BA CIEAMB GEOLOG

# 2013/14
# Tópicos de Matemática EC
# Lecture notes

Salvatore Cosentino

Departamento de Matemática e Aplicações - Universidade do Minho

Campus de Gualtar, 4710 Braga - PORTUGAL

gab B.4023, tel 253 604086

e-mail scosentino@math.uminho.pt

url http://w3.math.uminho.pt/~scosentino

November 9, 2014

**Abstract**

This is not a book! These are notes written for personal use while preparing lectures on "Análise Matemática" for student of BIOQ in the a.y. 2011/12, and then "Tópicos de Matemática EC" for students of CIEAMB and GEOLOG in the a.y. 20012/13 and 2013/14, and finally also for BA in the present a.y. They are rather informal and certainly contain mistakes. I tried to be as synthetic as I could, without missing the observations that I consider important.

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

References contain some introductory manuals that I like, some classics, and other books where I have learnt things in the past century. My favorite manuals are [Ba79] (for its examples and its informal style) and [Ap69] (for its rigor and simplicity). Besides, good material and further references can easily be found in the web, for example in Scholarpedia , in Wikipedia or in the MIT OpenCoureWare.

It would be nice to have time and places to do simulations, using some of the software at our disposal in laboratories: this includes proprietary software like `Mathematica®8` , `Matlab` and Maple , or open software like `Maxima` and GeoGebra . Occasionally, we may also use some `c++` code and `Java` applets. Some applets are in the bestiario in my web page, and everything about the course may be found in my page

`http://w3.math.uminho.pt/~scosentino/teaching/tm_BA_CIEAMB_GEOLOG_2013-14.html`

Pictures were made with *Grapher* on my MacBook, or taken from Wikipedia, or produced with `Matlab` or `Mathematica®8` . Sections about linear algebra (matrices, linear systems, determinants . . . ) are still missing.

# Contents

# Notations

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 in training classes.

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

red paragraphs are technical definitions, axioms or theorems which you may skip if you are not interested in proofs (but, of course, most following facts depend on them!).

QED or the symbol $\square$, means QUOD ERAT DEMOSTRANDUM, and indicates the end of a proof.

# 1 Numbers

**ref:** [Ap69, Ba79, RHB06, Wae91]

**The language of philosophy.** "... Signor Sarsi, la cosa non istà così. La filosofia è scritta in questo grandissimo libro che continuamente ci sta aperto innanzi agli occhi (io dico l'universo), ma non si può intendere se prima non s'impara a intender la lingua, e conoscer i caratteri, ne' quali è scritto. Egli è scritto in lingua matematica, e i caratteri son triangoli, cerchi, ed altre figure geometriche, senza i quali mezi è impossibile a intenderne umanamente parola; senza questi è un aggirarsi vanamente per un oscuro laberinto." [1]

**Counting and measuring.** We count finite collections of similar objects (as fingers in our hand, years, molecules in a mole of gas, baryons in the Universe) using the numbers

$$1, 2, 3, 4, 5, \ldots, 33, \ldots, 6 \times 10^{23}, \ldots, 10^{80}, \ldots$$

We may "sum" 33 goats and 66 goats, to get a flock of $33 + 66 = 99$ goats. Also, we may need a surface of $23 \times 23 = 529$ square meters to build our pyramid with side of 23 meters. Conversely, we may sell 2 of our 99 goats and stay with the remaining flock of $99 - 2 = 97$ goats. Or we may store the visible mass $\sim 4 \times 10^{41}$ kg of the Milky Way into $\sim (4 \times 10^{41})/(2 \times 10^{30}) = 2 \times 10^{11}$ stars of the same size of our Sun (estimated to be $\sim 2 \times 10^{30}$ kg).

**Peano axioms for the natural numbers.** We use the notation $\mathbb{N} := \{1, 2, 3, 4, 5, \ldots\}$ for the set of *natural numbers*. In order to be able to prove something, it is convenient to define $\mathbb{N}$ by a (minimal) set of "axioms", and this is what Giuseppe Peano [2] did back in 1889:

N1 any natural $n \in \mathbb{N}$ has a "successor" $n^+ \in \mathbb{N}$ (which, a posteriori, we think as $n+1$), different from $n$, and no two different naturals have the same successor;

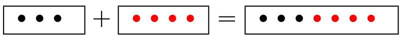
N2 there is an element, called "one" and denoted by $1 \in \mathbb{N}$, which is not the successor of any natural;

N3 (*induction principle*) a subset $A \subset \mathbb{N}$ which contains 1 and such that $n \in A$ implies $n^+ \in A$ is the whole $\mathbb{N}$.

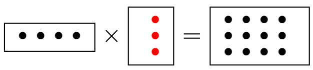
The third axiom is the key to prove that certain statements about numbers are valid for all naturals (since we humans have no time to check for all of them!). It is also the property that makes possible recursive definitions, as we'll see soon.

Once accepted the axioms, we set $2 := 1^+$, $3 := 2^+$, $4 := 3^+$, ... and so on (but of course any other list of symbols, as  ... would do).

**Sum and product.** We define *sums* inductively, starting from $n + 1 := n^+$, and setting $n + (m^+) := (n + m)^+$. The sum of two numbers represents a cardinality of an union. For example, $3 + 4 = 7$ means



We define *products* inductively, starting from $n \cdot 1 = n$, and setting $n \cdot (m^+) := n \cdot m + n$. Thus, $d \cdot a$ is the sum of $d$ times $a$, i.e. $\underbrace{a + a + \cdots + a}_{d \text{ times}}$, and actually represents an "area". For example, $4 \cdot 3 = 12$ means



If $a + b = c$, we say that $b$ is the *difference* between $c$ and $a$, and write $b = c - a$. Thus, for example, $7 = 13 - 6$.

If $q \cdot r = p$, we say that $r$ is the *ratio* between $p$ and $q$, and write $r = \frac{p}{q}$ or $p/q$. Thus, for example, $3 = 21/7$.

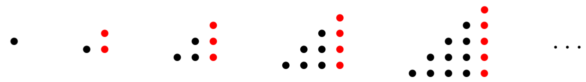
[1] Galileo Galilei, *Il Saggiatore*, 1623.

[2] G. Peano, *Arithmetices principia, nova methodo exposita*, 1889.

**e.g. Triangular numbers.** The sum of the first $n$ naturals is given by the formula

$$1 + 2 + 3 + \cdots + (n-1) + n = \frac{n(n+1)}{2},$$

which you may conjecture summing the last with the first numbers (hence $n+1$), then the last and the first of what remain (again $n - 1 + 2 = n + 1$), and so on, up to a total of $n/2$ such pairs, or, following the Greeks, observing the following picture (red bullets form the "gnomon") :

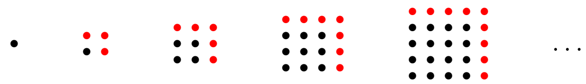


If you are not satisfied with that, you check the formula for $n = 1$ (this gives $1 = 1 \cdot 2/2$), assume it holds for $n$, sum the next term, which is $n+1$, and verifies that $n(n+1)/2+(n+1) = (n+1)(n+2)/2$.

**ex: Square numbers and ...** Show that the sum of the first $n$ odd numbers is

$$1 + 3 + 5 + 7 + \cdots + (2n-1) = n \cdot n$$

(i.e. $n^2$, but we have not introduced this notation yet!), as the following picture suggests (again, red bullets form the "gnomon"):



and guess a formula for the sum of the first $n$ even numbers

$$2 + 4 + 6 + \cdots + 2n = \,?$$

**Well-ordering principle.** We may define an order in $\mathbb{N}$ saying that $n < m$ ("$n$ is *smaller* than $m$") if there exists $x \in \mathbb{N}$ such that $n + x = m$. We say that $n \leqslant m$ ("$n$ is *not greater* than $m$") if $n < m$ or $n = m$. This relation is stable under sums and products: if $n \leqslant m$ then also

$$n + x \leqslant m + x \qquad \text{and} \qquad n \cdot x \leqslant m \cdot x$$

for all $x \in \mathbb{N}$. It is clear that 1 is the "smallest" of all the numbers, i.e. $1 \leqslant x$ for all $x \in \mathbb{N}$. The induction principle N3 is equivalent to the statement that any subset of the naturals has a smallest element:

WO *(well-ordering principle)* every subset $A \subset \mathbb{N}$ has a first element (or minimum), i.e. an element $a \in A$ such that $a \leqslant x$ for all $x \in A$.

**Integers.** It turns out (but this took quite a large time to mankind!) that even elementary problems are solved with much easy if we enlarge our numbers allowing *negative* numbers, like $-237$, hence a *zero* number, that we denote 0. The set thus obtained is the set of *integer numbers*

$$\mathbb{Z} := \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}.$$

(from the german *zahlen* = numbers) The two operations, $+$ and $\times$ (but we'd rather use "dots" for multiplication, like in $7 \cdot 3 = 21$, or even nothing when there is no possible confusion, like in $ab = a \cdot b$) are then characterized (i.e. defined!) by the following properties:

R1 (*associativity* of both $+$ and $\times$) $\quad (x + y) + z = x + (y + z) \qquad (x \cdot y) \cdot z = x \cdot (y \cdot z)$

R2 (*comutativity* of both $+$ and $\times$) $\quad x + y = y + x \qquad x \cdot y = y \cdot x$

R3 (exisistence of *neutral elements* 0 and 1) $\quad x + 0 = 0 \qquad x \cdot 1 = x$

R4 (existence of the *opposite* for $+$) $\quad \forall x$ there exists $-x$ such that $x + (-x) = 0$

R5 (*distributive law*) $\quad x \cdot (y + z) = x \cdot y + x \cdot z$

Mathematicians call a set with two such operations defined a *commutative ring*.

It is plain that all primary school arithmetical rules may be derived from these properties/axioms (but you should try to prove them by yourself!). In particular, you may derive the useful rule

$$a + x = a + y \quad \Rightarrow \quad x = y\,.$$

**Repeated sums and products.** It is useful to have a short notation for repeated sums

$$\sum_{n=1}^{N} x_n := x_1 + x_2 + \cdots + x_N$$

and products

$$\prod_{n=1}^{N} x_n := x_1 \cdot x_2 \cdot \cdots \cdot x_N$$

This is possible thanks to the associativity of both sums and products. The product of the first $n$ naturals is ubiquitous when counting cardinalities, and deserves a name: it is called $n$ *factorial*, and denoted by

$$n! := 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (n-1) \cdot n\,.$$

**Powers.** It is convenient to have a short notation for repeated products of a fixed number. For example, the product $x \cdot x$ is said "$x$ squared", and denoted by $x^2$ (for, if $x > 0$, it is the area of a square with side $x$). Similarly, $x \cdot x \cdot x$ is said "cube of $x$", and denoted by $x^3$ (if $x > 0$, it is the volume of a cube with side $x$). For integer $n = 1, 2, 3, \dots$, the $n$-th power of the (rational) number $x$ is defined by

$$x^n := \underbrace{x \cdot x \cdot \ \dots \ \cdot x}_{n \text{ times}}$$

(to be pedant, recursively according to $x^1 := x$ and $x^{n+1} := x^n \cdot x$ for $n \geqslant 1$). It is useful to set $x^0 := 1$.

**Clock arithmetics.** Less obvious is that there exist other commutative rings. For any integer $n \geqslant 2$, we may equip the quotient $\mathbb{Z}/n\mathbb{Z} := \{k + n\mathbb{Z}\,, \text{ with } k \in \mathbb{Z}\} \approx \{0, 1, \dots, n-1\}$ with the obvious ring structure inherited from $\mathbb{Z}$. Thus,

$$(a + n\mathbb{Z}) + (b + n\mathbb{Z}) = a + b + n\mathbb{Z} \qquad \text{and} \qquad (a + n\mathbb{Z}) \cdot (b + n\mathbb{Z}) = a \cdot b + n\mathbb{Z}\,.$$

**Combinatorial calculus.** Let $K \approx \{1, 2, \dots, k\}$ and $N \approx \{1, 2, \dots, n\}$ be finite sets with cardinalities $k$ and $n$, respectively. The cardinality of their Cartesian product is $|K \times N| = k \cdot n$. The cardinality of the space $N^K := \{\text{functions } K \to N\}$, isomorphic to the Cartesian product $N^k := \underbrace{N \times N \times \cdots \times N}_{k \text{ vezes}}$ is

$$\left|N^K\right| = n^k$$

The cardinality of the space $D_k^n := \{\text{injective functions } K \to N\}$ is

$$|D_k^n| = n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

provided $k \leqslant n$, where we define $0! = 1$. In particular, the cardinality of the space $D_n^n$ of *permutations* of $N$ is

$$|D_n^n| = n!$$

The cardinality of the space $C_k^n := \{\text{subsets } K \subset N \text{ with } |K| = k\}$, with $k \leqslant n$, is

$$|C_k^n| = \binom{n}{k} := \frac{n!}{k!(n-k)!}$$

since $C_k^n \approx D_k^n$ modulo $D_k^k$ (two injective functions $K \to N$ define the same subset $K \subset N$, their image, off they differ by a permutation of $K$).

**Division, factorization and primes.** We say that the number $a$ *divides* (or *is a divisor of*) the number $b$, and we write $a \mid b$, if there exists a $d \in \mathbb{N}$ such that $ad = b$. If $a$ does not divides $b$, we write $a \nmid b$. Given any $p \leqslant q$, either $p|q$, so that $q = dp$ for some $d \in \mathbb{N}$, or there exist a unique $d \in \mathbb{N}$ and a unique "rest" $0 < r < q$ such that

$$q = dp + r\,.$$

We say that a natural number $p$ is *prime* if it is not divided by any other natural but 1 and itself. Thus, 2, 3, 5, 7, 11, 13, ... are primes. It is a fundamental fact of arithmetic (and should be proved, of course!) that any natural $n$ can be uniquely factorized (up to order!) into prime factors, i.e. written as $n = p_1^{n_1} p_2^{n_2} \dots p_k^{n_k}$ for some primes $p_i$ and exponents $n_i \in \mathbb{N}$. Thus, primes are the building blocks with which all naturals are constructed.

Here is Proposition 20 of Book IX of the *Elements* by Euclid:

Οἱ πρῶτοι ἀριθμοὶ πλείους εἰσὶ παντὸς τοῦ προτεθέντος πλήθους πρώτων ἀριθμοί [3]

or, in modern language,

**Theorem 1.1. (Euclid's theorem)** *The set of prime numbers is not finite.*

Indeed, following Euclid, assume that $p_1, p_2, \dots, p_n$ are all the primes. We could take their product and sum one, i.e. form the number $x = p_1 p_2 \dots p_n + 1$, and observe that $x$ is not divisible by any of the $p_k$, since the rest of the division is always 1. Since $x$ is larger than any of the $p_k$, it must have a prime divisor larger than all of them. □

**Even and odd.** The smallest prime number is 2, and it divides the set of natural numbers into two classes: the *even* numbers, $2, 4, 6, 8, \dots$ and the *odd* numbers, $1, 3, 5, 7, \dots$.

**Greatest common divisor and smallest common multiple.** If $d$ divides both $a$ and $b$, it also divides their difference $b - a$. This observation gives rise to the *Euclid algorithm* to find $(a, b)$.

**e.g. Magicicadas.** Prime numbers may be selected by Nature as survival strategies. One example (popularized by Stephen J. Gould in [Gou77]) is that of *Magicicada*. They spend 13 or 17 years, depending on the species, under the ground as nimphs, and then get out for the few weeks or months of adult life (to mate, have offspring, and die).

A 17-year cicada, or Magicicada (from Wikipedia).

**e.g. Proportions.** If the recipe of a cake for 4 persons uses 6 eggs, and if you need the same cake for 12 guests, you must use $x = \frac{6}{4} \cdot 12$ eggs. That is, you must solve the "proportion"

$$6 : 4 = x : 12\,.$$

[3] "Prime numbers are more than any assigned multitude of prime numbers" [Euclid, *Elements*, Book IX, Proposition 20].

**Rationals and the four operations.** We may form quotients $p/q$ of integer numbers (the denominator $q$ not being 0), and define their sum and product as

$$\frac{a}{b} + \frac{c}{d} := \frac{ad + bc}{bd} \qquad \text{and} \qquad \frac{a}{b} \cdot \frac{c}{d} := \frac{ca}{bd}$$

In addition to the properties of a commutative ring, axioms R1-R5, the set of fractions also satisfies the following axiom

F6 (existence of the inverse for $\times$) $\qquad \forall x \neq 0$ there exists $x^{-1}$ such that $x \cdot x^{-1} = 1$

Indeed, the inverse of a non-zero fraction $p/q$ is simply $q/p$. The set of fractions is called *rational field*, and denoted by $\mathbb{Q}$.

An important consequence of F6 is the rule

$$\lambda x = \lambda y \quad \text{and} \quad \lambda \neq 0 \quad \Rightarrow \quad x = y\,.$$

**Order.** The field of rationals is an ordered field, i.e. may be "ordered". This means that we may define a subset $\mathbb{Q}^+ := \{p/q \, \text{with} \; p, q \in \mathbb{N}\}$ of *positive* rationals satisfying the "axioms of order"

O1 $0 \notin \mathbb{Q}^+$,

O2 if $a, b \in \mathbb{Q}^+$, then also $a + b \in \mathbb{Q}^+$ and $a \cdot b \in \mathbb{Q}^+$,

O3 $\forall\, x \neq 0$, either $x \in \mathbb{Q}^+$ or $-x \in \mathbb{Q}^+$.

We then define $\mathbb{Q}^- := \mathbb{Q} \backslash (\mathbb{Q}^+ \cup \{0\})$, the set of *negative* rationals. We say that $a < b$ if there exists a $c \in \mathbb{Q}^+$ such that $a + c = b$. We say that $a > b$ if $b < a$. In particular, all $a \in \mathbb{Q}^+$, as for example 1, are $a > 0$, and all $b \in \mathbb{Q}^-$ are $b < 0$. We also say that $a \leqslant b$ is $a < b$ or $a = b$, and then that $a \geqslant b$ if $a \leqslant b$. Clearly,

$$a < b \qquad \Rightarrow \qquad a + c < b + c$$

and also

$$a < b \quad \text{and} \quad c < d \qquad \Rightarrow \qquad a + c < b + d$$

Moreover,

$$a < b \qquad \Rightarrow \qquad \begin{cases} ad < bd & \text{if } d > 0 \\ ad > bd & \text{if } d < 0 \end{cases}$$

In particular,

$$a < b \qquad \Rightarrow \qquad -b < -a$$

Also, if $ab > 0$, then $a$ and $b$ are either both positive or both negative. Finally,

$$a \neq 0 \qquad \Rightarrow \qquad a \cdot a > 0$$

i.e. squares of non-zero numbers are positive.

ex: **Bernoulli inequality.** For any $n = 1, 2, 3, \ldots$ and any $x > -1$

$$(1 + x)^n \geqslant 1 + nx$$

Prove it using indutction.

**Absolute value and distance.** The *absolute value* (or *modulus*) of a number $x$ is

$$|x|. := \max\{x, -x\} = \begin{cases} x & \text{se } x \geqslant 0 \\ -x & \text{se } x < 0 \end{cases}\,.$$

The *distance* between $x$ and $y$ is then defined as $\mathrm{dist}(x, y) := |x - y|$. Thus, the distance between $x$ and $y$ is zero iff $x \neq y$, and we have the triangle inequalities

$$|x + y| \leqslant |x| + |y| \qquad \text{that is,} \qquad \mathrm{dist}(x, y) \leqslant \mathrm{dist}(x, z) + \mathrm{dist}(y, z)\,.$$

**Negative powers.** For $x \neq 0$, we also define negative powers according to

$$x^{-n} := \frac{1}{x^n}$$

for $n = 1, 2, 3, \ldots$. Then, for all $n, m \in \mathbb{Z}$, and all $x \neq 0$, we have

$$x^n \cdot x^m = x^{n+m}\,.$$

We observe that for even $n$, the $n$-th power of any number $x \neq 0$ is $x^n > 0$.

**Useful formulas.** The square of a binomial is

$$(a \pm b)^2 = a^2 \pm 2ab + b^2$$

Also,

$$(a + b)(a - b) = a^2 - b^2$$

Less obvious is that one can give a formula for the $n$-th power of a binomial. This has been found by Newton, and is

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

where the *binomial coefficient* is defined as

$$\binom{n}{k} := \frac{n!}{k!\,(n-k)!}$$

**ex:** Show that if $n^2$ is even then also $n$ is even.

**ex:** Find and prove a formula for the sum of the cubes of the first $n$ numbers

$$1^3 + 2^3 + 3^3 + \cdots + n^3 = \,?$$

**e.g. Surface area to volume ratio and shapes.** The volume and the surface area of an organism depend on the linear dimension $L$ according to $V = vL^3$ and $S = sL^2$, where $v$ and $s$ are certain constants that depend on the shape. For example, in a (ideal!) spherical cell, $V = \frac{4}{3}\pi \cdot L^3$ and $S = 4\pi \cdot L^2$. There follows that the *surface area to volume ratio* is

$$S : V = r \cdot L^{-1}$$

(and in particular decreases as the linear dimension increases). Nature selected a huge variety of shapes, hence of values of the constant $r = s/v$.[4]

(from Life at the Edge of Reef and Wikipedia)

For example, the "sahuaro" cactus (*Carnegiea gigantea*), from the Sonora desert of Mexico, optimize their surface area to volume ratio, hence minimize transpiration, assuming a cylindrical shape.

[4] K. Schmidt-Nielson, *Scaling: Why is Animal Size so Important?* Cambridge University Press, 1984.

**e.g. Gulliver in Liliput.** "... the Emperor stipulates to allow me a quantity of meat and drink sufficient for the support of 1728 Lilliputians. Some time after, asking a friend at court how they came to fix on that determinate number, he told me that his Majesty's mathematicians, having taken the height of my body by the help of a quadrant, and finding it to exceed theirs in the proportion of twelve to one, they concluded, from the similarity of their bodies, that mine must contain at least 1728 of theirs, and consequently would require as much food as was necessary to support that number of Lilliputians.". [5]

**Decimal notation.** Since we have 10 fingers, we like powers of 10, as 100, 1000, ... 1000000, ..., to the point that they also deserve their own names (hundreds, thousands, ..., millions, ...).

We decided to represent numbers using a "decimal" positional notation. This means that we chose 10 symbols to represent the first 9 numbers and the "zero" number, as

$$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\},$$

and then write numbers as

$$367,89 := 3 \cdot 10^2 + 6 \cdot 10 + 7 + 8 \cdot \frac{1}{10} + 9 \cdot \frac{1}{10^2}$$

All finite decimal expansions represent fractions (just multiply by a convenient power of 10). Some fractions do not terminate, and give rise to infinite periodic decimal expansions. Actually, rationals are exactly those real numbers that admit periodic (possibly finite) decimal expansions.

**ex:** Compute the following decimal expansions

$$1/20 \qquad 3/4 \qquad 5/100 \qquad 1/3 \qquad 1/7 \qquad 1/9 \qquad 1/111$$

**First degree equations.** In a field, like the rationals (or, as you will see, the reals), we are able to solve a first degree equation like

$$ax + b = 0$$

(as usual, the notation above means that we are given the numbers $a$ and $b$, and we want to find possible values for the "unknown" $x$). Indeed, we simply put $b$ on the right hand side, multiplying by $-1$, and then divide by $a$ (the case $a = 0$ being trivial: it is no equation at all!). The solutions, which is obviously unique, is

$$x = -b/a\,.$$

**Percentage.** A popular way to express ratios is using percentages: $p\%$ means $p/100$. For example, the 25% of a mass of 60 kg is $60 \cdot (25/100) = 15$ kg. Other popular expressions are "increase or decrease of some percentage". For example, a 20% increase means a factor $(1+20/100)$ multiplying the the given initial quantity.

**ex:**

- An increment of 20% followed by a farther increment of 20% correspond to a single increment of 40% or not?
- Does the order of increments matter? That is, an increment of 20% followed by an increment of 30% is the same thing as an increment of 30% followed by an increment of 20%?
- [Ba79] 1.3.2.

[5] Jonathan Swift, *Travels into Several Remote Nations of the World. In Four Parts. By Lemuel Gulliver, First a Surgeon, and then a Captain of several Ships*, 1726.

**e.g.** **Composite interests.** Assume your bank pays to yo an interest of $p\%$ each year. If you deposit a capital of $x_0$ euros, you'll get a gain of $x_0 \cdot p/100$ after one year, and therefore a total capital of $x_1 = x_0 \cdot (1 + p/100)$. The second year, the interest will be calculated on the capital $x_1$, thus leading to a total capital of $x_2 = x_1 \cdot (1 + p/100) = x_0 \cdot (1 + p/100)^2$. The total capital after $n$ years is therefore

$$x_n = x_0 \cdot \left(1 + \frac{p}{100}\right)^n$$

# 2 Real line

e.g. **Pythagora's theorem.** Take a rigth triangle, set to 1 the length of the hypotenuse, and call $\alpha$ and $\beta$ the lengths of the other sides. The altitude from the vertex opposed to the hypotenuse divides the latter into two pieces of lengths $\alpha^2$ and $\beta^2$, because they are sides of right triangles similar to the first one, having hypotenuses the two sides of length $\alpha$ and $\beta$, respectively. Therefore,

$$\alpha^2 + \beta^2 = 1 .$$

For example, the diagonal $\ell$ of a square with unit side satisfies $1 + 1 = \ell^2$, i.e. it is what we call $\sqrt{2}$.

e.g. **Babylonians-Heron method to compute square roots.** Consider the problem to find the side $\ell$ of a square given its area $a > 0$, that is, the number which we modern call $\ell = \sqrt{a}$. A method, described by Heron of Alexandria [6], but most probably already known to the Babylonians [7] [8], consists in constructing recursively rectangles with fixed area $a$ and sides which are nearer and nearer. If $x_1$ and $y_1$ are the base and the height of the first rectangle (chosen arbitrarily!), and therefore $x_1 y_1 = a$, then the second rectangle has for base the arithmetic mean $x_2 = (x_1 + y_1)/2$ and consequently height $y_2 = a/x_2$, the third rectangle has for base the arithmetic mean $x_3 = (x_2 + y_2)/2$, ... and so on. The recursive equation for the basis is

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) .$$

Observe that if the area $a$ and the initial conjecture $x_1$ are rationals, then all the $x_n$ are rationals too.

The algorithm converges, and quite fast. Consider, for example, $a = 2$, so that we are looking for $\sqrt{2}$. We could, as the Babylonians, start from an initial guess $x_1 = 3/2$ for $\sqrt{2}$ (since $1^2 < 2 < 2^2$), and find

$$x_2 = \frac{17}{12} \simeq 1.41666666666 \qquad x_3 = \frac{577}{408} \simeq 1.41421568627 \qquad x_4 = \frac{665857}{470832} \simeq 1.41421356237$$

As you see, the sequence stabilizes quite fast.

As a first attempt to explain this miracle, we could start looking at the recursive equations for the bases and the heights of the rectangles:

$$x_{n+1} = \frac{x_n + y_n}{2} \qquad 1/y_{n+1} = \frac{1/x_n + 1/y_n}{2}$$

(so, the next height is the "harmonic mean" of the base and height). We see that the $x_n$'s and the $y_n$'s form decreasing and increasing sequences, respectively (disregarding the first guess, of course), namely

$$y_2 \leqslant y_3 \leqslant \cdots \leqslant y_n \leqslant \quad \dots \quad \leqslant x_n \leqslant \cdots \leqslant x_3 \leqslant x_2 ,$$

The real root is somewhere between, namely $y_n \leqslant \sqrt{a} \leqslant x_n$. Hence, we have an explicit control of the error: the difference between $x_n$ (or $y_n$) and the real value of $\sqrt{a}$ is not greater than $|x_n - y_n|$. A computation shows that the lengths of those intervals, the differences $\varepsilon_n = x_n - y_n$ satisfy the recursion

$$\varepsilon_{n+1} < \frac{1}{2} \cdot \varepsilon_n$$

So, and initial "error" $\varepsilon_1 \leqslant 1$ (an easy achievement, since we easily recognize squares of integers) reduces to at least $\varepsilon_n \leqslant 2^{-n}$ after $n$ iterations. The true error is actually much smaller. Indeed, in our example we may compute

$$\varepsilon_2 = \frac{17}{12} - 2\frac{12}{17} = \frac{1}{204} \simeq 0.005 \qquad \text{and} \qquad \varepsilon_3 = \frac{577}{408} - 2\frac{408}{577} = \frac{1}{235416} \simeq 0.000004$$

So that the first improved guess $x_2$ has already one correct decimal, and the second, $x_3$ has already four correct decimals!

[6] Heron of Alexandria, *Metrica*, Book I.
[7] Carl B. Boyer, *A history of mathematics*, John Wiley & Sons, 1968.
[8] O. Neugebauer, *The exact sciences in antiquity*, Dover, 1969.

**ex:** Of course, more important is the “relative error”, which may be estimated as $|x_n - \sqrt{a}|/\sqrt{a} \simeq \varepsilon_n/y_n$

**ex:** **Heron formula.** According to *Heron formula*, the area of a triangle with sides $a$, $b$ e $c$, hence semi-perimeter $s = (a+b+c)/2$, is given by

$$\sqrt{s(s-a)(s-b)(s-c)}$$

Estimate the area of a triangle with sides 7, 8 e 9.

**ex:** Estimate $\sqrt{13}$ with an error $< 0.01$ or $< 0.001$ (without using your machine!).

**ex:** How many iterations are necessary to get the first $n$ correct decimals of $\sqrt{2}$ using the babylonians-Heron method?

**ex:** **A0, A1, ..., A4 paper.** Find the ratio $\lambda := b/a$ of the sides of a rectangle such that cutting it along the middle of the longer side, say $b$, you get two rectangles, with sides $b/2$ and $a$, similar to the original.

**Irrationals.** What Babylonians didn’t suspect is that if you start with a rational guess for $\sqrt{2}$, you get an infinite sequence of rational approximations, but the process never stops. This is due to

**Theorem 2.1. (Pythagoras theorem)** *There is no rational number whose square is equal to* 2 *(i.e. the square root of* 2 *is not rational).*

Indeed, assume that such a rational $p/q$ exists, and assume it is reduced. Squaring we get $(p/q)^2 = 2$, that is, $p^2 = 2q^2$. Therefore, $p^2$ is divisible by 2, hence by $2^2$ (because the factorization of a square must contain even exponents). But this implies the existence of an integer $r$ such that $2^2 r = 2q^2$, hence also $q^2$ is divisible by 2, contrary to our hypothesis that the fraction was reduced. □

It is clear that the same proof work with other square roots.

**The real line.** Pythagoras theorem suggests the need to enlarge the set $\mathbb{Q}$ of rational numbers and get the “reals” $\mathbb{R}$. This is done by admitting a new axiom, in addition to the axioms of field and order. This is a rather technical point, but it amounts to saying that the reals “have no holes”, and may be thought as a continuous line of points. Thus, once fixed an origin, called 0, a “positive” direction (typically pointing to the right) and a unit of measure (like meter, or feet, ..., fixing the point called 1), any real number $x \in \mathbb{R}$ corresponds to one and only one point on our line, the one at a distance $|x|$ from 0, on the right if $x > 0$ or on the left, if $x < 0$.

**The supremum axiom.** First, we need some terminology. A *upper bound* (*limite superior*) of a set $A$ is any number $M$ such that $a \leqslant M$ for any $a \in A$. If a upper bound of $A$ belongs to $A$ (and therefore is the unique one belonging to $A$!), than it is called a *maximum* of $A$, and denoted $\max A$. Similarly, a *lower bound* (*limite inferior*) of a set $A$ is any number $m$ such that $m \leqslant a$ for any $a \in A$. A lower bound which belongs to $A$ itself is called *minimum* of $A$, and denoted $\min A$.

Clearly, a set may have no upper and/or lower bound. A set of numbers $A$ is *bounded from above* if it admits an upper bound, and *bounded from below* if it admits a lower bound. It is called *bounded* if it admits both upper and lower bounds (i.e. if there exists a number $K$ such that $|a| \leqslant K$ for any $a \in A$). It may also happens that a set is bounded above and/or below without having maximum and/or minimum.

We define the *supremum* of $A$, notation $\sup A$, as the smallest of all the upper bounds of $A$. This means that $M = \sup A$ if $a \leqslant M$ for all $a \in M$, and if no $b < M$ is an upper bound for $A$. Similarly, we define the *infimum* of $A$, notation $\inf A$, as the largest of all lower bounds. Both supremum and infimum, if they exist, are clearly unique.

This is the final axiom, to be added to the field and order axioms, which entirely defines the real line:

S1 (*the supremum axiom*) Any not-empty subset $A \subset \mathbb{R}$ of the real line which is bounded from above has a supremum.

Of course, also any not-empty subset $B \subset \mathbb{R}$ which is bounded from below has a infimum (just reverse the signs of the numbers forming the set). The real line is the unique (up to isomorphism!) ordered and complete field, i.e. is characterized by the axioms R1-R5, F6, O1-O3 and S1.

e.g. **Existence of the square root of two.** So, for example, consider the set of decreasing rationals $\cdots \leqslant x_n \leqslant \cdots \leqslant x_3 \leqslant x_2$ obtained by the Heron method as basis of rectangles of area equal to 2. Since they all satisfy $x_n^2 > 2$, they admits an infimum, say $a$, which clearly satisfies $a^2 \geqslant 2$. Similarly, the heights $y_2 \leqslant y_3 \leqslant \cdots \leqslant y_n \leqslant \dots$ satisfy $y_n^2 < 2$, and therefore their supremum $b$ satisfies $b^2 \leqslant 2$. But the difference $|x_n - y_n|$ is arbitrarily small, since it is bounded by $1/2^n$. There follows that $a = b$ and therefore $a^2 = 2$.

ex: Find examples of unbounded sets, and of bounded sets with no maximum or minimum.

**Archimedean property of real numbers.** A first consequence of the supremum axiom is that the set of natural numbers $\mathbb{N} \subset \mathbb{R}$ is unbounded from above (if it were bounded it would have a supremum $s = \sup \mathbb{N}$, but then there would exist some natural $n > s-1$, and we could find another natural $n^+ = n+1 > s$, contradicting the assumption that $s$ is a upper bound for $\mathbb{N}$). There follows that any real $x \in \mathbb{R}$ is strictly less than some natural $n$ (and therefore of all its successors). Now, take any positive real number $\varepsilon > 0$. We claim that for any $x \in \mathbb{R}$ we can find an integer $n \in \mathbb{N}$ so large that

$$n \cdot \varepsilon > x \,,$$

for otherwise $x/\varepsilon$ would be an upper bound for $\mathbb{N}$. This property of numbers, that "multiples of a given positive quantity (no matter how small) may be as large as we want", is called *Archimedean property*.

**Intervals.** The set of numbers $a < x < b$ is called *interval* $(a, b)$, the set of numbers $a < x \leqslant b$ is called interval $(a, b]$, ... and so on. It is also useful to use the symbols $\pm\infty$ to denote intervals like $(a, \infty)$, the set of numbers $x > a$, ...

ex:

- [Ba79] 1.6.1., 1.6.2., 1.6.3., 1.6.4.
- Solve (i.e., find the value/s or interval/s of $x$)

$$3x - 1 > x + 5 \qquad |x| = 9 \qquad |x-1| = 2$$

$$x^2 \leqslant 4 \qquad (x-1)^2 > 1 \qquad |x| < 100$$

$$|x-3| \leqslant 2 \qquad |7x-2| = 3 \qquad (x-1)(x-2)(x-3) > 0$$

**Radicals and fractional powers.** The *square root* of a non-negative number $x \geqslant 0$ is the unique non-negative $y := \sqrt{x} := x^{1/2}$ such that $y^2 = x$ (the side of a square with area $x$). The *cubic root* of a non-negative number $x \geqslant 0$ is the unique non-negative $y := \sqrt[3]{x} := x^{1/3}$ such that $y^3 = x$ (the side of a cube with volume $x$). In general, the *n-th root* of a non-negative number $x \geqslant 0$ is the unique non-negative $y := \sqrt[n]{x} := x^{1/n}$ such that $y^n = x$.

Similarly, we define "fractional powers" of non-negative numbers $x \geqslant 0$ as follows: for $n, m \in \mathbb{Z}$, with $m \neq 0$, we define $x^{n/m}$ as the unique $y \geqslant 0$ such that $y^m = x^n$. Therefore, the rules

$$x^a \cdot x^b = x^{a+b} \qquad (x^a)^b = x^{ab} \qquad x^a \cdot y^a = (xy)^a$$

hold for all positive $x, y > 0$ and all rationals $a, b \in \mathbb{Q}$.

**e.g. Babilonians' problems with areas and perimeters** Besides square roots, typical problems considered by Babilonians were those involving rectangles. For example: find the sides $a$ and $b$ of a rectangle given its area $A = ab$ and its perimeter, or, equivalently, its semi-perimeter $P = a + b$. Using our modern language, we see that both sides $a$ and $b$ solve a quadratic equation, namely $x^2 + A = Px$. Observe that this means "finding the intersection between the line $x + y = P$ and the hyperbole $xy = A$". A recursive method [9] may be devised writing the problem as $x(x - P) + A = 0$, and therefore trying to solve simultaneously

$$x = P - \frac{A}{x} \qquad \text{and} \qquad x = \frac{A}{P - x} .$$

**Solving a quadratic equation.** We pose the problem to solve the quadratic equation

$$ax^2 + bx + c = 0$$

where, of course, $a \neq 0$ (for otherwise the equation would not be quadratic!). We divide the l.h.s. by $a \neq 0$, and "complete the square", as

$$\begin{aligned} x^2 + (b/a)x + c/a &= x^2 + 2(b/2a)x + (b/2a)^2 - (b/2a)^2 + c/a \\ &= (x + b/2a)^2 - b^2/4a^2 + c/a \end{aligned}$$

This is zero when

$$(x + b/2a)^2 = (b^2 - 4ac)/4a^2$$

Taking the square root, we see that two possible values of $x$ are given by the well known *resolvent fórmula*

$$x_\pm = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

In particular, we get two real solutions when the *discriminant* $\Delta := b^2 - 4ac$ is $\Delta > 0$, one real solution (to be interpreted as two coincident solutions!) when $\Delta = 0$, and no real solutions (but two complex conjugate solutions $z_\pm = (-b \pm i\sqrt{|\Delta|})/(2a)$) when $\Delta < 0$.

**ex:**

- Solve

$$x^2 - x - 1 = 0 \qquad x^2 + 3x = 0 \qquad 3x^2 - 6x + 2 = 0 \qquad x^2 + 6x + 9 = 0$$

- Find a quadratic polynomial with roots 2 and $-7$.
- Find the sum and the product of the solutions of $x^2 - 5x + 6 = 0$.
- Find the interval defined by $x^2 < x + 1$.

**Means.** The *arithmetic mean* of the numbers $a$ and $b$ is $\frac{a+b}{2}$. The *geometric mean* of the positive numbers $a$ and $b$ is $\sqrt{ab}$ (the side of a square with area equal to the area of the rectangle with sides $a$ and $b$).

- Show that the arithmetic mean between two positive numbers is never smaller that their geometric mean (compute the difference between the squares of both means)

[9] E.L. Lima, *Matemática e Ensino*, Gradiva, 2004.

**Scientific notation.** Se um observável chamado $x$ é observado/medido $n$ vezes, e se $x_1, x_2, x_3, \ldots, x_n$ são os valores obtidos nas $n$ observações, é natural estimar o "valor verdadeiro" de $x$ com a *média aritmética*

$$\boxed{\overline{x} := \frac{1}{n}\sum_{k=1}^{n} x_k = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)}$$

Umas medidas da dispersão dos dados são o *desvio quadrático médio*

$$\boxed{S^2 := \frac{1}{n}\sum_{k=1}^{n}(x_k - \overline{x})^2}$$

e o *desvio padrão* (*standard deviation*, ou *standard uncertainty*) $S := \sqrt{S^2}$. Se $n$ é grande, é razoável esperar que o valor verdadeiro esteja no intervalo

$$\boxed{\overline{x} \pm (S/\sqrt{n})}$$

com grande probabilidade. O "erro relativo" $(S/\sqrt{n})/\overline{x}$ indica a quantidade dos dígitos significativos, ou seja confiáveis, na estimação de $x$. Por exemplo, uma tabela das constantes da física tem este valor da constante de gravitação de Newton:

$$G = 6.673(10) \times 10^{-11}\mathrm{m}^3\mathrm{kg}^{-1}\mathrm{s}^{-2} \text{ with relative standard uncertainty } 1.5 \times 10^{-3}$$

Isto quer dizer que, embora a média observada seja $6.67310 \times 10^{-11}\mathrm{m}^3\mathrm{kg}^{-1}\mathrm{s}^{-2}$, só podemos confiar nos primeiros três dígitos decimais deste valor, e portanto escrever $G \simeq 6.673 \times 10^{-11}$. Os *algarismos significativos* são apenas os primeiros 4, ou seja, 6.673, e o expoente $-3$ é a *ordem de grandeza* de G (no sistema mks).

- A média aritmética $\overline{x}$ é o valor de $a$ que minimiza a soma

  $$(x_1 - a)^2 + (x_2 - a)^2 + \cdots + (x_n - a)^2$$

  dos quadrados dos "desvios" nas distintas observações.

- Transforme em notação científica os seguintes dados (de Wikipedia):
  a massa de um eletrão $\simeq$ 0.00000000000000000000000000000091093822 kg,
  a massa da Terra $\simeq$ 5973600000000000000000000 kg.
  a circunferência da Terra $\simeq$ 40000000 m.

- [Ba79] 1.10.8, 1.10.9., 1.10.10., 1.10.12., 1.10.12., 1.10.14, 1.10.15, 1.10.16. 1.10.21., 1.10.22. e 1.10.23.

# 3 Euclidean spaces

**A reta real.** Fixada uma origem (ou seja, um ponto 0), um unidade de medida e uma orientação (ou seja, uma direção "positiva"), é possivel representar cada ponto de uma reta/linha com um número real $x \in \mathbb{R}$. Vice-versa, ao número $x \in \mathbb{R}$ corresponde o ponto da reta posto a distância $\sqrt{x^2}$ da origem, na direção positiva se $x > 0$ e negativa se $x < 0$.

**O plano cartesiano.** O *plano cartesiano*[10] $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$ é o conjunto dos *pontos* $\mathbf{r} = (x, y)$, com $x, y \in \mathbb{R}$. A *origem* é o ponto $\mathbf{0} := (0, 0)$. O ponto $\mathbf{r} = (x, y)$ pode ser pensado como o *vetor* (i.e. o segmento orientado) entre a origem $(0, 0)$ e o ponto $(x, y)$. A *soma* dos vetores $\mathbf{r} = (x, y)$ e $\mathbf{r}' = (x', y')$ é o vetor

$$\mathbf{r} + \mathbf{r}' := (x + x', y + y')\,,$$

que representa a diagonal do paralelogramo de lados $\mathbf{r}$ e $\mathbf{r}'$. O *produto* do número/escalar $\lambda \in \mathbb{R}$ pelo vetor $\mathbf{r} = (x, y)$ é o vetor

$$\lambda\mathbf{r} := (\lambda x, \lambda y)$$

que representa uma dilatação/contração (e uma inversão se $\lambda < 0$) de razão $\lambda$ do vetor $\mathbf{r}$. Cada vetor pode ser representado de maneira única como soma

$$\mathbf{r} = (x, y) = x\mathbf{i} + y\mathbf{j}\,,$$

onde $\mathbf{i} := (1, 0)$ e $\mathbf{j} := (0, 1)$ denotam os vetores da base canónica.

Lugares geométricos (pontos, retas, circumferências, parábolas, . . . ) podem ser descritos/definidos por equações algébricas, ditas "equações cartesianas".

- Descreva as coordenadas cartesianas dos pontos da reta que passa por $(1, 2)$ e $(-1, 3)$.
- Descreva as coordenadas cartesianas do triângulo de vértices $(0, 0)$, $(1, 0)$ e $(0, 2)$.
- Esboce os lugares geométricos definidos pelas equações

$$xy = 1 \qquad y = 2x - 7 \qquad (x + 1)^2 + (y - 3)^2 = 9 \qquad x - 2y^2 = 3$$

e.g. **O espaço, o espaço-tempo e o espaço de fases da física newtoniana.** O espaço onde acontece a física newtoniana é o *espaço 3-dimensional* $\mathbb{R}^3 := \mathbb{R} \times \mathbb{R} \times \mathbb{R}$. A posição de uma partícula num referencial inercial é um vetor

$$\mathbf{r} = (x, y, z) := x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \in \mathbb{R}^3$$

onde $\mathbf{i} := (1, 0, 0)$, $\mathbf{j} := (0, 1, 0)$ e $\mathbf{k} := (0, 0, 1)$ denotam os vetores da base canónica.

A *lei horária/trajetória*, de uma partícula é uma função $t \mapsto \mathbf{r}(t)$ que associa a cada tempo $t \in I \subset \mathbb{R}$ a posição $\mathbf{r}(t) = (x(t), y(t), z(t)) \in \mathbb{R}^3$ da partícula no instante $t$. A *velocidade* da partícula no instante $t$ é o vetor $\mathbf{v}(t) := \dot{\mathbf{r}}(t) = (\dot{x}(t), \dot{y}(t), \dot{z}(t))$. A *aceleração* da partícula no instante $t$ é o vetor $\mathbf{a}(t) := \dot{\mathbf{v}}(t) = \ddot{\mathbf{r}}(t) = (\ddot{x}(t), \ddot{y}(t), \ddot{z}(t))$, determinado pela equação de Newton[11]

$$m\,\mathbf{a}(t) = \mathbf{F}(\mathbf{r}(t))$$

onde $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ é um campo de forças e $m > 0$ a massa da partícula.

O *espaço-tempo*[12] da física newtoniana é o produto cartesiano $\mathbb{R} \times \mathbb{R}^3 \approx \mathbb{R}^4$, o espaço dos *eventos* $(t, x, y, z) \in \mathbb{R}^4$, onde $\mathbf{r} = (x, y, z) \in \mathbb{R}^3$ representa uma posição num referencial inercial, e $t \in \mathbb{R}$ é o *tempo absoluto.*

O *estado* de uma partícula, a informação necessária e suficiente para determinar a trajetória futura (e passada), é um ponto $(\mathbf{r}, \mathbf{p}) \in \mathbb{R}^3 \times \mathbb{R}^3 = \mathbb{R}^6$ do *espaço dos estados/de fases* , onde $\mathbf{r}$ é a posição e $\mathbf{p} := m\mathbf{v}$ é o *momento (linear).*

[10] René Descartes, *La Géométrie* [em *Discourse de la Méthode*, 1637].

[11] Isaac Newton, *PhilosophiæNaturalis Principia Mathematica*, 1687.

[12] "Cette manière de considérer les quantités de trois dimensions est aussi exacte que l'autre, car les lettres peuvent toujours être regardées comme représentant des nombres rationnels ou non. J'ai dit plus haut qu'il n'était pas possible de concevoir plus de trois dimensions. Un homme d'esprit de ma connaaisance croit qu'on pourrait cependant regarder la durée comme une quatriéme dimension, et que le produit temps par la solidité serait en qualque manière un produit de quatre dimensions; cette idée peut être contestée, mais elle a, ce me semble, quelque mérite, quand ce ne serait que celui de la nouveauté." [Jean-le-Rond D'Alembert, *Encyclopédie*, Vol. 4, 1754.]

- Determine a "dimensão" do espaço de fases de um sistema composto por 8 planetas (como, por exemplo, Mercúrio, Vênus, Terra, Marte, Júpiter, Saturno, Urano, Netuno) e de um sistema composto por $6 \times 10^{23}$ moléculas.

e.g. **Reações químicas.** O estado de uma reação química

$$aA + bB + cC + \dots \quad \to \quad xX + yY + zZ + \dots$$

entre os $n$ reagentes $A$, $B$, $C$, ... e os $m$ produtos $X$, $Y$, $Z$, ... é descrito usando as concentrações $[A]$, $[B]$, $[C]$, ..., $[X]$, $[Y]$, $[Z]$, ..., e portanto $n+m$ números (entre 0 e 1).

**Spaces and coordinates.** The space where we think we live in is the 3-dimensional space $\mathbb{R}^3$. This means that we need 3 numbers, for example the Cartesian coordinates $x$, $y$ and $z$ in a fixed reference frame, to uniquely define/indicate the position of a planet at a given time. A rattlesnake in the Sonora desert thinks she lives in a plane, since she need just two coordinates, say $x$ and $y$, to say her friend where she lives. Similarly, a chemist who is describing a reaction like

$$A + 2B + 3C \leftrightarrows 4D + 5E + 6F + 7G$$

needs 7 numbers, the concentrations $a = [A]$, $b = [B]$, ..., $g = [G]$ of the seven reagents, to describe to his collegues the state of the reaction at a given time.

**O espaço vetorial $\mathbb{R}^n$.** O *espaço vetorial real* de *dimensão* $n$ é o espaço

$$\mathbb{R}^n := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n \text{ vezes}}$$

das $n$-uplas $\mathbf{x} = (x_1, x_2, \dots, x_n)$ de números reais, ditas *vetores* ou *pontos*, munido das operações *adição* : $\mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ , definida por

$$\boxed{\mathbf{x}, \mathbf{y} \mapsto \mathbf{x} + \mathbf{y} := (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)}$$

e *multiplicação por um escalar* : $\mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$ , definida por

$$\boxed{\lambda, \mathbf{x} \mapsto \lambda\mathbf{x} := (\lambda x_1, \lambda x_2, \dots, \lambda x_n)}$$

O *vetor nulo/origem* é o vetor $\mathbf{0} := (0, 0, \dots, 0)$, tal que $\mathbf{x} + \mathbf{0} = \mathbf{x}$ para todo $\mathbf{x} \in \mathbb{R}^n$. O *simétrico* do vetor $\mathbf{x} = (x_1, x_2, \dots, x_n)$ é o vetor $-\mathbf{x} := (-x_1, -x_2, \dots, -x_n)$, tal que $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$. Isto justifica a notação $\mathbf{x} - \mathbf{y} := \mathbf{x} + (-\mathbf{y})$.

A "combinação linear" dos vetores $\mathbf{v}_1$, $\mathbf{v}_2$, ..., $\mathbf{v}_k \in \mathbb{R}^n$ com "coeficientes" $\lambda_1$, $\lambda_2$, ..., $\lambda_k \in \mathbb{R}$ é o vetor

$$\sum_{i=1}^{k} \lambda_i \mathbf{v}_i := \lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_k \mathbf{v}_k \,.$$

A *base canónica* de $\mathbb{R}^n$ é o conjunto ordenado dos vetores

$$\boxed{\mathbf{e}_1 = (1, 0, \dots, 0) \quad \mathbf{e}_2 = (0, 1, 0, \dots, 0) \quad \dots \quad \mathbf{e}_n = (0, \dots, 0, 1)}$$

assim que cada vetor $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ é uma combinação linear única

$$\mathbf{x} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \cdots + x_n \mathbf{e}_n$$

dos vetores da base canónica. O número $x_k$ é chamado $k$-ésima *coordenada* do vetor $\mathbf{x}$. Outra notação usada nos manuais para os vetores é $\vec{x}$.

As coordenadas no plano Euclidiano ou no espaço 3-dimensional são também denotadas, conforme a tradição, por $\mathbf{r} = (x, y) = x\,\mathbf{i} + y\,\mathbf{j} \in \mathbb{R}^2$ ou $\mathbf{r} = (x, y, z) = x\,\mathbf{i} + y\,\mathbf{j} + z\,\mathbf{k} \in \mathbb{R}^3$.

**Euclidean spaces according Descartes.** O *produto interno Euclidiano* em $\mathbb{R}^n$, denotado por $\langle \mathbf{x}, \mathbf{y} \rangle$ ou $\mathbf{x} \cdot \mathbf{y}$, é

$$\langle \mathbf{x}, \mathbf{y} \rangle := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \, ,$$

e a *norma Euclidiana* é

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

Dois vetores $x$ e $y$ são ditos *ortogonais* se $\langle \mathbf{x}, \mathbf{y} \rangle = 0$. A desigualdade de Schwarz diz que

$$| \langle \mathbf{x}, \mathbf{y} \rangle | \leqslant \| \mathbf{x} \| \cdot \| \mathbf{y} \| \, .$$

(para provar a desigualdade no caso não trivial em que $\mathbf{x} \neq 0$ e $\mathbf{y} \neq 0$, basta definir $\mathbf{u} = \mathbf{x}/\|\mathbf{x}\|$ e $\mathbf{v} = \mathbf{y}/\|\mathbf{y}\|$, e observar que $0 \leqslant \|\mathbf{u} \pm \mathbf{v}\|^2 = 2(1 \pm \langle \mathbf{u}, \mathbf{v} \rangle)$, donde $-1 \leqslant \langle \mathbf{u}, \mathbf{v} \rangle \leqslant 1$). O *ângulo* $\theta \in [0, \pi]$ entre os vetores não nulos $\mathbf{x}$ e $\mathbf{y}$ é definido pela identidade $\langle \mathbf{x}, \mathbf{y} \rangle = \|\mathbf{x}\| \cdot \|\mathbf{y}\| \cdot \cos(\theta)$. A *distância Euclidiana* entre os pontos $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ é definida por

$$\mathrm{d}(\mathbf{x}, \mathbf{y}) := \| \mathbf{x} - \mathbf{y} \| \, .$$

Em particular, o comprimento do vetor $\mathbf{x}$, a distância entre $\mathbf{x}$ e 0, é dado pelo teorema de Pitágoras

$$\mathrm{d}(\mathbf{x}, 0) = \| \mathbf{x} \| = \sqrt{x_1^2 + \cdots + x_n^2} \, .$$

A desigualdade de Schwarz implica a desigualdade do triângulo

$$\mathrm{d}(\mathbf{x}, \mathbf{y}) \leqslant \mathrm{d}(\mathbf{x}, \mathbf{z}) + \mathrm{d}(\mathbf{z}, \mathbf{y})$$

(para provar a desigualdade, calcule $\|x + y\|^2$ e use a desigualdade de Schwarz).

A bola aberta de centro $\mathbf{a} \in \mathbb{R}^n$ e raio $r > 0$ é o conjunto $B_r(\mathbf{a}) := \{\mathbf{x} \in \mathbb{R}^n \text{ s.t. } \|\mathbf{x} - \mathbf{a}\| < r\}$. 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 $B_\varepsilon(a) \subset A$, com $\varepsilon > 0$ suficientemente pequeno.

**Lines and planes.** A reta que passa pelo ponto $\mathbf{a} \in \mathbb{R}^n$ na direção do vetor não nulo $\mathbf{v} \in \mathbb{R}^n$ é

$$\mathbf{a} + [\mathbf{v}] := \{\mathbf{a} + t\mathbf{v} \ \text{ com } \ t \in \mathbb{R}\} \, .$$

A reta perpendicular/normal ao vetor não nulo $n \in \mathbb{R}^2$ que passa pelo ponto $a \in \mathbb{R}^2$ é

$$\mathbf{a} + [\mathbf{n}]^\perp := \{\mathbf{x} \in \mathbb{R}^2 \ \text{ t.q. } \ \langle \mathbf{x} - \mathbf{a}, \mathbf{n} \rangle = 0\}$$

O plano gerado pelos vectores linearmente independentes $v$ e $w$ que passa pelo ponto $a \in \mathbb{R}^n$ é

$$\mathbf{a} + [\mathbf{v}, \mathbf{w}] := \{\mathbf{a} + t\mathbf{v} + s\mathbf{w} \ \text{ com } \ (t, s) \in \mathbb{R}^2\}$$

O plano ortogonal/perpendicular/normal ao vetor não nulo $\mathbf{n} \in \mathbb{R}^3$ que passa pelo ponto $\mathbf{a} \in \mathbb{R}^3$ é

$$\mathbf{a} + [\mathbf{n}]^\perp := \{\mathbf{x} \in \mathbb{R}^3 \ \text{ t.q. } \ \langle \mathbf{x} - \mathbf{a}, \mathbf{n} \rangle = 0\}$$

($\mathbf{n}$ é dito *vector normal* ao plano).

**Trigonometric functions.**

**Coordenadas polares.** As coordenadas polares $(\rho, \theta)$, com $\rho \in \mathbb{R}_+$ e $\theta \in [0, 2\pi[$, no plano estão definidas por

$$x = \rho \cos(\theta)$$
$$y = \rho \sin(\theta)$$

onde $x$ e $y$ são as coordenadas cartesianas de $\mathbb{R}^2$. Em particular, $\rho = \sqrt{x^2 + y^2}$ é a norma do vetor $(x, y)$.

# 4 Linear transformations and matrizes

STILL MISSING

# 5 Linear systems

STILL MISSING

# 6 Area, volume and determinants

STILL MISSING

# 7 Eigenvalues and eigenvectors

STILL MISSING

# 8 Sequences and series

**ref:** [Ap69] [EK05]

**e.g.** **Fibonacci numbers.** Consider the following problem, posed by Leonardo Pisano (a.k.a. Fibonacci, namely "filius Bonacci") in his *Liber Abaci*, 1202:

> *Quot paria cuniculorum in uno anno ex uno pario germinentur.*
> *Quidam posuit unum par cuniculorum in quodam loco, qui erat undique pariete circundatus, ut sciret, quot ex eo paria germinarentur in uno anno: cum natura eorum sit per singulum mensem aliud par germinare; et in secundo mense ab eorum nativitate germinant.*[13]

Let us denote by $f_n$ the number of pairs (of rabbits) in the $n$-th month. It is clear that the number $f_{n+1} - f_n$ of pairs of newborns in the $(n+1)$-th month is equal to the number of adult pairs in the $n$-th month, which is $f_{n-1}$. Therefore, we may write

$$f_{n+1} = f_n + f_{n-1}\,, \tag{8.1}$$

This is a law that recursively determine the values of $f_n$ given certain initial values $f_0$ and $f_1$.

Natural initial conditions are $f_0 = f_1 = 1$ (corresponding to Fibonacci's problem if the initial pair is made of newborn rabbits). The sequence grows quite fast, as you can see,

$$1,\ 1,\ 2\ ,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ 233,\ 377,\ 610,\ 987,\ 1597,\ 2584\ ,\ \dots$$

and the numbers soon become astronomically large. For example, after 10 years we get (assuming rabbits do not die or get murdered!)

$$f_{120} \simeq 8.67 \times 10^{24}\,,$$

larger than the Avogadro number!

**help:** An applet which computes the sequence is in my bestiario.

**help:** A Java or c++ recursive definition could be

```
int Fib(int n)
  {
    if (n==0) return 1;
    else if (n==1) return 1;
    else return Fib(n-1) + Fib(n-2);
  }
```

**e.g.** **Duplicação de células.** As experiências mostram que a população de uma colónia de bactérias, num período de tempo em que podemos considerar ilimitado o nutrimento e desprezáveis as toxinas produzidas, duplica-se em cada tempo característico $\tau > 0$. Assim, uma população inicial de $N_0$ células, dá origem a uma população de $N_1 = 2N_0$ células passado o tempo $\tau$, $N_2 = 4N_0$ células passado o tempo $2\tau$, ..., de

$$N_n = 2^n N_0$$

células passado o tempo $n\tau$. A lei recursiva que produz esta sucessão é

$$N_{n+1} = 2N_n\,.$$

Por exemplo, uma única célula dá origem a 1024 células passado um tempo $n\tau$ dado por $2^n = 1024$, ou seja, $n\tau = (\log_2 1024) \cdot \tau = 10 \cdot \tau$.

[13] Quantos pares de coelhos podem ser gerados por um par em um ano.
Alguém tem um par de coelhos, em um lugar inteiramente fechado, para descobrir quantos pares de coelhos podem ser gerados deste par em um ano: por natureza, cada par de coelhos gera cada mês outro par, e começa a procrear a partir do segundo mês após o nascimento.

**Sequences.** A (real valued) *sequence* is a collection $(x_n)_{n\in\mathbb{N}_0}$ of (real) numbers $x_n \in \mathbb{R}$, indexed (hence ordered) by a non-negative integer $n \in \mathbb{N}_0 := \{0, 1, 2, 3, \dots\}$. We may think of the index $n$ as "time", and therefore at the $n$-th term $x_n$ as the value of some "observable" (something that we may observe, i.e. measure) $x$ at time $n$. Clearly, we may as well define sequences with values in an arbitrary set $X$, for example in the Euclidean space $\mathbb{R}^d$.

Sequences may be defined as functions are. Indeed, a sequence with values in the set $X$ is nothing but a function $x : \mathbb{N}_0 \to X$, disguised by the notation $x_n := x(n)$ or $x[n]$. A second possibility is some recursive law

$$x_{n+1} = f(x_0, x_1, \dots, x_n)$$

prescribing the value of $x_{n+1}$ given the (past) values of $x_0$, $x_1$, ..., $x_n$. A third possibility, is using some property that the successive terms must have.

e.g. **Discrete-time signals.** Engineers (in digital signal processing) think at sequences as discrete-time "signals", and use the notation $x[n]$ for the value of the signal $x$ at "sample" $n$, which corresponds to a physical time $t = n\tau$, which is an integer multiple of a "sampling time" $\tau > 0$. Of course, one may also imagine a signal $x[n]$ which is defined for all samples $n \in \mathbb{Z}$, past and future.

e.g. **Arithmetic progression.** An *arithmetic progression*

$$x_n = a + nb$$

which may also be defined using the recursion $x_{n+1} = x_n + b$, with initial term $x_0 = a$. It represents the successive positions of a walk starting at $a$ with step $b$.

e.g. **The primes sequence.** The sequence

$$2,\ 3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23,\ \dots$$

whose generic term is the $n$-th prime number $p_n$. It is not clear what the recursive law could be.[14]

**Limits.** We say that the real sequence $(x_n)$ *converges* to some *limit* $a \in \mathbb{R}$, and we write $\lim_{n\to\infty} x_n = a$ or simply $x_n \to a$ (as $n \to \infty$), if for any "precision" $\varepsilon > 0$ there exists a time $\overline{n}$ such that $|x_n - a| < \varepsilon$ for all times $n \geqslant \overline{n}$. This means that the values $x_n$ are within an arbitrarily small neighborhood of $a$ as long as the time $n$ is sufficiently large.

The basic fact about limits in the real line $\mathbb{R}$ is that monotone (non-decreasing or non-increasing, i.e. satisfying $x_{n+1} \geqslant x_n$ or $x_{n+1} \leqslant x_n$, for any $n$, respectively) bounded (i.e. such that $|x_n| \leqslant M$ for some $M > 0$ and all $n$) sequences of real numbers do admit limit. For example, the limit of a bounded increasing sequence is simply the supremum of the set of values.

We also use the notation $x_n \to \pm\infty$ to say that given an arbitrarily large $K > 0$ we can find a time $\overline{n}$ such that $\pm x_n > K$ for all times $n \geqslant \overline{n}$.

Of course, there exist sequences which do not admit limits in either senses. These are, for example, oscillating sequences, as $x_n = (-1)^n$. We'll encounter sequences with much more wild behavior.

**Fundamental sequences.** A sequence $(x_n)$ is said *fundamental*, or *Cauchy sequence*, if for any precision $\varepsilon > 0$ there exists a time $\overline{n}$ such that

$$|x_n - x_m| < \varepsilon$$

for all times $n, m > \overline{n}$. Fundamental sequences are clearly bounded. It is obvious that a convergent sequence is fundamental (a triangular argument, since both $x_n$ and $x_m$ are $\varepsilon/2$-near to the limit for sufficiently large $n$ and $m$). A similar triangular argument shows that a fundamental sequence with a convergent subsequence is itself convergent. Less obvious is that any fundamental sequence in $\mathbb{R}$ is convergent. Indeed, let $X_n := \{x_k \text{ with } k \geqslant n\}$. It is clear that the $X_n$ are bounded, and therefore

[14] This is not the place to talk about it, but if you find it intriguing, you may take a look at the wonderful book by Marcus du Sautoy, The music of primes, Harper-Collins, 2003 [*A música dos números primos*, Zahar, 2008].

by the supremum axiom there exist the numbers $a_n := \inf X_n$. But the sequence $(a_n)$ is bounded and not decreasing, and therefore there exists $a = \lim_{n\to\infty} a_n$ (indeed, $a = \sup\{a_n \text{ with } n \in \mathbb{N}\}$). It is then easy to construct subsequences of $(x_n)$ which converge to $a$, and this implies that $(x_n)$ itself is convergent to $a$.

Thus, we may know that a sequence is convergent without knowing its limit! In general, convergence of all fundamental sequences is taken as a definition of (sequential) completeness of a metric space.

**Geometric progression.** The most important sequence is the *geometric progression*, defined starting from an initial term $x_0 = a$ using the recursion

$$x_{n+1} = \lambda x_n .$$

Thus, the sequence is

$$x_0 = a \qquad x_1 = a\lambda \qquad x_2 = a\lambda^2 \qquad \dots \qquad x_n = a\lambda^n \qquad \dots$$

The parameter $\lambda$ (which may be real or complex) is called *ratio*, since it is the ratio $x_{n+1}/x_n$ between successive terms of the sequence.

If $|\lambda| < 1$, it follows from Bernoulli inequality, applied to $x = 1/|\lambda| - 1 > 0$, that $|\lambda|^{-n} = (1+x)^n \geqslant 1 + nx > nx$, and therefore $0 < |\lambda|^n < 1/(nx)$. Thus, the geometric sequence converges to zero when $|\lambda| < 1$. It is constant, hence trivially convergent, when $\lambda = 1$. It also follows from Bernoulli inequality (taking $1 + x = |\lambda|$) that $|\lambda^n| \to \infty$ as $n \to \infty$ whenever $|\lambda| > 1$.

**ex:** Show that the term $x_n$ of a geometric progression is equal to the geometric mean $\sqrt{x_{n+1}x_{n-1}}$ of its neighbors.

**Computing limits.** Observe that $x_n \to a$ is equivalent to $x_n - a \to 0$. Therefore, we only need to understand how to "prove" that some sequence converges to zero, i.e. is "infinitesimal".

One possibility is to "compare" the sequence $(x_n)$ under investigation with a sequence with known behavior, as for example the geometric progression. Indeed, if $|x_n| \leqslant y_n$ for all $n$ sufficiently large, then $y_n \to 0$ implies $x_n \to 0$ too. More generally, if a sequence is bounded between two convergent sequences with common limit, then the first sequence too is convergent to the same limit, i.e.

$$\boxed{y_n \leqslant x_n \leqslant z_n \quad \text{and} \quad y_n \to a\,, \quad z_n \to a \qquad \Rightarrow \qquad x_n \to a}$$

In particular, the product of a bounded sequence times an infinitesimal one is infinitesimal too, i.e.

$$\boxed{x_n \to 0 \quad \text{e} \quad |y_n| \leqslant M \quad \Rightarrow \quad x_n \cdot y_n \to 0}$$

**Algebra of limits.** Limits are linear, namely,

$$\boxed{x_n \to a \text{ and } y_n \to b \quad \Rightarrow \quad x_n + y_n \to a + b \quad \text{and} \quad \lambda x_n \to \lambda a}$$

and behave nicely under multiplication and division, namely,

$$\boxed{x_n \to a \text{ and } y_n \to b \qquad \Rightarrow \qquad x_n \cdot y_n \to ab \quad \text{and} \quad x_n/y_n \to a/b \quad (\text{provided } b \neq 0)}$$

**ex:**

- Compute the limits when $n \to \infty$ of the following sequences, or show that they do not exist.

$$\frac{1}{n} \qquad \frac{(-1)^n}{n} \qquad (-1)^n \qquad 2^{-n} \qquad 3^n \qquad (-2)^n$$

$$\frac{10n^2+11}{n^3+n} \qquad \frac{n+1}{7n-3} \qquad \frac{9n^6-n^3}{7n^6+10^{23}n^5-3} \qquad \frac{3n+1}{n-2}\cdot\frac{2n+1}{6n-3}$$

$$\frac{\sin(n)}{n} \qquad \frac{\sin(1/n)}{n} \qquad \frac{\sin n}{\cos n} \qquad \sqrt{n+1}-\sqrt{n}$$

**Subsequences and sequential compactness.** A *subsequence* of a sequence $(x_n)$ is a sequence $(x_{n_i})$ obtained selecting only the values $x_{n_i}$ of the original sequence, where $i \mapsto n_i$ is an increasing map $\mathbb{N}_0 \to \mathbb{N}_0$.

The basic fact (that closed and bounded sets of the real line are *sequentially compact*) is that any bounded sequence admits a convergent subsequence.

**Limsup and liminf.** Sometimes we are only interested in a rough estimate of the growth of a sequence $(x_n)$. The "limsup" is the limit

$$\limsup_{n\to\infty} x_n := \lim_{n\to\infty} a_n \in \mathbb{R} \cup \{\infty\}$$

of the non-increasing sequence $a_n := \sup\{x_n, x_{n+1}, x_{n+2}, \dots\}$. The "liminf" is the limit

$$\liminf_{n\to\infty} x_n := \lim_{n\to\infty} b_n \in \mathbb{R} \cup \{-\infty\}$$

of the non-decreasing sequence $b_n := \inf\{x_n, x_{n+1}, x_{n+2}, \dots\}$.

e.g. **Tempo de meia-vida.** O decaimento de uma substância radioactiva pode ser caracterizado pelo "tempo de meia-vida" $\tau$, passado o qual aproximadamente metade dos núcleos inicialmente presentes terá decaido (dentro de uma amostra suficientemente grande). Se $q_n$ denota a quantidade de substância radioactiva presente no instante $n\tau$, com $n = 0, 1, 2, \dots$, então

$$q_{n+1} = \tfrac{1}{2}\, q_n \,.$$

Portanto a quantidade de substância radioactiva no instante $n\tau$ é $q_n = q_0 2^{-n}$, enquanto o produto do decaimento é $q_0 - q_n = q_0(1 - 2^{-n})$. Observe que $q_n \to 0$ quando $n \to \infty$.

Se a radiação solar produz núcleos radioactivos a uma taxa constante $\alpha > 0$ (i.e. $\alpha$ núcleos cada tempo $\tau$), a quantidade de núcleos radioactivos no instante $n\tau$ é dada pela lei recursiva

$$q_{n+1} = \tfrac{1}{2} q_n + \alpha \,. \tag{8.2}$$

Um equilíbrio é possível quando a quantidade inicial $q_0$ é igual a $\overline{q} := 2\alpha$, pois então $q_1 = \alpha + \alpha = q_0$, $q_2 = \alpha + \alpha = q_1 = q_0$, e assim a seguir, $q_n = \overline{q}$ para todos os $n \in \mathbb{N}$.

O que acontece se $q_0 \neq \overline{q}$ ? A equação recursiva diz que

$$\begin{aligned}
q_1 &= \tfrac{1}{2} q_0 + \alpha \\
q_2 &= \tfrac{1}{4} q_0 + \tfrac{1}{2}\alpha + \alpha \\
q_3 &= \tfrac{1}{8} q_0 + \tfrac{1}{4}\alpha + \tfrac{1}{2}\alpha + \alpha \\
&\vdots \\
q_n &= \tfrac{1}{2^n} q_0 + \left(\tfrac{1}{2^{n-1}} + \cdots + \tfrac{1}{8} + \tfrac{1}{4} + \tfrac{1}{2} + 1\right)\alpha
\end{aligned}$$

A primeira parcela $q_0/2^{n+1} \to 0$ quando $n \to \infty$, ou seja, o futuro é independente da condição inicial $q_0$. A segunda parcela tem limite $2\alpha$ quando $n \to \infty$ (uma prova está no parágrafo sobre a série geométrica!).

Uma fórmula (aparentemente) mais simples para os $q_n$ pode ser obtida usando a substituição $x_n := q_n - \overline{q}$, onde $\overline{q} = 2\alpha$ é a solução estacionária. De facto,

$$\begin{aligned}
x_{n+1} &= q_{n+1} - 2\alpha \\
&= \tfrac{1}{2} q_n + \alpha - 2\alpha \qquad \text{(usando a (8.2))} \\
&= \tfrac{1}{2}\, x_n \,,
\end{aligned}$$

ou seja, a diferênça entre $q_n$ e $\overline{q}$ é uma progressão geométrica de razão $1/2$. Portanto $x_n = x_0 2^{-n}$, donde

$$q_n = 2\alpha + (q_0 - 2\alpha) \cdot 2^{-n} \,.$$

É interessante observar que $x_n \to 0$, e de consequência $q_n \to \overline{q}$, quando $n \to \infty$. Ou seja, a quantidade de substância radioactiva converge para o valor estacionário, independentemente do valor inicial.

**ex:** O tempo de meia-vida do radiocarbono $^{14}C$ é $\tau \simeq 5730$ anos. Mostre como datar um fóssil, sabendo que a proporção de $^{14}C$ num ser vivente é conhecida.[15]

**e.g.** **Crescimento exponencial.** O crescimento exponencial de uma população num meio ambiente ilimitado é modelado pela equação recursiva

$$p_{n+1} = \lambda p_n \,,$$

onde $p_n$ representa a população no tempo $n$, dada uma certa população inicial $p_0$. Um significado do parâmetro $\lambda$ é o seguinte: em cada unidade de tempo o incremento $p_{n+1} - p_n$ da população é igual a soma de uma parcela $\alpha p_n$, onde $\alpha > 0$ é um coeficiênte de fertilidade, e uma parcela $-\beta p_n$, onde $\beta > 0$ é um coeficiênte de mortalidade.

Se a uma população que cresce segundo o modelo exponencial, é adicionada ou retirada uma certa quantidade $\beta$ em cada unidade de tempo, o modelo é

$$p_{n+1} = \lambda p_n + \beta \,,$$

onde $\beta$ é um parâmetro positivo ou negativo.

**ex:** Determine a solução estacionária de $p_{n+1} = \lambda p_n + \beta$, e a solução com condição inicial $p_0$ arbitrária (considere a substituição $x_n = p_n - \overline{p}$, onde $\overline{p}$ é a solução estacionária). Para quais valores dos parâmetros $\lambda$ e $\beta$ as soluções $p_n$ convergem para a solução estacionária quando o tempo $n \to \infty$?

**help:** An applet with the simulations is in exponentialgrowth.

**help:** A `Java` or `c++` cycle could be

```
for (int i = 0, i < n, i++)
  {
    population = lambda * population + beta;
  }
```

**e.g.** **Growth of Fibonacci numbers.** We want to understand how fast do Fibonacci numbers grow. We call $q_n := f_{n+1}/f_n$the quotients between successive Fibonacci numbers. They satisfy the recursive law

$$q_{n+1} = 1 + 1/q_n \tag{8.3}$$

which is an immediate consequence of (8.1). An applet with the sequence is in fibonacci. We compute:

$$1\,, \quad 2\,, \quad 3/2 = 1.5\,, \quad 5/3 \simeq 1.66666\,, \quad 8/5 = 1.6\,, \quad 13/8 = 1.625\,, \quad 21/13 \simeq 1.61538\,, \quad \dots$$

It turns out that the sequence of the $q_n$ converge, namely $q_n \to \phi$ as $n \to \infty$. Taking limits in the recursive equation $q_{n+1} = 1 + 1/q_n$ we see that $\phi = 1 + 1/\phi$, so that $\phi$ is a root (positive) of the polynomial $x^2 - x - 1$, i.e.

$$\phi = \frac{1+\sqrt{5}}{2} \simeq 1.6180339887498948482\dots$$

Hence, for large values of $n$ we may approximate Fibonacci law as

$$f_{n+1} \approx \phi f_n \,,$$

an exponential growth with rate $\phi$. In particular, we expect $f_n \sim \phi^n$.

The limit $\phi$ is a famous irrational, the Greeks' *ratio/proportion*. As described by Euclid[16]:

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

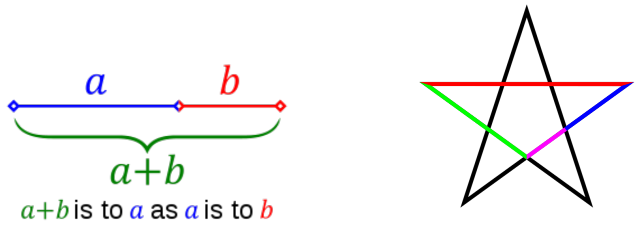
[16] Euclid, *Elements*, Book VI, Definition 3.

> *"A straight line is said to have been cut in extreme and mean ratio when, as the whole line is to the greater segment, so is the greater to the less."*

If $a$ is the greater part and $b$ the less of a line of lenght $a+b$, Euclid's requirement is

$$\frac{a+b}{a} = \frac{a}{b}$$

There follows that the ratio $\phi = a/b$ satisfies $1 + 1/\phi = \phi$. This division of an interval is used in Book IV of the *Elements* to construct a regular pentagon.




Extreme and mean ratio, and regular pentagon.
(from http://en.wikipedia.org/wiki/Golden_ratio)

e.g. **Invenção do xadrez.** Dizem que o sábio hindu Sissa inventou o jogo do xadrez e o ofereceu ao rei de Pérsia. Ao rei, que o convidou a escolher uma recompensa, pediu um grão de arroz (ou era trigo?) para o primeiro quadrado do tabuleiro, o dobro, ou seja, dois grãos, para o segundo quadrado, o dobro, ou seja, quatro grãos, pelo terceiro quadrado, e assim a seguir até o último dos quadrados do tabuleiro. O rei riu-se, num primeiro instante, mas ... a recompensa é

$$1 + 2 + 4 + 8 + \cdots + 2^{63} \simeq 1.84 \times 10^{19}$$

(see (8.6) below) grãos de arroz.

Se 1 Kg de arroz contém $\approx 30000$ grãos, isto significa algo como $6.13 \times 10^{11}$ toneladas de arroz (which you may want to compare with People's Republic of China's production in 2008, which has been, according to FAO, about $1.93 \times 10^8$ metric tons!).

**Sums.** Given a sequence $(x_n)$, one may compute the *partial sums*

$$X_n := \sum_{k=0}^{n} x_k = x_1 + x_2 + x_3 + \cdots + x_n$$

The partial sums are then obtained from the $x_n$'s by the recurrence

$$X_{n+1} = X_n + x_{n+1}\,,$$

given the initial value $X_0 = x_0$. Conversely, the original sequence is obtained from its sum computing a sort of "discrete (backward) derivative"

$$x_n = (\Delta_- x)_n := X_n - X_{n-1}\,,$$

where, of course, we must start with $X_{-1} = 0$.

Let $(x_n)$ and $(y_n)$ be two sequences, and $(X_n)$ and $(Y_n)$ be their partial sums. Rearranging the terms in the partial sums of the product sequence $x_n y_n$, we discover the *Abel transformation/formula*

$$\begin{aligned} \sum_{k=0}^{n} x_k y_k &= x_0 y_0 + \sum_{k=1}^{n} x_k (Y_k - Y_{k-1}) \\ &= x_0 y_0 - x_1 Y_0 + \sum_{k=1}^{n-1} (x_k - x_{k+1}) Y_k + x_n Y_n \qquad (8.4) \\ &= x_n Y_n - \sum_{k=0}^{n-1} (x_{k+1} - x_k) Y_k \end{aligned}$$

If we then substitute $x_k$ with its partial sum $X_k$ in Abel's formula (8.4), we get the *summation by parts formula*

$$\sum_{k=0}^{n} X_k y_k = X_n Y_n - \sum_{k=1}^{n} x_k Y_{k-1} \tag{8.5}$$

**Asymptotic averages.**

$$\frac{1}{n+1} X_n = \frac{x_0 + x_1 + \cdots + x_n}{n+1}$$

**Series.** A *series* is a formal infinite sum

$$\text{“} \sum_{n=0}^{\infty} x_n = x_0 + x_1 + x_2 + x_3 + \dots \text{”} ,$$

where the $x_n \in \mathbb{R}$ are elements of some given real (or complex) sequence. If the sequence $(X_n)$ of *partial sums*, defined as $X_n := \sum_{k=0}^{n} x_k$ (which are honest numbers) converges to some limit, say $\lim_{n\to\infty} X_n = s$, then we say the series is *convergent* (or *summable*), and that its *sum* is

$$\sum_{n=0}^{\infty} x_n := s .$$

A series $\sum_n x_n$ is *absolutely convergent* if the series $\sum_n |x_n|$, formed with the absolute values of its terms, is convergent. Of course, absolute convergence is stronger than mere convergence.

A series which is not absolutely convergent is quite a delicate object, since rearrangements of its terms may produce convergence to any real number (including infinite): i.e. the order matters! [Ap69]

e.g. **Arithmetic series.** The partial sums of an arithmetic sequence $x_n = a + nb$ is

$$\sum_{n=1}^{N} x_n = Na + \frac{N(N+1)}{2} b = \frac{N}{2}(x_N + x_1) .$$

In particular, the series diverges, unless $a = b = 0$.

e.g. **Harmonic series.** The *harmonic series* is the formal inifinite sum of the inverses of all natural numbers:

$$1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \dots$$

It is not convergent, since, if we ... to be completed!

**Geometric series.** A identidade

$$(1 + \lambda + \lambda^2 + \lambda^3 + ... + \lambda^n)(\lambda - 1) = \lambda^{n+1} - 1$$

(just multiply, and observe that all terms but the first and the last do cancel) mostra que, se $\lambda \neq 1$, a soma dos primeiros $n+1$ termos da progressão geométrica (com $a = 1$) é

$$1 + \lambda + \lambda^2 + \lambda^3 + \cdots + \lambda^n = \frac{\lambda^{n+1} - 1}{\lambda - 1} \tag{8.6}$$

Em particular, quando $|\lambda| < 1$, a *série geométrica* $\sum_{n=0}^{\infty} \lambda^n$ é (absolutamente) convergente, e a sua soma é

$$1 + \lambda + \lambda^2 + \lambda^3 + \cdots + \lambda^n + \cdots = \frac{1}{1 - \lambda} . \tag{8.7}$$

**e.g. Dichotomy paradox.** Using the above formula (8.7) for the sum of the geometric series, you may try to convince Zeno that

$$1/2 + 1/4 + 1/8 + 1/16 + 1/32 + \cdots = 1\,.$$

**e.g. Decimal expansions.** Also, you may convince yourself that $0.99999\ldots$, which by definition is the sum of the series

$$0.99999\ldots \;=\; \frac{9}{10} + \frac{9}{100} + \frac{9}{1000} + \frac{9}{10000} + \ldots\,,$$

is not "almost one" or "a bit less than one", as somebody says, but actually equal to

$$\begin{aligned} 0.99999\ldots &= \frac{9}{10}\left(1 + \frac{1}{10} + \frac{1}{100} + \frac{1}{100} + \ldots\right) \\ &= \frac{9}{10}\cdot\frac{1}{1-\frac{1}{10}} \\ &= 1\,. \end{aligned}$$

Moreover, you may learn how to recognize rational numbers as $0.33333\ldots$ or $1.285714285714\ldots$ from their (eventually) periodic expansion. Indeed, *a real number is rational if and only if its base* 10 *(or any other base $d \geqslant 2$) expansion is eventually periodic.*

**ex:** Diga se a seguintes séries são convergentes, e, se for o caso, calcule a soma.

$1 + 1/2 + 1/4 + 1/8 + 1/16 + \ldots \qquad 1 + 10 + 100 + 1000 + \ldots \qquad 1 + 1/10 + 1/100 + 1/1000 + \ldots$

$$\sum_{n=0}^{\infty}(4/5)^n \qquad 9/10 + 9/100 + 9/1000 + \ldots \qquad 0.3333\ldots \qquad 0.\overline{123}$$

**Convergence tests.** Deciding convergence or divergence of a series is not easy. The only tool at our disposal is comparison with known series, and essentially the only known non-trivial series is the geometric one. Comparison means the obvious observation that $0 \leqslant x_n \leqslant y_n$ for any $n$ sufficiently large implies the following two conclusions: $\sum_n y_n < \infty \Rightarrow \sum_n x_n < \infty$, and $\sum_n x_n = \infty \Rightarrow \sum_n y_n = \infty$.

Now, if $|x_n| \leqslant C\,\lambda^n$ for some constant $C > 0$ and any $n$ sufficiently large, then the partial sums of the series $\sum_n x_n$ are bounded by a constant times the partial sums of the geometric series $\sum_n \lambda^n$, therefore the series $\sum_n x_n$ is absolutely convergent whenever $|\lambda| < 1$. This happens when

- $\limsup_{n\to\infty} |x_n|^{1/n} < 1$ (*root test*)
- or when $\limsup_{n\to\infty} |x_{n+1}/x_n| < 1$ (*ratio test*).

[Ap69]

**e.g. The exponential.** Take $x_n = t^n/n!$. The series

$$\begin{aligned} \exp(t) &:= \sum_{n\geqslant 0}\frac{t^n}{n!} \\ &= 1 + t + \frac{t^2}{2} + \frac{t^3}{6} + \frac{t^4}{24} + \ldots \end{aligned}$$

is absolutely convergent for any $t \in \mathbb{R}$ (for example, by the ratio test). Therefore, it defines a function $\exp : \mathbb{R} \to \mathbb{R}$, which we call *exponential*, and also denote by $e^t$, if the *Neper constant* is defined by $e := \exp(1)$, i.e.

$$\begin{aligned} e &:= 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \ldots \\ &\simeq 2,718281828459045235360287 4\ldots \end{aligned}$$

Comparing the coefficients of the power series and using the binomial formula, one may show that the exponential "sends sums into products", namely

$$e^{t+s} = e^t e^s$$

for any $t, s \in \mathbb{R}$. Consequently, $e^{-t} = (e^t)^{-1}$, and in particular $e^t$ is never zero.

**Conditional convergence and rearrangements.** A series which is convergent but not absolutely convergent is called *conditionally convergent*. The standard example is the alternating harmonic series $\sum_n (-1)^n/n$. According to the *Riemann rearrangement theorem*, given any number $a \in \mathbb{R} \cup \{\pm\infty\}$, it is possible to rearrange the terms of a conditionally convergent series $\sum_n x_n$, i.e. to find a permutation $\sigma : \mathbb{N} \to \mathbb{N}$ of the naturals, and get a series $\sum_n x_{\sigma(n)}$ which converges to the desired $a$. Thus, a conditionally convergent series may simulate all convergent series!

# 9 Elementary functions

**Functions.** It turns out that a convenient language to do mathematics is that of functions, or operators. A *function* is a "law" $f : X \to Y$ that associates to any element/point $x$ of some set $X$, called *domain* of the function, a (unique!) element/point $f(x)$ of some set $Y$ (possibly different from $X$). We also use the notation

$$x \mapsto y = f(x) ,$$

that suggests that a function is a "machine", or "operator", that produces an $y$ out of any $x$. The element $y = f(x) \in Y$ is called *image* of $x \in X$, or also *value* of $f$ at $x$ if we are dealing with numbers. The *image* of the subset $A \subset X$ is the set $f(A) := \{f(a) \text{ with } a \in A\} \subset Y$. In particular, the *range* of the function $f : X \to Y$ is the set $f(X) := \{f(x) \text{ with } x \in X\} \subset Y$ of all its values. The *restriction* of the function $f : X \to Y$ to the subset $A \subset X$ is the function $f|_A : A \to Y$ defined by $f|_A(a) := f(a)$.

The *graph* of the function $f : X \to Y$ is the subset

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

of the Cartesian product of $X$ and $Y$.

The *identity* function $\text{id}_X : X \to X$ is defined by $\text{id}_X(x) = x$, and its graph is the *diagonal* $\{(x, x) \text{ with } x \in X\} \subset X \times X$.

The *compositon* of the functions $f : X \to Y$ and $g : f(X) \subset Y \to Z$ (in this temporal order!) is the function $g \circ f : X \to Z$ defined by $(g \circ f)(x) := g(f(x))$, that is, by the following sequence of operations:

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

A function $f : X \to Y$ is *into* if $x \neq x'$ implies $f(x) \neq f(x')$, and therefore the image $f(X)$ is a "copy" of $X$ inside $Y$. A function $f : X \to Y$ is *onto* if every $y \in Y$ is the image $y = f(x)$ of some $x \in X$, i.e. if $Y = f(X)$. A function $f : X \to Y$ is a *bijection/invertible* if it is into and onto, and therefore admits an *inverse* function $f^{-1} : Y \to X$, which satisfies $f^{-1}(f(x)) = x$ and $f(f^{-1}(y)) = y$ for all $x \in X$ and all $y \in Y$. Of course, an injective function$f : X \to Y$ may be considered as an invertible function $f : X \to f(X)$ onto its image.

**ex:**

- Let $f : \mathbb{N} \to \mathbb{N}$ be the function defined by $n \mapsto 2n$. Find its image $P := f(\mathbb{N})$. Find the restriction $g := f|_I : I \to \mathbb{Z}$ of $f$ to the subset $O := \{1, 3, 5, 7, \dots\} \subset \mathbb{N}$ of odd numbers, and its image $g(I)$. The function $f : \mathbb{N} \to f(\mathbb{N})$ is invertible?
- Is it true that $f \circ g$ is always equal to $g \circ f$? Never?

**Graphs of real values functions, curves.** In this section, we deal with "real functions of a real variable", that is, functions $f : X \subset \mathbb{R} \to \mathbb{R}$ defined in subsets, typically intervals, $X \subset \mathbb{R}$. Their graphs are "curves" in the Cartesian plane $x$-$y$.

**Monotone functions.** A real valued function of a real variable is *(strictly) (increasing* if $x < x'$ implies $f(x) < f(x')$, and *(strictly) decreasing* if $x < x'$ implies $f(x) > f(x')$. In both cases, it is said *(strictly) monotone.* A monotone function from an interval $X \subset \mathbb{R}$ to its image $f(X) \subset \mathbb{R}$ is invertible.

**ex:**

- Draw the graphs of

$$3 \qquad -3x \qquad |x| \qquad x-2 \qquad |x-1| \qquad |3x+5| \qquad |x-1| \pm |x-2|$$

**Linear functions.** The simplest relation between two observables, say $x$ and $y$, is proportionality:

$$y = \lambda x$$

for some non-zero $\lambda \in \mathbb{R}$, which we also abbreviate as $y \propto x$. The graph of the function $f(x) = \lambda x$ is a line through the origin, and $\lambda$ is its slope. Slightly more general is a affine dependence

$$y = \lambda x + \alpha$$

(which gives back a proportionality if the independent variable $x$ is substituted by the new variable $x' = x + \alpha/\lambda$). The parameter $\lambda$ is still the slope of the graph of $g(x) = \lambda x + \alpha$, and $\alpha$ is the value of the function for $x = 0$, i.e. the intersection of the graph with the $y$-axis. In implicit form, a linear relation is given by the law

$$ax + by = c\,,$$

which is the Cartesian equation of a generic line in the plane $x$-$y$ (including vertical lines $x = c$, which describe no relation at all!).

ex:

- Show that the substitution $x' = ax + b$ transforms the law $y = \lambda x + \alpha$ into a law $y = \lambda' x' + \alpha'$, and compute the new parameters $\lambda'$ and $\alpha'$.
- Show that the substitution $y' = ax + b$ transforms the law $y = \lambda x + \alpha$ into a law $y' = \lambda' x + \alpha'$, and compute the new parameters $\lambda'$ e $\alpha'$.
- Find the linear relation between $x$ and $y$ knowing that $y(3) = 2$ and $y(1) = 5$.

e.g. **Hubble law.** In 1929, Hubble[17] discovered the velocity-distance relation

$$v = Hd$$

of distant galaxies, suggesting the expansion of our Universe. A recent estimation of the *Hubble constant* gives the value $H = 73.8 \pm 2.4$ (km/s)/Mpc.

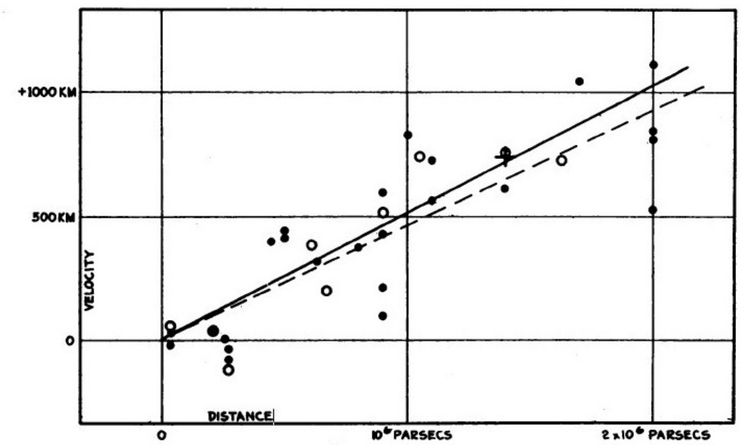



**FIGURE 1**

**Velocity-Distance Relation among Extra-Galactic Nebulae.**

Picture from the original paper by Hubble.

e.g. **Celsius, Fahrenheit and Kelvin degrees.** Temperature may be measured in Celsius ($C$), Fahrenheit ($F$) or Kelvin ($K$) degrees, and

$$F = 1.8 \cdot C + 32 \qquad K = (F + 459.67)/1.8$$

- Find the relation between Kelvin and Celsius degrees, and the ratio between one degree Kelvin and one degree Fahrenheit.
- Find the Celsius degrees of the cosmic background radiation, estimated around $3K$.

[17] E. Hubble, A relation between distance and radial velocity among extra-galactic nebulae, *Proc. N. A. S.* **15** (1929), 168-173.

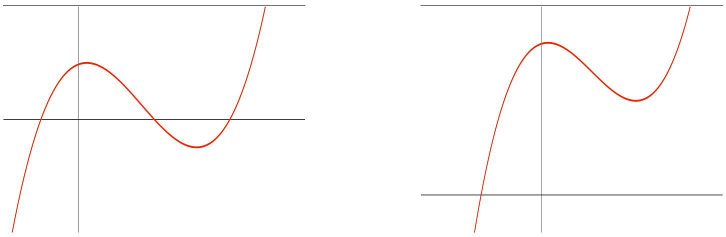
**Powers and polynomials.** A *(real) polynomial* of *degree* $n$ is a linear combination

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

of integer non-negative powers of $x$, with "coefficients" $a_0, a_1, \ldots, a_n \in \mathbb{R}$, and $a_n \neq 0$. A *root* of the polynomial $p(x)$ is a number $r$ such that $p(r) = 0$. A polynomial of degree $n \geqslant 1$ with roots $z_1, z_2, \ldots z_n$ is proportional to the "monic" polynomial

$$(x - z_1)(x - z_2)\ldots(x - z_n) = x^n - (z_1 + z_2 + \cdots + z_n)x^{n-1} + \cdots + (z_1 z_2 \ldots z_n)$$

A polynomial of degree $n$ has $n$ complex roots (some or all of which may coincide!), but a number $k \leqslant n$ of real roots (which may be zero!).



Graphs of two cubic polynomials

- Draw the graphs of

$$x^2 \qquad (x+1)^2 \qquad x^2 - 1 \qquad x^3 \qquad \sqrt{x} \qquad x^{2/3} \qquad x^{3/2} \qquad x \pm x^3$$

- Give examples of polynomials with roots 1, 2 and 3.
- Give examples of real polynomials without real roots.

**Rational functions.** Quotients of polynomials like

$$\frac{p(x)}{q(x)} = \frac{a_n x^n + \ldots a_1 x + a_0}{b_n x^m + \cdots + b_1 x + b_0}$$

defined outside the real roots of the denominator $q(x)$, are called *rational functions*.

**Cycles.** Many natural phenomena are periodic, or "quasi-periodic". A function $f(t)$ of a real variable $t$ is said *periodic* if

$$f(t + T) = f(t)$$

for all "times" $t \in \mathbb{R}$ and some minimal $T > 0$ called *period* (of the function $f$). The parameter $\omega := 1/T$ is then called *frequency* (it measures how many time a given value $f(t)$ recurs each unit of time).

**ex:**

- If $f(t)$ has period 3 and $g(t)$ has period 5, find the periods of the functions

$$2f(t+8) + 2 \qquad f(7t) \qquad f(t)^2 \qquad g(t/9) \qquad f(t) + g(t) \qquad f(t) \cdot g(t)$$

- What is the frequency of friday 13th's?

**Trigonometric functions.** The most important periodic functions (using which we may approximate all reasonable periodic functions with arbitrary precision!) are the "trigonometric functions" *sine* and *cosine*. If $\theta$ denotes the length of the arc between the point $(0,1)$ and the point $(x,y)$ along the unit circle $S^2 := \{(x,y) \in \mathbb{R}^2 \text{ t.q. } x^2+y^2=1\}$ of the Cartesian plane, then the coordinates of the final point are $x = \cos\theta$ and $y = \sin\theta$. By Pythagora's theorem,

$$\boxed{(\cos\theta)^2 + (\sin\theta)^2 = 1}$$

Sine and cosine are periodic functions, with period equal to the length of the unit circle, which is $2\pi$. Both are bounded between $\pm 1$. In particular, $\cos\theta = 0$ iff $\theta = \pi/2 + n\pi$ with $n \in \mathbb{Z}$, and $\sin\theta = 0$ iff $\theta = n\pi$, with $n \in \mathbb{Z}$. Sine e cosine satisfy the "sum and difference formulas"

$$\boxed{\cos(\theta \pm \phi) = \cos(\theta)\cos(\phi) \mp \sin(\theta)\sin(\phi) \qquad \sin(\theta \pm \phi) = \cos(\theta)\sin(\phi) \pm \sin(\theta)\cos(\phi)\,.}$$

Also useful is the *tangent* function, defined by $\tan\theta := (\sin\theta)/(\cos\theta)$, for values of $\theta \neq \pi/2 + n\pi$, with $n \in \mathbb{Z}$.

The restriction $\sin : [-\pi/2, \pi/2] \to [-1,1]$ is increasing, and therefore admits an inverse function, $\arcsin : [-1,1] \to [-\pi/2, \pi/2]$. The restriction $\cos : [0,\pi] \to [-1,1]$ is decreasing, and therefore admits an inverse function, $\arccos : [-1,1] \to [0,\pi]$.

ex:

- Verify that the functions $t \mapsto \sin(\omega t)$ and $t \mapsto \cos(\omega t)$ are periodic with period $2\pi/\omega$, and therefore frequency $\nu = \omega/(2\pi)$.
- Draw the graphs of
$$\sin(\theta \pm \pi/2) \qquad \cos(\theta) + \cos(2\theta) \qquad \sin(\theta)\cdot\sin(10\cdot\theta) \qquad \theta\cdot\sin\theta$$
- Verify that
$$(\cos\theta)^2 = \frac{1+\cos(2\theta)}{2} \qquad (\sin\theta)^2 = \frac{1-\cos(2\theta)}{2}$$
- Compute
$$\sin(\arcsin(-1/2)) \qquad \arcsin(\sin(7\pi/6)) \qquad \cos(\arccos(\sqrt{3}/2)) \qquad \arccos(\cos(-\pi/3))$$

**Exponentials and logarithms.** Given a *base* $b > 0$, it is possible to exten its fractional powers $b^{(p/q)}$ to irrational values of the exponent, and therefore define an *exponential* function

$$x \mapsto b^x$$

for all $x \in \mathbb{R}$ (a true definition will be given later!). The exponential satisfies

$$\boxed{b^x b^y = b^{x+y} \qquad b^{-x} = 1/b^x \qquad b^0 = 1}$$

In particular, it is always positive, i.e. $b^x > 0$. If the base is $b > 1$,

$$\boxed{\lim_{x\to-\infty} b^x = 0 \qquad \text{e} \qquad \lim_{x\to\infty} b^x = \infty}$$

When $b \neq 1$, the exponential $x \mapsto b^x$ is a monotone function (increasing if $b > 1$, decreasing if $0 < b < 1$). Its inverse function (defined for positive numbers!) is called base $b$ *logarithm*, and denoted by $\log_b : \mathbb{R}_+ \to \mathbb{R}$. Thus,

$$\log_b y = x \qquad \text{sse} \qquad y = b^x$$

The base 10 logarithm is also called simply $\log := \log_{10}$ by engineers. The logarithm satisfies the properties

$$\boxed{\log_b 1 = 0 \qquad \log_b xy = \log_b x + \log_b y \qquad \log_b(1/x) = -\log_b x}$$

$$\boxed{\log_b(x/y) = \log_b x - \log_b y \qquad \log_b x^y = y\log_b x}$$

Cjanging the base ...

**ex:**

- Compute

$$2^7 \qquad 3^4 \qquad 5^{-2} \qquad 10^{80} \times 10^{-27}$$

$$\log_2 16 \qquad \log_3 0.\overline{3} \qquad \log_{10} 10000 \qquad \log_{10} 0.00000001$$

**e.g. pH.** The concentration of $H_3O^+$ is measured in logarithmic scale, using the

$$pH := -\log_{10}[H_3O^+]$$

**e.g. Apparent luminosity of stars.** The *apparent luminosity* of stars is a function

$$m = m_0 - 2.5 \cdot \log_{10}(F/F_0)$$

of the flow $F$ (in a given frequency interval), where $F_0$ and $m_0$ are certain reference values.

**Logarithmic and semi-logarithmic scales.**

**Other elementary functions?**

**Limits.** Let $f : X \subset \mathbb{R} \to \mathbb{R}$ be a real valued function of a real variable, and let $a \in \mathbb{R}$ be an *accumulation point* of its domain $X$ (i.e., a point such that there exists a sequence $(x_n)$ of points of $X$ different from $a$ such that $x_n \to a$, that is, a point such that any interval $(a-\varepsilon, a+\varepsilon)$ with $\varepsilon > 0$ contains points of $X$ other than $a$). For example, $X$ may contain a union of intervals $(b,a) \cup (a,c)$, or an interval like $(b,a)$ or $(a,c)$. The number $A$ is the *limit* of $f$ when $x \to a$, notation $\lim_{x\to a} f(x) = A$ (or $\lim_{x\to a^\pm} f(x) = A$ if $X = (b,a)$ or $X = (a,c)$, respectively), if for any "precision" $\varepsilon > 0$ there exists a "tolerance" $\delta > 0$ such that an error $0 < |x-a| < \delta$, with $x \in X$, implies an error $|f(x) - A| < \varepsilon$ (observe that the actual value of $f(a)$, if any, is irrelevant!).

Limits obey the following algebraic rules:

$$\boxed{\lim_{x\to a} f(x) = F \text{ and } \lim_{x\to a} g(x) = G \quad \Rightarrow \quad \lim_{x\to a} f(x) \pm g(x) = F + G}$$

$$\boxed{\lim_{x\to a} f(x) = F \text{ and } \lim_{x\to a} g(x) = G \quad \Rightarrow \quad \lim_{x\to a} f(x) \cdot g(x) = F \cdot G}$$

$$\boxed{\lim_{x\to a} f(x) = F \text{ and } \lim_{x\to a} g(x) = G \quad \Rightarrow \quad \lim_{x\to a} f(x)/g(x) = F/G \quad (\text{ if } G \neq 0)}$$

A useful principle to compute limits is the following

$$\boxed{g(x) \leqslant f(x) \leqslant h(x) \text{ and } \lim_{x\to a} g(x) = \lim_{x\to a} h(x) = A \quad \Rightarrow \quad \lim_{x\to a} f(x) = A}$$

**ex:**

- Compute

$$\lim_{x\to 3^-} \frac{7}{x-3} \qquad \lim_{x\to 1} \frac{x^2-1}{x-1} \qquad \lim_{x\to 0^+} \frac{|x|}{x} \qquad \lim_{x\to\infty} \frac{3x^3 - 5x + 1}{5x^3 + 2x^2}$$

$$\lim_{x\to 8} \sqrt{1+x} \qquad \lim_{x\to 0} \frac{\sin x}{x} \qquad \lim_{x\to 0} x \cdot \sin(1/x) \qquad \lim_{\theta\to\pi/4} \frac{\tan\theta}{1-\cos\theta}$$

**Continuous functions.** A real valued function of a real variable $f : X \subset \mathbb{R} \to \mathbb{R}$ is *continuous* at the point $a \in X$ (the point $a$ must belong to the domain!) if for any "precision" $\varepsilon > 0$ there exists a "tolerance" $\delta > 0$ such that an error $|x - a| < \delta$, with $x \in X$, implies an error $|f(x) - f(a)| < \varepsilon$ (that is, the image of an interval of radius $\delta$ centered at $a$ belongs to an interval of radius $\varepsilon$ centered at $f(a)$). In particular, when $a$ is not an *isolated* point of $X$ (i.e. if there exists no interval $I = (a - \varepsilon, a + \varepsilon)$ with $\varepsilon > 0$ such that $X \cap I = \{a\}$), a function is continuous at $a \in X$ iff $\lim_{x \to a} f(x) = f(a)$. A *continuous* function is a function which is continuous at all points of its domain.

Powers, polynomials, trigonometric functions, exponentials and logarithms are continuous functions in their natural domains. Sums $f(x) \pm g(x)$, products $f(x) \cdot g(x)$ and quotients $f(x)/g(x)$ (where $g(x) \neq 0$) of continuous functions are continuous functions. A composition $(g \circ f)(x) = g(f(x))$ of two continuous functions $f(x)$ and $g(y)$ is a continuous function too.

**Intermediate value theorem.** A continuous function that takes positive (or negative) value $f(c) > 0$ at some point $c$ of its domain, remains positive (or negative) in some neighborhood $(c - \delta, c + \delta)$ of the point $c$. This is obvious taking, for example, $\varepsilon = |f(c)|/2$ in the definition of continuity at $c$. A consequence of this "stability of sign principle" and of the supremum axiom of the real line (the completeness axiom) is

**Theorem 9.1** (Bolzano)**.** *If a continuous function $f : [a, b] \to \mathbb{R}$ takes values $f(a)$ and $f(b)$ with opposite signs (i.e. $f(a) \cdot f(b) < 0$) then there exists a point $c \in ]a, b[$ where $f(c) = 0$*

*Proof.* Indeed, assume that $f(a) < 0$ and $f(b) > 0$ (the other case being analogous). The set $A = \{x \in [a, b]\,, \text{ s.t. } f(x) < 0\}$ is not-empty, since it contains $a$, and bounded from above, since $b$ is an upper bound. Let $c = \sup A$. The value $f(c)$ cannot be negative neither positive, for otherwise the function would be negative or positive in a whole neighborhood of $c$, and in both cases $c$ could not be the supremum of $A$. Thus, $f(c)$ must be equal to zero. □

A consequence is the

**Theorem 9.2** (Intermediate value theorem)**.** *A continuous function $f : [a, b] \to \mathbb{R}$ assumes all the values in the interval between $f(a)$ and $f(b)$, that is, if $f(a) < C < f(b)$, or if $f(b) < C < f(a)$, then there exists a point $c \in ]a, b[$ where $f(c) = C$.*

*Proof.* Just apply the Bolzano theorem to the continuous function $f(x) - C$. □

**ex:**

- Show that it is possible to solve $x^3 - x + 3 = 0$ in the interval $[-2, -1]$.
- Show that there exists a number $x$ in the interval $[0, \pi/2]$ such that $\cos x = x$ .

**Discontinuous functions.** Discontinuous functions which are useful in engineering and physics are the *integer part/floor*, defined by

$$[t] := \max \{n \in \mathbb{Z} \quad \text{s.t. } \ n \leqslant t\} \,,$$

and the *unit jump* at $\tau$, defined by

$$u_\tau(t) := \begin{cases} 0 & \text{if } t < \tau \\ 1 & \text{if } t \geqslant \tau \end{cases} .$$

- Draw the graphs of

$$f(t) = t - [t] \qquad 1 - u_0(t)$$

$$f(t) = \begin{cases} 0 & \text{if } [t] \text{ is even} \\ 1 & \text{if } [t] \text{ is odd} \end{cases} \qquad f(t) = \begin{cases} t - [t] & \text{if } [t] \text{ is even} \\ 1 + [t] - t & \text{if } [t] \text{ is odd} \end{cases}$$

**Extrema.** Uma função contínua $f : [a, b] \to \mathbb{R}$ definida num intervalo fechado e limitado possui (pelo menos) um mínimo e um máximo.

- Determine mínimos e máximos de $x(1 - x)$ no intervalo $[0, 1]$.
- Determine mínimos e máximos de $|x - 1| - |x - 2|$ no intervalo $[0, 3]$.
- Dê exemplos de funções contínuas definidas em $(0, \infty)$ ou em $(0, 1)$ sem máximos nem mínimos.

# 10 Discrete-time models and iterations

**ref:** [EK05, HK03]

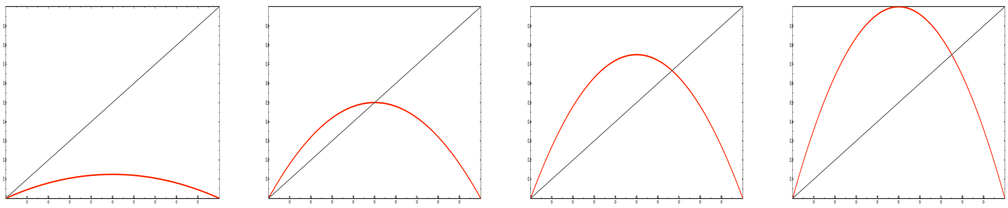
**e.g. Transformação logística.** Um modelo mais realista da dinâmica de uma população num meio ambiente limitado é

$$P_{n+1} = \lambda P_n \left(1 - P_n/M\right)$$

onde $P_n \geqslant 0$ é a população no tempo $n$, e a contante $M > 0$ é a maior população suportada pelo meio ambiente (observe que $P_{n+1} < 0$ quando $P_n > M$, o que pode ser interpretado como "extinção" no tempo $n+1$). A substituição $x_n = P_n/M$ transforma a lei acima na forma adimensional

$$x_{n+1} = \lambda x_n (1 - x_n)\,,$$

chamada *transformação logística*[18]. Se $0 \leqslant \lambda \leqslant 4$, a transformação logística $f_\lambda(x) := \lambda x(1-x)$ envia o intervalo unitário no intervalo unitário, i.e. $f_\lambda : [0,1] \to [0,1]$.



Gráficos da transformação logística $f_\lambda|_{[0,1]}$ quando $\lambda = 0.5$, $\lambda = 2$, $\lambda = 3$ e $\lambda = 4$.

Os ponto estacionários são o estado trivial 0 e

$$\overline{x} = \frac{\lambda - 1}{\lambda}$$

desde que $\lambda \geqslant 1$. Um applet `Java` com simulações do sistema está no meu bestiario.

**ex:** Discuta e interprete o comportamento das soluções para valores do parâmetro $0 < \lambda \leqslant 1$. Discuta e interprete o comportamento das soluções para valores do parâmetro $1 < \lambda \leqslant 3$. Observe os fenómenos que acontecem ao variar o parâmetro $\lambda$ entre 3 e 4. O que acontece quando $\lambda > 4$ ?

**Modelos discretos.** Um sistema dinâmico com tempo discreto é definido por uma equação/lei recursiva

$$x_{n+1} = f(x_n)\,, \tag{10.1}$$

onde $x_n \in X$ denota o *estado* (posição, população, concentração, temperatura, ...) do sistema no *tempo* $n \in \mathbb{N}_0 := \{0, 1, 2, 3, \dots\}$ (segundos, horas, meses, anos, ...). O *espaço dos estados* pode ser um intervalo $X \subset \mathbb{R}$ da recta real, um domínio $X \subset \mathbb{R}^d$ do espaço euclidiano de dimensão $d$, ou um conjunto mais exótico. A dinâmica é portanto determinada por uma *transformação* $f : X \to X$ do espaço dos estados em si mismo.

As *trajetórias* do sistema dinâmico são as sucessões $(x_n)_{n\in\mathbb{N}_0}$,

$$x_0 \quad\mapsto\quad x_1 := f(x_0) \quad\mapsto\quad x_2 := f(x_1) \quad\mapsto\quad \dots \quad\mapsto\quad x_{n+1} := f(x_n) \quad\mapsto\quad \dots\,,$$

definidas a partir de uma *condição/estado inicial* $x_0 \in X$ usando a recursão (10.1). A imagem de uma trajetória, o conjunto $\mathcal{O}(x_0) := \{x_0, x_1, x_2, \dots\} \subset X$, é dito *órbita* do estado inicial $x_0$.

**Equilíbrios e soluções periódicas.** As *soluções estacionárias*, ou *de equilíbrio*, são as trajetórias constantes $x_n = c$ para todos os tempos $n \in \mathbb{N}_0$, onde o *estado estacionário*, ou *de equilíbrio*, $c \in X$ é um "ponto fixo" da transformação $f : X \to X$, ou seja, um ponto tal que

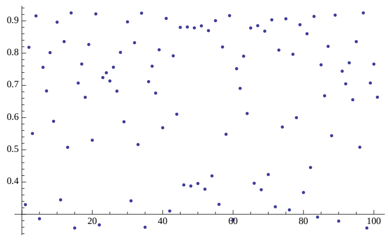
$$f(c) = c\,.$$

As *soluções periódicas* são as trajetórias $(x_n)$ tais que $x_{n+p} = x_n$ para todos os tempos $n$ e algum tempo minimal $p \geqslant 1$, dito *período*. Portanto, uma órbita periódica é um conjunto finito $\{x_0, x_1, \dots, x_{p-1}\} \subset X$ de pontos que são permutados pela transformação $f$.

[18] Robert M. May, Simple mathematical models with very complicated dynamics, *Nature* **261** (1976), 459-467.

**help:** As we already know, `Mathematica®8` may compute and plot trajectories of, for example, the logistic map $x_{n+1} = 3.7x_n(1 - x_n)$, with initial condition $x_0 = 0.33$, with the instructions

```
RecurrenceTable[{x[n + 1] == 3.7 x[n] (1- x[n]), x[0] == 0.33}, x, {n, 0, 100}]
ListPlot[%, PlotRange -> All]
```




**help:** Trajectories may be obtained with `Maxima` using the "evolution" command, as

```
(%i1) load("dynamics")$
(%i2) evolution(3.7*x*(1-x), 0, 100);
```

**Trajetórias convergentes.** Se uma trajectória $(x_n)$ é convergente e se a transformação $f : X \to X$ é contínua, então o limite $x_\infty = \lim_{n\to\infty} x_n$ é um estado estacionário, pois

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

**Solving a problem by recursion.** The above is a most useful idea in mathematics. If we are looking for a solution of some "equation" $g(x) = y$, we may try to rewrite it in the form $f(x) = x$ (in a naive way summing $x - y$ to both sides, or in some other clever way as we will encounter later), so that we are really looking for a fixed point of a transformation $f : X \to X$. Then, we may try to decide if some trajectory of the recurrence $x_{n+1} = f(x_n)$ converges. If this happens, the limit $x_\infty$ is one of the solutions we were after.

**Limits and continuity in Euclidean spaces.** Limits may be defined for sequences in any metric space $(X, d)$, simply replacing $|x_n - a|$ with the distance $\mathrm{dist}(x_n, a)$. A *metric space* is a set $X$ equipped with a *metric*, a symmetric non-negative function $\mathrm{dist} : X \times X \to [0, \infty)$ which is non-degenerate, i.e. $\mathrm{dist}(x, y) = 0$ iff $x = y$, and which satisfies the "triangular inequality"

$$\mathrm{dist}(x, y) \leqslant \mathrm{dist}(x, z) + \mathrm{dist}(z, y) \qquad \text{for any } x, y, z \in X\,.$$

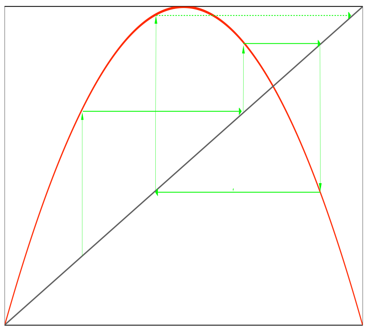
This is the case of the Euclidean space $\mathbb{R}^d$, the linear space of vectors $x = (x_1, \dots, x_d)$ equipped with the Euclidean distance $\mathrm{dist}(x, y) := \|x - y\|$, where the Euclidean norm is $\|x\| := \sqrt{\langle x, y\rangle}$ and the Euclidean inner product is $\langle x, y\rangle := x_1y_1 + \dots + x_dy_d$.

A function/map $f : X \to Y$ between two metric spaces $(X, \mathrm{dist}_X)$ and $(Y, \mathrm{dist}_Y)$ is *continuous* if whenever $x_n \to x$ in $X$ we also have $f(x_n) \to f(x)$ in $Y$ (that is, we are allowed to exchange limits with the map). Equivalently, if for any $x \in X$ and any "precision" $\varepsilon > 0$ there exists an allowed "error" $\delta > 0$ such that $\mathrm{dist}_X(x, x') < \delta$ implies $\mathrm{dist}_Y(f(x), f(x')) < \varepsilon$.

**help:** The RSolve command of Mathematica finds analytic solutions of recurrent sequations/systems, if possible.

**help:** The Nest command of Mathematica also does iterations.

**Análise gráfica.** Se o espaço dos estados é um intervalo $X \subset \mathbb{R}$, as trajetórias podem ser observadas no plano $x$-$y$ esboçando o caminho poligonal (*cobweb plot*)



$$(x_0, x_0) \mapsto (x_0, x_1) \mapsto (x_1, x_1) \mapsto (x_1, x_2) \mapsto (x_2, x_2) \mapsto (x_2, x_3) \mapsto \ldots$$

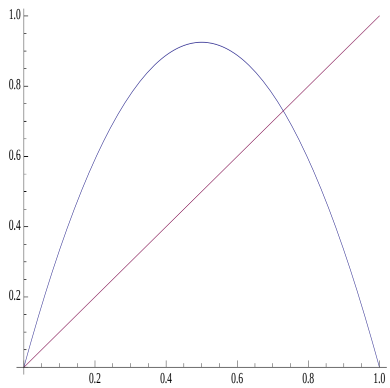
entre o gráfico da transformação, $y = f(x)$, e a diagonal, $y = x$.

**help:** You may find a GeoGebra code and the derived applet in my web page.

**help:** With Mathematica®8 , you may plot the graphs of both the map $y = f(x)$ and the identity $y = x$ with the "Plot" command

```
Plot[{3.7 x (1 - x), x}, {x, 0, 1}]
```

and get




**help:** A cobweb plot with Maxima is obtained with the "staircase" commands, as

```
(%i1) load("dynamics")$
(%i2) staircase(3*x*(1-x), 0, 10, [x, 0, 1]);
```

**ex:** Estude as trajetórias (ou seja, determine os estados de equilíbrio, as trajetórias periódicas, e o comportamento assimptótico de algumas das outras trajetórias) dos sistemas dinâmicos definidos pelas seguintes transformações do intervalo

$$f(x) = \frac{1}{3}x \qquad f(x) = 7x \qquad f(x) = -x$$

$$f(x) = 3x + 1 \qquad f(x) = 2x - 7 \qquad f(x) = \frac{1}{2}x + 5$$

$$f(x) = |1 - x| \qquad f(x) = x^2 - \frac{1}{4} \qquad f(x) = x^2 - 2$$

$$f(x) = x^3 \qquad f(x) = -x^3 \qquad f(x) = x^{1/3}$$

$$f(x) = x - x^3 \qquad f(x) = x + x^3$$

**e.g. Equilíbrio de Hardy-Weinberg.** Considere a transmissão hereditária de um gene com dois alelos, $A$ e $a$. Sejam $P_0$, $Q_0$ e $Z_0 = 1-(P_0+Q_0)$ as frequências dos genótipos $AA$, $aa$ e $Aa$, respectivamente, numa dada população inicial. Então as probabilidades de ter o alelo $A$ ou $a$ na formação de um gameta são

$$p_0 = P_0 + \frac{1}{2}Z_0 \qquad \text{e} \qquad q_0 = 1-p_0 = Q_0 + \frac{1}{2}Z_0\,,$$

respectivamente. Na fecundação, logo na primeira geração, teremos os genótipos $AA$, $aa$ e $Aa$ com probabilidades/frequências

$$P_1 = p_0^2\,, \qquad Q_1 = q_0^2 \quad \text{e} \qquad Z_1 = 2p_0q_0\,,$$

respectivamente. Sucessivamente, as probabilidades de ter os alelos $A$ ou $a$ na formação de um gameta na primeira geração são

$$p_1 = P_1 + \frac{1}{2}Z_1 \qquad \text{e} \qquad q_1 = Q_1 + \frac{1}{2}Z_1\,.$$

respectivamente. Mas $p_1 = p_0^2 + p_0q_0 = p_0$ e $q_1 = q_0^2 + p_0q_0 = q_0$. Consequentemente, as frequências dos três genótipos na segunda geração serão

$$P_2 = p_1^2 = P_1\,, \qquad Q_2 = q_1^2 = Q_1 \quad \text{e} \qquad Z_2 = 2p_1q_1 = Z_1\,.$$

Ou seja, a distribuição dos três genótipos atinge um valor estacionário a partir da primeira geração (*Hardy*[19]*-Weinberg*[20] *equilibrium/principle/law*) This is a physically interesting dynamical system with (mathematically) trivial dynamics.

**e.g. Seleção natural, modelo de Fisher, Wright e Haldane.** Um modelo simples de seleção natural, proposto por Ronald Fisher[21], Sewall Wright[22] e John Burdon Haldane[23], considera apenas um gene com dois alelos, $A$ e $a$. A vantagem ou desvantagem competitiva dos diferentes genótipos, $AA$, $Aa$ e $aa$, é modelada por coeficientes de "sucesso biológico" (*fitness*), $\phi_{AA}$, $\phi_{Aa}$ e $\phi_{aa}$, que determinam as diferentes taxas de sobrevivência, e portanto de reprodução. Sejam $0 \leqslant p_n \leqslant 1$ e $q_n = 1-p_n$ as frequências dos alelos $A$ e $a$, respetivamente, na $n$-ésima geração. Então a frequência do alelo $A$ na $(n+1)$-ésima geração é dada por

$$p_{n+1} = \frac{\alpha p_n^2 + p_nq_n}{\alpha p_n^2 + 2p_nq_n + \beta q_n^2}$$

onde $\alpha = \phi_{AA}/\phi_{Aa} > 0$ e $\beta = \phi_{aa}/\phi_{Aa} > 0$.

As soluções estacionárias são as soluções triviais 0 e 1, e, quando $\alpha$ e $\beta$ são os dois superiores ou os dois inferiores a 1,

$$\overline{p} = \frac{|\beta-1|}{|\alpha-1|+|\beta-1|}\,.$$

Quando $\alpha < 1 < \beta$ ou $\beta < 1 < \alpha$, na população assimptótica apenas sobrevive o alelo favorecido.

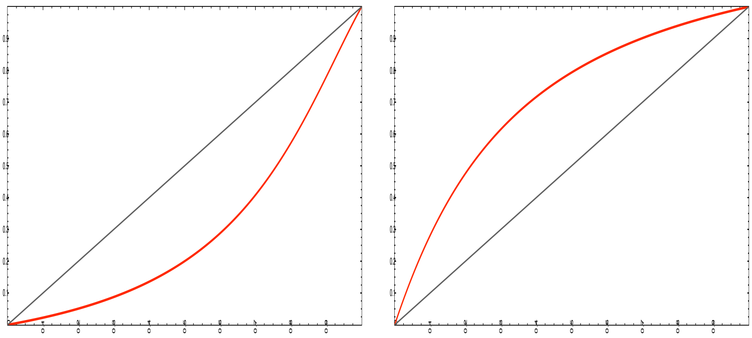
[19] G.H. Hardy, Mendelian proportions in a mixed population, *Science* **28** (1908), 49-50.

[20] W. Weinberg, Über den Nachweis der Vererbung beim Menschen, *Jahreshefte des Vereins für vaterländische Neturkunde in Württemberg* **64** (1908), 368-382.

[21] R.A. Fisher, *Genetical Theory of Natural Selection*, Clarendon Press, 1930.

[22] S. Wright, Evolution in Mendelian populations, *Genetics* **16** (1931), 97-159.

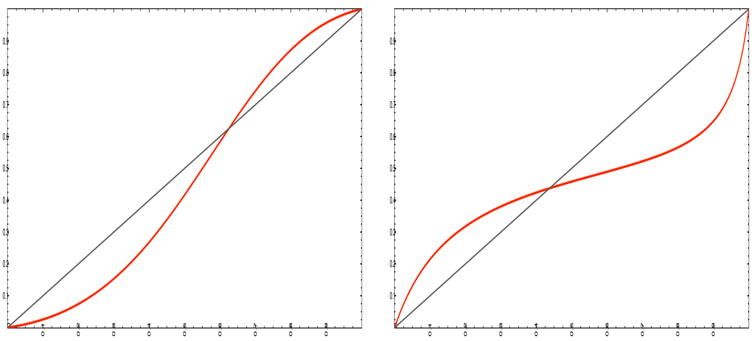
[23] J.B.S. Haldane, A Mathematical Theory of Natural and Artificial Selection (1924-1934). J.B.S. Haldane, The effect of variation on fitness, *Am. Nat.* **71** (1937), 337-349.



$\alpha < 1 < \beta$ $\qquad\qquad$ $\beta < 1 < \alpha$

Quando $\alpha > 1$ e $\beta > 1$ (ou seja, os genótipos $AA$ e $aa$ têm uma vantagem competitiva em relação ao genótipo $Aa$), o estado estacionário $\overline{p}$ é instável, e pequenas perturbações $x_0 = \overline{p} \pm \varepsilon$ do equilíbrio produzem comportamentos assimptóticos diferentes, a extinção de $A$ ou a extinção de $a$, dependendo do sinal de $\pm\varepsilon$ (*disruptive selection*).

Quando $\alpha < 1$ e $\beta < 1$ (ou seja, o genótipo $Aa$ é o favorecido), o estado estacionário $\overline{p}$ é estável, e portanto os dois alelos convivem na população assimptótica (*heterosis*).



*Disruptive selection*: $1 < \alpha < \beta$. $\qquad$ *Heterosis*: $\alpha < \beta < 1$.

**e.g. Hénon map.** The *Hénon map*[24] is the recursive map of the plane

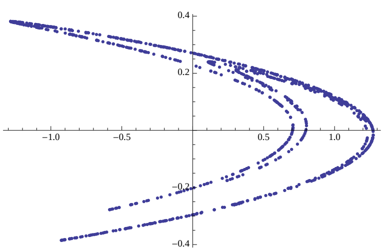
$$\begin{cases} x_{n+1} = 1 + y_n - \alpha x_n^2 \\ y_{n+1} = \beta x_n \end{cases}$$

Depending on the values of its parameters, its trajectories show regular, "intermittent" or "chaotic" behavior. If you choose the parameters $\alpha \simeq 1.4$ and $\beta \simeq 0.3$, you see the "Hénon attractor".

**help:** With `Mathematica®8` , you may use the commands

```
RecurrenceTable[{x[n + 1] == y[n] + 1 - 1.4 x[n]^2,
                                         y[n + 1] == 0.3 x[n], x[0] == 0.6, y[0] == 0.2},
                                         {x, y}, {n, 1, 1000}] // Short
ListPlot[%, PlotRange -> All]
```

to get the following picture of the "Henón attractor".




[24] M. Hénon, A two-dimensional mapping with a strange attractor, *Comm. Math. Phys.* **50** (1976), 69-77.

**help:** Two-dimensional orbits may be obtained with `Maxima` using the "evolution2d" command, as

```
(%i1) load("dynamics")$
(%i2) f:  1+y+1.4*x^2$
(%i3) g:  0.3*x$
(%i4) evolution2d([f,g], [x,y], [0,6, 0.2],  1000, [style,dots]);
```

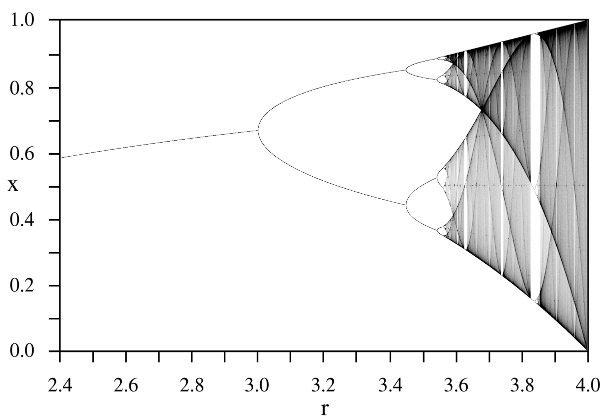
**e.g. Cigarras periódicas.** As cigarras passam quase toda a vida, um periódo de $1 \leqslant c \leqslant 17$ anos (dependendo da espécie), no chão como ninfas, e depois saem durante as poucas semanas ou meses de vida adulta (acasalar, pôr ovos e morrer). Se os predadores têm ciclos de vida de $p_k$ anos, a escolha de $c$ que minimiza os encontros é um número primo diferente dos $p_k$'s. As *magicidada* (uma cigarra americana) saem da terra cada 13 ou 17 anos, aproximadamente sincronizadas em diferentes lugares do continente. Modelos matemáticos que sugerem "explições" do fenómeno, descrito por Stephen Jay Gould em [Gou77], foram propostos a partir dos anos '70 [25] [26] [27].

A game on prime number and cicadas is in Marcus du Sautoy's page Music of the primes.

**Orbit diagram.** Consider a family of transformations

$$x_{n+1} = f_\lambda(x_n)\,,$$

depending on a parameter $\lambda$. The behavior of a typical orbit may change as $\lambda$ changes. An interesting picture is obtained if we plot the parameter $\lambda$, within some interval, versus a typical orbit of $f_\lambda$, say $\{x_{100}, x_{101}, \dots, x_{200}\}$ starting from a random initial point $x_0$.




Orbit diagram for the logistic family (from the Wikipedia).

**help:** A orbit diagram with `Maxima` is obtained with the "orbits" command, as

```
(%i1) load("dynamics")$
(%i2) orbits(a*x*(1-x), 0, 10, 100, [a, 0, 4], [style, dots]);
```

---

[25] F.C. Hoppensteadt and J.B. Keller, Syncronization of Periodical Cicada Emergences, *Sciences*, New Series, **194** (1976), 335-337.

[26] R.M. May, Periodical cicadas, *Nature*, **277** (1979), 347-349.

[27] E. Goles, O. Schulz and M. Markus, Prime number selection of cycles in a predator-prey mode, *Complexity* **6** (2001), 33-38.

# 11 Derivative

**ref:** [Ap69, RHB06]

**Slope.** We pose the problem to find, actually to "define", the slope of the graph of a real valued function $f(x)$ at some point $x_0$ of its domain. For $x \neq x_0$, we may compute the slope of the straight line passing through the points $(x_0, f(x_0))$ and $(x, f(x))$ of the Cartesian plane $x$-$y$, which is equal to the quotient

$$\frac{f(x) - f(x_0)}{x - x_0}$$

We define the *slope* of $f$ at $x_0$ as the limit, whenever it exists, of the above ratio when $x \to x_0$, namely

$$f'(x_0) := \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} . \tag{11.1}$$

We observe that a linear function $f(x) = \lambda x + a$ has constant slope equal to $f'(x_0) = \lambda$. In particular, a constant function $f(x) = a$ has zero slope $f'(x_0) = 0$ everywhere.

**Mean and instantaneous velocity.** When the the independent variable is a "time" $t$, and the function $r(t)$ represents a "position/distance" at different times, the quotient

$$\frac{r(t) - r(t_0)}{t - t_0}$$

represents the *mean velocity* between the times $t$ and $t_0$. The limit

$$v(t_0) := \lim_{t \to t_0} \frac{r(t) - r(t_0)}{t - t_0} . \tag{11.2}$$

is therefore the *(instantaneous) velocity* at time $t_0$.

e.g. **Movimento rectilíneo uniforme.** A lei do movimento rectilíneo uniforme num referencial inercial é

$$q(t) = q_0 + v_0 t \,,$$

onde $q(t) = (x(t), y(t), z(t))$ denota a posição no tempo $t$, $v_0 \in \mathbb{R}^3$ a velocidade e $q_0 \in \mathbb{R}^3$ a posição inicial.

- Determine a velocidade média no intervalo de tempos entre $t$ e $t+\varepsilon$, e a velocidade instantânea no tempo $t$.
- Determine a lei horária de uma partícula que viaja com velocidade de 3 m/s e que no instante $t = 10$ s está na posição $q(10) = 10$ m. Quando estava na origem?

e.g. **Aquiles e a tartaruga.** Aquiles (or Usain Bolt?) começa a correr com velocidade de 10 m/s em direcção de uma tartaruga que a sua vez foge com velocidade de 0.1 m/s. A distância inicial entre Aquiles e a tartaruga é de 100 m.

- Determine quanto tempo demora Aquiles a percorrer $1/2$, $1/2 + 1/4$, $1/2 + 1/4 + 1/8$, ..., da distância inicial, e passado quanto tempo chega ao ponto onde estava inicialmente a tartaruga.
- Determine a distância $d(t)$ entre Aquiles e a tartaruga no tempo $t$.
- Aquiles alcança a tartaruga? Se sim, em quanto tempo?

**e.g. Queda livre.** A queda livre de uma partícula próxima da superfície terrestre é modelada pela lei horária

$$z(t) = z_0 + v_0 t - \frac{1}{2} g t^2 \,,$$

onde $z(t)$ denota a altura da partícula no tempo $t$, $z_0$ é a altura inicial,$v_0$ é velocidade inicial, e $g \simeq 980$ cm/s$^2$ é a aceleração da gravidade próximo da superfície terrestre.

- Determine a velocidade média
$$\overline{v}_{t_0,t_1} := \frac{z(t_1) - z(t_0)}{t_1 - t_0}$$
no intervalo de tempos entre $t_0 = t$ e $t_1 = t + \varepsilon$, e a velocidade (instantânea), ou seja, o limite
$$v(t) := \lim_{\varepsilon \to 0} \frac{z(t+\varepsilon) - z(t)}{\varepsilon}$$
- Determine a aceleração da partícula, ou seja, o limite
$$a(t) := \lim_{\varepsilon \to 0} \frac{v(t+\varepsilon) - v(t)}{\varepsilon}$$
- Uma pedra é deixada cair do topo da torre de Pisa, que tem $\simeq 56$ metros de altura, com velocidade inicial nula. Calcule a altura da pedra após 1 segundo, o tempo necessário para a pedra atingir o chão e a sua velocidade no instante do impacto.
- 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?

**Derivative.** Let $f : I \to \mathbb{R}$ be a real valued function defined in some open interval $I = (a, b) \subset \mathbb{R}$. The function $f$ is *differentiable* at the point $x \in I$ if there exists the limit

$$f'(x) := \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon) - f(x)}{\varepsilon} \,, \tag{11.3}$$

called *derivative* of $f$ at $x$. Equivalently, the function $f$ is differentiable at the point $x$ if there exists a number $\lambda$, called derivative of $f$ at $x$ and denoted by $\lambda = f'(x)$, such that for all sufficiently small "variations" $\varepsilon$ we may write

$$f(x+\varepsilon) - f(x) = \lambda \cdot \varepsilon + r(\varepsilon) \tag{11.4}$$

where the "remainder" $r(\varepsilon)$ satisfies

$$\lim_{\varepsilon \to 0} \frac{r(\varepsilon)}{\varepsilon} = 0 \tag{11.5}$$

Thus, the derivative $\lambda = f'(x)$ is the "slope" of the best linear approximation

$$f(x+\varepsilon) \simeq f(x) + \lambda \cdot \varepsilon \tag{11.6}$$

to the function $f$ near the point $x$. Geometrically, this is as well the tangent line at the graph $\mathcal{G}_f := \{(x, f(x)) \,,\ x \in I\} \subset \mathbb{R}^2$ of $f$ at the point $(x, f(x))$.

Taking the limit as $\varepsilon \to 0$ in (11.4), we see that $f(x+\varepsilon) \to f(x)$. Thus, a function which is differentiable at $x$ is also continuous at $x$.

A function $f : I \to \mathbb{R}$ is *differentiable* if it admits a derivative $f'(x)$ at all points $x \in I$.

**Successive derivatives.** If $f : I \to \mathbb{R}$ admits derivatives for all $x$ in its domain, we may regard the derivative $f'$ as a function, say $f' : I \to \mathbb{R}$, hence try to compute its derivative. The derivative of the derivative of $f$ is called *second derivative* of $f$, and denoted by $f'' := (f')'$. In the same manner we may define the successive derivatives $f'''$, $f''''$ and so on (i.e. inductively as $f^{(k+1)} := (f^{(k)})'$), whenever they exist.

**Leibniz' notation.** We may write $y = f(x)$, hence denote by $\delta y := f(x+\delta x) - f(x)$ the variation of $y$ due to a variation $\delta x$ of $x$. Then the derivative is the limit

$$\lim_{\delta x \to 0} \frac{f(x+\delta x) - f(x)}{\delta x} = \lim_{\delta x \to 0} \frac{\delta y}{\delta x} := \frac{dy}{dx}$$

This is *Leibniz' notation* for derivatives. The definition (11.4) then reads $\delta y = \lambda \delta x + r(\delta x)$, with $r(\delta x)/\delta x \to 0$ as $\delta x \to 0$.

The second derivative is then $\frac{d^2 y}{dx^2}$, the third $\frac{d^3 y}{dx^3}$, and so on.

Observe that if we rescale both variables as $\tilde{x} = \lambda x$ and $\tilde{y} = \mu y$, with $\lambda > 0$ and $\mu > 0$, then the derivatives change according to

$$\frac{d\tilde{y}}{d\tilde{x}} = \frac{\mu}{\lambda} \frac{dy}{dx} \qquad \frac{d^2\tilde{y}}{d\tilde{x}^2} = \frac{\mu}{\lambda^2} \frac{d^2 y}{dx^2} \qquad \dots \qquad \frac{d^k\tilde{y}}{d\tilde{x}^k} = \frac{\mu}{\lambda^k} \frac{d^k y}{dx^k}$$

This explain the different use of the exponents in Leibniz' notation.

**Derivative as velocity, Newton/physicists' notation.** When the independent variable has the meaning of "time", hence is denoted by $t \in I \subset \mathbb{R}$, a function $t \mapsto x(t)$ represents a *trajectory*, or a *time law*, the way some observable called $x$ is changing in time. Its time derivative is then denoted using a "dot", as

$$\dot{x} := \frac{dx}{dt}$$

and referred to as a *velocity* $v := \dot{x}$, or "time variation" (newton called it "fluxione") The second derivative $a := \ddot{x}$ is also meaningful, and it is called *acceleration*. Very few (not to say none!) physical phenomena require the use of higher order time derivatives.

**ex:** Calcule as derivadas $f'(x)$ e $f''(x)$ de cada uma das seguintes funções $f(x)$ nos pontos onde existem.

$$f(x) = 2x - 3 \qquad f(x) = x^2 \qquad f(x) = |x| \qquad f(x) = \begin{cases} \frac{x}{|x|} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases}$$

**Derivatives of elementary functions.** It is clear that the derivative of a constant function $f(x) = c$ is $f'(x) = 0$. Moreover, in high school you learn to derive positive integer powers, $(x^n)' = nx^{n-1}$, and the trigonometric functions $\sin' = \cos$ and $\cos' = -\sin$.

**ex:** Use the binomial formula to prove the above formula for the derivative of powers.

**help:** With `Mathematica®8` , you may define a function $f(x) = e\ {-x^2} \cos(3x)$ with

```
f[x_] :=  Exp[-x^2] Cos[3 x]
```

and then derive with a "prime", as

```
f'[x]
```

to get

$$-2e^{-x^2} x \cos(3x) - 3e^{-x^2} \sin(3x)$$

**Algebra of derivatives.** It is clear that the derivative is linear, namely

$$(\lambda f)' = \lambda f' \qquad \text{and} \qquad (f+g)' = f' + g' \tag{11.7}$$

whenever $f$ and $g$ are differentiable functions (defined in a common interval) and $\lambda \in \mathbb{R}$ is an arbitrary constant.

The product $f \cdot g$ of two differentiable functions $f$ and $g$ is also differentiable, and its derivative is given by *Leibniz' rule*

$$(fg)' = f'g + fg' \tag{11.8}$$

Indeed, let $y = f(x)$ and $z = g(x)$. A small variation $\delta x$ induces variations $\delta y := f(x + \delta x) - f(x)$ and $\delta z := g(x + \delta x) - g(x)$. Summing and subtracting $y \cdot (z + \delta z)$ to the variation of $f \cdot g$ below, we get

$$\begin{aligned} \frac{(y+\delta y)\cdot(z+\delta z) - yz}{\delta x} &= \frac{(y+\delta y)\cdot(z+\delta z) - y\cdot(z+\delta z) + y\cdot(z+\delta z) - yz}{\delta x} \\ &= \frac{\delta y}{\delta x}\cdot(z+\delta z) + y\cdot\frac{\delta z}{\delta x} \\ &\to_{\delta x\to 0} \frac{dy}{dx}\cdot z + y\cdot\frac{dz}{dx} \end{aligned}$$

since $\delta z \to 0$ (because $g$ is continuous).

The quotient $f/g$ of two differentiable functions $f$ and $g$ is also differentiable where the denominator is $g(x) \neq 0$, and its derivative is given by the formula

$$(f/g)' = \frac{f'g - fg'}{g^2}\,. \tag{11.9}$$

To see this, we first compute the derivative of $1/g(x)$ at a point where $g(x) \neq 0$. If $z = g(x)$ and $\delta z = g(x + \delta x) - g(x)$, then

$$\frac{1/g(x+\delta x) - 1/g(x)}{\delta x} = \frac{1}{(z+\delta z)\cdot z}\cdot\frac{z-(z+\delta z)}{\delta x} \to_{\delta x\to 0} -\frac{1}{z^2}\frac{dz}{dx}\,.$$

Finally, we apply Leibniz' rule to the product $f(x) \cdot (1/g(x))$, to get (11.9).

**ex:** Calcule a derivada $f'(x)$ de cada uma das seguintes funções $f(x)$ nos pontos onde podem ser definidas.

$$\begin{array}{ccc} f(x) = 3x & f(x) = x\sin(x) & f(x) = 17 \\ f(x) = x^3 - 3x + 1 & f(x) = \sqrt{x} & f(x) = x^{-1} - x^{5/3} \\ f(x) = \dfrac{1}{x} & f(x) = \dfrac{x-1}{x^3+2} & f(x) = \dfrac{1}{\sqrt{x}} \\ \tan(x) = \dfrac{\sin(x)}{\cos(x)} & \sec(x) = \dfrac{1}{\cos(x)} & \mathrm{cosec}(x) = \dfrac{1}{\sin(x)} \end{array}$$

**ex:** Calcule as derivadas $P'(0)$, $P''(0)$, $P'''(0)$, ..., $P^{(n)}(0)$, $P^{(n+1)}(0)$, de um polinómio

$$P(x) = a_0 + a_1 x + a_2 x^2 + \dots + a_n x^n\,.$$

**ex:** Estime os seguintes valores, usando a aproximação linear $f(x + \varepsilon) \simeq f(x) + f'(x) \cdot \varepsilon$.

$$\sin(0.01) \qquad \sqrt{1.1} \qquad \cos(\pi - 0.03) \qquad \frac{1}{1+0.001}$$

**Chain rule.** Let $g : I \subset \mathbb{R} \to \mathbb{R}$ and $f : J \subset \mathbb{R} \to \mathbb{R}$ be differentiable functions, with $f(I) \subset J$, so that we may form the composition $(f \circ g) : I \to \mathbb{R}$, the function $x \mapsto f(g(x))$. If both $f$ and $g$ are differentiable, then $f \circ g$ also is differentiable, and its derivative is given by the *chain rule*

$$(f \circ g)'(x) = f'(g(x)) \cdot g'(x)\,. \tag{11.10}$$

Here Leibniz' notation is particularly meaningful. If $y = g(x)$ and $z = f(y) = f(g(x))$, then

$$\frac{dz}{dx} = \frac{dz}{dy}\cdot\frac{dy}{dx}$$

That is, you may act as if you could divide by "$dy$".

For the proof, we use Leibniz' notation again. For small $\delta x$, we define the corresponding variations $\delta y = g(x+\delta x) - g(x)$, hence $\delta z = f(y+\delta y) - f(y)$. By (11.4), there exists a function $e(\varepsilon)$, which converges to 0 when $\varepsilon \to 0$, such that

$$\delta z = \frac{dz}{dy} \cdot \delta y + e(\delta y) \cdot \delta y$$

for sufficiently small $\delta y$. Consequently,

$$\frac{\delta z}{\delta x} = \left( \frac{dz}{dy} + e(\delta y) \right) \cdot \frac{\delta y}{\delta x} \to_{\delta x \to 0} \frac{dz}{dy} \cdot \frac{dy}{dx}$$

since $\delta y \to 0$ when $\delta x \to 0$.

ex: Calcule a derivada $f'(x)$ de cada uma das seguintes funções $f(x)$ nos pontos onde podem ser definidas.

$$f(x) = \cos(x^2) \qquad f(x) = \sqrt{2x-1} \qquad f(x) = \sin\left(\sqrt{x}\right)$$

$$f(x) = (\sin(x))^2 \qquad f(x) = \sin(\cos(x^3)) \qquad f(x) = \frac{\cos(2x) - x}{\sqrt{x}}$$

**Derivatives of inverse function.** Let $f : I \subset \mathbb{R} \to \mathbb{R}$ be one-to-one function onto $J = f(I)$, and let $h : J \to I$ be its inverse, so that $h(f(x)) = x$ for all $x \in I$ and $f(h(y)) = y$ for all $y \in J$. If $f$ is differentiable at $x \in I$ and $h$ is continuous at $y = f(x)$, then $h$ is differentiable at $y$ if and only if $f'(x) \neq 0$, and if this is the case, its derivative is

$$h'(y) = \frac{1}{f'(h(y))} \,. \tag{11.11}$$

Indeed, if $h$ is differentiable at $y$, we may apply the chain rule to $h(f(x)) = x$ to get $h'(y) \cdot f'(x) = 1$, hence (11.11) whenever $f'(x) \neq 0$. Conversely, given a variation $\delta y$, let $\delta x = h(y+\delta y) - h(y)$ be the corresponding variation of $x$. Since $h$ is continuous at $y$, $\delta x \to 0$ whenever $\delta y \to 0$. Therefore,

$$\frac{\delta x}{\delta y} = \left( \frac{\delta y}{\delta x} \right)^{-1} \to_{\delta x \to 0} \left( \frac{dy}{dx} \right)^{-1}$$

provided $f'(x) \neq 0$.

ex:

- Show that the derivative of $x^{1/n}$, for $x > 0$ and $n = 1, 2, 3, \ldots$, is $\frac{1}{n} x^{1/n - 1}$.
- Calcue as derivadas das seguintes funções nos pontos onde podem ser definidas.

$$f(x) = \arcsin(x) \qquad f(x) = \arccos(x) \qquad f(x) = \arctan(x)$$

- Calcule a derivada da função inversa de $f(x) = x + x^3$ no ponto $y = 0$.

ex: **Taxas de variação.** Determine a taxa de variação

- $\frac{dA}{dr}$, onde $A$ é a área de uma circunferência e $r$ o seu raio,
- $\frac{dV}{dr}$, onde $V$ é o volume de uma bola e $r$ o seu raio,
- $\frac{dV}{d\ell}$, onde $V$ é o volume de um cubo e $\ell$ o seu lado.

e.g. **Growth of a spherical cell.** A spherical cell grows absorbing raw material (necessary for its metabolism) from its surface, hence at a rate $\alpha S$ proportional to the surface $S$. On the other hand, the rate of raw material necessary for the metabolism is proportional to the volume $V$ of the cell, say $\beta V$. Therefore, the cell can survive only when $\beta V \leqslant \alpha S$, i.e until its surface area to volume ratio (SA:V) ratio $S/V$ is grater than some lower limit $\beta/\alpha$. The SA:V of a sphere of radius $r$ is $3/r$, therefore the limit size of the cell is $\overline{r} \simeq 3\alpha/\beta$.

If the density $\rho$ is assumed constant, then the mass $\rho V$ of the cell grows according to $\rho\dot{V} = \alpha S$. If $r$ denotes the radius of the cell, we get $\rho\, 4\pi r^2 \dot{r} = \alpha\, 4\pi r^2$, i.e.

$$\dot{r} = \alpha/\rho$$

and therefore the radius increases linearly with time $t$ according to $r(t) = r(0) + (\alpha/\rho)\cdot t$. Correspondingly, the SA:V ratio decreases as $S/V \simeq 3/r(t)$, until a certain limit $3/\overline{r}$ when the incoming material can no longer support the cell metabolism.

**Derivative and growth.** A differentiable function $f(x)$ is striclty increasing in intervals where $f'(x) > 0$, strictly decreasing in intervals where $f'(x) < 0$, and constant in intervals where its derivative vanishes. Consequently, if $c$ is a local maximum or minimum of a differentiable function $f$ defined in a neighborhood $(c-\varepsilon, c+\varepsilon)$ of $c$, then $c$ is a *critical point* of $f$, i.e. a point where $f'(c) = 0$.

ex:

- Esboce os gráficos das seguintes funções, definidas em oportunos domínios.

$$f(x) = 1 - \frac{x^2}{2} \qquad f(x) = x + \frac{1}{x} \qquad f(x) = 1 + x + \frac{x^2}{2}$$

$$f(x) = \frac{1}{(x-1)(x-2)} \qquad f(x) = (x-1)(x-2)(x-3)$$

$$f(x) = x - \sin(x) \qquad f(x) = \sin(x) + \sin(2x) \qquad f(x) = \frac{\sin(x)}{x}$$

- Mostre que, entre todos os rectângulos de perímetro $P$ fixado, o quadrado é o que tem área maior.

- Mostre que, dados $n$ números $a_1, a_2, ..., a_n$, o valor de $x$ que minimiza a soma dos "erros quadráticos"

$$(x-a_1)^2 + (x-a_2)^2 + ... + (x-a_n)^2$$

  é a média aritmética

$$\overline{a} = \frac{a_1 + a_2 + ... + a_n}{n}\,.$$

**Mean value theorem and inequality.** Let $f : [a,b] \to \mathbb{R}$ be a continuous function with $f(a) = f(b)$, and assume that $f$ differentiable in $(a,b)$. By Weierstrass theorem, the function attains its maximum and minimum values, say $M$ and $m$, respectively. If both are attained at the endpoints, then the function is constant and its derivative is everywhere zero at $c(a,b)$. If, on the other side, its maximum or its minimum is attained at an internal point $c \in (a,b)$, then this must be a critical point. Therefore, there always exists a point $c \in (a,b)$ where $f'(c) = 0$. This is called *Rolle's theorem*. More generally,

**Theorem 11.1** (Mean value theorem)**.** *If $f : [a,b] \to \mathbb{R}$ is a continuous function which is differentiable in $(a,b)$, then there is a point $c \in (a,b)$ where*

$$f(b) - f(a) = f'(c)\cdot(b-a)\,.$$

*Proof.* Apply Rolle's theorem to the function $f(a) + \frac{f(b)-f(a)}{b-a}x$. □

In particular, if $f'(x) = 0$ for all $c \in (a, b)$, then $f(b) = f(a)$. Thus, a function with zero derivative for all points in an interval is constant. More interesting (and physically obvious!) is that a bound on the derivative implies a bound on the displacement: if $|f'(c)| \leqslant \lambda$ for all $c \in (a, b)$, then we have the inequality

$$|f(b) - f(a)| \leqslant \lambda \cdot |b - a| \, .$$

**ex:**

- Mostre que, se $f(x)$ é um polinómio de segundo grau, então a recta que une os pontos $(a, f(a))$ e $(b, f(b))$ é paralela à recta tangente ao gráfico de $f$ no ponto médio $\frac{a+b}{2}$.
- Mostre que para todos os $x$ e $y$

$$|\sin(x) - \sin(y)| \leqslant |x - y|$$

- Mostre que para todos $0 < y \leqslant x$

$$ny^{n-1}(x - y) \leqslant x^n - y^n \leqslant nx^{n-1}(x - y)$$

## 12 Aproximation

**ref:** [Ap69, Li06, RHB06]

**Interpolation.** (see Klein)

**Polynomial approximation.** The value $f(a)$ is the best constant approximation

$$f(x) \simeq f(a)$$

to a continuous function $f(x)$ near the point $a$, in the sense that the "error" $e(x-a) := f(x) - f(a)$ goes to $e(\varepsilon) \to 0$ as $\varepsilon := x - a \to 0$. The derivative $f'(a)$ is the slope of the best linear approximation

$$f(x) \simeq f(a) + f'(a)(x-a)$$

to a differentiable function $f(x)$ near the point $a$, since by the very definition of derivative the "error" $e(x-a) := f(x) - f(a) - f'(a)(x-a)$ is so small that $e(\varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0$.

If $f$ has $n$ derivatives at the point $a$, we may as well look for the polynomial of degree $\leqslant n$ which best approximate $f(x)$ for small $\varepsilon = x - a$, hoping to get better approximations. After all, the only functions that human beings and machines can compute are polynomials (since we only can do finite sums and multiplications), and we need means to estimate the other functions which, we believe, describe Nature.

The key observation is the following: a $n$-times differentiable function $e(\varepsilon)$ satisfies $\lim_{\varepsilon\to 0} e(\varepsilon)/\varepsilon^n = 0$ iff its value and its first $n$ derivatives vanish at zero, i.e. $e^{(k)}(0) = 0$ for all $k = 0, 1, \ldots, n$ (this is not trivial, and is a consequence of the mean value theorem, see [Li06]). If we apply this to the error $r(x-a) = f(x) - P(x-a)$, where $P$ is any polynomial of degree $\leqslant n$, we see that the error goes to zero as $r(\varepsilon)/\varepsilon^n \to_{\varepsilon\to 0} 0$ if and only if $P$ is the *Taylor polynomial* (of the function $f$ at the point $a$)

$$P_n(x-a) := f(a) + f'(a)\cdot(x-a) + \frac{f''(a)}{2}\cdot(x-a)^2 + \frac{f'''(a)}{6}\cdot(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}\cdot(x-a)^n .$$

If, moreover, the derivative $f^{(n)}$ is continuos in the closed interval $[a, x]$ (if $a < x$, or $[x, a]$ if $x < a$) and $f^{n+1}(y)$ exists for all $y \in (a, x)$, then there is a point $c \in (a, x)$ such that the error is

$$e_n(\varepsilon) := f(a+\varepsilon) - P_n(\varepsilon) = \frac{f^{(n+1)}(c)}{(n+1)!}\,\varepsilon^{n+1}$$

i.e.

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \frac{f^{(n+1)}(c)}{(n+1)!}\,(x-a)^{n+1}$$

This generalizes the mean value theorem.

**Integral formula for the error.**

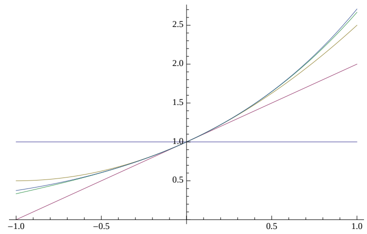
e.g. **The exponential.** We already saw that the exponential is the function defined by the power series

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!} .$$

If we limit the sum to finite degree $n$, we obtain a sequence of polynomial approximations

$$\exp x \simeq 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + \cdots + \frac{1}{n!}x^n .$$

This is what a machine computes when asked to produce $e^x$, once chosen an $n$ so large that the successive terms give no appreciable difference to the sum.




Taylor polynomials of the exponential near $a = 0$, with degrees $n = 0, 1, 2, 3, 4$.

e.g. **Trigonometric functions.** The Taylor polynomials of the trigonometric functions cos and sin centered at 0 start with

$$\cos(x) \simeq 1 - \frac{1}{2}x^2 + \frac{1}{24}x^4 + \dots \qquad \text{and} \qquad \sin(x) \simeq x - \frac{1}{6}x^3 + \frac{1}{120}x^5 + \dots$$

**ex:**

- Prove as seguintes aproximações, validas para $x$ suficientemente pequeno,

$$e^x \simeq 1 + x + \frac{x^2}{2} + \dots \qquad \log(1+x) \simeq x - \frac{x^2}{2} + \dots$$

$$\sin(x) \simeq x - \frac{x^3}{6} + \dots \qquad \cos(x) \simeq 1 - \frac{1}{2}x^2 + \dots$$

- e determine estas outras

$$\frac{1}{1-x} \simeq 1 + x + \dots \qquad \sqrt{1+x} \simeq 1 + \frac{1}{2}x + \dots$$

$$\log(1+x^2) \simeq \dots \qquad \sin(\pi e^{-x}) \simeq \dots$$

- Aproxime $e$, e estime o erro na sua aproximação, usando os polinómios de Taylor

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + r_n(x)$$

(observe que $1 \leqslant e^x \leqslant 3$ no intervalo $x \in [0,1]$).

**Contraction principle.** A *contraction* of an interval $X \subset \mathbb{R}$ is a transformation $f : X \to X$ such that there exists a constant $0 \leqslant \lambda < 1$ (strictly smaller than one!) such that

$$|f(x) - f(x')| \leqslant \lambda \cdot |x - x'| \tag{12.1}$$

for all $x, x' \in X$ (this definition extends to a generic metric space $(X, \mathrm{dist})$ if we replace the absolute value of the difference with the distance). For example, a differentiable transformation $f : X \to X$ of a (closed) interval $X \subset \mathbb{R}$ such that $|f'(x)| \leqslant \lambda < 1$ for all $x \in X$ is a contraction, since, by the mean value theorem,

$$|f(x) - f(y)| = |f'(c) \cdot (x - y)| \leqslant \lambda \cdot |x - y| \,.$$

where $c$ is some point between $x$ and $y$. Observe that a contraction is (uniformly) continuous, since for any $|x - y| < \delta = \varepsilon/\lambda$ we have $|f(x) - f(y)| < \lambda \cdot \delta < \varepsilon$.

**Proposition 12.1. (Contraction principle, or Banach fixed point theorem)** *A contraction $f : X \to X$ of a closed interval $X \subset \mathbb{R}$ (or a complete metric space) has one and only one fixed point $p$. Moreover, all trajectories defined recursively by $x_{n+1} = f(x_n)$ given an arbitrary initial condition $x_0 \in X$ converge exponentially fast to the fixed point $p$.*

Indeed, let $x_0 \in X$ and let $(x_n)$ be its trajectory, so that $x_{n+1} = f(x_n)$. If we iterate (12.1), we see that $|x_{k+1} - x_k| \leqslant \lambda^k \cdot |x_1 - x_0|$. Using $k$ times the triangular inequality, and then the convergence of the geometric series of ratio $\lambda$, we see that

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

Therefore, $(x_n)$ is a Cauchy (or fundamental) sequence. The limit $p = \lim_{n\to\infty} x_n$ exists because $X$ is a closed interval, and is a fixed point of $f$ because

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

by the continuity of $f$. Uniqueness of the fixed point is obvious, since if $p$ and $p'$ are both fixed points, then $|p - p'| = |f(p) - f(p')| \leqslant \lambda \cdot |p - p'|$ with $\lambda < 1$, and therefore $|p - p'| = 0$. On the other side, iteration of (12.1) implies that $|x_n - p| \leqslant \lambda^n \cdot |x_0 - p|$, so that the convergence $x_n \to p$ is exponential.

The last assertion suggests therefore a practical method to find, actually approximate, the fixed point: follow a trajectory!

**Estabilidade dos estados estacionários.** Seja $\overline{x}$ um estado estacionário da equação recursiva

$$x_{n+1} = f(x_n)$$

ou seja, um ponto tal que $f(\overline{x}) = \overline{x}$. Se a transformação $f(x)$ é diferenciável, o príncipio das contrações permite decidir sobre a estabilidade do estado estacionário.

Se $|f'(\overline{x})| < 1$, então o ponto fixo é *atrativo*. Existe uma vizinhança $B = [\overline{x} - \varepsilon, \overline{x} + \varepsilon]$ de $p$ tal que a restrição $f|_B : B \to B$ é uma contração, e $\overline{x}$ é o seu único ponto fixo. As trajetórias de todo o ponto $x_0$ suficientemente próximo de $\overline{x}$ (ou seja, em $B$) convergem exponencialmente para $\overline{x}$, ou seja $x_n \to \overline{x}$.

Se $|f'(\overline{x})| > 1$, então o ponto fixo é *repulsivo*: as trajetórias de todo o ponto $x_0 \neq \overline{x}$ numa vizinhança suficientemente pequena de $\overline{x}$ saem da vizinhança em tempo finito.

Se $f'(\overline{x}) = 0$, o ponto fixo $\overline{x}$ é dito *super-atrativo*. Usando o polinómio de Taylor de grau 1 com resto, vê-se que, se $x_0$ está numa vizinhança suficientemente pequena de $\overline{x}$, então a trajetória de $x_0$ converge para o ponto fixo $\overline{x}$ e a velocidade de convergência é "quadrática", ou seja,

$$|x_{n+1} - \overline{x}| \leqslant \beta \cdot |x_n - \overline{x}|^2$$

onde $\beta$ é uma constante.

**ex:**

- Estude a natureza dos pontos fixos das seguintes transformações

$$f(x) = \alpha x \qquad f(x) = \alpha x^3 \qquad f(x) = \alpha x + \beta x^2$$

  ao variar os parâmetros.

- Digite 0.1 na sua máquina de calcular, e pressione repetidamente a tecla "cos". O que acontece?

- Estude a natureza do ponto fixo não trivial do modelo logístico

$$x_{n+1} = \lambda x_n(1 - x_n)$$

  ao variar o parâmetro $\lambda$.

- Estude a natureza do ponto fixo no método de Heron para determinar a raiz quadrada de $a > 0$, dado pela iteração

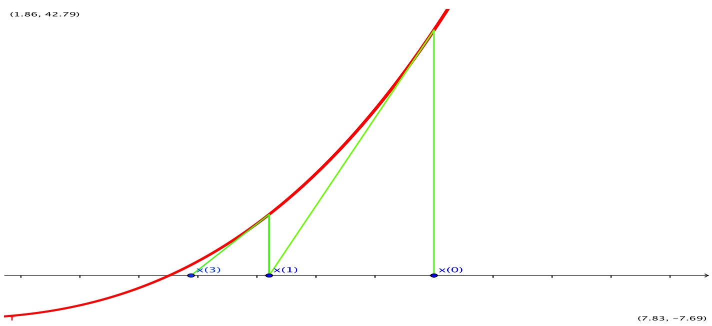
$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) .$$

**e.g. Newton-Raphson iterative scheme.** Finding $\sqrt{a}$ means solving the polynomial equation $z^2 - a = 0$. What about finding roots of a generic polynomial $p(x) \in \mathbb{R}[x]$ ?

Newton's idea consists in improving an initial guess $x_0$ using the root of the linear approximation $p(x) \simeq p(x_0) + p'(x_0)(x - x_0)$, which is $x_1 = x_0 - p(x_0)/p'(x_0)$. This amounts to the iterative scheme

$$x_{n+1} = x_n - \frac{p(x_n)}{p'(x_n)} \,. \tag{12.2}$$

If the trajectory converges, i.e. $x_n \to x_\infty$, and if $p'(x_\infty) \neq 0$, then clearly the limit $x_\infty$ is a root of $p$. On the other hand, if $c$ is a root where $p'(c) \neq 0$, then $c$ is a super-attractive fixed point of the map $x \mapsto f(x) ;= x - p(x)/p'(x)$. Therefore, an initial guess $x_0$ sufficiently near $c$ will produce a trajectory $(x_n)$ which converges to $c$ (quadratically fast, i.e. such that $|x_{n+1} - c| \leqslant \beta \cdot |x_n - c|^2$ for some constant $\beta > 0$).




Search for a root of $x^3 - 2x - 5$ using Newton iterations.

**help:** Mathematica®8 search for a root of an equation like $x^7 - 13x^5 + 9 = 0$ (or even more complicated equations, involving transcendental functions!) using the Newton iterative scheme starting with the initial guess $x_0 = 10$ with the instruction

```
FindRoot [x^7 - 13 x^5 + 9 == 0, {x, 10}]
{x -> 3.6035}
```

**ex: Exercícios.**

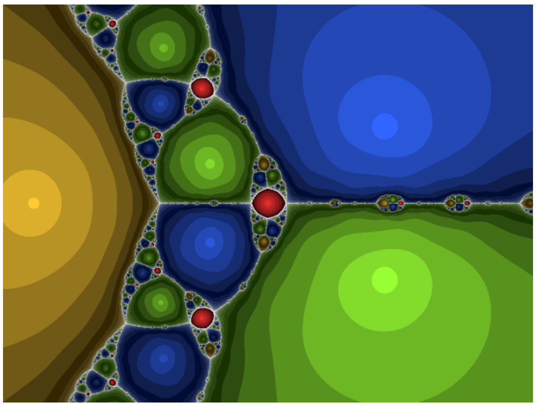
- Use Newton method to solve Newton's problem, i.e. find the roots of $x^3 - 2x - 5$.
- Show that Newton method to solve $x^2 - a = 0$ corresponds to babylonian-Heron iterative scheme.
- Use o método de Newton para aproximar a "razão", a raiz positiva de $x^2 - x - 1$. Then, compare with the babylonian-Heron method (i.e., estimate $\sqrt{5}$, then sum 1 and divide by 2).
- Write and implement Newton method to find $n$-th roots, i.e. to solve $x^n - a = 0$.
- Utilize o método de Newton para estimar raizes de

$$z^2 + 1 + z \qquad z^3 - z - 1 \qquad z^5 + z + 1 \qquad z^3 - 2z - 5$$

**e.g. Newton's fractals.** Em 1879 Cayley observou que o método pode ser utilizado também para aproximar raízes complexas de polinómios $p(z) \in \mathbb{C}[z]$. A receita consiste em iterar a função racional

$$f(z) = z - \frac{p(z)}{p'(z)}$$

O problema é decidir quando, ou seja para quais valores da conjetura inicial $z_0$, a sucessão $(z_n)$, com $z_{n+1} = f(z_n)$, converge para uma raiz de $p(z)$. As bacias de atração das diferentes raizes desenham padrões surprendentes no plano complexo



Basins of attraction of the three roots of $2z^3 - 2z + 2$ in the complex plane.
(from http://en.wikipedia.org/wiki/Newton_fractal).

**Iteração de funções racionais na esfera de Riemann.** É natural considerar iterações de funções racionais $f(z) \in \mathbb{C}(z)$ arbitrárias (os endomorfismos da esfera de Riemann $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$), e querer descrever as trajecórias definidas pela equação recursiva $z_{n+1} = f(z_n)$.

O exemplo mais estudado consiste nas iterações da família de polinómios quadráticos

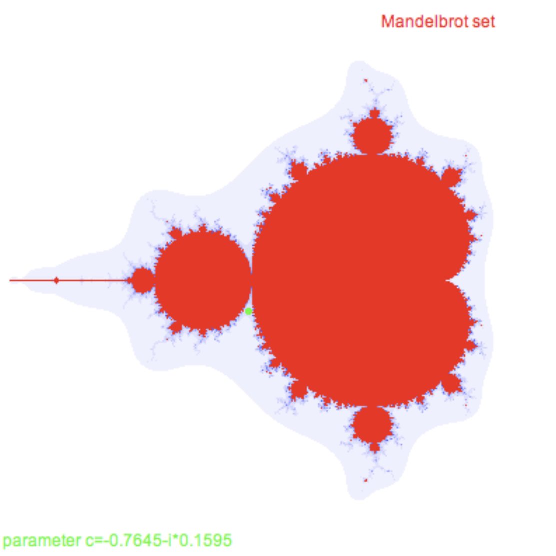
$$f(z) = z^2 + c$$

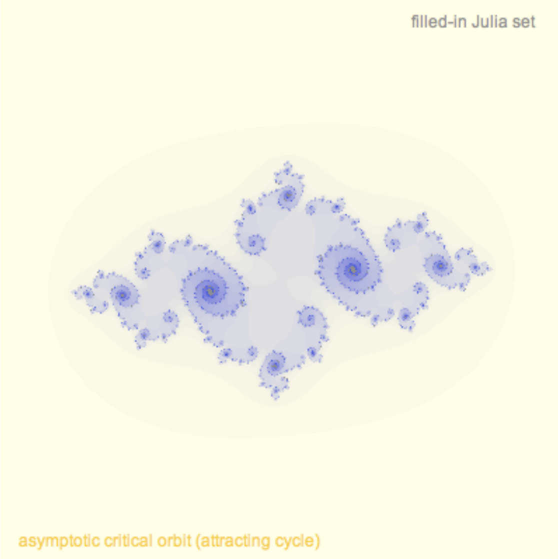
ao variar o parâmetro $c \in \mathbb{C}$. A sua beleza foi intuida por Gaston Julia[28] e Pierre Fatou[29] no princípio do século XX, desvendada com o auxílio dos computadores modernos por Benoît Madelbrot, e estudada por uma multidão de excelentes matemáticos (como Adrian Douady, Dennis Sullivan, John Milnor, Misha Lyubich, Jean-Christophe Yoccoz, Curtis McMullen, ...) a partir dos anos '80 do século passado.

**Nice pictures.** Em baixo, está uma imagem que nos tempos de Julia e Fatou apenas era possível ver com uns olhos matemáticos bem afinados (um applet Java que produz a figura está no meu bestiario). O laço de corações vermelhos à esquerda, chamado *Mandelbrot set*, consiste nos valores do parâmetro complexo $c$ tais que a órbita do ponto crítico $z_0 = 0$ permanece limitada. A região cinzenta à direita, chamada *filled-in Julia set*, consiste no conjunto das condições iniciais $z_0$ cuja órbita é limitada. As outras côres (que permitem ver os conjuntos "invisíveis" de Cantor) são escolhidas dependendo da velocidade com que as trajectórias $z_n$ fogem para o infinito.

[28] G. Julia, Mémoire sur l'iteration des fonctions rationnelles, *Journal de Mathématiques Pures et Appliquées*, **8** (1918), 47-245.

[29] P. Fatou, Sur les substitutions rationnelles, *Comptes Rendus de l'Académie des Sciences de Paris*, **164** (1917) 806-808, and **165** (1917), 992-995.




Mandelbrot set (left) and Julia set of the polynomial $z^2 + c$, with $c \simeq -0.7645 - i \cdot 0.1595$ (right).
(from http://w3.math.uminho.pt/~scosentino/bestiario/julia.html)

# 13 Area and integration

**The problem posed by Newton equation.** If you derive twice a trajectory $t \mapsto q(t) \in \mathbb{R}^3$, you get the velocity $v(t) := \dot{q}(t)$ and then then acceleration $a(t) := \ddot{q}(t)$. Physicists know the acceleration of a particle in an inertial frame, it is proportional to the force, according to Newton equation

$$m\ddot{q} = F(q, \dot{q}, t)\,.$$

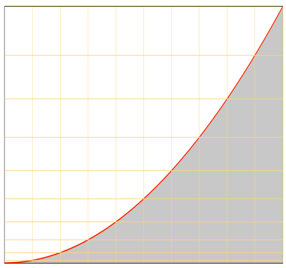
Therefore, they have the problem to deduce the trajectory from its second derivative.

**Work.** The work done by a constant force field $F$ to move a particle from the position $q_0$ to the position $q_1$, hence a distance $\delta q = q_1 - q_0$, is $\mathrm{W} = F \cdot \delta q$. If we make a path through the points $q_0$, $q_1$, ..., $q_n$, with increments $\delta q_k = q_k - q_{k-1}$, and assume that the force is piece-wise constant, we are led to an expression

$$\mathrm{W} = F(q_1) \cdot \delta q_1 + F(q_2) \cdot \delta q_2 + \cdots + F(q_n) \cdot \delta q_n\,.$$

This is a sum of signed areas (i.e. positive or negative depending on the sign of the force) of rectangles with bases $\delta q_k$ and heights $F(q_k)$. If we plot the graph of $F(q)$, this is the signed area of the region bounded by such a graph, the $q$-axis, and the vertical lines $q_0$ and $q_n$. For a generic force $F(q)$, say continuous, it is natural to call work such an area, and pose the problem to compute it.

e.g. **Area of a parabolic segment according to Eudoxo and Arquimedes.** O *método de exaustão*, utilizado por Eudoxo e Arquimedes, para calcular a área de uma figura geométrica consiste em aproximar a região com reuniões de figuras simples, como rectângulos e triângulos. Por exemplo, a área do "segmento parabólico"



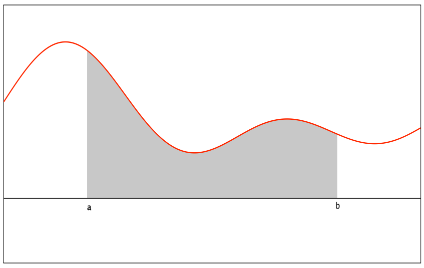
$$A = \{(x, y) \in \mathbb{R}^2 \text{ t.q. } 0 \leqslant x \leqslant 1 \text{ e } 0 \leqslant y \leqslant x^2\},$$

pode ser aproximada dividindo o intervalo $[0, 1]$ em $n$ subintervalos de comprimento $1/n$, e observando que área$(A)$ é superior à soma $s_n(A)$ das áreas dos rectângulos de bases $[\frac{k}{n}, \frac{k+1}{n}]$ e alturas $(k/n)^2$, e inferior à soma $S_n(A)$ das áreas dos rectângulos de bases $[\frac{k}{n}, \frac{k+1}{n}]$ e alturas $((k+1)/n)^2$, onde $k = 0, 1, 2, ..., n-1$. Ou seja,

$$\sum_{k=0}^{n-1} \frac{k^2}{n^3} = s_n(A) \leqslant \text{área}(A) \leqslant S_n(A) = \sum_{k=1}^{n} \frac{k^2}{n^3}$$

- Mostre que a diferença $S_n(A) - s_n(A) \to 0$ quando $n \to \infty$.
- Use a identidade
$$1^2 + 2^2 + 3^2 + ... + n^2 = \frac{n^3}{3} + \frac{n^2}{2} + \frac{n}{6}$$
para mostrar que, quando $n \to \infty$, as aproximações $s_n(A)$ e $S_n(A)$ convergem para 1/3.

**Riemann integral.** Let $f : [a, b] \to \mathbb{R}$ be a bounded function. We want to define the signed area between the graph of $y = f(x)$, the $x$-axis (i.e. $y = 0$), and the vertical lines $x = a$ and $x = b$. For example, if $f(x) \geqslant 0$, this is the area




$$\text{area}\left(\{(x,y) \in \mathbb{R}^2 \ \text{ t.q. } \ a \leqslant x \leqslant b \ \text{ e } \ 0 \leqslant y \leqslant f(x)\}\right) \quad ,$$

The strategy is to approximate the area from below and from above, namely fill and cover the region by unions of rectangles with smaller and smaller bases.

A *partition* of the interval $[a,b]$ is a finite collection $P \subset [a,b]$ of points (that we may order according to their natural order) $a = x_0 < x_1 < x_2 < \cdots < x_k < x_{k+1} < \cdots < x_n = b$ dividing the interval in a finite number (in this case $n$) of subintervals $[x_k, x_{k+1}]$ of lengths $\delta x_k := x_{k+1} - x_k$. Given a partition $P$, we denote by $m_k$ and $M_k$ the minimum and the maximum of $f$ in the subinterval $[x_k, x_{k+1}]$, respectively, hence define the *lower sum* and the *upper sum* of $f$ w.r.t. the partition $P$ as

$$s(f;P) := \sum_{k=0}^{n-1} m_k \cdot \delta x_k \qquad \text{and} \qquad S(f;P) := \sum_{k=0}^{n-1} M_k \cdot \delta x_k \,,$$

respectively. It is clear that the signed area (as the work, if $f$ represents a force) we are trying to compute should be somewhere between $s(f;P) \leqslant$ "area" $\leqslant S(f;P)$. It is also clear that if we "refine" the partition $P$, i.e. if we define a partition $P'$ containing more points than $P$ (hence $P \subset P'$ as subsets of $[a,b]$), then $s(f;P) \leqslant s(f;P')$ and $S(f;P') \leqslant S(f;P)$. In particular, we always have the inequality $s(f;P) \leqslant S(f;Q)$ for all partitions $P$ and $Q$ (just consider the common refinement $P \cup Q$ and use the previous observations).

We say that the bounded function $f$ is *(Riemann) integrable* in the interval $[a,b]$ if there exists a unique number $A$ such that $s(f;P) \leqslant A \leqslant S(f;Q)$ for all partitions $P$ and $Q$. Equivalently, if $\sup_P s(f;P) = A = \inf_P S(f;P)$ (if you know what sup and inf are). Equivalently, if for any precision $\varepsilon > 0$ one may find two partitions $P$ and $Q$ such that $S(f;Q) - s(f;P) < \varepsilon$, hence a partition $R$ (for example $R = P \cup Q$) such that $S(f;R) - s(f;R) < \varepsilon$. If this happens, we call such number "*integral* of $f$ in $[a,b]$", and denote it as

$$A := \int_a^b f(x)\, dx \,.$$

**About the notation.** The notation $\int_a^b f(x)\,dx$ reminds you that the integral should be thought as a sort of limit of the finite sums $\sum_k f(x_k) \cdot \delta x_k$ as the partition gets finer, i.e. as the maximal $|\delta x_k|$ goes to zero. Actually, the notation, to be compared with Leibniz notation $dy/dx$ for derivatives, is useful to state and remind some recipes to compute integrals, as will appear clear in the following.

Also, the variable $x$ inside the integral may be replaced by any other symbol, so that you can also write $\int_a^b f(t)\,dt$, or $\int_a^b f(\clubsuit)\,d\clubsuit$, or $\int_a^b f\,(\text{🦈})\,d\text{🦈}$, ...or whatever you want. The only forbidden symbols are those that you already used somewhere else: so, for example, you should avoid to write $\int_a^b f(b)\,db$.

**Integrability of continuous and monotone functions.** Which functions are Riemann integrable? The final answer is somehow technical. Here, you may be satisfied with knowing that continuous (or also piece-wise continuous) or monotone functions are.

**Theorem 13.1.** *Any continuos function $f(x)$ in a closed and bounded interval $[a,b]$ is integrable.*

*Proof.* Indeed, a continuous function in a closed and bounded (i.e. compact) interval is uniformly continuous. In particular, for any $\varepsilon > 0$ there exists a $\delta > 0$ such that $|x - x'| < \delta$ implies $|f(x) - f(x')| \leqslant \varepsilon/(b-a)$. Consequently, if $P$ is a partition of $[a,b]$ into $n$ subintervals of lengths $|x_{k+1} - x_k| < \delta$ for all $k$, we see that $S(f,P) - s(f,P) \leqslant n \cdot \delta \cdot \varepsilon/(b-a) \leqslant \varepsilon$. $\square$

It is clear that also a function $f : [a,b] \to \mathbb{R}$ with a finite number of discontinuities and finite limits on both sides is integrable (just repeating the argument in any closed subinterval where it is continuous).

**Theorem 13.2.** *Any monotone function $f(x)$ in a closed and bounded interval $[a,b]$ is integrable.*

*Proof.* Assume that $f(x)$ is non-decreasing (otherwise take $-f(x)$), and take any $\varepsilon > 0$. If $P$ is a partition of $[a,b]$ into $n$ subintervals of lengths $|x_{k+1} - x_k| \leqslant \varepsilon/(f(b) - f(a))$, then, since $m_k = f(x_k)$ and $M_k = f(x_{k+1})$, we see that

$$S(f,P) - s(f,P) \leqslant \frac{\varepsilon}{f(b) - f(a)} \sum_{k=0}^{n-1} (f(x_{k+1}) - f(x_k)) \leqslant \varepsilon\,,$$

because the above sum is telescopic and equal to $f(b) - f(a)$. □

**Elementary properties.** The following elementary properties of the integral are obvious for integrals of constant functions, namely for areas of rectangles. But the Riemann integral is defined using rectangles, so it is not surprising that they continue to hold for all integrable functions. You may want to draw pictures to understand their meanings and to convince yourself of their validity.

It is clear that the integral is linear, namely

$$\int_a^b (f(x) + g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx \tag{13.1}$$

and

$$\int_a^b \lambda \cdot f(x)\,dx = \lambda \cdot \int_a^b f(x)\,dx\ . \tag{13.2}$$

for all integrable function $f(x)$ and $g(x)$ and all constants $\lambda \in \mathbb{R}$. It is clear that

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx\,, \tag{13.3}$$

whenever $a < c < b$. If we define

$$\int_a^a f(x)\,dx := 0\,,$$

and

$$\int_b^a f(x)\,dx := -\int_a^b f(x)\,dx\,,$$

then formula (13.3) holds for all $a, b, c$, independently of their order. The integral, being a signed area, behaves well under translations and dilatations of the independent variable, namely:

$$\int_a^b f(x)\,dx = \int_{a+c}^{b+c} f(x-c)\,dx \tag{13.4}$$

for all $c \in \mathbb{R}$, and

$$\int_{\lambda a}^{\lambda b} f(x/\lambda)\,dx = \lambda \int_a^b f(x)\,dx \tag{13.5}$$

for all $\lambda > 0$. The integral is monotone:

$$f(x) \leqslant g(x) \quad \forall x \in [a,b] \qquad \Rightarrow \qquad \int_a^b f(x)\,dx \leqslant \int_a^b g(x)\,dx\,. \tag{13.6}$$

In particular,

$$\int_a^b f(x)\,dx \leqslant \int_a^b |f(x)|\,dx\,. \tag{13.7}$$

Finally, it can be (crudely) estimated from below and from above by the signed areas of two rectangles (the one which is contained and the one which contains the figure we are computing the area of) according to

$$m \leqslant f(x) \leqslant M \quad \forall x \in [a,b] \qquad \Rightarrow \qquad m \cdot (b-a) \leqslant \int_a^b f(x)dx \leqslant M \cdot (b-a)\,. \tag{13.8}$$

An interesting consequence is the

**Theorem 13.3** (Mean value theorem)**.** *If $f : [a,b] \to \mathbb{R}$ is a continuous function, then there is a point $c \in [a,b]$ such that*

$$\int_a^b f(x)\,dx = f(c) \cdot (b-a)\,.$$

Indeed, let $m$ and $M$ be the minimum and maximum of $f(x)$ in the interval $[a,b]$, respectively. From (13.8), we see that there exists some value $m \leqslant d \leqslant M$ such that $\int_a^b f(x)\,dx = d \cdot (b-a)$. By continuity (i.e. Bolzano theorem 9.1), there exists a point $c \in [a,b]$ where $f(c) = d$.

Observe that the value

$$f(c) = \frac{1}{b-a} \cdot \int_a^b f(x)\,dx$$

must be thought as an average of the values $f(x)$ for $a \leqslant x \leqslant b$ (the signed height of a rectangle with base $b-a$ and area $\int_a^b f(x)\,dx$).

**ex:** Compute the following integrals drawing a picture and using the elementary formulas for areas.

$$\int_0^1 3dx \qquad \int_{-2}^2 7\,dx \qquad \int_1^{10} x\,dx \qquad \int_{-2}^3 (-2x)\,dx$$

$$\int_{-2}^2 |x|\,dx \qquad \int_0^3 (5x-2)\,dx \qquad \int_{-33}^{33} (11-x)\,dx$$

$$\int_0^{n+1} [x]\,dx\ ^{30} \qquad \int_6^x 7t\,dt \qquad \int_x^{x^2} (1-t)\,dt$$

**Derivative of an integral.** Here is Newton's and Leibniz' discovery:

**Theorem 13.4** (fundamental theorem of calculus)**.** *Let $f(x)$ be a continuous function defined in some interval $I \subset \mathbb{R}$. Given a point $a \in I$, define the function $F(x)$ as the integral*

$$F(x) := \int_a^x f(t)\,dt\,,$$

*for $x \in I$. Then $F(x)$ is differentiable, and its derivative is $F'(x) = f(x)$, i.e*

$$\frac{d}{dx}\left(\int_a^x f(t)\,dt\right) = f(x)\,.$$

*Proof.* Indeed, the difference $F(x+\delta) - F(x)$ is equal to the integral $\int_x^{x+\delta} f(t)\,dt$. Therefore,

$$\frac{F(x+\delta)-F(x)}{\delta} - f(x) = \frac{1}{\delta}\int_x^{x+\delta} (f(t) - f(x))\,dt\,. \tag{13.9}$$

If $f$ is continuous at the point $x$ (just at the point $x$!), for any $\varepsilon > 0$ there exists $\delta > 0$ such that $|t-x| < \delta$ implies $|f(t) - f(x)| < \varepsilon$. Therefore, for such small $\delta > 0$, the r.h.s. of (13.9) above is bounded by $\frac{1}{\delta} \cdot \varepsilon \cdot \delta = \varepsilon$. Consequently, $(F(x+\delta) - F(x))/\delta \to f(x)$ when $\delta \to 0$. □

[30] $[x]$ denotes the "integer part of $x$", i.e. the unique integer $n \in \mathbb{Z}$ such that $n \leqslant x < n+1$.

**Logarithm and exponential.** The *logarithm* is the function $\log : \mathbb{R}_+ \to \mathbb{R}$ defined by the integral

$$\log(x) := \int_1^x \frac{dt}{t}\,.$$

By the fundamental theorem of calculus 13.4, the derivative is

$$\log'(x) = \frac{1}{x}$$

so that the logarithm is strictly increasing. It is clear that $\log(1) = 0$. Moreover, for any $x, y > 0$

$$\int_1^{xy} \frac{dt}{t} = \int_1^x \frac{dt}{t} + \int_x^{xy} \frac{dt}{t} = \int_1^x \frac{dt}{t} + \int_1^y \frac{ds}{s}\,,$$

(using (13.5) in the second integral above), therefore

$$\log(xy) = \log(x) + \log(y)\,. \tag{13.10}$$

Also, for any $x > 0$

$$\int_1^{1/x} \frac{dt}{t} = \int_x^1 \frac{ds}{s} = -\int_1^x \frac{ds}{s}$$

(using (13.5)), therefore

$$\log(1/x) = -\log(x)\,. \tag{13.11}$$

In particular, $\log(x) \to \infty$ when $x \to \infty$, and $\log(x) \to -\infty$ when $x \to 0$. Thus, $\log(\mathbb{R}_+) = \mathbb{R}$.

The *exponential* is the inverse function of the logarithm, the function $\exp : \mathbb{R} \to \mathbb{R}_+$ such that $\exp(\log x) = x$ for all $x \in \mathbb{R}_+$ and $\log(\exp(y)) = y$ for all $y \in \mathbb{R}$. In particular, $\exp(0) = 1$. The value $\exp(x)$ is also denoted by $e^x$, where $e = \exp(1)$, hence $\log(e) = 1$. The derivative of the exponential is

$$\exp'(x) = \exp(x)\,.$$

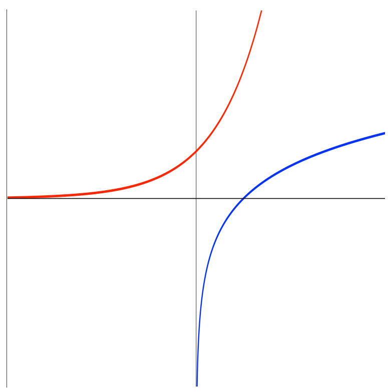
From (13.10) we get

$$\exp(x + y) = \exp(x)\exp(y) \tag{13.12}$$

for all $x, y \in \mathbb{R}$, and from (13.11) we get

$$\exp(-x) = 1/\exp(x) \tag{13.13}$$

for all $x \in \mathbb{R}$.



Graphs of the logarithm (blue) and the exponential (red).

**Primitives and integration.** A differentiable function $F(x)$ is called a *primitive* ("a" primitive, and not "the" primitive!) of the continuous function $f(x)$ if

$$F'(x) = f(x)$$

for any $x$ in a common interval of definition. Leibniz' notation for a primitive of $f(x)$ is $\int f(x)\,dx$. Thus, for example, $\log(x)$ is a primitive of $1/x$, and $e^x$ is a primitive of $e^x$ itself.

We already know a primitive of a continuous function $f(x)$, this is $G(x) = \int_a^x f(t)\,dt$, according to theorem 13.4. If $F(x)$ is any other primitive of the same function $f(x)$, then the difference $F(x) - G(x)$ has zero derivatives, hence is constant by the mean value theorem. Therefore, there is some constant $c \in \mathbb{R}$ such that $F(x) = G(x) + c$ for any $x$ (in some domain). In particular, if we take $x = a$ and then $x = b$, we see that $F(a) = c$ and $F(b) = G(b) + c$, so that $G(b) = F(b) - F(a)$. Therefore, we may state the following recipe to compute integrals, known as *Barrow formula*:

**Theorem 13.5** (Barrow)**.** *If $F(x)$ is any primitive of $f(x)$, then*

$$\int_a^b f(x)\,dx = [F(x)]_a^b := F(b) - F(a)\,. \tag{13.14}$$

Thus, to any derivative that you know there corresponds an integral that you can compute.

**ex:** Compute the following primitives.

$$\int dx \qquad \int x^2 dx \qquad \int \frac{1}{x^3} dx \qquad \int \sqrt{2x-1} dx$$

$$\int (x^2 - 2x + 5) dx \qquad \int \sin(\theta) d\theta \qquad \int \left(\cos(\pi x) - 2x^3\right) dx \qquad \int \frac{dx}{\sqrt{x}}$$

$$\int \frac{d\theta}{\cos^2(\theta)} \qquad \int \frac{dx}{1+x^2} \qquad \int \frac{dx}{\sqrt{1-x^2}}$$

**ex:** Compute the following integrals.

$$\int_0^3 (x-1) dx \qquad \int_{-1}^1 (1 - |x|) dx \qquad \int_0^{10} \sqrt{x} dx \qquad \int_{-\pi}^{\pi} \cos(x) dx$$

$$\int_{-3}^2 \sqrt{x^2} dx \qquad \int_{-\pi}^{\pi/2} \sin(2x) dx \qquad \int_1^2 \frac{1}{x^2} dx \qquad \int_3^5 \left(x^{1/3} - x^{1/5}\right) dx$$

$$\int_{-5}^5 \left(1 + 399x - x^2\right) dx \qquad \int_0^{2\pi} |\sin(x)| dx \qquad \int_{-1}^1 (33 - 11x)^{66}\, dx$$

$$\int_2^3 \frac{dx}{x} \qquad \int_{\log 1}^{\log 2} e^x dx \qquad \int \frac{dx}{x-1} \qquad \int_1^2 e^{x-1} dx$$

$$\int 2e^{3x} dx \qquad \int_0^7 e^{-x} dx \qquad \int \frac{1}{x(1-x)} dx$$

**ex:** Compute the derivative of

$$F(x) = \int_0^x \frac{dt}{1+t^2} \qquad F(x) = \int_0^{x^2} \sin(t) dt \qquad F(x) = \int_{2x}^{x^3} (t - t^2) dt$$

**ex:** Compute the area of the planar region bounded by the curves

$$y = x^2 \quad \text{e} \quad y = x^3\,, \qquad \text{com } 0 \leqslant x \leqslant 1$$

$$y = \sin(x) \quad \text{e} \quad y = -\sin(x)\,, \qquad \text{com } 0 \leqslant x \leqslant \pi$$

$$y = x^{1/3} \quad \text{e} \quad y = x^{1/2}\,, \qquad \text{com } 0 \leqslant x \leqslant 1$$

**e.g. Potential energy and work.** Let $f(x)$ be a continuous force field, defined in an interval of the real line. Any function $V(x)$ such that $V'(x) = -f(x)$ (i.e. minus a primitive of $f(x)$) is called *potential energy*. Theorem 13.5 says that the work done when displacing a particle from $a$ to $b$ is

$$W(a \to b) = \int_a^b f(x)\,dx = V(a) - V(b)\,.$$

thus equal to the difference between the potential energies of the initial and the final points. Observe that in dimension one all (continuous) forces are conservative! Indeed, any $-\int_a^x f(t)\,dt$ is a potential.

**Substitutions.** Let $F(x)$ be a primitive of the continuous function $f: I \to \mathbb{R}$. If $g: I \to \mathbb{R}$ is a continuous function, then by (13.14)

$$\int_{g(a)}^{g(b)} f(y)\, dy = F(g(b)) - F(g(a))\,.$$

If, moreover, $g$ is differentiable, by the chain rule $F(g(x))$ is a primitive of $f(g(x))\,g'(x)$, hence by (13.14) again, the r.h.s. above is also equal to

$$F(g(b)) - F(g(a)) = \int_a^b f(g(x))\, g'(x)\, dx\,.$$

Therefore, we may state the following recipe to transform one integral into another, hopefully simpler, integral: the "substitution" $y = g(x)$, with $dy = g'(x)dx$ (this means $dy/dx = g'(x)$), transforms $f(g(x))\, g'(x)\, dx$ into $f(y)\, dy$, so that

$$\int_a^b f(g(x))\, g'(x)\, dx = \int_{g(a)}^{g(b)} f(y)\, dy \tag{13.15}$$

**e.g.** For example, the substitution $y = x^2$, with $dy = 2xdx$, sends

$$\int_a^b xe^{x^2}\, dx = \frac{1}{2}\int_{a^2}^{b^2} e^y\, dy = \frac{e^{b^2} - e^{a^2}}{2}$$

**ex:** Compute

$$\int_0^1 xe^{x^2} dx \qquad \int \frac{\cos(\log x)}{x} dx \qquad \int \frac{\cos(\theta)}{\sqrt{5 + 2\sin(\theta)}} d\theta \qquad \int \tan(\theta) d\theta$$

$$\int 3x^2\cos(x^3) dx \qquad \int_\pi^{2\pi} \frac{\sin(\sqrt{x})}{\sqrt{x}} dx \qquad \int_{-1/2}^{1/2} \frac{x}{\sqrt{1-x^2}} dx$$

$$\int \cos(x) e^{\sin(x)} dx \qquad \int \frac{x}{x^2-1} dx \qquad \int_{\log 1}^{\log 2} \frac{e^x}{\sqrt{1+e^x}} dx \qquad \int \frac{\sin(\sqrt{x})}{\sqrt{x}} dx$$

**Integration by parts.** Let $f$ and $g$ be two differentiable functions. The derivative of the product $fg$ is $(fg)' = f'g + fg'$. Therefore, by (13.14),

$$\int_a^b f(x)g'(x)dx = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)dx\,. \tag{13.16}$$

This is useful when the integral on the left seems difficult, but the one on the right is simple.

**e.g.** For example, from $(x\sin x)' = \sin x + x\cos x$, we get

$$\int x\cos(x)\, dx = x\sin x - \int \sin(x)\, dx = x\sin(x) + \cos(x)\,.$$

**ex:** Calcule

$$\int_0^1 xe^{-x}\, dx \qquad \int \sin(\log(x))\, dx \qquad \int_1^{e^3} \log(x)\, dx$$

$$\int x\sin(x)\, dx \qquad \int x^2\sin(x)\, dx \qquad \int e^x\sin(x)\, dx$$

**ex:** Compute

$$\int_0^1 x e^{x^2} dx \qquad \int \frac{\cos(\log x)}{x} dx \qquad \int \frac{\cos(\theta)}{\sqrt{5+2\sin(\theta)}} d\theta \qquad \int \tan(\theta) d\theta$$

$$\int \cos(x) e^{\sin(x)} dx \qquad \int \frac{x}{x^2-1} dx \qquad \int_{\log 1}^{\log 2} \frac{e^x}{\sqrt{1+e^x}} dx$$

**help:** `Mathematica`®`8` computes primitives as

```
Integrate[1/ Sin[2 x], x]
-(1/2) Log[Cos[x]] + 1/2 Log[Sin[x]]
```

or (definite) integrals as

```
Integrate[x^2 - Sin[x], {x, 0, 3}]
8 + Cos[3]
```

**The problem of computing integrals.** Unlike computing derivatives, computing integrals is not a mechanical process. It rather resembles the job of Mr. Poirot, consisting in finding and analyzing clues that will eventually show (one of) the right path to the solution. It may also happens that no known method or primitive comes into ones mind, and indeed that the primitive of a given function cannot be expressed in terms of presently known functions. When this happens, the solution for physicists and engineers is to approximate the integral with some computational device (and many methods and techniques are available). Then, we may also give a name to the primitive (as we have done for the primitive of $1/x$, the "logarithm"), if it is recurrent and we judge it useful.

**e.g. Velocity/acceleration + initial conditions ⇒ time law)** If we know the velocity $v(t) = \dot{q}(t)$ of a particle (moving in one dimension) and its initial condition $q(0)$, we may find its trajectory as

$$q(t) = q(0) + \int_0^t v(s)\, ds\,.$$

If we know the acceleration $a(t) = \dot{v}(t)$ and the initial velocity $v(0)$, we may integrate once to get

$$v(t) = v(0) + \int_0^t a(s)\, ds\,,$$

and then integrate once again to get the trajectory $q(t)$ as above.

**e.g. Work of a perfect gas.** The work done by a perfect gas expanding from an initial volume $V_0$ to a final volume $V_1$ is given by the integral of the pressure $p(V)$

$$W = \int_{V_0}^{V_1} p(V)\, dV\,.$$

If the pressure is mantained constant, this is simply $W = p \cdot (V_1 - V_0)$. If the expansion occurs at constant temperature $T$, we get from the equation of state $pV = nRT$ (here $n$ is the number of moles, $R \simeq 8.314 \times 10^7$ J/K mol, and $T$ the absolute temperature)

$$W = nRT \int_{V_0}^{V_1} \frac{dV}{V} = nRT \log(V_1/V_0)\,.$$

**Improper integrals.** It is useful to integrate a function in an infinite domain like $[a, \infty)$, the definition being

$$\int_a^\infty f(x)\,dx := \lim_{K\to\infty} \int_a^K f(x)\,dx\,,$$

or in a domain $(a, b]$ bounded by a point $a$ where the function is not defined, the definition being

$$\int_{a^+}^b f(x)\,dx := \lim_{\varepsilon\searrow 0} \int_{a+\varepsilon}^b f(x)\,dx\,.$$

These limits are called *improper integrals*.

e.g. **Gaussian and error function.** An important function in many areas of mathematics and applied sciences is the *Gaussian* $g(x) := e^{-x^2}$ (and its variations, obtained by a linear change of coordinates, or multidimensional ones). One knows that the improper integral

$$\int_{-\infty}^\infty e^{-x^2}\,dx = \sqrt{\pi}$$

(we'll compute it later), thus $\frac{1}{\sqrt{\pi}}e^{x^2}$ is a probability distribution, indeed a most fundamental one. A primitive of the Gaussian cannot be computed in terms of elementary functions (polynomials, trigonometric, exponential, ...), hence deserves a name. It is called *error function*, and usually normalized according to

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{t^2}\,dt\,.$$

Thus, $\operatorname{erf}(x) \to \pm 1$ as $x \to \pm\infty$.

## Primitives

| | (function) | (“a” primitive) |
|---|---|---|
| | $f(x) = F'(x)$ | $\int f(x)\,dx = F(x)$ |
| (por substituição) | $f(y(x))\,y'(x)$ | $\int f(y(x))y'(x)\,dx = \int f(y)\,dy$ |
| (por partes) | $f(x)g'(x)$ | $\int f(x)g'(x)\,dx = f(x)g(x) - \int f'(x)g(x)\,dx$ |
| (constantes) | $\lambda$ | $\int \lambda\,dx = \lambda x$ |
| (potências, $\alpha \neq -1$) | $x^\alpha$ | $\int x^\alpha\,dx = \frac{1}{\alpha+1}x^{\alpha+1}$ |
| (logaritmo) | $1/x$ | $\int \frac{dx}{x} = \log\lvert x\rvert$ |
| (exponencial) | $e^x$ | $\int e^x\,dx = e^x$ |
| (seno) | $\sin(x)$ | $\int \sin(x)\,dx = -\cos(x)$ |
| (coseno) | $\cos(x)$ | $\int \cos(x)\,dx = \sin(x)$ |
| (tangente) | $\frac{1}{\cos^2(x)}$ | $\int \frac{dx}{\cos^2(x)} = \tan(x)$ |
| (cotangente) | $\frac{1}{\sin^2(x)}$ | $\int \frac{dx}{\sin^2(x)} = -\text{cotan}(x)$ |
| (arco cujo seno) | $\frac{1}{\sqrt{1-x^2}}$ | $\int \frac{dx}{\sqrt{1-x^2}} = \arcsin(x)$ |
| (arco cuja tangente) | $\frac{1}{1+x^2}$ | $\int \frac{dx}{1+x^2} = \arctan(x)$ |
| (exponencial × seno) | $e^{\alpha x}\sin(\beta x)$ | $\int e^{\alpha x}\sin(\beta x)\,dx = \frac{e^{\alpha x}(\alpha\sin(\beta x)-\beta\cos(\beta x))}{\alpha^2+\beta^2}$ |
| (exponencial × coseno) | $e^{\alpha x}\cos(\beta x)$ | $\int e^{\alpha x}\cos(\beta x)\,dx = \frac{e^{\alpha x}(\alpha\cos(\beta x)+\beta\sin(\beta x))}{\alpha^2+\beta^2}$ |
| (coseno × coseno, $n^2 \neq m^2$) | $\cos(nx)\cos(mx)$ | $\int \cos(nx)\cos(mx)\,dx = \frac{\sin((n+m)x)}{2(n+m)} + \frac{\sin((n-m)x)}{2(n-m)}$ |
| (seno × seno, $n^2 \neq m^2$) | $\sin(nx)\sin(mx)$ | $\int \sin(nx)\sin(mx)\,dx = -\frac{\sin((n+m)x)}{2(n+m)} - \frac{\sin((n-m)x)}{2(n-m)}$ |
| (seno × coseno, $n^2 \neq m^2$) | $\sin(nx)\cos(mx)$ | $\int \sin(nx)\cos(mx)\,dx = -\frac{\cos((n+m)x)}{2(n+m)} - \frac{\cos((n-m)x)}{2(n-m)}$ |
| ($x$ × coseno, $n \neq 0$) | $x\cos(nx)$ | $\int x\cos(nx)\,dx = \frac{\cos(nx)}{n^2} + \frac{x\sin(nx)}{n}$ |
| ($x$ × seno, $n \neq 0$) | $x\sin(nx)$ | $\int x\sin(nx)\,dx = \frac{\sin(nx)}{n^2} - \frac{x\cos(nx)}{n}$ |
| ($x^k$ × coseno, $n \neq 0$) | $x^k\cos(nx)$ | $\int x^k\cos(nx)\,dx = \frac{x^k\sin(nx)}{n} - \frac{k}{n}\int x^{k-1}\sin(nx)dx$ |
| ($x^k$ × seno, $n \neq 0$) | $x^k\sin(nx)$ | $\int x^k\sin(nx)\,dx = -\frac{x^k\cos(nx)}{n} + \frac{k}{n}\int x^{k-1}\cos(nx)dx$ |

# 14 Ordinary differential equations

e.g. **Free particle.** The trajectory $t \mapsto q(t) \in \mathbb{R}^3$ of a free particle of mass $m$ in an inertial frame is modeled by the Newton equation

$$\frac{d}{dt}(mv) = 0\,, \qquad \text{i.e., if } m \text{ is constant,} \qquad ma = 0\,,$$

where $v(t) := \dot{q}(t)$ denotes the *velocity* and $a(t) := \ddot{q}(t)$ denotes the *acceleration* of the particle. In particular, the *linear momentum* $p := mv$ is a constant of the motion (i.e. $\frac{d}{dt}p = 0$), in accordance with Galileo's principle of inertia or Newton's first law[31].

The solutions of Newton equation are the affine lines

$$q(t) = s + vt\,,$$

where $s, v \in \mathbb{R}^3$ are arbitrary vectors, the initial position and the initial velocity.

Thus, for example, the trajectory of a free particle starting at $q(0) = (3,2,1)$ with velocity $\dot{q}(0) = (1,2,3)$ is $q(t) = (3,2,1) + (1,2,3)t$.

e.g. **Free fall near the Earth surface.** The Newton equation

$$m\ddot{z} \simeq -G\frac{mM_\oplus}{R_\oplus^2}$$

models the free fall of a particle of mass $m$ near the Earth surface. Here $z(t)$ is the height of the particle at time $t$ (measured from some reference height, e.g. the sea level), $G \simeq 6.67$ m$^3$ kg$^{-1}$ s$^{-2}$ is the gravitational constant, $M_\oplus$ and $R_\oplus$ are the mass and radius of the Earth, respectively. We are assuming that $z \ll R_\oplus$. Since inertial and gravitational masses are (experimentally) equal, the mass $m$ cancels out and we get the equation

$$\ddot{z} = g\,,$$

where $g := GM_\oplus/R_\oplus^2 \simeq 9.8$ m s$^{-2}$ is the the gravitational acceleration near the Earth surface, independent on the falling object!

A function with constant second derivative equal to $-g$ is $-gt^2/2$. But it is not the unique solution. We may add to it any function with zero second derivative, that is any constant $s$ and any linear function $vt$. This means that also any

$$z(t) = s + vt - \frac{1}{2}gt^2$$

is a solution of our Newton equation, for any $s$ and any $v$. The first arbitrary constant $s$ is the initial height $z(0)$ (and this physically corresponds to the homogeneity of space: Newtonian physics is independent on the place where the laboratory is placed). The second arbitrary constant $v$ is the initial velocity $\dot{z}(0)$ (and this physically corresponds to Galilean invariance: we cannot distinguish between two inertial laboratories moving at constant speed one from each other).

The moral is that the Newton equation alone does not have a "unique" solution. It has a whole "family of solutions", depending on two parameters $s$ and $v$. On the other side, once we fix the initial position $z(0)$ and the initial velocity $\dot{z}(0)$, the solution turns out to be unique (we'll prove it soon! meanwhile, you may try to prove that the difference of any two solutions with the same initial conditions is constant and equal to zero). In other words, once known the initial "state" of the particle, i.e. its position and its velocity, the Newton equation uniquely determines the "future" and "past" history of the particle.

e.g. **A differential equation for the exponential function.** Consider the first order ODE

$$\dot{x} = x$$

[31] "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.]

where $\dot{x}$ denotes the derivative of $x(t)$ w.r.t. the real variable $t$.

An obvious solution is $x(t) = 0$. Besides, computation shows that the exponential function $e^t$ satisfies the equation. Indeed, the (natural) exponential is defined by the power series

$$\exp(t) := \sum_{n \geqslant 0} \frac{t^n}{n!} = 1 + t + \frac{t^2}{2} + \frac{t^3}{6} + \dots$$

(remember that $0! = 1$), which converges uniformly in any bounded interval. You may check, deriving the power series term by term, that $\exp' = \exp$.

But we can multiply it by any constant $b$ and still get a solution, hence any function $x(t) = be^t$ satisfies the above identity. If we set $t = 0$, we notice that $b$ is the value of $x(0)$.

We claim that $x(t) = x_0 e^t$ is the "unique" solution of the differential equation $\dot{x} = x$ with initial data $x(0) = x_0$. Indeed, let $y(t)$ be any other solution. Since the exponential is never zero, we can divide by $e^t$ and define the function $h(t) = y(t)e^{-t}$. Deriving we get

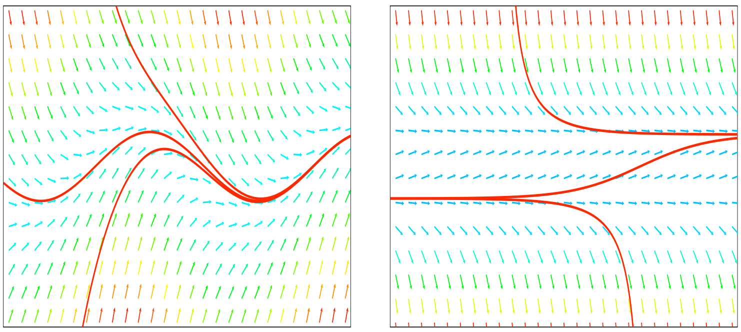
$$\dot{h} = (\dot{y} - y)\, e^{-t} \,.$$

But $y$ solves the equation, hence the first derivative of $h$ is everywhere zero. By the mean value theorem $h$ is a constant function, and, since $y(0) = x_0$ too, its value at the origin is $h(0) = y(0)e^{-0} = x_0$. This implies that $y(t)$ is indeed equal to $x(t)$.

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

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

para a trajectória $t \mapsto x(t) \in \mathbb{R}$ de um sistema dinâmico, onde $\dot{x} = \frac{dx}{dt}$ denota a derivada do observável $x$ em ordem ao tempo $t$, e $v(t, x)$ é um *campo de direções* dado (ou seja, uma recta com declive $v(t, x)$ para cada ponto $(t, x)$). Uma *solução* da EDO é uma função $t \mapsto x(t)$ tal que $\dot{x}(t) = v(t, x(t))$ para cada tempo $t$ num certo intervalo, ou seja, uma função cujo gráfico é tangente ao campo de direções em cada ponto $(t, x(t))$ do gráfico. Se o campo $v(t, x)$ é suficientemente regular (por exemplo, diferenciável), para cada ponto $(t_0, x_0)$ passa uma única solução com condição inicial $x(t_0) = x_0$.



Slope fields and some solutions of $\dot{x} = \sin(t) - x$ and of $\dot{x} = x(1 - x)$.

**ex:** A função $x(t) = t^3$ é solução da equação diferencial $t\dot{x} - 3x = 0$ ? E a função $x(t) = 0$ ?

**Simple ODEs.** The simplest case occurs when the velocity field $v$ does not depend on the phase space variable $x$, so the equation is

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

where $v(t)$ is some given function of time. This just says that $x$ is a primitive of $v$, and the fundamental theorem of calculus (i.e. Leibniz and/or Newton's discovery) tells us how to compute such a primitive: just integrate the function $v$ from some initial time $t_0$ up to a final time $t$. Indeed, provided $v$ is a continuous function, the derivative of $\int_{t_0}^{t} v(s)ds$ at the point $t$ is $v(t)$. This explains the current use of the expression "integrate" a differential equation instead of "solving" a differential equation, as well as the meaning of Newton's quoted anagram.

Primitives are not unique, but are defined modulo an additive constant. This arbitrary constant can be matched with the initial condition, so that the solution of $\dot{x} = v(t)$ with initial condition $x(t_0) = x_0$ is

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

Here you may observe that this class of ODEs have "symmetries". The line field does not depend on $x$, hence slopes of solutions are the same along horizontal lines ($t$ = constant) in the extended phase space. There follows that any translate $\varphi(t) + c$ of a solution $\varphi(t)$ is still a solution. This is but a geometrical interpretation of the arbitrary constant in the primitive of $v$.

ex: **Newtonian motion in a time dependent force field.** The one-dimensional motion of a particle of mass $m$ subject to a time-dependent force $F(t)$ is modeled by the Newton equation

$$m\ddot{x} = F(t) \,.$$

Call $v = \dot{x}$ the velocity of the particle, and derive the first order ODE satisfied by the velocity $v$. Solve the equation for the velocity, given a force $F(t) = F_0 \sin(\gamma t)$ and an initial condition $v(0) = v_0$. Use the above solution $v(t)$ to find the trajectory $x(t)$ of the particle, given an initial position $x(0) = x_0$.

ex: **Rockets.** Se um foguetão (no espaço vazio, sem forças gravitacionais!) expulsa combustível a uma velocidade relativa constante $-V$ e a uma taxa constante $\dot{m} = -\alpha$, então a sua trajectória num referencial inercial (uni-dimensional) é modelada pela equação de Newton

$$\frac{d}{dt}(mv) = \alpha(V - v) \,, \qquad \text{ou seja} \;, \qquad \dot{m}v + m\dot{v} = \alpha(V - v) \,.$$

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

$$\dot{v} = \frac{\alpha V}{m_0 - \alpha t}$$

(valida se $0 \leqslant t < m_0/\alpha$). Calcule a trajectória do foguetão com velocidade inicial $v(0) = v_0$ e posição inicial $q(0) = 0$, válida para tempos $t$ inferiores ao tempo necessário para acabar o combustível.

**Autonomous ODEs.** A first order ODE of the form

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

where the velocity field $v$ does not depend on time, is called *autonomous*. Most fundamental equations of physics (those describing closed systems, without external forces) can be written as autonomous first order ODEs, and this corresponds to time-invariance of physical laws.

Here you may notice symmetries again. The line field $v$ of an autonomous equation is constant along vertical lines ($x$ = constant) of the extended phase space. Hence any translate $\varphi(t + s)$ of a solution $\varphi(t)$ is still a solution. This is the manifestation of time-invariance of a law codified by an autonomous ODE. This also implies that there is no loss of generality in restricting to initial value problems with initial time $t_0 = 0$.

**Equilibrium solutions.** First, we observe that an autonomous equation may admit constant solutions. Indeed, if $x_0$ is a *singular point* of the vector field $v$, i.e. a point where $v(x_0) = 0$, then the constant function

$$x(t) = x_0$$

obviously solves the equation. Such solutions, which do not change with time, are called *equilibrium*, or *stationary*, solutions.

**Solutions near non-singular points.** Let $x_0$ be a *non-singular point* of the velocity field $v(x)$, i.e. a point $x_0$ where $v(x_0) \neq 0$. We want to solve $\dot{x} = v(x)$ with initial condition $x(t_0) = x_0$. First, rewrite the equation $dx/dt = v(x)$ formally as "$dx/v(x) = dt$" (multiply by $dt$ and divide by $v(x)$, so that all $x$'s are on the left and all $t$'s are on the right). Instead of trying to make sense to this last expression (which is possible, of course, and here you can appreciate the beauty of Leibniz' notation $dx/dt$ for derivatives!), observe that it is suggesting that $\int dx/v(x) = \int dt$. Now assume that the velocity field $v$ is continuous and let $J = (x_-, x_+)$ be the maximal interval containing $x_0$ where $v$ is different from zero. Integrating, from $x_0$ to $x \in J$ on the left and from $t_0$ to $t$ on the right, we obtain a differentiable function $x \mapsto t(x)$ defined as

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

for any $x \in J$. Now, observe that the derivative $dt/dx$ is equal to $1/v$. Since, by continuity, $1/v$ does not change its sign in $J$, our $t(x)$ is a strictly monotone continuously differentiable function. We can invoke the inverse function theorem and conclude that the function $t(x)$ is invertible. This prove that the above relation defines actually a continuously differentiable function $t \mapsto x(t)$ in some interval $I = t(J)$ of times around $t_0$. Finally, you may want to check that the function $t \mapsto x(t)$ solves the Cauchy problem: just compute the derivative (using the inverse function theorem),

$$\begin{aligned} \dot{x}(t) &= 1/\left(\frac{dt}{dx}(x(t))\right) \\ &= v(x)\,, \end{aligned}$$

and check the initial condition. Observe that the function $t(x) - t_0$ has then the interpretation of the "time needed to go from $x_0$ to $x$".

At the end of the story, if you are lucky enough and know how to invert the function $t(x)$, you'll get an explicit solution as

$$x(t) = F^{-1}\left(t - t_0 + F(x_0)\right)\,,$$

where $F$ is any primitive of $1/v$. Close inspection of the above reasoning shows that the local solution you've found is indeed the unique one. Namely, we have the following

**Proposition 14.1. (Existence and uniqueness theorem for autonomous ODEs near a non-singular point)** *Let $v(x)$ be a continuous velocity field and let $x_0$ be a point where $v(x_0) \neq 0$. Then there exist one and only one solution of $\dot{x} = v(x)$ with initial condition $x(t_0) = x_0$ in some sufficiently small interval $I$ around $t_0$. Moreover, the solution $x(t)$ is given implicitly by*

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

*defined in some small interval $J$ around $x_0$.*

**On the failure of uniqueness near singular points.** The interval $I = t(J)$ where the solution is defined need not be the entire real line: solutions may reach the boundary of $J$, i.e. one of the singular points $x_\pm$ of the velocity field, in finite time. Since singular points are themselves equilibrium solutions, this imply that solutions of the initial value problem at singular points may not be unique, under such mild conditions (continuity) for the velocity field. Picard's theorem prescribes stronger regularity conditions on $v$ under which the initial value problem admits unique solutions for any initial condition in the extended phase space.

e.g. **Two solutions with the same initial condition!** Both the curves $x(t) = 0$ and $x(t) = t^3$ solve the equation

$$\dot{x} = 3x^{2/3}$$

with initial condition $x(0) = 0$. The problem here is that the velocity field $v(x) = 3x^{2/3}$, although continuous, is not differentiable and not even Lipschitz at the origin. You may notice that the solution starting, for example, at $x_0 = 1$ reaches (or better comes from) the singular point $x_- = 0$ in finite time, since

$$\begin{aligned} t(x_-) - t(x_0) &= \int_1^0 \frac{1}{3} y^{-2/3} dy \\ &= -1 \,. \end{aligned}$$

**help:** O Mathematica® pode resolver analiticamente equações diferenciais. Por exemplo,

```
DSolve[x'[t] + 2 x[t] == Sin[t], x[t], t]
{{x[t] -> E^(-2 t) C[1] + 1/5 (-Cos[t] + 2 Sin[t])}}
```

e.g. **Radioactive decay.** Radioactive matter (such as $^{14}C$ or $^{238}U$) decay according to the law

$$\dot{N} = -\beta N$$

where $N(t)$ denotes the number of nuclei (assumed large so that the law of large number applies), and $1/\beta$ is the *mean life*, the average time life of one single nucleus. The solution with initial condition $N(0) = N_0 > 0$ is

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

In particular, the initial quantity is reduced to one half after a time $T = (\log 2)/\beta$, called *half-life*. For example, $^{14}C$ has an average time of $1/\beta \simeq 8033$ years.

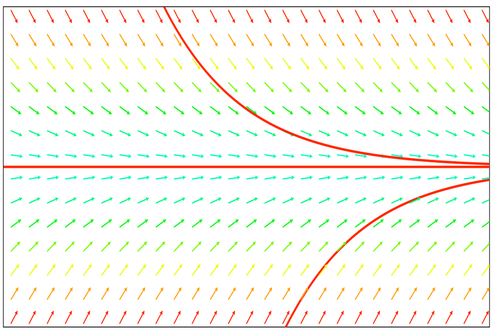
If cosmic radiation produces $^{14}C$ in Earth's atmosphere at a rate $\alpha$, then the quantity of $^{14}C$ in the atmosphere follows a law

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

The equilibrium is $\overline{N} = \alpha/\beta$. The difference $x(t) = N(t) - \overline{N}$ follows the law $\dot{x} = -\beta x$, hence $x(t) = x(0)e^{-\beta t}$, and therefore

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

In particular, $N(t) \to \overline{N}$ as $t \to \infty$, independently from the initial condition $N(0)$.



Direction field and some solutions of $\dot{x} = -2x + 1$.

e.g. **Exponential growth.** The growth of a population in a (virtually) unlimited medium is modeled by

$$\dot{N} = \lambda N$$

where $N(t)$ denotes the population size at time $t$, and $\lambda > 0$ is some growth rate. The solution is an exponential growth like

$$N(t) = N(0)e^{\lambda t} \,.$$

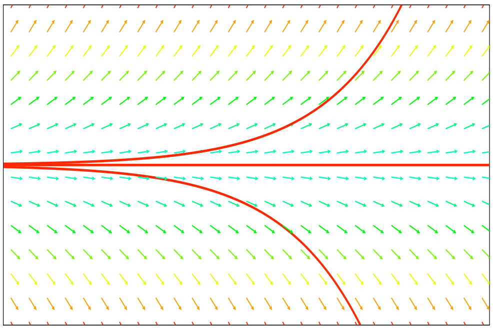
If we retire aportion of the population at constant rate $\alpha > 0$, we get

$$\dot{N} = \lambda N - \alpha$$

Now the stationary solution is $\overline{N} = \alpha/\lambda$, and the general solution

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

Now, the non-constant solutions diverge or disappear.



Direction field and some solutions of $\dot{x} = 2x - 1$.

**e.g. Logistic equation.** A more realistic model of the growth of a population in a limited environment is the *logistic equation*[32]

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

where $\lambda > 0$ and the number $M > 0$ is a maximal population. Observe that $\dot{N} \simeq \lambda N$ when $N \ll M$, and that $\dot{N} \to 0$ when $N \to M$. It is convenient to define the relative population $x(t) := N(t)/M$, which satisfies the adimensional logistic equation

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

Equilibrium solutions are $x = 0$ e $x = 1$. To find solutions with initial condition $x(0) = x_0 \neq 0, 1$, we may integrate

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

using the identity

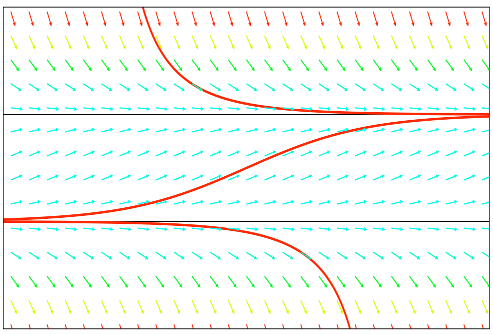
$$\frac{1}{y(1-y)} = \frac{1}{y} + \frac{1}{1-y} \, .$$

The result is

$$\log \left| \frac{x \cdot (x_0 - 1)}{x_0 \cdot (x - 1)} \right| = t$$

which may be solved for $x(t)$, giving, when $0 < x_0 < 1$,

$$x(t) = \frac{1}{1 + \left(\frac{1}{x_0} - 1\right) e^{-\lambda t}} \, .$$



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

Campo de direções e soluções da equação logística.

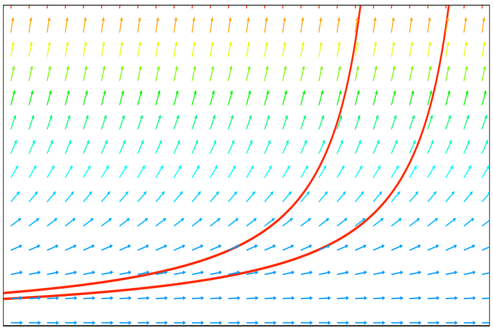
**e.g. Super-exponential growth.** Um outro modelo de dinâmica populacional em meio ilimitado é

$$\dot{N} = \alpha N^2 .$$

onde $\alpha > 0$. A solução com condição inicial $N(0) = N_0 > 0$ é

$$N(t) = \frac{1}{1/N_0 - \alpha t} .$$

Observe que as soluções não estão definidas para todos os tempos: este modelo prevê uma catástrofe (população infinita) após um intervalo de tempo finito (o tempo $\bar{t} = 1/(\alpha N_0)$) !



Campo de direções e soluções da equação $\dot{x} = x^2$.

**Separable ODEs.** A first order ODE $\dot{x} = v(t, x)$ is said *separable* when the velocity field $v$ is a product of a function which only depends on $t$ and another function which only depends on $x$. So it has the form

$$\dot{x} = \frac{f(x)}{g(t)}$$

for some known functions $f$ and $g$. We assume that both $f$ and $g$ are continuous functions on some intervals of the phase space and the real line, respectively, and that $g(t)$ does not vanishes. Observe that both simple ODEs like $\dot{x} = v(t)$ and autonomous ODEs like $\dot{x} = v(x)$ fall in this class.

If $x_0$ is a zero of $f$, then $x(t) = x_0$ is an equilibrium solution. The recipe to find other solutions is known as "separation of variables". Take point $x_0$ where $f(x_0) \neq 0$, and an initial time $t_0$ where $g(t_0) \neq 0$. Choose a maximal interval $J$ containing $x_0$ where $f$ is different from zero, rewrite the equation formally as "$dx/f(x) = dt/g(t)$", and then integrate from $x_0$ to $x \in J$ the r.h.s. and from $t_0$ to $t$ the l.h.s. You'll get

$$\int_{x_0}^{x} \frac{dy}{f(y)} = \int_{t_0}^{t} \frac{ds}{g(s)} .$$

As we did for autonomous equations, we can see that any continuously differentiable solution $t \mapsto x(t)$ of the equation passing through the non-singular point $(t_0, x_0)$ must satisfy the above relation, as long as $x$ is sufficiently near to $x_0$.

**e.g.** Solve $\dot{x} = tx^3$.

An obvious solution is the equilibrium solution $x(t) = 0$. For a positive initial condition $x(t_0) = x_0 > 0$, rewrite the equation as $dx/x^3 = tdt$ and integrate

$$\int_{x_0}^{x} \frac{dy}{y^3} = \int_{t_0}^{t} s ds$$

for $x > 0$. You'll find

$$1/x^2 - 1/x_0^2 = t^2 - t_0^2 ,$$

and, solving for $x$, the solution

$$x(t) = \frac{1}{\sqrt{t_0^2 + 1/x_0^2 - t^2}} \, .$$

defined for times $t$ in the interval $|t| < \sqrt{t_0^2 + 1/x_0^2}$. In the same way you'll find solutions with negative initial condition $x_0 < 0$.

**Linear first order ODEs.** A *first order linear differential equation* is a differential equation which can be written in the "canonical form"

$$\dot{x} + p(t)\, x = q(t) \, , \tag{14.1}$$

where the *coefficients* $p$ and $q$ are (known) functions of the real variable $t$ defined in some interval $I \subset \mathbb{R}$. We assume that both $p$ and $q$ are continuous functions, and we look for solutions $t \mapsto x(t)$ defined on $I$. Eventually we will want to solve the problem with some initial condition $x(t_0) = x_0$.

The equation

$$\dot{y} + p(t) y = 0 \tag{14.2}$$

is said the *homogeneous* equation associated with the general, hence *non-homogeneous*, equation (14.1) above. The word "homogeneous" is due to the fact that any constant multiple $\lambda \cdot y(t)$ of a solution $y$ of (14.2) is again a solution. Also, any linear combination (with real coefficients) $ay_1(t) + by_2(t)$ of solutions $y_1(t)$ and $y_2(t)$ of the homogeneous equation (14.2) is still a solution of the homogeneous equation. This means that the space of solutions of the homogeneous equation is a linear space, actually a one-dimensional vector space $\mathcal{H} \approx \mathbb{R}$, as a consequence of the following proposition 14.2.

Also interesting is that the difference $y(t) = x_1(t) - x_2(t)$ of any two solutions $x_1(t)$ and $x_2(t)$ of the non-homogeneous equation (14.1) is a solution of the associated homogeneous equation (14.2), hence belongs to the linear space $\mathcal{H}$. Therefore, the space of solutions of the non-homogeneous equation (14.1) is an affine space $x + \mathcal{H}$, where $x(t)$ is any (particular) solution of (14.1).

Solutions of the homogeneous equation are obtained separating the variables, and are given by the following

**Proposition 14.2. (Existence and uniqueness theorem for homogeneous first order linear ODEs)** *Let $p$ be a continuous function on some interval of the real line. Then the unique solution of the homogeneous equation $\dot{y} + p(t)y = 0$ with initial condition $y(t_0) = y_0$ is given by*

$$y(t) = y_0 e^{-\int_{t_0}^{t} p(s)ds} \, .$$

*In particular, the space of solutions of the homogeneous equation (14.2) is a real vector space of dimension one.*

Indeed, let $z(t)$ be a second solution of the Cauchy problem above, and define

$$h(t) = z(t) e^{\int_{t_0}^{t} p(s)ds} \, .$$

Its value for $t_0$ is $y_0$. Its derivative is

$$\dot{h}(t) = e^{\int_{t_0}^{t} p(s)ds} \left( \dot{z}(t) + p(t)z(t) \right) \, .$$

Since $z$ is supposed to solve the equation, the derivative of $h$ is equal to zero for any $t$ in the chosen interval, and the mean value theorem says that then $h(t)$ is constant and equal to $y_0$. There follows that $z(t)$ is indeed equal to our solution $y(t)$.

**e.g.** Solve $t\dot{x} - 2x = 0$ for $t \in (0, \infty)$ with initial condition $x(t_0) = x_0$.

If $x_0 = 0$, the solution is the equilibrium solution $x(t) = 0$. If $x_0 > 0$, write the equation as $dx/x = 2dt/t$, integrate

$$\int_{x_0}^{x} dy/y = \int_{t_0}^{t} 2ds/s \, ,$$

for positive $x$, obtain

$$\log x - \log x_0 = \log(t^2) - \log\left(t_0^2\right) ,$$

and solve it for $x$, the solution being

$$x(t) = \left(x_0/t_0^2\right) t^2 .$$

Finally observe that this formula gives the solutions for any initial condition $x_0$.

**Back to the non-homogeneous equation.** To solve the non-homogeneous equation

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

we use the following trick, a first and elementary instance of a much more general method named "variation of parameters" (or, sometimes, with the oxymoron "variation of constants"). We already know that any function proportional to $e^{-\int_a^t p(s)ds}$ solves the homogeneous equation. We look for a solution of the non-homogeneous equation having the form

$$x(t) = \lambda e^{-\int_{t_0}^t p(s)ds} ,$$

but, instead of treating the parameter $\lambda$ as a constant, we allow it to depend on $t$. Putting our guess into the non-homogeneous equation, we get

$$\frac{d}{dt}\left(\lambda(t)e^{-\int_{t_0}^t p(s)ds}\right) + p(t)\lambda(t)e^{-\int_{t_0}^t p(s)ds} = q(t) .$$

Computing the derivative, we get

$$\dot{\lambda}(t)e^{-\int_{t_0}^t p(s)ds} - \cancel{p(t)\lambda(t)e^{-\int_{t_0}^t p(s)ds}} + \cancel{p(t)\lambda(t)e^{-\int_{t_0}^t p(s)ds}} = q(t) ,$$

the two terms containing $p(t)$ do cancel, and we are left with

$$\dot{\lambda}(t)e^{-\int_{t_0}^t p(s)ds} = q(t) .$$

This can be solved for $\dot{\lambda}$ (because exponentials are never zero), and integration gives

$$\lambda(t) = \lambda(t_0) + \int_{t_0}^t e^{\int_{t_0}^s p(u)du} q(s)ds$$

for some constant $\lambda(t_0)$ equal to the value of $x(t_0)$ (this depends on our choice for $y(t)$, such that $y(t_0) = 1$). Finally, we get a solution

$$x(t) = \lambda(t)e^{-\int_{t_0}^t p(s)ds} ,$$

and you may check that it has initial value $x(t_0) = x_0$. Since the difference of any two solutions of the general equation is a solution of the associated homogeneous equation, and since (as follows from the uniqueness theorem above) the only solution of the homogeneous equation with initial condition $x(t_0) = 0$ is the zero solution, we just proved the following

**Proposition 14.3. (Existence and uniqueness theorem for first order linear ODEs)** *Let $p$ and $q$ be continuous functions in some interval $I$. Then the unique solution of the linear differential equation $\dot{x} + p(t)x = q(t)$ with initial condition $x(t_0) = x_0$ for $t_0 \in I$ is given by*

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

**Suggestion.** Perhaps, instead of fixing the unpleasant formula in the above theorem, you could simply remember the strategy used to derive it: find one non-trivial solution $y(t)$ of the associated homogeneous equation (which is separable!), and then make the conjecture $x(t) = \lambda(t)y(t)$ for some other unknown function $\lambda(t)$. You'll get a simple differential equation for $\lambda$, and integration gives you the solution.

**e.g.** Solve $t\dot{x} - 2x = t$ for $t \in (0, \infty)$ with initial condition $x(t_0) = x_0$.

You already know that the solution of the associated homogeneous equation $ty' - 2y = 0$ with initial condition $y(t_0) = 1$ is $y(t) = t^2/t_0^2$. Make the conjecture $x(t) = \lambda(t)t^2/t_0^2$, insert your guess into the non-homogeneous equation, and get

$$\dot{\lambda} = t_0^2/t^2 \,.$$

Integrate and find

$$\lambda(t) - \lambda(t_0) = t_0 - t_0^2/t \,,$$

and, since $\lambda(t_0) = x(t_0)$, finally find the solution

$$x(t) = \frac{x_0 + t_0}{t_0^2} t^2 - t \,.$$

**ex:** 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.

**ex:** Resolva os seguintes problemas nos intervalos indicados:

$$2\dot{x} - 3x = e^{2t} \qquad t \in (-\infty, \infty) \quad \text{com } x(0) = 1$$

$$\dot{x} + x = e^{3t} \qquad t \in (-\infty, \infty) \quad \text{com } x(1) = 2$$

$$t\dot{x} - x = t^3 \qquad t \in (0, \infty) \quad \text{com } x(1) = 3$$

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

$$dr/d\theta + r\tan\theta = \cos\theta \qquad t \in (-\pi/2, \pi/2) \quad \text{com } r(0) = 1$$

**e.g.** **Free fall with friction.** Friction may be modeled as a force $-kv$ proportional and contrary to velocity, where $k > 0$ is a friction coefficient (which depends on the shape of the falling body, and on many other things!). Therefore, free fall near the Earth's surface may be modeled by the Newton equation

$$m\dot{v} = -kv - mg$$

This is a linear ODE for the velocity, whose solution is

$$v(t) = \frac{gm}{k} + e^{-(k/m)t}\left(v(0) - \frac{gm}{k}\right) \,.$$

In particular, the velocity is asymptotic to the equilibrium value $\overline{v} = gm/k$.

**e.g.** **Circuito RL.** A corrente $I(t)$ num circuito RL, de resistência $R$ e indutância $L$, é determinada pela EDO

$$L\dot{I} + RI = V$$

onde $V(t)$ é a tensão que alimenta o circuito.

- Escreva a solução geral como função da corrente inicial $I(0) = I_0$.
- Resolva a equação para um circuito alimentado com tensão constante $V(t) = E$. Esboce a representação gráfica de algumas das soluções e diga o que acontece para grandes intervalos de tempo.
- Resolva a equação para um circuito alimentado com uma tensão alternada $V(t) = E\sin(\omega t)$. Se não conseguir, mostre que a solução com $I(0) = 0$ tem a forma

$$I(t) = \frac{E}{\sqrt{R^2 + \omega^2 L^2}} \sin(\omega t - \alpha) + \frac{E\omega L}{R^2 + \omega^2 L^2} e^{-\frac{R}{L}t}$$

onde $\alpha$ é uma constante que depende de $\omega$, $L$ e $R$.

e.g. **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 $M(t)$ segue a *lei do arrefecimento de Newton*

$$\dot{T} = -k\,(T - M(t))$$

onde $k$ é uma constante positiva (que depende do material do corpo).

- Escreva a solução $T(t)$ como função da temperatura inicial $T(0) = T_0$ e de $M(s)$ com $0 \leqslant s \leqslant t$.
- Resolva a equação quando a temperatura do meio ambiente é mantida constante $M(t) = M$. Esboce a representação gráfica de algumas das soluções e diga o que acontece para grandes intervalos de tempo.
- Uma chávena de café, com temperatura inicial de $100^o$C, é colocada numa sala cuja temperatura é de $20^o$C. Sabendo que o café atinge uma temperatura de $60^o$C em 10 minutos, determine a constante $k$ do café e o tempo necessário para o café atingir a temperatura de $40^o$C.

# 15 Curves

**Caminhos.** Um *caminho* em $\mathbb{R}^n$ é uma função $\mathbf{c} : I \to \mathbb{R}^n$,

$$t \mapsto \mathbf{c}(t) = (c_1(t), c_2(t), ..., c_n(t)),$$

definida num intervalo $I \subset \mathbb{R}$. Se $\mathbf{c}$ é uma função continua (ou seja, se as suas coordenadas $c_k : I \to \mathbb{R}$, com $k = 1, 2, \dots, n$, são funções contínuas), o caminho é dito *contínuo* e a sua imagem, o subconjunto $c(I) := \{\mathbf{c}(t), \text{ com } t \in I\} \subset \mathbb{R}^n$, é dita *curva*. Se $I = [a, b]$ é um intervalo fechado e $\mathbf{c}(a) = \mathbf{c}(b)$, então $\mathbf{c}$ é dito *caminho fechado*, ou *laço*.

Por exemplo, um caminho no plano $\mathbb{R}^2$ ou no espaço $\mathbb{R}^3$, é uma função

$$t \mapsto \mathbf{r}(t) = (x(t), y(t)) \in \mathbb{R}^2 \qquad \text{ou} \qquad t \mapsto \mathbf{r}(t) = (x(t), y(t), z(t)) \in \mathbb{R}^3$$

definida num intervalo (de tempos) $t \in I \subset \mathbb{R}$.

O parâmetro $t \in I$ tem a interpretação de um "tempo", o espaço $\mathbb{R}^n$ a interpretaçao dos possíveis "estados" de um sistema físico (a posição e o momento de um planeta, as concentrações dos $n$ reagentes de uma reação química, ...). Assim, o caminho $t \mapsto \mathbf{c}(t)$ representa uma "trajetória", ou "lei horária", uma lei que determina o estado $c(t)$ do sistema em cada tempo $t \in I$. A curva $\mathbf{c}(I)$, o conjunto dos estados pelos quais passa a trajetória, é dita "órbita" do sistema.

e.g. **Retas e segmentos.** A reta que passa pelo ponto $a \in \mathbb{R}^n$ (no tempo 0) na direção do vetor não nulo $v \in \mathbb{R}^n$ é o caminho

$$t \mapsto \mathbf{a} + t\mathbf{v} \qquad \text{com} \qquad t \in \mathbb{R}.$$

O segmento que une os pontos $\mathbf{a}$ e $\mathbf{b}$ de $\mathbb{R}^n$ é, por exemplo, o caminho

$$t \mapsto \mathbf{a} + (\mathbf{b} - \mathbf{a})t \qquad \text{com} \qquad t \in [0, 1].$$

**Space filling curves!** As strange as it may look, a generic continuous path may be much different from the idea we have in mind when drawing a curve in our blackboard. This was discovered by Giuseppe Peano, who shocked the mathematical community back in 1890 exhibiting a continuous image of the unit interval $[0, 1]$ which covered the entire unit square $[0, 1] \times [0, 1]$. More amazingly, you may want to know that the "obvious" statement that a closed curve without self-intersections divides the plane in two "pieces" requires a very long and delicate proof!, and deserves the name of *Jordan curve theorem*.

**Caminhos diferenciáveis.** Dado o caminho $\mathbf{c} : I \to \mathbb{R}^n$, o vetor $(\mathbf{c}(t + \varepsilon) - \mathbf{c}(t))/\varepsilon$ representa a velocidade média entre os "tempos" $t + \varepsilon$ e $t$. O caminho é dito *diferenciável* no ponto $t \in I$ quando existe o limite

$$\dot{\mathbf{c}}(t) := \lim_{\varepsilon \to 0} \frac{\mathbf{c}(t + \varepsilon) - \mathbf{c}(t)}{\varepsilon}.$$

O vetor $\frac{d\mathbf{c}}{dt}(t) := \dot{\mathbf{c}}(t) \in \mathbb{R}^n$ é dito *derivada* do caminho $c$ no ponto $t$, ou *velocidade* do caminho $c$ no tempo $t$.

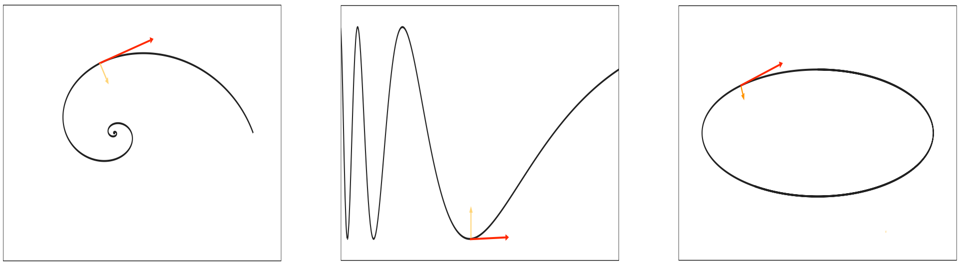
A diferenciabilidade do caminho em $t$, ou seja a existência do limite $\dot{\mathbf{c}}(t)$, é equivalente à diferenciabilidade das $n$ funções reais $t \mapsto c_k(t)$ em $t$, onde $k = 1, 2, ..., n$. A derivada $\dot{\mathbf{c}}(t)$ é portanto um vetor de coordenadas $\dot{\mathbf{c}}(t) = (\dot{c}_1(t), \dot{c}_2(t), ..., \dot{c}_n(t))$.

O caminho $\mathbf{c} : I \to X$ é dito *diferenciável* quando é diferenciável para todo tempo $t \in I$.

Se $t \mapsto \mathbf{c}(t) = (c_1(t), c_2(t), \dots, c_n(t)) \in \mathbb{R}^n$ é um caminho diferenciável, então a sua derivada $\dot{\mathbf{c}} : I \to \mathbb{R}^n$ é também um caminho, e faz sentido definir as derivadas sucessivas, como

$$\dot{\mathbf{c}} = \frac{d\mathbf{c}}{dt}, \qquad \ddot{\mathbf{c}} = \frac{d^2\mathbf{c}}{dt^2} := \frac{d}{dt}\left(\frac{d\mathbf{c}}{dt}\right), \qquad \dddot{\mathbf{c}} = \frac{d^3\mathbf{c}}{dt^3} := \frac{d}{dt}\left(\frac{d^2\mathbf{c}}{dt^2}\right), \qquad \dots$$

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



As curvas $(e^t\cos(3t), e^t\sin(3t))$, $(t, \sin(1/t))$, e $(2\cos(t), \sin(-t)$.

**Reparametrizatons.** A curve, seen as a subset of some $\mathbb{R}^n$, may have different paramerizations. Namely, if $\mathbf{c} : I \to \mathbb{R}^n$ is a continuous path, and $\varphi : J \to I$, sending $s \mapsto t(s)$, is a continuous function from the interval $J$ onto the interval $I$, then the composition $\mathbf{c} \circ \varphi : s \mapsto \mathbf{c}(t(s))$ is a continuous path from $J$ onto the same curve $\mathbf{c}(I)$. If both the path $c$ and the reparametrization $\varphi$ are differentiable, the velocity of the path $s \mapsto \mathbf{c}(t(s))$ is $\frac{d\mathbf{c}}{dt}(t(s)) \cdot \frac{dt}{ds}$.

**e.g. Uniform circular motion.** *Uniform circular motion* in the Euclidean plane is described by the path

$$t \mapsto \mathbf{r}(t) = (R\cos(\omega t), R\sin(\omega t)) \ .$$

Here $R > 0$ is a fixed radius, and $\omega > 0$ is an angular velocity. Indeed, the trajectory describes a circle $\{x^2 + y^2 = R^2\}$ of radius $R$ around the origin. The velocity is

$$\mathbf{v}(t) = \frac{d\mathbf{r}}{dt}(t) = (-R\omega\sin(\omega t), R\omega\cos(\omega t)) \ ,$$

and the acceleration is

$$\mathbf{a}(t) = \frac{d^2\mathbf{r}}{dt^2}(t) = \left(-R\omega^2\cos(\omega t), -R\omega^2\sin(\omega t)\right) \ .$$

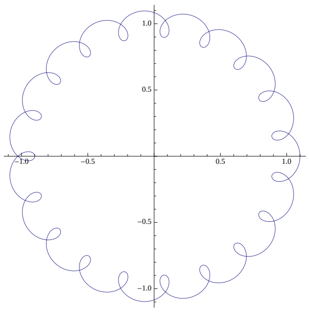
In particular, $\langle \mathbf{a}(t), \mathbf{v}(t)\rangle = 0$, i.e. the acceleration is orthogonal to the velocity, and it is directed towards the center of the orbit, since $\mathbf{a}(t) = -\omega^2 r(t)$. The quotient between the scalar velocity $v(t) = \|\mathbf{v}(t)\| = R\omega$ and the radius of the circle is the *angular velocity* $\omega$.

**e.g. Epicycles.** According to Aristotle and Plato, "all movements are combinations of circular uniform motions". This idea is at the basis of the cosmology of Hipparchus and Ptolemy, as transmitted to us in the *Almagest*. "Fixed" stars describe circles in the sky. "Wandering" (i.e. *planets*, from the greek πλανήτης) stars describe a circle (*epicycle*) around a circle, which again describe a circle around a circle, . . . , which describes a circle around a first circle (*deferent*).

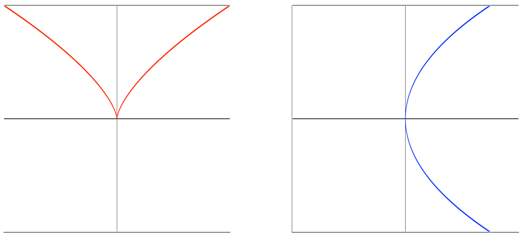
**help:** `Mathematica®8` plots parametric curves. For example, the command

```
ParametricPlot[{Cos[t] + 0.1 Cos[20 t], Sin[t] + 0.1 Sin[20 t]}, {t,  0, 2 Pi}]
```

produces the following pictures of an epicycle



**e.g. Cusps.** Differentiability depends on the parametrization of the path, i.e. on the time law, and not on the curve! For example, the path $t \mapsto (t^3, t^2)$, with $t \in [-1,1]$, describe the *cusp* $y^3 = x^2$ in the plane. Nevertheless, it is differentiable, and its velocity is the path $t \mapsto (3t^2, 2t)$. The apparent singularity at $t = 0$ is reached with zero velocity!



The cusp $t \mapsto (t^3, t^2)$ and its velocity.

**ex:** Esboce as seguintes curvas no plano, e calcule velocidade e aceleração, nos pontos onde podem ser definidas.

$$\mathbf{r}(t) = \left(t, t^2\right) \quad \text{com } t \in \mathbb{R}\,, \qquad \mathbf{r}(t) = \left(t^3, t^2\right) \quad \text{com } t \in \mathbb{R}\,,$$
$$\mathbf{r}(t) = (t, |t|) \quad \text{com } t \in [-1,1]\ , \qquad \mathbf{r}(t) = (\cos\theta, \sin\theta) \quad \text{com } \theta \in [0, 2\pi]\ ,$$
$$\mathbf{r}(t) = (t, [t]) \quad \text{com } t \in [-2,2]\ , \qquad \mathbf{r}(t) = (t, \sin(1/t)) \quad \text{com } t \in \,]0, \infty[\ .$$
$$\mathbf{r}(t) = (|\sin(5t)|\cos(2t), |\sin(5t)|\sin(2t)) \quad \text{com } t \in [0, 2\pi]\ ,$$
$$\mathbf{r}(t) = (\cos(t) + 0.1\cos(17t), \sin(t) + 0.1\sin(17t)) \quad \text{com } t \in [0, 2\pi]\,.$$

**ex:** Verifique que a trajetória

$$t \mapsto \mathbf{r}(t) = (a\cos t, b\sin t)\ ,$$

com $t \in \mathbb{R}$ e $a, b > 0$, descreve a elipse $x^2/a^2 + y^2/b^2 = 1$.

**ex:** Esboce a trajetória

$$t \mapsto \mathbf{r}(t) = (R\cos t, R\sin t, bt)\ ,$$

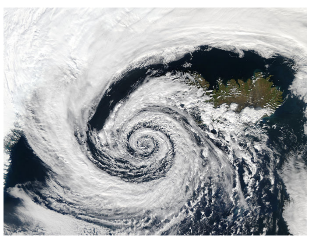
com $t \in \mathbb{R}$ e $R, b > 0$, descrita por uma partícula em movimento numa *hélice circular*.

**ex:** Determine umas equações paramétricas para a parábola $x = y^2 + 1$ e para a hipérbole $x^2 - y^2 = 1$ com $x > 0$ (lembre a identidade $\cosh^2\theta - \sinh^2\theta = 1$ entre as funções "hiperbólicas").

**Smooth paths.** A path is said of *class* $\mathcal{C}^0$ if it is continuous, of *class* $\mathcal{C}^1$ if its derivative is continuous. Using induction, it is said of class $\mathcal{C}^{k+1}$ if its derivative is of class $\mathcal{C}^k$. It is said of class $\mathcal{C}^\infty$ if it is of class $\mathcal{C}^k$ for any $k$, namely if all its derivatives are continuous.

Trajectories of physics used to have so many derivatives as we want (simply because most physical laws are written in terms of derivatives!), and we'll refer to them as "smooth", without specifying their regularity. Meanwhile, you must keep in mind that there are continuous paths which are nowhere differentiable. Actually, as shown by Weierstrass, almost all continuous paths are like that! They play a role in models of phenomena like the Brownian motion or turbolence ...

**e.g. Espiral logarítmica.** A recurrent pattern in Nature is the *logarithmic spiral*.

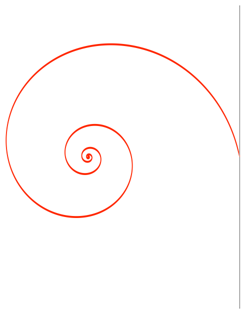




It is defined in polar coordinates, $r$ and $\theta$, by the law

$$r = c \cdot \lambda^{\theta} ,$$

for some constants $c > 0$ and $\lambda > 0$. We may parametrize the angle as $t \mapsto \theta(t) = \omega t$, for some angular velocity $\omega > 0$. Then, the logarithmic spiral is the curve drown by the path

$$t \mapsto (Ae^{-\alpha t}\cos(\omega t), Ae^{-\alpha t}\sin(\omega t)) .$$



**Theorem 15.1. (Teorema do valor médio)** *Seja* $\mathbf{c} : [a, b] \to \mathbb{R}^n$ *um caminho contínuo, diferenciável em* $]a, b[$ *e com derivada limitada por*

$$\|\dot{\mathbf{c}}\,(t)\| \leqslant K$$

*para todo* $t \in\, ]a, b[$. *Então*

$$\|\mathbf{c}\,(b) - \mathbf{c}\,(a)\| \leqslant K \cdot |b - a| \ .$$

De fato, o teorema do valor médio aplicado à função real $t \mapsto \langle \mathbf{c}\,(t)\,, \mathbf{c}\,(b) - \mathbf{c}\,(a)\rangle$, implica que existe um tempo $\bar{t} \in\, ]a, b[$ tal que

$$\|\mathbf{c}\,(b) - \mathbf{c}\,(a)\|^2 = \left\langle \dot{\mathbf{c}}\,(\bar{t})\,, \mathbf{c}\,(b) - \mathbf{c}\,(a)\right\rangle \cdot (b - a) \ .$$

Pela desigualdade de Cauchy-Schwarz

$$\|\mathbf{c}\,(b) - \mathbf{c}\,(a)\|^2 \leqslant K \cdot \|\mathbf{c}\,(b) - \mathbf{c}\,(a)\| \cdot |b - a| \ ,$$

e portanto, ou $\|\mathbf{c}\,(b) - \mathbf{c}\,(a)\| = 0$, ou $\|\mathbf{c}\,(b) - \mathbf{c}\,(a)\| \leqslant K \cdot |b - a|$.

**Comprimento de uma curva.** The length of the segment between the vectors $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{y} \in \mathbb{R}^n$, the curve $\overline{\mathbf{xy}} := \{\mathbf{x} + t\mathbf{y}\,,\, t \in [0, 1]\}$ is, by definition, the norm of the vector $\mathbf{y} - \mathbf{x}$, namely

$$\ell(\overline{\mathbf{xy}}) := \|\mathbf{y} - \mathbf{x}\| .$$

If $\mathbf{c} : [a, b] \to \mathbb{R}^n$ is a path made of straight segments between the points $\mathbf{x}_n = \mathbf{c}(t_n)$ and $\mathbf{x}_{n+1} = \mathbf{c}(t_{n+1})$, given the sequence of times $a = t_0 < t_1 < \dots t_n < t_{n+1} < \dots t_N = b$, it is natural to define its length as the sum $\sum_{n=0}^{N-1} \|\mathbf{c}(t_{n+1}) - \mathbf{c}(t_n)\|$. Therefore, a natural definition of *length* of a continuous path $\mathbf{c} : [a, b] \to \mathbb{R}^n$ is

$$\ell(\mathbf{c}) := \sup \sum_{n=0}^{N-1} \|\mathbf{c}(t_{n+1}) - \mathbf{c}(t_n)\| ,$$

where the "sup" is taken over all finite partitions $a = t_0 < t_1 < \dots t_n < t_{n+1} < \dots t_N = b$ of the time interval $[a, b]$. It is obvious from this definition that the length does not depend on the parametrization of the path. The curve is called *rectifiable* when $\ell(\mathbf{c}) < \infty$.

If we try to approximate a differentiable path $\mathbf{c} : [a, b] \to \mathbb{R}^n$ by a polygonal path between the points $\mathbf{c}(t_k)$ and $\mathbf{c}(t_{k+1})$, we may observe that the length of the portion of the path between $t_k$ and $t_{k+1} = t_k + dt$ is, to a first approximation, $\simeq \|\dot{\mathbf{c}}(t_k)\| \cdot dt$. It turns out (but it is not obvious!)

that the length of a differentiable curve may equivalently be computed/defined as the integral of its scalar velocity:

$$\ell(\mathbf{c}) = \int_a^b \|\dot{\mathbf{c}}(t)\| \, dt \, .$$

For example, a planar path like $t \mapsto \mathbf{r}(t) = (x(t), y(t))$, or a 3-dimensional path like $t \mapsto \mathbf{r}(t) = (x(t), y(t), z(t))$, with times $t \in [a,b]$, have length

$$\int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2} \, dt \qquad \text{or} \qquad \int_a^b \sqrt{\dot{x}(t)^2 + \dot{y}(t)^2 + \dot{z}(t)^2} \, dt \, .$$

**ex:** Calcule o comprimento ...
... do arco de circunferência $\theta \mapsto (\cos\theta, \sin\theta)$ com $\theta \in [\pi/2, 2\pi]$ ,
... da espiral logarítmica $t \mapsto (e^{-t}\cos t, e^{-t}\sin t)$ com $t \in [0, \infty[$ ,
... do arco de parábola $t \mapsto \left(t, t^2/2\right)$ com $t \in [0,1]$ (considere a susbtituição $t = \sinh s$).

**Comprimento de um gráfico.** Seja $f(t)$ uma função real com derivada contínua definida no intervalo $[a,b]$. O gráfico de $f$, o conjunto

$$\Gamma_f = \left\{(t, f(t)) \in \mathbb{R}^2 \ \text{ com } \ t \in [a,b]\right\} \subset \mathbb{R}^2 \, ,$$

é a imagem do caminho $t \mapsto (t, f(t))$ com $t \in [a,b]$. Em particular, o seu comprimento é

$$\ell\left(\Gamma_f\right) = \int_a^b \sqrt{1 + f'(t)^2} \, dt \, .$$

# 16 Scalar fields

**Scalar fields.** A *scalar field* is a real valued function $f : X \subset \mathbb{R}^n \to \mathbb{R}$ defined in some domain $X \subset \mathbb{R}^n$. We use both the notations $f(\mathbf{x}) = f(x_1, x_2, \dots, x_n)$ for the value of the field $f$ at the point $\mathbf{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$.

Thus, a scalar field is a number $f(\mathbf{x})$ attached to any point $\mathbf{x} \in X$. For example, the "coordinate functions" $\mathbf{x} = (x_1, x_2, \dots, x_n) \mapsto x_k$, for $k = 1, 2, \dots, n$, are scalar fields, which give the values of the different coordinates attached to a given point $\mathbf{x} \in \mathbb{R}^n$.

A scalar field $f : X \subset \mathbb{R}^n \to \mathbb{R}$ is said *continuous at the point* $\mathbf{x} \in X$ if for any $\varepsilon > 0$ there exists $\delta > 0$ such that $\|\mathbf{y} - \mathbf{x}\| < \delta$ implies $|f(\mathbf{y}) - f(\mathbf{x})| < \varepsilon$. This is the same saying that $f(\mathbf{x}_n) \to f(\mathbf{x})$ for any sequence $\mathbf{x}_n \to \mathbf{x}$. A scalar field $f$ is said *continuous* if it is continuous at all points $\mathbf{x} \in X$ of its domain.

e.g. **Temperature.** The temperature of a ideal gas, as a function of the pressure $P$ and the volume $V$, is

$$T(P, V) = \frac{1}{nR} PV .$$

where $n$ is the number of moles, and $R \simeq 8.314 \times 10^7$ J/K·mol. Curves with constant temperature are hyperbolas $PV =$ constant.

**Level sets.** Let $f : X \subset \mathbb{R}^n \to \mathbb{R}$ be a scalar field, and $\lambda$ one of the values of $f$. The $\lambda$-*level set* of $f$ is the subset

$$\Sigma_\lambda = f^{-1}(\{\lambda\}) := \{\mathbf{x} \in X \text{ such that } f(\mathbf{x}) = \lambda\} \subset X .$$

It may be one single point, or even all of $X$ (if $f$ is a constant function). For reasonable (i.e. sufficiently smooth) fields and generic values $\lambda$ (in some precise meaning), it is a *hypersurface*, a set of "dimension" $n - 1$ inside $\mathbb{R}^n$. The *graph* of $f$ is the set

$$\mathcal{G}_f := \{(\mathbf{x}, \lambda) \in X \times \mathbb{R} \text{ t.q. } f(\mathbf{x}) = \lambda\} \subset X \times \mathbb{R} .$$

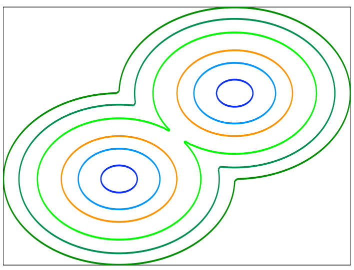
For example, if $f(x, y)$ is a smooth scalar field defined in $X \subset \mathbb{R}^2$, then

$$\Sigma_\lambda := \left\{(x, y) \in X \subset \mathbb{R}^2 \text{ t.q. } f(x, y) = \lambda\right\}$$

is, for generic values of $\lambda$, a *level curve*. The graph of $f$ is the surface

$$\mathcal{G}_f := \{(x, y, z) \in X \times \mathbb{R} \text{ t.q. } f(x, y) = z\} \subset \mathbb{R}^3 .$$

Of course, it is not easy to draw the graph of a function defined on $\mathbb{R}^n$ when $n \geqslant 3$ !



Curvas de nível e gráfico.

ex: Esboce as curvas de nível e os gráficos das seguintes funções, nos domínios onde podem ser definidas:

$$f(x, y) = x + y \qquad f(x, y) = xy \qquad f(x, y) = x^2 + 2y^2 \qquad f(x, y) = \sqrt{1 - x^2 - y^2}$$

$$f(x, y) = \log\left(x^2 + y\right) \qquad f(x, y) = x^2 - y^2 \qquad f(x, y) = \sin(xy)$$

**e.g. Equação de Van der Waals.**

$$\left(P + \frac{a}{V^2}\right)(V - b) = nRT$$

onde $b$ representa o efeito das dimensões finitas das moléculas e $a/V^2$ o efeito das forças moleculares de coesão.

**Directional and partial derivatives.** Let $f : X \subset \mathbb{R}^n \to \mathbb{R}$ be a scalar field. Given any point $\mathbf{x} \in \mathbb{R}^n$, a non-zero vector $\mathbf{v} \in \mathbb{R}^n$ defines a straight line $t \mapsto \mathbf{x} + t\mathbf{v}$ passing through the point $\mathbf{x}$ at time $t = 0$ with velocity $\mathbf{v}$. The *directional derivative* of the field $f$ at the point $\mathbf{x} \in X$ along the direction of the vector $\mathbf{v} \in \mathbb{R}^n$ is the derivative of the real valued function $t \mapsto f(\mathbf{x} + t\mathbf{v})$ computed at time $t = 0$, namely

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}) := \left.\frac{d}{dt} f(\mathbf{x} + t\mathbf{v})\right|_{t=0} .$$

Another notation for the directional derivative is $\mathcal{L}_{\mathbf{v}} f(\mathbf{x})$ (called *Lie derivative* of the scalar field $f$ along the constant vector field $\mathbf{v}$). Some authors reserve the name of directional derivative to the case when $\mathbf{v}$ is a unit vector, i.e. when $\|\mathbf{v}\| = 1$.

If we compute the directional derivative of $f$ w.r.t. the direction $\mathbf{v} = \mathbf{e}_k$, the $k$-th vector of the canonical basis $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ of $\mathbb{R}^n$, we get the *partial derivative* of $f$ at the point $\mathbf{x}$ with respect to the variable $x_k$, denoted as

$$\frac{\partial f}{\partial x_k}(x) := \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{e}_k) - f(\mathbf{x})}{t} .$$

Thus, in order to compute the partial derivative $\frac{\partial f}{\partial x_k}(\mathbf{x})$, you “freeze” all the remaining coordinates $x_i$, with $i \neq k$, to their values at the point $\mathbf{x}$, and compute the usual derivative of the real valued function $t \mapsto f(x_1, \dots, x_{k-1}, t, x_{k+1}, \dots, x_n)$ at the point $t = x_k$.

For example, the partial derivatives of the scalar field $f(x, y)$ defined in some domain of the Cartesian plane $\mathbb{R}^2$ with coordinates $(x, y)$ are the limits

$$\frac{\partial f}{\partial x}(x, y) = \lim_{\varepsilon \to 0} \frac{f(x + \varepsilon, y) - f(x, y)}{\varepsilon} \qquad \text{and} \qquad \frac{\partial f}{\partial y}(x, y) = \lim_{\varepsilon \to 0} \frac{f(x, y + \varepsilon) - f(x, y)}{\varepsilon} .$$

**Higher order derivatives and smooth fields.** Partial derivatives of a scalar field are themselves scalar fields, so it make sense to compute their partial derivatives,

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x} , \qquad \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2} , \qquad \dots$$

and so on.

A scalar field is said of *class* $\mathcal{C}^0$ if it is continuous, of *class* $\mathcal{C}^1$ if its partial derivatives are continuous. Using induction, it is said of class $\mathcal{C}^{k+1}$ if its partial derivatives are of class $\mathcal{C}^k$. It is said of class $\mathcal{C}^\infty$ if it is of class $\mathcal{C}^k$ for any $k$, namely if all its partial derivatives exist and are continuous. According to *Schwarz theorem*, if a scalar field $f$ is of class $\mathcal{C}^k$ in some domain, then its partial derivatives up to order $\leqslant k$ commute. Thus, for example,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$$

if the field $f$ is of class $\mathcal{C}^2$.

**Differentiable scalar fields.** A scalar field $f : X \subset \mathbb{R}^n \to \mathbb{R}$ is *differentiable* at the point $\mathbf{x} \in X$ if there exists a linear map $L : \mathbb{R}^n \to \mathbb{R}$ such that, for any $\mathbf{v} \in \mathbb{R}^n$ with sufficiently small norm,

$$f(\mathbf{x} + \mathbf{v}) = f(\mathbf{x}) + L \cdot \mathbf{v} + e(\mathbf{v})$$

where the “error” $e(\mathbf{v})$ is so small that

$$\lim_{\|\mathbf{v}\| \to 0} \frac{e(\mathbf{v})}{\|\mathbf{v}\|} = 0 .$$

The linear map $L$ is called *differential of $f$ at* $\mathbf{x}$, and denoted by $df(\mathbf{x})$ (or also $Df(\mathbf{x})$, or $f'(\mathbf{x})$). Above, we used the notation $L \cdot \mathbf{v} = L_1 v_1 + L_2 v_2 + \cdots + L_n v_n$ for the value of the linear map $L$ at the vector $\mathbf{v}$.

It is clear that a linear map $L$ as above, if it exists, must be unique. It also immediate to see that a differentiable field is continuous, since both $L \cdot \mathbf{v} \to 0$ and $e(\mathbf{v}) \to 0$, and consequently $f(\mathbf{x} + v) \to f(\mathbf{x})$, as $\|\mathbf{v}\| \to 0$.

If $f$ is differentiable at $\mathbf{x}$, its directional and partial derivatives may be computed as

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}) = df(\mathbf{x}) \cdot \mathbf{v} \qquad \text{and} \qquad \frac{\partial f}{\partial x_k}(\mathbf{x}) = df(\mathbf{x}) \cdot \mathbf{e}_k \, .$$

Therefore, the *differential* of a scalar field $f : X \subset \mathbb{R}^n \to \mathbb{R}$ at the point $\mathbf{x} \in X$ is the linear form $df(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ given in coordinates by

$$df(\mathbf{x}) := \frac{\partial f}{\partial x_1}(\mathbf{x})\, dx_1 + \frac{\partial f}{\partial x_2}(\mathbf{x})\, dx_2 + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{x})\, dx_n$$

where $dx_k$, the differential of the coordinate function $\mathbf{x} \mapsto x_k$, is the linear form which takes the vector $\mathbf{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ into the scalar $dx_k \cdot \mathbf{v} := v_k$.

**Gradient.** A convenient way to write the differential of a scalar field is the following. The *gradient* of the scalar field $f : X \subset \mathbb{R}^n \to \mathbb{R}$ at the point $\mathbf{x} \in X$ is the vector whose components are the partial derivatives of $f$ at $\mathbf{x}$, namely

$$\nabla f(\mathbf{x}) := \left( \frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \ldots, \frac{\partial f}{\partial x_n}(x) \right) .$$

An alternative notation, also used by physicists, is $\operatorname{grad} f(\mathbf{x})$.

In particular, the directional derivative of the differentiable field $f$ along the direction of $\mathbf{v} \in \mathbb{R}^n$ at the point $\mathbf{x}$ is

$$\frac{\partial f}{\partial \mathbf{v}}(x) = df(\mathbf{x}) \cdot v = \langle \nabla f(\mathbf{x}), \mathbf{v} \rangle \, .$$

If $\mathbf{v}$ is a unit vector, i.e. $\|\mathbf{v}\| = 1$, then the Schwarz inequality says that

$$-\|\nabla f(\mathbf{x})\| \leqslant \frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}) \leqslant \|\nabla f(\mathbf{x})\| \, .$$

More precisely,

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{x}) = \|\nabla f(\mathbf{x})\| \cdot \cos(\theta)$$

where $\theta$ is the angle between $\nabla f(\mathbf{x})$ and $\mathbf{v}$. Therefore, the directional derivative is the component of $\mathbf{v}$ along the direction of the gradient $\nabla f(\mathbf{x})$ . In particular, the directional derivative is maximal in the direction of the gradient, namely for $v = \nabla f(\mathbf{x}) / \|\nabla f(\mathbf{x})\|$, and minimal in the opposite direction, for $\mathbf{v} = -\nabla f(\mathbf{x}) / \|\nabla f(\mathbf{x})\|$. Thus, the gradient points to the direction along which the function increases most rapidly.

Computation of the gradient may be simplified using the following properties, easy consequences of the corresponding properties of the derivative:

$$\nabla f = 0 \qquad \text{if } f \text{ is constant}$$

$$\nabla (f + g) = \nabla f + \nabla f$$

$$\nabla (fg) = f\, \nabla g + g\, \nabla f$$

$$\nabla (f/g) = (g\, \nabla f - f\, \nabla g) / g^2$$

**Vector fields.** A *vector field* is a vector valued function $\mathbf{F} : X \subset \mathbb{R}^n \to \mathbb{R}^k$ defined in some domain $X \subset \mathbb{R}^n$, with coordinates $F_1(x), F_2(x), \dots, F_k(x)$ which are $k$ scalar fields. Continuity of a vector field is defined component-wise. Thus, a vector field $\mathbf{F}$ is continuous if all its coordinates $F_i$ are continuous scalar fields.

The gradient of a differentiable scalar field $f : X \subset \mathbb{R}^n \to \mathbb{R}$, thought as a function $\nabla f : X \subset \mathbb{R}^n \to \mathbb{R}^n$, sending $x \mapsto \nabla f(x)$, is an example (actually a most important one!) of a vector field.

A vector field $\mathbf{F} : X \subset \mathbb{R}^n \to \mathbb{R}^k$ is *differentiable* at the point $x \in X$ if there exists a linear map $\mathbf{L} : \mathbb{R}^n \to \mathbb{R}^k$ such that, for any $\mathbf{v} \in \mathbb{R}^n$ with sufficiently small norm,

$$\mathbf{F}(\mathbf{x}+\mathbf{v}) = \mathbf{F}(\mathbf{x}) + \mathbf{L}\cdot\mathbf{v} + \mathbf{E}(\mathbf{v})$$

where the error $\mathbf{E}(\mathbf{v})$ is so small that

$$\lim_{\|\mathbf{v}\|\to 0} \frac{\mathbf{E}(\mathbf{v})}{\|\mathbf{v}\|} = 0\,.$$

The linear map $\mathbf{L}$ is called *differential of* $\mathbf{F}$ *at* $\mathbf{x}$, and denoted by $D\mathbf{F}(\mathbf{x})$, or $\mathbf{F}'(\mathbf{x})$.

Thus, if $\mathbf{F}$ is differentiable at $\mathbf{x}$, its directional and partial derivatives may be computed as

$$\frac{\partial \mathbf{F}}{\partial \mathbf{v}}(\mathbf{x}) = D\mathbf{F}(\mathbf{x})\cdot\mathbf{v} \qquad \text{and} \qquad \frac{\partial \mathbf{F}}{\partial x_k}(\mathbf{x}) = D\mathbf{F}(\mathbf{x})\cdot\mathbf{e}_k\,.$$

Therefore, the matrix which represents the differential $D\mathbf{F}(\mathbf{x})$ in the canonical basis of $\mathbb{R}^n$ and $\mathbb{R}^k$ is the *Jacobian matrix*

$$J\mathbf{F}(\mathbf{x}) := \left(\frac{\partial F_i}{\partial x_j}(\mathbf{x})\right) \in \mathrm{Mat}_{k\times n}(\mathbb{R})\,.$$

**Differentiability classes.** The existence of partial derivatives does not implies differentiability. For example, the function $f(x,y)$ equal to 1 for $xy = 0$ (i.e. on the two coordinate axis) and equal to 0 for $xy \neq 0$ (i.e. outside the axis) does admit partial derivatives at the origin, but it is not even continuous there. Even the existence of directional derivatives for all non-zero directions does not implies differentiability.

More interesting is that the existence and continuity of all first partial derivatives in some domain does implies differentiability. The class of real valued functions having continuous partial derivatives inside the domain $X \subset \mathbb{R}^n$ is named the class of $\mathcal{C}^1(X,\mathbb{R})$ functions.

**Chain rule for scalar fields and paths.** Let $\mathbf{r} : I \subset \mathbb{R} \to X \subset \mathbb{R}^n$ be a differentiable path, given explicitly by $t \mapsto \mathbf{c}(t) = (x_1(t), x_2(t), \dots, x_n(t))$, and let $f : X \subset \mathbb{R}^n \to \mathbb{R}$ be a differentiable scalar field. The composite function $f \circ \mathbf{c} : I \to \mathbb{R}$ (which is a real valued function of a real variable) is differentiable and its derivative may be computed as

$$\begin{aligned} \frac{d}{dt} f(\mathbf{c}(t)) &= \langle \nabla f(\mathbf{c}(t)), \dot{\mathbf{c}}(t)\rangle \\ &= \frac{\partial f}{\partial x_1}(\mathbf{c}(t))\cdot\frac{dx_1}{dt}(t) + \frac{\partial f}{\partial x_2}(\mathbf{c}(t))\cdot\frac{dx_2}{dt}(t) + \dots + \frac{\partial f}{\partial x_n}(\mathbf{c}(t))\cdot\frac{dx_n}{dt}(t) \end{aligned}$$

Por exemplo, se $t \mapsto \mathbf{r}(t) = (x(t), y(t)) \in \mathbb{R}^2$ é um caminho com velocidade $\mathbf{v}(t) = (\dot{x}(t), \dot{y}(t))$, e $f(x,y)$ um campo escalar, então

$$\frac{d}{dt} f(\mathbf{r}(t)) = \langle \nabla f(\mathbf{r}(t)), \mathbf{v}(t)\rangle = \frac{\partial f}{\partial x}(\mathbf{r}(t))\cdot\dot{x}(t) + \frac{\partial f}{\partial y}(\mathbf{r}(t))\cdot\dot{y}(t)\,.$$

**e.g. Linear field.** Inner product by a fixed vector $\mathbf{w} \in \mathbb{R}^3$ (or in any other $\mathbb{R}^n$) defines a linear scalar field according to $f(\mathbf{r}) = \langle \mathbf{w}, \mathbf{r}\rangle$. One easily compute that

$$\nabla f(\mathbf{r}) = \mathbf{w} \qquad \text{and therefore} \qquad \frac{\partial f}{\partial \mathbf{v}}(\mathbf{r}) = \langle \mathbf{w}, \mathbf{v}\rangle$$

for any direction $\mathbf{v} \in \mathbb{R}^3$ and any point $\mathbf{r} \in \mathbb{R}^3$. Level surfaces of $f$ are the affine planes orthogonal to the vector $\mathbf{w} \neq 0$, namely

$$\Sigma_\lambda = \{\mathbf{x} \in \mathbb{R}^3 \text{ such that } \langle \mathbf{w}, \mathbf{x}-\mathbf{a}\rangle = 0\} = \mathbf{a} + \mathbf{w}^\perp\,,$$

if $\mathbf{a} \in \Sigma_\lambda$ is any point where $f(\mathbf{a}) = \langle \mathbf{w}, \mathbf{a}\rangle = \lambda$.

**e.g. Norm and its powers.** The norm may be viewed as a scalar field $\mathbf{r} \mapsto r := \|\mathbf{r}\|$. One computes, for any $k = 1, \dots, n$,

$$\frac{\partial r}{\partial x_k}(\mathbf{r}) = \frac{2x_k}{2\sqrt{x_1 + x_2^2 + \cdots + x_n^2}} = \frac{x_k}{r} \,,$$

and consequently

$$\nabla r(\mathbf{r}) = \frac{\mathbf{r}}{r}$$

for $\mathbf{r} \neq 0$. The gradient of the $N$-th power $f(\mathbf{r}) = r^N$ is therefore

$$\nabla f(\mathbf{r}) = N r^{N-1} \nabla r = N \frac{\mathbf{r}}{r^{2-N}} \,.$$

In particular, for $\varphi(\mathbf{r}) = \|\mathbf{r}\|^2$,

$$\nabla \varphi(\mathbf{r}) = 2\mathbf{r} \qquad \text{and therefore} \qquad \frac{\partial \varphi}{\partial \mathbf{v}}(\mathbf{r}) = 2\langle \mathbf{r}, \mathbf{v} \rangle \,.$$

for any direction $\mathbf{v} \in \mathbb{R}^3$. Observe that level surfaces of $\varphi(\mathbf{r}) = r^2$ in $\mathbb{R}^3$ are the spheres $\Sigma_\lambda = \{\mathbf{x} \in \mathbb{R}^3 \text{ such that } \|\mathbf{x}\|^2 = \lambda\}$ of radius $\sqrt{\lambda}$, for $\lambda \geqslant 0$.

Thus, if a particle moves inside a fixed sphere, i.e. if $t \mapsto \mathbf{r}(t)$ is a path with constant $\|\mathbf{r}(t)\|^2 = \lambda$, then $\langle \dot{\mathbf{r}}(t), \mathbf{r}(t) \rangle = 0$, so that the velocity $\mathbf{v}(t) = \dot{\mathbf{r}}(t)$ is orthogonal to the position vector $\mathbf{r}(t)$ at every time $t$.

**ex:** Calcule as derivadas parciais de primeira e segunda ordem das seguintes funções, nos domínios onde podem ser definidas:

$$f(x,y) = \sqrt{x^2 + y^2} \qquad f(x,y,z) = x^3 + y^2 + zxy \qquad f(x,y) = \log\left(x^2 + y^2\right)$$

$$f(x,y) = e^{x+y} \qquad f(x,y) = \frac{\sin(x^2)}{y} \qquad f(x,y) = e^{y \log x}$$

**ex:** Calcule o gradiente das seguintes funções, nos domínios onde podem ser definidas:

$$f(x,y) = \sqrt{x^2 + y^2 + z^2} \qquad f(x,y) = x^2 - y^2 \qquad f(x,y) = \sin\left(x^2 + y^2\right)$$

$$f(x,y) = e^{-x^2 - y^2} \qquad f(x,y,z) = xyz \qquad f(x,y) = e^{y \log x}$$

**ex:** Calcule a derivada $\frac{d}{dt} f(\mathbf{r}(t))$ dos seguintes campos $f(\mathbf{r})$ ao longo dos respetivos caminhos $t \mapsto \mathbf{r}(t)$ nos tempos indicados.

$$f(x,y) = x^3 y - xy^2 \qquad t \mapsto (t^2, t^3) \qquad t = 0 \,,$$

$$f(x,y) = xy \qquad t \mapsto (2e^t \cos(t), 2e^t \sin(t)) \qquad t = 1 \,,$$

$$f(x,y,z) = x^2 + y^2 + z^2 \qquad t \mapsto (\cos(t), \sin(t), t) \qquad t = \pi \,,$$

**e.g. Gravitational field.** The gravitational force field produced by a star of mass $M$ placed at the origin of $\mathbb{R}^3$ is

$$\mathbf{F}(\mathbf{r}) = -GM \frac{\mathbf{r}}{\|\mathbf{r}\|^3}$$

where $G \simeq 6.670 \times 10^{-8}$ dina-cm$^2$/gm$^2$. It is the gradient of the *gravitational potential*

$$\varphi(\mathbf{r}) = \frac{GM}{\|\mathbf{r}\|} \,.$$

**ex:** Mostre que o potencial Newtoniano $\varphi(\mathbf{r}) = 1/\|\mathbf{r}\|$ em $\mathbb{R}^3 \backslash \{(0,0,0)\}$ satisfaz a *equação de Laplace*

$$\frac{\partial \varphi}{\partial x^2} + \frac{\partial \varphi}{\partial y^2} + \frac{\partial \varphi}{\partial z^2} = 0 \,.$$

**ex:** A temperatura do mar num ponto $\mathbf{r} = (x, y, z)$ é dada por $T(x, y, z) = x^3 - xy + yz^2$. Uma sardinha encontra-se no ponto $\mathbf{a} = (3, 2, 1)$. Em que direcção e sentido a sardinha tem de nadar para arrefecer mais rapidamente?

**ex:** Seja $f(t)$ uma função real diferenciável. Mostre que a função $u(x, y) = f(xy)$ satisfaz a equação

$$x\frac{\partial u}{\partial x} - y\frac{\partial u}{\partial y} = 0$$

e que a função $v(x, y) = f(x/y)$ satisfaz a equação

$$x\frac{\partial v}{\partial x} + y\frac{\partial v}{\partial y} = 0$$

**ex: (aproximação linear)** Estime os seguintes valores, usando a aproximação linear

$$f(x + dx, y + dy) \simeq f(x, y) + \frac{\partial f}{\partial x}(x, y) \cdot dx + \frac{\partial f}{\partial y}(x, y) \cdot dy$$

$$e^{0.01}\sqrt{3.999} \qquad \frac{\log(1.01)}{1 + 0.001} \qquad \sqrt[3]{7.99}\sqrt{36.01}$$

**e.g. Kinetic energy and conservative systems.** Let $t \to \mathbf{r}(t) \in \mathbb{R}^2$ (or $\mathbb{R}^3$) the trajectory of a particle of mass $m > 0$, $\mathbf{v}(t) = \dot{\mathbf{r}}(t)$ its velocity and $\mathbf{a}(t) = \dot{\mathbf{v}}(t) = \ddot{\mathbf{r}}(t)$ its acceleration. The *kinetic energy* of the particle is

$$K := \frac{1}{2}m\,\|\mathbf{v}\|^2 .$$

Its time variation is

$$\frac{d}{dt}\left(\frac{1}{2}m\,\|\mathbf{v}(t)\|^2\right) = \langle m\,\mathbf{a}(t), \mathbf{v}(t)\rangle .$$

Thus, if the particle is subject to a force $\mathbf{F} = m\,\mathbf{a}$ which is orthogonal to the velocity (as a magnetic force acting on a moving charged particle) then the kinetic energy is a constant of the motion.

A force field $\mathbf{F}(\mathbf{r})$ is said *conservative* if there exists a scalar field $V(\mathbf{r})$, called potential, such that $\mathbf{F}(\mathbf{r}) = -\nabla V(\mathbf{r})$. The name is justified by the fact that the *(total) energy*, defined as

$$E := K + V = \frac{1}{2}m\,\|\mathbf{v}\|^2 + V(\mathbf{r}) ,$$

is a constant of the motion. Indeed

$$\begin{aligned} \frac{d}{dt}E(\mathbf{r}(t), \mathbf{v}(t)) &= \langle m\mathbf{a}(t), \mathbf{v}(t)\rangle + \langle \nabla V(\mathbf{r})(t), \mathbf{v}(t)\rangle \\ &= \langle m\mathbf{a}(t) - \mathbf{F}, \mathbf{v}(t)\rangle = 0 \end{aligned}$$

if the acceleration satisfies Newton equation $\mathbf{F} = m\mathbf{a}$.

**Tangent space to a level set.** Let $\Sigma_\lambda$ be a non-empty level set of the differentiable scalar field $f : X \subset \mathbb{R}^n \to \mathbb{R}$, and $\mathbf{x} \in \Sigma_\lambda$ one of its points. If $\mathbf{c} : ]-\varepsilon, \varepsilon[ \to \Sigma_\lambda$ is any differentiable curve lying entirely on the level set $\Sigma_\lambda$ and passing through $\mathbf{c}(0) = \mathbf{x}$ at time 0, then the composite function $t \mapsto f(\mathbf{c}(t))$ is constant and equal to $\lambda$, and therefore, by the chain rule,

$$0 = \left.\frac{d}{dt}f(\mathbf{c}(t))\right|_{t=0} = \langle \nabla f(\mathbf{x}), \dot{\mathbf{c}}(0)\rangle .$$

If the gradient of $f$ at $\mathbf{x}$ is different from the zero vector, i.e. $\nabla f(\mathbf{x}) \neq 0$, we deduce the the space of all such velocities $\dot{\mathbf{c}}(0)$, which we call *tangent space* to $\Sigma_\lambda$ at $\mathbf{x}$, is the normal space to the gradient of $f$ at $\mathbf{x}$.

For example, If $f(x,y,z)$ is a scalar field defined in some $X \subset \mathbb{R}^3$, then the tangent plane to the level surface $\Sigma_\lambda$ at some point $\mathbf{r}$ is the affine plane orthogonal to the gradient $\nabla f(\mathbf{r})$ and passing through the point $\mathbf{r}$, namely

$$\{\mathbf{v} \in \mathbb{R}^3 \text{ such that } \langle \nabla f(\mathbf{r}), \mathbf{v} - \mathbf{r} \rangle = 0\} .$$

The Cartesian equation of such a plane is

$$\frac{\partial f}{\partial x}(\mathbf{r}) \cdot (x-a) + \frac{\partial f}{\partial y}(\mathbf{r}) \cdot (y-b) + \frac{\partial f}{\partial z}(\mathbf{r}) \cdot (z-c) = 0 ,$$

where $\mathbf{v} = (x,y,z)$ and $\mathbf{r} = (a,b,c)$.

ex: Considere as seguintes funções:

$$f(x,y) = x^2 + y^2 \qquad f(x,y) = x^2 - y^2 \qquad f(x,y) = x^2$$

$$f(x,y) = xy \qquad f(x,y) = e^{x^2+y^2} \qquad f(x,y) = 1 - y - x^2$$

$$f(x,y,z,) = x^2 + y^2 + z^2 \qquad f(x,y,z) = x^2 + y^2 - z^2 \qquad f(x,y,z) = x^2 + y^2 - z$$

Calcule o gradiente num ponto genérico onde estão definidas. Determine a recta/superfície tangente à curva/superfície de nível no ponto $\mathbf{r} = (1,1)$ (ou $\mathbf{r} = (1,1,1)$).

**Critical points and local extrema.** Let $f : X \to \mathbb{R}$ be a differentiable scalar field defined in some domain $X \subset \mathbb{R}^n$. *Critical points* (or *stationary points*) of $f$ are points $\mathbf{a} \in X$ where the differential (hence the gradient) vanishes, i.e. where

$$df(\mathbf{a}) = 0.$$

Observe that this means that all partial derivatives vanish.

If $f$ has a local maximum or minimum at some interior point $\mathbf{a} \in X$ (as, for example, the origin for $f(x,y) = -x^2 - y^2$ or $f(x,y) = x^2 + y^2$), then it must be a critical point, since the directional derivative $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a})$ must vanishes there for any vector $\mathbf{v} \in \mathbb{R}^n$. The converse is, of course, false already in dimension one. Critical points such that in any neighborhood $B_\varepsilon(\mathbf{a})$ there exists points $\mathbf{x}, \mathbf{y}$ such that $f(\mathbf{x}) < f(\mathbf{a}) < f(\mathbf{y})$ are said *saddle points*. The simplest example in the plane is the origin for $f(x,y) = xy$.

To decide if a critical point $\mathbf{a}$ is indeed a local minimum or maximum we must look at least at the second derivatives of $f$, namely its *Hessian matrix*

$$\text{Hess} f(\mathbf{a}) := \left( \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}) \right) .$$

It follows from Schwarz theorem that, if $f$ is of class $\mathcal{C}^2$, this is a symmetric matrix. But this implies that $\text{Hess} f(\mathbf{a})$ is diagonalizable, namely that there exist $n$ linear independent eigenvectors $\mathbf{w}_1, \ldots, \mathbf{w}_n$, forming a base of $\mathbb{R}^n$, and corresponding eigenvalues $\lambda_1, ..., \lambda_n$, such that

$$\text{Hess} f(\mathbf{a}) \cdot \mathbf{w}_k = \lambda_k \mathbf{w}_k$$

Now, given any direction $\mathbf{v} = \sum_{k=1}^n v_k \mathbf{w}_k$, Taylor formula for the restriction $t \mapsto g(t) = f(a + t\mathbf{v})$ gives

$$g(t) = g(0) + \sum_i \lambda_k v_k^2 t^2 + \text{ higher order terms} .$$

There follows

**Proposition 16.1.** *Let $\mathbf{a} \in X$ be a critical point of a scalar field $f : X \subset \mathbb{R}^n \to \mathbb{R}$ of class $\mathcal{C}^2$. If all the eigenvalues of the Hessian matrix* $\text{Hess} f(\mathbf{a})$ *are positive/negative then $a$ is a local minimum/maximum of $f$. If the Hessian matrix has both positive and negative eigenvalues, then* $\mathbf{a}$ *is a saddle point.*

Observe also that if our scalar field is defined on the plane $\mathbb{R}^2$, then the Hessian matrix is two-by-two matrix, and the task to detect its signature is much easier. In this case we can state the recipe: a critical point of a scalar field $f(x,y)$ is a local extremum iff the determinant $\det(\text{Hess} f(\mathbf{a}))$ is positive; moreover, the local extremum is a maximum/minimum iff one of the diagonal entries of $\text{Hess} f(\mathbf{a})$ is negative/positive.

**ex:** Compute critical points of the following fields, and decide if they are maxima, minima or saddle points.

$$f(x,y) = (x-1)(y-2) \qquad f(x,y) = x^2 + (y-3)^2 \qquad f(x,y) = x^2 - y^2 + 7$$

$$f(x,y) = x^3 y^2 \qquad f(x,y = \sin(x)\cos(y) \qquad f(x,y) = e^{-x^2-y^2}$$

**e.g. Geometric center.** Given $N$ points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N \in \mathbb{R}^n$, we may try to minimize the sum

$$S(\mathbf{r}) = \sum_{k=1}^{N} \|\mathbf{x}_k - \mathbf{r}\|^2$$

of the square distances from a given point $\mathbf{r}$. The minimum is attained for $\mathbf{r}$ equal to the *(geometric) center*

$$\overline{\mathbf{x}} := \frac{1}{N} \sum_{k=1}^{N} \mathbf{x}_k \,.$$

**e.g. Least squares.** We measure $n$ times an observable $y$ in correspondence of another observable $x$, obtaining the set of data

$$x_1,\, y_1, \qquad x_2,\, y_2 \qquad \ldots \qquad x_n,\, y_n$$

We conjecture a law $y = f(x, \mathbf{a})$, depending on certain parameters $\mathbf{a} = (\alpha_1, \ldots, \alpha_k) \in \mathbb{R}^k$, and pose the problem to find the "best" values of the parameters that fit the experimental results. One popular answer, called *least square fitting*, consists in choosing those values of the parameters that minimize the sum

$$Q(\mathbf{a}) = \sum_{k=1}^{n} (y_k - f(x_k, \mathbf{a}))^2$$

of the squares of the errors. In general, the condition $\nabla Q(\mathbf{a}) = 0$ being nonlinear, cannot be solved by exact methods. Computational softwares, in particular statistics software, use to have routines dedicated to estimate a solution.

The answer is easy when we conjecture a linear law $y = \alpha + \beta x$, In this case, computing the partial derivatives $\partial Q/\partial\alpha$ and $\partial Q/\partial\beta$, we get the two equations

$$\sum_{k=1}^{n} (y_k - (\alpha + \beta x_k)) = 0 \qquad \text{and} \qquad \sum_{k=1}^{n} (y_k - (\alpha + \beta x_k))\, x_k = 0\,,$$

hence the system

$$\begin{cases} \beta\,\overline{x} + \alpha = \overline{y} \\ n\,\alpha + \beta\left(\overline{\sigma}^2_{xx} + n\overline{x}^2\right) = \overline{\sigma}^2_{xy} + n\overline{xy} \end{cases}$$

for $\alpha$ and $\beta$, where we used the notations $\overline{x} = \frac{1}{n}(x_1 + x_2 + \cdots + x_n)$ and $\overline{y} = \frac{1}{n}(y_1 + y_2 + \cdots + y_n)$ for the mean values, and

$$\overline{\sigma}^2_{xx} := \sum_{k=1}^{n} (x_k - \overline{x})^2 = \left(\sum_{k=1}^{n} x_k^2\right) - n\overline{x}^2$$

$$\overline{\sigma}^2_{xy} := \sum_{k=1}^{n} (x_k - \overline{x})(y_k - \overline{y}) = \left(\sum_{k=1}^{n} x_k y_k\right) - n\overline{xy}$$

for the covariances. After some rearrangement, we see that the critical point of $Q(\alpha, \beta)$, hence the answer according to the least squares principle, is given by the recipe

$$\beta = \frac{\overline{\sigma}^2_{xy}}{\overline{\sigma}^2_{xx}} \qquad \text{and} \qquad \alpha = \overline{y} - \beta\overline{x}\,.$$

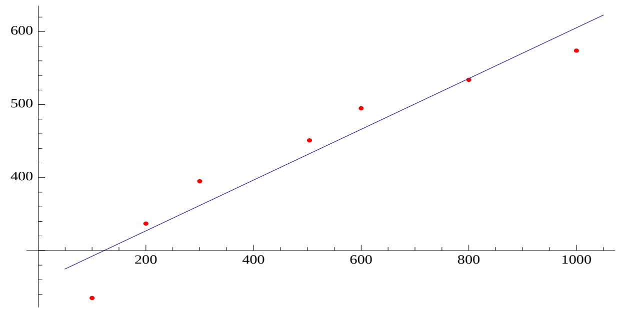
It must be said that minima of $Q(\mathbf{a})$ always exist, hence the method produces values of the parameters for all laws we may conjecture, true or false! The actual value of the minimum, together with some knowledge of the statistical errors in the data, gives a measure of the significance of the result. You may learn more in any good manual on statistics.

**help:** With Mathematica®8 you may define your data and fit a line with the commands

```
data = {{100, 235}, {200, 337}, {300, 395}, {504, 451}, {600,  495}, {800, 534}, {1000, 574}};
line = Fit[data, {1, x}, x]
```

and produce the picture

```
Show[ListPlot[data, PlotStyle -> Red], Plot[{line}, {x, 50, 1050}]]
```




**ex:** Na seguinte amostra, obtida por Galileo, foram registadas as coordenadas (altura $x$ e distância $y$) da trajectória de um objecto lançado com uma força horizontal,

| $x$ | 100 | 200 | 300 | 450 | 600 | 800 | 1000 |
|---|---|---|---|---|---|---|---|
| $y$ | 235 | 337 | 395 | 451 | 495 | 534 | 574 |

Ajuste uma recta.

**ex:** Na seguinte tabela, colecionada por Jaques Cassini, foram registadas as obliquidades da eclíptica (o ângulo entre o plano equatorial da Terra e o seu plano orbital) $(y + 23)^\circ$ em diferentes datas $t$,

| $t$ | $-140$ | $-140$ | 390 | 880 | 1070 | 1300 | 1460 |
|---|---|---|---|---|---|---|---|
| $y$ | 0.853 | 0.856 | 0.500 | 0.583 | 0.567 | 0.533 | 0.500 |

| $t$ | 1500 | 1500 | 1570 | 1570 | 1600 | 1656 | 1672 | 1738 |
|---|---|---|---|---|---|---|---|---|
| $y$ | 0.473 | 0.488 | 0.499 | 0.525 | 0.517 | 0.484 | 0.482 | 0.472 |

Ajuste uma recta. Retire os dados anteriores ao ano 1500, e ajuste outra recta. Discuta o resultado.

# 17 Continuous-time models and simulations

**Systems of ordinary differential equations.** Meaningful models of many physical, chemical, biological ... systems are written in the language of systems of differential equations

$$\dot{\mathbf{x}} = \mathbf{v}(t, \mathbf{x}) \tag{17.1}$$

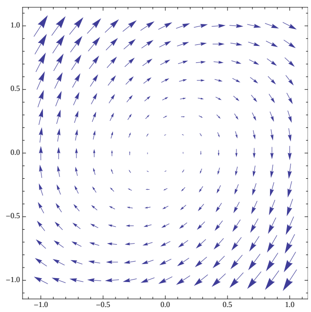
where $\mathbf{x}(t) = (x_1(t), x_2(t), \dots, x_n(t)) \in X \subset \mathbb{R}^n$ is a vector of values of certain observables at time $t$, and $\mathbf{v}(t, \mathbf{x})$ is a given direction field in the extended phase space $T \times X \subset \mathbb{R} \times \mathbb{R}^n$.

e.g. **Chemical reactions.** The modern approach to the kinetics of chemical reactions is discussed in the article Chemical reaction kinetics of the Scholarpedia.

**help:** O campo vetorial do oscilador harmónico com atrito pode ser desenhado, no Mathematica®, usando a instrução

```
VectorPlot[{y, -x + 0.5 y}, {x, -1, 1}, {y, -1, 1}]
```

O resultado é



**Simulations.** It is in general hopeless to find "exact" solutions of systems of differential equations, as long as they are not linear. For this reason, we must content with making simulations.

**Euler method.** Considere o problema de simular as soluções da EDO

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

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

$$\mathbf{x}(t + dt) - \mathbf{x}(t) \simeq \mathbf{v}(t, \mathbf{x}) \cdot dt\,,$$

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

$$x_{n+1} = x_n + v(t_n, x_n) \cdot dt\,, \tag{17.2}$$

onde $t_n = t_0 + n \cdot dt$. 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 ;
  }
```

**e.g. The exponential.** Considere a equação diferencial

$$\dot{x} = x$$

com condição inicial $x(0) = 1$. Mostre que, se o passo é $dt = \varepsilon$, então o método de Euler fornece a aproximação

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

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

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

**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

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

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

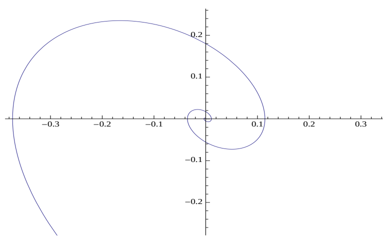
$$\boxed{k_1 = v(t_n, x_n) \quad k_2 = v\left(t_n + \tfrac{dt}{2}, x_n + \tfrac{dt}{2} \cdot k_1\right) \quad k_3 = v\left(t_n + \tfrac{dt}{2}, x_n + \tfrac{dt}{2} \cdot k_2\right) \quad k_4 = v(t_n + dt, x_n + dt \cdot k_3)}$$

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

**help:** 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



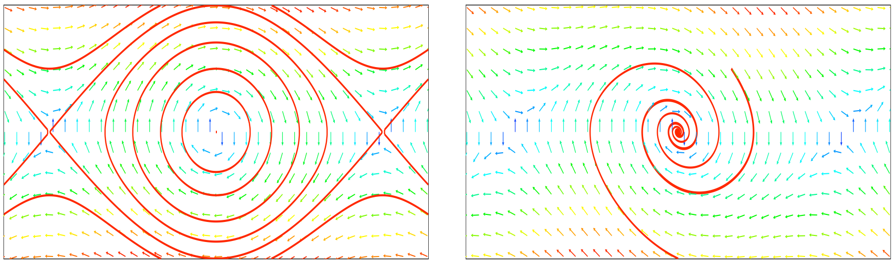

**e.g. Pêndulo matemático.** Considere a equação de Newton que modela as oscilações de um *pêndulo*,

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

onde $\omega = \sqrt{g/\ell}$, $g$ é a aceleração gravitacional, $\ell$ o comprimento do pêndulo, e $\alpha \geqslant 0$ um coeficiente de atrito. No espaço de fase, de coordenadas $\theta$ e $p = \dot{\theta}$, a equação assume a forma do sistema

$$\begin{aligned} \dot{\theta} &= p \\ \dot{p} &= -\omega^2 \sin(\theta) - \alpha p \end{aligned}$$

- Simule o sistema, e esboçe as trajectórias e as curvas de fase.



Retrato de fase do pêndulo (sem e com atrito).

**e.g.** **Oscilador harmónico.** As pequenas oscilações de um pêndulo em torno da posição de equilíbrio estável $\theta = 0$ são descrita pela equação do *oscilador harmónico*

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

onde $\omega$ é a frequência característica. No espaço de fase, de coordenadas $q$ e $p = \dot{q}$, a equação assume a forma do sistema

$$\begin{aligned} \dot{q} &= p \\ \dot{p} &= -\omega^2 q \end{aligned}$$

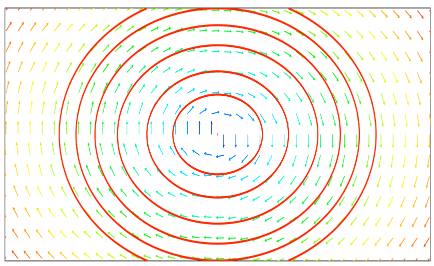
As tsoluç oes são

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

onde a amplitude $A$ e as fases $\varphi$ e $\phi$ dependem dos dados iniciais $q(0) = q_0$ e $\dot{q}(0) = v_0$. A energia

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

é uma constante do movimento, ou seja, $\frac{d}{dt} E\left(q(t), p(t)\right) = 0$.



Retrato de fase do oscilador harmónico.

**e.g.** **Circuito LRC.** 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.

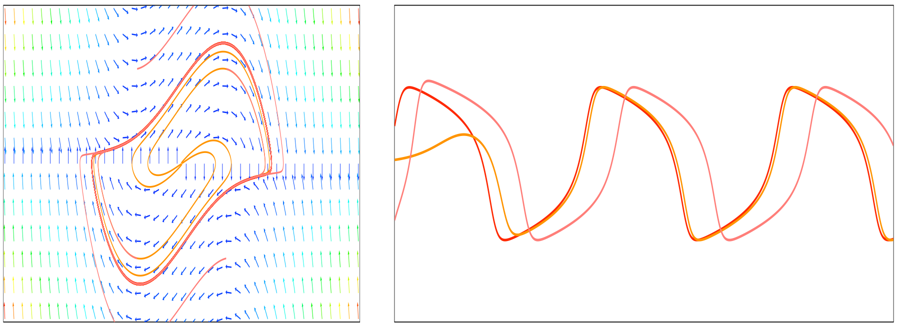
- Simule a corrente num circuito alimentado com uma tensão constante $V(t) = V_0$.
- Simule a corrente num circuito alimentado com uma tensão alternada $V(t) = V_0 \sin(\gamma t)$ (compare com a equação das oscilações forçadas amortecidas).

**e.g. Oscilador de van der Pol.** Considere o *oscilador de van der Pol*[33]

$$\ddot{q} - \mu(1 - q^2)\dot{q} + q = 0$$

que modela a corrente num circuito com um elemento não-linear.

- Simule o sistema e discuta o comportamento das soluções ao variar o parâmetro $\mu$.



Retrato de fase e trajectórias do oscilador de van der Pol.

- Simule o oscilador forçado

$$\ddot{q} - \mu(1 - q^2)\dot{q} + q = F_0 \sin(\omega t)$$

ao variar o parâmetro $\mu$ e a frequência $\omega$.

**e.g. Sistema de Lotka-Volterra.** Considere o *sistema de Lotka-Volterra*

$$\dot{x} = ax - bxy$$
$$\dot{y} = -cy + dxy$$

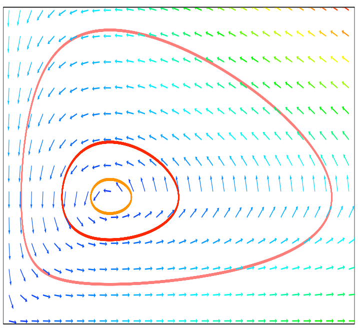
Foi proposto por Vito Volterra[34] para modelar a competição entre $x$ presas e $y$ predadores, e por Alfred J. Lotka[35] para modelar o comportamento cíclico de certas reacções químicas, como o esquema abstracto

$$A + X \to 2X \qquad X + Y \to 2Y \qquad Y \to B$$

Stationary solutions are found solving the system $\dot{x} = 0$ and $\dot{y} = 0$. This gives the trivial solution $(0, 0)$, and the point $(c/d, a/b)$. To understand the other solutions, one observes that the function

$$H(x, y) = dx + by - c \log x - a \log y$$

is a constant of the motion, i.e. $\frac{d}{dt} H(x(t), y(t)) = 0$. Therefore, orbits of the Lotka-Volterra system are contained in the level curves $H(x, y) = c$.



[33] B. van der Pol, A theory of the amplitude of free and forced triode vibrations, *Radio Review* **1** (1920), 701-710 and 754-762. B. van der Pol and J. van der Mark, Frequency demultiplication, *Nature* **120** (1927), 363-364.

[34] Vito Volterra, Variazioni e fluttuazioni del numero d'individui in specie di animali conviventi, *Mem. Acad. Lincei* **2** (1926), 31-113. Vito Volterra, *Leçons sur la Théorie Mathématique de la Lutte pour la Vie*, Paris 1931.

[35] Alfred J. Lotka, *J. Amer. Chem. Soc* **27** (1920), 1595. Alfred J. Lotka, *Elements of physical biology*, Williams & Wilkins Co. 1925.

Phase portrait of the Lotka-Volterra system.

e.g. **Rock-paper-scissor game.** Consider the reaction

$$X + Y \xrightarrow{\gamma} 2X \qquad Y + Z \xrightarrow{\alpha} 2Y \qquad Z + X \xrightarrow{\beta} 2Z$$

modeled by the system

$$\begin{aligned} \dot{x} &= x(\gamma y - \beta z) \\ \dot{y} &= y(\alpha z - \gamma x) \\ \dot{z} &= z(\beta x - \alpha y) \end{aligned}$$

e.g. **Double-negative feedback.** The interplay between two mutually repressing genes is described by the system[36]

$$\begin{aligned} \dot{x} &= \frac{\alpha}{1+y^\gamma} - x \\ \dot{y} &= \frac{\beta}{1+x^\delta} - y \end{aligned}$$

e.g. **Brusselator.** O *Brusselator* é um modelo autocatalítico proposto por Ilya Prigogine e colaboradores[37] que consiste na reacção abstracta

$$A \to X \qquad B + X \to Y + C \qquad 2X + Y \to 3X \qquad X \to D$$

- Simule o sistema

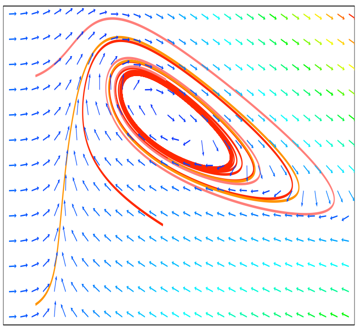
  $$\begin{aligned} \dot{x} &= \alpha - (\beta + 1)x + x^2 y \\ \dot{y} &= \beta x - x^2 y \end{aligned}$$

  para as concentrações das espécies catalíticas $X$ e $Y$, obtido quando as concentrações $[A] \sim \alpha$ e $[B] \sim \beta$ são mantidas constantes.

- Simule o sistema

  $$\begin{aligned} \dot{x} &= \alpha - (b + 1)x + x^2 y \\ \dot{y} &= bx - x^2 y \\ \dot{b} &= -bx + \delta \end{aligned}$$

  para as concentrações de $X$, $Y$ e $B$, obtido quando a concentração $[A] \sim \alpha$ é mantida constante e $B$ é injectado a uma velocidade constante $v \sim \delta$.



Rerato de fase do Brusselator.

[36] T.S. Gardner, C.R. Cantor and J.J. Collins, Construction of a genetic toggle switch in *Escherichia coli*, *Nature* **403** (2000) 339-342.

[37] I. Prigogine and R. Lefever, Symmetry breaking instabilities in dissipative systems, *J. Chem. Phys.* **48** (1968), 1655-1700. P. Glansdorff and I. Prigogine, *Thermodynamic theory of structure, stability and fluctuations*, Wiley, New York 1971. G. Nicolis and I. Prigogine, *Self-organization in non-equilibrium chemical systems*, Wiley, New York 1977.

**e.g. Reacção de Schnakenberg.** Considere a *reacção de Schnakenberg*[38]

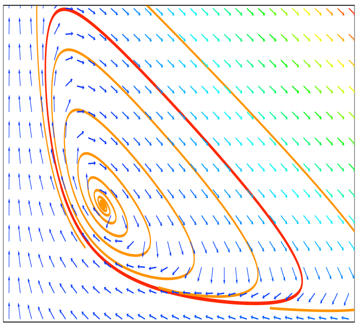
$$2X + Y \to 3X \qquad A \to Y \qquad X \to B$$

modelada pelo sistema

$$\begin{aligned} \dot{x} &= x^2 y - x + \beta \\ \dot{y} &= -x^2 y + \alpha \end{aligned}$$

para as concentrações $x \sim [X]$ e $y \sim [Y]$.

- Simule o sistema e discuta o comportamento das soluções ao variar so parâmetros.



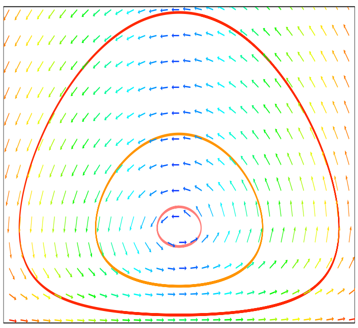
Retrato de fase do sistema de Schnakenberg.

**e.g. Oscilador bioquímico de Goodwin.** Um modelo de interações proteinas-mRNA proposto por Goodwin[39] é

$$\begin{aligned} \dot{M} &= \frac{1}{1+P} - \alpha \\ \dot{P} &= M - \beta \end{aligned}$$

onde $M$ e $P$ denotam as concentrações relativas de mRNA e proteina, respectivamente.

- Simule o sistema e discuta o comportamento das soluções ao variar so parâmetros.



Retrato de fase do sistema de Goodwin.

- Simule o sistema[40]

$$\begin{aligned} \dot{M} &= \frac{1}{1+P^n} - \alpha M \\ \dot{P} &= M^m - \beta P \end{aligned}$$

[38] J. Schnakenberg, Simple chemical reaction with limit cycle behavior, *J. Theor. Biol.* **81** (1979), 389-400.

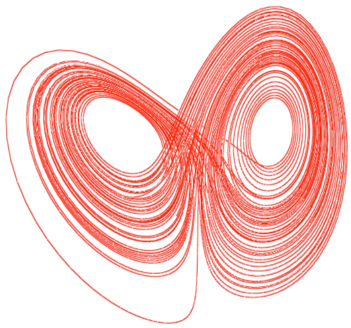
[39] B.C. Goodwin, *Temporal organization in cells*, Academic Press, London/New York 1963. B.C. Goodwin, Oscillatory behaviour in enzymatic control processes, *Adv. Enzyme Regul.* **3** (1965), 425-438.

[40] T. Scheper, D. Klinkenberg, C. Pennartz and J. van Pelt, A Mathematical Model for the Intracellular Cicardian Rhythm Generator, *J. Neuroscience* **19** (1999), 40-47.

**e.g.** **Atrator de Lorenz.** Considere o *sistema de Lorenz*[41]

$$
\begin{aligned}
\dot{x} &= \sigma(y - x) \\
\dot{y} &= x(\rho - z) - y \\
\dot{z} &= xy - \beta z
\end{aligned}
$$

- Analize o comportamento assimptótico das trajectórias ao variar os parâmetros $\sigma$, $\rho$ e $\beta$.
- Observe o comportamento das trajectórias quando $\sigma \simeq 10$, $\rho \simeq 28$ e $\beta \simeq 8/3$.



Atractor de Lorenz.

[41] E.N. Lorenz, Deterministic nonperiodic flow, *J. Atmspheric Science* **20** (1963), 130-141.

# References


[Ap69] T.M. Apostol, *Calculus*, John Wiley & Sons, 1969 [*Cálculo*, Editora Reverté, 1999].

[Ar78] V.I. Arnold, *Equações diferenciais ordinárias*, MIR, 1985.

[Be62] C. Kittel, W.D. Knight and M.A. Ruderman, *Berkeley Physics: Mechanics*, McGraw-Hill 1962.

[Ba79] E. Batschelet, *Introduction to Mathematics for Life Scientists*, Springer, 1979.

[Co04] V. Comincioli, *Modelli Matematici: elementi introduttivi*, Università degli studi di Pavia, 2004.

[CR48] R. Courant and H. Robbins, *What is mathematics?*, Oxford University Press, 1948. [*O que é Matemática?*, Editora Ciência Moderna, 2000].

[Ei88] M. Eisen, *Mathematical Methods and Models in the Biological Sciences*, Prentice Hall, 1988.

[EK05] L. Edelstein-Keshet, *Mathematical Models in Biology*, SIAM, 2005 [Random House, 1988].

[Go77] H.J. Gold, *Mathematical Modeling of Biological Systems - An Introduction Guidebook*, Wiley Interscinece, 1977.

[Gou77] S.J. Gould, *Ever Since Darwin: Reflections in Natural History*, Norton, New York, 1977.

[HK03] B. Hasselblatt and A. Katok, *A first course in dynamics, with a panorama of recent developments*, Cambridge University Press, 2003.

[HS74] M.W. Hirsch and S. Smale, *Differential equations, dynamical systems and linear algebra*, Academic Press, 1974.

[Kl09] F. Klein, *Matemática elementar de um ponto de vista superior*, Sociedade Portuguesa de Matemática, 2009. [*Elementarmathematik von höheren Standpunkte aus*, Teubner, 1908]

[KR85] D. Kleppner and N. Ramsey, *Quick calculus: a self-teaching guide*, Wiley 1985.

[La86] S. Lang, *A First Course in Calculus*, UTM Springer, 1986.

[Li06] E.L. Lima, *Análise Real*, Coleção Matemática Universitária, IMPA, 2006.

[MF05] Philip McCord Morse and Herman Feshbach, *Methods of Theoretical Physics*, McGraw-Hill 1953 or Feshbach Publishing, 2005.

[RHB06] K.F. Riley, M.P. Hobson and S. J. Bence, *Mathematical Methods for Physics and Engineering*, Cambridge University Press, 2006.

[Ri04] S.H. Rice, *Evolutionary Theory: Mathematical and Conceptual Foundations*, Sinauer, 2004.

[Ro04] J.C. Robinson, *An introduction to ordinary differential equations*, Cambridge University Press, 2004.

[Ros84] S.L. Ross, *Differential equations*, John Wiley & Sons, 1984.

[Si91] G.F. Simmons, *Differential equations with applications and historical notes*, McGraw-Hill, 1991.

[Wae91] B.L. van der Waerden, *Algebra*, Springer 1991.

[Wr69] S. Wright, *Evolution and the Genetic of Populations*, The University of Chicago Press, 1969.