5. Functions of several variables

5.1. $\mathbf{R}^n$ as a metric and linear space.

Definition. The set $\mathbf{R}^n$, $n \in \mathbf{N}$, is the set of all ordered n-tuples of real numbers.

Definition. *Euclidean metric on* $\mathbf{R}^n$ is the function $\rho\colon \mathbf{R}^n \times \mathbf{R}^n \to [0, +\infty)$ defined by

$$\rho(\boldsymbol{x}, \boldsymbol{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

The number $\rho(\boldsymbol{x}, \boldsymbol{y})$ is called *distance of the point $\boldsymbol{x}$ from the point $\boldsymbol{y}$.*

Theorem 5.1 (properties of Euclidean metric)**.** *Euclidean metric ρ has the following properties:*

(i) $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n\colon \rho(\boldsymbol{x}, \boldsymbol{y}) = 0 \Leftrightarrow \boldsymbol{x} = \boldsymbol{y}$,

(ii) $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n\colon \rho(\boldsymbol{x}, \boldsymbol{y}) = \rho(\boldsymbol{y}, \boldsymbol{x})$, *(symmetry)*

(iii) $\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbf{R}^n\colon \rho(\boldsymbol{x}, \boldsymbol{y}) \leq \rho(\boldsymbol{x}, \boldsymbol{z}) + \rho(\boldsymbol{z}, \boldsymbol{y})$, *(triangle inequality)*

(iv) $\forall \boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^n, \forall \lambda \in \mathbf{R}\colon \rho(\lambda\boldsymbol{x}, \lambda\boldsymbol{y}) = |\lambda|\, \rho(\boldsymbol{x}, \boldsymbol{y})$, *(homogeneity)*

(v) $\forall \boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbf{R}^n\colon \rho(\boldsymbol{x} + \boldsymbol{z}, \boldsymbol{y} + \boldsymbol{z}) = \rho(\boldsymbol{x}, \boldsymbol{y})$. *(translation invariance)*

Definition. Let $\boldsymbol{x} \in \mathbf{R}^n$, $r \in \mathbf{R}, r > 0$. The set $B(\boldsymbol{x}, r)$ defined by

$$B(\boldsymbol{x}, r) = \{\boldsymbol{y} \in \mathbf{R}^n;\ \rho(\boldsymbol{x}, \boldsymbol{y}) < r\}$$

is called *open ball with radius r centered at $\boldsymbol{x}$.*

Definition. Let $M \subset \mathbf{R}^n$. We say that $\boldsymbol{x} \in \mathbf{R}^n$ is an *interior point of M*, if there exists $r > 0$ such that $B(\boldsymbol{x}, r) \subset M$. The set $M \subset \mathbf{R}^n$ is *open in* $\mathbf{R}^n$, if each point of M is an interior point of M. We say that M is *closed in* $\mathbf{R}^n$, if its complement is closed.

Theorem 5.2 (properties of open sets)**.**

(i) The empty set and $\mathbf{R}^n$ are open in $\mathbf{R}^n$.

(ii) Let sets $G_\alpha \subset \mathbf{R}^n$, $\alpha \in A \neq \emptyset$, be open in $\mathbf{R}^n$. Then $\bigcup_{\alpha \in A} G_\alpha$ is open in $\mathbf{R}^n$.

(iii) Let sets G_i, $i = 1, \dots, m$, be open in $\mathbf{R}^n$. Then $\bigcap_{i=1}^{m} G_i$ is open in $\mathbf{R}^n$.

Theorem 5.3 (properties of closed sets)**.**

(i) The empty set and $\mathbf{R}^n$ are closed in $\mathbf{R}^n$.

(ii) Let sets $F_\alpha \subset \mathbf{R}^n$, $\alpha \in A \neq \emptyset$, be closed in $\mathbf{R}^n$. Then $\bigcap_{\alpha \in A} F_\alpha$ is closed in $\mathbf{R}^n$.

(iii) Let sets F_i, $i = 1, \dots, m$, are closed in $\mathbf{R}^n$. Then $\bigcup_{i=1}^{m} F_i$ is closed in $\mathbf{R}^n$.

Definition. Let $M \subset \mathbf{R}^n$ and $\boldsymbol{x} \in \mathbf{R}^n$. We say that $\boldsymbol{x}$ is a *boundary point of M*, if for each $r > 0$ we have $B(\boldsymbol{x}, r) \cap M \neq \emptyset$ and $B(\boldsymbol{x}, r) \cap (\mathbf{R}^n \setminus M) \neq \emptyset$.

Boundary of M is the set of all boundary points of M (notation $\operatorname{bd} M$).

Closure of M is the set $M \cup \operatorname{bd} M$ (notation $\overline{M}$).

Interior of M is the set of all interior points of M (notation $\operatorname{int} M$).

5.2. **Continuous functions of several variables.**

Definition. Let $\boldsymbol{x}^j \in \mathbf{R}^n$ for each $j \in \mathrm{N}$ and $\boldsymbol{x} \in \mathbf{R}^n$. We say that a sequence $\{\boldsymbol{x}^j\}_{j=1}^{\infty}$ *converges to* $\boldsymbol{x}$, if $\lim_{j\to\infty} \rho(\boldsymbol{x}, \boldsymbol{x}^j) = 0$. The vector $\boldsymbol{x}$ is called *limit of the sequence* $\{\boldsymbol{x}^j\}_{j=1}^{\infty}$.

Theorem 5.4. *Let $\boldsymbol{x}^j \in \mathbf{R}^n$ for each $j \in \mathrm{N}$ and $\boldsymbol{x} \in \mathbf{R}^n$. The sequence $\{\boldsymbol{x}^j\}_{j=1}^{\infty}$ converges to $\boldsymbol{x}$ if and only if for each $i \in \{1, \dots, n\}$ the sequence of real numbers $\{x_i^j\}_{j=1}^{\infty}$ converges to the real number x_i.*

Definition. Let $M \subset \mathbf{R}^n$, $\boldsymbol{x} \in M$, and $f: M \to \mathbf{R}$. We say that f is *continuous at $\boldsymbol{x}$ with respect to M*, if we have

$$\forall \varepsilon \in \mathbf{R}, \varepsilon > 0\ \exists \delta \in \mathbf{R}, \delta > 0\ \forall \boldsymbol{y} \in B(\boldsymbol{x}, \delta) \cap M: f(\boldsymbol{y}) \in B(f(\boldsymbol{x}), \varepsilon).$$

We say that f is *continuous at the point $\boldsymbol{x}$*, it it is continuous at $\boldsymbol{x}$ with respect to a neighborhood of $\boldsymbol{x}$, i.e.,

$$\forall \varepsilon \in \mathbf{R}, \varepsilon > 0\ \exists \delta \in \mathbf{R}, \delta > 0\ \forall \boldsymbol{y} \in B(\boldsymbol{x}, \delta): f(\boldsymbol{y}) \in B(f(\boldsymbol{x}), \varepsilon).$$

Remark. Let $M \subset \mathbf{R}^n$, $\boldsymbol{x} \in M$, $f: M \to \mathbf{R}$, $g: M \to \mathbf{R}$, and $c \in \mathbf{R}$. If f and g are continuous at the point $\boldsymbol{x}$ with respect to M, then the functions cf, $f + g$ a fg are continuous at $\boldsymbol{x}$ with respect to M. If the function g is nonzero at each point of M, then also the function f/g is continuous at $\boldsymbol{x}$ with respect to M.

Theorem 5.5 (Heine)**.** *Let $M \subset \mathbf{R}^n$, $\boldsymbol{x} \in M$, and $f: M \to \mathbf{R}$. Then the following are equivalent.*

(i) The function f is continuous at $\boldsymbol{x}$ with respect to M.

(ii) For each sequence $\{\boldsymbol{x}^j\}_{j=1}^{\infty}$ such that $\boldsymbol{x}^j \in M$ pro $j \in \mathrm{N}$ a $\lim\limits_{j\to\infty} \boldsymbol{x}^j = \boldsymbol{x}$, we have $\lim\limits_{j\to\infty} f(\boldsymbol{x}^j) = f(\boldsymbol{x})$.

Remark. Let $r, s \in \mathrm{N}$, $M \subset \mathbf{R}^s$, $L \subset \mathbf{R}^r$, and $\boldsymbol{y} \in M$. Let $\varphi_1, \dots, \varphi_r$ are functions defined on M, which are continuous at $\boldsymbol{y}$ with respect to M and $[\varphi_1(\boldsymbol{x}), \dots, \varphi_r(\boldsymbol{x})] \in L$ for each $\boldsymbol{x} \in M$. Let $f: L \to \mathbf{R}$ be continuous at the point $[\varphi_1(\boldsymbol{y}), \dots, \varphi_r(\boldsymbol{y})]$ with respect to L. Then the composed function $F: M \to \mathbf{R}$ defined by

$$F(\boldsymbol{x}) = f\big(\varphi_1(\boldsymbol{x}), \dots, \varphi_r(\boldsymbol{x})\big), \quad \boldsymbol{x} \in M,$$

is continuous at $\boldsymbol{y}$ with respect to M.

Definition. Let $M \subset \mathbf{R}^n$ a $f: M \to \mathbf{R}$. We say that f is *continuous on M*, if it is continuous at each point $\boldsymbol{x} \in M$ with respect to M.

Remark. The projection $\pi_j: \mathbf{R}^n \to \mathbf{R}$, $\pi_j(\boldsymbol{x}) = x_j$, $1 \le j \le n$, are continuous on $\mathbf{R}^n$.

Definition. We say that a set $M \subset \mathbf{R}^n$ is *compact*, if for each sequence of elements of M there exists a convergent subsequence with limit in M.

Theorem 5.6 (characterization of compact subsets of $\mathbf{R}^n$)**.** *The set $M \subset \mathbf{R}^n$ is compact if and only if M is bounded and closed.*

Definition. Let $M \subset \mathbf{R}^n$, $\boldsymbol{x} \in M$, and f be a function defined at least on M, i.e., $M \subset D_f$. We say that f attains at the point $\boldsymbol{x}$

- *maximum on M*, if for every $\boldsymbol{y} \in M$ we have $f(\boldsymbol{y}) \leq f(\boldsymbol{x})$,
- *local maximum with respect to M*, if there exists $\delta > 0$ such that for every $\boldsymbol{y} \in B(\boldsymbol{x}, \delta) \cap M$ we have $f(\boldsymbol{y}) \leq f(\boldsymbol{x})$,
- *sharp local maximum with respect to M*, if there exists $\delta > 0$ such that for every $\boldsymbol{y} \in \big(B(\boldsymbol{x}, \delta) \setminus \{x\}\big) \cap M$ we have $f(\boldsymbol{y}) < f(\boldsymbol{x})$.

The notions *minimum*, *local minimum*, and *sharp local minimum* with respect to M are defined analogically.

Definition. We say that a function f attains at the point $\boldsymbol{x} \in \mathbf{R}^n$ *local maximum*, if $\boldsymbol{x}$ is a local maximum with respect to some ball centered at the point $\boldsymbol{x}$. Similarly one can define *local minimum*, *sharp local maximum* and *sharp local minimum*.

Theorem 5.7 (attaining extrema)**.** *Let $M \subset \mathbf{R}^n$ be a nonempty compact set and $f: M \to \mathbf{R}$ be continuous on M. Then f attains on M its maximum and minimum.*

Corollary 5.8. *Let $M \subset \mathbf{R}^n$ be a nonempty compact set and $f: M \to \mathbf{R}$ be continuous on M. Then f is bounded on M.*

Definition. We say that function $f: \mathbf{R}^n \to \mathbf{R}$ has at a point $\boldsymbol{a} \in \mathbf{R}^n$ limit equal $A \in \mathbb{R}^*$, if we have

$$\forall \varepsilon \in \mathbf{R}, \varepsilon > 0 \; \exists \delta \in \mathbf{R}, \delta > 0 \; \forall \boldsymbol{x} \in B(\boldsymbol{a}, \delta) \setminus \{\boldsymbol{a}\}: f(\boldsymbol{x}) \in B(A, \varepsilon).$$

Remark.

- Each function has at a given point at most one limit. We write $\lim_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = A$.
- The function f is continuous at $\boldsymbol{a}$ if and only if $\lim_{\boldsymbol{x} \to \boldsymbol{a}} f(\boldsymbol{x}) = f(\boldsymbol{a})$.
- For functions of several variables one can prove similar theorems as for functions of one variable (arithmetics, sandwich theorem, ...).

Theorem 5.9. *Let $r, s \in \mathrm{N}$, $\boldsymbol{a} \in M \subset \mathbf{R}^s$, $L \subset \mathbf{R}^r$, $\varphi_1, \dots, \varphi_r$ be functions defined on M such that $\lim_{\boldsymbol{x} \to \boldsymbol{a}} \varphi_j(\boldsymbol{x}) = b_j$, $j = 1, \dots, r$, and $\boldsymbol{b} = [b_1, \dots, b_r] \in L$. Let $f: L \to \mathbf{R}$ be continuous at the point $\boldsymbol{b}$. We define a function $F: M \to \mathbf{R}$ by*

$$F(\boldsymbol{x}) = f(\varphi_1(\boldsymbol{x}), \varphi_2(\boldsymbol{x}), \dots, \varphi_r(\boldsymbol{x})), \qquad \boldsymbol{x} \in M.$$

Then $\lim_{\boldsymbol{x} \to \boldsymbol{a}} F(\boldsymbol{x}) = f(\boldsymbol{b})$.

5.3. **Partial derivatives.**

Definition. Let f be a function of n variables, $j \in \{1, \dots, n\}$, $\boldsymbol{a} \in \mathbf{R}^n$. Then the number

$$\begin{aligned} \frac{\partial f}{\partial x_j}(\boldsymbol{a}) &= \lim_{t \to 0} \frac{f(\boldsymbol{a} + t\boldsymbol{e}^j) - f(\boldsymbol{a})}{t} \\ &= \lim_{t \to 0} \frac{f(a_1, \dots, a_{j-1}, a_j + t, a_{j+1}, \dots, a_n) - f(a_1, \dots, a_n)}{t} \end{aligned}$$

is called *partial derivatives (of first order) of function f according to j-th variable at the point $\boldsymbol{a}$* (if it exists).

Theorem 5.10 (necessary condition of existence of local extremum)**.** *Let $G \subset \mathbf{R}^n$ be an open set, $\boldsymbol{a} \in G$, and a function $f: G \to \mathbf{R}$ have at the point $\boldsymbol{a}$ local extremum. Then for each $j \in \{1, \dots, n\}$ we have:*

The partial derivative $\frac{\partial f}{\partial x_j}(\boldsymbol{a})$ either does not exit or is zero.

Definition. Let $G \subset \mathbf{R}^n$ be a nonempty open set. Let a function $f: G \to \mathbf{R}$ have at each point of the set G all partial derivatives continuous (i.e., function $\boldsymbol{x} \mapsto \frac{\partial f}{\partial x_j}(\boldsymbol{x})$ are continuous on G for each $j \in \{1, \dots, n\}$). Then we say that f is of the *class $\mathcal{C}^1$ on G*. The set of all these functions is denoted by $\mathcal{C}^1(G)$.

Remark. If $G \subset \mathbf{R}^n$ is a nonempty open set and and $f, g \in \mathcal{C}^1(G)$, then $f + g \in \mathcal{C}^1(G)$, $f - g \in \mathcal{C}^1(G)$, and $fg \in \mathcal{C}^1(G)$. If moreover for each $\boldsymbol{x} \in G$ we have : $g(\boldsymbol{x}) \neq 0$, then $f/g \in \mathcal{C}^1(G)$.

Proposition 5.11 (Lagrange)**.** *Let $n \in \mathbf{N}$, $I_1, \dots, I_n \subset \mathbf{R}$ be open intervals, $I = I_1 \times I_2 \times \cdots \times I_n$, $f \in \mathcal{C}^1(I)$, $\boldsymbol{a}, \boldsymbol{b} \in I$. Then there exist points $\boldsymbol{\xi}^1, \dots, \boldsymbol{\xi}^n \in I$ with $\xi^i_j \in \langle a_j, b_j \rangle$ for each $i, j \in \{1, \dots, n\}$, such that*

$$f(\boldsymbol{b}) - f(\boldsymbol{a}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{\xi}^i)(b_i - a_i).$$

Definition. Let $G \subset \mathbf{R}^n$ be an open set, $\boldsymbol{a} \in G$, and $f \in \mathcal{C}^1(G)$. Then the graph of the function

$$T: \boldsymbol{x} \mapsto f(\boldsymbol{a}) + \frac{\partial f}{\partial x_1}(\boldsymbol{a})(x_1 - a_1) + \frac{\partial f}{\partial x_2}(\boldsymbol{a})(x_2 - a_2) + \cdots + \frac{\partial f}{\partial x_n}(\boldsymbol{a})(x_n - a_n), \quad \boldsymbol{x} \in \mathbf{R}^n,$$

is called *tangent hyperplane* to the graph of the function f at the point $[\boldsymbol{a}, f(\boldsymbol{a})]$.

Theorem 5.12. *Let $G \subset \mathbf{R}^n$ be an open set, $\boldsymbol{a} \in G$, $f \in \mathcal{C}^1(G)$, and T be a function, such that its graph is the tangent hyperplane of the function f at the point $[\boldsymbol{a}, f(\boldsymbol{a})]$. Then*

$$\lim_{\boldsymbol{x} \to \boldsymbol{a}} \frac{f(\boldsymbol{x}) - T(\boldsymbol{x})}{\rho(\boldsymbol{x}, \boldsymbol{a})} = 0.$$

Theorem 5.13. *Let $G \subset \mathbf{R}^n$ be an open nonempty set and $f \in \mathcal{C}^1(G)$. Then f is continuous on G.*

Theorem 5.14. *Let $r, s \in \mathbf{N}$, $G \subset \mathbf{R}^s$, $H \subset \mathbf{R}^r$ be open sets. Let $\varphi_1, \dots, \varphi_r \in \mathcal{C}^1(G)$, $f \in \mathcal{C}^1(H)$ and $[\varphi_1(\boldsymbol{x}), \dots, \varphi_r(\boldsymbol{x})] \in H$ for each $\boldsymbol{x} \in G$. Then the composed function $F: G \to \mathbf{R}$ defined by*

$$F(\boldsymbol{x}) = f\big(\varphi_1(\boldsymbol{x}), \varphi_2(\boldsymbol{x}), \dots, \varphi_r(\boldsymbol{x})\big), \quad \boldsymbol{x} \in G,$$

is of the class $\mathcal{C}^1$ on G. Let $\boldsymbol{a} \in G$ and $\boldsymbol{b} = [\varphi_1(\boldsymbol{a}), \dots, \varphi_r(\boldsymbol{a})]$. Then for each $j \in \{1, \dots, s\}$ we have

$$\frac{\partial F}{\partial x_j}(\boldsymbol{a}) = \sum_{i=1}^{r} \frac{\partial f}{\partial y_i}(\boldsymbol{b}) \frac{\partial \varphi_i}{\partial x_j}(\boldsymbol{a}).$$

Definition. Let $G \subset \mathbf{R}^n$ be an open set, $\boldsymbol{a} \in G$, and $f \in \mathcal{C}^1(G)$. *Gradient of f at the point $\boldsymbol{a}$* is defined as the vector

$$\nabla f(\boldsymbol{a}) = \left[\frac{\partial f}{\partial x_1}(\boldsymbol{a}), \frac{\partial f}{\partial x_2}(\boldsymbol{a}), \dots, \frac{\partial f}{\partial x_n}(\boldsymbol{a})\right].$$

Definition. Let $G \subset \mathbf{R}^n$ be an open set, $\boldsymbol{a} \in G$, $f \in \mathcal{C}^1(G)$, and $\nabla f(\boldsymbol{a}) = \boldsymbol{o}$. Then the point $\boldsymbol{a}$ is called *stationary* (or also *critical*) *point* of the function f.

Definition. Let $G \subset \mathbf{R}^n$ be an open set, $f: G \to \mathbf{R}$, $i, j \in \{1, \dots, n\}$, and $\frac{\partial f}{\partial x_i}(\boldsymbol{x})$ exists for each $\boldsymbol{x} \in G$. Then *partial derivative of the second order of the function f according to i-th and j-th variable at the point $\boldsymbol{a} \in G$* is defined by

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{a}) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\boldsymbol{a}).$$

If $i = j$ then we use the notation

$$\frac{\partial^2 f}{\partial x_i^2}(\boldsymbol{a}).$$

Theorem 5.15. *Let $i, j \in \{1, \dots, n\}$ and let both partial derivatives $\frac{\partial^2 f}{\partial x_i \partial x_j}$ and $\frac{\partial^2 f}{\partial x_j \partial x_i}$ be continuous at a point $\boldsymbol{a} \in \mathbf{R}^n$. Then we have*

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

Definition. Let $G \subset \mathbf{R}^n$ be an open set and $k \in \mathbf{N}$. We say that a function f is of the *class $\mathcal{C}^k$ on G*, if all partial derivatives of f till k-th order are continuous on G. The set of all these functions is denoted by $\mathcal{C}^k(G)$. We say that a function f is of the *class $\mathcal{C}^\infty$ on G*, if all partial derivatives of all orders of f are continuous on G. The set of all functions of the class $\mathcal{C}^\infty$ on G is denoted by $\mathcal{C}^\infty(G)$.

5.4. **Implicit function theorem.**

Theorem 5.16 (implicit function theorem)**.** *Let $G \subset \mathbf{R}^{n+1}$ be an open set, $F: G \to \mathbf{R}$, $\tilde{\boldsymbol{x}} \in \mathbf{R}^n$, $\tilde{y} \in \mathbf{R}$, $[\tilde{\boldsymbol{x}}, \tilde{y}] \in G$. Suppose that*

(1) $F \in \mathcal{C}^1(G)$,
(2) $F(\tilde{\boldsymbol{x}}, \tilde{y}) = 0$,
(3) $\dfrac{\partial F}{\partial y}(\tilde{\boldsymbol{x}}, \tilde{y}) \neq 0$.

Then there exist a neighborhood $U \subset \mathbf{R}^n$ of the point $\tilde{\boldsymbol{x}}$ and a neighborhood $V \subset \mathbf{R}$ of the point $\tilde{y}$ such that for each $\boldsymbol{x} \in U$ there exists unique $y \in V$ with the property $F(\boldsymbol{x}, y) = 0$. If we denote this y by $\varphi(\boldsymbol{x})$, then the resulting function φ is in $\mathcal{C}^1(U)$ and

$$\frac{\partial \varphi}{\partial x_j}(\boldsymbol{x}) = -\frac{\frac{\partial F}{\partial x_j}(\boldsymbol{x}, \varphi(\boldsymbol{x}))}{\frac{\partial F}{\partial y}(\boldsymbol{x}, \varphi(\boldsymbol{x}))} \quad \text{for } \boldsymbol{x} \in U,\ j \in \{1, \dots, n\}.$$

Theorem 5.17 (implicit function theorem)**.** *Let $m, n \in \mathbf{N}$, $k \in \mathbf{N} \cup \{\infty\}$, $G \subset \mathbf{R}^{n+m}$ be an open set, $F_j\colon G \to \mathbf{R}$ for $j = 1, \dots, m$, $\tilde{\boldsymbol{x}} \in \mathbf{R}^n$, $\tilde{\boldsymbol{y}} \in \mathbf{R}^m$, $[\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}] \in G$. Suppose that*

(1) $F_j \in \mathcal{C}^k(G)$ for each $j \in \{1, \dots, m\}$,
(2) $F_j(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) = 0$ for each $j \in \{1, \dots, m\}$,
(3) $\begin{vmatrix} \frac{\partial F_1}{\partial y_1}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) & \dots & \frac{\partial F_1}{\partial y_m}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) & \dots & \frac{\partial F_m}{\partial y_m}(\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}) \end{vmatrix} \neq 0.$

Then there exist a neighborhood $U \subset \mathbf{R}^n$ of the point $\tilde{\boldsymbol{x}}$ and a neighborhood $V \subset \mathbf{R}^m$ of the point $\tilde{\boldsymbol{y}}$ such that for each $\boldsymbol{x} \in U$ there exists unique $\boldsymbol{y} \in V$ with the property $F_j(\boldsymbol{x}, \boldsymbol{y}) = 0$ for each $j \in \{1, \dots, m\}$. If we denote coordinates of this $\boldsymbol{y}$ by $\varphi_j(\boldsymbol{x})$, $j = 1, \dots, m$, then the resulting functions φ_j are in $\mathcal{C}^k(U)$.

Remark. The symbol in the condition (3Summer semesterDoc-Start) of Theorem 5.17Summer semesterDoc-Start is called *determinant*. The definition will presented later on.

For $m = 1$ we have $\begin{vmatrix} a \end{vmatrix} = a$, $a \in \mathbf{R}$.

For $m = 2$ we have $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$, $a, b, c, d \in \mathbf{R}$.

5.5. **Lagrange multiplier theorem.**

Theorem 5.18 (Lagrange multiplier theorem)**.** *Let $G \subset \mathbf{R}^2$ be an open set, $f, g \in \mathcal{C}^1(G)$, $M = \{[x, y] \in G;\ g(x, y) = 0\}$, and $[\tilde{x}, \tilde{y}] \in M$ be a point of local extremum of f with respect to the set M. Then at least one of the following conditions holds:*

(1) $\nabla g(\tilde{x}, \tilde{y}) = \boldsymbol{o}$,
(2) there exists $\lambda \in \mathbf{R}$ satisfying

$$\frac{\partial f}{\partial x}(\tilde{x}, \tilde{y}) + \lambda \frac{\partial g}{\partial x}(\tilde{x}, \tilde{y}) = 0,$$
$$\frac{\partial f}{\partial y}(\tilde{x}, \tilde{y}) + \lambda \frac{\partial g}{\partial y}(\tilde{x}, \tilde{y}) = 0.$$

Theorem 5.19 (Lagrange multiplier theorem)**.** *Let $m, n \in \mathbf{N}$, $m < n$, $G \subset \mathbf{R}^n$ be an open set, $f, g_1, \dots, g_m \in \mathcal{C}^1(G)$,*

$$M = \{\boldsymbol{z} \in G;\ g_1(\boldsymbol{z}) = 0, g_2(\boldsymbol{z}) = 0, \dots, g_m(\boldsymbol{z}) = 0\}$$

and let $\tilde{\boldsymbol{z}} \in M$ be a point of local extremum of f with respect to the set M. Then at least one of the following conditions holds:

(1) the vectors

$$\nabla g_1(\tilde{\boldsymbol{z}}), \nabla g_2(\tilde{\boldsymbol{z}}), \dots, \nabla g_m(\tilde{\boldsymbol{z}})$$

are linearly dependent,

(2) there exist $\lambda_1, \lambda_2, \dots, \lambda_m \in \mathbf{R}$ satisfying

$$\nabla f(\tilde{\boldsymbol{z}}) + \lambda_1 \nabla g_1(\tilde{\boldsymbol{z}}) + \lambda_2 \nabla g_2(\tilde{\boldsymbol{z}}) + \dots + \lambda_m \nabla g_m(\tilde{\boldsymbol{z}}) = \boldsymbol{o}.$$

5.6. **Concave and quasiconcave functions.**

Definition. Let $M \subset \mathbf{R}^n$. We say that M is *convex*, if we have

$$\forall \boldsymbol{x}, \boldsymbol{y} \in M \ \forall t \in \langle 0, 1\rangle : t\boldsymbol{x} + (1-t)\boldsymbol{y} \in M.$$

Definition. Let $M \subset \mathbf{R}^n$ be a convex set and a function f be defined on M. We say that f is

- *concave on M*, if

$$\forall \boldsymbol{a}, \boldsymbol{b} \in M \ \forall t \in \langle 0, 1\rangle : f(t\boldsymbol{a} + (1-t)\boldsymbol{b}) \geq t f(\boldsymbol{a}) + (1-t) f(\boldsymbol{b}),$$

- *strictly concave on M*, if

$$\forall \boldsymbol{a}, \boldsymbol{b} \in M, \boldsymbol{a} \neq \boldsymbol{b} \ \forall t \in (0, 1) : f(t\boldsymbol{a} + (1-t)\boldsymbol{b}) > t f(\boldsymbol{a}) + (1-t) f(\boldsymbol{b}).$$

Theorem 5.20. *Let a function f be concave on an open convex set $G \subset \mathbf{R}^n$. Then f is continuous on G.*

Theorem 5.21. *Let a function f be concave on a convex set $M \subset \mathbf{R}^n$. Then for each $\alpha \in \mathbf{R}$ the set $Q_\alpha = \{\boldsymbol{x} \in M;\ f(\boldsymbol{x}) \geq \alpha\}$ is convex.*

Theorem 5.22 (characterization of concave functions of the class $\mathcal{C}^1$)**.** *Let $G \subset \mathbf{R}^n$ be a convex open set and $f \in \mathcal{C}^1(G)$. Then the function f is convex on G if and only if we have*

$$\forall \boldsymbol{x}, \boldsymbol{y} \in G : f(\boldsymbol{y}) \leq f(\boldsymbol{x}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{x})(y_i - x_i).$$

Corollary 5.23. *Let $G \subset \mathbf{R}^n$ be a convex open set and $f \in \mathcal{C}^1(G)$ be concave on G. If a point $\boldsymbol{a} \in G$ is a stationary point of f, then $\boldsymbol{a}$ is a point of maximum of f with respect to G.*

Theorem 5.24 (characterization of strictly concave functions of the class $\mathcal{C}^1$)**.** *Let $G \subset \mathbf{R}^n$ be a convex open set and $f \in \mathcal{C}^1(G)$. Then the function f is strictly concave on G if and only if we have*

$$\forall \boldsymbol{x}, \boldsymbol{y} \in G, \boldsymbol{x} \neq \boldsymbol{y} : f(\boldsymbol{y}) < f(\boldsymbol{x}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\boldsymbol{x})(y_i - x_i).$$

Definition. Let $M \subset \mathbf{R}^n$ be a convex set and f be defined on M. We say that f is

- *quasiconcave on M*, if

$$\forall \boldsymbol{a}, \boldsymbol{b} \in M \ \forall t \in [0, 1] : f(t\boldsymbol{a} + (1-t)\boldsymbol{b}) \geq \min\{f(\boldsymbol{a}), f(\boldsymbol{b})\},$$

- *strictly quasiconcave on M*, if

$$\forall \boldsymbol{a}, \boldsymbol{b} \in M, \boldsymbol{a} \neq \boldsymbol{b},\ \forall t \in (0, 1) : f(t\boldsymbol{a} + (1-t)\boldsymbol{b}) > \min\{f(\boldsymbol{a}), f(\boldsymbol{b})\}.$$

Remark. Let $M \subset \mathbf{R}^n$ be a convex set and f be a function defined on M.

- Let f be concave on M. Then f is quasiconcave on M.
- Let f be strictly concave on M. Then f is strictly quasiconcave on M.

Theorem 5.25 (on uniqueness of extremum)**.** *Let f be a strictly quasiconcave function on a convex set $M \subset \mathbf{R}^n$. Then there exists at most one point of maximum of f.*

Corollary 5.26. *Let $M \subset \mathbf{R}^n$ be a convex, bounded, closed and nonempty set. Let f be continuous and strictly quasiconcave function on M. Then f attains its maximum on M in a unique point.*

Theorem 5.27 (characterization of quasiconcave functions via level sets)**.** *Let $M \subset \mathbf{R}^n$ be a convex set and f be defined on M. The function f is quasiconcave on M if and only if for each $\alpha \in \mathbf{R}$ the set $Q_\alpha = \{\boldsymbol{x} \in M;\ f(\boldsymbol{x}) \geq \alpha\}$ is convex.*

6. Matrix calculus

6.1. **Basic operations with matrices.**

Definition. The scheme

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix},$$

where $a_{ij} \in \mathbf{R}$, $i = 1, \dots, m$, $j = 1, \dots, n$, is called a *matrix of the type* $m \times n$. We write $(a_{ij})_{\substack{i=1..m \\ j=1..n}}$. A matrix of type $n \times n$ is called *square matrix of the order* n. The set of all matrices of the type $m \times n$ is denoted $M(m \times n)$.

Definition. Let

$$\mathbb{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

The n-tuple $(a_{i1}, a_{i2}, \dots, a_{in})$, where

$i \in \{1, 2, \dots, m\}$, is called *i-th row* of the matrix $\mathbb{A}$. The m-tuple $\begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$, where

$j \in \{1, 2, \dots, n\}$, is called *j-th column* matrix $\mathbb{A}$.

Definition. We say that two matrices are equal, if they are of the same type and the corresponding elements are equal, i.e., if $\mathbb{A} = (a_{ij})_{\substack{i=1..m \\ j=1..n}}$ and $\mathbb{B} = (b_{uv})_{\substack{u=1..r \\ v=1..s}}$, then $\mathbb{A} = \mathbb{B}$ if and only if $m = r$, $n = s$ and $a_{ij} = b_{ij}$ for every $i \in \{1, \dots, m\}$, $j \in \{1, \dots, n\}$.

Definition. Let $\mathbb{A}, \mathbb{B} \in M(m \times n)$, $\mathbb{A} = (a_{ij})_{\substack{i=1..m \\ j=1..n}}$, $\mathbb{B} = (b_{ij})_{\substack{i=1..m \\ j=1..n}}$, $\lambda \in \mathbf{R}$. The *sum of* $\mathbb{A}$ *and* $\mathbb{B}$ is defined by

$$\mathbb{A} + \mathbb{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}.$$

Product of a real number λ *and the matrix* $\mathbb{A}$ is defined by

$$\lambda\mathbb{A} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{pmatrix}.$$

Proposition 6.1 (basic properties)**.**

- $\forall \mathbb{A}, \mathbb{B}, \mathbb{C} \in M(m \times n) \colon \mathbb{A} + (\mathbb{B} + \mathbb{C}) = (\mathbb{A} + \mathbb{B}) + \mathbb{C}$, *(associativity)*
- $\forall \mathbb{A}, \mathbb{B} \in M(m \times n) \colon \mathbb{A} + \mathbb{B} = \mathbb{B} + \mathbb{A}$, *(commutativity)*
- $\exists! \mathbb{O} \in M(m \times n)\ \forall \mathbb{A} \in M(m \times n) \colon \mathbb{A} + \mathbb{O} = \mathbb{A}$, *(existence of the zero element)*
- $\forall \mathbb{A} \in M(m \times n)\ \exists \mathbb{C}_{\mathbb{A}} \in M(m \times n) \colon \mathbb{A} + \mathbb{C}_{\mathbb{A}} = \mathbb{O}$,
- $\forall \mathbb{A} \in M(m \times n)\ \forall \lambda, \mu \in \mathbf{R} \colon (\lambda\mu)\mathbb{A} = \lambda(\mu\mathbb{A})$,
- $\forall \mathbb{A} \in M(m \times n) \colon 1 \cdot \mathbb{A} = \mathbb{A}$,
- $\forall \mathbb{A} \in M(m \times n)\ \forall \lambda, \mu \in \mathbf{R} \colon (\lambda + \mu)\mathbb{A} = \lambda\mathbb{A} + \mu\mathbb{A}$,
- $\forall \mathbb{A}, \mathbb{B} \in M(m \times n)\ \forall \lambda \in \mathbf{R} \colon \lambda(\mathbb{A} + \mathbb{B}) = \lambda\mathbb{A} + \lambda\mathbb{B}$.

Definition. Let $\mathbb{A} \in M(m \times n)$, $\mathbb{A} = (a_{is})_{\substack{i=1..m \\ s=1..n}}$, $\mathbb{B} \in M(n \times k)$, $\mathbb{B} = (b_{sj})_{\substack{s=1..n \\ j=1..k}}$. Then *the product of matrices* $\mathbb{A}$ and $\mathbb{B}$ is defined as $\mathbb{A}\mathbb{B} \in M(m \times k)$, $\mathbb{A}\mathbb{B} = (c_{ij})_{\substack{i=1..m \\ j=1..k}}$, where

$$c_{ij} = \sum_{s=1}^{n} a_{is} b_{sj}.$$

Theorem 6.2 (properties of matrix multiplication)**.** *Let* $m, n, k, l \in \mathbf{N}$. *Then we have:*

(i) $\forall \mathbb{A} \in M(m \times n)\ \forall \mathbb{B} \in M(n \times k)\ \forall \mathbb{C} \in M(k \times l) \colon \mathbb{A}(\mathbb{B}\mathbb{C}) = (\mathbb{A}\mathbb{B})\mathbb{C}$, *(associativity)*
(ii) $\forall \mathbb{A} \in M(m \times n)\ \forall \mathbb{B}, \mathbb{C} \in M(n \times k) \colon \mathbb{A}(\mathbb{B} + \mathbb{C}) = \mathbb{A}\mathbb{B} + \mathbb{A}\mathbb{C}$, *(left distributivity)*
(iii) $\forall \mathbb{A}, \mathbb{B} \in M(m \times n)\ \forall \mathbb{C} \in M(n \times k) \colon (\mathbb{A} + \mathbb{B})\mathbb{C} = \mathbb{A}\mathbb{C} + \mathbb{B}\mathbb{C}$, *(right distributivity)*
(iv) $\exists! \mathbb{I} \in M(n \times n)\ \forall \mathbb{A} \in M(n \times n) \colon \mathbb{I}\mathbb{A} = \mathbb{A}\mathbb{I} = \mathbb{A}$. *(identity matrix* $\mathbb{I}$*)*

Remark. Warning! Matrix multiplication is not commutative.

Definition. *Transpose matrix* for a matrix

$$\mathbb{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

is defined by

$$\mathbb{A}^T = \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ a_{13} & a_{23} & \dots & a_{m3} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix},$$

i.e., if $\mathbb{A} = (a_{ij})_{\substack{i=1..m \\ j=1..n}}$, then $\mathbb{A}^T = (b_{uv})_{\substack{u=1..n \\ v=1..m}}$, where $b_{uv} = a_{vu}$ for each $u \in \{1, \dots, n\}$, $v \in \{1, 2, \dots, m\}$.

Theorem 6.3 (properties of transpose matrix)**.** *We have*

(i) $\forall \mathbb{A} \in M(m \times n)\colon \left(\mathbb{A}^T\right)^T = \mathbb{A}$,

(ii) $\forall \mathbb{A}, \mathbb{B} \in M(m \times n)\colon (\mathbb{A} + \mathbb{B})^T = \mathbb{A}^T + \mathbb{B}^T$,

(iii) $\forall \mathbb{A} \in M(m \times n)\ \forall \mathbb{B} \in M(n \times k)\colon (\mathbb{A}\mathbb{B})^T = \mathbb{B}^T \mathbb{A}^T$.

6.2. **Regular matrices.**

Definition. Let $\mathbb{A} \in M(n \times n)$. We say that $\mathbb{A}$ is *regular* matrix, if there exists $\mathbb{B} \in M(n \times n)$ such that

$$\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A} = \mathbb{I}.$$

Definition. We say that $\mathbb{B} \in M(n \times n)$ is *inverse* to a matrix $\mathbb{A} \in M(n \times n)$, if $\mathbb{A}\mathbb{B} = \mathbb{B}\mathbb{A} = \mathbb{I}$.

Remark. A matrix $\mathbb{A} \in M(n \times n)$ is regular, if and only if $\mathbb{A}$ has its inverse matrix.

Theorem 6.4 (regularity and matrix operations)**.** *Let $\mathbb{A}, \mathbb{B} \in M(n \times n)$ be regular. Then we have:*

(i) $\mathbb{A}^{-1}$ *is regular and* $\left(\mathbb{A}^{-1}\right)^{-1} = \mathbb{A}$,

(ii) $\mathbb{A}^T$ *is regular and* $\left(\mathbb{A}^T\right)^{-1} = \left(\mathbb{A}^{-1}\right)^T$,

(iii) $\mathbb{A}\mathbb{B}$ *is regular and* $(\mathbb{A}\mathbb{B})^{-1} = \mathbb{B}^{-1}\mathbb{A}^{-1}$.

Definition. Let $\boldsymbol{v}^1, \dots, \boldsymbol{v}^k \in \mathbf{R}^n$ be vectors. *Linear combination* of vectors $\boldsymbol{v}^1, \dots, \boldsymbol{v}^k$ is an expression $\lambda_1 \boldsymbol{v}^1 + \dots + \lambda_k \boldsymbol{v}^k$, where $\lambda_1, \dots, \lambda_k \in \mathbf{R}$. *Trivial linear combination* of vectors $\boldsymbol{v}^1, \dots, \boldsymbol{v}^k$ we mean the linear combination $0 \cdot \boldsymbol{v}^1 + \dots + 0 \cdot \boldsymbol{v}^k$. Linear combination, which is not trivial, is called *nontrivial*.

Definition. We say that vectors $\boldsymbol{v}^1, \dots, \boldsymbol{v}^k$ are *linearly dependent*, if there exists their nontrivial linear combination, which is equal to the zero vector.

We say that vectors $\boldsymbol{v}^1, \dots, \boldsymbol{v}^k$ are *linearly independent*, if they are not linearly dependent, i.e., if $\lambda_1, \dots, \lambda_k \in \mathbf{R}$ satisfy $\lambda_1 \boldsymbol{v}^1 + \dots + \lambda_k \boldsymbol{v}^k = \boldsymbol{o}$, then $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$.

Definition. Let $\mathbb{A} \in M(m \times n)$. *Rank* of the matrix $\mathbb{A}$ is the maximal number of linearly independent row vectors of $\mathbb{A}$. Rank of $\mathbb{A}$ is denoted by $\operatorname{rk}(\mathbb{A})$.

Definition. We say that $\mathbb{A} \in M(m \times n)$ is in the *row echelon form*, if for each $i \in \{2, \dots, m\}$ we have, that i-th row of $\mathbb{A}$ is a zero vector or the number of zeros at the beginning of the row is bigger than the number of zeros at the beginning of $(i-1)$-st row.

Remark. The rank of row echelon matrix $\mathbb{A}$ is equal to the number of nonzero rows of $\mathbb{A}$.

Definition. *Elementary row transformations* of the matrix $\mathbb{A}$ are defined as:

(i) interchange of two rows,

(ii) multiplication of a row by a nonzero real number,

(iii) addition of a row to another row.

Definition. *Transformation* is defined as a finite sequence of elementary row transformation. If the matrix $\mathbb{B} \in M(m \times n)$ was created from $\mathbb{A} \in M(m \times n)$ applying a transformation T to $\mathbb{A}$, then this fact is denoted by $\mathbb{A} \overset{T}{\rightsquigarrow} \mathbb{B}$.

Theorem 6.5 (properties of transformation)**.**

(i) *Let* $\mathbb{A} \in M(m \times n)$*. Then there exists a transformation transforming* $\mathbb{A}$ *to a row echelon matrix.*

(ii) *Let* T_1 *be a transformation applicable to matrices of the type* $m \times n$*. Then there exists a transformation* T_2 *applicable to matrices of the type* $m \times n$ *such that if* $\mathbb{A} \overset{T_1}{\rightsquigarrow} \mathbb{B}$ *for some* $\mathbb{A}, \mathbb{B} \in M(m \times n)$*, then* $\mathbb{B} \overset{T_2}{\rightsquigarrow} \mathbb{A}$*.*

(iii) *Let* $\mathbb{A}, \mathbb{B} \in M(m \times n)$ *and there exist a transformation* T *such that* $\mathbb{A} \overset{T}{\rightsquigarrow} \mathbb{B}$*. Then* $\operatorname{rk}(\mathbb{A}) = \operatorname{rk}(\mathbb{B})$.

Theorem 6.6 (multiplication and transformation)**.** *Let* $\mathbb{A} \in M(m \times k)$*,* $\mathbb{B} \in M(k \times n)$*,* $\mathbb{C} \in M(m \times n)$ *and we have* $\mathbb{A}\mathbb{B} = \mathbb{C}$*. Let* T *be a transformation and* $\mathbb{A} \overset{T}{\rightsquigarrow} \mathbb{A}'$ *and* $\mathbb{C} \overset{T}{\rightsquigarrow} \mathbb{C}'$*. Then we have* $\mathbb{A}'\mathbb{B} = \mathbb{C}'$*.*

Lemma 6.7. *Let* $\mathbb{A} \in M(n \times n)$ *and* $\operatorname{rk}(\mathbb{A}) = n$*. Then there exists a transformation transforming* $\mathbb{A}$ *to* $\mathbb{I}$*.*

Theorem 6.8. *Let* $\mathbb{A} \in M(n \times n)$*. Then* $\mathbb{A}$ *is regular if and only if* $\operatorname{rk}(\mathbb{A}) = n$*.*

6.3. **Determinants.**

Definition. Let $\mathbb{A} \in M(n \times n)$. The symbol $\mathbb{A}_{ij}$ denotes the matrix of the type $(n-1) \times (n-1)$, which is created from $\mathbb{A}$ omitting i-th row and j-th column.

Definition. Let $\mathbb{A} = (a_{ij})_{i,j=1..n}$. *Determinant* of the matrix $\mathbb{A}$ is defined by

$$\det \mathbb{A} = \begin{cases} a_{11} & \text{then } n = 1, \\ \sum_{i=1}^{n} (-1)^{i+1} a_{i1} \det \mathbb{A}_{i1} & \text{then } n > 1. \end{cases}$$

For $\det \mathbb{A}$ we will use also the symbol

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}.$$

Theorem 6.9. *Let* $j, n \in \mathbf{N}$*,* $j \leq n$*, and matrices* $\mathbb{A}, \mathbb{B}, \mathbb{C} \in M(n \times n)$ *coincide at each row except* j*-th row. Let* j*-th row of* $\mathbb{A}$ *be equal to the sum of* j*-th rows of* $\mathbb{B}$ *and* $\mathbb{C}$*. Then we have* $\det \mathbb{A} = \det \mathbb{B} + \det \mathbb{C}$.

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{j-1,1} & \dots & a_{j-1,n} \\ u_1+v_1 & \dots & u_n+v_n \\ a_{j+1,1} & \dots & a_{j+1,n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{j-1,1} & \dots & a_{j-1,n} \\ u_1 & \dots & u_n \\ a_{j+1,1} & \dots & a_{j+1,n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{j-1,1} & \dots & a_{j-1,n} \\ v_1 & \dots & v_n \\ a_{j+1,1} & \dots & a_{j+1,n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix}$$

Theorem 6.10 (determinant and transformation)**.** *Let* $\mathbb{A}, \mathbb{A}' \in M(n \times n)$.

(i) Let $\mathbb{A}'$ *be created from* $\mathbb{A}$ *such that we interchanged two rows in* $\mathbb{A}$ *(i.e., we applied an elementary row transformation of the first kind). Then we have* $\det \mathbb{A}' = -\det \mathbb{A}$.

(ii) Let $\mathbb{A}'$ *be created from* $\mathbb{A}$ *such that a row in* $\mathbb{A}$ *is multiplied by* $\lambda \in \mathbf{R}$. *Then we have* $\det \mathbb{A}' = \lambda \det \mathbb{A}$.

(iii) Let $\mathbb{A}'$ *be created from* $\mathbb{A}$ *such that we added a row of* $\mathbb{A}$ *to another row of* $\mathbb{A}$ *(i.e., we applied an elementary row transformation of the third kind). Then we have* $\det \mathbb{A}' = \det \mathbb{A}$.

Corollary 6.11. *Let* $\mathbb{A}, \mathbb{A}' \in M(n \times n)$ *and* $\mathbb{A}'$ *be created from* $\mathbb{A}$ *applying a transformation. Then* $\det \mathbb{A}' \neq 0$ *if and only if* $\det \mathbb{A} \neq 0$.

Theorem 6.12 (determinant and transposition)**.** *Let* $\mathbb{A} \in M(n \times n)$. *Then we have* $\det \mathbb{A}^T = \det \mathbb{A}$.

Theorem 6.13 (determinant of product)**.** *Let* $\mathbb{A}, \mathbb{B} \in M(n \times n)$. *Then we have*

$$\det \mathbb{A}\mathbb{B} = \det \mathbb{A} \cdot \det \mathbb{B}.$$

Theorem 6.14. *Let* $\mathbb{A} = (a_{ij})_{i,j=1..n}$, $k \in \{1, \dots, n\}$. *Then*

$$\det \mathbb{A} = \sum_{i=1}^{n} (-1)^{i+k} a_{ik} \det \mathbb{A}_{ik},$$

$$\det \mathbb{A} = \sum_{j=1}^{n} (-1)^{k+j} a_{kj} \det \mathbb{A}_{kj}.$$

Definition. Let $\mathbb{A} = (a_{ij})_{i,j=1..n}$. We say that $\mathbb{A}$ is *upper triangular matrix* if we have $a_{ij} = 0$ for $i > j$, $i, j \in \{1, \dots, n\}$. We say that $\mathbb{A}$ is *lower triangular matrix*, if we have $a_{ij} = 0$ for $i < j$, $i, j \in \{1, \dots, n\}$.

Theorem 6.15. *Let* $\mathbb{A} = (a_{ij})_{i,j=1..n}$ *is upper (lower, respectively) triangular matrix. Then we have*

$$\det \mathbb{A} = a_{11} \cdot a_{22} \cdot \dots \cdot a_{nn}.$$

Theorem 6.16. *Let* $\mathbb{A} \in M(n \times n)$. *Then* $\mathbb{A}$ *is regular if and only if* $\det \mathbb{A} \neq 0$.

6.4. **Systems of linear equations.** The system of n equations with n unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{aligned} \tag{S}$$

Matrix form

$$\mathbb{A}\boldsymbol{x} = \boldsymbol{b},$$

where $\mathbb{A} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$ is called *matrix of the system*, $\boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ *vector of the right side* and $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ *vectors of unknowns.*

Theorem 6.17. *Let $\mathbb{A} \in M(n \times n)$. Then the following are equivalent.*

(i) The matrix $\mathbb{A}$ is regular.
(ii) The system (SSystems of linear equationsDoc-Start) *have for each $\boldsymbol{b}$ a unique solution.*
(iii) The system (SSystems of linear equationsDoc-Start) *have for each $\boldsymbol{b}$ at least one solution.*

Theorem 6.18 (Cramer's rule)**.** *Let $\mathbb{A} \in M(n \times n)$ be a regular matrix, $\boldsymbol{b} \in M(n \times 1)$, $\boldsymbol{x} \in M(n \times 1)$, and $\mathbb{A}\boldsymbol{x} = \boldsymbol{b}$. Then*

$$x_j = \frac{\begin{vmatrix} a_{11} & \dots & a_{1,j-1} & b_1 & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & & \vdots & & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & b_n & a_{n,j+1} & \dots & a_{nn} \end{vmatrix}}{\det \mathbb{A}}$$

for $j = 1, \dots, n$.

System of m equations with n unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n &= b_m \end{aligned} \tag{S'}$$

Matrix notation

$$\mathbb{A}\boldsymbol{x} = \boldsymbol{b},$$

where $\mathbb{A} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in M(m \times n)$, $\boldsymbol{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in M(m \times 1)$ a $\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in M(n \times 1)$.

Definition. The matrix

$$(\mathbb{A}|\boldsymbol{b}) = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right)$$

is called *extended matrix of the system* (S'Systems of linear equationsDoc-Start).

Theorem 6.19. *The system* (S'Systems of linear equationsDoc-Start) *has a solution if and only if the matrix has the same rank as the extended matrix of the system.*

6.5. **Matrix and linear mappings.**

Definition. We say that a mapping $f\colon \mathbf{R}^n \to \mathbf{R}^m$ is *linear* if

(i) $\forall \boldsymbol{u}, \boldsymbol{v} \in \mathbf{R}^n\colon f(\boldsymbol{u} + \boldsymbol{v}) = f(\boldsymbol{u}) + f(\boldsymbol{v})$,
(ii) $\forall \lambda \in \mathbf{R}\ \forall \boldsymbol{u} \in \mathbf{R}^n\colon f(\lambda \boldsymbol{u}) = \lambda f(\boldsymbol{u})$.

Definition. Let $i \in \{1, \dots, n\}$. The vector

$$\boldsymbol{e}^i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \dots\ i\text{-th coordinate}$$

is called i-th *canonical vector* of the space $\mathbf{R}^n$. The set $\{\boldsymbol{e}^1, \dots, \boldsymbol{e}^n\}$ of all canonical vectors in $\mathbf{R}^n$ is called *canonical basis of the space* $\mathbf{R}^n$.

The properties of canonical vectors:

(i) $\forall \boldsymbol{x} \in \mathbf{R}^n\ \exists \lambda_1, \dots, \lambda_n \in \mathbf{R}\colon \boldsymbol{x} = \lambda_1 \boldsymbol{e}^1 + \dots + \lambda_n \boldsymbol{e}^n$,
(ii) the vectors $\boldsymbol{e}^1, \dots, \boldsymbol{e}^n$ are linearly independent.

Theorem 6.20 (representation of linear mappings)**.** *The mapping $f\colon \mathbf{R}^n \to \mathbf{R}^m$ is linear if and only if there exists a matrix $\mathbb{A} \in M(m \times n)$ such that*

$$\forall \boldsymbol{u} \in \mathbf{R}^n\colon f(\boldsymbol{u}) = \mathbb{A}\boldsymbol{u} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}.$$

Theorem 6.21. *Let a mapping $f\colon \mathbf{R}^n \to \mathbf{R}^n$ be linear. Then the following are equivalent.*

(i) The mapping f is a bijection (i.e., f is an injective mapping $\mathbf{R}^n$ onto $\mathbf{R}^n$).
(ii) The mapping f is an injective mapping.
(iii) The mapping f is a mapping $\mathbf{R}^n$ onto $\mathbf{R}^n$.

Theorem 6.22. *Let $f\colon \mathbf{R}^n \to \mathbf{R}^m$ be a linear mapping represented by matrix $\mathbb{A} \in M(m \times n)$ a $g\colon \mathbf{R}^m \to \mathbf{R}^k$ be a linear mapping represented by a matrix $\mathbb{B} \in M(k \times m)$. Then the composed mapping $g \circ f\colon \mathbf{R}^n \to \mathbf{R}^k$ is linear and is represented by the matrix $\mathbb{B}\mathbb{A}$.*

7. Infinite series

7.1. **Basic notions.**

Definition. Let $\{a_n\}$ be a sequence of real numbers. Symbol $\sum_{n=1}^{\infty} a_n$ is called an *infinite series*. For $m \in \mathbb{N}$ we set

$$s_m = a_1 + a_2 + \cdots + a_m.$$

The number s_m is called *m-th partial sum* of the series $\sum_{n=1}^{\infty} a_n$. The element a_n is called *n-th member* of the series $\sum_{n=1}^{\infty} a_n$. The *sum* of infinite series $\sum_{n=1}^{\infty} a_n$ is defined as the limit of the sequence $\{s_m\}$, if such a limit exists. The sum of the series is denoted by the symbol $\sum_{n=1}^{\infty} a_n$. We say that a series *converges*, if its sum is a real number. In the opposite case, we say that the series *diverges*.

Theorem 7.1 (necessary condition)**.** *If a series $\sum_{n=1}^{\infty} a_n$ converges, then $\lim a_n = 0$.*

Remark. Suppose that $\alpha \in \mathbf{R}$ and a series $\sum_{n=1}^{\infty} a_n$ converges. Then the series $\sum_{n=1}^{\infty} \alpha a_n$ converges and it holds $\sum_{n=1}^{\infty} \alpha a_n = \alpha \sum_{n=1}^{\infty} a_n$. If $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ converge, then the series $\sum_{n=1}^{\infty}(a_n + b_n)$ convergens and if holds $\sum_{n=1}^{\infty}(a_n + b_n) = \sum_{n=1}^{\infty} a_n + \sum_{n=1}^{\infty} b_n$.

7.2. **Series with nonnegative members and absolute convergence.**

Theorem 7.2. *Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be series satisfying $0 \leq a_n \leq b_n$ for each $n \in \mathbb{N}$.*

(i) If $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.
(ii) If $\sum_{n=1}^{\infty} a_n$ diverges, then $\sum_{n=1}^{\infty} b_n$ diverges.

Theorem 7.3. *Let $\{a_n\}$ be a sequence of real numbers. If $\sum_{n=1}^{\infty} |a_n|$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.*

Definition. We say that $\sum_{n=1}^{\infty} a_n$ is *absolute convergent*, if $\sum_{n=1}^{\infty} |a_n|$ converges. If $\sum_{n=1}^{\infty} a_n$ converges but not absolutely, then $\sum_{n=1}^{\infty} a_n$ converges *nonabsolutely*.

Remark. Let $|a_n| \leq b_n$ for each $n \in \mathbb{N}$. If the series $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.

Theorem 7.4 (limit test)**.** *Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be series with nonnegative members.*

(i) Let

$$\lim_{n\to\infty} \frac{a_n}{b_n}$$

exists proper. If $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.

(ii) Let

$$\lim_{n\to\infty} \frac{a_n}{b_n} = c \in (0, +\infty).$$

Then $\sum_{n=1}^{\infty} a_n$ converges if and only if $\sum_{n=1}^{\infty} b_n$ converges.

Theorem 7.5 (Cauchy test). *Let* $\sum_{n=1}^{\infty} a_n$ *be a series. The we have*

(i) If $\lim \sqrt[n]{|a_n|} < 1$, *then* $\sum_{n=1}^{\infty} a_n$ *is absolutely convergent.*
(ii) If $\lim \sqrt[n]{|a_n|} > 1$, *then* $\sum_{n=1}^{\infty} a_n$ *diverges.*

Theorem 7.6 (d'Alembert test). *Let* $\sum_{n=1}^{\infty} a_n$ *be a series with nonzero members. Then we have*

(i) If $\lim |a_{n+1}/a_n| < 1$, *then* $\sum_{n=1}^{\infty} a_n$ *absolutely convergent.*
(ii) If $\lim |a_{n+1}/a_n| > 1$, *then* $\sum_{n=1}^{\infty} a_n$ *diverges.*

Theorem 7.7. *Let* $\alpha \in \mathbb{R}$. *The series* $\sum_{n=1}^{\infty} 1/n^{\alpha}$ *converges if and only if* $\alpha > 1$.

7.3. **Alternating series.**

Theorem 7.8 (Leibniz). *Let* $\sum_{n=1}^{\infty} (-1)^n a_n$ *be a series. Assume*

- $a_n \geq a_{n+1} \geq 0$ *for every* $n \in \mathbb{N}$,
- $\lim_{n\to\infty} a_n = 0$.

Then $\sum_{n=1}^{\infty} (-1)^n a_n$ *converges.*

8. INTEGRALS

8.1. Riemann integral.

Definition. A finite sequence $\{x_j\}_{j=0}^n$ is called a *partition of the interval* $[a,b]$, if we have

$$a = x_0 < x_1 < \cdots < x_n = b.$$

The points $x_0, \dots, x_n$ are called *partition points*.

By a *norm of partition* $D = \{x_j\}_{j=0}^n$ we mean

$$\nu(D) = \max\{x_j - x_{j-1};\ j = 1, \dots, n\}.$$

We say that a partition D' of an interval $[a,b]$ is a *refinement of the partition* D of the interval $[a,b]$, if each point of D is a partition point of D'.

Definition. Let f be a bounded function on an interval $[a,b]$ and $D = \{x_j\}_{j=0}^n$ be a partition of $[a,b]$. We denote

$$\overline{S}(f, D) = \sum_{j=1}^{n} M_j (x_j - x_{j-1}),\ \text{ where } M_j = \sup\{f(x);\ x \in [x_{j-1}, x_j]\},$$

$$\underline{S}(f, D) = \sum_{j=1}^{n} m_j (x_j - x_{j-1}),\ \text{ where } m_j = \inf\{f(x);\ x \in [x_{j-1}, x_j]\},$$

$$\overline{\int_a^b} f(x)\,\mathrm{d}x = \inf\{\overline{S}(f, D);\ D \text{ is a partition of the interval } [a,b]\},$$

$$\underline{\int_a^b} f(x)\,\mathrm{d}x = \sup\{\underline{S}(f, D);\ D \text{ is a partition of the interval } [a,b]\}.$$

Definition. We say that a bounded function f has *Riemann integral* over the interval $[a,b]$, if $\overline{\int_a^b} f(x)\,\mathrm{d}x = \underline{\int_a^b} f(x)\,\mathrm{d}x$. Then the value of the integral of f over the interval $[a,b]$ is equal to $\overline{\int_a^b} f(x)\,\mathrm{d}x$ and is denoted by $\int_a^b f(x)\,\mathrm{d}x$. If $a > b$, we define $\int_a^b f(x)\,\mathrm{d}x = -\int_b^a f(x)\,\mathrm{d}x$. If $a = b$, we define $\int_a^b f(x)\,\mathrm{d}x = 0$.

Theorem 8.1. *(i) Let a function f have Riemann integral over $[a,b]$ and let $[c,d] \subset [a,b]$. Then f has Riemann integral over $[c,d]$.*

(ii) Let $c \in (a,b)$ and a function f have Riemann integral over $[a,c]$ and $[c,b]$. Then f has Riemann integral over $[a,b]$ and we have

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x.$$

Theorem 8.2. *Let f and g be functions with Riemann integral over $[a,b]$ and let $\alpha \in \mathbf{R}$. Then*

(i) the function αf has Riemann integral over $[a,b]$ and it holds

$$\int_a^b \alpha f(x)\,\mathrm{d}x = \alpha \int_a^b f(x)\,\mathrm{d}x,$$

(ii) the function $f + g$ has Riemann integral over $[a, b]$ and it holds

$$\int_a^b \big(f(x) + g(x)\big)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x.$$

Theorem 8.3. *Let $a, b \in \mathbf{R}$, $a < b$, and let f and g be functions with Riemann integral over $[a, b]$.*

(i) If $f(x) \geq 0$ for each $x \in [a, b]$, then

$$\int_a^b f(x)\,\mathrm{d}x \geq 0.$$

(ii) If $f(x) \leq g(x)$ for each $x \in [a, b]$, then

$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x.$$

(iii) The function $|f|$ has Riemann integral over $[a, b]$ and it holds

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f(x)|\,\mathrm{d}x.$$

Theorem 8.4. *Let a function f be continuous on the interval $[a, b]$, $a, b \in \mathbf{R}$. Then f has Riemann integral over $[a, b]$.*

Theorem 8.5. *Let f be a continuous function on $[a, b]$ and let $c \in [a, b]$. If we denote $F(x) = \int_c^x f(t)\,\mathrm{d}t$ for $x \in (a, b)$, then $F'(x) = f(x)$ for each $x \in (a, b)$.*

8.3. **Primitive functions.**

Definition. Let a function f be defined on an open interval I. We say that a function F is a *primitive function of f on I*, if for each $x \in I$ there exists $F'(x)$ and $F'(x) = f(x)$.

Theorem 8.6. *Let F and G be primitive functions of f on an open interval I. Then there exists $c \in \mathbb{R}$ such that $F(x) = G(x) + c$ for each $x \in I$.*

Theorem 8.7. *Let f be a continuous function on an open interval I. Then f has on I a primitive function.*

Theorem 8.8. *Let f have on an open interval I a primitive function F, let a function g have on I a primitive function G, and $\alpha, \beta \in \mathbf{R}$. Then the function $\alpha F + \beta G$ is a primitive function of $\alpha f + \beta g$ on I.*

Theorem 8.9 (substitution). *(i) Let F be a primitive function of f on (a, b). Let φ be a function defined on an interval (α, β) with values in (a, b) and φ has at each point $t \in (\alpha, \beta)$ proper derivative. Then we have*

$$\int f(\varphi(t))\varphi'(t)\,dt \stackrel{c}{=} F(\varphi(t)) \text{ on } (\alpha, \beta).$$

(ii) Let a function φ have at each point of an interval (α, β) nonzero proper derivative and $\varphi((\alpha, \beta)) = (a, b)$. Let f be defined on an interval (a, b) and we have

$$\int f(\varphi(t))\varphi'(t)\,dt \stackrel{c}{=} G(t) \text{ on } (\alpha, \beta).$$

Then we have

$$\int f(x)\,dx \stackrel{c}{=} G(\varphi^{-1}(x)) \text{ on } (a, b).$$

Theorem 8.10 (integration per partes)**.** *Let I be an open interval and let functions f and g be continuous on I. Let F be a primitive function of f on I and G be a primitive function of g on I. Then we have*

$$\int g(x)F(x)\,dx = G(x)F(x) - \int G(x)f(x)\,dx \text{ na } I.$$

Definition. *Rational function* is a ratio of two polynomials, where the polynomial in denominator is not identically zero.

Theorem 8.11. *Let P, Q be polynomial functions with real coefficients such that*

(i) *degree of P is strictly smaller than degree of Q,*
(ii) $Q(x) = a_n(x - x_1)^{p_1} \dots (x - x_k)^{p_k}(x^2 + \alpha_1 x + \beta_1)^{q_1} \dots (x^2 + \alpha_l x + \beta_l)^{q_l}$,
(iii) $a_n,\ x_1, \dots x_k,\ \alpha_1, \dots, \alpha_l,\ \beta_1, \dots, \beta_l \in \mathbf{R},\ a_n \neq 0$,
(iv) $p_1, \dots, p_k,\ q_1, \dots, q_l \in \mathbf{N}$,
(v) *the polynomials $x - x_1,\ x - x_2, \dots, x - x_k,\ x^2 + \alpha_1 x + \beta_1,\ \dots, x^2 + \alpha_l x + \beta_l$ have no common root,*
(vi) *the polynomials $x^2 + \alpha_1 x + \beta_1,\ \dots, x^2 + \alpha_l x + \beta_l$ have no real root.*

Then there exist unique real numbers $A_1^1, \dots, A_{p_1}^1, \dots, A_1^k, \dots, A_{p_k}^k, B_1^1,\ C_1^1, \dots, B_{q_1}^1, C_{q_1}^1, \dots, B_1^l, C_1^l, \dots, B_{q_l}^l,\ C_{q_l}^l$ such that we have

$$\begin{aligned}
\frac{P(x)}{Q(x)} &= \frac{A_1^1}{(x - x_1)^{p_1}} + \dots + \frac{A_{p_1}^1}{(x - x_1)} \\
&+ \dots + \frac{A_1^k}{(x - x_k)^{p_k}} + \dots + \frac{A_{p_k}^k}{x - x_k} \\
&+ \frac{B_1^1 x + C_1^1}{(x^2 + \alpha_1 x + \beta_1)^{q_1}} + \dots + \frac{B_{q_1}^1 x + C_{q_1}^1}{x^2 + \alpha_1 x + \beta_1} + \dots \\
&+ \frac{B_1^l x + C_1^l}{(x^2 + \alpha_l x + \beta_l)^{q_l}} + \dots + \frac{B_{q_l}^l x + C_{q_l}^l}{x^2 + \alpha_l x + \beta_l}.
\end{aligned}$$