VSB – Technical University of Ostrava
Faculty of Electrical Engineering and Computer Science
Department of Applied Mathematics

Introduction to Differential Calculus of Several Variables

Jiří Bouchala, Bohumil Krajc

2022

EUROPEAN UNION
European Structural and Investment Funds
Operational Programme Research,
Development and Education

MINISTRY OF EDUCATION,
YOUTH AND SPORTS

The lecture notes 'Introduction to Differential Calculus of Several Variables' are based on the translation of the original Czech manuscript 'Matematika III pro bakalářské studium' by Jiří Bouchala available at `https://homel.vsb.cz/~bou10/archiv/ma3_bc.pdf`.

The translation was co-financed by the European Union and the Ministry of Education, Youth and Sports from the Operational Programme Research, Development and Education, project 'Technology for the Future 2.0', reg. no. CZ.02.2.69/0.0/0.0/18_058/0010212.

Ostrava, September 1, 2022.

Contents

Chapter 1

$\mathbb{R}^n$ as a vector and metric space

1.1 Calculation in $\mathbb{R}^n$, definition of the metric

We will denote the set of all ordered $n-$tuples of real numbers by the symbol $\mathbb{R}^n$. Elements of $\mathbb{R}^n$ will be denoted in the form

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

Real numbers $x_1, x_2, \dots, x_n$ (and $y_1, y_2, \dots, y_n$ respectively) are called <u>coordinates</u> or <u>components of the point</u> x (and y respectively)[1].

Define the following operations ($x, y \in \mathbb{R}^n$, $c \in \mathbb{R}$) in $\mathbb{R}^n$:

- $x + y := (x_1 + y_1, x_2 + y_2, \dots, x_n + y_n)$,
- $c \cdot x := (cx_1, cx_2, \dots, cx_n)$,

[1]We will omit using indices to distinguish coordinates of the vector, frequently. For example we will write

$$(x, y, z) \in \mathbb{R}^3, (u, v) \in \mathbb{R}^2, \dots .$$

$x - y := x + ((-1) \cdot y) = (x_1 - y_1, x_2 - y_2, \ldots, x_n - y_n)$.
(It is not difficult to show that $\mathbb{R}^n$ with the operations defined above represents the vector space.)

Moreover, define in the sequel:

- $x \cdot y := x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$ ... (euclidean) scalar product,
- $\|x\| := \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$ ... (euclidean) norm,
- $\varrho(x, y) := \|x - y\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2}$
  ... (euclidean) metric.

Example 1.

- $(2, 3, -1) + (0, 9, 12) = (2, 12, 11)$.
- $-6 \cdot (2, -1) = (-12, 6)$.
- $(1, 1, 1, 2) - (2, 2, 3, 4) = (-1, -1, -2, -2)$.
- $(1, 2) \cdot (3, -6) = -9$.
- $\|(2, 3)\| = \sqrt{13}$.
- $\varrho((1, 1), (-10, 0)) = \sqrt{122}$.

$\triangle$

Proposition 2. Metric ρ has the following properties for all $x, y, z \in \mathbb{R}^n$:

1. $\varrho(x, y) \in \mathbb{R}$, $\varrho(x, y) \geq 0$, $\varrho(x, y) = 0 \Leftrightarrow x = y$, (distinguish property)
2. $\varrho(x, y) = \varrho(y, x)$, (symmetry)
3. $\varrho(x, y) \leq \varrho(x, z) + \varrho(z, y)$. (triangle inequality)

Remark 3. Concern on the fact that for $n = 1$ (for $n = 2$ and 3, respectively) $\varrho(x, y)$ represents the usual distance of the points x, y on the real axis (on the real plane and the space respectively).

Remark 4. Let P is a nonempty set. Every mapping $\rho : P \times P \to \mathbb{R}$ for which three properties of the proposition above hold can be considered as an appropriate generalization of the notion of the distance. We will call such a mapping metric and the set P embedded with this metric as a metric space.

Since $\mathbb{R}^n$ is a metric space we are ready to state the notion of the convergence of the sequence in $\mathbb{R}^n$.

1.2 Convergence of sequences in $\mathbb{R}^n$

Remark 5. Similarly as in the one-dimensional case we will use the symbol (a_k) for sequence in $\mathbb{R}^n$. It means that for every sufficiently large $k \in \mathbb{N}$ there is the point $a_k \in \mathbb{R}^n$ assigned to k.

Definition 6. Let $b \in \mathbb{R}^n$ and the sequence (a_k) in $\mathbb{R}^n$ is given. Then (a_k) is said to be convergent to b ((a_k) has the limit b) if

$$\lim \varrho(a_k, b) = 0.$$

In this case we will write $\lim a_k = b$ or $a_k \to b$. ♡

Recall $\lim \varrho(a_k, b)$ represents the limit of the sequence of *real* numbers.

Proposition 7. (" convergence = convergence in components ")

Let $b \in \mathbb{R}^n$ and the sequence (a_k) in $\mathbb{R}^n$ is given. Denote by

$$a_k = (a_{k1}, a_{k2}, \dots, a_{kn}),\ b = (b_1, b_2, \dots, b_n)$$

coordinates of the points a_k, b. Then

$$\lim a_k = b$$

if and only if for every $i \in \{1, 2, \dots, n\}$:

$$\lim a_{ki} = b_i.$$

Example 8. The sequence $\left(k^2, 0, \frac{1}{k^2}\right)$ has not a limit since the related sequence (k^2) of the first coordinates has not the finite limit! $\triangle$

Exercise 9. Decide about the convergence of the sequences bellow. Compute the limits if exist:

- $\lim\left(\frac{1}{k}, (-1)^k, 3\right)$,
- $\lim\left(\frac{k^2}{3k-4k^2}, \left(\frac{1}{2}\right)^k\right)$,
- $\lim\left((1+\frac{1}{k})^k, \frac{2}{k^3}\sin(k^2+1)\right)$,
- $\lim\left(\sqrt[k]{k}, 4, k^2-k\right)$.

Chapter 2

Real function of n real variables

2.1 Definition, basic notions

Every mapping from $\mathbb{R}^n$ into $\mathbb{R}$ is said to be the real function of n real variables. Equivalently a function f is a rule that associates every element $x = (x_1, x_2, \ldots, x_n) \in Df \subset \mathbb{R}^n$ with exactly one value $f(x) = f(x_1, x_2, \ldots, x_n) \in Hf \subset \mathbb{R}$. ($Df$... the domain of f, Hf...the range of f). If f is a real function of n real variables, we write

$$f: \ \mathbb{R}^n \to \mathbb{R}.$$

Example 10.

- $f(x, y) = \sin(x + 2y),\ Df = \mathbb{R}^2,\ Hf = \langle -1, 1 \rangle$.
- $f(x, y, z) = \sqrt{x^2 + y^2 + z^2},\ Df = \mathbb{R}^3,\ Hf = \mathbb{R}^+ \cup \{0\}$.
- $f(x, y) = \begin{cases} 1, & \text{for } xy > 0, \\ -1, & \text{for } xy \leq 0, \end{cases} \quad Df = \mathbb{R}^2,\ Hf = \{1, -1\}$.
- $f(x, y) = 14,\ Df = \{(x, y) \in \mathbb{R}^2 :\ x^2 + y^2 \leq 9\},\ Hf = \{14\}$.

$\triangle$

2.2 Remark and convention.

Now we know that a function is determined by its domain and its rule which associates each element of the domain with exactly one value. We often determine a function $f:\ \mathbb{R}^n \to \mathbb{R}$ only by its rule; in this case the domain is a set of all elements in $\mathbb{R}^n$ for which the rule is meaningful. Let us illustrate this convention with the following example.

Example 11. Compute the domain of the function $f:\ \mathbb{R}^2 \to \mathbb{R}$ defined by the formula

$$f(x,y) := \sin(3x+y-6) + \sqrt{x+1} - \ln(y^2).$$

$$(x,y) \in Df \Leftrightarrow \big[x+1 \geq 0\ \wedge\ y^2 > 0\big],\ \text{hence}$$

$$Df = \{(x,y) \in \mathbb{R}^2:\ x \geq -1\ \wedge\ y \neq 0\}.$$

$\triangle$

Definition 12. A graph of a function $f:\ \mathbb{R}^n \to \mathbb{R}$ is defined as the following set: Graph $f :=$

$$= \{(x_1, x_2, \ldots, x_n, y) \in \mathbb{R}^{n+1}:\ (x_1, x_2, \ldots, x_n) \in Df\ \wedge\ y = f(x_1, x_2, \ldots, x_n)\}.$$

A contour line of a function f with the contour dimension $c \in \mathbb{R}$ is defined by

$$\mathrm{v}_f(c) := \{(x_1, x_2, \ldots, x_n) \in Df:\ f(x_1, x_2, \ldots, x_n) = c\}.$$

$\heartsuit$

2.3 Operations with functions

Definition 13. Let $f, g:\ \mathbb{R}^n \to \mathbb{R}$ and $c \in \mathbb{R}$. Functions

$$f+g,\ f-g,\ f\cdot g,\ \frac{f}{g},\ c\cdot f:\ \mathbb{R}^n \to \mathbb{R}$$

are defined by the following rules:

- $(f+g)(x) := f(x) + g(x)$,
- $(f-g)(x) := f(x) - g(x)$,
- $(f \cdot g)(x) := f(x) \cdot g(x)$,
- $\left(\frac{f}{g}\right)(x) := \frac{f(x)}{g(x)}$,
- $(c \cdot f)(x)" = cf(x)$.

$\heartsuit$

Definition 14. Consider $n+1$ functions

$$f: \ \mathbb{R}^m \to \mathbb{R}; \quad g_1, \ g_2, \ \ldots, \ g_m: \ \mathbb{R}^n \to \mathbb{R}.$$

A function $h: \ \mathbb{R}^n \to \mathbb{R}$ defined by the rule

$$h(x_1, \ldots, x_n) := f\big(g_1(x_1, \ldots, x_n), g_2(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n)\big)$$

is said to be <u>the function composed from the functions</u> f and $g_1, g_2, \ldots, g_m$. $\heartsuit$

Example 15. Let $f(u,v,w) := u^2 + 2v^2 + 3w^2$; $g_1(x,y) := x - y, g_2(x,y) := x + y, g_3(x,y) := 2x + y$. Then a function composed from the functions f and g_1, g_2, g_3 is the function

$$\begin{aligned} h(x,y) &:= f\big(g_1(x,y), g_2(x,y), g_3(x,y)\big) = (x-y)^2 + 2(x+y)^2 + 3(2x+y)^2 = \\ &= 15x^2 + 14xy + 6y^2. \end{aligned}$$

$\triangle$

Exercise 16.

1. Determine and draw in $\mathbb{R}^2$ domain of the function f defined by the formula $f(x,y) := \sqrt{9 - x^2 - y^2} - \sqrt{x^2 - y^2 - 1}$.

2. Sketch (in $\mathbb{R}^3$) graph of the function f defined bellow:

   (a) $f(x, y) := 4 - x^2 - y^2$,

   (b) $f(x, y) := 4 - x^2$,

   (c) $f(x, y) := 2x + 3y + 1$.

3. Determine and draw in $\mathbb{R}^2$ contour lines of the function f defined by

   (a) $f(x, y) := 1 - \frac{x^2}{4} - \frac{y^2}{9}$,

   (b) $f(x, y) := x^2 + y^2$,

   (c) $f(x, y) := y$,

   (d) $f(x, y) := 2x + 3y + 1$.

Chapter 3

Limit and continuity of functions of several variables

3.1 Limit of the function

Definition 17. Consider a point $x_0 \in \mathbb{R}^n$ and a sequence (x_k) in $\mathbb{R}^n$. Assume that for all sufficiently large $k \in \mathbb{N}$

$$x_k \in \mathbb{R}^n \setminus \{x_0\}$$

and

$$\lim x_k = x_0.$$

Then we will simply write

$$x_0 \neq x_k \to x_0.$$

♡

Definition 18. Let $f:\ \mathbb{R}^n \to \mathbb{R}$, $x_0 \in \mathbb{R}^n$ and $a \in \mathbb{R}^*$. Suppose that the following implication

$$x_0 \neq x_k \to x_0 \Rightarrow f(x_k) \to a$$

holds[1]. Then we will write

$$\lim_{x \to x_0} f(x) = a$$

and say that the function f takes the limit a at the point x_0. ♡

Proposition 19. Let $f : \mathbb{R}^n \to \mathbb{R}$, $x_0 \in \mathbb{R}^n$ and $a \in \mathbb{R}$. Then the following statements are true:

1. If the limit $\lim_{x \to x_0} f(x)$ exists then there is a positive δ such that

$$P(x_0, \delta) := \{x \in \mathbb{R}^n : \ 0 < \|x - x_0\| < \delta\} \subset Df$$
$$\ldots \underline{\text{deleted neighbourhood of}}\ x_0\ \underline{\text{with the radius}}\ \delta;$$

2.

$$\lim_{x \to x_0} f(x) = a \Leftrightarrow$$
$$\Leftrightarrow (\forall \varepsilon > 0)(\exists \delta > 0)(\forall x \in \mathbb{R}^n; 0 < \|x - x_0\| < \delta) : \ |f(x) - a| < \varepsilon;$$

3.

$$\lim_{x \to x_0} f(x) = +\infty \Leftrightarrow$$
$$\Leftrightarrow (\forall k \in \mathbb{R})(\exists \delta > 0)(\forall x \in \mathbb{R}^n; 0 < \|x - x_0\| < \delta) : \ f(x) > k;$$

4.

$$\lim_{x \to x_0} f(x) = -\infty \Leftrightarrow$$
$$\Leftrightarrow (\forall l \in \mathbb{R})(\exists \delta > 0)(\forall x \in \mathbb{R}^n; 0 < \|x - x_0\| < \delta) : \ f(x) < l.$$

Example 20.

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

△

[1] More specifically we mean that $f(x_k) \to a$ for every sequence (x_k) such that $x_0 \neq x_k \to x_0$

Proof. Define a function $f: \mathbb{R}^2 \to \mathbb{R}$ by the rule $f(x,y) := \frac{\sin(x^2+y^2)}{x^2+y^2}$. Then we can immediately confirm the validity of the following implications:

$$\big[(0,0) \neq (x_k, y_k) \to (0,0)\big] \Rightarrow {}^{2}$$

$$\Rightarrow \begin{bmatrix} x_k \to 0 \ \wedge \ x_k \to 0 \ \wedge \ y_k^2 + y_k^2 \neq 0 \\ \\ \text{for all sufficiently large } k \in \mathbb{N} \end{bmatrix} \Rightarrow$$

$$\Rightarrow \begin{bmatrix} x_k^2 \to 0 \ \wedge \ y_k^2 \to 0 \ \wedge \ x_k^2 + y_k^2 \neq 0 \\ \\ \text{for all sufficiently large } k \in \mathbb{N} \end{bmatrix} \Rightarrow$$

$$\Rightarrow \big[0 \neq z_k := x_k^2 + y_k^2 \to 0\big] \Rightarrow {}^{3}\Big[\frac{\sin z_k}{z_k} = f(x_k, y_k) \to 1\Big]$$

But it is exactly what we have to prove. □

Example 21. $\lim\limits_{(x,y)\to(0,2)} \left(y + \frac{1}{x}\right)$ does not exist. △

Proof. Define a function $g: \mathbb{R}^2 \to \mathbb{R}$ and a sequence $\big((x_k, y_k)\big)$ in $\mathbb{R}^2$ by the following rules:

$$g(x,y) := y + \frac{1}{x}, \quad (x_k, y_k) := \left(\frac{(-1)^k}{k}, 2\right).$$

Then we can see that

- $(0,2) \neq (x_k, y_k) \to (0,2)$,
- the sequence $\big(g(x_k, y_k)\big) = \big(2 + (-1)^k k\big)$ has not a limit.

Now, from the definition of the limit we conclude that the investigated limit does not exist. □

[2] See the theorem about convergence in components.
[3] We know that $\lim_{z\to 0} \frac{\sin z}{z} = 1$. It means under definition: $0 \neq z_k \to 0 \Rightarrow \frac{\sin z_k}{z_k} \to 1$.

3.2 Continuity of the function

Definition 22. A function $f: \mathbb{R}^n \to \mathbb{R}$ is said to be continuous at the point $x_0 \in \mathbb{R}^n$ if

$$\lim_{x \to x_0} f(x) = f(x_0).$$

♡

Proposition 23. Let $f: \mathbb{R}^n \to \mathbb{R}$ and $x_0 \in \mathbb{R}^n$. Then the following statements are true:

1. If f is continuous at x_0 then there is a positive δ such that

$$U(x_0, \delta) := \{x \in \mathbb{R}^n : \|x - x_0\| < \delta\} \subset Df$$

$$\ldots \underline{\text{neighbourhood of } x_0} \ \underline{\text{with the radius }} \delta;$$

2. f is continuous at x_0 if and only if

$$x_k \to x_0 \Rightarrow f(x_k) \to f(x_0).$$

Example 24.

1. A function $f: \mathbb{R}^2 \to \mathbb{R}$ given by the rule

$$f(x, y) := \begin{cases} \frac{\sin(x^2+y^2)}{x^2+y^2}, & \text{for } (x, y) \neq (0, 0), \\ 1, & \text{for } (x, y) = (0, 0), \end{cases}$$

is continuous at the point $(0, 0)$.

2. A function $g: \mathbb{R}^2 \to \mathbb{R}$ given by the rule

$$g(x, y) := \begin{cases} y + \frac{1}{x}, & \text{for } x \neq 0, \\ 1, & \text{for } x = 0, \end{cases}$$

is not continuous at the point $(0, 2)$.

$\triangle$

Theorem 25. Assume that functions $f, g:\ \mathbb{R}^n \to \mathbb{R}$ are continuous at the point $x_0 \in \mathbb{R}^n$. Then functions

$$f+g,\ f-g,\ f\cdot g,\ c\cdot f\ (c\in\mathbb{R})$$

are continuous at x_0. Moreover suppose that $g(x_0) \neq 0$. Then the division $\frac{f}{g}$ is continuous at x_0 as well. $\square$

Theorem 26. Consider the notation from the part devoted to the composed function. Let $h:\ \mathbb{R}^n \to \mathbb{R}$ is defined by

$$h(x) := f\big(g_1(x), g_2(x), \ldots, g_m(x)\big),$$

where $f:\ \mathbb{R}^m \to \mathbb{R}$, and $g_1,\ g_2,\ \ldots,\ g_m:\ \mathbb{R}^n \to \mathbb{R}$. Assume that functions $g_1, g_2, \ldots, g_m$ are continuous at x_0 and moreover f is continuous at $\big(g_1(x_0), g_2(x_0), \ldots, g_m(x_0)\big) \in \mathbb{R}^m$. Then h is also continuous at x_0.[4] $\square$

Example 27. A function $f:\ \mathbb{R}^2 \to \mathbb{R}$ given by the formula

$$f(x,y) := \sin(x+2y) - x^3 + \mathrm{e}^{x^2 y}$$

is continuous at every point $(x,y) \in \mathbb{R}^2$. $\triangle$

Definition 28. A <u>function</u> $f:\ \mathbb{R}^n \to \mathbb{R}$ is said to <u>be continuous on the set</u> $M \subset \mathbb{R}^n$ if [5]

$$M \ni x_k \to x_0 \in M \Rightarrow f(x_k) \to f(x_0).$$

$\heartsuit$

[4] The same in another words: Composition of continuous functions is continuous function again.

[5] By the notation $M \ni x_k \to x_0 \in M$ we mean that $x_0 \in M$ and $x_k \in M$ for all sufficiently large $k \in \mathbb{N}$.

Chapter 4

Derivative and differential of functions of several variables

4.1 Partial derivatives of the first order

Assume we want to study the behavior of a function $f:\ \mathbb{R}^2 \to \mathbb{R}$ defined on some neighborhood of the point $c = (c_1, c_2)$. It is very natural to consider appropriate cuts of the graph of this function. Hence the role of functions

$$g_1, g_2:\ \mathbb{R} \to \mathbb{R},\ g_1(x) := f(x, c_2), \quad g_2(y) := f(c_1, y)$$

seems to be crucial.

Suppose that g_1 admits finite derivative at c_1. Then this derivative is said to be the partial derivative with respect to x (or with respect to the first variable) of the function f at the point c. We will use the notation $\frac{\partial f(c)}{\partial x}$ for it.

Similarly if g_2 has finite derivative at c_2 we will speak about the partial derivative with respect to y (w. r. t. the second variable) at the point c and denote this number by the symbol $\frac{\partial f(c)}{\partial y}$.

Now we are ready to bring a generalization of these notions to the case of functions of several variables.

Definition 29. Let $f:\ \mathbb{R}^n \to \mathbb{R}$ and $k \in \{1, 2, \ldots, n\}$. Assume that the point

$$c = (c_1, c_2, \ldots, c_k, \ldots, c_n)$$

is an inner point of the domain Df (i. e. there exists $\delta > 0$ such that $U(c, \delta) \subset Df$[1]).

Consider a function $g_k:\ \mathbb{R} \to \mathbb{R}$ defined by the rule

$$g_k(x) := f(c_1, c_2, \ldots, c_{k-1}, x, c_{k+1}, \ldots, c_n).$$

If g_k has the finite derivative at the point c_k then this derivative is said to be the partial derivative of the function f with respect to the $k-$th variable at the point c and we will use the symbol

$$\frac{\partial f(c)}{\partial x_k} \text{ for it.}$$

(It means $\frac{\partial f(c)}{\partial x_k} := g_k'(c_k) \in \mathbb{R}$.)

Moreover a function $\frac{\partial f}{\partial x_k}:\ \mathbb{R}^n \to \mathbb{R}$ defined by the rule

$$\frac{\partial f}{\partial x_k}(x) := \frac{\partial f(x)}{\partial x_k}$$

represents a partial derivative of the function f with respect to the $k-$th variable.

♡

Example 30.

1. If

$$f(x, y) := \sin(2x + y)$$

then $\frac{\partial f}{\partial x}(\pi, 0) = g'(\pi)$, where $g(x) := f(x, 0) = \sin(2x)$. Because $g'(x) = 2\cos(2x)$ we conclude

$$\frac{\partial f}{\partial x}(\pi, 0) = 2\cos(2\pi) = 2.$$

[1]See the definition of the neighbourhood.

Similarly $\frac{\partial f}{\partial y}(\pi,0) = h'(0)$, where $h(y) := f(\pi,y) = \sin(2\pi+y) = \sin y$. Since $h'(y) = \cos y$, we have

$$\frac{\partial f}{\partial y}(\pi,0) = \cos 0 = 1.$$

2. Consider a function

$$f(x,y) = \begin{cases} 1, & \text{for } xy \neq 0, \\ 0, & \text{for } xy = 0\,. \end{cases}$$

Then

$$\frac{\partial f(0,0)}{\partial x} = 0 = \frac{\partial f(0,0)}{\partial y}\,.$$ [2]

3. If

$$f(x,y,z) := \sin x \cos(y+2z)$$

then

$$\begin{aligned} \frac{\partial f(x,y,z)}{\partial x} &= \cos x \cos(y+2z), \\ \frac{\partial f(x,y,z)}{\partial y} &= -\sin x \sin(y+2z), \\ \frac{\partial f(x,y,z)}{\partial z} &= -2\sin x \sin(y+2z). \end{aligned}$$

(The definition of the partial derivative w.r.t. the $k-$th variable bring an instruction for computing it at the same time. All variables except the $k-$th one must be seen as fixed parameters hence we simply compute the derivative of the function of one ($k-$th) variable).

$\triangle$

Exercise 31. Compute $\frac{\partial f}{\partial x}(x,y)$, $\frac{\partial f}{\partial y}(x,y)$ for the following functions:

[2] Concern on the fact that f is not continuous at the point $(0,0)$. Generally continuity of a function at a point c does not follow from the existence of partial derivatives at this point!

1. $f(x,y) := 2x^3y - 4xy^2 + 2x$;

2. $f(x,y) := \ln(x + \sqrt{x^2+y^2})$;

3. $f(x,y) := x^y$ (let us recall $x^y := \mathrm{e}^{y \ln x}$)

4.2 Directional derivative

Remark 32. Consider

$$f:\ \mathbb{R}^n \to \mathbb{R},$$
$$c = (c_1, c_2, \ldots, c_n) \in \mathbb{R}^n,$$
$$e_k := (0,\ldots,0, \underset{k\text{-th component}}{\to} 1, 0, \ldots, 0) \in \mathbb{R}^n.$$

Then[3]

$$\frac{\partial f(c)}{\partial x_k} = g_k'(c_k) = \lim_{t\to 0} \frac{g_k(c_k+t) - g_k(c_k)}{t} =$$
$$= \lim_{t\to 0} \frac{f(c_1,\ldots,c_{k-1},c_k+t,c_{k+1},\ldots,c_n) - f(c_1,\ldots,c_n)}{t} =$$
$$= \lim_{t\to 0} \frac{f(c+t\cdot e_k) - f(c)}{t}\,.$$

If we replace in the limit above the "direction" e_k by an arbitrarily chosen vector u (such that $||u|| = 1$) we obtain the derivative of the function f at the point c in the direction u.

Definition 33. Consider some function $f:\ \mathbb{R}^n \to \mathbb{R}$ and assume that $c \in \mathbb{R}^n$ is an inner point of a domain of f. Moreover fix some vector $u \in \mathbb{R}^n$, $\|u\| = 1$. Suppose the limit

$$\lim_{t\to 0} \frac{f(c+t\cdot u) - f(c)}{t}\,,$$

exists and belongs to $\mathbb{R}$. Then this limit is said to be <u>a derivative of the function</u>

[3] We use the notation from the definition of the partial derivative.

f at the point c in the direction u and it is denoted by the symbol

$$\frac{\mathrm{d}f(c)}{\mathrm{d}u}.$$

♡

Remark 34. From our investigation above we can observe that partial derivative under $k-$th variable represents a special case of the directional derivative, i. e.

$$\frac{\partial f(c)}{\partial x_k} = \frac{\mathrm{d}f(c)}{\mathrm{d}e_k}.$$

Remark 35. Computations of directional derivatives based on the limit from definition are brutal frequently. We promise to the reader that we will bring the more effective way for computing them in the future.

4.3 Partial derivatives of higher orders

Definition 36. Consider a function $f: \mathbb{R}^n \to \mathbb{R}$ and fix the couple of numbers $i, j \in \{1, 2, \ldots, n\}$. Let $\frac{\partial f}{\partial x_i}$ exists on some neighborhood of a point $c \in \mathbb{R}^n$. Suppose that the function $\frac{\partial f}{\partial x_i}$ admits a partial derivative with respect to the j variable at the point c. Then this derivative will be denoted by the symbol

$$\frac{\partial^2 f(c)}{\partial x_i \partial x_j}$$

and we will call it a partial derivative of the second order of the function f with respect to the $i-$th and $j-$th variable. (In the case $i = j$ we will briefly write $\frac{\partial^2 f(c)}{\partial x_i^2}$.)[4]

Generally a partial derivative of the $k-$order of f at c w.r.t. to variables $x_{i_1}, x_{i_2}, \ldots, x_{i_k}$

[4]It seems to be clear how we can define functions $\frac{\partial^2 f}{\partial x_i \partial x_j}$ and $\frac{\partial^2 f}{\partial x_i^2}$. Hence it is omitted in this text.

is given by the following rule:

$$\frac{\partial^k f(c)}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} := \frac{\partial \left(\frac{\partial^{k-1} f}{\partial x_{i_1} \dots \partial x_{i_{k-1}}} \right)(c)}{\partial x_{i_k}}, \qquad (i_1, \dots, i_k \in \{1, 2, \dots, n\}).$$

♡

Example 37.

1. Let

   $$f(x, y) := \sin(2x + 3y).$$

   Then

   $$\frac{\partial f}{\partial x}(x, y) = 2\cos(2x + 3y), \quad \frac{\partial f}{\partial y}(x, y) = 3\cos(2x + 3y),$$
   $$\frac{\partial^2 f}{\partial x^2}(x, y) = -4\sin(2x + 3y), \quad \frac{\partial^2 f}{\partial y^2}(x, y) = -9\sin(2x + 3y),$$
   $$\frac{\partial^2 f}{\partial x \partial y}(x, y) = -6\sin(2x + 3y) = \frac{\partial^2 f}{\partial y \partial x}(x, y).$$

2. Consider a function

   $$f(x, y) := \begin{cases} 1, & \text{for } x = 0, \\ 0, & \text{for } x \neq 0. \end{cases}$$

   Then $\frac{\partial f}{\partial y}(x, y) = 0$ (in $\mathbb{R}^2$) hence $\frac{\partial^2 f}{\partial y \partial x}(0, 0) = 0$.
   But from the other side $\frac{\partial f}{\partial x}(0, 0)$ does not exist and as a consequence $\frac{\partial^2 f}{\partial x \partial y}(0, 0)$ `does not exist`.

△

Theorem 38 (About commutability of partial derivatives)**.**
Let a function $f : \mathbb{R}^n \to \mathbb{R}$ admits partial derivatives $\frac{\partial f}{\partial x_i}, \frac{\partial f}{\partial x_j}, \frac{\partial^2 f}{\partial x_i \partial x_j} (i, j \in \{1, 2, \dots, n\})$ in some neighborhood of a point $c \in \mathbb{R}^n$. Moreover assume that a function $\frac{\partial^2 f}{\partial x_i \partial x_j}$ is continuous at the point c. Then $\frac{\partial^2 f(c)}{\partial x_j \partial x_i}$ exists and

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

□

Corollary 39. Assume that $M \subset \mathbb{R}^n$ is open (it means that every point of M has some neighborhood which is subset of M) and a function f admits `continuous` all partial derivatives up to the $k-$th order. Then these derivatives `are independent` with respect to `the sequence` of variables. They depend just on `the count of making derivatives` w. r. t. the given variable.

Exercise 40. Compute all partial derivatives of the second order of given functions:

1. $f(x, y) := \cos(2x + 3y)\sin(-3x)$;
2. $f(x, y) := \operatorname{arctg}(2x - y)$;
3. $f(x, y, z) := x^2 + 2y^3 + xyz^2$.

4.4 Differential

Definition 41. Assume $f:\ \mathbb{R}^n \to \mathbb{R}$ has continuous all partial derivatives of the first order at a point $c \in \mathbb{R}^n$. Then we say that f is (continuously) differentiable at the point c. Linear function $\mathrm{d}f_c:\ \mathbb{R}^n \to \mathbb{R}$ defined by the formula

$$\mathrm{d}f_c(h_1, \ldots, h_n) := \frac{\partial f(c)}{\partial x_1}h_1 + \frac{\partial f(c)}{\partial x_2}h_2 + \cdots + \frac{\partial f(c)}{\partial x_n}h_n$$

is said to be differential of f at c and vector

$$\operatorname{grad} f(c) := \left(\frac{\partial f(c)}{\partial x_1}, \frac{\partial f(c)}{\partial x_2}, \ldots, \frac{\partial f(c)}{\partial x_n}\right)$$

is gradient of f at c. ♡

Concern on the fact that if f is differentiable at c then

$$\mathrm{d}f_c(h) = \operatorname{grad} f(c) \cdot h\,, \text{for every } h \in \mathbb{R}^n\,.$$

Example 42.

1.
$$f(x,y) := \ln\sqrt{2x^2 - y^2},\ c = (3, -\sqrt{2});$$
$$\mathrm{d}f_c(h_1, h_2) = \frac{6}{16}h_1 + \frac{\sqrt{2}}{16}h_2.$$

2.
$$f(x,y,z) := x^{yz},\ c = (1, -2, 1);$$
$$\mathrm{d}f_c(h_1, h_2, h_3) = -2h_1.$$

$\triangle$

Exercise 43. Compute $\mathrm{d}f_c$:

1.
$$f(x,y) := \ln(x^2 + y^2),\ c = (1, 3);$$

2.
$$f(x,y,z,u) := \sin(x+y)\cos(z-u),\ c = (0,0,0,0).$$

Theorem 44. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function at a point $c \in \mathbb{R}^n$. Then

1.
$$\lim_{h \to (0,0,\ldots,0)} \frac{f(c+h) - f(c) - \mathrm{d}f_c(h)}{\|h\|} = 0;$$

2. f is continuous at c

3. If $u \in \mathbb{R}^n$ and $\|u\| = 1$ then $\frac{\mathrm{d}f(c)}{\mathrm{d}u}$ exists and
$$\frac{\mathrm{d}f(c)}{\mathrm{d}u} = \mathrm{grad} f(c) \cdot u;$$

4. Suppose grad $f(c) = (0, 0, \ldots, 0)$. Then $\frac{\mathrm{d}f(c)}{\mathrm{d}u} = 0$ for all $u \in \mathbb{R}^n$ such that $\|u\| = 1$.

If grad $f(c) \neq (0,0,\ldots,0)$ then the value of $\frac{\mathrm{d}f(c)}{\mathrm{d}u}$ is maximal for $u := \frac{\text{grad } f(c)}{\|\text{grad } f(c)\|} = u_1$ and minimal when $u := -\frac{\text{grad } f(c)}{\|\text{grad } f(c)\|} = u_2$. Moreover

$$\frac{\mathrm{d}f(c)}{\mathrm{d}u} = \|\text{grad } f(c)\|, \quad \frac{\mathrm{d}f(c)}{\mathrm{d}u_2} = -\|\text{grad} f(c)\|.$$

□

Example 45.

1.
$$f(x,y,z) := xyz, \ c = (5,1,2), \ u = \left(\frac{4}{5}, 0, -\frac{3}{5}\right);$$
$$\frac{\mathrm{d}f(c)}{\mathrm{d}u} = (2,10,5)\cdot\left(\frac{4}{5}, 0, -\frac{3}{5}\right) = 2\cdot\frac{4}{5} + 10\cdot 0 + 5\cdot\left(-\frac{3}{5}\right) = -\frac{7}{5}.$$

2.
$$f(x,y) := \arctan\frac{y}{x}, \ c = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \ u = \left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right);$$
$$\frac{\mathrm{d}f(c)}{\mathrm{d}u} = \left(-\frac{\sqrt{3}}{2}, \frac{1}{2}\right)\cdot\left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right) = -\frac{3}{4} + \frac{1}{4} = -\frac{1}{2}.$$

△

Exercise 46. Compute $\frac{\mathrm{d}f(c)}{\mathrm{d}u}$ for the following functions and points:

1.
$$f(x,y) := \ln(x^2 y), \ c = (1,4), \ u = (1,0);$$

2.
$$f(x,y) := \ln\sqrt{2x^2 - y^2}, \ c = (3, -\sqrt{2}), \ u = \left(\frac{\sqrt{3}}{2}, \frac{1}{2}\right).$$

4.5 Differentials of higher orders

Definition 47. Fix a point $c \in \mathbb{R}^n$ and consider a function $f:\ \mathbb{R}^n \to \mathbb{R}$ with continuous partial derivatives of the $k-$th order at c. Then the function

$$\mathrm{d}^k f_c:\ \mathbb{R}^n \to \mathbb{R}$$

defined by the formula

$$\mathrm{d}^k f_c(h_1, h_2, \ldots, h_n) := \sum_{i_1,\ldots,i_k=1}^{n} \frac{\partial^k f(c)}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}} h_{i_1} h_{i_2} \cdots h_{i_k}$$

is said to be the differential of the $k-$th order of the function f at the point c. ♡

Example 48. Let $f(x, y) := \sin(2x + y),\ c = (0, \pi)$. Then

1. $\mathrm{d}^1 f_c(h_1, h_2) = \mathrm{d}f_c(h_1, h_2) = -2h_1 - h_2,$
2. $\mathrm{d}^2 f_c(h_1, h_2) = 0,$
3. $\mathrm{d}^3 f_c(h_1, h_2) = 8h_1^3 + 12h_1^2 h_2 + 6h_1 h_2^2 + h_2^3.$

△

Remark 49. Observations and notes for computing higher-order differentials

Suppose that s function $f:\ \mathbb{R}^2 \to \mathbb{R}$ has at a point $c \in \mathbb{R}^2$ continuous all partial derivatives of the second order. Then (mixed partial derivatives of the second order are the same at c) we can write

$$\mathrm{d}^2 f_c(h_1, h_2) = \frac{\partial^2 f(c)}{\partial x^2} h_1^2 + \frac{\partial^2 f(c)}{\partial x \partial y} h_1 h_2 + \frac{\partial^2 f(c)}{\partial y \partial x} h_2 h_1 + \frac{\partial^2 f(c)}{\partial y^2} h_2^2 =$$

$$= \frac{\partial^2 f(c)}{\partial x^2} h_1^2 + 2\frac{\partial^2 f(c)}{\partial x \partial y} h_1 h_2 + \frac{\partial^2 f(c)}{\partial y^2} h_2^2 = \text{"} \left(\frac{\partial}{\partial x} h_1 + \frac{\partial}{\partial y} h_2 \right)^2 f(c)\text{"}\ .$$

For the general case we can conclude the similar observation. Assume that $f:\ \mathbb{R}^n \to \mathbb{R}$ has at $c \in \mathbb{R}^n$ continuous all partial derivatives up to the $k-$th order.

Then the following formal "equality" can be used for computing of the considered differential:

$$\mathrm{d}^k f_c(h_1, h_2, \ldots, h_n) = \text{"}\left(\frac{\partial}{\partial x_1}h_1 + \frac{\partial}{\partial x_2}h_2 + \cdots + \frac{\partial}{\partial x_n}h_n\right)^k f(c)\text{"}.$$

Let us show how this observation can be used practically.

Example 50. Compute $\mathrm{d}^3 f_c$ if $f(x, y) := y \ln x,\ c = (1, 3)$.

$$\mathrm{d}^3 f_c(h_1, h_2) = \text{"}\left(\frac{\partial}{\partial x}h_1 + \frac{\partial}{\partial y}h_2\right)^3 f(c)\text{"} =$$

$$= \text{"}\left(\frac{\partial^3}{\partial x^3}h_1^3 + 3\frac{\partial^3}{\partial x^2 \partial y}h_1^2 h_2 + 3\frac{\partial^3}{\partial x \partial y^2}h_1 h_2^2 + \frac{\partial^3}{\partial y^3}h_2\right)^3 f(c)\text{"} =$$

$$= \frac{\partial^3 f(c)}{\partial x^3}h_1^3 + 3\frac{\partial^3 f(c)}{\partial x^2 \partial y}h_1^2 h_2 + 3\frac{\partial^3 f(c)}{\partial x \partial y^2}h_1 h_2^2 + \frac{\partial^3 f(c)}{\partial y^3}h_2^3 =$$

$$= 6h_1^3 - 3h_1^2 h_2.$$

$\triangle$

Exercise 51. Compute $\mathrm{d}^k f_c$ if

1. $f(x, y) := \ln(1 + x)\ln(1 + y),\ c = (0, 0),\ k = 2;$
2. $f(x, y) := x^y,\ c = (1, 1),\ k = 3;$
3. $f(x, y, z) := xyz,\ c = (-1, 0, 1),\ k = 3.$

4.6 Derivatives of the composite functions

First of all, let us state the definition of the derivative of the composite function in general form. We will also discuss some useful special cases subsequently.

Theorem 52. Suppose that functions $g_1 : \mathbb{R}^n \to \mathbb{R}$, $g_2 : \mathbb{R}^n \to \mathbb{R}, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ have variables $x_1, x_2, \ldots, x_n$ and are differentiable at $c = (c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$. Let $d = (d_1, d_2, \ldots, d_m) = (g_1(c), g_2(c), \ldots, g_m(c))$ and assume that the function $f : \mathbb{R}^m \to \mathbb{R}$ with variables $y_1, y_2, \ldots, y_m$ is differentiable at d. Then the composite function $h : \mathbb{R}^n \to \mathbb{R}$,

$$h(x_1, x_2, \ldots, x_n) = f(g_1(x_1, x_2, \ldots, x_n), g_2(x_1, x_2, \ldots, x_n), \ldots, g_m(x_1, x_2, \ldots, x_n))$$

is differentiable at the point c and

$$\frac{\partial h(c)}{\partial x_l} = \frac{\partial f(d)}{\partial y_1} \cdot \frac{\partial g_1(c)}{\partial x_l} + \frac{\partial f(d)}{\partial y_2} \cdot \frac{\partial g_2(c)}{\partial x_l} + \cdots + \frac{\partial f(d)}{\partial y_m} \cdot \frac{\partial g_m(c)}{\partial x_l}, \tag{4.6.1}$$

for every $l \in \{1, 2, \ldots, n\}$. □

We can write the formula above in briefer form

$$\frac{\partial h(c)}{\partial x_l} = \sum_{j=1}^{m} \frac{\partial f(d)}{\partial y_j} \cdot \frac{\partial g_j(c)}{\partial x_l}, \text{ for every } l \in \{1, 2, \ldots, n\}.$$

Example 53. For the choice $f(y_1, y_2) := y_1^2 + y_2^2$, $g_1(x_1, x_2) := x_1 \cdot \cos x_2$, $g_2(x_1, x_2) := x_1 \cdot \sin x_2$ we have $h(x_1, x_2) = x_1^2 \cdot \cos^2 x_2 + x_1^2 \cdot \sin^2 x_2 = x_1^2$. Using notation

$$d = (d_1, d_2) = (g_1(c), g_2(c)) = (g_1(c_1, c_2), g_2(c_1, c_2)) = (c_1 \cdot \cos c_2, c_1 \cdot \sin c_2)$$

we obtain

$$\frac{\partial f(d)}{\partial y_1} = 2 \cdot d_1, \ \frac{\partial f(d)}{\partial y_2} = 2 \cdot d_2,$$

$$\frac{\partial g_1(c)}{\partial x_1} = \cos c_2, \ \frac{\partial g_2(c)}{\partial x_1} = \sin c_2, \quad \frac{\partial g_1(c)}{\partial x_2} = -c_1 \cdot \sin c_2, \ \frac{\partial g_2(c)}{\partial x_2} = c_1 \cdot \cos c_2$$

and compute

$$\frac{\partial h(c)}{\partial x_1} = 2 \cdot d_1 \cdot \cos c_2 + 2 \cdot d_2 \cdot \sin c_2 = 2 \cdot c_1 \cdot \cos^2 c_2 + 2 \cdot c_1 \cdot \sin^2 c_2 = 2c_1,$$

$$\frac{\partial h(c)}{\partial x_2} = -2 \cdot d_1 \cdot \sin c_2 + 2 \cdot d_2 \cdot \cos c_2 = 0.$$

The reader would be so kind and compute both derivatives of h directly. $\triangle$

Remark 54. Using notation $y = (y_1, y_{2,} \ldots, y_m)$, $x = (x_1, x_{2,} \ldots, x_n)$ (and assuming smoothness) we can write the following modification of the formula above (4.6.1):

$$\frac{\partial h(x_1, \ldots, x_n)}{\partial x_l} = \sum_{j=1}^{m} \frac{\partial f(g_1(x_1, \ldots, x_n), \ldots, g_m(x_1, \ldots, x_n))}{\partial y_j} \cdot \frac{\partial g_j(x_1, \ldots, x_n)}{\partial x_l},$$

for every $l \in \{1, 2, \ldots, n\}$.

The more comprehensive equivalent:

$$\frac{\partial h(x)}{\partial x_l} = \sum_{j=1}^{m} \frac{\partial f(g_1(x), \ldots, g_m(x))}{\partial y_j} \cdot \frac{\partial g_j(x)}{\partial x_l},$$

for every $l \in \{1, 2, \ldots, n\}$.

Now, let us have a look at aforementioned special cases:

Example 55. Let functions $g_1 : \mathbb{R} \to \mathbb{R}$, $g_2 : \mathbb{R} \to \mathbb{R}$ have the common variable x and be differentiable at c. Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ with the first variable denoted by y_1 and the second one denoted by y_2, differentiable at $d = (g_1(c), g_2(c))$, and let the composite function $h : \mathbb{R} \to \mathbb{R}$ be defined by the rule

$$h(x) := f(g_1(x), g_2(x)).$$

Then

$$h'(c) = \frac{\partial f(d)}{\partial y_1} \cdot g_1'(c) + \frac{\partial f(d)}{\partial y_2} \cdot g_2'(c).$$

For the choice $f(y_1, y_2) := y_1^2 \cdot y_2$, $g_1(x) := \sin x$, $g_2(x) := \mathrm{e}^x$ we have $h(x) = \sin^2 x \cdot \mathrm{e}^x$. Using notation $d = (d_1, d_2) = (g_1(c), g_2(c))$ we obtain

$$\frac{\partial f(d)}{\partial y_1} = 2 \cdot d_1 \cdot d_2, \ \frac{\partial f(d)}{\partial y_2} = d_1^2, \ g_1'(c) = \cos c, \ g_2'(c) = \mathrm{e}^c$$

and conclude from the formula above

$$h'(c) = 2 \cdot d_1 \cdot d_2 \cdot \cos c + d_1^2 \cdot \mathrm{e}^c = 2 \cdot \sin c \cdot \cos c \cdot \mathrm{e}^c + \sin^2 c \cdot \mathrm{e}^c.$$

△

Example 56. Let $f : \mathbb{R}^2 \to \mathbb{R}$ (given by some rule $f(y_1, y_2)$) be differentiable in every point of the plane. We want to study behaviour of f when a point moves on the axis of the first and the third quadrant ($y_1 = y_2$). So we can choose $g_1(x) := x$, $g_2(x) := x$. Thus $h(x) = f(x, x)$. Hence

$$h'(x) = \frac{\partial f(x,x)}{\partial y_1} \cdot 1 + \frac{\partial f(x,x)}{\partial y_2} \cdot 1 = \frac{\partial f(x,x)}{\partial y_1} + \frac{\partial f(x,x)}{\partial y_2}.$$

Furthermore, assume that f has continuous all partial derivatives of the second order in $\mathbb{R}^2$. Then we can use the formula for derivatives of the composed function once again and write

$$\begin{aligned} h''(x) &= \frac{\partial^2 f(x,x)}{\partial y_1^2} + \frac{\partial^2 f(x,x)}{\partial y_1 \partial y_2} + \\ &+ \frac{\partial^2 f(x,x)}{\partial y_2 \partial y_1} + \frac{\partial^2 f(x,x)}{\partial y_2^2} = \\ &= \frac{\partial^2 f(x,x)}{\partial y_1^2} + 2 \cdot \frac{\partial^2 f(x,x)}{\partial y_1 \partial y_2} + \frac{\partial^2 f(x,x)}{\partial y_2^2}. \end{aligned}$$

△

Example 57. Let $f : \mathbb{R}^2 \to \mathbb{R}$ (given by some rule $f(y_1, y_2)$) is differentiable in every point in $\mathbb{R}^2$. We want to study behaviour of f when a point moves on the unit circle, so $g_1(x) := \cos x$, $g_2(x) := \sin x$ seems to be natural choice. We have $h(x) = f(\cos x, \sin x)$. Hence

$$h'(x) = -\frac{\partial f(\cos x, \sin x)}{\partial y_1} \cdot \sin x + \frac{\partial f(\cos x, \sin x)}{\partial y_2} \cdot \cos x.$$

Furthermore, assume that f has continuous all partial derivatives of the second order in $\mathbb{R}^2$. Then we can use the formula for derivatives of the composite function once again and write

$$\begin{aligned}
h''(x) = {} & \frac{\partial^2 f(\cos x, \sin x)}{\partial y_1^2} \cdot \sin^2 x + \frac{\partial^2 f(\cos x, \sin x)}{\partial y_2^2} \cdot \cos^2 x + \\
& - 2 \cdot \frac{\partial^2 f(\cos x, \sin x)}{\partial y_1 \partial y_2} \cdot \sin x \cdot \cos x + \\
& - \frac{\partial f(\cos x, \sin x)}{\partial y_1} \cdot \cos x - \frac{\partial f(\cos x, \sin x)}{\partial y_2} \cdot \sin x.
\end{aligned}$$

△

Example 58. Suppose that functions $g_1 : \mathbb{R}^2 \to \mathbb{R}$, $g_2 : \mathbb{R}^2 \to \mathbb{R}$, $g_3 : \mathbb{R}^2 \to \mathbb{R}$ are differentiable at $c = (c_1, c_2)$. Assume that the function $f : \mathbb{R}^3 \to \mathbb{R}$ with variables denoted by y_1, y_2 and y_3 respectively is differentiable at the point $d = (d_1, d_2, d_3) = (g_1(c), g_2(c), g_3(c))$. Consider the composite function $h : \mathbb{R}^2 \to \mathbb{R}$, $h(x_1, x_2) := f(g_1(x_1, x_2), g_2(x_1, x_2), g_3(x_1, x_2))$. Then

$$\frac{\partial h(c)}{\partial x_1} = \frac{\partial f(d)}{\partial y_1} \cdot \frac{\partial g_1(c)}{\partial x_1} + \frac{\partial f(d)}{\partial y_2} \cdot \frac{\partial g_2(c)}{\partial x_1} + \frac{\partial f(d)}{\partial y_3} \cdot \frac{\partial g_3(c)}{\partial x_1},$$

$$\frac{\partial h(c)}{\partial x_2} = \frac{\partial f(d)}{\partial y_1} \cdot \frac{\partial g_1(c)}{\partial x_2} + \frac{\partial f(d)}{\partial y_2} \cdot \frac{\partial g_2(c)}{\partial x_2} + \frac{\partial f(d)}{\partial y_3} \cdot \frac{\partial g_3(c)}{\partial x_2}.$$

△

Example 59. Now, let us assume, that $f : \mathbb{R}^3 \to \mathbb{R}$ is differentiable everywhere in $\mathbb{R}^3$. If we want to study behaviour of f when a point moves on a unit ball surface, we can choose $g_1(x_1, x_2) := \cos x_1 \cos x_2$, $g_2(x_1, x_2) := \sin x_1 \cos x_2$, $g_3(x_1, x_2) := \sin x_2$. Using the formula from the previous example we obtain

$$\begin{aligned}
\frac{\partial h(x_1, x_2)}{\partial x_1} = {} & \frac{\partial f(\cos x_1 \cos x_2, \sin x_1 \cos x_2, \sin x_2)}{\partial y_1} \cdot (-\sin x_1 \cos x_2) + \\
& + \frac{\partial f(\cos x_1 \cos x_2, \sin x_1 \cos x_2, \sin x_2)}{\partial y_2} \cdot \cos x_1 \cos x_2,
\end{aligned}$$

similarly for $\frac{\partial h(x_1, x_2)}{\partial x_2}$. △

Exercise 60. Assuming smoothness of f write the related formula for derivatives of the related composite functions h at given points:

1. $h(x_1, x_2, x_3) := f(x_1^4 \cdot x_2, x_2 - 3x_3, \sin(x_1 + x_3)),\ (c_1, c_2, c_3) = (1, 2, -1);$
2. $h(r, t) := f(r \cdot \cos t, r \cdot \sin t),\ (c_1, c_2) = (0, \pi);$
3. $h(u, v) := f(u^3 - v^2, 2u - 3),\ (c_1, c_2) = (1, 2);$
4. $h(x) := f(\sin x, \sin^2 x, \sin^3 x),\ c = x \in \mathbb{R}.$

Chapter 5

Taylor theorem

Theorem 61 (Taylor)**.**
Fix $m \in \mathbb{N},\ c \in \mathbb{R}^n,\ f:\ \mathbb{R}^n \to \mathbb{R}$ and suppose there exists $\delta > 0$ such that

$$f \in C^{m+1}(U(c,\delta)).$$

Choose a vector $h = (h_1, h_2, \ldots, h_n) \in \mathbb{R}^n$ such that $c + h \in U(c,\delta)$. Then we can write

$$f(c+h) = T_m(c+h) + R_{m+1}(h),$$

where

$$T_m(c+h) = f(c) + \frac{1}{1!}\mathrm{d}f_c(h) + \frac{1}{2!}\mathrm{d}^2 f_c(h) + \cdots + \frac{1}{m!}\mathrm{d}^m f_c(h)$$

and

$$R_{m+1}(h) = \frac{1}{(m+1)!}\mathrm{d}^{m+1} f_{c+\xi\cdot h}(h) \quad \text{for some } \xi \in (0,1).$$

Moreover,

$$\lim_{h\to(0,\ldots,0)} \frac{R_{m+1}(h)}{\|h\|^m} = 0.$$

□

Remark 62. The formulas above can be explicitly written in the forms

$$T_m(c+h) = f(c) + \frac{1}{1!}\sum_{i=1}^{n} \frac{\partial f(c)}{\partial x_i} h_i + \frac{1}{2!}\sum_{i,j=1}^{n} \frac{\partial^2 f(c)}{\partial x_i \partial x_j} h_i h_j +$$

$$+\frac{1}{3!}\sum_{i,j,k=1}^{n} \frac{\partial^3 f(c)}{\partial x_i \partial x_j \partial x_k} h_i h_j h_k + \cdots + \frac{1}{m!}\sum_{i_1,\ldots,i_m=1}^{n} \frac{\partial^m f(c)}{\partial x_{i_1} \ldots \partial x_{i_m}} h_{i_1} \cdots h_{i_m},$$

$$R_{m+1}(h) = \frac{1}{(m+1)!}\sum_{i_1,\ldots,i_{m+1}=1}^{n} \frac{\partial^{m+1} f(c+\xi\cdot h)}{\partial x_{i_1} \ldots \partial x_{i_{m+1}}} h_{i_1} \cdots h_{i_{m+1}}.$$

Example 63. Choose $f(x, y) := \cos x \cos y$, $c = (\pi, 0)$, $m = 2$. Then we can compute that for every $h = (h_1, h_2) \in \mathbb{R}^2$

$$f(c+h) = f(\pi+h_1, h_2) = f(\pi, 0) + \frac{\partial f(\pi, 0)}{\partial x} h_1 + \frac{\partial f(\pi, 0)}{\partial y} h_2 +$$

$$+\frac{1}{2}\left(\frac{\partial^2 f(\pi, 0)}{\partial x^2} h_1^2 + 2\frac{\partial^2 f(\pi, 0)}{\partial x \partial y} h_1 h_2 + \frac{\partial^2 f(\pi, 0)}{\partial y^2} h_2^2\right) + R_3(h) =$$

$$= -1 + \frac{1}{2}h_1^2 + h_2^2 + R_3(h).$$

It means that for every "sufficiently close" point (x, y) to $(\pi, 0)$ we can write an approximation

$$f(x, y) \doteq -1 + \frac{1}{2}(x-\pi)^2 + \frac{1}{2}y^2 = T_2(x, y).$$

△

Definition 64. Let a function $f : \mathbb{R}^2 \to \mathbb{R}$ be differentiable at a point $c = (c_1, c_2) \in \mathbb{R}^2$. Then a plane

$$\tau := \{(x, y, z) \in \mathbb{R}^3 : z = f(c) + \frac{\partial f(c)}{\partial x}(x - c_1) + \frac{\partial f(c)}{\partial y}(y - c_2)\}$$

is said to be <u>the tangent plane of the graph of f at the point</u> $(c_1, c_2, f(c_1, c_2))$. (Observe that $\tau :\ z = f(c) + \mathrm{d}f_c(x - c_1, y - c_2) = T_1(x, y)$.) ♡

Example 65. Deduce from the example above that a plane

$$\vartheta := \{(x, y, z) \in \mathbb{R}^3 : \ z = -1\}$$

represents the tangent plane of the graph of the function $f(x, y) := \cos x \cos y$ at the point $(\pi, 0, -1)$. △

Exercise 66. Find equations of tangent planes for given functions and points.

1. $f(x, y) := x^2 + y^2, \ c = (c_1, c_2) = (1, 1)$;
2. $f(x, y) := x^3 - y^2 + 2x - 3, \ c = (c_1, c_2) = (1, 2)$.

Chapter 6

Implicitly defined function

Consider a function $f:\ \mathbb{R}^2 \to \mathbb{R}$ and the set

$$M = \left\{(x,y) \in \mathbb{R}^2 :\ f(x,y) = 0\right\}.$$

M represents the contour line of f and we will concern our attention on the following problem. Is it possible to describe M (or its part, at least) as a graph of some function $g:\ \mathbb{R} \to \mathbb{R}$, $y = g(x)$? Let us remark, we will use (not so much correct) notation "$y = g(x)$" in this part of the text. This notation helps us intuitively to understanding of the problem from the another point of view:
Is it possible to express from the "equation" $f(x,y) = 0$ y "as a function" of x?
Examples bellow show to us that generally the structure of M can be completely different.

Example 67.

1. $f(x,y) := x^2 + y^2 + 1; \quad M = \emptyset.$

2. $f(x,y) := x^2 + y^2; \quad M = \{(0,0)\}.$

3. $f(x,y) := 6x + 2y - 3; \quad M$ is line $\left\{(x,y) \in \mathbb{R}^2 :\ y = \frac{3-6x}{2}\right\}$
   (we can write $g(x) := \frac{3-6x}{2}$).

4. $f(x,y) := \sqrt{x^2 + 4xy + 4y^2} - x - 2y$; M is the half-plane $\left\{(x,y) \in \mathbb{R}^2 :\ y \geq -\frac{x}{2}\right\}$

(there is infinitively many different possibilities for chosing a function $y = g(x)$ such that its graph "lies" in M).

5. $f(x,y) := \sqrt{x^2+4xy+4y^2} - x - 2y$; M is the half-plane $\left\{(x,y) \in \mathbb{R}^2 : \ y \geq -\frac{x}{2}\right\}$ (there is infinitively many different possibilities for chosing a function $y = g(x)$ such that its graph "lies" in M).

6. $f(x,y) := x^2 + y^2 - 1$; M is the circle $\{(x,y) \in \mathbb{R}^2 : \ x^2+y^2 = 1\}$ (and there is infinitively many different possibilities for chosing a function $y = g(x)$ such that its graph "lies" in M again. However observe "the difference" between examples (4) and (5) ...).

$\triangle$

The examples above advice to us to concern not only on the existence of "adequated" function g but also on its uniqueness (locally, at least). We will investigate the continuity and the differentiability of g also. Let us come back to the Example (5) and choose $(a,b) \in M$ (i. e. $a^2 + b^2 = 1$) in such a way that $b \neq 0$. Moreover take $\delta > 0$ such that $(a-\delta, a+\delta) \subset (-1,1)$. Then it is easy to see that there exists uniquely one continuous (and differentiable) function $y = g(x)$ defined in $(a-\delta, a+\delta)$ and such that

$$f(x, g(x)) = 0, \quad \text{for all } x \in (a-\delta, a+\delta)$$

(i. e. Graph$g \subset M$) and

$$g(a) = b.$$

(Obviously $g(x) = \sqrt{1-x^2}$, for $b > 0$ and $g(x) = -\sqrt{1-x^2}$, for $b < 0$.)

Problems stay in neighbourhoods of the points $(-1,0)$,$(-1,0)$, the points with vertical "tangent lines". In this situation any proper function $y = g(x)$ does not exists [1].

Observation. Assuming that M has in some point $(a,b) \in M$ "vertical tangent line" we conclude $\frac{\partial f(a,b)}{\partial y} = 0$. [2]

---

[1]Situation changes when we try to find the function $x = g(y)$

[2]The reader should analyze this observation in detail. It is concerned with "geometrical

Theorem 68 (about implicitly defined function).
Suppose that the set $\Omega \subset \mathbb{R}^2$ is open and $(a,b) \in \Omega$. Assume that for some $k \in \mathbb{N}$ the function $f:\ \mathbb{R}^2 \to \mathbb{R}$ satisfies the following conditions:

$$f \in C^k(\Omega), \qquad f(a,b) = 0, \qquad \frac{\partial f(a,b)}{\partial y} \neq 0.$$

Then there exist real numbers $\delta > 0$ and $\eta > 0$ and a function $g:\ \mathbb{R} \to \mathbb{R}$ such that:

(i) $g \in C^k(a-\delta, a+\delta)$,

(ii) $g(x) \in (b-\eta, b+\eta)$, for all $x \in (a-\delta, a+\delta)$,

(iii) For all $x \in (a-\delta, a+\delta)$ and every $y \in (b-\eta, b+\eta)$:

$$f(x,y) = 0 \iff y = g(x),$$

(iv) For every $x \in (a-\delta, a+\delta)$:

$$g'(x) = -\frac{\frac{\partial f}{\partial x}(x, g(x))}{\frac{\partial f}{\partial y}(x, g(x))}.$$

(We say that the function g is done implicitly by the equation $f(x,y) = 0$ and the condition $g(a) = b$.) □

Example 69. Let $f(x,y) := x - y + 4\sin y,\ (a,b) = (0,0)$. It is easy to verify that

$$f \in f \in C^\infty(\mathbb{R}^2),\ f(0,0) = 0,\ \frac{\partial f(0,0)}{\partial y} \neq 0.$$

Since all assuptions of the theorem above are satisfied we know that there exists a function g which is done implicitly by the equation $x - y + 4\sin y = 0$ and the condition $g(0) = 0$. We do not know an explicit formula for g. However we can

interpretation" *that gradient is perpendicular to the contourline.*

use the property (iv) from the Theorem 68 and compute the value $g'(0)$:

$$g'(0) = -\frac{\frac{\partial f}{\partial x}(0, g(0))}{\frac{\partial f}{\partial y}(0, g(0))} = -\frac{\frac{\partial f}{\partial x}(0, 0)}{\frac{\partial f}{\partial y}(0, 0)} = -\frac{1}{3}.$$

Let us show different way how to compute requiered derivatives now. Theorem 68 say to us that there exists $\delta > 0$ such that $g \in C^\infty(a - \delta, a + \delta)$ and for every $x \in (a - \delta, a + \delta)$:

$$h(x) := f(x, g(x)) = x - g(x) + 4\sin(g(x)) = 0.$$

From this equality we conclude immediately that for every $k \in \mathbb{N}$

$$h(x) = h'(x) = h''(x) = \cdots = h^{(k)}(x) = 0$$

holds for all $x \in (a - \delta, a + \delta)$. Hence we can make derivatives step by step and substitute $x = 0$. In such a way we can compute the value of every derivative of g at the point 0.[3] Let us practice it now:

$$h'(x) = 1 - g'(x) + 4\cos(g(x))g'(x) = 0,$$

therefore

$$1 - g'(0) + 4\cos(g(0))g'(0) = 1 - g'(0) + 4g'(0) = 0 \Rightarrow g'(0) = -\frac{1}{3}.$$

$$h''(x) = -g''(x) - 4\sin(g(x))(g'(x))^2 + 4\cos(g(x))g''(x) = 0,$$

so

$$-g''(0) - 4\sin(g(0))(g'(0))^2 + 4\cos(g(0))g''(0) = 0 \Rightarrow g''(0) = 0.$$

$\triangle$

[3]The knowledge of derivatives allows to us make Taylor polynomial and we are able approximate g for x near to 0. Thus we can find approximation of solutions of the equation $f(x, y) = 0$ on some neighbourhood of the point $(0, 0)$.

Exercise 70. A function $y = g(x)$ is done implicitly by the equation

$$y - x - \ln y = 0$$

and the condition $g(\mathrm{e} - 1) = \mathrm{e}$. Compute $g'(\mathrm{e} - 1),\ g''(\mathrm{e} - 1)$.

Chapter 7

Extrema of functions of several variables

7.1 Local extrema

Definition 71. Assume the existence of $\delta > 0$ such that for a function $f:\ \mathbb{R}^n \to \mathbb{R}$ and a point $c \in \mathbb{R}^n$ the following implication holds:

$$x \in P(c,\delta) = \{x \in \mathbb{R}^n :\ 0 < \|x - c\| < \delta\} \quad \Rightarrow \quad f(x) \leq f(c).$$

Then we say that f has a local maximum at c. When the more strong implication

$$x \in P(c,\delta) = \{x \in \mathbb{R}^n :\ 0 < \|x - c\| < \delta\} \quad \Rightarrow \quad f(x) < f(c)$$

arise we say that f has a strict local maximum at c. Similarly for the inequality $f(x) \geq f(c)$ or $f(x) > f(c)$, respectively we talk about local minimum or strict local minimum, respectively. ♡

Remark 72. Suppose that $f:\ \mathbb{R}^n \to \mathbb{R}$ has at $c \in \mathbb{R}^n$ a local extremum (i. e. a local maximum or a local minimum). Then there exists $\delta > 0$ such that

$$U(c,\delta) = \{x \in \mathbb{R}^n :\ \|x - c\| < \delta\} \subset Df.$$

Example 73.

1. The function $f(x,y) := x^2+y^2$ has a strict local minimum at the point $(0,0)$. Observe $\operatorname{grad} f\,(0,0) = (0,0)$.

2. The function $f(x,y) := \sqrt{x^2+y^2}$ has a strict local minimum at the point $(0,0)$. Let us note that f is not differentiable at $(0,0)$(Partial derivatives does not exist at $(0,0)$, actually).

3. The function $f(x,y) := x^3$ has not a local extremum at the point $(0,0)$ although $\operatorname{grad} f\,(0,0) = (0,0)$.

$\triangle$

Theorem 74 (necessary condition of the existence of a local extremum)**.** Suppose that $f:\ \mathbb{R}^n \to \mathbb{R}$ has a local extremum at a point $c \in \mathbb{R}^n$ and $\frac{\mathrm{d}f(c)}{\mathrm{d}u}$ exists (for some $u \in \mathbb{R}^n,\ \|u\| = 1$). Then

$$\frac{\mathrm{d}f(c)}{\mathrm{d}u} = 0.$$

$\square$

Corollary. If f is differentiable at c additionally then

$$\operatorname{grad} f\,(c) = (0,0,\ldots,0)\,.$$

(At this case c is said to be <u>a stationary point of</u> f)

Remark 75. Let us recall sufficient conditions for local extrema of a function $f:\ \mathbb{R} \to \mathbb{R}$ at its stationary point c (i. e. $f'(c) = 0$).

If $f''(c) > 0$ then f has a strict local minimum at c. When $f''(c) < 0$ then it has a strict local maximum at c. For functions of several variables the multidimensional analogy of this conditions holds. The sign of $f'(c)$ will be replacing by requirements about (positive or negative) definiteness of the quadratic form $\mathrm{d}^2 f_c$.

Reminder

Consider a quadratic form $\varphi:\ \mathbb{R}^n \to \mathbb{R}$, $\quad \varphi(h) := \sum\limits_{i,j=1}^{n} a_{ij}h_ih_j$ ($a_{ij} = a_{ji} \in \mathbb{R}$, for all $i, j \in \{1, 2, \ldots, n\}$).

It is said to be positive definite if for every $h \in \mathbb{R}^n \setminus \{(0, 0, \ldots, 0)\}$

$$\varphi(h) > 0.$$

When

$$\varphi(h) < 0$$

for all $h \in \mathbb{R}^n \setminus \{(0, 0, \ldots, 0)\}$ then we call it the negative definite quadratic form.

Finally if there are vectors $k, l \in \mathbb{R}^n$ such that

$$\varphi(k) < 0 < \varphi(l),$$

the quadratic form φ is indefinite.

Theorem 76 (sufficient conditions of the existence of a local extremum)**.**
Assume that $c \in \mathbb{R}^n$ is a stationary point of the function $f:\ \mathbb{R}^n \to \mathbb{R}$ and f has continuous all second partial derivatives at c. Then the following implications are true.

1. If the quadratic form $\mathrm{d}^2 f_c$ is positively definite then f has at c a strict local minimum.
2. If the quadratic form $\mathrm{d}^2 f_c$ is negatively definite then f has at c a strict local maximum.
3. If the quadratic form $\mathrm{d}^2 f_c$ is indefinite then f has not any local extremum at c.

□

It is time to recall a practical tool for recognizing of the type of definiteness of quadratic forms, now.

Theorem 77 (Sylvester criterion)**.**
Consider a quadratic form

$$\varphi : \ \mathbb{R}^n \to \mathbb{R}, \quad \varphi(h) := \sum_{i,j=1}^{n} a_{ij} h_i h_j \ \ (a_{ij} = a_{ji}).$$

Denote in the sequel main sub-determinants of the related matrix:

$$\Delta_1 = a_{11}, \quad \Delta_2 = \begin{vmatrix} a_{11}, & a_{12} \\ a_{21}, & a_{22} \end{vmatrix}, \quad \ldots, \quad \Delta_n = \begin{vmatrix} a_{11}, & a_{12}, & \ldots, & a_{1n} \\ a_{21}, & a_{22}, & \ldots, & a_{2n} \\ & & \ldots & \\ a_{n1}, & a_{n2}, & \ldots, & a_{nn} \end{vmatrix}.$$

Then we can form the following statements.

1. φ is positive definite if and only if

$$\Delta_1 > 0, \ \Delta_2 > 0, \ \ldots, \ \Delta_n > 0.$$

2. φ is negative definite if and only if

$$\Delta_1 < 0, \ \Delta_2 > 0, \ \ldots, \ (-1)^n \Delta_n > 0.$$

3. Whenever $\Delta_n \neq 0$ and φ is not positive definite neither negative definite then φ is the indefinite quadratic form.

$\square$

Example 78. Find all local extrema of the function

$$f(x, y) := x^3 + y^3 - 18xy + 2007.$$

Evidently, $f \in C^\infty(\mathbb{R}^2)$ hence f admits local extrema only at stationary points.

So we should to solve the following system of equations:

$$\frac{\partial f}{\partial x}(x,y) = 3x^2 - 18y = 0,$$

$$\frac{\partial f}{\partial y}(x,y) = 3y^2 - 18x = 0.$$

Solution of this system consists of two points:

$$c_1 := (0,0), \qquad c_2 := (6,6).$$

We need to decide about behavior of our function at these points. We start with computing of the matrix of the quadratic form $\mathrm{d}^2 f_{(x,y)}$ at a general point $(x,y) \in \mathbb{R}^2$:

$$\mathrm{d}^2 f_{(x,y)} : \qquad \begin{pmatrix} \frac{\partial^2 f(x,y)}{\partial x^2}, & \frac{\partial^2 f(x,y)}{\partial x \partial y} \\ \frac{\partial^2 f(x,y)}{\partial y \partial x}, & \frac{\partial^2 f(c)}{\partial y^2} \end{pmatrix} = \begin{pmatrix} 6x, & -18 \\ -18, & 6y \end{pmatrix}.$$

We are ready to state a matrices related to the second differentials at $c_1,\ c_2$, now:

$$\mathrm{d}^2 f_{c_1} : \begin{pmatrix} 0, & -18 \\ -18, & 0 \end{pmatrix}; \qquad \mathrm{d}^2 f_{c_2} : \begin{pmatrix} 36, & -18 \\ -18, & 36 \end{pmatrix}.$$

Using Sylvester criterion we see immediately that $\mathrm{d}^2 f_{c_1}$ is the indefinite quadratic form ($\Delta_1 = 0,\ \Delta_2 = -18^2 \neq 0$) and $\mathrm{d}^2 f_{c_2}$ represents the positive definite quadratic form ($\Delta_1 = 36 > 0, \quad \Delta_2 = 36^2 - 18^2 > 0$). Finally, we conclude that f has not a local extremum at $c_1 = (0,0)$ and it has a strict local minimum at $c_2 = (6,6)$. $\triangle$

Example 79. Find all local extrema of the function

$$f(x,y) := (x^2 + y^2)^8.$$

It is easy to see that $f \in C^\infty(\mathbb{R}^2)$ and the point $c := (0,0)$ represents the unique stationary point of f. State a matrix of the quadratic form:

$$\mathrm{d}^2 f_c : \begin{pmatrix} 0, & 0 \\ 0, & 0 \end{pmatrix}.$$

We can see that it is not possible to use the theorems above for decision about extremum. But we are able to decide about extremum directly. Since for every $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$

$$f(x, y) > f(0, 0) = 0,$$

our function f has a strict local minimum at the point $c = (0, 0)$. △

Exercise 80. Find all local extrema of given functions:

1. $f(x, y) := (x + 1)^2 + y^2$;
2. $f(x, y) := 2x^3 - xy^2 + 5x^2 + y^2$;
3. $f(x, y) := 8x^3 + y^3 - 12xy - 41$;
4. $f(x, y) := (x - 2y + 1)^4$;
5. $f(x, y) := (x - 2y + 1)^3$;
6. $f(x, y, z) := x^2 + y^2 + z^2 + 2x + 4y - 6z + \sqrt{3}$;
7. $f(x, y, z) := x^2 + (2y - 1)^2 + (z + 2)^2$.

7.2 Global extrema

Definition 81. Consider a function $f : \mathbb{R}^n \to \mathbb{R}$ and a set $M \subset Df \subset \mathbb{R}^n$. Suppose the existence of a point $c \in M$ such that

$$f(c) = \max\{f(x) : x \in M\} =: \max_{x \in M} f(x).$$

Then we will say that f acquires its (global) maximum on the set M at c.

If

$$f(c) = \min\{f(x) : x \in M\} =: \min_{x \in M} f(x),$$

for $c \in M$ we will say that f has its (global) minimum on the set M at c. ♡

Example 82.

1. The function
$$f(x, y) := x^2 + y^2 + 1$$
has the minimum on $\mathbb{R}^2$ at the point $(0, 0)$. Maximum with respect to $\mathbb{R}^2$ does not exist.

2. The function
$$f(x, y) := x + y$$
acquires on the set
$$M = \langle 0, 1\rangle \times \langle 0, 1\rangle$$
its maximum at the point $(1, 1)$ and minimum at the point $(0, 0)$.
(Concern on the fact that global extrema of f do not exist on the set $N = (0, 1) \times (0, 1)$.)

△

Definition 83. A set $M \subset \mathbb{R}^n$ is said to be closed if the set $\mathbb{R}^n \setminus M$ is open[1]. We say that a set $M \subset \mathbb{R}^n$ is bounded if there exists $\delta > 0$ such that $\|x\| < \delta$ for every $x \in M$[2]. Closed and bounded subsets of $\mathbb{R}^n$ are called compact. ♡

Example 84.

1. The sets
$$\emptyset,\ \{(x, y) \in \mathbb{R}^2 :\ |x| + |y| \le 1\},\ \{(x, y) \in \mathbb{R}^2 :\ |x| + |y| = 1\}$$
are compact.

2. The sets
$$\mathbb{R}^2,\ \{(x, y) \in \mathbb{R}^2 :\ |x| + |y| < 1\},\ \{(x, y) \in \mathbb{R}^2 :\ |x| = 1\}$$
are not compact.

---

[1]See at Corollary 39 for the definition of an open set.
[2]In other words M is bounded if there exists $\delta > 0$ such that $M \subset U((0, 0, \ldots, 0), \delta)$.

$\triangle$

Theorem 85 (Weierstrass)**.** Let $M \subset \mathbb{R}^n$ is a nonempty and compact set. Assume that $f: \mathbb{R}^n \to \mathbb{R}$ is continuous on M. Then the numbers

$$\min_{x \in M} f(x), \quad \max_{x \in M} f(x)$$

exist. $\square$

Example 86. Find global extrema of f on M [3]if

$$f(x, y) := x^3 + y^3 - 3xy; \quad M = \langle 0, 2 \rangle \times \langle -1, 2 \rangle .$$

Evidently, $f \in C^\infty(\mathbb{R}^2)$ and $\emptyset \neq M$ is the compact set. From Weierstrass theorem we immediately deduce the existence of extrema

$$\min_{(x,y) \in M} f(x, y) \quad \text{and} \quad \max_{(x,y) \in M} f(x, y).$$

We will continue by computing all "suspicious" points and evaluating function values of f at these points. Finally we will solve the problem simply by comparing these values[4].

(a) Suspicious points inside M. If f has an extremum on M at some point

$$c \in \mathrm{int} M := (0, 2) \times (-1, 2),$$

then f has at the same point c local extremum also. Since $f \in C^\infty(\mathbb{R}^2)$, we can state necessary conditions

$$\frac{\partial f(c)}{\partial x} = \frac{\partial f(c)}{\partial y} = 0.$$

Solving the adequate system of equations

$$\frac{\partial f(x, y)}{\partial x} = 3x^2 - 3y = 0, \qquad \frac{\partial f(x, y)}{\partial y} = 3y^2 - 3x = 0,$$

[3]Understand it as a task to find all points such that f takes its maximum on M and its minimum on M, respectively.

[4]Concern on the fact that the information about the existence of extrema is necessary for this approach.

we obtain points $(0,0) \notin \mathrm{int}M$, $(1,1) \in \mathrm{int}M$. Therefore we have only one suspicious point inside M:

$$\underline{c_1 = (1,1)}.$$

(b) Suspicious points on the boundary of M. We will divide the boundary of our square into four parts.

(b1) Assume that f has its extremum on M at the point

$$c \in \partial M_1 := \langle 0,2\rangle \times \{-1\}.$$

Then $f(c)$ is the extremum of the set $\{f(x,-1):\ x \in \langle 0,2\rangle\}$ at the same time. It follows that one of the three cases will arise.

$c = \underline{c_2 = (0,-1)}$,

$c = \underline{c_3 = (2,-1)}$,

$c = (x,-1)$, where $x \in (0,2)$ represents a local extremum (and hence stationary point also) of the function

$$h_1(x) := f(x,-1) = x^3 - 1 + 3x.$$

But $h_1'(x) = 3x^2 + 3 > 0$ so we have not another suspicious points other then c_2, c_3 on ∂M_1.

An analogical approach can be applied on the rest parts of the boundary of M.

(b2) Candidates on extrema on

$$\partial M_2 := \langle 0,2\rangle \times \{2\}$$

are points $\underline{c_4 = (0,2)}$, $\underline{c_5 = (2,2)}$ and points $c = (x,2)$, where $x \in (0,2)$ is a stationary point of the function

$$h_2(x) := f(x,2) = x^3 + 8 - 6x.$$

Computing that

$$h_2'(x) = 3x^2 - 6 = 0 \Leftrightarrow \left[x = -\sqrt{2} \notin (0,2) \ \vee \ x = \sqrt{2} \in (0,2)\right],$$

we obtain another suspicious point $\underline{c_6 = (\sqrt{2}, 2)}$.

(b3) On the segment

$$\partial M_3 := \{0\} \times \langle -1, 2\rangle$$

are suspicious points c_3 c_4 and points $c = (0, y)$, where $y \in (-1, 2)$ is a stationary point of the function

$$h_3(y) := f(0, y) = y^3.$$

Hence we must add the another candidate $\underline{c_7 = (0,0)}$.

(b4) The last part of the boundary is

$$\partial M_4 := \{2\} \times \langle -1, 2\rangle.$$

Suspicious points are c_3, c_5 and point $c = (2, y)$ where $y \in (-1, 2)$ is a stationary point of the function

$$h_4(y) := f(2, y) = 8 + y^3 - 6y.$$

The last candidate is therefore the point $\underline{c_8 = (2, \sqrt{2})}$.

(c) Finally we are ready to compare function values at critical points i. e. the numbers

$$\begin{aligned} f(c_1) &= f(1,1) = -1, & f(c_5) &= f(2,2) = 4, \\ f(c_2) &= f(0,-1) = -1, & f(c_6) &= f(\sqrt{2},2) \doteq 2,3, \\ f(c_3) &= f(2,-1) = 13, & f(c_7) &= f(0,0) = 0, \\ f(c_4) &= f(0,2) = 8, & f(c_8) &= f(2,\sqrt{2}) \doteq 2,3. \end{aligned}$$

Since the greatest value is 13 we conclude that f has its maximum on M at the point $c_3 = (2, -1)$. Similarly, f has its minimum on M at the points $c_1 = (1,1)$, $c_2 = (0,-1)$. △

Exercise 87. Find global extrema of a function f on the set M if

1. $f(x,y) := x^2y,\ M = \{(x,y) \in \mathbb{R}^2 :\ x^2 + y^2 \leq 1\};$
2. $f(x,y) := x^2 + y^3 - 2x,\ M = \langle 0,2\rangle \times \langle -1,1\rangle\,;$
3. $f(x,y) := x^2 + 2xy - 4x + 8y,\ M = \langle 0,1\rangle \times \langle 0,2\rangle\,;$
4. $f(x,y) := x^2 - y^2,\ M = \{(x,y) \in \mathbb{R}^2 :\ x^2 + y^2 \leq 1\};$
5. $f(x,y) := x^2 - 2xy + y^2 + 13,\ M = \mathbb{R}^2;$
6. $f(x,y,z) := (x-1)^2 + (y-5)^5 + z^4,\ M = \{(x,y,z) \in \mathbb{R}^3 :\ y \geq 5\}.$

Bibliography

[1] J. Bouchala, M. Sadowská: Mathematical Analysis I,
https://homel.vsb.cz/ bou10/archiv/archiv.html

[2] H. Anton, I. Bivens, S. Davis: Calculus,
John Wiley & Sons, Inc., 2005

[3] W. Rudin: Principles of Mathematical Analysis, McGraw-Hill, Inc.,
1976