INFINITESIMAL CALCULUS 2

COURSE NOTES

(with self-assesment exercises)

Jaime Fe Marqués
ETSI Caminos - A Coruña

Purpose of these notes

The aim of this text is to be an aid for the students of the subject Infinitesimal Calculus 2.

These notes contain the fundamental ideas presented in the theory classes, accompanied by solved examples and proposed exercises. By collecting an important part of the content of the sessions, they allow the students to dedicate more attention to the explanation, thus facilitating the understanding of the subject. The solved examples and proposed exercises also facilitate the subsequent personal study.

The notes are intended as a support to the lectures, but not as a substitute for them. In the classroom, the ideas contained herein are developed and commented, concepts are related, exercises are solved and some graphs are drawn to complement the notes. Therefore, class attendance is strongly recommended to facilitate the mastery of the subject. As far as possible, it is desirable to read the notes before the classes, which will help to know in advance the main aspects of the subject to be covered.

Contents

Unit I. Integration (01.02.2024)

1 Primitive of a function

1.1 Definition

Let f be a function defined on $J = [a, b]$. We say that function $F : J \to \mathbb{R}$ is a primitive of f on J if F is differentiable on J and f is its derivative, that is $F' = f$. In this case we write

$$F(x) = \int f(x)dx$$

If F is a primitive of f, functions $F + K$, $\forall K \in \mathbb{R}$ (and only them) are primitives of f. Indeed:

a) If F is a primitive of f

$$(F + K)' = F' = f$$

then $F + K$ is also a primitive of f.

b) If G is a primitive of f, that is $G' = f$, then

$$(G - F)' = G' - F' = 0 \Longrightarrow G - F = K \Longrightarrow G = F + K$$

hence G is $F + K$.

1.2 Primitives of discontinuous functions

We will prove later that every continuous function has a primitive. But there are also discontinuous functions that have a primitive. For example, consider the function

$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

Outside the origin the derivative of f is $f'(x) = 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x}$. At the origin we obtain the derivative by the definition, obtaining

$$f'(0) = \lim_{x \to 0} \frac{x^2 \sin \dfrac{1}{x} - 0}{x - 0} = \lim_{x \to 0} x \sin \frac{1}{x} = 0$$

then f is also differentiable at $x = 0$. However, the derivative function

$$f'(x) = \begin{cases} 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

has no limit when $x \to 0$. It means that f' is discontinuous (at $x = 0$), but has a primitive f, $\forall x$. Then there are discontinuous functions that have a primitive.

Exercise. Study whether the above can be applied to the function

$$f(x) = \begin{cases} x^3 \cos \dfrac{1}{x^2}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

1.3 Necessary condition for the existence of primitive

As we have seen, there are discontinuous functions with primitive, but it does not mean that every discontinuous function has it. We will see below that *not having jump discontinuities on J is a necessary condition for a function to have a primitive on J.*

Proof. We are going to consider a differentiable function f. This function will therefore be a primitive of its derivative f'. We study the continuity of f'.

First we recall what was studied in "The derivative as a limit of derivatives" (Infinitesimal Calculus 1, unit IV, section D.5):

Let function f be defined on $J = [a, a + \delta)$. If f is continuous on J, differentiable on $J \setminus \{a\}$ and $\exists \lim_{x \to a^+} f'(x)$, then f is differentiable at a^+ and it holds

$$f'(a^+) = \lim_{x \to a^+} f'(x)$$

(and the corresponding statement for a^-).

That is, if function f' has a one-sided limit at a^+ (or a^-), the corresponding one-sided derivative of f will have the value of this limit.

Therefore, if f is differentiable on J, at each point of J the derivative will exist, so it will coincide with the one-sided derivatives at that point

$$\forall x_0 \in J \; \exists \, f'(x_0) = f'(x_0^+) = f'(x_0^-)$$

and there are two options:

a) $\forall x_0 \in J$ there exist the one-sided limits of $f'(x)$, which must coincide with the one-sided derivatives, so they will coincide with each other and the function f' is continuous on J.

b) At some point of J one or both one-sided limits do not exist, so f' has an essential discontinuity at that point.

But there cannot exist two different one-sided limits since they would be equal to the respective one-sided derivatives, which coincide.

Thus, if a function is differentiable, its derivative may be discontinuous at some point, because one or both one-sided limit do not exist; but it cannot have jump discontinuities (different one-sided limits). Therefore, **if a function has a jump discontinuity** at a point $a \in J$, **it has no primitive** on J.

Example. The continuous function

$$F(x) = \begin{cases} 0, & 0 \le x < 1 \\ x - 1, & 1 \le x < 2 \\ 2x - 3, & 2 \le x < 3 \end{cases}$$

is primitive of the floor function $y = \lfloor x \rfloor$ on the intervals (0,1),(1,2) and (2,3). But we have just seen that a function cannot have a primitive at the points where it has a jump discontinuity.

Indeed, function F is not differentiable at $x = 1$ and $x = 2$ (the points of discontinuity of the floor function), so at these points $y = \lfloor x \rfloor$ has no primitive, which agrees with the necessary condition.

Exercise 1. Given f, differentiable, what can we say about the continuity of f? And of f'?

Exercise 2. Consider the sign function defined on $[0, 3]$. Find its primitive on the intervals on which it exists.

2 The Riemann integral

2.1 Partition of an interval

Given an interval $J = [a, b]$, we divide it into n parts (subintervals) of equal or different size, inserting $n - 1$ points between a and b. A partition is the set of $n + 1$ points:

$$P(x_0, x_1 \dots x_n) \big/ a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

Relation between partitions. Let P and P' be two partitions of J. If every point of P belongs to P', then we say that P' is finer than P or that P is contained in P' ($P \subset P'$).

The **norm** (or diameter) of a partition P is the length of the largest of the subintervals into which J is divided:

$$\delta(P) = \max(x_i - x_{i-1}), \quad i = 1, 2, \dots n$$

The length of each subinterval is usually denoted as $\Delta x_i = x_i - x_{i-1}$.

Exercise. Prove that the relation $\subset$ between partitions of J is a partial order.

2.2 Darboux sums

Let f be a bounded function on $J = [a, b]$ and P a partition of J. On each subinterval $[x_{i-1}, x_i]$, let M_i and m_i be respectively the supremum and the infimum of the values of f.

$$M_i = \sup f(x), \quad m_i = \inf f(x), \quad x \in [x_{i-1}, x_i]_{i=1,2,\dots n}$$

Let M and m be respectively the supremum and the infimum of the values of f on J.

$$M = \sup f(x), \quad m = \inf f(x), \quad x \in [a, b]$$

Obviously the relation between the supremum and the infimum values of f on any subinterval and on the interval J is

$$m \le m_i \le M_i \le M, \quad i = 1, 2, \dots, n$$

where, multiplying each term by Δx_i and adding for $i = 1, 2, \dots, n$, the inequality becomes

$$\sum_{i=1}^{n} m\Delta x_i \le \sum_{i=1}^{n} m_i\Delta x_i \le \sum_{i=1}^{n} M_i\Delta x_i \le \sum_{i=1}^{n} M\Delta x_i \tag{1}$$

We now take as a common factor m in the first term and M in the fourth one; adding the values Δx_i in both terms, we obtain respectively $m(b - a)$ and $M(b - a)$.

The second and third terms are the sums of the products of the lengths of the subintervals by the infimum and supremum values of f in them. They are called respectively **lower and upper Darboux sum** of function f on the interval J with respect to the partition P.

$$\boxed{s(P) = \sum_{i=1}^{n} m_i\Delta x_i} \;\; ; \;\; \boxed{S(P) = \sum_{i=1}^{n} M_i\Delta x_i}$$

And the inequality (1) turns to the following expression, valid for any bounded function f and any partition P (the dependency on f is omitted for simplicity).

$$\boxed{m(b - a) \le s(P) \le S(P) \le M(b - a)} \tag{2}$$

Properties. The Darboux sums satisfy the following properties (the first is immediate and the third is proved):

a) S and s are bounded $\big($upper bound: $M(b-a)$, lower bound $m(b-a)\big)$.

b) Taking a partition P' finer than P, the upper sum decreases and the lower sum increases.

$$P \subset P' \Longrightarrow S(P) \geq S(P'), \quad s(P) \leq s(P')$$

c) Any upper sum is greater than or equal to any lower sum.

$$\forall P_1, P_2 \Longrightarrow S(P_1) \geq s(P_2)$$

P: We define $P = P_1 \cup P_2$, so $P_1 \subset P$ and $P_2 \subset P$. Applying relation (2) between $s(P)$ and $S(P)$ and property **b)** between sums corresponding to different partitions, it results:

$$S(P_1) \geq S(P) \geq s(P) \geq s(P_2) \Longrightarrow S(P_1) \geq s(P_2), \quad \forall P_1, P_2$$

In figure 1, the areas of the rectangles over and under the curve represent the upper and lower Darboux sums respectively.

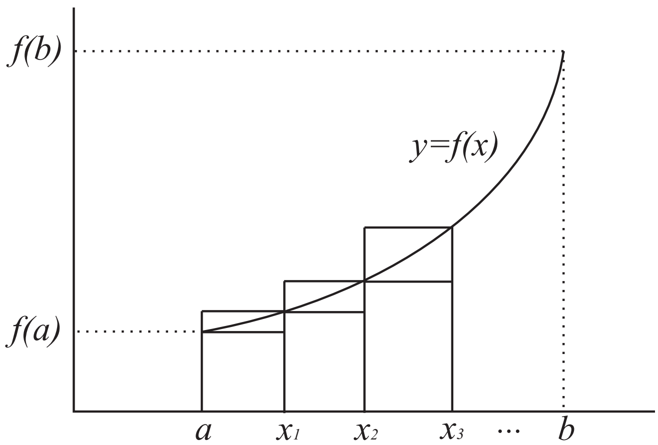


Figure 1: Darboux sums.

2.3 Riemann integrable function

We now take finer and finer partitions $P_1, P_2, \ldots, P_m$, so that $\delta(P_m) \to 0$ when $m \to \infty$. Thus we obtain two sequences of Darboux sums $\{S_m\}$ and $\{s_m\}$ that, applying property **b)** satisfy the following:

- The $\{S_m\}$ form a monotone decreasing sequence bounded from below, so they converge to a limit S, which is called the upper Darboux integral of f on J.

- The $\{s_m\}$ form a monotone increasing sequence bounded from above, so they converge to a limit s, which is called the lower Darboux integral of f on J.

Notice that m is the index of the different partitions and the corresponding Darboux sums and it does not represent the number of points of the partitions. But, if $m \to \infty$, the number of points also tends to ∞.

We say that f is Riemann integrable on J if and only if $S = s$. In this case, this common value is called the Riemann definite integral of f on $[a, b]$:

$$\boxed{I = \int_a^b f(x)dx}$$

From the process followed, we observe that, if f has a constant sign between a and b, the area between the curve $y = f(x)$, the lines $x = a$ and $x = b$ and the x axis is:

$$A_a^b = I \text{ (if } f > 0); \quad A_a^b = -I \text{ (if } f < 0)$$

If f does not have a constant sign, the integral gives us the value of the **net area**, where the parts above the x axis have a positive sign and those below have a negative sign.

Exercise. **a)** Prove that $y = \lfloor x \rfloor$ (floor function) is integrable on $[5/4, 7/4]$. **b)** Prove that the Dirichlet function is not integrable.

2.4 Necessary and sufficient condition of integrability

The stated integrability condition consists in the fact that, when taking increasingly finer partitions, the limits of S_m and s_m coincide, so the limit of their difference is null.

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} s_n = I \Longrightarrow \lim_{n\to\infty} (S_n - s_n) = 0$$

We can express the above by saying that the difference between the respective terms of both sequences can be made as small as we want, if the partition is sufficiently fine. That is

$$\forall \varepsilon > 0 \; \exists P \,\big/\, S(P) - s(P) < \varepsilon$$

(for any partition finer than P, the same condition will hold).
Taking into account the expressions of both sums, seen in **2.2**,

$$s(P) = \sum_{i=1}^{n} m_i \Delta x_i \; ; \; S(P) = \sum_{i=1}^{n} M_i \Delta x_i$$

we arrive at the condition we were looking for: *It is a necessary and sufficient condition for a function f, bounded on $[a, b]$, to be integrable on $[a, b]$ that*

$$\forall \varepsilon > 0, \, \exists P \,\Big/ \boxed{S(P) - s(P) = \sum_{i=1}^{n} (M_i - m_i)\Delta x_i < \varepsilon}$$

being $M_i = \sup f(x), \; m_i = \inf f(x), \; x \in [x_{i-1}, x_i]_{i=1,2,\dots n}$

2.5 Three sufficient conditions of integrability

Three particular cases in which the previous condition is satisfied are the following (the proofs can be seen in a supplementary document):

a) Every monotone function on $[a, b]$ is integrable on $[a, b]$. Example: the floor function.

b) Every continuous function on $[a, b]$ is integrable on $[a, b]$. Example: the sine function.

c) Every piecewise continuous function on $[a, b]$ is integrable on $[a, b]$. Example: the decimal part function.

(f is piecewise continuous on $[a, b]$ if it has a finite number of discontinuity points on $[a, b]$ and the one-sided limits exist and are finite at them).

Exercise. Find examples of integrable functions on $[a, b] \in \mathbb{R}$: a) continuous, not monotone; b) monotone, not continuous; c) continuous and monotone; d) neither continuous nor monotone.

2.6 Properties of the Riemann integral

a) Let f be integrable on $[a,b]$, $a<b$. We define:

$$\boxed{\int_b^a f(x)dx = -\int_a^b f(x)dx} \text{ from where } \boxed{\int_a^a f(x)dx = 0}$$

b) Let f be integrable on $[a,c]$ and on $[c,b]$. Then f is integrable on $[a,b]$ and its integral is the sum of the integrals.

$$\boxed{\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx}$$

Additivity with respect to the interval also holds when c is not between a and b. Indeed, let f be integrable on $[\alpha,\beta]$ and let $a,b,c \in [\alpha,\beta] \,/\, a<b<c$. Then:

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx \overset{\text{a)}}{=} \int_a^b - \int_c^b \Longrightarrow \int_a^b = \int_a^c + \int_c^b$$

c) Let f,g be integrables on $[a,b]$. It is satisfied:

- The linear combination of f and g is integrable on $[a,b]$. Its integral is the linear combination of the integrals.

$$\boxed{\int_a^b \big(\lambda f(x) + \mu g(x)\big)dx = \lambda \int_a^b f(x)dx + \mu \int_a^b g(x)dx} \quad (\lambda,\mu \in \mathbb{R})$$

- The product $\boxed{f \cdot g}$ is integrable on $[a,b]$.

- If $g(x) \neq 0$, $\forall x \in [a,b]$, then the quotient $\boxed{f/g}$ is integrable on $[a,b]$.

- If $f(x) \leq g(x)$, $\forall x \in [a,b]$, then

$$\boxed{\int_a^b f(x)dx \leq \int_a^b g(x)dx}$$

d) If f is integrable on $[a,b]$, its absolute value $|f|$ is also integrable and

$$\boxed{\left|\int_a^b f(x)dx\right| \leq \int_a^b |f(x)|\,dx}$$

3 Intermediate value theorem

Let f be a function continuous on $[a,b]$. It holds that:

$$\exists\, \xi \in [a,b] \,/\, \boxed{f(\xi) = \frac{1}{b-a}\int_a^b f(x)dx}$$

Graphic interpretation. There is a value of f which, multiplied by the length of the interval, gives us the net area under the curve (in fig. 2.a, the intermediate value is represented by t).

Proof. Since f is continuous on $J=[a,b]$, it follows that: 1) it reaches a maximum M and a minimum m on J (Weierstrass) and 2) it is integrable on J (sufficient condition of integrability). Then it exists the integral of f on $[a,b]$, which is bounded between the values $M(b-a)$ and $m(b-a)$ (see **2.2**). Therefore

$$m(b-a) \leq \int_a^b f(x)dx \leq M(b-a) \Longrightarrow m \leq \frac{1}{(b-a)}\int_a^b f(x)dx \leq M$$

The central term of the inequality is between two values (m and M) which are taken by the function. From Darboux property we know that, if f (continuous) reaches two values, it takes any value between them, that is

$$\exists\, \xi \in [a,b] \,/\, f(\xi) = \frac{1}{b-a}\int_a^b f(x)dx$$

Example. Let $f(x)=x^2-1$, $x\in[1,3]$. Using the Barrow rule (**4.3**), we obtain the value ξ:

$$f(\xi)=\xi^2-1=\frac{1}{3-1}\int_1^3 (x^2-1)dx = \frac{1}{2}\left(\frac{x^3}{3}-x\right)_1^3 = \frac{10}{3} \Longrightarrow \xi=\sqrt{13/3}$$

Exercise. If $f(x)=2x+1$ and $[a,b]=[1,3]$, find the intermediate value ξ and give a graphic interpretation of the result.

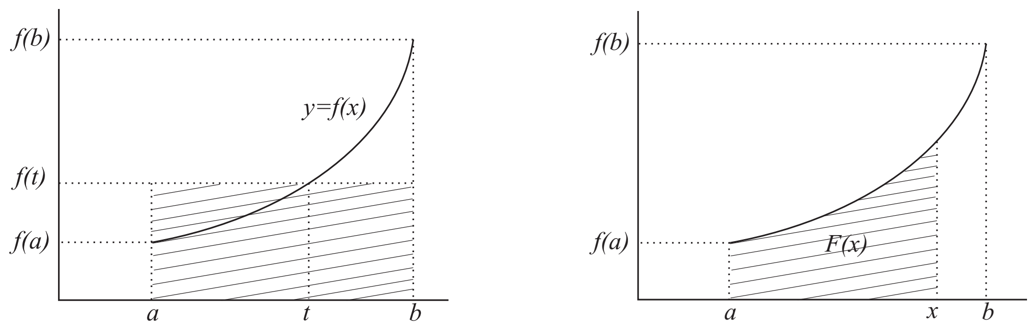


Figure 2: a) Interm. value theorem. b) Integral function.

4 First fundamental theorem of calculus.

4.1 The integral function

Let f be a function integrable on $[a,b]$. As we have seen, the net area bounded by the curve and the x-axis between two points is obtained by calculating the integral of f between those points. On the other hand, the area between the left endpoint a of the interval and any point x, will be a function of x. Then we define the integral function

$$F(x)=\int_a^x f(t)dt$$

which gives us the value of the net area from the endpoint a to the point x (fig. 2.b).

Continuity. The integral function is continuous on $[a,b]$. To prove it, let us see that it satisfies the continuity condition at any point x_0 of the interval.

$$F(x)=\int_a^x f(t)dt=\int_a^{x_0} f(t)dt+\int_{x_0}^x f(t)dt = F(x_0)+\int_{x_0}^x f(t)dt$$

The second addend is bounded since, from (**2.2**), we have

$$m(x-x_0) \leq \int_{x_0}^{x} f(t)dt \leq M(x-x_0), \quad m = \inf(f),\, M = \sup(f),\, t \in [a,b]$$

If, in addition, $x \to x_0$, both bounds tend to 0, so the integral also tends to 0. Then, taking limits, it turns out

$$\lim_{x\to x_0} F(x) = \lim_{x\to x_0} \left(F(x_0) + \int_{x_0}^{x} f(t)dt \right) = F(x_0), \quad \forall x_0 \in [a,b]$$

so $F(x)$ is continuous on $[a,b]$.

4.2 Theorem

If f is continuous on $[a,b]$, then F is differentiable on $[a,b]$ and it holds that

$$F' = f$$

hence the integral function F of a continuous function f is a primitive of f.

Proof: To find $F'(x)$, we first obtain ΔF.

$$F(x+h) - F(x) = \int_a^{x+h} f(t)dt - \int_a^x f(t)dt = \int_a^{x+h} f(t)dt + \int_x^a f(t)dt = \int_x^{x+h} f(t)dt$$

Dividing by h and applying the intermediate value theorem,

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t)dt = f(\xi), \quad \xi \in [x, x+h]$$

If we now make h tend to 0, then $x+h$ will tend to x. Since ξ is between x and $x+h$, it will also tend to x. But, since f is continuous, if $\xi \to x$, then $f(\xi) \to f(x)$.
Thus, taking limits, it turns out

$$\lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = \lim_{\xi\to x} f(\xi) = f(x) \Longrightarrow \boxed{F'(x) = f(x)}$$

As a consequence of this theorem, every continuous function has a primitive. But this primitive does not necessarily have to be formed by a finite number of terms: it can be a series of powers, as we will see in unit IV.

4.3 Corollary. Barrow rule

If f is continuous on $[a,b]$ and g is a primitive of f, then

$$\boxed{\int_a^b f(x)dx = g(b) - g(a)}$$

Proof: Since f is continuous, its integral function F is its primitive. By hypothesis, g is also a primitive of f. Then, the difference of both will be a constant, so

$$F(x) - g(x) = K \Longrightarrow F(b) - g(b) = F(a) - g(a) = 0 - g(a)$$

as $F(a) = \displaystyle\int_a^a f(x)dx = 0$. Therefore

$$F(b) = \int_a^b f(x)dx = g(b) - g(a)$$

Example. We find the integral function of $f(x) = \cos x$ on $[0.2\pi]$ and calculate the area bounded by the curve and OX between $x_1 = \pi/2$ and $x_2 = 3\pi/2$.

$$F(x) = \int_0^x \cos t\, dt = \sin t\Big|_0^x = \sin x - \sin 0 = \sin x$$

$$\int_{\frac{\pi}{2}}^{3\frac{\pi}{2}} f(x)dx = \sin 3\frac{\pi}{2} - \sin\frac{\pi}{2} = -1 - 1 = -2 \Longrightarrow A\Big|_{\frac{\pi}{2}}^{3\frac{\pi}{2}} = |-2| = 2$$

Since the cosine is negative between $\pi/2$ and $3\pi/2$, the area is the absolute value of the integral. We see that the net area between two points is the difference between the values of $F(x)$ at both points. This was to be expected, since $\displaystyle\int_b^c f(x)dx = \int_0^c f(x)dx - \int_0^b f(x)dx = F(c) - F(b)$.

Exercise. If $I = [0, \pi/4]$, find the integral function and the area between the endpoints:
a) $f_1(x) = 3x^2$ $\left(F_1(x) = x^3,\ A_1 = \pi^3/64\right)$; **b)** $f_2(x) = \tan x$ $\left(F_2(x) = -\ln\cos x,\ A_2 = \ln\sqrt{2}\right)$.

5 Second fundamental theorem of Calculus

If f is integrable on $[a, b]$ and it exists g such that $g' = f$, then $\displaystyle\int_a^b f(x)dx = g(b) - g(a)$.

This theorem can be considered a generalization of Barrow's rule, since it allows us to calculate the integral of f between a and b as the difference of the values of a primitive g at these points. But now it is only required that f is integrable and has a primitive, which does not need the continuity of f (see **1.2**). The theorem is proved in a supplementary document (in Spanish).

6 Improper integrals

We have obtained the definite integral of bounded functions on closed intervals (therefore compact). If the interval is not compact or f is not bounded on it, we call these integrals improper and solve them by taking a limit. That is, we obtain the value of the integral between two generic endpoints. Then we calculate the limit of the obtained expression when one of these endpoints tends to infinity or to the point where f would take an infinite value.

Example. Calculate the integral of $f(x) = 1/x^2$: a) between 0 and 1; b) between 1 and ∞.

$$\text{a)} \int_0^1 \frac{dx}{x^2} = \lim_{a\to 0}\int_a^1 \frac{dx}{x^2} = \lim_{a\to 0} -\frac{1}{x}\Big|_a^1 = \lim_{a\to 0}\left(-1 + \frac{1}{a}\right) = +\infty$$

$$\text{b)} \int_1^\infty \frac{dx}{x^2} = \lim_{b\to\infty}\int_1^b \frac{dx}{x^2} = \lim_{b\to\infty} -\frac{1}{x}\Big|_1^b = \lim_{b\to\infty}\left(-\frac{1}{b} + 1\right) = 1$$

The result of b) can be interpreted geometrically by saying that the mixtilinear triangle determined by the curve and the x-axis from $x = 1$ has an infinite perimeter, but a finite area. On the other hand, in case a) both the area and the perimeter of the figure are infinite.

Exercise 1. If $f(x) = 1/\sqrt{x}$, calculate: **a)** $\displaystyle\int_0^1 f(x)\,dx$ $(I = 2)$; **b)** $\displaystyle\int_1^\infty f(x)\,dx$ $(I = \infty)$.

Exercise 2. Calculate the integral of **a)** $f(x) = e^x$ between $-\infty$ and 0 $(I = 1)$; **b)** $g(x) = \ln x$ between 0 and 1 $(I = -1)$. Taking into account that the natural logarithm is the inverse function of the exponential of x, consider whether the obtained result was expected.

7 Self-assesment exercises

7.1 True/False exercise

Decide whether the following statements are true or false.

1. A discontinuous function on $I = [a, b]$ can have a primitive on I. The primitive must be continuous.
2. If $f : I \to \mathbb{R}$ has a jump discontinuity at $a \in I$, the function does not have a primitive on I. That is, it does not exist any function $F : I \to \mathbb{R}$ / $F' = f$ on I.
3. A function is integrable on $I = [a, b]$ if and only if it is continuous.
4. If f and g are integrable on $I = [a, b]$ and $f(x) \leq g(x)$, $\forall x \in I$, then it holds that $\displaystyle\int_a^b f(x)dx \leq \int_a^b g(x)dx$.
5. If f is continuous on $I = [a, b]$, its integral function F is primitive of f on I.
6. The Barrow rule can be applied to the floor function.

7.2 Question

We have studied the concept of definite integral for bounded functions on a compact. ¿How must we proceed to study the integral of a function if the interval is not bounded?

Apply it to the improper integral $\displaystyle\int_{-\infty}^{0} e^x dx$.

7.3 Solution to the True/False exercise

1. **T**. A discontinuous function on $I = [a, b]$ can have a primitive on I, as long as there is not a jump discontinuity (see **1.2.** Necessary condition for the existence of a primitive). If the function has a primitive, this must be derivable, so it will be continuous.
2. **T**. The derivative of a function can be discontinuous as long as the discontinuity is not a jump discontinuity. Then, if f has a jump discontinuity, f cannot be the derivative of a function, so it will not have a primitive (see **1.3**).
3. **F**. If f is continuous, then it is integrable, but not vice versa. For example, the fractional part function $\{x\} = x - \lfloor x \rfloor$ is discontinuous at points corresponding to integer values of x. However, it is integrable between a and b, $\forall a, b \in \mathbb{R}$ (see **2.5**).
4. **T**. It is stated in paragraph c) of **2.6**.
5. **T**. The first fundamental theorem of calculus proves it (see **4.2**).
6. **F**. The Barrow rule is stated for continuous functions (see **4.3**).

7.4 Solution to the question

In this cases we obtain the definite integral on a closed interval, replacing the not bounded endpoint by a generic value. Then we calculate the limit of the resulting expression when the endpoint tends to infinity. If a finite limit exists, the improper integral is convergent.

$$\int_{-\infty}^{0} e^x dx = \lim_{m \to -\infty} \int_m^0 e^x dx = \lim_{m \to -\infty} \left(e^0 - e^m\right) = 1 - \lim_{m \to -\infty} e^m = 1 - 0 = 1$$

Unit II. Vector functions (05.03.2024)

1 Introduction. Types of functions

So far we have been working with real functions (also called real-valued functions) of a real variable, what means that both the variable x and its image $f(x)$ are real numbers.

$$f : \mathbb{R} \to \mathbb{R} \quad \Longrightarrow \quad x \to f(x)$$

In this unit we will study vector functions (also called vector-valued functions) of a vector variable, in which both the variable $\vec{x}$ and its image $\vec{f}(\vec{x})$ are vectors.

$$\vec{f} : \mathbb{R}^n \to \mathbb{R}^m \quad \Longrightarrow \quad \vec{x} \to \vec{f}(\vec{x}), \text{ being } \vec{x} = \begin{Bmatrix} x_1 \\ \vdots \\ x_n \end{Bmatrix}, \quad \vec{f}(\vec{x}) = \begin{Bmatrix} f_1(x_1 \dots x_n) \\ \vdots \\ f_m(x_1 \dots x_n) \end{Bmatrix}$$

Before studying the general case, we will see two particular cases (the second one in detail):

a) Vector functions of a real variable.

$$\vec{f} : \mathbb{R} \to \mathbb{R}^m. \quad \text{Example: } \vec{r}(t) = \begin{Bmatrix} x(t) \\ y(t) \\ z(t) \end{Bmatrix} \text{ (Position vector)}$$

b) Real functions of a vector variable (commonly known as multivariable functions).

$$\vec{f} : \mathbb{R}^n \to \mathbb{R}. \quad \text{Example: } T(\vec{r}) = T(x, y, z) \text{ (function temperature)}$$

Remark. A vector can be denoted both in bold ($\mathbf{v}$) and with an arrow ($\vec{v}$). Likewise, its expanded expression can be written as a row or as a column. It is usually easier to write it as a row and clearer as a column, for example when developing the expression for $\vec{f}(\vec{x})$. Sometimes it will be convenient to write it as a column, for example when premultiplying it by a matrix.

2 Euclidean space

In this section we review some concepts of Algebra that will be used in the unit. Let $\mathbb{R}^p$ be the set of groupings of p real numbers

$$\mathbb{R}^p = \big\{ (x_1, \dots, x_p) \,/\, x_i \in \mathbb{R};\ i = 1, 2 \dots p \big\}$$

It can be shown that $\mathbb{R}^p$ is a vector space over $\mathbb{R}$. We define three applications.

2.1 Ordinary scalar product

It is an application of $\mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$, which associates a real number to each pair of vectors $\vec{x}, \vec{y} \in \mathbb{R}^p$.

$$\left.\begin{aligned} \vec{x} &= (x_1, \dots, x_p) \\ \vec{y} &= (y_1, \dots, y_p) \end{aligned}\right\} \Longrightarrow \vec{x} \cdot \vec{y} = x_1 y_1 + \dots + x_p y_p = \sum_{i=1}^{p} x_i\, y_i$$

Properties.

a) Positivity: $\vec{x} \cdot \vec{x} > 0, \forall \vec{x} \neq \vec{0}$; $\vec{0} \cdot \vec{0} = 0$.

b) Conmutativity: $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}, \forall \vec{x}, \vec{y} \in \mathbb{R}^p$.

c) Distributivity $(\cdot\,/\,+)$: $\vec{x} \cdot (\alpha \vec{y} + \beta \vec{z}) = \alpha(\vec{x} \cdot \vec{y}) + \beta(\vec{x} \cdot \vec{z}), \forall \vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^p; \forall \alpha, \beta \in \mathbb{R}$.

A vector space endowed with the ordinary scalar product is an **Euclidean space**

2.2 Euclidean norm

It is an application of $\mathbb{R}^p \to \mathbb{R}^+ \cup \{0\}$, which associates a positive or null real number to each vector $\vec{x} \in \mathbb{R}^p$. This number is also called the modulus or length of the vector.

$$\|\vec{x}\| = \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{(x_1)^2 + \cdots + (x_p)^2}$$

Properties.

a) Positivity: $\|\vec{x}\| > 0, \forall \vec{x} \neq \vec{0}; \quad \|\vec{0}\| = 0$.

b) Product by a escalar: $\|\lambda \vec{x}\| = |\lambda| \, \|\vec{x}\|, \forall \vec{x} \in \mathbb{R}^p, \forall \lambda \in \mathbb{R}$.

c) Triangular inequality: $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|, \forall \vec{x}, \vec{y} \in \mathbb{R}^p$.

2.3 Euclidean distance

It is an application of $\mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}^+ \cup \{0\}$, which associates a positive or null real number to each pair of vectors $\vec{x}, \vec{y} \in \mathbb{R}^p$.

$$d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_p - y_p)^2}$$

Properties.

a) Positivity: $d(\vec{x}, \vec{y}) > 0, \forall \vec{x} \neq \vec{y}; \quad d(\vec{x}, \vec{x}) = 0$.

b) Simmetry: $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x}), \forall \vec{x}, \vec{y} \in \mathbb{R}^p$.

c) Triangular inequality: $d(\vec{x}, \vec{z}) \leq d(\vec{x}, \vec{y}) + d(\vec{y}, \vec{z}), \forall \vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^p$.

3 Vector functions of a real variable

These functions are vectors of m components, each of them depending on a single variable, defined on a domain D. The generic point $a \in D$, used to define the limit of $\vec{f}$ when $x \to a$, must have points of D as close as desired, then it must be of accumulation of D.

$$\boxed{\vec{f} : D \to \mathbb{R}^m, \quad D \subset \mathbb{R}, \quad a \in D' \cap D}$$

Let us see that **a vector has a limit if and only if each of its m components has a limit**.

Proof. Let $\vec{x}(t) = \big(x_1(t), \ldots, x_m(t)\big)$ and let $\vec{a} = (a_1, \ldots, a_m)$. We say that $\vec{x}$ has a limit $\vec{a}$ (when $t \to t_0$), if $\|\vec{x} - \vec{a}\|$ can be made as small as we want (for t close enough to t_0); that is, if

$$\|\vec{x} - \vec{a}\| < \varepsilon, \forall \varepsilon > 0$$

On the other hand, each component $x_j(t)$ has a limit a_j (when $t \to t_0$), if $|x_j - a_j|$ can also be made as small as we want (for t sufficiently close to t_0); that is, if

$$|x_j - a_j| < \varepsilon, \forall \varepsilon > 0, \quad j = 1, 2, \ldots, m$$

We study the relation between $\|\vec{x} - \vec{a}\|$ and $|x_j - a_j|$.

$$\|\vec{x} - \vec{a}\| = \sqrt{(x_1 - a_1)^2 + \cdots + (x_m - a_m)^2} \geq \begin{cases} \sqrt{(x_1 - a_1)^2} = |x_1 - a_1| \\ \vdots \\ \sqrt{(x_m - a_m)^2} = |x_m - a_m| \end{cases}$$

That is

$$\|\vec{x}-\vec{a}\| \geq |x_j - a_j|\,, \quad j=1,2,\ldots,m$$

If $t \to t_0$, then it turns out that

1) If $\|\vec{x}-\vec{a}\| < \varepsilon$ then $|x_j - a_j| < \varepsilon$

2) If $|x_j - a_j| < \dfrac{\varepsilon}{\sqrt{m}}$, then $\|\vec{x}-\vec{a}\| = \sqrt{\displaystyle\sum_{i=1}^{m} (x_j - a_j)^2} \;<\; \sqrt{\displaystyle\sum_{i=1}^{m} \left(\frac{\varepsilon}{\sqrt{m}}\right)^2} \;=\; \varepsilon$

Therefore, vector $\vec{x}$ has a limit $\vec{a} = \{a_j\}_{j=1,\ldots,m}$ if and only if each component x_j has a limit a_j. This result allows us to reduce the study of the limit of a vector function to that of the limits of its components. Furthermore, the conditions of continuity and differentiability are expressed in limit form, so **a vector function will satisfy these conditions if and only if each of its components does it**.

3.1 Limit

We say that a function $\vec{f}$ has a limit $\vec{\varphi}$ at $x=a$ if, for values of x sufficiently close to a, $\vec{f}(x)$ is as close as we want to $\vec{\varphi}$.

$$\boxed{\lim_{x\to a} \vec{f}(x) = \vec{\varphi}} \Longleftrightarrow \boxed{\forall \varepsilon > 0\, \exists \delta > 0 \,\big/\, 0 < |x-a| < \delta \Longrightarrow \|\vec{f}(x) - \vec{\varphi}\| < \varepsilon}$$

As we have just seen, this limit exists if the limit of each component exists, then the limit of the vector function $\vec{f}$ is a vector whose components are the limits of the components of $\vec{f}$.

$$\boxed{\lim_{x\to a} \vec{f}(x) = \left\{\lim_{x\to a} f_j(x)\right\}_{j=1,\ldots,m}}$$

3.2 Continuity

The continuity condition in limit form says that $\vec{f}$ is continuous at a if its limit when $x \to a$ coincides with $\vec{f}(a)$. But this is equivalent to say that the limit of each component of $\vec{f}$ is the corresponding component of $\vec{f}(a)$, what means that each component must be continuous at a.

$$\boxed{\vec{f}\text{ continuous at } a} \Longleftrightarrow \lim_{x\to a} \vec{f}(x) = \vec{f}(a) \Longleftrightarrow \lim_{x\to a} f_j(x) = f_j(a)_{\,j=1,\ldots,m} \Longleftrightarrow \boxed{f_j \text{ continuous at } a}$$

3.3 Differentiability

Let a be an interior point of the domain D of $\vec{f}$. Each of the m components f_j of $\vec{f}$ is a real function of a real variable, so its differentiability condition at a is

$$f_j(x) - f_j(a) = [g_j + \varepsilon_j(x-a)](x-a) \quad \left(\text{being } \; g_j = \frac{df_j}{dx}(a) \text{ and } \lim_{x\to a} \varepsilon_j(x-a) = 0\right)$$

The differentiability condition of $\vec{f}$ at a is obtained by composing the conditions for the m components, that is, writing the condition for $j=1,2,\ldots,m$ and grouping the different terms f_j as $\vec{f}$, g_j as $\vec{g}$, etc. It results

$$\boxed{\vec{f}(x) - \vec{f}(a) = [\vec{g} + \vec{\varepsilon}(x-a)](x-a)} \quad \left(\lim_{x\to a} \vec{\varepsilon}\,(x-a) = \vec{0}\right)$$

Vector $\vec{g}$ is the derivative of $\vec{f}$. The differential of $\vec{f}$ is the product of $\vec{g}$ by the differential of x.

$$\boxed{\vec{g} = \frac{d\vec{f}}{dx}(a)}\,; \quad \boxed{d\vec{f} = \vec{g}\, dx}\,; \quad \{g_j\}_{j=1,\ldots,m} = \left\{\frac{df_j}{dx}\right\}_{j=1,\ldots,m}$$

Example. Let the position vector be $\vec{r}(t) = \{t, 2t^2, 3t^3\}$. We are going to obtain its limit and check the continuity at $t = 1$; we will calculate its first two derivatives and its differential.

a) $\lim_{t\to 1} \vec{r}(t) = \{1, 2, 3\} = \vec{r}(1)$. All components, and therefore $\vec{r}(t)$, are continuous functions.

b) $\vec{v}(t) = \dfrac{d\vec{r}}{dt} = \{1, 4t, 9t^2\}\,;\quad \vec{a}(t) = \dfrac{d\vec{v}}{dt} = \{0, 4, 18t\}\,;\quad d\vec{r} = \{dt, 4tdt, 9t^2dt\}.$

Exercise. Do the same with $\vec{r}(t) = \left\{\sqrt{t}, 2/t, \sin 3t\right\}$.

4 Real functions of a vector variable

We now study real functions whose variable is a vector of n components, so the domain is a subset of $\mathbb{R}^n$. As in **3**, $\vec{a} \in D$ must be an accumulation point of D.

$$\boxed{f : D \to \mathbb{R}, \quad D \subset \mathbb{R}^n, \quad a \in D' \cap D}$$

4.1 Limit.

a Limit of a function (functional limit).

We say that a function f has a limit φ at $\vec{x} = \vec{a}$ if, for values of $\vec{x}$ sufficiently close to $\vec{a}$, $f(\vec{x})$ is as close to φ as we want.

$$\boxed{\lim_{\vec{x}\to\vec{a}} f(\vec{x}) = \varphi} \iff \boxed{\forall \varepsilon > 0\, \exists \delta > 0 \big/ \, 0 < \|\vec{x} - \vec{a}\| < \delta \Longrightarrow |f(\vec{x}) - \varphi| < \varepsilon}$$

b Directional limit.

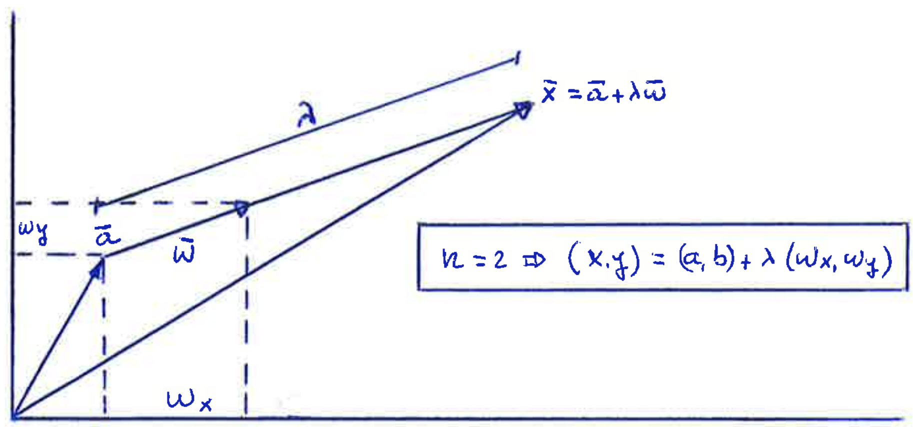
Since the set of variables has dimension $n > 1$, we can approach the point $\vec{a}$ through different subsets, that is, giving values to $\vec{x}$ only on these subsets. The simplest subsets are any of the lines that pass through $\vec{a}$, which we determine by means of its unit vector $\vec{\omega}$, that is

$$E = \left\{\vec{x} \in \mathbb{R}^n \big/ \vec{x} = \vec{a} + \lambda\vec{\omega}, \lambda \in \mathbb{R}, \|\vec{\omega}\| = 1\right\}$$

Then the directional limit of f at $\vec{a}$, along the direction given by $\vec{\omega}$ is

$$\boxed{\lim_{\vec{x}\to\vec{a},\, \vec{x}\in E} f(\vec{x}) = \lim_{\lambda\to 0} f(\vec{a} + \lambda\vec{\omega})}$$

To calculate it, we write $\vec{x}$ as $\vec{a} + \lambda\vec{\omega}$ and find the limit of f as $\lambda \to 0$.



Example. We study the directional limit at the origin of the function

$$f(x,y) = \begin{cases} \dfrac{x}{x+y}, & x+y \neq 0 \\ 0, & x+y = 0 \end{cases}$$

A generic point (x, y), written as a function of the unit vector (ω_x, ω_y) and the point $(a, b) = (0, 0)$, takes the form

$$(x,y) = (a,b) + \lambda(\omega_x, \omega_y) \Longrightarrow \begin{cases} x = \cancel{a} + \lambda\omega_x = \lambda\omega_x \\ y = \cancel{b} + \lambda\omega_y = \lambda\omega_y \end{cases}$$

with what

$$L = \lim_{\lambda\to 0} f(\vec{a} + \lambda\vec{\omega}) = \lim_{\lambda\to 0} f(\lambda\omega_x, \lambda\omega_y) = \lim_{\lambda\to 0} \frac{\lambda\omega_x}{\lambda\omega_x + \lambda\omega_y} = \lim_{\lambda\to 0} \frac{\cancel{\lambda}\omega_x}{\cancel{\lambda}(\omega_x + \omega_y)} = \frac{\omega_x}{\omega_x + \omega_y}$$

We see that the limit L depends on the direction. We obtain it for different cases:

- Axis OX: $\vec{\omega} = (1, 0) \Longrightarrow L = 1$.
- Axis OY: $\vec{\omega} = (0, 1) \Longrightarrow L = 0$.
- Line $y = x$: $\omega_x = \omega_y \Longrightarrow L = 1/2$.

Exercise. In the example above, the limit at the origin along the $y-$axis is equal to 0. Is there any other direction in which it is null? In which direction is the limit at $(0, 0)$ equal to $L = -1$?

c Relation between the functional limit and the directional limits.

If the functional limit condition holds, then it must hold when we approach the point by any particular subset. Hence:

- If there is a functional limit, all the directional limits exist and coincide.
- If the directional limits do not exist or do not coincide, there is no functional limit.
- If the directionals exist and coincide, the functional limit **may** exist. If it exists, it will take that value.

d Finding the functional limit in two variables.

As we see, in certain cases we can ensure that the functional limit does not exist; in others that, if it exists, it takes a certain value, which does not assure us its existence.

In the case of two variables we can prove the existence using polar coordinates with the pole at the point $\vec{a} = (a, b)$. Since in the limit condition we make $\vec{x}$ to be as close to $\vec{a}$ as we want, we impose the condition that $\rho \to 0$, for any value of the angle θ and study the limit of the function (ρ, θ). That is, we write

$$x - a = \rho\cos\theta, \quad y - b = \rho\sin\theta$$

and we calculate

$$\lim_{\rho\to 0} f(a + \rho\cos\theta, b + \rho\sin\theta)$$

If this limit exists $\forall\,\theta$, we have proved that there is a limit when approaching the point along any path, so the function has a limit at (a, b).

Example. We study at the origin the directional and functional limits of

$$f(x,y)=\begin{cases}\dfrac{y^3}{2x^2+y^2}, & (x,y)\neq(0,0)\\ 0, & (x,y)=(0,0)\end{cases}$$

- Directional limits. Since $(a,b)=(0,0)$, $(x,y)=(\lambda\omega_x,\lambda\omega_y)$ and

$$\lim_{\lambda\to 0} f(\lambda\omega_x,\lambda\omega_y)=\lim_{\lambda\to 0}\frac{\lambda^3\omega_y^3}{\lambda^2(2\omega_x^2+\omega_y^2)}=\lim_{\lambda\to 0}\frac{\lambda\,\omega_y^3}{1+\omega_x^2}=0$$

All the directional limits are 0, so there may be a (null) functional limit.

- Functional limit. Let $x=0+\rho\cos\theta,\ y=0+\rho\sin\theta$. We have, $\forall\theta$

$$\lim_{(x,y)\to(0,0)} f(x,y)=\lim_{\rho\to 0} f(\rho\cos\theta,\rho\sin\theta)=\lim_{\rho\to 0}\frac{\rho^3\sin^3\theta}{\rho^2(2\cos^2\theta+\sin^2\theta)}=\lim_{\rho\to 0}\frac{\rho\sin^3\theta}{1+\cos^2\theta}=0$$

Other examples can be seen in the supplementary documents.

Exercise. Check that the directional and functional limits of function g at $(0,0)$ are null.

$$g(x,y)=\begin{cases}\dfrac{x^2y}{x^2+3y^2}, & (x,y)\neq(0,0)\\ 0, & (x,y)=(0,0)\end{cases}$$

4.2 Continuity

We say that f is continuous at $\vec{x}=\vec{a}$ if its limit when $\vec{x}\to\vec{a}$ coincides with $f(\vec{a})$, that is

$$\boxed{f \text{ is continuous at } \vec{a} \Longleftrightarrow \lim_{\vec{x}\to\vec{a}} f(\vec{x})=f(\vec{a})}$$

4.3 Directional and partial derivatives

A first step to study the differentiability condition of a function is to analyze the concepts of directional and partial derivatives at an interior point of the domain, $\vec{a}\in\mathring{D}$.

a Directional derivative.

As in (**b**), we will approach the point $\vec{a}$ along one of the lines passing through $\vec{a}$, which gives rise to the concept of a directional derivative. We write $\vec{x}$ as

$$\vec{x}=\vec{a}+\lambda\vec{\omega},\quad \lambda\in\mathbb{R},\ \|\vec{\omega}\|=1$$

Thus, if $\lambda\to 0$ then $\vec{x}\to\vec{a}$ along the direction given by $\vec{\omega}$. The derivative of f at $\vec{a}$, along the direction given by $\vec{\omega}$, will be the limit of the quotient of the increment of f divided by λ:

$$\boxed{D_{\vec{\omega}}f(\vec{a})=\lim_{\lambda\to 0}\frac{f(\vec{a}+\lambda\vec{\omega})-f(\vec{a})}{\lambda}}$$

The difference $\vec{x}-\vec{a}$ is equal to $\lambda\vec{\omega}$, whose modulus is $|\lambda|$. If $\lambda>0$, then $\vec{x}$ is, with respect to $\vec{a}$, in the direction of $\vec{\omega}$. If $\lambda<0$, it is on the opposite direction. That is, we approach the point $\vec{a}$ along a line, from the two possible sides, as in the one-sided derivatives of functions of one variable. But now there are infinite directions determined by $\vec{\omega}$.

Notice that λ is the distance between $\vec{x}$ and $\vec{a}$ with a sign (positive if $\vec{x}$ is in the direction of $\vec{\omega}$ and negative otherwise).

Example. We obtain the directional derivative at $(0,0)$ of

$$f(x,y)=\begin{cases}\dfrac{x^3+y^3}{x^2+y^2}, & (x,y)\neq(0,0)\\ 0, & (x,y)=(0,0)\end{cases}$$

Since $\vec{a}=(0,0)$, $\vec{a}+\lambda\vec{\omega}=(\lambda\omega_x,\lambda\omega_y)$. And, knowing that $\omega_x^2+\omega_y^2=1$, the $D_{\vec{\omega}}f(\vec{a})$ is

$$\lim_{\lambda\to 0}\frac{f(\lambda\omega_x,\lambda\omega_y)-f(0,0)}{\lambda}=\lim_{\lambda\to 0}\left(\frac{(\lambda\omega_x)^3+(\lambda\omega_y)^3}{(\lambda\omega_x)^2+(\lambda\omega_y)^2}-0\right)\frac{1}{\lambda}=\lim_{\lambda\to 0}\frac{\lambda^3(\omega_x^3+\omega_y^3)}{\lambda^2(\omega_x^2+\omega_y^2)}\frac{1}{\lambda}=\omega_x^3+\omega_y^3$$

Remark. The directional derivative depends in general on the approaching direction to the point, unlike the directional limit of continuous functions, which does not depend.

Exercise. Find the directional derivative at $(0,0)$ of functions

$$\text{a) } f(x,y)=\begin{cases}\dfrac{2x^3-y^3}{x^2+y^2}, & (x,y)\neq(0,0)\\ 0, & (x,y)=(0,0)\end{cases};\quad \text{b) } g(x,y)=\begin{cases}\dfrac{x^2\sin y+y^2\sin x}{x^2+y^2}, & (x,y)\neq(0,0)\\ 0, & (x,y)=(0,0)\end{cases}$$

Solution. a) $D_{\vec{\omega}}f(0,0)=2\omega_x^3-\omega_y^3$; b) $D_{\vec{\omega}}\,g(0,0)=\omega_x^2\omega_y+\omega_y^2\omega_x$.

b Partial derivative

If we take as $\vec{\omega}$ any vector of the canonical base, we are approaching $\vec{a}$ from $\vec{x}$ modifying only one variable. This particular case of directional derivative is called partial derivative. In two or three dimensions, this means approaching the point $\vec{a}$ along the direction of one of the axes.
In the general case, the possible vectors and corresponding modified variables are:

$$\begin{aligned}\vec{\omega}_1&=(1,0,0\ldots0,0)\Longrightarrow\ \text{we modify } x_1\\ \vec{\omega}_2&=(0,1,0\ldots0,0)\Longrightarrow\ \text{we modify } x_2\\ &\vdots\\ \vec{\omega}_i&=(0,\ldots1\ldots,0)\Longrightarrow\ \text{we modify } x_i\\ &\vdots\\ \vec{\omega}_n&=(0,0,0\ldots0,1)\Longrightarrow\ \text{we modify } x_n\end{aligned}$$

so the partial derivative with respect to variable x_i is expressed as

$$\boxed{D_{\vec{\omega}_i}f(\vec{a})=\lim_{\lambda\to 0}\frac{f(a_1,\ldots,a_i+\lambda,\ldots a_n)-f(a_1,\ldots a_n)}{\lambda}=\frac{\partial f}{\partial x_i}(\vec{a})}$$

If $\dfrac{\partial f}{\partial x_i}$ exists for all $\vec{x}$ of the domain, it is called the partial derivative function of f with respect to x_i on D. It is also denoted f'_{x_i}

In practice, if a function is given by a single expression on all its domain, we can obtain the partial derivative with respect to variable x_i by deriving with respect to x_i and considering constant the other variables.

Example. Let $u=x^2yz^3$. Its partial derivatives are: $\dfrac{\partial u}{\partial x}=2xyz^3$, $\dfrac{\partial u}{\partial y}=x^2z^3$, $\dfrac{\partial u}{\partial z}=3x^2yz^2$.

Exercise. Find the partial derivatives at $P(-1,1)$ of functions

a) $\dfrac{y}{x^2}$ $\Big(\text{Sol: }(2,1)\Big)$; b) $x\arctan y$ $\Big(\text{Sol: }(\pi/4,-1/2)\Big)$; c) $\dfrac{y}{x}e^{xy}$ $\Big(\text{Sol: }(-2/e,0)\Big)$.

4.4 Differential

In real functions of one variable, we require as the differentiability condition that the increment of the function is well approximated by a linear function (the differential), which is the product of the derivative of the function by the increment (or differential) of the variable.

In functions of n variables we proceed in a similar way. We will see that the derivative with respect to the vector variable is a vector of n components and the differential (the product of this vector by the increment of the variable) is obtained as a scalar product.

Let $\vec{a}$ be an interior point of the domain D. We say that f is differentiable at $\vec{x} = \vec{a}$ if and only if the following condition is satisfied

$$\boxed{f(\vec{x}) - f(\vec{a}) = [\vec{g} + \vec{\varepsilon}\,(\vec{x} - \vec{a})] \cdot (\vec{x} - \vec{a})}$$

If f is differentiable, the vector $\vec{g}$ is called the total derivative of f or gradient vector. And the differential of f is the (scalar) product of the derivative of the function by the differential of the variable.

$$\boxed{\vec{g} = \left.\frac{df}{d\vec{x}}\right|_{\vec{x}=\vec{a}} \quad ; \quad df = \vec{g} \cdot d\vec{x}}$$

Before calculating the components of the total derivative, we will obtain the derivative of a differentiable function along a certain direction.

Directional derivative of a differentiable function. We replace in the formula of the directional derivative the increment of f by the expression of this increment in the differentiability condition (taking into account that $\vec{x} - \vec{a} = \lambda\omega$). It results

$$D_{\vec{\omega}} f(\vec{a}) = \lim_{\lambda \to 0} \frac{f(\vec{a} + \lambda\omega) - f(\vec{a})}{\lambda} = \lim_{\lambda \to 0} \frac{[\vec{g} + \vec{\varepsilon}\,(\lambda\vec{\omega})] \cdot (\cancel{\lambda}\vec{\omega})}{\cancel{\lambda}} = \vec{g} \cdot \vec{\omega}$$

expression that tells us that the derivative of a differentiable function along a certain direction is given by the scalar product of the gradient and the unit vector of the direction.

Components of the total derivative. To obtain each of them, we calculate the directional derivative with respect to the corresponding vector $\vec{\omega}_i$ of the canonical base (that is, the partial derivative with respect to x_i). We get

$$\frac{\partial f}{\partial x_i} = D_{\vec{\omega}_i} f = \vec{g} \cdot \vec{\omega}_i = (g_1, \ldots, g_i, \ldots, g_n) \cdot (0, 0, \ldots, 1, \ldots, 0) = g_i$$

we see that each component g_i is the partial derivative of f with respect to the variable x_i, so the total derivative of a differentiable function is formed by the n partial derivatives

$$\vec{g} = \frac{df}{d\vec{x}} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n} \right)$$

from which we get the differential of f

$$df = \vec{g} \cdot d\vec{x} = \left(\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n} \right) \cdot \big(dx_1, \ldots, dx_n\big) = \frac{\partial f}{\partial x_1} dx_1 + \cdots + \frac{\partial f}{\partial x_n} dx_n$$

Example. We calculate the differential of the function $u(x, y, z) = x^2 y z^3$ at $P(1, 1, 1)$.

$$du = \frac{\partial u}{\partial x} dx + \frac{\partial u}{\partial y} dy + \frac{\partial u}{\partial z} dz = 2xyz^3 dx + x^2 z^3 dy + 3x^2 y z^2 dz \stackrel{P}{=} 2dx + dy + 3dz$$

Exercise. Find the differential of the following three variable functions at $P(1, 1, \pi)$.

a) $f = \sin xyz \;\; (df_P = -\pi dx - \pi dy - dz)$; b) $g = \sqrt{xyz} \; \left(dg_P = \frac{\sqrt{\pi}}{2} dx + \frac{\sqrt{\pi}}{2} dy + \frac{dz}{2\sqrt{\pi}} \right)$

4.5 Gradient and contour lines. Geometric interpretation

An important feature of the gradient vector $\vec{g}$ (usually denoted as $\vec{\nabla} f$) is that it indicates the direction along which the value of the function increases at a fastest rate. That is, it tells us how to modify the variables so that f grows as fast as possible.

Proof (for 2 or 3 variables). The scalar product of two vectors $\vec{u}$ and $\vec{v}$ is the product of their modules (norms) and the cosine of the angle they form. Since the derivative of f along the direction of $\vec{\omega}$ is the scalar product of the gradient by $\vec{\omega}$, it turns out

$$D_{\vec{\omega}} f = \vec{\nabla} f \cdot \vec{\omega} = \|\vec{\nabla} f\| \, \|\vec{\omega}\| \cos\varphi = \|\vec{\nabla} f\| \cos\varphi$$

This product will be maximum if the cosine is equal to 1, which corresponds to an angle $\varphi = 0$ between $\vec{\nabla} f$ and $\vec{\omega}$ ($\vec{\omega}$ parallel to $\vec{\nabla} f$). The product will be zero if $\varphi = \pi/2$. Therefore, the variation of f is maximum along the direction of $\vec{\nabla} f$; and is null in the perpendicular direction.

Example in 2 variables. Consider the function $f(x,y) = x^2 + y^2$, whose graphic representation is a paraboloid with the vertex at the origin. Let $P = (1,1)$. The gradient vector is

$$\vec{\nabla} f = (f'_x, f'_y) = (2x, 2y) \stackrel{P}{=} (2,2)$$

This means that, when modifying the variables (x,y) –from their values at P– along the direction of the vector $(2,2)$, the increase in the value of f is maximum. Indeed, when moving in that direction, the corresponding point on the surface goes up the line of maximum slope.

On the contrary, if we move in the direction perpendicular to $\vec{\nabla} f$ at each point, we follow a curve in which the function does not vary, that is, it takes the same value as at $P(1,1)$. Its equation will be $f(x,y) = f(1,1)$, i.e. the circumference $x^2 + y^2 = 1^2 + 1^2 = 2$. We have thus obtained a **contour line** (*locus* of points where f is constant).

Thus the gradient of f at P is perpendicular to the contour line of f that passes through P.

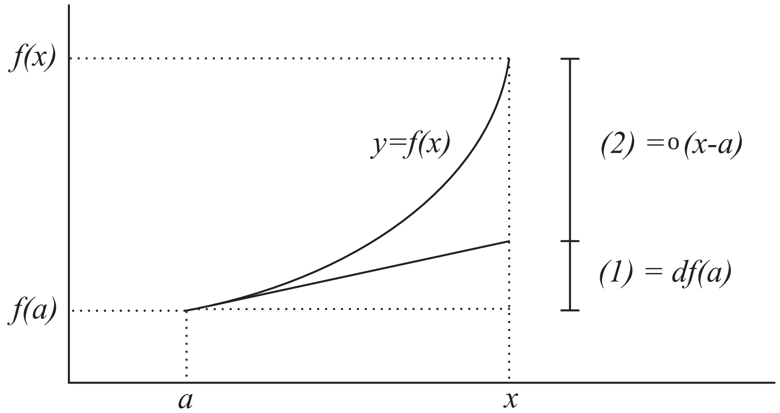
Geometric interpretation of the differential (functions of 2 variables). In functions of one variable, we interpreted geometrically the differential as the increase in the ordinate of the tangent line to the curve when the variable increases $x - a$ (see figure).

The equation of the tangent line to the curve $y = f(x)$ is

$$y = f(a) + f'(a)(x-a) \Longleftrightarrow y - f(a) = f'(a)(x-a) = df(a)$$

In two variables the interpretation is analogous: the differential represents the increment of the vertical coordinate of the tangent plane to the surface $z = f(x,y)$ when going from $P(a,b)$ to $P'(a+dx, b+dy)$. The equation of the plane tangent to the surface $z = f(x,y)$ is

$$z = f(a,b) + f'_x(a,b)(x-a) + f'_y(a,b)(y-b) \Leftrightarrow z - f(a,b) = f'_x(a,b)(x-a) + f'_y(a,b)(y-b) = df(a,b)$$



4.6 Differentiability theorems

Consider a function $f : D \to \mathbb{R},\ D \subset \mathbb{R}^n$. Let $\vec{a}$ be an interior point of the domain, $\vec{a} \in D$. The following theorems are very useful in the study of multivariable functions.

a Sufficient condition of differentiability

If f has partial derivatives on a neighborhood of $\vec{a}$ and they are continuous at $\vec{a}$, then f is differentiable at $\vec{a}$ (if the partial derivatives of f are continuous we say that $f \in C^1$).

Remark. This condition is sufficient, but not necessary. In fact, we can require even fewer conditions to ensure the differentiability of f. Indeed, it can be proved that, if the n partial derivatives exist on a neighborhood of $\vec{a}$ and $n-1$ of them are continuous at $\vec{a}$, then f is differentiable at $\vec{a}$.

b Necessary condition of differentiability

If f is differentiable at $\vec{a}$, its directional derivatives at $\vec{a}$ exist and take the value

$$\boxed{D_{\vec{\omega}} f(\vec{a}) = \vec{g} \cdot \vec{\omega}} \text{ being } \vec{g} = \left.\frac{df}{d\vec{x}}\right|_{\vec{x}=\vec{a}}$$

Directional derivatives depend in general on the direction, unlike directional limits of continuous functions. However, it can be proved that, if a differentiable function reaches an extremum at a point $\vec{a}$, interior of the domain, the directional derivatives of f at $\vec{a}$ are zero.

This condition is necessary but not sufficient. Since there is more than one variable, there are other ways of approaching a point, in addition to the lines that pass through it. Thus the existence of directional derivatives does not imply differentiability.

c Relation between differentiability and continuity

If f is differentiable at a point, it is continuous at that point, but not vice versa.

Indeed, taking limits on the differentiability condition,

$$f(\vec{x}) - f(\vec{a}) = [\vec{g} + \vec{\varepsilon}(\vec{x}-\vec{a})] \cdot (\vec{x}-\vec{a}) \Longrightarrow \boxed{\lim_{\vec{x}\to\vec{a}} f(\vec{x}) = f(\vec{a})}$$

The existence of directional derivatives, which does not ensure differentiability, does not ensure continuity either. If f admits derivative in one direction, the limit of f at $\vec{a}$ will coincide with $f(\vec{a})$ only in that direction.

Example. We will apply the previous theorems to three functions, obtaining their partial derivatives in different ways. Let the functions be $f_1,\ f_2$ and f_3:

$$f_1 = x^2 + y^2; \quad f_2 = \begin{cases} \dfrac{x^2 y}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}; \quad f_3 = \begin{cases} \dfrac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

We will see that:

1. f_1 satisfies the sufficient condition of differentiability at $\vec{a}\,(1,1)$. Hence it is differentiable at $\vec{a}$ and satisfies the necessary condition of differentiability at that point.

2. f_2 has directional derivatives at the origin, but it is not differentiable at that point, since it does not fulfill at it the necessary condition of differentiability.

3. f_3 has partial derivatives at the origin, although it is not continuous at that point.

Solution.

1. $f_1 = x^2 + y^2$. Since it is defined on all $\mathbb{R}^2$, we obtain its partial derivatives directly:
$$\frac{\partial f_1}{\partial x} = 2x; \quad \frac{\partial f_1}{\partial y} = 2y$$
The partial derivatives are continuous (sufficient condition), so f_1 is differentiable (we already knew it, since f_1 is a polynomial). Its gradient at $\vec{a}$ is equal to $\vec{\nabla} f_1(1,1) = (2,2)$.

   We obtain the directional derivative at $\vec{a}$.
$$D_{\vec{\omega}} f_1(\vec{a}) = \lim_{\lambda \to 0} \frac{f_1(\vec{a} + \lambda\vec{\omega}) - f_1(\vec{a})}{\lambda} = \lim_{\lambda \to 0} \frac{(1+\lambda\omega_x)^2 + (1+\lambda\omega_y)^2 - 2}{\lambda} = 2\omega_x + 2\omega_y$$
noticing that the necessary condition of differentiability is satisfied.
$$D_{\vec{\omega}} f_1(\vec{a}) = 2\omega_x + 2\omega_y = (2,2)\begin{Bmatrix}\omega_x \\ \omega_y\end{Bmatrix} = \vec{\nabla} f_1(\vec{a}) \cdot \vec{\omega}$$

2. $f_2 = \begin{cases} \dfrac{x^2 y}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$. We study its directional derivative at $(0,0)$.
$$D_{\vec{\omega}} f_2(0,0) = \lim_{\lambda \to 0} \frac{f_2(\lambda\omega_x, \lambda\omega_y) - f_2(0,0)}{\lambda} = \lim_{\lambda \to 0} \left(\frac{\lambda^3 \omega_x^2 \omega_y}{\lambda^2(\omega_x^2 + \omega_y^2)} - 0 \right) \frac{1}{\lambda} = \omega_x^2 \omega_y$$
from where we get the partial derivatives
$$\frac{\partial f_2}{\partial x} = D_{(1,0)} f_2(0,0) = 0, \quad \frac{\partial f_2}{\partial y} = D_{(0,1)} f_2(0,0) = 0$$
and we see that f_2 is not differentiable at $(0,0)$, since the necessary condition is not satisfied:
$$D_{\vec{\omega}} f_2(0,0) = \omega_x^2\,\omega_y \neq (0,0)\begin{Bmatrix}\omega_x \\ \omega_y\end{Bmatrix} = \vec{\nabla} f_2(0,0) \cdot \vec{\omega}$$

3. $f_3 = \begin{cases} \dfrac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$. We calculate its directional derivative at $(0,0)$
$$D_{\vec{\omega}} f_3(0,0) = \lim_{\lambda \to 0} \frac{f_3(\lambda\omega_x, \lambda\omega_y) - f_3(0,0)}{\lambda} = \lim_{\lambda \to 0} \left(\frac{\lambda^2 \omega_x \omega_y}{\lambda^2(\omega_x^2 + \omega_y^2)} - 0 \right) \frac{1}{\lambda} = \lim_{\lambda \to 0} \frac{\omega_x \omega_y}{\lambda}$$
limit that does not exist in general (unless one of the factors ω_x, ω_y is null).

   We obtain the partial derivatives at $(0,0)$ by using the unit vectors $(1,0)$ and $(0,1)$:
$$\frac{\partial f_3}{\partial x}(0,0) = \lim_{h \to 0} \frac{f_3(h,0) - f_3(0,0)}{h} = \lim_{h \to 0} \left(\frac{h\,0}{h^2+0} - 0 \right) \frac{1}{h} = 0, \quad \frac{\partial f_3}{\partial y}(0,0) = \cdots = 0$$
So f_3 has partial derivatives at $(0,0)$, but it can be verified that it has no limit at $(0,0)$, so it is not continuous.

We have seen then **three methods to obtain the partial derivatives**: 1) deriving the expression of the function, 2) as a particular case of the directional derivatives and 3) from the unit vectors of the axes.

Exercise. Apply what we have seen in the above example to
$$f_1 = x^2 + y^3; \quad f_2 = \begin{cases} \dfrac{xy^2}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}; \quad f_3 = \begin{cases} \dfrac{x^3 y}{x^4+y^4} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

d Necessary and sufficient condition of differentiability

The previous theorems help us to recognize differentiable functions (section **a**) or to be sure that they are not so (sections **b** and **c**). But there are continuous functions that do not verify the sufficient condition of differentiability but do verify the necessary one, so we need another condition. To do this we are going to modify the expression used in the definition of a differentiable function, writing it in the form of a limit.

Consider a function $f : D \to \mathbb{R},\ D \subset \mathbb{R}^n$. Let $\vec{a} \in D$. We say that f is differentiable at $\vec{a}$ if it holds:

$$f(\vec{x}) - f(\vec{a}) = [\vec{g} + \vec{\varepsilon}] \cdot (\vec{x} - \vec{a})$$

We operate the scalar product of the second member, remembering that the product of $\vec{g}$ (derivative of f) by the increment of the variable is the differential of f. And that the product of $\vec{\varepsilon}$ and $(\vec{x} - \vec{a})$ is an infinitesimal of higher order than $(\vec{x} - \vec{a})$. Then,

$$f(\vec{x}) - f(\vec{a}) = \vec{g} \cdot (\vec{x} - \vec{a}) + \vec{\varepsilon} \cdot (\vec{x} - \vec{a}) = df(\vec{a}) + o(\vec{x} - \vec{a})$$

from where

$$f(\vec{x}) - f(\vec{a}) - df(\vec{a}) = o(\vec{x} - \vec{a})$$

Since the term on the left is $o(\vec{x} - \vec{a})$, it must hold

$$\boxed{\lim_{\vec{x} \to \vec{a}} \frac{f(\vec{x}) - f(\vec{a}) - df(\vec{a})}{\|\vec{x} - \vec{a}\|} = 0}$$

Function of two variables. In this case we use the following notation (by f'_x, f'_y we indicate their values at $\vec{a}$):

$$\vec{x} = \begin{Bmatrix} x \\ y \end{Bmatrix};\ \vec{a} = \begin{Bmatrix} a \\ b \end{Bmatrix};\ \vec{x} - \vec{a} = \begin{Bmatrix} x - a \\ y - b \end{Bmatrix} = \begin{Bmatrix} h \\ k \end{Bmatrix};\ \vec{g} = \begin{Bmatrix} f'_x \\ f'_y \end{Bmatrix};\ \vec{df}(\vec{a}) = \begin{pmatrix} f'_x & f'_y \end{pmatrix} \begin{Bmatrix} h \\ k \end{Bmatrix}$$

And the (necessary and sufficient) condition of differentiability of f in $\vec{a}$ results in:

$$\boxed{\lim_{(h,k) \to (0,0)} \frac{f(a+h, b+k) - f(a,b) - hf'_x - kf'_y}{\sqrt{h^2 + k^2}} = 0}$$

Example. We study the differentiability at $(0,0)$ of $f(x,y) = \begin{cases} \dfrac{x^4 - y^4}{x^2 + y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$

The partial derivatives at the origin are:

$$f'_x = \lim_{h \to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h \to 0} \left(\frac{h^4 - 0}{h^2 + 0} - 0 \right) \frac{1}{h} = 0; \quad f'_y = \lim_{k \to 0} \frac{f(0,k) - f(0,0)}{k} = \cdots = 0$$

Since f, f'_x and f'_y are null at $(0,0)$, the condition takes the form: $\displaystyle\lim_{(h,k) \to (0,0)} \frac{h^4 - k^4}{\left(h^2 + k^2\right)^{3/2}} = 0$, that we solve using polar coordinates (see **d**):

$$\lim_{\rho \to 0} \frac{\rho^4 \left(\cos^4 \theta - \sin^4 \theta\right)}{\left[\rho^2 \left(\cos^2 \theta + \sin^2 \theta\right)\right]^{3/2}} = \lim_{\rho \to 0} \frac{\rho^4}{\rho^3} \left(\cos^4 \theta - \sin^2 \theta\right) = 0$$

hence f is differentiable at the origin.

5 Vector functions of a vector variable

We will now study the general case of vector functions of a vector variable, defining limit, continuity and differentiability. This will allow us to define the composite function and analyze its continuity and differentiability. We will then return to real functions of a vector variable to study the higher order derivatives, the Taylor expansion and the extrema of functions.

Consider a function $f : D \to \mathbb{R}^m,\ D \subset \mathbb{R}^n$. Let $\vec{a} \in D' \cap D$. As already mentioned (section **1**), the variable $\vec{x}$ and the function $\vec{f}(\vec{x})$ take the form

$$\vec{x} = \begin{Bmatrix} x_1 \\ \vdots \\ x_n \end{Bmatrix}; \qquad \vec{f}(\vec{x}) = \begin{Bmatrix} f_1(x_1 \dots x_n) \\ \vdots \\ f_m(x_1 \dots x_n) \end{Bmatrix}$$

5.1 Limit

From what was seen in **3.1**, it follows that in this type of function there is a limit if and only if it exists for each of the components, in which case

$$\lim_{\vec{x}\to\vec{a}} \vec{f}(\vec{x}) = \begin{Bmatrix} \lim\limits_{\vec{x}\to\vec{a}} f_1(\vec{x}) \\ \vdots \\ \lim\limits_{\vec{x}\to\vec{a}} f_m(\vec{x}) \end{Bmatrix} = \begin{Bmatrix} \lim\limits_{\vec{x}\to\vec{a}} f_1(x_1 \dots x_n) \\ \vdots \\ \lim\limits_{\vec{x}\to\vec{a}} f_m(x_1 \dots x_n) \end{Bmatrix}$$

5.2 Continuity

Taking into account the above, the condition of continuity at the point $\vec{a}$, in the form of a limit, is expressed as

$$\vec{f} \text{ is continuous at } \vec{a} \iff \lim_{\vec{x}\to\vec{a}} \vec{f}(\vec{x}) = \vec{f}(\vec{a}) \iff \lim_{\vec{x}\to\vec{a}} f_j(\vec{x}) = f_j(\vec{a}),\ j = 1, \dots m$$

Then $\vec{f}$ is continuous at $\vec{a}$ if and only if each of its m components is continuous.

5.3 Differentiability

As in the case of the limit, from **3.3** we conclude that these functions are differentiable if and only if each component satisfies the differentiability condition, that is

$$\vec{f} \text{ is differentiable at } \vec{a} \iff f_j(\vec{x}) - f_j(\vec{a}) = \left[\frac{df_j}{d\vec{x}} + \vec{\varepsilon}_j(\vec{x} - \vec{a})\right] \cdot (\vec{x} - \vec{a}),\ j = 1, \dots m$$

Grouping the m conditions, we obtain the differentiability condition at $\vec{a}$ for vector functions of a vector variable:

$$\vec{f}(\vec{x}) - \vec{f}(\vec{a}) = \Big[G_{m\times n} + \vec{\varepsilon}_{m\times n}(\vec{x} - \vec{a})\Big]\,(\vec{x} - \vec{a})$$

where

$$G_{m\times n} = \begin{Bmatrix} \dfrac{df_1}{d\vec{x}} \\ \vdots \\ \dfrac{df_m}{d\vec{x}} \end{Bmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix}$$

is the Jacobian matrix, whose rows are the total derivatives of the components of $\vec{f}$. Matrix $\vec{\varepsilon}_{m\times n}(\vec{x} - \vec{a})$ approaches the null matrix as $\vec{x} \to \vec{a}$.

If $\vec{f}$ is differentiable, its differential is obtained by multiplying the Jacobian matrix times the vector $d\vec{x}$, resulting

$$d\vec{f} = \begin{Bmatrix} df_1 \\ \vdots \\ df_m \end{Bmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1} & \cdots & \dfrac{\partial f_m}{\partial x_n} \end{pmatrix} \begin{Bmatrix} dx_1 \\ \vdots \\ dx_n \end{Bmatrix} = \begin{Bmatrix} \dfrac{\partial f_1}{\partial x_1}dx_1 + & \cdots & + \dfrac{\partial f_1}{\partial x_n}dx_n \\ \vdots & & \vdots \\ \dfrac{\partial f_m}{\partial x_1}dx_1 + & \cdots & + \dfrac{\partial f_m}{\partial x_n}dx_n \end{Bmatrix}$$

If the n partial derivatives of each component f_j of function $\vec{f}$ are continuous, f_j will be differentiable. Thus a sufficient condition of differentiability for $\vec{f}$ is that its $m \times n$ partial derivatives are continuous.

Example. We define a function $\vec{f} : \mathbb{R}^3 \to \mathbb{R}^3$, whose components are

$$u = f_1(x, y, z) = x^2 + xy + 1; \quad v = f_2(x, y, z) = x + y^2; \quad w = f_3(x, y, z) = y + z^2$$

We find the Jacobian matrix at $P(1, 1, 1)$. Using simplified notation,

$$J = \frac{d(u, v, w)}{d(x, y, z)} = \begin{pmatrix} \dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y} & \dfrac{\partial u}{\partial z} \\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y} & \dfrac{\partial v}{\partial z} \\ \dfrac{\partial w}{\partial x} & \dfrac{\partial w}{\partial y} & \dfrac{\partial w}{\partial z} \end{pmatrix} = \begin{pmatrix} 2x + y & x & 0 \\ 1 & 2y & 0 \\ 0 & 1 & 2z \end{pmatrix} \overset{P}{=} \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix}$$

Since all nine partial derivatives are continuous, $\vec{f}$ is differentiable. Its differential is

$$d\vec{f}_P = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 2 \end{pmatrix} \begin{Bmatrix} dx \\ dy \\ dz \end{Bmatrix} = \begin{Bmatrix} 3dx + dy \\ dx + 2dy \\ dy + 2dz \end{Bmatrix}$$

6 Composition of functions

6.1 Composite function. Continuity and differentiability

Given the sets $\mathbb{R}^{n_x}, \mathbb{R}^{n_y}, \mathbb{R}^{n_z}$, we consider the functions $\vec{f} : D_x \to \mathbb{R}^{n_y}$ and $\vec{g} : D_y \to \mathbb{R}^{n_z}$, being $D_x \subset \mathbb{R}^{n_x}$ and $D_y \subset \mathbb{R}^{n_y}$ their domains.
If $f(D_x) \subset D_y$, we define the composite function $\left(\vec{g} \circ \vec{f}\right)(\vec{x}) = \vec{g}\left(\vec{f}(\vec{x})\right) = \vec{\phi}(\vec{x})$. It holds that:

- **a)** *"If $\vec{f}$ is continuous at $\vec{a} \in D_x$ and $\vec{g}$ is continuous at $\vec{y} = \vec{f}(\vec{a}) \in D_y$, the composite function $\vec{g} \circ \vec{f}$ is continuous at $\vec{x} = \vec{a}$".*

- **b)** *"If $\vec{f}$ is differentiable at $\vec{a} \in D_x$ and $\vec{g}$ is differentiable at $\vec{y} = \vec{f}(\vec{a}) \in D_y$, the composite function $\vec{g} \circ \vec{f}$ is differentiable at $\vec{x} = \vec{a}$ ".*

Proof of differentiability: We start from the differentiability conditions of $\vec{f}$ and $\vec{g}$ at $\vec{x} = \vec{a}$ and $\vec{y} = \vec{b}$ respectively (with the notation $\underset{\sim}{M}$ we symbolize the matrix M):

$$\vec{f}(\vec{x}) - \vec{f}(\vec{a}) = \left[\left.\frac{d\vec{f}}{d\vec{x}}\right|_{\vec{x}=\vec{a}} + \underset{\sim}{\mathcal{E}_x}(\vec{x} - \vec{a}) \right] (\vec{x} - \vec{a}), \quad \text{being } \lim_{\vec{x}\to\vec{a}} \underset{\sim}{\mathcal{E}_x} = \underset{\sim}{\Omega} \tag{1}$$

$$\vec{g}(\vec{y}) - \vec{g}(\vec{b}) = \left[\left.\frac{d\vec{g}}{d\vec{y}}\right|_{\vec{y}=\vec{b}} + \underset{\sim}{\mathcal{E}_y}(\vec{y} - \vec{b}) \right] (\vec{y} - \vec{b}), \quad \text{being } \lim_{\vec{y}\to\vec{b}} \underset{\sim}{\mathcal{E}_y} = \underset{\sim}{\Omega} \tag{2}$$

Replacing $\vec{y}$ and $\vec{b}$ by $\vec{f}(\vec{x})$ and $\vec{f}(\vec{a})$ respectively in condition (2), we obtain

$$\vec{g}(\vec{f}(\vec{x})) - \vec{g}(\vec{f}(\vec{a})) = \left[\left.\frac{d\vec{g}}{d\vec{y}}\right|_{\vec{y}=\vec{f}\vec{a}} + \underset{\sim}{\mathcal{E}}_y \left(\vec{f}(\vec{x}) - \vec{f}(\vec{a})\right)\right] \left(\vec{f}(\vec{x}) - \vec{f}(\vec{a})\right)$$

Now we give the factor $\left(\vec{f}(\vec{x}) - \vec{f}(\vec{a})\right)$ its value in (1) and operate, It results a product of brackets, in which we omit the dependencies of $\varepsilon_{\vec{x}}$ and $\varepsilon_{\vec{y}}$ for simplicity:

$$\vec{g}(\vec{f}(\vec{x})) - \vec{g}(\vec{f}(\vec{a})) = \left[\left.\frac{d\vec{g}}{d\vec{y}}\right|_{\vec{y}=\vec{f}\vec{a}} + \underset{\sim}{\mathcal{E}}_y \right] \left[\left.\frac{d\vec{f}}{d\vec{x}}\right|_{\vec{x}=\vec{a}} + \underset{\sim}{\mathcal{E}}_x \right] (\vec{x} - \vec{a}) =$$

$$\left[\left.\frac{d\vec{g}}{d\vec{y}}\right|_{\vec{y}=\vec{f}\vec{a}} \left.\frac{d\vec{f}}{d\vec{x}}\right|_{\vec{x}=\vec{a}} + \left.\frac{d\vec{g}}{d\vec{y}}\right|_{\vec{y}=\vec{f}\vec{a}} \underset{\sim}{\mathcal{E}}_x + \underset{\sim}{\mathcal{E}}_y \left.\frac{d\vec{f}}{d\vec{x}}\right|_{\vec{x}=\vec{a}} + \underset{\sim}{\mathcal{E}}_y \underset{\sim}{\mathcal{E}}_x \right] (\vec{x} - \vec{a})$$

When $\vec{x} \to \vec{a}$, $\vec{f}(\vec{x}) \to \vec{f}(\vec{a})$ by continuity. Then the addends 2nd, 3rd and 4th of the bracket tend to the null matrix, so they are expressions of the form $\underset{\sim}{\mathcal{E}}(\vec{x} - \vec{a})$.

The first addend represents the derivative of the composite function, formed by the product of the (Jacobian) derivatives of $\vec{g}$ with respect to $\vec{y}$ and of $\vec{f}$ with respect to $\vec{x}$. If $\vec{g} \circ \vec{f} = \vec{\phi}$, we get the differentiability condition of $\vec{\phi}$

$$\vec{\phi}(\vec{x}) - \vec{\phi}(\vec{a}) = \left[\left.\frac{d\vec{\phi}}{d\vec{x}}\right|_{\vec{x}=\vec{a}} + \underset{\sim}{\mathcal{E}}(\vec{x} - \vec{a}) \right] (\vec{x} - \vec{a}),$$

being the derivative of the composite function at $\vec{x} = \vec{a}$,

$$\boxed{\left.\frac{d\vec{\phi}}{d\vec{x}}\right|_{\vec{x}=\vec{a}} = \left.\frac{d\vec{g}}{d\vec{y}}\right|_{\vec{y}=\vec{f}\vec{a}} \left.\frac{d\vec{f}}{d\vec{x}}\right|_{\vec{x}=\vec{a}}}$$

If the above is true $\forall \vec{x} \in D_x$, the composite function is differentiable on D_x.

6.2 Derivative of the composite function. Chain rule

We have seen that the derivative of the composite function $\vec{\phi} = \vec{g} \circ \vec{f}$ is a matrix obtained as the product of two: the derivative of $\vec{g}$ with respect to $\vec{y}$ and the derivative of $\vec{f}$ with respect to $\vec{x}$. That is, **the Jacobian of the composite function is the product of the Jacobians of the functions**.

$$\frac{d\vec{\phi}}{d\vec{x}} = \begin{pmatrix} \frac{\partial \phi_1}{\partial x_1} & \cdots & \frac{\partial \phi_1}{\partial x_{n_x}} \\ \vdots & & \vdots \\ \frac{\partial \phi_{n_z}}{\partial x_1} & \cdots & \frac{\partial \phi_{n_z}}{\partial x_{n_x}} \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_1}{\partial y_{n_y}} \\ \vdots & & \vdots \\ \frac{\partial g_{n_z}}{\partial y_1} & \cdots & \frac{\partial g_{n_z}}{\partial y_{n_y}} \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_{n_x}} \\ \vdots & & \vdots \\ \frac{\partial f_{n_y}}{\partial x_1} & \cdots & \frac{\partial f_{n_y}}{\partial x_{n_x}} \end{pmatrix}$$

We now develop the product of matrices. We observe that the element (i,k) of the first matrix (derivative of ϕ_i with respect to x_k) is obtained by multiplying the i-th row of the second by the $k-$th column of the third, that is:

$$\frac{\partial \phi_i}{\partial x_k} = \frac{\partial g_i}{\partial y_1}\frac{\partial f_1}{\partial x_k} + \frac{\partial g_i}{\partial y_2}\frac{\partial f_2}{\partial x_k} + \dots \frac{\partial g_i}{\partial y_{n_y}}\frac{\partial f_{n_y}}{\partial x_k} = \sum_{j=1}^{n_y} \frac{\partial g_i}{\partial y_j}\frac{\partial f_j}{\partial x_k}$$

which we can interpret by saying that *the derivative of ϕ_i with respect to x_k is the sum of the derivatives obtained through all the functions f_j that relate g_i with x_k*.

Example 1. Consider the functions $\vec{f}:\mathbb{R}^2 \longrightarrow \mathbb{R}^3$; $\vec{g}:\mathbb{R}^3 \longrightarrow \mathbb{R}^2$. From them we define the composite function

$$\vec{z}=\vec{g}(\vec{f}(\vec{x}))=\vec{\phi}(\vec{x})$$

We have:

$$\begin{Bmatrix} z_1 \\ z_2 \end{Bmatrix}=\begin{Bmatrix} g_1(y_1,y_2,y_3) \\ g_2(y_1,y_2,y_3) \end{Bmatrix};\quad \begin{Bmatrix} y_1 \\ y_2 \\ y_3 \end{Bmatrix}=\begin{Bmatrix} f_1(x_1,x_2) \\ f_2(x_1,x_2) \\ f_3(x_1,x_2) \end{Bmatrix} \Longrightarrow \begin{Bmatrix} z_1 \\ z_2 \end{Bmatrix}=\begin{Bmatrix} \phi_1(x_1,x_2) \\ \phi_2(x_1,x_2) \end{Bmatrix}$$

The derivatives of the functions ϕ_1,ϕ_2 with respect to the variables x_1,x_2 are:

$$\begin{aligned}
\frac{\partial \phi_1}{\partial x_1} &= \frac{\partial g_1}{\partial y_1}\frac{\partial f_1}{\partial x_1}+\frac{\partial g_1}{\partial y_2}\frac{\partial f_2}{\partial x_1}+\frac{\partial g_1}{\partial y_3}\frac{\partial f_3}{\partial x_1}\\
\frac{\partial \phi_1}{\partial x_2} &= \frac{\partial g_1}{\partial y_1}\frac{\partial f_1}{\partial x_2}+\frac{\partial g_1}{\partial y_2}\frac{\partial f_2}{\partial x_2}+\frac{\partial g_1}{\partial y_3}\frac{\partial f_3}{\partial x_2}\\
\frac{\partial \phi_2}{\partial x_1} &= \frac{\partial g_2}{\partial y_1}\frac{\partial f_1}{\partial x_1}+\frac{\partial g_2}{\partial y_2}\frac{\partial f_2}{\partial x_1}+\frac{\partial g_2}{\partial y_3}\frac{\partial f_3}{\partial x_1}\\
\frac{\partial \phi_2}{\partial x_2} &= \frac{\partial g_2}{\partial y_1}\frac{\partial f_1}{\partial x_2}+\frac{\partial g_2}{\partial y_2}\frac{\partial f_2}{\partial x_2}+\frac{\partial g_2}{\partial y_3}\frac{\partial f_3}{\partial x_2}
\end{aligned}$$

These expressions are greatly simplified by using the name of the variables to represent the functions, that is, replacing ϕ_i and g_i by z_i; and f_j by y_j. It results:

$$\begin{aligned}
\frac{\partial z_1}{\partial x_1} &= \frac{\partial z_1}{\partial y_1}\frac{\partial y_1}{\partial x_1}+\frac{\partial z_1}{\partial y_2}\frac{\partial y_2}{\partial x_1}+\frac{\partial z_1}{\partial y_3}\frac{\partial y_3}{\partial x_1}\\
\frac{\partial z_1}{\partial x_2} &= \frac{\partial z_1}{\partial y_1}\frac{\partial y_1}{\partial x_2}+\frac{\partial z_1}{\partial y_2}\frac{\partial y_2}{\partial x_2}+\frac{\partial z_1}{\partial y_3}\frac{\partial y_3}{\partial x_2}\\
\frac{\partial z_2}{\partial x_1} &= \frac{\partial z_2}{\partial y_1}\frac{\partial y_1}{\partial x_1}+\frac{\partial z_2}{\partial y_2}\frac{\partial y_2}{\partial x_1}+\frac{\partial z_2}{\partial y_3}\frac{\partial y_3}{\partial x_1}\\
\frac{\partial z_2}{\partial x_2} &= \frac{\partial z_2}{\partial y_1}\frac{\partial y_1}{\partial x_2}+\frac{\partial z_2}{\partial y_2}\frac{\partial y_2}{\partial x_2}+\frac{\partial z_2}{\partial y_3}\frac{\partial y_3}{\partial x_2}
\end{aligned}$$

It should be remembered that those who depend on the variables are the functions: the function g_i, on y_j; and the functions ϕ_i, f_j, on x_k. And that one variable can depend on another through different functions, which we must specify. Therefore, this notation is simpler but less rigorous.

Example 2. Let the function $z=f(x,y)=x^2-3y$, being $x=u-v;\quad y=2u+v^2$.

We find the derivatives of z with respect to u and v.

$$\begin{aligned}
\frac{\partial z}{\partial u} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial u}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial u}=2x\cdot 1+(-3)\,2=2(u-v)-6=2u-2v-6\\
\frac{\partial z}{\partial v} &= \frac{\partial z}{\partial x}\frac{\partial x}{\partial v}+\frac{\partial z}{\partial y}\frac{\partial y}{\partial v}=2x\,(-1)+(-3)\,2v=-2(u-v)-6v=-2u-4v
\end{aligned}$$

It is easy to check that we get the same result by writing z as a function of u and v

$$z=\phi(u,v)=\left. x^2-3y\right|_{y=2u+v^2}^{x=u-v}=u^2-2v^2-2uv-6u$$

and finding its partial derivatives.

Remark. An example of applying the chain rule can be seen in the supplementary document "Jacobian of the inverse function".

Exercise. Do the same operations as in example 2, being $z=xy+y^3$, $x=2u+1$ and $y=u^2-v$.

Particular case. There are functions that depend on several variables, one of which, in turn, is a function of one of the others. For example:

$$z = g(x, v), \text{ being } v = f(x, y)$$

where z depends on x directly (through g) and also indirectly (through f).
To obtain the derivative of the composite function with respect to x, we proceed as follows:

$$z = g(x, v)|_{v=f(x,y)} = \phi(x, y) \Longrightarrow \frac{\partial \phi}{\partial x} = \frac{\partial g}{\partial x} + \frac{\partial g}{\partial v}\frac{\partial f}{\partial x}$$

That is, we derive with respect to x the terms that depend directly on that variable and we take into account the "intermediate" function f for those that depend on x through v.

Example. Let $z = x^2 + v^3$, being $v = \sin(4x + y^2)$.

$$z = \phi(x, y) = x^2 + \sin^3(4x + y^2) \Longrightarrow \phi'_x = 2x + 3\sin^2(4x + y^2)\cos(4x + y^2)\,4$$

The derivative with respect to y is

$$\phi'_y = 3\sin^2(4x + y^2)\cos(4x + y^2)\,2y = 6y\sin^2(4x + y^2)\cos(4x + y^2)$$

Exercise 1. Derive with respect to x and to y the function $z = \tan x + v^2$, being $v = \cos(x+2y)$.
Solution. $\phi'_x = 1 + \tan^2 x - \sin(2x + 4y)$; $\phi'_y = -2\sin(2x + 4y)$.

Exercise 2. Derive with respect to x and to y the function $z = x\,e^v$, being $v = xy^2$.
Solution. $\phi'_x = e^{xy^2} + xy^2\,e^{xy^2}$; $\phi'_y = 2x^2y\,e^{xy^2}$.

7 Higher-order derivatives

7.1 Definition

Next we study the higher-order derivatives of real functions of a vector variable.
Let f be a real function of n variables $x_1, x_2, \dots x_n$, defined on a neighborhood of the point $\vec{a}$.

$$f : D \to \mathbb{R}, \quad D \subset \mathbb{R}^n, \quad \vec{a} \in U_{\vec{a}} \subset D$$

If f is differentiable with respect to the variable x_i on $U_{\vec{a}}$, on that neighborhood there will exist the function partial derivative of f with respect to x_i, which we denote $\dfrac{\partial f}{\partial x_i}$ or f'_{x_i}.
If function $\dfrac{\partial f}{\partial x_i}$ is differentiable with respect to the variable x_j at the point $\vec{a}$, at that point there will exist the second partial derivative of f with respect to x_i and x_j, that is

$$\exists\ \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)\bigg|_{\vec{x}=\vec{a}} = \frac{\partial^2 f}{\partial x_i\,\partial x_j}\bigg|_{\vec{x}=\vec{a}}$$

And if this second partial derivative exists for all $\vec{x}$ of the domain, we have obtained the second partial derivative function of f with respect to x_i and x_j

$$\frac{\partial^2 f}{\partial x_i\,\partial x_j} = f''_{x_i x_j}$$

Repeating this process, we can obtain the partial derivatives of f, of any order.
In the case of two variables we have the following four second derivatives:

$$\frac{\partial^2 f}{\partial x^2} = f''_{xx}\,; \quad \frac{\partial^2 f}{\partial x \partial y} = f''_{xy}\,; \quad \frac{\partial^2 f}{\partial y \partial x} = f''_{yx}\,; \quad \frac{\partial^2 f}{\partial y^2} = f''_{yy}$$

7.2 Cross partial derivatives

In functions of several variables we can derive with respect to the same variables in different order. A particular case in functions of two variables are the second cross derivatives, that is, the second derivatives with respect to x and y, in the two possible orders (f''_{xy} and f''_{yx}).

In the derivable functions that we commonly use, we observe that the derivation order does not seem to matter. For example:

$$f(x,y) = x^2y^5 \Longrightarrow f''_{x\,y} = f''_{y\,x} = 10xy^4$$

and the same happens with $g(x,y) = e^{xy^2}$ or $h(x,y) = x \sin y^2$.
However, if we consider the function

$$f(x,y) = \begin{cases} \dfrac{xy^3}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

we observe that, outside the origin, it is verified (check it, as an exercise)

$$f''_{x\,y} = f''_{y\,x} = \frac{y^6 + 6y^4x^2 - 3y^2x^4}{(x^2+y^2)^3}$$

but $f''_{x\,y}(0,0) = 1,\ f''_{y\,x}(0,0) = 0$. Therefore the cross derivatives do not always coincide.

Remark. The second derivatives of the previous case have been obtained as follows:

1. We obtain f'_x and f'_y outside the origin, differentiating. And their values at $(0,0)$, by the definition.

2. We calculate the second derivatives at the origin by the definition, based on what has been obtained above:

$$\left.\frac{\partial^2 f}{\partial x \partial y}\right|_{(0,0)} = \left.\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right|_{(0,0)} = \lim_{k\to 0} \frac{\dfrac{\partial f}{\partial x}(0,k) - \dfrac{\partial f}{\partial x}(0,0)}{k}$$

The following theorems establish sufficient conditions for this to occur.

Theorem. *Let f be defined on $U_{\vec{a}} \subset \mathbb{R}^n$. If the partial derivatives of order k at $\vec{a}$ are continuous, their value does not depend on the order of derivation.* For example, if $n = 2$ and $f \in C^3$, then

$$\frac{\partial^3 f}{\partial x \partial y^2} = \frac{\partial^3 f}{\partial y \partial x \partial y} = \frac{\partial^3 f}{\partial y^2 \partial x}; \quad \frac{\partial^3 f}{\partial x^2 \partial y} = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial y \partial x^2}$$

Remark. If the k-th partial derivatives of f are continuous, then we say $f \in C^k$.

Schwarz theorem. *Let f defined on $U_{\vec{a}} \subset \mathbb{R}^2$. If the following derivatives exist on $U_{\vec{a}}$*

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial^2 f}{\partial x \partial y}$$

and the second derivative is continuous at $\vec{a}$, then $\dfrac{\partial^2 f}{\partial y \partial x}$ exists at $\vec{a}$ and:

$$\frac{\partial^2 f}{\partial x \partial y}(\vec{a}) = \frac{\partial^2 f}{\partial y \partial x}(\vec{a})$$

Notice that Schwarz's theorem is not a simple particular case of the above for $n = 2$, but it requires fewer conditions.

7.3 Higher-order differentials

We know that the differential is a linear function that accurately approximates Δf from the value of f at $\vec{a}$, if we are close enough to the point. As we saw in **4.4**, the differential of f is obtained as a linear combination of the differentials of the variables, the coefficients being the corresponding partial derivatives:

$$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}dx_i$$

If we change the point, df varies, since -in general- the partial derivatives of f depend on the point considered. This means that we can consider df a function and calculate its differential, that is, the second differential of f:

$$d(df) = d^2 f$$

Repeating the process we can obtain the higher-order differentials. This will help us to better approximate the increment of a function in the surroundings of a point, as well as to simplify some expressions, for example in the Taylor series expansion of a function.

Differential of order p of a function of two variables. Let $f(x, y)$ a differentiable function. Its differential is

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy$$

If df is differentiable, we can obtain its differential by multiplying the partial derivatives of df by the corresponding differencials of the variables:

$$d^2 f = d(df) = d\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right)dx + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right)dy =$$
$$\left(\frac{\partial^2 f}{\partial x^2}dx + \frac{\partial^2 f}{\partial y \partial x}dy\right)dx + \left(\frac{\partial^2 f}{\partial x \partial y}dx + \frac{\partial^2 f}{\partial y^2}dy\right)dy$$

Grouping the terms containing the products of the differentials of x and y, we obtain

$$\frac{\partial^2 f}{\partial x^2}dx^2 + \left(\frac{\partial^2 f}{\partial x \partial y} + \frac{\partial^2 f}{\partial y \partial x}\right)dxdy + \frac{\partial^2 f}{\partial y^2}dy^2$$

and, if the cross derivatives coincide, it results

$$\boxed{d^2 f = \frac{\partial^2 f}{\partial x^2}dx^2 + 2\frac{\partial^2 f}{\partial x \partial y}dxdy + \frac{\partial^2 f}{\partial y^2}dy^2}$$

This expression, which reminds us of the expansion of the square of a sum, can be abbreviated by using the square in **symbolic** form:

$$d^2 f = \left[\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right]^{(2}$$

It is shown by induction that this notation is valid for a differential of any order, hence

$$\boxed{d^p f = \left[\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right]^{(p}}$$

Example. The 3rd. order differential of $f(x, y)$ (using the Newton's binomial expansion) is:

$$d^3 f = \left[\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy\right]^{(3} = \frac{\partial^3 f}{\partial x^3}dx^3 + 3\frac{\partial^3 f}{\partial x^2 \partial y}dx^2dy + 3\frac{\partial^3 f}{\partial x \partial y^2}dxdy^2 + \frac{\partial^3 f}{\partial y^3}dy^3$$

Differential of order p of a function of n variables. It is also shown that the symbolic notation for higher order differentials is valid for any number of variables. Thus, the differential of order p of a function of n variables can be expressed as

$$d^p f(x_1, x_2, \dots x_n) = \left[\frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \cdots + \frac{\partial f}{\partial x_n}dx_n\right]^{(p}$$

Hessian matrix. We have seen in the example a particular case of the differential (2 variables, differential of order 3). Setting $n = 3$, $p = 2$, we obtain the second differential of a function of three variables

$$d^2 f = \left[\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz\right]^{(2}$$

whose expansion, assuming that cross derivatives are equal, is

$$\frac{\partial^2 f}{\partial x^2}dx^2 + \frac{\partial^2 f}{\partial y^2}dy^2 + \frac{\partial^2 f}{\partial z^2}dz^2 + 2\frac{\partial^2 f}{\partial x\partial y}dxdy + 2\frac{\partial^2 f}{\partial x\partial z}dxdz + 2\frac{\partial^2 f}{\partial y\partial z}dydz$$

which can be written in matrix form as

$$d^2 f = \begin{pmatrix} dx & dy & dz \end{pmatrix} \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x\partial y} & \dfrac{\partial^2 f}{\partial x\partial z} \\ \dfrac{\partial^2 f}{\partial y\partial x} & \dfrac{\partial^2 f}{\partial y^2} & \dfrac{\partial^2 f}{\partial y\partial z} \\ \dfrac{\partial^2 f}{\partial z\partial x} & \dfrac{\partial^2 f}{\partial z\partial y} & \dfrac{\partial^2 f}{\partial z^2} \end{pmatrix} \begin{Bmatrix} dx \\ dy \\ dz \end{Bmatrix}$$

This matrix is called a Hessian matrix. It is symmetric if the cross derivatives are equal.

Exercise 1. Write the differential of order 4 of a function of 2 variables $f(x, y)$.

Exercise 2. Find the differential of order 2 of the function $f(x, y)$, $f(x, y) = x^2y + xy^2$, particularizing at the points $P(1, -1)$, $Q(1, 0)$.

Solution. $d^2f|_P = -2dx^2 + 2dy^2$; $d^2f|_Q = 2dy^2 + 4dxdy$

8 Taylor expansion

Given a function of n variables, we want to calculate its value at a point $\vec{x}$, from the values at point $\vec{a}$ of f and its derivatives. As we did in functions of one variable, we can approximate this value by means of the Taylor polynomial of degree k or express it exactly as the sum of the Taylor polynomial of degree k plus the remainder term of order k (limited expansion of order k of the function f).

$$f(\vec{x}) \approx P_k(\vec{x}) \ ; \ f(\vec{x}) = P_k(\vec{x}) + T_k(\vec{x})$$

8.1 General expression

We will see first (without proof) the case of two variables, which will help us to obtain a simplified expression of the addends of order k and the remainder term. Then we will generalize to the case of n variables.

Function of 2 variables. Let the function f be defined on $D \subset \mathbb{R}^2$. The points (a, b) and (x, y) are interiors of D and we assume that $f \in C^{k+1}$, so it is $k+1$ times differentiable $(k \geq 1))$.

To obtain the Taylor expansion, we add the derivatives in increasing order, multiplied by the increments of the corresponding variables. For example, the third derivative with respect to x twice and to y once is multiplied by $(x-a)^2(y-b)$. Then:

- The first term of the polynomial (degree 0) is $f(a, b)$ (derivative of order 0).
- The terms of 1st. degree are

$$\left[\frac{\partial f}{\partial x}(a, b)(x-a)+\frac{\partial f}{\partial y}(a, b)(y-b)\right]$$

- The second derivatives of f are continuous, so the second cross derivatives are equal. The terms of degree 2 result

$$\frac{1}{2!}\left[\frac{\partial^2 f}{\partial x^2}(a, b)(x-a)^2+2 \frac{\partial^2 f}{\partial x \partial y}(a, b)(x-a)(y-b)+\frac{\partial^2 f}{\partial y^2}(a, b)(y-b)^2\right]$$

  and we write them as a symbolic square of the sum of terms of degree 1 (see **7.3**):

$$\left[\frac{\partial f}{\partial x}(a, b)(x-a)+\frac{\partial f}{\partial y}(a, b)(y-b)\right]^{(2}$$

  It can be shown by induction that also the addends of degrees $3, 4, \ldots$ can be written as the corresponding symbolic powers of this sum.

- The remainder term of order k takes the form

$$T_k(x, y)=\frac{1}{(k+1)!}\left[\frac{\partial f}{\partial x}(\xi_1, \xi_2)(x-a)+\frac{\partial f}{\partial y}(\xi_1, \xi_2)(y-b)\right]^{(k+1}$$

  being (ξ_1, ξ_2) an intermediate point between (a, b) and (x, y), that is

$$(\xi_1, \xi_2)=(a, b)+\theta(x-a, y-b), \quad 0<\theta<1$$

Thus, if $[\cdot]$ represents the sum of the first degree terms, the **limited expansion of order k of the function f** is

$$\boxed{f(x, y)=f(a, b)+\frac{1}{1!}[\cdot]+\frac{1}{2!}[\cdot]^{(2}+\frac{1}{3!}[\cdot]^{(3}+\cdots+\frac{1}{k!}[\cdot]^{(k}+T_k(x, y)}$$

Tangent plane. Analogously to what happens in one variable with the tangent line to the curve $y=f(x)$ at $x=a$, the terms of the expansion of degree 0 and 1 represent the equation of the tangent plane to the surface $z=f(x, y)$ at $P(a, b)$.

$$z=f(a, b)+f_x^{\prime}(a, b)(x-a)+f_y^{\prime}(a, b)(y-b)$$

Function of n variables. In this case, the expression is analogous to that for 2 variables, but $[\cdot]$ now contains the n partial derivatives with respect to $x_1, x_2, \ldots x_n$, at $\vec{a}=(a_1, a_2, \ldots, a_n)$:

$$\boxed{f(\vec{x})=f(\vec{a})+\frac{1}{1!}[\cdot]+\frac{1}{2!}[\cdot]^{(2}+\frac{1}{3!}[\cdot]^{(3}+\cdots+\frac{1}{k!}[\cdot]^{(k}+T_k(\vec{x})}$$

being

$$[\cdot]=\frac{\partial f}{\partial x_1}(\vec{a})(x_1-a_1)+\frac{\partial f}{\partial x_2}(\vec{a})(x_2-a_2)+\cdots+\frac{\partial f}{\partial x_n}(\vec{a})(x_n-a_n)$$

8.2 Matrix expression.

In the case $n = 2$, the terms of degree equal to or less than 2, written in matrix form, are

$$f(a,b) + \left(\frac{\partial f}{\partial x} \quad \frac{\partial f}{\partial y} \right)\Bigg|_{(a,b)} \begin{Bmatrix} x-a \\ y-b \end{Bmatrix} + \frac{1}{2!} \left(x-a \quad y-b \right) \begin{pmatrix} \dfrac{\partial^2 f}{\partial x^2} & \dfrac{\partial^2 f}{\partial x \partial y} \\ \dfrac{\partial^2 f}{\partial y \partial x} & \dfrac{\partial^2 f}{\partial y^2} \end{pmatrix}\Bigg|_{(a,b)} \begin{Bmatrix} x-a \\ y-b \end{Bmatrix}$$

The second addend is the escalar product of the gradient of f by the increment $\{(x, y) - (a, b)\}$ of the variable. The third addend is the Hessian matrix pre and post multiplied by the increment of (x, y). Then, for any n, the expression of the terms of the expansion of degree less than or equal to 2 is

$$\boxed{f(\vec{x}) = f(\vec{a}) + \vec{\nabla} f(\vec{a}) \cdot (\vec{x} - \vec{a}) + \frac{1}{2!} (\vec{x} - \vec{a})^t \underset{\sim}{H}|_{\vec{x}=\vec{a}} \{\vec{x} - \vec{a}\} + \dots}$$

where we write $(\vec{x} - \vec{a})^t$ to indicate that we premultiply the Hessian matrix by a row matrix and postmultiply it by a column matrix $\{\vec{x} - \vec{a}\}$.

Example. We obtain the Taylor polynomial of degree 3, at point $P(1, -2)$ of the function

$$f(x, y) = x^2 y + 3y - 2$$

- First derivatives: $f'_x = 2xy$; $f'_y = x^2 + 3$.
- Second derivatives: $f''_{xx} = 2y$; $f''_{yy} = 0$; $f''_{xy} = f''_{yx} = 2x$.
- Third derivatives: $f'''_{xxx} = 0$; $f'''_{yyy} = 0$; $f'''_{xxy} = f'''_{xyx} = f'''_{yxx} = 2$; $f'''_{yyx} = f'''_{yxy} = f'''_{xyy} = 0$.

We observe that the cross derivatives, both of second and third order, are equal; and that all the derivatives of third order are constant, so those of order higher than 3 are null.
We particularize the values of the variables at point P and write the terms of degree less than or equal to 2 in matrix form. It results

$$f(P) + \left(f'_x \quad f'_y \right) \begin{Bmatrix} x-1 \\ y+2 \end{Bmatrix} + \frac{1}{2!} \left(x-1 \quad y+2 \right) \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{yx} & f''_{yy} \end{pmatrix} \begin{Bmatrix} x-1 \\ y+2 \end{Bmatrix} + \frac{3 f'''_{xxy} (x-1)^2 (y+2)}{3!} =$$

$$-10 + \left(-4 \quad 4 \right) \begin{Bmatrix} x-1 \\ y+2 \end{Bmatrix} + \frac{1}{2} \left(x-1 \quad y+2 \right) \begin{pmatrix} -4 & 2 \\ 2 & 0 \end{pmatrix} \begin{Bmatrix} x-1 \\ y+2 \end{Bmatrix} + \frac{3 \cdot 2\, (x-1)^2 (y+2)}{6}$$

Operating, we obtain the following polynomial, in powers of $(x - 1)$ and $(y + 2)$

$$P_3(x, y) = -10 - 4(x-1) + 4(y+2) - 2(x-1)^2 + 2(x-1)(y+2) + (x-1)^2(y+2)$$

If we operate and simplify, we arrive to the initial function. This was to be expected, since the function is a polynomial of degree 3, so its Taylor polynomial of degree 3 coincides with f.

Exercise. Find the Taylor polynomial of degree 3 of the function f at $P(-1, 1)$

$$f(x, y) = xy^2 - x^2$$

Verify that, operating the expression of the polynomial, the function f is obtained.

Sol. $z = -2 + 3(x+1) - 2(y-1) - (x+1)^2 + 2(x+1)(y-1) - (y-1)^2 + (x+1)(y-1)^2$

9 Local extrema

We are going to study the local (or relative) extrema of a function of several variables. We will start assuming that there is an extremum and we will look for the points that satisfy a certain condition (necessary condition of extremum). Next, we will analyze the first terms of the Taylor expansion and deduce whether, at the previous points, there is a maximum, a minimum or neither of the two; thus we will have a sufficient condition for the existence of an extremum.

9.1 Necessary condition

Let function f be

$$f : D \to \mathbb{R}, \quad D \subset \mathbb{R}^n, \quad f \in C^k, \, k \geq 2$$

Suppose that f has a relative minimum at $\vec{x} = \vec{a}$, being $\vec{a}$ an interior point of the domain (for a maximum we would proceed the same way). Then, there will exist a punctured neighborhood of the point at which the values of f will be greater than $f(\vec{a})$, that is

$$\exists\, U^*_{\vec{a}} \,/\, f(\vec{a}) < f(\vec{x}) \;\forall \vec{x} \in U^*_{\vec{a}}$$

(we are assuming strict minimum; otherwise we would use $\leq$).

If $\vec{a} = (a_1, a_2, \ldots, a_n)$ y $\vec{x} = (x_1, x_2, \ldots, x_n)$, we will have

$$f(a_1, a_2, \ldots, a_n) < f(x_1, x_2, \ldots, x_n) \;\; \forall (x_1, x_2, \ldots, x_n) \in U^*_{\vec{a}} \subset D$$

which means that f grows when modifying any of the variables x_i. Then, for each of them there will be an interval $U^*_{a_i}$ in which

$$\begin{array}{c} f(a_1, a_2, \ldots, a_n) < f(x_1, a_2, \ldots, a_n) = \phi_1(x_1), \;\forall x_1 \in U^*_{a_1} \\ f(a_1, a_2, \ldots, a_n) < f(a_1, x_2, \ldots, a_n) = \phi_2(x_2), \;\forall x_2 \in U^*_{a_2} \\ \vdots \\ f(a_1, a_2, \ldots, a_n) < f(a_1, a_2, \ldots, x_n) = \phi_n(x_n), \;\forall x_n \in U^*_{a_n} \end{array}$$

That is, we have n functions of one variable, $f(a_1, a_2, \ldots, x_i, \ldots a_n) = \phi_i(x_i)$, each with a minimum at an interior point.

The relative extremum theorem (for functions of one variable) says that *"if a function f is derivable on an interval I and it has an extremum at a point a, interior of I, then its derivative at a is null"*. Therefore,

$$\left.\frac{d\phi_1}{dx_1}\right|_{x_1=a_1} = \left.\frac{d\phi_2}{dx_2}\right|_{x_2=a_2} = \cdots = \left.\frac{d\phi_n}{dx_n}\right|_{x_n=a_n} = 0$$

But each of these derivatives is equivalent to the partial derivative of f with respect to the corresponding variable, particularized at $\vec{x} = \vec{a}$, hence

$$\left.\frac{\partial f}{\partial x_1}\right|_{\vec{x}=\vec{a}} = \left.\frac{\partial f}{\partial x_2}\right|_{\vec{x}=\vec{a}} = \cdots = \left.\frac{\partial f}{\partial x_n}\right|_{\vec{x}=\vec{a}} = 0$$

Or what is the same

$$\boxed{\vec{\nabla} f(\vec{a}) = \vec{0}}$$

which is a **necessary condition for the existence of an extremum** at the point $\vec{a}$. The points that verify this condition are called **critical points**.

Since we have assumed that f is differentiable and $\vec{a}$ an interior point, we must also check the **boundary points and those in which f is not differentiable**, where there could be an extremum (see the supplementary document "Problem of extrema with 2 variables").

9.2 Sufficient condition

To determine the type of extremum, we will use the first three terms of the Taylor expansion in matrix form. Since the second derivatives are continuous, the Hessian matrix is symmetric.

$$f(\vec{x}) = f(\vec{a}) + \vec{\nabla} f(\vec{a}) \cdot (\vec{x} - \vec{a}) + \frac{1}{2!} (\vec{x} - \vec{a})^t \underset{\sim}{H}|_{\vec{x}=\vec{a}} (\vec{x} - \vec{a}) + \dots$$

If we call $d\vec{x}$ the increment $\vec{x} - \vec{a}$ and take the $f(\vec{a})$ to the other member, we have

$$f(\vec{x}) - f(\vec{a}) = \vec{\nabla} f(\vec{a}) \cdot d\vec{x} + \frac{1}{2} d\vec{x}^{\,t} \underset{\sim}{H}|_{\vec{x}=\vec{a}} d\vec{x} + \dots$$

where we see that:

- The first member represents the increment of the value of the function between $\vec{a}$ y $\vec{x}$.
- The first term of the right hand side is the differential $df(\vec{a})$ of the function.
- The second term is the second differential $d^2 f(\vec{a})$, multiplied by $1/2$.
- The following terms are infinitesimals of higher order: if $\vec{x}$ is very close to $\vec{a}$, they are negligible compared to the second.

If the point $\vec{a}$ satisfies the necessary condition of extremum, $\vec{\nabla} f(\vec{a})$ is null, so the sign of Δf is given by the first non-zero addend, $d^2 f(\vec{a})$. Note that it does not give the exact value of Δf due to higher order infinitesimals. This term is discussed next.

The second differential $d^2 f$ is a quadratic form, which associates a real number to a vector

$$d\vec{x} \to d\vec{x}^{\,t} \, \underset{\sim}{H} \, d\vec{x}$$

We classify a quadratic form according to the sign that it takes $\forall \, d\vec{x} \neq 0$. Thus, $d^2 f$ will be:

- Positive definite $\iff d^2 f > 0$.
- Negative definite $\iff d^2 f < 0$.
- Positive semidefinite $\iff d^2 f \geq 0$.
- Negative semidefinite $\iff d^2 f \leq 0$.
- Indefinite $\iff d^2 f \gtrless 0$ (it may be null, but it is not necessary).

So, depending on the type of quadratic form that $d^2 f$ is at $\vec{a}$, we will have:

- If $d^2 f$ is positive definite, then $\Delta f > 0$ always; hence there is a minimum at $\vec{a}$.
- If $d^2 f$ is negative definite, then $\Delta f < 0$ always; hence there is a maximum at $\vec{a}$.
- If $d^2 f$ is indefinite, we can have $\Delta f \gtrless 0$; hence there is neither a maximum nor a minimum.

If $d^2 f$ is semidefinite, there are several options. We analyze it for PSD (with NSD it is analogous).

- In this case $d^2 f \geq 0$; that is, for some $d\vec{x}$, $d^2 f > 0 \Rightarrow \Delta f > 0$; but, for others, $d^2 f = 0$, in which case the sign of Δf is given by the higher order terms and Δf can be $\substack{> \\ = \\ <} 0$.
- We can only be sure that there is no maximum: since $d^2 f \geq 0$, for some $d\vec{x}$ it must be $d^2 f > 0 \Rightarrow \Delta f > 0$, so the function grows. Therefore there can be no maximum.
- A supplementary document discusses the function $f(x, y) = x^2 + ky^4$, which has a critical point at $(0, 0)$. It is shown that, depending on the sign of k, different results are obtained.

9.3 Determination of the type of quadratic form

We will see two ways to classify a quadratic form, from its associated matrix.

a) Sylvester's criterion. A quadratic form is:

- Positive definite if and only if the leading principal minors of the matrix are positive.

- Negative definite if and only if the signs of the leading principal minors are $-, +, -, \dots$

Remark. Given a $n \times n$ matrix, its leading principal minor of order k is the $k \times k$ minor obtained by deleting the last $n-k$ rows and columns. Thus, the l.p.m. of order $1, 2, 3, \dots$ are the upper left 1×1 corner, the upper left 2×2 corner, the upper left 3×3 corner,...

b) Diagonalization by congruence. It can be proved that every symmetric matrix H is diagonalizable by congruence, that is

$$H = H^t \Longrightarrow \exists\, C \left(|C| \neq 0 \right) / \, CHC^t = D$$

According to the signs of the elements of the main diagonal of D, the quadratic form will be:

- Positive definite $\Longleftrightarrow$ $d_i > 0$, $i = 1, 2, \dots, n$

- Negative definite $\Longleftrightarrow$ $d_i < 0$, $i = 1, 2, \dots, n$

- Positive semidefinite $\Longleftrightarrow$ $d_i \geq 0$, $i = 1, 2, \dots, n$

- Negative semidefinite $\Longleftrightarrow$ $d_i \leq 0$, $i = 1, 2, \dots, n$

- Indefinite $\Longleftrightarrow$ $d_i \gtreqless 0$ (there may be some null).

Remark. Square matrices satisfy that *"the determinant of the product is equal to the product of the determinants"* and *"the determinant of a matrix is equal to that of its transposed"*, so

$$|D| = |C|\,|H|\,\left|C^t\right| = |C|^2\,|H|$$

Hence the determinants of the Hessian matrix and its diagonal matrix have the same sign, which will be useful when classifying the quadratic form.

We show below the recommended steps in the study of extrema.

9.4 Search of extrema. Summary and examples

Consider a function $f : \mathrm{D} \to \mathbb{R}$, $\mathrm{D} \subset \mathbb{R}^n$, $f \in C^2$. To find its extrema on the interior of the domain $\overset{\circ}{D}$, we take into account the necessary and the sufficient conditions.

Necessary condition. Since the set is open and f is differentiable, its extrema will be at points that satisfy the null gradient condition. Then the possible extrema (critical points) are given by the solutions of the equation

$$\boxed{\vec{\nabla} f = \vec{0}}$$

Sufficient condition. To find out which critical points are extrema, the first step is to determine what type is the quadratic form $d^2 f$ at them. We know that:

- If it is definite, there is a minimum (positive) or a maximum (negative).

- If it is indefinite, there is neither maximum nor minimum (saddle point).

- If it is semidefinite, we only know that there can be no maximum (if it is positive SD) or there can be no minimum (if it is negative SD).

So, we follow these steps:

a) **¿Is H definite?** It is found by applying the **Sylvester criterion**. So, if it is positive definite, there exists a **minimum**. If it is negative definite, a **maximum**.

b) If H is **not definite**, studying its determinant there are two options:

b.1. $\boxed{|H| \neq 0}$. Since $|D| = |C|^2\,|H|$, it turns out that $|D| = d_1 d_2 \cdots \neq 0$, which means that no d_i is null, so they will have equal or different signs. If they had the same sign, H would be definite, which is not the case. Then there must exist some $d_i \gtrless 0$, so H is **indefinite** and we have a **saddle point**.

b.2. $\boxed{|H| = 0}$. In this case $|D| = 0$ and the d_i can be $\geq 0, \leq 0$ or $\lesseqgtr 0$ (some must be null), so H can be **semidefinite** or **indefinite**.

In these cases we can find out what type of quadratic form it is by diagonalizing by congruence.

In the particular case $n = 2$, there are only two d_i, one of which at least is null. Then it cannot be indefinite, so it will be semidefinite.

c) H **semidefinite**: To solve these inconclusive cases we can move from $\vec{a}$ along different straight lines: if f increases in one direction but decreases in another, it is a **saddle point**. The same happens if, along the same straight line, it increases in one direction and decreases in the other. In the case $n = 2$, we can move from (a, b) parallel to the axes, to the bisectors $y = x,\ y = -x$, etc.

Examples.

a) $f(x, y) = x^2 + y^2 + xy - x - 2y$.

Applying the necessary condition, we obtain the critical point $P(0, 1)$. The Hessian matrix $H_{(0,1)} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$ is positive definite (Sylvester), therefore it is a minimum.

b) $f(x, y) = xy + x - y - 1$.

Applying the necessary condition, the critical point is $P(1, -1)$ and the corresponding Hessian matrix is $H_{(1,-1)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. In this case H is not definite (Sylvester). Since $|H| \neq 0$, it turns out that H is indefinite, so we have a saddle point.

c) $f(x, y) = \sqrt{R^2 - x^2}$.

Applying the N.C., we get $P(0, y)$ (all points on the OY axis are critical). In all of them: $H = \begin{pmatrix} -\frac{1}{R} & 0 \\ 0 & 0 \end{pmatrix} \Longrightarrow |H| = 0 \Longrightarrow H$ semidefinite $(n = 2)$.

To solve it, we consider the second differential $d^2 f = -\dfrac{1}{R}\, dx^2 \leq 0$, which takes a negative value $\forall\, dx \neq 0$. Then the value of f decreases if we move, from the points $(0, y)$, in any direction, except the one given by $dx = 0$. But this means moving along the OY axis, where the value of the function is constant and its value is R. That is, at every point $(0, y)$ -the OY axis- there is a maximum of value $z = R$.

To see it graphically we write the equation in the form $z^2 + x^2 = R^2$, which corresponds to a cylinder of radius R and horizontal axis (the OY axis). The expression of $f(x)$ corresponds to the part of the cylinder above the plane XY $(z \geq 0)$.

d) $f(x,y) = y^3 - x^2 - 2x - 1$.

Applying the N.C., $P(-1,0)$. The Hessian matrix is

$H_{(-1,0)} = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}$. H is diagonal, so it is negative semidefinite.

To solve it we study -as in the previous case- the second differential $d^2f = -2dx^2$. This tells us that the value of f decreases as we move in any direction except the OY axis. Studying this case (variation only of y), we see that f increases if y increases and decreases if y decreases, therefore there is a saddle point at P.

Another option is to move in different directions starting from $P(-1,0)$, where f is null: for example in those of the axis, alternately increasing x and y by a value Δ. It results:

d.1. At any point of coordinates $(-1+\Delta, 0)$, the function takes the value $-\Delta^2$, then decreasing from its value in P, regardless of the sign of Δ.

d.2. At any point of coordinates $(-1, \Delta)$, the value of f is Δ^3. That is, it increases if we move from P in the positive direction of the OY axis ($\Delta > 0$) and decreases otherwise.

We conclude that, at $P(-1,0)$, f has neither a maximum nor a minimum (saddle point).

e) $f(x,y,z) = 3x^2 - 6x + y^4 - z^4$.

Applying the N. C., we get $P(1,0,0)$ and $H_{(1,0,0)} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

H is positive semidefinite and the second differential is $d^2f = 6dx^2$, so the function will grow whenever $dx \neq 0$. If, on the other hand, we move from P, parallel to the plane YZ ($dx = 0$), the second differential is null. To identify the type of extremum, we move parallel to YZ from P, modifying the value of one of the variables y, z each time:

e.1. Value of f at P: $f(1,0,0) = -3$.

e.2. We increase the value of y: $f(1,\Delta,0) = -3 + \Delta^4 \Longrightarrow \Delta f = \Delta^4$, so f increases.

e.3. We increase the value of z: $f(1,0,\Delta) = -3 - \Delta^4 \Longrightarrow \Delta f = -\Delta^4$, so f decreases.

It is, then, a saddle point. We have not tried to modify the value of x, because, as it has been said, the function grows whenever $dx \neq 0$.

10 Implicit function

From an equation that relates two variables, $F(x,y) = 0$, sometimes we can solve for one of them in terms of the other. For example:

$$x^3y - \sin x = 0 \Longrightarrow y = \frac{\sin x}{x^3},\ \forall x \neq 0$$

In other cases it is not possible to get an explicit expression of y in terms of x, e.g.:

$$x \sin y - y - x^2 = 0 \Longrightarrow \ ?$$

However, it may happen that, for all x, there is a unique y that verifies the equation. This means that y is a function of x, even though there is no explicit expression relating them. In this case we say that y is an implicit function of x.

Next we will give a definition of an implicit function and state the existence and differentiability theorem for two variables. Then we will generalize to the case of n variables.

10.1 Definition

$F(x,y) = 0$ defines y as an implicit function of x on the interval $I \subset \mathbb{R}$ if and only if, for all $x \in I$, there exists a unique value $\psi(x)$ that satisfies the equation:

$$\boxed{\forall x \in I \; \exists\, \psi(x) \in \mathbb{R} \,/\, F\big(x, \psi(x)\big) = 0}$$

10.2 Existence and differentiability theorem for two variables

We state it in two parts:

a) Let $A \subset \mathbb{R}^2$ be an open set. Let $F(x,y)$ be a continuous function on A, such that its derivative with respect to y is also continuous. We assume that there exists a point $P(a,b) \in A$ such that the function is null at P, but its derivative is not null there, that is

$$F(x,y),\, \frac{\partial F}{\partial y}, \text{ continuous on } A; \quad \exists (a,b) \,/\, F(a,b) = 0,\; \frac{\partial F}{\partial y}(a,b) \neq 0$$

In this case, the equation $F(x,y) = 0$ defines a **unique** continuous function $y = \psi(x)$ as an implicit function of x, on a neighborhood of $x = a$, $y = b$ (U_a).

b) If, in addition, the derivative of F with respect to x is continuous on A, then $y = \psi(x)$ is differentiable on U_a and -using the chain rule- it holds

$$F\big(x, \psi(x)\big) = \phi(x) = 0 \Longrightarrow \frac{d\phi}{dx} = \left.\frac{\partial F}{\partial x}\right|_{y=\psi(x)} + \left.\frac{\partial F}{\partial y}\frac{d\psi}{dx}\right|_{y=\psi(x)} = 0$$

from where

$$\boxed{\frac{d\psi}{dx} = -\left.\frac{\partial F/\partial x}{\partial F/\partial y}\right|_{y=\psi(x)}}$$

That is, although there may not be an explicit expression $y = \psi(x)$, we can know its derivative on a certain U_a (in particular at $x = a$). For example, we can obtain the tangent to the curve defined by $F(x,y) = 0$, without knowing the equation of the curve.

Example. Let $F(x,y) = x^2 + y^2 - 2 = 0$. Its partial derivatives are $F'_x = 2x, F'_y = 2y$. Doing $y = x$ in the equation, we obtain point $P\,(1,1)$ in which F is null and F'_y is not null. Then there exists a neighborhood of $x = 1$, $y = 1$, where

$$\frac{d\psi}{dx} = -\frac{F'_x}{F'_y} = -\frac{x}{y} \quad \left(\text{in particular } \left.\frac{d\psi}{dx}\right|_P = -1\right)$$

Note that $F(x,y) = 0$ is the equation of a circle, from which we obtain two possible functions $y = \psi(x)$, corresponding to the semicircles above and below OX.

We also see that all the points of the circle can be chosen as P, except $(1,0)$ and $(-1,0)$, in which $F'_y = 0$. The reason is that, in every neighborhood of those points, there are **two** implicit functions (the aforementioned semicircles). Then the theorem is not verified in them, since it guarantees the existence of a single implicit function on a certain neighborhood of the point.

Exercise 1. Given the equation $F(x,y) = y^2 \cos x - x^2 \cos y = 0$:

a) Check that an implicit f. exists on a neighborhood of $P(\pi/2, \pi/2)$ and find $\partial\psi/\partial x$ at P.

b) Identify the function, which in this case has a very simple explicit form.

c) At $Q(0,0)$ the condition $F'_y \neq 0$ is not satisfied. Find a reason for it.

Exercise 2. Apply the theorem to $F(x,y) = \sin y + \sec y - x = 0$, finding the point $P(a,b)$.

10.3 Generalization of the theorem to vector functions

We have studied one equation with two variables that defines one of them as an implicit function of the other, $y = \psi(x)$. We are going to state the theorem in the general case: m equations with $m + n$ variables that define m of them as an implicit (vector) function of the remaining n variables (vector variable). First we will see, as an introduction, a case with 3 variables.

Particular case. Consider the equation $g(u, v, y) = 0$ (to simplify, we will not take into account domains of existence). Analogously to what is defined in **10.1**, we can say that $g(u, v, y) = 0$ defines y as an implicit function of u and v if and only if

$$\forall(u, v)\, \exists\psi(u, v) \,/\, g\big(u, v, \psi(u, v)\big) = \phi(u, v) = 0$$

If the sufficient conditions for it are satisfied, $\psi(u, v)$ will be differentiable. Finding the partial derivatives of ϕ and operating we obtain the derivatives of ψ

$$\frac{\partial\phi}{\partial u} = \frac{\partial g}{\partial u} + \frac{\partial g}{\partial y}\frac{\partial\psi}{\partial u} = 0 \Longrightarrow \boxed{\frac{\partial\psi}{\partial u} = -\frac{\partial g/\partial u}{\partial g/\partial y}}$$
$$\frac{\partial\phi}{\partial v} = \frac{\partial g}{\partial v} + \frac{\partial g}{\partial y}\frac{\partial\psi}{\partial v} = 0 \Longrightarrow \boxed{\frac{\partial\psi}{\partial v} = -\frac{\partial g/\partial v}{\partial g/\partial y}}$$

provided that the partial derivative of g with respect to y is not null.

General case. We now consider a function of m components $(g_1, g_2, \dots g_m)$ and $m+n$ variables $(x_1, \dots x_n, y_1 \dots y_m)$, which we can express briefly as $\vec{g}(\vec{x}, \vec{y})$.
The conditions that $\vec{g}$ must satisfy are analogous to those required to the function of two variables $F(x, y)$ in **10.2**:

- $\vec{g}$ continuous; that is, the components g_k are continuous on $\mathrm{A} \subset \mathbb{R}^{n+m}$, $k = 1, \dots m$.
- Partial derivatives $\dfrac{\partial g_k}{\partial x_i}, \dfrac{\partial g_k}{\partial y_j}$ continuous on $\mathrm{A} \subset \mathbb{R}^{n+m}$, $i = 1, \dots n;\ j, k = 1, \dots m$.
- There exists a point $(a_1, \dots a_n, b_1, \dots b_m) = (\vec{a}, \vec{b}) \,/\, \vec{g}(\vec{a}, \vec{b}) = \vec{0}$, $\left|\dfrac{\partial\vec{g}}{\partial\vec{y}}\right|_{(\vec{a},\vec{b})} \neq 0$.

Theorem: Under the stated conditions, "*the equation $\vec{g} = \vec{0}$ defines, as an implicit function of $\vec{x}$, a unique function $\vec{y} = \vec{\psi}(\vec{x})$, on a neighborhood of $\vec{x} = \vec{a}$, $\vec{y} = \vec{b}$ $(U_{\vec{a}})$, differentiable on $U_{\vec{a}}$*".
This means that:

a) The implicit function $\vec{\psi}$ verifies the equation $\vec{g} = \vec{0}$:

$$\boxed{\vec{g}(\vec{x}, \vec{y}) = \vec{0} \Rightarrow \exists\, \vec{y} = \vec{\psi}(\vec{x}) \,/\, \vec{g}\left(\vec{x}, \vec{\psi}(\vec{x})\right) = \vec{\phi}(\vec{x}) = \vec{0}}$$

The above expression in vector form means that there exist m functions $\psi_1, \psi_2, \dots, \psi_m$, all of them of n variables $(x_1, \dots x_n)$, which satisfy the m equations $g_i = 0$. This gives rise to m functions of the same n variables which are null on $U_{\vec{a}}$. That is

$$\begin{gathered}
g_1(x_1, \dots x_n, \psi_1, \dots, \psi_m) = \phi_1(x_1, \dots x_n) = 0 \\
g_2(x_1, \dots x_n, \psi_1, \dots, \psi_m) = \phi_2(x_1, \dots x_n) = 0 \\
\vdots \\
g_m(x_1, \dots x_n, \psi_1, \dots, \psi_m) = \phi_m(x_1, \dots x_n) = 0
\end{gathered}$$

b) To calculate the derivatives of the functions ϕ_k with respect to the different variables we use the chain rule. The derivative of ϕ_k with respect to x_i will be

$$\frac{\partial \phi_k}{\partial x_i} = \frac{\partial g_k}{\partial x_i} + \frac{\partial g_k}{\partial y_1}\frac{\partial \psi_1}{\partial x_i} + \cdots + \frac{\partial g_k}{\partial y_m}\frac{\partial \psi_m}{\partial x_i} = \frac{\partial g_k}{\partial x_i} + \sum_{j=1}^{m} \frac{\partial g_k}{\partial y_j}\frac{\partial \psi_j}{\partial x_i} = 0$$

Taking k as row index and i as column index, we see that $\partial\phi_k/\partial x_i$ represents the element (k, i) of the matrix of the partial derivatives of $\vec{\phi}$ with respect to $\vec{x}$. In the same way, $\partial g_k/\partial x_i$ represents the element (k, i) of the matrix of the partial derivatives of $\vec{g}$ with respect to $\vec{x}$.

On the other hand, the sum between $j = 1$ and $j = m$ in the second member represents the product of a row times a column: the row k of the matrix of the derivatives of $\vec{g}$ with respect to $\vec{y}$ and the column i of the matrix of the derivatives of $\vec{\psi}$ with respect to $\vec{x}$.

Taking the above into account, we can write the expression in matrix form and operate, obtaining

$$\boxed{\frac{d\vec{\phi}}{d\vec{x}} = \frac{\partial \vec{g}}{\partial \vec{x}} + \frac{\partial \vec{g}}{\partial \vec{y}}\frac{d\vec{\psi}}{d\vec{x}} = \Omega_{m\times n}} \Longrightarrow \boxed{\frac{d\vec{\psi}}{d\vec{x}} = -\left(\frac{\partial \vec{g}}{\partial \vec{y}}\right)^{-1} \left.\frac{\partial \vec{g}}{\partial \vec{x}}\right|_{\vec{y}=\vec{\psi}(\vec{x})}}$$

which gives us the matrix of the derivatives of the implicit function $\vec{\psi}$ with respect to the variable $\vec{x}$. In these expressions, all matrices have dimension $m \times n$, except the derivative of $\vec{g}$ with respect to $\vec{y}$, which is square $(m \times m)$.

Note that the resulting vector expression corresponds exactly to the one we obtained in 2 variables, using vectors $\vec{x}$, $\vec{y}$ and $\vec{\psi}$ instead of scalars and the function $\vec{g}$ instead of F.

Example. We are going to apply the obtained expressions to the particular case that we solved as an introduction.

We express the equation $g(u, v, y) = 0$, as $\vec{g}(\vec{x}, \vec{y}) = \vec{0}$, where

$$\vec{x} = (u, v); \quad \vec{y} = y; \quad \vec{g} : \mathbb{R}^{n+m} \to \mathbb{R}^m \ \text{ with } n = 2,\ m = 1$$

The condition $\vec{g} = \vec{0}$ defines the implicit function $\vec{\psi}$

$$\vec{g}(\vec{x}, \vec{y}) = \vec{0} \Rightarrow \vec{y} = \vec{\psi}(\vec{x})$$

where $\vec{g}, \vec{y}, \vec{\psi}$ are now scalars.

Applying the general expression to our case, the matrix of derivatives of $\vec{g}$ with respect to $\vec{y}$ has a single element, so its inverse is $(\partial g/\partial y)^{-1}$. And the matrix of the derivatives of $\vec{g}$ with respect to $\vec{x}$ contains the derivatives of g with respect to u and v. Then the derivative of the implicit function $\vec{\psi}$ with respect to $\vec{x}$ is

$$\frac{d\vec{\psi}}{d\vec{x}} = \left(\frac{\partial \psi}{\partial u}, \frac{\partial \psi}{\partial v}\right) = -\frac{1}{\partial g/\partial y}\left(\frac{\partial g}{\partial u}, \frac{\partial g}{\partial v}\right) = -\left(\frac{\partial g/\partial u}{\partial g/\partial y}, \frac{\partial g/\partial v}{\partial g/\partial y}\right)$$

and the expressions of the derivatives of ψ with respect to u and v coincide with those obtained at the beginning. This suggests that sometimes the vector notation is not necessary and we can proceed as we did in the particular case.

It is recommended to see the problems solved in the supplementary document "Theorem of the implicit function. Examples".

11 Constrained extrema

In section **9** we have studied extrema of vector functions whose variables can "move" freely through their respective domains. Next we will study extrema of functions subject to some **equality constraints**, i.e. the variables of the function must verify certain conditions.

Suppose that $f(x,y)$ has an extremum at $P(a,b)$. If we impose a condition that relates x and y, it may happen that the extremum at P disappears and another one appears at P'.

Example. $f(x,y) = 1 - x^2 - y^2$ reaches its maximum value for $x = y = 0$. If we make the variables to satisfy the condition $x + y - 1 = 0$, the maximum is reached for $x = y = 1/2$.

11.1 Constraints in explicit form

We consider functions of $m+n$ variables, constrained by m explicit conditions such that m variables depend explicitly on the other n. That is, let the function f be

$$f(x_1,\dots,x_n,y_1\dots,y_m)$$

with the constraints

$$\begin{aligned} y_1 &= \psi_1(x_1,\dots,x_n) \\ &\ \vdots \\ y_m &= \psi_m(x_1,\dots,x_n) \end{aligned}$$

If we write the $m+n$ variables as $\vec{x} = (x_1,\dots,x_n)$ and $\vec{y} = (y_1\dots,y_m)$, the function becomes

$$f(\vec{x},\vec{y}), \ \text{ being } \ \vec{y} = \vec{\psi}(\vec{x})$$

It is immediate that, replacing in f the variables y_i by their respective expressions $\psi_i(\vec{x})$, the original function with the constraints results in a new function of only n variables

$$f(\vec{x},\vec{y})\big|_{\vec{y}=\vec{\psi}(\vec{x})} = f(\vec{x},\vec{\psi}(\vec{x})) = \phi(\vec{x})$$

from which we can obtain the extrema by the method described in **9**. This would be the case of the previous example, if we introduce the condition $y = 1 - x$ in $f(x,y) = 1 - x^2 - y^2$.

11.2 Constraints in implicit form

Now we want to obtain the extrema of the same function $f(\vec{x},\vec{y})$, but in this case the m conditions will be given implicitly by m equations

$$\begin{aligned} &g_1(x_1,\dots,x_n,y_1\dots,y_m) = 0 \\ &\ \vdots \\ &g_m(x_1,\dots,x_n,y_1\dots,y_m) = 0 \end{aligned}$$

that can be written more briefly as $\vec{g}(\vec{x},\vec{y}) = \vec{0}$.

Method of the Lagrange multipliers. It can be shown that there exist certain values $\lambda_i \in \mathbb{R}$, such that the extrema of the function f, with the m constraints between the variables, coincide with those of the Lagrangian function

$$L = f(\vec{x},\vec{y}) + \sum_{i=1}^{m} \lambda_i \, g_i(\vec{x},\vec{y})$$

which allows us to simplify the search of the extrema when the conditions are not explicit.

Necessary condition. We will justify it for a function of three variables subject to two constraints ($n = 1$, $m = 2$). We will see that, to obtain the critical points of f with the constraints $g_1 = 0$, $g_2 = 0$, we must satisfy the same conditions as for those of L. Let f, g_1, g_2 be

$$u = f(x, y, z); \quad g_1(x, y, z) = 0; \quad g_2(x, y, z) = 0$$

The Lagrangian function is

$$L = f(x, y, z) + \lambda_1 g_1(x, y, z) + \lambda_2\, g_2(x, y, z)$$

To find the critical points of L, the following conditions must be satisfied:

$$\begin{aligned} \frac{\partial L}{\partial x} &= \boxed{\frac{\partial f}{\partial x} + \lambda_1 \frac{\partial g_1}{\partial x} + \lambda_2 \frac{\partial g_2}{\partial x} = 0} \\ \frac{\partial L}{\partial y} &= \boxed{\frac{\partial f}{\partial y} + \lambda_1 \frac{\partial g_1}{\partial y} + \lambda_2 \frac{\partial g_2}{\partial y} = 0} \\ \frac{\partial L}{\partial z} &= \boxed{\frac{\partial f}{\partial z} + \lambda_1 \frac{\partial g_1}{\partial z} + \lambda_2 \frac{\partial g_2}{\partial z} = 0} \end{aligned} \qquad (*)$$

Let us now study the function f. By the implicit function theorem we know that, under certain conditions, the equalities $g_1 = 0$, $g_2 = 0$ define two implicit functions ψ_1, ψ_2 that satisfy them:

$$\left.\begin{array}{l} g_1(x, y, z) = 0 \\ g_2(x, y, z) = 0 \end{array}\right\} \Longrightarrow \exists\, \psi_1(x), \psi_2(x) \,/\, \left\{\begin{array}{l} g_1\big(x, \psi_1(x), \psi_2(x)\big) = 0 \\ g_2\big(x, \psi_1(x), \psi_2(x)\big) = 0 \end{array}\right\}$$

Replacing y and z by their corresponding functions ψ_1, ψ_2, f becomes a new function

$$f\big(x, \psi_1(x), \psi_2(x)\big) = \phi(x)$$

that we will derive with respect to x to find its possible extrema.
Since the functions $g_1\big(x, \psi_1(x), \psi_2(x)\big)$, $g_2\big(x, \psi_1(x), \psi_2(x)\big)$ are null for all x of a certain set (therefore, constant on it), their derivatives with respect to x will also be constant.

Using the chain rule to derive f, g_1 and g_2, we obtain

$$\begin{aligned} \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{d\psi_1}{dx} + \frac{\partial f}{\partial z}\frac{d\psi_2}{dx} &= 0 \\ \frac{\partial g_1}{\partial x} + \frac{\partial g_1}{\partial y}\frac{d\psi_1}{dx} + \frac{\partial g_1}{\partial z}\frac{d\psi_2}{dx} &= 0 \\ \frac{\partial g_2}{\partial x} + \frac{\partial g_2}{\partial y}\frac{d\psi_1}{dx} + \frac{\partial g_2}{\partial z}\frac{d\psi_2}{dx} &= 0 \end{aligned}$$

In these three equations there are only two unknowns, the derivatives of ψ_1 and ψ_2. For the system to have a solution, one of the equations must be a linear combination of the others. That is, there must exist $\alpha, \beta, \gamma \in \mathbb{R}$ such that

$$\alpha \begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix} + \beta \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} & \frac{\partial g_1}{\partial z} \end{pmatrix} + \gamma \begin{pmatrix} \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} & \frac{\partial g_2}{\partial z} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$$

The coefficient α must be not null. Otherwise the derivatives of g_1 and g_2 would be proportional and the condition "determinant of $\partial\vec{g}/\partial\vec{y}$ not null" would not be satisfied (see **10.3**). Dividing the equation by α and setting $\beta/\alpha = \lambda_1, \gamma/\alpha = \lambda_2$, it results

$$\begin{pmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{pmatrix} + \lambda_1 \begin{pmatrix} \frac{\partial g_1}{\partial x} & \frac{\partial g_1}{\partial y} & \frac{\partial g_1}{\partial z} \end{pmatrix} + \lambda_2 \begin{pmatrix} \frac{\partial g_2}{\partial x} & \frac{\partial g_2}{\partial y} & \frac{\partial g_2}{\partial z} \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \end{pmatrix}$$

Equating the components in both members, we arrive to the same necessary condition $(*)$ obtained for the Lagrangian function.

Application of the method. Let a function f be $f(x_1, \ldots, y_m)$ and the constraints

$$g_1(x_1, \ldots, y_m) = 0\,,\; g_2(x_1, \ldots, y_m) = 0\,,\; \ldots\; g_m(x_1, \ldots, y_m) = 0$$

The Lagrangian function is

$$L = f + \lambda_1\, g_1 + \cdots + \lambda_m\, g_m$$

Critical points. To find them, we apply the **necessary condition** by differentiating L with respect to the $m+n$ variables. We also add the m constraints:

$$\boxed{\frac{\partial L}{\partial x_1} = 0, \ldots, \frac{\partial L}{\partial x_n} = 0}\,; \quad \boxed{\frac{\partial L}{\partial y_1} = 0, \ldots, \frac{\partial L}{\partial y_m} = 0}\,; \quad \boxed{g_1 = 0, \ldots, g_m = 0}$$

We have obtained $2m+n$ equations. The $2m+n$ corresponding unknowns are the points $P(a_1, \ldots, a_n, b_1 \ldots b_m)$ and the multipliers $\lambda_1, \ldots, \lambda_m$.

Sufficient condition. At the critical points we can have a maximum, a minimum or a saddle point (neither maximum nor minimum) depending on the type of quadratic form of $d^2L|_{dg_i=0}$. The condition $dg_i = 0$ means that, when studying d^2L, it is necessary to take into account the relations between the differentials of the variables that result from the conditions $g_i = 0$ in differential form. Then

$$\boxed{\text{If } d^2L|_{dg_i=0} \text{ is } \begin{cases} \text{positive definite} & \Longrightarrow \text{ minimum} \\ \text{negative definite} & \Longrightarrow \text{ maximum} \\ \text{indefinite} & \Longrightarrow \text{ saddle point} \end{cases}}$$

Remark. If d^2L is definite, it is not necessary to study the conditions $dg_i = 0$, since they will not change the type of the quadratic form. On the other hand, when d^2L is semidefinite or indefinite, the relations between the differentials of the variables can influence the final result.

Example. We apply the method to the introductory example, in which

$$f(x, y) = 1 - x^2 - y^2, \quad g(x, y) = x + y - 1 = 0$$

The Lagrangian function is $L = 1 - x^2 - y^2 + \lambda(x + y - 1)$. Differentiating

$$\frac{\partial L}{\partial x} = -2x + \lambda = 0; \quad \frac{\partial L}{\partial y} = -2y + \lambda = 0 \Longrightarrow x = y$$

Equating x and y in condition $x + y - 1 = 0$ we obtain the point $P'(1/2, 1/2)$.
The second derivatives are

$$\frac{\partial^2 L}{\partial x^2} = \frac{\partial^2 L}{\partial y^2} = -2; \quad \frac{\partial^2 L}{\partial x \partial y} = \frac{\partial^2 L}{\partial y \partial x} = 0 \Longrightarrow H = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

H is a diagonal matrix, so it is immediate to see that it is negative definite. Applying the Sylvester criterion, the same result is obtained, since the principal minors are $-2, +4$. As a consequence, we have a maximum at $P'(1/2, 1/2)$

Exercise. Find the dimensions of the rectangle (base x and height y) of maximum area, with a given perimeter P.
Solution. It is the square of side $P/4$.

12 Self-assesment exercises

12.1 True/False exercises

Exercise 1 Decide whether the following statements are true or false.

1. The properties of the "euclidean norm" are: positivity, conmutativity and triangular inequality.
2. A vector function of a real variable is continuous at $x = a$, if and only if each one of its components is continuous at $x = a$.
3. Consider a function $f : \mathbb{R}^2 \to \mathbb{R}$. If the value of the limit of f at $\vec{a}$, along any direction given by $\vec{\omega}$, is φ, then the limite of f at $\vec{a}$ is φ.
4. If the directional derivatives of a real function of two variables exist, then the function is continuous.
5. The directional derivatives of a real function of several variables depends always on the direction.
6. A necessary condition of differentiability of a real function of several variables is that the partial derivatives are continuous.
7. The jacobian matrix of a vector function of 3 components and 2 variables is a matrix of $3^2 = 9$ elements.
8. We consider the functions $\vec{f} : \mathbb{R}^n \to \mathbb{R}^m$ and $\vec{g} : \mathbb{R}^m \to \mathbb{R}^p$ whose domains are $D_{\vec{f}} = \mathbb{R}^n$ and $D_{\vec{g}} = \mathbb{R}^m$. If $\vec{f}$ and $\vec{g}$ are differentiables on their respective domains, then $\vec{f} \circ \vec{g}$ is differentiable on $\mathbb{R}^n$.
9. Let $\vec{f}$ be a vector function of a vector variable. If $\vec{f}$ satisfies the Schwarz theorem of the cross derivatives, then the jacobian matrix of the function is symmetric.
10. A function $u = f(x, y, z),\ f \in C^2$ verifies $d^2u = \left(\dfrac{\partial f}{\partial x}\,dx + \dfrac{\partial f}{\partial y}\,dy + \dfrac{\partial f}{\partial z}\,dz\right)^{(2}$. If we expand this expression, the numerical coefficiente of the term that contains $dxdz$ is 2.

Exercise 2 Decide whether the following statements are true or false.

a) Consider a function $f(x, y),\ f \in C^2$ on its entire domain. We can state that:

1. Its Taylor polynomial of degree 2 at the point (a, b) is the equation of the tangent plane to the surface $z = f(x, y)$ at the point $(a, b, f(a, b))$.
2. If f is the function $f(x, y) = x^2y^2$, its derivatives of order five and higher are zero, so its fourth degree Taylor polynomial around the point $P(1, 1)$ corresponds exactly to the function.

b) We now study the local extrema of the function $f(x, y),\ f \in C^2$ at points where their partial derivatives are zero. We can state that:

3. The function will only have extrema if d^2f is definite (positive or negative)
4. If d^2f is not definite and $|H| \neq 0$, then it is a saddle point.
5. If the Hessian determinant is null, the second differential can only be semidefinite, despite which an extremum may exist.

6. If d^2f is indefinite at the critical points, we can ensure that the function has no extrema.

7. If the matrix H is semidefinite it is an inconclusive case. But, if it is positive semidefinite, there is no maximum and, if it is negative semidefinite, there is no minimum.

c) Consider the equation $F(x,y) = x^2 + y^2 + xy + 3y = 0$ and the points $P(0,0)$ and $Q(-1,-1)$.

8. The conditions of existence and differentiability of the implicit f. are satisfied at P.

9. The conditions of existence and differentiability of the implicit f. are satisfied at Q.

10. At P, the derivative of the implicit function is null.

d) We look for the extrema of $f(x,y,z)$, with the constraints $g_1(x,y,z) = 0$, $g_2(x,y,z) = 0$.

11. The extrema coincide with those of the function $L = f + \lambda g_1 + \mu g_2,\ \forall \lambda, \mu \in \mathbb{R}$.

12. The equations $g_1 = 0,\ g_2 = 0$ only influence the obtaining of the critical points, not the study of the type of extremum (sufficient condition).

13. Once the critical points are obtained, we will have a maximum if $d^2L|_{dg_i=0,\ i=1,2}$ takes negative values $\forall (dx,dy,dz) \neq (0,0,0)$.

14. In the finding of constrained extrema, the type of extremum is given by the sign of $d^2L|_{dg_i=0}$. But, if the matrix H is defined, it is not necessary to use the conditions $dg_i = 0$.

12.2 Questions

Question 1. Let f be a real function defined in $D \subset \mathbb{R}^n$. Define the directional limit and the directional derivative of f at a point $\vec{a}$ and reason the truth or falsity of these sentences:

a) If the directional limit at $\vec{a}$ depends on the direction, there is no functional limit at $\vec{a}$

b) If f is differentiable at $\vec{a}$, the directional derivative at $\vec{a}$ does not depend on the direction.

Question 2. Let f be a real function defined in $D \subset \mathbb{R}^n$. State the necessary condition and the sufficient condition of differentiability and reason the truth or falsity of the sentences:

a) If f satisfies the sufficient condition on D, then it satisfies the necessary one.

b) If f satisfies the necessary condition on D, then it satisfies the sufficient one.

Question 3. Consider the function

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0). \end{cases}$$

To obtain its directional derivatives at the origin, we calculate the limit:

$$D_{\vec{\omega}} f(0,0) = \lim_{\lambda \to 0} \frac{f(\lambda\omega_x, \lambda\omega_y) - f(0,0)}{\lambda} = \cdots = \lim_{\lambda \to 0} \frac{\omega_x \omega_y}{\lambda}$$

which has no finite value. However, calculating the partial derivatives at the origin by the definition, we obtain

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h \to 0} \frac{f(h,0) - f((0,0)}{h} = \lim_{h \to 0} \frac{0-0}{h} = 0, \quad \frac{\partial f}{\partial y}(0,0) = \cdots = 0$$

Since the partial derivatives (which exist) are a particular case of the directional derivatives (which seem not to exist), explain the reason for this apparent contradiction.

Question 4. Reason the truth or falsity of the following statement: "Let f be a real function of 2 variables, differentiable on $\mathbb{R}^2$. Its gradient at the point $P(a, b)$ is a vector of $\mathbb{R}^2$ that indicates the direction of minimum variation of f, from its value at P".

Question 5. State the theorem of existence and differentiability of the implicit function.

Question 6. Let $F(x, y) = x \sin y - y \sin x = 0$. It can be verified that the equation $F(x, y) = 0$ satisfies the conditions to define an implicit function on a neighborhood of the point $P(\pi/2, \pi/2)$ (which is none other than $y = x$). On the other hand, we see that the function $y = 0$, a straight line that intersects $y = x$ at (0,0), is also a solution to the equation.

This seems to mean that, on every neighborhood of $(0, 0)$, there are two implicit functions ($y = x$ and $y = 0$). But the theorem states that, under certain conditions, the i.f. defined by $F(x, y) = 0$ is unique on a certain neighborhood of $P(a, b)$. Can you explain this?

12.3 Solution to the True/False exercises

Exercise 1.

1. **F**. The three properties are: positivity, product by a scalar and triangular inequality.

2. **T**. See **3.2.**

3. **F**. The fact that all the directional limits exist and coincide does not mean that there is a limit when approaching the point along any curve. See an example of this case in the supplementary documents of the unit).

4. **F**. The existence of directional derivatives does not ensure differentiability, therefore it does not ensure continuity. If f admits derivative in one direction, the limit of f at $\vec{a}$ will coincide with $f(\vec{a})$ only in that direction.

5. **F**. Not always. For example, if a differentiable function has an extremum at a point inside the domain, all the directional derivatives will be null at that point .

6. **F**. It is a sufficient (not necessary) condition of differentiability.

7. **F**. The Jacobian matrix of a vector function of 3 components and 2 variables is a matrix of $3 \times 2 = 6$ elements.

8. **F**. We have defined the composite function by saying that $(\vec{f} \circ \vec{g})(\vec{x}) = \vec{f}(\vec{g}(\vec{x}))$. With this definition, $\vec{f} \circ \vec{g}$ may not be differentiable, since it may not even exist. Indeed, if $p \neq n$, the elements $\vec{g}(\vec{x})$ will not be variables of the function $\vec{f}$. On the other hand, $\vec{g} \circ \vec{f}$ exists and is differentiable (see **6.2**).

9. **F**. The Jacobian matrix of a vector function of m components and n variables, is the matrix $m \times n$ of the partial derivatives and does not even have to be square. On the contrary, the Hessian matrix of the second derivatives is symmetric, if the conditions of the Schwarz's theorem are satisfied.

10. **T**. The expansion of the **symbolic** square of the sum contains the sum of "squares" plus the sum of "double products". One of the latter is the one requested, $2\dfrac{\partial^2 f}{\partial x \partial z}\, dxdz$.

Exercise 2.

1. **F**. The equation of the tangent plane is given by the Taylor polynomial of degree 1 (see **8.1**).

2. **T**. Of the 16 fourth-order derivatives, only the 6 in which we derive two times with respect to x and two times with respect to y are different from zero (all are equal to 4). Therefore the derivatives of order greater than 4 are null. The result is a polynomial of degree 4, in powers of $(x-1)$ and $(y-1)$. If we operate and simplify we will obtain x^2y^2.

3. **F**. There may be extrema when d^2f is semidefinite. Example: $f(x,y) = x^2 + y^4$ at the point $P(0,0)$ (see section **9.2** and the supplementary document "Example of a semidefinite cuadratic form").

4. **T**. If $|H| \neq 0$, then no element of the diagonal matrix is null. Since the quadratic form is not defined, the elements cannot be both positive nor can they be both negative (Sylvester). Therefore they must have different signs, so the quadratic form is undefined and it will be a saddle point.

5. **T**. If $|\mathrm{H}| = 0$, at least one diagonal element is null. Since there are two, they cannot have different signs, so it is semidefinite, but there can be an extremum (see question 3).

6. **F**. It can have extrema at some point on the boundary, even though its partial derivatives are not zero at that point.

7. **T**. If the Hessian matrix is semidefinite, it is a inconclusive case with no unique solution, but we can affirm what follows:

   - If it is positive semidefinite, for some $d\vec{x}$ it will be $d^2f > 0 \Longrightarrow \Delta f > 0$ (the function increases as we move away from the point, at least in one direction), then it cannot be a maximum.

   - If it is negative semidefinite, for some $d\vec{x}$ it will be $d^2f < 0 \Longrightarrow \Delta f < 0$ (the function decreases as we move away from the point, at least in one direction), then it cannot be a minimum.

8. **T**. The function and its two partial derivatives are continuous at every point. At P, F is null ad its derivative with respect to y is not, then the conditions are satisfied.

9. **F**. The function and its two partial derivatives are continuous at every point. At Q F is null and so is its derivative with respect to y, so the conditions are not satisfied.

10. **T**. At P the $\dfrac{\partial F}{\partial x} = 0$, so $\dfrac{\partial \psi}{\partial x} = -\dfrac{\partial F/\partial x}{\partial F/\partial y} = 0$.

11. **F**. $\boxed{\exists \lambda, \mu \in \mathbb{R}}$ for which the statement holds, but it does not hold $\boxed{\forall \lambda, \mu \in \mathbb{R}}$.

12. **F**. If the second derivatives of the functions g_i are not zero, their differentials will be part of the second differential of the Lagrangian function, so the constraint conditions will influence the type of extremum.

13. **T**. See **11.2**.

14. **T**. See **11.2**.

12.4 Solution to the questions

Question 1. We define directional limit and directional derivative:

The **directional limit** is the value that f approaches as $\vec{x}$ approaches $\vec{a}$ through the subset $E = \{\vec{x} \in \mathbb{R}^n \ / \ \vec{x} = \vec{a} + \lambda\vec{\omega}\}$. That is, it is obtained as $\lim\limits_{\lambda\to 0} f(\vec{a} + \lambda\vec{\omega})$.

The **directional derivative** is given by the expression:

$$D_{\vec{\omega}} f(\vec{a}) = \lim_{\lambda\to 0} \frac{f(\vec{a} + \lambda\vec{\omega}) - f(\vec{a})}{\lambda}$$

From the above definitions,

a) The limit of f at $\vec{a}$ is the value to which f tends as $\vec{x} \to \vec{a}$ **through any path**. If it exists, any directional limit must coincide with it. If the directional limit depends on the direction, it will not have a unique value, then there is no functional limit. Therefore **the statement is true**.

b) If f is differentiable, then it has directional derivatives and $D_{\vec{\omega}} f = \vec{g} \cdot \vec{\omega}$ holds. So the directional derivative depends on the direction (given by $\vec{\omega}$) and **the statement is false**.

Question 2. We define the necessary and sufficient conditions:

Necessary condition: If f is differentiable on D, then there exist directional derivatives on D and take the value: $D_{\vec{\omega}} f = \vec{g} \cdot \vec{\omega}$.

Sufficient condition: If f has continuous partial derivatives on D, it is differentiable on D.

From this:

a) If f satisfies the sufficient condition, then it is differentiable, so it satisfies the necessary condition. **The statement is true**.

b) If f satisfies the necessary condition, then it has directional derivatives and therefore partial derivatives (which are a particular case). But we cannot ensure that these are continuous, so the sufficient condition is not satisfied. **The statement is false**.

Question 3. The statement says that the limit has no finite value, which is only true if the numerator is not null. If the numerator is zero, so is the limit. Indeed, if we calculate the limit taking for $\vec{\omega}$ the directions of the unit vectors $\vec{\omega}_1 = (1, 0)$ and $\vec{\omega}_2 = (0, 1)$, which correspond to the directions of the axes, we get the values

$$D_{(1,0)}\vec{f}(0,0) = \cdots = \lim_{\lambda\to 0} \frac{1\cdot 0}{\lambda} = 0, \quad D_{(0,1)}\vec{f}(0,0) = \cdots = \lim_{\lambda\to 0} \frac{0\cdot 1}{\lambda} = 0$$

which coincide with those of the partial derivatives, so there is no contradiction.

Question 4. If f has two variables x and y, the components of its gradient vector are the partial derivatives of f with respect to x and y, so it will be a vector of $\mathbb{R}^2$.

This vector does not indicate the direction of minimum variation of the function f, but that of the maximum variation. The variation will be minimum (null) in the direction perpendicular to the gradient vector at each point, which is the condition to obtain the contour lines. This can be justified from the directional derivative formula

$$D_{\vec{\omega}} f = \vec{g}^t \cdot \omega = \|\vec{g}\| \|\omega\| \cos\varphi = \|\vec{g}\| \cos\varphi,$$

where we see that the variation will be maximum in the direction given by the gradient vector: if $\varphi = 0$, the $\cos\varphi$ is equal to 1. It will be minimum (null) in the perpendicular direction: if $\varphi = \pi/2$, the $\cos\varphi$ is equal to 0.

Therefore, although its first part is true, **the statement is false**.

Question 5. The theorem states the following:

Let F be a continuous function of two variables. Consider a equation $F(x,y)=0$. If the partial derivatives $\partial F/\partial x$ and $\partial F/\partial y$ are continuous and there exists a point $P(a,b)$ at which

$$F(a,b)=0, \quad \frac{\partial F}{\partial y}(a,b)\neq 0$$

then $F(x,y)=0$ defines a unique continuous function $y=\psi(x)$ as an implicit function of x on a neighborhood of $x=a$, $y=b$, which is differentiable.

Question 6. Since the point $P(\pi/2,\pi/2)$ satisfies the conditions of the theorem, there exists a neighborhood of it* in which there is a unique implicit function (the function $y=x$). On the contrary, $Q(0,0)$ does not satisfy the "non-null derivative with respect to y" condition, so we cannot guarantee that in some neighborhood of it there is a unique function. In fact, on every neighborhood there are four solutions to the equation $F(x,y)=0$: the straight lines $y=\pm x$, $y=0$ and $x=0$, which intersect at the origin.

Take, for example, any point on the line $y=0$, different from the origin, for example $R(0,1)$). We see that the conditions of the theorem are satisfied at R and, therefore, on a certain neighborhood of it** there exists a unique implicit function (which is now the function $y=0$).

(*) There are an infinite number of them, e.g. the circles with center at P and radius $r<1$.

(**) There are an infinite number of them, e.g. the circles with center at R and radius $r<\frac{1}{2}$.

Unit III. Series of real numbers (02.04.2024)

1 Definition

Given a sequence of real numbers $\{a_n\}$, we want to obtain the sum of its terms. Since we are dealing with an infinite number of addends, we cannot perform the sum directly.

The way to do it is to define a new sequence whose elements are the sums of the terms of $\{a_n\}$ in increasing number; and calculate the limit, when that number tends to ∞.

That is, we define

$$\begin{aligned} S_1 &= a_1 \\ S_2 &= a_1 + a_2 \\ S_3 &= a_1 + a_2 + a_3 \\ &\vdots \\ S_n &= a_1 + a_2 + a_3 + \cdots + a_n \end{aligned}$$

$\{S_n\}$ is the **sequence of partial sums** associated to $\{a_n\}$, which we will call **series** or (to emphasize that there are an infinite number of terms) **infinite series**. The addends $a_1, a_2, \ldots$ are the **terms of the series** and a_n is the **general term**.

We will obtain the sum of the series by calculating the limit of the sequence of partial sums

$$\lim_{n\to\infty} S_n = \sum_{n=1}^{\infty} a_n$$

We will preferably represent a series as

$$\sum a_n$$

reserving for its sum the expression

$$\sum_1^{\infty} a_n$$

However, this second form has the advantage of indicating the first element considered and is also frequently used to refer to series.

Types of series. We will classify the series according to the convergence of the sequence of partial sums. We say that a series is:

a) Convergent, if and only if $\lim\limits_{n\to\infty} S_n = S \in \mathbb{R}$. The limit S is the **sum of the series**.

b) Divergent, if and only if $\lim\limits_{n\to\infty} |S_n| = \infty$.

c) Oscillating, if it neither converges nor diverges.

Examples.

- The harmonic series (the sum of the inverse of natural numbers): $\sum \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots$
- The series of the powers of (-1): $\sum (-1)^n = -1 + 1 - 1 + \ldots$
- The series of the powers of $1/2$: $\sum \frac{1}{2^n} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \ldots$
- The series of the inverse of the roots of the odd numbers: $\sum \frac{1}{\sqrt{2n-1}} = 1 + \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{5}} + \ldots$

2 Geometric series

Once series have been defined, we are going to study a particular case, that of the geometric series, in which each term is obtained by multiplying the previous one by a fixed non-zero number called **ratio**. The successive terms are:

$$\begin{aligned} a_2 &= a_1 r \\ a_3 &= a_2 r = a_1 r^2 \\ a_4 &= a_3 r = a_1 r^3 \\ &\vdots \\ a_n &= a_{n-1} r = a_1 r^{n-1} \end{aligned}$$

Convergence. To study it, we will first calculate the sum of n terms from the value of the first of them and the ratio; then we will make n tend to infinity. The sum S_n is

$$S_n = a_1 + a_2 + a_3 + \cdots + a_n$$

Multiplying it by r

$$S_n\, r = a_1 r + a_2 r + a_3 r + \cdots + a_n r$$

We now subtract both expressions, remembering that each term is equal to the previous one multiplied by r, so all the terms disappear, except two. We obtain

$$S_n\,(1-r) = a_1 - a_n r = a_1 - a_1 r^{n-1} r = a_1 - a_1 r^n$$

from where we solve for S_n

$$\boxed{S_n = a_1 \frac{1-r^n}{1-r} \qquad (r \neq 1)}$$

The following cases result

$$\begin{array}{rcll} r > 1 & \Longrightarrow & S_n = a_1 \dfrac{r^n - 1}{r-1} & \Longrightarrow S_n \to \infty \cdot \operatorname{sign}(a_1) \quad \text{(D)} \\ r = 1 & \Longrightarrow & S_n = n\, a_1 & \Longrightarrow S_n \to \infty \cdot \operatorname{sign}(a_1) \quad \text{(D)} \\ -1 < r < 1 & \Longrightarrow & r^n \to 0 & \Longrightarrow S_n \to \dfrac{a_1}{1-r} \quad \text{(C)} \\ r = -1 & \Longrightarrow & S_n = a_1 - a_1 + a_1 - \ldots & \Longrightarrow S_n = \begin{cases} a_1, & \text{odd}\, n \\ 0, & \text{even}\, n \end{cases} \quad \text{(O)} \\ r < -1 & \Longrightarrow & |S_n| = \left| \dfrac{a_1}{r-1} \right| |r^n - 1| & \Longrightarrow |S_n| \to \infty \quad \text{(D)} \end{array}$$

The geometric series only converges for $|r| < 1$. Otherwise, it diverges or oscillates.

3 Necessary condition of convergence

Theorem. *If a series converges, its general term tends to 0* (non-sufficient condition).

$$\boxed{\sum a_n \text{ converges} \Longrightarrow \lim_{n\to\infty} a_n = 0}$$

Proof. If $\sum a_n$ converges, the sequence of partial sums has a limit $(n \to \infty)$. This limit will be the same, whether we take n terms or $n-1$. Then

$$\lim_{n\to\infty} S_n = \lim_{n\to\infty} S_{n-1} = S \Longrightarrow \lim_{n\to\infty} a_n = \lim_{n\to\infty} (S_n - S_{n-1}) = S - S = 0$$

4 Properties of the series

1) *If we insert in the sequence $\{a_n\}_{n\in\mathbb{N}}$ a finite number of terms of sum b, the behavior of the series $\sum_n a_n$ does not change (it continues converging, diverging or oscillating). If it converges, its sum is increased by b.*

P: Let $b_1+b_2+\cdots+b_q=b$. Taking a partial sum sufficiently advanced, so that it contains all of them, it results

$$S'_{n+q}=S_n+b \Longrightarrow \lim_{n\to\infty} S'_{n+q}=\lim_{n\to\infty}(S_n+b)$$

Then, by the limit properties,

- If S_n converges to S, S'_n converges to $S+b$.
- If S_n diverges, S'_n diverges.
- If S_n oscillates, S'_n oscillates.

Remark: *If we remove from $\{a_n\}_{n\in\mathbb{N}}$ a finite number of terms of sum b, the behavior of the series $\sum_n a_n$ does not change. If it converges, its sum is decreased by b.* It can be proved analogously, adding the opposites of the terms that we want to remove.

2) *If we multiply all the terms of a series by a real number $\lambda \neq 0$, its behavior does not change. If it converges, its sum is multiplied by λ.*

P: The new sequence of partial sums is

$$\begin{aligned}
S'_1 &= \lambda a_1 && = \lambda S_1\\
S'_2 &= \lambda a_1+\lambda a_2 && = \lambda S_2\\
&\vdots\\
S'_n &= \lambda a_1+\cdots+\lambda a_n = \lambda S_n \Longrightarrow \lim_{n\to\infty} S'_n = \lim_{n\to\infty}(\lambda S_n)
\end{aligned}$$

Hence, by the properties of the limits, both series have the same behavior.

3) *If a series is* ***convergent or divergent****, we can replace several terms by their sum without changing the behavior (or the sum, if it converges).*

P: When replacing some terms by their sum, the new sequence of partial sums $\{S'_n\}$ will have fewer terms than the old one $\{S_n\}$, but all the terms of the new one will belong to the old one. That is, $S'_1, S'_2, \dots S'_n$ is a subsequence of $S_1, S_2, \dots S_n$. Hence:

a) If S_n converges, S'_n has the same limit (P. **7** of the limits of sequences).

b) If S_n diverges, so will S'_n. Indeed, since S'_n is a subsequence of S_n, the terms of S'_n also belong to S_n, therefore they satisfy the divergence condition.

4) If we remove the first n terms of a series, the resulting series is called **remainder of order** n: $R_n=a_{n+1}+a_{n+2}+\dots$ *The remainder of order n of a convergent series is convergent and its sum tends to 0 as $n\to\infty$.*

P: Given a convergent series, its partial sum of order p is $S_p=\sum_{i=1}^{p} a_i$ and the sum of the series $\big($which we denote by $\sum_{i=1}^{\infty} a_i\big)$ will be $S=\lim\limits_{p\to\infty} S_p$.

For a given n, the remainder series R_n is obtained by eliminating the first n terms, whose sum is $S_n=\sum_{i=1}^{n} a_i$. The partial sums of R_n will be those of the initial series, decreased by the value S_n, that is, $R_n^p=\sum_{i=n+1}^{p} a_i = S_p - S_n$.

So, making $p \to \infty$, the sum R_n^∞ of the series R_n will be

$$R_n^\infty = \sum_{i=n+1}^{\infty} a_i = \lim_{p\to\infty} (S_p - S_n) = S - S_n \in \mathbb{R}$$

so the remainder R_n is a convergent series.

If we now make $n \to \infty$, we get $\lim\limits_{n\to\infty} R_n^\infty = \lim\limits_{n\to\infty} (S - S_n) = S - S = 0.$

5) Given $\sum a_n$ and $\sum b_n$, their **linear combination** is the series whose general term is the linear combination of the general terms, $\sum(\alpha\, a_n + \beta\, b_n)$. It holds that *the linear combination of convergent series is convergent and its sum is the linear combination of the sums.*

P: If both series converge, their partial sums $S_n^a = \sum_{i=1}^n a_i$ and $S_n^b = \sum_{i=1}^n b_i$ will satisfy: $\{S_n^a\}_{n\in\mathbb{N}} \to S^a$, $\{S_n^b\}_{n\in\mathbb{N}} \to S^b$. The partial sum of their l.c. will be $\sum_{i=1}^n(\alpha a_i + \beta b_i) = \alpha \sum_{i=1}^n a_i + \beta \sum_{i=1}^n b_i = \alpha S_n^a + \beta S_n^b$ and we will have, $\forall \alpha, \beta \in \mathbb{R}$,

$$\lim_{n\to\infty} \left(\alpha S_n^a + \beta S_n^b\right) = \alpha \lim_{n\to\infty} S_n^a + \beta \lim_{n\to\infty} S_n^b = \alpha S^a + \beta S^b.$$

The next two properties refer only to **series of positive terms** (S.P.T.). As justified later (section **6**), these series are never oscillating.

6) *If we alter the order of the terms of a S.P.T., the behavior does not change (nor the sum, if it converges).*

P: Let $\sum a_n$ and $\sum a'_n$ be the same series with the terms in a different order.

Since both have the same terms (and they are non-negative), for all partial sum of $\sum a_n$ we can find a partial sum of $\sum a'_n$ that exceeds it: it is enough to take an index m such that S'_m contains all the terms of S_n, so that $n \leq m$.

We can now do the same with S'_m, finding an index p such that S_p contains all the terms of S'_m, so $m \leq p$. Then,

$$\forall n\ \exists\ m, p \big/ S_n \leq S'_m \leq S_p \quad (n \leq m \leq p)$$

Since $\lim\limits_{n\to\infty} S_n = \lim\limits_{p\to\infty} S_p$, it results

- If S_n converges, so does S'_m (P.**6** of the limits of sequences).
- If S_n diverges to ∞, since $S_n \leq S'_m$, so does S'_m.

Hence

$$\lim_{n\to\infty} S_n = \lim_{m\to\infty} S'_m \text{ (finite or infinite).}$$

7) *If we group a finite or infinite number of terms of a S.P.T. or we decompose them into the sum of* ***positive*** *terms, the behavior of the series is not altered (nor is the sum, if it converges).*

P: Since the S.P.T. are never oscillating, we can apply the P.**3** of the series. Then:

a) If we group terms of the series $\sum a_n$, neither the behavior nor the sum change.

b) If we decompose terms of $\sum a_n$ into sum of positive terms we will obtain a new series $\sum a'_n$. If, starting now from $\sum a'_n$, we regroup the terms that we decomposed before, we will obtain the initial series $\sum a_n$ again. But this will have the same behavior and sum as $\sum a'_n$ (property **3**).

Thus, whether we group terms or decompose them into a sum of positive terms, the behavior does not change (or the sum, if it converges).

5 Cauchy's criterion of convergence

Given a sequence $\{a_n\}$, we have defined its associated series $\sum a_n$ as a sequence of partial sums $\{S_n\}$, where

$$S_n = \sum_{i=1}^{n} a_i$$

so the series converges if and only if $\{S_n\}$ does. As we saw when studying real numbers, a sequence of real numbers converges if and only if it is a Cauchy sequence. Hence, $\sum a_n$ will converge if and only if $\{S_n\}$ is a Cauchy sequence. This can be expressed in two ways:

a) $\boxed{\sum a_n \text{ converges } \iff \forall \varepsilon > 0\ \exists n_0 \in \mathbb{N} \,/\, |S_p - S_q| < \varepsilon,\ \forall p, q \geq n_0}$

The series converges if the difference in absolute value of two sufficiently advanced partial sums is as small as we want.

Let $p > q$. Then $S_p - S_q$ can be written as: $\displaystyle\sum_{i=1}^{p} a_i - \sum_{i=1}^{q} a_i = \sum_{i=q+1}^{p} a_i$.

Doing in the previous expression $q = m$ and $p - q = t$, that is, $p = m + t$, condition **a)** becomes:

b) $\boxed{\sum a_n \text{ converges } \iff \forall \varepsilon > 0\ \exists n_0 \in \mathbb{N} \,/\, \left|\displaystyle\sum_{i=m+1}^{m+t} a_i\right| < \varepsilon,\ \forall m \geq n_0,\ \forall t \in \mathbb{N}}$

That is, the series $\sum a_n$ converges if and only if the sum of any number of terms, starting from a given one m, can be made as small as we want by taking m sufficiently high. This condition will be used in the first of the convergence criteria that we will see below (6.1. Criterion of the majorant and the minorant).

Saying that this expression can be made as small as we want for an m sufficiently high is equivalent to say that it tends to zero, if m tends to ∞. That is, the condition **b)** can also be written as:

$$\sum a_n \text{ converges } \iff \lim_{m\to\infty} \sum_{i=m+1}^{m+t} a_i = 0,\ \forall t \in \mathbb{N}$$

6 Series of positive terms. Convergence criteria

Definition. If the terms a_n are greater than or equal to zero, $a_n \geq 0$, we say that $\sum a_n$ is a series of positive terms (S.P.T.). In this case

$$S_n = \sum_{i=1}^{n} a_i,\ a_i \geq 0 \Longrightarrow S_1 \leq S_2 \leq \cdots \leq S_n \leq \ldots$$

that is, the sequence of partial sums is monotone increasing. Then there are two options:

a) $\{S_n\}$ is bounded, therefore it converges (bounded monotone increasing sequence).

b) $\{S_n\}$ is not bounded, therefore it diverges (its terms can eventually exceed any value).

Therefore, **a S.P.T. can only be convergent or divergent**, never oscillating.

Next we will study different convergence criteria for this type of series.

6.1 Majorant and minorant

Let two S.P.T. be $\sum a_n$ and $\sum b_n$. We say that $\sum b_n$ is a majorant of $\sum a_n$ if and only if

$$a_n \leq b_n \quad \forall n \geq n_1$$

If $\sum b_n$ is a majorant of $\sum a_n$, then we say that $\sum a_n$ is a minorant of $\sum b_n$.

Examples.

1. $\sum \dfrac{1}{n+7}$ is a minorant of $\sum \dfrac{1}{n}$, since $\dfrac{1}{n+7} < \dfrac{1}{n}\ \forall n$.
2. $\sum \dfrac{1}{n+7}$ is a majorant of $\sum \dfrac{1}{n^2}$, since , $\dfrac{1}{n+7} > \dfrac{1}{n^2}\ \forall n > 3$.

It holds that:

a) If the majorant is convergent, the minorant is convergent too.

b) If the minorant is divergent, the majorant is divergent too.

Proof. Applying the majorant condition $\forall m \geq n_1$, we have:

$$\left\{\begin{array}{c} a_{m+1} \leq b_{m+1} \\ a_{m+2} \leq b_{m+2} \\ \vdots \\ a_{m+t} \leq b_{m+t} \end{array}\right\} \Longrightarrow \sum_{i=m+1}^{m+t} a_i \leq \sum_{i=m+1}^{m+t} b_i,\ \forall m \geq n_1, \forall t \in \mathbb{N} \tag{1}$$

a) By Cauchy's general convergence criterion (section **5**), it results:

$$\sum b_n \text{ convergent} \Longrightarrow \forall \varepsilon > 0,\ \exists n_0 \,/\, \sum_{i=m+1}^{m+t} b_i < \varepsilon,\ \forall m \geq n_0,\ \forall t \in \mathbb{N} \tag{2}$$

Taking conditions (1) and (2) into account,

$$\exists n = \max(n_0, n_1) \,/\, \sum_{i=m+1}^{m+t} a_i \leq \sum_{i=m+1}^{m+t} b_i < \varepsilon,\ \forall m \geq n, \forall t \in \mathbb{N}$$

hence $\sum a_n$ is convergent.

b) If $\sum a_n$ is divergent, then $\sum b_n$ is divergent, by *reductio ad absurdum.*
Indeed, if $\sum b_n$ does not diverge, being a S.P.T., it has to converge. Hence, its minorant $\sum a_n$ also converges, against the hypothesis.

6.2 Riemann series

$\sum a_n$ is a Riemann series if its general term is $\boxed{a_n = \dfrac{1}{n^\alpha},\ \alpha > 0}$

From the previous criterion and the convergence of the geometric series, it can be shown that:

$$\boxed{\text{If } \alpha > 1 \text{, then the series converges. If } \alpha \leq 1 \text{, it diverges}}$$

Examples. $\sum \dfrac{1}{\sqrt{n}}$ and $\sum \dfrac{1}{n}$ (harmonic) diverge while $\sum \dfrac{1}{n^2}$ and $\sum \dfrac{1}{n^{3/2}}$ converge.

Proof. We write the expression of the partial sum S_n of the Riemann series:

$$S_n = \sum_{i=1}^{n} \frac{1}{i^\alpha} = \frac{1}{1^\alpha} + \frac{1}{2^\alpha} + \frac{1}{3^\alpha} + \cdots + \frac{1}{n^\alpha}$$

Since this is a series of positive terms, we can group them. We take groups of an increasing number of terms: 1, 2, 4, 8, 16. . . .

$$S_n = \frac{1}{1^\alpha} + \left(\frac{1}{2^\alpha} + \frac{1}{3^\alpha}\right) + \left(\frac{1}{4^\alpha} + \frac{1}{5^\alpha} + \frac{1}{6^\alpha} + \frac{1}{7^\alpha}\right) + \left(\frac{1}{8^\alpha} + \cdots + \frac{1}{15^\alpha}\right) + \ldots$$

We now write the expressions of a minorant S_n^- and a majorant S_n^+, of S_n

$$\begin{aligned} S_n^- &= \frac{1}{2^\alpha} + \left(\frac{1}{4^\alpha} + \frac{1}{4^\alpha}\right) + \left(\frac{1}{8^\alpha} + \frac{1}{8^\alpha} + \frac{1}{8^\alpha} + \frac{1}{8^\alpha}\right) + \left(\frac{1}{16^\alpha} + \cdots + \frac{1}{16^\alpha}\right) + \ldots \\ S_n^+ &= \frac{1}{1^\alpha} + \left(\frac{1}{2^\alpha} + \frac{1}{2^\alpha}\right) + \left(\frac{1}{4^\alpha} + \frac{1}{4^\alpha} + \frac{1}{4^\alpha} + \frac{1}{4^\alpha}\right) + \left(\frac{1}{8^\alpha} + \cdots + \frac{1}{8^\alpha}\right) + \ldots \end{aligned}$$

Between them, the following relation exists, $\forall n > 2$

$$S_n^- < S_n < S_n^+$$

We now study S_n^+, grouping the terms of equal value

$$S_n^+ = 1 + \frac{2}{2^\alpha} + \frac{4}{4^\alpha} + \frac{8}{8^\alpha} + \cdots = 1 + 2^{1-\alpha} + 4^{1-\alpha} + 8^{1-\alpha} + \ldots$$

and we see that S_n^+ is a geometric series of ratio $r = 2^{1-\alpha}$. If $\alpha > 1$, then

$$1 - \alpha < 0 \Longrightarrow 0 < 2^{1-\alpha} < 1$$

so the ratio is less than 1 in absolute value, hence S_n^+ converges (see section **2**). Therefore its minorant S_n converges as well (section **6.1**).

Doing the same with S_n^-

$$S_n^- = \frac{1}{2^\alpha} + \frac{2}{4^\alpha} + \frac{4}{8^\alpha} + \cdots = \frac{1}{2}\left(\frac{2}{2^\alpha} + \frac{4}{4^\alpha} + \frac{8}{8^\alpha} + \ldots\right)$$

and we have again obtained a geometric series of ratio $r = 2^{1-\alpha}$. If $\alpha \leq 1$, then

$$1 - \alpha \geq 0 \Longrightarrow 2^{1-\alpha} \geq 1$$

so the ratio is greater than or equal to 1, hence S_n^- diverges (see section **2**). Therefore its majorant S_n diverges as well, as seen in **6.1**.

Therefore

$$\alpha > 1 \Longrightarrow \sum \frac{1}{n^\alpha} \text{ converges}, \quad \alpha \leq 1 \Longrightarrow \sum \frac{1}{n^\alpha} \text{ diverges}$$

Exercises. Taking into account **6.1** and **6.2**, study the convergence of the following series

$$\sum \frac{|\sin n|}{n^\pi} \ \text{(C)}\,; \qquad \sum \frac{2 + \cos n}{n^{\pi-3}} \ \text{(D)}$$

6.3 Comparison of series

Given $\sum a_n$ and $\sum b_n,\ a_n,\, b_n > 0$, if $\boxed{\lim_{n\to\infty} \dfrac{a_n}{b_n} = k \neq 0}$ then both converge or both diverge.

Proof. The existence of the limit k means that

$$\forall \varepsilon > 0,\ \exists n \,/\, \left| \frac{a_m}{b_m} - k \right| < \varepsilon, \quad \forall m \geq n$$

that is:

$$k - \varepsilon < \frac{a_m}{b_m} < k + \varepsilon \iff (k-\varepsilon)\, b_m < a_m < (k+\varepsilon)\, b_m, \quad \forall m \geq n$$

From the previous relation, if $\sum b_n$ diverges, then $\sum (k-\varepsilon)\, b_n$ diverges, since its general term is that of $\sum b_n$ multiplied by a constant. Hence its majorant $\sum a_n$ diverges.

Likewise, if $\sum b_n$ converges, then $\sum (k+\varepsilon)\, b_n$ converges and its minorant $\sum a_n$ converges too.

We can conclude that both series converge or both diverge.

Practical application. When studying S.P.T. we can **replace the general term by an equivalent infinitesimal**, without changing the convergence or divergence of the series.

Remark. When the limit of the quotient of the general terms is not a $k \neq 0$, there are two cases in which the previous practical rule is still valid. From the definition of limit and the majorant criterion, it can be proved that:

a) If $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = 0$ and $\sum b_n$ converges, then $\sum a_n$ converges.

b) If $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = \infty$ and $\sum b_n$ diverges, then $\sum a_n$ diverges.

Example 1.

a) The series $\sum \tan \dfrac{1}{n}$ behaves like $\sum \dfrac{1}{n}$ (harmonic), therefore it diverges.

b) The series $\sum \left(e^{1/n^2} - 1 \right)$ behaves like $\sum \dfrac{1}{n^2}$ (Riemann, $\alpha = 2$), therefore it converges.

c) Let $a_n = \dfrac{\ln n}{n^2}$. Let $b_n = \dfrac{1}{n^{3/2}}$ ($\alpha > 1$, convergent). $\lim\limits_{n\to\infty} \dfrac{a_n}{b_n} = 0 \Longrightarrow \sum a_n$ converges.

As we can see, by combining **6.2** and **6.3** we can study the convergence of some series:

1. We find an infinitesimal a'_n, equivalent to a_n.
2. If $\sum a'_n$ has the same behavior as a Riemann series, from the value of α we can know whether $\sum a_n$ converges.

Example 2. We study the series of general term a_n:

a) $a_n = \sin \dfrac{2n+1}{n^3} \sim \dfrac{2n+1}{n^3} \sim \dfrac{2}{n^2}$. Since $2 \sum \dfrac{1}{n^2}$, converges ($\alpha > 1$), so does $\sum a_n$.

b) $a_n = \dfrac{1}{\sqrt{4n + 4\ln n}} = \dfrac{1}{\sqrt{4n}\sqrt{1 + (\ln n)/n}} \sim \dfrac{1}{2}\dfrac{1}{n^{1/2}} \Rightarrow \sum a_n$ diverges, since $\alpha = 1/2 < 1$.

Exercises. Study the convergence of: $\sum \dfrac{1}{\sqrt{n^3+n}}$ (C) ; $\sum \dfrac{k^2}{(n+k)^2}, k \in \mathbb{R}$ (C).

The following criteria (or tests) (**6.4** to **6.7**) are stated in two ways. **The first** one expresses a condition which must be satisfied from a given term and **is used to prove the criterion**. In **the second** one a limit is calculated and the test **is easier to apply**.

If the second condition holds, the first one holds too, as shown in **6.4**.

6.4 Root test (Cauchy-Hadamard)

a) $$\text{If } \forall m \geq n_0 \begin{cases} \sqrt[m]{a_m} \leq k < 1 \Longrightarrow \sum a_n \text{ is convergent} & \text{(a.1)} \\ \\ \sqrt[m]{a_m} \geq 1 \Longrightarrow \sum a_n \text{ is divergent} & \text{(a.2)} \end{cases}$$

b) $$\text{If } \exists \lim_{n\to\infty} \sqrt[n]{a_n} = l \begin{cases} l < 1 \Longrightarrow \sum a_n \text{ is convergent} & \text{(b.1)} \\ l > 1 \Longrightarrow \sum a_n \text{ is divergent} & \text{(b.2)} \\ l = 1 \Longrightarrow \text{ inconclusive (if } \sqrt[n]{a_n} \to 1^+\text{, D)} & \text{(b.3)} \end{cases}$$

Proof. (from condition **a**)

a.1. If $\forall m \geq n_0$, $\sqrt[m]{a_m} \leq k < 1$, then $a_m \leq k^m$, being $k < 1$.

Hence $\sum a_n$ is a minorant of a geometric series ($k < 1 \Rightarrow$ convergent), so it is convergent.

a.2. If $\forall m \geq n_0$, $\sqrt[m]{a_m} \geq 1$, then $a_m \geq 1$.

Hence $\sum a_n$ is a majorant of the series whose general term is $a_n = 1$, therefore divergent.

Remark. If the second condition holds, the first one holds too. Indeed:

b.1. If $\lim_{n\to\infty} \sqrt[n]{a_n} = l < 1$, we take a value beween l and 1 $\big($e.g. $k = (l+1)/2\big)$, hence $l < k < 1$.

Due to the properties of the limits, if the limit is less than k, the terms of the series, starting from a given one, will also be less than k, that is

$$\exists n_0 \big/ \sqrt[m]{a_m} < k < 1,\ \forall m \geq n_0$$

Then condition **a.1** is satisfied and the series $\sum a_n$ converges.

b.2. If $\lim_{n\to\infty} \sqrt[n]{a_n} = l > 1$, the terms will be greater than 1, starting from a given one, that is

$$\exists n_0 \big/ \sqrt[m]{a_m} > 1,\ \forall m \geq n_0$$

Then condition **a.2** is satisfied and the series $\sum a_n$ diverges.

b.3. If $\lim_{n\to\infty} \sqrt[n]{a_n} = 1^+$ (the terms tend to 1 from the right), then they will be greater than or equal to 1, starting from a given one, that is

$$\exists n_0 \big/ \sqrt[m]{a_m} \geq 1,\ \forall m \geq n_0$$

Condition **a.2** is also satisfied and the series $\sum a_n$ diverges.

Application. This test is especially suitable when a_n contains *nth* powers. It is also valid with factorial expressions, using Stirling's formula.

Example. To study $\sum \dfrac{P(n)}{2^n}$, we calculate $l = \lim_{n\to\infty} \sqrt[n]{a_n} = \dfrac{1}{2}$. Since $l < 1$, it converges.

Exercises. Study: $\sum \dfrac{\sqrt{n}}{3^n}$ (C) ; $\sum \dfrac{e^n}{n^3}$ (D) ; $\sum \dfrac{\ln n}{n^n}$ (C) ; $\sum \dfrac{n^2}{n!}$ (C).

6.5 Ratio test (D'Alembert)

a) $$\text{If } \forall m \geq n_0 \begin{cases} \dfrac{a_{m+1}}{a_m} \leq k < 1 \Longrightarrow \sum a_n \text{ is convergent} & \text{(a.1)} \\[2ex] \dfrac{a_{m+1}}{a_m} \geq 1 \Longrightarrow \sum a_n \text{ is divergent} & \text{(a.2)} \end{cases}$$

b) $$\text{If } \exists \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = l \begin{cases} l < 1 \Longrightarrow \sum a_n \text{ is convergent} \\ l > 1 \Longrightarrow \sum a_n \text{ is divergent} \\ l = 1 \Longrightarrow \text{ inconclusive (if } \dfrac{a_{n+1}}{a_n} \to 1^+\text{, D)} \end{cases}$$

Proof. (from condition **a**)

a.1. If $\forall m \geq n_0$, $\dfrac{a_{m+1}}{a_m} \leq k < 1$, then $a_{m+1} \leq k\, a_m$. Giving values to m, we obtain

$$\begin{array}{lclcl} a_{n_0+1} & \leq & k\, a_{n_0} & = & k\, a_{n_0} \\ a_{n_0+2} & \leq & k\, a_{n_0+1} & \leq & k^2\, a_{n_0} \\ & \vdots & & & \end{array}$$

Adding on both sides it yields $\sum_{n_0}^{\infty} a_n \leq a_{n_0}(1 + k + k^2 + \dots)$, so $\sum a_n$ converges, since $a_{n_0}(1 + k + k^2 + \dots)$ is a geometric series of ratio $k < 1$ (convergent).

a.2. If $\forall m \geq n_0$, $\dfrac{a_{m+1}}{a_m} \geq 1$, then $a_{m+1} \geq a_m$, hence $\{a_n\}$ is monotone increasing from $n = n_0$. Since a_{n_0} is positive (because the quotient exists for $m = n_0$), we have that

$$a_m \geq a_{n_0} > 0, \ \forall m \geq n_0$$

so a_n cannot have a null limit. The necessary condition of convergence is, therefore, not satisfied and $\sum a_n$ does not converge. Since it is of positive terms, it diverges.

Remark. The root law (CI1, unit III, 5.4) states that, if the limit of the quotient exists, so does the limit of the root and they coincide, that is

$$\text{If } \exists \lim_{n\to\infty} \frac{a_{n+1}}{a_n} = l \Longrightarrow \lim_{n\to\infty} \sqrt[n]{a_n} = l$$

So, if the limit of the quotient exists, the limit of the root exists too; but it may happen that the limit of the quotient does not exist and the limit of the root does. Therefore, the root test solves more cases than the ratio test.

Application. This test is especially suitable when the general term involves factorials and also gives good results with *nth* powers.

Examples. We study the following series.

1. $\sum \dfrac{2^n}{n!}$. We calculate $l = \lim_{n\to\infty} \dfrac{a_{n+1}}{a_n} = \lim_{n\to\infty} \dfrac{2^{n+1}}{(n+1)!} : \dfrac{2^n}{n!} = \lim_{n\to\infty} \dfrac{2}{n+1} = 0 < 1 \Rightarrow$ C.

2. $\sum \dfrac{\sqrt{n}}{n!}$. $l = \lim_{n\to\infty} \dfrac{a_{n+1}}{a_n} = \lim_{n\to\infty} \dfrac{\sqrt{n+1}}{(n+1)!} : \dfrac{\sqrt{n}}{n!} = \lim_{n\to\infty} \sqrt{\dfrac{n+1}{n}} \dfrac{1}{n+1} = 0 < 1 \Rightarrow$ C.

Exercises. Study the convergence of: $\sum \dfrac{\ln n}{n!}$ (C) ; $\sum \dfrac{\sqrt{n}}{n^n}$ (C) ; $\sum \dfrac{\ln n}{e^n}$ (C) ; $\sum \dfrac{n^n}{n!}$ (D).

6.6 Raabe's test

a) If $\forall m \geq n_0 \begin{cases} m\left(1-\dfrac{a_{m+1}}{a_m}\right) \geq k > 1 \Longrightarrow \sum a_n \text{ is convergent} \\[2ex] m\left(1-\dfrac{a_{m+1}}{a_m}\right) \leq 1 \Longrightarrow \sum a_n \text{ is divergent} \end{cases}$

b) If $\exists \lim\limits_{n\to\infty} n\left(1-\dfrac{a_{n+1}}{a_n}\right) = l \begin{cases} l > 1 \Longrightarrow \sum a_n \text{ is convergent} \\ l < 1 \Longrightarrow \sum a_n \text{ is divergent} \\ l = 1 \Longrightarrow \text{ inconclusive (if } n\big(1-\dfrac{a_{n+1}}{a_n}\big) \to 1^-\text{, D)} \end{cases}$

(The proof may be seen in J. Burgos, pg. 456).

Remark. In this test, contrary to those of Cauchy and D'Alembert, convergence is obtained for values of l greater than 1 and divergence for values less than 1.

Application. This test is less used than the previous two because the expression of the limit is somewhat more complex. It is particularly useful when there is a inconclusive case when applying D'Alembert, since the quotient between a_{n+1} and a_n has already been calculated.

Example. We study the series of general term $a_n = \dfrac{a(a+1)\dots(a+n)}{b(b+1)\dots(b+n)}$.

The expression of a_n suggests using the ratio test. We calculate the limit of the quotient:

$$\frac{a_{n+1}}{a_n} = \cdots = \frac{a+n+1}{b+n+1} \Longrightarrow \lim_{n\to\infty}\frac{a_{n+1}}{a_n} = 1$$

so it is an inconclusive case. Then, we apply Raabe's test:

$$\lim_{n\to\infty} n\left(1-\frac{a_{n+1}}{a_n}\right) = \lim_{n\to\infty} n\left(1-\frac{a+n+1}{b+n+1}\right) = \lim_{n\to\infty} n\left(\frac{b+n+1-(a+n+1)}{b+n+1}\right) = b-a$$

Hence:

- If $b-a>1$ ($\Leftrightarrow b>a+1$), the series converges.
- If $b-a<1$ ($\Leftrightarrow b<a+1$), the series diverges.
- If $b-a=1$ ($\Leftrightarrow b=a+1$), we arrive (again) to an inconclusive case, which we solve replacing in the series:

$$\sum a_n = \sum \frac{a}{n+a+1} = a\sum\frac{1}{n+a+1}$$

which has the same behavior as $\sum \dfrac{1}{n}$, divergent ($a_n \sim a'_n$, see **6.3**).

Therefore $\sum a_n$ converges if $b>a+1$ and diverges otherwise.

Exercise. Study the convergence of the series $\sum \dfrac{n!}{a(a+1)\dots(a+n-1)}$.

Solution. For $a>2$, it converges; for $a\leq 2$, it diverges.

6.7 Logarithmic test

a) If $\forall m \geq n_0 \begin{cases} \dfrac{\ln(1/a_m)}{\ln m} \geq k > 1 \Longrightarrow \sum a_n \text{ is convergent} & \text{(a.1)} \\ \dfrac{\ln(1/a_m)}{\ln m} \leq 1 \Longrightarrow \sum a_n \text{ is divergent} & \text{(a.2)} \end{cases}$

b) If $\exists \lim\limits_{n\to\infty} \dfrac{\ln(1/a_n)}{\ln n} = l \begin{cases} l > 1 \Longrightarrow \sum a_n \text{ is convergent} \\ l < 1 \Longrightarrow \sum a_n \text{ is divergent} \\ l = 1 \Longrightarrow \text{ inconclusive (if } \dfrac{\ln(1/a_n)}{\ln n} \to 1^-, \text{ D)} \end{cases}$

Proof. (from condition **a**)

a.1. If $\forall m \geq n_0$, $\dfrac{\ln(1/a_m)}{\ln m} \geq k > 1$, then $\ln(1/a_m) \geq k \ln m = \ln m^k$.

We raise e to the term on both sides of the inequality, obtaining $\dfrac{1}{a_m} \geq m^k$.

Solving for a_m, we get $a_m \leq \dfrac{1}{m^k}$, $k > 1$, $\forall m \geq n_0$.

Then $\sum a_n$ is a minorant of a Riemann series ($\alpha > 1 \Rightarrow$ convergent), so it converges.

a.2. Si $\forall m \geq n_0$, $\dfrac{\ln(1/a_m)}{\ln m} \leq 1$, entonces $\ln(1/a_m) \leq \ln m$.

We raise e to the term on both sides of the inequality, obtaining $\dfrac{1}{a_m} \leq m$.

Solving for a_m, we get $a_m \geq \dfrac{1}{m}$

Then $\sum a_n$ is a majorant of the harmonic series, so it diverges.

Remark. In this test, as in Raabe's, convergence is obtained for values of l greater than 1 and divergence for values less than 1.

Application. This tests gives usually good results when a_n contains logarithms.

Examples. We study the following series:

1. $\sum \dfrac{1}{2^{\ln n}}$. We calculate $\lim\limits_{n\to\infty} \dfrac{\ln(1/a_n)}{\ln n} = \lim\limits_{n\to\infty} \dfrac{\ln 2^{\ln n}}{\ln n} = \lim\limits_{n\to\infty} \dfrac{\ln n \ln 2}{\ln n} = \ln 2 < 1 \Rightarrow$ D.

2. $\sum \dfrac{1}{(\ln n)^{\ln n}}$, $n > 1$. We calculate $\lim\limits_{n\to\infty} \dfrac{\ln(1/a_n)}{\ln n} = \lim\limits_{n\to\infty} \dfrac{\ln(\ln n)^{\ln n}}{\ln n} = \lim\limits_{n\to\infty} \dfrac{\ln n \ln(\ln n)}{\ln n} =$ $\lim\limits_{n\to\infty} \ln(\ln n) = \infty > 1 \Rightarrow$ C.

Exercises. Study the converge of: $\sum \dfrac{1}{(\ln n)^q}$, $n > 1$ (D) ; $\sum \dfrac{1}{\ln(n^{\ln n})}$, $n > 1$ (D).

6.8 Condensation test (comparison with $\sum 2^n a_{2^n}$)

Each one of the previous tests **6.4** to **6.7** is especially appropriate in certain cases (a_n with nth powers, factorials or logarithms, or when we get an inconclusive case by D'Alembert). The condensation test is sometimes able to solve inconclusive cases after applying some of these four tests. Therefore, it will generally not be the first one that we apply.

If $\{a_n\}$ is a **monotone decreasing** sequence, the following holds:

$$\boxed{\sum a_n \text{ converges if and only if } \sum 2^n a_{2^n} \text{ converges}}$$

and the same can be said about divergence.

Proof. We want to find a majorant and a minorant of $\sum a_n$. To do it we take into account the following:

1. Since $\sum a_n$ is a S.P.T. we can group terms without affecting the convergence.
2. Being $\{a_n\}$ monotone decreasing, we have that $a_1 \geq a_2 \geq a_3 \geq \dots$

Then:

$$\begin{aligned}
\sum a_n &= a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + (a_8 + \cdots + a_{15}) + \dots \\
\sum a_n &\geq a_2 + (a_4 + a_4) + (a_8 + a_8 + a_8 + a_8) + (a_{16} + \cdots + a_{16}) + \dots \\
\sum a_n &\leq a_1 + (a_2 + a_2) + (a_4 + a_4 + a_4 + a_4) + (a_8 + \cdots + a_8) + \dots
\end{aligned}$$

Grouping the equal addends in the above expressions, we get

$$a_2 + 2a_4 + 4a_8 + \cdots \leq \sum a_n \leq a_1 + 2a_2 + 4a_4 + 8a_8 \dots$$

that is

$$\frac{1}{2}\sum 2^n a_{2^n} \leq \sum a_n \leq a_1 + \sum 2^n a_{2^n}$$

From these inequalities we deduce (from the majorant and minorant criterion):

1. If $\sum 2^n a_{2^n}$ converges, so does $\sum a_n$ (which is a minorant).
2. If $\sum 2^n a_{2^n}$ diverges, so does $\frac{1}{2}\sum 2^n a_{2^n}$, then $\sum a_n$ diverges (it is a majorant).

Hence both series converge or both diverge.

Remark. To obtain the general term of the series $\sum 2^n a_{2^n}$, we replace n by 2^n in the general term and multiply it by the factor 2^n.

Examples. We study the following series.

1. $\sum \frac{1}{n \ln n}$. We see that $\sum 2^n \frac{1}{2^n \ln 2^n} = \sum \frac{1}{n \ln 2} = \frac{1}{\ln 2} \sum \frac{1}{n}$, which is divergent.
2. $\sum \frac{1}{\sqrt{n}\ln(\ln n)}$. $\sum 2^n a_{2^n} = \sum 2^n \frac{1}{\sqrt{2^n}\ln(\ln 2^n)} = \sum \frac{\sqrt{2^n}}{\ln(n \ln 2)} = \sum \frac{(\sqrt{2})^n}{\ln n + \ln(\ln 2)}$,

   which diverges because its general term tends to ∞.

Exercises. Study the convergence of: $\sum \frac{1}{\sqrt{n}\ln n}$ (D) ; $\sum \frac{1}{n(\ln n)^2}$ (C).

7 Series of positive and negative terms

We have studied the series of positive terms. To study those that contain negative terms, we will distinguish the following cases:

a) If all the terms of a series are negative, we can take the minus sign as a common factor multiplied by a S.P.T., that can be studied by the criteria seen for this type of series.

b) If there is only a finite number of negative terms, they do not affect the convergence. They can be added apart and study the resulting S.P.T. (property 1 of the series, section **4**).

c) Similarly, if the terms are negative, except for a finite number of positive terms, we take the minus sign as a common factor and obtain a S.P.T., with a finite number of negative ones (case b).

d) Thus, **the only type of series that raises new problems is the one with infinitely many terms of each sign**. To study it, we will first give some definitions.

7.1 Absolute and unconditional convergence and divergence

Given a series $\sum a_n$, we will say that it is:

a) **Absolutely convergent (divergent)** if and only if the series $\sum |a_n|$, formed by the absolute values of the terms is convergent (divergent).

b) **Unconditionally convergent or divergent** if and only if, when rearranging its terms, the series remains convergent or divergent; and the sum does not change if it is convergent.

Otherwise the series is said to be conditional.

We will study the convergence of $\sum a_n$ by decomposing its partial sum

$$S_n = \sum_{i=1}^{n} a_i$$

into positive and negative subseries. For this, we define:

$$a_i^+ = \begin{cases} a_i & a_i > 0 \\ 0 & a_i \leq 0 \end{cases}; \qquad a_i^- = \begin{cases} a_i & a_i < 0 \\ 0 & a_i \geq 0 \end{cases}$$

In this way each a_i can be written as the sum $a_i^+ + a_i^-$. We are doing this to decompose a sum of n terms into two sums of n terms, S_n^+ and S_n^-, in each of which we replace by zero those that are not of the corresponding sign.

Thus

$$S_n = \sum_{i=1}^{n} a_i \overset{(1)}{=} \sum_{i=1}^{n} (a_i^+ + a_i^-) \overset{(2)}{=} \sum_{i=1}^{n} a_i^+ + \sum_{i=1}^{n} a_i^- \overset{(3)}{=} \sum_{i=1}^{n} \left|a_i^+\right| - \sum_{i=1}^{n} \left|a_i^-\right| \overset{(4)}{=} S_n^+ - S_n^-$$

where we have taken the following steps:

1. We decompose $a_i = a_i^+ + a_i^-$.

2. We split the sum of n pairs of items into a pair of sums.

3. We take into account that $a_i^+ = \left|a_i^+\right|$ and $a_i^- = -\left|a_i^-\right|$.

4. We represent the sums of the $\left|a_i^\pm\right|$ as $S_n^\pm$.

Once S_n is decomposed, we notice that the $S_n^{\pm}$ do not contain negative terms, so they satisfy two conditions:

1. They either converge or diverge (they cannot be oscillating).
2. They are unconditionally convergent or divergent (we can reorder them without affecting neither the convergence nor the sum).

There exist the following options: that both converge, that only one diverges, or that both diverge. Thus, taking limits in each case, it results

a) $S_n^+ \to S^+$ and $S_n^- \to S^- \Longrightarrow \lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty} \left(S_n^+ - S_n^-\right) = S^+ - S^-$.

The series converges unconditionally to $S^+ - S^-$.

b) $S_n^+ \to +\infty$ and $S_n^- \to S^- \Longrightarrow \lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty} \left(S_n^+ - S_n^-\right) = +\infty$.

The series diverges unconditionally to $+\infty$.

c) $S_n^+ \to S^+$ and $S_n^- \to +\infty \Longrightarrow \lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty} \left(S_n^+ - S_n^-\right) = -\infty$.

The series diverges unconditionally to $-\infty$.

d) $S_n^+ \to \infty$ and $S_n^- \to \infty \Longrightarrow \lim\limits_{n\to\infty} S_n = \lim\limits_{n\to\infty} \left(S_n^+ - S_n^-\right) = ?$ (indetermination $\infty - \infty$)

The series can converge, diverge, or oscillate (see examples below).

That is, **only in the first case the series converges unconditionally and only in the first three cases is unconditional**.

Examples of case d. Three series are shown below that are decomposed into divergent subseries. The first case is an alternating series that converges by Leibnitz's theorem (**7.4**). In the second case, taking either an odd or an even number of terms, the partial sums tend to ∞ and the series diverges. In the third, taking an even number of terms, the partial sums are null; and, taking an odd number, they tend to infinity; therefore the series oscillates. This shows that the indetermination $\infty - \infty$ in the decomposition $S_n^+ - S_n^-$ can produce different results.

1) $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} \ldots \longrightarrow \ln 2\,,$ being $\begin{cases} S_n^+ = 1 + 1/3 + 1/5 + \ldots \\ S_n^- = 1/2 + 1/4 + 1/6 + \ldots \end{cases}$

2) $1 - 1 + 2 - \frac{1}{2} + 3 - \frac{1}{3} + \ldots \longrightarrow \infty\,,$ being $\begin{cases} S_n^+ = 1 + 2 + 3 + \ldots \\ S_n^- = 1 + 1/2 + 1/3 + \ldots \end{cases}$

3) $1 - 1 + 2 - 2 + 3 - 3 + \ldots \longrightarrow \begin{cases} \infty \\ 0\,, \end{cases}$ being both subseries $1 + 2 + 3 + \ldots$

Exercise. Find three series, one convergent, another divergent, and the third oscillating, whose positive and negative subseries are divergent.

7.2 Riemann's theorem

As we have seen, if the two subseries diverge, the resulting series can be of any type. The following theorem shows that, by properly rearranging its terms, we can make it diverge, oscillate, or converge to any sum we want.

Theorem: *Let $\sum a_n$ be a series such that the positive and negative subseries $\sum a_n^+$ y $\sum a_n^-$ are divergent and $a_n \to 0$ when $n \to \infty$. By rearranging its terms we can make it diverge, oscillate or converge to any sum we choose.*

Proof:

a) Divergent. To form the rearranged series we proceed as follows:

- We take positive terms, in the initial order, until their sum exceeds twice the absolute value of the first negative term. We then take the first negative.
- Then we take positive terms, in the initial order, until their sum exceeds twice the absolute value of the second negative term. Next, the second negative.
- And so on...

The process can be repeated indefinitely, since the positive subseries diverges, so the sums of the groups of positive terms that we are taking can exceed any value. It results:

$$\sum a'_n = \underbrace{a_1^+ + a_2^+ + \cdots + a_\alpha^+}_{\geq 2\left|a_1^-\right|} - \left|a_1^-\right| + \underbrace{a_{\alpha+1}^+ + \cdots + a_{\alpha+\beta}^+}_{\geq 2\left|a_2^-\right|} - \left|a_2^-\right| + \cdots \geq \left|a_1^-\right| + \left|a_2^-\right| + \ldots$$

In this way, the rearranged series is a majorant of a divergent one, then it diverges.

b) Oscillating. We choose $k_1, k_2 \,/\, k_1 < k_2$ and, always maintaining the initial order,

- We take positive terms until their sum exceeds k_2. Negative terms follow until the total sum is less than k_1.
- We take more positive terms until the total sum exceeds k_2 again. And negative terms until the total sum is less than k_1 again.
- And so on...

In each cycle, the absolute value of the difference between the partial sum and the corresponding value k_1 or k_2 is less than or equal to the absolute value of the last term added. Being $a_n \to 0$, the partial sums can get as close to k_1 or k_2 as we want, so the series rearranged in this way is oscillating.

c) Convergent. We choose the value S for the sum. Maintaining the initial order:

- We take positive terms until the value S is exceeded. Then negative terms until the total sum is less than S.
- We take again positive terms until S is exceeded and then negative terms until the total sum is less than S.
- And so on...

In each cycle, the absolute value of the difference between the partial sum and the value S is less than or equal to the absolute value of the last term added. Since $a_n \to 0$, the partial sums can get as close to S as we want, so the reordered series converges to S.

Remark. Notice that in the first case we have not used the $a_n \longrightarrow 0$ condition.

7.3 Dirichlet's theorem

$\sum a_n$ is uncondicionally convergent if and only if it is absolutely convergent.

Proof: To prove the theorem we will see that both the unconditional and the absolute convergence of the series are equivalent to the convergence of the subseries of positive terms $\sum |a_n^+|$ and $\sum |a_n^-|$.

a) **Unconditional convergence**: In section **7.1** we studied the unconditional convergence of a series of positive and negative terms, analyzing the four possibilities for the convergence of the subseries

   We concluded that the series is unconditionally convergent if $\sum |a_n^+|$ and $\sum |a_n^-|$ are convergent and **only in that case**.

b) **Absolute convergence**: The series of absolute values $\sum |a_n|$ will converge if and only if $\sum |a_n^+|$ y $\sum |a_n^-|$ are convergent. Indeed, decomposing the general term $|a_i|$, the series becomes the sum of two subseries:

$$\sum_{i=1}^{n} |a_i| = \sum_{i=1}^{n} \left(|a_i^+| + |a_i^-|\right) = \sum_{i=1}^{n} |a_i^+| + \sum_{i=1}^{n} |a_i^-|$$

   Then, since both addends are series of positive terms, we have that

   - If $\sum |a_n^+|$ y $\sum |a_n^-|$ converges, then $\sum |a_n|$ converges to the sum of both.
   - If $\sum |a_n|$ converges, then both $\sum |a_n^+|$ and $\sum |a_n^-|$ converge. Otherwise, one of the two would diverge and $\sum |a_n|$ would then diverge, against the hypothesis.

Application. Thus, the theorem indicates the **first step that we should take** to study the convergence of a series of positive and negative terms: **to analyze the series** $\sum |a_n|$.

a) If $\sum |a_n|$ converges, $\sum a_n$ will be unconditionally convergent, so we can rearrange it, decompose it into positive and negative subseries, etc.

b) If $\sum |a_n|$ diverges, we check if only one of the two subseries does, in which case $\sum a_n$ will diverge unconditionally.

c) If, in addition to $\sum |a_n|$, both subseries diverge, we can only study a partial sum S_n or apply the Leibnitz's theorem, if certain conditions are satisfied (see **7.4**).

   In this case of divergent subseries, if in addition $a_n \to 0$, the series can converge, diverge or oscillate depending on the order we give to its terms (Riemann's theorem). For this reason we must study it in the order given in the statement.

Remark. The theorem states the equivalence between absolute and unconditional convergence, while it does not say anything about absolute divergence. In the next section we will see examples of absolutely divergent but conditionally convergent series.

Exercise. Prove by *reductio ad absurdum* that, *if $\sum a_n$ is unconditionally divergent, then it is absolutely divergent.*

7.4 Alternating series. Leibnitz's theorem

We call **alternating** the series whose terms are alternatively positive or negative. Its general term has one of this two expressions:

$$(-1)^n a_n \qquad \text{or} \qquad (-1)^{n+1} a_n$$

We are going to study a particular type, in which the absolute value of the terms is decreasing. A partial sum S_n will be:

$$S_n = a_1 - a_2 + a_3 - \cdots + (-1)^{n+1} a_n \,; \ a_n > 0, \forall n; \ \ \{a_n\} \text{ monotone decreasing}$$

The Leibnitz's theorem (alternating series test) is useful to study the convergence of these series:

Leibnitz's theorem. *Any alternating series of decreasing terms in absolute value, such that its general term tends to zero, is convergent.*

Remark. It is shown that the sum of the series lies between any two consecutive partial sums. Therefore, when taking the value of a partial sum S_n as the sum S of the series, the error made is less than or equal to the absolute value of the first neglected term, that is a_{n+1}.

Example 1. The alternating series of the inverse of the even numbers:

$$\sum \frac{(-1)^{n+1}}{2n} = \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \ldots \quad \left(S = \frac{1}{2}\ln 2\right)$$

Ejemplo 2. The alternating series of the inverse of the odd numbers:

$$\sum \frac{(-1)^{n+1}}{2n-1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots \quad \left(S = \frac{\pi}{4}\right)$$

Example 3. Calculation of Euler's constant γ. To do it, we study the convergence of the following alternating series

$$\sum a_n = 1 - \ln\frac{2}{1} + \frac{1}{2} - \ln\frac{3}{2} + \frac{1}{3} - \ln\frac{4}{3} + \cdots + \frac{1}{n} - \ln\frac{n+1}{n} + \frac{1}{n+1} - \ldots$$

We check that the conditions of the theorem are satisfied:

a) It decreases in absolute value. The sequence $\left(1 + 1/n\right)^n$ is monotone increasing of limit e, while $\left(1 + 1/n\right)^{n+1}$ is monotone decreasing, with the same limit. So, $\forall n \in \mathbb{N}$, it holds

$$\left(1 + \frac{1}{n}\right)^{n+1} > e > \left(1 + \frac{1}{n}\right)^n \overset{\ln}{\Longrightarrow} (n+1)\ln\left(1 + \frac{1}{n}\right) > 1 > n\ln\left(1 + \frac{1}{n}\right) \Longrightarrow$$

$$\frac{1}{n} > \ln\left(1 + \frac{1}{n}\right) \text{ and } \ln\left(1 + \frac{1}{n}\right) > \frac{1}{n+1} \Longrightarrow \boxed{\frac{1}{n} > \ln\left(1 + \frac{1}{n}\right) > \frac{1}{n+1}}$$

b) The two expressions of the general term tend to 0, hence $a_n \to 0$:

$$\lim_{n\to\infty} \ln\left(\frac{n+1}{n}\right) = \lim_{n\to\infty} \frac{1}{n} = 0 \Longrightarrow \lim_{n\to\infty} a_n = 0$$

As a consequence, the series $\sum a_n$ converges. We denote their sum by γ (Euler's constant). The theorem assures us that we can obtain γ with the desired precision, provided we add a sufficient number of terms. Its value is $\gamma = 0.577215\ldots$

Application. Sum of the first n terms of the harmonic series.

The harmonic series is divergent. To obtain the value of the sum of its first n terms, we do the following: **1)** we take the first $2n$ terms of the series of Example 3; **2)** we separately group the positive and negative terms; **3)** we denote the sum of the first positive n terms by H_n and **4)** we include the negatives in a single logarithm and simplify, that is

$$S_{2n} = 1 - \ln\frac{2}{1} + \frac{1}{2} - \ln\frac{3}{2} + \cdots + \frac{1}{n} - \ln\frac{n+1}{n} = \underbrace{\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right)}_{H_n} - \underbrace{\ln\left(\frac{\cancel{2}}{1}\cdot\frac{\cancel{3}}{\cancel{2}}\cdots\frac{n+1}{\cancel{n}}\right)}_{\ln(n+1)}$$

As we saw in Example 3, the sum of this series is γ. Hence its partial sum S_{2n} will be equal to the value of its limit γ plus one infinitesimal θ_n. On the other hand, as we have just seen, S_{2n} can be written as the difference of H_n and $\ln(n+1)$. Thus

$$\lim_{n\to\infty} S_{2n} = \gamma \Longrightarrow S_{2n} = \gamma + \theta_n = H_n - \ln(n+1)$$

Solving for H_n,

$$H_n = \ln(n+1) + \gamma + \theta_n = \ln\frac{n+1}{n}\,n + \gamma + \theta_n = \underbrace{\ln\frac{n+1}{n}}_{\theta'_n} + \ln n + \gamma + \theta_n = \ln n + \gamma + \theta_n + \theta'_n$$

Grouping θ_n and θ'_n into a single infinitesimal ε_n, we obtain the value of the sum of the first n terms of the harmonic series

$$\boxed{H_n = \left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) = \ln n + \gamma + \varepsilon_n}$$

This expression tells us that H_n is an infinite equivalent to $\ln n$. It also will be useful to sum some types of series (section **8.2**) and to solve limits formed by a number of addends that tends to ∞, as we will see below.

Example 1. We calculate $L = \lim\limits_{n\to\infty}\left(\dfrac{1}{n+1} + \dfrac{1}{n+2} + \cdots + \dfrac{1}{2n}\right)$.

We notice that the denominators are formed by the consecutive natural numbers from $n+1$ to $2n$, so we can write the sum as the difference of H_{2n} and H_n. Then,

$$L = \lim_{n\to\infty}(H_{2n} - H_n) = \lim_{n\to\infty}\big(\underbrace{\ln 2n}_{\ln 2 + \ln n} + \gamma + \varepsilon_{2n} - [\ln n + \gamma + \varepsilon_n]\big) = \lim_{n\to\infty}(\ln 2 + \varepsilon_{2n} - \varepsilon_n) = \ln 2$$

Example 2. We calculate $L = \lim\limits_{n\to\infty}\left(\dfrac{1}{n+1} + \dfrac{1}{n+2} + \cdots + \dfrac{1}{n^2}\right)$.

Now the denominators are the consecutive natural numbers from $n+1$ to n^2, so we can write the sum as the difference of H_{n^2} and H_n. It results

$$L = \lim_{n\to\infty}(H_{n^2} - H_n) = \lim_{n\to\infty}\big(\ln(n^2) + \gamma + \varepsilon'_n - [\ln n + \gamma + \varepsilon_n]\big) = \lim_{n\to\infty}(\ln n + \varepsilon'_n - \varepsilon_n) = \infty$$

Exercise 1. Solve the limit: $\lim\limits_{n\to\infty}\left(\dfrac{1}{n+1} + \dfrac{1}{n+2} + \cdots + \dfrac{1}{3n}\right)$ $(L = \ln 3)$.

Exercise 2. Solve the limit: $\lim\limits_{n\to\infty}\left(\dfrac{1}{n^2+1} + \dfrac{1}{n^2+2} + \cdots + \dfrac{1}{n^3}\right)$ $(L = \infty)$.

8 Methods for the summation of series

In this section we will study different practical methods to sum some series.

8.1 By decomposition

We apply this method to series whose general term can be decomposed into the sum of others, which are often simple fractions. It consists in taking a partial sum, decomposing its terms, simplifying the resulting expression of S_n and calculating its limit.

a. Telescopic series. They are a particular case of series, in which a_n can be written as

$$a_n = b_n - b_{n+1}$$

Thus, S_n becomes

$$S_n = \sum_{i=1}^{n} a_i = \sum_{i=1}^{n} (b_i - b_{i+1}) = b_1 - \cancel{b_2} + \cancel{b_2} - \cancel{b_3} + \cdots + \cancel{b_{n-1}} - \cancel{b_n} + \cancel{b_n} - b_{n+1} = b_1 - b_{n+1}$$

Therefore

$$\boxed{S = \lim_{n\to\infty} S_n = b_1 - \lim_{n\to\infty} b_{n+1}}$$

Example 1. Find $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n^2 - n}$. First we see that $a_n \sim \dfrac{1}{n^2}$, convergent (Riemann, $\alpha > 1$).

We decompose the general term: $a_n = \dfrac{1}{n(n-1)} = \dfrac{1}{n-1} - \dfrac{1}{n}$

so it is a telescopic series. We simplify S_n

$$S_n = \sum_{i=2}^{n} \left(\frac{1}{i-1} - \frac{1}{i} \right) = 1 - \cancel{\frac{1}{2}} + \cancel{\frac{1}{2}} - \cancel{\frac{1}{3}} + \cdots - \cancel{\frac{1}{n-1}} + \cancel{\frac{1}{n-1}} - \frac{1}{n} = 1 - \frac{1}{n}$$

We calculate the sum

$$S = \lim_{n\to\infty} S_n = 1$$

Example 2. Calculate $\displaystyle\sum_{n=1}^{\infty} \frac{\sqrt{n+1} - \sqrt{n}}{\sqrt{2n^2 + 2n}}$. The general term is decomposed into

$$a_n = \frac{1}{\sqrt{2}} \frac{\sqrt{n+1} - \sqrt{n}}{\sqrt{n}\,\sqrt{n+1}} = \frac{1}{\sqrt{2}} \left(\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} \right)$$

hence it is a telescopic series. Operating as in Example 1, we get $S = 1/\sqrt{2}$.

Remark. To study the convergence, we can multiply the numerator and the denominator by the conjugate of the numerator, $\left(\sqrt{n+1} + \sqrt{n}\right)$. We obtain an infinitesimal equivalent to a_n:

$$a'_n = \frac{1}{2\sqrt{2}} \frac{1}{n^{3/2}}$$

which corresponds to a convergent series (Riemann, $\alpha = 3/2 > 1$). If the statement does not ask for it and we only want to obtain the sum, it is not essential to do this, because when calculating the limit of S_n we will know whether it converges or not.

Exercise. Find the sum: **a)** $\displaystyle\sum_{n=1}^{\infty} \frac{4n+1}{n^2(n+1)^2}$ $(S = 2)$. **b)** $\displaystyle\sum_{n=2}^{\infty} \frac{\sqrt{3n} - \sqrt{3n-3}}{\sqrt{n^2 - n}}$ $(S = \sqrt{3})$.

b. Decomposition into simple fractions. If the general term is formed by a quotient of polynomials

$$a_n = \frac{P_k(n)}{Q_l(n)}, \quad k < l$$

we decompose it into simple fractions and simplify S_n. We will see examples of two of the most common cases.

Example 1. To calculate $\displaystyle\sum_{n=2}^{\infty} \frac{n-2}{n^3-n}$, we first decompose a_n, by identifying the coefficients:

$$a_n = \frac{n-2}{(n-1)n(n+1)} = \frac{A}{n-1} + \frac{B}{n} + \frac{C}{n+1} \Longrightarrow A = -\frac{1}{2},\ B = 2,\ C = -\frac{3}{2}$$

We notice that the coefficients add up to 0. The reason is that the degree of the numerator is equal to 1; so, after making common denominator, the coefficient of n^2 in the numerator of the right-hand side (that is, $A+B+C$) is zero.

To obtain S_n, we will decompose the sum of the terms a_i into three different sums. To ease the simplification we denote by H_n the sum of the first inverse of the naturals

$$H_n = \sum_{i=1}^{n} \frac{1}{i} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}$$

Then

$$S_n = \sum_{i=2}^{n} \left(-\frac{1/2}{i-1} + \frac{2}{i} - \frac{3/2}{i+1} \right) = -\frac{1}{2}\sum_{i=2}^{n} \frac{1}{i-1} + 2\sum_{i=2}^{n} \frac{1}{i} - \frac{3}{2}\sum_{i=2}^{n} \frac{1}{i+1} =$$

$$-\frac{1}{2}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n-1}\right) + 2\left(\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}\right) - \frac{3}{2}\left(\frac{1}{3} + \frac{1}{4} + \cdots + \frac{1}{n+1}\right) =$$

$$-\frac{1}{2}\left(\cancel{H_n} - \frac{1}{n}\right) + 2\left(\cancel{H_n} - 1\right) - \frac{3}{2}\left(\cancel{H_n} - 1 - \frac{1}{2} + \frac{1}{n+1}\right) = \frac{1}{2n} - 2 + \frac{9}{4} - \frac{3/2}{n+1}$$

Taking limits, it results $S = \lim\limits_{n\to\infty} S_n = 1/4$.

Example 2. Find $\displaystyle\sum_{n=2}^{\infty} \frac{1}{n^3-n^2}$, knowing that $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

We decompose a_n, separating the three resulting terms into two parts:

$$a_n = \frac{1}{n^2(n-1)} = \frac{A}{n} + \frac{B}{n^2} + \frac{C}{n-1} = \cdots = -\frac{1}{n} - \frac{1}{n^2} + \frac{1}{n-1} = \left(\frac{1}{n-1} - \frac{1}{n}\right) - \frac{1}{n^2}$$

We obtain S_n and take limits

$$S_n = \sum_{i=2}^{n} \left(\frac{1}{i-1} - \frac{1}{i}\right) - \sum_{i=2}^{n} \frac{1}{i^2} = \left(1 - \cancel{\frac{1}{2}} + \cancel{\frac{1}{2}} - \cancel{\frac{1}{3}} + \cdots + \cancel{\frac{1}{n-1}} - \frac{1}{n}\right) - \left(\sum_{i=1}^{n} \frac{1}{i^2} - 1\right) =$$

$$1 - \frac{1}{n} + 1 - \sum_{i=1}^{n} \frac{1}{i^2} \Longrightarrow S = \lim_{n\to\infty} S_n = 2 - \sum_{n=1}^{\infty} \frac{1}{i^2} = 2 - \frac{\pi^2}{6}$$

Exercise. Find the sum: **a)** $\displaystyle\sum_{n=2}^{\infty} \frac{n+2}{n^3-n}$ $(S = 5/4)$. **b)** $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^3+n^2}$ $\left(S = \frac{\pi^2}{6} - 1\right)$.

8.2 From the harmonic series

The harmonic series is divergent, but we can use it to sum different series of positive and negative terms. As in the previous section, we denote the sum of its first n terms by H_n, whose value (see **7.4**) is

$$\boxed{H_n = \sum_{i=1}^{n} \frac{1}{i} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} = \ln n + \gamma + \varepsilon_n}$$

If we consider the first $2n$ terms, H_{2n}, its sum is

$$H_{2n} = \sum_{i=1}^{2n} \frac{1}{i} = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{2n} = \ln(2n) + \gamma + \varepsilon_{2n}$$

We represent by P_n the sum of the first n inverses of the even numbers and obtain its value:

$$P_n = \sum_{i=1}^{n} \frac{1}{2i} = \frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2n} = \frac{1}{2}\left(1 + \frac{1}{2} + \cdots + \frac{1}{n}\right) = \frac{1}{2}H_n \Longrightarrow \boxed{P_n = \frac{1}{2}(\ln n + \gamma) + \varepsilon'_n}$$

I_n is the sum of the first n inverses of the odd numbers. To calculate its value, we take the first $2n$ terms of the harmonic and eliminate the first n inverses of the even numbers (P_n):

$$I_n = \sum_{i=1}^{n} \frac{1}{2i-1} = 1 + \frac{1}{3} + \cdots + \frac{1}{2n-1} = H_{2n} - P_n = \ln 2n + \gamma + \varepsilon_{2n} - \left[\frac{1}{2}(\ln n + \gamma) + \varepsilon'_n\right] =$$

$$\ln 2 + (\ln n + \gamma) + \varepsilon_{2n} - \frac{1}{2}(\ln n + \gamma) - \varepsilon'_n \Longrightarrow \boxed{I_n = \ln 2 + \frac{1}{2}(\ln n + \gamma) + \varepsilon''_n}$$

Thus, H_n, I_n and P_n are divergent. It is important to note that the subscript indicates the number of terms, while the type (inverses of natural, odd or even) is given by H, I or P.

Example 1. Sum of the alternating harmonic series.

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

We take $2n$ terms of the series, of which n will be from I_n and the other n from P_n:

$$S_{2n} = 1 - \frac{1}{2} + \frac{1}{3} - \cdots + \frac{1}{2n-1} - \frac{1}{2n} = \left(1 + \frac{1}{3} + \cdots + \frac{1}{2n-1}\right) - \left(\frac{1}{2} + \frac{1}{4} + \cdots + \frac{1}{2n}\right) =$$

$$I_n - P_n = \ln 2 + \cancel{\frac{1}{2}(\ln n + \gamma)} + \varepsilon''_n - \cancel{\frac{1}{2}(\ln n + \gamma)} - \varepsilon'_n \Longrightarrow S = \lim_{n\to\infty} S_{2n} = \ln 2$$

Example 2. Sum of $1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \cdots$

We see that the terms follow a certain pattern, which is repeated every three terms, so we take $3n$ terms, that is, n groups of 3. In each group there is one inverse of an odd number (+) and two inverses of even numbers (−), so there will be n odd and $2n$ even in total. Hence:

$$S_{3n} = I_n - P_{2n} = \ln 2 + \frac{1}{2}(\ln n + \gamma) + \varepsilon_1 - \frac{1}{2}(\ln 2n + \gamma) - \varepsilon_2 \Longrightarrow S = \lim_{n\to\infty} S_{3n} = \frac{1}{2}\ln 2$$

Exercise. Find the sum of: $1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \cdots \quad (S = \frac{3}{2}\ln 2)$

8.3 From the series expansion of e^x

The sum of $\sum \dfrac{P_k(n)}{n!}\,\alpha^n$, where $P_k(n)$ is a polynomial of degree k and $\alpha \in \mathbb{R}$, can be obtained from the series expansion of e^x

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \frac{x^0}{0!} + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

which will be studied in the next unit but is assumed to be known here.

Convergence. The parameter α can be negative, so we study $|a_n|$:

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty} \left|\frac{P_k(n+1)\,\alpha^{n+1}}{(n+1)!} \; \frac{n!}{P_k(n)\,\alpha^n}\right| = \lim_{n\to\infty} \left|\frac{P_k(n+1)}{P_k(n)} \, \frac{\alpha}{n+1}\right| = 0 < 1$$

hence, $\forall \alpha \in \mathbb{R}$, convergence is absolute, therefore unconditional (Dirichlet's theorem).

Sum. To simplify numerator and denominator, we write $P_k(n)$ as follows

$$P_k(n) = A_0 + A_1\, n + A_2\, n(n-1) + \cdots + A_k\, n(n-1)\cdots\big[n-(k-1)\big]$$

whose last addend -with k factors- has degree k. Now we decompose the general term $P_k(n)/n!$ into $k+1$ addends, so the series becomes a linear combination of convergent series, which we will add using the expansion of e^x.
In case that the denominator is not the factorial of n, but of $n \pm p$, the polynomial $P_k(n)$ will be decomposed into a linear combination of products of decreasing factors starting from $n \pm p$.

Example. We calculate the sum $\displaystyle\sum_{n=0}^{\infty} \frac{2n^2+n+1}{n!}\,2^n$.
We descompose $P_2(n)$: $2n^2+n+1 = A + Bn + Cn(n-1) \Longrightarrow A=1,\ B=3,\ C=2.$ Then

$$a_n = \frac{1+3n+2n(n-1)}{n!}\,2^n = \frac{1}{n!}\,2^n + \frac{3n}{n!}\,2^n + \frac{2n(n-1)}{n!}\,2^n$$

with which

$$S_n = \sum_{i=0}^{n} \frac{2^i}{i!} + 3\sum_{i=0}^{n} \frac{i\,2^i}{i!} + 2\sum_{i=0}^{n} \frac{i(i-1)\,2^i}{i!} = \sum_{i=0}^{n} \frac{2^i}{i!} + 3\sum_{i=1}^{n} \frac{i\,2^i}{i!} + 2\sum_{i=2}^{n} \frac{i(i-1)\,2^i}{i!}$$

(the addends with a null numerator have been eliminated in the second and third summations).
Simplifying the factors of the numerator with the factorials, we get

$$S_n = \sum_{i=0}^{n} \frac{2^i}{i!} + 3\cdot 2\sum_{i=1}^{n} \frac{2^{i-1}}{(i-1)!} + 2\cdot 2^2\sum_{i=2}^{n} \frac{2^{i-2}}{(i-2)!}$$

where the three summations have the same terms, $\dfrac{2^0}{0!} + \dfrac{2^1}{1!} + \dfrac{2^2}{2!} + \cdots$, and tend to e^2. Thus

$$S = \lim_{n\to\infty} S_n = \sum_{n=0}^{\infty} \frac{2^i}{i!} + 6\sum_{n=1}^{\infty} \frac{2^{i-1}}{(i-1)!} + 8\sum_{n=2}^{\infty} \frac{2^{i-2}}{(i-2)!} = e^2 + 6\,e^2 + 8\,e^2 = 15\,e^2$$

Exercise. Find **a)** $\displaystyle\sum_{n=1}^{\infty} \frac{n^2+n+1}{n!}$ $\quad(S = 4e - 1)$; **b)** $\displaystyle\sum_{n=0}^{\infty} \frac{2n^2+1}{(n+2)!}3^n$ $\quad(S = e^3 - 2)$

8.4 Hypergeometric series

Definition. A series is hypergeometric if its general term satisfies the condition

$$\boxed{\frac{a_{n+1}}{a_n} = \frac{\alpha\, n + \beta}{\alpha\, n + \gamma}}\ (\alpha > 0, \gamma \neq 0)$$

Convergence. We study it by applying the Raabe's test.

$$\lim_{n\to\infty} n\left(1 - \frac{a_{n+1}}{a_n}\right) = \lim_{n\to\infty} n\left(1 - \frac{\alpha\, n + \beta}{\alpha\, n + \gamma}\right) = \lim_{n\to\infty} n\left(\frac{\gamma - \beta}{\alpha\, n + \gamma}\right) = \frac{\gamma - \beta}{\alpha}$$

- If $\dfrac{\gamma - \beta}{\alpha} > 1 \Longrightarrow \boxed{\alpha + \beta < \gamma}$, the series converges.
- If $\dfrac{\gamma - \beta}{\alpha} < 1 \Longrightarrow \boxed{\alpha + \beta > \gamma}$, the series diverges.
- If $\dfrac{\gamma - \beta}{\alpha} = 1 \Longrightarrow \boxed{\alpha + \beta = \gamma}$, it is an inconclusive case (as we will see below, it diverges).

Sum. If $\alpha + \beta < \gamma$, we obtain the sum using the condition of hypergeometric series

$$\frac{a_{i+1}}{a_i} = \frac{\alpha\, i + \beta}{\alpha\, i + \gamma} \Longrightarrow a_{i+1}\,(\alpha\, i + \gamma) = a_i\,(\alpha\, i + \beta)\ \forall i \in \mathbb{N}$$

Giving values to i:

$$\begin{array}{ll}
i = 1: & a_2\,(1\cdot\alpha + \gamma) = a_1\,(1\cdot\alpha + \beta) \\
i = 2: & a_3\,(2\cdot\alpha + \gamma) = a_2\,(2\cdot\alpha + \beta) \\
i = 3: & a_4\,(3\cdot\alpha + \gamma) = a_3\,(3\cdot\alpha + \beta) \\
\vdots & \qquad\qquad \vdots \\
i = n-1: & a_n\big((n-1)\alpha + \gamma\big) = a_{n-1}\big((n-1)\alpha + \beta\big) \\
 & a_n\,(n\alpha + \beta) \qquad\quad = a_n\,(n\alpha + \beta) \quad \text{(we add an identity)}
\end{array}$$

We simplify the α on both sides and add the resulting equalities. If S_n is the sum of the first n terms, we obtain:

$$(S_n - a_1)\,\gamma + a_n(n\alpha + \beta) = S_n(\alpha + \beta) \Longrightarrow$$

$$S_n\big[\gamma - (\alpha + \beta)\big] = a_1\gamma - a_n(n\alpha + \beta) \Longrightarrow \tag{1}$$

$$S_n = \frac{a_1\gamma}{\gamma - (\alpha + \beta)} - \underbrace{\frac{a_n(n\alpha + \beta)}{\gamma - (\alpha + \beta)}}_{b_n} \qquad \big(\gamma - (\alpha + \beta) \neq 0\big) \tag{2}$$

As we assume that $\alpha + \beta < \gamma$, the series converges, so

$$\exists\, S = \lim_{n\to\infty} S_n = \frac{a_1\gamma}{\gamma - (\alpha + \beta)} - \lim_{n\to\infty} b_n$$

Since S exists, the $\lim\limits_{n\to\infty} b_n$ must exist. We will see that it is null by *reductio ad absurdum.*

Let us assume that the limit of b_n is not null. Operating from (2) we obtain for b_n the expression

$$b_n = \frac{a_n(n\alpha+\beta)}{\gamma-(\alpha+\beta)} = \frac{1}{\gamma-(\alpha+\beta)}\left(a_n : \frac{1}{n\alpha+\beta}\right)$$

and taking limits

$$\lim_{n\to\infty} b_n = \frac{1}{\gamma-(\alpha+\beta)}\lim_{n\to\infty}\left(a_n : \frac{1}{n\alpha+\beta}\right) = k \neq 0$$

Since the series

$$\sum \frac{1}{n\alpha+\beta}$$

diverges by comparison with the harmonic, $\sum a_n$ would diverge too (see **6.3**), against the hypothesis.

Thus we conclude that $b_n \to 0$ and the sum of the series is

$$\boxed{S = \frac{a_1\gamma}{\gamma-(\alpha+\beta)}} \quad (\text{ if } \alpha+\beta<\gamma)$$

Study of the inconclusive case. To solve the case $\alpha+\beta=\gamma$, we consider the equality (1) in the previous page. The left member is null and solving for a_n, we obtain

$$0 = a_1\gamma - a_n(n\alpha+\beta) \Longrightarrow a_n = a_1\gamma\,\frac{1}{n\alpha+\beta}$$

which corresponds to a divergent series.

Example. We calculate the sum of the following series, whose convergence was studied in section **6.6**.

$$\sum_{n=1}^{\infty} \frac{a(a+1)\cdots(a+n)}{b(b+1)\cdots(b+n)}$$

Simplifying the quotient

$$\frac{a_{n+1}}{a_n} = \cdots = \frac{a+n+1}{b+n+1} = \frac{n+a+1}{n+b+1}$$

we see that the series is hypergeometric, with

$$\alpha = 1,\ \beta = a+1,\ \gamma = b+1$$

As we know, the series will converge if

$$\gamma > \alpha+\beta \Longleftrightarrow b > a+1$$

To calculate the sum, we first obtain a_1 and replace in the formula of S.

$$a_1 = \frac{a\,(a+1)}{b\,(b+1)} \Longrightarrow S = \frac{a_1\gamma}{\gamma-(\alpha+\beta)} = \frac{a\,(a+1)}{b\,(b+1)}\,\frac{b+1}{b+1-(a+2)} = \frac{a\,(a+1)}{b\,(b-a-1)}$$

Exercise 1. Calculate the sums of the following series, checking first that they are hypergeometric (they can also be solved as telescopic).

a) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2n(2n+2)} \quad \left(S=\frac{1}{4}\right)$; **b)** $\displaystyle\sum_{n=1}^{\infty} \frac{1}{(2n-1)(2n+1)} \quad \left(S=\frac{1}{2}\right)$

Exercise 2. Study the convergence and calculate the sum of

$$\sum_{n=1}^{\infty} \frac{n!}{a(a+1)\cdots(a+n-1)} \quad \left(\text{If } a>2,\ \text{it converges to } S=\frac{1}{a-2}\right)$$

8.5 Convergence and sum of a series. Summary

To finish this section, the main steps that should be taken in the study of the convergence and sum of a series are shown below.

Series of positive terms. If $\sum a_n$ is a series of positive terms, we apply one of the criteria studied, trying to use the one that best suits the type of series in question. To sum it, if it converges, we can rearrange its terms, group them, or decompose them into a sum of positive terms, since the sum does not change (properties **3**, **6** and **7** of the series).

Series of positive and negative terms . If the series has infinitely many positive and infinitely negative terms, the first step is to study the series $\sum |a_n|$, which gives us two possible results:

a) **Convergent**. The series is absolutely convergent, so it is unconditionally convergent (Dirichlet). This allows us to rearrange or group its terms, e.g. separating positive from negative.

b) **Divergent**. If only one of the two subseries (positive or negative) diverges, the series diverges unconditionally to $\pm\infty$ (section **7.1**). If both diverge (the most frequent case), we distinguish two options:

- **b.1** If it is an alternating series, we apply Leibnitz's theorem.
- **b.2** Otherwise, we resort to the general method: to take a partial sum S_n (where we can group, simplify, etc.) and study its limit.

Remark. In both cases (whether alternating or not), if there is convergence, it is conditional (it holds for the given order of terms, but may not hold for others).

Particular case. Sometimes, to sum an absolutely convergent series, we decompose its general term a_n into sums or differences of terms $(b_n \pm c_n \pm \dots)$. By doing this (decomposition into positive and negative terms), the resulting series no longer has to be absolutely convergent and we have two options.

a) If the terms $b_n, c_n, \dots$ correspond to **convergent series**, we apply the property **5** ("*the c.l. of convergent series converges to the c.l. of the sums of the series*"). Since the series $\sum a_n$ is a linear combination of $\sum b_n, \sum c_n \dots$, its sum will be $S_b \pm S_c \pm \dots$, being $S_b, S_c, \dots$ the sums of the series $\sum b_n, \sum c_n, \dots$, respectively.

b) If **two or more** of the terms $b_n, c_n, \dots$ correspond to **divergent series**, we must analyze S_n, applying one of the methods studied: telescopic , decomposition into fractions, series I_n y P_n, etc. In any case, we cannot sum the different subseries $\sum b_n, \sum c_n, \dots$ separately.

Example: $\sum \frac{1}{n(n+1)}$ converges by comparison with $\sum \frac{1}{n^2}$. To sum it, its general term is decomposed into $\frac{1}{n} - \frac{1}{n+1}$ (difference of general terms of divergent series). Its sum can be calculated as a telescopic series or from the sum of n terms of the harmonic series.

Exercise. We have studied the cases **a)** (all convergent) and **b)** (two or more divergent). Reason that it cannot happen that **only one** of the terms $b_n, c_n, \dots$ corresponds to a divergent series.

9 Self-assesment exercises

9.1 True/False exercises

Exercise 1 Decide whether the following statements are true or false.

1. $S_n = \sum_{i=1}^{n} a_i$ is a partial sum of $\sum a_n$. The series diverges if and only if $\lim_{n\to\infty} S_n = \pm\infty$.
2. A geometric series converges if the ratio is less than 1.
3. If $\lim_{n\to\infty} a_n \neq 0$, the series is divergent.
4. We group the terms of a series of positive terms $\sum a_n$, obtaining $\sum a'_n$. Whether $\sum a_n$ converges or diverges, $\sum a'_n$ does the same.
5. If a series converges, the sum of any number of consecutive terms, starting from a given p, tends to zero if $p \to \infty$.
6. If a series does not have negative terms, it can only converge or diverge.
7. We say that $\sum a_n$ is a minorant of $\sum b_n$ if and only if $a_n < b_n$, $\forall n$.
8. The harmonic series $\sum \frac{1}{n}$ satisfies $\lim_{n\to\infty} a_n = 0$, so it converges.
9. The series $\sum \frac{1}{\sqrt{n^2+1}}$ is convergent, since the exponent of n is $\alpha = 2 > 1$.
10. The series $\sum 3^n/n^3$ satisfy the necessary condition of convergence, but is divergent by the root test.
11. If the limit of the quotient test exists, the limit of the nth root exists and they coincide.
12. If D'Alembert's test does not solve the character of a series, it does not make sense to apply Raabe's and it is preferable, in general, to use Cauchy's.
13. To know the character of $\sum \frac{1}{e^{2\ln n}}$ it is necessary to apply the logarithmic criterion.
14. According to the condensation test, the series $\sum a_n$ and $\sum 2^n a_{2^n}$ both converge or both diverge.
15. Study the convergence of the following series:

15.1.- $\sum \frac{1}{\alpha^n}$, $\alpha > 1$

15.2.- $\sum \frac{1}{n^\alpha}$, $\alpha > 1$

15.3.- $\sum \frac{1}{\ln n}$

15.4.- $\sum \frac{1}{n+\ln n}$

15.5.- $\sum \frac{1}{3^{\ln n}}$

15.6.- $\sum \frac{1}{\sqrt{2^n}} = \sum \frac{1}{2^{n/2}} = \sum \frac{1}{\left(\sqrt{2}\right)^n}$

Exercise 2 Decide whether the following statements are true or false.

1. Let a series be $\sum a_n$. Let a partial sum be $S_n = \sum_1^n a_i = \sum_1^n \left|a_i^+\right| - \sum_1^n \left|a_i^-\right| = S_n^+ - S_n^-$. $\sum a_n$ converges unconditionally if and only if S_n^+ and S_n^- converge.

2. If $\sum a_n$ is absolutely divergent, then it is unconditionally divergent.

3. The series $\sum(-1)^{n+1}/\ln n,\ n > 1$ is absolutely divergent, therefore we cannot guarantee anything about its convergence.

4. The series of general term $a_n = \dfrac{P_k(n)}{Q_k(n)}$ are convergent and their sum can be obtained by decomposition into simple fractions.

5. To sum $1 + \frac{1}{5} + \frac{1}{3} - \frac{1}{2} + \frac{1}{7} + \frac{1}{11} + \frac{1}{9} - \frac{1}{4} + \frac{1}{13} + \frac{1}{17} + \frac{1}{15} - \frac{1}{6} \dots$ we calculate the limit of the partial sum of $4n$ terms, which is equal to that of $I_{3n} - P_n$.

6. To sum $\sum \dfrac{P_3(n)}{(n+3)!}$, we have to perform the decomposition:

$$P_3(n) = A + B(n+3) + C(n+3)(n+2) + D(n+3)(n+2)(n+1)$$

7. Study the convergence of the following series finding, if possible, their sum and justifying it briefly:

   7.1. $\sum \dfrac{-1}{n}$

   7.2. $\sum \left(\dfrac{1}{\sqrt{n}} - \dfrac{1}{\sqrt{n+1}} \right)$

   7.3. $-1 + \frac{1}{2} + \frac{1}{4} - \frac{1}{3} + \frac{1}{6} + \frac{1}{8} - \frac{1}{5} + \frac{1}{10} + \frac{1}{12} - \dots$

   7.4. $1 - \frac{1}{3} + \frac{1}{9} - \frac{1}{27} + \frac{1}{81} - \dots$

9.2 Question

Let $\sum a_n$ be a series of positive terms. What can we say –and why– about whether $\sum a_n$ is convergent, divergent, oscillating, conditional, unconditional?

9.3 Solution to the True/False exercises

Exercise 1.

1. **F**. The series diverges if and only if $\lim\limits_{n\to\infty} |S_n| = \infty$.

2. **F**. A geometric series converges if the **absolute value** of the ratio is less than 1.

3. **F**. If $\lim\limits_{n\to\infty} a_n \neq 0$, the series does not converge, hence it can be divergent or oscillating.

4. **T**. Property 7 of the series.

5. **T**. Cauchy's criterion of convergence.

6. **T**. The series of positive terms never oscillate.

7. **F**. We say that $\sum a_n$ is a minorant of $\sum b_n$ if and only if $a_n < b_n$, $\boxed{\forall n \geq n_1}$.

8. **F**. The condition $\lim\limits_{n\to\infty} a_n = 0$ is necessary for convergence, not sufficient.

9. **F**. The condition $\alpha > 1$ is for the Riemann series $\sum \dfrac{1}{n^\alpha}$. In this case, $a_n \sim 1/n$, so it diverges as $\sum \dfrac{1}{n}$ (harmonic).

10. **F**. The series $\sum 3^n/n^3$ does not satisfy the necessary condition of convergence ($a_n \not\to 0$), although it is true that it is divergent by the root test.

11. **T**. From the "root law".

12. **F**. The Raabe's test is especially suitable for solving cases that are inconclusive by the quotient test.

13. **F**. It suffices to see that $a_n = \dfrac{1}{e^{\ln(n^2)}} = \dfrac{1}{n^2}$ (Riemann series, $\alpha = 2$ >1, convergent)

14. **F**. **If $\{a_n\}$ is monotone decreasing**, the series $\sum a_n$ and $\sum 2^n a_{2^n}$ both converge or both diverge.

15. Study the convergence of the following series:

15.1.- $\sum \dfrac{1}{\alpha^n}$, $\alpha > 1$: **C**, geometric series of ratio $\dfrac{1}{\alpha} < 1$.

15.2.- $\sum \dfrac{1}{n^\alpha}$, $\alpha > 1$: **C**, Riemann series, $\alpha > 1$.

15.3.- $\sum \dfrac{1}{\ln n}$: **D**, majorant of the harmonic.

15.4.- $\sum \dfrac{1}{n + \ln n}$: **D**, $a_n \sim \dfrac{1}{n}$, divergent as the harmonic.

15.5.- $\sum \dfrac{1}{3^{\ln n}}$: **C**, by the logarithmic test.

15.6.- $\sum \dfrac{1}{\sqrt{2^n}} = \sum \dfrac{1}{2^{n/2}} = \sum \dfrac{1}{\left(\sqrt{2}\right)^n}$: **C**, by the root test.

Exercise 2.

1. **T**. See section **7.1**.

2. **F**. For example, the alternate harmonic $\sum(-1)^{n+1}/n$ is absolutely divergent. But it converges (conditionally) to $\ln 2$. On the other hand, from Dirichlet's theorem it is shown that, if a series is unconditionally divergent, it is absolutely divergent.

3. **F**. It is alternate, the terms decrease in absolute value and $a_n \to 0$. Leibnitz's theorem assures that it is convergent in the order given by the statement (conditionally convergent).

4. **F**. If the numerator and denominator have the same degree, the general term does not tend to zero, so the series does not converge.

5. **T**. We calculate the limit of the partial sum of $4n$ terms. In each group of 4 terms there are 3 positive (inverse of odd numbers) and one negative (inverse of an even number), so in S_{4n} we have $3n$ positive (inverse of odd numbers) and n negative (inverse of an even number). We see that $S_{4n} = I_{3n} - P_n$ and, simplifying and taking limits, we obtain for the sum the value $S = \ln(2\sqrt{3})$.

6. **T**. See section **8.3**.

7. Study the convergence of the following series finding, if possible, their sum and justifying it briefly:

   7.1. $\displaystyle\sum_{n=1}^{\infty} \frac{-1}{n}$: It is the harmonic multiplied by -1, so divergent to $-\infty$.

   7.2. $\displaystyle\sum_{n=1}^{\infty} \left(\frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}}\right)$: It is telescopic, since $a_n = b_n - b_{n+1}$. Its partial sum is

   $$S_n = \frac{1}{\sqrt{1}} - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{3}} + \ldots \frac{1}{\sqrt{n}} - \frac{1}{\sqrt{n+1}} = \frac{1}{\sqrt{1}} - \frac{1}{\sqrt{n+1}} \Rightarrow \lim_{n\to\infty} S_n = 1$$

   The series converges (in the given order).

   7.3. $-1 + \frac{1}{2} + \frac{1}{4} - \frac{1}{3} + \frac{1}{6} + \frac{1}{8} - \frac{1}{5} + \frac{1}{10} + \frac{1}{12} - \ldots$:

   We take a partial sum of $3n$ terms. In each group of 3 terms there will be one negative and two positive. The denominators of the negative terms are consecutive odd numbers, while the denominators of the positive ones are consecutive even numbers. Then, in S_{3n} there will be n inverses of the odd, negative, and $2n$ inverses of the even, positive. We have: $S_{3n} = -I_n + P_{2n}$ and we obtain for the sum the value $S = -\ln\sqrt{2}$.

   7.4. $1 - \frac{1}{3} + \frac{1}{9} - \frac{1}{27} + \frac{1}{81} \ldots$: It is a geometric series of ratio $r = -1/3$, $|r| < 1$, therefore convergent. Its sum is $S = \dfrac{a_1}{1-r} = 3/4$.

9.4 Solution to the question

If $\sum a_n$ is a series of positive terms ($a_i \geq 0\ \forall i \in \mathbb{N}$), then the sequence of its partial sums

$$S_1 = a_1;\ S_2 = a_1 + a_2;\ \ldots\ S_n = a_1 + a_2 + \ldots a_n;\ \ldots$$

verifies

$$S_1 \leq S_2 \leq \cdots \leq S_n \leq S_{n+1} \leq \ldots$$

that is, it will be monotone increasing. If it is bounded, the series will be **convergent** (since every monotone increasing sequence bounded from above is convergent). Otherwise, it will be **divergent** to $+\infty$. That is, a series of positive terms is never oscillating.

On the other hand, property 6 of the series says that "if we alter the order of the terms of a series of positive terms, the behavior does not change, nor does the sum if it is convergent". This means that the series of positive terms are always **unconditional**.

We can therefore affirm that $\sum a_n$ can only be unconditionally convergent or unconditionally divergent.

Unit IV. Sequences and series of functions (17.04.2024)

1 Sequences of functions

1.1 Distance between functions

Let $\mathcal{F}_b(I, \mathbb{R})$ be the set of bounded real functions defined on an interval $I \subset \mathbb{R}$.
Given $f, g \in \mathcal{F}_b(I, \mathbb{R})$, we define the **distance** between them as

$$\boxed{d(f, g) = \sup |f(x) - g(x)|\,,\ x \in I}$$

that is, the supremum of the point-to-point distances between them, when x takes values on I.
With this definition of distance, it is verified that $\mathcal{F}_b(I, \mathbb{R})$ is a metric space.

Example 1. The distance between $f(x) = a + x$ and $g(x) = b + x$ is $d(f, g) = |a - b|$.

Example 2. The functions $f(x) = \dfrac{1}{1 + x^2}$ and $g(x) = \dfrac{4}{1 + x^2}$ verify:

$$|f(x) - g(x)| = \frac{3}{1 + x^2} \Longrightarrow d(f, g) = \sup \left.\frac{3}{1 + x^2}\right|_{x \in \mathbb{R}} = 3$$

This value is reached at $x = 0$, so in addition to the supremum it represents the maximum.

Remark. The supremum is not necessarily reached. For example, if we multiply the numerators of f and g by x^2, the distance between them is still 3, but the supremum corresponds to $x \to \infty$.

Exercise. Find the distance between $f(x) = \sqrt{x}$ and $g(x) = x^3$, $x \in [0, 1]$ $\left(d = 5 \cdot 6^{-6/5}\right)$.

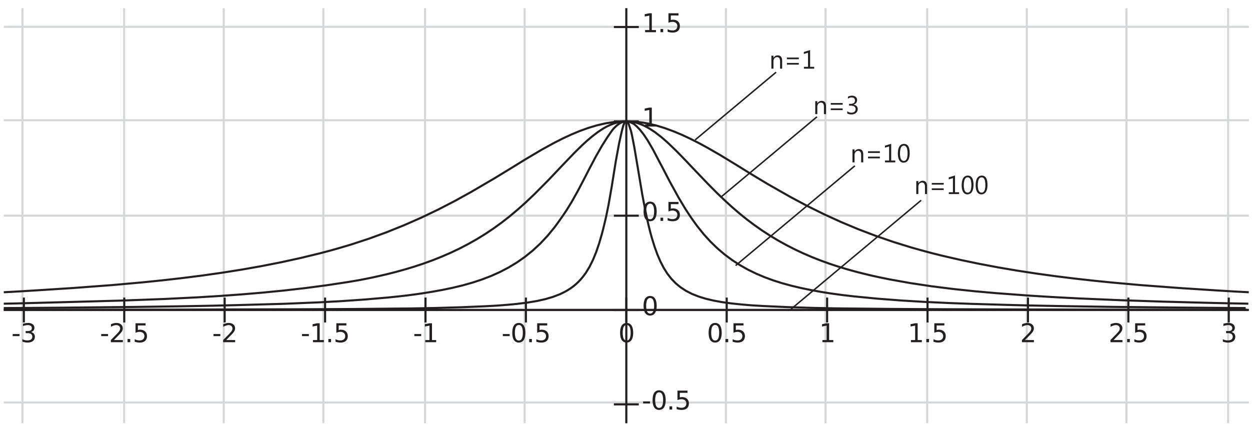
1.2 Sequence of functions

It is a sequence whose terms are real functions defined on a certain $I \in \mathbb{R}$.

$$\{f_n\} = f_1, f_2, \ldots f_n \ldots$$

Example. $f_n(x) = \dfrac{1}{1 + nx^2} \Longrightarrow \{f_n\} = \dfrac{1}{1 + x^2}, \dfrac{1}{1 + 2x^2}, \dfrac{1}{1 + 3x^2} \cdots \dfrac{1}{1 + nx^2} \cdots$

The curves corresponding to different values of n are shown in the figure. Notice that all of them pass through the point $(0, 1)$; and, for all non-zero x, they get closer to OX as n increases.



1.3 Pointwise convergence

Let $\{f_n\}$ be a sequence of functions. If we choose an $x \in I$, the resulting sequence $\{f_n(x)\}$ is a numerical sequence. If it converges, the limit will be a function of the x. Based on what was studied in numerical sequences, we say that the sequence has a limit $f(x)$ if it is satisfied:

$$\boxed{\lim_{n\to\infty} f_n(x) = f(x) \Longleftrightarrow \forall\varepsilon > 0\ \exists n \in \mathbb{N} \,\big/\, |f_m(x) - f(x)| < \varepsilon,\ \forall m \geq n}$$

that is, if the terms of the sequence get as close to the value $f(x)$ as we want, starting from a sufficiently advanced index.

If this occurs $\forall x \in I$, we say that the sequence of functions $\{f_n\}$ converges pointwise to the function f on I.

In the example above, all the curves pass through the point $(0, 1)$, while, outside the origin, they get as close as we want to the OX axis, if n is high enough. So we conclude that

$$f_n(x) = \frac{1}{1+nx^2} \Longrightarrow f(x) = \begin{cases} 1, & x = 0 \\ 0, & x \neq 0 \end{cases}$$

and we obtain a limit function discontinuous at the origin.

The set $\mathcal{C}$ of points where $\{f_n\}$ converges is called **Set of Convergence**.

Exercise. Study the sequence of functions $\{\sin^n x\}$, $x \in [0, \pi/2]$, and check that its limit function is similar to the one just obtained.

1.4 Uniform convergence

In the pointwise convergence it is required that, for every $\varepsilon > 0$, there exists an index n such that, starting from it, the difference $|f_m(x) - f(x)|$ becomes less than ε. The value of n will then depend on ε, but it can also depend on the chosen point x, that is to say

$$n = n(\varepsilon, x) \ \text{ (pointwise convergence)}$$

This allows the n obtained to be very different for the different points x and, as a consequence, the limit function f can be very different from the f_n (in the example, the f_n are continuous and f is not).

If, given an ε, we can find a maximum value n_0 for the different $n(\varepsilon, x)$, then

$$|f_m(x) - f(x)| < \varepsilon,\ \forall m \geq n_0,\ \forall x \in I$$

and, from a value n_0 of the index, the convergence condition is satisfied for all x. We have become independent of the point since the index now depends only on ε. In this case the convergence is said to be uniform and some properties of the functions f_n are inherited by f.

Definition. We say that $\{f_n\}$ converges uniformly to f on $I \in \mathbb{R}$ if and only if

$$\forall\varepsilon > 0\ \exists n(\varepsilon) \,\big/\, d(f_m, f) < \varepsilon,\ \forall m \geq n$$

Or what is the same

$$\boxed{f_n \text{ converges uniformly to } f \text{ on } I \Longleftrightarrow \lim_{n\to\infty} d(f_n, f) = 0}$$

(remember that $d(f_n, f) = \sup |f_n(x) - f(x)|\,,\ x \in I$).

The uniform convergence implies the pointwise one, since the former verifies the same requirements as the latter, plus some additional condition.

Example. To study the convergence of $f_n = \dfrac{x}{1+nx}$, $x \in [1,2]$, we take three steps. First we calculate the limit $f(x)$ for a given $x \in I$. Next we obtain the distance between $f_n(x)$ and $f(x)$. Finally we calculate the limit of this distance when $n \to \infty$.

1) $\lim\limits_{n\to\infty} f_n(x) = 0$, so f_n converges pointwise to $f(x) = 0$.

2) $d(f_n, f) = \sup \left| \dfrac{x}{1+nx} - 0 \right|_{x\in I} = \sup \left| \dfrac{1}{1/x+n} \right|_{x\in I} = \dfrac{1}{1/2+n}$

3) $\lim\limits_{n\to\infty} d(f_n, f) = \lim\limits_{n\to\infty} \dfrac{1}{1/2+n} = 0$, hence the convergence is uniform.

Exercise. Prove the uniform convergence of the sequences whose general terms are:

$$\text{a) } f_n(x) = \frac{1}{2n-\sqrt{x}},\ x \in [1,2]; \qquad \text{b) } f_n(x) = \frac{\cos x^2}{n^2},\ x \in \mathbb{R}$$

1.5 Sequence of continuous functions

Let f_n be a sequence of functions, defined on I. If $\{f_n\}$ converges uniformly to f on I and the f_n are continuous on I, then f is continuous on I. That is to say ,

$$\boxed{f_n \text{ continuous} + \text{uniform convergence} \Longrightarrow f \text{ continuous}}$$

The proof can be seen in a supplementary document.

Application. Given a sequence $\{f_n\}$, it is usually easy to find its limit function $f(x)$ as well as to find out if the f_n and f are continuous. But it is usually more complicated to analyze if the convergence is uniform.

The theorem solves this problem in many cases since, if a sequence of continuous functions converges to a discontinuous limit function, the convergence will not be uniform. If it were, $f(x)$ would be continuous, which does not happen.

Example. We study the pointwise and uniform convergence of the sequence of functions $\{x^n\}$, on the interval $I = [0,1]$.

- For $x = 1$, we have that $x^n = 1, \forall n$.
- For $x \in [0,1)$, $x^n \to 0$, if $n \to \infty$.

hence

$$\lim_{n\to\infty} f_n(x) = \begin{cases} 0, & x \in [0,1) \\ 1, & x = 1 \end{cases}$$

and the limit function $f(x)$ is discontinuous. Since the functions f_n are continuous $\forall n$, we can assure that the convergence is not uniform.

Exercise. Study the convergence of the sequences of functions:

$$\text{a) } \{\cos^n x\},\ x \in [0,\pi/2]; \qquad \text{b) } \left\{e^{-nx^2}\right\},\ x \in \mathbb{R}$$

2 Series of functions

2.1 Definition

To work with series of functions, we will follow the same steps that we took with the series of real numbers. Given a sequence $\{f_n\}$, we define the partial sum F_n as

$$F_n = \sum_{i=1}^{n} f_i = f_1 + f_2 + \cdots + f_n$$

The **series of functions associated to** $\{f_n\}$ is the sequence of partial sums $\{F_n\}$. We usually denote the series as $\sum f_n$. The addends f_i are the terms of the series and f_n is the general term.

Example. If $f_n = \sin \dfrac{x}{n}$, the series associated to $\{f_n\}$ will be $\sum \sin \dfrac{x}{n}$ and the partial sum is

$$F_n = \sum_{i=1}^{n} \sin \frac{x}{i} = \sin \frac{x}{1} + \sin \frac{x}{2} + \cdots + \sin \frac{x}{n}$$

2.2 Pointwise and uniform convergence

Let $\sum f_n$ be a series of functions. Reasoning as in **1.3**, if we choose an $x \in I$, we obtain a series of real numbers. If it converges, its sum will be a function of the chosen x.

Pointwise convergence. We say that the sum of the series is $F(x)$, if the terms $F_n(x)$ of the sequence of partial sums are as close as we want to the value $F(x)$.
If this holds for every $x \in I$, then the series $\sum f_n$ converges pointwise to the function F on I. The condition is:

$$\boxed{\lim_{n\to\infty} F_n(x) = F(x) \iff \forall \varepsilon > 0 \ \exists n \in \mathbb{N} \,\big/\, |F_m(x) - F(x)| < \varepsilon, \ \forall m \geq n}$$

In this case we write

$$\sum_{n=1}^{\infty} f_n(x) = F(x) \quad \big(F(x) \text{ is the sum function}\big)$$

Example. A very simple case of series of functions is the geometric series of ratio x:

$$x + x^2 + \ldots x^n + \ldots$$

This series converges if $|x| < 1$ and its sum is $S = x/(1-x)$ (see unit III, section 2).

Uniform convergence. As we saw in sequences of functions, the convergence will be uniform if n depens only on the chosen value of ε, $n = n(\varepsilon)$, that is

$$\forall \varepsilon > 0 \ \exists n\,(\varepsilon) \,\big/\, d(F_m, F) < \varepsilon, \ \forall m \geq n$$

Or, what is the same,

$$\boxed{\sum f_n \text{ converges uniformly to } F \text{ on } I \iff \lim_{n\to\infty} d(F_n, F) = 0}$$

being the distance between F_n and F the supremum of the point-to-point distances, when $x \in I$.
In the expression $d(F_n, F)$, F_n is the partial sum of the terms of the sequence, so the study of the uniform convergence of a series is more complicated than that of a sequence. For this reason, the criterion of the majorant, stated below, is very useful.

2.3 Criterion of the majorant (Weierstrass)

Let $\sum f_n$ be defined on I. Let $\sum |f_n(x)|$ be the series of the absolute values of its terms. If it has, as a majorant on I, a numerical series of positive terms, convergent, then $\sum f_n$ is uniformly convergent on I.

Example. The series $\sum \frac{1}{n^2}(\sin nx + \cos nx)$ satisfies $|f_n(x)| \le \frac{2}{n^2}$, $\forall n$, $\forall x \in \mathbb{R}$. Hence $\sum |f_n|$ is a minorant of a numerical convergent series, thus it converges uniformly on $\mathbb{R}$.

Exercise. Check the uniform convergence, in their definition intervals, of the series:

$$\text{a)} \sum \frac{x^n}{3^n},\ x \in [0,2]; \qquad \text{b)} \sum \frac{(1-\tan x)^n}{n^3},\ x \in [0, \pi/4]$$

2.4 Series of continuous functions

If a series $\sum f_n$, whose terms are continuous functions on $I = [a,b]$, converges uniformly on I to its sum function F, it is continuous on I.

Proof. We know (section **1.5**) that if a sequence of continuous functions converges uniformly to its limit function f, it is continuous.

Then, if the f_n are continuous on I, the partial sum $F_n = \sum_1^n f_i$ will also be continuous on I, since it is a sum of continuous functions.

Since F_n converges uniformly on I to its sum function F, it is continuous in I.

2.5 Integration of a series of functions

Our main interest in the integration and derivation of series of functions (**2.5** and **2.6**) is that they will be used later in the particular case of power series. Here we will just state the conditions for performing both operations and the properties that they verify.

If a series $\sum f_n$ of functions integrable on $I = [a,b]$, converges uniformly on I to its sum function F, it is integrable on I and its integral is the sum of the series of integrals.

$$\boxed{\sum_{i=1}^{\infty} f_i(x) \overset{\text{u.c.}}{=} F(x) \Longrightarrow \int_a^x F(t)dt \overset{\text{u.c.}}{=} \sum_{i=1}^{\infty} \int_a^x f_i(t)dt, \quad x \in [a,b]}$$

This can be briefly expressed by saying that *if a series of integrable functions converges uniformly to F, the integral of the sum is the sum of the series of integrals.*

2.6 Derivation of a series of functions

Let $I = [a,b]$. *Given a series $\sum f_n$, of differentiable functions on I, that converges at a point of I, such that $\sum f_n'$ converges uniformly on I, then $\sum f_n$ converges uniformly on I to its sum function F, which is differentiable on I and its derivative is the sum of the series of the derivatives.*

$$\boxed{\sum_{n=1}^{\infty} f_n(x_0) = F(x_0) \ \text{ y } \ \sum_{n=1}^{\infty} f_n'(x) \overset{\text{c.u.}}{=} G(x) \Longrightarrow \sum_{n=1}^{\infty} f_n(x) \overset{\text{c.u.}}{=} F(x) \ \text{ y } \ F'(x) = G(x)}$$

We can express it briefly by saying that, <u>under certain conditions</u>, *the derivative of the sum of a series of differentiable functions is the sum of the series of the derivatives.*

3 Power series

3.1 Definition

A power series is a particular case of a series of functions, in which $f_n(x)$ takes the form

$$f_n(x) = a_n(x-a)^n,\ a_n \in \mathbb{R}$$

In the most common case, with $a = 0$, the series expression is $\sum a_n x^n$.

In power series the term corresponding to $n = 0$ is usually included, so

$$\sum a_n x^n = a_0 + a_1 x + a_2 x^2 + \dots$$

3.2 Cauchy-Hadamard's theorem

For any power series $\exists\, r \,\big/\, 0 \le r \le \infty$ *(radius of convergence) such that:*

- *If* $|x| < r$*, the series is absolutely convergent.*
- *If* $|x| > r$*, the series is not convergent.*

Proof: As we saw in numerical series, absolute convergence is equivalent to unconditional convergence, so we apply the nth root test to $\sum |a_n x^n|$ (which becomes a numerical series for each $x \in \mathbb{R}$).

$$\lim_{n\to\infty} \sqrt[n]{|a_n x^n|} = \lim_{n\to\infty} \sqrt[n]{|a_n|}\,\sqrt[n]{|x|^n} = \lim_{n\to\infty} \sqrt[n]{|a_n|}\,|x| = l\,|x|$$

a) If $l \neq 0$ and $l \neq \infty$, we study two options:

a.1) If $|x| < \dfrac{1}{l} \Longrightarrow l\,|x| < 1 \Longrightarrow \sum |a_n x^n|$ convergent.

Hence the series $\sum a_n x^n$ is absolutely convergent.

a.2) If $|x| > \dfrac{1}{l} \Longrightarrow l\,|x| > 1 \Longrightarrow \lim\limits_{n\to\infty} \sqrt[n]{|a_n x^n|} > 1.$

If the limit is greater than 1, it means that

$$\exists n_0 \;/\; \sqrt[n]{|a_n x^n|} > 1\ \forall n \ge n_0 \Longrightarrow |a_n x^n| > 1\ \forall n \ge n_0$$

so the necessary condition of convergence is not satisfied and the series does not converge.

b) If $l = 0$, then $\forall x,\ l\,|x| = 0 < 1$, so we say that $r = \infty$.

c) If $l = \infty$ then the limit

$$\lim_{n\to\infty} \sqrt[n]{|a_n|}\,|x|$$

will be less than 1 only if $x = 0$, in which case all products $\sqrt[n]{|a_n|}\,|x|$ will be null and the limit will also be null. We then say that $r = 0$

That is, we have found a value r that verifies the condition of the statement, which proves the theorem. If $l = \lim\limits_{n\to\infty} \sqrt[n]{|a_n|}$, this value of r is:

$$\boxed{\textbf{a)}\ r = 0,\ \text{if } l = \infty\,;\quad \textbf{b)}\ r = \infty,\ \text{if } l = 0\,;\quad \textbf{c)}\ r = \frac{1}{l},\ \text{if } l \neq 0,\ l \neq \infty}$$

Remarks. It is important to note the following:

1. For $|x| = r$, the theorem does not say anything, so the series can be convergent or not and **we have to study the numerical series that results for** $x = \pm r$.

2. A sequence $\{\alpha_n\}$ has an **oscillation limit** $\alpha \in \mathbb{R}$ or $\alpha = \infty$, if there exists a subsequence of $\{\alpha_n\}$ which has limit α (or, what is equivalent, if in every neighborhood of α there are infinitely many elements of $\{\alpha_n\}$). This can happen, for example, if α_n does not have a unique expression, but is different for odd and even terms.

   A sequence of real numbers may not have a limit, but it must have some oscillation limit, finite or infinite (J. Burgos, p. 73). If it is also bounded, its oscillation limits will be finite (Bolzano-Weierstrass theorem for sequences). If we obtain different values, we will take for l the largest of them (upper limit of oscillation):

$$l = \overline{\lim_{n\to\infty}} \sqrt[n]{|a_n|}$$

3. We can also determine l as the limit of the quotient

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|$$

   because, if it exists, so does the limit of the nth root and takes the same value (root law).

4. From this theorem, it follows that the interval of convergence $\mathcal{C}$ of power series always takes one of these forms:

$$(-r, r),\ (-r, r],\ [-r, r),\ [-r, r]$$

Examples. We calculate the radius and the interval of convergence of the following series:

1. $\sum \dfrac{x^n}{n^n}$. $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} = \lim\limits_{n\to\infty} \sqrt[n]{\dfrac{1}{n^n}} = \lim\limits_{n\to\infty} \dfrac{1}{n} = 0 \Longrightarrow r = \infty$. Hence $\mathcal{C} = \mathbb{R}$.

2. $\sum n^n x^n$. $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} = \lim\limits_{n\to\infty} \sqrt[n]{n^n} = \lim\limits_{n\to\infty} n = \infty \Longrightarrow r = 0$. Hence $\mathcal{C} = \{0\}$.

3. $\sum \dfrac{x^n}{2^n}$. $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} = \lim\limits_{n\to\infty} \sqrt[n]{\dfrac{1}{2^n}} = \dfrac{1}{2} \Longrightarrow r = 2$. Hence $\mathcal{C} = (-2, 2)$, since:

   - $\sum \dfrac{2^n}{2^n} = \sum 1 \Longrightarrow$ the series diverges for $x = 2$.

   - $\sum \dfrac{(-2)^n}{2^n} = \sum (-1)^n \Longrightarrow$ the series oscillates for $x = -2$.

Exercises. Check that the intervals of convergence are the ones indicated:

1. $\sum n! x^n$, $\mathcal{C} = \{0\}$.

2. $\sum \dfrac{x^n}{n!}$, $\mathcal{C} = \mathbb{R}$.

3. $2x + 2x^2 + 2^3x^3 + 2^3x^4 + \ldots$, $\mathcal{C} = (-1/2, 1/2)$.

4. $1 + x + 25x^2 + 3x^3 + 625x^4 + \cdots + a_n x^n + \ldots$, where $\begin{cases} a_n = n, & n \text{ odd} \\ a_n = 5^n, & n \text{ even} \end{cases}$, $\mathcal{C} = \left(-\dfrac{1}{5}, \dfrac{1}{5} \right)$.

3.3 Continuity, derivation and integration

Let $\sum a_n x^n$ be a power series of radius of convergence $r > 0$ and sum $S(x)$. The function $S(x)$ verifies the following:

a) $S(x)$ is continuous on $(-r, r)$.

b) $S(x)$ is differentiable on $(-r, r)$ being its derivative $\boxed{S'(x) = \sum n a_n x^{n-1}}$

c) $S(x)$ is integrable on $[0, x], \forall x \in (-r, r)$. Its integral is $\boxed{\int_0^x S(t)dt = \sum \frac{a_n}{n+1} x^{n+1}}$

Therefore:

1. **Power series can be differentiated and integrated term by term**, being the derivative of the sum, the sum of the derivatives of the terms; and being the integral of the sum, the sum of the integrals of the terms.

   Note that the terms in **c)** are the integrals of the initial terms since we are integrating the addends $a_n t^n$ between 0 and x. The result coincides with the primitive of $\sum a_n x^n$.

2. As it is shown in the proof, when differentiating or integrating a series we obtain another with the **same radius of convergence**. But, since the interval is open, we cannot ensure if the series converges or not at the endpoints.

Proof: It can be seen in a supplementary document.

Application. The following two geometric series are very useful for applying the above properties to solving problems, as will be seen in the examples:

1. $\displaystyle\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + \cdots + x^n + \ldots$, whose sum is $S(x) = \dfrac{1}{1-x}$, $|x| < 1$.

2. $\displaystyle\sum_{n=0}^{\infty} (-1)^n x^n = 1 - x + x^2 - \cdots + (-1)^n x^n + \ldots$, whose sum is $S(x) = \dfrac{1}{1+x}$, $|x| < 1$.

Example 1. We calculate the sum of the numerical series $\displaystyle\sum_{n=1}^{\infty} \frac{n}{2^n}$.

We start from the series $\displaystyle\sum_{n=0}^{\infty} x^n$, whose sum is $S(x) = \dfrac{1}{1-x}$. To get the factor n in the numerator, we derive:

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots + x^n + \ldots \Longrightarrow \frac{1}{(1-x)^2} = 1 + 2x + 3x^2 + \cdots + nx^{n-1} + \ldots$$

In order to also have n in the exponent, we multiply both members by x, resulting

$$\frac{x}{(1-x)^2} = x + 2x^2 + 3x^3 + \cdots + nx^n + \cdots = \sum_{n=1}^{\infty} nx^n$$

Finally, giving to x the value $1/2$, we arrive to the sum we were seeking

$$\sum_{n=1}^{\infty} \frac{n}{2^n} = \frac{1/2}{(1-1/2)^2} = 2$$

Remark. We have applied the property **3.3.b)**.

Example 2. We now want to calculate the radius of convergence and obtain the sum of the power series

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots$$

We calculate the radius of convergence.

$$l = \lim_{n\to\infty} \sqrt[n]{|a_n|} = \lim_{n\to\infty} \sqrt[n]{\frac{1}{n}} = 1 \Longrightarrow r = \frac{1}{l} = 1$$

Differentiating the series, we obtain

$$1 - x + x^2 - x^3 + \dots$$

which is a geometric series of ratio $-x$. Its sum is:

$$S(x) = \frac{1}{1+x},\ |x| < 1$$

The integral of this function will be the sum of the series from which we started. Thus, we calculate the integral between 0 and x of $S(t)$, that is

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} = \int_0^x \frac{1}{1+t} dt = \ln(1+t)|_0^x = \ln(1+x) - \ln 1 = \ln(1+x),\ |x| < 1$$

We have applied the properties **3.3.b)** and **c)**, deriving the original series, adding the series that results and integrating the sum function obtained.

Exercise. Find the sum of the following series:

$$\textbf{a)} \sum_{n=1}^{\infty} \frac{2n}{3^n} \quad (S = 3/2)\,; \quad \textbf{b)} \sum_{n=1}^{\infty} \frac{n^2}{2^n} \quad (S = 6)$$

3.4 Abel's theorems.

The theorems we state next allow us to study the sum of a series at the endpoints of its interval of convergence, $x = \pm r$ (the proofs can be seen in J. Burgos, p. 548).

Let $\sum a_n x^n$ be a series and let $r > 0$.

1. *If $\sum a_n x^n$ converges at $x = r$ $(x = -r)$, it converges uniformly on $[0, r]$ $\big([-r, 0]\big)$.*
2. *If $\sum a_n x^n$ converges at $x = r$ $(x = -r)$, its sum function is continuous at $x = r$ $(x = -r)$ and integrable on $[0, r]$ $\big([-r, 0]\big)$.*

Application. Abel's second theorem allows us to obtain the value of the sum of a series at the endpoints of $\mathcal{C}$ from the expression on $(-r, r)$, by using the continuity of the sum function at $x = \pm r$. Indeed, if $\sum a_n x^n$ has a radius of convergence r and a sum $S(x)$, then we know that:

- The sum function $S(x)$ is continuous on $(-r, r)$, as seen in **3.3.a)**.
- If $\sum a_n x^n$ converges at $x = r$ $\big(\sum a_n r^n$ is convergent$\big)$, the theorem assures that $S(x)$ is continuous also at $x = r$. That is, we can calculate its value at $x = r$ by taking limits on the expression of $S(x)$ for $x \in (-r, r)$ (and analogously with $x = -r$) .

$$\boxed{\sum a_n r^n = S(r) = \lim_{x\to r} S(x)}$$

Remark. The fact that the sum has an expression as a function of x for $x \in (-r, r)$ and that at $x = r$ the series converges, does not ensure that we can use for the sum at the endpoints of the interval the expression valid for the interior. Indeed, the sum $S(x)$ could be defined differently on $(-r, r)$ and at $x = \pm r$, so that the expression for $(-r, r)$ might not even make sense at $x = r$.

Example 1. We study the uniform convergence of series $\sum \dfrac{x^n}{n}$.

- Applying the Cauchy-Hadamard theorem we see that it has a radius of convergence $r = 1$, therefore it converges on the interval $(-1, 1)$.
- Studying it at $x = \pm 1$, we see that it also converges at $x = -1$ (at $x = 1$ it diverges).
- Then, by Abel's first theorem, we know that it converges uniformly on the interval $[-1, 0]$ and also on every interval $[0, x_0]$, $\forall x_0 < 1$.
- Uniting both intervals, we can affirm that it converges uniformly on every interval.

$$[-1, x_0] \subset [-1, 1)$$

Example 2. We want to obtain the sum of the alternating series of the inverses of the odd numbers, as a particular case ($x = 1$) of the power series

$$\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \cdots \quad (1)$$

To do this, we start from the geometric series of ratio $\lambda = -x^2$

$$\sum_{n=0}^{\infty} (-1)^n x^{2n} = 1 - x^2 + x^4 - \cdots \quad (2)$$

which converges to

$$s(x) = \frac{1}{1+x^2}, \ |x| < 1$$

We see that, integrating series (2), we obtain series (1). So, taking into account **3.3.c)**, the sum $S(x)$ of series (1) will be the integral of the sum $s(x)$ of series (2). That is,

$$S(x) = \int_0^x s(t)dt = \int_0^x \frac{dt}{1+t^2} = \arctan t|_0^x = \arctan x$$

When integrating, the radius of convergence is maintained, so this sum function is valid for $|x| < 1$. But, in principle, we cannot use $x = 1$ in $S(x)$.

We know that the series (1) converges at $x = 1$ (Leibnitz's theorem). Then, applying Abel's second theorem, we can obtain the value of the sum at $x = 1$ by taking limits, when $x \to 1$, in function $S(x) = \arctan x$. Since the arctangent function is continuous, it results

$$\sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \cdots = \lim_{x \to 1} \arctan x = \arctan 1 = \frac{\pi}{4}$$

Exercise. Obtain the sum of the alternating series of the inverses of the even numbers. To do it, calculate the value of the sum funcion of the following power series, at the endpoints of the field of convergence.

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots$$

3.5 Power series expansion of a function. Taylor series

a Power series expansion

Consider the interval $\mathcal{C} = (-r, r)$, the power series $\sum a_n x^n$, convergent on $\mathcal{C}$, and the function f, defined on $\mathcal{C}$. If the sum of the series coincides with the value of the function on $\mathcal{C}$, we say that $\sum a_n x^n$ is a power series expansion of f on $\mathcal{C}$.

Example. We study a power series and a function defined respectively as:

$$\sum_{n=0}^{\infty} x^{2n+1} = x + x^3 + x^5 + \dots \qquad f(x) = \frac{x}{1-x^2}$$

The function exists on all of $\mathbb{R}$, except $x = \pm 1$. The series is geometric with a ratio of x^2, hence it converges if $|x| < 1$. Its sum is

$$S(x) = \frac{x}{1-x^2}$$

Since the sum function $S(x)$ coincides with $f(x)$ on the interval $(-1, 1)$, we say that $\sum x^{2n+1}$ is a power series expansion of $f(x)$ on $\mathcal{C} = (-1, 1)$.

Remark. As we see in the example, the domain of the function does not have to coincide with the interval of convergence of the series. The series will be a series expansion of the function only at the common points of its interval of convergence and the domain of f.

b Taylor series

When studying real functions, we called limited Taylor expansion of order n, of a sufficiently differentiable function, to the expression:

$$f(x) = \sum_{i=0}^{n} \frac{f^{(i}(0)}{i!}\, x^i + T_n(x)$$

that is, the sum of the Taylor polynomial plus the remainder term.

Suppose now that a function f admits a power series expansion, with certain coefficients a_n:

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \quad \text{on } (-r, r)$$

Since it is expressed as a sum of terms of the form $a_n x^n$, the function will be differentiable, its derivative being the sum of the derived series (section **3.3**):

$$f'(x) = \sum_{n=1}^{\infty} n\, a_n x^{n-1} \quad \text{on } (-r, r)$$

Functions $f', f'' \dots$ will also be differentiable. It is easy to prove that (J. Burgos, p. 510):

- The function f belongs to C^∞ on $(-r, r)$ (it is infinitely many times differentiable).
- Its kth derivative at the origin is $f^{(k}(0) = k!\, a_k$, from where $a_k = \dfrac{f^{(k}(0)}{k!}$, $\forall k \in \mathbb{N}$

We see that the coefficients a_n take the Taylor expression (reminded above). This allows us to calculate these coefficients for any function that admits a series expansion. In this case, we will call the **Taylor series expansion** of f to:

$$\boxed{f(x) = \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} \frac{f^{(n}(0)}{n!}\, x^n}$$

Example. We consider the geometric series $\sum x^n$. If, as in this case, we know its sum,

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \quad x \in (-1,1)$$

we can say that the series is the Taylor series expansion of $f(x) = 1/(1-x)$.

Usually we will proceed in the opposite direction, starting from a function and obtaining its expansion. If, in our example, we calculate the higher-order derivatives of f at $x = 0$ and, from them, the coefficients a_n, it turns out

$$f(x) = \frac{1}{1-x} \Longrightarrow f^{(k}(0) = k! \Longrightarrow a_k = \frac{f^{(k}(0)}{k!} = 1$$

so they all take the same value ($a_n = 1, \forall n$). The Taylor series obtained is $1 + x + x^2 + \ldots$, which logically coincides with the one we considered at the beginning. That is,

$$f(x) = \frac{1}{1-x} = \sum_{n=0}^{\infty} \frac{f^{(n}(0)}{n!} x^n = \sum_{n=0}^{\infty} x^n$$

c Necessary and sufficient condition

If $f \in C^{\infty}$, we can calculate its higher-order derivatives and, from them, the coefficients of the Taylor expansion. It may then seem that the condition of belonging to C^{∞} is sufficient for f to admit a Taylor series expansion, but it is not so: it is necessary, but not sufficient. We state the necessary and sufficient condition below.

Let $f \in C^{\infty}$. It is a necessary and sufficient condition for f to admit a Taylor series expansion on $(-r, r)$ that

$$\boxed{\lim_{n\to\infty} T_n(x) = 0, \ \forall x \in (-r,r)}$$

Proof. We write the limited Taylor expansion of the function f, from which we deduce the remainder term.

$$f(x) = \sum_{i=0}^{n} \frac{f^{(i}(0)}{i!} x^i + T_n(x) \Longrightarrow T_n(x) = f(x) - \sum_{i=0}^{n} \frac{f^{(i}(0)}{i!} x^i \qquad (1)$$

The Taylor series will converge to $f(x)$ on $(-r, r)$ if and only if the limit of the Taylor polynomial of degree n (which approximates the value of the function) is equal to $f(x)$, when $n \to \infty$.

$$\lim_{n\to\infty} \sum_{i=0}^{n} \frac{f^{(i}(0)}{i!} x^i = f(x)$$

But this is equivalent to saying that

$$\lim_{n\to\infty} \left(f(x) - \sum_{i=0}^{n} \frac{f^{(i}(0)}{i!} x^i \right) = 0 \overset{(1)}{\Longleftrightarrow} \lim_{n\to\infty} T_n(x) = 0$$

Remark. This result was to be expected, since the remainder term $T_n(x)$ is the difference between the exact value of the function and the approximate value obtained with the Taylor polynomial of degree n. Then, for the series to exactly represent the function, that difference must be made as small as we want by taking a sufficiently high number of terms. That is, $T_n(x)$ must have limit 0, when $n \to \infty$.

4 Self-assesment exercises

4.1 True/False exercise

Decide whether the following statements are true or false.

1. We have defined the distance between two functions as the maximum of their point-to-point distances.

2. If $f_n = \dfrac{x}{1+nx^2} + x$, then $\lim\limits_{n\to\infty} f_n = x$.

3. The sequence of general term $f_n(x) = \sin^n x,\ x \in \left[0, \dfrac{\pi}{2}\right]$ converges uniformly to

$$f(x) = \begin{cases} 1, & x = \pi/2 \\ 0, & x \neq \pi/2 \end{cases}$$

4. Let $\sum f_n$ be a series of functions defined on I. If $\sum |f_n(x)|$ has, as a majorant on I, a convergent numerical series of positive terms, then $\sum f_n$ is uniformly convergent on I.

5. The convergence interval of $\displaystyle\sum_{n=1}^{\infty} nx^n$ is $(-1, 1)$.

6. The convergence interval of $\displaystyle\sum_{n=1}^{\infty} \frac{x^n}{2n}$ is $[-1, 1)$.

7. A power series, the series of its derivatives and the series of its integrals have the same interval of convergence.

8. The sum of the power series $\ -1 - x - x^2 - x^3 \ldots$ is $\dfrac{1}{x-1},\ \ |x| < 1$.

9. Let f be a function defined $\forall x \in \mathbb{R}$. Sea $\sum a_n x^n$ a series convergent on a certain interval $\mathcal{C}$, such that $\sum a_n x^n = f(x)$. We can say that $\sum a_n x^n$ is a power series expansion of $f(x)$ on $\mathbb{R}$.

10. If a function f is infinitely many times differentiable, we say that $f \in C^\infty$. In this case we can obtain its derivatives of any order, so the function admits a Taylor series expansion.

4.2 Question

The power series expansion of function $f(x) = \ln(1+x)$ is

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}\, x^n$$

and its radius of convergence $r = 1$. Obtain, using the second Abel's theorem, the sum of the alternating harmonic series

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots$$

4.3 Solution to the True/False exercise

1. **F**. It has been defined as "the supremum of its point-to-point distances", which always exists, as they are bounded functions (property of the supremum), while the maximum may not exist.

   E.g. If $f(x) = 1/x,\ g(x) = 1$ and $I = [1, \infty)$, the set of point-to-point distances between f and g does not have a maximum. The supremum is 1 and corresponds to $x \longrightarrow \infty$.

2. **T**. For $x \neq 0$ the limit is x, since the denominator of the fraction tends to ∞ and the first addend to 0. For $x = 0$, f_n is null $\forall n$, so that the limit is 0, that is, x.

3. **F**. The limit function is correct, but the convergence is not uniform, because the functions f_n are continuous and f is not (see **1.5**).

4. **T**. See the criterion of the majorant (Weierstrass).

5. **T**. The limit $l = \sqrt[n]{|a_n|} = 1$, hence $r = 1/l = 1$. Studying the endpoints of $\mathcal{C}$ $(x = \pm 1)$, we see that in them the series diverges.

6. **T**. The limit $l = \sqrt[n]{|a_n|} = 1$, hence $r = 1/l = 1$. Studying the endpoints of $\mathcal{C}$ $(x = \pm 1)$, we see that the series diverges at $x = 1$ and converges at $x = -1$.

7. **F**. They have the same **radius** of convergence (see **3.3**).

8. **T**. $-1 - x - x^2 - x^3 \ldots$ is a geometric series of ratio $\lambda = x$. Its sum is

$$S(x) = \frac{-1}{1-x} = \frac{1}{x-1}, \quad |x| < 1$$

9. **F**. $\sum a_n x^n$ is a power series expansion of $f(x)$ on $\mathcal{C}$, where the series converges, not on $\mathbb{R}$.

10. **F**. That f belongs to C^∞ is a necessary condition, but not sufficient, for f to admit a Taylor series expansion. The necessary and sufficient condition is that $\lim\limits_{n\to\infty} T_n(x) = 0$, being $T_n(x)$ the remainder term of the limited Taylor expansion (section 3.5 of the unit).

4.4 Solution to the question

The series of real numbers, whose sum we want to obtain, converges by Leibnitz's theorem.

According to the statement, the sum of the given power series is the function $S(x) = \ln(1 + x)$, for values of $|x| < 1$. At the convergence interval endpoints $(x = \pm 1)$ it may or may not converge and must be studied in each case.

For $x = +1$, the power series becomes the series we want to sum: $1 - \dfrac{1}{2} + \dfrac{1}{3} - \dfrac{1}{4} + \ldots$.

Abel's second theorem states that "if a power series $\sum a_n x^n$ converges for $x = r$, its sum is a continuous function at $x = r$". Since we know the value of $S(x)\, \forall x \in (-1, 1)$, we can obtain the value of the sum of the series at $x = 1$ as the limit of $S(x)$ when $x \to 1$. Thus,

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \lim_{x\to 1} \ln(1 + x) = \ln(1 + 1) = \ln 2$$

as we already knew from unit III (Series of real numbers).

Remark. At the other endpoint of the convergence interval $(x = -1)$, the resulting series is

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} (-1)^n = \sum_{n=1}^{\infty} \frac{(-1)^{2n+1}}{n} = -\sum_{n=1}^{\infty} \frac{1}{n}$$

that is, the harmonic multiplied by -1, therefore divergent to $-\infty$.

Unit V. Complex numbers (26.04.2024)

1 Introduction

The objective of these notes is to give some basic notions about complex numbers (a rigorous axiomatic definition can be seen in J. Burgos, pg. 559).

As we know, there is a bijection between the points of the line and the real numbers, so that each element of one of the two sets corresponds to one, and only one, of the other. This allows us to find a solution in $\mathbb{R}$ for equations that do not have a solution in $\mathbb{Q}$, despite which there are still equations without solutions. One of the simplest is:

$$x^2 + 1 = 0$$

In order to solve it outside the real numbers, we define the number $i = \sqrt{-1}$, which we call **imaginary unit**. Thus, the solution of the equation is

$$x = \pm\sqrt{-1} = \pm i$$

Numbers of the form ki, with $k \in \mathbb{R}$ are called **imaginary** (or pure imaginary). If we consider the equation

$$x^2 + 2x + 2 = 0$$

we arrive at the following solution

$$x = -1 \pm \sqrt{-1} = -1 \pm i$$

Since this result is the sum of a real number and an imaginary number, we call it **complex number**. From the definition of this new type of number, we can solve any algebraic equation without solution in $\mathbb{R}$, such as $e^x = -1$ or $\cos x = 2$.

2 Definition. Rectangular form. Basic operations

A complex number in **rectangular form** is any element of the form $z = a + bi$, where a and b are real numbers and are called **real part** and **imaginary part** of z respectively.

Both real numbers and pure imaginary numbers are complex numbers. The former are complex with the imaginary part null ($b = 0$) while the latter have the real part null ($a = 0$). The set of complex numbers is denoted $\mathbb{C}$.

Given the complex numbers $z_1 = a_1 + b_1 i$ and $z_2 = a_2 + b_2 i$, we define the following operations.

a) Addition: $z_1 + z_2 = (a_1 + a_2) + (b_1 + b_2)i \in \mathbb{C}$.

b) Product: $z_1 \cdot z_2 = a_1a_2 + b_1b_2i^2 + (a_1b_2 + a_2b_1)i = a_1a_2 - b_1b_2 + (a_1b_2 + a_2b_1)i \in \mathbb{C}$.

c) Product by a $\lambda \in \mathbb{R}$: $\lambda \cdot z = \lambda a + \lambda bi \in \mathbb{C}$.

Two **complex** numbers are **equal** if their respective real and imaginary parts are equal:

$$\boxed{z_1 = z_2 \Longleftrightarrow a_1 = a_2,\ b_1 = b_2}$$

It is immediate to check that the **additive identity** (null element) in $\mathbb{C}$ is $0 = 0 + 0i$ and the **multiplicative identity** (unity element) is $1 = 1 + 0i$. From the above it can be shown that $\mathbb{C}(+, \cdot)$ has a field structure (it is also a two dimensional vector space over $\mathbb{R}$).

Just as sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ and $\mathbb{R}$ are each an extension of the previous one, $\mathbb{C}$ is an extension of $\mathbb{R}$. It is enough to observe that $\forall a \in \mathbb{R}$ we can write it as $a = a + 0i \in \mathbb{C}$, so $\mathbb{R} \subset \mathbb{C}$; and that the operations between numbers of the form $a + 0i$ are reduced to the operations between reals.

3 Polar form. Graphic representation

Considering the real (a) and imaginary (b) parts of a z as Cartesian coordinates, we observe that there is a bijection between $\mathbb{C}$ and $\mathbb{R}^2$ (in fact the expression (a, b) is the **Cartesian form** of z). The complex number z can be represented graphically by a point in the XY plane (called **affix**) as well as by a vector that joins the origin O with the affix.

The distance ρ from the origin to the affix (or modulus of the vector) is the **modulus** of z. The angle that the vector forms with the positive direction of the X axis (for z not null), is the **argument** θ of z. Then we have:

$$a = \rho \cos\theta, \quad b = \rho \sin\theta, \quad \rho = \sqrt{a^2 + b^2}, \quad \tan\theta = \frac{b}{a} \quad (a \neq 0)$$

For all $z \neq 0$ there exists a single value $\theta \in (-\pi, \pi]$ (**main argument**) and infinitely many $\theta + 2k\pi,\ k \in \mathbb{Z}$ values (arguments). Then the **polar** (or trigonometric) **form** of z is

$$z = a + bi = \rho\cos\theta + (\rho\sin\theta)\, i = \rho(\cos\theta + i\sin\theta)$$

and the equality condition of two complex numbers becomes

$$\boxed{z_1 = z_2 \Longleftrightarrow \rho_1 = \rho_2 \text{ and } \theta_1 = \theta_2 + 2k\pi,\ k \in \mathbb{Z}}$$

that is, two **complex numbers** are **equal** if and only if they have the same modulus and their arguments differ by an integer multiple of 2π.

Geometric interpretation of the sum of complex numbers. By representing the complex numbers by means of the two components (like vectors in the plane) we observe that their sum also follows the parallelogram rule. Thus, adding z to z_1 is equivalent to translating the affix of z_1 according to z; that is, moving it by a distance ρ, according to the direction given by θ.

Then, adding z to the complex numbers whose affixes form a geometric figure, is equivalent to translating it a distance ρ, along the direction given by θ, without deforming or rotating it.

Geometric interpretation of the product of complex numbers. Given $z_1 = \rho_1(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = \rho_2(\cos\theta_2 + i\sin\theta_2)$, operating we obtain (check it):

$$z_1 \cdot z_2 = \cdots = \rho_1\rho_2\big[\underbrace{\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2}_{\cos(\theta_1+\theta_2)} + i(\underbrace{\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2}_{\sin(\theta_1+\theta_2)})\big]$$

That is, when multiplying two complex numbers, the modulus of the product is the product of their moduli and the argument is the sum of their arguments. So

- By multiplying z_1 by $\rho \in \mathbb{R}$, we multiply its modulus by ρ, without changing the argument.
- By multiplying z_1 by $(\cos\theta + i\sin\theta)$, we add θ radians to its argument (i.e., we rotate it by an angle θ counterclockwise), without varying its modulus.

Therefore, multiplying by $z = \rho(\cos\theta + i\sin\theta)$ the complex numbers whose affixes form a geometric figure is equivalent to multiply its dimensions by ρ and rotate it θ radians counterclockwise.

Exercise. Given the complex numbers $z_1 = 2 - 2i,\ z_2 = 8 - 2i,\ z_3 = 8 + 2i,\ z_4 = 2 + 2i$.

a) If we multiply them by $z = i/2$ ($\rho = 1/2$ and $\theta = \pi/2$), check that the figure formed by their affixes reduces its dimensions by half and rotate 90 degrees counterclockwise.

b) What happens if we multiply $z_2 - z_1$ by i?

c) If the affixes of z_1 and z_2 are two consecutive vertices of a square located in the fourth quadrant, how can we get the other two vertices?

4 Complex conjugate, opposite and reciprocal. Quotient

Let $z = a + bi$ be a complex number. We define the complex conjugate, opposite and reciprocal.

a) The complex **conjugate** of z $(\bar{z})$ is obtained by changing the sign of the imaginary part, that is,

$$\boxed{\bar{z} = a - bi}$$

The property of conjugates is:

$$z \cdot \bar{z} = (a + bi) \cdot (a - bi) = a^2 - b^2 i^2 = a^2 + b^2 = |z|^2$$

b) The complex **opposite** of z $(-z)$ is the complex that, added to z, results in the null complex. It is immediate to see that

$$\boxed{-z = -a - bi}$$

c) The **reciprocal** of $z \neq 0$ (z^{-1}) is the complex that, multiplied by z, results in the unit complex. From the property of the conjugate, we have that

$$z \cdot \bar{z} = |z|^2 \Longrightarrow z \cdot \frac{\bar{z}}{|z|^2} = 1 \Longrightarrow \boxed{z^{-1} = \frac{\bar{z}}{|z|^2}}$$

d) The **quotient** of z_1 and z_2 $(z_2 \neq 0)$ is defined as the product of z_1 by the reciprocal of z_2.

$$\boxed{\frac{z_1}{z_2} = z_1 \cdot z_2^{-1} = z_1 \cdot \frac{\bar{z}_2}{|z_2|^2}}$$

It can be interpreted as the result of multiplying numerator and denominator by the conjugate of the denominator. In rectangular form,

$$\frac{a + bi}{c + di} = \frac{(a + bi)(c - di)}{(c + di)(c - di)} = \frac{(a + bi)(c - di)}{c^2 + d^2}$$

Geometric interpretation. We write z in polar form, $z = \rho(\cos\theta + i\sin\theta)$. Then:

a) $\bar{z} = \rho\big(\cos\theta - i\sin\theta\big) = \rho\big(\cos(-\theta) + i\sin(-\theta)\big)$.

$\bar{z}$ has the same modulus than z. Its argument is the opposite of the argument of z.

b) $z^{-1} = \dfrac{\bar{z}}{|z|^2} = \dfrac{1}{\rho^2} \cdot \rho\big(\cos(-\theta) + i\sin(-\theta)\big) = \dfrac{1}{\rho}\big(\cos(-\theta) + i\sin(-\theta)\big)$.

The modulus of z^{-1} is the inverse of the modulus of z. Its argument is the opposite of the argument of z.

c) $\dfrac{z_1}{z_2} = z_1 \cdot z_2^{-1} = \cdots = \dfrac{\rho_1}{\rho_2}\big(\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)\big)$.

The modulus of the quotient is the quotient ρ_1/ρ_2; the argument is the difference $\theta_1 - \theta_2$.

Exercise. Prove that:

1. Given two complex numbers, the conjugate of the sum is the sum of the conjugates and the conjugate of the product is the product of the conjugates.

2. A complex is pure imaginary if and only if its opposite is equal to its conjugate.

3. The inverse of z is obtained, as in $\mathbb{R}$, by means of the expression $1/z$.

5 Exponential of a complex number. Euler's formula

Given $z = a + bi$, we define the exponential of z as the (unique) complex of modulus e^a and argument b.

$$\boxed{\exp(z) = e^a(\cos b + i \sin b)}$$

The complex exponential is defined this way so that it preserves the properties that the function e^x holds in $\mathbb{R}$. In particular, the complex exponential of a real number, $\exp(a)$, $a \in \mathbb{R}$, will have the same value than the real exponential of a. Indeed, since the exponent is real,

$$a \in \mathbb{R} \Longrightarrow \exp(a) = \exp(a + 0i) = e^a(\cos 0 + i \sin 0) = e^a$$

so that, for example, the exponential of the complex null is equal to 1, as in $\mathbb{R}$.

The complex exponential is usually denoted by e^z in a simplified but inaccurate way, since this expression does not generally have a unique solution. Indeed, it is a particular case of the complex power of a complex number $z_1^{z_2}$ (not studied here).

From the definition, the following relations are also verified (prove them as an exercise):

a) $\boxed{e^{z_1} \cdot e^{z_2} = e^{z_1+z_2}}$, from where $\boxed{(e^z)^n = e^{nz}}$.

b) $\boxed{e^{z_1}/e^{z_2} = e^{z_1-z_2}}$, from where $\boxed{1/e^z = e^{-z}}$.

In the particular case $z = \theta i$, its exponential will be

$$e^z\big|_{z=0+\theta i} = e^0(\cos\theta + i \sin\theta)$$

from which the **Euler's formula** results

$$\boxed{e^{\theta i} = \cos\theta + i \sin\theta}$$

We see that $e^{\theta i}$ represents the complex number of module 1 and argument θ. Euler's formula helps us to remember the expression of the exponential of z, since

$$e^z = e^{a+bi} = e^a\, e^{bi} = e^a(\cos b + i \sin b)$$

We can now write a complex number in **exponential form**

$$\boxed{z = \rho\,(\cos\theta + i \sin\theta) = \rho e^{\theta i}}$$

In exponential form, the complex conjugate and the reciprocal of $\rho e^{\theta i}$ are

$$\boxed{\bar{z} = \rho e^{-\theta i}; \quad z^{-1} = \frac{1}{\rho} e^{-\theta i}}$$

Example. Let us see that $e^z = -1$ has a solution in $\mathbb{C}$. Indeed:

$$e^z = -1 \Longrightarrow e^{a+bi} = e^a\, e^{bi} = e^{\pi i} \Longrightarrow \left\{ \begin{array}{c} e^a = 1 \Longrightarrow a = 0 \\ b = \pi + 2k\pi, \ k \in \mathbb{Z} \end{array} \right\} \Longrightarrow z = (2k+1)\pi i, \ k \in \mathbb{Z}$$

The solutions are the infinitely many pure imaginary numbers of the form $z = (2k+1)\pi i, \ k \in \mathbb{Z}$.

Exercise. Represent in exponential form the complex numbers -1, i and $-i$. Find the values of x that satisfy the equation $e^{xi} + i = 0$.

6 Natural power. De Moivre's formula

By multiplying the complex number z by itself n times, we get the complex z^n. As we know from the properties of the product, its module will be ρ^n and its argument $n\theta$, that is,

$$z^n = [\rho(\cos\theta + i\sin\theta)]^n = \rho^n(\cos n\theta + i\sin n\theta)$$

In the particular case $\rho = 1$, we obtain De Moivre's formula:

$$\boxed{(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta}$$

Exercise. Calculate $\cos 3\theta$ y $\sin 3\theta$ as a function of $\sin\theta$ y $\cos\theta$.
Sol: $\cos 3\theta = \cos^3\theta - 3\sin^2\theta\cos\theta$; $\sin 3\theta = 3\sin\theta\cos^2\theta - \sin^3\theta$.

7 nth root of a complex

We say that $\omega \in \mathbb{C}$ is the nth root of $z \in \mathbb{C}$ if the power n of ω is equal to z.

$$\boxed{\sqrt[n]{z} = \omega \Longleftrightarrow \omega^n = z}$$

Let $z = \rho e^{\theta i}$ be a complex number and $\omega = r\, e^{\varphi i}$ its nth root. Then

$$\omega^n = r^n\, e^{n\varphi i} = \rho e^{\theta i} \Longrightarrow \left\{\begin{array}{l} r^n = \rho \\ n\varphi = \theta + 2k\pi,\ k \in \mathbb{Z} \end{array}\right\} \Longrightarrow \left\{\begin{array}{l} r = \rho^{\frac{1}{n}} \\ \varphi = \frac{\theta+2k\pi}{n},\ k \in \mathbb{Z} \end{array}\right\}$$

that is

$$\boxed{\sqrt[n]{z} = \rho^{\frac{1}{n}}\, e^{\frac{\theta+2k\pi}{n}} = \rho^{\frac{1}{n}}\, e^{\frac{\theta}{n}+k\,\frac{2\pi}{n}},\ k \in \mathbb{Z}}$$

Giving values to k, we obtain different solutions for the argument $\varphi = \frac{\theta}{n} + k\,\frac{2\pi}{n}$:

$$\begin{array}{ll}
k = 0 & \Longrightarrow \varphi_0 = \dfrac{\theta}{n} \\[2ex]
k = 1 & \Longrightarrow \varphi_1 = \dfrac{\theta}{n} + 1\,\dfrac{2\pi}{n} = \varphi_0 + \dfrac{2\pi}{n} \\[2ex]
k = 2 & \Longrightarrow \varphi_2 = \dfrac{\theta}{n} + 2\,\dfrac{2\pi}{n} = \varphi_0 + 2\,\dfrac{2\pi}{n} \\
\vdots & \\
k = n & \Longrightarrow \varphi_n = \dfrac{\theta}{n} + \cancel{n}\,\dfrac{2\pi}{\cancel{n}} = \varphi_0 + 2\pi \\[2ex]
k = n+1 & \Longrightarrow \varphi_{n+1} = \dfrac{\theta}{n} + (n+1)\,\dfrac{2\pi}{n} = \varphi_0 + \cancel{n}\dfrac{2\pi}{\cancel{n}} + \dfrac{2\pi}{n} = \varphi_1 + 2\pi
\end{array}$$

We see that, for $k = n$ and following, the arguments take a previous value increased by 2π, so the resulting complex is the same. Therefore we get only n different complex numbers and every non-zero complex has n nth roots.
All the roots have the same module, so the affixes will be in a circle of radius $\rho^{\frac{1}{n}}$. Since the angular difference between two consecutive roots is $\Delta\varphi = \frac{2\pi}{n}$, they will be equally spaced.
That is, the affixes of the n nth roots of z are the vertices of a regular polygon with n sides, inscribed in a circle with the center at the origin and radius $\rho^{\frac{1}{n}}$.

Exercise. Find the square, cubic, fourth and sixth roots of unity.

8 Fundamental theorem of Algebra

Every non-constant polynomial $P_n(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n,\ a_i \in \mathbb{C}$*, has some complex root.* As a consequence, a polynomial of degree $n \geq 1$ will have n complex roots.

If the coefficients a_i are real, the following holds:

$P_n(z)$ *has* n *complex roots, counting each one as many times as its order of multiplicity indicates, such that, for each complex root* $a + bi$*, its conjugate* $a - bi$ is also a root.

If there are p different real roots x_i, of order of multiplicity α_i and q non-real complex roots z_j, of order β_j, each one with its conjugate $\bar{z}_j$, decomposing into factors we obtain:

$$\boxed{P_n(z) = a_n(z - x_1)^{\alpha_1} \dots (z - x_p)^{\alpha_p}(z - z_1)^{\beta_1}(z - \bar{z}_1)^{\beta_1} \dots (z - z_q)^{\beta_q}(z - \bar{z}_q)^{\beta_q}}$$

being the degree of the polynomial $n = \alpha_1 + \cdots \alpha_p + 2\,(\beta_1 + \cdots \beta_q)$.
As a consequence, if n is odd, there will be at least one real root. If n is even, the number of real roots will be even or zero.

9 Self-assesment exercises

9.1 True/False exercise

Decide whether the following statements are true or false.

1. Let $z = 2e^{\frac{\pi}{2}i}$ be a complex number. If we multiply by z the complex numbers whose affixes form a square centered at the point $C(3, 0)$, the affixes of the resulting complex numbers will form a square of double size, rotated 90 degrees counterclockwise, centered at $C'(0, 6)$.

2. If we add $z = \sqrt{2}e^{\frac{\pi}{4}i}$ to the complex numbers whose affixes form a triangle of vertices the points $(-1, 0)$ $(0, 0)$ and $(0, 1)$, the vertex corresponding to the right angle will move a distance equal to the length of the hypotenuse of the triangle, parallel to it.

3. The complex conjugate and the reciprocal of a complex $z \neq 0$ coincide, as long as its module is less than 1.

4. The equality $e^{3\frac{\pi}{2}i} + i = 0$ holds.

5. The reciprocal of e^z verifies that: $(e^z)^{-1} = \dfrac{1}{e^z} = e^{-z}$.

6. The affixes of the 5th roots of 1 form a regular pentagon, with a vertex on the real axis.

7. Consider the roots of unity of index $2n, n \in \mathbb{N}$. The affixes form a regular polygon of $2n$ vertices, with two vertices on the imaginary axis.

8. The polynomial $P_7(z) = a_0 + a_1 z + \cdots + a_7 z^7$ has seven roots in $\mathbb{C}$. An odd number of them, between 1 and 7, are real.

9.2 Question

Let $z = x + yi$ be a complex number. In which cases the following equalities are verified?

a) $|e^z| = e^{|z|}$

b) $\overline{e^z} = e^{\bar{z}}$

9.3 Solution to the True/False exercise

1. **T**. The new center C' is obtained in the same way as the new vertices, multiplying by z the complex corresponding to the old one ($z_C = 3$). That is, $z_{C'} = 3 \cdot 2e^{\frac{\pi}{2}i} = 6e^{\frac{\pi}{2}i} = 6i$, which corresponds to the point $(0, 6)$ (see section **3**).

2. **T**. The new vertex is obtained by adding z to the complex corresponding to the origin, that is, $Z_{0'} = 0 + \sqrt{2}e^{\frac{\pi}{4}i} = 1 + i$, which corresponds to the point $(1, 1)$ (see section **3**).

3. **F**. They coincide if their modulus is equal to 1. Indeed, in that case, $z^{-1} = \dfrac{\overline{z}}{|z|^2} = \overline{z}$.

4. **T**. Since $e^{\frac{3\pi}{2}i} = \cos\dfrac{3\pi}{2} + i\sin\dfrac{3\pi}{2} = -i$.

5. **T**. The reciprocal of a complex w is represented by w^{-1} and also by $\dfrac{1}{w}$. In addition, it holds that $\dfrac{1}{e^z} = \dfrac{e^0}{e^z} = e^{0-z} = e^{-z}$.

6. **T**. It is true for any root of odd index of unity. There is always one (and only one) on the real axis, corresponding to the root $z = +1$.

7. **F**. They form a regular polygon of $2n$ vertices, with two of them on the real axis, corresponding to the roots $z = \pm 1$.

8. **T**. A polynomial P_k, of degree k, has k roots on $\mathbb{C}$ (Fundamental Theorem of Algebra). On the other hand, if z is a complex root of P_k, the conjugate one $\bar{z}$ is also a root. Hence the number of complex non-real roots of a polynomial is always even (non-real complex roots have a nonzero imaginary part).

   Since non-real complex roots always go two by two, there will be 0, 2, 4, or 6 of them. So, if the degree of the polynomial is odd (in this case 7), the number of real roots will be odd: at least one and at most seven.

9.4 Solution to the question

a) To solve this part, we expand both expressions:

From the definition of e^z, we know that its modulus is e^x. Besides, we can see it by operating.

$$|e^z| = |e^x(\cos y + i\sin y)| = e^x\,|\cos y + i\sin y| = e^x\sqrt{\cos^2 y + \sin^2 y} = e^x$$

On the other hand,

$$e^{|z|} = e^{|x+yi|} = e^{\sqrt{x^2+y^2}}$$

Equating the two results, we get:

$$e^x = e^{\sqrt{x^2+y^2}} \Longleftrightarrow x = \sqrt{x^2+y^2}$$

which only makes sense if $x \geq 0$. In that case, squaring

$$x^2 = x^2 + y^2,\ x \geq 0 \Longleftrightarrow y = 0,\ x \geq 0$$

The equality holds only if the imaginary part of z is zero and the real part is non-negative, that is, for $z \in \mathbb{R}^+ \cup \{0\}$.

b) We can get from one expression to the other $\forall z \in \mathbb{C}$, so they represent the same complex number.

$$\overline{e^z} \overset{(1)}{=} \overline{e^x(\cos y + i\sin y)} \overset{(2)}{=} e^x\left(\overline{\cos y + i\sin y}\right) \overset{(3)}{=} e^x(\cos y - i\sin y) \overset{(4)}{=} e^x\, e^{-yi} \overset{(5)}{=} e^{x-yi} \overset{(6)}{=} e^{\overline{z}}$$

where we have taken the following steps:

(1) By the definition of complex exponential e^z.

(2) Since e^x is a real factor and comes out of the complex conjugate.

(3) By the definition of complex conjugate .

(4) We write the complex conjugate in exponential form.

(5) By the properties of the complex exponential.

(6) By the definition of complex conjugate.

Remark. The above can be written more briefly, using the exponential notation for e^z and the properties of the complex exponential.

$$\overline{e^z} = \overline{e^{x+yi}} = \overline{e^x\, e^{yi}} = e^x\,\overline{e^{yi}} = e^x\, e^{-yi} = e^{x-yi} = e^{\overline{z}}$$