}$  os coeficientes de Fourier de uma função definida em  $E = [-\pi, \pi]$  portanto o operador linear definido na equação (688) pode transformar elementos de espaços bem diferentes.

Os espaços de Hilbert tem uma propriedade importante que é o teorema da representação de Riesz, caracterizando os funcionais lineares num espaço. Para os espaços de Hilbert o teorema retorna exatamente o produto escalar como equação para os funcionais. Considere um conjunto  $(E, \mu)$  um espaço de medida e conjunto das funções complexas definidas em  $E$  e temos de forma

---

<sup>36</sup>Eis o defeito, na verdade quando mencionei o exemplo do espaço das funções contínua usei uma expressão para definir produto escalar, a integral, que não precisa da continuidade para ser definida. Eis a correção...

canônica um espaço com produto escalar definido por

$$\langle f, g \rangle = \int_E f(x) \overline{g(x)} d\mu(x) \quad (689)$$

O espaço de Hilbert associado a este espaço com produto escalar é designado por  $L^2(E)$ . O funcional *avaliação*  $\delta_x; x \in E$ , pelo *teorema da representação de Riesz*, é dado pelo produto escalar com um dos elementos de  $L^2(E)$ :

$$\delta_x = \langle \cdot; f \rangle; f \in L^2(E); \quad (690)$$

Considere agora um operador linear  $J$  definido em  $L^2(E)$  por um núcleo  $K$  como na expressão da equação (688):

$$J(f) = \int_E f(x) K(x, y) d\mu(x) = \phi(y); \phi = J(f); \quad (691)$$

- **hipótese** É uma afirmação que se julga verdadeira e que não é possível encontrar uma demonstração a partir de outras hipóteses (axiomas) de uma teoria. A *geometria euclidiana* é considerada um das mais antigas construções lógicas estruturada como um conjunto de axiomas e os teoremas que podem ser deduzidos a partir dos axiomas. As hipóteses são afirmações que se espera poder demonstrar. Conjectura é uma classe de hipótese, mas as nem sempre se consideram estas duas palavras como sinônimas.

- **hipótese de Cantor** Georg Cantor é possivelmente um dos que melhor formulou a teoria dos Conjuntos que é considerada o principal fundamento da Matemática. Em seus estudos ele construiu a cardinalidade que é uma generalização do *princípio da contagem* com o qual se constrói o conjunto dos números naturais,  $\mathbf{N}$ . Todo número natural tem um sucessor (axioma de Peano) mas não podemos falar da “quantidade de elementos do conjunto dos naturais”, e Cantor inventou o conceito de cardinalidade para resolver este problema. Ao fazê-lo descobriu que haviam classes de conjuntos que se agrupam por cardinalidade mas que a cardinalidade dá saltos estabelecendo então a sua hipótese: *não há conjuntos com cardinalidade intermediária entre a cardinalidade de  $\mathbf{N} = \aleph_0$  e a cardinalidade de  $\mathbf{R} = c$  ou  $\aleph_1$* , o salto de cardinalidade, o primeiro dos 23 problemas de Hilbert. Em 1963 Cohen provou que não é possível provar a hipótese de Cantor a partir dos axiomas de *Zermelo-Fraenkel*, adotados para a teoria dos conjunto o que a tranforma em novo axioma.

- **Hölder, desigualdade** é uma generalização da desigualdade de Bounjakowsky-Cauchy-Schwarz definida em espaços vetoriais de sucessões ou funções integráveis:

$$\sum_{k=0}^{\infty} |a_k b_k| \leq \|a\|_p \|b\|_q; \frac{1}{p} + \frac{1}{q} = 1; \quad (692)$$

$$\int_{-\infty}^{\infty} |f(x)g(x)| dx \leq \|f\|_p \|g\|_q; \frac{1}{p} + \frac{1}{q} = 1; \quad (693)$$



Na desigualdade de Bounjakowsky-Cauchy-Schwarz  $p = q = 2$ .

A desigualdade na equação (693) vale para funções integráveis em intervalos limitados  $[a, b]$ . O caso  $(a, b)$  é equivalente ao caso **R**. Esta desigualdade segue sendo válidas se acrescentarmos pesos, para que as séries sejam convergentes, em particular se consideramos espaços de medida  $(X, \mu)$ , alias, uma soma é um tipo de integral!

Uma consequência importante da desigualdade de Hölder é que o espaço  $L^p$  se transforma num conjunto de funcionais lineares contínuos definidos no espaço  $L^q$  ou ainda, o dual do espaço  $L^p$  é o espaço  $L^q$ , com  $\frac{1}{p} + \frac{1}{q} = 1$  sendo o *operador multiplicação* a representação destes funcionais. O mesmo pode ser dito para os espaços de sucessões.

- **holomorfa, função** Uma função complexa é dita holomorfa, ou analítica, se satisfizer às equações de Cauchy-Riemann. O conjunto das funções analíticas num disco aberto do plano complexo tem a estrutura de anel com a soma e o produto de funções definido ponto a ponto e de espaço vetorial complexo. Elas são o conjunto solução das equações de Cauchy-Riemann que formam um sistema linear de equações diferenciais parciais. Confira *analítica, função*.

Uma das formas de definir uma função holomorfa é via *integral de Cauchy*

$$f(a) \text{Ind}_\gamma(a) = \frac{1}{2\pi i} \int_\gamma \frac{f(t)dt}{t-a}; \quad (694)$$

em que podemos identificar um *produto por convolução* de  $f$  pelo núcleo de Cauchy. Confira também *Cauchy, teorema*.

- **ideal** é uma subestrutura de um anel. Seja  $(A, +, *)$  um anel. Um subconjunto  $I$  de  $A$  é um ideal se  $(I, +)$  for um subgrupo de  $(A, +)$  e se

1.  $\forall x \in R; xI \subset I$ , um ideal à esquerda;
2.  $\forall x \in R; Ix \subset I$ , um ideal à direita;

O anel  $(A, +, *)$  não precisa ser comutativo o que faz com que as duas propriedade não precisem ser simultaneamente verdadeiras. Se o forem dizemos que  $I$  é um ideal bilateral.

Um caso muito comum são os conjuntos dos múltiplos de um inteiro positivo (não vale a pena considerar positivos e negativos...)  $2Z, 3Z, 4Z, \dots$ . Se o inteiro multiplicador for primo o ideal assim gerado não estará contido em nenhum outro e portanto é um *maximal na relação de ordem* da inclusão. Como os ideais são subgrupos então vale toda a teoria de subgrupos para os ideais, em particular que eles definem *classes quociente* descritas pela translações:  $aI + r; r \in A$ . No caso dos inteiros estas classe são em número finito e correspondem aos restos na divisão por  $\underline{a}$ . Se  $\underline{a}$  não for primo, então há pelo menos dois restos cujo produto é  $\underline{a}$  e como  $\underline{a}$  corresponde à classe nula, entre as classes quocientes, vemos que existem *divisores de zero* neste caso. Se  $\underline{a}$  for primo, o *anel quociente* não terá

divisores de zero sendo um corpo finito designado por  $Z/aZ$  ou mais comumente como  $Z_a$ .

Esta breve discussão mostra que é possível rerepresentar toda a aritmética usando o conceito de ideal e este é o método para generalizar a aritmética para um anel qualquer.

O caso mais simples que se segue aos inteiros é o dos polinômios a uma variável sobre o corpo dos reais,  $(\mathbf{R}[x], +, *)$ . Se um polinômio  $P$  não puder ser fatorado (sobre o corpo dos reais) ele é um polinômio primo e o ideal que ele gera,  $P\mathbf{R}[x]$  dos múltiplos por  $P$  é um ideal maximal.  $P(x) = 1 + x^2$  é um exemplo de polinômio primo (sobre o corpo dos reais) e os restos na divisão por  $P$  serão da forma  $(ad + bc)x + (bc - ac)$  em que podemos reconhecer como o resultado da multiplicação dos números complexos  $ai + b$  por  $ci + d$ , ou seja, as classes quocientes módulo  $P$  são isomorfas (como corpo) ao corpo dos números complexos. O ideal gerado por  $P$  é maximal.

### - implícita, teorema da função

É um teorema de existência!

Considere uma equação como

$$z = F(x, y) \quad (695)$$

em que  $F$  é uma função diferenciável pelo menos uma vez continuamente (tem derivadas parciais contínuas) em um domínio do plano. É possível derivar implicitamente a equação (695) para obter

$$dz = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy \quad (696)$$

A equação (696) é uma soma porque

1. a derivada de uma função, num ponto, é uma *função linear tangente* cuja matriz é a derivada  $J(F) = \left( \frac{\partial F}{\partial x} \quad \frac{\partial F}{\partial y} \right)$  chamada de *jacobiana* de  $F$ , a matriz das derivadas parciais calculadas em um ponto do domínio<sup>37</sup>;
2.  $dx, dy$  são duas novas variáveis, alguns autores caracterizam isto usando  $h, k$  em vez de  $dx, dy$ , e a expressão na equação (695) mostra a matriz aplicada a estas variáveis.

A equação (696) é um modelo que pode nos conduzir á equação da variedade linear tangente à variedade  $z = F(x, y)$  em um ponto  $(a, b, c) = (a, b, F(a, b))$  que seja conhecido do gráfico desta variedade com auxílio das substituições

$$dz := z - c; dx = x - a; dy = y - b; \quad (697)$$

$$A = \frac{\partial F}{\partial x}(a, b); B = \frac{\partial F}{\partial y}(a, b); \quad (698)$$

$$z - c = A(x - a) + B(y - b); \quad (699)$$

<sup>37</sup>No caso univariado vale a mesma explicação, mas a simplicidade deste caso esconde a matriz que parece ser um número. A derivada não exibe a variedade linear tangente, ela é um modelo que nos permite descobrir a variedade linear tangente a partir de uma função linear que este modelo constrói.

Se, na equação (699) eliminarmos a variável  $z$  considerando o seu valor constante,  $z = c$ , a dimensão desta variedade cai de uma unidade. Enquanto a variedade na equação (695) é uma variedade diferenciável de dimensão dois, uma superfície, com esta substituição temos uma variedade diferenciável de dimensão 1, uma curva, e a variedade linear tangente, também de dimensão 1, será uma reta.

Quando as retas não forem paralelas aos eixos, na notação padrão cartesianda que usamos, elas tem coeficiente angular e podem ser escritas com uma das alternativas, sempre coerente com a notação acima, uma reta que passe pelo ponto  $(a, b)$

$$dx = x - a; dy = y - b; \quad (700)$$

$$A = \frac{\partial F}{\partial x}(a, b); B = \frac{\partial F}{\partial y}(a, b); \quad (701)$$

$$A(x - a) + B(y - b) = 0; \quad (702)$$

$$y = -\frac{A}{B}(x - a); \quad (703)$$

$$x = -\frac{B}{A}(y - b); \quad (704)$$

Como é uma reta tangente a uma curva, com a hipótese de que a reta não seja paralela a nenhum dos eixos<sup>38</sup> então a curva que tem esta por tangente definida por  $F(x, y) = c$ , é o gráfico de  $y = g(x)$  no caso da equação (703), ou é o gráfico de  $x = h(y)$  no caso da equação (704).

No primeiro caso, o coeficiente angular da reta é a derivada de  $y = g(x)$  no ponto  $x = a$ , e podemos escrever

$$g'(a) = -\frac{A}{B} = -\frac{\frac{\partial F}{\partial x}(a, b)}{\frac{\partial F}{\partial y}(a, b)}; \quad (705)$$

No segundo caso, o coeficiente angular da reta é a derivada de  $x = h(y)$  no ponto  $y = b$ , e podemos escrever

$$h'(b) = -\frac{B}{A} = -\frac{\frac{\partial F}{\partial y}(a, b)}{\frac{\partial F}{\partial x}(a, b)}; \quad (706)$$

Embora não seja possível, em geral, encontrar uma equação para  $y = g(x)$  ou para  $x = h(y)$ , mas a existência de um gráfico com uma tangente como se pode ver na figura (90), página 234, não paralela aos eixos garante a existência de uma função com as derivadas calculadas acima numa vizinhança do ponto  $(a, b)$ . Se apenas uma das derivadas for diferente de zero, o teorema ainda se aplica com existência de apenas  $y = g(x)$ , quando  $\frac{\partial F}{\partial y}(a, b) \neq 0$  mas com  $g'(a) = 0$ , porque neste caso a reta é paralela ao eixo  $OX$ , ou no caso recíproco, existe  $x = h(y)$ , com  $h'(b) = 0$  uma vez que a reta tangente é paralela ao eixo  $OY$  o que elimina a restrição de que a reta não possa ser paralela a um dos eixos.

Com a notação do caso bivariado, o *teorema da função implícita* tem a seguinte redação

---

<sup>38</sup>Eu vou retirar esta restrição mais a frente.

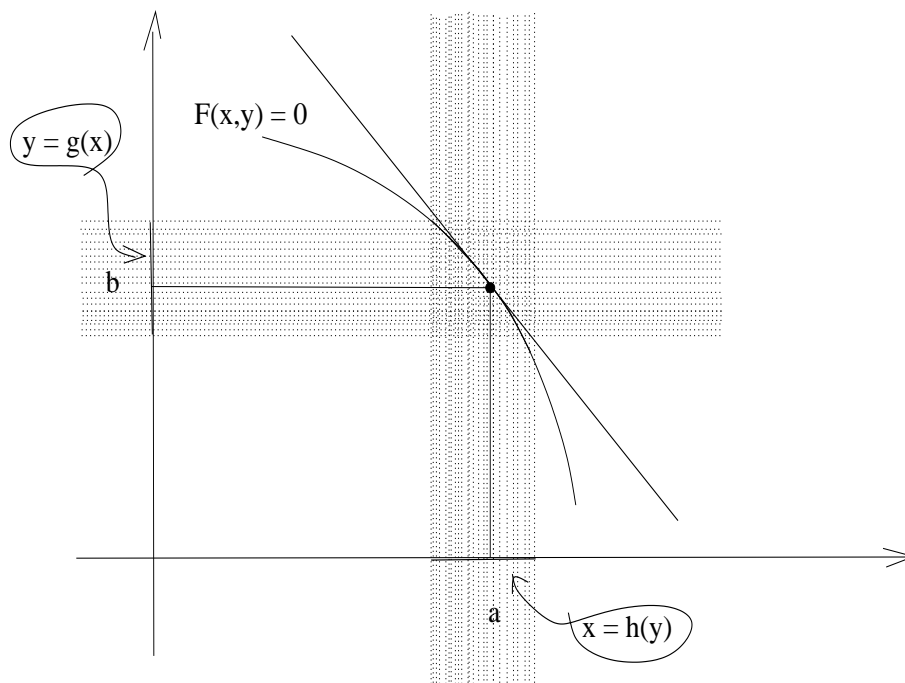


Figura 90: teorema da função implícita

**Teorema 11 (função implícita, teorema da )** *Teorema da função implícita*

Se  $y = F(x, y)$  for uma função continuamente derivável numa vizinhança do ponto  $(a, b)$  do seu domínio, então

- se  $\frac{\partial F}{\partial y}(a, b) \neq 0$  então existe uma função  $y = g(x)$ , definida numa vizinhança do ponto  $x = a$  tal que nesta vizinhança  $z = F(x, g(x))$  e

$$g'(a) = -\frac{\frac{\partial F}{\partial x}(a, b)}{\frac{\partial F}{\partial y}(a, b)};$$

- se  $\frac{\partial F}{\partial x}(a, b) \neq 0$  então existe uma função  $x = h(y)$ , definida numa vizinhança do ponto  $y = b$  tal que nesta vizinhança  $z = F(h(y), y)$  e

$$h'(b) = -\frac{\frac{\partial F}{\partial y}(a, b)}{\frac{\partial F}{\partial x}(a, b)};$$

- **índice** de um ponto  $\underline{a}$  relativamente a uma curva  $\gamma$  é uma função inteira que conta o número de voltas que uma urva der em torno de um ponto. Este número corresponde ao grau de  $\underline{a}$  como raiz de uma função complexa mas tem outras

aplicações em topologia diferencial e algébrica. Na teoria das funções analíticas o índice aparece na formula integral de Cauchy, teorema de Cauchy.

A definição do índice, na teoria das funções analíticas é

$$Ind_{\gamma}(a) = \frac{1}{2\pi i} \int_{\gamma} \frac{dt}{t-a} \quad (707)$$

O coeficiente  $\frac{1}{2\pi i}$  é uma consequência da homotopia, ou falando pedantemente, da “conjectura”<sup>39</sup> de Poincaré, pela qual todas as curvas que tenham índice 1 relativamente a todos os seus pontos interiores é *homotopicamente equivalente a um círculo* e o círculo trigonométrico tem por medida (perímetro)  $2\pi$ .

Na integral de Cauchy temos:

$$f(a)Ind_{\gamma}(a) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(t)dt}{t-a}; \quad (708)$$

em que podemos ver no segundo membro uma deformação da equação do índice com a presença da função  $f$  no numerador. Uma função complexa, como é o caso na integral de Cauchy, é uma deformação do plano que pode aumentar localmente o índice de um ponto relativamente a uma curva para caracterizar o grau de  $\underline{a}$  como raiz de  $f$ .

Observe o caso particular

$$\frac{1}{2\pi i} \int_{\gamma} \frac{dt}{t}; \quad (709)$$

quando 0 estiver no interior de  $\gamma$  e  $Ind_{\gamma}(a) = 1$  que irá resultar em  $2\pi$  que é o salto do logaritmo em volta da origem. Como  $\underline{a}$  tem índice 1 relativamente à  $\gamma$ <sup>40</sup> então o cálculo na equação (eq. 709) é equivalente a

$$\frac{1}{2\pi i} \int_{S^1} \frac{dt}{t} = \frac{1}{2\pi i} \int_0^{2\pi} \frac{ie^{it}dt}{e^{it}} = 2\pi \quad (710)$$

Esta é uma forma de justificar o coeficiente  $\frac{1}{2\pi i}$  na definição do índice.

- **indução finita** é um método para conduzir demonstrações baseado no conjunto dos números naturais. Para isto precisamos de uma afirmação  $P(k)$  que represente o teorema sob forma de uma expressão que dependa de um número natural  $k$ , por exemplo,

$$P(k) := \sqrt[k]{a_1 \cdots a_k} \leq \frac{1}{k} \sum_{j=1}^k a_j \quad (711)$$

<sup>39</sup>Demonstrada em 2003 por Perelman.

<sup>40</sup>embora a curva  $\gamma$  não precise ser homotópica a um círculo.

ou ainda que “a media geométrica é menor do que a média aritmética de  $k$  números positivos dados. Podemos facilmente provar que  $P(2)$  é verdadeira (como também  $P(1)$ ) e assim estabelecer a *hipótese de indução*: “a relação expressa na equação (711) é verdadeira”.

Se conseguirmos provar que  $P(k) \Rightarrow P(k+1)$ , então, pelo *teorema da indução finita*,  $P(n)$  é verdadeira para qualquer número natural  $n$ . A implicação, no exemplo acima, se obtém facilmente usando as propriedades do logaritmo, uma função convexa crescente.

Assim o método da *indução finita* tem duas etapas:

1. A demonstração da expressão para um valor especial, aqui no exemplo, mencionamos  $k = 2$ ;
2. A demonstração do encadeamento indutivo, a implicação

$$P(k) \Rightarrow P(k+1)$$

;

Em geral esta é a grande demonstração, o encadeamento indutivo. No exemplo, usamos um teorema difícil, que o logaritmo é crescente e convexo, para “terminar” a demonstração da desigualdade aritmético-geométrica.

Algumas fórmulas que podem ser provadas com indução finita.

$$1 + 2 + \cdots + n = \frac{n+1}{2}n \quad (712)$$

$$1 + 4 + \cdots + n^2 = \frac{n(n+1)(2n+1)}{6} \quad (713)$$

$$\sum_{k=1}^n k^3 = (1 + 2 + \cdots + n)^2 \quad (714)$$

$$\sum_{k=1}^n k^3 = \left(\frac{n+1}{2}n\right)^2 = \frac{(n+1)^2 n^2}{4} \quad (715)$$

$$\sum_{k=0}^{n-1} k^4 = \frac{6n^5 - 15n^4 + 10n^3 - n}{30} \quad (716)$$

$$\sum_{k=0}^{n-1} k^5 = \frac{2n^6 - 6n^5 + 5n^4 - n^2}{12} \quad (717)$$

$$\text{Soma dos ângulos num polígono com } n \text{ lados : } (n-2)\pi; \quad (718)$$

Se chamarmos de *progressão de grau  $m$*  a uma expressão  $P(k)$  que se pode expressar como um polinômio do grau  $\underline{m}$ , então “a soma dos termos de uma progressão de grau  $m$  é uma progressão de grau  $m+1$ ” é um exemplo de teorema que pode ser demonstrado por *indução finita*.

Observe que nas equações à direita esta a soma dum polinômio de grau  $p$  e a fórmula com que se obtém a soma é um polinômio do grau  $p+1$ . Por exemplo, a equação (712) é a soma dos termos dum polinômio do primeiro grau e a soma se obtém com um polinômio do segundo grau à direita da igualdade. Na equação (713), à esquerda se tem a soma dum polinômio do segundo grau cuja soma se obtém com um polinômio do terceiro grau.

É este fato que justifica a fórmula de integração de funções polinomiais que você pode encontrar em [5, volume 1], em que a integral dum polinômio de grau  $p$  é dada por um polinômio do grau  $p + 1$  como limite de *somas de Riemann* de expressões do grau  $p$ .

- **integral** Há várias formas de entender a *integral*, a mais elementar é como uma área. Se  $f$  for uma *função univariada integrável no sentido de Riemann* então o símbolo

$$\int_a^b f(x)dx \quad (719)$$

representa a área limitada pelo gráfico de  $f$ , pelo eixo OX desde o ponto  $a$  até o ponto  $b$ , quer dizer que é uma área algébrica porque o sinal muda se ela for calculada de  $b$  para  $a$ . Na figura (91) página 237, você pode ver a interpretação

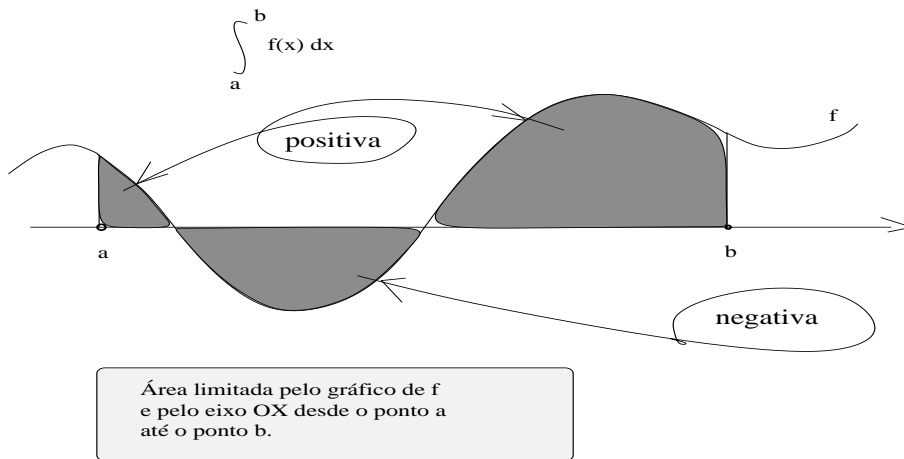


Figura 91:

geométrica da integral de  $f$  com a indicação de que algumas das “subáreas” são positivas ou negativas. A razão do sinal é que a integral, mesmo interpretada como área, é uma quantidade de um determinado fenômeno, quantidade calor, quantidade de movimento, distância percorrida dependendo do significado de  $f$ . Se  $f$  representar a velocidade com que um corpo se move, temos a distância percorrida pelo corpo entre os pontos  $a$  e  $b$ . A função  $f(x) = \frac{1}{x}$  não é integrável à Riemann se o ponto  $x = 0$  pertencer ao intervalo de integração, mas é possível definir uma “*medida adequada*” na reta que torne esta função integrável em qualquer subintervalo da reta. O conceito *primitiva* está associado à integral como inversa da derivada.

- **integral de Cauchy** Confira *Cauchy, teorema de, analítica, função, holo-*

*morfa, função*. Se identifica como *integral de Cauchy* a expressão

$$f(a)Ind_{\gamma}(a) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(t)dt}{t-a}; \quad (720)$$

em que podemos identificar um *produto por convolução* de  $f$  pelo núcleo de Cauchy.

Devido à regularização por convolução esta expressão define uma função altamente regular, uma função analítica que, entre outras propriedades, tem um desenvolvimento em série de potências.

---

- **integral de Lebesgue** De Riemann, 1866, até Lebesgue, 1941 a teoria da integração evoluiu de *área* a *funcional* embora não seja esta a formulação que Lebesgue deu a integração ao fazer parte importante do grupo que construiu a *teoria das medidas*. Com a *teoria das medidas* se entendem falhas na teoria de integração motivada pela área (que é uma métrica) produzida pelos contemporâneos de Riemann e da qual Riemann é um dos nomes mais significativos. Um resultado sutil da teoria da medida mostra que no intervalo unitário, o conjunto dos números irracionais tem medida 1 e o seu complementar, o conjunto dos números racionais tem medida zero. Isto faz com que a função característica do conjunto dos irracionais, deste intervalo, tenha medida de Lebesgue 1, mas esta função não é integrável à Riemann porque suas somas inferiores serão sempre zero e as superiores sempre 1. Podemos dizer que a medida de Lebesgue é uma completção da medida de Riemann na reta, ou no espaços de dimensão finita. Há várias extensões da teoria da integração de Riemann, uma delas denominada Riemann-Stieltjes é um passo intermediário para a teoria dita de Daniel em que a integral é vista como um funcional linear. Esta forma de descrever integração é muito mais poderosa do que o processo dito de Lebesgue e inclusive a integral de Lebesgue passa a ser representada pela função constante 1. Porém a teoria das medidas lança luz em outros aspectos que se perdem com a visão da integral como funcional linear, como por exemplo a *complexidade dos conjuntos* o que pode ser pressentido com a medida do conjunto dos números irracionais de módulo menor do que 1, acima mencionados. Não tem sentido falar em *quantidades*, aqui, mas podemos falar em *medidas*. A continuação deste processo levou à teoria das distribuições construída de forma independente, por L Schwartz, na França, e Sebastião Silva em Portugal em 1940.

---

- **integral de linha** É uma generalização das integrais *a uma variável* em que o integrando é uma função vetorial (com valores num espaço vetorial de dimensão maior do que 1) com variáveis reais. Um caso típico é o comprimento de arco de uma curva (embora nem sempre caracterizado como integral de linha). O teorema de Green é uma igualdade entre duas integrais em que uma delas é a integral sobre a fronteira da região sobre a qual a outra integral é calculada. A primeira é então uma integral de linha.

---

- **integravel a Riemann** Confira também *integral, integral de Lebesgue*.



**Definição 14 (integrável) a Riemann**

Uma função real de variável real,  $f$  é dita integrável a Riemann sobre um intervalo  $[a, b]$ , se as cadéias da classe de todas as partições do intervalo  $[a, b]$  definirem uma classe de equivalência de seqüências de Cauchy, via somas de Riemann, e então o símbolo  $\int_a^b f(x)dx$  representa o limite comum dessas seqüências de Cauchy.

Os pontos centrais para construir esta teoria são

- partições de  $[a, b]$  que formam uma Álgebra parcialmente ordenada de subconjuntos, intervalos, de  $[a, b]$ ;
- seqüências de Cauchy;
- somas de Riemann que são o instrumento para construir as seqüências de Cauchy.

Este artigo ainda precisa ser expandido.

- **inteligência artificial** É um ramo da computação científica que procura criar programas capazes de tomar decisões independentes daquelas que o programador tiver previsto e assim adquirir novas habilidades diferentes das originais. Há uma definição, devida Turing, do que seria *inteligência artificial* mas parece que ela nunca foi comprovada (a existência de inteligência usando a definição de Turing). Alan Mathison Turing é considerado um dos fundadores da computação teorica o que inclui os seus experimentos chamados de máquinas de Turing que fazem parte da Lógica. Há muita informação sobre Turing na wikipedia que seria inútil repetir aqui.

- **inversa, teorema da função** Este teorema estabelece condições para calcularmos a inversa de uma função que seja diferenciável. Suponha que  $f$  seja uma função diferenciável em um intervalo  $(a, b)$ , que  $c \in (a, b)$  e que  $f$  seja diferenciável. Se  $f'(c) \neq 0$  então existe uma vizinhança do ponto  $c$  em que podemos definir  $f^{-1}$ .

O exemplo gráfico mostra porque esta propriedade tem que ser local: válida numa vizinhança do ponto  $c$ , na figura (92) página 240, Na figura (92) você pode ver um pedaço do gráfico  $graf(f)$  que depois de espelhado relativamente à primeira bisettriz  $y = x$  corresponde ao gráfico de uma função. Nos pontos  $(a, f(a)), (b, f(b))$  a derivada é zero (tangente horizontal) que na imagem espelhada corresponde à tangentes verticais onde não há derivada. Assim o intervalo aberto  $(a, b)$  é uma vizinhança do ponto  $c$  onde se pode definir uma função inversa.

**Teorema 12 (da função inversa)** Se  $f$  for derivável com derivada diferente de zero no ponto  $c$ , existe uma vizinhança de um ponto  $c, \mathcal{V}(c)$  no domínio e uma vizinhança  $\mathcal{V}(f(c))$  em que podemos definir uma inversa

$$f^{-1} : \mathcal{V}(f(c)) \longrightarrow \mathcal{V}(c)$$

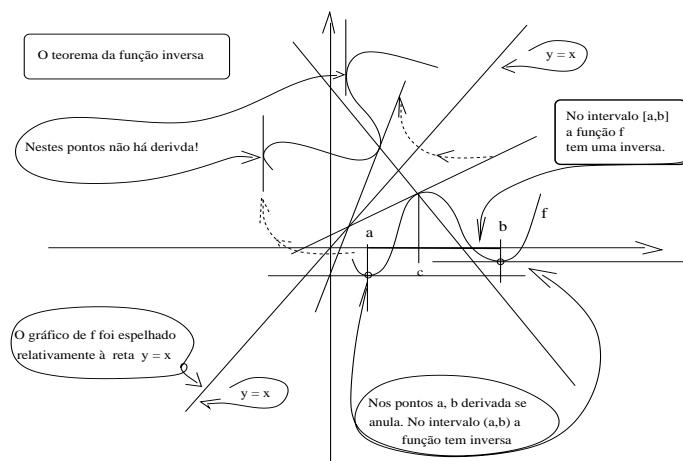


Figura 92: função inversa

É interessante uma extensão deste teorema para funções multivariadas. Suponha que  $F$  seja uma função vetorial de variável vetorial,

$$F : \mathbf{R}^n \times \mathbf{R}^m \longrightarrow \mathbf{R}^m$$

em que a escolha da dimensão nos espaços de saída e de chegada logo lhe vai ficar clara. Também numa redação mais precisa diríamos que  $F$  está definida num domínio  $\Omega \subset \mathbf{R}^n \times \mathbf{R}^m$ , entretanto a ideia essencial lhe será transmitida com esta versão.

Acompanhe as contas e os comentários que faremos logo em seguida:

$$\left\{ \begin{array}{l} x \in \mathbf{R}^n, y \in \mathbf{R}^m; \\ F : \mathbf{R}^n \times \mathbf{R}^m \longrightarrow \mathbf{R}^m; F(a, b) = c \in \mathbf{R}^m; \\ z = F(x, y) = (F_1(x, y) \dots, F_m(x, y)) \in \mathbf{R}^m; \\ J(F) = \left( \begin{array}{c} \left[ \begin{array}{ccc} F_{1,x_1} & \dots & F_{1,x_n} \\ \dots & & \dots \\ F_{m,x_1} & \dots & F_{m,x_n} \end{array} \right] \left[ \begin{array}{ccc} F_{1,y_1} & \dots & F_{1,y_m} \\ \dots & & \dots \\ F_{m,y_1} & \dots & F_{m,y_m} \end{array} \right] \\ D_x \quad D_y \end{array} \right) \end{array} \right. \quad (721)$$

Na quarta linha do sistema de equações (721) escrevemos a jacobiana de  $F$ ,  $J(F)$ , como uma matriz-blocos, novamente indicando isto na última linha em que estamos representando os dois blocos com índices  $x, y$  porque as derivadas parciais neles contidas, são respectivamente as derivadas parciais relativamente às coordenadas de cada um destes vetores sendo então razoável usar a notação de derivadas parciais com índices vetoriais para indicar estes dois blocos da jacobiana.

A razão desta subdivisão em matrizes-bloco logo ficará clara: precisaremos de uma matriz inversível, portanto precisamos de uma matriz quadrada. Uma questão pode ser levantada sobre a seleção das coordenadas: quem seriam  $x$

ou  $y$ , deixaremos de lado esta dificuldade que nos conduziria a um problema completamente diferente, mas é uma questão pertinente.

Vamos escrever a fórmula de Taylor do primeiro grau para  $F$  no ponto  $(a, b)$ :

$$F(x, y) = F(a, b) + J(F) \begin{pmatrix} x - a \\ y - b \end{pmatrix} + o \left( \begin{pmatrix} x - a \\ y - b \end{pmatrix} \right); (722)$$

$$F(x, y) = F(a, b) + \begin{pmatrix} D_x & D_y \end{pmatrix} \begin{pmatrix} x - a \\ y - b \end{pmatrix} + o \left( \begin{pmatrix} x - a \\ y - b \end{pmatrix} \right); (723)$$

$$F(x, y) = F(a, b) + D_x(x - a) + D_y(y - b) + o \left( \begin{pmatrix} x - a \\ y - b \end{pmatrix} \right); (724)$$

$$D_y(y - b) = F(x, y) - F(a, b) - D_x(x - a) - o \left( \begin{pmatrix} x - a \\ y - b \end{pmatrix} \right); (725)$$

$$D_y(y - b) = ({}_1F(x), {}_2F(y)) - ({}_1F(a), {}_2F(b)) - D_x(x - a) - o \left( \begin{pmatrix} x - a \\ y - b \end{pmatrix} \right); (726)$$

$$D_y(y - b) = ({}_1F(x) - {}_1F(a), {}_2F(y) - {}_2F(b)) - D_x(x - a) - o \left( \begin{pmatrix} x - a \\ y - b \end{pmatrix} \right); (727)$$

$$D_y(y - b) = ({}_1F(x) - {}_1F(a), G^{-1}(y) - G^{-1}(b)) - D_x(x - a) - o \left( \begin{pmatrix} x - a \\ y - b \end{pmatrix} \right) (728)$$

em que “ $o$ ” nas equações acima é o “*o pequeno de Landau*”. Também estamos agrupando a ação de  $F$  sobre  $x$  ou sobre  $y$  com a notação  ${}_1F(x), {}_2F(y)$ , idênticamente sobre  $a$  e  $b$ . As matrizes  $D_x, D_y$  são matrizes numéricas uma vez que as derivadas parciais foram calculadas no ponto  $(a, b)$  de acordo com a fórmula de Taylor, e neste ponto a notação está deficiente por não o indicar, mas a notação ficaria muito complicada e preferimos chamar sua atenção e seguir com a notação deficiente. Entretanto, vamos precisar de usar o fato de que  $D_y$  é a jacobiana calculada no ponto  $b$ .

Suponha agora que a matriz-bloco  $D_y$  seja inversível<sup>41</sup> então  $D_y$  é a jacobiana de uma função localmente inversível,

$$D_y = J(G)|_b; \mathbf{R}^m \xrightarrow{G} \mathbf{R}^m; F(a, b) = c;$$

em que  $G$  é a componente de  $F$  sobre o vetor  $y$ .

**Teorema 13 (da função inversa)** *Caso multivariado* Se

$$F : \mathbf{R}^n \times \mathbf{R}^m \longrightarrow \mathbf{R}^m$$

for derivável e pudermos identificar um bloco  $m \times m$ ,  $D$  da jacobiana  $J(F)$  cujo determinante seja diferente de zero em um ponto  $(a, b)$  do domínio, então existem abertos  $\mathcal{V}(a, b), \mathcal{V}(F(a, b))$  tal que a restrição

$$F|_{\mathcal{V}(a, b)} \longrightarrow \mathcal{V}(F(a, b))$$

é inversível, com  $J(F|_{\mathcal{V}(a, b)})^{-1} = D^{-1}$ .

<sup>41</sup>Corresponde a “derivada diferente de zero” na versão univariada do teorema da função inversa que estabelecemos logo acima. Agora corresponde dizer que determinante da jacobiana é diferente de zero.

Uma formulação ainda mais bonita deste teorema conduz de maneira natural ao conceito de *atlas*. Para isto considere que o domínio de  $F$  seja um conjunto compacto,  $\Omega$  de  $\mathbf{R}^{n+m}$ ,

$$F : \Omega \longrightarrow \mathbf{R}^m$$

a família dos determinantes maiores  $m \times m$  indexados nos pontos  $(a, b) \in \Omega$  que sejam diferente de zero, define uma cobertura por abertos de  $\Omega$  da qual se pode selecionar uma subcobertura finita (aqui está a razão da hipótese de compacidade de  $\Omega$ ). Esta família é um *atlas* de  $\Omega$  e  $F$  restrita a cada um dos *mapas* deste *atlas* é inversível. A demonstração é semelhante a que já fizemos acima.

Esta é a formulação que se pode usar para generalizar o teorema da função inversa para espaços abstratos da Análise Funcional.

---

- **isomorfismo** Dadas duas estruturas, aparentemente diferentes, é possível estabelecer entre elas um *functor* que ponha em relação os aspectos essenciais das duas teorias. Se este *functor* tiver na base uma função bijetiva entre os dois conjuntos das duas estruturas, temos um isomorfismo entre elas. Por exemplo, o espaço vetorial dos polinômios de grau menor ou igual à  $n$  é caracterizado essencialmente pelos coeficientes destes polinômios e assim podemos colocar em correspondência qualquer polinômio neste espaço com a enúpla de  $n+1$  números reais dos coeficientes. Estabelecemos assim uma correspondência bijetiva entre  $\mathbf{R}_n[x]$ , o espaço vetorial dos polinômios de grau menor ou igual à  $n$ , e  $\mathbf{R}^{n+1}$ . Se mostrarmos que a soma de um lado correspondente, pelo isomorfismo, à soma do outro lado, e semelhantemente, o produto por um escalar, então teremos mostrado que se trata de um isomorfismo de espaço vetorial e não haverá mais razões para não considerarmos estas duas estruturas como idênticas. Um isomorfismo é um tipo particular de *morfismo* e o texto acima pode ser lido substituindo *isomorfismo* por *morfismo* apenas não valeria a equivalência entre as estruturas, os morfismos são *mais fracos* que os isomorfismos.

---

- **jacobiana** é o nome adquirido a derivada  $DF$  de uma função de varias variáveis, a matriz das derivadas parciais.

---

- **kernel** É uma palavra alemã que significa núcleo, as duas palavras são usadas em português, em Matemática, com o mesmo sentido. Procure núcleo.

---

- **Lagrange, interpolação de** O algoritmo de interpolação polinomial conhecido como de Lagrange, possivelmente com a contribuição indireta de Waring, confira Waring, [12, Waring], consiste na construção dum polinômio de grau  $n$  que interpola  $n+1$  pontos do plano. Se for você tiver dois pontos para interpolar, o resultado será a reta que passa por estes 2 pontos, o polinômio do grau 1 determinado por estas retas, esta observação justifica o artigo definido caracterizando a unicidade do polinômio de interpolação de Lagrange, ele não é apenas o melhor, é o único.

Seria justo designar o algoritmo da interpolação de Lagrange como *Lagrange-*

Waring mas não pude encontrar suporte claro para esta denominação. Waring teve um vasto domínio de investigações mas parece ter sido muito pouco sistemático e sem a preocupação de registrar seus resultados.

O algoritmo da interpolação de Lagrange consiste na construção do polinômio de grau  $n$  que interpola  $n + 1$  pontos do plano. Os programas

`LagrangeWaring.gnuplot`; `LagrangeWaring.py`

que podem ser baixados de [6], lhe permitem construir e visualizar os gráficos do *polinômio de interpolação de Lagrange* e também os átomos cuja soma é o *polinômio de interpolação de Lagrange*. Estes programas foram usados na redação deste texto. O gráfico na figura (fig 93), página 243, mostra um polinômio de

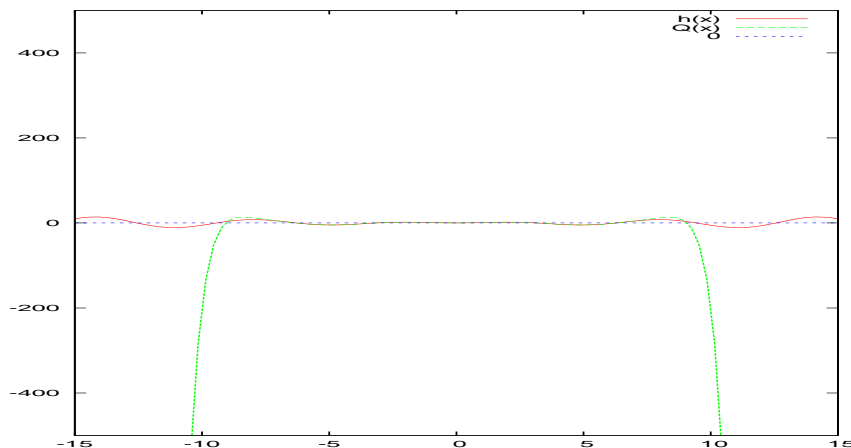


Figura 93: polinômio de interpolação de Lagrange

grau 11 interpolando 12 pontos escolhidos sobre o gráfico de  $h(x) = x \sin(x)$  e o programa mencionado acima lhe permite fazer várias experiências para entender e visualizar o método. No gráfico na figura (fig 94), página 244, você pode ver um zoom do gráfico anterior, (fig 93), mostrando a vizinhança tubular de raio 10 que contém o gráfico de  $h(x) = x \sin(x)$  no intervalo  $[-7, 7]$  mostrando o erro na aproximação do polinômio de Lagrange que fica mais acentuado nos intervalos extremos da partição do intervalo  $[-7, 7]$  sobre a qual foram escolhidos os pontos a serem interpolados.

Nos parágrafos acima há dois conceitos que é preciso separar, *interpolação* e *aproximação* que embora sejam usados frequentemente como sinônimos, eles têm significados, e definições, e objetivos distintos.

Na *interpolação* não se discute a *aproximação*, simplesmente se apresenta um modelo representado idealmente por uma curva que passa pelos pontos interpolados e há vários métodos de interpolação que criam curvas passando por uma coleção de pontos dados.

Quando consideramos *aproximação*, pretendemos estudar o erro da interpolação e neste caso é preciso apresentar uma distância que caracterize este erro. A figura (fig 94) sugere várias formas de medir este erro, por exemplo

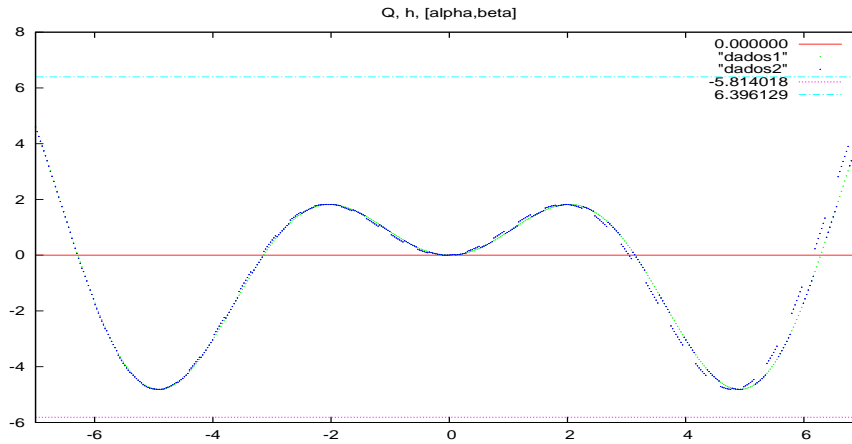


Figura 94: zoom do polinômio de interpolação de Lagrange

usando a integral do módulo da diferença  $\|P_n(x) - h(x)\|$ , ou a norma do supremo,

$$\|P_n - h\|_1 = \int_a^b |P_n(x) - h(x)| dx; \quad (729)$$

$$\|P_n - h\|_\infty = \sup_{x \in [a,b]} |P_n(x) - h(x)|; \quad (730)$$

O número  $n$  que aparece como índice do polinômio, indica o grau do polinômio que interpola os pontos e a função  $h$  é a “realidade” que está sendo *aproximada*.

Quando usamos a integral, como método para calcular distância, estamos nos preocupando com a energia do fenômeno e então as oscilações são *um problema menor*. A norma do supremo é mais restritiva uma vez que ela está associada com a convergência uniforme.

O erro na interpolação de Lagrange, vista como *aproximação*, pode ser diminuído se usarmos um truque com a *interpolação*: espalharmos os pontos de interpolação sobre um intervalo maior e com isto reduzimos o efeito da oscilação marcante que ocorre nos intervalos extremos da partição considerada no intervalo  $[a, b]$  onde queremos considerar a aproximação. As duas figuras seguintes ilustram o uso desta metodologia, mas antes de comentá-las vou introduzir a terminologia de *teoria da aproximação polinomial* para tornar a linguagem mais precisa.

São escolhidos  $n + 1$  pontos chamados *pontos de precisão* que você pode interpretar como sendo as medidas obtidas num experimento, seriam os pontos que caracterizam a “realidade”, enquanto que a curva que interpola estes pontos é um modelo para esta realidade. O polinômio de Lagrange passa por estes pontos e é isto que se chama de *interpolação*.

A figura (fig 95), página 245, foi feita para comparar com a que se lhe segue, e

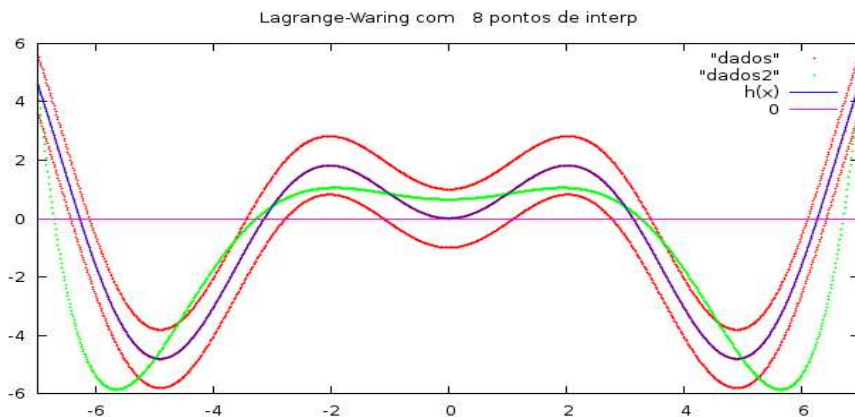


Figura 95: visão do polinômio de interpolação num subintervalo

foi construída a partir duma quantidade pequena de pontos, 8, que são os pontos em que o polinômio de grau 7 corta o gráfico da função escolhida. Na figura (fig 95) você pode ver o gráfico da função  $y = h(x)$  translataada de uma unidade, para cima e para baixo, para caracterizar uma vizinhança tubular centrada em  $h$ . O gráfico do polinômio de Lagrange é o que aparece em verde na figura.

Para obter a figura (fig 96), página 245, foram escolhidos 61 pontos distri-

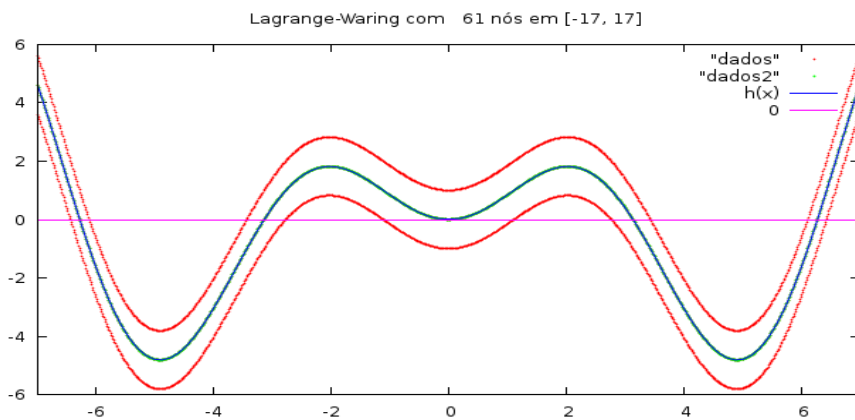


Figura 96: visão do polinômio de interpolação num subintervalo

buidos uniformemente no intervalo  $[-17, 17]$ , quando o nosso objetivo era obter uma boa aproximação no intervalo  $[-7, 7]$ , é o truque mencionado acima, e desta forma conseguimos uma aproximação uniforme excelente na vizinhança tubular

que aparece na figura (fig 96) porque expulsamos a oscilação mais intensa do polinômio para os extremos do intervalo  $[-17, 17]$ .

### O defeito da interpolação polinomial de Lagrange

O defeito da interpolação polinomial de Lagrange é que ela constrói um único polinômio, e no caso da figura (fig 96) temos um polinômio de grau 60 o que produz um “custo computacional elevado”. Hoje temos métodos de interpolação polinomial mais efetivos usando *splines*. Os *splines* são polinômios por pedaços o que permite trabalhar com um grau menor e obter uma aproximação superior. Observe que uma poligonal é um 0-splines e que os gráficos excelentes do *gnuplot* são feitos usando poligonais quando você não selecionar outro *método de interpolação*.

Caberia aqui salientar a contradição duma discussão dum algoritmo em que o próprio texto sugere que ele se encontra, tecnologicamente defasado, uma vez que os *splines* representam uma melhoria considerável sobre o a interpolação polinomial de Lagrange. O meu objetivo neste artigo é a demonstração do Teorema de Weierstrass e preciso de polinômios o que justifica o meu uso duma “*tecnologia defasada*”, apenas os avanços computacionais que se encontram à nossa disposição, nos permitem tirar o polinômio de Lagrange do arquivo mas para usá-lo na construção duma demonstração, ele é um instrumento teórico e nem sempre os instrumentos teóricos são ferramentas tecnológicas...

### O método

O método de Lagrange é muito engenhoso e vou descrever a metodologia dos programas mencionados acima o que vai permitir-lhe alterá-los facilmente para obter outros experimentos.

1. Os pontos de precisão Primeiro selecionamos os pontos de precisão, ou a “realidade” nos apresenta estes pontos ...

No caso do programa eles foram obtidos construindo, primeiro, uma progressão aritmética sobre o intervalo  $[a, b]$  onde desejamos ver o gráfico, ou no caso da figura (fig 96) sobre um intervalo maior para expulsar a oscilação grande do polinômio para fora do cenário em que desejamos a aproximação. Estes pontos são os nós da partição do intervalo  $[A, B]$ , e vou designá-los por  $a_i; i = 0, \dots, n$ , então  $A = a_0; a_n = B$ . Aqui estou usando uma dubiedade que as linguagens de programação modernas me oferecem quando consideram distintas as letras maiúsculas ou minúsculas. Nos programas,  $[a, b]$  é o intervalo onde quero ver a *aproximação* da sucessão de polinômios, enquanto que  $[A, B]; [a, b] \subset [A, B]$  é o intervalo mais amplo onde a interpolação está sendo construída, é o *truque mencionado*.

Depois selecionamos os pontos de precisão considerando as imagens destes nós pela função  $h(x) = x \sin(x)$  que é uma função *muito oscilante*, logo bem interessante para testar aproximação.



2. Um polinômio auxiliar, chamado  $P(x)$  no programa, tendo como raízes os nós da partição. Como é um produto de monômios, a derivada de  $P$ , chamada  $dP$  no programa, é a soma de  $n + 1$  produtos obtidos quando eliminamos cada um dos fatores, uma soma de polinômios de grau  $n$ , uma aplicação direta da *regra do produto para derivadas*.

Alguns textos que descrevem o algoritmo fazem referência aos produtos que compõem a *regra do produto para derivadas* como  $\frac{P(x)}{x-a_i}$ . Não dá para usar esta notação num programa de computação porque se produziria um erro de divisão por zero. Estes polinômios-parcelas são chamados no programa de  $P_i$  e foram obtidos por eliminação sucessiva dos monômios definidos pelas raízes de  $P$ . De fato

$$P_i(x) = \frac{P(x)}{x - a_i}; dP = \sum_{i=0}^n P_i = P';$$

A derivada  $dP$  não se anula nas raízes de  $P$  e registra a oscilação de  $P$  na vizinhança de cada uma das raízes.

*Aqui você pode comparar o planejamento límpido que nos permite uma linguagem de programação de alto nível como é Python comparado com o projeto de trabalho manual representado pelo script do gnuplot. Ao mesmo tempo é necessário deixar claro que esta comparação não deve lançar uma sombra negativa em gnuplot que é uma ferramenta para fazer gráficos que está sendo usada internamente pelo programa em Python. O meu trabalho para produzir este programa em Python seria imensamente maior se eu não tivesse gnuplot disponível.*

3. Os átomos do polinômio de Lagrange O polinômio de Lagrange vai ser a soma de  $n + 1$  polinômios de grau  $n$  indexados nas raízes definidos pela equação

$$Q_i = h(a_i)P_i/dP(a_i); \quad (731)$$

$$Q_i(x) = h(a_i)P_i(x)/dP(a_i); \quad (732)$$

Mas na linguagem dos programas:

$$(a) dP = \sum_{i=0}^n P_i$$

$$(b) P_i(a_j) \neq 0 \iff i = j$$

$$(c) P_i(a_j) = 0 \iff i \neq j$$

então

$$(a) Q_i(a_i) = h(a_i)P_i(a_i)/P_i(a_i) = h(a_i)$$

- (b)  $Q_j(a_j) = h(a_j) \iff i = j$   
 (c)  $Q_i(a_j) = 0 \iff i \neq j$   
 (d)  $Q = \sum_{i=0}^n Q_i$  coincide com  $h$  sobre os nós da partição que são os *pontos de precisão*.

$Q$  é o polinômio de interpolação de Lagrange que interpola os pontos

$$(a_i, h(a_i)); i = 0, \dots, n;$$

Como é um polinômio de grau  $n$ , fica univocamente determinado pelos  $n + 1$  pontos interpolados e portanto também pode ser obtido resolvendo-se uma equação linear de ordem  $n + 1$  cujas incógnitas são os  $n + 1$  coeficientes de  $Q$ .

- **laplaciano** é o nome do símbolo do operador de Laplace,  $\Delta$  ou  $\nabla^2$ . Ver *operador de Laplace*.

- **lei do coseno** Há duas leis que recebem esta denominação,

- uma delas é uma relação trigonométrica que generaliza o teorema de Pitágoras.
- também *lei do coseno* descreve as deformações de áreas de superfícies no espaço tridimensional.

O Teorema de Pitágoras estabelece que num triângulo retângulo com catetos  $a, b$  e a hipotenusa  $c$ , vale a relação<sup>42</sup>

$$c^2 = a^2 + b^2 \tag{733}$$

“a soma dos quadrados dos catetos é igual ao quadrado da hipotenusa”. A lei dos cossenos estabelece uma relação semelhante, válida para quaisquer sejam os lados de qualquer triângulo, conseqüentemente uma generalização do teorema de Pitágoras.

Dados dois vetores  $\vec{u}, \vec{v}$ , como na figura (fig 97), página 249, eles determinam um triângulo em que o terceiro lado é o segmento de reta *equipolente* à diferença entre eles,  $\vec{u} - \vec{v}$ , ou  $\vec{v} - \vec{u}$  cujo comprimento está indicado como  $h$  na figura (fig 97). A *lei dos cossenos* estabelece

$$h^2 = \|\vec{u}\|^2 + \|\vec{v}\|^2 - 2\|\vec{u}\|\|\vec{v}\|\cos(\alpha); \tag{734}$$

em que  $\alpha$  é o ângulo que os vetores  $\vec{u}, \vec{v}$  determinam no círculo trigonométrico  $\mathbf{S}^1$ .

Quando o ângulo  $\alpha = \frac{\pi}{2}$ , um ângulo “reto”, o termo  $2\|\vec{u}\|\|\vec{v}\|\cos(\alpha)$  se anula e “caímos” no teorema de Pitágoras.

<sup>42</sup>A expressão mais geral desta relação é o “último teorema de Fermat”  $a^n + b^n = c^n$  que somente se verifica para  $n = 2$  quando  $a, b, c \in \mathbf{N}$ .

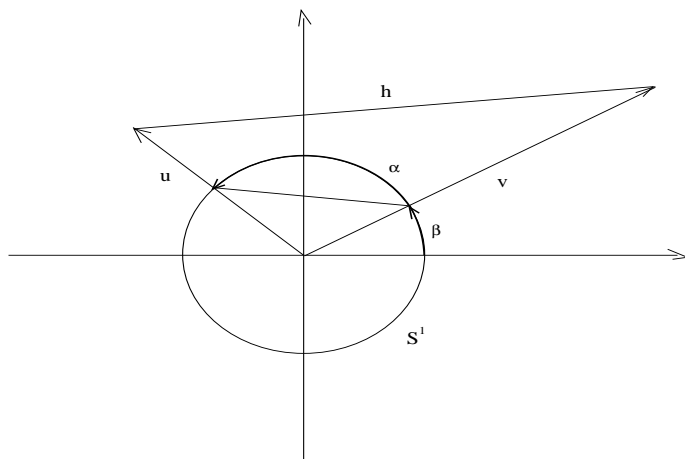


Figura 97: Lei dos cosenos

Observe que o termo  $2\|\vec{u}\|\|\vec{v}\|\cos(\alpha)$  contém a expressão geométrica<sup>43</sup> do produto escalar entre os vetores  $\vec{u}, \vec{v}$ , confira *produto escalar*.

Se o ângulo  $\alpha$  for agudo, o termo em que aparece o *produto escalar* será negativo e  $h^2$  é menor do que a soma dos quadrados dos módulos dos vetores  $\vec{u}, \vec{v}$ . Se  $\alpha$  for obtuso, medir mais do que  $\frac{\pi}{2}$ , então o termo em que aparece o *produto escalar* será positivo e  $h^2$  é maior do que a soma dos quadrados dos módulos dos vetores  $\vec{u}, \vec{v}$ . Se  $\alpha = \frac{\pi}{2}$ , que é o caso dos triângulos retângulos, então o termo em que aparece o *produto escalar* será *nulo* que é o caso do teorema de Pitágoras, em que  $h$  é a medida da hipotenusa.

Há várias formas de demonstrar a *lei do coseno*, e vou apresentar duas delas. Na primeira vou fazer uma transformação que irá simplificar as coordenadas dos vetores, serve também para exemplificar o uso da *Geometria Analítica* na obtenção dos teoremas da *Geometria Euclidiana* em particular eu vou fazer uso dum poderoso instrumento, o círculo trigonométrico,  $\mathbf{S}^1$ .

A figura (fig 97), página 249, mostra que *um tipo de projeção de um triângulo qualquer* no círculo trigonométrico pode ser feita com uma certa *deformação* “caindo” num triângulo cujos lados sejam vetores unitários. A *deformação* consiste em que o terceiro lado do triângulo original não é paralelo com o terceiro lado do triângulo-imagem, a versão transformada, e conseqüentemente os *triângulos não são semelhantes*. Considere esta projeção deformada do triângulo em  $\mathbf{S}^1$  apresentada na figura (fig 97).

O objetivo desta projeção foi apenas o determinar o ângulo do vetor  $\vec{v}$  para aplicar ao triângulo uma rotação que irá simplificar uma das coordenadas. A imagem do triângulo, embora deformada, diz-me que o ângulo do vetor  $\vec{v}$  é  $\beta$  e se eu der uma rotação de  $-\beta$  no triângulo, a imagem de  $\vec{v}$  terá ângulo 0 e a imagem de  $\vec{u}$  terá ângulo  $\alpha$ . Você pode ver na figura (fig 98), página 250, o

<sup>43</sup>Aqui havia um erro, corrigido!

resultado desta rotação

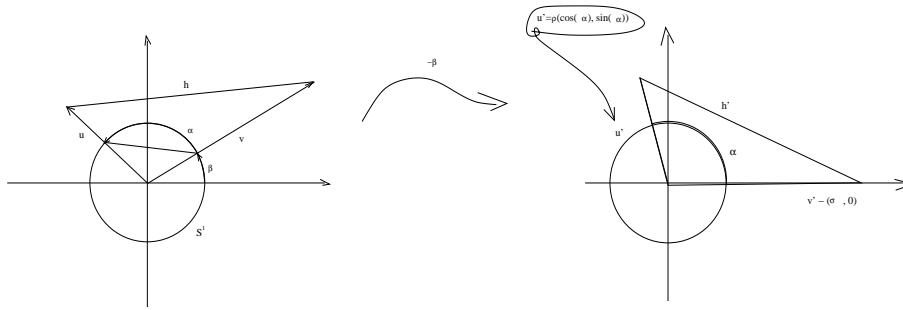


Figura 98: Imagem transformada por rotação de  $-\beta$

Ao fazer a rotação  $-\beta$  do triângulo obtive os vetores

$$\vec{u}' = \rho(\cos(\alpha), \sin(\alpha)); \vec{v}' = (\sigma, 0);$$

O uso do *círculo trigonométrico*,  $\mathbf{S}^1$ , nos permite simplificar a linguagem deixando de dizer “o ângulo que o vetor  $\vec{v}'$  faz com eixo  $OX$  é zero”, porque todo ponto sobre  $\mathbf{S}^1$  determina um ângulo, basta dizer “o ângulo do  $\vec{v}' = (\sigma, 0)$  é zero”. O ângulo entre  $\vec{u}'$ ,  $\vec{v}'$  é  $\alpha$ , o ângulo de  $\vec{u}'$  é  $\alpha$ .

Os dois triângulos assim transformados, por uma rotação, são semelhantes, e a razão entre seus lados homólogos é 1.

Vou usar as coordenadas e os módulos destes novos vetores e calcular o comprimento do vetor diferença, que é um segmento de reta que tem o mesmo comprimento que  $\vec{u} - \vec{v}$ . Acompanhe os cálculos nas equações:

$$\vec{u}' = \rho(\cos(\alpha), \sin(\alpha)) = (\rho \cos(\alpha), \rho \sin(\alpha)); \|\vec{u}'\| = \rho; \quad (735)$$

$$\vec{v}' = \sigma(1, 0) = (\sigma, 0); \|\vec{v}'\| = \sigma; \quad (736)$$

$$\vec{u}' - \vec{v}' = (\rho \cos(\alpha), \rho \sin(\alpha)) - (\sigma, 0); \quad (737)$$

$$\vec{u}' - \vec{v}' = (\rho \cos(\alpha) - \sigma, \rho \sin(\alpha)); \quad (738)$$

$$h' = \|\vec{u}' - \vec{v}'\| = \|(\rho \cos(\alpha) - \sigma, \rho \sin(\alpha))\|; \quad (739)$$

$$(h')^2 = (\rho \cos(\alpha) - \sigma)^2 + (\rho \sin(\alpha))^2; \quad (740)$$

$$(h')^2 = \rho^2 - 2\rho\sigma \cos(\alpha) + \sigma^2; \quad (741)$$

$$(h')^2 = \|\vec{u}'\|^2 - 2\|\vec{u}'\|\|\vec{v}'\| \cos(\alpha) + \|\vec{v}'\|^2; \quad (742)$$

e aparece a expressão da *lei do cosseno* na última equação.

Como *dei uma rotação  $-\beta$  no triângulo original* para obter o triângulo com lados  $\vec{u}'$ ,  $\vec{v}'$ , então vale a mesma relação entre os lados do triângulo primitivo.

$$h^2 = \|\vec{u}\|^2 - 2\|\vec{u}\|\|\vec{v}\| \cos(\alpha) + \|\vec{v}\|^2; \quad (743)$$

Observe que *rotação* é um exemplo de *codificação*<sup>44</sup>, que é uma *transformação* que fazemos uma “*informação*” com o objetivo de torná-la mais simples para guardá-la, transmiti-la, ou escondê-la (criptação) mas resguardada a possibilidade *decodificar* para retornar ao *caso* original como fiz com o triângulo.

A expressão da lei do cosseno para o triângulo sob rotação  $-\beta$  onde  $\beta$  é o “ângulo do vetor”  $\vec{v}$  original aparece na última equação.

Estou usando a “linguagem privativa” de  $\mathbf{S}^1$  para um vetor “qualquer” do plano, isto é possível porque qualquer vetor do plano tem uma projeção em  $\mathbf{S}^1$  e assim determina de maneira única um arco do círculo trigonométrico que é o seu *ângulo*.

Na segunda demonstração vou mostrar-lhe que podemos recuperar todas estas contas de uma forma muito mais simples usando o *produto escalar*:

$$h = \|\vec{u} - \vec{v}\| = \langle \vec{u} - \vec{v}, \vec{u} - \vec{v} \rangle = \quad (744)$$

$$= \langle \vec{u}, \vec{u} \rangle - \langle \vec{u}, \vec{v} \rangle - \langle \vec{v}, \vec{u} \rangle + \langle \vec{v}, \vec{v} \rangle = \quad (745)$$

$$= \langle \vec{u}, \vec{u} \rangle - 2 \langle \vec{u}, \vec{v} \rangle + \langle \vec{v}, \vec{v} \rangle = \quad (746)$$

$$= h^2 = \|\vec{u}\|^2 - 2 \langle \vec{u}, \vec{v} \rangle + \|\vec{v}\|^2; \quad (747)$$

$$= h^2 = \|\vec{u}\|^2 - 2\|\vec{u}\|\|\vec{v}\|\cos(\alpha) + \|\vec{v}\|^2; \quad (748)$$

onde podemos ler o resultado que aparece na equação (eq. 743) usando a definição geométrica do produto escalar, confira *produto escalar*.

Podemos tirar uma lição desta última demonstração, ao fazer abstrações as contas ficam mais sintéticas e elegantes porque toda uma teoria fica escondida . . . mas é esta a forma como o conhecimento<sup>45</sup> evolui. . .

- **lei do cosseno** Procure acima a *lei do cosseno*, como relação trigonométrica.

A lei do cosseno para áreas de superfícies fornece um valor local de distorção entre a área de uma superfície e a sua região de parametrização, no domínio. Pense numa região plana na encosta dum morro e na área que ela representa para agricultura. As plantas crescem “paralelamente” aos raios solares como resposta à tentativa da Natureza de otimizar a recepção da energia, então a área de agricultura nesta região plana na encosta do morro é menor do que a área geométrica que ela mede, confira a figura (fig 99), página 252, onde você pode ver a região plana associada ao seu “domínio de parametrização” e na região plana podemos ver o vetor normal  $\vec{N}$  que faz um ângulo  $\gamma$  relativamente à direção do eixo  $OZ$ . A área útil para plantação na região montanhosa é menor do que a área que ela realmente mede, e o coeficiente de correção é  $\cos(\gamma)$ .

A relação entre as duas áreas, a área da região plana na encosta do morro,  $S$ , e a área do domínio  $D$ , é

$$\text{Área}(D) = \cos(\gamma)\text{Área}(S) \quad (749)$$

<sup>44</sup>Esta frase é um exemplo de “*modem*” destes usados em telecomunicações, *codifiquei e decodifiquei* sem perder nada da informação. . . nas comunicações se perde sempre.

<sup>45</sup>Primeiro nós sabíamos programar em *assembler*, depois inventamos as linguagens “modernas” de programação que *abstraem* o *assembler*. Os computadores entendem *assembler* mas nós não sabemos mais programar em *assembler*.

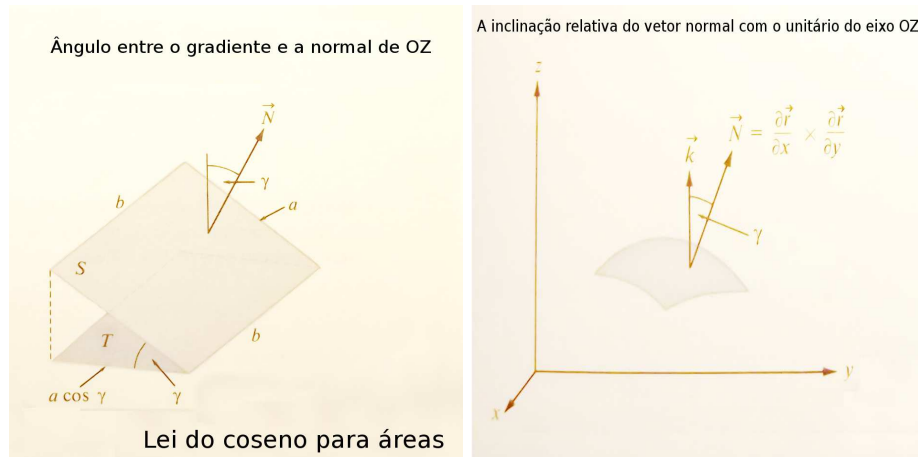


Figura 99: Lei do cosseno: área de superfície

Na figura em que aparece a área plana, as letras  $a, b$  fazem referência às medidas dos lados da região retangular inclinada, portanto os lados correspondentes no domínio medem  $a \cos(\gamma), b$  e desta forma a área aproveitável para agricultura<sup>46</sup> seria

$$ab \cos(\gamma)$$

calculada a partir das medições feitas na região montanhosa.

Em geral esta relação é escrita ao reverso porque se supõe conhecida a área do domínio, o que é inverídico no caso da montanha, onde se pode medir a área da superfície.

$$\frac{\text{Área}(S)}{\cos(\gamma)} = \text{Área}(D) \quad (750)$$

Mas é a equação (eq. 750) que irá aparecer na integral utilizada para o cálculo de superfícies porque neste caso se supõe conhecido o domínio o que é frequentemente verdadeiro em aplicações, por exemplo, se você quiser calcular o aproveitamento da energia solar irradiada sobre a superfície de placas dispostas no solo da Terra ou sobre algum satélite, neste caso é a área do domínio que é conhecida o que faz que a equação (eq. 750) seja a que melhor se adapte para o cálculo duma integral de superfície.

Vou desenvolver o cálculo da *integral de superfície* para mostrar o uso da *lei do cosseno*, e vou fazê-lo usando dois métodos.

Uma forma comum de representar a superfície  $S$  é na *forma implícita*

$$F(x, y, z) = 0$$

<sup>46</sup>A distorção no comprimento dos lados é dada pelos números  $s, t$  tal que  $st = \cos(\gamma)$ . No da figura  $s = \cos(\gamma), t = 1$  área( $D$ ) =  $(sa)(tb) = ab \cos(\gamma)$ .

da qual se pode deduzir, *nas condições do teorema da função implícita*, a expressão da superfície como função das variáveis  $(x, y)$  no domínio:

$$F(x, y, z) = 0 \Rightarrow z = f(x, y) \quad (751)$$

No cálculo da integral de superfície o vetor  $\vec{N}$ , que mede localmente a distorção de área da superfície relativamente ao domínio, em cada ponto de uma superfície, é obtido pelo *produto vetorial* de dois quaisquer vetores *linearmente independentes*, que se encontrem no *plano tangente* (que corresponderia à região plana na encosta da montanha):  $\vec{N} = \vec{r}_x \times \vec{r}_y$ .

Aqui há um defeito de notação, estou aproveitando uma figura que encontrei na Wikipédia, e me adaptando à notação da figura. O hábito é guardar o símbolo  $\vec{N}$ , em geral se usa minúscula, para o vetor normal, ortanto não está errado, mas é preciso alertar à leitora para não dificultar a compreensão:  $\vec{N}$  é apenas um vetor perpendicular ao plano tangente á superfície e obtido pela forma mais natural como produto vetorial de dois vetores calculados usando as derivadas parciais de  $f$ . As contas mostram que esta é a melhor opção.

Se admitirmos a equação (eq. 751) como dada, dois vetores contidos no plano tangente, em cada ponto do domínio seriam sob a suposição de que a derivada  $F_z \neq 0$  em todo o domínio  $D$ , (hipótese do *teorema da função implícita*),

$$F(x, y, z) = 0 \Rightarrow z = f(x, y); \quad (752)$$

$$\frac{\partial f}{\partial x} = -\frac{F_x}{F_z}; \frac{\partial f}{\partial y} = -\frac{F_y}{F_z}; \quad (753)$$

$$\vec{r}_x = (1, 0, \frac{\partial f}{\partial x}); \vec{r}_y = (0, 1, \frac{\partial f}{\partial y}); \quad (754)$$

$$\vec{N}(x, y) = (1, 0, \frac{\partial f}{\partial x}) \times (0, 1, \frac{\partial f}{\partial y}) = \vec{r}_x \times \vec{r}_y = \begin{vmatrix} i & j & k \\ 1 & 0 & \frac{\partial f}{\partial x} \\ 0 & 1 & \frac{\partial f}{\partial y} \end{vmatrix}; \quad (755)$$

$$\vec{N}(x, y) = (-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1); \|\vec{N}\|(x, y) = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1}; \quad (756)$$

$$\text{Área}(S) = \int_D \int \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2 + 1} dx dy = \int_D \int \|\vec{N}\|(x, y) dx dy; \quad (757)$$

Verificando, se  $S$  for uma área retângular, caso do morro,  $\text{Área}(S) > \text{Área}(D)$  e  $\|\vec{N}\|(x, y) = \frac{1}{\cos(\gamma)}(x, y) > 1$  e a segunda integral na equação (eq.757) é a área do domínio retângular multiplicada por um número maior do que 1. A função  $\frac{1}{\cos(\gamma)}(x, y)$  é o fator de correção para distorção da área da superfície relativamente ao seu domínio no plano, estou usando a *lei do coseno*.

O  $\cos(\gamma)$  está implícito no cálculo de  $\|\vec{N}\|$  uma vez que ele é perpendicular ao plano onde estão contidos os vetores  $\vec{r}_x, \vec{r}_y$  foi calculado usando dois vetores contidos no plano tangente ao gráfico de  $f$ , em cada ponto  $(x, y)$  do domínio de integração.

Podemos fazer estas contas sob outra ótica, suponha agora que a superfície  $S$  seja dada por um sistema de equações paramétricas<sup>47</sup>. Acompanhe os cálculos

<sup>47</sup>É o caso das contas anteriores com as equações paramétricas  $(x, y, f(x, y)) = (s_1(x, y), s_2(x, y), s_3(x, y))$ .

na sucessão de equações,

$$D \ni (s, t) \mapsto \vec{r}(s, t) = (s_1(s, t), s_2(s, t), s_3(s, t)) \in S \subset \mathbf{R}^3; \quad (758)$$

$$\frac{\partial \vec{r}}{\partial s} = \left( \frac{\partial s_1}{\partial s}, \frac{\partial s_2}{\partial s}, \frac{\partial s_3}{\partial s} \right); \frac{\partial \vec{r}}{\partial t} = \left( \frac{\partial s_1}{\partial t}, \frac{\partial s_2}{\partial t}, \frac{\partial s_3}{\partial t} \right); \quad (759)$$

$$\vec{N}(s, t) = \frac{\partial \vec{r}}{\partial s} \times \frac{\partial \vec{r}}{\partial t} = \vec{N}(s, t) = \begin{vmatrix} i & j & k \\ \frac{\partial s_1}{\partial s} & \frac{\partial s_2}{\partial s} & \frac{\partial s_3}{\partial s} \\ \frac{\partial s_1}{\partial t} & \frac{\partial s_2}{\partial t} & \frac{\partial s_3}{\partial t} \end{vmatrix} \quad (760)$$

$$\vec{N}(s, t) = \begin{vmatrix} \frac{\partial s_2}{\partial s} & \frac{\partial s_3}{\partial s} \\ \frac{\partial s_2}{\partial t} & \frac{\partial s_3}{\partial t} \end{vmatrix} i + \begin{vmatrix} \frac{\partial s_3}{\partial s} & \frac{\partial s_1}{\partial s} \\ \frac{\partial s_3}{\partial t} & \frac{\partial s_1}{\partial t} \end{vmatrix} j + \begin{vmatrix} \frac{\partial s_1}{\partial s} & \frac{\partial s_2}{\partial s} \\ \frac{\partial s_1}{\partial t} & \frac{\partial s_2}{\partial t} \end{vmatrix} k \quad (761)$$

$$\vec{N}(s, t) = \frac{\partial(s_2, s_3)}{\partial(s, t)} i + \frac{\partial(s_3, s_1)}{\partial(s, t)} j + \frac{\partial(s_1, s_2)}{\partial(s, t)} k \quad (762)$$

$$\text{Área}(S) = \iint_D \sqrt{\left( \frac{\partial(s_2, s_3)}{\partial(s, t)} \right)^2 + \left( \frac{\partial(s_3, s_1)}{\partial(s, t)} \right)^2 + \left( \frac{\partial(s_1, s_2)}{\partial(s, t)} \right)^2} ds dt \quad (763)$$

Você deve ver que as integrais na equação (eq. 757) e na equação (eq. 763) representam a mesma integral apenas parametrizadas de forma distinta. É verdade que agora fica muito difícil de ver  $\vec{N}(s, t) = \frac{1}{\cos(\gamma)} \big|_{(s, t)}$ .

### - lei do seno Lei do seno num triângulo inscrito em $\mathbf{S}^1$

Como três pontos determinam um círculo, dado um triângulo qualquer,  $PQR$ , confira a figura (fig 100), página 255, vou considerar este círculo como sendo o círculo trigonométrico,  $\mathbf{S}^1$ .

Estou usando o teorema seguinte:

#### **Teorema** 14 (Representação em $\mathbf{S}^1$ ) *representante em $\mathbf{S}^1$*

*Para qualquer que seja o triângulo  $PQR$ , existe uma sua representação inscrita no círculo trigonométrico,  $\mathbf{S}^1$ , e um número  $\rho$  que é o coeficiente de proporcionalidade entre os lados homólogos do triângulo original  $PQR$  e sua representação em  $\mathbf{S}^1$ .*

#### **Dem**:

*Dado um triângulo qualquer  $PQR$ , seja  $\mathbf{C}$  o círculo de raio  $R$  determinado pelos pontos  $P, Q, R$ , nesta ordem. Porque três pontos não colineares determinam de forma única um círculo<sup>48</sup>.*

*Execute duas operações geométrico-numéricas, confira produto geométrico*

- *Divida  $\mathbf{C}$  por  $R$ , obtendo um círculo de raio 1,  $\mathbf{C}'$*
- *translate  $\mathbf{C}'$  de modo que seu centro coincida com a origem dos eixos.*

*O resultado destas duas operações é  $\mathbf{S}^1$ , o círculo trigonométrico, com um triângulo inscrito que é semelhante ao triângulo primitivo  $PQR$ . O coeficiente de proporcionalidade entre os lados homólogos destes dois triângulos é  $\rho = R$ , o raio do círculo primitivo.*

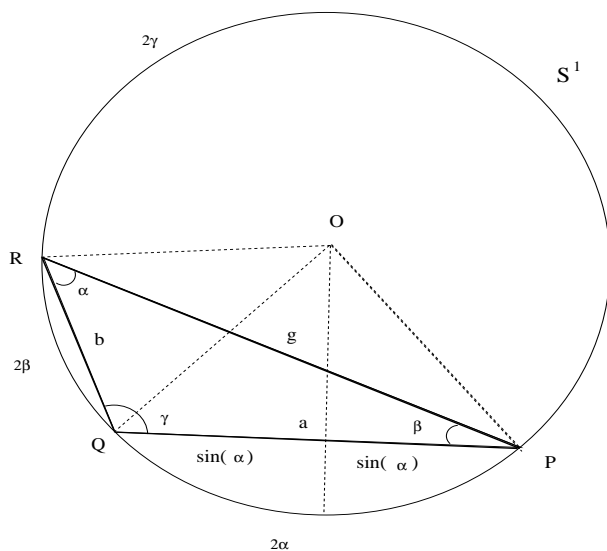
**q.e.d .**

Ou seja, podemos projetar qualquer triângulo em  $\mathbf{S}^1$ , de forma a ter uma representação (um triângulo equivalente) do mesmo como um triângulo inscrito no círculo trigonométrico.

Confira a figura (fig 100), página 255, em que o triângulo está inscrito

<sup>48</sup>E mesmo colineares, então o resultado é uma reta que é um círculo cujo raio é infinito. . .





$$a = 2\sin(\alpha); \quad b = 2\sin(\beta); \quad g = 2\sin(\gamma)$$

Figura 100: Lei dos senos

no círculo trigonométrico  $\mathbf{S}^1$ , e passo a chamar de  $PQR$  à *representação* do triângulo em  $\mathbf{S}^1$ .

Os lados do triângulo são as *cordas*,  $\overline{PQ}$ ,  $\overline{QR}$ ,  $\overline{RP}$ , que determinam em  $\mathbf{S}^1$  os arcos  $\alpha, \beta, \gamma$ . As cordas medem, respectivamente,

$$a = 2 * \sin(\alpha); \quad b = 2 * \sin(\beta); \quad g = 2 * \sin(\gamma) \quad (764)$$

porque o arco determinado no círculo pelo ângulo inscrito sobre o círculo mede o dobro do ângulo.

Vou mostrar-lhe como provar isto usando a figura (fig 101), página 256, também serve para lhe passar “técnicas de desenho geométrico” usando um editor de gráficos como `xfig`.

- Trace círculos com raios iguais em  $R$  e em  $P$ . Para fazê-lo copie o segmento  $\overline{RP}$  montando um paralelograma, e leve cada uma destas cópias para os pontos  $R, P$ .
- a reta  $s$  que passar pela interseção destes círculos, corta a corda  $RP$  perpendicularmente no ponto médio sendo então a mediatriz da nova versão do ângulo  $\gamma$  obtido ligando  $R$  e  $P$  ao ponto em que esta reta corta  $\mathbf{S}^1$ .
- $c$  é o dobro do  $\sin(\gamma) = \text{logo}$

$$\frac{c}{\sin(\gamma)} = \frac{2\sin(\gamma)}{\sin(\gamma)} = 2;$$

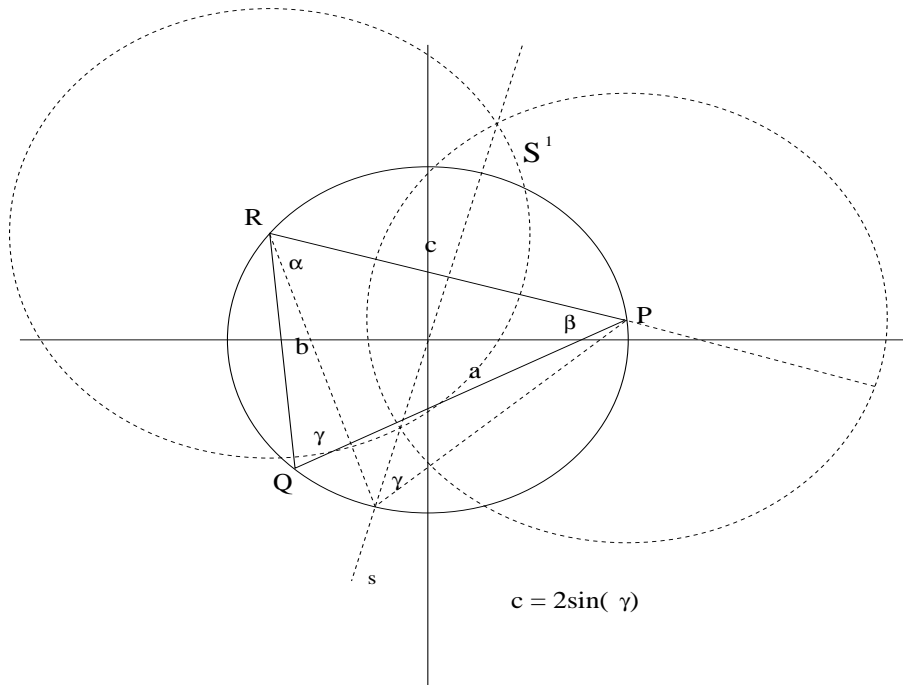


Figura 101: quando  $\gamma$  não for um ângulo reto

Na figura (fig 100) está detalhado o caso do ângulo  $\alpha$ . Então demonstramos assim a *lei do seno*

$$\frac{a}{\sin(\alpha)} = \frac{b}{\sin(\beta)} = \frac{c}{\sin(\gamma)} = 2 \quad (765)$$

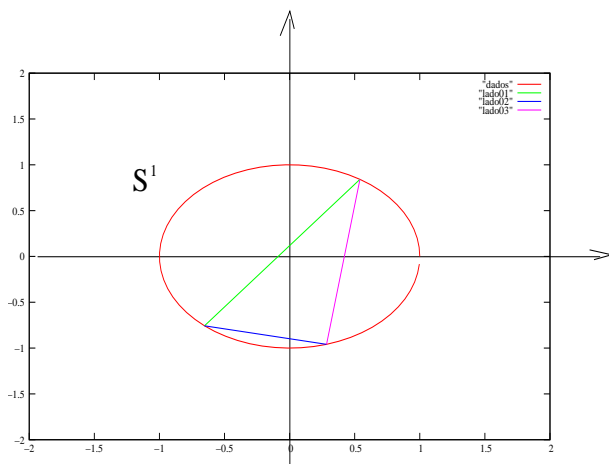
para o círculo trigonométrico. Como as medidas dos lados no triângulo original são recuperadas com o produto por  $R$ , o raio do círculo original, então a expressão da *lei do seno*, para um triângulo inscrito num círculo de raio  $R$  será

$$\frac{a}{\sin(\alpha)} = \frac{b}{\sin(\beta)} = \frac{c}{\sin(\gamma)} = 2R = D; \quad (766)$$

em que  $D = 2R$  é o diâmetro do círculo em que o triângulo estiver inscrito.

O valor comum destas razões dependem do raio do círculo em que o triângulo estiver inscrito, um *resultado colateral da lei do seno* é o valor do raio  $R$  dum círculo onde um triângulo esteja inscrito.

A figura (fig 102), página 257, mostra um exemplo de triângulo inscrito em  $\mathbf{S}^1$ . O gráfico foi obtido com o programa `LeiSeno.py` que pode ser baixado de [6, `LeiSeno.py`]. Este programa permite que você selecione valores arbitrários, dentro de  $\mathbf{S}^1$ , para os ângulos  $\alpha, \beta, \gamma \in \mathbf{S}^1$ , e vai lhe fazer o gráfico do triângulo inscrito em  $\mathbf{S}^1$  com estes ângulos. O programa também calcula o valor das razões na *lei do seno*.



Um triângulo inscrito no círculo trigonométrico  $S^1$

Figura 102: Triângulo inscrito em  $S^1$

No caso dum triângulo retângulo,  $PQR$ , inscrito em  $S^1$ , retângulo em  $Q$ ,  $\gamma = \frac{\pi}{2}$ , a hipotenusa,  $\overline{RP}$  mede  $c = 2$ , o tamanho do diâmetro do círculo trigonométrico, porque o arco que este ângulo subtende é  $2\gamma = \pi$ .

- **limite** é um funcional linear que produz um elemento completando um espaço. Esta não é a forma como este operador é tratado no Cálculo Diferencial e Integral onde ele aparece pela primeira vez nos estudos de Matemática.

No Cálculo Diferencial e Integral, limite aparece como uma metodologia para

- testar se uma função é contínua, confira *continuidade*;
- calcular derivadas, confira *derivada*;
- definir a integral de Riemann. confira integral de Riemann.
- mais avançado, limite de sucessões,
- em Análise Matemática, é uma forma de definir números reais.

Para entender como se processa esta completação, considere o exemplo do conjunto dos números racionais que é incompleto no sentido de que há sucessões de números racionais que são “convergentes”, e assim definem um número, mas este número não é um número racional. A sucessão  $P_n$  dos polígonos regulares inscritos num círculo de raio 1 permite ilustrar este caso de forma simples.

O quociente do perímetro de  $P_n$  pelo diâmetro do círculo se *aproxima arbitrariamente* dum número que os gregos chamaram de  $\pi$  que não é possível escrever como o quociente de dois inteiros e portanto não é um número racional. Desta forma

$$\lim_n \frac{P_n}{2} = \pi$$

e o operador  $\lim$  “*define*” um número que completa  $\mathbf{Q}$ . Confira real, número  $\mathbf{R}$

O conjunto de todos os números que completam  $\mathbf{Q}$  é o conjunto  $\mathbf{R}$  dos números reais, e, naturalmente,  $\mathbf{Q} \subset \mathbf{R}$ .

Outros exemplos mais simples de números não racionais são as raízes dos números naturais que, ou são números naturais, ou são número irracionais. O ou é exclusivo. Um algoritmo geométrico que faz esta construção pode ser encontrado em [8, Capítulo 5]

A definição rigorosa de limite foi feita por Bolzano, depois redescoberta por Weierstrass:

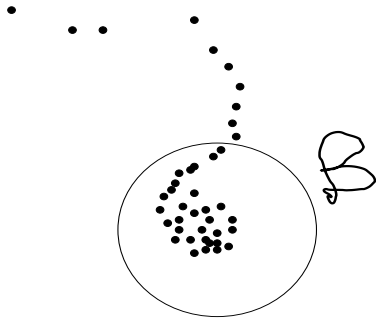
$$(\forall \epsilon > 0) (\exists N \in \mathbf{N}) (|n > N| \implies |s_n - l| < \epsilon) \quad (767)$$

quando dizemos sucintamente que  $s_n \rightarrow l$  ou também

$$\lim_{n=\infty} s_n = l \quad (768)$$

A igualdade com o símbolo  $\infty$ , na equação (eq. 768), é apenas uma maneira de indicar que  $n$  cresce indefinidamente.

A figura (fig 103), página 258, sugere o comportamento assintótico duma sucessão de pontos do plano, que, a partir do  $N$ -ésimo elemento, todos os elementos da sucessão se encontram dentro da bola  $\mathcal{B}$ . A partir da *índice*  $N$  todos



### comportamento assintótico duma sucessão de pontos do plano

Figura 103: Os pontos duma sucessão se acumulam numa bola

os pontos se encontram dentro duma bola de raio  $\epsilon$ . Se você exigir mais precisão, um valor menor para  $\epsilon$ , será preciso que  $N$  seja maior, comparando, seria

preciso que um computador rodasse mais *tempo* para encontrar um valor mais preciso.

## Alguns exemplos

1. Convergência para zero Se  $s_n = \frac{1}{n}$  então,  $\lim_n s_n = 0$ .  $N$  na (eq. 767) é um número natural maior do que  $\frac{1}{\epsilon}$ . Se  $\epsilon = 0.1$  então  $N \geq 11$ . Este exemplo é tão importante que vou reescrever a equação (eq. 767) para este caso:

$$(\forall \epsilon > 0) (\exists N \in \mathbf{N}) (|n > N| \implies |s_n| < \epsilon) \quad (769)$$

que define a convergência para zero.

2. Primeiro usando a aritmética  $s_n = \frac{n+1}{n}$  então,  $\frac{n+1}{n} = 1 + \frac{1}{n}$  pelas regras da aritmética. Pelas propriedades dos limite, confira o final do verbete, como  $\lim_n \frac{1}{n} = 0$  então  $\lim_n \frac{n+1}{n} = 1$  e vale a mesma relação entre  $\epsilon, N$  do item anterior.
3. Convergindo mais rápido para zero Se  $n > m > 0$  então  $\lim_k \frac{k^m}{k^n} = \lim_k \frac{1}{k^{n-m}}$  pelas regras da *aritmética*. Chame  $p = n - m$ , que é um número natural diferente de zero, então, para um inteiro  $k_0$

$$\frac{1}{k_0^p} < \epsilon \iff \frac{1}{\epsilon} < k_0^p; \quad (770)$$

$$(\exists N \in \mathbf{N}) \left( \frac{1}{\epsilon} < N \leq k_0^p \right); \quad (771)$$

$$(\exists N \in \mathbf{N}) \left( \sqrt[p]{\frac{1}{\epsilon}} < \sqrt[p]{N} \leq k_0 \right); \quad (772)$$

$$\sqrt[p]{\frac{1}{\epsilon}} < \sqrt[p]{N} \leq k_0 < k; \quad (773)$$

$$\frac{1}{\epsilon} < k^p \iff \frac{1}{k^p} < \epsilon; \quad (774)$$

$$(\forall k > \lceil \sqrt[p]{N} \rceil) \left( \frac{1}{k^p} < \epsilon \right) \quad (775)$$

O número  $\lceil \sqrt[p]{N} \rceil$  é o menor inteiro que é menor ou igual a  $\sqrt[p]{N}$ .

Então qualquer número inteiro positivo, maior do que  $\sqrt[p]{\frac{1}{\epsilon}}$ , serve no lugar de  $N$  na (eq. 767). Considere  $p = 2$ , e  $\epsilon = 0.01$  então

$$\sqrt{\frac{1}{\epsilon}} = \sqrt{100} \implies N \geq 11$$

Compare com o primeiro exemplo.

Observe a diferença entre os exemplos 1 e 3.

- No exemplo 1  $N \approx \frac{1}{\epsilon}$ ,

- enquanto que no exemplo 3,  $N \approx \sqrt[p]{\frac{1}{\epsilon}} < \frac{1}{\epsilon}$ .
- usando a comparação do computador, no exemplo 3 o computador precisa rodar *menos tempo* para conseguir a mesma precisão obtida no exemplo 1. Dizemos que no exemplo 3 a sucessão converge mais rápido para zero do que no exemplo 1.
- no exemplo 3 também estão sendo comparados os polinômios  $k^n, k^m$  em que  $n > m$ . Numa fração em que o polinômio do numerador tiver grau menor do que o polinômio do denominador se tem convergência para zero. Se o grau do numerador foi maior do que o grau do denominador se tem “convergência” para  $\infty$  e a maneira correta de falar é que se tem *divergência* porque  $\infty$  não é um número.

Estes exemplos mostram alguns aspectos interessantes no estudo dos limites. O primeiro é um exemplo dos mais comuns e relativamente imediato e junto com o terceiro, são exemplos da *classe do zero*, as sucessões que definem o número zero.

A importância da classe do zero é grande, o caso do exemplo 2 não é uma coincidência, com frequência precisamos de identificar sequências convergindo para zero quando tentamos descobrir qual é o limite duma sucessão.

A classe do zero é um grupo aditivo, quer dizer que se duas sucessões convergirem para zero, a soma delas também converge. É mais do que um grupo, é um ideal e volto logo a esta história.

É interessante observar, também, que a classe do 1 é um grupo multiplicativo! Quer dizer, se duas sucessões convergirem para 1 então o produto delas também converge para 1.

De forma semelhante é importante ter um estoque de exemplos de sucessões que definem inteiros ou números racionais porque nem sempre conseguimos calcular exatamente limites e uma saída consiste em encontrar um intervalo onde se encontre o limite estudado. Este “estoque” de limites conhecido é comumente referido como *limites notáveis*.

Foi isto que fiz ao estudar o segundo exemplo, usei o primeiro exemplo de uma forma que é preciso salientar:

- inicialmente foi usada a *aritmética* para que surgissem expressões cujos limites fossem conhecidos,
- depois foi produzido um *salto lógico*, usando as propriedades do limite (confira o final do verbete) para deduzir o limite da nova expressão que estava sendo estudada.

O *salto lógico* é o que caracteriza que o cálculo de limites não é uma operação da aritmética e neste cálculo podem aparecer entidades, como  $\infty$ , que não pertence à aritmética...

Um dos mais interessantes exemplos do uso repetido de operações aritméticas junto com saltos lógicos, é a determinação do número  $e \approx 2.71828182845904523536$ . Confira *limites notáveis*.

O conjunto dos números reais,  $\mathbf{R}$ , é a completção do conjunto dos números racionais e uma das formas de *construir* o conjunto  $\mathbf{R}$  consiste do seguinte programa:

- A imersão de  $\mathbf{Q}$  no conjunto das sucessões de números racionais. Identificamos  $\mathbf{Q}$  como um subconjunto do conjunto de todas as sucessões de números racionais associando-se  $a \in \mathbf{Q}$  com a sucessão constante  $s_n = a$  para todo  $n \in \mathbf{N}$ . Vou chamar de  $\mathcal{S}$  o conjunto de todas as sucessões de números racionais, então  $\mathbf{Q} \subset \mathcal{S}$ .
- Restrição do conjunto de todas as sucessões ao conjunto das sucessões que satisfizerem ao critério de Cauchy. Dizemos que uma sucessão  $(s_n)_{n \in \mathbf{N}}$  é *de Cauchy* se for verdadeiro

$$(\forall \epsilon > 0) (\exists N \in \mathbf{N}) n, m > N \implies |s_n - s_m| < \epsilon \quad (776)$$

Tais sucessões são *intrinsecamente* convergentes uma vez que a partir do índice  $N$  todos os termos da sucessão se encontram dentro da *bola* de centro  $s_n$  e raio  $\epsilon$ . Isto significa, exatamente, que  $s_n$  é um *valor aproximado do limite*.

- Vou reutilizar a notação  $\mathcal{S}$  para o conjunto de todas as sucessões de Cauchy de números racionais, então  $\mathbf{Q} \subset \mathcal{S}$ , porque uma sucessão constante satisfaz ao critério de Cauchy com qualquer valor estritamente positivo para  $\epsilon$ .
- Podemos mostrar que  $(\mathcal{S}, +, \cdot)$  é um anel que tem divisores de zero, mas podemos eliminar os divisores de zero passando ao quociente com *grupo das sucessões que convergem para zero*, a classe do zero. Como este grupo é um *ideal maximal* de  $\mathcal{S}$  o quociente é um corpo: o corpo  $\mathbf{R}$  dos números reais.

Observe que a última ação pode parecer refinada, mas ela é comum com os números racionais desde o ensino fundamental quando se definem frações equivalentes. . . aqui o objetivo é o mesmo apenas com uma pintura algébrica.

Este programa, que é atribuído à Cauchy, é bem natural pese que considerado *avançado*, possivelmente pelos aspectos algébricos que envolve. Mas o que se espera dum número real é que ele seja uma sucessão de medidas convergentes. Desta forma  $\mathbf{R}$  nada mais é do que o conjunto dos limites das sucessões de números racionais. Os limites são as etiquetas das classes quocientes mencionadas no *programa* acima. Como  $\mathbf{R}$  é um corpo, acabei de “demonstrar” os teoremas básicos do limite, as propriedades do limite.

## Propriedades do limite

- soma de limites é o limite da soma,
- produto de limites é o produto dos limites,
- como somente podemos dividir dois número reais se o segundo for diferente de zero, então o quociente de limites é o limite dos quocientes se o limite do quociente for diferente de zero.

O operador  $\lim$  é um funcional linear! Valem todas as propriedades duma transformação linear. As propriedades do limite, são as propriedades dos números reais.

O número  $s_n$ , na equação (eq. 776) pode ser considerado um valor aproximado do limite da sucessão a menos do erro  $\epsilon$ . Desta forma o critério de Cauchy é um *instrumento prático* para encontrar aproximadamente o limite duma sucessão quando ela for convergente e o critério de Cauchy serve para verificar se sucessão é convergente.

## Infinitésimos

Até o começo do século passado havia uma palavra muito frequente em Matemática, *infinitésimos*. Os livros de Cálculo se chamavam *Cálculo integral e infinitesimal* em que os *infinitésimos* eram as sucessões convergindo para zero. Quer dizer, as sucessões que definem o zero.

Foi preciso passar muito tempo até que os matemáticos compreendessem o simples que era: apenas uma classe de equivalência, a classe de equivalência do zero. O exemplos 1 e 3 trazem dois elementos desta classe e o que atrapalhava mesmo a compreensão é que se dividirmos o “zero”  $\frac{1}{k^m}$  pelo “zero”  $\frac{1}{k^n}$  o resultado pode ser “zero” ou  $\infty$ , dependendo se o maior grau estiver ou não, no denominador. Hoje apenas dizemos que são elementos com ordem de grandeza diferentes e a palavra *infinitésimos* está cuidadosamente guardada na história da Matemática.

E também ninguém deve dividir por zero... pode dar problema!

Confira também *limite superior*, *limite inferior*

- **limite inferior** é uma generalização do conceito de limite. Uma sucessão pode não ter limite mas ter um ponto de acumulação inferior que é o *limite inferior*. Para defini-lo se recorre a uma sucessão crescente dos ínfimos cujo limite é *limite inferior*. Considere uma sucessão  $(s_n)_{n \in \mathbf{N}}$  de números reais e defina

$$t_n = \inf_{k \geq n} s_k; \quad (777)$$

$$t_n \text{ é crescente; } \lim t_n = \lim \inf_n s_n \quad (778)$$



Obviamente que o limite inferior pode ser  $\infty$  e seria o caso da sucessão dos números naturais.

Por exemplo,  $\liminf_n (-1)^n = -1$  ou a sucessão  $\limsup_n (-1)^n = -\infty$ .

Se  $\liminf_n = \limsup_n$  então a sucessão tem limite que este valor comum.

- **limite superior** é uma generalização do conceito de limite. Uma sucessão pode não ter limite mas ter um ponto de acumulação superior que é o *limite superior*. Para defini-lo se recorre à sucessão decrescente dos supremos cujo limite é *limite superior*. Considere uma sucessão  $(s_n)_{n \in \mathbf{N}}$  de números reais e defina

$$t_n = \sup_{k \geq n} s_k; \quad (779)$$

$$t_n \text{ é decrescente}; \quad (780)$$

$$\lim_n t_n = \limsup_n s_n \quad (781)$$

Uma sucessão tem limite *se e somente se*

$$\limsup_n s_n = \overline{\lim}_n = \lim_{\underline{n}} s_n = \lim_n s_n \quad (782)$$

Obviamente que o limite superior pode ser  $\infty$  e seria o caso da sucessão dos números naturais e neste caso a sucessão  $t_n$  na equação (eq. 779) é constante igual a  $\infty$  e portanto “decrescente”... Neste caso valeria a igualdade da equação (eq. 782) permitindo-nos dizer que  $\infty$ . O que nos salva desta aberração é que não podemos escrever igualdades envolvendo o símbolo  $\infty$  porque ele não representa uma entidade aritmética.

Por exemplo,  $\limsup_n (-1)^n = 1$  ou a sucessão  $\limsup_n (-1)^n = \infty$ .

- **linear, função** É o *tipo natural* de função dum espaço vetorial, ou ainda é um morfismo de espaço vetorial.

Confira também *linear complex, função*, as funções lineares complexas tem propriedades importantes para as funções analíticas.

As características deste morfismo é colocar em correspondência os elementos da estrutura dum par de espaços vetoriais que são os objetos da categoria EV. Sejam  $E, F \in EV$  e  $T \in Mor_{E,F}$  então, supondo que sejam espaços vetoriais reais

$$x, y \in E; \alpha, \beta \in \mathbf{R} \implies T(\alpha x + \beta y) = \alpha T(x) + \beta T(y); \quad (783)$$

Com frequência separamos esta propriedade em duas partes:

- linearidade  $x, y \in E; \implies T(x + y) = T(x) + T(y);$
- homogeneidade  $x \in E; \alpha \in \mathbf{R} \implies T(\alpha x) = \alpha T(x)$

O núcleo de  $T$ , com frequência referência pelas notações  $Ker(T)$ ,  $Nuc(T)$  é o subespaço de  $E$  que é a pré-imagem de zero:

$$Ker(T) = T^{-1}(0) \subset E; \quad (784)$$

- **linear complexa, função.** Quando o corpo subjacente ao espaço vetorial for dos números complexos, as funções lineares complexas formam uma sub-algebra caracterizada pelas equações de Cauchy-Riemann.

No caso univariado, das funções lineares complexas de  $\mathbf{C}$ , que é um espaço vetorial de dimensão 1 sobre o corpo dos números complexos, as funções lineares complexas formam um sub-algebra de  $\mathcal{L}_{\mathbf{R}}(\mathbf{C}, \mathbf{C})$ .

Observe que  $\mathbf{C}$  é isomorfo a  $\mathbf{R}^2$  se o corpo subjacente for  $\mathbf{R}$  e as funções lineares complexas podem ser vistas como morfismos de  $\mathbf{R}^2$  mas como morfismo de  $\mathbf{C}$  elas são simples multiplicação por um número complexo o que leva às equações:

$$\mathcal{L}(\mathbf{R}^2) \ni \begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \approx (\alpha - i\beta) \in \mathbf{C} \quad (785)$$

$$\begin{bmatrix} \alpha & \beta \\ -\beta & \alpha \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \equiv (\alpha - i\beta)(x + iy); \quad (786)$$

A equação (eq.785) sendo consequência da (eq.786).

- **logaritmo** é o nome de uma família de funções estudadas intensivamente, pelo menos desde 1614, confira [12, logaritmo]. Os logaritmos estavam fundamentados na relação  $a^x a^y = a^{x+y}$  em que as potências transformam produto em adição. Foram usados de modo sistemático até a década de 70 do século passado quando as máquinas eletrônicas lhes tomaram o lugar como “máquinas de calcular” depois de 350 anos dum *serviço exemplar*.

No final do século 17 foi construída uma máquina de cálculo, a *régua de cálculo*, “*slide rule*”, em inglês, baseada no princípio da correspondência entre as duas progressões, apenas usando uma distribuição logarítmica para as marcações numéricas, como você pode ver nas figuras (fig 104), (fig 105) página 265.

Até 1960, as chamadas *tabelas de logaritmos*, eram indispensáveis nas escolas e todo estudante tinha a sua *tabela*, ou a sua *régua de cálculo*. Foi uma das invenções mais prolíficas da Matemática, foi a “máquina de calcular” usada pelos *calculistas do final Idade Média*, descritos no livro, *Mirifici Logarithmorum Canonis Descriptio* por John Napier, em 1614, e ainda estavam em uso até a primeira metade do século 20.

Os *calculistas* descobriram o *segredo dos logaritmos*: colocar em correspondência duas progressões, uma geométrica e a outra aritmética com a sincronização dos elementos neutros da multiplicação e da adição.

Um exemplo simples vem quando você escrever as sucessivas potências de um número  $a > 0$ , as potências formam uma progressão aritmética, e o resultado de elevar  $a$  a cada uma dessas potências é uma progressão geométrica.

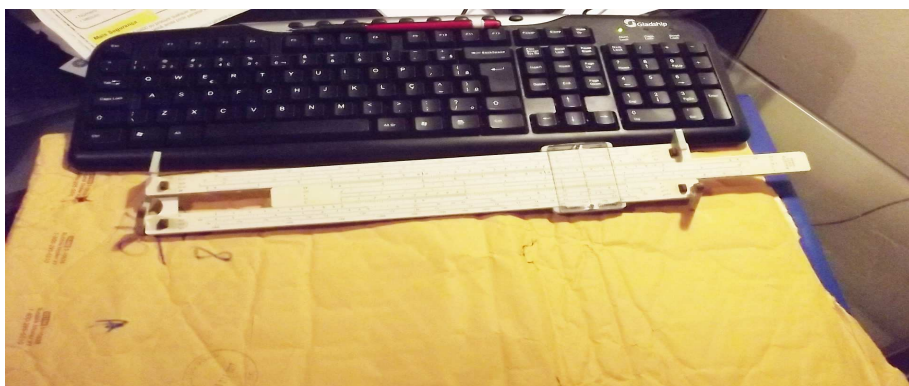
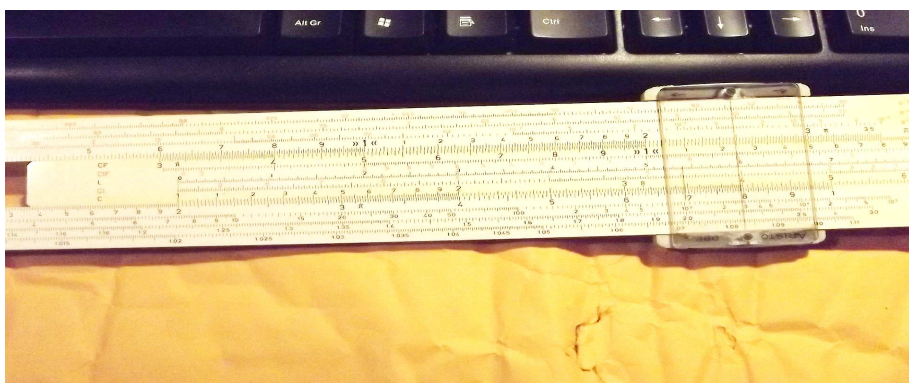


Figura 104: Régua de cálculo

Figura 105:  $2 \times 3 = 6$ 

Os cálculistas da Idade Média descobriram isto e começaram a colocar longas listas de p.g. sincronizadas com p.a. e depois somando os termos na p.a. aritmética podiam descobrir quanto valia o produto dos números em correspondência.

Experimente, na página 268 você encontra uma tabela de logaritmos . Se você quiser efetuar o produto dos números 1.1651390, 1.1973770, leia os logaritmos que lhe correspondem na coluna ao lado dos mesmos:

- $1.1651390 \mapsto 0.175 = \log(1.1651390)$
- $1.1973770 \mapsto 0.20625 = \log(1.1973770)$
- $0.175 + 0.20625 = 0.38125 \leftrightarrow 1.3951107 = 1.1651390 * 1.1973770$

Usando calc, uma linguagem de programação de domínio público, o resultado que encontro é

$$1.1651390 * 1.1973770 = 1.395110640403$$

portanto há um erro na sétima casa decimal usando esta tabela de logaritmos. Com uma tabela mais precisa este erro pode ser mais reduzido.

Uma tabela de logaritmos, essencialmente, é a listagem de duas progressões, uma *geométrica*, a coluna do  $x$  e a outra *aritmética*, a coluna de  $\log(x)$ , sincronizadas pela associação

$$1 \mapsto 0;$$

pondo em correspondência os elementos neutros da adição e da multiplicação. Qualquer tabela deste tipo é uma tabela de logaritmos e no ponto em que surgir a associação

$$a \mapsto 1;$$

se tem a *base*  $\underline{a}$  dos logaritmos desta tabela. E você não precisa saber qual é a *base* para usar os logaritmos, basta construir as duas progressões com grande precisão:

- Fixe a associação fundamental,  $1 \mapsto 0$ ;
- e na coluna do 1 divida ou multiplique por  $q > 0$ , *muito pequeno*;
- na coluna do 0 subtraia ou some um  $\Delta$  *muito pequeno, em valor absoluto*;
- e mantenha as duas sucessões *acopladas* na relação fundamental.

Não interessa nem mesmo se você alterar a ordem de crescimento entre as duas colunas isto significaria apenas que você inverteu a base do logaritmo, de  $a$ , para  $\frac{1}{a}$  . . .

Os dois números *muito pequenos*  $q > 0, \Delta$  representam a precisão da tabela construída. Se você puder encontrar  $a \mapsto 1$  você fica sabendo qual é a base destes logaritmos. Mas esta descoberta não interessa para fazer as contas com os logaritmos, e esta segunda associação não é sempre possível de ser obtida ao colocarmos em correspondência as duas progressões, o caso do número  $a = e$  ilustra bem esta dificuldade uma vez que  $\underline{e}$  é um dos poucos números que *não é algébrico* cuja *identidade nós conhecemos*. Não é possível encontrar-se um par de sucessões com as duas associações:

$$1 \mapsto 0; e \mapsto 1;$$

a segunda associação tem que ser feita com um número que representa uma aproximação do número  $\underline{e}$ .

As tabelas de *logaritmos decimais* foram as mais comuns, nelas se tinha o par de associações

$$1 \mapsto 0; 10 \mapsto 1;$$

e se você quiser construir uma tabela de logaritmos decimais, sincronize as progressões usando a relação  $10 \mapsto 1$  usando os dois números  $q, \Delta$  mencionados acima. Se quiser os logaritmos de base  $a$  sincronize as progressões usando a associação  $a \mapsto 1$ .

Na tabela , na página 268, você pode ver uma tabela de logaritmos que foi gerada por um programa em `python` que pode ser baixado de [6, log\_tabela\_py].

Para executar o programa, troque o nome de `log_tabela.py` para `log_tabela.py` para que o *interpretador* do `python` o reconheça.

Com este programa em `python` você pode construir uma tabela de logaritmos de alta precisão e bastante extensa. Imprima e guarde para quando não tivermos mais computadores disponíveis quando retornar aos métodos de cálculo da Idade Média ou de muito antes... se ainda soubermos ler, as tabelas serão úteis.

A correspondência entre progressões geométrica e aritmética mostra a impossibilidade de que se calcule  $p = \log(0)$  porque não é possível escrever a potência  $a^p = 0$ . Desta forma o domínio de qualquer função logarítmica real é o conjunto dos números reais estritamente positivos.

Na época em que Napier divulgou os logaritmos, eles foram objeto de diversas pesquisas um dos pontos altos certamente foi a demonstração por Euler dos dois *limites notáveis*

$$\begin{cases} e^x = \lim_n (1 + x/n)^n; \\ \log(x) = \lim_n n(x^{1/n} - 1); \end{cases} \quad (787)$$

definindo um par de funções inversas.

A forma moderna como se apresentam os logaritmos passa pela definição:

$$\log(x) = \int_1^x \frac{dt}{t} \quad (788)$$

com a qual é bem simples provar que  $\log(ab) = \log(a) + \log(b)$  para dois números reais positivos quaisquer, e, por definição mesmo,  $\log(1) = 0$ .

Depois é relativamente fácil encontrar a *solução aproximada* para a equação

$$\int_1^x \frac{dt}{t} = 1 \Rightarrow x = e \quad (789)$$

usando algum método para o cálculo aproximado da integral.

Como o integrando é uma função indefinidamente derivável então  $\log(x)$  também o é. Depois como a derivada de  $\log(x)$  é  $\frac{1}{x}$  então  $\log(x)$  é uma função crescente, logo inversível e podemos chamar a sua inversa de  $\exp(x)$  e com elas duas é possível obter as duas progressões, a aritmética e a geométrica, referidas acima com as associações:

$$1 \mapsto 0; e \mapsto 1;$$

apenas a segunda ficando em aberto...mas nós sabemos onde está o número  $e$ . Mesmo o *atalho moderno*, representado pela equação (eq. 788), não torna muito mais fácil provar as duas identidades da equação (787).

Por outro lado é a forma como se pode facilmente demonstrar as propriedades de diferenciabilidade e continuidade do *logaritmo natural* e conseqüentemente da sua função inversa  $y = e^x$ . Mas, depois de estabelecida, facilmente se mostra que ela leva de volta à correspondência entre as progressões da tabela do *logaritmo natural*, ou ainda chamado *logaritmo neperiano*.

x	log x	x	log x	x	log x	x	log x
1	0	1.2440142	0.25	1.5475715	0.5	1.9252010	0.75
1.0054735	0.00625	1.2508234	0.25625	1.5560421	0.50625	1.9357386	0.75625
1.0109769	0.0125	1.2576698	0.2625	1.5645591	0.5125	1.9463339	0.7625
1.0165105	0.01875	1.2645536	0.26875	1.5731228	0.51875	1.9569872	0.76875
1.0220744	0.025	1.2714752	0.275	1.5817333	0.525	1.9676988	0.775
1.0276687	0.03125	1.2784346	0.28125	1.5903909	0.53125	1.9784690	0.78125
1.0332937	0.0375	1.2854321	0.2875	1.5990959	0.5375	1.9892982	0.7875
1.0389495	0.04375	1.2924680	0.29375	1.6078486	0.54375	2.0001866	0.79375
1.0446362	0.05	1.2995423	0.3	1.6166492	0.55	2.0111347	0.8
1.0503540	0.05625	1.3066554	0.30625	1.6254979	0.55625	2.0221427	0.80625
1.0561031	0.0625	1.3138074	0.3125	1.6343951	0.5625	2.0332109	0.8125
1.0618837	0.06875	1.3209985	0.31875	1.6433410	0.56875	2.0443397	0.81875
1.0676959	0.075	1.3282290	0.325	1.6523358	0.575	2.0555294	0.825
1.0735400	0.08125	1.3354991	0.33125	1.6613799	0.58125	2.0667804	0.83125
1.0794160	0.0875	1.3428089	0.3375	1.6704735	0.5875	2.0780929	0.8375
1.0853242	0.09375	1.3501588	0.34375	1.6796169	0.59375	2.0894674	0.84375
1.0912647	0.1	1.3575489	0.35	1.6888103	0.6	2.1009041	0.85
1.0972378	0.10625	1.3649795	0.35625	1.6980540	0.60625	2.1124034	0.85625
1.1032435	0.1125	1.3724507	0.3625	1.7073483	0.6125	2.1239657	0.8625
1.1092822	0.11875	1.3799629	0.36875	1.7166935	0.61875	2.1355912	0.86875
1.1153538	0.125	1.3875161	0.375	1.7260899	0.625	2.1472804	0.875
1.1214587	0.13125	1.3951107	0.38125	1.7355376	0.63125	2.1590336	0.88125
1.1275971	0.1375	1.4027468	0.3875	1.7450371	0.6375	2.1708511	0.8875
1.1337690	0.14375	1.4104248	0.39375	1.7545886	0.64375	2.1827333	0.89375
1.1399747	0.15	1.4181448	0.4	1.7641924	0.65	2.1946805	0.9
1.1462143	0.15625	1.4259070	0.40625	1.7738487	0.65625	2.2066931	0.90625
1.1524881	0.1625	1.4337117	0.4125	1.7835579	0.6625	2.2187715	0.9125
1.1587963	0.16875	1.4415592	0.41875	1.7933202	0.66875	2.2309159	0.91875
1.1651390	0.175	1.4494496	0.425	1.8031360	0.675	2.2431269	0.925
1.1715164	0.18125	1.4573831	0.43125	1.8130054	0.68125	2.2554047	0.93125
1.1779287	0.1875	1.4653601	0.4375	1.8229289	0.6875	2.2677496	0.9375
1.1843761	0.19375	1.4733808	0.44375	1.8329068	0.69375	2.2801622	0.94375
1.1908588	0.2	1.4814454	0.45	1.8429392	0.7	2.2926427	0.95
1.1973770	0.20625	1.4895541	0.45625	1.8530266	0.70625	2.3051915	0.95625
1.2039308	0.2125	1.4977072	0.4625	1.8631691	0.7125	2.3178090	0.9625
1.2105206	0.21875	1.5059049	0.46875	1.8733672	0.71875	2.3304956	0.96875
1.2171464	0.225	1.5141475	0.475	1.8836211	0.725	2.3432515	0.975
1.2238084	0.23125	1.5224352	0.48125	1.8939311	0.73125	2.3560774	0.98125
1.2305070	0.2375	1.5307683	0.4875	1.9042976	0.7375	2.3689734	0.9875
1.2372422	0.24375	1.5391469	0.49375	1.9147208	0.74375	2.3819400	0.99375

x	log x	x	log x	x	log x	x	log x
2.3949776	1.0	2.9793863	1.25	3.7063991	1.5	4.6108134	1.75
2.4080865	1.00625	2.9956940	1.25625	3.7266861	1.50625	4.6360507	1.75625
2.4212672	1.0125	3.0120910	1.2625	3.7470842	1.5125	4.6614262	1.7625
2.4345200	1.01875	3.0285777	1.26875	3.7675939	1.51875	4.6869406	1.76875
2.4478454	1.025	3.0451546	1.275	3.7882159	1.525	4.7125946	1.775
2.4612437	1.03125	3.0618223	1.28125	3.8089507	1.53125	4.7383891	1.78125
2.4747154	1.0375	3.0785813	1.2875	3.8297990	1.5375	4.7643247	1.7875
2.4882607	1.04375	3.0954319	1.29375	3.8507615	1.54375	4.7904023	1.79375
2.5018803	1.05	3.1123748	1.3	3.8718387	1.55	4.8166226	1.8
2.5155743	1.05625	3.1294104	1.30625	3.8930312	1.55625	4.8429865	1.80625
2.5293434	1.0625	3.1465393	1.3125	3.9143398	1.5625	4.8694946	1.8125
2.5431878	1.06875	3.1637619	1.31875	3.9357650	1.56875	4.8961478	1.81875
2.5571079	1.075	3.1810788	1.325	3.9573074	1.575	4.9229470	1.825
2.5711043	1.08125	3.1984905	1.33125	3.9789678	1.58125	4.9498928	1.83125
2.5851773	1.0875	3.2159974	1.3375	4.0007467	1.5875	4.9769861	1.8375
2.5993273	1.09375	3.2336002	1.34375	4.0226449	1.59375	5.0042277	1.84375
2.6135547	1.1	3.2512994	1.35	4.0446629	1.6	5.0316184	1.85
2.6278600	1.10625	3.2690954	1.35625	4.0668014	1.60625	5.0591590	1.85625
2.6422437	1.1125	3.2869888	1.3625	4.0890611	1.6125	5.0868504	1.8625
2.6567060	1.11875	3.3049802	1.36875	4.1114426	1.61875	5.1146933	1.86875
2.6712475	1.125	3.3230701	1.375	4.1339466	1.625	5.1426886	1.875
2.6858686	1.13125	3.3412589	1.38125	4.1565738	1.63125	5.1708372	1.88125
2.7005698	1.1375	3.3595474	1.3875	4.1793249	1.6375	5.1991398	1.8875
2.7153514	1.14375	3.3779359	1.39375	4.2022005	1.64375	5.2275974	1.89375
2.7302139	1.15	3.3964250	1.4	4.2252013	1.65	5.2562107	1.9
2.7451577	1.15625	3.4150154	1.40625	4.2483280	1.65625	5.2849806	1.90625
2.7601834	1.1625	3.4337075	1.4125	4.2715812	1.6625	5.3139080	1.9125
2.7752913	1.16875	3.4525020	1.41875	4.2949618	1.66875	5.3429938	1.91875
2.7904819	1.175	3.4713993	1.425	4.3184703	1.675	5.3722387	1.925
2.8057556	1.18125	3.4904000	1.43125	4.3421075	1.68125	5.4016437	1.93125
2.8211129	1.1875	3.5095048	1.4375	4.3658741	1.6875	5.4312097	1.9375
2.8365543	1.19375	3.5287141	1.44375	4.3897707	1.69375	5.4609375	1.94375
2.8520803	1.2	3.5480286	1.45	4.4137982	1.7	5.4908280	1.95
2.8676912	1.20625	3.5674487	1.45625	4.4379572	1.70625	5.5208821	1.95625
2.8833875	1.2125	3.5869752	1.4625	4.4622484	1.7125	5.5511007	1.9625
2.8991698	1.21875	3.6066086	1.46875	4.4866726	1.71875	5.5814847	1.96875
2.9150384	1.225	3.6263494	1.475	4.5112304	1.725	5.6120351	1.975
2.9309939	1.23125	3.6461983	1.48125	4.5359227	1.73125	5.6427526	1.98125
2.9470367	1.2375	3.6661558	1.4875	4.5607501	1.7375	5.6736383	1.9875
2.9631674	1.24375	3.6862225	1.49375	4.5857134	1.74375	5.7046930	1.99375

A notação varia um pouco, alguma vezes o *logaritmo natural* é designado por  $\ln(x)$  mas entendemos que  $\log(x)$  quer dizer o *logaritmo neperiano*. Quando se tratar de outra *base* o hábito é indicá-lo:

- $\log_a(x)$  o logaritmo de  $x$  na base  $a$ ;
- $\log_{10}(x)$  o logaritmo decimal de  $x$ .

As *linguagens de programação “entendem”* que  $\log(x)$  é o logaritmo natural. Para obter os outros tipos de logaritmo podemos usar a equação:

$$\log_a(x) = \int_1^x \frac{k}{t} dt = k \log(x) \quad (790)$$

e podemos descobrir o número  $k$  que corresponde a uma *base de logaritmos* usando a própria base. Por exemplo, para encontrar a constante que irá reproduzir o  $\log_{10}$  escreva  $x = 10$  na equação (790):

$$\log_{10}(10) = 1 = \int_1^{10} \frac{k}{t} dt = k \log(10) = 2.625 * k \quad (791)$$

$$k = \frac{1}{\log(10)} = \frac{1}{\ln(10)}; \quad (792)$$

$$\log_{10}(x) = \frac{\ln(x)}{\ln(10)}; \ln(x) = \ln(10) \log_{10}(x); \quad (793)$$

Deste exemplo você vê que

$$\log_a(x) = C \ln(x); \quad (794)$$

$$\log_a(a) = 1 = C \ln(a) \Rightarrow C = \frac{1}{\ln(a)}; \quad (795)$$

$$\log_a(x) = \frac{\ln(x)}{\ln(a)}; \quad (796)$$

$$\log_a(x) = \frac{\ln(x)/\ln(b)}{\ln(a)/\ln(b)} = \frac{\log_b(x)}{\log_b(a)}; \quad (797)$$

$$\log_a(x) = \frac{\log_{10}(x)}{\log_{10}(a)}; \quad (798)$$

$$\log_a(x) = \frac{\log(x)}{\log(a)}; \quad (799)$$

a fórmula muito conhecida de transformação do logaritmo na base  $a$  a partir do *logaritmo decimal*. A multiplicação e divisão pelo mesmo número na equação (eq. 797) mostra que a mesma expressão vale para qualquer logaritmo em lugar do logaritmo decimal o que me permite escrever a fórmula ambígua, (eq. 799) mas correta, usando a notação  $\log$  querendo dizer “escolha a base que desejar neste fórmula”...

A definição na equação (eq. 788) dá um exemplo da *ruptura cultural* entre a Matemática atual e aquela que se fazia no século 15. Esta equação aparece *abruptamente* dentro dos livros de Cálculo sem aparente ligação com os sistemas de progressões aritméticas e geométricas do século 15. Pior, nada tem o que ver a forma arcaica com que as professoras apresentam logaritmo no Ensino



Médio em que se vêm forçadas a algumas ginásticas mentais para manter nesta etapa do ensino algo que simplesmente já deveria ter desaparecido uma vez que perdeu a sua razão de ser na forma como se tenta mostrar. Se apenas as alunas tivessem voz ativa e senso crítico, coisa que o *ensino* não objetiva estimular, a pergunta natural seria “*para que serve logaritmo?*” que deixaria as professoras entaladas com a perda de sentido calculatório do logaritmo. Ou talvez a resposta fosse: *porque eles aparecem obrigatoriamente nos exames oficiais do Ministério da Educação . . .* numa demonstração de que os burocratas que dirigem o *ensino* precisavam mesmo desaparecer. . . afinal quem entende de *ensino* são as professoras.

Os logaritmos perderam o seu posto como “*máquina de calcular*” mas adquiriram uma posição muito mais proeminente, eles descrevem diversas relações importantes para as ciências naturais, na Biologia, na Física, na Química e até mesmo na Economia. Esta outra forma de usar os logaritmos é que justifica que apareçam nos livros de Cálculo usando a equação (eq. 788) como definição e representando um caminho rápido para estudar  $\log$ ,  $\exp$  como um par de funções inversas, seus gráficos e suas propriedades, e até mesmo as tabelas de logaritmo. Mais outra forma de ver a *Idade Média* em que se encontra o *Ensino Médio*.

### Observação 6 (A questão) histórica

*A proposta de que o logaritmo desapareça do Ensino Médio pode se bater com a questão da história. Eu acho que existe historicismo em certas tentativas pedagógicas. É fácil demonstrar que é impossível construirmos o conhecimento usando o método histórico, porque não haveria tempo hábil para fazê-lo. . .*

*Os objetos históricos devem ser preservados em museus que são locais que devemos visitar regularmente e desta forma, com certeza, criar motivações para os estudos. Sala de aula não pode ser considerada parte do museu, mas pode, se houver razões, ser transferida para um ambiente de museu. O currículo é que não pode ser sacrificado para incluir todo o conhecimento acumulado pela Humanidade.*

*E logaritmo como máquina de calcular pertence à história, é um artigo de museu.*

E aqui vem um salto importante: podemos definir logaritmo para os números complexos?

É muito interessante a discussão e a definição das funções complexas, logaritmo e exponencial:

$$e^z, \log(z); z \in \mathbf{C} \quad (800)$$

mas, passando para as funções complexas, há um salto, dentro do currículo universitário de dois anos, pelo menos, de afastamento do ponto em que os logaritmos aparecem no Cálculo. Entretanto serão as funções complexas que vão mostrar mais claramente porque não temos logaritmos de números negativos na definição do caso real.

Ainda no caso complexo temos que fazer restrições ao domínio do logaritmo, por exemplo não podemos definir  $p = \log(0)$  porque também é impossível obter-se  $e^p = 0$ . A forma de definir logaritmos para os número complexos passa pela fórmula de Euler:  $e^{it} = \cos(t) + i \sin(t)$ . Com algum trabalho típico da Geometria Analítica podemos demonstrar que a fórmula de Euler se comporta

como uma exponencial. A sequência de equações

$$e^{i(a+b)} = \cos(a+b) + i \sin(a+b) \in \mathbf{S}^1; \quad (801)$$

$$e^{i(a+b)} = e^{i(a)} e^{i(b)} = (\cos(a) + i \sin(a))(\cos(b) + i \sin(b)); \quad (802)$$

$$e^{i(a+b)} = \cos(a) \cos(b) - \cos(a) \sin(b) + i (\cos(a) \sin(b) + \cos(b) \sin(a)) \quad (803)$$

$$\cos(a+b) = \cos(a) \cos(b) - \cos(a) \sin(b); \quad (804)$$

$$\sin(a+b) = \cos(a) \sin(b) + \cos(b) \sin(a); \quad (805)$$

$$(\cos(a) + i \sin(a)), (\cos(b) + i \sin(b)) \in \mathbf{S}^1; \quad (806)$$

são *verdades simples* à partir da fórmula de Euler. Como os números complexos de módulo 1 são *estáveis* frente à multiplicação, então o produto na equação (eq. 802) se mantém no círculo unitário  $\mathbf{S}^1$ , que assim se verifica ser um subgrupo multiplicativo do grupo multiplicativo dos complexos diferentes de zero. Depois, uma *ingeniosa construção geométrica*, na verdade da Geometria Analítica, mostra que o produto corresponde à soma de arcos da equação (eq. 803). Com isto se prova que a fórmula de Euler é uma *exponencial verdadeira*.

Agora posso terminar a definição da exponencial complexa

$$e^{x+iy} = e^x e^{iy} = e^x (\cos(y) + i \sin(y)); \quad (807)$$

que representa uma extensão da fórmula de Euler<sup>49</sup> ao plano complexo, é a definição da exponencial complexa.

Considerando a fórmula polar podemos escrever

$$w = \rho e^{i\theta} = e^{x+iy}; \text{Log}(w) = \log(\rho) + i\theta; \quad (808)$$

$$w = 1 - i; \text{Log}(w) = \frac{\log(2)}{2} - i\frac{\pi}{4}; \quad (809)$$

$$w = a \in \mathbf{R}^{++}; \text{Log}(w) = \log(a); \quad (810)$$

$$\text{Log}(w_1 w_2) = \log(\|w_1 w_2\|) + i \text{Arg}(w_1 w_2) \quad (811)$$

$$\text{Log}(w_1 w_2) = \log(\|w_1\|) + \log(\|w_2\|) + i (\text{Arg}(w_1) + \text{Arg}(w_2)) \quad (812)$$

$$\text{Log}(w_1 w_2) = \text{Log}(w_1) + \text{Log}(w_2) \quad (813)$$

A equação (eq. 808) é uma definição! A soma de arcos que aparece na equação (eq. 812), vem do produto de números complexos usando a forma polar, em outras palavras, propriedade da exponencial.

Algumas vezes se usa a notação  $\text{Log}(w)$  para fazer referência a expressão *principal* do logaritmo associado à faixa de largura  $[-\pi, \pi)$  da exponencial. Mas não há uma convenção estabelecida neste sentido. A parte imaginária do logaritmo é chamada de *argumento*.

Consequência da equação (eq. 807), a função exponencial é periódica na variável  $y$  se repetindo em faixas de largura  $2\pi$  portanto deixando de ser injetiva. Para que o logaritmo tenha inversa é preciso fazer uma restrição da exponencial a uma faixa de largura  $2\pi$ . A figura (fig 106), página 273, mostra alguma das "*relações geométricas*" que podemos obter de forma imediata a partir da (eq. 807):

<sup>49</sup>Também é a *fórmula polar* dum número complexo usualmente escrita como  $w = e^{x+iy} = e^x e^{iy} = \rho e^{i\theta}$

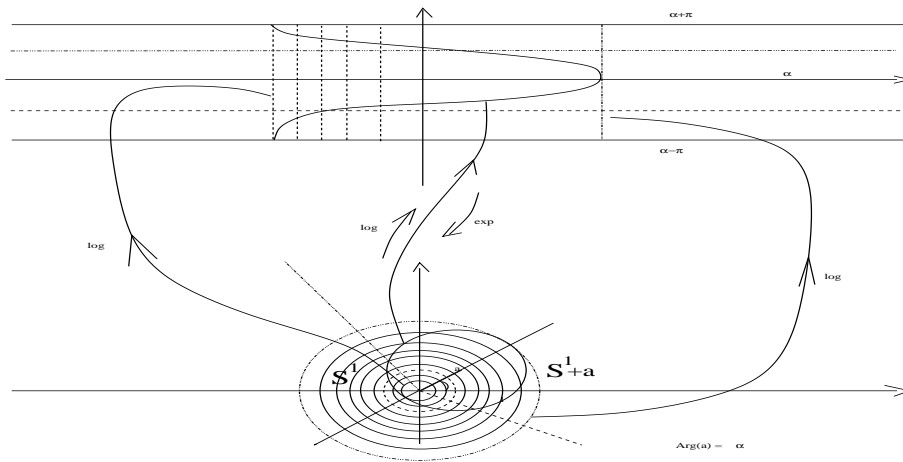


Figura 106: Exponencial e logaritmo no plano complexo

- A imagem duma semireta partindo da origem, pela função logaritmo, é uma reta horizontal no domínio da  $exp$  então para que se tenha uma função exponencial injetiva, que tenha inversa, temos que eliminar do domínio da função logaritmo uma semireta partindo de zero e inclusive o zero. . . desta forma podemos obter a função bijetiva,  $exp$ , de uma faixa de largura  $2\pi$ , sem a fronteira, no plano complexo de onde se retirou uma semireta.
- Você pode escolher qual semireta retirar, na figura (fig 106) foi retirada a semireta dos números reais negativos. Na próxima figura eu retirei a semireta que se opõe ao vetor  $\vec{a}$  que faz um ângulo  $\alpha$  com a semireta dos números reais positivos. Neste caso a  $exp$  estará definida na faixa que corresponden à variação  $[\alpha - \pi, \alpha + \pi]$ .
- Os círculos de centro na origem, no domínio do logaritmo, correspondem aos segmentos de reta verticais de comprimento  $2\pi$  na faixa-domínio da exponencial. Por exemplo, a imagem  $Log(\mathbf{S}^1)$  é o segmento de reta que corta a faixa no *argumento* zero.

A figura (fig 106), página 273, mostra como podemos obter, geometricamente, a imagem duma curva via logaritmo, no caso uma translação de  $\mathbf{S}^1$  por um vetor de módulo menor que 1.

Dominando a imagem de semiretas partindo da origem e de círculos concêntricos no ponto  $(0,0)$ , podemos obter a imagem de praticamente *qualquer curva* entre o plano complexo-logaritmo (sem uma semireta) e uma faixa-domínio da exponencial. A liberdade da *escolha duma semireta para ser eliminada* facilita este trabalho geométrico sem nenhuma particularização do problema: A função  $Log$  é definida assim

$$Log_{\alpha} : \mathbf{C} - \{0\} \rightarrow \mathbf{R} \times [\alpha - \pi, \alpha + \pi]$$

e poderíamos dizer que  $Log$  corresponderia à  $\alpha = 0$ , aqui não há uma notação padrão. Este é o caso da figura (fig 107), página 274, em que foi eliminada

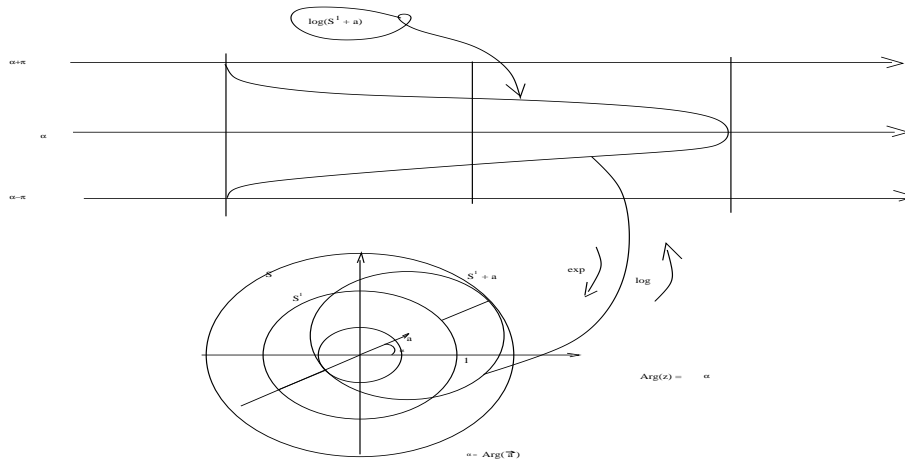


Figura 107: Imagem de  $\mathbf{S}^1 + \vec{a}$  pelo Log

a semireta que se opõe ao vetor  $\vec{a}$  usado para obter a translação  $\mathbf{S}^1 + \vec{a}$  cuja imagem aparece num retângulo da faixa-domínio da exponencial.

Há duas classes de gráficos (transformações) de  $\mathbf{S}^1 + \vec{a}$

#### Clasificación das transformações sob Logartimo

- Se o vetor  $\vec{a}$  tiver módulo menor do que 1 então  $\mathbf{S}^1 + \vec{a}$  estará contida em um círculo de centro na origem e de raio  $\|a\| + 1$ , tangenciando este círculo num ponto que se encontra sobre a reta  $r$  que passa por  $\vec{a}$  e também irá tangenciar o círculo de centro na origem de raio  $1 - \|a\|$ , também num ponto sobre a mesma reta  $r$ , analise a (fig 107). A particularidade que caracteriza esta classe é que  $0 \in U + \vec{a}$ , em que  $U$  é o disco unitário, tendo como resultado que  $Log_\alpha(\mathbf{S}^1 + \vec{a})$  é uma *curva aberta*.
- Se o vetor  $\vec{a}$  tiver módulo maior do que 1 então  $0 \notin U + \vec{a}$  e os dois círculos que “limitam”  $\mathbf{S}^1 + \vec{a}$  definem um retângulo que vai conter a imagem de  $Log(\mathbf{S}^1 + \vec{a})$ , agora uma curva fechada. Se o círculo de raio menor, como mostra a figura tiver raio menor do que um, então a imagem de  $Log(\mathbf{S}^1 + \vec{a})$  irá cortar a imagem de  $Log(\mathbf{S}^1)$  e aqui poderíamos ter duas subclasses de gráficos, quando  $Log(\mathbf{S}^1)$  estiver contida no retângulo que contém a imagem de  $Log(\mathbf{S}^1 + \vec{a})$  e o outro caso em que  $Log(\mathbf{S}^1)$  esteja à esquerda deste retângulo.

O ponto alto do raciocínio geométrico, que também envolve os conhecimentos do Cálculo sobre curvas tangentes e suas imagens por uma transformação que preserve ângulos, as chamadas *transformações conformes*, estudadas em *variável complexa*, mostra que  $Log(\mathbf{S}^1 + \vec{a})$  está contida no retângulo que é obtido como

imagem dos dois círculos tangentes, referidos acima, tendo a reta de altura  $\alpha$  como centro de simetria. A imagem estará contida entre as retas de alturas  $\alpha - \pi, \alpha + \pi$  nas duas classes descritas acima e em particular nas duas subclasses do segundo item.

Desenhando uma família de círculos “entre os círculos de raio menor e maior”, referidos acima, é possível obter uma quantidade significativa de pontos da imagem e assim desenhar esta imagem com precisão melhor do que esta que aparece nas figuras (fig 106), (fig 107).

Nas figuras (fig 108), página 275, você pode ver o gráfico da imagem via

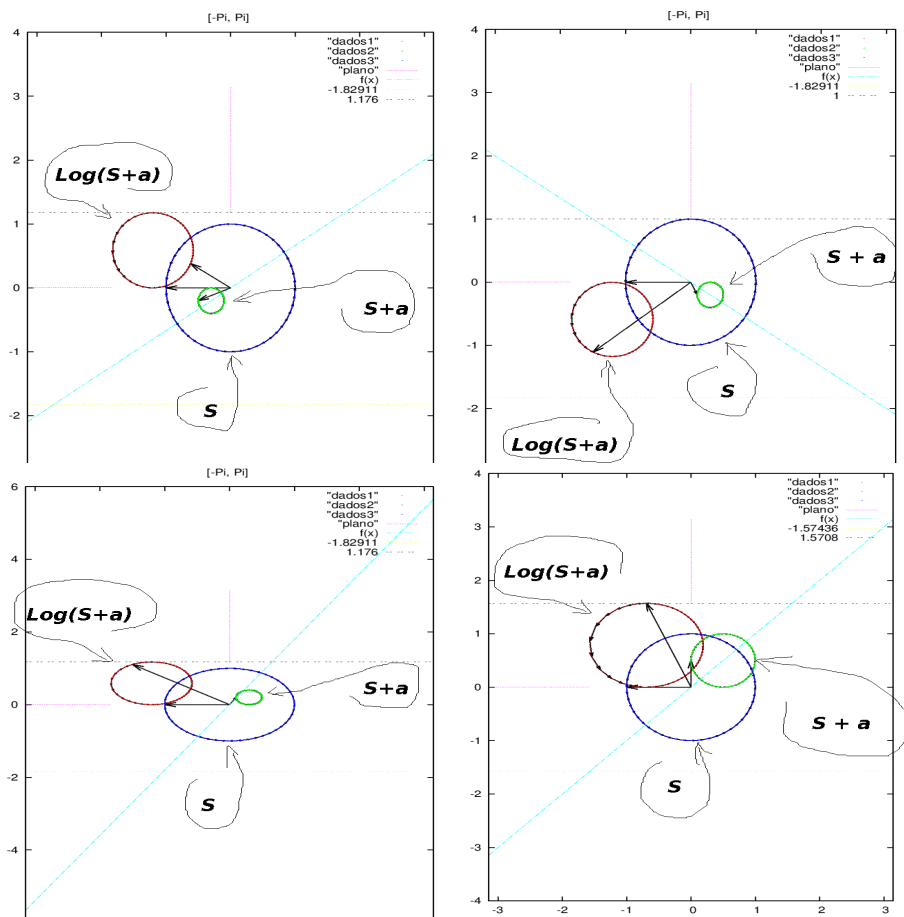


Figura 108: Imagem de  $S^1 + \vec{a}$  pelo Log

Log de  $S^1 + \vec{a}$  feita com o programa LogImage.cc que pode ser obtido em [6]. Este programa ainda não está completo, mas está perfeitamente operacional, ele gera o arquivo,  
transfere3

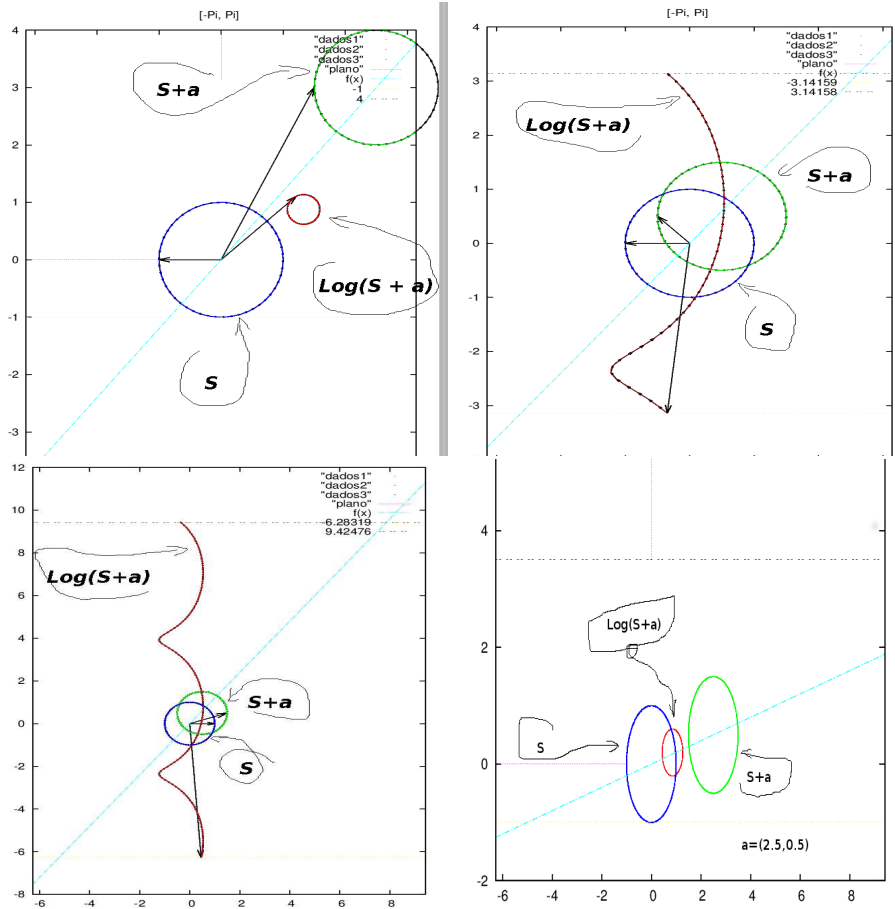


Figura 109:

com comandos para **gnuplot** que você pode editar para atender as suas necessidades e pode voltar a ver o gráfico executando num terminal

**gnuplot transfere3.**

O gráfico de  $y = f(x) = (b/a)x$  serve para mostrar a direção sobre a qual foi feita a translação  $S^1 + \vec{a}; \vec{a} = (a, b)$ . Na versão produzida pelo programa aparecem vetores mostrando a orientação das curvas, eles foram eliminados ao produzir a imagem nas figuras (fig 108)- (fig ??).

Derivada e primitiva

Os números complexos funcionam algébrica e topologicamente de forma semelhante aos números reais o que nos permite transferir para eles os conceitos de limite, derivada e integral do Cálculo mas surgem alguns fenômenos novos que de certa forma têm de ser esperados uma vez que  $\mathbf{C}$  se identifica com  $\mathbf{R}^2$  sob alguns aspectos, quer dizer, que  $\mathbf{C}$  é um espaço vetorial complexo de dimensão

1 e um espaço vetorial real de dimensão 2.

Confira o exemplo seguinte e as “coincidências” que ele apresenta

$$f(z) = z^2; f'(z) = 2z; \quad (814)$$

$$z = x + iy; f(z) = x^2 - y^2 + 2xyi; \quad (815)$$

$$f(z) = f(x + iy) = u(x, y) + iv(x, y) = x^2 - y^2 + 2xyi; \quad (816)$$

$$J(f) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix} = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} u_x & u_y \\ -u_y & u_x \end{pmatrix} = \begin{pmatrix} v_y & -v_x \\ v_x & v_y \end{pmatrix} \quad (817)$$

Como compatibilizar todas as expressões (eq. 814)-(eq. 817)?

A resposta vem da teoria das funções analíticas que identificou a  $J(f)$  com um único número complexo a partir das equações seguintes:

$$f'(z)dz = f'(z)(dx + idy) \equiv J(f) \begin{pmatrix} dx \\ dy \end{pmatrix}; \quad (818)$$

$$f'(z)dz = \begin{pmatrix} u_x & u_y \\ -u_y & u_x \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} = \begin{pmatrix} u_x dx + u_y dy \\ -u_y dx + u_x dy \end{pmatrix} \quad (819)$$

$$f'(z)dz \equiv u_x dx + u_y dy + i(-u_y dx + u_x dy) = (u_x - iu_y)(dx + idy) \quad (820)$$

$$f'(z)dz = \begin{pmatrix} v_y & -v_x \\ v_x & v_y \end{pmatrix} \begin{pmatrix} dx \\ dy \end{pmatrix} \quad (821)$$

$$f'(z)dz = (v_y dx - v_x dy) + i(v_x dx + v_y dy) \quad (822)$$

$$f'(z)dz = (v_y + iv_x)(dx + idy) \quad (823)$$

Na equação (eq. 819) identifiquei a  $J(f)$  usando apenas as derivadas de  $u$  e depois usando apenas as derivadas de  $v$  usando a primeira matriz.

Usei esta identificação para reescrever a diferencial nas equações (eq. 818)-(eq. 823). Os dois números  $u_x - iu_y$  ou  $v_y + iv_x$ , qualquer um deles, serve para identificar a jacobiana  $J(f)$  usando as identidades

$$u_x = v_y; u_y = -v_x; \quad (824)$$

chamadas *equações de Cauchy-Riemann* que valem para qualquer função complexa que seja diferenciável quando fizermos a identificação da função complexa com uma função de  $\mathbf{R}^2$  em  $\mathbf{R}^2$  usando as funções coordenadas  $u, v$ , portanto as seqüências de equações acima não foi nenhuma coincidência.

A conclusão é que a diferencial das funções complexas que forem diferenciáveis, e isto quer dizer, “satisfaçam às *equações de Cauchy-Riemann*”, se expressam da mesma forma como a diferencial no Cálculo a uma variável como o produto de números, a derivada univariada é um número. A *derivada complexa* é um número complexo e uma função complexa é derivável se sua derivada for um número complexo, o que é equivalente a dizer-se que a jacobiana da correspondente função de  $\mathbf{R}^2$  em  $\mathbf{R}^2$  satisfaz às *equações de Cauchy-Riemann*.

A função  $f(z) = e^z = e^{x+iy} = e^x(\cos(y) + i\sin(y))$  em que  $u(x, y) = e^x \cos(y); v(x, y) = e^x \sin(y)$  satisfaz às equações de Cauchy-Riemann e  $f'(z) = e^z$  a verificação é um bom exercício.

Da mesma forma  $\text{Log}(z) = \log\|z\| + i\text{Arg}(z) = u(x, y) + iv(x, y)$  também satisfaz às equações de Cauchy-Riemann em que  $\text{Arg}(z) = \text{atan}(y/x)$ ;  $z = x+iy$ ; se  $x \neq 0$  em cujo caso  $\text{Arg}(z) = \text{sign}(y)\frac{\pi}{2}$ , a verificação é um bom exercício.

A derivada de  $\text{Log}(z)$  é, como se deveria esperar,  $\frac{1}{z}$ , verifique as equações de Cauchy-Riemann.

### O efeito do zero

Nas figuras (fig 108)- (fig 109). você pode ver a evidência da classificação feita acima das transformações sob logaritmo complexo. Se o zero for ponto interior de uma curva fechada, a transformada desta curva, pelo logaritmo complexo será uma curva aberta, confira (fig 107) tangenciando a segmento de reta-imagem do círculo interior à curva nos seus pontos extremos. O padrão assim obtido repetir-se-á por diversas faixas paralelas se a curva der várias voltas em torno do zero. É isto que estabelece o *número de voltas* chamado de “*índice de um ponto curva relativamente a uma curva*”,  $\text{Ind}_\gamma(a)$ . Também tem o que ver com a “multiplicidade de uma raiz em uma equação”, ou de um polo.

É fácil provar isto interpretando geometricamente a integral

$$\oint_{\mathbf{S}^1} \frac{dz}{z} = \int_{-\pi}^{\pi} idt = 2\pi i \quad (825)$$

em vez de calcular a integral vou interpretá-la geometricamente e assim novamente obter o resultado da conta acima. Pelo teorema fundamental do Cálculo, o valor da integral é a diferença de uma primitiva calculada nos extremos do intervalo de integração:

$$\oint_{\mathbf{S}^1} \frac{dz}{z} = \text{Log}(z_1) - \text{Log}(z_0) = 2\pi i \quad (826)$$

porque  $\text{Log}$  transforma  $\mathbf{S}^1$  em uma curva aberta cujos extremos tangenciam a imagem de  $\mathbf{S}^1$  que é um segmento de reta perpendicular ao eixo  $OX$  limitado pela faixa-imagem do  $\text{Log}$  que tem largura  $2\pi$ . Este é o significado da integral na equação (eq. 825).

Este resultado não é particular da curva  $\mathbf{S}^1$ , vale para qualquer curva que lhe seja *homotópica a um ponto*, quer dizer, uma curva em que  $\mathbf{S}^1$  possa ser continuamente transformada com um mesmo ponto interior, no caso o zero.

A *homotopia* é uma relação de equivalência entre objetos, neste caso estou usando a homotopia entre curvas. Este conceito é usado no teorema de Green para definir *integrais independentes do caminho*.

---

- **matriz** é uma tabela retangular de números, reais ou complexos, que generaliza as enúplas de números porque com a adição satisfazem à estrutura de grupo comutativo. Elas foram introduzidas por Cayley para representar de forma



sintética um sistema de equações:

$$\begin{cases} ax_1 + bx_2 = p \\ cx_1 + dx_2 = q \end{cases} \quad (827)$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix} \quad (828)$$

mas um século depois da invenção de Cayley as matrizes adquiriram *independência* se tornando membros duma rica estrutura algébrica, a *álgebra das matrizes*.

O sistema de equações na equação (eq. 827) equivale ao produto das matrizes na equação (eq. 828) que foi a invenção de Cayley, transformando um sistema de equação numa única equação algébrica. Para isto ele teve que inventar um produto de matrizes fazendo as linhas atuarem sobre as colunas combinando os seus elementos exatamente como o produto escalar de vetores, neste caso *vetores-linha* multiplicados por *vetores-coluna*.

Observe que a equação (eq. 828) ainda pode ser sintetizada um pouco mais se atribuímos variáveis aos seus elementos:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}; x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}; b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} p \\ q \end{pmatrix}; \quad (829)$$

$$Ax = b; x = A^{-1}b; \quad (830)$$

se  $A$ , na (eq. 830) puder ser invertida, porque nem sempre uma matriz tem inverso. Não se caminha por esta trilha, entretanto, para resolver sistemas de equações porque o cálculo da inversa duma matriz é muito custoso e outros métodos são mais efetivos. Um dos mais comuns se chama *escalonamento de matrizes* que equivale à substituição de cada variável numa equação na equação seguinte com o objetivo de reduzir o número de variáveis até obter a equação mais simples. Confira este método e o algoritmo associado a ele.

Com frequência usamos programas de Álgebra Linear computacional, como *octave* ou *scilab* que são dois programas distribuídos com licenças igual ou semelhante à GPL, confira [15].

Eis um exemplo que pode ser rodado com *scilab* para resolver a equação matricial  $Ax = b$ ;

#### Exemplo 6 (sistema de equações) *Equação matricial*

```
a=3; b=2; c=-1;d=0;
A = [a,b;c,d];
b = [3;2]
x = linsolve(A,-b);
```

porque a função *linsolve()* do *scilab* foi projetada para resolver o sistema no formato  $Ax - b = 0$ . Se você copiar este exemplo num terminal do *scilab* o resultado virá com a notação

```
ans = [13;-2]
```

em que `ans` é uma variável muito usada pelos sistemas de *Álgebra Linear* para guardar a resposta portanto, executando agora

```
A*ans = b = [3;2];
```

Observe a sintaxe para definir uma matriz no `scilab` que é a mesma usada no `octave`<sup>50</sup>. Rode o exemplo num terminal do `scilab` para entender como funciona.

A equação matricial funciona, algébricamente, como uma equação numérica, se soubermos, e pudermos, calcular a inversa da matriz  $A$ , mas na prática o que se faz é obter a solução mecanicamente com um pacotes *Álgebra Linear* como `scilab` ou `octave`.

As matrizes se classificam pelo número de linhas ou colunas. Na próxima equação você pode ver três matrizes,  $A$  é uma matriz  $2 \times 3$ ,  $B$  é uma matriz  $3 \times 2$  e  $C$  é o produto  $AB$ , de  $A$  por  $B$ , é uma matriz  $2 \times 2$

$$A = \begin{pmatrix} -1 & 1 & 1 \\ 0 & -1 & 1 \end{pmatrix}; B = \begin{pmatrix} -1 & 1 \\ 1 & 1 \\ 0 & -1 \end{pmatrix}; C = \begin{pmatrix} 2 & -2 \\ -1 & -2 \end{pmatrix} = AB \quad (831)$$

Mas não é possível multiplicar  $BA$  uma vez que a regra de multiplicação exige que se faça o *produto escalar* da linha de ordem  $j$  da primeira matriz com a coluna de ordem  $k$  da segunda matriz, para obter o elemento de ordem  $j, k$  da matriz-produto.

Então é possível multiplicar uma matriz  $n \times p$  por uma matriz  $p \times m$  resultando numa matriz  $n \times m$  como se ve na equação (831).

Um tipo de matrizes em que o número de linhas e colunas é igual, é chamada de matriz quadrada, é o caso da matriz  $2 \times 2$  que aparece na equação (831).

No conjunto, das matrizes quadradas de ordem  $n$ , quer dizer, matrizes  $n \times n$ , tanto está definida a adição como a multiplicação. É comum denominarmos este conjunto com o símbolo  $\mathcal{M}_n$  então  $(\mathcal{M}_n, +, \cdot)$  satisfaz a estrutura de anel não comutativo por que em geral  $AB \neq BA$ , por exemplo

$$A = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}; B = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}; \quad (832)$$

$$AB = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} -1 & 1 \\ -1 & 2 \end{pmatrix} = BA; \quad (833)$$

As matrizes quadradas são um exemplo de uma estrutura algébrica semelhante a dos inteiros com adição e multiplicação mas a multiplicação não é comutativa e onde não há uma divisão definida. No anel das matrizes  $\mathcal{M}_n$  há divisores de zero, elementos que são diferentes de zero mas que o produto é nulo:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad (834)$$

<sup>50</sup>Há pequenas diferenças de sintaxe entre `scilab` e `octave`.

Este “fenômeno” também acontece no conjuntos dos restos na divisão por  $n$  se  $n$  não for um número primo, onde se têm também a estrutura de anel, comutativo, mas com divisores de zero, são os chamados números congruentes módulo  $n$ .

As matrizes são estudadas extensivamente numa disciplina chamada *álgebra linear* onde se pode mostrar que uma matriz  $n \times p$  representa uma função definida no espaço vetorial  $\mathbf{K}^p$ ,  $\mathbf{K}$  podendo ser  $\mathbf{R}$ ,  $\mathbf{C}$ .

1. Os vetores do espaço  $\mathbf{K}^p$  podem ser entendidos como matrizes  $p \times 1$ . Algumas vezes tais vetores são chamados de *vetor linha* ou *vetor coluna* dependendo da forma como seus elementos estiverem dispostos. Isto não é apenas uma *questão de gosto*, depende da necessidade de fazer o produto de matrizes à direita ou à esquerda.
2. estes vetores-coluna,  $p \times 1$ , podem ser multiplicados à esquerda por matrizes  $n \times p$  resultando em matrizes  $n \times 1$  que são os elementos de  $\mathbf{K}^n$
3. desta forma uma matriz  $\mathcal{A}; n \times p$  pode ser identificada como uma função  $\mathcal{A} : \mathbf{K}^p \rightarrow \mathbf{K}^n$ .

A *álgebra linear* é a disciplina que estuda as propriedades destas matrizes e a solução das equações algébricas que podem ser estabelecidas com elas que são chamadas de *sistemas de equações lineares*, ou simplesmente *equações lineares*.

Se uma matriz for quadrada de ordem  $n$  ela tem um *determinante* que é um número calculado com combinações multiplicativas dos elementos da matriz. Se  $\det(\mathcal{A}) \neq 0$  então  $\mathcal{A}$  tem uma matriz inversa  $\mathcal{B}$ , isto é,  $\mathcal{A}\mathcal{B} = \mathcal{B}\mathcal{A} = I$  é a matriz identidade de ordem  $n$ .

O conjunto das matrizes de ordem  $n$  cujo determinante seja diferente de zero é um grupo multiplicativo não comutativo. Dentre tais grupos tem um importante grupo das matrizes unitárias, aquelas cuja determinante é 1.

A inversa da matriz  $\mathcal{A}$ , se existir, é dada pela fórmula

$$\mathcal{A}^{-1} = \frac{1}{\det \mathcal{A}} \text{co}(\mathcal{A})^t \quad (835)$$

$$\det \mathcal{A} = \text{determinante de } \mathcal{A}; \quad (836)$$

$$\text{co}(\mathcal{A}) = \text{matriz dos cofatores de } \mathcal{A}; \quad (837)$$

$$\mathcal{B}^t = \text{matriz transposta de } \mathcal{B}; \quad (838)$$

em que vemos as operações  $\text{co}(\mathcal{A})$  é a matriz dos cofatores e a “potência”  $t$  representa a transposta da matriz.

A transposta de uma matriz real  $\mathcal{A}^t = (a_{ij})^t = (a_{ji})$

Os determinantes encontram uma explicação teórica relativamente simples na disciplina álgebra multilinear. A definição do determinantes diretamente para uma matriz é relativamente complicada e obscura se dividindo em regras artificiais relativas às ordens. As mais conhecidas e simples sendo para as matrizes de ordem 2, 3, 4. O cálculo do determinante de uma matriz de ordem 3

pode nos dar uma visão da teoria das “funções multilineares”

$$\mathcal{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}; \quad (839)$$

$$\det(\mathcal{A}) = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{vmatrix} = \quad (840)$$

$$a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \quad (841)$$

$$-(a_{31}a_{22}a_{13} + a_{32}a_{23}a_{11} + a_{33}a_{21}a_{12}) = \quad (842)$$

$$a_I + a_{(123)} + a_{(132)} - (a_{(13)} + a_{(23)} + a_{(12)}) \quad (843)$$

Na equação (840) apliquei a *regra de Sarrus* que consiste na duplicação das duas primeira colunas e nas equações (841) e (842) continuei com a *regra de Sarrus* multiplicando os elementos das diagonais com sinal negativo para as diagonais crescentes.

Agora observe que os índices destas matrizes são formados de pares  $(ij)$  tirados do conjunto  $\{1, 2, 3\}$  e podem assim ser identificados como funções deste conjunto nele mesmo, as permutações deste conjunto que são em número de 6 que é a quantidade termos que temos na “soma”. Vamos ver agora a razão pela qual alguns trazem o sinal negativo. Isto se apresenta na última equação (842) quando alterei a notação para representar as *permutações* de  $\{1, 2, 3\}$

$$I, (123), (132), (12), (13), (23) \quad (844)$$

$$a_I = a_{11}a_{22}a_{33}; \text{ todos elementos fixo, identidade} \quad (845)$$

$$a_{(123)} = a_{12}a_{23}a_{31}; \quad (846)$$

$$a_{(132)} = a_{13}a_{32}a_{21}; \quad (847)$$

$$a_{(13)} = a_{13}a_{31}a_{22}; \quad (848)$$

$$a_{(23)} = a_{23}a_{32}a_{11}; \quad (849)$$

$$a_{(12)} = a_{12}a_{21}a_{33}; \quad (850)$$

ou a notação alternativa,

$$a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)};$$

Quer dizer que podemos expressar a soma na equação (847) como

$$\det(\mathcal{A}) = \sum_{\sigma \in \text{sim}(3)} \text{ sinal}(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} a_{3\sigma(3)} \quad (851)$$

ou seja, consideramos todas as possíveis permutações dos índices da matriz para representar os produtos dos seus elementos e somamos estes produtos atribuindo a cada parcela o sinal da permutação  $\sigma$  que a gerou.

No caso de um determinante de uma matriz  $n \times n$  teremos

$$\det(\mathcal{A}) = \sum_{\sigma \in \text{sim}(n)} \text{ sinal}(\sigma) \prod_{i=1 \dots n} a_{i\sigma(i)} \quad (852)$$

Mas esta fórmula é raramente usada quando  $n$  for grande, observe que  $n = 4$  é pequeno e  $\text{card}(\text{sim}(4)) = 4! = 24$ . Em geral se opta por triangularizar as matrizes transformando o determinante no simples produto dos elementos da diagonal, é o método usado pelos programas de álgebra linear computacional, *octave* ou *scilab* ou em *cálculo numérico*

A pergunta agora é, qual é o *signal* de uma permutação. Quando construímos permutações o fazemos em cima de conjuntos com  $n$  elementos, conjuntos finitos e sempre então podemos dizer que se trata do conjunto

$$\{1, 2, 3, \dots, n\}$$

que é um conjunto ordenado. Então contamos o número de transposições que uma permutação produz, por exemplo, considerando

$$\{1, 2, 3, \dots, 9\}$$

e o produto de ciclos, confira *permutação*

$$(1357)(246)(89) \mapsto 1 = (-1)^{12}; 12 = o((1357)(246)(89)) \quad (853)$$

$$(13)(35)(57)(24)(46)(89) \quad (854)$$

$$\left\{ \begin{array}{l} 1 \mapsto 3 \\ 3 \mapsto 5 \\ 5 \mapsto 7 \\ 7 \mapsto 1 \end{array} \right. \quad (855)$$

$$\left\{ \begin{array}{l} 2 \mapsto 4 \\ 4 \mapsto 6 \\ 6 \mapsto 2 \end{array} \right. \quad (856)$$

$$\left\{ \begin{array}{l} 8 \mapsto 9 \\ 9 \mapsto 8 \end{array} \right. \quad (857)$$

tem 6 transposições. Olhando o caso do determinante de ordem 3, temos

$$(123) = (13)(12) \mapsto (-1)^2 = 1; \quad (858)$$

$$(132) = (12)(13) \mapsto (-1)^2 = 1; \quad (859)$$

$$I \mapsto (-1)^0 = 1 \text{ tem zero transposições;} \quad (860)$$

$$\left\{ \begin{array}{l} (12) \mapsto -1; \\ (13) \mapsto -1; \\ (23) \mapsto -1; \end{array} \right. \quad (861)$$

- **média** Este é um conceito que tem interpretações diferentes, a forma mais comum sendo a *média aritmética simples* cuja definição é, “dados dois números  $a, b$ , a média deles é  $\frac{a+b}{2}$ ”. Com frequência lidamos com a média aritmética ponderada e neste caso associamos um *peso* a cada um dos números e a fórmula passa a ser

$$\text{Med}_{A,B}(a, b) = \frac{Aa + Bb}{A + B} \quad (862)$$

e a média aritmética simples é um caso particular desta quando os pesos forem iguais:

$$\frac{Pa + Pb}{P + P} = \frac{a + b}{2} \quad (863)$$

A média aritmética está associada à interpolação linear, quer dizer, se considerarmos dois pontos no espaço, como mostra a figura (110), página 284, dois

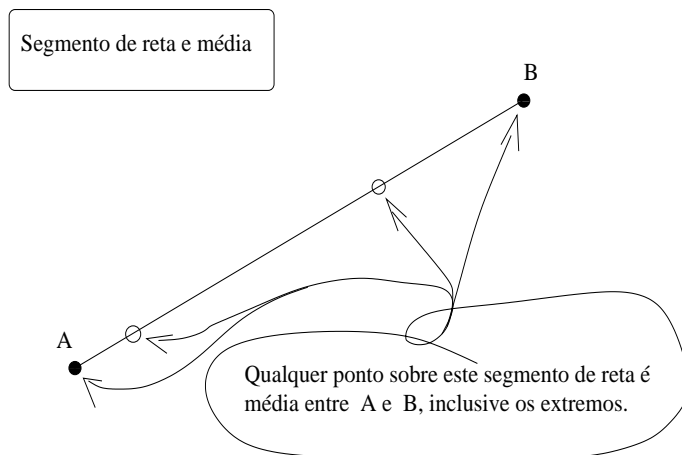
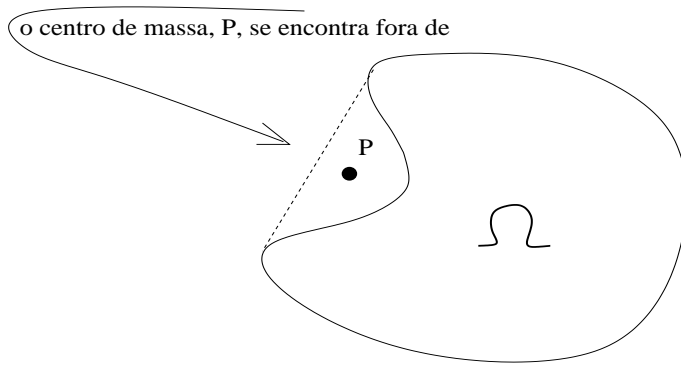


Figura 110: Segmento de reta e média

pontos  $A, B$  determinando um segmento de reta, qualquer ponto sobre este segmento de reta é uma média entre  $A$  e  $B$ , inclusive os dois extremos, obviamente, neste caso estamos pensando em *média aritmética ponderada*.

A média também está relacionada com o conceito de convexidade, e fica mais interessante analisar a convexidade no espaço de dimensão maior ou igual a dois, embora um segmento de reta, ou um intervalo de números reais já sejam exemplos de conjuntos convexos. A figura (111), página 285, mostra um conjunto que não é convexo e seu fecho convexo. O conjunto  $\Omega$  não é convexo, mas se acrescentarmos a parte limitada pelo segmento de reta, se obtem um conjunto convexo. O conceito de baricentro, ou centro de massa, é a forma que toma a média aritmética ponderada quando se pensa em figuras geométricas. Aqui é preciso pensar na integral para fazer o cálculo. A integral é uma forma de generalizar o conceito de medidas, *comprimento, área, volume, hipervolume*. . . mas a fórmula para calcular o *baricentro* é apenas uma generalização da fórmula apresentada mais acima para calcular a média aritmética ponderada: Calculamos uma integral da figura parametrizada sobre um domínio onde ela esteja definida e depois dividimos pela integral do domínio: o resultado é a média, apenas agora é habitual se referir a este média como *média integral*. Se a função que parametrizar o conjunto  $\Omega$  tiver variações ao longo do domínio, isto pode ser interpretado, ou pode interpretar, alterações da massa de  $\Omega$  ao longo do domínio. Usamos isto quando o objeto geométrico é feito de material não uniforme. Você



Um conjunto que não é convexo e seu fecho convexo

Figura 111: Conjunto convexo e média

pode ver um exemplo do baricentro de um objeto que seja feito de material não uniforme quando o “centro de massa” se encontra fora de  $\Omega$ , porém dentro do seu *fecho convexo*, na figura (111). O ponto  $P$  que é o centro de massa de  $\Omega$ , se encontra fora de  $\Omega$ , ou ainda, a *média aritmética ponderada* dos pontos de  $\Omega$  é  $P$  e dizemos *média integral*,

$$P = \frac{1}{\text{Area}(\mathcal{D})} \int_{\mathcal{D}} \int F(x, y) dx dy \quad (864)$$

em que  $F$  é a função que parametriza  $\Omega$  sobre o domínio  $\mathcal{D}$

Aqui podemos tomar o caminho inverso da generalização para lembrar o *valor médio integral* do Cálculo univariado:

$$\frac{1}{b-a} \int_a^b f(x) dx \quad (865)$$

e a comparação entre as fórmulas que aparecem nas equações (865) e (864) é instrutiva porque elas representam exatamente a mesma coisa apenas em dimensões diferentes, uma fórmula mais difícil que pode ser colocada também neste contexto mostrando que ela não precisa ser tão amedrontadora, a *integral de Riemann-Stieltjes* que precedeu a *integral de Lebesgue*

$$\int_a^b f(x) g'(x) dx \quad (866)$$

em que a função  $g'(x)$  representa a forma como a massa se distribui ao longo do intervalo  $[a, b]$  o que substitui a divisão por  $m([a, b]) = b - a$ . Um exemplo

simples deste caso é quando  $g'(x)$  representa uma distribuição de probabilidade, as probabilidades servem para calcular médias integrais e representam a forma como a “massa” do fenômeno considerado se distribui ao longo do espaço, em qualquer dimensão.

Na integral de Lebesgue  $g'(x)dx = d\mu(x)$  é uma medida que não precisa ser exatamente a derivada de uma função ou que a *derivada seja uma função*. . .

As *médias* nos acompanham até nos momentos mais íntimos de nossa vida, uma caixa d'água *representa uma média* uma vez que ao longo do tempo ela garante que o abastecimento médio d'água de uma residência se mantenha constante. Ou uma *fila única* num local de atendimento ao público também representa uma média tornando equitativo o atendimento para todas as pessoas que se encontram a espera.

---

- **medida** é uma generalização do conceito de área. Área é um tipo de medida. Há várias formas de fazer esta generalização, uma delas, que se deve ao matemático Francês Lebesgue, consiste em identificar uma família de subconjuntos de um determinado conjunto para construir com eles uma  $\sigma$ -álgebra de conjuntos e sobre esta  $\sigma$ -álgebra se pode definir uma função  $\mu$  que tem as mesmas propriedades da área. Esta função se chama “medida” e os subconjuntos que formam a  $\sigma$ -álgebra são os conjuntos mesuráveis. Desta forma se pode definir funções integráveis relativamente a esta  $\sigma$ -álgebra, ditas “integráveis no sentido de Lebesgue” em oposição à forma habitual do Cálculo de definir a integral que é dita “integração no sentido de Riemann”.

---

- **medida de Dirac** é uma *medida* que foi definida por P. Dirac em suas pesquisas em Física Quântica para isolar uma partícula, então ela é uma medida concentrada em um único ponto do espaço e pode ser definida pelo sistema de equações

$$\delta_a(S) = 1 \iff a \in S; \quad (867)$$

$$\delta_a(S) = 0 \iff a \notin S; \quad (868)$$

ou seja, a medida do conjunto  $S$  será 1 se  $a \in S$ . Ou em termos de probabilidade, “a probabilidade de  $S$  será 1 se  $a \in S$  que como seria do interesse de Dirac, no caso de partículas, em que, nesta linguagem  $a$  seria uma partícula e  $S$  uma região do espaço em que esta partícula poderia estar presente. É muito comum se fazer referência à *medida de Dirac* como “*função de Dirac*” pois ela teria sido definida, inicialmente, como uma *função de probabilidade* nula no espaço inteiro exceto num ponto  $a$  em que assumiria o valor  $\infty$  e cuja integral seria 1.

Uma função com estas propriedades não pode existir e foi preciso algum tempo para que os matemáticos, em particular Laurent Schwartz e Sebastião Silva, conseguissem generalizar o Cálculo Diferencial e Integral a um novo tipo de objetos, chamados funções generalizadas ou distribuições em que a medida de Dirac é um exemplo de distribuição ou função generalizada.

Na generalização do Cálculo Diferencial e Integral às distribuições, se ampliou o conceito de derivada, *derivada no sentido das distribuições*. Se uma



função for *classicamente* diferenciável, os dois conceitos coincidem, mas agora é possível derivar funções mesmo não contínuas. Por exemplo a função característica de um intervalo

$$\chi_{[a,b]}; \quad (869)$$

pode ser vista como uma alteração da função constantemente nula em lhe tenham sido incluídos dois pontos de salto,  $a, b$ . Em  $\{a\}$  um salto de amplitude 1, e em  $\{b\}$  um salto de amplitude -1. Se calcularmos a derivada no sentido das distribuições de  $\chi_{[a,b]}$  o resultado será

$$0 + \delta_a - \delta_b \quad (870)$$

em que aparece uma combinação linear de translações da *medida de Dirac* tendo por coeficientes as amplitudes dos saltos.

A medida de Dirac é frequentemente referida como um *pulso de energia 1* porque ela pode ser aproximada por funções diferenciáveis, com suporte de pequena medida em volta de um ponto e com integral 1.

É interessante que podemos obter aproximações computacionais mostrando que este *conceito generalizado* não se encontra *distanciado da realidade como pode parecer*. Para conseguir esta visão computacional, figura (112), página 287, siga os seguintes passos:

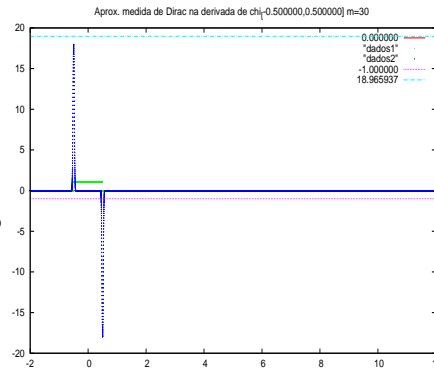


Figura 112: Medida de Dirac na derivada de função descontínua

$$(f * g)' = f * g' = f' * g \quad (871)$$

$$f * \delta_0 = f \quad (872)$$

$$(f * \delta_0)' = f' \quad (873)$$

1. A equação (871) mostra que a derivada de um produto de convoluções se distribui à direita ou à esquerda para qualquer dos fatores;
2. A equação (872) lembra que a medida de Dirac é a unidade na multiplicação por convolução;
3. Na equação (873) temos uma chave para obter uma aproximação das derivadas, substituindo a medida de Dirac por um núcleo, ou um sinal diferenciável com *suporte de pequena medida em volta de zero*, procure *regularização por convolução, sobre isto*. Se substituirmos:

$$(f * \delta_0)' = f'$$

por

$$(f * \eta)' = f * \eta' = f' * \eta \approx f'$$

e o resultado é o que você pode ver na figura (112), página 287. Se nesta figura *apagar a função característica*, verá dois *pulsos*, um com energia positiva 1 e outro com energia negativa -1 que são aproximações da medida de Dirac exatamente no sentido que Paul Dirac imaginava.

- **métrica** é a generalização do conceito de distância. Dado um conjunto  $M$  e uma função positiva

$$d : M \times M \rightarrow \mathbf{R}; \quad (874)$$

$$d(x, y) = d(y, x); \text{ simetria} \quad (875)$$

$$d(x, x) = 0; \quad d(x, y) \leq d(x, z) + d(z, y); \text{ desigualdade triangular} \quad (876)$$

$$(877)$$

dizemos que  $(M, d)$  é um *espaço métrico*.

Um exemplo trivial é a medida  $\delta$  definida pela “pela delta de Dirac”

$$d(x, y) = \delta_{x,y} = \begin{cases} x = y & 0 \\ x \neq y & 1 \end{cases} \quad (878)$$

Esta medida é comumente usada em conjuntos onde não se encontrem definidas operações algébricas e o espaço métrico resultante é chamado de espaço métrico discreto

Outro exemplo é a medida usual da Geometria Analítica, também chamada medida euclidiana

$$x, y \in \mathbf{R}^3; d(x, y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2} \quad (879)$$

- **mínimos, quadrados**, um método de análise numérica para encontrar a melhor curva que pode modelar uma massa de dados coletada.

Possivelmente o método surgiu na astronomia na tentativa de estabelecer as rotas dos astros a partir de dados coletados por observações astronômicas. Um caso célebre é o de Kepler que sucedeu a Tycho Brahe, na direção do observatório perto de Praga. Enquanto que Tycho Brahe mediu cuidadosamente as órbitas de planetas, foi finalmente Kepler que demonstrou que estas elipses.

Há basicamente dois métodos dos *quadrados mínimos*, o chamado linear, em que se procura uma reta, uma função linear afim, do primeiro grau, que melhor aproxime uma massa de dados, e o não linear, em que se procura uma expressão polinomial de grau maior do que 1. Estes dois métodos são semelhantes, mas o caso *linear* sendo um pouco mais elementar, é mais comumente usado, sobretudo quando a massa de dados sugere que haja uma solução “linear”.

Outras denominações deste método aparecem como *busca da melhor curva*, em inglês, *fitting curves*.

- **mínimos, quadrados**, o caso “linear”. É também chamado de *quadrados mínimos comum*.

O algoritmo se baseia na existência duma pesquisa em que os dados se encontram coletados numa tabela que vou aqui descrever como  $L = (x_k, y_k)_{k \in \{1, \dots, n\}}$  e se deseja descobrir um função do primeiro grau,  $P(x) = ax + b$ , tal que

$$y = P(x) = ax + b; \quad (880)$$

$$\vec{u} = (x_k, y_k)_{k \in \{1, \dots, n\}}; \vec{u} \in \mathbf{R}^{2n} \quad (881)$$

$$F(\vec{u}, a, b) = \sum_{k=1}^n (P(x_k) - y_k)^2; \sqrt{F(\vec{u}, a, b)} = \sqrt{\sum_{k=1}^n (P(x_k) - y_k)^2}; \quad (882)$$

$$F(\vec{u}, a, b) = \sum_{k=1}^n (ax_k + b - y_k)^2 = nb^2 + \sum_{k=1}^n (ax_k - y_k)^2 + 2b(ax_k - y_k) \quad (883)$$

em que  $\sqrt{F}$  mede a *dispersão padrão* do levantamento relativamente ao modelo  $P$ .

Se os pontos do levantamento satisfizerem à equação duma reta, então

$$y_k = \alpha x_k + \beta$$

e o valor mínimo de  $F$  será alcançado quando  $\alpha = a; \beta = b$ , mas esta situação ideal não deve ser esperada e o que se procura é uma reta que melhor represente o *levantamento de dados* no sentido de que a *dispersão padrão* seja mínima. *Dispersão padrão* é o nome dado em *Probabilidade e Estatística* à norma euclidiana,  $\|\cdot\|_2$ , que aparece na equação (eq. 882).

A equação (eq. 883) define um operador em que a variável, ou elemento do espaço onde o operador está definido, é o polinômio  $P$  representando pelos seus coeficientes  $a, b$  e o objetivo é encontrar um valor mínimo para este operador no espaço de todos os polinômios do primeiro grau  $P$ , é aqui que entra o ponto de vista do *Cálculo Variacional* sob cuja ótica estou resolvendo este problema.

De imediato vou aplicar as condições para obter o mínimo da função  $F$  de duas várias variáveis  $a, b$ , definida na equação (eq. 882), ou (eq. 883), o gradiente deve ser zero relativamente às variáveis  $a, b$

$$\frac{\partial F}{\partial a} = 2 \sum_{k=1}^n (x_k(ax_k - y_k) + bx_k) = 2 \sum_{k=1}^n x_k(ax_k - y_k + b) = 0; \quad (884)$$

$$\frac{\partial F}{\partial b} = 2nb + 2 \sum_{k=1}^n (ax_k - y_k) = 0; \quad (885)$$

$$nb^2 + \sum_{k=1}^n (ax_k - y_k)^2 + 2b(ax_k - y_k);$$

$$\frac{\partial F}{\partial a} = 2 \sum_{k=1}^n (x_k(ax_k - y_k) + bx_k) = 2 \sum_{k=1}^n x_k(ax_k - y_k + b) = 0; \quad (886)$$

$$\frac{\partial^2 F}{\partial a^2} = 2 \sum_{k=1}^n x_k^2; \quad (887)$$

$$\frac{\partial^2 F}{\partial b \partial a} = 2 \sum_{k=1}^n x_k; \quad (888)$$

$$\frac{\partial F}{\partial b} = 2nb + 2 \sum_{k=1}^n (ax_k - y_k); \quad (889)$$

$$\frac{\partial^2 F}{\partial a \partial b} = 2 \sum_{k=1}^n x_k; \quad (890)$$

$$\frac{\partial^2 F}{\partial b^2} = 2n; \quad (891)$$

e o determinante do Hessiano deve ser positivo onde o gradiente for zero, caracterizando que se tenha uma forma quadrática positiva definida:

$$H = \det \begin{pmatrix} \frac{\partial^2 F}{\partial a^2} & \frac{\partial^2 F}{\partial b \partial a} \\ \frac{\partial^2 F}{\partial a \partial b} & \frac{\partial^2 F}{\partial b^2} \end{pmatrix} = \frac{\partial^2 F}{\partial a^2} \frac{\partial^2 F}{\partial b^2} - \frac{\partial^2 F}{\partial b \partial a} \frac{\partial^2 F}{\partial a \partial b}; \quad (892)$$

$$H = 4n \sum_{k=1}^n x_k^2 - \left( 2 \sum_{k=1}^n x_k \right)^2 = 4 \left( n \sum_{k=1}^n x_k^2 - \sum_{k=1}^n x_k^2 - 2 \sum_{k=1}^n \sum_{j=1}^n x_k x_j \right) \quad (893)$$

$$H = 4 \left( (n-2) \sum_{k=1}^n x_k^2 + \sum_{k=1}^n x_k^2 - 2 \sum_{k=1}^n \sum_{j=1}^n x_k x_j \right); \quad (894)$$

$$H = 4 \left( (n-2) \sum_{k=1}^n x_k^2 + \left( \sum_{k=1}^n x_k \right)^2 \right) \geq 0; \quad (895)$$

Se houver solução em  $a, b$  ela será o mínimo requerido, então retornando às equações do gradiente igualado a zero tenho

$$\frac{\partial F}{\partial b} = 2nb + 2 \sum_{k=1}^n (ax_k - y_k) = 0 \Rightarrow b = -\frac{1}{n} \sum_{k=1}^n (ax_k - y_k); \quad (896)$$

$$b = -a\bar{x} + \bar{y}; \quad (897)$$

$$\frac{\partial F}{\partial a} = 2 \sum_{k=1}^n (x_k(ax_k - y_k) + bx_k) = 2 \sum_{k=1}^n x_k (ax_k - y_k + b) = 0; \quad (898)$$

$$\sum_{k=1}^n x_k (ax_k - y_k + b) = a \sum_{k=1}^n x_k^2 - \sum_{k=1}^n x_k y_k + b \sum_{k=1}^n x_k; \quad (899)$$

$$a\|x\|^2 - \langle x, y \rangle + b \sum_{k=1}^n x_k = 0; \quad (900)$$

$$b = \frac{\langle x, y \rangle - a\|x\|^2}{\sum_{k=1}^n x_k}; \quad (901)$$

$$\frac{\langle x, y \rangle - a\|x\|^2}{\sum_{k=1}^n x_k} = -a\bar{x} + \bar{y}; \quad (902)$$

$$\langle x, y \rangle - a\|x\|^2 = -a\bar{x} \sum_{k=1}^n x_k + \bar{y} \sum_{k=1}^n x_k; \quad (903)$$

$$a = \frac{\langle x, y \rangle - \bar{y} \sum_{k=1}^n x_k}{\|x\|^2 - \bar{x} \sum_{k=1}^n x_k}; \quad (904)$$

$$P(x) = ax + b; \quad (905)$$

e na equação (eq. 905) você tem a expressão do polinômio do primeiro grau melhor se insere nos dados do levantamento  $L$ .

Aqui está um programa escrito em `calc` para obter a função linear que pelo método dos mínimos quadrados melhor se insere num levantamento de dados. Em seguida coloco instruções de como fazer uso do programa.

```
mat y[] = {4, 1.5, 6, 8, 10, 7, 14.5, 15, 18, 20}
mat x[] = {-3, -1, 1, 3, 5, 7, 9, 11, 13, 15}
n = size(x);
define escalar(x,y) {
local n = size(x), k = 0, soma = 0;
while(k < n) {
soma = soma + x[k]*y[k];
k++;
}
return(soma);
}
define norma(x) {
local n = size(x), k = 0, soma = 0;
while(k < n) {
soma = soma + x[k]*x[k];
```

```

k++;
}
return(sqrt(soma));
}
define media(x) {
local n = size(x), k = 0, soma = 0;
while(k < n) {
    soma = soma + x[k];
    k++;
}
return(soma/n);
}
define Soma(x) {
    local n = size(x), k = 0, soma = 0;
    while(k < n) {
        soma = soma + x[k];
        k++;
    }
return(soma);
}

a = (escalar(x,y) - media(y)*Soma(x) )/(power(norma(x),2) - media(x)*Soma(x));
b = (escalar(x,y) - a*power(norma(x),2))/(Soma(x));

define EscreveTransfere() {
local transfere = fopen("transfere", "w");
fprintf(transfere, "P(x) = %2.3f*x + %2.3f \n", re(a),re(b));
fprintf(transfere, "plot \"MinimosQuadradosLinear.data\" with points,\
    P(x), 0 \n");
fprintf(transfere, "pause -2 \"aperte enter para terminar\" \n");
fclose(transfere);
}

define ChamaGnuplot() {
system("gnuplot transfere")
}

define Main() {
EscreveTransfere();
system("joe transfere");
ChamaGnuplot();
}
Main();

```

Este programa precisa que no arquivo “MinimosQuadradosLinear.data” estejam os dados do levantamento no formato que gnuplot possa entender, como

os dados distribuídos em duas colunas representando em cada linha, respectivamente, a abscissa e a ordenada dum ponto do plano conforme se pode ver na figura (fig 113), página 293,

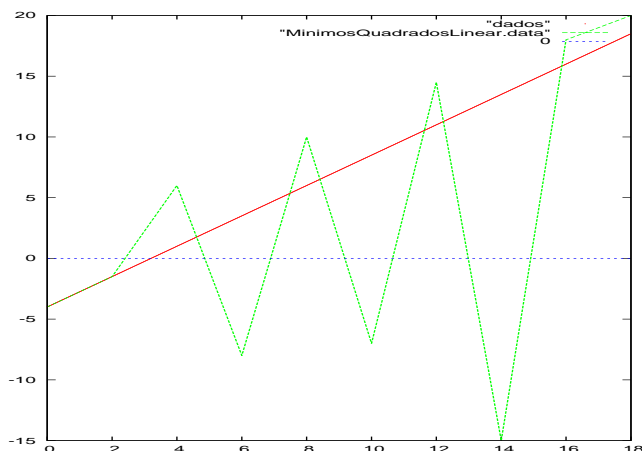


Figura 113: tabela de dados do levantamento

Abra um terminal do `calc` e nele digite:

```
read "MinimosQuadradosLinear.calc"
main();
```

tendo o cuidado de colocar no arquivo `MinimosQuadradosLinear.data` os valores que aparecem na matriz de pontos  $(x, y)$  no formato para que `gnuplot` possa entender, confira a figura (fig 113), página 293. Você irá ver a reta que melhor se enquadra nestes dados assim como sua equação estará guardada no arquivo `transfere` que será criado pelo programa.

Você pode ver na figura (fig 114), página 294, o gráfico que correspondem aos dados tabulados anteriormente.

- **mínimos quadrados, caso polinomial** é um método para encontrar uma curva que melhor modele uma massa de dados. Confira também o *caso linear*.

Esta é uma das formas do problema de regularização e também chamado de *regularização por quadrados mínimos*. A ideia é a de encontrar uma curva polinomial que se encontre centralizada relativamente a uma massa de dados relativamente a norma do  $\mathcal{L}^2([a, b])$  em que  $[a, b]$  é o intervalo da reta dentro dos qual foi feito um experimento. Vou desenvolver uma *linguagem* apropriada para construir o algoritmo.

Feita uma pesquisa são obtidos dados coletados numa tabela que vou aqui descrever como  $L = (x_k, y_k)_{k \in \{1, \dots, n\}}$  e se deseja descobrir um função polinomial  $y = P(x)$  que melhor descreva estes dados, ou cujo gráfico melhor aproxime os resultados representados pelo levantamento  $L$ . Pensando na massa de dados

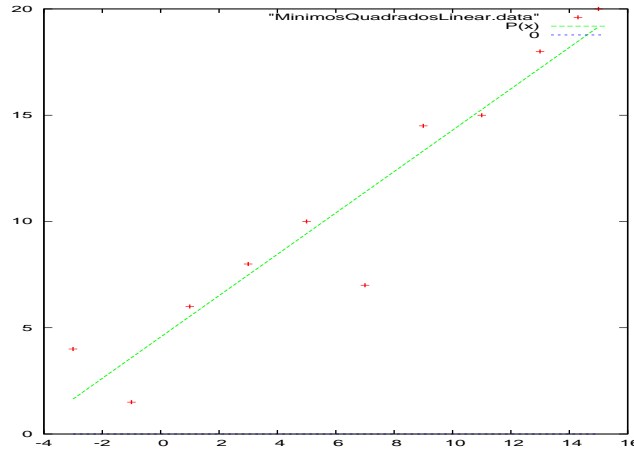


Figura 114: melhor reta para uma massa de dados

como uma “mancha” no plano (ou no espaço), como a que se pode obter fotografando uma cultura de bactérias, se deseja uma curva que possa melhor representar o centro desta mancha, ou também a fronteira da mesma que foi objetivo de Kepler ao tentar interpretar os dados coletados por Tycho Brahe concluindo que a rota dos planetas era elíptica, uma curva do segundo grau.

O conjunto destes pontos, interligados por segmentos de reta, define uma função poligonal  $y = f(x); y_k = f(x_k)$  que vou usar para desenvolver o algoritmo.

$$f : [a, b] \longrightarrow \mathbf{R}; x_k \mapsto y_k = f(x_k); \quad (906)$$

$$(x_k, y_k)_{k \in \{1, \dots, n\}} \subset [a, b]; a = x_1; b = x_n; \quad (907)$$

$$y = P(x); \text{ um polinômio de grau } m \leq n; \quad (908)$$

$$P \text{ minimiza } \sqrt{\int_a^b |f(x) - P(x)|^2 dx} = \|f - P\|_2 \quad (909)$$

Embora se pense num polinômio de grau menor ou igual a  $n$ , na verdade se tem por objetivo encontrar um polinômio de grau mínimo, dois ou três, que melhor se adapte aos dados  $L$ , porque um polinômio de grau  $n$  na verdade seria o polinômio de Lagrange que é o melhor polinômio que interpola estes dados, entretanto quanto maior for  $n$  mais acentuado será o erro nas extremidades do intervalo, o fenômeno de Runge, o que torna o polinômio de Lagrange menos interessante. Além do mais, polinômios de grau alto oferecem erros computacionais maiores. Uma das soluções para o problema de minimização por quadrados é um segmento de reta, confira o *caso linear*. Então uma das variáveis deste problema é  $m$  e sendo  $m < n$  existem muitos polinômios passando entre os pontos de  $L$  e o objetivo é obter também  $m$  mínimo que resolva o problema das equações (eq. 906)-(eq. 909).



O Na figura (fig 115), página 295, você pode ver uma simulação do que

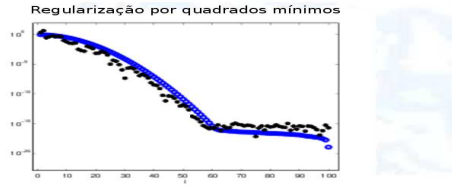


Figura 115: melhor curva adaptada

poderia ser um caso real e que você pode usar como apoio geométrico para ajudá-la a compreender o método.

Vou começar por definir um operador definido na *classe* de todos os polinômios de grau menor ou igual  $m < n$  com auxílio da equação (eq. 909). Desta forma vou transformar o problema numa questão de Cálculo Multivariado e usar as ferramentas do Cálculo para encontrar uma solução para o problema.

Para considerar este um problema definido no  $\mathbf{R}^n$  deixe-me introduzir a notação

$$\vec{u} = (x_1, \dots, x_n) \in \mathbf{R}^n; \quad (910)$$

$$P(x) = \sum_{k=0}^m a_k x^k; m < n; \quad (911)$$

$$\int_a^b (f(x) - P(x))^2 dx \approx \sum_{k=1}^n |f(x_k) - P(x_k)|^2 \Delta x_k = T_P(\vec{u}); \quad (912)$$

$$= \quad (913)$$

em que no segundo membro da equação (eq. 911) a leitora deve reconhecer uma função de  $n$  variáveis, definida com a *soma de Riemann* associada a partição do intervalo  $[a, b]$  representada pela seleção de pontos da pesquisa  $L$ , e que também depende do polinômio  $P$  que se procura oferecendo o valor mínimo para a expressão à direita na equação (eq. 911). O objetivo é encontrar um polinômio que minimize este problema. Se houver algum mínimo ele corresponde aos zeros do gradiente da função no segundo membro da equação (eq. 911).

Mas para calcular as derivadas parciais de  $T_P$  eu preciso acrescentar mais hipóteses ao problema:  $f$  tem que ser diferenciável. Assim vou substituir a função  $f$  na equação (eq. 906) por uma função qualquer satisfazendo a equação

(eq. 906) junto com ser diferenciável, posso agora calcular as derivadas parciais de  $T_P$  relativamente a cada uma das variáveis  $x_k$ :

(914)

- **modelo** modelo é uma palavra técnica que representa um método que dominamos, que sabemos implementar computacionalmente, por exemplo, e que deverá representar a realidade.

Há diversos tipos de modelos, para citar alguns, modelos polinômiais, com frequência representados pelos *splines*. Modelos aleatórios, em que se usam expressões obtidas usando *variáveis aleatórias* em lugar de coeficientes constantes.

Um exemplo simples de modelo (polinomial) são os polinômios de Taylor. em que uma certa função  $f$  da qual se conhecem poucos dados, é modelada por um polinômio.

- **morfismo** <sup>51</sup> As categorias são compostas de *objetos* e das *funções naturais entre estes objetos*, os morfismos. Confira *categoria, teoria das*. Em alguns textos é chamado de homomorfismo.

Um exemplo mostra melhor o que é um morfismo. Considere o grupo aditivo, um objeto da categoria *Grupo*, por exemplo, dos polinômios de grau menor ou igual à  $n$ ,  $(\mathbf{R}_n[x], +)$ . Este grupo é caracterizado, essencialmente, pelos coeficientes dos polinômios e assim podemos colocar em correspondência qualquer polinômio deste grupo com a enúpla de  $n + 1$  números reais dos coeficientes estabelecendo assim um *morfismo de grupo* entre  $(\mathbf{R}_n[x], +)$  e  $(\mathbf{R}_{n+1}, +)$ .

Se chamarmos esta correspondência de  $T$  temos:

$$P \in \mathbf{R}_n[x] \mapsto T(P) = (a_0, a_1, \dots, a_n) = p \in \mathbf{R}^{n+1} \quad (915)$$

$$P, Q \in \mathbf{R}_n[x] \mapsto T(P) = p, T(Q) = q \in \mathbf{R}^{n+1} \quad (916)$$

$$(P + Q) \in \mathbf{R}_n[x] \mapsto T(P) + T(Q) = p + q \in \mathbf{R}^{n+1} \quad (917)$$

$$0 \in \mathbf{R}_n[x] \mapsto (0, 0, \dots, 0) \in \mathbf{R}^{n+1} \quad (918)$$

então dizemos que  $T$  é um morfismo de grupos.

Um outro exemplo é o *logaritmo* entre  $(\mathbf{R}^{++}, \cdot)$  o grupo multiplicativo dos números reais estritamente positivos e  $(\mathbf{R}, +)$  o grupo aditivo de todos os números reais. Aqui teremos que escrever

$$(PQ) \xrightarrow{T} (\log(P) + \log(Q)) = p + q \quad (919)$$

e temos um *morfismo* entre o grupo multiplicativo dos números reais estritamente positivos e o grupo aditivo de todos os números reais.

A imagem inversa do elemento neutro do grupo imagem no grupo pré-imagem é um subgrupo chamado de *núcleo*. Aqui  $T^{-1}(0)$ , neste caso é o subgrupo com um único elemento 1, isto caracteriza que este morfismo é um

<sup>51</sup>Havia um erro, aqui, na definição de morfismo.

*isomorfismo.*  $\log \in \text{Mor}_{(R^{++}, \cdot), (R, +)}$ , o  $\log$  é um morfismo entre estes dois grupos.

Qualquer matriz  $m \times n$  define um *morfismo* entre os grupos aditivos de vetores  $\mathbf{R}^n, \mathbf{R}^m$ . Se  $m \neq n$  este morfismo não será um isomorfismo. Se  $n \leq m$  o núcleo, um subgrupo do espaço de saída  $\mathbf{R}^n$ , será o conjunto não trivial de soluções do sistema de equações com vetor de dados nulo e podemos com este sub-grupo construir a solução geral de qualquer sistema de equações definido com esta matriz com uma translação. Observe a semelhança de métodos, seja  $T \in \mathcal{L}(\mathbf{R}^n, \mathbf{R}^m)$ . Então  $\text{Ker}(T) = T^{-1}(0)$  é um subespaço do espaço de saída  $\mathbf{R}^n$  que é a solução do sistema homogêneo associado a  $T$ . Se este espaço se reduzir ao elemento neutro, um subespaço, então  $T \in \text{Mor}_{\mathbf{R}^n, \mathbf{R}^m}$  é um isomorfismo e  $n = m$ . O conjunto  $\text{Mor}_{\mathbf{R}^n, \mathbf{R}^m}$  é comumente chamado de  $\mathcal{L}_{n,m}$  é o espaço das matrizes  $n \times m$ .

Se estivermos estudando um categoria cujos elementos sejam funções, um tipo especial de morfismo pode ser construído, o *morfismo de avaliação* que em geral não será um isomorfismo. Em geral produz um exemplo de *functor esquecido*.

Este último exemplo mostra o poder de síntese que o uso de morfismos pode dar para o estudo das diversas estruturas algébricas, topológicas ou geométricas. Obviamente se acrescenta um nível maior de abstração com o que se consegue expressar de forma muito mais simples os problemas. Os autores da *teoria das categorias* a designaram no início como *general abstract nonsense*, ou seja *absurdo generalizado* e certamente é uma ironia que a metodologia de programação orientada a objetos esteja modelada pela *teoria das categorias*. Aparentemente *programação orientada a objetos* é uma derivação da *teoria das categorias*.

---

- **morfismo de avaliação**

Considere a estrutura de grupo aditivo da álgebra de Wiener, das séries trigonométricas absolutamente convergentes,  $(W(\mathbf{T}), +)$ . A função

$$W(\mathbf{T}) \xrightarrow{\delta_a} \mathbf{C}; \delta_a(f) = f(a) \quad (920)$$

em que  $f$  é a função definida por uma série trigonométrica absolutamente convergente, e  $\delta_a$  é o functor avaliação no ponto  $a \in \mathbf{T}$  é um exemplo de morfismo de grupo de  $W(\mathbf{T})$  no grupo dos números complexos:

$$(f + g)(a) = f(a) + g(a); 0(a) = 0 \quad (921)$$

Este morfismo é chamado de *morfismo complexo* da álgebra  $W(\mathbf{T})$  e todos os morfismos complexos desta álgebra são deste tipo.

---

- **multiplicação** Uma das quatro operações fundamentais da aritmética, as outras são *adição*, *subtração*, *divisão*.

- **multiplicação, operador** Em uma estrutura algébrica em que há uma operação caracterizada como *multiplicação*, quer dizer, não necessariamente comutativa podemos definir um *morfismo* desta estrutura fixando um elemento,

digamos  $\mathbf{a}$  e definindo

$$f \mapsto \mathbf{a}f; \text{ ou então } f \mapsto f\mathbf{a}; \quad (922)$$

Se a estrutura algébrica mencionada for um grupo, estes operadores, a *multiplicação à esquerda* ou *multiplicação à direita* são dois *isomorfismos de grupo*.

No exemplo anterior, no caso de um grupo multiplicativo, se tem um isomorfismo porque todo elemento de um grupo inversível, logo os inversos deste isomorfismos são, respectivamente

$$f \mapsto f\mathbf{a}^{-1}; \text{ ou então } f \mapsto \mathbf{a}^{-1}f; \quad (923)$$

**Definição 15 (Operador)** *Multiplicação*

**Notação**  $M_{\mathbf{a}} : f \mapsto \mathbf{a}f;$

*Em caso de ambiguidade se pode usar  $M_{\mathbf{a}, \mathbf{a}}M$  para representar estes operadores.*

Surgem fatos interessantes quando a inversibilidade do elemento  $\mathbf{a}$  pode ser posta em questão então os operadores multiplicação não são inversíveis: as funções não inversíveis (não injetivas) definem classes de equivalência não triviais no domínio com uma fatorização também não trivial:

$$M_{\mathbf{a}} = i_{\mathbf{a}}os_{\mathbf{a}}; \mathbf{a}M = \mathbf{a}io_{\mathbf{a}}s \quad (924)$$

em que  $s$  é uma sobrejeção, sobre o conjunto das classes de equivalências geradas pelo operador no domínio e  $i$  injeção sobre a imagem do operador. Esta decomposição (ou fatoração) vale para qualquer tipo de morfismo. É a fatoração canônica do morfismo. Em particular vale para os *morfismos de conjuntos*, as *funções*.

Alguns exemplos de aplicação:

1.  $M_{\mathbf{a}}l^p \implies l^p$  em que  $\mathbf{a}$  é um elemento de  $l^q$  com  $\frac{1}{p} + \frac{1}{q} = 1$  é um operador linear entre estes dois espaços (de Banach) de sucessões. A desigualdade de Hölder é que justifica este resultado.
2. Como o produto ponto a ponto de sucessões é transformado num produto por convolução em algum espaço de funções definidas em  $R$  então o exemplo anterior se apresenta com o operador multiplicação por convolução:  $M_{\mathbf{a}}L^p \implies L^p$  se  $\mathbf{a}$  for um elemento de  $L^q$  com  $\frac{1}{p} + \frac{1}{q} = 1$ , também é consequência da desigualdade de Hölder.
3. Se  $\mathbf{a} \in L^\infty$  então o operador multiplicação  $M_{\mathbf{a}}L^p \implies L^p$  estará bem definido e é contínuo para qualquer que seja  $1 \leq p < \infty$ . O mesmo pode ser dito no caso dos espaços de sucessões.

---

- **natural, número  $\mathbf{N}$**  é o conjunto dos números naturais.

Os números naturais são *fracamente*<sup>52</sup> definidos por nove axiomas de Peano.

---

<sup>52</sup>Fracamente porque os axiomas de Peano também se aplicam ao conjunto  $\{-3, -2, -1, 0, 1, 2, \dots\}$ . Na verdade Peano definiu a indução finita.

1. O conjunto dos números naturais não é vazio e portanto contém pelo menos um elemento cujo símbolo é 0. Para Peano este elemento seria 1. Este é um dos pontos críticos dos axiomas de Peano, nada nos impediria de considerar que o primeiro elemento de  $\mathbf{N}$  fosse o número inteiro  $-3$  ou  $3$  e toda a axiomática de Peano funcionaria. Na verdade os axiomas de Peano são a formulação do *princípio da indução finita* e  $\mathbf{N}$  é um dos possíveis resultados da aplicação deste princípio.

2. Da igualdade:

- (a) reflexividade  $\forall x \in \mathbf{N}; x = x$
- (b) transitividade  $\forall x, y, z \in \mathbf{N}; x = y$  e  $y = z$  implica  $x = z$ ;
- (c) simetria  $\forall x, y \in \mathbf{N}; x = y$  implica que  $y = x$ ;

3. Fechado para igualdade Se  $a \in \mathbf{N}$  e  $b = a$  então  $b \in \mathbf{N}$ .

O uso deste axioma ocorre em situação como operações com frações, se o resultado final for, por exemplo,  $\frac{3a}{a}$ , com  $a \neq 0$  então  $\frac{3a}{a} \equiv 3$ , e como  $3 \in \mathbf{N}$  não há razão para considerar estes dois objetos como diferentes e ambos pertencem ao conjunto  $\mathbf{N}$ .

4. A operação sucessor Existe uma operação,  $s$  com as seguintes propriedades:

- (a)  $s(n) \in \mathbf{N}$  para todo  $n \in \mathbf{N}; s(0) = 1$ ;
- (b)  $s(s(0)) = 2; s^n(0) = n$ ;
- (c) Para qualquer  $n \in \mathbf{N}$   $s(n) = 0$  é falso;
- (d)  $s$  é injetiva:  $s(n) = s(m) \Rightarrow n = m$ ;

5. O princípio da indução finita

Considere  $K$  um subconjunto de  $\mathbf{N}$  que tenha as propriedades seguintes:

- (a)  $0 \in K$ ;
- (b)  $q \in K \Rightarrow s(q) \in K$

então  $K = \mathbf{N}$ .

**Teorema 15** (infinitude dos números naturais)  $\mathbf{N}$  é infinito

**Dem**:

Como  $s(s(0)) = s(0)$  implica que  $s(0) = 0$  pela injetividade de  $s$  então, por absurdo,  $s(0) \neq 0$  ou, equivalentemente,  $s(s(0)) \neq s(0)$ . **q.e.d.**

Aplicação repetida, indefinidamente, deste teorema mostra que  $\mathbf{N}$  não tem um número finito de elementos e isto caracteriza que  $\mathbf{N}$  é um conjunto infinito.

A construção feita por Peano é minuciosa, ele pensou em todos os detalhes, e inclusive, a construção de Peano foi em parte produzida independentemente

por Frege sem que nenhum dos dois soubesse dos trabalhos do outro. Ainda assim é possível perceber que o conceito *conjunto infinito* se mantém vago uma vez que ele depende de uma *aplicação repetida, indefinidamente, de uma certa operação* e esta “ação” não poderia ser nunca executada. Observações deste tipo conduziram aos trabalhos de um dos fundadores da computação, Turing, e seu célebre teste, a *máquina de Turing*.

A Matemática, resolve problemas práticos, mas ela está longe de ser uma teoria perfeita ou exata! Talvez isto seja o seu aspecto mais forte, as suas falhas, uma permanente fonte de inspiração de nossas pesquisas.

---

- **núcleo** Tem vários sentidos em Matemática e é usada frequentemente com o mesmo sentido que kernel.

- **Em Teoria dos Grupos** (e nas supercategorias da categoria de grupo) representa o subgrupo obtido como imagem inversa do elemento neutro da adição por um morfismo (entre dois grupos não comutativos é a imagem inversa por um morfismo do elemento neutro da multiplicação do grupo-imagem). Desta forma se caracterizam os isomorfismos quando o núcleo se reduz ao grupo formado apenas pelo elemento neutro. O núcleo é um subgrupo normal.
- **Em teoria dos operadores** é o nome que se usa para as funções que ocupam o lugar das matrizes quando usamos integrais para definir funções lineares:

$$J(f)(y) = \int_E f(x)K(x, y)d\mu(x) = \phi(x); \phi = J(f) \quad (925)$$

define um operador linear entre dois espaços de funções definidas no (espaço de medida  $(E, \mu)$ ).

- **Com as séries de Fourier** usamos o kernel  $e^{inx}$  para transformar funções definidas em  $E = [-\pi, \pi]$  em funções definidas em  $\mathbf{Z}$  os coeficientes complexos de Fourier de uma função definida em  $E = [-\pi, \pi]$

$$\hat{f}(n) = \int_{-\pi}^{\pi} f(x)e^{inx} = c_n; c = \hat{f} \quad (926)$$

portanto o operador linear definido na equação (925) pode transformar elementos de espaços bem diferentes.

Usando o núcleo de Dirichlet podemos transformar este operador integral num operador por convolução.

- **Uma função positiva cuja integral é 1** é também chamada de núcleo e tais funções, quando forem contínuas e a medida do suporte seja pequena, são usadas em produtos por convolução para regularizar funções. Confira *convolução* para ver mais a este respeito.

- **núcleo de Dirichlet** a sucessão de funções

$$D_n(x) = \frac{\sin nx}{x} \quad (927)$$

convergem no sentido das distribuições para a distribuição de Dirac e desta formam representa uma *unidade aproximada* relativamente á operação de convolução,

$$D_n * f \approx f \quad (928)$$

em que a convolução é definida como uma integral sobre  $\mathbf{R}$ . Como a sucessão é *quase limitada* uma condição suficiente para que esta sucessão de integrais convirja é que  $f$  seja integrável. A dedução vem da expressão complexa da série de Fourier

$$f(x) \approx \sum_{k=-n}^n c_k e^{ikx} = \sum_{k=-n}^n \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-ikt} f(t) dt \right) e^{ikx} = \quad (929)$$

$$= f(x) \approx f * D_n(x); D_n(x) = \sum_{k=-n}^n e^{ikx} \quad (930)$$

$$D_n(x) = \frac{\sin(nx)}{\sin(x)} \quad (931)$$

$D_n$  é uma função  $2\pi$ -periódica e a medida que  $n$  cresce as oscilações no intervalo  $[-\pi, \pi]$  aumentam por que a aceleração  $n$  aumenta o número de raízes de  $\sin(nx)$ .

- **número** são os objetos básicos da Matemática que popularmente é chamada de “*ciências dos números*” embora hoje ela seja muito mais do que isto, confira *álgebra, análise, geometria, estatística e lógica* que são as divisões clássicas da Matemática embora seja difícil classificar todos os tópicos da Matemática apenas entre estas divisões. Confira *divisões da Matemática*.

Os números, em Matemática, se classificam entre as seguintes conjuntos:

$$\mathbf{N} \subset \mathbf{Z} \subset \mathbf{Q} \subset \mathbf{R} \subset \mathbf{C} \quad (932)$$

chamados, respectivamente, *naturais inteiros, racionais, reais, complexos*.

Os números naturais são um objeto mais próprio da *lógica matemática*, os inteiros são mais próprio da *álgebra*. Os números racionais, reais e complexos são estudados mais intensivamente na *análise matemática*.

Os números complexos criaram um ramo da análise matemática que é chamada de análise complexa e que está à base da *teoria das funções* e da *teoria das funções holomorfas* ou *teoria das funções analíticas*.

Confira também

- número algébrico, confira algébrico, número
- número complexo, confira complexo, número  $\mathbf{C}$

- número natural, confira natural, número **N**;
- número racional, confira racional, número **Q**;
- número real, confira real, número **R**;
- número transcendente, confira transcendente, número **R**

- **onda** A *equação diferencial parcial da onda*, frequentemente dita simplesmente *equação da onda* é uma das equações diferenciais parciais clássicas, no sentido das que fazem parte do conjunto das equações cujas soluções se conhecem desde o século 19. Ver *equação diferencial parcial* para maiores detalhes sobre este assunto.

A equação da onda é uma equação envolvendo derivadas parciais:

$$u_{tt} - c^2 u_{xx} = 0 \quad (933)$$

Ela traduz o movimento que um impulso de energia produz numa corda. Podemos ouvir música ao piano, de uma guitarra, violão, etc. . . porque este instrumento tem cordas presas nas duas extremidades e elas têm dois coeficientes, de *elasticidade* e de *tensão* que um músico usa para *afinar* o seu instrumento.

Quando o martelo, no piano, bate numa corda, ele produz, mecanicamente, na mesma um impulso de energia que em princípio ficaria a percorrê-la entre as extremidades *indefinidamente* fazendo-a vibrar com uma frequência definida pela *afinação* e uma intensidade definida pela força com que o martelo é acionado. Mas esta energia se dispersa, se transforma, nas extremidades e com a resistência do ar, ou pode ser imediatamente “apagada” pela ação de silenciadores, no piano, ou pelos dedos do músico no demais instrumentos de corda. A mesma corda de piano que se encontre solta não produzirá nenhum som se nela batermos com um martelo, isto sugere uma questão essencial em *equações diferenciais* que são as condições iniciais ou as condições de fronteira do problema em questão. Fazemos música, com a equação da onda.

A figura (fig 116), página 303, é uma sugestão do que ocorre numa corda de instrumento musical usando a função *seno* e uma função impulso positivo. A *função impulso positivo*, com gráfico em vermelho, representa o efeito do martelo, no piano, gerando uma onda na corda. As ondas do mar são outro exemplo e de certa forma muito semelhante à música, uma grande variedade de impulsos atuando sobre uma superfície também presa num suporte<sup>53</sup>, a água do mar, se apenas observarmos a massa d’água. Entretanto este exemplo está contido num fenômeno maior que é o tempo e a equação diferencial para tratá-lo são as equações de Navier–Stokes que descrevem o comportamento dos flúidos. Ainda assim, de pé na praia você pode observar um fenômeno que é semelhante ao da onda numa corda de instrumento musical. Um exemplo mais simples

<sup>53</sup>É como se água do mar estivesse dentro de uma garrafa com ar comprimido, e é isto que acontece com os mares, a pressão atmosférica representa o ar comprimido e as praias a fronteira onde a membrana está presa.



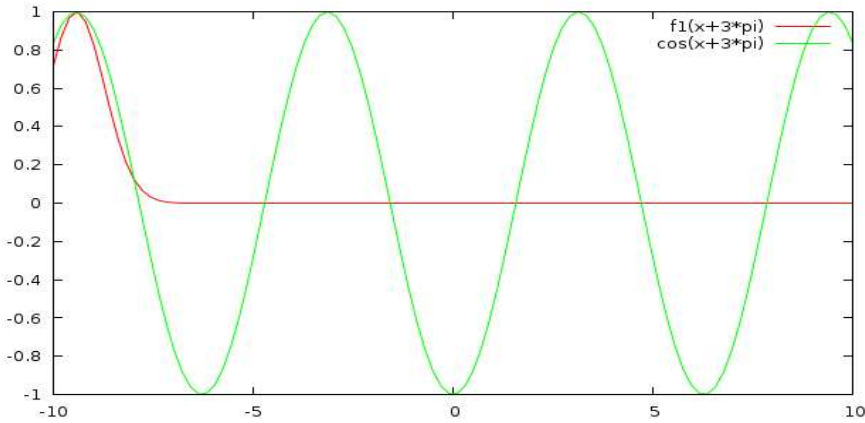


Figura 116:

do a superfície do mar você pode reproduzir em casa com um lençol sobre a cama. Prenda com as mãos uma lateral e dê um impulso, você vai ver uma onda tridimensional percorrer o lençol e morrer no lado oposto e aqui você pode observar a ausência da tensão que permite as vibrações na corda do instrumento musical. Um exemplo melhor é película de couro ou plástico de um tambor musical. Como ela se encontra presa ao longo de sua *fronteira* haverá vibração durante algum tempo transmitida ao ar e finalmente aos seus tímpanos. No caso do lençol da cama você induz uma onda que morre na outra extremidade.

A equação diferencial da onda costuma ser apresentada de forma mais completa, ou complexa do que a equação (933)

$$y = u(t, x) \quad (934)$$

$$u_{tt} - c^2 u_{xx} = 0; x \in [0, l] \quad (935)$$

$$u(0, x) = a(x); u_t(1, x) = b(x); \quad (936)$$

formando um *problema* com valores na fronteira, PVF ou na sigla ingles, BVP. Os valores na fronteira estão representados aqui pelas funções  $a, b$  que dão valores à expressão da onda quando o tempo é zero ou 1. O tempo foi escolhido variar no intervalo  $[0, 1]$ , mas isto nada mais é do que uma escolha, e  $x \in [0, l]$  em que  $l$  é o comprimento da corda. Este problema descreve o comportamento de uma corda de instrumento musical. A película do tambor tem que ser descrita por um número maior de variáveis que pode ser obtida ampliando a equação (933)

$$u_{tt} - c^2(u_{xx} + u_{yy} + u_{zz}) = 0; \quad (937)$$

Mas como já observei na discussão acima, uma equação como (937) é inútil, como no caso do lençol, se a “membrana” não estiver presa a uma determinada fronteira, um problema de valores na fronteira, PVF, nada acontece de muito interessante.

É fácil ver a importância do domínio associado a uma equação, mais exatamente a influência da fronteira, um experimento interessante e de fácil realização está descrito no verbete *PVF*, página 369 em que a “película” é assucar refinado. Ele mostra, e você mesmo pode repetir a experiência, que o formato da película depende da fronteira do domínio, claro que tem mais *condições de fronteira* envolvidas nesta questão, como a pressão atmosférica, a energia interna das moléculas do material utilizado.

- **operador integral** É o nome de uma transformação entre espaços de funções cuja equação se expressa com uma integral

$$T(f)(y) = \int_{\Omega} f(x)K(x, y)d\mu(x) = g(y) \quad (938)$$

em que  $\Omega$  é o espaço (medido) em que  $T$  está definido. A função  $K$  se chama, neste contexto, *kernel* do operador. Um exemplo importante e elementar é dado pelo núcleo de Dirichlet que é uma forma simples de expressar a transformada de Fourier.

Se você expressar uma aproximação da equação (938) num espaço de *funções simples* poderá ver que os operadores integrais são uma direta generalização das transformações matriciais com o *kernel* substituindo uma matriz.

- **operador de Laplace** é um operador linear definido num espaço de funções diferenciáveis:

$$-\Delta(F) = \frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} = 0 \quad (939)$$

no caso de funções de duas variáveis. O operador de Laplace é também chamado de *laplaciano*.

A semelhança formal da expressão do laplaciano com o “produto escalar” do *operador gradiente*,  $\nabla$ , por ele mesmo fez com que se usasse  $\nabla^2$  como notação para o operador de Laplace. Desta forma você pode encontrar as duas notações,  $\Delta(F) = \nabla^2(F)$  para representar o *laplaciano* aplicado à função  $F$ , mas a primeira notação está predominando na literatura. Esta é a expressão real, quando as variáveis são números reais, do operador, há uma outra forma de expressá-lo quando as variáveis são complexas.

A equação (939) é uma equação diferencial, com ela se procuram as funções duas vezes diferenciáveis continuamente que satisfaçam esta expressão. A partir de 1960 o hábito de escrever a equação de Laplace com o sinal negativo passou a predominar, é esta a razão da formulação na equação (939)

Como a derivação é um operador linear definido no espaço de funções suficientemente deriváveis continuamente, então o *laplaciano* é um operador linear e a equação (939) é um exemplo de *equação diferencial linear a derivadas parciais homogênea*, ou de forma mais curta, como é habitual, *equação diferencial parcial linear*, é homogênea porque o segundo membro é a função constante nula.

Esta equação, no caso homogêneo, como muitas outras equações diferenciais, foi resolvida indiretamente, como resultado do esforço imenso dos matemáticos

dos séculos 16, 17,18 para entender as equações diferenciais. Qualquer solução desta equação se diz uma *função harmônica* e um *resumo* da teoria das funções de variáveis complexas seria:

**Teorema 16 (funções ) analíticas**

Se  $F = u + iv$  for uma função analítica, então  $u, v$  são harmônicas.

$$\Delta(u) = 0 = \Delta(v) \quad (940)$$

Recíprocamente, se  $u$  for uma função harmônica então existe uma outra função harmônica,  $v$ , chamada complemento harmônico de  $u$  tal que  $u + iv$  é uma função analítica. **Dem**:

O caso ( $\Rightarrow$ ) é consequência direta das equações de Cauchy-Riemann e do teorema de Schwarz-Clairaut da igualdade das derivadas mistas.

O caso ( $\Leftarrow$ ) não é tão imediato. Estas duas funções harmônicas ainda são chamadas conjugados harmônicos, ou seja, se  $u, v$  forem conjugados harmônicos então  $u + iv$  é uma função analítica. Isto é, para toda função harmônica  $u$  existe uma única  $v$ , harmônica, a menos de uma constante, tal que  $u + iv$  é uma função analítica: esta é a solução do sistema de equações diferenciais de primeira ordem chamado de equações de Cauchy-Riemann tendo  $u$  como uma "constante dada".

**q.e.d .**

O estudo dos operadores diferenciais, e o operador de Laplace é um exemplo, tem como foco a busca de propriedades das soluções e neste sentido uma das técnicas importantes e a determinação dos *vetores próprios*. Confira *valor próprio* e *vetor próprio* da equação de Laplace.

**- ordem**

**Definição 16 (Relação) de ordem**

Seja  $A$  um conjunto. Uma relação binária  $S$  definida em  $A$  é de ordem se

1. for transitiva,  $S(x, y), S(y, z) \Rightarrow S(x, z)$ .
2. Se simétrica,  $S(x, y), S(y, x)$  então  $x = y$ .

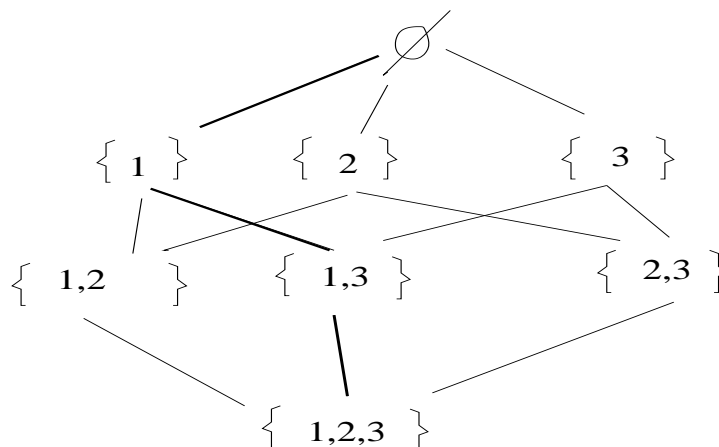
As relações de ordem podem ser:

1. *estritas* ou *largas*. É *larga* quando for reflexiva, isto é,  $S(x, x)$  é verdadeira para todo  $x$ . Nos conjuntos numéricos a ordem "menor ou igual" é *larga* e "menor do que" é *estrita*.
2. *total* ou *parcial*. Se dado um par  $(x, y) \in A^2$  uma das relações  $S(x, y), S(y, x)$  sempre for verdadeira, então  $(A, <)$  é totalmente ordenado. No caso contrário dizemos que é parcialmente ordenado. A figura (117), página 306, mostra o conjunto das partes de  $\{1, 3, 3\}$  como exemplo de um conjunto parcialmente ordenado.

No conjunto das partes de um conjunto qualquer não vazio, a inclusão é uma ordem *larga* que não é total porque há subconjuntos que não são comparáveis usando a inclusão.

A ordem nos conjuntos numéricos satisfaz à tricotomia portanto é *total*.

## Conjunto parcialmente ordenado



### Uma cadeia do conjunto $P(\{1,2,3\})$ marcada com linha mais grossa.

Figura 117: Conjunto parcialmente ordenado  $P(\{1,2,3\})$

Dizemos então que  $(A, <)$  é uma *estrutura de ordem*, ou simplesmente uma *ordem*.

Exemplos de relação de ordem:

1. A desigualdade habitual dos conjuntos numéricos  $\mathbf{N}$ ,  $\mathbf{Z}$ ,  $\mathbf{Q}$ ,  $\mathbf{R}$ . O conjunto dos números complexos,  $\mathbf{C}$  não é ordenado, habitualmente. Mas podemos definir em  $\mathbf{C}$  a ordem *alfabética* porque  $\mathbf{C}$  é um produto cartesiano de conjuntos ordenados:  $\mathbf{C} = \mathbf{R}^2$ , porém esta ordem não é compatível com a ordem de  $\mathbf{R}$  como subconjunto de  $\mathbf{C}$ .
2. Se  $(A, <)$  for ordenado, os produtos cartesianos de  $A$  têm uma ordem natural que é a alfabética:  
 $(x, y) << (a, b)$  sse  $x < a$  e  $y < b$
3. A relação “ $p$  divide  $q$ ” nos inteiros positivos é uma ordem larga. Este é um exemplo de ordem que não é total, há pares de números  $(p, q)$  para os quais não precisa ser verdade “ $p$  divide  $q$ ” e nem “ $q$  divide  $p$ ”.
4. A *implicação* é uma ordem numa classe de sentenças.

O último exemplo mostra a amplitude do conceito de ordem e vamos aqui nos restringir à ordem definida em conjuntos numéricos. Se  $(A, <)$  for uma

ordem, existem subconjuntos especiais de  $A$  relativamente a ordem que são as *cadeias*:

**Definição 17 (Cadéia)** *de um conjunto ordenado*

*Considere em  $(A, <)$  um subconjunto  $T \subset A$  tal que  $T$  é totalmente ordenado. Dizemos que  $T$  é uma cadeia de  $A$ .*

A figura (117), página 306, mostra uma cadeia do conjunto das partes de  $\{1, 2, 3\}$  indicada com uma linha mais grossa.

Vejamus uma propriedade que distingue  $\mathbf{Z}$  como conjunto numérico. Para entendê-la melhor, vamos começar com o conjunto  $\mathbf{Q}$  onde esta propriedade não vale. Considere um “intervalo limitado” de números racionais, um intervalo é uma *cadeia*. Há dois tipos, *intervalos abertos* e *intervalos fechados*. Os *intervalos abertos* não tem nem máximo e nem mínimo, ao passo que os *intervalos fechados* têm máximo e mínimo. Portanto em  $\mathbf{Q}$  existem cadeias limitadas que não têm mínimo (e nem máximo).

Em  $\mathbf{Z}$ , toda cadeia limitada inferiormente, tem um mínimo. Toda cadeia limitada superiormente, tem um máximo. Dizemos então que  $\mathbf{Z}$  é bem ordenado, e que  $\mathbf{Q}$  não é bem ordenado.  $\mathbf{R}$  também não é bem ordenado, devido aos intervalos abertos.

Se considerarmos o conjunto das partes de  $\{1, 2, 3\}$  temos um outro exemplo de conjunto bem ordenado uma vez que  $\{1, 2, 3\}$  é o mínimo de qualquer cadeia.

Mas se omitirmos o conjunto vazio e o próprio conjunto  $\{1, 2, 3\}$  teremos um conjunto que não é *bem ordenado*. Neste caso aparece um novo conceito: *elemento minimal*, há três elementos minimais:  $\{2, 3\}$ ,  $\{1, 3\}$ ,  $\{1, 2\}$  e três elementos maximais  $\{2\}$ ,  $\{1\}$ ,  $\{3\}$

- **Oscilação** é um funcional definido num espaço vetorial de funções reais ou complexas que generaliza o conceito de *limite*.

Considere uma função real ou complexa definida em num subconjunto  $\Omega$  de  $\mathbf{R}^n$ . A oscilação de  $f$  em  $x$  é a diferença

$$\text{caso real } \omega_f(x) = \limsup_x(f) - \liminf_x(f) \quad (941)$$

$$\text{caso complexo } \omega_{|f|}(x) \quad (942)$$

Por exemplo se

$$f(x) = x \sin\left(\frac{1}{x}\right)$$

que é contínua para qualquer número real diferente de zero, é interessante observar o que acontece com  $\omega_f(0)$  que é zero, logo esta função é uma função contínua.

Se  $f$  for uma função contínua então  $\omega_f$  é identicamente zero o que mostra a inutilidade deste conceito para tais funções. É interessante o conceito para funções que não sejam contínuas ou para verificar a continuidade duma função particular.

Um outro conceito que usa *oscilação* é a oscilação média, conhecida pela denominação inglesa BMO, bounded mean oscillation.

$$BMO_f(x) = \frac{1}{\Omega} \int_{\Omega} \omega_f(x) dx \quad (943)$$

é o valor médio integral de  $\omega_f(x)$  numa vizinhança de  $x$ . Pode-se provar que  $BMO(\Omega)$  é um espaço vetorial em que  $\Omega \subset \mathbf{R}^n$ .

- **pantógrafo** de Silvestre, é um instrumento de desenho baseado num paralelograma cujos ângulos que se opõem se transformam permitindo que as duas extensões dos seus lados paralelos leiam e copiem um desenho. Confira a figura (fig 118), página 308, em que aparece um *pantógrafo comercial*, mas que você

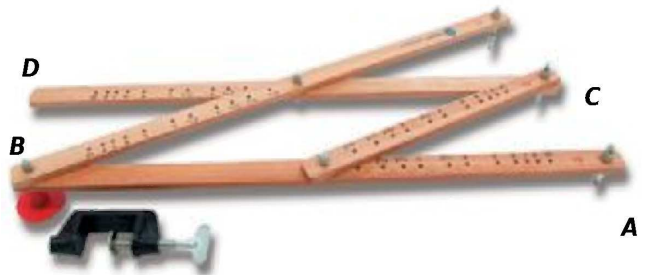


Figura 118: um pantógrafo de madeira

pode construir um, você mesma. Na figura (fig 119), página 309, aparece o *pantógrafo* de forma *abstrata*, substituídas as hastes por segmentos de reta. O pantógrafo é um polígono *não regular* que contém um *subpolígono de quatro lados* que é um retângulo<sup>54</sup>.

O *instrumento* se compõe de quatro hastes de madeira presas duas a duas formando, *os dois lados não isométricos do paralelograma*. Os dois vértices, “B” e “C”, confira a figura (fig 118), são presos com *parafusos e porca*, sem grande pressão para não estragar o mecanismo, mas eles devem ficar sempre presos com a possibilidade de movimento (rotação). Isto pode ser obtido, na prática, incluindo *arruelas lisas* entre as cabeças dos parafusos e as porcas.

É preciso que as hastes sejam de tamanhos distintos porque, numa delas, na maior, se praticam *furos*, a distâncias regulares, permitindo assim obter

<sup>54</sup>E não precisa ser regular, você pode selecionar os orifícios de forma a obter um trapézio e a consequência será uma deformação no desenho copiado. . .

*paralelogramas de tamanhos diversos* para reproduzir as figuras alterando-lhes os tamanhos. Desta forma o *pantógrafo* lhe permite:

1. copiar desenhos;
2. ampliar desenhos;
3. miniaturizar desenhos;
4. deformar desenhos se você escolher um trapézio em lugar dum paralelograma quando configurar os furos.

Se você for construir o seu pantógrafo, leve em conta que os furos de regulação nas hastes devem ser de diâmetro o menor possível para não enfraquecer as mesmas.

Uma das hastes tem seu extremo fixo, “A”, de modo que sua outra extremidade, “B”, um dos vértices do paralelograma, se mova em cima dum *círculo virtual*. Confira a figura (fig 120), página 312. Logo vou fazer uso deste *círculo virtual*.

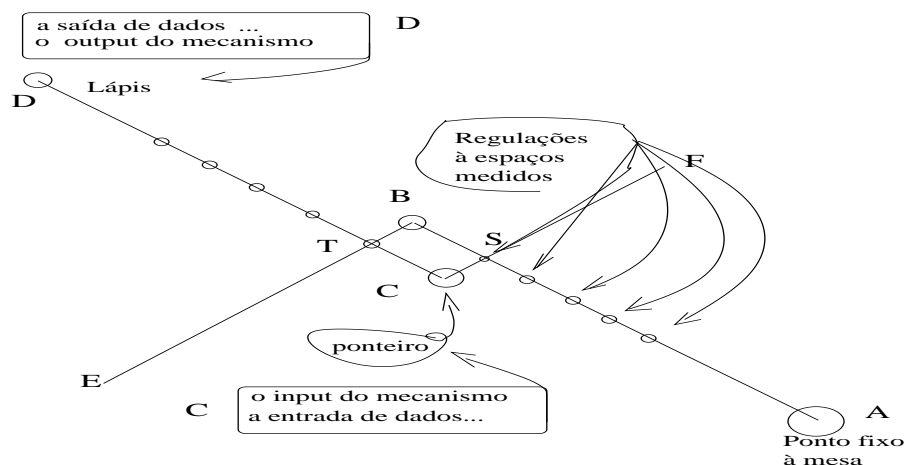


Figura 119: A forma *abstrata* do pantógrafo

*virtual*.

O vértice que se opõe ao “B”, “C”, é usado como um *ponteiro* para ler a figura a ser reproduzida. É o *input* do mecanismo...

No outro extremo do lado determinado pelo ponteiro, que é homólogo ao lado que tem o extremo fixo, “D”, é onde se encontra o *lápis* que irá reproduzir a figura. É o *output* do mecanismo... a *saída de dados*. Na figura (fig 119), página 309, estas funções estão etiquetadas.

Faça um, você mesma, e ficará convencida de que tem um *mecanismo* simples, mas importante, em suas mãos.

E se você quiser *inventar*, aqui você tem a oportunidade:

1. aproveite os dois pontos extra, você também pode usar o ponto “B” assim como os pontos ficam sem denominação nos *pantógrafos comerciais* chamados de “E” e “F”, confira a figura (fig 119), como saídas de dados, e obter mais de uma cópia. O ponto “B” não é usado nos *pantógrafos comerciais*, certamente porque fica muito próximo do ponteiro em “C”.
2. pouco aconselhável, aumente a quantidade de *pernas* do pantógrafo criando mais retângulos semelhantes e isto irá permitir-lhe de fazer várias cópias dum mesmo desenho numa única vez. Apenas *não exagere*, porque pode ficar muito difícil de manejar o mecanismo.

Os *pantógrafos comerciais*, feitos em madeira, costumam ter as letras A,B,C,D indicando os quatro vértices sendo

1. O vértice fixo à mesa é o “A”,
2. o vértice de *saída de dados* é o “D”, o *output*
3. os dois vértices presos por parafusos, ligando as hastes de tamanhos diferentes, são denominados “B” e “C”. Vou adotar a letra “C” para representar a entrada de dados, o *input*.

Observe que esta denominação das extremidades das hastes do pantógrafo pode variar de uma construção para outra e inclusive você pode encontrar *pantógrafos comerciais* sem estas letras inscritas. Mas vou usar esta denominação, consistentemente, neste artigo e também fazer referência aos pontos “E”, “F” como saída de dados extra, confira a figura (fig 119).

### Calibração do instrumento

Antes de usar o pantógrafo será preciso, *calibrar* o instrumento para garantir que o ponteiro consiga atingir todos os pontos da figura a ser copiada.

Nas hastes  $\overline{AB}$  e  $\overline{CD}$  se encontram os orifícios de regulação do pantógrafo para selecionar o tamanho do paralelograma. Estes orifícios vão ser usados aos pares de modo que apenas dois, em cada uma das hastes, estarão sendo usados. Vou designar:

- $S \in \overline{AB}$ ;
- $T \in \overline{CD}$ ;

confira a figura (fig 119). A única utilização dos pontos  $S, T$  é a calibração do pantógrafo.

A *calibração* do pantógrafo consiste em selecionar o ponto fixo “A” para que o ponteiro possa atingir todos os pontos da figura a ser transformada. O mais fácil será *calibrar* a figura, movendo-a para uma posição de alcance do pantógrafo, ou seja, dentro do *círculo virtual*, (fig 120). Vou fazer a *calibração*



*matemática* do pantógrafo virtual na próxima seção. A capacidade de acesso do pantógrafo é dado pelo círculo que tem raio  $\overline{CS}$  e a figura a ser copia tem estar dentro deste círculo, confira a figura (fig 119). Faz parte também da calibração a seleção dos pontos  $S, T$  para definir o tamanho do paralelograma e decidir pelo efeito do pantógrafo: *copiar, dilatar ou miniaturizar*.

### A Matemática do pantógrafo

A Matemática do pantógrafo é o teorema de Tales, apenas usado tão intensivamente que, na verdade, o pantógrafo produz *curvas homotópicas a um ponto*.

O *pantógrafo* é um instrumento de *topologia* aplicada! Confira a figura (fig 120), página 312.

Estou começando o *processo de abstração* com dois objetivos,

- descrever a Matemática do *pantógrafo*,
- preparar a construção dum *programa* que irá simular um *pantógrafo virtual*.

Como estou construindo uma abstração do pantógrafo, vou supor que o ponto fixo  $A$  fique dentro curva  $\gamma$ , o que seria impossível com o pantógrafo de madeira<sup>55</sup> e que o *círculo virtual* descrito por  $B$  seja o círculo trigonométrico  $\mathbf{S}^1$ . É a *calibração matemática* do pantógrafo virtual. O alcance do pantógrafo é o círculo com raio  $\overline{CS}$  e a curva  $\gamma$  deve estar contida neste círculo. Não projetei furos de regulação nas hastes  $\overline{CF}, \overline{BE}$ , mas obviamente que isto pode ser feito e irá permitir um alcance maior ao pantógrafo porque permitirá aumentar o raio  $\overline{CS}$ . Embora este detalhe seja importante para o *pantógrafo de madeira*, ele não tem maior importância para o *pantógrafo virtual* de que estou me ocupando de agora em diante.

Vou usar um parâmetro  $t$  para representar o *tempo* ao longo do qual a ação do pantógrafo vai operar sobre a curva  $\gamma$ .

Desta forma em cada “momento”  $t$  há um ponto,  $M(t)$ , de interseção do segmento  $\overline{AB}$  com a curva  $\gamma$ ,  $\gamma(t)$ , definindo  $B(t)$  como projeção de  $\gamma(t)$  sobre o  $\mathbf{S}^1$ , fica, assim, definida a correspondência

$$\gamma(t) \mapsto \alpha(t); \gamma \rightarrow \alpha \subset \mathbf{S}^1; \quad (944)$$

A curva  $\alpha$  assim obtida, contida em  $\mathbf{S}^1$ , é um “*equivalente homotópico a um ponto*” de  $\gamma$ , com a aparência de círculo, mas logo você verá que não é um círculo.

---

<sup>55</sup>Não seria impossível, apenas haveria que começar a transformação *zerando* a posição do pantógrafo ao encostá-lo no ponto fixo e seguir o acompanhamento da curva até encontrar o ponto fixo do outro lado da rotação, apenas não é uma metodologia prática. . .

Estou chamando esta curva de  $\alpha$  e ela é também um “*equivalente homotópico a um ponto*” de  $\beta$ , uma vez que  $\beta$  é um “*equivalente homotópico a um ponto*” de  $\gamma$ .

A relação “*equivalente homotópico a um ponto*” é uma relação de equivalência.

A equação de  $\alpha$  é

$$t \mapsto (\cos(2\pi f_1(t)), \sin(2\pi f_2(t))) = (\alpha_1(t), \alpha_2(t)); t \in [0, 1]; \quad (945)$$

$$f_1, f_2 : [0, 1] \rightarrow [0, 1]; \text{ diferenciáveis }; \quad (946)$$

parametrizada em  $[0, 1)$ .

O ponto  $M$  em que o segmento  $\overline{AB}$  corta a curva  $\gamma$  é transformado no ponto  $B \in \mathbf{S}^1$  para cada valor do parâmetro  $t \in [0, 1]$ . Podemos desta forma considerar todas as curvas,  $\alpha, \gamma, \beta$  sincronizadas com o mesmo intervalo de parametrização  $[0, 1)$ .

Confira a figura (fig 120), página 312.

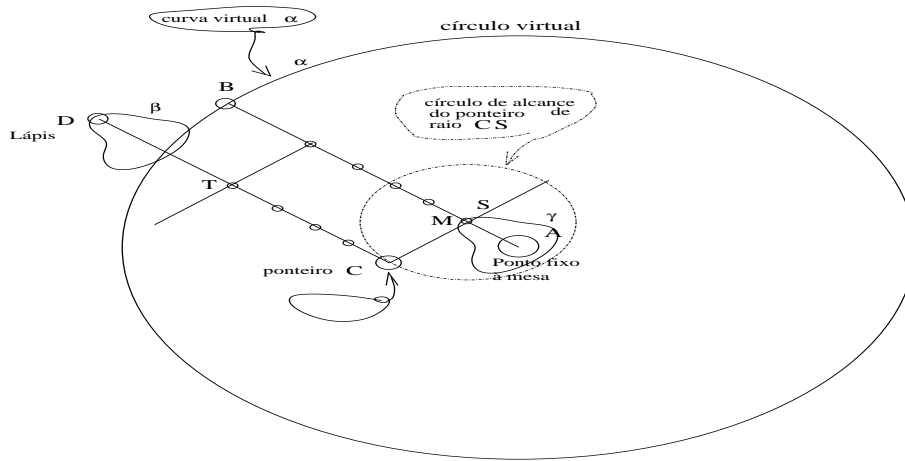


Figura 120: Um instrumento de topologia

A forma algorítmica obter as equações para estas curvas passa por uma aproximação:

- considere uma  $\epsilon$ -rede sobre a região onde opera o *pantógrafo* e portanto uma  $\epsilon$ -rede induzida sobre  $\gamma$  e sobre o intervalo  $[0, 1)$ ;
- defina aproximações polinomiais associadas a estas  $\epsilon$ -redes para obter uma aproximação para  $\alpha, \gamma$  e também para  $\beta$ .

A leitora pode argumentar que *descrevi o método* mas não o apresentei concretamente. Entretanto, vou *resolver concretamente* este problema dentro do programa anunciado como um dos objetivos deste artigo. Na verdade, o artigo representa o planejamento do programa.

Ao mesmo tempo o lápis em  $D$  estabelece uma correspondência biunívoca e contínua entre cada ponto de  $\gamma$  e a curva  $\beta$ , imagem de  $\gamma$ , de modo que podemos pensar no pantógrafo como uma transformação linear,  $T$ , do plano nele mesmo levando  $\gamma$  em  $\beta$ . A transformação  $\gamma \mapsto \alpha$  não é linear quem prova isto é a derivada da transformação.

Traduzindo em equações:

$$\alpha(t) = (\cos(2\pi f_1(t)), \sin(2\pi f_2(t))); t \in [0, 1]; \quad (947)$$

$$\alpha'(t) \left( -\sin(2\pi f_1(t))f_1', \cos(2\pi f_2(t))f_2' \right); \quad (948)$$

$$\gamma(t) = (\gamma_1(t), \gamma_2(t)); t \in [0, 1]; \quad (949)$$

$$\beta(t) = P(t) = T(\gamma(t)); \quad (950)$$

$$\beta(t) = \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix} \begin{pmatrix} \gamma_1(t) \\ \gamma_2(t) \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix} = P(t); \quad \beta'(t) = T'(\gamma(t))\gamma'(t) = T(\gamma(t))\gamma'(t) \quad (951)$$

descrevendo a segunda curva,  $\beta$ , que é a transformada de  $\gamma$ .

O número real  $\lambda$  é o coeficiente de dilatação entre a imagem  $\gamma$  e sua transformada  $\beta$  que pode ser um número real qualquer positivo. Quando os lados do paralelograma forem iguais  $\lambda = 1$  e neste caso o pantógrafo simplesmente copia a figura.

**Observação 7** ( $\alpha$ ) *pode não ser um círculo*

*Da derivada na equação (eq. 948) se vê que  $\alpha$  pode não ser um círculo uma vez que a derivada dum círculo é novamente um círculo apenas com um retardamento de  $\frac{\pi}{4}$  que irá garantir que a tangente seja perpendicular ao raio de tal forma que na segunda derivada, também um círculo a aceleração irá atuar na direção do raio representando a força centrípeta.*

Confira a figura (fig 120), página 312, em que representei estes conceitos graficamente.

Se considerarmos todas estas aplicações diferenciáveis, uma hipótese de trabalho, então

$$P(t) = T(\gamma(t)) + \begin{pmatrix} a \\ b \end{pmatrix}; \quad (952)$$

$$P'(t) = T'(\gamma(t))\gamma'(t) = T(\gamma(t))\gamma'(t); \quad (953)$$

porque a derivada de  $T$  é ela mesma<sup>56</sup>.

Agora, calculando o comprimento de  $\gamma \mapsto G \in \mathbf{R}$  podemos transformar a equação (eq. 952) em

$$t(\cos(2\pi f_1(t)), \sin(2\pi f_2(t))) = \alpha(t); \quad (954)$$

$$(\gamma_1(t), \gamma_2(t)) = \gamma(t); \quad (955)$$

com que obtivemos uma parametrização pelo comprimento de arco de  $\gamma$  fazendo com que  $\|\gamma'\| = 1$  na equação (eq. 955).

<sup>56</sup>As aplicações lineares tem uma derivada constante que coincide com elas mesmas, o que leva a dizer-se, no Cálculo univariado, que as funções do primeiro grau tem por derivada uma constante que é a matriz das funções lineares univariadas.

Esta é uma forma intuitiva de mostrar porque a derivada tem módulo 1 quando a parametrização for o *comprimento de arco*: na prática a curva passa a ser vista como se estivesse projetada sobre  $\mathbf{S}^1$ .

- **permutação** é um dos tipos de arranjo de  $n$  elementos tomados  $n$  a  $n$ , confira *arranjo*, que é um assunto da *Análise Combinatória*.

As permutações são um caso particular dos arranjos e a notação é a seguinte:

$$P_p = A_p^p; \quad (956)$$

- $P_0 = 1 = 0!$ ;
- $P_1 = 1 = 1!$ ;
- $P_2 = 2 = 2!$ ;
- $P_3 = 6 = 3!$ ;

Se você tiver  $n$  elementos,

- o primeiro pode ser escolhido dentre  $n = A_n^1$ ;
- mas para escolher o segundo, sem repetição, você já tem apenas  $n - 1$ :  
 $A_n^2 = n(n - 1)$ ;
- mas para escolher o terceiro, sem repetição, você já tem apenas  $n - 2$  possibilidades:  $A_n^3 = n(n - 1)(n - 2) = \frac{n!}{(n-3)!}$ ;
- e sucessivamente, para escolher o  $n$ -ésimo, so lhe resta uma escolha:  $A_n^n = P_n = n(n - 1)(n - 2) \dots 1 = n! = \frac{n!}{0!}$
- $P_n = n!$

Confira também *arranjo*, *arranjo com repetição*, *combinação*.

- **permutação** uma outra forma de entender, permutação é que ela é uma função injetiva (e conseqüentemente também bijetiva) dum conjunto com  $n$  elementos nele mesmo. Nesta forma de ver, *permutação* é um assunto da Álgebra, mais exatamente da *teoria dos grupos*.

Como a composição de duas tais funções é novamente uma permutação, então a operação “composição de permutações” é uma operação interna do *conjunto de todas as permutações de  $n$  elementos*. Como a composição de funções é associativa e a função identidade é também uma permutação, que deixa todos os elementos fixos, e finalmente, como toda função bijetiva tem inversa, então o conjunto de todas as permutações de um conjunto com  $n$  elementos,  $sim(n)$ , com a composição de funções é um grupo, não comutativo quando  $n \geq 3$ . Tais grupos se chamam “grupo de permutações” e são identificados pelo símbolo,  $sim(n)$ .

O que é importante nos grupos  $sim(n)$  é que *eles contém todos os grupos finitos*, no sentido de que *qualquer grupo finito aparece como subgrupo de algum*

grupo  $sim(n)$ , e a justificativa é simples e ao mesmo tempo um importante método na teoria dos grupos. Numa tabela qualquer de grupo, todas as linhas, ou colunas, são *arranjos simples* dos  $n$  elementos do grupo: permutações dos elementos do grupo. Vou usar isto logo no grupo das matrizes usando um elemento de  $sim(n)$  como a permutação de linhas ou colunas numa matriz.

Portanto estudar os grupos  $sim(n)$  significa estudar todos os grupos finitos. Isto justifica a busca de uma notação adequada para descrever as permutações. Entretanto, pelo próprio resultado enunciado acima, você não pode esperar que num verbete de dicionário se possa esgotar este assunto, você terá aqui apenas uma visão preliminar.

Uma notação efetiva para lidar com os elementos do grupo  $sim(n)$  é o ciclo. Considere a função que ponha em correspondência, *ciclicamente* os números 12345678 na equação

$$\left\{ \begin{array}{l} 1 \mapsto 2 \\ 2 \mapsto 3 \\ 3 \mapsto 4 \\ 4 \mapsto 5 \\ 5 \mapsto 6 \\ 6 \mapsto 7 \\ 7 \mapsto 8 \\ 8 \mapsto 1 \end{array} \right. ; \quad (957)$$

$$(12345678); \quad (958)$$

esta correspondência está sintetizada na equação (eq. 958). em que o último elemento, 8, tem como imagem o primeiro 1, fechando o ciclo.

Outro exemplo,  $(1357)(246)$ , o produto de dois ciclos, é a permutação que, circularmente, leva  $1 \mapsto 3 \mapsto 5 \mapsto 7 \mapsto 1$  fechando assim o ciclo  $(1357)$ , e  $2 \mapsto 4 \mapsto 6 \mapsto 2$ , fechando o outro ciclo  $(246)$ . No *produto de ciclos* os domínios são disjuntos:  $\{1, 3, 5, 7\} \cap \{2, 4, 6\} = \emptyset$ , o que torna o produto de ciclos comutativo:  $(1357)(246) = (246)(1357)$  mas damos preferência ao produto  $(1357)(246)$  que é “alfabeticamente” menor do que  $(246)(1357)$ .

Se o conjunto dos elementos permutados for  $\{1, 2, 3, 4, 5, 6, 7, 8\}$  esta notação indica que 8 é um ponto fixo da permutação. Fica óbvio do exemplo que não sabemos qual é o domínio desta permutação que também poderia ser o conjunto  $\{1, 2, 3, 4, 5, 6, 7, 8, 9\}$  e neste caso haveria dois pontos fixos a saber,  $\{8, 9\}$ . Mas ficaria tudo bem determinado se tivesse sido dito

$$(1357)(246) \in sim(9) \quad (959)$$

Ainda considerando este exemplo, a permutação  $(1357)(246)(89)$ , seria o produto de três ciclos e poderia ainda ser representada como

$$(1357)(246)(89) = \begin{pmatrix} 1 & \mapsto & 3 \\ 2 & \mapsto & 4 \\ 3 & \mapsto & 5 \\ 4 & \mapsto & 6 \\ 5 & \mapsto & 7 \\ 6 & \mapsto & 2 \\ 7 & \mapsto & 1 \\ 8 & \mapsto & 9 \\ 9 & \mapsto & 8 \end{pmatrix} \quad (960)$$

A *teoria dos grupos* convive com uma convenção: nos grupos *não comutativos* a operação é chamada de *multiplicação* e vou adotar esta convenção aqui. Então as permutações são decompostas em *produto de ciclos*. Desta forma os ciclos são os fatores de uma permutação e o produto deles é comutativo por que são funções com domínios disjuntos. Esta decomposição é única. Observe que o ciclo poderia ser apresentado com diferentes “notações”, como

$$(1357) = (3571) = (5713) = (7135)$$

então se os elementos permutados percenterem a um conjunto ordenado, é interessante selecionar a “notação” começando com o menor elemento, e assim se eliminam as demais notações.

Na teoria dos grupos finitos existe o conceito de ordem que posso aplicar às permutações. A ordem é a menor potência positiva que *eleva* um elemento do grupo à unidade, lembrando que a notação é multiplicativa. A ordem duma permutação  $\sigma$  que é o menor inteiro positivo tal que  $\sigma^n = I$ . Vou seguir usando exclusivamente a notação multiplicativa<sup>57</sup> como é o hábito fazer-se com permutações. Suponha que a fatoração de uma permutação em ciclos seja

$$\sigma = \sigma_1 \dots \sigma_p$$

como os produtos comutam então

$$\sigma^n = \sigma_1^n \dots \sigma_p^n = I$$

sendo igual a identidade então todos os fatores são iguais a  $I$  quer dizer que  $n$  é múltiplo da ordem de todos os fatores. Então o menor múltiplo comum serve como valor para  $n$ , considerando que numa permutação os seus fatores sejam os seus ciclos, demonstrei

**Teorema 17 (Ordem de uma permutação)** *ordem dos ciclos de uma permutação*

*A ordem de uma permutação é o menor múltiplo comum da ordem de seus ciclos.*

<sup>57</sup>Nos grupos aditivos: substituindo potência por “multiplicação por  $n$ ”  $n\sigma = \sigma + \dots + \sigma$  com  $n$  parcelas iguais.

Calculando as sucessivas potências dum ciclo, por exemplo, as potências de  $(123456) \in \text{sim}(9)$  tenho

$$(123456)(123456) = (351)(246) = (123456)^2; \quad (961)$$

$$((351)(246))(123456) = (14)(25)(36) = (123456)^3; \quad (962)$$

$$((14)(25)(36))(123456) = (153)(264) = (123456)^4; \quad (963)$$

$$((153)(264))(123456) = (654321) = (123456)^5; \quad (964)$$

$$(654321)(123456) = (1)(2)(3)(4)(5)(6) = I = (123456)^6; \quad (965)$$

$$(966)$$

não sendo nenhuma coincidência que  $6 = o((123456)) = \text{length}((123456))$ . Quando você testar as contas na equação (eq. 961) e seguintes, observe a ordem de aplicação das funções. Por exemplo

$$((351)(246))(123456) : x \mapsto ((351)(246))((123456)(x))$$

primeiro aplico a permutação  $(123456)$  à cuja imagem aplico a permutação  $((351)(246))$ . Observe ainda mais que o cálculo na equação (eq. 962) poderia ter sido feito assim:

$$((351)(246))(123456) = (351)((246))(123456)$$

pela associatividade, uma vez que a segunda potência de  $(123456)$  é um produto de ciclos.

Outro conceito importante das permutações é a *paridade*, ou *sinal*, quero mostrar que é possível estabelecer uma sobrejeção entre  $\text{sim}(n)$  e o grupo  $Z_2 = \{0, 1\}$  dos restos na divisão por 2, algumas vezes a paridade é referida como  $(-1)^p; p \in \{0, 1\}$  e neste caso uma sobrejeção com o *grupo multiplicativo binário*  $\{-1, 1\} \approx Z_2$ .

Esta forma de entender é útil na teoria dos determinantes, e neste caso a paridade é chamada de *sinal*. Como é uma sobrejeção haverá duas classes de permutações, aquelas com *sinal*  $-1$  e as demais com *sinal*  $1$ , respectivamente, as ímpares e as pares.

Para demonstrá-lo vou usar dois métodos. Primeiro vou introduzir um conceito: vou “*representar*” um grupo de permutações como um conjunto de matrizes, na verdade um grupo de matrizes, a *representação* é um isomorfismo de grupos.

Com isto vou mostrar que posso associar, facilmente, a cada permutação um dos números  $1, -1$ , usando o valor dum determinante. Com este conceito vou conseguir mostrar que existe uma *operação elementar* sobre as permutações que são as *transposições* ligadas às operações elementares com matrizes o que me dará, ao final, o outro método.

Além disto a *representação* matricial de permutações é uma forma computacional de calcular com permutações, vou mostrar um exemplo desta técnica.

Deixe-me considerar  $\text{sim}(n)$  que é o grupo das permutações de  $n$  objetos e vou então considerar as  $n$  colunas da matriz identidade,  $I = (a_{ij})$ , de ordem  $n$ , cujo determinante vale 1. Se  $\sigma \in \text{sim}(n)$  seja  $\sigma(I) = (a_{i\sigma(i)})$  a matriz obtida aplicando-se  $\sigma$  às colunas de  $I$ . Então, propriedade dos determinantes,  $\det(\sigma(I)) \in \{-1, 1\}$ , porque houve uma permutação das colunas dum determinante cujo valor inicial era 1, apenas uma *eventual* troca do sinal do determinante a cuja matriz foi aplicada a *mais elementar* das *operações elementares*: uma sucessão de troca de colunas.

**Definição 18 (permutação:)** *paridade de uma permutação*

Seja  $\sigma \in \text{sim}(n)$  e  $I$  a matriz identidade de ordem  $n$ . Definindo  $\sigma(I) = (a_{i\sigma(i)})$  a matriz obtida de  $I$  permutando suas colunas com  $\sigma$ . A paridade de  $\sigma \in \text{sim}(n)$  é o sinal de  $\det(\sigma(I))$ .

Como as permutações são bijeções, então com esta representação obtive um conjunto  $(a_{i\sigma(i)})_{\sigma \in \text{sim}(n)}$  que tem a mesma cardinalidade de  $\text{sim}(n)$ . É interessante explorar até que ponto estes dois conjuntos são semelhantes.

Seria razoável designar o novo conjunto com a notação  $\text{sim}(I)$ , observe que ele vem de  $\text{sim}(n)$  e que  $I$  é uma matriz  $n \times n$ , portanto está tudo muito bem definido, é o conjunto das permutações das  $n$  colunas da matriz  $I$ .

Agora posso dizer que  $I \in \text{sim}(I)$  e a identidade é o elemento neutro do produto de matrizes sendo imagem do elemento neutro  $e \in \text{sim}(n)$ . Se eu considerar um par de inversos em  $\text{sim}(n) \ni \sigma, \tau$  então

$$\sigma(I)\tau(I) = I = \tau(I)\sigma(I) \quad (967)$$

porque estou aplicando o “arranjo inverso” às colunas de  $I$ , em linguagem técnica, aplique a *operação elementar inversa*.

Aqui há um pequeno truque que é preciso esclarecer,  $\sigma(I)$ , aplicada a uma base do espaço  $\mathbf{R}^n$  arranja os elementos da base do espaço<sup>58</sup> de acordo com a permutação  $\sigma$  considerando a base do espaço numa certa ordem:

$$e_1, \dots, e_n \quad (968)$$

e sempre podemos definir uma ordenação dum conjunto finito, numa aplicação reiterada do *axioma da escolha*<sup>59</sup>.

Ou seja  $\sigma(I)$  é a matriz identidade relativamente à base rearranjada

$$e_{\sigma(1)}, \dots, e_{\sigma(n)} \quad (969)$$

ou ainda  $e_{\sigma(k)}$  são os auto-vetores de  $\sigma(I)$ , e os autovalores serão  $-1, 1$ .

As propriedades acima descritas sugerem que  $\text{sim}(I)$  pode ser um grupo isomorfo a  $\text{sim}(n)$ , apenas precisamos mostrar que o produto de matrizes é compatível com a composição de permutações:

$$(\sigma\tau)(I) \equiv \sigma(I)\tau(I) \quad (970)$$

<sup>58</sup>E aqui tem outra ideia escondida,  $I$ , como qualquer matriz  $n \times n$  é uma função linear de  $\mathbf{R}^n$  e se o determinante for diferente de zero é uma bijeção de  $\mathbf{R}^n$ , quer dizer, uma permutação de  $\mathbf{R}^n$  em particular uma permutação de qualquer base do  $\mathbf{R}^n$ .

<sup>59</sup>Ou lema de Zorn



são equivalentes, em que à direita temos um produto de matrizes e à esquerda temos a imagem de  $I$  por uma composição de permutações. Mas o produto de matrizes é uma composição de funções lineares cujas matrizes são elementos do conjunto  $\text{sim}(I)$  são funções lineares bijetivas, e são elementos do grupo das permutações de um conjunto de  $n$  elementos, as colunas da matriz  $I$ , que identificamos como  $\text{sim}(I)$ .

Então  $\text{sim}(I)$  é o grupo de permutação de  $n$  elementos isomorfo a  $\text{sim}(n)$ , que dizemos ser uma representação matricial do grupo  $\text{sim}(n)$  quando  $I$  for a matriz identidade de ordem  $n$ .

Tecnicamente esta construção mostra uma forma computacional de operar com permutações porque existem múltiplos programas de *álgebra linear computacional*, como `octave`, `scilab`. Mas infelizmente o tempo de processamento cresce muito mais do que  $n$ , pelo menos como  $n^2$ , ou talvez  $n^4$ .

O próximo exemplo mostra como construí a imagem de  $\text{sim}(3)$  como permutações das colunas da matriz identidade de ordem  $3 \times 3$ .

**Exemplo 7** ( $\text{sim}(3)$ ) ( $\text{sim}(3)$ )

*Use octave ou scilab, são ambos de domínio público e defina*

```
I = [1,0,0;0,1,0;0,0,1];
I_132=[0,0,1;1,0,0;0,1,0];
I_123=[0,1,0;0,0,1;1,0,0];
I_12=[0,1,0;1,0,0;0,0,1];
I_23=[1,0,0;0,0,1;0,1,0];
I_13=[0,0,1;0,1,0;1,0,0];
```

*então, quando você der “enter” depois de cada uma destas entradas, irá ver as matrizes, no terminal do scilab, como na sequência de equações abaixo.*

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}; I_{132} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}; I_{123} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}; \quad (971)$$

$$I_{12} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}; I_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}; I_{13} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}; \quad (972)$$

*Verifique que*

$o$	$I$	$I_{123}$	$I_{132}$	$I_{12}$	$I_{13}$	$I_{23}$
$I$	$I$	$I_{123}$	$I_{132}$	$I_{12}$	$I_{13}$	$I_{23}$
$I_{123}$	$I_{123}$	$I_{132}$	$I$	$I_{13}$	$I_{23}$	$I_{12}$
$I_{132}$	$I_{132}$	$I$	$I_{123}$	$I_{23}$	$I_{12}$	$I_{13}$
$I_{12}$	$I_{12}$	$I_{23}$	$I_{13}$	$I$	$I_{132}$	$I_{132}$
$I_{13}$	$I_{13}$	$I_{12}$	$I_{23}$	$I_{123}$	$I$	$I_{132}$
$I_{23}$	$I_{23}$	$I_{13}$	$I_{12}$	$I_{132}$	$I_{123}$	$I$

*Entre as operações elementares com matrizes as permutações de linhas ou colunas, são as mais simples. Aqui trabalhei apenas com permutações de colunas*

criando uma representação matricial do grupo das permutações  $sim(3)$  como imagem de subgrupo multiplicativo das matrizes com determinante diferente de zero, o das matrizes unitárias.

Calculando com *scilab* tenho:

```
det(I)   = det(I_132) = det(I_123) = 1;
sinal(I) = sinal(132) = sinal(123) = 1;
det(I_23) = det(I_13) = det(I_12) = -1;
sinal(23) = sinal(13) = sinal(12) = -1;
```

Retornando de  $sim(I)$  para  $sim(n)$  existe uma operação elementar com permutações chamada *transposição* designada por  $(ij)$  na notação de ciclos, é a permutação que intercambia os dois elementos  $i, j$ . Todo ciclo pode ser fatorado em transposições: consequência de que a troca de duas colunas é uma operação elementar com matrizes é uma transposição de colunas como elementos de um conjunto, uma *permutação elementar*.

- Posso ver o que acontece com um ciclo “abstrato”  $(a_1 \dots a_m)$  em que  $m \leq n$  no grupo  $sim(n)$  das permutações de  $n$  elementos.
- Mas sempre o ciclo pode ser escrito com começando com o menor elemento da ordem alfabética dos elementos do conjunto  $\{a_1, \dots, a_n\}$  mostrando que podemos reduzir tudo ao que acontece com o conjunto “menos abstrato”  $\{1, 2, \dots, n\}$ .
- O que diferencia os ciclos são as possíveis permutações dos elementos maiores do que o primeiro.
- Conseqüentemente basta-me analisar qual é o sinal de  $(12 \dots m)$ ;  $m \leq n$  que é a permutação que leva a coluna 1 para o lugar da coluna 2 e finalmente a coluna  $m$  para a posição da coluna 1 isto quer dizer que foi feita a seguinte composição de operações elementares:

$$(12 \dots m) = (1m)(1m-1) \dots (12) \quad (973)$$

em que cada parêntesis, à direita da igualdade, se refere a uma “transposição” de colunas da matriz. O sinal deste determinante será  $(-1)^{(m-1)}$  em que  $m$  é a ordem do ciclo que é o tamanho da sua *cadeia de caracteres* que é  $m$ .

O exemplo 7 mostra que as permutações de ordem 3 junto com a identidade, tem sinal positivo, e que as de ordem 2 tem sinal negativo.

**Exemplo 8** ( $sim(9)$ )  $sim(9)$

```
I=[1,0,0,0,0,0,0,0,0;
   0,1,0,0,0,0,0,0,0;
   0,0,1,0,0,0,0,0,0;
```

0,0,0,1,0,0,0,0,0;  
 0,0,0,0,1,0,0,0,0;  
 0,0,0,0,0,1,0,0,0;  
 0,0,0,0,0,0,1,0,0;  
 0,0,0,0,0,0,0,1,0;  
 0,0,0,0,0,0,0,0,1]

I\_123456=[0,0,0,0,0,1,0,0,0;  
 1,0,0,0,0,0,0,0,0;  
 0,1,0,0,0,0,0,0,0;  
 0,0,1,0,0,0,0,0,0;  
 0,0,0,1,0,0,0,0,0;  
 0,0,0,0,1,0,0,0,0;  
 0,0,0,0,0,0,0,1,0,0;  
 0,0,0,0,0,0,0,0,1,0;  
 0,0,0,0,0,0,0,0,0,1]

I\_123456\*I\_123456=I\_351\_246

I\_351\_246\*I\_123456=

*Este é um grupo grande e não poderia aparecer numa página de texto porque seriam necessárias  $9! = 362880$  colunas na tabela. Mas posso analisar alguns de seus elementos.*

$$(1357) = (17)(15)(13) \in \text{sim}(9); \mapsto -1 = (-1)^3 \quad (974)$$

$$(246) = (26)(24) \in \text{sim}(9); \mapsto 1 = (-1)^2 \quad (975)$$

$$(1357)(246)(89) = (89)(26)(24)(17)(15)(13) \in \text{sim}(9); \mapsto 1 \quad (976)$$

$$\begin{cases} (1357) = (17)(15)(13) \mapsto -1; \\ (246) = (26)(24) \mapsto 1 \\ (89) \mapsto -1; \end{cases} \quad (977)$$

$$(1357)(246)(89) \mapsto 1 = (-1)^3 * (-1)^2 * (-1)^1; \quad (978)$$

$$(1357)(246)(89) \text{ é uma permutação par } \mapsto 1; \quad (979)$$

$$(135)(246)(89) = (15)(13)((26)24)(89) \mapsto -1; \quad (980)$$

$$(14569) = (19)(16)(15)(14) \mapsto 1 = (-1)^4; \quad (981)$$

$$(123456789) = (18)(17)(16)(15)(14)(13)(12) \mapsto 1 = (-1)^8; \quad (982)$$

Observe que

$$(12)(13)(14)(15)(16)(17)(18) \neq (18)(17)(16)(15)(14)(13)(12)$$

*e depende da convenção sobre o significado de fog na composição de funções, nas equações (974)- usei a convenção  $\text{fog}(x) = f(g(x))$*

Sendo o *segundo método* no cálculo da paridade duma permutação a análise do comprimento dos seus ciclos em sua única decomposição em ciclos:

Posso agora *redefinir* a paridade duma partição:

**Definição 19 (permutação:)** *paridade de uma permutação*

A paridade, o sinal de  $\sigma \in \text{sim}(n)$  é o número  $(-1)^p \in \{-1, 1\}$  em que  $p$  é o número de transposições de sua fatoração e é igual ao produto das potências de  $-1$  a soma dos comprimentos dos ciclos cada um reduzido de uma unidade.

---

- **Pitágoras, teorema de** Se  $a, b, c$  forem, nesta ordem, os dois catetos e a hipotenusa de um triângulo retângulo, então

$$a^2 + b^2 = c^2 \quad (983)$$

Existe uma demonstração bem simples usando comparação de áreas que pode ser vista na figura (121), página 323,

A demonstração habitual usa relações métricas no triângulo.

O matemático (e advogado) Fermat, que viveu no século 15 deve ter se inspirado no teorema de Pitágora para enunciar o famoso teorema denominado “último teorema de Fermat” hoje se sugere que seja chamado Fermat-Wiles, associando-o ao nome do matemático inglês, Andrew Wiles, que entre 1993 e 1995 batalhou na versão final de uma demonstração que envolve praticamente toda a *teoria dos números* ou da *álgebra modular*. Embora a demonstração seja muito complicada, o teorema tem uma expressão muito simples. A equação (983) pode ser expressa como

$$a^n + b^n = c^n; a, b, c, n \in \mathbf{Z} \quad (984)$$

e a afirmação de Fermat foi a de que esta expressão somente tinha solução com os números pitagóricos e apenas no caso em que  $n = 2$ . Os números pitagóricos

$$\{(3, 4, 5), (6, 8, 10), \dots, (5, 12, 13), \dots\};$$

---

- **plano**. Confirma *equação do plano*.

---

- **plano cartesiano** confirma *eixos coordenados*

---

- **plano coordenado** também chamado *eixos coordenados*. é uma representação do plano numérico da Geometria Analítica. Confirma também *equação do plano*.

---

- **plano tangente** é uma *variedade linear* de dimensão dois que é *tangente* a uma *superfície*, uma *variedade* de dimensão dois. Estende o conceito de *reta tangente*.

É um conceito do *Cálculo Multivariado* e da *Geometria Diferencial*.

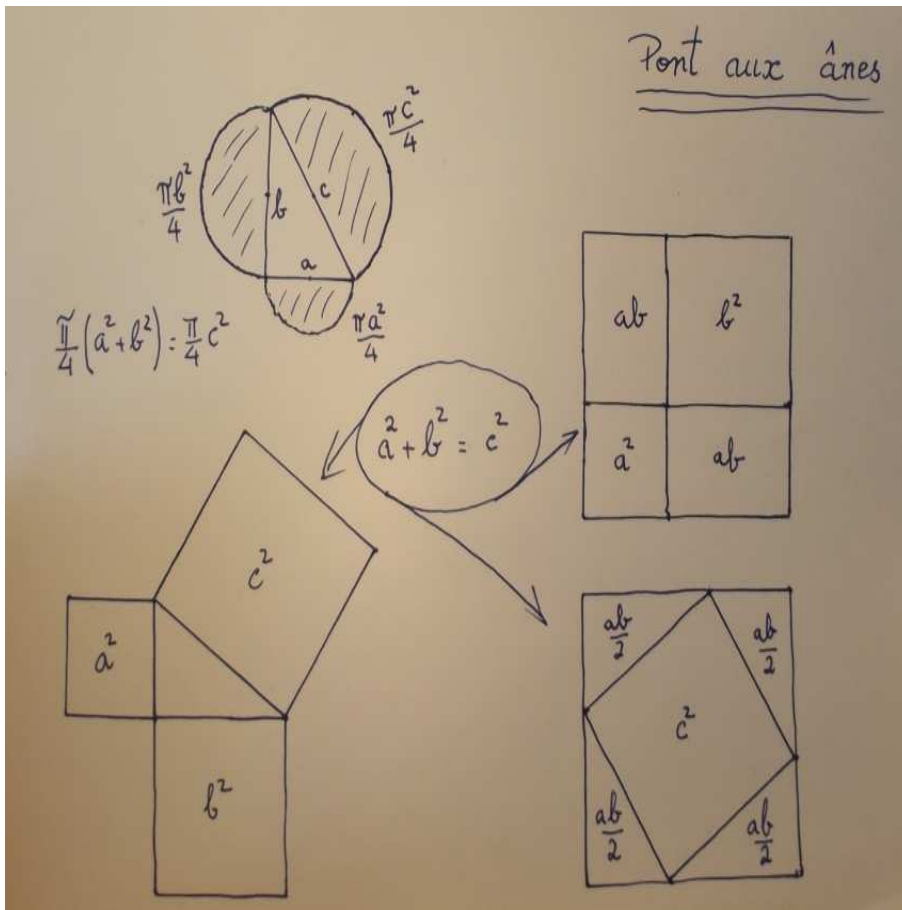


Figura 121: teorema de Pitágoras

Para o Cálculo Multivariado (bivariado) se considera uma função  $z = F(x, y)$  que seja diferenciável e um ponto  $\mathbf{P} = (a, b, c); c = F(a, b)$  então um plano de equação

$$P(x, y) = c + A(x - a) + B(y - b); \quad (985)$$

$$A = \frac{\partial F}{\partial x}|_{(a,b)}; B = \frac{\partial F}{\partial y}|_{(a,b)}; c = F(a, b); \quad (986)$$

$$P(x, y) = F(a, b) + \frac{\partial F}{\partial x}|_{(a,b)}(x - a) + \frac{\partial F}{\partial y}|_{(a,b)}(y - b); \quad (987)$$

é o plano tangente ao gráfico de  $z = F(x, y)$  no ponto  $\mathbf{P}$ .

O plano tangente,  $T_{\mathbf{P}}$ , na notação da *geometria diferencial*, é uma variedade de dimensão 2 obtida como *tangente* no ponto  $\mathbf{P} \in \mathcal{S}$  em que  $\mathcal{S}$  é uma variedade de dimensão 2. Na descrição feita para a função diferenciável  $z = F(x, y)$ , a variedade  $\mathcal{S}$  é o *graf*( $F$ ), o gráfico de  $F$ , e o ponto  $\mathbf{P} = (a, b, c) = (a, b, F(a, b))$  é o ponto de tangência.

Aqui há três conceitos associados:

1. O conceito *variedade* surgiu para suprir o defeito linguístico de nossa *prisão tridimensional* que culturalmente nos prende ao mundo tridimensional em que vivemos e a própria *geometria euclidiana* sofre deste mesmo defeito reconhecendo a geometria tridimensional como o *universo geométrico*. Como consequência desta *prisão tridimensional* a nossa linguagem não comporta os conceitos de uma geometria de dimensão maior do que três. Simplificando, posso dizer que uma variedade é um objeto geométrico, e depois qualificativos, ou *predicativos*, como *linear*, *dimensão* completam a ideia.

Confira *variedade*.

2. *tangência* é uma relação de equivalência entre figuras geométricas, *variedades*, criada pelo *Cálculo Diferencial e Integral* e depois generalizada em outras disciplinas da Matemática como a *Geometria Diferencial*. O Cálculo define esta relação de equivalência estabelecendo que se duas variedades  $\Omega, W$

- tiverem um ponto  $\mathbf{P}$  em comum,  $\mathbf{P} \in \Omega \cup W$ ;
- as derivadas de  $\Omega$  e de  $W$ , em  $\mathbf{P}$ , forem iguais;

então  $\Omega$  é tangente a  $W$  e reciprocamente. Foi esta formulação que usei acima na definição de plano tangente ao gráfico da função  $z = F(x, y)$  no ponto  $\mathbf{P} = (a, b, c) = (a, b, F(a, b))$ .

É uma relação de equivalência valendo as três propriedades para este tipo de relação. Assim se três variedades se encontrarem num ponto  $\mathbf{P}$  e neste ponto tiverem a mesma derivada, então serão tangentes duas a duas e esta relação é uma relação de equivalência.

3. *derivada* Obviamente que esta teoria está incompleta uma vez que lancei mão de mais um conceito: *derivada*. Para chegar a este conceito, o Cálculo Diferencial e Integral precisa de três capítulos, *limite*, *continuidade* e finalmente *derivada*. Vou criar um atalho usando exemplos que podem servir para que você, se inspirando neles, leia os três capítulos dum livro de Cálculo para completar o terceiro conceito deste verbete...

### Derivada

O Cálculo univariado constrói a derivada como um *coeficiente angular instâneo* do gráfico de uma função em um determinado ponto  $\mathbf{P}$ , enquanto que o Cálculo multivariado precisa de *coeficientes angulares instâneos parciais* a serem calculados nas distintas direções do espaço e que se chamam *derivadas parciais*. Vou mostrar-lhe com alguns exemplos que é a ideia de *tangência* que está sendo usada.

- A equação de uma reta, na Geometria Analítica, é uma expressão da forma  $y = m(x - a) + b$ , uma reta com coeficiente angular  $m$  e que passa no ponto  $\mathbf{P} = (a, b)$ . O número  $m$  é a *derivada* da reta.

- Retas são os gráficos das funções do primeiro grau sendo por esta razão que se diz que a *derivada* de uma função polinomial do primeiro grau é constante, porque retas podem ser definidas como as curvas cujo coeficiente angular é constante. Assim, para a função polinomial  $P(x) = b + m(x - a)$  se tem:  $P'(x) = m$ .
- Se uma função univariada  $y = f(x)$  tiver derivada no ponto  $a$  então ela terá uma *reta tangente* ao seu gráfico no ponto  $\mathbf{P} = (a, f(a))$  e como a equação desta reta será  $y = m(x - a) + b$  então:
  - $b = f(a)$ ;
  - $f'(a) = m$
  - $P(x) = f(a) + f'(a)(x - a)$

Esta forma de apresentar a *derivada* cria um método geométrico para o cálculo de funções univariadas: desenhe uma reta tangente ao gráfico da função, no ponto  $(a, f(a))$ , se for possível, calcule o coeficiente angular da reta, para assim obter  $f'(a)$ . Isto não é possível fazer com a função  $f(x) = |x|$  no ponto  $x = 0$  e então esta função não tem derivada no ponto  $x = 0$  porque não é possível traçar-se uma reta tangente no ponto  $(0, f(0))$ .

- Na parte final do capítulo sobre *derivada* do livro de Cálculo, você vai encontrar uma metodologia para calcular derivadas chamada *derivação implícita*. Ela vai nos permitir um salto, porque nem toda curva é uma função e isto pode produzir situações do tipo *múltiplas derivadas num ponto*. É o caso dum círculo cuja equação é dada por

$$(x - a)^2 + (y - b)^2 = r^2; \quad (988)$$

$$y = b \pm \sqrt{r^2 - (x - a)^2} = \pm f(x); \quad (989)$$

O Cálculo nos permite deduzir:

- derivada implícita

$$2(x - a)dx + 2(y - b)dy = 0 \equiv (x - a)dx + (y - b)dy = 0; \quad (990)$$

- Se um ponto  $\mathbf{P} = (p, q)$  pertencer ao círculo podem acontecer três situações:
  - \* a reta tangente em  $\mathbf{P}$  ser paralela ao eixo  $OX$  e neste caso diremos que a ela tem coeficiente zero, e há uma outra reta tangente, também com coeficiente angular zero, no ponto *antípoda* do círculo.
  - \* a reta tangente em  $\mathbf{P}$  ser paralela ao eixo  $OY$  e neste caso diremos que a ela não tem coeficiente angular, e também há uma outra reta tangente, também se coeficiente angular zero, no ponto *antípoda* do círculo.
  - \* nenhum dos casos anteriores, então a reta tangente<sup>60</sup> será  $y = m(x - p) + q$  e podemos efetuar um cálculo simbólico usando a derivada implícita para concluir que

<sup>60</sup>Também neste caso há uma reta paralela a esta situada no ponto *antípoda* do círculo.

- a reta é tangente a um semi-círculo que é gráfico duma função  $y = f(x)$ , e a tangente que se encontra no ponto antípoda é tangente, no caso do círculo à função  $y = -f(x)$ . A equação de cada uma das retas tangentes é uma polinômio do primeiro  $y = P(x)$  cujo gráfico é o da reta tangente.
- $\frac{dy}{dx} = m$
- $f'(p) = m = \frac{df}{dx}|_{x=p}$  que é a controversa, mas perfeitamente funcional notação de Leibniz para a derivada.
- $P(x) = f(a) + f'(a)(x - a)$

O círculo é tangente a retas que se tiverem coeficientes angulares corresponderão às derivadas de funções deduzidas do gráfico do círculo:  $\pm f$ .

- \* Se a função for bivariada,  $z = F(x, y)$  você irá chegar, usando derivação implícita, à expressão do polinômio do primeiro grau em duas variáveis que aparece na equação (eq. 987);

$$dz = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy; \quad (991)$$

$$\mathbf{P} = (a, b, c) = (a, b, F(a, b)); A = \frac{\partial F}{\partial x}|_{(a,b)}; B = \frac{\partial F}{\partial y}|_{(a,b)}; \quad (992)$$

$$dz := z - c; dx := x - a; dy := y - b; \quad (993)$$

$$z = P(x, y) = c + A(x - a) + B(y - b); \quad (994)$$

$$z - c = P(x, y) - F(a, b); \quad (995)$$

$$P(x, y) = F(a, b) + \frac{\partial F}{\partial x}|_{(a,b)}(x - a) + \frac{\partial F}{\partial y}|_{(a,b)}(y - b) \quad (996)$$

Na equação (eq.993) eu usei o símbolo “:=” para indicar que está sendo feita uma *modelagem*, não é um cálculo, mas sabemos que se fizermos esta substituição vamos obter a equação certa...

Observe que no sistema de equações (eq.991) - (eq.996) existem duas funções cujos gráficos são tangentes, no ponto  $\mathbf{P} = (a, b, c)$ , a função linear bivariada  $z = P(x, y)$  e a função diferenciável  $z = F(x, y)$ . Ter uma função linear tangente é o que caracteriza a *diferenciabilidade* duma função em qualquer dimensão. E o gráfico da função linear tangente é a variedade linear tangente que no caso de funções bivariadas é *plano tangente*.

---

- **polar** confira *coordenadas polares*.

---

- **poligonal** é uma curva formada por segmentos de reta. Desta forma uma poligonal é apenas um tipo particular de curva caracterizada por uma definição por pedaços em que cada pedaço é um segmento de reta. Uma poligonal é também um 0-splines.



Os elementos duma poligonal plana são (1) a sucessão de números reais,  $(a_i)_{i=0}^n$ , as abcissas dos nós e a (2) sucessão das ordenadas  $(b_i)_{i=0}^n$ , formando a sucessão dos vértices  $((a_i, b_i))_{i=0}^n$ , que determinam os segmentos de reta da poligonal, chamados de *lados* da poligonal.

Algumas vezes os *vértices* são chamados de *nós*, está linguagem é comum em *teoria dos grafos*. Um grafo é formado de varias poligonais se originando em um conjunto de vértices iniciais ou algumas vezes num único vértice. A figura (fig 122), página 327, lhe mostra uma *poligonal* plana.

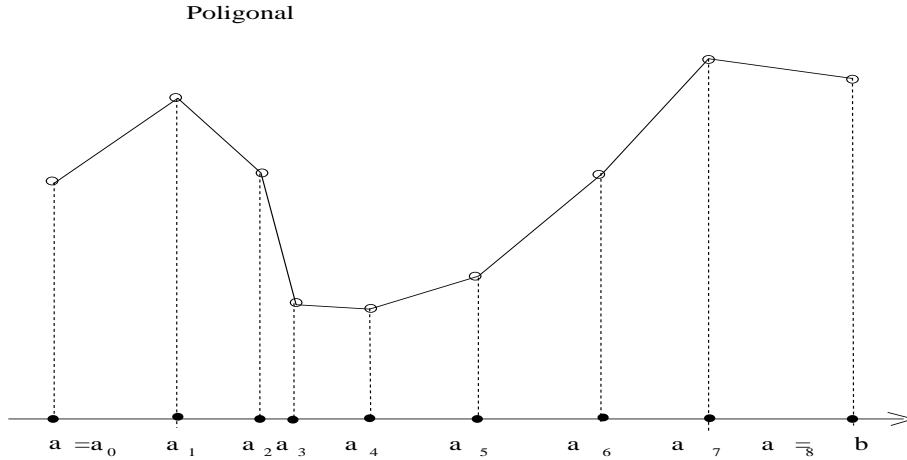


Figura 122: Poligonal

Um polígono é um exemplo de *poligonal fechada*, em que o primeiro vértice coincide com o último,

$$a_0 = a_n; b_0 = b_n; (a_0, b_0) = (a_n, b_n); \quad (997)$$

Como uma poligonal é apenas um tipo de curva, podemos usar muito do se diz sobre curvas para as poligonais e a importância delas está ligada ao caso particular dos polígonos. A discontinuidade da derivada duma poligonal nos vértices produz o conceito de ângulo, então à cada vértice, com exceção dos extremos na poligonal aberta, corresponde um ângulo.

No caso dos polígonos, por serem poligonais fechadas, o número de vértices coincide com o número de lados e com o número de ângulos.

Se a sucessão da ordenadas for crescente, então a poligonal define uma função no domínio  $[a, b] = [a_0, a_n]$ , é o caso da figura (fig 122).

Podemos ter poligonais em espaço de dimensão maior que dois, neste caso a sucessão das abcissas passa ser uma sucessão de vetores de dimensão  $p - 1$  de modo que a poligonal seja uma curva dum espaço de dimensão  $p$  sendo os seus vértices

$$(a_i, b_i)_{i=0}^n; b_i \in \mathbf{R}^{p-1}; (a_i, b_i) \in \mathbf{R}^p; p \geq 1; \quad (998)$$

Obviamente que podemos ter poligonais definidas em espaços de números complexos mas o uso de tais poligonais parece ser muito raro com exceção dos polígonos regulares convexos associados ao círculo unitário de  $\mathbf{C}$  que representam as raízes da unidade.

- um polígono plano é *convexo* se qualquer reta do plano o cortar em apenas dois pontos. A figura (fig 123), página 328,

O caso contrário, quando o polígono é côncavo, alguma reta do plano irá

**há uma reta que corta o polígono em mais de dois pontos**

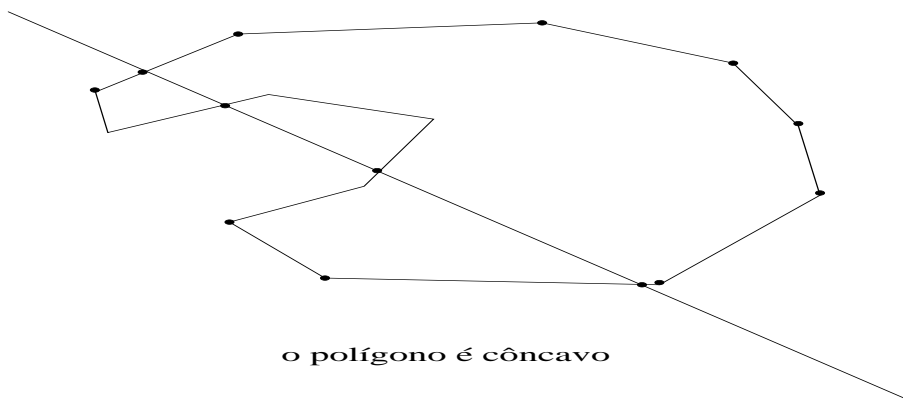


Figura 123: polígono côncavo cortado por uma reta

cortá-lo em mais de dois pontos.

- um *polígono regular* é aquele em que os lados e os ângulos têm todos a mesma medida, e como consequência um polígono regular plano tem os seus vértices em cima de alguma circunferência no plano. Não parece haver interesse de estudar polígonos em espaço de dimensão maior do que dois. A teoria dos polígonos planos é muito grande e envolve bonitas relações entre ângulos e lados. O exemplo mais comum é da soma dos ângulos dum triângulo que vale  $\pi$ . No caso dum polígono qualquer este valor pode ser obtido usando diagonais que subdividam o polígono em triângulos, no caso dum quadrilátero o valor é  $2\pi$ , num pentágono o valor é  $3\pi$ , no caso dum polígono de  $n$  lados o valor é  $(n-2)\pi$ . O número de diagonais corresponde às combinações dos vértices dois a dois,  $\binom{n}{2}$ .
- no conjunto dos números complexos, os polígonos regulares convexos cujos vértices estejam sobre o círculo unitário  $\mathbf{S}^1$  de tal modo que um deles seja o número real 1, formam o conjunto das raízes complexas da unidade tendo por equação

$$(e^{2ki\pi/n})_{k=0}^{n-1} = \sqrt[n]{1} \quad (999)$$

e você pode ver, na equação (eq. 999), a fórmula de Euler. A figura (fig. 124) mostra as raízes quintas da unidade. Cada vértice é um número complexo unitário que elevado a  $n$  reproduz 1. Na figura (fig 124), página 329,

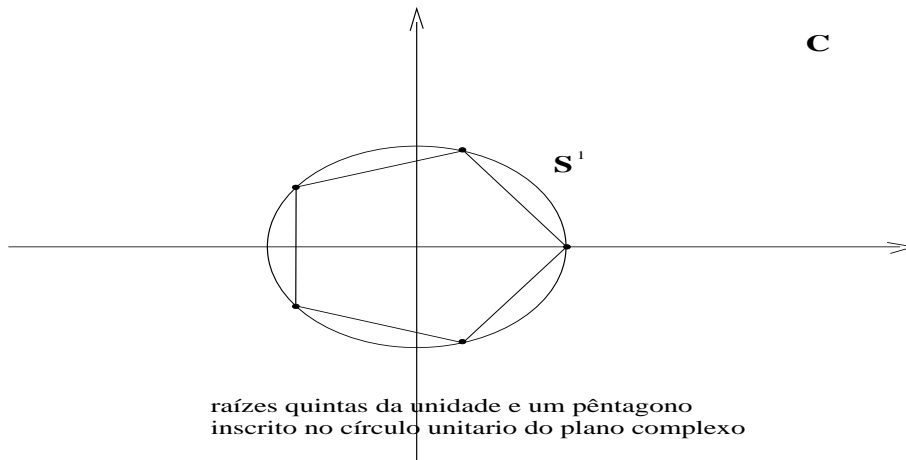


Figura 124: raízes quintas da unidade e um pentágono

- As raízes dum número complexo,  $z$ , qualquer são obtidas usando um círculo de raio  $r = |z|$  com um polígono regular convexo com  $n$  lados inscrito neste círculo tendo um dos vértices situado sobre a reta suporte de  $e^{i\theta/n}$  com  $\arg(z) = \theta$ . Cada vértice é um número complexo que elevado a  $n$  reproduz  $z$ .

- polinômio de Taylor Ver Taylor.

- polinômio trigonométrico é uma descoberta atribuída à Fourier mas na verdade usando os trabalhos dos o que precederam no uso de somas trigonométricas, como Leonhard Euler, Jean le Rond d'Alembert, and Daniel Bernoulli e com o mesmo objetivo, resolver equações diferenciais. Apenas Fourier em sua monografia apresentada à Academia de Ciências da França teve a ousadia de afirmar que *as séries trigonométricas poderiam representar quaisquer funções periódicas*.

Os polinômios trigonométricos, de que vou tratar aqui, são aproximações de funções, ou melhor, projeções de funções definidas num determinado espaço, com imagem num espaço de dimensão finita gerado pelos vetores linearmente independentes  $\sin(kx), \cos(kx)$  com  $k \leq N$ , portanto um espaço vetorial de dimensão  $2N + 1$ .

A leitora pode ficar intrigada com a dimensão  $2N + 1$ , de onde sai este número? Se  $k \in \{0, 1, \dots, N\}$  seria razoável pensar em  $2N + 2$ , entretanto se

$k = 0 \sin(kx) \equiv 0$  e o conjunto deixaria de ser uma base pois a Álgebra Linear nos ensina que zero não pode ] pertencer a um conjunto de vetores linearmente independentes. Então temos que retirar  $\sin(0)$  o nos deixa com o total indicado.

Falar das *séries de Fourier* me obrigaria discutir a convergência destas séries que tornaria um verbete de dicionário muito longo. A leitora interessada pode encontrar este tópico em outro lugar, aqui ficará um resumo do essencial. Aliás, para que fique claro que é um resumo, compare com o livro de Zygmund, *Trigonometric Series* que tem cerca de 750 páginas.

Há várias formas de apresentar os polinômios trigonométricos, a que vou preferir vem sob a forma duma aplicação da Álgebra Linear cujo sumário seria:

- Um espaço vetorial de funções,  $E$ , munido de produto escalar;
- Seleção de vetores ortonormais para estabelecer uma base para um subespaço de  $E$ , chamado aqui de  $\mathcal{F}$ .
- Definição de um projetor de  $E \rightarrow \mathcal{F}$

#### Descobrimo um produto interno

Uma integração por partes mostra que

$$\int_0^{2\pi} \sin(x) \cos(x) dx = 0 \quad (1000)$$

e se fizermos uma abstração podemos escrever

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x) dx \quad (1001)$$

que representa, se  $f, g$  forem duas funções integráveis sobre o intervalo  $[0, 2\pi]$ , uma fórmula que uma breve investigação mostra que tem as propriedades

$$f, g \text{ integráveis e } \alpha, \beta \in \mathbf{R}; \quad (1002)$$

$$\langle \alpha f_1 + \beta f_2, g \rangle = \alpha \langle f_1, g \rangle + \beta \langle f_2, g \rangle; \quad (1003)$$

$$\langle f, f \rangle \geq 0 \text{ e } \langle f, f \rangle = 0 \iff f = 0; \quad (1004)$$

sendo portanto um produto escalar como é definido na Álgebra Linear.

A equação (eq. 1004) não é verdadeira em geral e leva à definição de função quase sempre zero. Vou omitir este detalhe apenas com a observação de que uma discussão mais cuidadosa deste “fenômeno” leva à definição de uma relação de equivalência e à teoria dos espaços de *Lebesgue*, inclusive mesmo à medida de Lebesgue. Mas se  $f, g$  forem funções contínuas a equação (eq. 1004) é verdadeira e vou assumir esta hipótese para mais à frente mostrar que ela pode ser eliminada.

As equações (eq. 1002)- (eq. 1004) são as propriedades do *produto escalar* definido nos espaços vetoriais da Álgebra Linear isto nos permite usar a equação

(eq. 1001) como definição do produto escalar em um determinado espaço de funções que vou aqui definir, usando a hipótese que me permitiu considerar verdadeira a equação (eq. 1004), como sendo o espaço vetorial das funções contínuas definidas no intervalo  $[0, 2\pi]$ .

Posso agora recomeçar a introdução que fiz acima com dados mais concretos. O espaço vetorial  $E$  mencionado acima é  $\mathcal{C}([0, 2\pi])$  no qual a fórmula

$$\langle f, g \rangle = \int_0^{2\pi} f(x)g(x)dx \quad (1005)$$

define um produto escalar. Mas qualquer múltiplo por um número real será também um produto escalar e logo mais vou mostrar que esta não é a melhor fórmula para o produto escalar do espaço  $E$ .

Este espaço é “pequeno” para a teoria de Fourier, tendo sido este ponto a razão de críticas que foram feitas à teoria incompleta de Fourier que somente pode ser entendida completamente um século depois com a integração à Lebesgue, mas também por razões de espaço vou deixar estes detalhes para que a leitora interessada complete. Mais adiante vou deixar alguns dicas de como fazê-lo, apenas no momento uma observação, a continuidade é um requisito exagerado, bastava exigir que “ $f, g$  fossem integráveis” e assim já teríamos um espaço vetorial onde o produto escalar da equação (eq. 1005) estivesse definido.

A leitora pode ver aqui como, aos poucos, os detalhes em uma teoria podem ser acrescentados, e esta história foi sendo construída de 1700 a 1900, ou seja de Euler à Lebesgue e o livro de Zygmund, 1930, é um resumo significativo da teoria.

Estabelecido este cenário podemos calcular a projeção de um vetor  $f \in E$ , um vetor do espaço  $E$ , na direção de um dos vetores da base, por exemplo, a projeção de  $f \in E$  na direção de  $\cos_k(x) = \cos(kx)$  ou na direção de  $\sin_k(x) = \sin(kx)$

$$a_k = \langle f, \cos_k \rangle = \int_0^{2\pi} f(x) \cos_k(x) dx; k > 0; \quad (1006)$$

$$b_k = \langle f, \sin_k \rangle = \int_0^{2\pi} f(x) \sin_k(x) dx; k \geq 0; \quad (1007)$$

Os números,  $a_k, b_k$  são os chamados *coeficientes de Fourier*<sup>61</sup> de  $f$  e da mesma forma como fazemos na Álgebra Linear, ou na Geometria Analítica, quando descobrimos as projeções de um vetor relativamente aos vetores da base do espaço, podemos escrever o vetor como combinação linear dos vetores da

---

<sup>61</sup>Neste momento errados! Quando for corrigido o produto escalar serão apresentados os coeficientes corretos de Fourier.

base:

$$\hat{f}(x) = \sum_{k=0}^N a_k \cos(kx) + b_k \sin(kx) \approx f(x); \quad (1008)$$

$$\hat{f} = \sum_{k=0}^N a_k \cos_k + b_k \sin_k \approx f; \quad (1009)$$

Há uma diferença entre as equações (eq. 1008) e (eq. 1009), uma representa o que acontece em cada ponto e a outra é uma expressão global para os espaços. Para explicarmos o que significa a "aproximação" na equação e (eq. 1009) teríamos que discutir a topologia, ou neste caso mais simples, definir qual é a norma, ou distância usada no espaço  $\mathcal{C}([0, 2\pi])$ .

Novamente aqui ficam reticências para serem preenchidas pela leitora interessada.

### Um gráfico mostrando a aproximação

O próximo gráfico na figura (fig 125), página 332, mostra-lhe um polinômio

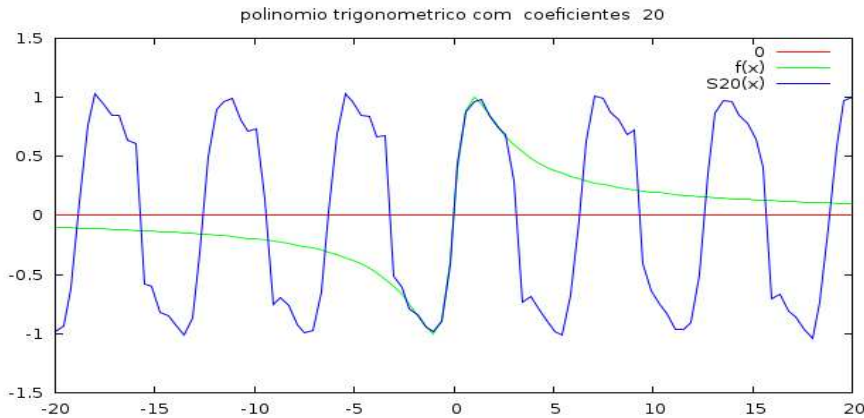


Figura 125: uma sucessão de polinômios trigonométricos

trigonométrico com  $N = 20$  que foi feito com o script do gnuplot, `fourier.gnuplot`, que você pode encontrar na página

<http://www.multivariado.sobralmatematica.org/programas/>

Rodando este programa foi você verá uma sucessão de gráficos de polinômios trigonométricos sendo o último este que aparece na figura (fig 125). Você também pode ler e alterar o programa selecionando outra função  $f$  para calcular os coeficientes de Fourier, o programa é distribuído sob a GPL.

Na figura (fig 125), o gráfico de cor verde é o da função  $f$  e em azul você pode ver um padrão que se repete é gráfico de uma função periódica, de período  $2\pi$  porque é uma combinação linear de senos e cosenos que são

funções  $2\pi$ -periódicas. No intervalo  $[-\pi, \pi]$ <sup>62</sup> o padrão se aproxima do gráfico de  $f$ , como disse Fourier, as *séries trigonométricas podem representar qualquer função periódica*.

Lendo o programa `fourier.gnuplot` você deve observar que não foi o produto escalar da fórmula (eq. 1005) que foi utilizado mas sim

$$\langle f, g \rangle = \frac{1}{\pi} \int_0^{2\pi} f(x)g(x)dx \quad (1010)$$

Experimente retirar este fator, no programa, e verá que não ha superposição dos gráficos, mais exatamente, você irá observar que houve um erro pela soma de uma constante. Ora a constante é o termo  $a_0$  e aqui começa uma das histórias mais intrigantes da Matemática: *uma fórmula com uma exceção com cerca de cem anos de investigação para deixar claro o porque da exceção*, de Fourier a Riesz-Fischer. Mas vou poder mostrar-lhe o porque, em algumas linhas, montado no trabalho que foi feito por centenas de matemáticos.

Vou começar por escrever a fórmula para o cálculo dos coeficientes de Fourier de  $f$  deixando a exceção em relevo:

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x)dx; \quad (1011)$$

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(kx)dx; k > 0; \quad (1012)$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(kx)dx k > 0; \quad (1013)$$

Sem o fator  $\frac{1}{\pi}$  já tínhamos a ortogonalidade de todos os vetores da base com exceção de um, quando  $k = 0$ ,  $\cos(kx)$  que é a função constante 1. Verificando  $\|\cos(0x)\|$  vamos encontrar o valor 2. É um vetor de norma 2 que é perpendicular a todos os outros. Valores experimentais, o resultado dos programas, mostram um erro ao consideramos a definição de  $a_0$  como  $\frac{1}{\pi} \int_0^{2\pi} f(x)dx$ .

Como já não é possível alterar o produto escalar, poderíamos alterar a definição da função  $\cos(0x)$ . A solução é considerá-la definida por:

$$\cos(0x) = \frac{1}{2}.$$

Mas esta definição não é aceitável, a solução é a de redefinir o valor do coeficiente  $a_0$

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x)dx \quad (1014)$$

---

<sup>62</sup>O programa `fourier.gnuplot` usa o intervalo  $[-\pi, \pi]$  esta é a razão de que o padrão que se observa é uma aproximação para o gráfico de  $f$  no intervalo  $[-\pi, \pi]$ .

Deixe-me usar o fato de que a *série de Fourier* é absolutamente convergente, vou poder enunciar dois resultados importantes que se completam, teorema de Riesz-Fischer e a identidade de Parseval, que estabelecem em conjunto que a transformada de Fourier é um isomorfismo entre os espaços  $\mathcal{L}^2([0, 2\pi])$ , o espaço das funções com quadrado integrável, e  $l^2(\mathbf{N})$ , o espaço das sucessões cujo quadrado define uma série convergente. Os vetores  $f \in \mathcal{L}^2([0, 2\pi])$  e  $\hat{f} \in l^2(\mathbf{N})$  têm a mesma norma.

É mais do que um isomorfismo, é uma isometria.

Vou usar este resultado para mostrar que a solução adotada de redefinir  $a_0$  não é tão exdrúxula. Supondo a *convergência absoluta* da série de Fourier podemos intercambiar série e integral para escrever:

$$\int_0^{2\pi} \left( \sum_{k=0}^{\infty} a_k \sin(kx) + b_k \cos(kx) \right) dx = \quad (1015)$$

$$\sum_{k=0}^{\infty} \int_0^{2\pi} (a_k \sin(kx) + b_k \cos(kx)) dx = \quad (1016)$$

$$2\pi a_0 = \int_0^{2\pi} f(x) dx; \quad (1017)$$

onde voltamos a encontrar o valor proposto para a fórmula de  $a_0$

$$a_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) dx; \quad (1018)$$

Foi este o rumo a história tomou.

Para tornar menos estranha esta solução forçada em que o coeficiente  $a_0$  tem uma definição diferente dos demais coeficientes, observe que a definição proposta tem o valor particularmente importante:  $a_0$  é o *valor médio integral* de  $f$  no intervalo  $[0, 2\pi]$ . Quando  $f$  for uma função descontínua, a sua série de Fourier será contínua e corrigirá  $f$  atribuindo-lhe o valor médio no ponto de discontinuidade.

Como já antecipei os teoremas de Riesz-Fischer e a identidade de Parseval, deixe-me antecipar também um resultado envolvendo a convolução e a série de Fourier que também mostra como a redefinição de  $a_0$  é bem apropriada. Podemos mostrar que a série de Fourier é uma convolução com o núcleo de Dirichlet resultando numa função regularizada<sup>63</sup> e contínua em que voltamos a ver o valor médio como a correção da descontinuidade de  $f$ , confira *convolução*.

---

- **Potência** ou potenciação é uma das operações básicas da Aritmética. Confira também *potência dum conjunto*. A Aritmética define a potência como uma multiplicação repetida

$$a^b = \underbrace{a \cdots a}_b; a, b \in \mathbf{N} \quad (1019)$$

---

<sup>63</sup>mais exatamente, de classe  $\mathcal{C}^\infty$ .



e podemos estender a definição para expoente em  $\mathbf{Z}$  mantendo o mesmo sentido de multiplicação repetida

$$a^b = \begin{cases} b > 0 & \underbrace{a \cdots a}_b; \\ b = 0 & 1; \\ b < 0 & \frac{1}{a^b} \end{cases} ; a > 0; \quad (1020)$$

Para  $a = 0$  a definição  $0^0$  é indefinida e para os expoentes negativos não existe porque implicaria no inverso do zero.

Todas as definições nas equações (eq. 1019) e (eq. 1020) funcionam para qualquer número real estritamente positivo  $a$  e dou assim a definição expandida para  $a \in \mathbf{R}^{++}$ . Para números negativos também funciona, levando em conta a regra do sinal para multiplicação:

$$\begin{cases} a < 0; a^b > 0 & \Leftarrow b \text{ par;} \\ a < 0; a^b < 0 & \Leftarrow b \text{ ímpar;} \end{cases} \quad (1021)$$

e como sempre o zero fica excluído como expoente, no caso da base ser zero.

O que é interessante é expandir a definição para o expoente  $b \in \mathbf{Q}$  o que passa pela definição

$$a^{\frac{1}{b}}; b > 0$$

e truque é considerar esta operação como inversa da potenciação:

$$a^{\frac{1}{b}} = \sqrt[b]{a}; (\sqrt[b]{a})^b = a; \quad (1022)$$

observe que posso incluir um número intermediário escrevendo:

$$\begin{cases} a \xrightarrow{\sqrt[b]{\phantom{x}}} c; & a^{\frac{1}{b}} = \sqrt[b]{a} = c; \\ c \xrightarrow{(\phantom{x})^b} a; & c^b = a; \end{cases} \quad (1023)$$

É aqui que surgem algoritmos para o cálculo das raízes quadrada, cúbica, construções muito antigas cuja história pode ser consultada na Wikipédia fazendo a busca com a palavra chave *algoritmo da raiz*. São construções extremamente bonitas e considerando a época em que foram feitas, verdadeira arte de calcular que nós hoje perdemos pelo apoio que a computação nos oferece e portanto, numa simples falta de energia elétrica nos veremos repentinamente sem poder executar alguns cálculos bem elementares.

Mas, logo a raiz quadrado de dois é um número irracional levando o algoritmo da raiz quadrado a um laço infinito... em programação, e com a presença da energia elétrica, isto se resolve estabelecendo um limite de tempo de processamento ou uma precisão adequada.

Com programação podemos pensar  $a^b$ ;  $a, b \in \mathbf{R}$ ;  $a \neq 0$  e simplesmente ignorar a raiz a partir da definição feita na equação (eq. 1023). Se você tiver acesso à linguagem `calc` que é distribuída de maneira livre sob GPL, [15], basta escrever num terminal do `calc`

```
tarcisio@casa:/pub/tarcisio/tex/dicionario$ calc
C-style arbitrary precision calculator (version 2.12.4.4)
Calc is open software. For license details type:  help copyright
[Type "exit" to exit, or "help" for help.]

; a = sqrt(5)
; pi = 4*atan(1);
; c = power(a,pi);
; c
12.52966660804931024735
;
quit;
tarcisio@casa:/pub/tarcisio/tex/dicionario$
```

para calcular  $(\sqrt{5})^\pi$ . O algoritmo está incluído na linguagem, mas é possível escrever um programa simples para obter este resultado, ou mesmo calcular qualquer seja a potência dum número arbitrário estritamente positivo.

Desta forma podemos dizer que a radiciação não precisa ser considerada uma vez que ela é apenas a operação inversa da potência. Ainda usando o exemplo do `calc` esta afirmação fica assim:

$$\sqrt[p]{a} = \text{power}\left(a, \frac{1}{p}\right); \quad (1024)$$

num terminal do `calc`

```
sqrt(5);
2.23606797749978969641
power(5,0.5);
2.23606797749978969641
c = power(5,(1/7));
power(c,7);
4.9999999999999999992
```

observe que o último resultado é um valor aproximado:

$$4.9999999999999999992 \approx 5;$$

Na linguagem `calc` não existe uma função para calcular raízes arbitrárias exatamente porque a potência resolve isto e podemos simplesmente ignorar as raízes assim como os magníficos algoritmos que os antigos construíram, sempre com esperança de que não nos venha a faltar a energia elétrica.

---

- **potência dum conjunto** se um conjunto for finito,  $\mathbf{A} = \{1, 2, \dots, n\}; n \in \mathbf{N}$ , então o conjunto das partes de  $\mathbf{A}$ , tem  $2^n$  elementos e por esta *coincidência* a operação

$$A \mapsto \mathbf{A}; X \subset A \Rightarrow X \in \mathbf{P}(\mathbf{A}) = \mathbf{2}^{\mathbf{A}}; \quad (1025)$$

	0	1	2	3	4	5	6	7	8	$k \mapsto 2^k$
$\{\}$	1									$1 \mapsto 2^0$
$\{1\}$	1	1								$2 \mapsto 2^1$
$\{1, 2\}$	1	2	1							$4 \mapsto 2^2$
$\{1, 2, 3\}$	1	3	3	1						$8 \mapsto 2^3$
$\{1, 2, 3, 4\}$	1	4	6	4	1					$16 \mapsto 2^4$
$\{1, 2, 3, 4, 5\}$	1	5	10	10	5	1				$32 \mapsto 2^5$
$\{1, 2, 3, 4, 5, 6\}$	1	6	15	20	15	6	1			$64 \mapsto 2^6$
$\{1, 2, 3, 4, 5, 6, 7\}$	1	7	21	35	35	21	7	1		$128 \mapsto 2^7$
$\{1, 2, 3, 4, 5, 6, 7, 8\}$	1	8	28	56	70	56	28	8	1	$256 \mapsto 2^8$

é chamada de potência do conjunto  $\mathbf{A}$ .

A última expressão na equação (eq. 1025) é apenas uma notação, não é um *cálculo*, uma vez que envolve um número e um conjunto.

O chamado triângulo de Pascal mostra em suas linhas a distribuição das partes do conjunto  $\mathbf{A}$  em função do número de elementos como mostra a tabela

O conjunto das partes de  $\mathbf{A}$  é um novo conjunto cujos elementos tem uma hierarquia maior do que os elementos de  $\mathbf{A}$  ou maior complexidade sendo assim vale a relação

$$X \subset \mathbf{A} \Rightarrow X \in \mathbf{P}(\mathbf{A}) \quad (1026)$$

Se o conjunto  $\mathbf{A}$  for finito, podemos dizer que  $2^{\mathbf{A}}$  tem mais elementos que  $\mathbf{A}$  e a cardinalidade é simplesmente a quantidade de elementos do conjunto que e o número  $2^n$  é a cardinalidade de  $2^{\mathbf{A}}$  se  $n$  for a cardinalidade de  $\mathbf{A}$ .

### Cardinalidade

Se o conjunto  $\mathbf{A}$  não for finito esta linguagem deixa de ter sentido porque a cardinalidade de conjuntos não finitos deixa de ser um número e passa a ser uma classificação dos conjuntos.

Se o conjunto  $\mathbf{A}$  não for finito, os dois conjuntos  $\mathbf{A}, 2^{\mathbf{A}}$  se encontram em classes de cardinalidade diferentes o que levou a Cantor enunciar o salto de cardinalidade, a ainda chamada *hipótese do contínuo*.

Assim, para o conjunto dos números naturais,  $\mathbf{N}, \mathbf{P}(\mathbf{N})$  estão em classes de cardinalidade diferentes não havendo uma classe de cardinalidade intermediária. A classe de cardinalidade de  $\mathbf{P}(\mathbf{N})$  é a mesma de  $\mathbf{R}$  que é a chamada de *cardinalidade do contínuo* e indicada com a letra  $c$ . É o chamado salto de cardinalidade.

Esta questão teve que espera cem para ser finalmente esclarecida por Paul Cohen completando a tese de Gödel e estabelecendo que não é possível ter em Matemática uma única teoria dos conjuntos e sim pelo menos duas, uma delas é a chamada teoria dos conjuntos de ZFC (Zermelo-Fraenkel).

Estas questões são tratadas na Lógica Matemática.

- primitiva Se uma função  $f$  for integrável no sentido de Riemann é possível

definir uma outra função associada a  $f$  usando a integral como “equação”

$$t \mapsto \int_a^t f(x)dx = F(t) \quad (1027)$$

que é *uma primitiva* de  $f$ . Podemos provar que  $F'(t) = f(t)$  e neste caso também é verdade que, se  $G(t) = F(t) + C$ , em que  $C$  é uma constante, então  $G'(t) = f(t)$  o que mostra que  $f$  tem uma infinidade de primitivas e portanto para que esta operação seja inversa da derivada é preciso selecionar uma constante, por exemplo,  $C = 0$ . A constante  $a$  na expressão da integral, recebe o nome “condição inicial” e esta relacionada com a constante  $C$ .

Este é um primeiro exemplo de *equação diferencial ordinária*, uma primitiva,  $F$  da função  $f$  é uma solução da equação diferencial

$$F' = f; \quad (1028)$$

Quando a função  $f$  for multivariada ela pode (ou não ter uma primitiva) que então será solução de uma *equação diferencial parcial*. Neste caso entra o conceito de integral dependente, ou independente de um caminho e o Teorema de Green tem uma versão trivial que separa estas águas.

---

- **primo** é um número natural positivo, maior do que 2, que não pode ser decomposto como um produto de outros números naturais. 7 é um número primo, assim como 5,3 e 2. Mas, como  $8 = 2^3$  então 8 não é um número primo. O *crivo de Eratóstenes* é a lista dos números primos até um certo “último número primo”, observando que *o conjunto dos números primos é infinito*, é o chamado teorema de Euclides, é um dos teoremas fáceis da Álgebra.

O teorema fundamental da Álgebra afirma que todo número inteiro maior do que 2 pode ser escrito de maneira única como um produto de fatores primos, a menos da ordem como estes fatores apareçam:

$$n \in \mathbf{Z}; n > 2; n = q_1^{\alpha_1} \cdots q_r^{\alpha_r} \quad (1029)$$

quando dizemos que  $n$  tem  $r$  fatores primos distintos.

O teorema seguinte mostra que a “densidade” dos números primos no conjunto dos números naturais é grande

$$\sum_{p \in P} \frac{1}{p} \quad (1030)$$

diverge, em que  $P$  é o conjunto de todos os números primos. Compare o conceito de “densidade” com a série obtida quando  $P$  for o conjunto das potências de qualquer inteiro, então uma série geométrica, logo, convergente.

É divertido observar que todos os conjuntos  $P$  mencionados têm a mesma cardinalidade apenas o salto entre os seus elementos aumenta com grande “celeridade” em todos os casos, exceto quando  $P$  for o conjunto dos números primos.

Isto justifica a intensa pesquisa que continua a ser feita para entender a estrutura do conjunto dos números primos.

Observe que o produto

$$\prod_{p \in \{2,3,5\}} \left(1 + \frac{1}{p}\right) = \quad (1031)$$

$$= 1 + \sum_{p \in \{2,3,5\}} p + \sum_{p,q \in \{2,3,5\}} pq + \sum_{p,q,r \in \{2,3,5\}} pqr \quad (1032)$$

$$= 1 + \sum_{p \in \{2,3,5\}} p + \sum_{p,q \in \{2,3,5\}} pq + \sum_{p,q,r \in \{2,3,5\}} pqr \quad (1033)$$

$$= 1 + \sum_{p \in \{2,3,5\}} p + \sum_{p,q \in \{2,3,5\}} pq + pqr \quad (1034)$$

pode ser escrito como somas de produtos com termos aumentando sucessivamente o número de fatores o que nos permite de chamar um “produto infinito” (que não existe, como também não existem “somas infinitas”) de *série*. O desenvolvimento nas equações (1031)- (1035) mostra que a *série*

$$\prod_{p \in P} \left(1 + \frac{1}{p}\right) \quad (1035)$$

em que  $P$  é o conjunto de todos os números primos, contém todos os termos da série na equação (1030) e a conclusão é a de que se esta divergir, também a série na equação (1035) diverge. Este é um resultado bem conhecido, ambas divergem!

Dois teoremas para fechar o verbete! O primeiro conhecido como *teorema dos números primos*

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{x/\ln(x)} = 1 \quad (1036)$$

A notação,  $\pi(x)$ , representa todos os números primos menores ou igual ao número primo  $x$ .

O segundo, o teorema de Mertens

$$\prod_{p < x} \left(1 - \frac{1}{p}\right) \approx \frac{e^{-\gamma}}{\ln(x)} \quad (1037)$$

quando  $x$  cresce indefinidamente e  $\gamma$  é a constante de Euler definida por

$$\gamma = \lim_{n \rightarrow \infty} \left( \sum_{k=1}^n \frac{1}{k} - \ln(n) \right) \quad (1038)$$

Os números primos são muito importantes na teoria das congruências que servem para classificar objetos, criar estruturas aritméticas finitas. Estes dois conceitos fazem parte da construção dos algoritmos de criptografia.

Dois números dizem-se *coprimos* se o conjunto dos seus fatores for disjuncto desconsiderando o fator 1. Por exemplo 21, 12 não são coprimos porque 3 é

um fator comum e  $15 = 3 \times 5$ ,  $14 = 2 \times 7$  são coprimos. Se uma fração for irredutível então o numerador e o denominador são *coprimos*, ou ainda como se diz, *primos entre si*.

O algoritmo da divisão sucessiva euclidiana, aplicado a dois números  $a, b$  podemos determinar se eles são coprimos: o resto da divisão tem que ser 1. Se o resto da divisão possível entre  $a, b$  for  $r \neq 1$  então  $r$  divide ambos os números que assim não serão coprimos. Mas este método é inútil na prática, é melhor a busca sucessiva de divisores dos números  $a, b$ . Porém é o método para construir uma demonstração o que dá um exemplo de uma situação muito comum em que os *métodos teóricos* podem ser destituídos de *aspectos práticos*. Conclusão, a visão *utilitarista* é anti-científica porque tolhe o desenvolvimento científico ante a barreira produzida pela *aplicabilidade* que é o *motor central* das famigeradas agências científicas. . . Confira Roger, Leonard James.

---

- **primos entre si** Dados dois números naturais,  $p, q$  dizemos que ele são *primos entre si* se ele não tiverem fatores comuns. Por exemplo 12, 15 não são *primos entre si* uma vez que 3 é um fator comum entre 12 e 15. Mas 21 e 25 são primos entre si. Um teorema de Dirichlet descreve uma lista de números primos, dados  $p, q$  primos entre si.

$$f(n) = pn + q; n \in \mathbf{N} \quad (1039)$$

produz uma lista infinita de números primos.

---

- **problema** A palavra *problema* tem um uso particular em equações diferenciais, representando a busca de uma solução com certa unicidade, ver *problemas, equações diferenciais*. Alguns *problemas* em franca atividade são

- problema de Cauchy, (Cauchy-Kovalevska); Ver *Cauchy, problema de*.
- problema de Riemann-Hilbert;

---

- **problema, mal posto** Não se trata de uma adjetivação *negativa*, mas a constatação da *realidade*. Quando se tenta modelar *problemas da vida real* se constata a grande dificuldade de *formalizar* a realidade. Se chama de *problema mal posto* aqueles que deixam difusa as condições em que acontece um certo fenômeno. Como é importante criar modelos o risco de criamos *problemas mal postos* é grande mas isto não deve ser desencorajador, apenas dever haver uma crítica na formulação (um aviso para o utilizador!). Há várias técnicas, diversos métodos, com distinto nível de abstração para corrigir a redação de problemas. Posso citar alguns (uma lista incompleta):

1. o da melhor curva, fitting curves Suponha que  $f$  seja uma função (bem posta) e se deseje resolver a equação  $f(x) = y$  em que  $x$  ou  $y$  seja *um problema mal posto*. Supondo-se que  $x$  ou  $y$  sejam vetores, quer dizer um conjunto dados coletados com imprecisões, uma saída seria recorrer

a “curva melhor adaptada”, uma curva que melhor represente estes dados usando uma distância apropriada nesta decisão. Um tipo de método comum é o dos mínimos quadrados, neste caso a distância é a distância euclidiana. A figura (126) página 341, ilustra aproximação por quadrados mínimos para encontrar uma curva que melhor se adapte a uma massa de dados discretos. Na figura, a curva em azul representa a *fitting curve*, a curva que melhor se adapta, calculada por quadrados mínimos dos dados representados pelos pontos em cor preta. A figura (126) foi obtida usando

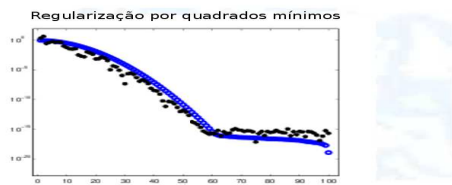


Figura 126: Aproximação por quadrados mínimos

um script para `gnuplot`.

2. Regularização de Tikhonov. Embora seja mencionada como um método, na verdade há uma pluralidade de métodos sob este nome. Com a mesma notação do item anterior,  $x$ , ou  $y$  são dados obtidos com erros, um dos métodos de regularização de Tikhonov envolve uma matriz de Tikhonov que é da forma  $\alpha I$ , uma matriz escalar em que a constante  $\alpha$  é calculada usando o desvio padrão,  $\alpha = \frac{\sigma_x}{\sigma_y}$ . Embora esta metodologia pareça nebulosa, ela se mostra, na prática, efetiva. Ela se baseia na Estatística Bayesiana que tem uma justificativa ingênua, mas fundamentada: se os dados foram colhidos por um experimento sério, e contém erros, é possível filtrar os erros procurando encontrar a melhor forma de corrigir os dados usando uma das probabilidades que a Natureza aprovou! O cálculo do desvio padrão envolve um valor médio que depende da seleção de uma probabilidade consequentemente o método consiste em testar distintas probabilidades, e naturalmente, verificar se a regularização produz bons resultados quando experimentada com dados reais, *se possível*.
3. Regularização de Tikhonov em espaços de funções. Procure *regularização*.

- **problema, equações diferenciais** As equações diferenciais, quando têm soluções, tem uma infinidade delas. O exemplo mais simples é o cálculo de uma primitiva, no Cálculo Diferencial e Integral:

$$f' = x \Rightarrow f(x) = \frac{x^2}{2} + C \quad (1040)$$

em que  $C$  é um número real qualquer. Porém se esta questão for apresentada como:

$$y' = x; \quad (1041)$$

$$x = 0 \Rightarrow y = 0; \text{ condição inicial} \quad (1042)$$

então o valor de  $C$  na equação (1040) fica determinado, vale zero, e a solução é única. Costuma-se chamar de um problema quando uma equação diferencial é posta desta forma com condições que particularizam a solução. Observe o nome “condição inicial” que provavelmente vem do significado físico de algumas equações diferenciais estudadas no século 15 quando elas representavam em geral problemas de velocidade, como é usada em Física, “velocidade inicial” é uma condição inicial de uma equação diferencial. Uma integral definida resume o problema:

$$\int_a^x t dt \quad (1043)$$

a condição inicial é  $x = a$  quando  $y = 0$  e o resultado do problema é  $y = \frac{(x-a)^2}{2}$ .

A equação estudada acima é de primeira ordem, se a ordem for maior será preciso um número maior de “condições” para obter uma solução única, porque uma equação de ordem  $n$  pode ser expressa como um sistema de equações de primeira ordem, então serão necessárias, em princípio,  $n$  condições para “criar um problema”, mas isto é equivalente a ter *uma condição vetorial*. Se a equação diferencial ordinária for linear podemos dizer que uma condição inicial (numérica, ou vetorial) é o necessário para fixar uma solução única.

Se passarmos às equações diferenciais parciais, como elas são postas em uma região do plano ou do espaço, então passamos a falar de “condições de fronteira”:

$$\frac{\partial \rho}{\partial t} + \nabla_u \rho = 0; t \in (a, b); (x, y, z) \in \Omega \subset \mathbf{R}^3 \quad (1044)$$

$$\rho \left[ \frac{\partial u}{\partial t} + \nabla_u \rho \right] = \rho F - \nabla p + \mu \nabla^2 u; \quad (1045)$$

$$\text{div}(v) = 0; u(x, y, z) = f(x, y, z) \text{ se } (x, y, z) \in \partial\Omega \quad (1046)$$

esta equação está definida numa região do  $\mathbf{R}^4$  que contém a região  $\Omega$  que é tridimensional e os valores de  $u$  estão sendo estipulados na fronteira de  $\Omega$  na equação (1046) pela função  $f$ , observe que  $f$  é um dado do problema, é conhecida são os valores que a função desconhecida,  $u$ , assume sobre a fronteira. Temos aqui um *problema com valores na fronteira*, (PVF) ou a sigla em inglês (BVP), também se costuma dizer *um problema com condição na fronteira*.



Um experimento simples pode dar lhe uma visão concreta do significado da fronteira na solução de uma equação diferencial parcial, a figura (127), página 369, lhe mostra duas regiões planas onde simulamos um problema com valor na fronteira e que você pode repetir usando regiões distintas, como retangulares ou mesmo com buracos em seu interior, e ver surgir distintos tipos de superfícies que serão produzidas como solução de uma certa equação diferencial. Se houver uma modificação na fronteira, a superfície-solução será diferente, como você pode ver nos dois exemplos de superfícies obtidas com dois domínios diferentes. As figuras mostram o resultado da experiência em que usamos “açúcar refinado” para simular um fluido. Se lhe parecer que o uso de um meio poroso é inadequado para tratar de uma questão de dinâmica de fluidos, fazemos referência à introdução de um dos clássicos do assunto, *Supersonic Flow and Shock Waves* de Courant e Friedrichs, em os autores que usam como um exemplo inicial para *ondas de choque* o comportamento de uma multidão que busca a saída dum prédio motivada pelo pânico... então podemos usar açúcar refinado para entender o que acontece com uma superfície submetida a uma condição de fronteira. Aqui não há nenhum choque envolvido, apenas mostramos que a fronteira influi no formato da superfície.

Se você usar distintos fluidos, terá também superfícies distintas. mas o melhor mesmo é que você gere sua própria experiência baseado no método que aqui lhe expomos: distintos domínios correspondem a distintas superfícies. Seria injusto não mencionar que vi esta experiência num evento de universidade aberta, Universidade de Upsala, apresentado por Gunnar Aronsson, não foi eu que a inventei. Gunnar usou areia de praia em vez de açúcar...

- **produto cartesiano** é uma operação entre dois conjuntos. Notação: se  $A, B$  forem dois conjuntos, então

$$A \times B = \{(x, y); x \in A \text{ e } y \in B\}; \quad (1047)$$

$$B \times A = \{(x, y); x \in B \text{ e } y \in A\}; \quad (1048)$$

consequentemente o *produto cartesiano* não é comutativo.

Os produtos cartesianos aparecem nas comunicações técnico-científicas muitas de forma implícita, por exemplo em gráficos estatísticos, de funções que são subconjuntos de algum produto cartesiano. O *plano cartesiano* na figura (fig 128), página 344, mostra dois gráficos, da função  $y = f(x)$  e da função  $y = g(x)$ .

A figura (fig 129), página 345, mostra o produto cartesiano do conjunto  $A = \{-5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5\}$  com ele mesmo,  $A \times A$  e nele foi selecionado o conjunto

$$C = \{(x, y) \in A \times A; x \leq y\}$$

que são os pares em que a ordenada é maior ou igual do que a abcissa. Um subconjunto de  $C$  seria

$$H = \{(x, y) \in A \times A; x = y\}$$

que é a diagonal que limite inferiormente o conjunto  $C$ .

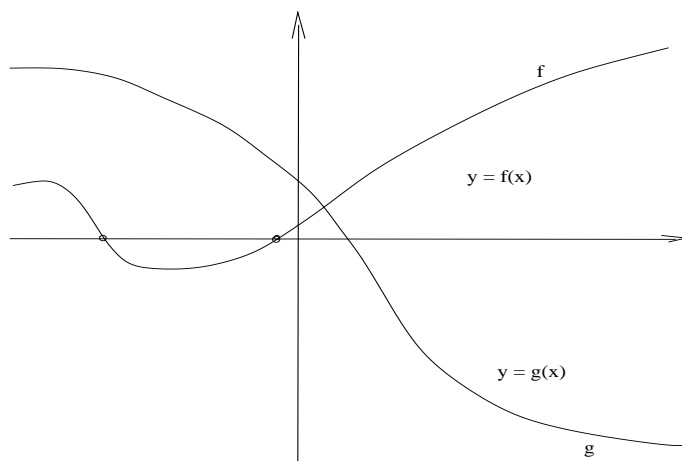


Figura 128: produto cartesiano  $\mathbf{R} \times \mathbf{R}$

- **produto escalar** Produto escalar, ou ainda *produto interno* é uma “falsa” operação binária, no sentido de que sai do espaço onde está definida. O resultado do produto escalar é um número, um escalar.

Mesmo “falsa” é de importância significativa! O produto escalar está envolvido em várias relações interessantes e junto com ele, o círculo trigonométrico  $\mathbf{S}^1$ , que é uma espécie *transferidor* no plano.

Na figura (fig. 130), página 346, um ponto qualquer sobre  $\mathbf{S}^1$  tem coordenadas  $(\cos(\alpha), \sin(\alpha))$  e assim todas as direções do plano estão, nele, anotadas:

Há duas definições equivalentes do produto escalar de dois vetores:

**Definição 20** *Produto escalar de dois vetores*

Dados dois vetores  $u = (a, b); v = (p, q)$  definimos o produto escalar entre eles por uma das duas formas:

$$\langle u, v \rangle = ap + bq; \quad (1049)$$

$$\langle u, v \rangle = \|u\| \|v\| \cos(\gamma); \gamma \text{ é o ângulo determinado por } u, v; \quad (1050)$$

Vou fazer uso de  $\mathbf{S}^1$  na demonstração do próximo teorema e prepare-se para ver  $\mathbf{S}^1$ , o círculo trigonométrico, aparecer com frequência em questões envolvendo ângulos, como eu disse, ele é uma espécie de transferidor.

Precisamos formalizar com um *teorema* que estas duas definições são equivalentes.

**Teorema 18** *Produto escalar* As definições para o produto escalar na definição 20, são equivalentes.

**Dem**:

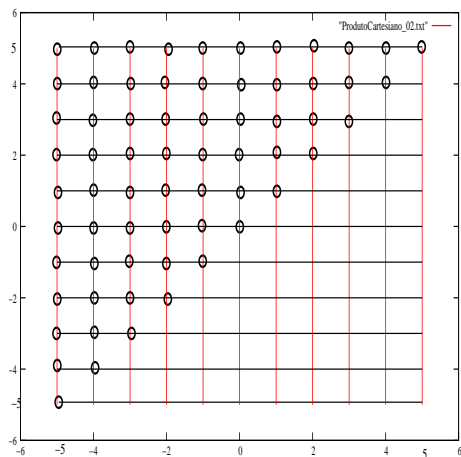


Figura 129: produto cartesiano de conjuntos finitos

Projete estes vetores no círculo trigonométrico  $\mathbf{S}^1$ , então  $u, v$  determinem os ângulos  $\alpha, \beta$ , respectivamente, com  $\alpha - \beta = \gamma$ , supondo que  $\alpha > \beta$ .

Podemos escrever

$$u = (a, b) = \|u\|(\cos(\alpha), \sin(\alpha)); a = \|u\| \cos(\alpha); b = \|u\| \sin(\alpha); \quad (1051)$$

$$v = (p, q) = \|v\|(\cos(\beta), \sin(\beta)); p = \|v\| \cos(\beta); q = \|v\| \sin(\beta); \quad (1052)$$

$$\langle u, v \rangle = \langle \|u\|(\cos(\alpha), \sin(\alpha)), \|v\|(\cos(\beta), \sin(\beta)) \rangle = \quad (1053)$$

$$\|u\|\|v\| \cos(\alpha) \cos(\beta) + \sin(\alpha) \sin(\beta) = \quad (1054)$$

$$\|u\|\|v\| \cos(\alpha - \beta) = \|u\|\|v\| \cos(\gamma); \quad (1055)$$

**q.e.d .**

Como dois vetores determinam um plano, esta demonstração vale para o produto escalar de dois vetores em qualquer dimensão que eles estejam. Usar uma definição ou a outra é uma questão de selecionar qual das duas traduz melhor o significado do produto escalar em cada situação.

Por exemplo, no próximo teorema, cuja demonstração é um exercício interessante, é mais prático usar a primeira definição em sua demonstração.

**Teorema 19** *Bilinearidade do produto escalar*

*O produto escalar tem as seguintes propriedades:*

$$u_1 = (a_1, b_1); u_2 = (a_2, b_2); v = (p, q); \quad (1056)$$

$$\langle u_1 + u_2, v \rangle = \langle u_1, v \rangle + \langle u_2, v \rangle; \quad (1057)$$

$$\langle v, u_1 + u_2 \rangle = \langle v, u_1 \rangle + \langle v, u_2 \rangle; \quad (1058)$$

$$\lambda \langle u, v \rangle = \langle \lambda u, v \rangle = \langle u, \lambda v \rangle; \quad (1059)$$

$$(1060)$$

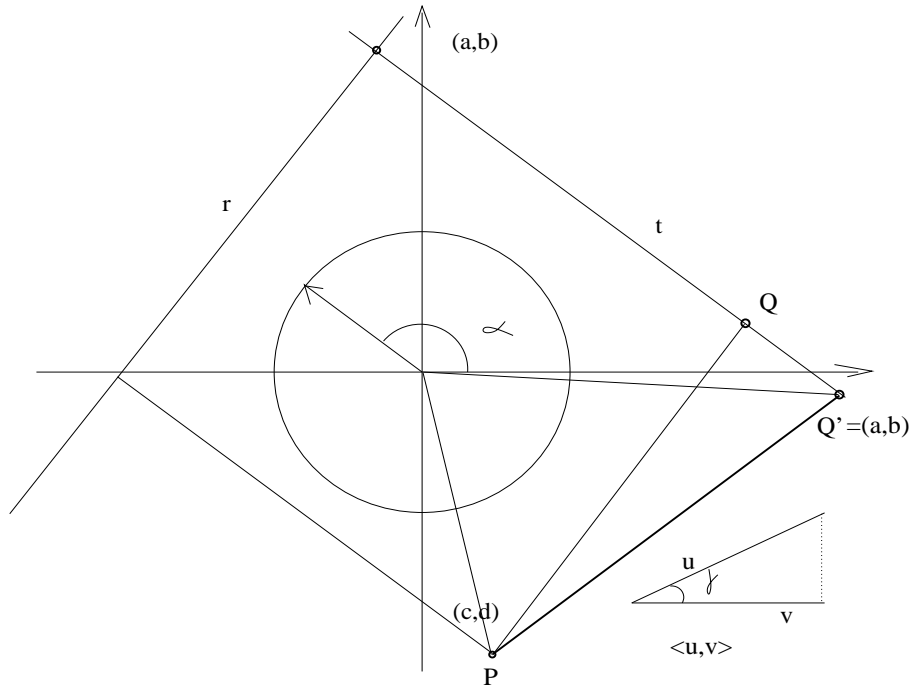


Figura 130: O círculo trigonométrico  $S^1$

A “linearidade” é a propriedade de preservar soma e produto na forma como aparece no teorema. E como vale para os dois lados da expressão então dizemos que é bilinear. Estas propriedades se demonstram usando a primeira forma da definição. A segunda forma da definição aparece em questões de projeção como o próximo teorema mostra.

**Teorema 20** *Projeção com produto escalar*

*Se o vetor  $u$  for unitário, então  $\langle u, v \rangle$  mede a projeção do vetor  $v$  na direção do vetor  $u$ .*

A demonstração é uma simples interpretação geométrica do significado do produto escalar usando a segunda definição na equação (eq. 1049).

O que é importante é chamar sua atenção sobre como usar esta propriedade, lembre-se que por repetidas vezes chameis o *círculo trigonométrico* de “transferidor”, coloque o centro dum círculo unitário na interseção de duas retas e aplique o produto escalar em sua segunda versão quando precisar de calcular a projeção dum vetor na direção de outro.

O produto escalar está envolvido na determinação da *equação da reta* e na *equação do plano*, na definição do módulo de um vetor, na distância de um ponto a uma reta, para mencionar algumas das aplicações mais diretas desta operação.

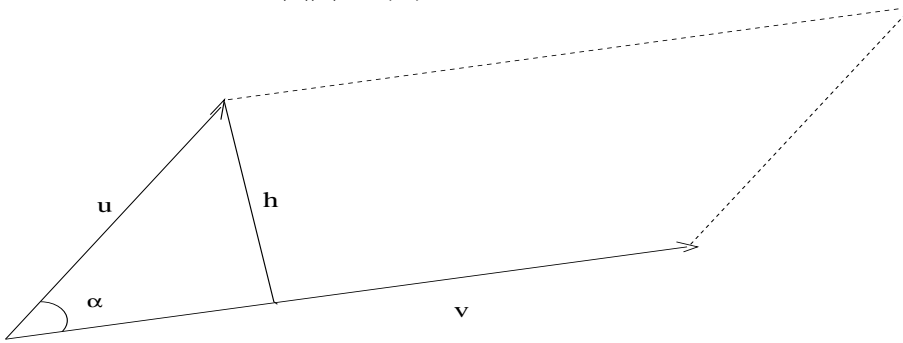
- **produto vetorial** de dois vetores  $\vec{u}, \vec{v}$  define um novo vetor  $\vec{w}$  com as seguintes propriedades:

**Definição 21 (Produto Vetorial)** *Produto Vetorial*

- $\|\vec{w}\| = \vec{u} \times \vec{v} = \|\vec{u}\|\|\vec{v}\| \sin(\alpha)$  em que  $\alpha$  é o ângulo que os vetores  $\vec{u}, \vec{v}$  determinam no círculo unitário do plano que eles determinam.
- $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$
- $\vec{u}, \vec{v}, \vec{u} \times \vec{v}$ , nesta ordem, é um triedro orientado positivamente.

Na figura (fig 156), página 391,  $u \times v = |u||v| \sin(\alpha)$  podemos ver que o

$$u \times v = |u||v| \sin(\alpha)$$



$$u \times v = \text{Área do losângulo}$$

Figura 131:  $\vec{u} \times \vec{v} = \text{Área do losângulo}$

significado geométrico do produto vetorial de dois vetores  $\vec{u}, \vec{v}$ . Eles determinam um losângulo quando copiamos paralelamente os segmentos que correspondem aos vetores e a projeção de  $\vec{u}$  sobre a perpendicular à reta suporte de  $\vec{v}$  vale  $\|\vec{u}\| \sin(\alpha)$ . O produto  $\|\vec{u}\|\|\vec{v}\| \sin(\alpha)$  é a área deste losângulo determinado pelos vetores  $\vec{u}, \vec{v}$ .

Temos os exemplos, com os vetores do triedro fundamental da Física:

$$\vec{i}, \vec{j}, \vec{k}; \text{triedro dos vetores unitários da Física}; \quad (1061)$$

$$\vec{i} \times \vec{j} = \vec{k}; \vec{j} \times \vec{k} = \vec{i}; \vec{k} \times \vec{i} = \vec{j}; \quad (1062)$$

$$(1063)$$

Se a dimensão do espaço  $E$  for maior ou igual a três, o *produto vetorial* é uma operação binária bilinear alternada (anticomutativa) em  $E$ . No  $\mathbf{R}^3$  munido com o triedro da Física,

---

- **programação** é usada com frequência nos verbetes deste dicionário baseadas nas linguagens `calc` ou `python` que são duas linguagens interpretadas e de domínio público no sentido de que são livremente distribuídas na Internet.

`python` pode ser encontrada no site [www.python.org](http://www.python.org) e está comumente incorporada a todas as distribuições `Gnu/Linux`.

`calc` pode também ser encontrada em algumas distribuições `Gnu/Linux`, como por exemplo em `Debian/Gnu/Linux` mas também pode ser livremente baixada do site do grupo que mantém

<http://www.isthe.com/chongo/tech/comp/calc/calc-what-is.html>

Também é feito uso do pacote computacional `gnuplot` que pode funcionar como uma linguagem muito rudimentar de programação, mas o seu objetivo, e no qual é muito forte, é fazer gráficos sendo inclusive possível chamar `gnuplot` de dentro dum programa escrito em qualquer linguagem de programação para obter um gráfico produzido com um programa da linguagem.

---

- **radiação** É um dos três métodos de transmissão de calor: condução, convecção e radiação. Ver *condução* como método de transmissão do calor.

É o caso em que dois corpos, de diferentes temperaturas, se encontram imersos em um fluido que conduz o calor entre eles (no vácuo não há transmissão de calor). A hipótese é que o fluido e o corpo mais frio se encontram a mesma temperatura. Na verdade aqui estamos em presença da *convecção* em que o calor se transmite do *corpo quente* para o fluido e deste para o *corpo frio*.

Isto é, na radiação nos interessam dois corpos, um dos quais é uma fonte de calor, que vou chamar de quente, e o outro que vou chamar de frio. Esta radiação vai ficar sujeita à difusividade térmica do fluido que envolve o corpo quente e o frio. Existem tabelas de difusividade térmica para diversos materiais.

---

- **radiciação** é a operação inversa da potência, ou potenciação. Confira potência. Como

$$\sqrt[p]{a} = a^{\frac{1}{p}}; \quad (1064)$$

$$c = \sqrt[p]{a} \Rightarrow c^p = a; \quad (1065)$$

então podemos simplesmente ignorar a operação raiz colocando em seu lugar a potenciação.

Este é o caminho mais curto usado pelas linguagens de programação evitando a construção dois algoritmos com a mesma funcionalidade.

---

- **raio de convergência**, associado às série de potências. Confira também o *Lema de Abel*. Uma série de potências é uma expressão da forma

$$S_n(z) = \sum_{k=0}^n a_k(z - a)^k; n \in \mathbf{N}; \quad (1066)$$

é uma sucessão de funções  $S_n$  cujo *termo geral* é uma soma.

**Observação 8 (erro) de linguagem**

Há um erro de linguagem que consiste em definir série como

$$\sum_{k=0}^{\infty} a_k (z - a)^k \quad (1067)$$

quando na verdade a equação (eq. 1067) é a expressão do limite da série na equação (eq. 1066) se este existir. Embora seja um erro de linguagem na prática é um hábito conviver com este erro.

O lema de Abel é um dos resultados mais importantes associados às séries de potência. Ele garante a existência do raio de convergência que é um processo de comparação duma série de potências com uma série geométrica. Confira *série geométrica*.

As séries geométricas são expressões da forma

$$S_n = \sum_{k=0}^n r^k; r \in \mathbf{R} \quad (1068)$$

e se  $\|r\| < 1$  converge para  $\frac{1}{1-r}$ . O lema de Abel diz que nada se pode dizer quando  $\|r\| = 1$ .

Como resultado desta comparação podemos provar que se o limite

$$\frac{1}{\limsup_n \sqrt[n]{\|a_n\|}} = \rho \quad (1069)$$

existir, então a série de potências converge no disco aberto  $D(a, \rho)$ , sendo  $\rho$  o raio de convergência da série na equação (eq.1066). Nada se pode garantir, de acordo com o lema de Abel, na fronteira do disco onde eventualmente pode haver convergência e isto é que dá uma saída para a expansão dos domínios de holomorfia, ver função holomorfa.

A equação (eq.1069) é conhecida como *teste da raiz* para o raio de convergência.

Há outras formas de calcular o raio de convergência sendo o mais comum o método do quociente:

$$\frac{1}{\limsup_n \frac{\|a_{n+1}\|}{\|a_n\|}} = \rho \quad (1070)$$

a demonstração deste método de comparação é uma direta consequência da comparação com uma série geométrica (cuja razão tem que ser, necessariamente, em módulo, menor do que 1, uma vez que “série” é um limite...). Com frequência a fórmula do raio de convergência aparece na forma mais amigável

$$\frac{1}{\rho} = \limsup_n \frac{\|a_{n+1}\|}{\|a_n\|} \quad (1071)$$

O teste na fórmula (eq. 1071), ou (eq.1070), é referido como *teste do quociente* para o cálculo do raio de convergência.

---

- **real, número** Os números reais completam a reta numérica de tal modo que podemos identificar a *reta numérica* com os números reais. Na Matemática fazemos uso de vários conjuntos numéricos que tem propriedades e objetivos distintos, como os *número naturais*,  $\mathbf{N}$ , os números racionais,  $\mathbf{Q}$ , os reais e os complexos. Os números reais completam os números racionais na *reta numérica*, uma outra forma de ver este números é com o conceito de *limite* confira *convergência*.. Confira a representação dos racionais e dos reais na *reta numérica*.

---

- **redundante** Que contem repetições de informações. Uma base ortogonal de vetores é um exemplo de sistema não redundante. Uma base de um espaço vetorial, por definição, não pode ser redundante porque o número dos seus elementos é que define a dimensão do espaço.

A redundância é uma *fraqueza* na teoria mas pode ser um aspecto positivo. Isto foi descoberto com a Teoria da Informação, o acréscimo de uma quantidade maior de elementos no *sistema de geradores* de uma estrutura acrescenta mais *informação* a ser utilizada nos processos estatísticos de correção dos dados transmitidos (ou na recuperação de dados compactados em processos de encriptação. Desta forma se criou um conceito alternativo ao de base que é *dicionário*. Um dicionário é um *sistema de geradores* que pode ser redundante...em geral é redundante.

---

- **regularização**

1. **regularização por convolução** A convolução é chamada, em um livro de H S Shapiro, *média viajante* o que se justifica da seguinte maneira:

- (a)  $f \mapsto f * \eta = \phi; \phi(x) = \int \eta(t)f(x-t)d\mu(t)$  em que esta integral é calculada sobre um domínio conveniente, procure *convolução* para ver um exemplo mais escalarecedor.
- (b) Na equação acima  $\eta$  é uma função positiva cuja integral é 1, portanto, o valor de  $f * \eta$  em cada ponto  $x$ ,  $\phi(x)$ , é um valor médio de  $f$  nas vizinhanças deste ponto  $x$ . Isto fica mais claro se acrescentarmos uma hipótese, que o suporte de  $\eta$  é um conjunto compacto com métrica “muito pequena”, por exemplo, se a integral  $f * \eta$  estiver sendo calculada sobre  $\mathbf{R}$  e o suporte de  $\eta$  for o intervalo  $[-\epsilon, \epsilon]$  com  $\epsilon$  “muito pequeno”.

Uma das consequências notáveis da convolução, se  $f$  for integrável, é que basta que o fator  $\eta$  seja também integrável para que  $f * \eta$  seja contínua, e muitas vezes *quase sempre* derivável. Se  $\eta$  for de classe  $\mathcal{C}^n$  então  $f * \eta$  será pelo menos de classe  $\mathcal{C}^n$ . O polinômio trigonométrico de ordem  $n$  de  $f$  é a convolução  $f$  com o núcleo de Dirichlet de ordem  $n$  e conseqüentemente, bastando que  $f$  seja integrável, o polinômio trigonométrico será de classe  $\mathcal{C}^\infty$ .



Mas a razão principal a que se refere a regularização por convolução, e que H S Shapiro a chame de *média viajante*, é que se  $\eta$  for positiva e com integral 1, e  $x = a$  for um ponto de *salto de primeira espécie* de  $f$ , então  $(f * \eta)(a)$  é o *valor médio integral* de  $f$  numa vizinhança de  $x = a$ .

## 2. regularização de Tikhonov em espaços de funções

### - regularização de Tikhonov em espaços de funções

- **relação de equivalência** é uma relação  $R$  binária definida em conjuntos satisfazendo às propriedades

**Definição 22 (relação de equivalência)** *relação de equivalência*

Seja  $S$  um conjunto e  $R$  uma relação binária, definida em  $S \times S$  tal que as seguintes propriedades sejam satisfeitas

1. reflexividade  $R(x, x)$  é verdadeira.
2. simetria  $R(x, y) \implies R(y, x)$ .
3. transitividade  $R(x, y)$  e  $R(y, z) \implies R(x, z)$ .

Dizemos que as relações de equivalência são reflexivas, simétricas e transitivas e toda relação binária que tenha estas propriedades é uma relação de equivalência. As relações de equivalência *generalizam* o conceito de igualdade, porque a igualdade tem estas três propriedades.

As relações de equivalência resolvem uma questão lógica básica, por exemplo em geometria, se fala de “igualdade de triângulos” quando na verdade se deveria falar de “equivalência de triângulos” com maior propriedade quando se faz referência a dois triângulos, tal que, transladando um é possível fazer com que dois lados consecutivos fiquem coincidentes assim como os ângulos que eles determinam, o caso LAL. Resolvem também a questão da identidade entre frações equivalentes que definem o mesmo número racional.

- **reta**. Confira *equação da reta*.

- **reta numérica**. O sistema de coordenadas cartesianas basicamente identifica os pontos da *reta numérica* com um número real e é isto que chamamos de *coordenada* dum ponto na reta.

Se identificarmos um ponto, numa reta qualquer, como sendo o *zero*, à direita do qual, por convenção se identifica um outro ponto como sendo o 1 como você pode ver na figura (fig 132), página 352, então criamos uma sistema para “numerizar a reta” ou “digitalizar a reta”. Desta forma a reta deixa de ser apenas um ente geométrico e passa agora a ser uma representação do conjunto dos números reais **R**.

- A escolha dum ponto para representar o zero dividiu a reta em duas semiretas.

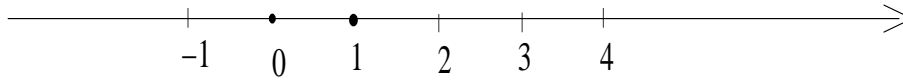


Figura 132: reta numérica, representação geométrica de  $\mathbf{R}$

- A escolha dum ponto para representar o número 1 selecionou uma das semiretas como a *semireta positiva*, mas fez mais do que isto definiu um segmento de reta com medida 1 o que nos permite propagar pela reta numérica todos os números inteiros usando, por exemplo, um compasso para marcar os inteiros positivos e negativos.
- Os números negativos são marcados na *semireta negativa*, naturalmente.

A escolha do 1, por convenção, à direita do zero, também define quais são os números positivos e a semi-reta dos números positivos, e consequentemente, também, a semi-reta dos números negativos, à esquerda do zero.

Vou mostrar que é possível fazer-se uma construção geométrica dos números reais portanto a afirmação que aparece na etiqueta da figura (fig 132) está correta, a *reta numérica* é uma representação dos conjunto dos números reais,  $\mathbf{R}$ .

Na figura (fig 133), página 352, você pode ver paralelas à reta que passa por

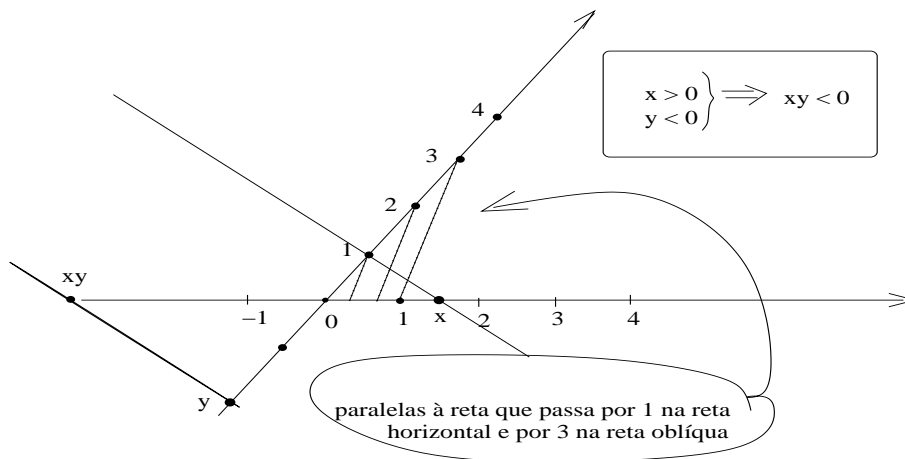


Figura 133: Determinação dos racionais na reta numérica

1, na reta horizontal e por 3 na reta oblíqua permitindo encontramos  $\frac{1}{3}, \frac{2}{3}$ , na reta horizontal, entre 0 e 1, com relativa exatidão. Da mesma forma, traçando uma reta passando por 1, na horizontal e por um número  $m$  qualquer na oblíqua, podemos determinar as *frações próprias* de denominador  $m$  entre 0 e 1, de volta na reta horizontal. Você pode ver isto na figura (fig 133).

A soma de números, na reta numérica, se faz com soma de segmentos de reta, desta forma temos uma adição geométrica definida na *reta numérica*. Com a

soma geométrica podemos expandir a representação de qualquer número racional na reta inteira a partir das frações próprias que acabamos de ver representadas entre 0 e 1. Em suma, qualquer número racional pode ser marcado na *reta numérica* com razoável precisão, usando os métodos do desenho geométrico.

Para fazer a multiplicação podemos usar semelhança de triângulos. Observe a figura (fig 134), página 353.

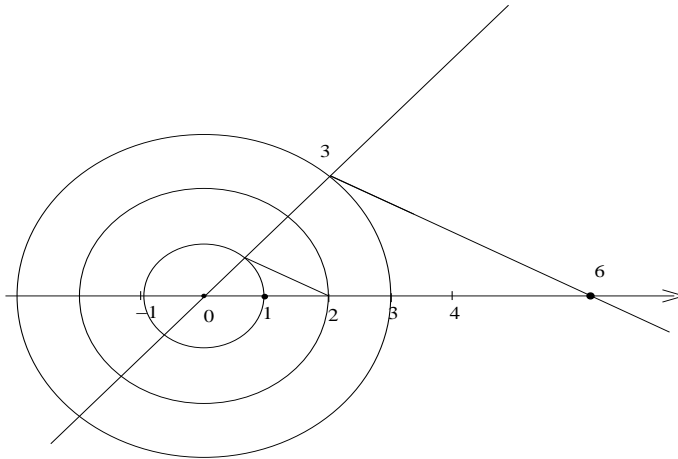


Figura 134: A multiplicação geométrica na reta

Com círculos concêntricos no ponto zero podemos transferir as marcas dos números na reta horizontal para a reta oblíqua passando pelo zero. Ligando 2 na reta horizontal com 1 na oblíqua, e depois passando uma reta paralela a esta pelo 3 na oblíqua, vamos encontrar 6 na horizontal e assim multiplicamos  $3 \times 2 = 6$ . Como podemos marcar qualquer número racional nestas retas, então definimos por semelhança de triângulos a multiplicação geométrica na reta.

Tendo adição e multiplicação na reta numérica onde podemos encontrar qualquer número racional nos mostra que temos uma representação geométrica dos números.

Porém na *reta numérica* existem números que não são racionais. A figura (fig 135), página 354, mostra como podemos calcular a  $\sqrt{m}$ ;  $m \in \mathbf{N}$ , a raiz quadrada de qualquer número inteiro positivo. Os círculos têm como raio  $\sqrt{m}$ , o primeiro corresponde à  $\sqrt{2}$ , o segundo  $\sqrt{3}$ , o terceiro  $\sqrt{4}$ . Basta levantar uma perpendicular à reta horizontal, em  $\sqrt{m}$ , quando ela encontrar a paralela que esta uma unidade acima, você tem o raio que corresponde à próxima raiz  $\sqrt{m+1}$ .

Isto mostra que qualquer raiz quadrada de número natural pode ser marcada com boa exatidão usando métodos geométricos, na *reta numérica*. Em particular temos  $\sqrt{2}$  que não é um número racional, mais exatamente,  $\sqrt{2}$  é um número irracional. A demonstração de que  $\sqrt{2}$  não é racional é relativamente simples, habitualmente é feita por contradição.

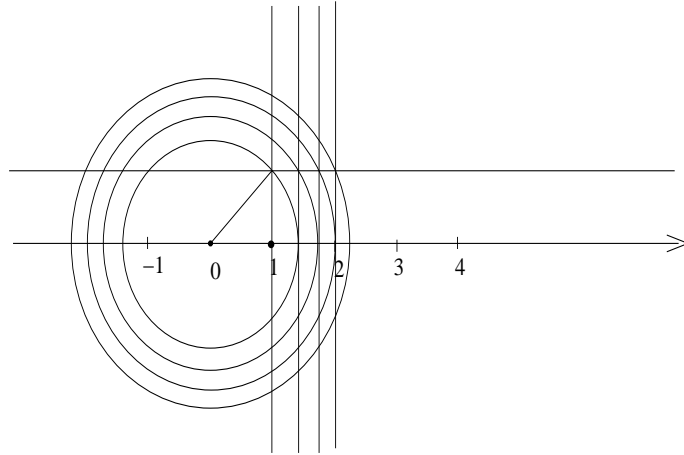


Figura 135: Calculando  $\sqrt{m}; m \in \mathbf{N}$

Suponha que  $\sqrt{2}$  seja um número racional, então pode ser escrito na forma mais simples:

$$\sqrt{2} = \frac{p}{q}; \quad p \text{ e } q \text{ primos entre si}; \quad (1072)$$

$$2 = \left(\frac{p}{q}\right)^2 \quad \text{elevando ambos os membros ao quadrado} \quad (1073)$$

$$2 = \frac{p^2}{q^2} \Rightarrow 2q^2 = p^2 \quad (1074)$$

A conclusão a partir da equação (1074) é que  $p^2$  é um número par, como é um número inteiro par e também um quadrado de número inteiro então 2 é fator de  $p^2$  o que mostra que  $q^2$  também é par.

Isto é uma contradição porque partimos da hipótese de que havíamos escrito o número racional  $\sqrt{2}$  em sua forma mais simples, com denominador e numerador primos entre si quando agora chegamos à conclusão de que ambos são números pares portanto divisíveis por 2.

A falsidade consiste na hipótese de que seria possível escrever  $\sqrt{2}$  como um número racional, portanto  $\sqrt{2}$  é irracional.

O mesmo pode ser feito com qualquer raiz de número inteiro cuja raiz não seja um inteiro, é um número irracional, e a conclusão é a de que na *reta numérica* podemos encontrar todos os números racionais e também os irracionais. Como a *reta numérica* é um conjunto de números uma vez que podemos efetuar as quatro operações com os seus elementos, então é um novo conjunto numérico que contém  $\mathbf{Q}$ , é o conjunto dos números reais.

Este é uma forma geométrica de construir os números reais que vi, pela primeira vez, no livro de Hilbert *Fundamentos da Geometria*. Uma outra forma de construir os números reais se atribui a Cauchy usando o conceito de *seqüências de Cauchy*, confira *convergência*.

Você talvez esteja curioso para ver como dividir dois números diferentes de zero. A figura (fig 136), página 355, mostra como fazê-lo. Primeiro calculo  $\frac{1}{b}$ ,

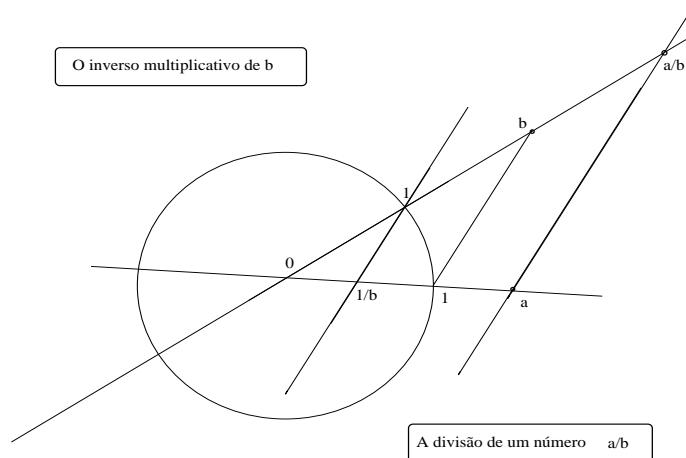


Figura 136: A divisão de um número  $a$  por  $b \neq 0$

traçando a reta que liga o neutro multiplicativo, 1 com  $b$  e uma paralela a ela passando pelo neutro multiplicativo. Agora multiplico  $\frac{1}{b}$  por  $a$  traçando uma paralela a qualquer dessas retas passando por  $a$ .

Alguns resultados são fáceis de serem deduzidos da multiplicação geométrica, como

$$a, b > 0 \Rightarrow ab > 0; \quad (1075)$$

$$b > 1 \Rightarrow \frac{1}{b} < 1; \quad (1076)$$

$$a, b > 1 \Rightarrow ab > 1; \quad (1077)$$

$$0 < a < b \text{ e } c > 0 \Rightarrow 0 < ac < bc; \quad (1078)$$

$$a < 0; b < 0 \Rightarrow ab > 0; \quad (1079)$$

$$a < 0; b > 0 \Rightarrow ab < 0; \quad (1080)$$

para obtê-los observe que “menor do que 1”, em módulo, significa “estar dentro do círculo unitário”, e que o zero divide a reta em duas classes de números: a *semi-reta dos números positivos*, e a *semi-reta dos números negativos*. Na figura (fig 137), página 356, transferi a posição de  $b$  para a outra reta usando o compasso e obtive  $ab$  na semi-reta negativa:  $ab < 0$ .

Esta operações geométricas, adição e multiplicação, podem ser generalizadas para obter-se a *álgebra dos números complexos* também definida geometricamente. É interessante que ao fazer esta generalização podemos identificar novamente o conjunto dos números reais como eixo  $OX$  do plano complexo e que as operações geométricas de números reais são um caso particular das operações geométricas dos números complexos.

- **reta orientada** é o mesmo que *reta numérica*.

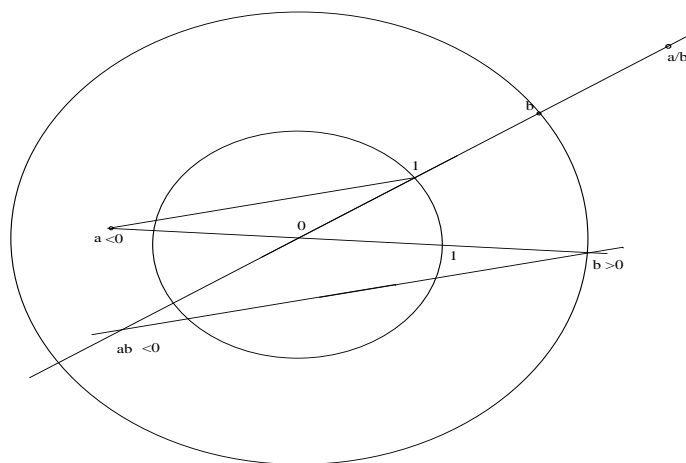


Figura 137: Produto de  $a < 0$  por  $b > 0$

- **reta tangente** Uma reta é uma *curva* caracterizada pelo coeficiente angular constante. Sendo neste sentido uma curva especial porque ela serve para determinar a direção instantânea, num determinado ponto, de outras curvas. A figura (fig 138), página 357, lhe mostra geometricamente o significado desta afirmação, nela você vê um círculo com o ponto central representando a mão de alguém que roda uma pedra presa a um cordão. Num dado momento se o cordão se quebra e a pedra segue pela tangente.

Como vivemos sob a *força gravitacional* da Terra, a real curva tangente será uma parábola. Mas você pode pensar numa outra situação, um foguete colocado em órbita em volta da Terra para num momento oportuno lançar outro foguete menor, uma *sonda espacial*, dirigida para um dos planetas do sistema solar. É o que lhe pretende mostrar a figura (fig 139), página 358.

Também neste caso a curva tangente não seria uma reta<sup>64</sup> mas durante um pequeno lapso de teríamos esta impressão, de reta tangente, como a órbita da *sonda espacial* seguindo na *direção ótima* em busca do planeta desejado.

Para resolvermos os problemas, como estes de lançamentos de foguetes, por exemplo, precisamos começar errando, admitindo que é possível haver *reta tangente*, como no caso da figura (fig 138), página 357, fazendo a suposição de que a pedra parte pela reta tangente.

Escolhemos uma direção ótima, e neste momento o foguete lançador dispara a sonda na direção da reta tangente à sua órbita.

O movimento inicial que o foguete lançador dará à sonda é de uma reta tangente à sua órbita. Depois a sonda passa a ser uma *prisioneira das forças gravitacionais do sistema solar* e se os cálculos tiverem sido bem feitos ela irá caminhar na direção do planeta desejado. A figura (fig 139), página 358, sugere

<sup>64</sup>Até porque retas são “abstrações”, não existe uma única reta em todo o Universo. . .

A pedra e cordão que se quebra

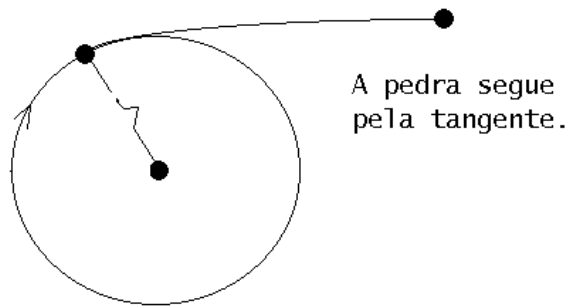


Figura 138: Reta tangente: quando o cordão se quebra

uma rota para a *sonda espacial* lançada dum foguete em órbita terrestre no *momento gravitacional ótimo*. Inicialmente a curva da sonda é tangente à órbita do foguete de lançamento, mas logo a força de gravitação solar começa a predominar e se *altera a curvatura* da trajetória da sonda espacial em busca do objetivo desejado.

Mas nós raramente sabemos escrever equações de trajetórias, estas serão determinadas pelo *campo gravitacional* do sistema solar ou de nossa galáxia. O que sabemos é equacionar retas que depois se transformam nas curvas que irão obedecer ao *campo gravitacional*. Por isto falamos de *retas tangentes*.

### Equação da reta tangente

Logo mais voltarei a equação da reta tangente que serve como órbita inicial da sonda no momento em que foguete lançador lhe dá impulso inicial que a irá “libertar” da *força gravitacional* da Terra.

Deixe-me mostrar-lhe um caso mais simples e para o qual você poderá usar um programa em **gnuplot** para ver as retas tangentes. A figura (fig 140), página 359, exhibe o gráfico de uma função  $y = f(x)$  no qual três pontos foram escolhidos,  $(a, f(a))$ ,  $(b, f(b))$ ,  $(c, f(c))$  e neles foram desenhadas retas tangentes de forma parecida com o caso da pedra da figura (fig 138).

Esta figura lhe mostra o gráfico de uma função  $y = f(x)$  cujo coeficiente angular, em cada ponto é medido pela derivada  $y = f'(x)$ . Assim, no ponto

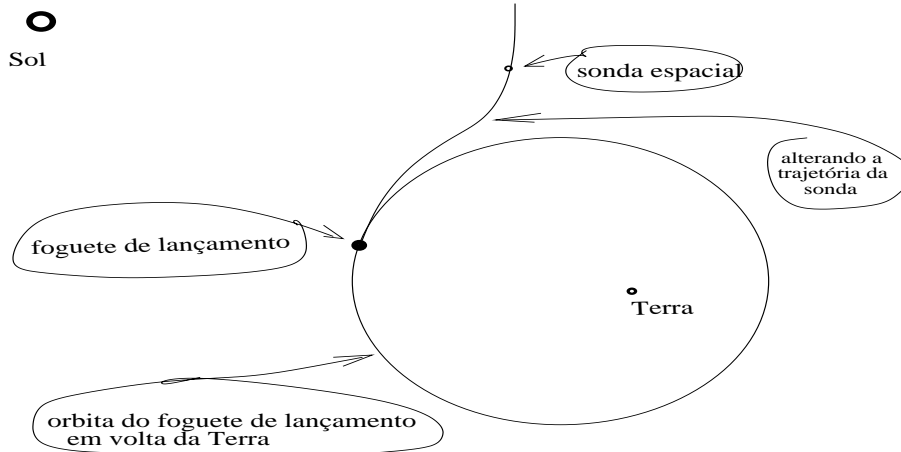


Figura 139: Curva tangente da sonda espacial lançada pelo foguete

$(a, f(a))$  o coeficiente angular “instantâneo” do gráfico de  $f$  é  $f'(a)$ . Podemos desta forma descobrir a equação da reta que passa no ponto  $(a, f(a))$  sabendo que seu coeficiente angular é  $m = f'(a)$ . Vou usar a equação da reta que passa por um ponto dado  $(a, b) = (a, f(a))$  com o coeficiente angular  $m = f'(a)$ :

$$y = b + m(x - a); \text{ reta que passa em } (a, b) \text{ com coeficiente angular } m \quad (1081)$$

$$y = f(a) + f'(a)(x - a); (a, b) = (a, f(a)); m = f'(a); \quad (1082)$$

Você pode fazer suas experiências usando **gnuplot** para conferir que a equação (eq. 1082) irá produzir o gráfico de uma reta tangente ao gráfico de

$$y = f(x) = x^3 + 3x^2 + x + 1$$

num ponto escolhido. Use o script

```
pow(x,n) = x**n; ## definindo potência em gnuplot
f(x) = pow(x,3) + 3*pow(x,2) + x + 1; ## equação da função
df(x) = 3*pow(x,2) + 6*x + 1; ## equação da derivada
a = -5; ## ponto escolhido para obtenção tangente
set xrange [a-2:a+2]; ## limita o gráfico ao intervalo [a-2, a+2]
set yrange [f(a-2):f(a+2)] ## limita o gráfico ao intervalo [f(a-2), f(a+2)]
r(x) = f(a) + df(a)*(x-a); ## equação da tangente
plot f(x), r(x), 0 ## comando para executar o gráfico
```

vai mostrar-lhe o gráfico da reta tangente no ponto  $(a, f(a)); a = -3$ . Troque a equação de  $y = f(x)$  para obter outro gráfico não esquecendo de atualizar a equação da derivada  $df(x)$ .

Ou troque o valor de  $a$  para obter outro gráfico de reta tangente ao gráfico da mesma função. Observe que você pode obter gráficos de várias retas tangentes, simultaneamente, usando “constantes”  $a, b, c, \dots$  com o risco de poluir tremendamente o gráfico.



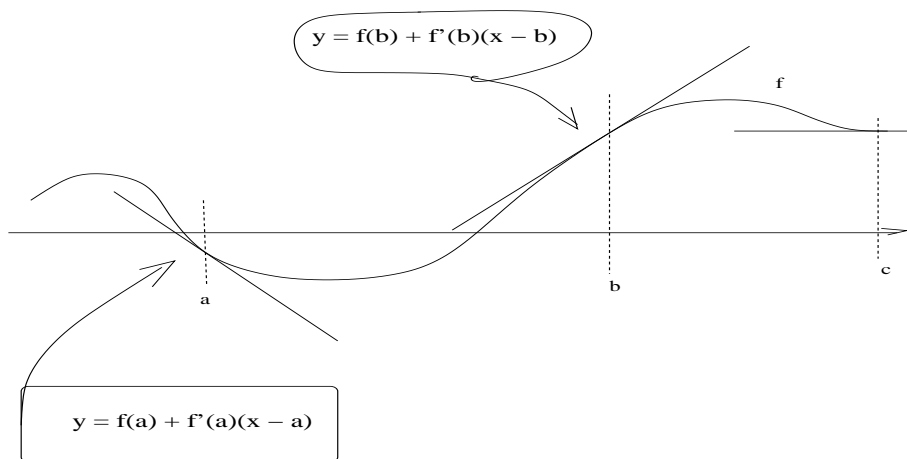


Figura 140: Reta tangente ao gráfico de  $y = f(x)$

Infelizmente não posso usar a notação da Matemática,  $f'(x)$  no script do *gnuplot* porque daria um erro. O símbolo de derivação, apóstrofe, na linguagem do *gnuplot*, é usado para demarcar textos, por isto usei a notação  $df(x)$  para representar a derivada.

O símbolo  $\#$  marca um comentário que *gnuplot* ignora, mas serve para informá-la do significado de cada linha.

A figura (fig 141), página 359, ainda lhe dá mais uma aplicação das retas tan-

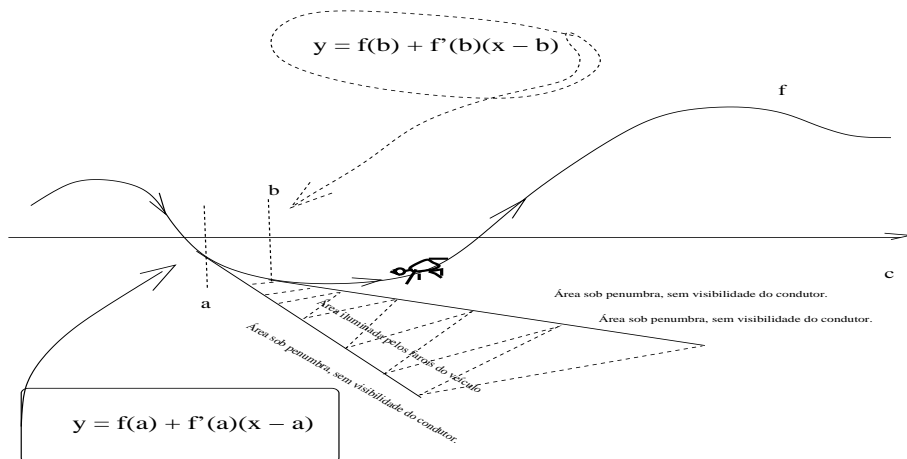


Figura 141: O cone de visibilidade

gentes. Suponha agora que o gráfico de  $y = f(x)$ , na figura (fig 141), represente uma estrada noturna.

Em dois pontos sucessivos do tempo,  $a, b$  correspondem dois pontos na estrada,  $(a, f(a)), (b, f(b))$ . As retas tangentes em cada um destes pontos mostram

- área iluminada pelos faróis do veículo.
- área sob penumbra, sem visibilidade do condutor.

O *cone de visibilidade* é a região entre as duas retas tangentes. Se houver um animal se movendo na estrada, fora do *cone de visibilidade*, haverá uma surpresa desagradável para o condutor em seguida.

### Equação vetorial da reta tangente

No caso de foguetes e sondas espaciais temos que trabalhar com equações paramétricas de curvas no espaço e com as derivadas destas curvas paramétricas. A equação do foguete lançador e a equação de uma reta tangente num ponto escolhido do tempo  $t$  se encontram nas equações

$$s(t) = (a \cos(t), b \sin(t)); \quad (1083)$$

$$s'(t) = (-a \sin(t), b \cos(t)); \quad (1084)$$

$$(a \cos(t), b \sin(t)) + \rho(-a \sin(t), b \cos(t)); \rho \in \mathbf{R}; \quad (1085)$$

(eq. 1083) (eq. 1085).

Apenas estas equações teriam que representar as *condições iniciais* em que a *sonda espacial* seria colocada dentro do *campo gravitacional solar* para determinação da sua órbita. Resolvemos este tipo de problemas com *equações diferenciais ordinárias*. A *Lua* é considerada uma excelente lançadora de *sondas espaciais*, se alguns problemas complementares puderem ser resolvidos, um deles seria combustível lunar . . . claro, a *Lua* poderia ser poluída com total liberdade pelas mega-empresas de combustível fóssil que pretendem agora destruir o Ártico. Algumas manchas de com quilómetros quadrados de extensão mal perturbariam os nossos luares, sobre tudo se ficarem do lado oculto da Lua.

---

- **retas reversas** é um item da Geometria Espacial. Num plano, dadas duas retas, existem apenas duas possibilidades:

- as retas são paralelas;
- as retas são concorrentes.

Uma reta, num plano, determina uma relação de equivalência entre todas as retas do plano que é a relação de paralelismo. Algumas vezes se acrescenta uma terceira possibilidade às duas possibilidades acima: *retas coincidentes*. Mas este terceiro caso está incluído na relação de equivalência *paralelismo*. Confira *relação de equivalência*.

Se o espaço tiver dimensão maior do que dois, então estas duas possibilidades acima são insuficientes para classificar todas as retas do espaço:

- as retas são paralelas;

- as retas são concorrentes;
- as retas são reversas, quando elas não determinarem um plano.

Duas retas paralelas, determinam um plano, e ainda as chamamos de coplanares. Duas retas concorrentes também são coplanares e determinam um ponto. Dadas duas retas num espaço de dimensão maior do que dois então precisamos destas três possibilidades para classificar as possibilidades entre elas.

É interessante observar a semelhança com os planos e a questão da dimensão. Seria uma especie de classificação dos planos no espaço de dimensão três ou maior. Num espaço de dimensão 3 somente é possível haver duas classes de planos

- os planos são paralelos;
- os planos são concorrentes;

Não é habitual usar-se esta linguagem *planos concorrentes*. Se dois planos não forem paralelos eles determinam um espaço de dimensão imediatamente inferior: uma reta.

Em dimensão quatro, ou maior, se repetem, para os planos, as possibilidades mencionadas para as retas num espaço de dimensão maior do 2. Num espaço de dimensão quatro, ou maior:

- os planos são paralelos;
- os planos são concorrentes;
- os planos são reversos.

Também não é habitual usar-se esta linguagem, *planos reversos*.

A teoria da dimensão, em Álgebra Linear, e na verdade a *teoria das equações lineares*, descreve de forma precisa esta questão dos objetos reversos, paralelos ou coincidentes. Planos, retas, pontos são apenas *espaços* de dimensão dois, um ou zero. A interseção de dois espaços produz um espaço de dimensão menor ou igual aos dois que se interceptam podendo ser o vazio também que é o caso das retas reversas.

---

- **Riemann, integrável a.** Confirma *integrável a Riemann*. É uma condição para que uma função seja integrável. Confirma também *integrável a Lebesgue*

---

- **Riesz, representação de** Se  $(E, \langle \rangle)$  for um espaço de Hilbert então qualquer que seja o funcional linear contínuo  $\Phi$  (elemento do dual) de  $E$  é dado pela expressão  $\Phi(f) = \langle f, g \rangle; g \in E$ .

Uma consequência disto é que os espaços de Hilbert são reflexivos:  $E' = E$  e também  $E'' = E$ . Esta propriedade torna os espaços de Hilbert muito atraentes, um exemplo simples mostra que há uma distância considerável já com os espaços de Banach: nos espaços de sucessões temos  $l^{p'} = l^q; \frac{1}{p} + \frac{1}{q} = 1$ . O unico espaço reflexivo de sucessões é  $l^2$ .

---

- **Roger, Leonard James**, (1862-1933), matemático inglês que havia enunciado e provado praticamente todas as célebres equações de Ramanujan em 1894. Ao morrer, Roger foi agraciado com um elogio fúnebre publicado na *Nature* em 1933, que em resumo diz “*Roger havia escrito uma matemática muito precisa mas, infelizmente, completamente*<sup>65</sup> *inútil.* George E. Andrews relata isto num artigo sobre *séries q* que é o tópico a que pertencem as chamadas equações de Roger-Ramanujan, observando que em 1979, portanto 85 anos após a descoberta por Roger das séries-*q*, ditas de Ramanujan, elas encontraram aplicações na mecânica estatística no chamado *modelo hexagonal rígido, hard-hexagonal model*, a Matemática inútil de Roger.

---

- **salto** Se diz, no caso de uma função de variável real, que seja descontínua no ponto  $x = a$ , é a diferença entre os limites laterais  $f(a^+) - f(a^-)$ .

---

- **salto de primeira espécie** em oposição à *salto de segunda espécie*. Num *salto de primeira espécie* existem os limites laterais  $f(a^+)$ ,  $f(a^-)$  e portanto se pode calcular o salto como  $f(a^+) - f(a^-)$ . Quando não houver um dos limites laterais  $f(a^+)$ ,  $f(a^-)$  tem-se um *salto de segunda espécie*.

---

- **série**

- **série**, a definição
- **série de Fourier**
- **série trigonométrica** é a série de Fourier.
- **série trigonométrica** absolutamente convergente, o conjunto de tais séries forma um *espaço de Banach* denominado  $A(T)$  em que  $T$  representa a fronteira do disco unitário do plano complexo e é chamado de “álgebra de Wiener”. A discussão sobre a denominação “álgebra” abre caminho para um outro verbete... está ligada com a transformação do produto de convolução,

$$\widehat{f * g} = \hat{f}\hat{g}; \widehat{f * g}(n) = \hat{f}(n)\hat{g}(n) \quad (1086)$$

num produto ponto-a-ponto. À esquerda, na equação (1086), estou transformando o produto de convolução de duas funções usando a *transformada de Fourier* cujo resultado é uma sucessão (dos coeficientes de Fourier) e à direita aparece o produto ponto-a-ponto das sucessões  $\hat{f}$ ,  $\hat{g}$  dos coeficientes de Fourier de  $f$  e de  $g$ . Se  $f, g \in \mathcal{L}^2(T) = A(T)$  vale

$$\|\widehat{f * g}\|_2 = \|\hat{f}\|_2 \|\hat{g}\|_2 \quad (1087)$$

---

<sup>65</sup>... though everything he wrote was marked by a certain distinction nothing else of first-rate importance was discovered.

estabelecendo uma isometria entre  $\mathcal{L}^2(T)$  e  $l_2(\mathbf{N})$ . Isto é um resumo do capítulo 4 do livro de análise complexa de Rudin. Neste contexto a transformada de Fourier recebe também o nome de transformada de Gelfand. Em  $A(T)$  temos adição e o produto por convolução que compõem as operações da álgebra de Wiener mas será preciso considerar as distribuições para que se tenha uma unidade nesta álgebra. Aqui surge um problema clássico da transformada de Fourier:  $\hat{\delta}_0 = 1 \notin l_2(\mathbf{N})$ , e  $l_2(\mathbf{N})$  também é uma álgebra sem unidade.

Uma série é um tipo de sucessão,  $S_n$  cujo termo geral se expressa com uma soma:

$$S_n = \sum_{k=0}^n a_k; \quad (1088)$$

$$S = \lim_n S_n = \sum_{k=0}^{\infty} a_k; \quad (1089)$$

A expressão de uma soma com um número infinito de termos, na equação (1089), é apenas um símbolo, o símbolo do limite desta sucessão, que é muito prático porque *simula uma soma* o que nos permite descrever propriedades, teoremas, sobre uma determinada série de forma bem compacta. Por exemplo se

$$S_n = \sum_{k=0}^n a_k; T_n = \sum_{k=0}^n b_k; \quad (1090)$$

$$S = \sum_{k=0}^{\infty} a_k; T = \sum_{k=0}^{\infty} b_k; \quad (1091)$$

$$S + T = \sum_{k=0}^{\infty} a_k + b_k; \quad (1092)$$

A expressão na (1091) apenas se refere à propriedade, “a soma dos limites é o limite da soma” se as duas séries envolvidas forem convergentes. É um exemplo de como a notação, apesar de envolver uma expressão aritmética impossível, se salvou dentro do crivo de perfeição matemática do século 20, como muitas outras notações igualmente críticas (porém magníficas) como esta.

O índice inicial da soma, nem sempre pode ser o zero, ou algum dos primeiros números naturais, e isto cria uma problema para a descrição geral de uma série, mas nas séries, assim como nas sucessões, nos interessa é o *comportamento assintótico* das mesmas e não valores particulares de termos, entendendo assim, salvamos a descrição geral acima.

Algumas vezes se usa dizer que  $\sum_{k=0}^{\infty} a_k$  é a série de termo geral  $a_k$ .

Uma série, como  $\sum_{k=0}^{\infty} a_k$ , pode ser vista como uma integral e a teoria da integração se aplica ao estudo das séries.

- **série de Fourier** Há duas formas das *séries de Fourier*, a complexa e a real. Aqui estou tratando da formulação real da série de Fourier.

Com a série de Fourier real podem ser aproximadas ondas de período arbitrário, pelas ondas básicas  $\sin(x)$ ,  $\cos(x)$  desde que devidamente transformadas por alteração dos seus parâmetros para corrigir o período isto porque as séries de Fourier servem para codificar funções periódicas. Identificado um período  $T$  é relativamente fácil traduzir toda a teoria escrita para o caso  $2\pi$  para o valor adequado do período.

Como esta correção do período é uma operação elementar, vou me fixar aqui no caso genérico em que se usam as ondas básicas convencionais  $\sin(x)$ ,  $\cos(x)$ . Esta forma de escrever simplifica a exposição mas esconde algumas propriedades, entretanto o interessado por recuperar a informação dentro de um texto específico sobre o assunto. Neste caso

$$S_n(x) = a_0 + \sum_{k=1}^n a_k \cos(kx) + b_k \sin(kx) \quad (1093)$$

é a reduzida de ordem  $n$  da série de Fourier que produz uma função (diferenciável) e periódico com período  $2\pi$ .

O espaço gerado pelas funções

$$x \mapsto \cos(kx); k \geq 0; x \mapsto \sin(kx); k \geq 1;$$

considerada a *norma oriunda* do produto escalar

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(x)g(x)dx \quad (1094)$$

é um espaço de Hilbert que contém elementos que não são funções diferenciáveis (de forma muito semelhante com o que acontece com os números irracionais relativamente ao conjunto  $\mathbf{R}$ , com que se obtém a completitude de  $\mathbf{Q}$  que é  $\mathbf{R}$ ). Este espaço completo é denominado  $\mathcal{L}^2([-\pi, \pi])$ . Isto é uma questão descoberta por *Paul Du Bois-Reymond*, em 1873, que descobriu uma falha na teoria construída por Fourier: *era preciso definir corretamente de que forma as séries trigonométricas convergiam*. Até então se pensava que era uma *simples convergência ponto-a-ponto*. Na verdade este pode ser um dos pontos iniciais para a construções dos diversos espaços de função cada um deles com uma topologia adequada.

Os coeficientes do polinômio trigonométrico, equação (1093), são as projeções de uma onda  $f$  na direção das ondas básicas como a Álgebra Linear explica. O coeficiente  $a_0$  tem a aparência de um caso particular, mas uma análise mais cuidadosa do mesmo mostra que ele resolve o caso das “*ondas não contínuas*”, quando surge alguma perturbação, ele memoriza o valor médio nas perturbações, e como equação (1093) define uma função de classe  $\mathcal{C}^\infty$  se produz assim uma regularização da perturbação com este coeficiente. Alguns autores evitam esta discussão definindo a equação (1093), usando  $\frac{a_0}{2}$ .

Esta descrição ultrapassa muito a visão que Fourier tinha em 1822 quando apresentou a sua monografia *Théorie analytique de la chaleur* à Academia Francesa de Ciências que foi recebida com reservas<sup>66</sup>... as consequências deste trabalho de Fourier foram muito profundas em Matemática criando ou alterando significativamente diversos setores não menos com o debate que se seguiu à publicação do seu trabalho. Fourier entendeu e incorporou em um trabalho, tudo que já vinha sendo escrito e usado por matemáticos como Euler e alguns dos Bernoulli que usaram somas de senos e cossenos para encontrar soluções de equações diferenciais.

Seria injusto não computar como consequência das séries de Fourier a invenção da década de 80 das *wavelets* que, de uma certa forma, generalizaram a conceituação das séries de Fourier liberalizando-a de uma onda particular além de considerar duas operações chamadas *translações* e *dilações* para descobrir a presença de modificações de uma certa *onda mãe* dentro de um sinal que é o que fazem as séries de Fourier relativamente à *onda mãe seno*, porque *coseno* é apenas uma translação do seno.

O polinômio trigonométrico, na equação (1094), coincide com a convolução  $f * D_n$  em que  $D_n$  é o núcleo de Dirichlet de ordem  $n$ , no seu intervalo de expansão, confira em *convolução*. Confira também as figuras (fig. 23), página 67, (fig 24), página 68.

- **sim(n)** abreviação de “simétrico n” designa o conjunto das permutações de  $n$  elementos. As permutações podem ser entendidas como funções bijetivas de um conjunto com  $n$  elementos nele mesmo e então, como operação de composição de funções se tem um grupo não comutativo quando  $n \geq 3$ . A notação de ciclos é muito prática para lidar com estes grupos. Por exemplo,  $sim(3)$  seria

$$sim(3) = \{I, (12), (13), (23), (123), (132)\} \quad (1095)$$

Entenda:

$$(12) = \begin{pmatrix} 1 & \mapsto & 2 \\ 2 & \mapsto & 1 \\ 3 & \mapsto & 3 \end{pmatrix} \quad (1096)$$

a função que tem o ponto fixo 3 permutando entre si  $\{1, 2\}$ , da mesma forma os dois outros 2-ciclos,  $(13), (23)$  têm 2, 1 como pontos fixos respectivamente. Em  $(123)$  costumamos ler “1 vai em 2, 2 vai em 3 e 3 vai em 1. Isto significa que  $(123) = (231) = (312)$  havendo o costume de sempre preferir a expressão começando com o “menor elemento” quando houve uma ordem possível entre os elementos que são permutados.

Os grupos de permutações são muito importantes do ponto de vista didático porque facilmente podemos construir grupos não comutativos com uma grande quantidade de sub-grupos.  $sim(4)$ , por exemplo, tem subgrupos de ordem 2,4,6,8,12

<sup>66</sup>Era um resultado muito avançado para sua época, seria preciso uma nova teoria da integração, que surgiu em 1900 com Lebesgue, para que se pudesse reescrever “corretamente” a teoria de Fourier.

isto permite construir exemplos simples das relações entre os subgrupos, exemplificar subgrupos normais e grupos quocientes. Uma aplicação importante destes grupos se encontra na descrição dos determinantes das matrizes quadradas, no grupo das transformações lineares e na teoria da representação linear. Confira  $sim(4)$ .

A tabela seguinte é da adição módulo 24 e como qualquer tabela de grupo traz em sua linhas (ou colunas) as permutações dos elementos do grupo, então a tabela abaixo apresenta todas as *permutações circulares* dos 24 elementos do

conjunto  $\{0, 1, 2, \dots, 23\}$ .

+	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
0	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0
2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1
3	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2
4	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3
5	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4
6	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5
7	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6
8	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7
9	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8
10	10	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9
11	11	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10
12	12	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11
13	13	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12
14	14	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13
15	15	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
16	16	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
17	17	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
18	18	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
19	19	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
20	20	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
21	21	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
22	22	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
23	23	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22

Usando um editor de textos e um programa em python substitui os elementos do conjunto  $\{0, \dots, 23\}$ , nesta ordem, pelas 24 primeiras letras do alfabeto para ob-

ter a tabela

I	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y
b	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	a
c	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	a	b
d	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	a	b	c
e	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	a	b	c	d
f	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	a	b	c	d	e
g	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	a	b	c	d	e	f
h	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	a	b	c	d	e	f	g
i	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	a	b	c	d	e	f	g	h
j	j	k	l	m	n	o	p	q	r	s	t	u	v	x	y	a	b	c	d	e	f	g	h	i
k	k	l	m	n	o	p	q	r	s	t	u	v	x	y	a	b	c	d	e	f	g	h	i	j
l	l	m	n	o	p	q	r	s	t	u	v	x	y	a	b	c	d	e	f	g	h	i	j	k
m	m	n	o	p	q	r	s	t	u	v	x	y	a	b	c	d	e	f	g	h	i	j	k	l
n	n	o	p	q	r	s	t	u	v	x	y	a	b	c	d	e	f	g	h	i	j	k	l	m
o	o	p	q	r	s	t	u	v	x	y	a	b	c	d	e	f	g	h	i	j	k	l	m	n
p	p	q	r	s	t	u	v	x	y	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
q	q	r	s	t	u	v	x	y	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p
r	r	s	t	u	v	x	y	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q
s	s	t	u	v	x	y	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r
t	t	u	v	x	y	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s
u	u	v	x	y	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t
v	v	x	y	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u
x	x	y	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v
y	y	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	x

e portanto todas as permutações circulares do conjunto  $\{a, b, c, \dots, x, y\}$ . Observe bem, *permutações circulares*, porque *todas as permutações* dos elementos dum conjunto com 24 elementos formaria um conjunto com

$$24! = 620448401733239439360000$$

elementos, impossível de colocar numa página de texto.

- **sistema linear** é um *sistema de equações lineares*, ou ainda, duas ou mais



equações das quais se pretenda uma solução comum. Por exemplo

$$\begin{cases} 3x + 2y = 4 \\ 2x - 3y = 5 \end{cases} \quad (1097)$$

é um sistema de duas equações lineares a duas incógnitas. A teoria dos *sistemas lineares* está associadas a teoria das matrizes porque um sistema como da equação (eq. 1097) é equivalente a uma equação matricial

$$\begin{pmatrix} 3 & 2 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}; \quad (1098)$$

$$\mathcal{M}X = A \quad (1099)$$

em que na equação (eq. 1099) eu resumi a equação (eq. 1098) com os símbolos  $\mathcal{M}, X, A$  para representar as matrizes e o sistema como um produto de matrizes.

O sistema de equações lineares (eq. 1097) terá solução, e uma única solução, se a matriz  $\mathcal{M}$  for inversível quando a solução do sistema será dado pelo produto de matrizes

$$X = \mathcal{M}^{-1}A; \quad (1100)$$

em que  $\mathcal{M}^{-1}$  é a inversa da matriz  $\mathcal{M}$ .

Uma das formas de verificar se uma matriz é inversível é o cálculo do seu determinante, que sendo diferente de zero equivale a que a matriz seja inversível. Se o determinante for zero, ainda assim é possível que o sistema de equações tenha solução, mas não será única e já não pode mais ser expresso numa equação como (eq. 1100).

Um exemplo simples pode ser apresentando com a geometria, o sistema de equações

$$\begin{cases} 3x + 2y + 4z = 4 \\ 2x - 3y - 5z = 5 \end{cases} \quad (1101)$$

representa, no  $\mathbf{R}^3$ , dois planos e a Geometria Analítica nos ensina que estes dois planos são perpendiculares, respectivamente, aos vetores

$$\vec{u} = (3, 2, 4); \vec{v} = (2, -3, -5); \langle u, v \rangle = -20 \quad (1102)$$

cujo produto escalar sendo  $-20$ , diferente de zero, implica que  $\vec{u}, \vec{v}$  não são colineares e logo os planos a eles perpendiculares se cortam segundo uma reta que é a solução do sistema de equações lineares da equação (eq. 1101). Este sistema, descrito matricialmente, é

$$\begin{pmatrix} 3 & 2 & 4 \\ 2 & -3 & -5 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}; \quad (1103)$$

quando nos referimos à matriz  $2 \times 3$

$$\begin{pmatrix} 3 & 2 & 4 \\ 2 & -3 & -5 \end{pmatrix} \quad (1104)$$

como a *matriz do sistema*,  $M$ , que não sendo uma matriz quadrada, não pode ser inversível porque o seu determinante é necessariamente nulo, mas o sistema tem solução, que é qualquer dos pontos que pertença a reta determinada pela interseção dos dois planos, uma infinidade de soluções.

A teoria dos espaços vetoriais e dos seus *morfismos* que são as transformações lineares, se refere ao núcleo do morfismo  $M$  que é um subespaço do espaço de saída definido como a solução do sistema homogêneo de equações obtido quando se anula a matriz de dados que se encontra no segundo membro.

A reta, que é solução do sistema na equação (eq. 1103), é uma translação do núcleo determinada por uma solução particular da equação (eq. 1103), ou ainda, é uma translação do núcleo por uma solução particular.

O núcleo é uma reta paralela à solução da equação (eq. 1103), passando na origem.

Este exemplo descreve exatamente como se resolve qualquer sistema de equações lineares, primeiro determinando o *núcleo do operador linear* que o define depois transladando-o por uma *solução particular* como “*coeficiente de translação*”. A figura (fig 142), página 368,

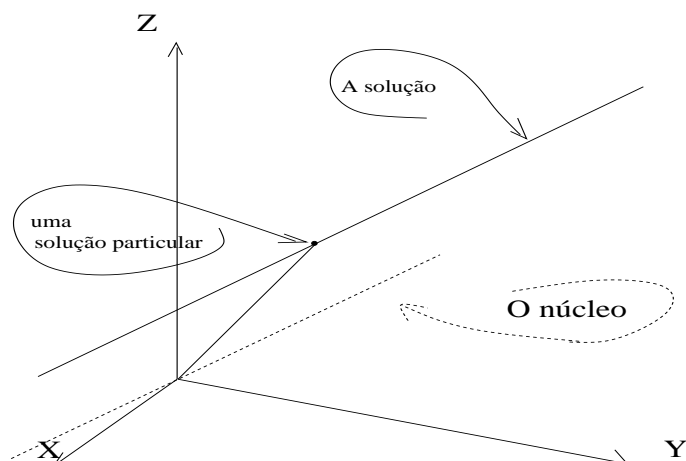


Figura 142: translação do núcleo com solução particular

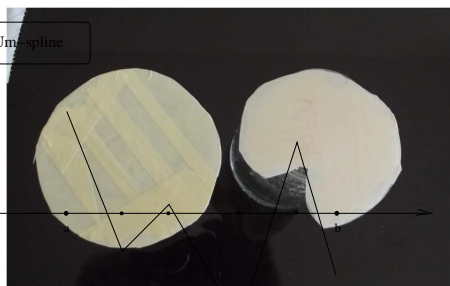
Apenas para mostrar a amplitude das aplicações desta teoria, o parágrafo anterior descreve de maneira precisa qual é a solução duma equação diferencial linear de ordem qualquer . . . *descreve* é o que foi dito, outra coisa é *apresentar explicitamente* as soluções.

A determinação do núcleo dum operador é no fundo equivalente à discussão sobre determinantes mencionada anteriormente. Entretanto os determinantes estão restritos aos sistemas de equações lineares definidos em espaços de dimensão finita, enquanto que os núcleos de operadores estão livres da questão dimensional e portanto, num certo sentido, generalizam o conceito de determinante.

- somável

- absolutamente somável
- convergência e comutatividade

- splines Um *splines* é uma função definida por pedaços que tem uma regularidade condicionada ao seu grau. Dizemos *n*-spline quando os pedaços são polinômios de grau menor ou igual a *n* e a classe de continuidade é *n* - 1. Como exemplo, a figura (fig. 144), página 369, mostra um



Um spline

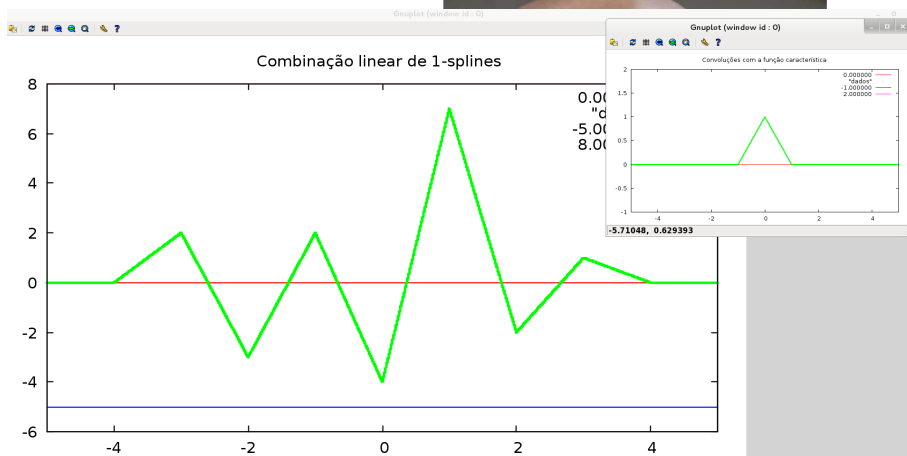


Figura 144: 1-splines - combinação linear de *T*

“1-splines”, é uma função contínua, de classe  $C^0$ , cujos pedaços são do primeiro grau (de grau menor ou igual a 1). Se você definir num terminal do gnuplot a função *T*, cujo gráfico aparece num detalhe da figura (fig. 144), com a equação:

$$T(x) = (x < -1) ? 0 : (x < 0) ? x + 1 : (x < 1) ? 1 - x : 0$$

o comando do gnuplot

```
plot 2*T(x+3)-3*T(x+2)+2*T(x+1)-4*T(x)+7*T(x-1)-2*T(x-2)+T(x-3)
```

irá produzir o gráfico principal da figura (fig. 144) e você pode selecionar os coeficientes que bem desejar para obter outras *poligonais* que são exemplos de

1-splines. Neste exemplo os coeficientes são: 2, -3, 2, -4, 7, -2, 1. Observe que, associado a cada coeficiente há um “*fator de translação*” que são, respectivamente, -3, -2, -1, 0, 1, 2, 3.

Como outro exemplo a figura (145), página 370, mostra um 29-spline junto

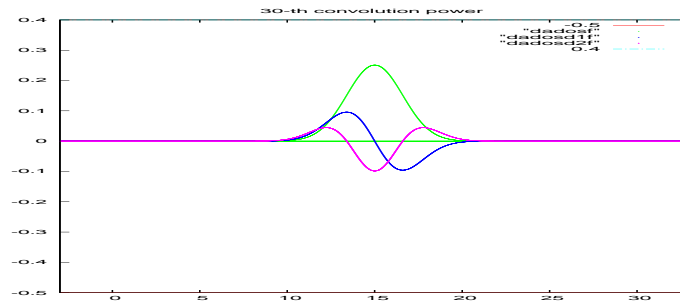


Figura 145: 29-spline e suas derivadas primeira e segunda

com suas derivadas primeira e segunda que então serão, respectivamente, um 28-spline e um 27-spline. Aqui os exemplos são de splines univariados e da mesma forma os podemos ter bivariados ou multivariados. Na linguagem da teoria de aproximação a que basicamente pertencem os splines, eles são pedaços de polinômios soldados nos pontos extremos de cada um dos segmentos, estes extremos são chamados de *nós*, e também se chamam de *nós* as projeções destes pontos no domínio do spline. Como a condição de diferenciabilidade é crucial sobre os *nós*, de uma certa forma podemos ver os splines como uma generalização dos polinômios de Taylor e muito provavelmente eles tenham surgido inspirados na fórmula de Taylor. A construção de splines a partir desta definição é penosa e existe um processo para obtê-los por convolução, mas de qualquer forma eles são uma ferramenta matemática para ser utilizada em programas de computador. A figura (146), página 371, mostra uma função e sua aproximação com um 5-spline no intervalo  $[-5, 5]$  em que foram usado 5 nós, portanto uma péssima aproximação para que você possa observar que existe uma aproximação...um erro! Se fosse usada uma precisão maior, 100 nós, (nós distanciados de  $\frac{1}{10}$ ), não seria possível ver, no gráfico, diferença entre o “modelo” e a “realidade”. Se você for crítico deve estar considerando estranho falar de “aproximação” de uma função tão bem conhecida como a função polinomial  $y = f(x) = (x + 5) * (x + 1) * (x - 1) * (x - 5)$ . Qual seria a razão de dar um exemplo deste? A resposta é simples, como podemos fazer o gráfico tanto de  $f$  como da sua aproximação com splines, então é possível compreender melhor como funciona o método de aproximação que está sendo estudado ou desenvolvido. Se funcionar bem em casos conhecidos então é muito provável que venha a funcionar para *interpolar* dados obtidos por um processo experimental criando um modelo computacional para representar o fenômeno medido pela experiência.

Uma outra forma de definir  $n$ -splines é feita associando-se a uma partição de um intervalo  $[a, b]$

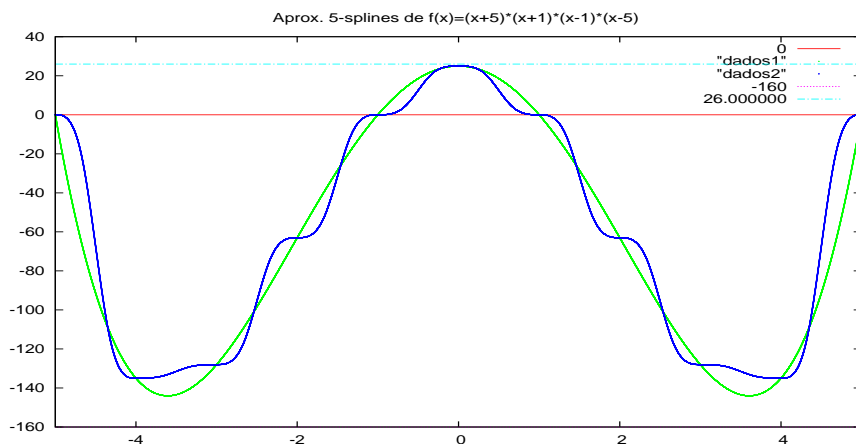


Figura 146: aproximação 5-splines de  $f(x) = (x + 5) * (x + 1) * (x - 1) * (x - 5)$

**Definição 23** (1 – splines) *Seja a partição do intervalo  $[a, b]$  definida pelos nós*

$$a = x_0, \dots, x_k, \dots, x_n = b \quad (1105)$$

*um conjunto de pontos do plano*

$$(x_0, y_0), \dots, (x_k, y_k), \dots, (x_n, y_n) \quad (1106)$$

*e a função  $f$ ;  $f|_{[x_k, x_{k+1}]} = f_k$  em que a função  $f_k$  é o polinômio de grau menor ou igual 1 cujo gráfico liga os pontos  $(x_k, y_k), (x_{k+1}, y_{k+1})$ .*

*Então o gráfico de  $f$  é uma poligonal, um 1 – splines.*

*Qualquer primitiva de  $f$  é um 2 – splines e assim as sucessivas primitivas serão 3 – splines ...  $n$  – splines.*

O gráfico na (fig. 144), página 369, foi obtido usando-se esta definição junto com um programa escrito na linguagem do `gnuplot`.

Mas esta definição é difícil de ser usada, e por isto cada autora tenta encontrar uma que se adeque melhor ao seu trabalho. Para algumas autoras *splines* é feminino, para outras é um *splines*. Ninguém sabe ao certo quando foram inventados e nem quem os inventou, simplesmente apareceram! Dizem que Courant fazia uso deles, e nunca encontrei nenhum texto do Courant mencionando *splines*, o que também não significa que não exista, mas diversos autores afirmam que ele teria usado *splines* e este fato coloca o aparecimento dos  $n$ -splines na década de 40 do século passado. As funções definidas por pedaços são antigas, remontam ao século 19.

Se  $y = f(x)$  for um  $n$  – splines então  $\frac{d^{(n-1)}f}{dx}$  será uma poligonal, uma função contínua, e  $\frac{d^{(n)}f}{dx}$  será uma *distribuição*, uma medida, combinação linear

de translações da *medida de Dirac*. Usando a definição 23, página 370,

$$\frac{d^{(n)}f}{dx} = \sum_{k=0}^{n-1} m_k \delta_{x_k} \quad (1107)$$

em que  $m_k$  é o coeficiente angular das retas que definem o 1-splines. Ou seja,  $f$  é de classe  $\mathcal{C}^{(n-1)}$ .

Se fala de splines *não polinomiais* em que os segmentos não precisariam ser algébricos, poderiam ser trigonométricos, por exemplo.

Aplicações de splines se encontram hoje em *desenho assistido por computador, modelagem em cima de dados discretos, soluções aproximadas de equações diferenciais, como ferramenta auxiliar na construção de wavelets*.

Quando se passa ao caso multivariado há uma conexão com *elementos finitos* que é um caso particular de spline: 1-spline. Como para os elementos finitos se desenvolveu uma ferramenta chamada *triangulação* que consiste em particionar o domínio com triângulos, na literatura se fala de spline multivariado como “funções polinômiais por pedaços definidas sobre domínios poligonais”.

- **Stone-Weierstrass** Este teorema estabelece a densidade de certas classes de funções dentro do espaço das funções contínuas, indiretamente. O conjunto das funções contínuas definidas num intervalo fechado da reta é um espaço vetorial. Seja  $[a, b]$  um intervalo da reta e consideremos o *espaço vetorial*  $(\mathcal{C}([a, b]), +, \cdot)$  em que a soma é a operação definida ponto a ponto e a multiplicação é o *produto por um escalar*.

Observando que os *escalares* são um tipo particular de função vemos que o este produto pode ser generalizado com a definição do produto ponto a ponto

$$f, g \in \mathcal{C}([a, b]); h = fg; x \mapsto f(x)g(x) = h(x); \quad (1108)$$

e agora temos duas operações internas no conjunto  $\mathcal{C}([a, b])$  uma vez que soma de funções contínuas é uma função contínua valendo o mesmo para o produto ponto a ponto.

Esta estrutura algébrica é um *anel com divisores de zero*: considere duas funções,  $f, g$ , com *suportes disjuntos* então  $fg$  é a função *identicamente nula*.

Mas como não perdemos o espaço vetorial original podemos nele definir uma *norma*. O fato de que o intervalo seja fechado e limitado, um conjunto compacto é crucial neste próximo passo, toda função em  $\mathcal{C}([a, b])$  tem máximo (e mínimo), e se  $f$  for contínua então a função  $|f|$  também é contínua e podemos assim definir

$$\|f\|_{\infty} = \sup_{x \in [a, b]} |f(x)| \quad (1109)$$

a chamada *norma do supremo*. Com isto temos um espaço vetorial normado que também é um anel e a norma e o produto satisfazem à desigualdade

$$\|fg\|_{\infty} \leq \|f\|_{\infty} \|g\|_{\infty} \quad (1110)$$

completando a definição duma *álgebra comutativa*, um espaço vetorial normado que é também um anel comutativo.

Há várias estruturas interessantes que “aparecem” dentro desta álgebra e que foram estudadas paralelamente, por exemplo, se  $[a, b] = [0, 2\pi]$  um polinômio trigonométrico é um exemplo de função contínua e aqui surgiu um problema interessante de convergência da *série de Fourier*. Converge? A resposta é sim, mas para um ponto externo ao espaço  $\mathcal{C}([a, b])$ , em geral... e foi isto que mostrou que a integral de Riemann era insuficiente também conduzindo à criação dos espaços de Lebesgue para conter os limites das séries de Fourier.

Este exemplo mostra que é preciso entender melhor esta álgebra e a figura (fig 147), página 373, mostra um caminho construtivo para atingir este obje-

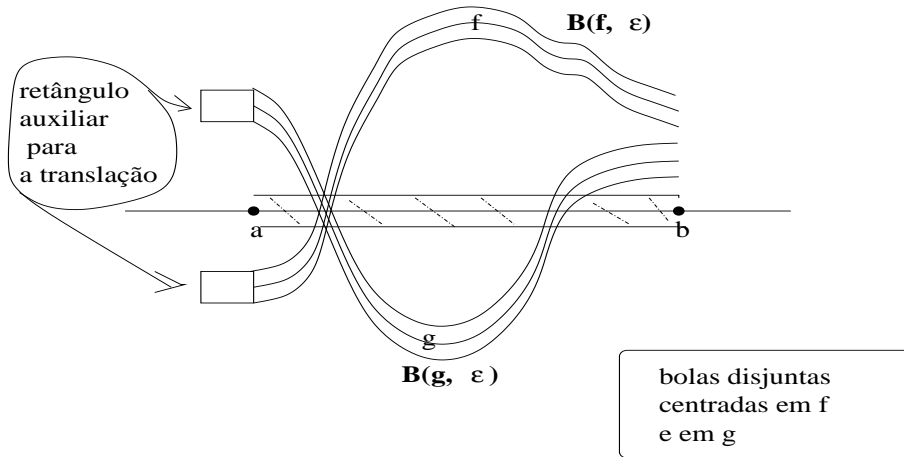


Figura 147: Vizinhança de zero em  $\mathcal{C}([a, b])$

tivo. Nela você pode ver uma vizinhança de zero, um tubo em volta da função constante zero, e as somas desta vizinhança produzindo vizinhanças de  $f$  e de  $g$ . Se duas funções contínuas forem diferentes é possível calcular a “distância entre os seus gráficos” e assim produzir uma vizinhanças de zero que transladada por  $f$  ou por  $g$ , produz dois abertos disjuntos, como estão mostrados na figura (fig 147). Isto mostra que  $\mathcal{C}([a, b])$  é um *espaço de Hausdorff*, um espaço topológico em que os pontos podem ser separados por abertos.

- **superfície** é a imagem de uma função da forma

$$\alpha(s, t) = (x_1(s, t), \dots, x_n(s, t)); n > 2; \quad (1111)$$

em que  $n$  é a dimensão do espaço onde reside o gráfico da superfície  $\alpha$ , e é comum chamar-se o gráfico de traço da superfície para tornar independentes os dois conceitos, a função  $\alpha$  e o gráfico desta função que é o objeto geométrico “*superfície*”, uma variedade de dimensão dois..

Um exemplo simples de superfície, ou variedade de dimensão dois, é o gráfico de uma função diferenciável,  $F$ , de duas variáveis:

$$[a, b] \times [c, d] \xrightarrow{F} \mathbf{R}; (x, y) \mapsto F(x, y) \in \mathbf{R} \quad (1112)$$

e podemos adaptar a notação da equação (eq. 1112) ao formato da equação (eq. 1111) escrevendo

$$(s, t) \mapsto (x_1(s, t), x_2(s, t)) = ((s, t), F(s, t)); \quad (1113)$$

`gnuplot` faz gráficos de superfícies definidas no formato da equação (eq. 1112) com o comando `splot`. Por exemplo, raspando e colando num terminal do `gnuplot`

```
pow(x,n) = x**n;
F(x,y) = pow(x,2) + 3*x*y + pow(y,3);
splot F(x,y), 0
```

vai produzir o gráfico da superfície  $graf(F)$  quando

$$z = F(x, y) = x^2 + 3xy + y^3;$$

- **sucessões** são funções definidas no conjunto  $\mathbf{N}$  dos números naturais e tomando valores num outro espaço que caracteriza o tipo de sucessão. Exemplo, se o conjunto de chegada for  $\mathbf{Q}$ , o conjunto dos números racionais, dizemos que se trata duma sucessão numérica, ou uma sucessão de números racionais.

Associado às sucessões existe um funcional linear, *limite*, que nos permite classificar as sucessões em duas grandes classes, *convergentes* e *divergentes*.

As sucessões representam resultados de experimentos e o comportamento da sucessão assim resultante caracteriza o experimento como um *sucesso*, se a sucessão for convergente, ou um *insucesso* no caso contrário.

Por exemplo, o conjunto dos números racionais é incompleto no sentido de que há sucessões de números racionais que são “convergentes” e assim definem um número, mas este número não é um número racional. Considere a sucessão  $P_n$  dos polígonos regulares inscritos num círculo de raio 1. O quociente do perímetro de  $P_n$  pelo diâmetro do círculo se *aproxima, arbitrariamente*, de um número que os gregos chamaram de  $\pi$  que não é possível escrever como o quociente de dois inteiros e portanto não é um número racional. Desta forma  $\lim_n \frac{P_n}{2} = \pi$  e o operador  $\lim$  fornece um número que completa  $\mathbf{Q}$ . O conjunto de todos os números que completam  $\mathbf{Q}$  é o conjunto  $\mathbf{R}$  dos números reais, e naturalmente  $\mathbf{Q} \subset \mathbf{R}$ .

Você tem aqui um exemplo de experimento, medir o perímetro de polígonos e comparar o resultado com o diâmetro. Os gregos ou possivelmente matemáticos antes deles, observaram que este quociente produzia números muito próximos de 3.14 e extrapolaram este resultado afirmando que no caso do círculo esta razão seria um número que chamaram de  $\pi$ .

A definição rigorosa de limite foi feita por Bolzano, depois redescoberta por Weierstrass:

$$(\forall \epsilon > 0) (\exists N \in \mathbf{N}) |n > N| \implies |s_n - l| < \epsilon \quad (1114)$$

quando dizemos suscintamente que  $s_n \rightarrow l$  ou também

$$\lim_{n \rightarrow \infty} s_n = l \quad (1115)$$



A igualdade na expressão (eq. 1115) apenas quer dizer que  $n$  cresce indefinidamente.

O conjunto dos números reais,  $\mathbf{R}$ , é a completção dos conjunto dos números racionais. Uma das formas de *construir* o conjunto  $\mathbf{R}$  consiste do seguinte *programa*:

- a imersão de  $\mathbf{Q}$  no conjunto das sucessões de números racionais, isto é na identificação de  $\mathbf{Q}$  como um subconjunto do conjunto de todas as sucessões de números racionais,
- restrição do conjunto de todas as sucessões ao conjunto das sucessões que satisfizerem ao critério de Cauchy. Dizemos que uma sucessão  $(s_n)_{n \in \mathbf{N}}$  é *de Cauchy* se for verdadeiro

$$(\forall \epsilon > 0) (\exists N \in \mathbf{N}) n, m > N \implies |s_n - s_m| < \epsilon \quad (1116)$$

Tais sucessões são *intrinsecamente* convergentes uma vez que a partir do índice  $N$  todos os termos da sucessão se encontram dentro da *bola* de centro  $s_n$  e raio  $\epsilon$ .

- Chame de  $\mathcal{C}$  este conjunto das sucessões de Cauchy de números racionais. Podemos mostrar que  $(\mathcal{C}, +, \cdot)$  é um anel que tem divisores de zero, mas podemos eliminar os divisores de zero passando ao quociente com grupo das sucessões que convergem para zero. Como este grupo é um *ideal maximal* de  $\mathcal{C}$  o quociente é um corpo: o corpo  $\mathbf{R}$  dos números reais.

Este programa que é atribuído à Cauchy é bem natural pese que considerado *avançado*, possivelmente pelos aspectos algébricos que envolve. Mas o que se espera dum número real é que ele seja uma sucessão de medidas convergentes. Desta forma  $\mathbf{R}$  nada mais é do que o conjunto dos limites das sucessões de números racionais. Os limites são as etiquetas das classes quocientes mencionadas no *programa* acima.

Considere o experimento mencionado acima da comparação do perímetro de uma sucessão de polígonos com o diâmetro do círculos em que eles estiverem inscritos. Em lugar de polígonos regulares inscritos, podemos considerar uma sucessão de polígonos regulares envolventes a um círculo e o resultado do experimento será equivalente, apenas os quocientes obtido serão maiores que  $\pi$ . Assim você vê dois exemplos de sucessões equivalentes, a *etiqueta* que registra a classe a que elas pertencem é  $\pi$ .

O número  $s_n$ , na equação (eq. 1116) pode ser considerado um valor aproximado do limite da sucessão a menos do erro  $\epsilon$ . Desta forma o critério de Cauchy é um *instrumento prático* para encontrar aproximadamente o limite duma sucessão.

Outros exemplos mais simples de números não racionais são as raízes dos números naturais que, ou são números naturais, ou são número irracionais. O ou é exclusivo. Um algoritmo geométrico que faz esta construção pode ser encontrado em [8, Capítulo 5]

Confira também *limite superior, limite inferior*

- **sucessões de Cauchy** são uma classe de sucessões definidas pelo *critério de Cauchy*. Dizemos que uma sucessão  $(s_n)_{n \in \mathbf{N}}$  é de *Cauchy* se for verdadeiro

$$(\forall \epsilon > 0) (\exists N \in \mathbf{N}) n, m > N \implies |s_n - s_m| < \epsilon \quad (1117)$$

As sucessões de Cauchy são *intrinsecamente* convergentes uma vez que a partir do índice  $N$  todos os termos da sucessão se encontram dentro da *bola* de centro  $s_n$  e raio  $\epsilon$ , confira a figura (fig 148), página 376.

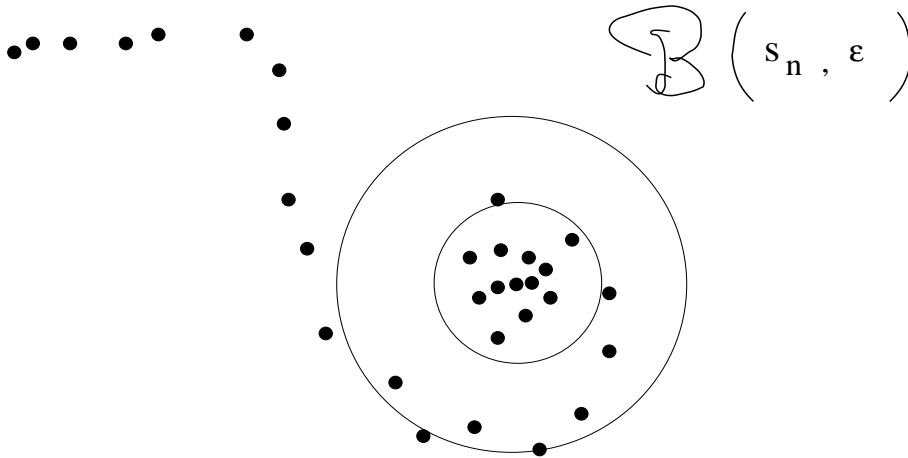


Figura 148: Bola contendo o rabo duma sucessão de Cauchy

Associado às sucessões de Cauchy existe um funcional linear, *limite* que nos permite classificar as sucessões de Cauchy em duas grandes classes, *convergentes* e *divergentes*.

As sucessões de Cauchy representam resultados de experimentos e o comportamento da sucessão assim resultante caracteriza o experimento como um *sucesso*, se a sucessão for convergente, ou um *insucesso* no caso contrário.

Por exemplo, o conjunto dos números racionais é incompleto no sentido de que há sucessões de Cauchy de números racionais que são “convergentes” e assim definem um número, mas este número não é um número racional. Considere a sucessão  $P_n$  dos polígonos regulares inscritos num círculo de raio 1. O quociente do perímetro de  $P_n$  pelo diâmetro do círculo se *aproxima arbitrariamente* de um número que os gregos chamaram de  $\pi$  que não é possível escrever como o quociente de dois inteiros e portanto não é um número racional. Desta forma  $\lim_n \frac{P_n}{2} = \pi$  e o operador  $\lim$  fornece um número que completa  $\mathbf{Q}$ . O conjunto de todos os números que completam  $\mathbf{Q}$  é o conjunto  $\mathbf{R}$  dos números reais, e naturalmente  $\mathbf{Q} \subset \mathbf{R}$ .

Você tem aqui um exemplo de experimento, medir o perímetro de polígonos e comparar o resultado com o diâmetro. Os gregos ou possivelmente matemáticos

antes deles, observaram que este quociente produzia números muito próximos de 3.14 e extrapolaram este resultado afirmando que no caso do círculo esta razão seria um número que chamaram de  $\pi$ .

A definição rigorosa de limite foi feita por Bolzano, depois redescoberta por Weierstrass:

$$(\forall \epsilon > 0) (\exists N \in \mathbf{N}) |n > N| \implies |s_n - l| < \epsilon \quad (1118)$$

quando dizemos sucintamente que  $s_n \rightarrow l$  ou também

$$\lim_{n \rightarrow \infty} s_n = l \quad (1119)$$

A igualdade na expressão (eq. 1119) apenas quer dizer que  $n$  cresce indefinidamente.

O conjunto dos números reais,  $\mathbf{R}$ , é a completção dos conjunto dos números racionais. Uma das formas de *construir* o conjunto  $\mathbf{R}$  consiste do seguinte *programa*:

- a imersão de  $\mathbf{Q}$  no conjunto das sucessões de Cauchy de números racionais, isto é na identificação de  $\mathbf{Q}$  como um subconjunto do conjunto de todas as sucessões de Cauchy de números racionais,
- 
- Chame de  $\mathcal{C}$  este conjunto das sucessões de Cauchy de números racionais. Podemos mostrar que  $(\mathcal{C}, +, \cdot)$  é um anel que tem divisores de zero, mas podemos eliminar os divisores de zero passando ao quociente com grupo das sucessões de Cauchy que convergem para zero. Como este grupo é um *ideal maximal* de  $\mathcal{C}$  o quociente é um corpo: o corpo  $\mathbf{R}$  dos números reais.

Este programa que é atribuído à Cauchy é bem natural pese que considerado *avanzado*, possivelmente pelos aspectos algébricos que envolve. Mas o que se espera dum número real é que ele seja uma sucessão de medidas convergentes. Desta forma  $\mathbf{R}$  nada mais é do que o conjunto dos limites das sucessões de Cauchy de números racionais. Os limites são as etiquetas das classes quocientes mencionadas no *programa* acima.

Considere o experimento mencionado acima da comparação do perímetro de uma sucessão de polígonos com o diâmetro do círculos em que eles estiverem inscritos. Em lugar de polígonos regulares inscritos, podemos considerar uma sucessão de polígonos regulares envolventes a um círculo e o resultado do experimento será equivalente, apenas os quocientes obtido serão maiores que  $\pi$ . Assim você vê dois exemplos de sucessões de Cauchy equivalentes, a *etiqueta* que registra a classe a que elas pertencem é  $\pi$ .

O número  $s_n$ , na equação (eq. 1117) pode ser considerado um valor aproximado do limite da sucessão a menos do erro  $\epsilon$ . Desta forma o critério de Cauchy é um *instrumento prático* para encontrar aproximadamente o limite duma sucessão.

Outros exemplos mais simples de números não racionais são as raízes dos números naturais que, ou são números naturais, ou são número irracionais. O ou é exclusivo. Um algoritmo geométrico que faz esta construção pode ser encontrado em [8, Capítulo 5]

Confira também *limite superior*, *limite inferior*

- **tabela verdade** Uma *tabela de verdade* é um algoritmo que permite comparar, e eventualmente demonstrar, a igualdade entre duas relações.

Dadas duas relações  $\mathcal{A}, \mathcal{B}$  existe, para cada uma delas, dois “*estados possíveis: verdade, falso*” o que produz um arranjo com repetição destes dois valores apresentados na tabela:

	$\mathcal{A}$	$\mathcal{B}$
	V	V
	V	F
	F	V
	F	F

São os arranjos com repetição *2-a-2*, dos dois valores possíveis V, F, ou dois estados possíveis se estas relações puderem ser testadas contra estes dois valores e obviamente que existem relações que não respondem a nenhum destes valores porque não podem ser testadas, mas vou deixar de lado este caso. Considerar esta possibilidade consiste em adotar a chamada *lógica fuzzy*.

A *teoria dos conjuntos* é uma perfeita realização da Lógica Matemática uma vez que podemos expressar qualquer sentença de *lógica* usando uma expressão da *teoria dos conjuntos*. Aqui vou usar a seguinte notação:

$$\mathcal{A}(x) \equiv x \in A; \quad (1120)$$

ou seja a sentença  $\mathcal{A}(x)$  é realizada pela sentença da teoria dos conjuntos  $x \in A$ . Consequentemente terei:

$$\mathcal{A}(x) \text{ e } \mathcal{B}(x) \equiv x \in (A \cap B); \quad (1121)$$

$$x \in (A \cap B) \equiv (x \in A) \text{ e } (x \in B); \quad (1122)$$

$$\mathcal{A}(x) \text{ ou } \mathcal{B}(x) \equiv x \in (A \cup B); \quad (1123)$$

$$x \in (A \cup B) \equiv (x \in A) \text{ ou } (x \in B); \quad (1124)$$

Confira o exemplo analisando a figura (fig 149), página 379, em que você pode ver dois conjuntos  $A, B$  representados como subconjuntos do conjunto  $U$ .

Uma questão interessante, muito comum e relativamente difícil de ser demonstrada, é “*qual seria o resultado da complementação  $(A \cup B)_U^c$ , o complementar de  $(A \cup B)$  em relação ao universo  $U$ , traduzido em termos de  $A$  e de  $B$ .*”

É o mesmo que decidir qual é a negação duma sentença do tipo “ou” em Lógica Matemática: “não ( $A(x)$  ou  $B(x)$ )”.

A resposta pode ser facilmente obtida analisando uma tabela de verdade que envolva também estas relações.

Agora temos mais cinco relações a serem consideradas que nos permite expandir a tabela anterior para

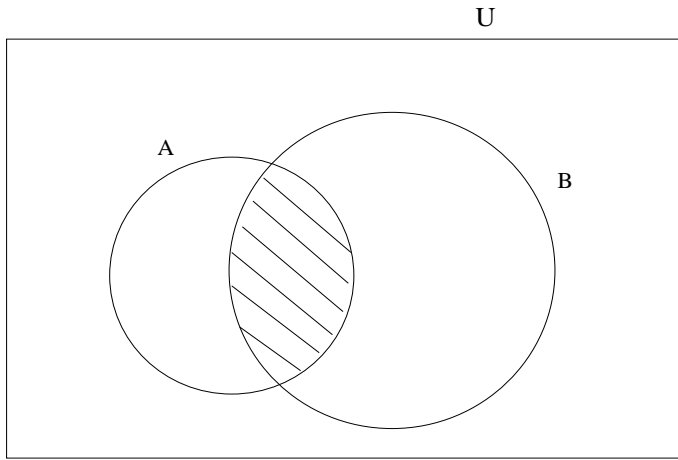


Figura 149:  $x \in A \cup B, x \in A \cap B$

$x \in A$	$x \in B$	$x \in A^c_U$	$x \in B^c_U$	$x \in (A \cup B)$	$x \in (A \cup B)^c_U$	$x \in A^c_U \cap B^c_U$
V	V	F	F	V	F	F
V	F	F	V	V	F	F
F	V	V	F	V	F	F
F	F	V	V	F	V	V

e podemos ver que os arranjos 4-a-4, ao longo das duas últimas colunas são idênticos: FFFV. O significado é que estas duas relações têm os mesmos estados e as podemos considerar idênticas.

Usando a expressão das sentenças de acordo com a convenção que estabeleci,  $\mathcal{A}(x) \equiv x \in A, \mathcal{B}(x) \equiv x \in B$  posso representar a tabela anterior como:

$\mathcal{A}(x)$	$\mathcal{B}(x)$	$\neg\mathcal{A}(x)$	$\neg\mathcal{B}(x)$	$(\mathcal{A}(x) \text{ ou } \mathcal{B}(x))$	$\neg(\mathcal{A}(x) \text{ ou } \mathcal{B}(x))$	$\neg\mathcal{A}(x) \text{ e } \neg\mathcal{B}(x)$
V	V	F	F	V	F	F
V	F	F	V	V	F	F
F	V	V	F	V	F	F
F	F	V	V	F	V	V

em que  $\neg\mathcal{A}(x)$  é a negação de  $\mathcal{A}(x)$ .

Demonstrei então que

$$\neg(\mathcal{A}(x) \text{ ou } \mathcal{B}(x)) = x \in (A \cup B)^c_U = x \in (A^c_U \cap B^c_U) = (\neg\mathcal{A}(x)) \text{ e } (\neg\mathcal{B}(x)); \tag{1125}$$

O complementar da união de dois conjuntos é a interseção dos complementares destes conjuntos. Também mostrei que a negação duma relação do tipo “ou” é uma relação do tipo “e”:

$$\neg(\mathcal{A}(x) \text{ ou } \mathcal{B}(x)) = (\neg\mathcal{A}(x)) \text{ e } (\neg\mathcal{B}(x)); \tag{1126}$$

Ou ainda, *trocando o estado da última equação*, posso concluir que

$$(\mathcal{A}(x) \text{ ou } \mathcal{B}(x)) = \neg((\neg\mathcal{A}(x)) \text{ e } (\neg\mathcal{B}(x))); \quad (1127)$$

que conduz a definição de “ou” em termos de “e”. Assim se chega à conclusão de que as duas operações lógicas básicas são  $\neg$ , e, podendo “e” ser trocado por “ou” para que as operações básicas sejam  $\neg$ , ou, evidentemente.

Os símbolos usados em lógica matemática para as conjunções “e,ou”, são, respectivamente,  $\wedge, \vee$ . Então, eu deveria ter escrito:

$$\mathcal{A}(x) \text{ ou } \mathcal{B}(x) = \mathcal{A}(x) \vee \mathcal{B}(x); \quad (1128)$$

$$\mathcal{A}(x) \text{ e } \mathcal{B}(x) = \mathcal{A}(x) \wedge \mathcal{B}(x); \quad (1129)$$

- **tangente** Duas funções  $f, g$  se dizem tangentes no ponto  $(a, b)$  sse

1.  $f(a) = g(a)$ , ou seja  $(a, b) \in \text{graf}(f) \cap \text{graf}(g)$ ;
2. Existe uma função linear  $K$  do espaço vetorial onde  $f, g$  estiverem definidas, tal que  $f(x) - g(x) = o(K(x - a))$  em que  $o$  é o *pequeno de Landau*, a ordem de grandeza de  $f(x) - g(x)$  é menor do que a ordem de grandeza de  $K(x - a)$  numa vizinhança de  $x = a$  (o limite do quociente pelo módulo de  $x - a$  existe).

Por exemplo, se  $f$  for derivável, então em cada ponto do seu domínio existe uma função linear tangente ao gráfico de  $f$ . A relação de tangência num ponto é uma relação de equivalência e a classe de  $f$ , se existir, é chamada de *germe* de  $f$ .

- **Taylor, polinômio** O polinômio de Taylor de uma função univariada e que tenha derivadas até a ordem  $n$ , conhecidas, num ponto  $x = a$  é a expressão polinômial

$$P(x) = a_0 + a_1(x - a) + a_2(x - a)^2 + \dots + a_n(x - a)^n \quad (1130)$$

com  $a_k = \frac{f^{(k)}(a)}{k!}$ . Os coeficientes são determinados pelo conjunto de equações

$$\begin{cases} P(a) = f(a) & \Rightarrow a_0 = f(a); \\ P'(a) = f'(a) & \Rightarrow a_1 = f'(a); \\ P^{(k)}(a) = f^{(k)}(a) & \Rightarrow a_k = \frac{f^{(k)}(a)}{k!}; \end{cases} \quad (1131)$$

Como  $0! = 1!$  e  $2! = 2$  então esta fórmula pode ser escrita de forma concisa como

$$P(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} \quad (1132)$$

Dois exemplos importantes da fórmula de Taylor, chamadas de McLaurin é quando aplicamos a Fórmula de Taylor ao *seno* ou ao *coseno*. Nós conhecemos

as derivadas de qualquer ordem destas funções em alguns pontos, na origem por exemplo.

As derivadas do *seno* na origem são

$$0, 1, 0, -1, \dots, 0, 1, 0, -1, \dots, \quad (1133)$$

$$dsen(n)(n\%4 == 0)?0 : (n\%4 == 1)?1 : (n\%4 == 2)?0 : -1; \quad (1134)$$

em que foi usado `if-else-compacto`, com a sintaxe da linguagem `C`, e o símbolo `%`, em `C`, é a função congruência módulo-um resto dos inteiros. Na equação (1134), você tem uma função inteira de período 4, então o polinômio de Taylor (ou de McLaurin) do seno é

$$P(x) = \sum_{k=0}^n dsen(k) \frac{f^{(k)}(a)}{k!} \quad (1135)$$

Usando a linguagem `calc`, usualmente distribuída com os sistemas Debian/Gnu/Linux, você pode implementar este algoritmo para obter o seno com alta precisão, porque `calc` é de precisão infinita (inteira) como também o são `Python` e em geral os dialetos da linguagem `LISP`, embora não seja necessário usar polinômios de grau muito alto definindo módulo  $\pi$ , por exemplo, com um polinômio de grau 17.

Na figura (150) página 381, você pode ver o gráfico da função seno, definida

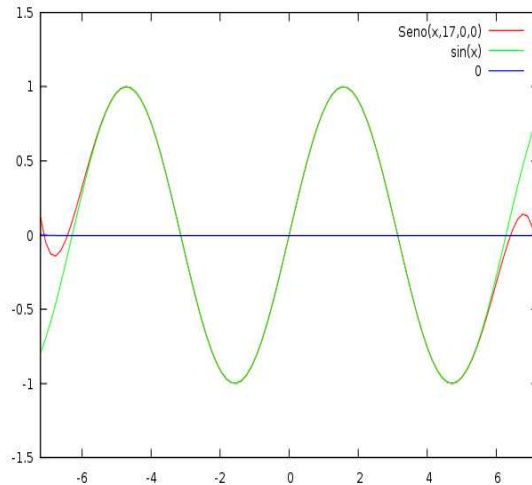


Figura 150: Polinômio de Taylor de grau 17 do seno na origem

algoritmicamente dentro do `gnuplot` e de um polinômio de Taylor de grau 17,

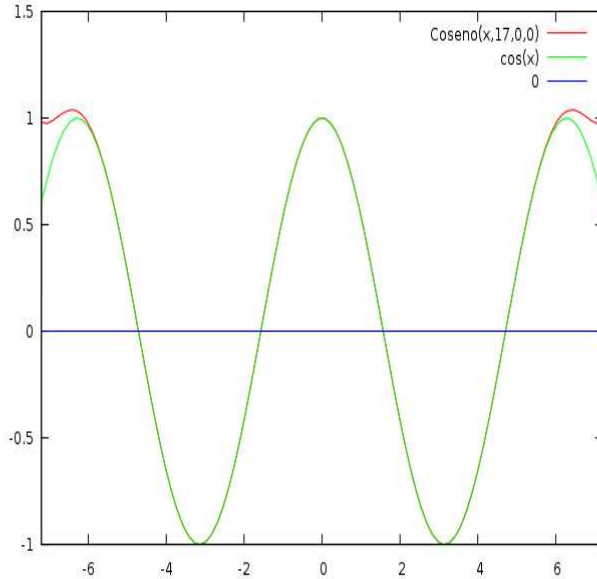


Figura 151: Polinômio de Taylor de grau 17 do cosseno na origem

do seno, no intervalo  $[-6, 6]$ . e na figura (151) página 382, também usando a expressão algorítmica do cosseno de `gnuplot` e do polinômio de Taylor de grau 17, cosseno, no intervalo  $[-6, 6]$ .

- **teorema espectral** É a generalização, na teoria dos operadores, do sistema de valores próprios e vetores próprios da Álgebra Linear no sentido de que um operador,  $T$  definido num espaço medido de funções, tem uma representação sob forma de integral num certo domínio  $\Omega$  pode ser reparametrizado (mudança de variável) para ser representado como uma integral sob um domínio  $spec(T)$  de tal modo que

$$T(f) = \int_{spec(T)} x d\mu(x) \quad (1136)$$

a integral da função identidade deste espaço de funções. Isto é a forma como se consegue colocar uma matriz diagonalizada usando os seus vetores próprios como base para o espaço vetorial. No caso das matrizes aparecem os valores próprios como multiplicadores do vetores próprios que representam a escala do operador na direção de cada um dos vetores próprios, aqui, é a medida do espaço que faz este papel “distribuindo” a massa no espaço. Desta forma a medida é definida no espaço por  $T$ . A transformada de Fourier é um exemplo de aplicação do teorema espectral e portanto uma versão da transformada de Gelfand. O operador linear definido na equação (1136) se chama *operador integral*.



---

- **T<sub>E</sub>X** é um programa feito por Donald Knuth, inicialmente para que ele pudesse produzir o seu livro (inacabado) *The art of computer programming* mas que terminou se tornando um objetivo em si próprio às custas do livro. . . O programa na verdade é uma linguagem *rudimentar* de programação que objetiva colocar texto em forma artística (no sentido que as Editoras entendem) em papel. Como Knuth, que assim é um dos pioneiros do *código aberto* deixou o seu programa em domínio público, uma grande coleção de outros programas e linguagens de programação de nível mais alto foram produzidas em cima do T<sub>E</sub>X, como, por exemplo L<sup>A</sup>T<sub>E</sub>X que é possivelmente a forma mais comum de usar T<sub>E</sub>X. Este dicionário está sendo redigido com L<sup>A</sup>T<sub>E</sub>X .

---

- **Topologia** é uma das grandes *divisões da Matemática*.

A *Topologia* consiste na busca das estruturas que permitam a definição de funções contínuas, neste caminho se procurou “limpar” o caminho na busca de uma melhor compreensão do que seria uma *função contínua* se chegando a relação entre “abertos” e a imagem inversa de funções destes “abertos” como forma de estabelecer o que é uma função contínua. A *continuidade* é então um conceito relativo à estrutura topológica que estiver definida entre dois espaços, se o espaço de saída for suficientemente rico de abertos então as funções nele definida tem mais “chance” de serem contínuas, e reciprocamente, quanto mais “pobre” em abertos for o espaço de chegada, maior “chance” têm as funções, que nele tomem valor, de serem contínuas.

Entretanto existem topologias “usuais” que de uma certa forma já foram aprovadas por uma certa prática, ao reduzir a quantidade de abertos se “enfraquece” a topologia (porque se diminuem as chances de que uma função, definida nesta topologia, seja contínua). Este processo de análise do enfraquecimento de topologias conduz à descoberta de propriedades interessantes de algumas funções ou classes de funções, é esta a pesquisa central na *Topologia*. Desta forma a *Topologia* é vista como uma pesquisa de estruturas, as estruturas topológicas dos espaços.

Há um outra forma de ver a *Topologia* como o estudo das propriedades locais de um espaço, independente (de certa forma) de funções definidas nele, mas na verdade analisando as funções definidas nele, nele mesmo, Em particular a identidade ou a inclusão em espaços de dimensão maior. Isto conduz a descoberta de objetos com formatos muito interessante e a chamada *conjectura de Poincaré* cai neste caso. Um exemplo entre os mais simples é a fita de Moebius que é o símbolo do IMPA.

---

- **transcendente, número** se ele não for algébrico sobre o anel dos números racionais, ou em outras palavras se ele não for um número real algébrico. Por exemplo sabemos, não é fácil demonstrar mas existem demonstrações de que  $\pi$ ,  $e$  são números reais transcendentos, o que significa que eles não algébricos.

O conceito de transcendência pode ser colocado de forma mais geral, contexto da teoria dos anéis, e me permitir o enunciado dum problema interessante que *não sei se já foi resolvido*.

Considere o anel  $A$  como um subanel do anel  $B$ ,  $A \subset B$ . Podemos então construir um outro anel

$$R = A[x_1, \dots, x_n] \quad (137)$$

$$P \in R; P \text{ é um polinômio em } n \text{ variáveis sobre } A; \quad (138)$$

Então  $R$  é o anel dos polinômios em  $n$  variáveis sobre  $A$ . Se  $A$  for um corpo, então  $R$  é um espaço vetorial de dimensão  $n$  sobre  $A$ . Podemos definir um morfismo de anel de  $R$  em  $B$  fixando um elemento  $b \in B$  e considerando o morfismo de avaliação:

$$R \xrightarrow{\phi} B; \phi(P) = P(b); \quad (139)$$

Se o  $\ker(\phi) = \{0\}$ , o núcleo trivial, então este morfismo é uma imersão de  $R$  em  $B$ , um isomorfismo de  $R$  num subanel de  $B$ , porque podemos identificar todo polinômio  $P \in R$  com sua única imagem  $P(b) = P(b_1, \dots, b_n) \in B$ , e não haverá nenhum polinômio com coeficientes em  $A$  tal que  $P(b_1, \dots, b_n) = 0$ . Neste caso  $\{b_1, \dots, b_n\}$  são independentes sobre  $A$ <sup>67</sup>.

**O problema:** não sei se  $\pi, e$  são independentes sobre  $\mathbf{Q}$ , ou ainda se há algum polinômio em duas variáveis, não identicamente nulo,  $P$ , tal que  $P(e, \pi) = 0$ .

Basta agora considerar a redação do problema para  $n = 1$  e não existe nenhum polinômio que não seja identicamente nulo, tal que  $P(e) = 0$  ou  $P(\pi) = 0$  que significa:  $e, \pi$  são números transcendententes.

---

- **Transformada de Fourier** Traz o nome de Joseph Fourier que entre 1807 e 1822 escreveu alguns trabalhos publicados nos anais da Academia Francesa de Ciências sobre a propagação do calor usando somas de senos e cossenos *acelerados* e *amplificados* para aproximar as ondas térmicas conseguindo assim descrever com grande precisão a propagação do calor. Estas somas de senos e cossenos acelerados e amplificados, hoje chamadas de *séries de Fourier* [?], já vinham sendo utilizadas por Euler e alguns dos irmãos Bernouilli na solução de equações diferenciais. As transformadas de Fourier (as séries ou a integral) redescrivem uma onda em termos das ondas mais simples, sin, cos que são assim os vetores próprios de um sub-espaço (medido) de funções sendo os coeficientes desta transformação os valores próprios que caracterizam uma determinada equação diferencial que esteja sendo estudada.

---

- **triângulos semelhantes**, é uma relação de equivalência entre triângulos da Geometria Euclidiana. Como um triângulo é uma figura plana, se trata dum problema do plano.

**Teorema 21 (semelhança de ) triângulos**

*Todos os triângulos do plano tem uma representação semelhante no círculo trigonométrico.*

---

<sup>67</sup>A leitora deve reconhecer que este é o caminho para definir a dimensão dum espaço vetorial.

**Dem**:

Considere que qualquer diâmetro divide o círculo trigonométrico em duas semiesferas então

- todos os triângulos retângulos correspondem a um triângulo inscrito tal que a corda coincide com o diâmetro,
- todos os triângulos acutângulos correspondem a um triângulo inscrito tal que a corda esteja numa semiesfera, determinando dois vértices, com o terceiro vértice na outra semiesfera, isto garante que nenhum ângulo será maior do que  $\frac{\pi}{2}$ , mas podemos obter qualquer ângulo menor do que  $\frac{\pi}{2}$ .
- todos os triângulos obtusângulos correspondem a triângulos inscritos na esfera  $\mathbf{S}^1$  tal que todos os três vértices estejam na mesma semiesfera, o que garante que um dos ângulos inscritos é maior do que  $\frac{\pi}{2}$ .

**q.e.d.**

Este teorema é obviamente um “teorema de existência”, porque não é construtivo. O método para obter um representante para um determinado triângulo consiste em obter uma sucessão de triângulos que se aproximem da imagem desejada.

Confira a figura (fig 152), página 385,

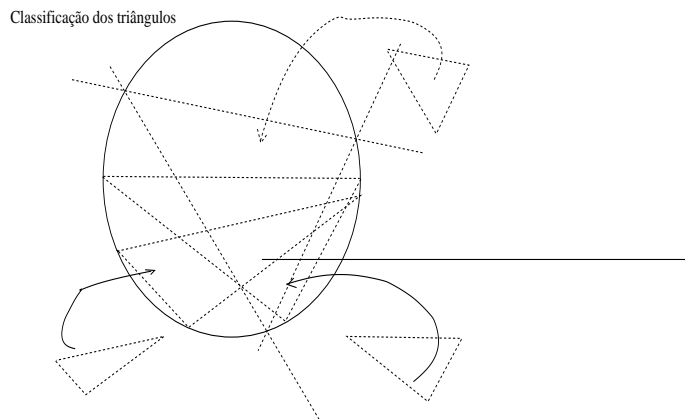


Figura 152: Representação dos triângulos em  $S$

Entretanto o teorema 21 omite um fato importante que é a multiplicidade das representações. Mas é possível obter um único representante para cada triângulo do plano observando que a imagem dos triângulos isósceles é exatamente a que se obtém usando como vértice o ponto obtido com a perpendicular à corda mencionada acima. Vou então introduzir uma linguagem que vai me levar de forma simples ao conjunto quociente desta relação.

Para encontrar a representação dum triângulo dado, selecione uma corda, e aqui começa a pluralidade de representações pois é possível escolher a corda de muitas maneiras. Então, no círculo trigonométrico,  $\mathbf{S}^1$  que existe uma infinidade de triângulos semelhantes a um dado triângulo do plano inscrito num círculo.

Vou expor uma forma de selecionar as distintas classes de equivalência para finalmente encontrar uma forma simples de representação dum triângulo no círculo trigonométrico  $\mathbf{S}^1$ .

Para justificar a linguagem, vou seguir mostrando a pluralidade e como contorná-la.

- aos triângulos retângulos correspondem à corda coincidindo com o diâmetro, e assim definindo dois vértices, que são pontos *antípodas* do círculo, ficando o terceiro vértice em uma das semi-esferas em que o diâmetro divide o círculo trigonométrico. Temos assim duas possibilidades para escolher ou para eliminar uma delas.

Se o triângulo for isósceles, as duas imagens que lhe correspondem serão encontradas selecionando o terceiro vértice nos pontos em que a corda vertical ao diâmetro encontrar  $\mathbf{S}^1$ . Como há dois tais triângulos isósceles, eu preciso de um método para eliminar um deles. Se eu chamar o diâmetro escolhido de *equador* os dois pontos que determinam os triângulos isósceles, serão chamados naturalmente de *Sul* e *Norte*. O *S* está no hemisfério *Sul* e portanto vou sempre escolher o terceiro ponto no *Sul* com isto identifiquei de forma único o *triângulo isósceles*, confira a figura (fig. 153), página 386, onde aparece o *triângulo isósceles* separando dois triângulos retângulos equivalentes.

### O triângulo isósceles divisor de classes

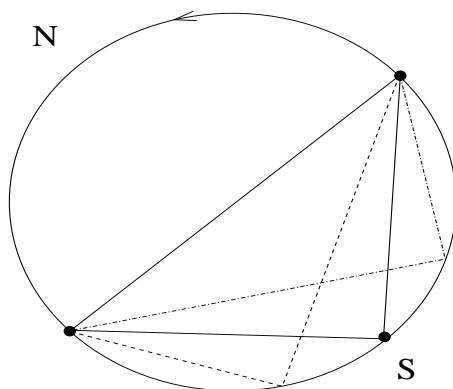


Figura 153: O triângulo isósceles

Se o triângulo não for isósceles, há dois triângulos equivalentes que podem ser obtidos escolhendo o terceiro vértice, no *Sul*, à direita ou à esquerda de *S* e assim eu poderei encontrar um único representante para qualquer triângulo retângulo. Mas para falar de *esquerda* ou *direita*, tenho que considerar o círculo  $\mathbf{S}^1$  orientado, o que farei como habitual considerando o sentido positivo, o antihorário, e então à *esquerda* significa *positivo*,

porque, eu, o observador, encontro-me em pé no centro do círculo tendo à minha frente o ponto inicial:  $e^{i0} = 1$ , o que em particular me agrada!

Este *triângulo isósceles* único, vai estar presente em toda a análise e portanto preciso que esta expressão seja introduzida para fazer-lhe referência. Na construção do método observei que ao selecionar uma corda existe um caso particular do *triângulo isósceles* que irá dividir todos os triângulos em duas classes idênticas e fiz uma seleção por uma dessas classes para estabelecer a unicidade da representação, confira a figura (153) página 386, onde se podem ver dois triângulos retângulos equivalentes com o terceiro vértice à esquerda ou à direita do  $S$  que é o terceiro vértice do *triângulo isósceles*. Este “*triângulo isósceles*” vai ser importante em toda a construção a seguir.

- aos acutângulos correspondem à corda numa semi-esfera, determinando dois vértices, com o terceiro vértice na outra semi-esfera, desta forma o ângulo que se opõe à corda é menor do que  $\frac{\pi}{2}$  sobrando a diferença para distribuir pelos dois outros ângulos. Fixando uma semi-esfera para nela considerar a corda, o terceiro vértice estará na outra semi-esfera. Vou sempre escolher o terceiro vértice no hemisfério Sul, a corda, no caso dos triângulos acutângulos será sempre escolhida no hemisfério Norte.

Agora existe um único triângulo isósceles associado à corda atendendo à condição de que o terceiro vértice seja escolhido no Sul. Mas haverá duas escolhas para o terceiro vértice se triângulo não for isósceles. Logo veremos como decidir isto de forma única. Aqui também o “*triângulo isósceles*” irá dividir todos os triângulos em duas classes idênticas.

- aos obtusângulos correspondem à corda numa semi-esfera, determinando dois vértices, com o terceiro vértice na mesma semi-esfera. O ângulo que corresponde à corda mencionada mede mais do que  $\frac{\pi}{2}$ . Como sempre vou escolher o terceiro vértice no Sul, agora a corda também vai ser escolhida no Sul. Como no caso anterior, dos triângulos acutângulos, há uma única possibilidade de construção do *triângulo isósceles* mas haverá duas no caso dum triângulo que não seja isósceles pela existência das duas classes.

Do exposto nos itens anteriores, se vê que é preciso uma notação para obter uma classificação mais simples e mais efetiva uma vez que a cada triângulo no plano correspondem múltiplos triângulos que lhe são semelhantes inscritos no círculo trigonométrico.

As palavras que preciso, ou os símbolos, são

- Estarei a todo momento mencionando três triângulos, um que será escolhido aleatoriamente no plano, outro que é a projeção deste em  $\mathbf{S}^1$ , e o terceiro que é o *triângulo isósceles*. Vou denominá-los, respectivamente, de *o triângulo* e *a imagem do triângulo*, este um triângulo inscrito em  $\mathbf{S}^1$ , e o terceiro é o triângulo isósceles que separa os triângulos associados a uma determinada corda em duas classes de triângulos equivalentes, vou referir-me a este triângulo como *o triângulo isósceles*.

- A origem de  $\mathbf{S}^1$ ,  $e^{i0} = 1$ , e sua orientação positiva, o sentido antihorário.
- Ao escolher um diâmetro para  $\mathbf{S}^1$  o estarei dividindo em duas semi-esferas que vou chamar de  $N$  e de  $S$ , sugerido pelos polos *norte e sul*, sem que o diâmetro pertença a nenhuma das semi-esferas. Temos assim duas semi-esferas abertas. Vou chamar, na continuação, este diâmetro escolhido de equador para continuar com a notação geográfica. A seleção do norte e do sul ficaram bem definidas pela orientação com  $S$  à direita da origem e  $N$  a esquerda da origem.

### A relação de equivalência

1. retângulos Para obter o representante de qualquer *triângulo retângulo do plano*, tomamos o *equador*, como corda, paralelo à hipotenusa, e o terceiro vértice na semi-esfera  $S$ .

Se o triângulo for isósceles sua imagem será o *triângulo isósceles*.

Se o triângulo não for isósceles, tome uma paralela ao cateto menor passando pela origem de  $\mathbf{S}^1$ . Esta paralela irá encontrar  $S$  no ponto  $1 > P > S$ , passe uma paralela ao outro cateto por  $P$  para encontrar o ponto  $Q$ . Com isto temos o *triângulo imagem* equivalente ao *triângulo* pelo método: *lado, ângulo, lado*.

Aqui vemos ressurgir o problema de invariância já citado acima, um problema, digamos, *epistemológico*, o triângulo, cuja representação estamos querendo encontrar em  $\mathbf{S}^1$ , e neste caso um triângulo retângulo, pode não ter a hipotenusa paralela ao equador, mas para isto, basta redesenhar  $\mathbf{S}^1$  e selecionar o equador paralelamente à hipotenusa. Este problema vai se repetir nos demais casos sem que eu me sinta obrigado a discutí-lo novamente. Esta questão também mostra a fragilidade de nossa comunicação oral ou escrita cuja solução passaria por uma linguagem extremamente sofisticada e difícil, e é melhor encontrar um meio termo entre o *bourbaquismo* e a imprecisão...

2. acutângulos Para obter o representante de qualquer *triângulo acutângulo*, selecione a corda, na semi-esfera  $N$ , paralela ao lado menor do triângulo, assim como também ao *equador*. A seleção do lado menor é consequência da seleção do *triângulo isósceles*. Se o *triângulo* for isósceles, o lado menor é exatamente o que é diferente dos outros dois. Deixe-me agora designar por  $l_1, l_2$  os dois outros lados com  $\|l_1\| \leq \|l_2\|$ , porque eles poderão ter mesmo comprimento no caso do triângulo isósceles.

O uso dum editor gráfico pode servir de forma magnífica nesta construção trazendo como exemplo um *problema de aproximação* e ilustração dum *teorema de existência*. Enquanto que na construção anterior, do triângulo retângulo, a simples escolha dum cateto paralelo encerrou o trabalho, aqui, a seleção da corda, paralela a um dos lados é insuficiente, mas é um *valor inicial dum problema iterativo de aproximação*.

Trace agora a paralela ao  $l_1$  passando pelo menor dos vértices da corda<sup>68</sup>. Esta paralela encontra  $\mathbf{S}^1$  no ponto  $P_1$ . Passe uma paralela a  $l_2$  pelo ponto  $P_1$ . Estas duas retas vão se encontrar no ponto  $Q_1$ , confira a figura (fig. 154), página 389,

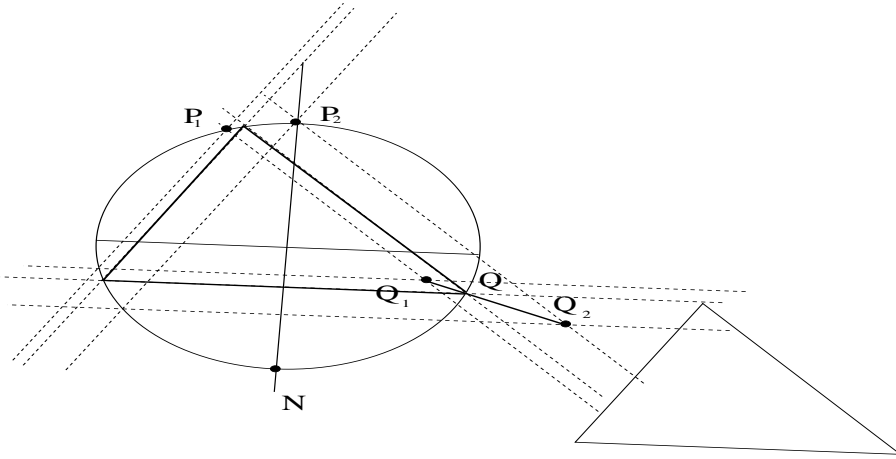


Figura 154: Primeira aproximação

Ao fazê-lo acontecerão exatamente dois eventos exclusivos:

- (a) o ponto  $Q_1$  cai sobre  $\mathbf{S}^1$  ou no exterior de  $\mathbf{S}^1$ ;
- (b) o ponto  $Q_1$  cai sobre  $\mathbf{S}^1$  ou no interior de  $\mathbf{S}^1$ ;

e cada um destes eventos corresponde a uma corda mais distante ou mais próxima do *equador* associando

$$P_1 \mapsto Q_1; P_2 \mapsto Q_2; \quad (1140)$$

o que prova que a solução existe. Confira a figura (fig. 154), página 389, onde se podem ver dois pontos pontos,  $P_1, P_2$  escolhidos pela determinação de duas *cordas* associados às imagens  $Q_1, Q_2$ .

A reta determinada por  $Q_1, Q_2$  encontra  $\mathbf{S}^1$  no ponto  $Q$  que não precisa ser a solução que procuramos mas vai ser o segundo valor dum algoritmo do tipo binário, em que estaremos sempre escolhendo um ponto médio de outros recentemente obtidos, e este é o caso,  $Q$  é um ponto médio entre  $Q_1, Q_2$ .

Na figura (fig. 154), página 389, ocorreu o evento “o ponto  $Q_1$  cai sobre  $\mathbf{S}^1$  ou no interior de  $\mathbf{S}^1$ ”.

Trace agora a corda paralela ao lado menor do triângulo passando por  $Q$ . Ela estará entre as duas outras portanto, ao iterar o método iremos

<sup>68</sup>Não esquecendo que  $\mathbf{S}^1$  está orientado portanto faz sentido falar no menor vértice.

encontrar como limite o ponto  $Q$  em que deve passar o terceiro lado de modo que reconstruindo, no sentido reverso, o segundo lado e o primeiro lado, encontramos, como um limite, a imagem do triângulo sobre  $\mathbf{S}^1$ .

Observe que a seleção do lado menor para construir a paralela nos leva a selecionar os pontos  $P_i$  à esquerda de  $S$  que seria o vértice do *triângulo isósceles* e portanto o método está consistente com o objetivo de usar o *triângulo isósceles* para delimitar o *conjunto quociente* da relação de equivalência.

*método: ângulo, ângulo, ângulo.*

3. obtusângulos Para obter o representante de qualquer triângulo obtusângulo, selecionamos a corda na semi-esfera  $S$ , paralela ao lado maior do triângulo seguindo-se depois pelo método descrito para triângulos acutângulos para obter os dois outros vértices na semi-esfera  $S$ . Na figura (155) página 390, você pode ver os pontos  $P_1, P_2$  e suas imagens,  $Q_1, Q_2$  assim como o ponto

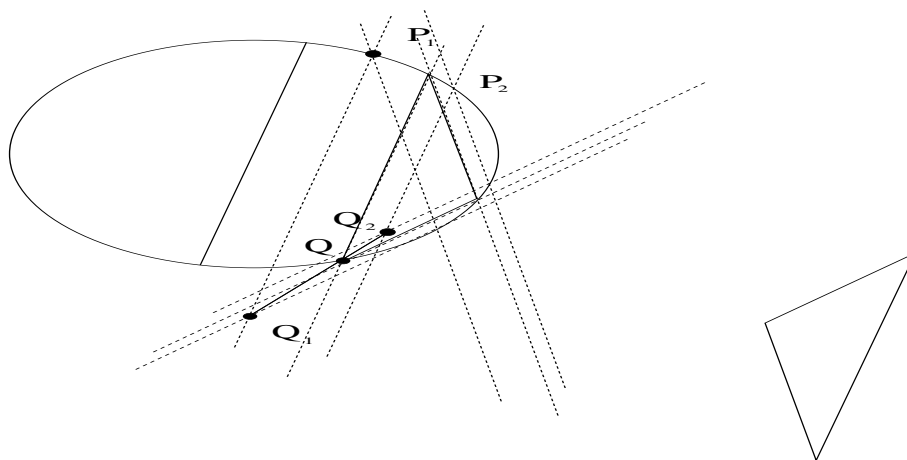


Figura 155: O algoritmo para obter imagem de obtusângulos

médio  $Q$  que foi usado no próximo passo do algoritmo iterativo. *método: ângulo, ângulo, ângulo.*

Como existem apenas três tipos de triângulos e todos foram estudados posso agora enunciar o *teorema das classes de equivalência dos triângulos do plano* cuja demonstração foi desenvolvida ao longo análise geométrica que fiz.

**Teorema 22 (Conjunto quociente)** *equivalência de triângulos*

*Considere no círculo trigonométrico um diâmetro, chamado equador relativamente ao qual  $\mathbf{S}^1$  fica dividida em duas semiesferas chamada de norte e sul. Considere também em  $\mathbf{S}^1$  a orientação positiva, em que o ponto  $e^{i0} = 1$  é a origem, então o polo sul  $S$  e o polo norte  $N$  são dois pontos antípodas sendo  $S < 1$  e  $N > 1$ .*



O conjunto quociente da relação de equivalência de triângulos do plano é o conjunto de todos os triângulos inscritos no círculo trigonométrico tendo um lado paralelo ao equador e o terceiro vértice no intervalo  $[S, 1)$ . No caso dos acutângulos o lado paralelo ao equador fica na semiesfera norte, para os retângulos este lado é exatamente o equador e os obtusângulos têm o lado paralelo na semiesfera sul.

A figura (fig 156), página 391, mostra três elementos do conjunto quociente, de cada um dos três tipos de triângulos. Em busca da figura de melhor qualidade estética terminei preferindo representar os *triângulos isósceles*, porque afinal são

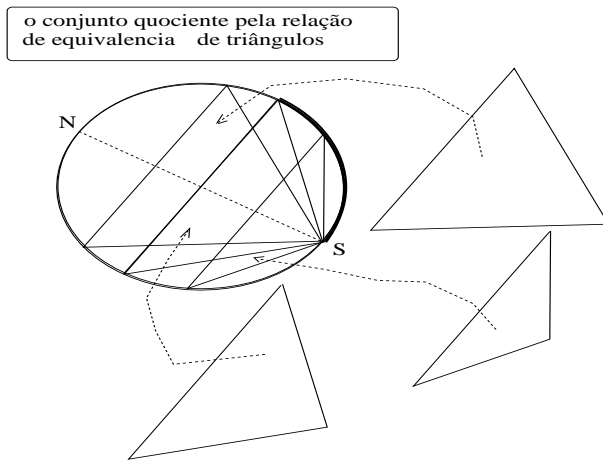


Figura 156: Três elementos do conjunto quociente

eles que determinam o início do intervalo que define o conjunto quociente.

Nestas condições para demonstrar que as bissetrizes num triângulo qualquer se encontram num mesmo ponto equidistante dos lados, basta fazê-lo para um triângulo qualquer inscrito no círculo trigonométrico, e verificar que a média dos pontos  $(\cos(\alpha), \sin(\alpha))$ ,  $(\cos(\beta), \sin(\beta))$  e  $(\cos(\gamma), \sin(\gamma))$  em que  $\alpha, \beta, \gamma \in [0, 2\pi)$  é o ponto de encontro das mediatrizes. É o que vou fazer na próxima seção.

É possível obter-se uma expressão mais simples para o teorema 22 observando que são apenas os vértices que interessa na descrição mas que é importante fixar um diâmetro para estabelecer uma referência. Precisamos então do *equador* e de cordas paralelas a este diâmetro com isto temos dois pontos. O terceiro irá se encontrar no intervalo  $[S, 1)$ . Assim temos, usando a linguagem estabelecida acima,

**Teorema 23 (classes de) equivalência dos triângulos**

Considere o equador como um diâmetro selecionado de  $\mathbf{S}^1$ . Qualquer triângulo do plano tem um único representante inscrito em  $\mathbf{S}^1$  determinado por três pontos, dois deles determinando uma corda paralela ao equador e outro escolhido no intervalo  $[S, 1)$ .

Não há nada a demonstrar, o teorema 23 é apenas uma reformulação mais simples do teorema 22. Nesta formulação simplificada está omissa a condição de se a corda paralela ao *equador*, determinando dois pontos que se encontram na semiesfera *norte* o triângulo obtido representa um triângulo isósceles. Se estiverem na semiesfera *sul* é um triângulo acutângulo que está representado, e se forem os pontos  $L, O$  que determinam o *equador* é um triângulo retângulo que está sendo representado.

- **trigonometria** é a parte elementar da Matemática em que se relacionam os ângulos num triângulo retângulo com seus catetos e a hipotenusa, a figura (157) página 392, mostra o círculo trigonométrico, o ângulo  $\alpha$  e as duas funções funda-

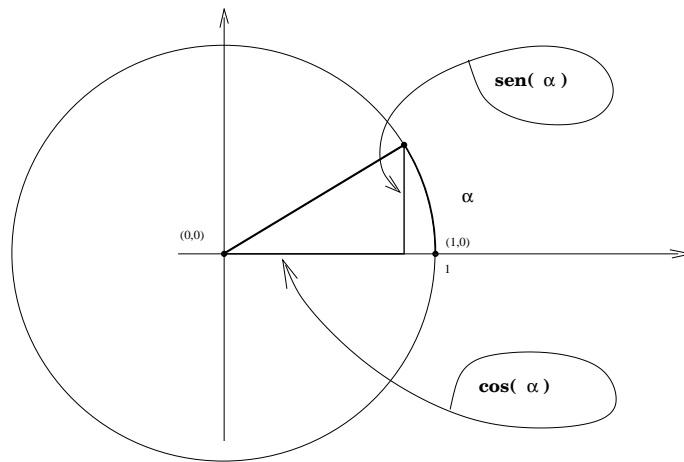


Figura 157: círculo trigonométrico  $S^1$

mentais,  $\sin(\alpha), \cos(\alpha)$  que podem ser calculadas geometricamente se o círculo for desenhado em papel milimetrado, por contagem das subunidades. Como  $\sin(\alpha), \cos(\alpha)$  são as coordenadas de um ponto no círculo unitário determinado pela origem  $(1, 0)$  do círculo trigonométrico, e a hipotenusa traçada da origem dos eixos  $(0, 0)$ , o teorema de Pitágoras nos fornece a relação fundamental da trigonometria

$$\sin^2(\alpha) + \cos^2(\alpha) = 1 \quad (1141)$$

a fórmula de *De Moivre-Euler-Abel*,

$$e^{i\alpha} = \cos(\alpha) + i \sin(\alpha) \quad (1142)$$

permite-nos descobrir rapidamente várias outras fórmulas fundamentais da trigonometria:

$$e^{i\alpha}e^{i\beta} = (\cos(\alpha) + i\sin(\alpha))(\cos(\beta) + i\sin(\beta)); \quad (1143)$$

$$e^{i\alpha}e^{i\beta} = e^{i(\alpha+\beta)} = \quad (1144)$$

$$= \cos(\alpha + \beta) + i\sin(\alpha + \beta) = \quad (1145)$$

$$\cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta) + i(\cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta)); \quad (1146)$$

$$\cos(\alpha + \beta) = \cos(\alpha)\cos(\beta) - \sin(\alpha)\sin(\beta); \quad (1147)$$

$$\sin(\alpha + \beta) = \cos(\alpha)\sin(\beta) + \sin(\alpha)\cos(\beta); \quad (1148)$$

$$\cos(\alpha - \beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta); \quad (1149)$$

$$\sin(\alpha - \beta) = \sin(\alpha)\cos(\beta) - \cos(\alpha)\sin(\beta); \quad (1150)$$

A passagem da equação (1146) para as equações (1147) e (1148), as chamadas equações do *coseno do ângulo soma* e *seno do ângulo soma*, é feita observando as partes real e imaginária na equação (1146) comparada com a equação (1145).

O nome de *De Moivre* está associado às potências de  $e^{i\alpha}$  que nos permitem descobrir diversas variantes de expressões trigonométricas associando as potências de  $\cos(\alpha) + i\sin(\alpha)$  com sua expressão expandida usando o binômio de Newton:

$$e^{in\alpha} = (\cos(\alpha) + i\sin(\alpha))^n = \sum_{k=0}^n \binom{n}{k} \cos^{n-k}(\alpha) i^k \sin^k(\alpha); \quad (1151)$$

$$e^{i3\alpha} = \cos^3(\alpha) + 3i\cos^2(\alpha)\sin(\alpha) - 3\cos(\alpha)\sin^2(\alpha) - i\sin^3(\alpha) = \quad (1152)$$

$$= \cos^3(\alpha) - 3\cos(\alpha)\sin^2(\alpha) + i(3\cos^2(\alpha)\sin(\alpha) - \sin^3(\alpha)); \quad (1153)$$

$$\cos(3\alpha) = \cos^3(\alpha) - 3\cos(\alpha)\sin^2(\alpha); \quad (1154)$$

$$\sin(3\alpha) = 3\cos^2(\alpha)\sin(\alpha) - \sin^3(\alpha); \quad (1155)$$

**Soma de cosenos:**  $f(x) = a\cos(x) + b\sin(x)$

Esta é uma expressão interessante que permite escrevermos  $f$  como um múltiplo do  $h(x) = \cos(x - \alpha)$ . Observe como se faz.

Compare  $f(x)$  com um produto escalar e aplique a desigualdade de Cauchy-Bunyakovsky-Schwarz:

$$f(x) = a\cos(x) + b\sin(x); \quad (1156)$$

$$f(x) = \langle (a, b), (\cos(x), \sin(x)) \rangle; \quad (1157)$$

$$|f(x)|^2 \leq a^2 + b^2; |(\cos(x), \sin(x))| = 1; \quad (1158)$$

$$|f(x)| \leq R = \sqrt{a^2 + b^2}; \quad (1159)$$

então existe um número,  $R = \sqrt{a^2 + b^2}$ , tal que  $|f(x)| \leq R$ , mais exatamente,  $R$  limita  $|f|$  o gráfico de  $f$  ao intervalo  $[-R, R]$ .

Como a expressão do *coseno da soma* oferece uma combinação linear de seno e coseno,

$$\cos(x - \alpha) = \cos(\alpha)\cos(x) + \sin(\alpha)\sin(x) = A\cos(x) + B\sin(x); \quad (1160)$$

comparando com a expressão  $R \cos(x - \alpha)$  tem-se

$$R \cos(x - \alpha) = R (\cos(x) \cos(\alpha) + \sin(x) \sin(\alpha)); \quad (1161)$$

$$R \cos(x - \alpha) = R \cos(\alpha) \cos(x) + R \sin(\alpha) \sin(x); \quad (1162)$$

$$R \cos(x - \alpha) = a \cos(x) + b \sin(x); \quad (1163)$$

$$a = R \cos(\alpha); b = R \sin(\alpha); \quad (1164)$$

$$\cos(\alpha) \neq 0 \Rightarrow \frac{b}{a} = \frac{R \sin(\alpha)}{R \cos(\alpha)} = \tan(\alpha); \quad (1165)$$

$$\alpha = \text{Atan}\left(\frac{b}{a}\right); \quad (1166)$$

que é uma expressão válida para uma grande variação de escolhas dos coeficientes  $a, b$ .

$$f(x) = a \cos(x) + b \sin(x) = R \cos(x - \alpha); \alpha = \text{Atan}\left(\frac{b}{a}\right); \quad (1167)$$

Com `gnuplot` você pode experimentar com vários valores para  $a, b$ , eis o código:

```
pow(x,n) = x**n;
a=2.0;b=-4.0;alpha = atan(b/a);R = sqrt(pow(a,2) + pow(b,2));
f(x) = a*cos(x) + b*sin(x);h(x) = R*cos(x-alpha);
plot h(x),f(x),0
```

---

### - trigonométrica, série

Procure **séries**. Uma *série trigonométrica* é um dos formatos em que as *transformadas de Fourier* podem se apresentar é a chamada *transformação discreta de Fourier*.

---

### - trigonométricas, séries absolutamente convergentes

Se uma série trigonométrica for absolutamente convergente, ela define uma função (se não for absolutamente convergente também define, mas pode não ser contínua e se enrique a teoria com novos aspectos). Vamos usar a notação complexa porque ela nos permite um texto mais resumido.

$$f(x) = c_0 + \sum_{k \in \mathbf{Z}} c_k e^{ikx} \quad (1168)$$

podemos mostrar que a equação

$$\| \| f \| \|_1 = |c_0| + \sum_{k \in \mathbf{Z}} |c_k| \quad (1169)$$

é uma *norma* e portanto o conjunto das séries trigonométricas absolutamente convergentes é um espaço vetorial normado..

Se “esquecermos” as funções  $e^{ikx}$  na expressão podemos identificar, na expressão de uma série trigonométrica absolutamente convergente, a série de termo geral  $a_k$ , a série dos termos em módulo sendo convergente o que nos permite

associação com um outro tipo de espaço vetorial, o das sucessões associadas à séries absolutamente convergentes que é o espaço vetorial normado  $l^1$  e os dois espaços vetoriais normados em questão serão isomorfos, este é o conteúdo do Lema de Wiener. Um século se passou antes que este *detalhe* fosse descoberto. A importância deste detalhe aparece num fato simples: é “fácil” provarmos que no espaço  $l^1$  das sucessões, existe mais uma operação: o produto de convolução das sucessões que neste caso é uma operação interna, o produto de duas sucessões cujas séries sejam absolutamente convergentes, é outra sucessão com uma série absolutamente convergente. Com o isomorfismo mencionado acima podemos retornar ao espaço das séries trigonométricas absolutamente convergentes e obter de forma relativamente simples que elas formam uma *álgebra de Banach*. O isomorfismo mencionado associa o produto de convolução do espaço  $l^1$  com o produto ponto a ponto das funções que as séries trigonométricas definem. Porém com um problema extra a unidade no álgebra de Banach das séries trigonométrica é a função constante que não tem série de Fourier. Em  $l^1$  é a sucessão  $\delta_0$ . Os *morfismos* são uma forma de descobrir problemas! Quer dizer os espaços de Banach são isomorfos mas não o são as álgebras de Banach, e não o deveriam?

É interessante como este problema, da falta de unidade na álgebra de Banach da séries trigonométricas absolutamente convergentes, a álgebra de Wiener, está associada com outras questões. Em teoria da informação e comunicações este problema é conhecido como a dualidade entre a limitação no *espaço da frequência vis a vis espaço do tempo*, ou, se uma das transformadas tiver suporte limitado a outra o terá não limitado. A resposta para existência da unidade seria uma imagem com suporte reduzido a um ponto, a distribuição de Dirac. Este é apenas um resumo, entretanto.

---

- **UML\*\*** Do inglês, Universal Modeling Language, uma das tentativas de se criar uma linguagem universal de processamento - independente de linguagens específicas de programação. Uma outra tentativa se chama Interlanguage Unification.

---

- **variedade** A palavra *variedade* foi inventada para nos liberar da prisão tri-dimensional em que nos encontramos tanto por razões físico-energéticas, como culturais uma vez que a nossa cultura geométrica, de origem dita grega, nos fixou o vocabulário dentro da dimensão três.

Falamos em *ponto, retas, planos, superfícies, volumes* que são todos objetos que ficam dentro do limite da dimensão três. Embora pontos, reta, planos sejam invisíveis para nós seres tri-dimensionais, temos a sensação de que os conhecemos e entendemos.

As necessidades científicas vão muito além da dimensão três e precisam das dimensões 0,1,2 às quais não podemos ter acesso por razões físicas, de energia mesmo. Para isto, e para completar o vocabulário, criamos a palavra *variedade* que modificada por um adjetivo dimensional, resolve a questão linguística.

Por exemplo, classificamos os objetos da geometria, como variedades,

- diremos uma variedade de dimensão 1, para fazermos referência aos segmentos de reta, às retas, às curvas. Todos estes objetos são variedades de dimensão 1.
- quando nos referirmos aos objetos de dimensão 2, diremos variedades de dimensão dois são os planos e as superfícies da geometria;
- há uma grande subclassificação das variedades - duas grandes classes:
  - As variedades lineares, segmentos de reta, retas, planos, as variedades lineares de dimensão 3, 4 para as quais não temos mais nomes geométricos.
  - As variedades não lineares, um círculo, uma parábola, são dois exemplos de variedades não lineares de dimensão 1. Uma superfície de tipo parabolóide, as parabólicas podem ser pensadas como sendo tal, são variedades não lineares de dimensão 2.
  - Um ponto você pode classificar como quiser, uma variedade linear de dimensão 0, ou uma variedade não linear de dimensão 0.
- Herdamos nomes particulares para alguns tipos de variedades de dimensão 1, círculos, retas, parábolas, ou simplesmente a palavra curva, variedades de dimensão 1.
- Também temos nomes para algumas variedades de dimensão dois, plano, superfície esférica. A palavra superfície quer dizer uma variedade de dimensão 2.
- O espaço todo em que estamos imersos é uma variedade linear de dimensão 3.
- O espaço-tempo da Física é uma variedade de dimensão 4. Linear? depende, se o tempo tiver uma condição inicial, não! porque neste caso seria um poliedro... poliedros, embora tenham fronteiras feitas de subconjuntos de variedades lineares, não são mais variedades lineares.
- Observe que uma reta qualquer é uma variedade linear afim, isto quer dizer, por exemplo, que a origem pode não pertencer a reta. Se a origem pertencer à reta, ela é uma variedade linear. Se a origem não pertencer à reta ela é uma variedade linear afim de dimensão 1.
- As variedades lineares afins de dimensão 2 são os planos que podem não passar pela origem. O qualificativo “afim” caracteriza que a variedade foi obtida por uma translação de uma variedade linear.

---

- **Venn** diagramas de. São figuras geométricas planas, que são regiões bem delimitadas do plano, como auxílio duma curva, para representar conjuntos.

Na figura (fig 158), página 397, você pode ver alguns diagramas de Venn

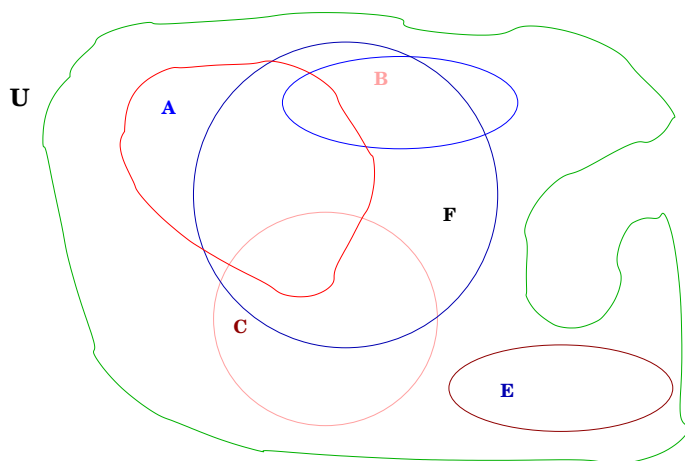


Figura 158: diagrama de Venn

incluídos num diagrama, chamado  $U$ , representando o *universo* relativo aos demais que nele estão contidos.

Os diagramas de Venn servem para destacar interseção de conjuntos e também a inclusão entre eles. Na (fig 158), página 397, você pode ver um conjunto  $U$  que é o *universo relativo aos demais* que estão contidos nele.

Os conjuntos  $A, B, F, C$  todos tem interseção não vazia uns com os outros, exceção: os conjuntos  $B$  e  $C$  são disjuntos.

O conjunto  $E$  tem interseção vazia com todos os conjuntos  $A, B, F, C$ .

Todos os conjuntos  $A, B, F, C, E$  são subconjuntos de  $U$  que assim é o *universo*.

Um outro diagrama, chamado de *Hasse* tem o aspecto dum tipo de *grafo*, *árvore*, e serve para representar inclusão. Confira Hasse, diagrama.

---

- **verdade, tabela** confira *tabela verdade*. Dada uma relação frequentemente é possível identificar dois estados que esta relação pode assumir caracterizados pelos dois valores **verdade**, **falso**. Há relações para às quais não é possível definir um destes estados para elas, o que conduz a uma classe de lógica chamada de lógica fuzzy em que se considera um terceiro estado de indecisão acrescentando-se um percentual para as possibilidade **verdade**, **falso**.

Estes conceitos fazem parte da lógica matemática.

---

- **vizinhança** é um aberto de um espaço topológico contendo um ponto  $x$ , se diz então uma vizinhança de  $x$ ,  $v_x$ . O conceito “vizinhança” pode ser usado como “conceito primitivo” para definir *topologia*, assim como *topologia* pode ser o “conceito primitivo” para definir vizinhança. Funcionam as duas formas de pensar.

Topologia ou vizinhança servem para definir convergência, entretanto este é um conceito mais fraco do que topologia no sentido de que nem todas as

propriedades de um espaço topológico podem ser obtidos a partir do conceito de convergência. Um exemplo interessante é o conjunto dos números reais em que a convergência usual define a sua topologia usual, e conseqüentemente uma função será contínua, se e somente se, for sequencialmente contínua. Quase todos os teoremas sobre limite e derivação do Cálculo ficam bastante simplificados se for usada continuidade sequencial em lugar da continuidade tradicional que é usada.

Há autores que admitem vizinhanças que não sejam abertas, aqui vou considerar apenas abertos como vizinhanças. Uma bola aberta, centrada num ponto  $x$  em um espaço métrico é uma vizinhança de  $x$  neste espaço métrico. Se usa a notação  $\mathcal{B}(x, \epsilon)$ , bola de centro  $x$  e raio  $\epsilon$ .

Uma forma de definir topologia parte do conceito de distância e das bolas abertas definidas com uma distância a serem usadas como *vizinhanças básicas* da topologia. Métrica é um sinônimo de distância.

Quando pudermos definir uma distância  $d$  em um conjunto  $X$ , diremos que se trata de um *espaço métrico*  $(X, d)$

**Definição 24 (distância)** *Distância ou métrica*

Uma distância é uma função positiva,  $d$ , definida para todos os pares  $(x, y)$  de elementos de um conjunto  $X$  satisfazendo as propriedades

1. positiva  $d(x, y) \geq 0$  e  $d(x, y) = 0$  se e somente se  $x = y$ ;
2. reflexividade  $d(x, y) = d(y, x)$ ;
3. desigualdade triangular  $d(x, y) + d(y, z) \geq d(x, z)$ ;

Os *espaços métricos* são casos particulares de *espaços topológicos*.

Um tipo particular de espaço métrico é o espaço das funções contínuas com a métrica do supremo (convergência uniforme) e na figura (159) página 399, você pode ver um exemplo de vizinhança tubular, uma bola da métrica do supremo no espaço  $\mathcal{C}([a, b])$  centrada em uma função contínua, o gráfico de  $f$  é o centro da faixa (vizinhança tubular). A topologia construída com esta métrica é chamada de *topologia da convergência uniforme*.

Outro exemplo de distância entre duas funções, no espaço  $f, g \in \mathcal{C}([a, b])$  pode ser definida como o módulo da diferença entre os valores destas funções em um dado ponto do espaço:  $|f(x) - g(x)|$ , neste caso uma vizinhança de  $f$ , pode ser vista na figura (160) página 400, é o conjunto de todas as funções cujos gráficos cortem o segmento vertical de medida  $2\epsilon$  contendo o ponto  $(c, f(c))$ . Todos os gráficos, na figura (160) se encontram a uma distância menor do que  $\epsilon$  de  $f$ . O sistema destas vizinhanças define a convergência ponto a ponto em  $\mathcal{C}([a, b])$ . A topologia construída com esta métrica é chamada de *topologia da convergência pontual*.

Uma outra forma de medir a distância entre funções é sugerida pela quantidade do fenômeno contido na função, por exemplo, a Física fala de *quantidade de movimento* e calcula a integral da velocidade sobre um intervalo considerado.

Podemos assim definir



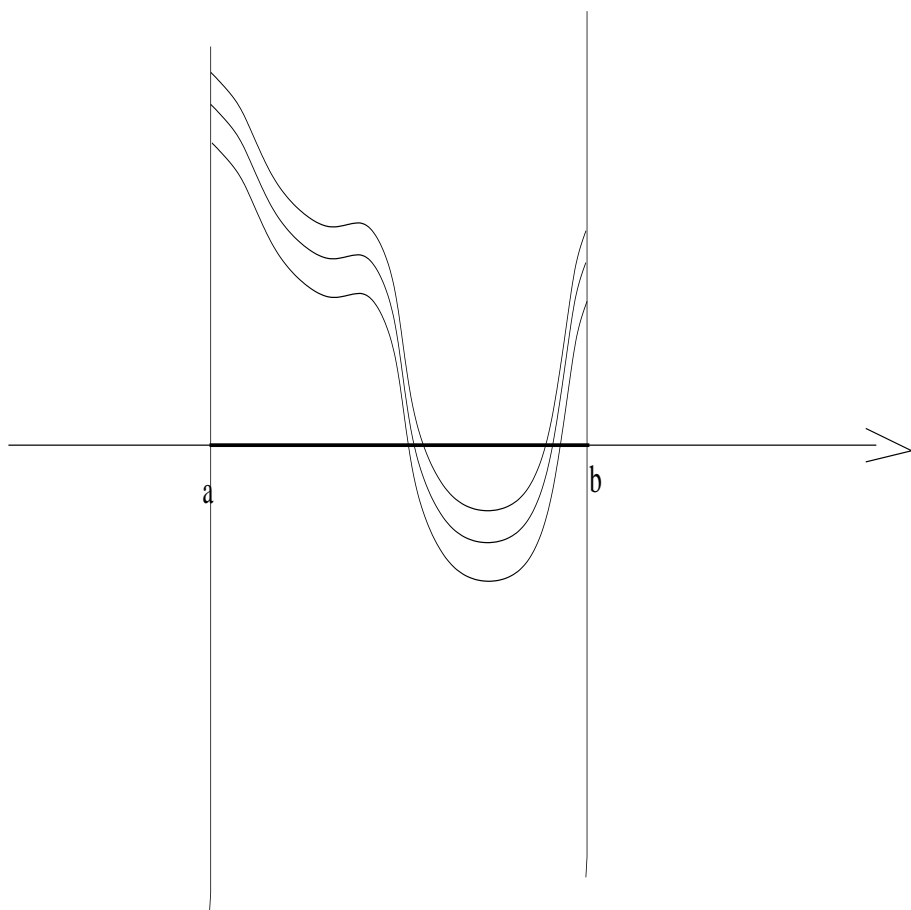


Figura 159: Vizinhança tubular em  $C([a, b])$

**Definição 25 (distância) entre funções**

Dadas duas funções integráveis,  $f, g$  definidas no intervalo  $[a, b]$  podemos definir

$$d(f, g) = \int_a^b |f(x) - g(x)| dx = \|f - g\|_1$$

O número "1" que aparece no símbolo  $\|f - g\|_1$  está registrando que foi usado um caso particular de distância entre funções e um caso mais geral seria

**Definição 26 (distância) entre funções**

Dadas duas funções integráveis,  $f, g$  definidas no intervalo  $[a, b]$  podemos

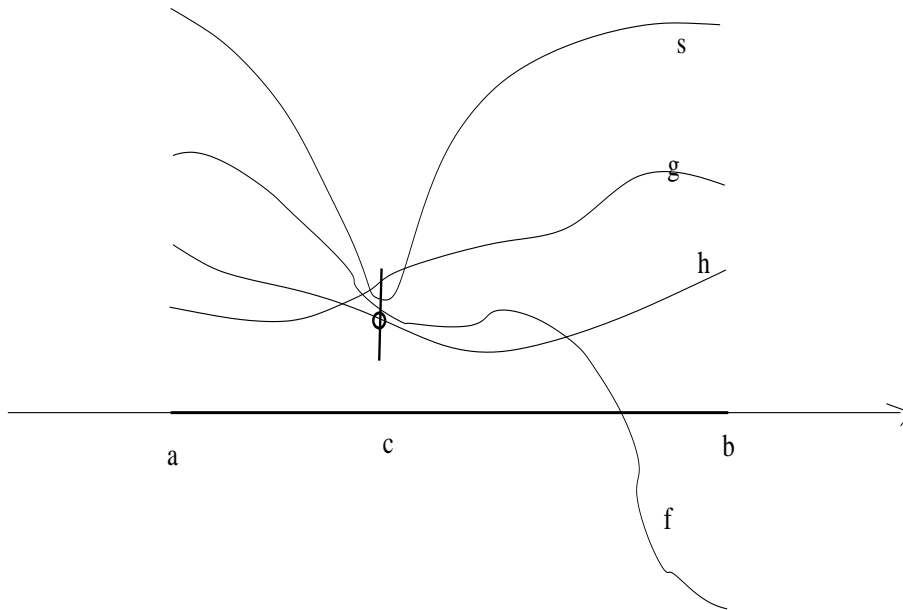


Figura 160: Distância pontual, da convergência pontual

definir

$$d(f, g) = \sqrt[p]{\int_a^b |f(x) - g(x)|^p dx} = \|f - g\|_p$$

com  $\infty > p \geq 1$ .

Se  $p < 1$  se perde a *desigualdade triangular* e dizemos que se trata de uma pseudo-métrica. As pseudo-métricas tem também as suas utilidades. Enquanto que as bolas, quando  $p \geq 1$  serão conjuntos convexos (onde vale a desigualdade triangular), as bolas das pseudo-métricas deixarão de ser convexas e é porisso que falha a desigualdade triangular.

Como no caso das distâncias entre pontos do  $\mathbf{R}^n$ , podemos definir a distância ou a norma do supremo

**Definição 27 (distância do supremo)** *entre funções*

Dadas duas funções limitadas,  $f, g$  definidas no intervalo  $[a, b]$  podemos definir

$$d_\infty(f, g) = \sup_{x \in [a, b]} |f(x) - g(x)| = \|f - g\|_\infty$$

com  $p = \infty$ .

A norma do *sup* é usada para definir *vizinhanças tubulares*, *convergência uniforme* e a *continuidade uniforme*.

Você encontra este conceito sendo usado em nível muito elementar, em Estatística, no *desvio padrão* que é uma aplicação distância-2 a uma massa de dados discreta.

Você pode se perguntar pela razão da variedade tão grande de topologias, e conseqüentemente, de tipos de convergência. Uma forma simples de justificar que é necessária esta variedade, vem com o seguinte exemplo. Suponha que desejemos medir a distância entre duas funções  $f, g$ , diferenciáveis,  $f, g \in \mathcal{C}^n([a, b])$ , tais funções não guardam a apenas informação do valor num determinado ponto  $c \in (a, b)$ , mas também das taxas de variação instantâneas em  $c$  até a ordem  $n$ , as sucessivas derivadas,

$$f(c), f'(c), \dots, f^{(n)}(c); g(c), g'(c), \dots, g^{(n)}(c);$$

e, para medir a distância entre  $f, g$  teremos que também incluir as distâncias entre suas derivadas o que nos leva a definir, por exemplo,

$$d_p(f, g) = \sum_{k=0}^n \sqrt[p]{\int_a^b |f^{(k)}(x) - g^{(k)}(x)|^p}; \quad (1170)$$

Agora foi feita uma medição mais fina das distâncias entre  $f, g$ , usando a norma  $p$  espalhada por todas as derivadas possíveis das funções  $f, g$ . Com supremo podemos definir

$$d_\infty(f, g) = \|f - g\|_\infty$$

se as derivadas de todas as ordens forem limitadas. Os espaços definidos usando estas métricas ou normas, levando em conta as derivadas, são chamados de espaços de Hardy com variantes chamados de espaços de Besov. Estes espaços aparecem na busca de soluções de equações diferenciais parciais.

Os exemplos de distância entre duas funções diferenciáveis, mostra que temos necessidade de distintos tipos de topologias, ou métricas para formalizar distintas situações, a distância entre soluções de equações diferenciais tem que ser medida usando uma norma- $p$  ou uma pseudo-distância- $p$ .

---

- **valor próprio** da Álgebra Linear, associado ao conceito de *vetor próprio*. Sinônimo: *autovalor*.

**Definição 28 (valor próprio)** *relativamente a um operador linear*

*Se diz que um vetor  $\vec{u} \neq 0$  é um vetor próprio relativamente a um operador linear  $T$  se houver um escalar  $\lambda$ , chamado valor próprio tal que*

$$T(u) = \lambda u$$

*Confira vetor próprio, equação característica*

---

- **vetor próprio** da Álgebra Linear, associado ao conceito de *valor próprio* ou *autovalor*. Sinônimo *autovetor*,

**Definição 29 (vetor próprio)** *relativamente a um operador linear*

Se diz que um vetor  $\vec{u} \neq 0$  é um vetor próprio relativamente a um operador linear  $T$  se houver um escalar  $\lambda$ , chamado valor próprio tal que

$$T(u) = \lambda u$$

Se  $u$  for um vetor próprio do operador  $T$  e  $\alpha \in \mathbf{R}$ , então  $\alpha u$  também será um vetor próprio, porque

$$T(\alpha u) = \alpha T(u) = \alpha \lambda u = \lambda(\alpha u)$$

o que gera uma aparente inconsistência na definição. Na verdade vamos ver que não é exatamente um vetor próprio que é a *informação* importante, mas sim o “espaço próprio de dimensão 1” que ele define, ou, como algumas vezes se diz, “as linhas próprias de um operador  $T$ ” para fazer referências aos espaços próprios de  $T$ .

O vetor nulo também responde à definição de “vetor próprio” e como os vetores nulos geram um subespaço “sem grande interesse”, precisamos eliminar esta possibilidade. Logo você vai ver que há outra razão importante para eliminar o vetor nulo: ele “destrói a base dos vetores próprios”.

Um vetor próprio representa uma *direção* do espaço em que o operador linear fica extremamente simples: *multiplicação por um escalar que é o seu valor próprio*. Confira *valor próprio*.

Uma alteração da equação na definição 29 vai nos levar a um método, (algoritmo), para encontrar os *valores próprios*.

$$T(u) = \lambda u = \lambda I(u);$$

em que  $I$  é o operador identidade, então podemos “colocar  $u$  em evidência”, escrevendo

$$T(u) = \lambda u = \lambda I(u) \Rightarrow T(u) - \lambda I(u) = 0; \quad (1171)$$

$$(T - \lambda I)u = 0 \Rightarrow \det(T - \lambda I) = 0; \quad (1172)$$

O *consequente*, a tese, na equação (eq. 1172), vem que de que precisamos de um vetor diferente de zero que seja solução da equação linear homogênea. Para tal é preciso que o determinante do sistema de equações lineares em (eq. 1172) seja nulo afim de que haja soluções não triviais: diferentes de zero.

A equação polinomial  $\det(T - \lambda I)$  é chamada de *equação característica* do operador linear  $T$ . O próximo teorema mostra a razão da exclusão do zero e a importância dos valores próprios e vetores próprios.

**Teorema 24 (indepêcia)** *dos vetores próprios*

Dado um operador linear  $T$ , se a equação (eq. 1172) tiver  $n$  soluções distintas,  $\{\lambda_1, \dots, \lambda_n\}$ , os correspondentes vetores próprios  $\{u_1, \dots, u_n\}$  de  $T$  formam um conjunto de vetores linearmente independentes, portanto uma base para espaço vetorial  $\mathbf{R}^n$ .

**Dem**:

Os valores próprios podem ser todos iguais o que equivale a dizer que  $T = \lambda I$  e neste caso qualquer que seja a base escolhida os vetores da base vão ser vetores próprios e portanto l.i..

Deixe-me agora supor que haja dois valores próprios diferentes,  $0 < \lambda_1 < \lambda_2$  e seus correspondentes vetores próprios,  $u_1, u_2$ , e escreva uma combinação linear nula não trivial,

$$\frac{\lambda_2}{\lambda_1} > 1; \text{ suponha } \alpha_1 \neq 0; \alpha_1 u_1 + \alpha_2 u_2 = 0; \quad (1173)$$

$$T(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 T(u_1) + \alpha_2 T(u_2) = \alpha_1 \lambda_1 u_1 + \alpha_2 \lambda_2 u_2 = 0; \quad (1174)$$

$$u_1 = -\frac{\alpha_2}{\alpha_1} \frac{\lambda_2}{\lambda_1} u_2; \quad (1175)$$

$$T(u_1) = \lambda_1 u_1 = -\frac{\alpha_2}{\alpha_1} \frac{\lambda_2^2}{\lambda_1} u_2; \quad (1176)$$

$$u_1 = -\frac{\alpha_2}{\alpha_1} \frac{\lambda_2^2}{\lambda_1} u_2; \quad (1177)$$

$$\frac{\lambda_2}{\lambda_1} = \frac{\lambda_2^2}{\lambda_1^2} \text{ impossível com } \frac{\lambda_2}{\lambda_1} > 1; \quad (1178)$$

Logo  $\alpha_1 = 0$ . Repetindo o argumento com a hipótese de que  $\alpha_2 \neq 0$  chega-se também a uma contradição portanto  $\alpha_1 = \alpha_2 = 0$  provando que  $u_1, u_2$  são l.i.. Resultado igual se obtém com a hipótese de  $0 < \lambda_2 < \lambda_1$ . Como os valores próprios podem ser negativos a demonstração deve ser corrigida usando o módulo de  $u_1$  valendo as mesmas contas.

Provei que os vetores próprios são l.i. dois a dois, o que implica na l.i. do conjunto de todos os vetores próprios.

**q.e.d.**

A hipótese de que a equação característica tenha  $n$  soluções não é necessária, o número de vetores próprios com frequência é  $m \leq n$  e formam um conjunto l.i.. Este “defeito” está ligado à incompletitude algébrica de  $\mathbf{R}$ , sobre o corpo  $\mathbf{C}$  dos complexos, a equação característica terá exatamente  $n$  soluções, entretanto com a possibilidade de raízes múltiplas trazendo uma outra informação sobre o espaço próprio associado às raízes múltiplas.

O vetor nulo  $\vec{0}$  satisfaz a esta definição para qualquer escalar, mas se for incluído tornaria l.d. o conjunto dos vetores próprios o que no fundo representa uma informação “desnecessária” que deve ser retirada do contexto..

A importância dos vetores próprios fica salientado neste caso particular em que  $T$  é um operador linear de  $\mathbf{R}^n$ :

**Teorema 25 (base) matriz de  $T$**

Se o operador linear  $T$  do  $\mathbf{R}^n$  tiver  $n$  vetores próprios a matriz de  $T$  relativamente à base dos vetores próprios é uma matriz diagonal tendo os valores próprios na diagonal.

**Dem**:

Basta escrever a matriz de  $T$  relativamente à base.

**q.e.d.**

Ou seja, os vetores próprios representam uma base em que o operador linear tem a matriz mais simples possível. Esta técnica é muito usada no estudo das equações diferenciais. Mesmo no caso em que os vetores próprios de  $T$  não se encontrem na quantidade  $n$  que corresponde à dimensão do espaço, ainda assim eles representam uma informação importante permitindo decompor uma matriz de  $T$  em blocos que salientam os sub-espacos que ficam invariantes por  $T$ .

O próximo teorema traz outro aspecto importante dos valores próprios.

**Teorema 26 (inversibilidade)** de um operador linear

Se o zero for solução da equação característica dum operador linear  $T$  então este operador não é inversível e reciprocamente, se  $T$  não for inversível então 0 é um valor próprio de  $T$ .

**Dem:**

Porque existe um vetor não nulo  $u$  tal que  $Tu = 0u = 0$ , ou ainda  $\text{Nuc}(T)$  é um espaço próprio.

Se  $T$  for inversível então sua matriz pode ser triangularizada e os valores próprios aparecem na diagonal desta matriz sendo  $\det(T) \neq 0$  o produto dos elementos da diagonal que não pode ser zero. **q.e.d.**

**Observação 9 (Notação:)** espectro de  $T$ 

O conjunto dos valores próprios de  $T$  é chamado de espectro de  $T$  e se usa a notação  $\sigma(T)$  para identificar o espectro de  $T$ .

Prese atenção de que excluímos o vetor zero na definição de vetor próprio porque ele atrapalharia a teoria. Mas não se exclui o zero como valor próprio e neste caso é uma informação importante: zero ser solução da equação característica.

Enquanto estivermos em dimensão finita, os valores próprios aparecem também em número finito menor ou igual à dimensão do espaço. Passando a espaços de dimensão não finita, por exemplo das ondas eletromagnéticas, e naturalmente considerando um operador  $T$  definido neste espaço, o espectro é um conjunto compacto dum espaço vetorial adequado e o estudo dos valores próprios se chama de *teoria espectral*. Aqui, novamente se vê a importância dos vetores próprios e dos valores próprios: o espectro de um operador linear é um espaço topológico sobre o qual podemos representar o operador linear sob a forma duma integral.

A teoria espectral é, assim, uma generalização da diagonalização de matrizes para o caso dos operadores integrais, ou dos operadores lineares que se possam expressar sob forma de integral cuja expressão mais simples seria sob a forma de uma integral sendo o espectro o *espaço medido* onde esta integral estará definida, “*domínio de integração*”. Se um operador integral  $T$  for inversível então o zero não pertence ao espectro de  $T$ ,  $\sigma(T)$ . A recíproca é verdadeira mas a demonstração é muito grande para este verbete... O caso mais intuitivo é o das matrizes  $n \times n$ , uma matriz diagonal (ou triangular) será inversível sse não houver zero na diagonal (no espectro).

---

- **Waring, Edward** um matemático inglês, (c. 1736 – 15 August 1798), que fez muitas conjecturas, em particular o chamado *problema de Waring* e cujo desenvolvimento matemático ficou todo contido num livro chamado *Miscellanea Analytica* publicado parcialmente em alguns momentos de sua vida como professor em Cambridge, no Magdalene College. Possivelmente suas pesquisas influenciaram Lagrange que finalizou algumas das suas conjecturas ligadas ao *problema de Waring*. Waring estudou Álgebra, teoria dos números, e Análise, interpolação polinomial mas era pouco sistemático sem a preocupação de publicar o que estudava. Possivelmente o chamado *algoritmo de interpolação de Lagrange* devesse ter o seu nome também associado.

- **wavelet** pode-se dizer que consiste de uma generalização dos *polinômios trigonométricos*, e como estes, é um exemplo de aplicação da *Álgebra Linear*.

As *wavelets* são vetores linearmente independentes e desta forma uma base para um espaço vetorial. Aqui os vetores são funções e o *objetivo* é obter *aproximações ótimas* de certas classes de funções pela seleção de uma base *apropriada*.

Se você concluir que esta é uma frase vazia, não estará em nada errada, mas foi este o meu objetivo: inicialmente chocá-la com o intuito de em seguida dar sentido aos aspectos difusos da frase anterior. Para isto preciso de um pouco da história do aparecimento das *wavelets*.

Observe que hoje vamos encontrar as raízes das *wavelets* no século 19 e certamente no século 18 se as considerarmos uma “alteração” dos polinômios trigonométricos. Com isto quero dizer que uma forma de analisar o surgimento das *wavelets* passa pela teoria de Fourier. Nesta forma de ver, se buscava uma alternativa às *ondas trigonométricas*  $\sin(kx)$ ,  $\cos(kx)$  para melhorar as comunicações das quais elas eram a *base*<sup>69</sup> até 1950 quando começaram a ser “alteradas” com *enjanelamento* para corrigir um *defeito básico*: elas são ondas que se propagam para o infinito com seu comportamento periódico o que atrapalhava sua função de captar sons com a finalidade de codificá-los para transmissão. Os sons, muito em particular a voz humana, tem um comportamento útil de “sinal”, quer dizer, uma função que se anula fora de um intervalo compacto.

Observe o significado do adjetivo “útil”, os sons se propagam indefinidamente no orbe, na Terra, uma vez que dependem de material que vibre para sua propagação, o ar. Mas do ponto de vista de quem os escuta, eles tem uma vida útil de suporte compacto, deixam de ser escutados depois de alguns segundos.

Na década de 80 diversos matemáticos desenvolveram experimentos em busca de alternativas para as *ondas trigonométricas*, entre eles o francês Y. Meyer, que criou o termo *ondelettes*, traduzido para o inglês como *wavelets* que se tornou o termo técnico predominante. A matemática americana, Daubechies, foi extremamente feliz em suas experiências descobrindo uma *wavelet* que melhor se adapta à percepção do ouvido humano e conseqüentemente a *wavelet* ideal para as comunicações.

Continuando a comparação com as *ondas trigonométricas*, temos uma base ortonormal de vetores com  $\sin(kx)$ ,  $\cos(kx)$  com um produto escalar adequado, confira *polinômios trigonométricos*. O parâmetro  $k$  corresponde à frequência das ondas eletromagnéticas ou sonoras. A invenção das *wavelets* se baseiou na busca de uma função  $\Phi$  que fosse à suporte compacto à qual se poderia aplicar dois tipos de parâmetros:

$$\Phi_{a,w}(x) = \Phi(w(x - a)); \quad (1179)$$

Os dois parâmetros,  $a$ ,  $w$  são chamados, respectivamente, de *translação* e *dilação*.

A função  $\Phi$  é a *onda mãe*, chamada em inglês, *mother wave*, e  $\Phi_{a,w}$  é uma

<sup>69</sup>E aqui a palavra “base” pode ser literalmente tomada no sentido da *Álgebra Linear*.

*translação-dilacionada* de  $\Phi$ . Desta forma podemos obter resultados superiores ao que seria possível com as “velhas ondas trigonométricas” uma vez que podemos fazer análises locais usando *translações-dilacionadas* da onda mãe.

A figura (fig 161), página 406, mostra o significado do enjanelamento em

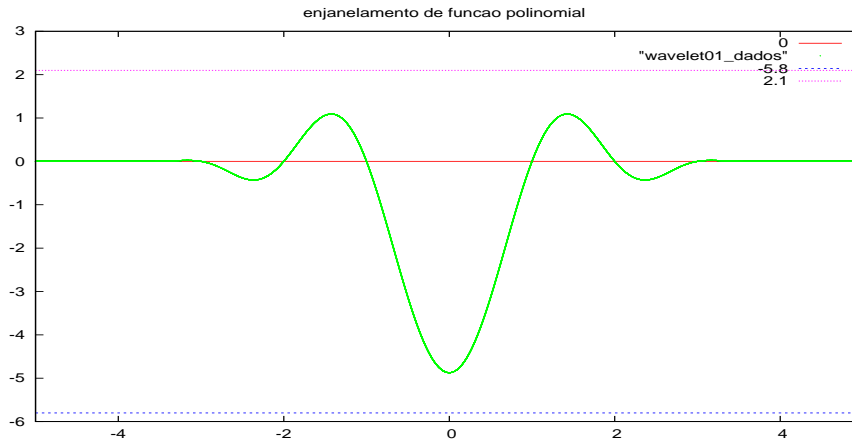


Figura 161: enjanelamento dum polinômio

que usei uma função à suporte compacto para enjanelar uma função polinomial no intervalo  $[-3, 3]$ , quer dizer que a função polinomial foi zerada fora deste intervalo e preservada dentro do mesmo intervalo. Observe que as funções polinomiais não constantes, crescem, ou decrescem arbitrariamente para grande valores de  $x$ , com o enjanelamento se “guarda” a parte do gráfico da função na região de enjanelamento.

A figura (fig 162), página 407, mostra uma *onda mãe* que é utilizada na figura (fig 163) para ilustrar uma análise local.

O gráfico na figura (fig 163), página 408, mostra a função polinomial enjanelada, uma *onda mãe*, e uma dilação-translação da *onda mãe* “dirigida” para analisar o que ocorre no ponto  $x = 1$ . Desta forma é possível analisar o comportamento local de um sinal, ou, no caso da voz humana, é possível captar melhor, em distintos momentos, os sinais da voz humana o que permite uma codificação e posterior decodificação, mais precisa e sobretudo, com uma menor quantidade de “levantamentos”, em suma, *mais barato* e *mais preciso*.

Há varios aspectos em que a Álgebra Linear é utilizada para produzir as nuances de um determinado fenômeno, o assunto específico são os *autovetores* e seus correspondentes *autovalores*. Comparando, a *onda-mãe* de Daubechies, na verdade suas dilações-translações, são os *autovetores* de melhor qualidade para captar os sons que possam ser escutados pelo ouvido humano.

É preciso corrigir o aspecto de “fora de uso” das “velhas ondas trigonométricas”, é verdade que, para as comunicações telefônicas, elas ficaram ultrapassadas, entretanto sempre que formos estudar um fenômeno eletromagnético, as “velhas ondas trigonométricas” se encontram no foco principal e é preciso não



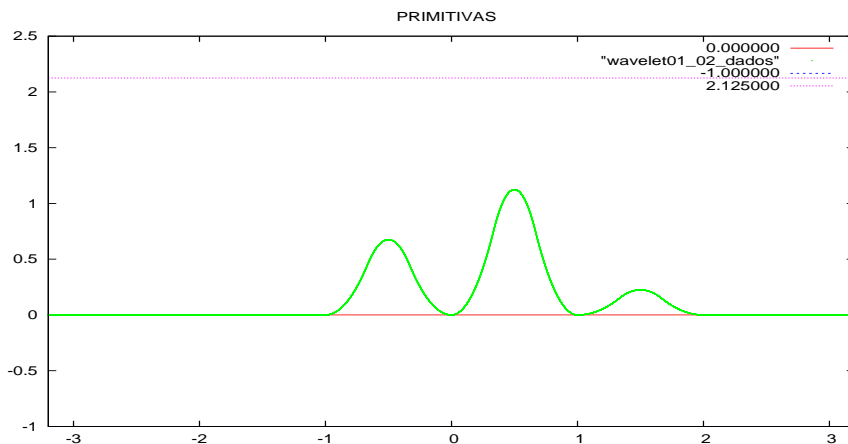


Figura 162: uma wavelet - onda mãe

esquecer que parte das comunicações dependem do campo eletromagnético para serem conduzidas. Desta forma contínua, eminentemente pedagógico, começar pela teoria dos polinômios trigonométricos, até porque é mais fácil produzir exemplos computacionais com eles para ilustrar a teoria. Assim existem duas grandes áreas a *análise harmônica* dedicada à teoria de Fourier, e *análise não harmônica* que é a extensão da *análise harmônica* para estudar as propriedades de ondas-mãe particulares.

- **Weierstrass, teorema de** Este teorema estabelece a densidade do espaço vetorial dos polinômios dentro do espaço vetorial das funções contínuas. Há duas versões deste teorema, a complexa e a real.

Confira também a generalização deste teorema, o teorema de Stone-Weierstrass, pela semelhança de métodos.

Vou começar pela versão real. Seja  $[a, b]$  um intervalo da reta e consideremos o *espaço vetorial real*  $(\mathcal{C}([a, b]), +, \cdot)$  das funções reais definidas e contínuas no intervalo  $[a, b]$ , em que a soma é a operação definida ponto a ponto e a multiplicação é o *produto por um escalar*. Como quero discutir a densidade dum subconjunto, dos polinômios, dentro do espaço de todas as funções contínuas, então preciso dum estrutura de espaço métrico e neste caso a *métrica usual* é a induzida pela *norma do supremo*:

$$\|f\|_{\infty} = \sup_{x \in [a, b]} |f(x)| \quad (1180)$$

a escolha desta métrica foi feita por Weierstrass uma vez que ele enunciou o teorema como “*aproximação uniforme de funções contínuas por polinômios*” o que significa usar o supremo sobre um intervalo para medir a distância.

As bolas abertas relativamente a esta norma são faixas ou tubos a figura (fig 164), página 409, mostra uma bola centrada na função identicamente zero e

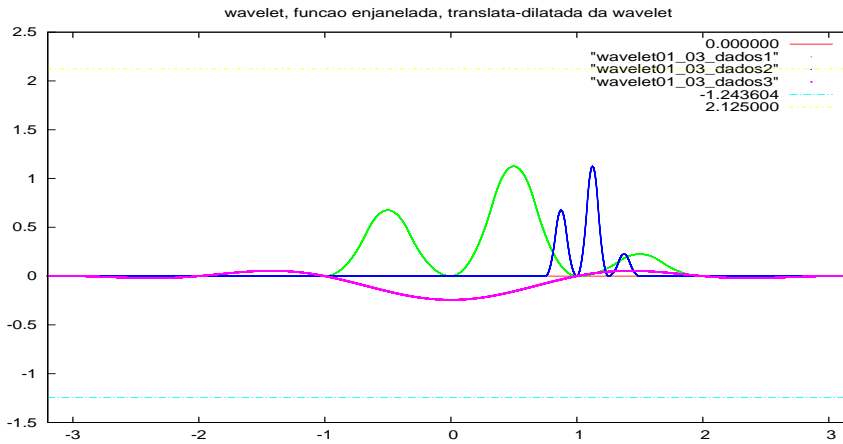


Figura 163: função enjanelada, wavelet, wavelet-translatada-dilatada

duas outras bolas disjuntas, centradas nas funções  $f, g$ . As bolas centradas em  $f$  e  $g$  foram obtidas por translação da bola  $\mathcal{B}(0, \epsilon)$  portanto são definidas por

$$\mathcal{B}(f, \epsilon) = \mathcal{B}(0, \epsilon) + f; \mathcal{B}(g, \epsilon) = \mathcal{B}(0, \epsilon) + g; \quad (1181)$$

Com estas ideias é muito fácil mostrar, construtivamente, a existência dum polinômio arbitrariamente próximo de zero no espaço  $\mathcal{C}([a, b])$  usando o polinômio auxiliar,  $P$ , da construção do algoritmo de interpolação polinômial de Lagrange, confira “*Lagrange, interpolação de*”,

$$P(x) = \prod_{k=0}^n (x - a_k); a = a_0; b = a_n; \quad (1182)$$

uma vez que sabemos, pelo fenômeno de Runge, que o máximo  $M$  deste polinômio ocorre nos intervalos extremos da partição de  $[a, b]$  produzida pelas raízes de  $P$ , então

$$\frac{\epsilon}{2M} P \in \mathcal{B}(0, \epsilon)$$

em que  $M$  é o máximo de  $P$  em  $[a, b]$ . A tentação seguinte seria a de transladar,  $\mathcal{B}(0, \epsilon) + f$ , para ter uma translação de  $P$  à volta de  $f$ . Mas

$$(P + f) \in \mathcal{B}(0, \epsilon) + f$$

e  $P + f$  não é um polinômio *viciando* de forma irremediável a demonstração ...

Mas posso ainda aproveitar parte da ideia contida na construção acima fazendo uma *deformação* no algoritmo de Lagrange, usar uma *versão alterada* do algoritmo do polinômio Lagrange, para construir um polinômio dentro duma bola de raio  $\epsilon$  centrada na função  $f$ , para um valor arbitrário do raio  $\epsilon$ , e vou incluir um programa de computador que irá gerar o gráfico.

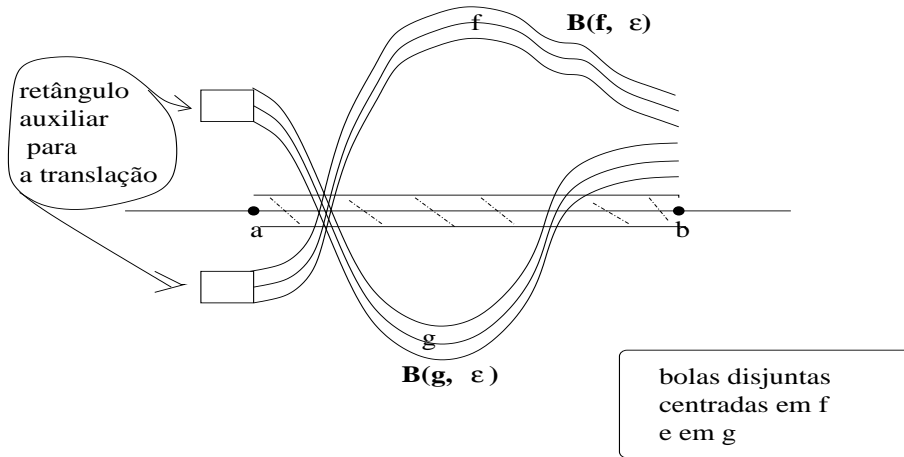


Figura 164: Bolas na norma do supremo em  $\mathcal{C}([a, b])$

Considere um intervalo auxiliar  $[A, B]$  e uma partição uniforme de  $[A, B]$ :

$$\Pi_n([A, B]) = \{x_0 = A, \dots, x_n = B\} \cup \{a, b\}; [a, b] \subset [A, B] \quad (1183)$$

em que a partição de  $[A, B]$  está representada pelos nós que formam uma progressão aritmética de pontos do intervalo  $[A, B]$  entre os quais incluí os dois extremos  $a, b$  do intervalo  $[a, b]$  para garantir um controle maior da oscilação sobre este intervalo.

Seja  $\tilde{f} \in \mathcal{C}([A, B])$  uma extensão contínua de  $f \in \mathcal{C}([a, b])$  e  $Q$  o polinômio de Lagrange interpolando os pontos

$$(x_0, \tilde{f}(x_0)), \dots, (a, \tilde{f}(a)), \dots, (b, \tilde{f}(b)), \dots, (x_n, \tilde{f}(x_n)),$$

que é o polinômio de grau  $n+2$  que interpola estes pontos. Como  $Q$  é uma função contínua, então sejam,  $m_Q, M_Q$ , respectivamente, o mínimo e o máximo de  $Q$  em  $[A, B]$ , que pelo fenômeno de Runge ocorrem, nos dois intervalos extremos da partição de  $[A, B]$  e  $R_Q = M_Q - m_Q$  a amplitude de  $Q$ .

Agora eu posso levar este polinômio para origem, como eu faria com qualquer vetor da Geometria Analítica para lá reduzir sua amplitude, mas vou levar junto uma vizinhança de  $Q$ :

$$\mathcal{B}(Q, \epsilon) - Q = \mathcal{V}(0); \quad (1184)$$

$$M \in \mathcal{V}(0); M((x_0) \cdots = M(a) = \dots M(b) = \dots M((x_n)); \quad (1185)$$

que vou deformar homogeneamente dividindo pela amplitude de  $Q$  ainda multiplicada por  $\frac{\epsilon}{2}, \frac{\epsilon}{2R_Q}$

$$\frac{\epsilon}{2R_Q} \mathcal{V}(0) = \mathcal{V}_\delta(0) \quad (1186)$$

Considere agora o polinômio de Lagrange,  $Q$ , interpolando os nós da partição considerada,

$$(x_0, 0), \dots, (x_n, 0)$$

e sejam,  $m_Q, M_Q$ , respectivamente, o mínimo e o máximo de  $Q$  em  $[a, b]$  e  $R_Q = M_Q - m_Q$  a amplitude de  $Q$ , e  $r = \min\{\frac{1}{R_P}, \frac{1}{R_Q}\}$  então  $M = \epsilon r Q \in \mathcal{B}(0, \epsilon)$  e  $M = \epsilon r Q$  é um polinômio que se anula em todos os nós da partição considerada.

Então  $M = \epsilon r Q + P$  também interpola os pontos

$$(x_0, f(x_0)), \dots, (x_n, f(x_n))$$

Considere agora as seguintes operações sobre os vetores  $f, P$ :

$$\mathcal{B}(0, \epsilon); \tag{1187}$$

$$Q = P - f \text{ translação de } P \text{ por } f; \tag{1188}$$

$$Q_1 = \epsilon Q/R; \Rightarrow Q_1 \in \mathcal{B}(0, \epsilon); \tag{1189}$$

$$\tag{1190}$$

estou usando uma partição uniforme mas neste contexto você pode ignorar esta restrição que será usada apenas pelo programa com que vou produzir os próximos gráficos.

Demonstrei assim o teorema:

**Teorema 27 (densidade)** *dos polinômios em  $\mathcal{C}([a, b])$  O conjunto dos polinômios, como funções definidas em  $[a, b]$  é denso no espaço vetorial normado  $(\mathcal{C}([a, b], \mathbf{R}), \|\cdot\|_\infty)$  das funções reais definidas em  $[a, b]$  com a norma do supremo.*

### - Wiener, álgebra de

O espaço das funções definidas por uma série trigonométrica absolutamente convergente é fechado para somas e produtos ponto a ponto e para o produto por um escalar (complexo ou real, são dois casos). Pensando no caso complexo, a notação é  $W(\mathbf{T})$  em que  $\mathbf{T}$  é o grupo dos números complexos de módulo 1, porque as funções assim definidas podem ser restringidas ao círculo unitário.  $W(\mathbf{T})$  é um espaço de Banach isomorfo a  $l^1$ . O produto ponto a ponto de  $W(\mathbf{T})$  corresponde ao produto de convolução em  $l^1$  o que mostra um defeito na álgebra de Banach de Wiener que não tem uma unidade, a função identidade, que seria esta unidade, pode ser aproximada arbitrariamente por séries trigonométricas absolutamente convergentes, sendo portanto um ponto de acumulação com uma topologia convenientemente definida.

Os dois espaços de Banach,  $W(\mathbf{T}), l^1$  são isomorfos e podemos trazer para  $W(\mathbf{T})$  um resultado que é fácil de ser demonstrado em  $l^1$  e agora expresso com a notação da álgebra de Wiener:  $\|fg\| \leq \|f\| \|g\|$  este produto ponto a ponto se transforma em  $l^1$  num produto por convolução de sucessões valendo a mesma desigualdade com a correspondente expressão. Esta desigualdade é chamada de *desigualdade de Wiener*.

# Índice Remissivo

- $\aleph_0$ , 230
- $\aleph_1$ , 230
- $A(T)$ , 362
- $\mathbf{C}$ , 41
- $\mathcal{C}([a, b])$ 
  - $\mathcal{C}((a, b))$ , 226
- $\mathcal{C}^n$ , 58
- $\mathcal{C}(\mathbf{R})$ , 226
- $\mathcal{L}^2([0, 2\pi])$ , 334
- $\mathcal{L}^2([a, b])$ , 229
- $e$ , 110, 260
- $GL_n(K)$ , 214
- $GL_n(R)$ , 214, 215
- $i$ , 41
- álgebra, 301
  - comutativa, 372
- álgebra de Wiener, 362
- álgebra linear, 364, 382
  - octave, 280
  - scilab, 280
- ângulo, 5
  - centesimal, 6
  - central, 7
  - dum vetor, 250, 251
  - hexadecimal, 6
  - inscrito, 7
  - radiano, 6
- ângulo soma
  - coseno, 393
  - seno, 393
- área
  - agricultura, 251, 252
  - superfície, 252
- árvore, 397
- índice
  - ponto, curva, 278
- índice,  $Ind_\gamma(a)$ , 21
- $Im$ , 42
- $Ind_\gamma(a)$ , 21, 234
- $\infty$ , 263
- $l^1$ , 395
- L<sup>A</sup>T<sub>E</sub>X, 3
- $\mathbf{S}^{n-1}$ , 76
- $l^1$ , 395, 410
- $l^2(\mathbf{N})$ , 334
- $M_n(K)$ , 214
- modn*, 48
- $\mathbf{N}$ , 298, 350
- $\bar{\partial}$ , 90
- $\partial$ , 90
- $\partial\Omega$ , 153
- $\pi$ , 9, 258, 374, 376
- $\pi(x)$ , 339
- $\mathbf{Q}$ , 185, 350
  - representação geométrica, 187
- $Re$ , 42
- $\sigma$ -álgebra, 286
- $\sigma(T)$ , 404
- $Sim(n)$ , 212, 214
- $W(\mathbf{T})$ , 410
- Abel, Lema de, 4
- aberta
  - aplicação, 89
- absolutamente convergente, 4
- absolutamente somável, 4
- aceleração, 58
- acumulação
  - ponto de, 56
- adição
  - geométrica, 352
- afinar, 302

- algébrica
  - computação, 172
  - curva, 77
  - geometria, 122, 132, 144, 156, 201, 205
  - variedade, 205
- Algébrica
  - Computação, 85
- algébrico
  - número, 266
- algébrico, número, 4
- algorítmica
  - diferenciação, 86
- algoritmo, 4
- alternada
  - forma multilinear, 93
- analítica, 89
  - função, 22, 90, 231, 305
  - geometria, 5, 77, 105, 122, 132, 143–145, 156, 159, 191, 197, 206
- analíticas
  - funções, 88
- anel, 231
  - matrizes, 280
- anel quociente, 231
- análise
  - harmônica, 407
  - não harmônica, 407
- análise combinatória, 10, 12, 26
- análise matemática, 301
- anecedente, 9, 53
- antípodas, 386
- aproximada
  - unidade, 64, 69
- aproximação, 9, 243
  - função, 405
- aritmética
  - computação, 49
- arranjo, 9, 10, 12, 26, 314
  - com repetição, 12
- arranjos
  - simples, 34
- artificial
  - inteligência, 86, 239
- assembler, 251
- assíntota, 138
- assintótico
  - comportamento, 363
- atlas, 12, 13, 242
  - círculo, 13
  - mapa, 13
  - mundi, 13
- autovalor, 13, 17, 401, 406
- auto valor, 318
- autovetor, 13, 406
- auto vetor, 318
- axioma
  - da escolha, 172, 318
- axioma da escolha, 18
- Baire
  - categorias de, 13
  - teorema de, 13
- Banach
  - álgebra, 395
  - espaço, 361
  - espaço de, 14, 363
- Banach-Steinhaus
  - teorema de, 14
- baricentro, 284
- base, 98
  - dimensão, 99
- bem ordenado, 307
- Bernoulli, 365, 384
- Bernoulli, Daniel, 329
- bilinear, 345, 346
- binários, 49
- binomial
  - coeficiente, 15
- bloco
  - matriz, 403
- bola
  - convexa, 400
  - não convexa, 400
- Bolzano, 374, 377
  - continuidade, 54
- borelianos, 15
- Bounjakowsky
  - Cauchy-Schwarz-, 230
- Bourbaki, N, 15, 110
- braquistocrona, 191

- Brouwer, 16  
 BVP, 21  
     PVF, 303, 342  
 BVP, PVF, 118
- cabelo  
     oscilador, 24
- cadeia, 307
- cadéia  
     regra da, 86
- calc, 348  
     programação, 112
- calculistas, 264
- calor  
     equação do, 119  
     transmissão, 48, 61, 348
- calor, difusão, 117
- caminho  
     integral depende do, 209  
     integral independente do, 209  
     integral que depende do, 154
- campo  
     escalar, 16  
     gravitacional, 357  
     vetorial, 16
- campos  
     conservativos, 209  
     não conservativos, 209
- Cantor  
     conjectura, 18, 50  
     hipótese de, 40, 230
- característica  
     equação, 17, 401, 402
- cardinalidade, 17, 39, 318, 337  
     do contínuo, 40  
     salto, 18, 230  
     salto de, 337  
     saltos, 40
- Cartan, 19  
     Henri, 88
- cartesiana  
     equação, 147  
     equação da reta, 147
- cartesianas  
     coordenadas, 191
- categoria, 18, 61, 297, 300  
     computação, 19  
     EV, 263  
     Grupo, 296  
     teoria das, 296  
     Teoria das, 19
- Cauchy  
     critério de, 261, 375, 376  
     integral de, 77, 237  
     integral de , 21  
     problema de, 119  
     Schwarz-Bounjakowsky, 230  
     sequências de, 354  
     sucessão de, 63  
     teorema de, 21  
     teste de, 63
- Cauchy, Augustin-Louis, 109
- Cauchy-Bunyakovsky-Schwarz  
     desigualdade, 226, 227  
     desigualdade, 108, 199
- Cauchy-Kovalesvkaya, 20
- Cauchy-Riemann, 4, 231  
     equações, 22, 277, 305  
     equações de, 23, 89, 90
- centro  
     dum grupo, 213
- centro de massa, 284
- Chasles  
     lei de, 97
- Churyumov  
     Gerasimenko, 149
- ciclo, 23  
     duma permutação, 315  
     permutação, 316  
     produto, 283  
     produto de, 316
- cálculo, 16  
     multivariado, 322  
     régua, 264
- Cálculo  
     Diferencial, 54  
     Integral, 54
- círculo, 24, 76, 156  
     corda, 255  
     equação do, 121  
     trigonométrico, 159, 249  
     unitário, 159

- círculo trigonométrico, 392
- código aberto, 383
- côncavo
  - polígono, 328
- cônicas, 145, 191, 202
- classe do um, 260
- classe do zero, 260
- classes
  - de equivalência, 169, 186
- classes quociente, 168, 232
- clima
  - EDP, 119
- codificação, 127
- coeficiente
  - angular, 150, 196
  - linear, 150, 196
- coeficiente angular, 78
- coeficiente binomial, 24
- Cohen
  - Paul, 337
- Cohen, Paul, 230
- combinação, 26, 224
- combinatória, 10, 12, 26
  - análise, 314
  - lei de Stifel, 28
- combinatório
  - número, 224
- cometa
  - Churyumov–Gerasimenko, 149
- cometa Ison, 36
- compacto, 242
  - localmente, 174
- complexa
  - exponencial, 111, 271
  - função linear, 21
  - logaritmo, 271
- complexas
  - variáveis, 305
- complexidade, 39
- complexo
  - argumento, 179
  - conjugado, 46
  - forma polar, 46, 272
  - número, 40, 179, 183
  - produto, 46
  - repres. geom., 42
- comprimento, de curva, 190
- comprimento de onda, 23
- computacional
  - álgebra linear, 94, 283
  - matemática, 172
- computador
  - assistido, 372
- Computação
  - Algébrica, 85
- comutatividade
  - séries, 4
- comutativo
  - corpo, 50
- condição inicial, 113, 114, 153, 338
  - problema, 19
- condições
  - de fronteira, 302
  - iniciais, 302
- condições de contorno, 154
- condução, 48, 61, 348
- cone, 48
  - diretriz, 134
  - eixo, 131, 134
  - geratriz, 131
  - visibilidade, 360
- conforme
  - transformação, 274
- congruência, 48, 168, 339
  - módulo  $n$ , 281
- conjectura, 50, 230
  - Cantor, 18, 50
  - Fermat, 50
  - Poincaré, 50
  - sucessões
    - suporte finito, 70
- conjunto, 50
  - agregado, 51
  - elemento, 52
  - finito, 17
  - infinito, 17
  - lista, 51
  - vazio, 52
- conjunto convexo, 284
- conjunto mesurável, 286
- conjuntos
  - ferramenta, 51



- linguagem, 51
- conjuntos, teoria, 39
- consequente, 9, 53, 402
- conservativo
  - campo vetorial, 153
- constante
  - de Euler, 339
- contínuo
  - hipótese do, 337
  - hypótese, 40, 230
- continuidade, 61
  - Bolzano, 54
  - classe de, 58
  - global, 55
  - integral de Riemann, 226
  - local, 55
  - sequencial, 59, 60
    - função real, 60
- Continuidade, 54
- contorno
  - integral que depende do, 154
- convecção, 48, 61, 348
- convergência, 61, 260, 350, 354
  - absoluta, 4
  - comutatividade, 4
  - ponto a ponto, 398
  - raio de, 349
    - teste da raiz, 349
    - teste do quociente, 349
  - sucessão, 60
  - uniforme, 398
- convexa
  - variedade, 165
- convexo, 284
  - conjunto, 284
  - fecho, 284, 285
  - polígono, 328
- convolução, 300, 334
  - aproximação, 66
  - discreta, 69, 71
  - produto, 21, 231, 238
  - produto de, 395
  - produto por, 410
  - regularização por, 65, 350
- convolução, produto de, 63, 109
- coordenadas cartesianas, 72
- coordenadas esféricas, 72
- coordenadas polares, 72
- Copêrnico, 122
- coprimo, 340
- corpo, 186, 227
  - dos complexos, 88
  - dos números complexos, 232
  - finito, 50
- coseno, 380
  - lei, 181
- cosenos
  - soma de, 393
- cosenos diretores, 76, 165
- crescimento
  - curva, 114
- criptografia, 169, 339
- critério
  - de Cauchy, 261, 375, 376
- curva, 76, 120
  - de Jordan, 78
  - de Peano, 76
  - filling curves, 76
  - melhor adaptada, 341
  - número de voltas, 6
- curva algébrica, 77
- curva, comprimento, 190
- curva retificável, 190
- curvas
  - homotópicas, 311
- d'Alembert, 329
- Darwin, 54
- Darwin, C., 60
- Daubechies
  - wavelet, 405
- Debian/GNU/Linux, 3
- decodificação, 127
- De Moivre, 393
- dependentes
  - linearmente, 97
- derivada, 78, 114, 358
  - complexa, 88, 277
  - da função composta, 86
  - do produto, 85
  - do quociente, 86
  - funções polinomiais, 85

- integral, 338
  - primitiva, 338
  - regra do produto, 247
- derivada implícita, 77
- derivada parcial, 91
- derivadas parciais, 17
- derivação
  - regras de, 85
- Descartes, 191
- descontinuidade
  - salto, 55
- desencriptar, 170
- desenvolvimento
  - curvas de, 114
- desigualdade
  - artimético-geométrica, 236
  - Cauchy-Bunyakovsky-Schwarz, 226, 393
  - de Hölder, 230, 298
  - triangular, 63, 400
- desigualdade de Hardy, 92
- desigualdade de Wiener, 410
- desvio padrão, 401
- determinante, 92, 93, 281, 282, 317, 368
- determinantes
  - ordem 3, 282
- determinante, 367
- diagrama
  - Hasse, 224
- dicionário, 350
- difeomorfismo, 13
- diferenciabilidade, 94, 326
- diferenciação
  - algorítmica, 86
- diferenciais
  - equações, exemplos, 154
- diferenciais, equações, 384
- diferencial
  - equação, 152, 360
  - geometria, 12, 94, 322
  - topologia, 94
- diferencial, equação, 152
- diferenciável, 95
- difusividade
  - térmica, 348
- dilação, 405
- dilações, 365
- dimensão, 95, 98, 395
  - base, 99
  - infinita, 98, 99, 102
  - quarta, 121
  - segunda, 120
  - terceira, 120
  - um, 120
  - zero, 119
- dinâmico
  - sistema, 123, 135
- dinâmicos
  - sistemas, 115
- Dirac
  - delta de, 63
  - distribuição de, 64
  - função de, 63
  - medida, 64
  - medida de, 372
- Dirac, medida, 100
- Dirac, Paul Adrien Maurice, 109
- Dirac, Paul A. M., 63
- Dirichlet
  - núcleo, 65, 68, 100, 300, 301, 304
  - núcleo de, 65, 301, 334, 365
  - problema de, 119
  - teorema de, 340
- discreta
  - convolução, 63, 69
- dispersão padrão, 289
- distância, 100–102, 156, 237, 398
  - entre dois pontos, 100, 105, 197
  - entre funções, 399, 400
  - euclidiana, 100, 101, 341
  - plano à origem, 104
  - reta à origem, 102, 104
  - reta origem, 164
  - variedade linear
    - à origem, 105
- distribuição, 108, 286
- distribuições
  - teoria das, 64, 238
  - Teoria das, 155
- divergência, 260
- dividendo, 109

- divisão
  - euclidiana, 109
  - algoritmo da, 109
  - geométrica, 355
- divisões da Matemática, 383
- divisor, 109
- divisor de zero, 280, 372
- divisores de zero, 231
- domínio, 110
- dual
  - de um espaço, 231
- Du Bois-Reymond, 364
  
- EDO, 112, 152
- EDP, 117, 152, 154
  - clima, 119
  - EDPL, 304
- EDP, lineares, 155
- EDPL, LPDE, 119
- EDP, não lineares, 155
- Eilenberg, 19
- eixos coordenados, 119
- elemento
  - conjunto, 52
- elementos finitos, 155
  - triangulação, 372
- eletromagnética
  - onda, 98
- elipse, 76
  - equação da, 122
  - excentricidade, 130
  - excentricidade, 122
- encriptar, 170
- endomorfismo
  - canônico, 213
- energia
  - solar, 252
- enumeráveis, 17
- enumerável, 99
- equação
  - Cauchy-Riemann, 23
  - círculo
    - trigonométrico, 200
  - da onda, 119
  - da reta, 143
  - de Laplace, 91, 118
  - de Poisson, 119
  - do calor, 118, 119
  - do plano, 159
  - funcional, 112
    - segundo grau, 40
  - equação algébrica, 115
  - equação da elipse, 122
  - equação da hipérbole, 134
  - equação diferencial, 152, 342, 384
    - linear, 115
    - não linear, 115
    - ordem, 152
    - ordem  $n$ , 342
    - ordinária, 152, 155, 338
    - parcial, 114, 152, 156, 304, 338
    - problema, 342
    - solução computacional, 118
  - equação do círculo, 131, 156
  - equação normal
    - reta, 163
  - equação polar
    - reta, 165
  - equações
    - de Cauchy-Riemann, 22
    - de Maxwell, 119
    - de Navier-Stokes, 119
    - diferenciais, 365
    - hiperbólicas, 155
    - lineares, 281
    - parabólicas, 155
  - equações diferenciais, parciais, 172
- equipolência, 250
- equivalência, 167
  - classe, 186
  - classe de, 63
  - classes, 169
  - homotópica, 312
  - módulo tangência, 380
  - relação, 77
  - relação de, 48, 186, 324, 351
- Eratóstenes, 169
  - crivo de, 173, 338
- erro
  - no dicionário, 60–62, 153, 168, 296
- escada
  - função, 175

- escalar
  - produto, 227, 330, 344, 364
  - propriedades, 345, 346
- escolha
  - axioma da, 18
- esfera, 171
  - dimensão  $n - 1$ , 76
- esféricas
  - coordenadas, 72
- espaço
  - Banach, 361
  - completo, 226
  - com produto interno, 107, 198
  - de Besov, 401
  - de Hardy, 401
  - de sucessões, 71, 102
  - métrico, 225, 398
  - topológico, 60, 398
  - vetorial topológico, 225
- espaço de Hausdorff, 224, 373
- espaço de Hilbert, 365
  - reproduzindo kernel, 229
- espaço métrico, 101, 106, 198, 288, 398
  - de Hausdorff, 225
- espaço-tempo, 396
- espaço vetorial, 171
- espaço vetorial topológico, 171
- espectral
  - teoria, 404
- espectral, teorema, 382
- espectro, 24, 382
  - operador, 404
- esquecer
  - funtor, 394
- essencialmente, 15
- estatística
  - Bayesiana, 341
- estrutura
  - ordem, 306
- euclidiana
  - geometria, 105, 197
- Euclides, 338
- euclidiana
  - divisão, 109
  - geometria, 160, 182, 230
  - norma, 289
- Euler, 267, 365, 384
  - fórmula, 6, 111, 179, 180, 271, 272
  - fórmula de, 46, 329
- Euler, Leonhard, 329
- avaliação
  - morfismo de, 297
- exata
  - equação diferencial, 154
  - Matemática, 300
- existência
  - teorema de, 172
- falso
  - verdade, 397
- fatoração, 172
  - de morfismos, 298
- Fermat
  - último teorema, 248
  - último teorema de, 206
  - o último teorema de, 50, 174
- fórmula
  - de Bhaskara, 40
  - de Euler, 6
  - de Euler, 179
  - Euler, 111, 214
  - trigonométrica, 179
- Fields
  - medalha, 50
- figura
  - $\mathcal{L}^1$ , 101
  - ângulo, 6, 107, 199
  - ângulo central, 7
  - ângulo inscrito, 8
  - 1-splines, 369
  - 30-spline, 370
  - aproximação
    - por convolução, 69
  - arco, soma, 182
  - base hexadecimal, 25
  - binômio de Newton, 25
  - bolas
    - distância  $p$ , 102
    - norma do sup, 409
  - cardinal, 19
  - Cauchy, critério de, 376
  - círculo, 132, 156

- trigonométrico, 346
- círculo unitário, 392
- cônicas, 125, 136, 205, 206
- coeficiente angular, 148, 151
- coeficiente linear, 148
- cometa, 37, 39
- complexa
  - exponencial, 273
  - logaritmo, 273
- complexo
  - conjugado, 47
  - produto, 42
- cone
  - a elipse, 204
  - a hipérbole, 203
  - a parábola, 203
  - visibilidade, 359
- cone de duas folhas, 49
- conjunto
  - interseção, 52
- continuidade, 54
- convexo
  - fecho, 285
- convolução, 64, 65
  - Potência de, 70
- coordenadas, 194
  - polares, 73
- coseno
  - lei, 183
- cosenos diretores, 166
- curva, 78
- curva adaptada, 295
- curva de crescimento, 114
- curva plana, 201
- curva tangente, 358
- derivada, 79, 88
- descontinuidade, 57
- diagrama
  - Hasse, 225
- Dirichlet
  - núcleo, 68
- distância, 100
  - entre dois pontos, 106, 197
  - origem, 104
  - ponto a ponto, 400
- divisão
  - geométrica, 355
- dois vetores, 96
- domínio, 110
- eixos coordenados, 121
- elipse, 123, 130
- elipse no cone, 127
- equação da reta, 103, 150
- equação diferencial, 113, 114
  - solução, 117
- equivalência de triângulos, 391
- frações, 193
- frações equivalentes, 186
- função inversa, 240
- Green, teorema de, 210
- hipérbole, 135, 141–144
- hipérbole plana, 139
- impulso sonoro, 303
- integral, 237
- Ison, 36
- Lagrange
  - interpolação, 243–245
- lei
  - cosenos, 249
  - senos, 255
- Lei
  - seno, 257
- limite, 258
- Logaritmo, 276
- logaritmo complexo, 274, 275
- massa
  - centro de, 285
- matriz de pontos, 293
- Medida de Dirac, 287
- medidas da elipse, 128
- média
  - convexidade, 285
  - ponderada, 284
  - segmento de reta, 284
- multiplicação
  - geométrica, 353
- multiplicação geométrica, 188
- núcleo
  - solução particular, 368
- números complexos
  - geometria, 44
- obsângulo, 390

- onda, 303
- ordem parcial, 306
- pantógrafo, 308, 309
- Pascal
  - triângulo, 30
- perpendicular, 161
- plano cartesiano, 121
- polares
  - coordenadas, 73
- polígono côncavo, 328
- poligonal, 327
- Polinômio de Taylor, 381, 382
- polinômio trigon., 68
- produto
  - cartesiano, 344, 345
  - geométrico, 356
  - vetorial, 347
- produto de convolução, 66
- programa, 202
- PVF, 369
  - visão experimental, 369
- quadrados mínimos, 294, 341
- racionais e a reta, 187
- raízes da unidade, 329
- raíz quadrada, 354
- regra
  - paralelograma, 45
- regularização
  - fitting curve, 341
  - por convolução, 67
- reta
  - numérica, 192, 352
  - tangente, 359
- reta, equação, 150
- reta no plano, 146
- reta orientada, 193
- reta tangente, 80, 81, 357, 358
- régua
  - cálculo, 265
- Seno
  - Lei, 256
- sobrejeção, 19
- soma de frações, 185
- spline, 369
- splines, 371
- tabela
  - verdade, 379
- taxa de variação, 83
- teorema
  - função implícita, 234
- teorema de Pitágoras, 323
- topologia, 312
- triângulo
  - de Pascal, 71
- trigonométrico
  - polinômio, 332
- triângulo isósceles, 386, 389
- triângulos
  - semelhança, 385
- Venn
  - diagrama de, 397
- verdade
  - tabela, 379
- vetor normal, 164
- vizinhança de zero, 373
- vizinhança tubular, 399
- wavelet
  - dilação, 408
  - janela, 406
  - onda mãe, 407
- figure
  - atlas
    - círculo, 13
- filtro
  - crescente, 174
- finita
  - indução, 299
- finito
  - conjunto, 17
  - corpo, 50
- finitos
  - elementos, 372
- fitnito
  - conjunto, 336
- fitting curve, 288, 340, 341
- flúido
  - dinâmica, 119
- forma
  - bilinear, 227
- forma multilinear, 93
- Formula de McLaurin, 380
- Fourier, 329, 365

- coeficientes, 68, 229, 300, 362
- coeficientes de  $f$ , 68
- série, 229, 300, 373
- séries, 228, 300
- séries de, 98, 330
- teoria de, 405
- transformada, 65, 304
- transformada de, 64
- transformação discreta, 394
- Fourier, séries de, 384
- Fourier, transformada de, 182, 384
- fraca
  - convergência, 69
  - solução, 109
- fração, 182
  - imprópria, 182
  - irredutível, 184
  - própria, 182, 352
- frações
  - equivalentes, 186
- frequência, 98
- fronteira, 153, 154, 209, 303
  - $\partial\Omega$ , 153
- FSF, 3
- funcional, 189
- funcional linear, 190
- funcional não linear, 190
- função, 188
  - aritmética, 189
  - bijetiva, 189
  - conjuntos dos valores, 188
  - contradomínio, 188
  - de Euler, 189
  - domínio, 188
  - generalizada, 155
  - injetiva, 188
  - sobrejetiva, 189
  - totalmente aritmética, 189
- função analítica, 90
- função constante, 188
- função de Dirac, 100
- função de probabilidade, 286
- função generalizada, 189, 286
- função harmônica, 91
- função implícita
  - teorema da, 115, 234
- função integrável, 237
- função linear
  - complexa, 21
- funções
  - teoria das, 90
- funções analíticas, 4
- funtor, 18, 189, 242
  - esquecido, 18, 297
- fuzzy
  - lógica, 378
- Gauss
  - plano de, 186
  - teorema da divergência, 208
- gaussiana, 75
- Gelfand
  - transformada de, 363
  - transformada de, 64
- Gelfand, transformada de, 382
- geometria
  - analítica, 143, 159, 191
  - diferencial, 94
  - euclidiana, 160, 161
  - não euclidiana, 144, 146, 161
  - números complexos, 44
- geometria algébrica, 132, 144, 156, 206
- geometria analítica, 77, 132, 144, 145, 156, 206
- geometria euclidiana, 182
- geral linear
  - grupo, 214
- Gerasimenko
  - Churyumov, 149
- geratriz
  - do cone, 204
- germe, 380
- Gibbs
  - fenômeno de, 68
- GNU, 206
- gnuplot, 80, 146, 206, 348
- Gödel, 15, 337
  - teorema, 172
- gordo, 13
- GPL, 280
- gradiente, 207
- grafico

- números complexos, 43
- grafo, 327, 397
- grafos
  - teoria dos, 327
- Gram-Schmidt
  - ortonormalização de, 167
- Gram-Schmidt, 207
- grau, 6
- gravitação
  - campo, 357
- Green
  - teorema, 154
  - teorema de, 208, 238
  - teorema de , 153
  - Teorema de, 338
- grupo, 186, 210, 300
  - $Sim(n)$ , 211
  - centro, 213
  - clases
    - mod  $n$ , 215
  - classe lateral, 215
  - codificação, 212
  - comutativo, 213
  - conjugação, 212, 215
  - das rotações, 214
  - de matrizes, 281, 320
  - dos inteiros, 213
  - endomorfismo, 212
  - estrutura algébrica, 314
  - finito, 214, 315
  - geral linear, 214
  - grupo quociente, 215
  - inteiros mod  $n$ , 215
  - isomorfismo, 212
  - matrizes, 214
  - não comutativo, 215, 314, 365
  - normal, 215
  - numérico, 214
  - ordem, 211, 215
  - permutação, 212
  - permutações, 211, 214, 282
  - quociente, 63, 216, 223
  - representação, 215
  - representação de grupos, 212
  - rotação
    - matrizes, 214
  - subgrupo, 211, 212
    - normal, 216
    - próprio, 212
  - topológico, 223
  - transformações lineares, 366
- guitarra, 302
- Hardy, desigualdade de, 92
- harmônica
  - análise, 407
  - função, 91, 305
- harmônico
  - complemento, 305
  - conjugado, 91, 305
- Hasse
  - diagrama, 224, 397
- Hausdorff
  - espaço, 224
  - exemplo negativo, 225
- Hilbert, 16, 354
  - 23 problemas, 230
  - espaço de, 14, 63, 226, 361, 364
    - reproduzindo kernel, 229
  - Space, 226
- hiperplano, 163
  - equação, 145, 159
- hipervolume, 176
- hipérbole, 76
  - degenerada, 126, 138
  - equação da, 122
- hipótese, 230
  - do contínuo, 18
- hipótese de Cantor, 230
- hipótese do contínuo, 40, 230
- Hölder
  - desigualdade de, 230
- holomorfa
  - função, 90, 231
- holomorfas, 4
- homotópicas
  - curvas, 311
- homogênea
  - equação diferencial, 304
- homomorfismo, 296
- homotopia
  - equivalência, 278, 312



- html, 224
- HTML, 224
- ideal, 231
  - à direita, 231
  - à esquerda, 231
  - bilateral, 231
  - lateral, 231
- ideal maximal, 261, 375, 377
- ideias básicas, 51
- igualdade, 351
- imaginária
  - número complexo
    - parte imaginária, 42
- imaginária
  - raíz, 41
  - raiz, 41
  - unidade, 41
- imaginário
  - complexo, 41
  - puro, 41
- IMPA, 383
- implicação, 9, 53
- implicação lógica, 168
- implícita
  - derivada, 207
  - teorema da função, 232
- impulso positivo
  - função, 302
- independentes
  - linearmente, 97
- independência
  - linear, 97
- indução
  - finita, 298
- indução finita, 34, 173, 235, 299
  - teorema da, 236
- infinita
  - dimensão, 98, 99
  - precisão, inteira, 381
- infinitésimo, 87, 88, 262
- infinito, 98
  - clases de, 99
  - conjunto, 17
- inicial
  - condição, 113
- inútil
  - Matemática, 362
- input, 309
- integração
  - a Daniel, 238
  - a Lebesgue, 238
  - a Riemann, 238
  - a Riemann-Stieltjes, 238
- integral, 237
  - de Cauchy, 237
  - dependente do caminho, 338
  - derivada, 338
  - independente do caminho, 338
  - Riemann–Stieltjes, 176
  - Riemann–Stieltjes, 286
  - valor médio, 284
- integral de linha, 153, 238
- integravel a Riemann, 238, 239
- integrável
  - à Riemann, 176
- inteira
  - precisão infinita, 381
- inteligência artificial, 86, 239
- interno
  - produto, 344
- interpolação, 243, 244
  - Lagrange, 409
  - linear, 284
- intuicionismo, 16
- invariante
  - subespaço, 403
- invariância, 139
- inversa
  - da derivada, 237
  - de uma matriz, 281
- inversão, 212
- inversível
  - matriz, 367
- irracional
  - número, 354
- isometria, 334
- isomorfismo, 242, 297, 300, 334, 395
  - de grupos, 317–319
- Ison, 36
- jacobiana, 16, 17, 91, 232, 242

- matriz, 89
- Jordan
  - curva de, 77
- Kepler, 122
- kernel, 66, 242, 300, 304
- Knuth, Donald, 383
- Kovalesvkaya, 20
- Lagrange
  - interpolação de, 404
  - o erro de, 161
  - polinômio, 408
- Lagrange, interpolação de, 242
- Landau, 380
- Laplace
  - operador, 304
- Laplace, equação de, 117
- laplaciano, 91, 248, 304
- latex,  $\text{\LaTeX}$ , 383
- Laundau
  - o pequeno de, 241
- Laurent Schwartz, 286
- Lebesgue
  - espaços, 373
  - espaços de, 330
  - integração á, 286
  - integral, 238
  - integral de, 177
  - media, 330
  - medida de, 330
- lei
  - coseno, 181, 248, 253
  - superfícies, 251
  - seno, 254, 256
- Leibniz
  - Gottfried, 54
  - notação de, 326
- Leibniz, G., 60
- Leibnz
  - notação, 116
- lei de Chasles, 97
- Lema de Abel, 4
- lógica, 239, 337
  - fuzzy, 378
  - operações básicas, 380
- lógica fuzzy, 397
- lógica matemática, 301, 378, 397
- limite, 54, 56, 257
  - inferior, 56, 263
  - notável, 86, 112, 260
  - propriedades, 261, 262
  - superior, 56, 263
- limite inferior, 262
- limite superior, 263
- linear, 346
  - álgebra, 281
  - EDP, EDPL, 155
  - equação diferencial, 115, 119, 368
  - funcional, 231
  - função, 263
  - independência, 97
  - interpolação, 284
  - operador, 119, 368
  - sistema, 367
  - sistema de equações, 281, 368
  - variedade, 159
- linear complexa
  - função, 264
- linear, funcional, 190
- linha
  - integral de, 153
- linha, integral de, 238
- Linux, 173
  - factor, 173
- LISP, 381
- logaritmo
  - natural, 267
- logaritmo, 214, 264, 271
  - argumento, 272
  - complexo, 181, 271
  - decimal, 266
  - domínio, 267
  - fórmula fundamental, 270
  - integral, 267
  - morfismo, 296
  - neperiano, 270
  - p.a., 264
  - p.g., 264
  - precisão, 266
- logaritmos
  - tabelas, 267

- magro, 13
- make, 3
- mal posto
  - problema, 340
- mapa
  - dum atlas, 13
- mapas, 242
- markup language, 224
- massa
  - centro de, 284, 285
  - distribuição de, 285
- Matemática
  - divisões, 110, 383
- Mathematics
  - Subject Classification
    - MSC, 110
- matriz, 278
  - bloco, 403
  - cofatora, 281
  - diagonal, 403
  - equação, 279
  - função, 281
  - identidade, 281
  - inversa, 281
  - operação elementar, 318, 319
  - quadrada, 280
  - semelhança, 17
  - transposta, 281
  - triangular, 94
  - triangularizar, 283
- maximal, 232, 307
- Maxwell
  - equações de, 119
- McLaurin
  - Formula de, 380
- mecânica quântica, 109
- medalha Fields, 50
- medida, 190, 237, 286
  - Dirac, 64
  - espaço de, 229, 300
  - euclidiana, 288
- medida de Dirac, 100, 286
- medidas
  - teoria das, 177, 238
- medidas, teoria das, 191
- melhor curva
  - busca da, 288, 340
- menor inteiro
  - $\lceil \sqrt[N]{N} \rceil$ , 259
- mesurável
  - conjunto, 177
- mesurável,conjunto, 286
- Meyer
  - ondelettes, 405
- mãe
  - onda, 405
- máquina
  - de Turing, 300
- máximo, 307
- média, 283
  - aritmética ponderada, 283
  - aritmética simples, 283
  - caixa d'água, 286
  - fila única, 286
  - integral, 284, 285
  - ponderada, 284
  - probabilidade, 286
- método
  - história, 271
- métrica, 101, 288, 398
  - pseudo, 400
- métrico
  - espaço, 55, 101, 106, 198, 225, 398
  - espaço discreto, 288
- métrico, espaço, 288
- mínimo, 307
- mínimos quadrados
  - caso polinomial, 293
- mínimos, quadrados
  - linear, 288
- módulo
  - produto, 180
- módulo n
  - aritmética, 50
- música, 302
- minimal, 307
- Miscellanea Analytica, 404
- modeling language, 395
- modelo, 244, 296
- modular

- álgebra, 322
- Moebius, fita, 383
- morfismo, 296, 300, 395
  - complexo, 297
  - de avaliação, 297
  - de grupo, 296
  - espaço vetorial, 263
  - fatoração canônica, 298
  - grupo, 181
  - logartimo, 296
  - núcleo, 368
- morfismos, 368
- movimento, 58
- movimentos rígidos
  - rotação, 139
  - translação, 139
- multíndice, 91
- multilinear
  - álgebra, 281
- multiplicação, 297
  - à direita, 298
  - à esquerda, 298
  - geométrica, 353
  - operador, 231, 297, 298
- museu, 271
- Napier
  - John, 264
- naturais
  - números, 298
- Navier-Stokes
  - equações de, 119
- neperiano
  - logaritmo, 111
- Newton
  - binômio de, 393
- não linear
  - equação diferencial, 115
- não lineares, EDP, 155
- não linear, funcional, 190
- núcleo, 64, 66, 242, 297, 300
  - de Cauchy, 21, 231, 238
- núcleo de Dirichlet, 68, 301
- núcleo dum morfismo, 368
- número, 301
  - complexo, 40
    - parte imaginária, 42
    - parte real, 42
  - real, 350, 352, 354
- número algébrico, 302
- número complexo, 302
- número natural, 302
- número racional, 302
- números
  - naturais, 298
  - reais, 63
  - teoria dos, 189
- números binomiais, 24
- números complexos, 301
- números inteiros, 301
- números naturais, 301
- números racionais, 301
- números reais, 301
- número transcendente, 302
- noção primeira, 51
- noções básicas, 51
- nonsense
  - general abstract, 19
- norma, 14, 394
  - do sup, 400
  - do supremo, 244
- normado
  - espaço vetorial, 14, 225, 394
- norma do supremo, 372
- norma euclidiana, 289
- normal
  - vetor, 163
- notação de Leibniz, 116
- notável
  - limite, 111
- n-splines, 68
- numerizado
  - plano, 120
- numérica
  - análise, 94
  - reta, 120, 351
- numérico
  - cálculo, 94, 283
  - plano, 120
- objeto
  - programação a, 19

- octave, 280, 283, 319
- onda
  - comprimento, 23
  - eletromagnética, 98
  - equação da, 119, 302
- ondelettes, 405
- OOP
  - object o. programming, 297
- operador, 190
  - multiplicação, 211, 297, 298
- operadores
  - teoria dos, 300
- operadores, teoria dos, 190
- operador integral, 304, 382
- ordem, 305, 306
  - alfabética, 306
  - EDO, 152
  - EDP, 152
  - estrita, 305, 306
  - larga, 305, 306
  - parcial, 305
  - reflexiva, 305
  - relação, 305
  - relação de, 193
  - total, 305, 307
- ordinária
  - equação diferencial, 152, 155, 342
- ordinárias
  - equação diferencial, 152
- origem
  - dos eixos, 200
- ortonormalização
  - Gram-Schmidt, 167
  - Gram-Schmidt, 208
- oscilador
  - cabelo, 24
  - pêndulo, 24
- oscilação, 56
- Oscilação, 307
- output, 309
  
- pantógrafo, 308, 314
  - calibragem, 309
  - comercial, 308
  - furos, calibragem, 309
  - topologia, 311
- virtual, 311
- parciais, equações diferenciais, 172
- parcial
  - derivada, 91
  - equação diferencial, 152, 156, 338, 342
- paridade
  - permutação, 317, 318, 322
- parábola, 76
  - equação da, 122
- Parseval
  - identidade, 334
- partes
  - conjunto das, 39
- Pascal
  - triângulo, 24, 337
  - triângulo de, 28, 337
- Pascal, Blaise, 182
- Peano, 230
  - Giuseppe, 298
- Perelman, Grigori, 50
- permutação, 10, 26, 33, 283, 314, 366
  - circular, 366
  - computação, 317, 319
  - impar, 317
  - par, 317
  - paridade, 317
  - representação matricial, 317
  - sinal, 283, 317
  - transposição, 283, 317
- permutações, 217, 282, 365
  - grupo, 314
- PFF, 304
- piano, 302
- pêndulo
  - oscilador, 24
- pitagóricos
  - números, 174, 322
- Pitágoras
  - teorema, 44, 392
- plano
  - digitalizar, 193
  - equação, 145, 159
  - equação do, 122
  - feixe de planos, 125, 137
  - numerizar, 193

- tangente, 207
- plano cartesiano, 121, 322
- Plano Cartesiano, 119
- plano coordenado, 194
- Plano Coordenado, 119
- planos
  - classes, 361
  - concorrentes, 361
  - paralelos, 361
  - reversos, 361
- plano tangente, 322
- Poincaré
  - conjectura, 50, 383
- polar
  - coordenadas, 326
- polares
  - coordenadas, 72
- polígono, 327
  - irregular, 308
  - subpolígono, 308
- poligonal, 326
  - ângulos, 327
  - 1-splines, 371
  - lados, 327
  - vértices, 327
- polinômio, 98
  - de Taylor, 329
  - Lagrange, 409
  - produto, 70
  - trigonométrico, 329
- polinômio de Taylor, 380
- polinômios
  - trigonométricos, 405
- polo, 278
- polos
  - raízes, 23
- ponto, 120
- POO
  - programação o. objetos, 297
- potencial, 153
- potência
  - dum conjunto, 336
  - por convolução, 70
  - séries de, 349
- Potência, 334
- precisão
  - pontos de, 244
- preço, de terreno, 190
- própria
  - linha, 402
- próprio
  - espaço, 402
- primitiva, 114, 153, 237, 337
- primo, 338
  - gêmeo, 169
  - número, 172, 174
- primo com, 189
- primos
  - entre si, 340
- princípio
  - indução finita, 299
- principal
  - valor, 109
- prisão
  - tridimensional, 395
- problema, 118, 340, 384
  - condição inicial, 19
  - de Cauchy, 19, 119, 340
  - de Dirichlet, 119
  - equação diferencial, 342
  - mal posto, 340
  - PCI, 19
  - PVF, 303
  - valor inicial, 342
  - valor na fronteira, 342, 369
    - visão experimental, 343
  - variacional, 191
- produto
  - cartesiano, 343
  - ciclos, 283
  - comutativo, 92
  - de convolução, 63
  - de matrizes, 280
  - escalar, 107, 198, 248, 344
  - geométrico, 254
  - interno, 107, 198, 344
  - não comutativo, 93
  - por convolução, 64, 67
  - vetorial, 347
- produto de convolução, 109, 395
- produto escalar, 364
  - definição geométrica, 251

- produto interno
  - espaço com, 107, 198
- programa, 4
  - teorema
    - de Weierstrass, 408
- programação, 348
  - calc, 348
  - gnuplot, 348
  - python, 348
- progressão
  - aritmética, 264
  - de grau  $m$ , 236
  - geométrica, 264
- projeção, 331, 364
- proporção, 167, 186
- proporções
  - lei das, 145
- prova
  - dos 9, 168
- pulso, 287, 288
- PVF, 21
  - BVP, 303, 342
- PVF, BVP, 118
- Python, 381
  
- quadrado integrável
  - funções, 334
- quadrados
  - dos mínimos, 341
- quações
  - elípticas, 155
- quantidade
  - de calor, 237
  - de movimento, 237
- quantidade de movimento, 398
- quaternions, 183
- quase
  - sempre, 350
- quase sempre, 15
- quase-sempre
  - nula, 330
- Quântica
  - Física, 286
- quociente
  - teste
    - convergência, 349
- racional
  - número, 182
- Rademacher, função de, 109
- radiação, 48, 61, 348
- radiciação, 348
- raíz, 278
  - teste
    - convergência, 349
- raízes
  - número complexo, 329
  - polos, 23
- raio de convergência, 349
- Ramanujan, 362
- reais
  - números, 63
- real
  - número, 193, 350, 354
  - número complexo
    - parte real, 42
- redundante, 350
- reflexividade, 168, 351
- reflexivo
  - espaço, 361
- regra
  - da cadéia, 86
- regras de derivação, 85
- regularização, 350
  - convolução, 22
  - por convolução, 65, 67, 350
  - Tikhonov, de, 341
- relação de equivalência, 351
- relação de ordem, 231
- representação
  - geométrica, 43
- representação
  - dum grupo, 317, 318
  - em  $\mathbf{S}^1$ , 254
  - linear, 219
    - teoria da, 366
  - matricial de  $\text{sim}(n)$ , 319
- reproduzindo kernel
  - espaço de Hilbert, 229
- resíduos, 49
- resto
  - adição módulo 19, 171
  - equação, 170

- na divisão, 281
- produto módulo 19, 171
- tabela de adição, 170
- tabela de multiplicação, 169
- reta, 351
  - coeficiente angular, 149
  - coeficiente linear, 149
  - digitalizar, 192
  - equação, 143
    - paramétrica, 147
    - vetorial, 147
  - equação normal, 163
  - equação polar, 165
  - geometria euclidiana, 144
  - geométrica, 193
  - numerizar, 192
  - numérica, 193, 350–352
  - proporção, 145
  - segmento de, 120
  - tangente, 149
- reta numérica, 194
- retas
  - concorrentes, 361
  - coplanares, 361
  - paralelas, 360
  - reversas, 360, 361
- reta tangente, 94, 95, 356
- retificável, curva, 190
- régua
  - de cálculo, 264
- rígido
  - modelo hexagonal, 362
- Riemann
  - integração á, 286
  - integral, 373
  - integral de, 175
  - soma, 175
- Riemann–Stieltjes, 191
- Riemann, integrável a, 361
- Riesz
  - representação, 229
  - representação de, 361
- Riesz-Fischer
  - teorema, 334
- Roger
  - Leonard James, 362
- Roger-Ramanujan
  - equações, 362
- Rosetta, 123, 135
  - nave espacial, 149, 165
- rotação, 139
- rule
  - slide, 264
- Runge
  - fenômeno, 294
  - fenômeno de, 408, 409
- saca rolha
  - regra, 107, 198
- salto, 362
  - de cardinalidade, 337
  - descontinuidade, 55
  - finito, 55
  - infinito, 55
  - ponto de, 55
  - primeira espécie, 362
  - segunda espécie, 362
- saltos, 60
- Sarrus
  - regra de, 282
- Schwartz
  - Laurent, 108, 155
- Schwartz, Laurent, 238
- Schwarz-Clairaut
  - derivadas mistas, 305
- scilab, 280, 283, 319
- Sebastião Silva, 286
- semelhantes
  - triângulos, 169
- semi-espaco, 163
- seno, 380
- sequencial
  - continuidade, 60
- sequências de Cauchy, 239
- ser vivo
  - equação diferencial, 114
- série, 362
  - assintótico
    - comportamento, 363
  - de Fourier, 229
  - termo geral, 363
- série de Fourier, 364



- séries de Fourier, 384
- séries-q, 362
- série trigonométrica, 394
- Silva
  - Sebastião e, 108, 155
- Silva, Sebastião e, 238
- sim(3), 319
- sim(9), 320
- simetria, 168, 351
- sim(I)
  - sim(n), 318
- sim(n), 314, 318, 365
  - sim(I), 318
- simples
  - função, 304
  - funções, 177
- sinal
  - permutação, 317
- sistema
  - equações lineares, 94
- sistema linear, 366
- sistemas dinâmicos, 115
- slide
  - rule, 264
- somas de Riemann, 239
- somável, 369
- spline
  - desenho automatizado, 372
  - equação diferencial, 372
- splines, 68, 246, 296, 326, 369, 370, 372
  - não polinomiais, 372
  - por convolução, 68
- Stieltjes, 176
- Stieltjes, Riemann, 191
- Stifel
  - lei de, 28
- Stokes
  - teorema de, 208
- Stone-Weierstrass
  - teorema, 372
  - teorema de, 373
- subgrupo, 63
- sucessão
  - convergente, 63, 226
  - de Cauchy, 226, 261
  - divergente, 63
- sucessões, 374
  - limite, 178
- sucessões de Cauchy, 376
- superfície, 373
- suporte, 65
  - compacto, 405
  - finito, 70
- suporte finito, 70
- supremo
  - norma do, 244, 372, 407
- tabela
  - verdade, 378
- tangente, 380
  - função linear, 95
  - plano, 322, 326
  - reta, 356
  - variedade linear, 326
- tangência, 324
  - relação de equivalência, 77, 95
- Taylor
  - polinômio, 296, 380
  - polinômio de, 79
- Taylor, polinômio, 329
- teorema
  - da aproximação
    - de Weierstrass, 407
  - da função implícita, 19
  - da função implícita, 115, 172, 234
  - da Função Implícita, 232
  - da função inversa, 239, 240, 242
  - da representação de Riesz, 230
  - de existência, 232
  - de Gödel, 172
  - de Green, 23, 153, 154, 208
  - de Jordan, 78
  - de Mertens, 339
  - de Rolle, 172
  - de Weierstrass
    - programa, 408
  - dos números primos, 339
  - espectral, 382
  - Fermat, o último, 322
  - Fermat-Wiles, 322
  - fundamental da Álgebra, 338
  - fundamental do Cálculo, 153, 154

- função implícita, 253
- identidade
  - de Parseval, 334
- Pitágoras, 248, 249
- Riesz-Fischer, 334
- Schwartz-Clairaut, 208
- Schwartz-Clairaut, 91
- Stone-Weierstrass, 373
- valor médio da derivada, 172
- teoria
  - das distribuições, 64, 155, 238
  - das medidas, 238
  - dos conjuntos, 39
  - dos números, 322
- teoria das medidas, 191
- teoria espectral, 404
- terreno, preço, 190
- TeX, T<sub>E</sub>X, 383
- térmica
  - difusividade, 348
- Tikhonov
  - regularização, 341
- topológico
  - espaço, 60, 223
  - espaço vetorial, 171, 225
  - grupo, 223
- topologia, 50, 311, 383, 397
  - compacto-aberta, 226
  - diferencial, 94
  - fraca, 226
  - induzida, 226
  - tubular-compacta, 226
- Topologia, 54
- topologia usual, 60
- transcendente, número, 383
- transferidor, 6, 344
- transformada de Fourier, 384
- transitividade, 168, 351
- translação, 139, 405
  - fator de, 370
- translações, 365
- triângulo
  - de Pascal, 71
- tricotomia, 305
- triedro
  - positivo, 107, 198
- trigonometria, 392
  - ângulo soma, 393
  - círculo unitário, 392
  - coseno
    - soma, 393
  - fórmulas, 393
- trigonométrica
  - série, 394
- trigonométricas
  - séries
    - absolutamente convergentes, 14, 394
- trigonométrico
  - círculo, 5, 200
  - polinômio, 68
- trigonométrico, polinômio, 329
- trigonométricos
  - polinômios, 98, 228, 405
- triângulo
  - de Pascal, 182
- triângulo de Pascal, 24, 28
- triângulos
  - classificação, 384
  - semelhantes, 384
- tubular
  - vizinhança, 400
- Turin, Alan, 169
- Turing, 300
  - Alan Mathison, 239
  - máquina de, 239
- Tycho Brae, 122
- UML, 395
- unidade
  - aproximada, 64, 301
  - convolução, produto, 64
  - raízes da, 328, 329
- unidades aproximadas, 65
- uniforme
  - continuidade, 400
  - convergência, 400
  - limitação, 15
- unitária
  - matriz, 320
- unitárias
  - matrizes, 281

- usual
  - topologia, 60
- valor
  - lateral, 55
- valor médio
  - integral, 284, 285, 334
- valor próprio, 17, 401, 402
- variacional
  - Cálculo, 289
- variacional, problema, 191
- variedade, 76, 77, 205, 233, 324, 395
  - dim 2, 160, 373
  - dim 2, 160
  - linear tangente, 165
  - maximal, 163
  - tangente, 232
- variedade algébrica, 205
- variedade linear
  - hiperplano, 159
  - plano, 159
  - reta, 159
- variedades
  - dimensão 1, 396
  - dimensão 2, 396
  - lineares, 396
  - não lineares, 396
  - topologia, 12
- variedade tangente, 77
- variáveis aleatórias, 296
- velocidade, 58
- verdade
  - falso, 397
  - tabela, 378
- vetor
  - coluna, 281
  - linha, 281
- vetores
  - orientação, 107, 198
- vetorial
  - espaço, 171
- vetor próprio, 384, 401, 402
- vida
  - comprimento médio, 24
- violão, 302
- visibilidade
  - cone, 360
- vizinhança, 80, 101, 397
  - tubular, 398, 400
- voltas
  - número de, 77
- Waring
  - Edward, 404
  - problema de, 404
- wavelet, 365, 405
  - de Daubechies, 405
  - dilação, 405
  - translação, 405
- wavelets, 67, 372
- Weierstrass
  - teorema de, 407
- Wiener
  - álgebra de, 297, 395, 410
  - álgebra de, 362
  - desigualdade de, 410
  - Lema de, 395
- Wiles
  - Andrew, 206
- winding
  - number, 77
- xfig, 146
- Zermelo-Fraenkel, 230, 337
- zero
  - divisor, 280
  - divisor de, 280, 372
- Zorn
  - lema, 318
  - lema de, 318
- Zygmund, 330



# Referências Bibliográficas

- [1] American Mathematical Society. 2010 mathematics subject classification. <http://www.ams.org/mathscinet/msc/msc2010.html>.
- [2] R. C. Boyce, William E e DiPrima. *Equações diferenciais elementares e problemas de valores de contorno*. Editora: LTC - ISBN-13: 9788521614999, 2006.
- [3] J. Dieudonné. *Calcul Infinitésimal*. Herman Éditeurs, 1968.
- [4] S.V. Gelfand, I.M. e Fomin. *Calculus of variations*. Dover, 2000.
- [5] Richard Courant. *Differential and Integral Calculus I*. Interscience Publishers Wiley classics library, 1988.
- [6] T Praciano-Pereira. Programas para cálculo numérico. Technical report, 2009. <http://www.calculo-numerico.sobralmatematica.org/programas/>.
- [7] T Praciano-Pereira. Programando em gnuplot. Préprints da Sobral Matemática no 2008.1 - 2008 [http://www.sobralmatematica.org/preprints/programando\\_gnuplot.pdf](http://www.sobralmatematica.org/preprints/programando_gnuplot.pdf), 01 2008.
- [8] Tarcisio Praciano-Pereira Stálio Rodrigues dos Santos. *Introdução à Matemática Universitária*. Sobral Matemática, 2009.
- [9] G.F. Simmons. *Introduction to Topology and Modern Analysis*. McGraw-Hill - Book Company, 1968.
- [10] G.F. Simmons. *Differential Equations with Applications and Historical Notes*. McGraw-Hill - Book Company, 1979.
- [11] Stephen Smale Morris W. Hirsch. *Differential Equations, Dynamical Systems, and Linear Algebra*. Academic Press, 1974.
- [12] the free encyclopedia in the Internet Wikipedia. Wikipedia, the free encyclopedia in the internet. <http://www.wikipedia.org>.
- [13] Wikipedia. *Wikipedia, a free encyclopedia*. <http://pt.wikipedia.org/wiki/>.

- [14] the free encyclopedia in the Internet Wikipedia. Wikipedia, the free encyclopedia in the internet. [http://en.wikipedia.org/wiki/Joseph\\_Fourier](http://en.wikipedia.org/wiki/Joseph_Fourier).
- [15] Foundation for Free Software. Gpl - general public license. Technical report, <http://www.FSF.org>, 2011.