FOREWORD

This book contains the subjects taught, at FILS, along many years, by the authors in the first semester of Calculus. The main topics presented are: series of numbers, series of functions and power series, Taylor formulas and the differential calculus of the functions of several variables (partial derivatives, differential, local extrema). Some theorems are completely proved, some proofs are only sketched and other theorems are only stated, trying to keep a balance between theory and practice.

We tried to be precise, to go to the point as straight as possible. Standard exercises are added. The only prerequisite for the book is a good knowledge of college mathematics but we recall many definitions and facts to make the reading easier.

Any suggestion for improving the text will be welcome.

Acknowledgments. We thank FILS for giving, in the new curriculum, the opportunity for a natural teaching of Calculus.

The authors

CALCULUS I
Theory, worked-out examples, exercises

Paul Flondor, Mircea Olteanu

Contents

Chapter 1

Sequences and Series of numbers

1.1 Real numbers

As a general fact, we shall suppose the basic properties of the (usual) ***operations*** (addition, substraction, multiplication, division) with real numbers well known. We shall also suppose that the ***order relation*** between reals and the idea of representing real numbers as points of a line are familiar, too.
As for notation: $\mathbb{N}$ will be the set of ***natural numbers***, $\mathbb{Z}$ the set of the integers, $\mathbb{Q}$ the set of the ***rationals*** and R the set of the ***reals***. We have the following strict inclusions: $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

We intend, in the following, to introduce some basic concepts concerning the order relation. The notation for the order will be " $\leq$ ": "$x \leq y$" means "x is less than y or equal to y". Strict inequality $x < y$ means $x \leq y$ and $x \neq y$. As it is well known, the inequalities are also denoted by $\geq, >$.

Definitions
Let A be a non empty set of $\mathbb{R}$ ($A \subseteq \mathbb{R}$, $\emptyset \neq A$).
A number $b \in \mathbb{R}$ is an ***upper bound*** of A if $x \leq b$ for every $x \in A$.
A set having at least one upper bound is called ***bounded from above***.
A number $a \in \mathbb{R}$ is called a ***lower bound*** of A if $a \leq x$ for every $x \in A$.
A set having at least one lower bound is called ***bounded from bellow***.

A set which is (both) bounded from above and from bellow is called ***bounded***. So the set A is bounded if there are $a, b \in \mathbb{R}$ such that $a \leq x \leq b$ for every $x \in A$.
We shall consider the empty set $\emptyset$ as bounded by definition.

Exercise
Prove that a set $A \subseteq \mathbb{R}$ is bounded if and only if there exists $M > 0$ such that $|x| \leq M$ for every $x \in A$.

Remark
From now on we shall use the notations:
"iff" for "if and only if", "$\forall$" for "for every", "$\exists$" for "there exists" and "s.t." for "such that".

Exercise
Prove that:
(i) Every finite set is bounded.
(ii) If A, B are bounded, then $A \cup B$ is bounded.
(iii) If $A \subseteq B$ and B is bounded, then A is bounded.

Definition
The number $s \in \mathbb{R}$ is said to be ***the least upper bound*** (l.u.b.) of $A \subseteq \mathbb{R}$ if:
(i) s is an upper bound of A.
(ii) for every upper bound b of A one has $s \leq b$.
We shall use the notation $s = \sup A$.

Exercise
(i) Prove the uniqueness of the l.u.b. of a set in case of existence.
(ii) Find the l.u.b. of the sets $[0, 1)$ and $[0, 1]$.

Definition
The number $m \in \mathbb{R}$ is said to be ***the greatest lower bound*** (g.l.b.) of $A \subseteq \mathbb{R}$ if:
(i) m is a lower bound of A.
(ii) for every lower bound a of A one has $a \leq m$.
We shall use the notation $m = \inf A$.

Exercise
(i) Prove the uniqueness of the g.l.b of a set in case of existence.
(ii) Find the g.l.b of the sets $[0, 1)$ and $[0, 1]$.

Remark
The l.u.b. (g.l.b.) is sometimes called also ***supremum (infimum)***, explaining the notations above.

We can now state a fundamental fact about real numbers which we accept as an ***axiom***.

The L.u.b. Axiom
Every non empty, bounded from above subset of $\mathbb{R}$ has l.u.b.

Remark
This axiom might not seem very intuitive. Loosely speaking the meaning of it could be "the real numbers fill a whole line, there are no "holes", etc."
Maybe the following remark could help: it is easy to see that the definitions above can be given in the set $\mathbb{Q}$ (we only need an order relation). So we can ask if the L.u.b. Axiom holds if our "universe" is the set of rationals, i.e.: is it true that a non empty, bounded from above set of rationals has a ***rational*** l.u.b. ? The answer is ***no***. Think about the set of non negative rationals of square less than 2, (see exercise 4).

Exercise
Every non empty, bounded from bellow subset of $\mathbb{R}$ has g.l.b.

The following property is useful:
Proposition
Let $A \neq \emptyset$ be a set of real numbers. Then:
$s = \sup A$ iff the following assertions are (both) true:
(i) s is an upper bound of A.
(ii) $\forall \alpha \in \mathbb{R}$ s.t. $\alpha < s$, $\exists x_\alpha \in A$ s.t. $\alpha < x_\alpha$.
Proof First, let us suppose $s = \sup A$. We only need to prove the condition (ii). If this would be false then one can find $\alpha \in \mathbb{R}$, $\alpha < s$,

s.t. $x \leq \alpha$, $\forall x \in A$; but this is a contradiction (we have found an upper bound of A strictly less than the least one).
Let now suppose that s satisfies conditions (i) and (ii). We only need to prove condition (ii) in the definition of the l.u.b. Let $b \in \mathbb{R}$ be an upper bound of A. If $b < s$ then using the condition (ii) we can find $x_b \in A$ s.t. $b < x_b$. This contradicts the fact that b is an upper bound of A; so we have to accept that $s \leq b$.

As a first application of the L.u.b. Axiom let us prove the following:

Theorem (Archimedes)

Let $a, b \in \mathbb{R}, a > 0$. Then there exists $n \in \mathbb{N}$ s.t. $na > b$.

Proof Let us suppose, by contradiction, that $na \leq b, \forall n \in \mathbb{N}$ and let $A = \{na \; ; \; n \in \mathbb{N}\}$. Obviously $A \neq \emptyset$ and A is bounded from above (b is an upper bound of A). Let $s = \sup A$ (according to l.u.b. Axiom); then $s - a < s$ (because $(a > 0)$ and by using the above proposition we can find $n \in \mathbb{N}$ s.t. $s - a < na$. Consequently this means $s < (n+1)a$, which is a contradiction since $(n+1)a \in A$.

Corollary

$\mathbb{N}$ is not bounded; take $a = 1$ above. And so is $\mathbb{R}$.

Exercises

1. Using Archimedes' Theorem prove that if $x, a \in \mathbb{R}, a > 0$ then there exists an unique $n \in \mathbb{Z}$ s.t. $na \leq x < (n+1)a$.

Hint By applying Archimedes'Theorem to $x \in \mathbb{R}$ and $a > 0$ one gets $m \in N$ s.t. $x < ma$. Analogously, there exists $p \in \mathbb{N}$ s.t. $-x < pa$, hence $-pa < x < ma$. It results that:

$$x \in [-pa, (-p+1)a) \cup [(-p+1)a, (-p+2)a) \cup ... \cup [(m-1)a, ma)$$

Moreover, since any two intervals (in the above union) are disjoints, x must be in exactly one of them.

2. Prove that if $a, b \in \mathbb{R}$, $a < b$, then there exists $x \in \mathbb{Q}$ s.t. $a < x < b$ (one says that $\mathbb{Q}$ is dense in $\mathbb{R}$).

Hint Let $a < b$; by applying Archimedes' Theorem to $1 \in \mathbb{R}$ and $b - a > 0$ there exists $n \in \mathbb{N}$ s.t. $1 < n(b-a)$, hence $\frac{1}{n} < b - a$.

According to the previous exercise applied to $a \in \mathbb{R}$ and $\frac{1}{n} > 0$ there exists $m \in \mathbb{Z}$ s.t. $\frac{m}{n} \leq a < \frac{m+1}{n}$; since $\frac{m+1}{n} \in \mathbb{Q}$, the proof is over if $\frac{m+1}{n} < b$. This results from the inequalities:

$$\frac{m+1}{n} - a \leq \frac{m+1}{n} - \frac{m}{n} = \frac{1}{n} < b - a.$$

3. Using the fact that $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$, prove that if $a, b \in \mathbb{R},\ a < b$, then there exists $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ s.t. $a < \alpha < b$, i.e. $\mathbb{R} \setminus \mathbb{Q}$ is dense in $\mathbb{R}$. Remember that for two sets A, B, then $A \setminus B = \{x;\ x \in A \text{ and } x \notin B\}$.
Hint Let first observe that if $r, s \in \mathbb{Q}$, $s \neq 0$, then $r + s\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$. Now let $a, b \in \mathbb{R}$ and let apply Archimedes' Theorem to $\sqrt{2} \in \mathbb{R}$ and $b - a > 0$; it results that there exists $m \in \mathbb{N}$ s.t. $\sqrt{2} < m(b - a)$, hence $ma + \sqrt{2} < mb$. Let n be the largest integer s.t. $n \leq mb$, hence $n + \sqrt{2} \leq ma + \sqrt{2} < mb$. Since n is the largest integer s.t. $n \leq mb$, it results that $ma < n + \sqrt{2}$, so $ma < n + \sqrt{2} < mb$; finally $a < \frac{n+\sqrt{2}}{m} < b$ and $\frac{n+\sqrt{2}}{m} \in \mathbb{R} \setminus \mathbb{Q}$.

4. Let $A = \{p \in \mathbb{Q}\ ;\ 0 < p,\ p^2 < 2\}$. Prove that A is bounded in $\mathbb{Q}$ (adapt the definitions of bounds to $\mathbb{Q}$) and that A has no l.u.b. in $\mathbb{Q}$.
Hint An upper bound of A is 2. We first prove that if it would exist a rational $r = \sup A$, then $r^2 = 2$. Indeed, if $r^2 < 2$, then $r \in A$. Let $h \in \mathbb{R}$ s.t. $0 < h < 1$ and $h < \frac{2-r^2}{2r+1}$; then $t = r + h$ has the properties: $r < t$ and $t \in A$, contradiction. If $r^2 > 2$, then the number $q = \frac{1}{r} + \frac{r}{2}$ has the properties: $0 < q < r$ and $q^2 > 2$, contradiction. So $r^2 = 2$. It is well known that there is no rational with this property.

5. Let $a \in \mathbb{R}$ be given. Show that a is the l.u.b. of the set of rationals which are strictly less than a. Same for the irrationals. State and prove the corresponding result for the g.l.b.'s.
Hint Let $A = \{x \in \mathbb{Q}\ ;\ x < a\}$; then $A \neq \emptyset$ (why ?). Obviously a is an upper bound of A; let $m \in \mathbb{R}$ be another upper bound of A. If $m < a$ then there exists $y \in \mathbb{Q}$ s.t $m < y < a$ ($\mathbb{Q}$ is dense), hence $y \in A$, which is a contradiction, etc.

6. If $A, B \subseteq \mathbb{R}$ are two non empty sets, then define
$A + B = \{x + y\ ;\ x \in A, y \in B\}$. Prove that if A and B are bounded

from above, then so is $A+B$ and $\sup(A+B) = \sup A + \sup B$.

7. Let $\emptyset \neq A \subseteq \mathbb{R}$ be bounded from bellow and let $-A = \{-x\ ;\ x \in A\}$; prove that $-A$ is bounded from above and that $\inf(A) = -\sup(-A)$.

8. Suppose A and B are two non empty sets of real numbers and $A \subseteq B$. Prove that:
(i) If B is bounded from above so is A and $\sup A \leq \sup B$.
(ii) If B is bounded from bellow so is A and $\inf A \geq \inf B$.

9. Let $A = \{\frac{1}{n}\ ;\ n \in \mathbb{N}, n \neq 0\}$, $B = \{\frac{2x}{x^2+1}\ ;\ x \in \mathbb{R}\}$. Prove that A and B are bounded and compute $\inf A$, $\sup A$, $\inf B$, $\sup B$.

1.2 Cauchy sequences

Let us recall the notion of a "sequence".
Definition
Let X be a non empty set. A function $f : \mathbb{N} \mapsto X$ is called a ***sequence of elements of*** X or shorter a ***sequence in*** X. The classical notation for a sequence is $(x_n)_n$, where $x_n = f(n)$. Sometimes we need to consider functions defined only for $n \geq k$, where k is a fixed natural; we call them sequences, too. If $(x_n)_n$ is a sequence in X, then x_n is called the ***term*** (of rank n) of the sequence.

Going back to the set $\mathbb{R}$ we can state the following important:
Theorem (of nested intervals)
Let $[a_0, b_0] \supseteq [a_1, b_1] \supseteq ... \supseteq [a_n, b_n] \supseteq ...$ be a sequence of closed, bounded intervals in $\mathbb{R}$. Then $\bigcap\limits_{n \in N} [a_n, b_n] \neq \emptyset$.
Proof Define $A = \{a_0, a_1, ..., a_n, ...\}$ and $B = \{b_0, b_1, ..., b_n, ...\}$; obviously A and B are non empty sets. It is clear that A is bounded from above and B is bounded from bellow (in fact they are both bounded). Let $a = \sup A$, $b = \inf B$. It is easy to check that $a \leq b$ and $[a, b] \subseteq \bigcap\limits_{n \in N} [a_n, b_n]$.

Remark
If the sequence $(b_n - a_n)_n$ converges to 0 (see the following definition) then $\bigcap\limits_n [a_n, b_n]$ has only one element (prove it).

The notion of a convergent sequence of real numbers was studied at school as well as computations of limits. We shall recall the basic definition of the limit, but we shall consider the elementary computations with limits as known.

Definition
Let $(a_n)_n$ be a sequence in R. The number $a \in \mathbb{R}$ is said to be the ***limit*** of the sequence $(a_n)_n$ if:

$$(\star) \qquad \forall \varepsilon > 0, \ \exists N_\varepsilon \ \text{ s.t. } \forall n \geq N_\varepsilon \ \text{ implies } \ |a_n - a| < \varepsilon.$$

In this case we write $\lim\limits_{n\to\infty} a_n = a$ or $a_n \longrightarrow a$. A sequence having a (real) limit is called ***convergent***.

Remark
(i) If $a \in \mathbb{R}$ satisfy the definition $(\star)$ (for a fixed sequence $(a_n)_n$) then it is unique with this property; (that's why we say "the limit", etc).
(ii) The geometric interpretation of the previous definition is: for every interval $(a - \varepsilon, a + \varepsilon)$ it is possible to find (a natural number, or rank) N_ε such that for all $n \geq N_\varepsilon, \ a_n \in (a - \varepsilon, a + \varepsilon)$.
(iii) In $(\star)$ one can replace $n \geq N_\varepsilon$ by $n > N_\varepsilon$ or $|a_n - a| < \varepsilon$ by $|a_n - a| \leq \varepsilon$.

Definition
Define the ***distance*** between two real numbers as $d(x, y) = |x - y|$. The definition $(\star)$ becomes:

$$(\star\star) \qquad \forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N} \text{ s.t. } \forall n \geq N_\varepsilon \ \ d(a_n, a) < \varepsilon.$$

$(\star\star)$ has the advantage of "generality": once one has a "distance", one can define convergent sequences.

Exercise (the basic properties of the distance)
Prove that $\forall x, y, z \in \mathbb{R}$:

(i) $d(x, y) \geq 0,\ d(x, y) = 0 \Leftrightarrow x = y$
(ii) $d(x, y) = d(y, x)$
(iii) $d(x, z) \leq d(x, y) + d(y, z)$.

Let us now see how the L.u.b. axiom can be used to prove a very useful result.
First remember that the sequence $(a_n)_n$ is ***increasing*** if:
$a_n \leq a_{n+1}, \forall n \in \mathbb{N}$ (and ***strictly increasing*** if $a_n < a_{n+1},\ \forall n \in \mathbb{N}$).
Changing " $\leq$ " into " $\geq$ " one obtain ***decreasing*** (***strictly decreasing***) sequences. A sequence is said to be ***monotone*** (***strictly monotone***) if it is either increasing or decreasing (strictly, etc).

Theorem
Every monotone, bounded sequence is convergent.
Proof Remember that a sequence of real numbers $(a_n)_n$ is bounded if there are $a, b \in \mathbb{R}$ s.t. $a \leq a_n \leq b,\ \forall n \in \mathbb{N}$ (or, equivalently: $\exists M > 0$ s.t. $|a_n| \leq M, \forall n \in \mathbb{N}$.) Now suppose that $(a_n)_n$ is an increasing bounded sequence. We write $a = \sup_n a_n$ for $a = \sup A$, where $A = \{a_0, a_1, ...\}$. Let $\varepsilon > 0$ be fixed; then there is $N_\varepsilon \in \mathbb{N}$ s.t. $a - \varepsilon < a_{N_\varepsilon} \leq a$ (see proposition pg. 5). The sequence being increasing we get $a - \varepsilon < a_n \leq a,\ \forall n \geq N_\varepsilon$ and so $|a_n - a| < \varepsilon,\ \forall n \geq N_\varepsilon$.

Example
It is useful to remember that the sequence $a_n = (1 + \frac{1}{n})^n, n \geq 1$ is increasing and bounded. Its limit is Euler's famous number e.

Definition
Let X be a non empty set and let $(x_n)_{n\in\mathbb{N}}$ be a sequence in X. If $n_0 < n_1 < ... < n_k < ...$ is a strictly increasing sequence of natural numbers, then the sequence $(x_{n_k})_k$ is called a ***subsequence*** of the sequence $(x_n)_n$.
For example, if $n_k = 2k$, one obtains the subsequence of "even terms" of $(x_n)_n$, usually denoted by $(x_{2n})_n$; analogously $(x_{2n+1})_n$ is the subsequence of "odd terms".

Theorem (Cesaro)
Let $(a_n)_n$ be a bounded sequence in $\mathbb{R}$. Then there exists at least one

convergent subsequence of $(a_n)_n$.

Proof Let $a, b \in \mathbb{R}$ be such that $a \leq a_n \leq b, \forall n \in \mathbb{N}$. Let us denote $[a, b] = [a_0, b_0]$ and take $a_{n_0} \in [a_0, b_0]$. Divide the interval $[a_0, b_0]$ into two subintervals of equal length. At least one of these two subintervals contains infinitely many terms of the sequence $(a_n)_n$; denote by $[a_1, b_1]$ such an interval (if both subintervals do, then choose, for instance the left one). Now pick $n_1 \in \mathbb{N}, n_1 > n_0$ s.t. $a_{n_1} \in [a_1, b_1]$. In this way, we can construct, by ***induction*** a sequence of nested intervals:

$$[a_0, b_0] \supset [a_1, b_1] \supset ... \supset [a_k, b_k] \supset ...$$

s.t. $b_k - a_k = \frac{b-a}{2^k}$ and a subsequence of $(a_n)_n$, denoted $(a_{n_k})_k$ s.t. $a_{n_k} \in [a_k, b_k], \forall k \in \mathbb{N}$. Using the nested intervals theorem we have $\bigcap\limits_{k \in \mathbb{N}} [a_k, b_k] = \{c\}$. It is easy to check that $a_{n_k} \longrightarrow c$, simply because $|a_{n_k} - c| \leq \frac{b-a}{2^k}$.

In the definition of a convergent sequence the limit of a sequence plays a central role. But if one would need to use this definition for checking the convergence of a (given) sequence it would be necessary first to "guess" the limit and then check the condition $(\star)$. It would be very useful to have a test for convergence involving only the given sequence.

Definition

A sequence $(a_n)_n$ is said to be a ***Cauchy sequence*** if:

$$(\star\star\star) \qquad \forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N} \text{ s.t. } \forall n, m \geq N_\varepsilon \ |a_n - a_m| < \varepsilon.$$

Exercise

(i) The sequence $(a_n)_n$ is a Cauchy sequence iff:

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N} \text{ s.t. } \forall n \geq N_\varepsilon, \forall k \in \mathbb{N} \quad |a_{n+k} - a_n| < \varepsilon.$$

(ii) Prove that Cauchy sequences are bounded (remember that convergent sequences are bounded too).

Proposition

Convergent sequences are Cauchy sequences.

Proof Let $(a_n)_n$ be a convergent sequence, say $a_n \longrightarrow a$. Let $\varepsilon > 0$ be given; then $\exists N_\varepsilon \in \mathbb{N}$ s.t. $\forall n \geq N_\varepsilon$ $|a_n - a| < \frac{\varepsilon}{2}$. Then for every $n, m \geq N_\varepsilon$:

$$|a_n - a_m| = |a_n - a + a - a_m| \leq |a_n - a| + |a - a_m| < \varepsilon.$$

Theorem (Cauchy criterion)
Let $(a_n)_n$ be a Cauchy sequence in $\mathbb{R}$. Then $(a_n)_n$ is a convergent sequence.
Proof By using the second exercise above one obtains that $(a_n)_n$ is bounded. Now by Cesaro's theorem let $a_{n_k} \longrightarrow a$ be a convergent subsequence of $(a_n)_n$. Let $\varepsilon > 0$; then we can find $N'_\varepsilon \in \mathbb{N}$ s.t. $\forall n, m \geq N'_\varepsilon$ $|a_n - a_m| < \frac{\varepsilon}{2}$ and N''_ε s.t. $\forall n_k \geq N''_\varepsilon \quad |a_{n_k} - a| < \frac{\varepsilon}{2}$. Now take $N_\varepsilon = \max\{N'_\varepsilon, N''_\varepsilon\}$; for every $n \geq N_\varepsilon$ and $n_k \geq N_\varepsilon$ we get:

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

We say that $\mathbb{R}$ is a ***complete metric space*** (the meaning being exactly the theorem above).

Exercises

1. Consider the sequence of intervals

$$(0,1] \supset \left(0, \frac{1}{2}\right] \supset ... \supset \left(0, \frac{1}{n}\right] \supset ...$$

Compute $\bigcap_n \left(0, \frac{1}{n}\right]$. Does this contradict the nested intervals theorem?
Hint The intersection is the empty set; the intervals are not closed.

2. Let $(x_n)_n$ be a sequence of real numbers s.t. the subsequences $(x_{2k})_k$, $(x_{2k+1})_k$, $(x_{3k})_k$ are convergent. Prove that $(x_n)_n$ is convergent.
Hint Let $x_{2k} \longrightarrow a$, $x_{2k+1} \longrightarrow b$, $x_{3k} \longrightarrow c$. Consider the subsequence $(x_{6k})_k$; it is a subsequence of both sequences $(x_{2k})_k$ and $(x_{3k})_k$, hence $x_{6k} \longrightarrow a$ and $x_{6k} \longrightarrow c$. It results that $a = c$ (the limit is unique !);

analogously, by using the subsequence $(x_{6k+3})_k$, we get $b = c$, so $a = b$.

3. Compute $\sup x_n$, $\inf x_n$ for the sequence $x_n = n^{(-1)^n}$.
Hint $x_{2k} = 2k$ and $x_{2k+1} = \frac{1}{2k+1}$; if $A = \{x_n \ ; \ n \in \mathbb{N}\}$, it follows that $\inf A = 0$ and A is not bounded from above.

4. Every real number a is the limit of a sequence of rationals strictly less than a and of a sequence of rationals strictly greater than a.
Hint Let $a \in \mathbb{R}$; for every natural number n, there exists a rational $x_n \in (a - \frac{1}{n}, a)$. Obviously $x_n < a \ \ \forall n \in \mathbb{N}$ and $|x_n - a| < \frac{1}{n}$.

5. Let a, b, c distinct real numbers; find a sequence s.t. it contains convergent subsequences to a, b and c.
Hint For example $a, b, c, a, b, c, a....$

6. Find a sequence s.t. for every natural number n there is a subsequence convergent to n.

7. Let $A \neq \emptyset$ be bounded from above (bellow) set of real numbers. Prove that there exists a sequence in A convergent to $\sup A$ ($\inf A$).
Hint Suppose A is bounded from above and let $s = \sup A$.
Then $\forall n \in \mathbb{N}$, $\exists x_n \in A$ s.t. $s - \frac{1}{n} < x_n \leq s$.

8. Let $x > 0$ and let $a_n = \frac{1}{n}[nx], n \geq 1$. Is the sequence $(a_n)_n$ bounded ?
Hint By the definition of the integer part, we have:

$$\frac{1}{n}(nx - 1) < \frac{1}{n}[n\,x] \leq \frac{1}{n}nx.$$

9. Let a and x_0 be two strictly positive real numbers. Let x_n be the sequence defined by:

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

Prove that x_n is convergent and compute its limit.
Hint We shall prove that $(x_n)_n$ is a bounded (from bellow) decreasing sequence; first, obviously $x_n > 0, \forall n \in \mathbb{N}$. The definition of x_n can

be written as: $x_{n-1}^2 - 2x_n x_{n-1} + a = 0$. The equation must have real solutions, hence : $x_n^2 - a \geq 0$. Now we test the monotony:

$$x_n - x_{n-1} = \frac{1}{2}\left(x_{n-1} + \frac{a}{x_{n-1}}\right) - x_{n-1} = \frac{1}{2x_{n-1}}\left(a - x_{n-1}^2\right) \geq 0, \forall n \geq 1$$

It results that $(x_n)_n$ is decreasing and $x_n \geq \sqrt{a}$, so it is convergent, etc.

1.3 Complex numbers

Definition and Notations
We shall denote the set of complex numbers by $\mathbb{C}$. A typical element of $\mathbb{C}$ is written as $z = x + iy,\ x, y \in \mathbb{R}$; x is the ***real part*** of z (denoted by $\Re z$) and y is the ***imaginary part*** of z (denoted by $\Im z$). So $z = \Re z + i\Im z$. Two complex numbers z_1 and z_2 are ***equal*** iff $\Re z_1 = \Re z_2$ and $\Im z_1 = \Im z_2$. We suppose the algebraic operations (and their properties, $i^2 = -1$) with complex numbers well known. We consider $\mathbb{R} \subset \mathbb{C}$ by identifying the real number x with the complex number $x + i0$. The algebraic properties of $\mathbb{C}$ are synthetically expressed by saying that $\mathbb{C}$ is a (commutative) field containing $\mathbb{R}$ as a subfield. We ***do not*** consider any ***order relation*** on $\mathbb{C}$.

The complex numbers are represented as points in the (cartesian) plane by associating the number $z = x + iy$ to the point of coordinates (x, y).
Remember that the ***modulus*** of $z = x + iy$ is $|z| = \sqrt{x^2 + y^2}$. The geometric interpretation of $|z|$ is the (euclidean) distance from z to the origin. We define the ***distance*** between two complex numbers z_1 and z_2 to be $d(z_1, z_2) = |z_1 - z_2|$. In other words $d(z_1, z_2)$ is the ***length*** of the line segment joining (the points) z_1, z_2 (direct computation). This length is the same as in elementary analytic geometry.

The basic properties of the modulus are ($\forall z, z_1, z_2 \in \mathbb{C}$):
(i) $|z| \geq 0$, $|z| = 0 \Longleftrightarrow z = 0$.
(ii) $|z_1\, z_2| = |z_1|\, |z_2|$.
(iii) $|z_1 + z_2| \leq |z_1| + |z_2|$.

The properties of the distance are consequences of the properties of the modulus ($\forall z_1, z_2, z_3 \in \mathbb{C}$):
(i) $d(z_1, z_2) \geq 0,\ d(z_1, z_2) = 0 \iff z_1 = z_2$.
(ii) $d(z_1, z_2) = d(z_2, z_1)$
(iii) $d(z_1, z_3) \leq d(z_1, z_2) + d(z_2, z_3)$.
The property (iii) is called "triangle inequality" for obvious reasons.

By using the distance (modulus) we can elegantly describe the following sets:
the ***open disk*** centered at $z_0 \in \mathbb{C}$ and of radius $r > 0$:
$D(z_0, r) = \{z \in \mathbb{C}\ ;\ |z - z_0| < r\}$;
the ***closed disk*** centered at $z_0 \in \mathbb{C}$ and of radius $r > 0$:
$D'(z_0, r) = \{z \in \mathbb{C}\ ;\ |z - z_0| \leq r\}$;
the ***circle*** centered at $z_0 \in \mathbb{C}$ and of radius $r > 0$:
$\mathcal{C}(z_0, r) = \{z \in \mathbb{C}\ ;\ |z - z_0| = r\}$.
Obviously, $D'(z_0, r) = D(z_0, r) \bigcup \mathcal{C}(z_0, r)$. We shall term ***unit disk*** (***circle***) the case $z_0 = 0$ and $r = 1$.

Exercise
Prove that if $z_1, z_2 \in \mathbb{C},\ z_1 \neq z_2$ then there are $r_1 > 0, r_2 > 0$ s.t. $D(z_1, r_1) \bigcap D(z_2, r_2) = \emptyset$.

Remark
It would be useful to keep in mind that by "proof" we mean to use the definitions and properties of complex numbers, not "geometrical intuitions" (which make the previous exercise more or less trivial). Generally speaking, geometric intuition (or the use of a picture) is useful as a starting point but it will be never considered as a proof.

Definition
(i) A subset $A \subseteq \mathbb{C}$ is called ***bounded*** if either $A = \emptyset$ or :
$\exists M > 0$ s.t. $|z| \leq M,\ \forall z \in A$.
(ii) A sequence $(z_n)_n$ in $\mathbb{C}$ is called ***bounded*** if:
$\exists M > 0$ s.t. $|z_n| \leq M,\ \forall n \in \mathbb{N}$.

Exercise
What is the geometric interpretation of boundness? (in terms of disks).

Definition
The sequence $(z_n)_n$ in $\mathbb{C}$ has the limit $z \in \mathbb{C}$ if:

$$\forall \varepsilon \ \exists N_\varepsilon \in \mathbb{N} \text{ s.t. } \forall n \geq N_\varepsilon \ \ |z_n - z| < \varepsilon.$$

We write $\lim_{n\to\infty} z_n = z$, or $z_n \longrightarrow z$.
The sequences having limit are called **convergent**.
The geometric interpretation of the previous definition is: $z_n \longrightarrow z$ iff for every open disk $D(z, \varepsilon)$, the terms z_n with $n \geq N_\varepsilon$ belong to the disk ($z_n \in D(z, \varepsilon)$, $\forall n \geq N_\varepsilon$).

Remark
It is useful to keep in mind that $z_n \longrightarrow z$ in $\mathbb{C}$ iff $|z_n - z| \longrightarrow 0$ in $\mathbb{R}$.

Exercise (the uniqueness of the limit)
Prove that if $z_n \longrightarrow z$ and $z_n \longrightarrow w$, then $z = w$. For the proof one can use the exercise after the definition of disks (pg. 16).

Exercise
Prove that convergent sequences in $\mathbb{C}$ are bounded.

Exercise
If $x_n \longrightarrow x$ in $\mathbb{R}$, then $x_n \longrightarrow x$ in $\mathbb{C}$.

Example
Let $z \in \mathbb{C}$ and consider the sequence of the powers of z: $(z^n)_n$. We want to find the complex numbers z for which $(z^n)_n$ is convergent. Consider the following cases:
(i) $|z| < 1$; in this case it is obvious that $z^n \longrightarrow 0$, because:
$|z^n| = |z|^n \longrightarrow 0$ (in $\mathbb{R}$).
(ii) $|z| > 1$; in this case $(z^n)_n$ is not bounded (again one can reduce the problem to the real case), so not convergent.
(iii) $|z| = 1$ (z is on the unit circle); we shall use that $z^{n+1} = z^n z$ (this relation could be an inductive definition of powers of z). Let us suppose that $z^n \longrightarrow \ell$; it is clear that $z^{n+1} \longrightarrow \ell$ and $z^n z \longrightarrow \ell z$. So we obtain $\ell = \ell z$. Consequently, we get $\ell = 0$ or $z = 1$. Finally we obtain only

that $z = 1$, simply because ℓ cannot be 0. After all, the only complex number on the unit circle having a convergent sequence of powers is $z = 1$.

We shall need the following inequalities (the proof is trivial and will be left to the reader):

$$(\star) \quad \max\{\,|\,\Re z\,|\,,\,|\,\Im z\,|\,\} \le |z| \le |\,\Re z\,| + |\,\Im z\,|,\ \forall z \in \mathbb{C}.$$

Let now $(z_n)_n$ be a sequence in $\mathbb{C}$; there are two sequences $(x_n)_n$ and $(y_n)_n$ of real numbers such that $z_n = x_n + iy_n$. Viceversa, two sequences of real numbers define a sequence of complex numbers (more precisely the first sequence will be the sequence of the real parts and the second the imaginary parts). We want to investigate the connections between the nature (convergence or not) of these sequences.

Proposition
Let $(z_n)_n$ be a sequence in $\mathbb{C}$, $z_n = x_n + iy_n$ and let $z = x + iy \in \mathbb{C}$; then:

$$z_n \longrightarrow z \ \text{ (in } \mathbb{C}) \ \text{ iff } \ x_n \longrightarrow x \ \text{ and } \ y_n \longrightarrow y \ \text{ (in } \mathbb{R}).$$

Proof Use $(\star)$ to obtain:

$$\max\{\,|x_n - x\,|\,,\ |y_n - y|\,\} \le |z_n - z| \le |x_n - x| + |y_n - y|,\ \forall n \in \mathbb{N}$$

and then apply the definition of convergence.

Remark
In conclusion the convergence of a sequence in $\mathbb{C}$ can be reduced to the convergence of two sequences in $\mathbb{R}$. One can use this to compute limits of convergent sequences in $\mathbb{C}$.

Example
Test the convergence of $z_n = \dfrac{1}{n+i}$ and compute the limit (if any).

Solution $z_n = \dfrac{n}{n^2+1} - i\dfrac{1}{n^2+1}$, so $\Re z_n = \dfrac{n}{n^2+1} \to 0$ and

$\Im z_n = -\dfrac{1}{n^2+1} \to 0$. Another method: $|z_n| = \dfrac{1}{\sqrt{n^2+1}} \longrightarrow 0$.

Definition
A sequence $(z_n)_n$ in $\mathbb{C}$ is called a ***Cauchy sequence*** if:

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N} \text{ s.t. } \forall n, m \geq N_\varepsilon \;\; |z_n - z_m| < \varepsilon.$$

Proposition
Let $(z_n)_n$ be a sequence in $\mathbb{C}$, $z_n = x_n + iy_n$. Then $(z_n)_n$ is a Cauchy sequence in $\mathbb{C}$ iff $(x_n)_n$ and $(y_n)_n$ are Cauchy sequences in $\mathbb{R}$.
The proof is similar to the proof of the previous proposition and will be left to the reader.

Theorem
Let $(z_n)_n$ be a sequence in $\mathbb{C}$; then:

$$(z_n)_n \text{ is convergent iff } (z_n)_n \text{ is a Cauchy sequence.}$$

Proof It obviously follows from the corresponding result in $\mathbb{R}$ and the above two propositions.
So we can conclude (as we did for $\mathbb{R}$) that $\mathbb{C}$ is a ***complete metric space***.

The following proposition is easy.

Proposition
(i) If $z_n \longrightarrow z$ and $w_n \longrightarrow w$ then $z_n + w_n \longrightarrow z + w$.
(ii) If $z_n \longrightarrow z$ and $w_n \longrightarrow w$ then $z_n w_n \longrightarrow zw$.
(iii) If $z_n \longrightarrow z$ and $w_n \longrightarrow w, w \neq 0$ then $\frac{z_n}{w_n} \longrightarrow \frac{z}{w}$. (the sequence $\frac{z_n}{w_n}$ is defined from a rank N s.t. $w_n \neq 0, \forall n \geq N$).

Exercises

1. Prove the last proposition.

2. Draw graphic representations for the following sets of complex numbers:
(i) $\{z \in \mathbb{C} \; ; \; |z - i| \leq 1\}$.

(ii) $\{z \in \mathbb{C}\,;\ 1 \le |z| < 2\}$.
(iii) $\{z \in \mathbb{C}\,;\ 1 \le \Re z \le 2\}$
(iv) $\{z \in \mathbb{C}\,;\ |z| < 1$ and $|\Im z| < \frac{1}{2}\}$.
(v) $\{z \in \mathbb{C}\,;\ |z - z_1| = |z - z_2|\}$, z_1, z_2 are fixed complex numbers.

3. If $(x_n)_n$ is a sequence in $\mathbb{R}$ and $x_n \longrightarrow z$ in $\mathbb{C}$ then $z \in \mathbb{R}$.

4. If $z_n \longrightarrow z$ then $|z_n| \longrightarrow |z|$.
Hint It can be proved directly or by using the inequality (prove it):

$$|\,|z| - |w|\,| \le |z - w|,\ \forall z, w \in \mathbb{C}$$

5. Is the converse of the previous exercise true? What if $z = 0$?
Hint Obviously, it is false. However, the assertion:
$|z_n| \longrightarrow 0 \implies z_n \longrightarrow 0$ is true.

6. Is Cesaro's theorem true in $\mathbb{C}$? (remember that Cesaro's theorem in $\mathbb{R}$ is about convergent subsequences of bounded sequences).
Hint The answer is yes: if z_n is a bounded sequence in $\mathbb{C}$, then $\Re z_n$ and $\Im z_n$ are bounded sequences in $\mathbb{R}$; apply now Cesaro's theorem (in $\mathbb{R}$) for them. Be careful in choosing the common indices of the subsequences!

7. If $z = x + iy$ then, by definition, its ***conjugate*** is $\overline{z} = x - iy$. Prove that $z_n \longrightarrow z$ iff $\overline{z_n} \longrightarrow \overline{z}$.

8. Suppose that $z_n \longrightarrow z$; let $w \in \mathbb{C}$ and $r > 0$.
(i) Prove that if $z_n \in D'(w, r), \forall n \in \mathbb{N}$ then $z \in D'(w, r)$.
(ii) What if $z_n \in D(w, r), \forall n \in \mathbb{N}$? Is it true that $z \in D(w, r)$?
(iii) Prove that if $z_n \in \mathcal{C}(w, r), \forall n \in \mathbb{N}$ then $z \in \mathcal{C}(w, r)$.
Hint (i) We can suppose $w = 0$. If $|z_n| \le r$ and $z_n \longrightarrow z$, then:

$$|z| = \lim_n |z_n| \le r.$$

(ii) The answer is no, in general; take for example $z_n = 1 - \frac{1}{n}$. Then $z_n \in D(0, 1), \forall n \in \mathbb{N}$, $z_n \longrightarrow 1$, but $1 \notin D(0, 1)$.

9. Study the convergence of the sequences:
(i) $z_n = \frac{2^n}{n!} + i\frac{n^2}{2^n}$
(ii) $w_n = n^{-1} + in$
(iii) $u_n = \frac{1+in}{1-in}$
(iv) $v_n = 1 + i + i^2 + ... + i^n$
Hint $z_n \longrightarrow 0$ and $u_n \longrightarrow -1$; w_n and v_n are not convergent.

10. If $z_n \longrightarrow z$ then $\frac{1}{n}(z_1 + z_2 + ... + z_n) \longrightarrow z$.
Hint Use the corresponding property for real sequences.

11. By using the sequence $(z^n)_n$, study if the sequences $(\sin n)_n$ and $(\cos n)_n$ can be both convergent. But at least one of them?
Hint Let $z = \cos 1 + \sin 1$; then $z^n = \cos n + i \sin n$, and z^n is divergent (because $|z| = 1$ and $z \neq 1$).
Now the second part. Suppose that $\sin n \longrightarrow \ell$; then $\cos^2 n \longrightarrow 1 - \ell^2$, so $\cos 2n \longrightarrow 1 - 2\ell^2$. It results that $z^{2n} \longrightarrow (1 - 2\ell^2) + i\ell^2$, contradiction because the sequence $z^{2n} = (z^2)^n$ is not convergent.

1.4 Series of numbers

Definition
Let u_n be a sequence in $\mathbb{C}$; form a new sequence $(S_n)_n$ defined by:

$$S_n = u_0 + u_1 + ... + u_n.$$

The ***pair of sequences*** $((u_n)_n, (S_n)_n)$ is called a ***series*** and is denoted $\sum_n u_n$. The terms of the sequence $(u_n)_n$ are ***the terms of the series*** and the terms of the sequence $(S_n)_n$ the ***partial sums of the series***.
Convergence The series $\sum_n u_n$ is said to be ***convergent*** if the sequence of partial sums converges. In this case, say, $S_n \longrightarrow S$, then (the complex number) S is called the ***sum*** of the (convergent) series and we write $\sum_n u_n = S$. The notation $\sum_n u_n$ is used for both the series and

for its sum (if convergent); it will be clear from the context what is the case.
Not convergent series are called ***divergent***. By the ***nature*** of a series we mean convergence (or divergence).

Remark
(i) It is sometimes useful to consider $\sum_{n\geq 1} u_n$ with corresponding partial sums $S_n = u_1 + u_2 + ...u_n$; the results can be easily adapted to this case.
(ii) If $(u_n)_n$ is a sequence in $\mathbb{R}$ then $(S_n)_n$ is a sequence of real numbers and if $\sum_n u_n$ converges then the sum S is a real number.
(iii) We shall use also the notation $u_0 + u_1 + ... + u_n + ...$ for $\sum_n u_n$.

Theorem
Let $\sum_n u_n$ be a convergent series; then $\lim_n u_n = 0$.
Proof Obviously $u_n = S_n - S_{n-1}, \forall n \geq 1$, etc.

Remark It is important to notice that the condition $u_n \longrightarrow 0$ is a ***necessary*** condition for convergence; it is not ***sufficient*** for the convergence, as the following example shows.

Example
Let $u_n = \sqrt{n+1} - \sqrt{n}$; obviously $u_n = \frac{1}{\sqrt{n+1}+\sqrt{n}} \longrightarrow 0$ but $s_n = \sqrt{n+1}$ is not convergent.

An important series is given by:
Example (the geometric series)
Let $z \in \mathbb{C}$ and consider the series

$$1 + z + z^2 + ... = \sum_{n\geq 0} z^n$$

If $z = 1$, then $s_n = n + 1$, hence the series is divergent.
If $z \neq 1$, then $s_n = \dfrac{1 - z^{n+1}}{1 - z}$ and by using the result about the convergence of the sequence $(z^n)_n$ (see the previous section) we get that the ***geometric series converges iff*** $|z| < 1$; in this case $\sum_{n\geq 0} z^n = \dfrac{1}{1-z}$.

Proposition

(i) If $\sum_n u_n$, $\sum_n v_n$ are two series s.t. $\exists n_0 \in \mathbb{N}$ s.t. $u_n = v_n$, $\forall n \geq n_0$, then they have the same nature (both convergent or both divergent).
(ii) Suppose $\sum_n u_n = S$, $\sum_n v_n = T$; then $\sum_n (u_n + v_n) = S + T$.
(iii) Suppose $\sum_n u_n = S$ and let $\alpha \in \mathbb{C}$; then $\sum_n \alpha u_n = \alpha S$.
Proof Easy.

Remark

Assertion (i) of the above proposition may be used if a condition about a series holds from an index on, more precisely, if we are interested (only) about its nature, then we can suppose that the condition holds for all the terms.

Theorem (Cauchy criterion)

The series $\sum_n u_n$ converges iff:

$$\forall \varepsilon > 0, \exists N_\varepsilon, \forall n \geq N_\varepsilon, \forall p \in \mathbb{N}^\star \ \ |u_{n+1} + u_{n+2} + ... + u_{n+p}| < \varepsilon.$$

Proof trivial application of the Cauchy criterion for sequences to $(S_n)_n$.

Example

Let consider the ***harmonic series***: $1 + \frac{1}{2} + \frac{1}{3} + ... = \sum_{n \geq 1} \frac{1}{n}$.
Direct computation gives:

$$S_{2n} - S_n = \frac{1}{n+1} + \frac{1}{n+2} + ... + \frac{1}{2n} \geq \frac{1}{2}.$$

The harmonic series is **divergent** because $(S_n)_n$ is not a Cauchy sequence. We remark that the sequence of partial sums is (in this case) strictly increasing so the divergence of the harmonic series is equivalent to the fact that $(S_n)_n$ is not bounded (from above).

Definition

A series $\sum_n u_n$ is said to be a ***series with positive terms*** (s.p.t.) if $u_n \geq 0$, $\forall n \in \mathbb{N}$. When needed, we shall suppose that $u_n > 0$ (for example, when considering $\frac{u_{n+1}}{u_n}$, etc). The series with positive terms are easier to handle because of the following:

Proposition

If $\sum_n u_n$ is a s.p.t. then its convergence is equivalent to the boundness (from above) of the sequence $(S_n)_n$.

Proof The sequence $(S_n)_n$ is increasing, etc.

The nature of series is decided by using ***tests of convergence***. In the next we give some usual tests.

The Comparison Test

Let $\sum_n u_n$, $\sum_n v_n$ be two s.p.t. Suppose that $u_n \leq v_n, \forall n \geq n_0$. Then:

(i) If $\sum_n v_n$ converges so does $\sum_n u_n$.

(ii) If $\sum_n u_n$ diverges so does $\sum_n v_n$.

Proof Let $(S_n)_n$ and $(T_n)_n$ be the sequences of partial sums of the series $\sum_n u_n$, $\sum_n v_n$. We can suppose $u_n \leq v_n$, $\forall n \in \mathbb{N}$. (why ?). Then $S_n \leq T_n$, $\forall n \in \mathbb{N}$ and the result follows by using the previous proposition.

Example

The series $\sum_{n\geq 2} \dfrac{1}{2^n \ln n}$ is convergent because $\dfrac{1}{2^n \ln n} \leq \dfrac{1}{2^n}$, $n \geq 3$.

The Limit Comparison Test

Let $\sum_n u_n$, $\sum_n v_n$ be two s.p.t. Suppose that $\lim_{n\to\infty} \dfrac{u_n}{v_n} = k \in \mathbb{R}$. Then:

(i) If $\sum_n v_n$ converges so does $\sum_n u_n$.

(ii) If $k \neq 0$ and $\sum_n u_n$ converges so does $\sum_n v_n$; (for $k \neq 0$ the series have the same nature).

Proof (i) One obtains $\dfrac{u_n}{v_n} \leq k+1, \forall n \geq n_0$ and so $u_n \leq (k+1)v_n$, $\forall n \geq n_0$. Now apply the (previous) comparison test.

(ii) We get that $\lim_{n\to\infty} \dfrac{u_n}{v_n} = \dfrac{1}{k}$; now reason like in (i).

Example

Consider the series $\sum_{n\geq 1} \frac{1}{n^2}$ and $\sum_{n\geq 1} \frac{1}{n(n+1)}$; as $\lim_{n\to\infty} \frac{n(n+1)}{n^2} = 1$, the series have the same nature. But the second series is convergent because: $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$, etc. So the first series is convergent too.

The Root Test

Let $\sum_n u_n$ be a s.p.t.

(i) If $\exists k \in (0,1)$ s.t. $\sqrt[n]{u_n} \leq k$, $\forall n \geq n_0$, then $\sum_n u_n$ is convergent.

(ii) If $\sqrt[n]{u_n} \geq 1$ for an infinity of indices, then $\sum_n u_n$ is divergent.

Proof (i) If $\sqrt[n]{u_n} \leq k$ then $u_n \leq k^n$; now compare with the geometric series of ratio k.

(ii) Obviously $u_n \not\to 0$.

The Root Test (Limit form)

Let $\sum_n u_n$ be a s.p.t. and suppose that $\lim_{n\to\infty} \sqrt[n]{u_n} = \ell$.

(i) If $\ell < 1$ the series $\sum_n u_n$ converges.

(ii) If $\ell > 1$ then $\sum_n u_n$ diverges.

(iii) If $\ell = 1$ then no conclusion can be drawn.

Proof (i) (sketch) Take $\ell < k < 1$ and remark that $\sqrt[n]{u_n} \leq k$, $\forall n \geq n_0$.

(ii) We leave it as an exercise; (iii) the meaning of "no conclusion" is that there are examples of convergent series and of divergent series satisfying (iii); take for example $u_n = \frac{1}{n}$ and $v_n = \frac{1}{n^2}$.

The Ratio Test (Limit form)

Let $\sum_n u_n$ be a s.p.t. and suppose that $\lim_n \frac{u_{n+1}}{u_n} = \ell$; then:

(i) if $\ell < 1$ the series $\sum_n u_n$ converges;

(ii) if $\ell > 1$ the series $\sum_n u_n$ diverges;

(iii) if $\ell = 1$ no conclusion can be drawn. We omit the proof (see [4], [8]).

Exercise

(i) Prove that the series $\sum_{n\geq 0} \frac{1}{n!}$ is convergent.

(ii) Prove that $\sum_{n\geq 0} \frac{1}{n!} = \lim_n \left(1+\frac{1}{n}\right)^n = e.$

Proof (i) We have that $\lim_{n\to\infty} \frac{u_{n+1}}{u_n} = 0$, so the series is convergent.

(ii) Let S be the sum of the series; we have:

$$\left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!\,n^k} = \sum_{k=0}^{n} \frac{1}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)...\left(1-\frac{k-1}{n}\right).$$

It results that $\left(1+\frac{1}{n}\right)^n < \sum_{k=0}^{n} \frac{1}{k!}$, hence $S \geq e$. For the other inequality, let $p \in \mathbb{N}$, $n \geq p$; then:

$$\left(1+\frac{1}{n}\right)^n \geq \sum_{k=0}^{p} \frac{1}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)...\left(1-\frac{k-1}{n}\right) >$$

$$> \left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)...\left(1-\frac{p-1}{n}\right)\left(\sum_{k=0}^{p}\frac{1}{k!}\right).$$

It results that

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \geq \left(\sum_{k=0}^{p}\frac{1}{k!}\right), \forall p \in \mathbb{N}, \text{ so } S \leq e.$$

The Integral Test

Suppose that $f : [1,\infty) \mapsto \mathbb{R}$ is a ***decreasing*** and ***positive*** function. Then the series $\sum_n f(n)$ is convergent iff the sequence $\left(\int_1^n f(x)dx\right)_n$ is convergent.

Proof The function f being decreasing it results that all (Riemann) integrals $\int_1^n f(x)dx$ exist. Moreover, the sequence $\left(\int_1^n f(x)dx\right)_n$ is positive and increasing. By integrating the inequalities:

$$f(k+1) \leq \ f(x) \leq f(k), \ \forall x \in [k, k+1],$$

we get:

$$f(k+1) \leq \int_k^{k+1} f(x)dx \leq f(k),\ \forall k \in \mathbb{N},\ k \geq 1.$$

By summing all these from $k = 1$ to $k = n$, it results:

$$f(2) + f(3) + ... + f(n+1) \leq \int_1^{n+1} f(x)dx \leq f(1) + f(2) + ... + f(n).$$

If $(s_n)_n$ is the sequence of partial sums of the series $\sum\limits_n f(n)$ one obtains:

$$s_{n+1} - f(1) \leq \int_1^{n+1} f(x)dx \leq s_n,\ \forall n \in \mathbb{N}, n \geq 1.$$

Now $\sum\limits_n f(n)$ being a s.p.t. and $\left(\int_1^n f(x)dx\right)_n$ being increasing, one easily obtains the result.

Example (The Riemann Series)
Let $\alpha \in \mathbb{R}$; the real ***Riemann series*** $\sum\limits_{n \geq 1} \frac{1}{n^\alpha}$ converges iff $\alpha > 1$.
Proof Obviously, if $\alpha \leq 0$ the series is divergent. For $\alpha > 0$, apply the integral test to the function $f(x) = \frac{1}{x^\alpha}$.

In exercises the reader will find other useful tests for the convergence of s.p.t.

We come back to general series (not necessarily with positive terms).

Abel's Theorem
Let $\sum\limits_n \alpha_n u_n$ be a series of complex numbers s.t:
(i) $\alpha_n \searrow 0$ (the sequence $(\alpha_n)_n$ is a sequence of positive numbers decreasing to zero);
(ii) the sequence $\sigma_n = u_0 + u_1 + ... + u_n$ is bounded.
Then $\sum\limits_n \alpha_n u_n$ is convergent.
For the proof, see [1], [4].

Definition
A Leibniz series (alternating series) is a series of real numbers of the form

$$\alpha_0 - \alpha_1 + \alpha_2 - \alpha_3 + ... + (-1)^n \alpha_n + ...,$$

where $(\alpha_n)_n$ is a sequence s.t. $\alpha_n \searrow 0$.

Theorem
Leibniz series are convergent.
Proof Apply Abel's theorem by taking $u_n = (-1)^n$.

Example
The ***alternating harmonic series*** $\sum_{n \geq 1} \frac{(-1)^{n+1}}{n}$ is convergent.

Definition
The series $\sum_n u_n$ is said to be ***absolutely convergent*** if the series $\sum_n |u_n|$ is convergent.

Proposition
Absolutely convergent series are convergent.
Proof Apply Cauchy criterion (for series) and use the inequality:

$$|u_{n+1} + u_{n+2} + ... + u_{n+p}| \leq |u_{n+1}| + |u_{n+2}| + ... + |u_{n+p}|$$

Remark
(i) Generally, convergent series are not absolutely convergent; an example is the harmonic alternating series.
(ii) Absolutely convergent series are important because they are, in some sense, commutatively convergent; more precisely:
if $\sum_n u_n$ is absolutely convergent with the sum S then for every bijective map $\tau : \mathbb{N} \mapsto \mathbb{N}$ (permutation) the series $\sum_n u_{\tau(n)}$ is absolutely convergent with sum S.
This no more true for convergent but non absolutely convergent series. Again, the alternating harmonic series $\sum_{n \geq 1} \frac{(-1)^{n+1}}{n}$ is an example.

Indeed, let $S > 0$ (why ?) be the sum:

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ... = S$$

By multiplying the above equality by $\frac{1}{2}$, it results:

$$\frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + ... = \frac{1}{2}S$$

Now sum the above equalities and associate the terms as follows:

$$1 + \left(-\frac{1}{2} + \frac{1}{2}\right) + \frac{1}{3} + \left(-\frac{1}{4} - \frac{1}{4}\right) + \frac{1}{5} + \left(-\frac{1}{6} + \frac{1}{6}\right) + \frac{1}{7} + ... = \frac{3}{2}S.$$

After computing the parentheses one gets:

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - ... = \frac{3}{2}S,$$

which is a permutation of the initial series.

Our final result of this section concerns the approximation of the sum of a Leibniz series; the idea is to replace it with a partial sum and compute the (absolute) error.

Proposition
Let $\sum_{n\geq 0}(-1)^n\alpha_n$ be a Leibniz series and let S be its sum. Then

$$\left|\sum_{k=0}^{n}(-1)^k\alpha_k - S\right| \leq \alpha_{n+1}$$

Proof Remember that $\alpha_n \searrow 0$. Let:

$$S_0 = \alpha_0, S_2 = \alpha_0 - \alpha_1 + \alpha_2 = \alpha_0 - (\alpha_1 - \alpha_2) \leq S_0, \text{ etc}$$

It is easy to observe that the subsequence $S_{2k} \searrow S$; analogously, $S_{2k+} \nearrow S$. Consequently, one gets $S_{2n+} \leq S \leq S_{2n}$, so:
$0 \leq S_{2n} - S \leq S_{2n} - S_{2n+1} = \alpha_{2n+1}$, etc.

Exercises

1. Let $\sum_n u_n$ be a series in $\mathbb{C}$ and $u_n = x_n + iy_n$.
Prove that $\sum_n u_n$ converges iff $\sum_n x_n$ and $\sum_n y_n$ converge.
Is the series $\sum_n \frac{1}{n+i}$ convergent ?
Hint Apply the corresponding result for sequences to the partial sums.
$\frac{1}{n+i} = \frac{n}{n^2+1} - i\frac{1}{n^2+1}$; the series is divergent because the series of the real parts is divergent.

2. Study the nature of the series:

$$\sum_n \frac{2\sqrt{n^4+1}}{n^3-10n}$$

$$\sum_n \frac{\sqrt{n+1}-\sqrt{n}}{\sqrt{n+2}},$$

$$\sum_n \frac{\sqrt[3]{n+1}-\sqrt[3]{n}}{n^p},\ p \in \mathbb{R}$$

$$\sum_n \frac{1}{\sqrt{n}} \ln\left(1+\frac{1}{\sqrt{n^3+1}}\right)$$

$$\sum_n \sin\frac{n}{n^2+3}$$

$$\sum_n \frac{\ln(n+2)}{\sqrt{n^3+1}}.$$

Hint We apply the limit comparison test.
$\lim_n n^\alpha \frac{2\sqrt{n^4+1}}{n^3-10n} = 2$ if $\alpha = 1$, so the series has the same nature as the Riemann series with $\alpha = 1$ (divergent).
$\lim_n n^\alpha \frac{\sqrt{n+1}-\sqrt{n}}{\sqrt{n+2}} = 1$ if $\alpha = \frac{3}{2}$, so the series converges.
$\lim_n n^\alpha \frac{1}{\sqrt{n}} \ln\left(1+\frac{1}{\sqrt{n^3+1}}\right) = 1$ if $\alpha = 2$, so the series converges.

$\lim\limits_n n^\alpha \dfrac{\ln(n+2)}{\sqrt{n^3+1}} = 0$, $\forall \alpha \in \left(1, \frac{3}{2}\right)$, so the series converges.

3. Study the nature of the series:
$\sum\limits_n \dfrac{n!}{n^n}$, $\sum\limits_n \dfrac{n}{4^n}$, $\sum\limits_n \left(\dfrac{2n+1}{3n+1}\right)^n$, $\sum\limits_n \dfrac{(n!)^2}{(2n)!}$.

Hint Apply ratio test or root test:

$$\lim_n \frac{(n+1)!}{(n+1)^{n+1}} \frac{n^n}{n!} = \left(\frac{n}{n+1}\right)^n = \frac{1}{e},$$

so the series is convergent, etc.

4. Find an example of a series s.t. the ratio test (the limit form) doesn't decide but the root test (limit form) does.
Hint Consider the series defined by the sequence:

$$x_n = \begin{cases} \frac{1}{n} a^n & \text{if } n \text{ is even} \\ n a^n & \text{if } n \text{ is odd} \end{cases}, \ a > 0$$

By applying the root test, one gets $\lim\limits_n \sqrt[n]{x_n} = a$, so if $a < 1$ the series is convergent, if $a > 1$ the series is divergent; if $a = 1$ the series is divergent ($x_n \not\to 0$).

5. Study the nature of series defined by the sequences:

$$x_n = \frac{1 \cdot 3 \cdot 5 \cdot ...(2n-1)}{2 \cdot 4 \cdot 6 ... 2n}, \quad y_n = \frac{1 \cdot 3 \cdot 5 \cdot ...(2n-1)}{2 \cdot 4 \cdot 6 ... 2n} \frac{1}{2n+1},$$
$$u_n = n! \left(\frac{a}{n}\right)^n, a > 0, \quad v_n = \frac{n!}{(a+1)(a+2)...(a+n+1)}, a > -1.$$

Hint We shall use **Raabe test**:
Let $\sum\limits_n u_n$ be a s.p.t. and let $\ell = \lim\limits_n n\left(\dfrac{u_n}{u_{n+1}} - 1\right)$.
(i) If $\ell > 1$, the series is convergent.
(ii) If $\ell < 1$, the series is divergent.
(iii) If $\ell = 1$, no conclusion.
Apply Raabe test for x_n, y_n and v_n.
For u_n, by applying ratio test one gets: if $a < e$ the series is convergent

and if $a > e$ the series diverges. For $a = e$ ratio test does'nt decide so we apply Raabe test:

$$\lim_n n\left(\frac{x_n}{x_{n+1}} - 1\right) = \lim_n n\left(\left(\frac{n+1}{n}\right)^n \frac{1}{e} - 1\right) =$$

$$= n\left(\left(1 + \frac{1}{n}\right)^n \frac{1}{e} - 1\right) = \frac{1}{e}\lim_n \frac{\left(1+\frac{1}{n}\right)^n - e}{\frac{1}{n}}.$$

For the last limit apply L'Hopital rule:

$$\lim_{x\to 0} \frac{(1+x)^{\frac{1}{x}} - e}{x} = \lim_{x\to 0} \frac{(1+x)^{\frac{1}{x}-1}[x - (1+x)\ln(1+x)]}{x^2} = -\frac{e}{2},$$

so the series is divergent.

6. Study the nature of the series defined by the sequences:

$$x_n = \left(1 - \frac{3\ln n}{2n}\right)^n, \quad y_n = \frac{\ln n \cdot \ln(1+\frac{1}{n})}{n}, \quad z_n = (\ln n)^{-\ln(\ln n)}.$$

Hint We shall use the ***logarithmic test***:
Let $\sum\limits_n u_n$ be a s.p.t. and let $\ell = \lim\limits_n \dfrac{-\ln u_n}{\ln n}$.
(i) If $\ell > 1$, the series is convergent.
(ii) If $\ell < 1$ the series is divergent.
(iii) If $\ell = 1$, no conclusion.

7. If $\sum\limits_n u_n$ is a convergent s.p.t then $\sum\limits_n u_n^2$ is convergent.
Hint $u_n \longrightarrow 0$ hence $u_n^2 < u_n, \forall n \geq n_0$.

8. Study the nature of the series $\sum\limits_{n\geq 1} \dfrac{\sin nx}{n}$, $x \in \mathbb{R}, x \neq 2k\pi, k \in \mathbb{Z}$.
Hint Apply Abel test: $\alpha_n = \frac{1}{n} \searrow 0$ and $u_n = \frac{\sin nx}{n}$; then prove that:

$$\sin x + \sin 2x + ... + \sin nx = \frac{\sin\frac{(n+1)x}{2}\sin\frac{nx}{2}}{\sin\frac{x}{2}}, \forall n \in N^\star.$$

The series is not absolutely convergent; if it would be, then the series $\sum_n \frac{1-\cos(2nx)}{2n}$ should be convergent (comparison test).

9. Study the convergence and the absolute convergence of the series defined by the sequences:

$$u_n = \frac{(-1)^n n + 2}{n^2},\ v_n = \sin\left(\frac{n^2+n+1}{n+1}\,\pi\right).$$

Hint The first is convergent (sum of two convergent series) but not absolutely convergent. For the second:

$$v_n = \sin\left(\frac{n^2+n+1}{n+1}\pi\right) = (-1)^n \sin\frac{\pi}{n+1}.$$

It results that the series is convergent, but it is not absolutely convergent (comparison limit test).

10. Study the nature of series defined by the sequences:

$$x_n = \frac{(n!)^2}{(2n)!}$$

$$x_n = (2n+1)\left(\frac{a(a-1)...(a-n+1)}{(a+1)(a+2)...(a+n+1)}\right)^2, a \notin \mathbb{Z}$$

$$x_n = \frac{\sqrt[3]{n+1}-\sqrt[3]{n}}{n^a},\ a \in \mathbb{R}$$

$$x_n = \frac{1}{n\ln^2 n}$$

$$x_n = \frac{z^n}{n}, z \in C$$

$$x_n = \frac{a+(-1)^n\sqrt{n}}{n},\ a \in \mathbb{R}.$$

Chapter 2

Sequences and Series of functions, Elementary functions

2.1 Sequences of functions

For the definitions we shall consider real-valued functions defined on a (nonempty) set X. Most results have obvious extensions to complex-valued functions; this fact will be just mentioned, without any details. For the applications, the basic case considers X an interval of real numbers.

Let $(f_n)_n$ be a sequence of functions $f_n : X \mapsto \mathbb{R}, \forall n \in \mathbb{N}$.
For every fixed $x \in X$ we get a sequence $(f_n(x))_n$ in $\mathbb{R}$ (by evaluating the the functions of the sequence at x).

Definition (pointwise convergence)
The sequence $(f_n)_n$ ***converges pointwise*** (or it is ***pointwise convergent***) $f : X \mapsto \mathbb{R}$ if:

$$\forall x \in X \quad \lim_{n\to\infty} f_n(x) = f(x).$$

If this is true, then the function f is said to be the ***pointwise limit*** of the sequence $(f_n)_n$. We shall use the notation $f_n \xrightarrow{p} f$.

Remark
(i) The same definition works for complex-valued functions.
(ii) Loosely speaking, checking the pointwise convergence of a sequence $(f_n)_n$ means to fix $x \in X$ and to compute the limit (if any) $\lim_{n\to\infty} f_n(x)$; if the limits exist for every $x \in X$ then one defines the (pointwise limit) function $f : X \mapsto \mathbb{R}$ as $f(x) = \lim_{n\to\infty} f(x)$.
(iii) The pointwise limit of a sequence (if any) is ***unique***.

Example
Let $X = [0, 1]$ and $f_n(x) = x^n$; by applying the above "algorithm" one obtains that $f_n \xrightarrow{p} f$ with $f(x) = \begin{cases} 0 & \text{if} \quad 0 \le x < 1 \\ 1 & \text{if} \quad x = 1 \end{cases}$

The following proposition is obvious (a restatement of the definition).

Proposition
The following assertions are equivalent:
(i) $f_n \xrightarrow{p} f$
(ii) $\forall x \in X, \forall \varepsilon > 0, \exists N_{\varepsilon,x} \in \mathbb{N}$ s.t. $\forall n \ge N_{\varepsilon,x}$ $|f_n(x) - f(x)| \le \varepsilon$.

Remark
It is crucial to understand that $N_{\varepsilon,x}$ depends both on x and on ε; in other terms, for a given $\varepsilon > 0$ it will be, generally, not possible to find a N_ε "good" for all $x \in X$.

Let us now give the definition of uniform convergence.

Definition (uniform convergence)
The sequence of functions $(f_n)_n$ is said to ***converge uniformly*** to the function f if:

$$(\star) \qquad \forall \varepsilon > 0,\ \exists N_\varepsilon \in \mathbb{N} \text{ s.t. } \forall n \ge N_\varepsilon\ \ |f_n(x) - f(x)| \le \varepsilon,\ \ \forall x \in X.$$

We shall denote this by $f_n \xrightarrow{u} f$.

Remark
(i) The new fact is that N_ε depends only on ε (it is the same for all

$x \in X$). So the condition for uniform convergence is ***stronger*** than that of pointwise convergence: if $f_n \xrightarrow{u} f$ then $f_n \xrightarrow{p} f$.
(ii) Same definition for complex-valued functions.
(iii) The uniform limit of a sequence is ***unique***.

Example
Consider $X = [0, 1)$, $f_n(x) = x^n$. It is clear that $f_n \xrightarrow{p} 0$ (the constant function 0). Let us show that $f_n \not\xrightarrow{u} 0$ (the sequence is not uniformly convergent). Let us suppose, by contradiction, that $f_n \xrightarrow{u} 0$. Then taking $\varepsilon = \frac{1}{2}$ we could find $N \in \mathbb{N}$ s.t. $\forall n \geq N \;\; x^n \leq \frac{1}{2}$, $\forall x \in [0, 1)$, so in particular $x^N \leq \frac{1}{2}$, $\forall x \in [0, 1)$. But $\lim\limits_{x \to 1} x^N = 1$, contradiction.

Exercise
Prove that $f_n \xrightarrow{p} 0 \iff f_n - f \xrightarrow{u} 0$; remember that for two functions $f, g : X \mapsto \mathbb{R}$, then $f - g : X \mapsto \mathbb{R}$ is the function defined by $(f - g)(x) = f(x) - g(x)$, $\forall x \in X$.

Theorem (Cauchy criterion for uniform convergence)
A sequence of functions $(f_n)_n$ is uniformly convergent iff

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N} \quad \text{s.t. } \forall n, m \geq N_\varepsilon \;\; |f_n(x) - f_m(x)| \leq \varepsilon, \forall x \in X.$$

Proof It will be left as an exercise (not completely trivial).

In order to obtain some useful tests for uniform convergence let us remind that $\overline{\mathbb{R}} = \mathbb{R} \bigcup \{-\infty, +\infty\}$ (the extended real line), where $-\infty, +\infty$ are not real numbers; the order in $\mathbb{R}$ is extended to $\overline{\mathbb{R}}$ by defining $-\infty < x < +\infty$, $\forall x \in \mathbb{R}$, $-\infty < \infty$. We suppose that the reader is familiar with sequences ***having limit*** $-\infty$ (or $+\infty$). We keep the notion "convergent" for sequences (in $\mathbb{R}$) with real limit. In $\overline{\mathbb{R}}$ every non empty set has l.u.b. (g.l.b.); for example if $A \subseteq \mathbb{R}$ is not bounded from above then $\sup A = +\infty$, etc. As usual, we shall also write "∞" instead of "$+\infty$".

Now let us go back to a sequence of functions $(f_n)_n$ and $f : X \mapsto \mathbb{R}$. Put $m_n = \sup\{|f_n(x) - f(x)| \;\; ; \;\; x \in X\} \in \overline{\mathbb{R}}$ (so m_n always exists, possibly ∞).

Proposition
$f_n \xrightarrow{u} f \iff \lim_n m_n = 0.$
Proof It is enough to see that $m_n \leq \varepsilon \iff |f_n(x) - f(x)| \leq \varepsilon, \ \forall x \in X$ and apply the definitions.

The above result is useful in applications if (of course) one manages to compute m_n.

Example
Check if $f_n : [0,1] \mapsto \mathbb{R}, \ f_n(x) = x^n - x^{n+1}$ is a uniform convergent sequence.
The first step Let us check the pointwise convergence of the sequence. It is easy to observe that $f_n \xrightarrow{p} 0$.
The second step We have the alternative: $f_n \xrightarrow{u} 0$ or $(f_n)_n$ is not uniformly convergent (why ?). We shall compute m_n; the functions f_n are all continuous on the closed and bounded interval $[0,1]$, so (by the Weierstrass theorem - known from the college) they all have a maximum value which is m_n. As $f_n(0) = f_n(1) = 0$, the maximum value is taken somewhere in the open interval $(0,1)$.
The derivative $f_n'(x) = nx^{n-1} - (n+1)x^n$ is zero for $x_n = \frac{n}{n+1}$. It follows that $m_n = f_n(x_n)$ (Fermat's theorem). But:

$$\lim_n m_n = \lim_n \left(\frac{n}{n+1}\right)^n \left(1 - \frac{n}{n+1}\right) = 0.$$

Consequently, $f_n \xrightarrow{u} 0$.

Proposition
If there exists a sequence $(x_n)_n \in X$ s.t. $f_n(x_n) \not\to 0$ then $f_n \overset{u}{\not\to} 0$.
Proof Obviously, $m_n \geq |f_n(x_n)| \not\to 0$, so $m_n \not\to 0$.

Example
Let us prove that the sequence $f_n(x) = \frac{2nx}{1+n^2x^2}, \ x \in (0,1]$ does not converge uniformly.
First, $f_n \xrightarrow{p} 0$. If $x_n = \frac{1}{n} \in (0,1]$, then $f_n(x_n) = 1, \forall n \in \mathbb{N}$, so $f_n(x_n) \not\to 0$ and we can apply the previous proposition.

We now recall the notion of a ***continuous function***.

Definition
Let $I \subseteq \mathbb{R}$ be a non empty interval, $f : I \mapsto \mathbb{R}$ be a function and $a \in I$; we say that f is ***continuous at*** a if:

$$\forall \varepsilon > 0, \exists \delta_\varepsilon > 0 \text{ s.t. if } x \in I,\ |x-a| < \delta_\varepsilon \text{ then } |f(x) - f(a)| < \varepsilon.$$

If f is continuous at every point of I then f is said to be ***continuous on*** I.

Remark
An equivalent condition of continuity at $a \in I$ is:

for every sequence $(x_n)_n$ in I, s.t. $x_n \longrightarrow a$ one has $f(x_n) \longrightarrow f(a)$.

As the first example of this section shows, the pointwise limit of a sequence of continuous function is not necessarily continuous. The notion of uniform convergence is proved to be useful in "preserving continuity":

Theorem (transfer of continuity)
Let $(f_n)_n$ be a sequence of functions $f_n : I \mapsto \mathbb{R}$ s.t.:
(i) $f_n \xrightarrow{u} f$.
(ii) the functions f_n are continuous at $a \in I$.
Then f is continuous at a.
Proof Let $\varepsilon > 0$ be given. Then:

$$f_n \xrightarrow{u} f \quad \Rightarrow \quad \exists N_\varepsilon \text{ s.t. if } n \geq N_\varepsilon\ |f_n(x) - f(x)| \leq \frac{\varepsilon}{3}, \forall \in I.$$

f_{N_ε} is continuous at a, hence :

$$\exists \delta_\varepsilon > 0 \text{ s.t. if } x \in I, |x-a| < \delta_\varepsilon \text{ then } |f_{N_\varepsilon}(x) - f_{N_\varepsilon}(a)| < \frac{\varepsilon}{3}.$$

Now if $|x - a| < \delta_\varepsilon$:

$$|f(x)-f(a)| \leq |f(x)-f_{N_\varepsilon}(x)|+|f_{N_\varepsilon}(x)-f_{N_\varepsilon}(a)|+|f_{N_\varepsilon}(a)-f(a)| < \varepsilon.$$

Remark
(i) If $f_n \xrightarrow{u} f$ and all the functions f_n are continuous on I then f is continuous on I.
(ii) The uniform convergence is not a necessary condition for the continuity of the (pointwise) limit; take $f_n(x) = x^n$, $x \in [0,1)$.

Definition
If $(f_n)_n$ is a sequence of functions $f_n : [a,b] \mapsto \mathbb{R}$ such that f_n has the limit f (pointwise or uniform) then we say that it can be ***term by term integrated*** if $\int\limits_a^b f_n \longrightarrow \int\limits_a^b f$.
If the functions f, f_n are differentiable then we say that the sequence $(f_n)_n$ can be ***term by term differentiated*** if $f_n' \longrightarrow f'$.

Theorem (term by term integration)
Let $(f_n)_n$ be a sequence of continuous functions $f_n : [a,b] \mapsto \mathbb{R}$ s.t. $f_n \xrightarrow{u} f$. Then $\int_a^b f_n \longrightarrow \int_a^b f$.
Proof By using the previous theorem, f is continuous and so integrable; moreover:

$$\left| \int_a^b f_n - \int_a^b f \right| = \left| \int_a^b (f_n - f) \right| \leq \int_a^b |f_n - f|.$$

Let $\varepsilon > 0$; the uniform convergence gives $N_\varepsilon \in \mathbb{N}$ s.t.
$\forall n \geq N_\varepsilon \ \ |f_n(x) - f(x)| \leq \frac{\varepsilon}{b-a}$, $\forall x \in [a,b]$. Then:

$$\left| \int_a^b f_n - \int_a^b f \right| \leq \frac{\varepsilon}{b-a}(b-a) = \varepsilon, \ \forall n \geq N_\varepsilon.$$

Remark
(i) Generally, the pointwise convergence is not sufficient for the integration term by term. An example is the sequence of piecewise linear functions: $f_n(x) = \begin{cases} n^2 x & \text{if } 0 \leq x \leq \frac{1}{n} \\ n^2\left(\frac{2}{n} - x\right) & \text{if } \frac{1}{n} \leq x \leq \frac{2}{n} \\ 0 & \text{if } \frac{2}{n} \leq x \leq 1 \end{cases}$
The sequence is pointwise convergent to 0, but $\lim\limits_n \int_0^1 f_n = 1$.

(ii) The uniform convergence is not a necessary condition for the integration term by term; take $f_n(x) = x^n$, $x \in [0,1]$.
Then $f_n \xrightarrow{p} f,\ f(x) = \begin{cases} 0 & \text{if } 0 \le x < 1 \\ 1 & \text{if } x = 1 \end{cases}$, $f_n \overset{u}{\not\longrightarrow} f$ and
$\lim\limits_n \int_0^1 f_n = \lim\limits_n \frac{1}{n+1} = 0 = \int_0^1 f.$

Theorem (term by term differentiation)
Let $(f_n)_n$ be a sequence of functions $f_n : I \mapsto \mathbb{R}$ s.t:
(i) $f_n \xrightarrow{p} f$
(ii) $f_n' \xrightarrow{u} g$.
Then f is differentiable and $f' = g$.

Proof Take $a \in I$; it is enough to prove that $\lim\limits_{x\to a} \dfrac{f(x) - f(a)}{x - a} = g(a)$.
Let us define

$$\varphi_n(x) = \begin{cases} \frac{f_n(x)-f_n(a)}{x-a} & \text{if } x \ne a \\ f_n'(a) & \text{if } x = a \end{cases} \quad \text{and} \quad \varphi(x) = \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{if } x \ne a \\ g(a) & \text{if } x = a \end{cases}$$

It is clear that $\varphi_n \xrightarrow{p} \varphi$ and that the functions φ_n are continuous. It is enough to prove that $\varphi_n \xrightarrow{u} \varphi$. Let $\varepsilon > 0$; then, by the Cauchy criterion applied to f_n', we have:

$$\exists N_\varepsilon \in \mathbb{N} \text{ s.t. if } n, m \ge N_\varepsilon \ \ |f_n'(x) - f_m'(x)| \le \varepsilon,\ \forall x \in I.$$

Now apply the Lagrange mean value theorem to $f_n - f_m$ on the interval $[x, a]$ (or $[a, x]$) with $x \in I$. Finally we get:

$$|(f_n(x) - f_m(x)) - (f_n(a) - f_m(a))| \le \varepsilon\, |x - a|.$$

Dividing by $x - a$, $(x \ne a)$ it results:

$$|\varphi_n(x) - \varphi_m(x)| \le \varepsilon,\ \forall x \ne a, \forall n, m \ge N_\varepsilon.$$

The continuity of the functions φ_n and φ_m at a implies

$$|\varphi_n(x) - \varphi_m(x)| \le \varepsilon,\ \forall x \in I, \forall n, m \ge N_\varepsilon.$$

It is now clear that $\varphi_n \xrightarrow{u} \varphi$ and the theorem is proved.

Exercises

1. Let $f_n : X \mapsto \mathbb{R}$ and let $(a_n)_n$ be a sequence s.t. $a_n \longrightarrow 0$.
If $|f_n(x) - f(x)| \leq a_n$, $\forall x \in X$ then $f_n \xrightarrow{u} f$.
Hint $\sup\limits_{x \in X} |f_n(x) - f(x)| \leq a_n$.

2. Test for uniform convergence the sequences:
(i) $f_n(x) = x^n,\ x \in [0, a],\ 0 < a < 1$.
(ii) $g_n(x) = (x + n)^{-1},\ x \in (0, \infty)$.
Hint Both sequences are uniformly convergent to 0.

3. Study the pointwise and uniform convergence of the following sequences of functions:
(i) $f_n : (0, 1) \mapsto R,\ f_n(x) = (nx + 1)^{-1}, n \geq 0$.
(ii) $g_n : [0, 1] \mapsto R\ ;\ f_n(x) = x^n - x^{2n}\ , n \geq 0$.
Hint (i) Let $x > 0$; $\lim\limits_{n \to \infty} f_n(x) = \lim\limits_{n \to \infty} (nx + 1)^{-1} = 0$, hence f_n converges pointwise to the null function.
Obviously, $\sup\limits_{x \in (0,1)} |f_n(x)| = \sup\limits_{x \in (0,1)} \left|(nx + 1)^{-1}\right| = 1$ hence f_n is not uniformly convergent.
(ii) $g_n \xrightarrow{p} 0$. The uniform convergence:

$$\sup_{x \in (0,1)} |g_n(x)| = \sup_{x \in (0,1)} |x^n - x^{2n}| = g_n\left(\frac{1}{\sqrt[n]{2}}\right) = \frac{1}{4},$$

hence g_n is not uniformly convergent.

4. Test for uniform convergence:
(i) $f_n(x) = \sin \frac{x}{n},\ x \in \mathbb{R}$.
(ii) $g_n(x) = \sqrt{x^2 + \frac{1}{n^2}},\ x \in \mathbb{R}$.
Hint (i) Take $x_n = n\frac{\pi}{2}$; then $f_n(x_n) \not\longrightarrow 0$.
(ii) g_n is uniformly convergent to the function $g(x) = |x|$.

5. Let $u_n : R \mapsto R,\ u_n(x) = x + \frac{1}{n}$.
Study the pointwise and uniform convergence of u_n and u_n^2.
Hint The sequence u_n is uniformly convergent to $u(x) = x$, while u_n^2 is pointwise convergent to $u^2(x) = x^2$, but fails to be uniformly

convergent:

$$\lim_{n\to\infty}\sup_{x\in R}|u_n^2(x)-x^2| = \lim_{n\to\infty}\sup_{x\in\mathbb{R}}\left|2\frac{x}{n}+\frac{1}{n^2}\right| = \infty.$$

6. Let $f_n(x) = nxe^{-nx^2}$, $x \in [0,1]$. Can this sequence be integrated term by term on $[0,1]$?

Hint First, $f_n \stackrel{p}{\longrightarrow} 0$; the sequence of the integrals $\int_0^1 f_n \longrightarrow 2^{-1}$.

7. Let $f_n(x) = \frac{1}{n}\arctan x^n,\ x \in \mathbb{R}$. Can this sequence be differentiated term by term ?

Hint First, $f_n \stackrel{u}{\longrightarrow} 0$; the sequence f_n' is not pointwise convergent (check at $x = -1$).

8. Let $u_n : (0,\infty) \mapsto R,\ u_n(x) = e^{-nx}$. Test the pointwise and uniform convergence of u_n and u_n'.

Hint Both u_n and u_n' are pointwise convergent to 0 but none of them is uniformly convergent.

9. Let $f_n(x) = \dfrac{\sin nx}{n}$ and $g_n(x) = \dfrac{\sin nx}{n^2}$.

Can these sequences be differentiating term by term?

Hint The answer is affirmative for g_n, but negative for f_n.

10. Test for pointwise and uniform convergence:

(i) $f_n(x) = x^n e^{-nx},\ x \geq 0$.

(ii) $g_n(x) = nx(1+n+x)^{-1},\ x \in [0,1]$.

(iii) $h_n(x) = (e^{nx}-1)(e^{nx}+1)^{-1},\ x < 0$.

2.2 Series of functions and power series

The theory of convergence of series of functions can be reduced to that of sequences of functions.

If $(f_n)_n$ is a sequence of functions, $f_n : X \mapsto \mathbb{R}$ then we consider the sequence of ***partial sums*** $(S_n)_n$ defined by $S_n = f_0 + f_1 + ...f_n$;

the series $\sum_n f_n$ is the pair of these two sequences $(f_n)_n$ and $(S_n)_n$.

Definition
The series $\sum_n f_n$ is said to be **pointwise convergent** if the sequence $(S_n)_n$ is pointwise convergent. If $S_n \xrightarrow{p} S$ then S is the ***pointwise sum*** of the series and is denoted by $\sum_n f_n$.
The series $\sum_n f_n$ is said to be ***uniformly convergent*** if the sequence $(S_n)_n$ is uniformly convergent; if $S_n \xrightarrow{u} S$ then S is the ***uniform sum*** of the series and is denoted by the same symbol, $\sum_n f_n$.

Theorem (Cauchy Criterion)
The series $\sum_n f_n$ is uniformly convergent iff:

$$\forall \varepsilon > 0,\ \exists N_\varepsilon \in \mathbb{N} \text{ s.t. if } \forall k \in \mathbb{N}^\star,\ \forall n \geq N_\varepsilon \quad \left|\sum_{j=1}^{k} f_{n+j}(x)\right| \leq \varepsilon,\ \forall x \in X.$$

Using this theorem a very useful test for uniform convergence of series of functions is obtained.

Proposition (test for uniform convergence of series)
Let $\sum_n f_n$ be a series of functions and let $\sum_n a_n$ be a convergent series of positive numbers and suppose that $|f_n(x)| \leq a_n$, $\forall n \in \mathbb{N}$, $\forall x \in X$. Then $\sum_n f_n$ is uniformly convergent.
Proof We have $|f_{n+1}(x) + f_{n+2} + ... + f_{n+k}(x)| \leq a_{n+1} + ... + a_{n+k}$, $\forall x \in X$; now apply Cauchy criterion.

Example
Let $f_n(x) = \dfrac{\sin nx}{n^3}$, $x \in \mathbb{R}$. The series is uniformly convergent because one can take $a_n = \frac{1}{n^3}$ in the above proposition.

Theorem (transfer of continuity)
Let $I \subseteq \mathbb{R}$ be an interval, let $a \in I$ and let $f_n : I \mapsto \mathbb{R}$ be continuous

functions at a. If $\sum_n f_n$ is uniformly convergent then its sum is continuous at a.
Proof Apply the corresponding theorem for sequences to $(S_n)_n$.

Example
The sum of the series $\sum_n \frac{\sin nx}{n^3}$ is continuous (see the above example).

Theorem (term by term integration)
If $\sum_n f_n$ is a uniformly convergent series of continuous functions on $[a, b]$ with sum S then $\int_a^b S = \sum_n \int_a^b f_n$.
Proof Apply the theorem on integration term by term for sequences of functions to the sequence of partial sums, $(S_n)_n$.

Example
Let $S(x) = \sum_{n\geq 1} \frac{\sin nx}{n^3}$; as the hypothesis of the above theorem are fulfilled, we have:
$\int_0^\pi S(x)dx = \sum_{n\geq 1} \int_0^\pi \frac{\sin nx}{n^3}\, dx = \sum_{n\geq 1} \frac{1-(-1)^n}{n^4}$.

Theorem (term by term differentiation)
If $\sum_n f_n$ is a series of differentiable functions on an interval I s.t:
$\sum_n f_n$ is pointwise convergent to f and
$\sum_n f_n'$ is uniformly convergent to g, then f is differentiable and $f' = g$.
Proof Apply the corresponding theorem for sequences of functions.

Example
If $S(x) = \sum_{n\geq 1} \frac{\sin nx}{n^3}$, then $S'(x) = \sum_{n\geq 1} \left(\frac{\sin nx}{n^3}\right)' = \sum_{n\geq 1} \frac{\cos nx}{n^2}$ because the series $\sum_{n\geq 1} \frac{\cos nx}{n^2}$ is uniformly convergent (why ?).

Definition (power series)
A ***power series*** is a series of the form

$$(\star) \qquad \sum_{n\geq 0} a_n z^n = a_0 + a_1 z + a_2 z^2 + ... + a_n z^n + ...$$

Clearly, this is a series of functions defined by $f_n(z) = a_n z^n,\ z \in \mathbb{C}$. The complex numbers a_n are the ***coefficients*** of the power series. If $a_n \in \mathbb{R},\ \forall n \in \mathbb{N}$ and $x \in \mathbb{R}$ the power series

$$(\star\star) \qquad \sum_{n\geq 0} a_n x^n = a_0 + a_1 x + a_2 x^2 + ... + a_n x^n + ...$$

is a ***power series in*** $\mathbb{R}$.

We could also consider power series in $z - z_0$ ($z_0 \in \mathbb{C},$ fixed):

$$\sum_{n\geq 0} a_n (z-z_0)^n = a_0 + a_1(z-z_0) + a_2(z-z_0)^2 + ... + a_n(z-z_0)^n + ...$$

Setting $w = z - z_0$ one can reduce the study of such series to $(\star)$.

We shall consider mainly series in $\mathbb{R}$, but there are reasons to start with power series in $\mathbb{C}$. Of course, we are interested in the convergence of power series.

Proposition
Let $0 \neq z_0 \in \mathbb{C}$ be such the sequence $(a_n z_0^n)_n$ is bounded and let $0 < r < |z_0|$; then:
(i) for every $z \in \mathbb{C},\ |z| < |z_0|$, the series $(\star)$ is absolutely convergent (as a series of numbers).
(ii) the series $(\star)$ is uniformly convergent on the closed disk
$D'(0,r) = \{z \in \mathbb{C}\ ;\ |z| \leq r\}$.
Proof (i) Let $M > 0$ be s.t. $|a_n z_0^n| \leq M,\ \forall n \in \mathbb{N}$ and let $z \in \mathbb{C}$ s.t. $|z| < |z_0|$; then:

$$|a_n z^n| \leq \left| a_n\ z_0^n\ \frac{z^n}{z_0^n} \right| \leq M \left|\frac{z}{z_0}\right|^n$$

Now observe that $\left|\dfrac{z}{z_0}\right| < 1$ and apply the comparison test.
(ii) Let $0 < r < |z_0|$; if $|z| < r$ then $|a_n\, z^n| \le M \left|\frac{r}{z_0}\right|^n$; now apply the test for uniform convergence of series (it obviously works for complex valued functions too).

The convergence (pointwise and uniform) of the power series $(\star)$ is clarified by:

Theorem (radius of convergence)
Given the series $(\star)$ then there is an ***unique*** $R \in [0, \infty]$ s.t:
(i) if $R = 0$ the series $(\star)$ converges only for $z = 0$ and if $R = \infty$ the series is absolutely convergent for every $z \in \mathbb{C}$;
(ii) if $0 < R < \infty$ then for $|z| < R$ the series $(\star)$ is absolutely convergent and for $|z| > R$ the series is divergent;
(iii) if $0 < r < R$ the series $(\star)$ is uniformly convergent on the closed disk $D'(0, r)$.
R is called the ***radius of convergence*** of the power series $(\star)$.
Proof Take $R = \sup\{r\ ;\ r \ge 0$ s.t. the series $\sum_n |a_n| r^n$ is convergent$\}$ (this set is non empty); then $R \in [0, \infty]$.
(i) and (ii) If $R = 0$ the result is obvious. If $0 \ne |z| < R$ then we find $|z| < r < R$ s.t. the series $\sum_n |a_n| r^n$ be convergent, so $\sum_n |a_n z^n|$ is convergent as well.
If $|z| > R$ and if the series $\sum_n a_n z^n$ would be convergent then $z \ne 0$ and the sequence $(a_n z^n)_n$ is bounded. Taking $|z| > r > R$ and applying the above proposition the series $\sum_n |a_n| r^n$ would be convergent, contradicting the definition of R.
(iii) We leave the proof as an exercise.

Remark
(i) The above theorem gives ***no information*** about points on the circle $|z| = R$ (if $0 < R < \infty$.
(ii) For the real series $(\star\star)$ we have:
for $-R < x < R$ the series is absolutely convergent;
for $|x| > R$ the series is divergent;
if $0 < r < R$ the series is uniformly convergent on $[-r, r]$.

We shall accept the following formulas for computing the radius of convergence:

Proposition
(i) If the limit $\ell = \lim_n \left| \frac{a_n}{a_{n+1}} \right|$ exists (in $[0, \infty]$), then $R = \ell$.
(ii) If the limit $\ell = \sqrt[n]{|a_n|}$ exists (in $[0, \infty]$) then $R = \ell^{-1}$, with the conventions $\frac{1}{0} = \infty$ and $\frac{1}{\infty} = 0$.

Example
(i) Consider the geometric series $1 + z + z^2 + ... + z^n + ...$; obviously, the radius of convergence is $R = 1$ (apply the above proposition (i)).
(ii) Consider the (very important) power series $1 + \frac{z}{1!} + \frac{z^2}{2!} + ... + \frac{z^n}{n!} + ...$
Then the radius of convergence is $R = \infty$.

From now on, up to the end of this section, we consider power series of type $(\star\star)$. The following theorem is crucial.

Theorem
Let the radius of convergence of the power series $\sum_n a_n x^n$ be $R > 0$.
If f is the (pointwise) sum of the series on $(-R, R)$, then:
(i) f is continuous.
(ii) f has derivatives of ***any order*** and these derivatives can be obtained by term by term differentiation .
(iii) The series can be term by term integrated on every compact interval $[a, b] \subset (-R, R)$.
Proof (i) Take $x_0 \in (-R, R)$ and let $0 < r < R$ s.t. $x_0 \in (-r, r)$. By applying the transfer of continuity on $[-r, r]$ we obtain the result.
(ii) First, let us observe (exercise !) that the (power) series of the derivatives $\sum_{n \geq 1} n a_n x^{n-1}$ has ***the same radius of convergence*** as the initial series. Using this fact and reasoning as in (i) we obtain the result by term by term differentiation theorem.
(iii) A trivial application of the term by term integration theorem.

Remark
Generally, we do not have uniform convergence on $(-R, R)$ but only on compact intervals. That's why we need the trick in the proof of (i) above; of course, (ii) implies (i), but we prefer to state it apart.

Definition
Let $\sum_{n\geq 0} a_n x^n$ be a power series with radius of convergence $R > 0$ and let $f : (-R, R) \mapsto \mathbb{R}$ be its sum. In this case we shall say, also, that $\sum_{n\geq 0} a_n x^n$ is the ***expansion*** of the function f in a power series.

Proposition
Let $f(x) = \sum_{n\geq 0} a_n x^n$, $x \in (-R, R)$. Then $a_n = \frac{f^{(n)}(0)}{n!}$, $\forall n \in \mathbb{N}$; as usual, $f^{(n)}$ denotes the $n-$derivative of f.
Proof Obviously $a_0 = f(0)$; then by applying the previous theorem $f'(x) = \sum_{n\geq 1} n a_n x^{n-1}$, so $a_1 = f'(0)$. The result follows by induction.

Remark
An important consequence of the above result is the fact that the expansion (in a power series) of a function is unique (if it exists !).

Examples
(i) We start with the geometric series:

$$\frac{1}{1-x} = 1 + x + x^2 + ... + x^n + ..., |x| < 1.$$

Changing x into $-x$ we get:

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + ... + (-1)^n x^n + ..., |x| < 1$$

By integrating term by term (on what interval ?) we get:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + ... + (-1)^n \frac{x^{n+1}}{n+1} + ..., |x| < 1$$

(ii) Analogously, starting with:

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + ... + (-1)^n x^{2n} + ..., |x| < 1,$$

we obtain:

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + ... + (-1)^n \frac{x^{2n+1}}{2n+1} + ..., |x| < 1.$$

So we have expanded the functions $\ln(1+x)$ and $\arctan x$ in power series (on the corresponding intervals).

We mention (without proof) an important theorem of Abel (see [8]):

Theorem (Abel)
If the series of numbers $\sum_n a_n$ is convergent and if
$f : (-1, 1) \mapsto \mathbb{R},\ f(x) = \sum_n a_n x^n$, then $\lim_{x \nearrow 1} f(x) = \sum_n a_n$.

Example
Starting with the alternating harmonic series, the expansion:

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + ... + (-1)^n \frac{x^{n+1}}{n+1} + ..., |x| < 1$$

gives that:

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ... + (-1)^n \frac{1}{n+1} + ...,$$

so we have computed the sum of the alternating harmonic series.

Exercises
1. Compute $\arctan 1 = \frac{\pi}{4}$ with an absolute error less than 10^{-1}. Deduce an approximation of π with error less than 10^{-2}.
Hint Starting with the convergent series $\sum_{n \geq 0} (-1)^n \frac{1}{2n+1}$ and by applying Abel's theorem to the expansion

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + ... + (-1)^n \frac{x^{2n+1}}{2n+1} + ..., |x| < 1$$

we obtain the sum $\sum_{n\geq 0}(-1)^n\frac{1}{2n+1} = \arctan 1 = \frac{\pi}{4}$. Now approximate the sum of the alternating series of the left side.

2. Study the pointwise and uniform convergence of the series $\sum_{n\geq 0} e^{-nx}$ on $(0,\infty)$.
Hint Put $e^{-x} = t$, etc.

3. Can be the series $\sum_{n=1}^{\infty}\frac{\sin nx}{2^n}$ term by term differentiated on $\mathbb{R}$?
Hint $\left|\frac{\sin nx}{2^n}\right| \leq \frac{1}{2^n}$,$\forall x \in \mathbb{R}$, so the series is uniformly convergent on $\mathbb{R}$. Same is true for the series of the derivatives, hence the answer is affirmative.

4. Consider the series $\sum_{n\geq 1}\frac{\sin nx}{n^2}$, $x \in \mathbb{R}$. Can it be term by term differentiated ?
Hint The answer is negative (the series of the derivatives is not pointwise convergent).

5. Let $a, b \in \mathbb{R}, 0 < a < b$; check if the series $\sum_{n\geq 1}\frac{(x+n)^2}{n^4}$ can be term by term differentiated on $[a, b]$.
Hint We have $(x+n)^2 \leq (b+n)^2, \forall x \in [a, b]$, so

$$\sum_{n\geq 1}\frac{(x+n)^2}{n^4} \leq \sum_{n\geq 1}\frac{(b+n)^2}{n^4},\ \forall x \in [a, b].$$

Analogously, the series of the derivatives is uniformly convergent on $[a, b]$, so the answer is positive.

6. Consider the series $\sum_{n\geq 1}(-1)^n\frac{x^2+n}{n^2}$, $x \in R$.
(i) Prove that the series is pointwise convergent for every $x \in \mathbb{R}$.
(ii) Study the uniform convergence on every closed and bounded interval.

(iii) Study the uniform convergence on $\mathbb{R}$.
(iv) Is the sum a continuous function ?
(v) Same questions for the series of the derivatives.
(vi) Can be the series term by term differentiated?
Hint Study the series $\sum_n (-1)^n \frac{x^2}{n^2}$ and $\sum_n (-1)^n \frac{1}{n}$.

7. Let $\sum_{n\geq 1} \frac{nx^n + x}{n^2+1}, x \in (-1,1)$.
(i) Study the pointwise and uniform convergence
(ii) Can the series be term by term differentiated?
Hint Study the series $x \sum_{n\geq 1} \frac{1}{n^2+1}$ and $\sum_{n\geq 1} \frac{n}{n^2+1} x^n$.

8. Compute the radius of convergence and study the convergence of the series: $\sum_{n\geq 1} \frac{z^n}{n}$, $\sum_{n\geq 1} \frac{(-1)^n}{n} z^n$ and $\sum_{n\geq 1} n! z^n$.
Hint The first two series have the radius of convergence $R = 1$. The first one is divergent if $z = 1$ and convergent for $|z| = 1, z \neq 1$. Analogously for the second one. The third series has the radius of convergence 0, so it it convergent only for $z = 0$.

9. Compute the sum of the series $\sum_{n\geq 0} \frac{(-1)^n}{3n+1}$ by term by term integration of the power series $\sum_{n\geq 0} \frac{(-1)^n}{3n+1} x^{3n+1}$.
Hint By applying Abel theorem we get:

$$\sum_{n\geq 0} \frac{(-1)^n}{3n+1} = \lim_{x\to 1} \sum_{n\geq 0} \frac{(-1)^n}{3n+1} x^{3n+1} =$$

$$= \lim_{x\to 1} \int_0^x \sum_{n\geq 0} (-1)^n x^{3n}\, dx = \lim_{x\to 1} \int_0^x \frac{dx}{1+x^3}$$

10. Compute the sum of the series $\sum_{n\geq 0} \frac{n+1}{4^n}$ by term by term differentiation of the power series $\sum_{n\geq 0} x^{n+1}$.

Hint If $|x| < 1$, then:

$$\sum_{n\geq 0}(n+1)x^n = \left(\sum_{n\geq 0} x^{n+1}\right)' = \frac{1}{(1-x)^2}, \text{ and put } x = \frac{1}{4}$$

2.3 Elementary functions

The aim of this section is to define (by using power series) some basic transcendental elementary functions, such as the exponential, sine (sin) and cosine (cos).
We start by defining the exponential in the complex domain taking advantage of the possibility of obtaining the (real) sine and cosine from it. This point of view (due to Euler) is, not only elegant, but very useful in showing the connection between the exponential and the trigonometric functions.

The exponential function
Define the ***complex exponential function*** by:

$$(\star) \quad \exp(z) = 1 + \frac{z}{1!} + \frac{z^2}{2!} + ... + \frac{z^n}{n!} + ...$$

As the radius of convergence of the series is ∞ so $\exp : \mathbb{C} \mapsto \mathbb{C}$.

The basic properties of exp are:
1. $\exp(0) = 1$
2. $\exp(z+w) = \exp(z)\,\exp(w),\ \forall z, w \in \mathbb{C}$.
The first property is trivial. To obtain the second, we can use the following theorem about the multiplication of series (which is interesting by itself):

Theorem
Let $\sum_n a_n$, $\sum_n b_n$ be two series of complex numbers and define the ***product series*** $\sum_n c_n$ by $c_n = a_0b_n + a_1b_{n-1} + ... + a_nb_0 = \sum_{k=0}^{n} a_kb_{n-k}$.

If $\sum_n a_n$ is absolutely convergent and $\sum_n b_n$ is convergent then the product series is convergent and $\left(\sum_n a_n\right) \cdot \left(\sum_n b_n\right) = \sum_n c_n.$
For the proof see, for example, [4].
Now use this theorem to prove the above second property (exercise).

A consequence of properties 1 and 2 is that

$$\exp(z)\exp(-z) = \exp(0) = 1, \text{ so } \exp(z) \neq 0, \forall z \in \mathbb{C}.$$

Remark
We can give an algebraic form of property 2 by considering the abelian groups $(\mathbb{C}, +)$ and $(\mathbb{C}^\star, \cdot)$. Then $\exp : \mathbb{C} \mapsto \mathbb{C}^\star$ is a group homomorphism.

The ***real exponential function*** is the restriction of exp to $\mathbb{R}$:

$$(\star\star) \quad \exp(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + ... + \frac{x^n}{n!} + ..., \ \forall x \in \mathbb{R}.$$

Using the basic properties of power series we get:
(i) $\exp(x) > 0, \ \forall x \in \mathbb{R}$.
In fact $\exp(x) \neq 0, \ \forall x \in \mathbb{R}$, so the continuity of exp implies that it keeps a constant sign and $\exp(0) = 1$, so we get the result.
(ii) The exponential has derivatives of any order and these derivatives can be obtained by term by term differentiation; obviously, by differentiating the equality $(\star\star)$ we get

$$\exp' = \exp$$

and by induction $\exp^{(n)} = \exp$.
(iii) The exponential is ***strictly increasing*** and $\lim_{x\to\infty} \exp(x) = \infty$; in fact $\exp' = \exp > 0$. Moreover, $\exp(x) > x, \ \forall x > 0$.
(iv) $\lim_{x\to-\infty} \exp(x) = 0$.
(v) The exponential $\exp : \mathbb{R} \mapsto (0, \infty)$ is an ***isomorphism*** between the groups $(\mathbb{R}, +)$ and $((0, \infty), \cdot)$. The inverse of this isomorphism is denoted by ln (logarithm). The properties of the logarithm are easily deduced from those of the exponential.

Remark

(i) One can ask if the function exp defined by $(\star)$ is the same as the exponential e^x already studied at the college. The answer is yes because there exists an unique function $f : \mathbb{R} \mapsto \mathbb{R}$ satisfying the conditions $f' = f$ and $f(0) = 1$; in fact, if f is such a function then

$$\left(\frac{f}{\exp}\right)' = \frac{f' \exp - \exp\, f}{\exp^2} = 0,$$

so there is $k \in \mathbb{R}$ s.t. $f = k \exp$. But $f(0) = \exp(0) = 1$ so $f = \exp$.
(ii) The previous remark gives a hint for the proof of the basic property $\exp(x + y) = \exp(x)\exp(y)$ in the real case. Actually, for $c \in \mathbb{R}$ we have: $(\exp(c - x)\exp(x))' = 0$, so $\exp(c-x)\exp(x) = \exp(c)$, $\forall x \in \mathbb{R}$; take now $c = x + y$.
(iii) Due to these remarks, we use the notation $e^z = \exp(z), z \in \mathbb{C}$.

The trigonometric functions

From $(\star)$ we obtain $e^{\overline{z}} = \overline{e^z}$, $\forall z \in \mathbb{C}$. So if $x \in \mathbb{R}$ we have that $|e^{ix}|^2 = e^{ix}\, e^{-ix} = 1$ so $|e^{ix}| = 1$, $\forall x \in \mathbb{R}$.

Definition

Define the trigonometric functions ***sine*** and ***cosine*** by:

$$(\star\star\star) \quad \sin x = \Im(e^{ix}) \;\text{ and }\; \cos x = \Re(e^{ix}),\; \forall x \in \mathbb{R}$$

So, by the very definition, we have the Euler formula:

$$e^{ix} = \cos x + i \sin x,\; \forall x \in \mathbb{R}$$

As $e^{ix} = 1,\; \forall x \in \mathbb{R}$, we obtain:

$$\sin^2 x + \cos^2 x = 1 \;\text{ so }\; |\sin x|, |\cos x| \le 1,\; \forall x \in \mathbb{R}$$

Replacing $z = ix$ in $(\star)$ and separating the real and imaginary parts we get the (real) power series expansions of cos and sin:

$$(\star\star\star\star) \quad \cos x = \sum_{n\ge 0} \frac{(-1)^n}{(2n)!}\, x^{2n},\;\; \sin x = \sum_{n\ge 0} \frac{(-1)^n}{(2n+1)!}\, x^{2n+1},\; \forall x \in \mathbb{R}.$$

Clearly, cos and sin have derivatives of any order and these derivatives can be obtained by term by term differentiation. We have and: $\cos' = -\sin$ and $\sin' = \cos$.

The number π

As $\cos 0 = 1$ (and cos is continuous as a sum of a power series), then the function cos is strictly positive in a small interval centered at 0. Define the set $M = \{x\ ;\ x > 0,\ \cos x = 0\}$. We want to prove that $M \neq \emptyset$. Let us first observe that this is not obvious from the above definitions (all we have until now).

Proposition

The set M is non empty.

Proof Suppose, by contradiction, that $M = \emptyset$; then cos has to keep a constant sign (the continuity !) on $(0, \infty)$, so $\cos x > 0, \forall x > 0$ due to $\cos 0 = 1$. It results that sin is strictly increasing and as $\sin 0 = 0$, we have that $\sin x > 0, \forall x > 0$. Take $a > 0$ and consider the function $f : [a, \infty)$, $f(x) = \cos x + x \sin a$; the derivative $f'(x) = -\sin x + \sin a < 0$, $\forall x > a$, so f is strictly decreasing on $[a, \infty)$. But as cos is a bounded function we get that $\lim\limits_{x \to \infty} f(x) = \infty$ and this contradicts the fact that f is decreasing; so $M \neq \emptyset$.

We now define the number π; take $m = \inf M$. Then $m > 0$ and $m \in M$ (prove these statements). Define $\pi = 2m$. In this way, $\frac{\pi}{2}$ is the least positive zero of cosine.

Remark

π is just "a name" for a real number whose existence was proved. Now (from the definition): $e^{i\frac{\pi}{2}} = \cos\dfrac{\pi}{2} + i \sin\dfrac{\pi}{2} = i$ (why is $\sin\frac{\pi}{2} = 1$?).

From the property (ii) of the exponential we obtain $e^{i\pi} = -1$ and $e^{2\pi i} = 1$, so $e^{z+2\pi i} = e^z$, $\forall z \in \mathbb{C}$. It results that the complex exponential is ***periodic*** on $\mathbb{C}$ (not on the real line).

By analyzing the functions cos and sin on $[0, 2\pi]$ we can easily obtain the well known properties of trigonometric functions (periodicity, etc). We leave this analysis to the reader.

The binomial series
For $x > 0$ and $\alpha \in \mathbb{R}$ define $x^\alpha = e^{\alpha \ln x}$. An important power series expansion is given bellow:

Proposition
One has for $\alpha \in \mathbb{R}$:

$$(1+x)^\alpha = 1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+...+\frac{\alpha(\alpha-1)...(\alpha-n+1)}{n!}x^n+..., |x| < 1$$

The power series in the right side is called the ***binomial series***. We do not prove this formula (see [1]), just observe that it is a generalization of a well known formula in combinatorics.

Example
Find the power expansion of the function arcsin : $(-1, 1) \mapsto (-\frac{\pi}{2}, \frac{\pi}{2})$.
Solution We have $(\arcsin x)' = \left(1 - x^2\right)^{-\frac{1}{2}}$, $|x| < 1$; now apply the binomial series for $\alpha = -\frac{1}{2}$, replace x by $-x^2$ and integrate term by term.

Exercises

1. By using Euler formula prove:
$\sin(x + y) = \sin x \cos y + \sin y \cos x$,
$\cos(x + y) = \cos x \cos y - \sin x \sin y, \forall x, y \in \mathbb{R}$.
Hint Use $e^{i(x+y)} = e^{ix}e^{iy}$, $\forall x, y \in \mathbb{R}$.

2. Complex Riemann function
For $z \in \mathbb{C}$, consider the series $\sum_{n\geq 1} \frac{1}{n^z}$, (by definition: $n^z = e^{z \ln n}$).
Prove the absolutely convergence of the series for $\Re z > 1$.
Hint $\left|\frac{1}{n^z}\right| = \frac{1}{n^{\Re z}}$. The sum of the series is the ***Riemann function.***

3. Compute (by using power series) the integral $\int_0^1 \frac{\sin x}{x}\, dx$, with an error $\varepsilon < 10^{-2}$.

Hint Use the expansion of sine and integrate term by term:

$$\int_0^1 \frac{\sin x}{x}\,dx = \int_0^1 \sum_{n\geq 0} \frac{(-1)^n}{(2n+1)!} x^{2n}\,dx = \sum_{n\geq 0} \frac{(-1)^n}{(2n+1)!\,(2n+1)}.$$

Now approximate the alternating series.

4. Compute the integral $\int_0^1 \frac{\arctan x}{x}\,dx$ with an error $\varepsilon < 10^{-2}$.
Hint Use the expansion of arctan.

5. Compute the integral $\int_0^1 \frac{\ln(1+x)}{x}\,dx$ with an error $\varepsilon < 10^{2}$.
Hint Use the expansion of $\ln(1+x)$.

6. Expand in a power series arcsin and prove:

$$1 + \sum_{n\geq 1} \frac{1\cdot 3\cdot 5 \ldots \cdot (2n-1)}{2\cdot 4\cdot 6 \ldots (2n)} \cdot \frac{1}{2n+1} = \frac{\pi}{2}.$$

Hint By using the binomial series ($\alpha = -\frac{1}{2}$) we get:

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}} = 1 + \sum_{n\geq 1} \frac{1\cdot 3\cdot 5 \ldots \cdot (2n-1)}{2\cdot 4\cdot 6 \ldots (2n)} x^{2n}, \forall\, |x| < 1.$$

Term by term integration gives:

$$\arcsin x = x + \sum_{n\geq 1} \frac{1\cdot 3\cdot 5 \ldots (2n-1)}{2\cdot 4\cdot 6 \ldots (2n)} \cdot \frac{1}{2n+1} x^{2n+1}, \forall\, x \in (-1,1).$$

The series is convergent for $x = 1$ (Raabe's test); now apply Abel theorem.

7. Expand in a power series $\ln(x + \sqrt{1+x^2})$ and prove:

$$1 + \sum_{n\geq 1} (-1)^n \frac{1\cdot 3\cdot 5 \ldots (2n-1)}{2\cdot 4\cdot 6 \ldots (2n)} \cdot \frac{1}{2n+1} = \ln(1+\sqrt{2}).$$

Hint Same method as above.

2.4 Taylor formulas

We shall use the following notations:
Let $\emptyset \neq I$ be an interval (of real numbers) not reduced to a point and let $f : I \mapsto \mathbb{R}$ be a map. We say that $f \in \mathcal{C}^0(I)$ if f is continuous; if $n \in \mathbb{N}, n \geq 1$ we say that $f \in \mathcal{C}^n(I)$ if f has derivatives up to the order n (included) and these derivatives are continuous. So, for example, $f \in \mathcal{C}^1(I)$ if f is differentiable on I and f' is a continuous function. If I is not open then at the end points of I we mean one sides derivatives.

Let $f : I \mapsto \mathbb{R}$ and $a \in I$ and suppose that $f^{(n)}(a)$ exists (finite). We define the ***Taylor polynomial of order*** n ***for*** f ***at*** a by:

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

The basic property of the Taylor polynomial is (as it is easy to check) that the values at a of f and of its derivatives coincide up to the order n with the corresponding values of $P_n(a,x,f)$:

$$f^{(k)}(a) = P_n^{(k)}(a,x,f),\ \forall k = 0, 1, ..., n.$$

One can say that $P_n(a,x,f)$ is a kind of approximation of f around a. Let us define $R_n(a,x,f) = f(x) - P_n(a,x,f)$, $x \in I$, so, trivially:

$$(\star) \qquad f(x) = P_n(a,x,f) + R_n(a,x,f), \forall x \in I.$$

A ***Taylor formula*** is a relation of type $(\star)$ as above together with an ***estimation*** of $R_n(a,x,f)$ (called the **remainder** ***of order*** n in Taylor formula).

We shall give, without proofs (see [1], [4], [5], [8]) two Taylor formulas which are very useful in the local study and approximation of functions.

The Taylor-Lagrange formula
This is a generalization of the mean-value theorem formula of Lagrange.

Theorem
Let $f \in \mathcal{C}^n[a,b]$ and $f \in \mathcal{C}^{n+1}(a,b)$. Then there exists $c \in (a,b)$ s.t:

$$(\star\star) \quad f(b) = f(a) + \frac{f'(a)}{1!}(b-a) + ... + \frac{f^{(n)}(a)}{n!}(b-a)^n + \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}$$

In this way the remainder can be written as:

$$R_n(a,b,f) = \frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}$$

So, for example, if there exists $M > 0$ s.t. $\left|f^{(n+1)}(x)\right| \leq M, \forall x \in (a,b)$ then we obtain the estimation :

$$|R_n(a,b,f)| \leq M\frac{(b-a)^{n+1}}{(n+1)!}$$

Remark
Of course one can apply $(\star\star)$ on every interval $[a,x]$, $\forall\, a < x \leq b$; (what if we consider intervals $[x,a]$?)

The Taylor-Young formula
Suppose $f : I \mapsto \mathbb{R}$ and that $f^{(n)}(a)$ exists $(n \geq 1)$.

Proposition
In the above conditions we have:

$$(\star\star\star) \quad \lim_{x \to a} \frac{f(x) - f(a) - \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k}{(x-a)^n} = 0$$

For $n = 1$ this is the definition of the derivative $f'(a)$.

Remark
Intuitively, $(\star\star\star)$ says that the remainder in the Taylor formula tends to zero as x tends to a ***faster*** than $(x-a)^n$ (the meaning faster is $(\star\star\star)$).
To obtain the Taylor-Young formula we put:

$$\frac{f(x) - f(a) - \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k}{(x-a)^n} = \rho(x), x \neq a$$

Then we obtain the Taylor-Young formula:

Theorem

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \rho(x)(x-a)^n, x \neq a$$

and $\lim_{x \to a} \rho(x) = 0$.
Equivalently, if we define $\rho(a) = 0$ we can write:

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + \rho(x)(x-a)^n, \forall x \in I$$

and ρ is continuous and vanishes at a.

Landau's symbols $\boldsymbol{O}$ (big O) and $\boldsymbol{o}$ (small o)
Let φ be a map defined on an open interval I and $x_0 \in I$.
We define the ***sets*** $\boldsymbol{O}(\varphi)$ and $\boldsymbol{o}(\varphi)$ as follows:

(i) $f \in \boldsymbol{O}(\varphi)$ if the map f is defined on a neighborhood of x_0 and there exists a neighborhood V of x_0 and $K \geq 0$ s.t.

$$|f(x)| \leq K\,|\varphi(x)|, \forall x \in V.$$

(ii) $f \in \boldsymbol{o}(\varphi)$ if the map f is defined on a neighborhood of x_0 and

$$\forall \varepsilon > 0,\ \exists \delta_\varepsilon > 0 \text{ s.t. if } |x - x_0| < \delta_\varepsilon \text{ then } |f(x)| \leq \varepsilon\,|\varphi(x)|.$$

One can also write $f = \boldsymbol{O}(\varphi)$ or $f = \boldsymbol{o}(\varphi)$.

Proposition
Let φ and x_0 be as above and $\varphi(x) \neq 0$ if $x \neq x_0$. Then:
(i) $f \in \boldsymbol{O}(\varphi) \iff \dfrac{f}{\varphi}$ is bounded in a neighborhood of x_0 (without x_0).
(ii) $f \in \boldsymbol{o}(\varphi) \iff \lim_{x \to x_0} \dfrac{f(x)}{\varphi(x)} = 0$.

We now can restate Taylor-Young formula by using the symbols $\boldsymbol{O}$ and $\boldsymbol{o}$ as follows:

$$f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \boldsymbol{o}\left((x-a)^n\right).$$

and (if $f^{(n+1)}(a)$ exists):

$$f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \boldsymbol{O}\left((x-a)^{n+1}\right).$$

The first formula is obvious. To prove the second one, we have:

$$f(x) = \sum_{k=0}^{n} \frac{(x-a)^k}{k!} f^{(k)}(a) + \frac{f^{(n+1)}(a)}{(n+1)!} (x-a)^{n+1} + \boldsymbol{o}\left((x-a)^{n+1}\right)$$

and obviously:

$$\frac{f^{(n+1)}(a)}{(n+1)!} (x-a)^{n+1} + \boldsymbol{o}\left((x-a)^{n+1}\right) = \boldsymbol{O}\left((x-a)^{n+1}\right)$$

Examples

The usual power expansions can be restated with $\boldsymbol{O}$:

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + ... + \frac{x^n}{n!} + \boldsymbol{O}\left(x^{n+1}\right)$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + ... + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \boldsymbol{O}\left(x^{2n+3}\right)$$

$$\cos x = 1 - \frac{x^2}{2!} + ... + (-1)^n \frac{x^{2n}}{(2n)!} + \boldsymbol{O}\left(x^{2n+2}\right)$$

Extrema of functions of one variable

Let $I \subseteq \mathbb{R}$ be an interval, $a \in I$ and $f : I \mapsto \mathbb{R}$.
Then a is said to be a local maximum (minimum) point for f if there is a neighborhood U of a s.t. $f(x) \leq f(a)$ $(f(x) \geq f(a))$, $\forall x \in U \cap I$.
Local minima or maxima points are called local extrema.
We suppose known (from the school) ***Fermat's theorem***:

If I is an open interval, f is differentiable at a and a is local extremum point for f, then $f'(a) = 0$.

The converse is false, so we need also sufficient conditions for a critical point ($f'(a) = 0$) to be an extremum. Below we obtain such conditions based on Taylor's formula.

Let $I \subseteq \mathbb{R}$ be an open interval, f of class $\mathcal{C}^k, (k \geq 2)$, on I and $a \in I$ a critical point for f. Let us suppose the first $k-1$ $(k \geq 2)$ derivatives of f at a are zero:

$$f'(a) = f''(a) = ... = f^{(k-1)}(a) = 0$$

and $f^{(k)}(a) \neq 0$. Applying Taylor-Young formula we obtain:

$$f(x) - f(a) = \frac{f^{(k)}(a)}{k!}(x-a)^k + \rho(x)(x-a)^k =$$

$$= \left(\frac{f^{(k)}(a)}{k!} + \rho(x)\right)(x-a)^k,$$

ρ being a continuous function vanishing at a.

So we get the following **test**:

1. If k is even, then a is an extreme point, more precisely:
(i) if $f^{(k)}(a) > 0$, then a is a local minimum point.
(ii) if $f^{(k)}(a) < 0$, then a is a local maximum point.

2. If k is odd, then a is not an extreme point.

Exercises

1. Compute the Taylor polynomial of order n at 0 of $f(x) = e^x$.
Hint Obviously, $f^{(n)}(x) = e^x$.

2. (i) Compute the Taylor polynomial of order $2n+1$ at 0 of the function $g(x) = \sin x$.
(ii) Compute the Taylor polynomial of order $2n$ at 0 for $\cos x$.
Hint For every $x \in \mathbb{R}$ and $n \in \mathbb{N}$ one can prove (by induction):
$\sin^{(n)}(x) = \sin\left(x + n\frac{\pi}{2}\right)$, $\cos^{(n)}(x) = \cos\left(x + n\frac{\pi}{2}\right)$, etc.

4. Find the ***affine*** and ***quadratic approximations*** of the map $f(x) = x \ln x$ about $a = 1$.
Hint Affine and quadratic approximations are Taylor polynomials of first degree and second degree, respectively. We have:
$P_1(x) = f(1) + f'(1)(x-1) = x - 1$ and

$P_2(x) = f(1) + f'(1)(x-1) + \frac{1}{2!}f''(1)(x-1)^2 = (x-1) + \frac{1}{2}(x-1)^2.$

5. Prove **Taylor's Theorem**:
If $f \in \mathcal{C}^\infty[a,b] = \bigcap_{k\geq 0} \mathcal{C}^k[a,b]$ and if there exists $M > 0$ s.t.

$$|f^n(x)| \leq M,\ \forall x \in [a,b],\ \forall n \in \mathbb{N},$$

then the **Taylor series of f about a**

$$f(a) + \frac{f'(a)}{1!}(x-a) + ... + \frac{f^{(n)}(a)}{n!}(x-a)^n + ...$$

is uniformly convergent on $[a,b]$ to the function f.
Hint Apply the Taylor-Lagrange formula.

6. Prove that there are functions $f \in \mathcal{C}^\infty(\mathbb{R})$ for which the associated Taylor series has not the sum f.
Hint An usual example is: $f : R \mapsto R,\ f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$

7. Approximate $\sqrt{101}$ with an error $\varepsilon < 10^{-2}$.
Hint We have $\sqrt{101} = 10\sqrt{1+\frac{1}{100}}$ and use the Taylor polynomial at 0 of the function $f(x) = \sqrt{1+x}$.

8. Compute e^{-1} with error less than 10^{-3}.
Hint Applying Taylor-Lagrange formula, for every $x \in \mathbb{R}$ there is $\xi \in (0,x)$ (or $(x,0)$) s.t:

$$e^x = 1 + \frac{x}{1!} + ... + \frac{x^{n+1}}{n!} + \frac{\xi^{n+1}}{(n+1)!}$$

So there is $\xi \in (-1,0)$ st:

$$e^{-1} = 1 - \frac{1}{1!} + \frac{1}{2!} - ... + \frac{(-1)^n}{n!} + (-1)^{n+1}\frac{e^\xi}{(n+1)!}$$

But $e^\xi < 1$, so:

$$|R_n| = \frac{e^\xi}{(n+1)!} < \frac{1}{(n+1)!}$$

The least $n \in N$ s.t. $\frac{1}{(n+1)!} < 10^{-3}$ is $n = 6$, etc.

9. Find an upper bound for the error in the approximation

$$e^2 \approx 1 + \frac{2}{1!} + \frac{2^2}{2!} + \frac{2^3}{3!} + \frac{2^4}{4!} + \frac{2^5}{5!}$$

Hint If in Taylor-Lagrange formula for e^x we put $x = 2$ and $n = 5$, the remainder satisfies:

$$R_5 = \frac{e^\xi}{6!} \leq \frac{e^2}{6!} < \frac{9}{6!} = 0.0125$$

10. Find the least n for which the Taylor polynomial of order n approximates the function e^x on the interval $[-1, 1]$ with error less than 10^{-3}.
Hint The remainder of order n is

$$R_n = \frac{e^\xi}{(n+1)!} x^{n+1},$$

with an appropriate $\xi \in [-1, 1]$; we have:

$$|R_n| \leq \frac{e}{(n+1)!} < \frac{3}{(n+1)!}, \quad \text{etc.}$$

11. Compute the limits (by using Taylor polynomials)
(i) $\lim_{x \to 0} \frac{e^x - 1}{x}$.
(ii) $\lim_{x \to 0} \frac{\sin x}{x}$.
(iii) $\lim_{x \to 0} \frac{1 - \cos x}{x^2}$.

12. Find the extrema of the following functions:
(i) $f(x) = e^x x^{-e}$, $x > 0$; decide who is larger: e^π or π^e?
(ii) $f(x) = \arcsin \frac{2x}{1 + x^2}$, $x \in \mathbb{R}$.
(iii) $f(x) = |x| \ln |x|$ if $x \neq 0$ and $f(0) = 0$.

Chapter 3

Functions of several variables

3.1 The Euclidean space $\mathbb{R}^n$

Definition
For every natural number n define:

$$\mathbb{R}^n = \{x = (x_1, x_2, ..., x_n) \; ; \; x_k \in \mathbb{R}, \forall k = 1, 2, ..., n\},$$

hence $\mathbb{R}^n$ is the set of all (ordered) n-tuples of real numbers.
For a better understanding, if $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n)$ then $x = y$ iff $x_1 = y_1, x_2 = y_2, ..., x_n = y_n$.
For $n = 1$, then $\mathbb{R}^1 = \mathbb{R}$, for $n = 2$ then $\mathbb{R}^2$ is the usual Euclidean "plane" and for $n = 3$, then $\mathbb{R}^3$ is the usual Euclidean "space". The elements of $\mathbb{R}^n$ are called ***points***. If $x = (x_1, x_2, ..., x_n)$, then the real numbers $x_1, x_2, ...x_n$ are called the ***components*** of x (more precisely the first component, the second component, ..., the n-th component). So equality in $\mathbb{R}^n$ means equality of the (corresponding) components.

As usually, we shall denote points in $\mathbb{R}^2$ by (x, y) and in $\mathbb{R}^3$ by (x, y, z).

The vector structure on $\mathbb{R}^n$
For every $x = (x_1, x_2, ..., x_n)$, $y = (y_1, y_2, ..., y_n)$ and $\alpha \in \mathbb{R}$ define:

$$x + y = (x_1 + y_1, x_2 + y_2, ..., x_n + y_n)$$

$$\alpha x = (\alpha x_1, \alpha x_2, ..., \alpha x_n)$$

It is easy to check that in this way $\mathbb{R}^n$ becomes a ***real vector space***. Sometimes we shall write 0 for the null vector $(0, 0, ..., 0)$.
The dimension of the vector space $\mathbb{R}^n$ is n. The vectors:

$$e_1 = (1, 0, 0, ...0), e_2 = (0, 1, 0, ...0), ..., e_n = (0, 0, ..., 0, 1)$$

form the ***canonical basis*** of $\mathbb{R}^n$. We shall use only elementary properties of the vector space structure of $\mathbb{R}^n$.

Remark
The elements of $\mathbb{R}^n$ have now two names: points or vectors. The name "points" is generally used in describing "topological" properties of $\mathbb{R}^n$ (such as distance, open sets, etc); the name "vectors" is generally used in describing tangent vectors, etc. We do not bother with such distinctions at this level.

The dot product
If $x = (x_1, x_2, ..., x_n)$, $y = (y_1, y_2, ..., y_n)$ are vectors in $\mathbb{R}^n$ then we define the ***dot (scalar)*** product:

$$x \cdot y = x_1 y_1 + x_2 y_2 + ... + x_n y_n$$

Another notation for the dot product is $< x, y >$.
It is easy to check the properties:
(i) $x \cdot y = y \cdot x$.
(ii) $x \cdot x \geq 0,\ x \cdot x = 0$ iff $x = 0$.
(iii) $(x + y) \cdot z = (x \cdot z) + (y \cdot z)$.
(iv) $(\alpha x) \cdot y = \alpha(x \cdot y),\ \ \forall x, y, z \in \mathbb{R}^n,\ \forall \alpha \in \mathbb{R}$.
The vector space $\mathbb{R}^n$ together with the above dot product is called the n-th ***dimensional Euclidean space***.

The Cauchy-Schwarz inequality
For every $x = (x_1, x_2, ..., x_n)$ and $y = (y_1, y_2, ..., y_n) \in \mathbb{R}^n$ we have:

$$(\star) \qquad (x \cdot y)^2 \leq \left(\sum_{k=1}^{n} x_k^2 \right) \left(\sum_{k=1}^{n} y_k^2 \right)$$

Proof If $y = 0$ then no proof is needed (0=0).

Suppose that $y \neq 0$; then $\sum_{k=1}^{n}(\lambda y_k - k_k)^2 \geq 0, \forall \lambda \in \mathbb{R}$. We get:

$$\lambda^2 \sum_{k=1}^{n} y_k^2 - 2\lambda \sum_{k=1}^{n} x_k y_k + \sum_{k=1}^{n} x_k^2 \geq 0,\ \forall \lambda \in \mathbb{R}$$

By using the well known properties of the quadratic function we obtain the required inequality.

The Euclidean norm

For $x \in \mathbb{R}^n$ we define $\| x \| = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$ and call it the ***(Euclidean) norm*** of x.
The properties of the norm are $(\forall x, y \in \mathbb{R}^n,\ \forall \alpha \in \mathbb{R})$:
(i) $\| x \| \geq 0,\ \| x \| = 0$ iff $x = 0$;
(ii) $\| \alpha x \| = |\alpha|\ \| x \|$;
(iii) $\| x + y \| \leq \| x \| + \| y \|$.
Only property (iii) is not trivial to check (one can use Cauchy-Schwarz inequality).
Obviously, for $n = 1$ the euclidean norm is the same as the absolute value. The vector space $\mathbb{R}^n$ together with the norm is a ***normed vector space***.

The Euclidean distance

For $x, y \in \mathbb{R}^n$ we define the (***Euclidean***) ***distance*** by

$$d(x, y) = \| x - y \|$$

It is clear that $(\forall x, y, z \in \mathbb{R}^n)$:
(i) $d(x, y) \geq 0,\ d(x, y) = 0$ iff $x = y$.;
(ii) $d(x, y) = d(y, x)$;
(iii) $d(x, z) \leq d(x, y) + d(y, z)$ (the triangle inequality).
R^n together with the Euclidean distance is a ***metric space***.

Remark

In the metric spaces R^2 and R^3 the Euclidean distance coincide with the distance (studied at school) of analytic geometry.

Sequences in $\mathbb{R}^p$

During this section we use the letter p for the dimension of the Euclidean space, the letter n being used for sequences.

Definition

The sequence $(x_n)_n$ in R^p ***converges*** to $x \in \mathbb{R}^p$ if:

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N} \text{ s.t. if } \quad n \geq N_\varepsilon, \quad \text{then} \quad \| x_n - x \| < \varepsilon.$$

If so, then x is called the ***limit*** of the sequence $(x_n)_n$ and the sequence is said to be ***convergent*** to x. We write $\lim\limits_{n \to \infty} x_n = x$ or $x_n \longrightarrow x$.

Remark

If $x_n \longrightarrow x$ and $x_n \longrightarrow y$ then $x = y$ (if the limit exists, then it is unique).

It is not difficult to prove the following:

Proposition

(i) If $x_n \longrightarrow x$ and $y_n \longrightarrow y$ then $x_n + y_n \longrightarrow x + y$.
(ii) If $x_n \longrightarrow x$ and $\alpha_n \longrightarrow \alpha$ (in $\mathbb{R}$) then $\alpha x_n \longrightarrow \alpha x$.
We say that the operations of the vector space $\mathbb{R}^p$ are continuous.

Let us consider a sequence $(x_n)_n$ in $\mathbb{R}^p$; we have $x_n = (x_{n1}, x_{n2}, ..., x_{np})$, so the notation is a little bit clumsy. Anyway, it is clear that to give a sequence in $\mathbb{R}^p$ is equivalent to give p sequences of real numbers. To be more specific, let us take the case of $\mathbb{R}^2$; for a sequence $(x_n, y_n)_n$ there are two sequences (of real numbers) $(x_n)_n$ and $(y_n)_n$ (the first component and the second component, respectively).

Theorem

Let $(x_n)_n$ be a sequence in $\mathbb{R}^p$ and let $x \in \mathbb{R}^p$.
Then $x_n \longrightarrow x$ iff the sequences of the components of $(x_n)_n$ converge to the corresponding components of x.
For example, in $\mathbb{R}^2$, $(x_n, y_n) \longrightarrow (x, y)$ iff $x_n \to x$ and $y_n \to y$ (in $\mathbb{R}$).
Proof It's easy if one uses the inequalities:

$$|x_k| \ \leq \| x \| \leq \ |x_1| + |x_2 + ... + |x_p|, \ \forall k = 1, 2, ..., p.$$

Remark

We can say that convergence in $\mathbb{R}^p$ is ***componentwise convergence***.

Example

In $\mathbb{R}^2$ the sequence $\left(\frac{1}{n}\, ,\, \left(1+\frac{1}{n}\right)^n\right)_n$ converges to $(0, e)$ simply because $\frac{1}{n} \longrightarrow 0$ and $\left(1+\frac{1}{n}\right)^n \longrightarrow e$ (in $\mathbb{R}$).

Definition

A ***Cauchy sequence*** in $\mathbb{R}^p$ is a sequence $(x_n)_n$ s.t:

$$\forall \varepsilon > 0, \exists N_\varepsilon \in \mathbb{N} \text{ s.t. if } \ n, m \geq N_\varepsilon, \ \text{ then } \ \| x_n - x_m \| < \varepsilon.$$

It is easy to observe (by using the previous discussion) that for a sequence in $\mathbb{R}^p$ to be a Cauchy sequence is (again) a componentwise property.

Theorem

In $\mathbb{R}^p$ a sequence is convergent iff it is a Cauchy sequence (one says that the metric space $\mathbb{R}^p$ is a ***complete metric space***).

Proof The problem can be reduced to sequences in $\mathbb{R}$ (by using the componentwise philosophy).

Exercise

A sequence $(x_n)_n$ in $\mathbb{R}^p$ is said to be ***bounded*** if there exists $M > 0$ s.t. $\| x_n \| \leq M$, $\forall n \in \mathbb{N}$.

Prove that convergent sequences are bounded.

Exercise

Does the Cesaro theorem hold in $\mathbb{R}^p$?

Hint The answer is affirmative, but be carefull at the choice of the convergent subsequence of each component.

Closed sets, open sets

Definition

A set $F \subseteq \mathbb{R}^p$ is said to be ***closed*** (in $\mathbb{R}^p$) if, either $F = \emptyset$, or: if for every sequence $(x_n)_n$ in F converging to $x \in \mathbb{R}^p$, we have $x \in F$.

Loosely speaking F is closed to taking limits of sequences in F.

Example
(i) $\mathbb{R}^p$ is closed (why ?).
(ii) Every interval $[a, b] \subseteq \mathbb{R}$ is closed (in $\mathbb{R}$).
(iii) Every ***rectangle*** $[a, b] \times [c, d] \subseteq \mathbb{R}^2$ is closed.
Proof We only prove (iii); if (x_n, y_n) is a convergent sequence in F s.t. $(x_n, y_n) \longrightarrow (x, y)$ in $\mathbb{R}^2$, then $x_n \in [a, b]$, $x_n \longrightarrow x$, and $y_n \in [c, d]$, $y_n \longrightarrow y$. We get that $x \in [a, b]$ and $y \in [c, d]$, so $(x, y) \in [a, b] \times [c, d]$.

Definition
A set $D \subseteq \mathbb{R}^p$ is said to be ***open*** (in $\mathbb{R}^p$) if its complement $\mathbb{R}^p \setminus D$ is closed.

Example
(i) $\mathbb{R}^p$ is open; also $\emptyset$ is open. So, $\mathbb{R}^p$ and $\emptyset$ are both open and closed.
(ii) $D = (a, b) \subseteq \mathbb{R}$ is open (in $\mathbb{R}$).
(iii) $D = (a, b) \times (c, d) \subseteq \mathbb{R}^2$ is open (in $\mathbb{R}^2$).

Definition
Let $a \in \mathbb{R}^p$ and $r > 0$; ***the open ball*** centered at a and radius r is the set:

$$B(a, r) = \{x \in \mathbb{R}^p \ ; \ \| x - a \| < r\}$$

The ***closed ball*** centered at a and radius r is the set:

$$B'(a, r) = \{x \in \mathbb{R}^p \ ; \ \| x - a \| \leq r\}$$

We leave to the reader to see what the open and closed balls are in $\mathbb{R}$, $\mathbb{R}^2$ and $\mathbb{R}^3$. For example, in $\mathbb{R}^2$ the open balls are open disks.

In order to understand what open sets mean it is useful to prove the following:

Proposition
A non empty set $D \subseteq \mathbb{R}^p$ is open iff for every $a \in D$ there exists $r > 0$ s.t. $B(a, r) \subseteq D$.
Proof Suppose D is open and let $a \in D$. If the conclusion of the proposition would be false, then for every $n \in \mathbb{N}$ there is $x_n \in B(a, \frac{1}{n})$

s.t. $x_n \notin D$. It is clear that $x_n \in \mathbb{R}^p \setminus D$ and that $x_n \longrightarrow a$. As $\mathbb{R}^p \setminus D$ is closed, it results that $a \in \mathbb{R}^p \setminus D$, which is a contradiction. Analogously one can prove the converse.

Remark

The intuition of open sets given by the above proposition is that a point of an open set D has "around" it (in any "direction") points of D as close to it as we want.

Theorem (basic properties of open sets)

(i) $\mathbb{R}^p$ and $\emptyset$ are open.
(ii) If $(D_i)_{i \in J}$ is a family of open sets (not necessarily finite), then $\bigcup_{i \in J} D_i$ is open.
(iii) Let $n \in \mathbb{N}$; if $D_1, D_2, ..., D_n$ are open, then $D_1 \bigcup D_2 \bigcup ... \bigcup D_n$ is open.

Theorem (basic properties of closed sets)

(i) $\mathbb{R}^p$ and $\emptyset$ are closed.
(ii) If $(F_i)_{i \in J}$ is a family of closed sets, then $\bigcap_{i \in J} F_i$ is closed.
(iii) If $F_1, F_2, ..., F_n$ are closed sets, then $F_1 \bigcup F_2 \bigcup ... \bigcup F_n$ is closed.

The previous theorems are easy to prove by using the characterization of open sets (with open balls); for the closed sets, take the complements and use De Morgan laws.

Exercise

Prove that open balls are open sets and closed balls are closed sets.

Compact sets

Definition

A set $K \subseteq \mathbb{R}^p$ is ***compact*** if it is closed and bounded; remember that a set K is bounded if $\exists M > 0$ s.t. $\parallel x \parallel \leq M$, $\forall x \in K$. The empty set is compact by definition.

Example
(i) Closed balls are compact.
(ii) $\mathbb{R}^p$ is not compact.

Theorem
A set $K \subseteq \mathbb{R}^p$ is compact iff for every sequence $(x_n)_n$ in K there is a convergent subsequence $x_{n_k} \longrightarrow x \in K$.
So compactness is somehow a Cesaro-type property together with a closeness condition.
Proof Let us prove only one implication and this one for $p = 1$. Let $\emptyset \neq K \subseteq$ be a compact set and let x_n be a sequence in K; then, due to the boundness of K, the sequence $(x_n)_n$ will be bounded. By Cesaro theorem there exists a subsequence $x_{n_k} \longrightarrow x \in \mathbb{R}$; the set K being closed, $x \in K$.

Exercises
1. If $K_1, K_2, ..., K_p$ are compact subsets of $\mathbb{R}$, then the set $K_1 \times K_2 \times ... \times K_p \subseteq \mathbb{R}^p$ is compact.
Hint Use Cesaro theorem in $\mathbb{R}^p$ (see the exercise on pg.69).

2. If K_1 and K_2 are compact sets in $\mathbb{R}^p$ then $K_1 \bigcup K_2$ is compact.

3. Let $K \subseteq H \subseteq \mathbb{R}^p$; if H is compact and K is closed then K is compact.

4. Let $(x_n)_n$ be a convergent sequence in $\mathbb{R}^p$ and let a be its limit. Prove that the set $\{x_n \, ; \, n \in \mathbb{N}\} \bigcup \{a\}$ is compact (in $\mathbb{R}^p$).
Hint Prove the statement for $p = 1$, then apply exercise 1.

5. Check if the following sets are open, closed or compact:
$A = \{(x, y) \in R^2 \mid 0 < x^2 + y^2 \leq 1\}$,
$B = \{(x, y) \in R^2 \mid 1 \leq x^2 + y^2 \leq 2\}$,
$D = \{(x, y) \in R^2 \mid x^2 + y^2 < 1, x > 0, y > 0\}$,
$E = \{(x, y) \in R^2 \mid ax + by + c = 0, a, b, c \in R \; a^2 + b^2 + c^2 \neq 0\}$,
$F = \left\{ \left(\frac{1}{n}, 1 \right) \in R^2 \mid n \in N \right\}$, $G = F \cup \{(0, 1)\}$.

3.2 Continuity

We shall study functions $f : \mathbb{R}^n \mapsto \mathbb{R}^m$, or, more generally, $f : E \mapsto \mathbb{R}^m$, $E \subseteq \mathbb{R}^n$. Such functions are called ***vector-valued functions of several variables*** (more specific, "vector valued" if $m > 1$ and "real-valued" if $m = 1$).

Of great significance for the differential calculus are the ***linear functions***. Remember that a function $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is ***linear*** if:

$$f(x+y) = f(x) + f(y) \text{ and } f(\alpha x) = \alpha f(x),\ \forall x, y \in \mathbb{R}^n,\ \forall \alpha \in \mathbb{R}$$

A linear map is completely determined by the values it takes on the canonical basis of $\mathbb{R}^n$, so by the vectors $f(e_1), f(e_2), ..., f(e_n)$; in fact, if $x = (x_1, x_2, ..., x_n)$ then $x = x_1e_1 + x_2e_2 + ... + x_ne_n$, hence $f(x) = x_1f(e_1) + x_2f(e_2) + ... + x_nf(e_n)$. It is well known that to every linear function one can associate a (unique) matrix with respect to the canonical bases. Consequently, computing the values of the function reduces to multiplication of matrices.

Examples
(i) Every linear map $f : \mathbb{R} \mapsto \mathbb{R}$ has the form $f(x) = \alpha x$ for a given $\alpha \in \mathbb{R}$. We can identify α with the matrix of f (in the canonical bases).
(ii) Every linear map $f : \mathbb{R}^2 \mapsto \mathbb{R}$ has the form $f(x, y) = \alpha x + \beta y$ for given $\alpha, \beta \in \mathbb{R}$; the matrix of f is $(\alpha\ \beta)$, so $f(x, y) = (a\ b)\begin{pmatrix} x \\ y \end{pmatrix}$.

An important example of linear functions are the ***canonical projections*** of $\mathbb{R}^n$. They are denoted by $\pi_1, \pi_2, ..., \pi_n$ and defined by $\pi_k(x_1, x_2, ..., x_n) = x_k,\ k = 1, 2, ..., n$.

By using the canonical projections we can better describe a function $f : \mathbb{R}^n \mapsto \mathbb{R}^m$. In fact let $f_k = \pi_k \circ f,\ \ k = 1, 2, ..., n$; then $\forall (x_1, x_2, ..., x_n) \in \mathbb{R}^n$:

$$f(x_1, x_2, ..., x_n) = (f_1(x_1, x_2, ..., x_n), f_2(x_1, x_2, ..., x_n), ..., f_n(x_1, x_2, ..., x_n))$$

The real valued functions $f_1, f_2, ..., f_m$ are called the ***components*** of f and we write $f = (f_1, f_2, ..., f_m)$.

If we denote by $y = (y_1, y_2, ..., y_m)$ a generic point $\mathbb{R}^m$, the function f can be described (in a traditional notation) as:

$$y_1 = f_1(x_1, x_2, ..., x_n), y_2 = f_2(x_1, x_2, ..., x_n), ..., y_m = f_m(x_1, x_2, ..., x_n)$$

So vector valued functions have components and are determined by these ones. Viceversa given m functions $f_1, f_2, ..., f_m : \mathbb{R}^n \mapsto \mathbb{R}$ there is a unique function $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ s.t. $f = (f_1, f_2, ..., f_m)$. We shall see that the study of f can be sometimes reduced to the study of its components.
Of course the components of a function $f : E \mapsto \mathbb{R}^m$ are defined in the same way.

Exercise
A map $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is linear iff its components are linear.

Definition
Let $f : E \mapsto \mathbb{R}^m$, $E \subseteq \mathbb{R}^n$, $a \in E$.
The function f is ***continuous*** at a if:

$\forall \varepsilon > 0 \; \exists \delta_\varepsilon > 0$ s.t. if $x \in E$ and $\| x - a \| < \delta_\varepsilon$ then $\| f(x) - f(a) \| < \varepsilon$.

Here $\| \ \|$ is used to denote the euclidian norm both in $\mathbb{R}^n$ and $\mathbb{R}^m$.
The function f is ***continuous*** if continuous at every point of E.

An equivalent condition for continuity is:
Proposition
f is continuous at a iff for every sequence $(x_k)_k$ in E, $x_k \longrightarrow a$ one has $f(x_k) \longrightarrow f(a)$.
We admit this result without proof (it is a good exercise to try it).
Generally, the above proposition is easier to apply in concrete situations to prove continuity.

Example
Let $f : \mathbb{R}^2 \mapsto \mathbb{R}$ be $f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$
Let us prove that f is continuous at every point $(a,b) \neq (0,0)$ (here

"a" is the first component of the point) and not continuous (so discontinuous) at $(0,0)$.
Let $(a,b) \neq (0,0)$ be given and take $(x_k, y_k) \longrightarrow (a,b)$. We can suppose that $(x_k, y_k) \neq (0,0)$ for every n (why ?) Then $x_k \longrightarrow a$ and $y_k \longrightarrow b$ so $x_k y_k \longrightarrow ab$, $x_k^2 + y_k^2 \longrightarrow a^2 + b^2$ and

$$f(x_k, y_k) = \frac{x_k y_k}{x_k^2 + y_k^2} \longrightarrow \frac{ab}{a^2 + b^2}$$

by well-known results about sequences of real numbers. Now, for the case of the point $(0,0)$, remark that if $x \neq 0$, $y = tx$, $t \in \mathbb{R}$ then $f(x, tx) = \frac{t}{1+t^2}$ (the function f is constant along lines passing through the origin). So, if $x_k \longrightarrow 0$, $x_k \neq 0$ then $f(x_k, tx_k) \longrightarrow \frac{t}{1+t^2} \neq 0$ for $t \neq 0$, so f is not continuous at $(0,0)$. In fact the proof shows more: there is no value at $(0,0)$ making f continuous at $(0,0)$.

Exercise
Linear functions are continuous.

Proposition
Let $f : E \mapsto \mathbb{R}^m$, $E \subseteq \mathbb{R}^n$, $f = (f_1, f_2, ..., f_m)$ and $a \in E$. Then f is continuous at a iff $f_1, f_2, ..., f_m$ are all continuous at a. (So continuity is componentwise).
Proof Trivial using sequences and componentwise convergence.

Remark
In proving continuity we can reduce the case of vector valued functions to that of real valued functions.

Remember that for $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ and $B \subseteq \mathbb{R}^m$ we define

$$f^{-1}(B) = \{x \in \mathbb{R}^n \ ; \ f(x) \in B\}.$$

We have the following elegant description of continuity:

Theorem
$f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is continuous iff for every ***closed set*** $F \subseteq \mathbb{R}^m$, the set $f^{-1}(F)$ is ***closed***.
Proof

First, suppose f is continuous and F is closed (and non empty, the case $F = \emptyset$ being trivial). Let $(x_k)_k$ be a convergent sequence in

$f^{-1}(F)$, $x_k \longrightarrow a$ (in $\mathbb{R}^n$). Then $f(x_k) \in F$, $\forall k$ and, by continuity, $f(x_k) \longrightarrow f(a)$. But F being closed $f(a) \in F$ so $a \in f^{-1}(F)$ proving that $f^{-1}(F)$ is closed.
Conversely, let $f^{-1}(F)$ be closed (in $\mathbb{R}^n$) for every closed set $F \subseteq \mathbb{R}^m$; let $a \in \mathbb{R}^n$. if f would not be continuous at a then:
$\exists \varepsilon > 0$ s.t. $\forall \delta > 0$, $\exists x_\delta \in \mathbb{R}^n$ s.t. $\| x_\delta - a \| < \delta$ and $\| f(x_\delta) - f(a) \| \geq \varepsilon$.
Take now $F = \{y \in \mathbb{R}^m \; ; \; \| y - f(a) \| \geq \varepsilon\}$; F is closed (why?) and of course, $a \notin f^{-1}(F)$. But if we take $\delta = \frac{1}{k}$, $k \geq 1$ we can find a sequence $(x_k)_k$, $x_k = x_{\frac{1}{k}}$; then $x_k \in F$ and $x_k \longrightarrow a$. So $f^{-1}(F)$ would not be closed, which is a contradiction.

Exercise
$f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is continuous iff for every open set $D \subseteq \mathbb{R}^m$ $f^{-1}(D)$ is open (in $\mathbb{R}^n$).

We shall now consider a basic property of continuous functions with respect to compact sets.

Theorem
If $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is ***continuous*** and $K \subseteq \mathbb{R}^n$ is ***compact*** then $f(K)$ is ***compact***.
Proof Remember first that $f(K) = \{f(x) \; ; \; x \in K\} \subseteq \mathbb{R}^m$. Now let $(y_k)_k$ be a sequence in $f(K)$; we can obtain a sequence $(x_k)_k$ s.t. $x_k \in K$ and $f(x_k) = y_k$. But K being compact there is a convergent subsequence $x_{k_p} \longrightarrow x \in K$. From the continuity of f we deduce that $f(x_{k_p}) \longrightarrow f(x) \in K$. So $(y_k)_k$ has a convergent subsequence (to a point of $f(K)$) and so $f(K)$ is compact.

In particular for the case $m = 1$ we have:

Proposition
If $f : \mathbb{R}^n \mapsto \mathbb{R}$ is continuous and $K \subseteq \mathbb{R}^n$ is compact then f is ***bounded*** and attains its ***minimum*** and ***maximum*** values on K.
Proof By the ***bounds*** of f on K we mean

$$\sup_K f = \sup\{f(x) \; ; \; x \in K\} = \sup f(K)$$

$$\inf_K f = \inf\{f(x) \; ; \; x \in K\} = \inf f(K)$$

By "attains its bounds" we mean that $\sup f$ and $\inf f$ are values of f (they belong to $f(K)$). Now, for the proof, we know that $f(K)$ is a compact set so, closed and bounded. Being bounded, $\sup_K$ and $\inf_K$ are real numbers. There are sequences of $f(K)$ converging to $\sup f(K)$ and $\inf f(K)$ (why?). So $f(K)$ being closed, then $\sup f(K)$ and $\inf f(K)$ belong to $f(K)$.

Remark
(i) It is no need for f to be defined on the whole $\mathbb{R}^n$; the same result holds for f defined just on the compact set K.
(ii) The proposition is very important for the so called "optimization problems" (finding the cheapest, the least, etc). It tells us that on compact sets the optimization problem has solutions. Unfortunately it gives no method for finding (computing) them.

Exercises

1. Find the minimum and maximum values of the function $f(x, y) = \frac{xy}{x^2+y^2}$ on the unit circle $S = \{(x, y)\ ;\ x^2 + y^2 = 1\}$.
Hint One can use polar coordinates $x = \rho\cos\varphi,\ y = \rho\sin\varphi$; we get $f(\varphi) = \frac{1}{2}\sin 2\varphi$, etc.

2. If $K \subseteq \mathbb{R}$ is not compact then there is a continuous not bounded function $f : K \mapsto \mathbb{R}$.
Hint If K is not bounded, take $f(x) = x$; if K is not closed, let $a \notin K$ s.t. there is a sequence $(x_n)_n$ in K such that $x_n \longrightarrow a$ and take $f(x) = \frac{1}{|x-a|}$.

Study the continuity of the following functions (exercises 3 - 7):

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

Hint $\lim\limits_{(x,y)\to(0,0)} f(x, y) = \lim\limits_{(x,y)\to(0,0)} \frac{x^2y}{x^2+y^2} = 0$, because $|\frac{x^2y}{x^2+y^2}| \leq |x|$, so f is continuous at $(0, 0)$.

4. $f(x, y) = \begin{cases} \frac{x^2y}{x^4+y^2}, & (x, y) \neq (0, 0) \\ 0, & (x, y) = (0, 0) \end{cases}$

Hint Use the sequences: $(x_n, y_n) = (\frac{1}{n}, \frac{1}{n^2}) \to (0,0)$ and $(x'_n, y'_n) = (\frac{1}{n}, \frac{1}{n}) \to (0,0)$; it results f is not continuous at $(0,0)$. One can observe that $\lim\limits_{x\to 0} f(x, mx) = 0, \forall m \in \mathbb{R}$ and $\lim\limits_{\rho\to 0} f(\rho\cos\varphi, \rho\sin\varphi) = 0, \forall\varphi \in \mathbb{R}$.

5. $f(x,y) = \begin{cases} (x^2+y^2)\sin\frac{1}{xy}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$

Hint Using the inequality $\left|(x^2+y^2)\sin\frac{1}{xy}\right| \leq x^2+y^2$ we obtain:

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0$$

6. $f(x,y) = \begin{cases} \frac{ye^{-\frac{1}{x^2}}}{y^2+e^{-\frac{2}{x^2}}}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$

Hint Use the sequences: $(x_n, y_n) = (\frac{1}{\sqrt{\ln n}}, \frac{1}{n}) \to (0,0)$ and $(x'_n, y'_n) = (\frac{1}{\sqrt{\ln n}}, \frac{1}{n^2}) \to (0,0)$ to see f is not continuous.

7. $f(x,y) = \begin{cases} \frac{1}{xy}\sin\frac{x^3y^2}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$

8. Find a continuous not bounded function on $A = [-2,2] \setminus \{0,1\}$.

9. Find an example of a continuous function f on $\mathbb{R}$ and an open set $D \subseteq \mathbb{R}$ with $f(D)$ not open.
Hint For example $f(x) = x^2$.

10. If $f : D \subseteq \mathbb{R}^n \mapsto \mathbb{R}^m$, then by definition the ***graph*** of f is $\{(x, f(x)) \in \mathbb{R}^n \times \mathbb{R}^m \; ; \; x \in D\}$.
Give a function $f : [0,1] \mapsto \mathbb{R}$ that has a closed graph (in $\mathbb{R}^2$) but is not continuous.
Hint Take $f(x) = \ln x$ for $x \in (0,1]$ and $f(0) = 0$.

Chapter 4

Differentiable functions

4.1 Partial derivatives and the differential

Partial derivatives

A vector $v \in \mathbb{R}^n$ is called a ***direction*** if $\| v \| = 1$.

Examples
(i) In $\mathbb{R}$ there are only two directions: 1 and -1.
(ii) In $\mathbb{R}^2$ the directions can be identified with the points of the unit circle $x^2 + y^2 = 1$.
(iii) In $\mathbb{R}^3$ the directions can be identified with the points of the unit sphere $x^2 + y^2 + z^2 = 1$.
(iv) In $\mathbb{R}^n$ the vectors $e_1, e_2, ...e_n$ of the canonical basis are directions (sometimes called the "positive" directions of the coordinate axes).

Let now $\Omega \subseteq \mathbb{R}^n$ be a non empty open set, $f : \Omega \mapsto \mathbb{R}$ and $a \in \Omega$. Remark that for ***any direction*** v and "enough small" $t \in \mathbb{R}$ ($|t| < \varepsilon$), then $a + tv \in \Omega$, so $f(a + tv)$ is well-defined.

Defintion
We say that f is ***differentiable at*** a ***in the direction*** v if the limit:

$$\lim_{t \to 0} \frac{f(a + tv) - f(a)}{t}$$

exists and is finite (it's a real number). If this is the case then we denote this limit by $\frac{df}{dv}(a)$ and call it the ***derivative of f in the direction v at a*** (***directional derivative***). So:

$$\lim_{t\to 0}\frac{f(a+tv)-f(a)}{t}=\frac{df}{dv}(a)$$

Intuitively we can consider the function $t\mapsto f(a+tv)$ as the restriction of f to the line passing through a and of direction v, giving information about the variation of f along the mentioned line.

The derivatives of f in the directions $e_1, e_2, ..., e_n$ (if they exist) are called ***partial derivatives*** and we shall use the notations

$$\frac{df}{de_k}(a)=\frac{\partial f}{\partial x_k}(a)$$

(an exception in the case of $\mathbb{R}$ for which the well-known notation $\frac{df}{dx}(a)$, or $f'(a)$ will be used). So we have:

$$\frac{\partial f}{\partial x_k}(a)=\lim_{t\to 0}\frac{f(a+te_k)-f(a)}{t}=\lim_{t\to 0}\frac{f(a_1,...,a_k+t,a_{k+1},...,a_n)}{t}$$

We see that $\frac{\partial f}{\partial x_k}(a)$ is the derivative of the one single variable $x_k\mapsto f(a_1,a_2,...,x_k,a_{k+1},...,a_n)$ at the point a_k. ***So partial derivatives follow the usual rules of computing derivatives***.

If $\frac{\partial f}{\partial x_k}(a)$ exists for every $a\in\Omega$ we say that $\frac{\partial f}{\partial x_k}$ exists on Ω; if $\frac{\partial f}{\partial x_1},\frac{\partial f}{\partial x_2},...,\frac{\partial f}{\partial x_n}$ exist on Ω we say that f has (first order) partial derivatives on Ω.

Example

(i) Let $f:\mathbb{R}^2\mapsto\mathbb{R}$ be given by $f(x,y)=\begin{cases}\frac{xy}{x^2+y^2} & \text{if } (x,y)\neq(0,0)\\ 0 & \text{if } (x,y)=(0,0)\end{cases}$

Then $f(x,0)=0=f(0,y)$, so $\frac{\partial f}{\partial x}(0,0)=0=\frac{\partial f}{\partial y}(0,0)$; f has partial derivatives at $(0,0)$ but we know (see the previous section) that f is not continuous at $(0,0)$. If we want to compute the partial derivatives at a point $(x,y) \neq (0,0)$ we can use the usual rules of differentiation: for $\dfrac{\partial f}{\partial x}$ keep y constant and compute the derivative with respect to x, etc.

$$\frac{\partial f}{\partial x}(x,y) = y\,\frac{x^2+y^2-2x^2}{(x^2+y^2)^2} = \frac{y(y^2-x^2)}{(x^2+y^2)^2},$$
$$\frac{\partial f}{\partial y}(x,y) = \frac{x(x^2-y^2)}{(x^2+y^2)^2},$$

for $(x,y) \neq (0,0)$. We obtain the partial derivatives of f:

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \frac{y(y^2-x^2)}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} \frac{x(x^2-y^2)}{(x^2+y^2)^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

(ii) For the same function, take a direction $v = (\cos\theta, \sin\theta)$.
Then $f(0+tv) = \cos\theta\,\sin\theta,\ t \neq 0$; we obtain:

$$\lim_{t\to 0}\frac{f(0+tv)-f(0)}{t} = \lim_{t\to 0}\frac{\cos\theta\,\sin\theta}{t},$$

so $\dfrac{df}{dv}(0,0)$ exists only for the directions $(1,0),(0,1),(-1,0),(0,-1)$, being 0 in these cases.

The differential

We first recall the definition of the derivative in the one variable case and restate this definition in a suitable way for an extension to several variables.

Let $I \subseteq \mathbb{R}$ be an open interval, $a \in I$ and $f : I \mapsto \mathbb{R}$. We know that f is ***differentiable*** at a if the limit:

$$\lim_{x\to a}\frac{f(x)-f(a)}{x-a}$$

exists and is finite (a real number). In this case the limit is the ***derivative*** $\left(\frac{df}{dt}(a) \text{ or } f'(a) \right)$ of f at a.
Now put $x = a + h$:

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = f'(a)$$

Equivalently:

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - f'(a)h}{h} = 0$$

The ***linear*** function $\mathbb{R} \ni h \mapsto f'(a)h$ is called the ***differential*** of f at a. This way we can restate the definition of differentiability of f at a as follows:

The function f is ***differentiable*** at a iff there exists a ***linear map*** $\lambda : \mathbb{R} \mapsto \mathbb{R}$ s.t:

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

The connection between the differential and the derivative is very simple (and elegant): $f'(a)$ is (identified to) the matrix of λ in the canonical basis of $\mathbb{R}$ (which is 1).

In the last form the definition of the differential can be extended to several variables as follows:

Definition

Let $f : \Omega \mapsto \mathbb{R}^m$, $\Omega \subseteq \mathbb{R}^n$, open and $a \in \Omega$. f is said to be ***differentiable*** at a if there exists a linear map $\lambda : \mathbb{R}^n \mapsto \mathbb{R}^m$ s.t.

$$(\star) \qquad \lim_{h \to 0} \frac{\| f(a+h) - f(a) - \lambda(h) \|}{\| h \|} = 0$$

Both norms in $\mathbb{R}^n$ and $\mathbb{R}^m$ were denoted by $\| \ \|$.

Remark

In $\mathbb{R}^n$, $h \longrightarrow 0$ iff $\| h \| \longrightarrow 0$.

Going on with the definition, call λ the ***differential*** of f at a (it will be proved to be ***unique*** if any) and denote it by $Df(a)$; the matrix of $Df(a)$ in the canonical bases is denoted by $f'(a)$ (it is an $m \times n$ matrix) and it is called the ***Jacobian matrix*** of f at a. In the case $m = n$ the determinant $\det f'(a)$ of $f'(a)$ is called the ***Jacobian*** of f at a and is denoted by $J_f(a)$. If f is differentiable at every $a \in \Omega$ we say that f is differentiable on Ω.

Remark

The notation $Df(a)$ is somehow clumsy as $Df(a)$ is a (linear) function. So $Df(a) : \mathbb{R}^n \mapsto \mathbb{R}^m$ and $Df(a)(x) \in \mathbb{R}^m$, $\forall x \in \mathbb{R}^n$.

Proposition

If the linear functions λ and μ satisfy $(\star)$ for f at a then $\lambda = \mu$. So, the differential at a (if exists) is unique.

Proof Let $\Delta f(h) = f(a+h) - f(a)$; then:

$$\frac{\| \lambda(h) - \mu(h) \|}{\| h \|} \leq \frac{\| \lambda(h) + \Delta f(h) - \Delta f(h) - \mu(h) \|}{\| h \|} \leq$$

$$\leq \frac{\| \Delta f(h) - \lambda(h) \|}{\| h \|} + \frac{\Delta f(h) - \mu(h) \|}{\| h \|}$$

So $\lim\limits_{h \to 0} \dfrac{\| \lambda(h) - \mu(h) \|}{\| h \|} = 0$. Now take $x \in \mathbb{R}^n$, $x \neq 0$ and a sequence $t_n \longrightarrow 0$, $t_n \neq 0$ in $\mathbb{R}$; of course $t_n x \longrightarrow 0$ in $\mathbb{R}^n$ so

$$\lim_{n \to \infty} \frac{\| \lambda(t_n x) - \mu(t_n x) \|}{\| t_n x \|} = \lim_{n \to \infty} \frac{|t_n| \ \| \lambda(x) - \mu(x) \|}{|t_n| \ \| x \|} =$$

$$= \lim_{n \to \infty} \frac{\| \lambda(x) - \mu(x) \|}{\| x \|} = 0$$

(the linearity of λ and μ was used). We obtain that $\| \lambda(x) - \mu(x) \| = 0$, so $\lambda(x) = \mu(x), \forall x \in \mathbb{R}^n$, $x \neq 0$; but $\lambda(0) = \mu(0) = 0$, so $\lambda = \mu$.

By using ($\star$) and the fact that linear functions are continuous we easily obtain:

Proposition
If f is differentiable at a, then it is continuous at a.

Remark
We could think about ($\star$) as a possibility of approximating f, around a, by the much more simpler function of type $f(a)+\lambda$ (affine function). The sense of the approximation is exactly ($\star$): the difference between f and $f(a)+\lambda$ tends to zero faster than h.

Examples
(i) Let $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a ***constant function***, $f(x) = c$, $\forall x \in \mathbb{R}^n$. Then f is differentiable on $\mathbb{R}^n$ and $Df(a) = 0$, $\forall a \in \mathbb{R}^n$; here 0 is the null (linear) function.
In fact it is enough to check ($\star$): $f(a+h) - f(a) - 0(h) = 0, \forall h \in \mathbb{R}^n$, so ($\star$) is true.
(ii) Let $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ be ***linear***. Then f is differentiable on $\mathbb{R}^n$ and $Df(a) = f$, $\forall a \in \mathbb{R}^n$.
In fact $f(a+h) - f(a) - f(h) = 0, \forall h \in \mathbb{R}^n$, etc.
(iii) Take $s : \mathbb{R}^2 \mapsto \mathbb{R},\ s(x, y) = x + y$. Then

$$Ds(a, b) = s,\ \forall (a, b) \in \mathbb{R}^2.$$

In fact s is linear.
(iv) Consider $p : \mathbb{R}^2 \mapsto \mathbb{R},\ p(x, y) = xy$. Then:

$$Dp(a, b)(x, y) = bx + ay,\ \forall (a, b) \in \mathbb{R}^2.$$

(the Jacobian matrix of p at (a, b) is $(b\ a)$):

$$p(a+h, b+k) - p(a, b) - bh - ak = (a+h)(b+k) - ab - bh - ak = hk,$$

so replacing in ($\star$) we get:

$$\lim_{(h,k)\to(0,0)} \frac{|hk|}{\sqrt{h^2 + k^2}} = 0,$$

because of the inequality: $|hk| \le h^2 + k^2$.

Basic properties of the differential

We shall admit, without proof the following:

Theorem (chain rule)

Suppose $\mathbb{R}^n \overset{f}{\mapsto} \mathbb{R}^m \overset{g}{\mapsto} \mathbb{R}^p$ s.t.:
f is differentiable at $a \in \mathbb{R}^n$ and g is differentiable at $f(a)$.
Then $g \circ f$ is differentiable at a and

$$D(g \circ f)(a) = Dg(f(a)) \circ Df(a)$$

In the language of Jacobian matrices: $(g \circ f)'(a) = g'(f(a)) \cdot f'(a)$.

The theorem says that the ***composition*** of differentiable functions is differentiable and that the differential of the composition is the composition of the differentials. The result seems quite natural and elegant (for a proof see [4]). The relation between the Jacobian matrices generalizes the well-known result for $m = n = p = 1$.

We shall apply the chain rule to prove:

Proposition

Let $f : \mathbb{R}^n \mapsto \mathbb{R}^m$, $f = (f_1, f_2, ..., f_m)$ and $a \in \mathbb{R}^n$. Then f is differentiable at a iff $f_1, f_2, ..., f_m$ are differentiable at a. In this case $Df(a) = (Df_1(a), Df_2(a), ..., Df_m(a))$ (the components of the differential are the differentials of the components). In terms of Jacobian matrices: the lines of $f'(a)$ are (with obvious identifications) the Jacobian matrices of the components.

Proof The proposition is highly plausible due to the fact the limits are componentwise. We restrict ourselves to prove one half of the result; so, we suppose f to be differentiable at a. If π_k are the canonical projections, then $f_k = \pi_k \circ f$, $\forall k = 1, 2, ..., m$. The canonical projections are linear and so differentiable at every point, so using the chain rule we obtain that the components f_k are differentiable, $\forall k = 1, 2, ..., m$. Moreover, $Df_k(a) = D\pi_k(f(a)) \circ Df(a)$; but (see a previous example) $D\pi_k(f(a)) = \pi_k$, so $Df_k = \pi_k \circ Df(a)$, which means that the components of the differential are the differentials of the components. The translation to Jacobian matrices is trivial.

A connection between the differential and the partial derivatives is given in the following:

Theorem

Let $f : \mathbb{R}^n \mapsto \mathbb{R}^m,\ f = (f_1, f_2, ..., f_m$ be differentiable at $a \in \mathbb{R}^n$. Then the partial derivatives $\dfrac{\partial f_i}{\partial x_j}(a)$ exist $\forall i = 1, 2, ..., m,\ j = 1, 2, ..., n$ and the Jacobian matrix of f at a is:

$$f'(a) = \left(\frac{\partial f_i}{\partial x_j}\right)_{1 \le i \le m\ ,\ 1 \le j \le n}$$

Proof Due to the previous proposition it will be enough to take the case $m = 1$; denote $\lambda = Df(a)$. If (general form of linear maps) $\lambda(x_1, x_2, ..., x_n) = \alpha_1 x_1 + \alpha_2 x_2 + ... + \alpha_n x_n$ for fixed $\alpha_1, \alpha_2, ..., \alpha_n$, then we have to prove that $\alpha_j = \frac{\partial f}{\partial x_j}(a),\ \forall j = 1, 2, ..., n$. From the very definition of the differentiability we have (for a given j):

$$\lim_{h_j \to 0} \frac{|f(a_1, a_2, ..., a_j + h_j, a_{j+1}, ..., a_n)|}{|h_j|} = 0,$$

where $h = (0, 0, ..., h_j, 0, ...0)$. But this exactly means $\alpha_j = \frac{\partial f}{\partial x_j}(a)$, etc.

Remark

The theorem is to be applied as follows: if you are asked about the differentiability of a function at a point then first you check the existence of the partial derivatives of all components with respect to all variables. If at least one of these derivatives does not exist then f is not differentiable. If all of them exist then we do not know yet if f is differentiable but we have a ***candidate*** for the differential, namely the linear map having the matrix of partial derivatives and we can apply the definition to see what is the case.

Example

Let $f : \mathbb{R}^2 \mapsto \mathbb{R}, f(x, y) = \begin{cases} (x^2 + y^2)\sin\frac{1}{x^2+y^2} & \text{if} \quad (x, y) \ne (0, 0) \\ 0 & \text{if} \quad (x, y) = (0, 0) \end{cases}$

We want to prove that f is differentiable at $(0, 0)$.

First we compute the partial derivatives at the origin:

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{x \to 0} \frac{f(x, y) - f(0, 0)}{x} = \lim_{x \to 0} \frac{x^2 \sin\frac{1}{x^2}}{x} = 0,$$

and by symmetry $\frac{\partial f}{\partial y}(0,0) = 0$. So the candidate for the differential at $(0,0)$ is the constant null function. We have:

$$\lim_{(x,y)\to(0,0)} \frac{f(x,y) - f(0,0) - 0}{\sqrt{x^2+y^2}} =$$

$$= \lim_{(x,y)\to(0,0)} \sqrt{x^2+y^2} \sin \frac{1}{x^2+y^2} = 0,$$

so f is differentiable at $(0,0)$ and $Df(0,0) = 0$.

Proposition
Let $f : \Omega \mapsto \mathbb{R}$, Ω a non empty open subset in $\mathbb{R}^n$ and $a \in \Omega$. If f is differentiable at a then for every direction $v = (v_1, v_2, ..., v_n)$, f is differentiable in the direction v at a and:

$$\frac{df}{dv}(a) = v_1 \frac{\partial f}{\partial x_1}(a) + v_2 \frac{\partial f}{\partial x_2}(a) + ... + v_n \frac{\partial f}{\partial x_n}(a)$$

Proof Let $g(t) = f(a + tv)$ be defined in a neighborhood of $t = 0$. As f is differentiable at a, it results that g is differentiable at 0. By the very definition, it results that f is differentiable in the direction v at a and (using the chain rule):

$$\frac{\partial f}{\partial v}(a) = g'(0) = v_1 \frac{\partial f}{\partial x_1}(a) + v_2 \frac{\partial f}{\partial x_2}(a) + ... + v_n \frac{\partial f}{\partial x_n}(a)$$

The vector

$$\left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), ..., \frac{\partial f}{\partial x_n}(a) \right)$$

is called the ***gradient*** of f at a and is usually denoted by $\text{grad}_a f$.
So the above formula can be written by using the dot product as

$$\frac{df}{dv} = (\text{grad}_a f) \cdot v$$

Remark
For a differentiable function $f : \mathbb{R}^n \mapsto \mathbb{R}$ at $\in \mathbb{R}^n$ we can write:

$$Df(a)(x) = \left(\frac{\partial f}{\partial x_1}(a), ..., \frac{\partial f}{\partial x_n}(a)\right) \begin{pmatrix} x_1 \\ . \\ . \\ . \\ x_n \end{pmatrix} =$$

$$= \frac{\partial f}{\partial x_1}(a)x_1 + \frac{\partial f}{\partial x_2}(a)x_2 + ... + \frac{\partial f}{\partial x_n}(a)x_n.$$

It is a tradition (in differential calculus) to use the notation dx_k for the canonical projection π_k, so we can rewrite:

$$Df(a)(x) = \frac{\partial f}{\partial x_1}(a)dx_1(x) + \frac{\partial f}{\partial x_2}(a)dx_2(x) + ... + \frac{\partial f}{\partial x_n}(a)dx_n(x),$$

or simply:

$$Df(a) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a)dx_k$$

Generally, the existence of partial derivatives at a point is not enough to assure differentiability (it does not assure even continuity !) But we have in this respect:

Theorem
Let $f : \mathbb{R}^n \mapsto \mathbb{R}^m,\ f = (f_1, f_2, ..., f_m),\ a \in \mathbb{R}^n$ be such that the partial derivatives $\dfrac{\partial f_i}{\partial x_j}$ exist in a neighborhood (open ball centered at a) of $a,\ \forall i = 1, 2, ..., n,\ \forall j = 1, 2, ..., m$ and suppose they are continuous at a. Then f is differentiable at a.
Proof Again we can consider only the case $m = 1$. The candidate for the differential is the linear function of matrix $\left(\dfrac{\partial f}{\partial x_1}(a), ..., \dfrac{\partial f}{\partial x_n}(a)\right)$.
The idea is to use the trick:

$$f(a_1 + h_1, a_2 + h_2, ..., a_n + h_n) - f(a_1, a_2, ..., a_n) =$$

$$= f(a_1 + h_1, a_2, ..., a_n) - f(a_1, a_2, ..., a_n)+$$

$$+f(a_1+h_1,a_2+h_2,a_3,...,a_n)-f(a_1+h_1,a_2,...,a_n)+$$

$$..$$

$$+f(a_1+h_1,a_2+h_2,...,a_n+h_n)-f(a_1+h_1,...,a_{n-1}+h_{n-1},a_n)=$$

$$=\frac{\partial f}{\partial x_1}(\xi_1,a_2,...,a_n)h_1+\frac{\partial f}{\partial x_2}(a_1+h_1,\xi_2,...,a_n)h_2+...+\frac{\partial f}{\partial x_n}(a_1+h_1,...,\xi_n)h_n,$$

by Lagrange mean theorem (with appropriate $\xi_1,\xi_2,...,\xi_n$). By using the last relation in the definition of the differentiability, we get:

$$\frac{\left|\frac{\partial f}{\partial x_1}(\xi_1,a_2,...,a_n)h_1+...+\frac{\partial f}{\partial x_n}(a_1+h_1,...,\xi_n)h_n-\frac{\partial f}{\partial x_1}(a)h_1-...-\frac{\partial f}{\partial x_n}(a)h_n\right|}{\sqrt{h_1^2+h_2^2+...+h_n^2}}\leq$$

$$\leq\left|\frac{\partial f}{\partial x_1}(\xi_1,a_2,...,a_n)-\frac{\partial f}{\partial x_1}(a)\right|+...+\left|\frac{\partial f}{\partial x_n}(a_1+h_1,...,\xi_n)-\frac{\partial f}{\partial x_n}(a)\right|,$$

we used $\frac{|h_i|}{\|h\|}\leq 1$; but the last sum of absolute values tends to zero as $h\longrightarrow 0$ due to the continuity at a of the partial derivatives.

Computing partial derivatives of composite functions

Using the chain rule and the form of the Jacobian matrix described above we obtain the basic rule for computing partial derivatives of composite functions.

Proposition
Suppose $f:\mathbb{R}^n\mapsto\mathbb{R}^m$, $f=(f_1,f_2,...,f_m)$, $g:\mathbb{R}^m\mapsto\mathbb{R}$ be s.t. f is differentiable at $a\in\mathbb{R}^n$ and g is differentiable at $f(a)$. Let the variables be $x=(x_1,x_2,...,x_n)\in\mathbb{R}^n$ and $y=(y_1,y_2,...,y_m)\in\mathbb{R}^m$. Then if $F=g\circ f$ one has:

$$(\star)\ \frac{\partial F}{\partial x_i}(a)=\frac{\partial g}{\partial y_1}(f(a))\frac{\partial f_1}{\partial x_i}(a)+\frac{\partial g}{\partial y_2}(f(a))\frac{\partial f_2}{\partial x_i}(a)+...+\frac{\partial g}{\partial y_m}(f(a))\frac{\partial f_m}{\partial x_i}(a)$$

Proof F will be differentiable at a by the chain rule.
Moreover $F'(a)=g'(f(a))\cdot f'(a)$. So we get:

$$\left(\frac{\partial F}{\partial x_1}(a),...,\frac{\partial F}{\partial x_i}(a),...,\frac{\partial F}{\partial x_n}(a)\right)=$$

$$= \left(\frac{\partial g}{\partial y_1}(f(a)), ..., \frac{\partial g}{\partial y_m}(f(a)) \right) \left(\begin{array}{ccc} \frac{\partial f_1}{\partial x_1}(a) & ..... & \frac{\partial f_1}{\partial x_n}(a) \\ ..... & ..... & ..... \\ \frac{\partial f_m}{x_1}(a) & ..... & \frac{\partial f_m}{\partial x_n}(a) \end{array} \right)$$

We trivially obtain $(\star)$ from the matrices multiplication rule.

Example

Let $F(x, y) = g(xy, x+y)$, where $g : \mathbb{R}^2 \mapsto \mathbb{R}$ is a differentiable function (all over $\mathbb{R}^2$). Denote the variables of g by (u, v); then

$$\frac{\partial F}{\partial x} = \frac{\partial g}{\partial u} y + \frac{\partial g}{\partial v} \ , \ \frac{\partial F}{\partial y} = \frac{\partial g}{\partial u} x + \frac{\partial g}{\partial v},$$

where we omitted to specify the (corresponding) points.

The inverse function theorem

Let $\Omega \subseteq \mathbb{R}^n$ be a non empty open set and $f : \Omega \mapsto \mathbb{R}$ be a function. We say that f is of class $\mathcal{C}^1$ on Ω and we write $f \in \mathcal{C}^1(\Omega)$ if all the (first order) partial derivatives of f exist on Ω and are continuous. More generally, a function $f : \Omega \mapsto \mathbb{R}^m$ is of class $\mathcal{C}^1$ if all its components are of class $\mathcal{C}^1$. Using this language we can rephrase a previous theorem by saying that $\mathcal{C}^1$ functions are differentiable at all points of the domain of definition.

If $a \in \mathbb{R}^n$ we define a ***neighborhood*** of a as being an open set U s.t. there is an open ball $B(a, r) \subseteq U$.

The following theorem is the central result of the differential calculus showing the "power" of defining the differential as a linear function.

Theorem (of the inverse function)

Suppose $f : \mathbb{R}^n \mapsto \mathbb{R}^n$, $a \in \mathbb{R}^n$, s.t. f is $\mathcal{C}^1$ in a neighborhood of a and suppose $Df(a)$ is an isomorphism (a linear bijection). Then there are neighborhoods U of a and V of $f(a)$ s.t.the restriction of f to U is bijection on V and the inverse function of this bijection is of class $\mathcal{C}^1$. So if $Df(a)$ has an inverse so does f locally around a. This result is strong enough because it is quite easy to check the bijectivity of $Df(a)$ (it is enough to see that the determinant of $f'(a)$ is not zero) but, generally, it is very difficult to check the local invertibility of a function. We do not prove this theorem (see for ex [5], [8]).

Exercises

1. Compute, by using the definition the partial derivatives of $f : R^2 \mapsto R,\ f(x,y) = 2x^3y - e^{x^2}$ at $(-1,1)$.
Hint We have:

$$\frac{\partial f}{\partial x}(-1,1) = \lim_{x\to -1} \frac{f(x,1) - f(-1,1)}{x+1} = \lim_{x\to -1} \frac{2x^3 - e^{x^2} + 2 + e}{x+1} = 6+2e.$$

$$\frac{\partial f}{\partial y}(-1,1) = \lim_{y\to 1} \frac{f(-1,y) - f(-1,1)}{y-1} = -2.$$

2. Study the existence of the partial derivatives at the origin of the functions: $f(x,y) = \sqrt{x^2+y^2}$ and $g(x,y) = x\sqrt{x^2+y^2}$.
Hint f has not partial derivatives at the origin, while g does.

3. Compute the derivative in the direction $v = \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{3}\right)$ of the function $f(x,y,z) = x^2 - yz$ at $(1,1,2)$.
Hint By the definition we have:

$$\frac{df}{dv}(1,1,2) = \lim_{t\to 0} \frac{f\left((1,1,2) + t(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})\right) - f(1,1,2)}{t} =$$

$$= \lim_{t\to 0} \frac{-1 - \frac{4}{3}t - \frac{1}{3}t^2 - (-1)}{t} = -\frac{4}{3}.$$

We can compute also by using the gradient:

$$\frac{df}{dv}(1,1,2) = \left(\mathrm{grad}_{(1,1,2)} f\right) \cdot v = (2,-2,-1) \cdot \left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right) = -\frac{4}{3}.$$

4. Let $f(x,y) = \begin{cases} (x^2+y^2)\sin\frac{1}{x^2+y^2} & \text{if} \quad (x,y) \neq (0,0) \\ 0 & \text{if} \quad (x,y) = (0,0) \end{cases}$
Is f of class $\mathcal{C}^1$ on $\mathbb{R}^2$?

5. Is the following function differentiable at the origin?

$$f(x,y) = \begin{cases} \frac{xy^2}{\sqrt{x^2+y^4}} & \text{if} \quad (x,y) \neq (0,0) \\ 0 & \text{if} \quad (x,y) = (0,0) \end{cases}$$

6. Let $f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^4} & \text{if} \quad (x,y) \neq (0,0) \\ 0 & \text{if} \quad (x,y) = (0,0) \end{cases}$

(i) Compute the partial derivatives.
(ii) Study the continuity.

7. Study the continuity and compute the partial derivatives (where they exist) for the function: $g(x,y) = \begin{cases} e^{-\left(\frac{x^2}{y^2}+\frac{y^2}{x^2}\right)} & \text{if} \quad xy \neq 0 \\ 0 & \text{if} \quad xy = 0 \end{cases}$

8. Let $f(x,y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & \text{if} \quad (x,y) \neq (0,0) \\ 0 & \text{if} \quad (x,y) = (0,0) \end{cases}$

(i) Compute the partial derivatives.
(ii) Is f of class $\mathcal{C}^1$?
(iii) Study the differentiability.

9. Same questions for $f(x,y) = \begin{cases} \frac{xy^2}{x^2+y^2} & \text{if} \quad (x,y) \neq (0,0) \\ 0 & \text{if} \quad (x,y) = (0,0) \end{cases}$

10. A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is called ***homogenous of order*** α ($\alpha \in \mathbb{R}$) iff $\forall x \in \mathbb{R}^n,\ \forall t \in \mathbb{R},\ t > 0$, we have $f(tx) = t^\alpha f(x)$.
Prove that if f is ***homogenous of order*** α then the following ***Euler's formula*** holds:

$$\alpha\, f(x_1, x_2, ..., x_n) = x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} + ... + x_n \frac{\partial f}{\partial x_n}$$

Hint Let $u_1 = tx_1, u_2 = tx_2, ..., u_n = tx_n$.
By differentiating the relation $f(u_1, u_2, ..., u_n) = t^\alpha\, f(x_1, x_2, ..., x_n)$ with respect to t we get:

$$\frac{\partial f}{\partial x_1}(u_1, ..., u_n)\frac{\partial u_1}{\partial x_1} + ... + \frac{\partial f}{\partial x_n}(u_1, ..., u_n)\frac{\partial u_n}{\partial x_n} = \alpha\, t^{\alpha-1}\, f(x_1, ..., x_n)$$

Or, equivalently ($\frac{\partial u_1}{\partial t} = x_1$, etc):

$$x_1 \frac{\partial f}{\partial x_1}(u_1, ..., u_n) + ... + x_n \frac{\partial f}{\partial x_n}(u_1, ..., u_n) = \alpha\, t^{\alpha-1}\, f(x_1, ..., x_n)$$

In the last relation put $t = 1$ and get the result.

11. Let $f : \mathbb{R}^2 \mapsto \mathbb{R}$ be differentiable at $(a, b) \in \mathbb{R}^2$. Find the directions $v \in \mathbb{R}^2$ s. t.

$$\frac{df}{dv}(a,b) = \inf\left\{\frac{df}{du}(a,b)\ ;\ u \text{ direction in } \mathbb{R}^2\right\}$$

Same question for sup. Generalize to $\mathbb{R}^n$.
Hint Use Cauchy-Schwarz inequality ,etc.

4.2 Local extrema

Higher order partial derivatives

Let $f : \Omega \mapsto \mathbb{R}$ be a function defined on a nonempty open set $\Omega \subseteq \mathbb{R}^2$. Suppose $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ to exist on Ω. Then obviously $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ define functions from Ω to $\mathbb{R}$ so we can ask if they (at their turn) have partial derivatives, etc. Define the ***partial derivatives of second order*** by

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}\ ,\ \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}\ ,\ \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2}$$

(read "d two over d x square", etc).
If these derivatives exist, we can go on by defining higher order derivatives, etc. Of course we can extend this discussion to functions of more than two variables in an obvious manner.
Sometimes it is easier to use the following notations:

$$\frac{\partial f}{\partial x} = f'_x,\ \frac{\partial f}{\partial y} = f'_y$$

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

$f''_{x\,y}$ and $f''_{y\,x}$ are called ***mixed*** partial derivatives and they are not necessary equal (the order of taking derivatives could matter). The following theorem gives conditions for the equality of the mixed partial derivatives.

Theorem
Let $f''_{x\,y}$ and $f''_{y\,x}$ exist in a neighborhood of the point $(a, b) \in \Omega$ and suppose they are continuous at (a, b). Then $f''_{x\,y} = f''_{y\,x}$.
Proof We fix an open disk centered at (a, b) where the mixed partial derivatives exist and let (x, y) be a point in this disk s.t. $x \neq a,\ y \neq b$. Consider:

$$R(x, y) = \frac{f(x, y) - f(x, b) - f(a, y) + f(a, b)}{(x - a)(y - b)}$$

Now define $\varphi(t) = \frac{f(t,y)-f(t,b)}{y-b}$ and apply to φ the Lagrange mean value theorem on $[a, x]$ or $[x, a]$ as $a < x$ or $x < a$). There exist ξ between a and x s.t:

$$\frac{\varphi(x) - \varphi(a)}{x - a} = \varphi'(\xi)$$

But $\frac{\varphi(x)-\varphi(a)}{x-a} = R(x, y)$ and $\varphi'(\xi) = \frac{f'_x(\xi,y)-f'_x(\xi,b)}{y-b}$, so

$$R(x, y) = \frac{f'_x(\xi, y) - f'_x(\xi, b)}{y - b}$$

Apply again the Lagrange theorem mean value theorem to the function $\psi(u) = f'_x(\xi, u)$ and find η between b and y s.t:

$$R(x, y) = \psi'(\eta) = f''_{xy}(\xi, \eta)$$

Changing the role of x and y and using the same method we can find ξ' between a and x and η' between b and y s.t:

$$R(x, y) = f''_{yx}(\xi', \eta'),$$

so $f''_{xy}(\xi, \eta) = f''_{yx}(\xi', \eta')$.
Now take a sequence $(x_n, y_n) \longrightarrow (a, b),\ x_n \neq a,\ y_n \neq b, \forall n \in \mathbb{N}$.
We obtain that

$$f''_{xy}(\xi_n, \eta_n) = f''_{yx}(\xi'_n, \eta'_n)$$

for convenient ξ_n, ξ'_n, η_n, η'_n as before.
It is clear that $(\xi_n, \eta_n) \longrightarrow (a, b)$, $(\xi'_n, \eta'_n) \longrightarrow (a, b)$ and by using the continuity of f''_{xy} and f''_{yx} at (a, b) we obtain $f''_{xy}(a, b) = f''_{yx}(a, b)$.

This theorem can be extended to more than two variables and to higher order partial derivatives without any difficulty. We shall use the following terminology:

If $f : \Omega \mapsto \mathbb{R}$, Ω an open non empty set of $\mathbb{R}^n$ we say that f is of class $\mathcal{C}^2$ on Ω if all partial derivatives of second order of f exist on Ω and they are continuous. One can define functions of class $\mathcal{C}^k$; the class $\mathcal{C}^\infty$ will be the intersection of all $\mathcal{C}^k$, $\forall k \in \mathbb{N}$. It follows that the order of taking the derivatives (up to order k, for a function in class $\mathcal{C}^k$, $k \geq 2$) is immaterial.
The general symbolism for partial derivatives in the case of independence of the order is the following: take $(\alpha_1, \alpha_2, ..., \alpha_n)$ a n-tuple of natural numbers and consider $\dfrac{\partial^{\alpha_1+\alpha_2+...+\alpha_n} f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} ... \partial x_n^{\alpha_n}}$ meaning that we take partial derivatives of f α_1-times with respect to x_1, α_2-times with respect to x_2 and so on (if they exist).

Exercise
Let $u : \mathbb{R}^2 \mapsto \mathbb{R}$ be a function of class $\mathcal{C}^2$ on $\mathbb{R}^2$ and let
$f(x, y) = u(xy, x + y)$. Compute $\frac{\partial^2 f}{\partial x \partial y}$.
Solution Let denote by (u, v) the variables of g; we first compute (we omit the point):

$$\frac{\partial f}{\partial y} = x\frac{\partial g}{\partial u} + \frac{\partial g}{\partial v}$$

Now (we use that g is of class $\mathcal{C}^2$):

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}\left(x\frac{\partial g}{\partial u} + \frac{\partial g}{\partial v}\right) =$$

$$= \frac{\partial g}{\partial u} + y\frac{\partial^2 g}{\partial u^2} + \frac{\partial^2 g}{\partial v \partial u} + y\frac{\partial^2 g}{\partial u \partial v} + \frac{\partial^2 g}{\partial v^2} = \frac{\partial g}{\partial u} + y\frac{\partial^2 g}{\partial u^2} + (1+y)\frac{\partial^2 g}{\partial u \partial v} + \frac{\partial^2 g}{\partial v^2}$$

Taylor formulas

Let us now prove a Taylor formula for functions of several variables which will be useful in the discussions of local extrema. We shall limit ourselves to $\mathcal{C}^2$ functions.

Definition
If $a, b \in \mathbb{R}^n$ define the (line) segment $[a, b] = \{a + t(b - a) \; ; \; t \in [0, 1]\}$. One can observe that the segment $[a, b]$ is the range (image) of the map $\varphi : [0, 1] \mapsto \mathbb{R}^n, \; \varphi(t) = a + t(b - a)$.

Theorem (Taylor - Lagrange formula)
Let $f : \Omega \mapsto \mathbb{R}$, Ω a non empty set of $\mathbb{R}^n$, $a \in \Omega$ and suppose f of class $\mathcal{C}^2$ on Ω. Then, for every $x \in \Omega$ with $[a, x] \subseteq \Omega$, there exists $\xi \in (0, 1)$ s.t:

$$(\star) \qquad f(x) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) +$$

$$+ \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a + \xi(x - a))\,(x_i - a_i)\,(x_j - a_j)$$

Proof Let x be as above and let $\psi : [0, 1] \mapsto \mathbb{R}, \; \psi(t) = f(a+t(x-a))$. We can apply to ψ the Taylor Lagrange formula, so there is $\xi \in (0, 1)$ s.t:

$$\psi(1) = \psi(0) + \frac{\psi'(0)}{1!} + \frac{\psi''(\xi)}{2!}$$

We have $\psi(0) = f(a), \; \psi(1) = f(x)$ and $(a = (a_1, a_2, ..., a_n))$:

$$\psi'(t) = \frac{\partial f}{\partial x_1}(a + t(x - a))(x_1 - a_1) + ... + \frac{\partial f}{\partial x_n}(a + t(x - a))(x_n - a_n),$$

so:

$$\psi'(0) = \frac{\partial f}{\partial x_1}(a)(x_1 - a_1) + ... + \frac{\partial f}{\partial x_n}(a)(x_n - a_n) = Df(a)(x - a)$$

The second derivative of ψ:

$$\psi''(t) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a + t(x - a))(x_i - a_i)(x_j - a_j),$$

so

$$\psi''(\xi) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a + \xi(x-a))(x_i - a_i)(x_j - a_j)$$

and we obtain $(\star)$.

Theorem (Taylor- Young formula)

We can transform the previous formula into a Taylor-Young formula as follows.
Put $a + \xi(x-a) = \eta$ and let ω_{ij} be s.t:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\eta) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a) + \omega_{ij}(x)$$

In fact ω_{ij} is a notation; we have (f is of class $\mathcal{C}^2$): $\lim_{x \to a} \omega_{ij}(x) = 0$. So:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\eta)(x_i - a_i)(x_j - a_j) =$$

$$= \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(x_i - a_i)(x_j - a_j) + \omega_{ij}(x)(x_i - a_i)(x_j - a_j)$$

Summing this up we obtain:

$$\sum_{ij=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\eta)(x_i - a_i)(x_j - a_j) =$$

$$= \sum_{ij=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)(x_i - a_i)(x_j - a_j) + \sum_{ij=1}^{n} \omega_{ij}(x)(x_i - a_i)(x_j - a_j)$$

Denote by:

$$\omega(x) = \frac{\sum_{ij=1}^{n} \omega_{ij}(x)(x_i - a_i)(x_j - a_j)}{\| x - a \|^2}, \; x \neq a$$

We have that:

$$\omega(x) \leq \frac{\sum_{ij=1}^{n} |\omega_{ij}(x)| \, |(x_i - a_i)| \, |(x_j - a_j)|}{\| x - a \|^2}$$

and as $\dfrac{|x_i - a_i|}{\| x - a \|} \leq 1$, we get

$$|\omega(x)| \leq \sum_{ij=1}^{n} |\omega_{ij}(x)|$$

so $\omega(x) \longrightarrow 0$ as $x \longrightarrow a$.
We obtain the Taylor-Young formula:

$$(\star\star) \qquad f(x) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) +$$

$$+\frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\,(x_i - a_i)\,(x_j - a_j) + \frac{1}{2}\,\omega(x)\ \ \| x - a \|^2,$$

$x \neq a$, $\omega(x) \longrightarrow 0$ as $x \longrightarrow a$. If we put $\omega(0) = 0$, then ω is continuous and zero at a and $(\star\star)$ holds for every x with $[a, x] \subseteq \Omega$.
Using the Landau notation we get:

$$\left[f(x) - \left(f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)(x_i - a_i) + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a)\,(x_i - a_i)\,(x_j - a_j) \right) \right] =$$

$$= \boldsymbol{o}\left(\| x - a \|^2 \right)$$

Local extrema

In this section we study the local extrema of functions defined on open sets. Let $\Omega \subseteq \mathbb{R}^n$ be a non empty set open set and $f : \Omega \mapsto \mathbb{R}$.

Definition

A point $a \in \Omega$ is said to be:
(i) a ***local minimum point*** for f if there is an open ball $B(a, r) \subseteq \Omega$ s.t. $f(x) \geq f(a)$, $\forall x \in B(a, r)$.
(ii) ***a local maximum point*** for f if there is an open ball $B(a, r) \subseteq \Omega$ s.t. $f(x) \leq f(a)$, $\forall x \in B(a, r)$.
A local minimum (maximum) point for f is said to be a ***local extremum point*** for f. We say also that f has a local minimum (maximum, extremum) at a.

Remark
A point a is a local minimum (local maximum) point for f iff in a neighborhood of a one has $f(x) - f(a) \geq 0$ ($f(x) - f(a) \leq 0$). So being a local extremum point is a matter of "sign" of $f(x) - f(a)$ in a neighborhood of a.

Example
(i) $f : \mathbb{R}^2 \mapsto \mathbb{R}$, $f(x, y) = x^2 + y^2$; it is clear that $(0, 0)$ is a local minimum of f. The ***graph*** of f is the set $\{(x, y, f(x, y))\ ;\ (x, y) \in \mathbb{R}^2\} \subseteq \mathbb{R}^3$. Draw the graph of the function f.
(ii) $g : \mathbb{R}^2 \mapsto \mathbb{R}$, $g(x, y) = xy$; then $(0, 0)$ is not a local extremum for f. In fact in every neighborhood of $(0, 0)$ there are points (x, y) with $g(x, y) > 0$ and points with $g(x, y) < 0$. Use a computer to draw the graph of g.

We now remind the Fermat theorem for functions in one variable, which was already stated in the last section of the second chapter.

Theorem (Fermat)
If $\Omega \subseteq \mathbb{R}$ is an ***open set***, $f : \Omega \mapsto \mathbb{R}$ and $a \in \Omega$ s.t. f is ***differentiable*** at a and has a ***local extremum*** at a then $f'(a) = 0$.

Remark
All the underlined above conditions are essential (necessary) in the Fermat theorem. For example $f(x) = |x|$ has a local minimum at 0 but it is not differentiable at 0, etc.

The Fermat theorem generalizes to several variables as follows:
Theorem (Fermat)
If $\Omega \subseteq \mathbb{R}^n$ is an ***open set***, $f : \Omega \mapsto \mathbb{R}$ and $a \in \Omega$ are s.t. f is differentiable at a and has a local extremum at a then:

$$\frac{\partial f}{\partial x_i}(a) = 0\, \forall i = 1, 2, ..., n.$$

Proof We know that $\frac{\partial f}{\partial x_i}$ is the derivative of the function (of one variable) $t \mapsto f(a_1, a_2, ..., a_{i-1}, t, a_{i+1}, ..., a_n)$; it's clear that this function has a local extremum at a_i so the Fermat theorem for one variable functions can be applied.

Remark
The condition $Df(a) = 0$ is only ***necessary*** for local extrema of differentiable functions (see example (ii) above, or think to the one variable case). In other words, for differentiable functions on open sets the local extreme points satisfy the condition $\frac{\partial f}{\partial x_i} = 0, \forall i = 1, 2, ..., n$, but, generally, a point satisfying this condition could be not a local extremum point.

Let us introduce the terminology: for a differentiable function $f : \Omega \mapsto \mathbb{R}$ the solutions of the system:

$$(\star) \qquad \frac{\partial f}{\partial x_1}(x) = 0, \frac{\partial f}{\partial x_2}(x) = 0, ..., \frac{\partial f}{\partial x_n}(x) = 0$$

are called ***critical points*** of f.
So the Fermat theorem states that (under appropriate conditions) local extrema are critical points (but not viceversa, generally). In the remaining of this section we shall find sufficient conditions for local extrema points.

Definition
A function $\varphi : \mathbb{R}^n \mapsto \mathbb{R}$ is called a ***quadratic form*** if φ is of the form:

$$\varphi(x) = \sum_{i,j=1}^{n} a_{ij}\, x_i\, x_j, \quad a_{ij} \in \mathbb{R},\ a_{ij} = a_{j\,i},\ \forall i, j = 1, 2, ..., n$$

The matrix $(a_{ij})_{ij}$ is the matrix of the quadratic form.

Examples
(i) $\varphi(x) = \| x \|^2 = x_1^2 + x_2^2 + ... + x_n^2$ is a quadratic form; its matrix is the identity matrix I_n.
(ii) $\psi(x) = -x_1^2 - x_2^2$ is a quadratic form.
(iii) A basic example of a quadratic form is the following. If f is of class $\mathcal{C}^2$ on Ω and $a \in \Omega$ then

$$D^2 f(a)(x) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i x_j}(a)\, x_i\, x_j$$

is a quadratic form (the left hand is a notation) The matrix of $D^2f(a)$ is called the ***hessian*** of f at a. As expected, $D^2f(a)$ will play an essential role in the study of local extrema points of f.

Definition

The quadratic form $\varphi(x) = \sum_{i,j=1}^{n} a_{ij}\, x_i\, x_j$ is ***positive*** (***negative***) definite if $\varphi(x) > 0,\ \forall x \neq 0$ ($\varphi(x) < 0,\ \forall x \neq 0$).

Proposition

Suppose φ be a positive definite quadratic form. Then there exists $\mu > 0$ s.t. $\varphi(x) \geq \mu \ \| x \|^2,\ \forall x \in \mathbb{R}^n$.

Proof For $x = 0$ the inequality is trivial. Let $S = \{x \in \mathbb{R}^n\ ;\ \| x \|= 1\}$; S is a compact set. As φ is clearly continuous then φ is bounded and reaches its bounds on S.

Let $\mu = \inf_S \varphi$; then $\mu > 0$ (being a value of φ and only $\varphi(0) = 0$, but $0 \notin S$). So $\varphi(y) \geq \mu,\ \forall y \in S$. If $x \neq 0$, then $\frac{x}{\|x\|} \in S$, so $\varphi\left(\frac{x}{\|x\|}\right) \geq \mu$, but $\varphi\left(\frac{x}{\|x\|}\right) = \frac{1}{\|x\|^2}\ \varphi(x)$ and we obtain $\varphi(x) \geq \mu \ \| x \|^2$ as desired.

We now can state our basic result on sufficient conditions for extrema points.

Theorem

Let $f : \Omega \mapsto \mathbb{R}$ be of class $\mathcal{C}^2$ on the open set $\Omega \subseteq \mathbb{R}^n$ and $a \in \Omega$ be a critical point of f. Suppose that $D^2f(a)$ is positive (negative) definite. Then a is a local minimum (maximum) for f.

Proof It is enough to consider the positive definite case.

The Taylor-Young formula gives

$$f(x) = f(a) + \frac{1}{2}D^2f(a)(x-a) + \frac{1}{2}\ \omega(x)\ \ \| x - a \|^2, \quad \lim_{x\to a} \omega(x) = 0$$

But $D^2f(a)(x) \geq \mu \ \| x \|^2$ for a convenient $\mu > 0$ so

$$f(x) - f(a) \geq \frac{1}{2}\ (\mu + \omega(x))\ \ \| x - a \|^2$$

As $\lim_{x\to a} \omega(x) = 0$ and $\mu > 0$, the function $\mu + \omega(x) \geq 0$ in a neighborhood of a. This proves the result.

Remark
It can be proved that if a is a critical point and $D^2 f(a)$ takes both strictly positive and strictly negative values then a ***is not*** a local extremum point.

The problem now is to find methods for checking the positive (negative) definiteness of quadratic forms.

Proposition
The quadratic form φ is positive (negative) definite iff all the eigenvalues of the hessian matrix are strictly positive (negative).

We do not prove this result but deduce an elementary test for the case of functions of two variables. For doing this let us use some traditional notations:

$$r = \frac{\partial^2 f}{\partial x^2}, \quad s = \frac{\partial^2 f}{\partial x \partial y}, \quad t = \frac{\partial^2 f}{\partial y^2}$$

(other notations, not to be used now: $p = \frac{\partial f}{\partial x}$, $q = \frac{\partial f}{\partial y}$). Using these notations we obtain:

$$D^2 f(a)(x, y) = r(a, b)\, x^2 + 2s(a, b)\, x\, y + t(a, b)\, y^2$$

The study of the sign of $D^2 f(a)(x, y)$ is closely related to the elementary study of quadratic functions. We get (the discriminant, etc):

$D^2 f(a, b)$ is definite (positive or negative) iff $r(a, b)\, t(a, b) - s^2(a, b) > 0$

If this condition holds, then:
$D^2 f(a, b)$ is positive (negative) definite iff $r(a, b) > 0$ (< 0).

Resuming all the discussions we have the following "algorithm" for searching local extrema of $\mathcal{C}^2$ - functions of two variables on open sets:
1. Find the critical points of f.
2. Apply to every critical point the above test "$rt - s^2$:
(i) if $rt - s^2 > 0$ then we have a local extremum point;
(ii) if $rt - s^2 < 0$ then we do not have a local extremum point (see the

above remark);
(iii) if $rt - s^2 = 0$ some other considerations are needed (and we do not enter into details).

Example
Take $f : \mathbb{R}^2 \mapsto \mathbb{R},\ f(x,y) = xy(a-x-y),\ a > 0$. Let us find the local extrema of f.
Solution Of course, f is of class $\mathcal{C}^2$ on $\mathbb{R}^2$.
1. To find the critical points we solve the system $\frac{\partial f}{\partial x} = 0,\ \frac{\partial f}{\partial y} = 0$. But:

$$\frac{\partial f}{\partial x} = y(a - 2x - y),\ \frac{\partial f}{\partial y} = x(a - x - 2y),$$

so the system becomes: $\left\{ \begin{array}{rcl} y(a-2x-y) & = & 0 \\ x(a-x-2y) & = & 0 \end{array} \right.$.
The solutions are $(0,0),\ (0,a),\ (a,0),\ \left(\frac{a}{3},\frac{a}{3}\right)$. We shall test (if it is extremum or not) only the point $\left(\frac{a}{3},\frac{a}{3}\right)$.
2. Compute:

$$r = \frac{\partial^2 f}{\partial x^2} = -2y,\ s = \frac{\partial^2 f}{\partial x \partial y} = a - 2x - 2y,\ t = \frac{\partial^2 f}{\partial y^2} = -2x$$

Now compute:

$$r\left(\frac{a}{3},\frac{a}{3}\right) = -\frac{2a}{3},\ s\left(\frac{a}{3},\frac{a}{3}\right) = -\frac{a}{3},\ t\left(\frac{a}{3},\frac{a}{3}\right) = -\frac{2a}{3}$$

Then $rt - s^2 = \frac{a^2}{3} > 0$, so $\left(\frac{a}{3},\frac{a}{3}\right)$ is a local extremum point. As $r\left(\frac{a}{3},\frac{a}{3}\right) = -\frac{2a}{3} < 0$, $\left(\frac{a}{3},\frac{a}{3}\right)$ is a local maximum for f.
Let us give a geometric interpretation of the above result (justifying somehow the interest of the point $\left(\frac{a}{3},\frac{a}{3}\right)$). If $x,y,z > 0$ then xyz can be thought as the volume of a parallelepiped of dimensions x,y,z. Consider the functions $f(x,y) = xy(a-x-y)$ only for $x > 0,\ y > 0,\ a - x - y > 0$, so f is defined on the open triangle T defined by the points $(0,0),\ (a,0),\ (0,a)$. We keep the notation f (although the set of definition is changed). Now the geometric interpretation of f is: f is the volume of a parallelepiped of dimensions $x, y, a-x-y$. But

$x + y + (a - x - y) = a$, so all the parallelepipedes considered have the ***same perimeter***. A nice problem could be stated: from all the parallelepipedes of a given perimeter find that one having maximum volume.
We already know that $\left(\frac{a}{3}, \frac{a}{3}\right)$ is the only critical point of f in T and we know it is a local maximum point (corresponding to the ***cube***). Can we declare this point to be the solution of the stated geometrical problem ? For the moment no, simply because it is only a ***local*** maximum. But we can consider the function f on the ***closed*** triangle T′: $x \geq 0,\ y \geq 0,\ a - x - y \geq 0$. This time the geometric insight is lost (at least one dimension could be 0), but T′ is a compact set so f (being continuous) has a maximum value on T′. It is easy to see that the maximum value is not to be reached on the sides, hence it is valued at $\left(\frac{a}{3}, \frac{a}{3}\right)$. So in fact the cube is the solution to our problem. The above reasoning is a useful "combination" of a local study (Fermat theorem, open sets) with a global one (continuous functions on compact sets).

An important application of computing extrema of functions of several variables is the following method to approximate functions of one variable with affine functions.

The least squares method
Suppose that for a function $f : I \subseteq \mathbb{R} \mapsto \mathbb{R}$ the values at (distinct) points $x_0, x_1, ..., x_p$ are known:

$$f(x_0) = y_0,\ f(x_1) = y_1, ..., f(x_p) = y_p$$

Generally, the points $M_i(x_i, y_i)$ are not collinear, which means that there are not $a, b \in \mathbb{R}$ s.t. $y_i - ax_i - b = 0,\ \forall i = 0, 1, ..., p$. Instead, we look for $a, b \in \mathbb{R}$ s.t. the sum of squares

$$\sum_{i=0}^{p} (y_i - ax_i - b)^2$$

be as small as possible. More precisely, consider the function

$$E : \mathbb{R}^2 \mapsto \mathbb{R},\ E(a, b) = \sum_{i=0}^{p} (y_i - ax_i - b)^2$$

The problem is to find the minima points of E. By using the above algorithm, we first solve the system: $\dfrac{\partial E}{\partial a} = 0, \quad \dfrac{\partial E}{\partial b} = 0$; the linear system:

$$\sum_{i=0}^{p}(y_i - ax_i - b)x_i = 0, \quad \sum_{i=0}^{p}(y_i - ax_i - b) = 0$$

has a unique solution (a_0, b_0). Now compute

$$r = \frac{\partial^2 E}{\partial a^2} = 2\sum_{i=0}^{p} x_i^2, \quad s = \frac{\partial^2 E}{\partial a \partial b} = 2\sum_{i=0}^{p} x_i, \quad t = \frac{\partial^2 E}{\partial b^2} = 2(p+1)$$

Using Schwartz inequality it can be check $rt - s^2 > 0$ and $r > 0$, so (a_0, b_0) is a minimum point for E. The line $y = a_0 x + b_0$ is called the ***regression line*** of the data $(x_0, y_0), (x_1, y_1), ..., (x_p, y_p)$.

Exercises

1. Let $f(x,y) = \begin{cases} xy \sin \frac{x^2-y^2}{x^2+y^2} & \text{if} \quad (x,y) \neq (0,0) \\ 0 & \text{if} \quad (x,y) = (0,0) \end{cases}$

(i) Prove f is of class $\mathcal{C}^1$ on R^2.

(ii) Prove that f has second order partial derivatives on R^2 and compute the mixed second order partial derivatives at the origin; is f of class $\mathcal{C}^2$ on R^2 ?

Hint (i) The first order partial derivatives:

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} y \sin \frac{x^2-y^2}{x^2+y^2} + \frac{4x^2y^3}{(x^2+y^2)^2} \cos \frac{x^2-y^2}{x^2+y^2} & \text{if} \quad (x,y) \neq (0,0) \\ 0 & \text{if} \quad (x,y) = (0,0) \end{cases}$$

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} x \sin \frac{x^2-y^2}{x^2+y^2} - \frac{4y^2x^3}{(x^2+y^2)^2} \cos \frac{x^2-y^2}{x^2+y^2} & \text{if} \quad (x,y) \neq (0,0) \\ 0 & \text{if} \quad (x,y) = (0,0) \end{cases}$$

It can be proved that $f \in \mathcal{C}^1(\mathbb{R}^2)$.

(ii) We have:

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{x \to 0} \frac{x \sin 1}{x} = \sin 1; \quad \frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{y \to 0} \frac{y \sin(-1)}{y} = -\sin 1.$$

So f is not of class $\mathcal{C}^2$ on R^2.

2. Same questions as above for $f(x,y) = \begin{cases} \frac{xy^3}{x^2+y^2} & \text{if} \quad (x,y) \neq (0,0) \\ 0 & \text{if} \quad (x,y) = (0,0) \end{cases}$

3. For a function $f \in \mathcal{C}^2(\Omega)$, Ω an open non empty subset in $\mathbb{R}^n$, the ***Laplacian*** is by definition:

$$\Delta f = \frac{\partial^2 f}{\partial x_1^2} + \frac{\partial^2 f}{\partial x_2^2} + ... + \frac{\partial^2 f}{\partial x_n^2}.$$

A function with $\Delta f = 0$ is said to be a ***harmonic*** function.
Prove the following are harmonic functions:
(i) $f : R^2 \setminus \{(0,0)\} \mapsto R, f(x,y) = \ln(x^2+y^2)$.
(ii) $g : R^3 \setminus \{(0,0,0)\} \mapsto R$, $k(x,y,z) = \dfrac{1}{\sqrt{x^2+y^2+z^2}}$.

4. Let $f \in \mathcal{C}^2(R^2)$ and $g : R^2 \mapsto R$, $g(x,y) = f(x^2+y^2, x^2-y^2)$. Compute the second order partial derivatives of g.
Hint Let $u = x^2+y^2$ and $v = x^2-y^2$; the partial derivatives of u and v are : $\dfrac{\partial u}{\partial x} = 2x$, $\dfrac{\partial u}{\partial y} = 2y$, $\dfrac{\partial v}{\partial x} = 2x$, $\dfrac{\partial v}{\partial y} = -2y$. So we have:

$$\frac{\partial g}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} = 2x\left(\frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}\right)$$

$$\frac{\partial g}{\partial y} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial y} = 2y\left(\frac{\partial f}{\partial u} - \frac{\partial f}{\partial v}\right).$$

The second order partial derivatives (g is of class $\mathcal{C}^2(R^2)$):

$$\frac{\partial^2 g}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial g}{\partial x}\right) = \frac{\partial}{\partial x}\left(2x\left(\frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}\right)\right) =$$

$$= 2\left(\frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}\right) + 2x\left(\frac{\partial^2 f}{\partial u^2}\frac{\partial u}{\partial x} + \frac{\partial^2 f}{\partial v \partial u}\frac{\partial v}{\partial x} + \frac{\partial^2 f}{\partial u \partial v}\frac{\partial u}{\partial x} + \frac{\partial^2 f}{\partial v^2}\frac{\partial v}{\partial x}\right) =$$

$$= 2\left(\frac{\partial f}{\partial u} + \frac{\partial f}{\partial v}\right) + 4x^2\left(\frac{\partial^2 f}{\partial u^2} + 2\frac{\partial^2 f}{\partial u \partial v} + \frac{\partial^2 f}{\partial v^2}\right), \quad \text{etc.}$$

5. Let $f(x,y) = e^x \sin y$; write Taylor formula at $(0,0)$ using Landau notations.
Hint $f(x,y) = y + xy + \boldsymbol{o}(x^2 + y^2)$.

6. Write Taylor formula at $(1,1)$ using Landau notations for the function $f(x,y) = y^x$.
Hint $f(x,y) = 1 + (y-1) + (x-1)(y-1) + \boldsymbol{o}((x-1)^2 + (y-1)^2)$.

7. Find the local extrema of the functions:
(i) $f : R^2 \mapsto R,\ f(x,y) = x^3 + y^3 - 6xy$.
(ii) $g : R^2 \mapsto R,\ g(x,y) = x^3 + 8y^3 - 2xy$.
Hint (i) The critical points of f are the solutions of the system:

$$\begin{cases} \frac{\partial f}{\partial x} &= 0 \\ \frac{\partial f}{\partial y} &= 0 \end{cases} \iff \begin{cases} 3x^2 - 6y &= 0 \\ 3y^2 - 6x &= 0 \end{cases}$$

We obtain two critical points: $(0,0)$ and $(2,2)$.
Now $r(0,0) = t(0,0) = 0,\ s(0,0) = -6$, so $(0,0)$ is not an extremum point and $r(2,2) = t(2,2) = 12,\ s(2,2) = -6$, so $(2,2)$ is a local minimum.

8. Find the local extrema points of $f : R^2 \mapsto R,\ f(x,y) = x^2ye^{2x+3y}$.
Hint The set of critical points is $\{(0,y)\,|\,y \in R\} \cup \{(-1, -\frac{1}{3})\}$. It is easy to check $(-1, -\frac{1}{3})$ is a local minimum. For $(0,y)$ we obtain $rt - s^2 = 0$, so we need to evaluate the sign of

$$f(x,y) - f(0,y) = x^2ye^{2x+3y}.$$

The origin is not an extremum, while the points $(0,y),\ y > 0$ are local minima and the points $(0,y),\ y < 0$ are local maxima.

9. Find the local extrema of the function:
$f : (0, 2\pi) \times (0, 2\pi) \mapsto R, f(x,y) = \sin x \sin y \sin(x+y)$.
Hint The critical points:

$$\begin{cases} \frac{\partial f}{\partial x} &= \cos x \sin y \sin(x+y) + \sin x \sin y \cos(x+y) = \sin y \sin(2x+y) = 0 \\ \frac{\partial f}{\partial y} &= \sin x \cos y \sin(x+y) + \sin x \sin y \cos(x+y) = \sin x \sin(x+2y) = 0 \end{cases}$$

We obtain: $(x_1, y_1) = (\pi, \pi)$ and $(x_2, y_2) = \left(\frac{\pi}{3}, \frac{\pi}{3}\right)$.
Second order partial derivatives:

$$\frac{\partial^2 f}{\partial x^2} = 2 \sin y \cos(2x + y),$$

$$\frac{\partial^2 f}{\partial y^2} = 2 \sin x \cos(x + 2y),$$

$$\frac{\partial^2 f}{\partial x \partial y} = \sin(2x + 2y).$$

The point (x_2, y_2) is a local maximum:

$$rt - s^2 =$$

$$= \frac{\partial^2 f}{\partial x^2}\left(\frac{\pi}{3}, \frac{\pi}{3}\right) \frac{\partial^2 f}{\partial y^2}\left(\frac{\pi}{3}, \frac{\pi}{3}\right) - \left(\frac{\partial^2 f}{\partial x \partial y}\left(\frac{\pi}{3}, \frac{\pi}{3}\right)\right)^2 = \frac{9}{4} > 0,\ r = -\sqrt{3} < 0.$$

For $(x_1, y_1) = (\pi, \pi)$, we need to evaluate the sign of $f(x, y) - f(\pi, \pi) = \sin x \sin y \sin(x + y)$ in a neighborhood of (π, π). Finally, this is not an extremum point.

10. Find the extrema of the functions:
(i) $f(x, y, z) = \dfrac{1}{x} + \dfrac{x}{y} + \dfrac{y}{z} + z,\ x \neq 0, y \neq 0, z \neq 0.$
(ii) $g : (0, \pi)^3 \mapsto R, g(x, y, z) = \sin x + \sin y + \sin z - \sin(x + y + z).$
Hint (i) The critical points are $(1, 1, 1)$ and $(-1, 1, -1)$. The hessian of f:

$$H_h(x, y, z) = \begin{pmatrix} \frac{2}{x^3} & -\frac{1}{y^2} & 0 \\ -\frac{1}{y^2} & \frac{2x}{y^3} & -\frac{1}{z^2} \\ 0 & -\frac{1}{z^2} & \frac{2y}{z^3} \end{pmatrix}$$

The point$(1, 1, 1)$ is a local minimum (all the eigenvalues of the hessian are strictly positive) and $(-1, 1 - 1)$ is a local maximum (all the eigenvalues of the hessian are strictly negative).

4.3 Implicit functions, conditional extrema and Lagrange multipliers

Implicit functions

Let $f : \mathbb{R}^2 \mapsto \mathbb{R}$ and consider the equation:

$$(\star) \qquad f(x, y) = 0$$

Intuitively the set of the solutions of $(\star)$ defines a relation between x and y. The problem is if this relation can take, at least locally, the form of a functional dependence. This means that if, for example, $f(a, b) = 0$ is it possible to find (open) neighborhoods A of a and B of b and a function $g : A \mapsto B$ s.t the equality $f(x, y) = 0,\ x \in A,\ y \in B$ be equivalent to the equality $y = g(x)$? In this case we say that the function g is ***implicitly defined*** by the equation $(\star)$. Loosely speaking one can solve $(\star)$ with respect to the unknown y (as a function of x). In other words in $A \times B$ the set of solutions of $(\star)$ is the graph of the function g. Generally, it is the ***existence*** (and eventually, ***uniqueness***) of g which matters because finding g explicitly is (in most cases) impossible.

Example
Consider the equation $x^2 + y^2 - 1 = 0$ and a point $(a, b),\ b > 0$ s.t. $a^2 + b^2 - 1 = 0$. Then it is obvious that the function $g(x) = \sqrt{1 - x^2}$ uniquely satisfies the above conditions in some neighborhood of a. But for the point $(1, 0)$ such neighborhoods are impossible to find.

Theorem (implicit function theorem for 2 variables)
Let $f : \mathbb{R}^2 \mapsto \mathbb{R}$ be a function of class $\mathcal{C}^1$ in a neighborhood of (a, b). Suppose that:
(i) $f(a, b) = 0$
(ii) $\frac{\partial f}{\partial y}(a, b) \neq 0$.
Then there are open neighborhoods A of a and B of b and an unique function $g : A \mapsto B$ of class $\mathcal{C}^1$ on A s.t:

$$x \in A,\ y \in B,\ f(x, y) = 0 \quad \text{iff} \quad y = g(x)$$

(it follows that $g(a) = b$).

We omit the proof of this theorem (see [4], [5], [8]).

Although the function g cannot be found explicitly, it is possible to compute its derivative. In fact we have that

$$f(x, g(x)) = 0, \ \forall x \in A$$

Differentiating with respect to x, we obtain:

$$f'_x(x, g(x)) + f'_y(x, g(x))\, g'(x) = 0, \ \forall x \in A$$

In particular for $x = a, \ y = b$ we have:

$$f'_x(a, b) + f'_y(a, b)\, g'(a) = 0$$

As $f'_y(a, b) \neq 0$ we finally get:

$$g'(a) = -\frac{f'_x(a, b)}{f'_y(a, b)}$$

Remark
In fact as $f'_y(x, g(x)) \neq 0$ in a neighborhood of a we obtain:

$$g'(x) = -\frac{f'_x(x, g(x))}{f'_y(x, g(x))}, \ \text{ in a neighborhood of } \ a.$$

Example
Compute the extrema of the solutions $y = y(x)$ of the equation:

$$x^3 + y^3 - 3xy = 0.$$

Hint Let $f(x, y) = x^3 + y^3 - 3xy$. The function $y = y(x)$ exists around the points satisfying $\frac{\partial f}{\partial y} \neq 0$, or $3y^2 - 3x \neq 0$. The derivative of y is:

$$3x^2 + 3y^2 y' - 3y - 3xy' = 0 \Rightarrow y'(x) = \frac{y - x^2}{y^2 - x}.$$

The critical point of f are:

$$\begin{cases} y' &= 0 \\ f &= 0 \\ \frac{\partial f}{\partial y} &\neq 0 \end{cases} \Rightarrow \begin{cases} y - x^2 &= 0 \\ x^3 + y^3 - 3xy &= 0 \\ y^2 - x &\neq 0 \end{cases}.$$

The only solution is $(x,y) = (\sqrt[3]{2}, \sqrt[3]{4})$. We now compute $y''(\sqrt[3]{2})$:

$$2x + 2y\,(y')^2 + y^2y' - y' - y' - xy'' = 0, \Rightarrow y''(\sqrt[3]{2}) = -2 < 0,$$

so $x = \sqrt[3]{2}$ is a local maximum point for y and $y(\sqrt[3]{2}) = \sqrt[3]{4}$.

We now state (without proof) the general form of the implicit function theorem.
First some notations: if $f : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^m,\ f = (f_1, f_2, ..., f_m)$ and $x \in \mathbb{R}^n,\ y \in \mathbb{R}^m$. Suppose f is differentiable at $(a,b) \in \mathbb{R}^n \times \mathbb{R}^m,\ (a \in \mathbb{R}^n,\ b \in \mathbb{R}^m)$. Define $f'_y(a,b)$ by the matrix:

$$\begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a,b) & . & . & . & . & \frac{\partial f_1}{\partial y_m}(a,b) \\ . & . & . & . & . & . \\ \frac{\partial f_m}{\partial y_1}(a,b) & . & . & . & . & \frac{\partial f_m}{\partial y_n}(a,b) \end{pmatrix}$$

Theorem (general implicit function theorem)
Let $f : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^m$ be a function of class $\mathcal{C}^1$ in a neighborhood of $(a,b) \in \mathbb{R}^n \times \mathbb{R}^m$. Suppose:
(i) $f(a,b) = 0$
(ii) $\det f'_y(a,b) \neq 0$.
Then there are open neighborhoods A of a and B of b and an unique function $g : A \mapsto B$ of class $\mathcal{C}^1$ s.t.

$$\text{for } x \in A,\ y \in B,\ f(x,y) = 0 \quad \text{iff} \quad y = g(x).$$

In the example below we show how derivatives can be computed in the case $n = 1$, $m = 2$.

Example
Let $f : \mathbb{R} \times (\mathbb{R} \setminus \{0\}) \times \mathbb{R} \mapsto \mathbb{R}^2, f(x,y,z) = (f_1(x,y,z), f_2(x,y,z))$, where:

$$f_1(x,y,z) = yz + 2y^2 + xy + y + z - 2,$$

$$f_2(x,y,z) = \arctan\frac{z}{y} + \ln\frac{x^2 + 2y^2 + z^2}{2}$$

The system of two equations $f = (0,0)$ defines implicitly two functions $y = y(x)$ and $z = z(x)$. Let first observe that $f(0,1,0) = (0,0)$; in

the above notations, $a = 0,\ b = (1,0),\ x_1 = x,\ (y_1, y_2) = (y, z)$. The problem is to compute $y'(0)$ and $z'(0)$.
First we check the condition

$$\begin{vmatrix} \frac{\partial f_1}{\partial y}(0,1,0) & \frac{\partial f_1}{\partial z}(0,1,0) \\ \\ \frac{\partial f_2}{\partial y}(0,1,0) & \frac{\partial f_2}{\partial z}(0,1,0) \end{vmatrix} = 1 \neq 0$$

is fulfilled, so the functions y and z exist around $x = 0$ and
$y(0) = 1,\ z(0) = 0$. By differentiating the system with respect to x we obtain (y and z are functions of x):

$$y'z + yz' + y + xy' + y' + z' = 0$$

$$\frac{yz' - y'z}{y^2 + z^2} + \frac{2x + 2zz' + 4yy'}{x^2 + 2y^2 + z^2} = 0$$

For $x = 0,\ y = 1,\ z = 0$ we obtain the system:

$$y'(0) + z'(0) = -1, \quad 2y'(0) + z'(0) = 0$$

We finally get: $y'(0) = 1$ and $z'(0) = -2$.

Conditional extrema

Let $f : \mathbb{R}^n \mapsto \mathbb{R}$ and $M \subseteq \mathbb{R}^n$ a non empty set. The point $a \in M$ is a ***local conditional extremum*** for f with constraint M if there is an open ball $B(a,r)$ s.t. $f(x) - f(a) \geq 0$ (or ≤ 0) on $B(a,r) \bigcap M$
The set M being not necessarily open, Fermat theorem cannot be applied to conditional local extrema. Instead, a version of Fermat theorem (called Lagrange multipliers theorem) is valid (for the proof, see [4], [5]):

Theorem (Lagrange multipliers)
Let $g : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^m,\ g = (g_1, g_2, ..., g_m)$ and let

$$M = \{(x,y) \in \mathbb{R}^n \times \mathbb{R}^m \ ;\ g(x,y) = 0\}$$

Let $f : \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}$ and $(a,b) \in M$ be s.t:
(i) f and g are of class $\mathcal{C}^1$ in a neighborhood of (a,b).

(ii) (a, b) is a local conditional extremum for f constrained by M;
(iii) $\det g'_y(a, b) \neq 0$.
Then there are (unique) $\lambda_1, \lambda_2, ..., \lambda_m \in R$ (called ***Lagrange multipliers***) s.t. the function $F = f + \lambda_1 g_1 + \lambda_2 g_2 + ... + \lambda_m g_m$ satisfies the conditions:

$$\frac{\partial F}{\partial x_i}(a, b) = 0,\ \frac{\partial F}{\partial y_j}(a, b) = 0,\ i = 1, 2, ..., n,\ j = 1, 2, ..., m$$

To apply this theorem (we suppose the hypothesis are fulfilled) one proceeds as follows:
(i) Form the function $F = f + \sum_{k=1}^{m} \lambda_k g_k$ (with unknowns $\lambda_1, \lambda_2, ..., \lambda_m$).
(ii) Solve the system:

$$\frac{\partial F}{\partial x_i}(x, y) = 0,\ \frac{\partial F}{\partial y_j}(x, y) = 0,\ g_k(x, y) = 0,\ i = 1, 2, ..., n,\ j, k = 1, 2, ..., m$$

There are $n + 2m$ equations and $n + 2m$ unknowns:

$$x = (x_1, ..., x_n),\ y = (y_1, ..., y_m),\ \lambda = (\lambda_1, ..., \lambda_m)$$

(iii) For every solution (λ, x, y), the point (x, y) is a possible local conditional extremum.

Example
Let $f : \mathbb{R}^2 \mapsto \mathbb{R},\ f(x, y) = x+y$ and $g : \mathbb{R}^2 \mapsto \mathbb{R},\ g(x, y) = x^2+y^2-1$.
We want to compute the maximum and minimum values of f on
$M = \{(x, y) \in \mathbb{R}^2\ ;\ g(x, y) = 0\}$.
Solution First observe that:
(i) the problem has solutions (f is continuous and the constrain M is compact).
(ii) the equation $g(x, y) = 0$ satisfies implicit function theorem at every point (with respect to x or to y).
So one can use Lagrange multipliers theorem to find the local conditional extrema candidates; then decide the maximum and the minimum.

Define $F(x,y) = x + y + \lambda(x^2 + y^2 - 1)$; the system:

$$\begin{cases} \frac{\partial F}{\partial x}(x,y) = & 1 + 2\lambda x = 0 \\ \frac{\partial F}{\partial y}(x,y) = & 1 + 2\lambda y = 0 \\ g(x,y) = & x^2 + y^2 - 1 = 0 \end{cases}$$

has the solutions:
$\lambda_1 = \frac{\sqrt{2}}{2}$, $x_1 = -\frac{\sqrt{2}}{2}$, $y_1 = -\frac{\sqrt{2}}{2}$ and $\lambda_2 = -\frac{\sqrt{2}}{2}$, $x_2 = \frac{\sqrt{2}}{2}$, $y_2 = \frac{\sqrt{2}}{2}$
So being exactly two local conditional extrema candidates, they are the points where f attains its extreme values:

$$\inf_M f = f(x_1, y_1) = -\sqrt{2}, \quad \sup_M f = f(x_2, y_2) = \sqrt{2}$$

We shall not consider the problem of finding sufficient conditions for local conditional extrema (see [1], [7]).

Exercises

1. Find the extrema of the solutions $y = y(x)$ defined by the equation $x^3 + y^3 - 2xy = 0$.
Hint Let $f(x,y) = x^3 + y^3 - 2xy$; the condition $\frac{\partial f}{\partial y} \neq 0$ is $3y^2 - 2x \neq 0$. If this condition is fulfilled we compute the derivative of $y = y(x)$:

$$3x^2 + 3y^2y' - 2y - 2xy' = 0 \Rightarrow y'(x) = \frac{2y - 3x^2}{3y^2 - 2x}.$$

Now compute the critical points, etc.

2. Find the extrema of the solutions $y = y(x)$ defined by the equation $x^3 + y^3 - 3x^2y - 3 = 0$.

3. The function $z = z(x,y)$ is implicitly defined by the equation:

$$(y + z)\sin z - y(x + z) = 0.$$

Compute:

$$E = (z \sin z)\frac{\partial z}{\partial x} - y^2 \frac{\partial z}{\partial y}.$$

Hint Let $f(x, y, z) = (y+z) \sin z - y(x+z)$. The hypothesis of implicit function theorem is:

$$\sin z + (y + z) \cos z - y \neq 0.$$

By differentiating with respect to x and y we obtain:

$$\frac{\partial z}{\partial x} \sin z + (y + z) \cos z \frac{\partial z}{\partial x} - y \left(1 + \frac{\partial z}{\partial x}\right) = 0,$$

so:

$$\frac{\partial z}{\partial x} = \frac{y}{\sin z + (y + z) \cos z - y}.$$

$$\left(1 + \frac{\partial z}{\partial y}\right) \sin z + (y + z) \cos z \frac{\partial z}{\partial y} - (x + z) - y\frac{\partial z}{\partial y} = 0,$$

so

$$\frac{\partial z}{\partial y} = \frac{x + z - \sin z}{\sin z + (y + z) \cos z - y}.$$

Replacing $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ in E we obtain:

$$E = \frac{yf(x, y, z)}{\sin z + (y + z) \cos z - y} = 0.$$

4. Compute the extrema of $z = z(x, y)$, implicitly defined by the equation: $z^3 + z + 20(x^2 + y^2) - 8(xy + x + y) = 0$.
Hint Let $f(x, y, z) = z^3 + z + 20(x^2 + y^2) - 8(xy + x + y)$; the implicit function theorem can be applied at every $(x, y, z) \in \mathbb{R}^3$. The first order partial derivatives of z:

$$3z^2\frac{\partial z}{\partial x} + \frac{\partial z}{\partial x} + 40x - 8(y + 1) = 0 \Rightarrow \frac{\partial z}{\partial x} = \frac{8(y + 1) - 40x}{3z^2 + 1}.$$

$$3z^2\frac{\partial z}{\partial y} + \frac{\partial z}{\partial y} + 40y - 8(x + 1) = 0 \Rightarrow \frac{\partial z}{\partial y} = \frac{8(x + 1) - 40y}{3z^2 + 1}.$$

The critical points of z are the solutions of the system:

$$\begin{cases} \frac{\partial z}{\partial x} & = 0 \\ \frac{\partial z}{\partial y} & = 0 \\ f(x,y,z) & = 0 \end{cases} \Rightarrow \begin{cases} \frac{8(y+1)-40x}{3z^2+1} & = 0 \\ \frac{8(x+1)-40y}{3z^2+1} & = 0 \\ z^3+z+20(x^2+y^2)-8(xy+x+y) & = 0 \end{cases}$$

The unique critical point is $(x,y,z)=(\frac{1}{4},\frac{1}{4},1)$. The second order partial derivatives of z:

$$6z\left(\frac{\partial z}{\partial x}\right)^2+3z^2\frac{\partial^2 z}{\partial x^2}+\frac{\partial^2 z}{\partial x^2}+40=0 \Rightarrow \frac{\partial^2 z}{\partial x^2}=-\frac{40}{3z^2+1}.$$

$$6z\left(\frac{\partial z}{\partial y}\right)^2+3z^2\frac{\partial^2 z}{\partial y^2}+\frac{\partial^2 z}{\partial y^2}+40=0 \Rightarrow \frac{\partial^2 z}{\partial y^2}=-\frac{40}{3z^2+1}.$$

$$6z\frac{\partial z}{\partial y}\frac{\partial z}{\partial x}+3z^2\frac{\partial^2 z}{\partial x\partial y}+\frac{\partial^2 z}{\partial x\partial y}-8=0 \Rightarrow \frac{\partial^2 z}{\partial x\partial y}=\frac{8}{3z^2+1}.$$

We get: $\frac{\partial^2 z}{\partial x^2}\left(\frac{1}{4},\frac{1}{4}\right)=-10$, $\frac{\partial^2 z}{\partial y^2}\left(\frac{1}{4},\frac{1}{4}\right)=-10$, $\frac{\partial^2 z}{\partial x\partial y}\left(\frac{1}{4},\frac{1}{4}\right)=2$, so $(\frac{1}{4},\frac{1}{4})$ is a local maximum point for z.

5. Compute $y'(1)$, y being the solution of the equation

$$x^3-y-\cos y=0, \quad \text{with} \quad y(1)=0.$$

6. Compute $y'(0)$ and $y''(0)$, if $y(0)=0$ and

$$e^{x^2-y^2}=\sin(x+2y)+1.$$

7. Compute y' and y'' if $y=x+\ln y$. (Make this more precise).

8. Compute $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ if: $x\cos y+y\cos z+z\cos x=1$.

9. Compute the extreme values of $f(x,y)=2x^2+2y^2+2x$ on the set $D'=\{(x,y)\in \mathbb{R}^2 \; ; \; x^2+y^2\leq 1\}$.
Hint We observe the problem has solutions (as f is continuous and D' is compact). We solve the problem in two steps: first we find the local (unconditional, free) extrema on the open unit disk and then we find

the local conditional extrema on the unit circle. For the first step we compute the critical points of f:

$$\frac{\partial f}{\partial x} = 4x + 2 = 0, \ \frac{\partial f}{\partial y} = 4y = 0,$$

so there is only one critical point $\left(-\frac{1}{2}, 0\right)$.
Now $r = 4, \ t = 4, \ s = 0$, so $\left(-\frac{1}{2}, 0\right)$ is a local minimum point and $f\left(-\frac{1}{2}, 0\right) = -\frac{1}{2}$.
For the second step we consider $F(x, y) = 2x^2 + 2y^2 + 2x + \lambda(x^2 + y^2 - 1)$.
The system:

$$\begin{cases} \frac{\partial F}{\partial x}(x, y) = & 2(x(\lambda + 2) + 1) = 0 \\ \frac{\partial F}{\partial y}(x, y) = & 2y(\lambda + 2) = 0 \\ g(x, y) = & x^2 + y^2 - 1 = 0 \end{cases}$$

has the solutions:
$\lambda_1 = -1, \ x_1 = -1, \ y_1 = 0$ and $\lambda_2 = 1, \ x_2 = 1, \ y_2 = 0$.
By comparing the values of f we find:

$$f\left(-\frac{1}{2}, 0\right) = -\frac{1}{2}, \ f(-1, 0) = 0, \ f(1, 0) = 4,$$

so $\inf\limits_{D'} f = -\frac{1}{2}$ and $\sup\limits_{D'} f = 4$

10. Compute the extreme values of $f(x, y) = xy$ on the ellipse of equation $x^2 + 2y^2 = 1$.
Hint Using Lagrange multipliers method we consider
$F(x, y) = xy + \lambda(x^2 + 2y^2 - 1)$. We finally obtain (the constrain is compact !):
$\inf f = -\frac{\sqrt{2}}{4}$ and $\sup f = \frac{\sqrt{2}}{4}$.

11. Let $f(x, y) = x^2 + y^2 - xy + x + y$. Compute $\inf\limits_{K} f$ and $\sup\limits_{K} f$, if $K = \{(x, y) \in \mathbb{R}^2 \ ; \ x \leq 0, \ y \leq 0, \ x + y \geq -3\}$.
Hint Same method as in exercise 9. Inside the triangle we have the critical point $(-1, -1)$ with $f(-1, -1) = -1$. On the edges of the triangle we have 3 cases: on $x = 0$, on $y = 0$ and on $x + y = -3$, etc.

Bibliography

[1] K.R. Davidson, A.P. Donsig, *Real Analysis with Real Applications*, Prentice Hall, 2002.

[2] B.P. Demidovitch, *Culegere de probleme si exercitii de analiza matematica*, Ed Tehnica, 1956. (other editions, Mir Publishers).

[3] N. Donciu, D. Flondor, *Analiza matematica, culegere de probleme*, Ed ALL, 2004.

[4] P. Flondor, O. Stanasila, *Lectii de analiza matematica si exercitii rezolvate*, Ed ALL, 2004.

[5] G.B. Folland, *Advanced Calculus*, Prentice Hall, 2002.

[6] M. Olteanu, *Analiza Matematica - notiuni teoretice si probleme rezolvate*, Ed. Printech, 2005.

[7] Gh. Oprisan, O. Stanasila, *Analiza matematica in 14 lectii*, Ed. Printech, 2006.

[8] W. Rudin, *Principles of Mathematical Analysis*, Mc Graw-Hill, 1976.