Matematická analýza pro fyziky II

Robert Černý & Milan Pokorný

11. května 2016

Obsah

Kapitola 8

Číselné řady

V dalším se budeme zabývat otázkou konvergence číselných řad. Podobně jako u konvergence Newtonova integrálu je konvergence řad jen mezivýsledek, který nám později umožní určovat součty řad za pomoci metod, které konvergenci vyžadují. Na druhou stranu, někdy nám výsledek, zda řada konverguje či ne, dává přesně tu informaci, kterou potřebujeme, a přesná hodnota součtu řady není až tak důležitá. Někdy vystačíme s více či méně přesným odhadem součtu řady.

Od této kapitoly výklad teorie poněkud zrychlíme. Při citování použitých vět již nebudeme uvádět podrobné ověření jejich předpokladů mimo situace, kdy je ověření obtížné, jinak práci přenecháme (často bez varování) čtenáři. Dále budeme v odhadech používat C jako (nejčastěji multiplikativní) neškodnou konstantu, která může z řádku na řádek měnit svoji hodnotu (vzpomeňte si na důkaz aritmetiky limit, kde jsme chtěli vždy zkoumanou veličinu odhadnout násobkem ε a na velikosti multiplikativní konstanty nezáleželo). Dále symbol $+\infty$ budeme zkracovat na ∞, kdykoliv bude jasné, že pracujeme na $\mathbb{R}^*$.

8.1 Základní pojmy

Definice 8.1.1 (Řada)**.** Nechť $\{a_k\} \subset \mathbb{R}$ je posloupnost. Symbol $\sum_{k=1}^{\infty} a_k$ budeme nazávat *řadou.* Pro $k \in \mathbb{N}$ se číslo a_k nazývá *k-tý člen*, číslo $s_n := \sum_{k=1}^{n} a_k$ se nazývá *n-tý částečný součet* a $\{s_n\}$ nazveme *posloupností částečných součtů* řady $\sum_{k=1}^{\infty} a_k$.

Existuje-li vlastní $s := \lim_{n\to\infty} s_n$, říkáme, že řada *konverguje.* Pokud je uvedená limita nevlastní, řada *diverguje* a pokud limita částečných součtů neexistuje, řada *osciluje.*

V prvních dvou případech číslo s nazýváme *součtem řady* a píšeme $\sum_{k=1}^{\infty} a_k = s$.

Poznámka 8.1.2. V případě, kdy s existuje, má symbol $\sum_{k=1}^{\infty} a_k$ vlastně dva významy. Jednak zastupuje posloupnost, kterou se snažíme sečíst, jednak její součet (tedy číslo). Bývá zvykem v takovéto situaci přednostně chápat $\sum_{k=1}^{\infty} a_k$ jako číslo s.

Poznámka 8.1.3. V některých situacích bude přirozené pracovat s $\sum_{k=0}^{\infty} a_k$. Nazývejme posloupností rovněž zobrazení z $\mathbb{N}_0$ do $\mathbb{R}$ (opět budeme psát $\{a_k\}_{k=0}^{\infty}$ či jen $\{a_k\}$).

Poznámka 8.1.4. Řady komplexních čísel se definují analogicky. Nebude-li řečeno jinak, v dalším se budeme zabývat řadami reálných čísel. Odvození podobných výsledků pro komplexní řady přenecháváme čtenáři jako cvičení, popřípadě budou okomentovány zvlášť.

Příklad 8.1.5. (i) Nechť $q \in \mathbb{C}$ a $a_0 \in \mathbb{C} \setminus \{0\}$. Pro každé $k \in \mathbb{N}_0$ definujme $a_k = a_0 q^k$. Vzniklá řada se nazývá *geometrická řada* a díky identitě

$$(1 + q + \cdots + q^n)(1 - q) = 1 - q^{n+1} \qquad \text{platné pro každé } n \in \mathbb{N}$$

její částečné součty splňují pro $q \neq 1$

$$s_n = a_0 \frac{1 - q^{n+1}}{1 - q}.$$

Platí-li $|q| < 1$, řada konverguje a dostáváme $\sum_{k=0}^{\infty} a_0 q^k = \frac{a_0}{1-q}$. Pokud $q = 1$, pracujeme s řadou $\sum_{k=0}^{\infty} a_0$ a dostáváme $\sum_{k=0}^{\infty} a_0 = \infty \in \mathbb{C}^*$. Pokud $|q| = 1$ a $q \neq 1$, řada osciluje. Konečně, pro $|q| > 1$ dostáváme $\sum_{k=0}^{\infty} a_0 q^k = \infty \in \mathbb{C}$ (reálný případ vyžaduje ohlídání jak $\operatorname{sign} a_0$ tak $\operatorname{sign} q$, pro $q < -1$ řada osciluje).
(ii) Uvažme *harmonickou řadu* $\sum_{k=1}^{\infty} \frac{1}{k}$. Její částečné součty tvoří monotonní posloupnost, mají tedy limitu v $\mathbb{R}^*$. Platí pro ně

$$s_1 = 1 \qquad s_2 = 1 + \frac{1}{2} = \frac{3}{2} \qquad s_4 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} > 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} = \frac{4}{2}$$
$$s_8 = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} > 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{5}{2}$$

a indukcí lze získat $s_{2^n} > \frac{n+2}{2}$. Odtud $\sum_{k=1}^{\infty} \frac{1}{k} = \infty$. Připomeňme ještě, že v kapitole o určitém integrálu jsme diverenci této řady už ukázali takto

$$s_n = \sum_{k=1}^{n} \frac{1}{k} > (\mathcal{R}) \int_1^{n+1} \frac{1}{x} \, dx = (\mathcal{N}) \int_1^{n+1} \frac{1}{x} \, dx = [\log]_1^{n+1} = \log(n+1) \overset{n \to \infty}{\to} \infty.$$

(iii) Dalším typem řad, které umíme sečíst, jsou *teleskopické řady*. Například řada $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$, pro kterou máme

$$\begin{aligned} s_n = \sum_{k=1}^{n} \frac{1}{k(k+1)} &= \sum_{k=1}^{n} \Big(\frac{1}{k} - \frac{1}{k+1} \Big) = \frac{1}{1} - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \cdots + \frac{1}{n} - \frac{1}{n+1} \\ &= 1 - \frac{1}{n+1} \overset{n \to \infty}{\to} 1. \end{aligned}$$

Obecně pro teleskopickou řadu typu

$$a_k = b_k - b_{k+m} \qquad \text{kde } m \in \mathbb{N} \text{ a } \lim_{k \to \infty} b_k = 0$$

máme

$$s_n = \sum_{k=1}^{n} a_k = \sum_{k=1}^{n} (b_k - b_{k+m}) = b_1 + \cdots + b_n - (b_{m+1} + \cdots + b_{m+n}) \overset{n\to\infty}{\to} b_1 + \cdots + b_m.$$

(iv) Uvažme řadu $\sum_{k=1}^{\infty} \frac{1}{k^2}$. Opět se jedná o řadu s nezápornými členy, proto jsou částečné součty monotonní a existuje jejich limita. Navíc máme

$$s_n = \sum_{k=1}^{n} \frac{1}{k^2} = 1 + \sum_{k=2}^{n} \frac{1}{k^2} \le 1 + \sum_{k=2}^{n} \frac{1}{k(k-1)} = 1 + \sum_{k=2}^{n} \Big(\frac{1}{k-1} - \frac{1}{k} \Big) = 2 - \frac{1}{n},$$

odkud dostáváme konvergenci. Dalo by se také postupovat přes $(\mathcal{N}) \int_1^{\infty} \frac{1}{x^2}\,\mathrm{d}x$.

(v) Uvažme řadu $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$. Částečné součty si přepišme do tvaru

$$s_{2n} = \Big(-1 + \frac{1}{2} \Big) + \Big(-\frac{1}{3} + \frac{1}{4} \Big) + \ldots \Big(-\frac{1}{2n-1} + \frac{1}{2n} \Big)$$

a

$$s_{2n+1} = -1 + \Big(\frac{1}{2} - \frac{1}{3} \Big) + \Big(\frac{1}{4} - \frac{1}{5} \Big) + \ldots \Big(\frac{1}{2n} - \frac{1}{2n+1} \Big).$$

Odtud vidíme, že $\{s_{2n}\}$ a $\{s_{2n+1}\}$ jsou monotonní posloupnosti s členy v intervalu $[-1, 0]$ (neboť vždy $-1 < s_{2n+1} < s_{2n} < 0$). Obě tedy musí být konvergentní. Navíc

$$s_{2n+1} - s_{2n} = -\frac{1}{2n+1} \overset{n\to\infty}{\to} 0,$$

a proto mají obě limity stejnou hodnotu. Zkoumaná řada tedy konverguje. Rozmyslete si, že v této situaci není možné použít přístup přes určitý integrál.

Poznámka 8.1.6. Konvergence řady byla definována jako konvergence jejích částečných součtů. Nabízí se tedy myšlenka, že budeme-li studovat limitní chování posloupnosti s_k, získáme tím nejen informaci o konvergenci studované řady, ale i její součet. Žádnou teorii pro řady by pak nebylo nutné budovat, neboť vystačíme s teorií pro limity posloupností. Velice často však bývá obtížné či nemožné z předpisu pro k-tý člen a_k získat vzorec pro s_k (ve vzácných případech se podle chování prvních několika členů posloupnosti $\{s_k\}$ dá odhadnout správný vzorec, ten se pak dokáže indukcí). V dalším se nebudeme snažit vzorce pro s_k hledat a budeme budovat teorii pracující jen s předpisem pro člen a_k.

Poznámka 8.1.7. Povšimněte si, že na konvergenci řady nemá vliv přidání, vynechání či změna hodnoty u konečného počtu členů.

Nejprve si uvedeme kritérium, pomocí něhož konvergenci vylučujeme.

Věta 8.1.8 (Nutná podmínka konvergence řady)**.** *Nechť řada* $\sum_{k=1}^{\infty} a_k$ *konverguje. Pak* $\lim_{k\to\infty} a_k = 0$.

Důkaz. Označme $L := \sum_{k=1}^{\infty} a_k$. Pro částečné součty pak platí $\lim_{k\to\infty} s_k = L$, a proto

$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} (s_k - s_{k-1}) = \lim_{k\to\infty} s_k - \lim_{k\to\infty} s_{k-1} = L - L = 0.$$

□

Silnějším nástrojem je B-C podmínka (jedná se jen o přepis B-C podmínky pro posloupnosti), která dává ekvivalentní charakterizaci konvergence číselných řad.

Věta 8.1.9 (B-C podmínka pro řady)**.** *Číselná řada $\sum_{k=1}^{\infty} a_k$ konverguje právě tehdy, když splňuje B-C podmínku*

$$\forall \varepsilon > 0\, \exists n_0 \in \mathbb{N}\, \forall n \in \mathbb{N} \cap [n_0, \infty)\, \forall p \in \mathbb{N} \qquad \sum_{k=n+1}^{n+p} a_k < \varepsilon.$$

Cvičení 8.1.10. Dokažte tuto větu přepisem na standardní Bolzano–Cauchyovu podmínku pro konvergenci číselných posloupností.

Příklad 8.1.11. (i) Nechť $a_1, d \in \mathbb{R}$. Definujme *aritmetickou posloupnost* předpisem $a_k = a_1 + (n-1)d$. S výjimkou případu $a_1 = d = 0$ odpovídající řada nemůže konvergovat podle nutné podmínky.
(ii) Harmonická řada $\sum_{k=1}^{\infty} \frac{1}{k}$ nesplňuje B-C podmínku díky vlastnosti

$$\frac{1}{2^n+1} + \frac{1}{2^n+2} + \cdots + \frac{1}{2^n+2^n} > \frac{1}{2}.$$

Poznámka 8.1.12. (i) Později si představíme ještě několik dalších kritérií pro vyloučení konvergence řady. Tato kritéria budou však pracovat jen s řadami, jejichž členy nemění znaménko (myslíme nekonečněkrát, nezapomínejme na poznámku o konečném počtu změn).
(ii) Nutná podmínka je jen speciálním případem B-C podmínky, v němž vlastně uvažujeme $p = 1$, tedy zkoumáme $\sum_{k=n+1}^{n+1} a_k = a_{n+1}$.
(iii) Ve světle předchozích dvou částí této poznámky bude B-C podmínka jediným naším kritériem pro vyloučení konvergence řady u řad s takzvanými obecnými členy (kde nekontrolujeme znaménkové změny).

Z aritmetiky (nevlastních) limit aplikované na částečné součty dostáváme okamžitě následující výsledek.

Věta 8.1.13 (Aritmetika řad)**.** *Nechť $\sum_{k=1}^{\infty} a_k = A \in \mathbb{R}^*$, $\sum_{k=1}^{\infty} b_k = B \in \mathbb{R}^*$, $\alpha, \beta \in \mathbb{R}$. Pak*

$$\sum_{k=1}^{\infty} (\alpha a_k + \beta b_k) = \alpha A + \beta B,$$

kdykoliv má pravá strana smysl.

Příklad 8.1.14. (i) Řada $\sum_{k=1}^{\infty} \left(\frac{1}{k} + \frac{1}{k^2}\right)$ diverguje, neboť její členy jsou součty členů divergentní a konvergentní řady.
(ii) Řada $\sum_{k=1}^{\infty} \left((-1)^k + \frac{1}{k^2}\right)$ osciluje. Skutečně, pokud by konvergovala, musela by konvergovat i

$$\sum_{k=1}^{\infty} (-1)^k = \sum_{k=1}^{\infty} \left((-1)^k + \frac{1}{k^2} - \frac{1}{k^2}\right) = \sum_{k=1}^{\infty} \left((-1)^k + \frac{1}{k^2}\right) - \sum_{k=1}^{\infty} \frac{1}{k^2}.$$

Divergenci vyvrátíme podobně.

Při našem budoucím studiu nás bude jen zřídka zajímat konvergence řady $\sum_{k=1}^{\infty} a_k$. Podstatně důležitější pro nás bude konvergence $\sum_{k=1}^{\infty} |a_k|$. Proto zavádíme následující pojmy.

Definice 8.1.15 (Absolutní a neabsolutní konvergence)**.** Říkáme, že číselná řada $\sum_{k=1}^{\infty} a_k$ *konverguje absolutně*, jestliže konverguje $\sum_{k=1}^{\infty} |a_k|$. Říkáme, že řada $\sum_{k=1}^{\infty} a_k$ *konverguje neabsolutně*, jestliže konverguje $\sum_{k=1}^{\infty} a_k$ ale nekonverguje $\sum_{k=1}^{\infty} |a_k|$.

Poznámka 8.1.16. Řada $\sum_{k=1}^{\infty} |a_k|$ má monotonní částečné součty. Může tedy jen konvergovat a divergovat, nikoliv oscilovat.

Příklad 8.1.17. Již jsme si ukázali, že $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ konverguje a $\sum_{k=1}^{\infty} \frac{1}{k}$ diverguje. Proto $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ konverguje neabsolutně.

Věta 8.1.18 (Absolutní konvergence implikuje konvergenci)**.** *Jestliže číselná řada $\sum_{k=1}^{\infty} a_k$ konverguje absolutně, pak konverguje klasicky.*

Důkaz. Splnění B-C podmínky pro $\sum_{k=1}^{\infty} |a_k|$ implikuje splnění B-C podmínky pro $\sum_{k=1}^{\infty} a_k$, neboť pro všechna $n, p \in \mathbb{N}$ máme

$$\Big|\sum_{k=n+1}^{n+p} a_k\Big| \le \sum_{k=n+1}^{n+p} |a_k|.$$

□

Stejná myšlenka důkazu nám dává následující kritérium.

Věta 8.1.19 (Srovnávací kritérium I)**.** *Nechť pro všechna $k \in \mathbb{N}$ platí $a_k \in \mathbb{R}$, $b_k \ge 0$ a $|a_k| \le b_k$. Jestliže $\sum_{k=1}^{\infty} b_k$ konverguje, pak $\sum_{k=1}^{\infty} a_k$ konverguje (dokonce absolutně).*

Důkaz. Splnění B-C podmínky pro $\sum_{k=1}^{\infty} b_k$ implikuje splnění B-C podmínky pro $\sum_{k=1}^{\infty} |a_k|$, neboť pro všechna $n, p \in \mathbb{N}$ máme

$$\sum_{k=n+1}^{n+p} |a_k| \le \sum_{k=n+1}^{n+p} |b_k|.$$

Podle předchozí věty proto také $\sum_{k=1}^{\infty} a_k$ konverguje. □

Příklad 8.1.20. (i) Řada $\sum_{k=1}^{\infty} \frac{1}{k^3}$ konverguje, neboť $\frac{1}{k^3} \le \frac{1}{k^2}$ pro všechna $k \in \mathbb{N}$ a $\sum_{k=1}^{\infty} \frac{1}{k^2}$ konverguje.
(ii) Řada $\sum_{k=1}^{\infty} \frac{1}{\log(k-\frac{1}{2})k^2}$ konverguje, neboť $|\frac{1}{\log(k-\frac{1}{2})k^2}| \le \frac{1}{k^2}$ pro všechna $k \in \mathbb{N} \cap [4, \infty)$, $\sum_{k=1}^{\infty} \frac{1}{k^2}$ konverguje a konvergence řady nezávisí na chování konečného počtu členů (můžeme kupříkladu první tři členy studované řady nahradit nulou).
(iii) Řada $\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}$ diverguje, neboť díky nezáporným členům nemůže oscilovat a kdyby konvergovala, konvergovala by i $\sum_{k=1}^{\infty} \frac{1}{k}$ (používáme $\frac{1}{\sqrt{k}} \ge \frac{1}{k}\ \forall k \in \mathbb{N}$), což není pravda.

Další kritérium je založené na našich myšlenkách z důkazu konvergence řady $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$.

Věta 8.1.21 (Leibnizovo kritérium)**.** *Nechť $\{a_n\}$ je nezáporná nerostoucí posloupnost. Pak $\sum_{k=1}^{\infty}(-1)^k a_k$ konverguje právě tehdy, když $\lim_{k\to\infty} a_k = 0$.*

Důkaz. "$\Rightarrow$"Tato implikace plyne z nutné podmínky konvergence.
"$\Leftarrow$"Částečné součty si přepišme do tvaru

$$s_{2n} = (-a_1 + a_2) + (-a_3 + a_4) + \ldots (-a_{2n-1} + a_{2n})$$

a

$$s_{2n+1} = -a_1 + (a_2 - a_3) + (a_4 - a_5) + \ldots (a_{2n} - a_{2n+1}).$$

Odtud vidíme, že $\{s_{2n}\}$ a $\{s_{2n+1}\}$ jsou monotonní posloupnosti s členy v intervalu $[-a_1, 0]$ (neboť vždy $-a_1 \leq s_{2n+1} < s_{2n} \leq 0$). Obě tedy musí být konvergentní. Navíc

$$s_{2n+1} - s_{2n} = -a_{2n+1} \overset{n\to\infty}{\to} 0,$$

a proto mají obě posloupnosti stejnou limitu. Zkoumaná řada tedy konverguje. □

Poznámka 8.1.22. Předchozí kritérium by se dalo aplikovat také na řadu

$$1 + \frac{1}{2} - \frac{1}{3} - \frac{1}{4} + \frac{1}{5} + \frac{1}{6} - \frac{1}{7} - \frac{1}{8} + \ldots,$$

pokud bychom pracovali s $b_k := a_{2k-1} + a_{2k}$ (dostáváme nové členy střídající znaménko s nerostoucími absolutními hodnotami). Časem si představíme Dirichletovo kritérium, které bude zobecňovat Leibnizovo kritérium tímto směrem.

8.2 Řady s nezápornými členy

Připomeňme, že v této situaci má posloupnost částečných součtů vždy limitu a proto může řada jen konvergovat nebo divergovat. Představíme si zde další kritéria konvergence. Ještě připomeňme, že se zabýváme případem, kdy není znám obecný předpis pro s_n a proto musíme pracovat s předpisem pro a_k. V některých případech vznikají jednoduché formule z výrazů $\frac{a_{k+1}}{a_k}$ či $\sqrt[k]{a_k}$. Naše kritéria budou připravena pracovat i s těmito výrazy. Často budeme využívat skutečnost, že změna konečného počtu členů neovlivní konvergenci řady.

Poznamenejme ještě, že všechny naše výsledky v této části textu lze také chápat jako výsledky pro absolutní konvergenci řad.

Věta 8.2.1 (Srovnávací kritérium II)**.** *Nechť $\{a_k\}, \{b_k\} \subset [0, \infty)$, $k_0 \in \mathbb{N}$ a je splněna alespoň jedna z podmínek*
(i) $a_k \geq b_k \quad \forall k \geq k_0$
(ii) $\frac{a_{k+1}}{a_k} \geq \frac{b_{k+1}}{b_k} \quad \forall k \geq k_0$ *(tedy $\{a_k\}, \{b_k\} \subset (0, \infty)$)*
a $\sum_{k=1}^{\infty} a_k$ konverguje. Pak $\sum_{k=1}^{\infty} b_k$ konverguje.

Důkaz. U obou podmínek můžeme předpokládat, že $k_0 = 1$, jinak vhodným způsobem změníme prvních $k_0 - 1$ členů zkoumaných řad. Platí-li podmínka (i), výsledek plyne ze Srovnávacího kritéria I. Nechť platí podmínka (ii), pak pro všechna $k \in \mathbb{N}$ máme

$$a_k = \frac{a_k}{a_{k-1}} \frac{a_{k-1}}{a_{k-2}} \dots \frac{a_2}{a_1} a_1 \geq \frac{b_k}{b_{k-1}} \frac{b_{k-1}}{b_{k-2}} \dots \frac{b_2}{b_1} a_1 = \frac{a_1}{b_1} b_k.$$

Odtud $b_k \leq \frac{b_1}{a_1} a_k$ pro všechna $k \in \mathbb{N}$. Napravo máme členy konvergentní řady, proto lze užít první část věty a jsme hotovi. □

Poznámka 8.2.2. (i) První podmínku ve větě je možno také nahradit podmínkou $Ca_k \geq b_k$ (díky aritmetice řad).
(ii) Předchozí věta také říká, že pokud dvě řady s nezápornými členy splňují (i) nebo (ii), a $\sum_{k=1}^{\infty} b_k$ diverguje, pak $\sum_{k=1}^{\infty} a_k$ diverguje.
(iii) Podmínka (ii) se dá přepsat do tvaru $\frac{a_k}{a_{k+1}} \leq \frac{b_k}{b_{k+1}}$. S tímto tvarem se příjemně pracuje v případě řad typu $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ (velice brzy budeme umět charakterizovat konvergenci těchto řad v závislosti na $\alpha \in \mathbb{R}$ a pak je budeme velice často používat ve srovnávacích kritériích), neboť $\frac{a_k}{a_{k+1}} = \frac{\frac{1}{k^\alpha}}{\frac{1}{(k+1)^\alpha}} = (1 + \frac{1}{k})^\alpha$ a pravá strana se dá ještě upravovat pomocí Taylorova rozvoje.

Z první části Srovnávacího kritéria II se snadno získá další užitečný nástroj.

Věta 8.2.3 (Limitní srovnávací kritérium)**.** *Nechť $\{a_k\}, \{b_k\} \subset (0, \infty)$ a dále nechť $\lim_{k\to\infty} \frac{a_k}{b_k} \in (0, \infty)$. Pak $\sum_{k=1}^{\infty} a_k$ konverguje právě tehdy, když $\sum_{k=1}^{\infty} b_k$ konverguje.*
Jestliže $\{a_k\}, \{b_k\} \subset (0, \infty)$ a $\lim_{k\to\infty} \frac{a_k}{b_k} \in [0, \infty)$ a $\sum_{k=1}^{\infty} b_k$ konverguje, pak $\sum_{k=1}^{\infty} a_k$ konverguje.

Důkaz. Nejprve dokažme první část kritéria. Označme $L := \lim_{k\to\infty} \frac{a_k}{b_k}$. Z definice limity existuje $k_0 \in \mathbb{N}$ takové, že

$$\frac{L}{2} b_k \leq a_k \leq 2L b_k \qquad \text{pro } k \geq k_0.$$

Nyní již stačí použít první část Srovnávacího kritéria II.

Důkaz druhé části je podobný, používáme nerovnost $a_k \leq (L+1) b_k$. □

Poznámka 8.2.4. Limitní srovnávací kritérium je díky použití limity v předpokladech poměrně rychlý nástroj. Na druhou stranu není tak silný jako jeho původní nelimitní verze, která existenci limity nepožaduje a umožňuje díky tomu třeba ukázat konvergenci řady $\sum_{k=1}^{\infty} \frac{1+(-1)^k}{k^2}$ pomocí konvergence řady $\sum_{k=1}^{\infty} \frac{1}{k^2}$.

Nyní si značně rozšíříme množství známých řad, s nimiž budeme vyšetřované řady srovnávat (zejména o řady typu $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$).

Věta 8.2.5 (Integrální kritérium)**.** *Nechť $a \in \mathbb{N}$ a $f\colon \mathbb{R} \to \mathbb{R}$ je spojitá, kladná a nerostoucí na $[a, \infty]$. Pak*

$$\sum_{k=a}^{\infty} f(k) \quad \textit{konverguje} \qquad \iff \qquad (\mathcal{N}) \int_a^{\infty} f \, \mathrm{d}x \in \mathbb{R}.$$

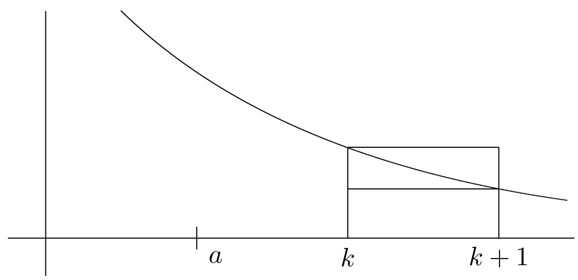
Důkaz. Díky monotonii funkce f máme

$$f(k+1) \le \int_k^{k+1} f \,\mathrm{d}x \le f(k).$$

Proto pro libovolné $n \in \mathbb{N}$, $n > a$, platí

$$\sum_{k=a+1}^{n+1} f(k) = \sum_{k=a}^{n} f(k+1) \le \int_a^{n+1} f \,\mathrm{d}x \le \sum_{k=a}^{n} f(k).$$

Pokud Newtonův integrál konverguje, je (neklesající) primitivní funkce omezená a nutně pak jsou podle levé části našeho odhadu omezené (neklesající) částečné součty řady $\sum_{k=a}^{\infty} f(k)$. Tato řada proto konverguje. Naopak, omezenost částečných součtů řady $\sum_{k=a}^{\infty} f(k)$ implikuje omezenost (neklesající) primitivní funkce, ta proto musí mít vlastní limitu v nekonečnu. □



Obrázek 8.1: Integrální kritérium: odhady integrálu.

Příklad 8.2.6. (i) Funkce $x \mapsto \frac{1}{x^\alpha}$ splňuje pro $\alpha > 0$ předpoklady Integrálního kritéria (Věta 8.2.5), a protože

$$\int_1^{\infty} \frac{1}{x^\alpha}\,\mathrm{d}x = \begin{cases} [\frac{1}{1-\alpha}x^{1-\alpha}]_1^{\infty} = \infty & \text{pro } \alpha < 1 \\ [\log x]_1^{\infty} = \infty & \text{pro } \alpha = 1 \\ [\frac{1}{1-\alpha}x^{1-\alpha}]_1^{\infty} = \frac{1}{1-\alpha} & \text{pro } \alpha > 1, \end{cases}$$

dostáváme (pro $\alpha \le 0$ je dokonce porušena nutná podmínka konvergence řad)

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \quad \text{konverguje} \qquad \Longleftrightarrow \qquad \alpha > 1.$$

(ii) Uvažme řadu typu $\sum_{k=2}^{\infty} \frac{1}{k^\alpha \log^\beta k}$, kde $\alpha, \beta \in \mathbb{R}$ (řadu sčítáme až od druhého členu, neboť první není definován). Pokud $\alpha > 1$ a $\beta \in \mathbb{R}$, Limitní srovnávací kritérium aplikované na naši řadu a řadu $\sum_{k=1}^{\infty} \frac{1}{k^{\frac{\alpha+1}{2}}}$ $\left(\frac{k^{\frac{\alpha+1}{2}}}{k^\alpha \log^\beta k} \to 0\right)$ spolu s předchozí částí příkladu dávají konvergenci naší řady.

Pokud $\alpha < 1$ a $\beta \in \mathbb{R}$, srovnání s $\sum_{k=1}^{\infty} \frac{1}{k^{\frac{\alpha+1}{2}}}$ dává divergenci. Pokud $\alpha = 1$, Limitní srovnávací kritérium (Věta 8.2.3) kombinované s první částí příkladu je nepoužitelné. Na druhou stranu, pro $\alpha = 1$ umíme funkce tvaru $x \mapsto \frac{1}{x^\alpha \log^\beta x}$ snadno integrovat a máme

$$\int_2^\infty \frac{1}{x \log^\beta x} \, \mathrm{d}x = \begin{cases} [\frac{1}{1-\beta} \log^{1-\beta} x]_2^\infty = \infty & \text{pro } \beta < 1 \\ [\log(\log x)]_2^\infty = \infty & \text{pro } \beta = 1 \\ [\frac{1}{1-\beta} \log^{1-\beta} x]_2^\infty = \frac{1}{\beta-1} \log^{1-\beta} 2 & \text{pro } \beta > 1. \end{cases}$$

Integrální kritérium aplikujeme na $[a, \infty)$, kde $a > 2$ je dost velké, aby zde platilo

$$\Big(\frac{1}{x} \log^{-\beta} x\Big)' = \frac{-\log^{-\beta} x}{x^2} + \frac{-\beta \log^{-\beta-1} x}{x^2} = \frac{\log^{-\beta-1} x}{x^2}(-\log x - \beta) < 0.$$

Celkově dostáváme

$$\sum_{k=2}^\infty \frac{1}{k^\alpha \log^\beta k} \quad \text{konverguje} \qquad \iff \qquad \alpha > 1 \vee (\alpha = 1 \wedge \beta > 1).$$

(iii) Mohli bychom náš postup použít i na případ $\sum_{k=3}^\infty \frac{1}{k^\alpha \log^\beta k \log^\gamma(\log k)}$. S výjimkou případu $\alpha = \beta = 1$ se dají opět kombinovat předchozí výsledky spolu s Limitním srovnávacím kritériem (Věta 8.2.3). Ve vyloučeném případě se naopak dobře integruje. Celkově se dostane

$$\sum_{k=3}^\infty \frac{1}{k^\alpha \log^\beta k \log^\gamma(\log k)} \quad \text{konverguje}$$
$$\iff \qquad \alpha > 1 \vee (\alpha = 1 \wedge \beta > 1) \vee (\alpha = 1 \wedge \beta = 1 \wedge \gamma > 1).$$

Poznámka 8.2.7. (i) Povšimněme si, že například ke zkoumání konvergence řad typu $\sum_{k=2}^\infty \frac{1}{k^\alpha \log^\beta k}$ pro $\alpha \neq 1$ nám stačí znalost chování řad $\sum_{k=2}^\infty \frac{1}{k \log^\beta k}$, neboť pro $\alpha_1 < 1 < \alpha_2$, $\beta_1, \beta_2 \in \mathbb{R}$ a $k \in \mathbb{N}$ dostatečně velké máme

$$\frac{1}{k^{\alpha_2} \log^{\beta_2} k} < \frac{1}{k \log^\beta k} < \frac{1}{k^{\alpha_1} \log^{\beta_1} k}$$

($\sum_{k=2}^\infty \frac{1}{k \log^2 k}$ konverguje, $\sum_{k=2}^\infty \frac{1}{k \log k}$ diverguje).
(ii) Nejčastěji budeme studované řady srovnávat s řadami

$$\sum_{k=1}^\infty q^k, \quad \sum_{k=1}^\infty \frac{1}{k^\alpha}, \quad \sum_{k=2}^\infty \frac{1}{k \log^\alpha k}, \quad \sum_{k=3}^\infty \frac{1}{k \log k \log^\alpha(\log k)}, \quad \dots.$$

Je výhodné si pamatovat, že ve všech výše uvedených typech řad je číslo jedna hraniční hodnotou parametru (q či α) z hlediska konvergence řady.

Poznámka 8.2.8. Nikdy nebudeme mít natolik univerzální kritérium, aby nám o každé řadě řeklo, zda konverguje či diverguje. Jednak je to tím, že se nám nepodařilo najít „hraniční řadu“ takovou, že by řady s většími členy divergovaly a s menšími konvergovaly (taková řada ani existovat nemůže, ať už by konvergovala či divergovala, neboť aritmetika řad, konkrétně násobení kladným číslem, by nám dala spor). Navíc členy řad nemusí mít srovnatelný pokles s nějakou důležitou řadou uvedenou výše. Lze třeba vymyslet konvergentní i divergentní řady splňující

$$a_k \leq \frac{1}{k^2} \quad \text{pro nekonečně mnoho } k \qquad \text{a} \qquad a_k \geq \frac{1}{k} \quad \text{pro nekonečně mnoho } k.$$

Nyní si uvedeme dvě kritéria založená na srovnání s geometrickou řadou.

Věta 8.2.9 (Cauchyovo odmocninové kritérium)**.** *Nechť* $\{a_k\} \subset [0,\infty)$ *a* $k_0 \in \mathbb{N}$.
(i) *Jestliže existuje* $q \in [0,1)$ *takové, že* $\sqrt[k]{a_k} \leq q$ *pro všechna* $k \geq k_0$*, pak* $\sum_{k=1}^{\infty} a_k$ *konverguje. Speciálně, pokud* $\lim_{k\to\infty} \sqrt[k]{a_k} < 1$*, řada konverguje.*
(ii) *Jestliže* $\sqrt[k]{a_k} \geq 1$ *pro všechna* $k \geq k_0$*, pak* $\sum_{k=1}^{\infty} a_k$ *diverguje. Speciálně, pokud* $\lim_{k\to\infty} \sqrt[k]{a_k} > 1$*, řada diverguje.*

Důkaz. Dokažme (i). V prvním případě máme $a_k \leq q^k$ pro $q \in [0,1)$ a $k \geq k_0$, přičemž řada $\sum_{k=1}^{\infty} q^k$ je konvergentní. Výsledek tedy plyne ze Srovnávacího kritéria II (Věta 8.2.1). Pokud $\lim_{k\to\infty} \sqrt[k]{a_k} < 1$, stačí zafixovat $q \in (\lim_{k\to\infty} \sqrt[k]{a_k}, 1)$. Najdeme k_0 tak, že platí $\sqrt[k]{a_k} \leq q$ pro všechna $k \geq k_0$ a jsme v situaci jako výše.

Dokažme (ii). Zde máme odhad $a_k \geq 1$ a pro všechna $k \geq k_0$ máme porušenu nutnou podmínku konvergence číselných řad. Předpoklad $\lim_{k\to\infty} \sqrt[k]{a_k} > 1$ vede na tutéž situaci. $\square$

Příklad 8.2.10. Studujme konvergenci řady $\sum_{k=1}^{\infty} (\frac{k}{k+2})^{k^2}$. Máme

$$\lim_{k\to\infty} \sqrt[k]{a_k} = \lim_{k\to\infty} \Big(1 - \frac{2}{k+2}\Big)^k = \mathrm{e}^{-2} < 1$$

(lze využít $\lim_{k\to\infty} (1 + \frac{1}{k})^k = \mathrm{e}$, nebo si přepsat obecnou mocninu pomocí funkce exp, což vede na limitu standardní obtížnosti).

Poznámka 8.2.11. (i) Cauchyovo odmocninové kritérium (Věta 8.2.9) si neporadí s žádnou z řad $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$, $\alpha > 0$, neboť $\lim_{k\to\infty} \sqrt[k]{\frac{1}{k^\alpha}} = 1$. Zároveň vidíme, že případ $\lim_{k\to\infty} \sqrt[k]{a_k} = 1$ připouští jak konvergentní, tak divergentní řady.
(ii) Přestože je odmocninové kritérium poměrně slabé, nachází uplatnění v situacích, kdy se zápis členu a_k značně zjednoduší po aplikaci k-té odmocniny. Aplikace mocného integrálního kritéria na předchozí příklad by jistě příjemná nebyla.

Další kritérium je opět slabé, leč leckdy uživatelsky velice příjemné.

Věta 8.2.12 (d'Alembertovo podílové kritérium)**.** *Nechť* $\{a_k\} \subset (0,\infty)$ *a* $k_0 \in \mathbb{N}$.
(i) *Jestliže existuje* $q \in [0,1)$ *takové, že* $\frac{a_{k+1}}{a_k} \leq q$ *pro všechna* $k \geq k_0$*, pak* $\sum_{k=1}^{\infty} a_k$ *konverguje. Speciálně, pokud* $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} < 1$*, řada konverguje.*
(ii) *Jestliže* $\frac{a_{k+1}}{a_k} \geq 1$ *pro všechna* $k \geq k_0$*, pak* $\sum_{k=1}^{\infty} a_k$ *diverguje. Speciálně, pokud* $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} > 1$*, řada diverguje.*

Důkaz. Dokažme (i). V prvním případě máme pro libovolné $k > k_0$

$$a_k = \frac{a_k}{a_{k-1}} \frac{a_{k-1}}{a_{k-2}} \dots \frac{a_{k_0+1}}{a_{k_0}} a_{k_0} \leq q^{k-k_0} a_{k_0} = Cq^k$$

a konvergence studované řady je důsledkem konvergence geometrické řady. V případě, že $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} < 1$, pro zafixované $q \in (\lim_{k\to\infty} \frac{a_{k+1}}{a_k}, 1)$ vždy najdeme $k_0 \in \mathbb{N}$ tak, že máme $\frac{a_{k+1}}{a_k} \leq q$ pro všechna $k \geq k_0$ a jsme v situaci jako výše.

Dokažme (ii). V tomto případě máme pro libovolné $k > k_0$

$$a_k = \frac{a_k}{a_{k-1}} \frac{a_{k-1}}{a_{k-2}} \dots \frac{a_{k_0+1}}{a_{k_0}} a_{k_0} \geq a_{k_0},$$

je tedy porušena nutná podmínka konvergence. Předpoklad $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} > 1$ vede na tutéž situaci. □

Příklad 8.2.13. Studujme konvergenci řady $\sum_{k=1}^{\infty} \frac{(k!)^2}{(2k)!}$. Máme

$$\frac{a_{k+1}}{a_k} = \frac{((k+1)!)^2 (2k)!}{(k!)^2 (2k+2)!} = \frac{(k+1)^2}{(2k+2)(2k+1)} \overset{k\to\infty}{\to} \frac{1}{4}.$$

Naše řada proto konverguje podle d'Alembertova podílového kritéria (Věta 8.2.12).

Poznámka 8.2.14. (i) Ani toto kritérium nefunguje na řady typu $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$, $\alpha > 0$, či obecně v případě $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = 1$. Oceníme jej zejména v situacích, kdy se dobře počítá $\lim_{k\to\infty} \frac{a_{k+1}}{a_k}$ a nerovná se jedné.
(ii) Poznamenejme ještě, že ve výrazu $\frac{a_{k+1}}{a_k}$ dochází ke značnému zjednodušení faktoriálu, který se často vyskytuje v Taylorových řadách.

Poznámka 8.2.15. Limitní verze kritéria je opět rychlejší, ale slabší než nelimitní. Stačí uvážit řadu

$$1 + \frac{1}{2} + \frac{1}{8} + \frac{1}{16} + \frac{1}{64} + \frac{1}{128} + \dots$$

(střídá se $\frac{a_{k+1}}{a_k} = \frac{1}{2}$ a $\frac{a_{k+1}}{a_k} = \frac{1}{4}$).

Poznámka 8.2.16. Přestože obě výše dokázaná kritéria jsou shodně založena na vlastnostech geometrické řady, fungují odlišně. Například s řadou

$$1 + \frac{1}{4} + \frac{1}{2} + \frac{1}{8} + \frac{1}{4} + \frac{1}{16} + \frac{1}{8} + \dots$$

(střídá se $\frac{a_{k+1}}{a_k} = \frac{1}{4}$ a $\frac{a_{k+1}}{a_k} = 2$), si odmocninové kritérium poradí, podílové nikoliv.

Podílové kritérium se dá zobecnit tak, že si spočítáme $\frac{a_{k+1}}{a_k}$ pro řady typu $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ a získaný výsledek budeme kombinovat s druhou částí Srovnávacího kritéria II (Věta 8.2.1).

Věta 8.2.17 (Raabeho kritérium)**.** *Nechť* $\{a_k\} \subset (0, \infty)$ *a* $k_0 \in \mathbb{N}$.
(i) *Existuje-li* $q > 1$ *tak, že* $k(\frac{a_k}{a_{k+1}} - 1) \geq q$ *pro všechna* $k \geq k_0$, *pak řada* $\sum_{k=1}^{\infty} a_k$ *konverguje. Speciálně, jestliže* $\lim_{k\to\infty} k(\frac{a_k}{a_{k+1}} - 1) > 1$, *řada konverguje.*
(ii) *Jestliže* $k(\frac{a_k}{a_{k+1}} - 1) \leq 1$ *pro všechna* $k \geq k_0$, *pak řada* $\sum_{k=1}^{\infty} a_k$ *diverguje. Speciálně, jestliže* $\lim_{k\to\infty} k(\frac{a_k}{a_{k+1}} - 1) < 1$, *řada diverguje.*

Důkaz. V prvním případě provedeme srovnání s konvergentní řadou $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$, kde zafixujeme $\alpha \in (1, q)$. Položme tedy $b_k = \frac{1}{k^\alpha}$ pro $k \in \mathbb{N}$ a zafixujme ještě $\beta \in (\alpha, q)$. Pro k dostatečně velké dostáváme odhad

$$\frac{b_k}{b_{k+1}} = \left(\frac{\frac{1}{k}}{\frac{1}{k+1}}\right)^\alpha = \left(1 + \frac{1}{k}\right)^\alpha \leq 1 + \frac{\beta}{k}.$$

Skutečně, Taylorův rozvoj funkce $(1 + x)^\alpha$ v počátku a Lagrangeův tvar zbytku dávají

$$(1 + x)^\alpha = 1 + \alpha x + \frac{1}{2}\alpha(\alpha - 1)(1 + \xi)^{\alpha-2}x^2,$$

kde $\xi \in (0, x)$. Pro x dostatečně blízko k počátku proto můžeme poslední člen pravé strany odhadnout libovolně malým násobkem předposledního.

Z předchozích odhadů a předpokladu $k(\frac{a_k}{a_{k+1}} - 1) \geq q$ máme

$$\frac{a_k}{a_{k+1}} \geq 1 + \frac{q}{k} > 1 + \frac{\beta}{k} \geq \frac{b_k}{b_{k+1}}.$$

Druhá část Srovnávacího kritéria II (Věta 8.2.1) nám dává konvergenci $\sum_{k=1}^{\infty} a_k$.

Nechť nyní $k(\frac{a_k}{a_{k+1}} - 1) \leq 1$ pro všechna $k \geq k_0$. Odtud

$$\frac{a_k}{a_{k+1}} \leq 1 + \frac{1}{k} = \frac{\frac{1}{k}}{\frac{1}{k+1}}$$

a druhá část Srovnávacího kritéria II nám dává divergenci, neboť $\sum_{k=1}^{\infty} \frac{1}{k}$ diverguje. $\square$

Poznámka 8.2.18. (i) Raabeho kritérium se používá v situacích, kdy je zápis $\frac{a_{k+1}}{a_k}$ jednodušší než zápis a_k, ale podílové kritérium je v dané situaci příliš slabé. Typicky se k Raabeho kritériu přechází po neúspěšné aplikaci podílového kritéria (mějte ovšem na paměti, že jedno z kritérií pracuje s výrazem $\frac{a_{k+1}}{a_k}$, zatímco druhé s $\frac{a_k}{a_{k+1}}$).
(ii) Raabeho kritérium není v žádném případě všemocné. Ověřte si sami, že si neporadí s řadami typu $\sum_{k=1}^{\infty} \frac{1}{k \log^\alpha k}$, $\alpha > 0$.

Další krůček ve zjemnění práce s výrazem $\frac{a_k}{a_{k+1}}$ nám dává následující kritérium.

Věta 8.2.19 (Gaussovo kritérium)**.** *Nechť* $\{a_k\} \subset (0, \infty)$. *Nechť existují* $p, q \in \mathbb{R}$ *a* $\varepsilon, C > 0$ *tak, že*

$$\frac{a_k}{a_{k+1}} = p + \frac{q}{k} + \frac{t_k}{k^{1+\varepsilon}}, \qquad \textit{kde } |t_k| \leq C.$$

(i) *Jestliže* $p > 1$, *řada* $\sum_{k=1}^{\infty} a_k$ *konverguje. Jestliže* $p < 1$, *řada diverguje.*
(ii) *Jestliže* $p = 1$ *a* $q > 1$, *řada konverguje.*
(iii) *Jestliže* $p = 1$ *a* $q \leq 1$, *řada diverguje.*

Důkaz. Všechny případy, kdy $p \neq 1$ nebo $q \neq 1$ nám dává Raabeho kritérium. Uvažme zbývající případ $p = q = 1$. Definujme $b_k = \frac{1}{k \log k}$, $k \in \mathbb{N} \setminus \{1\}$. Pak

$$\frac{b_k}{b_{k+1}} = \frac{(k+1)\log(k+1)}{k \log k} = \frac{(k+1)\big(\log k + \log(1+\frac{1}{k})\big)}{k \log k}$$
$$= 1 + \frac{1}{k} + \frac{\log(1+\frac{1}{k})}{\log k} + \frac{\log(1+\frac{1}{k})}{k \log k}.$$

Protože pro dostatečně velké k máme odhad $\log(1+\frac{1}{k}) \geq \frac{1}{2k}$ (využíváme známou limitu $\lim_{x\to 0} \frac{\log(1+x)}{x} = 1$), dostáváme

$$\frac{b_k}{b_{k+1}} \geq 1 + \frac{1}{k} + \frac{1}{2k \log k} \qquad \text{pro } k \text{ dostatečně velké.}$$

Celkově i s předpokladem $\frac{a_k}{a_{k+1}} \leq 1 + \frac{1}{k} + \frac{C}{k^{1+\varepsilon}}$ proto máme pro k dostatečně velké

$$\frac{a_k}{a_{k+1}} \leq 1 + \frac{1}{k} + \frac{C}{k^{1+\varepsilon}} \leq 1 + \frac{1}{k} + \frac{1}{2k \log k} \leq \frac{b_k}{b_{k+1}}$$

a Srovnávací kritérium II (Věta 8.2.1) nám dává divergenci studované řady, neboť $\sum_{k=2}^{\infty} \frac{1}{k \log k}$ diverguje. □

Poznámka 8.2.20. Přestože jsme v důkazu používali řadu $\sum_{k=2}^{\infty} \frac{1}{k \log k}$, s touto řadou si Gaussovo kritérium neporadí, neboť pro žadnou volbu $p, q \in \mathbb{R}$ a $\varepsilon > 0$ zbytkový člen $\frac{t_k}{k^{1+\varepsilon}}$ nemá omezený čitatel (podívejte se na „nejnadějnější" je případ $p = q = 1$ v předchozím důkazu). Dokonce nepomůže ani zesílená verze Gaussova kritéria z Cvičení 8.2.21 níže.

Cvičení 8.2.21. Dokažte, že řada $\sum_{k=1}^{\infty} a_k$ diverguje i za předpokladu, že pro $\alpha > 1$ a $k \geq k_0$

$$\frac{a_k}{a_{k+1}} = 1 + \frac{1}{k} + \frac{t_k}{k \log^{\alpha} k},$$

kde $|t_k| \leq C$, nezávisle na k.

Příklad 8.2.22. Nechť $a, b \in \mathbb{R}$. Zkoumejme konvergenci $\sum_{k=1}^{\infty} \frac{(k+a)(k-1+a)\dots a}{k! k^b}$. Platí (ověřte si sami pomocí Taylorova rozvoje, že r_k, s_k, t_k jsou v dalším omezené)

$$\frac{a_k}{a_{k+1}} = \frac{k+1}{k+1+a}\Big(\frac{k+1}{k}\Big)^b = \Big(1 - \frac{a}{k+1+a}\Big)\Big(1+\frac{1}{k}\Big)^b$$
$$= \Big(1 - \frac{a}{k} + \frac{r_k}{k^2}\Big)\Big(1 + \frac{b}{k} + \frac{s_k}{k^2}\Big) = 1 + \frac{b-a}{k} + \frac{t_k}{k^2}.$$

Zkoumaná řada tedy konverguje právě tehdy, když $b - a > 1$.

Poznámka 8.2.23. Celá výše probíraná teorie se dá aplikovat na řady se zápornými členy (vytkneme znaménko mínus, nebo ve všech kritériích nahradíme a_k za $|a_k|$). Vzhledem k tomu, že změna konečného počtu členů neovlivní konvergenci řady, naše teorie se dá rozšířit i na všechny řady, které nemají zároveň nekonečně mnoho kladných členů a nekonečně mnoho záporných členů.

8.3 Dodatek k řadám s nezápornými členy: kondenzační kritérium

Někdy se používá ještě následující kritérium, zejména v případě, kdy se teorie číselných řad vykládá dříve než teorie integrálu.

Věta 8.3.1 (Lobačevského kondenzační kritérium)**.** *Nechť $\{a_k\} \subset [0,\infty)$ je nerostoucí posloupnost. Pak*

$$\sum_{k=1}^{\infty} a_k \quad \textit{konverguje} \qquad \Longleftrightarrow \qquad \sum_{k=1}^{\infty} 2^k a_{2^k} \quad \textit{konverguje.}$$

Důkaz. Implikace „$\Leftarrow$“ plyne z odhadu (používáme monotonii)

$$(a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \cdots \leq 2a_2 + 4a_4 + 8a_8 + \dots .$$

Implikace „$\Rightarrow$“ plyne z odhadu (opět používáme monotonii)

$$\begin{aligned}(a_2) + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \ldots &\geq a_2 + 2a_4 + 4a_8 + \ldots \\ &= \frac{1}{2}\Big(2a_2 + 4a_4 + 8a_8 + \ldots\Big).\end{aligned}$$

□

Příklad 8.3.2. (i) Konvergence řady $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$ je podle Lobačevského kritéria ekvivalentní konvergenci řady

$$\sum_{k=1}^{\infty} 2^k \frac{1}{(2^k)^\alpha} = \sum_{k=1}^{\infty} \frac{1}{(2^k)^{\alpha-1}} = \sum_{k=1}^{\infty} \frac{1}{(2^{\alpha-1})^k},$$

což nastává právě tehdy, když $\alpha > 1$.
(ii) Konvergence řady $\sum_{k=2}^{\infty} \frac{1}{k \log^\alpha k}$ je podle Lobačevského kritéria ekvivalentní (připomeňme, že konečný počet členů není schopen ovlivnit konvergenci, proto nám stačí monotonie od jistého $k_0 \in \mathbb{N}$) konvergenci řady

$$\sum_{k=2}^{\infty} 2^k \frac{1}{2^k \log^\alpha(2^k)} = \log^{-\alpha} 2 \sum_{k=2}^{\infty} \frac{1}{k^\alpha}$$

a opět dostáváme nám již známý výsledek.

Poznámka 8.3.3. (i) Pokud bychom předchozí příklad zkoumali z pohledu Integrálního kritéria (Věta 8.2.5), zjistili bychom, že Lobačevského kritérium (Věta 8.3.1) vlastně jen pod integrálem provádí logaritmickou substituci.
(ii) Lobačevského kritérium nám podobně jako Integrální kritérium umožní určit konvergenci několika důležitých (a obtížných) řad. Na druhou stranu si neporadí s řadami, kde není předpis pro a_k velice jednoduchý.
(iii) Zhruba se dá říci, že je jedno, které ze dvojice Integrální a Lobačevského kritérium ovládáme, obě zaberou na důležité řady typů $\sum_{k=1}^{\infty} \frac{1}{k^\alpha}$, $\sum_{k=2}^{\infty} \frac{1}{k \log^\alpha k}$, atd. Vyšetřením konvergence těchto řad obě kritéria splnila svou úlohu a už je s největší pravděpodobností čtenář nikdy nevyužije.
(iv) Lobačevského kritérium by se dalo přeformulovat a dokázat rovněž pro pomocnou řadu $\sum_{k=1}^{\infty} 3^k a_{3^k}$. Nic nového bychom tím ovšem nezískali, stále by se jednalo o ekvivalent jedné logaritmické substituce pod integrálem. Skutečný přínos přináší teprve iterování Lobačevského kritéria, tedy například

$$\sum_{k=1}^{\infty} a_k \text{ konverguje} \iff \sum_{k=1}^{\infty} 2^k a_{2^k} \text{ konverguje} \iff \sum_{k=1}^{\infty} 2^k 2^{2^k} a_{2^{2^k}} \text{ konverguje.}$$

Zde jsme provedli operaci odpovídající dvěma logaritmickým substitucím.

8.4 Řady s obecnými členy

Nyní se budeme zabývat řadami, jejichž členy nekonečněkrát změní znaménko, neboli nekonečně mnoho členů má znaménko kladné a nekonečně mnoho záporné. Tato situace je provázena hned několika jevy, které se u řad s kladným znaménkem nevyskytovaly. Jednak kromě konvergence a divergence nyní může nastat i oscilace. Dalším jevem je neabsolutní konvergence. Absolutní konvergence znamenala, že je vhodným způsobem kontrolována velikost členů studované řady. V případě neabsolutní konvergence již nemusí velikost (absolutní hodnota) členů řady splňovat tak přísné podmínky, je-li to kompenzováno dostatečným vzájemným vyrušením kladných a záporných členů (uvažte $\sum_{k=2}^{\infty} \frac{(-1)^k}{\log k}$). V této situaci už informace typů třeba $a_k \le b_k$, $|a_k| \le |b_k|$, $\frac{a_{k+1}}{a_k} \le \frac{b_{k+1}}{b_k}$ a $\lim_{k\to\infty} \frac{a_k}{b_k} \in (0,\infty)$ neimplikují žádný vztah mezi konvergencí řad $\sum_{k=1}^{\infty} a_k$ a $\sum_{k=1}^{\infty} b_k$ (sami si zkonstruujte příklady jako třeba $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ a $\sum_{k=1}^{\infty} \left(\frac{(-1)^k}{k} + \frac{1}{\sqrt{k}}\right)$).

Věta 8.4.1 (Abelovo a Dirichletovo kritérium)**.** *Nechť $\{a_n\}, \{b_n\} \subset \mathbb{R}$ a $\{a_n\}$ je monotonní.*
(Dirichlet) *Jestliže $a_k \to 0$ a $\{b_k\}$ má omezené částečné součty, pak $\sum_{k=1}^{\infty} a_k b_k$ konverguje.*
(Abel) *Jestliže $\{a_n\}$ je omezená a $\sum_{k=1}^{\infty} b_k$ konverguje, pak $\sum_{k=1}^{\infty} a_k b_k$ konverguje.*

Důkaz. Nejprve předpokládejme Dirichletovy podmínky a ukažme, že zkoumaná řada splňuje B-C podmínku. Zvolme $\varepsilon > 0$. Bez újmy na obecnosti můžeme předpokládat, že $\{a_n\}$ je nerostoucí. Ve znění nerovnosti z B-C podmínky si členy posloupnosti $\{b_k\}$ vyjádříme pomocí částečných součtů této posloupnosti, které budeme značit B_n, a máme

$$\begin{aligned}\sum_{k=n+1}^{n+p} a_k b_k &= a_{n+1}b_{n+1} + a_{n+2}b_{n+2} + \cdots + a_{n+p-1}b_{n+p-1} + a_{n+p}b_{n+p} \\ &= a_{n+1}(B_{n+1} - B_n) + a_{n+2}(B_{n+2} - B_{n+1}) \\ &\qquad + \cdots + a_{n+p-1}(B_{n+p-1} - B_{n+p-2}) + a_{n+p}(B_{n+p} - B_{n+p-1}) \\ &= -B_n a_{n+1} + B_{n+1}(a_{n+1} - a_{n+2}) + B_{n+2}(a_{n+2} - a_{n+3}) \\ &\qquad + \cdots + B_{n+p-1}(a_{n+p-1} - a_{n+p}) + a_{n+p}B_{n+p}.\end{aligned}$$

Odtud s využitím monotonie $\{a_k\}$, omezenosti $\{B_n\}$ a vlastnosti $a_k \to 0$ dostáváme pro n dostatečně velká následující odhad

$$\begin{aligned}\Big|\sum_{k=n+1}^{n+p} a_k b_k\Big| &\le |-B_n a_{n+1}| + |B_{n+1}|(a_{n+1} - a_{n+2}) + |B_{n+2}|(a_{n+2} - a_{n+3}) \\ &\qquad + \cdots + |B_{n+p-1}|(a_{n+p-1} - a_{n+p}) + |a_{n+p}B_{n+p}| \\ &\le C\varepsilon + C(a_{n+1} - a_{n+2}) + C(a_{n+2} - a_{n+3}) \\ &\qquad + \cdots + C(a_{n+p-1} - a_{n+p}) + C\varepsilon \\ &= C\varepsilon + C(a_{n+1} - a_{n+p}) + C\varepsilon \le C\varepsilon + Ca_{n+1} + C\varepsilon \le 3C\varepsilon.\end{aligned}$$

Ověřili jsme B-C podmínku pro limitu částečných součtů řady $\sum_{k=1}^{\infty} a_k b_k$ a jsme v prvním případě hotovi.

Nyní předpokládejme Abelovy podmínky. Protože posloupnost $\{a_k\}$ je monotonní a omezená, má vlastní limitu. Označme ji A. Pak

$$\sum_{k=1}^{\infty} a_k b_k = \sum_{k=1}^{\infty} A b_k + \sum_{k=1}^{\infty} (a_k - A) b_k,$$

kde první řada napravo konverguje díky aritmetice řad a druhá splňuje předpoklady Dirichletova kritéria. Proto díky aritmetice řad konverguje i řada nalevo. □

Příklad 8.4.2. (i) Z Dirichletova kritéria plyne Leibnizovo kritérium (tedy Věta 8.1.21), neboť posloupnost $\{(-1)^k\}$ má omezené částečné součty (střídají se hodnoty -1 a 0).
(ii) Často se dá kombinovat Dirichletovo kritérium s Abelovým, jak nám ukazuje příklad $\sum_{k=1}^{\infty} \frac{(-1)^k}{\sqrt{k}} \arctan k$, kde nejprve použijeme Dirichletovo kritérium k ověření konvergence $\sum_{k=1}^{\infty} \frac{(-1)^k}{\sqrt{k}}$ a pak využijeme právě získanou konvergenci spolu s monotonií a omezeností posloupnosti $\{\arctan k\}$ při aplikaci Abelova kritéria.
(ii) Abelovo kritérium se dá použít i vícekrát za sebou. Uvažme například řadu $\sum_{k=1}^{\infty} \frac{(-1)^k}{\sqrt{k}} \frac{k}{k+1} \arctan k$, kde jedna aplikace Dirichletova kritéria a jedna aplikace Abelova kritéria dávají konvergenci $\sum_{k=1}^{\infty} \frac{(-1)^k}{\sqrt{k}} \arctan k$ (bylo výše) a pak

díky Abelovu kritériu ještě můžeme do řady přidat omezený monotonní činitel $\frac{k}{k+1} = 1 - \frac{1}{k+1}$.

Poznámka 8.4.3. (i) Dirichletovo kritérium oproti Abelovu má přísnější podmínky na $\{a_k\}$ (konvergence k nule implikuje omezenost) a volnější podmínky na $\{b_k\}$ (konvergence řady implikuje omezenost jejích částečných součtů). Není možné vzít jen omezenost $\{a_k\}$ a omezenost částečných součtů $\{b_k\}$, jak ukazuje volba $a_k := 1$, $b_k := (-1)^k$.
(ii) Není radno zapomínat na monotonii posloupnosti $\{a_k\}$. Jinak Dirichletovo ani Abelovo kritérium neplatí ($a_k := \frac{(-1)^k}{\sqrt{k}} \to 0$, $b_k := \frac{(-1)^k}{\sqrt{k}}$ má konvergentní řadu, ale celkově $\sum_{k=1}^{\infty} a_k b_k = \sum_{k=1}^{\infty} \frac{1}{k}$ diverguje).
(iii) Komplexní varianta Abelova a Dirichletova kritéria vypadá tak, že $\{a_k\}$ je reálná monotonní posloupnost, $\{b_k\}$ je komplexní posloupnost a zbytek znění je stejný jako v reálném případě. Důkaz se získá rozkladem posloupnosti $\{b_k\}$ na reálnou a imaginární složku, případně se zopakuje důkaz Věty 8.4.1 pro komplexní částečné součty. Nemůže platit varianta s $\{a_k\}, \{b_k\} \in \mathbb{C}$, neboť pak bychom neměli pojem monotonie a bez něho Věta 8.4.1 nemůže platit, jak bylo ukázáno výše.

Poznámka 8.4.4. Povšimněte si, že v případě řad s nezápornými členy nám ani Abelovo ani Dirichletovo kritérium nenabízí nic, co by nám nedalo Srovnávací kritérium I (Věta 8.1.19).

Představíme si ještě dva typy posloupností s omezenými částečnými součty.

Tvrzení 8.4.5. *Nechť $a \in \mathbb{R}$. Pak posloupnost $k \mapsto \sin(ak)$ má omezené částečné součty. Posloupnost $k \mapsto \cos(ak)$ má omezené částečné součty právě tehdy, když a není násobkem čísla 2π.*

Důkaz. Pokud a není násobkem 2π, máme

$$\begin{aligned} s_n := \sum_{k=1}^{n} \sin(ak) &= \sum_{k=1}^{n} \frac{\mathrm{e}^{\mathrm{i}ak} - \mathrm{e}^{-\mathrm{i}ak}}{2\mathrm{i}} = \frac{\sum_{k=1}^{n} \mathrm{e}^{\mathrm{i}ak} - \sum_{k=1}^{n} \mathrm{e}^{-\mathrm{i}ak}}{2\mathrm{i}} \\ &= \frac{1}{2\mathrm{i}} \mathrm{e}^{\mathrm{i}a} \frac{1 - \mathrm{e}^{\mathrm{i}a(k+1)}}{1 - \mathrm{e}^{\mathrm{i}a}} - \frac{1}{2\mathrm{i}} \mathrm{e}^{-\mathrm{i}a} \frac{1 - \mathrm{e}^{-\mathrm{i}a(k+1)}}{1 - \mathrm{e}^{-\mathrm{i}a}}. \end{aligned}$$

Odtud

$$\begin{aligned} |s_n| &\le \frac{1}{|2\mathrm{i}|} |\mathrm{e}^{\mathrm{i}a}| \frac{|1| + |\mathrm{e}^{\mathrm{i}a(k+1)}|}{|1 - \mathrm{e}^{\mathrm{i}a}|} + \frac{1}{|2\mathrm{i}|} |\mathrm{e}^{-\mathrm{i}a}| \frac{|1| + |\mathrm{e}^{-\mathrm{i}a(k+1)}|}{|1 - \mathrm{e}^{-\mathrm{i}a}|} \\ &\le \frac{1}{2} \cdot 1 \cdot \frac{1+1}{|1 - \mathrm{e}^{\mathrm{i}a}|} + \frac{1}{2} \cdot 1 \cdot \frac{1+1}{|1 - \mathrm{e}^{-\mathrm{i}a}|}, \end{aligned}$$

tedy částečné součty posloupnosti $k \mapsto \sin(ak)$ jsou omezené. Pokud a je násobkem 2π, sčítáme samé nulové členy a výsledek platí triviálně. Při práci s posloupností $k \mapsto \cos(ak)$ použijeme vzorec $\cos(ak) = \frac{\mathrm{e}^{\mathrm{i}ak} + \mathrm{e}^{-\mathrm{i}ak}}{2}$. Je-li a násobkem 2π, máme $\cos(ak) \equiv 1$ a částečné součty nejsou omezené. □

Cvičení 8.4.6. Postupem ukázaným výše ukažte, že číselné posloupnosti $\{\sin^3 k\}$, $\{\cos^3 k\}$, $\{(-1)^k \sin^3 k\}$, $\{(-1)^k \cos^3 k\}$ mají omezené částečné součty (při výpočtu budete vždy pracovat se čtveřicí konvergentních geometrických řad). Tímto postupem se dá rovněž ukázat, že $\sin^2 k$ nemá omezené částečné součty (postup výše vede na součet dvou konvergentních geometrických řad a řady reálných konstant).

Poznámka 8.4.7. Samozřejmě, pokud čtenář umí zacházet se součtovými vzorci pro goniometrické funkce a všimne si, že $(-1)^k = \cos(k\pi)$, lze předchozí cvičení vyřešit mnohem snadněji použitím Věty o aritmetice řad (Věta 8.1.13).

Příklad 8.4.8. Řady $\sum_{k=1}^{\infty} \frac{\sin k}{k}$, $\sum_{k=1}^{\infty} \frac{\sin^3 k}{k}$ a $\sum_{k=1}^{\infty} \frac{(-1)^k \sin^3 k}{k}$ jsou konvergentní podle Dirichletova kritéria.

Pro důkaz toho, že zkoumaná řada nekonverguje, máme jedinou přímou metodu a sice porušení B-C podmínky (případně porušení nutné podmínky konvergence, což je ovšem speciální případ B-C podmínky).

Příklad 8.4.9. Ukažme, že nekonverguje řada $\sum_{k=1}^{\infty} \frac{\sin\sqrt{k}}{\sqrt{k}}$. K porušení B-C podmínky využijeme toho, že pro velká k jsou řetězce členů stejného znaménka velmi dlouhé. Předně si povšimněme, že

$$\sqrt{k+1} - \sqrt{k} = \frac{1}{\sqrt{k+1} + \sqrt{k}} \le \frac{1}{2\sqrt{k}}.$$

Pro každé $m \in \mathbb{N}$ zvolme $k_m \in \mathbb{N}$ takové, že $\sqrt{k_m} \in [2m\pi + \frac{\pi}{6}, 2m\pi + \frac{\pi}{4}]$ (aspoň jedno takové číslo existovat musí, neboť pro $m \ge 1$ pracujeme napravo od bodu 2π, tedy $\sqrt{k} \ge 6$, odtud $\sqrt{k+1} - \sqrt{k} \le \frac{1}{2\sqrt{k}} \le \frac{1}{12}$, a proto není možné, aby dvojice $\sqrt{k}$ a $\sqrt{k+1}$ „přeskočila“ interval délky $\frac{\pi}{12} > \frac{1}{12}$). Z odhadu výše také vidíme, že

$$\sqrt{k_m + j} \in [2m\pi + \tfrac{\pi}{6}, 2m\pi + \tfrac{5\pi}{6}] \qquad \text{pro } j = 0, 1, 2, \ldots, 2[\sqrt{k_m}].$$

Odtud

$$\sum_{k=k_m}^{k_m+2[\sqrt{k_m}]} \frac{\sin\sqrt{k}}{\sqrt{k}} \ge \sum_{k=k_m}^{k_m+2[\sqrt{k_m}]} \frac{\frac{1}{2}}{2\sqrt{k_m}} = \frac{1}{4\sqrt{k_m}}(2[\sqrt{k_m}] + 1) \ge \frac{1}{2\sqrt{k_m}}\sqrt{k_m} = \frac{1}{2}.$$

Nedá se proto splnit B-C podmínka s volbou $\varepsilon = \frac{1}{2}$.

Výpočet spojený s porušením B-C podmínky bývá často zdlouhavý. Občas se proto vyplatí jít na příklad oklikou.

Příklad 8.4.10. Ukažme, že řada $\sum_{k=1}^{\infty} \frac{\sin^2 k}{k}$ nekonverguje. Máme

$$\sum_{k=1}^{\infty} \frac{\sin^2 k}{k} = \sum_{k=1}^{\infty} \Big(\frac{1}{2k} - \frac{\cos(2k)}{2k}\Big).$$

Protože řada $\sum_{k=1}^{\infty} \frac{\cos(2k)}{2k}$ konverguje podle Dirichletova kritéria (Věta 8.4.1), pokud by konvergovala naše řada, podle aritmetiky řad by konvergovala i řada $\sum_{k=1}^{\infty} \frac{1}{2k}$ a tím bychom dostali spor.

8.5 Přerovnávání řad a součin řad

V dalším si ukážeme, že na součet absolutně konvergentní řady nemá přerovnání členů žádný vliv. Naproti tomu u neabsolutně konvergentních řad může mít tato operace závažné následky.

Definice 8.5.1 (Přerovnání řady)**.** Nechť $\{a_k\} \subset \mathbb{R}$ a $\varphi\colon \mathbb{N} \to \mathbb{N}$ je bijekce. Pak řadu $\sum_{k=1}^{\infty} a_{\varphi(k)}$ nazveme *přerovnáním* řady $\sum_{k=1}^{\infty} a_k$ (odpovídajícím bijekci φ).

Definice 8.5.2 (Kladná a záporná část)**.** Nechť $x \in \mathbb{R}$. *Kladnou část* čísla x definujeme jako $x^+ := \max\{x, 0\}$ a *zápornou část* jako $x^- := \max\{-x, 0\}$.

Příklad 8.5.3. Pro $x \geq 0$ máme $x^+ = x$ a $x^- = 0$, pro $x \leq 0$ máme $x^+ = 0$ a $x^- = -x = |x|$. Vždy platí $x = x^+ - x^-$ a $|x| = x^+ + x^-$.

Věta 8.5.4 (Charakterizace absolutní a neabsolutní konvergence)**.** *Nechť* $\{a_k\} \subset \mathbb{R}$. *Pak*
(i) $\sum_{k=1}^{\infty} a_k$ *konverguje absolutně* $\iff \sum_{k=1}^{\infty} a_k^+$ *a* $\sum_{k=1}^{\infty} a_k^-$ *konvergují.*
(ii) $\sum_{k=1}^{\infty} a_k$ *konverguje neabsolutně* $\implies \sum_{k=1}^{\infty} a_k^+ = \sum_{k=1}^{\infty} a_k^- = \infty$.

Důkaz. V části (i) plyne implikace „$\Rightarrow$“ z odhadů $0 \leq a_k^+ \leq |a_k|$ a $0 \leq a_k^- \leq |a_k|$. Implikace „$\Leftarrow$“ plyne z identity $|a_k| = a_k^+ + a_k^-$ a aritmetiky konvergentních řad.

V části (ii) máme $\infty = \sum_{k=1}^{\infty} |a_k| = \sum_{k=1}^{\infty} (a_k^+ + a_k^-)$. Alespoň jedna z řad na pravé straně implikace proto musí mít nekonečný součet. Protože zároveň $\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} (a_k^+ - a_k^-)$ konverguje, podle aritmetiky řad nemůže mít nekonečný součet právě jedna řada na pravé straně dokazované implikace. $\square$

Poznámka 8.5.5. Implikace v části (ii) se nedá otočit, jak ukazuje řada

$$1 - 2 + 1 - 2 + 1 - 2 + 1 - 2 + \dots.$$

Věta 8.5.6 (O přerovnání absolutně konvergentní řady)**.** *Nechť* $\{a_k\} \subset \mathbb{R}$ *a řada* $\sum_{k=1}^{\infty} a_k$ *konverguje absolutně. Pak každé její přerovnání konverguje absolutně a má stejný součet.*

Důkaz. Nechť $\sum_{k=1}^{\infty} b_k$ je přerovnáním $\sum_{k=1}^{\infty} a_k$. Nejprve uvažme jednoduchý případ $\{a_k\} \subset [0, \infty)$. Je-li $n \in \mathbb{N}$, pak existuje $k_0 \in \mathbb{N}$ takové, že $\{b_1, \dots, b_n\} \subset \{a_1, \dots, a_{k_0}\}$, a proto

$$\sum_{k=1}^{n} b_k \leq \sum_{k=1}^{k_0} a_k \leq \sum_{k=1}^{\infty} a_k \in \mathbb{R}.$$

Odtud $\sum_{k=1}^{\infty} b_k$ má omezené monotonní částečné součty, tedy (absolutně) konverguje a platí $\sum_{k=1}^{\infty} b_k \leq \sum_{k=1}^{\infty} a_k$. Prohozením rolí $\{a_k\}$ a $\{b_k\}$ dostáváme obrácenou nerovnost. Proto jsou součty obou řad stejné.

V obecném případě si napíšeme $\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} b_k^+ - \sum_{k=1}^{\infty} b_k^-$. Pro každou z řad na pravé straně platí výsledek dokázaný výše. Odtud

$$\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} b_k^+ - \sum_{k=1}^{\infty} b_k^- = \sum_{k=1}^{\infty} a_k^+ - \sum_{k=1}^{\infty} a_k^- = \sum_{k=1}^{\infty} a_k.$$

Absolutní konvergence plyne z konvergence řad $\sum_{k=1}^{\infty} b_k^+$, $\sum_{k=1}^{\infty} b_k^-$. $\square$

Věta 8.5.7 (Riemannova věta o přerovnání neabsolutně konvergentní řady)**.** *Nechť* $\{a_k\} \subset \mathbb{R}$ *a řada* $\sum_{k=1}^{\infty} a_k$ *konverguje neabsolutně. Pak pro každé* $S \in \mathbb{R}^*$ *existuje přerovnání řady* $\sum_{k=1}^{\infty} a_k$ *se součtem* S*.*

Důkaz. Máme $\sum_{k=1}^{\infty} a_k^+ = \sum_{k=1}^{\infty} a_k^- = \infty$ a $a_k \to 0$ pro $k \to \infty$. Nechť nejprve $S \in \mathbb{R}$. Zvolme $n_1 \in \mathbb{N}$ jako nejmenší přirozené číslo splňující

$$S_1 := a_1^+ + a_2^+ + \cdots + a_{n_1}^+ > S.$$

Dále vezmeme $m_1 \in \mathbb{N}$ jako nejmenší přirozené číslo splňující

$$S_2 := a_1^+ + a_2^+ + \cdots + a_{n_1}^+ - a_1^- - a_2^- - \ldots - a_{m_1}^- < S.$$

Nyní zase zvolíme $n_2 > n_1$ jako nejmenší číslo splňující

$$S_3 := a_1^+ + a_2^+ + \cdots + a_{n_1}^+ - a_1^- - a_2^- - \ldots - a_{m_1}^+ + a_{n_1+1}^+ + \cdots + a_{n_2}^+ > S.$$

Dále pokračujeme indukcí. Konstrukce se nikdy nezastaví, neboť $\sum_{k=1}^{\infty} a_k^+ = \sum_{k=1}^{\infty} a_k^- = \infty$. Navíc pro libovolné $j \in \mathbb{N}$, $j \geq 2$ máme

$$|S_{2j-1} - S| \leq a_{n_j}^+ \overset{j\to\infty}{\to} 0 \qquad a \qquad |S_{2j} - S| \leq a_{m_j}^- \overset{j\to\infty}{\to} 0,$$

neboť vždy $n_i \geq i$, $m_i \geq i$ a $a_k \to 0$. Celkově dostáváme přerovnanou řadu se součtem S.

Pokud $S = \infty$, provedeme variantu konstrukce s $S_1 > 1$, $S_2 < 1$, $S_3 > 2$, $S_4 < 2$, $S_5 > 3$, atd. Pro $S = -\infty$ pracujeme podobně. □

Naším dalším cílem je studovat součiny řad. Pro absolutně konvergentní řady $\sum_{k=1}^{\infty} a_k$ a $\sum_{k=1}^{\infty} b_k$ dostaneme vzorec

$$\Big(\sum_{k=1}^{\infty} a_k\Big)\Big(\sum_{k=1}^{\infty} b_k\Big) = \sum_{i,j=1}^{\infty} a_i b_j.$$

Výraz napravo obsahuje zápis sumy, s nímž jsme se dosud nesetkali a neumíme s ním pracovat. Začneme tedy opatrně definicí.

Definice 8.5.8 (Zobecněná řada a její konvergence)**.** Nechť M je spočetná množina (existuje bijekce mezi M a $\mathbb{N}$). Řekneme, že *zobecněná řada* $\sum_{m\in M} a_m$ konverguje, jestliže existuje taková bijekce $\varphi\colon \mathbb{N} \to M$, že $\sum_{k=1}^{\infty} a_{\varphi(k)}$ je absolutně konvergentní. Pak definujeme $\sum_{m\in M} a_m := \sum_{k=1}^{\infty} a_{\varphi(k)}$.

Poznámka 8.5.9. Protože absolutně konvergentní řady mají součet stabilní vůči přerovnání, pokud existuje jedna bijekce s vlastností z definice, všechny ostatní bijekce mezi $\mathbb{N}$ a M dávají absolutně konvergentní řady se stejným součtem.

Příklad 8.5.10. Uvažme $M = \mathbb{N}^2$ a řadu $\sum_{(i,j)\in\mathbb{N}^2} 2^{-(i+j)}$. Uvažme bijekci φ, která prvky $\mathbb{N}^2$ seřadí do posloupnosti

$$(1,1), (2,1), (1,2), (3,1), (2,2), (1,3), (4,1), \ldots.$$

V tomto případě máme

$$S := \sum_{k=1}^{\infty} a_{\varphi(k)} = 2^{-2} + 2^{-3} + 2^{-3} + 2^{-4} + 2^{-4} + 2^{-4} + 2^{-5} + \dots .$$

Dostáváme absolutně konvergentní řadu (porovnejte částečné součty naší řady s částečnými součty řady $\sum_{k=1}^{\infty} k2^{-k-1}$, jejíž konvergenci umíte ověřit pomocí odmocninového kritéria). Odtud $\sum_{(i,j)\in\mathbb{N}^2} 2^{-(i+j)} = S$ (zatím S neumíme vyčíslit, ale již brzy to umět budeme) a tento výsledek nezávisí na volbě bijekce mezi $\mathbb{N}^2$ a $\mathbb{N}$.

Poznámka 8.5.11. Často se pro $\sum_{(i,j)\in\mathbb{N}^2} a_{(i,j)}$ používá značení $\sum_{i,j=1}^{\infty} a_{i,j}$ nebo $\sum_{i,j=1}^{\infty} a_{ij}$.

Věta 8.5.12 (Cauchyova věta o součinu řad)**.** *Nechť* $\{a_k\}, \{b_k\} \subset \mathbb{R}$ *a nechť řady* $\sum_{k=1}^{\infty} a_k$ *a* $\sum_{k=1}^{\infty} b_k$ *konvergují absolutně. Pak je řada* $\sum_{i,j=1}^{\infty} a_i b_j$ *absolutně konvergentní a platí*

$$\sum_{i,j=1}^{\infty} a_i b_j = \Big(\sum_{k=1}^{\infty} a_k\Big)\Big(\sum_{k=1}^{\infty} b_k\Big).$$

Důkaz. Definujme $\widetilde{A}_n := \sum_{k=1}^{n} |a_k|$, $\widetilde{B}_n := \sum_{k=1}^{n} |b_k|$ pro každé $n \in \mathbb{N}$. Bijekci $\varphi\colon \mathbb{N} \to \mathbb{N}^2$ (jednotlivé složky budeme později značit φ_1 a φ_2) tentokrát zaveďme konstrukcí

$$(1,1), (2,1), (2,2), (1,2), (3,1), (3,2), (3,3), (2,3), (1,3), (4,1), \dots .$$

Dále definujme $S_n := \sum_{k=1}^{n} a_{\varphi_1(k)} b_{\varphi_2(k)}$ a $\widetilde{S}_n := \sum_{k=1}^{n} |a_{\varphi_1(k)}||b_{\varphi_2(k)}|$ pro každé $n \in \mathbb{N}$.

Absolutní konvergence $\sum_{i,j=1}^{\infty} a_i b_j$ plyne z toho, že $\{\widetilde{S}_n\}$ je neklesající posloupnost splňující

$$\widetilde{S}_{n^2} = \sum_{k=1}^{n^2} |a_{\varphi_1(k)}||b_{\varphi_2(k)}| = \sum_{i=1}^{n} |a_i| \sum_{j=1}^{n} |b_j| = \widetilde{A}_n \widetilde{B}_n \le \sum_{i=1}^{\infty} |a_i| \sum_{j=1}^{\infty} |b_j|.$$

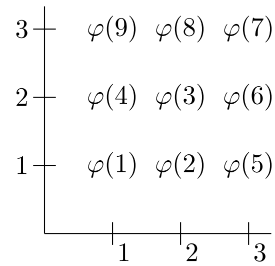
Proto je také $\sum_{k=1}^{\infty} a_{\varphi_1(k)} b_{\varphi_2(k)}$ konvergentní a platí

$$\begin{aligned}\sum_{i,j=1}^{\infty} a_i b_j :&= \sum_{k=1}^{\infty} a_{\varphi_1(k)} b_{\varphi_2(k)} = \lim_{n\to\infty} S_n = \lim_{n\to\infty} S_{n^2} = \lim_{n\to\infty} \sum_{k=1}^{n^2} a_{\varphi_1(k)} b_{\varphi_2(k)} \\ &= \lim_{n\to\infty} \Big(\sum_{i=1}^{n} a_i\Big)\Big(\sum_{j=1}^{n} b_j\Big) = \sum_{i=1}^{\infty} a_i \sum_{j=1}^{\infty} b_j.\end{aligned}$$

□

Příklad 8.5.13. Díky Větě 8.5.12 dostáváme, že

$$\sum_{(i,j)\in\mathbb{N}^2} 2^{-(i+j)} = \sum_{i=1}^{\infty} 2^{-i} \sum_{j=1}^{\infty} 2^{-j} = 1.$$

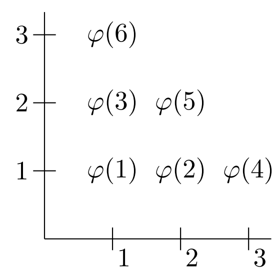


Obrázek 8.2: Částečné znázornění bijekce z důkazu Cauchyovy věty.

Poznámka 8.5.14. Někdy se používá pro součin řad jiná bijekce, která se zapisuje jako

$$\sum_{i,j=1}^{\infty} a_i b_j = \sum_{n=1}^{\infty} \Big(\sum_{i+j=n+1} a_i b_j \Big).$$

Tento vztah se také někdy nazývá Cauchyův vzorec.



Obrázek 8.3: Částečné znázornění bijekce z Poznámky 8.1.7.

8.6 Metoda aritmetických průměrů a cesarovské součty

Nyní se budeme zabývat otázkou, zda je možné nekonvergentní řadě přiřadit číslo, které bude mít alespoň částečně vlastnosti jejího součtu. Náš přístup bude založen na následující konstrukci.

Lemma 8.6.1 (O konvergenci aritmetických průměrů)**.** *Nechť* $\{a_k\} \subset \mathbb{R}$ *splňuje* $\lim_{k\to\infty} a_k = A \in \mathbb{R}^*$*. Definujme posloupnost* $\{b_k\}$ *předpisem*

$$b_1 = a_1,\ b_2 = \frac{a_1 + a_2}{2},\ b_3 = \frac{a_1 + a_2 + a_3}{3},\ \ldots, \qquad \textit{neboli} \qquad b_j = \frac{1}{j} \sum_{k=1}^{j} a_k.$$

Pak $\lim_{k\to\infty} b_k = A$.

Důkaz. Budeme se zabývat jen případem $A \in \mathbb{R}$, v ostatních případech se použije podobná myšlenka. Zvolme $\varepsilon > 0$. Pak existuje $k_0 \in \mathbb{N}$ tak, že $A-\varepsilon < a_k < A+\varepsilon$ pro $k > k_0$. Je-li potom $k > k_0$ dostatečně velké, dostáváme

$$b_k = \frac{a_1 + \cdots + a_k}{k} = \frac{a_1 + \cdots + a_{k_0}}{k} + \frac{a_{k_0+1} + \cdots + a_k}{k-k_0}\,\frac{k-k_0}{k} < \varepsilon + A + \varepsilon$$

a

$$\begin{aligned} b_k &= \frac{a_1 + \cdots + a_{k_0}}{k} + \frac{a_{k_0+1} + \cdots + a_k}{k-k_0} - \frac{k_0}{k}\,\frac{a_{k_0+1} + \cdots + a_k}{k-k_0} \\ &> -\varepsilon + A - \varepsilon - \varepsilon(|A| + \varepsilon). \end{aligned}$$

Proto $\lim_{k\to\infty} b_k = A$. □

Poznámka 8.6.2. Obrácená implikace neplatí. Abychom to demonstrovali, uvažme posloupnost $\{a_k\} = \{(-1)^k\}$. Tato posloupnost limitu nemá. Pro aritmetické průměry však platí

$$\{b_k\} = \{-1, 0, -\tfrac{1}{3}, 0, -\tfrac{1}{5}, \dots\} \qquad a \qquad \lim_{k\to\infty} b_k = 0.$$

Příklad 8.6.3. Z teorie Taylorových rozvojů víme, že

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \quad \text{na } (-1,1).$$

Jinými slovy, máme posloupnost polynomů $\{P_k\}$ takových, že st $P_k = k$ pro každé $k \in \mathbb{N}_0$ a $P_k(x) \to \frac{1}{1-x}$ pro každé $x \in (-1,1)$ (a v žádném jiném bodě to neplatí). Pokud však definujeme polynomy

$$Q_k := \frac{1}{k+1} \sum_{j=0}^{k} P_j,$$

dostáváme posloupnost polynomů stupně k s o něco lepší aproximační vlastností $Q_k(x) \to \frac{1}{1-x}$ pro kažké $x \in [-1,1)$.

Poznámka 8.6.4. Výsledek předchozího příkladu, tedy získání konvergence v jednom bodě navíc, není příliš oslnivý. Později se budeme zabývat teorií Fourierových řad, tedy rozvojů typu

$$f(x) = \sum_{k=0}^{\infty} \big(c_k \cos(kx) + d_k \sin(kx)\big)$$

($\{a_k\}, \{b_k\}$ jsou posloupnosti reálných koeficientů), kde metoda aritmetických průměrů přináší podstatně zajímavější výsledky. Poznamenejme ještě, že Fourierovy řady mají široké uplatnění od teorie parciálních diferenciálních rovnic až třeba po zpracování zvukového záznamu.

Metoda aritmetických průměrů aplikovaná na částečné součty posloupnosti $\{a_k\}$ má vlastní stručnou terminologii danou následující definicí.

Definice 8.6.5 (Cesarovská sčítatelnost)**.** Nechť $\{a_n\} \subset \mathbb{R}$. Pro všechna $n \in \mathbb{N}$ definujme $S_n = \sum_{k=1}^{n} a_k$ a $\sigma_n = \frac{1}{n} \sum_{k=1}^{n} S_k$. Řekneme, že $\sum_{k=1}^{\infty} a_k$ je *cesarovsky sčítatelná*, jestliže $\lim_{n\to\infty} \sigma_n \in \mathbb{R}$. Číslo $A := \lim_{n\to\infty} \sigma_n$ pak nazveme *cesarovským součtem* řady $\sum_{k=1}^{\infty} a_k$ a píšeme $(C,1) \sum_{k=1}^{\infty} a_k = A$.

Poznámka 8.6.6. Podle Lemmatu o konvergenci aritmetických průměrů (Lemma 8.6.1) je každá konvergentní řada cesarovsky sčítatelná a součty v obou smyslech jsou totožné.

Poznámka 8.6.7. Metoda aritmetických průměrů se dá iterovat. Pro posloupnost $\{a_k\}$ definujme (použijeme trochu odlišné značení od definice)

$$s_n^0 := s_n = \sum_{k=1}^{n} a_k, \qquad s_n^1 := \frac{1}{n} \sum_{k=1}^{n} s_k^0, \qquad s_n^2 := \frac{1}{n} \sum_{k=1}^{n} s_k^1, \qquad \dots$$

(pozor, horní index není mocnina). Řadu $\sum_{k=1}^{\infty} a_k$ nazveme (C,r)-sčítatelnou, jestliže $\lim_{n\to\infty} s_n^r \in \mathbb{R}$. Pak píšeme $(C,r) \sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} s_n^r$. Je-li řada (C,r)-sčítatelná, je i (C,s)-sčítatelná pro každé $s \geq r$ (podle Lemmatu o konvergenci aritmetických průměrů). Tato implikace se nedá obrátit.

Příklad 8.6.8. Položme $\{a_k\} = \{1, -2, 3, -4, 5, \dots\}$. Pak posloupnost částečných součtů je $\{S_k\} = \{s_k^0\} = \{1, -1, 2, -2, 3, \dots\}$, a proto $\sum_{k=1}^{\infty} a_k$ nekonverguje (klasická konvergence je totéž, co $(C,0)$-sčítatelnost). Dále máme $\{s_k^1\} = \{1, 0, \frac{2}{3}, 0, \frac{3}{5}, \dots\}$. Cesarovské součty (podle definice) řady $\sum_{k=1}^{\infty} a_k$ také nekonvergují. Povšimneme-li si však, že pro každé $m \in \mathbb{N}$ platí $s_{2m}^1 = 0$ a $s_{2m-1}^1 = \frac{m}{2m-1} \to \frac{1}{2}$, plyne odsud, že

$$\lim_{n\to\infty} s_n^2 = \frac{1}{4}, \qquad \text{a proto} \qquad (C,2) \sum_{k=1}^{\infty} a_k = \frac{1}{4}.$$

8.7 Dodatek k číselným řadám: nekonečné součiny

Nechť $\{p_k\} \subset (0, \infty)$ je posloupnost. Symbol $\prod_{k=1}^{\infty} p_k$ nazýváme *nekonečným součinem*. K jeho vyčíslení definujme $P_n = p_1 p_2 \dots p_n$. Nekonečný součin nazveme *konvergentní*, jestliže existuje vlastní nenulová $\lim_{n\to\infty} P_n =: P$. Pak píšeme

$$\prod_{k=1}^{\infty} p_k := P.$$

Věta 8.7.1 (Nutná podmínka konvergence)**.** *Nechť posloupnost* $\{p_k\} \subset (0, \infty)$ *a nechť* $\prod_{k=1}^{\infty} p_k$ *konverguje. Pak* $\lim_{n\to\infty} p_n = 1$.

Důkaz. Jestliže $\prod_{k=1}^{\infty} p_k$ konverguje, máme

$$p_{k+1} = \frac{P_{k+1}}{P_k} \to \frac{P}{P} = 1.$$

□

Poznámka 8.7.2. Vynechání, přidání či změna konečného počtu činitelů neovlivní konvergenci nekonečného součinu.

Pokud $P_n \to P \in (0, \infty)$, máme ze spojitosti funkce log

$$\begin{aligned} \log P &= \log(\lim_{n\to\infty} P_n) = \lim_{n\to\infty} \log(P_n) \\ &= \lim_{n\to\infty} \log(p_1 \dots p_n) = \lim_{n\to\infty} \sum_{k=1}^{n} \log p_k = \sum_{k=1}^{\infty} \log p_k, \end{aligned}$$

tedy $\sum_{k=1}^{\infty} \log p_k$ konverguje. Tato implikace se dá zřejmě otočit. Existuje ale ještě jednodušší charakterizace konvergence nekonečného součinu, v níž pracujeme s $u_k := p_k - 1 \in (-1, \infty)$.

Věta 8.7.3 (Ekvivalentní charakterizace konvergence součinu)**.** *Nechť posloupnost $\{u_k\} \subset (0, \infty)$ nebo $\{u_k\} \subset (-1, 0)$. Pak nekonečný součin $\prod_{k=1}^{\infty}(1 + u_k)$ konverguje právě tehdy, když konverguje řada $\sum_{k=1}^{\infty} u_k$.*

Důkaz. Nejprve si povšimněme, že díky tomu, že $\lim_{x\to 0} \frac{\log(1+x)}{x} = 1$, existuje $\delta > 0$ takové, že

$$\frac{1}{2} \le \frac{\log(1 + x)}{x} \le 2 \qquad \text{pro } x \in (-\delta, \delta) \setminus \{0\}.$$

V dalším uvažujme případ $\{u_k\} \subset (0, \infty)$. Dokažme implikaci „$\Rightarrow$“. Pokud konverguje $\prod_{k=1}^{\infty}(1 + u_k)$, z nutné podmínky konvergence součinu dostáváme $u_k \to 0$. Proto $u_k < \delta$ od jistého $k_0 \in \mathbb{N}$. Navíc jsme si výše ukázali, že $\sum_{k=1}^{\infty} \log(1 + u_k)$ konverguje. Celkově

$$-\infty < \frac{1}{2} \sum_{k=k_0}^{\infty} \log(1 + u_k) \le \sum_{k=k_0}^{\infty} u_k \le 2 \sum_{k=k_0}^{\infty} \log(1 + u_k) < \infty.$$

Protože konvergence řady nezávisí na chování konečného počtu členů, máme konvergenci $\sum_{k=1}^{\infty} u_k$.

Dokažme nyní implikaci „$\Leftarrow$“. Pokud konverguje $\sum_{k=k_0}^{\infty} u_k$, z nutné podmínky konvergence řady dostáváme, že $u_k < \delta$ od jistého $k_0 \in \mathbb{N}$. Důkaz dokončíme pomocí nerovností

$$-\infty < \frac{1}{2} \sum_{k=k_0}^{\infty} u_k \le \sum_{k=k_0}^{\infty} \log(1 + u_k) \le 2 \sum_{k=k_0}^{\infty} u_k < \infty.$$

V případě $\{u_k\} \subset (-1, 0)$ postupujeme podobně. □

Příklad 8.7.4 (Cantorovo discontinuum kladné délky)**.** V kapitole z prvního dílu skript věnované hlubším vlastnostem spojitých a diferencovatelných funkcí jsme si představili Cantorovo discontinuum $C \subset [0,1]$. Získali jsme jej tak, že jsme z intervalu $[0,1]$ nejprve vynechali prostřední třetinu. V dalším kroku vynecháme prostřední třetinu v každém ze vzniklých podintervalů a takto pokračujeme dále. Získáme neprázdnou množinu C (například 0, 1, $\frac{1}{3}$, $\frac{2}{3}$ leží v C). Dá se nahlédnout, že každý bod této množiny se dá ztotožnit s nekonečnou posloupností nul a jedniček. Množina C je tedy nespočetná a má stejnou mohutnost jako $[0,1]$. Na druhou stranu vzniklá množina je v jistém smyslu velice malá, neboť v každém okolí libovolného bodu z $[0,1]$ najdeme otevřený interval, který má s C prázdný průnik (srovnejte s racionálními čísly, která mají s každým otevřeným intervalem neprázdný průnik, třebaže jsou spočetná). Navíc celková délka vynechaných intervalů je

$$\frac{1}{3}+\frac{2}{3}\frac{1}{3}+\frac{2}{3}\frac{2}{3}\frac{1}{3}+\cdots=\frac{1}{3}\sum_{k=0}^{\infty}\left(\frac{2}{3}\right)^k=\frac{1}{3}\frac{1}{1-\frac{2}{3}}=1.$$

Pokud bychom nevynechávali prostřední třetinu, ale interval délky $q \in (0,1)$, dostali bychom stejné vlastnosti, neboť

$$q+(1-q)q+(1-q)^2q+\cdots=q\sum_{k=0}^{\infty}(1-q)^k=q\frac{1}{1-(1-q)}=1.$$

Podívejme se na věc nyní trochu jinak, po k-tém kroku má ořezaná množina celkovou délku q^n. Pokud budeme v jednotlivých krocích vhodně měnit délku vynechaných částí, můžeme dostat odlišný výsledek. Skutečně, protože například $\sum_{k=2}^{\infty}\frac{1}{k^2}$ konverguje, konverguje rovněž nekonečný součin $\prod_{k=2}^{\infty}(1-\frac{1}{k^2})$. Pokud tedy v prvním kroku vynecháme prostřední čtvrtinu intervalu $[0,1]$, ve druhém kroku prostřední devítinu vzniklých intervalů, atd., získáme Cantorovo discontinuum kladné délky.

Kapitola 9

Mocninné řady

V následujícím textu se budeme zabývat zobecněním Taylorových rozvojů. Bude nás zajímat, na jakých množinách studované řady konvergují, ale i hlubší výsledky, jako jsou rovnosti typu $(\sum_{k=0}^{\infty} a_k x^k)' = \sum_{k=0}^{\infty} k a_k x^{k-1}$. Tyto rovnosti neplynou z aritmetiky derivace (ta si v kombinaci s matematickou indukcí poradí jen s konečnými součty), dokonce ani nemají šanci platit obecně pro řady funkcí, jak si později ukážeme v kapitole o stejnoměrné konvergenci.

V celé kapitole budeme pracovat především v komplexním oboru. Poznamenejme, že v komplexním oboru absolutní konvergenci definujeme jako konvergenci řady (komplexních) velikostí jejích členů. Snadno se ověří, že absolutní konvergence opět implikuje konvergenci.

9.1 Základní vlastnosti mocninných řad

Definice 9.1.1 (Mocninná řada)**.** Nechť $\{a_n\} \subset \mathbb{C}$ a $z_0 \in \mathbb{C}$. Pak řadu

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$

nazýváme *mocninnou řadou* se *středem* z_0. Čísla a_k nazýváme *koeficienty* mocninné řady.

Často je výhodné použít přeznačení $w := z - z_0$, které vede na mocninnou řadu se středem v počátku (a hlavně s jednodušším zápisem). Proto nám stačí veškeré výsledky formulovat pro řady se středem v počátku. Pro $\varrho > 0$ a $z_0 \in \mathbb{C}$ budeme značit

$$B_\varrho(z_0) := \{z \in \mathbb{C} \colon |z - z_0| < \varrho\} \qquad \text{a} \qquad B_\varrho := B_\varrho(0) = \{z \in \mathbb{C} \colon |z| < \varrho\}.$$

Věta 9.1.2 (O konvergenci mocninné řady)**.** *Nechť* $\{a_k\} \subset \mathbb{C}$*. Položme*

$$R := \frac{1}{\limsup_{k\to\infty} \sqrt[k]{|a_k|}} \qquad \textit{s konvencí } \frac{1}{0} = \infty \textit{ a } \frac{1}{\infty} = 0.$$

Pak
(i) *řada* $\sum_{k=0}^{\infty} a_k z^k$ *konverguje absolutně na* B_R
(ii) *řada* $\sum_{k=0}^{\infty} a_k z^k$ *nekonverguje na* $\{z \in \mathbb{C}\colon |z| > R\}$
(iii) *existuje-li* $\lim_{k\to\infty} |\frac{a_k}{a_{k+1}}|$, *pak se rovná* R
(iv) *existuje-li* $\lim_{k\to\infty} \sqrt[k]{|a_k|}$, *pak se rovná* $\frac{1}{R}$.

Důkaz. Část (iv) je zřejmá a část (i) snadno plyne z odmocninového kritéria. Dokažme (ii) v případě $R \in (0,\infty)$. V tomto případě pišme $|z| = (1+\delta)R$, kde $\delta > 0$, a máme

$$|a_k z^k| = |a_k||z|^k = \Big((1+\delta)\frac{\sqrt[k]{|a_k|}}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}}\Big)^k.$$

Z definice veličiny $\limsup$ musí být vnitřek závorky větší než 1 pro nekonečně mnoho indexů. Je tedy porušena nutná podmínka konvergence. Pro $R = \infty$ není co dokazovat a pokud $R = 0$, pro každé $z \neq 0$ najdeme nekonečně mnoho indexů, kde $\sqrt[k]{|a_k|} \geq \frac{1}{|z|}$, což opět vede na porušení nutné podmínky.

Dokažme zbývající část (iii). Položme $r := \lim_{k\to\infty} |\frac{a_k}{a_{k+1}}|$. Pro zafixované $z \in \mathbb{C}$ pak máme

$$\frac{|a_{k+1}z^{k+1}|}{|a_k z^k|} = \frac{|z|}{|\frac{a_k}{a_{k+1}}|} \overset{k\to\infty}{\to} \frac{|z|}{r}.$$

Podle podílového kritéria naše řada konverguje absolutně pro $|z| < r$, naopak pro $|z| > r$ absolutně konvergovat nemůže. Podle výsledků (i) a (ii) tedy máme $r = R$. □

Definice 9.1.3 (Poloměr konvergence mocninné řady)**.** Číslo $R \in [0,\infty]$ z minulé věty se nazývá *poloměr konvergence* mocninné řady $\sum_{k=0}^{\infty} a_k z^k$.

Poznámka 9.1.4. Předchozí věta nic neříká o konvergenci pro $|z| = R$. Situace se zde liší případ od případu.

Příklad 9.1.5. (i) Řada $\sum_{k=0}^{\infty} z^k$ má $R = 1$ ($\forall k \in \mathbb{N} \quad \sqrt[k]{a_k} = 1$) a pro $|z| = 1$ je vždy porušena nutná podmínka konvergence.
(ii) Řada $\sum_{k=0}^{\infty} \frac{z^k}{k^2}$ má $R = 1$ ($\frac{a_k}{a_{k+1}} \to 1$) a pro $|z| = 1$ vždy absolutně konverguje, neboť

$$\sum_{k=0}^{\infty} \frac{|z^k|}{k^2} = \sum_{k=0}^{\infty} \frac{1}{k^2} \in \mathbb{R}.$$

(iii) Řada $\sum_{k=0}^{\infty} \frac{z^k}{k}$ má $R = 1$ ($\frac{a_{k+1}}{a_k} \to 1$). Pro $z = 1$ dostáváme harmonickou řadu. Pro $|z| = 1$, $z \neq 1$, dává Dirichletovo kritérium neabsolutní konvergenci (zde $z = \mathrm{e}^{i\theta}$, kde $\theta \in (0, 2\pi)$, a $\{z^k\} = \{\mathrm{e}^{i\theta k}\}$ je geometrická posloupnost s omezenými částečnými součty).
(iv) Řada $\sum_{k=0}^{\infty} \frac{z^k}{k!}$ má $R = \infty$, neboť $\frac{a_k}{a_{k+1}} = \frac{(k+1)!}{k!} = k \to \infty$.
(v) Řada $\sum_{k=0}^{\infty} k! z^k$ má $R = 0$, neboť $\frac{a_k}{a_{k+1}} = \frac{k!}{(k+1)!} = \frac{1}{k} \to 0$.

Nyní se budeme zabývat otázkou derivování mocninných řad. Budeme potřebovat verzi aritmetiky limit pro $\limsup$.

Cvičení 9.1.6. Nechť $\{b_k\}, \{c_k\} \subset [0,\infty)$.
(i) Dokažte, že $\limsup_{k\to\infty} b_k c_k \le \limsup_{k\to\infty} b_k \limsup_{k\to\infty} c_k$, pokud má pravá strana smysl v $\mathbb{R}^*$.
(ii) Ukažte, že výše obecně neplatí rovnost.
(iii) Ukažte, že má-li alespoň jedna z posloupností limitu, platí rovnost.

Protože $\sqrt[k]{k} \to 1$ pro $k \to \infty$, z předchozího cvičení okamžitě dostáváme následující výsledek.

Lemma 9.1.7. *Nechť $\{a_k\} \subset \mathbb{C}$. Pak mocninné řady $\sum_{k=0}^{\infty} a_k z^k$ a $\sum_{k=1}^{\infty} k a_k z^{k-1}$ mají stejný poloměr konvergence.*

Ještě potřebujeme jeden jednoduchý odhad.

Lemma 9.1.8. *Nechť $\alpha, \beta > 0$ a $k \in \mathbb{N}$, $k \ge 2$. Pak*

$$(\alpha+\beta)^k - \alpha^k - k\alpha^{k-1}\beta \le \frac{k(k-1)(\alpha+\beta)^{k-2}}{2}\beta^2.$$

Důkaz. Definujme $f(x) = (\alpha + x)^k$. Pro odpovídající Taylorův polynom stupně 1 máme

$$T_1(\beta) = \alpha^k + k\alpha^{k-1}\beta$$

a pro Lagrangeův tvar zbytku R_2 v bodě β platí ($\xi \in (0,\beta)$)

$$R_2(\beta) = \frac{f''(\alpha+\xi)}{2!}\beta^2 = \frac{k(k-1)(\alpha+\xi)^{k-2}}{2}\beta^2 \le \frac{k(k-1)(\alpha+\beta)^{k-2}}{2}\beta^2.$$

Odtud plyne dokazovaná nerovnost. □

Věta 9.1.9 (Derivace mocninné řady)**.** *Nechť $\{a_k\} \subset \mathbb{C}$. Pak pro $x \in (-R,R)$, kde $R > 0$ je poloměr konvergence příslušné mocninné řady, platí*

$$\Big(\sum_{k=0}^{\infty} a_k x^k\Big)' = \sum_{k=1}^{\infty} k a_k x^{k-1}.$$

Důkaz. Provedeme přímé ověření definice derivace. Zafixujme $x \in (-R,R)$ a $\delta < R - |x|$. Pak pro $h \in (-\delta,\delta) \setminus \{0\}$ má podle aritmetiky konvergentních řad dobrý smysl veličina

$$\begin{aligned}\Psi(x,h) :&= \frac{\sum_{k=0}^{\infty} a_k (x+h)^k - \sum_{k=0}^{\infty} a_k x^k}{h} - \sum_{k=1}^{\infty} k a_k x^{k-1} \\ &= \sum_{k=2}^{\infty} \Big(a_k \frac{(x+h)^k - x^k}{h} - k a_k x^{k-1}\Big).\end{aligned}$$

Naším cílem je ukázat, že $\Psi(x,h) \to 0$ pro $h \to 0$. Použijme Lemma 9.1.8 a skutečnost, že $\sum_{k=2}^{\infty} k(k+1)|a_k|(|x|+\delta)^{k-2}$ konverguje (neboť $|x| + \delta < R$), tedy

Lemma 9.1.7. Potom

$$
\begin{aligned}
|\Psi(x,h)| &= \frac{1}{|h|}\Big|\sum_{k=2}^{\infty} a_k\big((x+h)^k - x^k - kx^{k-1}h\big)\Big| \\
&= \frac{1}{|h|}\Big|\sum_{k=2}^{\infty} a_k\Big(\binom{k}{2}x^{k-2}h^2 + \dots + \binom{k}{k}h^k\Big)\Big| \\
&\le \frac{1}{|h|}\sum_{k=2}^{\infty} |a_k|\Big(\binom{k}{2}|x|^{k-2}|h|^2 + \dots + \binom{k}{k}|h|^k\Big) \\
&= \frac{1}{|h|}\sum_{k=2}^{\infty} |a_k|\Big((|x|+|h|)^k - |x|^k - k|x|^{k-1}|h|\Big) \\
&\le \frac{1}{|h|}\sum_{k=2}^{\infty} |a_k|\frac{k(k-1)(|x|+|h|)^{k-2}}{2}|h|^2 \\
&\le \frac{1}{2}|h|\sum_{k=2}^{\infty} |a_k|k(k-1)(|x|+\delta)^{k-2} = C|h| \to 0.
\end{aligned}
$$

□

Poznámka 9.1.10. Mohli bychom pokračovat indukcí s dalšími derivacemi a řadami $\sum_{k=2}^{\infty} k(k-1)a_k x^{k-2}$, $\sum_{k=3}^{\infty} k(k-1)(k-2)a_k x^{k-3}$, atd.

Důsledek 9.1.11. *Každá mocninná řada na svém kruhu konvergence definuje nekonečněkrát spojitě diferencovatelnou funkci.*

Mocninné řady můžeme na jejich kruhu konvergence rovněž integrovat.

Věta 9.1.12 (Integrace mocninné řady)**.** *Nechť* $\{a_k\} \subset \mathbb{C}$.
(i) *Pro* $x \in \mathbb{R}$ *ležící uvnitř konvergenčního kruhu platí*

$$
\int \Big(\sum_{k=0}^{\infty} a_k x^k\Big)\,\mathrm{d}x = \sum_{k=0}^{\infty} \frac{a_k}{k+1}x^{k+1} + C.
$$

(ii) *Jestliže* $a, b \in (-R, R)$*, kde* R *je poloměr konvergence řady* $\sum_{k=0}^{\infty} a_k z^k$*, pak*

$$
\begin{aligned}
(\mathcal{R})\int_a^b \Big(\sum_{k=0}^{\infty} a_k x^k\Big)\,\mathrm{d}x &= (\mathcal{N})\int_a^b \Big(\sum_{k=0}^{\infty} a_k x^k\Big)\,\mathrm{d}x \\
= \sum_{k=0}^{\infty}(\mathcal{N})\int_a^b a_k x^k\,\mathrm{d}x &= \sum_{k=0}^{\infty}(\mathcal{R})\int_a^b a_k x^k\,\mathrm{d}x.
\end{aligned}
$$

Důkaz. První část věty je jen důsledkem Věty o derivaci mocninné řady (Věta 9.1.9). Dokažme druhou část věty. Předpokládejme $a < b$ (pro $b < a$ pracujeme podobně, pro $a = b$ je důkaz triviální). První a třetí rovnost ve druhé části věty plynou z toho, že Riemannův a Newtonův integrál se pro omezené spojité funkce rovnají (pro rovnost nalevo spojitost plyne z diferencovatelnosti a omezenost plyne ze spojitosti na omezeném uzavřeném intervalu $[a,b]$). Nyní podle první části věty máme, že

$$F(x) := \sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}$$

je primitivní funkce k $\sum_{k=0}^{\infty} a_k x^k$ na $(-R,R)$, a proto díky aritmetice konvergentních řad

$$\begin{aligned}(\mathcal{N}) \int_a^b \Big(\sum_{k=0}^{\infty} a_k x^k\Big)\,\mathrm{d}x &= F(b) - F(a) \\ = \sum_{k=0}^{\infty} \frac{a_k}{k+1}(b^{k+1} - a^{k+1}) &= \sum_{k=0}^{\infty} (\mathcal{N}) \int_a^b a_k x^k \,\mathrm{d}x.\end{aligned}$$

□

Příklad 9.1.13. Sečtěme mocninnou řadu $\sum_{k=1}^{\infty} kx^k$. Poloměr konvergence je roven jedné, proto budeme pracovat na intervalu $(-1,1)$. Platí

$$\sum_{k=1}^{\infty} kx^k = x\sum_{k=1}^{\infty} kx^{k-1} = x\sum_{k=1}^{\infty} (x^k)' = x\Big(\sum_{k=1}^{\infty} x^k\Big)'.$$

Odtud

$$\sum_{k=1}^{\infty} kx^k = x\Big(\frac{x}{1-x}\Big)' = x\frac{1\cdot(1-x) - x\cdot(-1)}{(1-x)^2} = \frac{x}{(1-x)^2} \qquad \text{pro } x \in (-1,1).$$

Příklad 9.1.14. Sečtěme řadu $\sum_{k=1}^{\infty} \frac{k^2}{2^k}$. Povšimněme si, že výsledek získáme dosazením $x = \frac{1}{2}$ do mocniné řady $\sum_{k=1}^{\infty} k^2 x^k$ s poloměrem konvergence rovným jedné. Na intervalu $(-1,1)$ proto platí (využijeme výsledek předchozího příkladu)

$$\begin{aligned}\sum_{k=1}^{\infty} k^2 x^k &= x\sum_{k=1}^{\infty} k^2 x^{k-1} = x\sum_{k=1}^{\infty} (kx^k)' = x\Big(\sum_{k=1}^{\infty} kx^k\Big)' = x\Big(\frac{x}{(1-x)^2}\Big)' \\ &= x\frac{(1-x)^2 + 2x(1-x)}{(1-x)^4} = \frac{x(1+x)}{(1-x)^3}.\end{aligned}$$

Odtud

$$\sum_{k=1}^{\infty} \frac{k^2}{2^k} = \frac{\frac{1}{2}\frac{3}{2}}{(\frac{1}{2})^3} = 6.$$

Příklad 9.1.15. Sečtěme řadu $\sum_{k=1}^{\infty} \frac{x^{k-1}}{k}$. Poloměr konvergence je roven jedné a na $(-1,1) \setminus \{0\}$ platí

$$\begin{aligned}\sum_{k=1}^{\infty} \frac{x^{k-1}}{k} &= \frac{1}{x}\sum_{k=1}^{\infty} \frac{x^k}{k} = \frac{1}{x}\sum_{k=1}^{\infty} \int x^{k-1}\,\mathrm{d}x = \frac{1}{x}\int \sum_{k=1}^{\infty} x^{k-1}\,\mathrm{d}x = \frac{1}{x}\int \frac{1}{1-x}\,\mathrm{d}x \\ &= \frac{1}{x}\Big(-\log(1-x) + C\Big) = -\frac{\log(1-x)}{x} + \frac{C}{x}.\end{aligned}$$

Protože funkce úplně nalevo se v počátku rovná jedné, musí jít funkce napravo spojitě dodefinovat v počátku stejnou hodnotou. Odtud $C = 0$ a

$$\sum_{k=1}^{\infty} \frac{x^{k-1}}{k} = \begin{cases} -\frac{\log(1-x)}{x} & \text{pro } x \in (-1,1) \setminus \{0\} \\ 1 & \text{pro } x = 0. \end{cases}$$

9.2 Dodatek k teorii mocninných řad: derivace funkce komplexní proměnné

Pokud rozšíříme pojem derivace na funkce komplexní proměnné, můžeme zesílit výsledky obdržené v předchozí sekci.

Definice 9.2.1. Nechť $f\colon \mathbb{C} \to \mathbb{C}$ a $z \in \mathbb{C}$. Řekneme, že funkce f má v bodě z *derivaci* $A \in \mathbb{C}$, jestliže

$$\lim_{h\to 0} \frac{f(z+h) - f(z)}{h} = A.$$

Pak píšeme $f(z) = A$.

Poznámka 9.2.2. (i) Limita v komplexním oboru je opět definována za pomocí okolí (tedy $\lim_{z\to z_0} f(z) = A \Leftrightarrow \forall \varepsilon > 0\, \exists \delta > 0 \quad z \in \mathcal{P}_\delta(z_0) \Rightarrow f(z) \in \mathcal{U}_\varepsilon(A)$) a ještě připomeňme, že pro $z_0 \in \mathbb{C}$ jsme definovali $\mathcal{P}_\delta(z_0) = \{z \in \mathbb{C}\colon 0 < |z - z_0| < \delta\}$ a $\mathcal{U}_\delta(z_0) = \{z \in \mathbb{C}\colon |z - z_0| < \delta\}$.
(ii) V komplexním oboru nezavádíme nevlastní derivaci (podobně jako jsme nezavedli nevlastní parciální derivace).
(iii) Pro komplexní derivaci se dají dokazovat podobné výsledky jako pro derivaci reálnou, například aritmetika derivace. Ovšem pozor na absenci uspořádání na $\mathbb{C}$.
(iv) Rozmyslete si, že má-li komplexní funkce reálnou derivaci v bodě s nulovou imaginární složkou, má i restrikce naší funkce na reálnou osu v odpovídajícím bodě stejnou derivaci.

Příklad 9.2.3. (i) Pokud $f(z) = z$, máme

$$f'(z) = \lim_{h\to 0} \frac{f(z+h) - f(z)}{h} = \lim_{h\to 0} \frac{z + h - z}{h} = 1.$$

(ii) Pokud $f(z) = z^2$, máme

$$f'(z) = \lim_{h\to 0} \frac{f(z+h) - f(z)}{h} = \lim_{h\to 0} \frac{z^2 + 2hz + h^2 - z^2}{h} = \lim_{h\to 0} 2z + \lim_{h\to 0} h = 2z.$$

(iii) Pokud $f(z) = z^n$, kde $n \in \mathbb{N}$, analogicky jako výše dostáváme $f'(z) = nz^{n-1}$.
(iv) Rovnost $(x^2)' = 2x$ plyne z rovnosti $(z^2)' = 2z$.

Připomeňme, že Lemma 9.1.7 platí i pro komplexní případ, proto také Větu 9.1.9 lze formulovat i pro komplexní případ.

Věta 9.2.4 (Derivace mocninné řady)**.** *Nechť* $\{a_n\} \subset \mathbb{C}$*. Pak uvnitř konvergenčního kruhu platí*

$$\left(\sum_{k=0}^{\infty} a_k z^k\right)' = \sum_{k=1}^{\infty} k a_k z^{k-1}.$$

Důkaz lze provést analogicky reálnému případu.

Cvičení 9.2.5. Proveďte důkaz Věty 9.2.4 podrobně.

9.3 Vztah mezi mocninnými řadami a Taylorovými rozvoji

V kapitole o hlubších vlastnostech spojitých a diferencovatelných funkcí jsme si u několika elementárních funkcí zkonstruovali nekonečné Taylorovy rozvoje, které měly tvar mocninné řady. Zde si postupně ukážeme, že každá mocninná řada je uvnitř svého konvergenčního kruhu Taylorovým rozvojem nějaké nekonečněkrát diferencovatelné funkce (v předchozím textu jsme již získali nekonečněnásobnou diferencovatelnost, nyní nás už zajímá jen shoda s Taylorovým rozvojem).

Definice 9.3.1 (Taylorova řada)**.** Nechť $f\colon \mathbb{R} \to \mathbb{R}$ je někonečněkrát diferencovatelná v $x_0 \in \mathbb{R}$. Pak řadu

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

nazveme *Taylorovou řadou* funkce f v bodě x_0.

Příklad 9.3.2. (i) V kapitole o Taylorových rozvojích jsme si ukázali, že $\log(1 + x) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k}$ na $(-1, 1]$.
(ii) Funkce $x \mapsto \mathrm{e}^{-\frac{1}{x^2}}$ dodefinovaná nulou v počátku je nekonečněkrát diferencovatelná na $\mathbb{R}$ a má v počátku všechny derivace nulové. Odpovídající Taylorova řada v počátku se s původní funkcí shoduje jen v počátku.

Věta 9.3.3 (O vztahu mocninných a Taylorových řad)**.** *Nechť* $\{a_k\} \subset \mathbb{R}$*,* $x_0 \in \mathbb{R}$ *a existuje* $\delta > 0$ *takové, že mocninná řada* $\sum_{k=0}^{\infty} a_k(x - x_0)^k$ *konverguje na* $(x_0 - \delta, x_0 + \delta)$*. Pak je Taylorovou řadou svého součtu v bodě* x_0*.*

Důkaz. Podle Věty o derivaci mocninné řady (Věta 9.1.9) víme, že funkce $f(x) := \sum_{k=0}^{\infty} a_k(x - x_0)^k$ je C^∞-funkce na $(x_0 - \delta, x_0 + \delta)$ pro jisté $\delta > 0$ a pro každé $n \in \mathbb{N}_0$ platí

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1)\dots(k-n+1)a_k(x - x_0)^{k-n} \qquad \text{na } (x_0 - \delta, x_0 + \delta).$$

Odtud $f^{(n)}(x_0) = n!a_n$, proto $a_n = \frac{f^{(n)}(x_0)}{n!}$. Naše mocninná řada má tedy stejné koeficienty jako Taylorova řada. $\square$

Předchozí věta nám umožňuje hledat Taylorovy řady i jinými metodami, než je postupné derivování zadané funkce. To si teď ukážeme na příkladech.

Příklad 9.3.4. (i) Platí

$$(\log(1+x))' = \frac{1}{1+x} = \sum_{k=0}^{\infty}(-x)^k \qquad \text{na } (-1,1).$$

Odtud s využitím $\log(1+0)=0$

$$\log(1+x) = \sum_{k=0}^{\infty}(-1)^k \frac{x^{k+1}}{k+1} + C \overset{\log(1+0)=0}{=} \sum_{n=1}^{\infty}(-1)^{n+1}\frac{x^n}{n} \qquad \text{na } (-1,1).$$

Na zkoumání vztahu součtu řady a funkce $\log(1+x)$ na konvergenční kružnici (v kapitole o Taylorových rozvojích jsme rovněž uvažovali $x=1$) nám zatím teorie mocninných řad nedává žádné nástroje.
(ii) Platí

$$(\arctan x)' = \frac{1}{1+x^2} = \sum_{k=0}^{\infty}(-x^2)^k \qquad \text{na } (-1,1).$$

Odtud s využitím $\arctan 0 = 0$

$$\arctan x = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{2k+1} + C = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{2k+1} \qquad \text{na } (-1,1).$$

(iii) Najděme ještě rozvoj funkce $\log(1+x)$ v bodě $x_0>0$. Máme

$$\begin{aligned}(\log(1+x))' &= \frac{1}{1+x} = \frac{1}{1+x_0+x-x_0} = \frac{1}{1+x_0}\frac{1}{1+\frac{x-x_0}{1+x_0}} \\ &= \frac{1}{1+x_0}\sum_{k=0}^{\infty}\Big(-\frac{x-x_0}{1+x_0}\Big)^k = \sum_{k=0}^{\infty}\frac{(-1)^k}{(1+x_0)^{k+1}}(x-x_0)^k\end{aligned}$$

a pro poloměr konvergence platí $R = 1+x_0$. Odtud

$$\log(1+x) = \sum_{k=0}^{\infty}\frac{(-1)^k}{(1+x_0)^{k+1}}\frac{1}{k+1}(x-x_0)^{k+1} + \log(1+x_0) \qquad \text{pro } |x-x_0|<1+x_0.$$

Taylorovy řady se také dají získat z již známých Taylorových řad pomocí aritmetických operací a skládání.

Věta 9.3.5 (Aritmetika Taylorových řad)**.** *Nechť posloupnosti* $\{a_k\},\{b_k\}\subset\mathbb{R}$ *a mocninné řady* $f(x) := \sum_{k=0}^{\infty} a_k x^k$ *a* $g(x) := \sum_{k=0}^{\infty} b_k x^k$ *konvergují na jistém okolí počátku. Pak existuje okolí počátku, kde platí*
(i) $(f+g)(x) = \sum_{k=0}^{\infty}(a_k+b_k)x^k$
(ii) $(fg)(x) = \sum_{k=0}^{\infty}(\sum_{m=0}^{k} a_m b_{k-m})x^k$
(iii) *jestliže* $a_0 = 0$*, pak* $g(f(x)) = \sum_{n=0}^{\infty} b_n(\sum_{k=0}^{\infty} a_k x^k)^n = \sum_{k=0}^{\infty} d_k x^k$*, kde*

$$d_k = \sum_{n=0}^{\infty} b_n \sum_{\substack{k_1,k_2,\dots,k_n\in\mathbb{N}_0\\ k_1+k_2+\dots+k_n=k}} a_{k_1}a_{k_2}\dots a_{k_n}.$$

Důkaz. První část plyne z aritmetiky řad. Druhá část plyne z Cauchyovy věty o součinu řad, neboť každá z našich řad je absolutně konvergentní uvnitř svého kruhu konvergence. Dokažme část (iii). Funkce f má nulovou hodnotu v počátku a je tam spojitá. Proto pro dostatečně malá $|x|$ platí (používáme vícenásobně Cauchyovu větu o součinu řad)

$$\begin{aligned}
g(f(x)) &= \sum_{n=0}^{\infty} b_n \Big(\sum_{k=0}^{\infty} a_k x^k\Big)^n \\
&= \sum_{n=0}^{\infty} b_n \Big(\sum_{k_1=0}^{\infty}\sum_{k_2=0}^{\infty}\cdots\sum_{k_n=0}^{\infty} a_{k_1}a_{k_2}\dots a_{k_n} x^{k_1+k_2+\dots+k_n}\Big) \\
&= \sum_{n=0}^{\infty} b_n \sum_{m=0}^{\infty} \sum_{\substack{k_1,k_2,\dots,k_n\in\mathbb{N}_0\\ k_1+k_2+\dots+k_n=m}} a_{k_1}a_{k_2}\dots a_{k_n} x^m \\
&= \sum_{m=0}^{\infty} \Big(\sum_{n=0}^{\infty} b_n \sum_{\substack{k_1,k_2,\dots,k_n\in\mathbb{N}_0\\ k_1+k_2+\dots+k_n=m}} a_{k_1}a_{k_2}\dots a_{k_n}\Big) x^m.
\end{aligned}$$

□

Výše jsme si dokázali, že mocninně řady na svém kruhu konvergence definují C^∞-funkce. Výše jsme si také uvedli C^∞-funkce, které se mocninnou řadou vyjádřit nedají. To nás vede k definici následující podmnožiny množiny všech C^∞-funkcí.

Definice 9.3.6 (Reálně analytické funkce)**.** Nechť $I \subset \mathbb{R}$ je interval a $f\colon I \to \mathbb{R}$. Řekneme, že f je *reálně analytická* na I, jestliže se dá na okolí každého bodu I vyjádřit Taylorovou řadou se středem v tomto bodě.

Poznámka 9.3.7. Podle předchozí věty jsou součty, součiny reálně analytických funkcí opět reálně analytické.

Na závěr si ještě uveďme důležitou větu o konvergenci na konvergenční kružnici, kterou si dokážeme později (v kapitole o stejnoměrné konvergenci).

Věta 9.3.8 (Abelova věta)**.** *Nechť* $\{a_k\} \subset \mathbb{C}$ *a příslušná mocninná řada* $f(z) := \sum_{k=0}^{\infty} a_k z^k$ *má poloměr konvergence* $R \in (0,\infty)$*. Je-li* $\varphi \in [0,2\pi)$ *takové, že pro* $z = Re^{i\varphi}$ *konverguje řada* $\sum_{k=0}^{\infty} a_k z^k$*, pak* $t \mapsto f(te^{i\varphi})$ *je spojitá na* $[0,R]$.

Poznámka 9.3.9. Abelova věta se používá tak, že (je-li to možné) sečteme mocninnou řadu uvnitř konvergenčního kruhu a s výsledkem dokonvergujeme do bodu na konvergenční kružnici, v němž mocninná řada konverguje.

Příklad 9.3.10. (i) Sečtěme řadu $S := \sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}$. Podle Dirichletova kritéria se jedná o konvergentní řadu. Zároveň $S = f(1)$, kde

$$f(x) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k} \quad \text{na } (-1,1] \qquad \text{a} \qquad f(x) = \log(1+x) \quad \text{na } (-1,1)$$

(k aplikaci Abelovy věty v našem případě stačí umět mocninnou řadu sečíst jen pro reálné argumenty v nějakém levém prstencovém okolí bodu 1). Podle Abelovy věty dostáváme

$$f(1) = \lim_{x \to 1_-} f(x) = \lim_{x \to 1_-} \log(1 + x) = \log 2.$$

(ii) Sečtěme řadu $S := \sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1}$. Podle Dirichletova kritéria se jedná o konvergentní řadu. Zároveň $S = f(1)$, kde

$$f(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \quad \text{na } (-1, 1] \qquad \text{a} \qquad f(x) = \arctan x \quad \text{na } (-1, 1).$$

Podle Abelovy věty dostáváme

$$f(1) = \lim_{x \to 1_-} f(x) = \lim_{x \to 1_-} \arctan x = \arctan 1 = \frac{\pi}{4}.$$

9.4 Řešení diferenciálních rovnic pomocí řad

Mocninné řady se dají někdy použít při řešení diferenciálních rovnic. Používáme metodu neurčitých koeficientů.

Příklad 9.4.1. Řešme počáteční úlohu $y' = y$, $y(0) = 1$. Hledejme řešení ve tvaru $y(x) = \sum_{k=0}^{\infty} a_k x^k$. Dosazením dostáváme (pokud lze hledanou funkci vyjádřit pomocí mocninné řady s jistým kruhem konvergence, lze ji tam i derivovat člen po členu)

$$\sum_{k=1}^{\infty} k a_k x^{k-1} = \sum_{k=0}^{\infty} a_k x^k.$$

Zároveň máme

$$1 = y(0) = \sum_{k=0}^{\infty} a_k 0^k = a_0,$$

a proto porovnáváním koeficientů u jednotlivých mocnin (podle Věty o vztahu mocninných a Taylorových řad, tedy Věty 9.3.3, musí být stejné) postupně dostáváme $a_k = \frac{1}{k!}$ pro všechna $k \in \mathbb{N}_0$. Odtud

$$y(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} \qquad \text{na } \mathbb{R},$$

neboť poloměr konvergence u naší řady splňuje $R = \infty$. Tedy jak již víme, $y(x) = \mathrm{e}^x$.

Příklad 9.4.2 (Besselova rovnice)**.** Nechť $n \in \mathbb{N}_0$. Řešme úlohu $x^2 y'' + x y' + (x^2 - n^2) y = 0$. Hledejme řešení ve tvaru $y(x) = \sum_{k=0}^{\infty} a_k x^k$. Dosazením dostáváme

$$a_0(-n^2) + a_1(1 - n^2)x + \sum_{k=2}^{\infty} \big(a_k k(k-1) + a_k k - n^2 a_k + a_{k-2}\big) x^k = 0.$$

Odtud porovnáním koeficientů u jednotlivých mocnin

$$n^2 a_0 = 0, \qquad (1-n^2)a_1 = 0 \qquad \text{a} \qquad (k^2-n^2)a_k + a_{k-2} = 0 \quad \text{pro všechna } k \geq 2.$$

V případě $n = 0$, volíme jako počáteční podmínku $a_0 = y(0)$. Dále vyjde $a_{2m-1} = 0$ a $a_{2m} = -\frac{a_{2m-2}}{(2m)^2}$, kdykoliv $m \in \mathbb{N}$. Odtud

$$y(x) = a_0 \sum_{m=0}^{\infty} \frac{(-1)^m}{(m!)^2} \left(\frac{x}{2}\right)^{2m}.$$

Pokud $n = 1$, dostáváme $a_0 = 0$ a a_1 volíme jako počáteční podmínku. Dále vyjde $a_{2m} = 0$ a $a_{2m+1} = -\frac{a_{2m-1}}{(2m+1)^2-1}$ pro všechna $m \in \mathbb{N}$. Proto z rovnosti

$$(2m+1)^2 - 1 = 4m^2 + 4m = 4m(m+1)$$

dostáváme

$$y(x) = 2a_1 \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(m+1)!} \left(\frac{x}{2}\right)^{1+2m}.$$

Pro $n \geq 2$ vyjde $a_0 = a_1 = \cdots = a_{n-1} = 0$ a a_n volíme jako počáteční podmínku. Dále vyjde $a_{n+2m-1} = 0$ a $a_{n+2m} = -\frac{a_{n+2m-2}}{(2m+n)^2-n^2}$ Proto z rovnosti

$$(2m+n)^2 - n^2 = 4m^2 + 4mn = 4m(m+n)$$

dostáváme

$$y(x) = n!2^n a_n \sum_{m=0}^{\infty} \frac{(-1)^m}{m!(m+n)!} \left(\frac{x}{2}\right)^{n+2m}.$$

Pokud volíme pro $n \in \mathbb{N}_0$ hodnotu $a_n = \frac{1}{n!2^n}$, dostáváme kanonický tvar Besselových funkcí 1. druhu.

9.5 Zavedení funkcí sin, cos a exp

V kapitole o elementárních funkcích jsme zůstali dlužni důkazy následujících dvou vět.

Věta 9.5.1 (O funkcích sin a cos)**.** *Existují právě dvě funkce* $\sin, \cos \colon \mathbb{R} \to \mathbb{R}$ *a jediné iracionální číslo* π *tak, že platí*
(i) $\sin(x+y) = \sin x \cos y + \cos x \sin y \qquad \forall x, y \in \mathbb{R}$
(ii) $\cos(x+y) = \cos x \cos y - \sin x \sin y \qquad \forall x, y \in \mathbb{R}$
(iii) $\sin(-x) = -\sin x$ *a* $\cos(-x) = \cos x \qquad \forall x \in \mathbb{R}$
(iv) $\sin$ *je rostoucí na* $[0, \frac{\pi}{2}]$
(v) $\sin 0 = 0$ *a* $\sin \frac{\pi}{2} = 1$
(vi) $\sin'(0) = 1$.

Věta 9.5.2 (O exponenciále)**.** *Existuje právě jedna funkce* $\exp\colon \mathbb{C} \to \mathbb{C}$ *taková, že*
(i) $\exp(z_1 + z_2) = \exp z_1 \exp z_2 \qquad \forall z_1, z_2 \in \mathbb{C}$
(ii) $\exp(x + \mathrm{i}y) = \exp x(\cos y + \mathrm{i}\sin y) \qquad \forall x, y \in \mathbb{R}$
(iii) $\exp 0 = 1$
(iv) *restrikce funkce* $\exp$ *na* $\mathbb{R}$ *(reálná funkce) je rostoucí a jejím oborem hodnot je* $(0, \infty)$
(v) *restrikce funkce* $\exp$ *na* $\mathbb{R}$ *splňuje* $\exp' x = \exp x \qquad \forall x \in \mathbb{R}$.

Společný důkaz pro obě věty. Krok 1: definice funkcí
Definujme funkce z $\mathbb{C}$ do $\mathbb{C}$ předpisy

$$\exp z = \sum_{k=0}^{\infty} \frac{z^k}{k!}, \qquad \sin z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k+1}}{(2k+1)!} \qquad \text{a} \qquad \cos z = \sum_{k=0}^{\infty} (-1)^k \frac{z^{2k}}{(2k)!}.$$

Všechny tři řady mají poloměr konvergence roven nekonečnu, proto jsou všechny tři funkce definované na celém $\mathbb{C}$.
Krok 2: vztahy mezi funkcemi
Z definice funkce exp máme pro každé $z \in \mathbb{C}$

$$\exp(\mathrm{i}z) = \sum_{k=0}^{\infty} \frac{(\mathrm{i}z)^k}{k!} = \sum_{m=0}^{\infty} (-1)^m \frac{z^{2m}}{(2m)!} + \mathrm{i} \sum_{m=0}^{\infty} (-1)^m \frac{z^{2m+1}}{(2m+1)!} = \cos z + \mathrm{i}\sin z$$

a

$$\exp(-\mathrm{i}z) = \sum_{k=0}^{\infty} \frac{(-\mathrm{i}z)^k}{k!} = \sum_{m=0}^{\infty} (-1)^m \frac{z^{2m}}{(2m)!} - \mathrm{i} \sum_{m=0}^{\infty} (-1)^m \frac{z^{2m+1}}{(2m+1)!} = \cos z - \mathrm{i}\sin z.$$

Odtud

$$\cos z = \frac{\exp(\mathrm{i}z) + \exp(-\mathrm{i}z)}{2} \qquad \text{a} \qquad \sin z = \frac{\exp(\mathrm{i}z) - \exp(-\mathrm{i}z)}{2\mathrm{i}}.$$

Krok 3: součtové vzorce
Podle Cauchyovy věty o součinu řad (rozmyslete si, že platí i v komplexním případě) máme pro libovolná $z_1, z_2 \in \mathbb{C}$

$$\begin{aligned}\exp(z_1)\exp(z_2) &= \sum_{k=0}^{\infty} \frac{z_1^k}{k!} \sum_{j=0}^{\infty} \frac{z_2^j}{j!} = \sum_{n=0}^{\infty} \sum_{j=0}^{n} \frac{z_1^j z_2^{n-j}}{j!(n-j)!} = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{j=0}^{n} \binom{n}{j} z_1^j z_2^{n-j} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!}(z_1 + z_2)^n = \exp(z_1 + z_2).\end{aligned}$$

Odtud pro $x, y \in \mathbb{R}$

$$\exp(x + \mathrm{i}y) = \exp x \exp(\mathrm{i}y) = \exp x(\cos y + \mathrm{i}\sin y).$$

Dále

$$\begin{aligned}&\cos(x + y) + \mathrm{i}\sin(x + y) = \exp(\mathrm{i}(x + y)) = \exp(\mathrm{i}x)\exp(\mathrm{i}y) \\ &= (\cos x + \mathrm{i}\sin x)(\cos y + \mathrm{i}\sin y).\end{aligned}$$

Roznásobením pravé strany poslední rovnosti a porovnáním reálné a imaginární složky dostáváme součtové vzorce pro sinus a kosinus.
Krok 4: zbylé vlastnosti sinu a kosinu
Přímo z předpisu dostáváme sudost kosinu, lichost sinu, $\sin 0 = 0$, $\cos 0 = 1$. Derivací řady pro sinus člen po členu dostáváme $\sin' = \cos$ (speciálně $\sin' 0 = 1$). Dále

$$\begin{aligned}
\cos 2 &= \sum_{k=0}^{\infty} (-1)^k \frac{2^{2k}}{(2k)!} = 1 - \frac{2^2}{2!} + \frac{2^4}{4!} - \frac{2^6}{6!} + \dots \\
&\leq 1 - \frac{2^2}{2!} + \frac{2^4}{4!}\left(1 + \frac{2^2}{5 \cdot 6} + \frac{2^4}{5 \cdot 6 \cdot 7 \cdot 8} + \dots\right) \\
&\leq 1 - \frac{2^2}{2!} + \frac{2^4}{4!}\left(1 + \left(\frac{2}{5}\right)^2 + \left(\frac{2}{5}\right)^4 + \dots\right) = 1 - 2 + \frac{16}{24}\frac{1}{1 - \frac{4}{25}} \\
&= -1 + \frac{16}{24}\frac{25}{21} < 0.
\end{aligned}$$

Podle Darbouxovy věty proto interval obsahuje alespoň jeden nulový bod funkce kosinus. Nejmenší z nich (minimum existuje díky spojitosti) označme $\frac{\pi}{2}$. Hned také dostáváme, že sinus je rostoucí na $[0, \frac{\pi}{2}]$. Konečně, volba $x = -y$ v součtovém vzorci pro kosinus dává

$$1 = \cos 0 = \cos^2 x + \sin^2 x.$$

Odtud a z vlastnosti $\cos \frac{\pi}{2} = 0$ dostáváme $\sin \frac{\pi}{2} = 1$.
Krok 5: jednoznačnost sinu a kosinu (na $\mathbb{R}$)
Z šesti vlastností sinu a kosinu uvedených ve větě jsme v kapitole o elementárních funkcích odvodili další vlastnosti, které nám později daly Taylorovy rozvoje

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \qquad \text{a} \qquad \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}.$$

Tyto řady mají jednoznačně určené součty.
Krok 6: zbylé vlastnosti exponenciály
Vlastnost $\exp 0 = 1$ se získá dosazením. Zjevně také pro $x \in \mathbb{R}$ je $\exp x \in \mathbb{R}$. Vztah pro derivaci dostaneme derivováním mocninné řady člen po členu. Dále

$$\exp x = \exp \tfrac{x}{2} \exp \tfrac{x}{2} \geq 0,$$

proto je exponenciála nezáporná a proto také neklesající. Ze zápisu pomocí řady zřejmě plyne $\lim_{x \to \infty} \exp x = \infty$, a protože

$$\exp 0 = \exp(x - x) = \exp x \exp(-x),$$

exponenciála se nule nerovná nikdy a máme $\lim_{x \to -\infty} \exp x = 0$. Pomocí Darbouxovy věty již snadno získáme dokazovaný obor hodnot.
Krok 7: jednoznačnost exponenciály

Jednoznačnost na $\mathbb{C}$ plyne z vlastnosti (ii). Jednoznačnost na $\mathbb{R}$ plyne díky stejnému argumentu jako pro funkce sinus a kosinus.
Krok 8: iracionalita čísla π
Pro spor předpokládejme, že $\pi = \frac{p}{q}$, kde $p, q \in \mathbb{N}$ (už víme, že $\pi > 0$). Pro každé $n \in \mathbb{N}$ definujme funkci

$$f_n(x) = \frac{q^n}{n!} x^n (\pi - x)^n.$$

Máme

$$\begin{aligned}
0 < I_n &:= (\mathcal{R}) \int_0^\pi f_n(x) \sin x \, \mathrm{d}x \le (\mathcal{R}) \int_0^\pi f_n(x) \, \mathrm{d}x = \frac{q^n}{n!} (\mathcal{R}) \int_0^\pi x^n (\pi - x)^n \, \mathrm{d}x \\
&= \frac{q^n}{n!} (\mathcal{R}) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (t + \tfrac{\pi}{2})^n (\tfrac{\pi}{2} - t)^n \, \mathrm{d}t = \frac{q^n}{n!} (\mathcal{R}) \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\tfrac{\pi^2}{4} - t^2)^n \, \mathrm{d}t \\
&= 2 \frac{q^n}{n!} (\mathcal{R}) \int_0^{\frac{\pi}{2}} (\tfrac{\pi^2}{4} - t^2)^n \, \mathrm{d}t \le \frac{q^n}{n!} \left(\frac{\pi}{2} \right)^{2n} \overset{n \to \infty}{\to} 0.
\end{aligned}$$

Speciálně existuje $n \in \mathbb{N}$ tak, že $0 < I_n < 1$.

Na druhou stranu, f_n je polynom stupně $2n$. Integrál I_n lze počítat $(2n + 1)$-násobnou aplikací per partes a s využitím $\sin 0 = \sin \pi = 0$, $1 = \cos 0 = -\cos \pi$ a $f_n(x) = f_n(\pi - x)$ (odtud $f_n^{(k)}(\pi) = (-1)^k f_n^{(k)}(0)$) dostáváme

$$\begin{aligned}
I_n &= [-f_n(x) \cos x]_0^\pi + (\mathcal{R}) \int_0^\pi f_n' \cos x \, \mathrm{d}x \\
&= [-f_n(x) \cos(x)]_0^\pi + [f_n'(x) \sin x]_0^\pi - (\mathcal{R}) \int_0^\pi f_n'' \sin x \, \mathrm{d}x \\
&= \cdots = [-f_n(x) \cos(x)]_0^\pi - [-f_n''(x) \cos(x)]_0^\pi + \cdots + (-1)^n [-f_n^{(2n)}(x) \cos(x)]_0^\pi \\
&= 2 f_n(0) - 2 f_n''(0) + 2 f_n^{IV}(0) + \cdots + 2(-1)^n f_n^{(2n)}(0).
\end{aligned}$$

Spor získáme tak, že ukážeme, že všechna čísla na posledním řádku jsou celá (odtud $I_n \in \mathbb{Z}$ a nemůže platit výše dokázané $0 < I_n < 1$). Podle Leibnizova pravidla máme

$$f^{(k)}(x) = \frac{q^n}{n!} \sum_{j=0}^k \binom{k}{j} (x^n)^{(j)} \big((\pi - x)^n \big)^{(k-j)}.$$

Protože navíc

$$(x^n)^{(j)} \equiv 0 \quad \text{pro } j > n, \qquad (x^n)^{(n)} \equiv n! \qquad \text{a} \qquad (x^n)^{(j)} = C x^{n-j} \quad \text{pro } j < n$$

(v posledním případě po dosazení nuly dostaneme nulu), jediný nenulový člen v $f^{(k)}(0)$ je tedy (vyskytuje se jen pro $n \le k \le 2n$)

$$\frac{q^n}{n!} \binom{n}{n} n! \pi^n = q^n \left(\frac{p}{q} \right)^n \in \mathbb{N} \qquad \text{pro } k = n,$$

respektive pro $n < k \leq 2n$

$$\begin{aligned}&\frac{q^n}{n!}\binom{k}{n} n! n(n-1)\ldots(n-(k-n)+1)\pi^{n-(k-n)} \\ &\qquad\qquad = q^n\binom{k}{n} n(n-1)\ldots(2n-k+1)\left(\frac{p}{q}\right)^{2n-k} \in \mathbb{N}\end{aligned}$$

(mocnina čísla q je kladná, neboť $n < k \leq 2n$ implikuje $2n - k < n$). Tím jsme získali požadovaný spor a důkaz je dokončen. $\square$

Kapitola 10

Obyčejné diferenciální rovnice

V kapitole o primitivních funkcích jsme udělali krátkou exkurzi do problematiky diferenciálních rovnic. V této kapitole značně rozšíříme počet typů diferenciálních rovnic, které umíme řešit a navíc přidáme i důkazy, které jsme zůstali dlužni.

Poznamenejme, že v této kapitole budeme teorii budovat od začátku a nebudeme používat výsledky o diferenciálních rovnicích z kapitoly o primitivních funkcích, čtenář si je tedy nemusí zopakovat. Na druhou stranu, pokud čtenář vyloženě nespěchá, doporučujeme mu, aby se nejprve seznámil s částí kapitoly o metrických prostorech, kde jsou zadefinovány a studovány některé důležité pojmy z teorie funkcí více proměnných, které zde budeme používat (zejména limitu a spojitost).

Pro ostatní čtenáře alespoň tyto základní pojmy stručně představíme.

10.1 Limita a spojitost funkcí více proměnných

Zde si uvedeme jen nejdůležitější definice a výsledky (bez důkazů, podrobnosti jsou v kapitole o metrických prostorech). Na $\mathbb{R}^N$ budeme používat eukleidovskou vzdálenost

$$|x-y| := \sqrt{(x_1-y_1)^2+\cdots+(x_N-y_N)^2}.$$

Pro bod $x \in \mathbb{R}^N$ a $\varepsilon > 0$ definujeme ε-ové *okolí* bodu x jako

$$\mathcal{U}_\varepsilon(x) := \{y \in \mathbb{R}^N : |y-x| < \varepsilon\}.$$

Prstencové ε-ové *okolí* zavádíme předpisem $\mathcal{P}_\varepsilon(x) = \mathcal{U}_\varepsilon(x) \setminus \{x\}$. Je-li $A \subset \mathbb{R}^N$, o bodu $x_0 \in \mathbb{R}^N$ říkáme, že je *hromadným bodem* množiny A, jestliže každé jeho prstencové okolí má neprázdný průnik s A.

Definice 10.1.1 (Otevřená množina)**.** Řekneme, že $\Omega \subset \mathbb{R}^N$ je *otevřená* množina, jestliže ke každému jejímu bodu existuje okolí, které je celé obsažené v Ω.

Definice 10.1.2 (Limita funkce)**.** Nechť $f\colon \mathbb{R}^N \to \mathbb{R}$, $x_0 \in \mathbb{R}^N$ je hromadným bodem D_f a $y_0 \in \mathbb{R}$. Řekneme, že zobrazení f má v bodě x_0 *limitu* y_0, jestliže

$$\forall \varepsilon > 0 \,\exists \delta > 0 \qquad x \in \mathcal{P}_\delta(x_0) \cap D_f \implies f(x) \in \mathcal{U}_\varepsilon(y_0).$$

V takovém případě píšeme $\lim_{x\to x_0} f(x) = y_0$ nebo $f(x) \to y_0$ pro $x \to x_0$.

Definice 10.1.3 (Spojitost)**.** Nechť $f\colon \mathbb{R}^N \to \mathbb{R}$ a $x_0 \in D_f$. Řekneme, že zobrazení f je v bodě x_0 *spojité*, jestliže

$$\forall \varepsilon > 0 \,\exists \delta > 0 \qquad x \in \mathcal{U}_\delta(x_0) \cap D_f \implies f(x) \in \mathcal{U}_\varepsilon(f(x_0)).$$

Cvičení 10.1.4. Rozmyslete si, že zobrazení $f\colon (x, y) \mapsto x$ je spojité všude na $\mathbb{R}^2$.

Věta 10.1.5 (Aritmetika limit)**.** *Nechť $f, g\colon \mathbb{R}^N \to \mathbb{R}$ a $x_0 \in \mathbb{R}^N$ je hromadný bod $D_f \cap D_g$. Nechť $\lim_{x\to x_0} f(x) = A \in \mathbb{R}$ a $\lim_{x\to x_0} g(x) = B \in \mathbb{R}$. Pak*
(i) $\lim_{x\to x_0} \big(f(x) + g(x)\big) = A + B$
(ii) $\lim_{x\to x_0} f(x)g(x) = AB$
(iii) *pokud $B \neq 0$, platí* $\lim_{x\to x_0} \frac{f(x)}{g(x)} = \frac{A}{B}$.

Poznámka 10.1.6. Automaticky také platí aritmetika spojitosti.

Poznámka 10.1.7. Pojmy limita a spojitost se dají rozšířit rovněž na $f\colon \mathbb{R}^N \to \mathbb{R}^m$. Spojitost zadefinovaná pomocí okolí je ekvivalentní tomu, že je spojitá každá složka f. Podobně pro limitu.

Věta 10.1.8 (O spojitosti složeného zobrazení)**.** *Nechť $f\colon \mathbb{R}^N \to \mathbb{R}^m$, $g\colon \mathbb{R}^m \to \mathbb{R}^k$ a $x_0 \in \mathbb{R}^N$. Je-li f spojité v $x_0 \in \mathbb{R}^N$ a g spojité v $f(x_0)$, pak $g \circ f$ je spojité v x_0.*

Věta 10.1.9 (O limitě složeného zobrazení)**.** *Nechť $f\colon \mathbb{R}^N \to \mathbb{R}^m$, $g\colon \mathbb{R}^m \to \mathbb{R}^k$ a $x_0 \in \mathbb{R}^N$ je hromadným bodem $D_{g\circ f}$. Nechť $\lim_{x\to x_0} f(x) = y_0 \in \mathbb{R}^m$, $\lim_{y\to y_0} g(y) = z_0 \in \mathbb{R}^k$ a je splněna alespoň jedna z podmínek:*
(i) *existuje prstencové okolí bodu x_0, kde vnitřní zobrazení f nenabývá své limitní hodnoty y_0*
(ii) *vnější zobrazení g je spojité v bodě y_0.*
Pak $\lim_{x\to x_0} (g \circ f)(x) = z_0$.

Připomeňme ještě, že pro $f\colon \mathbb{R}^N \to \mathbb{R}$, $x_0 \in \mathbb{R}^N$ a $i \in \{1, \ldots, N\}$ definujeme i-tou *parciální derivaci* funkce f v bodě x předpisem

$$\frac{\partial f(x)}{\partial x_i} = \lim_{h\to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + h, x_{i+1}, \ldots, x_N) - f(x)}{h},$$

pokud existuje vlastní limita napravo. V případě vektorové funkce $f\colon \mathbb{R}^N \to \mathbb{R}^m$ se analogicky zavádí i-tá parciální derivaci j-té složky funkce f a značí se $\frac{\partial f_j}{\partial x_i}$.

10.2 Základní pojmy

Definice 10.2.1 (Obyčejná diferenciální rovnice)**.** Nechť $n \in \mathbb{N}$ a $f\colon \mathbb{R}^{n+2} \to \mathbb{R}$. Pak

$$f(x, y, y', \dots, y^{(n)}) = 0 \tag{10.2.1}$$

se nazývá *skalární obyčejná diferenciální rovnice* n-tého řádu.

Příklad 10.2.2. (i) Do našeho případu spadá třeba rovnice matematického kyvadla $y'' + y = 0$.
(ii) Rovnice vedení tepla $\frac{\partial u(x,t)}{\partial t} = \frac{\partial^2 u(x,t)}{\partial x^2}$ je *parciální diferenciální rovnice*. Tímto typem rovnic se zatím nebudeme zabývat. Poznamenejme alespoň, že teorie parciálních diferenciálních rovnic je komplikovanější než u rovnic obyčejných.
(iii) Model *dravec–kořist* s parametry $a, b, c, d > 0$

$$\begin{aligned} x'(t) &= ax(t) - bx(t)y(t) \\ y'(t) &= -cy(t) + dx(t)y(t) \end{aligned}$$

je systém obyčejných diferenciálních rovnic 1. řádu. Systémy se někdy dají převádět na jednu rovnici vyššího řádu. Tím se budeme také zabývat.

Definice 10.2.3 (Řešení obyčejné diferenciální rovnice)**.** Funkci $y\colon (a, b) \to \mathbb{R}$ nazveme *řešením obyčejné diferenciální rovnice* (10.2.1), jestliže

- y má na (a, b) vlastní derivace n-tého řádu
- pro všechna $x \in (a, b)$ platí (10.2.1).

Často bude možné rovnici (10.2.1) přepsat do *tvaru rozřešeného vzhledem k nejvyšší derivaci*

$$y^{(n)} = g(x, y, y', \dots, y^{(n-1)}), \tag{10.2.2}$$

kde $g\colon \mathbb{R}^{n+1} \to \mathbb{R}$.

Budeme zde také uvažovat systémy obyčejných diferenciálních rovnic, ale pouze prvního řádu a rozřešené vzhledem k první derivaci. Pro vektorovou funkci $y = (y_1, y_2, \dots, y_m)\colon \mathbb{R} \to \mathbb{R}^m$ budeme používat značení $y' = (y_1', \dots, y_m')$.

Definice 10.2.4 (Systém obyčejných diferenciálních rovnic prvního řádu)**.** Nechť $F\colon \mathbb{R}^{m+1} \to \mathbb{R}^m$. Pak

$$y' = F(x, y)$$

je *systém obyčejných diferenciálních rovnic prvního řádu* (rozřešený vzhledem k první derivaci) pro m neznámých funkcí $y_1, \dots, y_m$. Jeho řešením na (a, b) nazveme $y = (y_1, \dots, y_m)\colon (a, b) \to \mathbb{R}^m$ splňující

- $y_1, \dots, y_m$ mají na (a, b) vlastní derivace (prvního řádu)
- pro všechna $x \in (a, b)$ platí $y'(x) = F(x, y(x))$.

Uveďme si nyní několik příkladů. Budou ilustrovat, že obyčejné diferenciální rovnice mohou sloužit k popisu různých reálných problémů.

Příklad 10.2.5. (i) Rovnice popisující radioaktivní rozpad
Nechť $N = N(t)$ je počet radioaktivních částic v čase t. Nechť jejich úbytek (tedy počet radioaktivních rozpadů) je přímo úměrný jejich počtu, tedy

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = -\lambda N(t), \quad \lambda > 0.$$

Zřejmě tuto rovnici řeší $N(t) = C\mathrm{e}^{-\lambda t}$, kde C je libovolná (vzhledem k fyzikální úloze kladná) konstanta. Známe-li počet částic v čase $t = t_0$, tedy $N(t_0) = N_0$, pak

$$N(t) = N_0 \mathrm{e}^{-\lambda(t-t_0)}$$

řeší naši úlohu. Později si ukážeme, že toto řešení je jediné a řeší rovnici pro libovolné $t \in \mathbb{R}$, i když v dané úloze nás typicky zajímají pouze hodnoty $t \geq t_0$.

(ii) Rovnice popisující růst počtu obyvatel
(a) *neomezený růst*
V tomto případě je přírůstek počtu obyvatel přímo úměrný počtu obyvatel v daném okamžiku, úloha je tedy dosti podobná předchozí úloze o radioaktivním rozpadu. Potom

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = \alpha N(t), \quad \alpha > 0.$$

Analogicky jako výše, pokud známe $N_0 = N(t_0)$, pak

$$N(t) = N_0 \mathrm{e}^{\alpha(t-t_0)}.$$

Všimněme si, že $N(t) \to \infty$ pro $t \to \infty$.
(b) *omezený růst*
Předchozí úloha měla tu nepříjemnou vlastnost, že počet obyvatel mohl růst nade všechny meze, což není příliš reálné. Proto se častěji předpokládá, že počet obyvatel nemůže překročit jistou předem danou mez $P_{\max}$. Odpovídající rovnice je potom například

$$\frac{\mathrm{d}N(t)}{\mathrm{d}t} = \alpha N(t)\big(P_{\max} - N(t)\big), \quad \alpha > 0.$$

Je-li opět $N_0 = N(t_0) \in (0, P_{\max}]$ (jinak úloha nemá rozumný smysl), lze ukázat, že

$$N(t) = \frac{P_{\max} N_0}{N_0 + (P_{\max} - N_0)\mathrm{e}^{-P_{\max}\alpha(t-t_0)}}.$$

Dostali jsme takzvanou logistickou křivku. Zřejmě $N(t) \to P_{\max}$ pro $t \to \infty$. V obou případech je řešení určeno jednoznačně.

(iii) Pohyb hmotného bodu
(a) *jednodimenzionální případ*
Jestliže se hmotný bod může pohybovat pouze ve směru osy x a označíme-li $x(t)$ jeho polohu v čase t, potom $v(t) = \frac{\mathrm{d}x(t)}{\mathrm{d}t}$ je jeho okamžitá rychlost a $a(t) = \frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2}$ jeho okamžité zrychlení. Působí-li na hmotný bod síla $f(t,x,v) = f\Big(t, x(t), \frac{\mathrm{d}x(t)}{\mathrm{d}t}\Big)$, dostáváme z Newtonova pohybového zákona rovnici

$$m\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2} = f\Big(t, x(t), \frac{\mathrm{d}x(t)}{\mathrm{d}t}\Big),$$

kde m je hmotnost částice. Aby mohlo být řešení dáno jednoznačně, je třeba předepsat dvě hodnoty (rovnice je druhého řádu), tedy

$$x(t_0) = x_0, \qquad \frac{\mathrm{d}x(t_0)}{\mathrm{d}t} = v_0.$$

Je možno ukázat, že za jistých (fyzikálně rozumných) předpokladů na funkci f je řešení skutečně dáno jednoznačně. Ovšem nalézt řešení není jednoduché a záleží na tvaru této funkce. Dokonce v mnohých případech nelze čekat, že by řešení šlo vyjádřit pomocí elementárních funkcí.
(b) *třídimenzionální případ*
V tomto případě již nevystačíme s jednou obyčejnou diferenciální rovnicí (tedy skalární rovnicí) a musíme uvažovat systém tří obyčejných diferenciálních rovnic druhého řádu. Dostáváme

$$\begin{aligned} m\frac{\mathrm{d}^2x_1(t)}{\mathrm{d}t^2} &= F_1\Big(t, \vec{x}(t), \frac{\mathrm{d}\vec{x}(t)}{\mathrm{d}t^2}\Big) \\ m\frac{\mathrm{d}^2x_2(t)}{\mathrm{d}t^2} &= F_2\Big(t, \vec{x}(t), \frac{\mathrm{d}\vec{x}(t)}{\mathrm{d}t^2}\Big) \\ m\frac{\mathrm{d}^2x_3(t)}{\mathrm{d}t^2} &= F_3\Big(t, \vec{x}(t), \frac{\mathrm{d}\vec{x}(t)}{\mathrm{d}t^2}\Big). \end{aligned}$$

Případně je možno tento systém zapsat vektorově

$$m\frac{\mathrm{d}^2\vec{x}(t)}{\mathrm{d}t^2} = \vec{F}\Big(t, \vec{x}(t), \frac{\mathrm{d}\vec{x}(t)}{\mathrm{d}t^2}\Big),$$

kde $\vec{F}\colon \mathbb{R}^7 \to \mathbb{R}^3$. Opět je třeba zadat počáteční polohu a počáteční rychlost, tedy

$$\vec{x}(t_0) = \vec{x}_0, \qquad \frac{\mathrm{d}\vec{x}(t_0)}{\mathrm{d}t} = \vec{v}_0.$$

Řešení této úlohy je již velice komplikované i pro relativně jednoduchou funkci $\vec{F}$.

Ukažme si ještě, jak se převádějí systémy rovnic prvního řádu na jednu rovnici vyššího řádu a naopak. Jde to ale jen někdy, nicméně rovnici typu (10.2.2) lze na systém rovnic prvního řádu rozřešený vzhledem k první derivaci převést vždy.

Příklad 10.2.6. (i) Uvažme rovnici $y''' + 2y'' + y' = 2yx$. Označme $u = y$, $v = y'$ a $w = y''$. Dostáváme

$$\begin{aligned} u' &= v \\ v' &= w \\ w' &= -2w - v + 2xu. \end{aligned}$$

Podobně rovnici typu (10.2.2) lze na systém rovnic prvního řádu (rozřešený vzhledem k derivaci 1. řádu) převést vždy. Stačí použít analogický postup jako výše.

Obecně pak u rovnice n-tého řádu tvaru $f(x, y, y', \ldots, y^{(n)}) = 0$ definujeme $u_k = y^{(k-1)}$ pro $k = 1, \ldots, n$. Prvních $n-1$ rovnic má tvar $u_k' = u_{k+1}$ a poslední

je $f(x, u_1, u_2, \ldots, u_n, u_n') = 0$.
(ii) Uvažme systém

$$\begin{aligned}
u_1' &= u_1 + u_2 + u_3\\
u_2' &= u_1 - u_2 - u_3\\
u_3' &= u_1 + u_2 - u_3.
\end{aligned}$$

Budeme se postupně zbavovat u_2 a u_3. Nejprve si z první rovnice vyjádříme u_2 a ještě jej zderivujeme.

$$u_2 = u_1' - u_1 - u_3 \qquad \text{a} \qquad u_2' = u_1'' - u_1' - u_3'.$$

Výsledek dosadíme do zbývajících rovnic

$$\begin{aligned}
u_1'' - u_1' - u_3' &= u_1 - (u_1' - u_1 - u_3) - u_3\\
u_3' &= u_1 + (u_1' - u_1 - u_3) - u_3,
\end{aligned}$$

což po zjednodušení dává

$$\begin{aligned}
u_3' &= u_1'' - 2u_1\\
u_3' &= u_1' - 2u_3.
\end{aligned}$$

Odtud se dá vyjádřit u_3, které opět zderivujeme

$$u_3 = \frac{1}{2}(-u_1'' + u_1' + 2u_1) \qquad \text{a} \qquad u_3' = \frac{1}{2}(-u_1''' + u_1'' + 2u_1').$$

To dosadíme do poslední rovnice

$$\frac{1}{2}(-u_1''' + u_1'' + 2u_1') = u_1' - (-u_1'' + u_1' + 2u_1).$$

Teď už stačí jen výsledek zjednodušit a dostáváme

$$u_1''' + u_1'' - 2u_1' - 4u_1 = 0.$$

Postup z bodu (ii) z předchozího příkladu lze použít v případě soustavy rovnic

$$u_i' = \sum_{j=1}^{n} a_{ij} u_j + f_i(x) \qquad i = 1, \ldots, n$$

s konstantními koeficienty $a_{ij} \in \mathbb{R}$ a f_i majícími konečnou derivaci řádu $n-1$. Nefunguje vždy. Například soustava

$$\begin{aligned}
u_1' &= u_1\\
u_2' &= u_2
\end{aligned}$$

se takto přepsat nedá (v tomto jednoduchém případě je však nesmyslné se o takový přepis pokoušet).

10.3 Základní existenční věty

V následujícím textu budeme teorii budovat pro systémy rovnic prvního řádu ve tvaru

$$y' = F(x, y)$$

(kde $F\colon \mathbb{R}^{n+1} \to \mathbb{R}^n$ a hledáme $y\colon \mathbb{R} \to \mathbb{R}^n$). Vzhledem k úzkému vztahu mezi systémem a rovnicí vyššího řádu (jak jsme si naznačili výše) se získané výsledky po odpovídající transformaci dají snadno převést i na skalární rovnice vyššího řádu rozřešené vzhledem k nejvyšší derivaci (tj. ve tvaru (10.2.2)).

Definice 10.3.1 (Cauchyova úloha)**.** *Cauchyovou úlohou* pro rovnici $y' = F(x, y)$ na (a, b), kde $F\colon \mathbb{R}^{n+1} \to \mathbb{R}^n$, rozumíme hledání vektorové funkce $y\colon \mathbb{R} \to \mathbb{R}^n$ splňující $y'(x) = F(x, y(x))$ na (a, b) a $y(x_0) = y_0$, kde $x_0 \in (a, b)$ a $y_0 \in \mathbb{R}^n$ jsou zadané hodnoty.

Poznámka 10.3.2. Řešení y_1 na (a_1, b_1) a y_2 na (a_2, b_2) budeme považovat za stejná jen v případě, že $(a_1, b_1) = (a_2, b_2)$ a $y_1 = y_2$ na (a_1, b_1). Naopak, pokud $(a_1, b_1) \neq (a_2, b_2)$, řešení považujeme za různá i v případě, že $y_1 = y_2$ na $(a_1, b_1) \cap (a_2, b_2)$. Tento přístup má dobrý důvod: řešení diferenciálních rovnic totiž občas v některých bodech ztrácejí jednoznačnost (mohou se „rozvětvit“).

Definice 10.3.3 (Prodloužení řešení)**.** Nechť y_1 řeší rovnici $y' = F(x, y)$ na intervalu (a_1, b_1) a y_2 ji řeší na intervalu (a_2, b_2). Jestliže $(a_1, b_1) \subset (a_2, b_2)$ a $y_1 = y_2$ na (a_1, b_1), řekneme, že y_2 je *prodloužením řešení* y_1 (na interval (a_2, b_2)). Řešení se nazývá *maximální*, jestliže se nedá prodloužit.

Existenci a jednoznačnost dávají následující dva výsledky, na jejichž důkaz zatím nejsme vybaveni (oba důkazy jsou uvedeny v kapitole o metrických prostorech).

Věta 10.3.4 (Peanova existenční věta)**.** *Nechť $F\colon \mathbb{R}^{n+1} \to \mathbb{R}^n$ je spojitá na otevřené množině $\Omega \subset \mathbb{R}^{n+1}$ a $(x_0, y_0) \in \Omega$. Pak existuje řešení Cauchyovy úlohy pro systém rovnic $y' = F(x, y)$ s počáteční podmínkou $y(x_0) = y_0$.*

Věta 10.3.5 (Picard–Lindelöfova existenční věta)**.** *Nechť $F\colon \mathbb{R}^{n+1} \to \mathbb{R}^n$ je spojitá na otevřené množině $\Omega \subset \mathbb{R}^{n+1}$, $(x_0, y_0) \in \Omega$ a F je na Ω lokálně lipschitzovská vzhledem k poslední n-tici proměnných. Pak existuje $\delta > 0$ tak, že na intervalu $(x_0 - \delta, x_0 + \delta)$ existuje právě jedno řešení Cauchyovy úlohy pro systém rovnic $y' = F(x, y)$ s počáteční podmínkou $y(x_0) = y_0$.*

Poznámka 10.3.6. Jednoznačnost ve větě výše (a také všude dál) chápeme v následujícím smyslu. Libovolné jiné řešení procházející bodem (x_0, y_0) se shoduje s daným řešením na průniku definičních oborů.

Poznámka 10.3.7. Lokální lipschitzovskost vzhledem k poslední n-tici proměnných znamená, že pro každé $(x_0, y_0) \in \Omega$ existují $K > 0$ a $\delta > 0$ taková, že

$$|F(x, y_1) - F(x, y_2)| \leq K|y_1 - y_2| \qquad \text{kdykoliv } (x, y_1), (x, y_2) \in \mathcal{U}_\delta((x_0, y_0)).$$

Poznámka 10.3.8. Lipschitzovskost je podstatně přísnější podmínka než lokální lipschitzovskost. Uvažte třeba funkci $x \mapsto x^2$.

Poznámka 10.3.9. Připomeneme-li si konstrukci, pomocí níž se z úlohy $y^{(n)} = g(x, y, y', \ldots, y^{(n-1)})$ získá systém rovnic prvního řádu, máme

$$F(x, u) = (u_2, u_3, \ldots, u_n, g(x, u)).$$

Pro aplikaci Picard–Lindelöfovy existenční věty (Věta 10.3.5) se pak hodí následující pozorování. Funkce F je spojitá na Ω právě tehdy, když g je spojitá na Ω. Podobně pro lokální lipschitzovskost v poslední n-tici proměnných. Správná sada počátečních podmínek (kompatibilní s teorií systémů rovnic) je $y(x_0) = y_0, y'(x_0) = y_1, \ldots, y^{(n-1)}(x_0) = y_{n-1}$.

Jak bylo řečeno výše, obě existenční věty budou dokázány v příští kapitole. My si v následujícím příkladu budeme ilustrovat metodu důkazu na jednoduché úloze

$$y' = y, \quad y(0) = 1. \tag{10.3.1}$$

Příklad 10.3.10. (i) Nejprve ilustrujme na řešení úlohy (10.3.1) důkaz Peanovy existenční věty (Věta 10.3.4). Předpoklady věty jsou zřejmě splněny. Vezměme interval $[0, 1]$ a rozdělme ho na n stejných dílků délky $\frac{1}{n}$. Vezměme interval $\left[0, \frac{1}{n}\right]$ a řešme

$${}^{1}_{n}y' = 1, \quad {}^{1}_{n}y(0) = 1.$$

Volba pravé strany rovnice 1 souvisela s tím, že díky počáteční podmínce víme, že $y(0) = 1$. Zjevně

$${}^{1}_{n}y(x) = x + 1, \quad x \in \left[0, \frac{1}{n}\right].$$

Dále uvažujme interval $\left[\frac{1}{n}, \frac{2}{n}\right]$, uvědomíme si, že z předchozího kroku máme, že aproximace řešení má splňovat $y(\frac{1}{n}) = 1 + \frac{1}{n}$, a řešme

$${}^{2}_{n}y' = 1 + \frac{1}{n}, \quad {}^{2}_{n}y\left(\frac{1}{n}\right) = 1 + \frac{1}{n}.$$

Dostáváme

$${}^{2}_{n}y(x) = \left(1 + \frac{1}{n}\right)x + \left(1 - \frac{1}{n}\right)\left(1 + \frac{1}{n}\right) \quad x \in \left[\frac{1}{n}, \frac{2}{n}\right].$$

Tedy pro $1 \leq k < n$ máme na intervalu $\left[\frac{k}{n}, \frac{k+1}{n}\right]$ (připomeňme, že ${}^{k}_{n}y(\frac{k}{n}) = (1 + \frac{1}{n})^k$)

$${}^{k+1}_{n}y' = \left(1 + \frac{1}{n}\right)^k, \quad {}^{k+1}_{n}y\left(\frac{k}{n}\right) = \left(1 + \frac{1}{n}\right)^k.$$

Proto

$${}^{k+1}_{n}y(x) = \left(1 + \frac{1}{n}\right)^k x + \left(1 - \frac{k}{n}\right)\left(1 + \frac{1}{n}\right)^k \quad x \in \left[\frac{k}{n}, \frac{k+1}{n}\right].$$

Například tedy

$${}^{n}_{n}y(1) = \left(1 + \frac{1}{n}\right)^n \to \mathrm{e}^1 \quad \text{pro } n \to \infty.$$

Analogicky pak pro ${}_n y$ definováné jako ${}^k_n y$ na intervalu $[\frac{k-1}{n}, \frac{k}{n}]$, $1 \le k \le n$

$$ {}_n y(z) \sim \left(1 + \frac{1}{n}\right)^{zn} \to \mathrm{e}^z \qquad \text{pro } n \to \infty $$

pro libovolné $z \in [0,1]$. V celém důkazu Peanovy existenční věty je nejtěžší ověřit, že pro spojitou funkci F konverguje ${}_n y$ k nějaké funkci y, která je spojitá a diferencovatelná na daném intervalu a splňuje danou rovnici.

(ii) Nyní si na úloze (10.3.1) ilustrujme důkaz Picard–Lindelöfovy existenční věty (Věta 10.3.5). Zřejmě platí

$$ y(x) = \int_0^x y'(s)\,\mathrm{d}s + y(0). $$

Protože $y' = y$ a $y(0) = 1$, máme

$$ y(x) = \int_0^x y(s)\,\mathrm{d}s + 1. $$

Není těžké ověřit, že spojitá funkce y řeší úlohu výše právě tehdy, když řeší (10.3.1). Nyní položme $y_0 \equiv 1$ a definujme

$$ y_n(x) = \int_0^x y_{n-1}(s)\,\mathrm{d}s + 1. $$

To těžké na důkazu Picard–Lindelöfovy existenční věty je ověřit, že taková posloupnost má limitu, která řeší úlohu (10.3.1). V našem případě ale máme

$$ \begin{aligned} y_1(x) &= x + 1 \\ y_2(x) &= \frac{x^2}{2} + x + 1 \\ &\vdots \\ y_n(x) &= \frac{x^n}{n!} + \frac{x^{n-1}}{(n-1)!} + \cdots + x + 1. \end{aligned} $$

Tedy

$$ y_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!} \to \mathrm{e}^x $$

pro $n \to \infty$.

Tvrzení 10.3.11 (O slepování řešení)**.** *Nechť y_1 řeší úlohu $y' = F(x,y)$ na (a,b) a y_2 řeší tutéž úlohu na (b,c). Pokud navíc platí*

$$ \lim_{x \to b_+} y_1(x) = \lim_{x \to b_-} y_2(x) = z \in \mathbb{R}^n, $$

(limitu vektorové funkce počítáme zvlášť po jednotlivých složkách) a F je spojitá v bodě (b,z), pak vektorová funkce

$$ y(x) = \begin{cases} y_1(x) & \text{pro } x \in (a,b) \\ z & \text{pro } x = b \\ y_2(x) & \text{pro } x \in (b,c) \end{cases} $$

řeší rovnici $y' = F(x, y)$ na (a, c).

Důkaz. Nejprve si povšimněme, že stačí ukázat $y'(b) = F(b, z)$. Spojitost F v bodě (b, z) a předpoklad $\lim_{x \to b_+} y_1(x) = \lim_{x \to b_-} y_2(x) = z \in \mathbb{R}^n$ implikují

$$\lim_{x \to b_+} F(x, y_1(x)) = \lim_{x \to b_-} F(x, y_2(x)) = F(b, z).$$

Podle věty o limitě derivací pak dostáváme $y'(b) = F(b, z)$ a jsme hotovi. $\square$

Poznámka 10.3.12. Otázkou je, kdy můžeme řešení prodlužovat a získat řešení maximální. Nepůjdeme zde do detailů a jen si stručně představíme možné scénáře.
(i) „Narazíme“ na hranici oblasti, na které má úloha smysl. Tedy pravá strana rovnice přestává být spojitá či vůbec definovaná a nemůžeme pokračovat.
(ii) Zůstaneme někde uvnitř oblasti spojitosti, tedy naše řešení $y(x)$ je omezené a pravá strana rovnice má smysl na nějakém okolí „koncového bodu“. V takovém případě můžeme řešení „prodloužit“ do koncového bodu (v něm je řešení nejen spojité, ale i diferencovatelné) a můžeme z toho bodu nalézt pokračování řešení, ať už pomocí Peanovy či Picard–Lindelöfovy existenční věty. Podle předchozího tvrzení můžeme obě řešení napojit a řešení tedy prodloužíme.
(iii) Řešení nám „uteče“ do nekonečna, tedy $|y(x)| \to \infty$. Toto je typický scénář pro některé nelineární rovnice, pro lineární rovnice se tohle nestane. Řešení pak samozřejmě nelze prodloužit.

10.4 Metody řešení vybraných skalárních rovnic prvního řádu

10.4.1 Rovnice $y' = f(x)$

Úloha

$$\begin{aligned} y' &= f(x) \\ y(x_0) &= y_0 \end{aligned}$$

má v případě, kdy f je spojitá na (a, b) a $x_0 \in (a, b)$, jednoznačné řešení tvaru

$$y(x) = y_0 + (\mathcal{R}) \int_{x_0}^{x} f(s)\,\mathrm{d}s.$$

Existence a jednoznačnost plynou například z Picard–Lindelöfovy existenční věty (v našem případě je $F(x, y) = f(x)$). Platnost vzorečku plyne okamžitě z teorie Riemannnova integrálu (Takzvaná hlavní věta diferenciálního a integrálního počtu, tedy Věta 7.5.12). V tomto případě jsme také mohli jednoznačnost dokázat pomocí Věty o nejednoznačnosti primitivní funkce (Věta 4.1.4), existenci pomocí Věty o existenci primitivní funkce ke spojité funkci (Věta 7.5.13). Zde zatím nebylo důležité, že jsme ve skalárním případě, v případě systému rovnic tohoto typu bychom postupovali po složkách.

10.4.2 Rovnice $y' = g(y)$

V této situaci je řešením

$$y = G^{-1}(x+C), \qquad \text{kde } G(y) := \int \frac{\mathrm{d}y}{g(y)}.$$

Výsledný vzorec si lze pamatovat pomocí zjednodušené myšlenky jeho důkazu (ale pozor, tato myšlenka kupříkladu používá nedefinované znaky a neřeší existenci inverze, není tedy důkazem)

$$y' = \frac{\mathrm{d}y}{\mathrm{d}x} = g(y) \qquad \rightsquigarrow \qquad \frac{\mathrm{d}y}{g(y)} = \mathrm{d}x \qquad \rightsquigarrow \qquad \int \frac{\mathrm{d}y}{g(y)} = \int \mathrm{d}x = x + C.$$

Věta 10.4.1 (O řešení rovnice $y' = g(y)$)**.** *Nechť $g\colon \mathbb{R} \to \mathbb{R}$ je spojitá a nenulová na $(\alpha,\beta) \subset \mathbb{R}$. Nechť G je primitivní funkce k $y \mapsto \frac{1}{g(y)}$ na (α,β). Pak na intervalu $G((\alpha,\beta))$ existuje inverzní funkce G^{-1} a každé maximální řešení v $\Omega = \mathbb{R} \times (\alpha,\beta)$ má tvar*

$$y(x) = G^{-1}(x+C),$$

kde $C \in \mathbb{R}$, a je definováno na intervalu

$$I_C := \{x \in \mathbb{R}\colon \exists y \in (\alpha,\beta) \quad G(y) = x + C\} = \{x = z - C\colon z \in G((\alpha,\beta))\}.$$

Navíc každým bodem $(x_0, y_0) \subset \Omega$ prochází právě jedno maximální řešení (v Ω).

Důkaz. Krok 1: existence a diferencovatelnost G^{-1}.
Díky předpokladům na g je $\frac{1}{g}$ spojitá, nenulová a nemění znaménko. Díky tomu existuje ryze monotonní G (primitivní funkce k $\frac{1}{g}$), G' je nenulová a nemění znaménko. Podle Věty o derivaci inverzní funkce (verze pro funkci s nenulovou derivací neměnící znaménko, tj. Věta 3.3.22) je $G^{-1}\colon G((\alpha,\beta)) \to (\alpha,\beta)$ diferencovatelná. Navíc funkce $x \mapsto G^{-1}(x+C)$ zobrazuje interval I_C na (α,β).

Krok 2: $y(x) = G^{-1}(x+C)$ řeší úlohu $y' = g(y)$ na I_C.
Na I_C podle prvního kroku můžeme funkci $x \mapsto G^{-1}(x+C)$ derivovat a dostáváme

$$y'(x) = \frac{\mathrm{d}}{\mathrm{d}x} G^{-1}(x+C) = \frac{1}{G'(G^{-1}(x+C))} = g(G^{-1}(x+C)) = g(y(x)).$$

Krok 3: maximalita řešení tvaru $y(x) = G^{-1}(x+C)$.
Odvodíme spor v situaci, kdy $g > 0$ na (α,β) (tedy G^{-1} je rostoucí) a řešení umíme prodloužit doprava mimo interval I_C (v ostatních případech postupujeme analogicky). Pišme $(a,b) := I_C$. Nejprve si povšimněme, že díky monotonii funkce G^{-1} máme

$$\lim_{x \to b_-} G^{-1}(x+C) = \sup_{x \in (a,b)} G^{-1}(x+C) = \sup(\alpha,\beta) = \beta.$$

Dále, je-li možné řešení y prodloužit za bod b, y je spojité v bodě b, a proto $y(b) = \beta$. Ale $\beta \notin (\alpha,\beta)$, není tedy v definičním oboru funkce g. To je ve sporu s tím, že by mělo platit $y'(b) = g(y(b))$.

Krok 4: jednoznačnost.
Uvažme například situaci, kdy $g > 0$ na (α, β). Nechť η je řešení s definičním oborem (a, b). Pak

$$\eta'(x) = g(\eta(x)) > 0 \qquad \text{pro } x \in (a, b).$$

Funkce η proto má diferencovatelnou inverzi na $\eta((a, b))$. Označme ji μ. Pak pro všechna $y \in \eta((a, b))$ platí

$$\mu'(y) = \frac{1}{\eta'(\mu(y))} = \frac{1}{g(\eta(\mu(y)))} = \frac{1}{g(y)}.$$

Odtud $\mu(y) = G(y) - C$ na $\eta((a, b))$. Pro $x \in (a, b)$ tedy dostáváme

$$\mu(\eta(x)) = G(\eta(x)) - C \qquad \Longleftrightarrow \qquad x + C = G(\eta(x)).$$

Následně $x+C$ leží v oboru hodnot funkce G, neboli $(a, b) \subset I_C$, a $\eta(x) = G^{-1}(x+C)$.

Krok 5: existence řešení pro každou počáteční podmínku.
Nechť $(x_0, y_0) \in \Omega = \mathbb{R} \times (\alpha, \beta)$. Protože G^{-1} zobrazuje $G((\alpha, \beta))$ na (α, β), existuje $\xi \in G((\alpha, \beta))$ takové, že $y_0 = G^{-1}(\xi)$. Stačí proto položit $C := \xi - x_0$ a pak funkce

$$y(x) := G^{-1}(x + C) = G^{-1}(x + \xi - x_0)$$

splňuje $y(x_0) = G^{-1}(\xi) = y_0$. □

Příklad 10.4.2. (i) Nechť $\alpha \in \mathbb{R} \setminus \{0\}$ a $A \in \mathbb{R}$. Uvažme úlohu

$$\begin{aligned} y'(x) &= \alpha y(x) \\ y(0) &= A. \end{aligned}$$

Předchozí větu můžeme používat v situacích $\Omega = \mathbb{R} \times (-\infty, 0)$ a $\Omega = \mathbb{R} \times (0, \infty)$. Zde shodně máme

$$G(y) = \int \frac{\mathrm{d}y}{g(y)} = \int \frac{\mathrm{d}y}{\alpha y} = \frac{1}{\alpha} \log |y| = x + C.$$

Odtud, píšeme-li $\mathrm{e}^{\alpha C} = K > 0$,

$$|y| = \mathrm{e}^{\alpha(x+C)} = K\mathrm{e}^{\alpha x}.$$

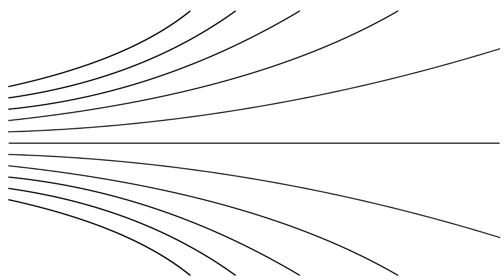
Při počáteční podmínce $y(0) = A > 0$ používáme větu na $\Omega = \mathbb{R} \times (0, \infty)$ a dostáváme jednoznačné maximální řešení tvaru

$$y = A\mathrm{e}^{\alpha x}$$

definované na celém $\mathbb{R}$. Analogicky se postupuje v případě $A < 0$. Opět dostaneme jednoznačné řešení popsané vzorečkem $y = A\mathrm{e}^{\alpha x}$ na $\mathbb{R}$. Povšimněme si ještě, že bychom podobně mohli postupovat i v případě počáteční podmínky tvaru $y(x_0) = A$, pro $A \neq 0$, jen by nám vyšla odlišná multiplikativní konstanta.

V situaci $A = 0$ předchozí větu použít nemůžeme. Snadno však nahlédneme, že funkce $y \equiv 0$ je řešením. Podobně v případě počáteční podmínky $y(x_0) = 0$. I tato řešení jsou jednoznačná. Kdyby totiž nejaké řešení zároveň splňovalo třeba $y(x_1) = 0$ a $y(x_2) > 0$, z našich výsledků pro $\Omega = \mathbb{R} \times (0, \infty)$ by plynulo, že se toto řešení musí shodovat s nějakým řešením tvaru $K\mathrm{e}^{\alpha x}$, ale pak by nemohlo platit $y(x_1) = 0$.

Poznamenejme ještě, že existenci a (lokální) jednoznačnost pro jakékoliv $A \in \mathbb{R}$ nám také zaručuje Picard–Lindelöfova existenční věta, neboť funkce $(x, y) \mapsto \alpha y$ je spojitá a lipschitzovská ve druhé složce.



Obrázek 10.1: Náčrt částí několika větví obecného řešení rovnice $y' = \alpha y$. Povšimněte si, že aditivní konstanta z integrace se projevuje jinak, než jak jsme tomu byli zvyklí u primitivních funkcí (tentokrát má roli multiplikativní konstanty).

(ii) Nechť $A \in \mathbb{R}$. Uvažme úlohu

$$\begin{aligned} y'(x) &= y^2(x) \\ y(0) &= A. \end{aligned}$$

Předchozí větu můžeme používat v situacích $\Omega = \mathbb{R} \times (-\infty, 0)$ a $\Omega = \mathbb{R} \times (0, \infty)$. Zde shodně máme

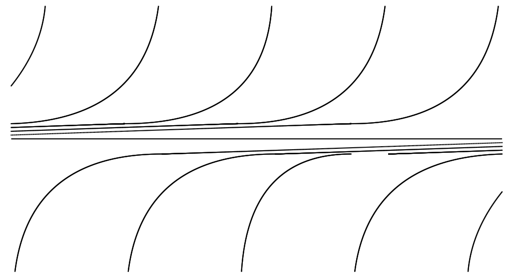
$$G(y) = \int \frac{\mathrm{d}y}{g(y)} = \int \frac{\mathrm{d}y}{y^2} = -\frac{1}{y} = x + C.$$

Z počáteční podmínky dostáváme $C = -\frac{1}{A}$, a proto máme jednoznačná maximální řešení

$$\begin{aligned} A > 0 \quad &\Longrightarrow \quad y = \frac{-1}{x - \frac{1}{A}} \quad \text{na } \left(-\infty, \frac{1}{A}\right) \\ A < 0 \quad &\Longrightarrow \quad y = \frac{-1}{x - \frac{1}{A}} \quad \text{na } \left(\frac{1}{A}, \infty\right). \end{aligned}$$

Prodloužení mimo popsané intervaly není možné díky nevlastním limitám.

Pokud $A = 0$, možným řešením je $y \equiv 0$. Jeho jednoznačnost se dá dokázat jako u předchozí rovnice. To opět není nijak překvapivé z pohledu Picard–Lindelöfovy existenční věty, neboť funkce $(x, y) \mapsto \alpha y$ je spojitá a lokálně lipschitzovská ve druhé složce.



Obrázek 10.2: Náčrt částí několika větví obecného řešení rovnice $y' = y^2$.

(iii) Uvažme úlohu

$$y'(x) = 3y^{\frac{2}{3}}$$
$$y(0) = 1.$$

Předchozí větu můžeme používat v situacích $\Omega = \mathbb{R} \times (-\infty, 0)$ a $\Omega = \mathbb{R} \times (0, \infty)$. Zde shodně máme

$$G(y) = \int \frac{\mathrm{d}y}{g(y)} = \int \frac{\mathrm{d}y}{3y^{\frac{2}{3}}} = y^{\frac{1}{3}} = x + C.$$

V případě počáteční podmínky $y(0) = 1$ pracujeme na $\Omega = \mathbb{R} \times (0, \infty)$, a proto máme jednoznačné maximální řešení

$$y = (x+1)^3 \qquad \text{na } (-1, \infty).$$

Jednoznačnost a maximalita se týkají množiny $\Omega = \mathbb{R} \times (0, \infty)$. Nepřekvapí tedy, že naše řešení je možné prodloužit předpisem

$$y(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, -1] \\ (x+1)^3 & \text{pro } x \in [-1, \infty). \end{cases}$$

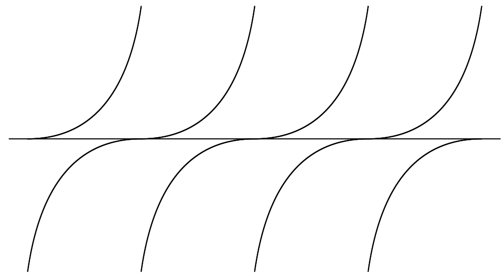
nebo

$$y(x) = \begin{cases} (x-\alpha)^3 & \text{pro } x \in (-\infty, \alpha) \\ 0 & \text{pro } x \in (-\alpha, -1] \\ (x+1)^3 & \text{pro } x \in [-1, \infty), \end{cases}$$

kde $\alpha \in (-\infty, 1]$. Jiné případy nenastanou díky jednoznačnosti, kterou nám dává předchozí věta pro $\Omega = \mathbb{R} \times (-\infty, 0)$.

Při počáteční podmínce $y(0) = 0$ je situace ještě o něco složitější.

Z pohledu Peanovy existenční věty existuje alespoň jedno řešení pro každou počáteční podmínku $y(x_0) = y_0$ ($(x_0, y_0) \in \mathbb{R}^2$), neboť funkce $(x, y) \mapsto 3y^{\frac{2}{3}}$ je spojitá na $\mathbb{R}^2$. Picard–Lindelöfova existenční věta se nedá aplikovat na žádné otevřené podmnožině $\mathbb{R}^2$, která obsahuje x-ovou osu, neboť pak nemáme lokální lipschitzovskost funkce $(x, y) \mapsto 3y^{\frac{2}{3}}$ ve druhé proměnné.



Obrázek 10.3: Náčrt částí několika větví obecného řešení rovnice $y' = 3y^{\frac{2}{3}}$.

V předchozích příkladech se často vyskytovala situace, kdy pro jisté $\theta \in \mathbb{R}$ platí $g(\theta) = 0$. Pak automaticky $y \equiv \theta$ je řešení rovnice $y' = g(y)$. Toto řešení se nazývá *triviální řešení*. Viděli jsme, že v některých případech se dá napojit triviální řešení na řešení získaná pomocí Věty o řešení rovnice $y' = g(y)$ (Věta 10.4.1). Této problematice se nyní budeme věnovat podrobněji. Nechť v dalším je θ nulový bod funkce g a $y = G^{-1}(x+C)$ je řešení získané pomocí Věty o řešení rovnice $y' = g(y)$ na $\Omega = (\theta, \tau) \times \mathbb{R}$ pro jisté $\tau \in (\theta, \infty]$. Tedy g je nenulová a spojitá na (θ, τ) a G je monotonní na (θ, τ). V této situaci vždy existují $\lim_{y \to \theta_+} G(y)$, $\lim_{y \to \tau_-} G(y)$ a řešení $y = G^{-1}(x+C)$ máme na intervalu s krajními body $a := \lim_{y \to \theta_+} G(y) - C$ a $b := \lim_{y \to \tau_-} G(y) - C$, neboť v naší situaci platí

$$y = G^{-1}(x + C) \qquad \Longleftrightarrow \qquad x = G(y) - C.$$

Pokud jsou obě výše uvedené limity nevlastní, řešení máme definované na celém $\mathbb{R}$, což je jednak nejlepší možný výsledek z hlediska definičního oboru řešení, zároveň odpadá možnost slepit řešení s řešením triviálním. V dalším se tedy budeme zabývat třeba situací $\lim_{y \to \theta_+} G(y) \in \mathbb{R}$.

Tvrzení 10.4.3 (O slepování řešení)**.** *Nechť v situaci uvedené výše je $a \in \mathbb{R}$ a funkce g je spojitá zprava v bodě θ. Pak lze v bodě a řešení $y = G^{-1}(x+C)$ slepit s triviálním řešením identicky rovným θ.*

Důkaz. Pro jednoduchost značení uvažujme jen případ $g > 0$ na (θ, τ). Ukažme, že

$$y = \begin{cases} \theta & \text{pro } x \le a \\ G^{-1}(x + C) & \text{pro } x \in (a, b) \end{cases}$$

řeší úlohu $y' = g(y)$ na $(-\infty, b)$. Připomeňme, že platí

$$\lim_{x \to a_+} G^{-1}(x + C) = \theta,$$

proto je y spojitá funkce. Dále ze spojitosti g zprava v bodě θ dostáváme

$$\frac{\mathrm{d}}{\mathrm{d}x} G^{-1}(x + C) = \frac{1}{G'(G^{-1}(x + C))} = g(G^{-1}(x + C)) \overset{x \to a_+}{\to} g(\theta) = 0.$$

Proto Věta o limitě derivací (Věta 6.3.9) dává $y'(a) = 0$ a platí $y'(a) = 0 = g(\theta) = g(y(a)))$. □

Pokud je funkce g lipschitzovská na nějakém pravém okolí bodu θ, automaticky nastane situace, že a je nevlastní.

Tvrzení 10.4.4 (O nevlastní mezi definičního oboru řešení)**.** *Nechť v situaci uvedené výše je funkce g lipschitzovská na $[\theta, \theta+\delta)$ pro jisté $\delta > 0$. Pak je a nevlastní.*

Důkaz. Opět se zabývejme jen případem $g > 0$ na (θ, τ). Pak G je rostoucí na $(\theta, \theta + \delta)$ a pro libovolná $\theta < \xi_1 < \xi_2 < \theta + \delta$ máme

$$G(\xi_2) - G(\xi_1) = \int_{\xi_1}^{\xi_2} \frac{\mathrm{d}y}{g(y)} \geq \int_{\xi_1}^{\xi_2} \frac{\mathrm{d}y}{g(\theta) + K(y-\theta)} = \int_{\xi_1}^{\xi_2} \frac{\mathrm{d}y}{K(y-\theta)}$$
$$= \frac{1}{K}[\log(y-\theta)]_{\xi_1}^{\xi_2} \overset{\xi_1 \to \theta_+}{\to} \infty.$$

Z toho zřejmě plyne dokazovaný výsledek. □

Lipschitzovskost je pro výše popsaný jev podmínkou postačující, nikoliv nutnou.

Příklad 10.4.5. Definujeme-li

$$g(y) = \begin{cases} y \log|y| & \text{pro } y \neq 0 \\ 0 & \text{pro } y = 0, \end{cases}$$

máme funkci, která je spojitá (podle Peanovy existenční věty máme řešení pro libovolnou počáteční podmínku $y(0) = x_0$, kde $x_0 \in \mathbb{R}, y_0 \in \mathbb{R} \setminus \{-1, 1\}$) a platí

$$\int_0^\delta \frac{\mathrm{d}y}{y \log y} = [\log|\log(y)|]_0^\delta = \infty$$

pro každé $\delta \in (0, 1)$ (projde tedy konstrukce z předchozího důkazu), ale g není lipschitzovská na žádném okolí počátku.

10.4.3 Rovnice $y' = f(x)g(y)$

Této rovnici se říká *rovnice se separovanými proměnnými.* Jejím řešením je

$$y = G^{-1}(F(x) + C), \qquad \text{kde } G(y) = \int \frac{\mathrm{d}y}{g(y)} \text{ a } F(x) = \int f(x)\,\mathrm{d}x.$$

Výsledný vzorec si lze pamatovat pomocí zjednodušené myšlenky jeho důkazu (tato myšlenka je ale opět zjednodušená a nelze ji považovat za důkaz)

$$y' = \frac{\mathrm{d}y}{\mathrm{d}x} = f(x)g(y) \rightsquigarrow \frac{\mathrm{d}y}{g(y)} = f(x)\,\mathrm{d}x \rightsquigarrow \int \frac{\mathrm{d}y}{g(y)} = \int f(x)\,\mathrm{d}x = F(x) + C.$$

Věta 10.4.6 (O řešení rovnice se separovanými proměnnými). *Nechť $f\colon \mathbb{R} \to \mathbb{R}$ je spojitá na $(a,b) \subset \mathbb{R}$ a $g\colon \mathbb{R} \to \mathbb{R}$ je spojitá a nenulová na $(\alpha,\beta) \subset \mathbb{R}$. Nechť F je primitivní funkce k f na (a,b) a G je primitivní funkce k $y \mapsto \frac{1}{g(y)}$ na (α,β). Pak na intervalu $G((\alpha,\beta))$ existuje inverzní funkce G^{-1} a každé maximální řešení v $\Omega = (a,b) \times (\alpha,\beta)$ má tvar*

$$y(x) = G^{-1}(F(x) + C),$$

kde $C \in \mathbb{R}$, a je definováno na otevřeném intervalu

$$I := \{x \in (a,b)\colon \exists y \in (\alpha,\beta) \quad G(y) = F(x) + C\}.$$

Navíc každým bodem $(x_0, y_0) \subset \Omega$ prochází právě jedno maximální řešení (v Ω).

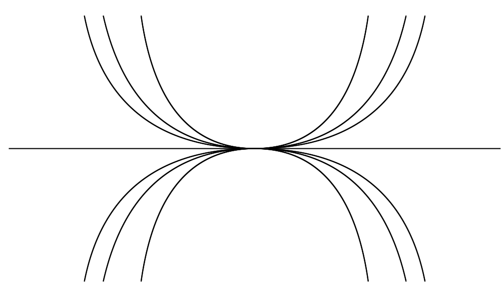
Důkaz. Důkaz se získá nenáročnou modifikací důkazu Věty o řešení rovnice $y' = g(y)$ (Věta 10.4.1), kterou přenecháváme čtenáři jako cvičení. Jen důkaz toho, že I je otevřený interval, vyžaduje hlubší zamyšlení. $\square$

Příklad 10.4.7. Hledejme obecné řešení úlohy $y' = \frac{2y}{x}$. Po integraci dostáváme

$$\log|y| = \log x^2 + C \qquad \Longrightarrow \qquad y = Kx^2.$$

Tento postup jsme mohli aplikovat na $\Omega = (-\infty,0) \times (-\infty,0)$, $\Omega = (-\infty,0) \times (0,\infty)$, $\Omega = (0,\infty) \times (-\infty,0)$, nebo $\Omega = (0,\infty) \times (0,\infty)$. V případě počáteční podmínky $y(x_0) = 0$ je řešením

$$y \equiv 0 \quad \text{na } (-\infty,0) \qquad \text{pro } x_0 < 0 \qquad \text{a} \qquad y \equiv 0 \quad \text{na } (0,\infty) \qquad \text{pro } x_0 > 0.$$



Obrázek 10.4: Náčrt částí několika větví obecného řešení rovnice $xy' = 2y$.

Slepování řešení zde není možné. Pokud bychom však pracovali s rovnicí $xy' = 2y$, slepovat bychom mohli v počátku (všechna uvažovaná řešení mají v počátku nulovou limitu a rovněž nulovou limitu derivací) a dostali bychom obecné řešení

$$y(x) = \begin{cases} Kx^2 & \text{pro } x \le 0 \\ Lx^2 & \text{pro } x \ge 0, \end{cases}$$

kde $K, L \in \mathbb{R}$. Jednoznačnost by nám daly počáteční podmínky $y(x_1) = y_1$ a $y(x_2) = y_2$, kde $x_1 < 0 < x_2$ a $y_1, y_2 \in \mathbb{R}$. Zajímavou variantou předchozího příkladu je úloha $xy' = y$. Opět je porušena jednoznačnost v počátku, ale řešení není možné slepovat v počátku dle libosti.

10.4.4 Homogenní diferenciální rovnice

Budeme se zabývat rovnicí tvaru

$$y' = f(x, y),$$

kde funkce $f\colon \mathbb{R}^2 \to \mathbb{R}$ je definovaná na $\mathbb{R}^2 \setminus \{0, 0\}$ a platí zde pro každé $\lambda \neq 0$

$$f(\lambda x, \lambda y) = f(x, y).$$

Poznámka 10.4.8. (i) Funkce f je tedy konstantní na jednotlivých přímkách vycházejících z počátku (neuvažujeme hodnotu v počátku).
(ii) Název tohoto typu rovnic je odvozen od následující terminologie. Funkce $g\colon \mathbb{R}^2 \to \mathbb{R}$ definovaná na celém $\mathbb{R}^2$ je *homogenní stupně* $\alpha \in \mathbb{N}$ (nebo také α-*homogenní*), jestliže pro všechna $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ a $\lambda > 0$ platí

$$g(\lambda x, \lambda y) = \lambda^\alpha g(x, y).$$

Podíl dvou homogenních funkcí stejného řádu vede na typ funkce, který uvažujeme v naší diferenciální rovnici, jestliže jmenovatel je nenulový na každém paprsku vycházejícím z počátku.

Rovnici řešíme pro $x \neq 0$ (nakonec se získané výsledky pokusíme slepit) následujícím postupem. Nejprve si přepišme

$$f(x, y) = f\Big(x \cdot 1, x\frac{y}{x}\Big) = f\Big(1, \frac{y}{x}\Big) := g\Big(\frac{y}{x}\Big)$$

a definujeme pomocnou funkci $z(x) = \frac{y(x)}{x}$. Opět si napišme stručné (poněkud nekorektní) schéma dořešení úlohy. Postupujeme následovně

$$y(x) = xz(x) \rightsquigarrow y'(x) = z(x) + xz'(x) = g\Big(\frac{y(x)}{x}\Big) = g(z(x)) \rightsquigarrow z'(x) = \frac{g(z) - z}{x}$$

a úplně napravo jsme získali úlohu se separovanými proměnnými, kterou už umíme řešit. Přesněji, podle Věty o řešení rovnice se separovanými proměnnými (Věta 10.4.6) můžeme použít zmíněnou metodu v situaci, kdy funkce $z \mapsto f(1, z) - z$ je spojitá a nenulová na jistém $(\alpha, \beta) \subset \mathbb{R}$ a pracujeme na množinách $\Omega = (-\infty, 0) \times (\alpha, \beta)$ a $\Omega = (0, \infty) \times (\alpha, \beta)$ (vychází nám výsledek tvaru $z(x) = H^{-1}(\log|x| + C)$, kde H^{-1} je inverzní funkcí k $h\colon z \mapsto \frac{1}{f(1,z)-z}$). Protože výsledné řešení z je diferencovatelné, je diferencovatelné i $y(x) = xz(x)$ a splňuje požadovaný vztah

$$y'(x) = z(x) + xz'(x) = g(z(x)) = g\Big(\frac{y(x)}{x}\Big) = f\Big(1, \frac{y(x)}{x}\Big) = f(x, y)$$

na intervalech získaných při řešení úlohy $z'(x) = \frac{g(z)-z}{x}$ (žádný z těchto intervalů neobsahuje počátek). Zbývá ještě ukázat, že úloha $y' = f(x,y)$ nemůže mít jiná řešení než ta, která jsme získali výše. To už je ale snadné. Pokud je totiž y řešením úlohy $y' = f(x,y)$ na nějakém intervalu neobsahujícím počátek, je také diferencovatelná funkce $z := \frac{y}{x}$ a dostáváme

$$z' = \frac{xy' - y}{x^2} = \frac{xf(x,y) - xz}{x^2} = \frac{f(1,z) - z}{x}.$$

Příklad 10.4.9. Řešme úlohu $y' = \frac{x}{y} + \frac{y}{x}$. Po použití rovnosti $y = xz$ pro $x \neq 0$ máme

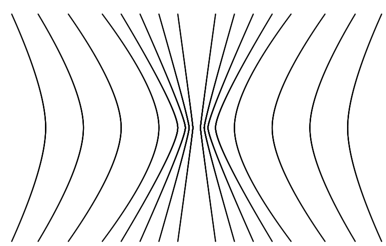
$$z + xz' = \frac{1}{z} + z \qquad \Longleftrightarrow \qquad z' = \frac{1}{x}\frac{1}{z}.$$

Na jednotlivých kvadrantech pak máme

$$\frac{z^2}{2} = \int z \,\mathrm{d}z = \int \frac{\mathrm{d}x}{x} = \log|x| + C.$$

Odtud

$$z^2 = 2(\log|x| + C) = \log x^2 + 2C \quad \Longrightarrow \quad y^2 = x^2(\log x^2 + 2C) \qquad \text{pro } |x| > \mathrm{e}^{-C}.$$



Obrázek 10.5: Náčrt částí několika větví obecného řešení rovnice $y' = \frac{x}{y} + \frac{y}{x}$. Pozor, dvojice netriviálních řešení $y = \pm|x|\sqrt{\log x^2 + 2C}$ není možné slepit.

Příklad 10.4.10. Řešme úlohu $yy' + x = \sqrt{x^2 + y^2}$. Po úpravě (dočasně se omezíme na $x \neq 0$, $y \neq 0$) dostáváme

$$y' = \frac{1}{y}|x|\sqrt{1 + \frac{y^2}{x^2}} - \frac{x}{y}.$$

Odtud po použití rovnosti $y = xz$ máme

$$z + xz' = \operatorname{sign} x \frac{1}{z}\sqrt{1 + z^2} - \frac{1}{z}$$

a po úpravě dostáváme

$$z' = \frac{1}{x}\frac{\operatorname{sign} x\sqrt{1+z^2} - 1 - z^2}{z}.$$

Pro $x < 0$ a $z < 0$ nebo $z > 0$ máme

$$\log(1 + \sqrt{1+z^2}) = \int \frac{z\mathrm{d}z}{\sqrt{1+z^2} + 1 + z^2} = -\int \frac{\mathrm{d}x}{x} = -\log|x| + C.$$

Odtud pro $K > 0$ a $x \in (-\frac{K}{2}, 0)$

$$\sqrt{1+z^2} + 1 = \frac{K}{|x|} \quad\Longrightarrow\quad \sqrt{1+\frac{y^2}{x^2}} + 1 = \frac{K}{|x|} \quad\Longrightarrow\quad y^2 = \Big(\frac{K}{|x|} - 1\Big)^2 x^2 - x^2.$$

Po úpravě obdržíme

$$y^2 = K^2 - 2K|x|.$$

Na $\Omega = (-\infty, 0) \times (-\infty, 0)$ máme obecná řešení tvaru

$$y = -\sqrt{K^2 + 2Kx} \qquad \text{pro } x \in \Big(-\frac{K}{2}, 0\Big).$$

Na $\Omega = (-\infty, 0) \times (0, \infty)$ máme obecná řešení tvaru

$$y = \sqrt{K^2 + 2Kx} \qquad \text{pro } x \in \Big(-\frac{K}{2}, 0\Big).$$

Pro $x > 0$ a $z < 0$ nebo $z > 0$ máme

$$\log(\sqrt{1+z^2} - 1) = \int \frac{z\mathrm{d}z}{1 + z^2 - \sqrt{1+z^2}} = -\int \frac{\mathrm{d}x}{x} = \log x + C.$$

Odtud pro $K > 0$ a $x > 0$

$$\sqrt{1+z^2} - 1 = \frac{K}{x} \quad\Longrightarrow\quad \sqrt{1+\frac{y^2}{x^2}} - 1 = \frac{K}{x} \quad\Longrightarrow\quad y^2 = \Big(\frac{K}{x} + 1\Big)^2 x^2 - x^2.$$

Po úpravě získáváme

$$y^2 = K^2 + 2Kx.$$

Na $\Omega = (0, \infty) \times (-\infty, 0)$ máme obecná řešení tvaru

$$y = -\sqrt{K^2 + 2Kx} \qquad \text{pro } x \in (0, \infty).$$

Na $\Omega = (0, \infty) \times (0, \infty)$ máme obecná řešení tvaru

$$y = \sqrt{K^2 + 2Kx} \qquad \text{pro } x \in (0, \infty).$$

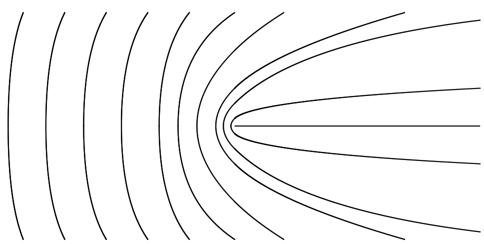
Pro $x = 0$ můžeme části řešení získané na $(-\frac{K}{2}, 0)$ a $(0, \infty)$ slepit, neboť funkce $x \mapsto \sqrt{K^2 + 2Kx}$ je třídy $C^1((-\frac{K}{2}, \infty))$ (výraz $yy' + x - \sqrt{x^2 + y^2}$ je tedy spojitý na

$(-\frac{K}{2}, \infty)$, a proto jeho nulovost na $(-\frac{K}{2}, \infty) \setminus \{0\}$ implikuje nulovost i v počátku). Pro $y = 0$ požaduje původní rovnice $x = \sqrt{x^2}$. Na $(-\infty, 0)$, proto neexistuje řešení protínající osu x. Na intervalu $[0, \infty)$ rovnost platí vždy, máme tedy „triviální“ řešení

$$y \equiv 0 \qquad \text{na } (0, \infty)$$

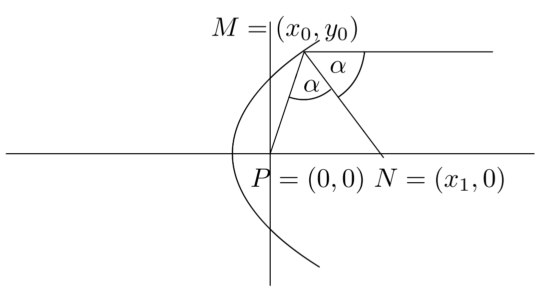
(pojem řešení jsme definovali pro otevřený interval) a toto řešení se zřejmě nedá slepit s žádným řešením dříve získaným. Celkově tedy máme výše popsané triviální řešení a řešení tvaru

$$y = \sqrt{K^2 + 2Kx} \qquad \text{a} \qquad y = -\sqrt{K^2 + 2Kx} \qquad \text{definovaná na } \left(-\frac{K}{2}, \infty\right).$$



Obrázek 10.6: Náčrt částí několika větví obecného řešení rovnice $yy' + x = \sqrt{x^2 + y^2}$. Pozor, dvojice netriviálních řešení $y = \pm\sqrt{K^2 + 2Kx}$ není možné slepit.

Poznámka 10.4.11. Předchozí příklad odpovídá hledání tvaru zrcadla, pro které jsou paprsky rovnoběžné s optickou osou po odrazu soustředěny do jednoho bodu.



Obrázek 10.7: Odvození rovnice $yy' + x = \sqrt{x^2 + y^2}$ z úlohy o hledání tvaru zrcadla.

Přesněji, pokud si na Obrázku 10.7 označíme souřadnice bodu M (tedy průsečíku dráhy paprsku rovnoběžného s optickou osou zrcadla a zrcadla) jako (x_0, y_0), pak bod $N = (x_1, 0)$ (průsečík osy úhlu určeného dráhou paprsku a optické osy zrcadla) leží na normále k zrcadlu, tedy na přímce

$$y - y_0 = -\frac{1}{y'(x_0)}(x - x_0).$$

Dosazením bodu N máme

$$-y_0 = -\frac{1}{y'(x_0)}(x_1 - x_0),$$

tedy

$$x_1 = y'(x_0)y_0 + x_0.$$

Protože trojúhelník MNP je rovnoramenný, je

$$x_0^2 + y_0^2 + |MP|^2 = |NP|^2 = x_1^2,$$

proto

$$y'(x_0)y_0 + x_0 = \sqrt{x_0^2 + y_0^2}.$$

Výsledná rovnice zrcadla je

$$yy' + x = \sqrt{x^2 + y^2},$$

dostáváme tedy rovnici z Příkladu 10.4.10.

10.4.5 Rovnice, které lze převést na homogenní diferenciální rovnici

Uvažujme rovnici typu

$$y' = f\Big(\frac{ax + by + c}{\alpha x + \beta y + \gamma}\Big)$$

a předpokládejme, že $a\beta \neq b\alpha$. Pak má soustava

$$\begin{aligned} ax + by + c &= 0 \\ \alpha x + \beta y + \gamma &= 0 \end{aligned}$$

jednoznačné řešení $(x_0, y_0) \in \mathbb{R}^2$. Položme

$$\xi := x - x_0 \qquad \text{a} \qquad \eta := y - y_0$$

(vlastně se jedná jen o posunutí souřadných os do průsečíku přímek $ax+by+c = 0$ a $\alpha x + \beta y + \gamma = 0$). Okamžitě pro funkci $\eta\colon \xi \mapsto \eta(\xi) := y(\xi + x_0) - y_0$ dostáváme

$$\frac{\mathrm{d}\eta}{\mathrm{d}\xi}(\xi) = \frac{\mathrm{d}y}{\mathrm{d}x}(\xi + x_0) = f\Big(\frac{a(\xi + x_0) + b(\eta + y_0) + c}{\alpha(\xi + x_0) + \beta(\eta + y_0) + \gamma}\Big) = f\Big(\frac{a\xi + b\eta}{\alpha\xi + \beta\eta}\Big),$$

což je homogenní diferenciální rovnice.

Pokud není splněna podmínka $a\beta \neq b\alpha$, jedná se o snazší úlohu, která převod na homogenní rovnici nevyžaduje. Abychom si toto ukázali, rozlišíme tři případy.

V prvním případě je $\beta = 0$ a $\alpha \neq 0$, vlastnost $a\beta = b\alpha$ pak dává $b = 0$ a máme úlohu

$$y' = f\Big(\frac{ax + c}{\alpha x + \gamma}\Big),$$

která se řeší integrací.

Ve druhém případě máme $(\alpha, \beta) = (0, 0)$, formule má pak smysl jen pro $\gamma \neq 0$ a můžeme psát

$$y' = f\Big(\frac{a}{\gamma}x + \frac{b}{\gamma}y + \frac{c}{\gamma}\Big).$$

Pokud je nyní $b = 0$, úlohu opět řešíme integrací. Pokud $b \neq 0$, definujeme novou funkci $z = \frac{a}{\gamma}x + \frac{b}{\gamma}y + \frac{c}{\gamma}$ a dostáváme

$$z' = \frac{a}{\gamma} + \frac{b}{\gamma}y' = \frac{a}{\gamma} + \frac{b}{\gamma}f(z),$$

což je rovnice se separovanými proměnnými.

Konečně, ve třetím případě máme $\beta \neq 0$ a vlastnost $a\beta = b\alpha$ implikuje $a = \frac{b}{\beta}\alpha$. Odtud

$$\frac{ax + by + c}{\alpha x + \beta y + \gamma} = \frac{\frac{b}{\beta}\alpha x + \frac{b}{\beta}\beta y + c}{\alpha x + \beta y + \gamma} = \frac{b}{\beta} + \frac{c - \frac{b}{\beta}\gamma}{\alpha x + \beta y + \gamma}$$

a definice nové funkce $z = \alpha x + \beta y + \gamma$ dává

$$z' = \alpha + \beta y' = \alpha + \beta f\Big(\frac{b}{\beta} + \frac{c - \frac{b}{\beta}\gamma}{z}\Big),$$

což je opět rovnice se separovanými proměnnými.

Poznámka 10.4.12. U právě probíraného typu diferenciálních rovnic jsme čtenáři nenabídli přehlednou větu, která by popisovala přesný tvar obecného řešení a jeho definičních oborů. Máme k tomu dva důvody. Jednak se dá snadno nahlédnout, že ve všech výše uvedených situacích, které vyžadovaly zavedení pomocné funkce z či η, jsou všechna řešení původní úlohy pro y jednoznačně určena všemi řešeními úlohy pro pomocnou funkci. Na druhou stranu, kupříkladu v situaci $\alpha = \beta = 0$ a $b \neq 0$ na definiční obor řešení mají zásadní vliv množiny nenulovosti funkce $z \mapsto \frac{a}{\gamma} + \frac{b}{\gamma}f(z)$, které nezávisí jen na chování funkce f ale i na parametrech a, b, γ. To činí obecnou charakterizaci množin nenulovosti poměrně složitou, třebaže v konkrétních úlohách se typicky o obtížný problém nejedná.

Příklad 10.4.13. Řešme úlohu

$$y' = 2\Big(\frac{y + 2}{x + y - 1}\Big)^2.$$

Soustava

$$\begin{aligned} y_0 + 2 &= 0 \\ x_0 + y_0 - 1 &= 0 \end{aligned}$$

má řešení $x_0 = 3$ a $y_0 = -2$. Závádíme tedy novou proměnnou $\xi := x - 3$, novou funkci $\eta(\xi) := y(\xi + 3) + 2$ a dostáváme

$$\eta' = 2\Big(\frac{\eta}{\xi + \eta}\Big)^2.$$

Dostali jsme homogenní diferenciální rovnici. Rozlišujeme případy $\xi \in (-\infty, 0)$ a $\xi \in (0, \infty)$. Zavádíme pomocnou funkci $z := \frac{\eta}{\xi}$ (tedy $\eta = \xi z$ a $\eta' = z + \xi z'$) a máme

$$z + \xi z' = 2\Big(\frac{z}{z+1}\Big)^2 \qquad \Longleftrightarrow \qquad z' = -\frac{1}{\xi}\frac{z^3 + z}{z^2 + 2z + 1}.$$

Tím jsme přešli k rovnici se separovanými proměnnými. Triviální řešení je $z \equiv 0$ a dále dostáváme

$$2\arctan z + \log|z| = \int \frac{z^2 + 2z + 1}{z^3 + z}\,\mathrm{d}z = -\int \frac{\mathrm{d}\xi}{\xi} = -\log|\xi| + C.$$

Označíme-li výraz na levé straně jako $\Phi(z)$, není těžké ověřit, že funkce Φ je klesající na $(-\infty, 0)$ a zobrazuje tento interval na $\mathbb{R}$. Dále Φ je rostoucí na $(0, \infty)$ a zobrazuje tento interval také na $\mathbb{R}$. Označme Ψ_1 inverzi k restrikci Φ na $(-\infty, 0)$ a Ψ_2 inverzi k restrikci Φ na $(0, \infty)$. Pro každé $C \in \mathbb{R}$ proto máme čtveřici řešení

$$\begin{aligned}
z_1 &= \Psi_1(-\log|\xi| + C) && \text{na } (-\infty, 0)\\
z_2 &= \Psi_1(-\log|\xi| + C) && \text{na } (0, \infty)\\
z_3 &= \Psi_2(-\log|\xi| + C) && \text{na } (-\infty, 0)\\
z_4 &= \Psi_2(-\log|\xi| + C) && \text{na } (0, \infty).
\end{aligned}$$

Odtud máme po přidání triviálního řešení (získaného ze $z \equiv 0$)

$$\begin{aligned}
\eta_1 &= \xi\Psi_1(-\log|\xi| + C) && \text{na } (-\infty, 0)\\
\eta_2 &= \xi\Psi_1(-\log|\xi| + C) && \text{na } (0, \infty)\\
\eta_3 &= \xi\Psi_2(-\log|\xi| + C) && \text{na } (-\infty, 0)\\
\eta_4 &= \xi\Psi_2(-\log|\xi| + C) && \text{na } (0, \infty)\\
\eta_5 &\equiv 0 && \text{na } (-\infty, 0)\\
\eta_6 &\equiv 0 && \text{na } (0, \infty).
\end{aligned}$$

Konečně dostáváme (rovnou uvádíme množiny, kde lze k zadané počáteční podmínce nalézt jednoznačné C)

$$\begin{aligned}
y_1 &= -2 + (x-3)\Psi_1(-\log|x-3| + C) && \text{pro } \Omega_1 = (-\infty, 3) \times (-2, \infty)\\
y_2 &= -2 + (x-3)\Psi_1(-\log|x-3| + C) && \text{pro } \Omega_2 = (3, \infty) \times (-\infty, -2)\\
y_3 &= -2 + (x-3)\Psi_2(-\log|x-3| + C) && \text{pro } \Omega_3 = (-\infty, 3) \times (-\infty, -2)\\
y_4 &= -2 + (x-3)\Psi_2(-\log|x-3| + C) && \text{pro } \Omega_4 = (3, \infty) \times (-2, \infty)\\
y_5 &\equiv -2 && \text{na } (-\infty, 3)\\
y_6 &\equiv -2 && \text{na } (3, \infty).
\end{aligned}$$

Příklad 10.4.14. Řešme úlohu $y' = \cos(y - x)$. To odpovídá situaci $\alpha = \beta = 0$. Definujeme pomocnou funkci $z = y - x$ a dostáváme

$$z' = y' - 1 = \cos z - 1$$

(povšimněte si ještě, že y řeší úlohu $y' = \cos(y - x)$ na nějakém intervalu (a, b) právě tehdy, když $z = y - x$ řeší úlohu $z' = \cos z - 1$ na (a, b)). Druhá úloha je rovnice se separovanými proměnnými řešitelná na množinách tvaru $\Omega_k := \mathbb{R} \times (2k\pi, 2(k+1)\pi)$, $k \in \mathbb{Z}$, kde dostáváme

$$\cot \frac{z}{2} = -\int \frac{\mathrm{d}z}{2\sin^2 \frac{z}{2}} = \int \frac{\mathrm{d}z}{\cos z - 1} = \int \mathrm{d}x = x + C.$$

Na Ω_k pak máme

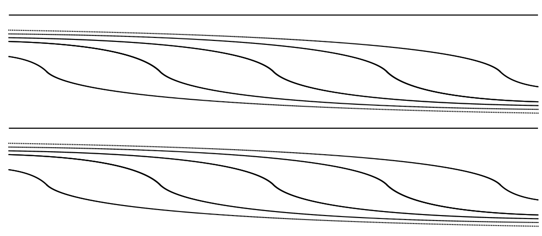
$$z = 2\,\mathrm{arccot}(x + C) + 2k\pi.$$

Mimo $\bigcup_{k\in\mathbb{Z}} \Omega_k$ ještě získáváme triviální řešení

$$z \equiv 2k\pi.$$

Pro původní úlohu tedy dostáváme řešení typů

$$y = 2k\pi + x \qquad a \qquad y = 2\,\mathrm{arccot}(x + C) + 2k\pi + x.$$



Obrázek 10.8: Náčrt částí několika větví obecného řešení rovnice $z' = \cos z - 1$.

10.4.6 Lineární diferenciální rovnice prvního řádu

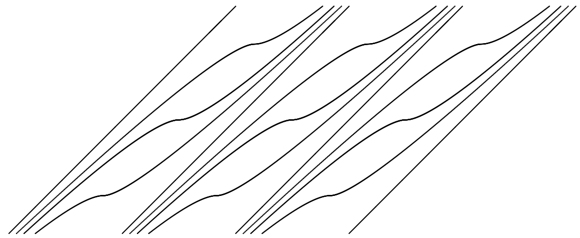
Lineární diferenciální rovnice prvního řádu je rovnice typu

$$y'(x) + p(x)y(x) = f(x),$$

pro kterou uvažujeme počáteční podmínku

$$y(x_0) = y_0.$$

Předpokládáme, že $p, f \in C((a, b))$, $x_0 \in (a, b)$ a $y_0 \in \mathbb{R}$. Tento typ rovnic jsme si už představili v kapitole o primitivních funkcích. Naučili jsme se řešení *metodou*



Obrázek 10.9: Náčrt částí několika větví obecného řešení rovnice $y' = \cos(y - x)$.

integračního faktoru a dokázali existenci a jednoznačnost řešení. Pro úplnost si metodu integračního faktoru stručně připomeňme. Původní rovnici přenásobíme výrazem $\mathrm{e}^{\int p(x)\,\mathrm{d}x}$ a povšimneme si, že levou stranu je pak možné napsat jako derivaci součinu

$$\Big((y(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x}\Big)' = y'(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x} + p(x)y(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x} = f(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x}.$$

Odtud

$$y(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x} = \int f(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x}\,\mathrm{d}x + C$$

a dostáváme

$$y(x) = C\mathrm{e}^{-\int p(x)\,\mathrm{d}x} + \mathrm{e}^{-\int p(x)\,\mathrm{d}x}\int f(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x}\,\mathrm{d}x.$$

Díky počáteční podmínce nyní určíme jednoznačné $C \in \mathbb{R}$ (neboť $\mathrm{e}^{-\int p(x)\,\mathrm{d}x} > 0$ na (a, b)), aby platilo $y(x_0) = y_0$. Všechny primitivní funkce existují díky spojitosti integrandů.

Poznámka 10.4.15. (i) Integrační faktor není určen jednoznačně. Můžeme použít jakýkoliv jeho násobek, což je ostatně důvod, proč se při jeho hledání nestaráme o aditivní konstantu po integraci.
(ii) Díky tomu, že už umíme řešit rovnici se separovanými proměnnými, už není nutné si vzorec $\mathrm{e}^{\int p(x)\,\mathrm{d}x}$ pamatovat, ale můžeme si jej odvodit. Skutečně, hledáme-li integrační faktor Q tak, aby

$$Q(y' + py) = (yQ)'$$

řešíme

$$Qy' + pQy = Qy' + Q'y \qquad \Longleftrightarrow \qquad pQy = Q'y \qquad \Longleftarrow \qquad Q' = pQ.$$

Rovnice úplně napravo má separované proměnné a dostáváme

$$\log|Q| = \int \frac{\mathrm{d}Q}{Q} = \int p(x)\,\mathrm{d}x + C,$$

tedy pro $K \in \mathbb{R}$ (případ $K = 0$ nám dalo triviální řešení)

$$Q(x) = K\mathrm{e}^{\int p(x)\,\mathrm{d}x}.$$

Navíc podle Věty o řešení rovnice se separovanými proměnnými (Věta 10.4.6) dostáváme vždy řešení na (a, b) (díky spojitosti $p(x)$ na (a, b) a tomu, že v roli funkce G^{-1} zde vystupuje exponenciála, která nezúží definiční obor řešení).

Poznamenejme ještě, že tento typ rovnic je také možné řešit metodou variace konstant. Připomeňme, že tuto metodu jsme si představili v kapitole o primitivních funkcích pro lineární obyčejné diferenciální rovnice 2. řádu. Nejprve se vyřeší úloha s nulovou pravou stranou

$$y_h' + p(x)y_h = 0,$$

což je úloha se separovanými proměnnými, pro kterou máme

$$\log|y_h| = \int \frac{\mathrm{d}y_h}{y_h} = -\int p(x)\,\mathrm{d}x + C.$$

Odtud

$$y_h(x) = K\mathrm{e}^{-\int p(x)\,\mathrm{d}x}$$

je řešení na (a, b) pro každé $K \in \mathbb{R}$ (Věta o řešení rovnice se separovanými proměnnými, tedy Věta 10.4.6, zde funguje vždy, podobně jako v předchozí poznámce). Pokud nyní do původní rovnice dosadíme $y_p(x) = c(x)\mathrm{e}^{-\int p(x)\,\mathrm{d}x}$, dostáváme

$$\begin{aligned} f(x) = y' + p(x)y &= c'(x)\mathrm{e}^{-\int p(x)\,\mathrm{d}x} - c(x)p(x)\mathrm{e}^{-\int p(x)\,\mathrm{d}x} + p(x)c(x)\mathrm{e}^{-\int p(x)\,\mathrm{d}x} \\ &= c'(x)\mathrm{e}^{-\int p(x)\,\mathrm{d}x}. \end{aligned}$$

Odtud

$$c(x) = \int f(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x}\,\mathrm{d}x$$

(aditivní konstanta po integraci nehraje žádnou roli, neboť už je obsařena v y_h) a díky linearitě diferenciálního operátoru na levé straně řešené rovnice dostáváme

$$y = y_h + y_p = K\mathrm{e}^{-\int p(x)\,\mathrm{d}x} + \mathrm{e}^{-\int p(x)\,\mathrm{d}x}\int f(x)\mathrm{e}^{\int p(x)\,\mathrm{d}x}\mathrm{d}x,$$

což je stejný výsledek, jaký dává metoda integračního faktoru.

Příklad 10.4.16. Hledejme obecné řešení rovnice $y' + \frac{1}{x}y = 3x$. Nejprve si ukažme řešení metodou integračního faktoru. Pokud bychom si vzoreček pro integrační faktor nepamatovali, řešili bychom nejprve úlohu

$$Qy' + Q\frac{1}{x}y = (Qy') \qquad \Longleftrightarrow \qquad Q\frac{1}{x}y = Q'y \qquad \Longleftarrow \qquad Q' = \frac{1}{x}Q,$$

kde dostáváme

$$\log|Q| = \int \frac{\mathrm{d}Q}{Q} = \int \frac{\mathrm{d}x}{x} = \log|x| + C.$$

Pokud bychom nyní postupovali striktně podle výše uvedených postupů, pracovali bychom s $Q = |x|$. Výhodnější je použít $Q(x) = x$ (ze zadání je jasné, že musíme pracovat zvlášť na intervalech $(-\infty, 0)$ a $(0, \infty)$, nijak tedy nevadí, když na těchto intervalech použijeme odlišné multiplikativní konstanty pro integrační faktor). Celkově tedy máme rovnici

$$(xy)' = 3x^2$$

a dostáváme

$$xy = x^3 + C \qquad \Longrightarrow \qquad y = x^2 + \frac{C}{x} \quad \text{na } (-\infty, 0) \text{ nebo na } (0, \infty).$$

Pokud bychom zvolili metodu variace konstant, nejprve bychom řešili

$$y_h' = -\frac{1}{x} y_h.$$

Odtud

$$\log |y_h| = \int \frac{\mathrm{d}y_h}{y_h} = -\int \frac{1}{x}\,\mathrm{d}x = -\log|x| + C$$

a po přidání triviálního řešení, úpravě a přeznačení multiplikativních konstant na jednotlivých intervalech (jako výše) máme

$$y_h = K\frac{1}{x}.$$

Variaci konstant provádíme v podobě $y_p = c(x)\frac{1}{x}$ a po dosazení máme

$$3x = y_p' + \frac{1}{x} y_p = c'(x)\frac{1}{x} - c(x)\frac{1}{x^2} + \frac{1}{x}c(x)\frac{1}{x} = c'(x)\frac{1}{x}.$$

Proto

$$c'(x) = 3x^2 \qquad \Longrightarrow \qquad c(x) = x^3$$

a celkově

$$y = y_h + y_p = K\frac{1}{x} + x^2 \quad \text{na } (-\infty, 0) \text{ nebo na } (0, \infty).$$

Poznámka 10.4.17. (i) Obě metody řešení obsahují dvě integrace, takže při standardním provedení bývají přibližně stejně dlouhé. Nicméně metoda integračního faktoru se dá občas urychlit tím, že se v jednoduchých situacích integrační faktor uhodne.
(ii) Uvedení dvou metod řešení u jednoho typu rovnice se může zdát jako plýtvání časem čtenáře. Uvedené metody však budou mít další uplatnění a lineární rovnice prvního řádu nabízejí jednoduchý typ problémů, kde se tyto metody dají procvičit. Metodu variace konstant budeme později používat u řešení lineárních rovnic vyššího řádu (v kapitole o primitivní funkci jsme ji používali na rovnice druhého řádu s konstantními koeficienty, tuto metodu však lze použít i na rovnice vyššího řádu a dokonce koeficienty nemusí být konstantní ale zastoupené spojitou funkcí). Metodu integračního faktoru naopak použijeme u rovnic *ve tvaru totálního diferenciálu*, což jsou rovnice prvního řádu s komplikovanější závislostí y' na x a y, než s jakou pracujeme zde. Tento typ diferenciálních rovnic však nejprve vyžaduje důkladnou přípravu z diferenciálního počtu funkcí více proměnných.

Cvičení 10.4.18. Dokažte existenci a jednoznačnost řešení lineární rovnice prvního řádu pomocí Picard–Lindelöfovy existenční věty.

10.4.7 Bernoulliova rovnice

Bernoulliovou rovnicí se nazývá

$$y'(x) + p(x)y(x) = f(x)y^{\alpha}.$$

Opět uvažujeme počáteční podmínku

$$y(x_0) = y_0.$$

Předpokládáme $p, f \in C((a,b))$, $x_0 \in (a,b)$, $y_0 \in \mathbb{R}$ a $\alpha \in \mathbb{R} \setminus \{0,1\}$.

Předně poznamenejme, že vynechání případů $\alpha = 0$ a $\alpha = 1$ nám není na škodu, neboť v takových případech se jedná o lineární rovnici prvního řádu, kterou už umíme řešit.

Pokud je $\alpha < 0$, Bernoulliova rovnice nemá dobrý smysl na x-ové ose, je tedy ještě nutné požadovat $y_0 \neq 0$ a úloha se řeší zvlášť na množinách $(a,b) \times (-\infty, 0)$ a $(a,b) \times (0,\infty)$ (bez možnosti slepení; navíc druhá z množin připadá v úvahu jen pro některé hodnoty exponentu α: celočíselné exponenty či exponenty odpovídající lichým odmocninám). Naopak, pro $\alpha \in (0,\infty) \setminus \{1\}$ máme vždy triviální řešení $y \equiv 0$. Netriviální řešení hledáme opět na množinách $(a,b) \times (-\infty, 0)$ a $(a,b) \times (0,\infty)$ (vyžaduje to metoda, kterou si uvedeme níže; druhá z množin připadá v úvahu opět jen pro některá α) a případné slepení diskutujeme na konci výpočtu.

Když máme zaručenou nenulovost y, naše úloha je ekvivalentní (rovnici dělíme y^{α})

$$y'y^{-\alpha} + p(x)y^{1-\alpha} = f(x).$$

Definujme novou funkci $z(x) = y^{1-\alpha}(x)$, máme $z' = (1-\alpha)y^{-\alpha}y'$ (oprávněnost zderivování vysvětlíme níže), a proto

$$\frac{1}{1-\alpha}z' + pz = f. \tag{10.4.1}$$

Po přenásobení konstantou $1-\alpha$ dostáváme lineární diferenciální rovnici prvního řádu, která má za našich předpokladů na p, f jednoznačné řešení pro každou počáteční podmínku z $(a,b) \times \mathbb{R}$.

Nyní každé počáteční podmínce $(x_0, y_0) \in (a,b) \times (0,\infty)$ pro původní rovnici odpovídá počáteční podmínka $(x_0, z_0) = (x_0, y_0^{1-\alpha}) \in (a,b) \times (0,\infty)$. Dále, řeší-li y původní rovnici na $(a,b) \times (0,\infty)$, pak y je diferencovatelné, platí

$$\frac{\mathrm{d}}{\mathrm{d}x}(y^{1-\alpha}(x)) = (1-\alpha)y^{-\alpha}y' \qquad \text{na } (a,b)$$

a $z = y^{1-\alpha}$ řeší rovnici (10.4.1) na $(a,b) \times (0,\infty)$. Naopak, pokud z řeší rovnici (10.4.1) na $(a,b) \times (0,\infty)$, pak je

$$\frac{\mathrm{d}}{\mathrm{d}x}z^{\frac{1}{1-\alpha}} = \frac{1}{1-\alpha}z^{\frac{\alpha}{1-\alpha}}z' \quad \text{na } (a,b)$$

a $y = z^{\frac{1}{1-\alpha}}$ splňuje

$$y' = \frac{\mathrm{d}}{\mathrm{d}x} z^{\frac{1}{1-\alpha}} = \frac{1}{1-\alpha} z^{\frac{\alpha}{1-\alpha}} z' = z^{\frac{\alpha}{1-\alpha}} (f - pz) = y^\alpha f - y^\alpha p y^{1-\alpha} = f y^\alpha - py,$$

tedy y řeší původní rovnici. Celkově jsme z existence a jednoznačnosti rovnice (10.4.1) na $(a,b)\times(0,\infty)$ získali existenci a jednoznačnost Bernoulliovy rovnice na $(a,b)\times(0,\infty)$. Podobně se postupuje na $(a,b)\times(-\infty,0)$ (pro exponenty α připouštějící práci se zápornými čísly). Obecně však může řešení Bernoulliovy rovnice na $(a,b)\times(-\infty,0)$ být také odvozeno z řešení rovnice (10.4.1) na $(a,b)\times(0,\infty)$ (podívejte se na Příklad 10.4.20). Povšimněme si, že podle Picard–Lindelöfovy věty (Věta 10.3.5) slepování s triviálním řešením může nastat jen pro $\alpha \in (0,1)$.

Příklad 10.4.19. Řešme úlohu

$$y' + 2y = y^2 \mathrm{e}^x$$

se dvěma variantami počáteční podmínky

$$y(0) = \frac{1}{2} \qquad \text{a} \qquad y(0) = -\frac{1}{2}.$$

Postupem popsaným výše dostáváme na množinách $\mathbb{R}\times(-\infty,0)$ a $\mathbb{R}\times(0,\infty)$

$$y' y^{-2} + 2y^{-1} = \mathrm{e}^x.$$

Pro $z = y^{-1}$ pak máme $z' = -y^{-2}y'$, a proto má rovnice (10.4.1) tvar

$$-z' + 2z = \mathrm{e}^x \qquad \Longleftrightarrow \qquad z' - 2z = -\mathrm{e}^x.$$

Integrační faktor e^{-2x} dává

$$z = \mathrm{e}^{2x} \int (-\mathrm{e}^{-x})\,\mathrm{d}x = \mathrm{e}^x + C\mathrm{e}^{2x}.$$

V případě počáteční podmínky $y(0) = \frac{1}{2}$ pracujeme na $\mathbb{R}\times(0,\infty)$ (na stejné množině probíhalo řešení pomocné diferenciální rovnice) a z možných větví obecného řešení

$$y(x) = \frac{1}{\mathrm{e}^x + C\mathrm{e}^{2x}}$$

nám počáteční podmínka vybírá větev

$$y(x) = \frac{1}{\mathrm{e}^x + \mathrm{e}^{2x}} \qquad \text{pro } x \in \mathbb{R}.$$

V případě počáteční podmínky $y(0) = -\frac{1}{2}$ pracujeme na $\mathbb{R}\times(-\infty,0)$ a dostáváme

$$y(x) = \frac{1}{\mathrm{e}^x - 3\mathrm{e}^{2x}} \qquad \text{pro } x \in (-\log 3, \infty).$$

Příklad 10.4.20. Řešme počáteční úlohu

$$y' + \frac{3}{2}y = \frac{3}{2}xy^{\frac{1}{3}} \qquad y(2) = 1.$$

Nejprve si povšimněme, že jako obecné řešení připadá také v úvahu triviální řešení $y \equiv 0$. Na množinách $\mathbb{R} \times (-\infty, 0)$ a $\mathbb{R} \times (0, \infty)$ postupujeme standardně a dostáváme pro pomocnou funkci $z = y^{1-\frac{1}{3}} = y^{\frac{2}{3}}$

$$y'y^{-\frac{1}{3}} + \frac{3}{2}y^{\frac{2}{3}} = \frac{3}{2}x \qquad \Longleftrightarrow \qquad z' + z = x.$$

Pro rovnici výše lze zřejmě použít integrační faktor e^x a dostáváme obecné řešení na $\mathbb{R} \times (0, \infty)$

$$z = \mathrm{e}^{-x} \int x\mathrm{e}^x \, \mathrm{d}x = \mathrm{e}^{-x}((x-1)\mathrm{e}^x + C) = x - 1 + C\mathrm{e}^{-x}.$$

Odtud máme obecné řešení Bernoulliovy rovnice na $\mathbb{R} \times (0, \infty)$ tvaru

$$y = (x - 1 + C\mathrm{e}^{-x})^{\frac{3}{2}} \qquad \text{pro } x \text{ z intervalu, kde } x - 1 + C\mathrm{e}^{-x} > 0$$

a na $\mathbb{R} \times (-\infty, 0)$

$$y = -(x - 1 + C\mathrm{e}^{-x})^{\frac{3}{2}} \qquad \text{pro } x \text{ z intervalu, kde } x - 1 + C\mathrm{e}^{-x} > 0.$$

Díky počáteční podmínce $y(2) = 1$ z množiny $\mathbb{R} \times (0, \infty)$ vybíráme jednoznačnou větev $y = (x-1)^{\frac{3}{2}}$. Protože však pro tuto funkci platí $y'(x) \to 0$ pro $x \to 1_+$, lze toto řešení slepit v bodě $(1, 0)$ s řešením triviálním. Možným řešením je tedy

$$y(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 1] \\ (x-1)^{\frac{3}{2}} & \text{pro } x \in (1, \infty). \end{cases}$$

Podobně se dá ověřit, že máme i další řešení, například

$$y(x) = \begin{cases} (x - 1 + \mathrm{e}^{-x})^{\frac{3}{2}} & \text{pro } x \in (-\infty, 0) \\ 0 & \text{pro } x \in [0, 1] \\ (x-1)^{\frac{3}{2}} & \text{pro } x \in (1, \infty). \end{cases}$$

a

$$y(x) = \begin{cases} -(x - 1 + \mathrm{e}^{-x})^{\frac{3}{2}} & \text{pro } x \in (-\infty, 0) \\ 0 & \text{pro } x \in [0, 1] \\ (x-1)^{\frac{3}{2}} & \text{pro } x \in (1, \infty). \end{cases}$$

S Bernoulliovou rovnicí úzce souvisí Riccatiova rovnice

$$y'(x) + p(x)y(x) + q(x)y^2(x) = f(x),$$

kde $p, q, f \in C((a,b))$. Obecný postup řešení není znám, ale v případě znalosti jednoho řešení y_1 dostáváme pro $z = y - y_1$

$$\begin{aligned} z' + y_1' + pz + py_1 + qy_1^2 + 2qy_1z + qz^2 &= f \\ \iff \qquad z' + pz + 2qy_1z + qz^2 &= 0 \\ \iff \qquad z' + (p + 2qy_1)z &= -qz^2. \end{aligned}$$

Tím jsme Riccatiovu rovnici převedli na rovnici Bernoulliovu, kterou jsme se naučili řešit výše. Protože jsme od Riccatiovy rovnice přešli k Bernoulliově rovnici jen přičtením diferencovatelné funkce, je zřejmé, že pokud nalezneme nějaké řešení Riccatiovy rovnice, existence a jednoznačnost řešení příslušné počáteční úlohy je potom ekvivalentní existenci a jednoznačnosti získané Bernoulliovy rovnice na příslušném intervalu (musí se ale přepočítat počáteční podmínka).

Poznámka 10.4.21. Řešení y_1 se obvykle hledá metodou uhodnutí kombinovanou s metodou neurčitých koeficientů.

Příklad 10.4.22. Na intervalu $(0, \infty)$ řešme rovnici

$$y' + \frac{1}{x}y + y^2 = \frac{4}{x^2}.$$

Pokud se pokusíme typnout řešení ve tvaru $y_1 := ax + b$, dostáváme

$$a + a + \frac{b}{x} + a^2x^2 + 2abx + b^2 = \frac{4}{x^2},$$

což je rovnost, kterou nemůžeme splnit žádnou volbou koeficientů, neboť žádný člen levé strany není násobkem $\frac{1}{x^2}$.

Pokud zkusíme volbu $y_1 := \frac{a}{x} + b$, dostáváme

$$-\frac{a}{x^2} + \frac{a}{x^2} + \frac{b}{x} + \frac{a^2}{x^2} + \frac{2ab}{x} + b^2 = \frac{4}{x^2},$$

čemuž vyhovíme například volbou $a = 2$ a $b = 0$ (šlo by i $a = -2$ a $b = 0$, odpovídající řešení získáme spolu se ostatními pomocí Bernoulliovy rovnice). Dostáváme tedy Bernoulliovu rovnici (píšeme $y_1 := \frac{2}{x}$ a rovnou dosazujeme do vzorce získaného v obecném případě)

$$z' + \Big(\frac{1}{x} + 2 \cdot 1 \cdot \frac{2}{x}\Big)z = -z^2 \qquad \iff \qquad z' + \frac{5}{x}z = -z^2.$$

Tuto Bernoulliovu rovnici řešme na množinách $(0,\infty) \times (-\infty, 0)$ a $(0,\infty) \times (0,\infty)$ (vzhledem k zadání příkladu), kde dostáváme pro $w := \frac{1}{z}$

$$\frac{-z'}{z^2} - \frac{5}{x}\frac{1}{z} = 1 \quad \iff \quad w' - \frac{5}{x}w = 1 \quad \iff \quad \Big(\frac{1}{x^5}w\Big)' = \frac{1}{x^5}w' + \frac{-5}{x^6}w = \frac{1}{x^5}.$$

Odtud

$$w = x^5 \int \frac{1}{x^5}\,\mathrm{d}x = x^5\Big(-\frac{1}{4x^4} + C\Big) = -\frac{x}{4} + Cx^5.$$

Proto
$$z = \frac{1}{-\frac{x}{4} + Cx^5} = \frac{-4}{x - 4Cx^5}.$$
Tato řešení máme jen na intervalech, kde je výraz $x - 4Cx^5$ nenulový. Navíc máme ještě triviální řešení $z \equiv 0$. Konečně dostáváme obecné řešení Riccatiovy rovnice
$$y = \frac{-4}{x - 4Cx^5} + \frac{2}{x} \qquad \text{respektive} \qquad y = \frac{2}{x}.$$

Příklad 10.4.23. Na $\mathbb{R}$ řešme rovnici
$$y' = y^2 - (2x+1)y + (1 + x + x^2).$$
Zkusme hledat řešení y_1 ve tvaru $y_1 := ax + b$. Pak máme
$$\begin{aligned} a &= (ax+b)^2 - (2x+1)(ax+b) + 1 + x + x^2 \\ &= a^2x^2 + 2abx + b^2 - (2ax^2 + 2bx + ax + b) + 1 + x + x^2 \\ &= (a^2 - 2a + 1)x^2 + (2ab - 2b - a + 1)x + b^2 - b + 1. \end{aligned}$$
Porovnání koeficientů u x^2 dává $a = 1$. Dále pozorujeme, že buď $b = 0$ nebo $b = 1$. Položme tedy třeba $y_1 := x$ a následně $z := y - x$. Pak po úpravě dostáváme rovnici
$$z' + z = z^2.$$
Tuto Bernoulliovu rovnici řešme na množinách $\mathbb{R} \times (-\infty, 0)$ a $\mathbb{R} \times (0, \infty)$, kde dostáváme pro $w := \frac{1}{z}$
$$\frac{-z'}{z^2} - \frac{1}{z} = -1 \quad \Longleftrightarrow \quad w' - w = -1 \quad \Longleftrightarrow \quad \left(\mathrm{e}^{-x}w\right)' = \mathrm{e}^{-x}w' - \mathrm{e}^{-x}w = -\mathrm{e}^{-x}.$$
Odtud
$$w = \mathrm{e}^x \int (-\mathrm{e}^{-x})\,\mathrm{d}x = \mathrm{e}^x\left(\mathrm{e}^{-x} + C\right) = 1 + C\mathrm{e}^x.$$
Proto
$$z = \frac{1}{1 + C\mathrm{e}^x} \qquad a \qquad y = x + \frac{1}{1 + C\mathrm{e}^x}$$
na intervalech, kde $1 + C\mathrm{e}^x \neq 0$. Triviální řešení Bernoulliovy rovnice $z \equiv 0$ ještě dává řešení $y = x$, ale to jsme již znali.

10.5 Lineární rovnice n-tého řádu

Nechť $f, a_0, a_1, \dots, a_n \in C((a,b))$. Rovnice
$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y = f(x)$$
se nazývá *lineární rovnice n-tého řádu.* Funkcím $a_0, \dots, a_n$ říkáme *koeficienty* a f je *pravá strana.* Pro jednoduchost zápisu budeme v dalším používat značení
$$Ly := a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \dots + a_1(x)y' + a_0(x)y.$$

Naše úloha má tedy tvar $Ly = f$. Rovnice

$$a_n(x)y^{(n)} + a_{n-1}(x)y^{(n-1)} + \cdots + a_1(x)y' + a_0(x)y = 0 \qquad \text{neboli} \qquad Ly = 0$$

se nazývá *homogenní rovnice* příslušející k $Ly = f$.

Základní existenční výsledek je vybudován na Picard–Lindelöfově existenční větě (Věta 10.3.5).

Věta 10.5.1 (Globální existence a jednoznačnost pro rovnici n-tého řádu)**.** *Nechť $f, a_0, a_1, \ldots, a_n \in C((a,b))$ a $a_n \neq 0$ na (a,b). Pak pro každé $x_0 \in (a,b)$ a každé $(y_0, y_1, \ldots, y_{n-1}) \in \mathbb{R}^n$ existuje jednoznačné řešení rovnice $Ly = f$ splňující*

$$y(x_0) = y_0, \qquad y'(x_0) = y_1, \qquad \ldots, \qquad y^{(n-1)}(x_0) = y_{n-1}.$$

Důkaz. Podělíme-li rovnici $Ly = f$ nenulovým výrazem $a_n(x)$ a přepíšeme-li si výsledek jako soustavu rovnic prvního řádu (používáme postup ze začátku kapitoly), dostáváme soustavu $u' = G(x,u)$ ve tvaru

$$\begin{aligned} u_1' &= u_2 \\ u_2' &= u_3 \\ &\vdots \\ u_{n-1}' &= u_n \\ u_n' &= -\frac{a_0(x)}{a_n(x)}u_1 - \frac{a_1(x)}{a_n(x)}u_2 - \ldots - \frac{a_{n-1}(x)}{a_n(x)}u_n + \frac{f(x)}{a_n(x)} \end{aligned} \tag{10.5.1}$$

s počáteční podmínkou

$$u_1(x_0) = y_0, \qquad u_2(x_0) = y_1, \qquad \ldots, \qquad u_n(x_0) = y_{n-1}. \tag{10.5.2}$$

Tato úloha splňuje předpoklady Picard–Lindelöfovy existenční věty a dostáváme lokální existenci a jednoznačnost na $(x_0-\delta, x_0+\delta)$ pro jisté $\delta > 0$. Zbývá ukázat, že řešení dokážeme jednoznačně prodloužit na celé (a,b). Prodloužení zkonstruujeme v několika krocích. Podrobný důkaz provedeme jen pro prodloužení doleva do bodu a, pro prodloužení do bodu b se postupuje analogicky.

Předpokládejme, že máme řešení na intervalu $(z, x_0 + \delta)$, kde $a < z \leq x_0 - \delta$ (v první fázi pracujeme s $z := x_0 - \delta$) a $x_0 + \delta < b$.

Krok 1: omezenost řešení a jeho derivací na $(z, x_0 + \delta)$.
Nejprve odhadneme eukleidovskou velikost vektoru $u = (u_1, \ldots, u_n)$ pro $x \in (z, x_0 + \delta)$. Pro derivaci její druhé mocniny ze soustavy (10.5.1) vyčteme rovnost, na kterou následně použijeme Youngovu nerovnost spolu s omezeností spojitých

koeficientů a_j a pravé strany f na omezeném uzavřeném intervalu $[z, x_0 + \delta]$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}|u|^2 &= \frac{\mathrm{d}}{\mathrm{d}x}\sum_{i=1}^{n} u_i^2 = 2\sum_{i=1}^{n} u_i u_i' \\ &= 2\sum_{i=1}^{n-1} u_i u_{i+1} - 2\sum_{j=0}^{n-1} \frac{a_j(x)}{a_n(x)} u_{j+1} u_n + 2\frac{f(x)}{a_n(x)} u_n \\ &\le \sum_{i=1}^{n-1} (u_i^2 + u_{i+1}^2) + \sum_{j=0}^{n-1} \max_{[z,x_0+\delta]} \Big|\frac{a_j(x)}{a_n(x)}\Big| (u_{j+1}^2 + u_n^2) \\ &\qquad + \max_{[z,x_0+\delta]} \frac{1}{|a_n(x)|}\Big(\max_{[z,x_0+\delta]} f^2 + u_n^2\Big) \\ &\le C_1|u|^2 + C_2,\end{aligned}$$

kde $C_1, C_2 > 0$. Proto také platí pro všechna $x \in (z, x_0 + \delta)$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\Big(\mathrm{e}^{-C_1(x-x_0)}|u|^2\Big) &= \mathrm{e}^{-C_1(x-x_0)}\frac{\mathrm{d}}{\mathrm{d}x}|u|^2 - \mathrm{e}^{-C_1(x-x_0)} C_1|u|^2 \\ &= \mathrm{e}^{-C_1(x-x_0)}\Big(\frac{\mathrm{d}}{\mathrm{d}x}|u|^2 - C_1|u|^2\Big) \\ &\le C_2 \mathrm{e}^{-C_1(x-x_0)} \le C_3.\end{aligned}$$

Odtud

$$\begin{aligned}|u(x)|^2 &= \mathrm{e}^{C_1(x-x_0)}|u(x_0)|^2 + \mathrm{e}^{C_1(x-x_0)}\Big(\mathrm{e}^{-C_1(x-x_0)}|u(x)|^2 - |u(x_0)|^2\Big) \\ &= \mathrm{e}^{C_1(x-x_0)}|u(x_0)|^2 + \mathrm{e}^{C_1(x-x_0)}\Big(\int_{x_0}^{x} \frac{\mathrm{d}}{\mathrm{d}t}\Big(\mathrm{e}^{-C_1(t-x_0)}|u|^2(t)\Big)\,\mathrm{d}t\Big) \\ &\le C + C\Big|\int_{x_0}^{x} C_3\,\mathrm{d}t\Big| \le C.\end{aligned}$$

Vektor $u(x)$ má díky tomu omezené složky na $(z, x_0 + \delta)$. Následně jsou zde také omezené derivace jednotlivých složek (máme omezené pravé strany v (10.5.1)). Z toho také plyne, že řešení je lipschitzovské na $(z, x_0 + \delta)$.

Krok 2: limity v krajním bodě z.

Zafixujme $i \in \{1, \dots, n\}$. Zvolme posloupnost $\{x_j\} \subset (z, x_0 + \delta)$ takovou, aby $x_j \to z$. Protože funkce u_i je podle předchozího kroku omezená na $(z, x_0 + \delta)$, je omezená i posloupnost $\{u_i(x_j)\}$. Po přechodu k podposloupnosti dostáváme $U_i \in \mathbb{R}$ tak, že $u_i(x_{j_k}) \to U_i$. Ke zvolenému $\varepsilon > 0$ díky lipschitzovskosti u_i (konstantu lipschitzovskosti označme L) snadno obdržíme pro $k \in \mathbb{N}$ dost velké a $x \in (z, x_0+\delta)$ dost blízko k z

$$|u_i(x)-U_i| \le |u_i(x)-u_i(x_{j_k})|+|u_i(x_{j_k})-U_i| \le L|x-x_{j_k}|+|u_i(x_{j_k})-U_i| \le L\varepsilon+\varepsilon.$$

Odtud

$$\lim_{x\to z_+} u_i(x) = U_i.$$

Krok 3: platnost (10.5.1) v krajním bodě z.
Podle předchozího kroku lze vektorovou funkci u spojitě dodefinovat v bodě z. Ukažme, že pro takto dodefinovanou funkci je soustava rovnic (10.5.1) splněna i v bodě z, uvažujeme-li jednostranné derivace (z vnitřní strany intervalu). To však okamžitě plyne z Věty o limitě derivací (Věta 6.3.9) a spojitosti jednotlivých složek vektorové funkce u na $[z, x_0+\delta)$ a spojitosti koeficientů a pravé strany.

Krok 4: prodloužení řešení za bod z.
Uvažujme soustavu rovnic (10.5.1) tentokrát s počáteční podmínkou odpovídající našim výsledkům ze druhého kroku

$$u_1(z) = U_1, \qquad u_2(z) = U_2, \qquad \dots, \qquad u_n(z) = U_n.$$

Podle Picard–Lindelöfovy existenční věty má nová úloha jednoznačné řešení na jistém $(z-\tau, z+\tau)$. Toto řešení se na množině $[z, z+\tau) \cap [z, x_0+\delta)$ musí shodovat s řešením, o kterém jsme hovořili v předchozích krocích (kdyby tomu tak nebylo, mohli bychom díky třetímu kroku v bodě z slepit naše původní řešení s restrikcí nového řešení na $(z-\tau, z]$ a byla by porušena jednoznačnost z Picard–Lindelöfovy existenční věty). Umíme tedy prodloužit řešení za bod z.

Krok 5: prodloužení řešení na $(a, x_0+\delta)$.
Definujme množinu intervalů

$$\mathcal{M} := \{(\alpha, x_0+\delta) \subset (a,b) \colon \text{na } (\alpha, x_0+\delta) \text{ existuje řešení (10.5.1) a (10.5.2)}\}.$$

Množina $\mathcal{M}$ je neprázdná, neboť $(x_0-\delta, x_0+\delta) \in \mathcal{M}$. Dále z Picard–Lindelöfovy existenční věty plyne, že kdykoliv $(\alpha_1, x_0+\delta), (\alpha_2, x_0+\delta) \in \mathcal{M}$, jim odpovídající řešení se shodují na průniku těchto intervalů (v Kroku 4 jsme provedli podrobné zdůvodnění v analogické situaci). Definujme

$$A := \inf_{(\alpha, x_0+\delta) \in \mathcal{M}} \alpha.$$

Zřejmě platí $a \le A \le x_0 - \delta$. Ukažme, že platí $A = a$. Pokud by tomu tak nebylo, využili bychom toho, že z definice infima umíme zkonstruovat posloupnost $\{\alpha_n\} \subset (A, x_0+\delta)$ takovou, že $\alpha_n \to A$ a $(\alpha_n, x_0+\delta) \in \mathcal{M}$. Díky tomu a jednoznačnosti zmíněné výše bychom měli řešení úlohy (10.5.1) a (10.5.2) definované na intervalu $(A, x_0+\delta)$ (skutečně, každý bod $x \in (A, x_0+\delta)$ leží v některém z intervalů $(\alpha_n, x_0+\delta)$). Ale v takové situaci bychom mohli použít Kroky 1 až 4 na prodloužení našeho řešení až za bod A (opět se využije toho, že $[A, x_0+\delta]$ je omezený uzavřený interval). Tím ale dostáváme spor s definicí A.

Proto $A = a$ a výše uvedenou konstrukcí využívající posloupnost $\{\alpha_n\} \subset (A, x_0+\delta)$ splňující $\alpha_n \to A$ dostaneme řešení (10.5.1) a (10.5.2) na $(a, x_0+\delta)$. Prodloužení řešení na celé (a,b) se provede analogicky. □

Tvrzení 10.5.2 (Tvar řešení při nulových datech)**.** *Jestliže pravá strana splňuje $f \equiv 0$ na (a,b) a pro počáteční podmínky platí $y_0 = \dots = y_{n-1} = 0$, pak řešení úlohy $Ly = f$ platí $y \equiv 0$ na (a,b).*

Důkaz. Zřejmě identicky nulová funkce je řešením s předepsanými vlastnostmi. Dále jednoznačnost daná předchozí větou zaručuje, že jiné řešení není. □

Podobně jako u lineárních rovnic druhého řádu s konstantními koeficienty, s nimiž jsme se seznámili v kapitole o primitivní funkci, i zde budeme hledat všechna řešení rovnice $Ly = f$ za pomoci linearity diferenciálního operátoru L. Opět nalezneme všechna řešení y_h jednodušší *homogenní rovnice* $Ly = 0$ a pak budeme hledat jedno *partikulární* řešení y_p původní rovnice $Ly = f$. Celkově pak předpis $y = y_h + y_p$, kde y_h probíhá všechna řešení homogenní rovnice, bude dávat všechna řešení rovnice $Ly = f$.

10.5.1 Homogenní rovnice: obecné výsledky

V dalším se budeme snažit nalézt všechna řešení homogenní úlohy $Ly = 0$. Tato řešení díky linearitě operátoru L tvoří vektorový prostor (podprostor prostoru $C^n((a, b))$). Strukturu na tomto prostoru nám dá vhodná definice lineární nezávislosti.

Definice 10.5.3 (Lineární nezávislost funkcí)**.** Řekneme, že $u_1, \dots, u_n \colon \mathbb{R} \to \mathbb{R}$ definované na $(a, b) \subset \mathbb{R}$ jsou *lineárně nezávislé* na (a, b), jestliže pro každou n-tici $(\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$ platí

$$\sum_{i=1}^{n} \alpha_i u_i \equiv 0 \qquad \text{na } (a, b) \qquad \Longrightarrow \qquad \alpha_1 = \dots = \alpha_n = 0.$$

V opačném případě říkáme, že $u_1, \dots, u_n$ jsou *lineárně závislé* na (a, b).

Věta 10.5.4 (O prostoru řešení homogenní rovnice)**.** *Množina všech řešení homogenní rovnice $Ly = 0$ tvoří n-dimenzionální podprostor prostoru $C^n((a, b))$.*

Důkaz. Každé řešení je podle definice pojmu řešení n-krát diferencovatelné. Navíc přepis rovnice $Ly = 0$ do tvaru

$$y^{(n)}(x) = -\sum_{i=0}^{n-1} \frac{a_i(x)}{a_n(x)} y^{(i)}(x)$$

zaručuje, že dokonce máme $y \in C^n((a, b))$.

Z linearity operátoru L navíc plyne, že řešení tvoří podprostor $C^n((a, b))$. Zbývá zjistit jeho dimenzi.

Zafixujme libovolné $x_0 \in (a, b)$. Podle Věty o globální existenci a jednoznačnosti pro rovnici n-tého řádu (Věta 10.5.1) existují funkce $u_0, \dots, u_{n-1} \in C^n((a, b))$ takové, že u_i splňuje úlohu $Ly = 0$ s počátečními podmínkami

$$u_i^{(j)}(x_0) = \delta_{ij}$$

(δ_{ij} je *Kroneckerovo delta*). Tato řešení jsou lineárně nezávislá, neboť pokud máme

$$\sum_{i=1}^{n} \alpha_i u_i(x) \equiv 0 \qquad \text{na } (a, b),$$

nutně totéž platí pro všechny derivace funkce nalevo a po dosazení $x = x_0$ z počátečních podmínek postupně získáme $\alpha_1 = \cdots = \alpha_n = 0$. Odtud vidíme, že dimenze prostoru řešení je alespoň n. Ukažme ještě, že každé řešení úlohy $Ly = 0$ je lineární kombinací funkcí $u_0, \ldots, u_{n-1}$. Nechť tedy y řeší $Ly = 0$ na (a, b). Pro $i = 0, \ldots, n-1$ definujme $\beta_i = y^{(i)}(x_0)$. Pak funkce

$$w := y - \sum_{i=0}^{n-1} \beta_i u_i$$

je lineární kombinací řešení homogenní rovnice, a proto ji řeší také. Dále v bodě x_0 platí

$$w^{(j)}(x_0) = 0 \qquad \text{pro všechna } j \in \{0, \ldots, n-1\}.$$

Odtud podle Tvrzení o tvaru řešení při nulových datech (Tvrzení 10.5.2) dostáváme $w \equiv 0$ na (a, b), tedy $y = \sum_{i=0}^{n-1} \beta_i u_i$ na (a, b). □

Postupně si odvodíme další výsledky o prostoru řešení homogenní rovnice. K tomu potřebujeme ještě zadefinovat dva nové pojmy.

Definice 10.5.5 (Fundamentální systém)**.** Množina $u_1, \ldots, u_n$ se nazývá *fundamentální systém* rovnice $Ly = 0$ na (a, b), jestliže funkce $u_1, \ldots, u_n$ řeší $Ly = 0$ na (a, b) a jsou zde lineárně nezávislé.

Definice 10.5.6 (Wronskián)**.** *Wronského determinant* (častěji zkráceně *wronskián*) funkcí $u_1, \ldots, u_n \in C^{(n-1)}((a, b))$ v bodě $x \in (a, b)$ je

$$W_{[u_1, u_2, \ldots, u_n]}(x) := \det \begin{pmatrix} u_1(x) & u_2(x) & \cdots & u_n(x) \\ u_1'(x) & u_2'(x) & \cdots & u_n'(x) \\ \vdots & \vdots & \ddots & \vdots \\ u_1^{(n-1)}(x) & u_2^{(n-1)}(x) & \cdots & u_n^{(n-1)}(x) \end{pmatrix}.$$

Výše uvedená matice se nazývá *Wronského matice.*

Věta 10.5.7 (Obecný vztah lineární závislosti a wronskiánu)**.** *Jsou-li* $u_1, \ldots, u_n \in C^{(n-1)}((a, b))$ *lineárně závislé na* (a, b)*, pak* $W_{[u_1, u_2, \ldots, u_n]} \equiv 0$ *na* (a, b)*.*

Důkaz. Lineární závislost implikuje, že matice z definice wronskiánu má lineárně závislé sloupce, a proto je wronskián nulový. □

Obrácená implikace v předchozí větě obecně neplatí. Skutečně, funkce

$$u_1(x) = \begin{cases} 0 & \text{pro } x \in (-\infty, 0] \\ x^3 & \text{pro } x \in (0, \infty) \end{cases} \qquad \text{a} \qquad u_2(x) = \begin{cases} x^3 & \text{pro } x \in (-\infty, 0] \\ 0 & \text{pro } x \in (0, \infty) \end{cases}$$

jsou lineárně nezávislé na $\mathbb{R}$, ale snadno se spočítá, že $W_{[u_1, u_2]} \equiv 0$ na $\mathbb{R}$. Všimněte si, že obě funkce jsou třídy $C^2(\mathbb{R})$.

Na druhou stranu, řešení rovnice $Ly = 0$ jsou velmi speciální funkce, pro které zmíněná obrácená implikace platí vždy. Dokonce stačí ověřovat nenulovost wronskiánu v jediném bodě. Pro získání tohoto výsledku potřebujeme ještě jedno pomocné tvrzení.

Lemma 10.5.8 (Derivace wronskiánu)**.** *Nechť* $u_1, \dots, u_n$ *řeší* $Ly = 0$ *na* (a, b). *Označme* $W(x) = W_{[u_1, u_2, \dots, u_n]}(x)$. *Pak*

$$W'(x) = -\frac{a_{n-1}(x)}{a_n(x)} W(x) \qquad \text{na } (a, b).$$

Speciálně, pro libovolná $x_0, x \in (a, b)$ *máme*

$$W(x) = W(x_0) \mathrm{e}^{-\int_{x_0}^{x} \frac{a_{n-1}(t)}{a_n(t)} \,\mathrm{d}t}.$$

Důkaz. Platí následující rovnosti, které si zdůvodníme pod výpočtem

$$\begin{aligned}
W' &= \frac{\mathrm{d}}{\mathrm{d}x} \det \begin{pmatrix} u_1 & \cdots & u_n \\ u_1' & \cdots & u_n' \\ \vdots & \ddots & \vdots \\ u_1^{(n-1)} & \cdots & u_n^{(n-1)} \end{pmatrix} \\
&= \det \begin{pmatrix} u_1' & \cdots & u_n' \\ u_1' & \cdots & u_n' \\ \vdots & \ddots & \vdots \\ u_1^{(n-1)} & \cdots & u_n^{(n-1)} \end{pmatrix} + \det \begin{pmatrix} u_1 & \cdots & u_n \\ u_1'' & \cdots & u_n'' \\ \vdots & \ddots & \vdots \\ u_1^{(n-1)} & \cdots & u_n^{(n-1)} \end{pmatrix} \\
&\qquad + \cdots + \det \begin{pmatrix} u_1 & \cdots & u_n \\ u_1' & \cdots & u_n' \\ \vdots & \ddots & \vdots \\ u_1^{(n)} & \cdots & u_n^{(n)} \end{pmatrix} \\
&= \det \begin{pmatrix} u_1 & \cdots & u_n \\ u_1' & \cdots & u_n' \\ \vdots & \ddots & \vdots \\ -\frac{1}{a_n} \sum_{i=0}^{n-1} a_i u_1^{(i)} & \cdots & -\frac{1}{a_n} \sum_{i=0}^{n-1} a_i u_n^{(i)} \end{pmatrix} \\
&= \det \begin{pmatrix} u_1 & \cdots & u_n \\ u_1' & \cdots & u_n' \\ \vdots & \ddots & \vdots \\ -\frac{a_{n-1}}{a_n} u_1^{(n-1)} & \cdots & -\frac{a_{n-1}}{a_n} u_n^{(n-1)} \end{pmatrix} = -\frac{a_{n-1}}{a_n} W.
\end{aligned}$$

Nejprve jsme si uvědomili, že při počítání determinantu se sčítají členy obsahující po jednom prvku z každého řádku. To nám s ohledem na pravidla derivování součinu více funkcí dává první rovnost. Získali jsme tím součet determinantů, v němž všechny matice s výjimkou té poslední obsahují dvojici totožných řádků, což má za následek nulovost odpovídajícího determinantu. V jediném zbývajícím determinantu jsme na posledním řádku využili toho, že jednotlivé funkce u_i řeší $Lu_i = 0$. Zároveň z předpisu pro operátor L zjišťujeme, že poslední řádek je součtem násobků řádků předchozích (což opět vede na nulový příspěvek do determinantu) a členů s $(n-1)$-tými derivacemi funkcí u_i. Celkově jsme tedy dostali násobek wronskiánu. Speciální část lemmatu plyne z první části po přenásobení integračním faktorem $\mathrm{e}^{\int_{x_0}^{x} \frac{a_{n-1}(t)}{a_n(t)}\,\mathrm{d}t}$. □

Poznámka 10.5.9. Ze speciální části předchozího lemmatu okamžitě plyne, že pro n-tici řešení rovnice $Ly = 0$ se wronskián buď rovná nule ve všech bodech intervalu (a,b) a nebo je ve všech bodech intervalu (a,b) nenulový.

Věta 10.5.10 (Vztah lineární závislosti a wronskiánu pro řešení)**.** *Nechť funkce $u_1,\dots,u_n$ řeší $Ly = 0$ na (a,b). Pak $u_1,\dots,u_n$ jsou lineárně nezávislá právě tehdy, když existuje $x_0 \in (a,b)$ splňující $W_{[u_1,u_2,\dots,u_n]}(x_0) \neq 0$ na (a,b).*

Důkaz. Dle Věty o obecném vztahu lineární závislosti a wronskiánu (Věta 10.5.10) implikuje lineární závislost $u_1,\dots,u_n$ nulovost wronskiánu ve všech bodech. S ohledem na předchozí poznámku tedy zbývá ukázat, že pokud $W_{[u_1,u_2,\dots,u_n]}(x_0) = 0$ pro nějaké $x_0 \in (a,b)$, pak $u_1,\dots,u_n$ jsou lineárně závislá.

Podmínka $W_{[u_1,u_2,\dots,u_n]}(x_0) = 0$ však znamená, že v bodě x_0 má Wronského matice lineárně závislé sloupce, a proto existuje netriviální n-tice $(\alpha_1,\dots,\alpha_n) \in \mathbb{R}^n$ tak, že

$$\sum_{i=1}^{n} \alpha_i u_i^{(k)}(x_0) = 0 \qquad \text{pro všechna } k = 0,1,\dots,n-1.$$

Pokud tedy definujeme $u := \sum_{i=1}^{n} \alpha_i u_i$, díky linearitě L je u řešením úlohy $Ly = 0$ se sadou počátečních podmínek

$$y(x_0) = 0, \qquad y'(x_0) = 0, \quad \dots, \quad y^{(n-1)}(x_0) = 0.$$

Tuto úlohu zároveň zřejmě řeší i $y \equiv 0$. Díky globální jednoznačnosti řešení proto dostáváme

$$\sum_{i=1}^{n} \alpha_i u_i \equiv 0 \qquad \text{na } (a,b)$$

a jsme hotovi. □

Naše výsledky týkající se vlastností wronskiánu se dají využívat hned několika způsoby. Jednak jde o ověření lineární nezávislosti sady řešení homogenní rovnice získané kupříkladu uhodnutím. Druhou možnou aplikací je doplnění chybějících funkcí do neúplné sady. Třetí aplikací bude zanedlouho představená metoda *variace konstant* pro hledání partikulárního řešení nehomogenní rovnice. Čtvrtou aplikací je určení konstant v předpisu pro obecné řešení, abychom zajistili splnění počátečních podmínek.

Příklad 10.5.11. V kapitole o primitivních funkcích jsme se krátce zabývali lineárními rovnicemi druhého řádu s konstantními koeficienty. Představili jsme si tři možné dvojice řešení homogenní úlohy v závislosti na kořenech charakteristického polynomu. Jednalo se o $\{\mathrm{e}^{\lambda_1 x}, \mathrm{e}^{\lambda_2 x}\}$ pro $\lambda_1, \lambda_2 \in \mathbb{R}$ různé, $\{\mathrm{e}^{\lambda x}, x\mathrm{e}^{\lambda x}\}$ pro $\lambda \in \mathbb{R}$ a $\{\mathrm{e}^{\mu x}\cos\nu x, \mathrm{e}^{\mu x}\sin\nu x\}$ pro $\mu, \nu \in \mathbb{R}$. Snadno se ověří, že každá z těchto dvojic má nenulový wronskián, jedná se tedy pokaždé o fundamentální systém.

Příklad 10.5.12. (i) Předpokládejme, že se nám k rovnici

$$y'' + p(x)y' + q(x)y = 0,$$

kde $p, q \in C((a,b))$, podařilo nalézt jedno řešení $u \in C^2((a,b))$, které splňuje $u \neq 0$ na (a,b). Podle Věty o prostoru řešení homogenní rovnice (Věta 10.5.4) pak musí existovat ještě jedno lineárně nezávislé řešení v. Pak máme podle Lemmatu o derivaci wronskiánu (Lemma 10.5.8), kde můžeme předpokládat, že $W(x_0) = 1$ (neboť libovolný nenulový násobek řešení homogenní rovnice je rovněž řešením),

$$\mathrm{e}^{-\int_{x_0}^{x} p(t)\,\mathrm{d}t} = W(x_0)\mathrm{e}^{-\int_{x_0}^{x} p(t)\,\mathrm{d}t} = W(x) = uv' - u'v.$$

Odtud díky nenulovosti u dostáváme

$$\left(\frac{v}{u}\right)' = \frac{v'u - vu'}{u^2} = \frac{1}{u^2}\mathrm{e}^{-\int_{x_0}^{x} p(t)\,\mathrm{d}t}.$$

Proto

$$v(x) = u(x)\int \frac{1}{u^2(x)}\mathrm{e}^{-\int_{x_0}^{x} p(t)\,\mathrm{d}t}\,\mathrm{d}x.$$

Poznamenejme, že aditivní konstanta z integrace napravo není podstatná, neboť se ve výsledku projeví jako násobek nám již známé funkce u.

(ii) Řešme konkrétně rovnici

$$y'' + \frac{2}{x}y' + y = 0 \qquad \text{na } (0,\infty)$$

a předpokládejme, že se nám nějak podařilo uhodnout, že řešením je $u(x) = \frac{\sin x}{x}$ (ověřte dosazením). Pak předchozí obecný postup dává (vezmeme $x_0 = 1$ a nejjednodušší aditivní konstantu) na intervalu $(0,\pi)$

$$\begin{aligned} v(x) &= u(x)\int \frac{1}{u^2(x)}\mathrm{e}^{-\int_{x_0}^{x} p(t)\,\mathrm{d}t}\,\mathrm{d}x = \frac{\sin x}{x}\int \frac{x^2}{\sin^2 x}\mathrm{e}^{-\int_1^x \frac{2}{t}\,\mathrm{d}t}\,\mathrm{d}x \\ &= \frac{\sin x}{x}\int \frac{x^2}{\sin^2 x}\mathrm{e}^{-2[\log t]_1^x}\,\mathrm{d}x = \frac{\sin x}{x}\int \frac{x^2}{\sin^2 x}\frac{1}{x^2}\,\mathrm{d}x \\ &= \frac{\sin x}{x}(-\cot x) = -\frac{\cos x}{x}. \end{aligned}$$

Díky linearitě L není znaménko podstatné. Snadno ověříme, že v je řešením naší rovnice na $(0,\pi)$. Nenulovost wronskiánu na $(0,\pi)$ implikuje jeho nenulovost na $(0,\infty)$. Jedním z možných fundamentálních systémů je tedy $\{\frac{\sin x}{x}, \frac{\cos x}{x}\}$.

(iii) U lineárních rovnic druhého řádu s konstantními koeficienty je v případě rovnice s dvojnásobným kořenem charakteristického polynomu, tedy

$$y'' - 2ay' + a^2y = 0,$$

kde $a \in \mathbb{R}$, fundamentální systém tvořen funkcemi e^{ax} a $x\mathrm{e}^{ax}$. První z těchto funkcí jsme v kapitole o primitivní funkci získali pokusným dosazením funkce $x \mapsto \mathrm{e}^{\lambda x}$,

$\lambda \in \mathbb{R}$. Druhou funkci jsme tenkrát prozradili a ověřili, že je řešením, ale neřekli, jak se na ni přijde. K tomu lze užít předchozí metodu. Při volbě $u(x) = \mathrm{e}^{ax}$ totiž dostáváme

$$\begin{aligned} v(x) = u(x) \int \frac{1}{u^2(x)} \mathrm{e}^{-\int_{x_0}^{x} p(t)\,\mathrm{d}t}\,\mathrm{d}x &= \mathrm{e}^{ax} \int \mathrm{e}^{-2ax} \mathrm{e}^{\int_{x_0}^{x} 2a\,\mathrm{d}t}\,\mathrm{d}x \\ = \mathrm{e}^{ax} \int \mathrm{e}^{-2ax_0}\,\mathrm{d}x = \mathrm{e}^{-2ax_0} x \mathrm{e}^{ax}. \end{aligned}$$

Tedy až na (nepodstatnou) multiplikativní konstantu jsme dostali nám známý výsledek.
(iv) V případě rovnice $y'' + 1 = 0$ a znalosti řešení $u(x) = \cos x$, dostáváme jako druhou funkci

$$\begin{aligned} v(x) = u(x) \int \frac{1}{u^2(x)} \mathrm{e}^{-\int_{x_0}^{x} p(t)\,\mathrm{d}t}\,\mathrm{d}x &= \cos x \int \frac{1}{\cos^2 x} \mathrm{e}^{-\int_{x_0}^{x} 0\,\mathrm{d}t}\,\mathrm{d}x \\ &= \cos x \int \frac{1}{\cos^2 x}\,\mathrm{d}x = \cos x \tan x = \sin x. \end{aligned}$$

Výsledek jsme získali na intervalech, kde $\cos \neq 0$. Pak použijeme lepení, či ověříme přímo zderivováním, že jsme získali opět nám známý fundamentální systém.

Cvičení 10.5.13. Rovnici $x^2y'' - 4xy' + 6y = 0$ řeší lineárně nezávislé funkce $y_1(x) = x^2$ a $y_2(x) = x^3$. Jim odpovídající wronskián je však v počátku nulový. Proč tento jev není ve sporu s výše získanými výsledky?

10.5.2 Variace konstant

Metoda variace konstant, se kterou jsme se setkali u lineárních rovnic druhého řádu s konstantními koeficienty v kapitole o primitivních funkcích a také u lineárních rovnic prvního řádu, umožňuje najít partikulární řešení i v případě naší úlohy $Ly = f$ (poznamenejme, že tato úloha v sobě zahrnuje oba zmíněné případy). Předpokládejme, že máme fundamentální systém $\{u_1, \ldots, u_n\}$. Řešení homogenní rovnice jsou pak dána předpisem

$$y(x) = \sum_{i=1}^{n} C_i u_i(x),$$

kde $C_1, \ldots, C_n$ jsou reálné konstanty. Metoda *variace konstant* spočívá v tom, že tyto konstanty nahradíme (neznámými) $C^1((a,b))$-funkcemi $c_1(x), \ldots, c_n(x)$ (na následujícím postupu bude vidět, že skutečně nepotřebujeme vyšší derivace než první). U lineární rovnice prvního řádu jsme viděli, že už toto přepsání náš problém natolik zprůhlednilo, že jsme po dosazení nového tvaru do diferenciální rovnice okamžitě dostali předpis pro $c_1'(x)$. Zde se do podobné situace dostaneme až poté, co si uměle přidáme $n-1$ podmínek, které zvolíme co nejvýhodněji. Postupně derivujme předpis pro y. Dostáváme

$$\begin{aligned}
y(x) &= \sum_{i=1}^{n} c_i u_i \\
y'(x) &= \sum_{i=1}^{n} c_i u_i' + \sum_{i=1}^{n} c_i' u_i & &\text{s podmínkou} & \sum_{i=1}^{n} c_i' u_i &\equiv 0 \\
y''(x) &= \sum_{i=1}^{n} c_i u_i'' + \sum_{i=1}^{n} c_i' u_i' & &\text{s podmínkou} & \sum_{i=1}^{n} c_i' u_i' &\equiv 0 \\
y'''(x) &= \sum_{i=1}^{n} c_i u_i''' + \sum_{i=1}^{n} c_i' u_i'' & &\text{s podmínkou} & \sum_{i=1}^{n} c_i' u_i'' &\equiv 0 \\
&\vdots & & & &\vdots \\
y^{(n-1)}(x) &= \sum_{i=1}^{n} c_i u_i^{(n-1)} + \sum_{i=1}^{n} c_i' u_i^{(n-2)} & &\text{s podmínkou} & \sum_{i=1}^{n} c_i' u_i^{(n-2)} &\equiv 0 \\
y^{(n)}(x) &= \sum_{i=1}^{n} c_i u_i^{(n)} + \sum_{i=1}^{n} c_i' u_i^{(n-1)}.
\end{aligned} \tag{10.5.3}$$

Právě odvozené vztahy pro derivace funkce y dosadíme do rovnice $Ly = f$. Protože funkce u_i splňují $Lu_i = 0$, máme také

$$0 = a_n \sum_{i=1}^{n} c_i u_i^{(n)} + a_{n-1} \sum_{i=1}^{n} c_i u_i^{(n-1)} + \cdots + a_0 \sum_{i=1}^{n} c_i u_i. \tag{10.5.4}$$

Proto se po dosazení odpovídajíci členy na levé straně vyruší a zůstane jen

$$Ly = a_n \sum_{i=1}^{n} c_i' u_i^{(n-1)} = f.$$

Celkově máme soustavu

$$\sum_{i=1}^{n} c_i' u_i \equiv 0, \quad \sum_{i=1}^{n} c_i' u_i' \equiv 0, \quad \ldots, \quad \sum_{i=1}^{n} c_i' u_i^{(n-2)} \equiv 0, \quad \sum_{i=1}^{n} c_i' u_i^{(n-1)} = \frac{f}{a_n}.$$

Jedná se o soustavu reprezentovanou Wronského maticí. Protože wronskián odpovídající fundamentálnímu systému je vždy nenulový, dostáváme pro každé $x \in (a, b)$ jednoznačně danou hodnotu $c_1'(x), \ldots, c_n'(x)$. Aplikujeme-li na soustavu Cramerovo pravidlo, dostáváme

$$c_i' = \frac{1}{W} \det \begin{pmatrix} u_1 & \cdots & u_{i-1} & 0 & u_{i+1} & \cdots & u_n \\ u_1' & \cdots & u_{i-1}' & 0 & u_{i+1}' & \cdots & u_n' \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ u_1^{(n-1)} & \cdots & u_{i-1}^{(n-1)} & \frac{f}{a_n} & u_{i+1}^{(n-1)} & \cdots & u_n^{(n-1)} \end{pmatrix}.$$

Ze spojitosti použitých funkcí vidíme, že funkce $x \mapsto c_i'(x)$ jsou spojité a jsou tedy skutečně derivacemi nějakých $C^1((a,b))$-funkcí c_i splňujících námi požadované podmínky.

Poznámka 10.5.14. (i) Zatímco u lineárních rovnic prvního řádu při variaci konstant dochází k tomu, že se na levé straně rovnice vyruší dva členy ze tří, u rovnic vyššího řádu je do vzájemného vyrušování reprezentovaného formulí (10.5.4) zapojeno často velmi velké množství členů levé strany a je poměrně časově náročné dohledávání vzájemně se vyrušujících skupin. Proto čtenáři silně doporučujeme, aby si buď pamatoval tvar výsledné soustavy nebo princip celého postupu, kdy se ve schématu (10.5.3) (po přenásobení jednotlivých řádků odpovídajícími koeficienty a_i) na pravé straně vyrušily všechny první sumy a druhé sumy zmizely díky námi vytvořeným požadavkům na jejich nulovost. Proto po dosazení do rovnice $Ly = f$ levou stranu zastupuje jen druhá suma z výrazu pro $y^{(n-1)}$ (vynásobená koeficientem a_n).
(ii) Metoda variace konstant bývá časově velmi náročná. Je-li to možné, určitě se vyplatí upřednostnit metodu pro speciální pravou stranu (zatím jsme ji viděli u lineárních rovnic druhého řádu s konstantními koeficienty v kapitole o primitivních funkcích, brzy se ji však naučíme používat i u rovnic vyššího řadu), kde stačí vyřešit jen jedinou soustavu rovnic pro neurčité koeficienty (u variace konstant řešíme různé soustavy rovnic pro každé $x \in (a,b)$ a výsledek navíc musíme integrovat).
(iii) V praxi se při aplikaci metody variace konstant používá Cramerovo pravidlo spíše výjimečně. Většinou se nám podaří vyjádřit derivace funkcí c_i méně pracně (zejména u rovnic s konstantními koeficienty, kde se ve fundamentálním systému vyskytují jen velmi speciální funkce). Kupříkladu u rovnice

$$y'' + y = f(x)$$

je fundamentální systém tvořen funkcemi cos a sin. Proto při variaci konstant přicházíme k soustavě

$$\begin{aligned} c_1'(x)\cos x + c_2'(x)\sin x &= 0 \\ -c_1'(x)\sin x + c_2'(x)\cos x &= f(x). \end{aligned}$$

Vynásobíme-li první řádek výrazem $\sin x$, druhý výrazem $\cos x$ a pak řadky sečteme, máme $c_2'(x) = f(x)\cos x$. Podobně se získá $c_1'(x) = -f(x)\sin x$.

10.5.3 Splnění počátečních podmínek

Podle Věty o globální existenci a jednoznačnosti řešení (Věta 10.5.1) pro rovnici $Ly = f$ víme, že při zadaných počátečních podmínkách

$$y(x_0) = y_0, \qquad y'(x_0) = y_1, \qquad \ldots, \qquad y^{(n-1)}(x_0) = y_{n-1}$$

je řešení jednoznačné. Zde si stručně vysvětlíme, že k jeho určení stačí jen vyřešit soustavu lineárních rovnic s regulární maticí. Skutečně, je-li fundamentální systém tvořen funkcemi $u_1, \ldots, u_n$ a partikulární řešení značíme y_p, pak obecné řešení má tvar

$$y(x) = C_1 u_1(x) + \cdots + C_n u_n(x) + y_p(x) \qquad \text{na } (a, b).$$

Počáteční podmínky nám pak dávají soustavu pro neznámé konstanty $C_1, \ldots, C_n$

$$\begin{array}{ccccccc} C_1 u_1(x_0) & + & \ldots & + & C_n u_n(x_0) & = & y_0 - y_p(x_0) \\ C_1 u_1'(x_0) & + & \ldots & + & C_n u_n'(x_0) & = & y_1 - y_p'(x_0) \\ \vdots & & \ddots & & \vdots & & \vdots \\ C_1 u_1^{(n-1)}(x_0) & + & \ldots & + & C_n u_n^{(n-1)}(x_0) & = & y_0 - y_p^{(n-1)}(x_0). \end{array}$$

Levá strana je reprezentována Wronského maticí v bodě x_0. Tato matice je vždy regulární (odpovídá fundamentálnímu systému).

Příklad 10.5.15. Řešme počáteční úlohu

$$y''' - y'' + y' - y = 1 - x \qquad y(0) = 1, \quad y'(0) = 2, \quad y''(0) = 3.$$

Obecné řešení homogenní rovnice má tvar (srovnejte s Příkladem 10.5.26)

$$y_h = C_1 \mathrm{e}^x + C_2 \cos x + C_3 \sin x \qquad \text{na } \mathbb{R},$$

kde $C_1, C_2, C_3 \in \mathbb{R}$. Partikulární řešení v tomto jednoduchém případě není obtížné uhodnout, lze vzít $y_p = x$. Pak máme obecné řešení

$$y = C_1 \mathrm{e}^x + C_2 \cos x + C_3 \sin x + x \qquad \text{na } \mathbb{R}.$$

Počáteční podmínky dávají soustavu (partikulární řešení přesouváme na pravou stranu)

$$\begin{array}{ccccc} C_1 \mathrm{e}^0 + C_2 \cos 0 + C_3 \sin 0 & = & C_1 + C_2 & = & 1 \\ C_1 \mathrm{e}^0 - C_2 \sin 0 + C_3 \cos 0 & = & C_1 + C_3 & = & 1 \\ C_1 \mathrm{e}^0 - C_2 \cos 0 - C_3 \sin 0 & = & C_1 - C_2 & = & 3. \end{array}$$

Odtud $C_1 = 2$, $C_2 = C_3 = -1$ a dostáváme řešení počáteční úlohy

$$y = 2\mathrm{e}^x - \cos x - \sin x + x \qquad \text{na } \mathbb{R}.$$

10.5.4 Homogenní rovnice s konstantními koeficienty

V případě, že koeficienty a_i, $i = 0, \ldots, n$, jsou konstantní funkce, je známa metoda hledání fundamentálního systému rovnice $Ly = 0$. Tuto metodu si zde předvedeme. V dalším budeme symboly a_i používat k označení konstantní hodnoty koeficientů (tedy a_i pro nás budou reálná čísla, nikoliv funkce).

Zkusme hledat řešení úlohy $Ly = 0$ ve tvaru $y = \mathrm{e}^{\lambda x}$, kde $\lambda \in \mathbb{C}$. Dostáváme

$$\mathrm{e}^{\lambda x} \text{ splňuje } Ly = 0 \qquad \Longleftrightarrow \qquad \sum_{i=0}^{n} a_i \lambda^i = 0.$$

Výraz $p(\lambda) := \sum_{i=0}^n a_i\lambda^i$ (přesněji, funkce $\lambda \mapsto \sum_{i=0}^n a_i\lambda^i$) se nazývá *charakteristický polynom* rovnice $Ly = 0$. V dalším se budeme zabývat otázkou, zda se dá z případné znalosti kořenů charakteristického polynomu vytvořit fundamentální systém řešení. Budeme tedy chtít získat n lineárně nezávislých funkcí splňujících $Ly = 0$.

Poznámka 10.5.16. Naše teorie z předchozích částí kapitoly vychází z Picard–Lindelöfovy existenční věty (Věta 10.3.5), kterou budeme dokazovat pouze v reálném případě. Není tedy jasné, že v komplexním oboru je n správný počet lineárně nezávislých řešení. Nás skutečně budou zajímat jen řešení reálná, ale budeme postupovat tak, že nejprve zkonstruujeme n lineárně nezávislých komplexních řešení a později z nich získáme n lineárně nezávislých reálných řešení.

Poznámka 10.5.17. Nalezneme-li n různých kořenů charakteristického polynomu, není těžké přímo dokázat, že funkce $\{e^{\lambda_i x}\}_{i=1}^n$ tvoří fundamentální systém dané rovnice. Zřejmě totiž

$$\begin{aligned} W_{[e^{\lambda_1 x}, e^{\lambda_2 x}, \dots, e^{\lambda_n x}]}(x) &= e^{\sum_{i=1}^n \lambda_i x} \det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix} \\ &= e^{\sum_{i=1}^n \lambda_i x} \prod_{i>j} (\lambda_i - \lambda_j) \neq 0. \end{aligned}$$

Příslušný determinant se nazývá Vandermondův determinant a jeho hodnotu lze nalézt použitím matematické indukce.

Cvičení 10.5.18. Dokažte vztah pro Vandermondův determinant, tedy ukažte, že pro libovolné $n \in \mathbb{N}$, $n \geq 2$, platí

$$\det \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \cdots & \lambda_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix} = e^{\sum_{i=1}^n \lambda_i x} \prod_{i>j} (\lambda_i - \lambda_j).$$

Pokud ale některé kořeny jsou vícenásobné, musíme postupovat jinak. Využijeme toho, že polynom stupně n má právě n komplexních kořenů, započítáváme-li násobnost. Vícenásobné kořeny přispějí odpovídajícím počtem funkcí díky následujícímu výsledku.

Tvrzení 10.5.19 (O vícenásobných kořenech charakteristického polynomu)**.** *Nechť $\lambda_j \in \mathbb{C}$ je k-násobným kořenem charakteristického polynomu. Pak funkce*

$$e^{\lambda_j x}, \quad x e^{\lambda_j x}, x^2 e^{\lambda_j x}, \quad \dots, \quad x^{k-1} e^{\lambda_j x}$$

řeší $Ly = 0$.

Důkaz. Pokud $\lambda_i = 0$ je k-násobným kořenem polynomu $p(\lambda)$, platí

$$p(\lambda) = \sum_{i=k}^{n} a_i \lambda^i \qquad \text{a odtud} \qquad Ly = \sum_{i=k}^{n} a_i y^{(i)}.$$

Proto funkce $1, x, \dots, x^{k-1}$ řeší $Ly = 0$.

Pro $\lambda_j \neq 0$ pišme $y(x) = z(x)\mathrm{e}^{\lambda_j x}$. Po dosazení máme

$$Ly = a_n(z\mathrm{e}^{\lambda_j x})^{(n)} + a_{n-1}(z\mathrm{e}^{\lambda_j x})^{(n-1)} + \cdots + a_1(z'\mathrm{e}^{\lambda_j x} + z\lambda_j \mathrm{e}^{\lambda_j x}) + a_0 z\mathrm{e}^{\lambda_j x}.$$

Ze všech členů na pravé straně se dá vytknout $\mathrm{e}^{\lambda_j x}$, a proto dostáváme, že

$$Ly = \mathrm{e}^{\lambda_j x} Mz,$$

kde M je lineární diferenciální operátor s konstantními koeficienty závislými na $a_n, a_{n-1}, \dots, a_0, n, \lambda_j$. Nechť q je charakteristický polynom operátoru M. Pro libovolné $\mu \in \mathbb{C}$ pak máme

$$q(\mu) = \frac{Me^{\mu x}}{\mathrm{e}^{\mu x}} = \frac{L(e^{\mu x}\mathrm{e}^{\lambda_j x})}{\mathrm{e}^{\mu x}\mathrm{e}^{\lambda_j x}} = \frac{Le^{(\mu+\lambda_j)x}}{e^{(\mu+\lambda_j)x}} = p(\mu + \lambda_j).$$

Odtud, protože λ_j je k-násobným kořenem polynomu p, 0 je k-násobným kořenem polynomu q. Podle počátku důkazu tedy funkce $1, x, \dots, x^{k-1}$ řeší $Mz = 0$, což znamená, že funkce $\mathrm{e}^{\lambda_j x}, x\mathrm{e}^{\lambda_j x}, \dots, x^{k-1}\mathrm{e}^{\lambda_j x}$ řeší $Ly = 0$. $\square$

V dalším nás bude zajímat lineární nezávislost právě získaných funkcí. K tomu využijeme následující obecný výsledek.

Lemma 10.5.20. *Nechť $\lambda_1, \dots, \lambda_m$ jsou různá komplexní čísla a $P_1, \dots, P_m$ jsou polynomy s komplexními koeficienty. Jestliže platí*

$$\sum_{i=1}^{m} P_i(x)\mathrm{e}^{\lambda_i x} \equiv 0 \qquad \textit{na } (a,b),$$

pak $P_i \equiv 0$ na (a,b) pro každé $i \in \{1, \dots, m\}$.

Důkaz. Postupujme indukcí přes $m \in \mathbb{N}$. Pro $m = 1$ je výsledek zřejmý. Nechť dokazovaný výrok platí pro $m - 1 \in \mathbb{N}$ a máme

$$\sum_{i=1}^{m} P_i(x)\mathrm{e}^{\lambda_i x} \equiv 0 \qquad \text{na } (a,b).$$

Aritmetickou úpravou dostáváme

$$-P_m(x) = \sum_{i=1}^{m-1} P_i(x)\mathrm{e}^{(\lambda_i - \lambda_m)x} \qquad \text{na } (a,b).$$

Obě strany nyní ($\operatorname{st} P_m + 1$)-krát zderivujeme a vychází nám (připomeňme, že $\lambda_i \neq \lambda_m$, $i = 1, 2, \ldots, m-1$)

$$0 \equiv \sum_{i=1}^{m-1} Q_i(x) \mathrm{e}^{(\lambda_i - \lambda_m)x} \qquad \text{na } (a, b),$$

kde koeficienty polynomu Q_i závisí na koeficientech polynomu P_i, čísle $\lambda_i - \lambda_m$ a čísle ($\operatorname{st} P_m + 1$). Podle indukčního předpokladu proto máme $Q_i \equiv 0$ na (a, b) pro každé $i \in \{1, \ldots, m-1\}$. Zároveň si ale povšimněme, že pro $\lambda \neq 0$ a P polynom platí

$$(P(x)\mathrm{e}^{\lambda x})' = (P'(x) + \lambda P(x))\mathrm{e}^{\lambda x},$$

neboli při derivování výrazu $P(x)\mathrm{e}^{\lambda x}$ neklesá stupeň polynomu doprovázejícího funkci $\mathrm{e}^{\lambda x}$ a dokonce u nejvyššího koeficientu je nenulový násobek koeficientu původního. Díky tomuto pozorování a výsledku $Q_i \equiv 0$ na (a, b) dostáváme, že $P_i \equiv 0$ na (a, b) pro každé $i \in \{1, \ldots, m-1\}$. Proto mutně $P_m \equiv 0$ na (a, b) a jsme hotovi. □

Tvrzení 10.5.21 (O nezávislosti komplexních řešení homogenní rovnice)**.** *Nechť $\lambda_1, \ldots, \lambda_m$ jsou různé kořeny charakteristického polynomu příslušejícího rovnici $Ly = 0$ a $k_1, \ldots, k_m$ jsou jejich násobnosti. Pak*

$$\begin{array}{llll}
\mathrm{e}^{\lambda_1 x}, & x\mathrm{e}^{\lambda_1 x}, & \ldots, & x^{k_1-1}\mathrm{e}^{\lambda_1 x}, \\
\mathrm{e}^{\lambda_2 x}, & x\mathrm{e}^{\lambda_2 x}, & \ldots, & x^{k_2-1}\mathrm{e}^{\lambda_2 x}, \\
\vdots & \vdots & \ddots & \vdots \\
\mathrm{e}^{\lambda_m x}, & x\mathrm{e}^{\lambda_m x}, & \ldots, & x^{k_m-1}\mathrm{e}^{\lambda_m x}
\end{array}$$

jsou nezávislá řešení úlohy $Ly = 0$ (neexistuje netriviální sada komplexních konstant dávající $\sum_{i=1}^n c_i y_i \equiv 0$ na (a, b), kde y_i jsou výše popsané funkce a $n = k_1 + \cdots + k_m$).

Důkaz. Dle Tvrzení o vícenásobných kořenech charakteristického polynomu (Tvrzení 10.3.4) všechny uvedené funkce řeší $Ly = 0$. Pokud pro $c_1, \ldots, c_n \in \mathbb{C}$ platí

$$0 \equiv \sum_{i=1}^{n} c_i y_i =: \sum_{j=1}^{m} P_j(x) \mathrm{e}^{\lambda_j x},$$

aplikací předchozího lemmatu dostáváme $P_1 \equiv \cdots \equiv P_m \equiv 0$, tedy $c_1 = \cdots = c_n = 0$ a máme dokázanou lineární nezávislost. □

Zbývá ještě případná komplexní řešení nahradit stejným počtem řešení reálných. Připomeňme, že má-li polynom s reálnými koeficienty nereálný kořen, kořenem je i číslo komplexně sdružené a dokonce má stejnou násobnost (tento výsledek jsme si dokázali v kapitole o primitivních funkcích). Proto kdykoliv jsme v dosavadním postupu získali komplexní řešení ve tvaru

$$y = x^l \mathrm{e}^{\lambda x} = x^l \mathrm{e}^{\operatorname{Re} \lambda x}(\cos(\operatorname{Im} \lambda x) + \mathrm{i} \sin(\operatorname{Im} \lambda x)),$$

máme také řešení

$$\tilde{y} = x^l \mathrm{e}^{\bar{\lambda}x} = x^l \mathrm{e}^{\mathrm{Re}\,\lambda x}(\cos(\mathrm{Im}\,\lambda x) - \mathrm{i}\sin(\mathrm{Im}\,\lambda x)).$$

Navíc díky linearitě operátoru L jsou funkce

$$u := \frac{y + \tilde{y}}{2} = x^l \mathrm{e}^{\mathrm{Re}\,\lambda x}\cos(\mathrm{Im}\,\lambda x)$$

a

$$v := \frac{y - \tilde{y}}{2i} = x^l \mathrm{e}^{\mathrm{Re}\,\lambda x}\sin(\mathrm{Im}\,\lambda x)$$

opět řešeními úlohy $Ly = 0$ a už se jedná o reálné funkce. Konečně, touto záměnou jsme si nemohli zkazit lineární nezávislost. Skutečně, pokud by nová sada řešení byla lineárně závislá, z předpisu pro u a v daných výše bychom dokázali spočítat netriviální (komplexní) koeficienty dávající lineární závislost původní sady řešení. Celkově jsme dostali následující výsledek.

Věta 10.5.22 (O fundamentálním systému rovnice s konstantními reálnými koeficienty)**.** *Nechť charakteristický polynom rovnice $Ly = 0$ s konstantními reálnými koeficienty má reálné kořeny $\lambda_1, \dots, \lambda_m$ a komplexní kořeny $\lambda_{m+1}, \dots, \lambda_l$, $\bar{\lambda}_{m+1}, \dots, \bar{\lambda}_l$ s násobnostmi $k_1, \dots, k_l$. Pak funkce*

$$\begin{array}{lll}
\mathrm{e}^{\lambda_1 x}, & \dots, & x^{k_1-1}\mathrm{e}^{\lambda_1 x}, \\
\vdots & \ddots & \vdots \\
\mathrm{e}^{\lambda_m x}, & \dots, & x^{k_m-1}\mathrm{e}^{\lambda_m x}, \\
\mathrm{e}^{\mathrm{Re}\,\lambda_{m+1}x}\cos(\mathrm{Im}\,\lambda_{m+1}x), & \dots, & x^{k_{m+1}-1}\mathrm{e}^{\mathrm{Re}\,\lambda_{m+1}x}\cos(\mathrm{Im}\,\lambda_{m+1}x), \\
\mathrm{e}^{\mathrm{Re}\,\lambda_{m+1}x}\sin(\mathrm{Im}\,\lambda_{m+1}x), & \dots, & x^{k_{m+1}-1}\mathrm{e}^{\mathrm{Re}\,\lambda_{m+1}x}\sin(\mathrm{Im}\,\lambda_{m+1}x), \\
\vdots & \ddots & \vdots \\
\mathrm{e}^{\mathrm{Re}\,\lambda_l x}\cos(\mathrm{Im}\,\lambda_l x), & \dots, & x^{k_l-1}\mathrm{e}^{\mathrm{Re}\,\lambda_l x}\cos(\mathrm{Im}\,\lambda_l x), \\
\mathrm{e}^{\mathrm{Re}\,\lambda_l x}\sin(\mathrm{Im}\,\lambda_l x), & \dots, & x^{k_l-1}\mathrm{e}^{\mathrm{Re}\,\lambda_l x}\sin(\mathrm{Im}\,\lambda_l x)
\end{array}$$

tvoří fundamentální systém rovnice $Ly = 0$.

Příklad 10.5.23. Uvažme rovnici

$$y^{VII} + 2y^{V} + y^{III} = 0.$$

Charakteristický polynom pak má tvar

$$p(\lambda) = \lambda^7 + 2\lambda^5 + \lambda^3 = (\lambda^2 + 1)^2\lambda^3.$$

Číslo 0 je jeho trojnásobným kořenem a čísla $\pm\mathrm{i}$ jsou shodně dvojnásobné kořeny. Fundamentální systém je proto tvořen funkcemi

$$1, \quad x, \quad x^2, \quad \cos x, \quad \sin x, \quad x\cos x \quad \text{a} \quad x\sin x \qquad \text{na } \mathbb{R}.$$

Obecné řešení na $\mathbb{R}$ je tedy dáno vztahem

$$y(x) = C_1 + C_2 x + C_3 x^2 + C_4\cos x + C_5\sin x + C_6 x\cos x + C_7 x\sin x,$$

kde $C_1, \dots, C_7$ jsou reálné konstanty.

10.5.5 Metoda speciální pravé strany pro rovnice s konstantními koeficienty

Partikulární řešení nehomogenní rovnice rovnice $Ly = f$ jsme se při znalosti fundamentálního systému naučili řešit pomocí metody variace konstant. V případě konstantních koeficientů lze pro některé typy funkce f volit podstatně jednodušší postup.

Tvrzení 10.5.24 (O speciální pravé straně). *Nechť*

$$f(x) = e^{\mu x}\big(P_1(x)\cos(\nu x) + P_2(x)\sin(\nu x)\big),$$

kde $\mu, \nu \in \mathbb{R}$ a P_1, P_2 jsou polynomy. Pak existují polynomy Q_1, Q_2 stupně nejvýše $\max\{\operatorname{st} P_1, \operatorname{st} P_2\}$ *takové, že funkce*

$$y_p = e^{\mu x}x^k\big(Q_1(x)\cos(\nu x) + Q_2(x)\sin(\nu x)\big),$$

kde $k \in \mathbb{N}_0$ je násobnost čísla $\mu + \nu i$ jakožto kořene charakteristického polynomu, řeší nehomogenní rovnici $Ly = f$.

Poznámka 10.5.25. (i) Polynomy Q_1, Q_2 se hledají metodou neurčitých koeficientů.
(ii) V tvrzení připouštíme, že $\mu + \nu \mathrm{i}$ není kořenem charakteristického polynomu a v tom případě je $k = 0$.
(iii) Pozor, i když je jeden z polynomů P_1, P_2 nulový, nemůžeme obecně předpokládat nulovost kteréhokoliv z polynomů Q_1, Q_2 (uvažte třeba rovnici $y''' - y' = \sin x$, kterou jistě neřeší žádný násobek funkce sin, naopak ji řeší násobek funkce cos).
(iv) Tvrzení o speciální pravé straně se na základních kurzech matematické analýzy nedokazuje. Bývá zahrnuto do případu uhodnutí (kdykoliv najdeme koeficienty polynomů Q_1, Q_2, ukázali jsme, že Tvrzení o speciální pravé straně platí přinejmenším v našem případě).
(v) Metoda speciální pravé strany se u složitějších pravých stran aplikuje jen na některé části (sčítance) funkce f s využitím toho, že operátor L je lineární, a proto

$$Ly_1 = f_1 \quad \wedge \quad Ly_2 = f_2 \qquad \Longrightarrow \qquad L(y_1 + y_2) = f_1 + f_2.$$

Tomuto jevu se říká *princip superpozice.*

Příklad 10.5.26. Hledejme obecné řešení úlohy

$$y''' - y'' + y' - y = 1 + \cos x + x\mathrm{e}^{-x}.$$

Charakteristický polynom má tvar

$$\lambda^3 - \lambda^2 + \lambda - 1 = (\lambda - 1)(\lambda^2 + 1).$$

Jednonásobné kořeny $1, -\mathrm{i}, \mathrm{i}$ dávají fundamentální systém

$$\{\mathrm{e}^x, \cos x, \sin x\}.$$

Obecné řešení homogenní rovnice má proto tvar

$$y_h = C_1\mathrm{e}^x + C_2\cos x + C_3\sin x \qquad \text{na } \mathbb{R},$$

kde $C_1, C_2, C_3 \in \mathbb{R}$. Pravá strana jako celek nemá tvar popsaný v Tvrzení o speciální pravé straně (Tvrzení 10.5.24), ale je součtem tří funkcí, které jednotlivě požadovaný tvar mají. Použijeme tedy princip superpozice a budeme hledat partikulární řešení úloh $Ly_1 = 1$, $Ly_2 = \cos x$, $Ly_2 = x\mathrm{e}^{-x}$.

V prvním případě snadno uhodneme $y_1 \equiv -1$. Ve druhém případě hledáme řešení ve tvaru (číslo i je jednonásobný kořen charakteristického polynomu)

$$y_{p_2} = Ax\cos x + Bx\sin x.$$

Tento tvar dosadíme do řešené rovnice, použijeme Leibnizovo pravidlo a upravujeme

$$\begin{aligned}
\cos x =& \Big(-3A\cos x + Ax\sin x - 3B\sin x - Bx\cos x\Big)\\
&- \Big(-2A\sin x - Ax\cos x + 2B\cos x - Bx\sin x\Big)\\
&+ \Big(A\cos x - Ax\sin x + B\sin x + Bx\cos x\Big) - \Big(Ax\cos x + Bx\sin x\Big)\\
=& \,(2A - 2B)\sin x + (-2A - 2B)\cos x.
\end{aligned}$$

Dostáváme $A = B = -\frac{1}{4}$ a odtud

$$y_{p_2} = -\frac{1}{4}x\cos x - \frac{1}{4}x\sin x.$$

Třetí partikulární řešení hledáme ve tvaru (číslo -1 není kořenem charakteristického polynomu)

$$y_{p_3} = (Ax + B)\mathrm{e}^{-x}.$$

Opět dosadíme

$$\begin{aligned}
x\mathrm{e}^{-x} &= \Big(3A\mathrm{e}^{-x} - (Ax+B)\mathrm{e}^{-x}\Big) - \Big(-2A\mathrm{e}^{-x} + (Ax+B)\mathrm{e}^{-x}\Big)\\
&\quad + \Big(A\mathrm{e}^{-x} - (Ax+B)\mathrm{e}^{-x}\Big) - (Ax+B)\mathrm{e}^{-x}\\
&= (6A - 4(Ax+B))\mathrm{e}^{-x}.
\end{aligned}$$

Dostáváme $A = -\frac{1}{4}$, $B = -\frac{3}{8}$ a odtud

$$y_{p_3} = \Big(-\frac{1}{4}x - \frac{3}{8}\Big)\mathrm{e}^{-x}.$$

Obecné řešení úlohy na $\mathbb{R}$ má proto tvar

$$\begin{aligned}
y &= y_h + y_{p_1} + y_{p_2} + y_{p_3}\\
&= C_1\mathrm{e}^x + C_2\cos x + C_3\sin x - 1 - \frac{1}{4}x\cos x - \frac{1}{4}x\sin x - \Big(\frac{1}{4}x + \frac{3}{8}\Big)\mathrm{e}^{-x}.
\end{aligned}$$

10.5.6 Eulerova rovnice

Jednou z úloh s nekonstantními koeficienty, kde je znám postup řešení, je *Eulerova rovnice*
$$\sum_{i=0}^{n} a_i x^i y^{(i)}(x) = f(x),$$
kde $a_0, \dots, a_n$ jsou reálné konstanty a $f \in C((a,b))$. U této úlohy máme k dispozici metodu variace konstant. Naším dalším úkolem je nalezení fundamentálního systému. Budeme pracovat zvlášť pro $x > 0$ a $x < 0$. Pro $x > 0$ zavádíme novou proměnnou předpisem $\xi = \log x$ (tedy $x = \mathrm{e}^{\xi}$) a pomocnou funkci $z(\xi) := y(x(\xi))$. Pak máme
$$\begin{aligned} z'(\xi) &= y'(x(\xi))\mathrm{e}^{\xi} && = xy' \\ z''(\xi) &= y''(x(\xi))\mathrm{e}^{2\xi} + y'(x(\xi))\mathrm{e}^{\xi} && = x^2y'' + xy' \end{aligned}$$
a tak dále. Je vidět, že po dosazení získaných vztahů do rovnice $Ly = 0$ dostaneme pro funkci $\xi \mapsto z(\xi)$ rovnici n-tého řádu s konstantními koeficienty, které závisí na číslech $a_0, \dots, a_n$.

Pro $x < 0$ zavádíme $\xi = \log|x|$ (tedy $x = -\mathrm{e}^{\xi}$) a opět $z(\xi) = y(x(\xi))$. Pak máme
$$\begin{aligned} z'(\xi) &= y'(x(\xi))(-\mathrm{e}^{\xi}) && = xy' \\ z''(\xi) &= y''(x(\xi))(-\mathrm{e}^{\xi})^2 + y'(x(\xi))(-\mathrm{e}^{\xi}) && = x^2y'' + xy' \end{aligned}$$
a tak dále. Vyjde nám pomocná diferenciální rovnice se stejnými koeficienty jako v předchozím případě. Po vyřešení pomocné diferenciální rovnice se vrátíme k proměnné x a funkci y. Nakonec se ještě pokusíme řešení slepit v počátku.

Příklad 10.5.27. Hledejme obecné řešení rovnice
$$x^2y'' + y = x^3.$$
Nejprve řešme homogenní úlohu. Pro $x > 0$ (totéž vyjde i pro $x < 0$) máme podle vzorců odvozených výše
$$0 = x^2y'' + y = x^2y'' + xy' - xy' + y = z'' - z' + z.$$
Charakteristický polynom nové rovnice má tvar $p(\lambda) = \lambda^2 - \lambda + 1$ a kořeny $\frac{1}{2} \pm \frac{\sqrt{3}}{2}\mathrm{i}$. Odtud dostáváme
$$z(\xi) = C_1 \mathrm{e}^{\frac{1}{2}\xi} \cos\Big(\frac{\sqrt{3}}{2}\xi\Big) + C_2 \mathrm{e}^{\frac{1}{2}\xi} \sin\Big(\frac{\sqrt{3}}{2}\xi\Big) \qquad \text{na } \mathbb{R}.$$
Proto máme pro $x > 0$ a $x < 0$ řešení (píšeme $\xi = \log|x|$ a lehce upravujeme)
$$y(x) = C_1\sqrt{|x|}\cos\Big(\frac{\sqrt{3}}{2}\log|x|\Big) + C_2\sqrt{|x|}\sin\Big(\frac{\sqrt{3}}{2}\log|x|\Big).$$
Partikulární řešení zkusíme uhodnout ve tvaru $y_p = Ax^3$, což okamžitě dává $A = \frac{1}{7}$. Dostali jsme obecné řešení
$$y(x) = C_1\sqrt{|x|}\cos\Big(\frac{\sqrt{3}}{2}\log|x|\Big) + C_2\sqrt{|x|}\sin\Big(\frac{\sqrt{3}}{2}\log|x|\Big) + \frac{1}{7}x^3.$$

Slepení řešení v počátku je možné v případě $C_1 = C_2 = 0$ (v počátku spojitě dodefinujeme nulou), v ostatních případech se dá nahlédnout, že po spojitém dodefinování v počátku nemá získaná funkce derivaci v počátku.

Poznámka 10.5.28. (i) Na Eulerovu rovnici se dají aplikovat všechny naše obecné výsledky (kromě výsledků pro rovnice s konstantními koeficienty). Připomeňme však, že pro tyto výsledky potřebujeme, aby byl činitel u nejvyšší derivace nenulový, což Eulerova rovnice nesplňuje v počátku. Není tedy žádným překvapením, že jsme předchozí příklad vyřešili právě na intervalech $(-\infty, 0)$ a $(0, \infty)$, zatímco jsme měli potíže se slepením v počátku.
(ii) Funkce $u_1(x) = \sqrt{|x|}\cos(\frac{\sqrt{3}}{2}\log|x|)$ a $u_2(x) = \sqrt{|x|}\sin(\frac{\sqrt{3}}{2}\log|x|)$ z minulého příkladu tvoří fundamentální systém. Skutečně, výpočtem jsme ověřili, že se jedná o řešení. Dále, pokud by existovala netriviální dvojice konstant C_1, C_2 tak, že

$$C_1u_1 + C_2u_2 \equiv 0,$$

tatáž dvojice konstant by u funkcí $\mathrm{e}^{\frac{1}{2}\xi}\cos(\frac{\sqrt{3}}{2}\xi)$ a $\mathrm{e}^{\frac{1}{2}\xi}\sin(\frac{\sqrt{3}}{2}\xi)$ vyvracela lineární nezávislost, kterou jsme dokázali v předchozí kapitole.
(iii) Díky předchozí části poznámky máme k dispozici variaci konstant, kdykoliv se nám podaří vyřešit homogenní rovnici.
(iv) Řešení homogenní Eulerovy rovnice se dá někdy uhodnout tak, že položíme $y = |x|^\lambda$ a funkci zkusíme dosadit do rovnice. Tento přístup vede na jakousi charakteristickou rovnici. Zatím však nevíme, že získaná řešení jsou lineárně nezávislá a neumíme ani zajistit jejich dostatečný počet (problémy dělají komplexní a vícenásobné kořeny). Můžeme si ale povšimnout (podrobně si rozmyslete sami), že zmíněná charakteristická rovnice je tožná s charakteristickou rovnicí pro úlohu s funkcí z. Díky tomu víme, že u vícenásobných kořenů máme do fundamentálního systému brát funkce typu

$$|x|^\lambda, \quad \log|x||x|^\lambda, \quad \log^2|x||x|^\lambda, \quad \dots$$

a v případě komplexního kořene λ do fundamentálního systému přijdou funkce

$$|x|^{\operatorname{Re}\lambda}\cos(\operatorname{Im}\lambda\log|x|), \quad |x|^{\operatorname{Re}\lambda}\sin(\operatorname{Im}\lambda\log|x|), \quad \log|x||x|^{\operatorname{Re}\lambda}\cos(\operatorname{Im}\lambda\log|x|)$$

a tak dále. Tato úvaha nám dokonce umožňuje používat Tvrzení o speciální pravé straně (Tvrzení 10.5.24), pokud ve všech jeho formulích nahradíme proměnnou x výrazem $\log|x|$.

Příklad 10.5.29. Řešme rovnici

$$x^3y''' + 4x^2y'' + 2xy' = x + \frac{1}{x} + x\log x.$$

Do homogenní rovnice dosadíme $y = x^\lambda$ pro $x > 0$ (pro záporná x pravá strana rovnice nemá smysl, jinak bychom pracovali s $|x|^\lambda$ a vyšla by tatáž charakteristická rovnice, neboť při každém derivování získáme činitel $\operatorname{sign} x$, který pak spolu s x dá $|x|$) a máme

$$0 = x^3\lambda(\lambda-1)(\lambda-2)x^{\lambda-3} + 4x^2\lambda(\lambda-1)x^{\lambda-2} + 2x\lambda x^{\lambda-1} = x^\lambda\lambda^2(\lambda+1).$$

Dostáváme fundamentální systém

$$\left\{1, \log x, \frac{1}{x}\right\}.$$

Modifikované Tvrzení o speciální pravé straně (Tvrzení 10.5.24) (pro $x > 0$ nemusíme psát absolutní hodnoty) požaduje pravou stranu ve tvaru

$$f(x) = x^{\mu}\Big(P_1(\log x)\cos(\nu \log x) + P_2(\log x)\sin(\nu \log x)\Big)$$

a nabízí partikulární řešení ve tvaru

$$y_p(x) = x^{\mu} \log^k x\Big(Q_1(\log x)\cos(\nu \log x) + Q_2(\log x)\sin(\nu \log x)\Big),$$

kde k je násobnost čísla $\mu + \mathrm{i}\nu$ jako kořene charakteristického polynomu. Využijme principu superpozice a řešme napřed úlohu $Ly_{p_1} = x$. Řešení hledáme ve tvaru

$$y_{p_1} = Ax.$$

Vychází nám $y_{p_1} = \frac{1}{2}x$. Dále řešme $Ly_{p_2} = \frac{1}{x}$. Protože číslo -1 je jednonásobným kořenem charakteristického polynomu, hledáme řešení ve tvaru

$$y_{p_2} = A\frac{\log x}{x}.$$

Dostáváme

$$\frac{1}{x} = Ax^3\Big(\frac{11}{x^4} - \frac{6\log x}{x^4}\Big) + 4Ax^2\Big(-\frac{3}{x^3} + \frac{2\log x}{x^3}\Big) + 2Ax\Big(\frac{1}{x^2} - \frac{\log x}{x^2}\Big) = \frac{A}{x}.$$

Proto $y_{p_2} = \frac{\log x}{x}$. Konečně, řešme $Ly_{p_3} = x\log x$. Řešení hledáme ve tvaru

$$y_{p_3} = Ax\log x + Bx.$$

Dostáváme

$$\begin{aligned} x\log x &= x^3\Big(-\frac{A}{x^2}\Big) + 4x^2\Big(\frac{A}{x}\Big) + 2x\Big(A\log x + A + B\Big) \\ &= (5A + 2B)x + 2Ax\log x. \end{aligned}$$

Odtud $A = \frac{1}{2}$, $B = -\frac{5}{4}$ a $y_{p_3} = \frac{1}{2}x\log x - \frac{5}{4}x$. Obecným řešením naší úlohy tedy je

$$y = C_1 + C_2\log x + C_3\frac{1}{x} + \frac{1}{2}x + \frac{\log x}{x} + \frac{1}{2}x\log x - \frac{5}{4}x \qquad \text{na } (0,\infty).$$

10.6 Další typy rovnic vyšších řádů

Nyní si představíme další typy rovnic, které umíme přímo vyřešit nebo alespoň převést na rovnice prvního řádu.

10.6.1 Rovnice tvaru $y^{(n)} = f(x)$

Úlohu

$$y^{(n)} = f(x) \qquad y(x_0) = y_0, \quad y'(x_0) = y_1, \quad \dots, \quad y^{(n-1)}(x_0) = y_{n-1},$$

kde $f \in C((a,b))$, už umíme řešit, neboť je případem lineární diferenciální rovnice n-tého řádu s konstantními koeficienty (dokonce s velmi jednoduchým fundamentálním systémem $\{1, x, \dots, x^{n-1}\}$). Existuje však rychlejší postup, který zde umí zastoupit variaci konstant. Je založen na postupném integrování, které je zároveň možné přepsat do vzorce obsahujícího jediný integrál.

Tvrzení 10.6.1. *Je-li f spojitá na (a,b), pak řešení počáteční úlohy splňuje*

$$\begin{aligned} y(x) &= \int_{x_0}^{x}\int_{x_0}^{\tau_1}\dots\int_{x_0}^{\tau_{n-1}} f(\tau_n)\,\mathrm{d}\tau_n \dots \mathrm{d}\tau_2\,\mathrm{d}\tau_1 + \sum_{k=0}^{n-1}\frac{y_k}{k!}(x-x_0)^k \\ &= \frac{1}{(n-1)!}\int_{x_0}^{x}(x-\tau)^{n-1}f(\tau)\,\mathrm{d}\tau + \sum_{k=0}^{n-1}\frac{y_k}{k!}(x-x_0)^k. \end{aligned}$$

Důkaz. První rovnost snadno ověříme postupným derivováním (indukcí přes $n \in \mathbb{N}$). I druhou rovnost ověříme indukcí. Pro $n = 1$ rovnost zřejmě platí. Pro $n > 1$ stačí ukázat klíčovou identitu

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_{x_0}^{x}(x-t)^{n-1}f(t)\,\mathrm{d}t = (n-1)\int_{x_0}^{x}(x-t)^{n-2}f(t)\,\mathrm{d}t.$$

Zafixujme $x \in (x_0, b)$. Budeme postupovat podobně jako v důkazu takzvané hlavní věty diferenciálního a integrálního počtu (Věta 7.5.12). Označme

$$\Phi(x) := \int_{x_0}^{x}(x-t)^{n-1}f(t)\,\mathrm{d}t \qquad \text{a} \qquad \varphi(x) := (n-1)\int_{x_0}^{x}(x-t)^{n-2}f(t)\,\mathrm{d}t.$$

Zafixujme ještě $y \in (x, b)$. Pak díky spojitosti f na $[x_0, y] \subset (a,b)$ máme $K > 0$ takové, že

$$t \in [x_0, y] \qquad \Longrightarrow \qquad |f(t)| \le K.$$

Dále máme pro $h \in (0, y-x)$

$$\begin{aligned} \frac{\Phi(x+h)-\Phi(x)}{h} &= \frac{1}{h}\Big(\int_{x_0}^{x+h}(x+h-t)^{n-1}f(t)\,\mathrm{d}t - \int_{x_0}^{x}(x-t)^{n-1}f(t)\,\mathrm{d}t\Big) \\ &= \frac{1}{h}\int_{x_0}^{x}\big((x+h-t)^{n-1}-(x-t)^{n-1}\big)f(t)\,\mathrm{d}t \\ &\qquad + \frac{1}{h}\int_{x}^{x+h}(x+h-t)^{n-1}f(t)\,\mathrm{d}t =: I_1 + I_2. \end{aligned}$$

Odtud

$$|I_2| \le \frac{1}{h}\int_{x}^{x+h} h^{n-1}K\,\mathrm{d}t = h^{n-1}K \overset{h\to 0+}{\to} 0.$$

Dále máme díky předchozím vzorcům a binomické větě

$$\begin{aligned}
|I_1 - \varphi(x)| &= \Big| \int_{x_0}^{x} \Big(\frac{(x-t+h)^{n-1} - (x-t)^{n-1}}{h} - (n-1)(x-t)^{n-2} \Big) f(t)\,\mathrm{d}t \Big| \\
&= \Big| \int_{x_0}^{x} \frac{1}{h} \sum_{j=2}^{n-1} \binom{n-1}{j} h^j (x-t)^{n-1-j} f(t)\,\mathrm{d}t \Big| \\
&\leq \int_{x_0}^{x} \sum_{j=2}^{n-1} \binom{n-1}{j} h^{j-1} (x-x_0)^{n-1-j} K\,\mathrm{d}t \leq \sum_{j=2}^{n-1} C h^{j-1} \overset{h\to 0+}{\to} 0.
\end{aligned}$$

Celkově máme $\Phi'_+(x) = \varphi(x)$. Podobně se získá $\Phi'_-(x) = \varphi(x)$. Případy $x \in (a, x_0)$ a $x = x_0$ jsou analogické. Tím je dokázána klíčová identita pro (teď už snadnou) matematickou indukci. □

10.6.2 Rovnice tvaru $y^{(n)} = f(x, y^{(n-1)})$

U tohoto typu zavedeme pomocnou funkci $z(x) = y^{n-1}(x)$. Pro ni potom máme pomocnou úlohu

$$z' = f(x, z),$$

což je typ úlohy, pro který jsme si již představili několik metod (v závislosti na tvaru funkce f). Po nalezení funkce z ještě potřebujeme vyřešit úlohu

$$y^{n-1} = z,$$

kterou jsme řešili před chvílí. Poznamenejme, že řešení úlohy pomocné a úlohy původní si vzájemně odpovídají (přechodem k pomocné úloze žádná řešení nezískáme ani neztratíme). Skutečně, pokud nějaká funkce řeší úlohu původní, její derivace řádu $(n-1)$ zřejmě řeší úlohu pomocnou. Naopak, pokud nějaká funkce řeší úlohu pomocnou, je nutně spojitá, můžeme ji proto $(n-1)$-krát integrovat a dostaneme řešení úlohy původní.

Příklad 10.6.2. Řešme úlohu

$$y^V = \frac{2y^{IV}}{x} \qquad y(1) = 1, \quad y'(1) = 0, \quad y''(1) = 1, \quad y'''(1) = 0, \quad y^{IV}(1) = 1.$$

Položme $z := y^{IV}$. Tím jsme přešli k úloze

$$z' = \frac{2z}{x} \qquad z(1) = 1$$

(díky včasné aplikaci počátečních podmínek se nebudeme muset zabývat všemi případy). Jedná se o rovnici se separovanými proměnnými a máme pro ni

$$\log|z| = \int \frac{\mathrm{d}z}{z} = \int \frac{2\mathrm{d}x}{x} = 2\log|x| + C,$$

odkud s využitím počáteční podmínky dostáváme řešení

$$z = x^2 \qquad \text{na } (0, \infty).$$

Toto řešení je jednoznačné na $(0,\infty)\times(0,\infty)$, slepení (v počátku) zadání nepřipouští. Celkově pak na $(0,\infty)$ máme řešení

$$y(x) = \Phi(x) + 1 + \frac{(x-1)^2}{2} + \frac{(x-1)^4}{24},$$

kde funkci Φ můžeme určit pomocí předchozího tvrzení. První možností jsou čtyři postupné integrace, které vzhledem k dolní mezi $x_0 = 1$ nejsou příliš příjemné. Druhou možností je použít jedinou integraci se vzorcem

$$\Phi(x) = \frac{1}{6}\int_1^x (x-t)^3 t^2 \,\mathrm{d}t.$$

Cvičení 10.6.3. Řešte úlohu $y^{IV} = x^2$ pomocí Tvrzení o speciální pravé straně (Tvrzení 10.5.24) a pomocí variace konstant kombinované s Cramerovým pravidlem.

10.6.3 Rovnice tvaru $y^{(n)} = f\big(y^{(n-2)}\big)$

Zde pokládáme $z(x) = y^{n-2}(x)$ a tím dostáváme

$$z'' = f(z).$$

Nyní je výhodné obě strany vynásobit činitelem $2z'$ (nejedná se o ekvivalentní úpravu, novou rovnici řeší širší třída funkcí než rovnici původní), čímž obdržíme

$$(z'^2)' = 2z'z'' = 2z'f(z).$$

Má-li funkce f primitivní funkci F, potom máme

$$z'^2 = 2F(z) + C.$$

Uvážíme oba případy $z' = \pm\sqrt{2F(z)+C}$, nalezneme z a pak y (žádný z těchto kroků neprojde zcela obecně, nicméně třeba případná spojitost f zaručuje existenci všech zúčastněných primitivních funkcí).

Poznámka 10.6.4. Podobně by se postupovalo v případě rovnice

$$y^{(n)}(x) = f(x, y^{(n-2)}(x)).$$

Jako výše bychom přešli k rovnici

$$z''(x) = f(x, z(x)),$$

po vynásobení $2z'(x)$ potom

$$(z'^2)'(x) = 2z'(x)f(x, z(x)).$$

Označíme-li $F(x,z) = \int f(x,z)\,\mathrm{d}z$, dostáváme podobně jako výše rovnici

$$z'(x) = \pm\sqrt{2F(x,z)+C},$$

kterou v některých případech můžeme být schopni řešit. Poslední krok je potom stejný jako výše.

Příklad 10.6.5. Řešme úlohu

$$y''' = \frac{1}{4\sqrt{y'}} \qquad y(0) = 0, \quad y'(0) = 1, \quad y''(0) = 1.$$

Položme $z = y'$. Pak máme

$$z'' = \frac{1}{4\sqrt{z}} \qquad z(0) = 1, \quad z'(0) = 1$$

(díky včasné aplikaci počátečních podmínek se nebudeme muset zabývat všemi případy). Pak po přenásobení $2z'$ dostáváme

$$(z'^2)' = 2z'\frac{1}{4\sqrt{z}} \qquad \Longrightarrow \qquad z'^2 = \sqrt{z} + C.$$

Z počátečních podmínek dále plyne

$$z' = \sqrt[4]{z}.$$

Tuto úlohu řešíme separací proměnných a získáváme

$$\frac{4}{3}z^{\frac{3}{4}} = \int \frac{\mathrm{d}z}{\sqrt[4]{z}} = \int \mathrm{d}x = x + C.$$

Odtud díky počáteční podmínce $z(0) = 1$ dostáváme

$$z = \left(\frac{3}{4}x + 1\right)^{\frac{4}{3}}.$$

Všimněme si, že tato funkce není v bodě $x = -\frac{4}{3}$ dvakrát spojitě diferencovatelná, což souvisí s tím, že $z'(-\frac{4}{3}) = 0$. V takovém případě nebyla výše provedená úprava ekvivalentní. Musíme se tedy omezit na interval $(-\frac{4}{3}, \infty)$.

Podle teorie rovnic se separovanými proměnnými je výše uvedené řešení jednoznačné na $\mathbb{R} \times (0, \infty)$. Slepování s triviálním řešením $z \equiv 0$ či řešeními na $\mathbb{R} \times (-\infty, 0)$ nepřipadá v úvahu díky původní formulaci $z'' = \frac{1}{4\sqrt{z}}$. Nyní

$$y = \int \left(\frac{3}{4}x + 1\right)^{\frac{4}{3}} \mathrm{d}x = \frac{4}{7}\left(\frac{3}{4}x + 1\right)^{\frac{7}{3}} + C \overset{y(0)=0}{=} \frac{4}{7}\left(\frac{3}{4}x + 1\right)^{\frac{7}{3}} - \frac{4}{7}$$

pro $x \in (-\frac{4}{3}, \infty)$.

10.6.4 Rovnice tvaru $y^{(n)} = f(x, y^{(k)}, y^{(k+1)}, \dots, y^{(n-1)})$

Definujeme-li pomocnou funkci $z := y^{(k)}$, dostáváme pro ni

$$z^{(n-k)} = f(x, z, z', \dots, z^{(n-k-1)}),$$

což je rovnice nižšího řádu.

Příklad 10.6.6. Úlohu

$$y^{XXII} - 2y^{XXI} + y^{XX} = \mathrm{e}^{2x} \qquad y(0) = y'(0) = \cdots = y^{XX}(0) = 0, \quad y^{XXI}(0) = 1$$

můžeme řešit standardním způsobem. Jde také zavést pomocnou funkci $z = y^{XX}$ a řešit

$$z'' - 2z' + z = \mathrm{e}^{2x} \qquad z(0) = 0, \quad z'(0) = 1.$$

Obecné řešení má tvar

$$z = C_1\mathrm{e}^x + C_2x\mathrm{e}^x + \mathrm{e}^{2x} \qquad \text{na } \mathbb{R}.$$

Po aplikaci počátečních podmínek pro z dostáváme $z = \mathrm{e}^{2x} - \mathrm{e}^x$. Zbývá dořešit úlohu

$$y^{XX} = \mathrm{e}^{2x} - \mathrm{e}^x \qquad y(0) = y'(0) = \cdots = y^{XIX}(0) = 0,$$

což je typ, kterým jsme se již podrobně zabývali (v tomto případě se dá poměrně rychle provést dvacet integrací za sebou) a dostaneme součet polynomu a funkce $2^{-20}\mathrm{e}^{2x} - \mathrm{e}^x$.

10.6.5 Rovnice tvaru $y^{(n)} = f(y, y', \ldots, y^{n-1})$

V tomto případě se používá následující postup ke snížení řádu o jedna. Pokud má rovnice řešení, které splňuje

$$y'(x) = p(y(x))$$

pro nějakou $(n-1)$-krát diferencovatelnou funkci p, pak pro ni máme

$$\begin{aligned} y'(x) &= p(y(x)) \\ y''(x) &= p'(y(x))y'(x) = p'(y(x))p(y(x)) \\ y'''(x) &= p''(y(x))y'(x)p(y(x)) + p'(y(x))p'(y(x))y'(x) \\ &= p''(y(x))p^2(y(x)) + (p'(y(x)))^2p(y(x)) \qquad \text{atd.} \end{aligned}$$

Po dosazení do původní rovnice a zjednodušení zápisu položením $z := y(x)$ dostáváme pro hledanou funkci z diferenciální rovnici $(n-1)$-tého řádu.

Zamysleme se ještě nad tím, jakým způsobem si odpovídají řešení původní a pomocné úlohy. Jak jsme již odvodili výše, pokud nějaké řešení původní úlohy má výše uvedenou speciální vlastnost, funkce p řeší pomocnou rovnici. Naopak, pokud funkce p řeší pomocnou rovnici a funkce y řeší $y'(x) = p(y(x))$, pak z výpočtu výše plyne (máme k dispozici hladkost potřebnou pro všechna složená derivování), že y řeší rovnici původní. Musíme být ale opatrní s definičními obory funkcí p a y.

Příklad 10.6.7. Uvažme úlohu

$$y'' = 2y'y \qquad y(0) = 1, \quad y'(0) = 1$$

a pokusme se nalézt alespoň jedno její řešení. Popsaný postup nám dává rovnici

$$p'(z)p(z) = 2zp(z).$$

Tuto rovnici splňuje jednak $p \equiv 0$ (tedy $y' \equiv 0$, což nevyhovuje počáteční podmínce $y'(0) = 1$). Druhou možností je řešení rovnice

$$p'(z) = 2z,$$

tedy $p(z) = z^2 + C$, neboli $y' = y^2 + C$. Počáteční podmínka pak vyžaduje $C = 0$. Určením p jsme tedy získali diferenciální rovnici

$$y' = y^2.$$

Jedná se o rovnici se separovanými proměnnými, proto pokračujeme

$$-\frac{1}{y} = \int \frac{\mathrm{d}y}{y^2} = \int \mathrm{d}x = x + C.$$

Odtud

$$y = \frac{-1}{x + C}$$

a počáteční podmínky dávají

$$y = \frac{1}{1 - x} \qquad \text{na } (-\infty, 1).$$

Příklad 10.6.8. Zkusme naši novou metodu aplikovat na několik lineárních rovnic druhého řádu, kde jsme fundamentální systém získali pomocí uhodnutého řešení ve tvaru $\mathrm{e}^{\lambda x}$.
(i) Uvažme rovnici

$$y'' = 0.$$

Naše metoda dává, že pokud existuje řešení splňující $y' = p(y)$, pak (připomeňme $y = z$, $y' = p(z)$, $y'' = p'(z)p(z)$)

$$p'p = 0.$$

Tuto rovnici zřejmě řeší $p \equiv C$. Odtud $y' \equiv C$, a proto $y = Cx + D$ je řešením původní úlohy.
(ii) Uvažme rovnici

$$y'' + y = 0.$$

Tentokrát dostáváme

$$p'p = -z.$$

Jedná se o rovnici se separovanými proměnnými, ale rychlejší je přepis

$$\frac{1}{2}(p^2)' = -z.$$

Odtud

$$p(z) = \pm\sqrt{-z^2 + C} \qquad \text{tedy} \qquad y' = \pm\sqrt{-y^2 + C}.$$

Smysl má jen případ $C > 0$ a odpovídající rovnici se separovanými proměnnými řešíme postupem

$$\arcsin \frac{y}{\sqrt{C}} = \int \frac{\mathrm{d}y}{\sqrt{-y^2 + C}} = \pm \int \mathrm{d}x = \pm x + D.$$

Odtud (pro $\pm x + D \in (-1, 1)$)

$$y = \sqrt{C} \sin(\pm x + D).$$

Vhodnou volbou konstant a slepováním dostaneme sinus i kosinus ($\cos x = \sin(x + \frac{\pi}{2})$).
(iii) Pro rovnici

$$y'' - y = 0.$$

dostáváme

$$p'p = z.$$

Přepis $\frac{1}{2}(p^2)' = z$ vede na

$$p(z) = \pm\sqrt{z^2 + C} \qquad \text{tedy} \qquad y' = \pm\sqrt{y^2 + C}.$$

Pro $C = 0$ máme $y' = \pm y$ a nám známá řešení $\mathrm{e}^x, \mathrm{e}^{-x}$. Ani případ $C > 0$ nedává nepoužitelný výsledek, neboť máme

$$\operatorname{argsinh} \frac{x}{\sqrt{C}} = \int \frac{\mathrm{d}y}{\sqrt{y^2 + C}} = \pm \int \mathrm{d}x = \pm x + D$$

a odtud

$$y = \sqrt{C} \sinh(\pm x + D).$$

Naopak pro $C < 0$ máme

$$\operatorname{argcosh} \frac{x}{\sqrt{|C|}} = \int \frac{\mathrm{d}y}{\sqrt{y^2 - |C|}} = \pm \int \mathrm{d}x = \pm x + D$$

a odtud

$$y = \sqrt{|C|} \cosh(\pm x + D).$$

Poslední dva výsledky dávají opět řešení a dají se z nich získat funkce $\mathrm{e}^x, \mathrm{e}^{-x}$.
(iv) V případě rovnice

$$y'' - 2y' + y = 0$$

(charakteristická rovnice má tvar $(\lambda - 1)^2 = 0$ a fundamentální systém je $\{\mathrm{e}^x, x\mathrm{e}^x\}$) dostáváme

$$p'p - 2p + z = 0 \qquad \text{tedy} \qquad p' = \frac{2p - z}{p}.$$

Jedná se o homogenní rovnici, kterou umíme řešit. Jen poznamenejme, že jedním z řešení je zřejmě $p(z) = z$, což vede na rovnici $y' = y$ a řeší ji například e^x. Díky této znalosti a metodě doplňování fundamentálního systému pomocí vlastností wronskiánu se pak získá i $x\mathrm{e}^x$.

Kapitola 11

Metrické prostory

V předchozích kapitolách jsme se (poměrně úspěšně) zabývali budováním teorie reálných funkcí jedné reálné proměnné. Podobnou teorii bychom měli rádi k dispozici i ve složitějších případech, jako jsou funkce více proměnných, či pro zobrazení, která přiřazují reálné číslo prvkům z vhodné množiny funkcí. Posledně jmenovaný případ má uplatnění například v teorii deformací či v některých pokročilých postupech řešení parciálních diferenciálních rovnic. Vzpomeňme si také na starověkou úlohu o brachystochroně, která vyžaduje takto obecný přístup.

Základním pojmem v diferenciálním počtu je limita, která se neobejde bez pojmu okolí a ten zase používá pojem vzdálenosti (v $\mathbb{R}$ nám takto sloužila absolutní hodnota). Připomeňme si eukleidovskou vzdálenost na $\mathbb{R}^N$, se kterou jsme již několikrát pracovali. Ta se pro $x, y \in \mathbb{R}^N$ definuje předpisem

$$\varrho(x,y) = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_N - y_N)^2}.$$

Pro všechna $x, y, z \in \mathbb{R}^N$ platí

$$\begin{aligned} &\text{(i) } \varrho(x,y) \geq 0 \quad \text{a} \quad \varrho(x,y) = 0 \Longleftrightarrow x = y \\ &\text{(ii) } \varrho(x,y) = \varrho(y,x) \\ &\text{(iii) } \varrho(x,z) \leq \varrho(x,y) + \varrho(y,z). \end{aligned} \tag{11.0.1}$$

Platnost prvních dvou vlastností je zřejmá. Třetí vlastnost se nazývá *trojúhelníková nerovnost* a dokáže se snadno pomocí První Cauchy–Schwarzovy nerovnosti.

Později zjistíme, že uvedené tři vlastnosti jsou klíčové k vybudování pojmu limity a odvození jejích standardních vlastností, na které jsme zvyklí z předchozích kapitol.

11.1 Základní pojmy

Předchozí úvahy nás vedou k následující definici.

Definice 11.1.1 (Metrika a metrický prostor)**.** Nechť P je množina. Zobrazení $\varrho\colon P\times P\to[0,\infty)$ se nazývá *metrika*, jestliže má vlastnosti (i), (ii) a (iii) ze vztahu (11.0.1). V takovém případě se dvojice (P,ϱ) nazývá *metrický prostor*.

Příklad 11.1.2. (i) Na $\mathbb{R}$ je možné uvažovat například metriky

$$\varrho(x,y)=|x-y|,\qquad \varrho(x,y)=\frac{1}{3}|x-y|\qquad\text{a}\qquad \varrho(x,y)=\arctan|x-y|$$

(při důkazu trojúhelníkové nerovnosti pro poslední metriku se využije nerovnost $\arctan(\alpha+\beta)\le\arctan\alpha+\arctan\beta$ platná pro $\alpha,\beta\ge 0$, která se získá z Lagrangeovy věty o přírůstku funkce díky konkávnosti funkce arctan).
(ii) Na $\mathbb{R}^N$ se často pracuje s metrikami:

$$\begin{aligned}\varrho_1(x,y)&=\sum_{i=1}^N|x_i-y_i|\\ \varrho_2(x,y)&=\sqrt{\sum_{i=1}^N|x_i-y_i|^2}\\ \varrho_\infty(x,y)&=\max_{i\in\{1,\dots,N\}}|x_i-y_i|.\end{aligned}$$

Metrikami dokonce jsou pro $p\in[1,\infty)$

$$\varrho_p(x,y)=\Big(\sum_{i=1}^N|x_i-y_i|^p\Big)^{\frac{1}{p}}$$

(důkaz trojúhelníkové nerovnosti využívající obecnou Youngovu nerovnost $\alpha\beta\le\frac{\alpha^p}{p}+\frac{\beta^{\frac{p}{p-1}}}{\frac{p}{p-1}}$ platnou pro $\alpha,\beta\ge 0$ a $p\in(1,\infty)$ uvedeme v kapitole o Lebesgueově integrálu pro obecnější výsledek zvaný Minkowského nerovnost).
(iii) Je-li $[a,b]\subset\mathbb{R}$, na prostoru spojitých funkcí $C([a,b])$ lze zavést třeba metriky

$$\varrho_\infty(f,g)=\max_{[a,b]}|f-g|\qquad\text{a}\qquad \varrho_p(f,g)=\Big(\int_a^b|f-g|^p\,\mathrm{d}x\Big)^{\frac{1}{p}}\quad\text{pro } p\in[1,\infty)$$

(pro $p\in(1,\infty)$ je u metrik ϱ_p je opět poměrně obtížné dokázat trojúhelníkovou nerovnost, která se nazývá *Minkowského nerovnost*).
(iv) Představme si ještě některé prostory posloupností a jejich metriky

$$\ell_\infty=\{\{x_i\}\subset\mathbb{R}\colon \exists C>0\quad |x_i|\le C\,\forall i\in\mathbb{N}\},\qquad \varrho_\infty(\{x_i\},\{y_i\})=\sup_{i\in\mathbb{N}}|x_i-y_i|,$$

$$\ell_p=\Big\{\{x_i\}\subset\mathbb{R}\colon \sum_{i=1}^\infty|x_i|^p<\infty\,\forall i\in\mathbb{N}\Big\},\qquad \varrho_p(\{x_i\},\{y_i\})=\Big(\sum_{i=1}^\infty|x_i-y_i|^p\Big)^{\frac{1}{p}}.$$

Opět zde uvažujeme $p\in[1,\infty)$ a pro $p\ne 1,2$ je důkaz trojúhelníkové nerovnosti obtížný (a plyne z Minkowského nerovnosti).

(v) Na jakékoliv množině splňuje vlastnosti metriky *diskrétní metrika*

$$\varrho(x,y) = \begin{cases} 1 & \text{pro } x \neq y \\ 0 & \text{pro } x = y. \end{cases}$$

(vi) Vlastnosti metriky má také třeba vzdušná vzdálenost dvou míst na mapě, doba chůze mezi dvěma body, cena jízdného na železnici, atd. Ale například doba jízdy autem mezi dvěma body nemusí být metrika (kvůli existenci jednosměrných ulic nemusí být symetrická).

Poznámka 11.1.3. Je přirozené se ptát, zda má nějaký smysl zavádět více než jednu metriku na dané množině, jak jsme to udělali třeba na $\mathbb{R}^N$. Zde je vhodné poznamenat, že obecně různé metriky dávají různé pojmy konvergence, a proto každá metrika na daném prostoru nemusí odpovídat studovanému problému. Na druhou stranu, později si ukážeme, že třeba na $\mathbb{R}^N$ dávají metriky ϱ_2 a ϱ_∞ stejný pojem limity. Metrika ϱ_2 se používá nejčastěji, ale díky odmocnině v její definici se s ní nepracuje příliš pohodlně. Proto ji bývá někdy výhodné nahradit metrikou ϱ_∞.

Naším cílem není jen pojem limity, ale i hlubší výsledky pracující s (vícedimenzionální) analogií pojmu derivace (či přesněji diferenciál). Takové výsledky již potřebují, aby množina P měla lineární strukturu a abychom pracovali s metrikou, která je s touto lineární strukturou kompatibilní. Proto zavádíme ještě další pojem.

Definice 11.1.4 (Norma a normovaný lineární prostor)**.** Nechť V je vektorovým prostorem nad $\mathbb{R}$ (či $\mathbb{C}$). Zobrazení $\|\cdot\|\colon V \to [0,\infty)$ se nazývá *norma*, jestliže pro všechna $u, v \in V$ a $\lambda \in \mathbb{R}$ $(\mathbb{C})$ splňuje

$$\begin{aligned} &\text{(i) } \|u\| \geq 0 \quad \text{a} \quad \|u\| = 0 \Longleftrightarrow u = 0 \\ &\text{(ii) } \|\lambda u\| = |\lambda|\|u\| \\ &\text{(iii) } \|u+v\| \leq \|u\| + \|v\|. \end{aligned} \tag{11.1.1}$$

Dvojice $(V, \|\cdot\|)$ se pak nazývá *normovaný lineární prostor*.

Poznámka 11.1.5. Domluvme se, že počátek (nulový prvek) ve vektorových prostorech budeme zapisovat zkráceně jako 0.

Snadno se dokáže následující výsledek.

Tvrzení 11.1.6. *Nechť $(V, \|\cdot\|)$ je normovaný lineární prostor. Pak zobrazení $\varrho\colon V \times V \to [0,\infty)$ definované předpisem*

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

je metrika na V.

Příklad 11.1.7. (i) Na $\mathbb{R}^N$ máme třeba normy

$$\|x\|_p = \Big(\sum_{i=1}^{N} |x_i|^p\Big)^{\frac{1}{p}} \quad \text{pro } p \in [1,\infty) \quad \text{a} \quad \|x\|_\infty = \max_{i\in\{1,\dots,N\}} |x_i|.$$

(ii) Na $C([a,b])$ mají vlastnosti normy

$$\|f\|_\infty = \max_{[a,b]} |f| \qquad \text{a} \qquad \|f\|_p = \Big(\int_a^b |f|^p \,\mathrm{d}x\Big)^{\frac{1}{p}} \quad \text{pro } p \in [1,\infty).$$

(iii) Na ℓ_∞ je normou $\|x\|_\infty = \sup_{i\in\mathbb{N}} |x_i|$ (zkráceně píšeme $x := \{x_i\}$) a pro $p \in [1,\infty)$ je na ℓ_p normou

$$\|x\|_p = \Big(\sum_{i=1}^{\infty} |x_i|^p\Big)^{\frac{1}{p}}.$$

(iv) Neexistuje norma, ze které by vznikla diskrétní metrika (s výjimkou triviálních případů $V = \emptyset$ a $V = \{0\}$).

Poznámka 11.1.8. (i) Snadno se dá nahlédnout, že na $\mathbb{R}^N$ platí

$$\|\cdot\|_\infty \le \|\cdot\|_2 \le \|\cdot\|_1 \le N\|\cdot\|_\infty.$$

Dobrou představu pro porovnání velikostí norem nám dávají *jednotkové sféry* v jednotlivých normách, tedy množiny $\{x \in \mathbb{R}^N : \|x\| = 1\}$. Pro případ $N = 2$ jsou zobrazeny na Obrázku 11.1.
(ii) Není těžké ukázat, že

$$1 \le p < q \le \infty \qquad \Longrightarrow \qquad \|\cdot\|_p \ge \|\cdot\|_q.$$

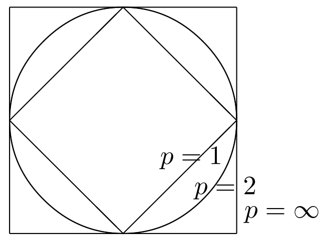
(Návod: díky vlastnosti (ii) v definici normy stačí uvažovat jen body tvaru $(1,t)$ a $(t,1)$, kde $t > 0$. Dále funkce $p \mapsto (1+t^p)^{\frac{1}{p}}$ je klesající, což zjistíme zderivováním.)
(iii) V nekonečné dimenzi jsou vztahy mezi normami často složitější. Kupříkladu pro funkce $f_n = n\chi_{(0,\frac{1}{n})}$ a $g_n = \frac{1}{n}\chi_{(0,n)}$ platí

$$\|f_n\|_1 = 1 \quad \text{ale } \|f_n\|_\infty = n \to \infty \qquad \text{a} \qquad \|g_n\|_1 = 1 \quad \text{ale } \|g_n\|_\infty = \frac{1}{n} \to 0.$$

Mezi normovanými lineárními prostory budeme pracovat nejraději s těmi, které umožňují zavedení *skalárního součinu*, což nám umožní používat Pythagorovu větu, její zobecnění a důsledky.

Definice 11.1.9 (Skalární součin)**.** Nechť V je vektorový prostor. Potom zobrazení $(\cdot,\cdot)\colon V \times V \to \mathbb{R}$ ($\mathbb{C}$) nazveme *skalárním součinem*, jestliže pro všechna $x, y \in V$ a $\lambda \in \mathbb{R}$ ($\mathbb{C}$) platí

$$\begin{aligned} &\text{(i) } (x,x) \ge 0 \quad \text{a} \quad (x,x) = 0 \Longleftrightarrow x = 0 \\ &\text{(ii) } (x,y) = \overline{(y,x)} \\ &\text{(iii) } (\lambda x, y) = \lambda(x,y) \quad \text{a} \quad (x+y,z) = (x,z) + (y,z). \end{aligned} \tag{11.1.2}$$



Obrázek 11.1: Tvar jednotkové sféry v závislosti na volbě normy na $\mathbb{R}^2$.

Poznámka 11.1.10. (i) Obecně platí

$$(x, \lambda y) = \overline{(\lambda y, x)} = \overline{\lambda (y, x)} = \bar{\lambda}\overline{(y, x)} = \bar{\lambda}(x, y)$$

a

$$(u, v + w) = \overline{(v + w, u)} = \overline{(v, u) + (w, u)} = \overline{(v, u)} + \overline{(w, u)} = (u, v) + (u, w).$$

(ii) Nad reálným vektorovým prostorem máme samozřejmě jednodušší vztahy

$$(x, y) = (y, x) \qquad \text{a} \qquad (x, \lambda y) = \lambda (x, y).$$

Věta 11.1.11 (Cauchy–Schwarzova nerovnost)**.** *Nechť P je prostor se skalárním součinem. Pak pro všechna $x, y \in P$ platí*

$$|(x, y)| \leq \sqrt{(x, x)}\sqrt{(y, y)}.$$

Důkaz. Pokud $y = 0$, tvrzení zřejmě platí. Nechť je v dalším $y \neq 0$ (tedy $(y, y) > 0$). Pro všechna $\lambda \in \mathbb{R}$ ($\mathbb{C}$) máme

$$0 \leq (x - \lambda y, x - \lambda y) = (x, x) - \lambda \overline{(x, y)} - \bar{\lambda}(x, y) + \lambda \bar{\lambda}(y, y).$$

Speciálně volba $\lambda = \frac{(x,y)}{(y,y)}$ dává

$$\begin{aligned} 0 &\leq (x, x) - \frac{(x, y)\overline{(x, y)}}{(y, y)} - \frac{\overline{(x, y)}(x, y)}{(y, y)} + \frac{(x, y)\overline{(x, y)}}{(y, y)} = (x, x) - \frac{(x, y)\overline{(x, y)}}{(y, y)} \\ &= (x, x) - \frac{|(x, y)|^2}{(y, y)}, \end{aligned}$$

odkud plyne dokazovaná nerovnost. $\square$

Věta 11.1.12 (Skalární součin generuje normu)**.** *Nechť P je prostor se skalárním součinem. Pak zobrazení $\|\cdot\| \colon V \to [0, \infty)$ definované předpisem*

$$\|x\| := \sqrt{(x, x)}$$

je normou na P.

Důkaz. Ověření prvních dvou vlastností normy je snadné cvičení. Ověřme ještě trojúhelníkovou nerovnost za pomoci Cauchy–Schwarzovy nerovnosti

$$\begin{aligned}\|x+y\|^2 &= (x+y, x+y) = \|x\|^2 + \|y\|^2 + (x,y) + \overline{(x,y)} \\ &= \|x\|^2 + \|y\|^2 + 2\operatorname{Re}(x,y) \le \|x\|^2 + \|y\|^2 + 2|(x,y)| \\ &\le \|x\|^2 + \|y\|^2 + 2\|x\|\|y\| = (\|x\| + \|y\|)^2.\end{aligned}$$

□

Poznámka 11.1.13. Pro právě zavedenou normu má Cauchy–Schwarzova nerovnost tvar $|(x,y)| \le \|x\|\|y\|$.

Poznámka 11.1.14. Je-li P prostor se skalárním součinem, snadno se ověří, že platí *rovnoběžníkové pravidlo*

$$\|x+y\|^2 + \|x-y\|^2 = 2(\|x\|^2 + \|y\|^2) \qquad \forall x, y \in P.$$

Naopak, je-li P reálný normovaný lineární prostor, jehož norma splňuje rovnoběžníkové pravidlo, lze na něm definovat skalární součin předpisem

$$(x,y) := \frac{1}{4}(\|x+y\|^2 - \|x-y\|^2).$$

V komplexním případě je situace složitější.

Příklad 11.1.15. (i) Na $\mathbb{R}^N$ či $\mathbb{C}^N$ má vlastnost skalárního součinu $(x,y) = \sum_{i=1}^N x_i \bar{y}_i$ (díky tomu je První Cauchy–Schwarzova nerovnost z úvodní kapitoly (Tvrzení 2.2.44) jen speciálním případem výše dokázané Cauchy–Schwarzovy nerovnosti). Podobně ℓ_2 je prostor se skalárním součinem $(x,y) = \sum_{i=1}^\infty x_i \bar{y}_i$.
(ii) Na $C([a,b])$ má vlastnosti skalárního součinu zobrazení

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

Poznámka 11.1.16. Trojúhelníková nerovnost normy a metriky mají za následek nerovnosti

$$|\|x\| - \|y\|| \le \|x-y\|$$

a

$$|\varrho(x,y) - \varrho(x,z)| \le \varrho(y,z).$$

První nerovnost plyne z odhadů

$$\|x\| - \|y\| = \|x-y+y\| - \|y\| \le \|x-y\| + \|y\| - \|y\| = \|x-y\|$$

a

$$\|y\| - \|x\| = \|y-x+x\| - \|x\| \le \|y-x\| + \|x\| - \|x\| = \|x-y\|.$$

Druhá zase z nerovností

$$\varrho(x,y) - \varrho(x,z) \le \varrho(x,z) + \varrho(z,y) - \varrho(x,z) = \varrho(y,z)$$

a

$$\varrho(x,z) - \varrho(x,y) \le \varrho(x,y) + \varrho(y,z) - \varrho(x,y) = \varrho(y,z).$$

11.2 Konvergence posloupnosti v metrickém prostoru

Definice 11.2.1 (Konvergence v metrickém prostoru)**.** Nechť (P, ϱ) je metrický prostor a $\{x_n\} \subset P$ je posloupnost. Řekneme, že x_n *konverguje* k $x \in P$ pro $n \to \infty$, jestliže $\varrho(x_n, x) \to 0$. V takovém případě píšeme $x_n \to x$.

Cvičení 11.2.2. Dokažte si, že posloupnost v metrickém prostoru může mít nejvýše jednu limitu.

Příklad 11.2.3. (i) Na $\mathbb{R}$ s metrikou generovanou absolutní hodnotou dostáváme obvyklou definici konvergence posloupnosti reálných čísel.
(ii) Pokud však na $\mathbb{R}$ uvážíme diskrétní metriku, ke konvergenci dochází právě tehdy, když se členy posloupnosti od jistého indexu rovnají limitní hodnotě.

Poznámka 11.2.4. (i) V případě normovaného lineárního prostoru (a metriky odpovídající normě) konvergence posloupnosti prvků znamená

$$\|x_n - x\| = \varrho(x_n, x) \to 0.$$

Bývá tedy zvykem rovněž říkat, že x_n konvergují k x *v normě*.
(ii) V případě normovaného lineárního prostoru je totéž $x_n \to x$ a $x_n - x \to 0$. Umíme tedy jakoukoliv konvergenci převést na konvergenci k nulovému prvku.
(iii) Bude-li se na metrickém prostoru nabízet více metrik, k údaji $x_n \to x$ budeme ještě doplňovat, ve které metrice tuto konvergenci uvažujeme.

Příklad 11.2.5. (i) Není těžké ověřit, že konvergence na $\mathbb{R}^N$ v kterékoliv z norem $\|\cdot\|_p$, kde $p \in [1, \infty]$, znamená, že jednotlivé složky bodů x_n konvergují k odpovídající složce bodu x v $\mathbb{R}$, tedy

$$x_n \to x \qquad \Longleftrightarrow \qquad (x_n)_i \to (x)_i \, \forall i \in \{1, \dots, N\}.$$

(ii) Konvergence na $C([a, b])$ v normě $\|\cdot\|_\infty$ se nazývá *stejnoměrná konvergence.* Této konvergenci se budeme časem věnovat podrobně, neboť má řadu pěkných vlastností (například se dozvíme, že stejnoměrná limita posloupnosti spojitých funkcí je rovněž spojitá funkce), a proto se s ní příjemně pracuje.

První část předchozího příkladu ukázala, že v některých případech mohou dát různé normy stejný typ konvergence. Tomuto jevu se nyní budeme věnovat podrobněji.

Definice 11.2.6 (Ekvivalentní normy a metriky)**.** Nechť ϱ_1, ϱ_2 jsou metriky na P. Řekneme, že tyto metriky jsou *ekvivalentní*, jestliže existují $c_1, c_2 > 0$ takové, že na P platí

$$c_1 \varrho_1(x, y) \le \varrho_2(x, y) \le c_2 \varrho_1(x, y).$$

Analogicky se definují ekvivalentní normy podmínkou

$$c_1 \|x\|_1 \le \|x\|_2 \le c_2 \|x\|_1.$$

Poznámka 11.2.7. Ekvivalence dvou norem zřejmě implikuje ekvivalenci jim odpovídajících metrik.

Příklad 11.2.8. (i) Výše jsme si ukázali, že všechny normy, které jsme si představili na $\mathbb{R}^N$, jsou ekvivalentní.
(ii) Na ℓ_∞ mají jednotkovou normu posloupnosti

$$\{1,0,0,\dots\},\quad \{1,1,0,0,\dots\},\quad \{1,1,1,0,0,\dots\},\quad \dots.$$

Norma těchto posloupností v prostoru ℓ_1 však odpovídá počtu jedniček v dané posloupnosti. Proto ℓ_1-norma a ℓ_∞-norma nemohou být ekvivalentní.

Věta 11.2.9 (Ekvivalentní metriky generují stejnou konvergenci)**.** *Nechť P je množina a ϱ_1, ϱ_2 jsou dvě ekvivalentní metriky na P. Pak pro každou posloupnost $\{x_n\} \subset P$ a $x \in P$ platí*

$$x_n \to x \text{ v metrice } \varrho_1 \qquad \Longleftrightarrow \qquad x_n \to x \text{ v metrice } \varrho_2.$$

Důkaz. Ekvivalence metrik má za následek, že pro všechna $\varepsilon > 0$ existuje $\tilde{\varepsilon} > 0$ tak že

$$\varrho_2(x_n, x) < \tilde{\varepsilon} \qquad \Longrightarrow \qquad \varrho_1(x_n, x) < \varepsilon.$$

Odtud konvergence v metrice ϱ_2 implikuje konvergenci v metrice ϱ_1. Obrácená implikace se dokáže analogicky. □

Vztah mezi konvergencí v normě a po složkách, kterou jsme pozorovali v příkladu výše, platí obecně ve všech konečnědimenzionálních lineárních prostorech.

Věta 11.2.10 (O vztahu konvergence v normě ke konvergenci po složkách)**.** *Nechť P je normovaný lineární prostor, $\{e_1, \dots, e_N\}$ je jeho báze, $x_n = \sum_{i=1}^N \alpha_i^n e_i$ a $x = \sum_{i=1}^N \alpha_i e_i$, kde $\alpha_i^n, \alpha_i \in \mathbb{R}$ ($\mathbb{C}$) pro $i \in \{1, \dots, N\}$ a pro $n \in \mathbb{N}$. Pak*

$$x_n \to x \qquad \Longleftrightarrow \qquad \alpha_i^n \to \alpha_i \quad \forall i \in \{1, \dots, N\}.$$

Důkaz. „$\Leftarrow$“ Tato implikace plyne z odhadu (využijeme trojúhelníkovou nerovnost)

$$\|x_n - x\| = \left\| \sum_{i=1}^N (\alpha_i^n - \alpha_i) e_i \right\| \le \sum_{i=1}^N |\alpha_i^n - \alpha_i| \|e_i\| \le \max_{i \in \{1,\dots,N\}} \|e_i\| \sum_{i=1}^N |\alpha_i^n - \alpha_i|.$$

„$\Rightarrow$“ Bez újmy na obecnosti můžeme předpokládat, že $x = 0$. Pro spor předpokládejme, že $x_n = \sum_{i=1}^N \alpha_i^n e_i \to 0$, ale alespoň pro jedno $i \in \{1, \dots, N\}$ máme $\alpha_i^n \nrightarrow 0$. Odtud dostáváme existenci $\delta > 0$ takového, že pro nekonečně mnoho indexů n platí

$$\max\{|\alpha_1^n|, |\alpha_2^n|, \dots, |\alpha_N^n|\} \ge \delta.$$

Přechodem k podposloupnosti dosáhneme toho, že pro jisté $j \in \{1, \dots, N\}$ a všechna $n \in \mathbb{N}$ máme

$$|\alpha_j^n| = \max\{|\alpha_1^n|, |\alpha_2^n|, \dots, |\alpha_N^n|\} \ge \delta.$$

Definujme nyní

$$y_n = \sum_{i=1}^{N} \beta_i^n e_i := \sum_{i=1}^{N} \frac{\alpha_i^n}{|\alpha_j^n|} e_i = \frac{1}{|\alpha_j^n|} x_n.$$

Odtud $\beta_j^n = 1$ pro všechna $n \in \mathbb{N}$ a $\beta_i^n \in [-1,1]$ pro všechna $n \in \mathbb{N}$ a $i \neq j$. Aplikujeme-li $(n-1)$-násobně Weierstrassovu větu, po přechodu k podposloupnosti β_i^n konvergují k $\beta_i \in [-1,1]$ pro všechna $i \in \{1,\ldots,N\}$, speciálně $\beta_j = 1$. Odtud podle již dokázané implikace „$\Leftarrow$“ máme

$$y_n \to y \neq 0.$$

Na druhou stranu, z definice bodů y_n a vlastnosti $x_n \to 0$ plyne

$$\|y_n\| = \left\| \frac{x_n}{|\alpha_j^n|} \right\| \leq \left\| \frac{x_n}{\delta} \right\| \to 0$$

a máme spor. □

Věta 11.2.11 (O ekvivalenci norem v konečné dimenzi)**.** *Na konečnědimenzionálním lineárním prostoru jsou libovolné dvě normy ekvivalentní.*

Důkaz. Pokud by tomu tak nebylo, našli bychom posloupnost netriviálních prvků $\{x_n\}$ a dvojici norem $\|\cdot\|_1, \|\cdot\|_2$ tak, že $\|x_n\|_1 \geq n\|x_n\|_2$. Bez újmy na obecnosti můžeme předpokládat, že $\|x_n\|_1 = 1$ pro všechna $n \in \mathbb{N}$ (jinak přejdeme k prvkům $\frac{x_n}{\|x_n\|_1}$). Odtud

$$\|x_n\|_2 \leq \frac{1}{n} \to 0.$$

Podle předchozí věty dostáváme konvergenci k nulovému prvku po složkách a další aplikací předchozí věty obdržíme $\|x_n\|_1 \to 0$, což je ve sporu s $\|x_n\|_1 \equiv 1$. □

Věta 11.2.12 (O spojitosti metriky, normy a skalárního součinu)**.** (i) *Jestliže $x_n \to x$ a $y_n \to y$ v metrickém prostoru (P,ϱ), pak $\varrho(x_n,y_n) \to \varrho(x,y)$.*
(ii) *Jestliže $x_n \to x$ v normovaném lineárním prostoru, pak $\|x_n\| \to \|x\|$.*
(iii) *Jestliže $x_n \to x$ a $y_n \to y$ v prostoru se skalárním součinem, pak $(x_n,y_n) \to (x,y)$.*

Důkaz. Druhé tvrzení plyne z poznámky o důsledcích trojúhelníkové nerovnosti pro normu (Poznámka 11.1.16). První část se dokazuje podobně, neboť máme (poznámku o důsledcích trojúhelníkové nerovnosti pro metriku použijeme v druhém kroku odhadu)

$$\begin{aligned}|\varrho(x_n,y_n) - \varrho(x,y)| &= |\varrho(x_n,y_n) - \varrho(x_n,y) + \varrho(x_n,y) - \varrho(x,y)| \\ &\leq |\varrho(x_n,y_n) - \varrho(x_n,y)| + |\varrho(x_n,y) - \varrho(x,y)| \\ &\leq \varrho(y_n,y) + \varrho(x_n,x) \to 0.\end{aligned}$$

Důkaz třetího tvrzení využívá Cauchy–Schwarzovu nerovnost a omezenost $\|x_n\|$ (plyne z (ii))

$$\begin{aligned}|(x_n,y_n) - (x,y)| &= |(x_n,y_n) - (x_n,y) + (x_n,y) - (x,y)| \\ &\leq |(x_n,y_n) - (x_n,y) + (x_n,y) - (x,y)| \\ &\leq \|x_n\|\|y_n - y\| + \|x_n - x\|\|y\| \to 0.\end{aligned}$$

□

Důsledek 11.2.13. *Konvergentní posloupnost v normovaném lineárním prostoru má omezené normy.*

11.3 Základní vlastnosti podmnožin metrického prostoru

Definice 11.3.1 (Okolí v metrickém prostoru)**.** Nechť (P, ϱ) je metrický prostor, $x_0 \in P$ a $\varepsilon > 0$. Množina

$$\mathcal{U}_\varepsilon(x_0) := \{x \in P \colon \varrho(x, x_0) < \varepsilon\}$$

se nazývá ε-ovým *okolím* bodu x_0. *Prstencové* ε-ové *okolí* bodu x_0 definujeme jako

$$\mathcal{P}_\varepsilon(x_0) := \mathcal{U}_\varepsilon(x_0) \setminus \{x_0\}.$$

Příklad 11.3.2. (i) Bereme-li na $\mathbb{R}$ metriku generovanou absolutní hodnotou, předchozí definice dává nám známý pojem okolí.
(ii) Podobně na $\mathbb{C}$ s obvyklou eukleidovskou vzdáleností.
(iii) Při diskrétní metrice máme pro $\varepsilon \leq 1$

$$\mathcal{U}_\varepsilon(x_0) = \{x_0\} \qquad \text{a} \qquad \mathcal{P}_\varepsilon(x_0) = \emptyset.$$

Pokud je však $\varepsilon > 1$, dostáváme

$$\mathcal{U}_\varepsilon(x_0) = P \qquad \text{a} \qquad \mathcal{P}_\varepsilon(x_0) = P \setminus \{x_0\}.$$

(iv) Na $\mathbb{R}^N$ s normou $\|\cdot\|_2$ mají okolí tvar koulí, s normou $\|\cdot\|_\infty$ mají tvar krychlí.

Definice 11.3.3 (Otevřené a uzavřené množiny)**.** Nechť (P, ϱ) je metrický prostor. Množina $G \subset P$ se nazývá *otevřená*, jestliže ke každému jejímu bodu existuje okolí, které leží v G. Množina $F \subset P$ se nazývá *uzavřená*, jestliže je doplňkem otevřené množiny.

Příklad 11.3.4. (i) Na $\mathbb{R}$ s obvyklou metrikou je každý otevřený interval otevřená množina. Skutečně, je-li $x \in (a, b)$, kde $a, b \in \mathbb{R}$, máme $(x - \varepsilon, x + \varepsilon) \subset (a, b)$, kdykoliv $\varepsilon \leq \min\{x - a, b - x\}$. Pokud je některá z mezí nevlastní, je důkaz ještě jednodušší.
(ii) Každý omezený uzavřený interval v $\mathbb{R}$ je uzavřená množina. Uzavřené množiny jsou dále intervaly typu $[a, \infty)$ a $(-\infty, b]$.

Cvičení 11.3.5. (i) Rozmyslete si, které množiny jsou otevřené a které uzavřené na $\mathbb{R}$ s diskrétní metrikou.
(ii) Rozmyslete si, které množiny jsou otevřené a které uzavřené na $(-1, 1)$ s obvyklou metrikou.
(iii) Rozmyslete si, že $(0, 1)^N \subset \mathbb{R}^N$ je otevřená jak v normě $\|\cdot\|_2$, tak v normě $\|\cdot\|_\infty$ (ke každému bodu z $(0, 1)^N$ umíme sestrojit jak kouli tak krychličku, že jsou v tomto bodě centrované a leží v $(0, 1)^N$).

Poznámka 11.3.6. (i) V každém metrickém prostoru (P, ϱ) jsou $\emptyset$ a P otevřené množiny. Následně jsou $\emptyset$ a P také uzavřené. Takovýmto množinám se říká *obojetné*. Obecně mohou existovat i jiné obojetné množiny než $\emptyset$ a P (uvažte třeba $P = (0,1) \cup (2,3)$ s obvyklou metrikou).
(ii) Okolí je vždy otevřená množina, což snadno plyne z trojúhelníkové nerovnosti.
(iii) Existují množiny (a v obvykle používaných metrických prostorech je jich dokonce většina), které nejsou ani otevřené ani uzavřené (na $\mathbb{R}$ s obvyklou metrikou uvažte třeba $[0,1)$).
(iv) Podle definice je množina uzavřená právě tehdy, když je doplňkem otevřené. Podobně množina je otevřená právě tehdy, když je doplňkem uzavřené (neboť $P \setminus (P \setminus G) = G$).

Poznámka 11.3.7. Otevřené množiny jsou důležité například při budování diferenciálního počtu (vzpomeňte si, že při definici derivace jsme potřebovali mít funkci definovanou na okolí všech bodů, kde jsme derivaci počítali). Uzavřené množiny jsou zase důležité při studiu extrémů (vzpomeňte si na Větu o existenci extrémů (Věta 6.1.7), tedy že spojitá funkce nabývá na omezeném uzavřeném intervalu svého maxima i minima). Oba typy množin jsou „hezké“ v teorii Lebesgueova integrálu a jim odpovídající charakteristické funkce bude možné integrovat (analogie k tomu, jak si teorie Riemannova integrálu dobře rozuměla s omezenými spojitými funkcemi).

Určování otevřenosti či uzavřenosti množin se často neprovádí z definice, ale pomocí vhodných nástrojů. Ty si zde budeme postupně uvádět.

Věta 11.3.8 (O sjednocení a průniku otevřených a uzavřených množin)**.** *Sjednocení libovolného systému otevřených množin je otevřené, průnik konečného systému otevřených množin je otevřený. Průnik libovolného systému uzavřených množin je uzavřený, sjednocení konečného systému uzavřených množin je uzavřené.*

Důkaz. Nechť I je indexová množina a G_α je otevřená pro každé $\alpha \in I$. Nechť $x_0 \in \bigcup_{\alpha \in I} G_\alpha$. Pak existuje $\alpha_0 \in I$ tak, že $x_0 \in G_{\alpha_0}$. Protože G_{α_0} je otevřená, dostáváme $\varepsilon > 0$ takové, že

$$\mathcal{U}_\varepsilon(x_0) \subset G_{\alpha_0} \subset \bigcup_{\alpha \in I} G_\alpha.$$

Odtud $\bigcup_{\alpha \in I} G_\alpha$ je otevřená.

Pokud je I konečná a $x_0 \in \bigcap_{\alpha \in I} G_\alpha$, pak pro každé $\alpha \in I$ existuje $\varepsilon_\alpha > 0$ tak, že

$$\mathcal{U}_{\varepsilon_\alpha}(x_0) \subset G_\alpha.$$

Navíc, protože I je konečná, existuje kladné

$$\varepsilon := \min_{\alpha \in I} \varepsilon_\alpha.$$

Odtud

$$\mathcal{U}_\varepsilon(x_0) \subset \mathcal{U}_{\varepsilon_\alpha}(x_0) \subset G_\alpha \quad \forall \alpha \in I \qquad \Longrightarrow \qquad \mathcal{U}_\varepsilon(x_0) \subset \bigcap_{\alpha \in I} G_\alpha.$$

Proto $\bigcap_{\alpha\in I} G_\alpha$ je otevřená.

Nechť F_α je uzavřená pro každé $\alpha \in I$ a $x_0 \notin \bigcap_{\alpha\in I} F_\alpha$. Pak existuje $\alpha_0 \in I$ tak, že $x_0 \notin F_{\alpha_0}$. Protože $P \setminus F_{\alpha_0}$ je otevřená, dostáváme $\varepsilon > 0$ takové, že

$$\mathcal{U}_\varepsilon(x_0) \subset P \setminus F_{\alpha_0} \subset P \setminus \bigcap_{\alpha\in I} F_\alpha.$$

Odtud $P \setminus \bigcap_{\alpha\in I} F_\alpha$ je otevřená a $\bigcap_{\alpha\in I} F_\alpha$ je uzavřená.

Pokud je I konečná a $x_0 \in P \setminus \bigcup_{\alpha\in I} F_\alpha$, pak pro každé $\alpha \in I$ existuje $\varepsilon_\alpha > 0$ tak, že

$$\mathcal{U}_{\varepsilon_\alpha}(x_0) \subset P \setminus F_\alpha.$$

Navíc, protože I je konečná, existuje kladné

$$\varepsilon := \min_{\alpha\in I} \varepsilon_\alpha.$$

Odtud

$$\mathcal{U}_\varepsilon(x_0) \subset \mathcal{U}_{\varepsilon_\alpha}(x_0) \subset P \setminus F_\alpha \quad \forall \alpha \in I \qquad \Longrightarrow \qquad \mathcal{U}_\varepsilon(x_0) \subset P \setminus \bigcup_{\alpha\in I} F_\alpha.$$

Proto $\bigcup_{\alpha\in I} F_\alpha$ je uzavřená. $\square$

Poznámka 11.3.9. Tvrzení o uzavřených množinách šla také dokázat za pomoci tvrzení o otevřených množinách a de Morganových vzorců.

Příklad 11.3.10. Spočetný systém otevřených množin $G_i = (-\frac{1}{i}, \frac{1}{i})$, $i \in \mathbb{N}$, má jednobodový průnik $\{0\}$, což není otevřená množina. Spočetný systém uzavřených množin $F_i = [\frac{1}{i}, 1-\frac{1}{i}]$, $i \in \mathbb{N}\setminus\{1\}$, má sjednocení $(0,1)$, což není uzavřená množina.

Cvičení 11.3.11. Nechť F je uzavřená a G je otevřená. Ukažte, že $F \setminus G$ je uzavřená a $G \setminus F$ je otevřená.

Představme si ještě další základní pojmy.

Definice 11.3.12 (Vnitřní, vnější a hraniční body)**.** Nechť (P, ϱ) je metrický prostor a $A \subset P$. Řekneme, že bod $x_0 \in A$ je *vnitřním bodem* množiny A, jestliže existuje jeho okolí ležící v A. Bod $x_0 \in P$ se nazývá *vnějším* bodem množiny A, jestliže je vnitřním bodem jejího doplňku. Bod $x_0 \in P$ se nazývá *hraničním* bodem množiny A, jestliže není vnitřním ani vnějším bodem množiny A.

Množina všech vnitřních bodů se nazývá *vnitřek* množiny A a značí se A°. Množina vnějších bodů se nazývá *vnějšek* množiny A, množina hraničních bodů se nazývá *hranice* množiny A a značí se ∂A. Množinu $\bar{A} := A \cup \partial A$ nazýváme *uzávěr* množiny A.

Příklad 11.3.13. (i) Nechť $A = (0,1] \subset \mathbb{R}$ s obvyklou metrikou. Pak

$$A^\circ = (0,1), \quad \bar{A} = [0,1] \quad \text{a} \qquad \partial A = \{0,1\}.$$

(ii) Nechť $A = (0,1) \cap \mathbb{Q} \subset \mathbb{R}$ s obvyklou metrikou. Pak

$$A^\circ = \emptyset, \quad \bar{A} = [0,1] \quad \text{a} \qquad \partial A = [0,1].$$

Věta 11.3.14 (Charakterizace vnitřku a uzávěru pomocí inkluze). *Nechť (P, ϱ) je metrický prostor a $A \subset P$. Pak A° je největší otevřená podmnožina A a $\bar{A}$ je nejmenší uzavřená nadmnožina A.*

Důkaz. Nejprve ukažme, že A° je otevřená. Zvolme $x_0 \in A^\circ$. Podle definice existuje $\varepsilon > 0$ takové, že $\mathcal{U}_\varepsilon(x_0) \subset A$. Zvolme libovolné $x \in \mathcal{U}_\varepsilon(x_0)$. Pak pro $\delta := \varepsilon - \varrho(x, x_0)$ platí

$$y \in \mathcal{U}_\delta(x) \quad \Longrightarrow \quad \varrho(y, x_0) \le \varrho(y, x) + \varrho(x, x_0) < \varepsilon \quad \Longrightarrow \quad y \in \mathcal{U}_\varepsilon(x_0) \subset A.$$

Odtud $\mathcal{U}_\delta(x) \subset A$, a proto $x \in A^\circ$. Protože $x \in \mathcal{U}_\varepsilon(x_0)$ bylo libovolné, máme $\mathcal{U}_\varepsilon(x_0) \subset A^\circ$. Protože $x_0 \in A^\circ$ bylo libovolné, A° je otevřená.

Dále zřejmě platí $A^\circ \subset A$. Konečně, pokud je $G \subset A$ otevřená, každý její bod má okolí ležící v $G \subset A$, tedy tento bod je vnitřním bodem A, a musí proto ležet v A°. Tedy $G \subset A^\circ$.

Podle definice je $\bar{A} = P \setminus (P \setminus A)^\circ$. Je to tedy doplněk největší otevřené podmnožiny $P \setminus A$, z čehož plynou dokazované vlastnosti (uzavřenost a $A \subset \bar{A}$ jsou jasné; pokud by existovala uzavřená $F \supset A$ splňující $\bar{A} \setminus F \neq \emptyset$, pak by $P \setminus (\bar{A} \cap F)$ dávala spor s první částí tvrzení aplikovanou na $P \setminus A$). □

Poznámka 11.3.15. Podobně jako je vnitřek otevřená množina, je i vnějšek (vnitřek doplňku) otevřená množina. Hranice je proto uzavřená (doplněk sjednocení dvou otevřených množin).

Věta 11.3.16 (Charakterizace hranice pomocí okolí). *Nechť (P, ϱ) je metrický prostor a $A \subset P$. Pak $x \in \partial A$ právě tehdy, když v každém jeho okolí leží alespoň jeden bod z A a alespoň jeden bod z $P \setminus A$.*

Důkaz. „$\Rightarrow$“ Pokud by existovalo okolí x, kde by chyběl bod z doplňku, platilo by $x \in A^\circ$, tedy $x \notin \partial A$. Analogicky pro případ s okolím neprotínajícím A.
„$\Leftarrow$“ Pokud platí výrok napravo, nemohou platit výroky $x \in A^\circ$ a $x \in (P \setminus A)^\circ$. Odtud

$$x \in P \setminus (A^\circ \cup (P \setminus A)^\circ) = \partial A.$$

□

Cvičení 11.3.17. Sami si dokažte následující výsledky

$$A \subset B \Longrightarrow \bar{A} \subset \bar{B}, \quad A \subset B \Longrightarrow A^\circ \subset B^\circ, \quad \overline{A \cup B} = \bar{A} \cup \bar{B}, \quad \bar{\bar{A}} = \bar{A},$$
$$\overline{\{x\}} = \{x\}, \quad \partial A = \bar{A} \cap \overline{P \setminus A}, \quad A^\circ = A \setminus \overline{P \setminus A}, \quad (A \cap B)^\circ = A^\circ \cap B^\circ.$$

Definice 11.3.18 (Izolovaný a hromadný bod množiny). Nechť (P, ϱ) je metrický prostor a $A \subset P$. Bod $x_0 \in A$ se nazývá *izolovaný bod* množiny A, jestliže má prstencové okolí neprotínající A. Bod $x_0 \in P$ se nazývá *hromadný bod* množiny A, jestliže každé jeho prstencové okolí protíná A. Množina všech hromadných bodů množiny A se nazývá *derivace* množiny A a značí se A'.

Příklad 11.3.19. (i) Nechť $A = (0,1) \cup \{2\} \subset \mathbb{R}$ s obvyklou metrikou. Pak jediným izolovaným bodem je číslo 2. Dále $A' = [0,1]$.
(ii) Jestliže $A = \mathbb{Q}$, pak $A' = \mathbb{R}$ a žádný bod A není izolovaný.
(iii) V diskrétní metrice jsou všechny body dané množiny izolované.

Poznámka 11.3.20. (i) Izolovaný bod je automaticky hraničním bodem podle Věty o charakterizaci hranice množiny pomocí okolí (Věta 11.3.16).
(ii) Každý bod množiny A je buď jejím hromadným bodem, nebo je izolovaný. Navíc ještě některé body z jejího doplňku mohou být jejími hromadnými body.

Věta 11.3.21 (O vztahu hromadných a hraničních bodů)**.** *Nechť (P, ϱ) je metrický prostor a $A \subset P$. Pak*

$$\partial A \setminus A = A' \setminus A \qquad a \qquad \partial A \cup A = A' \cup A = \bar{A}.$$

Důkaz. Dokažme inkluzi „$\subset$“ v první rovnosti. Nechť $x_0 \in \partial A \setminus A$. Pak podle Věty o charakterizaci hranice pomocí okolí (Věta 11.3.16) musí být $x_0 \in A'$ (připomeňme, že $x_0 \notin A$).

Dokažme inkluzi „$\supset$“ v první rovnosti. Nechť $x_0 \in A' \setminus A$. Opět stačí na x_0 aplikovat Větu o charakterizaci hranice pomocí okolí.

Druhá rovnost plyne z první, neboť

$$\partial A \cup A = (\partial A \setminus A) \cup A = (A' \setminus A) \cup A = A' \cup A$$

a podle definice uzávěru je $\partial A \cup A = \bar{A}$. □

Věta 11.3.22 (Charakterizace uzavřené množiny)**.** *Nechť (P, ϱ) je metrický prostor a $A \subset P$. Pak*

$$A \textit{ je uzavřená} \iff A = \bar{A} \iff \partial A \subset A \iff A' \subset A.$$

Důkaz. Je-li $A = \bar{A}$, musí být A uzavřená, neboť uzávěr je vždy uzavřený. Je-li A uzavřená, musí platit $A = \bar{A}$, neboť $\bar{A}$ je nejmenší uzavřená množina obsahující A. Tím je dokázána první ekvivalence. Druhá ekvivalence plyne z definice uzávěru a třetí z Věty o vztahu hromadných a hraničních bodů (Věta 11.3.21). □

V dalším se budeme zabývat vztahem nově zavedených pojmů a konvergence posloupností.

Věta 11.3.23 (Charakterizace hromadných bodů pomocí posloupností)**.** *Nechť (P, ϱ) je metrický prostor a $A \subset P$. Pak*

$$x_0 \in A' \qquad \iff \qquad \exists \{x_n\} \subset A \setminus \{x_0\} \quad x_n \to x_0.$$

Důkaz. „$\Rightarrow$“ Posloupnost zkonstruujeme tak, že bereme $x_n \in A \cap \mathcal{P}_{\frac{1}{n}}(x_0)$ pro každé $n \in \mathbb{N}$.
„$\Leftarrow$“ Tato implikace je zřejmá. □

Věta 11.3.24 (Charakterizace uzávěru pomocí posloupností)**.** *Nechť (P,ϱ) je metrický prostor a $A \subset P$. Pak*

$$x_0 \in \bar{A} \qquad \Longleftrightarrow \qquad \exists\{x_n\} \subset A \quad x_n \to x_0.$$

Důkaz. „$\Rightarrow$“ Pokud $x_0 \in A$, stačí uvážit konstantní posloupnost $x_n \equiv x_0$. Pokud $x_0 \in \partial A$, bereme $x_n \in A \cap \mathcal{U}_{\frac{1}{n}}(x_0)$ pro každé $n \in \mathbb{N}$.
„$\Leftarrow$“ Platí-li výrok na pravé straně, x_0 nemůže být vnějším bodem množiny A. $\square$

Věta 11.3.25 (Charakterizace hranice pomocí posloupností)**.** *Nechť (P,ϱ) je metrický prostor a $A \subset P$. Pak*

$$x_0 \in \partial A \qquad \Longleftrightarrow \qquad \exists\{x_n\} \subset A, \{y_n\} \subset P \setminus A \quad x_n \to x_0, y_n \to x_0.$$

Důkaz. „$\Rightarrow$“ Posloupnosti zkonstruujeme tak, že bereme $x_n \in A \cap \mathcal{U}_{\frac{1}{n}}(x_0)$ a $y_n \in \mathcal{U}_{\frac{1}{n}}(x_0) \setminus A$ pro každé $n \in \mathbb{N}$.
„$\Leftarrow$“ Platí-li výrok na pravé straně, x_0 nemůže být ani vnitřním ani vnějším bodem množiny A. $\square$

Věta 11.3.26 (Charakterizace uzavřenosti pomocí posloupností)**.** *Nechť (P,ϱ) je metrický prostor a $A \subset P$. Pak A je uzavřená právě tehdy, když každá konvergentní (v P) posloupnost prvků z A má limitu v A.*

Důkaz. „$\Rightarrow$“ Je-li A je uzavřená, je $P \setminus A$ otevřená. Každý její bod má proto okolí neprotínající A a není tedy možné k němu dokonvergovat posloupností prvků A.
„$\Leftarrow$“ V tomto případě má každá posloupnost z Věty o charakterizaci uzávěru pomocí posloupností (Věta 11.3.24) limitu z A. Odtud $\bar{A} \subset A$, proto $\bar{A} = A$ a A je uzavřená. $\square$

Poznámka 11.3.27. Pokud je $P = (0,1)$ s obvyklou metrikou pro reálná čísla, je $(0,1)$ uzavřená množina (celý prostor je obojetná množina). Posloupnost $\{\frac{1}{n}\}$ v předchozí větě žádné problémy nezpůsobí, neboť není konvergentní v P. Pokud bychom uvážili případ $P = \mathbb{R}$ a $A = (0,1)$, množina A není uzavřená, což se dá zdůvodnit právě pomocí posloupnosti $\{\frac{1}{n}\}$.

Poznámka 11.3.28. Řadu nových pojmů a jevů jsme si ilustrovali pomocí příkladů, kde jsme používali například diskrétní metriku či třeba prostor $P = (0,1)$. Pro takovéto metriky a prostory nebudeme mít v dalším výkladu využití. Naším hlavním cílem však budou prostory funkcí, kde bývá situace mnohdy značně složitá. Proto jsme raději pro rychlou ilustraci volili zmíněné jednoduché prostory, ačkoliv nemají v matematice a fyzice takový význam.

11.4 Hustota a separabilita

V matematické analýze se velice často přistupuje k tomu, že obecný výsledek dokazujeme jen v méně obecné situaci, která je však pro platnost obecného výroku rozhodující. Kupříkladu k důkazu omezenosti funkčních hodnot shora postačuje znalost hodnoty jejich maxima, při aplikaci Lagrangeovy věty často uvažujeme

„nejhorší“ bod $\xi \in (a, b)$ a aplikace teorie Taylorova polynomu těží z toho, že s polynomy se dobře pracuje a zároveň je to dostatečně bohatá třída funkcí, aby nám poskytla libovolně přesnou aproximaci vyšetřované (typicky reálně analytické) funkce. Právě aproximační schopnosti „hezkých“ prvků studovaného metrického prostoru se kupříkladu v teorii Lebesgueova integrálu či parciálních diferenciálních rovnic používají velice často. To nás vede k následující definici.

Definice 11.4.1 (Hustá podmnožina)**.** Nechť (P, ϱ) je metrický prostor a $A \subset P$. Říkáme, že množina A je *hustá* v P, jestliže v každém okolí každého prvku z P leží prvek z A.

Poznámka 11.4.2. Hustota se dá také charakterizovat tak, že pro každé $x \in P$ existuje posloupnost $\{a_n\} \subset A$ splňující $a_n \to x$. Ekvivalentní je též $\bar{A} = P$.

Příklad 11.4.3. Vezmeme-li $\mathbb{R}$ s obvyklou metrikou, pak jsou zde hustá racionální čísla. Podobně pro iracionální čísla.

Definice 11.4.4 (Separabilní prostor)**.** Řekneme, že metrický prostor je *separabilní*, jestliže v něm existuje spočetná hustá množina.

Poznámka 11.4.5. (i) Separabilita v podstatě znamená, že prostor není příliš velký. Separabilita nevylučuje nekonečnou dimenzi, proto obecně nezaručuje kupříkladu ekvivalenci norem. Přesto si později ukážeme, že separabilní prostory si stále ještě zachovávají řadu příjemných vlastností (oproti prostorům neseparabilním).
(ii) S ohledem na předchozí část poznámky se jeví přirozené v definici používat raději termín „nejvýše spočetná“. Existuje-li však v metrickém prostoru konečná hustá množina, zřejmě musí být celý prostor roven této množině a obsahuje jen konečný počet prvků. To je ale vlastnost, kterou žádný z běžně užívaných metrických prostorů nemá.

Poznámka 11.4.6. Občas budeme hovořit o separabilních množinách. Budeme tím opět myslet, že mají (nejvýše) spočetnou hustou podmnožinu. Případ konečných množin je opět nezajímavý.

Příklad 11.4.7. Množina $\mathbb{R}$ s obvyklou metrikou je separabilní díky hustotě racionálních čísel.

Cvičení 11.4.8. Nechť $n \in N$ a $[a, b] \subset \mathbb{R}$. Dokažte, že v prostoru polynomů stupně nejvýše n na $[a, b]$ je hustá množina polynomů s reálnými koeficienty a stupněm nejvýše n, a proto je původní prostor separabilní.

Lemma 11.4.9 (O separabilitě podmnožin)**.** *V separabilním metrickém prostoru je každá množina separabilní. Obecněji, podmnožina separabilní množiny je separabilní.*

Důkaz. Nechť (P, ϱ) je separabilní metrický prostor a $\{x_n\}$ je jeho spočetná hustá podmnožina a $A \subset P$. Pro každou dvojici $m, n \in \mathbb{N}$ zvolme

$$y_{m,n} \in A \cap \mathcal{U}_{\frac{1}{m}}(x_n),$$

je-li průnik napravo neprázdný, jinak prvek $y_{m,n}$ nedefinujeme. Zřejmě jsme tím definovali nejvýše spočetnou množinu. Hustota systému $\{y_{m,n}\}$ v A plyne nyní z toho, že ke zvolenému $x \in A$ umíme najít libovolně blízký prvek $x_n \in P$ a jemu se zase umíme přinejmenším srovnatelně přesně přiblížit prvkem $y_{m,n}$ (číslo $\frac{1}{m}$ musí být dost malé, ale zase ne tak malé, aby $\mathcal{U}_{\frac{1}{m}}(x_n) \cap A = \emptyset$, čemuž ovšem zabráníme, pokud $\frac{1}{m} > \varrho(x, x_n)$). $\square$

Příklad 11.4.10. Prostor ℓ_∞ není separabilní. Abychom si toto uvědomili, uvažme množinu $\{x_G\}_{G\subset\mathbb{N}} \subset \ell_\infty$, kde i-tý člen posloupnosti x_G definujeme jako 1 pokud $i \in G$ a 0 jinak. Množina $\{x_G\}_{G\subset\mathbb{N}}$ je nespočetná (má mohutnost kontinua, neboť má tolik prvků, kolik je podmnožin reálných čísel) a každá dvojice jejích různých prvků má vzdálenost rovnu 1.

11.5 Hustota polynomů v $C([a,b])$ a separabilita $C([a,b])$

Nyní si ukážeme, že v maximové metrice je spojitou funkci na omezeném uzavřeném intervalu možné libovolně přesně aproximovat polynomem (na $\mathbb{R}$ to nejde, uvažte exponenciálu a to, že pro libovolný polynom P platí $\lim_{x\to\infty}\big(f(x) - P(x)\big) = \infty$). Přípravu začneme aproximací charakteristické funkce intervalu.

Lemma 11.5.1. *Nechť $[a,b] \subset \mathbb{R}$, $c \in (a,b)$, $\eta \in (0,1)$ a $\delta \in (0, c-a)$. Pak existuje polynom T tak, že*

$$T(x) \in \begin{cases} (1-\eta, 1] & \textit{na } [a, c-\delta] \\ [0,1] & \textit{na } [c-\delta, c] \\ [0,\eta) & \textit{na } [c,b]. \end{cases}$$

Důkaz. Krok 1.: hrubá aproximace funkce $\chi_{[c-\frac{1}{2}\delta, b]}$.
Položme

$$Q(x) = \frac{1}{2} + \frac{x - (c - \frac{1}{2}\delta)}{2(b-a)}.$$

Pak Q je polynom splňující $Q \in [0, \frac{1}{2}]$ na $[a, c-\frac{1}{2}\delta]$ a $Q \in [\frac{1}{2}, 1]$ na $[c - \frac{1}{2}\delta, b]$.
Krok 2.: zpřesnění aproximace umocněním.
Pro zafixované $n \in \mathbb{N}$ položme

$$T(x) = (1 - Q^n(x))^{2^n}.$$

Ve zbytku důkazu ukážeme, že je-li n dostatečně velké, pak T má požadované vlastnosti. Předně zřejmě máme $T \in [0,1]$ na $[a,b]$. Dále Bernoulliova nerovnost na $[a, c-\delta]$ dává spolu s $\max_{[a,c-\delta]} Q = Q(c-\delta) < Q(c - \frac{1}{2}\delta) = \frac{1}{2}$

$$\min_{[a,c-\delta]} T \geq 1 - 2^n \max_{[a,c-\delta]} Q^n = 1 - (2 \max_{[a,c-\delta]} Q)^n \overset{n\to\infty}{\to} 1.$$

Konečně, díky $\min_{[c,b]} Q = Q(c) > \frac{1}{2}$, na $[c,b]$ platí

$$\begin{aligned}\max_{[c,b]} T &= (1-Q^n(c))^{2^n} = \frac{1}{(1+Q^n(c))^{2^n}}(1-Q^{2n}(c))^{2^n} \le \frac{1}{(1+Q^n(c))^{2^n}}\\ &\le \frac{1}{1+2^nQ^n(c)} \overset{n\to\infty}{\to} 0.\end{aligned}$$

Proto pro n dostatečně velké máme polynom s požadovanými vlastnostmi. □

Věta 11.5.2 (Weierstrassova aproximační věta)**.** *Nechť $[a,b] \subset \mathbb{R}$, $f \in C([a,b])$ a $\varepsilon > 0$. Pak existuje polynom P takový, že*

$$\max_{[a,b]} |f - P| < \varepsilon.$$

Důkaz. K zadanému $\varepsilon > 0$ a f definujme množinu

$$S = \Big\{ t \in [a,b] \colon \text{existuje polynom } Q \text{ splňující } \max_{[a,t]} |f - Q| < \varepsilon \Big\}.$$

Označme $s = \sup S$. Protože f je zprava spojitá v a, je $|f - f(a)| < \varepsilon$ na jistém pravém okolí bodu a. Odtud $S \neq \emptyset$ a $s > a$ (aproximovali jsme konstantním polynomem). Pokud by platilo $s \ge b$, byli bychom hotovi. V dalším se tedy zabývejme případem $s \in (a,b)$ a odvodíme spor.

Ze spojitosti f v bodě s dostáváme $\delta \in (0, \frac{1}{2}\min\{s-a, b-s\})$ tak, že

$$|f - f(s)| < \frac{1}{2}\varepsilon \qquad \text{na } [s-2\delta, s+2\delta].$$

Definujme $c := s - \delta$. Pak $c \in S$ a existuje polynom Q tak, že

$$m := \max_{[a,c]} |f - Q| < \varepsilon.$$

Definujme dále $M := \sup_{[a,b]} |f - Q| + \sup_{[a,b]} |f - f(s)|$. Konečně zafixujme $\eta > 0$ tak malé, že

$$m + M\eta < \varepsilon \qquad \text{a} \qquad M\eta < \frac{1}{2}\varepsilon.$$

Použijme nyní předchozí lemma k získání polynomu T pro parametry a, b, c, δ, η a definujme

$$P(x) := f(s) + (Q(x) - f(s))T(x).$$

Funkce P je polynom. Na intervalu $[a, c-\delta]$ platí

$$|f(x) - P(x)| \le |f(x) - Q(x)|T(x) + |f(x) - f(s)|(1 - T(x)) \le m + M\eta < \varepsilon.$$

Na intervalu $[c-\delta, c]$ platí

$$\begin{aligned}|f(x) - P(x)| &\le |f(x) - Q(x)|T(x) + |f(x) - f(s)|(1 - T(x))\\ &\le mT(x) + \frac{1}{2}\varepsilon(1 - T(x)) < \varepsilon.\end{aligned}$$

Konečně, na $[c, s+2\delta)$ dostáváme

$$|f(x) - P(x)| \leq |f(x) - Q(x)|T(x) + |f(x) - f(s)|(1 - T(x)) \leq M\eta + \frac{1}{2}\varepsilon < \varepsilon.$$

Celkově máme $|f - P| < \varepsilon$ na $[a, s+2\delta)$, z čehož plyne

$$\max_{[a,s+2\delta]} |f - P| < \varepsilon$$

($f - P$ je spojitá funkce na $[a, s+2\delta]$, a proto zde nabývá maxima) a to dává spor s $s = \sup S$. □

Důsledek 11.5.3. *Prostor $C([a,b])$ s maximovou metrikou je separabilní.*

Důkaz. Každou spojitou funkci umíme aproximovat polynomem a ten zase polynomem s racionálními koeficienty (promyslete si podrobnosti). Množina polynomů pevného stupně s racionálními koeficienty je spočetná, neboť konečný kartézský součin spočetných množin je spočetný (v kapitole o mohutnosti jsme si ukázali kontrukci odpovídající posloupnosti). Množinu všech polynomů s racionálními koeficienty získáme sjednocením předchozích množin přes všechny stupně. Je tedy spočetná, neboť sjednocení spočetného systému spočetných množin je spočetné. □

Příklad 11.5.4. Dokažme, že pro $f \in C([0, 2\pi])$ platí

$$\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x \overset{n\to\infty}{\to} 0.$$

Pro funkce z $C^1([0, 2\pi])$ výsledek platí, neboť můžeme použít integraci per partes a ta dává

$$\begin{aligned} \Big|\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x\Big| &= \Big|\Big[f(x)\frac{\sin(nx)}{n}\Big]_0^{2\pi} - \int_0^{2\pi} f'(x)\frac{\sin(nx)}{n}\,\mathrm{d}x\Big| \\ &= \Big|\int_0^{2\pi} f'(x)\frac{\sin(nx)}{n}\,\mathrm{d}x\Big| \leq \int_0^{2\pi} \max_{[0,2\pi]} |f'| \frac{1}{n}\,\mathrm{d}x \overset{n\to\infty}{\to} 0. \end{aligned}$$

V obecném případě k $f \in C([a,b])$ a $\varepsilon > 0$ vezmeme polynom P z Weierstrassovy aproximační věty, pro který platí

$$\max_{[0,2\pi]} |f - P| < \varepsilon.$$

Díky této vlastnosti a tomu, že $P \in C^1([0, 2\pi])$, pro $n \in \mathbb{N}$ dostatečně velké dostáváme rozhodující odhad

$$\begin{aligned} \Big|\int_0^{2\pi} f(x)\cos(nx)\,\mathrm{d}x\Big| &\leq \Big|\int_0^{2\pi} (f(x) - P(x))\cos(nx)\,\mathrm{d}x\Big| + \Big|\int_0^{2\pi} P(x)\cos(nx)\,\mathrm{d}x\Big| \\ &\leq 2\pi\varepsilon + \varepsilon. \end{aligned}$$

Cvičení 11.5.5. Nechť $f \in C([a,b])$ a

$$\int_a^b f(x)x^k \, \mathrm{d}x = 0 \qquad \text{pro každé } k \in \mathbb{N}_0.$$

Ukažte, že pak $f \equiv 0$ na $[a,b]$. (Návod: nejprve ukažte, že $\int_a^b f(x)P(x)\,\mathrm{d}x = 0$ pro každý polynom P, pak použijte Weierstrassovu aproximační větu podobně jako v minulém příkladu k důkazu $\int_a^b f^2(x)\,\mathrm{d}x = 0$. Nakonec ukažte, že je-li f nenulová v nějakém bodě, má díky spojitosti srovnatelnou hodnotu na nějakém jeho okolí, a proto integrál z její druhé mocniny nemůže být nulový.)

Poznámka 11.5.6. (i) Zde bychom čtenáře rádi upozornili na závažnost právě představeného postupu. Doposud jsme si práci na důkazech usnadňovali jen používáním různých symetrií (kupříkladu všechny výsledky pro konkávnost se dají získat drobnou modifikací důkazů analogických tvrzení pro konvexitu či jen přechodem k funkci $-f$) a vyšetřováním nejhorších případů (při důkazu aritmetiky nevlastních limit je nejtěžším případem dvojice vlastních limit, protože s výrazy typu „$\infty - \infty$“ věta odmítá pracovat, v případě „$\infty + \infty$“ je důkaz velice jednoduchý, atd.). Naše nová technika funguje přesně obráceně. Umožňuje nám pracovat ve velice příznivém případě, třebaže tyto případy jsou oproti všem ostatním velice vzácné (funkce z $C^1([a,b])$ jsou mezi funkcemi z $C([a,b])$ skutečně velmi vzácné, vlastnost mít derivaci znamená mít obě jednostranné derivace a tyto derivace se musejí rovnat, což u dvojice reálných čísel nastává velice zřídka). Cenou, kterou za tento luxus musíme zaplatit je ověření, že dokazovaná vlastnost se zachovává při konvergenci ve studovaném metrickém prostoru a že používaná množina „hezkých“ prvků je skutečně hustá (ale to u běžně používaných metrických prostorů bývají dobře známé výsledky, které se dokazují mezi prvními, když se takový prostor zavede).
(ii) Na obvykle používaných metrických prostorech tvořených funkcemi nad otevřenou množinou $\Omega \subset \mathbb{R}^N$ se k výše popsaným účelům nejčastěji používá množina $C_0^\infty(\Omega)$, což je množina nekonečněkrát diferencovatelných funkcí s kompaktním nosičem, přičemž nosič funkce je definován jako uzávěr podmnožiny definičního oboru, kde je funkce nenulová. Do $C_0^\infty(\mathbb{R}^N)$ patří třeba funkce

$$f(x) = \begin{cases} \exp\left(\frac{1}{\|x\|_2^2 - 1}\right) & \text{pro } \|x\|_2 < 1 \\ 0 & \text{jinak} \end{cases}$$

(zde symbol $\|x\|_2$ značí eukleidovskou normu bodu $x \in \mathbb{R}^N$).

11.6 Úplné metrické prostory

Nyní se budeme zabývat jemnějším přístupem ke konvergenci posloupnosti. Vzpomeňme si na příklad posloupnosti $\{\frac{1}{n}\}$, která konverguje na $\mathbb{R}$ v obvyklé metrice, ale na prostoru $P = (0,1)$ s toutéž metrikou konvergentní není. Konvergence posloupnosti je tedy závislá na dvou jevech. Jednak se musí členy posloupnosti někde

„hromadit“ (připomeňme, že tuto vlastnost nám na $\mathbb{R}$ charakterizuje B-C podmínka), jednak musí prostor P obsahovat limitní bod. To nás vede k následujícím definicím.

Definice 11.6.1 (Cauchyovská posloupnost)**.** Nechť (P, ϱ) je metrický prostor. Řekneme, že posloupnost $\{x_n\} \subset P$ je *cauchyovská* v (P, ϱ), jestliže

$$\forall \varepsilon > 0\, \exists n_0 \in \mathbb{N} \quad m, n > n_0 \implies \varrho(x_m, x_n) < \varepsilon.$$

Definice 11.6.2 (Úplný metrický prostor)**.** Řekneme, že metrický prostor(P, ϱ) je *úplný*, jestliže každá cauchyovská posloupnost jeho prvků v něm konverguje.

Příklad 11.6.3. (i) Prostor $\mathbb{R}$ s obvyklou metrikou je úplný, podobně $[0, 1]$. Naopak $(0, 1)$, $(0, 1]$ či $\mathbb{Q}$ úplné nejsou.
(ii) Prostor $\mathbb{R}^N$ opatřený kteroukoliv normou, které jsme si na něm představili, je úplný. Skutečně, cauchyovskost posloupnosti v kterékoliv z uvažovaných norem implikuje cauchyovskost jednotlivých složek (vůči metrice generované absolutní hodnotou). Jednotlivé složky tedy konvergují v $\mathbb{R}$ a Věta o vztahu konvergence v normě ke konvergenci po složkách (Věta 11.2.10) dává konvergenci v normě.
(iii) Podle Věty o vztahu konvergence v normě ke konvergenci po složkách je každý konečnědimenzionální normovaný lineární prostor úplný.
(iv) Uvažme prostor $C([-1, 1])$ s metrikou $\varrho(f, g) = \int_{-1}^{1} |f - g| \, \mathrm{d}x$ a posloupnost $\{f_n\} \subset C([-1, 1])$, kde

$$f_n(x) = \begin{cases} 0 & \text{pro } x \in [-1, 0] \\ nx & \text{pro } x \in [0, \frac{1}{n}] \\ 1 & \text{pro } x \in [\frac{1}{n}, 1]. \end{cases}$$

Pro $m > n$ platí

$$\begin{aligned} \varrho(f_n, f_m) &= \int_{-1}^{1} |f_n - f_m| \, \mathrm{d}x = \int_{0}^{\frac{1}{n}} |f_n - f_m| \, \mathrm{d}x \\ &\leq \int_{0}^{\frac{1}{n}} \Big(\max_{[0, \frac{1}{n}]} |f_n| + \max_{[0, \frac{1}{n}]} |f_m| \Big) \, \mathrm{d}x = \int_{0}^{\frac{1}{n}} (1 + 1) \, \mathrm{d}x = \frac{2}{n}. \end{aligned}$$

Proto je posloupnost $\{f_n\}$ cauchyovská. Pokud by $\{f_n\}$ měla v našem prostoru limitu f, muselo by pro ni platit $f = 0$ na $(-1, 0)$. Skutečně, pokud by například existovalo $x_0 \in (-1, 0)$ tak, že $f(x_0) > 0$, ze spojitosti by platilo $f > \frac{1}{2} f(x_0)$ na $\mathcal{U}_\delta(x_0)$ pro jisté $\delta \in (0, \min\{|x_0|, |x_0 + 1|\})$ a odtud

$$\begin{aligned} \varrho(f_n, f) &= \int_{-1}^{1} |f_n - f| \, \mathrm{d}x \geq \int_{x_0 - \delta}^{x_0 + \delta} |f_n - f| \, \mathrm{d}x = \int_{x_0 - \delta}^{x_0 + \delta} |f| \, \mathrm{d}x \\ &\geq \int_{x_0 - \delta}^{x_0 + \delta} \frac{1}{2} f(x_0) \, \mathrm{d}x = \delta f(x_0). \end{aligned}$$

Proto by nemohlo platit $f_n \to f$. Analogicky se dostane, že $f \equiv 1$ na $(0, 1)$. Celkově limitní funkce není spojitá v počátku, což je spor. Proto $\{f_n\}$ není konvergentní a

náš metrický prostor není úplný.
(v) Lehkou modifikací předchozích výpočtů se dá ukázat, že prostor $C([-1,1])$ není úplný s žádnou z integrálních norem, které jsme si na něm představili.
(vi) Ukažme, že prostor $C([a,b])$ opatřený maximovou normou úplný je. Nechť $\{f_n\} \subset C([a,b])$ je cauchyovská posloupnost. Díky definici maximové normy okamžitě dostáváme, že pro každé zafixované $x \in [a,b]$ je posloupnost $\{f_n(x)\} \subset \mathbb{R}$ cauchyovská v $\mathbb{R}$ (s obvyklou metrikou, tedy v úplném prostoru). Označme limitní prvek $f(x)$.

V dalším ukážeme, že zobrazení $x \mapsto f(x)$ je spojité a $\|f_n - f\|_\infty \to 0$. Zafixujme $\varepsilon > 0$. Díky cauchyovskosti $\{f_n\}$ pak existuje $n_0 \in \mathbb{N}$ takové, že

$$\max_{[a,b]} |f_n - f_{n+p}| < \varepsilon \qquad \text{pro všechna } n > n_0 \text{ a } p \in \mathbb{N}.$$

Pro každé $x \in [a,b]$ pevné a $n > n_0$ pevné nám limitní přechod $p \to \infty$ dává (vzpomeňte si na větu o zachování neostré nerovnosti při limitním přechodu)

$$|f_n(x) - f(x)| \le \varepsilon.$$

To má pro nás dva důsledky. Jednak okamžitě dostáváme

$$\|f_n - f\| \le \varepsilon \qquad \text{pro všechna } n > n_0,$$

čímž jsme dokázali $\|f_n - f\|_\infty \to 0$. Dále pro zafixovaná $n > n_0$ a $x_0 \in [a,b]$ je f_n spojité v x_0 a existuje tedy $\delta > 0$ tak, že

$$x \in [a,b] \cap \mathcal{U}_\delta(x_0) \qquad \Longrightarrow \qquad |f_n(x) - f_n(x_0)| < \varepsilon.$$

Celkově pro $x \in [a,b] \cap \mathcal{U}_\delta(x_0)$ dostáváme

$$|f(x) - f(x_0)| \le |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| < 3\varepsilon,$$

čímž jsme dokázali, že funkce f je spojitá v x_0. Úplnost $C([a,b])$ s maximovou normou je dokázána.

Věta 11.6.4 (O vztahu konvergence a cauchyovskosti)**.** *Nechť (P,ϱ) je metrický prostor. Je-li posloupnost $\{x_n\} \subset P$ konvergentní, pak je cauchyovská.*

Důkaz. Důkaz plyne okamžitě z trojúhelníkové nerovnosti. □

Definice 11.6.5 (Banachův a Hilbertův prostor)**.** Úplnému normovanému prostoru říkáme *Banachův prostor*. Úplnému prostoru se skalárním součinem říkáme *Hilbertův prostor*. (Úplnost samozřejmě bereme vůči metrice vzniklé z uvedené normy respektive skalárního součinu.)

Poznámka 11.6.6. Zřejmě každý Hilbertův prostor je i Banachův.

Příklad 11.6.7. (i) Prostor $C([a,b])$ opatřený maximovou normou je Banachův prostor.
(ii) Prostor $\mathbb{R}$ s klasickou metrikou je Hilbertův (úplnost nám dává Věta o B-C

podmínce, skalární součin zavádíme jako $(x, y) := xy$).
(iii) Prostor $\mathbb{R}^N$ s eukleidovskou metrikou je Hilbertův. Skalární součin je zde

$$(x, y) = \sum_{i=1}^{N} x_i y_i,$$

ten generuje právě eukleidovskou normu a metriku. Úplnost už jsme měli (plyne z konvergence po složkách).
(iv) Připomeňme si prostor

$$\ell_2 = \Big\{\{x_i\} \in \mathbb{R} \colon \sum_{i=1}^{\infty} |x_i|^2 < \infty \, \forall i \in \mathbb{N}\Big\}, \qquad \varrho(\{x_i\}, \{y_i\}) = \Big(\sum_{i=1}^{\infty} |x_i - y_i|^2\Big)^{\frac{1}{2}}.$$

Uvedená metrika vznikla ze skalárního součinu (sami si ověřte splnění vlastností skalárního součinu z definice)

$$(\{x_i\}, \{y_i\}) = \sum_{i=1}^{\infty} x_i y_i.$$

Dokažme si ještě úplnost. Nechť $\{x^j\} = \{x_1^j, x_2^j, \dots\} \subset \ell_2$ je cauchyovská posloupnost. To znamená, že pro každé $\varepsilon > 0$ existuje $n_0 \in \mathbb{N}$ tak, že

$$\sum_{i=1}^{\infty} |x_i^m - x_i^n|^2 < \varepsilon^2 \qquad \text{kdykoliv } m, n > n_0. \tag{11.6.1}$$

Proto je každá složka cauchyovská v $\mathbb{R}$, tedy konvergentní a dostáváme $x = \{x_1, x_2, \dots\}$ takové, že $\{x^j\}$ konverguje k x po složkách.

Ukažme, že $\varrho(x_n, x) \to 0$. Zvolme $\varepsilon > 0$. Pak existuje $n_0 \in \mathbb{N}$ tak, že platí (11.6.1) a odtud pro zafixované $N \in \mathbb{N}$ máme

$$\sum_{i=1}^{N} |x_i^m - x_i^n|^2 < \varepsilon^2 \qquad \text{kdykoliv } m, n > n_0.$$

Při zafixovaném $n > n_0$ proveďme limitní přechod $m \to \infty$ (využíváme konvergenci po složkách, kterých je již jen konečný počet)

$$\sum_{i=1}^{N} |x_i^n - x_i|^2 \le \varepsilon^2.$$

Protože N bylo libovolné a součet nekonečné řady je limitou jejích částečných součtů, dostáváme

$$\sum_{i=1}^{\infty} |x_i^n - x_i|^2 \le \varepsilon^2,$$

což je $\varrho(x_n, x) \le \varepsilon$. Ukázali jsme tedy, že $\varrho(x_n, x) \to 0$.

Zbývá dokázat, že $x \in \ell_2$. Vezměme $n \in \mathbb{N}$ tak, aby $\varrho(x_n, x) \leq 1$. Pak pro libovolné $N \in \mathbb{N}$ platí díky trojúhelníkové nerovnosti

$$\begin{aligned}\Big(\sum_{i=1}^{N} |x_i|^2\Big)^{\frac{1}{2}} &\leq \Big(\sum_{i=1}^{N} |x_i^n - x_i|^2\Big)^{\frac{1}{2}} + \Big(\sum_{i=1}^{N} |x_i^n|^2\Big)^{\frac{1}{2}} \\ &\leq 1 + \Big(\sum_{i=1}^{\infty} |x_i^n|^2\Big)^{\frac{1}{2}} = 1 + \|x^n\|_2.\end{aligned}$$

Protože tento odhad platí pro každé $N \in \mathbb{N}$, máme

$$\|x\|_2 = \Big(\sum_{i=1}^{\infty} |x_i|^2\Big)^{\frac{1}{2}} \leq 1 + \|x^n\|_2 \qquad \Longrightarrow \qquad x \in \ell_2.$$

Celkově ℓ_2 je Hilbertův prostor.

Věta 11.6.8 (Charakterizace úplných podprostorů úplného prostoru)**.** *Nechť je (P, ϱ) úplný metrický prostor a $A \subset P$. Pak*

$$(A, \varrho) \text{ je úplný metrický prostor} \qquad \Longleftrightarrow \qquad A \text{ je uzavřená (v } P).$$

Důkaz. Věta snadno plyne z Věty o charakterizaci uzavřenosti pomocí posloupností (Věta 11.3.26). □

Příklad 11.6.9. (i) V $\mathbb{R}$ s obvyklou metrikou je $[0, \infty)$ uzavřená množina, a proto je jí odpovídající prostor úplný.
(ii) V $(0, \infty)$ s obvyklou metrikou je $(0, 1]$ uzavřená množina, ale jí odpovídající prostor není úplný. To není ve sporu s větou, neboť původní prostor $(0, \infty)$ není úplný.
(iii) Protože je každý konečnědimenzionální normovaný lineární prostor úplný, je uzavřenou podmnožinou v každém větším úplném prostoru se stejnou normou.

11.7 Omezenost a kompaktnost

Jedním z hlavních výsledků diferenciálního počtu v jedné dimenzi byla existence globálních extrémů na omezeném uzavřeném intervalu. Podobný výsledek budeme chtít získat i pro obecnější zobrazení z metrického prostoru do $\mathbb{R}$. Jako první krok budeme hledat správný typ množin, které by nahradily omezený uzavřený interval, na němž se dá aplikovat Weierstressova věta. To nás vede k následující definici.

Definice 11.7.1 (Kompaktní množina)**.** Nechť (P, ϱ) je metrický prostor a $A \subset P$. Řekneme, že A je *kompaktní*, jestliže z každé posloupnosti jejích prvků lze vybrat podposloupnost konvergentní v A.

Poznámka 11.7.2. V $\mathbb{R}$ má tuto vlastnost omezený uzavřený interval, neboť díky omezenosti lze aplikovat Weierstrassovu větu a uzavřenost zaručí, že i limitní hodnota leží v našem intervalu. Zřejmě má stejnou vlastnost jakákoliv omezená uzavřená množina v $\mathbb{R}$.

Cvičení 11.7.3. Dokažte si, že funkce, která je spojitá na kompaktní podmnožině $\mathbb{R}$, zde nabývá svých globálních extrémů.

Není těžké ověřit, že množina v $\mathbb{R}$ je kompaktní právě tehdy, když je omezená a uzavřená. Tato informace je pro nás výhodná, neboť kompaktnost je přímo z definice neověřitelná (museli bychom otestovat všechny možné posloupnosti), naproti tomu máme nástroje na ověření omezenosti a uzavřenosti (uzavřená množina je doplněk otevřené, průnik uzavřených je uzavřený, atd.). Nás bude v dalším zajímat, zda lze zjistit kompaktnost pomocí jiných (snáze ověřitelných) vlastností i v obecnějších metrických prostorech.

Definice 11.7.4 (Omezená množina)**.** Nechť (P, ϱ) je metrický prostor a $A \subset P$. Řekneme, že A je *omezená*, jestliže existují $x_0 \in P$ a $R > 0$ tak, že $A \subset \mathcal{U}_R(x_0)$.

Poznámka 11.7.5. Díky trojúhelníkové nerovnosti lze v normovaném lineárním prostoru omezenost také charakterizovat existencí $K > 0$ takového, že $\|x\| \leq K$ pro každé $x \in A$.

Poznámka 11.7.6. (i) Budeme také používat pojem *omezená posloupnost*. Jako obvykle tím myslíme, že je omezená množina jejích členů.
(ii) Každá cauchyovská posloupnost je omezená (zvolte $\varepsilon = 1$ v definici cauchyovskosti, pak už je důkaz snadný).

Věta 11.7.7 (Kompaktnost implikuje omezenost a uzavřenost)**.** *Nechť (P, ϱ) je metrický prostor a $A \subset P$ je kompaktní. Pak A je omezená a uzavřená.*

Důkaz. Nejprve pro spor předpokládejme, že A není omezená. Zvolme $x_0 \in A$ pak díky neomezenosti musí existovat $\{x_n\} \subset A$ tak, že $\varrho(x_n, x_0) > n$ pro každé $n \in \mathbb{N}$. Díky kompaktnosti A můžeme po přechodu k podposloupnosti předpokládat, že $x_n \to y$. Díky spojitosti metriky celkově máme

$$\infty \leftarrow \varrho(x_n, x_0) \to \varrho(y, x_0) < \infty$$

a máme spor.

Nyní dokažme uzavřenost pomocí Věty o charakterizaci uzavřenosti pomocí posloupností (Věta 11.3.25). Nechť $\{x_n\} \subset A$ a $x_n \to x \in P$. Díky kompaktnosti existuje podposloupnost $\{x_{n_k}\} \subset \{x_n\}$ taková, že $x_{n_k} \to y \in A$. Nutně potom $x = y$, a proto $x \in A$. $\square$

Obecně se implikace v předchozí větě nedá otočit.

Příklad 11.7.8. V prostoru ℓ_2 vezměme množinu

$$A = \{x_1, x_2, \dots\} := \big\{\{1, 0, 0, \dots\}, \{0, 1, 0, \dots\}, \{0, 0, 1, 0, \dots\}, \dots\big\}.$$

Pak $\|x_n\|_2 = 1$ pro všechna $n \in \mathbb{N}$, tedy A je omezená. Dále pro $m \neq n$ platí $\|x_m - x_n\|_2 = \sqrt{2}$. To má dva důsledky. Jednak posloupnosti prvků A jsou konvergentní právě tehdy, když jsou od určitého členu konstantní. Konvergentní posloupnosti prvků A mají tedy vždy limitu v A a A je uzavřená. Druhým důsledkem je, že uvážíme-li přímo posloupnost $\{x_n\}$, žádná její podposloupnost není cauchyovská, natož aby konvergovala. Množina A proto není kompaktní.

V konečné dimenzi se ovšem implikace otočit dá.

Věta 11.7.9 (Omezenost a uzavřenost implikují kompaktnost v konečné dimenzi)**.** *V konečnědimenzionálním normovaném lineárním prostoru je každá omezená a uzavřená množina kompaktní.*

Důkaz. Nechť A je omezená uzavřená množina a $\{e_1, e_2, \dots, e_N\}$ je báze našeho prostoru. Zvolme posloupnost $\{x_n\} = \{\sum_{j=1}^N \alpha_j^n e_j\} \subset A$. Pak jsou omezené i posloupnosti koeficientů jednotlivých složek $\{\alpha_1^n\}, \{\alpha_2^n\}, \dots, \{\alpha_N^n\}$ (to není zcela zřejmé, neboť nepracujeme s ortogonální bází. Stačí však navázat na postup z důkazu Věty o vztahu konvergence v normě ke konvergenci po složkách, tedy Věty 11.2.10, kde jsme dokázali, že konvergence k nule v normě implikuje konvergenci koeficientů jednotlivých složek k nule. Pokud by totiž po přechodu k podposloupnosti šly třeba koeficienty α_1^n do ∞, pak by prvky $\frac{x_n}{\alpha_1^n}$ šly v normě k nule, ale koeficient jejich první složky by byl konstantně roven jedné.)

Posloupnosti koeficientů jsou tedy omezené. Postupným přecházením k podposloupnosti (aplikací Weierstrassovy věty) dosáhneme toho, že tyto koeficienty konvergují k číslům $\alpha_1, \alpha_2, \dots, \alpha_N$. Proto odpovídající vybraná posloupnost $\{x_{n_k}\}$ konverguje po složkách k $x := \sum_{j=1}^N \alpha_j e_j$. Podle Věty o vztahu konvergence v normě ke konvergenci po složkách (Věta 11.2.10) pak máme $x_{n_k} \to x$, a protože A je uzavřená, platí $x \in A$. Tím je důkaz dokončen. $\square$

Věta 11.7.10 (Kompaktnost implikuje separabilitu)**.** *Každá kompaktní podmnožina metrického prostoru je separabilní.*

Důkaz. Nechť A je kompaktní podmnožina (P, ϱ). Nejprve ukažme, že pro každé $\varepsilon > 0$ existuje konečná množina $K_\varepsilon \subset A$ taková, že pro každé $x \in A$ existuje $y \in K_\varepsilon$ tak, že $\varrho(x, y) < \varepsilon$. Pokud by tomu tak nebylo, našli bychom posloupnost $\{x_n\} \subset K$ splňující

$$\varrho(x_m, x_n) \geq \varepsilon \qquad \text{kdykoliv } m \neq n.$$

Totiž, k libovolně (ale pevně) zvolenému prvku $x_1 \in A$ dle předpokladu existuje $x_2 \in A$ tak, že $\varrho(x_1, x_2) \geq \varepsilon$. Nyní opět k množině $\{x_1, x_2\}$ existuje prvek $x_3 \in A$ tak, že $\varrho(x_3, x_1) \geq \varepsilon$ a $\varrho(x_3, x_2) \geq \varepsilon$. Dále postupujeme indukcí.

Z takové posloupnosti bychom však nemohli vybrat konvergentní podposloupnost, což by byl spor s kompaktností A.

Spočetnou hustou podmnožinu K nyní získáme konstrukcí $K_1 \cup K_{\frac{1}{2}} \cup K_{\frac{1}{3}} \cup \dots$. $\square$

Věta 11.7.11 (Cantorova věta o průniku kompaktů)**.** *Nechť $K_1 \supset K_2 \supset \dots$ je posloupnost neprázdných kompaktů v metrickém prostoru. Pak $\bigcap_{n\in\mathbb{N}} K_n$ je neprázdný kompakt.*

Důkaz. Pokud $\{x_n\} \subset \bigcap_{n\in\mathbb{N}} K_n$, zároveň platí $\{x_n\} \subset K_1$ a po přechodu k podposloupnosti díky kompaktnosti K_1 máme $x_{n_k} \to x$ v K_1. Nutně pak také $x \in \bigcap_{n\in\mathbb{N}} K_n$, neboť $\bigcap_{n\in\mathbb{N}} K_n$ je uzavřená množina jakožto průnik uzavřených množin. Celkově jsme ověřili kompaktnost $\bigcap_{n\in\mathbb{N}} K_n$.

Ukažme ještě neprázdnost. Sestrojme „diagonální“ posloupnost $\{x_n\}$ tak, že $x_n \subset K_n$ pro každé $n \in \mathbb{N}$. Po přechodu k podposloupnosti pak máme $x_{n_k} \to x \in K_1$. Protože od druhého indexu všechny členy leží v K_2, máme zároveň konvergenci na K_2, tedy $x \in K_2$. Takto můžeme pokračovat dále a dostáváme $x \in K_n$ pro každé $n \in \mathbb{N}$. Proto $x \in \bigcap_{n\in\mathbb{N}} K_n$ a množina $\bigcap_{n\in\mathbb{N}} K_n$ je neprázdná. $\square$

Cvičení 11.7.12. Dokažte, že každá uzavřená podmnožina kompaktní množiny je kompaktní (použijte Větu o charakterizaci uzavřenosti pomocí posloupností, tedy Větu 11.3.26).

11.8 Pokrývací věty

Budou nás zajímat ještě další vlastnosti kompaktních množin. K jejich odvození použijeme nové nástroje, jimž se říká *pokrývací věty*, a které mají široké uplatnění i v jiných částech matematické analýzy.

Definice 11.8.1 (Pokrytí množiny)**.** Nechť (P, ϱ) je metrický prostor, $A \subset P$, I je indexová množina a $\{M_\alpha\}_{\alpha\in I}$ je systém podmnožin P. Řekneme, že $\{M_\alpha\}$ je *pokrytí* A, jestliže $A \subset \bigcup_{\alpha\in I} M_\alpha$. Jsou-li všechny množiny v systému $\{M_\alpha\}$ otevřené, hovoříme o *otevřeném pokrytí.*

Věta 11.8.2 (Lindelöfova pokrývací věta)**.** *Nechť (P, ϱ) je metrický prostor a $A \subset P$ je separabilní. Pak lze z každého otevřeného pokrytí množiny A vybrat nejvýše spočetné podpokrytí (podsystém stále pokrývající A).*

Důkaz. Nechť $\{M_\alpha\}_{\alpha\in I}$ je pokrytí A. Nechť dále $\{x_n\}$ je spočetná hustá podmnožina A.

Nejprve tvrdíme, že pro každé $x \in A$ existují $\alpha \in I$ a $m, n \in \mathbb{N}$ tak, že $x \in \Omega_x := \mathcal{U}_{\frac{1}{m}}(x_n) \subset M_\alpha$. Skutečně, díky tomu, že $\{M_\alpha\}$ je otevřené pokrytí, existují α a $R > 0$ tak, že $\mathcal{U}_R(x) \subset M_\alpha$. Zafixujme dále $m > \frac{2}{R}$ (tj. $\frac{1}{m} < \frac{R}{2}$). Díky hustotě $\{x_n\}$ existuje navíc $n \in \mathbb{N}$ tak, že $\varrho(x, x_n) < \frac{1}{m} < \frac{R}{2}$, a proto z trojúhelníkové nerovnosti dostáváme

$$x \in \mathcal{U}_{\frac{1}{m}}(x_n) \subset \mathcal{U}_R(x) \subset M_\alpha.$$

Konečně, systém $\{\Omega_x\}_{x\in A}$ pokrývá A a zároveň je nejvýše spočetný (je to podsystém $\{\mathcal{U}_{\frac{1}{m}}(x_n)\}_{m,n\in\mathbb{N}}$). Ke každé množině z tohoto systému přiřadíme právě jednu množinu ze systému $\{M_\alpha\}$, aby ji obsahovala. Tímto přiřazováním jsme z $\{M_\alpha\}$ získali podsystém požadovaných vlastností. $\square$

Věta 11.8.3 (Borelova pokrývací věta)**.** *Nechť (P, ϱ) je metrický prostor a $A \subset P$ je kompaktní. Pak lze z každého otevřeného pokrytí množiny A vybrat konečné podpokrytí.*

Důkaz. Protože kompaktnost implikuje separabilitu (Věta 11.3.22), pomocí Lindelöfovy pokrývací věty (Věta 11.8.2) můžeme od obecného otevřeného pokrytí přejít ke spočetnému pokrytí, které v dalším značme $\{M_n\}$. Definujme $G_n =$

$\bigcup_{i=1}^{n} M_i$, $n \in \mathbb{N}$. Stačí zřejmě ukázat, že existuje $n_0 \in \mathbb{N}$ tak, že $A \subset G_{n_0}$. Pro spor předpokládejme, že to není pravda. Pak jsou $F_n := A \setminus G_n$, $n \in \mathbb{N}$, neprázdné kompakty splňující $F_n \subset F_m$, kdykoliv $m < n$. Cantorova věta o průniku kompaktů dává, že

$$F := \bigcap_{n\in\mathbb{N}} F_n = \bigcap_{n\in\mathbb{N}} A \setminus G_n = A \setminus \bigcup_{n\in\mathbb{N}} G_n$$

je neprázdná kompaktní množina. Proto $\{M_n\}$ nepokrývá A a máme spor. $\square$

Příklad 11.8.4. Dokažme, že je-li funkce f lokálně lipschitzovská na $[a, b]$, pak je zde dokonce lipschitzovská. Podle předpokladu pro každé $x \in [a, b]$ existují $\delta_x > 0$ a $L_x > 0$ taková, že

$$|f(y) - f(z)| < L_x|y - z| \qquad \text{pro každé } y, z \in [a, b] \cap (x - \delta_x, x + \delta_x).$$

Systém $\{(x - \delta_x, x + \delta_x)\}_{x\in[a,b]}$ tvoří otevřené pokrytí intervalu $[a, b]$. Podle Borelovy pokrývací věty existují body $x_1, \dots, x_m \in [a, b]$ takové, že $[a, b] \subset \bigcup_{i=1}^{m}(x_i - \delta_{x_i}, x_i + \delta_{x_i})$. Položíme-li nyní

$$L := \max\{L_{x_1}, \dots, L_{x_m}\}$$

je již snadné nahlédnout, že L je konstanta lipschitzovskosti funkce f pro celé $[a, b]$.

Poznámka 11.8.5. (i) Funkce $f(x) = \frac{1}{x}$ na $(0, 1]$ a $g(x) = x^2$ na $\mathbb{R}$ ukazují, že v předchozím příkladě byla podstatná jak omezenost tak uzavřenost intervalu $[a, b]$.
(ii) Předchozí výsledek se nedá získat postupným plížením, tedy konstrukcí, kdy nejprve vezmene δ_a a L_a, pak vezmeme bod $x_1 \in (a, a + \delta_a)$ a jemu odpovídající konstanty δ_{x_1} a L_{x_1} (teď už víme, že funkce f je na $[a, x_1 + \delta_{x_1})$ lipschitzovská s konstantou $\max\{L_a, L_{x_1}\}$), pak vezmeme $x_2 \in (x_1, x_1 + \delta_{x_1})$ a jemu odpovídající konstanty δ_{x_2} a L_{x_2}, atd. Tento proces nemusí fungovat, protože čísla δ_{x_i} se mohou rychle zkracovat a nám se pak nepodaří dosáhnout bodu b v konečném počtu kroků.

Pokrývací věty patří k velmi často používaným nástrojům pokročilé matematické analýzy. Jejich účinnost si můžete vyzkoušet na následujících důkazech souvisejících s látkou předchozího semestru, kde jsme leckdy museli postupovat trikově, zatímco pokrývací věty nám nyní umožňují přímočařejší přístup.

Cvičení 11.8.6. (i) Pomocí Borelovy pokrývací věty dokažte, že je-li f omezená na okolí každého $x \in [a, b]$, pak je omezená na $[a, b]$ (speciálně, je-li f spojitá na $[a, b]$, pak je zde omezená).
(ii) Pomocí Borelovy pokrývací věty dokažte, že je-li $f' > 0$ (i nevlastní) na (a, b), pak f je rostoucí na (a, b).
(iii) Pomocí Borelovy pokrývací věty dokažte, že je-li f rostoucí ve všech bodech intervalu (a, b), pak je rostoucí na (a, b).
(iv) Pomocí Borelovy pokrývací věty dokažte Cantorovu větu o stejnoměrné spojitosti.

Dalším pojmem, který chceme zobecnit do vyšší dimenze, je interval.

Definice 11.8.7 (Interval)**.** Intervalem v $\mathbb{R}^N$ nazveme množinu tvaru $I = I_1 \times I_2 \times \cdots \times I_N$, kde $I_1, \dots, I_N$ jsou intervaly v $\mathbb{R}$. Je-li množina I otevřená, hovoříme o otevřeném intervalu. Podobně se definují uzavřené, omezené a kompaktní intervaly.

Snadno se dokáže následující charakterizace právě zavedených pojmů.

Tvrzení 11.8.8. *Interval $I = I_1 \times I_2 \times \cdots \times I_N \subset \mathbb{R}^N$ otevřený právě tehdy, když je každý z intervalů $I_1, \dots, I_N$ otevřený. Analogická tvrzení platí pro uzavřenost, omezenost a kompaktnost.*

Věta 11.8.9 (Charakterizace otevřené množiny pomocí otevřených intervalů)**.** *Nechť A je otevřená množina v $\mathbb{R}^N$. Pak existuje spočetný systém otevřených intervalů $\{I_n\}$ takový, že $A = \bigcup I_n$.*

Důkaz. Nechť $A \subset \mathbb{R}^N$. Pracujme s maximovou normou (okolí v této normě jsou otevřené krychle, tedy otevřené intervaly). Podle definice otevřené množiny pro každý bod $x \in A$ najdeme jeho okolí, které leží celé v A. Systém všech těchto okolí je otevřeným pokrytím A. Pomocí Lindelöfovy pokrývací věty z něj vybereme spočetný systém a jsme hotovi. □

11.9 Banachova věta o kontrakci

V dalším si dokážeme větu, která umožňuje řešit některé poměrně obtížné problémy. Nejprve si zadefinujme základní pojmy.

Definice 11.9.1 (Kontraktivní zobrazení)**.** Nechť (P, ϱ) je metrický prostor a $T\colon P \to P$ je zobrazení definované na celém P. Řekneme, že T je *kontraktivní* zobrazení (nebo stručně *kontrakce*), jestliže existuje $q \in [0, 1)$ tak, že

$$\varrho(Tx, Ty) \leq q\varrho(x, y) \qquad \text{pro všechna } x, y \in P.$$

Bod x_0 se nazývá *pevný bod* zobrazení T, jestliže $Tx_0 = x_0$.

Věta 11.9.2 (Banachova věta o kontrakci)**.** *Nechť (P, ϱ) je úplný metrický prostor a $T\colon P \to P$ je kontraktivní zobrazení definované na celém P. Pak má právě jeden pevný bod.*

Dokonce pro každou posloupnost $\{x_n\} \subset P$ splňující $x_{n+1} = Tx_n$ pro všechna $n \in \mathbb{N}$ ($x_1 \in P$ je libovolné) platí $x_n \to x_0$, kde x_0 je zmíněný pevný bod zobrazení T.

Důkaz. Jednoznačnost plyne z toho, že pro dvojici pevných bodů x_1, x_2 máme

$$\varrho(x_1, x_2) = \varrho(Tx_1, Tx_2) \leq q\varrho(x_1, x_2).$$

Existence plyne z druhé části věty, kterou dokážeme nyní. Nechť $\{x_n\}$ je jako ve znění věty. Díky kontaktivitě T máme pro každé $n \in \mathbb{N}$, $n \geq 3$

$$\begin{aligned}\varrho(x_{n+1}, x_n) &= \varrho(Tx_n, Tx_{n-1}) \leq q\varrho(x_n, x_{n-1}) = q\varrho(Tx_{n-1}, Tx_{n-2}) \leq \dots \\ &\leq q^{n-1}\varrho(x_2, x_1).\end{aligned}$$

Odtud pro libovolná $p, n \in \mathbb{N}$ máme

$$\begin{aligned}\varrho(x_{n+p}, x_n) &\le \varrho(x_{n+p}, x_{n+p-1}) + \varrho(x_{n+p-1}, x_{n+p-2}) + \cdots + \varrho(x_{n+1}, x_n)\\ &\le q^{n+p-2}\varrho(x_2, x_1) + \cdots + q^{n-1}\varrho(x_2, x_1) = \frac{q^{n-1}(1-q^p)}{1-q}\varrho(x_2, x_1)\\ &\le \frac{q^{n-1}}{1-q}\varrho(x_2, x_1).\end{aligned}$$

Posloupnost $\{x_n\}$ je proto cauchyovská, úplnost (P, ϱ) zaručuje, že $x_n \to x_0$ pro jisté $x_0 \in P$. Bod x_0 je pevný bod, neboť díky spojitosti metriky máme

$$\varrho(x_0, Tx_0) \leftarrow \varrho(x_{n+1}, Tx_0) = \varrho(Tx_n, Tx_0) \le q\varrho(x_n, x_0) \to 0.$$

□

Příklad 11.9.3. Nechť $c_0 \in \mathbb{R}$. Ukažme, že existuje právě jedno řešení úlohy

$$x = \frac{1}{2}\sin x + c_0.$$

Pracujeme na $\mathbb{R}$ s obvyklou metrikou, což je úplný metrický (dokonce normovaný lineární) prostor a zobrazení T definujme jako

$$Tx = \frac{1}{2}\sin x + c_0.$$

Zobrazení T je kontrakce, neboť pro $x < y$ díky Lagrangeově větě o přírůstku funkce máme (níže $\xi \in (x, y)$)

$$|Ty - Tx| = \frac{1}{2}|\sin y - \sin x| = \frac{1}{2}\cos\xi|y - x| \le \frac{1}{2}|y - x|.$$

Podle Banachovy věty o kontrakci tedy existuje jednoznačně určené řešení zadaného problému, které lze zkonstruovat metodou postupných aproximací.

Příklad 11.9.4. Díky Banachově větě o kontrakci se dá snadno ukázat, že pro $a \in [0, 1]$ iterační metoda zadaná předpisem

$$x_0 = 0, \qquad x_{n+1} = x_n - \frac{1}{2}(x_n^2 - a),$$

definuje posloupnost konvergující k $\sqrt{a}$. Skutečně (v dalším se zabýváme jen případem $a \in (0, 1]$, pro $a = 0$ je vše zřejmé), zadefinujeme-li funkci $f\colon \mathbb{R} \to \mathbb{R}$ předpisem

$$f(t) = t - \frac{1}{2}(t^2 - a),$$

pak číslo $\sqrt{a}$ je jejím pevným bodem. Dále

$$f'(t) = 1 - t.$$

Odtud pro libovolné $x \in [0, \sqrt{a})$ máme podle Lagrangeovy věty o přírůstku funkce

$$f(x) - f(a) = f'(\xi)(x - a) = (1 - \xi)(x - a).$$

Pro libovolné zafixované $\delta > 0$ je proto f kontrakcí na $[\delta, \sqrt{a}]$ a podle Banachovy věty o kontrakci má jediný pevný bod $\sqrt{a}$. Na naši posloupnost tento výsledek aplikujeme následovně. Podle výpočtu uvedeného výše platí $x_1 \in (0, \sqrt{a})$. Nyní zafixujeme $\delta \in (0, x_1)$ a Banachovu větu o kontrakci použijeme na posloupnost $\{x_n\}_{n=1}^{\infty}$, úplný metrický prostor $C([\delta, \sqrt{a}])$ a kontrakci f.

Poznámka 11.9.5. (i) Předpoklad o úplnosti není možné vypustit. Kupříkladu $f(x) = \frac{x}{2}$ je kontrakce na $(0, 1)$, ale nemá zde pevný bod.
(ii) Funkce $f(x) = \frac{x}{2}$ nemá pevný bod na úplném prostoru $[1, 2]$. Není to totiž kontrakce na tomto prostoru (špatný obor hodnot).
(iii) Kontraktivitu není možné nahradit neexpanzivitou (ta připouští $q \leq 1$), jak ukazuje volba $f(x) = x + 1$ na $\mathbb{R}$.

11.10 Existenční věty pro ODR 1.řádu

Významnou aplikací Banachovy věty o kontrakci je důkaz Picard–Lindelöfovy existenční věty (existence a jednoznačnost řešení systémů lineárních obyčejných diferenciálních rovnic prvního řádu, Věta 10.3.5). Tento důležitý výsledek z teorie obyčejných diferenciálních rovnic si nyní dokážeme. Nejprve si připomeňme znění věty.

Věta 11.10.1 (Picard–Lindelöfova existenční věta)**.** *Nechť $F\colon \mathbb{R}^{n+1} \to \mathbb{R}^n$ je spojitá na otevřené množině $\Omega \subset \mathbb{R}^{n+1}$, $(x_0, y_0) \in \Omega$ a F je na Ω lokálně lipschitzovská vzhledem k poslední n-tici proměnných. Pak existuje $\delta > 0$ tak, že na intervalu $(x_0 - \delta, x_0 + \delta)$ existuje právě jedno řešení Cauchyovy úlohy pro systém rovnic $y' = F(x, y)$ s počáteční podmínkou $y(x_0) = y_0$.*

Důkaz. Abychom měli co nejjednodušší značení, podrobný důkaz uvedeme jen pro případ jedné obyčejné diferenciální rovnice. Na jeho konci uvedeme seznam změn, které vyžaduje důkaz v případě soustavy rovnic.

V dalším tedy uvažujme úlohu

$$y' = f(x, y) \qquad y(x_0) = x_0, \tag{11.10.1}$$

kde $f\colon \mathbb{R}^2 \to \mathbb{R}$, f je spojitá na otevřené množině $\Omega \subset \mathbb{R}^2$, $(x_0, y_0) \in \Omega$ a existují $\delta, L > 0$ tak, že $[x_0 - \delta, x_0 + \delta] \times [y_0 - \delta, y_0 + \delta] \subset \Omega$ a

$$|f(x, y_1) - f(x, y_2)| \leq L|y_1 - y_2| \tag{11.10.2}$$

kdykoliv $(x, y_1), (x, y_2) \in (x_0 - \delta, x_0 + \delta) \times (y_0 - \delta, y_0 + \delta)$.

Krok 1: integrální formulace úlohy.
Zkoumejme vztah

$$y(x) = y_0 + \int_{x_0}^{x} f(t, y(t))\,\mathrm{d}t. \tag{11.10.3}$$

Tvrdíme, že je-li $\tau > 0$, pak spojitá funkce y řeší úlohu (11.10.1) na $(x_0-\tau, x_0+\tau)$ právě tehdy, když splňuje (11.10.3) na $(x_0-\tau, x_0+\tau)$. Skutečně, pokud y řeší úlohu (11.10.1) na $(x_0-\tau, x_0+\tau)$, snadno nahlédneme, že $y \in C^1((x_0-\tau, x_0+\tau))$ a po integraci máme

$$y(x) - y_0 = y(x) - y(0) = \int_{x_0}^{x} y'(t)\,\mathrm{d}t = \int_{x_0}^{x} f(t, y(t))\,\mathrm{d}t.$$

Obrácenou implikaci získáme zderivováním (11.10.3) (můžeme zderivovat pravou stranu, protože $t \mapsto f(t, y(t))$ je spojitá funkce).

Krok 2: volba metrického prostoru (P, ϱ).
Díky spojitosti f na Ω existuje $K > 0$ splňující

$$|f(x,y)| \leq K \qquad \text{na } [x_0-\delta, x_0+\delta] \times [y_0-\delta, y_0+\delta].$$

Nyní zafixujme $\tau \in (0, \delta)$ tak malé, aby platilo

$$L\tau < 1 \qquad \text{a} \qquad [y_0 - K\tau, y_0 + K\tau] \subset [y_0-\delta, y_0+\delta].$$

Konečně, definujme prostor (P, ϱ) tak, že

$$P := \{\varphi \in C([x_0-\tau, x_0+\tau]) \colon \varphi(x) \in [y_0 - K\tau, y_0 + K\tau] \text{ na } [x_0-\tau, x_0+\tau]\}$$

a ϱ je maximová metrika.

Krok 3: úplnost metrického prostoru (P, ϱ).
Předpokládejme, že $\{\varphi_n\}$ je cauchyovská posloupnost v (P, ϱ). V důkazu úplnosti prostoru $C([a,b])$ opatřeného maximovou normou (Příklad 11.6.3 (vi)) jsme ukázali, že posloupnost, která je cauchyovská v maximové normě, vždy v této normě konverguje ke spojité funkci φ a platí pro ni

$$\varphi(x) = \lim_{n\to\infty} \varphi_n(x) \qquad \text{pro všechna } x \in [a,b].$$

Z toho již plyne úplnost našeho prostoru.

Krok 4: volba kontraktivního zobrazení.
Na prostoru (P, ϱ) definujme

$$\Phi \colon y \mapsto y_0 + \int_{x_0}^{x} f(t, y(t))\,\mathrm{d}t.$$

Pak platí $\Phi \colon P \to P$ (skutečně, díky vlastnostem integrálu s proměnnou horní mezí napravo vždy získáváme spojitou funkci, navíc má díky odhadu $|f(x,y)| \leq K$ požadovaný obor hodnot). Dále pro libovolná $y, z \in P$ dostáváme s využitím (11.10.2)

$$\begin{aligned}
\varrho\big(\Phi(y), \Phi(z)\big) &= \max_{[x_0-\tau, x_0+\tau]} \left| \int_{x_0}^{x} f(t, y(t))\,\mathrm{d}t - \int_{x_0}^{x} f(t, z(t))\,\mathrm{d}t \right| \\
&\leq \max_{[x_0-\tau, x_0+\tau]} \left| \int_{x_0}^{x} |f(t, y(t)) - f(t, z(t))|\,\mathrm{d}t \right| \\
&\leq |x_0 - x| L \max_{[x_0-\tau, x_0+\tau]} |y - z| \leq \tau L \varrho(y, z).
\end{aligned}$$

Protože $\tau L < 1$, dostali jsme, že Φ je kontrakce.

Krok 5: existence a jednoznačnost řešení.
Podle třetího a čtvrtého kroku můžeme na prostor (P, ϱ) a zobrazení Φ aplikovat Banachovu větu o kontrakci a dostáváme jednoznačné $y \in P$ splňující (11.10.3). Díky prvnímu kroku a skutečnosti, že jsme celý postup mohli provádět na libovolném podintervalu intervalu $[x_0 - \tau, x_0 + \tau]$ obsahujícím bod x_0, máme jednoznačné řešení úlohy (11.10.1).

Zbývá vysvětlit, jak bychom postupovali v případě soustavy rovnic

$$y' = F(x, y) \qquad y(x_0) = y_0,$$

kde $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}^n$, $F\colon \mathbb{R}^{n+1} \to \mathbb{R}^n$. Na vektorových funkcích bychom zavedli metriku

$$\varrho(y, z) = \max_{\substack{i \in \{1,\dots,n\} \\ x \in [x_0 - \tau, x_0 + \tau]}} |y_i(x) - z_i(x)|$$

a integrál bychom chápali vektorově, tedy

$$\int_{x_0}^{x} F(t, y(t))\,\mathrm{d}t = \Big(\int_{x_0}^{x} F_1(t, y(t))\,\mathrm{d}t, \dots, \int_{x_0}^{x} F_n(t, y(t))\,\mathrm{d}t \Big).$$

Ostatní modifikace důkazu jsou zřejmé. □

Poznámka 11.10.2. V důkazu Banachovy věty o kontrakci jsme pevný bod získali postupnými aproximacemi. Pro naši diferenciální rovnici máme například následující posloupnost přibližných řešení. Napřed zvolíme konstantní funkci

$$\varphi_0 \equiv y_0.$$

Pak postupně zavádíme pro každé $n \in \mathbb{N}$

$$\varphi_{n+1}(x) = y_0 + \int_{x_0}^{x} f(t, \varphi_n(t))\,\mathrm{d}t.$$

Příklad 11.10.3. Připomeňme si Příklad 10.3.10, část (ii). V případě rovnice $y' = y$ s počáteční podmínkou $y(0) = 1$ máme

$$\varphi_0 \equiv 1 \qquad \text{a} \qquad \varphi_{n+1}(x) = 1 + \int_0^x \varphi_n(t)\,\mathrm{d}t$$

na $(-\tau, \tau)$, kdykoliv $\tau < 1$ (v našem případě je $L = 1$). Proto

$$\begin{aligned}
\varphi_1(x) &= 1 + \int_0^x 1\,\mathrm{d}t = 1 + x \\
\varphi_2(x) &= 1 + \int_0^x (1 + t)\,\mathrm{d}t = 1 + x + \frac{1}{2}x^2 \\
\varphi_3(x) &= 1 + \int_0^x (1 + t + \frac{1}{2}t^2)\,\mathrm{d}t = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3,
\end{aligned}$$

atd. Postupně dostáváme Taylorův rozvoj exponenciály. Pokud nás zajímá, jak moc dobře funkce φ_n aproximuje skutečné řešení, stačí použít odhad z odvození cauchyovskosti v důkazu Banachovy věty o kontrakci, který v našem případě dává (v důkazu Picard–Lindelöfovy věty jsme viděli, že pro konstantu z definice kontrakce zde platí $q = L\tau$)

$$\varrho(\varphi_{n+p}, \varphi_n) \leq \frac{q^n}{1-q}\varrho(\varphi_1, \varphi_0) = \frac{\tau^n}{1-\tau} \max_{[-\tau,\tau]} |(1+x) - 1| = \frac{\tau^{n+1}}{1-\tau}.$$

Limitním přechodem $p \to \infty$ odtud dostáváme

$$\max_{[-\tau,\tau]} |\varphi - \varphi_n| \leq \frac{\tau^{n+1}}{1-\tau}.$$

Poznámka 11.10.4. Získaný odhad poněkud pokulhává za odhadem, který nám dává Lagrangeův tvar zbytku Taylorova polynomu funkce e^x

$$\max_{[-\tau,\tau]} |\varphi - \varphi_n| = \max_{[-\tau,\tau]} |R_{n+1}| = \max_{x\in[-\tau,\tau]} \frac{1}{(n+1)!}\mathrm{e}^{\xi_x}|x|^{n+1} \leq \frac{\mathrm{e}^\tau}{(n+1)!}\tau^{n+1}.$$

Je však nutné podotknout, že přístup přes odhad zbytku Taylorova polynomu požaduje znalost explicitního zápisu řešení, což je ale situace, kdy toto řešení nepotřebujeme aproximovat funkcemi φ_n.

Nyní si dokážeme Peanovu větu, tedy existenci řešení pro systémy obyčejných diferenciálních rovnic prvního řádu pro spojitou pravou stranu. Budeme potřebovat několik pomocných výsledků. Začneme další vlastností prostoru $C([a,b])$ opatřeného maximovou metrikou.

Definice 11.10.5. Nechť $M \subset C([a,b])$ je množina. Řekneme, že M je *stejně omezená* (neboli funkce z M jsou *stejně omezené*) na $[a,b]$, jestliže existuje $K > 0$ takové, že

$$|f(x)| \leq K \qquad \text{kdykoliv } x \in [a,b] \text{ a } f \in M.$$

Řekneme, že M je *stejně stejnoměrně spojitá* (neboli funkce z M jsou *stejně stejnoměrně spojité*) na $[a,b]$, jestliže pro každé $\varepsilon > 0$ existuje $\delta > 0$ takové, že

$$|f(x_1) - f(x_2)| \leq \varepsilon \qquad \text{kdykoliv } x_1, x_2 \in [a,b], |x_1 - x_2| < \delta \text{ a } f \in M.$$

Není těžké si uvědomit, že posloupnost funkcí z $C([a,b])$ konvergující v maximové metrice je stejně omezená a stejně stejnoměrně spojitá (připomeňte si důkaz úplnosti prostoru $C([a,b])$). Do určité míry platí i opačná implikace, kterou později použijeme.

Věta 11.10.6 (Arzelà–Ascoliho věta)**.** *Nechť $\{f_n\} \subset C([a,b])$ jsou stejně omezené a stejně stejnoměrně spojité funkce. Pak existuje podposloupnost $\{f_{n_k}\} \subset \{f_n\}$ konvergentní v maximové metrice.*

Důkaz. Seřaďme všechna racionální čísla z intervalu $[a, b]$ do posloupnosti $\{x_k\}$. Díky stejnoměrné omezenosti $\{f_n\}$ je posloupnost $\{f_n(x_1)\}$ omezená a proto lze podle Weierstrassovy věty vybrat podposloupnost $\{f_{1,n}\} \subset \{f_n\}$ tak, že $f_{1,n}(x_1)$ konverguje. Pokračujeme s bodem x_2 a podposloupností $\{f_{2,n}\} \subset \{f_{1,n}\}$ takovou, že $f_{2,n}(x_2)$ konverguje a tak dále. Zřejmě pak diagonální posloupnost $\{f_{n,n}\}$ splňuje

$$\lim_{n\to\infty} f_{n,n}(x_k) \qquad \text{existuje vlastní pro každé } k \in \mathbb{N}.$$

Ukažme, že tento výsledek implikuje cauchyovskost posloupnosti $\{f_{n,n}\}$ v maximové metrice. Zvolme $\varepsilon > 0$. Nechť $\delta > 0$ odpovídá číslu ε v definici stejné stejnoměrné spojitosti. Dále zřejmě existují racionální čísla $y_1, \ldots, y_l$ tak, že $\{(y_i - \delta, y_i + \delta)\}_{i=1}^l$ tvoří pokrytí $[a, b]$. Navíc podle předchozího pro každé $i \in \{1, \ldots, l\}$ existuje $n_i \in \mathbb{N}$ takové, že

$$|f_{n,n}(y_i) - f_{m,m}(y_i)| < \varepsilon \qquad \text{pro } n, m > n_i.$$

Položme $n_0 = \max\{n_1, \ldots, n_l\}$. Zafixujeme-li nyní libovolné $x \in [a, b]$ a vezmemeli k němu y_{i_0} tak, že $|x - y_{i_0}| < \delta$, díky předchozím výsledkům a stejné stejnoměrné spojitosti dostáváme pro $m, n > n_0$

$$\begin{aligned} |f_{n,n}(x) - f_{m,m}(x)| &\le |f_{n,n}(x) - f_{n,n}(y_{i_0})| \\ &\quad + |f_{n,n}(y_{i_0}) - f_{m,m}(y_{i_0})| + |f_{m,m}(y_{i_0}) - f_{m,m}(x)| < 3\varepsilon. \end{aligned}$$

Proto $\max_{[a,b]} |f_{n,n} - f_{m,m}| < 3\varepsilon$, tedy $\{f_{n,n}\}$ je cauchyovská v maximové metrice a úplnost $C([a, b])$ implikuje dokazovaný výsledek. □

Lemma 11.10.7. *Nechť $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ a $f : \mathbb{R} \to \mathbb{R}$ je spojitá na $[a, b]$ a afinní na $[t_{i-1}, t_i]$ pro všechna $i \in \{1, \ldots, n\}$. Označme*

$$k_i = \frac{f(t_i) - f(t_{i-1})}{t_i - t_{i-1}} \qquad \text{pro } i \in \{1, \ldots, n\}.$$

Pak pro všechna $x, y \in [a, b]$, $x \neq y$, platí

$$\min\{k_1, \ldots, k_n\} \le \frac{f(x) - f(y)}{x - y} \le \max\{k_1, \ldots, k_n\}.$$

Důkaz. Pokud $x \le t_j < t_{j+1} < \cdots < t_{l-1} < t_l \le y$, máme

$$\begin{aligned} f(y) - f(x) &= f(y) - f(t_l) + f(t_l) - f(t_{l-1}) + \cdots + f(t_{j+1}) - f(t_j) \\ &\quad + f(t_j) - f(x) \\ &= k_{l+1}(y - t_l) + k_l(t_l - t_{l-1}) + \cdots + k_{j+1}(t_{j+1} - t_j) + k_j(t_j - x) \\ &\le \max\{k_j, \ldots, k_{l+1}\}(y - x) \le \max\{k_1, \ldots, k_n\}(y - x) \end{aligned}$$

a tak dále. □

Nyní přistoupíme k důkazu Peanovy existenční věty. Uvedeme lehce odlišné (výstižnější) znění, ve kterém je uvedeno, na jakém okolí bodu reprezentujícího počáteční podmínku řešení existuje. Znění a důkaz uvedeme jen pro $n = 1$, modifikace pro vícerozměrný případ je jasná (interval $[y_0 - b, y_0 + b]$ se ve znění nahradí krychlí a v důkazu se každé složce vektorové funkce y věnujeme zvlášť).

Věta 11.10.8 (Peanova existenční věta)**.** *Nechť $a, b > 0$, $x_0 \in \mathbb{R}$, $y_0 \in \mathbb{R}$ a $F\colon \mathbb{R}^2 \to \mathbb{R}$ je spojitá na $[x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b]$. Nechť $M > 0$ splňuje*

$$M \geq \max_{[x_0-a,x_0+a]\times[y_0-b,y_0+b]} |F(x, y)|.$$

Položme $h := \min\{a, \frac{b}{M}\}$. Pak na intervalu $(x_0 - h, x_0 + h)$ existuje řešení Cauchyovy úlohy pro rovnici $y' = F(x, y)$ s počáteční podmínkou $y(x_0) = y_0$.

Důkaz. Krok 1: Konstrukce aproximativních řešení φ_n.
Na chvíli zafixujme $n \in \mathbb{N}$. Interval $[x_0 - h, x_0 + h]$ rozdělíme na $2n$ stejných dílků zavedením dělících bodů

$$t_j := x_0 + j\frac{h}{n} \qquad \text{pro } j = -n, -n + 1, \dots, -1, 0, 1, \dots, n - 1, n.$$

Aproximativní řešení φ_n konstruujeme jako spojitou po částech afinní funkci definovanou následovně. Na $[t_0, t_1]$ definujeme

$$\varphi_n(x) := y_0 + F(t_0, y_0)(x - t_0)$$

(připomeňme, že $t_0 = x_0$). Položme ještě $y_1 := \varphi_n(t_1)$. Na $[t_1, t_2]$ definujeme

$$\varphi_n(x) := y_1 + F(t_1, y_1)(x - t_1) \qquad \text{a} \qquad y_2 := \varphi_n(t_2).$$

Postupujeme indukcí a pro $j \in \{3, \dots, n\}$ definujeme na $[t_{j-1}, t_j]$

$$\varphi_n(x) := y_{j-1} + F(t_{j-1}, y_{j-1})(x - t_{j-1}) \qquad \text{a} \qquad y_j := \varphi_n(t_j).$$

Na intervalu $[x_0 - h, x_0]$ pracujeme podobně, jen se v každém kroku posouváme o jeden dílek doleva.

Krok 2: Konvergence aproximativních řešení.
Z konstrukce je vidět, že každá z funkcí φ_n splňuje $\varphi_n(x_0) = y_0$ a navíc je lipschitzovská s konstantou lipschitzovskosti rovnou M. Proto jsou splněny předpoklady Arzelà–Ascoliho věty a po přechodu k podposloupnosti máme $\varphi \in C([x_0 - h, x_0 + h])$ takovou, že

$$\max_{[x_0-h,x_0+h]} |\varphi_{n_k} - \varphi| \overset{k\to\infty}{\to} 0. \tag{11.10.4}$$

Krok 3: Funkce φ řeší diferenciální rovnici.
Protože $\varphi_{n_k}(x_0) = y_0$ pro všechna $k \in \mathbb{N}$, (11.10.4) implikuje $\varphi(x_0) = y_0$. Zbývá ukázat, že

$$\varphi'(x) = F(x, \varphi(x)) \qquad \text{na } (x_0 - h, x_0 + h).$$

Zafixujme $\varepsilon > 0$ a $x \in (x_0, x_0 + h)$ (případ $x \in (x_0 - h, x_0)$ je analogický, případ $x = x_0$ dokonce o trochu jednodušší). Ze spojitosti F existuje $\delta > 0$ tak, že

$$|t - x| \leq \delta \wedge |y - \varphi(x)| \leq \delta \qquad \Longrightarrow \qquad |F(t, y) - F(x, \varphi(x))| < \varepsilon. \tag{11.10.5}$$

Označme $h_1 := \min\{\frac{\delta}{2}, \frac{\delta}{6M}\}$. Dále zafixujme libovolné $s \in (-h_1, h_1) \setminus \{0\}$. Nyní ještě zafixujme $k \in \mathbb{N}$ tak velké, aby

$$\max_{[x_0-h,x_0+h]} |\varphi_{n_k} - \varphi| < \min\Big\{|s|\varepsilon, \frac{\delta}{2}\Big\} \tag{11.10.6}$$

(použili jsme (11.10.4)) a aby pro vzdálenost dělících bodů funkce φ_{n_k} platilo $\frac{h}{n_k} < h_1$. Díky druhé vlatnosti existuje index j tak, že odpovídající dělící bod splňuje

$$x - 2h_1 < t_j \leq x - h_1.$$

Odtud

$$|t_j - x| < 2h_1 \qquad \text{a} \qquad |t_j - (x+s)| < 3h_1.$$

Dále díky M-lipschitzovskosti funkce φ_{n_k} a nerovnosti (11.10.6) plyne pro každé $t \in [t_j, \max\{x, x+s\}]$

$$\begin{aligned}|\varphi(x) - \varphi_{n_k}(t)| &\leq |\varphi(x) - \varphi_{n_k}(x)| + |\varphi_{n_k}(x) - \varphi_{n_k}(t)| \\ &\leq \frac{\delta}{2} + M|t-x| \leq \frac{\delta}{2} + 3Mh_1 \leq \frac{\delta}{2} + 3M\frac{\delta}{6M} = \delta.\end{aligned}$$

Díky tomuto odhadu, Lemmatu 11.10.7 před větou (používáme jej na intervalu $[t_j, \max\{x, x+s\}]$) a metodě konstrukce funkce φ_{n_k} máme

$$\begin{aligned}\min_{[x-\delta,x+\delta]\times[\varphi(x)-\delta,\varphi(x)+\delta]} F(t,y) &\leq \frac{\varphi_{n_k}(x+s) - \varphi_{n_k}(x)}{s} \\ &\leq \max_{[x-\delta,x+\delta]\times[\varphi(x)-\delta,\varphi(x)+\delta]} F(t,y).\end{aligned} \tag{11.10.7}$$

Celkově dostáváme z (11.10.6) a z (11.10.7) kombinovaného s (11.10.5)

$$\begin{aligned}&\left|\frac{\varphi(x+s) - \varphi(x)}{s} - F(x,\varphi(x))\right| \\ &\qquad \leq \left|\frac{\varphi(x+s) - \varphi(x)}{s} - \frac{\varphi_{n_k}(x+s) - \varphi_{n_k}(x)}{s}\right| \\ &\qquad\qquad + \left|\frac{\varphi_{n_k}(x+s) - \varphi_{n_k}(x)}{s} - F(x,\varphi(x))\right| \\ &\qquad < 2\varepsilon + \varepsilon = 3\varepsilon.\end{aligned}$$

Odtud plyne $\varphi'(x) = F(x, \varphi(x))$. □

11.11 Limita a spojitost zobrazení na metrických prostorech

Definice 11.11.1 (Limita zobrazení)**.** Nechť (P_1, ϱ_1) a (P_2, ϱ_2) jsou metrické prostory, $\varphi\colon P_1 \to P_2$, $x_0 \in P_1$ je hromadným bodem D_φ a $y_0 \in P_2$. Řekneme, že zobrazení φ má v bodě x_0 *limitu* y_0, jestliže

$$\forall \varepsilon > 0\, \exists \delta > 0 \qquad x \in \mathcal{P}_\delta(x_0) \cap D_\varphi \implies \varphi(x) \in \mathcal{U}_\varepsilon(y_0).$$

V takovém případě píšeme $\lim_{x \to x_0} \varphi(x) = y_0$, nebo $\varphi(x) \to y_0$ pro $x \to x_0$.

Poznámka 11.11.2. (i) V případě $\mathbb{R}$ s obvyklou metrikou se nová definice shoduje s naší starou definicí vlastní limity ve vlastních bodech.
(ii) Nevlastními limitami se zde nezabýváme, neboť $\mathbb{R}^*$ není metrický prostor (prvky $\pm\infty$ nemají definovanou konečnou vzdálenost od ostatních prvků).

Definice 11.11.3 (Spojitost zobrazení)**.** Nechť (P_1, ϱ_1) a (P_2, ϱ_2) jsou metrické prostory, $\varphi\colon P_1 \to P_2$ a $x_0 \in D_\varphi$. Řekneme, že zobrazení φ je v bodě x_0 *spojité*, jestliže

$$\forall \varepsilon > 0\, \exists \delta > 0 \qquad x \in \mathcal{U}_\delta(x_0) \cap D_\varphi \implies \varphi(x) \in \mathcal{U}_\varepsilon(\varphi(x_0)).$$

Poznámka 11.11.4. Limita a spojitost jsou invariantní vůči přechodu k ekvivalentní metrice jak ve vzoru, tak v obraze.

Poznámka 11.11.5. Je-li x_0 izolovaný bod D_φ, je v něm φ automaticky spojité. V opačném případě je spojitost ekvivalentní podmínce $\lim_{x\to x_0} f(x) = f(x_0)$.

Příklad 11.11.6. (i) Uvažme $\mathbb{R}^2$ s eukleidovskou metrikou, $\mathbb{R}$ s obvyklou metrikou a definujme $\varphi\colon \mathbb{R}^2 \to \mathbb{R}$ předpisem

$$\varphi(x,y) = \begin{cases} \frac{xy^2}{x^2+y^2} & \text{pro } (x,y) \neq (0,0) \\ 0 & \text{pro } (x,y) = (0,0). \end{cases}$$

Pomocí Youngovy nerovnosti dostáváme pro $(x,y) \neq (0,0)$

$$0 \le |\varphi(x,y)| = \left|\frac{xy^2}{x^2+y^2}\right| \le \frac{1}{2}|y| \le \frac{1}{2}\sqrt{x^2+y^2}.$$

Odtud vidíme, že $\lim_{(x,y)\to(0,0)} \varphi(x,y) = 0$ a φ je spojité v počátku.
(ii) Uvažme stejné prostory a zobrazení

$$\varphi(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{pro } (x,y) \neq (0,0) \\ 0 & \text{pro } (x,y) = (0,0). \end{cases}$$

Pro libovolné $x \in \mathbb{R} \setminus \{0\}$ máme

$$\varphi(x,0) = 0 \qquad \text{a} \qquad \varphi(x,x) = \frac{1}{2}.$$

Protože se v každém prstencovém okolí bodu $(0,0)$ vyskytují body výše uvedených typů, $\lim_{(x,y)\to(0,0)} \varphi(x,y)$ neexistuje a φ není spojité v počátku.
(iii) Uvažme prostor $C^1([0,1])$ s normou

$$\|f\|_{C^1} = \max_{[0,1]} |f'| + \max_{[0,1]} |f|$$

a prostor $C([0,1])$ s maximovou normou, tedy $\|f\|_\infty = \max_{[0,1]} f$. Definujme zobrazení $\varphi\colon C^1([0,1]) \to C([0,1])$ předpisem $\varphi(f) = f'$. Pak φ je spojité v každém $f \in C^1([0,1])$, neboť pro každé $g \in C^1([0,1])$ máme při volbě $\delta := \varepsilon$

$$\|g-f\|_{C^1} < \delta \implies \max_{[0,1]} |g'-f'| < \delta \iff \|g'-f'\|_\infty < \varepsilon \iff \|\varphi(g)-\varphi(f)\|_\infty < \varepsilon.$$

Poznámka 11.11.7. Pro zobrazení $\varphi\colon P_1 \to P_2$, kde (P_1, ϱ_1) a (P_2, ϱ_2) jsou metrické prostory, a množinu $A \subset P_1$ se někdy používají pojmy limita vzhledem k

A a spojitost vzhledem k A. Ty se zavádějí tak, že nejprve zúžíme definiční obor na $D_\varphi \cap A$ a na výsledné zobrazení teprve pak použijeme definici limity či spojitosti.

Povšimněte si, že spojitost vzhledem k A není totéž co spojitost na A. Například Dirichletova funkce je spojitá vzhledem k $\mathbb{Q}$ ve kterémkoliv bodě z $\mathbb{Q}$ (z definičního oboru jsme vyloučili všechna iracionální čísla, zbyly jen body s funkční hodnotou 1). Dirichletova funkce však není (klasicky) spojitá v žádném bodě.

Těmto novým pojmům nemusíme věnovat v budování teorie větší pozornost, vše se dá vyřešit zúžením definičního oboru. Na to jsme v naší teorii připraveni, neboť připouštíme zobrazení, jejichž definiční obor není celá množina P_1.

V dalším si uvedeme několik vět, jejichž důkazy se získají jen minimální modifikací důkazů odpovídajících tvrzení z teorie funkcí jedné reálné proměnné.

Věta 11.11.8 (Heineho věta)**.** *Nechť (P_1, ϱ_1) a (P_2, ϱ_2) jsou metrické prostory, $\varphi\colon P_1 \to P_2$, $x_0 \in P_1$ je hromadným bodem D_φ a $y_0 \in P_2$. Pak*
(i) $\lim_{x\to x_0} \varphi(x) = y_0$ *právě tehdy, když pro každou posloupnost $\{x_n\} \subset D_\varphi \setminus \{x_0\}$ splňující $x_n \to x_0$ platí $\varphi(x_n) \to x_0$.*
(ii) $\lim_{x\to x_0} \varphi(x)$ *existuje právě tehdy, když pro každou posloupnost $\{x_n\} \subset D_\varphi \setminus \{x_0\}$ splňující $x_n \to x_0$ existuje limita posloupnosti $\{\varphi(x_n)\}$.*

Věta 11.11.9 (O B-C podmínce)**.** *Nechť (P_1, ϱ_1) a (P_2, ϱ_2) jsou metrické prostory, (P_2, ϱ_2) je úplný, $\varphi\colon P_1 \to P_2$ a $x_0 \in P_1$ je hromadným bodem D_φ. Pak $\lim_{x\to x_0} \varphi(x)$ existuje právě tehdy, když zobrazení φ splňuje B-C podmínku ($\forall \varepsilon > 0\, \exists \delta > 0 \quad x, y \in \mathcal{P}_\delta(x_0) \cap D_\varphi \Rightarrow \varrho_2(\varphi(x), \varphi(y)) < \varepsilon$).*

Věta 11.11.10 (O spojitosti složeného zobrazení)**.** *Nechť (P_1, ϱ_1) , (P_2, ϱ_2) a (P_3, ϱ_2) jsou metrické prostory, $\varphi\colon P_1 \to P_2$ a $\psi\colon P_2 \to P_3$. Je-li φ spojité v $x_0 \in P_1$ a ψ spojité v $\varphi(x_0)$, pak $\psi \circ \varphi$ je spojité v x_0.*

Věta 11.11.11 (O limitě složeného zobrazení)**.** *Nechť (P_1, ϱ_1) , (P_2, ϱ_2) a (P_3, ϱ_2) jsou metrické prostory, $\varphi\colon P_1 \to P_2$ a $\psi\colon P_2 \to P_3$ a nechť $x_0 \in P_1$. Nechť $\lim_{x\to x_0} \varphi(x) = y_0 \in P_2$, $\lim_{y\to y_0} \psi(y) = z_0 \in P_3$, x_0 je hromadným bodem $D_{\psi\circ\varphi}$ a je splněna alespoň jedna z podmínek*
(i) *existuje prstencové okolí bodu x_0, kde vnitřní zobrazení φ nenabývá své limitní hodnoty y_0*
(ii) *vnější zobrazení ψ je spojité v bodě y_0.*
Pak $\lim_{x\to x_0} (\psi \circ \varphi)(x) = z_0$.

Věta 11.11.12 (Cantorova věta o stejnoměrné spojitosti)**.** *Nechť (P_1, ϱ_1), (P_2, ϱ_2) jsou metrické prostory, $\varphi\colon P_1 \to P_2$ je spojité na $A \subset P_1$ a A je kompaktní. Pak φ je stejnoměrně spojité na A ($\forall \varepsilon > 0\, \exists \delta > 0 \quad x, y \in A \wedge \varrho_1(x, y) < \delta \Rightarrow \varrho_2(\varphi(x), \varphi(y)) < \varepsilon$).*

Ve speciálním případě zobrazení z $\mathbb{R}^N$ do $\mathbb{R}$ (na $\mathbb{R}$ uvažujeme obvyklou metriku, na $\mathbb{R}^N$ metriku odvozenou od libovolné normy) máme ještě další výsledky, jejichž důkaz se opět získá jen drobnou modifikací důkazu z jednodimenzionálního případu.

Věta 11.11.13 (Aritmetika limit)**.** *Nechť* $\varphi, \psi\colon \mathbb{R}^N \to \mathbb{R}$*, kde oba prostory bereme s obvyklými metrikami, a nechť* $x_0 \in \mathbb{R}^N$ *je hromadný bod* $D_\varphi \cap D_\psi$*. Nechť* $\lim_{x\to x_0} \varphi(x) = A \in \mathbb{R}$ *a* $\lim_{x\to x_0} \psi(x) = B \in \mathbb{R}$*. Pak*
(i) $\lim_{x\to x_0} \big(\varphi(x) + \psi(x)\big) = A + B$
(ii) $\lim_{x\to x_0} \varphi(x)\psi(x) = AB$
(iii) *pokud* $B \neq 0$*, platí* $\lim_{x\to x_0} \frac{\varphi(x)}{\psi(x)} = \frac{A}{B}$.

Poznámka 11.11.14. Automaticky také platí aritmetika spojitosti.

Věta 11.11.15 (O dvou strážnících)**.** *Nechť* $\varphi, \psi, \eta\colon \mathbb{R}^N \to \mathbb{R}$*, kde oba prostory bereme s obvyklými metrikami,* $D_\varphi = D_\psi = D_\eta$*, a* $x_0 \in \mathbb{R}^N$ *je hromadný bod* D_φ*. Nechť* $\lim_{x\to x_0} \varphi(x) = \lim_{x\to x_0} \eta(x) = A \in \mathbb{R}$ *a* $\varphi \le \psi \le \eta$ *na* $\mathcal{P}_\delta(x_0) \cap D_\varphi$ *pro jisté* $\delta > 0$*. Pak* $\lim_{x\to x_0} \psi(x) = A$*.*

Věta 11.11.16 (O nabývání extrémů spojitou funkcí)**.** *Nechť* $\varphi\colon \mathbb{R}^N \to \mathbb{R}$*, kde oba prostory bereme s obvyklými metrikami, je spojité na kompaktní množině* $A \subset \mathbb{R}^N$*. Pak zde nabývá svého maxima a minima (existují* $x_1, x_2 \in A$ *tak, že* $\max_A f = f(x_1)$ *a* $\min_A f = f(x_2)$*).*

Poznámka 11.11.17. Je-li $\varphi\colon \mathbb{R}^N \to \mathbb{R}$ spojité v bodě $a = [a_1, a_2, \dots, a_N] \in \mathbb{R}^N$, snadno se dá ověřit (použijte maximovou normu), že funkce φ je spojitá v jednotlivých proměnných v odpovídajících bodech, přesněji $t \mapsto \varphi(t, a_2, a_3, \dots, a_N)$ je spojitá v bodě a_1, $t \mapsto \varphi(t, a_2, a_3, \dots, a_N)$ je spojitá v bodě a_2 a tak dále pro všech N složek funkce φ. Implikace se obrátit nedá. Definujeme-li totiž

$$\varphi(x,y) = \begin{cases} 0 & \text{pro } xy \neq 0 \\ 1 & \text{pro } xy = 0 \end{cases}$$

(funkce má hodnotu 1 na osovém kříži, všude jinde je nulová), pak naše funkce není spojitá v bodě $[0,0]$. Naproti tomu podle proměnné x i y zvlášť funkce spojitá v počátku (teď už jednodimenzionálním) je, neboť tyto funkce pracují pouze s hodnotami na osovém kříži.

Podobně jako u funkcí jedné proměnné, i ve vyšší dimenzi se spojitost nejčastěji ověřuje pomocí aritmetiky spojitosti (počítání limit přichází na řadu pouze v problematických bodech).

Příklad 11.11.18. Ukažme, že funkce $\varphi(x,y) = x^2 + xy + \sin(xy)$ je spojitá na $\mathbb{R}^2$. Předně si dokážeme, že funkce $\eta\colon (x,y) \mapsto x$ je spojitá na $\mathbb{R}^2$. Zvolme $(x_0, y_0) \in \mathbb{R}^2$. Jednodimenzionální funkce $\psi\colon x \mapsto x$ je spojitá na celém $\mathbb{R}$. Speciálně k zadanému $\varepsilon > 0$ existuje $\delta > 0$ takové, že

$$x \in \mathcal{P}_\delta(x_0) \qquad \Longrightarrow \qquad \psi(x) \in \mathcal{U}_\varepsilon(\psi(x_0)).$$

Odtud díky tomu, že $\eta(x,y) = \psi(x)$ na $\mathbb{R}^2$, máme

$$(x,y) \in \mathcal{P}_\delta(x_0) \times \mathbb{R} \qquad \Longrightarrow \qquad \eta(x,y) \in \mathcal{U}_\varepsilon(\psi(x_0)).$$

Bez ohledu na normu zvolenou na $\mathbb{R}^2$ dostáváme

$$\mathcal{P}_\delta((x_0, y_0)) \subset \mathcal{P}_\delta(x_0) \times \mathbb{R}$$

(nalevo $\mathcal{P}$ značí dvoudimenzionální prstencová okolí, napravo jednodimenzionální). Předchozí úvahy se zřejmě dají zobecnit: pokud funkce více proměnných nezávisí na některých proměnných, její spojitost je ekvivalentní spojitosti funkce získané její restrikcí do prostoru nižší dimenze odpovídajícímu proměnným, na kterých funkce závisí.

Nyní již stačí použít aritmetiku spojitosti na první dva členy v zadání funkce φ, spojitost třetího členu plyne z Věty o spojitosti složeného zobrazení (Věta 11.11.10). Celkově díky aritmetice spojitosti je φ spojitá na $\mathbb{R}^2$.

Představme si ještě jednu charakterizaci spojitosti.

Věta 11.11.19 (O vzoru otevřených množin při spojitém zobrazení)**.** *Mějme metrické prostory (P_1, ϱ_1) a (P_2, ϱ_2), $\varphi\colon P_1 \to P_2$ a $D_\varphi = P_1$. Pak φ je spojité na P_1 právě tehdy, když pro každou otevřenou množinu $A \subset P_2$ je její vzor $\varphi^{-1}(A)$ otevřený.*

Důkaz. „$\Rightarrow$“ Zvolme $A \subset P_2$ otevřenou. Pokud $\varphi^{-1}(A) = \emptyset$, jsme hotovi. V opačném případě zafixujme $x \in \varphi^{-1}(A)$. Odtud $\varphi(x) \in A$ a otevřenost A dává $\varepsilon > 0$ takové, že $\mathcal{U}_\varepsilon(\varphi(x)) \subset A$. Ze spojitosti existuje $\delta > 0$ takové, že

$$\varphi(\mathcal{U}_\delta(x)) \subset \mathcal{U}_\varepsilon(\varphi(x)) \subset A.$$

Ukázali jsme tedy, že $\mathcal{U}_\delta(x) \subset \varphi^{-1}(A)$. Protože $x \in \varphi^{-1}(A)$ bylo libovolné, dokázali jsme otevřenost $\varphi^{-1}(A)$.
„$\Leftarrow$“ Pro každé $x \in P_1$ a $\varepsilon > 0$ je $\varphi^{-1}(\mathcal{U}_\varepsilon(\varphi(x)))$ otevřená množina. Existuje proto $\delta > 0$ takové, že

$$\mathcal{U}_\delta(x) \subset \varphi^{-1}(\mathcal{U}_\varepsilon(\varphi(x))) \qquad \Longrightarrow \qquad \varphi(\mathcal{U}_\delta(x)) \subset \mathcal{U}_\varepsilon(\varphi(x)).$$

Dokázali jsme implikaci $y \in \mathcal{U}_\delta(x) \;\Rightarrow\; f(y) \in \mathcal{U}_\varepsilon(\varphi(x))$ a tím jsme ověřili spojitost φ v x. $\square$

Poznámka 11.11.20. Připomeňme, že vzor množiny $A \subset P_2$ je definován jako

$$\varphi^{-1}(A) = \{x \in P_1 \colon \varphi(x) \in A\}.$$

Množina A nemusí být podmnožinou oboru hodnot.

Poznámka 11.11.21. (i) Předchozí věta se často používá k důkazu otevřenosti množin. Stačí najít vhodné spojité zobrazení a vhodnou otevřenou množinu tak, aby naše množina byla vzorem nalezené množiny při nalezeném zobrazení.
(ii) Je-li φ definováno na celém P_1, pak pro libovolnou množinu $A \subset P_2$ je doplněk jejího vzoru zároveň vzorem jejího doplňku ($\varphi^{-1}(P_2 \setminus A) = P_1 \setminus \varphi^{-1}(A)$). Odtud dostáváme ještě třetí ekvivalentní výrok a sice, že vzor každé uzavřené množiny je uzavřený.

Příklad 11.11.22. Ukažme, že množina $M := \{(x,y) \in \mathbb{R}^2 \colon x^2 < \sin(xy)\}$ je otevřená. Definujme $\varphi\colon (x,y) \mapsto x^2 - \sin(xy)$. Pomocí aritmetiky spojitosti snadno ověříme, že se jedná o spojitou funkci na celém $\mathbb{R}^2$. Pak $M = \varphi^{-1}((-\infty,0))$, a proto je otevřená ($(-\infty,0)$ je otevřená).

Poznámka 11.11.23. (i) Ve větě je důležitý předpoklad $D_\varphi = P_1$. Bez tohoto předpokladu věta neplatí. Stačí uvážit

$$\{x \in \mathbb{R}\colon \log^2 x \geq 0\} = (0,\infty),$$

což není uzavřená množina, třebaže je to vzor uzavřené množiny $[0,\infty)$ při spojitém zobrazení $x \mapsto \log^2 x$. Tímto bychom čtenáře také rádi varovali před často se vyskytujícím mylným názorem, že množiny zavedené pomocí neostré nerovnosti jsou uzavřené a množiny zavedené pomocí ostré nerovnosti jsou otevřené.
(ii) Pokud při aplikaci věty nezvolíme reprezentaci šťastně, některé informace se nedozvíme. Uvážíme-li například

$$\sin^{-1}([-1,1]) = \sin^{-1}((-2,2)) = \sin^{-1}([-3,3)) = \sin^{-1}(\mathbb{R}) = \mathbb{R},$$

pak první reprezentace reálných čísel nám pomohla odhalit jen jejich uzavřenost, druhá jen otevřenost, třetí žádnou z těchto vlastností, čtvrtá obě vlastnosti.

Poznámka 11.11.24. Analogie předchozí věty neplatí pro obraz. Například interval $(-\pi,\pi)$ je otevřená množina, ale $\sin((-\pi,\pi)) = [-1,1]$ je uzavřená množina.

V případě kompaktních množin, však analogie pro obraz při spojitém zobrazení platí.

Věta 11.11.25 (O obrazu kompaktní množiny při spojitém zobrazení)**.** *Nechť (P_1,ϱ_1) a (P_2,ϱ_2) jsou metrické prostory, $\varphi\colon P_1 \to P_2$, $D_\varphi = P_1$, $A \subset P_1$ je kompaktní a φ je spojité na A. Pak $\varphi(A)$ je kompaktní.*

Důkaz. Nechť $\{y_n\} \subset \varphi(A)$, pak existují body $x_n \in A$ takové, že $y_n = \varphi(x_n)$ pro každé $n \in \mathbb{N}$. Díky kompaktnosti A dostáváme $\{x_{n_k}\} \subset \{x_n\}$ takovou, že x_{n_k} konverguje k jistému $x \in A$. Díky spojitosti φ dostáváme $y_{n_k} = \varphi(x_{n_k}) \to \varphi(x)$. Posloupnost $\{y_n\}$ má tedy podposloupnost konvergentní v A a jsme hotovi. □

Poznámka 11.11.26. Vzor kompaktní množiny při spojitém zobrazení být kompaktní nemusí. Uvažme například

$$\mathbb{R} = \sin^{-1}([-1,1]).$$

Poznámka 11.11.27. Pokud je v předchozí větě zobrazení φ navíc prosté na A, je $\varphi^{-1}\colon \varphi(A) \to A$ spojitá na A. Skutečně, pokud by tomu tak nebylo, měli bychom posloupnost $\{y_n\} \subset \varphi(A)$ a $y_0 \in \varphi(A)$ takové, že $y_n \to y_0$ v P_2, ale pro $x_n := \varphi^{-1}(y_n)$ s $x_0 := \varphi^{-1}(y_0)$ by platilo

$$\varrho_1(x_n,x_0) \geq \varepsilon$$

pro jisté $\varepsilon > 0$. Na druhou stranu, díky kompaktnosti A umíme najít $x \in A$ a $\{x_{n_k}\} \subset \{x_n\}$ tak, aby $x_{n_k} \to x$. Spojitost metriky pak dává $\varrho(x_0,x) \geq \varepsilon$ (tedy $x \neq x_0$) a spojitost φ dává $y_{n_k} = \varphi(x_{n_k}) \to \varphi(x)$. Proto $\varphi(x) = y_0 = \varphi(x_0)$ a máme spor s prostotou φ na A.

Poznámka 11.11.28. Předchozí věta se dá také používat k elegantním důkazům spojitosti zobrazení získaných pomocí zobrazení, jejichž spojitost už máme dokázanou. Nechť například $(P_1, \varphi_1), (P_2, \varphi_2), (P_3, \varphi_3)$ jsou metrické prostory, $\varphi\colon P_1 \to P_2$ je spojité na P_1 a $\psi\colon P_2 \to P_3$ je spojité na P_2. Je-li $A \subset P_3$ libovolná otevřená množina, pak díky spojitosti ψ je $\psi^{-1}(A)$ také otevřená. Proto díky spojitosti φ je otevřená také množina $(\psi \circ \varphi)^{-1}(A) = \varphi^{-1}(\psi^{-1}(A))$. Dokázali jsme, že $\psi \circ \varphi$ je spojité na P_1.

Poznámka 11.11.29. (i) Podobor matematické analýzy, který se zabývá podrobněji metrickými prostory a jim příbuznými tématy, se nazývá funkcionální analýza. Funkcionální analýza navíc navazuje na lineární algebru a topologii.
(ii) Topologie se zabývá vztahem mezi vlastnostmi prostoru a vlastnostmi sady jeho otevřených množin (připomeňme, že pro metrické prostory umíme konvergenci zavést dvěma způsoby: jednak přes metriku pomocí podmínky $\varrho(x_n, x) \to 0$, ale také pomocí okolí, což jsou otevřené koule v uvažovaném metrickém prostoru). Oproti našim zvyklostem, v topologii bývá zvykem okolím bodu nazývat jakoukoliv množinu obsahující nějakou otevřenou kouli centrovanou v uvedeném bodě.

Kapitola 12

Diferenciální počet funkcí více proměnných

V kapitole o metrických prostorech jsme se již zabývali limitou a spojitostí funkcí více proměnných. Nyní se budeme zabývat problémy souvisejícími s derivováním.

12.1 Parciální derivace, derivace ve směru, totální diferenciál

Nejprve si připomeneme definici parciální derivace, kterou jsme si již představili. Pak přistoupíme k novým pojmům.

Definice 12.1.1 (Parciální derivace)**.** Nechť $a \in \mathbb{R}^N$, $i \in \{1, \dots, N\}$ a $f : \mathbb{R}^N \to \mathbb{R}$ je definována na množině $\{a_1\} \times \cdots \times \{a_{i-1}\} \times (a_i - \delta, a_i + \delta) \times \{a_{i+1}\} \times \cdots \times \{a_N\}$ pro jisté $\delta > 0$. Jestliže existuje vlastní limita

$$\lim_{h \to 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_N) - f(a)}{h},$$

pak se nazývá *parciální derivace* funkce f podle i-té proměnné v bodě a a značí se $\frac{\partial f}{\partial x_i}(a)$ nebo $f_{x_i}(a)$. *Druhá parciální derivace* podle proměnných x_i a x_j je definována vztahem $\frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i})(a)$ (pokud má výraz smysl) a značí se $\frac{\partial^2 f}{\partial x_j \partial x_i}(a)$ nebo $f_{x_j x_i}(a)$. Pokud $i = j$, první verze zápisu se zkracuje na $\frac{\partial^2 f}{\partial x_i^2}(a)$. Pokud $i \neq j$, hovoříme o *smíšené parciální derivaci.* Analogicky pro vyšší parciální derivace.

Pro $\Omega \subset \mathbb{R}^N$ otevřenou a $k \in \mathbb{N}_0$ značí $C^k(\Omega)$ množinu funkcí, které mají spojité všechny parciální derivace až do řádu k na Ω. Opět zavádíme

$$C^\infty(\Omega) = \bigcap_{k \in \mathbb{N}_0} C^k(\Omega).$$

Poznámka 12.1.2. Je-li $\{e_1, \dots, e_n\}$ kanonická báze v $\mathbb{R}^N$, definiční vztah pro parciální defivaci lze psát jako

$$\frac{\partial f}{\partial x_i}(a) = \lim_{h\to 0} \frac{f(a+he_i) - f(a)}{h}.$$

Poznámka 12.1.3. Protože pojem parciální derivace byl zaveden za pomoci funkce jedné reálné proměnné a její derivace, okamžitě máme k dispozici aritmetiku (vlastní derivace) a v některých jednoduchých (jednodimenzionálních z hlediska vnitřní funkce) případech máme také k dispozici Větu o derivaci složené funkce (Věta 3.3.14.) Pro obecný případ derivace složené funkce si později odvodíme takzvané *řetízkové pravidlo*.

Příklad 12.1.4. (i) Nechť $f(x,y) = x^2 + y^2$ na $\mathbb{R}^2$. Pak zde platí

$$\frac{\partial f}{\partial x} = 2x, \qquad \frac{\partial f}{\partial y} = 2y, \qquad \frac{\partial^2 f}{\partial x^2} = 2, \qquad \frac{\partial^2 f}{\partial y^2} = 2, \qquad \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 0$$

a rovněž libovolná parciální derivace řádu 3 a více je identicky nulová. Vidíme také, že $f \in C^\infty(\mathbb{R}^2)$.
(ii) Nechť $f(x,y) = x\sin(xy)$ na $\mathbb{R}^2$. Pak zde platí

$$\frac{\partial f}{\partial x} = \sin(xy) + xy\cos(xy) \qquad a \qquad \frac{\partial f}{\partial y} = x^2\cos(xy).$$

Opět se dá nahlédnout, že $f \in C^\infty(\mathbb{R}^2)$.
(iii) Nechť $f(x,y) = (xy)^{\frac{1}{3}}$ na $\mathbb{R}^2$. Pak přímý výpočet (pomocí aritmetiky derivace) dává

$$\frac{\partial f}{\partial x} = \frac{1}{3}x^{-\frac{2}{3}}y^{\frac{1}{3}} \qquad \text{kdykoliv } x \neq 0.$$

Označíme-li nyní pro zafixované $y \in \mathbb{R}$ funkci $g\colon x \mapsto f(x,y)$, pak g je spojitá na $\mathbb{R}$ a Věta o limitě derivací (Věta 6.3.9) dává

$$g'(0) = \lim_{x\to 0} \frac{1}{3}x^{-\frac{2}{3}}y^{\frac{1}{3}} = \begin{cases} 0 & \text{pro } y = 0 \\ \operatorname{sign} y \cdot \infty & \text{pro } y \neq 0. \end{cases}$$

Protože pojem parciální derivace připouští jen konečná čísla, máme $\frac{\partial f}{\partial x}(0,0) = 0$ a $\frac{\partial f}{\partial x}(0,y)$ neexistuje pro $y \neq 0$. Uvedené parciální derivace jsme také mohli počítat přímo z definice a dostat

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

a pro $y \neq 0$

$$\frac{\partial f}{\partial x}(0,y) = \lim_{h\to 0} \frac{f(h,y) - f(0,y)}{h} = \lim_{h\to 0} \frac{(yh)^{\frac{1}{3}} - 0}{h} = y^{\frac{1}{3}} \cdot \infty,$$

tedy tato parciální derivace neexistuje. Pro druhou parciální derivaci podle x máme

$$\frac{\partial^2 f}{\partial x^2} = -\frac{2}{9}x^{-\frac{5}{3}}y^{\frac{1}{3}} \qquad \text{kdykoliv } x \neq 0.$$

Snadno opět získáme $\frac{\partial^2 f}{\partial x^2}(0,0) = 0$. V bodech tvaru $(0,y)$, $y \neq 0$, neexistuje $\frac{\partial^2 f}{\partial x^2}$ už jenom proto, že tam neexistuje $\frac{\partial f}{\partial x}$. Analogické výsledky platí pro parciální derivaci podle y. Snadno se nahlédne, že $f \in C(\mathbb{R}^2)$ a pro každou otevřenou množinu $\Omega \subset \mathbb{R}^2$ platí

$$f \in C^1(\Omega) \quad \Longleftrightarrow \quad f \in C^\infty(\Omega) \quad \Longleftrightarrow \quad \Omega \text{ neprotíná osový kříž.}$$

Zobecněním parciální derivace je následující pojem.

Definice 12.1.5 (Derivace ve směru)**.** Nechť $v \in \mathbb{R}^N$, $a \in \mathbb{R}^N$ a $f\colon \mathbb{R}^N \to \mathbb{R}$ je definovaná na množině $\{a + hv\colon h \in (-\delta, \delta)\}$ pro jisté $\delta > 0$. Pak definujeme *derivaci* funkce f *ve směru* v v bodě a předpisem

$$\frac{\partial f}{\partial v}(a) := \lim_{h\to 0} \frac{f(a+hv) - f(a)}{h},$$

pokud limita na pravé straně existuje a je vlastní.

Příklad 12.1.6. (i) Nechť $f(x,y) = x^2 + 2y^2$, $v_1 = (1,1)$, $v_2 = (2,0)$ a $a = (1,1)$. Pak

$$\frac{\partial f}{\partial x}(1,1) = 2x|_{\{(x,y)=(1,1)\}} = 2 \qquad \text{a} \qquad \frac{\partial f}{\partial y}(1,1) = 4y|_{\{(x,y)=(1,1)\}} = 4.$$

Dále

$$\frac{\partial f}{\partial v_1}(1,1) = \lim_{h\to 0} \frac{(1+h)^2 + 2(1+h)^2 - 3}{h} = \lim_{h\to 0} \frac{6h + 3h^2}{h} = 6$$

a

$$\frac{\partial f}{\partial v_2}(1,1) = \lim_{h\to 0} \frac{(1+2h)^2 + 2(1+0)^2 - 3}{h} = \lim_{h\to 0} \frac{4h + 4h^2}{h} = 4.$$

(ii) Nechť

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

Díky tomu, že funkce je konstantně rovná nule na osovém kříži, okamžitě vidíme, že v počátku jsou obě parciální derivace nulové. Stejně zde vyjdou nulové směrové derivace pro směry rovnoběžné se souřadnými osami. Pokud však počítáme směrovou derivaci pro ostatní směry, nespojitost naší funkce v počátku se projeví tím, že přímo z definice odpovídající směrová derivace neexistuje (vyjdou rozdílné nevlastní jednostranné limity).

(iii) Nechť $f(x,y) = \sqrt{|xy|}\operatorname{sign} x$. Díky nulovosti na osovém kříži opět dostáváme

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

Pro $v = (1,1)$ dále máme

$$\frac{\partial f}{\partial v}(0,0) = \lim_{h\to 0} \frac{\sqrt{h^2}\operatorname{sign} h - 0}{h} = 1.$$

Poznámka 12.1.7. (i) V první části předchozího příkladu jsme viděli, že pojem směrová derivace je poněkud zavádějící, neboť hodnota této veličiny závisí nejen na směru vektoru, vůči kterému se počítá, ale také na jeho délce (porovnejte $\frac{\partial f}{\partial x}(1,1)$ a $\frac{\partial f}{\partial v_2}(1,1)$).
(ii) Funkce z první části předchozího příkladu naznačuje, že by mohl platit vzoreček

$$\frac{\partial f}{\partial v} = \sum_{i=1}^{N} \frac{\partial f}{\partial x_i} v_i, \tag{12.1.1}$$

nicméně zbylé části příkladu ukazují, že tento vzoreček platit nemusí či dokonce odpovídající směrová derivace nemusí existovat (třebaže existují obě parciální derivace).
(iii) Existence $\frac{\partial f}{\partial v}(a)$ zřejmě zaručuje, že restrikce funkce f na množinu $\{a+tv : v \in (-\delta,\delta)\}$ je spojitá v bodě a. Druhá část předchozího příkladu ukazuje, že ve spojitost v dalších směrech obecně doufat nemůžeme.

Než přistoupíme k teorii, která se bude zabývat platností vzorce (12.1.1), ukažme si ještě, že ani vztah existence derivace (zde parciální, či dokonce ve všech směrech) a spojitosti není tak jednoduchý, jako tomu bylo v jednodimenzionálním případě.

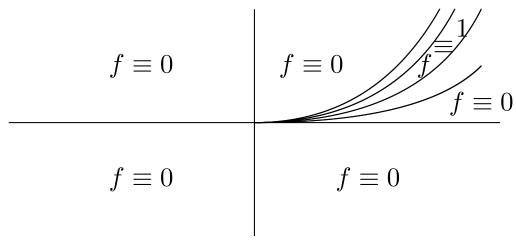
Příklad 12.1.8. Nechť $f : \mathbb{R}^2 \to \mathbb{R}^2$ je definována následovně (používáme zkrácené značení $\{x \le 0\} := \{(x,y) \in \mathbb{R}^2 : x \le 0\}$, atd., srovnejte též s Obrázkem 12.1)

$$f(x,y) = \begin{cases} 0 & \text{na } \{x \le 0\} \cup \{y \le x^2\} \cup \{y \ge 4x^2\} \\ 1 & \text{na } \{x > 0 \wedge 2x^2 \le y \le 3x^2\} \\ \text{afinní v } y & \text{na } \{x > 0 \wedge x^2 \le y \le 2x^2\} \\ \text{afinní v } y & \text{na } \{x > 0 \wedge 3x^2 \le y \le 4x^2\}. \end{cases}$$

Výsledná funkce je spojitá všude kromě počátku. Zřejmě má v počátku nulové parciální derivace (díky nulové hodnotě na osovém kříži). Dokonce jsou v počátku nulové derivace ve všech směrech. Skutečně, každý paprsek vycházející z počátku začíná částí kladné délky, která leží v množině, kde $f \equiv 0$ (a nespojitost v počátku souvisí s tím, že tyto části jsou na jednotlivých paprscích nestejně dlouhé). Příklad by šel předělat tak, aby i v ostatních bodech existovaly derivace ve všech směrech (pokud bychom „obrousili hrany“).

Definice 12.1.9 (Gradient)**.** Nechť $f : \mathbb{R}^N \to \mathbb{R}$, $a \in \mathbb{R}^N$ a existují všechny parciální derivace funkce f v bodě a. Pak *gradient* funkce f v bodě a je definován předpisem

$$\nabla f(a) := \Big(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \dots, \frac{\partial f}{\partial x_N}(a)\Big).$$



Obrázek 12.1: Ilustrace k definici funkce, která má v počátku všechny směrové derivace, ale není tam spojitá.

Analogicky pro $f\colon \mathbb{R}^N \to \mathbb{R}^m$ zavádíme gradient

$$\nabla f(a) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \cdots & \frac{\partial f_1}{\partial x_N}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \cdots & \frac{\partial f_2}{\partial x_N}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \cdots & \frac{\partial f_m}{\partial x_N}(a) \end{pmatrix},$$

existují-li jednotlivé parciální derivace. Funkce $f\colon \mathbb{R}^N \to \mathbb{R}^m$ je třídy $C^k(\Omega;\mathbb{R}^m)$ pro $k \in \mathbb{N}_0 \cup \{\infty\}$, jsou-li všechny její složky třídy $C^k(\Omega)$. Vždy se zkracuje $C^k(\Omega;\mathbb{R}^1)$ na $C^k(\Omega)$, $C^0(\Omega;\mathbb{R}^m)$ na $C(\Omega;\mathbb{R}^m)$.

Definice 12.1.10 (Totální diferenciál)**.** Nechť $f\colon \mathbb{R}^N \to \mathbb{R}$ je definovaná na nějakém okolí bodu $a \in \mathbb{R}^N$. Řekneme, že funkce f má v bodě a *totální diferenciál*, jestliže existuje taková lineární funkce $L\colon \mathbb{R}^N \to \mathbb{R}$, že

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - Lh}{\|h\|} = 0.$$

Uvedenou lineární funkci L nazýváme totálním diferenciálem funkce f v bodě a a značíme ji $\mathrm{d}f(a)$.

Poznámka 12.1.11. (i) Protože na $\mathbb{R}^N$ jsou všechny normy ekvivalentní, je jedno, kterou z nich v definici totálního diferenciálu používáme.
(ii) Pro $N = 1$ jsme si již totální diferenciál představili. Existuje právě tehdy, když existuje vlastní $f'(a)$, a totálním diferenciálem je zobrazení $h \mapsto f'(a)h$. Všechny tyto poznatky plynou ze vztahu

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

(iii) V definici totálního diferenciálu je nutné trvat na slově „lineární“ v pravém slova smyslu (nulová hodnota v počátku), nikoliv „afinní“.
(iv) Protože lineární funkce na $\mathbb{R}^N$ se dají charakterizovat jako skalární součin

argumentu s jistým $A \in \mathbb{R}^N$ (tedy $Lh = A \cdot h$ pro všechna $h \in \mathbb{R}^N$), existence totálního diferenciálu je ekvivalentní existenci $A \in \mathbb{R}^N$ splňujícího

$$\lim_{h\to 0} \frac{f(x+h) - f(x) - A \cdot h}{\|h\|} = 0.$$

V tomto případě je totální diferenciál $\mathrm{d}f(a)$ dán předpisem $\mathrm{d}f(a)(h) = A \cdot h$.
(v) Podobně jako se dá existence vlastní derivace v jednodimenzionálním případě geometricky interpretovat existencí tečny ke grafu, totální diferenciál souvisí s existencí tečné nadroviny ke grafu.
(vi) Totální diferenciál se dá zavést i pro $f\colon \mathbb{R}^N \to \mathbb{R}^m$. Výše uvedenou definici použijeme na jednotlivé složky. Totální diferenciál je v takovém případě lineární zobrazení z $\mathbb{R}^N$ do $\mathbb{R}^m$.

V dalším se nemusíme případem $f\colon \mathbb{R}^N \to \mathbb{R}^m$ příliš zabývat, následující teorie se vybuduje pro případ $m = 1$ a v případě potřeby se použije na jednotlivé složky vektorových polí.

Věta 12.1.12 (O vlastnostech plynoucích z existence totálního diferenciálu)**.** *Nechť $f\colon \mathbb{R}^N \to \mathbb{R}$ má totální diferenciál v bodě $a \in \mathbb{R}^N$. Pak*
(i) *v bodě a existují všechny parciální derivace a platí $\mathrm{d}f(a) = \nabla f(a)$ (přesněji $\mathrm{d}f(a)h = \nabla f(a) \cdot h$ pro všechna $h \in \mathbb{R}^N$)*
(ii) *existují derivace ve všech směrech a platí pro ně $\frac{\partial f}{\partial v}(a) = \nabla f(a) \cdot v$*
(iii) *funkce f je spojitá v bodě a.*

Důkaz. Nejprve dokažme části (i) a (ii). Zafixujme $v \in \mathbb{R}^N$. Pak z definice totálního diferenciálu a jeho linearity máme

$$\begin{aligned}\frac{\partial f}{\partial v}(a) &= \lim_{h\to 0} \frac{f(a+hv) - f(a)}{h} \\ &= \lim_{h\to 0} \frac{f(a+hv) - f(a) - \mathrm{d}f(a)(hv)}{\|hv\|}\frac{\|hv\|}{h} + \frac{\mathrm{d}f(a)(hv)}{h} \\ &= 0 + \mathrm{d}f(a)(v).\end{aligned}$$

Proto existuje $\frac{\partial f}{\partial v}(a)$. Volba $v = e_i$ dává v předchozím výpočtu $\frac{\partial f}{\partial x_i}(a) = \mathrm{d}f(a)(e_i)$ a z linearity $\mathrm{d}f(a)$ pak dostáváme

$$\begin{aligned}\frac{\partial f}{\partial v}(a) &= \mathrm{d}f(a)(v) = \mathrm{d}f(a)\Big(\sum_{i=1}^N v_i e_i\Big) = \sum_{i=1}^N v_i \mathrm{d}f(a)(e_i) = \sum_{i=1}^N v_i \frac{\partial f}{\partial x_i}(a) \\ &= \nabla f(a) \cdot v.\end{aligned}$$

Třetí část plyne z výpočtu

$$\lim_{h\to 0} \big(f(a+h) - f(a)\big) = \lim_{h\to 0}\Big(\frac{f(a+h) - f(a) - \mathrm{d}f(a)(h)}{\|h\|}\|h\| + \mathrm{d}f(a)(h)\Big) = 0 + 0,$$

kde jsme využili následující odhad založený na Cauchy–Schwarzově nerovnosti

$$|\mathrm{d}f(a)(h)| = |\nabla f(a) \cdot h| \leq \|\nabla f(a)\|\|h\| = C\|h\|.$$

$\square$

Příklad 12.1.13. (i) Zkoumejme existenci totálního diferenciálu v počátku pro funkci $f(x,y) = x^2 + y^2$. Pokud totální diferenciál existuje, podle předchozí věty úzce souvisí s gradientem. Spočítejme proto

$$\frac{\partial f}{\partial x}(0,0) = 0 \qquad \text{a} \qquad \frac{\partial f}{\partial y}(0,0) = 0.$$

Jediným kandidátem na totální diferenciál je proto zobrazení $L\colon (h_1,h_2) \mapsto 0 \cdot h_1 + 0 \cdot h_2 = 0$. Máme pro něj

$$\begin{aligned}\lim_{h\to 0} \frac{f(0+h_1, 0+h_2) - f(0,0) - Lh}{\|h\|} &= \lim_{h\to 0} \frac{h_1^2 + h_2^2 - 0 - 0}{\|h\|} \\ &= \lim_{h\to 0} \frac{h_1^2 + h_2^2}{\sqrt{h_1^2 + h_2^2}} = 0.\end{aligned}$$

Tím jsme ověřili, že se skutečně jedná o totální diferenciál. Tedy $\mathrm{d}f(0,0) \equiv 0$.
(ii) Zkoumejme existenci totálního diferenciálu v počátku pro funkci $f(x,y) = \sqrt[3]{xy}$. Přímo z definice spočítáme

$$\frac{\partial f}{\partial x}(0,0) = 0 \qquad \text{a} \qquad \frac{\partial f}{\partial y}(0,0) = 0.$$

Pro jediného kandidáta $L \equiv 0$ na totální diferenciál pak máme

$$\frac{f(0+h_1, 0+h_2) - f(0,0) - Lh}{\|h\|} = \frac{\sqrt[3]{h_1 h_2}}{\|h\|}.$$

Limita posledního výrazu pro $h \to 0$ však nemůže být nulová, což snadno nahlédneme v situaci $h_1 = h_2$. Totální diferenciál proto neexistuje.

Druhá část předchozího příkladu nám právě ukázala, že v poslední větě se implikace nedají obrátit. Postačující podmínka pro existenci totálního diferenciálu proto vyžaduje více než jen existenci parciálních derivací a spojitost.

Věta 12.1.14 (O postačující podmínce pro existenci totálního diferenciálu)**.** *Nechť $f\colon \mathbb{R}^N \to \mathbb{R}$ a $a \in \mathbb{R}^N$.*
(i) *Má-li f na jistém okolí bodu a omezené parciální derivace, pak je v něm spojitá.*
(ii) *Má-li f na jistém okolí bodu a parciální derivace a ty jsou spojité v bodě a, pak v něm má totální diferenciál.*

Důkaz. Z důvodu zjednodušení zápisu důkaz provedeme jen pro $N = 2$. Ve vyšší dimenzi se postupuje analogicky.

Dokažme nejprve část (i). Nechť $G := (a_1 - \delta, a_1 + \delta) \times (a_2 - \delta, a_2 + \delta)$ je podmnožinou okolí uvedeného ve znění věty a $K > 0$ je konstanta, která na něm omezuje všechny parciální derivace. Pokud $y \in G$ a $y_1 \neq a_1$ a $y_2 \neq a_2$, pak podle Lagrangeovy věty o přírůstku funkce (Věta 6.3.3) máme

$$\begin{aligned}|f(y_1,y_2) - f(a_1,a_2)| &= |f(y_1,y_2) - f(a_1,y_2) + f(a_1,y_2) - f(a_1,a_2)| \\ &= \Big|\frac{\partial f}{\partial x_1}(\xi_1, y_2)(y_1 - a_1) + \frac{\partial f}{\partial x_2}(a_1, \xi_2)(y_2 - a_2)\Big| \\ &\leq K|y_1 - a_1| + K|y_2 - a_2| \leq 2K\|y - a\|\end{aligned}$$

(rozmyslete si, proč jsme mohli Lagrangeovu větu použít). V případě, kdy y a a se liší jen v jedné souřadnici, snadno dostaneme $|f(y_1, y_2) - f(a_1, a_2)| \leq K\|y - a\|$. Z našich odhadů plyne spojitost f v bodě a.

Dokažme část (ii). Zvolme $\varepsilon > 0$. Pak existuje $\delta > 0$ takové, že pro $y \in G := (a_1 - \delta, a_1 + \delta) \times (a_2 - \delta, a_2 + \delta)$ platí

$$\Big|\frac{\partial f}{\partial x_1}(y) - \frac{\partial f}{\partial x_1}(a)\Big| < \varepsilon \qquad \text{a} \qquad \Big|\frac{\partial f}{\partial x_2}(y) - \frac{\partial f}{\partial x_2}(a)\Big| < \varepsilon.$$

Pro libovolné $y \in G$ takové, že $y_1 \neq a_1$ a $y_2 \neq a_2$ pak podle Lagrangeovy věty máme

$$\begin{aligned}
&\Big|f(y_1, y_2) - f(a_1, a_2) - \Big(\frac{\partial f}{\partial x_1}(a)(y_1 - a_1) + \frac{\partial f}{\partial x_2}(a)(y_2 - a_2)\Big)\Big| \\
&= \Big|f(y_1, y_2) - f(a_1, y_2) + f(a_1, y_2) - f(a_1, a_2) \\
&\qquad - \Big(\frac{\partial f}{\partial x_1}(a)(y_1 - a_1) + \frac{\partial f}{\partial x_2}(a)(y_2 - a_2)\Big)\Big| \\
&= \Big|\frac{\partial f}{\partial x_1}(\xi_1, y_2)(y_1 - a_1) + \frac{\partial f}{\partial x_2}(a_1, \xi_2)(y_2 - a_2) \\
&\qquad - \frac{\partial f}{\partial x_1}(a)(y_1 - a_1) + \frac{\partial f}{\partial x_2}(a)(y_2 - a_2)\Big| \\
&= \Big|\Big(\frac{\partial f}{\partial x_1}(\xi_1, y_2) - \frac{\partial f}{\partial x_1}(a_1, a_2)\Big)(y_1 - a_1) \\
&\qquad + \Big(\frac{\partial f}{\partial x_2}(a_1, \xi_2) - \frac{\partial f}{\partial x_2}(a_1, a_2)\Big)(y_2 - a_2)\Big| \\
&\leq \varepsilon|y_1 - a_1| + \varepsilon|y_2 - a_2| \leq 2\varepsilon\|y - a\|.
\end{aligned}$$

V případě, kdy y a a se liší jen v jedné souřadnici, stejný odhad získáme ještě jednodušeji. Protože ε bylo libovolné, ověřili jsme definici totálního diferenciálu funkce f v bodě a. $\square$

Příklad 12.1.15. (i) Pro funkci $f(x, y) = x^2 + y^2$ platí

$$\frac{\partial f}{\partial x} = 2x \qquad \text{a} \qquad \frac{\partial f}{\partial y} = 2y.$$

Protože funkce $(x, y) \mapsto 2x$ a $(x, y) \mapsto 2y$ jsou spojité na $\mathbb{R}^2$, funkce f má totální diferenciál na $\mathbb{R}^2$.

(ii) Pro funkci $f(x, y) = \sqrt[3]{xy}$ platí (používáme Příklad 12.1.13)

$$\frac{\partial f}{\partial x}(a) \begin{cases} = \frac{\sqrt[3]{a_2}}{3\sqrt[3]{a_1^2}} & \text{pro } a_1 \neq 0 \\ = 0 & \text{pro } a_1 = 0 \text{ a } a_2 = 0 \\ \text{neexistuje} & \text{pro } a_1 = 0 \text{ a } a_2 \neq 0 \end{cases}$$

a

$$\frac{\partial f}{\partial y}(a) \begin{cases} = \frac{\sqrt[3]{a_1}}{3\sqrt[3]{a_2^2}} & \text{pro } a_2 \neq 0 \\ = 0 & \text{pro } a_1 = 0 \text{ a } a_2 = 0 \\ \text{neexistuje} & \text{pro } a_1 \neq 0 \text{ a } a_2 = 0. \end{cases}$$

Mimo osový kříž jsou parciální derivace spojité, a proto zde existuje totální diferenciál. Na osovém kříži vyjma počátku vždy jedna parciální derivace neexistuje, a proto zde neexistuje totální diferenciál. Zbývá vyšetřit existenci totálního diferenciálu v počátku, ale to už jsme učinili v Příkladu 12.1.13.

Protože totální diferenciál je reprezentován gradientem, pro který platí aritmetika derivace, čtenáře jistě nepřekvapí, že podobná aritmetika platí i pro totální diferenciál (úplně zřejmé to ale není, musíme ověřit požadovanou aproximační vlastnost).

Tvrzení 12.1.16 (Aritmetika totálního diferenciálu)**.** *Nechť* $f, g\colon \mathbb{R}^N \to \mathbb{R}$, $a \in \mathbb{R}^N$ *a existují totální diferenciály* $\mathrm{d}f(a)$ *a* $\mathrm{d}g(a)$. *Pak*
(i) $\mathrm{d}(f+g)(a) = \mathrm{d}f(a) + \mathrm{d}g(a)$
(ii) $\mathrm{d}(fg)(a) = g(a)\mathrm{d}f(a) + f(a)\mathrm{d}g(a)$
(iii) *pokud navíc* $g(a) \neq 0$, *pak* $\mathrm{d}\frac{f}{g}(a) = \frac{g(a)\mathrm{d}f(a) - f(a)\mathrm{d}g(a)}{g^2(a)}$.

Důkaz. Důkaz první části tvrzení je triviální. Druhá část plyne z identity

$$\begin{aligned}
&f(a+h)g(a+h) - f(a)g(a) - g(a)\mathrm{d}f(a)h - f(a)\mathrm{d}g(a)h \\
&= f(a+h)g(a+h) - f(a+h)g(a) + f(a+h)g(a) - f(a)g(a) \\
&\qquad - g(a)\mathrm{d}f(a)h - f(a)\mathrm{d}g(a)h \\
&= f(a+h)(g(a+h) - g(a)) - f(a)\mathrm{d}g(a)h + g(a)(f(a+h) - f(a) - \mathrm{d}f(a)h) \\
&= f(a+h)(g(a+h) - g(a) - \mathrm{d}g(a)h) + (f(a) - f(a+h))\mathrm{d}g(a)h \\
&\qquad + g(a)(f(a+h) - f(a) - \mathrm{d}f(a)h)
\end{aligned}$$

(u druhého členu na předposledním řádku připomeňme, že existence totálního diferenciálu implikuje spojitost, dále činitel $\frac{\mathrm{d}g(a)h}{\|h\|}$ je omezený díky linearitě totálního diferenciálu).

Důkaz třetí části, díky již dokázané druhé části, plyne z rovností

$$\frac{1}{g(a+h)} - \frac{1}{g(a)} + \frac{\mathrm{d}g(a)h}{g^2(a)} = \frac{g^2(a) - g(a)g(a+h) + g(a+h)\mathrm{d}g(a)h}{g^2(a)g(a+h)}$$

a

$$\begin{aligned}
&g^2(a) - g(a)g(a+h) + g(a+h)\mathrm{d}g(a)h \\
&= g(a)(g(a) - g(a+h) + \mathrm{d}g(a)h) + (g(a+h) - g(a))\mathrm{d}g(a)h.
\end{aligned}$$

□

Připomeňme, že celá dosavadní teorie se dala dělat pro zobrazení z $\mathbb{R}^N$ do $\mathbb{R}^m$. Postupovali bychom po jednotlivých složkách. Tato metoda se však nedá použít na výsledky o složených zobrazeních.

Věta 12.1.17 (O totálním diferenciálu složeného zobrazení)**.** *Nechť* $f\colon \mathbb{R}^N \to \mathbb{R}^m$ *má totální diferenciál v bodě* $a \in \mathbb{R}^N$ *a* $g\colon \mathbb{R}^m \to \mathbb{R}^k$ *má totální diferenciál v bodě* $f(a)$. *Pak funkce* $g \circ f$ *má totální diferenciál v bodě* a *a platí pro něj*

$$\mathrm{d}(g \circ f)(a) = \mathrm{d}g(f(a)) \circ \mathrm{d}f(a)$$

(neboli $\mathrm{d}(g \circ f)(a)(h) = \mathrm{d}g(f(a))(\mathrm{d}f(a)(h))$ *pro všechna* $h \in \mathbb{R}^N$*).*

Důkaz. Pro jednoduchost značení pišme $b = f(a)$. Díky existenci dílčích totálních diferenciálů existují funkce $\varphi\colon \mathbb{R}^N \to \mathbb{R}^m$ a $\psi\colon \mathbb{R}^m \to \mathbb{R}^k$ splňující $\varphi(0) = 0$, $\psi(0) = 0$ (pozor, pokaždé se jedná o počátek v jiném prostoru),

$$\begin{aligned} f(a+h) &= f(a) + \mathrm{d}f(a)(h) + \varphi(h) && \text{pro všechna } h \in \mathbb{R}^N \\ g(b+l) &= g(b) + \mathrm{d}g(b)(l) + \psi(l) && \text{pro všechna } l \in \mathbb{R}^m \end{aligned}$$

a

$$\lim_{h\to 0} \frac{\|\varphi(h)\|_{\mathbb{R}^m}}{\|h\|_{\mathbb{R}^N}} = 0 \qquad \text{a} \qquad \lim_{l\to 0} \frac{\|\psi(l)\|_{\mathbb{R}^k}}{\|l\|_{\mathbb{R}^m}} = 0. \tag{12.1.2}$$

Proto máme

$$\begin{aligned} g(f(a+h)) &= g(b + \mathrm{d}f(a)(h) + \varphi(h)) \\ &= g(b) + \mathrm{d}g(b)(\mathrm{d}f(a)(h) + \varphi(h)) + \psi(\mathrm{d}f(a)(h) + \varphi(h)) \\ &= g(b) + \mathrm{d}g(b)(\mathrm{d}f(a)(h)) + \mathrm{d}g(b)(\varphi(h)) + \psi(\mathrm{d}f(a)(h) + \varphi(h)). \end{aligned}$$

Díky linearitě zobrazení $\mathrm{d}g(b)$ a první části (12.1.2) dostáváme

$$0 \le \lim_{h\to 0} \frac{\|\mathrm{d}g(b)(\varphi(h))\|_{\mathbb{R}^k}}{\|h\|_{\mathbb{R}^N}} \le C \lim_{h\to 0} \frac{\|\varphi(h)\|_{\mathbb{R}^m}}{\|h\|_{\mathbb{R}^N}} = 0.$$

Konečně, pokud $\mathrm{d}f(a)(h) + \varphi(h) \neq 0$ (v opačném případě není co dokazovat), pak

$$\frac{\|\psi(\mathrm{d}f(a)(h) + \varphi(h))\|_{\mathbb{R}^k}}{\|h\|_{\mathbb{R}^N}} = \frac{\|\psi(\mathrm{d}f(a)(h) + \varphi(h))\|_{\mathbb{R}^k}}{\|\mathrm{d}f(a)(h) + \varphi(h)\|_{\mathbb{R}^m}} \frac{\|\mathrm{d}f(a)(h) + \varphi(h)\|_{\mathbb{R}^m}}{\|h\|_{\mathbb{R}^N}}.$$

Snadno se nahlédne, že druhý činitel na pravé straně poslední rovnosti je omezený a první činitel konverguje do nuly pro $\|h\|_{\mathbb{R}^N} \to 0$. Celkově jsme ukázali

$$\lim_{\|h\|_{\mathbb{R}^N}\to 0} \frac{\|g(f(a+h)) - g(f(a)) - \mathrm{d}g(b)(\mathrm{d}f(a)(h))\|_{\mathbb{R}^k}}{\|h\|_{\mathbb{R}^N}} = 0.$$

□

Protože totální diferenciál je reprezentován gradientem a skládání lineárních zobrazení odpovídá součinu reprezentujících matic, z předchozí věty okamžitě dostáváme následující výsledek.

Věta 12.1.18 (Řetízkové pravidlo)**.** *Nechť $f\colon \mathbb{R}^N \to \mathbb{R}^m$ má totální diferenciál v bodě $a \in \mathbb{R}^N$ a $g\colon \mathbb{R}^m \to \mathbb{R}$ má totální diferenciál v bodě $f(a)$. Pak pro každé $i \in \{1, \ldots, N\}$ platí*

$$\frac{\partial (g\circ f)}{\partial x_i}(a) = \sum_{j=1}^{m} \frac{\partial g}{\partial y_j}(f(a)) \frac{\partial f_j}{\partial x_i}(a).$$

Příklad 12.1.19. (i) Nechť $F(x_1, x_2) = \operatorname{arctan} \sqrt{1 + x_1^2 + x_2^2}$. Pak

$$\frac{\partial F}{\partial x_1} = \frac{1}{1 + 1 + x_1^2 + x_2^2} \frac{\partial \sqrt{1 + x_1^2 + x_2^2}}{\partial x_1} = \frac{1}{2 + x_1^2 + x_2^2} \frac{x_1}{\sqrt{1 + x_1^2 + x_2^2}} \qquad \text{na } \mathbb{R}^2.$$

(ii) Nechť $f(x_1,x_2)=(x_1+x_2^2,x_1x_2,x_1)$ a $g(y_1,y_2,y_3)=y_1y_2y_3$. Pak na $\mathbb{R}^2$ máme

$$\begin{aligned}\frac{\partial(g\circ f)}{\partial x_1}(a)&=\frac{\partial g}{\partial y_1}(f(a))\frac{\partial f_1}{\partial x_1}(a)+\frac{\partial g}{\partial y_2}(f(a))\frac{\partial f_2}{\partial x_1}(a)+\frac{\partial g}{\partial y_3}(f(a))\frac{\partial f_3}{\partial x_1}(a)\\&=a_1^2a_2\cdot 1+(a_1^2+a_1a_2^2)\cdot a_2+(a_1^2a_2+a_1a_2^3)\cdot 1.\end{aligned}$$

(iii) Převedeme si gradient do polárních souřadnic. Nechť $u\in C^1(\mathbb{R}^2)$ pracuje s proměnnými x,y a máme $x=r\cos\varphi$, $y=r\sin\varphi$ pro $r>0$ a $\varphi\in(0,2\pi)$. Spočítáme Δu pomocí $\tilde{u}(r,\varphi)=u(x(r,\varphi),y(r,\varphi))$. Platí

$$\begin{aligned}\frac{\partial\tilde{u}}{\partial r}&=\frac{\partial u}{\partial x}\frac{\partial x}{\partial r}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial r}=\frac{\partial u}{\partial x}\cos\varphi+\frac{\partial u}{\partial y}\sin\varphi\\\frac{\partial\tilde{u}}{\partial\varphi}&=\frac{\partial u}{\partial x}\frac{\partial x}{\partial\varphi}+\frac{\partial u}{\partial y}\frac{\partial y}{\partial\varphi}=\frac{\partial u}{\partial x}(-r\sin\varphi)+\frac{\partial u}{\partial y}r\cos\varphi.\end{aligned}$$

Vynásobíme-li první vztah výrazem $\cos\varphi$ a druhý výrazem $-\frac{\sin\varphi}{r}$, po sečtení získáme vyjádření $\frac{\partial u}{\partial x}$. Podobný trik použijeme pro získání $\frac{\partial u}{\partial y}$. Dostáváme

$$\begin{aligned}\frac{\partial u}{\partial x}&=\frac{\partial\tilde{u}}{\partial r}\cos\varphi-\frac{1}{r}\frac{\partial\tilde{u}}{\partial\varphi}\sin\varphi\\\frac{\partial u}{\partial y}&=\frac{\partial\tilde{u}}{\partial r}\sin\varphi+\frac{1}{r}\frac{\partial\tilde{u}}{\partial\varphi}\cos\varphi.\end{aligned}$$

Dalším klasickým výsledkem, který se dá rozšířit do vyšší dimenze, je Lagrangeova věta o přírůstku, někdy též Věta o střední hodnotě (Věta 6.3.3.) V jednodimenzionální sitiuaci jsme pracovali na intervalu. Zde se jako přirozený ukáže následující typ množiny.

Definice 12.1.20 (Konvexní množina)**.** Řekneme, že množina $A\subset\mathbb{R}^N$ je *konvexní*, jestliže pro každá $x,y\in A$ a $\lambda\in[0,1]$ platí $\lambda x+(1-\lambda)y\in A$.

Poznámka 12.1.21. (i) Platí $\lambda x+(1-\lambda)y=y+\lambda(x-y)$.
(ii) Jsou-li x,y pevně zvolené a $x\neq y$, pak možina $\{\lambda x+(1-\lambda)y\}$ je úsečka spojující body x a y. Prvky z uvedené množiny nazýváme *konvexní kombinace* prvků x a y.
(iii) Rozmyslete si, že konvexita je také ekvivalentní podmínce

$$k\in\mathbb{N}\wedge x_1,\dots,x_k\in A\wedge\lambda_1,\dots,\lambda_k\in[0,1]\wedge\sum_{j=1}^{k}\lambda_j=1\quad\Longrightarrow\quad\sum_{j=1}^{k}\lambda_kx_k\in A.$$

(iv) Konvexita funkce $f\colon\mathbb{R}\to\mathbb{R}$ je ekvivalentní tomu, že je konvexní její nadgraf, tedy množina všech $(x,y)\in\mathbb{R}^2$ takových, že $y>f(x)$.

Věta 12.1.22 (O střední hodnotě)**.** *Nechť $A\subset\mathbb{R}^N$ je otevřená konvexní množina a $f\colon\mathbb{R}^N\to\mathbb{R}$ má totální diferenciál na A. Pak pro všechna $a,b\in A$, $a\neq b$, existuje $\theta\in(0,1)$ takové, že*

$$\begin{aligned}f(b)-f(a)&=\mathrm{d}f(a+\theta(b-a))(b-a)=\nabla f(a+\theta(b-a))\cdot(b-a)\\&=\sum_{j=1}^{N}\frac{\partial f}{\partial x_j}(a+\theta(b-a))(b_j-a_j).\end{aligned}$$

Důkaz. Definujme pomocnou funkci

$$F(t) = f(a + t(b - a)).$$

Pak $F = f \circ g$, kde $g\colon t \mapsto a + t(b - a)$ je nekonečněkrát spojitě diferencovatelná na $[0, 1]$. Potom díky Větě o totálním diferenciálu složeného zobrazení (Věta 12.1.17) má funkce F totální diferenciál (tedy vlastní derivaci) na $[0, 1]$ a platí

$$F'(t) = \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}(g(t)) \frac{\partial g_j}{\partial t}(t) = \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}(a + t(b - a))(b_j - a_j).$$

Podle Lagrangeovy věty o střední hodnotě pak existuje $\theta \in (0, 1)$ takové, že

$$f(b) - f(a) = F(1) - F(0) = F'(\theta) = \sum_{j=1}^{N} \frac{\partial f}{\partial x_j}(a + \theta(b - a))(b_j - a_j).$$

□

12.2 Derivace a totální diferenciály vyšších řádů, Taylorův vzorec

Jak uvidíme později, vyšší (zejména druhé) parciální derivace mají velký význam při klasifikaci lokálních extrémů funkcí více proměnných. V předchozím textu si na příkladech čtenář mohl povšimnout, že často platí $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$. Tomuto jevu se říká *záměnnost parciálních derivací* a při výpočtech nám může ušetřit mnoho času. Naneštěstí neplatí obecně.

Příklad 12.2.1. Definujme

$$f(x, y) = \begin{cases} xy & \text{pro } |x| < |y| \\ 0 & \text{jinak.} \end{cases}$$

Pak snadno nahlédneme, že

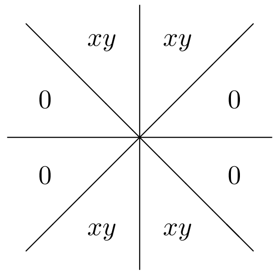
$$\frac{\partial f}{\partial x}(0, 0) = 0 = \frac{\partial f}{\partial y}(0, 0), \qquad \frac{\partial f}{\partial x}(0, y) = y \qquad \text{a} \qquad \frac{\partial f}{\partial y}(x, 0) = 0.$$

Odtud

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = 1 \neq 0 = \frac{\partial^2 f}{\partial x \partial y}(0, 0).$$

Dostatečně hladké funkce však záměnnost parciálních derivací splňují.

Věta 12.2.2 (O záměnnosti parciálních derivací)**.** *Nechť $\Omega \subset \mathbb{R}^2$ je otevřená množina a $f \in C^2(\Omega)$. Pak má f na Ω záměnné druhé parciální derivace.*



Obrázek 12.2: Náčrt k definici funkce porušující záměnnost parciálních derivací v počátku.

Důkaz. Bez újmy na obecnosti předpokládejme, že pracujeme v počátku. Existuje $\delta > 0$ takové, že $(-\delta, \delta) \times (-\delta, \delta) \subset \Omega$. Pro $h, k \in (0, \delta)$ definujme funkce

$$\begin{aligned}
Q(h,k) &= \frac{1}{hk}\Big(f(h,k) - f(h,0) - f(0,k) + f(0,0)\Big)\\
\varphi_k(h) &= \frac{f(h,k) - f(h,0)}{k}\\
\psi_h(k) &= \frac{f(h,k) - f(0,k)}{h}.
\end{aligned}$$

Zafixujme nyní $h, k \in (0, \delta)$. Pak podle Lagrangeovy věty existují $\xi_1 \in (0, h)$ a $\xi_2 \in (0, k)$ taková, že platí

$$\begin{aligned}
Q(h,k) &= \frac{1}{h}\Big(\frac{f(h,k) - f(h,0)}{k} - \frac{f(0,k) - f(0,0)}{k}\Big) = \frac{1}{h}\Big(\varphi_k(h) - \varphi_k(0)\Big)\\
&= \varphi_k'(\xi_1) = \frac{1}{k}\Big(\frac{\partial f}{\partial x}(\xi_1, k) - \frac{\partial f}{\partial x}(\xi_1, 0)\Big) = \frac{\partial^2 f}{\partial y \partial x}(\xi_1, \xi_2).
\end{aligned}$$

Podobně dostaneme

$$\begin{aligned}
Q(h,k) &= \frac{1}{k}\Big(\frac{f(h,k) - f(0,k)}{h} - \frac{f(h,0) - f(0,0)}{h}\Big) = \frac{1}{k}\Big(\psi_h(k) - \psi_h(0)\Big)\\
&= \psi_h'(\xi_3) = \frac{1}{h}\Big(\frac{\partial f}{\partial y}(h, \xi_3) - \frac{\partial f}{\partial y}(0, \xi_3)\Big) = \frac{\partial^2 f}{\partial x \partial y}(\xi_4, \xi_3).
\end{aligned}$$

Ze spojitosti druhých parciálních derivací v počátku nyní plyne, že jsou-li h, k dostatečně malé, máme

$$\Big|\frac{\partial^2 f}{\partial y \partial x}(\xi_1, \xi_2) - \frac{\partial^2 f}{\partial y \partial x}(0,0)\Big| < \varepsilon, \qquad \Big|\frac{\partial f}{\partial x \partial y}(\xi_4, \xi_3) - \frac{\partial f}{\partial x \partial y}(0,0)\Big| < \varepsilon.$$

Protože jsme výše ukázali $\frac{\partial^2 f}{\partial x \partial y}(\xi_1, \xi_2) = \frac{\partial^2 f}{\partial y \partial x}(\xi_4, \xi_3)$, nyní již snadno obdržíme dokazovaný výsledek. $\square$

Důsledek 12.2.3. *Nechť $\Omega \subset \mathbb{R}^N$ a $f \in C^k(\Omega)$ pro $k \geq 2$. Pak jsou všechny parciální derivace k-tého řádu záměnné.*

Důkaz. Předchozí věta dává záměnnost druhých parciálních derivací v $\mathbb{R}^N$ (skutečně, při výpočtu $\frac{\partial^2 f}{\partial x_i \partial x_j}$ jsou zafixované všechny proměnné až na x_i a x_j). Dále podle předchozího výsledku při postupném parciálním derivování umíme prohodit pořadí dvou po sobě následujících parciálních derivací. $\square$

Věta 12.2.4 (Taylorův vzorec)**.** *Nechť $a \in \mathbb{R}^N$, $\delta > 0$, $m \in \mathbb{N}$, $f \in C^{m+1}(\mathcal{U}_\delta(a))$ a $a + h \in \mathcal{U}_\delta(a)$. Pak existuje $\theta \in (0,1)$ takové, že*

$$
\begin{aligned}
f(a+h) = f(a) &+ \sum_{i_1=1}^{N} \frac{\partial f}{\partial x_{i_1}}(a) h_{i_1} + \frac{1}{2!} \sum_{i_1,i_2=1}^{N} \frac{\partial^2 f}{\partial x_{i_1} \partial x_{i_2}}(a) h_{i_1} h_{i_2} \\
&+ \cdots + \frac{1}{m!} \sum_{i_1,\ldots,i_m=1}^{N} \frac{\partial^m f}{\partial x_{i_1} \ldots \partial x_{i_m}}(a) h_{i_1} \ldots h_{i_m} \\
&+ \frac{1}{(m+1)!} \sum_{i_1,\ldots,i_{m+1}=1}^{N} \frac{\partial^{m+1} f}{\partial x_{i_1} \ldots \partial x_{i_{m+1}}}(a + \theta h) h_{i_1} \ldots h_{i_{m+1}}.
\end{aligned}
$$

Důkaz. Pro $t \in [0,1]$ definujme $\varphi(t) = f(a + th)$. Pak už výsledek plyne z řetízkového pravidla (Věta 12.1.18) a (jednodimenzionální) Věty o odhadu chyby Taylorova polynomu (Věta 6.8.10). $\square$

Definice 12.2.5 (Totální diferenciál řádu k)**.** Nechť $f \in C^k(\mathcal{U}_\delta(a))$. *Totálním diferenciálem* řádu k příslušejícím funkci f v bodě a nazýváme k-lineární funkci (zobrazuje $\mathbb{R}^{Nk}$ do $\mathbb{R}$)

$$
\mathrm{d}^k f(a)(h^1, \ldots, h^k) = \sum_{i_1,\ldots,i_k=1}^{N} \frac{\partial^k f}{\partial x_{i_1} \ldots \partial x_{i_k}}(a) h^1_{i_1} \ldots h^k_{i_k}.
$$

Definice 12.2.6 (Multiindex)**.** Nechť $N \in \mathbb{N}$ je pevné. *Multiindexem* nazýváme N-tici nezáporných celých čísel $\alpha = (\alpha_1, \ldots, \alpha_N)$. Číslo

$$
|\alpha| := \alpha_1 + \cdots + \alpha_N
$$

se nazývá *výška multiindexu.* Pro multiindex α, $f \in C^{|\alpha|}(\Omega)$ a $x \in \mathbb{R}^N$ zavádíme značení

$$
D^\alpha f := \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots \partial x_N^{\alpha_N}}, \qquad x^\alpha := x_1^{\alpha_1} \ldots x_N^{\alpha_N} \qquad \text{a} \qquad \binom{|\alpha|}{\alpha} = \frac{|\alpha|!}{\alpha_1! \ldots \alpha_N!}.
$$

Poznámka 12.2.7. Taylorův rozvoj lze při našem novém značení psát ve tvaru

$$\begin{aligned} f(a+h) = f(a) &+ \sum_{|\alpha|=1} \binom{1}{\alpha} D^\alpha f(a) h^\alpha + \frac{1}{2!} \sum_{|\alpha|=2} \binom{2}{\alpha} D^\alpha f(a) h^\alpha + \dots \\ &+ \frac{1}{m!} \sum_{|\alpha|=m} \binom{m}{\alpha} D^\alpha f(a) h^\alpha \\ &+ \frac{1}{(m+1)!} \sum_{|\alpha|=m+1} \binom{m+1}{\alpha} D^\alpha f(a+\theta h) h^\alpha. \end{aligned}$$

Jinou možností je zápis pomocí diferenciálů vyšších řádů

$$f(a+h) = f(a) + \sum_{i=1}^{k} \frac{1}{i!} \mathrm{d}^i f(a)(h,\dots,h) + \frac{1}{(k+1)!} \mathrm{d}^{k+1} f(a+\theta h)(h,\dots,h).$$

Cvičení 12.2.8. Dokažte si zobecnění binomické věty

$$(h_1 + h_2 + \dots + h_N)^n = \sum_{|\alpha|=n} \binom{n}{\alpha} h^\alpha \qquad \text{se značením } h := (h_1,\dots,h_N).$$

12.3 Potenciál vektorového pole

V dalším si představíme analogii pojmu primitivní funkce pro funkce více proměnných. Nový pojem budeme dále studovat a ukážeme si, jak se používají jeho vlastnosti při řešení diferenciálních rovnic.

Definice 12.3.1 (Potenciál vektorového pole)**.** Vektorové pole $T = (T_1,\dots,T_N)\colon \mathbb{R}^N \to \mathbb{R}^N$ nazýváme *potenciální* na otevřené množině $\Omega \subset \mathbb{R}^N$, jestliže existuje funkce $U\colon \mathbb{R}^N \to \mathbb{R}$ taková, že

$$\frac{\partial U}{\partial x_i} = T_i \qquad \text{na } \Omega \text{ pro všechna } i \in \{1,\dots,N\}.$$

Funkci U pak nazýváme *potenciálem* vektorového pole T.

Je zřejmé, že přičteme-li k potenciálu vektorového pole T libovolnou aditivní konstantu, výsledná funkce je také potenciálem vektorového pole T. Na druhou stranu pokud vezmeme $\Omega = \Omega_1 \cup \Omega_2$, kde $\Omega_1 = (0,1) \times (0,1)$ a $\Omega_2 = (2,3) \times (0,1)$ a $T \equiv (0,\dots,0)$, pak jakákoliv funkce tvaru

$$U = \begin{cases} C_1 & \text{na } \Omega_1 \\ C_2 & \text{na } \Omega_2, \end{cases}$$

kde $C_1, C_2 \in \mathbb{R}$, je potenciálem vektorového pole T. Není tedy obecně pravda, že by potenciál byl určen jednoznačně až na (jednu) aditivní konstantu, jak jsme tomu byli zvyklí u primitivní funkce (která je jednorozměrným případem potenciálu). Tento typ jednoznačnosti nám nabídne vhodně zvolený typ množin.

Definice 12.3.2 (Souvislá množina)**.** Množina $\Omega \subset \mathbb{R}^N$ se nazývá *souvislá*, jestliže každé její dva body lze spojit lomenou čarou tvořenou konečným počtem úseček, které celé leží v Ω. Množina $\Omega \subset \mathbb{R}^N$ se nazývá *oblast*, je-li otevřená a souvislá. Oblast $\Omega \subset \mathbb{R}^N$ se nazývá *jednoduše souvislá*, je-li souvislý také její doplněk.

Věta 12.3.3 (O nejednoznačnosti potenciálu na oblasti)**.** *Nechť vektorové pole T má na oblasti $\Omega \subset \mathbb{R}^N$ potenciály U_1 a U_2. Pak existuje $C \in \mathbb{R}$ takové, že $U_2 = U_1 + C$.*

Důkaz. Stačí ukázat, že funkce s nulovými parciálními derivacemi na Ω musí být konstantní. Zvolíme-li libovolné dva body v Ω a aplikujeme-li Větu o střední hodnotě (Věta 12.1.22) na jednotlivých segmentech jim odpovídající lomené čáry z definice souvislé množiny, okamžitě dostáváme, že tyto body mají stejnou funkční hodnotu. □

Poznámka 12.3.4. Připomeňme, že v teorii primitivních funkcí jsme pracovali jen na otevřených intervalech v $\mathbb{R}$. Jednalo se tedy o oblasti a dosažené výsledky stran nejednoznačnosti primitivní funkce odpovídají našim výsledkům o nejednoznačnosti potenciálu.

Podobně jako nemusí existovat primitivní funkce, nemusí existovat ani potenciál. Skutečně, protože je pojem parciální derivace odvozen od derivace klasické, snadno si rozmyslíme, že pro existenci potenciálu je nutné, aby funkce

$$x_1 \mapsto f(x_1, \dots, x_N), \quad x_2 \mapsto f(x_1, \dots, x_N), \quad \dots, \quad x_N \mapsto f(x_1, \dots, x_N)$$

měly Darbouxovu vlastnost na intervalech odpovídajících „propíchnutí“ množiny Ω v předpisech uvedených funkcí. Máme však ještě jednu nutnou podmínku, která požaduje jistou provázanost jednotlivých složek vektorového pole T.

Věta 12.3.5 (Nutná podmínka existence potenciálu)**.** *Nechť $\Omega \subset \mathbb{R}^N$ je otevřená množina a $T \in C^1(\Omega; \mathbb{R}^N)$. Má-li T potenciál na Ω, pak platí*

$$\frac{\partial T_i}{\partial x_j} = \frac{\partial T_j}{\partial x_i} \qquad \text{na } \Omega \qquad \text{pro všechna } i, j \in \{1, \dots, N\}.$$

Důkaz. Pokud existuje potenciál, podle předpokladu má spojité všechny parciální derivace druhého řádu. Ty jsou proto záměnné a máme

$$\frac{\partial T_i}{\partial x_j} = \frac{\partial^2 U}{\partial x_j \partial x_i} = \frac{\partial^2 U}{\partial x_i \partial x_j} = \frac{\partial T_j}{\partial x_i}$$

na Ω pro všechna $i, j \in \{1, \dots, N\}$. □

Nyní si vyslovíme a dokážeme postačující podmínku pro existenci potenciálu.

Věta 12.3.6 (Postačující podmínka existence potenciálu)**.** *Nechť $\Omega \subset \mathbb{R}^N$ je otevřený interval (tedy kvádr v případě omezenosti), $T \in C^1(\Omega; \mathbb{R}^N)$ a*

$$\frac{\partial T_i}{\partial x_j} = \frac{\partial T_j}{\partial x_i} \qquad \text{na } \Omega \qquad \text{pro všechna } i, j \in \{1, \dots, N\}.$$

Pak má vektorové pole T potenciál na Ω.

Důkaz. Podrobný důkaz provedeme jen pro případ $N = 2$. Jeho rozšíření do obecné dimenze okomentujeme na konci.

Krok 1: konstrukce potenciálu v $\mathbb{R}^2$.

Označme $(\alpha, \beta) \times (\sigma, \tau) := \Omega$ a zafixujme bod $(a, b) \in \Omega$. Definujme pro každé $(x, y) \in \Omega$

$$U(x, y) = \int_a^x T_1(s, b) \, \mathrm{d}s + \int_b^y T_2(x, t) \, \mathrm{d}t.$$

Definice je korektní, protože funkce $s \mapsto T_1(s, b)$ je spojitá na (α, β). Podobně pro druhý integrál. Zbývá ukázat, že $\frac{\partial U}{\partial x} = T_1$ a $\frac{\partial U}{\partial y} = T_2$ na Ω. Tento výsledek plyne z rovností

$$\frac{\partial}{\partial x} \int_a^x T_1(s, b) \, \mathrm{d}s = T_1(x, b), \quad \frac{\partial}{\partial x} \int_b^y T_2(x, t) \, \mathrm{d}t = T_1(x, y) - T_1(x, b) \tag{12.3.1}$$

a

$$\frac{\partial}{\partial y} \int_a^x T_1(s, b) \, \mathrm{d}s = 0, \quad \frac{\partial}{\partial y} \int_b^y T_2(x, t) \, \mathrm{d}t = T_2(x, y), \tag{12.3.2}$$

jejichž platností se budeme zabývat ve druhém kroku.

Krok 2: důkaz pomocných formulí (12.3.1) a (12.3.2).

První rovnost v (12.3.1) a obě rovnosti v (12.3.2) okamžitě plynou z Takzvané hlavní věty diferenciálního a integrálního počtu (Věta 7.5.12). Dokažme zbývající rovnost. Díky předpokladu $\frac{\partial T_i}{\partial x_j} = \frac{\partial T_j}{\partial x_i}$ máme

$$T_1(x, y) - T_1(x, b) = \int_b^y \frac{\partial T_1}{\partial y}(x, t) \, \mathrm{d}t = \int_b^y \frac{\partial T_2}{\partial x}(x, t) \, \mathrm{d}t.$$

Zbývá tedy ukázat

$$\frac{\partial}{\partial x} \int_b^y T_2(x, t) \, \mathrm{d}t = \int_b^y \frac{\partial T_2}{\partial x}(x, t) \, \mathrm{d}t,$$

neboli

$$\lim_{h \to 0} \int_b^y \Big(\frac{T_2(x + h, t) - T_2(x, t)}{h} - \frac{\partial T_2}{\partial x}(x, t) \Big) \, \mathrm{d}t = 0.$$

V poslední rovnosti můžeme na zlomek uvnitř integrandu použít Lagrangeovu větu o přírůstku funkce (Věta 6.3.3) a pro každé t ležící mezi b a y dostaneme ξ_t ležící mezi 0 a h takové, že

$$\frac{T_2(x + h, t) - T_2(x, t)}{h} = \frac{\partial T_2}{\partial x}(x + \xi_t, t).$$

Nyní si stačí uvědomit, že spojitost $\frac{\partial T_2}{\partial x}$ implikuje stejnoměrnou spojitost $\frac{\partial T_2}{\partial x}$ na kompaktech (v našem případě za kompaktní množinu vezmeme třeba $[x - \delta, x + \delta] \times [\min\{b, y\}, \max\{b, y\}]$ pro $\delta > 0$ dostatečně malé). Ukázali jsme, že k zadanému $\varepsilon > 0$ máme pro h hostatečně blízko k počátku

$$\begin{aligned} &\Big| \int_b^y \Big(\frac{T_2(x + h, t) - T_2(x, t)}{h} - \frac{\partial T_2}{\partial x}(x, t) \Big) \, \mathrm{d}t \Big| \\ &\qquad = \Big| \int_b^y \Big(\frac{\partial T_2}{\partial x}(x + \xi_t, t) - \frac{\partial T_2}{\partial x}(x, t) \Big) \, \mathrm{d}t \Big| \le \Big| \int_b^y \varepsilon \, \mathrm{d}t \Big| = |y - b| \varepsilon. \end{aligned}$$

Z toho plyne požadovaný výsledek.

Krok 3: Modifikace pro obecnou dimenzi.
V obecném případě potenciál konstruujeme předpisem

$$\begin{aligned}U(x_1,\ldots,x_N) = \int_{a_1}^{x_1} T_1(s_1,a_2,\ldots,a_n)\,\mathrm{d}s_1 + \int_{a_2}^{x_2} T_2(x_1,s_2,a_3,\ldots,a_n)\,\mathrm{d}s_2 \\ + \cdots + \int_{a_N}^{x_N} T_N(x_1,\ldots,x_{N-1},s_N)\,\mathrm{d}s_N.\end{aligned}$$

Zbytek důkazu v obecné dimenzi přenecháváme čtenáři na rozmyšlenou (podrobně se zabývejte případem $N = 3$ a vše Vám bude jasné). □

V předchozím výsledku není možné interval nahradit za libovolnou oblast, jak ukazuje následující příklad.

Příklad 12.3.7. Nechť vektorové pole $T : \mathbb{R}^2 \to \mathbb{R}^2$ je na $\mathbb{R}^2 \setminus \{(0,0)\}$ je definováno předpisem

$$T(x,y) = \Big(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2}\Big).$$

Pak $T \in C^\infty(\mathbb{R}^2 \setminus \{(0,0)\}; \mathbb{R}^2)$ a navíc

$$\frac{\partial T_1}{\partial y} = \frac{y^2-x^2}{(x^2+y^2)^2} = \frac{\partial T_2}{\partial x}.$$

Pokud by existoval potenciál U na $\mathbb{R}^2 \setminus \{(0,0)\}$, muselo by pro něj platit podle Newtonovy formule

$$\begin{aligned}0 &= U(1,1) - U(-1,1) + U(-1,1) - U(-1,-1) + U(-1,-1) - U(1,-1) \\ &\qquad + U(1,-1) - U(1,1) \\ &= \int_{-1}^{1} T_1(s,1)\,\mathrm{d}s + \int_{-1}^{1} T_2(-1,t)\,\mathrm{d}t - \int_{-1}^{1} T_1(s,-1)\,\mathrm{d}s - \int_{-1}^{1} T_2(1,t)\,\mathrm{d}t \\ &= \int_{-1}^{1} \frac{-1}{1+s^2}\,\mathrm{d}s + \int_{-1}^{1} \frac{-1}{1+t^2}\,\mathrm{d}t - \int_{-1}^{1} \frac{1}{1+s^2}\,\mathrm{d}s - \int_{-1}^{1} \frac{1}{1+t^2}\,\mathrm{d}t \\ &= 4\int_{-1}^{1} \frac{-1}{1+s^2}\,\mathrm{d}s = -4\frac{\pi}{2} = -2\pi\end{aligned}$$

a máme spor.

Ukažme si ještě techniku hledání potenciálu.

Příklad 12.3.8. Nechť vektorové pole $T : \mathbb{R}^2 \to \mathbb{R}^2$ je na $\mathbb{R}^2$ definováno předpisem

$$T(x,y) = (5x^4y + 2x^3y^2, x^5 + x^4y + 2y\Big).$$

Protože $T \in C^1(\mathbb{R}^2; \mathbb{R}^2)$ a navíc zde platí

$$\frac{\partial(5x^4y + 2x^3y^2)}{\partial y} = 5x^4 + 4x^3y = \frac{\partial(x^5 + x^4y + 2y)}{\partial x},$$

pole T má na $\mathbb{R}^2$ potenciál U. Protože $\frac{\partial U}{\partial x} = 5x^4y + 2x^3y^2$, pro zafixované $y \in \mathbb{R}$ máme

$$U(x,y) = \int (5x^4y + 2x^3y^2)\,\mathrm{d}x = x^5y + \frac{1}{2}x^4y^2 + C(y).$$

Konstanta $C(y)$ závisí na $y \in \mathbb{R}$ a lze ji chápat jako funkci $y \mapsto C(y)$. Abychom ji určili, využijeme druhou podmínku

$$x^5 + x^4y + 2y = \frac{\partial U}{\partial y} = \frac{\partial\big(x^5y + \frac{1}{2}x^4y^2 + C(y)\big)}{\partial y} = x^5 + x^4y + C'(y).$$

Odtud $C'(y) = 2y$, a proto $C(y) = y^2 + C$, kde $C \in \mathbb{R}$. Celkově máme

$$U(x,y) = x^5y + \frac{1}{2}x^4y^2 + y^2 + C \qquad \text{na } \mathbb{R}^2.$$

Mohli jsme také postupovat integrací obou složek T_1, T_2

$$\begin{aligned} U(x,y) &= \int (5x^4y + 2x^3y^2)\,\mathrm{d}x = x^5y + \frac{1}{2}x^4y^2 + C(y),\\ U(x,y) &= \int (x^5 + x^4y + 2y)\,\mathrm{d}y = x^5y + \frac{1}{2}x^4y^2 + y^2 + D(x), \end{aligned}$$

a odtud porovnáním výsledků

$$C(y) = y^2 + C, \qquad D(x) = C.$$

V našich budoucích aplikacích na diferenciální rovnice nebudeme přímo potřebovat potenciálnost studovaného vektorového pole. Postačí nám, bude-li toto vektorové pole potenciální po přenásobení vhodnou (nám známou) funkcí.

Definice 12.3.9 (Integrační faktor)**.** Nechť $T\colon \mathbb{R}^N \to \mathbb{R}^N$ je vektorové pole a $\Omega \subset \mathbb{R}^N$ je otevřená množina. Řekneme, že funkce $\mu\colon \mathbb{R}^N \to \mathbb{R}$ je *integračním faktorem* vektorového pole T na množině Ω, jestliže μT je potenciální na Ω.

Z Věty o nutné podmínce existence potenciálu (Věta 12.3.5) okamžitě dostáváme nutnou podmínku pro integrační faktor.

Věta 12.3.10 (Nutná podmínka pro integrační faktor)**.** *Nechť $\Omega \subset \mathbb{R}^N$ je otevřená množina a $T \in C^1(\Omega;\mathbb{R}^N)$ a $\mu \in C^1(\Omega)$ je integrační faktor pole T na Ω. Pak platí*

$$\frac{\partial(\mu T_i)}{\partial x_j} = \frac{\partial(\mu T_j)}{\partial x_i} \qquad \textit{na } \Omega \qquad \textit{pro všechna } i,j \in \{1,\dots,N\}.$$

Jsou známé také postačující podmínky pro existenci integračního faktoru. Protože však tyto výsledky neposkytují explicitní vzorec pro integrační faktor, nebudeme pro ně mít využití.

Integrační faktor se hledá pomocí metody uhodnutí kombinované s právě uvedenou nutnou podmínkou. Podrobnosti si brzy ukážeme v oddíle o řešení diferenciálních rovnic ve tvaru totálního diferenciálu. Všimněme si, že pro $N = 1$ máme jednu podmínku (tedy jednu diferenciální rovnici) na jednu neznámou funkci μ, zatímco pro $N \geq 2$ máme na jednu funkci μ více než jednu podmínku, proto lze očekávat (a skutečně tomu tak je), že tato úloha je přeurčená a obecně v tomto případě integrační faktor nemusí existovat. Naopak, pro $N = 1$ lze ukázat, že integrační faktor za rozumných předpokladů na hladkost všech funkcí existuje (ovšem jen lokálně, tedy na jistém okolí daného bodu).

12.4 Věta o implicitní funkci

V našem dosavadním výkladu jsme se zatím zabývali pouze explicitně zadanými funkcemi. Často se však v matematice setkáme se situací, že přímý předpis pro studovanou funkci neumíme získat. V takovou chvíli typicky nemáme k dispozici nic z dosud probrané teorie. Tuto teorii nám však může zpřístupnit mocný nástroj, který se nazývá Věta o implicitní funkci. Uveďme si nejprve několik příkladů, abychom si udělali představu o situacích, kterými se budeme zabývat.

Příklad 12.4.1. (i) Uvážíme-li diferenciální rovnici

$$y' = \frac{2x}{3y^2 + 1},$$

jedná se o rovnici se separovanými proměnnými, jejíž řešení je na $\mathbb{R}$ dáno vztahem

$$y^3 + y = \int (3y^2 + 1)\,\mathrm{d}y = \int 2x\,\mathrm{d}x = x^2 + C.$$

Funkce $y \mapsto y^3 + y$ je invertovatelná na celém $\mathbb{R}$. Inverzi nazvěme třeba Φ. Přechod k zápisu

$$y = \Phi(x^2 + C)$$

nám však žádné velké výhody nepřinese, protože pokud bychom chtěli výsledek třeba derivovat, vzoreček z Věty o derivaci inverzní funkce, tedy Věty 3.3.16, (který stejně vede jen na $y' = \frac{2x}{3y^2+1}$) požaduje, abychom k zadanému bodu $x \in \mathbb{R}$ znali jeho funkční hodnotu $y(x)$.
(ii) Uvažme podmnožinu $\mathbb{R}^2$ danou předpisem

$$x^2 + y^2 = 1$$

(tedy jednotkovou kružnici v $\mathbb{R}^2$). Pokud nás zajímá, zda uvedený vztah definuje nějakou funkci, uvědomme si, že pro $x \notin [-1, 1]$ neexistuje $y \in \mathbb{R}$ splňující požadovaný vztah, číslům $x = \pm 1$ jednoznačně odpovídá $y = 0$ a pro $x \in (-1, 1)$ máme dvě řešení $y = \pm\sqrt{1 - x^2}$. Přestože se na první pohled může zdát, že nejlepší situace je v bodech $x = \pm 1$, protože zde máme jednoznačně určené řešení, z hlediska aplikace diferenciálního počtu je mnohem výhodnější případ $x \in (-1, 1)$, kdy za pomoci vhodné počáteční podmínky určíme jednoznačnou (alespoň na malém okolí) větev řešení a tu pak můžeme třeba derivovat.
(iii) Uvažme funkci $f(x) = x^2$ a definujme $L(r)$ jako délku grafu funkce f uvnitř

kružnice se středem v počátku a o poloměru $r > 0$. Například pro $r = \sqrt{2}$ máme (používáme $f(-1) = f(1) = 1$)

$$L(\sqrt{2}) = \int_{-1}^{1} \sqrt{1 + f'^2(x)}\,\mathrm{d}x = 2\int_{0}^{1} \sqrt{1 + 4x^2}\,\mathrm{d}x.$$

Výpočet integrálu není příliš složitý (položí se $2x = \sinh t$), ale situaci komplikuje přepočet mezi poloměrem a délkou intervalu $(-z, z)$ splňujícího

$$\sqrt{z^2 + z^4} = r,$$

přes který integrujeme. Tento problém by opět znesnadnil třeba výpočet

$$L'(r) = \frac{\mathrm{d}}{\mathrm{d}r}\int_{-z(r)}^{z(r)} \sqrt{1 + 4x^2}\,\mathrm{d}x = 2\sqrt{1 + 4z^2(r)}z'(r).$$

(Tento typ úloh se vyskytuje v teorii minimálních ploch.)

Všechny tři části předchozího příkladu můžeme shrnout do situace, že studujeme funkci $x \mapsto y(x)$, kterou nemáme danou explicitně, nýbrž implicitně vztahem

$$F(x, y(x)) = 0,$$

kde F je funkce dvou proměnných (v první části předchozího příkladu máme $F(x,y) = y^3 + y - x^2 - C$, ve druhé $F(x,y) = x^2 + y^2 - 1$ a ve třetí $F(r,z) = \sqrt{z^2 + z^4} - r$).

Náš první výsledek nám říká, kdy je možné vztahem $F(x,y) = 0$ zadefinovat funkci $x \mapsto y(x)$.

Věta 12.4.2 (O implicitní funkci (základní verze))**.** *Nechť* $F\colon \mathbb{R}^{N+1} \to \mathbb{R}$, $a \in \mathbb{R}^N$ *a* $b \in \mathbb{R}$*. Nechť* $F(a,b) = 0$ *a existuje okolí bodu* (a,b)*, kde* F *je spojitá a funkce* $y \mapsto F(x,y)$ *je ryze monotonní (pro všechna* x *stejným způsobem). Pak existují* $\delta, \Delta > 0$ *taková, že pro každé* $x \in \mathcal{U}_\delta(a)$ *existuje právě jedno* $y_x \in \mathcal{U}_\Delta(b)$ *splňující* $F(x, y_x) = 0$*. Navíc funkce* $x \mapsto y_x$ *je spojitá na* $\mathcal{U}_\delta(a)$*.*

Důkaz. Nechť $\tau > 0$ je takové, že funkce $y \mapsto F(x,y)$ je na $\mathcal{U}_\tau((a,b))$ rostoucí (případ klesající funkce se vyřídí analogicky) a F je zde spojitá.

Krok 1: existence y_x.

Zvolme $\Delta > 0$ tak, aby $(a, b-\Delta), (a, b+\Delta) \in \mathcal{U}_\tau((a,b))$. Proto

$$F(a, b-\Delta) < F(a,b) = 0 < F(a, b+\Delta).$$

Odtud díky spojitosti funkce F v bodech $(a, b-\Delta)$ a $(a, b+\Delta)$ dále existuje $\delta > 0$ takové, že pro každé $x \in \mathcal{U}_\delta(a)$ platí $(x, b-\Delta), (x, b+\Delta) \in \mathcal{U}_\tau((a,b))$ a

$$F(x, b-\Delta) < 0 < F(x, b+\Delta).$$

Proto Darbouxova věta (Věta 6.2.1) dává $y_x \in (b-\Delta, b+\Delta)$ splňující $F(x, y_x) = 0$. Na intervalu $(b-\Delta, b+\Delta)$ je y_x jednoznačné díky ryzí monotonii funkce $y \mapsto F(x,y)$.

Krok 2: spojitost $x \mapsto y_x$.
Spojitost dokážeme drobnou modifikací předchozího postupu. Nechť $x \in \mathcal{U}_\delta(a)$ a $\varepsilon > 0$. Případným zmenšením ε dosáhneme toho, že $(x, y_x - \varepsilon), (x, y_x + \varepsilon) \in \mathcal{U}_\tau((a,b))$. Odtud díky spojitosti funkce F v bodech $(x, y_x - \varepsilon)$ a $(x, y_x + \varepsilon)$ a ryzí monotonii funkce $y \mapsto F(x,y)$ existuje $\eta > 0$ takové, že pro každé $\xi \in \mathcal{U}_\eta(a)$ platí $(\xi, y_x - \varepsilon), (\xi, y_x + \varepsilon) \in \mathcal{U}_\tau((a,b))$ a

$$F(\xi, y_x - \varepsilon) < 0 < F(\xi, y_x + \varepsilon).$$

Nutně pak (díky Darbouxově větě a již dokázané jednoznačnosti) platí

$$y_x - \varepsilon < y_\xi < y_x + \varepsilon.$$

□

Poznámka 12.4.3. V aplikacích se nejčastěji setkáme se situací, kdy je monotonie funkce $y \mapsto F(x,y)$ zaručena podmínkami

$$\frac{\partial F}{\partial y}(a,b) \neq 0, \quad \frac{\partial F}{\partial y} \text{ existuje na okolí bodu } (a,b) \text{ a je v něm spojitá.}$$

Příklad 12.4.4. V případě úlohy s $F(x,y) = x^2 + y^2 - 1$ máme

$$\frac{\partial F}{\partial y} = 2y.$$

Předchozí teorii je tedy možné aplikovat kdekoliv na jednotkové kružnici s výjimkou bodů $(-1,0)$ a $(1,0)$.

Poznámka 12.4.5. Věta o implicitní funkci (Věta 12.3.6) se často aplikuje na geometrické problémy. Kupříkladu v rovině nemá x-ová osa nijak odlišnou funkci od osy y-ové. Proto v předchozím příkladu není nijak nepřirozené pokušet se vyjadřovat proměnnou x pomocí proměnné y. V takově situaci o aplikovatelnosti Věty o implicitní funkci rozhoduje podmínka

$$\frac{\partial F}{\partial x} = 2x \neq 0$$

a teorii je možné aplikovat na jednotkové kružnici s výjimkou bodů $(0,-1)$ a $(0,1)$.

Věta 12.4.6 (O derivaci implicitní funkce (základní verze))**.** *Nechť $F \colon \mathbb{R}^{N+1} \to \mathbb{R}$, $k \in \mathbb{N} \cup \{\infty\}$, $a \in \mathbb{R}^N$, $b \in \mathbb{R}$. Nechť $F(a,b) = 0$, existuje okolí bodu (a,b), kde F je třídy C^k a $\frac{\partial F}{\partial y}(a,b) \neq 0$ (parciální derivace podle poslední proměnné). Pak existují $\delta, \Delta > 0$ taková, že pro každé $x \in \mathcal{U}_\delta(a)$ existuje právě jedno $y_x \in \mathcal{U}_\Delta(b)$ splňující $F(x,y_x) = 0$ a funkce $\varphi \colon x \mapsto y_x$ je třídy C^k na $\mathcal{U}_\delta(a)$. Navíc*

$$\frac{\partial \varphi}{\partial x_j}(x) = -\frac{\frac{\partial F}{\partial x_j}(x, \varphi(x))}{\frac{\partial F}{\partial y}(x, \varphi(x))} \qquad \textit{pro všechna } j \in \{1, \ldots, N\} \textit{ a } x \in \mathcal{U}_\delta(a).$$

Důkaz. Podle Poznámky 12.4.3 jsou splněny předpoklady předchozí věty. Proto je funkce φ definována na $\mathcal{U}_\delta(a)$, je zde spojitá a tuto množinu zobrazuje do intervalu $(b-\Delta, b+\Delta)$. Případným zmenšením δ dosahneme toho, že pro všechna $x \in \mathcal{U}_\delta(a)$ je $(x,\varphi(x))$ uvnitř okolí, kde je F třídy C^k. Zafixujme $x \in \mathcal{U}_\delta(a)$. Funkce F má v bodě $(x,\varphi(x))$ totální diferenciál. Pro $h \in \mathbb{R}^N$ a $s \in \mathbb{R}$ dostatečně blízko k počátku (ve své dimenzi) pak máme

$$F(x+h,\varphi(x)+s) - F(x,\varphi(x)) = \sum_{i=1}^N \frac{\partial F}{\partial x_i}(x,\varphi(x))h_i + \frac{\partial F}{\partial y}(x,\varphi(x))s + \eta(h,s),$$

kde η splňuje

$$\frac{\eta(h,s)}{\|(h,s)\|} \to 0 \qquad \text{pro } (h,s) \to 0.$$

Poslední vlastnost navíc umožňuje přepis

$$\begin{aligned}
\eta(h,s) &= \frac{\eta(h,s)}{\|(h,s)\|}\frac{h_1^2+\cdots+h_N^2+s^2}{\|(h,s)\|} \\
&= \frac{\eta(h,s)}{\|(h,s)\|}\frac{h_1}{\|(h,s)\|}h_1 + \cdots + \frac{\eta(h,s)}{\|(h,s)\|}\frac{h_N}{\|(h,s)\|}h_N + \frac{\eta(h,s)}{\|(h,s)\|}\frac{s}{\|(h,s)\|}s \\
&=: \xi_1(h,s)h_1 + \cdots + \xi_N(h,s)h_N + \xi_{N+1}(h,s)s,
\end{aligned}$$

kde

$$\xi_i(h,s) \to 0 \qquad \text{pro } (h,s)\to 0 \qquad \text{pro každé } i \in \{1,\dots,N,N+1\}. \tag{12.4.1}$$

Položme nyní výše $h = te_j$, kde e_j je j-tý bázový vektor a $t \in \mathbb{R}\setminus\{0\}$ je dostatečně malé. Pak pro $s := \varphi(x+h)-\varphi(x)$ dostáváme

$$\begin{aligned}
&F(x+te_j,\varphi(x+te_j)) - F(x,\varphi(x)) \\
&\quad= \frac{\partial F}{\partial x_j}(x,\varphi(x))t + \frac{\partial F}{\partial y}(x,\varphi(x))(\varphi(x+te_j)-\varphi(x)) \\
&\qquad\qquad + \xi_j(te_j,\varphi(x+te_j)-\varphi(x))t \\
&\qquad\qquad + \xi_{N+1}(te_j,\varphi(x+te_j)-\varphi(x))(\varphi(x+te_j)-\varphi(x)).
\end{aligned}$$

Odtud díky tomu, že oba členy levé strany jsou nulové, máme

$$\frac{\varphi(x+te_j)-\varphi(x)}{t} = -\frac{\frac{\partial F}{\partial x_j}(x,\varphi(x)) + \xi_j(te_j,\varphi(x+te_j)-\varphi(x))}{\frac{\partial F}{\partial y}(x,\varphi(x)) + \xi_{N+1}(te_j,\varphi(x+te_j)-\varphi(x))}.$$

Provedeme-li nyní limitní přechod $t \to 0$ (připomeňme, že φ je spojitá funkce a máme (12.4.1)), dostáváme dokazovaný vzorec ze znění věty. Protože parciální derivace na pravé straně získaného vzorce jsou spojité na okolí bodu (a,b) a φ je spojitá na okolí bodu a, je rovněž spojitá $\frac{\partial\varphi}{\partial x_j}$ na odpovídajícím okolí bodu a. Tím je důkaz dokončen v případě, že $k=1$.

Pokud $k=2$, pro libovolné $i \in \{1,\dots,N\}$ máme díky spojitosti druhých parciálních derivací funkce F a spojitosti prvních parciálních derivací funkce φ použitím řetízkového pravidla (na třetím řádku z důvodu lepší čitelnosti vynecháváme argument $(x, \varphi(x))$)

$$\begin{aligned}
\frac{\partial^2 \varphi}{\partial x_i \partial x_j}(x) &= -\frac{\partial}{\partial x_i}\left(\frac{\frac{\partial F}{\partial x_j}(x,\varphi(x))}{\frac{\partial F}{\partial y}(x,\varphi(x))}\right) \\
&= -\frac{\frac{\partial}{\partial x_i}(\frac{\partial F}{\partial x_j}(x,\varphi(x)))\frac{\partial F}{\partial y}(x,\varphi(x)) - \frac{\partial F}{\partial x_j}(x,\varphi(x))\frac{\partial}{\partial x_i}(\frac{\partial F}{\partial y}(x,\varphi(x)))}{(\frac{\partial F}{\partial y}(x,\varphi(x)))^2} \\
&= -\frac{(\frac{\partial^2 F}{\partial x_i \partial x_j} + \frac{\partial^2 F}{\partial y \partial x_j}\frac{\partial \varphi}{\partial x_i})\frac{\partial F}{\partial y} - \frac{\partial F}{\partial x_j}(\frac{\partial^2 F}{\partial x_i \partial y} + \frac{\partial^2 F}{\partial y^2}\frac{\partial \varphi}{\partial x_i})}{(\frac{\partial F}{\partial y})^2}.
\end{aligned}$$

Z výsledného vzorce také vidíme, že spojitost druhých parciálních derivací funkce F a spojitost prvních parciálních derivací funkce φ implikují spojitost druhých parciálních derivací funkce φ. Pro $k \geq 3$ pokračujeme indukcí. Při počítání parciálních derivací l-tého řádu se využijí parciální derivace prvního až l-tého řádu funkce F a parciální derivace prvního až $(l-1)$-tého řádu funkce φ. Spojitost zdůvodníme ze získaného vzorce pomocí aritmetiky spojitosti. □

Poznámka 12.4.7. Když už máme dokázanou existenci parciálních derivací funkce φ, dá se při jejich výpočtu využívat následující pohodlnější postup (který se v praxi upřednostňuje před derivováním podílu). Vyjdeme z toho, že na $\mathcal{U}_\delta(a)$ platí rovnost $F(x, \varphi(x)) = 0$ a tu postupně derivujeme (pozor, pracujeme v $\mathbb{R}^N$, nikoliv v $\mathbb{R}^{N+1}$)

$$0 = \frac{\partial 0}{\partial x_j} = \frac{\partial}{\partial x_j}F(x,\varphi(x)) = \frac{\partial F}{\partial x_j}(x,\varphi(x)) + \frac{\partial F}{\partial y}(x,\varphi(x))\frac{\partial \varphi}{\partial x_j}(x,\varphi(x)).$$

Z této rovnosti je možné vyjádřit $\frac{\partial \varphi}{\partial x_j}(x,\varphi(x))$, neboť $\frac{\partial F}{\partial y}(x,\varphi(x)) \neq 0$ na $\mathcal{U}_\delta(a)$. Získanou rovnost (pro $F \in C^2$ na okolí bodu (a,b)) můžeme dále derivovat a dostáváme (opět pro přehlednost vynecháváme argument $(x,\varphi(x))$)

$$\begin{aligned}
0 &= \frac{\partial}{\partial x_i}\Big(\frac{\partial F}{\partial x_j} + \frac{\partial F}{\partial y}\frac{\partial \varphi}{\partial x_j}\Big) \\
&= \frac{\partial^2 F}{\partial x_i \partial x_j} + \frac{\partial^2 F}{\partial y \partial x_j}\frac{\partial \varphi}{\partial x_i} + \frac{\partial^2 F}{\partial x_i \partial y}\frac{\partial \varphi}{\partial x_j} + \frac{\partial^2 F}{\partial y^2}\frac{\partial \varphi}{\partial x_i}\frac{\partial \varphi}{\partial x_j} + \frac{\partial F}{\partial y}\frac{\partial^2 \varphi}{\partial x_i \partial x_j}.
\end{aligned}$$

Takto bychom mohli pokračovat dále. Parciální derivace funkce φ nejvyššího řádu (ta nás zajímá) se vyskytuje vždy v jediném členu, je vždy vynásobena nenulovým výrazem $\frac{\partial F}{\partial y}$ (takže tímto výrazem můžeme dělit). V identitě se vyskytují parciální derivace funkce F nejvýše stejného řadu a dále parciální derivace funkce φ nižších řádů, které jsme si vyjádřili v předchozích krocích.

Poznámka 12.4.8. Samozřejmě není vůbec podstatné, že vyjadřujeme poslední proměnnou pomocí ostatních. Takto lze vyjadřovat kteroukoliv z proměnných, je-li parciální derivace funkce F vůči této proměnné ve studovaném bodě nenulová.

Příklad 12.4.9. Ukažme, že identita

$$\log\sqrt{x^2+y^2} = \arctan\frac{y}{x}$$

na jistém okolí bodu $(x_0, y_0) = (1, 0)$ jednoznačně určuje funkci $y = y(x)$ a spočítejme $y'(1)$ a $y''(1)$. Definujme funkci

$$F(x,y) = \log\sqrt{x^2+y^2} - \arctan\frac{y}{x}.$$

Okamžitě vidíme, že $F(1, 0) = \log 1 - \arctan 0 = 0$. Dále

$$\frac{\partial F}{\partial x} = \frac{x}{x^2+y^2} - \frac{-\frac{y}{x^2}}{1+\frac{y^2}{x^2}} = \frac{x+y}{x^2+y^2}$$

a

$$\frac{\partial F}{\partial y} = \frac{y}{x^2+y^2} - \frac{\frac{1}{x}}{1+\frac{y^2}{x^2}} = \frac{y-x}{x^2+y^2}.$$

Odtud jednak vidíme, že $\frac{\partial F}{\partial y}(1, 0) \neq 0$. Dále z pravidel pro derivaci podílu dostáváme, že F je třídy C^∞ na jakékoliv podmnožině $\mathbb{R}^2$ neobsahující počátek. Jsou tedy splněny předpoklady předchozích dvou vět, které nám dávají funkci $y = y(x)$ a můžeme ji derivovat. Použijeme postup z předchozí poznámky. Budeme derivovat rovnost

$$0 = \log\sqrt{x^2+y^2(x)} - \arctan\frac{y(x)}{x}.$$

Na dostatečně malém okolí bodu 1 máme

$$0 = \frac{1}{2(x^2+y^2)}(2x+2yy') - \frac{1}{1+\frac{y^2}{x^2}}\Big(\frac{-y}{x^2}+\frac{y'}{x}\Big). \tag{12.4.2}$$

Tedy v bodě $x = 1$ (připomeňme $y(1) = 0$) platí

$$0 = \frac{1}{2}(2+0\cdot y'(1)) - 1\cdot(0+y'(1)) \qquad \Longrightarrow \qquad y'(1) = 1.$$

Hodnotu $y''(1)$ zjistíme dalším derivováním vztahu (12.4.2) (pro pohodlnější derivování jej napřed upravíme do podoby $0 = \frac{x+y+(y-x)y'}{x^2+y^2}$) a máme

$$0 = \frac{\big(1+y'+(y'-1)y'+(y-x)y''\big)(x^2+y^2) - \big(x+y+(y-x)y'\big)(2x+2yy')}{(x^2+y^2)^2}.$$

Do poslední identity dosadíme $x = 1$ (také $y(1) = 0$ a již spočítaný výsledek $y'(1) = 1$)

$$\begin{aligned} 0 &= \frac{\big(1+1+(1-1)\cdot 1+(0-1)y''(1)\big)(1+0) - (1+0+(0-1)\cdot 1)(2+0)}{(1+0)^2} \\ &= 1+1-y''(1)-0\cdot 2. \end{aligned}$$

Proto $y''(1) = 2$.

Posledně představená metoda derivování implicitně zadané funkce je natolik uživatelsky příjemná, že se často vyplatí přistupovat přes implicitní funkce k problémům, které svou povahou s implicitně zadanými funkcemi mají pramálo společného.

Příklad 12.4.10. Za pomoci teorie implicitních funkcí spočítejme druhou a třetí derivaci inverzní funkce. Nechť tedy máme standardní situaci, kdy funkce $f\colon \mathbb{R} \to \mathbb{R}$ má na okolí bodu $y_0 \in \mathbb{R}$ derivaci neměnící znaménko (všude kladná, nebo všude záporná). Nechť na tomto okolí existují také druhá a třetí derivace funkce f. Vztah $f^{-1}(x) = y$ na vhodných množinách odpovídá vztahu $x = f(y)$, který využijeme ke konstrukci funkce F. Položme

$$F(x, y) = x - f(y)$$

(tedy $x = f(y)$ je splněno právě tehdy, když $F(x, y) = 0$). Označme $x_0 = f(y_0)$. Pak zřejmě $F(x_0, y_0) = 0$. Fukce F je třídy C^k právě tehdy, když funkce f je třídy C^k. Navíc $\frac{\partial F}{\partial y}(x_0, y_0) = -f'(y_0) \neq 0$. Podle předchozí věty vztah $F(x, y)$ definuje C^k-funkci $x \mapsto y(x)$ (námi studovanou funkci f^{-1}). Postupným derivováním dostáváme z rovnosti $0 = x - f(y(x))$

$$\begin{aligned}
0 &= 1 - f'(y(x))y'(x)\\
0 &= -f''(y(x)){y'}^2(x) - f'(y(x))y''(x)\\
0 &= -f'''(y(x)){y'}^3(x) - 2f''(y(x))y'(x)y''(x) - f''(y(x))y'(x)y''(x)\\
&\qquad - f'(y(x))y'''(x)\\
&= -f'''(y(x)){y'}^3(x) - 3f''(y(x))y'(x)y''(x) - f'(y(x))y'''(x).
\end{aligned}$$

Z první rovnosti snadnou úpravou dostáváme nám znamý vzorec

$$y'(x) = \frac{1}{f'(y(x))}.$$

Druhá rovnost spolu s právě získaným výsledkem dávají

$$y''(x) = -\frac{f''(y(x)){y'}^2(x)}{f'(y(x))} = -\frac{f''(y(x))}{{f'}^3(y(x))}$$

a podobně ze třetí rovnosti dostáváme

$$\begin{aligned}
y'''(x) &= -\frac{f'''(y(x)){y'}^3(x) + 3f''(y(x))y'(x)y''(x)}{f'(y(x))}\\
&= -\frac{f'''(y(x))\frac{1}{{f'}^3(y(x))} - 3f''(y(x))\frac{1}{f'(y(x))}\frac{f''(y(x))}{{f'}^3(y(x))}}{f'(y(x))}\\
&= \frac{3{f''}^2(y(x)) - f'''(y(x))f'(y(x))}{{f'}^5(y(x))}.
\end{aligned}$$

Poznámka 12.4.11. K derivování inverzní funkce přístup přes implicitní funkce není potřeba. Je možné derivovat základní vzoreček $y' = \frac{1}{f'(y(x))}$. Pokud pro přehlednost nebudeme psát argumenty, postup vypadá následovně

$$
\begin{aligned}
y'' &= \frac{-f''y'}{f'^2} = \frac{-f''}{f'^3} \\
y''' &= \frac{-f'''y'f'^3 + 3f''f'^2f''y'}{f'^6} = \frac{-f'''f'^2 + 3f''^2f'}{f'^6} = \frac{3f''^2 - f'''f'}{f'^5}.
\end{aligned}
$$

Naše věty o implicitních funkcích nám umožňují pracovat s derivací, což je nejdůležitější veličina pro vyšetřování průběhu funkce.

Příklad 12.4.12. Uvažme vztah

$$
\mathrm{e}^y + \log x + xy = 0. \tag{12.4.3}
$$

Levá strana má smysl pro $x > 0$ a $y = \mathbb{R}$. Definujme $F(x, y) = \mathrm{e}^y + \log x + xy$. Pro každé $x > 0$ pevné platí

$$
\lim_{y\to-\infty} F(x, y) = -\infty, \qquad \lim_{y\to\infty} F(x, y) = \infty \qquad \text{a} \qquad \frac{\partial F}{\partial y} = \mathrm{e}^y + x > 0.
$$

Odtud okamžitě vidíme, že pro každé $x > 0$ existuje jednoznačné $y(x)$ splňující $F(x, y(x)) = 0$. Ze základní verze Věty o derivaci implicitní funkce (Věta 12.4.6) dostáváme, že zobrazení $\varphi\colon x \mapsto y(x)$ splňuje $\varphi \in C^\infty((0, \infty))$. Povšimněme si dále, že platí

$$
F(x, 0) = 0 \qquad \Longleftrightarrow \qquad x = \frac{1}{\mathrm{e}}.
$$

Navíc

$$
\varphi'(\tfrac{1}{\mathrm{e}}) = -\frac{\frac{\partial F}{\partial x}(\frac{1}{\mathrm{e}}, 0)}{\frac{\partial F}{\partial y}(\frac{1}{\mathrm{e}}, 0)} = -\frac{\frac{1}{x} + y|_{(x,y)=(\frac{1}{\mathrm{e}},0)}}{\mathrm{e}^y + x|_{(x,y)=-(\frac{1}{\mathrm{e}},0)}} = -\frac{\mathrm{e} + 0}{1 + \frac{1}{\mathrm{e}}} < 0.
$$

Odtud $\varphi > 0$ na $(0, \frac{1}{\mathrm{e}})$ a $\varphi < 0$ na $(\frac{1}{\mathrm{e}}, \infty)$. Tento výsledek spolu se vzorcem

$$
\varphi'(x) = -\frac{\frac{1}{x} + \varphi(x)}{\mathrm{e}^{\varphi(x)} + x}
$$

zaručují, že φ je klesající na $(0, \frac{1}{\mathrm{e}})$. Navíc si ve formuli (12.4.3) povšimněme, že pro $x \to 0_+$ musí platit $y(x) \to \infty$, tedy

$$
\lim_{x\to 0_+} \varphi(x) = \infty.
$$

Podobně pozorováním formule (12.4.3) v případě, že $x \to \infty$, objevíme

$$
\lim_{x\to\infty} \varphi(x) = 0.
$$

Zabývejme se dále otázkou, kdy platí $\varphi'(x) = 0$. Pro takový bod musí platit $\frac{1}{x} + y(x) = 0$ a s využitím (12.4.3) přicházíme k podmínce

$$
\psi(x) := \mathrm{e}^{-\frac{1}{x}} + \log x - 1 = 0.
$$

Funkce ψ splňuje

$$\psi < 0 \quad \text{na } (\tfrac{1}{\mathrm{e}}, 1], \qquad \lim_{x\to\infty} \psi(x) = \infty \quad \text{a} \quad \psi'(x) = -\mathrm{e}^{-\frac{1}{x}} \frac{1}{x^2} + \frac{1}{x} > 0 \quad \text{na } (1, \infty).$$

Odtud vidíme, že na intervalu $(\frac{1}{\mathrm{e}}, \infty)$ existuje právě jeden bod x_0, kde má funkce φ nulovou derivaci. Navíc spojitost φ' spolu s $\varphi'(\frac{1}{\mathrm{e}}) < 0$ a $\varphi'(x_0) = 0$ implikují, že $\varphi' < 0$ na $(\frac{1}{\mathrm{e}}, x_0)$ a φ je zde klesající. Naopak o intervalu (x_0, ∞) víme, že zde derivace nemá nulový bod, tedy φ je ryze monotonní a limitní chování v nekonečnu spolu se záporností funkce připouštějí pouze možnost, že funkce φ je zde rostoucí. Navíc jsme také zjistili, že v bodě x_0 je globální minimum.

Plná verze Věty o implicitní funkci pracuje s m-ticí funkcí $F_1, \dots, F_m \colon \mathbb{R}^{N+m} \to \mathbb{R}$ a nabízí vyjádření poslední m-tice proměnných pomocí proměnných ostatních.

Věta 12.4.13 (O implicitní funkci (plná verze))**.** *Nechť $N, m \in \mathbb{N}$, $k \in \mathbb{N} \cup \{\infty\}$, $F \colon \mathbb{R}^{N+m} \to \mathbb{R}^m$, $a \in \mathbb{R}^N$, $b \in \mathbb{R}^m$. Nechť $F(a,b) = (0, \dots, 0)$, existuje okolí bodu (a,b), kde všechny složky zobrazení F jsou třídy C^k, a*

$$\det \begin{pmatrix} \frac{\partial F_1}{\partial y_1}(a,b) & \frac{\partial F_1}{\partial y_2}(a,b) & \cdots & \frac{\partial F_1}{\partial y_m}(a,b) \\ \frac{\partial F_2}{\partial y_1}(a,b) & \frac{\partial F_2}{\partial y_2}(a,b) & \cdots & \frac{\partial F_2}{\partial y_m}(a,b) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1}(a,b) & \frac{\partial F_m}{\partial y_2}(a,b) & \cdots & \frac{\partial F_m}{\partial y_m}(a,b) \end{pmatrix} \neq 0.$$

Pak existují $\delta, \Delta > 0$ taková, že pro každé $x \in \mathcal{U}_\delta(a)$ existuje právě jedno $y_x \in \mathcal{U}_\Delta(b)$ splňující $F(x, y_x) = (0, \dots, 0)$ a pro zobrazení $\varphi \colon x \mapsto y_x$ platí, že $\varphi \in C^k(\mathcal{U}_\delta(a); \mathbb{R}^m)$.

Důkaz. Důkaz provedeme indukcí přes $m \in \mathbb{N}$. Případ $m = 1$ plyne z předchozích dvou vět. Předpokládejme, že věta platí pro nějaké $m \in \mathbb{N}$. Dokážeme platnost věty pro $m + 1$. Označme matici ze znění věty

$$A := \begin{pmatrix} \frac{\partial F_1}{\partial y_1}(a,b) & \cdots & \frac{\partial F_1}{\partial y_{m+1}}(a,b) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_{m+1}}{\partial y_1}(a,b) & \cdots & \frac{\partial F_{m+1}}{\partial y_{m+1}}(a,b) \end{pmatrix}.$$

Krok 1: případ jednotkové matice A.
Myšlenka důkazu je následující. Nejprve použijeme funkci F_{m+1} k tomu, abychom vyjádřili proměnnou y_{m+1} pomocí ostatních proměnných, což nám umožňují předchozí věty. Pak použijeme funkce $F_1, \dots, F_m$ k tomu, abychom vyjádřili proměnné $y_1, \dots, y_m$ pomocí $x_1, \dots, x_N$, což nám umožňuje indukční předpoklad.

Přistupme k podrobnému důkazu. Protože $F_{m+1}(a,b) = 0$ a $\frac{\partial F_{m+1}}{\partial y_{m+1}} = 1 \neq 0$, podle předchozích dvou vět existují $\delta_1, \Delta_1 > 0$ taková, že pro každé

$$(x_1, \dots, x_N, y_1, \dots, y_m) \in \mathcal{U}_{\delta_1}((a_1, \dots, a_N, b_1, \dots, b_m))$$

existuje právě jedno $y_{m+1} \in \mathcal{U}_{\Delta_1}(b_{m+1})$ takové, že

$$F_{m+1}(x_1, \dots, x_N, y_1, \dots, y_m, y_{m+1}) = 0.$$

Navíc zobrazení $\varphi_{m+1}\colon (x_1,\dots,x_N,y_1,\dots,y_m)\mapsto y_{m+1}$ splňuje

$$\varphi_{m+1}\in C^k(\mathcal{U}_{\delta_1}((a_1,\dots,a_N,b_1,\dots,b_m))).$$

Nyní definujme pro každé $i\in\{1,\dots,m\}$ pomocnou funkci

$$\begin{aligned} &H_i(x_1,\dots,x_N,y_1,\dots,y_m)\\ &\qquad := F_i(x_1,\dots,x_N,y_1,\dots,y_m,\varphi_{m+1}(x_1,\dots,x_N,y_1,\dots,y_m)). \end{aligned}$$

Podle řetízkového pravidla platí v bodě $(a_1,\dots,a_N,b_1,\dots,b_m)$ pro každé $i,j\in\{1,\dots,m\}$

$$\frac{\partial H_i}{\partial y_j}=\frac{\partial F_i}{\partial y_j}+\frac{\partial F_i}{\partial y_{m+1}}\frac{\partial\varphi_{m+1}}{\partial y_j}=\delta_{ij}+0\cdot\frac{\partial\varphi_{m+1}}{\partial y_j}=\delta_{ij},$$

tedy odpovídající matice parciálních derivací je jednotková (hlavně regulární). Pomocí řetízkového pravidla, hladkosti F_i a φ_{m+1} snadno nahlédneme, že

$$H_i\in C^k(\mathcal{U}_{\delta_1}((a_1,\dots,a_N,b_1,\dots,b_m))).$$

Na funkce H_i proto můžeme aplikovat indukční předpoklad a dostáváme $\delta_2,\Delta_2>0$ taková, že pro každé $(x_1,\dots,x_N)\in\mathcal{U}_{\delta_2}((a_1,\dots,a_N))$ existuje jednoznačné $(y_1,\dots,y_m)\in\mathcal{U}_{\Delta_2}((b_1,\dots,b_m))$ splňující pro každé $i\in\{1,\dots,m\}$

$$\begin{aligned} 0&=H_i(x_1,\dots,x_m,y_1,\dots,y_m)\\ &=F_i(x_1,\dots,x_N,y_1,\dots,y_m,\varphi_{m+1}(x_1,\dots,x_N,y_1,\dots,y_m)) \end{aligned}$$

a zobrazení $\varphi_i\colon (x_1,\dots,x_m)\mapsto y_i$ jsou třídy C^k na $\mathcal{U}_{\delta_2}((a_1,\dots,a_N))$.

Z toho plynou všechny požadované výsledky až na to, že nemáme správný tvar okolí vůči proměnným $y_1,\dots,y_{m+1}$. Pokud si však připomeneme, že platí $\varphi_i(a_1,\dots,a_N)=b_i$ a funkce φ_i jsou spojité na dostatečně malých okolích bodu $(a_1,\dots,a_N)$, patřičným zmenšením δ_2 dosáhneme i tohoto výsledku.

Krok 2: případ obecné matice A.
Protože A je v obecném případě regulární, existuje inverzní matice A^{-1}. Nechť nyní $L\colon\mathbb{R}^{m+1}\to\mathbb{R}^{m+1}$ je lineární (tedy C^∞) zobrazení reprezentované maticí A^{-1}. Definujme zobrazení $T:=L\circ F$. Zřejmě se jedná o C^k-zobrazení na okolí bodu (a,b). Navíc díky tomu, že řetízkové pravidlo odpovídá součinu matic, máme

$$\begin{pmatrix} \frac{\partial T_1}{\partial y_1}(a,b) & \frac{\partial T_1}{\partial y_2}(a,b) & \cdots & \frac{\partial T_1}{\partial y_{m+1}}(a,b)\\ \frac{\partial T_2}{\partial y_1}(a,b) & \frac{\partial T_2}{\partial y_2}(a,b) & \cdots & \frac{\partial T_2}{\partial y_{m+1}}(a,b)\\ \vdots & \vdots & \ddots & \vdots\\ \frac{\partial T_{m+1}}{\partial y_1}(a,b) & \frac{\partial T_{m+1}}{\partial y_2}(a,b) & \cdots & \frac{\partial T_{m+1}}{\partial y_{m+1}}(a,b) \end{pmatrix}=\begin{pmatrix} 1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Podle předchozích kroků existují $\delta,\Delta>0$ taková, že pro každé $x\in\mathcal{U}_\delta(a)$ existuje právě jedno $z_x\in\mathcal{U}_\Delta(b)$ splňující $T(x,z_x)=(0,\dots,0)$ a jednotlivé složky zobrazení $\psi\colon x\mapsto z_x$ jsou třídy C^k na $\mathcal{U}_\delta(a)$. Podle definice zobrazení T zřejmě platí

$$T(x,z_x)=(0,\dots,0)\qquad\Longleftrightarrow\qquad F(x,z_x)=(0,\dots,0).$$

Odtud bod $z_x\in\mathbb{R}^{m+1}$ je hledaný bod y_x, hledané zobrazení φ je ψ a má požadované vlastnosti. $\square$

Příklad 12.4.14. Rovnice

$$x + y + z = 0$$
$$x^2 + y^2 + z^2 = 1$$

jsou splněny v bodě $(\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$. Zabývejme se otázkou, zda na okolí tohoto bodu uvedené rovnice jednoznačně určují proměnné y, z pomocí proměnné x. Píšeme $(a_1, b_1, b_2) := (\frac{1}{\sqrt{2}}, 0, -\frac{1}{\sqrt{2}})$ a zavedeme C^∞-funkce

$$F_1(x, y, z) = x + y + z$$
$$F_2(x, y, z) = x^2 + y^2 + z^2 - 1.$$

Máme

$$\begin{pmatrix} \frac{\partial F_1}{\partial y} & \frac{\partial F_1}{\partial z} \\ \frac{\partial F_2}{\partial y} & \frac{\partial F_2}{\partial z} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 2y & 2z \end{pmatrix}.$$

Odtud

$$\det \begin{pmatrix} \frac{\partial F_1}{\partial y}(a_1, b_1, b_2) & \frac{\partial F_1}{\partial z}(a_1, b_1, b_2) \\ \frac{\partial F_2}{\partial y}(a_1, b_1, b_2) & \frac{\partial F_2}{\partial z}(a_1, b_1, b_2) \end{pmatrix} = \det \begin{pmatrix} 1 & 1 \\ 0 & -\sqrt{2} \end{pmatrix} = -\sqrt{2} \neq 0.$$

Můžeme tedy použít plnou verzi Věty o implicitní funkci (Věta 12.4.13) a ta nám dává vyjádření $z = z(x)$ a $y = y(x)$ jako nekonečněkrát diferencovatelných funkcí na nějakém okolí bodu a_1. Zde máme (protože je zde $N = 1$, jedná se o klasickou derivaci)

$$0 = (F_1(x, y(x), z(x)))' = 1 + y' + z'$$
$$0 = (F_2(x, y(x), z(x)))' = 2x + 2yy' + 2zz'.$$

Odtud dostáváme

$$y' = -1 - z'$$
$$0 = 2x + 2y(-1 - z') + 2zz' = 2x - 2y + (2z - 2y)z'.$$

Proto

$$z' = \frac{y - x}{z - y} \qquad \text{a} \qquad y' = -1 - \frac{y - x}{z - y} = \frac{x - z}{z - y}.$$

Speciálně

$$z'(a_1) = \frac{-\frac{1}{\sqrt{2}}}{-\frac{1}{\sqrt{2}}} = 1 \qquad \text{a} \qquad y'(a_1) = \frac{\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{2}}}{-\frac{1}{\sqrt{2}}} = -2.$$

Pokud by nás zajímaly derivace druhého řádu, můžeme dále počítat

$$0 = (F_1(x, y(x), z(x)))'' = (1 + y' + z')' = y'' + z''$$
$$0 = (F_2(x, y(x), z(x)))'' = (2x + 2yy' + 2zz')' = 2 + 2{y'}^2 + 2yy'' + 2{z'}^2 + 2zz''.$$

Odtud $y'' = -z''$ a

$$0 = 2 + 2{y'}^2 - 2yz'' + 2{z'}^2 + 2zz'' = 2 + 2{y'}^2 + 2{z'}^2 + (2z - 2y)z''.$$

Proto

$$z'' = \frac{1 + {y'}^2 + {z'}^2}{y - z} = \frac{(y - z)^2 + (x - z)^2 + (y - x)^2}{(y - z)^3} = -y''.$$

Poznámka 12.4.15. (i) Povšimněme si, že nenulovost determinantu ze znění věty je v našem případě ekvivalentní podmínce $y \neq z$. Proto se dá plná verze Věty o implicitní funkci (Věta 12.4.13) používat na okolí všech bodů splňujících

$$x + y + z = 0 \qquad \text{a} \qquad x^2 + y^2 + z^2 = 1$$

s výjimkou bodů, kde

$$x + y + z = 0 \qquad \text{a} \qquad x^2 + y^2 + z^2 = 1 \qquad \text{a} \qquad y = z.$$

(ii) Po krátkém výpočtu se dá zjistit, že zakázány jsou body $\left(\sqrt{\frac{2}{3}}, -\sqrt{\frac{1}{6}}, -\sqrt{\frac{1}{6}}\right)$ a $\left(-\sqrt{\frac{2}{3}}, \sqrt{\frac{1}{6}}, \sqrt{\frac{1}{6}}\right)$. Rovnice $x^2 + y^2 + z^2 = 1$ popisuje jednotkovou sféru v $\mathbb{R}^3$ a rovnice $x + y + z = 0$ popisuje jednu z rovin procházejících počátkem. Obě rovnice jsou splněny na průniku těchto množin, což je kružnice o jednotkovém poloměru.

12.5 Rovnice ve tvaru totálního diferenciálu

V dalším budeme uvažovat rovnice typu

$$M(x, y) + N(x, y)\frac{\mathrm{d}y}{\mathrm{d}x} = 0 \qquad \text{respektive} \qquad M(x, y)\frac{\mathrm{d}x}{\mathrm{d}y} + N(x, y) = 0.$$

Tyto rovnice se často prezentují ve tvaru

$$M(x, y)\,\mathrm{d}x + N(x, y)\,\mathrm{d}y = 0. \tag{12.5.1}$$

Doposud jsme v našem výkladu nikdy nezavedli symboly „dx“ a „dy“ samostatně, nejedná se tedy o matematicky korektní zápis. To alespoň v této situaci napravíme.

Definice 12.5.1 (Rovnice ve tvaru totálního diferenciálu)**.** Rovnici (12.5.1) nazveme *rovnicí ve tvaru totálního diferenciálu* na otevřené množině $\Omega \subset \mathbb{R}^2$, jestliže existuje $U\colon \mathbb{R}^2 \to \mathbb{R}$ takové, že levá strana rovnice (12.5.1) je totálním diferenciálem funkce U na Ω, neboli pro všechna $(x, y) \in \Omega$ a $(h_1, h_2) \in \mathbb{R}^2$ platí

$$\mathrm{d}U(x, y)(h_1, h_2) = M(x, y)h_1 + N(x, y)h_2.$$

Funkci U v takovém případě nazýváme *potenciálem* rovnice (12.5.1).

Věta 12.5.2 (O řešení rovnice ve tvaru totálního diferenciálu)**.** *Nechť U je potenciálem rovnice* (12.5.1) *na otevřené množině $\Omega \subset \mathbb{R}^2$, $M, N \in C(\Omega)$ a $N \neq 0$ na Ω. Pak každým bodem $(x_0, y_0) \in \Omega$ prochází právě jedno řešení rovnice*

$$M(x, y) + N(x, y)\frac{\mathrm{d}y}{\mathrm{d}x} = 0$$

a je implicitně dáno vztahem

$$U(x, y) = U(x_0, y_0).$$

Pokud $M \neq 0$ na Ω, platí analogický výsledek pro rovnici

$$M(x, y)\frac{\mathrm{d}x}{\mathrm{d}y} + N(x, y) = 0.$$

Důkaz. Nejprve dokažme existenci. Podle základní verze Věty o implicitní funkci (Věta 12.4.2) vztah

$$U(x,y) - U(x_0,y_0) = 0$$

na jistém okolí bodu x_0 definuje funkci $y(x)$. Skutečně,

$$F(x,y) := U(x,y) - U(x_0,y_0) \in C^1(\Omega)$$

a

$$\frac{\partial F}{\partial y}(x_0,y_0) = \frac{\partial U}{\partial y}(x_0,y_0) = N(x_0,y_0) \neq 0.$$

Podle základní verze Věty o derivaci implicitní funkce (Věta 12.4.6) dále na tomto okolí platí

$$y'(x) = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}} = -\frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}} = -\frac{M(x,y)}{N(x,y)}.$$

Odtud

$$M(x,y) + N(x,y)y' = 0.$$

Nyní dokažme jednoznačnost. Pokud y splňuje $y(x_0) = y_0$, na nějakém okolí bodu x_0 řeší rovnici

$$M(x,y) + N(x,y)y' = 0$$

a U je potenciál této rovnice, pak podle řetízkového pravidla máme

$$\frac{\mathrm{d}}{\mathrm{d}x}U(x,y(x)) = \frac{\partial U}{\partial x} + \frac{\partial U}{\partial y}\frac{\mathrm{d}y}{\mathrm{d}x} = M(x,y) + N(x,y)y' = 0.$$

Podle Věty o nejednoznačnosti primitivní funkce (Věta 4.1.4) na odpovídajícím okolí musí platit $U(x,y(x)) = U(x_0,y_0)$. Řešení se proto na dostatečně malém okolí shoduje s jednoznačnou funkcí danou základní verzí Věty o implicitní funkci (Věta 12.4.2).

Pro druhou rovnici a podmínku $M \neq 0$ na Ω je postup analogický. □

Příklad 12.5.3. Uvažme rovnici

$$5x^4y + 2x^3y^2 + (x^5 + x^4y + 2y)y' = 0, \qquad y(0) = 1.$$

Přepis do tvaru před hledáním potenciálu je

$$(5x^4y + 2x^3y^2)\,\mathrm{d}x + (x^5 + x^4y + 2y)\,\mathrm{d}y = 0.$$

V Příkladu 12.3.8 jsme nalezli potenciál

$$U(x,y) = x^5y + \frac{1}{2}x^4y^2 + y^2 + C \qquad \text{na } \mathbb{R}^2.$$

Díky tomu, že funkce $N(x,y) := x^5 + x^4y + 2y$ splňuje $N(0,1) = 2 \neq 0$, můžeme použít předchozí větu a řešení rovnice je na okolí počátku dáno vztahem

$$0 = U(x,y) - U(0,1) = x^5y + \frac{1}{2}x^4y^2 + y^2 + C - (1+C) = x^5y + \frac{1}{2}x^4y^2 + y^2 - 1$$

(povšimněte si, jak se aditivní konstanta C z nejednoznačnosti potenciálu vyrušila). U naší rovnice navíc nejsme závislí jen na základní verzi Věty o implicitní funkci (Věta 12.4.2). Z přepisu

$$\Big(\frac{1}{2}x^4 + 1\Big)y^2 + x^5y - 1 = 0$$

dostáváme

$$y(x) = \frac{-x^5 \pm \sqrt{x^{10} + 2(x^4 + 2)}}{x^4 + 2},$$

přičemž počáteční podmínka $y(0) = 1$ připouští jen jednu větev řešení

$$y(x) = \frac{-x^5 + \sqrt{x^{10} + 2(x^4 + 2)}}{x^4 + 2} \qquad \text{na } \mathbb{R}.$$

Poznamenejme ještě, že získané řešení je jednoznačné na $\mathbb{R}$. Skutečně, pokud by tomu tak nebylo, musely by v nějakém bodě na grafu výše uvedené funkce být porušeny předpoklady základní verze Věty o implicitní funkci (Věta 12.4.2). V našem případě by muselo nastat

$$0 = N(x, y(x)) = x^5 + x^4y(x) + 2y(x) \qquad \Longleftrightarrow \qquad y(x) = \frac{-x^5}{x^4 + 2}.$$

Pokud poslední podmínku porovnáme se vzorcem pro $y(x)$, vidíme, že taková situace nenastane nikdy.

Poznámka 12.5.4. K existenci a jednoznačnosti v předchozím příkladu se dá přistupovat také přes Picard–Lindelöfovu větu (Věta 10.3.5). Rovnici si přepíšeme do tvaru

$$y' = -\frac{5x^4y + 2x^3y^2}{x^5 + x^4y + 2y}.$$

Funkce dvou proměnných na pravé straně je spojitá na množině

$$\mathbb{R}^2 \setminus \Big\{(x, y) \in \mathbb{R}^2 \colon y = \frac{-x^5}{x^4 + 2}\Big\}.$$

Lokální lipschitzovskost ověříme zderivováním pravé strany podle y a opět nám vyjde podmínka $y \neq \frac{-x^5}{x^4+2}$ (není potřeba derivaci provádět, stačí si uvědomit, jak funguje vzoreček pro derivaci podílu dvou funkcí). Picard–Lindelöfovu větu (Věta 10.3.5) proto můžeme používat na množinách

$$\Omega_1 = \Big\{(x, y) \in \mathbb{R}^2 \colon x \in \mathbb{R} \wedge y < \frac{-x^5}{x^4 + 2}\Big\}$$

a

$$\Omega_2 = \Big\{(x, y) \in \mathbb{R}^2 \colon x \in \mathbb{R} \wedge y > \frac{-x^5}{x^4 + 2}\Big\}.$$

Připomeňme, že vektorové pole mít potenciál nemusí. Existence potenciálu je spíše vzácností a souvisí s podmínkou

$$\frac{\partial M}{\partial y} = \frac{\partial N}{\partial x},$$

která je v případě dostatečné hladkosti funkcí M, N a vhodného tvaru množiny Ω zároveň podmínkou nutnou i postačující.

Pokud uvedená podmínka není splněna, můžeme se ještě pokusit nalézt integrační faktor, který rovnici do vhodného tvaru převede. Zde se v praxi postupuje metodou částečného uhodnutí. Integrační faktor hledáme ve tvaru

$$\mu(x, y) = m(\Phi(x, y)),$$

kde funkci $\Phi(x, y)$ zkoušíme nastřelit a funkce m nám v případě šťastného nástřelu vyjde z Věty o nutné podmínce pro integrační faktor (Věta 12.3.5), tedy

$$\frac{\partial}{\partial y}\Big(\mu(x, y)M(x, y)\Big) = \frac{\partial}{\partial x}\Big(\mu(x, y)N(x, y)\Big).$$

Zde díky tvaru $\mu(x, y) = m(\Phi(x, y))$ máme (pro přehlednost vynecháváme argument (x, y) u funkcí M, N, Φ)

$$m'(\Phi(x, y))\frac{\partial \Phi}{\partial y}M + m(\Phi(x, y))\frac{\partial M}{\partial y} = m'(\Phi(x, y))\frac{\partial \Phi}{\partial x}N + m(\Phi(x, y))\frac{\partial N}{\partial x}.$$

K tomu stačí splnit

$$\frac{m'(\Phi(x, y))}{m(\Phi(x, y))} = \frac{\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}}{M\frac{\partial \Phi}{\partial y} - N\frac{\partial \Phi}{\partial x}} =: \Psi(x, y).$$

Pokud je možné přepsat $\Psi(x, y)$ do tvaru $H(\Phi(x, y))$ (toto je kritická část postupu, která se podaří jen málokdy), máme úlohu se separovanými proměnnými

$$\frac{m'(z)}{m(z)} = H(z)$$

a tu řeší

$$m(z) = \mathrm{e}^{\int H(z)\,\mathrm{d}z}.$$

Tím je nalezen požadovaný integrační faktor. Mezi nejčastěji používané volby patří

$$\Phi(x, y) = x, \quad \Phi(x, y) = y, \quad \Phi(x, y) = xy \quad \text{a} \quad \Phi(x, y) = x + y.$$

Aditivní konstanta při integraci v tomto případě nic zajímavého nepřináší.

Příklad 12.5.5. Uvažme úlohu

$$2xy + x^2y + \frac{y^3}{3} + (x^2 + y^2)y' = 0.$$

Nejedná se přímo o rovnici ve tvaru totálního diferenciálu, neboť

$$\frac{\partial M}{\partial y} = 2x + x^2 + y^2 \neq 2x = \frac{\partial N}{\partial x}.$$

Pokusme se hledat integrační faktor. Nejprve volíme $\Phi(x, y) = y$. Pak máme

$$\begin{aligned}\frac{m'(\Phi(x,y))}{m(\Phi(x,y))} &= \frac{\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}}{M\frac{\partial \Phi}{\partial y} - N\frac{\partial \Phi}{\partial x}} = \frac{2x - (2x + x^2 + y^2)}{(2xy + x^2y + \frac{y^3}{3}) \cdot 1 - (x^2 + y^2) \cdot 0} \\ &= \frac{-x^2 - y^2}{2xy + x^2y + \frac{y^3}{3}}.\end{aligned}$$

Výraz úplně napravo není možné zapsat jako funkci závisející pouze na y, čímž pro nás tento pokus končí neúspěchem.

Zkusme položit $\Phi(x, y) = x$. Pak

$$\frac{m'(\Phi(x,y))}{m(\Phi(x,y))} = \frac{\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}}{M\frac{\partial \Phi}{\partial y} - N\frac{\partial \Phi}{\partial x}} = \frac{2x - (2x + x^2 + y^2)}{(2xy + x^2y + \frac{y^3}{3}) \cdot 0 - (x^2 + y^2) \cdot 1} = 1.$$

To už je funkce závislá na zvolené proměnné x a odtud

$$m(x) = \mathrm{e}^{\int 1\,\mathrm{d}x} = \mathrm{e}^x.$$

Nyní přecházíme k rovnici ve tvaru totálního diferenciálu

$$\mathrm{e}^x\left(2xy + x^2y + \frac{y^3}{3}\right)\mathrm{d}x + \mathrm{e}^x(x^2 + y^2)\mathrm{d}y = 0.$$

Proto

$$U(x, y) = \int \mathrm{e}^x\left(2xy + x^2y + \frac{y^3}{3}\right)\mathrm{d}x = \mathrm{e}^x\left(x^2y + \frac{y^3}{3}\right) + \varphi(y)$$

a

$$\mathrm{e}^x(x^2 + y^2) = \frac{\partial U}{\partial y} = \mathrm{e}^x(x^2 + y^2) + \varphi'(y) \qquad \Longrightarrow \qquad \varphi'(y) = 0.$$

Celkově má naše rovnice řešení dané implicitně předpisem

$$\mathrm{e}^x\left(x^2y + \frac{y^3}{3}\right) = C,$$

kde konstantu C určíme z počáteční podmínky a tu lze volit jakkoliv kromě případu $0 = N(x_0, y_0) = x_0^2 + y_0^2$.

Příklad 12.5.6. Uvažme úlohu

$$xy^2 + (x^2y - x)y' = 0.$$

Nejedná se přímo o rovnici ve tvaru totálního diferenciálu, neboť (ověření provádíme rovnou ve tvaru, který odpovídá čitateli ze vzorečku pro hledání integračního faktoru)

$$\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} = 2xy - 1 - 2xy = -1 \neq 0.$$

Pokusme se hledat integrační faktor. Nejprve volíme $\Phi(x,y)=x$. Pak máme

$$\frac{m'(\Phi(x,y))}{m(\Phi(x,y))}=\frac{\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}}{M\frac{\partial \Phi}{\partial y}-N\frac{\partial \Phi}{\partial x}}=\frac{-1}{(xy^2)\cdot 0-(x^2y-x)\cdot 1}=\frac{1}{x^2y-x}$$

a hned vidíme, že jsme neuspěli. Nyní zkusme položit $\Phi(x,y)=y$. Pak

$$\frac{m'(\Phi(x,y))}{m(\Phi(x,y))}=\frac{\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}}{M\frac{\partial \Phi}{\partial y}-N\frac{\partial \Phi}{\partial x}}=\frac{-1}{(xy^2)\cdot 1-(x^2y-x)\cdot 0}=\frac{-1}{xy^2}$$

a opět jsme neuspěli. Zkusme dále volbu $\Phi(x,y)=x+y$. Ta dává

$$\frac{m'(\Phi(x,y))}{m(\Phi(x,y))}=\frac{\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}}{M\frac{\partial \Phi}{\partial y}-N\frac{\partial \Phi}{\partial x}}=\frac{-1}{(xy^2)\cdot 1-(x^2y-x)\cdot 1}=\frac{-1}{xy^2-x^2y+x},$$

kde se výsledek nedá zapsat jako funkce pracující pouze s $x+y$ (pokud by to šlo, dvojice $(1,0)$ a $(0,1)$ by musely dostat stejnou funkční hodnotu).

Zkusme ještě $\Phi(x,y)=xy$. Dostáváme

$$\frac{m'(\Phi(x,y))}{m(\Phi(x,y))}=\frac{\frac{\partial N}{\partial x}-\frac{\partial M}{\partial y}}{M\frac{\partial \Phi}{\partial y}-N\frac{\partial \Phi}{\partial x}}=\frac{-1}{(xy^2)\cdot x-(x^2y-x)\cdot y}=\frac{-1}{xy}.$$

Konečně jsme našli funkci závislou pouze na $\Phi(x,y)$. Dostáváme diferenciální rovnici

$$\frac{m'(z)}{m(z)}=-\frac{1}{z}$$

a tu řeší

$$m(z)=\mathrm{e}^{-\int\frac{1}{z}\,\mathrm{d}z}=\mathrm{e}^{-\log|z|}=\frac{C}{|z|}.$$

Volíme integrační faktor $\mu(x,y)=\frac{1}{xy}$ (a dále už nepracujeme na souřadných osách). Máme rovnici ve tvaru totálního diferenciálu

$$y\,\mathrm{d}x+\left(x-\frac{1}{y}\right)\mathrm{d}y=0.$$

Pro ni dosáváme

$$U(x,y)=\int y\,\mathrm{d}x=xy+\varphi(y)$$

a

$$x-\frac{1}{y}=\frac{\partial U}{\partial y}=x+\varphi'(y)\qquad\Longrightarrow\qquad\varphi'(y)=-\frac{1}{y}.$$

Celkově má naše rovnice řešení dané implicitně předpisem

$$xy-\log|y|=C,$$

kde konstantu C určíme z počáteční podmínky. Správné dořešení příkladu by požadovalo dále zkoumat možnost lepení na souřadných osách. Poznamenejme ještě, že formule $xy - \log|y| = C$ sice neumožňuje pohodlně vyjádřit y, ale máme

$$x = \frac{1}{y}(C + \log|y|).$$

Pokud by původní zadání bylo $(xy^2)x' + x^2y - x = 0$, měli bychom explicitní vzorec pro řešení. V některých aplikacích (zejména z geometrie) se skutečně stává, že nám příliš nezáleží na tom, zda řešíme úlohu $M(x,y) + N(x,y)y' = 0$ nebo $M(x,y)x' + N(x,y) = 0$.

Příklad 12.5.7. Uvažme lineární rovnici prvního řádu

$$y' + p(x)y = f(x) \qquad \Longleftrightarrow \qquad p(x)y - f(x) + y' = 0.$$

Pak $M(x,y) = p(x)y - f(x)$ a $N(x,y) \equiv 1$. Platí

$$\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y} = 0 - p(x) = -p(x).$$

Pokud je tedy $p \equiv 0$, máme rovnici ve tvaru totálního diferenciálu

$$-f(x)\,\mathrm{d}x + 1\,\mathrm{d}y = 0.$$

Potenciál získáme z formnulí (F je primitivní funkce k f)

$$U(x,y) = \int -f(x)\,\mathrm{d}x = -F(x) + \varphi(y)$$

a

$$1 = \frac{\partial U}{\partial y} = \varphi'(y).$$

Proto je řešení dáno vztahem

$$-F(x) + y = C \qquad \Longleftrightarrow \qquad y = F(x) + C.$$

Neplatí-li $p \equiv 0$, zkusme hledat integrační faktor pomocí volby $\Phi(x,y) = x$. Odtud

$$\frac{m'(\Phi(x,y))}{m(\Phi(x,y))} = \frac{\frac{\partial N}{\partial x} - \frac{\partial M}{\partial y}}{M\frac{\partial \Phi}{\partial y} - N\frac{\partial \Phi}{\partial x}} = \frac{-p(x)}{(p(x)y - f(x)) \cdot 0 - 1 \cdot 1} = p(x).$$

Dostáváme integrační faktor $\mu(x) = \mathrm{e}^{P(x)}$, kde P je primitivní funkce k p. Rovnice má po přenásobení tvar

$$\mathrm{e}^{P(x)}(p(x)y - f(x))\,\mathrm{d}x + \mathrm{e}^{P(x)}\,\mathrm{d}y = 0.$$

Potenciál získáme z formnulí

$$U(x,y) = \int \mathrm{e}^{P(x)}\,\mathrm{d}y = \mathrm{e}^{P(x)}y + \psi(x)$$

a

$$e^{P(x)}(p(x)y - f(x)) = \frac{\partial U}{\partial x} = e^{P(x)}p(x)y + \psi'(x) \qquad \Longleftrightarrow \qquad \psi'(x) = -f(x)e^{P(x)}.$$

Celkově je řešení dáno implicitně vztahem

$$e^{P(x)}y - \int f(x)e^{P(x)} = C.$$

Odtud

$$y(x) = e^{-P(x)}\Big(\int f(x)e^{P(x)} + C\Big).$$

Poznámka 12.5.8. Poslední úloha se dá interpretovat tak, že naše nová metoda je schopna vyřešit lineární rovnice prvního řádu pouhou volbou $\Phi(x, y) = x$. Protože máme k dispozici i mnoho dalších funkcí dosaditelných za $\Phi(x, y)$, dá se říci, že metody řešení rovnic ve tvaru totálního diferenciálu jsou nesrovnatelně mocnější nástroj, než je metoda integračního faktoru pro lineární rovnice prvního řádu.

12.6 Lokální extrémy funkcí více proměnných

Definice 12.6.1. Nechť $f\colon \mathbb{R}^N \to \mathbb{R}$ je definována na $M \subset \mathbb{R}^N$. Řekneme, že f má v bodě $a \in M$ *lokální maximum* vzhledem k M, jestliže existuje $\delta > 0$ takové, že

$$f(x) \leq f(a) \qquad \text{na } \mathcal{U}_\delta(a) \cap M.$$

Lokální minimum se definuje analogicky. Ostré lokální maximum a minimum definujeme pomocí ostrých nerovností a prstencových okolí.

Snadnou modifikací jednorozměrného důkazu obdržíme nutnou podmínku pro lokální extrém.

Věta 12.6.2 (Nutná podmínka pro lokální extrém)**.** *Nechť $f\colon \mathbb{R}^N \to \mathbb{R}$ je definována na $M \subset \mathbb{R}^N$, $a \in M$ je vnitřní bod množiny M a $i \in \{1, \dots, N\}$. Má-li f v bodě a lokální extrém (vzhledem k M) a existuje-li $\frac{\partial f}{\partial x_i}(a)$, pak $\frac{\partial f}{\partial x_i}(a) = 0$.*

Body splňující $\frac{\partial f}{\partial x_i}(a) = 0$ pro každé $i \in \{1, \dots, N\}$ se nazývají *stacionární body*. Už v jednorozměrném případě jsme viděli, že ve stacionárním bodě nemusí být lokální extrém (uvažte funkci $x \mapsto x^3$). Nicméně díky Taylorovu rozvoji jsme dokázáli, že stacionarita bodu spolu s vhodnou kontrolou znaménka druhé derivace extrém v tomto bodě zaručují. Obdobný výsledek se pokusíme získat i ve vícerozměrném případě.

Definice 12.6.3 (Klasifikace kvadratických forem)**.** Nechť A je symetrická matice typu $N \times N$ a $Q\colon \mathbb{R}^N \mapsto \mathbb{R}$ je jí odpovídající kvadratická forma, tedy

$$Q(h) = (Ah, h) = \sum_{i,j=1}^{N} a_{ij}h_ih_j \qquad \text{pro všechna } h \in \mathbb{R}^N.$$

Tato kvadratická forma se nazývá
- *pozitivně definitní*, jestliže $Q(h) > 0$ pro všechna $h \in \mathbb{R}^N \setminus \{0\}$
- *negativně definitní*, jestliže $Q(h) < 0$ pro všechna $h \in \mathbb{R}^N \setminus \{0\}$
- *pozitivně semidefinitní*, jestliže $Q(h) \geq 0$ pro všechna $h \in \mathbb{R}^N$
- *negativně semidefinitní*, jestliže $Q(h) \leq 0$ pro všechna $h \in \mathbb{R}^N$
- *indefinitní*, jestliže existují $h, l \in \mathbb{R}^N$ taková, že $Q(h) < 0 < Q(l)$.

Připomeňme si metody určování definitnosti kvadratických forem. První metodou je diagonalizace procvičovaná v lineární algebře. Druhou metodou je převod na čtverec.

Příklad 12.6.4. Nechť

$$Q(h) = h_1^2 + 4h_2^2 + 7h_3^2 + 4h_1h_2 + 2h_1h_3 + 16h_2h_3.$$

Tento předpis postupně upravujeme

$$\begin{aligned} Q(h) &= (h_1 + 2h_2 + h_3)^2 + 6h_3^2 + 12h_2h_3 \\ &= (h_1 + 2h_2 + h_3)^2 + (\sqrt{6}h_2 + \sqrt{6}h_3)^2 - 6h_2^2. \end{aligned}$$

Kvadratická forma je indefinitní, neboť máme

$$Q(-1, 1, -1) = 0 + 0 - 6 \qquad \text{a} \qquad Q(1, 0, 0) = 1 + 0 + 0.$$

Třetím nástrojem je Sylvesterovo kritérium, podle něhož je Q pozitivně definitní právě tehdy, když všechny hlavní subdeterminanty matice A jsou kladné. Přechodem k matici, jejíž prvky mají obrácené znaménko, dostáváme, že Q je negativně definitní právě tehdy, když znaménko hlavních subdeterminantů je $(-1)^k$, kde k je pořadí subdeterminantu, neboli počet řádků hlavní submatice, z níž právě počítáme determinant.

Příklad 12.6.5. Nechť

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -4 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Pak

$$D_1 := \det(-1) = -1 < 0, \qquad D_2 := \det \begin{pmatrix} -1 & 1 \\ 1 & -4 \end{pmatrix} = 4 - 1 = 3 > 0$$

a

$$D_3 := \det \begin{pmatrix} -1 & 1 & 0 \\ 1 & -4 & 0 \\ 0 & 0 & -1 \end{pmatrix} = -4 + 0 + 0 - (-1) - 0 - 0 = -3 < 0.$$

Kvadratická forma je proto negativně definitní. Pokud bychom naopak pracovali s maticí $-A$, měli bychom

$$D_1 = 1 > 0, \qquad D_2 = 3 > 0, \qquad D_3 = 3 > 0$$

a pozitivně definitní kvadratickou formu.

Občas (zejména v nízké dimenzi) se hodí charakterizace definitnosti pomocí znamének vlastních čísel a skutečnosti, že determinant se rovná součinu vlastních čísel.

Příklad 12.6.6. Nechť
$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$
Pak $\det A = -3$. Jedno vlastní číslo je proto kladné a druhé záporné. Odpovídající kvadratická forma je proto indefinitní.

Pro naše záměry se bude hodit ještě jedna charakterizace pozitivní definitnosti.

Lemma 12.6.7. *Kvadratická forma Q je pozitivně definitní právě tehdy, když existuje $\alpha > 0$ splňující $Q(h) \geq \alpha\|h\|^2$ pro všechna $h \in \mathbb{R}^N$.*

Důkaz. Implikace „$\Leftarrow$“ je zřejmá. Dokažme „$\Rightarrow$“. Definujme
$$\alpha := \inf\{Q(h) \colon \|h\| = 1\}.$$
Díky pozitivní definitnosti počítáme infimum ze samých kladných čísel, proto je $\alpha \geq 0$. Navíc Q je spojitá funkce a ta na jednotkové sféře (kompaktní množina) nabývá svého minima. Proto je $\alpha > 0$. Nyní pro libovolné $h \in \mathbb{R}^N \setminus 0$ máme
$$Q(h) = Q\Big(\|h\|\frac{h}{\|h\|}\Big) = \|h\|^2 Q\Big(\frac{h}{\|h\|}\Big) \geq \alpha\|h\|^2.$$
Pro $h = 0$ je nerovnost zřejmě splněna. □

Poznámka 12.6.8. Negativní definitnost se dá obdobně charakterizovat pomocí podmínky $Q(h) \leq -\alpha\|h\|^2$.

Věta 12.6.9 (Postačující podmínka pro lokální extrém)**.** *Nechť $f \colon \mathbb{R}^N \to \mathbb{R}$, $a \in \mathbb{R}^N$ je stacionární bod funkce f a existuje $\delta > 0$ takové, že $f \in C^3(\mathcal{U}_\delta(a))$. Definujme kvadratickou formu $Q \colon \mathbb{R}^N \to \mathbb{R}$ předpisem*
$$Q(h) := \mathrm{d}^2 f(a)(h) = \sum_{i,j=1}^{N} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j.$$
Je-li Q pozitivně definitní, f má v bodě a ostré lokální minimum.
Je-li Q negativně definitní, f má v bodě a ostré lokální maximum.
Je-li Q indefinitní, f nemá v bodě a lokální extrém.

Důkaz. Podle Taylorova vzorce na $\mathcal{U}_\delta(a)$ platí
$$\begin{aligned} f(a+h) = f(a) &+ \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(a) h_i + \frac{1}{2} \sum_{i,j=1}^{N} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j \\ &+ \frac{1}{6} \sum_{i,j,k=1}^{N} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(a+\theta h) h_i h_j h_k \\ = f(a) &+ 0 + \frac{1}{2} Q(h) + \frac{1}{6} \sum_{i,j,k=1}^{N} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(a+\theta h) h_i h_j h_k. \end{aligned}$$

Je-li nyní Q pozitivně definitní, pak podle předchozího lemmatu pro $\|h\|$ dostatečně malé máme

$$\begin{aligned} f(a+h) &= f(a) + \frac{1}{2}Q(h) + \frac{1}{6}\sum_{i,j,k=1}^{N} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(a+\theta h)h_i h_j h_k \\ &\geq f(a) + \frac{1}{2}\alpha\|h\|^2 - \sum_{i,j,k=1}^{N} \max_{\overline{\mathcal{U}}_{\frac{\delta}{2}}(a)} \Big|\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}\Big| |h_i||h_j||h_k| \\ &\geq f(a) + \frac{1}{2}\alpha\|h\|^2 - C\|h\|^3 = f(a) + \Big(\frac{1}{2}\alpha - C\|h\|\Big)\|h\|^2 \\ &\geq f(a) + \frac{1}{4}\alpha\|h\|^2. \end{aligned}$$

Analogicky postupujeme v případě negativně definitní formy. Je-li Q indefinitní, pak existují $h, l \in \mathbb{R}^N$ taková, že $Q(h) < 0 < Q(l)$. Pro $t \in (0,1)$ dostatečně malé pak máme ($\theta \in (0,1)$ závisí na t i h)

$$\begin{aligned} f(a+th) &= f(a) + \frac{1}{2}Q(th) + \frac{1}{6}\sum_{i,j,k=1}^{N} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(a+\theta th)t^3 h_i h_j h_k \\ &\leq f(a) + \frac{1}{2}t^2 Q(h) + Ct^3 < f(a). \end{aligned}$$

Podobně $f(a+tl) > f(a)$ pro $t \in (0,1)$ dostatečně malé. □

Poznámka 12.6.10. (i) Stacionární bod, v němž je Q indefinitní, se nazývá *sedlový bod*. Typickým případem je počátek pro funkci $f(x,y) = x^2 - y^2$.
(ii) Matici druhých parciálních derivací se říká *Hessova matice*. Značíme ji $H_f(x)$.
(iii) Pokud je ve stacionárním bodě Q netriviální a pozitivně semidefinitní, pomocí konce předchozího důkazu nahlédneme, že v tomto bodě nemůže být lokální maximum. Nicméně lokální minimum v něm být může a nemusí. Uvažte počátek a funkce $f(x,y) = x^2 \pm y^4$. Podobně pro pro netriviální negativně semidefinitní formu.
(iv) Pokud je kvadratická forma příslušející druhému diferenciálu pozitivně semidefinitní na nějakém okolí stacionárního bodu, v bodě je (obecně neostré) lokální minimum. To nám dá následující modifikace finálního výpočtu z předchozího důkazu

$$\begin{aligned} f(a+h) &= f(a) + \sum_{i=1}^{N} \frac{\partial f}{\partial x_i}(a)h_i + \frac{1}{2}\sum_{i,j=1}^{N} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+\theta h)h_i h_j \\ &= f(a) + \frac{1}{2}\sum_{i,j=1}^{N} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+\theta h)h_i h_j \geq f(a). \end{aligned}$$

Podobně pro negativně definitní kvadratickou formu na okolí stacionárního bodu.
(v) V případě, kdy nám předchozí věta poskytuje neplnohodnotnou nebo žádnou informaci, nezbývá, než se pokusit použít elementární prostředky.

Příklad 12.6.11. Zkoumejme lokální extrémy funkce

$$f(x, y, z) = x^3 + y^2 + z^2 + 12xy + 2z$$

na $\mathbb{R}^3$. Jedná se o $C^\infty(\mathbb{R}^3)$-funkci, jejíž gradient a Hessova matice mají tvar

$$\nabla f(x, y, z) = (3x^2 + 12y, 2y + 12x, 2z + 2) \quad \text{a} \quad H_f(x, y, z) = \begin{pmatrix} 6x & 12 & 0 \\ 12 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Z tvaru gradientu snadno nalezneme stacionární body

$$a = (0, 0, -1) \qquad \text{a} \qquad b = (24, -144, -1).$$

V bodě b se pokusíme aplikovat Sylvesterovo pravidlo. Máme

$$D_1 = 6 \cdot 24 > 0, \quad D_2 = 6 \cdot 24 \cdot 2 - 12^2 > 0 \quad \text{a} \quad D_3 = 6 \cdot 24 \cdot 2 \cdot 2 - 2 \cdot 12^2 > 0.$$

Hessova matice je zde pozitivně definitní. Proto f má v bodě b ostré lokální minimum. V bodě a máme

$$D_1 = 0.$$

Odtud vidíme, že Hessova matice už nemůže být ani pozitivně, ani negativně definitní. Navíc

$$D_3 = -2 \cdot 12^2 \neq 0.$$

Všechna vlastní čísla jsou proto nenulová, a proto je Hessova matice indefinitní a v bodě a je sedlový bod.

Příklad 12.6.12. Zkoumejme lokální extrémy funkce

$$f(x, y) = x^2 + y^4 - y^3$$

na $\mathbb{R}^2$. Jedná se o $C^\infty(\mathbb{R}^2)$-funkci, jejíž gradient a Hessova matice mají tvar

$$\nabla f(x, y) = (2x, 4y^3 - 3y^2) \quad \text{a} \quad H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 12y^2 - 6y \end{pmatrix}.$$

Z tvaru gradientu snadno nalezneme stacionární body $a = (0, 0)$ a $b = (0, \frac{3}{4})$. V bodě b máme

$$H_f(0, \tfrac{3}{4}) = \begin{pmatrix} 2 & 0 \\ 0 & \frac{27}{4} - \frac{18}{4} \end{pmatrix}.$$

Hessova matice je zde pozitivně definitní. Proto f má v bodě b ostré lokální minimum. V bodě a máme

$$H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Hessova matice je netriviální a pozitivně semidefinitní. V počátku proto nemůže být lokální maximum. Pokud si však povšimneme chování funkce f na y-ové ose, tedy

$$f(0, y) = y^4 - y^3 = y^3(y - 1)$$

(pro $y \in (-1, 0)$ máme kladné funkční hodnoty, pro $y \in (0, 1)$ záporné), okamžitě dostáváme, že v počátku není lokální extrém.

Příklad 12.6.13. Zkoumejme lokální extrémy funkce

$$f(x,y) = x^2 + y^4 - y^5$$

na $\mathbb{R}^2$. Jedná se o $C^\infty(\mathbb{R}^2)$-funkci, jejíž gradient a Hessova matice mají tvar

$$\nabla f(x,y) = (2x, 4y^3 - 5y^4) \quad \text{a} \quad H_f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 12y^2 - 20y^3 \end{pmatrix}.$$

Z tvaru gradientu snadno nalezneme stacionární body $a = (0,0)$ a $b = (0, \frac{4}{5})$. V bodě b máme

$$H_f(0, \tfrac{4}{5}) = \begin{pmatrix} 2 & 0 \\ 0 & \frac{192}{25} - \frac{256}{25} \end{pmatrix}.$$

Hessova matice je zde indefinitní. Proto f má v bodě b sedlo.

V bodě a máme

$$H_f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Hessova matice je netriviální a pozitivně semidefinitní. V počátku proto nemůže být lokální maximum. Na druhou stranu, máme

$$f(x,y) = x^2 + y^4 - y^5 = x^2 + y^4(1-y).$$

Odtud $f > 0$ na $(-1,1)^2 \setminus \{0,0\}$, a proto je v počátku ostré lokální minimum.

12.7 Globální extrémy funkcí více proměnných

Základní existenční nástroj pro existenci globálních extrémů spojitých funkcí jsme už měli a to sice existenci extrémů spojitých funkcí na kompaktních množinách. Výsledek se dá lehce rozšířit, využijeme-li ještě základní vlastnosti limity funkce.

Tvrzení 12.7.1. *Nechť $f\colon \mathbb{R}^N \to \mathbb{R}$ je spojitá na $M \subset \mathbb{R}^m$.*
(i) *Je-li M omezená a uzavřená, f zde nabývá svého maxima a minima.*
(ii) *Je-li $M = \mathbb{R}^N$ a $\lim_{\|x\|\to\infty} f(x) = \infty$, f zde nabývá svého minima. Podobně pro maximum.*
(iii) *Je-li $M = \mathbb{R}^N$, $\lim_{\|x\|\to\infty} f(x) = 0$ a existuje bod, v němž má f zápornou hodnotu, pak na M nabývá svého minima. Podobně pro maximum.*

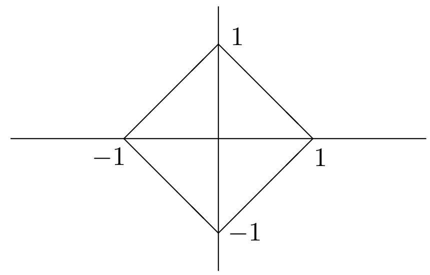
Důkaz. První výsledek již známe, neboť omezenost a uzavřenost v konečnědimenzionálním prostoru implikují kompaktnost. Dokažme druhý výsledek. Z definice limity existuje $R > 0$ takové, že

$$f > f(0) \qquad \text{na } \mathbb{R}^N \setminus \mathcal{U}_R(0).$$

Navíc $\overline{\mathcal{U}}_R(0)$ je kompaktní, proto zde f nabývá svého minima a platí

$$\min_M f = \min_{\overline{\mathcal{U}}_R(0)} f \le f(0).$$

Z toho již plyne požadovaný výsledek. Ostatní tvrzení dokážeme pomocí podobné myšlenky. $\square$



Obrázek 12.3: Ilustrace množiny $M = \{(x, y) \in \mathbb{R}^2 \colon |x| + |y| \leq 1\}$.

Příklad 12.7.2. Uvažme funkci $f(x, y) = x^2 + y^2$ na množině $M = \{(x, y) \in \mathbb{R}^2 \colon |x| + |y| \leq 1\}$.

Jedná se o spojitou funkci na kompaktní množině, proto zde musí nabývat svého maxima a minima, které se nyní pokusíme najít. Předně každý globální extrém je i lokálním extrémem, proto můžeme použít nutnou podmínku pro lokální extrém pro vyloučení velkého počtu bodů. Skutečně, funkce f splňuje

$$\nabla f(x, y) = (2x, 2y).$$

Proto ve vnitřku množiny M s výjimkou počátku nikde globální extrém být nemůže. Hranice množiny M je tvořená čtyřmi úsečkami. Na úsečce v prvním kvadrantu se dají funkční hodnoty popsat pomocí

$$\varphi_1(x) := f(x, 1 - x) = x^2 + (1 - x)^2 = 2x^2 - 2x + 1.$$

Je jasné, že pokud funkce φ_1 nemá lokální extrém v nějakém bodě $x \in (0, 1)$, pak ani funkce f nemá lokální extrém (vůči M) v bodě $(x, 1 - x)$. Protože

$$\varphi_1'(x) = 4x - 2 \qquad \text{na } (0, 1),$$

funkce f nemá globální extrém v žádném bodě tvaru $(x, 1-x)$, kde $x \in (0, 1) \setminus \{\frac{1}{2}\}$.

Ve druhém kvadrantu podobně používáme funkci

$$\varphi_2(x) := f(x, 1 + x) = 2x^2 + 2x + 1 \qquad \text{na } (-1, 0).$$

Pro její derivaci máme

$$\varphi_2'(x) = 4x + 2 \qquad \text{na } (-1, 0).$$

To má za následek, že ani v bodech tvaru $(x, 1 + x)$, kde $x \in (-1, 0) \setminus \{-\frac{1}{2}\}$, není globální extrém.

Ve třetím kvadrantu používáme funkci

$$\varphi_3(x) := f(x, -1 - x) = 2x^2 + 2x + 1 \qquad \text{na } (-1, 0).$$

Pro její derivaci máme

$$\varphi_3'(x) = 4x + 2 \qquad \text{na } (-1, 0).$$

Proto ani v bodech tvaru $(x, -1-x)$, kde $x \in (-1, 0) \setminus \{-\frac{1}{2}\}$, není globální extrém.

Konečně, ve čtvrtém kvadrantu používáme funkci

$$\varphi_4(x) := f(x, x-1) = 2x^2 - 2x + 1 \qquad \text{na } (0, 1).$$

Pro její derivaci máme

$$\varphi_4'(x) = 4x - 2 \qquad \text{na } (0, 1).$$

Tedy ani v bodech tvaru $(x, x-1)$, kde $x \in (0, 1) \setminus \{\frac{1}{2}\}$, není globální extrém. Shrneme-li dosavadní výsledky, pak víme, že funkce f svých globálních extrémů nabývá, ale nemůže to být v žádném bodě s výjimkou bodů $(0, 0)$, $(1, 0)$, $(0, 1)$, $(-1, 0)$, $(0, -1)$ a $(-\frac{1}{2}, -\frac{1}{2})$, $(\frac{1}{2}, -\frac{1}{2})$, $(-\frac{1}{2}, \frac{1}{2})$, $(\frac{1}{2}, \frac{1}{2})$. Do těchto bodů stačí již jen dosadit

$$f(0,0) = 0, \qquad f(1,0) = f(0,1) = f(-1,0) = f(0,-1) = 1,$$
$$f(-\tfrac{1}{2}, -\tfrac{1}{2}) = f(\tfrac{1}{2}, -\tfrac{1}{2}) = f(-\tfrac{1}{2}, \tfrac{1}{2}) = f(\tfrac{1}{2}, \tfrac{1}{2}) = \tfrac{1}{2}$$

a již víme vše, co potřebujeme.

Poznámka 12.7.3. Samozřejmě šlo využít symetrie a uvažovat pouze první kvadrant. Pak ovšem musíme navíc uvažovat uměle vzniklou hranici.

Poznámka 12.7.4. Při právě použité metodě nevadí, pokud nebudeme příliš důsledni při vylučování bodů, kde extrém být nemůže. Poslední fáze (dosazování) si s takovými body poradí.

Ve vyšší dimenzi může být poměrně komplikovaný rozklad na množiny, kde ověřujeme nutnou podmínku lokálního extrému pro funkci vyšetřovanou pomocí funkcí pomocných. Někdy bývá výhodný popis takových množin pomocí výroků. Dále si musíme kompaktní množinu někdy uměle vyrobit a pak ji porovnat s množinou původní.

Příklad 12.7.5. Vyšetřujme extrémy funkce $F(x, y, z) = \frac{x+y}{1+z\sqrt{z}}$ na množině

$$M = \{(x, y, z) \in \mathbb{R}^3 \colon z \geq x^2 \wedge z \geq |y|\}.$$

Nejprve si povšimněme toho, že

$$(\pm 1, 0, 1) \in M \qquad \text{a} \qquad f(\pm 1, 0, 1) = \pm \frac{1}{2},$$

a pro $(x, y, z) \in M$ takové, že $z \geq 25$ platí

$$\begin{aligned} |f(x,y,z)| &\leq \left| \frac{x}{1 + z\sqrt{z}} \right| + \left| \frac{y}{1 + z\sqrt{z}} \right| \leq \frac{\sqrt{z}}{1 + z\sqrt{z}} + \frac{z}{1 + z\sqrt{z}} \\ &\leq \frac{1}{z} + \frac{1}{\sqrt{z}} \leq \frac{6}{25} < \frac{1}{2}. \end{aligned}$$

Proto na kompaktní množině (omezená množina, která je průnikem tří uzavřených množin)

$$N := \{(x, y, z) \in \mathbb{R}^3 \colon z \geq x^2 \wedge z \geq |y| \wedge z \leq 25\}$$

funkce f nabývá svého maxima a minima a platí

$$\min_N f = \min_M f \qquad \text{a} \qquad \max_N f = \max_M f.$$

Stačí tedy vyšetřovat extrémy na množině N. Množinu N si rozdělíme na deset kusů

$$\begin{aligned}
N_1 &= \{z \in (0, 25) \wedge x \in (-\sqrt{z}, \sqrt{z}) \wedge y \in (-z, z)\} \\
N_2 &= \{z \in (0, 25) \wedge x \in (-\sqrt{z}, \sqrt{z}) \wedge y = -z\} \\
N_3 &= \{z \in (0, 25) \wedge x \in (-\sqrt{z}, \sqrt{z}) \wedge y = z\} \\
N_4 &= \{z \in (0, 25) \wedge x = -\sqrt{z} \wedge y \in (-z, z)\} \\
N_5 &= \{z \in (0, 25) \wedge x = \sqrt{z} \wedge y \in (-z, z)\} \\
N_6 &= \{z \in (0, 25) \wedge x = -\sqrt{z} \wedge y = -z\} \\
N_7 &= \{z \in (0, 25) \wedge x = -\sqrt{z} \wedge y = z\} \\
N_8 &= \{z \in (0, 25) \wedge x = \sqrt{z} \wedge y = -z\} \\
N_9 &= \{z \in (0, 25) \wedge x = \sqrt{z} \wedge y = z\} \\
N_{10} &= \{(0, 0, 0)\} \cup (N \cap \{z = 25\})\}.
\end{aligned}$$

Podle předchozích výpočtů a podle $f(0, 0, 0) = 0$ hledané extrémy nemohou být na množině N_{10}. Na množině N_1 (otevřená množina) máme

$$\frac{\partial f}{\partial x} = \frac{1}{1 + z\sqrt{z}},$$

proto extrémy nejsou ani zde. Pro studium chování funkce f na N_2 použijeme pomocnou funkci

$$\varphi_2(x, z) := \frac{x - z}{1 + z\sqrt{z}}$$

definovanou na otevřené množině $\{z \in (0, 25) \wedge x \in (-\sqrt{z}, \sqrt{z})\}$. Díky tomu, že $\frac{\partial \varphi_2}{\partial x} = \frac{1}{1+z\sqrt{z}}$, hledané extrémy nejsou ani zde. Podobně pro množinu N_3.

V případě množiny N_4 je rozhodující chování

$$\varphi_4(y, z) := \frac{-\sqrt{z} + y}{1 + z\sqrt{z}}$$

na otevřené množině $\{z \in (0, 25) \wedge y \in (-z, z)\}$. Díky tomu, že $\frac{\partial \varphi_4}{\partial y} = \frac{1}{1+z\sqrt{z}}$, ani zde hledané extrémy nejsou. Podobně pro množinu N_5.

Přistupme k množinám N_6, N_7, N_8, N_9. Zde zkoumáme funkce typu

$$\frac{\pm\sqrt{z} \pm z}{1 + z^2} \qquad \text{pro } z \in (0, 25).$$

Z dosavadních výpočtů a právě uvedených formulí platí

$$-\min_M f = \max_M f = \max_{z\in(0,25)} \frac{\sqrt{z}+z}{1+z\sqrt{z}} = \max_{t\in(0,5)} \frac{t+t^2}{1+t^3} = \max_{t\in(0,5)} \frac{t}{t^2-t+1} = 1.$$

Úplně napravo jsme použili

$$\Big(\frac{t}{t^2-t+1}\Big)' = \frac{t^2-t+1-t(2t-1)}{(t^2-t+1)^2} = \frac{1-t^2}{(t^2-t+1)^2}.$$

Při popisu chování funkce na hranici se také mohou hodit polární nebo sférické souřadnice.

Příklad 12.7.6. Zkoumejme extrémy funkce $f(x,y) = 4x+3y$ na uzavřeném jednotkovém kruhu. Extrémů se opět nabývá. Na vnitřku kruhu platí

$$\frac{\partial f}{\partial x} = 3,$$

extrémy proto musí být na hranici. Tu popišme pomocí polárních souřadnic

$$x = \cos t \quad \text{a} \quad y = \sin t \qquad \text{pro } t \in [0, 2\pi).$$

pro funkci $\varphi(t) = f(\cos t, \sin t)$ máme díky Cauchy–Schwarzově nerovnosti pro skalární součin

$$|f(t)| = |4\cos t + 3\sin t| = |(4,3)\cdot(\cos t, \sin t)| \le \|(4,3)\|\|(\cos t, \sin t)\| = 5\cdot 1 = 5,$$

přičemž hodnoty na pravé straně je dosaženo, jsou-li vektory ve skalárním součinu rovnoběžné. Odtud

$$\min_M f = -5 \qquad \text{a} \qquad \max_M f = 5.$$

V některých případech může být výhodnější (nebo jedině možné) popsat hranici množiny implicitně. Tehdy je třeba použít níže uvedený nástroj. Ten může být někdy výhodný i za situací, kdy lze vyjádřit hranici explicitně, ale příslušné vyjádření je příliš komplikované.

Věta 12.7.7 (O Lagrangeových multiplikátorech)**.** *Nechť $\Omega \subset \mathbb{R}^N$ je otevřená množina, $m \in \mathbb{N}$, $m < N$ a $f, g_1, \dots, g_m \in C^1(\Omega)$. Označme*

$$M = \{x \in \Omega \colon g_i(x) = 0 \ \textit{pro všechna} \ i \in \{1,\dots,m\}\}.$$

Nechť matice $\{\frac{\partial g_i}{\partial x_j}\}_{i=1,j=1}^{m,N}$ má všude na M hodnost rovnu m. Jestliže f má v $a \in M$ lokální extrém vzhledem k M, pak existují čísla $\lambda_1,\dots,\lambda_m \in \mathbb{R}$ taková, že

$$\nabla f(a) = \sum_{i=1}^m \lambda_i \nabla g_i(a).$$

Důkaz. Označme $s = N - m$. Díky předpokladu o hodnosti matice popsané ve větě lze v matici $\{\frac{\partial g_i}{\partial x_j}(a)\}_{i=1,j=1}^{m,N}$ najít s takových sloupců, že jejich vynecháním získáme regulární čtvercovou matici. Bez újmy na obecnosti předpokládejme, že jsme vynechali prvních s sloupců. Aplikací plné verze Věty o implicitní funkci (Věty 12.4.13) dostáváme $\delta, \Delta > 0$ taková, že pro každý bod $(x_1, \dots, x_s) \in \mathcal{U}_\delta((a_1, \dots, a_s))$ existuje právě jeden bod $(x_{s+1}, \dots, x_N) \in \mathcal{U}_\Delta((a_{s+1}, \dots, a_N))$, který budeme v dalším značit $\varphi(x_1, \dots, x_s)$, splňující pro všechna $i \in \{1, \dots, m\}$

$$0 = g_i(x_1, \dots, x_s, \varphi_1(x_1, \dots, x_s), \dots, \varphi_m(x_1, \dots, x_s)) := \eta_i(x_1, \dots, x_s).$$

Navíc $\varphi \in C^1(\mathcal{U}_\delta((a_1, \dots, a_s)); \mathbb{R}^m)$. Definujme $\psi\colon \mathbb{R}^s \to \mathbb{R}$ na $\mathcal{U}_\delta((a_1, \dots, a_s))$ předpisem

$$\psi(x_1, \dots, x_s) = f(x_1, \dots, x_s, \varphi_1(x_1, \dots, x_s), \dots, \varphi_m(x_1, \dots, x_s)).$$

Pak podle řetízkového pravidla $\psi \in C^1(\mathcal{U}_\delta((a_1, \dots, a_s)))$ a podle předpokladu věty má v bodě $(a_1, \dots, a_s)$ lokální extrém. Proto pro všechna $j \in \{1, \dots, s\}$ platí podle řetízkového pravidla

$$\begin{aligned} 0 &= \frac{\partial \psi}{\partial x_j}(a_1 \dots, a_s) \\ &= \frac{\partial f}{\partial x_j}(a_1 \dots, a_N) + \sum_{k=1}^{m} \frac{\partial f}{\partial x_{s+k}}(a_1 \dots, a_N) \frac{\partial \varphi_k}{\partial x_j}(a_1 \dots, a_s). \end{aligned} \tag{12.7.1}$$

Navíc platí pro všechna $j \in \{1, \dots, s\}$ a všechna $i \in \{1, \dots, m\}$

$$\begin{aligned} 0 &= \frac{\partial 0}{\partial x_j}(a_1 \dots, a_s) = \frac{\partial \eta_i}{\partial x_j}(a_1 \dots, a_s) \\ &= \frac{\partial g_i}{\partial x_j}(a_1 \dots, a_N) + \sum_{k=1}^{m} \frac{\partial g_i}{\partial x_{s+k}}(a_1 \dots, a_N) \frac{\partial \varphi_k}{\partial x_j}(a_1 \dots, a_s). \end{aligned} \tag{12.7.2}$$

Definujme nyní lineárně nezávislé vektory v_j, $j \in \{1, \dots, s\}$, předpisem

$$v_j = \Big(\delta_{1j}, \delta_{2j}, \dots, \delta_{sj}, \frac{\partial \varphi_1}{\partial x_j}(a_1 \dots, a_s), \dots, \frac{\partial \varphi_m}{\partial x_j}(a_1 \dots, a_s)\Big).$$

Definujme ještě

$$H = \operatorname{span}\{v_1, \dots, v_s\}$$

a nechť $H^\perp$ je ortogonální doplněk H v $\mathbb{R}^N$. Podle (12.7.1) máme

$$\nabla f(a) \cdot v_j = 0 \qquad \text{pro každé } j \in \{1, \dots, s\}$$

a podle (12.7.2) zase

$$\nabla g_i(a) \cdot v_j = 0 \qquad \text{pro každé } j \in \{1, \dots, s\} \text{ a } i \in \{1, \dots, m\}.$$

Odtud $\nabla f(a) \in M^\perp$ a $\nabla g_i(a) \in M^\perp$ pro každé $\{1, \dots, m\}$. Dokonce musí platit

$$\operatorname{span}\{\nabla g_1(a), \dots, \nabla g_m(a)\} = M^\perp,$$

neboť podle předpokladu o matici je nalevo m-dimenzionální prostor a prostor napravo má dimenzi $N - s = m$ (podívejte se na definici H a $H^{\perp}$). Proto musí být $\nabla f(a)$ lineární kombinací $\nabla g_i(a)$, $i \in \{1, \ldots, m\}$. $\square$

Poznámka 12.7.8. (i) Čísla $\lambda_1, \ldots, \lambda_m$ se nazývají *Lagrangeovy multiplikátory.*
(ii) Jednotlivé funkce $g_1, \ldots, g_m$ se nazývají *vazby.* Odpovídajícím extrémům se říká *vázané extrémy.*
(iii) Přestože se znění věty týká lokálních extrémů, tato věta se v praxi takřka výhradně používá na hledání extrémů globálních. Na druhou stranu, existují obsáhlejší verze věty umožňující vytvořit kvadratickou formu, jejíž definitnost určuje typ lokálního extrému (zde je potřeba postupovat velmi opatrně, musí dojít ke snížení dimenze za pomoci plné verze Věty o implicitní funkci (Věta 12.4.13).
(iv) Předchozí věta se v praxi používá podobně jako nutná podmínka k vyloučení bodů, v nichž lokální extrém být nemůže. Tentokrát nám ale zbudou dva druhy bodů. Jednak to jsou body, které porušují podmínku o hodnosti matice, dále body, které splňují vazební podmínky a existují pro ně Lagrangeovy multiplikátory. Ve druhém případě máme soustavu

$$\begin{aligned} g_i(a) &= 0 && \text{pro } i \in \{1, \ldots, m\} \\ \frac{\partial f}{\partial x_j}(a) &= \sum_{i=1}^{m} \lambda_i \frac{\partial g_i}{\partial x_j}(a) && \text{pro } j \in \{1, \ldots, N\}. \end{aligned}$$

Celkově v tomto případě máme $m + N$ (obecně nelineárních) rovnic pro $N + m$ neznámých $a_1, \ldots, a_N, \lambda_1, \ldots, \lambda_m$. Hodnotu konstant $\lambda_1, \ldots, \lambda_m$ znát nepotřebujeme.
(v) Výše uvedený postup je také možno interpretovat tak, že si definujeme pomocnou funkci

$$L(x_1, \ldots, x_N) := f(x) - \sum_{i=1}^{m} \lambda_i g_i(x)$$

a zkoumáme, které body z množiny M mohou být stacionárními body funkce L alespoň pro jednu sadu parametrů $\lambda_1, \ldots, \lambda_m$. Ekvivalentně můžeme definovat dokonce

$$L(x_1, \ldots, x_N, \lambda_1, \ldots, \lambda_m) := f(x) - \sum_{i=1}^{m} \lambda_i g_i(x)$$

a zkoumat stacionární body této funkce $N + m$ proměnných.

Nyní si novou metodu ukážeme na příkladech. Čtenář si jistě všimne, že všechny níže uvedené příklady se dají řešit rychleji bez použití Lagrangeových multiplikátorů (třeba pomocí polárních souřadnic). Upřímní autoři zde musí přiznat, že neznají jediný příklad, u kterého by měl výpočet za použití Lagrangeových multiplikátorů rozumnou délku, a zároveň příklad nešel řešit elementárnějším a rychlejším způsobem.

Příklad 12.7.9. Hledejme globální extrémy funkce $f(x, y) = x^3 + y^3$ na množině $M = \{x^2 + y^2 \leq 1\}$. Máme spojitou funkci na kompaktu, proto se extrémů nabývá.

Na vnitřku množiny M použijeme nutnou podmínku, která nám spolu s

$$\nabla f = (3x^2, 3y^2)$$

dává počátek jako jediný bod podezřelý z extrému. V tomto bodě však extrém být nemůže, protože se v každém jeho okolí nalézají jak body s kladnou funkční hodnotou, tak body se zápornou funkční hodnotou, zatímco v počátku je funkční hodnota nulová. Na hranici množiny M použijeme metodu Lagrangeových multiplikátorů. Definujme

$$g(x, y) = x^2 + y^2 - 1.$$

Předně $\nabla g = (2x, 2y)$. Proto je podmínka o hodnosti matice porušena pouze v počátku. Ten ale nesplňuje vazební podmínku. Nyní položme

$$L(x, y) = x^3 + y^3 - \lambda(x^2 + y^2 - 1).$$

Pak

$$\begin{aligned} \frac{\partial L}{\partial x} &= 3x^2 - 2\lambda x \\ \frac{\partial L}{\partial y} &= 3y^2 - 2\lambda y. \end{aligned}$$

Celkově získáváme soustavu

$$\begin{aligned} 0 &= 3x^2 - 2\lambda x \\ 0 &= 3y^2 - 2\lambda y \\ 1 &= x^2 + y^2. \end{aligned}$$

Nyní máme tři možnosti. Pokud $x = 0$, dostáváme podezřelé body $(0, \pm 1)$. Pokud $y = 0$, dostáváme podezřelé body $(\pm 1, 0)$. Konečně, pokud $x \neq 0$ a $y \neq 0$, máme

$$3x - 2\lambda = 0 = 3y - 2\lambda \qquad \Longrightarrow \qquad x = y \qquad \Longrightarrow \qquad (x, y) = \left(\pm\tfrac{1}{\sqrt{2}}, \pm\tfrac{1}{\sqrt{2}}\right).$$

V ostatních bodech být extrémy nemohou. Nyní stačí dosadit

$$f(0, \pm 1) = \pm 1, \quad f(\pm 1, 0) = \pm 1, \quad f\left(-\tfrac{1}{\sqrt{2}}, -\tfrac{1}{\sqrt{2}}\right) = -\tfrac{1}{\sqrt{2}}, \quad f\left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}\right) = \tfrac{1}{\sqrt{2}}$$

a víme vše.

Příklad 12.7.10. Hledejme globální extrémy funkce $f(x, y) = xy + xz$ na množině $M = \{x^2 + y^2 \leq 1 \wedge x^2 + z^2 \leq 1\}$. Máme spojitou funkci na kompaktu, proto se extrémů nabývá. Předně si povšimněme toho, že globální minimum musí být záporné a globální maximum kladné (uvažte $f(\pm\frac{1}{10}, \frac{1}{10}, \frac{1}{10})$). Na vnitřku množiny M použijeme nutnou podmínku a díky

$$\nabla f = (y + z, x, x)$$

dostáváme, že pokud by se uvnitř množiny M v nějakém bodě nabývalo extrému, muselo by pro něj platit $x = 0$, čemuž ale odpovídá nulová funkční hodnota, která však nemůže být extrémem, jak jsme zjistili výše. Hramici množiny M si rozdělíme na tři podmnožiny

$$\partial M = M_1 \cup M_2 \cup M_3,$$

kde

$$\begin{aligned} M_1 &= \{x^2 + y^2 = 1 \wedge x^2 + z^2 \leq 1\} \\ M_2 &= \{x^2 + y^2 \leq 1 \wedge x^2 + z^2 = 1\} \\ M_3 &= \{x^2 + y^2 = 1 \wedge x^2 + z^2 = 1\}. \end{aligned}$$

Na množině M_1 máme vazbu $g(x, y, z) = x^2 + y^2 - 1$. Díky tomu, že $\nabla g = (2x, 2y)$, je podmínka na hodnost matice porušena jen pokud $x = y = 0$, kde je ovšem pro nás nezajímavá nulová funkční hodnota. Dále pracujeme s funkcí

$$L(x, y, z) = xy + xz - \lambda(x^2 + y^2 - 1).$$

Po zderivování máme

$$\begin{aligned} \frac{\partial L}{\partial x} &= y + z - 2\lambda x \\ \frac{\partial L}{\partial y} &= x - 2\lambda y \\ \frac{\partial L}{\partial z} &= x. \end{aligned}$$

Odtud dostáváme soustavu

$$\begin{aligned} 0 &= y + z - 2\lambda x \\ 0 &= x - 2\lambda y \\ 0 &= x \\ 1 &= x^2 + y^2. \end{aligned}$$

Tuto soustavu není nutné řešit, protože podmínka $x = 0$ zaručuje, že můžeme dostat pouze body s nezajímavou nulovou funkční hodnotou. Podobně postupujeme na množině M_2. Ze symetrie úlohy v proměnných y a z je hned vidět, že opět globální extrém nezískáme.

Přistupme konečně k množině M_3. Zde máme dvě vazby $g(x, y, z) = x^2 + y^2 - 1$ a $h(x, y, z) = x^2 + z^2 - 1$. Podmínka na hodnost se týká matice

$$\begin{pmatrix} 2x & 2y & 0 \\ 2x & 0 & 2z \end{pmatrix}.$$

Hodnost této matice nebude rovna dvěma, pokud budou řádky lineárně závislé. Pokud vyloučíme nezajímavý případ $x = 0$ (vede na nulovou funkční hodnotu), lineární závislost řádků může nastat už jen v případě $y = z = 0$, který opět vede na nulovou funkční hodnotu. Definujme pomocnou funkci

$$L(x, y, z) = xy + xz - \lambda(x^2 + y^2 - 1) - \mu(x^2 + z^2 - 1).$$

Po zderivování máme

$$\frac{\partial L}{\partial x} = y + z - 2\lambda x - 2\mu x$$
$$\frac{\partial L}{\partial y} = x - 2\lambda y$$
$$\frac{\partial L}{\partial z} = x - 2\mu z.$$

Odtud dostáváme soustavu

$$\begin{aligned} 0 &= y + z - 2\lambda x - 2\mu x \\ 0 &= x - 2\lambda y \\ 0 &= x - 2\mu z \\ 1 &= x^2 + y^2 \\ 1 &= x^2 + z^2. \end{aligned}$$

Poslední dvě rovnice dávají, že buď $y = -z$ nebo $y = z$. První případ opět vede na nulovou funkční hodnotu. Zabývejme se případem druhým a to jen v případě $x \neq 0$ (jediný zajímavý). Zde máme $\lambda = \mu \neq 0$. Ze druhé rovnice vyjádříme x a první rovnice pak má tvar

$$0 = 2y - 4\lambda x = 2y - 8\lambda^2 y = 2y(1 - 4\lambda^2).$$

Zbývá nám vyšetřit tři možnosti. Pokud $y = 0$, máme rovněž $z = 0$ a nulovou funkční hodnotu. Pokud $\lambda = \pm\frac{1}{2}$, platí $x = \pm y$. Odtud

$$1 = x^2 + y^2 = 2y^2 \qquad \Longrightarrow \qquad y^2 = \frac{1}{2}.$$

Proto

$$f(x, y, z) = f(\pm y, y, y) = \pm 2y^2 = \pm 1.$$

Odtud $\min_M f = -1$ (nabývá se ho ve dvou bodech $(\mp\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}})$) a dále $\max_M f = 1$ (nabývá se ho ve dvou bodech $(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}})$).

12.8 Věta o regulárním zobrazení

V dalším se budeme zabývat otázkou invertovatelnosti zobrazení $f\colon \mathbb{R}^N \to \mathbb{R}^N$. Výsledky, které získáme, jsou důležité například při vícerozměrné integraci nebo při řešení parciálních diferenciálních rovnic pomocí vhodné změny souřadného systému.

V jednorozměrném případě nám k invertovatelnosti na otevřeném intervalu stačí, když má funkce spojitou a nenulovou derivaci. Zde se pokusíme obdržet vícerozměrnou analogii takového výsledku.

Definice 12.8.1 (Jacobián)**.** Nechť všechny složky zobrazení $f\colon \mathbb{R}^N \to \mathbb{R}^N$ mají v bodě $a \in \mathbb{R}^N$ parciální derivace. Pak matici

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_N}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial x_1}(a) & \cdots & \frac{\partial f_N}{\partial x_N}(a) \end{pmatrix}$$

nazýváme Jacobiho maticí zobrazení f v bodě a. Její determinant se nazývá *Jacobiho determinant* nebo též *jacobián* a značí se $\mathrm{J}_f(a)$.

Definice 12.8.2 (Regulární zobrazení)**.** Nechť $f\colon \mathbb{R}^N \to \mathbb{R}^N$ a $\Omega \subset \mathbb{R}^N$. Řekneme, že f je *regulární zobrazení* na Ω, jestliže

(i) množina Ω je otevřená

(ii) $f \in C^1(\Omega, \mathbb{R}^N)$

(iii) $\mathrm{J}_f \neq 0$ na Ω.

Příklad 12.8.3. (i) Připomeňme si polární souřadnice definované vztahy

$$\begin{aligned} x &= r\cos\varphi \\ y &= r\sin\varphi. \end{aligned}$$

Nejčastěji se pracuje se zobrazením $f\colon (r,\varphi) \mapsto (x,y)$, které zobrazuje $(0,\infty) \times (-\pi,\pi)$ na $\mathbb{R}^2 \setminus \{x \le 0 \wedge y = 0\}$. Jacobiho matice má tvar

$$\begin{pmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{pmatrix}.$$

Proto $\mathrm{J}_f = r \neq 0$ na studované množině a jedná se o regulární zobrazení.

(ii) Sférické souřadnice jsou definované vztahy

$$\begin{aligned} x &= r\cos\psi\cos\varphi \\ y &= r\cos\psi\sin\varphi \\ z &= r\sin\psi. \end{aligned}$$

Nejčastěji se pracuje se zobrazením $f\colon (r,\psi,\varphi) \mapsto (x,y,z)$, které zobrazuje $(0,\infty)\times(-\frac{\pi}{2},\frac{\pi}{2}) \times (-\pi,\pi)$ na $\mathbb{R}^3 \setminus \{x \le 0 \wedge y = 0\}$. Jacobiho matice má tvar

$$\begin{pmatrix} \cos\psi\cos\varphi & -r\sin\psi\cos\varphi & -r\cos\psi\sin\varphi \\ \cos\psi\sin\varphi & -r\sin\psi\sin\varphi & r\cos\psi\cos\varphi \\ \sin\psi & r\cos\psi & 0 \end{pmatrix}.$$

Proto $\mathrm{J}_f = r^2\cos\psi \neq 0$ na studované množině a jedná se o regulární zobrazení.

(iii) Válcové souřadnice jsou definované vztahy

$$\begin{aligned} x &= r\cos\varphi \\ y &= r\sin\varphi \\ z &= h. \end{aligned}$$

Nejčastěji se pracuje se zobrazením $f\colon (r,\varphi,h)\mapsto(x,y,z)$, které zobrazuje $(0,\infty)\times(-\pi,\pi)\times\mathbb{R}$ na $\mathbb{R}^3\setminus\{x\le 0\wedge y=0\}$. Jacobiho matice má tvar

$$\begin{pmatrix} \cos\varphi & -r\sin\varphi & 0\\ \sin\varphi & r\cos\varphi & 0\\ 0 & 0 & 1 \end{pmatrix}.$$

Proto $\mathrm{J}_f = r\neq 0$ na studované množině a jedná se o regulární zobrazení.

Cvičení 12.8.4. Rozmyslete si, jak se zavádí sférické souřadnice v $\mathbb{R}^N$.

Věta 12.8.5 (O inverzi (lokální verze))**.** *Nechť $f\colon\mathbb{R}^N\to\mathbb{R}^N$ je regulární zobrazení na $\mathcal{U}_\tau(a)$ pro jistá $a\in\mathbb{R}^N$ a $\tau>0$. Pak existuje $\sigma>0$ s následujícími vlastnostmi:*
(i) *f je prosté na $\mathcal{U}_\sigma(a)$*
(ii) *$f(\mathcal{U}_\sigma(a))$ je otevřená množina v $\mathbb{R}^N$ (s jakoukoliv normou)*
(iii) *značí-li g inverzní zobrazení k $f|_{\mathcal{U}_\sigma(a)}$, pak $g\in C^1(f(\mathcal{U}_\sigma(a));\mathbb{R}^N)$*
(iv) *$\mathrm{J}_g(f(x))=\frac{1}{\mathrm{J}_f(x)}$ pro všechna $x\in\mathcal{U}_\sigma(a)$*
(v) *pokud $k\in\mathbb{N}\cup\{\infty\}$ a $f\in C^k(\mathcal{U}_\tau(a),\mathbb{R}^N)$, pak $g\in C^k(f(\mathcal{U}_\sigma(a));\mathbb{R}^N)$.*

Důkaz. Důkaz je založen na plné verzi Věty o implicitní funkci (Věta 12.4.13), kde prohodíme roli proměnných $x_1,\dots,x_N$ a $y_1,\dots,y_N$. Pro $x\in\mathcal{U}_\tau(a)$ a $y\in\mathbb{R}^N$ definujme funkce

$$F_i(x,y)=y_i-f_i(x)\qquad\text{pro } i\in\{1,\dots,N\}.$$

Pro tyto funkce platí

$$F_i(a,f(a))=f_i(a)-f_i(a)=0\qquad\text{pro všechna } i\in\{1,\dots,N\},$$

$F_i\in C^1(\mathcal{U}_\tau(a)\times\mathbb{R}^N)$ a

$$\det\begin{pmatrix} \frac{\partial F_1}{\partial x_1}(a,f(a)) & \cdots & \frac{\partial F_1}{\partial x_N}(a,f(a))\\ \vdots & \ddots & \vdots\\ \frac{\partial F_N}{\partial x_1}(a,f(a)) & \cdots & \frac{\partial F_N}{\partial x_N}(a,f(a)) \end{pmatrix}=(-1)^N\mathrm{J}_f(a)\neq 0.$$

Můžeme proto aplikovat plnou verzi Věty o implicitní funkci (Věta 12.4.13) a dostáváme $\delta,\Delta>0$ taková, že pro každé $y\in\mathcal{U}_\delta(f(a))$ existuje právě jedno $x\in\mathcal{U}_\Delta(a)$ splňující

$$0=F_i(x,y)=y_i-f_i(x)\qquad\text{pro všechna } i\in\{1,\dots,N\}.$$

Označíme-li toto x jako $\varphi(y)$, pak $\varphi\in C^1(\mathcal{U}_\delta(f(a));\mathbb{R}^N)$. Ze spojitosti f v bodě a plyne existence $\sigma\in(0,\Delta)$ splňujícího

$$f(x)\in\mathcal{U}_\delta(f(a))\qquad\text{pro každé } x\in\mathcal{U}_\sigma(a).$$

Díky tomu je f prosté zobrazení na $\mathcal{U}_\sigma(a)$. Skutečně, f zobrazuje $\mathcal{U}_\sigma(a)$ do $\mathcal{U}_\delta(f(a))$ a tam každému bodu $y\in\mathcal{U}_\delta(f(a))$ odpovídá právě jeden bod $x\in\mathcal{U}_\Delta(a)$ splňující $y=f(x)$ podle výsledků získaných výše.

Dokažme nyní otevřenost $f(\mathcal{U}_\sigma(a))$. Nechť $y \in f(\mathcal{U}_\sigma(a))$ a $x \in \mathcal{U}_\sigma(a)$ je takový bod, že $f(x) = y$, neboli $x = \varphi(y)$. Z toho, že $y \in \mathcal{U}_\delta(f(a))$, a spojitosti φ na $f(\mathcal{U}_\delta(a))$ plyne existence $\eta > 0$ splňujícího

$$\mathcal{U}_\eta(y) \subset \mathcal{U}_\delta(f(a)) \qquad \text{a} \qquad \varphi(\mathcal{U}_\eta(y)) \subset \mathcal{U}_\sigma(a).$$

Z poslední vlastnosti díky předchozí konstrukci dostáváme $\mathcal{U}_\eta(y) \subset f(\mathcal{U}_\sigma(a))$. To dává otevřenost $f(\mathcal{U}_\sigma(a))$.

Tvrzení (iii) a (v) plynou z toho, že funkce F_i mají stejnou hladkost jako funkce f_i a tuto hladkost mají podle plné verze Věty o implicitní funkci (Věta 12.4.13) rovněž jednotlivé složky zobrazení φ.

Konečně, tvrzení (iv) získáme derivováním vztahu $\varphi(f(x)) = x$ na $\mathcal{U}_\sigma(a)$. Skutečně, podle řetízkového pravidla máme

$$\nabla(\varphi(f(x))) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial y_1}(f(x)) & \cdots & \frac{\partial \varphi_1}{\partial y_N}(f(x)) \\ \vdots & \ddots & \vdots \\ \frac{\partial \varphi_N}{\partial y_1}(f(x)) & \cdots & \frac{\partial \varphi_N}{\partial y_N}(f(x)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_N}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial x_1}(x) & \cdots & \frac{\partial f_N}{\partial x_N}(x) \end{pmatrix}$$

a to se má rovnat jednotkové matici. □

Poznámka 12.8.6. Nenulovost jakobiánu spolu s hladkostí (na rozdíl od jednodimenzionálního případu) nejsou globálně schopny zaručit prostotu. To si snadno uvědomíme, pokud uvážíme polární souřadnice s proměnnými (r, φ) z množiny $(0, \infty) \times \mathbb{R}$.

Věta 12.8.7 (O inverzi (globální verze))**.** *Nechť $f: \mathbb{R}^N \to \mathbb{R}^N$ je regulární zobrazení na otevřené množině $\Omega \subset \mathbb{R}^N$. Pak*
(i) *$f(\Omega)$ je otevřená množina*
(ii) *f je lokálně prosté, neboli ke každému bodu z Ω existuje okolí, kde f je prosté.*
Je-li navíc f prosté na M, odpovídající inverzní zobrazení je regulární na $f(\Omega)$ a pro každé $x \in \Omega$ je Jacobiho matice f v bodě x inverzní maticí k Jakobiho matici zobrazení f^{-1} v bodě $f(x)$.
Je-li f prosté na M, $k \in \mathbb{N} \cup \{\infty\}$ a $f \in C^k(\Omega; \mathbb{R}^N)$, pak $f^{-1} \in C^k(f(\Omega); \mathbb{R}^N)$.

Důkaz. Tvrzení (i) a (ii) plynou z předchozí věty (připomeňme, že sjednocení otevřených množin je otevřená množina). Zbylá tvrzení plynou rovněž z předchozí věty, neboť mají lokální charakter. □

Poznámka 12.8.8. Předchozí větě se také říká Věta o regulárním zobrazení (přesněji, týká se to dvou tvrzení pro f prosté).

Příklad 12.8.9. Převeďme do polárních souřadnic a tím vyřešme úlohu

$$x\frac{\partial u}{\partial y} - y\frac{\partial u}{\partial x} = 0.$$

Definujme pomocné zobrazení $\tilde{u}(r, \varphi) = u(x(r, \varphi), y(r, \varphi))$ pro $(r, \varphi) \in (0, \infty) \times (-\pi, \pi)$, neboli $\tilde{u} = u \circ f$, kde f bylo představeno v první části Příkladu 12.8.3.

Podle globální verze Věty o inverzi (Věta 12.8.7) funkce $u \in C^1(\mathbb{R}^2 \setminus \{x \le 0 \wedge y = 0\})$ řeší původní úlohu na $\mathbb{R}^2 \setminus \{x \le 0 \wedge y = 0\}$ právě tehdy, když nová funkce $\widetilde{u} \in C^1((0,\infty) \times (-\pi,\pi))$ řeší na $(0,\infty) \times (-\pi,\pi)$ novou úlohu, kterou získáme přepočítáním pomocí řetízkového pravidla. Máme

$$\begin{aligned}
\frac{\partial \widetilde{u}}{\partial r} &= \frac{\partial u}{\partial x}\frac{\partial f_1}{\partial r} + \frac{\partial u}{\partial y}\frac{\partial f_2}{\partial r} = \frac{\partial u}{\partial x}\cos\varphi + \frac{\partial u}{\partial y}\sin\varphi \\
\frac{\partial \widetilde{u}}{\partial \varphi} &= \frac{\partial u}{\partial x}\frac{\partial f_1}{\partial \varphi} + \frac{\partial u}{\partial y}\frac{\partial f_2}{\partial \varphi} = -\frac{\partial u}{\partial x} r\sin\varphi + \frac{\partial u}{\partial y} r\cos\varphi,
\end{aligned}$$

a proto

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \frac{\partial \widetilde{u}}{\partial r}\cos\varphi - \frac{1}{r}\frac{\partial \widetilde{u}}{\partial \varphi}\sin\varphi \\
\frac{\partial u}{\partial y} &= \frac{\partial \widetilde{u}}{\partial r}\sin\varphi + \frac{1}{r}\frac{\partial \widetilde{u}}{\partial \varphi}\cos\varphi.
\end{aligned}$$

Původní rovnice dostává nový tvar

$$0 = r\cos\varphi\Big(\frac{\partial \widetilde{u}}{\partial r}\sin\varphi + \frac{1}{r}\frac{\partial \widetilde{u}}{\partial \varphi}\cos\varphi\Big) - r\sin\varphi\Big(\frac{\partial \widetilde{u}}{\partial r}\cos\varphi - \frac{1}{r}\frac{\partial \widetilde{u}}{\partial \varphi}\sin\varphi\Big) = \frac{\partial \widetilde{u}}{\partial \varphi}.$$

Proto jsou řešením naší rovnice funkce tvaru $\widetilde{u}(r,\varphi) = \widetilde{U}(r)$ pro jakoukoliv funkci $U \in C^1((0,\infty))$. Odtud původní úlohu řeší

$$u(x,y) = \widetilde{U}(\sqrt{x^2+y^2}) =: U(x^2+y^2) \qquad \text{na } \mathbb{R}^2 \setminus \{x \le 0 \wedge y = 0\}.$$

Toto je zřejmě řešení na $\mathbb{R}^2 \setminus \{(0,0)\}$. Je-li možné funkci U v počátku spojitě (zprava) dodefinovat tak, aby existovala $U'_+(0)$, máme řešení na celém $\mathbb{R}^2$.

Kapitola 13

Variační počet

13.1 Úvod

Stejně jako bylo přirozené rozšířit teorii extrémů funkcí jedné reálné proměnné na teorii extrémů funkcí více proměnných, existují aplikace vyžadující ještě obecnější přístup. Budeme se zabývat extrémy v nekonečnědimenzionálních prostorech, jejichž prvky jsou funkce. Jako motivaci si uveďme několik příkladů.

Příklad 13.1.1 (Úloha o brachystochroně)**.** Nechť $a, b, A, B \in \mathbb{R}$, $a < b$, $A > B$. Naším cílem je nalézt trajektorii, po níž se hmotný bod vlivem působení gravitace co nejrychleji dostane z bodu (a, A) do bodu (b, B) (jiná interpretace: mezi oběma body vyrobíme skluzavku a necháme po ní sklouznout kuličku, přičemž zanedbáváme její vlastní moment setrvačnosti). Uvažujeme-li jen trajektorie, které je možné popsat jako graf $C^1([a, b])$-funkce, pak pro celkový čas máme

$$T(y) = \int_a^b \frac{\sqrt{1 + {y'}^2(x)}}{v(x)} \,\mathrm{d}x.$$

Vyjádření rychlosti v získáme ze zákona zachování mechanické energie

$$\frac{1}{2} m v^2(x) + m g y(x) = m g A \qquad \Longrightarrow \qquad v(x) = \sqrt{2g(A - y(x))}.$$

Celkově minimalizujeme funkcionál

$$T(y) = \frac{1}{\sqrt{2g}} \int_a^b \frac{\sqrt{1 + {y'}^2(x)}}{\sqrt{A - y(x)}} \,\mathrm{d}x$$

přes funkce z $C^1([a, b])$ splňující $y(a) = A$ a $y(b) = B$.

Definice 13.1.2 (Funkcionál)**.** Zobrazení z normovaného lineárního prostoru do $\mathbb{R}$ se nazývá *funkcionál*.

Příklad 13.1.3. (i) Délka grafu funkce

$$F(y) = \int_a^b \sqrt{1 + y'(x)}\,\mathrm{d}x$$

je funkcionál na $C^1([a,b])$ (připomeňme ještě, že na $C^1([a,b])$ je zvykem zavádět normu $\|y\|_{C^1([a,b])} := \max_{[a,b]} |y| + \max_{[a,b]} |y'|$).
(ii) Riemannův (či Newtonův) integrál je dokonce lineární funkcionál na $C([a,b])$.
(iii) Funkcionálem na $C^1([a,b])$ je také

$$F(y) = \sqrt{\int_a^b y'^2(x)\,\mathrm{d}x}.$$

V dalším si nejprve vybudujeme abstraktní teorii pro klasifikaci lokálních extrémů, která se velmi podobá teorii pro lokální extrémy funkcí více proměnných. Později při aplikaci této teorie zjistíme, že kupříkladu ověřování pozitivní definitnosti druhého diferenciálu v nekonečné dimenzi není vůbec snadné a vyžaduje vybudování nových nástrojů. Právě toto rozšíření abstraktní teorie bude těžistěm této kapitoly. Nakonec se budeme zabývat několika klasickými úlohami, jako jsou již zmíněná úloha o brachystochroně či úloha o zavěšeném řetězu.

Naše teorie bude pracovat s prostorem $C^1([a,b])$. Na rozdíl od případu extrémů funkcí více proměnných se nám podaří získat jen velmi slabé výsledky ohledně existence globálních extrémů. Konkrétně bude zcela chybět výsledek, který by byl svou užitečností srovnatelný s Větou o nabývání extrémů spojitou funkcí (Věta 11.11.16), tj. že spojitá funkce na omezené a uzavřené (tedy kompaktní) množině nabývá svého maxima a minima (v nekonečnědimenzionálním prostoru omezenost a uzavřenost neimplikují kompaktnost). Moderní matematická analýza z těchto důvodů prostor $C^1([a,b])$ nahrazuje takzvanými *Sobolevovými prostory*, které jsou vybudovány na teorii *Lebesgueova integrálu* a z hlediska variačního počtu nabízejí silnější výsledky. Tyto partie jdou ale nad rámec těchto skript.

13.2 Abstraktní teorie

Protože v prostorech funkcí nemáme přirozeně danou kanonickou bázi, základním pojmem diferenciálního počtu je derivace ve směru.

Definice 13.2.1 (Gâteauxův a Fréchetův diferenciál)**.** Nechť X je normovaný lineární prostor, $F\colon X \to \mathbb{R}$ je funkcionál a $a \in D_F$.
(i) Nechť $h \in X$ a existuje $\delta > 0$ takové, že $\{a + th\colon |t| < \delta\} \subset D_F$. Řekneme, že F má v bodě a *Gâteauxův diferenciál* ve směru h (nebo též *Gâteauxovu derivaci* ve směru h), jestliže existuje vlastní limita

$$\lim_{t\to 0} \frac{F(a+th) - F(a)}{t} = \frac{\mathrm{d}}{\mathrm{d}t} F(a+th)|_{t=0}.$$

Tuto limitu pak značíme $\delta F(a;h)$ a nazýváme ji *Gâteauxovým diferenciálem* funkcionálu F v bodě a ve směru h.

(ii) Nechť existuje $\delta > 0$ takové, že $\mathcal{U}_\delta(a) \subset D_F$. Řekneme, že F má v bodě a *Fréchetův diferenciál*, jestliže existuje spojitý lineární funkcionál $L\colon X \to \mathbb{R}$ splňující

$$\lim_{h\to 0} \frac{F(a+h) - F(a) - Lh}{\|h\|} = 0.$$

Zmíněný lineární funkcionál pak značíme $\mathrm{d}F(a)$ a nazýváme jej *Fréchetovým diferenciálem* funkcionálu F v bodě a.

Poznámka 13.2.2. (i) Snadno se ověří, že Gâteauxův diferenciál (v daném bodě a směru) a Fréchetův diferenciál (v daném bodě) jsou v případě existence určeny jednoznačně.
(ii) Pro funkce více proměnných je Gâteauxův diferenciál derivací ve směru a Fréchetův diferenciál je totální diferenciál.
(iii) Existence Fréchetova diferenciálu zřejmě implikuje existenci Gâteauxova diferenciálu pro každý směr a pak platí

$$\delta F(a;h) = \mathrm{d}F(a)(h).$$

Toho se využívá při hledání Fréchetova diferenciálu. Spočítáme Gâteauxovy diferenciály ve všech směrech a máme jediného kandidáta na Fréchetův diferenciál.
(iv) Existence Fréchetova diferenciálu implikuje spojitost.
(v) Zatímco lineární zobrazení mezi konečnědimenzionálními prostory jsou vždy spojitá, v nekonečné dimenzi toto není pravda. Stačí uvážit prostor ℓ_2 a na něm funkcionál

$$Lx := x_1 + 2x_2 + 3x_3 + \dots,$$

kde $x = \{x_1, x_2, x_3, \dots\}$. Tento funkcionál je zřejmě lineární, ale pro vektory $\mathrm{e}_k = \{\delta_{ik}\}_{i=1}^{\infty}$ platí

$$L\Big(\frac{\mathrm{e}_k}{k}\Big) = 1, \qquad \frac{\mathrm{e}_k}{k} \to 0 \qquad \text{a} \qquad L(0) = 0.$$

Příklad 13.2.3. Uvažujme zobrazení

$$F(y) = \int_0^1 xy^2(x)\,\mathrm{d}x$$

na $C([0,1])$. Pro každé $h \in C([0,1])$ a $t \in \mathbb{R} \setminus \{0\}$ máme

$$\frac{F(y+th) - F(y)}{t} = \frac{1}{t}\int_0^1 x((y+th)^2 - y^2)\,\mathrm{d}x = \int_0^1 2xyh\,\mathrm{d}x + t\int_0^1 xh^2\,\mathrm{d}x.$$

Odtud (pro zafixované h jsou oba integrály konečné)

$$\delta F(y;h) = \int_0^1 2xyh\,\mathrm{d}x \qquad \text{pro všechna } y \in C([0,1]) \text{ a } h \in C([0,1]).$$

Tím jsme zároveň dostali i jediného kandidáta na Fréchetův diferenciál. Zbývá ověřit podmínku z definice. Máme

$$\frac{F(y+h) - F(y) - \delta F(y;h)}{\|h\|} = \frac{1}{\|h\|}\int_0^1 x((y+h)^2 - y^2 - 2yh)\,\mathrm{d}x = \frac{1}{\|h\|}\int_0^1 xh^2\,\mathrm{d}x,$$

kde stačí využít

$$\Big|\frac{1}{\|h\|}\int_0^1 xh^2\,\mathrm{d}x\Big| \le \frac{1}{\|h\|}\int_0^1 1\cdot \max_{[0,1]}{}^2|h|\,\mathrm{d}x = \frac{1}{\|h\|}\max_{[0,1]}{}^2|h| = \|h\|$$

a odtud již snadno obdržíme existenci Fréchetova diferenciálu. Proto $\mathrm{d}F(y)(h) = \int_0^1 2xyh\,\mathrm{d}x$.

Cvičení 13.2.4. (i) Uvažte zobrazení

$$F(y) = \int_0^1 xy'^2(x)\,\mathrm{d}x$$

na $C^1([0,1])$. Stejným postupem jako v předchozím příkladu ukažte, že

$$\mathrm{d}F(y)(h) = \delta F(y)(h) = \int_0^1 2xy'h'\,\mathrm{d}x \qquad \forall y \in C^1([0,1]) \text{ a } h \in C^1([0,1]).$$

(ii) Podobně pro

$$F(y) = \int_0^1 \big(xy^2(x) + xy'^2(x)\big)\,\mathrm{d}x$$

získejte

$$\mathrm{d}F(y)(h) = \delta F(y)(h) = 2\int_0^1 (xyh + xy'h')\,\mathrm{d}x \quad \forall y \in C^1([0,1]) \text{ a } h \in C^1([0,1]).$$

(iii) Předchozí příklady naznačují, jak vypadá základní aritmetika Gâteauxova diferenciálu (případně Fréchetova diferenciálu). Čemu se rovná $\delta(\alpha F + \beta G)(y)(h)$ pro $\alpha, \beta \in \mathbb{R}$ a funkcionály F, G? Vyslovte odpovídající větu a dokažte ji.

Definice 13.2.5 (Lokální minimum)**.** Nechť X je normovaný lineární prostor a $F\colon X \to \mathbb{R}$ je funkcionál. Řekneme, že $a \in D_F$ je *bodem lokálního minima* funkcionálu F, jestliže existuje $\varepsilon > 0$ takové, že

$$F(a) \le F(x) \qquad \text{pro všechna } x \in \mathcal{U}_\varepsilon(a) \cap D_F.$$

V případě ostré nerovnosti na $\mathcal{P}_\varepsilon(a) \cap D_F$ hovoříme o ostrém lokálním minimu. Analogicky se definuje lokální maximum.

Definice 13.2.6 (Stacionární bod)**.** Nechť X je normovaný lineární prostor a nechť $F\colon X \to \mathbb{R}$ je funkcionál. Řekneme, že $a \in D_F$ je *stacionárním bodem* (nebo *extremálou* nebo *kritickým bodem*) funkcionálu F, jestliže

$$\delta F(a;h) = 0 \qquad \text{pro všechna } h \in X.$$

Základním tvrzením právě budované teorie je následující nutná podmínka.

Věta 13.2.7 (Eulerova nutná podmínka)**.** *Nechť X je normovaný lineární prostor, funkcionál $F\colon X \to \mathbb{R}$ má lokální extrém v bodě $a \in X$ a $h \in X$. Pokud existuje $\delta F(a;h)$, pak $\delta F(a;h) = 0$.*

Důkaz. Vše plyne z chování funkce $t \mapsto F(a+th)$, neboť toto zobrazení má v bodě $t = 0$ extrém. $\square$

Příklad 13.2.8. Uvážíme-li funkcionál z Příkladu 13.2.3, pak je Eulerova nutná podmínka splněna zřejmě pro $y \equiv 0$. Naopak, nutná podmínka není splněna pro žádnou jinou spojitou funkci. Skutečně, je-li $y \in C([0,1])$ netriviální, stačí zvolit $h := y$ a dostaneme nenulový integrál (pokud existuje $x_0 \in [0,1]$ takové, že $f(x_0) > 0$, díky spojitosti na jistém okolí platí $f(x) > \frac{1}{2} f(x_0)$, atd.).

Poznámka 13.2.9. V předchozím příkladu jsme měli velice jednoduchý funkcionál, a proto bylo poměrně snadné určit, které body jsou stacionární a které nikoliv. V praxi se většinou pracuje s integrálními funkcionály, kde integrand závisí nejen na x a y, ale také na y' (podívejte se na funkcionál vystupující v úloze o brachystochroně). K hledání stacionárních bodů takových funkcionálů byly vyvinuty techniky, které si představíme později.

Nyní si vybudujeme analogickou teorii ke klasifikaci lokálních extrémů pomocí definitnosti kvadratických forem příslušejícím druhým diferenciálům funkcí více proměnných. Pro jednoduchost značení všechny výsledky vyslovíme jen pro lokální minimum.

Definice 13.2.10 (Druhý Gâteauxův diferenciál)**.** Nechť X je normovaný lineární prostor, $F\colon X \to \mathbb{R}$ je funkcionál, $a, h, k \in X$ a existuje $\delta F(a; h)$. Nechť existuje vlastní
$$\delta^2 F(a; h, k) := \lim_{t\to 0} \frac{\delta F(a + tk; h) - \delta F(a; h)}{t}.$$
Pak $\delta^2 F(a; h, k)$ nazýváme *druhým Gâteauxovým diferenciálem* ve směrech h a k.

Poznámka 13.2.11. Pokud $k = h$, pak máme
$$\delta^2 F(a; h, h) = \frac{\mathrm{d}^2}{\mathrm{d}t^2} F(a + th)|_{t=0}.$$

Věta 13.2.12 (Lagrangeova nutná podmínka)**.** *Nechť X je normovaný lineární prostor, funkcionál $F\colon X \to \mathbb{R}$ má lokální minimum v bodě $a \in X$ a $h \in X$. Pokud existuje $\delta^2 F(a; h, h)$, pak $\delta^2 F(a; h, h) \geq 0$.*

Důkaz. Definujme $g(t) = F(a + th)$. Máme $g'(0) = 0$ (podle Eulerovy nutné podmínky, tedy Věty 13.2.7). Pokud by platilo $g''(0) = \delta^2 F(a; h, h) = -\theta < 0$, pak bychom pro $t > 0$ dostatečně malé měli díky Lagrangeově větě o přírůstku funkce
$$g(t) - g(0) = tg'(\xi) = t\frac{g'(\xi) - g'(0)}{\xi}\xi \leq t\frac{-\theta}{2}\xi < 0,$$
což je spor s tím, že v bodě a je lokální minimum. $\square$

Příklad 13.2.13. Pro funkcionál z Příkladu 13.2.3 máme pro každé $y \in C([0,1])$
$$\begin{aligned}\delta^2 F(y; h, h) &= \lim_{t\to 0} \frac{\delta F(y + th; h) - \delta F(y; h)}{t}\\ &= \lim_{t\to 0} \frac{2}{t}\Big(\int_0^1 x(y + th)h \,\mathrm{d}x - \int_0^1 xyh \,\mathrm{d}x\Big) = 2\int_0^1 xh^2 \,\mathrm{d}x.\end{aligned}$$

Lagrangeova nutná podmínka (Věta 13.2.12) je tedy splněna pro všechna $y \in C([0,1])$. Pro hledání lokálního minima se ukázala jako užitečnější dříve použitá Eulerova nutná podmínka. Na druhou stranu, Lagrangeova nutná podmínka pro lokální maximum (s obrácenou nerovností) nám říká, že náš funkcionál nemá žádné lokální maximum na $C([0,1])$.

Postačující podmínka má podobný tvar jako u funkcí více proměnných. Uvedeme si slabší verzi (oproti standardním výsledkům v literatuře), abychom byli schopni výsledek dokázat elementárními prostředky.

Věta 13.2.14 (Lagrangeova postačující podmínka (zeslabená verze))**.** *Nechť X je normovaný lineární prostor a $a \in X$ je stacionárním bodem funkcionálu $F\colon X \to \mathbb{R}$. Jestliže existuje okolí bodu a, kde platí $\delta^2 F(x;h,h) \geq 0$, pak F má v bodě a lokální minimum.*

Důkaz. Nechť podmínka $\delta^2 F(x;h,h) \geq 0$ platí na $\mathcal{U}_\eta(a)$. Pro libovolné $y \in \mathcal{U}_\eta(a)$ definujme

$$g(t) = F(a + t(y-a)) \qquad \text{pro } t \in [0,1].$$

Pak máme pro $\xi \in (0,1)$

$$\begin{aligned} F(y) - F(a) &= g(1) - g(0) = g'(0) + \frac{1}{2} g''(\xi) \\ &= \delta F(a; y-a) + \frac{1}{2}\delta^2 F(a + \xi(y-a); y-a, y-a) \geq 0 + 0 = 0. \end{aligned}$$

□

Poznámka 13.2.15. Standardní znění namísto nezápornosti druhých Gâteauxových diferenciálů na okolí požaduje spojitost druhého Fréchetova diferenciálu v bodě a a existenci $\alpha > 0$ splňujícího

$$\delta^2 F(a;h,h) \geq \alpha \|h\|^2 \qquad \text{pro všechna } h \in X.$$

Takové znění pak nabízí ostré lokální minimum v bodě a. Důkaz tohoto tvrzení je trochu těžší a navíc se pro naši situaci, kdy pracujeme s prostorem $C^1([a,b])$, se výše uvedený předpoklad špatně ověřuje.

Příklad 13.2.16. Pro funkcionál z Příkladu 13.2.3 jsme již nalezli jediný stacionární bod $y \equiv 0$ a navíc jsme již také ukázali

$$\delta^2 F(y;h,h) = \int_0^1 x h^2 \,\mathrm{d}x \geq 0.$$

Náš stacionární bod je tedy bodem lokálního minima.

Cvičení 13.2.17. Ukažte, že v předchozím příkladu není možné pro žádné $\alpha > 0$ splnit podmínku

$$\delta^2 F(a;h,h) \geq \alpha \|h\|^2 \qquad \text{pro všechna } h \in X.$$

K tomu stačí sestrojit posloupnost funkcí $\{h_k\} \subset C([0,1])$ splňující

$$\|h_k\| = \max_{[0,1]} |h_k| = 1 \qquad \text{a} \qquad \int_0^1 x h_k^2 \,\mathrm{d}x \overset{k\to\infty}{\to} 0.$$

V jednodimenzionálním případě se výše uvedená kritéria týkají nezápornosti druhé derivace a ta zase souvisí s konvexitou. I ve variačním počtu je konvexita postačující podmínkou pro existenci globálního minima ve stacionárním bodě.

Definice 13.2.18 (Konvexita funkcionálu)**.** Nechť X je normovaný lineární prostor, $M \subset X$ je konvexní a $F\colon X \to \mathbb{R}$ je funkcionál. Řekneme, že funkcionál F je *konvexní* na M, jestliže

$$F(\lambda x + (1-\lambda)y) \le \lambda F(x) + (1-\lambda)F(y) \qquad \text{pro všechna } x, y \in M \text{ a } \lambda \in [0,1].$$

Věta 13.2.19 (Postačující podmínka pro konvexní funkcionál)**.** *Nechť X je normovaný lineární prostor a $F\colon X \to \mathbb{R}$ je konvexní funkcionál. Pak každý jeho stacionární bod je bodem globálního minima F na X.*

Důkaz. Pro spor předpokládejme, že $a \in X$ je stacionární bod, ale existuje $b \in X$ takové, že $F(b) < F(a)$. Díky konvexitě pak máme pro každé $t \in (0,1]$

$$\frac{F(a + t(b-a)) - F(a)}{t} \le \frac{(1-t)F(a) + tF(b) - F(a)}{t} = F(b) - F(a).$$

Odtud $\delta F(a; b-a) \le F(b) - F(a) < 0$ a bod a není stacionárním bodem, což je spor. □

Příklad 13.2.20. Předchozí věta se dá aplikovat na funkcionál z Příkladu 13.2.3. Ten je totiž konvexní, což snadno plyne z odhadu založeného na Youngově nerovnosti

$$\begin{aligned}(\lambda y + (1-\lambda)z)^2 &= \lambda^2 y^2 + 2\lambda(1-\lambda)yz + (1-\lambda)^2 z^2 \\ &\le \lambda^2 y^2 + \lambda(1-\lambda)(y^2 + z^2) + (1-\lambda)^2 z^2 \\ &= \lambda y^2 + (1-\lambda)z^2.\end{aligned}$$

Také bylo možné využít konvexitu funkce $t \mapsto t^2$.

13.3 Teorie pro funkcionály reprezentované integrálem

V dalším se budeme zabývat lokálními extrémy funkcionálů tvaru

$$F(y) = \int_a^b f(x, y(x), y'(x)) \,\mathrm{d}x$$

na množině

$$M := \{y \in C^1([a,b])\colon y(a) = A, y(b) = B\},$$

kde $a, b, A, B \in \mathbb{R}$, $a < b$ a $f \in C^2([a,b]\times\mathbb{R}^2)$. Argument funkce f značíme (x, y, z).

Předchozí teorii zde nemůžeme aplikovat přímo, neboť množina M obecně není lineární prostor (vyjma případu $A = B = 0$). Tento problém vyřešíme tím, že si zadefinujeme pomocnou funkci v jako afinní funkci splňující $v(a) = A$ a $v(b) = B$ (tedy $v(x) = A + \frac{B-A}{b-a}(x-a)$), píšeme $y = u + v$ a pracujeme s funkcionálem

$$\Phi(u) := F(u + v)$$

na množině

$$X := \{u \in C^1([a,b]) \colon u(a) = 0, u(b) = 0\},$$

což už je lineární prostor. Používáme zde normu

$$\|u\|_{C^1([a,b])} = \max_{[a,b]} |u| + \max_{[a,b]} |u'|.$$

V dalším budeme vždy pod symbolem X rozumět právě zavedený prostor.

Domluvme se ještě, že z důvodu zjednodušení budou v dalším výrazy $y'(a)$ a $y'(b)$ znamenat jednostranné derivace v těchto krajních bodech z vnitřní strany intervalu (a, b). Podobně, budeme-li hovořit o splnění nějaké diferenciální rovnice na $[a, b]$, myslíme tím, že v krajních bodech tuto rovnici splňují odpovídající jednostranné derivace.

Pro aplikaci předchozí teorie potřebujeme znát tvar $\delta\Phi(u; h)$ a $\delta^2\Phi(u; h, h)$. V zájmu přehlednosti zápisu budeme dále používat značení $f_z = \frac{\partial f}{\partial z}$, atd.

Lemma 13.3.1 (O tvaru $\delta\Phi(u; h)$ a $\delta^2\Phi(u; h, h)$)**.** *Nechť F, Φ, v jsou jako výše a $u, h \in X$. Pak (používáme $y := u + v$)*

$$\delta\Phi(u; h) = \int_a^b \big(f_y(x, y, y')h + f_z(x, y, y')h'\big)\,\mathrm{d}x$$

a

$$\delta^2\Phi(u; h, h) = \int_a^b \big(f_{yy}(x, y, y')h^2 + 2f_{yz}(x, y, y')hh' + f_{zz}(x, y, y')h'^2\big)\,\mathrm{d}x.$$

Důkaz. První výsledek získáme tak, že provedeme úpravu

$$\begin{aligned}
&\delta\Phi(u; h) - \int_a^b \big(f_y(x, y, y')h + f_z(x, y, y')h'\big)\,\mathrm{d}x \\
&\quad= \lim_{t\to 0} \int_a^b \Big(\frac{f(x, y + th, y' + th') - f(x, y, y')}{t} - f_y(x, y, y')h - f_z(x, y, y')h'\Big)\,\mathrm{d}x \\
&\quad= \lim_{t\to 0} \int_a^b \Big(\frac{f(x, y + th, y' + th') - f(x, y, y' + th')}{t} - f_y(x, y, y')h\Big)\,\mathrm{d}x \\
&\qquad\qquad + \lim_{t\to 0} \int_a^b \Big(\frac{f(x, y, y' + th') - f(x, y, y')}{t} - f_z(x, y, y')h'\Big)\,\mathrm{d}x \\
&\quad=: \lim_{t\to 0} I_1 + \lim_{t\to 0} I_2
\end{aligned}$$

a ukážeme, že

$$\lim_{t\to 0} I_1 = \lim_{t\to 0} I_2 = 0.$$

Protože $f \in C^1([a,b]\times\mathbb{R}^2)$, podle Lagrangeovy věty o přírůstku funkce (Věta 6.3.3) pro každé $t > 0$ exitstuje $\theta \in (0,1)$ splňující

$$\frac{f(x,y+th,y'+th') - f(x,y,y'+th')}{t} = f_y(x,y+\theta th, y'+th')h.$$

Dále, $y, h \in X$ jsou pevně zvolená (tudíž $\max_{[a,b]}|h| + \max_{[a,b]}|h'| < \infty$) a f_y je spojitá na $[a,b]\times\mathbb{R}^2$, tedy stejnoměrně spojitá na kompaktních podmnožinách $[a,b]\times\mathbb{R}^2$. Proto ke každému $\varepsilon > 0$ dostáváme t_0 takové, že pro $t \in (0,t_0)$ platí

$$\begin{aligned}
|I_1| &= \Big|\int_a^b \big(f_y(x,y+\theta th,y'+th')h - f_y(x,y,y')h\big)\,\mathrm{d}x\Big| \\
&\le \int_a^b \max_{[a,b]}|h||f_y(x,y+\theta th,y'+th') - f_y(x,y,y')|\,\mathrm{d}x \\
&\le \int_a^b \max_{[a,b]}|h|\varepsilon\,\mathrm{d}x = C\varepsilon.
\end{aligned}$$

Podobně pro $t < 0$. Dokázali jsme, že $\lim_{t\to 0} I_1 = 0$. Výsledek $\lim_{t\to 0} I_2 = 0$ získáme analogicky. Tím je dokázán vztah pro $\delta\Phi(y;h)$. Vztah pro $\delta^2\Phi(u;h,h)$ obdržíme tak, že předchozí metodu aplikujeme na vztah pro $\delta\Phi(u;h)$. Tentokrát využíváme předpoklad $f \in C^2([a,b]\times\mathbb{R}^2)$. Navíc použijeme $f_{zy} = f_{yz}$. □

13.3.1 Euler–Lagrangeova rovnice

Dalším problémem je, že Eulerova nutná podmínka (Věta 13.2.7) díky právě získaným výsledkům získává poněkud nepřehledný tvar

$$\int_a^b \big(f_y(x,y,y')h + f_z(x,y,y')h'\big)\,\mathrm{d}x = 0 \qquad \text{pro všechna } h \in X$$

a my z ní potřebujeme získat stacionární body. V tom nám pomůže integrace per partes kombinovaná s následujícím výsledkem.

Lemma 13.3.2 (DuBois–Reymondovo lemma)**.** *Nechť pro funkci $g \in C([a,b])$ platí*

$$\int_a^b gh'\,\mathrm{d}x = 0 \qquad \textit{pro všechna } h \in C^1([a,b]) \textit{ splňující } h(a) = h(b) = 0.$$

Pak g je konstantní na $[a,b]$.

Důkaz. Nejprve si povšimněme, že pokud funkce $\varphi \in C([a,b])$ splňuje $\int_a^b \varphi\,\mathrm{d}t = 0$, pak pro funkci $\Phi(x) = \int_a^x \varphi(t)\,\mathrm{d}t$ máme $\Phi \in C^1([a,b])$ a $\Phi(a) = \Phi(b) = 0$. Díky tomu podmínka ze znění věty implikuje

$$\int_a^b g\varphi\,\mathrm{d}x = 0 \qquad \text{pro všechna } \varphi \in C([a,b]) \text{ splňující } \int_a^b \varphi\,\mathrm{d}x = 0$$

(tato podmínka je dokonce podmínce ze znění věty ekvivalentní, neboť derivace všech funkcí z X mají výše popsané vlastnosti).

Položme

$$\alpha := \frac{1}{b-a}\int_a^b g\,\mathrm{d}t.$$

Díky linearitě integrálu a předchozím výsledkům máme

$$\int_a^b (g-\alpha)\varphi\,\mathrm{d}x = 0 \qquad \text{pro všechna } \varphi \in C([a,b]) \text{ splňující } \int_a^b \varphi\,\mathrm{d}x = 0.$$

Speciálně lze volit $\varphi := g - \alpha$ (od spojité funkce odečítáme přesně tu konstantu, která zařídí nulovost integrálu přes (a,b)) a dostáváme

$$\int_a^b (g-\alpha)^2\,\mathrm{d}x = 0.$$

Nyní již snadno díky spojitosti a nezápornosti integrandu obdržíme $g \equiv \alpha$ na $[a,b]$. □

V nejjednodušších případech, kdy funkce f nezávisí na třetí proměnné z (třeba u funkcionálu z Příkladu 13.2.3), je výhodnější používat následující výsledek.

Lemma 13.3.3 (Fundamentální lemma variačního počtu)**.** *Nechť pro funkci $g \in C([a,b])$ platí*

$$\int_a^b gh\,\mathrm{d}x = 0 \qquad \textit{pro všechna } h \in C^1([a,b]) \textit{ splňující } h(a) = h(b) = 0.$$

Pak $g \equiv 0$ na $[a,b]$.

Důkaz. Nechť existuje $x_0 \in (a,b)$ takové, že $g(x_0) > 0$ (díky spojitosti plyne předchozí i za situací $g(a) > 0$ a $g(b) > 0$, případ záporné funkční hodnoty se řeší podobně, budeme tedy umět vyřešit všechny možné případy). Pak existuje $\delta \in (0, \min\{x_0 - a, b - x_0\})$ takové, že $g(x) > \frac{1}{2}g(x_0)$ na $(x_0-\delta, x_0+\delta)$. Definujme

$$h(x) = \begin{cases} 0 & \text{pro } x \in [a,b] \setminus (x_0-\delta, x_0+\delta) \\ \cos^2(\frac{\pi}{2\delta}(x-x_0)) & \text{pro } x \in (x_0-\delta, x_0+\delta). \end{cases}$$

Pak $h \in C^1([a,b])$ a $h(a) = h(b) = 0$. Navíc

$$\begin{aligned} \int_a^b gh\,\mathrm{d}x &= \int_{x_0-\delta}^{x_0+\delta} g(x)\cos^2(\tfrac{\pi}{2\delta}(x-x_0))\,\mathrm{d}x \\ &\geq \frac{g(x_0)}{2}\int_{x_0-\delta}^{x_0+\delta} \cos^2(\tfrac{\pi}{2\delta}(x-x_0))\,\mathrm{d}x > 0. \end{aligned}$$

□

Naše dosavadní výsledky nám dávají následující tvar nutné podmínky pro lokální extrém.

Věta 13.3.4 (Euler–Lagrangeova rovnice)**.** *Nechť* $f \in C^2([a,b] \times \mathbb{R}^2)$ *a* $y_0 \in M$ *je stacionárním bodem funkcionálu* F*. Pak funkce*

$$x \mapsto f_z(x, y_0(x), y_0'(x))$$

je spojitě diferencovatelná na $[a,b]$ *a* y_0 *splňuje Euler–Lagrangeovu rovnici*

$$f_y(x, y_0(x), y_0'(x)) - \frac{\mathrm{d}}{\mathrm{d}x} f_z(x, y_0(x), y_0'(x)) = 0 \qquad \text{na } [a,b].$$

Důkaz. Jak již bylo ukázáno výše, odečtením vhodné afinní funkce můžeme přejít k funkci $u_0 \in X$ a funkcionálu Φ. Pro každé $h \in X$ pak máme splněnu Eulerovu nutnou podmínku (Věta 13.2.7)

$$0 = \delta\Phi(u_0, h) = \int_a^b \big(f_y(x, y_0, y_0')h + f_z(x, y_0, y_0')h' \big) \, \mathrm{d}x.$$

První část integrálu napravo se dá pomocí integrace per partes přepsat do tvaru (hraniční členy zmizí díky $h(a) = h(b) = 0$)

$$\int_a^b f_y(x, y_0, y_0')h \, \mathrm{d}x = -\int_a^b \Big(\int_a^x f_y(t, y_0(t), y_0'(t)) \, \mathrm{d}t \Big) h' \, \mathrm{d}x.$$

Proto celkově máme pro všechna $h \in X$

$$0 = \int_a^b \Big(\int_a^x f_y(t, y_0(t), y_0'(t)) \, \mathrm{d}t - f_z(x, y_0(x), y_0'(x)) \Big) h' \, \mathrm{d}x.$$

Du Bois–Reymondsovo lemma (Lemma 13.3.2) nám proto dává

$$\int_a^x f_y(t, y_0(t), y_0'(t)) \, \mathrm{d}t - f_z(x, y_0(x), y_0'(x)) \equiv C \qquad \text{na } [a,b].$$

Zderivováním dostáváme Euler–Lagrangeovu rovnici. Navíc funkce

$$x \mapsto f_z(x, y_0(x), y_0'(x))$$

musí mít spojitou derivaci na $[a,b]$ (v krajních bodech jednostrannou), protože ostatní členy derivované rovnosti mají spojitou derivaci. □

Poznámka 13.3.5. Právě uvedený tvar Euler–Lagrangeovy rovnice má jednu nevýhodu, která je vidět na následujícím příkladu. Uvažujme funkcionál

$$F(y) = \int_0^1 y^2 {y'}^2 \, \mathrm{d}x \qquad \text{na } X.$$

Euler–Lagrangeova rovnice má tvar

$$f_y(x, y, y') - \frac{\mathrm{d}}{\mathrm{d}x} f_z(x, y, y') = 2y{y'}^2 - \frac{\mathrm{d}}{\mathrm{d}x}(2y^2 y') = 0 \qquad \text{na } [0,1].$$

Tento tvar není vhodný pro další práci (dokonce z něj ani není okamžitě vidět, že řešením je $y \equiv 0$, což je podle zadání zřejmě funkce, která dává globální minimum). Rádi bychom provedli úpravu

$$0 = 2yy'^2 - (2y^2y')' = 2yy'^2 - (4yy'^2 + 2y^2y'') = -2y(y'^2 + yy'') = -2y(yy')'$$

a tím dostali rovnici, se kterou už se dá pracovat. Zmíněnou úpravu však obecně provést nemůžeme, neboť nevíme, zda existuje y''.

Slabinou Euler–Lagrangeovy rovnice je tedy předpoklad $y_0 \in C^2([a,b])$, který obecně nemusí extremála splňovat.

Věta 13.3.6 (O regularitě minimizéru)**.** *Nechť $f \in C^2([a,b] \times \mathbb{R}^2)$, $y_0 \in M$ je stacionárním bodem funkcionálu F a $x_0 \in (a,b)$ je takové, že*

$$f_{zz}(x_0, y_0(x_0), y_0'(x_0)) \neq 0.$$

Pak existuje $\delta > 0$ takové, že $y_0 \in C^2((x_0 - \delta, x_0 + \delta))$.

Důkaz. Pokud je $y_0 \in M$ stacionárním bodem, podle předchozí věty splňuje Euler–Lagrangerovu rovnici, jejíž integrací dostáváme $\alpha \in \mathbb{R}$ splňující

$$\alpha \equiv f_z(x, y_0(x), y_0'(x)) - \int_a^x f_y(t, y_0(t), y_0'(t))\,\mathrm{d}t \qquad \text{na } [a,b].$$

Definujme nyní funkci $\Psi\colon [a,b] \times \mathbb{R}$ předpisem

$$\Psi(x,w) = f_z(x, y_0(x), w) - \int_a^x f_y(t, y_0(t), y_0'(t))\,\mathrm{d}t - \alpha.$$

Zřejmě $\Psi(x_0, y_0'(x_0)) = 0$, $\Psi \in C^1((a,b) \times \mathbb{R})$ (podle řetízkového pravidla, $f \in C^2([a,b] \times \mathbb{R}^2)$ a $y_0 \in C^1([a,b])$) a

$$\frac{\partial \Psi}{\partial w}(x_0, y_0'(x_0)) = f_{zz}(x, y_0(x_0), y_0'(x_0)) \neq 0.$$

Můžeme proto aplikovat Větu o implicitní funkci (základní verze), tedy Větu 12.4.2, a dostáváme $\delta_1, \Delta > 0$ taková, že pro každé $x \in (x_0 - \delta_1, x_0 + \delta_1)$ existuje právě jedno $w =: \varphi(x) \in (y'(x_0) - \Delta, y'(x_0) + \Delta)$ splňující $\Psi(x, \varphi(x)) = 0$ a $\varphi \in C^1((x_0 - \delta_1, x_0 + \delta_1))$. Protože funkce $w\colon x \mapsto y_0'(x)$ také splňuje $\Psi(x, w(x)) = 0$, z její spojitosti ($y_0 \in C^1([a,b])$) a jednoznačnosti dané Větou o implicitní funkci (základní verze), tedy Větou 12.4.2, dostáváme $\delta \in (0, \delta_1)$ takové, že

$$\varphi(x) = w(x) \qquad \text{na } (x_0 - \delta, x_0 + \delta).$$

Protože na levé straně máme funkci z $C^1((x_0 - \delta, x_0 + \delta))$, máme takovou funkci i na straně pravé, a proto $y_0 \in C^2((x_0 - \delta, x_0 + \delta))$. □

Příklad 13.3.7. (i) Uvažujme funkcionál

$$F(y) = \int_0^1 xy^2(x)\,\mathrm{d}x \qquad \text{s okrajovými podmínkami } y(0) = 0 \text{ a } y(1) = 0.$$

Euler–Lagrangeova rovnice má tvar

$$f_y(x, y, y') - \frac{\mathrm{d}}{\mathrm{d}x} f_z(x, y, y') = 2xy = 0.$$

Jediným řešením je zřejmě $y \equiv 0$ (globální minimum). Globální maximum existovat nemůže, jak ukazují funkce $x \mapsto C \sin(\pi x)$.

Pokud bychom pracovali na množině $M = \{y \in C^1([0,1]) \colon y(0) = 0, y(1) = 1\}$, Euler–Lagrangeova rovnice má opět tvar $2xy = 0$ a nám se jí díky počátečním podmínkám nepodaří splnit. Tedy žádná extremála neexistuje. Dokonce se dá ukázat přímo, že neexistuje globální minimum, neboť funkce

$$y_k(x) := \begin{cases} 0 & \text{pro } x \in [0, 1 - 2^{-k}] \\ \cos^4(\pi 2^{k-1}(x-1)) & \text{pro } x \in [1 - 2^{-k}, 1] \end{cases}$$

splňují $F(y_k) \to 0$, zatímco pro žádnou funkci z M nedostaneme $F(y) = 0$.
(ii) Uvažujme funkcionál

$$J(y) = \int_a^b y'^2\,\mathrm{d}x.$$

Pak Euler–Lagrangeova rovnice má tvar

$$f_y(x, y, y') - \frac{\mathrm{d}}{\mathrm{d}x} f_z(x, y, y') = 0 - \frac{\mathrm{d}}{\mathrm{d}x}(2y') = 0 \qquad \text{na } [a, b].$$

Zde můžeme jednak postupovat bez využití Věty o regularitě minimizéru (Věta 13.3.6), když si uvědomíme, že naše Euler–Lagrangeova rovnice je ekvivalentní podmínce

$$-2h' = C \qquad \text{na } [a, b]$$

a odtud $h = -\frac{C}{2}x + D$ (konstanty C, D budou zde jednoznačně určeny okrajovými podmínkami z definice množiny M). Větu o regularitě minimizéru (Věta 13.3.6) zde použít můžeme, neboť $f_{zz}(x, y, y') = 2 > 0$ (zde je velice příjemné, že je podmínka z Věty o regularitě minimizéru splněna pro všechna $y \in M$). To nám dává diferenciální rovnici

$$-2y'' = 0,$$

kterou umíme snadno vyřešit.
(iii) Uvažujme funkcionál

$$J(y) = \int_a^b y'(1 + x^2 y')\,\mathrm{d}x.$$

Povšimněme si, že $f_{zz}(x,y,y') = 2x^2$. Pokud bychom věděli, že $0 \notin (a,b)$, mohli bychom použít Větu o regularitě minimizéru (Věta 13.3.6). V opačném případě musíme postupovat opatrně

$$f_y(x,y,y') - \frac{\mathrm{d}}{\mathrm{d}x} f_z(x,y,y') = 0 - \frac{\mathrm{d}}{\mathrm{d}x}(1 + 2x^2 y') = 0 \qquad \text{na } [a,b].$$

To je ekvivalentní

$$1 + 2x^2 y' = C \qquad \Longleftrightarrow \qquad x^2 y' = \frac{C-1}{2} =: D.$$

Pokud $0 \notin (a,b)$, máme

$$y' = \frac{D}{x^2} \qquad \Longleftrightarrow \qquad y = -\frac{D}{x} + E,$$

kde $E \in \mathbb{R}$. Pokud $0 \in (a,b)$, musí platit $D = 0$. To zase implikuje, že $y' = 0$ na $[a,0) \cup (0,b]$. Protože $y \in C^1([a,b])$ celkově dostáváme $y \equiv E$.
(iv) Uvažujme funkcionál

$$J(y) = \int_{-1}^{1} x^4 {y'}^2 \,\mathrm{d}x \qquad \text{s okrajovými podmínkami } y(-1) = -1 \text{ a } y(1) = 1.$$

Platí $f_{zz}(x,y,y') = 2x^4$, a proto nemůžeme použít Větu o regularitě minimizéru (Věta 13.3.6) v počátku. Opět se dá postupovat následovně

$$f_y(x,y,y') - \frac{\mathrm{d}}{\mathrm{d}x} f_z(x,y,y') = 0 - \frac{\mathrm{d}}{\mathrm{d}x}(2x^4 y') = 0 \qquad \text{na } [-1,1].$$

To je ekvivalentní

$$2x^4 y' = C.$$

Mimo počátek dostáváme $y = -\frac{-\frac{C}{6}}{x^3} + D$. Funkce tohoto typu je možné slepovat v počátku (hledáme funkci z $C^1([-1,1])$) jen pokud $C = 0$, ale pak dostáváme konstantní řešení, které nesplňuje okrajové podmínky. Celkově neexistuje žádná extremála a funkcionál F na M nenabývá lokálního minima (ani globálního, podobně pro maximum).
(v) Uvažujme funkcionál

$$J(y) = \int_{-1}^{1} \left(2y^2 + x^2 {y'}^2\right) \mathrm{d}x \quad \text{s okrajovými podmínkami } y(-1) = -1 \text{ a } y(1) = 1.$$

Pak Euler–Lagrangeova rovnice má tvar

$$f_y(x,y,y') - \frac{\mathrm{d}}{\mathrm{d}x} f_z(x,y,y') = 4y - \frac{\mathrm{d}}{\mathrm{d}x}(2x^2 y') = 0 \qquad \text{na } [-1,1].$$

Protože $f_{zz}(x,y,y') = 2x^2$, nemůžeme použít Větu o regularitě minimizéru (Věta 13.3.6) v počátku. Použijme ji alespoň zvlášť na intervalech $(-1,0)$ a $(0,1)$. Na těchto intervalech dostáváme

$$4y - 4xy' - 2x^2 y'' = 0 \qquad \Longleftrightarrow \qquad x^2 y'' + 2xy'' - 2y = 0.$$

Jedná se o Eulerovu rovnici, jejíž obecné řešení má v našem případě tvar

$$y = \alpha x + \frac{\beta}{x^2}.$$

Protože extremála musí být z prostoru $C^1([a,b])$, z uvedených řešení můžeme v počátku slepit jen dvojici lineárních funkcí se stejnou směrnicí a díky okrajovým podmínkám celkově dostáváme jedinou extremálu $y_0 = x$ na $[-1,1]$.
(vi) Uvažujme funkcionál

$$F(y) = \int_0^1 x{y'}^2(x)\,\mathrm{d}x \qquad \text{s okrajovými podmínkami } y(0)=0 \text{ a } y(1)=0.$$

Euler–Lagrangeova rovnice má tvar

$$f_y(x,y,y') - \frac{\mathrm{d}}{\mathrm{d}x} f_z(x,y,y) = 0 - \frac{\mathrm{d}}{\mathrm{d}x}(2xy') = 0.$$

Odtud $y' = \frac{C_1}{2x}$, a proto

$$y = \frac{C_1}{2}\log x + C_2.$$

Podmínku $y \in C^1([0,1])$ spolu s okrajovými podmínkami zde splňuje jen $y_0 \equiv 0$. Zřejmě se jedná o globální minimum. Snadno si čtenář sám zkonstruuje příklady ukazující, že pro žádné okrajové podmínky se nenabývá globálního maxima. Obecně v případě okrajových podmínek $y(0) = y(1) = C_2$ se globální minimum nabývá pro $y \equiv C_2$. Pokud máme různé okrajové podmínky $y(0) \neq y(1)$, globálního minima se nenabývá. To se dá také ukázat pomocí následující konstrukce. Předně si povšimněme, že žádná $C^1([0,1])$-funkce nemůže zároveň splňovat takové okrajové podmínky a $F(y) = 0$. Bez újmy na obecnosti v dalším předpokládejme, že máme okrajové podmínky $y(0) = 0$ a $y(1) = 1$ (pokud pro tyto podmínky umíme zkonstruovat posloupnost $\{y_k\} \subset X$ splňující $F(y_k) \to 0$, umíme to pro jakoukoliv sadu různých okrajových podmínek, stačí si rozmyslet vztah mezi $F(y)$ a $F(\alpha y + \beta)$). Nechme se inspirovat řešením Euler–Lagrangeovy rovnice a definujme spojité funkce

$$y_k = \begin{cases} 0 & \text{pro } x \in [0, \frac{1}{k}] \\ \frac{\log x}{\log k} + 1 & \text{pro } x \in [\frac{1}{k}, 1]. \end{cases}$$

Pak platí

$$\begin{aligned} \int_0^{\frac{1}{k}} x{y_k'}^2(x)\,\mathrm{d}x + \int_{\frac{1}{k}}^1 x{y_k'}^2(x)\,\mathrm{d}x &= 0 + \int_{\frac{1}{k}}^1 x\Big(\frac{1}{x\log k}\Big)^2\,\mathrm{d}x \\ &= \Big(\frac{1}{\log k}\Big)^2 \int_{\frac{1}{k}}^1 \frac{1}{x}\,\mathrm{d}x = \frac{1}{\log k} \overset{k\to\infty}{\to} 0. \end{aligned}$$

Není těžkě si vymyslet drobnou modifikaci funkce y_k na malém okolí bodu $\frac{1}{k}$, aby vznikla funkce z $C^1([0,1])$ a hodnoty integrálů spočítané výše se změnily jen nepatrně.

Poznámka 13.3.8. (i) Euler–Lagrangeova rovnice má klíčové postavení ve *variačních metodách* důkazu existence řešení diferenciálních rovnic. Zde se postupuje obráceně. Máme zadanou diferenciální rovnici. Pokusíme se nalézt jí odpovídající funkcionál (pro který je naše rovnice Euler–Lagrangeovou rovnicí) a ukázat, že tento funkcionál má stacionární bod dostatečné hladkosti. Pak okamžitě dostáváme, že tento stacionární bod je řešením zkoumané rovnice. Například v poslední části předchozího příkladu má Euler–Lagrangeova rovnice tvar (po rozderivování)

$$2y' + 2xy'' = 0 \qquad y(0) = y(1) = 0.$$

Protože odpovídající funkcionál

$$F(y) = \int_0^1 x{y'}^2(x)\,\mathrm{d}x$$

má na množině $\{y \in C^1([0,1])\colon y(0) = y(1) = 0\}$ globální minimum v $y \equiv 0$ ($C^2([0,1])$-funkce), okamžitě dostáváme, že tato funkce řeší uvažovanou diferenciální rovnici.

(ii) V praxi se při aplikaci variačních metod podaří nalézt stacionární bod jen málokdy. Většinou se jen dokáže jeho existence za pomoci geometrických vlastností zkoumaného funkcionálu (pak se aplikují numerické metody pro přibližná řešení diferenciálních rovnic). K nejdůležitějším z těchto vlastností patří konvexita (sama o sobě ovšem existenci minima nezaručuje, uvažte konvexní funkci exp) doprovázená dalšími pojmy a technikami, jako jsou *koercivita*, *slabá polospojitost zdola*, *slabá konvergence* či dalšími pojmy, které přesahují teorii těchto skript. Hlavním problémem, se kterým se musí příslušná teorie vyrovnat, je skutečnost, že v nekonečnědimenzionálních prostorech nejsou omezené množiny kompaktní, čímž přicházíme o možnost vybírání konvergentních posloupností z posloupností omezených.

13.3.2 Euler–Lagrangeova rovnice pro funkcionály speciálních typů

Řešení Euler–Lagrangeovy rovnice může být velice obtížné. Jistá zjednodušení, nebo alespoň alternativní postupy, se nabízejí v situacích, kdy funkce f nezávisí na jedné nebo více proměnných. Těmto případům se nyní budeme věnovat.

Nejprve se zabývejme případem, že funkce f nezávisí na proměnné z (neboli $f = f(x, y)$). Euler–Lagrangeova rovnice má jednodušší tvar

$$f_y(x, y) = 0,$$

což je rovnice v níž se žádné derivace hledané funkce nevyskytují. Jedná se tedy o podstatně jednodušší úlohu než v obecném případě. Řešení však nemusí být vždy snadné, neboť hledaná funkce je zde zadaná implicitně. Poznamenejme ještě, že v tomto případě není nutné požadovat, aby $f \in C^2([a,b] \times \mathbb{R}^2)$, ale stačí $f \in C^1([a,b] \times \mathbb{R}^2)$. Zde totiž k odvození výše uvedené podmínky stačí kombinovat první část Lemmatu o tvaru $\delta\Phi(u;h)$ a $\delta^2\Phi(u;h,h)$ (Lemma 13.3.1) a Fundamentální lemma variačního počtu (Lemma 13.3.3). S typickým funkcionálem tohoto typu jsme se setkali třeba v Příkladu 13.2.3. Navíc v těchto případech platí zjednodušená formule

$$\delta^2\Phi(u;h,h) = \int_a^b f_{yy}(x,y)h^2\,\mathrm{d}x.$$

To nám často umožňuje přímo používat výsledky abstraktní teorie týkající se druhého diferenciálu.

Pokud funkce f nezávisí na proměnné y (neboli $f = f(x,y')$), Euler–Lagrangeova rovnice má tvar

$$-\frac{\mathrm{d}}{\mathrm{d}x} f_z(x,y') = 0$$

a její řešení se dá popsat podmínkou $f_z(x,y_0') = C$. Tím jsme dostali diferenciální rovnici prvního řádu.

Poznamenejme ještě, že v tomto případě navíc platí zjednodušená formule

$$\delta^2\Phi(u;h,h) = \int_a^b f_{zz}(x,y')h'^2\,\mathrm{d}x,$$

která je opět často vhodná pro aplikaci abstraktní teorie týkající se druhého diferenciálu.

Příklad 13.3.9. Mezi křivkami, které lze popsat grafem C^1-funkce, hledejme tu, která spojuje bod (a,A) s bodem (b,B), kde $a < b$, a je z takových křivek nejkratší. Minimalizujeme tedy funkcionál

$$F(y) = \int_a^b \sqrt{1+y'^2(x)}\,\mathrm{d}x$$

přes funkce z $C^1([a,b])$ splňující $y(a) = A$ a $y(b) = B$. Pro řešení Euler–Lagrangeovy rovnice proto máme

$$\frac{y'}{\sqrt{1+y'^2(x)}} = C \qquad \Longleftrightarrow \qquad y'^2 = C(1+y'^2).$$

Odtud vidíme, že extremálami mohou být jen afinní funkce, což spolu s okrajovými podmínkami dává jediného kandidáta na lokální minimum $y_0(x) = A + \frac{B-A}{b-a}(x-a)$. Dále pro libovolné y z uvažované třídy funkcí platí

$$\delta^2\Phi(u;h,h) = \int_a^b f_{zz}(x,y')h'^2\,\mathrm{d}x = \int_a^b \frac{1}{(1+y'^2(x))^{\frac{3}{2}}}h'^2\,\mathrm{d}x \geq 0.$$

Díky tomu lze aplikovat Lagrangeovu postačující podmínku (Věta 13.2.14) a dostáváme, že naše extremála je bodem lokálního minima. Dokonce se dá dvojím zderivováním snadno ověřit, že funkce $t \mapsto \sqrt{1+t^2}$ je konvexní. Proto je funkcionál F konvexní a jedná se o globální minimum díky Větě o postačující podmínce pro konvexní funkcionál (Věta 13.2.19).

Dalším důležitým případem je $f = f(y, y')$. Zde máme následující popis extremál.

Tvrzení 13.3.10 (Nutná podmínka řešení Euler–Lagrangeovy rovnice pro autonomní úlohu)**.** *Jestliže $y_0 \in M \cap C^2([a,b])$ řeší Euler–Lagrangeovu rovnici, pak*

$$f(y_0, y_0') - y_0' f_z(y_0, y_0') \equiv C.$$

Důkaz. Přenásobíme Euler–Lagrangeovu rovnici výrazem y_0' a dostáváme

$$\begin{aligned} 0 &= y_0' f_y(y_0, y_0') - y_0' \frac{\mathrm{d}}{\mathrm{d}x} f_z(y_0, y_0') \\ &= y_0' f_y(y_0, y_0') - \frac{\mathrm{d}}{\mathrm{d}x}(y_0' f_z(y_0, y_0')) + y_0'' f_z(y_0, y_0') \\ &= \frac{\mathrm{d}}{\mathrm{d}x}(f(y_0, y_0')) - \frac{\mathrm{d}}{\mathrm{d}x}(y_0' f_z(y_0, y_0')). \end{aligned}$$

Odtud

$$f(y_0, y_0') - y_0' f_z(y_0, y_0') \equiv C.$$

□

Poznámka 13.3.11. (i) Jedná se skutečně pouze o podmínku nutnou, neboť přenásobení výrazem y_0' v důkazu není ekvivalentní operací.
(ii) Výhoda právě dokázaného výsledku spočívá v tom, že standardní Euler–Lagrangeovu rovnici, což je rovnice druhého řádu, převádí na rovnici prvního řádu, čímž se zvyšuje pravděpodobnost, že úlohu budeme schopni vyřešit. V některých případech je však přechod k nové rovnici nevýhodný, jak uvidíme na příkladech uvedených níže.
(iii) Přestože je výsledná rovnice prvního řádu, při jejím odvození se používala podmínka $y_0 \in M \cap C^2([a,b])$. Využití nové metody proto opět vyžaduje předchozí aplikaci Věty o regularitě minimizéru (Věta 13.3.6).
(iv) Této metodě se ve fyzice často říká metoda hledání prvního integrálu.

Příklad 13.3.12. (i) V případě výše zmíněné úlohy o nejkratší spojnici dvojice bodů jsme pracovali s funkcionálem

$$F(y) = \int_a^b \sqrt{1 + y'^2(x)}\, \mathrm{d}x.$$

Protože $f_{zz}(y, y') = \frac{1}{(1+y'^2(x))^{\frac{3}{2}}} > 0$, můžeme díky Větě o regularitě minimizéru (Věta 13.3.6) používat jak standardní Euler–Lagrangeovu rovnici, která nám dala podmínku

$$\frac{y'}{\sqrt{1 + y'^2(x)}} = C,$$

tak alternativní rovnici

$$f(y, y') - y' f_z(y, y') = \sqrt{1 + y'^2} - \frac{y'^2}{\sqrt{1 + y'^2(x)}} = \frac{1}{\sqrt{1 + y'^2(x)}} \equiv C,$$

ze které je snáze vidět, že y' musí být konstantní.
(ii) Uvažujme funkcionál

$$F(y) = \int_a^b \frac{1+y^2}{{y'}^2}\,\mathrm{d}x.$$

Protože $f_{zz} = 6\frac{1+y^2}{{y'}^4}$, případná extremála patří do $C^2((a,b))$. Euler–Lagrangeova rovnice zde má tvar

$$\begin{aligned} 0 = f_y(x,y,y') - \frac{\mathrm{d}}{\mathrm{d}x} f_z(x,y,y) &= \frac{2y}{{y'}^2} + 2\frac{\mathrm{d}}{\mathrm{d}x}\Big(\frac{1+y^2}{{y'}^3}\Big) \\ = \frac{2y}{{y'}^2} + 2\frac{2y{y'}^2 - 3(1+y^2)y''}{{y'}^4} &= \frac{6}{{y'}^4}\Big(y{y'}^2 - (1+y^2)y''\Big). \end{aligned}$$

Závorka úplně napravo dává

$$y{y'}^2 - (1+y^2)y'' = 0,$$

což není žádný ze základních typů diferenciálních rovnic, které umíme řešit. Na druhou stranu, Tvrzení o nutné podmínce řešení Euler–Lagrangeovy rovnice pro autonomní úlohu (Tvrzení 13.3.10) nám dává rovnici

$$C \equiv f(y,y') - y' f_z(y,y') = \frac{1+y^2}{{y'}^2} + 2y'\frac{1+y^2}{{y'}^3} = 3\frac{1+y^2}{{y'}^2}.$$

To mohou řešit jen funkce splňující

$$y' = K\sqrt{1+y^2}, \quad K \in \mathbb{R}\setminus\{0\},$$

a máme řešitelnou rovnici se separovanými proměnnými (integrace povede na funkci argcosh).
(iii) Uvažujme funkcionál

$$F(y) = \int_a^b \big({y'}^2 - yy' + y^2\big)\,\mathrm{d}x.$$

Protože $f_{zz} = 2$, případná extremála patří do $C^2((a,b))$. Euler–Lagrangeova rovnice zde má tvar

$$\begin{aligned} 0 &= f_y(x,y,y') - \frac{\mathrm{d}}{\mathrm{d}x} f_z(x,y,y) = -y' + 2y - \frac{\mathrm{d}}{\mathrm{d}x}(2y' - y) \\ &= -y' + 2y - (2y'' - y') = -2(y'' - y). \end{aligned}$$

Extremálami proto jsou funkce tvaru

$$y_0 = C\mathrm{e}^x + D\mathrm{e}^{-x}.$$

Přístup využívající Tvrzení o nutné podmínce řešení Euler–Lagrangeovy rovnice pro autonomní úlohu (Tvrzení 13.3.10) nám dává rovnici

$$C \equiv f(y,y') - y' f_z(y,y') = {y'}^2 - yy' + y^2 - y'(2y' - y) = -{y'}^2 + y^2 = -{y'}^2 + y^2.$$

Tato rovnice rozhodně není řešitelsky příjemnější než výše uvedený standardní tvar Euler–Lagrangeovy rovnice. Navíc tato rovnice není ekvivalentní s Euler–Lagrangeovou rovnicí, neboť připouští kupříkladu všechna konstantní řešení.
(iv) V případě, že f závisí jen na poslední proměnné z, Euler–Lagrangeova rovnice okamžitě dává

$$f_z(y') = C.$$

Naproti tomu Tvrzení o nutné podmínce řešení Euler–Lagrangeovy rovnice pro autonomní úlohu (Tvrzení 13.3.10) jednak vyžaduje vyšší regularitu extremály a navíc dává tvar

$$f(y') - y' f_z(y') = C,$$

který bývá často složitější a navíc obecně není ekvivalentní s Euler–Lagrangeovou rovnicí.

13.3.3 Klasifikace extremál založená na chování druhého diferenciálu

Nyní se budeme zabývat podmínkami odvozenými od chování druhého diferenciálu.

Věta 13.3.13 (Lagrangeova nutná podmínka pro integrální funkcionál)**.** *Nechť $f \in C^2([a,b] \times \mathbb{R}^2)$ a $y_0 \in M$ je bodem lokálního minima funkcionálu F. Pak*

$$f_{zz}(x, y_0(x), y_0'(x)) \geq 0 \qquad \textit{na } [a,b].$$

Důkaz. Pokud by existovalo $x_0 \in (a,b)$ takové, že

$$-A := f_{zz}(x, y_0(x_0), y_0'(x_0)) < 0,$$

pro $\varepsilon > 0$ dostatečně malé bychom z volby

$$h = \begin{cases} \varepsilon \cos^2(\frac{x-x_0}{\varepsilon}) & \text{pro } |x - x_0| \leq \frac{\pi}{2}\varepsilon \\ 0 & \text{jinak} \end{cases}$$

dostali pomocí Lemmatu o tvaru $\delta^2\Phi(y; h, h)$ a $\delta^2\Phi(u; h, h)$ (Lemma 13.3.1) a předpokladů $f \in C^2([a,b] \times \mathbb{R}^2)$ a $y_0 \in C^1([a,b])$ (na konci výpočtu používáme ještě jednoduchou substituci)

$$\begin{aligned}
\delta^2\Phi(u_0; h, h) &= \int_a^b \Big(f_{yy}(x, y_0, y_0')h^2 + 2f_{yz}(x, y_0, y_0')hh' + f_{zz}(x, y_0, y_0'){h'}^2\Big)\,\mathrm{d}x \\
&\leq \int_{x_0 - \frac{\pi\varepsilon}{2}}^{x_0 + \frac{\pi\varepsilon}{2}} \Big(Ch^2 + C|h||h'| - \frac{A}{2}{h'}^2\Big)\,\mathrm{d}x \\
&\leq \int_{x_0 - \frac{\pi\varepsilon}{2}}^{x_0 + \frac{\pi\varepsilon}{2}} \big(C\varepsilon^2 + C\varepsilon\big)\,\mathrm{d}x - \int_{x_0 - \frac{\pi\varepsilon}{2}}^{x_0 + \frac{\pi\varepsilon}{2}} \frac{A}{2}\sin^2(2\tfrac{x-x_0}{\varepsilon})\,\mathrm{d}x \\
&= \pi\varepsilon(C\varepsilon^2 + C\varepsilon) - \frac{A}{2}\varepsilon \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^2(2t)\,\mathrm{d}t < 0.
\end{aligned}$$

To je spor s Lagrangeovou nutnou podmínku z abstraktního případu, tedy s Větou 13.2.12. Celkově proto máme

$$f_{zz}(x, y_0(x), y_0'(x)) \geq 0 \qquad \text{na } (a, b).$$

Navíc levá strana předchozí nerovnosti je díky předpokladům spojitá na $[a, b]$, a proto dokazovaná nerovnost platí na celém $[a, b]$. □

Věta 13.3.14 (Legendreova postačující podmínka)**.** *Nechť $f \in C^2([a, b] \times \mathbb{R}^2)$ a $y_0 \in M$ je stacionárním bodem funkcionálu F. Jestliže existují $\alpha, \delta > 0$ taková, že*

$$\delta^2\Phi(u_0; h, h) \geq \alpha \|h\|^2_{C^1([a,b])} \qquad \textit{pro všechna } h \in X \textit{ splňující } \|h\|_{C^1([a,b])} \leq \delta,$$

pak F má v bodě y_0 ostré lokální minimum.

Důkaz. Zafixujme $h \in X$ splňující $0 < \|h\|_{C^1([a,b])} \leq \delta$ a definujme

$$\varphi(t) := F(y_0 + th) \qquad \text{pro } t \in \mathbb{R}.$$

Pak z Taylorova rozvoje máme (v následujícím výpočtu číslo $\theta \in (0, 1)$ závisí na t a h)

$$\begin{aligned} F(y_0 + th) - F(y_0) &= \varphi(t) - \varphi(0) = \varphi'(0) + \frac{1}{2}\varphi''(\theta t) \\ &= 0 + \frac{1}{2}\varphi''(0) + \frac{1}{2}(\varphi''(\theta t) - \varphi''(0)) \\ &\geq \frac{\alpha}{2}\|h\|^2_{C^1([a,b])} + \frac{1}{2}(\varphi''(\theta t) - \varphi''(0)). \end{aligned}$$

Dokazovaný výsledek proto plyne z následujícího odhadu platného pro $|t|$ dostatečně malé ($|t| \leq t_0$, kde t_0 nezávisí na volbě h; používáme Lemma o tvaru $\delta\Phi(y_0; h)$ a $\delta^2\Phi(y_0; h, h)$ (Lemma 13.3.1) a stejnoměrnou spojitost druhých parciálních derivací funkce f na kompaktech)

$$\begin{aligned} |\varphi''(\theta t) - \varphi''(0)| \leq \int_a^b \Big[&|f_{yy}(x, y_0 + \theta th, y_0' + \theta th') - f_{yy}(x, y_0, y_0')|h^2 \\ &+ 2|f_{yz}(x, y_0 + \theta th, y_0' + \theta th') - f_{yz}(x, y_0, y_0')||h||h'| \\ &+ |f_{zz}(x, y_0 + \theta th, y_0' + \theta th') - f_{zz}(x, y_0, y_0')|{h'}^2 \Big] \,\mathrm{d}x \\ \leq \frac{\alpha}{2}\|h\|^2_{C^1([a,b])}. \end{aligned}$$

□

Věta 13.3.15 (Lagrangeova postačující podmínka pro integrální funkcionál)**.** *Nechť $f \in C^2([a, b] \times \mathbb{R}^2)$ a $y_0 \in M$ je stacionárním bodem funkcionálu F. Jestliže existuje $\delta > 0$ takové, že pro každé $h \in X$ splňující $\|h\|_{C^1([a,b])} \leq \delta$ má funkce*

$$\varphi(t) := F(y_0 + th)$$

vlastnost

$$\varphi''(t) \geq 0 \qquad \textit{pro } t \in (0, 1),$$

pak F má v bodě y_0 lokální minimum. V případě, že předchozí vlastnost platí s ostrou nerovností, jedná se o ostré lokální minimum.

Důkaz. Podle Lagrangeovy věty o přírůstku funkce (Věta 6.3.3) máme

$$F(y_0 + h) - F(y_0) = \varphi(1) - \varphi(0) = \varphi'(\xi) = \varphi'(\xi) - \varphi'(0) = \varphi''(\theta\xi)\xi,$$

z čehož vše plyne. □