ENGFIS FIS
2017/18

Complementos de Cálculo e de Geometria Analítica

Salvatore Cosentino
Departamento de Matemática e Aplicações - Universidade do Minho
Campus de Gualtar - 4710 Braga - PORTUGAL
gab: CG - Edifício 6 - 3.48, tel: 253 604086
e-mail scosentino@math.uminho.pt
url http://w3.math.uminho.pt/~scosentino

22 de Maio de 2018

Resumo

This is not a book! These are notes written for personal use while preparing lectures on "Complementos de Cálculo e de Geometria Analítica" for students of FIS and MIENGFIS during the a.y.'s 2013/14, 2014/15 and 2017/18. They are rather informal and certainly contain mistakes (indeed, they are constantly actualized). I tried to be as synthetic as I could, without missing the observations that I consider important.

Most probably I will not lecture all I wrote, and did not write all I plan to lecture. So, I included sketched or even paragraphs, about material that I think should/could be lectured within the same course, given enough time.

References contain some introductory manuals that I like, some classics, books where I have learnt things in the past century, recent books which I find interesting. Almost all material can be found in the Calculus book [Ap69] by Apostol. More classical references are Lang's [La87, La97] for linear algebra, and [BDP92] for differential equations. Good material and further references can easily be found in the web, for example in Scholarpedia, in Wikipedia, or in the MIT OpenCoureWare.

Everything about the course may be found in my web pages

`http://w3.math.uminho.pt/~scosentino/salteaching.html`

The notation is as follows:

e.g. means EXEMPLI GRATIA, that is, "for example", and is used to introduce important or (I hope!) interesting examples.

ex: means "exercise", to be solved at home or in the classroom.

ref: means "references", places where you can find and study what follows inside each section.

Orange paragraphs form the main text.

Blue paragraphs deal with applications and ideas relevant in physics, engineering or other sciences. They are the real reason why all this maths is worth studying.

Red paragraphs (mostly written in english) are more advanced or non trivial facts and results which may be skipped in a first (and also second) reading.

□ indicates the end of a proof.

Pictures were made with *Grapher* on my MacBook, or taken from Wikipedia, or produced with `Matlab` or `Mathematica`®`8` .

Conteúdo

Notações

Conjuntos. $a \in A$ quer dizer que a é um elemento do conjunto A. $A \subset B$ quer dizer que o conjunto A é um subconjunto do conjunto B. $A \cap B$ é a interseção dos conjuntos A e B, e $A \cup B$ é a reunião dos conjuntos A e B. $A \times B$ é o produto cartesiano dos conjuntos A e B, o conjunto dos pares ordenados (a, b) com $a \in A$ e $b \in B$.

Números. $\mathbb{N} := \{1, 2, 3, \dots\}$ denota o conjunto dos números naturais. $\mathbb{Z} := \{0, \pm 1, \pm 2, \pm 3, \dots\}$ denota o anel dos números inteiros. $\mathbb{Q} := \{p/q \text{ com } p, q, \in \mathbb{Z}\,,\ q \neq 0\}$ denota o corpo dos números racionais. $\mathbb{R}$ e $\mathbb{C}$ são os corpos dos númeors reais e complexos, respetivamente.

Funções. Uma *função* $f : X \to Y$, com *domíno* o conjunto X e *conjunto de chegada* o conjunto Y, é um subconjunto $R \subset X \times Y$ tal que para cada $x \in X$ existe um único $y := f(x) \in Y$, dito *imagem* de x, tal que $(x, y) \in R$. Quando domínio e contradomínio são claros, uma função pode ser denotada apenas por $x \mapsto f(x)$, ou seja, identificada com a "regra" que determina $y = f(x)$ a partir de x. A *imagem* do subconjunto $A \subset X$ é o conjunto $f(A) := \{f(a) \text{ com } a \in A\} \subset Y$. Em particular, a *imagem/contradomínio* da função $f : X \to Y$ é o conjunto $f(X) := \{f(x) \text{ com } x \in X\} \subset Y$ dos valores da função. O *gráfico* da função $f : X \to Y$ é o subconjunto

$$\mathrm{Graph}(f) := \{(x, y) \in X \times Y \text{ t.q. } y = f(x)\} \subset X \times Y$$

do produto cartesiano do domínio e o conjunto de chegada. A função *identidade* $I_X : X \to X$ é definida por $I_X(x) = x$, e o seu gráfico é a *diagonal* $\{(x, x) \text{ com } x \in X\} \subset X \times X$.

A *restrição* da função $f : X \to Y$ ao subconjunto $A \subset X$ é a função $f|_A : A \to Y$ definida por $f|_A(a) := f(a)$.

A *composição* das funções $f : X \to Y$ e $g : f(X) \subset Y \to Z$ é a função $g \circ f : X \to Z$ definida por $(g \circ f)(x) := g(f(x))$, ou seja,

$$x \mapsto y = f(x) \mapsto z = g(y) = g(f(x))$$

Uma função $f : X \to Y$ é *injetiva* se $x \neq x'$ implica $f(x) \neq f(x')$, e portato a imagem $f(X)$ é uma "cópia" de X. Uma função $f : X \to Y$ é *sobrejetiva* se todo $y \in Y$ é imagem $y = f(x)$ de algum $x \in X$, ou seja, se $Y = f(X)$. Uma função $f : X \to Y$ é *bijetiva/invertível* se é injetiva e sobrejetiva, e portanto admite uma função *inversa* $f^{-1} : Y \to X$, que verifica $f^{-1}(f(x)) = x$ e $f(f^{-1}(y)) = y$ para todos os $x \in X$ e $y \in Y$.

Espaço euclidiano. $\mathbb{R}^n$ denota o espaço euclidiano de dimensão n. Fixada a base canónica $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \dots)$, ..., $\mathbf{e}_n = (0, \dots, 0, 1)$, os pontos de $\mathbb{R}^n$ são os vetores

$$\mathbf{x} = (x^1, x^2, \dots, x^n) := x^1\mathbf{e}_1 + x^2\mathbf{e}_2 + \cdots + x^n\mathbf{e}_n$$

de coordenadas $x^i \in \mathbb{R}$, com $i = 1, 2, \dots, n$.

O *produto interno Euclidiano/canónico* é definido por

$$\mathbf{x} \cdot \mathbf{y} := x^1y^1 + x^2y^2 + \cdots + x^ny^n = \sum_{ij} \delta_{ij} x^i x^j \,.$$

(onde (δ_{ij}) é a matriz/símbolo de Kronecker igual a $\delta_{ii} = 1$ na diagonal e $\delta_{ij} = 0$ se $i \neq i$). O produto interno Euclidiano realiza um isomorfismo entre o espaço dual (algébrico) $(\mathbb{R}^n)^* := \mathrm{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R})$ e o próprio $\mathbb{R}^n$: o valor da forma linear (ou co-vetor) $\boldsymbol{\xi} \in (\mathbb{R}^n)^* \approx \mathbb{R}^n$ no vetor $\mathbf{x} \in \mathbb{R}^n$ é $\boldsymbol{\xi}(\mathbf{x}) = \boldsymbol{\xi} \cdot \mathbf{x}$.

A norma Euclidiana do vetor $\mathbf{x} \in \mathbb{R}^n$ é $\|\mathbf{x}\| := \sqrt{\mathbf{x} \cdot \mathbf{x}}$. A distância Euclidiana entre os pontos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ é definida pelo teorema de Pitágoras

$$\mathrm{d}(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\| = \sqrt{(x^1 - y^1)^2 + \cdots + (x^n - yn)^2}\,.$$

A bola aberta de centro $a \in \mathbb{R}^n$ e raio $\varepsilon > 0$ é o conjunto $B_\varepsilon(\mathbf{a}) := \{\mathbf{x} \in \mathbb{R}^n \text{ s.t. } \|\mathbf{x} - \mathbf{a}\| < \varepsilon\}$. Um subconjunto $A \subset \mathbb{R}^n$ é aberto em $\mathbb{R}^n$ se cada seu ponto $\mathbf{a} \in A$ é o centro de uma bola $B_\varepsilon(\mathbf{a}) \subset A$, com $\varepsilon > 0$ suficientemente pequeno.

Os pontos e as relativas coordenadas no plano Euclidiano $\mathbb{R}^2$ ou no espaço Euclidiano 3-dimensional $\mathbb{R}^3$ (ou seja, as posições dos pontos materiais da física) são também denotados, conforme a tradição, pelas letras $\mathbf{r} = (x, y) \in \mathbb{R}^2$ ou $\mathbf{r} = (x, y, z) \in \mathbb{R}^3$. Então $r := \|\mathbf{r}\|$ denota o comprimento do vetor $\mathbf{r}$, ou seja, a distância do ponto $\mathbf{r}$ da origem do referencial.

Caminhos. Se $t \mapsto \mathbf{x}(t) = (x^1(t), x^2(t), \dots, x^n(t)) \in \mathbb{R}^n$ é uma função diferenciável do "tempo" $t \in I \subset \mathbb{R}$, ou seja, um caminho diferenciável definido num intervalo de tempos $I \subset \mathbb{R}$ com valores no espaço Euclidiano $\mathbb{R}^n$, então as suas derivadas são denotadas por

$$\dot{\mathbf{x}} := \frac{d\mathbf{x}}{dt}\,, \qquad \ddot{\mathbf{x}} := \frac{d^2\mathbf{x}}{dt^2}\,, \qquad \dddot{\mathbf{x}} := \frac{d^3\mathbf{x}}{dt^3}\,, \qquad \dots$$

Em particular, a primeira derivada $\mathbf{v}(t) := \dot{\mathbf{x}}(t)$ é dita "velocidade", e a sua norma $v(t) := \|\mathbf{v}(t)\|$ é dita "velocidade escalar" (*speed*, em inglês). A segunda derivada $\mathbf{a}(t) := \ddot{\mathbf{x}}(t)$ é dita "aceleração".

Campos. Um *campo escalar* é uma função real $u : X \subset \mathbb{R}^n \to \mathbb{R}$ definida num domínio $X \subset \mathbb{R}^n$. Um *campo vetorial* é uma função $\mathbf{F} : X \subset \mathbb{R}^n \to \mathbb{R}^k$, $\mathbf{F}(\mathbf{x}) = (F^1(\mathbf{x}), F^2(\mathbf{x}), \dots, F^k(\mathbf{x}))$, cujas coordenadas $F^i(\mathbf{x})$ são k campos escalares.

A derivada do campo diferenciável $\mathbf{F} : X \subset \mathbb{R}^n \to \mathbb{R}^k$ no ponto $\mathbf{x} \in X$ é a aplicação linear $\mathrm{d}\mathbf{F}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^k$ tal que

$$\mathbf{F}(\mathbf{x} + \mathbf{v}) = \mathbf{F}(\mathbf{x}) + \mathrm{d}\mathbf{F}(\mathbf{x})\,\mathbf{v} + o(\|\mathbf{v}\|)$$

para todos os vetores $\mathbf{v} \in \mathbb{R}^n$ de norma $\|\mathbf{v}\|$ suficientemente pequena, definida em coordenadas pela matriz Jacobiana $\mathrm{Jac}\,\mathbf{F}(\mathbf{x}) := \left(\partial F^i/\partial x^j(\mathbf{x})\right) \in \mathrm{Mat}_{k\times n}(\mathbb{R})$. Em particular, o *diferencial* do campo escalar $u : X \subset \mathbb{R}^n \to \mathbb{R}$ no ponto $\mathbf{x} \in X$ é a forma linear $\mathrm{d}u(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$,

$$\mathrm{d}u(\mathbf{x}) := \frac{\partial u}{\partial x^1}(\mathbf{x})\,dx^1 + \frac{\partial u}{\partial x^2}(\mathbf{x})\,dx^2 + \cdots + \frac{\partial u}{\partial x^n}(\mathbf{x})\,dx^n$$

(onde $\mathrm{d}x^k$, o diferencial da função coordenada $\mathbf{x} \mapsto x^k$, é a forma linear que envia o vetor $\mathbf{v} = (v^1, v^2, \dots, v^n) \in \mathbb{R}^n$ na sua k-ésima coordenada $\mathrm{d}x^k(\mathbf{v}) := v^k$). A derivada do campo escalar diferenciável $u : X \subset \mathbb{R}^n \to \mathbb{R}$ na direção do vetor $\mathbf{v} \in \mathbb{R}^n$ (aplicado) no ponto $\mathbf{x} \in X \subset \mathbb{R}^n$, é igual, pela regra da cadeia, a

$$(\pounds_{\mathbf{v}} u)(\mathbf{x}) := \left.\frac{d}{dt} u(\mathbf{x} + t\mathbf{v})\right|_{t=0} = \langle \mathrm{d}u(\mathbf{x}), \mathbf{v}\rangle\,.$$

O *gradiente* do campo escalar diferenciável $u : X \subset \mathbb{R}^n \to$ é o campo vetorial $\nabla u : X \subset \mathbb{R}^n \to \mathbb{R}^n$ tal que

$$\mathrm{d}u(\mathbf{x})\,\mathbf{v} = \nabla u(\mathbf{x}) \cdot \mathbf{v}$$

para todo os vetores (tangentes) $\mathbf{v} \in \mathbb{R}^n$ (aplicados no ponto $\mathbf{x} \in X$).

1 Equações diferenciais ordinárias

[Ap69] *Vol. 1*, **8.1-7**

1. (partícula livre) A trajetória $t \mapsto \mathbf{r}(t) = (x(t), y(t), z(t)) \in \mathbb{R}^3$ de uma partícula livre de massa $m > 0$ num referencial inercial é modelada pela equação de Newton

$$\frac{d}{dt}(m\,\mathbf{v}) = 0\,,$$

ou seja, se a massa m é constante,

$$m\,\mathbf{a} = 0\,, \tag{1.1}$$

onde $\mathbf{v}(t) := \dot{\mathbf{r}}(t)$ denota a *velocidade* e $\mathbf{a}(t) := \ddot{\mathbf{r}}(t)$ denota a *aceleração* da partícula. Em particular, o *momento linear* $\mathbf{p} := m\mathbf{v}$ é uma constante do movimento, de acordo com o princípio de inércia de Galileo[1] ou a primeira lei de Newton[2]. As soluções da equação de Newton (1.1) da partícula livre, ou seja, as trajetórias com aceleração nula, são as retas afins

$$\mathbf{r}(t) = \mathbf{s} + \mathbf{v}t\,,$$

onde $\mathbf{s} = \mathbf{r}(0) \in \mathbb{R}^3$ é a posição inicial e $\mathbf{v} = \dot{\mathbf{r}}(0) \in \mathbb{R}^3$ é a velocidade (inicial).

- Determine a trajetória de uma partícula livre que passa, no instante $t_0 = 0$, pela posição $\mathbf{r}(0) = (3, 2, 1)$ com velocidade $\dot{\mathbf{r}}(0) = (1, 2, 3)$.
- Determine a velocidade inicial da trajetória de uma partícula livre que passa pela posição $\mathbf{r}(0) = (0, 1, 2)$ no instante $t_0 = 0$ e pela posição $\mathbf{r}(2) = (3, 4, 5)$ no instante $t_1 = 2$.

2. (queda livre) A queda livre de uma partícula próxima da superfície terrestre é modelada pela equação de Newton

$$m\ddot{q} = -mg \tag{1.2}$$

onde $q(t) \in \mathbb{R}$ denota a altura da partícula no instante t, $m > 0$ é a massa da partícula, e $g \simeq 980$ cm/s^2 é a aceleração da gravidade próximo da superfície terrestre. É um fato experimental que a "massa inercial" (o factor de $\ddot{q}$ na (1.2)) e a "massa gravitacional" (o fator de g na (1.2)) são iguais. De consequência, a equação de Newton reduz-se a $\ddot{q} = -g$, ou seja, a lei horária da queda livre não depende da massa da partícula! As soluções da equação de Newton (1.2) da queda livre, ou seja, as trajetórias com aceleração constante, são as parábolas

$$q(t) = s + v_0 t - \tfrac{1}{2}gt^2\,,$$

onde $s = q(0) \in \mathbb{R}$ é a altura inicial e $v_0 = \dot{q}(0) \in \mathbb{R}$ é a velocidade inicial.

- Uma pedra é deixada cair do topo da torre de Pisa, que tem cerca de 56 metros de altura, com velocidade inicial nula. Calcule a altura da pedra após 1 segundo e determine o tempo necessário para a pedra atingir o chão.
- Com que velocidade inicial deve uma pedra ser atirada para cima de forma a atingir a altura de 20 metros, relativamente ao ponto inicial?
- Com que velocidade inicial deve uma pedra ser atirada para cima de forma a voltar de novo ao ponto de partida ao fim de 10 segundos?
- Determine soluções da equação de Newton

$$\ddot{q} = 1\,.$$

[1] "...il mobile durasse a muoversi tanto quanto durasse la lunghezza di quella superficie, né erta né china; se tale spazio fusse interminato, il moto in esso sarebbe parimenti senza termine, cioè perpetuo" [Galileo Galilei, *Dialogo sopra i due massimi sistemi del mondo*, 1623.]

[2] "Lex prima: Corpus omne perseverare in statu suo quiescendi vel movendi uniformiter in directum, nisi quatenus a viribus impressis cogitur statum illum mutare" [Isaac Newton, *Philosophiae Naturalis Principia Mathematica*, 1687.]

3. (o exponencial) O *exponencial* (real) é a função $\exp : \mathbb{R} \to \mathbb{R}$, $t \mapsto \exp(t) = e^t$, definida pela série de potências

$$\boxed{\begin{aligned} e^t &:= \sum_{n=0}^{\infty} \frac{t^n}{n!} \\ &= 1 + t + \frac{t^2}{2} + \frac{t^3}{6} + \frac{t^4}{24} + \dots \end{aligned}} \tag{1.3}$$

que converge uniformemente em cada intervalo limitado da reta real.

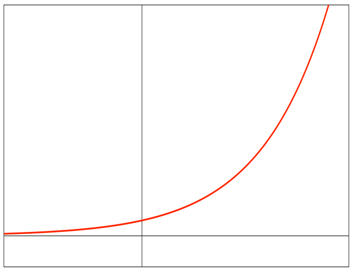


Gráfico do exponencial.

É imediato verificar que $e^0 = 1$, e que $e^{t+s} = e^t e^s$ para todos os $t, s \in \mathbb{R}$ (ou seja, exp define um homomorfismo do grupo aditivo $\mathbb{R}$ no grupo multiplicativo $\mathbb{R}_+^\times$ dos números reais positivos). Em particular, $e^t \neq 0$ para todos os $t \in \mathbb{R}$, e $e^{-t} = (e^t)^{-1}$.

A derivada do exponencial é o próprio exponencial, como se pode ver derivando a série de potências

$$\frac{d}{dt}\left(1 + t + \frac{t^2}{2} + \frac{t^3}{6} + \frac{t^4}{24} + \dots\right) = 0 + 1 + t + \frac{t^2}{2} + \frac{t^3}{6} + \dots$$

(e usando resultados sobre a derivação de séries convergentes). Em outras palavras, a função exponencial $x(t) = e^t$ satisfaz a equação diferencial

$$\dot{x} = x$$

com condição inicial $x(0) = 1$. Mais em geral,

Teorema 1.1. *A função $x(t) = x_0 e^{\lambda t}$, com x_0 constante, é a única solução da equação diferencial*

$$\dot{x} = \lambda x \tag{1.4}$$

(i.e. a única função diferenciável cuja derivada é igual a λ vezes a própria função) com condição inicial $x(0) = x_0$ (i.e. tal que o seu valor quando $t = 0$ é igual a x_0).

Demonstração. Se $y(t)$ é uma (outra?) solução de (1.4) com condição inicial $y(0) = x_0$, então o quociente $q(t) = y(t)/e^{\lambda t}$ tem derivada $\dot{q} = (\dot{y} - \lambda y)e^{-\lambda t} = 0$. Pelo teorema do valor médio, $q(t)$ é constante e, em particular, igual ao seu valor em $t = 0$, que é x_0. De consequência, $y(t) = x_0 e^{\lambda t}$. □

- Determine as soluções de

$$\dot{x} = 3x$$

tais que $x(0) = 1/e$ ou $x(2) = e$.

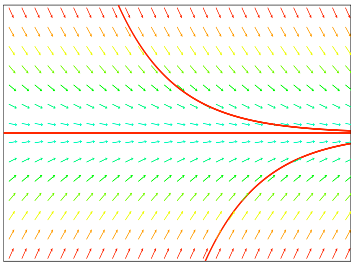
4. (decaimento radioativo) A taxa de decaimento de matéria radioativa é proporcional à quantidade de matéria existente, desde que a amostra seja suficientemente grande. Quer isto dizer que a quantidade $N(t)$ de matéria radioativa existente no instante t satisfaz a lei

$$\dot{N} = -\beta N\,, \tag{1.5}$$

onde o parâmetro $1/\beta > 0$ é a "vida média" dos núcleos[3]. A solução de (1.5) com condição inicial $N(0) = N_0 > 0$ é

$$N(t) = N_0 \, e^{-\beta t} \, ,$$

e "decai" para o equilíbrio $\overline{N} = 0$ quando $t \to \infty$.



Se a radiação solar produz radiocarbono na atmosfera terrestre a uma taxa constante $\alpha > 0$, então a quantidade de radiocarbono na atmosfera segue a lei (decaimento com reposição)

$$\dot{N} = -\beta N + \alpha \, . \tag{1.6}$$

A solução de equilíbrio de (1.6) é $\overline{N} = \alpha/\beta$. A diferença $x(t) := N(t) - \overline{N}$ satisfaz a equação diferencial $\dot{x} = -\beta x$ (ou seja, a (1.5)), e portanto a solução de (1.6) é

$$N(t) = (N(0) - \overline{N})e^{-\beta t} + \overline{N} \, .$$

Observe que $N(t) \to \overline{N}$ quando $t \to \infty$, independentemente da condição inicial $N(0)$ (por exemplo no instante da creação do Universo!).

- O tempo de *meia-vida* de uma matéria radioativa é o tempo τ necessário até a quantidade de matéria se reduzir a metade da quantidade inicial (ou seja, $N(\tau) = \frac{1}{2}N(0)$). Mostre que o tempo de meia-vida não depende da quantidade inicial $N(0)$, e determine a relação entre o tempo de meia-vida τ e o parâmetro β.
- O radiocarbono ^{14}C tem vida média $1/\beta \simeq 8033$ anos. Mostre como datar um fóssil, assumindo que a proporção de radiocarbono num ser vivente é conhecida[4].

5. (crescimento exponencial) Um modelo do crescimento de uma população num meio ambiente ilimitado é

$$\dot{N} = \lambda N \, , \tag{1.7}$$

onde $N(t)$ é a quantidade de exemplares existentes no instante t, e $\lambda > 0$ (se α é a taxa de natalidade e β é a taxa de mortalidade, então $\lambda = \alpha - \beta$). A solução estacionária é a solução trivial $N(t) = 0$ (população ausente). A solução com condição inicial $N(0) = N_0 > 0$ é

$$N(t) = N_0 \, e^{\lambda t}$$

e diverge quanto $t \to \infty$ (explosão demográfica!).

- Se a população de uma bactéria duplica numa hora, quanto aumentará em duas horas?
- Se de uma população que cresce exponencialmente é retirada uma parte a uma taxa constante $\gamma > 0$, então a população segue a lei

$$\dot{N} = \lambda N - \gamma \, .$$

Determine o estado estacionário, e discuta o comportamento assimptótico das outras soluções (veja a solução do problema do decaimento com reposição).

[3] O tempo de vida de cada núcleo é modelado por uma variável aleatória exponencial X, com lei $\text{Prob}(X \leq t) = 1 - e^{-\beta t}$ se $t \geq 0$, e 0 se $t < 0$, e média $\mathbf{E}X := \int_0^\infty t \, d\text{Prob}(X \leq t) = 1/\beta$. A equação diferencial, quando a quantidade N de núcleos é grande, é uma consequência da lei dos grandes números.

[4] J.R. Arnold and W.F. Libby, Age determinations by Radiocarbon Content: Checks with Samples of Known Ages, *Sciences* **110** (1949), 1127-1151.

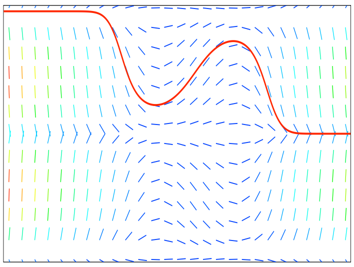
6. (equações diferenciais ordinárias) Uma *equação diferencial ordinária (EDO)* de primeira ordem (resolúvel para a derivada) é uma lei

$$\dot{\mathbf{x}} = \mathbf{v}(t, \mathbf{x}) \tag{1.8}$$

para a trajetória $t \mapsto \mathbf{x}(t)$ de um sistema com *espaço de fases* $X \subset \mathbb{R}^n$, onde $\mathbf{x}(t) \in X$ denota o *estado* do sistema no instante $t \in T \subset \mathbb{R}$, $\dot{\mathbf{x}} := \frac{d\mathbf{x}}{dt}$ denota a derivada do observável/is x em ordem ao tempo t, e $\mathbf{v} : T \times X \to \mathbb{R}^n$ é um *campo de direções* dado.

Uma *solução (local)* da EDO (1.8) é um caminho diferenciável $t \mapsto \mathbf{x}(t)$ cuja velocidade satisfaz $\dot{\mathbf{x}}(t) = \mathbf{v}(t, \mathbf{x}(t))$ para cada tempo t num intervalo $I \subset T$, ou seja, uma função $\mathbf{x} : I \to X$ cujo gráfico $\Gamma := \{(t, \mathbf{x}(t)) \in I \times X \text{ com } t \in I\}$, dito *curva integral* de (1.8), é tangente ao campo de direções $\mathbf{v}(t, \mathbf{x})$ em cada ponto $(t, \mathbf{x}(t)) \in \Gamma$. Uma solução $\mathbf{x} : I \to X$ é dita *maximal* se não pode ser extendida a um intervalo maior $J \supset I$. Uma solução definida para todos os tempos $t \in \mathbb{R}$ é dita *solução global*.

Dados um tempo $t_0 \in T$ e um ponto $\mathbf{x}_0 \in X$, uma solução da EDO (1.8) com *condição inicial* $\mathbf{x}(t_0) = \mathbf{x}_0$ (solução do "Problema de Valores Iniciais", P.V.I., ou solução do "problema de Cauchy") é uma solução definida numa vizinhança de t_0 cujo gráfico contém o ponto $(t_0, \mathbf{x}_0) \in T \times X$.



Campo de direções e uma solução de $\dot{x} = \sin(x)(1 - t^2)$.

O *teorema de Peano* [5] [6] afirma que, se o campo $\mathbf{v}(t, \mathbf{x})$ é contínuo, então existem sempre soluções locais, i.e. definidas em vizinhanças suficientemente pequenas do tempo inicial, do problema de Cauchy. Por outro lado, a continuidade do campo de direções não é suficiente para garantir a unicidade das soluções. O *teorema de Picard-Lindelöf* [7] afirma que, se o campo $\mathbf{v}(t, \mathbf{x})$ é contínuo e localmente Lipschitziano [8] (por exemplo, diferenciável e com derivada contínua) na variável $\mathbf{x}$, então para cada ponto $(t_0, \mathbf{x}_0) \in T \times X$ passa uma única solução com condição inicial $\mathbf{x}(t_0) = \mathbf{x}_0$.

A forma mais geral de uma EDO de primeira ordem é

$$F(t, \mathbf{x}, \dot{\mathbf{x}}) = 0 \, .$$

Então também existem EDOs que não têm solução por razões triviais, como por exemplo $(\dot{x})^2 + 1 = 0$.

- Esboce o campo de direções das EDOs

$$\dot{x} = t \qquad \dot{x} = -x + t \qquad \dot{x} = \sin(t)$$

e conjeture sobre o comportamento qualitativo das soluções.

[5] G. Peano, Sull'integrabilità delle equazioni differenziali del primo ordine, *Atti Accad. Sci. Torino* **21** (1886), 677-685.

[6] G. Peano, Demonstration de l'intégrabilité des équations différentielles ordinaires, *Mathematische Annalen* **37** (1890) 182-228.

[7] M. E. Lindelöf, Sur l'application de la méthode des approximations successives aux équations différentielles ordinaires du premier ordre, *Comptes rendus hebdomadaires des séances de l'Académie des sciences* **114** (1894), 454-457.

[8] A função $f : U \to \mathbb{R}^m$ é *Lipschitziana* no domínio $U \subset \mathbb{R}^n$ se

$$\exists \, L > 0 \quad \text{t.q.} \qquad \|f(x) - f(y)\| \leq L \cdot \|x - y\| \qquad \forall \, x, y \in U \, .$$

- A função $x(t) = t^3$ é solução da equação diferencial $\dot{x} = 3x^{2/3}$ com condição inicial $x(0) = 0$? E a função $x(t) = 0$?
- Determine umas EDOs de primeira ordem e de segunda ordem que admitam como solução a Gaussiana $\varphi(t) = e^{-t^2/2}$.

7. (quase todas as EDOs têm ordem um!) A EDO de ordem $n \geq 2$ (resolúvel para a n-ésima derivada)
$$y^{(n)} = F\left(t, y, \dot{y}, \ddot{y}, ..., y^{(n-1)}\right)$$
para o observável $y(t) \in \mathbb{R}$ é equivalente à EDO (ou sistema de EDOs) de primeira ordem
$$\dot{\mathbf{x}} = \mathbf{v}\left(t, \mathbf{x}\right)$$
para o observável $\mathbf{x} = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ definido por
$$x_1 := y \qquad x_2 := \dot{y} \qquad x_3 := \ddot{y} \qquad ... \qquad x_n := y^{(n-1)}\,,$$
onde o campo de direções é $\mathbf{v}(t, \mathbf{x}) := (x_2, ..., x_{n-1}, F\left(t, x_1, x_2, ..., x_n\right))$.
   - Determine os sistemas de ODEs de ordem 1 que traduzem as seguintes ODEs de ordem > 1
   $$\ddot{x} = -x \qquad \ddot{x} + \dot{x} = 0 \qquad \ddot{x} + \dot{x} + x = 0$$
   $$\ddot{x} = t - x \qquad \ddot{x} + \dot{x} + x = 0 \qquad \dddot{x} = x$$
   - Determine uma EDO de segunda ordem que admita como solução a Gaussiana $\varphi(t) = e^{-t^2/2}$.
   - Mostre que soluções da EDO $x^{(n)} = 0$ são os polinómios de grau $n - 1$.

8. (soluções de Chandrasekhar da equação de Lane-Emden) Um modelo do perfil de equilíbrio hidrostático de uma estrela é a *equação de Lane-Emden*
$$\frac{1}{\xi^2}\frac{d}{d\xi}\left(\xi^2\frac{d\theta}{d\xi}\right) = -\theta^p\,, \tag{1.9}$$
onde $\xi \geq 0$ é uma "distância adimensional" do centro da estrela, $\theta(\xi)$ é proporcional à densidade, e p é um parâmetro que depende da equação de estado $P = K\rho^{1+1/p}$ do gás que forma a estrela. O problema físico é determinare a solução com condições iniciais $\theta(0) = 1$ e $d\theta/d\xi(0) = 0$, e o menor zero de $\theta(\xi)$ com $\xi > 0$ é interpretado como sendo o raio da estrela.
   - Verifique que
   $$\theta(\xi) = 1 - \frac{1}{6}\xi^2\,, \qquad \theta(\xi) = \frac{\sin\xi}{\xi} \qquad \text{e} \qquad \theta(\xi) = \frac{1}{\sqrt{1 + \frac{1}{3}\xi^2}}$$
   são soluções da equação de Lane-Emden (1.9) quando $p = 0$, 1 e 5, respetivamente [9].

[9] S. Chandrasekhar, *An Introduction to the Study of Stellar Structure*, Dover, 1958.

2 Integração numérica e simulações

1. (problema) Exceto poucos casos importantes, resolvidos por matemáticos e físicos famosos (como a queda livre, a lei de Hooke/oscillador harmónico, o problema de Kepler, ...), não há nenhuma esperança de resolver "analiticamente" as equações diferenciais que descrevem fenómenos interessantes do mundo real (leis da física, problemas de engenharia, ...).

2. (método de Euler/diferênças finitas) Considere o problema de simular as soluções da EDO

$$\dot{x} = v(t,x) \tag{2.1}$$

O *método de Euler* consiste em utilizar recursivamente a aproximação linear

$$x(t+dt) \simeq x(t) + v(t,x) \cdot dt \,,$$

dado um "passo" dt suficientemente pequeno. Portanto, a solução $x(t_n)$, nos tempos $t_n = t_0 + n \cdot dt$, com condição inicial $x(t_0) = x_0$ é estimada pela sucessão $(x_n)_{n \in \mathbb{N}_0}$ definida recursivamente por

$$\boxed{x_{n+1} = x_n + v(t_n, x_n) \cdot dt \,.}$$

Numa linguagem como `c++` ou `Java`, o ciclo para obter uma aproximação de $x(t)$, dado $x(t_0) = $ `x`, é

```
while (time < t)
  {
    x += v(time, x) * dt ;
    time += dt ;
  }
```

- Considere a equação diferencial

$$\dot{x} = x$$

com condição inicial $x(0) = 1$. Mostre que, se o passo é $dt = \varepsilon$ e o tempo final é $t = n\varepsilon$ com $n \in \mathbb{N}$, então o método de Euler fornece a aproximação

$$x(t) \simeq x_n = (1+\varepsilon)^n$$

onde $n = t/\varepsilon$ é o número de passos. Deduza que, no limite quando o passo $\varepsilon \to 0$, as aproximações convergem para a solução $x(t) = e^t$, pois

$$\lim_{\varepsilon \to 0} (1+\varepsilon)^{t/\varepsilon} = \lim_{n \to \infty} \left(1 + \frac{t}{n}\right)^n = e^t \,.$$

- Simule a solução da EDO $\dot{x} = (1-2t)\,x$ com condição inicial $x(0) = 1$. Compare o resultado com o valor exacto $x(t) = e^{t-t^2}$, usando passos diferentes, por exemplo 0.01, 0.001, 0.0001 ...

- Aproxime, usando o método de Euler, a solução do oscilador harmónico

$$\begin{cases} \dot{q} = p \\ \dot{p} = -q \end{cases}$$

com condição inicial $q(0) = 1$ e $p(0) = 0$. Compare o valor de $q(1)$ com o valor exacto $q(1) = \cos(1)$, usando passos diferentes, por exemplo 0.1, 0.01, 0.001, 0.0001 ...

3. (a priori bounds for the finite-difference method) It is important to have some control on the difference between the true solution $x(t)$ of a autonomous ODE

$$\dot{x} = v(x)$$

and the conjecture x_n obtained with the finite-difference method with time-step $\tau > 0$, hence $t = n\tau$ (thus, n is the number of steps). A crude bound is (see [**?**], page 544)

$$|x(t) - x_n| \leq \tfrac{1}{2}\, C\, \tau \left(e^{Lt} - 1\right)$$

where

$$C := \sup_{x\in\Omega} |v(x)| \qquad \text{and} \qquad L := \sup_{x\in\Omega} \left|\frac{\partial v}{\partial x}(x)\right|$$

(C is the maximal speed, and L is the Lipschitz constant of the velocity field) and $\Omega \subset \mathbb{R}$ is a convex region where we know "a priori" the solution is confined to. Observe that for small times $t \ll 1/L$, the error is $|x(t) - x_n| \approx \frac{1}{2} CL\tau^2 n$.

4. (método RK-4) O *método de Runge-Kutta* (de ordem) *4* para simular a solução de

$$\dot{x} = v(t, x) \qquad \text{com condição inicial} \qquad x(t_0) = x_0$$

consiste em escolher um "passo" dt, e aproximar $x(t_0 + n \cdot dt)$ com a sucessão (x_n) definida recursivamente por

$$\boxed{x_{n+1} = x_n + \tfrac{dt}{6}\left(k_1 + 2k_2 + 2k_3 + k_4\right)}$$

onde $t_n = t_0 + n \cdot dt$, e os coeficientes k_1, k_2, k_3 e k_4 são definidos recursivamente por

$$\boxed{k_1 = v(t_n, x_n) \qquad k_2 = v\left(t_n + \tfrac{dt}{2}, x_n + \tfrac{dt}{2} \cdot k_1\right)}$$

$$\boxed{k_3 = v\left(t_n + \tfrac{dt}{2}, x_n + \tfrac{dt}{2} \cdot k_2\right) \qquad k_4 = v(t_n + dt, x_n + dt \cdot k_3)}$$

- Implemente um código para simular sistemas de EDOs usando o método RK-4.

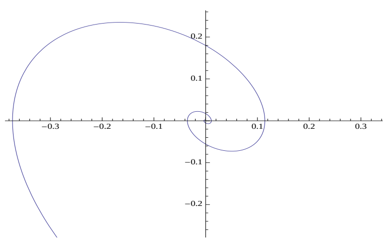
5. (simulações com software proprietário) Existem software proprietários que permitem resolver analiticamente, quando possível, ou fazer simulações numéricas de equações diferenciais ordinarias e parciais. Por exemplo, a função `ode45` do MATLAB, ou a função `NDSolve` do Mathematica, calculam soluções aproximadas de EDOs $\dot{x} = v(t, x)$ utilizando variações do método de Runge-Kutta.

   - Verifique se os PC do seu Departamento/da sua Universidade têm accesso a um dos software proprietários MATLAB ou Mathematica.
   - Em caso afirmativo, aprenda a usar as funções `ode45` ou `NDSolve`.

     Por exemplo, o pêndulo com atrito pode ser simulado, no Mathematica, usando as instruções

```
s = NDSolve[{x'[t] == y[t], y'[t] == -Sin[x[t]] - 0.7 y[t],
   x[0] == y[0] == 1}, {x, y}, {t, 20}]
ParametricPlot[Evaluate[{x[t], y[t]} /. s], {t, 0, 20}]
```

     O resultado é



3 Teoremas de existência e unicidade

1. (iterações de Picard) Uma função diferenciável $t \mapsto \varphi(t)$, definida num intervalo $I \subset \mathbb{R}$ e com valores num domínio $X \subset \mathbb{R}^n$, é solução da equação diferencial $\dot{x} = v(t,x)$ com condição inicial $\varphi(t_0) = x_0$ se e só se

$$\varphi(t) = x_0 + \int_{t_0}^{t} v\left(s, \varphi(s)\right) ds \,,$$

ou seja, se $\varphi(t)$ é um ponto fixo do *mapa de Picard* $\mathcal{P} : C(I,X) \to C(I,X)$, que envia uma função $\phi(t)$ na função

$$\boxed{\left(\mathcal{P}\phi\right)(t) := x_0 + \int_{t_0}^{t} v\left(s, \phi(s)\right) ds \,.} \tag{3.1}$$

Se a sucessão de funções ϕ, $\mathcal{P}\phi$, $\mathcal{P}^2\phi := \mathcal{P}\left(\mathcal{P}\phi\right)$, ..., $\mathcal{P}^n\phi := \mathcal{P}\left(\mathcal{P}^{n-1}\phi\right)$, ..., obtidas iterando o mapa de Picard a partir de uma função inicial ϕ, é convergente (numa topologia apropriada definida num subespaço $\mathcal{C} \subset C(I,X) := \{\phi : I \to X \text{ contínua}\}$ tal que $\mathcal{P} : \mathcal{C} \to \mathcal{C}$ seja contínua), então o limite $x(t) = \lim_{n\to\infty} \left(\mathcal{P}^n\phi\right)(t)$ é um ponto fixo do mapa de Picard, e portanto uma solução da equação diferencial $\dot{x} = v(t,x)$ com a condição inicial dada $x(t_0) = x_0$.

   - Se o campo de velocidades apenas depende do tempo, ou seja o problema é a EDO simples $\dot{x} = v(t)$, então o mapa de Picard envia toda função inicial $\phi(t)$ na solução

   $$\left(\mathcal{P}\phi\right)(t) = x_0 + \int_{t_0}^{t} v\left(s\right) ds$$

   com $x(t_0) = x_0$.
   - Queremos resolver $\dot{x} = x$ com condição inicial $x(0) = 1$. Começamos pela conjetura $\phi(t) = 1$, e depois calculamos

   $$\left(\mathcal{P}\phi\right)(t) = 1 + t \qquad \left(\mathcal{P}^2\phi\right)(t) = 1 + t + \frac{t^2}{2} \qquad \left(\mathcal{P}^3\phi\right)(t) = 1 + t + \frac{t^2}{2} + \frac{t^3}{3!} \qquad \dots$$

   Então a sucessão converge (uniformemente em intervalos limitados) para a série de Taylor da função exponencial

   $$\left(\mathcal{P}^n\phi\right)(t) = 1 + t + \frac{t^2}{2} + \dots + \frac{t^n}{n!} \quad \to \quad e^t \,,$$

   que é a solução que já conhecemos.

2. (contrações e teorema de ponto fixo de Banach) Seja (X, d) um espaço métrico. Uma transformação $f : X \to X$ é dita *contração* se é Lipschitz e tem constante de Lipschitz inferior a um, ou seja, se existe $0 \leq \lambda < 1$ tal que para todos $x, x' \in X$

$$\mathrm{d}(f(x), f(x')) \leq \lambda \cdot \mathrm{d}(x, x') \,.$$

As *trajetórias* da transformação $f : X \to X$ são as sucessões (x_n) definidas recursivamente por $x_{n+1} = f(x_n)$, se $n \geq 0$, a partir de uma condição inicial $x_0 \in X$. Os *pontos fixos* de f são os pontos $p \in X$ tas que $f(p) = p$.

Teorema 3.1 (princípio das contrações, teorema de ponto fixo de Banach)**.** *As trajetórias de uma contração $f : X \to X$ são sucesssões de Cauchy, e a distância entre cada duas trajetórias decai exponencialmente no tempo. Se X é completo, então f admite um único ponto fixo p, e a trajetória de todo ponto converge exponencialmente para o ponto fixo.*

Demonstração. Seja $f : X \to X$ uma λ-contração. Seja $x_0 \in X$ um ponto arbitrário, e seja (x_n) a sua trajetória, a sucessão definida recursivamente por $x_{n+1} = f(x_n)$. Usando k-vezes a contratividade ve-se que $\mathrm{d}\,(x_{k+1}, x_k) \leq \lambda\, \mathrm{d}(x_1, x_0)^k$, e portanto que

$$\begin{aligned} \mathrm{d}(x_{n+k}, x_n) &\leq \sum_{j=0}^{k-1} \mathrm{d}(x_{n+j+1}, x_{n+j}) \leq \mathrm{d}(x_1, x_0) \cdot \sum_{j=0}^{k-1} \lambda^{n+j} \\ &\leq \mathrm{d}(x_1, x_0) \cdot \lambda^n \cdot \sum_{j=0}^{\infty} \lambda^j \leq \frac{\lambda^n}{1-\lambda} \cdot \mathrm{d}(x_1, x_0)\,. \end{aligned}$$

Em particular, (x_n) é uma sucessão de Cauchy, pois para todo $\varepsilon > 0$ existe $\overline{n}$ tão grande que se $n \geq \overline{n}$ então $\lambda^n/(1-\lambda) < \varepsilon$.

Se (y_n) é a trajetória de um outro ponto $y_0 \in Y$, a contratividade também implica que

$$\mathrm{d}(x_n, y_n) \leq \lambda^n \cdot \mathrm{d}(x_0, y_0)\,,$$

ou seja que que a distância entre duas trajetórias decai exponencialmente.

O limite $p = \lim_{n\to\infty} x_n$, que existe se X é completo, é um ponto fixo de f, porque f é contínua e de consequência

$$f(p) = f\left(\lim_{n\to\infty} x_n\right) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = p\,.$$

Se p e p' são pontos fixos, então $d(p,p') = d(f(p), f(p')) \leq \lambda\, d(p,p')$ com $\lambda < 1$ implica que $d(p,p') = 0$, o que mostra que o ponto fixo é único. Comparando a trajetória de x_0 e do ponto fixo p (que é a sucessão constante), ve-se que

$$\mathrm{d}(x_n, p) \leq \lambda^n \cdot \mathrm{d}(x_0, p)\,,$$

ou seja que a convergência $x_n \to p$ é exponencial. $\square$

- Utilize o teorema do valor médio para mostrar que uma função $f : \mathbb{R}^n \to \mathbb{R}^n$ de classe $\mathcal{C}^1$ é uma contração sse existe $\lambda < 1$ tal que $|f'(x)| \leq \lambda$ para todo $x \in \mathbb{R}^n$.
- Mostre que uma transformação $f : X \to X$ tal que

  $$d(f(x), f(x')) < d(x, x')$$

  para todos $x, x' \in X$ distintos pode não ter pontos fixos, mesmo se o espaço métrico X for completo.

3. (teorema de Picard-Lindelöf[10]) O teorema de existência e unicidade básico para equações diferenciais ordinárias é o seguinte.

Teorema 3.2 (Picard-Lindelöf)**.** *Seja $v(t,x)$ um campo de velocidades contínuo definido num domínio D do espaço de fases extendido $\mathbb{R} \times X$. Se v é localmente Lipschitziana (por exemplo, diferenciável com continuidade) com respeito a segunda variável $x \in X \subset \mathbb{R}^n$, então existe uma e uma única solução local da equação diferencial $\dot{x} = v(t,x)$ que passa por cada ponto $(t_0, x_0) \in D$.*

[10]M. E. Lindelöf, Sur l'application de la méthode des approximations successives aux équations différentielles ordinaires du premier ordre, *Comptes rendus hebdomadaires des séances de l'Académie des sciences* **114** (1894), 454-457.

Demonstração. Seja $I \times B = [t_0 - \varepsilon, t_0 + \varepsilon] \times \overline{B}_\delta\,(x_0)$ uma vizinhança sufiientemente pequena de (t_0, x_0), onde $B = \overline{B}_\delta\,(x_0)$ denota o disco fechado de centro x_0 e raio δ in X. Pela continuidsde do campo de velocidades $v(t.x)$ existe $K > 0$ tal que $|v(t,x)| \leq K$ se $(t,x) \in I \times B$. Pela condição de Lipschitz, existe $M > 0$ tal que $|v(t,x) - v(t,y)| \leq M|x-y|$ se $t \in I$ e $x, y \in B$. Podemos reduzir, se necessário, o raio ε de I de maneira tal que $K\varepsilon \leq \delta$ e $M\varepsilon < 1$. Seja $\mathcal{C} = C^0(I, B)$ o espaço das funções contínuas $t \mapsto \phi(t)$ que enviam I em B. Munido da norma do supremo, $\|\phi - \varphi\|_\infty := \sup_{t \in I} |\phi(t) - \varphi(t)|$ este é um espaço métrico completo (de fato, um espaço de Banach). O mapa de Picard (3.2) envia $\mathcal{C}$ em $\mathcal{C}$, pois

$$|\left(\mathcal{P}\phi\right)(t) - x_0| \leq \int_{t_0}^{t} |v\left(s, \phi(s)\right)| \, ds \leq K\varepsilon \leq \delta\,.$$

Finalmente, dadas duas funções $\phi, \varphi \in \mathcal{C}$, acontece que

$$\begin{aligned} |\left(\mathcal{P}\phi\right)(t) - \left(\mathcal{P}\varphi\right)(t)| &\leq \int_{t_0}^{t} |v\left(s, \phi(s)\right) - v\left(s, \varphi(s)\right)| \, ds \\ &\leq M\varepsilon \cdot \sup_{t \in I} |\phi(t) - \varphi(t)|\,, \end{aligned}$$

e portanto

$$\|\mathcal{P}\phi - \mathcal{P}\varphi\|_\infty < M\varepsilon \cdot \|\phi - \varphi\|_\infty\,.$$

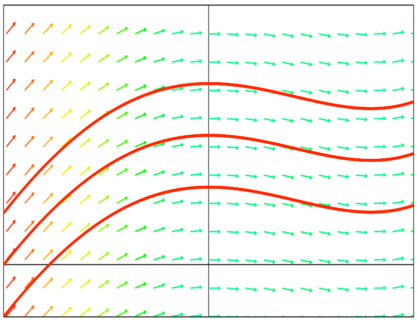
Sendo $M\varepsilon < 1$, o mapa de Picard é uma contração. O teorema segue do princípio das contrações 3.1. $\square$

4 EDOs simples e autónomas

1. (integração de EDOs simples) A equação diferencial mais simples é do género

$$\dot{x} = v(t) \tag{4.1}$$

onde o campo de direções $v(t)$ depende apenas do tempo t, e não da própria função incógnita x. Se $x(t)$ é solução de (4.1) então também $x(t)+c$ é solução, para todas as constantes $c \in \mathbb{R}$. De consequência, as soluções diferem por uma constante aditiva, determinada pela condição inicial.



Três soluções da EDO simples $\dot{x} = t^3 \sin(t)$ que diferem por uma constante aditiva.

O *teorema (fundamental do cálculo) de Newton e Leibniz*[11] afirma que a derivada do integral indefinido $F(t) := \int_a^t f(s)\,ds$ de uma função contínua $f(t)$ existe e é igual a $F'(t) = f(t)$. Portanto,

Teorema 4.1. *Se $v(t)$ é um campo de direções contínuo definido num intervalo de tempos, então a solução da EDO simples (4.1) com condição inicial $x(t_0) = x_0$ é determinada por meio de uma integração, ou seja, é dada por*

$$\boxed{x(t) = x_0 + \int_{t_0}^{t} v(s)\,ds} \tag{4.2}$$

(donde a tradição de dizer "integrar" uma equação diferencial em vez de "resolver").

Demonstração. Pelo teorema fundamental do cálculo, a derivada da (4.2) é $\dot{x}(t) = v(t)$, e o seu valor no instante t_0 é $x(t_0) = x_0$, pois o integral $\int_{t_0}^{t_0} v(s)\,ds$ é nulo. A unicidade é um exercício. □

- Integre (ou seja, determine a solução geral) as seguintes EDOs, definidas em oportunos intervalos de tempo

$$\dot{x} = 2 - t + 3t^2 + 5t^6 \qquad \dot{x} = e^{-t} \qquad \dot{x} = \cos(3t) \qquad \dot{x} = 1/t$$

- Determine $x(t)$ sabendo que

$$\dot{x} = e^{2t} \quad \text{e} \quad x(0) = 6$$

$$\dot{x} = \sin(t) \quad \text{e} \quad x(\pi) = 0$$

[11] A solução do anagrama

6accdae13eff7i3l9n4o4qrr4s8t12vx

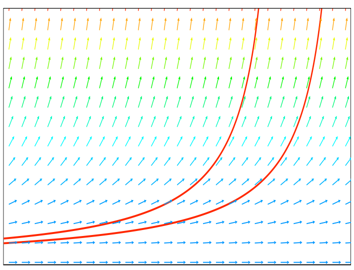
contido numa carta de Isaac Newton dirigida a Gottfried Leibniz em 1677, é "*Data aequatione quotcunque fluentes quantitates involvente fluxiones invenire et vice versa*".

- Mostre que a solução do teorema 4.2 é única (considere duas soluções, calcule a derivada da diferença, e utilize o teorema do valor médio ...)

2. (foguetão) Se um foguetão de massa $m(t)$ no espaço vazio (ou seja, sem forças gravitacionais!) expulsa combustível a uma velocidade relativa constante $-V$ e a uma taxa constante $\dot{m} = -\alpha$, (com $\alpha > 0$) então a sua trajetória num referencial inercial é modelada pela equação de Newton
$$\frac{d}{dt}(mv) = \dot{m}(v - V)\,, \qquad \text{e portanto}\ , \qquad m\dot{v} = -\alpha V\,.$$
onde $v(t) := \dot{q}(t)$ é a velocidade e $q(t)$ a posição.

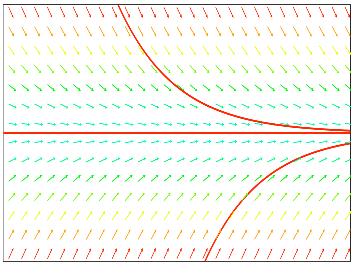
- Resolva a EDO $\dot{m} = -\alpha$ para a massa do foguetão, com massa inicial $m(0) = m_0$, e substitua o resultado na equação de Newton, obtendo
$$\dot{v} = \frac{\alpha V}{m_0 - \alpha t}$$
(desde que $0 \leq t < m_c/\alpha$, onde $m_c < m_0$ é a massa inicial do carburante). Calcule a trajetória do foguetão com velocidade inicial $v(0) = 0$ e posição inicial $q(0) = 0$, válida para tempos t inferiores ao tempo necessário para acabar o combustível.
- Se $q(t)$ representa a altura e o foguetão está sujeito à forção gravitacional próximo da superfícia da Terra, então a equação de movimento fica
$$m\dot{v} = -\alpha V - mg$$
Calcule a trajetória do foguetão com velocidade inicial $v(0) = 0$ e posição inicial $q(0) = 0$, e determone a alura atingida no instante $\bar{t} = m_c/\alpha$ em que o combustível acaba.

3. (campos de vetores e EDOs autónomas na reta) Um campo de vetores $v : X \to \mathbb{R}$, definido num intervalo $X \subset \mathbb{R}$, define uma EDO *autónoma*
$$\dot{x} = v(x)\,. \tag{4.3}$$
A palavra "autónoma" indica que o campo $v(x)$, portanto a dinâmica modelada pela equação diferencial, não depende explicitamente do tempo. Se $x(t)$ é solução de (4.3), então também $x(t - c)$ é solução, para todos os tempos $c \in \mathbb{R}$. De consequência, a física modelada por uma EDO autónoma é invariante para translações no tempo.



Duas soluções da EDO autónoma $\dot{x} = x^2$ que diferem por uma translação no tempo.

Se x_0 é um *ponto singular* de $v(x)$, i.e. um ponto onde $v(x_0) = 0$, então a trajetória constante
$$\boxed{x(t) = x_0}$$
para todos os tempos $t \in \mathbb{R}$ é uma solução *estacionária*, ou *de equilíbrio*, da equação diferencial autónoma (4.3).



Equilíbrio e outras duas soluções da EDO autónoma $\dot{x} = -x$.

Se x_0 é um ponto não singular do campo contínuo $v(x)$, i.e. se $v(x_0) \neq 0$ (e portanto, pela continuidade, $v(x)$ continua diferente de zero numa vizinhança de x_0), então uma solução local de (4.3) com condição inicial $x(t_0) = x_0$ pode ser determinanda "separando as variáveis", ou seja, fazendo formalmente

$$\frac{dx}{dt} = v(x) \qquad \Rightarrow \qquad \frac{dx}{v(x)} = dt$$

e integrando os dois membros,

$$\int \frac{dx}{v(x)} = \int dt$$

entre limites de integração apropriados. Ou seja,

Teorema 4.2. *Se x_0 é um ponto regular do campo $v(x)$, então uma solução da EDO autónoma (4.3) com condição inicial $x(t_0) = x_0$ é dada implicitamente por*

$$\boxed{\int_{x_0}^{x} \frac{dy}{v(y)} = t - t_0} \tag{4.4}$$

Se o campo $v(x)$ é diferenciável, esta solução é única.

Demonstração. Assumimos que o campo de velocidades v é contíuo, e seja $J = (x_-, x_+)$ o intervalo maximal contendo x_0 onde v é diferente de zero. Definimos a função n $H : \mathbb{R} \times J \to \mathbb{R}$ como

$$H(t, x) = t - t_0 - \int_{x_0}^{x} \frac{dy}{v(y)} .$$

Se $t \mapsto \varphi(t)$ é uma solução do problema de Cauchy, então um cálculo mostra que $\frac{d}{dt} H(t, \varphi(t)) = 0$ para todo tempo t. De consequência H é constante ao longo das soluções do problema de Cauchy. Sendo $H(t_0, x_0) = 0$, concluimos que o gráfico de toda solução pertence ao conjunto de nível $\Sigma = \{(t, x) \in \mathbb{R} \times J \text{ s.t. } H(t, x) = 0\}$. A função H é derivável, e o seu diferencial $dH = dt + dx/v(x)$ não é nulo. De fato, as duas derivadas parciais $\partial H/\partial t$ e $\partial H/\partial x$ são sempre diferentes de zero. Pelo teorema da função implícita o conjunto de nível Σ é, numa vizinhança $I \times J$ do ponto (t_0, x_0), o gráfico de uma função diferenciável $x \mapsto t(x)$, assim como o gráfico de uma função diferenciável $t \mapsto x(t)$, a função iversa de $t(x)$, que é uma solução do problema de Cauchy. De fato, a sua derivada é, pelo teorema da função inversa,

$$\dot{x}(t) = 1/\left(\frac{dt}{dx}(x(t))\right) = v(x)$$

e a condição inicial é $x(t_0) = x_0$. Observe que a função $t(x) - t_0$ tem a interpretação do "tempo necessário para ir de x_0 até x". $\square$

- Considere as seguintes EDOs autónomas

$$\dot{x} = -3x \qquad \dot{x} = x - 1 \qquad \dot{x} = x^2 \qquad \dot{x} = \sqrt{x}$$

$$\dot{x} = (x-1)(x-2) \qquad \dot{x} = e^x \qquad \dot{x} = (x-1)(x-2)(x-3)$$

definidas em intervalos convenientes. Encontre, caso existam, as soluções estacionárias. Desenhe os respectivos campos de vetores e conjeture sobre o comportamento das soluções. Integre, quando possível, as equações e calcule soluções. Determine, quando possível, umas fórmulas para a solução do problema de Cauchy comcondição inicial $x(0) = x_0$ e esboce a representação gráfica de algumas das soluções encontradas.

4. (atrito e tempo de relaxamento) O atrito pode ser modelado como sendo uma força proporcional e contrária à velocidade. Portanto, a equação de Newton (em dimensão 1) de uma partícula livre de massa m em presença de atrito é

$$m\ddot{q} = -\gamma\dot{q}$$

onde $\gamma > 0$ é o "coeficiente de atrito".

- Mostre que a velocidade $v := \dot{q}$ satisfaz

$$\dot{v} = -\frac{1}{\tau}v$$

onde $\tau = m/\gamma > 0$ é um "tempo de relaxamento". Resolva a equação dada uma velocidade inicial $v(0) = v_0 > 0$. Deduza a trajectória $q(t)$ com posição inicial $q(0) = 0$.

- Mostre que a energia cinética $T := \frac{1}{2}mv^2$ da partícula satisfaz

$$\dot{T} = -\frac{2}{\tau}T\,,$$

e portanto decresce exponencialmente com tempo de relaxamento $\tau/2$.

5. (paraquedista) Um modelo da queda de um paraquedista é

$$m\dot{v} = -\alpha v|v| - mg\,,$$

onde $v(t) := \dot{q}(t)$, $q(t) \in \mathbb{R}$ é a altura no instante t, $m > 0$ é a massa, $g \simeq 9.80$ m/s^2 é a aceleração da gravidade próximo da superfície terrestre, e $\alpha > 0$ é uma constante que depende da atmosfera e do paraqueda (um valor realístico é $\alpha \simeq 30$ kg/m).

- Mostre que a velocidade $v(t)$ converge para o valor estacionário $\overline{v} = \sqrt{mg/\alpha}$ quando $t \to \infty$.

6. (circuito/filtro RC) A tensão sobre o capacitor num filtro/circuito RC é modelada pela lei de Kirchoff

$$C\dot{V} + \frac{1}{R}V = 0$$

- Verifique que a solução com tensão inicial $V(0) = V_0$ é

$$V(t) = V_0\,e^{-t/\tau}$$

onde o tempo caraterístico do filtro é $\tau = 1/(RC)$.

7. (logística) Um modelo mais realista da dinâmica de uma população $N(t)$ num meio ambiente limitado é dado pela *equação logística*[12]

$$\dot{N} = \lambda\,N\,(1 - N/M)$$

[12] Pierre François Verhulst, Notice sur la loi que la population pursuit dans son accroissement, *Correspondance mathématique et physique* **10** (1838), 113-121.

onde a constante positiva M é a população máxima permitida e $\lambda > 0$. Observe que $\dot{N} \simeq \lambda N$ se $N \ll M$, e que $\dot{N} \to 0$ quando $N \to M$. A "população relativa" $x(t) := N(t)/M$ satisfaz a equação logística "adimensional"

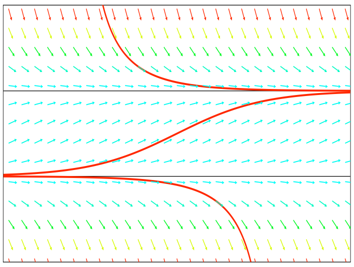
$$\dot{x} = \lambda\, x\,(1-x)\,. \tag{4.5}$$

As soluções de equilíbrio são $\overline{x} = 0$ (população ausente) e $\overline{x} = 1$ (ou seja, $\overline{N} = M$, população máxima). A solução de (4.5) com condição inicial $x(0) = x_0 \neq 0, 1$ pode ser determinada separando as variáveis e integrando, e é dada em forma implícita por

$$\int_{x_0}^{x} \frac{dy}{y(1-y)} = \int_0^t \lambda\, ds$$

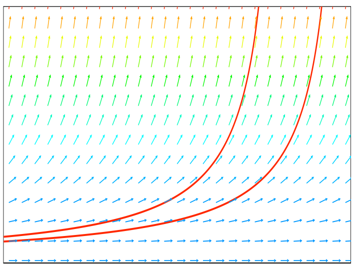
- Use a identidade
$$\frac{1}{y(1-y)} = \frac{1}{y} + \frac{1}{1-y}$$
e deduza que a solução com condição inicial $x(0) = x_0 \in (0,1)$ (fora deste intervalo o modelo não faz sentido físico) é
$$x(t) = \frac{1}{1 + \left(\frac{1}{x_0} - 1\right) e^{-\lambda t}}\,.$$

- Discuta o comportamento assimptótico das soluções da equação logística.



8. (epidemias) Num surto epidémico, a taxa de crescimento do número $I(t)$ de indivíduos infetados, dentro de uma população total constante N, é proporcional ao produto do número de indivíduos infetados e o número $S(t) = N - I(t)$ de indivíduos saudáveis (e portanto susceptíveis de serem infetados), ou seja,
$$\dot{I} = \lambda I(N - I)$$
com $\lambda > 0$.
   - Determine a lei de crescimento da população infectada relativa $x(t) := I(t)/N$, e discuta o comportamento assimptótico de $x(t)$.

9. (crescimento super-exponencial/explosão) Um outro modelo de dinâmica de uma população em meio ilimitado é
$$\dot{N} = \lambda N^2\,,$$
ou seja, a taxa de crescimento é proporcional aos pares de indivíduos contidos na população. A solução estacionária é a solução trivial $N(t) = 0$.



- Mostre que a solução com condição inicial $N(0) = N_0 > 0$ é

$$N(t) = \frac{1}{N_0^{-1} - \lambda t},$$

definida para tempos $t < 1/\lambda N_0$. Este modelo prevê uma catástrofe (população infinita, explosão) após um intervalo de tempo finito!

10. (draining a tank). Some liquid is contained in a tank which has section $S(h)$ in correspondence with height h. A hole of section s is opened at the base of the tank, and liquid start to drain. Torricelli's law says that the velocity of the dropping liquid at time t should be $v = -\sqrt{2gh}$, where $h(t)$ is the height of the liquid at time t (since the potential energy mgh gained by liquid particles falling from the liquid surface down to the hole will be transformed into a kinetic energy $mv^2/2$). Actually, due to some friction around the hole, the observed velocity is $-\gamma\sqrt{2gh}$ for some coefficient $\gamma < 1$ (which is experimentally seen to be of order 0.6 for usual liquids in usual conditions). There follows that the flow of dropping liquid is $\gamma s\sqrt{2gh}$, hence the volume $V(t)$ of liquid in the tank at time t decreases as

$$\dot{V} = -\gamma s\sqrt{2gh}\,.$$

   - Write the volume as $V(t) = \int_0^{h(t)} S(x)dx$ and show that $h(t)$ satisfies the autonomous first order ODE
   $$S(h)\,\dot{h} = -\gamma s\sqrt{2gh}\,.$$
   - Solve the equation for a cylindrical tank with constant section $S(h) = S$, and say what time does it take to drain a tank filled up to a height h_0.
   - Solve the equation for a funnel, a conical tank having section $S(h) = s + kh$ for some positive k, and answer the same question as above.

11. (fazer modelos) Escreva equações diferenciais que modelem cada uma das seguintes situações. O que pode dizer sobre as soluções?

   - A taxa de variação da temperatura de uma chávena de chá é proporcional à diferença entre a temperatura do quarto, suposta constante, e a temperatura do chá.
   - A velocidade vertical de um foguetão é inversamente proporcional à altura atingida.
   - A taxa de crescimento da massa de um cristal cúbico é proporcional à sua superfície.
   - Uma esfera de gelo derrete a uma taxa proporcional à sua superfície.
   - A taxa de crescimento de uma população de marcianos é proporcional ao número de trios que é possível formar com a dada população.

12. (campos de vetores e EDOs autónomas) Um *campo de vetores* $\mathbf{v} : X \to \mathbb{R}^n$ no espaço de fases $X \subset \mathbb{R}^n$ define uma equação diferencial ordinária *autónoma* (que não depende explicitamente do tempo, como todas as leis fundamentais da física)

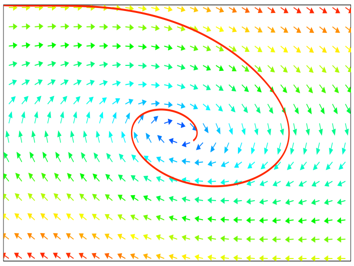
$$\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x})\,. \tag{4.6}$$

As imagens $\mathbf{x}(I) = \{\mathbf{x}(t) \text{ com } t \in I\} \subset X$ das soluções/trajetórias $\mathbf{x} : I \to X$ no espaço de fases são ditas *órbitas*, ou *curvas de fases*, do sistema autónomo.

Se $\overline{\mathbf{x}} \in X$ é um ponto singular do campo de vetores, i.e. um ponto onde $\mathbf{v}(\overline{\mathbf{x}}) = 0$, então o caminho constante $\mathbf{x}(t) = \overline{\mathbf{x}}$ para todos os tempos $t \in \mathbb{R}$ é uma solução da EDO autónoma (4.6), dita solução de *equilíbrio*, ou *estacionária*.

Soluções periódicas são soluções globais tais que $\mathbf{x}(t + T) = \mathbf{x}(t)$ para todo t e algum $T > 0$ minimal, dito *periódo*. As órbitas correspondentes são curvas fechadas.

Se a EDO (4.6) satisfaz um teorema de existência e unicidade (pelo teorema de Picard-Lindelöf 3.2, é suficiente que o campo seja Lipschitziano, por exemplo diferenciável com continuidade), então para cada ponto do espaço de fases passa uma e uma única órbita (que pode ser o próprio ponto no caso de uma solução estacionária). Em particular, órbitas diferentes não têm interseções, e portanto as órbitas definem uma "partição" do espaço de fases.



Campo de vetores e uma curva de fases do pêndulo com atrito,
$\dot{q} = p$, $\dot{p} = -\sin(q) - p/2$.

- Esboce o campo de direções e o campo de vetores das EDOs autónomas

$$\dot{x} = -x \qquad \dot{x} = x - 1 \qquad \dot{x} = x(1-x)$$

$$\dot{x} = (x-1)(x-2)(x-3) \qquad \dot{x} = (x-1)^2(x-2)^2$$

$$\begin{cases} \dot{q} = p \\ \dot{p} = -q \end{cases} \qquad \begin{cases} \dot{q} = 2q \\ \dot{p} = -p/2 \end{cases} \qquad \begin{cases} \dot{q} = q - p \\ \dot{p} = p - q \end{cases}$$

  determine as soluções de equilíbrio, e conjeture sobre o comportamento qualitativo das (outras) soluções.

13. (campos completos e fluxos de fases) Seja $\mathbf{v} : X \subset \mathbb{R}^n \to \mathbb{R}^n$ um campo de vetores definido num domínio $X \subset \mathbb{R}^n$ (ou numa variedade diferenciável). Se por cada ponto $\mathbf{x}_0 \in X$ do espaço de fases passa uma e uma única solução global (i.e. definida para todos os tempos) $\varphi : \mathbb{R} \to X$, $t \mapsto \varphi(t)$, com condição inicial $\varphi(0) = \mathbf{x}_0$, então o campo de vetores é dito *completo*. Um campo completo define/gera portanto um *fluxo de fases*, um grupo de transformações $\Phi_t = e^{t\mathbf{v}} : X \to X$, com $t \in \mathbb{R}$, tais que

$$\Phi_t \circ \Phi_s = \Phi_{t+s} \qquad \text{e} \qquad \Phi_0 = \mathrm{id}_X \qquad \forall\, t, s, \in \mathbb{R}\,.$$

O ponto $\Phi_t(\mathbf{x}_0)$ é o estado $\varphi(t)$ no tempo t da solução que passa por $\mathbf{x}_0$ no instante 0. Vice-versa, um fluxo de fases diferenciável define um campo de vetores

$$\mathbf{v}(\mathbf{x}) := \lim_{t \to 0} \frac{\Phi_t(\mathbf{x}) - \mathbf{x}}{t}\,,$$

dito "gerador infinitesimal" do grupo de transformações. As curvas $t \mapsto \Phi_t(\mathbf{x}_0)$ são as soluções de $\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x})$ com condição inicial $\mathbf{x}(0) = \mathbf{x}_0$.

- Determine os campos de vetores que geram os seguintes fluxos no plano $\mathbb{R}^2$

$$\Phi_t(x, y) = (e^{\lambda t} x\,,\, e^{\mu t} y)$$

$$\Phi_t(x, y) = (\cos(t)\, x - \sin(t)\, y\,,\, \sin(t)\, x + \cos(t)\, y)$$

$$\Phi_t(x, y) = (x + ty\,,\, y)$$

5 Sistemas conservativos

1. (constantes do movimento) Seja $\mathbf{v} : X \subset \mathbb{R}^n \to \mathbb{R}^n$, com coordenadas $\mathbf{v}(\mathbf{x}) = (v_1(\mathbf{x}), \dots, v_n(\mathbf{x}))$, um campo de vetores que define a EDO autónoma

$$\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x})$$

num espaço de fases $X \subset \mathbb{R}^n$. Os *observáveis* são as funções $\varphi : X \to \mathbb{R}$. Os observáveis que assumem valores $\varphi(\mathbf{x}(t))$ constantes ao longo das soluções $\mathbf{x}(t)$ de $\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x})$ são ditos *constantes do movimento*, ou *integrais primeiros*. Pela regra da cadeia, o observável diferenciável $\varphi : X \to \mathbb{R}$ é uma constante do movimento se e só se a *derivada de Lie* de φ ao longo do campo $\mathbf{v}$ é igual a zero, ou seja,

$$(\pounds_{\mathbf{v}}\varphi)(\mathbf{x}) := \sum_{k=1}^{n} \frac{\partial \varphi}{\partial x_k}(\mathbf{x}) \cdot v_k(\mathbf{x}) = 0\,.$$

para todos os pontos $\mathbf{x} \in X$.

As órbitas/curvas de fases estão contidas nas hiperfícies de nível $\Sigma_c := \{\mathbf{x} \in X \text{ t.q. } \varphi(\mathbf{x}) = c\}$ das constantes do movimento. Se o sistema admite k constantes do movimento independentes $\varphi_1, \varphi_2, \dots, \varphi_k$ (ou seja, tais que os diferenciais $d\varphi_i(\mathbf{x})$ são linearmente independentes em cada ponto $\mathbf{x}$), então as órbitas do sistema estão contidas nas interseções das k hiperfícies de nível, umas sub-variedades de co-dimensão k. Em particular, a existência de $n-1$ constantes do movimento independentes permite determinar as órbitas.

- Verifique que o sistema $\dot{\mathbf{x}} = A\mathbf{x}$, onde A é uma matriz diagonal com $\det A \neq 0$ (ou seja, todos os valores próprios são $\neq 0$) não admite constantes do movimento não triviais (i.e. constantes).

2. (sistemas conservativos) [LL78] A trajetória $t \mapsto \mathbf{r}(t) \in \mathbb{R}^3$ de uma partícula de massa $m > 0$ (suposta constante!) num campo de forças conservativo é modelada pela equação de Newton

$$m\ddot{\mathbf{r}} = \mathbf{F}$$

onde a força é $\mathbf{F}(\mathbf{r}) = -\nabla U(\mathbf{r})$, e $U(\mathbf{r})$ é um(a energia) *potencial*. Um sistema isolado de N pontos materiais, com posições $\mathbf{r}_\alpha(t) \in \mathbb{R}^3$ e massas $m_\alpha > 0$, com $\alpha = 1, 2, \dots, N$, é modelado pelas equações de Newton

$$m_\alpha \ddot{\mathbf{r}}_\alpha = \mathbf{F}_\alpha \qquad \alpha = 1, 2, \dots, N$$

onde a força que atua sobre o α-ésimo ponto material é

$$\mathbf{F}_\alpha = -\nabla_{\mathbf{r}_\alpha} U := -\left(\frac{\partial U}{\partial x_\alpha}, \frac{\partial U}{\partial y_\alpha}, \frac{\partial U}{\partial z_\alpha}\right),$$

i.e. o oposto do gradiente, com respeito às coordenadas $\mathbf{r}_\alpha = (x_\alpha, y_\alpha, z_\alpha)$, de uma energia potencial $U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$. A *energia cinética* é

$$K := \sum_{\alpha=1}^{N} \frac{1}{2} m_\alpha \|\mathbf{v}_\alpha\|^2$$

onde $\mathbf{v}_\alpha = \dot{\mathbf{r}}_\alpha$ é a velocidade do α-ésimo ponto material.

- Verifique que a *energia* (energia cinética + energia potencial)

$$E := K + U = \sum_{\alpha=1}^{N} \frac{1}{2} m_\alpha \|\mathbf{v}_\alpha\|^2 + U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$$

é uma constante do movimento, ou seja, que $\frac{d}{dt}E = 0$ ao longo das trajetórias.

3. (mecânica lagrangiana e hamiltoniana) Um sistema mecânico é descrito por um espaço das configurações M, tipicamente um aberto de $\mathbb{R}^n$ ou, em geral, uma variedade diferenciável, com coordenadas locais $\mathbf{q} = (q_1, q_2, \dots, q_n) \in \mathbb{R}^n$, e uma *lagrangiana* $L : TM \to \mathbb{R}$, uma função $L(\mathbf{q}, \dot{\mathbf{q}})$ que depende das coordenadas e das velocidades generalizadas $\dot{\mathbf{q}} = (\dot{q_1}, \dot{q_2}, \dots, \dot{q_n})$. Por exemplo, o espaço das configurações de um sistema de N pontos materiais é o espaço dos vetores $\mathbf{q} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) \in \mathbb{R}^{3N}$, onde $\mathbf{r}_\alpha \in \mathbb{R}^3$, com $\alpha = 1, 2, \dots, N$, representa a posição do α-ésimo ponto. A lagrangiana é

$$L(\mathbf{q}, \dot{\mathbf{q}}) = \sum_\alpha \frac{m_\alpha}{2} \|\dot{\mathbf{r}}_\alpha\|^2 - U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_\alpha) \, .$$

A *ação* de uma trajetória $[t_0, t_1] \mapsto \mathbf{q}(t) \in M$ entre a posição $\mathbf{q}(t_0)$ e a posição $\mathbf{q}(t_1)$ é o integral

$$S[t \mapsto \mathbf{q}(t)] := \int_{t_0}^{t_1} L(\mathbf{q}(t), \dot{\mathbf{q}}(t)) \, dt \, .$$

O *princípio de mínima ação (de Hamilton)* afirma que as trajetórias físicas são os pontos críticos da ação. A variação δS, dada uma variações infinitésimas $\mathbf{q}(t) + \delta\mathbf{q}(t)$ da trajetória com $\delta\mathbf{q}(t_0) = \delta\mathbf{q}(t_1) = 0$, é dada por

$$\begin{aligned} \delta S &= \int_{t_0}^{t_1} \sum_i \left(\frac{\partial L}{\partial q_i} \delta q_i(t) + \frac{\partial L}{\partial \dot{q}_i} \delta \dot{q}_i(t) \right) dt \\ &= \int_{t_0}^{t_1} \sum_i \left(\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} \right) \delta q_i(t) \, dt \end{aligned}$$

(integrando por partes a segunda soma e usando as condições de fronteira). Portanto, os pontos críticos da ação são as soluções das *equações de Euler-Lagrange*

$$\boxed{\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = \frac{\partial L}{\partial q_i} \qquad i = 1, 2, \dots, n \, .}$$

A forma linear $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2 \dots, \mathbf{p}_N) = p^1 dq_1 + p^2 dq_2 + \dots + p^n dq_n \in T^*_{\mathbf{q}} M$, de coordenadas $p^i := \partial L / \partial \dot{q}_i(\mathbf{q})$, é dito *momento*. O espaço $X = T^*M \simeq \mathbb{R}^n \times (\mathbb{R}^n)^*$, com coordenadas $(\mathbf{q}, \mathbf{p})$, é dito *espaço de fases* do sistema mecânico. As equações de Euler-Lagrange são equivalentes às *equações de Hamilton*

$$\boxed{\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} \qquad i = 1, 2, \dots, n \, ,}$$

onde a *hamiltoniana* do sistema, $H : X \to \mathbb{R}$, é a "transformada de Legendre" da lagrangiana, definida por

$$\begin{aligned} H(\mathbf{q}, \mathbf{p}) &:= \sup_v \left(\mathbf{p} \cdot \dot{\mathbf{q}} - L(\mathbf{q}, \dot{\mathbf{q}}) \right) \\ &= \sum_\alpha \frac{1}{2m} \|\mathbf{p}_\alpha\|^2 + U(\mathbf{q}) \, . \end{aligned}$$

- O espaço das configurações de um sistema de N pontos materiais é o espaço dos vetores $\mathbf{q} = (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) \in \mathbb{R}^{3N}$, onde $\mathbf{r}_\alpha \in \mathbb{R}^3$, com $\alpha = 1, 2, \dots, N$, representa a posição do α-ésimo ponto. A lagrangiana é

  $$L(\mathbf{q}, \dot{\mathbf{q}}) = \sum_\alpha \frac{m_\alpha}{2} \|\dot{\mathbf{r}}_\alpha\|^2 - U(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_\alpha) \, ,$$

  onde U é a energia potencial da interação. Verifique que as equações de Euler-Lagrange são equivalentes às equações de Newton $m_\alpha \ddot{\mathbf{r}_\alpha} = \mathbf{F}_\alpha$.

- Mostre que a hamiltoniana é uma constante do movimento, ou seja, que

$$\frac{d}{dt}H(\mathbf{q}(t),\mathbf{p}(t)) = 0$$

ao longo das soluções das equações de Hamilton. Deduza que as órbitas do sistema no espaço de fases X estão contidas nas curvas/superfícies de nível $\{H(\mathbf{q},\mathbf{p}) = c\}$ da hamiltoniana.

4. (one-dimensional Newtonian motion in a time independent force field) The one-dimensional motion of a particle of mass m subject to a force $F(x)$ that does not depend on time is described by the Newton equation

$$m\ddot{x} = -\frac{dU}{dx}(x)\,,$$

where the potential $U(x) = -\int F(x)dx$ is some primitive of the force. The total energy

$$E\,(x,\dot{x}) = \frac{1}{2}m\dot{x}^2 + U(x)$$

(which of course is defined up to an arbitrary additive constant) of the system is a constant of the motion, i.e. is constant along solutions of the Newton equation. In particular, once a value E of the energy is given (depending on the initial conditions), the motion takes place in the region where $U(x) \leq E$, since the kinetic energy $\frac{1}{2}m\dot{x}^2$ is non-negative. Conservation of energy allows to reduce the problem to the first order ODE

$$\dot{x}^2 = \frac{2}{m}\left(E - U(x)\right)\,,$$

which has the unpleasant feature to be quadratic in the velocity $\dot{x}$. Meanwhile, if we are interested in a one-way trajectory going from some x_0 to x, say with $x > x_0$, we may solve for $\dot{x}$ and find the first order autonomous ODE

$$\dot{x} = \sqrt{\frac{2}{m}\left(E - U(x)\right)}\,.$$

There follows that the time needed to go from x_0 to x is

$$t(x) = \int_{x_0}^{x} \frac{dy}{\sqrt{\frac{2}{m}\left(E - U(y)\right)}}\,.$$

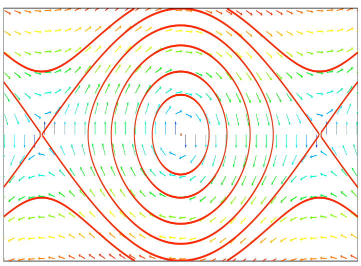
The inverse function of the above $t(x)$ will give the trajectory $x(t)$ with initial position $x(0) = x_0$ and initial positive velocity $\dot{x}(0) = \sqrt{\frac{2}{m}\left(E - U(x_0)\right)}$, at least for sufficiently small times t.

5. (pêndulo matemático) A equação de Newton que modela as oscilações de um *pêndulo* é

$$\ddot{\theta} = -\omega^2 \sin(\theta)\,,$$

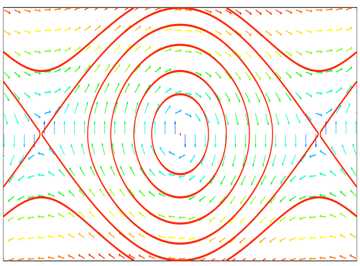
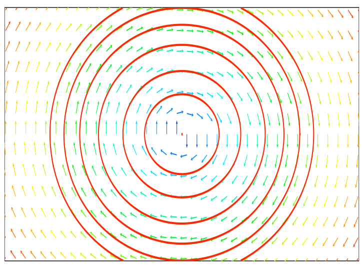
onde $\omega = \sqrt{g/\ell}$, $g \simeq 980$ cm s^{-2} é a aceleração gravitacional, ℓ o comprimento do pêndulo e θ é o angulo que o pêndulo forma com a vertical. No espaço de fases, de coordenadas θ e $p := \dot{\theta}$, a equação assume a forma do sistema

$$\left\{ \begin{array}{l} \dot{\theta} = p \\ \dot{p} = -\omega^2 \sin(\theta) \end{array} \right. .$$



- Verifique que a energia
$$H(\theta,p) := \frac{1}{2}p^2 + \omega^2(1-\cos(\theta))$$
é uma constante do movimento.
- Esboçe as curvas de energia constante e o campo de velocidades, e conjeture sobre as trajetórias.
- Show that the motion with energy E is given by
$$t = \int \frac{d\theta}{\sqrt{2(E-\cos(\theta))}}$$
- Define the new variable $x := \sqrt{\frac{2}{E+1}}\sin(\theta/2)$ and the square energy $K := \sqrt{\frac{E+1}{2}}$, and show that the motion reads
$$\dot{x} = \sqrt{(1-x^2)(1-K^2x^2)}$$
Deduce that time is given by the so called *Jacobi's elliptic integral of the first kind*
$$t = \int \frac{dx}{\sqrt{(1-x^2)(1-K^2x^2)}}$$
whose solution (i.e. x as a function of time t) is "defined" as the *elliptic function* $x(t) = \mathrm{sn}(t,K)$ (see [?]).

6. (oscilador harmónico/lei de Hooke) As pequenas oscilações de um péndulo à volta da posição de equilíbrio $\theta = 0$, ou as oscilações de uma partícula sujeita à *lei de Hooke*, são modeladas pela equação do *oscilador harmónico*
$$\ddot{q} = -\omega^2 q\,.$$
No espaço de fases, de coordenadas q e $p := \dot{q}$, a equação assume a forma do sistema
$$\left\{\begin{array}{l} \dot{q} = p \\ \dot{p} = -\omega^2 q \end{array}\right.$$
- Verifique que a energia
$$H(q,p) := \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2$$
é uma constante do movimento.
- Esboce as curvas de energia constante e o campo de velocidades, e conjeture sobre as trajetórias.
- Fixed a positive energy E, the motion takes place in the interval (x_-, x_+) with $x_\pm = \pm\sqrt{2E}/\omega$, and the velocity $\dot{x}$ satisfies the quadratic equation
$$\dot{x}^2 = \omega\sqrt{(|x_\pm|^2 - x^2)}\,.$$
Find the trajectory from x_- to any $x \le x_+$.
- Compute the time needed to go from x_- to x_+, and show that it does not depend on the energy E.

Retratos de fases do pêndulo matemático e do oscilador harmónico.

7. (real gravity and second cosmic velocity) The distance r of a particle of mass m from the center of the Earth satisfies the Newton equation

$$m\ddot{r} = -mg\frac{R^2}{r^2}\,,$$

where R is the radius of the Earth (and, of course, $r \geq R$). Here we are considering the real gravitational force produced by the Earth, but we are disregarding the gravitational influence of the Sun and other celestial bodies.

- Find the potential $U(r)$ of the gravitational field and write the expression for the total energy of the system.
- Write the integral that represents the time needed to send a particle from the Earth surface $r_0 = R$ up to a height $h = r - R > 0$ from the Earth surface, given an initial energy $E > gR$.
- Find the minimum upward velocity necessary to escape from the Earth gravitational field, i.e. to reach an infinite distance.

6 EDOs lineares de primeira ordem

[Ap69] *Vol. 1*, **8.1-7**

1. (EDOs lineares de primeira ordem) Uma EDO *linear de primeira ordem* é uma lei

$$\dot{x} + p(t)\, x = q(t) \tag{6.1}$$

para o observável $x(t)$, onde os "coeficientes" $p(t)$ e $q(t)$ são funções contínuas definidas num intervalo de tempos (por exemplo, em toda a reta real).

Se $x_1(t)$ e $x_2(t)$ são duas soluções da EDO linear de primeira ordem (6.1), então a diferença $y(t) = x_1(t) - x_2(t)$ é uma solução da *equação homogénea associada*

$$\dot{y} + p(t)\, y = 0 \tag{6.2}$$

O espaço das soluções da equação homogénea (6.2) é um espaço vetorial de dimensão 1, uma reta $\mathcal{H} \approx \mathbb{R}$ gerada, por exemplo, pela solução

$$\boxed{y_1(t) = e^{-\int_{t_0}^t p(s)\, ds}} \tag{6.3}$$

que vale $y_1(t_0) = 1$ no instante inicial t_0. Portanto, o espaço das soluções da equação linear (6.1) é uma reta afim $z + \mathcal{H}$, onde $z(t)$ é uma solução particular de (6.1).

A solução da EDO linear (6.1) com condição inicial $x(t_0) = x_0$ pode ser determinada usando o método da "variação das constantes/dos parâmetros". O primeiro passo consiste em determinar uma solução não-trivial $y(t)$ da equação homogénea (6.2) (por exemplo, a solução (6.3), que tem valor 1 no instante inicial). O segundo passo consiste em substituir a "conjetura"

$$x(t) = \lambda(t)\, y(t)$$

(o fator λ é o parâmetro que varia!) na equação não-homogénea (6.1), deduzir a EDO simples

$$\dot{\lambda} y + \lambda \dot{y} + p\lambda y = q \qquad \Rightarrow \qquad \dot{\lambda} y = q$$

(porque $\lambda\dot{y} + p\lambda y = 0$, sendo y uma solução da homogénea) para o parâmetro $\lambda(t)$, e integrar

$$\lambda(t) = \lambda(t_0) + \int_{t_0}^t \frac{q(s)}{y(s)}\, ds\,,$$

usando a condição inicial $x_0 = \lambda(t_0)\, y(t_0)$. Observe que se $y(t_0) = 1$, como sugerido, então $x_0 = \lambda(t_0)$. O resultado é a seguinte receita (mas é mais fácil lembrar o método!).

Teorema 6.1. *A solução da* (6.1) *com condição inicial* $x(t_0) = x_0$ *é*

$$\boxed{x(t) = e^{-\int_{t_0}^t p(u)\, du} \left(x_0 + \int_{t_0}^t e^{\int_{t_0}^s p(u)\, du} q(s) ds \right)\,.}$$

- Determine a solução geral das EDOs lineares de primeira ordem

$$2\dot{x} - 6x = e^{2t} \qquad \dot{x} + 2x = t \qquad \dot{x} + x/t^2 = 1/t^2 \qquad \dot{x} + tx = t^2$$

definidas em oportunos intervalos da recta real.

- Resolva os seguintes problemas de Cauchy nos intervalos indicados:

$$2\dot{x} - 3x = e^{2t} \qquad t \in (-\infty, \infty) \quad \text{com } x(0) = 1$$
$$\dot{x} + x = e^{3t} \qquad t \in (-\infty, \infty) \quad \text{com } x(1) = 2$$
$$t\dot{x} - x = t^3 \qquad t \in (0, \infty) \quad \text{com } x(1) = 3$$
$$\dot{x} + tx = t \qquad t \in (-\infty, \infty) \quad \text{com } x(0) = 0$$
$$dr/d\theta + r \tan\theta = \cos\theta \qquad t \in (-\pi/2, \pi/2) \quad \text{com } r(0) = 1$$

2. (queda livre com atrito) Um modelo mais realista da queda livre de uma partícula próxima da superfície terrestre deve ter em conta a resistência do ar. A resistência pode ser modelada como sendo uma força proporcional e contrária à velocidade, assim que a equação de Newton escreve-se
$$m\ddot{q} = -\gamma\dot{q} - mg$$
onde $\gamma > 0$ é um coeficiente de atrito. Portanto, a velocidade $v := \dot{q}$ satisfaz a EDO linear de primeira ordem
$$m\dot{v} = -\gamma v - mg\,.$$
   - Resolva o problema com condição inicial $v(0) = 0$.
   - Mostre que a velocidade $v(t)$ converge para um valor assimptótico $\overline{v}$ quando $t \to \infty$, independentemente do seu valor inicial, e determine este valor.
   - Utilize a solução encontrada para determinar a trajectória $q(t)$ com condição inicial $q(0) = q_0 > 0$.

3. (circuito RL) A corrente $I(t)$ num circuito RL, de resistência R e indutância L, é determinada pela EDO
$$L\dot{I} + RI = V(t)$$
onde $V(t)$ é a tensão que alimenta o circuito.

   As soluções da equação homogénea, ou seja, com $V(t) = 0$ (circuito desligado), são proporcionais a $e^{-(R/L)t}$, e portanto decrescem exponencialmente com tempo de relaxamento $\tau = L/R$. Se o circuito é alimentado com tensão constante $V(t) = E$, então a solução estacionária é a $\overline{I} = E/R$ (lei de Ohm). A diferença $x(t) = I(t) - \overline{I}$ é solução de $\dot{x} = -(R/L)x$, e portanto a solução com corrente inicial $I(0) = I_0$ é
$$I(t) = \overline{I} + e^{-\frac{R}{L}t}\left(I_0 - \overline{I}\right)\,,$$
assimptótica à lei de Ohm.

   Quando a tensão que alimenta o circuito é variável, então a solução com corrente inicial $I(0) = I_0$ é
$$I(t) = e^{-\frac{R}{L}t}\left(I_0 + \tfrac{1}{L}\int_0^t e^{\frac{R}{L}\tau}\,V(\tau)\,d\tau\right)\,.$$
   - Resolva a equação para um circuito alimentado com uma tensão alternada $V(t) = E\sin(\omega t)$. Verifique que a solução com $I(0) = 0$ é
$$I(t) = \frac{E}{\sqrt{R^2+\omega^2L^2}}\sin(\omega t - \phi) + \frac{E\omega L}{R^2+\omega^2L^2}e^{-\frac{R}{L}t}$$
onde ϕ é uma fase que depende de ω, L e R.

4. (lei do arrefecimento de Newton) Numa primeira aproximação, a temperatura $T(t)$ no instante t de um corpo num meio ambiente cuja temperatura no instante t é $M(t)$ pode ser modelada pela *lei do arrefecimento de Newton*
$$\dot{T} = -k\,(T - M(t))$$
onde $k > 0$ é uma constante positiva (que depende do material do corpo). A solução com condição inicial $T(0) = T_0$ é
$$T(t) = e^{-kt}\left(T_0 + k\int_0^t e^{k\tau}M(\tau)\,d\tau\right)\,.$$
   - Se a temperatura do meio ambiente é mantida constante $M(t) = M$, então a diferença $x(t) := T(t) - M$ satisfaz a EDO
$$\dot{x} = -kx\,.$$
     Determine $T(t)$ e diga o que acontece quando $t \to \infty$

- Determine a solução assimptótica (ou seja, quando t é grande) quando a temperatura do meio ambiente é a função periódica $M(t) = M \sin(\omega t)$.
- Uma chávena de café, com temperatura inicial de 100^oC, é colocada numa sala cuja temperatura é de 20^oC. Sabendo que o café atinge uma temperatura de 60^oC em 10 minutos, determine a constante k do café e o tempo necessário para o café atingir a temperatura de 40^oC.

5. (equações de Bernoulli) Uma EDO da forma

$$\dot{x} + p(t)\, x = q(t)\, x^n,$$

onde p e q são funções contínuas num intervalo $I \subset \mathbb{R}$ e $n \neq 0, 1$ (caso contrário trata-se de uma normal equação linear da primeira ordem), é dita *equação de Bernoulli*.

- Verifique que $x(t) = 0$ é uma solução de equilíbrio da equação de Bernoulli.
- Seja $k = 1 - n$. Mostre que $x(t)$ é uma solução positiva da equação de Bernoulli com condição inicial $x(t_0) = x_0 > 0$ se e só se a função $y(t) = x(t)^k$ é uma solução da EDO linear

$$\dot{y} + k\, p(t)\, y = k\, q(t)$$

com condição inicial $y(t_0) = (x_0)^{1/k}$.
- Resolva os seguintes problemas de Cauchy para equações de Bernoulli:

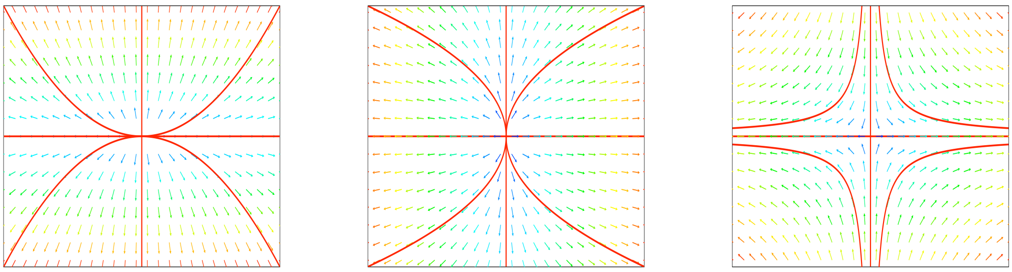
$$\dot{x} + x = x^2 (\cos t - \sin t) \qquad t \in (-\infty, \infty) \quad \text{com } x(1) = 2$$

$$t\dot{x} + e^{t^2} x = x^2 \log t \qquad t \in (0, \infty) \quad \text{com } x(3) = 0$$

$$\dot{x} - x/t = t\sqrt{x} \qquad t \in (0, \infty) \quad \text{com } x(1) = 1$$

7 EDOs separáveis e homogéneas

1. (produto direto de EDOs) O *produto directo* das EDOs autónomas $\dot{x} = v(x)$ e $\dot{y} = w(y)$ é o sistema autónomo
$$\begin{cases} \dot{x} = v(x) \\ \dot{y} = w(y) \end{cases}$$
As soluções do sistema são os caminhos $t \mapsto (x(t), y(t))$, onde $x(t)$ e $y(t)$ são as soluções das EDOs autónomas $\dot{x} = v(x)$ e $\dot{y} = w(y)$, respetivamente. A curva de fases que passa pelo ponto $(x_0, y_0) \in \mathbb{R}^2$, onde $v(x_0) \neq 0$ (ou onde $w(y_0) \neq 0$), é (localmente) o gráfico de uma função $x \mapsto y(x)$ (ou $y \mapsto x(y)$) que satisfaz a EDO
$$\frac{dy}{dx} = \frac{w(y)}{v(x)} \qquad \left(\text{ou} \quad \frac{dx}{dy} = \frac{v(y)}{w(x)} \right) .$$



   - Determine as soluções e as curvas de fases do sistema
$$\begin{cases} \dot{x} = x \\ \dot{y} = \lambda y \end{cases}$$
quando $\lambda = 0, \pm 1, 2, \ldots$ e quando $\lambda = 1/2, 1/3, \ldots$.

2. (EDOs separáveis) A solução da uma EDO *separável*
$$\frac{dy}{dx} = \frac{w(y)}{v(x)} \tag{7.1}$$
com condição inicial $y(x_0) = y_0$ tal que $v(x_0) \neq 0$ e $w(y_0) \neq 0$, é dada em forma implícita por
$$\boxed{\int_{x_0}^{x} \frac{d\xi}{v(\xi)} = \int_{y_0}^{y} \frac{d\eta}{w(\eta)}}$$

   - Resolva as seguintes EDOs separáveis definidas em oportunos domínios.
$$\frac{dy}{dx} = -x/y \qquad \frac{dy}{dx} = x/y \qquad \frac{dy}{dx} = kx^\alpha y^\beta \qquad \frac{dy}{dx} = \frac{\sin x}{\sin y}$$
$$\dot{x} = tx^3 \qquad t\dot{x} + t = t^2 \qquad \dot{x} = t^3/x^2 \qquad x\dot{x} = e^{x+3t^2}t \qquad \dot{x} = e^{t-x}$$
$$\dot{x} = \frac{t-1}{x^2} \qquad \frac{x-1}{t}\dot{x} + \frac{x-x^2}{t^2} = 0 \qquad \left(t^2+1\right)\dot{x} = 2tx \qquad \dot{x} = t\left(x^2 - x\right)$$

3. (allometric laws) If two organs/tissues/components of a living body/organism/community grow with different (but both constant!) relative growth rates α and β, say
$$\dot{x} = \alpha x \qquad \text{and} \qquad \dot{y} = \beta y$$
(the independent variable t may be time, or a linear dimension, or something else), then they satisfy the relation
$$\frac{1}{\beta y}\frac{dy}{dt} = \frac{1}{\alpha x}\frac{dx}{dt}$$

Eliminating "dt", we get the linear/separable/homogeneous ODE

$$\frac{dy}{dx} = (\beta/\alpha)\,\frac{y}{x}\,,$$

Its solution is the *allometric law* [13] [14]

$$y = c \cdot x^{\gamma} \qquad \text{or, equivalently,} \qquad \log y = \gamma \cdot \log x + \log c\,,$$

with "scaling exponent" $\gamma = \beta/\alpha$, and some constant $c = x_0/y_0$ related to the initial conditions $x(t_0) = x_0$ and $y(t_0) = y_0$.

- A famous example is *Kleiber's law* [15] (*mouse-to-elephant curve*)

$$\mathrm{BMR} = c \cdot M^{3/4}$$

which relates the basal metabolic rate BMR to the mass M of an animal.

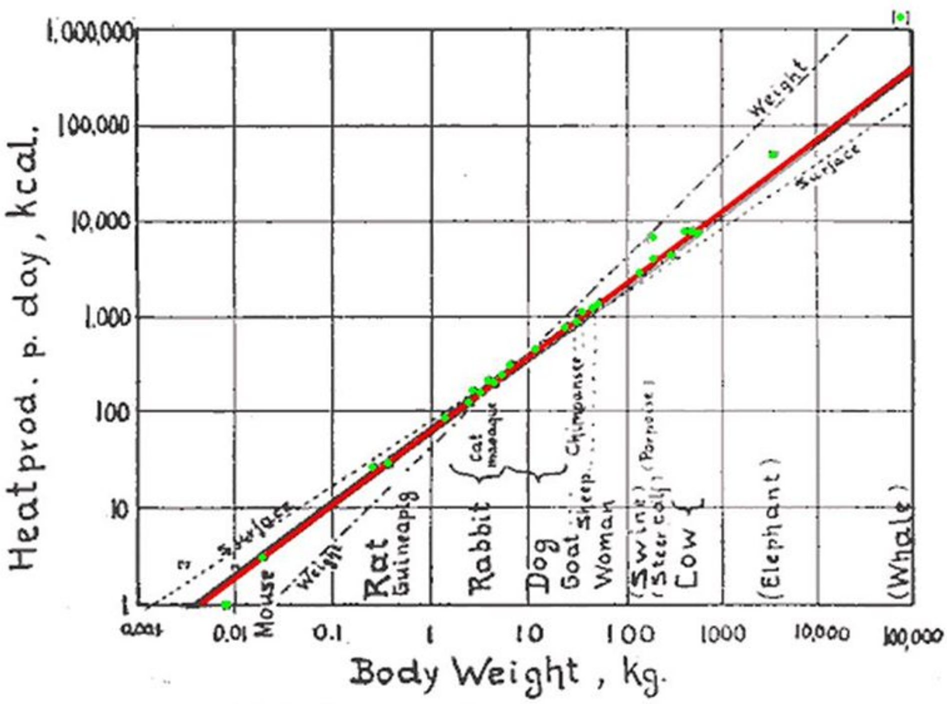


Fig. 1. Log. metabol. rate/log body weight

Original graph of body size versus metabolic rate hand-drawn by Max Kleiber (source Wikipedia)

.

- The heart rate T and the mass M of an animal are related by the allometric law

$$T = c \cdot M^{1/4}$$

4. (homotetias e funções homogéneas) As *homotetias (positivas)* do espaço Euclidiano $\mathbb{R}^n$ são as transformações $\mathbf{x} \mapsto \lambda\mathbf{x}$, com $\lambda \in \mathbb{R}_+ := (0, \infty)$. Uma função $f : D \to \mathbb{R}$ (ou um campo vetorial $\mathbf{F} : D \to \mathbb{R}^m$), definida num domínio "homogéneo" (i.e. invariante para homotetias) $D \subset \mathbb{R}^n \backslash \{0\}$, é dita *homogénea de grau k* se

$$f(\lambda\mathbf{x}) = \lambda^k f(\mathbf{x}) \qquad \forall\, \lambda \in \mathbb{R}_+ \quad \text{e} \quad \forall\, \mathbf{x} \in D\,,$$

[13] W. D'Arcy Thompson, *On Growth and Form*, 1917, 2nd ed. 1942 [Cambridge University Press, 1992].

[14] Julian S. Huxley, *Problems of Relative Growth (2nd ed.)*, Dover, 1972.

[15] M. Kleiber, Body size and metabolism, *Hilgardia* **6** (1932), 315-351. M. Kleiber, Body size and metabolic rate, *Physiological Reviews* **27** (1947), 511-541.

e é dita *homogénea* (de grau 0) se é invariante para homotetias, ou seja, se

$$f(\lambda \mathbf{x}) = f(\mathbf{x}) \qquad \forall \lambda \in \mathbb{R}_+ \quad \text{e} \quad \forall \mathbf{x} \in D\,.$$

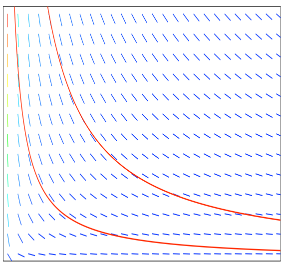
O *teorema de Euler* afirma que uma função diferenciável $f : D \subset \mathbb{R}^n \backslash \{0\} \to \mathbb{R}$ é homogénea de grau k se e só se $\langle \mathbf{x}, \nabla f(\mathbf{x}) \rangle = k\, f(\mathbf{x})$ para todos os pontos $\mathbf{x} \in D$.

- Prove o teorema de Euler (calcule as derivadas em ordem a λ dos dois termos $f(\lambda \mathbf{x})$ e $\lambda^k f(\mathbf{x})$ quando $\lambda = 1$).
- Determine os polinómios homogéneos de grau 1, de grau 2 e de grau 3 no plano $\mathbb{R}^2$.
- Mostre que as únicas funções homogéneas e contínuas definida em todo o espaço $\mathbb{R}^n$ são as constantes (observe que as funções homogéneas são constantes ao longo das semi-retas que saem da origem, logo, se a origem está no domínio da função ...).
- Determine o grau de homogeneidade dos campos de forças elástico e gravitacional/elétrico, definidos por
$$\mathbf{F}(\mathbf{r}) = -\mathbf{r} \qquad \text{e} \qquad \mathbf{F}(\mathbf{r}) = -\frac{\mathbf{r}}{\|\mathbf{r}\|^3}\,,$$
respetivamente (o segundo definido para $\mathbf{r} \in \mathbb{R}^3 \backslash \{0\}$).
- Diga se as seguintes funções $f(x,y)$, definidas em oportunos domínios do plano, são homogéneas:
$$x/y \qquad e^{x-y} \qquad \frac{x^2 - xy}{xy + 3y^2} \qquad \sin(y)\cos(x)$$

5. (EDOs homogéneas) Uma EDO *homogénea* é uma equação diferencial

$$\dot{x} = v(t,x)$$

definida, num domínio $D \subset \mathbb{R}^2$ do plano de coordenadas (t,x), por um campo de direções homogéneo, ou seja, tal que $v(\lambda t, \lambda x) = v(t,x)$ para todos os $\lambda > 0$. As homotetias $(t,x) \mapsto (\lambda t, \lambda x)$, com $\lambda \in \mathbb{R}_+$, enviam curvas integrais de uma EDO homogénea em curvas integrais.



A mudança de variável $y(t) := x(t)/t$, num domínio onde $t > 0$ ou $t < 0$, transforma uma EDO homogénea $\dot{x} = v(t,x)$ numa EDO separável $y + t\dot{y} = v(1,y)$. Ou seja,

$$\boxed{\dot{x} = v(1, x/t) \qquad \Rightarrow \qquad y + t\dot{y} = v(1,y) \qquad \text{se } \; y = x/t}$$

- Seja $\dot{x} = v(t,x)$ uma EDO homogénea. Mostre que, se $\varphi(t)$ é uma solução e $\lambda > 0$, então também $\phi(t) := \lambda \cdot \varphi(t/\lambda)$ é uma solução.
- Seja $\varphi(t)$ uma solução da EDO homogénea $\dot{x} = v(t,x)$ tal que $\varphi(1) = 5$ e $\varphi(2) = 7$. Se $\phi(t)$ é uma outra solução tal que $\phi(3) = 15$, quanto vale $\phi(6)$?
- Resolva as seguintes EDOs homogéneas
$$\dot{x} = -t/x \qquad \dot{x} = \frac{x-t}{x+t} \qquad \dot{x} = 1 + x/t$$
$$\dot{x} = x/t \qquad \dot{x} = 2\frac{t}{x}e^{x/t} + \frac{x}{t} \qquad \frac{dy}{dx} = y/x + \sin(y/x)\,,$$
definidas em oportunos domínios, e esboce a representação gráfica de algumas das soluções.

6. (EDOs quase-homogéneas) [Ar89] pag. 16.

7. (equação de Newton com forças homogéneas) [Ar89, LL78] Considere a equação de Newton

$$m\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r})$$

para a trajetória $t \mapsto \mathbf{r}(t) \in \mathbb{R}^3$ de uma partícula sujeita a una força homogénea de grau k, ou seja, tal que $\mathbf{F}(\lambda\mathbf{r}) = \lambda^k\mathbf{F}(\mathbf{r})$ para todos os $\lambda > 0$ e todo os pontos $\mathbf{r} \in \mathbb{R}^3\backslash\{0\}$. As "quase-homotetias"

$$(t, \mathbf{r}) \mapsto (\lambda^\alpha t, \lambda^\beta \mathbf{r}),$$

com $\lambda > 0$, enviam curvas integrais em curvas integrais se os "pesos" α e β satisfazem a relação

$$\beta(1-k) = 2\alpha .$$

Em particular, uma órbita fechada de dimensão linear L e período de revolução T é enviada numa órbita fechada de dimensão linear $L' = \lambda^\beta L$ e período de revolução $T' = \lambda^\alpha T$, e portanto o quociente T^β/L^α é constante.

- Considere uma força constante (e.g. a gravidade próximo da superfície da terra)

$$F(x) \propto 1 ,$$

  e determine a relação entre espaço percorrido e tempo necessário.

- Considere a forçã elástica (e.g. lei de Hooke, oscilador harmónico)

$$F(x) \propto -x ,$$

  e deduza que os períodos das órbitas fechadas não dependem das amplitudes das oscilações.

- Considere uma força elástica "fraca"

$$F(x) \propto -x^3 ,$$

  Determine o período das pequenas oscilações em quanto função da amplitude.

- Considere a força gravitacional

$$\mathbf{F}(\mathbf{r}) \propto -\frac{\mathbf{r}}{\|\mathbf{r}\|^3} ,$$

  e deduza a *terceira lei de Kepler*[16]: "os quadrados dos períodos de revolução T são proporcionais aos cubos das distâncias médias L do Sol aos planetas, ou seja, T^2/L^3 é uma constante".

[16] Johannes Kepler, *Harmonices mundi*, 1619.

8 EDOs exatas e campos conservativos

1. (EDOs exatas, diferenciais exatos e campos conservativos) O differencial $p(x,y)dx+q(x,y)dy$ é dito *exato* no domínio $\Omega \subset \mathbb{R}^2$ se existe uma função $U : \Omega \to \mathbb{R}$ de classe C^1, dita *primitiva*, tal que $dU = p\,dx + q\,dy$, ou seja,

$$\frac{\partial U}{\partial x} = p \qquad \text{e} \qquad \frac{\partial U}{\partial y} = q\,.$$

A primitiva $U(x,y)$ pode ser pensada como o *potencial* do campo de vetores $\mathbf{F} := -\nabla U = -(p,q)$. Se c é um valor *regular* de U (i.e. se $\nabla U \neq 0$ nos pontos onde $U(x,y) = c$), então a curva de nível

$$\Sigma_c := \{(x,y) \in \Omega \subset \mathbb{R}^2 \text{ t.q. } U(x,y) = c\}\,,$$

ortogonal ao campo de vetores $\mathbf{F}$, é uma solução implícita da *equação diferencial exata*

$$p(x,y)\,dx + q(x,y)\,dy = 0\,.$$

Por exemplo, numa vizinhança de um ponto onde $U(x,y) = c$ define implicitamente uma função $y = y(x)$, ao derivar esta relação descobrimos que $y(x)$ resolve a equação diferencial

$$p + q\frac{dy}{dx} = 0\,.$$

O *teorema de Euler-Poincaré* (caso particular do *teorema de Stokes*) afirma que o diferencial $p(x,y)dx + q(x,y)dy$, definido num domínio convexo[17] (é suficiente que seja "simplesmente conexo") $\Omega \subset \mathbb{R}^2$, é exato se e só se é "fechado", i.e. se

$$\left(\frac{\partial q}{\partial x} - \frac{\partial p}{\partial y}\right) dx \wedge dy = 0 \qquad \text{ou seja, se} \qquad \frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$$

Neste caso, um potencial é dado pelo integral de linha

$$U(x,y) = \int_\gamma \left(p(x,y)\,dx + q(x,y)\,dy\right)\,,$$

onde $\gamma : [0,1] \to D \subset \mathbb{R}^2$ é um caminho seccionalmente diferenciável entre $\gamma(0) = (x_0,y_0)$ e $\gamma(1) = (x,y)$ (ou seja, o (oposto do) trabalho feito pela força $\mathbf{F}$ para deslocar uma partícula do ponto (x_0,y_0) ao ponto (x,y)). Por exemplo, se Ω é um retângulo, é possível escolher um caminho horizontal de (x_0,y_0) até (x,y_0), e depois um caminho vertical de (x,y_0) até (x,y), e definir um potencial

$$\boxed{U(x,y) := \int_{x_0}^{x} p(t,y_0)\,dt + \int_{y_0}^{y} q(x,t)\,dt}$$

- Diga quais dos seguintes diferenciais,

$$dx + dy \qquad (2y+3x)\,dx + (y+2x)\,dy$$
$$e^{xy}dx + e^{xy}dy \qquad \frac{x}{y}dy + (1+\log y)\,dx\,,$$

definidos em oportunos retângulos, são exatos, e esboce algumas curvas integrais da correspondente equação diferencial.

- Diga quais das seguintes EDOs

$$5 + 3\frac{dx}{dt} = 0 \qquad (x-t)\,\frac{dx}{dt} + e^x = 0 \qquad \frac{1}{x} + t - \frac{t}{x^2}\frac{dx}{dt} = 0$$

$$\left(4x+3y^2\right)+2xy\frac{dy}{dx} = 0 \qquad 2x^2+4t^3+(4tx+1)\,\frac{dx}{dt} = 0 \qquad \left(r^2+1\right)\cos\theta+2r\sin\theta\frac{dr}{d\theta} = 0\,,$$

definidas em oportunos retângulos, são exatas, e resolva-as.

[17] Uma região $\Omega \subset \mathbb{R}^n$ é *convexa* se $a,b \in \Omega$ implica $ta + (1-t)b \in \Omega$ para todos os $t \in [0,1]$.

- Verifique que, se a forma $pdx + qdy$ está definida numa bola $B_R(0)$ de raio $0 < R \leq \infty$, e satisfaz a condição $\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$ do teorema de Euler-Poincaré, então uma primitiva é também dada pelo integral

$$U(x,y) = \int_0^1 \left(x\, p(tx,ty) + y\, q(tx,ty)\right) \, dt \, .$$

2. (fatores integrantes) Um diferencial arbitrário $pdx + qdy$ pode ser transformado num diferencial exato $(p\mu)dx + (q\mu)dy$ por meio de um *fator integrante*, uma função $\mu(x,y)$ tal que

$$\frac{\partial(\mu p)}{\partial y} = \frac{\partial(\mu q)}{\partial x} \, .$$

Não existem métodos gerais para determinar fatores integrantes. No entanto, fatores integrantes que dependem de apenas uma variável podem ser determinados, quando existem!, por meio de uma integração. Por exemplo, um fator integrante $\mu(x)$ é solução da EDO linear de primeira ordem

$$\mu \frac{\partial p}{\partial y} = \mu' q + \mu \frac{\partial q}{\partial x} \qquad \text{ou seja,} \qquad \mu'/\mu = \frac{1}{q}\left(\frac{\partial p}{\partial y} - \frac{\partial q}{\partial x}\right)$$

e portanto é igual a

$$\mu = e^{\int \frac{1}{q}\left(\frac{\partial p}{\partial y} - \frac{\partial q}{\partial x}\right) dx}$$

desde que $\frac{1}{q}\left(\frac{\partial p}{\partial y} - \frac{\partial q}{\partial x}\right)$ não depende de y.

- Considere as equações diferenciais

$$\left(4x + 3y^2\right) + 2xy\frac{dy}{dx} = 0 \qquad\qquad \left(2x^2 + y\right) + \left(x^2 y - x\right)\frac{dy}{dx} = 0 \, .$$

Mostre que não são exactas.
Determine um fator integrante da forma x^n com n inteiro. Multiplique as equações pelos respetivos fatores integrantes e resolva as equações resultantes.

3. (curvas ortogonais) Se a famíia $\mathcal{C}$ de curvas no plano é definida como sendo as curvas integrais da equação diferencial

$$p(x,y)\, dx + q(x,y)\, dy = 0 \, ,$$

então a família $\mathcal{C}^{\perp}$ de curvas ortogonais é composta pelas curvas integrais da equação diferencial

$$p(x,y)\, dy - q(x,y)\, dx = 0 \, .$$

(o operador que envia o diferencial $\omega = pdx + qdx$ no diferencial $*\omega = pdy - qdx$ é chamado "Hodge star operator" no plano euclidiano).

- Determine e esboce as curvas ortogonais ...
 ... à família de círculos $x^2 + y^2 = c$,
 ... à família de hipérboles $xy = c$,
 ... e à família de parábolas $y^2 = cx$.

4. (campos conservativos) O *trabalho* efectuado pelo campo de vetores/forças $\mathbf{F}(x,y) \in \mathbb{R}^2$ ou $\mathbf{F}(x,y,z) \in \mathbb{R}^3$, definido num domínio $D \subset \mathbb{R}^2$ ou $\mathbb{R}^3$, ao longo do caminho diferenciável $\gamma : [0,1] \to D$ é o integral de linha

$$\mathrm{W}[\mathbf{F},\gamma] := \int_\gamma \mathbf{F}(\gamma(t)) \cdot \dot{\gamma}(t) \, dt \, .$$

O campo de vectores $\mathbf{F}$ é *conservativo* se o trabalho apenas depende dos pontos inicial e final do caminho, ou seja, se o campo admite um *potencial*, i.e. uma função diferenciável $U : D \to \mathbb{R}$ tal que $\mathbf{F} = -\nabla U$, e portanto

$$\int_\gamma \mathbf{F}(\gamma(t)) \cdot \dot{\gamma}(t) \, dt = U(\gamma(1)) - U(\gamma(0)) \, .$$

- Diga quais dos seguintes campos de forças

$$\mathbf{F}(x,y) = (3,2) \qquad \mathbf{F}(x,y) = (x,y) \qquad \mathbf{F}(x,y) = (x,-y)$$

$$\mathbf{F}(x,y) = (-y,x) \qquad \mathbf{F}(x,y) = \left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2} \right) ,$$

  definidos em oportunos domínios do plano, são conservativos, e determine as curvas equipotenciais.

- Diga se o campo de forças

$$\mathbf{F}(x,y) = \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \right) ,$$

  definido em $\mathbb{R}\backslash\{(0,0)\}$, é conservativo.

9 EDOs lineares homogéneas com coeficientes constantes

ref: [Ap69] *Vol. 1*, **8.8-14** ; [MW85] *Vol. 2*, **12.6-7**

1. (equação de Newton num potencial quadrático) Considere a equação de Newton
$$\ddot{q} = -\beta q \tag{9.1}$$
que determina a trajetória $t \mapsto q(t) \in \mathbb{R}$ de uma partícula (de massa unitária) num potencial quadrático $U(q) = \frac{1}{2}\beta q^2$.
   - Verifique que $q(t) = 0$ é uma solução de equilíbrio.
   - Verifique que, se $\beta = 0$, as soluções da equação de Newton (9.1)
   $$\ddot{q} = 0$$
   (que neste caso é a equação da partícula livre) são $q(t) = a + bt$, com $a, b \in \mathbb{R}$ constantes arbitrárias.
   - Verifique que, se $\beta = -k^2 < 0$, as soluções da equação de Newton
   $$\ddot{q} = k^2 q$$
   são $q(t) = ae^{kt} + be^{-kt}$, com $a, b \in \mathbb{R}$ constantes arbitrárias.
   - Verifique que, se $\beta = \omega^2 > 0$, as soluções da equação (do oscilador harmónico)
   $$\ddot{q} = -\omega^2 q$$
   são $q(t) = a\cos(\omega t) + b\sin(\omega t)$, com $a, b \in \mathbb{R}$ constantes arbitrárias.

2. (partícula num potencial quadrático com atrito) A função $q(t) := e^{-\alpha t}y(t)$ é uma solução da equação de Newton de uma partícula num potencial quadrático $U(q) = \frac{1}{2}\beta q^2$ com uma força de atrito $-2\alpha\dot{q}$ (se $\alpha > 0$),
$$\ddot{q} = -2\alpha\dot{q} - \beta q\,, \tag{9.2}$$
se $y(t)$ é uma solução da equação de Newton
$$\ddot{y} = -\left(\beta - \alpha^2\right) y$$
de uma partícula num potencial quadrático $U(y) = \frac{1}{2}\left(\beta - \alpha^2\right) y^2$. A solução de (9.2) é uma combinação linear
$$x(t) = c_+ e^{-\alpha t}\varphi_+(t) + c_- e^{-\alpha t}\varphi_-(t)\,,$$
once c_+ e c_- são coeficientes, e φ_+ e φ_- é um par de *soluções fundamentais* (i.e. linearmente independentes) da EDO $\ddot{y} = \delta y$ com $\delta = \alpha^2 - \beta$, por exemplo:
$$\begin{aligned} \varphi_+(t) = 1 \quad &\text{e} \quad \varphi_-(t) = t\,, & \text{se} \quad &\delta = 0\,, \\ \varphi_+(t) = e^{kt} \quad &\text{e} \quad \varphi_-(t) = e^{-kt}\,, & \text{se} \quad &\delta = k^2 > 0\,, \\ \varphi_+(t) = \cos(\omega t) \quad &\text{e} \quad \varphi_-(t) = \sin(\omega t)\,, & \text{se} \quad &\delta = -\omega^2 < 0\,. \end{aligned}$$
   - Verifique a afirmação acima (calcule $\dot{q}$ e $\ddot{q}$ em quanto funções das derivadas de y ...) .
   - Existem soluções de equilíbrio da equação de Newton (9.2) ?
   - Mostre, nos três casos, que φ_+ e φ_- são linearmente independentes.

3. (existence and uniqueness theorem for homogeneous second order linear ODEs with constant coefficients) A generic *second order linear homogeneous ODE*
$$\ddot{x} + 2\alpha\dot{x} + \beta x = 0 \tag{9.3}$$
has at least two linearly independent solutions, as shown by the preceding examples.

Teorema 9.1. *The space of solutions of (9.3) is a 2-dimensional linear space $\mathcal{H} \approx \mathbb{R}^2$. Equivalently, the solution of (9.3) with any given initial conditions $x(0) = x_0$ and $\dot{x}(0) = v_0$ exists and is unique, and may be obtained as a linear combination of two independent solutions.*

Demonstração. It is sufficient to prove the result for the linear ODE $\ddot{x} = \delta x$. Also, by linearity, it is sufficient to show that the only solution of $\ddot{x} = \delta x$ with zero initial data $x(0) = 0$ and $\dot{x}(0) = 0$ is the trivial solution $x(t) = 0$.

The starting observation is that solutions of $\ddot{x} = \delta x$ are analytic functions (their Taylor series converges to the function). The way you prove it is an elementary instance of a strategy, called "bootstrap", which works for eigenfunctions of a Laplacian or more generally of any elliptic differential operator. The equation $\ddot{x} = \delta x$ implies that x admits derivatives all orders, and we can actually compute them. Indeed, $\ddot{x}' = (\ddot{x})' = \delta \dot{x}$, $x^{(4)} = (\ddot{x}')' = (\delta \dot{x})' = \delta \ddot{x} = \delta^2 x$, ..., and by induction you see that

$$x^{(2n)} = \delta^n x \quad \text{and} \quad x^{(2n+1)} = \delta^n \dot{x} \,.$$

Since x are $\dot{x}$ are bounded on a bounded interval (because they are continuous), the derivatives of x grow at most polynomially, say $x^{(k)}(t) \leq CK^k$ for some constants C and K and any t in a fixed bounded interval. Now you use the fact that a polynomial bound for the derivatives of a function in some bounded interval implies (by the Taylor formula with error, or, if you want, because the series is bounded by the Taylor series of an exponential) absolute convergence of the Taylor series.

Now, assume that $x(t)$ is a solution of $\ddot{x} = \delta x$ with initial conditions $x(0) = 0$ and $\dot{x}(0) = 0$. The above formulae show that all the derivatives of x at the origin are zero. There follows from analyticity that x is identically equal to zero on any bounded interval around the origin. □

4. (EDOs lineares homogéneas com coeficientes constantes, polinómio caraterístico) O uso dos exponenciais complexos permite uma leitura unificada dos três casos tratados acima. A conjectura $x(t) = e^{zt}$ é uma solução (complexa) da EDO *linear homogénea de segunda ordem com coeficientes constantes*

$$\ddot{x} + 2\alpha\dot{x} + \beta x = 0 \tag{9.4}$$

se z é igual a uma das raízes $z_\pm = -\alpha \pm \sqrt{\alpha^2 - \beta}$ do *polinómio caratéristico*

$$\boxed{P(z) := z^2 + 2\alpha z + \beta \,.}$$

Duas soluções (reais) independentes de (9.4) podem ser obtidas calculando a parte real e a parte imaginária das soluções complexas, e são

$e^{(-\alpha+k)t}$	e	$e^{(-\alpha-k)t}$	se $z_\pm = -\alpha \pm k$, com $k > 0$ (raízes reais e distintas)
$e^{-\alpha t}\cos(\omega t)$	e	$e^{-\alpha t}\sin(\omega t)$	se $z_\pm = -\alpha \pm i\omega$, com $\omega > 0$ (raízes complexas conjugadas)
$e^{-\alpha t}$	e	$te^{-\alpha t}$	se $z_\pm = -\alpha$ (raiz dupla)

O espaço das soluções (reais) da EDO homogénea (9.4) é um espaço linear $\mathcal{H} \approx \mathbb{R}^2$, de dimensão 2. Se $\phi_+(t)$ e $\phi_-(t)$ formam uma base de $\mathcal{H}$, então a "solução geral" é $x(t) = c_+\phi_+(t) + c_-\phi_-(t)$, onde $c_\pm \in \mathbb{R}$ são constantes arbitrárias.

- Verifique que o espaço das soluções da EDO homogénea (9.4) é um espaço linear, ou seja, que uma combinação linear $c_+x_+(t) + c_-x_-(t)$, com $c_\pm \in \mathbb{R}$, de soluções $x_\pm(t)$ é uma solução.
- Determine a solução geral das seguintes EDOs homogéneas:

$$\ddot{x} - 2x = 0 \qquad \ddot{x} + \pi^2 x = 0 \qquad 3\ddot{x} + \dot{x} = 0 \qquad \ddot{x} - \dot{x} = 0$$

$$\ddot{x} + 2\dot{x} - x = 0 \qquad \ddot{x} + 2\dot{x} + x = 0 \qquad \ddot{x} + 4\dot{x} + 5x = 0 \qquad \ddot{x} - 4\dot{x} + x = 0 \,.$$

- Resolva os seguintes problemas de Cauchy:

$$\ddot{x} + 2x = 0 \quad \text{com } x(0) = 0 \text{ e } \dot{x}(0) = 2$$

$$\ddot{x} + \dot{x} = 0 \quad \text{com } x(0) = 1 \text{ e } \dot{x}(0) = 0$$

$$\ddot{x} + 4\dot{x} + 5x = 0 \quad \text{com } x(0) = 2 \text{ e } \dot{x}(0) = -1$$

$$\ddot{x} - 17\dot{x} + 13x = 0 \quad \text{com } x(3) = 0 \text{ e } \dot{x}(3) = 0$$

$$\ddot{x} - 2\dot{x} - 2x = 0 \quad \text{com } x(0) = 0 \text{ e } \dot{x}(0) = 9$$

$$\ddot{x} - 4\dot{x} - x = 0 \quad \text{com } x(1) = 2 \text{ e } \dot{x}(1) = 1\,.$$

- Determine umas equações diferenciais de segunda ordem que admitem como soluções os seguintes pares de funções:

$$e^{2t} \quad \text{e} \quad e^{-2t}\,, \qquad e^{-t}\sin(2\pi t) \quad \text{e} \quad e^{-t}\cos(2\pi t)\,, \qquad \sinh(t) \quad \text{e} \quad \cosh(t)\,,$$

$$e^{-3t} \quad \text{e} \quad te^{-3t}\,, \qquad \sin(2t+1) \quad \text{e} \quad \cos(2t+2)\,, \qquad 3 \quad \text{e} \quad 5t\,.$$

5. (independência linear e Wronskiano) O *(determinante) Wronskiano* entre as funções $f(t)$ and $g(t)$, definidas num intervalo $I \subset \mathbb{R}$, é a função

$$\boxed{W_{f,g}(t) := \det \begin{pmatrix} f(t) & \dot{f}(t) \\ g(t) & \dot{g}(t) \end{pmatrix} = f(t)\dot{g}(t) - \dot{f}(t)g(t)}$$

Se $W_{f,g}(t) = 0$ para todos os tempos $t \in I$ então o quociente g/f (ou f/g) é constante no intervalo I. Consequentemente, se $f(t)$ e $g(t)$ são linearmente independentes, então o Wronskiano $W_{f,g}(t) \neq 0$ em algum ponto $t \in I$.

Se ϕ_+ e ϕ_- são duas soluções da mesma EDO linear $\ddot{x} + p(t)\dot{x} + q(t)x = 0$, definidas num intervalo $I \subset \mathbb{R}$, e $t_0 \in I$, então o Wronskiano satisfaz a equação diferencial

$$\frac{d}{dt} W_{\phi_+,\phi_-} = -p(t)\, W_{\phi_+,\phi_-}\,,$$

e portanto a *identidade de Abel*

$$W_{\phi_+,\phi_-}(t) = W_{\phi_+,\phi_-}(t_0) \cdot e^{-\int_{t_0}^{t} p(s)ds}\,.$$

De consequência, ϕ_+ e ϕ_- são linearmente independente se e só se $W_{\phi_+,\phi_-}(t) \neq 0$ num ponto (e portanto em todos os pontos) $t_0 \in I$.

- Calcule

$$W_{e^{-\alpha t},te^{-\alpha t}}\,, \quad W_{e^{-\alpha t}e^{kt},e^{-\alpha t}e^{-kt}} \quad \text{and} \quad W_{e^{-\alpha t}\sin(\omega t),e^{-\alpha t}\cos(\omega t)}\,.$$

6. (equação de Schrödinger estacionária) Considere a *equação de Schrödinger estacionária*

$$-\frac{\hbar^2}{2m}\frac{d^2\psi}{dx^2} = E\psi$$

para a função de onda $\psi(x)$ de uma partícula livre, onde m é a massa da partícula, $\hbar = h/2\pi$ é a constante de Planck reduzida, $h \simeq 6.262... \times 10^{-34}$ J·s.

- Determine para quais valores E da energia existem soluções não triviais da equação no intervalo $x \in [0,\ell]$ com condições de fronteira $\psi(0) = 0$ e $\psi(\ell) = 0$ (partícula numa caixa).

7. (EDOs equidimensionais) Uma equação diferencial da forma

$$ax^2\frac{d^2y}{dx^2} + bx\frac{dy}{dx} + cy = 0$$

é dita *equidimensional* (é invariante pela transformação $x \mapsto \lambda x$ com $\lambda > 0$).

- Mostre que a substituição $x = e^t$ transforma a equação equidimensional para $y(x)$ numa equação com coeficientes constantes para $z(t) := y(x(t))$.
- Resolva a equação
$$x^2 \frac{d^2y}{dx^2} + x\frac{dy}{dx} - 4y = 0\,,$$
na semirecta $x > 0$.

8. (EDOs de Riccati) Uma equação diferencial da forma
$$\ddot{y} + p(t)\,y^2 + q(t)\,y + r(t) = 0$$
é dita *equação de Riccati*.

- Mostre que a variável $x(t)$, solução da EDO simples
$$\dot{x} = -p(t)\,y(t)\,,$$
satisfaz a EDO linear
$$\ddot{x} + \left(q(t) + \frac{\dot{p}(t)}{p(t)}\right)\dot{x} + p(t)\,r(t)x = 0\,.$$

10 Números complexos e oscilações

ref: [Ap69] *Vol. 1*, **9.1-10** ; [MW85] *Vol. 2*, **12.6.**

1. (o plano dos números complexos) O corpo dos *números complexos* é o conjunto $\mathbb{C} \approx \mathbb{R}^2$ dos pontos/números $z = x + iy \approx (x, y)$, com $x, y \in \mathbb{R}$, munido das operações "soma", definida por
$$(x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2)$$
(que corresponde à soma dos vetores (x_1, y_1) e (x_2, y_2) do plano $\mathbb{R}^2$), e "multiplicação", definida por
$$(x_1 + iy_1) \cdot (x_2 + iy_2) = (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1)\,.$$
Em particular, se $i := 0 + i \cdot 1 \in \mathbb{C}$, então $i \cdot i = -1$, ou seja, $i = \sqrt{-1}$. O *conjugado* de $z = x + iy$ é $\overline{z} := x - iy$. O *módulo* de $z = x + iy$ é
$$|z| := \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}\,.$$
Os números reais
$$x = \Re(z) := \frac{z + \overline{z}}{2} \qquad \text{e} \qquad y = \Im(z) := \frac{z - \overline{z}}{2i}$$
são ditos *parte real* e *parte imaginária* do número complexo $z = x + iy$. A *representação polar* do número complexo $z = x + iy \approx (x, y) \in \mathbb{R}^2$ é
$$\boxed{z = \rho\, e^{i\theta}}$$
onde $\rho = |z| \geq 0$ é o módulo de z, $\theta \in \mathbb{R}$ é "um" *argumento* de z, ou seja, um "ângulo" $\arg(z) = \theta + 2\pi n$, com $n \in \mathbb{Z}$, tal que $x = \rho\cos(\theta)$ e $y = \rho\sin(\theta)$, e o número complexo $e^{i\theta}$ é definido pela *fórmula de Euler*
$$\boxed{e^{i\theta} := \cos(\theta) + i\sin(\theta)}$$
   - Verifique que o inverso multiplicativo de um número complexo $z \neq 0 := 0 + i0$ é
   $$1/z = \overline{z}/|z|^2$$
   - Represente na forma $x + iy$ os seguintes números complexos
   $$1/i \qquad \frac{2-i}{1+i} \qquad \frac{1-i}{1+i} \cdot \frac{i}{2+i} \qquad (1-i3)^2$$
   - Resolva as seguintes equações
   $$z^2 - 2z + 2 = 0 \qquad z^2 + z + 1 = 0$$
   - Verifique que, se $z_1 = \rho_1 e^{i\theta_1}$ e $z_2 = \rho_2 e^{i\theta_2}$, então
   $$z_1z_2 = \rho_1\rho_2 e^{i(\theta_1+\theta_2)} \qquad \text{e} \qquad \frac{z_1}{z_2} = \frac{\rho_1}{\rho_2} e^{i(\theta_1-\theta_2)} \quad (\text{se } \rho_2 \neq 0)\,.$$
   Deduza que a multiplicação por $z = \rho e^{i\theta}$, no plano $\mathbb{C} \simeq \mathbb{R}^2$, corresponde a uma dilatação/contração por ρ e uma rotação de um ângulo θ. Em particular, a multiplicação por $i = e^{i\pi/2}$ é uma rotação de um ângulo $\pi/2$, uma "raiz quadrada" da rotação $z \mapsto e^{i\pi}z = -z$ de um ângulo π.
   - Use a fórmula de Euler para provar as fórmulas
   $$\cos(\theta \pm \phi) = \cos(\theta)\cos(\phi) \mp \sin(\theta)\sin(\phi)$$
   e
   $$\sin(\theta \pm \phi) = \cos(\theta)\sin(\phi) \pm \sin(\theta)\cos(\phi)\,.$$

- Use a representação polar e a fórmula de Euler para provar a *fórmula de de Moivre*

$$(\cos(\theta) + i\sin(\theta))^n = \cos(n\theta) + i\sin(n\theta)\,.$$

  Deduza as fórmulas

$$\cos(n\theta) = \dots \qquad \text{e} \qquad \sin(n\theta) = \dots$$

- Verifique que o conjugado de $z = \rho e^{i\theta}$ é $\overline{z} = \rho e^{-i\theta}$.
- Calcule

$$\sqrt{i} \qquad \sqrt{-i} \qquad \sqrt{1+i}$$

- Resolva as equações $z^3 = 1$, $z^5 = 1$ e $z^3 = 81$.
- Mostre que se ω é uma raiz n-ésima não trivial da unidade (ou seja, $\omega^n = 1$ e $\omega \neq 1$) então

$$1 + \omega + \omega^2 + \omega^3 + \dots + \omega^{n-1} = 0\,.$$

  (multiplique por $1 - \omega$ ...).

2. (exponencial complexo e funções trigonométricas) A *função exponencial* $\exp(z) := e^z$, é a função inteira $\exp : \mathbb{C} \to \mathbb{C}$ definida pela série de potências

$$\boxed{e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \dots}$$

- Verifique a fórmula de adição $e^{z+w} = e^z e^w$, e deduza que $e^z \neq 0$ para todo o $z \in \mathbb{C}$.
- Verifique que e^z é igual à sua derivada, ou seja, $\exp'(z) = \exp(z)$.
- Mostre que, se $\theta \in \mathbb{R}$, então o conjugado de $e^{i\theta}$ é $e^{-i\theta}$, e portanto $|e^{i\theta}| = 1$. Defina as funções reais de variável real "cos" e "sin" usando a fórmula de Euler $e^{i\theta} = \cos\theta + i\sin\theta$, ou seja,

$$\cos(\theta) := \frac{e^{i\theta} + e^{-i\theta}}{2} \qquad \text{e} \qquad \sin(\theta) := \frac{e^{i\theta} - e^{-i\theta}}{2i}$$

  e deduza as suas expansões em série de potências em torno de 0.
- Deduza que, se $\alpha, \theta \in \mathbb{R}$,

$$e^{\alpha + i\theta} = e^{\alpha}\left(\cos\theta + i\sin\theta\right)$$

3. (oscilações complexas e sobreposições) A função $t \mapsto z(t) = e^{i\omega t}$ descreve um ponto que percorre a circunferência unitária $\mathbb{S}^1 := \{z \in \mathbb{C} \text{ t.q. } |z| = 1\}$ do plano complexo no sentido anti-horário com "frequência angular" $\omega > 0$, i.e. uma rotação cada período $T = 2\pi/\omega$, e portanto frequência $\nu = \omega/(2\pi)$ (medida em Hertz, rotações por segundo).

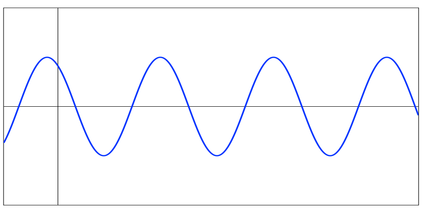
- Verifique qua a função $z(t) = e^{i\omega t}$ satisfaz as equações diferenciais lineares

$$\dot{z} = i\omega z \qquad \text{e} \qquad \ddot{z} = -\omega^2 z\,.$$

- Deduza que a parte real (e a parte imaginária) de $z(t) = z(0)e^{i\omega t}$, com $z(0) = \rho e^{i\varphi}$,

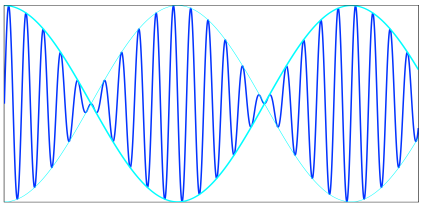
$$q(t) := \Re\left[z(t)\right] = \rho\cos(\omega t + \varphi)$$

  é uma solução (real) do oscilador harmónico $\ddot{q} = -\omega^2 q$. Identifique as condições iniciais $q(0)$ e $\dot{q}(0)$ em quanto funções de $z(0) = \rho e^{i\varphi}$.



Oscilação $q(t) = \rho\cos(\omega t + \varphi)$.

- Observe que a sobreposição das oscilações $z_1(t) = e^{i\omega_1 t}$ e $z_2(t) = e^{i\omega_2 t}$,
$$z(t) = e^{i\omega_1 t} + e^{i\omega_2 t} ,$$
é máxima quando $\omega_1 t = \omega_2 t$ (módulo 2π), e mínima quando $\omega_1 t - \omega_2 t = \pi$ (módulo 2π).
- Observe que, se $\omega_1 = \omega + \varepsilon$ e $\omega_2 = \omega - \varepsilon$, a sobreposição das duas oscilações $z_1(t) = e^{i\omega_1 t}$ e $z_2(t) = e^{i\omega_2 t}$ pode ser representada como
$$z(t) = e^{i\omega t} \left(e^{i\varepsilon t} + e^{-i\varepsilon t} \right) = 2e^{i\omega t} \cos(\varepsilon t)$$
Em particular, se $|\varepsilon| \ll |\omega|$, então a sobreposição consiste numa modulação lenta (com período $2\pi/\varepsilon \gg 2\pi/\omega$) da frequência fundamental $\omega \simeq \omega_1 \simeq \omega_2$.

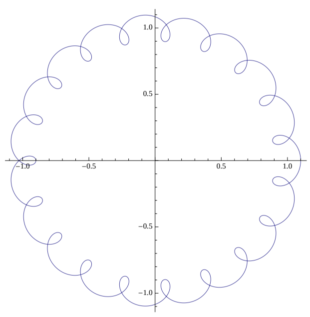


Sobreposição $z(t) = \sin(0.95 \cdot t) + \sin(1.05 \cdot t)$.

4. (epiciclos e deferentes) Números e circunferências (os objectos matemáticos mais elementares e as figuras geométricas mais perfeitas!) podem descrever os movimentos dos estrelas fixas e errantes com uma precisão espectacular. A ideia de Aristóteles e Platão, de que "todos os movimentos são combinações de movimentos circulares uniformes" está na base dos calendários calculados por Iparco e Ptolomeu, e transmitidos até nós pelos árabes no *Almagesto*.

   Nestas cosmologias, cada corpo celeste descreve uma circunferência, dita *epiciclo*, à volta de uma circunferência, que por sua vez descreve uma circunferência à volta de uma circunferência, ... , que por sua vez descreve uma circunferência à volta de uma circunferência inicial, dita *deferente*, centrada na Terra. Até o sistema de Nicholas Copernicus, apesar do escândalo que gerou, funciona assim: a única novidade, que também não uma novidade porque os próprios gregos consideraram esta possibilidade!, é que o centro do deferente é posto no Sol, ideia que pareceu simplificar muito o modelo.

   Indeed, any quase-periodic[18] (planar) motion may be approximated with arbitrary precision by a "superposition"
$$z(t) = a_0 e^{i\omega_0 t} + a_1 e^{i\omega_1 t} + \cdots + a_n e^{i\omega_n t}$$
   of a finite number of uniform circular motions. This is the content of modern Fourier analysis.



[18]G. Gallavotti, quase periodic motions from Hipparchus to Kolmogorov, *Rendiconti Lincei - Matematica e Applicazioni*, Series 9, Band 12, No. **2** (2001), 125-152 (http://arxiv.org/abs/chao-dyn/9907004).

11 Variação das constantes e coeficientes indeterminados

ref: [Ap69] *Vol. 1*, **8.15-19** ; [MW85] *Vol. 2*, **12.6-7**

1. (EDOs de segunda ordem lineares com coeficientes constantes) Uma *EDO de segunda ordem linear com coeficientes constantes* é uma lei
$$\ddot{x} + 2\alpha\dot{x} + \beta x = f(t) \tag{11.1}$$
para a trajetória $x(t)$, onde $\alpha, \beta \in \mathbb{R}$ são coeficientes constantes, e $f(t)$ é uma função dada (uma força externa dependente do tempo) definida num intervalo de tempos $I \subset \mathbb{R}$.
Se $x_1(t)$ e $x_2(t)$ são duas soluções de (11.1), então a diferença $y(t) = x_2(t) - x_1(t)$ é uma solução da *EDO homogénea associada*
$$\ddot{y} + 2\alpha\dot{y} + \beta y = 0\,, \tag{11.2}$$
obtida de (11.1) ao fazer $f(t) = 0$ (i.e. força nula). Portanto, a solução geral de (11.1) pode ser representada como uma soma
$$x(t) = z(t) + y(t)\,,$$
onde $z(t)$ é uma (apenas uma!) "solução particular" de (11.1) e $y(t) = c_+\phi_+(t) + c_-\phi_-(t)$ é a solução geral da EDO homogénea associada (11.2), combinação linear de duas soluções independentes $\phi_\pm(t)$ com coeficientes arbitrários $c_\pm \in \mathbb{R}$. Em particular, o espaço das soluções da EDO linear (11.1) é um plano afim $z + \mathcal{H}$, modelado sobre o espaço linear $\mathcal{H} \approx \mathbb{R}^2$ das soluções da EDO homogénea associada (11.2).
A procura de uma solução particular de (11.1) pode ser simplificada usando o *princípio de sobreposição*, consequência da linearidade do problema: se $x_1(t), x_2(t), \ldots, x_n(t)$ são soluções das EDOs lineares
$$\ddot{x_k} + 2\alpha\dot{x_k} + \beta x_k = f_k(t) \qquad \text{com} \qquad k = 1, 2, \ldots, n\,,$$
(observe que α e β são sempre os mesmos!) então a "sobreposição"
$$x(t) = x_1(t) + x_2(t) + \cdots + x_n(t)$$
é solução da EDO linear
$$\ddot{x} + 2\alpha\dot{x} + \beta x = f_1(t) + f_2(t) + \cdots + f_n(t)\,.$$

   - Determine a solução geral de
$$\ddot{x} = \sin(t) \qquad \ddot{x} + \dot{x} = t \qquad \ddot{x} + x = e^{-t}$$
   - Determine uma solução particular de
$$\ddot{x} = 1 + t + t^2 \qquad \ddot{x} + \dot{x} = e^{-t} - e^{-2t}$$

2. (variação das constantes) Uma solução particular da EDO linear não homogénea
$$\ddot{x} + 2\alpha\dot{x} + \beta x = f(t) \tag{11.3}$$
pode ser determinada usando a conjetura
$$z(t) = \lambda_+(t)\,\phi_+(t) + \lambda_-(t)\,\phi_-(t) \tag{11.4}$$
onde $\phi_+(t)$ e $\phi_-(t)$ são duas soluções independentes da equação homogénea $\ddot{y} + 2\alpha\dot{y} + \beta y = 0$, e $\lambda_\pm(t)$ são "coeficientes/parâmetros" variáveis. Um cálculo mostra que
$$\ddot{z} + 2\alpha\dot{z} + \beta z = \frac{d}{dt}\left(\dot{\lambda_+}\phi_+ + \dot{\lambda_-}\phi_-\right) + 2\alpha\left(\dot{\lambda_+}\phi_+ + \dot{\lambda_-}\phi_-\right) + \left(\dot{\lambda_+}\dot{\phi_+} + \dot{\lambda_-}\dot{\phi_-}\right)$$

Em particular, (11.4) é solução de (11.3) se (mas não só se!) as derivadas $\dot{\lambda}_\pm$ dos coeficientes satisfazem o sistema linear

$$\begin{cases} \dot{\lambda}_+\phi_+ + \dot{\lambda}_-\phi_- &= 0 \\ \dot{\lambda}_+\dot{\phi}_+ + \dot{\lambda}_-\dot{\phi}_- &= f \end{cases}$$

O determinante da matriz 2×2 que define o sistema é o Wronskiano

$$W_{\phi_+,\phi_-}(t) = \phi_+(t)\dot{\phi}_-(t) - \dot{\phi}_+(t)\phi_-(t)\,,$$

que é diferente de zero porque as $\phi_\pm$ são independentes. A única solução do sistema é

$$\dot{\lambda}_+ = -\frac{\phi_- f}{W_{\phi_+,\phi_-}} \qquad \dot{\lambda}_- = \frac{\phi_+ f}{W_{\phi_+,\phi_-}}\,,$$

e portanto os coeficientes podem ser umas primitivas

$$\boxed{\lambda_+ = -\int \phi_-(t)\frac{f(t)}{W_{\phi_+,\phi_-}(t)}dt\,, \qquad \lambda_-(t) = \int \phi_+(t)\frac{f(t)}{W_{\phi_+,\phi_-}(t)}dt\,,}$$

definidas a menos de constantes aditivas arbitrárias (que correspondem, em (11.4), a somar soluções da EDO linear homogénea).

- Determine uma solução particular das seguintes EDOs lineares, definidas em oportunos domínios, utilizando o método de variação dos parâmetros:

$$\ddot{x} + x = 1/\sin(t) \qquad \ddot{x} + 2\dot{x} + x = e^{-t} \qquad \ddot{x} + 4\dot{x} + 4x = e^{-2t}\log t\,.$$

$$\ddot{x} + x = \frac{\sin(t)}{\cos^2(t)} \qquad \ddot{x} + x = \tan(t) \qquad \ddot{x} - 4\dot{x} + 8x = \frac{e^{2t}}{\cos(2t)} \qquad .$$

3. (coeficientes indeterminados) O *método dos coeficientes indeterminados* permite determinar soluções particulares de uma EDO linear

$$\ddot{x} + 2\alpha\dot{x} + \beta x = f(t) \tag{11.5}$$

(ou, mais em geral, de uma EDO linear com coeficientes costantes de ordem n arbitrária) quando o segundo membro (a força) $f(t)$ é um quase-polinómio.

Um *quase-polinómio* é um produto $p(t)e^{\lambda t}$ de um polinómio $p(t) = p_0 + p_1 t + \dots + p_k t^k$ vezes um exponencial (real ou complexo, ou seja, um produto de um exponencial real $e^{\rho t}$ e funções trigonométricas, $\cos(\omega t)$ e/ou $\sin(\omega t)$, se $\lambda = \rho \pm i\omega$). Observe que os polinómios correspondem ao expoente $\lambda = 0$, e portanto são quase-polinómios.

Um operador diferencial linear com coeficientes constantes $L = \sum_n a_n d^n/dt^n$ envia um quase-polinómio $P(t)e^{\lambda t}$ de grau g num quase-polinómio $p(t)e^{\lambda t}$ com o mesmo expoente λ e grau $\leq g$. De consequência, se a força em (11.5) é um quase-polinómio

$$f(t) = e^{\rho t}\left(p(t)\cos(\omega t) + q(t)\sin(\omega t)\right)\,,$$

onde $p(t)$ e $q(t)$ são polinómios de grau $\leq g$, então a EDO (11.5) admite uma solução particular

$$z(t) = t^n\, e^{\rho t}\left(P(t)\cos(\omega t) + Q(t)\sin(\omega t)\right)\,,$$

onde $Q(t)$ e $P(t)$ são polinómios de grau $\leq g$, se $\lambda = \rho + i\omega$ é uma raiz do polinómio característico $z^2 + 2\alpha z + \beta$ da equação homogénea com multiplicidade $n \leq 2$ (se λ não for uma raiz, basta considerar $n = 0$). Os "coeficientes indeterminados" dos polinómios $P(t)$ e $Q(t)$ são obtidos ao igualar os termos de mesmo grau na (11.5), e portanto ao resolver um sistema linear de $2(g+1)$ equações em $2(g+1)$ incógnitas (no caso geral em que $\omega \neq 0$, ou a metade se $\omega = 0$).

Usando o princípio de sobreposição, é possível determinar soluções particulares quando o segundo membro $f(t)$ é uma combinação linear de quase-polinómios.

- Determine a solução geral das seguintes EDOs lineares utilizando o método dos coeficientes indeterminados.

$$\ddot{x} + x = t \qquad \ddot{x} - \dot{x} = t^2 \qquad \ddot{x} + 4\dot{x} + 3x = t^2 - 1 \qquad \ddot{x} - 4x = e^{-2t}$$

$$\ddot{x} + 2\dot{x} + x = t^3 e^{-t} + e^t \qquad \ddot{x} + x = \sin(t) \qquad \ddot{x} + 4x = 2t\cos(t)$$

$$\ddot{x} + 9x = \sin(\pi t) \qquad \ddot{x} + 4x = \cos(2t) \qquad \ddot{x} - 4x = te^{-2t} \qquad \ddot{x} + 4x = te^{-t}\cos(2t)\,.$$

4. (representação integral da resposta de um oscilador) Mostre que uma solução particular (com condições iniciais triviais) da equação do oscilador harmónico forçado $\ddot{x} + \omega^2 x = f(t)$ é

$$x(t) = \frac{1}{\omega}\int_0^t f(\tau)\,\sin\left(\omega(t-\tau)\right)\,d\tau\,. \tag{11.6}$$

Calcule o limite quando $\omega \to 0$, e deduza que uma solução particular da equação de Newton $\ddot{x} = f(t)$ é

$$x(t) = \int_0^t f(\tau)\,(t-\tau)\,d\tau\,.$$

Verifique e que uma solução particular da equação $\ddot{x} - k^2 x = f(t)$ é

$$x(t) = \frac{1}{k}\int_0^t f(\tau)\,\sinh\left(k(t-\tau)\right)\,d\tau\,.$$

5. (partícula num campo de forças dependente do tempo) Considere a equação de Newton

$$m\ddot{q} = -2\alpha\dot{q} + F(t)$$

de uma partícula de massa m sujeita a uma força $F(t)$, onde $2\alpha := 1/\tau \geq 0$ é um coeficiente de atrito. Sabendo que $q(0) = q_0$ e $\dot{q}(0) = v_0$, determine a trajectória quando a força é

- $F(t) = g$, ou seja, constante,
- $F(t) = -t^2$,
- $F(t) = F_0\cos(\gamma t)$,
- $F(t) = \sum_{i=1}^n F_i\cos(\gamma_i t)$.

12 Oscilador harmónico

ref: [MW85] *Vol. 2*, **12.6-7**

1. (oscilador harmónico) As pequenas oscilações de um pêndulo $\ddot{\theta} = -\omega^2 \sin(\theta)$ em torno da posição de equilíbrio estável $\theta = 0$ são descritas pela equação de Newton do *oscilador harmónico*

$$\ddot{q} = -\omega^2 q\,, \tag{12.1}$$

onde $\omega > 0$ é a "frequência (angular) característica". Esta é uma equação universal, pois descreve as pequenas oscilações de qualquer sistema Newtoniano unidimensional numa vizinhança de um equilíbrio estável genérico.[19] No espaço de fases $X = \mathbb{R}^2$, de coordenadas q e $p := \dot{q}$, a equação assume a forma do sistema

$$\begin{cases} \dot{q} = p \\ \dot{p} = -\omega^2 q \end{cases} . \tag{12.2}$$

- Mostre que a solução com condições iniciais $q(0) = q_0$ e $\dot{q}(0) = v_0$ é

$$q(t) = q_0 \cos(\omega t) + \frac{v_0}{\omega} \sin(\omega t)\,.$$

- Mostre que as trajectórias $q(t) = a\cos(\omega t) + b\sin(\omega t)$ podem ser escritas como

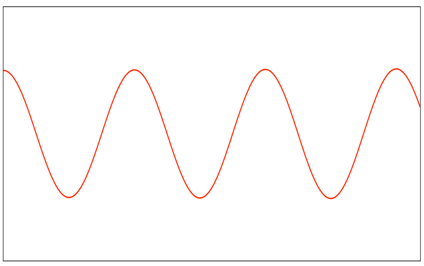
$$q(t) = A \sin(\omega t + \varphi) \qquad \text{ou} \qquad A\cos(\omega t + \phi)\ ,$$

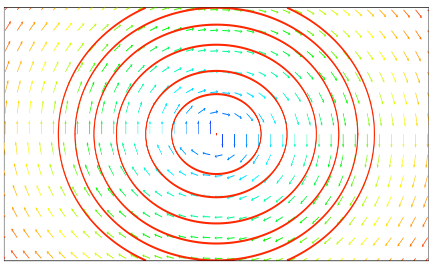
onde a amplitude A e as fases φ e ϕ dependem dos coeficientes a e b, ou seja, dos dados iniciais $q(0) = q_0$ e $\dot{q}(0) = v_0$ (use as fórmulas $\cos(a \pm b) = \cos(a)\cos(b) \mp \sin(a)\sin(b)$ e $\sin(a \pm b) = \sin(a)\cos(b) \pm \cos(a)\sin(b)$).

- Mostre que a energia

$$E(q,p) := \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2$$

é uma constante do movimento, ou seja que, se $(q(t), p(t))$ é uma solução do oscilador harmónico, então $\frac{d}{dt}E\left(q(t), p(t)\right) = 0$ para todo o tempo t. De consequência, as órbitas do oscilador harmónico estão contidas (de fato, são iguais!) nas curvas de nível da energia E, que são elipses.





Uma trajectória e retrato de fases do oscilador harmónico.

[19] "The harmonic oscillator, which we are about to study, has close analogs in many other fields; although we start with a mechanical example of a weight on a spring, or a pendulum with a small swing, or certain other mechanical devices, we are really studying a certain *differential equation*. This equation appears again and again in physics and other sciences, and in fact is a part of so many phenomena that its close study is well worth our while. Some of the phenomena involving this equation are the oscillations of a mass on a spring; the oscillations of charge flowing back and forth in an electrical circuit; the vibrations of a tuning fork which is generating sound waves; the analogous vibrations of the electrons in an atom, which generate light waves; the equations for the operation of a servosystem, such as a thermostat trying to adjust a temperature; complicated interactions in chemical reactions; the growth of a colony of bacteria in interaction with the food supply and the poison the bacteria produce; foxes eating rabbits eating grass, and so on; . . . "

Richard P. Feynman [Fe63]

- Mostre que a variável complexa $z := p + i\omega q$ satisfaz a EDO de primeira ordem

$$\dot{z} = i\omega z\,,$$

cuja solução é $z(t) = z(0)e^{i\omega t}$. Verifique que a energia do oscilador é dada por $E = \frac{1}{2}|z|^2$.

- Determine a energia em quanto função da amplitude e da frequência das oscilações.
- "Elimine" dt no sistema (12.2), e mostre que as curvas de fases são soluções da EDO exacta

$$p\,dp + \omega^2 q\,dq = 0\,,$$

equivalente a $dE = 0$, cujas soluções implícitas são as curvas de nível de E.

2. (oscilações amortecidas) Considere a equação das *oscilações amortecidas*

$$\ddot{q} = -2\alpha\dot{q} - \omega^2 q\,,$$

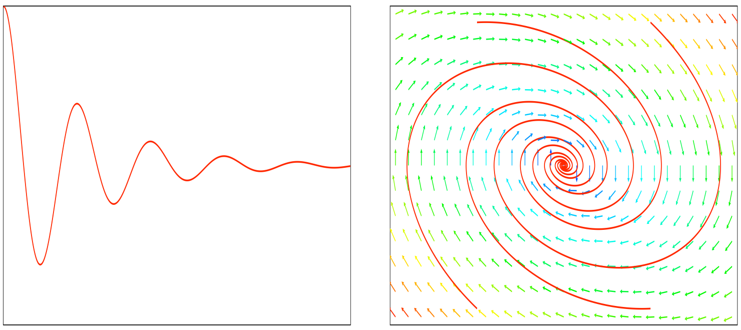
onde $2\alpha := 1/\tau > 0$ é um coeficiente de atrito (τ é o tempo de relaxamento). No espaço de fases, de coordenadas q e $p := \dot{q}$, a equação assume a forma do sistema

$$\begin{cases} \dot{q} = p \\ \dot{p} = -\omega^2 q - 2\alpha p \end{cases}.$$

- Mostre que as soluções do sistema "sub-crítico", ou seja, com $\alpha^2 < \omega^2$, são

$$q(t) = Ae^{-\alpha t}\sin\left(\sqrt{\omega^2 - \alpha^2}\,t + \varphi\right)$$

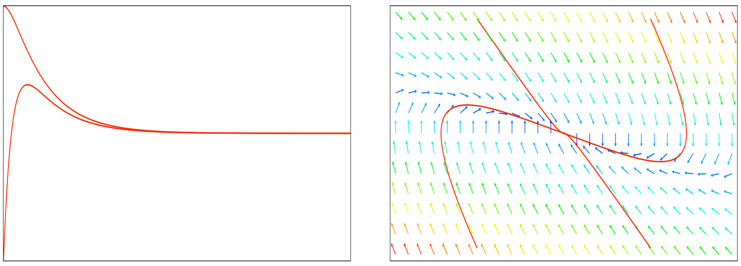
Observe que a frequência é $\sqrt{\omega^2 - \alpha^2} \simeq \omega - \frac{\alpha^2}{2\omega} + \dots$ se $\alpha \ll \omega$, mas tende para zero (e, de consequência, o período das oscilações tende para o ∞) quando $\alpha \to \omega$.



Trajectórias e retrato de fases do oscilador amortecido sub-crítico.

- Mostre que as soluções do sistema "super-crítico", ou seja, com $\alpha^2 > \omega^2$, são

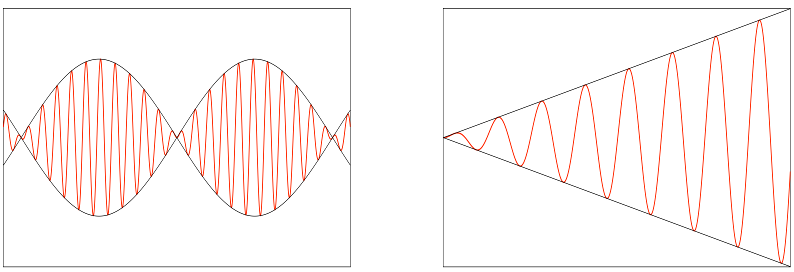
$$q(t) = Ae^{-\alpha t}\sinh\left(\sqrt{\alpha^2 - \omega^2}\,t + \varphi\right)$$



Trajectórias e retrato de fases do oscilador amortecido super-crítico.

- Mostre que as soluções do sistema "crítico", ou seja, com $\alpha^2 = \omega^2$ (uma condição muito difícil de observar!), são
$$q(t) = (a + bt)e^{-\alpha t} .$$
- Mostre que a energia
$$E(q,p) := \frac{1}{2}p^2 + \frac{1}{2}\omega^2 q^2$$
decresce a uma taxa proporcional à energia cinética, pois
$$\frac{d}{dt}E = -2\alpha\, p^2 \leq 0 .$$
Uma medida da perda de energia ao longo de um período é o *Q-factor* $Q := \omega/2\alpha$. Mostre que, se $\Delta E = E(t+T) - E(t)$, onde $T = 2\pi/\omega$, então
$$\frac{\Delta E}{E} \simeq \frac{2\pi}{Q} .$$

3. (oscilações forçadas, batimentos e ressonância) Considere a equação das *oscilações forçadas*
$$\ddot{q} = -\omega^2 q + F_0 \cos(\gamma t) .$$
- Mostre que, quando $\gamma^2 \neq \omega^2$, a solução geral é
$$q(t) = A\cos(\omega t + \phi) + \frac{F_0}{\omega^2 - \gamma^2}\cos(\gamma t)$$
onde A e ϕ são constantes arbitrárias.
- Verifique que a solução com condições iniciais triviais pode ser escrita
$$\begin{aligned} q(t) &= \frac{F_0}{\omega^2 - \gamma^2}\left(\cos(\gamma t) - \cos(\omega t)\right) \\ &= \frac{F_0}{\omega^2 - \gamma^2} 2\sin\left(\frac{\omega - \gamma}{2}t\right) \cdot \sin\left(\frac{\omega + \gamma}{2}t\right) \end{aligned}$$
Quando a diferença $2\varepsilon := \omega - \gamma$ é pequena, ou seja $|\varepsilon| \ll |\omega|$, e portanto $\frac{\omega+\gamma}{2} \simeq \omega$, podemos estimar
$$q(t) \simeq \frac{F_0}{2\omega\varepsilon}\sin(\varepsilon t) \cdot \sin(\omega t) .$$
Portanto, a resposta do oscilador à força externa é uma "modulação" lenta (de período $2\pi/\varepsilon \gg 2\pi/\omega$) de uma oscilação com frequência fundamental ω. Este fenómeno é chamado "batimentos" (em inglês, *beats*). Calcule o limite da resposta $q(t)$ quando $\varepsilon \to 0$.
- Mostre que, quando $\gamma^2 = \omega^2$, a solução com condições iniciais triviais é
$$q(t) = \frac{F_0}{2\omega}\, t\, \sin(\omega t) .$$
Este fenómeno, uma resposta cuja amplitude cresce linearmente no tempo, é chamado "ressonância".



Batimentos e ressonância.

4. (oscilações forçadas em notação complexa) Considere a equação das *oscilações forçadas*

$$\ddot{q} = -\omega^2 q + F(t)\,.$$

- Verifique que a variável complexa $z := p + i\omega q$ satisfaz a EDO linear de primeira ordem

$$\dot{z} - i\omega z = F(t)\,.$$

Uma solução não trivial da EDO homogénea associada $\dot{y} - i\omega y = 0$ é $y(t) = e^{i\omega t}$. Use o método da variação das constantes para determinar a solução na forma de um produto $z(t) = \lambda(t)e^{i\omega t}$, onde λ é solução de $\dot{\lambda} = F(t)e^{-i\omega t}$. Deduza que

$$z(t) = e^{i\omega t}\left(z(t_0) + \int_{t_0}^{t} F(s)\, e^{-i\omega\tau}\, d\tau\right)\,.$$

- Verifique que a energia cedida ao oscilador por uma força $F(t)$ num intervalo de tempos $(-\infty, \infty)$ é dada por [LL78]

$$E = \tfrac{1}{2}\,|z(\infty)|^2 = \tfrac{1}{2}\left|\int_{-\infty}^{\infty} F(t)\, e^{-i\omega t}\, dt\right|^2\,.$$

5. (oscilações forçadas amortecidas) Considere a equação das *oscilações forçadas amortecidas*

$$\ddot{q} = -2\alpha\dot{q} - \omega^2 q + F(t)\,,$$

onde $2\alpha := 1/\tau > 0$ é um coeficiente de atrito, e a força é $F(t) = F_0 \sin(\gamma t)$.

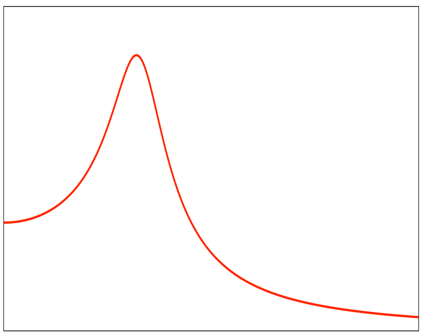
- Mostre que, se $\alpha^2 < \omega^2$ (ou seja, se o sistema não forçado é sub-crítico), a solução geral é

$$q(t) = A\, e^{-\alpha t}\, \sin\left(\sqrt{\omega^2 - \alpha^2}\, t + \varphi\right) + R(\gamma)\, F_0\, \sin\left(\gamma t + \phi\right)\,,$$

onde a amplitude A e a fases φ dependem dos dados iniciais,

$$R(\gamma) = \frac{1}{\sqrt{\left(\omega^2 - \gamma^2\right)^2 + 4\alpha^2\gamma^2}} \qquad \text{e} \qquad \tan\phi = -\frac{2\alpha\gamma}{\omega^2 - \gamma^2}\,.$$

A primeira parcela da solução representa um "regime transitório" (transiente), desprezável para grandes valores do tempo (i.e. para $t \gg \tau$). A segunda é dita "solução estacionária", e representa a resposta sincronizada, mas desfasada, do sistema à força periódica. A função $R(\gamma)$ é dita *curva de ressonância* do sistema, pois representa o factor de proporcionalidade entre a amplitude da força e a amplitude da resposta.



Um exemplo de curva de ressonância $R(\gamma)$.

- Mostre que a curva de ressonância $R(\gamma)$ atinge um máximo para o valor

$$\gamma_r = \sqrt{\omega^2 - 2\alpha^2}$$

da frequência, chamada *frequência de ressonância*. Observe que, se $\omega\tau \gg 1$, então $\gamma_r \simeq \omega(1 - 1/(4\tau^2\omega^2) + \dots)$.

- Discuta também os casos $\alpha^2 = \omega^2$ e $\alpha^2 > \omega^2$.

6. (circuito RLC) A corrente $I(t)$ num circuito RLC, de resistência R, indutância L e capacidade C, é determinada pela EDO
$$L\ddot{I} + R\dot{I} + \frac{1}{C}I = \dot{V} ,$$
onde $V(t)$ é a tensão que alimenta o circuito.

   - Determine a corrente $I(t)$ num circuito alimentado com uma tensão constante $V(t) = V_0$, e esboce as soluções (compare com a equação das oscilações amortecidas).
   - Determine a corrente $I(t)$ num circuito alimentado com uma tensão alternada $V(t) = V_0 \sin(\gamma t)$ (compare com a equação das oscilações forçadas amortecidas).
   - Determine a frequência de ressonância do circuito.

13 Espaços lineares

ref: [Ap69] *Vol. 2*, **1.1-10** ; [La87] *Ch. I*

1. (espaços lineares/vetoriais) Um *espaço linear/vetorial real* (ou seja, sobre o corpo $\mathbb{R}$ dos números reais) é um conjunto $\mathbf{V}$ munido de duas operações:

   a "adição" : $\mathbf{V} \times \mathbf{V} \to \mathbf{V}$
   $$\mathbf{v}\,,\,\mathbf{w} \mapsto \mathbf{v} + \mathbf{w}\,,$$
   que satisfaz os axiomas

   EL1 *(propriedade associativa)* $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$,
   EL2 *(propriedade comutativa)* $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$,
   EL3 *(existência do elemento neutro)* existe $\mathbf{0} \in \mathbf{V}$, tal que $\mathbf{0} + \mathbf{v} = \mathbf{v}$
   EL4 *(existência do simétrico)* $\forall \mathbf{v} \in \mathbf{V}$ existe $-\mathbf{v} \in \mathbf{V}$ tal que $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

   e a "multiplicação por escalares/números" : $\mathbb{R} \times \mathbf{V} \to \mathbf{V}$
   $$\lambda\,,\,\mathbf{v} \mapsto \lambda\mathbf{v} \in \mathbf{V}$$
   que satisfaz os axiomas

   EL5 *(propriedade associativa)* $(\lambda\mu)\mathbf{v} = \lambda(\mu\mathbf{v})$,
   EL6 *(propriedade distributiva para a adição em $\mathbb{R}$)* $(\lambda + \mu)\mathbf{v} = \lambda\mathbf{v} + \mu\mathbf{v}$,
   EL7 *(propriedade distributiva para a adição em $\mathbf{V}$)* $\lambda(\mathbf{v} + \mathbf{w}) = \lambda\mathbf{v} + \lambda\mathbf{w}$,
   EL8 *(existência do elemento neutro)* $1\mathbf{v} = \mathbf{v}$,

   Um *isomorfismo* entre os espaços lineares $\mathbf{V}$ e $\mathbf{V}'$ é uma aplicação bijectiva $f : \mathbf{V} \to \mathbf{V}'$, $\mathbf{v} \leftrightarrow \mathbf{v}'$, que respeita as operações, i.e. tal que se $\mathbf{v} \leftrightarrow f(\mathbf{v}) = \mathbf{v}'$ e $\mathbf{w} \leftrightarrow f(\mathbf{w}) = \mathbf{w}'$, então
   $$\mathbf{v} + \mathbf{w} \leftrightarrow \mathbf{v}' + \mathbf{w}' \qquad \text{e} \qquad \lambda\mathbf{v} \leftrightarrow \lambda\mathbf{v}' \qquad \forall \lambda \in \mathbb{R}\,.$$
   Se substituimos o corpo $\mathbb{R}$ dos números reais pelo corpo $\mathbb{C}$ dos números complexos, obtemos a definição de *espaço linear/vetorial complexo*.

   - Verifique que $\mathbb{R}$, munido das operações usuais "+" e "·", é um espaço vetorial real.
   - Verifique que $\mathbb{R}^n$, munido das operações "adição" e "produto por um escalar" definidas no exercício 1 do capítulo 1, é um espaço vetorial real.
   - Mostre que o elemento neutro $\mathbf{0}$ é único.
   - Mostre que o simétrico $-\mathbf{v}$ de cada $\mathbf{v} \in \mathbf{V}$ e único.
   - Mostre que $0\mathbf{v} = \mathbf{0}$ para todo $\mathbf{v} \in \mathbf{V}$ e que $\lambda\mathbf{0} = \mathbf{0}$ para todo $\lambda \in \mathbb{R}$.
   - Mostre que $\lambda\mathbf{v} = \mathbf{0}$ implica $\mathbf{v} = \mathbf{0}$ ou $\lambda = 0$.
   - Mostre que $\lambda\mathbf{v} = \mu\mathbf{v}$ implica $\mathbf{v} = \mathbf{0}$ ou $\lambda = \mu$.
   - Mostre que $\lambda\mathbf{v} = \lambda\mathbf{w}$ implica $\mathbf{v} = \mathbf{w}$ ou $\lambda = 0$.
   - [Ap69] **15.5**.

2. (o espaço linear complexo $\mathbb{C}^n$) O *espaço vetorial complexo* de *dimensão* n é o espaço
   $$\mathbb{C}^n := \underbrace{\mathbb{C} \times \mathbb{C} \times \cdots \times \mathbb{C}}_{n \text{ vezes}}$$
   das n-uplas $\mathbf{z} = (z^1, z^2, \ldots, z^n)$ de números complexos, munido das operações *adição* $+ : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$, definida por
   $$\boxed{\mathbf{z}, \mathbf{w} \mapsto \mathbf{z} + \mathbf{w} := (z^1 + w^1, z^2 + w^2, \ldots, z^n + w^n)}$$
   e *multiplicação por um escalar* $\cdot : \mathbb{C} \times \mathbb{C}^n \to \mathbb{C}^n$, definida por
   $$\boxed{\lambda, \mathbf{z} \mapsto \lambda\mathbf{z} := (\lambda z^1, \lambda z^2, \ldots, \lambda z^n)}$$

3. (espaço afim) Um *espaço afim* modelado sobre o espaço vetorial $\mathbf{V}$ é um conjunto $\mathcal{A}$ munido de uma aplicação $\mathcal{A} \times \mathcal{A} \to \mathbf{V}$
$$P, Q \mapsto \vec{PQ}$$
que satisfaz os axiomas

   EA1 para cada $P \in \mathcal{A}$ e cada $\mathbf{v} \in \mathbf{V}$ existe um único $Q \in \mathcal{A}$ tal que $\vec{PQ} = \mathbf{v}$

   EA2 para quaisquer $P, Q, R \in \mathcal{A}$,
$$\vec{PQ} + \vec{QR} = \vec{PR}$$

   - Verifique que o conjunto $\mathbb{R}^n$ munido da aplicação $P, Q \mapsto Q - P$ é um espaço afim modelado sobre o espaço vetorial real $\mathbb{R}^n$.

4. (espaços de funções) Os espaços inrteressantes em análise, em física e em engenharia, são espaços de funções, chamados "espaços funcionais".

   Sejam X um conjunto e $\mathbb{R}^X = \mathcal{F}(X, \mathbb{R})$ ou $\mathbb{C}^X = \mathcal{F}(X, \mathbb{C})$ os espaços das funções reais ou complexas definidas em X, cujos elementos são denotados por $x \mapsto f(x)$. Os espaços $\mathcal{F}(X, \mathbb{R})$ e $\mathcal{F}(X, \mathbb{C})$, munidos das operações "adição" e "produto por um escalar" definidas por
$$(f+g)(x) := f(x) + g(x) \qquad \text{e} \qquad (\lambda f)(x) := \lambda f(x)\,,$$
são espaços vetoriais reais e complexos, respetivamente. Também interessantes são os 'campos vetoriais", funções com valores num espaço vetorial como $\mathbb{R}^m$ (campos de força, de velocidade, campo eletro-magnético, ...

   Engenheiros e físicos estão por exemplo interessados em "sinais" $f(t)$ (a intensidade de uma onda de som, onde t é o tempo), ou "funções de onda" $\psi(\mathbf{r}, t)$, em mecânica quânticas, ou outros "campos" $u(\mathbf{r}, t)$ (um deslocamento, um campo de velocidades, o campo eletro-magnético, ...) que resolvem certas equações diferenciais parciais como a equação de onda, de calor, de Laplace, de Schrödinger, ...).

   Um sinal contínuo $f(t)$ podes ser observado apenas em múltiplos inteiros de um tempo de amostragem $\tau > 0$, e transformado num sinal discreto $x[n] := f(n\tau)$, com $n \in \mathbb{Z}$ ou $\mathbb{N}$, que é uma sequência.

   - Mostre que $\mathcal{F}(\mathbb{R}, \mathbb{R})$ é um espaço vetorial real, determine o elemento neutro, e dê exemplos de elementos de $\mathcal{F}(\mathbb{R}, \mathbb{R})$.
   - Seja $\mathbb{R}^\infty := \mathbb{R}^\mathbb{N} = \mathcal{F}(\mathbb{N}, \mathbb{R})$ o espaço das sucessões reais $\mathbf{x} = (x_1, x_2, \dots, x_n, \dots)$, com $x_n \in \mathbb{R}$. Descreva a sua estrutura de espaço linear real. Mostre que o espaço b das sucessões limitadas, o espaço c das sucessões convergentes, o espaço c_0 das sucessões que convergem para 0, e o espaço ℓ das sucessões com suporte compacto (i.e. tais que $x_n = 0$ se n é suficientemente grande) são subespaços vetoriais de $\mathbb{R}^\infty$, e que
$$\ell \subset c_0 \subset c \subset b \subset \mathbb{R}^\infty\,.$$
   - Seja $\mathcal{P}(\mathbb{R}) := \mathbb{R}[t]$ o espaço dos polinómios com coeficientes reais na variável t, e $\mathcal{P}_{\leq n}(\mathbb{R})$ o espaço dos polinómios reais de grau $\leq n$. Verifique que $\mathcal{P}(\mathbb{R})$ e $\mathcal{P}_{\leq n}(\mathbb{R})$ são espaços lineares.
   - Verifique que são espaços vetoriais, reais ou complexos, o espaço $\mathcal{C}^0(\mathbb{R})$ das funções contínuas $f : \mathbb{R} \to \mathbb{R}$ ou $\mathbb{C}$, os espaços $\mathcal{C}^k(\mathbb{R})$ das funções com k derivadas contínuas, o espaço $\mathcal{C}^\infty(\mathbb{R})$ das funções infinitamente deriváveis. As inclusões sendo
$$\mathcal{C}^\infty(\mathbb{R}) \subset \cdots \subset \mathcal{C}^{k+1}(\mathbb{R}) \subset \mathcal{C}^k(\mathbb{R}) \subset \cdots \subset \mathcal{C}^0(\mathbb{R})$$
   - [Ap69] **15.5**.

5. (superposition principle for linear ODEs) Consider a "linear homogeneous" ordinary differential equation, for example with constant coefficients and of second order like
$$\ddot{y} + \alpha \dot{y} + \beta y = 0\,. \qquad (13.1)$$

A superposition $ay_1(t) + by_2(t)$ of any two solutions, $y_1(t)$ and $y_2(t)$, is still a solution. Therefore, the space $\mathbf{H}$ of solutions of (13.1) is a linear space (a subspace of the infinite-dimensional space $\mathcal{C}^2(\mathbb{R})$ of twice differentiable functions on the real line). Actually, as you will see in analysis, this space is two-dimensional, so that $\mathbf{H} \approx \mathbb{R}^2$, and therefore any solution is a superposition

$$y(t) = c_+\, y_+(t) + c_-\, y_-(t)$$

of two basic independent solutions $y_\pm(t)$, forming a base of $\mathbf{H}$. We now add a time-dependent force $f(t)$, and consider the linear ordinary differential equation

$$\ddot{x} + \alpha\dot{x} + \beta x = f(t)\,. \tag{13.2}$$

The difference $x_1(t) - x_2(t)$ between any two solutions of (13.2) is a solution of (13.1). Therefore, if $z(t)$ is any (particular) solution of (13.2), then the space of all solutions of (13.2) is the affine space $z + \mathbf{H}$.

6. (quantum superposition principle) If a physical system can be prepared in each of the states $|1\rangle$, $|2\rangle$, $|3\rangle$, ..., for example corresponding to certain values $\lambda_1, \lambda_2, \lambda_3, \dots$ of a certain observable L, then the most general state is a *superposition* [20]

$$|\psi\rangle = \psi_1\,|1\rangle + \psi_2\,|2\rangle + \psi_3\,|3\rangle + \dots$$

with complex coefficients $\psi_n \in \mathbb{C}$. The observation of the observable L in the state $|\psi\rangle$ will then give one of the value λ_n (and not a value in between!) with relative frequency (if the experience is repeated a large number of times) proportional to $|\psi_n|^2$. States of a quantum system therefore live in a complex linear space $\mathbf{H}$ (which is infinite dimensional). Actually, proportional states must be considered as equal, which amounts to say that we may only consider normalized states, those with $|\psi_1|^2 + |\psi_2|^2 + |\psi_3|^2 + \cdots = 1$, and do not distinguish between those which differ by a multiplicative phase $e^{i\phi}$ (although this ambiguity is then central in quantum field theory). Therefore the state space of a quantum system is a projective space $\mathbb{P}\mathbf{H}$ of a complex linear space $\mathbf{H}$.

7. (subespaços e geradores) Seja $\mathbf{V}$ um espaço vetorial real ou complexo. Um subconjunto não vazio $W \subset \mathbf{V}$ que é também um espaço vetorial (ou seja, tal que $w + w' \in W$ e $\lambda w \in W$ para todos os $w, w' \in W$ e $\lambda \in \mathbb{R}$) é dito *subespaço (linear/vetorial)* de $\mathbf{V}$. Se $S \subset \mathbf{V}$ é um subconjunto de $\mathbf{V}$, o conjunto $\mathrm{Span}(S)$ das combinações lineares finitas

$$\lambda^1\mathbf{s}_1 + \lambda^2\mathbf{s}_2 + \cdots + \lambda^m\mathbf{s}_m \qquad\qquad \mathbf{s}_i \in S\,,\ \lambda^i \in \mathbb{R} \text{ ou } \mathbb{C}$$

é um subsespaço de $\mathbf{V}$, dito subespaço *gerado* por S. O espaço linear $\mathbf{V}$ tem "dimensão finita" se admite um conjunto finito de geradores.

Se X e Y são subespaços de $\mathbf{V}$, então a "soma"

$$X + Y := \{\mathbf{x} + \mathbf{y} \text{ com } \mathbf{x} \in X,\, \mathbf{y} \in Y\}$$

é um subespaço de $\mathbf{V}$. Se cada vetor $\mathbf{v} \in X + Y$ admite uma única representação $\mathbf{v} = \mathbf{x} + \mathbf{y}$ com $\mathbf{x} \in X$ e $\mathbf{y} \in Y$, então a soma é chamada *soma direta*, e denotada por $X \oplus Y$.

[20] "Any state may be considered as the result of a superposition of two or more other states, and indeed in an infinite number of ways. Conversely any two or more states may be superposed to give a new state ... The non-classical nature of the superposition process is brought out clearly if we consider the superposition of two states, A and B, such that there exists an observation which, when made on the system in state A, is certain to lead to one particular result, a say, and when made on the system in state B is certain to lead to some different result, b say. What will be the result of the observation when made on the system in the superposed state? The answer is that the result will be sometimes a and sometimes b, according to a probability law depending on the relative weights of A and B in the superposition process. It will never be different from both a and b [i.e, either a or b]. The intermediate character of the state formed by superposition thus expresses itself through the probability of a particular result for an observation being intermediate between the corresponding probabilities for the original states, not through the result itself being intermediate between the corresponding results for the original states."
.

P.A.M. Dirac [Di47]

- Se $X_1, X_2, \dots$ são subespaços de $\mathbf{V}$, então
$$\bigcap_i X_i := \{\mathbf{x} \text{ t.q. } \mathbf{x} \in X_i \,\forall i\}$$
é um subespaço de $\mathbf{V}$.
- Dados os números $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{R}$, o conjunto dos vetores $\mathbf{x} \in \mathbb{R}^n$ tais que
$$\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0$$
é um subespaço linear de $\mathbb{R}^n$.
- Se V é um subespaço do espaço euclidiano $\mathbb{R}^n$, então
$$V^\perp := \{\mathbf{x} \in \mathbb{R}^n \text{ t.q. } \mathbf{x} \cdot \mathbf{v} = 0 \quad \forall \mathbf{v} \in V\}$$
é um subespaço de $\mathbb{R}^n$.
- Verifique que $\mathcal{C}(\mathbb{R}, \mathbb{R})$ é um subespaço de $\mathcal{F}(\mathbb{R}, \mathbb{R})$.
- Verifique que $\mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ é um subespaço de $\mathcal{C}^{k+1}(\mathbb{R}, \mathbb{R})$, que é um subespaço de $\mathcal{C}^k(\mathbb{R}, \mathbb{R})$, que é um subespaço de $\mathcal{C}(\mathbb{R}, \mathbb{R})$.
- Verifique que $\mathcal{P}_{\leq n}(\mathbb{R})$ é um supespaço de $\mathcal{P}(\mathbb{R})$, que é um subespaço de $\mathcal{C}(\mathbb{R}, \mathbb{R})$. O espaço $\mathcal{P}_n(\mathbb{R})$ dos polinómios de grau (exatamente) n é um suberspaço vetorial de $\mathcal{P}(\mathbb{R})$?
- Determine os subespaços gerados por
$$(1,2) \qquad (-1,-2) \qquad \text{em } \mathbb{R}^2$$
$$(0,1,0) \qquad (0,0,1) \qquad \text{em } \mathbb{R}^3$$
$$\mathbf{i} \quad \mathbf{j} \qquad \mathbf{i}+\mathbf{j} \qquad \text{em } \mathbb{R}^3$$
$$\{1, t, t^2, \dots, t^n, \dots\} \qquad \text{em } \mathbb{R}^\mathbb{R}$$
$$t \qquad t^2 \qquad \text{em } \mathcal{P}(\mathbb{R})$$
$$e^t \qquad e^{-t} \qquad \text{em } \mathcal{F}(\mathbb{R}, \mathbb{R})$$
- O conjunto b das sucessões $(x_n)_{n\in\mathbb{N}}$ limitadas é um subespaço do espaço das sucessões reais $\mathbb{R}^\infty := \mathbb{R}^\mathbb{N}$? E o conjunto c das sucessões convergentes? E o conjunto c_0 das sucessões convergentes tais que $x_n \to 0$?
- O espaço das funções não negativas, i.e. tais que $f(t) \geq 0 \; \forall t$, é um subespaço do espaço $\mathcal{F}(\mathbb{R}, \mathbb{R})$?
- Os conjuntos das funções pares e ímpares, definidos por
$$\mathcal{F}_\pm(\mathbb{R}, \mathbb{R}) = \{f \in \mathcal{F}(\mathbb{R}, \mathbb{R}) \quad \text{t.q. } f(t) = \pm f(-t) \quad \forall t \in \mathbb{R}\},$$
são subespaços de $\mathcal{F}(\mathbb{R}, \mathbb{R})$? Mostre que cada $f \in \mathcal{F}(\mathbb{R}, \mathbb{R})$ é uma soma $f = f_+ + f_-$ com $f\pm \in \mathcal{F}_\pm(\mathbb{R}, \mathbb{R})$.
- [Ap69] **15.9**.

8. (conjuntos livres/linearmente independentes) Seja $\mathbf{V}$ um espaço linear. O conjunto $S \subset \mathbf{V}$ é *livre/(linearmente) independente* se gera cada vetor de $\operatorname{Span}(S)$ duma única maneira, ou seja, se gera o vetor nulo $\mathbf{0}$ duma única maneira, i.e. se quaisquer que sejam os elementos distintos $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_m \in S$,
$$\boxed{\lambda^1 \mathbf{s}_1 + \lambda^2 \mathbf{s}_2 + \cdots + \lambda^m \mathbf{s}_m = \mathbf{0} \qquad \Longrightarrow \qquad \lambda^i = 0 \quad \forall i = 1, 2, \dots, m}$$
Caso contrário, o conjunto é dito *(linearmente) dependente*.

- Verifique se os seguintes conjuntos de vetores são linearmente independentes

$$(1,2) \qquad (-1,2) \qquad \text{em } \mathbb{R}^2$$
$$(1,1,0) \qquad (0,1,1) \qquad (1,0,1) \qquad \text{em } \mathbb{R}^3$$
$$(1,2,3) \qquad (2,3,4) \qquad (3,4,5) \qquad \text{em } \mathbb{R}^3$$
$$(1,0,0,0) \qquad (1,1,0,0) \qquad (1,1,1,0) \qquad (1,1,1,1) \qquad \text{em } \mathbb{R}^4$$
$$\cos t \qquad \sin t \qquad \text{em } \mathcal{F}(\mathbb{R},\mathbb{R})$$
$$\cos^2 t \qquad \sin^2 t \qquad 1/2 \qquad \text{em } \mathcal{F}(\mathbb{R},\mathbb{R})$$
$$1 \qquad t \qquad t^2 \qquad \text{em } \mathcal{P}(\mathbb{R})$$
$$(1,-1,1,-1,1,-1,\dots) \qquad (1,1,1,1,1,1,\dots) \qquad \text{em } \mathbb{R}^\infty$$

- [Ap69] **15.9.**

9. (bases e dimensão) Seja $\mathbf{V}$ um espaço linear de dimensão finita, real ou complexo. Uma *base* de $\mathbf{V}$ é um conjunto livre de geradores de $\mathbf{V}$, ou seja, um conjunto ordenado $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ de vetores tal que cada $\mathbf{v} \in \mathbf{V}$ admite uma e uma única representação

$$\mathbf{v} = v^1\mathbf{b}_1 + v^2\mathbf{b}_2 + \cdots + v^n\mathbf{b}_n$$

Os números v^i são as *componentes*. do vetor $\mathbf{v}$ relativamente à base $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$. A *dimensão* $\dim_{\mathbb{R}} \mathbf{V}$ do espaço linear $\mathbf{V}$ (de dimensão finita) é o número de elementos de uma base. Se $(\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ é uma base ordenada do espaço linear $\mathbf{V}$, então a aplicação

$$\mathbf{v} = v^1\mathbf{b}_1 + v^2\mathbf{b}_2 + \cdots + v^n\mathbf{b}_n \mapsto (v^1, v^2, \dots, v^n)$$

define um isomorfismo $\mathbf{V} \approx \mathbb{R}^n$ ou $\mathbb{C}^n$ (dependendo se o espaço linear é real ou complexo).

- Determine uma base e a dimensão dos seguintes subespaços de $\mathbb{R}^2$ ou $\mathbb{R}^3$

$$\{(x,y) \in \mathbb{R}^2 \text{ t.q. } x = y\} \subset \mathbb{R}^2$$
$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x = y\} \subset \mathbb{R}^3$$
$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x + y + z = 0\} \subset \mathbb{R}^3$$
$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x + y = 0 \text{ e } y + z = 0\} \subset \mathbb{R}^3$$
$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x - 2y + z = 0\} \subset \mathbb{R}^3$$
$$\{(x,y,z) \in \mathbb{R}^3 \text{ t.q. } x + y + z = 0 \text{ e } z = 0\} \subset \mathbb{R}^3$$

- Verifique que
$$1 \qquad t \qquad t^2 \qquad t^3 \qquad \dots \qquad t^n$$
é uma base do espaço $\mathcal{P}_n(\mathbb{R})$ dos polinómios de grau $\le n$. Qual a dimensão de $\mathcal{P}_n(\mathbb{R})$?
- Determine as coordenadas do polinómio $f(t) = (1-t)^2$ relativamente à base ordenada $(1, t, t^2)$ de $\mathcal{P}_2(\mathbb{R})$.
- [Ap69] **15.9.**

10. (rational linear independence on the line) The real numbers x_1, x_2, $\dots$, x_n are said *linear independent over the field of rationals* if the only solution of

$$k_1x_1 + k_2x_2 + \cdots + k_nx_n = 0$$

with $k_i \in \mathbb{Z}$ is the trivial solution $k_1 = k_2 = \cdots = k_n = 0$.

- Show that (any finite subset of) the sequence
$$\log 2\,, \quad \log 3\,, \quad \log 5\,, \quad \log 7\,, \quad \dots \quad \log p\,, \quad \dots$$
of natural logarithms of prime numbers is linear independent over the rationals (use the unique factorization of an integer into prime factors) [21].

[21] H. Bohr, 1910.

14 Formas lineares

ref: [La87] *Ch. III*

1. (linearidade) Se cada kilo de ♡ custa A euros e cada kilo de ♠ custa B euros, então a kilos de ♡ e b kilos de ♠ custam $aA + bB$ euros. Ou seja, a função "preço" P satisfaz

   $$P(a\heartsuit + b\spadesuit) = a \cdot P(\heartsuit) + b \cdot P(\spadesuit)$$

   Esta propriedade é chamada *linearidade*.

   Por outro lado, a superfície e o volume de um cubo de lado 2ℓ são 4 e 8 vezes a superfície e o volume de um cubo de lado ℓ, respetivamente (e esta é uma das razões pela existência de dimensões típicas de animais e plantas, como explicado por D'Arcy Thompson[22]). São funções não lineares.

   - Dê exemplos de funções lineares.
   - Dê exemplos de funções não lineares.

2. (formas lineares, espaço dual) Seja $\mathbf{V}$ um espaço linear real (ou complexo). Uma função real $f : \mathbf{V} \to \mathbb{R}$ (ou complexa $f : \mathbf{V} \to \mathbb{C}$) é dita *aditiva* se $\forall\, \mathbf{v}, \mathbf{w} \in \mathbf{V}$

   $$\boxed{f(\mathbf{v} + \mathbf{w}) = f(\mathbf{v}) + f(\mathbf{w})}$$

   e é dita *homogénea* se $\forall\, \mathbf{v} \in \mathbf{V}$ e $\forall\, \lambda \in \mathbb{R}$ (ou $\forall\, \lambda \in \mathbb{C}$)

   $$\boxed{f(\lambda\mathbf{v}) = \lambda f(\mathbf{v})}$$

   Uma função real $\boldsymbol{\xi} : \mathbf{V} \to \mathbb{R}$ (ou complexa $\boldsymbol{\xi} : \mathbf{V} \to \mathbb{C}$) aditiva e homogénea, ou seja, tal que

   $$\boxed{\boldsymbol{\xi}(\lambda\mathbf{v} + \mu\mathbf{w}) = \lambda\, \boldsymbol{\xi}(\mathbf{v}) + \mu\, \boldsymbol{\xi}(\mathbf{w})}$$

   é dita *forma linear*, ou *covetor* (ou *funcional linear* quando $\mathbf{V}$ é um espaço de funções). Uma notação simétrica para o valor da forma linear $\boldsymbol{\xi}$ sobre o vetor $\mathbf{v}$ é

   $$\langle \boldsymbol{\xi}, \mathbf{v} \rangle := \boldsymbol{\xi}(\mathbf{v})\,.$$

   O espaço $\mathbf{V}^* := \mathrm{Hom}_{\mathbb{R}}(\mathbf{V}, \mathbb{R})$ (ou $\mathrm{Hom}_{\mathbb{C}}(\mathbf{V}, \mathbb{C})$) das formas lineares, dito *espaço dual (algébrico)* de $\mathbf{V}$, é um espaço linear real (ou complexo) se a adição e o produto por um escalar são definidos por

   $$\langle \boldsymbol{\xi} + \boldsymbol{\eta}, \mathbf{v} \rangle := \langle \boldsymbol{\xi}, \mathbf{v} \rangle + \langle \boldsymbol{\eta}, \mathbf{v} \rangle \qquad \text{e} \qquad \langle \lambda\boldsymbol{\xi}, \mathbf{v} \rangle := \langle \boldsymbol{\xi}, \lambda\mathbf{v} \rangle$$

   respetivamente, e a forma nula $\mathbf{0}^* \in \mathbf{V}^*$ é definida por $\langle \mathbf{0}^*, \mathbf{v} \rangle = 0\ \forall\, \mathbf{v} \in \mathbf{V}$.

   Se o espaço vetorial $\mathbf{V}$ tem dimensão finita, e se $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ é uma base de $\mathbf{V}$ (por exemplo a base canónica de $\mathbf{V} \approx \mathbb{R}^n$), então cada forma linear $\boldsymbol{\xi} \in \mathbf{V}^*$ é determinada pelos seus valores $\xi_i := \langle \boldsymbol{\xi}, \mathbf{e}_i \rangle$ nos vetores da base, pois

   $$\langle \boldsymbol{\xi}, \mathbf{v} \rangle = \left\langle \boldsymbol{\xi}, v^1\mathbf{e}_1 + v^2\mathbf{e}_2 + \cdots + v^n\mathbf{e}_n \right\rangle = \xi_1 v^1 + \xi_2 v^2 + \cdots + \xi_n v^n\,.$$

   Portanto, também $\mathbf{V}^*$ tem dimensão finita, e uma base de $\mathbf{V}^*$, dita *base dual*, é o conjunto ordenado dos covetores $\mathbf{e}^1, \mathbf{e}^2, \dots, \mathbf{e}^n$ definidos por [23]

   $$\left\langle \mathbf{e}^i, \mathbf{e}_j \right\rangle = \delta^i{}_j$$

[22] D'Arcy Wentworth Thompson, *On growth and form*, 1917 and 1942.

[23] O *símbolo de Kronecker* é definido por

$$\delta^i{}_j = \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j \end{cases}$$

As coordenadas da forma linear $\boldsymbol{\xi} = \xi_1 \mathbf{e}^1 + \xi_2 \mathbf{e}^2 + \dots + \xi_n \mathbf{e}^n$ relativamente à base dual são os números $\xi_i = \langle \boldsymbol{\xi}, \mathbf{e}_i \rangle$, assim que $\langle \boldsymbol{\xi}, \mathbf{v} \rangle = \sum_i \xi_i v^i$.

Para cada $\mathbf{v} \in \mathbf{V}$, a função $\xi \mapsto \langle \boldsymbol{\xi}, \mathbf{v} \rangle$ é uma forma linear em $\mathbf{V}^*$, e portanto existe um homomorfismo injetivo de $\mathbf{V}$ em $(\mathbf{V}^*)^*$. Se $\mathbf{V}$ tem dimensão finita, então todas as formas lineares $g \in (\mathbf{V}^*)^*$ podem ser representadas como $g(\boldsymbol{\xi}) = \langle \boldsymbol{\xi}, \mathbf{v} \rangle$ para algum $\mathbf{v} \in \mathbf{V}$ (basta definir $\mathbf{v} = (v^1, v^2, \dots, v^n)$ com $v^i = g(\mathbf{e}^i)$), e portanto o espaço dual do espaço dual é isomorfo a $(\mathbf{V}^*)^* \approx \mathbf{V}$ (mas o isomorfismo não é canónico, depende da escolha de uma base!).

- Mostre que uma função homogénea $f : \mathbb{R} \to \mathbb{R}$ é $f(x) = \lambda x$, com $\lambda = f(1) \in \mathbb{R}$. Deduça que uma função homogénea $f : \mathbb{R} \to \mathbb{R}$ é também aditiva, e portanto linear.
- Diga se as seguintes funções $f : \mathbb{R}^n \to \mathbb{R}$ são ou não lineares.

$$x \mapsto 3x \qquad x \mapsto 2x - 1 \qquad x \mapsto \sin(2\pi x)$$

$$(x, y) \mapsto 3x - 5y \qquad (x, y) \mapsto x^2 - xy$$

$$(x, y, z) \mapsto 2x - y + 3z \qquad (x, y, z) \mapsto 2x - y + 3z + 8$$

$$(x, y, z) \mapsto 0 \qquad (x, y, z) \mapsto \sqrt{3}$$

$$(x_1, x_2, \dots, x_n) \mapsto 0 \qquad (x_1, x_2, \dots, x_n) \mapsto 66$$

- Mostre que, dados $a_1, a_2, \dots, a_n \in \mathbb{R}$, a aplicação

$$(x_1, x_2, \dots, x_n) \mapsto a_1 x_1 + a_2 x_2 + \dots + a_n x_n$$

  é uma forma linear em $\mathbb{R}^n$.

3. (wild additive functions in the real line) For real valued functions of a real variable $f : \mathbb{R} \to \mathbb{R}$, homogeneity implies additivity, since $x + y = (1 + y/x)\, x$ (if $x \neq 0$, of course), and therefore an homogeneous function satisfies

$$f(x + y) = f((1 + y/x)\, x) = (1 + y/x) f(x) = f(x) + (y/x) f(x) = f(x) + f(y) \,.$$

Surprisingly, there exist additive functions which are not homogeneous (hence not linear), at least if we accept the axiom of choice. Indeed, additivity only implies linearity on "rational lines" $\mathbb{Q}x \subset \mathbb{R}$, i.e.

$$f(rx) = r f(x) \qquad \forall\, r = p/q \in \mathbb{Q} \text{ and } \forall\, x \in \mathbb{R} \,.$$

Therefore, if we could choose a different slope $\lambda_{\mathbb{Q}x}$, hence a different homogeneous function $rx \mapsto \lambda_{\mathbb{Q}x} rx$, for any orbit $\mathbb{Q}x$ of the quotient space $\mathbb{Q} \backslash \mathbb{R}$, the resulting function $f : \mathbb{R} \to \mathbb{R}$ would be additive but not homogeneous. Any such wild additive but not homogeneous function $f : \mathbb{R} \to \mathbb{R}$ cannot be continuous, and indeed has a dense graph.

- Let $f : \mathbb{R} \to \mathbb{R}$ be an homogeneous function, and $x \in \mathbb{R}$. For $r = p/q \in \mathbb{Q}$ with $p, q \in \mathbb{Z}$ and $q \neq 0$, show that $f(rx) = f(px/q) = pf(x/q)$ and also $f(x) = f(qx/q) = qf(x/q)$. Deduce that $f(rx) = r f(x)$ for all $r \in \mathbb{Q}$.
- Let $f : \mathbb{R} \to \mathbb{R}$ be an additive but not homogeneous function, so that there exists two points $a, b \in \mathbb{R}$ where $f(a)/a \neq f(b)/b$. This implies that $\mathbf{v} = (a, f(a))$ and $\mathbf{w} = (b, f(b))$ are linearly independent vectors, hence a basis, of $\mathbb{R}^2$. From $\mathbb{R}\mathbf{v} + \mathbb{R}\mathbf{w} = \mathbb{R}^2$, deduce that the "rational plane" $\mathbb{Q}\mathbf{v} + \mathbb{Q}\mathbf{w}$ is dense in $\mathbb{R}^2$, i.e. ant point $\mathbf{r} = (x, y) \in \mathbb{R}^2$ may be approximated with arbitrary precision by a point in $\mathbb{Q}\mathbf{v} + \mathbb{Q}\mathbf{w}$, i.e. for any $\mathbf{r} = (x, y) \in \mathbb{R}^2$ and any $\varepsilon > 0$ there exist rationals $\lambda, \lambda' \in \mathbb{Q}$ such that $\|\mathbf{r} - (\lambda \mathbf{v} + \lambda' \mathbf{w})\| < \varepsilon$. Use again the additivity of f to show that this implies the existence of a point $c \in \mathbb{R}$ such that $\|\mathbf{r} - (c, f(c)\| < \varepsilon$.

4. (formas lineares no espaço euclidiano $\mathbb{R}^n$) Uma forma linear $\boldsymbol{\xi} \in (\mathbb{R}^n)^*$ é determinada pelos seus valores nos vetores de uma base de $\mathbb{R}^n$. Por exemplo, se $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ é a base canónica de $\mathbb{R}^n$, e $\xi_i := \langle \boldsymbol{\xi}, \mathbf{e}_i \rangle$, com $i = 1, 2, \ldots, n$, então o valor da forma $\boldsymbol{\xi} \in (\mathbb{R}^n)^*$ sobre o vetor genérico $\mathbf{x} = x^1\mathbf{e}_1 + x^2\mathbf{e}_2 + \cdots + x^n\mathbf{e}_n \in \mathbb{R}^n$ é

$$\begin{aligned} \langle \boldsymbol{\xi}, \mathbf{x} \rangle &= \left\langle \boldsymbol{\xi}, x^1\mathbf{e}_1 + x^2\mathbf{e}_2 + \cdots + x^n\mathbf{e}_n \right\rangle \\ &= x^1 \langle \boldsymbol{\xi}, \mathbf{e}_1 \rangle + x^2 \langle \boldsymbol{\xi}, \mathbf{e}_2 \rangle + \cdots + x^n \langle \boldsymbol{\xi}, \mathbf{e}_n \rangle \\ &= x^1\xi_1 + x^2\xi_2 + \cdots + x^n\xi_n \,. \end{aligned}$$

Portanto, se $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_n) \in \mathbb{R}^n$, então

$$\langle \boldsymbol{\xi}, \mathbf{x} \rangle = \boldsymbol{\xi} \cdot \mathbf{x}$$

A correspondência que associa a forma $\langle \boldsymbol{\xi}, \mathbf{x} \rangle = \boldsymbol{\xi} \cdot \mathbf{x}$ ao vetor $\boldsymbol{\xi} \in \mathbb{R}^n$ é um isomorfismo $\mathbb{R}^n \approx (\mathbb{R}^n)^*$ entre o espaço euclidiano $\mathbb{R}^n$ e o seu dual $(\mathbb{R}^n)^*$ (que depende da estrutura euclidiana, ou seja, do produto escalar euclidiano!). Em coordenadas, a correspondência é $\xi_i = \sum_j \delta_{ij}\xi^j$.

- Determine o vetor $\boldsymbol{\xi} = (\xi_1, \xi_2, \ldots, \xi_n) \in \mathbb{R}^n$ que define as seguintes formas lineares

$$x \mapsto 3x \qquad (x, y) \mapsto 0$$

$$(x, y) \mapsto 5x + 9y \qquad (x, y, z) \mapsto -3x + 7y - z$$

$$(x_1, x_2, \ldots, x_n) \mapsto 3x_k \qquad \text{com } 0 \leq k \leq n$$

- $f : \mathbb{R}^2 \to \mathbb{R}$ é uma função linear tal que $f(\mathbf{i}) = 5$ e $f(\mathbf{j}) = -2$. Determine $f(x, y)$.
- $f : \mathbb{R}^3 \to \mathbb{R}$ é uma função linear tal que $f(\mathbf{i}) = -3$, $f(\mathbf{j}) = 1$ e $f(\mathbf{k}) = 7$. Determine $f(x, y, z)$.

5. (núcleo e hiperplanos) O *núcleo/espaço nulo* (em inglês *kernel*) da forma linear $\boldsymbol{\xi} \in \mathbf{V}^*$ é o subespaço vetorial

$$\boxed{\ker(\boldsymbol{\xi}) := \{\mathbf{x} \in \mathbf{V} \text{ t.q. } \langle \boldsymbol{\xi}, \mathbf{x} \rangle = 0\} \subset \mathbf{V} \,.}$$

Se $\boldsymbol{\xi} \neq \mathbf{0}'$ e $\mathbf{v} \in \mathbf{V} \backslash \ker(\boldsymbol{\xi})$ (i.e. um vetor tal que $\langle \boldsymbol{\xi}, \mathbf{v} \rangle \neq 0$), então cada vetor $\mathbf{x} \in \mathbf{V}$ pode ser representado de uma única maneira como

$$\mathbf{x} = \lambda\mathbf{v} + \mathbf{w}$$

com $\lambda \in \mathbb{R}$ e $\mathbf{w} \in \ker(\xi)$. De fato, a condição $\mathbf{x} - \lambda\mathbf{v} \in \ker(\xi)$, ou seja, $\langle \xi, \mathbf{x} - \lambda\mathbf{v} \rangle = 0$, obriga a escolher $\lambda = \langle \boldsymbol{\xi}, \mathbf{x} \rangle / \langle \boldsymbol{\xi}, \mathbf{v} \rangle$. Portanto, o núcleo $\ker(\boldsymbol{\xi})$ de uma forma $\boldsymbol{\xi} \neq \mathbf{0}'$ é um *hiperplano* do espaço linear $\mathbf{V}$, ou seja, um subespaço linear di "co-dimensão" 1. Em particular, se $\mathbf{V}$ tem dimensão finita,

$$\boxed{\dim_{\mathbb{R}} \ker(\boldsymbol{\xi}) + 1 = \dim_{\mathbb{R}} \mathbf{V}}$$

O *hiperplano afim* que passa pelo ponto $\mathbf{a} \in \mathbf{V}$ e é paralelo ao hiperplano $\ker(\boldsymbol{\xi})$ é

$$\mathbf{a} + \ker(\boldsymbol{\xi}) = \{\mathbf{v} \in \mathbf{V} \text{ t.q. } \langle \boldsymbol{\xi}, \mathbf{v} \rangle = \lambda\} \qquad \text{onde } \lambda = \langle \boldsymbol{\xi}, \mathbf{a} \rangle \,.$$

- Mostre que o núcleo de uma forma linear $\boldsymbol{\xi} \in \mathbf{V}^*$ é um subespaço linear de $\mathbf{V}$.
- Uma "equação linear"
$$a_1x^1 + a_2x^2 + \cdots + a_nx^n = b$$
define um hiperplano afim em $\mathbb{R}^n$. Mostre que o hiperplano afim é um subespaço linear de $\mathbb{R}^n$ sse $b = 0$.

6. (interseções de hiperplanos/sistemas homogéneos) Se $\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \dots, \boldsymbol{\xi}^m \in \mathbf{V}^*$ são m formas lineares independentes num espaço linear de dimensão finita $\mathbf{V} \approx \mathbb{R}^n$ (em particular, $m \leq n$), então a interseção dos núcleos $\mathbf{W} = \cap_{k=1}^m \ker(\boldsymbol{\xi}^k)$, ou seja o conjunto dos vetores $\mathbf{x} \in \mathbf{V}$ que resolvem o "sistema homogéneo"

$$\langle \boldsymbol{\xi}^1, \mathbf{x} \rangle = 0\,, \quad \langle \boldsymbol{\xi}^2, \mathbf{x} \rangle = 0\,, \qquad \dots \qquad \langle \boldsymbol{\xi}^m, \mathbf{x} \rangle = 0\,,$$

é um subespaço vetorial de co-dimensão m, e portanto de dimensão $n-m$. De fato, se completamos o sistema de formas independentes até obter uma base $\boldsymbol{\xi}^1, \boldsymbol{\xi}^2, \dots, \boldsymbol{\xi}^m, \boldsymbol{\xi}^{m+1}, \dots, \boldsymbol{\xi}^n$ de $\mathbf{V}^*$, e se $\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n$ denota a base dual de $\mathbf{V}$ (assim que $\langle \boldsymbol{\xi}^i, \mathbf{b}_j \rangle = \delta^i{}_j$), então nas coordenadas relativas a esta base o sistema homogéneo é

$$x^1 = 0\,, \quad x^2 = 0\,, \qquad \dots \qquad x^m = 0\,,$$

e as soluções são todos os vetores $\mathbf{x} = \sum_{k=m+1}^n x^k \mathbf{b}_k \in \mathbf{W} \approx \mathbb{R}^{n-m}$.

7. (integral) O integral

$$f \mapsto I(f) := \int_a^b f(t)\, dt$$

é uma forma linear no espaço $\mathcal{C}^0([a,b])$ das funções contínuas no intervalo $[a,b] \subset \mathbb{R}$. O seu núcleo é o hiperplano das funções com média nula no intervalo.

   - Toda função contínua no intervalo pode ser representada de forma única como soma $f(x) = c + g(x)$ onde $c = I(f)$ é uma constante (a média de f vezes o comprimento do intervalo) e $g(x) = f(x) - I(f)$ é uma função com média nula.

8. (delta de Dirac) A *delta de Dirac* (no ponto 0), definida por

$$f \mapsto \delta(f) := \text{``} \int_{-\infty}^{\infty} \delta(t) f(t)\, dt \text{''} = f(0)\,,$$

é uma forma linear no espaço $\mathcal{C}^0(\mathbb{R})$ das funções contínuas $f : \mathbb{R} \to \mathbb{R}$. O núcleo de δ é o conjunto das funções (contínuas) tais que $f(0) = 0$, que é um hiperplano do espaço $\mathcal{C}^0(\mathbb{R})$.

9. (plane waves and Pontryagin dual) A *plane wave* is a complex valued function $e_\xi : \mathbb{R}^n \to \mathbb{C}$ defined by

$$e_\xi(\mathbf{x}) := e^{2\pi i \boldsymbol{\xi} \cdot \mathbf{x}}$$

for some "wave vector" $\boldsymbol{\xi} \in (\mathbb{R}^n)^* \approx \mathbb{R}^n$ (an alternative definition omits the factor 2π). For example, a plane wave

$$e^{2\pi i (\omega t - \mathbf{p} \cdot \mathbf{r})}$$

in the space-time $\mathbb{R} \times \mathbb{R}^3$ with coordinates $(t, \mathbf{r})$, describes a (transversal) wave traveling in the direction of the vector $\mathbf{p} \in \mathbb{R}^3$ with frequency ω (observe that vectors and covectors have different dimensions, e.g. if t is time and $\mathbf{r}$ is a length, then ω is a frequency while $\mathbf{p}$ is the inverse of a length, hence the distinction between a linear space and its dual is real!).

   Any plane wave is a continuous (actually infinitely differentiable) homeomorphism from the abelian additive group $\mathbb{R}^n$ into the abelian multiplicative group $\mathbb{S} = \{z \in \mathbb{C} \text{ s.t. } |z| = 1\}$ of unit complex numbers, i.e.

$$e_\xi(\mathbf{x} + \mathbf{y}) = e_\xi(\mathbf{x})\, e_\xi(\mathbf{y})$$

   The set $\widehat{\mathbb{R}^n}$ of all plane waves, or, technically, the set of all continuous homomorphisms $e_\xi : \mathbb{R}^n \to \mathbb{S}$, equipped with the group law

$$(e_\xi \cdot e_{\xi'})(\mathbf{x}) = e_\xi(\mathbf{x})\, e_{\xi'}(\mathbf{x})\,,$$

   is called *Pontryagin/topological dual* of the abelian (topological) group $\mathbb{R}^n$, and, as an abelian group, it is isomorphic to $\widehat{\mathbb{R}^n} \approx (\mathbb{R}^n)^*$, the isomorphism being $e_\xi \leftrightarrow \boldsymbol{\xi}$.

10. (lattices and reciprocal lattices) A *(Bravais) lattice* is an additive subgroup

$$\begin{aligned}\Lambda &= \mathbb{Z}\mathbf{v}_1 + \mathbb{Z}\mathbf{v}_2 + \cdots + \mathbb{Z}\mathbf{v}_n \\ &:= \{n_1\mathbf{v}_1 + n_2\mathbf{v}_2 + \ldots n_n\mathbf{v}_n \ \text{ with } n_1, n_2, \ldots, n_n \in \mathbb{Z}\} \subset \mathbb{R}^n\end{aligned}$$

of the additive group $\mathbb{R}^n$, generated by n linearly independent vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n \in \mathbb{R}^n$, called *primitive vectors*. A plane wave $e_{\boldsymbol{\xi}}(\mathbf{x}) := e^{2\pi i \boldsymbol{\xi}\cdot\mathbf{x}}$ in $\mathbb{R}^n$ is Λ-periodic, i.e. satisfies $e_{\boldsymbol{\xi}}(\mathbf{x}+\mathbf{v}) = e_{\boldsymbol{\xi}}(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^n$ and all $\mathbf{v} \in \Lambda$, provided the wave vector $\boldsymbol{\xi}$ belongs to the *reciprocal lattice*

$$\Lambda_* := \{\boldsymbol{\xi} \in (\mathbb{R}^n)^* \text{ s.t. } \boldsymbol{\xi}\cdot\mathbf{v} \in \mathbb{Z} \quad \forall \mathbf{v} \in \Lambda\} \subset (\mathbb{R}^n)^* ,$$

isomorphic to the *dual subgroup*

$$\Lambda^\perp := \{e_{\xi} \in \widehat{\mathbb{R}^n} \text{ s.t. } e_{\xi}(\mathbf{v}) = 1 \quad \forall \mathbf{v} \in \Lambda\} \subset \widehat{\mathbb{R}^n} .$$

For example, the reciprocal lattice of the one-dimensional lattice $\Lambda = \lambda\mathbb{Z} \subset \mathbb{R}$, with $\lambda > 0$, is $\Lambda_* = \lambda^{-1}\mathbb{Z} \subset \mathbb{R} \approx \mathbb{R}^*$. Physicists are mainly interested in lattices $\Lambda = \mathbb{Z}\mathbf{v}_1 + \mathbb{Z}\mathbf{v}_2 + \mathbb{Z}\mathbf{v}_3 \subset \mathbb{R}^3$ generated by three independent vectors $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in \mathbb{R}^3$, since they describe (the positions of atoms in) crystals. The volume of a *fundamental domain* (or *primitive unit cell*) for such a lattice $\Lambda \subset \mathbb{R}^3$ (a domain $F \subset \mathbb{R}^3$ such that $\cup_{\mathbf{v}\in\Lambda} T_{\mathbf{v}}(F) = \mathbb{R}^3$, and such that the different images $T_{\mathbf{v}}(F)$ and $T_{\mathbf{v}'}(F)$ for $\mathbf{v} \neq \mathbf{v}'$ in Λ have disjoint interiors), or, equivalently, the volume of the quotient space $\mathbb{R}^3/\Lambda$, is

$$\mathrm{Vol}(\mathbb{R}^3/\Lambda) = |\mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3)| .$$

The reciprocal lattice is the lattice $\Lambda_* = \mathbb{Z}\boldsymbol{\xi}^1 + \mathbb{Z}\boldsymbol{\xi}^2 + \mathbb{Z}\boldsymbol{\xi}^3 \subset (\mathbb{R}^3)^*$ generated by the co-vectors

$$\boldsymbol{\xi}^1 = \frac{\mathbf{v}_2 \times \mathbf{v}_3}{\mathbf{v}_1 \cdot (\mathbf{v}_2 \times \mathbf{v}_3)} \qquad \boldsymbol{\xi}^2 = \frac{\mathbf{v}_3 \times \mathbf{v}_1}{\mathbf{v}_2 \cdot (\mathbf{v}_3 \times \mathbf{v}_1)} \qquad \boldsymbol{\xi}3 = \frac{\mathbf{v}_1 \times \mathbf{v}_2}{\mathbf{v}_3 \cdot (\mathbf{v}_1 \times \mathbf{v}_2)}$$

(under the identification $(\mathbb{R}^3)^* \approx \mathbb{R}^3$ induced by the Euclidian scalar product).

11. (conjuntos convexos) Um subconjunto $C \subset \mathbf{V}$ de um espaço vetorial real é *convexo* se contém o segmento entre cada par de seus pontos, i.e. se

$$\mathbf{x}, \mathbf{y} \in C \qquad \Rightarrow \qquad (1-t)\mathbf{x} + t\mathbf{y} \in C \quad \forall t \in [0,1]$$

O menor convexo que contém (ou seja, a interseção de todos os convexos que contêm) um subconjunto $A \subset \mathbb{R}^n$ é chamado *envoltória/invólucro/fecho convexa/o* de A, e denotado por $\mathrm{Conv}(A)$. Em particular, o menor convexo que contém o conjunto finito de pontos $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \mathbf{V}$ é

$$\mathrm{Conv}(\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m\}) := \{t_1\mathbf{x}_1 + t_2\mathbf{x}_2 + \cdots + t_m\mathbf{x}_m \quad \text{com } t_i \geq 0 \text{ e } t_1 + t_2 + \cdots + t_m = 1\}$$

- As bolas $B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^n \text{ t.q. } \|\mathbf{y}-\mathbf{x}\| < r\}$ e $\overline{B_r(\mathbf{x})} = \{\mathbf{y} \in \mathbb{R}^n \text{ t.q. } \|\mathbf{y}-\mathbf{x}\| \leq r\}$ são convexas.
- Dados um vetor $\mathbf{n} \in \mathbb{R}^n$ e um escalar $b \in \mathbb{R}$, os semi-espaços $\{\mathbf{x} \in \mathbb{R}^n \text{ t.q. } \mathbf{n}\cdot\mathbf{x} \geq b\}$ são convexos.
- Dada uma forma $\boldsymbol{\xi} \in \mathbf{V}^*$ e um escalar $b \in \mathbb{R}$, os semiespaços $\{\mathbf{v} \in \mathbf{V} \text{ t.q. } \langle\boldsymbol{\xi}, \mathbf{v}\rangle \geq b\}$ são convexos.
- A envoltória convexa de três pontos, A, B e C do plano $\mathbb{R}^2$ é o triângulo de vértices A, B e C.
- Translações e homotetias preservam os convexos, i.e. se $C \subset \mathbf{V}$ é convexo, então também $T_{\mathbf{a}}(C) = C + \mathbf{a}$ e $H_\lambda(C) = \lambda C$ são convexos, $\forall \mathbf{a} \in \mathbf{V}$ e $\forall \lambda \in \mathbb{R}$.
- Se $\boldsymbol{\xi} \in \mathbf{V}^*$, então $\{\mathbf{v} \in \mathbf{V} \text{ t.q. } \boldsymbol{\xi}(\mathbf{v}) < 0\}$ é convexo.

12. (medidas de probabilidades) Uma *(medida de) probabilidade* num "espaço dos acontecimentos" finito $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ é uma função $\mathbb{P} : 2^\Omega := \{\text{subconjuntos } A \subset \Omega\} \to [0,1]$ aditiva, i.e. tal que
$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) \qquad \text{se } A \cap B = \emptyset\,,$$
(a probabilidade do evento "A ou B" é igual à probabilidade do evento A mais a probabilidade do evento B se A e B são eventos mutuamente exclusivos) que verifica $\mathbb{P}(\emptyset) = 0$ (a probabilidade do "evento impossível" é nula) e $\mathbb{P}(\Omega) = 1$ (a probabilidade do "evento certo" é um).

    Cada vetor $\mathbf{p} = (p_1, p_2, \ldots, p_n) \in \mathbb{R}^n$ cujas coordenadas estão limitadas por $0 \le p_i \le 1$ e tal que $p_1 + p_2 + \cdots + p_n = 1$ define uma probabilidade $\mathbb{P}$, por meio de
$$\mathbb{P}(A) = \sum_{\omega_i \in A} p_i$$
ou seja, $p_i = \mathbb{P}(\{\omega_i\})$. Portanto, o espaço das medidas de probabilidades em Ω é o fecho convexo dos vetores da base canónica de $\mathbb{R}^n$, chamado *"simplex"*
$$\Delta^{n-1} := \{\mathbf{p} = (p_1, p_2, \ldots, p_n) \in \mathbb{R}^n \text{ com } 0 \le p_i \le 1 \text{ e } p_1 + p_2 + \cdots + p_n = 1\} \subset \mathbb{R}^n\,.$$

15 Espaços euclidianos

ref: [Ap69] *Vol. 2*, **1.11-17** ; [La87] *Ch. V*

1. (espaços euclidianos) Um *espaço euclidiano* é um espaço vetorial $\mathbf{E}$, real ou complexo, munido de um *produto interno* (também dito *hermítico* se o espaço é complexo), uma aplicação que associa a cada par de vetores $\mathbf{x}, \mathbf{y} \in \mathbf{E}$ um escalar $\langle \mathbf{x}, \mathbf{y} \rangle \in \mathbb{R}$ ou $\mathbb{C}$, satisfazendo os axiomas

   E1 *(simetria hermítica)* $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$, $\qquad \forall \mathbf{x}, \mathbf{y} \in \mathbf{E}$

   E2 *(linearidade)* $\langle \lambda\mathbf{x} + \mu\mathbf{z}, \mathbf{y} \rangle = \lambda \langle \mathbf{x}, \mathbf{y} \rangle + \mu \langle \mathbf{z}, \mathbf{y} \rangle \qquad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{E}$ e $\forall \lambda, \mu \in \mathbb{R}$ ou $\mathbb{C}$.

   E3 *(positividade)* $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ se $\mathbf{x} \neq \mathbf{0}$.

   Se o espaço é real o axioma E1 diz simplesmente que $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$. O axioma E2 diz que o produto interno é linear na primeira variável, e o axioma E1 então implica que é "anti-linear" na segunda variável, ou seja, que $\langle \mathbf{x}, \lambda\mathbf{y} \rangle = \overline{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle$ (mas há textos, em particular de física matemática, onde acontece o contrário).

   Os arquétipos de espaços euclidianos de dimensão finita são: o espaço euclidiano real $\mathbb{R}^n$, munido do produto interno "usual" $\langle \mathbf{x}, \mathbf{y} \rangle = x_1y_1 + \cdots + x_ny_n$; o espaço euclidiano complexo $\mathbb{C}^n$, munido do produto interno "usual" $\langle \mathbf{x}, \mathbf{y} \rangle = x_1\overline{y_1} + \cdots + x_n\overline{y_n}$.

   A *norma* do vetor $\mathbf{x} \in \mathbf{E}$ é o número real não negativo definido por

   $$\boxed{\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}}$$

   Pelas E2 e E3, o único vetor com norma 0 é o vetor nulo $\mathbf{0}$. Também é imediato observar que, pelas propriedades E1 e E2, $\|\lambda x\| = |\lambda| \, \|\mathbf{x}\|$, ou seja, a norma é "homogénea".

   Um vetor é dito *unitário* se a sua norma é um. Todo vetor não nulo $\mathbf{x}$ é proporcional a um vetor unitário, por exemplo $\mathbf{u} = \mathbf{x}/\|\mathbf{x}\|$ (se o espaço é real, apenas podemos multiplicar $\mathbf{u}$ por ± 1, em quanto se o espaço é complexo temos ainda a liberdade de multiplicar $\mathbf{u}$ por uma fase arbitrária $e^{i\theta}$, com $\theta \in \mathbb{R}$)

   Os vetores $\mathbf{x}$ e $\mathbf{y}$ de $\mathbf{E}$ são ditos *ortogonais/perpendiculares* quando $\langle \mathbf{x}, \mathbf{y} \rangle = 0$, e uma notação é $\mathbf{x} \perp \mathbf{y}$. Uma consequência imediata é o *teorema de Pitágoras*: se $\mathbf{x}$ e $\mathbf{y}$ são ortogonais então

   $$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 .$$

   Seja $\mathbf{v} \in \mathbf{E}$ um vetor não nulo. Cada vetor $\mathbf{x} \in \mathbf{E}$ pode ser representado de maneira única como soma

   $$\mathbf{x} = \lambda\mathbf{v} + \mathbf{w}$$

   de um vetor $\lambda\mathbf{v}$ proporcional a $\mathbf{v}$ e um vetor $\mathbf{w}$ ortogonal a $\mathbf{v}$. De fato, a condição de ortogonalidade $\langle \mathbf{x} - \lambda\mathbf{v}, \mathbf{v} \rangle = 0$ obriga a escolher

   $$\lambda = \frac{\langle \mathbf{x}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2} .$$

   O vetor $\lambda\mathbf{v}$ é dito *projeção (ortogonal)* do vetor $\mathbf{x}$ sobre (a reta definida pel)o vetor $\mathbf{v}$, e o coeficiente λ é dito *componente* de $\mathbf{x}$ ao longo de $\mathbf{v}$. Em particular, a componente de $\mathbf{x}$ ao longo de um vetor unitário $\mathbf{u}$ é o produto escalar $\langle \mathbf{x}, \mathbf{u} \rangle$.

   A positividade da norma da diferença $\mathbf{x} - \lambda\mathbf{v}$ entre o vetor $\mathbf{x}$ e a sua projeção $\lambda\mathbf{v}$ sobre $\mathbf{y} \neq 0$ implica a fundamental *desigualdade de Schwarz*.

   Teorema 15.1 (desigualdade de Schwarz)**.** *O módulo do produto escalar entre dois vetores* $\mathbf{x}, \mathbf{y} \in \mathbf{E}$ *é limitado por*

   $$\boxed{|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \|\mathbf{y}\|}$$

   e a igualdade verifica-se sse os vetores $\mathbf{x}$ *e* $\mathbf{y}$ *são linearmente dependentes (ou seja, se um é proporcional ao outro).*

Se o espaço euclidiano é real, é então possível definir o *ângulo* θ entre dois vetores não nulos $\mathbf{x}$ e $\mathbf{y}$ pela identidade

$$\cos\theta := \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\|\,\|\mathbf{y}\|} ,$$

sendo o último um número real entre -1 e 1.

Consequência importante da desigualdade de Schwarz é que a norma satisfaz a *desigualdade do triângulo*

$$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| .$$

A *distância* entre os vetores/pontos $\mathbf{x}$ e $\mathbf{y}$ de $\mathbf{E}$ é definida por

$$\boxed{\mathrm{d}(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|}$$

que satisfaz a desigualdade do triângulo

$$\mathrm{d}(\mathbf{x}, \mathbf{y}) \leq \mathrm{d}(\mathbf{x}, \mathbf{z}) + \mathrm{d}(\mathbf{z}, \mathbf{y}) .$$

O axioma E2 implica que, fixado $\mathbf{y} \in \mathbf{E}$, a função $\boldsymbol{\xi}_{\mathbf{y}} : \mathbf{x} \mapsto \langle \mathbf{x}, \mathbf{y} \rangle$ é uma forma linear, e pela desigualdade de Schwarz 15.1 a sua norma é $\|\mathbf{y}\|$. Em particular, a correspondência $\mathbf{y} \mapsto \boldsymbol{\xi}_{\mathbf{y}} := \langle \cdot, \mathbf{y} \rangle$ define uma inclusão $\mathbf{E} \subset \mathbf{E}^*$, que é anti-linear (ou seja, $\xi_{\lambda \mathbf{y}} = \overline{\lambda} \xi_{\mathbf{y}}$). Se o espaço euclidiano tem dimensão finita, esta inclusão é claramente uma bijeção.

- Verifique que $\langle \mathbf{x}, \lambda \mathbf{y} \rangle = \overline{\lambda} \langle \mathbf{x}, \mathbf{y} \rangle$. Em particular, se o espaço é real, o produto escalar é linear também na segunda variável.
- Mostre que $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ para todos os $\mathbf{y} \in \mathbf{E}$ sse $\mathbf{x} = 0$.
- Prove a desigualdade de Schwarz (aplique o axioma E3, ou seja, a positividade, ao vetor $\mathbf{x} - \lambda \mathbf{y}$ com $\lambda = \langle \mathbf{x}, \mathbf{y} \rangle / \|\mathbf{y}\|^2$ se $\mathbf{y} \neq 0$).
- Verifique que $\mathbf{x}, \mathbf{y} \mapsto \sum_i^n x_i y_i$ é um produto escalar em $\mathbb{R}^n$.
- Verifique que $\mathbf{z}, \mathbf{z'} \mapsto \sum_i^n z_i \overline{z'_i}$ é um produto escalar hermiteano em $\mathbb{C}^n$.
- Mostre que num espaço euclidiano real

  $$\|\mathbf{x} + \mathbf{y}\|^2 - \|\mathbf{x} - \mathbf{y}\|^2 = 4 \langle \mathbf{x}, \mathbf{y} \rangle$$

  e deduza que $\|\mathbf{x} + \mathbf{y}\| = \|\mathbf{x} - \mathbf{y}\|$ sse $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.
- Prove o *teorema de Pitágoras*: se $\mathbf{x}$ e $\mathbf{y}$ são ortogonais então

  $$\|\mathbf{x} + \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2$$

- Verifique a *identidade do paralelogramo*

  $$\|\mathbf{x} + \mathbf{y}\|^2 + \|\mathbf{x} - \mathbf{y}\|^2 = 2\|\mathbf{x}\|^2 + 2\|\mathbf{y}\|^2$$

- Mostre que a norma, e portanto a distância, satisfazem a *desigualdade do triângulo*

  $$\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \qquad \mathrm{d}(\mathbf{x}, \mathbf{y}) \leq \mathrm{d}(\mathbf{x}, \mathbf{z}) + \mathrm{d}(\mathbf{z}, \mathbf{y})$$

  (calcule $\|\mathbf{x} + \mathbf{y}\|^2$ e use a desigualdade de Schwarz ...).
- [Ap69] **15.12**.

2. (matrizes simétricas e positivas) Os axiomas E1 e E2 dizem que um produto interno no espaço $\mathbb{R}^n$ é

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i,j} g_{ij} x^i y^j$$

onde $G = (g_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é uma matriz simétrica, i.e. tal que $g_{ij} = g_{ji}$. O axioma E3 diz que a matriz G é *positiva*, i.e. $\sum_{i,j} g_{ij}x^i x^j > 0$ se $\mathbf{x} \neq \mathbf{0}$. Em particular, considerando vetores $\mathbf{x}$ iguais aos vetores da base canónica, $g_{ii} > 0$ para todo i.

Numa base em que a matriz G é diagonal, com valores próprios $\lambda_i > 0$, o produto interno é

$$\langle \mathbf{x}, \mathbf{y}\rangle = \lambda_1 x^1 y^1 + \lambda_2 x^2 y^2 + \cdots + \lambda_n x^n y^n$$

Por exemplo, o produto escalar euclidiano em $\mathbb{R}^n$ é definido pela matriz identidade $I = (\delta_{ij})$ numa base ortonormada (como a base canónica).

- Verifique que
$$\langle \mathbf{r}, \mathbf{r}'\rangle = 2xx' + xy' + yx' + yy'$$
é um produto interno em $\mathbb{R}^2$.
- Verifique que a *métrica de Lorentz/Minkowski*
$$\langle (t, \mathbf{r}), (t', \mathbf{r}')\rangle = -c^2 tt' + \mathbf{r} \cdot \mathbf{r}'$$
não é um produto interno em $\mathbb{R}^4 = \mathbb{R} \times \mathbb{R}^3$.

3. (ℓ^2 space) The simplest infinite dimensional Euclidean space is the space ℓ^2 of those infinite sequences $\mathbf{x} = (x_k)_{k\in\mathbb{N}} = (x_1, x_2, x_3, \dots) \in \mathbb{C}^{\mathbb{N}}$ of complex numbers $x_k \in \mathbb{C}$ with finite norm $\|\mathbf{x}\|^2 := \sum_k |x_k|^2 < \infty$, equipped with the inner product
$$\langle \mathbf{x}, \mathbf{y}\rangle := \sum_{k=1}^{\infty} x_k\, \overline{y_k}$$
(convergence of the inner product comes from Cauchy-Schwarz inequality applied to finite sums and convergence of the infinite sums defining the norms of the two vectors). The infinite system $\mathbf{e}_1 = (1, 0, \dots)$, $\mathbf{e}_2 = (0, 1, 0, \dots)$, ... is made of pairwise orthogonal unitary vectors. By definition, any vector $\mathbf{x} \in \ell^2$ is a sum $\sum_{n=1}^{\infty} \langle \mathbf{x}, \mathbf{e}_n\rangle\, \mathbf{e}_n$.

   In some precise sense, any reasonable (complete and separable) infinite dimensional complex Euclidean space is isomorphic to ℓ^2.

   Also useful are spaces of two-sided sequences $(x_k)_{k\in\mathbb{Z}} = (\dots, x_{-2}, x_{-1}, x_0, x_1, x_2, \dots)$, equipped with the obvious inner product.

4. (L^2 spaces) Typical Euclidean spaces of interest in analysis of PDEs (and in physics) are spaces $\mathbf{E}$ made of continuous (or at least integrable) functions $f : X \to \mathbb{R}$ or $\mathbb{C}$, defined in some domain X which may be an interval like $[-\pi, \pi]$, the whole real line $\mathbb{R}$, some domains $\Omega \subset \mathbb{R}^n$ (or more sophisticated objects called "manifolds", and so on), equipped with the L^2 inner product defined as
$$\langle f, g\rangle := \int_X f(t)\, \overline{g(t)}\, dt\,.$$
For this to make sense we must require finiteness of the norm, i.e.
$$\|f\|^2 = \int_X |f(t)|^2\, dt\, < \infty\,.$$
In some sense this space is too small, and one should "complete" it in order to include the limits of its fundamental sequences, and then one should restore positivity identifying functions differing by something that has zero norm. The result is a complete Euclidean space called $L^2(X)$. In some other sense it is too large, since interesting linear operators, such as differential operators, may only be defined in strict subspaces made of those functions which admit a sufficient number of derivatives. For example, one may consider the subspace $\mathcal{C}_c^\infty(X) \subset L^2(X)$ of infinitely differentiable functions with compact support. It turns out that this subspace is "dense" in $L^2(X)$ (this means that for any $f \in L^2(X)$ and any precision $\varepsilon > 0$ we may find a smooth function with compact support $g \in \mathcal{C}_c^\infty(X)$ such that $\|f - g\| < \varepsilon$).

When $f(t)$ is a time-dependent "signal", as for example an electric current or the amplitude of a sound wave, then its squared modulus $|f(t)|^2$ is a "power", hence its time integral $\int_{\mathbb{R}} |f(t)|^2 \, dt$ has the meaning of an "energy".

If a continuous-time signal $f(t)$ is "sampled" at integer multiples of some sampling time $\tau > 0$, we get a discrete-time signal $z_n := f(n\tau)$, with $n \in \mathbb{N}$ or $\mathbb{Z}$, which is a sequence. The L^2 inner product is then approximated/replaced by the ℓ^2 inner product.

5. (ortogonalização de Gram-Schmidt) Num espaço euclidiano, há uma relação simples entre ortogonalidade e independência.

Teorema 15.2. *Um conjunto ortogonal de elementos não nulos de um espaço euclidiano* $\mathbf{E}$ *é independente.*

Demonstração. Sejam $\mathbf{v}_1, \dots, \mathbf{v}_n$ vetores não nulos e (dois a dois) ortogonais, i.e. tais que $\langle \mathbf{v}_i, \mathbf{v}_j \rangle = 0$ se $i \neq j$ e $\|\mathbf{v}_k\| \neq 0$ para todo k. A igualdade $\lambda \mathbf{v}_1 + \lambda \mathbf{v}_2 + \dots + \lambda_n \mathbf{v}_n = 0$ implica, ao calcular os produtos internos com os vetores $\mathbf{v}_k$, que $\lambda_k \|\mathbf{v}_k\|^2 = 0$ para todos os k, ou seja, sendo os $\|\mathbf{v}_k\| > 0$, que todos os λ_k são nulos. □

Em particular, todo o conjunto ortogonal de n vetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ não nulos num espaço euclidiano $\mathbf{E} \approx \mathbb{R}^n$ ou $\mathbb{C}^n$ de dimensão n é uma base. Os vetores $\mathbf{u}_k := \mathbf{v}_k / \|\mathbf{v}_k\|$ formam então uma base *ortonormada*. (ou seja, formada por vetores unitários ortogonais).

Conjuntos independentes, não necesariamente finitos, podem ser feitos ortogonais, de acordo com o

Teorema 15.3 (ortogonalização de Gram-Schmidt)**.** *Seja* $\mathbf{v}_1, \mathbf{v}_2, \dots,$ *um conjunto independente de vetores do espaço euclidiano* $\mathbf{E}$*. Então existe um conjunto ortogonal* $\mathbf{u}_1, \mathbf{u}_2, \dots$ *tal que os espaços* $[\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n]$ *gerados pelo primeiros* n *vetores coincidem com o espaços* $[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$*, para todo o* n*.*

Demonstração. Basta definir o conjunto $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n, \dots$ recursivamente por

$$\mathbf{u}_1 = \mathbf{v}_1 \qquad \mathbf{u}_2 = \mathbf{v}_2 - \frac{\langle \mathbf{v}_2, \mathbf{u}_1 \rangle}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 \qquad \dots$$

$$\mathbf{u}_{n+1} = \mathbf{v}_{n+1} - \left(\frac{\langle \mathbf{v}_{n+1}, \mathbf{u}_1 \rangle}{\|\mathbf{u}_1\|^2} \mathbf{u}_1 + \frac{\langle \mathbf{v}_{n+1}, \mathbf{u}_2 \rangle}{\|\mathbf{u}_2\|^2} \mathbf{u}_2 + \dots + \frac{\langle \mathbf{v}_{n+1}, \mathbf{u}_n \rangle}{\|\mathbf{u}_n\|^2} \mathbf{u}_n \right) \qquad \dots$$

Ou seja, $\mathbf{u}_{n+1}$ é obtido retirando de $\mathbf{v}_{n+1}$ a soma das suas projeções sobre os $\mathbf{u}_1, \dots, \mathbf{u}_n$. É imediato verificar que $\mathbf{u}_{n+1}$ é ortogonal ao subespaço $[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$, e portanto a todos os $\mathbf{u}_k$ com $k \leq n$. □

Em particular,

Teorema 15.4. *Todo espaço euclidiano de dimensão finita admite uma base ortonormada, e de consequência é isomorfo a* $\mathbb{R}^n$ *ou* $\mathbb{C}^n$*, dependendo se real ou complexo, munidos do produto interno usual.*

6. (polinómios de Legendre) Ortonormalize a família $v_0(t) = 1$, $v_1(t) = t$, $v_2(t) = t^2$, ..., $v_n(t) = t^n$, ... de vetores no espaço $\mathbf{E}$ das funções contínuas $x : [-1,1] \to \mathbb{R}$, munido do produto interno
$$\langle x, y\rangle := \int_{-1}^{1} x(t)\, y(t)\, dt\,.$$
O resultado é a família dos *polinómios de Legendre*,
$$\ell_0(t) = 1, \quad \ell_1(t) = t, \quad \ell_2(t) = t^2 - 1/3, \quad \ell_3(t) = t^3 - , \dots$$

7. (complemento ortogonal, projeção ortogonal) Seja $\mathbf{E}$ um espaço euclidiano, real ou complexo. O complemento ortogonal do subconjunto $X \subset \mathbf{E}$ (não necesariamente um subespaço) é o subespaço linear
$$\boxed{X^\perp := \{\mathbf{v} \in \mathbf{E} \ \text{ t.q. } \ \langle \mathbf{v}, \mathbf{x}\rangle = 0 \ \ \forall \mathbf{x} \in X\}}$$
Seja $S \subset \mathbf{E}$ um subespaço de dimensão finita do espaço euclidiano $\mathbf{E}$. Então cada vetor $\mathbf{v} \in \mathbf{E}$ pode ser representado de maneira única como soma
$$\mathbf{v} = \mathbf{s} + \mathbf{t}$$
de um vetor $\mathbf{s} \in S$ e um vetor $\mathbf{t} \in S^\perp$, ou seja, o espaço $\mathbf{E}$ é uma soma direta $\mathbf{E} = S \oplus S^\perp$. De facto, se $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ é uma base ortonormada de S (que existe pelo teorema de Gram-Schmidt), basta escolher
$$\mathbf{s} = \sum_{k=1}^{n} \langle \mathbf{v}, \mathbf{e}_k\rangle\, \mathbf{e}_k$$
e verificar que $\mathbf{t} := \mathbf{v} - \mathbf{s}$ é ortogonal a todos os vetores $\mathbf{e}_1, \dots, \mathbf{e}_n$, e portanto a todos os vetores de S. O vetor $\mathbf{s}$ é dito *projeção (ortogonal)* de $\mathbf{v}$ sobre S, e o operador $P_S : \mathbf{E} \to \mathbf{E}$, definido por $P_S(\mathbf{v}) = \mathbf{s}$, é dito *projeção ortogonal* sobre o subespaço S.

Teorema 15.5 (teorema de aproximação). *A projecção rortogonal realiza a distância entre* $\mathbf{v}$ *e* S*, ou seja, para todo* $\mathbf{s}' \in S$*,*
$$\|\mathbf{v} - P_S(\mathbf{v})\| \le \|\mathbf{v} - \mathbf{s}'\|\,.$$

Demonstração. A vetor $\mathbf{s}' - \mathbf{s}$, que pertence ao subespaço S, é ortogonal ao vetor $\mathbf{t} = \mathbf{v} - \mathbf{s}$, e portanto, pelo teorema de Pitágoras, $\|\mathbf{v} - \mathbf{s}'\|^2 = \|\mathbf{v} - \mathbf{s}\|^2 + \|\mathbf{s}' - \mathbf{s}\|^2 \ge \|\mathbf{v} - \mathbf{s}\|^2$. □

Em dimensão infinita, este resultado é falso em geral. Projeções ortogonais, e portanto decomposiçoes como somas diretas $\mathbf{E} = S \oplus S^\perp$ são possíveis apenas quando o subespaço S é fechado.

- Mostre que $X^\perp$ é um subespaço linear de $\mathbf{E}$.
- Verifique que $\|\mathbf{v}\|^2 = \|\mathbf{s}\|^2 + \|\mathbf{t}\|^2$.
- Verifique que $P_S(\mathbf{E}) = S$, e que $P_S P_S = P_S$.

8. (bases ortonormadas e coeficientes de Fourier) Seja $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n, \dots$ um conjunto/sistema ortonormado do espaço euclidiano $\mathbf{E}$ (ou seja, $\langle \mathbf{e}_i, \mathbf{e}_j\rangle = 0$ se $i \neq j$, e $\|\mathbf{e}_i\| = 1$). Os *coeficientes de Fourier* do vetor $\mathbf{x} \in \mathbf{E}$ (relativamente ao conjunto ortonormado $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n, \dots$) são os escalares
$$\boxed{x_k := \langle \mathbf{x}, \mathbf{e}_k\rangle}$$
A projeção ortogonal de um vetor $\mathbf{x}$ sobre o subespaço de dimensão finita $\mathbf{E}_N \subset \mathbf{E}$, gerado pelos $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N$, é o vetor $\mathbf{x}_N = \sum_{k=1}^{N} x_k \mathbf{e}_k$. Da desigualdade $\|\mathbf{x} - \mathbf{x}_N\|^2 \ge 0$ segue que $\sum_{k=1}^{N} |x_k|^2 \le \|\mathbf{x}\|^2$. Em particular, a série de termos não-negativos $\sum_k |x^k|^2$ é convergente. No limite quando $N \to \infty$, segue a

Teorema 15.6 (desigualdadde de Bessel)**.** *Seja* $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n, \dots$ *um sistema ortonormado do espaço euclidiano* $\mathbf{E}$*. Se* $\mathbf{x} \in \mathbf{E}$ *e* $x_k = \langle \mathbf{x}, \mathbf{e}_k \rangle$ *são os seus coeficientes de Fourier, então*

$$\boxed{\sum_{k=1}^{\infty} |x_k|^2 \leq \|\mathbf{x}\|^2}$$

Se $\mathbf{E}$ tem dimensão finita e se $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ é uma base ortonormada, então cada vetor $\mathbf{x}$ pode ser representado de maneira única como $\mathbf{x} = \sum_{k=1}^{n} x_k \mathbf{e}_k$. Então, vale o análogo do eorema de Pitágoras,

$$\boxed{\|\mathbf{x}\|^2 = \sum_{i=1}^{n} |x_k|^2}$$

Mais m geral, o produto escalar entre os vetores $\mathbf{x}$ e $\mathbf{y}$ é dado pela *identidade de Parseval*

$$\boxed{\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_k \, \overline{y_k}}$$

- Mostre que se $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ é uma base ortonormada do espaço euclidiano de dimensão finita $\mathbf{E} \approx \mathbb{R}^n$ ou $\mathbb{C}^n$, então todo o vetor $\mathbf{x} \in \mathbf{E}$ é igual a

$$\mathbf{x} = \sum_{k=1}^{n} x_k \mathbf{e}_k$$

  onde $x_k = \langle \mathbf{x}, \mathbf{e}_k \rangle$.
- [Ap69] **15.12**.

9. (séries de Fourier) O produto interno L^2 no espaço $\mathbf{H}$ das funções contínuas (ou simplesmente integráveis) $f : [-\pi, \pi] \to \mathbb{C}$ é

$$\langle f, g \rangle := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \, \overline{g(t)} \, dt \, .$$

Não é difícil verificar que $1, \cos(t), \cos(2t), \dots, \sin(t), \sin(2t), \dots$ é um conjunto ortogonal. Mais fácil é verificar que a família $(\mathbf{e}_n)_{n \in \mathbb{Z}}$ das "harmónicas"

$$\mathbf{e}_n(t) := e^{int} \, ,$$

com $n \in \mathbb{Z}$, é um conjunto ortonormado.

Uma combinação linear $\sum_{-N}^{N} c_n e^{int}$ é chamada "polinómio trigonométrico" de grau N, e o espaço $\mathbf{E}_N$ dos polinómios trigonométricos de grau N é um subespaço de dimensão finita de $\mathbf{E}$, gerado pelas harmónicas $\mathbf{e}_n$ com $|n| \leq N$.

Seja $f(t)$ uma função integrável (por exemplo, seccionalmente contínua) no intervalo $[-\pi, \pi]$. A sua projeção ortogonal sobre $\mathbf{E}_N$ é o polinómio trigonométrico

$$S_N f(t) := \sum_{-N}^{N} \widehat{f}(n) \, e^{int} \, ,$$

onde os *coeficientes de Fourier* de f são definidos por

$$\boxed{\widehat{f}(n) := \langle f, \mathbf{e}_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \, e^{-int} \, dt \, .}$$

A *série de Fourier* de $f(t)$ é a série formal

$$\sum_{-\infty}^{\infty} \widehat{f}(n) \, e^{int} = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nt) + b_n \sin(nt))$$

(a segunda é obtida da primeira usando a fórmula de Euler, e os coeficientes a_n e b_n são combinações lineares dos $\widehat{f}(\pm n)$). Se a função $f(t)$ é suficientemente regular (por exemplo, diferenciável com continuidade), a série de Fourier converge para a própria função.

10. (Hilbert spaces and Dirac's notation) Physicists are interested in certain infinite dimensional complex Euclidean spaces $\mathbf{H}$ called *Hilbert spaces*. They are characterized by a "completeness condition" and by the fact of having an "infinite countable basis" $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \dots$ Each vector may be written as an infinite sum $\mathbf{x} = \sum_k x_k \, \mathbf{e}_k$, with Fourier coefficients $x_k = \langle \mathbf{x}, \mathbf{e}_k \rangle$, and (the square of) its norm is the sum of the series $\|\mathbf{x}\|^2 = \sum_k |x_k|^2$. It happens that the inner product induces an isomorphism between $\mathbf{H}$ and its dual $\mathbf{H}^*$. Paul Dirac, one of the fathers os quantum mechanics, invented the following notation[24]: a generic vector is denoted by $|x\rangle$ and called *ket*, while a generic co-vector is denoted by $\langle y|$ and called *bra*, so that their pairing $\langle y|x\rangle$, the inner product, is a *bra-ket*.

11. (quadrados mínimos e regressão linear) Sejam $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ e $B \in \mathbb{R}^m$. Uma *solução de quadrados mínimos* do sistema $AX = B$ (que pode ser inconsistente!) é um vetor $\overline{X} \in \mathbb{R}^n$ que minimiza a soma
$$(\epsilon^1)^2 + (\epsilon^2)^2 + \cdots + (\epsilon^m)^2$$
dos quadrados dos "erros" $\epsilon^i := \left(\sum_j a^i{}_j \, x^j\right) - b^i$, ou seja, o quadrado da norma $\|AX - B\|^2$ em $\mathbb{R}^m$. Se as colunas de A são linearmente independentes, então a matriz $A^\top A$ é invertível, e a única solução de quadrados mínimos é
$$\boxed{\overline{X} = (A^\top A)^{-1} A^\top B}$$
   - Verifique que $\overline{X} = (\overline{x}^1, \overline{x}^2, \dots, \overline{x}^n) \in \mathbb{R}^n$ é um mínimo de $\|AX - B\|^2$ se $A^\top A\overline{X} = A^\top B$ (calcule as derivadas parciais em ordem aos $\overline{x}^j$'s).
   - Mostre que se $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ tem caraterística n então $A^\top A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ também tem caraterística n, e portanto é invertível.

[24]P.A.M. Dirac, A new notation for quantum mechanics, *Mathematical Proceedings of the Cambridge Philosophical Society* **35** (3) (1939), 416-418.
P.A.M. Dirac, *The Principles of Quantum Mechanics*, Oxford University Press, 1930.

16 Transformações lineares

ref: [Ap69] *Vol 2*, **2.1-9** ; [La87] *Ch. III*

1. (transformações lineares) Uma *transformação/aplicação/operador linear* entre os espaços vetoriais reais (ou complexos) $\mathbf{V}$ e $\mathbf{W}$ é uma função $L : \mathbf{V} \to \mathbf{W}$ aditiva e homogénea, ou seja, tal que
$$L(\mathbf{v}+\mathbf{v}') = L(\mathbf{v}) + L(\mathbf{v}') \qquad \text{e} \qquad L(\lambda \mathbf{v}) = \lambda\, L(\mathbf{v})$$
ou seja, tal que
$$\boxed{L(\lambda\mathbf{v}+\lambda'\mathbf{v}') = \lambda L(\mathbf{v}) + \lambda' L(\mathbf{v}')}$$
$\forall\, \mathbf{v}, \mathbf{v}' \in \mathbf{V}$ e $\forall \lambda, \lambda' \in \mathbb{R}$ (ou $\in \mathbb{C}$). É usual omitir as parêntesis, e denotar a imagem do vetor $\mathbf{v}$ simplesmente por $L\mathbf{v}$.

   O espaço $\mathcal{L}(\mathbf{V},\mathbf{W}) := \mathrm{Hom}_{\mathbb{R}}(\mathbf{V},\mathbf{W})$ (ou $\mathrm{Hom}_{\mathbb{C}}(\mathbf{V},\mathbf{W})$) das transformações lineares de $\mathbf{V}$ em $\mathbf{W}$ é um espaço linear real (ou complexo) se a adição e a multiplicação por um escalar são definidas por
$$(L+L')(\mathbf{v}) := L(\mathbf{v}) + L'(\mathbf{v}) \qquad\qquad (\lambda L)(\mathbf{v}) := \lambda L(\mathbf{v})$$
Em particular, é um espaço linear o espaço $\mathrm{End}(\mathbf{V}) := \mathcal{L}(\mathbf{V},\mathbf{V})$ dos *endomorfismos* de $\mathbf{V}$. Uma transformação linear bijetiva $L : \mathbf{V} \to \mathbf{W}$ é dita *isomorfismo (linear)* entre os espaços lineares $\mathbf{V}$ e $\mathbf{W}$.

   - Se $L : \mathbf{V} \to \mathbf{W}$ é uma transformação linear, então $L(\mathbf{0}) = \mathbf{0}$ e $L(-\mathbf{v}) = -L(\mathbf{v})$.
   - Uma transformação linear $L : \mathbf{V} \to \mathbf{W}$ envia retas afins $\mathbf{a} + \mathbb{R}\mathbf{v} \subset \mathbf{V}$ em retas afins $\mathbf{b} + \mathbb{R}\mathbf{w} \subset \mathbf{W}$, se $\mathbf{b} = L(\mathbf{a})$ e $\mathbf{w} = L(\mathbf{v}) \neq \mathbf{0}$, ou em pontos $\mathbf{b}$, se $L(\mathbf{v}) = \mathbf{0}$. Em particular, envia retas $\mathbb{R}\mathbf{v} \subset \mathbf{V}$ passando pela origem em retas $\mathbb{R}\mathbf{w} \subset \mathbf{W}$ passando pela origem.
   - A inversa de uma transformação linear $L : \mathbf{V} \to \mathbf{W}$ bijetiva é uma transformação linear $L^{-1} : \mathbf{W} \to \mathbf{V}$.
   - Uma transformação $T : \mathbb{R}^n \to \mathbb{R}^m$, definida por
$$(x^1, x^2, \dots, x^n) \mapsto (T^1(x^1, x^2, \dots, x^n), T^2(x^1, x^2, \dots, x^n), \dots, T^m(x^1, x^2, \dots, x^n))$$
é linear sse todas as suas "coordenadas" $T^i : \mathbb{R}^n \to \mathbb{R}$, com $i = 1, 2, \dots, m$, são lineares.
   - A transformação *identidade* $I_n : \mathbb{R}^n \to \mathbb{R}^n$, definida por $\mathbf{x} \mapsto \mathbf{x}$, e transformação *nula* $\mathbf{0}_n : \mathbb{R}^n \to \mathbb{R}^n$, definida por $\mathbf{x} \mapsto \mathbf{0}$, são lineares.
   - Diga se as seguintes aplicações de $\mathbb{R}^n$ em $\mathbb{R}^m$ são lineares.
$$(x,y) \mapsto (3x-5y, x-y) \qquad (x,y) \mapsto (x^2, xy)$$
$$(x,y,z) \mapsto (x, y+z, 2) \qquad (x,y,z) \mapsto (x, y+z, 0)$$
$$(x,y,z) \mapsto (x-y, z+y, 3x+2y-z) \qquad (x,y,z) \mapsto (1,2,3)$$
$$(x_1, x_2, \dots, x_n) \mapsto (0,0,1) \qquad (x_1, x_2, \dots, x_n) \mapsto (0,0,\dots,0) \in \mathbb{R}^m$$
   - Diga se as seguintes aplicações de $T : \mathbb{R}^2 \to \mathbb{R}^2$ são lineares.

     T transforma cada ponto no seu simétrico em relação à reta $y = 0$

     T transforma cada ponto no seu simétrico em relação à reta $y = x$

     T transforma o ponto de coordenadas polares (r, θ) no ponto de coordenadas polares$(2r, \theta)$

     T transforma o ponto de coordenadas polares (r, θ) no ponto de coordenadas polares $(r, \theta+\pi/2)$
   - Seja $L : \mathbb{R}^2 \to \mathbb{R}^2$ uma transformação linear tal que $L(1,1) = (1,4)$ e $L(2,-1) = (-2,3)$. Determine $L(5,-1)$ (observe que $1 + 2 \cdot 2 = 5$ e $1 + 2 \cdot (-1) = -1$).

- Determine a transformação linear $T : \mathbb{R}^2 \to \mathbb{R}^2$ tal que
$$T(\mathbf{i}) = 2\mathbf{i} + \mathbf{j} \qquad \text{e} \qquad T(\mathbf{j}) = \mathbf{i} - 3\mathbf{j}$$
- Se $X \subset \mathbf{V}$ é um subespaço linear de $\mathbf{V}$ e $L : \mathbf{V} \to \mathbf{W}$ é uma transformação linear, então a imagem $L(X)$ é um subespaço linear de $\mathbf{W}$.
- As translações de $\mathbb{R}^n$, as transformações $\mathbf{v} \mapsto \mathbf{v} + \mathbf{a}$ com $\mathbf{a} \in \mathbb{R}^n$, são transformações lineares?
- As homotetias de $\mathbb{R}^n$, as transformações $\mathbf{v} \mapsto \lambda\mathbf{v}$ com $\lambda \in \mathbb{R}$, são transformações lineares?
- [Ap69] **16.4**.

2. (núcleo e imagem) Seja $L : \mathbf{V} \to \mathbf{W}$ uma transformação linear. O *núcleo/espaço nulo* (em inglês, *kernel*) de L é o subespaço vetorial
$$\boxed{\ker(L) := \{\mathbf{v} \in \mathbf{V} \ \text{ t.q. } \ L(\mathbf{v}) = \mathbf{0}\} \subset \mathbf{V}}$$
A *imagem* de L é o subespaço vetorial
$$\boxed{\operatorname{im}(L) := L(\mathbf{V}) = \{L(\mathbf{v}) \ \text{ com } \ \mathbf{v} \in \mathbf{V}\} \subset \mathbf{W}}$$
A dimensão do núcleo é dita *nulidade* de L, e a dimensão da imagem é dita *ordem* de L.

Teorema 16.1. *Seja $L : \mathbf{V} \to \mathbf{W}$ uma transformação linear. Se $\mathbf{V}$ tem dimensão finita, então também a imagem $L(\mathbf{V})$ tem dimensão finita e*
$$\boxed{\dim_{\mathbb{R}} \ker(L) + \dim_{\mathbb{R}} \operatorname{im}(L) = \dim_{\mathbb{R}} \mathbf{V}}$$

Demonstração. Seja $\dim_{\mathbb{R}}(\mathbf{V}) = n$. Se os vetores $\mathbf{v}_1, , \dots, \mathbf{v}_k$ formam uma base de $\ker(L)$, e se juntamente com os vetores $\mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ formam uma base de $\mathbf{V}$ (se $k < n$, caso contrário o teorema é trivial), então é imediato verificar que os vetores $\mathbf{w}_1 = L(\mathbf{v}_{k+1}), \dots, \mathbf{w}_{n-k} = L(\mathbf{v}_n)$ geram $\operatorname{im}(L)$ e são independentes. $\square$

- Mostre que $\ker(L)$ é um subespaço de $\mathbf{V}$ e que $\operatorname{im}(L)$ é um subespaço de $\mathbf{W}$.
- Determine o núcleo, a imagem, a nulidade e a ordem das seguintes transformações lineares
$$L(x,y) = (x+y, x-y) \qquad L(x,y) = (y, -x) \qquad L(x,y) = (x+y, 3x-2y)$$
$$L(x,y,z) = (2x, 3y, 0) \qquad L(x,y,z) = (x+y, y+z, z+x)$$
$$L(x,y,z) = (0,0,0) \qquad L(x,y,z) = (x,y)$$
- Determine o núcleo, a imagem, a nulidade e a ordem da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto no seu simétrico em relação à reta $x = 0$
- Determine o núcleo, a imagem, a nulidade e a ordem da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto no seu simétrico em relação à origem.
- [Ap69] **16.14**.

3. (operadores) Seja $\mathbf{V}$ um espaço linear (de dimensão não necessariamente finita!). Um *operador (linear)* em $\mathbf{V}$ é uma transformação linear $A : \mathbf{D} \to \mathbf{V}$ definida num subespaço linear $\mathbf{D} \subset \mathbf{V}$, dito *domínio* do operador A (esta definição é importante em análise, quando $\mathbf{V}$ é um espaço de dimensão infinita e o operador apenas pode ser definido num subespaço próprio de $\mathbf{V}$). Um subespaço linear $\mathbf{W} \subset \mathbf{D}$ é *invariante* se $A(\mathbf{W}) \subset \mathbf{W}$ (ou seja, se $\mathbf{v} \in \mathbf{W}$ implica $A\mathbf{v} \in \mathbf{W}$), e portanto a restrição de A a $\mathbf{W}$ é um endomorfismo $A|_{\mathbf{W}} : \mathbf{W} \to \mathbf{W}$.

4. (operadores derivação e multiplicação) O operador *derivação* envia uma função derivável $f(x)$ na função
$$(Df)(x) := f'(x)\,.$$
Pode ser pensado como um operador $D : \mathcal{C}^{k+1}(\mathbb{R}) \to \mathcal{C}C^k(\mathbb{R})$, ou também como un endomorfismo $D : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$ do espaço linear das funções $f(x)$ infinitamente diferenciáveis. O operador *multiplicação* envia uma função $f(x)$ na função
$$(Xf)(x) := x \cdot f(x)\,.$$
   - Determine o núcleo e a imagem do operador derivação $D : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$,
   - O subespaço $\mathcal{P}$ dos polinómio é um subespaço invariante de $D : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$. O subespaço $\mathcal{P}_{\leq n}(\mathbb{R}) \subset \mathcal{P}(\mathbb{R})$ dos polinómios de grau $\leq n$ é um subespaço invariante do operador derivação $D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$?

5. (operador primitivação) O operador *primitivação* envia uma função integrável (na reta real ou num intervalo da reta) $f(x)$ na função
$$(Pf)(x) := \int_c^x f(t)\,dt$$
Pode ser pensado como um endomorfismo $P : \mathcal{C}^0(\mathbb{R}) \to \mathcal{C}^0(\mathbb{R})$ do espaço linear das funções contínuas.
   - Determine o núcleo $\ker(P)$ e a imagem $\mathrm{im}(P) = P(\mathcal{C}^0(\mathbb{R}))$.

6. (Laplaciano, equação de Laplace, funções harmónicas) O *Laplaciano* (ou *operador de Laplace*) no espaço euclidiano $\mathbb{R}^n$ é o operador diferencial $\Delta := \mathrm{div} \circ \mathrm{grad}$, definido, em coordenadas cartesianas (ou seja, relativamente à uma base ortonormal), por
$$(\Delta f)(\mathbf{x}) := \frac{\partial^2 f}{\partial (x^1)^2}(\mathbf{x}) + \frac{\partial^2 f}{\partial (x^2)^2}(\mathbf{x}) + \cdots + \frac{\partial^2 f}{\partial (x^n)^2}(\mathbf{x})$$
se $f(\mathbf{x}) = f(x^1, x^2, \dots, x^n)$ é uma função real de classe C^2 defnida num domínio $\Omega \subset \mathbb{R}^n$. O Laplaciano pode ser pensado como um operador $\Delta : \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$. A *equação de Laplace* para o campo escalar $f(\mathbf{x})$, definido num domínio $\Omega \subset \mathbb{R}^n$, é
$$\Delta f = 0$$
As soluções da equação de Laplace, ou seja, os campos escalares contidos no núcleo do Laplaciano, são ditas *funções harmónicas.*
   - Quais funções satisfazem a equação de Laplace $f''(x) = 0$ na reta?
   - Determine as soluções da equação de Laplace $f''(x) = 0$ no intervalo $[a, b] \subset \mathbb{R}$ com condições de fronteira $f(a) = A$ e $f(b) = B$.
   - Verifique que $f(\mathbf{r}) = \log \|\mathbf{r}\|$ é uma solução da equação de Laplace em $\mathbb{R}^2 \backslash \{\mathbf{0}\}$.
   - Verifique que (o potencial elétrico/gravitacional gerado por uma carga/massa unitária colocada na origem do espaço 3-dimensional) $f(\mathbf{r}) = 1/\|\mathbf{r}\|$. é uma solução da equação de Laplace em $\mathbb{R}^3 \backslash \{\mathbf{0}\}$.

7. (composição e álgebra dos endomorfismos) A composição de duas transformações lineares $L : \mathbf{V} \to \mathbf{W}$ e $M : \mathbf{W} \to \mathbf{Z}$, definida por $(ML)(\mathbf{v}) := M(L(\mathbf{v}))$, é uma transformação linear. Em particular, a composição de dois endomorfismos de um espaço vetorial $\mathbf{V}$ é um endomorfismo de $\mathbf{V}$. A n-ésima iterada do endomorfismo $L \in \mathrm{End}(\mathbf{V})$ é o endomorfismo E^n definido indutivamente por
$$E^0 = I_{\mathbf{V}}\,, \qquad E^{n+1} = EE^n \qquad \text{se } n \geq 1\,.$$
A composição de transformações lineares satisfaz as propriedades distributivas
$$(L + M)N = LN + MN \qquad L(M + N) = LM + LN$$

(que justificam a notação "multiplicativa" LM em vez de $L \circ M$).

A composição não é comutativa! Ou seja, em geral, não há razão para que LM seja igual a ML. Os endomorfismos $L, M \in \mathrm{End}(\mathbf{V})$ *comutam entre si/são permutáveis* se $LM = ML$. A obstrução é o *comutador*, definido por

$$[L, M] := LM - ML\,,$$

que é igual a transformação nula sse L e M comutam.

- Calcule a composição ML quando

$$L(x, y) = (x + y, x - y) \qquad M(x, y) = 2x - 3y$$

$$L(x, y, z) = (x - y + z, z - y) \qquad M(x, y) = (x, y, x + y)$$

- Verifique que cada endomorfismo comuta com si próprio, ou seja, $[L, L] = 0$
- Calcule o comutador entre os endomorfismos do plano $E_+(x, y) = (y, 0)$ e $E(x, y) = (x, -y)$.
- Se $[L, M] = 0$ e $[M, N] = 0$, é verdade que $[L, N] = 0$?
- Determine todos os endomorfismos do plano que comutam com $I(x, y) = (x, y)$.

8. (álgebra de Weyl-Heisenberg-Schrödinger) A álgebra do grupo de Weyl-Heisenberg[25] (em dimensão um) é gerada pelos operadores *multiplicação* e *derivação*

$$(Qf)(x) := x\, f(x) \qquad \text{e} \qquad (\partial f)(x) := f'(x)\,,$$

definidos, por exemplo, no *espaço de Schwarz* $\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R})$ das funções complexa $f(x)$ infinitamente diferenciáveis tais que todas as derivadas decrescem mais rapidamente do que o inverso de qualquer polinómio[26]. Os operadores Q e $P := -i\hbar\partial$ (onde $\hbar$ é a constante de Planck reduzida) representam a *posição* e o *momento*, respetivamente, de uma partícula quântica livre na "representação de Schrödinger". Os operadores P e Q não comutam, pois

$$[Q, P] = i\hbar$$

e está é a causa do "princípio de incerteza de Heisenberg".

9. (equações diferenciais ordinárias lineares) Uma *equação diferencial ordinária linear* de ordem n é uma lei

$$a_n \frac{d^n x}{dt^n} + a_{n-1} \frac{d^{n-1} x}{dt^{n-1}} + \cdots + a_1 \frac{dx}{dt} + a_0 x = f(t)$$

para a função $x(t)$, onde a_k são coeficientes (não necessariamente constantes) e $f(t)$ é uma função dada (uma força se $n = 2$). Pode ser escrita como

$$Lx = f$$

se $L : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$ é o operador diferencial $L := a_n D^n + a_{n-1} D^{n-1} + \cdots + a_1 D + a_0$. A equação *homogénea* associada é a equação diferencial $Ly = 0$, ou seja,

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \cdots + a_1 \frac{dy}{dt} + a_0 y = 0\,.$$

O espaço das soluções da equação homogénea $Lz = 0$ é um subespaço vetorial de dimensão finita, tipicamente $\mathbf{H} = \ker(L) \approx \mathbb{R}^n$. Por exemplo, se os coeficientes a_k's são constantes, uma base de $\mathbf{H}$ é formada pelos exponenciais complexos $y_k(t) = e^{z_k t}$, onde $z_1, z_2, \ldots, z_n$ são as

[25] Os físicos dizem "grupo de Weyl", que era um matemático (colega de Einstein em Zurich e depois em Princeton), e os matemático dizem "grupo de Heisenberg", que era um físico, um dos pais da mecânica quântica.

[26] Ou seja, uma função f está no espaço de Schwartz se para todos $n, m \in \mathbb{N}$ existem constantes $C_{n,m}$ tais que $\|Q^n \partial^m f\|_\infty \leq C_{n,m}$. De consequência, se $f \in \mathcal{S}(\mathbb{R})$ então também $Q^n \partial^m f \in \mathcal{S}(\mathbb{R})$ para todos n, m.

raizes (distintas, no caso genérico!) do polinómio caraterístico $a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$. A "solução geral", ou seja, o ponto genérico de $\mathbf{H}$, é portanto uma combinação linear

$$c_1 e^{z_1 t} + c_2 e^{z_2 t} + \cdots + c_n e^{z_n t} .$$

com $c_1, c_2, \ldots, c_n \in \mathbb{R}$ coeficientes arbitrários (a cada raiz z de multiplicidade m corresponde um "quase-polinómio" $(c_1 + c_2 t + \cdots + c_p t^{m-1}) e^{zt}$, assim que a solução geral no caso não-genérico de raizes múltiplas é uma soma de quase-polinómios). Se $z(t)$ é uma solução (apenas uma!) de $Lz = f$, então o espaço das todas as soluções de $Lx = f$ é o espaço afim $z + \mathbf{H}$, formado pelos pontos

$$z(t) + c_1 e^{z_1 t} + c_2 e^{z_2 t} + \cdots + c_n e^{z_n t} .$$

17 Transformações lineares e matrizes

ref: [Ap69] *Vol 2*, **2.10-16** ; [La87] *Ch. IV*

1. (espaço linear das matrizes) Uma *matriz* real (ou complexa) $m \times n$ é uma tabela

$$\boxed{A = ({a^i}_j) = \begin{pmatrix} {a^1}_1 & {a^1}_2 & \dots & {a^1}_n \\ {a^2}_1 & {a^2}_2 & \dots & {a^2}_n \\ \vdots & \vdots & \ddots & \vdots \\ {a^m}_1 & {a^m}_2 & \dots & {a^m}_n \end{pmatrix}}$$

de $m \cdot n$ números reais (ou complexos) dispostos em m linhas e n colunas. O número real (ou complexo) ${a^i}_j$ é dito *elemento/componente/entrada* ij da matriz A. O co-vetor e o vetor

$$\boxed{\mathbf{a}^i = ({a^i}_1, {a^i}_2, \dots, {a^i}_n) \in (\mathbb{R}^n)^* \approx \mathbb{R}^n \qquad \text{e} \qquad \mathbf{a}_j = \begin{pmatrix} {a^1}_j \\ {a^2}_j \\ \vdots \\ {a^m}_j \end{pmatrix} \in \mathbb{R}^m}$$

são ditos *i-ésima linha* e *j-ésima coluna* da matriz A, respetivamente. Se $n = m$, a matriz é dita "quadrada".

O espaço $\mathrm{Mat}_{m\times n}(\mathbb{R}) \approx \mathbb{R}^{m\cdot n}$ (ou $\mathrm{Mat}_{m\times n}(\mathbb{C}) \approx \mathbb{C}^{m\cdot n}$) das matrizes reais (ou complexas) $m \times n$ é um espaço linear real (ou complexo) se a adição e a multiplicação por um escalar são definidas por

$$A + B := ({a^i}_j + {b^i}_j) \qquad \text{e} \qquad \lambda A := (\lambda {a^i}_j)$$

se $A = ({a^i}_j)$ e $B = ({b^i}_j) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, e $\lambda \in \mathbb{R}$. O elemento neutro é a "matriz nula" $0 = (0) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, cujas entradas são todas nulas, que satisfaz $A + 0 = A$ para toda a matriz A. A matriz "oposta" da matriz $A = ({a^i}_j)$ é a matriz $-A := (-{a^i}_j)$, tal que $A + (-A) = 0$. Então, podemos definir $A - B := A + (-B)$.

- Calcule $A + B$, $A - B$, $2A - 3B$ quando

$$A = \begin{pmatrix} 1 & 3 \\ -2 & 0 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} -7 & 0 \\ 2 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 3 \\ -2 & 0 \\ 5 & 6 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} -7 & 0 \\ 2 & 1 \\ 2 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 3 & 4 \\ -2 & 0 & 1 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} -7 & 0 & 9 \\ 2 & 1 & -3 \end{pmatrix}$$

- Verifique que

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

  é uma base do espaço linear $\mathrm{Mat}_{2\times 2}(\mathbb{R})$.

- [Ap69] **16.16**.

2. (álgebra das matrizes) Se $A = ({a^i}_j) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ e $B = ({b^i}_j) \in \mathrm{Mat}_{n\times s}(\mathbb{R})$, o *produto (linhas por colunas)* AB é a matriz $AB = C = ({c^i}_j) \in \mathrm{Mat}_{m\times s}(\mathbb{R})$ definida por

$$\boxed{{c^i}_j = \textstyle\sum_{k=1}^n {a^i}_k\, {b^k}_j}$$

(ou seja, o elemento $c^i{}_j$ de $C = AB$ é o valor $c^i{}_j = \langle \mathbf{a}^i, \mathbf{b}_j \rangle$ da i-ésima linha de A sobre a j-ésima coluna de B). A "matriz identidade" $I_n \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é a matriz

$$I_n = (\delta^i{}_j) := \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & 1 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 \end{pmatrix},$$

que satisfaz $I_n A = A$ e $BI_n = B$ para todas as matrizes A e B. O produto é "associativo",

$$\boxed{A(BC) = (AB)C}$$

e satisfaz as "propriedades distributivas" à esquerda e à direita,

$$\boxed{A(B+C) = AB + AC \qquad \text{e} \qquad (A+B)C = AC + BC}$$

- Calcule AB quando

$$A = \begin{pmatrix} 1 & 3 \\ -2 & 0 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} -7 & 0 \\ 2 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 3 \\ -2 & 0 \\ 5 & 6 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} -7 & 0 & 1 \\ 2 & 1 & -1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 3 & 4 \\ -2 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \qquad \text{e} \qquad B = I_3$$

$$A = \begin{pmatrix} 1 & 3 & 5 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 5 & 3 \\ -3 & -5 \\ 1 & 1 \end{pmatrix}$$

- Existem matrizes $A \neq 0$ e $B \neq 0$ tais que $AB = 0$?
- [Ap69] **16.16**.

3. (transformações lineares e matrizes) Uma transformação linear $L : \mathbf{V} \to \mathbf{W}$ definida num espaço de dimensão finita $\mathbf{V}$ é determinada pelos seus valores nos vetores de uma base. De fato, se $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ é uma base ordenada de $\mathbf{V} \approx \mathbb{R}^n$ (por exemplo, a base canónica de $\mathbb{R}^n$), e definimos os vetores $\mathbf{w}_j := L(\mathbf{b}_j) \in \mathbf{W}$, com $j = 1, 2, \dots, n$, então o valor de L sobre o vetor genérico $\mathbf{x} = x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \dots + x^n\mathbf{b}_n$ é

$$\begin{aligned} L(\mathbf{x}) &= L(x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \dots + x^n\mathbf{b}_n) \\ &= x^1 L(\mathbf{b}_1) + x^2 L(\mathbf{b}_2) + \dots + x^n L(\mathbf{b}_n) \\ &= x^1\mathbf{w}_1 + x^2\mathbf{w}_2 + \dots + x^n\mathbf{w}_n \,. \end{aligned}$$

Em particular, os vetores $\mathbf{w}_1, \dots, \mathbf{w}_n$ geram $\mathrm{im}(L) \subset \mathbf{W}$. Se também $\mathbf{W}$ tem dimensão finita, e $\mathcal{C} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m)$ é uma base ordenada de $\mathbf{W} \approx \mathbb{R}^m$ (por exemplo, a base canónica de $\mathbb{R}^m$), e definimos os números $a^i{}_j$ como sendo as coordenadas dos $\mathbf{w}_j$' s relativamente à base $\mathcal{C}$, i.e.

$$L(\mathbf{b}_j) = \mathbf{w}_j = a^1{}_j\,\mathbf{c}_1 + a^2{}_j\,\mathbf{c}_2 + \dots + a^m{}_j\,\mathbf{c}_m$$

com $j = 1, 2, \dots, n$, então o valor de L sobre o vetor genérico $\mathbf{x} = x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \dots + x^n\mathbf{b}_n$ é

$$L(\mathbf{x}) = L\left(\sum_j x^j\mathbf{b}_j\right) = \sum_{j=1}^{m}\left(\sum_{i=1}^{n} a^i{}_j\, x^j\right)\mathbf{c}_i\,,$$

Portanto, as coordenadas do vetor $\mathbf{y} = L(\mathbf{x})$ são

$$\boxed{y^i = \sum_{j=1}^n a^i{}_j\, x^j \qquad\qquad i = 1, 2, \dots, m\,.}$$

Se X e Y denotam os "vetores coluna" (matrizes com apenas uma coluna!)

$$X := \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} \in \mathbb{R}^n \qquad \text{e} \qquad Y := \begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix} \in \mathbb{R}^m$$

e A denota a matriz

$$A := \begin{pmatrix} a^1{}_1 & a^1{}_2 & \dots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \dots & a^2{}_n \\ \vdots & \vdots & \ddots & \vdots \\ a^m{}_1 & a^m{}_2 & \dots & a^m{}_n \end{pmatrix} \in \mathrm{Mat}_{m\times n}(\mathbb{R}) \,,$$

então a transformação linear L é representada pela equação matricial

$$\boxed{Y = AX}$$

ou seja, explicitamente,

$$\begin{pmatrix} y^1 \\ y^2 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} a^1{}_1 & a^1{}_2 & \dots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \dots & a^2{}_n \\ \vdots & \vdots & \ddots & \vdots \\ a^m{}_1 & a^m{}_2 & \dots & a^m{}_n \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}$$

Se uma segunda matriz $B \in \mathrm{Mat}_{p\times m}(\mathbb{R})$ define a transformação linear $M : \mathbf{W} \to \mathbf{Z}$ relativamente às bases $\mathcal{C} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m)$ de $\mathbf{W} \approx \mathbb{R}^m$ e $\mathcal{D} = (\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_p)$ de $\mathbf{Z} \approx \mathbb{R}^p$, então a composição $ML : \mathbf{V} \to \mathbf{Z}$, que é também uma transformação linear, é definida pela matriz produto $BA \in \mathrm{Mat}_{p\times n}(\mathbb{R})$. De fato, se as coordenadas de $\mathbf{y} = L(\mathbf{x})$ são $y^k = \sum_j a^k{}_j \, x^j$ e as coordenadas de $\mathbf{z} = M(\mathbf{y})$ são $z^i = \sum_k b^i{}_k \, y^k$, então

$$z^i = \sum_{j,k} b^i{}_k \, a^k{}_j \, x^j \qquad i = 1, 2, \dots, p \,,$$

Em notaçþa matricial,

$$\boxed{\text{se} \quad Y = AX \quad \text{e} \quad Z = BY \quad \text{então} \quad Z = BAX}$$

- Determine a matriz que define a transformação

$$L(x, y) = (x - y, 2x - 3y) \qquad L(x, y, z) = (3x + y - z, -x + 2y + z)$$

$$L(x, y, z) = (3x, 3y, 3z) \qquad L(x, y) = (x + y, x - y, 2x - 7y)$$

$$L(x, y, z) = (x, y) \qquad L(x, y, z) = (x, z)$$

- Determine a transformação linear $L : \mathbb{R}^n \to \mathbb{R}^m$ definida pela matriz

$$\begin{pmatrix} 1 & 3 \\ -2 & 0 \end{pmatrix} \qquad \begin{pmatrix} -5 & 0 & -1 \\ 3 & 1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 0 & -1 \\ 0 & 8 \\ -1 & -3 \end{pmatrix} \qquad \begin{pmatrix} 1 & 3 \\ -2 & 0 \end{pmatrix}$$

- Determine a matriz 2×2 que define a transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que

transforma cada ponto no seu simétrico em relação à reta $x = 0$

transforma cada ponto no seu simétrico em relação à reta $y = -x$

transforma o ponto de coordenadas polares (r, θ) no ponto de coordenadas polares$(r/2, \theta)$

transforma o ponto de coordenadas polares (r, θ) no ponto de coordenadas polares $(r, \theta{-}\pi/2)$

- [Ap69] **16.12**.

4. (Einstein's notation) It is often important to take care of the distinction between vectors and co-vectors, and then different types of tensors, as well as to shorten formulas and computations. One possibility is the convention introduced by Einstein. We denote vectors using upper indices as $\mathbf{x} = (x^i) \in \mathbb{R}^n$, and denote co-vectors using lower indices as $\boldsymbol{\xi} = (\xi_j) \in (\mathbb{R}^n)^*$. The pairing between a co-vector $\boldsymbol{\xi}$ and a vector $\mathbf{x}$ reads $\langle \boldsymbol{\xi}, \mathbf{x} \rangle = \xi_1 x^1 + \xi_2 x^2 + \cdots + \xi_n x^n$, and is shortened as $\langle \boldsymbol{\xi}, \mathbf{x} \rangle = \xi_i x^i$, Einstein's convention being that a repeated index which appears once as an upper index and once as a lower index implies summation. Using Einstein's sum convention, the coordinates of the image of a linear transformation represented by the matrix $T = (t^i{}_j)$ are given by $y^i = t^i{}_j \, x^j$. The composition of $T = (t^i{}_j)$ followed by $S = (s^i{}_j)$ is then represented by the matrix $ST = (s^i{}_k \, t^k{}_j)$.

5. (endomorfismos e matrizes quadradas) Fixada uma base (por exemplo, a base canónica de $\mathbb{R}^n$), o espaço linear $\mathrm{End}(\mathbf{V}) := \mathcal{L}(\mathbf{V}, \mathbf{V})$ dos *endomorfismos* de um espaço linear $\mathbf{V} \approx \mathbb{R}^n$ é isomorfo ao espaço linear $\mathrm{Mat}_{n\times n}(\mathbb{R})$ das matrizes "quadradas" $n \times n$, os seus elementos sendo as transformações lineares

$$\boxed{X \mapsto Y = AX}$$

ou seja, $x^i \mapsto y^i = \sum_j a^i{}_j \, x^j$, com $X \in \mathbb{R}^n$ e $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$. O caso complexo é análogo.

À composição de dois endomorfismos corresponde o produto de matrizes. Em particular, a n-ésima iterada $L^n = L \circ L \circ \cdots \circ L$ (n vezes) do endomorfismo $L : X \mapsto AX$ é representada pela *potência* n-ésima de A, definida recursivamente por

$$A^0 = I_n \qquad \text{e} \qquad A^{n+1} = AA^n \qquad \text{se } n \geq 0\,.$$

A matriz quadrada A (ou o endomorfismo $L : X \mapsto AX$) é *nilpotente* se existe um inteiro $k \in \mathbb{N}$ tal que $A^k = 0$. A matriz quadrada A é *unipotente* se $A - I_n$ é nilpotente, e portanto existe um inteiro $k \in \mathbb{N}$ tal que $(A - I_n)^k = 0$.

A *diagonal* da matriz quadrada $A = (a^i{}_j) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é o conjunto ordenado dos elementos "diagonais" $a^1{}_1, a^2{}_2, \ldots, a^n{}_n$. O *traço* (em inglês, *trace*) de A é a soma dos elementos da diagonal,

$$\boxed{\mathrm{tr}(A) := \textstyle\sum_i a^i{}_i = a^1{}_1 + a^2{}_2 + \cdots + a^n{}_n}$$

É imediato verificar que $\mathrm{tr}(AB) = \mathrm{tr}(BA)$.

Uma matriz quadrada é uma *matriz diagonal* se os elementos que não pertencem à diagonal são nulos, ou seja, se é da forma

$$\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) := \begin{pmatrix} \lambda_1 & 0 & 0 & \ldots \\ 0 & \lambda_2 & 0 & \ldots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

- Determine a matriz da transformação "identidade" $\mathbf{x} \mapsto \mathbf{x}$ e da transformação "nula" $\mathbf{x} \mapsto \mathbf{0}$.
- Determine a matriz da homotetia $\mathbf{x} \mapsto \lambda\mathbf{x}$, com $\lambda \in \mathbb{R}$, e calcule o seu traço.
- Mostre que
$$\mathrm{tr}(AB) = \mathrm{tr}(BA)$$
- Calcule A^0, A^1, A^2, A^3, ... , A^n , ... quando
$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$
$$A = \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

- Determine as matrizes $A \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$ tais que $A^2 = 0$.
- Determine as matrizes $A \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$ tais que $A^2 = I_2$.

6. (comutador) A composição de transformações lineares, e portanto o produto de matrizes, não são comutativos! Ou seja, em geral, $AB \neq BA$. As matrizes quadradas $A, B \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ (e portanto os endomorfismos que representam), *comutam entre si/são permutáveis* se

$$AB = BA\,.$$

A obstrução é o *comutador*, definido por

$$\boxed{[A,B] := AB - BA}$$

O comutador satisfaz a *identidade de Jacobi*

$$\boxed{[[A,B],C] + [[B,C],A] + [[C,A],B] = 0}$$

- Mostre que cada matriz quadrada A comuta com si própria, i.e. $[A,A] = 0$.
- Mostre que duas matrizes diagonais comutam.
- Verifique se A e B comutam quando

$$A = \begin{pmatrix} 1 & 3 & 2 \\ -2 & 0 & 1 \\ 5 & 6 & 7 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} -7 & -1 & 3 \\ 2 & 1 & -1 \\ 0 & -2 & 3 \end{pmatrix}$$

- Considere as matrizes 2×2

$$E = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad E_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad E_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

  Calcule $[E, E_+]$, $[E, E_-]$ e $[E_+, E_-]$.

7. (affine transformations) The affine transformation of the plane $\mathbf{r} \mapsto \mathbf{r}' = A\mathbf{r} + \mathbf{a}$, where $A = (a_{ij}) \in \mathrm{Mat}_{2\times 2}(\mathbb{R})$ and $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ may be represented with the aid of a 3×3 matrix as

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_1 \\ a_{21} & a_{22} & a_2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

8. (shearing) An *horizontal shearing* of the plane is a linear transformation

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

Similarly, one defines a *vertical shearing* according to

$$\begin{pmatrix} x \\ y \end{pmatrix} \mapsto \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

- Try to figure out the effect of shearings on simple plane shapes.

9. (operadores e matrizes transpostas) O *operador transposto* do operador linear $L : \mathbb{R}^n \to \mathbb{R}^m$ é o operador linear $L^\top : (\mathbb{R}^m)^* \to (\mathbb{R}^n)^*$ definido por

$$\boxed{\langle L^\top \boldsymbol{\xi}, \mathbf{x} \rangle := \langle \boldsymbol{\xi}, L\mathbf{x} \rangle}$$

se $\mathbf{x} \in \mathbb{R}^n$ e $\boldsymbol{\xi} \in (\mathbb{R}^m)^*$. Se $A = (a^i{}_j) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ é a matriz que representa L relativamente às bases canónicas de $\mathbb{R}^n$ e $\mathbb{R}^m$, então a matriz que representa $L^\top \in L((\mathbb{R}^m)^*, (\mathbb{R}^n)^*) \approx L(\mathbb{R}^m, \mathbb{R}^n)$ relativamente às bases duais é a *matriz transposta* $A^\top = (b_i{}^j) \in \mathrm{Mat}_{n\times m}(\mathbb{R})$, definida por ${b_i}^j = {a^j}_i$ (ou seja, as linhas de $A^\top$ são as colunas de A e vice-versa). De fato, se $\mathbf{y} = L(\mathbf{x})$ tem coordenadas $y^i = \sum_j {a^i}_j \, x^j$, então

$$\left\langle A^\top \boldsymbol{\xi}, \mathbf{x} \right\rangle = \langle \boldsymbol{\xi}, A\mathbf{x} \rangle = \sum_i \xi_i \left(\sum_j {a^i}_j \, x^j \right) = \sum_j \left(\sum_i {a^i}_j \xi_i \right) x^j \, .$$

Uma matriz quadrada $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é *simétrica* se $A = A^\top$ e *anti-simétrica* se $A^\top = -A$.

- Verifique que $(A^\top)^\top = A$.
- Mostre que $\mathrm{tr}(A^\top) = \mathrm{tr}(A)$.
- Calcule o operador transposto de $L(x, y) = (x + y, x - y)$.
- Mostre que, se A é uma matriz quadrada, então $A + A^\top$ é simétrica, e $A - A^\top$ é anti-simétrica. Deduza que cada matriz quadrada pode ser decomposta como soma $A = A_\mathrm{s} + A_\mathrm{a}$ de uma matriz simétrica com uma matriz anti-simétrica.
- Mostre que o traço de uma matriz (quadrada) anti-simétrica é nulo.

18 Automorfismos e matrizes invertíveis

ref: [Ap69] *Vol 2*, **2.19-21**

1. (transformações lineares invertíveis) A transformação linear $L : \mathbf{V} \to \mathbf{W}$ é *invertível* se é *biunívoca* (i.e. a equação $L(\mathbf{x}) = \mathbf{y}$ admite uma e uma única solução para cada $\mathbf{y} \in \mathrm{im}(L)$), $\Leftrightarrow$ existe uma transformação linear $L^{-1} : \mathrm{im}(L) \to \mathbf{V}$ (a *inversa* de L) tal que $L^{-1}L = I_{\mathbf{V}}$ e $LL^{-1} = I_{\mathrm{im}(L)}$ $\Leftrightarrow$ o núcleo de L é trivial, i.e. $\ker(L) = \{\mathbf{0}\}$.

   Em particular, se $\mathbf{V}$ tem dimensão finita, a transformação linear $L : \mathbf{V} \to \mathbf{W}$ é invertível $\Leftrightarrow$ $\dim_{\mathbb{R}} \mathrm{im}(L) = \dim_{\mathbb{R}} \mathbf{V}$ $\Leftrightarrow$ L transforma vetores independentes de $\mathbf{V}$ em vetores independentes de $\mathbf{W}$ $\Leftrightarrow$ L transforma bases de $\mathbf{V}$ em bases de $\mathrm{im}(L)$.

   Os *automorfismos* $\mathrm{Aut}(\mathbf{V})$ de um espaço linear $\mathbf{V}$ são os endomorfismos $L : \mathbf{V} \to \mathbf{V}$ invertíveis.

   - Diga se L é biunívoca e, caso afirmativo, determine a imagem $\mathrm{im}(L)$ e a transformação inversa L^{-1}.
   $$L(x,y) = (x,x) \qquad L(x,y) = (y,x)$$
   $$L(x,y) = (x-y, x+y) \qquad L(x,y) = (0,y)$$
   $$L(x,y,z) = (x+y, y+z, z+x) \qquad L(x,y,z) = (3x, 2y, z)$$
   $$L(x,y,z) = (y,z,0) \qquad L(x,y,z) = (x+y+z, y, z)$$
   $$L(x,y) = (x,0,y) \qquad L(x,y) = (x-x, x+y, 0)$$
   - [Ap69] **16.8**.
   - Mostre que se $L : \mathbf{V} \to \mathbf{V}$ é invertível então também L^n é invertível e $(L^n)^{-1} = (L^{-1})^n$.
   - Mostre que se $L : \mathbf{V} \to \mathbf{V}$ e $M : \mathbf{V} \to \mathbf{V}$ comutam, então também as inversas L^{-1} e M^{-1} comutam, e $(LM)^n = L^n M^n$.
   - Considere o *operador derivação* $D : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$, definido por $f(t) \mapsto f'(t)$, e o *operador integração* $I : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$, definido por $f(t) \mapsto \int_0^t f(s)ds$. Mostre que $DI = I_{\mathcal{P}(\mathbb{R})}$ mas $ID \neq I_{\mathcal{P}(\mathbb{R})}$. Descreva o núcleo e a imagem de ID.
   - O operador $T : \mathcal{P}(\mathbb{R}) \to \mathcal{P}(\mathbb{R})$, definido por $f(t) \mapsto t \cdot f(t)$, é invertível?
   - O operador *deslocamento* $\sigma : \mathbb{R}^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$, definido por
   $$(x_1, x_2, x_3, \dots) \mapsto (x_2, x_3, x_4, \dots)\,,$$
   é invertível?

2. (inversão de matrizes 2×2) A matriz 2×2
$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
representa o endomorfismo genérico do plano $L(x,y) = (ax+by, cx+dy)$. A transformação L é invertível se para cada vetor $(\alpha, \beta) \in \mathbb{R}^2$ é possível encontrar um vetor $(x,y) \in \mathbb{R}^2$ tal que $L(x,y) = (\alpha, \beta)$, ou seja, resolver o sistema linear
$$\left\{ \begin{array}{rcl} ax+by & = & \alpha \\ cx+dy & = & \beta \end{array} \right. \qquad \Rightarrow \qquad \left\{ \begin{array}{rcl} (ad-bc)\,x & = & d\alpha - c\beta \\ (ad-bc)\,y & = & a\beta - c\alpha \end{array} \right.$$
(o segundo sistema é obtido ao retirar b vezes a segunda equação de d vezes a primeira equação, e depois ao retirar c vezes a primeira equação de a vezes a segunda equação). Portanto, a transformação L é invertível sse $\det A := ad - bc \neq 0$, e a sua inversa é a transformação linear
$$L^{-1}(\alpha, \beta) = \frac{1}{ad-bc}(d\alpha - c\beta, a\beta - c\alpha)\,,$$
representada pela matriz
$$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

3. (automorfismos e matrizes invertíveis) A matriz quadrada $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é *invertível* (ou *não-singular*, ou *regular*) se existe uma matriz quadrada $A^{-1} \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, dita *inversa* de A, tal que

$$\boxed{A^{-1}A = AA^{-1} = I_n}$$

A transformação linear $L : \mathbb{R}^n \to \mathbb{R}^n$, representada pela equação matricial $X \mapsto Y = AX$, é invertível sse a matriz A é invertível, e a sua inversa é a transformação linear $Y \mapsto X = A^{-1}Y$. Se $A = (a^i{}_j)$, então as entradas da inversa $A^{-1} = (b^i{}_j)$ satisfazem as n^2 equações lineares

$$\sum_k b^i{}_k \, a^k{}_j = \delta^i{}_j$$

Se A e B são invertíveis, então também AB é invertível e a sua inversa é

$$\boxed{(AB)^{-1} = B^{-1}A^{-1}}$$

Se A é invertível então também $A^\top$ é invertível e

$$\boxed{(A^\top)^{-1} = (A^{-1})^\top}$$

- Mostre que, se $A, B \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, então $BA = I_n \Rightarrow AB = I_n$ (ou seja, uma inversa esquerda é também uma inversa direita, logo uma inversa).
- Diga se as seguintes matrizes são invertíveis e, caso afirmativo, calcule a inversa.

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad \begin{pmatrix} 1 & -3 \\ -2 & 6 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 3 & 2 & 0 \\ 5 & 4 & 3 \end{pmatrix} \quad \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ a & 0 & 1 \end{pmatrix}$$

- [Ap69] **16.20**.
- [La87] ... $A^2 = I_n$ é invertível ...

4. (mudança de bases/coordenadas) Sejam $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$ e $\mathcal{B}' = (\mathbf{b}'_1, \mathbf{b}'_2, \dots, \mathbf{b}'_n)$ duas bases do espaço vetorial $\mathbf{V} \approx \mathbb{R}^n$. Então existe uma matriz invertível $U = (u^i{}_j) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, com inversa $U^{-1} = (v^i{}_j)$, tal que

$$\mathbf{b}'_j = \sum_i u^i{}_j \, \mathbf{b}_i \qquad \text{e} \qquad \mathbf{b}_j = \sum_i v^i{}_j \, \mathbf{b}'_i \, .$$

Se (x^i) são as coordenadas do vetor $\mathbf{x} \in \mathbf{V}$ relativamente à base $\mathcal{B}$, então as coordenadas do vetor $\mathbf{x}$ relativamente à base $\mathcal{B}'$ são

$$\boxed{x'^i = \textstyle\sum_k v^i{}_j \, x^j \qquad \text{ou seja} \qquad x^i = \sum_j u^i{}_j \, x'^j}$$

pois $\mathbf{x} = \sum_j x^j \mathbf{b}_j = \sum_{j,k} v^i{}_j x^j \mathbf{b}'_i$. A matriz U, cujas colunas são as coordenadas dos vetores da base $\mathcal{B}'$ relativamente à base $\mathcal{B}$, é a matriz que realiza a mudança de coordenadas. As matrizes U e U^{-1} podem ser pensadas como matrizes das derivadas parciais, pois $u^i{}_j = \partial x^i / \partial x'^j$ e $v^i{}_j = \partial x'^i / \partial x^j$. Nesta notação, as fórmulas para a mudança de coordenadas parecem tautológicas (o que explica o valor da notação de Leibniz para as derivadas):

$$x'^i = \sum_j \frac{\partial x'^i}{\partial x^j} \, x^j \qquad \text{e} \qquad x^i = \sum_j \frac{\partial x^i}{\partial x'^j} \, x'^j \, .$$

Na notação matricial, se X e X' são os vetores coluna de coordenadas x_i's e x_i''s respetivamente, a mudança de coordenadas assume a forma

$$X = UX' \qquad \text{ou seja,} \qquad X' = U^{-1}X\,.$$

Sejam $\mathcal{C} = (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_m)$ e $\mathcal{C}' = (\mathbf{c}_1', \mathbf{c}_2', \dots, \mathbf{c}_m')$ duas bases do espaço vetorial $\mathbf{W} \approx \mathbb{R}^m$, e seja $S = ({s^i}_j) \in \mathrm{Mat}_{m\times m}(\mathbb{R})$, com inversa $S^{-1} = (t^i_j)$, a matriz que realiza a mudança de coordenadas, ou seja, $\mathbf{c}_j' = \sum_i {s^i}_j\,\mathbf{c}_i$ e $\mathbf{c}_j = \sum_i {t^i}_j\,\mathbf{c}_i'$. Se $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ é a matriz da transformação L relativamente às bases $\mathcal{B}$ e $\mathcal{C}$, então a matriz da transformação L relativamente às bases $\mathcal{B}'$ e $\mathcal{C}'$ é

$$\boxed{A' = S^{-1}AU}$$

De fato, se $y^i = \sum_j {a^i}_j\,x^j$, então $y'^i = \sum_k {t^i}_k\,y^k = \sum_{k,\ell} {t^i}_k\,{a^k}_\ell\,x^\ell = \sum_{k,\ell,j} {t^i}_k\,{a^k}_\ell\,{u^\ell}_j\,x'^j$.

Em particular, se $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é a matriz do endomorfismo $L \in \mathrm{End}(\mathbf{V})$ relativamente à base $\mathcal{B}$, então a matriz de L relativamente à base $\mathcal{B}'$ é

$$\boxed{A' = U^{-1}AU}$$

e, de consequência,

$$\boxed{A = UA'U^{-1}\,.}$$

Matrizes A e A' relacionadas pelas identidades acima são ditas *semelhantes*.

- Mostre que $\sum_k {v^i}_k\,{u^k}_j = {\delta^i}_j$ e $\sum_k {u^i}_k\,{v^k}_j = {\delta^i}_j$.
- Verifique que se $A' = U^{-1}AU$ então $A = UA'U^{-1}$.
- Use $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ para mostrar que $\mathrm{tr}(U^{-1}AU) = \mathrm{tr}\,A$, ou seja, o traço de uma matriz quadrada apena depende da transformação linear definida pela matriz, e não da base usada.
- Determine a matriz de $L(x,y) = (3x, 2y)$ relativamente à base $\mathbf{b}_1 = (1,1)$ e $\mathbf{b}_2 = (1,-1)$.
- Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a reflexão na reta $y = x$. Determine a matriz de T relativamente à base canónica e relativamente à base $(1,1)$ e $(1,-1)$.
- Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a reflexão na reta $y = 3x$. Determine a matriz de T relativamente à base canónica.
- Seja $T : \mathbb{R}^2 \to \mathbb{R}^2$ a projecção ortogonal sobre a reta $x + y = 0$. Determine a matriz de T relativamente à base canónica.
- Determine a matriz que representa o operador derivação $D : \mathcal{P}_{\leq 3}(\mathbb{R}) \to \mathcal{P}_{\leq 2}(\mathbb{R})$, definido por $(Df)(t) := f'(t)$, relativamente às bases ordenadas $(1, t, t^2, t^3)$ e $(1, t, t^2)$. Determine umas bases de $\mathbf{V} = \mathcal{P}_{\leq 3}(\mathbb{R})$ e $\mathbf{W} = \mathcal{P}_{\leq 3}(\mathbb{R})$ tal que o operador $D : \mathbf{V} \to \mathbf{W}$ seja diagonal.

5. (rotações no plano e funções trigonométricas) Seja R_θ a matriz 2×2 que representa uma rotação $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$ de um ângulo θ no sentido anti-horário, relatvamente à base canónica do plano. Pela definição das funções trigonométricas,

$$R_\theta \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \qquad \text{e} \qquad R_\theta \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

e portanto, usando a linearidade,

$$R_\theta = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Ou seja, a rotação de um ângulo θ é o automorfismo do plano cartesiano definido por $R_\theta(x,y) = ((\cos\theta)x - (\sin(\theta)y, (\sin\theta)x + (\cos\theta)y)$.

- Observe que $R_\theta R_{\pm\phi} = R_{\theta\pm\phi}$ e deduza as fórmulas de adição para as funções trigonométricas

$$\begin{aligned} \cos(\theta \pm \phi) &= \cos(\theta)\cos(\phi) \mp \sin(\theta)\sin(\phi) \\ \sin(\theta \pm \phi) &= \sin(\theta)\cos(\phi) \pm \cos(\theta)\sin(\phi) \end{aligned}$$

- Observe que $R_\theta R_{-\theta}$ é a transformação "identidade", e calcule a matriz inversa de R_θ.

19 Valores e vetores próprios

ref: [Ap69] *Vol 2*, **4.1-10** ; [La87] *Ch. VIII*

1. (subespaços invariantes) Seja $L : \mathbf{D} \to \mathbf{V}$ um operador linear definido num subespaço $\mathbf{D} \subset \mathbf{V}$ do espaço linear $\mathbf{V}$, real ou complexo. Um subespaço linear $\mathbf{W} \subset \mathbf{D}$ é *invariante* (ou *estável*) se $L(\mathbf{W}) \subset \mathbf{W}$, ou seja, se $\mathbf{v} \in \mathbf{W}$ implica $L\mathbf{v} \in \mathbf{W}$. Subespaços invariantes triviais são o subespaço nulo $\{\mathbf{0}\}$, e o próprio $\mathbf{V}$ quando $\mathbf{D} = \mathbf{V}$.

   - Mostre que o núcleo $\ker(L)$ e a imagem $\mathrm{im}(L)$ (se $\mathbf{D} = \mathbf{V}$) são subespaços invariantes.
   - Determine os subespaços invariantes da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto (x, y) no seu simétrico em relação à reta $y = x$.
   - Determine os subespaços invariantes da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto (x, y) na sua projeção ortogonal sobre a reta $y = x$.
   - Determine os subespaços invariantes do "cisalhamento" (em inglês, *shear*) horizontal de fator $\mu \neq 0$, a transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ definida por $T(x, y) = (x + \mu y, y)$.

2. (valores e vetores próprios) Seja $L : \mathbf{D} \to \mathbf{V}$ um operador linear definido num subespaço $\mathbf{D} \subset \mathbf{V}$ do espaço linear $\mathbf{V}$, real ou complexo. Um *vetor próprio/autovetor* (em inglês *eigenvector*, do alemão *eigen* = próprio) de L é um vetor não nulo $\mathbf{v} \in \mathbf{D}$ que gera um subespaço invariante $\mathbb{R}\mathbf{v}$ (ou $\mathbb{C}\mathbf{v}$) de dimensão um para L, ou seja, pela linearidade, tal que

   $$\boxed{L\mathbf{v} = \lambda\mathbf{v}}$$

   onde $\lambda \in \mathbb{R}$ (ou $\lambda \in \mathbb{C}$) é um escalar, chamado *valor próprio/autovalor* (em inglês *eigenvalue*) do operador L (associado ao vetor próprio $\mathbf{v}$).

   Teorema 19.1. *Se $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ são vetores próprios de $L : \mathbf{D} \to \mathbf{V}$ e se os correspondentes valores próprios $\lambda_1, \lambda_2, \dots, \lambda_k$ são dois a dois distintos (i.e. $\lambda_i \neq \lambda_j$ se $i \neq j$), entaõ os vetores $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ são linearmente independentes.*

   Demonstração. A prova é por indução. O caso $k = 1$ é trivial. Seja $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_{k+1}\mathbf{v}_{k+1} = \mathbf{0}$. Aplicando o operador L ou multiplicando por λ_{k+1}, e depois calculando a diferença, obtemos

   $$c_1(\lambda_1 - \lambda_{k+1})\mathbf{v}_1 + c_2(\lambda_2 - \lambda_{k+1})\mathbf{v}_2 + \cdots + c_k(\lambda_k - \lambda_{k+1})\mathbf{v}_k = \mathbf{0}$$

   Pela hipótese indutiva todos os coeficientes são nulos, e portanto, sendo os valores próprios distintos, $c_1 = c_2 = \cdots = c_k = 0$. Então também $c_{k+1} = 0$ (pois $\mathbf{v}_{k+1} \neq \mathbf{0}$). □

   Se λ é um valor próprio de $L : \mathbf{V} \to \mathbf{V}$, o conjunto

   $$\mathbf{V}_\lambda := \{\mathbf{v} \in \mathbf{V} \text{ t.q. } L\mathbf{v} = \lambda\mathbf{v}\} = \ker(\lambda - L)$$

   ($\lambda - L$ denota o operador $\mathbf{v} \mapsto \lambda\mathbf{v} - L(\mathbf{v})$) é um subespaço invariante, diferente do espaço trivial $\{\mathbf{0}\}$, dito *subespaço próprio* associado ao valor próprio λ. A restrição do operador linear L a cada espaço próprio $\mathbf{V}_\lambda$ é uma homotetia $\mathbf{v} \mapsto \lambda\mathbf{v}$.

   Em particular, se existe uma decomposição de $\mathbf{V}$ como soma direta finita $\mathbf{V} = \oplus_k \mathbf{V}_k$ de espaços próprios $\mathbf{V}_k$ associados aos valores próprios distintos λ_k de L (ou seja, todo $\mathbf{v} \in \mathbf{V}$ é uma sobreposição única $\mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 + \cdots + \mathbf{v}_n$ de vetores $\mathbf{v}_k \in \mathbf{V}_k$ tais que $L\mathbf{v}_k = \lambda_k\mathbf{v}_k$), então o operador é uma soma direta $L = \oplus_k \lambda_k$ de homotetias. Em outras palavras, se $P_k : \mathbf{V} \to \mathbf{V}_k$ denota a projeção sobre $\mathbf{V}_k$ (definida por $P_k(\mathbf{v}) := \mathbf{v}_k$), então $L = \sum_k \lambda_k P_k$. Tais operadores são chamados "diagonalizáveis". De fato, se $\mathbf{V}$ tem dimensão finita, são representados por matrizes diagonais numas bases formadas por vetores próprios.

- Todo vetor $\mathbf{v} \neq \mathbf{0}$ é um vetor própro da trasformação identidade $I_{\mathbf{V}}\mathbf{v} := \mathbf{v}$, com valor próprio $\lambda = 1$. Todo vetor $\mathbf{v} \neq \mathbf{0}$ é um vetor própro da trasformação nula $0_{\mathbf{V}}\mathbf{v} := \mathbf{0}$, com valor próprio $\lambda = 0$. Em geral, todo vetor $\mathbf{v} \neq 0$ é um vetor próprio de uma homotetia $L\mathbf{v} = \lambda\mathbf{v}$, de valor próprio λ.
- Determine valores e vetores próprios da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto (x, y) no seu simétrico em relação à reta $y = 2x$.
- Determine valores e vetores próprios da transformação $T : \mathbb{R}^2 \to \mathbb{R}^2$ que transforma cada ponto (x, y) na sua projeção ortogonal sobre a reta $3y = x$.
- Uma rotação $R_\theta : \mathbb{R}^2 \to \mathbb{R}^2$, definida no espaço vetorial real $\mathbb{R}^2$ por

  $$(x, y) \mapsto (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta) \ ,$$

  não admite vetores próprios se o ângulo θ não é um múltiplo inteiro de π. No entanto, a mesma rotação pode ser pensada como a transformação $z \mapsto e^{i\theta}z$ definida no espaço vetorial complexo $\mathbb{C} \approx \mathbb{R}^2$, onde $z = x + iy \approx (x, y)$. Neste caso, todo vetor $z \neq 0$ é um vetor próprio, de valor próprio $e^{i\theta}$.
- Determine os valores e os vetores próprios das transformações

  $$L(x, y) = (x, 0) \qquad L(x, y) = (x/2, 3y) \qquad L(x, y) = (-y, x)$$

  $$L(x, y) = (x, x + y) \qquad L(x, y) = (x + \lambda y, y) \qquad L(x, y) = (x + \alpha y, y)$$

  $$L(x, y, z) = (0, y, -z) \qquad L(x, y, z) = (y, z, x)$$
- Se λ é um valor próprio de $L : \mathbf{V} \to \mathbf{V}$ e $k \in \mathbb{N}$, então λ^k é um valor próprio de L^k. No entanto, as potências L^k de um operador podem ter mais valores próprios que o próprio L. Por exemplo, uma rotação $R_{\pi/2}$ de um ângulo $\pi/2$ no plano não tem valores próprios, mas o seu quadrádo $R_{\pi/2}R_{\pi/2} = R_\pi$ admite um valor próprio, -1 (cuja raiz quadrada não é um número real!).
- Em particular, os valores próprios de um operador nilpotente (um operador tal que alguma potência L^n é o operador nulo) não podem ser diferentes de zero.
- Se $L : \mathbf{V} \to \mathbf{V}$ é um automorfismo (i.e. uma transformação linear invertível) e $\mathbf{v} \in \mathbf{V}$ é um vetor próprio de L, então $\mathbf{v}$ é um vetor próprio de L^{-1}
- [Ap69] **4.4**.

3. (resolvente e espetro) Se λ é um valor próprio do operador linear $L : \mathbf{V} \to \mathbf{V}$, então o operador

   $$L_\lambda := \lambda - L$$

   não é invertível (pois um autovetor $\mathbf{v}$ associado a λ está no seu núcleo). O *espetro* do operador L é o conjunto

   $$\sigma(L) := \{\lambda \in \mathbb{C} \ \text{t.q.} \ \ \lambda - L \ \text{ não é invertível } \} \subset \mathbb{C} \,,$$

   ou seja, o conjunto dos valores complexos de $z \in \mathbb{C}$ fora dos quais existe o operador *resolvente*

   $$R_z := (z - L)^{-1}$$

   O número $\rho(L) := \sup_{z\in\sigma(L)} |z|$ é dito *raio espetral* do operador L. Em particular, se λ é um valor próprio de L, então o seu módulo é limitado por $|\lambda| \leq \rho(L)$.

   Quando $\mathbf{V}$ tem dimensão finita, o espetro $\sigma(L)$ coincide com o conjunto (finito) dos valores próprios, pois se L_λ não é invertível então o seu núcleo $\ker(L_\lambda)$ não é vazio, e um vetor não nulo do núcleo de L_λ é por definição um vetor próprio de L.

   Em geral, em dimensão infinita, o espetro pode conter números que não são valores próprios.

   - O operador *deslocamento* (em ingês, *shift*)

     $$L : (x_1, x_2, x_3, \dots) \mapsto (x_2, x_3, x_4, \dots)$$

     definido no espaço das sucessões reais $\mathbf{V} = \mathbb{R}^{\mathbb{N}}$ não é invertível, portanto $0 \in \sigma(L)$. Determine o espaço próprio $\mathbf{V}_0$.

4. (operadores derivação e multiplicação) O operador *derivação* envia uma função derivável $f(x)$ na função
$$(\partial f)(x) := f'(x)\,.$$
Pode ser pensado como un endomorfismo de $\mathcal{C}^\infty(\mathbb{R})$, o espaço linear das funções infinitamente diferenciáveis definidas em $\mathbb{R}$ com valores reais ou complexos. O produto $P := -i\hbar\,\partial$, onde $\hbar \simeq 1.054\cdots \times 10^{-34}$ J $\cdot$ s é a constante de Planck reduzida, é o *operador momento* na "representação de Schrödinger" da mecânica quântica. O operador *multiplicação* envia $f(x)$ em
$$(Xf)(x) := x\, f(x)\,.$$
Observe que ∂ e X não comutam. De fato, $\partial X - X\partial$ é o operador identidade.

Todo $\lambda \in \mathbb{C}$ é um valor próprio do operador derivação, e o espaço próprio associado ao valor próprio λ é gerado pela função exponencial $e^{\lambda x}$. Se $\lambda_1, \dots, \lambda_n$ são distintos, então as funções $e^{\lambda_1 x}, \dots, e^{\lambda_n x}$ são linearmente independentes.

- Calcule valores e vetores próprios dos operadores ∂ e ∂^2 no espaço $\mathcal{C}^\infty(\mathbb{R})$ das funções reais infinitamente diferenciáveis.
- O subespaço $\mathcal{P}(\mathbb{R}) \subset \mathcal{C}^\infty(\mathbb{R})$ dos polinómios é um subespaço invariante de ∂. O subespaço $\mathcal{P}_{\leq n}(\mathbb{R}) \subset \mathcal{P}(\mathbb{R})$ dos polinómios de grau $\leq n$ é um subespaço invariante do operador derivação? E do operador multiplicação?
- Mostre que os vetores próprios do operador $L = X\partial$, definido em $\mathcal{P}(\mathbb{R})$, são os monómios $f(x) = x^n$.
- Fixada uma "frequência" λ, real ou complexa, o espaço dos quase-polinómios $p(x)e^{\lambda x}$, com $p \in \mathcal{P}(\mathbb{R})$, é um subespaço invariante para qualquer operador diferencial com coeficientes constantes $L = a_n\partial^n + \cdots + a_1\partial + a_0$. Esta observação justifica o método dos coeficientes indeterminados para encontrar uma solução particular de uma EDO linear $Lf = g$ quando a "força" $g(x)$ é um quase-polinómio.

5. (operador primitivação) O operador *primitivação* envia uma função integrável $f(x)$ numa das suas primitivas $F(x)$, por exemplo a função
$$(Pf)(x) := \int_0^x f(t)\,dt$$
Pode ser pensado como um endomorfismo do espaço linear real $\mathcal{C}^\infty(\mathbb{R})$ das fuções infinitamente diferenciáveis.

- Mostre que o operador $P : \mathcal{C}^\infty(\mathbb{R}) \to \mathcal{C}^\infty(\mathbb{R})$ não tem valores próprios (derive a identidade $Pf = \lambda f$ para obter uma equação diferencial para o suposto vetor próprio $f(x)$, e observe que a mesma identidade também implica uma condição inicial $f(0)\ \dots$).
- Mostre que ∂P é o operador identidade. Calcule o comutador $\partial P - P\partial$.

6. (operadores diferenciais, translações e ondas planas) Sejam $\Omega \subset \mathbb{R}^n$ um aberto e $\mathbf{V} = \mathcal{C}^\infty(\Omega)$ o espaço vetorial complexo das funções $f : \Omega \to \mathbb{C}$ infinitamente diferenciáveis. Dado um multi-índice $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$, de grau $|\alpha| := \alpha_1 + \alpha_2 + \cdots + \alpha_n$, o operador diferencial $\partial^\alpha : \mathbf{V} \to \mathbf{V}$ é definido por
$$(\partial^\alpha f)(\mathbf{x}) := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}(\mathbf{x})\,.$$

- Verifique que as ondas planas $e_{\boldsymbol{\xi}}(\mathbf{x}) := e^{i\boldsymbol{\xi}\cdot\mathbf{x}}$, com $\boldsymbol{\xi} \in (\mathbb{R}^n)^* \approx \mathbb{R}^n$, são funções próprias dos operadores diferenciais ∂^α, com $\alpha \in \mathbb{N}^n$, com valores próprios $(i\boldsymbol{\xi})^\alpha := (i\xi_1)^{\alpha_1}(i\xi_2)^{\alpha_2}\dots(i\xi_n)^{\alpha_n}$, ou seja,
$$\partial^\alpha e_{\boldsymbol{\xi}} = (i\boldsymbol{\xi})^\alpha e_{\boldsymbol{\xi}}\,.$$

- O operador de *translação* $T_{\mathbf{a}}$, com $\mathbf{a} \in \mathbb{R}^n$, é definido por

$$(T_{\mathbf{a}} f)(\mathbf{x}) := f(\mathbf{x} + \mathbf{a}) .$$

Verifique que as ondas planas $e_{\boldsymbol{\xi}}(\mathbf{x}) = e^{i\boldsymbol{\xi} \cdot \mathbf{x}}$ são também funções próprias dos operadores de translação com valores próprios $\lambda_{\mathbf{a}}(\boldsymbol{\xi}) = e^{i\boldsymbol{\xi} \cdot \mathbf{a}}$, ou seja,

$$T_{\mathbf{a}} e_{\boldsymbol{\xi}} = e^{i\boldsymbol{\xi} \cdot \mathbf{a}} e_{\boldsymbol{\xi}} .$$

- O operador de *modulação* $M_{\boldsymbol{\xi}}$, com $\boldsymbol{\xi} \in (\mathbb{R}^n)^* \approx \mathbb{R}^n$, é definido por

$$(M_{\boldsymbol{\xi}} f)(\mathbf{x}) := e^{i\boldsymbol{\xi} \cdot \mathbf{x}} f(\mathbf{x}) .$$

Mostre que $T_{\mathbf{a}} \circ M_{\boldsymbol{\xi}} = e^{i\boldsymbol{\xi} \cdot \mathbf{a}} M_{\boldsymbol{\xi}} \circ T_{\mathbf{a}}$. Os operadores translação e modulação geram o *grupo de Heisenberg*.

7. (PDEs e símbolo de um operador diferencial) Um *operador diferencial linear* de grau $\leq k$, definido num domiínio $\Omega \subset \mathbb{R}^n$, é um polinómio

$$P(\partial, \mathbf{x}) = \sum_{|\alpha| \leq k} a_\alpha(\mathbf{x}) \, \partial^\alpha ,$$

em ∂, com "coeficientes" $a_\alpha(\mathbf{x})$ que são funções $a_\alpha : \Omega \to \mathbb{R}$. O operador comuta com as translações de $\mathbb{R}^n$, os operadores $T_{\mathbf{a}}$, se os coeficientes a_α não dependem do ponto $\mathbf{x}$, i.e. são constantes. As "ondas planas" $e_{\boldsymbol{\xi}}(\mathbf{x}) := e^{i\boldsymbol{\xi} \cdot \mathbf{x}}$ são funções próprias do operador linear com coeficientes constantes $L = P(\partial)$, ou seja, satisfazem

$$L e_{\boldsymbol{\xi}} = \sigma(\boldsymbol{\xi}) \, e_{\boldsymbol{\xi}} ,$$

onde o valor próprio, o polinómio $\sigma(\boldsymbol{\xi}) := P(i\boldsymbol{\xi})$, é dito *símbolo* do operador L. O *símbolo principal* é o termo (polinómio homogéneo) de grau máximo, $\sigma_p(\boldsymbol{\xi}) = \sum_{|\alpha|=k} a_\alpha \cdot (i\boldsymbol{\xi})^\alpha$.

8. (polinómio caraterístico) Se λ é um valor próprio do operador linear $L : \mathbf{V} \to \mathbf{V}$, então o operador $L_\lambda := \lambda - L$ não é invertível, e os vetores próprios com valor próprio λ são os vetores não triviais do núcleo de L_λ. Se $\mathbf{V} \approx \mathbb{R}^n$ ou $\mathbb{C}^n$ tem dimensão finita, então, fixada uma base, o operador L é $X \mapsto AX$, onde $A \in \mathrm{Mat}_{n \times n}(\mathbb{R})$ ou $\mathrm{Mat}_{n \times n}(\mathbb{C})$ é uma matriz quadrada (que depende da base escolhida). O escalar $\lambda \in \mathbb{R}$ ou $\mathbb{C}$ é um valor próprio da matriz quadrada A (ou do operador L) sse a matriz $\lambda I - A$ é singular, ou seja, sse

$$\det(\lambda I - A) = 0 .$$

O *polinómio caraterístico* da matriz quadrada $A \in \mathrm{Mat}_{n \times n}(\mathbb{R})$ é

$$\boxed{P_A(t) := \det(tI - A)}$$

e portanto

Teorema 19.2. *O escalar λ é um valor próprio d(o operador definido pel)a matriz A sse é uma raiz do polinómio caraterístico, i.e. se $P_A(\lambda) = 0$.*

O espaço próprio associado ao valor próprio λ é $\mathbf{V}_\lambda = \ker(\lambda - L)$. Em particular, os vetores próprios associados ao valor próprio λ são as soluções não triviais do sistema homogéneo $(\lambda I - A)V = 0$.

- Mostre que A e $A^\top$ têm o mesmo polinómio caraterístico.
- Verifique que o polinómio caraterístico da matriz 2×2

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

é $P_A(t) = t^2 - (a+d)t + (ad - bc)$. Verifique que $P_A(A) = 0$, ou seja, que $A^2 - (\mathrm{tr} A)A + (\det A)I_2 = 0$.

- Os valores próprios de uma matriz nilpotente (uma matriz A tal que alguma potência $A^n = 0$) não podem ser diferentes de zero.
- A matriz quadrada A é unipotente (ou seja, $A^n = I$ para algum $n \geq 1$) sse o seu polinómio caraterístico é uma potência de $t-1$.
- Determine valores e vetores próprios dos endomorfismos definidos pela seguintes matrizes

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 1/3 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \qquad A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

$$A = \begin{pmatrix} 2 & 5 & -1 \\ 0 & -3 & 1 \\ 0 & 0 & 7 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 5 & -1 \\ 0 & -2 & 1 \\ -4 & 0 & 3 \end{pmatrix}$$

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 2 & 3 & 2 \\ 3 & 3 & 4 \end{pmatrix} \qquad A = \begin{pmatrix} 7 & 5 & 1 \\ 0 & -2 & 1 \\ 20 & 0 & 3 \end{pmatrix} \qquad A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}$$

- [Ap69] **4.10**.

9. (matrizes semelhantes e diagonalização) As matrizes quadradas $A, B \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ são ditas *semelhantes* se existe uma matriz invertível $U \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ tal que

$$B = U^{-1}AU\,,$$

ou seja, se representam a mesma transformação linear $L : \mathbb{R}^n \to \mathbb{R}^n$ em bases que podem ser diferentes. Se A e B são semelhantes então $\det A = \det B$. Matrizes semelhantes têm o mesmo polinómio caraterístico, pois

$$\begin{aligned} \det\left(tI_n - U^{-1}AU\right) &= \det\left(tU^{-1}I_nU - U^{-1}AU\right) \\ &= \det\left(U^{-1}(tI_n - A)U\right) = \det\left(tI_n - A\right) \end{aligned}$$

e portanto os mesmos valores próprios.

A matriz quadrada $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é *diagonalizável* se é semelhante a uma matriz diagonal. Se (o operador linear $L : \mathbb{R}^n \to \mathbb{R}^n$ definido na base canónica pel)a matriz quadrada $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ admite n vetores próprios linearmente independentes $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, com valores próprios $\lambda_1, \lambda_2, \dots, \lambda_n$, respetivamente (não necessariamente distintos), e se $U \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é a matriz (invertível, pela independência dos $\mathbf{v}_k$'s) cujas colunas são os vetores $\mathbf{v}_k$, então $U^{-1}AU$ é a matriz diagonal

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}.$$

Se $\lambda_1, \lambda_2, \dots, \lambda_n$ são as raizes (em geral, complexas e distintas) do polinómio caraterístico da matriz quadrada $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$, então

$$\begin{aligned} P_A(t) &= (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n) \\ &= t^n - (\lambda_1 + \lambda_2 + \dots + \lambda_n)\,t^{n-1} + \dots + (-1)^n(\lambda_1\lambda_2\cdots\lambda_n) \end{aligned}$$

Mas $P_A(0) = \det A$, e portanto

$$\boxed{\det A = \lambda_1\lambda_2\cdots\lambda_n}$$

Também acontece que

$$\boxed{\mathrm{tr} A = \lambda_1 + \lambda_2 + \dots + \lambda_n}$$

(pois $\mathrm{tr}(AB) = \mathrm{tr}(BA)$, e portanto $\mathrm{tr}(\Lambda) = \mathrm{tr}(U^{-1}AU) = \mathrm{tr}(A)$)

- Mostre que se A e B são semelhantes então $\det A = \det B$.
- As matrizes
$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \qquad \text{e} \qquad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
são semelhantes?
- Diagonalize as seguintes matrizes, ou mostre que não é possível.
$$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 1 & 3 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 2 & 1 \\ -1 & 4 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

10. (diagonalização simultânea) Sejam $L : \mathbf{V} \to \mathbf{V}$ e $M : \mathbf{V} \to \mathbf{V}$ dois operadores lineares que comutam (i.e. $[L, M] = 0$) definidos no espaço linear $\mathbf{V}$. Se λ é um valor próprio de L, então o espaço próprio $\mathbf{V}_\lambda = \{\mathbf{v} \in \mathbf{V} \text{ t.q. } L\mathbf{v} = \lambda\mathbf{v}\}$ é um subespaço invariante para o operador M, i.e. $M(\mathbf{V}_\lambda) \subset \mathbf{V}_\lambda$. De fato, se $\mathbf{v} \in \mathbf{V}_\lambda$, então
$$LM\mathbf{v} = ML\mathbf{v} = \lambda M\mathbf{v}$$
e portanto $M\mathbf{v} \in \mathbf{V}_\lambda$. ...

11. (gato de Arnold) Considere a transformação $L : \mathbb{R}^2 \to \mathbb{R}^2$ definida pela matriz
$$A = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$
ou seja, $L(x, y) = (2x + y, x + y)$. O polinómio caraterístico é $P_A(t) = t^2 - 3t - 5$, e portanto os valores próprios são $\lambda_\pm = (3 \pm \sqrt{5})/2$. Vetores próprios correspondentes, que satisfazem $L(\mathbf{v}_\pm) = \lambda_\pm \mathbf{v}_\pm$, são (por exemplo)
$$\mathbf{v}_+ = (\varphi, 1) \qquad \text{e} \qquad \mathbf{v}_- = (1, -\varphi)$$
onde $\varphi = (1+\sqrt{5})/2 \simeq 1.6180339887\dots$ é a "razão" dos gregos. A matriz que representa L na base $(\mathbf{v}_+, \mathbf{v}_-)$ é a matriz diagonal $A' = \mathrm{diag}(\lambda_+, \lambda_-)$. Observem que $\lambda_+ > 1$ e $0 < \lambda_- < 1$, e portanto L estica os vetores da reta $\mathbb{R}\mathbf{v}_+$ e contrae os vetores da reta $\mathbb{R}\mathbf{v}_-$. Observem também que $\det A = \lambda_+ \lambda_- = 1$. Em particular, L preserva as áreas, ou seja, a área de uma região (mensurável!) $R \subset \mathbb{R}^2$ é igual à área da sua imagem $L(R)$. A matriz A é invertível, e a sua inversa A^{-1} tem entradas inteira (pois $\det A = 1$). Em particular, L preserva o subgrupo $\mathbb{Z}^2 \subset \mathbb{R}^2$, ou seja, $L(\mathbb{Z}^2) = \mathbb{Z}^2$. Isto implica que L define uma transformação invertível $T : \mathbb{T}^2 \to \mathbb{T}^2$ do "toro" $\mathbb{T}^2 := \mathbb{R}^2/\mathbb{Z}^2$, definida por
$$(x, y) + \mathbb{Z}^2 \mapsto (2x + y, x + y) + \mathbb{Z}^2$$
A "dinâmica" da transfomação T, ou seja, o comportamento das trajetórias $\mathbf{v} \mapsto T(\mathbf{v}) \mapsto T(T(\mathbf{v})) \mapsto \dots$, é particularmente interessante, e T é o arquétipo de uma classe importante de transformações, ditas "hiperbólicas".

12. (Cayley-Hamilton theorem) If $p \in \mathbb{C}[t]$ is a polynomial in the indeterminate t, for example $p(t) = a_k t^k + \dots + a_1 t + a_0$, and A an $n \times n$ matrix, real or complex, one defines the matrix $p(A)$ as
$$p(A) := a_k A^k + \dots + a_1 A + a_0 I_n \, .$$
For example, one may consider the characteristic polynomial $P_A(t)$ of a square matrix A, and try to compute $P_A(A)$.

If $\Lambda = \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$ is a diagonal matrix, a simple computation (using the fact that its powers are also diagonal with entries which are powers of the λ_i's) shows that $P_\Lambda(\Lambda) = 0$. A diagonalizable matrix as $D = U^{-1}\Lambda U$ has its powers of the form $D^k = U^{-1}\Lambda^k U$, and consequently $P_D(D) = U^{-1}P_\Lambda(\Lambda)U = 0$, since $P_D = P_\Lambda$.

The set of diagonalizable complex matrices is dense in the space $\mathrm{Mat}_{n\times n}(\mathbb{C}) \approx \mathbb{C}^{n^2}$. Indeed, it is the set of those matrices such that the characteristic polynomial $P_A(t)$ has n distinct complex roots. In particular, for any $n \times n$ matrix A, real or complex, we may find a sequence $(D_m)_{m\in\mathbb{N}}$ of diagonalizable complex matrices such that $D_m \to A$ (w.r.t. some compatible norm). By continuity we get

Teorema 19.3 (Cayley-Hamilton)**.** *Any square matrix A satisfies its characteristic equation, i.e.*

$$\boxed{P_A(A) = 0}$$

- If A is an invertible $n \times n$ matrix, then we may multiply the identity $P_A(A) = 0$ by A^{-1}, and obtain a formula for the inverse matrix A^{-1} as a function of the powers A^0, A, A^2, ..., A^{n-1}. For example, show that the inverse of an invertible 2×2 matrix A is

$$A^{-1} = \frac{1}{\det A}\left((\mathrm{tr}A)I_2 - A\right) .$$

20 Operadores hermíticos e unitários

ref: [Ap69] *Vol 2*, **5.1-5**, **5.7-8**, **5.10**, **5.19** ; [La87] *Ch. VII*

1. (representação matricial de operadores em espaços euclidianos de dimensão finita) Seja $\mathbf{H}$ um espaço euclidiano, real ou complexo, de dimensão finita. Fixada uma base ortonormada $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$, podemos representar, cada vetor $\mathbf{x} \in \mathbf{H} \approx \mathbb{C}^n$ ou $\mathbb{R}^n$ de maneira única como combinação linear $\mathbf{x} = \sum_{i=1}^n x^i \mathbf{e}_i$, onde as "coordenadas" de $\mathbf{x}$ na base escolhida são os coeficientes de Fourier $x^i := \langle \mathbf{x}, \mathbf{e}_i \rangle$. O vetor coluna que representa $\mathbf{x}$ é então a matriz

$$X = \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}$$

O produto interno entre os vetores $\mathbf{x}$ e $\mathbf{y}$ do espaço euclidiano complexo $\mathbf{H} \approx \mathbb{C}^n$, definido por

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_i \sum_j x^i \delta_{ij} \overline{y^j} = \sum_i x \overline{y^i}$$

(onde $\delta_{ij} = 0$ se $i \neq j$ e $\delta_{ii} = 1$) pode ser representado como o produto linhas por colunas

$$\begin{pmatrix} \overline{y^1} & \overline{y^2} & \dots & \overline{y^n} \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}$$

do transposto conjugado do vetor coluna Y que representa $\mathbf{y}$ (que é portanto um vetor linha!) vezes o vetor coluna X que representa $\mathbf{x}$, ou seja, , em notação matricial,

$$\boxed{\langle \mathbf{x}, \mathbf{y} \rangle = \overline{Y}^\top X}$$

(esta troca entre "esquerda" e "direita" é uma consequência da tradição de definir os produtos AB entre duas matrizes A e B como sendo produtos "linhas pos colunas", (o que corresponde à interpretação do produto em quanto composição das transformações lineares definidas pelas matrizes), e não "colunas por linhas", e da tradição de pedir que os produtos internos $\langle \mathbf{x}, \mathbf{y} \rangle$ sejam lineares na variável à esquerda $\mathbf{x}$ e conjugados-lineares na variável à direita $\mathbf{y}$). Se $\mathbf{H} \approx \mathbb{R}^n$ é um espaço euclidiano real, então o produto interno é simplesmente $\langle \mathbf{x}, \mathbf{y} \rangle = Y^\top X$.

Seja $T : \mathbf{H} \to \mathbf{H}$ um operador linear. A imagem de cada $\mathbf{e}_j$, com $1 \leq j \leq n$, é uma combinação linear $T(\mathbf{e}_j) = \sum_i a^i{}_j \, \mathbf{e}_i$ dos elementos da base, onde os coeficientes $a^1{}_j$, $a^2{}_j$, $\dots$, $a^n{}_j$ são os coeficientes de Fourier

$$a^i{}_j := \langle T(\mathbf{e}_j), \mathbf{e}_i \rangle \tag{20.1}$$

de $T(\mathbf{e}_j)$. Então a matriz que representa T na base $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ é a matriz quadrada

$$A = (a^i{}_j) = \begin{pmatrix} a^1{}_1 & a^1{}_2 & \dots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \dots & a^2{}_n \\ \vdots & \vdots & \ddots & \vdots \\ a^n{}_1 & a^n{}_2 & \dots & a^n{}_n \end{pmatrix}$$

De fato, a imagem do vetor $\mathbf{x} = \sum_j x^j \, \mathbf{e}_j$ pelo operador T é o vetor

$$T(\mathbf{x}) = T\left(\sum_j x^j \, \mathbf{e}_j \right) = \sum_j x^j \, T(\mathbf{e}_j) = \sum_j x^j \, \sum_i a^i{}_j \, \mathbf{e}_i = \sum_i \left(\sum_j a^i{}_j \, x^j \right) \mathbf{e}_i$$

Em notação matricial, se X é o vetor coluna que representa $\mathbf{x}$, então $T(\mathbf{x})$ é representado pelo vetor coluna

$$AX = \begin{pmatrix} a^1{}_1 & a^1{}_2 & \dots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \dots & a^2{}_n \\ \vdots & \vdots & \ddots & \vdots \\ a^n{}_1 & a^n{}_2 & \dots & a^n{}_n \end{pmatrix} \begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix}$$

Seja agora $\mathbf{y}$ um outro vetor de $\mathbf{H}$. O produto interno $\langle T(\mathbf{x}), \mathbf{y}\rangle$, em notação matricial, é o escalar

$$\langle T\mathbf{x}, \mathbf{y}\rangle = \overline{Y}^\top AX$$

Mas $\overline{Y}^\top AX = \overline{\left(\overline{A}^\top Y\right)}^\top X$, ou seja,

$$\langle T\mathbf{x}, \mathbf{y}\rangle = \; = \langle \mathbf{x}, T^*\mathbf{y}\rangle$$

se o operador $T^* : \mathbf{H} \to \mathbf{H}$ é definido, na base ortonormada escolhida, pela matriz "transposta conjugada"

$$\boxed{A^* := \overline{A}^\top}$$

Se o espaço euclidiano $\mathbf{H}$ é real, e portanto A é uma matriz real, então o operador T^* é definido pela matriz transposta $A^\top$.

2. (operador adjunto) Seja $\mathbf{H}$ um espaço euclidiano, real ou complexo, de dimensão finita. O *adjunto* do operador linear $T \in \mathrm{End}(\mathbf{H})$ é o único operador $T^* \in \mathrm{End}(\mathbf{H})$ (a notação dos físicos é $T^\dagger$) tal que

$$\boxed{\langle T\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, T^*\mathbf{y}\rangle}$$

para todos $\mathbf{x}, \mathbf{y} \in \mathbf{H}$.

Se o espaço linear é real, fixada uma base ortonormada $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$, podemos considerar que $\mathbf{H} \approx \mathbb{R}^n$ munido do produto interno canónico $\langle \mathbf{x}, \mathbf{y}\rangle = \sum_i x^i y^i$. Se $A = (a^i{}_j) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é a matriz que representa o operador T na base escolhida, então a *matriz transposta* $A^\top = (b^i{}_j)$, definida por $b^i{}_j := a_i^j$ (ou seja, as linhas de $A^\top$ são as colunas de A e vice-versa), define o operador adjunto T^*. Se o espaço linear é complexo, fixada uma base ortonormada, podemos considerar que $\mathbf{H} \approx \mathbb{C}^n$ munido do produto hermítico $\langle \mathbf{x}, \mathbf{y}\rangle = \sum_i x^i \overline{y^i}$. Se $A = (a^i{}_j) \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ é a matriz que representa o operador T na base escolhida, então a matriz do operador adjunto T^* é a matriz *conjugada transposta* $\overline{A}^\top = (b^i{}_j)$, definida por $b^i{}_j = \overline{a^j{}_i}$. Isto prova a existência do adjunto. A unicidade é um exercício.

Por outro lado, a existência do adjunto em espaços euclidianos de dimensão infinita é problemática.

É imediato verificar que

$$(T^*)^* = T \qquad (T+S)^* = T^* + S^* \qquad (\lambda T)^* = \overline{\lambda} T^* \qquad (TS)^* = S^* T^*$$

- Mostre que o adjunto de um operador linear é único (lembre que o único vetor ortogonal a todos os vetores de um espaço euclidiano é o vetor nulo). Observe que esta prova não usa a dimensão finita do espaço, mas apenas a existência do adjunto.
- Mostre que $\ker T^* = (\mathrm{im} T)^\perp$ e que $(\ker T^*)^\perp = \mathrm{im} T$.
- Se T é invertível, então $(T^{-1})^* = (T^*)^{-1}$.
- Um subespaço $V \subset \mathbf{H}$ é invariante para o operador T sse o subespaço ortogonal $V^\perp$ é invariante para o operador adjunto T^*.

3. (operadores auto-adjuntos, valores médios e valores próprios) Um operador linear $T : \mathbf{H} \to \mathbf{H}$, definido num espaço euclidiano $\mathbf{H}$, é dito *auto-adjunto* ou *hermítico* (em inglês, *self-adjoint* ou *hermitian*) se

$$\boxed{\langle \mathbf{x}, T\mathbf{y} \rangle = \langle T\mathbf{x}, \mathbf{y} \rangle} \tag{20.2}$$

para todos $\mathbf{x}, \mathbf{y} \in \mathbf{H}$, ou seja, se existe o operador adjunto T^* e se $T^* = T$ (mas observe que para definir um operador auto-adjunto não é necessário definir o adjunto de um operador!).

Se $\mathbf{H} \approx \mathbb{C}^n$, e se $A = (a^i{}_j) \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ é a matriz que representa o operador T numa base ortonormada, então T é auto-adjuntos sse a matriz é "hermítica", i.e. se $\overline{A}^\top = A$, ou seja, $a^i{}_j = \overline{a^j{}_i}$. Um operador auto-adjunto de um espaço euclidiano real $\mathbf{H} \approx \mathbb{R}^n$ é também dito *simétrico*, pois é representado, numa base ortonormada, por uma matriz $A = (a^i{}_j) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ que é *simétrica*, ou seja, tal que $A^\top = A$, ou seja, $a^i{}_j = a^j{}_i$.

Dado um operador linear $T \in \mathrm{End}(\mathbf{H})$ e um vetor $\mathbf{v} \in \mathbf{H}$, o número

$$\boxed{\langle T\mathbf{v}, \mathbf{v} \rangle}$$

é chamado *valor médio* do T sobre $\mathbf{v}$. Se $\mathbf{H}$ tem dimensão finita e A é matriz que representa o operador T numa base ortonormada, então o seu valor médio sobre o vetor coluna V é

$$A[V] := \overline{V}^\top AV$$

Se $\mathbf{v}$ é um vetor próprio com valor próprio λ do operador T, ou seja, $T\mathbf{v} = \lambda\mathbf{v}$ e $\|\mathbf{v}\| \neq 0$, então $\langle T\mathbf{v}, \mathbf{v} \rangle = \lambda\|\mathbf{v}\|^2$, e portanto o valor próprio é dado pela expressão

$$\lambda = \frac{\langle T\mathbf{v}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2} \,. \tag{20.3}$$

Consequências importante desta fórmula são as seguintes.

Teorema 20.1. *O único operador auto-adjunto $T : \mathbf{H} \to \mathbf{H}$ que tem valores médios nulos $\langle T\mathbf{v}, \mathbf{v} \rangle = 0$ para todos os $\mathbf{v} \in \mathbf{H}$ é o operador trivial $T = 0$.*

Demonstração. Se T é auto-adjunto, aplicando a identidade de polarização

$$\langle T(\mathbf{x}+\mathbf{y}), \mathbf{x}+\mathbf{y} \rangle - \langle T(\mathbf{x}-\mathbf{y}), \mathbf{x}-\mathbf{y} \rangle = 2\left(\langle T\mathbf{x}, \mathbf{y} \rangle + \langle T\mathbf{y}, \mathbf{x} \rangle\right)$$

aos pares de vetores $\mathbf{x}$, $\mathbf{y}$ e depois $\mathbf{x}$, $i\mathbf{y}$, obtemos $\Re\langle T\mathbf{x}, \mathbf{y} \rangle = 0$ e $\Im\langle T\mathbf{x}, \mathbf{y} \rangle = 0$ para todos $\mathbf{x}, \mathbf{y} \in \mathbf{H}$, respetivamente, e portanto $T\mathbf{x} = 0$ para todo $\mathbf{x} \in \mathbf{H}$. □

Os operadores auto-adjuntos são precisamente os operadores com valores médios reais.

Teorema 20.2. *Um operador $T : \mathbf{H} \to \mathbf{H}$ é auto-adjunto sse o valor médio $\langle T\mathbf{v}, \mathbf{v} \rangle$ é real para todos os $\mathbf{v} \in \mathbf{H}$. Em particular, os valores próprios de um operador auto-adjunto são reais.*

Demonstração. Se T é auto-adjunto, então $\langle T\mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v}, T\mathbf{v} \rangle = \overline{\langle T\mathbf{v}, \mathbf{v} \rangle}$. Vice-versa, se $\langle T\mathbf{v}, \mathbf{v} \rangle$ é real para todos os $\mathbf{v} \in \mathbf{H}$, então $\langle T\mathbf{v}, \mathbf{v} \rangle = \overline{\langle T\mathbf{v}, \mathbf{v} \rangle} = \langle \mathbf{v}, T\mathbf{v} \rangle = \langle T^*\mathbf{v}, \mathbf{v} \rangle$, ou seja, $\langle (T-T^*)\mathbf{v}, \mathbf{v} \rangle = 0$ para todos os $\mathbf{v} \in \mathbf{H}$. Sendo $i(T-T^*)$ auto-adjunto, o teorema 20.1 então implica que $T = T^*$. □

Um operador linear $S : \mathbf{H} \to \mathbf{H}$, definido num espaço euclidiano $\mathbf{H}$, é dito *hemi-hermítico* ou *anti-hermítico* (em inglês, *skew-hermitian*) se

$$\boxed{\langle \mathbf{x}, S\mathbf{y} \rangle = - \langle S\mathbf{x}, \mathbf{y} \rangle} \tag{20.4}$$

para todos $\mathbf{x}, \mathbf{y} \in \mathbf{H}$, ou seja, se $S^* = -S$. Se $\mathbf{H} \approx \mathbb{C}^n$, e se $B = (b^i{}_j) \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ é a matriz que representa o operador S numa base ortonormada, então S é hemi-hermítico se $B^* = -B$, ou seja, $b^i{}_j = -\overline{b^j{}_i}$. Se o espaço é complexo, $S = iT$ é hemi-hermítico sse T é hermítico. Em particular, pelo teorema 20.2, os valores próprios de um operador hemi-hermítico num espaço euclidiano complexo são imaginários puros, i.e. estão no eixo imaginário $i\mathbb{R} \subset \mathbb{C}$. Um operador hemi-hermítico de um espaço euclidiano real é também dito *anti-simétrico*, pois a sua matriz B numa base ortonormada satisfaz $B^\top = -B$, ou seja, $b^i{}_j = -b^j{}_i$. Um operador anti-simétrico de um espaço euclidiano real não pode ter valores próprios diferentes de zero.

Teorema 20.3. *Vetores próprios com valores próprios distintos de um operador auto-adjunto (ou hemi-hermítico) são ortogonais.*

Demonstração. Se $T\mathbf{v} = \lambda\mathbf{v}$ e $T\mathbf{w} = \mu\mathbf{w}$, e se o operador A é hermítico ou hemi-hermítico, então

$$\langle A\mathbf{v}, \mathbf{w} \rangle = \pm \langle \mathbf{v}, A\mathbf{w} \rangle \qquad \Rightarrow \qquad (\lambda \mp \overline{\mu}) \langle \mathbf{v}, \mathbf{w} \rangle = 0 \,,$$

onde os sinais $\pm$ dependem do operador A ser hermítico ou anti-hermítico, respetivamente. Nos dois casos, obervando que os valores próprios são reais ou imaginários puros, respetivamente, $\lambda \neq \mu$ implica $\langle \mathbf{v}, \mathbf{w} \rangle = 0$. □

- A identidade é um operador auto-adjunto.
- Uma homotetia $\mathbf{x} \mapsto \lambda\mathbf{x}$ com λ real é hermítica ou simétrica, dependendo se definida num espaço euclidiano complexo ou real, respetivamente.
- Se T é auto-adjunto, então também αT, com α real, é autoadjunto. Se T e S são auto-adjuntos, então também $\alpha T + \beta S$, com α e β reais, é auto-adjunto.
- Se T é auto-adjunto, então também as potências T^n, com $n = 1, 2, \dots$, são operadores auto-adjuntos. De consequência, se $p(t) = a_n t^n + \cdots + a_1 t + a_0$ é um polinómio com coeficientes reais, então o operador $p(T) := a_n T^n + \cdots + a_1 T + a_0$ é auto-adjunto.
- Se T é auto-adjunto e invertível, então o inverso T^{-1} é também auto-adjunto.
- Se $T : \mathbf{H} \to \mathbf{H}$ é auto-adjunto e $\mathbf{V} \subset \mathbf{H}$ é um subespaço invariante (por exemplo, o espaço próprio $\mathbf{H}_\lambda$ dos vetores $\mathbf{v} \in \mathbf{H}$ tais que $T\mathbf{v} = \lambda\mathbf{v}$), então também o complemento ortogonal $\mathbf{V}^\perp$ é invariante.
- Sejam T e S dois operadores hermíticos de um espaço euclidiano $\mathbf{E}$. Em geral, a composição TS não é hermítico (dê exemplos). A composição TS é um operador hermítico se T e S comutam, ou seja, se $TS = ST$.
- Se S é um operador anti-simétrico de um espaço vetorial real $\mathbf{E}$, então $\langle S\mathbf{x}, \mathbf{x} \rangle = 0$ para todo o $\mathbf{x} \in \mathbf{E}$.
- Dado um operador linear arbitrário $T \in \mathrm{End}(\mathbf{E})$ num espaço euclidiano de dimensão finita $\mathbf{E}$, o operador T^*T é auto-adjunto. Se $\mathbf{v}$ é um vetor próprio de T com valor próprio λ, então $\langle T^*T\mathbf{v}, \mathbf{v} \rangle = |\lambda|^2 \|\mathbf{v}\|^2$.
- Cada operador $T \in \mathrm{End}(\mathbf{H})$ de um espaço euclidiano de dimensão finita pode ser decomposto, de maneira única, como soma

$$T = X + iY$$

de um operador hermítico X e um operador hemi-hermítico iY (ou seja, i vezes um operador hermítico Y). Basta escolher $X = (T + T^*)/2$ e $Y = (T - T^*)/2i$. Observe que $T^* = X - iY$ (uma fórmula que lembra o conjugado de um número complexo, ou seja, X e Y são moralmente a parte real e a parte imaginária do operador T!)
Mostre que T é *normal*, ou seja, comuta com o próprio adjunto (i.e. $TT^* = T^*T$), sse $XY = YX$.

- Verifique a *identidade de polarização*

$$\langle A(\mathbf{x}+\mathbf{y}), \mathbf{x}+\mathbf{y}\rangle - \langle A(\mathbf{x}-\mathbf{y}), \mathbf{x}-\mathbf{y}\rangle = 2\left(\langle A\mathbf{x}, \mathbf{y}\rangle + \langle A\mathbf{y}, \mathbf{x}\rangle\right) \tag{20.5}$$

  válida para um operador genérico $A \in \mathrm{End}(\mathbf{H})$.

- A inversão $T(x, y, z) = (-x, y, z)$ relativamente ao plano y-z é simétrica? E a inversão relativamente a um plano genérico de $\mathbb{R}^2$ passando pela origem?
- [Ap69] *Vol. 2* **5.5**.

4. (adjoint and self-adjointness in infinite dimension) When the Euclidian space $\mathbf{H}$ has infinite dimension, interesting linear operators may be defined only in proper subspaces of well behaved vectors, i.e. are linear maps $L : \mathbf{D}(L) \to \mathbf{H}$ from a "domain" $\mathbf{D}(L) \subset \mathbf{H}$ into $\mathbf{H}$. When $\mathbf{H}$ is an Hilbert space, an operator is said *symmetric* or *hermitian* if $\langle L\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, L\mathbf{y}\rangle$ for all $\mathbf{x}$, $\mathbf{y} \in \mathbf{D}(L)$ (typically we also ask that such domain is "dense" in $\mathbf{H}$, i.e. that its closure is the whole space).

   The term *self-adjoint* is reserved to operators which are defined in the whole Hilbert space, satisfying $\langle T\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, T\mathbf{y}\rangle$ for all $\mathbf{x}$, $\mathbf{y} \in \mathbf{H}$. It turns out that such operators are actually "bounded", i.e. satisfy a bound $\|T\mathbf{v}\| \leq K\,\|\mathbf{v}\|$ for all $\mathbf{v} \in \mathbf{H}$ and some minimal $K := \|T\| < \infty$ called "norm" of T.

   Conversely, any bounded operator $T : \mathbf{H} \to \mathbf{H}$ has an adjoint operator T^*, defined, as in the finite dimensional case, by the identity $\langle Tx, y\rangle = \langle x, T^*y\rangle$ (its existence is not obvious, and follows from the Riesz representation theorem). Thus, self-adjointness means $T = T^*$, as in the finite dimensional case.

   Nevertheless, differential operators and most physically relevant operators (position, momentum, energy, ...) are not bounded! Self-adjointness for unbounded operators is a much subtler issue.

5. (multiplication and differential operators) Physicists and engineers interested in solving ordinary or partial differential equations (like the wave equation, the heat equation, the Schrödinger equation, ...) must deal with differential and multiplication operators defined in spaces of functions, which are infinite dimensional linear spaces.

   For example, we may consider the Euclidean space $\mathbf{E}$ of continuous complex or real valued functions $f : [a, b] \to \mathbb{C}$, equipped with the inner product $\langle f, g\rangle := \int_a^b f(t)\,\overline{g(t)}\,dt$. The bounded interval $[a, b]$ may be substituted by an unbounded interval like the whole real line $(-\infty, \infty)$, with some care ...

   Any continuous function $m(t)$ induces a *multiplication operator* $M \in \mathrm{End}(\mathbf{E})$ defined as

$$(Mf)(t) := m(t)\,f(t)$$

   It is clear that M satisfies $\langle f, Mg\rangle = \langle Mf, g\rangle$ for all $f, g, \in \mathbf{E}$, i.e. is self-adjoint, if the function $m(t)$ has real values.

   Differential operators can only operate on smaller domains, for example in the subspaces $\mathbf{E}^k \subset \mathbf{E}$ of those functions which admit $k = 1, 2, \ldots,$ or ∞ continuous derivatives. There, we may define the *derivative operator*

$$(\partial f)(t) := f'(t)\,,$$

   its powers ∂^k, and therefore polynomials $L = \sum_{k=0}^N a_k \partial^k$. Looking for the adjoint of ∂, we observe that integration by parts gives the identity

$$\langle f, \partial g\rangle = f(b)\overline{g(b)} - f(a)\overline{g(a)} - \langle \partial f, g\rangle$$

Choosing appropriate "boundary conditions", for example restricting ∂ to the subspace $\mathbf{E}^1_0 \subset \mathbf{E}^1$ of those functions satisfying $f(a) = f(b)$, this operator ∂ becomes skew-symmetric. Most important is the *Laplacian* $\Delta := \partial^2$, which therefore is symmetric when defined on $\mathbf{E}^2_0$.

- Given a continuous function $q(t)$ and continuously differentiable function $p(t)$, consider the subspace $\mathbf{E}^2_p \subset \mathbf{E}^2$ of those twice continuously differentiable functions satisfying the boundary conditions $p(a)f(a) = p(b)f(b) = 0$, and the *Sturm-Liouville operator* $L : \mathbf{E}^2_p \to \mathbf{E}$, defined by
$$Lf := (pf')' + qf \, .$$
For example, if $q = 0$ e $p = 1$, we recover the Laplacian $\Delta = D^2$. Show that L is symmetric.
- Let $P : \mathbf{E} \to \mathbf{E}$ be the *primitive operator*, defined as
$$(Pf)(t) = \int_a^t f(\tau)\, d\tau \, .$$
Show that it is skew-symmetric when restricted to the kernel $\mathbf{V} = \ker(I)$ of the linear form $I(f) := \int_a^b f(t)\, dt$.

6. (observáveis e valores médios em mecânica quântica) O espaço dos estados de uma partícula quântica (não relativística) é um espaço de Hilbert $\mathbf{H} \approx \ell^2$, i.e. um espaço euclidiano complexo de dimensão infinita que admite um subconjunto denso e numerável, e portanto uma base numerável. Os "observáveis" limitados são operadores auto-adjuntos A (o caso de observáveis não limitados, como "posição", "momento", "energia", . . . é mais delicado), cujos vetores próprios $|1\rangle, |2\rangle, |3\rangle, \dots$ formam uma base de $\mathbf{H}$. Os valores próprios $\lambda_1, \lambda_2, \lambda_3, \dots$ (que são números reais) são os possíveis resultados das observações. O valor médio esperado, ao repetir a observação do observável A no estado unitário arbitrário $|\psi\rangle = \sum_n \psi_n |n\rangle$ um número grande de vezes, é o valor médio
$$\langle\psi| A |\psi\rangle = \sum_n \lambda_n |\psi_n|^2 \, ,$$
sendo $\|\psi\|^2 = \sum_n |\psi_n|^2 = 1$.

7. (isometrias e operadores ortogonais/unitários) Uma transformação linear $T : \mathbf{E} \to \mathbf{F}$ de um espaço euclidiano $\mathbf{E}$ num espaço euclidiano $\mathbf{F}$, reais ou complexos, é uma *isometria (linear)* se preserva os produtos internos, i.e. se
$$\langle T\mathbf{x}, T\mathbf{y}\rangle_{\mathbf{F}} = \langle \mathbf{x}, \mathbf{y}\rangle_{\mathbf{E}}$$
para todos $\mathbf{x}, \mathbf{y} \in \mathbf{E}$. É evidente que uma isometria preserva a ortogonalidade (i.e. se $\mathbf{x} \perp \mathbf{y}$ então também $T\mathbf{x} \perp T\mathbf{y}$), preserva as normas (i..e. $\|T\mathbf{x}\|_{\mathbf{F}} = \|x\|_{\mathbf{E}}$ para todos $\mathbf{x} \in \mathbf{E}$), e de consequência as distâncias. Toda isometria é injetiva (pois preserva as normas), de consequência invertível na sua imagem, e a sua inversa $T^{-1} : \mathrm{im}(T) \to \mathbf{E}$ é também uma isometria.

Se existir uma isometria bijetiva $T : \mathbf{E} \to \mathbf{F}$, os espaços euclidianos $\mathbf{E}$ e $\mathbf{F}$ são ditos *isomorfos*.

- Uma "inclusão" $(x, y) \mapsto (x, y, 0)$ é uma isometria de $\mathbb{R}^2$ em $\mathbb{R}^3$, que não é sobrejetiva.
- Uma projeção ortogonal $(x, y, z) \mapsto (x, y, 0)$ não é uma isometria de $\mathbb{R}^3$ em $\mathbb{R}^3$.
- Sejam $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ e $(\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n)$ duas bases ortonormadas dos espaços euclidianos de dimensão finita $\mathbf{E}$ e $\mathbf{F}$, respetivamente, e seja $T : \mathbf{E} \to \mathbf{F}$ o operador tal que $T(\mathbf{e}_k) = \mathbf{f}_k$ para todo $k = 1, \dots, n$. Mostre que T é uma isometria bijetiva. Em particular, todo espaço euclidiano de dimensão finita é isomorfo a $\mathbb{C}^n$ ou $\mathbb{R}^n$, dependendo se real ou complexo.

Se $T : \mathbf{H} \to \mathbf{H}$, e se existe o operador adjunto T^* (por exemplo, se $\mathbf{H}$ tem dimensão finita), então a condição $\langle T\mathbf{x}, T\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ para todos $\mathbf{x}, \mathbf{y} \in \mathbf{H}$ é equivalente a $T^*T = I$. Um operador linear $U \in \mathrm{End}(\mathbf{H})$ é dito *unitário* se
$$\boxed{U^*U = UU^* = I \, ,}$$

ou seja, se é uma isometria bijetiva, e portanto $U^{-1} = U^*$. Um operador unitário de um espaço euclidiano real é também dito *ortogonal*.

Se $\mathbf{H}$ é um espaço euclidiano de dimensão finita, então toda a isometria $U \in \mathrm{End}(\mathbf{H})$ é unitária, pois envia uma base (finita) ortonormada numa base ortonormada, e portanto é bijetiva. Se o espaço é complexo, $\mathbf{H} \approx \mathbb{C}^n$, então a matriz $A = (a^i{}_j) \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ que define um operador numa base ortonormada é uma "matriz unitária", ou seja, tal que $A^*A = AA^* = I$. Se o espaço é real, $\mathbf{H} \approx \mathbb{R}^n$, então a matriz A que define um operador ortogonal numa base ortonormada é uma "matriz ortogonal", ou seja, tal que $A^\top A = AA^\top = I$.

Teorema 20.4. *Os valores próprios de um operador unitário $U \in \mathrm{End}(\mathbf{H})$ estão na circunferência unitária $\mathbb{S} = \{z \in \mathbb{C} \quad t.q. \quad |z| = 1\}$, i.e. têm valor absoluto $|\lambda| = 1$. Em particular, um operador ortogonal só pode ter valores próprios ± 1.*

Demonstração. Se U é unitário e $U\mathbf{v} = \lambda\mathbf{v}$, então $|\lambda|^2 \|\mathbf{v}\|^2 = \langle U\mathbf{v}, U\mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{v}\rangle = \|\mathbf{v}\|^2$. □

- Mostre que $\|T\mathbf{x}\| = \|x\|$ para todo $\mathbf{x} \in \mathbf{H}$ implica $\langle T\mathbf{x}, T\mathbf{y}\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ para todos $\mathbf{x}$, $\mathbf{y} \in \mathbf{H}$ (aplique a primeira identidade aos vetores $\mathbf{x} + \mathbf{y}$ e $\mathbf{x} + i\mathbf{y}$).
- Se $\mathbf{V} \subset \mathbf{H}$ é um subespaço invariante para o operador unitário $U : \mathbf{H} \to \mathbf{H}$, então também o subespaço ortogonal $\mathbf{V}^\perp$ é invariante.
- As translações $T_\mathbf{a} : \mathbf{x} \mapsto \mathbf{x} + \mathbf{a}$ são "isometrias" do espaço euclidiano $\mathbb{R}^n$, no sentido em que preservam as distâncias (pois $\|T_\mathbf{a}(\mathbf{x}) - T_\mathbf{a}(\mathbf{y})\| = \|\mathbf{x} - \mathbf{y}\|$), mas não são transformações lineares, se $\mathbf{a} \neq \mathbf{0}$.
- Em dimensão infinita existem isometrias $T \in \mathrm{End}(\mathbf{H})$ não sobrejetivas, e portanto não unitárias! Por exemplo, o *deslocamento* (em inglês, *shift*) $T : \ell^2 \to \ell^2$, que envia
$$T(x_1, x_2, x_3, \dots) = (0, x_1, x_2, \dots) \,.$$
- Sejam $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$ e $(\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_n)$ duas bases ortonormadas do espaço euclidiano de dimensão finita $\mathbf{H}$, e seja $U : \mathbf{H} \to \mathbf{H}$ o operador tal que $U(\mathbf{e}_k) = \mathbf{f}_k$ para todo $k = 1, \dots, n$, definido pela matriz cujas colunas são as componentes dos $\mathbf{f}_k$ relativamente à base $(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n)$. Mostre que U é unitário, se $\mathbf{H} \approx \mathbb{C}^n$ é um espaço euclidiano complexo, ou ortogonal, se $\mathbf{H} \approx \mathbb{R}^n$ é um espaço euclidiano real.
- Mostre que um operador $U \in \mathrm{End}(\mathbf{H})$ de um spaço euclidiano de dimensão finita $\mathbf{H}$ é unitário sse envia bases ortonormadas em bases ortonormadas.
- [Ap69] *Vol. 2* **5.11.**

21 Teorema espetral

ref: [Ap69] *Vol 2*, **5.6**, **5.11**, **5.20** ; [La87] *Ch. VIII*

1. (diagonalização em espaços euclidianos) Any linear operator $L : \mathbf{E} \to \mathbf{F}$ between finite dimensional linear spaces is "diagonal" (i.e. is represented by a matrix $A = (a_{ij})$ such that $a_{ij} = 0$ if $i \neq j$ and $a_{kk} = 1$ or 0) if one chooses appropriate basis of $\mathbf{E}$ and $\mathbf{F}$ (see [Ap69], *Vol. 2*, **2.11**).

   Also, a "generic" linear operator $L : \mathbb{C}^n \to \mathbb{C}^n$ is diagonalizable. Indeed, a generic complex polynomial of degree n (the characteristic polynomial of the matrix which represents L in the canonical basis) factorizes as a product $(z - \lambda_1)(z - \lambda_2)\dots(z - \lambda_n)$ with distict roots $\lambda_1, \lambda_2, \dots, \lambda_n$, and eigenvectors with distinct eigenvalues are linearly independent.

   The spectral theorem, in its various versions, addresses a more interesting and useful problem. To find out which operators $T : \mathbf{H} \to \mathbf{H}$, from an Euclidean space $\mathbf{H}$ into itself, admit an orthonormal basis of eigenvectors, and therefore, if $\mathbf{H}$ has finite dimension, are represented in this orthonormal basis by a square diagonal matrix

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

   with eigenvalues $\lambda_1, \dots, \lambda_n$ (which are in general complex numbers).

   The answer, in finite dimension, is the class of "normal operators", those operators which commute with their adjoint (this is clearly the case of operators defined by diagonal matrices). This include the important subclasses of self-adjoint and unitary operators, which have real and unitary eigenvalues, respectively.

   Besides a simple induction argument, which uses the geometric properties of normal operators, the main ingredient of the proof of the spectral theorem is the existence of at least one eigenvalue. Here is where analysis is needed. In the form of the fundamental theorem of algebra (any non constant polynomial has a root in the complex plane), or in the apparently simpler form of the Weierstrass theorem (a continuous function defined on a compact set attains its maximum) together with some simple differential calculus.

2. (teorema espectral para operadores auto-adjuntos em dimensão finita) A forma mais simples de enunciar o teorema espetral é a seguinte.

   Teorema 21.1 (teorema espetral)**.** *Seja A é um operador auto-adjunto de um espaço euclidiano de dimensão finita $\mathbf{H}$. Existe uma base ortonormada de $\mathbf{H}$ formada por vetores próprios de A, com valores próprios reais.*

   De consequência, o espaço euclidiano é uma soma direta ortogonal $\mathbf{H} = \oplus_{\lambda \in \sigma(A)} \mathbf{H}_\lambda$ de espaços próprios, e se $P_\lambda : \mathbf{H} \to \mathbf{H}_\lambda$ denotam a projeções ortogonais de $\mathbf{H}$ sobre os espaços próprios $\mathbf{H}_\lambda = \{\mathbf{v} \in \mathbf{H} \,:\, A\mathbf{v} = \lambda \mathbf{v}\}$, com $\lambda \in \sigma(A)$, então o operador é uma soma direta de homotetias

$$A = \sum_{\lambda \in \sigma(A)} \lambda \, P_\lambda$$

   Se $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ é uma base ortonormada de vetores próprios, com valores próprios reais $\lambda_1, \lambda_2, \dots, \lambda_n$ (não necessáriamente distintos), então a matriz que representa o operador A nesta base é a matriz diagonal

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Demonstração. Seja $\mathbf{H} \approx \mathbb{C}^n$, e seja A a matriz que representa o operador numa base ortonormada. Pelo teorema fundamental da álgebra, o polinómio caraterístico $P_A(t) = \det(tI - A)$ possui (pelo menos) uma raiz λ, e portanto o operador admite um vetor próprio $\mathbf{v}$ com valor próprio λ, que é real pelo teorema 20.2. Mas o espaço ortogonal $(\mathbb{C}\mathbf{v})^\perp \approx \mathbb{C}^{n-1}$ é um subespaço invariante de A, e a restrição de A a este espaço é ainda um operador auto-adjunto. O teorema espetral segue por indução finita, sendo trivial num espaço de dimensão um $\mathbf{H} \approx \mathbb{C}$. □

Em particular, se a matriz auto-adjunta $A = (a^i{}_j) \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ representa o operador auto-adjunto A numa base ortonormada arbitrária (por exemplo, a base canónica), então existe uma matriz unitária U tal que

$$A = U\,\Lambda\,U^* .$$

onde $\Lambda = U^*\,A\,U$ é uma matriz diagonal $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ cujas entradas são os valores próprios (reais) de A (as colunas de U são os vetores próprios de A na base ortonormal inicial). Se o espaço euclidiano é real, então existe uma matriz ortogonal O tal que

$$A = O\,\Lambda\,O^\top ,$$

onde $\Lambda = O^\top\,A\,O$ é diagonal.

O "teorema fundamental da álgebra" (o resultado de análise que diz que um polinómio não constante tem pelo menos uma raiz no plano complexo) foi utilizado para deduzir a existência de pelo menos um valor próprio. Pode ser substituido pelo teorema de Weierstrass, que afirma que uma função contínua num compacto admite um máximo, e um pouco de cálculo elementar. Esta ideia é utilizada na caracterização dos valores próprios mínimo e máximo de Rayleigh e Ritz.

- Verifique que um operador auto-adjunto L num espaço euclidiano de dimensão um $\mathbf{H} \approx \mathbb{C}$ ou $\mathbb{R}$ é uma homotetia $L\mathbf{v} = \lambda\mathbf{v}$ com $\lambda \in \mathbb{R}$.
- Seja $\mathbf{v}$ um vetor próprio do operador auto-adjunto L. Mostre que $(\mathbb{C}\mathbf{v})^\perp$ é um subespaço invariante de L.
- Mostre que as potências de $L = \sum_{\lambda\in\sigma(A)} \lambda P_\lambda$ são $L^n = \sum_{\lambda\in\sigma(L)} \lambda^n P_\lambda$.
- Observe que um operador $L = iT$ é hemi-hermítico sse T é hermítico. Enuncie o correspondente teorema espetral para operadores hemi-hermíticos.
- Diagonalize, se possível, a transformação $T(z, w) = (z - iw, iz - 2w)$ do espaço euclidiano $\mathbb{C}^2$.
- [Ap69] *Vol. 2* **5.11.**

3. (diagonalização simultânea) Sejam L e M são dois operadores lineares definidos num espaço linear $\mathbf{V}$. Se L e M comutam, e se $\mathbf{V}_\lambda$ é um espaço próprio de L associado ao valor próprio λ, então $\mathbf{V}_\lambda$ é invariante para M, i.e. $M(\mathbf{V}_\lambda) \subset \mathbf{V}_\lambda$. De fato, se $\mathbf{v} \in \mathbf{V}_\lambda$, então

$$L(M\mathbf{v}) = M(L\mathbf{v}) = \lambda M\mathbf{v} ,$$

o que significa que $M\mathbf{v} \in \mathbf{V}_\lambda$. Em particular,

Teorema 21.2. *Sejam L e M são dois operadores auto-adjuntos, definidos num espaço euclidiano de dimensão finita $\mathbf{H}$. Existe uma base ortonormada de vetores próprios para os dois operadores sse L e M comutam.*

Demonstração. Se L e M comutam, a restrição de M a cada espaço próprio de L é um operador auto-adjunto, que pelo teoema espetral admite uma base de vetores próprios. A outra implicação é trivial, pois duas matrizes diagonais comutam. □

De consequência, se as matrizes hermíticas A e B comutam, existe uma matriz unitária U tal que U^*AU e U^*BU são diagonais.

4. (Rayleigh-Ritz quotient and Courant-Fischer min-max principle) Let $A \in \mathrm{End}(\mathbf{H})$ be a self-adjoint operator on a finite dimensional complex or real Euclidean space $\mathbf{H}$ of dimension n. We order its eigenvalues, which are real, according to the non-decreasing order

$$\lambda_1(A) \leq \lambda_2(A) \leq \cdots \leq \lambda_n(A)$$

(repeating the same eigenvalue as many time as its geometric multiplicity). We also denote by $\lambda_{\min}(A) := \lambda_1(A)$ and $\lambda_{\max} := \lambda_n(A)$ the minimal and maximal eigenvalues, respectively.

The eigenvalues are the critical points of the (restriction of the) quadratic form

$$Q_A(\mathbf{x}) := \langle A\mathbf{x}, \mathbf{x} \rangle$$

on the unit sphere $\mathbb{S}^{n-1} = \{\mathbf{x} \in \mathbf{H} \,:\, \|\mathbf{x}\| = 1\}$. Indeed, the gradient of Q_A is $2A\mathbf{x}$, while the gradient of the functions $\|\mathbf{x}\|^2$, whose level one set is the unit sphere, is $2\mathbf{x}$. They are proportional precisely at points $\mathbf{x} \in \mathbb{S}^{n-1}$ where $A\mathbf{x} = \lambda\mathbf{x}$ for some λ (which is therefore an eigenvalue), and at such points one has $Q_A(\mathbf{x}) = \lambda\|\mathbf{x}\|^2$. In particular, since the unit sphere is compact, the minumum and the maximum are attained. We have proved the following variational characterization of the smallest and largest eigenvalues.

Teorema 21.3 (Rayleigh-Ritz)**.** *For any $\mathbf{x} \in \mathbf{H}$ one has*

$$\lambda_{\min}(A)\, \|\mathbf{x}\|^2 \leq \langle \mathbf{x}, A\mathbf{x} \rangle \leq \lambda_{\max}(A)\, \|\mathbf{x}\|^2 \,.$$

Moreover, the smallest and the largest eigenvalues are

$$\lambda_{\min}(A) = \min_{\|\mathbf{x}\|=1} \langle \mathbf{x}, A\mathbf{x} \rangle \qquad \text{and} \qquad \lambda_{\max}(A) = \max_{\|\mathbf{x}\|=1} \langle \mathbf{x}, A\mathbf{x} \rangle \,.$$

The *Rayleigh-Ritz quotient*[27] [28] is the function $R_A : \mathbf{H}\backslash\{0\} \to \mathbb{R}$ defined by

$$R_A(\mathbf{x}) := \frac{\langle A\mathbf{x}, \mathbf{x} \rangle}{\|\mathbf{x}\|^2}$$

Observe that the Reyleigh-Ritz quotient computed at en eigenvector is the associated eigenvalue, i.e. if $A\mathbf{v} = \lambda\mathbf{v}$ then $R_A(\mathbf{v}) = \lambda$. Theorem 21.3 then says that

$$\lambda_{\min}(A) \leq R_A(\mathbf{x}) \leq \lambda_{\max}(A) \,,$$

and that the smallest and largest eigenvalues may also be computed minimizing or maximizing the Rayleigh-Ritz quotient over all non-zero vectors.

Others eigenvalues also may be defined/computed by a variational principle, as follows.

Teorema 21.4 (Courant-Fischer)**.** *The eigenvalues of a self-adjoint operator A on the Euclidean space $\mathbf{H}$ of dimension n are*

$$\lambda_k(A) = \min_{V \subset \mathbf{H}\,,\, \dim V = k} \quad \max_{\mathbf{x} \in V\,,\, \mathbf{x} \neq 0} \quad R_A(\mathbf{x}) \tag{21.1}$$

or also

$$\lambda_k(A) = \max_{V \subset \mathbf{H}\,,\, \dim V = n-k+1} \quad \min_{\mathbf{x} \in V\,,\, \mathbf{x} \neq 0} \quad R_A(\mathbf{x})\,. \tag{21.2}$$

Again, the Rayleigh quotient may be substituted by the quadratic form $\langle A\mathbf{x}, \mathbf{x} \rangle$ in the Courant-Fischer variational principle, provided we restrict to unitary vectors $\mathbf{x}$, i.e. add the conditon $\|\mathbf{x}\| = 1$ in all the above formulas.

[27] J.W. Strutt (later Lord Rayleigh), In Finding the Correction for the Open End of an Organ-Pipe, *Phil. Trans.* **161** (1870) 77.

[28] W. Ritz, Uber eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik, *J. reine angew. Math.* **135** (1908).

Demonstração. According to the spectral theorem 21.1 there exist a basis $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ of $\mathbf{H}$ formed by the eigenvalues of A, i.e. $A\mathbf{u}_k = \lambda_k \mathbf{u}_k$. For dimensional reasons, any subspace $V \subset \mathbf{H}$ of dimension k contains at least one non-trivial vector belonging to the span of the $\mathbf{u}_k, \mathbf{u}_{k+1}, \dots, \mathbf{u}_n$'s, i.e. a vector of the form $\mathbf{x} = x_k\mathbf{u}_k + x_{k+1}\mathbf{u}_{k+1} + \cdots + x_n\mathbf{u}_n$ with at least one coefficient $x_i \neq 0$. It is clear that value of the Rayleigh-Ritz quotient on such a vector is $R_A(\mathbf{x}) \geq \lambda_k$. On the other side, this minimal value is attained at the vector $\mathbf{u}_k$, which belongs to the particular subspace of dimension k spanned by the first $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k$. This proves the first variational principle (21.1). The dual principle, (21.2), is obtained substituting A with $-A$, observing that $\lambda_k(-A) = \lambda_{n-k+1}(A)$. $\square$

The above variational principles are also at the basis of "Monte Carlo algorithms" to find the eigenvalues of a self-adjoint operator. One may generate a huge number of random points in the unit sphere, and find the maximum of the quadratic form $\mathcal{Q}$. This gives the greatest eigenvalue λ_n and its eigenvector. Then one consider the lower dimensional unit sphere orthogonal to this eigenvector, and find λ_{n-1} ... And so on.

5. (Weyl's inequalities) A consequence of the Courant-Fischer min-max principle, theorem 21.4, are the *Weyl's (additive) inequalities*

$$\lambda_k(A) + \lambda_{\min}(B) \leq \lambda_k(A+B) \leq \lambda_k(A) + \lambda_{\max}(B)$$

(but Weyl deduced much more similar inequalities...) for a pair of self-adjont operators A and B.

In particular, if we observe that the spectral radius of the self-adjoint operator B is the maximal modulus of its eigenvalues, i.e. $\rho(B) = \max\{|\lambda_{\min}(B)|, |\lambda_{\max}(B)|\}$, we deduce the inequality

$$|\lambda_k(A+B) - \lambda_k(A)| \leq \rho(B)\,.$$

We may think that $A+B$ is obtained adding a (small) perturbation B the unperturbed operator A. Then the above inequality says the eigenvalues are stable under small perturbations (i.e. change in a controlled way).

6. (Laplacian in a bounded interval) Let $\mathbf{E}$ be the space of continuous functions $f : [0,\pi] \to \mathbb{R}$ or $\mathbb{C}$, equipped with the inner product $\langle f, g\rangle = \int_0^\pi f(x)\,\overline{g(x)}\,dx$. Recall that the *Laplacian* is the differential operator $\Delta = \partial^2$, namely

$$(\Delta f)(x) := f''(x)\,.$$

Let $\mathbf{E}_0^\infty \subset \mathbf{E}$ be the space of infinitely differentiable functions $f : [0,\pi] \to \mathbb{R}$ satisfying the boundary conditions $f(0) = f(\pi) = 0$ (the space of transversal displacements of a vibrating string). As an operator $\Delta : \mathbf{E}_0^\infty \to \mathbf{E}$, the Laplacian is symmetryic, i.e. $\langle \Delta f, g\rangle = \langle f, \Delta g\rangle$ for all $f, g, \in \mathbf{E}_0^\infty$. It is a simple exercise on ordinary differential equations to show that the eigenvalues of the Laplacian are $\lambda_n = -n^2$, with $n = 1, 2, 3, \dots$, and the corresponding eigenfunctions are, for example,

$$\mathbf{v}_n(x) = \sin(nx)\,.$$

Observe that the Laplacian is not bounded! One may actually show that such eigenfunctions form a "basis" of the Hilbert space $\mathbf{H}$ obtained from $\mathbf{E}$ by completion.

- Show that the "positive definite" Laplacian $-\Delta$ is positive, namely

$$\langle -\Delta f, f\rangle > 0$$

for all non-trivial $f \in \mathbf{E}_0^\infty$. Indeed, integrating by parts and using the boundary conditions, $f(0) = f(\pi) = 0$, show that

$$\langle -\Delta f, f\rangle = \langle \nabla f, \nabla f\rangle = \|\nabla f\|^2\,.$$

7. (self-adjoint operators with no eigenvalues and Dirac delta functions). The multiplication operator $(Xf)(x) := x\, f(x)$ is self-adjoint and bounded on the Euclidean space $\mathbf{E}$ as above, but has no eigenvalues. Indeed, the eigenvalue equation reads $x\, f(x) = \lambda\, f(x)$, and says that $f(x) = 0$ for all $x \neq \lambda$, and therefore that $f(x)$ is the zero vector of $\mathbf{E}$.

   Indeed, eigenvectors of the multiplication operators are *Dirac delta functions* $\delta(x-\lambda)$, with $0 \leq \lambda \leq \pi$, defined by the paradoxical equation

$$\int_0^\pi \delta(x-\lambda)\, f(x)\, dx := f(\lambda)$$

   (saying that $\delta(x-\lambda)$ is equal to zero for all $x \neq \lambda$ and so large at the single point $x = \lambda$ that its integral over the whole interval is equal to one!), and live outside the space $\mathbf{E}$. Laurent Schwartz theory of distributions (worth a Fields medal in 1950) will give a precise mathematical meaning to this physicists' idea.

8. (momentum and Laplacian on the line) Let $\mathbf{H}$ be the complex Euclidean space of square-integrable functions $f : \mathbb{R} \to \mathbb{C}$, equipped with the inner product $\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)\, \overline{g(x)}\, dx$.

   The "momentum operator" (in quantum mechanics) $P = -i\partial$ is symmetric when defined, for example, on the subspace $\mathcal{C}_0^\infty(\mathbb{R}) \subset \mathbf{H}$ of infinitely differentiable functions with compact support (since ∂ is skew-symmetric). Nevertheless, it does not have any eigenvalue. Indeed, functions satisfying $Pf = \lambda f$ are the plane waves $e^{i\lambda x}$, but no such exponential is square-integrable (indeed, they can't even satisfy boundary conditions like $f(0) = f(\pi) = 0$ if we were in a bounded interval). Thus, the candidates to be eigenvectors of the momentum operator are outside the space where the operator is naturally defined.

   The "kinetic energy" $P^2 = -\Delta$, the positive-definite Laplacian, is symmetric too, and of course should share its eigenvectors with P, the plane waves $e^{i\lambda x}$, with non-negative eigenvalues $\lambda^2 \geq 0$.

   The theory of "Fourier transform" (to be studied next year) will reveal in which sense the Laplacian is unitarily equivalent to a multiplication operator (the prototype of a self-adjoint operator). Functions in $\mathbf{H}$ will be integrals

$$f(x) = \int_{-\infty}^{\infty} \widehat{f}(\lambda)\, e^{i\lambda x}\, d\lambda$$

   of "eigenfunctions" of the Laplacian, and the Laplacian $-\Delta$ (or also the derivative operator ∂ as well as any constant coefficients differential operator) will act on their "coefficients/coordinates" $\widehat{f}(\lambda)$ as a multiplication operator $\widehat{f}(\lambda) \mapsto \lambda^2\, \widehat{f}(\lambda)$.

9. (spectral theorem for normal operators) The largest class of operators for which the spectral theorem holds is the class of normal operators.

   Let $\mathbf{H} \approx \mathbb{C}^n$ be a finite dimensional complex Euclidean space. A linear operator $N \in \mathrm{End}(\mathbf{H})$ is said *normal* if it commutes with its adjoint, i.e. if

$$\boxed{N^*\, N = N\, N^*}$$

   Unitary operators are normal, since satisfy the stronger $UU^* = U^*U = I$. Hermitian and skew-hermitian operators are also examples of normal operators, since their adjoint is proportional to themselves. Indeed, an operator T is normal iff the Hermitian operator X and the skew-Hermitian operator iY of the decomposition $T = X + iY$ commute (since $T^* = X - iY$, and therefore $[T, T^*] = -2i\,[X, Y]$).

   - Show that N is normal iff N^* is normal.
   - Show, using theorem 20.1, that N is normal iff

$$\|N\mathbf{v}\| = \|N^*\mathbf{v}\|$$

     for all vectors $\mathbf{v} \in \mathbf{H}$.

- Deduce (applying the above observation to the operator $N - \lambda$ and its adjoint $N^* - \overline{\lambda}$) that if N is a normal operator, then $\mathbf{v}$ is an eigenvector of N with eigenvalue λ iff it is an eigenvector of N^* with eigenvalue $\overline{\lambda}$.
- Deduce from the previous exercise that two eigenvectors of a normal operator corresponding to different eigenvalues are orthogonal.

It is clear that an operator which admits an orthonormal basis of eigenvectors, with arbitrary complex eigenvalues, is normal, since the adjoint of a complex diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ is a diagonal matrix $\Lambda^* = \mathrm{diag}(\overline{\lambda_1}, \dots, \overline{\lambda_n})$ with conjugate entries, and any two diagonal matrices commute. The converse is the most general spectral theorem, which reads as follows.

Teorema 21.5 (spectral theorem)**.** *A normal operator N on a finite dimensional complex Euclidean space $\mathbf{H}$ admits an orthonormal basis of eigenvectors, and therefore is a directed sum $N = \oplus_k \lambda_k$ of homoteties w.r. to an orthogonal decomposition $\mathbf{H} = \oplus_k \mathbf{H}_k$, where the $\mathbf{H}_k$'s are the eigenspaces of the distinct eigenvalues λ_k's (which are just complex, not necessarily real!).*

Although it clearly follows from the spectral theorem for self-adjoint operators 21.1 (which holds also for hemi-hermitian operators and therefore for operators $T = X + iY$ such that their hermitian part X and hemi-hermitian part iY commute, according to theorem 21.2), we give an independent proof.

Demonstração. The proof, again, is by induction. The operator N admits an eigenvalue λ, by the fundamental theorem of algebra. The corresponding eigenspace $\mathbf{H}_\lambda$ is also an eigenspace of the adjoint operator N^* with eigenvalue $\overline{\lambda}$. But then the orthogonal subspace $(\mathbf{H}_\lambda)^\perp$ is N-invariant (for if $\mathbf{v} \in \mathbf{H}_\lambda$, then $\langle \mathbf{w}, \mathbf{v} \rangle = 0$ implies $\langle N\mathbf{w}, \mathbf{v} \rangle = \langle \mathbf{w}, N^*\mathbf{v} \rangle = \lambda \langle \mathbf{w}, \mathbf{v} \rangle = 0$ too), and has dimension strictly smaller than the dimension of $\mathbf{H}$. The theorem follows by induction observing that it is trivial in dimension one. □

Thus, the analogy between operators and complex numbers is the following. If we look at normal operators as the universe of all operators which are diagonalizable in an orthonormal basis, hence as a complex plane (their possible eigenvalues), self-adjoint operators correspond to the real line (being those normal operators with real eigenvalues), while unitary operators correspond to the unit circle (being those normal operators with unitary eigenvalues).

Observe that the conclusion of the above theorem does not hold for ortogonal operators in real Euclidean spaces, since the eigenvalues of a unitary operator, which belong to the unit circle of the complex plane, does not need to be real.

10. (transformações de Cayley) A *transformação de Cayley*

$$z \mapsto \frac{z-i}{z+i}$$

é um automorfismo da esfera de Riemann $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$. Em particular, envia o semiplano superior $\mathbb{H} := \{z \in \mathbb{C} \quad \text{t.q.} \quad \Im(z) > 0\}$ no disco unitário $\mathbb{D} := \{z \in \mathbb{C}\,\text{t.q.} \quad |z| < 1\}$, e a "circunferência" $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\} \subset \overline{\mathbb{C}}$ (a fronteira ideal $\partial\mathbb{H}$ do "plano hiperbólico") na circunferência unitária $\mathbb{S} = \{z \in \mathbb{C} \quad \text{t.q.} \quad |z| = 1\}$. Esta ideia admite muitas generalizações.

Por exemplo, se S é uma matriz hemi-hermítica (ou hemi-simétrica), e portanto os seus valores próprios são imaginários puros, então $I \pm S$ são invertíveis. Em uma base ortonormada em que S é diagonal, $I \pm S$ são diagonais com valores próprios correspondentes (i.e. associados aos mesmos vetores próprios) da forma $1 \pm it$ com t real, ou seja, complexos conjugados. Mas os quocientes $(1 - it)/(1 + it)$ são unitários. De consequência, a *transformada de Cayley*[29]

$$U := (I - S)(I + S)^{-1}$$

[29] A. Cayley, Sur quelques propriétés des déterminants gauches, *Journal für die reine und angewandte Mathematik* **32** (1846), 119-123.

é uma matriz unitária (respetivamente, ortogonal).

- Mostre que se U é uma matriz unitária e se $I+U$ é invertível (i.e. se U não tem valor próprio -1) então $S=(I-U)(I+U)^{-1}$ é hemi-hermítica.
- Mostre que se A é uma matriz auto-adjunta então $U=(A-iI)^{-1}(A+iI)$ é unitária.

11. (forma normal de operadores ortogonais) Operadores unitários num espaço euclidiano real de dimensão finita são representados por matrizes ortogonais, satisfazendo $O^\top = O^{-1}$. O teorema espetral para operadores normais implica o seguinte resultado sobre a estrutura de uma matriz ortogonal.

Teorema 21.6. *Seja $U:\mathbf{E}\to\mathbf{E}$ um operador unitário de um espaço euclidiano real de dimensão finita $\mathbf{E}\approx\mathbb{R}^n$, representado numa base ortonormada por uma matriz ortogonal O. Então o espaço é uma soma direta $\mathbf{E}=\mathbf{E}_1\oplus\cdots\oplus\mathbf{E}_k$ de subespaços invariantes ortogonais de dimensão 1 ou 2, em cada um dos quais U é o operador $\pm 1:\mathbb{R}\to\mathbb{R}$ ou uma rotação $R_\theta:\mathbb{R}^2\to\mathbb{R}^2$ do plano de um ângulo $\theta\notin\pi\mathbb{Z}$.*

Existe portanto uma base ortonormada na qual U é representado pela matriz ortogonal

$$O=\begin{pmatrix} O_1 & 0 & \dots & 0 \\ 0 & O_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & O_k \end{pmatrix}$$

onde cada bloco não nulo O_i, de dimensão um ou dois, é uma das matrizes

$$(1) \qquad (-1) \qquad \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

Em particular, um operador unitário do espaço euclidiano tri-dimensional $\mathbb{R}^3$ (ou de um espaço de dimensão ímpar) admite sempre uma reta invariante, ou seja, um "eixo de rotação".

Demonstração. A prova é por indução. A matriz real O define também um operador unitário no espaço euclidiano complexo $\mathbb{C}^n$ (a complexificação de $\mathbb{R}^n$), munido do produto interno usual, poi $\overline{O}^\top = O^\top = O^{-1}$. O teorema espetral para operadores normais diz que $\mathbb{C}^n$ admite uma base ortonormada de vetores próprios de O, e os seus valores próprios são unitários, i.e. satisfazem $|\lambda|=1$.

Seja $Z=X+iY\in\mathbb{C}^n\approx\mathbb{R}^n+i\mathbb{R}^n$ um vetor próprio com valor próprio $\lambda=e^{i\theta}$, ou seja,

$$O(X+iY)=e^{i\theta}(X+iY)\,.$$

Se o valor próprio é real, ou seja $\lambda=e^{i\theta}=\pm 1$, então ou X ou Y (pois um dos dois não é nulo) é um vetor próprio real V com valor próprio λ. A restrição de O a reta $\mathbf{E}_1$ gerada por V é então uma multiplicação por ± 1, e o espaço ortogonal a esta reta é um subespaço invariante de dimensão $n-1$.

Se o valor próprio $\lambda=e^{i\theta}=\cos\theta+i\sin\theta$ não é real (ou seja, θ não é um múltiplo inteiro de π), então é imediato ver que também $\overline{Z}=X-iY$ é um valor próprio de O, com valor próprio $\overline{\lambda}=e^{-i\theta}$. Sendo $\lambda\neq\overline{\lambda}$, também é imediato ver que X e Y são linearmente independentes. Mas a parte real e a parte imaginária da equação $OZ=e^{i\theta}Z$ dizem que

$$OX=\cos\theta\, X-\sin\theta\, Y$$
$$OY=\sin\theta\, X+\cos\theta\, Y$$

A restrição de O ao plano $\mathbf{E}_1\approx\mathbb{R}^2$ gerado por X e Y é portanto uma rotação de um ângulo θ, e o espaço ortogonal a este plano é um subsespaço invariante de dimensão $n-2$.

Nos dois casos, a indução pode funcionar. □

12. (operadores positivos e raízes quadradas) Um operador auto-adjunto $A : \mathbf{H} \to \mathbf{H}$ definido num espaço euclidiano, real ou complexo, de dimensão finita é dito *positivo* ou *não-negativo*, notação $A \succ 0$ ou $A \succeq 0$, respetivamente, se os seus valores médios são positivos ou não negativos, ou seja, se
$$\langle A\mathbf{v}, \mathbf{v}\rangle > 0 \qquad \text{ou} \qquad \langle A\mathbf{v}, \mathbf{v}\rangle \geq 0$$
para todos os $\mathbf{v} \in \mathbf{H}$ (atenção, esta definição/condição não faz sentido para operadores que não são auto-adjuntos!). Pelo teorema espetral, estas condições são equivalentes a A ter valores próprios $\lambda > 0$ ou $\lambda \geq 0$, respetivamente. É evidente que um operador positivo é invertível.

    Se T é um operador arbitrário, então $A = T^*T$ é um operador auto-adjunto cujos valores médios são $\langle A\mathbf{v}, \mathbf{v}\rangle = \langle T\mathbf{v}, T\mathbf{v}\rangle = \|T\mathbf{v}\|^2$. Portanto $A = T^*T \succeq 0$, e $A \succ 0$ se T é invertível.

    De fato, todo operador não-negativo $A \succeq 0$ é desta forma. Numa base ortonormada em que a matriz que representa A é uma matriz diagonal $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ com $\lambda_k \geq 0$, podemos definir um operador $P := \sqrt{A}$ representado pela matriz $\mathrm{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n})$. Então é claro que P é auto-adjunto, não-negativo, comuta com A, i.e. $AP = PA$, e é uma "raiz quadrada" de A, i.e. $P^2 = A$. De forma análoga, é possível construir raizes quadradas positivas de operadores positivos.

    - Seja $T : \mathbf{H} \to \mathbf{H}$ um operador arbitrário. Então $T^*T \succeq 0$ e de consequência admite uma raiz quadrada $\sqrt{T^*T}$, que é auto-adjunta. Mostre que
    $$\|T\mathbf{v}\|^2 = \|\sqrt{T^*T}\mathbf{v}\|^2$$
    para todo $\mathbf{v} \in \mathbf{H}$.
    - Mostre que um operador A é (auto-adjunto e) positivo sse é da forma $A = T^*T$ com T invertível.
    - Mostre que se A é (auto-adjunto e) positivo então A^2 e A^{-1} são também auto-adjuntos e positivos.
    - Calcule uma raiz quadrada das seguintes matrizes hermíticas e não-negativas
    $$\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \qquad \dots$$
    - Seja $A : \mathbf{E} \to \mathbf{E}$ um operador linear positivo no espaço euclidiano real $\mathbf{E}$. Mostre que
    $$(\mathbf{x}, \mathbf{y}) := \langle A\mathbf{x}, \mathbf{y}\rangle$$
    é um produto interno em $\mathbf{E}$. Em particular, se o espaço $\mathbf{E}$ tem dimensão finita, existe uma base $\mathbf{e}_1, \dots, \mathbf{e}_n$ de $\mathbf{E}$ (ortonormada para o produto interno $(\cdot,\cdot)$) em que o operador A é representado pela matriz identidade I. De consequência, se $\mathbf{x} = x^1\mathbf{e}_1 + \cdots + x^n\mathbf{e}_n$, então
    $$\langle A\mathbf{x}, \mathbf{x}\rangle = x_1^2 + \cdots + x_n^2 \,.$$

13. (decomposição polar) Uma consequência da existência de raízes quadradas é a seguinte *decomposição polar* de um operador arbitrário num espaço euclidiano de dimensão finita.

    Teorema 21.7 (decomposição polar)**.** *Todo operador $T : \mathbf{H} \to \mathbf{H}$ de um espaço euclidiano de dimensão finita $\mathbf{H} \approx \mathbb{R}^n$ ou $\mathbb{C}^n$ pode ser decomposto como produto*
    $$T = UP$$
    de um operador unitário U e um operador auto-adjunto não-negativo $P = \sqrt{T^*T} \succeq 0$.

Demonstração. Seja $P = \sqrt{T^*T}$. Sendo $\|T\mathbf{v}\|^2 = \|\sqrt{T^*T}\mathbf{v}\|^2$ para todo $\mathbf{v} \in \mathbf{H}$ (exercício acima), temos que $\ker(T) = \ker(P)$. Em particular, se T é invertível, também P é invertível, e é imediato verificar que o operador $U := TP^{-1}$ é unitário (exercício). No caso geral, podemos definir um operador $U' : \mathrm{im}(P) \to \mathrm{im}(T)$ de acordo com $U'(P\mathbf{v}) := T\mathbf{v}$. Então para todo $\mathbf{u} = P\mathbf{v} \in \mathrm{im}(P)$, com $\mathbf{v} \in \mathbf{H}$,

$$\|U'\mathbf{u}\|^2 = \|T\mathbf{v}\|^2 = \|\sqrt{T^*T}\mathbf{v}\|^2 = \|P\mathbf{v}\|^2 = \|\mathbf{u}\|^2 \,,$$

ou seja, U' é uma isometria de $\mathrm{im}(P)$ sobre $\mathrm{im}(T)$ (pois têm a mesma dimensão). O operador U' pode ser extendido a um operador unitário $U = U' \oplus U''$ de $\mathbf{H} = \mathrm{im}(P) \oplus \mathrm{im}(P)^\perp$ sobre $\mathbf{H} = \mathrm{im}(T) \oplus \mathrm{im}(T)^\perp$ escolhendo uma isometria arbitrária U'' de $\mathrm{im}(P)^\perp$ sobre $\mathrm{im}(T)^\perp$ (que existe porque têm a mesma dimensão). $\square$

Em particular, toda matriz $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ invertível pode ser decomposta como o produto

$$A = OP$$

de uma matriz simétrica e positiva $P = \sqrt{A^*A} \succ 0$ e uma matriz ortogonal $O = AP^{-1}$.

14. (corda vibrante e harmónicas) As pequenas vibrações transversais de uma corda de comprimento ℓ, tensão k e densidade linear ρ são modeladas pela equação de onda

$$\frac{\partial^2 u}{\partial t^2} - c^2 \frac{\partial^2 u}{\partial x^2} = 0 \tag{21.3}$$

com condições de fronteira $u(0,t) = u(\ell,t) = 0$, onde $u(x,t)$ denota o deslocamento transversal da corda na posição $x \in [0,\ell]$ e no tempo t, e $c = \sqrt{k/\rho}$.

O produto $u(x,t) = X(x)\,T(t)$ é uma solução "separável" de (21.3) se $XT'' = c^2X''T$, e portanto se existe uma constante $\lambda \in \mathbb{R}$ tal que

$$X'' = \lambda X \qquad \text{e} \qquad T'' = \lambda c^2 T$$

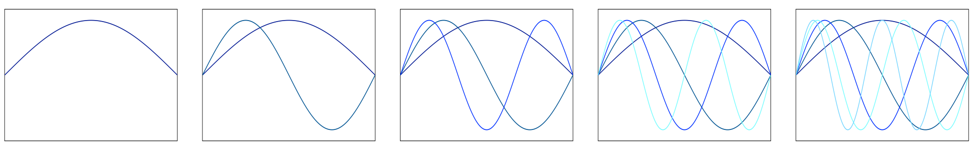
As únicas soluções não triviais do problema de Sturm-Liouville $X'' = \lambda X$ no intervalo $[0,\ell]$ com condições de fronteira nulas $X(0) = X(\ell) = 0$ são proporcionais a $X_n(x) = \sin(\pi n x/\ell)$ e têm valores próprios $\lambda_n := -\pi^2 n^2/\ell^2$ com $n = 1, 2, 3, \ldots$. As soluções separáveis do problema da corda vibrante são portanto as *ondas estacionárias*

$$\begin{aligned} u_n(x,t) &:= \Big(a_n \cos(2\pi\nu_n t) + b_n \sin(2\pi\nu_n t)\Big) \sin(2\pi x/\ell_n) \\ &= A_n \sin(2\pi\nu_n t + \tau_n) \sin(2\pi x/\ell_n) \,, \qquad \text{com } n = 1, 2, 3, \ldots \end{aligned}$$

onde os coeficientes a_n e b_n, ou a amplitude $A_n = \sqrt{a_n^2 + b_n^2}$, e a fase $\tau_n = \arctan(a_n/b_n)$ são constantes arbitrárias, e as *frequências próprias* e os *comprimentos de onda* são

$$\nu_n := \frac{c}{2\ell}\,n \qquad \text{e} \qquad \ell_n := \frac{2\ell}{n} \,, \qquad\qquad \text{com } n = 1, 2, 3, \ldots \,,$$

respetivamente. A primeira frequência, $\nu_1 = c/\ell_1$, é dita *som* (ou *tom*, ou *modo*) *fundamental*, e as outras, $\nu_n = n\nu_1 = c/\ell_n$, com $n = 2, 3, 4, \ldots$, são ditas n-ésimas *harmónicas* (ou *overtones*) da corda.



Primeiras 5 harmónicas de uma corda vibrante.

Por exemplo, se a fundamental é o A_4 de 440 Hz (é o caso da segunda corda de um violino), então a segunda harmónica é o A_5 de 880 Hz, a terceira está próxima do E_6 de 1318.5 Hz, a quarta é o A_6 de 1760 Hz, a quinta está próxima do $C\sharp_7$ de 2217.5 Hz, a sexta está próxima do E_7 de 2637 Hz, a sétima está próxima do G_7 de 3136 Hz, ... Em particular, as primeiras harmónicas contêm a "fundamental" A, a "quinta justa" E a "terça maior" C♯, as três notas ("tríade maior") do "acorde maior"!

Primeiras 12 harmónicas de uma corda cuja fundamental é C.

- A primeira corda de um violino, que tem comprimento 325 mm e costuma ser afinada com uma tensão de 70 N (ou seja, $\simeq 7.1$ Kg), vibra com frequências 660 Hz, 1320 Hz, 1980 Hz, ... O que deve fazer um violinista para obter o Lá5 de 880 Hz com esta corda?

15. (equação de Schrödinger estacionária) Considere a *equação de Schrödinger estacionária*

$$-\frac{\hbar^2}{2m}\Delta\psi = E\psi$$

para a função de onda $\psi(x)$ de uma partícula livre, onde m é a massa da partícula, e $\hbar = h/2\pi$ é a constante de Planck reduzida ($h \simeq 6.262\cdots \times 10^{-34}$ J·s).

- Determine para quais valores E da energia existem soluções não triviais da equação no intervalo $x \in [0, \ell]$ com condições de fronteira $\psi(0) = 0$ e $\psi(\ell) = 0$ (partícula numa caixa).

16. (determinant and zeta function) Let L be a self-adjoint operator with eigenvalues $\lambda_1, \lambda_2, \dots$. The logarithm of its determinant is the sum of the logarithms of its eigenvalues, since

$$\log\det L = \log(\lambda_1\lambda_2\dots) = \log\lambda_1 + \log\lambda_2 + \dots$$

(which of course may be divergent!). Define the *zeta function* of the operator L as the Dirichlet series

$$\zeta_L(s) := \sum_k \frac{1}{\lambda_k^s}\,,$$

which is, formally, $\zeta_L(s) = \mathrm{tr}(L^{-s})$. Here s is a complex variable, and it is understood that the zeta function if defined in a half-plane $\Re(s) > \sigma$ where the above sum is absolutely convergent. This is not an issue if the operator is finite-dimensional, since then $\zeta_L(s)$ is a finite sum, hence an entire function. In this case, a computation shows the remarkable identity

$$\log\det L = -\zeta_L'(0)\,.$$

Physicists use this identity to "extract" a finite value out of an infinite product that would be otherwise divergent, and call this "zeta function regularization".

22 Formas quadráticas e cónicas

ref: [Ap69] *Vol 2*, **5.12-18** ; [La87] *Ch. VIII*

1. (formas quadráticas reais) Uma *forma quadrática* em n variáveis reais é um polinómio homogéneo $\mathcal{Q}(x^1, x^2, \dots, x^n)$ de grau 2 nas coordenadas de $\mathbf{x} = (x^1, x^2, \dots, x^n) \in \mathbb{R}^n$, ou seja, um polinómio $\mathcal{Q} : \mathbb{R}^n \to \mathbb{R}$ de grau 2 tal que $\mathcal{Q}(\lambda\mathbf{x}) = \lambda^2\mathcal{Q}(\mathbf{x})$ para todo o escalar $\lambda \in \mathbb{R}$ e todo o $\mathbf{x} \in \mathbb{R}^n$. Esta condição diz que são nulos os coeficientes dos termos de grau um e zero, e que portanto o polinómio é da forma

$$\mathcal{Q}(\mathbf{x}) = \sum_{i,j} a_{ij} x^i x^j$$

com a_{ij} coeficientes reais.

Uma matriz quadrada $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ (e a escolha de uma base de $\mathbb{R}^n$, por exemplo a base canónica) define uma forma quadrática $A[X] := X^\top AX$, onde X é o vetor coluna que representa $\mathbf{x}$, ou seja,

$$\mathcal{Q}(\mathbf{x}) = X^\top AX$$

Dada uma forma quadrática $\mathcal{Q}$, existem muitas matrizes A tais que $\mathcal{Q}(\mathbf{x}) = X^\top AX$, sendo apenas fixados os elementos diagonais a_{ii}, que são os coeficientes de $(x^i)^2$, e as somas $a_{ij}+a_{ji}$, que são os coeficientes de $x^ix^j = x^jx^i$. Em particular, subtituindo A por $(A + A^\top)/2$, toda forma quadrática é definida por uma única matriz simétrica. O espaço das formas quadráticas em n variáveis reais pode portanto ser identificado com o espaço $\mathrm{Sym}_n(\mathbb{R})$ das matrizes simétricas reais $n \times n$.

A mudança de coordenadas $X \mapsto Y = U^{-1}X$, onde $X = UY$ e U é a matriz invertível cujas colunas são os vetores da nova base, transforma a forma quadrática $X^\top AX$ em

$$X^\top AX = (UY)^\top A\,(UY) = Y^\top\,(U^\top AU)\,Y = Y^\top BY$$

com $B := U^\top AU$. Ou seja, a matriz quadrada simétrica que representa a forma quadrática $X^\top AX$ na nova base é a matriz

$$\boxed{B = U^\top AU}$$

Duas matrizes quadradas A e B assim relacionadas são ditas *conjugadas*, e representam a mesma forma quadrática em sistemas de coordenadas diferentes.

- Determine a matriz simétrica que define as seguintes formas quadráticas

$$\mathcal{Q}(x, y) = x^2 - 2xy - y^2 \qquad \mathcal{Q}(x, y) = 2x^2 + 6xy + 7y^2$$

$$\mathcal{Q}(x, y) = xy \qquad \mathcal{Q}(x, y, z) = 2x^2 - yz$$

- Mostre que se A é simétrica então também $U^\top AU$ é simétrica.
- Seja $\mathcal{Q} : \mathbb{R}^n \to \mathbb{R}$ uma forma quadrática. Mostre que $\mathcal{B} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, definida por

$$\mathcal{B}(\mathbf{x}, \mathbf{y}) := \tfrac{1}{2}\left(\mathcal{Q}(\mathbf{x} + \mathbf{y}) - \mathcal{Q}(\mathbf{x}) - \mathcal{Q}(\mathbf{y})\right)$$

é uma "forma bilinear", ou seja, é linear em cada variável, e que a forma quadrática pode ser reconstruida como $\mathcal{Q}(\mathbf{x}) = \mathcal{B}(\mathbf{x}, \mathbf{x})$.

2. (diagonalização de Jacobi e lei de inércia de Sylvester) Seja $\mathcal{Q}(\mathbf{x}) = X^\top AX$ uma forma quadrática no espaço euclidiano real $\mathbb{R}^n$, definida pela matriz simétrica $A \in \mathrm{Sym}_n(\mathbb{R})$ relativamente à base canónica $\mathbf{e}_1, \dots, \mathbf{e}_n$ (ou outra base ortonormada), onde $\mathbf{x} = x^1\mathbf{e}_1 + \dots + x^n\mathbf{e}_n$ e X denota o vetor coluna $(x^1, \dots, x^n)$. Pelo teorema espetral, existe uma base ortonormada de vetores próprios $\mathbf{v}_1, \dots, \mathbf{v}_n$ de A, com valores próprios (reais) $\lambda_1, \dots, \lambda_n$, e portanto uma matriz ortogonal $U = (u^i{}_j) \in \mathbf{O}(n, \mathbb{R})$, cujas colunas são os vetores próprios (ou seja, $\mathbf{v}_j = \sum_i u^i{}_j\mathbf{e}_i$), tal que

$$U^\top A\,U = \Lambda := \mathrm{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)\,.$$

Nesta base, a forma quadrática é "diagonal". De fato, se $\mathbf{x} = y^1\mathbf{v}_1 + y^2\mathbf{v}_2 + \cdots + y^n\mathbf{v}_n$ e portanto o vetor coluna das novas coordenadas é $Y = U^\top X$, então

$$\begin{aligned}\mathcal{Q}(\mathbf{x}) &= X^\top AX = (UY)^\top A\,(UY) \\ &= Y^\top \Lambda Y = \lambda_1\,(y^1)^2 + \lambda_2\,(y^2)^2 + \cdots + \lambda_n\,(y^n)^2\,.\end{aligned}$$

Mais uma mudança de coordenadas $y^k \mapsto z^k := \sqrt{|\lambda_k|}\,y^k$, se $\lambda_k \neq 0$, e $y^k \mapsto z^k := y^k$ se $\lambda_k = 0$, definida por $Y \mapsto Z = C^{-1}Y$ com C matriz (diagonal) positiva (que não é ortogonal se pelo menos um dos valores próprios tem módulo $|\lambda_k| \neq 1$), pode ainda transformar a matriz Λ numa matriz diagonal $\Lambda' = C^\top \Lambda C$ com valores próprios 0 ou ± 1. Assim,

Teorema 22.1. *Seja $\mathcal{Q}(\mathbf{x}) = X^\top AX$ uma forma quadrática em $\mathbb{R}^n$. Existe uma base ortogonal (não necessariamente ortonormada!) na qual a forma quadrática é*

$$\mathcal{Q}(\mathbf{x}) = ((z^1)^2 + (z^2)^2 + \cdots + (z^{n_+})^2) - ((z^{n_++1})^2 + (z^{n_++2})^2 + \cdots + (z^{n_++n_-})^2)$$

É claro que as dimensões $n_\pm$ (o número de valores próprios positivos e negativos, respetivamente) e de consequência $n_0 := n - (n_+ + n_-)$ (a dimensão do núcleo de A), são invariantes da forma quadrática (*lei de inêrcia de Sylvester*).

Uma forma quadrática $\mathcal{Q}(\mathbf{x}) = X^\top AX$ é *positiva* se $\mathcal{Q}(\mathbf{x}) > 0$ para todo vetor $\mathbf{x} \neq 0$. É imediato verificar que a forma quadrática definida pela matriz simétrica A é positiva sse todos os valores próprios de A são positivos, e portanto se existe uma base ortogonal (mas não necessariamente ortonormada) na qual a forma quadrática assume a forma

$$\mathcal{Q}(\mathbf{x}) = (z^1)^2 + (z^2)^2 + \cdots + (z^n)^2\,.$$

Em outras palavras, a forma quadrática $\mathcal{Q}(\mathbf{x}) = X^\top AX$ é positiva sse existe uma matriz invertível C tal que $C^\top AC = I$.

- Diagonalize as seguintes formas quadráticas:

$$\mathcal{Q}(x,y) = x^2 + 2y^2 \qquad \mathcal{Q}(x,y) = xy \qquad \mathcal{Q}(x,y) = x^2 - 2xy + y^2$$

$$\mathcal{Q}(x,y,z) = x^2 - 2xy + 4yz + 6xz - 3z^2$$

- Mostre que a forma bilinear associada a uma forma quadrática positiva é um produto escalar.
- [Ap69] *Vol. 2* **5.15.**

3. (quociente de Rayleigh) Seja $A \in \mathrm{End}(\mathbf{E})$ um operador simétrico do espaço euclidiano de dimensão finita $\mathbf{E} \approx \mathbb{R}^n$, definido numa base ortonormada (e.g. a base canónica) pela matriz simétrica $A = (a^i{}_j) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, e seja

$$\mathcal{Q}_A(\mathbf{x}) := \langle A\mathbf{x}, \mathbf{x}\rangle$$

a forma quadrática associada. Os pontos críticos da restrição da forma quadrática $\mathcal{Q}_A$ à esfera unitária $\mathbb{S}^{n-1} := \{\mathbf{x} \in \mathbb{R}^n \text{ t.q. } \|\mathbf{x}\|^2 = 1\}$ são os pontos onde o gradiente de $\mathcal{Q}_A(\mathbf{x})$, que é igual a $\nabla\mathcal{Q}_A(\mathbf{x}) = \frac{1}{2}A\mathbf{x}$, é proporcional ao gradiente de $\|\mathbf{x}\|^2$ (que é $\frac{1}{2}\mathbf{x}$), ou seja os pontos onde existe $\lambda \in \mathbb{R}$ (o multiplicador de Lagrange) tal que $A\mathbf{x} = \lambda\mathbf{x}$. Em particular, um ponto $\mathbf{v} \in \mathbb{S}^{n-1}$ onde $\mathcal{Q}$ é máximo/mínimo é um vetor próprio de A com valor próprio $\lambda = \mathcal{Q}_A(\mathbf{v})$, e este valor próprio é o maior/menor dos valores próprios de A. Em particular, uma matriz real/operador simétrica/o admite pelo menos um valor próprio (pois a esfera é compacta, e a função $\mathcal{Q}_A$ é contínua).

O *quociente de Rayleigh* da matriz/operador simétrico A é a função

$$\mathcal{R}_A(\mathbf{x}) := \frac{\mathcal{Q}_A(\mathbf{x})}{\|\mathbf{x}\|^2} = \frac{\langle A\mathbf{x}, \mathbf{x}\rangle}{\langle \mathbf{x}, \mathbf{x}\rangle}$$

definida em $\mathbb{R}^n\backslash\{0\}$.

- Use os multiplicadores de Lagrange para determinar os extremos locais de $\mathcal{Q}_A$ na esfera unitária.
- Em alternativa (se não sabe o que são os multiplicadores de Lagrange), assuma que $\mathbf{x} \in \mathbb{S}^{n-1}$ é um extremo local de $\mathcal{Q}_A$ na esfera unitária. Os "equatores" que passam por $\mathbf{x}$ são os caminhos $\mathbf{r}(t) = \cos(t)\,\mathbf{x} + \sin(t)\,\mathbf{v} \in \mathbb{S}^{n-1}$, onde $\mathbf{v}$ é um vetor unitário do hiperplano $\mathbf{x}^\perp$ (observe que $\mathbf{x} + \mathbf{x}^\perp$ é o plano tangente à esfera unitária no ponto $\mathbf{x}$). Calcule a derivada da função $t \mapsto f(t) := \mathcal{Q}_A(\mathbf{r}(t))$ no instante $t = 0$ (usando a simetria de A). Mostre que se $\mathbf{x}$ é um extremo local, então a derivada $f'(0)$ é igual a zero para todos os $\mathbf{v} \in \mathbf{x}^\perp$, e isto implica que $A\mathbf{x}$ é proporcional a $\mathbf{x}$. Deduza que $\mathbf{x}$ é um vetor próprio de A.
- Mostre que o quociente de Rayleigh $\mathcal{R}_A(\mathbf{x})$ de um vetor próprio $\mathbf{x}$ é igual ao valor próprio associado λ.
- Deduza que o quociente de Rayleigh atinge máximo e mínimo, e que estes valores são o maior e o menor valor próprio de A, respetivamente.

4. (forma normal de um par de formas quadráticas) Sejam $\mathcal{K}(\mathbf{x}) = X^\top KX$ e $\mathcal{P}(\mathbf{x}) = X^\top PX$ duas formas quadráticas em $\mathbb{R}^n$, representadas (por exemplo, na base canónica) pelas matrizes simétricas K e P, respetivamente. Se $\mathcal{K}$ é positiva, então a matriz simétrica K define um produto interno $\langle \mathbf{x}, \mathbf{y} \rangle := Y^\top KX$ em $\mathbb{R}^n$, e portanto existe uma base (ortonormada para este produto interno) na qual K é representada pela matriz identidade. Ou seja, existe uma matriz invertível C tal que, se $X = CY$, então $C^\top KC = I$ e portanto $X^\top KX = Y^\top Y$. Nesta base, a segunda forma quadrática é representada pela uma matriz simétrica $P' = C^\top PC$, i.e. $X^\top PX = Y^\top P'Y$. Pelo teorema espetral, existe uma matriz ortogonal O que diagonaliza P', i.e. tal que $\Lambda = O^\top P'O$ é diagonal. Observe que a primeira forma quadrática continua sendo definida pela matriz identidade, pois $O^\top IO = I$. Portanto, nas coordenadas $Z = (CO)^{-1}X$, as formas quadráticas $\mathcal{K}$ e $\mathcal{P}$ são $Z^\top Z$ e $Z^\top \Lambda Z$, respetivamente.

   Teorema 22.2 (forma normal de um par de formas quadráticas)**.** *Sejam $\mathcal{K}$ e $\mathcal{Q}$ duas formas quadrática em $\mathbb{R}^n$. Se $\mathcal{K}$ é positiva, então existe uma base de $\mathbb{R}^n$ em que as formas são definidas pela matriz identidade e por uma matriz diagonal, respetivamente, ou seja, existem coordenadas $z_1, \dots, z_n$ nas quais as formas quadráticas assumem a forma*

   $$\mathcal{K}(\mathbf{z}) = z_1^2 + z_2^2 + \cdots + z_n^2 \qquad e \qquad \mathcal{Q}(\mathbf{z}) = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \cdots + \lambda_n z_n^2$$

   onde $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_n$.

   Cuidado! Este resultado não diz que as matrizes K e P são simultaneamente diagonalizáveis (caso que implica $KP = PK$), mas apenas que as fomas quadráticas que definem são diagonais numa base comum! O que acontece é que são diagonais simultaneamente as matrizes $(CO)^\top K(CO)$, que de fato é a matriz identidade, e $(CO)^\top P(CO)$ (e estas fórmulas usam transpostas de CO e não inversas!). A confusão pode nascer do fato que a matriz que define uma forma quadrática não é conceptualmente igual a uma matriz que define uma endomorfismo de $\mathbb{R}^n$ ...

5. (pequenas oscilações e frequências próprias) Numa vizinhança de um ponto de equilíbrio (um ponto onde a força, ou seja, o gradiente do potencial, é nula), e que podemos assumir ser a origem do sistema de coordenadas generalizadas, a energia potencial de um sistema mecânico pode ser aproximada por uma forma quadrática

   $$\mathcal{U} \simeq \frac{1}{2}\,Q^\top AQ\,,$$

   onde $A = (a_{ij}) := (\partial^2\mathcal{U}/\partial q_i \partial q_j(0))$ é a matriz Hessiana do potencial $\mathcal{U}(\mathbf{q})$ na origem (que é positiva se a origem é um mínimo local). Por outro lado, a energia cinética é uma forma quadrática positiva

   $$\mathcal{K} = \frac{1}{2}\,\dot{Q}^\top K\dot{Q}$$

nas velocidades generalizadas, definida por uma matriz simétrica e positiva $K = (k_{ij})$. A Lagrangiana do sistema é $\mathcal{L} = \mathcal{K} - \mathcal{U}$, e as equações de Euler-Lagrange (ou seja, as equações de movimento de Newton) são, nesta aproximação,

$$K\ddot{Q} = -AQ$$

Pelo teorema 22.2, existe uma transformação linear (não necessariamente ortogonal!) de coordenadas $Q \mapsto Z = C^{-1}Q$, tal que $C^\top KC = I$ e $C^\top AC = \Lambda$ é uma matriz diagonal, cujos valores próprios $\lambda_1 \leq \cdots \leq \lambda_n$ são positivos se a origem for um mínimo local não degenerado do potencial. De consequência, as equações de movimento assumem a forma

$$\ddot{Z} = -\Lambda Z \,,$$

ou seja, ficam decompostas nas n equações independentes

$$\ddot{z}_k = -\omega_k^2 \, z_k \qquad k = 1, 2, \dots, n$$

que são n osciladores harmónicos *frequências próprias* $\omega_k := \sqrt{\lambda_k}$. As trajetórias são combinações lineares de n oscilações

$$\mathbf{r}(t) = A_1 \cos(\omega_1 t + \phi_1)\, \mathbf{e}_1 + A_2 \cos(\omega_2 t + \phi_2)\, \mathbf{e}_2 + \cdots + A_n \cos(\omega_n t + \phi_n)\, \mathbf{e}_3$$

com amplitudes A_k e fases iniciais ϕ_k, sendo os $\mathbf{e}_k$ os vetores próprios de Λ. Os movimentos são periódicos ou quase-periódicos, dependendo se as frequências próprias ω_1, ω_2 $\dots \omega_n$ são racionalmente dependentes ou não.

6. (tensor de inércia e eixos principais) A energia cinética de rotação de um corpo rígido é

$$E = \frac{1}{2} \langle \boldsymbol{\omega}, I\boldsymbol{\omega} \rangle$$

onde $\boldsymbol{\omega}$ é a velocidade angular e I é o *tensor de inércia*, uma matriz simétrica (e positiva). Pelo teorema espetral existe um referencial em que o tensor de inércia é diagonal. Os vetores próprios são os “axes principais”, e os valores próprios são os momentos de inércia principais.

...

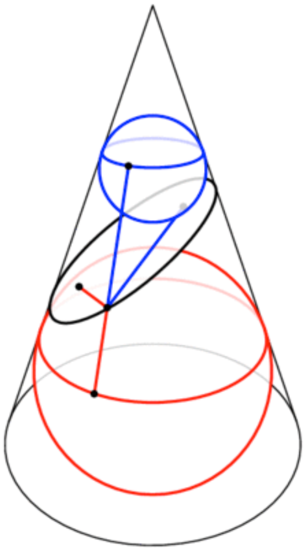
7. (pseudo-métrica de Minkowsky)

$$ds^2 = c^2 dt^2 - (dx^2 + dy^2 + dz^2)$$

...

8. (conic sections) The first, tautological, definition is the following: a *conic section* is the intersection between a (right circular) cone $C \subset \mathbb{R}^3$ and a plane $P \subset \mathbb{R}^3$ (not passing through the vertex, for otherwise we have the degenerate cases of a point, a line or two lines).

There are other definitions, much more useful in physical applications, which were already known to Apollonius of Perga and Pappus of Alexandria. The modern route to the understanding of them passes through the construction of the *Dandelin spheres*. These are spheres tangent to both the cone C and the plane P, inside the cone itself. There are two of them, say $S_\pm$, in the case of an ellipse (one on each side of the plane) or an hyperbola (one in each branch of the hyperbola), and only one for a parabola (say, the one with “+”). The points where the Dandelin spheres touch the plane P, say $F_\pm := S_\pm \cap P$, are called *foci* of the conic section. It is clear that when the plane P is orthogonal to the axis of the cone, the two foci coincide and the conic section is a circle. Each Dandelin sphere touches the cone at a circle $C_\pm$, belonging to a certain plane $P_\pm$ (perpendicular to the axis of the cone, hence horizontal in the picture), and the intersection of each of those planes with P determines a line $D_\pm := P_\pm \cap P$, called *directrix* of the conic section.



Dandelin spheres of an ellipse.

Second definition: *focal properties*. Consider a moving point $\mathbf{r}$ in the Euclidean plane $P \approx \mathbb{R}^2$. Let $f_\pm := \mathrm{d}(\mathbf{r}, F_\pm)$ denote the distances between $\mathbf{r}$ and a foci $F_\pm$, and let $\delta_\pm := \mathrm{d}(\mathbf{r}, D_\pm)$ denote the distances between $\mathbf{r}$ and the directrices $D_\pm$. Then, the moving point $\mathbf{r}$ describes

- an *ellipse* iff $f_+ + f_- = \text{constant}$;
- an *hyperbola* iff $|f_+ - f_-| = \text{constant}$;
- a *parabola* iff $f_+ = \delta_+$.

Third definition: eccentricity. The three conditions above may be merged into one single condition relating the distances between the moving point and one focus or one directrix, respectively. A moving point $\mathbf{r}$ in the plane describes a conic section if the ratio $e := f_+/\delta_+$ is constant, i.e.

$$f_+ = e\,\delta_+ \,. \tag{22.1}$$

The constant ratio e is called *eccentricity* of the conic section. One has an ellipse if $e < 1$, an hyperbola if $e > 1$, and a parabola if $e = 1$.

Polar equation. Consider the Euclidean plane $\mathbb{R}^2$, with coordinates x-y. Modulo a translation, we may assume that one of the foci is at the origin, say $F_+ = (0, 0)$. Modulo a rotation, we may also assume that the directrix is a vertical line $D_+ = \{x = d\}$, for some $d = \mathrm{d}(F_+, D_+) > 0$. Then, if the moving point has polar representation $\mathbf{r} = re^{i\theta} = (r\cos(\theta), r\sin(\theta))$, the defining equation (22.1) reads

$$r = e\,|r\cos(\theta) - d|\,.$$

Solving for r, we get

$$\boxed{r = \frac{ed}{e\cos(\theta) \pm 1}}$$

If $e \leq 1$, the only solution is the one with the $+$ sign, and the curve is an ellipse ($e < 1$) or a parabola ($e = 1$). If $e > 1$, the curve is a hyperbola, with the two solutions corresponding to its two branches. It is this form of the conics which appears when solving Kepler's problem.

Cartesian equations. Good looking Cartesian equations are those which are symmetric w.r.t. to the origin, i.e. such that $\mathbf{r}$ and $-\mathbf{r}$ both belong to the curve. The canonical form of an ellipse or a hyperbola is then

$$\frac{x^2}{a^2} + \frac{y^2}{a^2(1-e^2)} = 1\,.$$

The foci are $F_\pm = (\pm ae, 0)$, while the directrices are the vertical lines $D_\pm = \{x = \pm a/e\}$. If $e < 1$, hence it is an ellipse, we may write

$$\boxed{\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1} \tag{22.2}$$

where $a > 0$ and $b = a\sqrt{1-e^2} \leq a$ are the semi-axis, and $F_\pm = (\pm c, 0)$, with $c = \sqrt{a^2 - b^2}$, are the foci. The case $e = 0$, hence $a = b$ and $c = 0$, is a circle centered at the origin (an ellipse seen from a large distance). If $e < 1$, hence it is a hyperbola, we may write

$$\boxed{\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1} \tag{22.3}$$

with $a > 0$ and $b = a\sqrt{e^2 - 1}$. The foci are $F_\pm = (\pm c, 0)$ with $c = \sqrt{a^2 + b^2}$. The canonical form of a parabola, with focus at $F = (h, 0)$ and directrix $D = \{x = -d\}$ is

$$\boxed{y^2 = 4dx} \tag{22.4}$$

- Use the Dandelin spheres to prove that the tautological definition of the conic sections implies their focal properties.
- Check that the canonical Cartesian equations (22.2), (22.3) and (22.4) of the conic sections satisfy the focal properties as well as (22.1).
- Compute the area of the region bounded by the ellipse of (22.2).

9. (degree 2 equations in the plane and conics) Consider a degree two equation in two real variables, x and y, like

$$ax^2 + 2bxy + cy^2 + \alpha x + \beta y + \gamma = 0\,, \tag{22.5}$$

where $a, b, \ldots, \gamma$ are real coefficients. Our task is to change coordinates and reduce (22.5) to canonical form.

According to the spectral thorem, the quadratic form $\mathcal{Q}(x, y) = ax^2 + 2bxy + cy^2$ may be diagonalized by a orthogonal change of variables. In the new variables, say x' and y', related to the old ones by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

for some orthogonal matrix $C = \left(\begin{smallmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{smallmatrix}\right)$, the equation (22.5) reads

$$\lambda_1 (x')^2 + \lambda_2 (y')^2 + \alpha' x' + \beta' y' + \gamma' = 0\,,$$

where λ_1 and λ_2 are the eigenvalues of the symmetric matrix

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

inducing $\mathcal{Q}$. Observe that the product of the eigenvalues is equal to the determinant of the symmetric matrix: $\lambda_1\lambda_2 = ac - b^2 = \det A$. Completing the squares, and after a further translation of the form $x'' = x' - \xi$ and $y'' = y' - \eta$, this finally gives

$$\lambda_1 (x'')^2 + \lambda_2 (y'')^2 = \delta\,,$$

if both $\lambda_1 \neq 0$ and $\lambda_2 \neq 0$. This is an ellipse if $\lambda_1\lambda_2 > 0$ and δ has the right (opposite) sign, a hyperbola if $\lambda_1\lambda_2 < 0$ and $\delta \neq 0$, or some degenerate conics like a point (as $x^2 + y^2 = 0$), a couple of lines (as $x^2 - y^2 = 0$) or the empty set (as $x^2 + y^2 = -1$). If one of the eigenvalues vanishes, say $\lambda_1 = 0$, we are left with a parabola

$$\lambda_2 (y'')^2 = \delta x''\,,$$

or a line if also $\delta = 0$. If both the eigenvalues are zero, i.e. $\lambda = \mu 0$, we are left with a degree one equation $\alpha' x' + \beta' y' + \gamma' = 0$, hence an affine line.

In particular, if we know a priori that the conic defined by (22.5) is non-degenerate, its type is determined by the sign of the determinant $d := \det A = ac - b^2 = \lambda_1 \lambda_2$ of the matrix of the quadratic form: we have an ellipse if $d > 0$, and hyperbola if $d < 0$, and a parabola if $d = 0$.

- Write a Cartesian equation of a right circular cone $C \subset \mathbb{R}^3$. ...
- Find the conics defined by the following equations:

$$x^2 + xy + 2x = 0 \qquad \text{... more exercises ...}$$

- [Ap69], *vol. 2*, **5.15.**

10. (quadrics) *Quadrics* are surfaces defined, in the Euclidean space $\mathbb{R}^3$, by a degree 2 Cartesian equation like

$$ax^2 + by^2 + cz^2 + 2dxy + 2eyz + 2fxz + \alpha x + \beta y + \gamma z + \delta = 0$$

A similar procedure, diagonalization of the quadratic form $\mathcal{Q}(x, y, z)$ (the degree 2 part of the polynomial above) and then a translation (completing squares), shows that any non-degenerate quadric (not reduced to the empty set, a point, lines or planes, conics times lines, ...) is equivalent, modulo an isometry, to one of the following models:

an *ellipsoid*

$$\boxed{\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1}$$

an *hyperbolic hyperboloid*

$$\boxed{\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1}$$

and *elliptic hyperboloid*

$$\boxed{\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1}$$

an *elliptic paraboloid*

$$\boxed{\frac{x^2}{a^2} + \frac{y^2}{b^2} - z = 0}$$

an *hyperbolic paraboloid*

$$\boxed{\frac{x^2}{a^2} - \frac{y^2}{b^2} - z = 0}$$

(above, of course, all the parameters a, b, c are positive numbers).

- You may play with surfer to visualize quadrics, or create new amazing surfaces.

11. (motion in a central force and Kepler problem) Consider the Newton equation

$$m\ddot{\mathbf{r}} = F(r)\,\frac{\mathbf{r}}{r} \tag{22.6}$$

describing the motion of a particle/planet of mass $m > 0$ in a central force field $\mathbf{F}(\mathbf{r}) = F(r)\,\frac{\mathbf{r}}{r}$. Above, we use the traditional notation $r := \|\mathbf{r}\|$ for the length of the vector $\mathbf{r} \in \mathbb{R}^3$. If $\mathbf{v} = \dot{\mathbf{r}}$ denotes the velocity vector, then a computation shows that the angular momentum $\mathbf{L} := \mathbf{r} \times \mathbf{v}$ is a constant of the motion. If at some (initial) time the vectors $\mathbf{r}$ and $\mathbf{v}$ are not parallel, then $\mathbf{L} \neq \mathbf{0}$ and the motion occurs in the plane orthogonal to $\mathbf{L}$. We may therefore choose a reference Cartesian system $(\mathbf{i}, \mathbf{j}, \mathbf{k})$ in which $\mathbf{L} = L\,\mathbf{k}$ for some $L > 0$, and write the position vector as $\mathbf{r}(t) = \rho\cos(\theta)\,\mathbf{i} + \rho\sin(\theta)\,\mathbf{j}$ for some time-dependent angle θ and lenght $\rho = \|\mathbf{r}\|$. In polar coordinates Newton equation (22.6) reeds

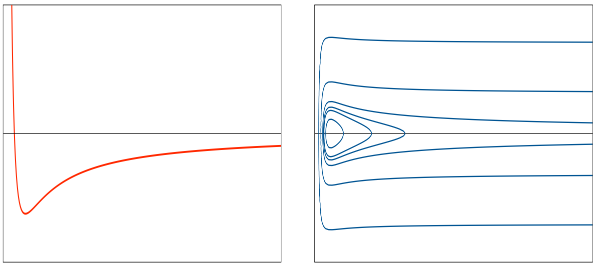
$$\begin{aligned} \ddot\rho - \rho\dot\theta^2 &= F(\rho)/m \\ \rho\ddot\theta + 2\dot\rho\dot\theta &= 0\,. \end{aligned} \tag{22.7}$$

The second equation (22.7) says that the "areal velocity" ("velocidade areal") $\ell := \rho^2\dot\theta$ is a constant of the motion, and this is *Kepler's second law* (which therefore holds for all central forces). We specialize now to Newton's gravitational force $F(\rho) = -\frac{GmM}{\rho^2}$, where M is the mass of the Sun and G is the gravitational constant. It may be observed that the first equation (22.7) then reads

$$m\ddot\rho = -\frac{\partial}{\partial\rho}V_\ell\,(\rho)\ ,$$

where we defined the "effective potential energy" as

$$V_\ell\,(\rho) := \tfrac{1}{2}m\ell^2/\rho^2 - GmM/\rho\,.$$



Kepler's effective potential and some energy level sets.

The conserved energy is therefore

$$E = \tfrac{1}{2}m\dot\rho^2 + \tfrac{1}{2}m\ell^2/\rho^2 - GmM/\rho\,.$$

Now we set $\rho = 1/x$ and look for a differential equation for x as a function of θ. Computation shows that $dx/d\theta = -\dot\rho/\ell$, and, using conservation of ℓ, that $d^2x/d\theta^2 = -\rho^2\ddot\rho/\ell^2$. There follows that the first Newton equation (22.7) reads

$$\frac{d^2x}{d\theta^2} + x = -\frac{GM}{\ell^2}\,.$$

This is an harmonic oscillator with unit frequency forced by a constant force, and its general solution is

$$x\,(\theta) = \frac{GM}{\ell^2}\,(1 + e\cos{(\theta - \theta_0)})$$

for some constants e and θ_0. Back to the original radial variable we get the solution

$$\rho\,(\theta) = \frac{\ell^2/GM}{1 + e\cos{(\theta-\theta_0)}}\ ,$$

Hence, orbits are conic sections with eccentricity e and focus at the origin. We get an ellipse for $0 \le e < 1$, corresponding to negative energy, hence to planets, and this is *Kepler's first law*. We get a parabola for $e = 1$, corresponding to zero energy, or an hyperbola for $e > 1$, corresponding to positive energy.

12. (hodograph and Hamilton's theorem) [30] [31]

13. (Kepler & Hooke) [32] [33]

[30] W.R. Hamilton, The hodograph or a new method of expressing in symbolic language the Newtonian law of attraction, *Proc. Roy. Irish Acad.* **3** (1846), 344-353.

[31] J. Milnor, On the geometry of the Kepler problem, *Amer. Math. Monthly* **90** (1983), 353-365.

[32] V.I. Arnold, *Huyghens & Barrow, Newton & Hooke*, Birkhäuser, 1990.

[33] T. Levi-Civita, Sur la régularisation du problème des trois corps, *Acta Math.* **42** (1920), 99-144.

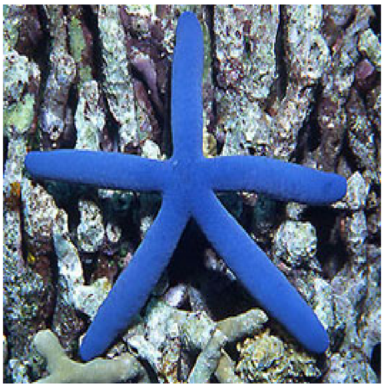
23 Grupos de matrizes

ref: [Ap69] *Vol 2*, **5.11**, **5.20**

1. (symmetries and transformations) Symmetries are, roughly speaking, movements of the background space (whatever it means) which leave unchanged the figure/pattern we are interesting in. The concept that capture this vague idea is that of a "group" (discovered/invented by Lagrange, Abel, Ruffini, Galois, ... in the early XIX century to understand polynomials, and rooted in modern geometry by the work of Klein, Hilbert, Poincaré, ...).

   The basic and concrete idea of a group is that of a family of "transformations", of some space or whatever, that can be "composed" (i.e. may act one after another) to produce new transformations, and that can be "undone". The transformation which undo a given one is called its "inverse". In particular, a transformation and its inverse may be composed to give the "identity" transformation, the transformation which doesn't do anything.

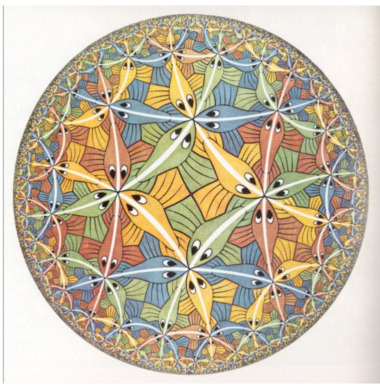
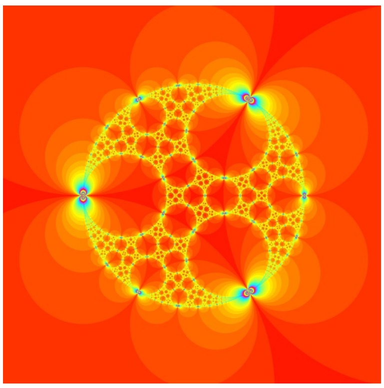
   - Describe the symmetries of the following figures.

   - Describe the symmetries of the following figures.

   - Try to describe the symmetries of those other figures (but don't feel discouraged if you don't succeed!).

2. (symmetries for laws of physics) Symmetries also play a fundamental role in our description and understanding of Nature.

   In classical mechanics, they are responsible for conservation laws, as follows from the celebrated Noether's theorem[34]. For example, time invariance of the Lagrangian implies conservation of energy, translation invariance implies conservation of the linear momentum, rotational invariance implies conservation of the angular momentum, and so on ... (see [LL78])

   Certain groups of isometries of the Euclidean 3-dimensional space appear as symmetries of crystals (and are therefore named "crystallographic groups"). Euclidean symmetries also appear when looking at snowflakes, or other chemical or biological structures ...

   Even more fundamental is the role played by (gauge) symmetries in quantum field theory. For example, in the Standard Model there appears the group $SU(3) \times SU(2) \times U(1) \dots$

3. (grupos) Um *grupo* é um conjunto G munido de uma (operação binária) lei de composição interna $G \times G \to G$, que associa a cada par ordenado $(g, g') \in G \times G$ um elemento $g \cdot g' \in G$ (ou simplesmente gg'), que verifica os seguintes axiomas:

   G1 *(propriedade associativa)* $(gg')g'' = g(g'g'')$, para todos $g, g', g'' \in G$.

   G2 *(existência do elemento neutro)* existe um elemento $e \in G$, chamado "elemento neutro", tal que $eg = ge = g$ para todo $g \in G$.

   G3 *(existência do inverso)* para todo $g \in G$ existe um elemento $g^{-1} \in G$, chamato "inverso de g", tal que $gg^{-1} = g^{-1}g = e$.

   Sei a lei de composição satisfaz também o axioma

   G4 *(comutatividade)* $gg' = g'g$, para todos $g, g' \in G$,

   então o grupo é dito *comutativo*, ou *abeliano*. Em um grupo abeliano costuma ser usada a notação "aditiva" $g + g'$ para a composição de g e g', $-g$ para o inverso de g, e consequentemente o elemento neutro costuma ser denotado por 0.

   Em um grupo (abeliano ou não), as equações (na incógnita x, dados $a, b \in G$)

   $$ax = b \qquad \text{e} \qquad ya = b$$

   admitem sempre soluções únicas, dadas por $x = a^{-1}b$ e $y = ba^{-1}$, respetivamente (que são iguais se o grupo é abeliano). Em particular, o elemento neutro é único, assim como o inverso de cada $g \in G$.

   Dado um elemento g de um grupo, é possível definir as suas potências g^n, com $n \in \mathbb{Z}$. basta definir $g^0 = e$ e $g^{n+1} := g \cdot g^n$ se $n \geq 0$, e as potências negativas por $g^{-n} := (g^{-1})^n$. É imediato verificar que vale a lei dos expoentes $g^n g^m = g^{n+m}$.

   Na prática, os elementos de um grupo podem ser "parametrizados" por um conjunto de "parâmetros" α, que podem ser números, ângulos, ..., assim que um grupo aparece como um conjunto $G = \{g_\alpha\}$, com α em um certo espaço $\mathcal{A}$. Em particular, pode acontecer que um grupo seja *finito*, ou seja, formado por um número finito de elementos, $G = \{g_0 = e, g_1, \dots, g_{n-1}\}$. Neste caso, o número $|G| := n$ é chamado *ordem* do grupo (finito) G. A lei de composição de um grupo finito pode ser apresentada na forma de uma "tabuada de multiplicar", uma matriz quadrada $n \times n$ onde no elemento da i-ésima linha e da j-ésima coluna aparece o produto $g_i g_j$.

   - Os conjuntos $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ e $\mathbb{C}$, munidos da lei "adição" $a, b \mapsto a + b$, são grupos abelianos.
   - Todo espaço vetorial $\mathbf{V}$ é um grupo abeliano, se munido da lei "adição" $\mathbf{v}, \mathbf{w} \mapsto \mathbf{v} + \mathbf{w}$.
   - Os conjuntos $\mathbb{Q}^\times := \mathbb{Q}\backslash\{0\}$, $\mathbb{R}^\times := \mathbb{R}\backslash\{0\}$ e $\mathbb{C}^\times := \mathbb{C}\backslash\{0\}$, munidos da lei "multiplicação" $a, b \mapsto a \cdot b$, são grupos abelianos.
   - Determine a tabuada de multiplicar do mais simples dos grupos não triviais, um grupo formado por apenas dois elementos, a identidade e e um outro elemento $g \neq e$.

[34] E. Noether, Invariante Variationsprobleme, *Nachr. D. König. Gesellsch. D. Wiss. Zu Göttingen, Math-phys. Klasse.* (1918), 235-257.

- Mostre que na tabuada de multiplicar de um grupo finito cada coluna ou cada linha contém cada elemento do grupo exatamente uma vez.
- As rotações do plano $\mathbb{R}^2$ formam um grupo comutativo. De fato, a composição de duas rotações R_θ e R_ϕ, de ângulos θ e ϕ, é uma rotação $R_\theta\, R_\phi = R_{\theta+\phi}$ de um ângulo $\theta+\phi$.
- As rotações do espaço euclidiano $\mathbb{R}^3$ também formam um grupo. Mostre, com um exemplo, que o grupo das rotações em dimensão 3 não é comutativo.

4. (permutações e grupo simétrico) Seja X um conjunto não vazio. As *permutações* de X são as transformações invertíveis $f : X \to X$. O conjunto $\mathrm{Per}(X)$ das permutações, munido da lei de composição $fg := f \circ g$, definida por $(f \circ g)(x) = f(g(x))$, é um grupo (a associatividade do produto num gropo é a propriedade natural da composição entre transformações). A identidade de $\mathrm{Per}(X)$ é a transformação identidade, definida por $1(x) = x$ para todo o $x \in X$. O elemento inverso de f é a transformação inversa $f^{-1} : X \to X$, tal que $f^{-1}(f(x)) = x$ e $f(f^{-1}(x)) = x$ para todo $x \in X$.

   Por exemplo, se $X_n \approx \{1, 2, \dots, n\}$ é um conjunto finito formado por n elementos, então $S_n := \mathrm{Per}(X_n)$ é um grupo finito chamado *grupo simétrico*. É claro que o grupo. simétrico S_n é formado por $n! = n \cdot (n-1) \dots 2 \cdot 1$ elementos (n possibilidades para escolher a imagem de 1, $n-1$ possibilidades para escolher a imagem de 2, ...) .

   - Calcule a tabuada de multiplicar do grupo S_2.
   - Mostre que o grupo S_3 não é comutativo.

5. (grupo diedral) O grupo das simetrias de um polígono regular de $n \geq 3$ lados é chamado *grupo diedral* (em inglês, *dihedral*), e denotado por D_n. Contém as n rotações de ângulos $2\pi k/n$, com $k = 0, 1, 2, \dots, n-1$ (ou seja, a rotação r de um ângulo $2\pi/n$ e as suas potências $r^2, r^3, \dots, r^n = e$), e n reflexões nos eixos de simetria do polígono (as n medianas, quando n é ímpar, ou as $n/2$ diagonais e as $n/2$ retas que unem os pontos médios dos lados opostos, quando n é par). De consequência, D_n é um grupo finito de ordem $2n$.

   - Verifique que a composição de duas reflexões é uma rotação, possivelmente trivial (pois as reflexões são "involuções", i.e. o quadrado de uma reflexão g é $g^2 = e$).
   - Verifique que o grupo diedral não é comutativo.

6. (grupo afim) Translações $z \mapsto z + a$ e homotetias $z \mapsto \lambda z$ geram o *grupo afim* do plano complexo, o grupo $\mathrm{Aff}(\mathbb{C})$ das transformações

$$z \mapsto g_{\lambda,a}(z) := \lambda z + a\,,$$

   com $a \in \mathbb{C}$ e $\lambda \in \mathbb{C}^\times$.

   - Calcule o inverso de $g_{\lambda,a}$, e a composição $g_{\lambda,a}\, g_{\mu,b}$.
   - Verifique que o grupo afim não é comutativo.

7. (grupo linear geral) O grupo $\mathrm{Aut}(\mathbb{C}^n)$ dos *automorfismos* do espaço linear $\mathbb{C}^n$ é o conjunto das aplicações lineares invertíveis $L : \mathbb{C}^n \to \mathbb{C}^n$, munido da lei de composição $L, M \mapsto L \circ M$. O elemento neutro é a aplicação identidade. Fixada uma base de $\mathbb{C}^n$ (por exemplo, a base canónica), uma transformação linear invertível é definida por $Z \mapsto AZ$, onde A é uma matriz $n \times n$ invertível e $Z \in \mathbb{C}^n$ é um vetor coluna. A composição das transformações $Z \mapsto BZ$ e depois $Z \mapsto AZ$ é a transformação $Z \mapsto ABZ$, e corresponde portanto ao produto "linhas por colunas" entre as matrizes A e B.

   $\mathrm{Aut}(\mathbb{C}^n)$ pode portanto ser identificado (tecnicamente é "isomorfo", no sentido explicado a seguir) ao conjunto $\mathbf{GL}(n, \mathbb{C})$ das matrizes $n \times n$ invertíveis A, munido do produto "linhas por colunas"

$$A, B \mapsto AB\,,$$

chamado *grupo linear geral complexo* em dimensão n. A condição que decide se uma matriz quadrada A é invertível é $\det A \neq 0$, portanto

$$\mathbf{GL}(n,\mathbb{C}) := \{A \in \mathrm{Mat}_{n\times n}(\mathbb{C}) \ \text{ t.q. } \ \det A \neq 0\}\,.$$

Todo grupo formado por matrizes $n \times n$ é um "subgrupo" de $\mathbf{GL}(n,\mathbb{C})$, no sentido, também explicado a seguir, que é um subconjunto que é ele próprio um grupo. Um exemplo é o *grupo linear geral real* $\mathbf{GL}(n,\mathbb{R})$ das matrizes invertíveis reais $n \times n$.

8. (homomorfismos e isomorfismos) Um *homomorfismo* do grupo G no grupo H é uma transformação $\Phi : G \to H$ que "envia produtos em produtos", ou seja, tal que

$$\Phi(g)\,\Phi(g') = \Phi(gg')$$

para todos os $g, g' \in G$. É evidente que a imagem da identidade é. a identidade, i.e. $\Phi(e) = e$. Um homomorfismo invertível é chamado *isomorfismo*. A existência de um isomorfismo entre dois grupos, G e H, é uma relação de equivalência, denotada por $G \approx H$. Grupos isomorfos são indistinguíveis do ponto de vista da estrutura de grupo.

Um elemento a de um grupo G define duas transformações $L_a : G \to G$ e $R_a : G \to G$ (multiplicação à esquerda e à direita), definidas por

$$L_a(g) := ag \qquad \text{e} \qquad R_a(g) = ga\,,$$

respetivamente. Em particular, sendo L_a uma permutação de G, todo grupo G é um subgrupo de um grupo de permutações, por exemplo de $\mathrm{Per}(G)$. As multiplicações à esquerda e à direita comutam, e a composição $R_{a^{-1}}L_a$, que envia

$$g \mapsto aga^{-1}$$

é um homomorfismo, chamado *conjugação* (observe que é a mesma fórmula que define matrizes semelhantes, ou seja, a mudança de coordenadas para uma matriz que representa um operador linear!).

- O exponencial define um isomorfismo $\exp : \mathbb{R} \to \mathbb{R}^\times_+$ entre o grupo aditivo $\mathbb{R}$ e o grupo multiplicativo $\mathbb{R}^\times_+$ dos reais positivos, e um homomorfismo $\exp : \mathbb{C} \to \mathbb{C}^\times$ do grupo aditivo $\mathbb{C}$ no grupo multiplicativo $\mathbb{C}^\times$.

9. (subgrupos e quocientes) Um *subgrupo* do grupo G é um subconjunto $H \subset G$ que forma um grupo com respeito à lei de composição de G. Para que o subconjunto H seja um subgrupo é suficiente que $ab \in H$ e $a^{-1} \in H$ para todos $a, b \in H$. Subgrupos triviais são o próprio G e o subgrupo minimal $\{e\}$.

O *núcleo* do homomorfismo $\Phi : G \to H$ é $\Phi^{-1}\{e\}$, o conjunto dos $g \in G$ tais que $\Psi(g)$ é a identidade em H. É imediato verificar que o núcleo de Ψ é um subgrupo de G.

Seja $H \subset G$ um subgrupo. A *coclasse à esquerda* e *à direita* (em inglês, *left* or *right coset*) de um elemento $g \in G$ são os subconjuntos

$$gH := \{gh\,,\ \text{with } h \in H\} \qquad \text{e} \qquad Hg := \{hg\,,\ \text{with } h \in H\}\,,$$

respetivamente. Pertencer a mesma coclasse (esquerda ou direita) é uma relação de equivalência, ou seja, particiona o grupo em classes de equivalência. De consequência, é possível definir os espaços *quociente* G/H ou $H\backslash G$ das classes de equivalência esquerdas ou direitas, respetivamente. Em geral, os espaços quociente não possuem uma estrutura natural de grupo. Isto acontece, ou seja, é possível definir um produto

$$(aH)\cdot(bH) := (ab)H$$

em G/H e verificar que este produto faz de G/H um grupo, quando H é um subgrupo *normal*, ou seja, quando $gH = Hg$ para todo $g \in G$. O mesmo acontece para o espaço quociente $H\backslash G$. É claro que todo subgrupo de um grupo abeliano é normal.

O *centro* do grupo G é o conjunto $Z(G)$ dos elementos $h \in G$ que comutam, i.e. satizfazem $gh = hg$, com todos os elementos $g \in G$. É imediato verificar que $Z(G)$ é um subgrupo normal de G. É claro também que um grupo é abeliano sse é igual ao próprio centro.

- Mostre que o subconjunto $H \subset G$ é um subgrupo de G se $ab^{-1} \in H$ para todos $a, b \in H$.
- Seja H um subgrupo finito do grupo G. Mostre que cada classe gH de G/H contém exatamente $|H|$ elementos (ou seja, existe uma bijeção de H sobre cada classe gH). Deduza que se G é também finito, então $|H|$ divide $|G|$ (*teorema de Lagrange*). A cardinalidade de G/H, ou seja, o número de coclasses gH e G, é chamada *índice* de H em G, e denotada por $[G : H]$. O teorema de Lagrange assume então a forma "tautológica"
$$|G| = [G : H] \cdot |H| \,.$$
Se H é um subgrupo normal do grupo finito G, então $[G : H] = |G/H|$.
- Mostre que um subgrupo $H \subset G$ é normal sse $h \in H$ implica $ghg^{-1} \in H$ para todo $g \in G$
- $\mathbb{Z}$ é um subgrupo do grupo aditivo $\mathbb{R}$. Em geral, $\mathbb{Z}^n$ é um subgrupo do grupo aditivo $\mathbb{R}^n$, formado pelos vetores de coordenadas inteiras. O quociente $\mathbb{T}^n := \mathbb{R}^n/\mathbb{Z}^n$ é chamado *toro* de dimensão n. Cada ponto do toro admite um (único) representante no "domínio fundamental" $[0, 1)^n$. Por exemplo, o toro $\mathbb{T}^2$ é obtido do quadrado $[0, 1] \times [0, 1]$ ao identificar os lados opostos da forma natural. É portanto a superfície de um "doughnut".
- A circunferência unitária $\mathbb{S} := \{z \in \mathbb{C} \text{ s.t. } |z| = 1\}$ é um subgrupo do grupo multiplicativo $\mathbb{C}^\times$. A aplicação "exponencial" $\exp : x \mapsto e^{2\pi i x}$ é um homomorfismo do grupo aditivo $\mathbb{R}$ sobre o grupo $\mathbb{S}$. O núcleo é o subgrupo $\mathbb{Z} \subset \mathbb{R}$. Portanto, $\mathbb{R}/\mathbb{Z} \approx \mathbb{S}$.
- Seja $n \in \mathbb{N}$. Os múltiplos inteiros de n formam um subgrupo $n\mathbb{Z}$ do grupo aditivo $\mathbb{Z}$, e o quociente $Z_n := \mathbb{Z}/n\mathbb{Z}$ é um grupo finito de ordem n. Os elementos são as classes $[k] = k + n\mathbb{Z}$ com $k = 0, 1, \ldots, n-1$, e o elemento neutro é a classe $[0]$. A lei "adição" é $(a + n\mathbb{Z}) + (b + n\mathbb{Z}) = (a + b) + n\mathbb{Z}$.
- Seja $G_n := \{z \in \mathbb{C} \text{ t.q. } z^n = 1\} \subset \mathbb{S} \subset \mathbb{C}$ o grupo das raizes n-ésimas da unidade. A aplicação "exponencial" $\exp : k + n\mathbb{Z} \mapsto e^{2\pi i k/n}$ realiza um isomorfismo de $Z_n \approx G_n$.
- A multiplicação em $\mathbb{Z}$ também passa ao quociente, e portanto define uma lei de composição interna em $\mathbb{Z}/n\mathbb{Z}$. Se p úm número primo, então o conjunto $(\mathbb{Z}/p\mathbb{Z})^\times$ dos elementos não nulos (i.e. diferentes da classe $0 + p\mathbb{Z}$) de $\mathbb{Z}/p\mathbb{Z}$ é um grupo abeliano multiplicativo se munido da lei "multiplicação" $(a + p\mathbb{Z}) \cdot (b + p\mathbb{Z}) = (a \cdot b) + p\mathbb{Z}$.
- Mostre que, se $n \geq 3$, o grupo cíclico Z_n é um subgrupo do grupo diedral (gerado pela rotação de um ângulo $2\pi/n$), que é um subgrupo do grupo simétrico, i.e.
$$Z_n \subset D_n \subset S_n \,.$$

10. (grupo linear especial) O determinante de uma matriz quadrada, a aplicação $A \mapsto \det A$, é um homomorfismo
$$\det : \mathbf{GL}(n, \mathbb{C}) \to \mathbb{C}^\times$$
do grupo linear geral complexo sobre o grupo multiplicativo dos números complexos diferentes de zero (pois o determinante de um produto é igual ao produto dos determinantes). O seu núcleo, o conjunto
$$\mathbf{SL}(n, \mathbb{C}) := \{A \in \mathbf{GL}(n, \mathbb{C}) \text{ t.q. } \det A = 1\}$$
das matrizes com determinante igual a um, é portanto um subgrupo do grupo linear, chamado *grupo linear especial*.

Da mesma forma, o grupo *linear especial real* $\mathbf{SL}(n, \mathbb{R})$ é o subgrupo do grupo linear real formado pelas matrizes com $\det A = 1$. O determinante de uma matriz quadrada real A é igual ao quociente
$$\det A = \pm \frac{\mathrm{Vol}(A(\Box))}{\mathrm{Vol}(\Box)}$$
entre o volumes do hipercubo $\Box := [0, 1]^n$ e o volume da sua imagem $A(\Box)$ pela transformação linear definida por A, com sinal positivo ou negativo dependendo se A preserva ou inverte a orientação. De consequência, $\mathbf{SL}(n, \mathbb{R})$ representa o grupo os automorfismos de $\mathbb{R}^n$ que preservam o volume e também a orientação.

11. (isometrias) Seja X um espaço munido de uma *métrica*, ou seja, uma função "distância" $\mathrm{d} : X \times X \to [0, \infty)$ que satisfaz os axiomas:

    M1 *(simetria)* $\mathrm{d}(x, y) = \mathrm{d}(y, x)$;

    M2 *(positividade)* $\mathrm{d}(x, x) = 0$, e $\mathrm{d}(x, y) > 0$ se $x \neq y$;

    M3 *(desigualdade do triângulo)* $\mathrm{d}(x, y) \leq \mathrm{d}(x, z) + \mathrm{d}(z, y)$.

    Uma permutação $f : X \to X$ é dita *isometria* se preserva as distâncias, i.e. se

    $$\mathrm{d}(f(x), f(y)) = \mathrm{d}(x, y)$$

    para todos os $x, y \in X$. É imediato verificar que a composição de duas isometrias é ainda uma isometria. De conseqência, o espaço $\mathrm{Isom}(X)$ das isometrias de X é um subgrupo do grupo das permutações de X.

12. (translações) O espaço vetorial $\mathbb{R}^n$ (ou $\mathbb{C}^n$) é um grupo abeliano relativamente à operação "soma", $\mathbf{a}+\mathbf{b}$. O elemento neutro é a origem $\mathbf{0}$. Os seua elementos podem ser pensados como "translações" do próprio espaço $\mathbb{R}^n$. De fato, cada vetor $\mathbf{a} \in \mathbb{R}^n$ define uma transformação invertível $T_{\mathbf{a}} : \mathbb{R}^n \to \mathbb{R}^n$, definida por

    $$T_{\mathbf{a}}(\mathbf{x}) := \mathbf{x} + \mathbf{a}\,,$$

    com inversa $(T_{\mathbf{a}})^{-1} = T_{-\mathbf{a}}$. A composição é $T_{\mathbf{a}} \circ T_{\mathbf{b}} = T_{\mathbf{a}+\mathbf{b}}$. O elemento neutro $\mathbf{0}$ define a translação trivial $\mathbf{x} \mapsto T_{\mathbf{0}}(\mathbf{x}) = \mathbf{x}$. O grupo $\mathbf{T}(n)$ das translações de $\mathbb{R}^n$ representa as mudanças da origem do referencial. Se $\mathbb{R}^n$ é munido da distância euclidiana $\mathrm{d}(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|$, então o grupo das translações é um subgrupo do grupo das isometrias de $\mathbb{R}^n$.

13. (grupo ortogonal e rotações) O *grupo das isometrias lineares* do espaço euclidiano $\mathbb{R}^n$ é o subgrupo dos automorfismos de $\mathbb{R}^n$ que preservam o produto escalar canónico. Fixada a base canónica (ou outra base ortonormada), a transformação linear $X \mapsto AX$, onde $X \in \mathbb{R}^n$ é um vetor coluna e $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$, é uma isometria sse $A^\top A = AA^\top = I$ (ou seja, se A é invertível e a sua inversa é $A^{-1} = A^\top$). Portanto, o grupo das isometrias lineares de $\mathbb{R}^n$ é isomorfo ao *grupo ortogonal*

    $$\mathbf{O}(n) := \{A \in \mathbf{GL}(n, \mathbb{R}) \;\text{ t.q. }\; A^\top A = AA^\top = I\}\,.$$

    O *grupo especial ortogonal* é o subgrupo $\mathbf{SO}(n)$ formado pelas matrizes ortogonais com $\det A = 1$. É chamado grupo das *rotações*, ou grupo das isometrias lineares "diretas" de $\mathbb{R}^n$ (ou seja, isometrias lineares que preservam a orientação).

    - Mostre que o determinante de uma matriz $A \in \mathbf{O}(n)$ é $\det A = \pm 1$.
    - Mostre que uma matriz $A \in \mathbf{SO}(2)$ é

      $$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

      com $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, e portanto define uma rotação de um ângulo θ no sentido anti-horário.
    - Se $B \in \mathbf{O}(2)$ tem determinante $\det B = -1$, então $A = BJ \in \mathbf{SO}(2)$ se $J = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$. Deduza que

      $$B = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$$

      com $\theta \in \mathbb{R}/2\pi\mathbb{Z}$.
    - Diga se as seguintes matrizes são ortogonais.

      $$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

    - O conjunto das matrizes $A \in \mathbf{O}(n)$ com $\det A = -1$ é um subgrupo de $\mathbf{O}(n)$?

14. (rotações do espaço de dimensão 3) O grupo das isometrias lineares do espaço euclidiano $\mathbb{R}^3$ (o espaço onde acontece a física newtoniana) é o grupo dos operadoes ortogonais, representados numa base ortonormada por matrizes do grupo ortogonal $\mathbf{O}(3)$.

    Uma matriz ortogonal A, pensada como matriz complexa, é unitária. De consequência, os seus valores próprios (em quanto matriz complexa!) satisfazem $|\lambda| = 1$, ou seja, estão na circunferência unitária do plano complexo. O polinómio caraterístico de A é um polinómio de grau 3 com coeficientes reais. Portanto, pelo menos uma das raízes é real, por exemplo $\lambda = \det A = \pm 1$ (os únicos pontos reais da circunferência unitária), e as outras duas raízes são complexas conjugadas, por exemplo $\lambda_\pm = e^{\pm i\theta}$, com $\theta \in \mathbb{R}/2\pi\mathbb{Z}$.

    Se $\lambda = \det A = -1$, então $\det(-A) = 1$, e portanto $-A \in \mathbf{SO}(3)$, ou seja, A é a composição $A = JR$ de uma rotação $R \in \mathbf{SO}(3)$ e a inversão $J := -I \in \mathbf{O}(3)$. Em particular, o grupo $\mathbf{O}(3)$ é a reunião disjunta de $\mathbf{SO}(3)$ e $-\mathbf{SO}(3) := J(\mathbf{SO}(3)$. Em quanto grupo, é isomorfo ao produto $\mathbf{SO}(3) \times C_2$ do grupo das rotações vezes o grupo cíclico $C_2 = \{\pm I\}$. É suficiente portanto compreender as rotações.

    Seja $R \in \mathbf{SO}(3)$. Se $\mathbf{n}$ é um vetor próprio unitário de R associado ao valor próprio $\lambda = 1$, então o complemento ortogonal $\mathbf{n}^\perp \approx \mathbb{R}^2$ é um subespaço invariante para R, e a restrição $R|_{\mathbf{n}^\perp}$ é uma rotação, representada por uma matriz de $\mathbf{SO}(2)$. Portanto, em um referencial ortonormado $\mathbf{i}'$, $\mathbf{j}'$, $\mathbf{k}'$ tal que $\mathbf{k}' = \mathbf{n}$, a rotação R é representada pela matriz

$$\begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

    Se o ângulo θ não é um múltiplo de π, então a reta $\mathbf{n}\mathbb{R}$ é o único espaço próprio de R, e é chamada "eixo de rotação".

    A mudança de coordenadas ortogonal que envia $\mathbf{n}$ no eixo $\mathbf{k}'$ é determinada por dois ângulos, por exemplo a longitude $\varphi \in [0, 2\pi]$ e a latidude $\psi \in [-\pi/2, \pi/2]$ de $\mathbf{n}$ (que é um ponto da esfera unitária $\mathbb{S}^2 = \{\mathbf{v} \in \mathbb{R}^3 : \|\mathbf{v}\| = 1\} \subset \mathbb{R}^3$). De consequência, as matrizes do grupo $\mathbf{SO}(3)$ podem ser parametrizadas por 3 ângulos.

    - Diga se as seguintes matrizes são ortogonais.

$$\begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

15. (isometrias euclidianas) É possível mostrar que toda a isometria do espaço euclidiano $\mathbb{R}^n$ é do género $X \mapsto GX + A$ com $A \in \mathbb{R}^n$ e $G \in \mathbf{O}(n)$, ou seja, é a composição de uma "isometria linear" (uma isometria que preserva a origem) e uma translação. Observem que translações e isometrias lineares não comutam. A composição de $X \mapsto GX + A$ e depois $X \mapsto HX + B$, com $G.H \in \mathbf{O}(n)$ e $A, B \in \mathbb{R}^n$, é

$$X \mapsto HGX + (HA + B).$$

    - Verifique que o grupo $\mathbf{T}(n)$ das translações $X \mapsto X + A$ é um subgrupo normal de $\mathrm{Isom}(\mathbb{R}^n)$, e que o quociente $\mathrm{Isom}(\mathbb{R}^n)/\mathbf{T}(n)$ é isomorfo ao grupo $\mathbf{O}(n)$.
    - Verifique que as isometrias $X \mapsto AX+C$ e $X \mapsto BX+D$, com $A, B \in \mathbf{O}(n)$ e $C, D \in \mathbb{R}^n$, nõa comutam.

16. (grupo de Galilei)

17. (corpos rígidos e ângulos de Euler)

18. (precession, nutation and intrinsic rotation of the Earth)

19. (grupos unimodulares $\mathbf{SL}(2,\mathbb{R})$ e $\mathbf{SL}(2,\mathbb{C})$ e isometrias hiperbólicas)

20. (grupo modular $\mathbf{SL}(2,\mathbb{Z})$ e números primos)

21. (grupo unitário) O *grupo dos automorfismos unitários* do espaço hermítico $\mathbb{C}^n$ é o subgrupo dos automorfismos que preservam o produto hermítico. A transformação $Z \mapsto AZ$, onde $Z \in \mathbb{C}^n$ é um vetor coluna e $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$, é unitária sse $A^*A = AA^* = I$. O grupo dos automorfismo unitários de $\mathbb{C}^n$ é portanto isomorfo ao *grupo unitário*

$$\mathbf{U}(n) := \{A \in \mathbf{GL}(n,\mathbb{C}) \text{ t.q. } A^*A = A^*A = I\}\,.$$

O determinante de uma matriz unitária satisfaz $|\det A| = 1$. O *grupo especial unitário* é o subgrupo $\mathbf{SU}(n)$ formado pelas matrizes unitárias com determinante $\det A = 1$.

- O grupo $\mathbf{U}(1)$ é o grupo multiplicativo dos números complexos $z = x + iy$ de módulo $|z| = 1$, ou seja, a circunferência unitária $\mathbb{S} \subset \mathbb{C}$. Todo número complexo de módulo um pode ser representado por $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, com $\theta \in \mathbb{R}/2\pi\mathbb{Z}$. Verifique que a correspondência
$$e^{i\theta} \mapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$
é um isomorfismo $\mathbf{U}(1) \approx \mathbf{SO}(2)$.
- Identifique $\mathbf{SU}(1)$.
- Mostre que o grupo $\mathbf{SU}(2)$ é o grupo das matrizes complexas
$$A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$
com $|\alpha|^2 + |\beta|^2 = 1$.

22. (conformal maps and Möbius group)

23. (quaternions and spin)

24. (matrizes de Pauli) As *matrizes de Pauli* são as matrizes complexas

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \text{e} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,.$$

- Veifique que as matrizes de Pauli geram a álgebra $\mathrm{Mat}_{2\times 2}(\mathbb{C})$.
- Verifique que
$$\det\sigma_i = -1 \qquad \text{e} \qquad \mathrm{tr}\,\sigma_i = 0 \qquad \forall i = 1,2,3$$
e que
$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = -i\sigma_1\sigma_2\sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$
- Determine os valores próprios das matrizes de Pauli.

25. (grupo de Lorentz) O *espaço-tempo de Minkowski* (o espaço-tempo da teoria da relatividade restrita) é o produto cartesiano $\mathbf{M} = \mathbb{R} \times \mathbb{R}^3$, de coordenadas $(t,x,y,z) = (t,\mathbf{r})$ (ou seja, uma coordenada temporal t e 3 coordenadas espaciais $\mathbf{r} = (x,y,z)$), munido da *métrica de Minkowski*, a métrica pseudo-euclidiana definida por

$$\langle (t,\mathbf{r}),(t',\mathbf{r}')\rangle := c^2tt' - \mathbf{r}\cdot\mathbf{r}'$$

onde $c > 0$ denota a "velocidade da luz". A pseudo-norma do vetor $(t,\mathbf{r})$ é portanto

$$\|(t,\mathbf{r})\|^2 = c^2t^2 - (x^2+y^2+z^2)\,.$$

Um ponto $(t,\mathbf{r})$ é de *tipo tempo* se $\|(t,\mathbf{r})\| > 0$, de *tipo espaço* se $\|(t,\mathbf{r})\| < 0$, e de *tipo luz*, ou *nulo*, se $\|(t,\mathbf{r})\| = 0$.

O grupo das isometrias de $\mathbf{M}$, é chamado *grupo de Poincaré*. As isometrias lineares, as transformações lineares de $\mathbb{R} \times \mathbb{R}^3$ que preservam a pseudo-norma de Minkowski, formam o *grupo de Lorentz* $\mathbf{O}(1,3)$. É claro que as isometrias do espaço Euclidiano $\mathbb{R}^3$, de coordenadas $\mathbf{r} = (x, y, z)$, são transformações de Lorentz, ou seja, $\mathbf{O}(3) \subset \mathbf{O}(1,3)$. Mais interessantes são as transformações de Lorentz chamadas *boosts* (ou seja, "empurrões"), que descrevem as coordenadas (t', x', y', z') em um referencial em movimento retilíneo uniforme relativamente a um referencial inercial (t, x, y, z). Se o eixo x é escolhido na direção da velocidade v, então a transformação de Lorentz é

$$\begin{aligned} t' &= \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}} \\ x' &= \frac{x - vt}{\sqrt{1 - v^2/c^2}} \\ y' &= y \\ z' &= z \end{aligned}$$

onde $|v| < c$. Em particular, a transformação linear que envia (ct, x) em (ct', x') é definida pela matriz dois por dois H_φ tal que

$$\begin{pmatrix} ct' \\ x' \end{pmatrix} = \begin{pmatrix} \cosh\varphi & -\sinh\varphi \\ -\sinh\varphi & \cosh\varphi \end{pmatrix} \begin{pmatrix} ct \\ x \end{pmatrix}$$

onde o "ângulo hiperbólico" φ (chamado, em inglês, *rapidity*) é definido por $\tanh\varphi = v/c$, ou seja,

$$\cosh\varphi = \frac{1}{\sqrt{1 - v^2/c^2}} \qquad \text{e} \qquad \sinh\varphi = \frac{v/c}{\sqrt{1 - v^2/c^2}} \,.$$

Observe que $-\infty < \varphi < \infty$ se $-c < v < c$, e que $\varphi = 0$ se $v = 0$, que corresponde a $H_0 = I$. A matriz H_φ define uma "rotação hiperbólica" de um ângulo hiperbólico φ.

- Verifique a lei de composição $H_\varphi H_\psi = H_{\varphi+\psi}$, que corresponde à lei de adição das velocidades (escalares)

$$(u, v) \mapsto w = \frac{u + v}{1 + uv/c^2}$$

(e verifique que se $|u| < c$ e $|v| < c$ então também $|w| < c$).

26. (grupo de Weyl/Heisenberg) Em mecânica quântica, assim como em análise de Fourier, são importantes os operadores de *translação* $T_\mathbf{q}$, com $\mathbf{q} \in \mathbb{R}^n$, definidos por

$$(T_\mathbf{q} f)(\mathbf{x}) := f(\mathbf{x} + \mathbf{q}) \,.$$

e os operadores de *modulação* $M_\mathbf{p}$, com $\mathbf{p} \in (\mathbb{R}^n)^* \approx \mathbb{R}^n$, definidos por

$$(M_\mathbf{p} f)(\mathbf{x}) := e^{i\mathbf{p}\cdot\mathbf{x}} f(\mathbf{x}) \,.$$

O domínio dos operdores pode ser o espaço $\mathcal{C}^\infty(\mathbb{R}^n)$ (onde vivem as funções próprias, que são as ondas planas), ou subespaços de $L^2(\mathbb{R}^n)$, onde é possível definir o produto interno L^2, como o espaço das *funções de prova* $\mathcal{D}(\mathbb{R}^n)$ (as funções infinitamente diferenciáveis com suporte compacto) ou o *espaço de Schwartz* $\mathcal{S}(\mathbb{R}^n)$ (as funções infinitamente diferenciáveis tais que todas as derivadas decaem mais rapidamente que o inverso de qualquer polinómio quando $\|\mathbf{x}\| \to \infty$). Estes operadores são unitários relativamente ao produto L^2, pois

$$\int_{\mathbb{R}^n} |f(\mathbf{x} + \mathbf{q})|^2 \, dx_1 \dots dx_n = \int_{\mathbb{R}} |f(\mathbf{x})|^2 \, dx_1 \dots dx_n$$

e

$$\int_{\mathbb{R}^n} \left| e^{i\mathbf{p}\cdot\mathbf{x}} f(\mathbf{x}) \right|^2 \, dx_1 \dots dx_n = \int_{\mathbb{R}^n} |f(\mathbf{x})|^2 \, dx_1 \dots dx_n \,.$$

A composição $W(\mathbf{q},\mathbf{p}) = M_{\mathbf{p}}T_{\mathbf{q}}$ é chamada *operador de Weyl*. Os operadores translação e modulação geram o *grupo de Weyl*, ou *grupo de Heisenberg reduzido*, $\mathrm{Wey}_n \approx \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{S}$, parametrizado por $\mathbf{p} \in \mathbb{R}^n$, $\mathbf{q} \in \mathbb{R}^n$ e $z \in \mathbb{S}$, de acordo com a identificação

$$M_{\mathbf{p}}T_{\mathbf{q}}\, z \approx (\mathbf{p},\mathbf{q},z)\,.$$

O produto, que corresponde a composição dos operadores $M_{\mathbf{p}}T_{\mathbf{q}}\, z$, com $z \in \mathbb{S}$, é dado por

$$(\mathbf{p},\mathbf{q},z)\cdot(\mathbf{p}',\mathbf{q}',z') \mapsto \left(\mathbf{p}+\mathbf{p}'\,,\ \mathbf{q}+\mathbf{q}'\,,\ zz'e^{i\mathbf{p}\cdot\mathbf{q}'}\right)$$

O *grupo de Heisenberg* (não reduzido) é o conjunto $\mathrm{Heis}_n \approx \mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}$, munido do produto

$$(\mathbf{p},\mathbf{q},t)\cdot(\mathbf{p}',\mathbf{q}',t') \mapsto (\mathbf{p}+\mathbf{p}'\,,\ \mathbf{q}+\mathbf{q}'\,,\ t+t'+\mathbf{p}\cdot\mathbf{q}')$$

e é claro que a aplicação $(\mathbf{p},\mathbf{q},t) \mapsto (\mathbf{p},\mathbf{q},e^{it})$ é um homomorfismo do grupo de Heisenberg sobre o grupo de Heisenberg reduzido. É possível realizar o grupo de Heisenberg como um grupo de matrizes. De fato, a correspondência $M : \mathrm{Heis}_n \to \mathrm{Mat}_{(n+2)\times(n+2)}(\mathbb{R})$, que envia $(\mathbf{q},\mathbf{p},t) \in \mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}$ na matriz

$$M(\mathbf{p},\mathbf{q},t) = \begin{pmatrix} 1 & p_1 & \dots & p_n & t \\ 0 & 1 & \dots & 0 & q_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & q_n \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

é um homomorfismo do grupo de Heisenberg Heis_n no grupo $\mathbf{GL}(n+2,\mathbb{R})$.

- Verifique que $T_{\mathbf{q}}M_{\mathbf{p}} = e^{i\mathbf{p}\cdot\mathbf{q}}M_{\mathbf{p}}T_{\mathbf{q}}$.
- Verifique que o centro do grupo de Heisenberg é o subgrupo $Z(\mathrm{Heis}_n) \approx \mathbb{R}$ dos elementos da forma $(0,0,t)$, e que o quociente $\mathrm{Heis}_n/Z(\mathrm{Heis}_n)$ é (isomorfo a)o grupo aditivo abeliano $\mathbb{R}^n\times\mathbb{R}^n$.

27. (matrizes de Dirac) As *matrizes de Dirac* são as matrizes $\gamma^\mu \in \mathrm{Mat}_{4\times4}(\mathbb{C})$, com $\mu = 0,1,2,3$, definidas por

$$\gamma^0 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \qquad \gamma^1 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix}$$

$$\gamma^2 = \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \end{pmatrix} \quad , \quad \gamma^3 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

- Verifique que o anticomutator[35]

$$\{\gamma^\mu,\gamma^\nu\} = 2\eta^{\mu\nu}I_4$$

onde $\eta^{\mu\nu}$ é a métrica de Minkowski com signatura $(+---)$.

[35] $\{A,B\} = AB+BA$

24 Exponencial

ref: [Ap69] *Vol. 2*, **7.1-10**, **7.12**

1. (systems of linear differential equations & exponential) Given a linear operator L, in finite dimension a square matrix A, one would like to give a meaning to an expression like $f(L)$, where $f(t)$ is some complex valued function of a complex variable. For polynomials $p(t) = a_n t^n + \cdots + a_1 t + a_0$, this is straightforward, since integer powers of operators are well defined (they are compositions) and linear combinations too. Thus, one simply set
$$p(L) := a_n L^n + \cdots + a_1 L + a_0 I\,.$$
More interesting are transcendental functions as the exponentials e^t or e^{it}. To define them (even when the argument is a real number!), we need Taylor series, hence some analysis to understand convergence.
The exponential $x(t) = e^{\lambda t}$ is the unique solution of the differential equation $\dot{x} = \lambda x$ with initial condition $x(0) = 1$. Moreover, it satisfies the functional equation $x(t+s) = x(t)\,x(s)$, which says that x defines a homomorphism from the additive group $\mathbb{R}$ into the multiplicative group $\mathbb{C}^\times$. If we try to solve a vector differential equation (i.e. a system of linear homogeneous differential equations) like
$$\dot{X} = AX\,,$$
with $X \in \mathbb{C}^n$ and $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$, we are tempted to look for a solution as
$$X(t) = e^{tA} X(0)\,.$$
In the following, we will give a meaning to such an expression, and prove that it really solve our problem. The functional equation will say that e^{tA} is a one-parameter subgroup of $\mathbf{GL}(n,\mathbb{C})$. The practical computation of the exponential of a matrix will make use of diagonalization, commutativity, and related considerations. More important, some qualitative aspects of the solutions of the system will derive simply from considerations on the spectrum of A, the set of its eigenvalues.

2. (normas) Uma *norma* no espaço linear $\mathbf{V}$, real ou complexo, é uma função real não-negativa $\|\cdot\| : \mathbf{V} \to [0,\infty)$ que satisfaz os axiomas

   N1 *(positividade)* $\|\mathbf{x}\| > 0 \qquad \forall \mathbf{x} \neq 0$.

   N2 *(homogeneidade)* $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\| \qquad \forall \lambda \in \mathbb{R}$ ou $\mathbb{C}$ e $\forall \mathbf{x} \in \mathbf{V}$.

   N3 *(desigualdade do triângulo)* $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\| \qquad \forall \mathbf{x},\mathbf{y} \in \mathbf{V}$.

   Um *espaço normado* é um espaço vetorial $\mathbf{V}$ munido de uma norma. A *distância* entre os pontos/vetores de $\mathbf{V}$ é definida por $\mathrm{d}(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\|$.
   é imediato verificar que
$$\|\mathbf{x}\|_\infty := \max_{1\leq k\leq n} |x_k| \qquad \text{e} \qquad \|\mathbf{x}\|_1 := \sum_{k=1}^n |x_k|$$
   são normas no espaço $\mathbb{R}^n$ ou $\mathbb{C}^n$. De fato, estas normas são os dois extremos de uma família de normas parametrizadas por um número $1 \leq p \leq \infty$, definidas por
$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$$
   Quando $p = 2$ esta é a norma euclidiana usual $\|\mathbf{x}\|_2 = \sqrt{\langle \mathbf{x},\mathbf{x}\rangle}$, definida à custa do produto interno usual.

   - Use N3 para deduzir a desigualdade do triângulo
$$\mathrm{d}(\mathbf{x},\mathbf{y}) \leq \mathrm{d}(\mathbf{x},\mathbf{z}) + \mathrm{d}(\mathbf{z},\mathbf{y})$$
     para a distância.

- Uma norma $\|\cdot\|$ é euclidiana (ou seja, é definida à custa de um produto interno $\langle\cdot,\cdot\rangle$ usando $\|\mathbf{x}\| = \langle\mathbf{x},\mathbf{x}\rangle^{1/2}$) sse satisfaz a identidade do paralelogramo.
- Dada uma norma $\|\cdot\|$ no espaço vetorial $\mathbb{R}^n$ (ou $\mathbb{C}^n$), e um operador linear invertível $A \in \mathrm{Aut}(\mathbb{R}^n)$, verifique que $\|\mathbf{v}\|_A := \|A\mathbf{v}\|$ é uma norma.

3. (normas equivalentes) Duas normas $\|\cdot\|_\alpha$ e $\|\cdot\|_\beta$ no mesmo espaço vetorial $\mathbf{V}$ são *equivalentes* se existem constantes $c, C > 0$ tais que para todo $\mathbf{x} \in \mathbf{V}$

$$c\,\|\mathbf{x}\|_\beta \leq \|\mathbf{x}\|_\alpha \leq C\,\|\mathbf{x}\|_\beta$$

Normas eqivalentes definem a mesma noção de limite, ou seja, a mesma "topologia".

Teorema 24.1. *Todas as normas num espaço vetorial de dimensão finita são equivalentes.*

Demonstração. Fixada uma base (arbitrária) $\mathbf{e}_1, \dots, \mathbf{e}_n$ de $\mathbf{V} \approx \mathbb{R}^n$ ou $\mathbb{C}^n$, é possível definir uma norma declarando que os $\mathbf{e}_k$'s formam uma base ortonormada. Basta dizer que a norma do vetor $\mathbf{v} = \sum_k v_k\mathbf{e}_k$ é $\|\mathbf{v}\|_2 := \left(\sum_k |v_k|^2\right)^{1/2}$ (esta é uma norma euclidiana, induzida pelo produto interno tal que $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$). Seja $\|\cdot\|$ uma outra norma em $\mathbf{V}$. Se $\mathbf{v} = \sum_k v_k\mathbf{e}_k$, então

$$\|\mathbf{v}\| \leq \sum_k |v_k|\,\|\mathbf{e}_k\| \leq \left(\sum_k |v_k|^2\right)^{1/2} \left(\sum_k \|\mathbf{e}_k\|^2\right)^{1/2} = M\,\|\mathbf{v}\|_2$$

onde usamos a desigualdade de Cauchy-Schwarz, e onde a constante $M > 0$ é

$$M = \left(\sum_k \|\mathbf{e}_k\|^2\right)^{1/2}.$$

Por outro lado, a esfera unitária $S := \{\|\mathbf{v}\|^2 = 1\}$ é compacta (porque estamos em um espaço de dimensão finita!), logo a função contínua $\|\cdot\|$ atinge um mínimo em S, que é necessariamente positivo. Ou seja, existe $m > 0$ tal que $\|\mathbf{v}\| \geq m$ se $\mathbf{v} \in S$. Se $\mathbf{v} \neq 0$ é um vetor genérico com norma $\|\mathbf{v}\|_2 = \lambda > 0$, então $\mathbf{v}/\lambda \in S$, e portanto

$$\|\mathbf{v}\| = \|\lambda\,(\mathbf{v}/\lambda)\| = \lambda\,\|\mathbf{v}/\lambda\| \geq \lambda\,m = m\,\|\mathbf{v}\|_2$$

Logo, $m\,\|\mathbf{v}\|_2 \leq \|\mathbf{v}\| \leq M\,\|\mathbf{v}\|_2$. □

Este resultado é importante porque diz que a topologia de um espaço linear de dimensão finita, ou seja, os conceitos de limites e vizinhanças, não dependem da particular norma usada na definição (desde que seja usada uma norma!).

4. (normas na álgebra das matrizes quadradas) O espaço $\mathrm{Mat}_{n\times n}(\mathbb{C})$ das matrizes quadradas $n\times n$ é um espaço linear isomorfo a $\mathbb{C}^{n^2}$. Como tal, admite muitas normas, todas equivalentes. Por exemplo, a norma de uma matriz $A = (a_{ij})$ pode ser

$$\|A\|_1 := \sum_{i,j} |a_{ij}| \qquad \|A\|_2 := \sum_{i,j} |a_{ij}|^2 \qquad \dots \qquad \|A\|_\infty := \max_{i,j} |a_{ij}|$$

Outra possibilidade importante é considerar a transformação linear $Z \mapsto AZ$ do espaço euclidiano $\mathbb{C}^n$, e definir a *norma do operador* de A como

$$\|A\| := \sup_{0\neq Z\in\mathbb{C}^n} \frac{\|AZ\|}{\|Z\|}$$

onde $\|Z\|$ denota a norma euclidiana de um vetor $Z \in \mathbb{C}^n$. Ou seja, $\|A\|$ é o menor raio R tal que a imagem da bola unitária $B_1(0) = \{\|Z\| \leq 1\}$ pela transformação linear definida por A está contida na bola $B_R(0) = \{\|Z\| \leq R\}$.

O produto linhas por colunas (ou seja, a composição das transformações lineares) faz do espaço das matrizes quadradas uma álgebra. É fácil provar que algumas de estas normas são "sub-multiplicativas", ou seja,

$$\|AB\|_1 \leq \|A\|_1 \, \|B\|_1 \qquad \text{e} \qquad \|AB\| \leq \|A\| \, \|B\| \tag{24.1}$$

Em particular, considerando $A = B$ e iterando,

$$\|A^k\|_1 \leq \|A\|_1^k \qquad \text{e} \qquad \|A^k\| \leq \|A\|^k \,.$$

- Mostre que $\|A\|_\infty \leq \|A\|_1 \leq n^2 \, \|A\|_\infty$.
- Mostre que $\|AB\|_1 \leq \|A\|_\infty \, \|B\|_1$. Deduza a primeira das desigualdades (24.1).
- Prove a segunda das desigualdades (24.1).
- Mostre que $\|AB\|_\infty \leq n \, \|A\|_\infty \, \|B\|_\infty$. Deduza que

$$\|A^k\|_\infty \leq n^k \, \|A\|_\infty^k \,.$$

- Prove que a norma do operador pode ser também definida como

$$\|A\| = \sup_{Z \in \mathbb{C}^n \,,\, \|Z\|=1} \|AZ\|$$

5. (exponencial) O *exponencial* da matriz quadrada $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{C})$ é a matriz quadrada e^A, ou $\exp(A)$, definida pela série de potências

$$\boxed{\begin{aligned} e^A &:= \sum_{k=0}^{\infty} \frac{1}{k!} A^k \\ &= I + A + \frac{1}{2}A^2 + \frac{1}{6}A^3 + \dots \end{aligned}} \tag{24.2}$$

A definição é bem posta porque cada entrada de e^A é a soma de uma série absolutamente convergente. De facto, se $a_{ij}^{(k)}$ são as entradas de A^k, então

$$|a_{ij}^{(k)}| \leq \|A^k\|_\infty \leq n^k \|A\|_\infty^k \,.$$

Portanto, as séries dos valores absolutos das entradas de e^A são limitadas pela série convergente

$$\left(\delta_{ij} + |a_{ij}| + \frac{|a_{ij}^{(2)}|}{2} + \frac{|a_{ij}^{(3)}|}{6} + \dots \right) \leq \sum_{k=0}^{\infty} \frac{n^k \, \|A\|_\infty^k}{k!} = e^{n\|A\|_\infty} \,.$$

Se A e B são matrizes semelhantes, ou seja, $A = U^{-1}BU$ com $U \in \mathbf{GL}(n, \mathbb{R})$, então também os exponenciais são semelhantes, pois as potências são $A^n = U^{-1}B^nU$ para todo $n \geq 0$, e portanto

$$\begin{aligned} e^A &= I + U^{-1}BU + \tfrac{1}{2}U^{-1}B^2U + \dots \\ &= U^{-1}\left(I + B + \tfrac{1}{2}B^2 + \dots\right) U \\ &= U^{-1} \, e^B \, U \,. \end{aligned} \tag{24.3}$$

De consequência, se L é um operador linear do espaço de dimensão finita $\mathbf{V} \approx \mathbb{C}^n$ ou $\mathbb{R}^n$, representado em uma base fixada pela matriz A, então a fórmula (24.2) define um operador linear

$$e^L = I + L + \frac{1}{2}L^2 + \frac{1}{6}L^3 + \dots$$

Pela fórmula (24.3), esta definição não depende da base escolhida.

Por exemplo, se $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ é uma matriz diagonal, então o seu exponencial também é diagonal e

$$e^{\Lambda} = \mathrm{diag}\left(e^{\lambda_1}, \dots, e^{\lambda_n}\right) .$$

Em particular, se A é diagonalizável, ou seja, $A = U^{-1}\Lambda U$ com Λ diagonal, então o seu exponencial é semelhante a matriz diagonal e^{Λ}, ou seja,

$$e^{A} = U^{-1}e^{\Lambda}U .$$

Uma consequência importante é a relação entre o exponencial e os invariantes principais de uma matriz quadrada, o determinante e o traço:

$$\boxed{\det\left(e^{A}\right) = e^{\mathrm{tr}A}} \tag{24.4}$$

Esta fórmula é evidente se A é diagonalizável, e segue por continuidade no caso geral (porque o conjunto das matrizes diagonalizáveis é denso no conjuntos $\mathrm{Mat}_{n\times n}(\mathbb{C})$ das matrizes quadradas complexas).

- Mostre que

$$\left(e^{A}\right)^{\top} = e^{A^{\top}} \qquad \text{e} \qquad \left(e^{A}\right)^{*} = e^{A^{*}}$$

6. (equação diferencial & subgrupos a um parâmetro) Dada uma matriz $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$, podemos construir a família das matrizes

$$G(t) := e^{tA}, \qquad \text{com } \ t \in \mathbb{R} .$$

As séries de funções $t \mapsto (e^{tA})_{ij}$ que definem as entradas de e^{tA} convergem uniformemente em cada intervalo limitado da reta real, assim como as séries das derivadas das entradas. Em particular, as derivadas em ordem a t podem ser calculadas derivando cada termo. O resultado é que

$$\frac{d}{dt}G(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}A^{k+1} = A\,G(t) = G(t)\,A \tag{24.5}$$

Em particular, isto diz que A comuta com $G(t)$. É imediato ver que $G(0) = I$.

A derivada de $F(t) := e^{tA}e^{-tA}$ é igual, pela regra de Leibniz (aplicada a cada entrada do produto), a

$$F'(t) = AF(t) - F(t)A = 0$$

porque A comuta com $G(t)$. Pelo teorema do valor médio, $F(t) = F(0) = I$.

De consequência, $G(t) = e^{tA}$ é invertível, e a sua inversa é

$$\boxed{(e^{tA})^{-1} = e^{-tA} .}$$

Ou seja, o exponencial envia $\exp : \mathrm{Mat}_{n\times n}(\mathbb{C}) \to \mathbf{GL}(n, \mathbb{C})$.

Teorema 24.2. *Seja $A \in \mathrm{Mat}_{n\times n}(\mathbb{C})$. A única solução da equação diferencial*

$$\dot{X} = AX \qquad \textit{ou} \qquad \dot{X} = XA$$

com condição inicial $X(0) = X_0 \in \mathbf{GL}(n, \mathbb{C})$, é

$$X(t) = e^{tA}\,X_0 \qquad \textit{ou} \qquad X(t) = X_0\,e^{tA} ,$$

respetivamente.

Demonstração. É evidente que $e^{tA}X_0$ ou X_0e^{tA} são soluções. Para provar a unicidade, basta observar que, se $X(t)$ é uma solução, então a matriz $X(t)\,e^{-tA}$ (ou $e^{-tA}X(t)$ no segundo caso) não depende do tempo, pois a sua derivada é zero, e portanto é constante e igual ao seu valor inicial X_0. □

Em geral, e^{A+B} é diferente de e^Ae^B, e e^Ae^B é diferente de e^Be^A. O que é verdade é o seguinte

Teorema 24.3. *Se A e B comutam, ou seja, se $AB = BA$, então*

$$e^{A+B} = e^Ae^B = e^Be^A\,.$$

Demonstração. Se A comuta com B, então todas as potências A^k comutam com todas as potências B^j, e de consequência com os exponenciais e^{tB} e e^{tA}, respetivamente. Então a derivada de

$$H(t) = e^{t(A+B)} - e^{tA}e^{tB}$$

é, usando as fórmulas 24.5,

$$H'(t) = (A+B)\,e^{t(A+B)} - A\,e^{tA}e^{tB} - e^{tA}e^{tB}\,B = (A+B)\,H(t)$$

Pelo teorema 24.2, $H(t) = e^{t(A+B)}H(0)$. Mas $H(0) = 0$, logo $H(t) = 0$ para todo t, em particular para $t = 1$. □

Em particular (sendo que todos os múltiplos tA de uma matriz quadrada A comutam), a família dos $G(t) = e^{tA}$, com $t \in \mathbb{R}$, é um *subgrupo a um parâmetro.* do grupo $\mathbf{GL}(n, \mathbb{C})$, ou seja, satisfaz

$$\boxed{e^{0A} = I \qquad \text{e} \qquad e^{tA}e^{sA} = e^{(t+s)A}\,.}$$

Em outras palavras, a correspondência $t \mapsto e^{tA}$ é um homomorfismo do grupo aditivo $\mathbb{R}$ no grupo $\mathbf{GL}(n, \mathbb{C})$. A sua imagem é uma curva no grupo linear, passando pela identidade quando $t = 0$, que resolve a equação diferencial $\dot{G} = AG$.

A matriz A é dita *gerador* do subgrupo $\{G(t)\}_{t\in\mathbb{R}}$, e pode ser obtida calculando o limite

$$A = \lim_{t\to 0} \frac{G(t) - I}{t}\,.$$

É a derivada, ou seja, a "velocidade", da curva $G(t)$ no instante $t = 0$.

Teorema 24.4. *Se A é uma matriz real anti-simétrica, entã e^A é uma matriz ortogonal. Se A é uma matriz hemi-hermítica, então e^A é uma matriz unitária.*

A prova é um exercício.

- Verifique que as identidades $e^{A+B} = e^Ae^B = e^Be^A$ são em geral falsas quando A e B não comutam. Por exemplo, considere

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \text{e} \qquad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

- Se A é simétrica/hermítica então e^A também é simétrica/hermítica.
- Se A é hemi-simétrica/hemi-hermítica então e^A é ortogonal/unitária. Em particular, se H é uma matriz auto-adjunta, então a família das

$$e^{itH}\,,$$

com $t \in \mathbb{R}$, é um grupo a um parâmetro de matrizes unitárias, i.e. $e^{itH} \in \mathbf{U}(n)$.

- A identidade (24.4) implica que o exponencial de uma matriz A com traço nulo, $\mathrm{tr}A = 0$, é uma matriz do grupo linear especial, i.e. $e^A \in \mathbf{SL}(n, \mathbb{C})$.
- [Ap69] *vol. 2,* **7.12.**

7. (cálculo do exponencial) O exponencial de uma matriz diagonal $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ é também uma matriz diagonal $e^\Lambda = \mathrm{diag}(e^{\lambda_1}, \dots, e^{\lambda_n})$. Também sabemos calcular o exponencial de uma matriz diagonalizável $A = C^{-1}\Lambda C$, que é $e^A = C^{-1}e^\Lambda C$. De consequência, sabemos calcular o exponencial de quase todas as matrizes complexas, pois o espaço das matrizes diagonalizáveis é um aberto denso em $\mathrm{Mat}_{n\times n}(\mathbb{C})$.

   No entanto, nas aplicações é necessário calcular o exponencial de matrizes particulares.

   Se $A = \lambda I + B$, então $e^A = e^\lambda e^B$, porque λI comuta com todas as matrizes. Em particular, o exponencial de um bloco de Jordan $A = \lambda I + N$, onde $N = (n_{ij})$ é uma matriz nilpotente com $n_{i,i+1} = 1$ e as outras entradas nulas, é

$$e^A = e^\lambda e^N$$

   e o exponencial e^N pode ser facilmente calculado, sendo um polinómio em N. De fato, se N é nilpotente de ordem k, ou seja, verifica $N^k = 0$ (subentendido, com $N^j \neq 0$ se $j < k$), então o seu exponencial é um polinómio de grau $k-1$ em N,

$$e^N = I + N + \tfrac{1}{2}N^2 + \cdots + \tfrac{1}{(k-1)!}N^{k-1} .$$

   O exponencial de uma matriz anti-simétrica é uma matriz ortogonal. O caso mais simples é em dimensão 2. Se

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} ,$$

   então um cálculo mostra que

$$e^{\theta A} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

   uma rotação do plano de um ângulo θ no sentido anti-horário.

   - Calcule o exponencial das seguintes matrizes

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

   - Verifique que

$$A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} \qquad \Rightarrow \qquad e^A = \begin{pmatrix} e^\lambda & e^\lambda \\ 0 & e^\lambda \end{pmatrix}$$

   - Verifique que

$$A = \begin{pmatrix} 0 & -\theta \\ \theta & 0 \end{pmatrix} \qquad \Rightarrow \qquad e^A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

     e

$$B = \lambda I + A = \begin{pmatrix} \lambda & -\theta \\ \theta & \lambda \end{pmatrix} \qquad \Rightarrow \qquad e^B = e^\lambda \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

   - Se P é uma projeção, i.e. satisfaz $P^2 = P$ (é "idempotente"), então

$$e^P = I + (e-1)P .$$

   - Verifique que

$$A = \begin{pmatrix} \rho & 1 \\ 0 & \rho \end{pmatrix} \qquad \Rightarrow \qquad e^{tA} = e^{\rho t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$

     e que

$$A = \begin{pmatrix} \rho & 1 & 0 \\ 0 & \rho & 1 \\ 0 & 0 & \rho \end{pmatrix} \qquad \Rightarrow \qquad e^{tA} = e^{\rho t} \begin{pmatrix} 1 & t & \frac{1}{2}t^2 \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}$$

     (observe que A é a soma de uma matriz nilpotente e um múltiplo da identidade).

- Verifique que

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \Rightarrow \qquad e^{tA} = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$$

- [Ap69] *vol. 2,* **7.12.**

8. (álgebra de Lie de $\mathbf{SL}(2,\mathbb{R})$) O grupo $\mathbf{SL}(2,\mathbb{R})$ é o subgrupo das matrizes invertíveis $A \in \mathbf{GL}(2,\mathbb{R})$ com determinante $\det A = 1$. O espaço $\mathrm{Mat}_{2\times 2}(\mathbb{R})$ das matrizes reais dois por dois $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ é um espaço linear real isomorfo a $\mathbb{R}^4$, com coordenadas a, b, c e d, e o grupo $\mathbf{SL}(2,\mathbb{R})$ é portanto a quádrica definida pela equação cartesiana $ad - bc = 1$. As matrizes

$$X = \begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}, \qquad H^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \text{e} \qquad H^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

formam uma base do espaço linear $\mathfrak{sl}(2,\mathbb{R}) \approx \mathbb{R}^3$ das matrizes reais dois por dois com traço nulo, o hiperplano definido pela equação cartesiana $a + d = 0$ em $\mathrm{Mat}_{2\times 2}(\mathbb{R})$. Pela (24.4), os exponenciais

$$g(t) := e^{tX} = \begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix}$$

e

$$h^+(x) := e^{xH^+} = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix} \qquad \text{e} \qquad h^-(y) := e^{yH^-} = \begin{pmatrix} 1 & 0 \\ y & 1 \end{pmatrix},$$

com $t, x, y \in \mathbb{R}$, definem subgrupos a um parâmetro de $\mathbf{SL}(2,\mathbb{R})$. O exponencial

$$k(\theta) := e^{\theta(H^- - H^+)} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

com $\theta \in \mathbb{R}/2\pi\mathbb{Z}$, define o subgrupo a um parâmetro $\mathbf{SO}(2,\mathbb{R}) \subset \mathbf{SL}(2,\mathbb{R})$, o grupo das rotações.

- Verifique as relações de comutação

$$[G, H^\pm] = \pm H^\pm \qquad [H^+, H^-] = 2G$$

e

$$g(t)\, h^\pm(x)\, g(-t) = h^\pm(e^{\pm t}x)$$

- Calcule $e^{t(H^+ + H^-)}$.

9. (rotações infinitesimais) Uma rotação de um ângulo θ (no sentido anti-horário) é representada, na base canónica do plano, pela matriz $R_\theta \in \mathbf{SO}(2,\mathbb{R})$ definida por

$$R_\theta = e^{\theta A} \qquad \text{onde o gerador é} \qquad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Uma rotação $R \in \mathbf{SO}(3,\mathbb{R})$ de um ângulo θ en torno do eixo definido pelo vetor unitário $\mathbf{n} = (x, y, z)$ é $R = e^{\theta A}$ onde o gerador é

$$A = \begin{pmatrix} 0 & -z & y \\ z & 0 & -x \\ -y & x & 0 \end{pmatrix}$$

. . . continua . . .

10. (geradores infinitesimais de grupos unitários) Seja A um operador auto-adjunto de um espaço de Hilbert $\mathcal{H}$. Os operadores
$$U_t := e^{itA}$$
com t real (por exemplo, um "tempo") formam uma família a um parâmetro de operadores unitários, ou seja, U_0 é a identidade, e $U_t U_s = U_{t+s}$. Isto é evidente em dimensão finita, mas não trivial em dimensão infinita (o caso interessante), onde a própria definição de exponencial precisa de uma versão do teorema espetral para operadores auto-adjuntos não necessariamente limitados.

    O *teorema de Stone*[36] afirma que toda família a um parâmetro de operadores unitários (U_t), desde que "fortemente contínua", é gerada por um operador auto-adjunto, possivelmente não limitado. Fortemente contínua significa que $\lim_{t\to s} \|U_t v - U_s v\| = 0$ para todo $v \in \mathcal{H}$. O *gerador infinitesimal* é definido pela fórmula natural
$$Av := \lim_{t\to 0} \frac{U_t v - v}{it}$$
no domímio $D_A \subset \mathcal{H}$ onde o limite existe (e o que não é trivial é provar que este subespaço é suficientemente grande, ou seja, denso).

    O teorema de Stone é particularmente importante em mecânica quântica, onde, por exemplo, o gerador da evolução temporal é o operador Hamiltoniano H. Na notação de Dirac,
$$|\psi(t)\rangle = e^{itH/\hbar} |\psi(0)\rangle \ .$$

11. (geradores infinitesimais do grupo de Weyl/Heisenberg) Outro exemplo clássico, importante em mecânica quântica ou mais em geral em análise de Fourier (vejam, por exemplo, [Fo89]), é o grupo de Weyl/Heisenberg, gerado pelas translações
$$(T_q f)(x) := f(x+q)$$
e pelas modulações
$$(M_p f)(x) := e^{ipx} f(x)\,,$$
com $q \in \mathbb{R}$ e $p \in \mathbb{R}^* \approx \mathbb{R}$. Translações e modulações são grupos a um parâmetro de operadores unitários de $L^2(\mathbb{R})$, pois são invertíveis e é claro que
$$\int_{\mathbb{R}} |f(x+q)|^2\, dx = \int_{\mathbb{R}} |f(x)|^2\, dx \qquad \text{e} \qquad \int_{\mathbb{R}} \left|e^{ipx} f(x)\right|^2\, dx = \int_{\mathbb{R}} |f(x)|^2\, dx\,.$$
É possível provar que o gerador infinitesimal das translações é o operador momento $P = -i\partial$, definido, por exemplo, no espaço de Schwartz $\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R})$. Ou seja,
$$\boxed{T_q = e^{iqP} = e^{q\partial}}$$
Este fato é razoável, interpretando a fórmula de Taylor como
$$\begin{aligned} f(x+q) &= f(x) + q\, f'(x) + \frac{q^2}{2}\, f''(x) + \frac{q^3}{6}\, f'''(x) + \dots \\ &= \left(\left(1 + q\,\partial + \frac{q^2}{2}\,\partial^2 + \frac{q^3}{6}\,\partial^3 + \dots \right) f \right)(x) \end{aligned}$$
(naturalmente, esta não é uma demonstração!). O gerador infinitesimal das modulações é o operador *posição* Q, definido por $(Qf)(x) := x\,f(x)$, também no espaço de Schwartz $\mathcal{S}(\mathbb{R}) \subset L^2(\mathbb{R})$. Ou seja, como parece natural,
$$\boxed{M_p = e^{ipQ}}$$

[36] M.H. Stone, On one-parameter unitary groups in Hilbert Space, *Annals of Mathematics* **33** (1932), 643-648.

Assim como T_q e M_p , os operadores momento e posição não comutam. De fato, satisfazem a relação de comutação

$$[P, Q] = -iI$$

e geram a "álgebra de Lie" do grupo de Heisenberg.

- Determine a matriz $D \in \mathrm{Mat}_{(n+1)\times(n+1)}(\mathbb{R})$ que representa o operador ∂ na base $e_k(x) := x^k/k!$ de $\mathrm{Pol}_{\leq n}(\mathbb{R})$, com $k = 0, 1, \ldots, n$. Calcule e^{qD}, e verifique que

$$e^{q\partial} = T_q \, .$$

25 Sistemas lineares

ref: [Ap69] *Vol. 2*, **7.9** , **7.16-17**

1. (sistemas lineares homogéneos com coeficientes constantes) Um *sistema linear homogéneo* com coeficientes constantes (real) é uma EDO autónoma

$$\dot{\mathbf{x}} = \mathbf{v}(\mathbf{x})$$

para uma variável $\mathbf{x}(t) \in \mathbb{R}^n$, definida por um campo de vetores linear $\mathbf{v} \in \mathrm{End}(\mathbb{R}^n)$. A origem é uma solução de equilíbrio do sistema (pois $\mathbf{v}(0) = 0$). Fixada uma base de $\mathbb{R}^n$, o sistema pode ser escrito em notação matricial

$$\dot{X} = AX\,,$$

onde $X(t) \in \mathbb{R}^n$ é um vetor coluna e $A = (a_{ij}) \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ é a matriz que representa o campo linear $\mathbf{v}$ na base escolhida. Pelo teorema 24.2 (ou melhor, pela demonstração deste teorema), a solução do sistema com condição inicial $X(0) = X_0 \in \mathbb{R}^n$ é

$$X(t) = e^{tA}X_0\,.$$

Se A é diagonalizável, e possui n valores próprios reais $\lambda_1, \lambda_2, \dots, \lambda_n$ com vetores próprios $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ respetivamente, então a solução é uma sobreposição

$$\mathbf{x}(t) = \sum_{k=1}^{n} e^{t\lambda_k} a_k \mathbf{v}_k$$

(onde $\mathbf{a} = (a_1, \dots, a_n)$ é o vetor das condições iniciais $\mathbf{x}(0)$ na base dos vetores próprios).

O comportamento qualitativo assimptótico das soluções é determinado pelos sinais dos valores próprios. Por exemplo, se todos os valores próprios são negativos, i.e. $\lambda_k < 0$ para todo $k = 1, \dots, n$, então todas as soluções decaem, i.e. $\mathbf{x}(t) \to 0$ quando $t \to \infty$ (ou seja, a origem é um equilíbrio assimptoticamente estável). Se todos os valores próprios são positivos, i.e. $\lambda_k > 0$ para todo $k = 1, \dots, n$, então todas as soluções com condição inicial $\mathbf{x}(0) \neq 0$ se afastam da origem, e de fato $\|\mathbf{x}(t)\| \to \infty$ quando $t \to \infty$ (ou seja, a origem é um equilíbrio assimptoticamente instável).

2. (sistemas lineares no plano) O caso mais simples e não trivial é um sistema linear no plano $\dot{X} = AX$, definido por uma matriz quadrada real 2×2

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

O produto dos valores próprios é $q = \det(A) = ad - bc$, e a soma dos valores próprios é $p = \mathrm{tr}(A) = a + d$. Os valores próprios são então

$$\lambda_\pm = (p \pm \sqrt{\Delta})/2\,.$$

onde $\Delta = p^2 - 4q$.

Se a matriz é diagonalizável (sobre os reais, ou seja, se $\Delta > 0$), com valores próprios reais $\rho_\pm$, então o sistema é equivalente a

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} \rho_+ & 0 \\ 0 & \rho_- \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

A solução é

$$\begin{pmatrix} e^{\rho_+ t} & 0 \\ 0 & e^{\rho_- t} \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

A origem é dita *nodo estável* se $\rho_\pm < 0$, *nodo instável* se $\rho_\pm > 0$, *ponto de sela* se $\rho_- < 0 < \rho_+$.

Se a matriz admite apenas um vetor próprio real (ou seja, se $\Delta = 0$), com valor próprio $\rho \in \mathbb{R}$, então o sistema é equivalente a

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} \rho & 1 \\ 0 & \rho \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

A solução é

$$e^{\rho t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

A origem é dita *nodo degenerado*, estável ou instável, dependendo do sinal de ρ.

Se a matriz não tem valores próprios reais, então admite, em quanto matriz complexa, dois valores próprios conjugados e distintos, i.e. λ e $\overline{\lambda}$. Se os valores próprios são imaginários puros $\lambda_\pm = \pm i\omega$ (com $\omega > 0$, ou seja, se $p = 0$ e $\Delta < 0$), então o sistema é equivalente a

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

ou seja, ao oscilador harmónico $\ddot{x} = -\omega^2 x$ de frequência ω. A solução é

$$\begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

As órbitas são elipses, e a origem é chamada *foco*. Trajetórias que começam próximas da origem ficam próximas para todos os tempos, mesmo não sendo assimptóticas.

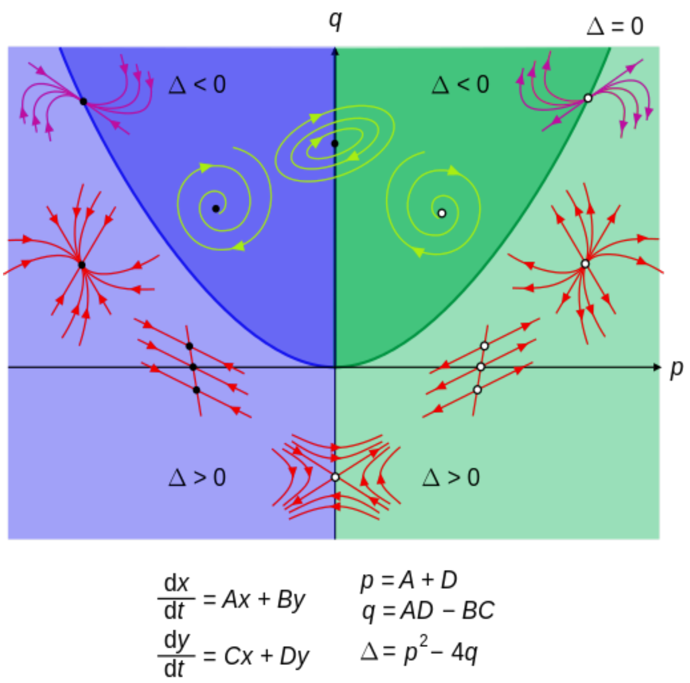
O caso genérico é uma matriz com valores próprios $\lambda_\pm = \rho \pm i\omega$, com parte real $\rho \neq 0$ (quando $p \neq 0$ e $\Delta < 0$). O sistema é equivalente a

$$\begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = \begin{pmatrix} \rho & \omega \\ -\omega & \rho \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

A solução é

$$e^{\rho t} \begin{pmatrix} \cos(\omega t) & \sin(\omega t) \\ -\sin(\omega t) & \cos(\omega t) \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

As órbitas são espirais logarítmicas, que saem da origem ou entram na origem, dependendo do sinal de ρ. A origem é dita *foco estável* se $\rho < 0$, *foco instável* se $\rho > 0$.



By Maschen, from Wikimedia Commons

É claro que a estabilidade ou instabilidade de nodos ou focos é preservada para pequenas perturbações dos parâmetros.

- Considere o oscilador invertido
$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= q \end{aligned}$$
Determine a natureza do equilíbrio. Determine a solução com condições iniciais $q(0) = 1$ e $p(0) = 0$.
- Considere o sistema linear
$$\begin{aligned} \dot{x} &= x - y \\ \dot{y} &= x + y \end{aligned}$$
Determine a solução com condições iniciais $x(0) = 1$ e $y(0) = 0$.
- Considere o sistema linear
$$\begin{aligned} \dot{x} &= 2x + y \\ \dot{y} &= x + y \end{aligned}$$
Determine a solução com condições iniciais $x(0) = 1$ e $y(0) = 1$.
- Discuta os casos degerados quando um dos valores próprios é nulo (ou seja, $\det(A) = 0$).
- [Ap69] *Vol. 2,* **7.12**

3. (ciclotrão) A força de Lorentz que age sobre uma partícula de carga q em um campo magnético $\mathbf{B}$ é igual ao produto vetorial $\mathbf{F} = (q/c)\,\mathbf{v} \times \mathbf{B}$, onde $\mathbf{v}$ é a velocidade da partícula e $c = 299,792,458$ m/s denota a velocidade da luz. De consequência, a equação do movimento (de Newton) da partícula de massa m é
$$m\dot{\mathbf{v}} = \frac{q}{c}\,\mathbf{v} \times \mathbf{B}$$
Se o campo é homogéneo e constante, por exemplo $\mathbf{B} = B\mathbf{k}$ com B independente da posição e do tempo, então a componente vertical v_z da velocidade $\mathbf{v} = (v_x, v_y, v_z)$ é constante (pois tem derivada nula), e as componentes v_x e v_y satisfazem o sistema de EDOs lineares
$$\begin{aligned} \dot{v_x} &= \omega\, v_y \\ \dot{v_y} &= -\omega\, v_x \end{aligned}$$
onde $\omega = qB/mc$ é chamada *frequência de ciclotrão*. As trajetórias da partícula são hélices
$$\mathbf{r}(t) = \mathbf{a} + (A\cos(\omega t + \varphi), A\sin(\omega t + \varphi), v_z t)$$
onde os parâmetros $\mathbf{a}, v_z, A, \varphi$ dependem da posição e da velocidade iniciais.

4. (complexificação)

   ... continua ...

5. (fatorização de operadores diferenciais lineares com coeficientes constantes) De acordo com o teorema de Picard, o espaço das soluções da EDO linear homogénea com coeficientes constantes de ordem n
$$x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_2\ddot{x} + a_1\dot{x} + a_0 x = 0 \tag{25.1}$$
é um subespaço linear $\mathcal{H} \subset \mathcal{C}^\infty(\mathbb{R})$ de dimensão $\dim \mathcal{H} = n$ do espaço $\mathcal{C}^\infty(\mathbb{R})$ das funções $x : \mathbb{R} \to \mathbb{C}$ infinitamente deriváveis. Pode ser caracterizado como sendo o núcleo $\mathcal{H} = \ker L$ do operador diferencial com coeficientes constantes
$$L = \partial^n + a_{n-1}\partial^{n-1} + \cdots + a_2\partial^2 + a_1\partial + a_0\,.$$
As soluções de (25.1) são combinações lineares (finitas) de "quase-polinómios". De fato, a conjetura $x(t) = e^{zt}$ é uma solução de (25.1) se z é uma raiz do *polinómio caratéristico*
$$P(z) := z^n + a_{n-1}z^{n-1} + \cdots + a_2 z^2 + a_1 z + a_0\,.$$

Pelo teorema fundamental da álgebra, o polinómio $P(z)$ é um produto

$$P(z) = (z - \lambda_1)^{n_1} (z - \lambda_2)^{n_2} \dots (z - \lambda_\ell)^{n_\ell}$$

de fatores $(z - \lambda_k)^{n_k}$, onde os $\lambda_k \in \mathbb{C}$ são as raízes (em geral, complexas) do polinómio, e os inteiros $n_k \geq 1$ as respetivas multiplicidades algébricas. As multiplicidades algébricas satisfazem $n_1 + n_2 + \dots + n_\ell = n$.

A correspondência entre o polinómio $P(z)$ e o operador $L = P(\partial)$ é um isomorfismo da álgebra dos polinómios em uma variável z sobre a álgebra dos operadores diferenciais com coeficientes constantes. Ou seja, combinações lineares de polinómios são enviadas em combinações lineares de operadores diferenciais, e ao produto pontual entre dois polinómios corresponde o produto, i.e. a composição, dos operadores (que é comutativa quando os coeficientes são constantes). De consequência, o operador L é também um produto

$$L = (\partial - \lambda_1)^{n_1} (\partial - \lambda_2)^{n_2} \dots (\partial - \lambda_\ell)^{n_\ell} .$$

Os operadores $(\partial - \lambda_k)^{n_k}$, com λ_k's diferentes, comutam. De consequência, o núcleo de L contém os núcleos de cada um destes operadores. O núcleo de $(\partial - \lambda_k)^{n_k}$ é o espaço de dimensão n_k dos quase-polinómios $p(t)e^{\lambda_k t}$, onde $p(t)$ é um polinómio de grau $\deg(p) < n_k$. Portanto, a solução geral de (25.1) é uma sobreposição

$$x(t) = \sum_k p_k(t)\, e^{\lambda_k t} ,$$

onde $\lambda_k \in \mathbb{C}$ são as raízes do polinómio caraterístico, com multiplicidades algébricas n_k, e os $p_k \in \mathbb{C}[t]$ são polinómios arbitrários de grau $\deg(p_k) < n_k$.

Se os coeficientes a_k's da equação diferencial linear (25.1) são números reais (como acontece frequentemente na física!), então o polinómio caraterístico possui raízes reais ou pares de raízes complexas conjugadas, e é possível construir um espaço de dimensão real n de soluções reais. A cada raiz real $\lambda_k \in \mathbb{R}$ com multiplicidade algébrica $n_k \geq 1$ está associado o espaço de dimensão real n_k dos quase-polinómios reais que anulam o operador $(\partial - \lambda_k)^{n_k}$, ou seja, o espaço linear de dimensão real n_k dos

$$p(t)\, e^{\lambda_k t}$$

com $p \in \mathbb{R}[t]$ e $\deg(p_k) < n_k$. A cada par de raízes complexas conjugadas $\lambda_k = \rho_k + i\omega_k$ e $\overline{\lambda_k} = \rho_k - i\omega_k \in \mathbb{C}$ (com $\omega_k > 0$), de multiplicidade algébrica n_k, ou seja, a cada fator

$$(\partial - \lambda_k)^{n_k} (\partial - \overline{\lambda_k})^{n_k} = \left((\partial - \rho_k)^2 + \omega_k^2 \right)^{n_k}$$

de L, está associado o espaço linear de dimensão real $2n_k$ dos quase-polinómios

$$p(t)\, e^{\rho t} \cos(\omega_k) + q(t)\, e^{\rho_k t} \sin(\omega_k t)$$

onde $p, q \in \mathbb{R}[t]$ são polinómios reais de grau $\deg(p) < n_k$ e $\deg(p) < n_k$.

- Verifique que o espaço das soluções de $\partial^n x = 0$ é o espaço linear $\mathcal{P}_{<n}$ dos polinómios de grau $< n$.
- Verifique que o núcleo do operador linear $L_\lambda := \partial - \lambda$ é o espaço linear de dimensão 1 gerado pela função $e^{\lambda t}$.
- Seja M_λ o operador *modulação*, definido por $(M_\lambda x)(t) := e^{\lambda t} x(t)$. Verifique que

  $$\partial - \lambda = M_\lambda\, \partial\, {M_\lambda}^{-1}$$

  (ou seja, M_λ realiza uma conjugação entre os operadores ∂ e $\partial - \lambda$), e portanto

  $$(\partial - \lambda)^n = M_\lambda\, \partial^n\, {M_\lambda}^{-1} .$$

  Deduza que o núcleo da potência $(\partial - \lambda)^n$, com $n \geq 1$, é o espaço linear (de dimensão n) dos quase-polinómios $p(t)e^{\lambda t}$ de grau $\deg(p) < n$.

- Determine a solução geral das seguintes ODEs lineares

$$\dddot{x} - \ddot{x} - 4\dot{x} + 4x = 0 \qquad \ddddot{x} = x \qquad \ddddot{x} + 2\ddot{x} + x = 0 \qquad \dddot{x} + 2\ddot{x} + \dot{x} = 0$$

- [Ap69] *Vol. 2,* **6.9.**

6. (variação das constantes) Um *sistema linear não homogéneo* (real) é uma lei

$$\dot{X} = A\,X + Q(t) \tag{25.2}$$

para o vetor $X(t) \in \mathbb{R}^n$, onde $A \in \mathrm{Mat}_{n\times n}(\mathbb{R})$ e $Q(t) \in \mathbb{R}^n$ é um vetor dado, dependente do tempo (uma força externa). As soluções da equação homogénea associada $\dot{Y} = A\,Y$ são $Y(t) = e^{tA}\,Z$, onde Z é um vetor constante. A conjetura $X(t) = e^{tA}\,Z(t)$ (obtida ao fazer "variar as constantes" da solução da homogénea) é solução de (25.2) sse

$$A\,e^{tA}\,Z + e^{tA}\,\dot{Z} = A\,e^{tA}\,Z + Q\,,$$

e portanto sse $Z(t)$ é solução do sistema simples

$$\dot{Z} = e^{-tA}\,Q(t)\,.$$

Pelo teorema fundamental do cálculo, aplicado a cada entrada da matriz,

$$Z(t) = Z(t_0) + \int_{t_0}^{t} e^{-\tau A}\,Q(\tau)\,d\tau\,.$$

Sendo $Z(t_0) = e^{-t_0 A} X(t_0)$, a solução de (25.2) com condição inicial $X(t_0) = X_0$ é

$$\boxed{X(t) = e^{(t-t_0)A}\,X_0 + \int_{t_0}^{t} e^{(t-\tau)A}\,Q(\tau)\,d\tau\,.}$$

Mais uma vez, mais do que lembrar esta fórmula é útil lembrar o método.

- Considere o oscilador harmónico forçado

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -q + f(t) \end{aligned}$$

Determine uma fórmula integral para a solução com condição inicial arbitrária $q(0) = q_0$ e $p(0) = p_0$. Compare a resposta com a fórmula (11.6).

- [Ap69] *Vol. 2,* **7.17.**

Referências

[Ap69] T.M. Apostol, *Calculus*, John Wiley & Sons, 1969 [*Cálculo*, Editora Reverté, 1999].

[Ar85] V.I. Arnold, *Equações diferenciais ordinárias*, MIR, 1985.

[Ar89] V.I. Arnold, *Metodi geometrici della teoria delle equazioni differenziali ordinarie*, Editori Riuniti - MIR, 1989.

[Ba77] F. Banino, *Geometria per fisici*, Feltrinelli, 1977.

[BDP92] W.E. Boyce and R.C. DiPrima, *Elementary Differential Equations and Boundary Value Problems*, John Wiley, 1992.

[Bo89] N. Bourbaki, *Elements of Mathematics, Algebra I*, Springer, 1989.

[BR98] T.S. Blyth and E.F. Robertson, *Basic Linear Algebra*, McGraw Hill, 1998.

[Ch00] T.L. Chow, *Mathematical Methods for Physicists: A concise introduction*, Cambridge University Press, 2000.

[CR48] R. Courant and H. Robbins, *What is mathematics?*, Oxford University Press, 1948. [*O que é Matemática?*, Editora Ciência Moderna, 2000].

[Di47] P.A.M. Dirac, *The Principles of Quantum Mechanics* (2nd edition), Clarendon Press, 1947

[Fe63] R.P. Feynman, R.B. Leighton and M. Sands, *The Feynman lectures on physics*, Addison-Wesley, Reading, 1963.

[Fo89] G.B. Folland, *Harmonic analysis in phase space*, Princeton University Press, 1989.

[Go96] R. Godement, *Cours d'algébre* (Troisième édition mise à jour), Hermann Éditeurs, 1996.

[Ha58] P.R. Halmos, *Finite dimensional vector spaces*, Van Nostrand, 1958.

[HS74] M.W. Hirsch and S. Smale, *Differential equations, dynamical systems and linear algebra*, Academic Press, 1974.

[KKR62] C. Kittel, W.D. Knight and M.A. Ruderman, *Berkeley Physics*, McGraw-Hill, 1962.

[La87] S. Lang, *Linear Algebra*, Third Edition, UTM Springer, 1987.

[La97] S. Lang, *Introduction to Linear Algebra*, Second Edition, UTM Springer, 1997.

[LL78] L.D. Landau e E.M. Lifshitz, *Mecânica*, MIR, 1978.

[MW85] J.E. Marsden and A. Weinstein, *Calculus I & II*, Springer, 1985.

[Me00] C.D. Meyer, *Matrix Analysis and Applied Linear Algebra*, SIAM, 2000.

[MB99] S. MacLane and G. Birkhoff, *Algebra (Third Edition)*, AMS Chelsea Publishing, 1999.

[Na06] P.J. Nahin, *Dr. Euler's fabulous formula: cures many mathematical ills*, Princeton University Press, 2006.

[Pe05] R. Penrose, *The Road to Reality: A Complete Guide to the Laws of the Universe*, Knopf, 2005.

[RHB06] K.F. Riley, M.P. Hobson and S.J. Bence, *Mathematical Methods for Physics and Engineering*, Cambridge University Press, 2006.

[Ro04] J.C. Robinson, *An introduction to ordinary differential equations*, Cambridge University Press, 2004.

[Se89] E. Sernesi, *Geometria 1*, Bollati Boringhieri, 1989.

[SG04] M. Stone and P. Goldbart, *Mathematics for Physics*, Cambridge University Press, 2004.

[St98] G. Strang, *Linear Algebra and its Applications*, Hartcourt Brace Jonovich Publishers, 1998.

[St09] G. Strang, *Introduction to Linear Algebra*, fourth edition, Wellesley-Cambridge Press and SIAM 2009.
http://math.mit.edu/linearalgebra/ , MIT Linear Algebra Lectures

[Tr13] W.F. Trench, *Elementary Differential Equations*, 2013. Books and Monographs. Book 8.
http://digitalcommons.trinity.edu/mono/8

[Wa91] B.L. van der Waerden, *Algebra*, Springer, 1991 [*Moderne Algebra*, 1930-1931].

[We52] H. Weyl, *Space Time Matter*, Dover, 1952 [*Raum Zeit Materie*, 1921].

[Ze16] A. Zee, *Group Theory in a Nutshell for Physicists*, Princeton University Press, 2016.