5 Functions of several variables

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

Let $n \in \mathbb{N}^*$. Consider the set $\mathbb{R}^n = \{(x_1, x_2, \ldots x_n) \mid x_1, x_2, \ldots, x_n \in \mathbb{R}\}$ ($\mathbb{R}^n$ consists of all ordered n-tuples of real numbers). The elements of $\mathbb{R}^n$ are called *vectors* or *points*. If $x \in \mathbb{R}^n$, then $x = (x_1, x_2, \ldots, x_n)$ and $x_1, x_2, \ldots, x_n$ are called the *coordinates* (or *components*) of the vector x (for $i = 1, 2, \ldots, n$, x_i is the i^{th} coordinate (or component) of x). Note that, for $n \geq 2$, $\mathbb{R}^n$ can be viewed as "products" of lower-dimensional spaces (for instance, putting side by side the first $(n-1)$ coordinates and then the n^{th} one: $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R} = \underbrace{\mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R}}_{n \text{ times}}$).

Sum of two vectors (vector addition): If $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, then $x + y = (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)$.
Multiplication of a vector by a number (scalar multiplication): If $x = (x_1, x_2, \ldots, x_n)$ and $\alpha \in \mathbb{R}$, then $\alpha x = (\alpha x_1, \alpha x_2, \ldots, \alpha x_n)$.
$\mathbb{R}^n$ together with the vector addition and the scalar multiplication

$$\forall (x, y) \in \mathbb{R}^n \times \mathbb{R}^n \longmapsto x + y \in \mathbb{R}^n$$
$$\forall (\alpha, x) \in \mathbb{R} \times \mathbb{R}^n \longmapsto \alpha x \in \mathbb{R}^n$$

is a vector space over the field $\mathbb{R}$ of real numbers. The zero vector (origin) of this vector space is the point $0_n = (0, 0, \ldots, 0)$ (should not be confused with the number 0) and the additive inverse of a point $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ is the point denoted by $-x = (-x_1, -x_2, \ldots, -x_n) = (-1)x$.
Notation:

$$\begin{aligned} e^1 &= (1, 0, 0, \ldots, 0) \in \mathbb{R}^n \\ e^2 &= (0, 1, 0, \ldots, 0) \in \mathbb{R}^n \\ &\vdots \\ e^n &= (0, 0, 0, \ldots, 1) \in \mathbb{R}^n \end{aligned}$$

Thus, for $1 \leq i, j \leq n$, $e^i_j = \delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j. \end{cases}$

Note that δ_{ij} is called the Kronecker delta function. The set $\{e^1, e^2, \ldots e^n\}$ is a basis of the vector space $\mathbb{R}^n$ called the *standard (canonical) basis* of $\mathbb{R}^n$. If $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, then $x = x_1 e^1 + x_2 e^2 + \ldots + x_n e^n$.

Let $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$. The real number defined by

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n$$

is called the *scalar product of x and y*. The nonnegative number

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{(x_1)^2 + (x_2)^2 + \ldots + (x_n)^2}$$

is called the *Euclidean norm of x*. The *Euclidean distance between x and y* is given by $\|x - y\|$.

Remark 5.1. For $x \in \mathbb{R}^n$, $\|x\|$ represents the Euclidean distance between x and 0_n.

Proposition 5.1 (Properties of the scalar product in $\mathbb{R}^n$)**.**

(i) $\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle, \quad \forall x, y, z \in \mathbb{R}^n$.

(ii) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle, \quad \forall \alpha \in \mathbb{R}, \forall x, y \in \mathbb{R}^n$.

(iii) $\langle x, y\rangle = \langle y, x\rangle, \quad \forall x, y \in \mathbb{R}^n$.

(iv) $\langle x, x\rangle > 0, \quad \forall x \in \mathbb{R}^n \setminus \{0_n\}$.

(v) $\langle 0_n, x\rangle = 0, \quad \forall x \in \mathbb{R}^n$.

(vi) $\langle x, x\rangle = 0 \Leftrightarrow x = 0_n$.

Proposition 5.2 (Cauchy-Buniakowski-Schwarz inequality)**.**

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\|, \quad \forall x, y \in \mathbb{R}^n.$$

Proposition 5.3 (Properties of the Euclidean norm)**.**

(i) $\|x\| = 0 \Leftrightarrow x = 0_n$.

(ii) $\|\alpha x\| = |\alpha| \cdot \|x\|, \quad \forall \alpha \in \mathbb{R}, \forall x \in \mathbb{R}^n$.

(iii) $\|x+y\| \leq \|x\| + \|y\|, \quad \forall x, y \in \mathbb{R}^n$ *(the triangle inequality).*

Remark 5.2. A vector $x \in \mathbb{R}^n$ is called a *unit vector* if $\|x\| = 1$. If $y \in \mathbb{R}^n \setminus \{0_n\}$, we can define the vector $\dfrac{1}{\|y\|}y$ which is a unit vector (we say that we *normalize* y).

Definition 5.1. Let $x_0 \in \mathbb{R}^n$ and $r > 0$. The set

- $B(x_0, r) = \{x \in \mathbb{R}^n \mid \|x - x_0\| < r\}$ is the *open ball of radius r centered at x_0*.
- $\overline{B}(x_0, r) = \{x \in \mathbb{R}^n \mid \|x - x_0\| \leq r\}$ is the *closed ball of radius r centered at x_0*.

Remark 5.3. Let $x_0 \in \mathbb{R}^n$, $r_1 > r > 0$. Then
(i) $x_0 \in B(x_0, r)$.
(ii) $B(x_0, r) \subseteq \overline{B}(x_0, r) \subseteq B(x_0, r_1) \subseteq \overline{B}(x_0, r_1)$.
(iii) $\forall x \in B(x_0, r)$, $B\left(x, r - \|x_0 - x\|\right) \subseteq B(x_0, r)$.

Definition 5.2. A *neighborhood* of $x \in \mathbb{R}^n$ is a set $V \subseteq \mathbb{R}^n$ for which $\exists r > 0$ s.t. $B(x, r) \subseteq V$. We denote by $\mathcal{V}(x)$ the family of all neighborhoods of x.

Sequences in $\mathbb{R}^n$

Notation: $(x^k)_{k\geq 1}$, $(x^k)_{k\in\mathbb{N}^*}$, or (x^k) (we do not index this sequence by n since n is the dimension of $\mathbb{R}^n$; we use an upper index notation since lower indexes are used for vector coordinates). Written explicitly,

$$\begin{aligned}
x^1 &= (x_1^1, x_2^1, \ldots, x_n^1) \in \mathbb{R}^n \\
x^2 &= (x_1^2, x_2^2, \ldots, x_n^2) \in \mathbb{R}^n \\
&\vdots \\
x^k &= (x_1^k, x_2^k, \ldots, x_n^k) \in \mathbb{R}^n \\
&\vdots
\end{aligned}$$

The sequences of real numbers $(x_1^k)_{k\in\mathbb{N}^*}$, $(x_2^k)_{k\in\mathbb{N}^*}$, ..., $(x_n^k)_{k\in\mathbb{N}^*}$ are called the *component sequences* of the sequence (x^k).

Definition 5.3. A sequence (x^k) in $\mathbb{R}^n$ is said to *converge* (or *tend*) to $x \in \mathbb{R}^n$ if

$$\forall \varepsilon > 0, \exists K = K(\varepsilon) \in \mathbb{N}^* \text{ such that } \forall k \geq K \text{ we have } \|x^k - x\| < \varepsilon.$$

Remark 5.4. A sequence in $\mathbb{R}^n$ cannot converge to two distinct vectors in $\mathbb{R}^n$.

Definition 5.4. If a sequence (x^k) in $\mathbb{R}^n$ converges to some $x \in \mathbb{R}^n$, we say that (x^k) is *convergent*. The vector x is called the *limit* of (x^k) and we write

$$\lim_{k\to\infty} x^k = x \quad \text{or} \quad x^k \to x.$$

If (x^k) does not converge to any vector in $\mathbb{R}^n$, we say that (x^k) is *divergent*.

Remark 5.5. A sequence (x^k) in $\mathbb{R}^n$ is
(i) convergent if and only if $\exists x \in \mathbb{R}^n, \forall \varepsilon > 0, \exists K \in \mathbb{N}^*, \forall k \geq K, \|x^k - x\| < \varepsilon$.
(ii) divergent if and only if $\forall x \in \mathbb{R}^n, \exists \varepsilon > 0,\ \forall K \in \mathbb{N}^*,\ \exists k \geq K, \|x^k - x\| \geq \varepsilon$.

Theorem 5.1. *Let (x^k) be a sequence in $\mathbb{R}^n$ and let $x \in \mathbb{R}^n$. Then*

$$\lim_{k\to\infty} x^k = x \iff \forall V \in \mathcal{V}(x), \exists K = K(V) \in \mathbb{N}^* \textit{ such that } \forall k \geq K \textit{ we have } x^k \in V$$

(every neighborhood of x contains all terms of (x^k) except a finite number).

The following result gives a characterization of the limit of a sequence in $\mathbb{R}^n$ in terms of the limits of the component sequences.

Theorem 5.2. *Let (x^k) be a sequence in $\mathbb{R}^n$ with $x^k = (x_1^k, x_2^k, \ldots, x_n^k)$ and let $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Then*

$$\lim_{k\to\infty} x^k = x \iff \forall j \in \{1, 2, \ldots, n\}, \lim_{k\to\infty} x_j^k = x_j.$$

Open and closed sets. Interior, closure, boundary and derived set.

Definition 5.5. Let $A \subseteq \mathbb{R}^n$. A point $a \in A$ is called an *interior point* of A if there exists $r > 0$ such that $B(a, r) \subseteq A$. The set of all interior points of A is called the *interior* of A and is denoted by $\mathrm{int}A$ (sometimes we write $\mathrm{int}(A)$).

Definition 5.6. Let $A \subseteq \mathbb{R}^n$. The set $A \subseteq \mathbb{R}^n$ is called

- *open*: if every $a \in A$ is an interior point of A.
- *closed*: if $\mathbb{R}^n \setminus A$ is open.

Theorem 5.3. *A set $A \subseteq \mathbb{R}^n$ is closed if and only if for every sequence (x^k) in A which converges to some $c \in \mathbb{R}^n$, we have that $c \in A$.*

Remark 5.6. (i) A set in $\mathbb{R}^n$ may be neither open, nor closed.
(ii) $\mathbb{R}^n$ and $\emptyset$ are both open and closed (in fact these are the only sets that are both open and closed).

Definition 5.7. Let $A \subseteq \mathbb{R}^n$. A point $c \in \mathbb{R}^n$ is called an *adherent point* of A if for every $r > 0$, $B(c, r) \cap A \neq \emptyset$. The set of all adherent points of A is called the *closure* of A and is denoted by $\mathrm{cl}A$ (sometimes we write $\mathrm{cl}(A)$ or $\overline{A}$).

Remark 5.7. Let $A \in \mathbb{R}^n$. Then $\mathrm{cl}A = \big\{c \in \mathbb{R}^n \mid \exists (x^k) \text{ a sequence in } A \text{ which converges to } c\big\}$.

Remark 5.8 (Interior vs. closure)**.** Let $A \in \mathbb{R}^n$.
(i) $\mathrm{int}A \subseteq A$ and $\mathrm{int}A = A$ if and only if A is open; $A \subseteq \mathrm{cl}A$ and $A = \mathrm{cl}A$ if and only if A is closed.
(ii) $\mathrm{int}A$ is the largest open set contained in A; $\mathrm{cl}A$ is the smallest closed set containing A.
(iii) $\mathrm{int}\,(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus \mathrm{cl}A$; $\mathrm{cl}\,(\mathbb{R}^n \setminus A) = \mathbb{R}^n \setminus \mathrm{int}A$.

Example 5.1. (i) $n = 1$:

- $A = (0, 1)$: $\mathrm{int}A = A$, A open, $\mathrm{cl}A = [0, 1]$.
- $A = [0, 1]$: $\mathrm{int}A = (0, 1)$, $\mathrm{cl}A = A$, A closed.
- $A = [0, 1)$: $\mathrm{int}A = (0, 1)$, $\mathrm{cl}A = [0, 1]$, A neither closed, nor open.
- $A = \mathbb{R}^*$: $\mathrm{int}A = A$, A open, $\mathrm{cl}A = \mathbb{R}$.
- $A = \mathbb{N}$: $\mathrm{int}A = \emptyset$, $\mathrm{cl}A = A$, A closed.

(ii) $n = 2$:

- $A = [0, 1] \times [0, 2] \setminus \{0_2\}$: $\mathrm{int}A = (0, 1) \times (0, 2)$, $\mathrm{cl}A = [0, 1] \times [0, 2]$, A neither closed, nor open.
- $A = \big\{(x, y) \in \mathbb{R}^2 \mid x > 0, y \neq 0\big\}$: $\mathrm{int}A = A$, A open, $\mathrm{cl}A = \big\{(x, y) \in \mathbb{R}^2 \mid x \geq 0\big\}$.

(iii) Any open ball in $\mathbb{R}^n$ is an open set (by Remark 5.3.(iii)).
(iv) Any closed ball in $\mathbb{R}^n$ is a closed set.

Proof. Let $x \in \mathbb{R}^n$, $r > 0$. We show that $\mathbb{R}^n \setminus \overline{B}(x, r)$ is open. Let $y \in \mathbb{R}^n \setminus \overline{B}(x, r)$. Then $r_y = \|x - y\| - r > 0$. For any $z \in B(y, r_y)$, $\|z - y\| < r_y = \|x - y\| - r \leq \|x - z\| + \|z - y\| - r$. Thus, $\|x - z\| > r$, so $z \in \mathbb{R}^n \setminus \overline{B}(x, r)$. Hence, $B(y, r_y) \subseteq \mathbb{R}^n \setminus \overline{B}(x, r)$. It follows that $\mathbb{R}^n \setminus \overline{B}(x, r)$ is open, which means that $\overline{B}(x, r)$ is closed. □

Definition 5.8. Let $A \in \mathbb{R}^n$. A point $c \in \mathbb{R}^n$ is called a *boundary point* of A if for every $r > 0$, $B(c, r) \cap A \neq \emptyset$ and $B(c, r) \cap (\mathbb{R}^n \setminus A) \neq \emptyset$. The set of all boundary points of A is called the *boundary* of A and is denoted by $\mathrm{bd}A$ (sometimes we write $\mathrm{bd}(A)$).

Remark 5.9. Let $A \in \mathbb{R}^n$.
(i) $\mathrm{bd}A = \mathrm{cl}A \setminus \mathrm{int}A = \mathrm{cl}A \cap \mathrm{cl}\,(\mathbb{R}^n \setminus A)$.
(ii) $\mathrm{bd}A = \mathrm{bd}\,(\mathbb{R}^n \setminus A)$.
(iii) $\mathrm{bd}A$ is closed.

Definition 5.9. Let $A \in \mathbb{R}^n$. A point an $c \in \mathbb{R}^n$ is called an *accumulation point* of A if for every $r > 0$, $B(c, r) \cap (A \setminus \{c\}) \neq \emptyset$. The set of all accumulation points of A is called the *derived set* of A and is denoted by A' (sometimes we write $(A)'$).

Remark 5.10. Let $A \in \mathbb{R}^n$. Then $A' = \big\{c \in \mathbb{R}^n \mid \exists (x^k)$ a sequence in $A \setminus \{c\}$ which converges to $c\big\}$.

Remark 5.11. Let $A \in \mathbb{R}^n$.
(i) $\mathrm{int}A \subseteq A' \subseteq \mathrm{cl}A$.
(ii) $\mathrm{cl}A = A' \cup A$.
(iii) A is closed if and only if $A' \subseteq A$.
(iv) A' is closed.

Example 5.2. (i) $n = 1$: $A = \{0\} \cup [1, 2] \cup (3, 4)$: $\mathrm{int}A = (1, 2) \cup (3, 4)$, $\mathrm{cl}A = \{0\} \cup [1, 2] \cup [3, 4]$, A neither closed, nor open, $\mathrm{bd}A = \{0, 1, 2, 3, 4\}$, $A' = [1, 2] \cup [3, 4]$.
(ii) $n = 2$: $A = \{(0, 2)\} \cup (\{1\} \times [0, 2])$: $\mathrm{int}A = \emptyset$, $\mathrm{cl}A = A$, A is closed, $\mathrm{bd}A = A$, $A' = \{1\} \times [0, 2]$.

5.2 Real-valued functions of several variables

Definition 5.10. Let $A \subseteq \mathbb{R}^n$. A function $f : A \to \mathbb{R}$ is called a *real-valued function of* n *variables.*

Example 5.3. (i) $V : [0, +\infty) \times [0, +\infty) \to \mathbb{R}$, $V(x_1, x_2) = \pi(x_1)^2 x_2$ or $V(x, y) = \pi x^2 y$ (the volume of a right circular cylinder of radius x and height y).
(ii) $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x_1, y_1, z_1) = (x_1)^2 + \sin(x_1 y_1 z_1) + y_1 z_1$ or $f(x, y, z) = x^2 + \sin(x\,y\,z) + y\,z$.
(iii) $f : \mathbb{R}^n \to \mathbb{R}$, $f(x_1, \ldots, x_n) = \|(x_1, \ldots, x_n)\| = \sqrt{(x_1)^2 + \ldots + (x_n)^2}$ (the Euclidean norm).

Limits of functions

Definition 5.11. Let $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$, $c \in A'$ and $L \in \overline{\mathbb{R}}$. We say that f *has limit* L *at* c if

$$\forall V \in \mathcal{V}(L), \exists U \in \mathcal{V}(c) \text{ such that } \forall x \in U \cap (A \setminus \{c\}) \text{ we have } f(x) \in V.$$

Remark 5.12. Let $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$, $c \in A'$.
(i) f has at most one limit at c.
(ii) f *has a limit at* c if $\exists L \in \overline{\mathbb{R}}$, $\forall V \in \mathcal{V}(L)$, $\exists U \in \mathcal{V}(c)$, $\forall x \in U \cap (A \setminus \{c\})$, $f(x) \in V$. In this case L is called the *limit of* f *at* c and we write

$$\lim_{x \to c} f(x) = L \quad \text{or} \quad \lim_{\substack{x_1 \to c_1 \\ \vdots \\ x_n \to c_n}} f(x_1, \ldots, x_n) = L \quad \text{or} \quad f(x) \to L \text{ as } x \to c.$$

We also sometimes say that $f(x)$ *approaches* L *as* x *approaches* c.
(iii) f *has no limit at* c if $\forall L \in \overline{\mathbb{R}}$, $\exists V \in \mathcal{V}(L)$, $\forall U \in \mathcal{V}(c)$, $\exists x \in U \cap (A \setminus \{c\})$, $f(x) \notin V$.

Theorem 5.4 (ε-δ characterization of limits)**.** *Let* $A \subseteq \mathbb{R}^n$*,* $f : A \to \mathbb{R}$*,* $c \in A'$*,* $L \in \mathbb{R}$*. Then*

(i) $\lim_{x \to c} f(x) = L \iff \forall \varepsilon > 0,\ \exists \delta = \delta(\varepsilon) > 0,\ \forall x \in A,\ 0 < \|x - c\| < \delta,\ |f(x) - L| < \varepsilon.$

(ii) $\lim_{x \to c} f(x) = +\infty\ (-\infty) \iff \forall \alpha \in \mathbb{R},\ \exists \delta = \delta(\alpha) > 0,\ \forall x \in A,\ 0 < \|x - c\| < \delta,\ f(x) > \alpha$ $(f(x) < \alpha)$.

Example 5.4. Let $f : \mathbb{R}^3 \setminus \{0_3\} \to \mathbb{R}$, $f(x, y, z) = \dfrac{x\,y\,z}{x^2 + y^2 + z^2}$. Note that $0_3 \in (\mathbb{R}^3 \setminus \{0_3\})'$. Then $\lim_{(x,y,z) \to 0_3} f(x, y, z) = 0$.
To see this, let $\varepsilon > 0$. Take $\delta = 2\varepsilon$. Then, $0 < \|(x, y, z) - 0_3\| = \sqrt{x^2 + y^2 + z^2} < \delta$ yields

$$|f(x, y, z) - 0| = \left| \frac{x\,y\,z}{x^2 + y^2 + z^2} \right| = \frac{|x\,y|}{x^2 + y^2 + z^2} |z| \le \frac{|x\,y|}{x^2 + y^2} |z| \le \frac{|z|}{2} \le \frac{\sqrt{x^2 + y^2 + z^2}}{2} < \frac{\delta}{2} = \varepsilon.$$

Theorem 5.5 (Sequential characterization of limits, Heine)**.** *Let* $A \subseteq \mathbb{R}^n$*,* $f : A \to \mathbb{R}$*,* $c \in A'$*,* $L \in \overline{\mathbb{R}}$*. Then*

$$\lim_{x \to c} f(x) = L \iff \forall \text{ sequence } (x^k) \text{ in } A \setminus \{c\} \text{ with } \lim_{k \to \infty} x^k = c \text{ we have that } \lim_{k \to \infty} f(x^k) = L.$$

Remark 5.13. (i) Since, by the above result, limits of functions of several variables can be characterized using limits of sequences, limit theorems and rules for functions of several variables can be derived from corresponding ones for sequences. For instance, there exists a Sandwich Theorem for functions of several variables.
(ii) If the limit L exists, then the same value L for the limit must be obtained along all paths to c. If there are paths to c which do not yield the same value, then the limit does not exist.

Example 5.5. Let $A = ((0, +\infty) \times \mathbb{R}) \setminus \{(1, 0)\}$, $f : A \to \mathbb{R}$, $f(x, y) = \dfrac{(x-1)^2 \ln x}{(x-1)^2 + y^2}$. Note that $(1, 0) \in A'$. Then $\lim_{(x,y) \to (1,0)} f(x, y) = 0$.

Note that for any $(x, y) \in A$,

$$0 \le \left| \frac{(x-1)^2 \ln x}{(x-1)^2 + y^2} \right| \le |\ln x|$$

and $\lim_{(x,y) \to (1,0)} |\ln x| = 0$. Apply the Sandwich Theorem.

Example 5.6. Let $f : \mathbb{R}^2 \setminus \{0_2\} \to \mathbb{R}$, $f(x, y) = \dfrac{x^2 - y^2}{x^2 + y^2}$. Note that $0_2 \in (\mathbb{R}^2 \setminus \{0_2\})'$. We show that f has no limit at 0_2.
To see this, consider the sequences (a^k) and (b^k) defined for $k \in \mathbb{N}^*$ by $a^k = (1/k, 1/k)$ and $b^k = (1/k, 0)$. Then $\lim_{k \to \infty} a^k = 0_2 = \lim_{k \to \infty} b^k$, but $\lim_{k \to \infty} f(a^k) = 0$ and $\lim_{k \to \infty} f(b^k) = 1$. By Theorem 5.5, f does not have a limit at 0_2.

Continuous functions

Definition 5.12. Let $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$, $c \in A$. We say that f is *continuous at c* if

$$\forall V \in \mathcal{V}\left(f(c)\right), \exists U \in \mathcal{V}(c) \text{ such that } \forall x \in U \cap A \text{ we have } f(x) \in V.$$

If B is a subset of A, we say that f is *continuous on B* if it is continuous at every point of B. If f is continuous on A, then f is simply called continuous.

Theorem 5.6 (Characterizations of continuity)**.** *Let $A \subseteq \mathbb{R}^n$, $f : A \to \mathbb{R}$, $c \in A$. Then f is continuous at c if and only if one of the following conditions are met:*

(i) $\forall \varepsilon > 0$, $\exists \delta = \delta(\varepsilon) > 0$, $\forall x \in A$, $\|x - c\| < \delta$, $|f(x) - f(c)| < \varepsilon$.

(ii) $\forall$ sequence (x^k) in A with $\lim\limits_{k\to\infty} x^k = c$ we have that $\lim\limits_{k\to\infty} f(x^k) = f(c)$.

Example 5.7. (i) Let $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \begin{cases} (x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}}, & \text{if } (x,y) \neq 0_2 \\ 0, & \text{if } (x,y) = 0_2. \end{cases}$

We show that f is continuous at 0_2.
To see this, let $\varepsilon > 0$. Take $\delta = \sqrt{\varepsilon}$. Then, $\|(x,y) - 0_2\| = \sqrt{x^2+y^2} < \delta$ yields

$$|f(x,y) - f(0_2)| = \left|(x^2+y^2)\sin\frac{1}{\sqrt{x^2+y^2}}\right| \leq x^2+y^2 < \delta^2 = \varepsilon.$$

(ii) Let $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, & \text{if } (x,y) \neq 0_2 \\ 0, & \text{if } (x,y) = 0_2. \end{cases}$

We show that f is not continuous at 0_2.
To see this, consider the sequence (a^k) defined for $k \in \mathbb{N}^*$ by $a^k = (1/k, 1/k)$. Then $\lim\limits_{k\to\infty} a^k = 0_2$ and $\lim\limits_{k\to\infty} f(a^k) = 1/2$. Since $f(0_2) = 0$, it follows that f is not continuous at 0_2.

Theorem 5.7. *Let $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}$, $a \in A$, $f : A \to B$ and $g : B \to \mathbb{R}$. If f is continuous at $a \in A$ an g is continuous at $f(a)$, then $g \circ f : A \to \mathbb{R}$ is continuous at a.*

Remark 5.14. (i) Polynomials in n several variables are continuous on $\mathbb{R}^n$.
$n = 2$: $P(x,y) = xy^2 + 7x^3y + y - 3$.
$n = 3$: $P(x,y,z) = 4x^2y^3 + 3x^2y^2z^2 - 5x + 4z + 1$.
(ii) Rational functions (a quotient of two polynomials) are continuous on their maximal domain of definition.
$n = 2$: $f : \mathbb{R}^2 \setminus \{(x,y) \in \mathbb{R}^2 \mid x + y = 0\} \to \mathbb{R}$, $f(x,y) = \dfrac{x^2+5y}{x+y}$.
$n = 3$: $f : \mathbb{R}^3 \setminus \{0_3\} \to \mathbb{R}$, $f(x,y,z) = \dfrac{x^3+x-y}{x^2+y^2+z^2}$.
(iii) Sums, products and quotients (when defined) of continuous real-valued functions of several varables are continuous.
(iv) One can construct continuous functions of several variables by taking, for instance, g in Theorem 5.7 to be an elementary function:
$f : \mathbb{R}^n \to \mathbb{R}$, $f(x_1,\ldots,x_n) = (x_1)^2 + \ldots + (x_n)^2$, $g : [0,\infty) \to \mathbb{R}$, $g(u) = \sqrt{u}$. Then $g \circ f : \mathbb{R}^n \to \mathbb{R}$, $(g \circ f)(x_1,\ldots,x_n) = \sqrt{(x_1)^2 + \ldots + (x_n)^2} = \|(x_1,\ldots,x_n)\|$ is continuous on $\mathbb{R}^n$.

Partial derivatives

Definition 5.13. Let $A \subseteq \mathbb{R}^n$, $c = (c_1,\ldots,c_n) \in \mathrm{int}A$ and $j \in \{1,\ldots,n\}$. A function $f : A \to \mathbb{R}$ is called *partially differentiable w.r.t. x_j at c* if the limit

$$\lim_{x_j \to c_j} \frac{f(c_1,\ldots,c_{j-1},x_j,c_{j+1},\ldots,c_n) - f(c_1,\ldots,c_n)}{x_j - c_j}$$

exists in $\mathbb{R}$. In this case, the above limit is called the *partial derivative of f w.r.t. x_j at c* and is denoted by $\dfrac{\partial f}{\partial x_j}(c)$ (or $f'_{x_j}(c)$, $D_j f(c)$).

If for all $j \in \{1, \dots, n\}$, f is partially differentiable w.r.t all variables x_j at c, then f is called *partially differentiable at c*. In this case, the vector

$$\left(\frac{\partial f}{\partial x_1}(c), \dots, \frac{\partial f}{\partial x_n}(c)\right) \in \mathbb{R}^n$$

is called the *gradient of f at c* and is denoted by $\nabla f(c)$.

If B is an open subset of A, we say that f is *partially differentiable w.r.t. x_j on B* if it is partially differentiable w.r.t. x_j at every point of B. In this case, the function

$$\frac{\partial f}{\partial x_j} : B \to \mathbb{R}, \quad x \in B \mapsto \frac{\partial f}{\partial x_j}(x) \in \mathbb{R}$$

is called the *partial derivative of f w.r.t. x_j on B*.

At the same tine, f is called *partially differentiable on B* if it is partially differentiable at every point of B. If A is open and f is partially differentiable on A, then f is simply called partially differentiable.

Remark 5.15. (i) Since $c \in \mathrm{int}A$, we can move a small distance in all directions from c while not leaving the set.

(ii) Partial differentiation means taking the ordinary derivative w.r.t. a single variable while keeping all other variables constant. Thus, we can apply all rules of differentiation.

Example 5.8. (i) Let $f : \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x^3 + x\sin(yz) + y^2 e^z$. Let $(x, y, z) \in \mathbb{R}^3$. Then

$$\frac{\partial f}{\partial x}(x, y, z) = 3x^2 + \sin(yz), \quad \frac{\partial f}{\partial y}(x, y, z) = xz\cos(yz) + 2ye^z, \quad \frac{\partial f}{\partial z}(x, y, z) = xy\cos(yz) + y^2 e^z.$$

The partial derivatives of f at the point $(1, 2, 0)$ are:

$$\frac{\partial f}{\partial x}(1, 2, 0) = 3, \quad \frac{\partial f}{\partial y}(1, 2, 0) = 4, \quad \frac{\partial f}{\partial z}(1, 2, 0) = 6.$$

Thus, the gradient of f at $(1, 2, 0)$ is $\nabla f(1, 2, 0) = (3, 4, 6) \in \mathbb{R}^3$.

(ii) Example 5.7.(ii) revisited: let $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & \text{if } (x, y) \neq 0_2 \\ 0, & \text{if } (x, y) = 0_2. \end{cases}$ Since

$$\lim_{x \to 0} \frac{f(x, 0) - f(0, 0)}{x - 0} = 0 \quad \text{and} \quad \lim_{y \to 0} \frac{f(0, y) - f(0, 0)}{y - 0} = 0,$$

we have that f is partially differentiable at 0_2 and $\dfrac{\partial f}{\partial x}(0, 0) = 0 = \dfrac{\partial f}{\partial y}(0, 0)$. Note that we already proved that f is not continuous at 0_2, thus partial differentiability at a given point does not imply continuity at that point. In fact, f is partially differentiable on $\mathbb{R}^2$ (HW: find the partial derivatives w.r.t. x and w.r.t. y on $\mathbb{R}^2$).

Remark 5.16. The function given in Example 5.7.(i) is continuous, partially differentiable, but its partial derivatives are not continuous at 0_2 (HW). If $A \subseteq \mathbb{R}^n$ is open, a function $f : A \to \mathbb{R}$ is called *continuously partially differentiable* if it is partially differentiable and all partial derivatives are continuous. In this case we write $f \in C^1(A)$.

Higher order partial derivatives

Definition 5.14. Let $A \subseteq \mathbb{R}^n$, $c \in \operatorname{int} A$, $i, j \in \{1, \ldots, n\}$ and $f : A \to \mathbb{R}$. We say that f is *twice partially differentiable w.r.t.* (x_i, x_j) *at* c if $\exists V \in \mathcal{V}(c)$, V open, $V \subseteq A$ such that f is partially differentiable w.r.t. x_i on V and the function

$$\frac{\partial f}{\partial x_i} : V \to \mathbb{R}, \quad x \in V \mapsto \frac{\partial f}{\partial x_i}(x) \in \mathbb{R} \tag{1}$$

is partially differentiable w.r.t. x_j at c. The partial derivative of the function (1) w.r.t. x_j at c is called the *second order partial derivative of* f *w.r.t.* (x_i, x_j) *at* c and is denoted by $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(c)$ (or $f''_{x_i x_j}(c)$). If $i = j$ we use the notation $\dfrac{\partial^2 f}{\partial x_i^2}(c)$ (or $f''_{x_i^2}(c)$). If for all $i, j \in \{1, \ldots, n\}$, f is twice partially differentiable w.r.t (x_i, x_j) at c, then f is called *twice partially differentiable at* c. Inductively, one can define partial derivatives of arbitrary order.

Remark 5.17. (i) $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(c) = \dfrac{\partial}{\partial x_j}\left(\dfrac{\partial f}{\partial x_i}\right)(c)$, $f''_{x_i x_j}(c) = \left(f'_{x_i}\right)'_{x_j}(c)$.
Note that f has n^2 second order partial derivatives.
(ii) Higher order partial derivatives w.r.t. two or more different variables are also called *mixed partial derivatives.*
(iii) Partial derivatives introduced in Definition 5.13 will also be called *first-order partial derivatives* (in order to distinguish them from higher order partial derivatives).
(iv) As in Definition 5.13, one can introduce the notions of twice partial differentiability and second order partial derivative (as a function) on open sets. In particular, if A is open, then f is called twice partially differentiable if f is twice partially differentiable at every point of A.

Remark 5.18. Mixed partial derivatives of a function are not always equal. Consider $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x,y) = \begin{cases} \frac{x^3 y}{x^2+y^2}, & \text{if } (x,y) \neq 0_2 \\ 0, & \text{if } (x,y) = 0_2. \end{cases}$
Since

$$\lim_{x\to 0} \frac{f(x,0) - f(0,0)}{x - 0} = 0 \quad \text{and} \quad \lim_{y \to 0} \frac{f(0,y) - f(0,0)}{y - 0} = 0,$$

we have that f is partially differentiable at 0_2 and $\dfrac{\partial f}{\partial x}(0,0) = 0 = \dfrac{\partial f}{\partial y}(0,0)$. For $(x,y) \in \mathbb{R}^2 \setminus \{0_2\}$ we have that

$$\frac{\partial f}{\partial x}(x,y) = \frac{x^2 y(x^2 + 3y^2)}{(x^2+y^2)^2} \quad \text{and} \quad \frac{\partial f}{\partial y}(x,y) = \frac{x^3(x^2 - y^2)}{(x^2+y^2)^2}.$$

Note that

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

while

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{x \to 0} \frac{\frac{\partial f}{\partial y}(x,0) - \frac{\partial f}{\partial y}(0,0)}{x - 0} = \lim_{x \to 0} \frac{x^5/x^4}{x} = 1.$$

Note that the mixed second order partial derivatives $\dfrac{\partial^2 f}{\partial y \partial x}$ and $\dfrac{\partial^2 f}{\partial x \partial y}$ are not continuous at 0_2.
If $A \subseteq \mathbb{R}^n$ is open, a function $f : A \to \mathbb{R}$ is called *twice continuously partially differentiable* if it is twice partially differentiable and all first and second order partial derivatives are continuous. In this case we write $f \in C^2(A)$.

Theorem 5.8 (Schwarz)**.** *Let $A \subseteq \mathbb{R}^n$ be open and $f \in C^2(A)$. Then for every $i, j \in \{1, \ldots, n\}$,*

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial^2 f}{\partial x_i \partial x_j}.$$

Let $A \subseteq \mathbb{R}^n$ be open, $c \in A$ and $f : A \to \mathbb{R}$. If f is twice partially differentiable at c, we can build the $n \times n$ matrix

$$H_f(c) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(c) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(c) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(c) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(c) & \frac{\partial^2 f}{\partial x_2^2}(c) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(c) \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(c) & \frac{\partial^2 f}{\partial x_n \partial x_2}(c) & \dots & \frac{\partial^2 f}{\partial x_n^2}(c) \end{pmatrix},$$

which is called the *Hessian matrix (or Hessian) of f at c* (denoted also by $Hess_f(c)$ or $\nabla^2 f(c)$). If f is twice partially differentiable, then we can consider the Hessian matrix at all points of A. Note that if $f \in C^2(A)$, by Theorem 5.8, $H_f(c)$ is symmetric at every $c \in A$.

Local extrema and partial derivatives

Definition 5.15. Let $A \subseteq \mathbb{R}^n$ and $f : A \to \mathbb{R}$. We say that f *attains a local maximum (local minimum)* at $c \in A$ if there exists $V \in \mathcal{V}(c)$ such that for every $x \in V \cap A$,

$$f(c) \geq f(x) \quad (f(c) \leq f(x)). \tag{2}$$

In this case c is called a *local maximum point (local minimum point)* of f. We say that f *attains a local extremum* at c if it attains either a local maximum or a local minimum at c. Local maximum points and local minimum points are called *local extremum points.*
If (2) holds for every $x \in A$, then we say that f attains its maximum (minimum) at c and we call c a *maximum point (minimum point)* of f (sometimes c is also said to be a *global maximum point (global minimum point)* of f). Global maximum points and global minimum points are called *global extremum points.*

Definition 5.16. A set $A \subseteq \mathbb{R}^n$ is called *bounded* if there exists $r > 0$ such that $A \subseteq B(0_n, r)$.

Theorem 5.9 (Maximum-Minimum Theorem, Weierstrass)**.** *Let $A \subseteq \mathbb{R}^n$ be nonempty, closed and bounded and let $f : A \to \mathbb{R}$ be continuous. Then f attains both its maximum and minimum on A.*

Theorem 5.10 (Fermat)**.** *Let $A \subseteq \mathbb{R}^n$ be open and $f : A \to \mathbb{R}$. If $c \in A$, f is partially differentiable at c and f attains a local extremum at c, then $\nabla f(c) = 0_n$.*

Definition 5.17. Let $A \subseteq \mathbb{R}^n$ be open and $f : A \to \mathbb{R}$. A point $c \in A$ at which f is partially differentiable is called a *stationary point* (or *critical point*) of f if $\nabla f(c) = 0_n$.

Remark 5.19. Local extremum points of a function which is defined on an open set and which is partially differentiable are found among its stationary points, but not all stationary points are local extremum points: take $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = x^2 - y^2$. Then 0_2 is a stationary point of f, but not a local extremum point of f.

Remark 5.20. Let $A \subseteq \mathbb{R}^n$ be closed, bounded and with nonempty interior. Note that $A = \mathrm{cl}A = \mathrm{bd}A \cup \mathrm{int}A$. Suppose that $f : A \to \mathbb{R}$ is continuous on A and partially differentiable on $\mathrm{int}A$. By Theorem 5.9, there exist $m = \min f(A)$ and $M = \max f(A)$. Since $\mathrm{bd}A$ is also a closed and bounded set, again by Theorem 5.9, there exist $m' = \min f(\mathrm{bd}A)$ and $M' = \max f(\mathrm{bd}A)$. Consider $S = \{x \in \mathrm{int}A \mid \nabla f(x) = 0\}$. Suppose that S is nonempty. If S is finite, then there exist $m'' = \min f(S)$ and $M'' = \max f(S)$. Thus, $m = \min\{m', m''\}$ and $M = \max\{M', M''\}$. If S is empty, then $m = m'$ and $M = M'$.

Definition 5.18. Let $C = (c_{ij})_{1 \leq i,j \leq n}$ be a symmetric $n \times n$ matrix of real numbers. The function $\Phi_C : \mathbb{R}^n \to \mathbb{R}$ defined by

$$\Phi_C(h) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} h_i h_j, \quad \forall h = (h_1, \dots, h_n) \in \mathbb{R}^n$$

is called the *quadratic form associated to C.*
We say that Φ_C (or, equivalently, C) is

- *positive definite (negative definite)*: if $\forall h \in \mathbb{R}^n \setminus \{0_n\}$, $\Phi_C(h) > 0$ ($\Phi_C(h) < 0$).
- *positive semidefinite (negative semidefinite)*: if $\forall h \in \mathbb{R}^n$, $\Phi_C(h) \geq 0$ ($\Phi_C(h) \leq 0$).
- *indefinite*: if $\exists a, b \in \mathbb{R}^n$, $\Phi_C(a) < 0 < \Phi_C(b)$.

Theorem 5.11 (Sylvester)**.** *Let $C = (c_{ij})_{1 \leq i,j \leq n}$ be a symmetric $n \times n$ matrix of real numbers. For every $k \in \{1, \dots, n\}$, let*

$$\Delta_k = \det(c_{ij})_{1 \leq i,j \leq k} = \begin{vmatrix} c_{11} & \dots & c_{1k} \\ \vdots & & \vdots \\ c_{k1} & \dots & c_{kk} \end{vmatrix}.$$

Then

(i) C is positive definite $\Leftrightarrow \Delta_k > 0$, $\forall k \in \{1, \dots, n\}$.

(ii) C is negative definite $\Leftrightarrow (-1)^k \Delta_k > 0$, $\forall k \in \{1, \dots, n\}$.

Theorem 5.12. *Let $A \subseteq \mathbb{R}^n$ be open, $c \in A$ and $f \in C^2(A)$. Then*

(i) if c is a local minimum (local maximum) point of f, then $\nabla f(c) = 0_n$ and $H_f(c)$ positive semidefinite (negative semidefinite).

(ii) if $\nabla f(c) = 0_n$ and $H_f(c)$ is positive definite (negative definite), then c is a local minimum (local maximum) point of f.

Remark 5.21. If $\nabla f(c) = 0_n$ and $H_f(c)$ is indefinite, then c is not a local extremum point of f.

5.3 Vector-valued functions of several variables

Let $n, m \in \mathbb{N}^*$, $m \geq 2$. For $j \in \{1, \dots, m\}$, consider the projection mapping $pr_j : \mathbb{R}^m \to \mathbb{R}$, $pr_j(y) = y_j$, $\forall y = (y_1, \dots, y_m) \in \mathbb{R}^m$.

Definition 5.19. Let $A \subseteq \mathbb{R}^n$. A function $f : A \to \mathbb{R}^m$ is called a *vector-valued function of n variables*. The *components of f* are the real-valued functions $f_1, \dots, f_m : A \to \mathbb{R}$ defined by $f_j = pr_j \circ f$, $\forall j \in \{1, \dots, m\}$ and we write $f = (f_1, \dots, f_m)$.

Remark 5.22. Let $A \subset \mathbb{R}^n$, $f = (f_1, \dots, f_m) : A \to \mathbb{R}^m$.
(i) If $c \in A'$ and $y^0 = (y^0_1, \dots, y^0_m) \in \mathbb{R}^m$, then $\lim\limits_{x \to c} f(x) = y^0 \Leftrightarrow \forall j \in \{1, \dots, m\}$, $\lim\limits_{x \to c} f_j(x) = y^0_j$.
(ii) If $c \in A$, then f is continuous at $c \Leftrightarrow \forall j \in \{1, \dots, m\}$, f_j is continuous.
(iii) If $c \in \text{int}A$, then f is partially differentiable at $c \Leftrightarrow \forall j \in \{1, \dots, m\}$, f_j is partially differentiable at c.