Obsah

Předmluva

Tato publikace je určena především pro studenty druhého ročníku specializace matematika.

Hlavní částí publikace je kapitola 2 s názvem „Diferenciální počet funkcí více proměnných“. Tato část má ráz učebnice, důkazy jsou úplné a je ji v podstatě možno studovat zcela samostatně. Má obvyklý tradiční obsah, který mírně překračuje sylabus přednášky. Jak je v novějších učebnicích běžné, základním pojmem je diferenciál (derivace) zobrazení mezi eukleidovskými prostory a teorie vázaných extrémů je založena na základech teorie k-rozměrných ploch v $\mathbb{R}^n$. Výklad je veden na základě přednášek, které jsem měl několikrát na MFF UK. Při jejich přípravě jsem používal řadu zdrojů, zejména [D II], [Fi], [Zo], [Fe] a [KF].

Výklad diferenciálního počtu funkcí více proměnných se již tradičně opírá o elementární teorii metrických prostorů, jejíž výklad lze nalézt například v [D II] a [Če]. Také proto kapitola 1 obsahuje pouze přehled základních výsledků teorie metrických prostorů a je míněna hlavně jako zdroj odkazů pro důkazy v kapitole 2. Důkazy vět jsou uvedeny jen výjimečně. Výjimkou je výklad o lineárních (a multilineárních) zobrazeních, který není obsažen v [D II], a proto je veden většinou s důkazy.

Kapitola 3 obsahuje úvod do diferenciálního počtu v normovaných lineárních prostorech. Jde skutečně jen o úvod do teorie; úplnější výklad lze nalézt například v [Ca] a [FM]. Některé důkazy jsou pouze naznačeny, což je vždy zmíněno. Při psaní jsem vycházel hlavně z [Ca], ale také ze [Zo], [FM] a [Fe].

Kapitola 4 obsahuje stručnou teorii Fourierových řad a velmi krátký (klasicky vedený) úvod do teorie Fourierova integrálu a Fourierovy transformace.

Kapitola 5 obsahuje jistou „minimální verzi“ teorie plošného a křivkového integrálu v $\mathbb{R}^n$. Záměrem bylo

a) co nejjednodušeji přesně formulovat a dokázat Gaussovu a Greenovu větu tak, aby je bylo možno přesně a pohodlně aplikovat na konkrétní „po částech hladké“ plochy a křivky, a

b) vysvětlit (pomocí klasického postupu) čtenáři intuitivní smysl základních pojmů a vět.

Výklad vychází do značné míry z [Kop], [ČM], [LM], [Ru1], [Zo] a z jedné mé výběrové přednášky na MFF UK. K jeho pochopení je nutné (a stačí) znát teorii míry a integrálu v rozsahu přednášeném ve 3. semestru oboru matematika.

Gaussova (a Greenova) věta je v hlavním textu dokázána pouze v obvyklém speciálním případě. Důkaz obecného případu je obsažen v dodatku.

Poslední kapitola 6 obsahuje řadu dodatků, mj. ohledně používané terminologie.

I poměrně systematické kapitoly 2 a 4 jsou míněny hlavně jako doplněk k přednáškám. Jejich obsah jen mírně přesahuje to, co lze stihnout na přednášce, a také výklad není veden v plné obecnosti. V žádném případě nemůže ovšem nahradit obsažné učebnice jako jsou např. [D II], [J II], [Fi] nebo [Zo].

Celý výklad publikace je ovšem silně ovlivněn tradicí na MFF UK (utvářenou hlavně dílem V. Jarníka, ale také například skripty a přednáškami J. Maříka, J. Lukeše, J. Miloty a J. Kopáčka).

Děkuji kolegům doc. P. Holickému, prof. M. Huškovi, dr. J. Jelínkovi, dr. O. Kalendovi, prof. J. Lukešovi, prof. J. Malému, doc. J. Staré, doc. J. Veselému, dr. M. Zelenému a zejména dr. J. Kolářovi, kteří četli některé části rukopisu, upozornili mě na velké množství nedostatků a dali mi řadu cenných rad.

Dále děkuji P. Charvátovi za zhotovení obrázků.

Budu vděčen za upozornění na chyby a kritické připomínky (např. na e-mailovou adresu zajicek ?? karlin.mff.cuni.cz); nalezené chyby budu zveřejňovat na své webové stránce.

Duben 2003

1. Metrické prostory

1.1 Pojem metrického prostoru

Hlavní náplní této publikace je diferenciální počet funkcí více proměnných. Je zcela zřejmé, že bez vyšetřování těchto funkcí se neobejdeme například v geometrii a ve fyzice. Například teplota T v bodě prostoru s kartézskými souřadnicemi x, y, z v čase t se popisuje funkcí 4 proměnných: $T = f(x, y, z, t)$.

Reálnou funkci n proměnných $f(x_1, \dots, x_n)$ přirozeně chápeme jako zobrazení z množiny $\mathbb{R}^n$ do $\mathbb{R}$. Přitom $\mathbb{R}^n$ je množina všech uspořádaných n-tic reálných čísel. Každou takovou n-tici zapisujeme symbolem $(x_1, \dots, x_n)$ a nazýváme ji bodem prostoru $\mathbb{R}^n$; reálná čísla $x_1, \dots, x_n$ nazýváme jeho složkami (nebo také souřadnicemi).

Chceme-li vybudovat teorii funkcí více proměnných, musíme pro ně ovšem nejprve definovat pojmy limity a spojitosti, které již v teorii funkcí jedné proměnné jsou zcela základní.

Tento úkol můžeme řešit tak, že si nejprve v $\mathbb{R}^n$ přirozeným způsobem zavedeme pojem vzdálenosti dvou bodů. Prostory $\mathbb{R}^2$ a $\mathbb{R}^3$ si představujeme jako „geometrickou rovinu“ a „geometrický (trojrozměrný) prostor“, ve kterých jsou zavedeny systémy kartézských souřadnic. Například prvek $(x, y, z) \in \mathbb{R}^3$ si představujeme jako bod „geometrického prostoru“ s kartézskými souřadnicemi (x, y, z). Poznatky analytické geometrie nám pak umožňují přirozeným způsobem definovat přímky, roviny a také vzdálenost dvou bodů v $\mathbb{R}^3$. Speciálně eukleidovskou vzdáleností dvou bodů $x = (x_1, x_2, x_3)$, $y = (y_1, y_2, y_3)$ nazýváme reálné číslo

$$\rho(x, y) := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + (x_3 - y_3)^2}.$$

Tím jsme vedeni k následující definici vzdálenosti dvou bodů v $\mathbb{R}^n$.

1.1 Definice. *Eukleidovskou vzdáleností dvou bodů $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ a $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ nazýváme reálné číslo*

$$\rho_2(x, y) := \sqrt{\sum_{k=1}^{n} (x_k - y_k)^2} = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

Množinu $\mathbb{R}^n$, na které uvažujeme eukleidovskou vzdálenost, nazýváme n-rozměrným eukleidovským prostorem.

1.2 Poznámka. Axiomatická výstavba geometrie je obsažena již v Eukleidově díle (3. stol. př. n. l.). Zcela přesná axiomatická teorie však vznikla až na přelomu 19. a 20. století.

Pomocí pojmu vzdálenosti pak můžeme přirozeně definovat pojem ε-ového okolí bodu $a \in \mathbb{R}^n$ jako $U_\varepsilon(a) \ := \{x \in \mathbb{R}^n\colon\ \rho_2(x,a) < \varepsilon\}$ a následně pojem spojitosti a limity funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ (a obecněji zobrazení $f\colon \mathbb{R}^n \to \mathbb{R}^k$). Základní vlastnosti limity a spojitosti lze pak přímočaře zobecnit z případu funkce jedné proměnné na případ funkcí více proměnných.

Zcela stejně však lze vybudovat teorii spojitosti a limity (a dalších důležitých pojmů) pro mnohem obecnější prostory X, než jsou eukleidovské prostory. Stačí, aby byl na množině X definován pojem vzdálenosti $\rho(x,y)$ bodů $x,y \in X$, který má několik velmi jednoduchých a přirozených vlastností.

Budeme proto pracovat v těchto obecnějších, tzv. metrických prostorech. Teorie metrických prostorů má mnoho různorodých důležitých aplikací a lze říci, že je jedním ze základů moderní analýzy. Přitom vybudování základů teorie metrických prostorů není podstatně složitější, než příslušná teorie v eukleidovských prostorech.

1.3 Poznámka. V některých partiích moderní analýzy se i pojem metrického prostoru ukazuje být málo obecný a pracuje se s obecnějším pojmem *topologického prostoru.* Poznamenejme, že v obecném topologickém prostoru není definována vzdálenost mezi jeho body, je však definován pojem spojitosti zobrazení mezi topologickými prostory. Pro více informací viz Poznámka 1.23 níže.

1.4 Definice. *Nechť X je množina a $\rho\colon X \times X \to \mathbb{R}$ je funkce taková, že pro všechna $x,y,z \in X$ platí:*

(i) $\rho(x,y) \geq 0 \quad$ a $\quad \rho(x,y) = 0 \iff x = y;$
(ii) $\rho(x,y) = \rho(y,x);$
(iii) $\rho(x,z) \leq \rho(x,y) + \rho(y,z).$

Pak funkci ρ nazýváme metrikou na množině X a uspořádanou dvojici (X,ρ) nazýváme metrickým prostorem.

Metrický prostor je tedy jakákoliv množina X (případně i prázdná), na které je zadána metrika ρ. Pokud je ze souvislosti jasné, o kterou metriku jde, mluvíme často o metrickém prostoru X místo o (X,ρ). Číslo $\rho(x,y)$ se nazývá *vzdáleností* bodů $x,y \in X$. Někdy se metrika ρ nazývá také „vzdálenost na X".

Jak již bylo řečeno, metrické prostory se přirozeně vyskytují v matematice v řadě různých situací. Nyní krátce popíšeme tři z nich.

(i) Základní aplikací metrických prostorů v této publikaci je teorie funkcí více proměnných. K tomu potřebujeme mít na množině $\mathbb{R}^n$ metriku. Motivováni analytickou geometrií, zavedli jsme na $\mathbb{R}^n$ eukleidovskou vzdálenost, která je v jistém smyslu nejpřirozenější. Ještě jsme ale nedokázali, že eukleidovská vzdálenost je skutečně metrika, tj. splňuje axiomy metrického prostoru z Definice 1.4. To zdůvodníme v dalším oddílu pomocí skalárního součinu – užijeme teorii unitárních prostorů. V teorii funkcí více proměnných je však v některých případech pohodlnější pracovat s jinými metrikami na $\mathbb{R}^n$, zvláště s „maximovou" metrikou ρ_∞ a „součtovou" metrikou ρ_1:

$$\rho_\infty(x,y) \ := \max\{|x_1 - y_1|, \dots, |x_n - y_n|\}, \quad \rho_1(x,y) \ := |x_1 - y_1| + \dots + |x_n - y_n|.$$

Je zcela jednoduché dokázat, že jde skutečně o metriky. V Příkladu 1.10 snadno ukážeme, že určují stejný pojem spojitosti, jako metrika eukleidovská.

(ii) Geometrická intuice nám správně napovídá, že lze přirozeně definovat pojem „vnitřní vzdálenosti“ dvou bodů na „hladké ploše“ v $\mathbb{R}^3$ (tedy nikoliv vzdálenost „vzdušnou čarou“, ale vzdálenost, kterou je třeba „urazit po ploše“). Ukazuje se, že tato „vnitřní vzdálenost“ má vlastnosti metriky a takto definovaná metrika má značný význam v geometrii.

(iii) Velmi důležité je, že metriku lze přirozeným způsobem zavést na mnoha „nekonečně rozměrných“ množinách, které se přirozeně vyskytují v matematické analýze. Prvky těchto množin (prostorů) jsou většinou funkce nebo posloupnosti; tyto prostory si již nejsme schopni dostatečně jasně představit („ještě méně“ než $\mathbb{R}^n$ pro $n \geq 4$). Na tyto prostory jsme však schopni velmi úspěšně aplikovat teorii metrických prostorů. Přitom v teorii metrických prostorů jsme přirozeně vedeni naší geometrickou intuicí z roviny a prostoru ($\mathbb{R}^2$ a $\mathbb{R}^3$). Tím dochází k jisté „geometrizaci analýzy“, která nám velmi pomáhá pochopit složité analytické problémy. V některých těchto prostorech je možno dokonce přirozeným způsobem zavést i pojem skalárního součinu (a tedy i kolmosti a úhlu sevřeného vektory) a používat geometrickou intuici v ještě větším rozsahu.

Uvažujme například množinu X všech reálných funkcí definovaných a spojitých na intervalu $[0,1]$ a položme si otázku, jak přirozeně definovat vzdálenost (odchylku) dvou funkcí f, $g \in X$. Zcela přirozeně se nabízí „maximová vzdálenost“ definovaná předpisem

$$\rho_\infty(f,g) = \max\{|f(x) - g(x)|\colon\ x \in [0,1]\}.$$

Víme, že maxima se skutečně nabývá, a je snadné dokázat, že ρ_∞ je metrika na X. Tato metrika je analogická metrice ρ_∞ na $\mathbb{R}^n$. Na X však můžeme přirozeně definovat také metriky, které jsou analogické metrikám ρ_2 a ρ_1 na $\mathbb{R}^n$:

$$\rho_2(f,g) = \sqrt{\int_0^1 (f(x) - g(x))^2\,\mathrm{d}x}, \quad \rho_1(f,g) = \int_0^1 |f(x) - g(x)|\,\mathrm{d}x.$$

Tyto „integrální odchylky“ jsou také metrikami na X; na rozdíl od analogických metrik na $\mathbb{R}^n$ však tyto tři metriky určují rozdílné pojmy spojitosti.

1.2 Normované lineární prostory a unitární prostory

V mnoha důležitých případech se v analýze pracuje s metrickými prostory, které jsou zároveň lineárními (vektorovými) prostory. Přitom je často metrika translačně invariantní (tj. $\rho(x+a, y+a) = \rho(x, y)$), takže $\rho(x, y) = \rho(0, y-x)$; k určení metriky pak stačí znát velikost („normu“) $\|z\| := \rho(z, 0)$ libovolného vektoru z, přičemž tato funkce $\|\cdot\|$ má vlastnosti z následující definice.

1.5 Definice. (normovaný lineární prostor) *Nechť X je lineární (vektorový) prostor nad tělesem $\mathbb{T}$ reálných nebo komplexních čísel. Pak funkce $\|\cdot\|: X \to \mathbb{R}$, která každému $x \in X$ přiřazuje číslo $\|x\| \in \mathbb{R}$, se nazývá norma na X, jestliže pro každé dva body x, $y \in X$ a každé číslo $\alpha \in \mathbb{T}$ platí*

(i) $\|x\| \geq 0$ *a* $\|x\| = 0 \iff x = 0$;
(ii) $\|\alpha \cdot x\| = |\alpha| \cdot \|x\|$;
(iii) $\|x + y\| \leq \|x\| + \|y\|$.

Uspořádanou dvojici $(X, \|\cdot\|)$ pak nazýváme normovaným lineárním prostorem.

Každá norma na lineárním prostoru X přirozeně indukuje (určuje) na X metriku rovností

$$\rho(x, y) := \|x - y\|.$$

Je zcela snadné dokázat, že ρ má všechny tři vlastnosti metriky. Navíc je zřejmě každá metrika indukovaná normou translačně invariantní a homogenní v tom smyslu, že $\rho(\alpha\, x, \alpha\, y) = |\alpha|\, \rho(x, y)$.

Hovoříme-li o normovaném lineárním prostoru jako o metrickém prostoru, jde o metriku ρ indukovanou normou, není-li řečeno jinak. Je-li metrický prostor (X, ρ) úplný (viz Definice 1.80), říkáme, že $(X, \|\cdot\|)$ je *Banachův prostor.*

1.6 Poznámka. Pokud není obava z nedorozumění, je běžné *označovat normy na různých prostorech stejným symbolem* $\|\cdot\|$. Například hovoříme (formálně nekorektně) o zobrazení $f: (X, \|\cdot\|) \to (Y, \|\cdot\|)$.

V řadě důležitých případů je norma indukovaná skalárním součinem. Připomeňme proto pojem unitárního prostoru.

1.7 Definice. (unitární prostor) *Nechť X je lineární (vektorový) prostor nad tělesem $\mathbb{T}$ reálných nebo komplexních čísel. Řekneme, že funkce $\langle\cdot,\cdot\rangle: X \times X \to \mathbb{T}$, která každé dvojici $(x, y) \in X \times X$ přiřazuje číslo $\langle x, y\rangle \in \mathbb{T}$, je skalární součin na X, jestliže pro všechny vektory x, y, $z \in X$ a každé číslo $\alpha \in \mathbb{T}$ platí*

(1) $\langle x, y\rangle = \overline{\langle y, x\rangle}$;
(2) $\langle \alpha \cdot x, y\rangle = \alpha \cdot \langle x, y\rangle$;
(3) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$;
(4) $x \neq 0 \Longrightarrow \langle x, x\rangle > 0$.

Uspořádanou dvojici $(X, \langle\cdot,\cdot\rangle)$ *pak nazýváme unitárním prostorem (nebo prostorem se skalárním součinem).*

Základy teorie unitárních prostorů se probírají v lineární algebře; připomeňme, že z axiomů (1)–(4) snadno plynou rovnosti $\langle x, y+z\rangle = \langle x,y\rangle + \langle x,z\rangle$, $\langle 0,x\rangle = \langle x,0\rangle = 0$, $\langle x, \alpha\cdot y\rangle = \overline{\alpha}\cdot\langle x,y\rangle$ a v případě reálného lineárního prostoru X také symetrie skalárního součinu ($\langle x,y\rangle = \langle y,x\rangle$).

Každý skalární součin na lineárním prostoru X přirozeně indukuje (určuje) na X normu rovností

$$||x|| := \sqrt{\langle x,x\rangle}.$$

Vlastnost (i) normy okamžitě plyne ze (4) a toho, že $\langle 0,0\rangle = 0$. Dále

$$||\alpha\cdot x|| = \sqrt{\langle \alpha\cdot x, \alpha\cdot x\rangle} = \sqrt{\alpha\overline{\alpha}}\cdot||x|| = |\alpha|\cdot||x||.$$

„Trojúhelníková nerovnost" (iii) je snadným důsledkem (viz [Be] či [Bi]) důležité Cauchy – Schwartzovy nerovnosti

$$|\langle x,y\rangle| \le ||x||\cdot||y||.$$

Na každém unitárním prostoru X je tedy kanonicky zadána norma a tudíž i metrika. Je-li příslušný metrický prostor úplný (tj. je-li příslušný normovaný lineární prostor Banachův), řekneme, že unitární prostor X je Hilbertův.

1.8 Poznámka. Terminologie kolísá. Někdy se při definici Hilbertova prostoru požaduje, aby šlo o komplexní lineární prostor, případně aby byl nekonečně rozměrný.

1.9 Příklad. (prostory $\mathbb{R}^n$ a $\mathbb{C}^n$) Nejjednodušími příklady unitárních prostorů jsou reálný lineární prostor $\mathbb{R}^n$ se skalárním součinem

$$\langle (x_1,\ldots,x_n),(y_1,\ldots,y_n)\rangle := x_1\,y_1 + \cdots + x_n\,y_n$$

a komplexní lineární prostor $\mathbb{C}^n$ se skalárním součinem

$$\langle (x_1,\ldots,x_n),(y_1,\ldots,y_n)\rangle := x_1\,\overline{y_1} + \cdots + x_n\,\overline{y_n}.$$

Skalární součin na $\mathbb{R}^n$ indukuje normu

$$||(x_1,\ldots,x_n)||_2 = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Tato norma $||\cdot||$ se nazývá eukleidovská norma a zřejmě indukuje eukleidovskou metriku. (Z toho vyplývá, že eukleidovská metrika je skutečně metrikou.)

Speciálně $\mathbb{R} = \mathbb{R}^1$ a $\mathbb{C} = \mathbb{C}^1$ chápeme jako unitární (a tedy i normované lineární) prostory. Zřejmě $||x|| = |x|$ v $\mathbb{R}$ i v $\mathbb{C}$.

1.10 Příklad. (normy na $\mathbb{R}^n$) Pro $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$ a $1 \le p < \infty$ klademe

$$||x||_p := \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}} \quad \text{a} \quad ||x||_\infty := \max(|x_1|,\ldots,|x_n|).$$

OBR. 1.1.

Ověření toho, že $||\cdot||_1$ a $||\cdot||_\infty$ jsou normy, je snadné. Pro $||\cdot||_p$, $1 < p < \infty$, jsou vlastnosti normy (i) a (ii) triviálně splněny. Trojúhelníková nerovnost pro tyto normy a body $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ s nezápornými složkami se nazývá Minkowského nerovnost (viz [D II], Věta 105), což je jedna z poměrně hlubokých klasických nerovností. Pro obecné $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$ trojúhelníkovou nerovnost ihned dostaneme, aplikujeme-li Minkowského nerovnost na vektory $x^* = (|x_1|, \ldots, |x_n|)$, $y^* = (|y_1|, \ldots, |y_n|)$. Označení maximové normy symbolem $||\cdot||_\infty$ je motivováno tím, že

$$||x||_\infty = \lim_{p\to\infty} ||x||_p, \qquad x \in \mathbb{R}^n.$$

Nejčastěji se v $\mathbb{R}^n$ užívají normy $||\cdot||_2$, $||\cdot||_\infty$ a $||\cdot||_1$, kterým se většinou říká eukleidovská, maximová a součtová norma, a proto se také označují symboly $||\cdot||_e$, $||\cdot||_m$ a $||\cdot||_s$. Je zřejmé, že pořadě indukují eukleidovskou metriku ρ_2, maximovou metriku ρ_∞ a součtovou metriku ρ_1, o kterých se již hovořilo v předcházejícím oddílu. Pro $x \in \mathbb{R}^n$ zřejmě platí

$$\max(|x_1|, \ldots, |x_n|) \le \sqrt{|x_1|^2 + \cdots + |x_n|^2} \le$$
$$\le |x_1| + \cdots + |x_n| \le n \cdot \max(|x_1|, \ldots, |x_n|),$$

tj.
$$||x||_\infty \le ||x||_2 \le ||x||_1 \le n \cdot ||x||_\infty.$$

Označíme-li pro $p \in \{1, 2, \infty\}$ symbolem $U_\varepsilon^p(a)$ ε-ové okolí bodu a vzhledem k metrice ρ_p, platí $U_\varepsilon^p(a) = \{x \in \mathbb{R}^n \colon ||x - a||_p < \varepsilon\}$, takže z předchozích nerovností dostáváme

$$U_\varepsilon^1(a) \subset U_\varepsilon^2(a) \subset U_\varepsilon^\infty(a) \subset U_{n\varepsilon}^1(a),$$

(viz obr. 1.1, na kterém je vidět tvar těchto okolí pro $n = 2$). Z těchto inkluzí snadno vyplývá, že metriky ρ_2, ρ_1, ρ_∞ určují na $\mathbb{R}^n$ stejný „pojem spojitosti“. Později ukážeme (viz Věta 1.132), že stejný pojem spojitosti určuje každá metrika na $\mathbb{R}^n$, která je indukována normou.

Přirozeným nekonečně dimenzionálním zobecněním unitárních prostorů $\mathbb{R}^n$ (resp. $\mathbb{C}^n$) je reálný (resp. komplexní) prostor ℓ_2.

1.11 Příklad. (prostor ℓ_2) Symbolem ℓ_2 (někdy také symbolem ℓ^2) se označuje množina všech posloupností $(a_n)_1^\infty$ reálných (resp. komplexních) čísel takových, že $\sum_{n=1}^{\infty} |a_n|^2 < \infty$. Protože $|a_n + b_n|^2 \le (|a_n| + |b_n|)^2 \le 2|a_n|^2 + 2|b_n|^2$, je snadno vidět, že ℓ_2 je lineární podprostor prostoru všech reálných (resp. komplexních) posloupností (s přirozenou lineární strukturou). Na ℓ_2 definujeme skalární součin podobně jako v $\mathbb{C}^n$; pro $x = (x_1, x_2, \dots) \in \ell_2$ a $x = (y_1, y_2, \dots) \in \ell_2$ klademe

$$(1.1) \qquad < x, y > := \sum_{i=1}^{\infty} x_i \, \overline{y_i}.$$

Protože $|x_i \, \overline{y_i}| = |x_i| \cdot |y_i| \le |x_i|^2 + |y_i|^2$, vidíme, že řada v (1.1) absolutně konverguje. Je snadné ověřit, že jde skutečně o skalární součin. Není těžké dokázat, že prostor ℓ_2 (reálný nebo komplexní) je Hilbertův. Tento fakt je také speciálním případem úplnosti prostoru $L_2(X, \mu)$, kde (X, μ) je prostor s mírou (viz [LM]). Je totiž $\ell_2 = L_2(\mathbb{N}, \mu)$, kde μ je aritmetická (počítací) míra na $\mathbb{N}$.

1.12 Příklad. (normy a skalární součin na $C[a, b]$) Množinu spojitých reálných funkcí na intervalu $[a, b]$ budeme označovat symbolem $C[a, b]$; tato množina zřejmě tvoří vektorový prostor, definujeme-li součet funkcí a násobení funkce číslem obvyklým způsobem. V předcházejícím oddílu se již hovořilo o metrikách ρ_∞, ρ_2 a ρ_1 na $C[a, b]$. Ukážeme, že tyto metriky jsou indukovány normami.

První z nich je maximová (supremová) norma definovaná předpisem

$$||f||_\infty := \max\{|f(x)|: \ x \in [a, b]\}.$$

Je zcela snadné ověřit, že jde skutečně o normu, která indukuje metriku ρ_∞. *Budeme-li dále hovořit o $C[a, b]$ jako o normovaném lineárním prostoru, uvažujeme tuto normu, není-li řečeno jinak.* Je snadno vidět, že ρ_∞ „popisuje stejnoměrnou konvergenci“: $f_n \to f$ v $C[a, b]$ (viz Definice 1.17), právě když $f_n \rightrightarrows f$. Ze známých vět o stejnoměrné konvergenci snadno plyne, že $C[a, b]$ je Banachův (tj. úplný) prostor.

Pokud pro $f \in C[a, b]$ položíme

$$||f||_2 := \sqrt{\int_a^b |f(x)|^2 \, dx},$$

je ρ_2 zřejmě indukována normou $||\cdot||_2$. To, že $||\cdot||_2$ je skutečně norma, vyplývá ze snadno ověřitelné skutečnosti, že

$$\langle f, g \rangle := \int_a^b f(x) g(x) \, dx$$

je skalární součin na $C[a, b]$, který určuje normu $||\cdot||_2$.

Metrika ρ_1 na $C[a, b]$ je zřejmě indukovaná normou

$$||f||_1 := \int_a^b |f(x)| \, dx.$$

Ověření toho, že jde skutečně o normu, je zcela snadné.

Vzájemný vztah těchto tří metrik (norem) je diskutován v Příkladě 1.57 níže. Poznamenejme, že prostor $C[a, b]$ s metrikou ρ_2 (resp. ρ_1) není úplný prostor.

Nakonec poznamenejme, že i na komplexním prostoru $C[a, b]$ tvořeném komplexními funkcemi lze uvažovat odpovídající tři normy a metriky. Norma $||\cdot||_2$ je pak však zadána skalárním součinem $\langle f, g \rangle := \int_a^b f(x)\overline{g(x)} \, dx$.

1.3 Základní pojmy teorie metrických prostorů

V tomto oddílu zavedeme některé základní pojmy z teorie metrických prostorů a uvedeme řadu jejich důležitých vlastností. Pojmům spojitosti a limity se však budeme věnovat až v dalším oddílu.

Poznamenejme, že řadu pojmů (např. otevřené množiny, uzavřené množiny, uzávěru, hranice) lze definovat různými přirozenými ekvivalentními způsoby; volba definice pak závisí na vkusu autora a je v každé učebnici trochu jiná.

1.13 Definice. *Nechť (X, ρ) je metrický prostor a $M \subset X$. Položíme-li $\rho_M := \rho\restriction_{M\times M}$, tj. klademe-li $\rho_M(x,y) := \rho(x,y)$ pro $x, y \in M$, je zřejmě ρ_M metrika na M. Metrický prostor (M, ρ_M) nazýváme podprostorem prostoru (X, ρ).*

1.14 Poznámka.

(i) Někdy se místo (M, ρ_M) píše (M, ρ). Tato úmluva je sice formálně nekorektní, ale protože zkracuje zápisy a nevede k omylům, často se užívá.
(ii) Máme-li definovanou nějakou vlastnost V metrických prostorů a řekneme, že *množina* $M \subset (X, \rho)$ má tuto vlastnost, myslíme tím, že podprostor (M, ρ_M) má vlastnost V.

Je-li $(X, \|\cdot\|)$ normovaný lineární prostor a Y je lineární podprostor prostoru X, pak $(Y, \|\cdot\|)$ (přesněji $(Y, \|\cdot\|_Y)$, kde $\|\cdot\|_Y$ je restrikce $\|\cdot\|$ na Y) je opět normovaný lineární prostor, který nazýváme podprostorem $(X, \|\cdot\|)$. Je zřejmé, že $(Y, \|\cdot\|)$ určuje metrický prostor, který je podprostorem metrického prostoru určeného $(X, \|\cdot\|)$.

1.15 Definice. *Nechť (X, ρ) je metrický prostor, $a \in X$ a $\varepsilon > 0$. Pak množinu $B(a, \varepsilon) := \{x \in X\colon \rho(x,a) < \varepsilon\}$ nazýváme otevřenou koulí o středu a a poloměru ε. Tuto množinu nazýváme také ε-okolím bodu a; při užití této terminologie ji budeme značit $U_\varepsilon(a)$. Platí tedy*

$$B(a, \varepsilon) = U_\varepsilon(a) = \{x \in X\colon\ \rho(x,a) < \varepsilon\}.$$

Množinu $U \subset X$ nazýváme okolím bodu a, existuje-li $\delta > 0$ takové, že platí $U_\delta(a) \subset U$.

1.16 Poznámka.

(i) Někteří autoři při definici okolí U bodu a požadují, aby U byla otevřená množina (viz Definice 1.19).
(ii) Množině $U_\varepsilon(a) \setminus \{a\}$ se někdy říká redukované (případně prstencové) ε-ové okolí bodu a. Protože tato terminologie není běžná, budeme ji používat jen výjimečně.

1.17 Definice. *Nechť (X, ρ) je metrický prostor, $x \in X$ a $x_n \in X$, $n = 1, 2, \dots$. Říkáme, že x je limitou posloupnosti (x_n) a píšeme*

$$\lim_{n\to\infty} x_n = x \qquad \textit{nebo} \qquad x_n \to x,$$

jestliže

$$\lim_{n\to\infty} \rho(x_n, x) = 0, \quad \textit{tj.}$$

$$\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n > n_0\colon\ \ \rho(x_n, x) < \varepsilon.$$

(Poslední podmínkou definujeme limitu posloupnosti (x_n) i tehdy, jestliže konečný počet jejích členů není definován.) Má-li posloupnost (x_n) limitu, říkáme, že je konvergentní (a konverguje ke své limitě), v opačném případě říkáme, že je divergentní.

Snadno je vidět, že posloupnost v metrickém prostoru může mít *nejvýše jednu limitu.*

1.18 Poznámka. Pojem limity tedy definujeme i pro posloupnosti, které jsou zobrazeními $x\colon (\mathbb{N} \setminus K) \to X$, kde $K \subset \mathbb{N}$ je konečná. To umožňuje jednodušší formulaci řady tvrzení (např. Věty 1.35 níže). Kdykoliv dále napíšeme symbol (x_n), myslíme tím takovou „posloupnost v širším smyslu“.

1.19 Definice. *Nechť (X, ρ) je metrický prostor, $a \in X$ a $M \subset X$.*

(a) *Řekneme, že a je vnitřním bodem množiny M, existuje-li $\varepsilon > 0$ takové, že $U_\varepsilon(a) \subset M$; tj. když M je okolím bodu a. Množinu všech vnitřních bodů množiny M nazýváme vnitřkem množiny M a značíme ji M^o nebo $\operatorname{int} M$.*
(b) *Uzávěrem množiny M rozumíme množinu $\overline{M}$ bodů $a \in X$ takových, že pro každé $\varepsilon > 0$ platí $U_\varepsilon(a) \cap M \neq \emptyset$.*
(c) *Řekneme, že a je hraničním bodem množiny M, jestliže pro každé $\varepsilon > 0$ platí $U_\varepsilon(a) \cap M \neq \emptyset$ a také $U_\varepsilon(a) \setminus M \neq \emptyset$. Množinu všech hraničních bodů množiny M nazýváme hranicí množiny M a značíme ji ∂M.*
(d) *Řekneme, že a je hromadným bodem množiny M, jestliže pro každé $\varepsilon > 0$ platí $(U_\varepsilon(a) \setminus \{a\}) \cap M \neq \emptyset$. Množinu všech hromadných bodů množiny M nazýváme derivací množiny M a značíme ji symbolem M'.*
(e) *Je-li $a \in M \setminus M'$, nazýváme a izolovaným bodem množiny M.*
(f) *Říkáme, že M je otevřená množina, jestliže $M = M^o$.*
(g) *Říkáme, že M je uzavřená množina, jestliže $M = \overline{M}$.*
(h) *Říkáme, že M je hustá množina, jestliže $\overline{M} = X$.*

Následující tvrzení snadno plynou z definic.

1.20 Tvrzení. *Nechť (X, ρ) je metrický prostor, $a \in X$ a $M \subset X$. Pak platí následující tvrzení.*

(a) $$M^o \subset M \subset \overline{M}, \quad M^o = X \setminus \overline{X \setminus M}, \quad \overline{M} = X \setminus (X \setminus M)^o,$$

$$\partial M = \overline{M} \setminus M^o, \quad \overline{M} = M \cup M'.$$

(b) *Bod a leží v uzávěru $\overline{M}$ množiny M, právě když existuje posloupnost $x_1, x_2, \dots$ bodů z M, která konverguje k bodu a.*

(c) *Bod a leží v derivaci M' množiny M, právě když existuje posloupnost $x_1, x_2, \dots$ bodů z $M \setminus \{a\}$, která konverguje k bodu a.*

(d) *Bod a leží v derivaci M' množiny M, právě když každá koule $B(a, \varepsilon)$, $\varepsilon > 0$, obsahuje nekonečně mnoho bodů množiny M.*

(e) *Vnitřek M^o množiny M je největší otevřená množina obsažená v množině M ($G \subset M^o$ pro každou otevřenou $G \subset M$).*

(f) *Uzávěr $\overline{M}$ množiny M je nejmenší uzavřená množina obsahující množinu M ($\overline{M} \subset F$ pro každou uzavřenou $F \supset M$).*

(g) *Množina M je uzavřená, právě když její doplněk $X \setminus M$ je otevřená množina. Množina M je otevřená, právě když její doplněk $X \setminus M$ je uzavřená množina.*

(h) *Množina M je uzavřená, právě když platí implikace*

$$(x_n \in M,\ n = 1, 2, \dots;\ x_n \to a) \Longrightarrow a \in M.$$

(i) *Množiny M' a ∂M jsou uzavřené.*

Také důkaz následující věty je snadný.

1.21 Věta. (vlastnosti systémů otevřených a uzavřených množin) *V libovolném metrickém prostoru X platí tato tvrzení.*

(i) *Množiny $\emptyset$ a X jsou zároveň otevřené i uzavřené.*

(ii) *Sjednocení libovolného systému otevřených množin je otevřená množina a průnik konečného systému otevřených množin je otevřená množina.*

(iii) *Průnik libovolného systému uzavřených množin je uzavřená množina a sjednocení konečného systému uzavřených množin je uzavřená množina.*

Není-li z kontextu jasné, v jakém metrickém prostoru pracujeme, používáme symboly $\overline{A}^{(X,\rho)}$, $\overline{A}^{\rho}$, $\partial_\rho(A)$ apod. Speciálně, uzávěr množiny $A \subset Y$ v podprostoru (Y, ρ_Y) prostoru (X, ρ) označujeme symbolem $\overline{A}^{Y}$. Zřejmě platí důležitá rovnost $\overline{A}^{Y} = \overline{A} \cap Y$, ze které snadno vyplývá následující užitečné tvrzení.

1.22 Tvrzení. (otevřenost a uzavřenost v podprostoru) *Buď (X, ρ) metrický prostor a nechť $A \subset Y \subset X$. Pak A je uzavřená (resp. otevřená) podmnožina podprostoru (Y, ρ), právě když existuje uzavřená (resp. otevřená) podmnožina B prostoru (X, ρ), pro kterou $A = B \cap Y$.*

Speciálně, je-li Y uzavřená (resp. otevřená) v X, pak množina A je uzavřená (resp. otevřená) v (Y, ρ), právě když je uzavřená (resp. otevřená) v (X, ρ).

1.23 Poznámka. (o topologických prostorech) Pojem topologického prostoru je důležitým zobecněním pojmu metrického prostoru. V obecném topologickém prostoru se zavádějí pojmy otevřené množiny, okolí bodu, konvergence posloupnosti, spojité funkce a mnoha dalších topologických pojmů (srov. str. 23), nezavádějí se však pojmy netopologické (např. omezenosti množiny, cauchyovské posloupnosti, lipschitzovské funkce). Nejběžnější způsob zavedení topologického prostoru je pomocí zadání systému otevřených množin, po kterém požadujeme pouze to, aby měl základní vlastnosti (z Věty 1.21) systému otevřených množin v metrickém prostoru.

Nechť X je množina a $\mathcal{G}$ je systém podmnožin X s těmito vlastnostmi:

(i) $\emptyset \in \mathcal{G}$, $X \in \mathcal{G}$.
(ii) *Průnik konečně mnoha prvků z $\mathcal{G}$ leží opět v $\mathcal{G}$.*
(iii) *Sjednocení libovolného systému množin z $\mathcal{G}$ leží opět v $\mathcal{G}$.*

Pak říkáme, že $(X, \mathcal{G})$ je topologický prostor a $\mathcal{G}$ je systém všech otevřených množin v tomto topologickém prostoru (sytému $\mathcal{G}$ se někdy říká „topologie“).

Za okolí bodu x pak prohlásíme takové množiny U, pro které existuje $G \in \mathcal{G}$ splňující $x \in G \subset U$. A pomocí pojmu okolí můžeme v topologickém prostoru přirozeně definovat všechny další důležité topologické pojmy z teorie metrických prostorů.

Největší význam mají (Hausdorffovy) topologické prostory, které navíc splňují následující přirozený „oddělovací axiom“:

(iv) *Jsou-li $x \neq y$ dva body z X, pak existují $G_x \in \mathcal{G}$, $G_y \in \mathcal{G}$ takové, že platí $x \in G_x$, $y \in G_y$ a $G_x \cap G_y = \emptyset$.*

Z Věty 1.21 vyplývá, že každá metrika na X indukuje na X topologii. Je snadno vidět, že tato topologie je Hausdorffova, tj. splňuje (iv).

Pro aplikace je důležité, že pomocí pojmu topologického prostoru „lze zachytit“ pojem bodové konvergence posloupnosti funkcí, což pomocí pojmu metrického prostoru možné není. Například na $C[0,1]$ existuje (Hausdorffova) topologie $\mathcal{G}$ taková, že $f_n \to f$ v $(C[0,1], \mathcal{G})$, právě když $f_n(x) \to f(x)$ pro každé $x \in [0,1]$; žádná topologie indukovaná metrikou však tuto vlastnost nemá.

1.24 Definice. (vzdálenost bodu od množiny a vzdálenost dvou množin) *Nechť (X, ρ) je metrický prostor, $a \in X$, $A \subset X$, $B \subset X$. Pak definujeme vzdálenost bodu a od množiny A vzorcem*

$$\rho(a, A) = \operatorname{dist}(a, A) := \inf_{x \in A} \rho(a, x)$$

a vzdálenost množin A a B předpisem

$$\rho(A, B) = \operatorname{dist}(A, B) := \inf_{x \in A, y \in B} \rho(x, y).$$

Zřejmě platí $\rho(a, A) = \rho(\{a\}, A)$. Dále $\rho(a, \emptyset) = \rho(A, \emptyset) = \rho(\emptyset, B) = \infty$; jinak jsou vždy čísla $\rho(a, A)$, $\rho(A, B)$ reálná a nezáporná. Snadno je vidět, že $\rho(a, A) = 0$, právě když $a \in \overline{A}$.

1.25 Definice. (průměr množiny) *Nechť (X, ρ) je metrický prostor. Průměr (diametr) množiny $\emptyset \neq A \subset X$ definujeme vzorcem*

$$\operatorname{diam} A := \sup_{x, y \in A} \rho(x, y)$$

a klademe $\operatorname{diam} \emptyset := 0$. *Jestliže* $\operatorname{diam} A < \infty$, *řekneme, že A je omezená množina. Řekneme, že X je omezený prostor, jestliže X je omezená množina.*

1.4 Funkce a zobrazení na metrických prostorech

Nyní budeme definovat pojem spojitosti a limity pro zobrazení mezi metrickými prostory. Jde o přímočaré zobecnění těchto pojmů z teorie reálných funkcí reálné proměnné; terminologie však trochu kolísá, pokud nejde o zobrazení definované na celém metrickém prostoru.

Již pro reálné funkce není terminologie zcela jednotná. Například podle nejběžnější terminologie reálná funkce reálné proměnné $\sqrt{x}$ není spojitá v bodě 0, protože spojitost se chápe vzhledem k celé reálné přímce a $\sqrt{x}$ není definována na žádném levém okolí 0. Někteří autoři však chápou spojitost vzhledem k definičnímu oboru; pak ovšem funkce $\sqrt{x}$ v bodě 0 spojitá je.

Začněme proto s nejjednodušším případem spojitosti zobrazení definovaného na celém metrickém prostoru, kde k žádnému kolísání terminologie nedochází.

1.26 Definice. *Nechť (X,ρ), (Y,σ) jsou metrické prostory, $f\colon X \to Y$ je zobrazení a $a \in X$. Potom:*

(i) *Řekneme, že zobrazení f je spojité v bodě a, jestliže pro každé $\varepsilon > 0$ existuje $\delta > 0$ takové, že (pro každý bod $x \in X$) platí implikace*

$$\rho(x,a) < \delta \quad \Longrightarrow \quad \sigma(f(x), f(a)) < \varepsilon.$$

(ii) *Řekneme, že zobrazení f je spojité, je-li spojité v každém bodě prostoru X.*

(iii) *Řekneme, že zobrazení f je homeomorfismus, je-li bijektivní a obě zobrazení f, f^{-1} jsou spojitá.*

1.27 Poznámka.

(i) Homeomorfismu se také říká homeomorfní (nebo topologické) zobrazení.

(ii) Řekneme-li (výjimečně), že $f\colon (X,\rho) \to (Y,\sigma)$ je homeomorfní, i když f není bijekce, myslíme tím, že f je prosté a bijekce $f\colon (X,\rho) \to (f(X),\sigma)$ je homeomorfismus.

V praxi (zvláště při vyšetřování funkcí více proměnných) potřebujeme následující obecnější pojem *spojitosti vzhledem k množině.*

1.28 Definice. *Nechť (X,ρ), (Y,σ) jsou metrické prostory, f je zobrazení z X do Y, nechť $A \subset X$ a $a \in X$. Potom:*

(i) *Řekneme, že zobrazení f je spojité v bodě a vzhledem k množině A, jestliže $a \in A$ a pro každé $\varepsilon > 0$ existuje $\delta > 0$ takové, že platí implikace*

$$(x \in A,\ \rho(x,a) < \delta) \quad \Longrightarrow \quad \sigma(f(x), f(a)) < \varepsilon.$$

Řekneme, že f je spojité v bodě a, je-li spojité v bodě a vzhledem k $A = X$.

(ii) *Řekneme, že f je spojité na (v) množině A, jestliže f je spojité vzhledem k množině A v každém bodě množiny A.*

1.29 Poznámka.

(i) Je-li f spojité v bodě a vzhledem k množině A, pak zřejmě existuje $\delta > 0$, pro které $U_\delta(a) \cap A \subset D_f$.

(ii) Zobrazení f je zřejmě spojité na množině A právě tehdy, když jeho restrikce $f\restriction_A: (A, \rho) \to (Y, \sigma)$ je spojité zobrazení.

(iii) Je-li f zobrazení z X do Y a $a \in D_f$, pak pro pojem spojitosti f v bodě a se nabízejí dvě přirozené možnosti; buď se uvažuje spojitost vzhledem k $A := D_f$ nebo spojitost vzhledem k $A := X$. My jsme zde zvolili druhou možnost, která je v teorii funkcí více proměnných častější (a přirozenější). Řada autorů však volí druhou možnost (která je přirozenější v jiných teoriích).

(iv) Výše definované pojmy týkající se spojitosti se ovšem netýkají pouze zobrazení f; abychom je mohli definovat, musíme mít zadány také prostory (X, ρ), (Y, σ) (a množinu A). Změníme-li například některou z metrik, může se změnit pojem spojitosti. Na $\mathbb{R}^n$ bereme vždy eukleidovskou metriku, není-li řečeno jinak.

Nyní vyslovíme některá základní tvrzení o spojitosti, která lze snadno dokázat z definice.

1.30 Tvrzení. *Nechť $f: (X, \rho) \to (Y, \sigma)$, $a \in X$. Pak následující tvrzení jsou po dvou ekvivalentní.*

(i) *f je spojité v bodě a.*

(ii) $\forall \varepsilon > 0 \, \exists \delta > 0$: $f(U_\delta(a)) \subset U_\varepsilon(f(a))$.

(iii) $\forall \varepsilon > 0 \, \exists \delta > 0$: $U_\delta(a) \subset f^{-1}\left(U_\varepsilon(f(a))\right)$.

(iv) *Je-li V okolí bodu $f(a)$, pak $f^{-1}(V)$ je okolí bodu a.*

(v) *Pro každou posloupnost (x_n) v X platí implikace*
$$x_n \to a \quad \Longrightarrow \quad f(x_n) \to f(a).$$

1.31 Věta. *Nechť je dáno zobrazení $f: (X, \rho) \to (Y, \sigma)$. Pak následující tvrzení jsou ekvivalentní.*

(i) *Zobrazení f je spojité.*

(ii) *Pro každou otevřenou podmnožinu G prostoru (Y, σ) je množina $f^{-1}(G)$ otevřená podmnožina prostoru (X, ρ).*

(iii) *Pro každou uzavřenou podmnožinu F prostoru (Y, σ) je množina $f^{-1}(F)$ uzavřená podmnožina prostoru (X, ρ).*

1.32 Tvrzení. *Nechť $f: (X, \rho) \to (Y, \sigma)$ je bijekce. Pak následující tvrzení jsou ekvivalentní.*

(i) *Zobrazení f je homeomorfismus.*

(ii) *Množina $A \subset X$ je otevřená, právě když $f(A)$ je otevřená v Y.*

(iii) *Množina $A \subset X$ je uzavřená, právě když $f(A)$ je uzavřená v Y.*

1.33 Definice. *Nechť (X, ρ), (Y, σ) jsou metrické prostory, f je zobrazení z X do Y, $A \subset X$ a nechť $a \in X$ je hromadným bodem množiny A . Řekneme, že $b \in Y$ je limita zobrazení f v bodě a vzhledem k množině A a píšeme*

$$\lim_{x \to a,\, x \in A} f(x) = b,$$

jestliže pro každé $\varepsilon > 0$ existuje $\delta > 0$ takové, že platí implikace

$$(x \in A,\ 0 < \rho(x, a) < \delta) \implies \sigma(f(x), b) < \varepsilon. \tag{1.2}$$

Je-li $A = X$, říkáme, že f má v bodě a limitu b a píšeme $\lim\limits_{x \to a} f(x) = b$.

1.34 Poznámka. Nechť X, Y, A, a jsou jako v předchozí definici a je dán systém f_ω $(\omega \in \Omega \neq \emptyset)$ zobrazení z X do Y. Jestliže pro každé $\omega \in \Omega$ existuje

$$\lim_{x \to a,\, x \in A} f_\omega(x) = b_\omega \tag{1.3}$$

a ke každému $\varepsilon > 0$ existuje $\delta > 0$ takové, že pro každé $\omega \in \Omega$ platí implikace (1.2), ve které místo f, b píšeme f_ω, b_ω, pak říkáme, že limita (1.3) je stejnoměrná vzhledem k $\omega \in \Omega$.

Důkazy následujících vět o limitách a o spojitosti jsou zcela analogické odpovídajícím důkazům pro reálné funkce reálné proměnné.

1.35 Věta. (Heineho věta) *Nechť (X, ρ), (Y, σ) jsou metrické prostory, f je zobrazení z X do Y, $A \subset X$, $b \in Y$ a nechť $a \in X$ je hromadným bodem množiny A. Pak $\lim\limits_{x \to a,\, x \in A} f(x) = b$, právě když platí následující (Heineho) podmínka:*

Pro každou posloupnost (x_n) bodů z množiny $A \setminus \{a\}$ platí implikace
$$x_n \to a \implies f(x_n) \to b.$$

Poznamenejme, že při této formulaci Heineho věty je nutno posloupnost $(f(x_n))$ v předchozí větě chápat jako posloupnost „v širším smyslu“ (viz Poznámka 1.18), jinak bychom museli předpokládat, že $x_n \in D_f \cap (A \setminus \{a\})$.

1.36 Tvrzení. *Nechť (X, ρ), (Y, σ) jsou metrické prostory, f je zobrazení z X do Y, $A \subset X$ a nechť $a \in A \cap A' \cap D_f$. Pak zobrazení f je spojité v bodě a vzhledem k množině A, právě když $\lim\limits_{x \to a,\, x \in A} f(x) = f(a)$.*

1.37 Poznámka. Předpoklad $a \in A'$ v předchozím tvrzení je nutný. Pokud $a \in A \cap D_f$, ale $a \notin A'$, pak f je zřejmě spojité v bodě a vzhledem k množině A, ale limita zobrazení f v bodě a vzhledem k A není definovaná.

1.38 Věta. (věta o spojitosti složeného zobrazení v bodě) *Nechť (X, ρ), (Y, σ), (Z, ω) jsou metrické prostory, f je zobrazení z X do Y a g je zobrazení z Y do Z. Nechť $A \subset X$, $a \in A$, $B \subset Y$, $f(a) \in B$ a existuje $\delta > 0$ takové, že $f(U_\delta(a) \cap A) \subset B$.*

Nechť dále zobrazení f je spojité v bodě a vzhledem k A a g je spojité v bodě $f(a)$ vzhledem k B. Pak zobrazení $g \circ f$ je spojité v bodě a vzhledem k A.

1.39 Věta. (věta o spojitosti složeného zobrazení) *Nechť (X,ρ), (Y,σ), (Z,ω) jsou metrické prostory a $f: X \to Y$, $g: Y \to Z$ jsou spojitá zobrazení. Pak zobrazení $g \circ f: X \to Z$ je spojité.*

1.40 Věta. (věta o limitě složeného zobrazení) *Nechť (X,ρ), (Y,σ), (Z,ω) jsou metrické prostory, f je zobrazení z X do Y a g je zobrazení z Y do Z. Nechť $A \subset X$, $a \in A'$, $B \subset Y$, $b \in B'$, $c \in Z$ a existuje $\delta > 0$ takové, že $f\left((U_\delta(a) \setminus \{a\}) \cap A\right) \subset B$. Nechť dále*

$$\lim_{x \to a,\, x \in A} f(x) = b, \qquad \lim_{y \to b,\, y \in B} g(y) = c$$

a platí jedna z následujících podmínek:

(i) *Existuje $\eta > 0$ takové, že $b \notin f\left((U_\eta(a) \setminus \{a\}) \cap A\right)$.*
(ii) *Zobrazení g je spojité v bodě b vzhledem k B.*

Pak
$$\lim_{x \to a,\, x \in A} (g \circ f)(x) = c.$$

Je-li f reálná funkce na metrickém prostoru (X,ρ), je přirozené chápat f jako zobrazení $f: (X,\rho) \to (\mathbb{R},\rho_2)$, takže předchozí definice dávají pojem spojitosti f (v bodě a na množině) a také pojem *vlastní* limity. Takto však nedostaneme pojem nevlastní limity (jehož „správná“ definice je ovšem zřejmou analogií definice v případě funkce jedné reálné proměnné). Ale i případ nevlastních limit můžeme chápat jako speciální případ Definice 1.33, chápeme-li f jako zobrazení $f: (X,\rho) \to (\mathbb{R}^*,\sigma)$, kde σ je vhodná (tzv. redukovaná) metrika na $\mathbb{R}^*$, srov. Tvrzení 1.59.

Je zřejmé, že na případ limity reálné funkce na metrickém prostoru lze zobecnit (se „stejným“ důkazem) všechna základní pravidla o počítání s limitami (vlastními i nevlastními). Například zůstávají v platnosti všechny věty o limitě součtu, součinu a podílu funkcí, stejně jako věty o vztahu limity a nerovností.

Pro reálné funkce na metrickém prostoru můžeme také zřejmým způsobem definovat symboly o, O, $\sim$, $\asymp$, týkající se „asymptotického chování“ funkcí.

1.41 Definice. *Nechť (X,ρ) je metrický prostor, $a \in X'$ a f, g jsou reálné funkce definované na podmnožinách X. Pak:*

(i) *Píšeme $f(x) = o(g(x))$, $x \to a$, jestliže $\lim\limits_{x \to a} \dfrac{f(x)}{g(x)} = 0$. (Symbol „o“ zde čteme „malé o“.)*
(ii) *Píšeme $f(x) = O(g(x))$, $x \to a$, jestliže existuje $\delta > 0$ takové, že funkce $\frac{f(x)}{g(x)}$ je definovaná a omezená na „redukovaném okolí“ $U_\delta(a) \setminus \{a\}$. (Symbol „O“ zde čteme „velké O“.)*
(iii) *Píšeme $f(x) \sim g(x)$, $x \to a$, jestliže $\lim\limits_{x \to a} \dfrac{f(x)}{g(x)} = 1$. Říkáme, že f a g jsou silně ekvivalentní v bodě a.*

(iv) *Píšeme* $f(x) \asymp g(x),\ x \to a,$ *jestliže* $f(x) = O(g(x)),\ x \to a$ *a také* $g(x) = O(f(x)),\ x \to a$. *Říkáme, že* f *a* g *jsou slabě ekvivalentní v bodě* a.

1.42 Poznámka.

(i) Ve všech čtyřech případech musí existovat $\delta > 0$ takové, že f i g jsou definované na množině $U_\delta(a) \setminus \{a\}$ a funkce g je na ní *nenulová*. Někdy se v definici předpokládá (viz [D II]), že g je kladná.

(ii) Je zřejmé, jak by se definoval např. symbol $f(x) = o(g(x)),\ x \to a,\ x \in A$, v případě, že $a \in A'$.

(iii) Uvedli jsme klasickou definici, která stačí pro většinu aplikací. Někteří autoři používají obecnější definici, která nevyžaduje nenulovost g ve smyslu (i). Při této definici platí, že $f(x) = o(g(x)),\ x \to a\ \ (f(x) = O(g(x)),\ x \to a;\ f(x) \sim g(x),\ x \to a;\ f(x) \asymp g(x);\ x \to a)$, právě když existuje funkce h taková, že platí $f = h \cdot g$ a $h(x) = o(1),\ x \to a$ podle klasické definice $(h(x) = O(1),\ x \to a;\ h(x) \sim 1,\ x \to a;\ h(x) \asymp 1,\ x \to a)$.

Tyto obecnější definice (které lze přirozeným způsobem přeformulovat na definice neužívající pojem limity), jsou přirozenější např. proto, že relace $\sim$ je v tomto obecném pojetí relací ekvivalence (na množině funkcí definovaných na nějakém redukovaném okolí bodu a), zatímco v klasickém pojetí relace $\sim$ není reflexivní.

Analogicky jako pro funkce jedné reálné proměnné se definují také důležité pojmy limes superior a limes inferior funkce v bodě. Nejrychlejší definice je tato:

1.43 Definice. *Nechť* (X, ρ) *je metrický prostor,* $a \in X'$ *a* f *je reálná funkce definovaná na podmnožině* $D_f \subset X$ *takové, že* $D_f \cup \{a\}$ *je okolím bodu* a. *Pak klademe*

$$\begin{aligned}\limsup_{x\to a} f(x) &:= \lim_{\delta\to 0+} \sup\{f(x)\colon\ x \in U_\delta(a) \setminus \{a\}\},\\ \liminf_{x\to a} f(x) &:= \lim_{\delta\to 0+} \inf\{f(x)\colon\ x \in U_\delta(a) \setminus \{a\}\}.\end{aligned}$$

Pokud je hodnota $s(\delta) := \sup\{f(x)\colon\ x \in U_\delta(a) \setminus \{a\}\}$ konečná pro některé $\delta = \delta_0 > 0$, pak funkce s je konečná a neklesající na $(0, \delta_0)$, takže existuje limita $\lim_{\delta\to 0+} s(\delta) \in \mathbb{R} \cup \{-\infty\}$. V opačném případě $s(\delta) = \infty$ pro každé $\delta > 0$. V tom případě ovšem klademe $\lim_{\delta\to 0+} s(\delta) := \infty$.

1.44 Poznámka. Analogicky se definuje limes superior a limes inferior funkce f v bodě $a \in (X, \rho)$ vzhledem k množině $A \subset X$ za předpokladu, že $a \in A'$ a D_f obsahuje množinu $(U_\delta(a) \setminus \{a\}) \cap A$ pro některé $\delta > 0$. Tento pojem lze ovšem také definovat jako limes superior v podprostoru (Y, ρ), kde $Y := A \cup \{a\}$, funkce $f^* := f\restriction_{Y \cap D_f}$ v bodě a.

Následující tvrzení, které ukazuje dvě charakterizace pojmu limes superior funkce, je analogické známému tvrzení o pojmu limes superior posloupnosti reálných čísel.

1.45 Tvrzení. *Nechť (X, ρ) je metrický prostor, $a \in X'$, f je reálná funkce definovaná na podmnožině $D_f \subset X$ takové, že $D_f \cup \{a\}$ je okolím bodu a a $H \in \mathbb{R}^*$. Pak jsou následující tvrzení ekvivalentní.*

(i) $\limsup\limits_{x \to a} f(x) = H$.

(ii) *Existuje posloupnost (x_n) bodů z $X \setminus \{a\}$ taková, že $x_n \to a$, $f(x_n) \to H$ a H je největší prvek $\mathbb{R}^*$ s touto vlastností.*

(iii) *H je jediný prvek $\mathbb{R}^*$ s těmito dvěma vlatnostmi: Jestliže $-\infty < c < H$, pak $a \in \big(f^{-1}((c,\infty))\big)'$. Jestliže $H < c < \infty$, pak $a \notin \big(f^{-1}((c,\infty))\big)'$.*

Analogické (symetrické, duální) tvrzení platí ovšem pro limes inferior. Poznamenejme, že prvkům s vlastností z (ii) se říká hromadná (resp. limitní) hodnota (nebo také hromadný bod) funkce f v bodě a. Podmínkou (ii) je vysvětlen latinský název limes superior (největší limita). V řadě jazyků se limes superior nazývá „horní limita“.

1.46 Poznámka. Pro názorné pochopení podmínky (iii) je vhodné si uvědomit toto:

(i) $a \in \big(f^{-1}((c,\infty))\big)'$ platí právě tehdy, když v každém redukovaném okolí $U_\delta(a) \setminus \{a\}$, $\delta > 0$, existuje bod x, pro který $f(x) > c$.

(ii) $a \notin \big(f^{-1}((c,\infty))\big)'$ platí právě tehdy, když existuje $\delta > 0$ takové, že $f(x) \le c$ pro každý bod $x \in U_\delta(a) \setminus \{a\}$.

Limes superior a inferior funkce mají zcela analogické vlastnosti jako limes superior a inferior posloupnosti. Například platí tato důležitá věta.

1.47 Věta. *Nechť (X, ρ) je metrický prostor, $a \in X'$ a f je reálná funkce definovaná na podmnožině $D_f \subset X$ takové, že $D_f \cup \{a\}$ je okolím bodu a. Pak $\lim_{x \to a} f(x)$ existuje, právě když*

$$\limsup_{x \to a} f(x) = \liminf_{x \to a} f(x) =: S \in \mathbb{R}^*;$$

v tom případě $\lim_{x \to a} f(x) = S$.

V Dodatku 6.1. je ukázáno obecné pojetí, které zahrnuje případ funkcí i posloupností; tamtéž lze nalézt více informací o hromadných hodnotách.

Také zesílení pojmu spojitosti zobrazení — stejnoměrná spojitost a lipschitzovskost — jsou přímočarými zobecněními těchto pojmů z případu reálných funkcí reálné proměnné.

1.48 Definice. (stejnoměrně spojité a lipschitzovské zobrazení) *Nechť jsou dány metrické prostory $(X, \rho), (Y, \sigma)$ a zobrazení $f \colon X \to Y$.*

(i) *Řekneme, že f je stejnoměrně spojité, jestliže platí podmínka*

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall u, v \in X\colon \quad \rho(u,v) < \delta \Longrightarrow \sigma(f(u), f(v)) < \varepsilon.$$

(ii) *Řekneme, že f je lipschitzovské, existuje-li $K \ge 0$ takové, že*

$$\forall u, v \in X\colon \quad \sigma(f(u), f(v)) \le K \cdot \rho(u,v).$$

Platí-li předchozí výrok, říkáme, že f je lipschitzovské s konstantou K.

Říkáme, že f je stejnoměrně spojité (resp. lipschitzovské) na $A \subset X$, jestliže restrikce $f\restriction_A\colon (A,\rho) \to (Y,\sigma)$ je stejnoměrně spojitá (resp. lipschitzovská).

Je zřejmé, že každé lipschitzovské zobrazení je stejnoměrně spojité; obrácená implikace však obecně neplatí.

1.49 Definice. *Nechť (X,ρ) a (Y,σ) jsou metrické prostory a nechť $f\colon X \to Y$ je bijekce. Řekneme, že f je izometrie (izometrické zobrazení), jestliže*

$$\forall u,v \in X\colon \quad \sigma(f(u),f(v)) \;=\; \rho(u,v).$$

Zobrazení f je tedy izometrie, právě když obě zobrazení f, f^{-1} jsou lipschitzovská s konstantou 1. Speciálně každá izometrie je homeomorfismus.

1.50 Poznámka. Někdy se izometrickými nazývají i zobrazení $f\colon (X,\rho) \to (Y,\sigma)$, která nejsou bijektivní, ale zachovávají vzdálenost (tj. $f\colon (X,\rho) \to (f(X),\sigma)$, je izometrie ve smyslu Definice 1.49).

1.51 Definice. *Řekneme, že metrické prostory (X,ρ) a (Y,σ) jsou homeomorfní (resp. izometrické), existuje-li bijekce $f\colon X \to Y$, která je homeomorfismus (resp. izometrie).*

Z trojúhelníkové nerovnosti ihned vyplývá, že v každém metrickém prostoru (X,ρ) je pro libovolný bod $a \in X$ funkce $\rho(\cdot,a)$ lipschitzovská s konstantou 1 (a je tedy i spojitá). Obecněji:

1.52 Tvrzení. *Nechť (X,ρ) je metrický prostor a A je jeho neprázdná podmnožina. Pak funkce $\rho(\cdot,A)$ je lipschitzovská s konstantou 1.*

Nakonec uvedeme větu, která ukazuje, že spojitost a limity zobrazení z metrického prostoru do eukleidovského prostoru $\mathbb{R}^n$ lze vyšetřovat „po složkách".

Nechť (X,ρ) je metrický prostor a f je zobrazení z X do $\mathbb{R}^n$. Je tedy $f\colon D \to \mathbb{R}^n$, kde $D \subset X$. Je zřejmé, že rovností $f(x) = (f_1(x),\dots,f_n(x))$ jsou definovány funkce $f_1\colon D \to \mathbb{R},\dots,f_n\colon D \to \mathbb{R}$. Říkáme, že $f_1,\dots,f_n$ jsou *složky* zobrazení f a píšeme $f = (f_1,\dots,f_n)$. (Zřejmě $f_i(x) = \langle f(x), e_i\rangle$, kde e_i je i-tý prvek kanonické báze $\mathbb{R}^n$.)

1.53 Věta. *Nechť (X,ρ) je metrický prostor, $A \subset X$ a $f = (f_1,\dots,f_n)$ je zobrazení z X do $\mathbb{R}^n$. Potom platí:*

(i) *Nechť $a \in A$. Pak zobrazení f je spojité v bodě a vzhledem k množině A, právě když každá složka f_i, $i = 1,\dots,n$, je spojitá v bodě a vzhledem k množině A.*

(ii) *Zobrazení f je spojité na A, právě když každá složka f_i, $i = 1,\dots,n$, je spojitá na A.*

(iii) *Nechť $a \in A'$ a $b = (b_1,\dots,b_n) \in \mathbb{R}^n$. Pak*

$$\lim_{x\to a,\, x\in A} f(x) = b \iff \lim_{x\to a,\, x\in A} f_i(x) = b_i,\ i = 1, \dots, n.$$

1.5 Vztahy mezi metrikami

Často se setkáváme se situací, kdy na množině jsou zadány dvě různé metriky. Pak je užitečná následující terminologie.

1.54 Definice. *Nechť ρ a σ jsou metriky na množině X a $I: X \to X$ je identické zobrazení. Potom řekneme, že:*

(i) *Metrika ρ je silnější než metrika σ, jestliže $I: (X, \rho) \to (X, \sigma)$ je spojité zobrazení.*
(ii) *Metrika ρ je slabší než metrika σ, jestliže σ je silnější než ρ.*
(iii) *Metriky ρ a σ jsou ekvivalentní, jestliže metrika ρ je zároveň silnější i slabší než metrika σ.*
(iv) *Metriky ρ a σ jsou lipschitzovsky ekvivalentní, jestliže $I: (X, \rho) \to (X, \sigma)$ je lipschitzovské zobrazení a také $I: (X, \sigma) \to (X, \rho)$ je lipschitzovské.*

Přímo z definic a z Tvrzení 1.30 snadno dostáváme následující tvrzení.

1.55 Tvrzení. *Nechť ρ a σ jsou metriky na množině X. Potom platí:*

(i) *Metrika ρ je silnější než metrika σ, právě když platí implikace*
$$x_n \to x \quad \text{v } (X, \rho) \implies x_n \to x \quad \text{v } (X, \sigma).$$
(ii) *Metriky ρ a σ jsou ekvivalentní, právě když platí ekvivalence*
$$x_n \to x \quad \text{v } (X, \rho) \iff x_n \to x \quad \text{v } (X, \sigma).$$
(iii) *Metriky ρ a σ jsou ekvivalentní právě tehdy, když identické zobrazení $I: (X, \rho) \to (X, \sigma)$ je homeomorfismus.*
(iv) *Metriky ρ a σ jsou lipschitzovsky ekvivalentní, právě když existují reálná čísla $c, d > 0$ taková, že pro každé dva body $x, y \in X$ platí*
$$c \cdot \sigma(x, y) \le \rho(x, y) \le d \cdot \sigma(x, y).$$

Pojmy definované v metrických prostorech, které nemění svůj význam při záměně metriky za metriku s ní ekvivalentní, se nazývají *topologické pojmy.* Z Tvrzení 1.55 (ii) je vidět, že kterýkoliv pojem, který je možno definovat jen s užitím pojmu konvergence posloupnosti (aniž by se hovořilo o metrice), je topologický pojem. Speciálně otevřenost a uzavřenost množiny, uzávěr a vnitřek množiny a také pojem okolí a hromadného bodu jsou topologické pojmy.

Naproti tomu například pojem omezenosti množiny není topologický pojem. Tento pojem a řada jiných pojmů (např. pojem úplnosti metrického prostoru) má však následující slabší vlastnost: nemění svůj význam při záměně metriky za metriku s ní lipschitzovsky ekvivalentní.

1.56 Definice. *Řekneme, že normy $\|\cdot\|_1$, $\|\cdot\|_2$ na lineárním prostoru X jsou ekvivalentní, jestliže jsou ekvivalentní jimi indukované metriky ρ_1, ρ_2.*

Níže (Důsledek 1.125) dokážeme, že $\|\cdot\|_1$, $\|\cdot\|_2$ jsou ekvivalentní, právě když existují čísla $K > 0$, $C > 0$ taková, že $K\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$ pro $x \in X$.

1.57 Příklad. Budeme zkoumat vzájemný vztah tří norem $\|\cdot\|_\infty$, $\|\cdot\|_2$, $\|\cdot\|_1$ a jimi indukovaných metrik ρ_∞, ρ_2, ρ_1 na $C[a,b]$, viz Příklad 1.12.

Použijeme-li pro skalární součin z Příkladu 1.12 Cauchyovu nerovnost na funkce $|f|$ a 1, dostáváme

$$\|f\|_1 = \int_a^b |f| \cdot 1 = \langle |f|, 1\rangle \leq \|f\|_2 \cdot \|1\|_2 = \|f\|_2\sqrt{b-a}.$$

Platí tedy $\rho_1(g,h) \leq \rho_2(g,h)\sqrt{b-a}$, z čehož ihned vyplývá, že ρ_2 je silnější než ρ_1. Podobně dostáváme, že ρ_∞ je silnější než ρ_2, protože

$$\|f\|_2 = \sqrt{\int_a^b |f|^2} \leq \sqrt{(b-a)\|f\|_\infty^2} = \sqrt{(b-a)}\|f\|_\infty.$$

Položíme-li (v případě $a = 0, b = 1$), $f_n(x) := x^n$ a $g_n(x) := \sqrt{n}x^n$ pro $x \in [0,1]$, snadný výpočet ukazuje, že f_n konvergují k nulové funkci vzhledem k metrice ρ_2, ne však vzhledem k metrice ρ_∞, a g_n konvergují k nulové funkci vzhledem k metrice ρ_1, ne však vzhledem k metrice ρ_2. Žádné dvě ze tří metrik ρ_∞, ρ_2, ρ_1 (resp. norem $\|\cdot\|_\infty$, $\|\cdot\|_2$, $\|\cdot\|_1$) nejsou tedy ekvivalentní.

Přenesení struktury

Jestliže $f\colon X \to Y$ je bijekce a na Y máme dánu nějakou strukturu (např. metrického prostoru, normovaného lineárního prostoru, tělesa, grupy apod.), je intuitivně jasné, že pomocí bijekce f můžeme zavést na X strukturu stejného typu (tím, že „ztotožníme" prvek x s prvkem $f(x)$).

Toto „metamatematické" pozorování můžeme pro případ metrických prostorů formulovat jako matematické tvrzení (jehož důkaz je zřejmý) takto:

1.58 Tvrzení. (o přenesení metriky) *Nechť (Y,ρ) je metrický prostor a nechť $f\colon X \to (Y,\rho)$ je bijekce. Položíme-li pro x_1, $x_2 \in X$*

$$\sigma(x_1,x_2) := \rho(f(x_1), f(x_2)),$$

je σ metrika na X; je to jediná metrika na X, při které $f\colon(X,\sigma) \to (Y,\sigma)$ je izometrie.

Pomocí přenesení metriky se definuje důležitá *redukovaná metrika* na $\mathbb{R}^*$, která umožňuje chápat nevlastní limitu a limitu v nevlastním bodě jako speciální případ limity zobrazení mezi metrickými prostory.

Pro definici redukované metriky zvolíme spojitou rostoucí funkci $\varphi\colon \mathbb{R} \to \mathbb{R}$, jejímž oborem hodnot je omezený interval (α,β). Zobrazení φ přirozeně rozšíříme na zobrazení $\varphi^*\colon \mathbb{R}^* \to [\alpha,\beta]$ tím, že položíme $\varphi^*(\infty) := \beta$, $\varphi^*(-\infty) := \alpha$. Metriku σ, kterou dostaneme přenesením eukleidovské metriky z $[\alpha,\beta]$ na $\mathbb{R}^*$ pomocí zobrazení

Obr. 1.2. Znázornění redukované metriky.

φ^*, budeme nazývat *redukovanou metrikou* na $\mathbb{R}^*$ (určenou funkcí φ). Klademe tedy $\sigma(x, y) := |\varphi^*(x) - \varphi^*(y)|$; (viz obr. 5.18, který odpovídá volbě $(\alpha, \beta) := (-\pi/2, \pi/2)$, $\varphi(x) := \operatorname{arctg} x$).

Není těžké dokázat následující tvrzení.

1.59 Tvrzení. *Nechť σ je redukovaná metrika na $\mathbb{R}^*$. Jestliže a, $x_1, x_2, \dots$ jsou reálná čísla, pak platí:*

(i) $x_n \to a \iff x_n \to a$ v $(\mathbb{R}^*, \sigma)$.

(i) $x_n \to \pm\infty \iff x_n \to \pm\infty$ v $(\mathbb{R}^*, \sigma)$,

kde konvergence na levých stranách má klasický význam.

Každá jiná metrika $\widetilde{\sigma}$ s těmito vlastnostmi je se σ ekvivalentní.

Je zřejmé, že redukovaná metrika je na $\mathbb{R}$ ekvivalentní s eukleidovskou metrikou, není s ní však lipschitzovsky ekvivalentní.

1.6 Součin metrických prostorů; dvojná a dvojnásobná limita

Uvažujme nyní n metrických prostorů $(X_1, \rho_1), \dots, (X_n, \rho_n)$ a kartézský součin $X := X_1 \times \cdots \times X_n$. Položme si otázku, jak na množině X „nejpřirozeněji" definovat metriku. Vedeni analogií se speciálním případem $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$, uvažujme metriky σ_2, σ_∞, σ_1, které dvěma bodům $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n)$ z množiny X přiřazují vzdálenosti

$$\begin{aligned} \sigma_2(x,y) &:= \sqrt{(\rho_1(x_1,y_1))^2 + \cdots + (\rho_n(x_n,y_n))^2}, \\ \sigma_\infty(x,y) &:= \max\{\rho_1(x_1,y_1), \dots, \rho_n(x_n,y_n)\}, \\ \sigma_1(x,y) &:= \rho_1(x_1,y_1) + \cdots + \rho_n(x_n,y_n). \end{aligned}$$

Označíme-li $v := (\rho_1(x_1,y_1), \dots, \rho_n(x_n,y_n)) \in \mathbb{R}^n$, máme

$$\sigma_2(x,y) = \|v\|_2, \quad \sigma_\infty(x,y) = \|v\|_\infty, \quad \sigma_1(x,y) = \|v\|_1.$$

Z těchto rovností a vlastností tří užitých norem na $\mathbb{R}^n$ (srov. Poznámka 1.61) pak již snadno plyne, že σ_2, σ_∞ a σ_1 jsou skutečně metriky na X a jsou po dvou lipschitzovsky ekvivalentní. Všechny topologické pojmy a řada dalších tedy nezávisí na tom, kterou z těchto tří metrik na X uvažujeme. Ani jeden z výběrů není „kanonický", pro určitost však zvolíme σ_∞.

1.60 Definice. *Kartézským součinem metrických prostorů* $(X_1, \rho_1), \dots, (X_n, \rho_n)$ *rozumíme kartézský součin* $X := X_1 \times \cdots \times X_n$ *opatřený metrikou* σ_∞. *Tento metrický prostor označujeme symbolem* $(X_1, \rho_1) \times \cdots \times (X_n, \rho_n)$.

1.61 Poznámka.

(i) Na X můžeme pro každé $1 \le p \le \infty$ definovat metriku σ_p předpisem $\sigma_p(x,y) := \|v\|_p$. Není pravda, že pro každou normu $\|\cdot\|$ na $\mathbb{R}^n$ je předpisem $\sigma(x,y) := \|v\|$ určena metrika na X. Pro σ totiž nemusí platit trojúhelníková nerovnost; ta je však splněna (jak ukazuje krátký přímočarý výpočet), je-li $\|\cdot\|$ monotónní v tom smyslu, že

$$\|(v_1, \dots, v_n)\| \le \|(w_1, \dots, w_n)\| \quad \text{kdykoliv} \quad 0 \le v_i \le w_i, \quad i = 1, \dots, n.$$

Je zřejmé, že normy $\|\cdot\|_p$ tuto „vlastnost monotonie" mají.

(ii) Protože na X není kanonicky určena metrika, někdy se poněkud neurčitě říká, že součin metrických prostorů je množina X s některou z metrik ρ_p. Všechny tyto metriky určují na X stejné topologické pojmy a také ty pojmy, které se nemění při přechodu k lipschitzovsky ekvivalentní metrice.

Pro zobrazení f z metrického prostoru (Y, σ) do součinu $X = X_1 \times \cdots \times X_n$ metrických prostorů platí zobecnění Věty 1.53: spojitost a limita se vyšetřuje „po složkách". Složky $f_1, \dots, f_n$ zobrazení f jsou ovšem (na D_f) opět definovány rovností $f(x) = (f_1(x), \dots, f_n(x))$.

1.62 Věta. *Nechť (Y, σ) je metrický prostor, $A \subset Y$ a $f = (f_1, \dots, f_n)$ je zobrazení z Y do $X = X_1 \times \cdots \times X_n$. Potom platí:*

(i) *Nechť $a \in A$. Pak zobrazení f je spojité v bodě a vzhledem k množině A, právě když každá složka f_i, $i = 1, \dots, n$, je spojitá v bodě a vzhledem k množině A.*

(ii) *Zobrazení f je spojité na A, právě když každá složka f_i, $i = 1, \dots, n$, je spojitá na A.*

(iii) *Nechť $a \in A'$ a $b = (b_1, \dots, b_n) \in X$. Pak*
$$\lim_{x \to a,\, x \in A} f(x) = b \iff \lim_{x \to a,\, x \in A} f_i(x) = b_i,\ i = 1, \dots, n.$$

Je-li na součinu dvou metrických prostorů $(X, \rho) \times (Y, \sigma)$ zadána reálná funkce (nebo obecněji zobrazení do metrického prostoru) f, pak funkci f chápeme jako funkci „dvou proměnných" $x \in X$, $y \in Y$. Limita funkce f v bodě $(x_0, y_0) \in X \times Y$ (tedy vzhledem k oběma proměnným) se často nazývá *dvojná limita.* Pokud nejprve vypočteme limitu podle jedné proměnné a potom podle druhé, hovoří se většinou o *opakované* nebo také *dvojnásobné* limitě. Základní (snadný) vztah mezi těmito pojmy udává následující věta (pro důkaz viz [D II]).

1.63 Věta. (o dvojné limitě) *Nechť $(X, \rho), (Y, \sigma)$ a (Z, ω) jsou metrické prostory a nechť f je zobrazení z $X \times Y$ do Z. Nechť $x_0 \in X'$, $y_0 \in Y'$. Položme $M := \{(x, y) \in X \times Y\colon\ x \neq x_0, y \neq y_0\}$ a předpokládejme, že*

(i) *existuje dvojná limita*

$$\lim_{(x,y) \to (x_0, y_0), (x,y) \in M} f(x, y) = z \in Z; \tag{1.4}$$

(ii) *existuje $\delta > 0$ takové, že pro každé $x \in U_\delta(x_0) \setminus \{x_0\}$ existuje*

$$\lim_{y \to y_0} f(x, y) =: \varphi(x) \in Z.$$

Potom existuje i dvojnásobná limita

$$\lim_{x \to x_0} \varphi(x) = \lim_{x \to x_0} \left(\lim_{y \to y_0} f(x, y) \right) = z. \tag{1.5}$$

1.64 Poznámka.

(i) Je zřejmé, jakou „symetrickou podmínkou" je třeba nahradit podmínku (ii), aby z předpokladu (i) plynula existence druhé dvojnásobné limity

$$\lim_{y \to y_0} \left(\lim_{x \to x_0} f(x, y) \right) = z.$$

(ii) Z existence dvojnásobné limity (1.5) existence dvojné limity (1.4) obecně neplyne. Pokud je však navíc limita $\lim_{y \to y_0} f(x, y)$ stejnoměrná (viz Poznámka 1.34) vzhledem k x z nějakého „redukovaného okolí“ $U_\delta(x_0) \setminus \{x_0\}$, $\delta > 0$, není těžké dokázat existenci (1.4).

1.7 Separabilní a totálně omezené prostory

1.65 Definice. (separabilní prostor) *Metrický prostor se nazývá separabilní, jestliže v něm existuje spočetná hustá podmnožina.*

Je zřejmé, že separabilita je topologický pojem. Pro práci s pojmem separabilního prostoru je výhodné zavést pojem báze otevřených množin.

1.66 Definice. (báze otevřených množin) *Nechť $\mathcal{B}$ je systém otevřených podmnožin metrického prostoru X. Řekneme, že $\mathcal{B}$ je báze otevřených množin prostoru X, jsou-li splněny následující podmínky, které jsou ekvivalentní.*

(i) *Pro každou otevřenou množinu $G \subset X$ existuje systém $\mathcal{B}^* \subset \mathcal{B}$ takový, že $G = \bigcup \mathcal{B}^*$.*

(ii) *Je-li $G \subset X$ otevřená množina a $x \in G$, pak existuje $H \in \mathcal{B}$ taková, že $x \in H \subset G$.*

1.67 Věta. *Metrický prostor je separabilní, právě když existuje spočetná báze otevřených množin prostoru X.*

Z této charakterizace separabilních prostorů a z Tvrzení 1.22 například okamžitě dostáváme, že *podprostor separabilního metrického prostoru je opět separabilní* a také následující hlubší větu.

1.68 Věta. (separabilní metrický prostor je Lindelöfův) *Nechť (X, ρ) je separabilní metrický prostor a $X = \bigcup_{\alpha \in A} G_\alpha$, kde všechny množiny G_α jsou otevřené. Pak existuje spočetná $S \subset A$ taková, že $X = \bigcup_{\alpha \in S} G_\alpha$. (Stručněji: z každého otevřeného pokrytí separabilního prostoru lze vybrat spočetné pokrytí.)*

Důkaz. (Náznak.) Nechť $\mathcal{B}$ je spočetná báze otevřených množin prostoru X a $\mathcal{B}^*$ je množina těch $B \in \mathcal{B}$, které jsou částí některé množiny G_α, $\alpha \in A$. Pro každé $B \in \mathcal{B}^*$ zvolme α_B takové, že $B \subset G_{\alpha_B}$. Je snadno vidět, že stačí položit $S := \{\alpha_B \colon B \in \mathcal{B}^*\}$.

1.69 Poznámka. Snadným důsledkem předchozí věty je toto její zesílení:

Z otevřeného pokrytí libovolné podmnožiny separabilního metrického prostoru lze vybrat spočetné pokrytí této množiny.

1.70 Tvrzení. *Součin $X := X_1 \times \cdots \times X_n$ separabilních metrických prostorů $X_1, \ldots X_n$ je opět separabilní prostor.*

1.71 Definice. *Nechť X je metrický prostor a $A \subset X$.*

Jestliže $A' = \emptyset$, řekneme, že množina A je diskrétní.

Jestliže $A' \cap A = \emptyset$, řekneme, že A je izolovaná.

1.72 Poznámka.

(a) Množina A je izolovaná, právě když každý její bod je jejím izolovaným bodem. Množina A je diskrétní, právě když je izolovaná a uzavřená.

(b) Terminologie kolísá, dokonce někteří autoři diskrétní (resp. izolovanou) množinu ve smyslu Definice 1.71 nazývají izolovanou (resp. diskrétní).

1.73 Tvrzení. *Nechť X je metrický prostor. Pak následující tvrzení jsou ekvivalentní.*

(i) Prostor X je separabilní.

(ii) Každá izolovaná množina $A \subset X$ je spočetná.

(iii) Každá diskrétní množina $A \subset X$ je spočetná.

(iv) Každý disjunktní systém otevřených neprázdných podmnožin X je spočetný.

1.74 Definice. *Nechť (X, ρ) je metrický prostor a $\varepsilon > 0$. Řekneme, že $M \subset X$ je ε-síť v X, jestliže pro každý bod $x \in X$ existuje bod $y \in M$ takový, že $\rho(x, y) < \varepsilon$.*

1.75 Definice. (totálně omezený prostor) *Metrický prostor (X, ρ) se nazývá totálně omezený, jestliže pro každé $\varepsilon > 0$ existuje konečná ε-síť v X. Množina $Y \subset X$ se nazývá totálně omezená, jestliže podprostor (Y, ρ) je totálně omezený.*

Není těžké dokázat následující tvrzení.

1.76 Tvrzení. *Nechť X je totálně omezený metrický prostor. Pak platí:*

(i) *X je omezený prostor.*

(ii) *X je separabilní prostor.*

(iii) *Každá množina $Y \subset X$ je totálně omezená.*

1.77 Tvrzení. *Součin $X := X_1 \times \cdots \times X_n$ totálně omezených metrických prostorů $X_1, \ldots X_n$ je opět totálně omezený prostor.*

Je snadno vidět, že pojem totálně omezeného prostoru *není topologický pojem*, nemění se však při přechodu k lipschitzovsky ekvivalentní metrice.

1.8 Úplné metrické prostory

Pojem úplného prostoru je velmi důležitý, zejména při práci s nekonečně dimenzionálními (Banachovými) prostory, kdy většinou nelze použít teorii (speciálnějších) kompaktních prostorů. Zvlášť důležité jsou aplikace Banachovy věty o kontrakci (Věta 1.87) a Baireovy věty.

1.78 Definice. *Posloupnost (x_n) bodů metrického prostoru (X, ρ) se nazývá cauchyovská, jestliže platí:*

$$\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n, m \in \mathbb{N}: \quad (n \geq n_0, m \geq n_0) \implies \rho(x_n, x_m) < \varepsilon. \tag{1.6}$$

1.79 Poznámka. Cauchyovské posloupnosti se říká také *fundamentální posloupnost*. Podmínka (1.6) je Bolzano-Cauchyova podmínka.

Snadno je vidět, že pokud je posloupnost (x_n) konvergentní v X, je také cauchyovská.

1.80 Definice. (úplný prostor) *Řekneme, že metrický prostor X je úplný, jestliže každá cauchyovská posloupnost bodů prostoru X je konvergentní v X.*

1.81 Poznámka. Není těžké nahlédnout, že pojem úplného prostoru *není topologický pojem*, nemění se však při přechodu k lipschitzovsky ekvivalentní metrice. Často se užívá následující snadné tvrzení.

1.82 Tvrzení. (úplnost podprostoru) *Nechť (X, ρ) je metrický prostor a $M \subset X$. Pak platí následující tvrzení.*

- (i) *Je-li (M, ρ) úplný prostor, pak M je uzavřená podmnožina prostoru X.*
- (ii) *Nechť (X, ρ) je úplný. Pak prostor (M, ρ) je úplný, právě když M je uzavřená podmnožina prostoru X.*

Snadný je i důkaz této věty:

1.83 Věta. *Součin $X := X_1 \times \cdots \times X_n$ úplných metrických prostorů $X_1, \ldots X_n$ je opět úplný prostor.*

1.84 Poznámka. Protože $\mathbb{R}$ je úplný prostor, je podle předchozí věty $(\mathbb{R}^n, \rho_\infty)$ také úplný. Je tedy (viz Poznámka 1.81) úplný také eukleidovský prostor $(\mathbb{R}^n, \rho_2)$.

Metrický prostor $\mathbb{C}$ je zřejmě izometrický s $(\mathbb{R}^2, \rho_2)$, takže je úplný. Protože metrika (indukovaná skalárním součinem) na $\mathbb{C}^n$ je lipschitzovsky ekvivalentní se součinovou (maximovou) metrikou na $\mathbb{C}^n$, je úplný i prostor $\mathbb{C}^n$. (Jiný argument: $\mathbb{C}^n$ je izometrický s $\mathbb{R}^{2n}$.)

Následující věta je snadným zobecněním klasického Cantorova principu vložených intervalů.

1.85 Věta. *Nechť X je úplný prostor a $F_1 \supset F_2 \supset \dots$ je posloupnost neprázdných uzavřených množin v X, pro kterou $\lim\limits_{n\to\infty} \operatorname{diam} F_n = 0$. Pak $\bigcap\limits_{n=1}^{\infty} F_n$ je jednobodová množina.*

Nyní uvedeme důležitou Banachovu větu o pevném bodě (o kontrakci), která je jistým „abstraktním zachycením“ metody postupných aproximací (ta se ovšem používá v důkazu věty).

1.86 Definice. (pevný bod, kontrakce)

(i) *Řekneme, že bod x je pevný bod zobrazení $f\colon X \to X$, jestliže $f(x) = x$.*

(ii) *Nechť (X, ρ) je metrický prostor. Řekneme, že zobrazení $f\colon X \to X$ je kontrakce, existuje-li číslo $0 \le q < 1$ takové, že pro každé dva body $x, y \in X$ platí*

$$\rho(f(x), f(y)) \le q \cdot \rho(x, y).$$

1.87 Věta. (Banachova věta o pevném bodě, princip kontrakce) *Nechť (X, ρ) je úplný prostor, $X \neq \emptyset$ a $f : X \to X$ je kontrakce. Pak f má právě jeden pevný bod.*

Abychom zformulovali „Baireovu větu o kategorii“, potřebujeme zavést několik pojmů.

1.88 Definice. *Nechť (X, ρ) je metrický prostor. Řekneme, že množina $A \subset X$ je řídká (v X), jestliže je množina $X \setminus \overline{A}$ hustá.*

1.89 Poznámka. Anglický termín pro řídkou množinu je „nowhere dense set“. Tato názornější terminologie odpovídá podmínce (iii) z následujícího tvrzení.

1.90 Tvrzení. *Nechť A je podmnožina metrického prostoru (X, ρ). Pak následující podmínky jsou ekvivalentní.*

(i) *A je řídká.*

(ii) *$(\overline{A})^o = \emptyset$.*

(iii) *Je-li $G \subset X$ neprázdná otevřená množina, pak $A \cap G$ není hustá v (G, ρ).*

(iv) *Je-li $G \subset X$ neprázdná otevřená množina, pak existuje otevřená neprázdná $H \subset G$ taková, že $H \cap A = \emptyset$.*

Pomocí podmínky (iv) snadno dostáváme, že *konečné sjednocení řídkých množin je řídká množina.*

1.91 Definice. *Řekneme, že podmnožina A metrického prostoru (X, ρ) je množina 1. kategorie (v X), je-li A sjednocením spočetně mnoha množin řídkých v X.*

Množina, která není 1. kategorie, se nazývá množina 2. kategorie. Doplněk (v X) množiny 1. kategorie se nazývá reziduální množina.

1.92 Věta. (Baireova) *Nechť X je úplný prostor a $G \neq \emptyset$ je otevřená podmnožina X. Pak G není 1. kategorie v X (tj. G je 2. kategorie v X).*

1.93 Poznámka.

(i) Snadno je vidět, že Baireovu větu lze ekvivalentně formulovat takto: *Reziduální množina v úplném metrickém prostoru je hustá.*

(ii) Baireova věta se často formuluje ve slabší formě; tvrdí se, že úplný metrický prostor není (v sobě) 1. kategorie, tj. každá jeho reziduální podmnožina je neprázdná. V této formě se také nejčastěji používá v důkazech existence *Baireovou metodou kategorií*, ve kterých se postupuje tímto způsobem:

Chceme dokázat existenci objektu, který má vlastnost V. Zvolíme úplný metrický prostor X (ten se většinou přirozeně nabízí) a dokážeme, že množina $\{x \in X\colon x$ má vlastnost $V\}$ je reziduální; tj. $\{x \in X\colon x$ nemá vlastnost $V\}$ je 1. kategorie. Podle Baireovy věty pak víme, že hledaný objekt existuje.

Tuto metodu lze použít například pro důkaz existence spojité funkce na $[0,1]$, která nemá v žádném bodě derivaci, bere-li se za X prostor $\mathcal{C}[0,1]$.

Pojem úplného prostoru je také velmi důležitý při vyšetřování existence limity zobrazení a existence spojitého rozšíření zobrazení.

1.94 Věta. *Nechť f je zobrazení z metrického prostoru (X,ρ) do úplného metrického prostoru (Y,σ). Nechť $A \subset X$ a $a \in A'$. Pak limita $\lim_{x\to a, x\in A} f(x)$ existuje, právě když je splněna následující (Bolzano-Cauchyova) podmínka*

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x,y \in A \cap (U_\delta(a) \setminus \{a\}):\quad \sigma(f(x),f(y)) < \varepsilon. \tag{1.7}$$

Níže (Věta 6.8) je dokázáno zobecnění předchozí věty.

1.95 Věta. *Nechť (X,ρ), (Y,σ) jsou metrické prostory, Y je úplný, $A \subset X$ a $f\colon (A,\rho) \to (Y,\sigma)$ je stejnoměrně spojité zobrazení. Pak existuje právě jedno spojité zobrazení $f^*\colon (\overline{A},\rho) \to (Y,\sigma)$, které je rozšířením zobrazení f. Přitom f^* je stejnoměrně spojité.*

Důkaz. (Náznak.) Zobrazení f zřejmě splňuje v každém bodě $a \in \overline{A} \setminus A$ Bolzano-Cauchyovu podmínku (1.7). Podle předchozí věty tedy existuje $\lim_{x\to a, x\in A} f(x) =: f^*(a)$. Položíme-li ještě $f^*(x) := f(x)$ pro $x \in A$, je snadné z definice ověřit, že f^* je stejnoměrně spojité rozšíření f. Jednoznačnost je zřejmá.

1.96 Definice. *Řekneme, že metrický prostor Q je úplným obalem metrického prostoru P, jestliže*

(i) *P je podprostorem Q,*
(ii) *Q je úplný prostor a*
(iii) *P je hustý v Q.*

1.97 Poznámka.

(a) Úplnému obalu prostoru P se také říká *zúplnění* P.

(b) Je-li P podprostorem úplného prostoru Q^*, je $\overline{P}^{Q^*}$ zřejmě úplným obalem P. Existence úplného obalu je tedy snadno ekvivalentní s existencí úplného nadprostoru.

Následující věta říká, že úplný obal existuje a je určen (co do struktury) jednoznačně.

1.98 Věta. *Každý metrický prostor má úplný obal. Jsou-li Q_1 a Q_2 úplné obaly P, pak existuje izometrie $f\colon Q_1 \to Q_2$ taková, že $f(x) = x$ pro každý bod $x \in P$.*

1.99 Poznámka. Prostor reálných čísel je ovšem zúplněním svého podprostoru všech racionálních čísel. Obvyklá konstrukce úplného obalu je zcela analogická Cantorově metodě konstrukce reálných čísel z čísel racionálních: prvky úplného obalu prostoru P jsou třídy ekvivalence cauchyovských posloupností v prostoru P.

Jinou možností je nalézt izometrii $f\colon P \to Q^*$ do vhodného úplného prostoru Q^*; není těžké si rozmyslet, že pak je sestrojení úplného obalu prostoru P snadné. (Je-li P omezený prostor, lze $f(x)$ definovat jako funkci $y \mapsto \operatorname{dist}(y, x)$, která je prvkem (úplného) prostoru všech omezených funkcí na P se supremovou metrikou).

1.9 Kompaktní prostory

Kompaktní metrické prostory mají řadu důležitých vlastností uzavřeného omezeného intervalu. Dají se charakterizovat například svou vlastností, která je velmi užitečná pro aplikace: každá spojitá reálná funkce na kompaktním prostoru nabývá svého maxima a minima. Teorie kompaktních metrických prostorů má velmi důležité aplikace zejména při práci s konečně rozměrnými prostory. V nekonečně rozměrných prostorech se aplikuje úspěšně hlavně pojem kompaktního *topologického prostoru.*

Nejobvyklejší je následující „pokrývací“ definice (pomocí které se definuje i pojem kompaktního topologického prostoru).

1.100 Definice. (kompaktní prostor) *Řekneme, že metrický prostor (X, ρ) je kompaktní, jestliže z každého pokrytí prostoru X otevřenými množinami lze vybrat konečné pokrytí. Jinými slovy: jestliže $X = \bigcup_{\alpha \in A} G_\alpha$ a všechny množiny G_α jsou otevřené, pak existuje konečná množina $K \subset A$ taková, že $X = \bigcup_{\alpha \in K} G_\alpha$.*
Řekneme, že množina $M \subset (X, \rho)$ je kompaktní, jestliže podprostor (M, ρ) je kompaktní.

Pojem kompaktního prostoru je zřejmě topologický pojem.

Z Tvrzení 1.22 okamžitě plyne, že $M \subset (X,\rho)$ je kompaktní, právě když z každého pokrytí množiny M množinami *otevřenými v X* lze vybrat konečné pokrytí. Tato vlastnost kompaktní množiny zobecňuje vlastnost intervalu $[a,b]$ známou z Borelovy věty.

Často se užívají vlastnosti ekvivalentní s kompaktností, které jsou obsaženy v následující větě.

1.101 Věta. *Nechť (X,ρ) je metrický prostor. Pak následující tvrzení jsou ekvivalentní.*

(i) *Prostor X je kompaktní.*
(ii) *Z každé posloupnosti bodů z X lze vybrat konvergentní posloupnost.*
(iii) *Každá nekonečná množina $M \subset X$ má v X aspoň jeden hromadný bod (tj. $M' \neq \emptyset$).*
(iv) *Je-li $X \supset F_1 \supset F_2 \supset F_3 \supset \dots$ nerostoucí posloupnost neprázdných uzavřených množin, pak*

$$\bigcap_{n=1}^{\infty} F_n \neq \emptyset.$$

(v) *Z každého spočetného pokrytí prostoru X otevřenými množinami lze vybrat konečné pokrytí.*

1.102 Poznámka. V případě topologického prostoru X podmínky (ii)–(v) plynou z kompaktnosti X, ale ne naopak. Topologické prostory splňující (ii) se nazývají sekvenciálně kompaktní.

Snadno je vidět, že podmnožina kompaktního prostoru je kompaktní, právě když je uzavřená. O něco hlubší je následující věta.

1.103 Věta. *Součin $X := X_1 \times \cdots \times X_n$ kompaktních metrických prostorů $X_1, \dots X_n$ je opět kompaktní prostor.*

Pro klasické aplikace kompaktnosti jsou důležité zejména následující dvě věty.

1.104 Věta. (kompaktní podmnožiny $\mathbb{R}^n$) *Množina $M \subset \mathbb{R}^n$ je kompaktní právě tehdy, je-li v $\mathbb{R}^n$ uzavřená a omezená.*

1.105 Věta. *Spojitá reálná funkce na neprázdném kompaktním metrickém prostoru nabývá svého maxima a minima (a je tudíž omezená).*

Nejdůležitější vlastnosti spojitých zobrazení definovaných na kompaktním prostoru jsou shrnuty v následující větě.

1.106 Věta. *Nechť $f\colon (X,\rho) \to (Y,\sigma)$ je spojité zobrazení a metrický prostor (X,ρ) je kompaktní. Pak platí:*

(i) *Zobrazení f je stejnoměrně spojité.*
(ii) *Množina $f(X)$ je kompaktní.*
(iii) *Je-li f bijekce, pak f je homeomorfismus.*

Aplikujeme-li vlastnost (iii) na identické zobrazení, dostáváme:

1.107 Tvrzení. *Nechť na kompaktním metrickém prostoru (X, ρ) je dána druhá metrika σ, která je slabší než ρ. Pak ρ a σ jsou ekvivalentní.*

Následující dvě věty ukazují souvislost mezi kompaktností, úplností a totální omezeností.

1.108 Věta. *Metrický prostor je kompaktní, právě když je úplný a totálně omezený.*

1.109 Věta. *Nechť X je úplný metrický prostor a $A \subset X$. Pak A je totálně omezená množina, právě když $\overline{A}$ je kompaktní množina.*

1.110 Poznámka. Podmnožině metrického prostoru se někdy říká *prekompaktní*, má-li kompaktní uzávěr. Předchozí věta ukazuje, že v *úplných prostorech* pojem prekompaktní množiny splývá s pojmem totálně omezené množiny.

1.10 Souvislé prostory

1.111 Definice. *Nechť (X, ρ) je metrický prostor. Řekneme, že $A \subset X$ je obojetná množina, je-li zároveň otevřená i uzavřená.*

1.112 Lemma. *Nechť metrický prostor (X, ρ) je sjednocením disjunktních neprázdných množin A, B. Pak následující tvrzení jsou ekvivalentní.*

(i) *A, B jsou otevřené množiny.*
(ii) *A, B jsou uzavřené množiny.*
(iii) *A je obojetná množina.*
(iv) $\overline{A} \cap B = \overline{B} \cap A = \emptyset$.
(v) $\overline{A} \cap \overline{B} = \emptyset$.
(vi) *$M \subset X$ je otevřená množina, právě když $M \cap A$ je otevřená v (A, ρ) a $M \cap B$ je otevřená v (B, ρ).*
(vii) *Pro každý metrický prostor (Y, σ) a každé zobrazení $f: X \to Y$ je f spojité, právě když obě zobrazení $f\restriction_A: (A, \rho) \to Y$, $f\restriction_B: (B, \rho) \to Y$ jsou spojitá.*
(viii) *Charakteristická funkce $C_A: X \to R$ množiny A je spojitá.*

Zejména tvrzení (vi) a (vii) ukazují, že za daných podmínek je prostor X rozložen na dva prostory, které spolu „nesouvisí“. Následující definice je tedy zcela přirozená.

1.113 Definice. *Metrický prostor se nazývá souvislý, není-li sjednocením dvou disjunktních otevřených neprázdných podmnožin.*

Každá z ekvivalentních podmínek z předchozího lemmatu vede ovšem k ekvivalentní definici souvislého prostoru.

1.114 Definice. *Řekneme, že podmnožina M metrického prostoru (X, ρ) je souvislá, je-li souvislý podprostor (M, ρ).*

Z Lemmatu 1.112 okamžitě dostáváme následující charakterizaci souvislých množin, která neužívá pojem podprostoru.

1.115 Tvrzení. *Nechť (X, ρ) je metrický prostor a $M \subset X$. Pak následující tvrzení jsou ekvivalentní.*

(i) *Množina M je souvislá.*
(ii) *Je-li $M = A \cup B$ a $\overline{A} \cap B = \overline{B} \cap A = \emptyset$, pak $A = \emptyset$ nebo $B = \emptyset$.*
(iii) *Je-li $M = A \cup B$ a $\overline{A} \cap \overline{B} \cap M = \emptyset$, pak $A = \emptyset$ nebo $B = \emptyset$.*

1.116 Tvrzení. *Neprázdná množina $M \subset \mathbb{R}$ je souvislá, právě když M je interval nebo jednobodová množina.*

Důkaz. Nechť M je interval a předpokládejme, že M není souvislá množina. Podle Lemmatu 1.112 (viii) existují disjunktní neprázdné množiny A, B takové, že $M = A \cup B$ a charakteristická funkce C_A je spojitá v M. Má tedy C_A Darbouxovu vlastnost na intervalu M. V některém bodě proto nabývá hodnoty $1/2$, a to je spor. Není-li M interval ani jednobodová množina, položme $\alpha := \inf M$, $\beta := \sup M$. Protože M není interval a $\alpha < \beta$, existuje číslo $c \in (\alpha, \beta) \setminus M$. Položíme-li $A := (-\infty, c) \cap M$ a $B := (c, \infty) \cap M$, pak zřejmě $M = A \cup B$, $A \neq \emptyset$, $B \neq \emptyset$ a $\overline{A} \cap \overline{B} \cap M = \emptyset$. Podle Tvrzení 1.115 (iii) není M souvislá množina, což je spor.

1.117 Věta. *Nechť (X, ρ) je metrický prostor. Pak platí:*

(i) *Je-li $A \subset X$ souvislá a $A \subset M \subset \overline{A}$, pak M je také souvislá.*
(ii) *Jsou-li A_ω, $\omega \in \Omega$, souvislé podmnožiny X a $\bigcap_{\omega \in \Omega} A_\omega \neq \emptyset$, pak množina $\bigcup_{\omega \in \Omega} A_\omega$ je také souvislá.*

1.118 Věta. *Nechť $f: (X, \rho) \to (Y, \sigma)$ je spojité zobrazení. Je-li X souvislý, pak $f(X)$ je souvislá množina.*

1.119 Definice. *Řekneme, že metrický prostorX je křivkově (obloukově) souvislý, jestliže pro každé dva body a, $b \in X$ existuje spojité zobrazení $\varphi: [0, 1] \to X$ takové,* že $\varphi(0) = a$ a $\varphi(1) = b$.

1.120 Tvrzení.

(i) *Každý křivkově souvislý metrický prostor je souvislý.*
(ii) *Je-li $G \subset \mathbb{R}^n$ otevřená množina, pak G je souvislá, právě když je křivkově souvislá.*

Tvrzení (i) ihned vyplývá z Tvrzení 1.116 a Věty 1.118. Tvrzení (ii) lze přirozeně zobecnit a zesílit: Každé dva body otevřené souvislé podmnožiny G normovaného lineárního prostoru lze spojit „lomenou čarou" ležící v G.

1.121 Definice. *Řekneme, že množina $C \subset X$ je komponenta (souvislosti) metrického prostoru (X, ρ), jestliže C je maximální souvislá množina v X (tj. v X neexistuje její souvislá vlastní nadmnožina).*

Komponentou množiny $M \subset X$ rozumíme komponentu podprostoru (M, ρ).

Z Věty 1.117 snadno vyplývá následující tvrzení.

1.122 Tvrzení. *Nechť $X \neq \emptyset$ je metrický prostor. Pak:*

(i) *Systém všech komponent prostoru X tvoří rozklad X na neprázdné uzavřené souvislé množiny.*
(ii) *Je-li $G \subset \mathbb{R}^n$ otevřená množina, pak komponenty množiny G jsou otevřené množiny.*

Každý bod $x \in X$ tedy leží v právě jedné komponentě prostoru X; té se někdy říká *komponenta souvislosti bodu x*.

1.11 Lineární zobrazení mezi normovanými lineárními prostory

Nechť $L: \mathbb{R}^n \to \mathbb{R}^k$ je lineární zobrazení. Z lineární algebry víme, že L je reprezentováno maticí typu $k \times n$, která je určena jednoznačně (a kterou v této publikaci označujeme symbolem $[L]$). V maticovém zápisu (ve kterém prvek $\mathbb{R}^p$ ztotožňujeme s maticí typu $1 \times p$ a A^T je matice transponovaná k matici A) platí

$$L(v)^T = [L] \cdot v^T.$$

Je-li $[L] = (a_{ij})$, $i = 1, \dots, k$, $j = 1, \dots, n$, a $L = (L_1, \dots, L_k)$, pak pro vektor $v = (v_1, \dots, v_n)$ platí

$$L_i(v_1,\dots,v_n)=\sum_{j=1}^{n} a_{ij}v_j,\quad i=1,\dots,k.$$

Protože projekce $v \mapsto v_j$ jsou spojité reálné funkce na $\mathbb{R}^n$, má zobrazení L zřejmě spojité složky, a je tedy spojité. Zcela snadno jsme tedy dokázali tvrzení:

Každé lineární zobrazení $L: \mathbb{R}^n \to \mathbb{R}^k$ je spojité.

V tomto oddílu toto snadné tvrzení zesílíme a zobecníme — dokážeme, že zobrazení L je dokonce lipschitzovské a místo eukleidovských prostorů lze brát libovolné konečně dimenzionální normované lineární prostory. Nejdříve však dokážeme základní větu o spojitých lineárních zobrazeních mezi zcela obecnými normovanými lineárními prostory (nad stejným tělesem). V její formulaci používáme úmluvu z Poznámky 1.6.

1.123 Věta. *Nechť X, Y jsou normované lineární prostory a $L: X \to Y$ je lineární zobrazení. Pak následující tvrzení jsou ekvivalentní:*

(a) *L je spojité v 0.*
(b) *L je spojité zobrazení.*
(c) *L je lipschitzovské zobrazení.*
(d) *Existuje taková konstanta $C \ge 0$, že $\|L(x)\| \le C\|x\|$ pro každý bod $x \in X$.*
(e) $\sup\{\|L(x)\|\colon \|x\| \le 1\} < \infty$.

Pokud tyto podmínky platí, pak reálné číslo $\|L\| := \sup\{\|L(x)\|\colon \|x\| \le 1\}$ se nazývá norma lineárního zobrazení L. Přitom $\|L\|$ je nejmenší konstanta, s kterou je L lipschitzovské. Dále platí nerovnost

$$\|L(x)\| \le \|L\|\cdot\|x\|,\qquad x\in X \tag{1.8}$$

a $\|L\|$ je nejmenší číslo, které má vlastnost čísla C z (d).

Důkaz. Budeme postupovat podle schématu

$$(e) \Longrightarrow (d) \Longrightarrow (c) \Longrightarrow (b) \Longrightarrow (a) \Longrightarrow (e).$$

Nechť $\|L\| := \sup\{\|L(x)\|\colon \|x\| \le 1\} < \infty$ a $x \in X, x \ne 0$. Protože $\|x/\|x\|\| = 1$, máme

$$\|L(x)\| = \left\|\|x\|\cdot L\left(\frac{x}{\|x\|}\right)\right\| = \|x\|\cdot\left\|L\left(\frac{x}{\|x\|}\right)\right\| \le \|x\|\cdot\|L\|,$$

takže podmínka (d) platí s $C := \|L\|$.

Je-li splněna podmínka (d) a jsou dány body $a, b \in X$, pak

$$\|L(b)-L(a)\| = \|L(b-a)\| \le C\|b-a\|,$$

takže L je lipschitzovské s konstantou C.

Implikace $(c) \Longrightarrow (b)$ a $(b) \Longrightarrow (a)$ jsou triviální.

Nechť platí podmínka (a). Protože $L(0)=0$, podle definice spojitosti můžeme zvolit $\delta > 0$ takové, že pro každé $z \in X$ platí implikace

$$\|z\| < \delta \Longrightarrow \|L(z)\| < 1.$$

Je-li nyní $x \in X,\ 0 < \|x\| \leq 1$, pak pro bod $z := \frac{\delta}{2}x$ platí $\|z\| = \frac{\delta}{2}\|x\| < \delta$, takže

$$\|L(x)\| = \left\|\frac{2}{\delta}L\left(\frac{\delta}{2}x\right)\right\| = \frac{2}{\delta}\|L(z)\| < \frac{2}{\delta}.$$

Platí tedy podmínka (e).

Z důkazů implikací $(e) \Longrightarrow (d)$ a $(d) \Longrightarrow (c)$ vyplývá, že pokud platí (e), pak platí (1.8) a L je lipschitzovské s konstantou $\|L\|$.

Je-li L lipschitzovské s konstantou C, pak $\|L(x)\| = \|L(x) - L(0)\| \leq C\|x\|$, takže C má vlastnost z (d). A má-li C tuto vlastnost, pak $\|L(x)\| \leq C$ pro $x \in X,\ \|x\| \leq 1$, takže $\|L\| \leq C$. Tím jsou dokázány obě charakterizace čísla $\|L\|$ z druhé části věty.

1.124 Poznámka.

(i) Hodnota normy $\|L\|$ zobrazení L se ovšem může změnit, pokud změníme normu na prostoru X nebo Y.

(ii) Je snadno vidět, že $\|L\| = \sup\{\|L(x)\|\colon\ \|x\| = 1\}$.

Z Věty 1.123 snadno dostáváme:

1.125 Důsledek. *Nechť $\|\cdot\|_1$, $\|\cdot\|_2$ jsou normy na lineárním prostoru X a ρ_1, ρ_2 jsou jimi indukované metriky. Pak následující podmínky jsou ekvivalentní.*

(i) *Normy $\|\cdot\|_1$, $\|\cdot\|_2$ jsou ekvivalentní.*

(ii) *Existují čísla $K > 0$, $C > 0$ taková, že $K\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1$ pro $x \in X$.*

(iii) *Metriky ρ_1, ρ_2 jsou ekvivalentní.*

(iv) *Metriky ρ_1, ρ_2 jsou lipschitzovsky ekvivalentní.*

1.126 Lemma. *Nechť $(Z, \|\cdot\|)$ je reálný normovaný lineární prostor a nechť $F\colon \mathbb{R}^n \to Z$ je lineární zobrazení. Uvažujme na $\mathbb{R}^n$ eukleidovskou normu. Pak platí:*

(i) *Zobrazení F je spojité.*

(ii) *Je-li F bijekce, pak F je homeomorfismus.*

Důkaz. (i) Nechť $x = (x_1, \dots, x_n) \in \mathbb{R}^n,\ \|x\| \leq 1$. Pak

$$\begin{aligned}\|F(x)\| &= \|F(x_1e_1 + \cdots + x_ne_n)\| = \|x_1F(e_1) + \cdots + x_nF(e_n)\| \leq \\ &\leq \|x_1F(e_1)\| + \cdots + \|x_nF(e_n)\| = |x_1| \cdot \|F(e_1)\| + \cdots + |x_n| \cdot \|F(e_n)\| \leq \\ &\leq \|F(e_1)\| + \cdots + \|F(e_n)\|.\end{aligned}$$

Podle Věty 1.123 je tedy zobrazení F spojité.

(ii) Protože norma $\|\cdot\|$ je spojitá na Z, je také funkce $\varphi\colon x \mapsto \|F(x)\|$ spojitá na $\mathbb{R}^n$. Protože F je lineární bijekce, platí $F(x) \neq 0$, a tedy $\varphi(x) > 0$ pro každý bod x sféry $S := \{x \in \mathbb{R}^n\colon\ \|x\| = 1\}$. Protože S je zřejmě omezená a uzavřená, je podle Věty 1.104 kompaktní. Podle Věty 1.105 tedy φ na S v nějakém bodě x_0 nabývá svého minima, takže

$$c := \min\{\varphi(x) \colon x \in S\} = \varphi(x_0) > 0.$$

Dostáváme tudíž pro každý bod $0 \neq x \in \mathbb{R}^n$ odhad

$$\|F(x)\| = \|x\| \left\| F\left(\frac{x}{\|x\|}\right) \right\| \geq c\|x\|.$$

Je-li nyní $0 \neq z \in Z$ a položíme $x := F^{-1}(z)$, máme $\frac{1}{c}\|z\| \geq \|F^{-1}(z)\|$, takže zobrazení F^{-1} je podle Věty 1.123 spojité.

1.127 Poznámka. Není těžké ukázat, že tvrzení lemmatu platí i v případě, že Z je komplexní prostor a místo $\mathbb{R}^n$ uvažujeme $\mathbb{C}^n$.

1.128 Definice. *Nechť X, Y jsou normované lineární prostory a $f \colon X \to Y$ je bijekce. Řekneme, že f je izomorfismus (normovaných lineárních prostorů), jestliže je lineární a homeomorfní. Množinu všech izomorfismů $f \colon X \to Y$ budeme značit symbolem* $\mathrm{Izom}(X, Y)$. *Řekneme, že X a Y jsou izomorfní, jestliže existuje izomorfismus $f \colon X \to Y$.*

Je zřejmé, že bijekce $f \colon X \to Y$ je izomorfismus, právě když $f^{-1} \colon Y \to X$ je izomorfismus.

1.129 Poznámka.

(i) Izomorfismus normovaných lineárních prostorů ovšem nezachovává strukturu normovaných lineárních prostorů, ale pouze lineární strukturu a topologii; v tom je ustálená terminologie trochu zavádějící. Izomorfismu se někdy (ještě méně logicky) říká lineární izomorfismus. Někteří autoři tyto termíny nepoužívají a hovoří o lineárním homeomorfismu.

(ii) Jsou-li X a Y obecné normované lineární prostory a $f \colon X \to Y$ je spojitá lineární bijekce, f^{-1} nemusí být izomorfismus. (Například identické zobrazení $I \colon (C[a,b], \|\cdot\|_\infty) \to (C[a,b], \|\cdot\|_2)$ je spojité, ale jeho inverze spojitá není, viz Příklad 1.57).

Pokud však navíc X a Y jsou úplné (tj. jsou Banachovy), dosti hluboká Banachova věta dává, že f je nutně izomorfismus.

Bijekce f prostoru $(X, \|\cdot\|_1)$ na $(Y, \|\cdot\|_2)$ zachovává strukturu normovaných prostorů, jestliže je lineární a zachovává normu, tj. $\|f(x)\|_2 = \|x\|_1$ pro každé $x \in X$. Tyto dvě vlastnosti platí, právě když f je lineární izometrie.

To je vidět z toho, že pokud f je lineární a $x, y \in X$, pak $\|f(x) - f(y)\|_2 = \|f(x-y)\|_2$ a $\|f(x)\|_2 = \|f(x) - f(0)\|_2$.

Řekneme, že X a Y jsou *izometricky izomorfní*, jestliže existuje lineární izometrie $f \colon X \to Y$.

1.130 Věta. *Nechť X, Y jsou normované lineární prostory, X je konečně dimenzionální a $L: X \to Y$ je lineární zobrazení. Pak L je spojité. Je-li L navíc bijekce, pak L je izomorfismus.*

Důkaz. Důkaz provedeme pouze pro případ reálných prostorů. Označme $n := \dim X$ a zvolme lineární bijekci $\varphi: \mathbb{R}^n \to X$. Pak podle Lemmatu 1.126 je φ homeomorfismus. Položme $M := L \circ \varphi$. Zobrazení $M: \mathbb{R}^n \to Y$ je lineární, a proto je podle Lemmatu 1.126 spojité. Protože $L = M \circ \varphi^{-1}$, je spojité i L. Je-li L bijekce, z lineární algebry víme, že zobrazení $L^{-1}: Y \to X$ je lineární a Y je konečně dimenzionální. Podle první části věty je tedy L^{-1} spojité.

1.131 Důsledek. *Libovolné dva normované lineární prostory (nad stejným tělesem) stejné konečné dimenze jsou izomorfní.*

Užijeme-li předchozí větu na identické zobrazení $I : X \to X$, dostáváme následující větu.

1.132 Věta. *Libovolné dvě normy na konečně dimenzionálním lineárním prostoru X jsou ekvivalentní.*

Množinu všech spojitých lineárních zobrazení normovaného lineárního prostoru X do normovaného lineárního prostoru Y budeme označovat symbolem $\mathcal{L}(X, Y)$ a položíme $\mathcal{L}(X) := \mathcal{L}(X, X)$. Ve funkcionální analýze se prvky $\mathcal{L}(X, Y)$ často nazývají spojité lineární operátory (nebo také omezené lineární operátory).

Je snadné dokázat, že pokud $f \in \mathcal{L}(X, Y)$, $g \in \mathcal{L}(X, Y)$ a $\alpha \in \mathbb{R}$, pak zobrazení $f + g : x \mapsto f(x) + g(x)$ a $\alpha \cdot f: x \mapsto \alpha \cdot f(x)$ jsou také spojitá a lineární a $\mathcal{L}(X, Y)$ s těmito operacemi tvoří lineární prostor. Na $\mathcal{L}(X, Y)$ vždy uvažujeme kanonickou („operátorovou“) normu:

1.133 Věta. *Nechť X, Y jsou normované lineární prostory. Pak funkce*

$$||L|| = \sup\{||L(x)||\colon\ ||x|| \leq 1\}, \qquad L \in \mathcal{L}(X, Y),$$

je norma na $\mathcal{L}(X, Y)$. Je-li prostor Y úplný, je úplný také prostor $\mathcal{L}(X, Y)$.

Důkaz. Podle Věty 1.123 je $||\cdot||$ reálná funkce. Jestliže $||L|| = 0$, pak podle (1.8) je L identicky nulové zobrazení, takže vlastnost (i) z Definice 1.5 platí. Jsou-li L_1, L_2 prvky $\mathcal{L}(X, Y)$ a $\alpha \in \mathbb{R}$, pak

$$||\alpha \cdot L|| = \sup\{||\alpha \cdot L(x)||\colon\ ||x|| \leq 1\} = |\alpha| \cdot \sup\{||L(x)||\colon\ ||x|| \leq 1\} = |\alpha| \cdot ||L||,$$

$$\begin{aligned} ||L_1 + L_2|| &= \sup\{||L_1(x) + L_2(x)||\colon\ ||x|| \leq 1\} \\ &\leq \sup\{||L_1(x)|| + ||L_2(x)||\colon\ ||x|| \leq 1\} \leq ||L_1|| + ||L_2||, \end{aligned}$$

takže jsou splněny i axiomy (ii), (iii).

Důkaz tvrzení o úplnosti nebudeme potřebovat a jen ho *naznačíme*. Předpokládejme, že (L_n) je cauchyovská posloupnost v prostoru $\mathcal{L}(X,Y)$. Zvolme bod $x \in X$. Protože pro $n, m \in \mathbb{N}$ platí $\|L_n(x) - L_m(x)\| \le \|L_n - L_m\| \cdot \|x\|$, vidíme, že posloupnost $(L_n(x))$ je cauchyovská v úplném prostoru Y, a proto má limitu, kterou označíme $L(x)$. Přímočaře lze nyní dokázat, že takto definované zobrazení $L: X \to Y$ patří do $\mathcal{L}(X,Y)$ a $L_n \to L$ v $\mathcal{L}(X,Y)$.

Je-li X reálný (resp. komplexní) normovaný lineární prostor, pak (Banachův) prostor $X^* := \mathcal{L}(X,\mathbb{R})$ (resp. $X^* := \mathcal{L}(X,\mathbb{C})$) se nazývá (topologický) *duální prostor* k prostoru X a jeho prvky se nazývají *spojité lineární funkcionály*, případně spojité lineární formy.

1.134 Poznámka. V lineární algebře se (algebraickým) duálním prostorem k lineárnímu prostoru X rozumí lineární prostor všech lineárních forem na X (a označuje se také někdy symbolem X^*). Algebraický duální prostor v analýze označujeme jinak, například $X^\#$. Pro normovaný lineární prostor X platí $X^* = X^\#$ (rovnost množin), právě když X je konečně dimenzionální.

Uveďme ještě toto snadné, ale důležité tvrzení.

1.135 Tvrzení. *Nechť X, Y Z jsou normované lineární prostory a nechť jsou dány $f \in \mathcal{L}(X,Y)$, $g \in \mathcal{L}(Y,Z)$. Pak*

$$g \circ f \in \mathcal{L}(X,Z) \qquad \text{a} \qquad \|g \circ f\| \le \|g\| \cdot \|f\|.$$

Důkaz. Tvrzení okamžitě plyne z Věty 1.123 a nerovností

$$\|(g \circ f)(x)\| \le \|g\| \cdot \|f(x)\| \le \|g\| \cdot \|f\| \cdot \|x\|, \qquad x \in X.$$

Na závěr tohoto oddílu uvedeme bez důkazu dva ze základních výsledků o spojitých lineárních funkcionálech. Pak ještě připojíme dva příklady konkrétních spojitých lineárních zobrazení.

1.136 Věta. *Nechť X je normovaný lineární prostor a $x \in X$. Pak existuje lineární funkcionál $f \in X^*$ takový, že $\|f\| = 1$ a $f(x) = \|x\|$.*

V mnoha důležitých Banachových prostorech existuje jednoduchý popis všech spojitých lineárních funkcionálů.

Je-li například X unitární prostor a $x \in X$, z Cauchyovy nerovnosti snadno vyplývá, že $f := \langle \cdot, x\rangle$ je spojitý lineární funkcionál na X a $\|f\| = \|x\|$. Podstatně těžší je dokázat následující (Rieszovu) větu „o reprezentaci“.

1.137 Věta. *Nechť X je Hilberův prostor. Pak každý spojitý lineární funkcionál f na X je tvaru $f = \langle \cdot, x\rangle$. Navíc zobrazení $x \mapsto \langle \cdot, x\rangle$ je lineární izometrie prostoru X na X^*.*

1.138 Příklad. Nechť $g \in C[0,1]$. Jestliže pro $x \in C[0,1]$ položíme $F(x) := g \cdot x$

(tj. $F(x)(t) = g(t) \cdot x(t)$), je zřejmě $F: C[0,1] \to C[0,1]$ lineární zobrazení. Protože

$$\|F(x)\| = \sup_{t\in[0,1]} |g(t) \cdot x(t)| \le \|g\| \cdot \|x\|,$$

je F také spojité (takže $F \in \mathcal{L}(C[0,1])$). Podle Věty 1.123 platí $\|F\| \le \|g\|$. Uážíme-li, že pro konstantní funkci $x(t) = 1$ platí $\|F(x)\| = \|g\|$, vidíme, že $\|F\| = \|g\|$.

1.139 Příklad. Nechť $g \in C[0,1]$. Jestliže pro funkci $x \in C[0,1]$ položíme $f(x) := \int_0^1 g(t)\, x(t)\, dt$, je f zřejmě lineární forma na $C[0,1]$. Protože

$$\left| \int_0^1 g(t)\, x(t)\, dt \right| \le \sup_{t\in[0,1]} |g(t)x(t)| \le \|g\| \cdot \|x\|,$$

je f spojitý lineární funkcionál na $C[0,1]$ (tj. $f \in (C[0,1])^*$) a $\|f\| \le \|g\|$. (Lze ukázat, že $\|f\| = \int_0^1 |g(t)|\, dt$.)

Izometricky izomorfní normované lineární prostory mají stejné všechny metrické vlastnosti. Izomorfní normované lineární prostory prostory mají stejné topologické vlastnosti, ale i řadu jiných. Například platí:

1.140 Tvrzení. *Nechť $(X, \|\cdot\|_1)$ a $(Y, \|\cdot\|_2)$ jsou izomorfní normované lineární prostory. Pak X je úplný, právě když je úplný Y.*

Důkaz. Nechť Y je úplný a $f: X \to Y$ je izomorfismus. Pro $x \in X$ položme $\|x\|_3 := \|f(x)\|_2$. Snadno je vidět, že $\|\cdot\|_3$ je norma na X a $f: (X, \|\cdot\|_3) \to (Y, \|\cdot\|_2)$ je lineární izometrie. Z toho vyplývá, že $(X, \|\cdot\|_3)$ je také úplný a identita $I: (X, \|\cdot\|_1) \to (X, \|\cdot\|_3)$ je izomorfismus. Podle Důsledku 1.125 jsou metriky indukované normami $\|\cdot\|_1$ a $\|\cdot\|_3$ lipschitzovsky ekvivalentní, takže (viz Poznámka 1.81) $(X, \|\cdot\|_1)$ je úplný.

Z Poznámky 1.84, Důsledku 1.131 a Tvrzení 1.140 tedy ihned vyplývá:

1.141 Důsledek. *Každý konečně dimenzionální normovaný prostor je úplný.*

1.142 Tvrzení. *Nechť X je n-rozměrný reálný normovaný lineární prostor, (Y, ρ) je metrický prostor a $y \in Y$. Nechť $\varphi_1, \dots, \varphi_n$ je báze duálního prostoru X^*. Pak zobrazení $f: (Y, \rho) \to X$ je spojité v bodě y, právě když všechny funkce*

$$\varphi_i \circ f: (Y, \rho) \to \mathbb{R}, \qquad i = 1, \dots, n,$$

jsou spojité v bodě y.

Důkaz. Zobrazení $\varphi = (\varphi_1, \dots, \varphi_n): X \to \mathbb{R}^n$ je lineární bijekce (to plyne například z [Be; Věta 21.25]; snadno se totiž ověří, že „duální homomorfismus“ k φ je bijekce), takže φ je izomorfismus podle Věty 1.130. Je tedy f spojité v bodě y, právě když $\varphi \circ f$ je spojité v bodě y, což je podle Věty 1.53 splněno, právě když všechny složky $(\varphi \circ f)_i = \varphi_i \circ f: (Y, \rho) \to \mathbb{R}$, $i = 1, \dots, n$, jsou spojité v bodě y.

1.12 Bilineární a multilineární zobrazení

Nechť $(X_1, \|\cdot\|_1), \dots, (X_n, \|\cdot\|_n)$ jsou normované lineární prostory (nad $\mathbb{T}$). Definujeme-li na $X := X_1 \times \cdots \times X_n$ sčítání a násobení číslem $\alpha \in \mathbb{T}$ po složkách, tj.

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n),$$
$$\alpha \cdot (x_1, \dots, x_n) := (\alpha x_1, \dots, \alpha x_n),$$

pak X je zřejmě lineární prostor. Položíme-li

$$\|x\| := \max(\|x_1\|_1, \dots, \|x_n\|_n) \quad \text{pro} \quad x = (x_1, \dots, x_n),$$

je snadné ověřit, že $\|\cdot\|$ je norma na X. Tuto normu nazveme součinovou normou a lineární normovaný prostor $(X, \|\cdot\|)$ *součinem normovaných lineárních prostorů* $(X_1, \|\cdot\|_1), \dots, (X_n, \|\cdot\|_n)$.

1.143 Poznámka.

(i) Zavedená (maximová) součinová norma není „kanonická". Stejně dobře lze použít (součtovou) normu $\|x\|_s := \|x_1\| + \cdots + \|x_n\|$ nebo („eukleidovskou") normu $\|x\|_e := \sqrt{\|x_1\|^2 + \cdots + \|x_n\|^2}$, které jsou s (maximovou) součinovou normou lipschitzovsky ekvivalentní.

(ii) Metrický prostor X (s metrikou indukovanou součinovou normou) je zřejmě součinem metrických prostorů $X_1, \dots, X_n$ (s metrikami indukovanými příslušnými normami).

(iii) Součinová norma (resp. metrika) na $\mathbb{R}^n = \mathbb{R} \times \cdots \times \mathbb{R}$ je ovšem maximová norma (resp. metrika) na $\mathbb{R}^n$.

1.144 Definice. *Nechť $X_1, \dots, X_n, Y$, $(n \geq 2)$, jsou lineární prostory. Zobrazení $F: X_1 \times \cdots \times X_n \to Y$ se nazývá multilineární (nebo n-lineární), jestliže je lineární ve všech proměnných, tj. jestliže všechna parciální zobrazení*

$$F(a_1, \dots, a_{i-1}, \cdot, a_{i+1}, \dots, a_n), \qquad 1 \leq i \leq n, \ a_1 \in X_1, \dots, a_n \in X_n,$$

jsou lineární zobrazení.

Zobrazení, která jsou 2-lineární, se nazývají bilineární zobrazení. V případě $X_1 = X_2 = \cdots = X_n =: X$ říkáme, že F je n-lineární zobrazení na X. Jestliže $Y = \mathbb{R}$ (resp. $Y = \mathbb{C}$), místo o n-lineárním zobrazení hovoříme o n-lineární (multilineární) formě.

Pro multilineární zobrazení platí následující věta analogická Větě 1.123 o lineárních zobrazeních.

1.145 Věta. *Nechť $X_1, X_2, \dots, X_n, Y$ jsou normované lineární prostory, $X := X_1 \times \cdots \times X_n$ a $F: X \to Y$ je multilineární zobrazení. Pak následující tvrzení jsou ekvivalentní:*

(a) *F je spojité v 0.*

(b) *F je spojité zobrazení.*

(c) *F je lipschitzovské na každé omezené množině $A \subset X$.*
(d) *Existuje taková konstanta $C \geq 0$, že $\|F(x_1, \ldots, x_n)\| \leq C \cdot \|x_1\| \cdot \|x_2\| \cdots \|x_n\|$.*
(e) $\sup\{\|F(x)\|\colon \|x\| \leq 1\} < \infty$.

Pokud tyto podmínky platí, pak reálné číslo $\|F\| := \sup\{\|F(x)\|\colon \|x\| \leq 1\}$ se nazývá norma multilineárního zobrazení F. Platí nerovnost

$$\|F(x)\| \leq \|F\| \cdot \|x_1\| \cdot \|x_2\| \cdots \|x_n\|, \qquad x = (x_1, \ldots, x_n) \in X \tag{1.9}$$

a $\|F\|$ je nejmenší číslo, které má vlastnost čísla C z (d).

Důkaz. (Náznak.) Důkazy implikací $(a) \Longrightarrow (e) \Longrightarrow (d)$ jsou zcela analogické důkazům odpovídajících implikací z Věty 1.123; stačí přímočaře využít multilinearitu F.

K důkazu implikace $(d) \Longrightarrow (c)$ předpokládejme, že C je číslo z podmínky (d) a $\|x\| \leq L$ pro každé $x \in A$. Uvažujme body $x = (x_1, \ldots, x_n)$, $h = (h_1, \ldots, h_n)$ z $X_1 \times \cdots \times X_n$ takové, že $x \in A$ a $x + h \in A$. Máme tedy

$$\|x\| \leq L \quad \text{a} \quad \|h\| \leq \|x + h\| + \|-x\| \leq 2L. \tag{1.10}$$

Jestliže mnohonásobně použijeme multilinearitu F při úpravě („roznásobení") výrazu $F(x + h) = F(x_1 + h_1, \ldots, x_n + h_n)$, není těžké nahlédnout, že diference $D := F(x_1 + h_1, \ldots, x_n + h_n) - F(x_1, \ldots, x_n)$ je součtem $2^n - 1$ prvků Y, z nichž každý má normu nejvýše $C(2L)^{n-1}\|h\|$. Například pro $n = 3$ máme

$$\begin{aligned} D = {} & F(x_1, x_2, h_3) + F(x_1, h_2, x_3) + F(x_1, h_2, h_3) \\ & + F(h_1, x_2, x_3) + F(h_1, x_2, h_3) + F(h_1, h_2, x_3) + F(h_1, h_2, h_3), \end{aligned}$$

takže v tomto případě uvedená vlastnost D snadno plyne z vlastnosti C a (1.10). Dostáváme tedy $\|D\| \leq (2^n - 1)C(2L)^{n-1}\|h\|$, takže F je na A lipschitzovské s konstantou $(2^n - 1)C(2L)^{n-1}$.

Implikace $(b) \Longrightarrow (a)$ je zřejmá; implikace $(c) \Longrightarrow (b)$ a tvrzení o minimalitě $\|F\|$ v (1.9) jsou snadná.

1.146 Poznámka.

(a) Platí zobecnění Věty 1.130: Jsou-li $X_1, \ldots, X_n$ a Y normované lineární prostory, $X_1, \ldots, X_n$ jsou konečně dimenzionální a $F\colon X_1 \times \cdots \times X_n \to Y$ je multilineární zobrazení, pak F je spojité.
(b) Je-li $F\colon X_1 \times \cdots \times X_n \to Y$ spojité multilineární zobrazení, z (1.9) ihned vyplývá, že každé parciální zobrazení $F(a_1, \ldots, a_{i-1}, \cdot, a_{i+1}, \ldots, a_n)$ je spojité lineární.

Jsou-li $X_1, \ldots, X_n$ a Y normované lineární prostory, budeme množinu všech spojitých multilineárních zobrazení $F\colon X_1 \times \cdots \times X_n \to Y$ označovat symbolem

$$\mathcal{L}(X_1, \ldots, X_n; Y).$$

V případě $X_1 = \cdots = X_n =: X$ klademe $\mathcal{L}_n(X, Y) := \mathcal{L}(X_1, \ldots, X_n; Y)$.

Pro $F \in \mathcal{L}(X_1, \dots, X_n; Y)$, $G \in \mathcal{L}(X_1, \dots, X_n; Y)$ a $\alpha \in \mathbb{T}$ zobrazení $(F+G)(x) := F(x)+G(x)$ a $(\alpha F)(x) := \alpha F(x)$ patří zřejmě také do prostoru $\mathcal{L}(X_1, \dots, X_n; Y)$. S těmito operacemi je pak $\mathcal{L}(X_1, \dots, X_n; Y)$ lineární prostor. Navíc zřejmě platí

$$\|F+G\| = \sup\{\|F(x)+G(x)\|\colon \|x\| \le 1\} \le \|F\| + \|G\|,$$

$$\|\alpha F\| = \sup\{\|\alpha F(x)\|\colon \|x\| \le 1\} = \alpha\|F\|,$$

takže snadno vidíme, že norma $\|\cdot\|$ je skutečně norma na $\mathcal{L}(X_1, \dots, X_n; Y)$.

Dále tedy chápeme $\mathcal{L}(X_1, \dots, X_n; Y)$ jako normovaný lineární prostor. Pokud Y je úplný, není těžké dokázat, že také $\mathcal{L}(X_1, \dots, X_n; Y)$ je úplný.

V Kapitole 3 budeme potřebovat následující tvrzení.

1.147 Tvrzení. *Nechť $n \ge 2$, $X_1, \dots, X_n$ a Y jsou normované lineární prostory a $F \in \mathcal{L}(X_1, \mathcal{L}(X_2, X_3, \dots, X_n; Y))$. Pak zobrazení $F_*\colon X_1 \times \dots \times X_n \to Y$ dané předpisem*

$$F_*(x_1, \dots, x_n) = F(x_1)\,(x_2, \dots, x_n)$$

patří do prostoru $\mathcal{L}(X_1, \dots, X_n; Y)$.

Navíc zobrazení $\varphi\colon F \mapsto F_*$ je lineární izometrie prostoru $\mathcal{L}(X_1, \mathcal{L}(X_2, \dots, X_n; Y))$ na prostor $\mathcal{L}(X_1, \dots, X_n; Y)$.

Důkaz. (Náznak) Multilinearita F_* je zřejmá. Je snadno vidět, že

$$\begin{aligned}\|F\| = \sup\bigl\{\sup\{\|F(x_1)\,(x_2, \dots, x_n)\|\colon \|(x_2, \dots, x_n)\| \le 1\}\colon \|x_1\| \le 1\bigr\}\\ = \sup\{\|F_*(x_1, x_2, \dots, x_n)\|\colon \|(x_1, \dots, x_n)\| \le 1\},\end{aligned}$$

takže $F_* \in \mathcal{L}(X_1, \dots, X_n; Y)$ a φ zachovává normu. Linearita φ je zřejmá. Ověření toho, že φ je bijekce, je také zcela přímočaré.

Je obvyklé prostory $\mathcal{L}(X_1, \mathcal{L}(X_2, \dots, X_n; Y))$ a $\mathcal{L}(X_1, \dots, X_n; Y)$ pomocí lineární izometrie φ ztotožňovat (ztotožňujeme F a $F_* = \varphi(F)$). Pak můžeme psát

$$\mathcal{L}(X_1, \mathcal{L}(X_2, Y)) = \mathcal{L}(X_1, X_2; Y),$$

$$\mathcal{L}(X_1, \mathcal{L}(X_2, \mathcal{L}(X_3, Y))) = \mathcal{L}(X_1, \mathcal{L}(X_2, X_3; Y)) = \mathcal{L}(X_1, X_2, X_3; Y)$$

a obecně

$$\mathcal{L}(X_1, \mathcal{L}(X_2, \dots, \mathcal{L}(X_n, Y))\dots) = \mathcal{L}(X_1, \dots, X_n; Y).$$

1.148 Definice. *Nechť X, Y jsou lineární prostory a $F\colon X^n \to Y$ je n-lineární zobrazení na X. Jestliže pro libovolné vektory $x_1, \dots, x_n$ a libovolné dva indexy $1 \le i < j \le n$ platí*

$$F(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = F(x_1, \dots, x_j, \dots, x_i, \dots, x_n),$$

řekneme, že F je symetrické n-lineární zobrazení.

Platí-li vždy, že

$$F(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = -F(x_1, \dots, x_j, \dots, x_i, \dots, x_n),$$

řekneme, že F je antisymetrické n-lineární zobrazení.

1.149 Příklad. Nechť X je *reálný* unitární prostor. Potom skalární součin $s\colon (x,y) \mapsto \langle x,y \rangle$ je zřejmě (symetrická) bilineární forma na X. Podle Cauchyovy nerovnosti $|s(x,y)| \le \|x\| \cdot \|y\|$, takže z Věty 1.145 vyplývá, že $s \in \mathcal{L}_2(X,\mathbb{R})$ a $\|s\| \le 1$. Je-li $X \ne \{0\}$, pro $x \ne 0$ platí $|s(x,x)| = \|x\| \cdot \|x\|$, takže $\|s\| = 1$.

1.150 Příklad. Nechť X, Y, Z jsou normované lineární prostory. Pro $x \in X$, $f \in \mathcal{L}(X,Y)$ a $g \in \mathcal{L}(Y,Z)$ položme

$$\varphi_1(x,f) := f(x) \qquad \text{a} \qquad \varphi_2(f,g) := g \circ f.$$

Pomocí nerovností (viz Věta 1.123 a Tvrzení 1.135)

$$\|f(x)\| \le \|f\| \cdot \|x\| \qquad \text{a} \qquad \|g \circ f\| \le \|g\| \cdot \|f\|$$

snadno dostáváme, že φ_1 a φ_2 jsou spojitá bilineární zobrazení:

$$\varphi_1 \in \mathcal{L}(X, \mathcal{L}(X,Y);Y) \quad \text{a} \quad \varphi_2 \in \mathcal{L}(\mathcal{L}(X,Y), \mathcal{L}(Y,Z); \mathcal{L}(X,Z)).$$

1.151 Příklad. Funkce $f(v_1,\ldots,v_n) := \det[v_1,\ldots,v_n]$ je n-lineární forma na $\mathbb{R}^n$. V lineární algebře se dokazuje, že f je jediná n-lineární forma na $\mathbb{R}^n$, která je antisymetrická a pro kterou $f(e_1,\ldots,e_n) = 1$.

2. Diferenciální počet funkcí více proměnných

2.1 Parciální derivace a totální diferenciál reálné funkce

Nechť f je reálná funkce n reálných proměnných, která je definovaná alespoň v bodě $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$. Zvolme pevně $1 \le i \le n$ a zkoumejme, jak „rychle se mění hodnoty funkce f“, měníme-li málo pouze i-tou souřadnici bodu a. Jinými slovy, vyšetřujme „rychlost změny funkce f v bodě a ve směru vektoru e_i“. Přesněji: uvažujme „parciální diferenci“

$$f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_n) = f(a + he_i) - f(a),$$

která odpovídá částečné (parciální) změně bodu a (jen v i-té souřadnici), a ptejme se, jak velká je tato diference ve srovnání s číslem $0 \neq h \in \mathbb{R}$, které je (v absolutní hodnotě) velmi malé. Tím jsme vedeni ke zkoumání limity

$$\lim_{h \to 0} \frac{f(a + he_i) - f(a)}{h} = \lim_{h \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}.$$

Tato limita však není nic jiného, než derivace funkce

$$g(x) := f(a_1, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n) = f(a + (x - a_i)e_i)$$

v bodě a_i. Skutečně,

$$\begin{aligned} g'(a_i) &= \lim_{h \to 0} \frac{g(a_i + h) - g(a_i)}{h} \\ &= \lim_{h \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{h}. \end{aligned}$$

O funkci g často mluvíme jako o (i-té) *parciální funkci* a nejčastěji pro ni používáme zápis $g = f(a_1, \ldots, a_{i-1}, \cdot, a_{i+1}, \ldots, a_n)$ (srov. obr. 2.3).

Mělo by být tedy jasné, proč v následující definici mluvíme o *parciální* derivaci.

OBR. 2.3. Znázornění parciálních funkcí $f(x_0, \cdot)$ a $f(\cdot, y_0)$.

2.1 Definice. *Nechť f je reálná funkce n proměnných, $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ a $1 \le i \le n$. Pak parciální derivaci funkce f podle i-té proměnné v bodě a definujeme jako limitu*

$$\begin{aligned}\frac{\partial f}{\partial x_i}(a) &:= \lim_{h \to 0} \frac{f(a + he_i) - f(a)}{h} = \left(f(a_1, \dots, a_{i-1}, \cdot, a_{i+1}, \dots, a_n)\right)'(a_i) \\ &= \lim_{h \to 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_n)}{h}.\end{aligned}$$

Symbolem $\dfrac{\partial f}{\partial x_i}$ označujeme parciální derivaci funkce f podle i-té proměnné: funkci definovanou předpisem $\dfrac{\partial f}{\partial x_i}\colon x \mapsto \dfrac{\partial f}{\partial x_i}(x)$.

Jak je všeobecným zvykem, není-li výslovně řečen opak, připouštíme pouze *konečné* parciální derivace. Definičním oborem funkce $\dfrac{\partial f}{\partial x_i}$ je tedy množina bodů $x \in \mathbb{R}^n$, pro které $\dfrac{\partial f}{\partial x_i}(x)$ existuje a je vlastní.

Značení parciálních derivací velmi kolísá. Pro zápis $\dfrac{\partial f}{\partial x_i}(a)$ se používají také následující symboly:

$$\frac{\partial f(a)}{\partial x_i},\ D_i f(a),\ f_i'(a),\ f_i(a)\ , f_{x_i}(a), f_{x_i}'(a), \partial_i f(a),$$

které budeme občas používat i my – zvlášť výhodný je krátký zápis $f_i(a)$ v případech, kdy nemůže dojít k nedorozumění. Pro případ $n = 1$ je ovšem $\dfrac{\partial f(a)}{\partial x_1}$ obyčejná derivace, takže dáváme přednost obvyklému zápisu $f'(a)$ nebo $\dfrac{\mathrm{d}f(a)}{\mathrm{d}x}$.

Při praktických výpočtech jsou většinou jednotlivé proměnné označeny písmeny, které mají často svůj geometrický nebo fyzikální význam; tomu pak odpovídá i značení parciálních derivací. Například pro funkci $f(x,y,z)$ derivaci podle druhé proměnné $D_2 f$ často zapisujeme i těmito symboly:

$$\frac{\partial f}{\partial y},\ D_y f,\ f'_y,\ f_y,\ \partial_y f.$$

Pokud parciální derivace $\dfrac{\partial f}{\partial x_i}(a)$ existuje, musí být ovšem i-tá parciální funkce definovaná na nějakém okolí bodu a_i; jinými slovy: musí existovat $\delta > 0$ takové, že funkce f je definovaná na otevřené úsečce $\{a + he_i\colon\ h \in (-\delta,\delta)\}$. Definiční obor D_f však nemusí obsahovat žádné okolí bodu a a existence a hodnota $\dfrac{\partial f}{\partial x_i}(a)$ závisí jen na hodnotách funkce f na (jakkoliv malé) úsečce výše uvedeného typu.

2.2 Příklad. Uvažujme funkci $f(x,y) = \sqrt{-x^2y^2}$. Snadno vidíme, že definiční obor f je sjednocením souřadnicových os, tj. $D_f = \{(x,y)\colon\ x = 0 \text{ nebo } y = 0\}$, a funkce f je na D_f nulová. Z definice ihned vidíme, že definičním oborem parciální derivace $\dfrac{\partial f}{\partial x}$ je x-ová osa a $\dfrac{\partial f}{\partial x}$ je na svém definičním oboru nulová. Všimněme si, že $D_{\frac{\partial f}{\partial x}} \neq D_f$, $(0,0) \in D_{\frac{\partial f}{\partial x}}$, ale neexistuje takové $\delta > 0$, pro které $U_\delta(0,0) \subset D_f$.

Protože parciální derivace je derivace parciální funkce – což je funkce jedné proměnné – dovedeme snadno počítat parciální derivace funkcí, které jsou explicitně zadány „jednoduchou formulí“.

2.3 Příklad. Vyšetřujme parciální derivaci podle první proměnné funkce dvou proměnných $f(x,y) = \sin(xy^3)$. Zvolíme-li $(x_0,y_0) \in \mathbb{R}^2$, je $\dfrac{\partial f}{\partial x}(x_0,y_0)$ derivace parciální funkce $g\colon x \mapsto \sin(xy_0^3)$ v bodě x_0. Protože $g'(x) = \cos(xy_0^3)\,y_0^3$, máme

$$\frac{\partial f}{\partial x}(x_0,y_0) = g'(x_0) = \cos(x_0y_0^3)\,y_0^3.$$

Je tedy $\dfrac{\partial f}{\partial x}$ definována ve všech bodech roviny a platí

$$\frac{\partial f}{\partial x}(x,y) = \cos(xy^3)y^3.$$

V praxi ovšem místo (x_0,y_0) píšeme hned (x,y), takže vzorec pro $\dfrac{\partial f}{\partial x}(x,y)$ dostáváme formálním derivováním výrazu $\sin(xy^3)$, ve kterém x chápeme jako proměnnou a y jako konstantu. Podobně zpaměti snadno počítáme

$$\frac{\partial f}{\partial y}(x,y) = 3\cos(xy^3)\,xy^2.$$

2.4 Poznámka. Protože parciální derivace je (obyčejná) derivace parciální funkce, platí pro počítání s parciálními derivacemi stejná pravidla, jako pro obyčejné derivace. Například každá z rovností

$$\frac{\partial(f+g)}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a) + \frac{\partial g}{\partial x_i}(a),\ \ \frac{\partial(cf)}{\partial x_i}(a) = c\frac{\partial f}{\partial x_i}(a),$$

$$\frac{\partial(fg)}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a)g(a) + \frac{\partial g}{\partial x_i}(a)f(a)$$

platí, má-li její pravá strana smysl; tj. když (konečné) parciální derivace na pravé straně rovnosti existují.

Následující *nepřesné* úvahy mají za cíl motivovat definici totálního diferenciálu. Pojem diferenciálu již čtenář asi zná pro případ funkce jedné proměnné – v tom případě však užitečnost tohoto pojmu není zřejmá, takže se jeho podstata špatně chápe.

Zakladateli diferenciálního počtu byl diferenciál chápán jako „nekonečně malá diference“ což byl pojem s velmi nejasným významem. Nyní se definuje diferenciál jako „*hlavní lineární část diference*“, což je pojem, který je snadné definovat jako lineární formu, která v jistém přesně definovaném smyslu dobře aproximuje diferenci (přírůstek funkce).

Uvažujme funkci dvou proměnných $f(x,y)$ a bod $(a,b) \in \mathbb{R}^2$. Při motivaci parciálních derivací jsme hovořili o parciálních diferencích funkce, které odpovídají *parciální* (částečné) změně; tj. změně jen v jedné proměnné.

Vyšetřujme nyní *totální* diferenci $f(a+h, b+k) - f(a,b)$, která odpovídá malé změně v obou proměnných. Zkoumáme tedy chování funkce

$$\Delta(h,k) \ := f(a+h, b+k) - f(a,b)$$

v blízkosti bodu $(0,0)$. Podívejme se na konkrétní případ $f(x,y) = x^2 y$, $(a,b) = (2,2)$. Po dosazení máme

$$\Delta(h,k) = (2+h)^2(2+k) - 8 = 8h + 4k + 4hk + 2h^2 + h^2 k.$$

Diference Δ sice v tomto případě není lineární forma, zdá se ale, že má „hlavní lineární část“ $8h + 4k$. To není náhoda. Ukážeme, že „hlavní lineární část“ mají diference většiny funkcí, které nás zajímají. To souvisí s tím, že většinou pracujeme s „hladkými funkcemi“. Nepřesný pojem hladkosti lze precizovat různými způsoby; naše geometrická intuice nám však jistě napovídá, že graf hladké funkce dvou proměnných „lokálně nerozeznáme od roviny“. Jak ukážeme dále při výkladu o tečné nadrovině, je tato vlastnost hladkých funkcí (po přirozeném upřesnění) ekvivalentní s existencí diferenciálu (hlavní lineární části diference).

Přesný význam pojmu „hlavní lineární části“ diference (tj. diferenciálu) je dán v následující definici.

2.5 Definice. *Nechť f je funkce n proměnných, $a \in \mathbb{R}^n$ a nechť $L\colon \mathbb{R}^n \to R$ je lineární forma. Řekneme, že L je totální diferenciál funkce f v bodě a (nebo také derivace funkce f v bodě a), jestliže*

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

(Totálnímu diferenciálu budeme často říkat pouze diferenciál.)

Velikost chyby, které se dopustíme, nahradíme-li diferenci $f(a+h) - f(a)$ hodnotou diferenciálu $L(h)$, je tedy mnohem menší než velikost $||h||$ přírůstku nezávisle proměnné, je-li $||h||$ velmi malé číslo.

Diferenciál (derivace) je lineární forma, tj. funkce tvaru

$$L(h_1, \ldots, h_n) = A_1 h_1 + \cdots + A_n h_n, \qquad \text{kde} \quad A_1, \ldots, A_n \in \mathbb{R}.$$

Za chvíli dokážeme (Důsledek 2.11), že pokud totální diferenciál funkce f v bodě a existuje, je určen jednoznačně. Následující definice je tedy korektní.

2.6 Definice. *Existuje-li diferenciál funkce f v bodě a, označujeme jej symbolem $\mathrm{d}f(a)$. Někdy jej také budeme označovat $f'(a)$; při tomto označení mu budeme říkat derivace funkce f v bodě a.*

Diferenciál $\mathrm{d}f(a) = f'(a)$ je tedy lineární forma na $\mathbb{R}^n$. Hodnota této lineární formy v bodě $h \in \mathbb{R}^n$ se v literatuře označuje různými způsoby. Před zápisy $\mathrm{d}f(a)(h)$, $f'(a)(h)$ se dává často přednost přehlednějším kratším zápisům

$$\mathrm{d}_h f(a),\ f'(a)h,\ f'(a;h),$$

které také budeme někdy používat.

2.7 Poznámka.

(i) V klasické literatuře se hovoří o (totálním) diferenciálu, v moderní literatuře převážně o derivaci. Přirozené zobecnění tohoto pojmu pro zobrazení mezi (nekonečně dimenzionálními) normovanými lineárními prostory se nazývá Fréchetova derivace (viz Definice 3.3).

(ii) Říkáme-li lineární formě L derivace, vzniká pro případ funkce z $\mathbb{R}$ do $\mathbb{R}$ jistá kolize s klasickou terminologií, při které je derivace $f'(a)$ funkce f v bodě a *reálné číslo*. Tato kolize se odstraní tím, že „ztotožňujeme" reálné číslo A s lineární formou $x \mapsto Ax$. Protože víme (viz DI, kap. VIII, par. 4), že A je derivací f v bodě a v klasickém smyslu, právě když lineární forma $x \mapsto Ax$ je diferenciálem v klasickém smyslu (tj. derivací v moderním smyslu), toto ztotožnění nevede ke sporu.

(iii) Místo symbolu $\mathrm{d}f(a)$ se často používá také symbol $Df(a)$.

(iv) Existuje-li diferenciál $\mathrm{d}f(a)$, říkáme, že funkce f *je diferencovatelná* v bodě a.

(v) Z definice okamžitě vidíme, že pokud je f diferencovatelná v a, musí být definovaná aspoň na nějakém okolí bodu a.

Je užitečné výslovně zformulovat následující tvrzení, které zcela jasně ukazuje pojetí diferenciálu jako „hlavní lineární části" diference a je jen snadným přepisem definice totálního diferenciálu.

2.8 Tvrzení. *Nechť f je funkce n proměnných, $a \in \mathbb{R}^n$ a nechť $L\colon \mathbb{R}^n \to R$ je lineární forma. Nechť $r(h)$ je funkce určená rovnicí*

$$f(a+h) - f(a) = L(h) + r(h).$$

Pak následující tvrzení jsou ekvivalentní.

(i) *L je totální diferenciál funkce f v bodě a.*

(ii) $r(h) = o(\|h\|),\ h \to 0.$

(iii) *Existuje reálná funkce n proměnných ω spojitá v $0 \in \mathbb{R}^n$, pro kterou platí $\omega(0) = 0$ a $r(h) = \|h\|\omega(h)$.*

Důkaz. Rovnost $f(a+h)-f(a) = L(h)+r(h)$ je ovšem chápána jako rovnost funkcí (proměnné h) a je ekvivalentní rovnosti $r(h) = f(a+h) - f(a) - L(h)$. Je tedy nulovost limity z Definice 2.5 jen jiný zápis vztahu (ii). Platí-li (ii), je nutně $a \in D_f$ (jinak by funkce r měla prázdný definiční obor), takže $r(0) = f(a) - f(a) - L(0) = 0$. Položíme-li tedy $\omega(h) := r(h)/||h||$ pro $h \neq 0$ a $\omega(0) := 0$, vidíme, že platí (iii). Implikace $(iii) \Longrightarrow (ii)$ je zřejmá.

2.9 Poznámka.

(i) Výrok (ii) z Tvrzení 2.8 se často zapisuje krátce takto:

$$f(a+h) - f(a) = L(h) + o(||h||), \ h \to 0.$$

(ii) Výhodnost podmínky (iii) z 2.8 spočívá v tom, že při jejím užití lze někdy výhodně použít větu o limitě složené funkce s předpokladem spojitosti vnější funkce. Poznamenejme ještě, že funkce ω není (bez požadavku $\omega(0) = 0$) rovnicí $f(a+h) - f(a) = L(h) + ||h||\omega(h)$ určena jednoznačně; vadí bod $h = 0$.

(iii) Věta 1.40 o limitě složeného zobrazení (položíme-li $h = x - a$) snadno dává, že lineární forma L na $\mathbb{R}^n$ je diferenciálem f v bodě a, právě když

$$\lim_{x \to a} \frac{f(x) - f(a) - L(x-a)}{||x-a||} = 0,$$

jinými slovy

$$f(x) - f(a) = L(x-a) + o(||x-a||), \ x \to a.$$

(iv) Uvažujme na $\mathbb{R}^n$ dvě ekvivalentní normy $||\cdot||_1, ||\cdot||_2$ (takže existují čísla $K > 0$, $C > 0$ taková, že $K \le ||h||_1/||h||_2 \le C$ pro $h \neq 0$; viz 1.125). Jestliže $\lim_{h\to 0} r(h)/||h||_1 = 0$, je také

$$\lim_{h \to 0} \frac{r(h)}{||h||_2} = \lim_{h \to 0} \frac{r(h)}{||h||_1} \cdot \frac{||h||_1}{||h||_2} = 0.$$

Podobně $\lim_{h\to 0} r(h)/||h||_1 = 0$, jestliže $\lim_{h\to 0} r(h)/||h||_2 = 0$. Smysl Definice 2.5 by se tedy nezměnil, kdybychom v ní místo eukleidovské normy brali maximovou nebo součtovou normu. (Podle Věty 1.132 bychom dokonce mohli brát libovolnou normu na $\mathbb{R}^n$.)

(v) Jestliže f je afinní funkce na $\mathbb{R}^n$, pak ji lze jednoznačně zapsat ve tvaru $f(x) = c + L(x)$, kde $c \in \mathbb{R}$ a L je lineární forma. Protože zřejmě platí rovnost $f(a+h) - f(a) = L(h)$, podle definice diferenciálu $\mathrm{d}f(a) = L$ pro každý bod $a \in \mathbb{R}^n$. Speciálně pro konstantní funkci f platí $\mathrm{d}f(a) = 0$ pro každý bod $a \in \mathbb{R}^n$.

2.10 Věta. *Nechť L je diferenciál funkce f v bodě $a \in \mathbb{R}^n$. Pak existují parciální derivace $\frac{\partial f}{\partial x_1}(a), \dots \frac{\partial f}{\partial x_n}(a)$ a platí*

$$L(h_1, \dots, h_n) = \frac{\partial f}{\partial x_1}(a)\, h_1 + \dots + \frac{\partial f}{\partial x_n}(a) h_n.$$

Důkaz. Protože L je lineární forma, je dána předpisem

$$L(h_1, \dots, h_n) = A_1 h_1 + \dots + A_n h_n,$$

kde $A_1 = L(e_1), \dots, A_n = L(e_n)$. Podle Tvrzení 2.8(ii) můžeme psát rovnost $f(a+h) - f(a) = L(h) + r(h)$, přičemž $r(h) = o(||h||), \ h \to 0$. Máme tedy

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t\to 0} \frac{f(a + te_i) - f(a)}{t} = \lim_{t \to 0} \frac{L(te_i)}{t} + \lim_{t\to 0} \frac{r(te_i)}{t}$$
$$= L(e_i) + \lim_{t \to 0} \frac{r(te_i)}{||te_i||} \operatorname{sgn}(t) = A_i.$$

Při poslední úpravě jsme mohli použít Větu 1.40 o limitě složeného zobrazení, protože pro zobrazení $h(t) = te_i$ platí $\lim_{t\to 0} h(t) = 0$ a $h(t) \neq 0$ pro $t \neq 0$.

2.11 Důsledek. *Existuje-li diferenciál funkce f v bodě a, je určen jednoznačně.*

2.12 Poznámka. Nechť f je funkce n proměnných. Diferenciálem $\mathrm{d}f$ pak rozumíme zobrazení z $\mathbb{R}^n$ do $(\mathbb{R}^n)^*$, které bodům $x \in \mathbb{R}^n$, ve kterých f má totální diferenciál, přiřazuje lineární formu $\mathrm{d}f(x)$. (Poznamenejme, že ve starší literatuře se symbol $\mathrm{d}f$ chápal jako funkce z $\mathbb{R}^n \times \mathbb{R}^n$ do $\mathbb{R}$ definovaná předpisem $\mathrm{d}f(x,h) := \mathrm{d}f(x)(h)$ pro ty body (x,h), pro které má pravá strana rovnosti smysl.)

Existence všech parciálních derivací funkce f v bodě $a \in \mathbb{R}^n$ je tedy nutnou podmínkou pro existenci diferenciálu $\mathrm{d}f(a)$. Může se však stát, že f má v bodě a všechny parciální derivace, a přesto nemá totální diferenciál.

2.13 Příklad. Pro funkci $f(x,y) = \sqrt{-x^2y^2}$ z Příkladu 2.2 existují obě parciální derivace $\frac{\partial f}{\partial x}(0,0)$, $\frac{\partial f}{\partial y}(0,0)$, ale f není definovaná na žádném okolí bodu $(0,0)$, takže nemůže mít v bodě $(0,0)$ totální diferenciál. Kdybychom funkci f dodefinovali ve všech bodech z $\mathbb{R}^2 \setminus D_f$ hodnotou 1, dostali bychom funkci $\tilde{f}$ definovanou v celé rovině, která má v bodě $(0,0)$ parciální derivace, ale zřejmě není v tomto bodě spojitá. Následující snadná věta ukazuje, že $\tilde{f}$ nemá v bodě $(0,0)$ diferenciál.

2.14 Věta. *Má-li funkce f v bodě $a \in \mathbb{R}^n$ totální diferenciál, je f v bodě a spojitá.*

Důkaz. Nechť $L = \mathrm{d}f(a)$. Podle Tvrzení 2.8 (iii) můžeme psát

$$f(a+h) = f(a) + L(h) + \omega(h)\|h\|,$$

přičemž $\lim_{h\to 0} \omega(h) = 0$. Protože L je spojitá funkce a $L(0) = 0$, dostáváme $\lim_{h\to 0} f(a+h) = f(a)$, takže f je spojitá v bodě a.

Existence diferenciálu hovoří o možnosti „lokální aproximace" *přírůstku (diference)* funkce *lineární formou.* Nyní ukážeme, že možnost „lokální aproximace" *funkce* pomocí *afinní funkce* je pouze snadnou „přeformulací" existence diferenciálu.

Poznamenejme, že afinní funkcí na $\mathbb{R}^n$ rozumíme funkci A, která je tvaru $A(x) = L(x) + c$, kde L je lineární forma na $\mathbb{R}^n$ a $c \in \mathbb{R}$. (Takové funkce se pro $n = 1$ a někdy i pro $n > 1$ nazývají „lineární funkce".) Je snadno vidět, že L, c jsou afinní funkcí A určeny jednoznačně a pro libovolný bod $a \in \mathbb{R}^n$ platí $A(x) = A(a) + L(x-a)$.

2.15 Tvrzení. *Nechť $a \in \mathbb{R}^n$ a f je funkce n proměnných. Pak následující tvrzení jsou ekvivalentní.*

(i) *Existuje* $\mathrm{d}f(a)$.
(ii) *Existuje afinní funkce A na $\mathbb{R}^n$ taková, že*

$$A(a) = f(a) \quad \text{a} \quad \lim_{x\to a} \frac{f(x) - A(x)}{\|x-a\|} = 0. \tag{2.1}$$

OBR. 2.4. H je tečná nadrovina ke G_f v bodě $c = (a, f(a)$; grafem diferenciálu je nadrovina H^*

Pokud tato tvrzení platí, pak podmínka (2.1) *je splněna pouze pro afinní funkci* $A(x) := f(a) + \mathrm{d}f(a)(x-a)$.

Důkaz. Nechť platí (i); položme $A(x) := f(a) + \mathrm{d}f(a)(x-a)$. Pak zřejmě $A(a) = f(a)$ a druhá část podmínky (2.1) platí podle Poznámky 2.9 (iii).

Dále předpokládejme, že A je afinní funkce splňující (2.1). Nechť L je lineární forma, pro kterou $A(x) = A(a) + L(x-a)$; protože $A(a) = f(a)$, platí $A(x) = f(a) + L(x-a)$. Protože platí (2.1), podle Poznámky 2.9 (iii) dostáváme $L = \mathrm{d}f(a)$.

Nechť f je funkce n proměnných, která je definovaná na nějakém okolí bodu $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. Uvažujme její graf

$$G_f = \{(x_1, \dots, x_n, y) \in \mathbb{R}^{n+1}\colon\ y = f(x_1, \dots, x_n)\}.$$

Položme si otázku, zda existuje nadrovina v $\mathbb{R}^{n+1}$ (tj. afinní podprostor dimenze n), která prochází bodem $c := (a_1, \dots, a_n, f(a))$ a blízko bodu c se „velmi těsně přimyká“ ke grafu G_f. V případě, že $n = 1$ a existuje $f'(a) \in \mathbb{R}$, taková „tečná nadrovina“ (v tomto případě přímka) existuje; je to tečna ke G_f v bodě c (s rovnicí $y = f(a) + f'(a)(x-a)$). Naše geometrická intuice nám říká, že pokud $n = 2$ a graf G_f je „hladká plocha“, pak taková „tečná nadrovina“ existuje – stojíme-li na hladké ploše, v naší těsné blízkosti se nám tato plocha jeví jako rovina.

Předpokládejme nyní, že $n \in \mathbb{N}$ a existuje totální diferenciál $\mathrm{d}f(a)$. Podle Tvrzení 2.15 pro afinní funkci $A(x) = f(a) + \mathrm{d}f(a)(x-a)$ platí

$$f(x) - A(x) = o(\|x-a\|),\ x \to a, \tag{2.2}$$

takže afinní funkce A „velmi dobře aproximuje“ funkci f v blízkosti bodu a (srov. obr. 2.4).

Graf H afinní funkce A se tedy v jistém smyslu velmi těsně přimyká ke grafu funkce f v blízkosti bodu $(a, f(a))$. Tím jsme motivovali následující definici. (Pojem tečné nadroviny více objasňuje Věta 2.17.)

Obr. 2.5. K důkazu o geometrickém významu tečné nadroviny.

2.16 Definice. *Nechť funkce f má diferenciál v bodě $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. Pak definujeme tečnou nadrovinu ke grafu funkce f v bodě $(a_1, \dots, a_n, f(a)) \in \mathbb{R}^{n+1}$ jako graf afinní funkce $A(x) = f(a) + \mathrm{d}f(a)(x - a)$.*

Z Věty 2.10 vyplývá, že tato tečná nadrovina má rovnici

$$x_{n+1} = f(a) + \frac{\partial f}{\partial x_1}(a)(x_1 - a_1) + \cdots + \frac{\partial f}{\partial x_n}(a)(x_n - a_n).$$

2.17 Věta. (geometrický smysl tečné nadroviny) *Nechť f je funkce n proměnných, která je definována aleaspoň na nějakém okolí bodu $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. Nechť $G_f \subset \mathbb{R}^{n+1}$ je graf funkce f, $c := (a, f(a)) \in \mathbb{R}^{n+1}$ a nechť $H \subset \mathbb{R}^{n+1}$ je nadrovina (tj. afinní podprostor dimenze n), která prochází bodem c a je grafem afinní funkce n proměnných. Pak následující tvrzení jsou ekvivalentní.*

(i) *H je tečná nadrovina ke G_f v bodě c.*

(ii) $$\lim_{z \to c, z \in G_f} \frac{\operatorname{dist}(z, H)}{\|z - c\|} = 0.$$

Důkaz. (Náznak.) Všechny níže použité normy a vzdálenosti (v $\mathbb{R}^n$ i $\mathbb{R}^{n+1}$) jsou eukleidovské. Nechť H je grafem afinní funkce $A\colon \mathbb{R}^n \to \mathbb{R}$. Předpokládejme nejprve, že platí (i). Podle Definice 2.16 a Tvrzení 2.15 dostáváme

$$\lim_{x \to a} \frac{|f(x) - A(x)|}{\|x - a\|} = 0. \tag{2.3}$$

Je-li $z = (x, f(x)) \in G_f$, pak vždy položíme $z^* := (x, A(x))$. Zřejmě $\|x - a\| \le \|z - c\|$ a $\operatorname{dist}(z, H) \le \|z - z^*\| = |f(x) - A(x)|$. Dostáváme tedy nerovnost

$$\frac{\operatorname{dist}(z, H)}{\|z - c\|} \le \frac{|f(x) - A(x)|}{\|x - a\|},$$

která spolu s nerovností $\|x - a\| \le \|z - c\|$ a (2.3) snadno dává (ii).

Důkaz implikace (ii) $\Longrightarrow$ (i) je založen na tom, že existuje číslo $\kappa > 0$ nezávislé na x, pro které

$$\operatorname{dist}(z, H) = \kappa\,|f(x) - A(x)|. \tag{2.4}$$

Platí $\kappa = |\cos\alpha|$, kde α je úhel sevřený vektorem e_{n+1} a normálovým vektorem k nadrovině H. To „je vidět“ z obr. 2.5, nebudeme to však formálně dokazovat.

Z (ii) a (2.4) snadno vyplývá, že existuje $\delta > 0$ takové, že pokud $z \in G_f$ a platí $0 < ||z-c|| < \delta$, pak $||z - z^*|| < (1/2)||z - c||$ a tedy

$$||z-c|| \leq ||z^* - c|| + ||z - z^*|| \leq ||z^* - c|| + \frac{1}{2}||z - c||,$$

takže $||z - c|| \leq 2||z^* - c||$.

Protože afinní funkce A je lipschitzovská s nějakou konstantou lipschitzovskosti $K > 0$, pro taková z dostáváme

$$\begin{aligned} ||z^* - c|| = ||(x, A(x)) - (a, A(a))|| &= \sqrt{||x-a||^2 + ||A(x) - A(a)||^2} \\ &\leq \sqrt{||x-a||^2 + K^2||x-a||^2} \leq ||x - a||\sqrt{1+K^2}, \end{aligned}$$

takže $||z-c|| \leq 2\sqrt{1+K^2}\,||x-a||$. Z této nerovnosti, (2.4) a (ii) okamžitě vyplývá platnost (2.3). Podle Tvrzení 2.15 a Definice 2.16 je H tečná nadrovina ke G_f v bodě c.

2.18 Poznámka. Nechť $n = 1$, $f(x) = \sqrt[3]{x}$, $a = 0$ a nadrovina H je osa y. Je snadno vidět, že pak (ii) platí, ale (i) neplatí. V předchozí větě tedy nelze vynechat předpoklad, že H je grafem afinní funkce.

V dalším budeme potřebovat následující tvrzení, které lze chápat jako jisté zobecnění Lagrangeovy věty o přírůstku funkce (první věty o střední hodnotě). Větou o přírůstku funkce (střední hodnotě) pro funkce více proměnných se ale nazývá jiná, důležitější věta (Věta 2.60), ve které za silnějších předpokladů odvodíme silnější tvrzení.

2.19 Věta. *Nechť f je reálná funkce n proměnných, která má všechny parciální derivace v každém bodě otevřeného intervalu $I \subset \mathbb{R}^n$, a nechť jsou dány body $a = (a_1, \ldots, a_n) \in I, b = (b_1, \ldots, b_n) \in I$. Pak v intervalu I existují body $\xi^1, \ldots, \xi^n$, pro které platí*

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

Důkaz. 1) Nejprve uvažujme případ $n = 1$. Pokud $a = b$, lze zvolit $\xi^1 \in I$ libovolně. Pokud $b > a$, existuje hledaný bod $\xi^1 \in (a, b) \subset I$ podle Lagrangeovy věty o přírůstku funkce. Pokud $b < a$, táž věta dává existenci bodu $\xi^1 \in (b, a) \subset I$, pro který $f(a) - f(b) = (a-b)f'(\xi^1)$, takže $f(b) - f(a) = (b-a)f'(\xi^1)$.

2) Nyní provedeme (geometricky zcela jasný) důkaz pro $n = 2$. Nechť $I = I_1 \times I_2$. Uvažujme pomocný bod $p := (b_1, a_2)$ (viz obr. 2.6).

Pak platí $f(b) - f(a) = (f(p) - f(a)) + (f(b) - f(p))$. Máme

$$f(p) - f(a) = f(b_1, a_2) - f(a_1, a_2) = f(\cdot, a_2)(b_1) - f(\cdot, a_2)(a_1).$$

Obr. 2.6.

Podle kroku 1) existuje $\eta_1 \in I_1$ takové, že

$$f(p) - f(a) = (f(\cdot, a_2))'(\eta_1)(b_1 - a_1) = \frac{\partial f}{\partial x_1}(\eta_1, a_2)(b_1 - a_1).$$

Zcela obdobně dostáváme, že existuje číslo $\eta_2 \in I_2$ takové, že platí $f(b) - f(p) = \dfrac{\partial f}{\partial x_2}(b_1, \eta_2)(b_2 - a_2)$. Můžeme tedy položit $\xi^1 := (\eta_1, a_2)$ a $\xi^2 := (b_1, \eta_2)$.

3) V obecném případě uvažujme pro $k = 0, 1, \ldots, n$ body

$$p_k = a + \sum_{i=1}^{k} (b_i - a_i) e_i = (b_1, \ldots, b_k, a_{k+1}, \ldots, a_n).$$

Pak $p_0 = a,\ p_n = b$ a $f(b) - f(a) = \sum_{k=1}^{n} (f(p_k) - f(p_{k-1}))$. Zcela obdobně jako v kroku 2) dostáváme existenci bodů $\xi^k \in I, k = 1, \ldots, n$ takových, že

$$f(p_k) - f(p_{k-1}) = f(p_{k-1} + (b_k - a_k)e_k) - f(p_{k-1}) = \frac{\partial f}{\partial x_k}(\xi^k)(b_k - a_k),$$

což dokazuje tvrzení.

2.20 Věta. *Nechť f je funkce n proměnných, $a \in \mathbb{R}^n$ a všechny parciální derivace $\dfrac{\partial f}{\partial x_1}, \ldots, \dfrac{\partial f}{\partial x_n}$ jsou funkce spojité v bodě a. Pak f má v bodě a totální diferenciál.*

Důkaz. Podle Věty 2.10 může být totálním diferenciálem $\mathrm{d}f(a)$ jedině funkce

$$L(h_1, \ldots, h_n) = \frac{\partial f}{\partial x_1}(a)\, h_1 + \cdots + \frac{\partial f}{\partial x_n}(a) h_n.$$

Zda L je diferenciálem, budeme zjišťovat z definice, ve které tentokrát použijeme součtovou normu $\|h\|_1 = |h_1| + \cdots + |h_n|$ (což podle Poznámky 2.9 (iv) můžeme udělat). Položíme-li tedy

$$r(h) = f(a + h) - f(a) - \left(\frac{\partial f}{\partial x_1}(a)\, h_1 + \cdots + \frac{\partial f}{\partial x_n}(a) h_n \right),$$

chceme dokázat, že $\lim\limits_{h\to 0} \dfrac{r(h)}{\|h\|_1} = 0$; to provedeme z definice limity. Zvolme libovolné $\varepsilon > 0$. Protože f má spojité parciální derivace v bodě a, můžeme najít $\delta > 0$ takové, že pro každé $1 \le i \le n$ a $x \in I := (a_1 - \delta, a_1 + \delta) \times \cdots \times (a_n - \delta, a_n + \delta)$ platí $\left|\dfrac{\partial f}{\partial x_i}(x) - \dfrac{\partial f}{\partial x_i}(a)\right| < \varepsilon$. Je-li nyní h takový bod, že $0 < \|h\|_\infty < \delta$, je zřejmě $b := a + h \in I$, takže podle Věty 2.19 existují body $\xi^1, \dots, \xi^n \in I$, pro které platí

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

Platí tedy

$$|r(h)| = \left|\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\xi^i) h_i - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a) h_i\right| \le \sum_{i=1}^{n} |h_i| \left|\frac{\partial f}{\partial x_i}(\xi^i) - \frac{\partial f}{\partial x_i}(a)\right| < \varepsilon \|h\|_1,$$

takže máme $|r(h)/\|h\|_1| < \varepsilon$ a důkaz je dokončen.

2.21 Poznámka. Existuje-li $\mathrm{d}f(a)$, víme, že platí

$$\mathrm{d}f(a)(h_1, \dots, h_n) = \frac{\partial f}{\partial x_1}(a)\, h_1 + \cdots + \frac{\partial f}{\partial x_n}(a) h_n.$$

Často se používá zápis

$$\mathrm{d}f(a) = \frac{\partial f}{\partial x_1}(a)\, \mathrm{d}x_1 + \cdots + \frac{\partial f}{\partial x_n}(a) \mathrm{d}x_n. \tag{2.5}$$

Původně se „diferenciály“ $\mathrm{d}x_1, \dots, \mathrm{d}x_n$ chápaly jako „nekonečně malé diference nezávisle proměnných“ s nejasným významem. Vzorec má ale smysl i v současném přesném pojetí (kdy diferenciál chápeme jako lineární formu), domluvíme-li se, že symbolem $\mathrm{d}x_i$ rozumíme lineární formu $L_i\colon (h_1, \dots, h_n) \mapsto h_i$. Vzorec (2.5) pak hovoří o rovnosti dvou lineárních forem n proměnných. Má-li f v každém bodě svého definičního oboru diferenciál, lze při této úmluvě psát také (srov. Poznámka 2.12)

$$\mathrm{d}f = \frac{\partial f}{\partial x_1}\, \mathrm{d}x_1 + \cdots + \frac{\partial f}{\partial x_n} \mathrm{d}x_n; \tag{2.6}$$

pak jde o rovnost dvou zobrazení, která každému bodu $x \in D_f$ přiřazují lineární formu $\mathbb{R}^n \to \mathbb{R}$ (rovnou $\mathrm{d}f(x)$). (Obecně může mít ale zobrazení na pravé straně větší definiční obor!)

Protože však symbol $\mathrm{d}f$ chápeme jako zobrazení $x \mapsto \mathrm{d}f(x)$ a x_i jako souřadnicovou funkci $(x_1, \dots, x_n) \mapsto x_i$ (tedy $x_i = L_i$), je přirozenější symbol $\mathrm{d}x_i$ chápat jako konstantní zobrazení

$$\mathrm{d}x_i\colon \mathbb{R}^n \to (\mathbb{R}^n)^*, \qquad x \mapsto \mathrm{d}x_i(x) = L_i.$$

Při tomto *obvyklém chápání* symbolu $\mathrm{d}x_i$ ovšem vzorec (2.5) *neplatí*, vzorec (2.6) však stále platí.

Je-li funkce f označena písmenem y, dostáváme klasický vzorec

$$dy = \frac{\partial y}{\partial x_1}\, \mathrm{d}x_1 + \cdots + \frac{\partial y}{\partial x_n} \mathrm{d}x_n.$$

Přibližný výpočet hodnoty diference funkce pomocí hodnoty diferenciálu (která se počítá mnohem pohodlněji) se často s úspěchem používá v aplikacích. Tam totiž v roli diferencí nezávisle proměnných často vystupují chyby v měření, které bývají velmi malé; viz následující příklad.

2.22 Příklad. Chceme změřit vzdálenost a bodů B, C v rovině, mezi kterými není přímá viditelnost; ta je však mezi body A, B i mezi body A, C. Známe-li vzdálenosti $b = ||C - A||$, $c = ||B - A||$ a úhel α sevřený vektory $C - A$ a $B - A$, jsme schopni spočítat vzdálenost a podle známého vzorce

$$a = a(b, c, \alpha) = \sqrt{b^2 + c^2 - 2bc\cos\alpha}.$$

Uvažujme konkrétní případ, kdy naměřené hodnoty jsou (délka je v kilometrech a úhel v radiánech) $b_n = 5,\ c_n = 8,\ \alpha_n = \pi/3,$ takže nám vychází hodnota
$a_n = \sqrt{25 + 64 - 80 \cdot (1/2)} = 7$. Předpokládejme, že víme, že chyby, kterých jsme se při měření dopustili, nepřesahují 1% naměřených hodnot, a ptejme se, jaké maximální chyby jsme se mohli dopustit při výpočtu vzdálenosti a. Jinými slovy: chceme odhadnout velikost chyby $|\Delta a| = |a - a_n|$, víme-li, že $|\Delta b| = |b - b_n| \le 5 \cdot 10^{-2}$,

$$|\Delta c| = |c - c_n| \le 8 \cdot 10^{-2},\ |\Delta\alpha| = |\alpha - \alpha_n| \le (\pi/3) \cdot 10^{-2}.$$

Po výpočtu parciálních derivací funkce $a(b, c, \alpha)$ z Věty 2.20 a Věty 2.10 snadno dostáváme, že

$$\delta a := da(b_n, c_n, \alpha_n)(\Delta b, \Delta c, \Delta\alpha) = \frac{1}{7}\Delta b + \frac{11}{14}\Delta c + \frac{20}{7}\sqrt{3}\Delta\alpha, \tag{2.7}$$

takže

$$|\delta a| \le \frac{1}{7} \cdot 5 \cdot 10^{-2} + \frac{11}{14} \cdot 8 \cdot 10^{-2} + \frac{20}{7} \cdot \sqrt{3} \cdot \frac{\pi}{3} \cdot 10^{-2} \approx 12 \cdot 10^{-2},$$

přičemž lepší horní odhad zřejmě nelze dostat. Protože čísla Δb, Δc a $\Delta\alpha$ se zdají být „dostatečně malá“, je pravděpodobná domněnka, že pokud nahradíme diferenci Δa diferenciálem δa, dopustíme se „zanedbatelné chyby“. Ze vzorce (2.7) vidíme, že chyba Δc má na chybu Δa (překvapivě) více než pětkrát větší vliv než chyba Δb, což lze jistě v praxi využít.

Parciální derivaci $\frac{\partial f}{\partial x_i}(a)$ lze interpretovat jako „rychlost změny funkce f v bodě a ve směru vektoru e_i“.

2.23 Poznámka. Směrem zde rozumíme „orientovaný směr“; dva vektory w, v mají (určují) stejný směr, pokud $w = \lambda v$, $\lambda > 0$. Formálně se (orientovaný) směr určený vektorem v definuje jako množina $\{\lambda v\colon\ \lambda > 0\}$. Každý jednotkový vektor tedy určuje právě jeden směr a naopak.

Je ovšem přirozené zkoumat „rychlost změny“ i v jiných směrech. Pak místo e_i bereme obecný jednotkový vektor v a zkoumáme velikost změny $f(a + tv) - f(a)$ v závislosti na t. Jsme tedy vedeni k následující definici, ve které ale *nepředpokládáme*, že v je jednotkový vektor.

2.24 Definice. (derivace podle vektoru) *Nechť f je funkce n proměnných, $a \in \mathbb{R}^n$ a $v \in \mathbb{R}^n$. Pak derivací funkce f v bodě a podle vektoru v rozumíme (vlastní) limitu*

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

Někdy se o derivaci podle vektoru hovoří jako o „směrové derivaci“. *Derivací ve směru* však rozumíme derivaci podle jednotkového vektoru daného směru. Tedy derivace f v bodě a ve směru vektoru $w \neq 0$ je rovna derivaci $D_v f(a)$ podle jednotkového vektoru $v := w/||w||$. Terminologie kolísá.

2.25 Poznámka.

(i) Pokud $D_v f(a)$ existuje, musí být ovšem nejen $a \in D_f$, ale D_f musí obsahovat otevřenou úsečku $\{a + tv\colon\ t \in (-\delta, \delta)\}$ pro některé $\delta > 0$.

(ii) Zřejmě $D_{e_i} f(a) = \dfrac{\partial f}{\partial x_i}(a)$.

(iii) Podobně jako parciální derivace je derivací parciální funkce, platí zřejmě $D_v f(a) = g'(0)$, kde $g(t) := f(a + tv)$.

(iv) Místo $D_v f(a)$ se také často píše $\dfrac{\partial f}{\partial v}(a)$ nebo $\partial_v f(a)$.

(v) Je-li $\lambda \neq 0$, pak $D_{\lambda v} f(a) = \lambda D_v f(a)$, má-li jedna strana rovnosti smysl. To je snadno vidět po úpravě

$$D_{\lambda v} f(a) = \lim_{t\to 0} \frac{f(a + t\lambda v) - f(a)}{t} = \lim_{t\to 0} \lambda \frac{f(a + t\lambda v) - f(a)}{t\lambda}.$$

Hledejme nyní „směr největšího růstu“ funkce f v bodě a; tj. zkoumejme, pro který jednotkový vektor v je $D_v f(a)$ největší. Začněme s definicí gradientu.

2.26 Definice. *Nechť f je funkce n proměnných, $a \in \mathbb{R}^n$ a nechť f má totální diferenciál v bodě a. Pak definujeme gradient funkce f v bodě a jako vektor*

$$\operatorname{grad} f(a) := \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right) \in \mathbb{R}^n.$$

2.27 Poznámka.

(i) Místo $\operatorname{grad} f(a)$ se často píše $\nabla f(a)$. Toto značení se používá zejména ve fyzice; symbol ∇ se čte „nabla“.

(ii) Podle naší (v moderní literatuře běžné) definice k existenci $\operatorname{grad} f(a)$ nestačí existence parciálních derivací f v a; musí existovat dokonce diferenciál funkce f v a.

(iii) Z Věty 2.10 okamžitě vyplývá, že $\mathrm{d}f(a)(h) = \langle \operatorname{grad} f(a), h \rangle$, má-li jedna strana rovnosti smysl.

2.28 Věta. *Nechť f je funkce n proměnných, $a \in \mathbb{R}^n$ a $v \in \mathbb{R}^n$. Nechť existuje $\mathrm{d}f(a)$. Pak platí:*

(i) $D_v f(a) = \mathrm{d}f(a)(v) = \langle \operatorname{grad} f(a), v \rangle$.

(ii) $\max\{D_v f(a)\colon\ \|v\| = 1\} = \|\operatorname{grad} f(a)\|$.

Pokud $\operatorname{grad} f(a) \neq 0$, *pak se maxima nabývá jen pro vektor* $v = \frac{\operatorname{grad} f(a)}{\|\operatorname{grad} f(a)\|}$.

Důkaz. Označme $L := \mathrm{d}f(a)$. Podle Tvrzení 2.8 (iii) můžeme psát

$$f(a + h) - f(a) = L(h) + \omega(h)\|h\|,$$

přičemž $\lim_{t\to 0} \omega(t) = \omega(0) = 0$. Je tedy

$$D_v f(a) = \lim_{t\to 0} \frac{f(a + tv) - f(a)}{t} = \lim_{t\to 0} \left(\frac{L(tv)}{t} + \omega(tv) \frac{\|tv\|}{t} \right) = L(v).$$

Je-li $\operatorname{grad} f(a) = 0$, je platnost (ii) zřejmá. Pokud $\operatorname{grad} f(a) \neq 0$, užijeme Cauchyovu nerovnost. Ta nám dává, že kdykoliv $\|v\| = 1$, pak

$$|\langle \operatorname{grad} f(a), v\rangle| \leq \|\operatorname{grad} f(a)\| \cdot \|v\| = \|\operatorname{grad} f(a)\|,$$

přičemž rovnost platí, právě když jsou vektory $v, \operatorname{grad} f(a)$ lineárně závislé, tj. platí

$$v = \frac{\operatorname{grad} f(a)}{\|\operatorname{grad} f(a)\|} \quad \text{nebo} \quad v = -\frac{\operatorname{grad} f(a)}{\|\operatorname{grad} f(a)\|}. \tag{2.8}$$

Platí tedy vždy $\langle \operatorname{grad} f(a), v\rangle \leq \|\operatorname{grad} f(a)\|$ a rovnost může nastat jen tehdy, když platí (2.8). Po dosazení ihned vidíme, že vyhovuje pouze případ $v = \frac{\operatorname{grad} f(a)}{\|\operatorname{grad} f(a)\|}$.

2.29 Poznámka.

(i) Ukázali jsme, že $D_v f(a) = \mathrm{d}_v f(a)$, má-li pravá strana rovnosti smysl. Může se ale stát, že $D_v f(a)$ existuje, ale $\mathrm{d}_v f(a)$ nikoliv (srov. Příklad 2.2 pro $v = e_1$).

(ii) Z Věty 2.28 okamžitě vyplývá, že $\|\mathrm{d}f(a)\| = \|\operatorname{grad} f(a)\|$, kde nalevo je norma $\mathrm{d}f(a)$ jako lineárního zobrazení (srov. Věta 1.123) a napravo je eukleidovská norma vektoru $\operatorname{grad} f(a)$.

(iii) Známe-li definici a *vlastnosti* úhlu sevřeného dvěma nenulovými vektory v $\mathbb{R}^n$, můžeme (ii) snadno a přirozeně dokázat ze vzorce

$$D_v f(a) = \langle \operatorname{grad} f(a), v\rangle = \|\operatorname{grad} f(a)\| \cdot \|v\| \cos\alpha = \|\operatorname{grad} f(a)\| \cos\alpha,$$

kde α je úhel sevřený vektory $\operatorname{grad} f(a)$ a v.

Ukázali jsme, že (pokud $\operatorname{grad} f(a) \neq 0$), funkce f v bodě a nejrychleji roste ve směru gradientu $\operatorname{grad} f(a)$; jeho směr je tedy *směrem největšího růstu.* Přitom rychlost růstu v tomto směru je dána normou gradientu. V opačném směru funkce nejrychleji klesá. Ve směrech kolmých na gradient se funkce mění „s nulovou rychlostí“. Je-li $\operatorname{grad} f(a) = 0$, je ovšem $D_v f(a) = 0$ pro každé v. Proto je přirozené zavést následující terminologii.

2.30 Definice. *Nechť f je funkce n proměnných a $a \in \mathbb{R}^n$. Řekneme, že a je stacionárním bodem funkce f, jestliže* $\operatorname{grad} f(a) = 0$.

Pro stacionární bod se někdy užívá název *kritický bod.*

2.31 Poznámka. Následující tvrzení jsou zřejmě ekvivalentní:

(i) Bod a je stacionárním bodem funkce f.
(ii) Diferenciál $\mathrm{d}f(a)$ je nulová lineární forma.
(iii) Diferenciál $\mathrm{d}f(a)$ existuje a všechny parciální derivace f v bodě a jsou nulové.

2.32 Poznámka. Uvažujme mapu, na které máme zvoleny kartézské souřadnice, a „hladkou“ funkci $f(x,y)$, která udává nadmořskou výšku bodu povrchu hornatého terénu, jehož obraz na mapě má souřadnice (x,y). Uvažujme obvyklé přibližné znázornění funkce f pomocí vrstevnic. Jdeme-li „po vrstevnici“, zřejmě nestoupáme ani neklesáme; příslušná směrová derivace (ve směru tečném k vrstevnici) je tedy nulová. Z toho usuzujeme, že vektor $\operatorname{grad} f(a)$ je kolmý na vrstevnici v bodě a. Přesné matematické zachycení těchto nepřesných názorných představ bude provedeno později (Věta 2.161 (iii)).

Existuje-li $\mathrm{d}f(a)$, pak $D_v f(a) = \mathrm{d}f(a)(v)$, takže funkce $v \mapsto D_v f(a)$ je lineární forma. Může se však stát, že $D_v f(a)$ existuje pro každé $v \in \mathbb{R}^n$, ale přesto není funkce $v \mapsto D_v f(a)$ lineární. Tuto vlastnost má třeba funkce $f(x,y)$, která je na ose x definovaná předpisem $f(x,y) = x$ a jinak je nulová.

Následující příklad ukazuje, že i když funkce $v \mapsto D_v f(a)$ je lineární, nemusí existovat $\mathrm{d}f(a)$. V tomto příkladu bude dokonce $D_v f(a) = 0$ pro každé v, ale bod a nebude stacionární.

2.33 Příklad.

a) Nechť $g(x,y) = 0$, pokud $y \neq x^2$ a $g(x,x^2) = x$ pro každé $x \in \mathbb{R}$. Pak lze snadno dokázat, že $D_v g(0,0) = 0$ pro každý vektor $v \in \mathbb{R}^2$, g je v bodě $(0,0)$ spojitá, ale diferenciál $\mathrm{d}g(0,0)$ neexistuje.

b) Položme $g^*(x,y) := \operatorname{sgn}(g(x,y))$. Opět platí $D_v g^*(0,0) = 0$ pro každý vektor $v \in \mathbb{R}^2$, ale g^* není dokonce v bodě $(0,0)$ spojitá.

c) Položíme-li $h(0,0) = 0$ a

$$h(x,y) = \frac{yx^2\sqrt{x^2+y^2}}{y^2+x^4} \quad \text{pro} \quad (x,y) \neq (0,0),$$

pak h je spojitá v každém bodě roviny, $dh(a)$ existuje pro $a \neq (0,0)$, $dh(0,0)$ neexistuje a $D_v h(0,0) = 0$ pro každý $v \in \mathbb{R}^2$.

2.34 Poznámka. Podle terminologie z funkcionální analýzy (viz Definice 3.7) řekneme, že lineární forma $L\colon \mathbb{R}^n \to \mathbb{R}$ je Gâteauxova derivace funkce f v bodě a, jestliže pro každý vektor $v \in \mathbb{R}^n$ platí $L(v) = D_v f(a)$. Předchozí příklad ukazoval, že z existence („slabé") Gâteauxovy derivace v bodě a nevyplývá existence („silné") Fréchetovy derivace v bodě a a dokonce ani spojitost v tomto bodě.

2.2 Derivace zobrazení mezi eukleidovskými prostory

Pro zobrazení F budeme definovat pojmy parciální derivace, derivace podle vektoru i diferenciálu zcela obdobně jako pro reálné funkce n proměnných. Z formálního (nikoliv však didaktického) hlediska bylo tedy logičtější definovat tyto pojmy jen jednou rovnou pro zobrazení.

2.35 Definice. *Nechť F je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^k$, $i \in \{1,\dots,n\}, a \in \mathbb{R}^n$ a $v \in \mathbb{R}^n$. Pak definujeme*

$$\frac{\partial F}{\partial x_i}(a) := \lim_{h\to 0} \frac{F(a+he_i)-F(a)}{h}, \quad D_v F(a) := \lim_{h\to 0} \frac{F(a+hv)-F(a)}{h}.$$

Zřejmě $\dfrac{\partial F}{\partial x_i}(a) = D_{e_i}F(a)$. Je-li $F = (F_1, \dots, F_k)$, pak

$$\frac{F(a+hv) - F(a)}{h} = \left(\frac{F_1(a+hv) - F_1(a)}{h}, \dots, \frac{F_k(a+hv) - F_k(a)}{h} \right),$$

a protože limita zobrazení do $\mathbb{R}^k$ se „počítá po složkách" (Věta 1.53), ihned vidíme, že

$$D_vF(a) = (D_vF_1(a), \dots, D_vF_k(a)),$$

má-li jedna strana rovnosti smysl. Speciálně

$$\frac{\partial F}{\partial x_i}(a) = \left(\frac{\partial F_1}{\partial x_i}(a), \dots, \frac{\partial F_k}{\partial x_i}(a) \right).$$

Nyní přirozeným způsobem zobecníme pojem derivace (neboli diferenciálu) reálné funkce n proměnných na pojem derivace zobrazení mezi eukleidovskými prostory.

Zcela obdobně se definuje pojem (Fréchetovy) derivace i pro zobrazení mezi (nekonečně dimenzionálními) normovanými lineárními prostory (viz Definice 3.3 níže). Zde jen poznamenejme, že pro případ zobrazení F mezi nekonečně dimenzionálními prostory je třeba v definici *požadovat spojitost* zobrazení $L = F'(a)$, která pro zobrazení mezi konečně dimenzionálními prostory vyplývá z jeho linearity (Věta 1.130).

Jak je zvykem, pro zobrazení budeme používat moderní „derivační terminologii"; místo o diferenciálu hovoříme o derivaci.

2.36 Definice. *Nechť F je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^k$, $a \in \mathbb{R}^n$ a nechť $L: \mathbb{R}^n \to \mathbb{R}^k$ je lineární zobrazení. Řekneme, že L je derivace funkce F v bodě a, jestliže*

$$\lim_{h \to 0} \frac{F(a+h) - F(a) - L(h)}{||h||} = 0. \tag{2.9}$$

Derivaci zobrazení F v bodě a označujeme $F'(a)$.

Označení $F'(a)$ bude ovšem korektní, až (za chvíli a snadno) dokážeme jednoznačnost derivace. Existuje-li $F'(a)$, říkáme, že zobrazení F je *diferencovatelné* v bodě a.

2.37 Poznámka.

(i) V (2.9) symbol 0 má ovšem dva různé významy – limitu provádíme v počátku $0 \in \mathbb{R}^n$ a její hodnota je $0 \in \mathbb{R}^k$. Místo (2.9) lze zřejmě („ekvivalentně") psát

$$\lim_{h \to 0} \frac{||F(a+h) - F(a) - L(h)||}{||h||} = 0,$$

kde nyní hodnota limity je reálné číslo 0, norma v čitateli je eukleidovská norma v $\mathbb{R}^k$ a ve jmenovateli eukleidovská norma v $\mathbb{R}^n$.

(ii) I v nyní zkoumaném případě, kdy F je zobrazení se pro derivaci $F'(a)$ používá někdy název diferenciál a označení $\mathrm{d}F(a)$ nebo $DF(a)$.

(iii) V případě, že F je zobrazení z $\mathbb{R}$ do $\mathbb{R}^k$, symbolem $F'(a)$ se běžně označuje vektor $\lim_{t\to 0} \frac{F(a+t)-F(a)}{t}$. (Tutéž limitu můžeme podle Definice 2.35 zapsat také jako $\frac{\partial F}{\partial x_1}(a)$.) Vzniká tedy kolize – symbol $F'(a)$ označuje dva různé objekty. Tato kolize se odstraní tím, že ztotožňujeme vektor $w \in \mathbb{R}^k$ a lineární zobrazení $L_w\colon \mathbb{R} \to \mathbb{R}^k$ dané předpisem $L_w(t) = tw$. Je snadno vidět, že zobrazení $w \mapsto L_w$ je lineární izometrie prostorů $\mathbb{R}^k$ a $\mathcal{L}(\mathbb{R}, \mathbb{R}^k)$. Z Tvrzení 2.38 vyplývá, že existence $F'(a)$ v obou smyslech znamená totéž.

(iv) Přímo z definice derivace okamžitě vyplývá, že pokud $F\colon \mathbb{R}^n \to \mathbb{R}^k$ je lineární zobrazení, pak $F'(a) = F$ pro každý bod $a \in \mathbb{R}^n$.

Následující tvrzení ukazuje, že pojem derivace zobrazení se snadno převádí na pojem derivace funkce – derivaci lze počítat „po složkách“.

2.38 Tvrzení. *Nechť $F = (F_1, \dots, F_k)$ je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^k$, $a \in \mathbb{R}^n$ a nechť $L = (L_1, \dots, L_k)$ je lineární zobrazení $\mathbb{R}^n$ do $\mathbb{R}^k$. Pak $L = F'(a)$, právě když platí $L_1 = F_1'(a), \dots, L_k = F_k'(a)$.*

Jinými slovy: $F'(a) = (F_1'(a), \dots, F_k'(a))$, má-li jedna strana rovnosti smysl. Existuje-li tedy derivace $F'(a)$, je určena jednoznačně.

Důkaz. Protože

$$\frac{F(a+h) - F(a) - L(h)}{||h||} =$$

$$= \left(\frac{F_1(a+h) - F_1(a) - L_1(h)}{||h||}, \dots, \frac{F_k(a+h) - F_k(a) - L_k(h)}{||h||} \right)$$

a limita zobrazení do $\mathbb{R}^k$ se „počítá po složkách“, (2.9) platí, právě když

$$\lim_{h\to 0} \frac{F_i(a+h) - F_i(a) - L(h)}{||h||} = 0, \qquad \text{tj.} \qquad L_i = F_i'(a), \qquad i = 1, \dots, k.$$

Jednoznačnost derivace $F'(a)$ nyní okamžitě vyplývá z Důsledku 2.11.

Definici derivace pro zobrazení můžeme přepsat zcela analogicky, jako jsme to udělali v Tvrzení 2.8 pro případ funkcí. Zcela obdobný snadný důkaz vynecháme.

2.39 Tvrzení. *Nechť F je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^k$, $a \in \mathbb{R}^n$ a nechť $L\colon \mathbb{R}^n \to \mathbb{R}^k$ je lineární zobrazení. Nechť $r(h)$ je zobrazení určené rovnicí*

$$F(a+h) - F(a) = L(h) + r(h).$$

Pak následující tvrzení jsou ekvivalentní.

(i) $L = F'(a)$.

(ii) $||r(h)|| = o(||h||),\ h \to 0$.

(iii) *Existuje zobrazení ω z $\mathbb{R}^n$ do $\mathbb{R}^k$, které je spojité v $0 \in \mathbb{R}^n$, a pro které platí $\omega(0) = 0$ a $r(h) = ||h||\omega(h)$.*

2.40 Poznámka. Z Tvrzení 2.38 a Poznámky 2.9 (iv) vyplývá, že pojem derivace $F'(a)$ zobrazení F z $\mathbb{R}^n$ do $\mathbb{R}^k$ nezávisí na volbě normy v $\mathbb{R}^n$. Protože pojem limity zobrazení do $\mathbb{R}^k$ nezávisí na volbě normy v $\mathbb{R}^k$, nezávisí na volbě normy v $\mathbb{R}^k$ ani pojem derivace $F'(a)$ (to vidíme přímo z definice derivace).

Z Tvrzení 2.38 a Věty 2.14 a Věty 2.28 snadno dostáváme následující výsledek. (Je ovšem také možno použít Tvrzení 2.39 a zopakovat důkazy již provedené pro funkce.)

2.41 Věta. *Nechť F je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^k$, $a \in \mathbb{R}^n$ a nechť existuje $F'(a)$. Pak*

(i) *F je spojité v bodě a.*

(ii) *Pro každé $v \in \mathbb{R}^n$ a $i = 1, \ldots, n$ platí*

$$D_v F(a) = F'(a)v \quad \text{a} \quad \frac{\partial F}{\partial x_i}(a) = F'(a)e_i.$$

2.42 Definice. *Nechť $F = (F_1, \ldots, F_k)$ je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^k$, $a \in \mathbb{R}^n$ a nechť existuje $F'(a)$. Pak matici $[F'(a)]$ tohoto lineárního zobrazení nazýváme Jacobiho matice zobrazení F v bodě a. Pokud $k = n$, pak se determinant Jacobiho matice $[F'(a)]$ nazývá Jacobiho determinant (krátce jacobián) zobrazení F v bodě a a budeme jej označovat symboly*

$$J_F(a), \quad \frac{D(F_1, \ldots, F_n)}{D(x_1, \ldots, x_n)}(a).$$

2.43 Poznámka.

(i) Jacobián se často nazývá také *funkční determinant* funkcí $F_1, \ldots, F_n$.

(ii) Značení Jacobiho matice a jacobiánu dosti kolísá. Někteří autoři ztotožňují lineární zobrazení $F'(a)$ a jeho matici $[F'(a)]$. Někdy se také Jacobiho matice $[F'(a)]$ označuje podobným symbolem jako jacobián, například symbolem

$$\frac{\mathcal{D}(F_1, \ldots, F_k)}{\mathcal{D}(x_1, \ldots, x_n)}(a).$$

Souvislost derivace $F'(a)$ s parciálními derivacemi složek $F_1, \ldots, F_k$ udává následující věta.

2.44 Věta. *Nechť $F = (F_1, \ldots, F_k)$ je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^k$, $a \in \mathbb{R}^n$. Pak platí následující tvrzení.*

(i) *Existuje-li $F'(a)$, pak Jacobiho matice má tvar*

$$[F'(a)] = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(a) & \ldots & \frac{\partial F_1}{\partial x_n}(a) \\ \frac{\partial F_2}{\partial x_1}(a) & \ldots & \frac{\partial F_2}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial F_k}{\partial x_1}(a) & \ldots & \frac{\partial F_k}{\partial x_n}(a) \end{pmatrix}, \tag{2.10}$$

neboli

$$[F'(a)] = \left(\frac{\partial F_i}{\partial x_j}(a)\right)_{j=1,\dots,n}^{i=1,\dots,k}.$$

(ii) *Nechť všechny parciální derivace* $\dfrac{\partial F_i}{\partial x_j}$, $i = 1,\dots,k$, $j = 1,\dots,n$, *jsou funkce spojité v bodě* a. *Pak existuje* $F'(a)$.

Důkaz. (i) Podle Tvrzení 2.38 platí $F'(a) = (F_1'(a),\dots,F_k'(a))$, takže podle Věty 2.10 existují parciální derivace $\dfrac{\partial F_i}{\partial x_j}(a)$, $i = 1,\dots,k$, $j = 1,\dots,n$. Zvolme nyní index j, $1 \le j \le n$. Dostáváme

$$F'(a)e_j = (F_1'(a)e_j,\dots,F_k'(a)e_j) = \left(\frac{\partial F_1}{\partial x_j}(a),\dots,\frac{\partial F_k}{\partial x_j}(a)\right).$$

Vektor $\left(\dfrac{\partial F_1}{\partial x_j}(a),\dots,\dfrac{\partial F_k}{\partial x_j}(a)\right) = \dfrac{\partial F}{\partial x_j}(a)$ je tedy skutečně j-tým sloupcovým vektorem Jacobiho matice $[F'(a)]$.

(ii) Z předpokladů a Věty 2.20 vyplývá existence derivací složek $F_1'(a),\dots,F_k'(a)$. Podle Tvrzení 2.38 existuje také $F'(a)$.

2.45 Poznámka.

(i) Někteří autoři definují Jacobiho matici (a jacobián) zobrazení F pomocí vzorce (2.10) kdykoliv existují parciální derivace složek. My však v definici požadujeme – jak je nyní obvyklejší – existenci derivace.

(ii) Protože j-tý sloupcový vektor Jacobiho matice je roven $\frac{\partial F}{\partial x_j}(a)$, platí

$$[F'(a)] = \left[\frac{\partial F}{\partial x_1}(a),\dots,\frac{\partial F}{\partial x_n}(a)\right].$$

Protože i-tý řádkový vektor $[F'(a)]$ je gradient složky F_i, někdy píšeme

$$[F'(a)] = \begin{pmatrix} \operatorname{grad} F_1(a) \\ \vdots \\ \operatorname{grad} F_k(a) \end{pmatrix}.$$

(iii) Je-li F zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^n$ a $a \in \mathbb{R}^n$, pak

$$\frac{D(F_1,\dots,F_n)}{D(x_1,\dots,x_n)}(a) = \begin{vmatrix} \frac{\partial F_1}{\partial x_1}(a) & \dots & \frac{\partial F_1}{\partial x_n}(a) \\ \frac{\partial F_2}{\partial x_1}(a) & \dots & \frac{\partial F_2}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1}(a) & \dots & \frac{\partial F_n}{\partial x_n}(a) \end{vmatrix},$$

má-li levá strana rovnosti smysl.

Je přirozené a potřebné definovat jacobián, i když funkce $F_1, \dots, F_k$ závisí na více než k souřadnicích.

Například jacobián $\dfrac{D(xyz, x^2y^2z)}{D(x,z)}(1,2,3)$ přirozeně počítáme tak, že y považujeme za konstantu ($y = 2$); tj. počítáme jej jako $\dfrac{D(2xz, 4x^2z)}{D(x,z)}(1,3)$. Následující definice vypadá tedy složitě jen kvůli nepřehlednosti přesného formálního zápisu.

2.46 Definice. (rozšíření pojmu Jacobiho determinantu) *Nechť je dáno zobrazení $F = (F_1, \dots, F_k)$ z $\mathbb{R}^n$ do $\mathbb{R}^k$ $(k \le n)$, bod $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ a přirozená čísla $1 \le i_1 < i_2 < \dots < i_k \le n$. Uvažujme parciální zobrazení*

$$F^*(t_1, \dots, t_k) \ := F(a_1, \dots, a_{i_1-1}, t_1, a_{i_1+1}, \dots, a_{i_k-1}, t_k, a_{i_k+1}, \dots, a_n)$$

z $\mathbb{R}^k$ do $\mathbb{R}^k$. Diferenciál (derivace) $\mathrm{d}F^*(a_{i_1}, \dots, a_{i_k})$ parciálního zobrazení se nazývá parciální diferenciál (derivace) zobrazení F v bodě a podle proměnných $x_{i_1}, \dots, x_{i_k}$. Pokud tento parciální diferenciál existuje, klademe

$$\frac{D(F_1, \dots, F_k)}{D(x_{i_1}, \dots, x_{i_k})}(a) \ := J_{F^*}(a_{i_1}, \dots, a_{i_k}).$$

2.47 Poznámka.

(i) Z definice diferenciálu je snadno vidět, že pokud existuje diferenciál $\mathrm{d}F(a)$, má F v bodě a všechny parciální diferenciály. Ohledně vztahu mezi diferenciálem a parciálními diferenciály viz Poznámka 2.58 níže.

(ii) Podle Věty 2.44 (i) platí rovnost

$$\frac{D(F_1, \dots, F_k)}{D(x_{i_1}, \dots, x_{i_k})}(a) = \begin{vmatrix} \frac{\partial F_1}{\partial x_{i_1}}(a) & \dots & \frac{\partial F_1}{\partial x_{i_k}}(a) \\ \frac{\partial F_2}{\partial x_{i_1}}(a) & \dots & \frac{\partial F_2}{\partial x_{i_k}}(a) \\ \vdots & & \vdots \\ \frac{\partial F_k}{\partial x_{i_1}}(a) & \dots & \frac{\partial F_k}{\partial x_{i_k}}(a) \end{vmatrix},$$

má-li levá strana rovnosti smysl.

2.3 Derivace složeného zobrazení a složené funkce

Nyní zobecníme klasickou větu o derivaci složené funkce na případ zobrazení mezi eukleidovskými prostory.

Nejdříve provedeme (nepřesnou) heuristickou úvahu. Nechť f je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^m$ a g je zobrazení z $\mathbb{R}^m$ do $\mathbb{R}^s$. Nechť f má derivaci v bodě $a \in \mathbb{R}^n$ a nechť g má derivaci v bodě $b := f(a)$. Prvky v prostorech $\mathbb{R}^n, \mathbb{R}^m, \mathbb{R}^s$ označujme pořadě x, y, z. Malé změně Δx odpovídá malá změna

$$\Delta y = f(a + \Delta x) - f(a) \approx f'(a)(\Delta x).$$

Této změně pak odpovídá změna

$$\Delta z = g(f(a + \Delta x)) - g(f(a)) = g(b + \Delta y) - g(b)$$

$$\approx g'(b)(\Delta y) \approx g'(b)\left(f'(a)(\Delta x)\right) = g'(b) \circ f'(a)(\Delta x).$$

Je tedy přirozené se domnívat, že derivace $(g \circ f)'(a)$ existuje a rovná se složení derivací $g'(f(a)) \circ f'(a)$. Přesný důkaz však dá trochu práce. Budeme potřebovat následující snadné tvrzení.

2.48 Tvrzení. *Nechť f je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^m$, které má derivaci $f'(a)$ v bodě $a \in \mathbb{R}^n$. Pak*

$$||f(a + h) - f(a)|| = O(||h||), \quad h \to 0.$$

Důkaz. Podle Tvrzení 2.39 lze psát $f(a+h) - f(a) = f'(a)h + r(h)$, kde $||r(h)|| = o(||h||)$, $h \to 0$. Existuje tedy $\delta > 0$ takové, že $||r(h)|| < ||h||$ kdykoliv $0 < ||h|| < \delta$. Pro taková h tedy platí (užíváme (1.8) z Věty 1.123)

$$||f(a+h) - f(a)|| = ||f'(a)h + r(h)|| \leq ||f'(a)|| \cdot ||h|| + ||r(h)|| \leq (||f'(a)|| + 1)||h||.$$

2.49 Poznámka. Zopakujme, že podmínka $||f(a + h) - f(a)|| = O(||h||)$, $h \to 0$ podle definice znamená, že podíl $||f(a+h) - f(a)||/||h||$ je omezený na nějakém redukovaném okolí bodu a. Jiný ekvivalentní zápis této podmínky je

$$\limsup_{h \to 0} \frac{||f(a + h) - f(a)||}{||h||} < \infty.$$

Je-li tato podmínka – která je zřejmě silnější než spojitost funkce f v bodě a – splněna, říká se někdy také, že f je lipschitzovská v bodě a.

2.50 Věta. (věta o derivaci složeného zobrazení) *Nechť f je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^m$ a g je zobrazení z $\mathbb{R}^m$ do $\mathbb{R}^s$. Nechť f má derivaci v bodě $a \in \mathbb{R}^n$ a nechť g má derivaci v bodě $b := f(a)$. Pak zobrazení $g \circ f$ má derivaci v bodě a a platí*

$$(g \circ f)'(a) = g'(b) \circ f'(a).$$

Důkaz. Podle Tvrzení 2.39 lze psát $f(a+h)-f(a) = f'(a)h + r_1(h)$, kde $r_1(h) = \omega_1(h)\|h\|$ a $\lim_{h\to 0}\omega_1(h) = \omega_1(0) = 0$. Podobně $g(b+k) - g(b) = g'(b)k + r_2(k)$, kde $r_2(k) = \omega_2(k)\|k\|$ a $\lim_{k\to 0}\omega_2(k) = \omega_2(0) = 0$.

Snadno dostáváme

$$\begin{aligned}(g\circ f)(a+h) - (g\circ f)(a) &= g(b + (f(a+h) - f(a))) - g(b)\\ &= g'(b)(f(a+h)-f(a)) + r_2(f(a+h)-f(a))\\ &= g'(b)(f'(a)h + r_1(h)) + r_2(f(a+h)-f(a))\\ &= g'(b)\circ f'(a)(h) + g'(b)(r_1(h)) + r_2(f(a+h)-f(a)).\end{aligned}$$

Podle Tvrzení 2.39 (ii) tedy stačí dokázat, že pro

$$r(h) := g'(b)(r_1(h)) + r_2(f(a+h)-f(a))$$

platí $\|r(h)\| = o(\|h\|),\ h\to 0$. K tomu zřejmě stačí ověřit, že $\|g'(b)(r_1(h))\| = o(\|h\|),\ h\to 0$ a také $\|r_2(f(a+h)-f(a))\| = o(\|h\|),\ h\to 0$. Podle Věty 1.123 platí

$$\|g'(b)(r_1(h))\| \le \|g'(b)\|\cdot\|r_1(h)\| = \|g'(b)\|\cdot\|\omega_1(h)\|\cdot\|h\|,$$

takže dostáváme $\|g'(b)(r_1(h))\| = o(\|h\|),\ h\to 0$.

Podle Věty 1.40 o limitě složeného zobrazení (zde podstatně používáme spojitost zobrazení ω_2 v 0) platí $\lim_{h\to 0}\omega_2(f(a+h)-f(a)) = 0$. S pomocí Tvrzení 2.48 (srov. Poznámka 2.49) dostáváme

$$\lim_{h\to 0}\frac{\|r_2(f(a+h)-f(a))\|}{\|h\|} = \lim_{h\to 0}\frac{\|f(a+h)-f(a)\|}{\|h\|}\cdot\|\omega_2(f(a+h)-f(a))\| = 0,$$

takže také $\|r_2(f(a+h)-f(a))\| = o(\|h\|),\ h\to 0$.

2.51 Důsledek. *Nechť $f = (f_1,\dots,f_m)$ je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^m$ a $g = (g_1,\dots,g_s)$ je zobrazení z $\mathbb{R}^m$ do $\mathbb{R}^s$. Nechť f má derivaci v bodě $a\in\mathbb{R}^n$ a nechť g má derivaci v bodě $b := f(a)$. Položme ještě $h = (h_1,\dots,h_s) := g\circ f$. Pak platí*

$$[h'(a)] = [g'(b)]\cdot[f'(a)], \quad \textit{tj.}$$

$$\frac{\partial h_i}{\partial x_j}(a) = \sum_{k=1}^{m}\frac{\partial g_i}{\partial y_k}(b)\,\frac{\partial f_k}{\partial x_j}(a),\quad 1\le i\le s,\ 1\le j\le n. \tag{2.11}$$

Pokud $n = m = s$, pak

$$J_h(a) = J_g(b)\cdot J_f(a).$$

Důkaz. Vzorec pro Jacobiho matici h plyne okamžitě z rovnosti $h'(a) = g'(b)\circ f'(a)$ a z věty o matici složení dvou lineárních zobrazení známé z lineární algebry. Podle Věty 2.44 (i) platí tedy rovnost

$$\begin{pmatrix}\frac{\partial h_1}{\partial x_1}(a) & \dots & \frac{\partial h_1}{\partial x_n}(a)\\ \vdots & & \vdots\\ \frac{\partial h_s}{\partial x_1}(a) & \dots & \frac{\partial h_s}{\partial x_n}(a)\end{pmatrix} = \begin{pmatrix}\frac{\partial g_1}{\partial y_1}(b) & \dots & \frac{\partial g_1}{\partial y_m}(b)\\ \vdots & & \vdots\\ \frac{\partial g_s}{\partial y_1}(a) & \dots & \frac{\partial g_s}{\partial y_m}(b)\end{pmatrix}\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a)\\ \vdots & & \vdots\\ \frac{\partial f_m}{\partial x_1}(a) & \dots & \frac{\partial f_m}{\partial x_n}(a)\end{pmatrix},$$

která je ekvivalentní s (2.11).

Poslední rovnost plyne z věty o determinantu součinu (čtvercových) matic.

2.52 Poznámka. Jestliže zobrazení F z $\mathbb{R}^n$ do $\mathbb{R}^m$ je diferencovatelné v bodě $a \in \mathbb{R}^n$ a $L: \mathbb{R}^m \to \mathbb{R}^s$ je lineární, pak z Poznámky 2.37 (iv) a Věty 2.50 okamžitě vyplývá rovnost $(L \circ F)'(a) = L \circ F'(a)$.

2.53 Poznámka. Ve starších učebnicích se často píše $y(x_1, \dots, x_n)$ místo f, $z(y_1, \dots, y_m)$ místo g a $z(x_1, \dots, x_n)$ místo h. V případě $n = m = s$ se pak vzorec pro jacobián složeného zobrazení zapisuje ve tvaru

$$\frac{D(z_1, \dots, z_n)}{D(x_1, \dots, x_n)}(a) = \frac{D(z_1, \dots, z_n)}{D(y_1, \dots, y_n)}(b)\,\frac{D(y_1, \dots, y_n)}{D(x_1, \dots, x_n)}(a),$$

který zobecňuje klasický vzorec $\frac{\mathrm{d}z}{\mathrm{d}x}(a) = \frac{\mathrm{d}z}{\mathrm{d}y}(b)\,\frac{\mathrm{d}y}{\mathrm{d}x}(a)$ pro funkce jedné proměnné. Tento názorný vzorec však není formálně korektní: symbolem $z = (z_1, \dots, z_n)$ jsou označeny dvě různá zobrazení a symboly $y_1, \dots, y_n$ jsou označeny jak nezávislé proměnné, tak funkce, které za ně dosazujeme.

Nejčastěji se vzorec (2.11) používá v případě $s = 1$, tj. když vnější zobrazení g je reálná funkce. Proto tento nejdůležitější speciální případ zformulujeme jako samostatnou větu.

2.54 Věta. (o derivaci složené funkce, „řetízkové pravidlo")
Nechť každá z funkcí $f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n)$ *má totální diferenciál v bodě* $a \in \mathbb{R}^n$ *a nechť funkce* $g(y_1, \dots, y_m)$ *má totální diferenciál v bodě* $b := (f_1(a), \dots, f_m(a))$. *Pak složená funkce*

$$h(x) := g(f_1(x), \dots, f_m(x))$$

má v bodě a *totální diferenciál a platí*

$$\frac{\partial h}{\partial x_i}(a) = \frac{\partial g}{\partial y_1}(b)\,\frac{\partial f_1}{\partial x_i}(a) + \dots + \frac{\partial g}{\partial y_m}(b)\,\frac{\partial f_m}{\partial x_i}(a), \qquad i = 1, \dots, n. \tag{2.12}$$

Důkaz. Stačí položit $f := (f_1, \dots, f_m)$, uvážit, že f má derivaci v bodě a (Tvrzení 2.38) a použít (2.11) pro případ $s = 1$.

2.55 Poznámka. („invariantnost formy totálního diferenciálu") Za předpokladů předchozí věty platí

$$dh(a) = \frac{\partial g}{\partial y_1}(b)\,\mathrm{d}f_1(a) + \dots + \frac{\partial g}{\partial y_m}(b)\,\mathrm{d}f_m(a). \tag{2.13}$$

Tato jedna rovnost je ekvivalentní s n rovnostmi (2.12); jde totiž o rovnost dvou lineárních forem (nezávisle proměnných $k_1, \dots, k_n$)

$$\sum_{i=1}^{n} \frac{\partial h}{\partial x_i}(a)\,k_i = \frac{\partial g}{\partial y_1}(b) \sum_{i=1}^{n} \frac{\partial f_1}{\partial x_i}(a)\,k_i + \dots + \frac{\partial g}{\partial y_m}(b) \sum_{i=1}^{n} \frac{\partial f_m}{\partial x_i}(a)\,k_i,$$

přičemž rovnost koeficientů u k_i těchto dvou lineárních forem je i-tá rovnost z (2.12).

Označme funkce $f_1, \dots, f_m$ symboly $y_1^*, \dots, y_m^*$, funkci g symbolem z a funkci h symbolem z^*. Pak $z^*(x) = z(y_1^*(x), \dots, y_m^*(x))$ a rovnost (2.13) má formu

$$dz^*(a) = \frac{\partial z}{\partial y_1}(b)\,\mathrm{d}y_1^*(a) + \dots + \frac{\partial z}{\partial y_m}(b)\,\mathrm{d}y_m^*(a),$$

která se velmi podobá formě zápisu totálního diferenciálu, ve kterém $y_1, \dots, y_m$ nejsou funkce, ale nezávisle proměnné (srov. Poznámka 2.21). Ve starší literatuře se často místo z^*, y_i^* psalo (formálně nekorektně) z, y_i a hovořilo se o „invariantnosti formy totálního diferenciálu“.

2.56 Poznámka. Často používáme „řetízkové pravidlo“ ve speciálním případě, kdy $f = (f_1, \dots, f_m)$ je zobrazení z $\mathbb{R}$ do $\mathbb{R}^m$ a g je funkce z $\mathbb{R}^m$ do $\mathbb{R}$. Pokud $a \in \mathbb{R}$ a existují derivace $f'(a)$ a $g'(b)$ pro $b := f(a)$, pak $h := g \circ f$ je reálná funkce reálné proměnné, která má v bodě a derivaci

$$h'(a) = \frac{\partial g}{\partial y_1}(b)\, f_1'(a) + \dots + \frac{\partial g}{\partial y_m}(b)\, f_m'(a).$$

Chápeme-li tedy $f'(a)$ jako prvek $\mathbb{R}^m$ (viz Poznámka 2.37 (iii)), můžeme psát

$$h'(a) = \langle \operatorname{grad} g(b), f'(a) \rangle.$$

2.57 Příklad. Nechť reálné funkce reálné proměnné f_1, f_2 mají derivaci v bodě $a \in \mathbb{R}$. Pak lze psát

$$f_1(x) \cdot f_2(x) = S(f_1(x), f_2(x)) = S(f(x)),$$

kde $f = (f_1, f_2)$ a $S\colon \mathbb{R}^2 \to \mathbb{R}$, $S(u,v) = uv$. Podle předpokladů existuje $f'(a) = (f_1'(a), f_2(a))$. Protože parciální derivace $\dfrac{\partial S}{\partial u} = v$ a $\dfrac{\partial S}{\partial v} = u$ jsou spojité na $\mathbb{R}^2$, je S na $\mathbb{R}^2$ diferencovatelná. Položíme-li $b := f(a)$, platí $\operatorname{grad} S(b) = (f_2(a), f_1(a))$, takže podle předchozí poznámky máme

$$(f \cdot g)'(a) = \langle (f_2(a), f_1(a)), (f_1'(a), f_2'(a)) \rangle = f_1'(a)\, f_2(a) + f_2'(a)\, f_1(a).$$

Zcela analogicky lze odvodit vzorec pro derivaci podílu funkcí.

2.58 Poznámka. Nejčastěji se parciální diferenciály (viz Definice 2.46) používají, vyšetřujeme-li zobrazení f z $\mathbb{R}^n \times \mathbb{R}^k = \mathbb{R}^{n+k}$ do $\mathbb{R}^p$. Pro $c = (a,b) \in \mathbb{R}^n \times \mathbb{R}^k$ uvažujme parciální diferenciály

$$f_x'(c) := d\,f(\cdot, b)(a) = (f(\cdot, b))'(a), \quad f_y'(c) := d\,f(a, \cdot)(b) = (f(a, \cdot))'(b).$$

(Někdy se místo o parciálních diferenciálech hovoří o parciálních derivacích a také se označují symboly $\dfrac{\partial f}{\partial x}(c)$, $\dfrac{\partial f}{\partial y}(c)$.) Existuje-li $f'(c)$, pak z definice derivace snadno dostáváme, že oba parciální diferenciály v bodě c existují a pro $h_1 \in \mathbb{R}^n$, $h_2 \in \mathbb{R}^k$ platí

$$f_x'(c)h_1 = f'(c)(h_1, 0), \quad f_y'(c)h_2 = f'(c)(0, h_2)$$

a tedy

$$f'(c)(h_1, h_2) = f'(c)(h_1, 0) + f'(c)(0, h_2) = f_x'(c)\, h_1 + f_y'(c)\, h_2. \tag{2.14}$$

Předpokládejme navíc, že je dáno zobrazení $\omega = (\omega_1, \omega_2)$ z $\mathbb{R}^s$ do $\mathbb{R}^n \times \mathbb{R}^k$ a bod $d \in \mathbb{R}^s$, pro který $\omega(d) = c$ a existuje $\omega'(d)$. Není těžké dokázat, že platí $\omega'(d) = (\omega_1'(d), \omega_2'(d))$. Z Věty 2.50 a ze vzorce (2.14) dostáváme

$$(f \circ \omega)'(d) = f_x'(c) \circ \omega_1'(d) + f_y'(c) \circ \omega_2'(d). \tag{2.15}$$

2.4 Věta o přírůstku funkce

Následující tvrzení je snadným a přímočarým zobecněním Lagrangeovy věty.

2.59 Tvrzení. *Nechť f je funkce n proměnných a $a \neq b$ jsou dva body z $\mathbb{R}^n$. Předpokládejme, že f má směrovou derivaci $D_{b-a}f(x)$ v každém bodě x otevřené úsečky $P := \overline{ab} \setminus \{a, b\}$ a je spojitá v bodech a, b vzhledem k úsečce $\overline{ab}$. Pak existuje bod $\xi \in P$ takový, že*

$$f(b) - f(a) = D_{b-a}f(\xi).$$

Důkaz. Položme $h := b - a$ a $g(t) := f(a + th)$, $t \in [0, 1]$. Protože zobrazení $\varphi: t \mapsto a + th$ je spojité na $\mathbb{R}$, $\varphi(0) = a$, $\varphi(1) = b$ a $\varphi([0, 1]) = \overline{ab}$, z Věty 1.38 dostáváme, že funkce g je spojitá v bodech a, b vzhledem k intervalu $[0, 1]$ (tj. zprava a zleva). Je-li $t \in (0, 1)$, máme

$$g'(t) = \lim_{u \to 0} \frac{g(t + u) - g(t)}{u} = \lim_{u \to 0} \frac{f(a + th + uh) - f(a + th)}{u} = D_h f(a + th).$$

Funkce g tedy splňuje na intervalu $[0, 1]$ předpoklady Lagrangeovy věty. Existuje proto $\theta \in (0, 1)$ takové, že $g(1) - g(0) = g'(\theta)$, takže

$$f(b) - f(a) = g(1) - g(0) = g'(\theta) = D_h f(a + \theta h).$$

Stačí tedy položit $\xi := a + \theta h$.

Věta o přírůstku funkce pro funkce více proměnných se většinou uvádí ve slabších verzích, které jsou však vhodnější pro aplikace.

2.60 Věta. (o přírůstku funkce) *Nechť f je funkce n proměnných, která má diferenciál v každém bodě otevřené množiny G. Nechť $a, b \in G$ jsou takové body, že $\overline{ab} \subset G$. Pak existuje bod $\xi \in \overline{ab}$, pro který*

$$f(b) - f(a) = D_{b-a}f(\xi) = \mathrm{d}_{b-a}f(\xi) = f'(\xi)(b - a) = \langle \operatorname{grad} f(\xi), b - a \rangle.$$

Důkaz. Je-li $a = b$, bod $\xi := a = b$ má zřejmě požadované vlastnosti. Pokud $a \neq b$, stačí aplikovat předchozí tvrzení (přitom užíváme Věty 2.14, 2.28).

Hlavní význam věty o přírůstku funkce spočívá v tom, že jsme schopni odhadnout shora velikost přírůstku funkce pomocí derivace (diferenciálu, gradientu). Jedním z okamžitých důsledků je například následující tvrzení. Zopakujme (viz Poznámka 2.29), že pokud reálná funkce f je diferencovatelná v bodě $x \in \mathbb{R}^n$, pak $\|\operatorname{grad} f(x)\| = \|f'(x)\|$.

2.61 Tvrzení. *Nechť f je funkce n proměnných, která má totální diferenciál na otevřené konvexní množině $C \subset \mathbb{R}^n$ a nechť*

$$\sup\{\|f'(x)\|\colon\ x \in C\} = \sup\{\|\operatorname{grad} f(x)\|\colon\ x \in C\} =: K < \infty.$$

Pak f je lipschitzovská na C s konstantou K, tj.

$$|f(b) - f(a)| \leq K\,\|b - a\|, \qquad a, b \in C.$$

Důkaz. Nechť $a, b \in C$. Protože C je konvexní, platí $\overline{ab} \subset C$. Podle Věty 2.60 existuje tedy bod $\xi \in \overline{ab}$ takový, že $f(b) - f(a) = \langle \operatorname{grad} f(\xi), b - a\rangle$. Z Cauchyovy nerovnosti pak dostáváme

$$|f(b) - f(a)| = |\langle \operatorname{grad} f(\xi), b - a\rangle| \leq \|b - a\| \cdot \|\operatorname{grad} f(\xi)\| \leq K\|b - a\|.$$

Aplikujeme-li předchozí tvrzení pro $K = 0$, dostáváme, že každá funkce s nulovým diferenciálem na otevřené konvexní množině C je na C konstantní.

Toto pozorování nyní snadno zobecníme z případu otevřené konvexní množiny na případ souvislé otevřené množiny.

2.62 Věta. *Nechť f je funkce n proměnných, $G \subset \mathbb{R}^n$ je otevřená souvislá množina a $\mathrm{d}f(x) = 0$ pro každý bod $x \in G$. Pak funkce f je na množině G konstantní.*

Důkaz. Zvolme $a \in G$ a položme

$$A := \{x \in G\colon\ f(x) = f(a)\}, \quad B := \{x \in G\colon\ f(x) \neq f(a)\}.$$

Předpokládejme, že f není na G konstantní, tj. $B \neq \emptyset$. Protože G je souvislá a $A \cup B = G$, existuje bod $c \in G \cap \overline{A} \cap \overline{B}$ (viz 1.115). Zvolme $\delta > 0$ tak malé, aby $C := U_\delta(c) \subset G$. Protože C je konvexní, je f na C konstantní. To je ale spor, protože $C \cap A \neq \emptyset$ i $C \cap B \neq \emptyset$.

Nyní se budeme zabývat otázkou, zda lze větu o přírůstku funkce zobecnit na případ vektorových funkcí více proměnných, tj. na případ, kdy f je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^k$.

Následující přirozený příklad ukazuje, že tomu tak není – přímočaré zobecnění Věty 2.60 neplatí již pro zobrazení z $\mathbb{R}$ do $\mathbb{R}^2$.

2.63 Příklad. Položme $f(t) := (\cos t, \sin t)$ pro $t \in \mathbb{R}$. Pak $f\colon \mathbb{R} \to \mathbb{R}^2$ je zřejmě zobrazení třídy C^∞ a $f'(t) = (-\sin t, \cos t)$. (Při tomto zápisu je derivací prvek $\mathbb{R}^2$, který však ztotožňujeme s lineárním zobrazením $L\colon \mathbb{R} \to \mathbb{R}^2$ daným předpisem $L(h) = h(-\sin t, \cos t)$; viz Poznámka 2.37 (iii)). Zřejmě $\|f'(t)\| = 1$ pro každé $t \in \mathbb{R}$. Pro body $a = 0$, $b = 2\pi$ máme $f(b) - f(a) = (1,0) - (1,0) = (0,0)$. Pro každý bod $\xi \in \mathbb{R}$ však platí $\|f'(\xi)(b-a)\| = 2\pi$, takže $f(b) - f(a) = f'(\xi)(b-a)$ neplatí pro žádné $\xi \in \mathbb{R}$.

Tento příklad má názornou „kinematickou" interpretaci. Vektorová funkce f totiž popisuje rovnoměrný pohyb hmotného bodu po kružnici (s jednotkovou rychlostí; $\|f'(t)\| = 1$). Kdyby platilo zobecnění Věty 2.60, musela by se „průměrná (vektorová) rychlost" $(f(b) - f(a))/(b-a)$ rovnat okamžité (vektorové) rychlosti $f'(\xi)$ v nějakém bodě ξ. To však není možné, protože v našem případě je průměrná rychlost nulový vektor a okamžitá rychlost je všude jednotková.

Přestože přímé zobecnění Věty 2.60 neplatí, lze snadno zobecnit její důsledek – Tvrzení 2.61, které hovoří o *odhadu shora* velikosti přírůstku funkce pomocí velikosti její derivace.

2.64 Věta. (věta o přírůstku funkce pro zobrazení) *Nechť $n, k \in \mathbb{N}$, $K \in \mathbb{R}$, $C \subset \mathbb{R}^n$ je otevřená konvexní množina, $F\colon C \to \mathbb{R}^k$ je zobrazení diferencovatelné na množině C a nechť*

$$\sup\{\|F'(x)\|\colon\ x \in C\} \le K.$$

Pak F je na C lipschitzovské s konstantou K, tj.

$$\|F(b) - F(a)\| \le K\,\|b-a\|, \qquad \text{kdykoliv } a, b \in C.$$

Důkaz. Jestliže $F(a) = F(b)$, nerovnost $\|F(b) - F(a)\| \le K\|b-a\|$ zřejmě platí. V opačném případě položme $v := \dfrac{F(b) - F(a)}{\|F(b) - F(a)\|}$ a uvažujme funkci $f\colon C \to \mathbb{R}$ danou předpisem $f(x) := \langle F(x), v\rangle$. Pak zřejmě

$$|f(b) - f(a)| = |\langle F(b) - F(a), v\rangle| = \|F(b) - F(a)\|.$$

Protože funkce $g := \langle \cdot, v\rangle$ je lineární na $\mathbb{R}^k$, podle Poznámky 2.52 pro $x \in C$ existuje $f'(x)$ a platí $f'(x)h = \langle F'(x)h, v\rangle$. Jestliže $\|h\| \le 1$, dostáváme

$$|f'(x)h| = |\langle F'(x)h, v\rangle| \le \|F'(x)h\| \cdot \|v\| \le \|F'(x)\| \le K.$$

Je tedy $\|f'(x)\| \le K$ pro každý bod $x \in C$, takže podle Tvrzení 2.61 platí

$$\|F(b) - F(a)\| = |f(b) - f(a)| \le K\|b-a\|.$$

2.65 Poznámka.

(i) Kdybychom Tvrzení 2.61 aplikovali pouze na funkce $\langle F(x), e_j\rangle$ dané „základními souřadnicovými funkcemi“ $\langle\cdot, e_j\rangle$ na $\mathbb{R}^k$, dostali bychom slabší odhad $\|F(b) - F(a)\| \leq \sqrt{k}K\|b - a\|$. Idea důkazu je v tom, že se použijí i jiné „souřadnicové funkce“ $\langle\cdot, v\rangle$ pro vhodná v.

(ii) Není těžké dokázat (např. analogicky jako předchozí větu), že pokud zobrazení F z $\mathbb{R}^n$ do $\mathbb{R}^k$ je spojité na úsečce $\overline{ab}$ a $\|D_{b-a}F(x)\| \leq K$ pro každé $x \in \overline{ab} \setminus \{a, b\}$, pak $\|F(b) - F(a)\| \leq K$. Dokonce lze dokázat (viz [Ca]), že místo $x \in \overline{ab}\setminus\{a, b\}$ lze psát $x \in \overline{ab}\setminus S$, kde S je libovolná spočetná množina.

2.5 Parciální derivace vyšších řádů a funkce třídy C^k

Nechť f je funkce n proměnných a nechť $i, j \in \{1, \ldots, n\}$. Pak $\dfrac{\partial f}{\partial x_i}$ je opět funkcí n proměnných (která může mít menší definiční obor než f).

Parciální derivaci této funkce podle j-té proměnné pak označujeme symbolem $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$. Tato funkce se nazývá *parciální derivací druhého řádu* funkce f; pokud $i \neq j$, jde o tzv. *smíšenou* parciální derivaci druhého řádu. Jinak řečeno, pro $a \in \mathbb{R}^n$ klademe

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

Obdobně definujeme parciální derivace vyšších řádů, například parciální derivaci třetího řádu

$$\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k} := \frac{\partial\left(\frac{\partial^2 f}{\partial x_j \partial x_k}\right)}{\partial x_i}.$$

Obecně je každá parciální derivace řádu $k+1$ parciální derivací (prvního řádu) parciální derivace řádu k.

Parciální derivace vyššího řádu se označují různými způsoby; například $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ se zapisuje často pomocí symbolů

$$D_{ji}f,\ f_{ji},\ f_{j,i},\ f''_{ji},\ f_{x_j x_i},\ f''_{x_j x_i}, \partial_{ji}f.$$

Obdobná symbolika se používá, jsou-li proměnné označeny písmeny. Například pro funkci $f(x, y, z)$ se parciální derivace druhého řádu $f_{2,3}$ označuje také symboly

$$\frac{\partial^2 f}{\partial y \partial z},\ D_{yz}f,\ f_{yz},\ f''_{yz},\ \partial_{yz}f.$$

Místo $\dfrac{\partial^3 f}{\partial x_2 \partial x_2 \partial x_3}$ píšeme většinou $\dfrac{\partial^3 f}{\partial x_2^2 \partial x_3}$, $\dfrac{\partial^2 f}{\partial y^2 \partial z}$ apod.

2.66 Poznámka. (důležitá) V klasické literatuře (včetně [D II]) se používala odlišná symbolika s „obráceným pořadím proměnných", podle které například

$$\frac{\partial^2 f}{\partial x \partial y} = f_{xy} := \frac{\partial \left(\frac{\partial f}{\partial x}\right)}{\partial y}.$$

Změna symboliky (poměrně nedávná) asi souvisí s operátorovým zápisem parciálních derivací, který je popsán níže. Změnou symboliky bohužel vznikl zmatek, který však naštěstí nevede k vážnějším chybám, protože (jak záhy dokážeme) v důležitých případech na pořadí derivování nezáleží.

Podobně není v literatuře jednotnost při definici symbolů f''_{yz}, $f''(a)(h,k)$ (viz oddíl 2.7) apod. Proto se v *této publikaci* řídíme jednotným pravidlem, že *operace, které odpovídají symbolům stojícím více napravo, se provádějí dříve*, které ve většině případů souhlasí se symbolikou, která je v současnosti nejčastější.

Operátorem se většinou rozumí zobrazení, které funkcím přiřazuje opět funkce (v moderní terminologii má pojem „operátor" širší význam).

Zobrazení, které každé funkci n proměnných f přiřazuje funkci $\frac{\partial f}{\partial x_i}$ je nejjednodušší parciální diferenciální operátor prvního řádu, který označujeme symbolem $\frac{\partial}{\partial x_i}$. Můžeme tedy psát

$$\frac{\partial}{\partial x_i} \colon f \mapsto \frac{\partial f}{\partial x_i} \qquad \text{neboli} \qquad \tfrac{\partial}{\partial x_i}(f) = \tfrac{\partial f}{\partial x_i}.$$

Operátor $\frac{\partial}{\partial x_i}$ se také označuje symbolem D_i nebo ∂_i. V operátorovém zápisu zřejmě platí

$$\frac{\partial}{\partial x_j}\left(\frac{\partial}{\partial x_i}(f)\right) = \frac{\partial^2 f}{\partial x_j \partial x_i}. \tag{2.16}$$

Parciální diferenciální operátor druhého řádu $\dfrac{\partial^2}{\partial x_j \partial x_i} = D_{ji} = \partial_{ji}$ je ovšem definován jako zobrazení

$$f \mapsto \frac{\partial^2 f}{\partial x_j \partial x_i},$$

takže platí

$$\frac{\partial^2}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j} \circ \frac{\partial}{\partial x_i},$$

kde napravo stojí složení dvou zobrazení (operátorů).

2.67 Poznámka. Důvodem změny klasické symboliky byla asi snaha, aby ve formuli (2.16) byl symbol ∂x_i napravo od symbolu ∂x_j na obou stranách rovnosti.

2.68 Poznámka. Při takto obecné definici ovšem může být hodnotou parciálního diferenciálního operátoru prázdná funkce. Většinou však tento operátor zužujeme na takový prostor funkcí, že po provedení operátoru na funkci z tohoto prostoru dostáváme funkci se stejným definičním oborem.

2.69 Příklad. Uvažujme funkci $f(x,y) = \sqrt{-x^2y^2}$ z Příkladu 2.2. Víme, že $\frac{\partial f}{\partial x}$ má definiční obor menší než f; definičním oborem funkce $\frac{\partial f}{\partial x}$ je x-ová osa. Proto zřejmě $\frac{\partial^2 f}{\partial y \partial x}$ neexistuje v žádném bodě, takže operátor $\frac{\partial^2}{\partial y \partial x}$ zobrazuje funkci f na prázdnou funkci (s definičním oborem $\emptyset$). Naproti tomu zřejmě $\frac{\partial^2 f}{\partial x^2} = \frac{\partial f}{\partial x}$.

Při výpočtech často používáme běžnou úmluvu, podle které například funkci $f\colon (x,y) \mapsto y^3 \sin^2 x$ označujeme symbolem $y^3 \sin^2 x$, pokud je z kontextu jasné, že jde o funkci dvou proměnných (a ne o číslo nebo třeba o parciální funkci $x \mapsto y^3 \sin^2 x$).

2.70 Příklad. Funkce $f(x,y) = x^y = e^{y \ln x}$ má definiční obor $G = (0,\infty) \times \mathbb{R}$. Snadno počítáme:

$$\frac{\partial x^y}{\partial x} = x^y \frac{y}{x} = x^{y-1} y, \quad \frac{\partial^2 (x^y)}{\partial y \partial x} = x^{y-1} y \ln x + x^{y-1},$$

$$\frac{\partial x^y}{\partial y} = x^y \ln x, \quad \frac{\partial^2 (x^y)}{\partial x \partial y} = x^y \frac{y}{x} \ln x + x^y \frac{1}{x}.$$

Vidíme tedy, že $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$ na G.

Brzy ale uvidíme (Příklad 2.78 níže), že „záměnnost parciálních derivací“ nenastává vždy.

Ve většině důležitých klasických vět, které pracují s pojmem parciální derivace p-tého řádu, se předpokládá, že všechny parciální derivace až do řádu p jsou spojité. Proto zavádíme následující terminologii.

2.71 Definice. *Nechť $G \subset \mathbb{R}^n$ je otevřená množina, je dána funkce $f\colon G \to \mathbb{R}$ a $p \in \mathbb{N}$. Řekneme, že f je třídy C^p, jestliže všechny parciální derivace funkce f až do řádu p jsou spojité na G. Množinu všech funkcí $f\colon G \to \mathbb{R}$ třídy C^p označujeme $C^p(G)$ a klademe $C^\infty(G) := \bigcap_{p=1}^\infty C^p(G)$. O funkci g řekneme, že je třídy C^p na G $(p \in \mathbb{N} \cup \{\infty\})$, jestliže $g \restriction_G \in C^p(G)$.*

2.72 Poznámka.

(i) $C^\infty(G)$ je tedy systém všech funkcí na G, které mají na G spojité parciální derivace všech řádů.

(ii) Je-li f třídy C^1 na G, má podle Věty 2.20 diferenciál v každém bodě množiny G, speciálně je na G spojitá.

(iii) Funkcemi třídy C^0 na G se rozumí funkce spojité na G.

(iv) V definici funkce třídy $C^p(G)$ ovšem stačí požadovat, aby f měla na G spojité parciální derivace řádu p. Pak má každá parciální derivace řádu $p-1$ spojité parciální derivace a je tedy spojitá (srov. (ii)). Postupně dostáváme spojitost parciálních derivací všech řádů $1 \le k \le p-1$.

Pojem parciálních derivací vyššího řádu zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^k$ (a následně zobrazení třídy C^p) bychom mohli definovat zcela obdobně jako pro funkce. Tento pojem ale nebudeme potřebovat a zobrazení třídy C^p budeme (ekvivalentně) definovat „po složkách“.

2.73 Definice. *Nechť $G \subset \mathbb{R}^n$ je otevřená množina, $p \in \mathbb{N} \cup \{\infty\}$ a $f = (f_1, \dots, f_k)$ je zobrazení $f: G \to \mathbb{R}^k$. Řekneme, že f je zobrazení třídy C^p, jestliže jeho složky $f_1, \dots, f_k$ jsou třídy C^p.*

2.74 Poznámka. Podle Poznámky 2.72(ii) a Tvrzení 2.38 vidíme, že každé zobrazení třídy C^p $(p \geq 1)$ na G má derivaci v každém bodě množiny G.

2.75 Příklad. Funkce $S(u, v) = uv$ je třídy C^∞ na $\mathbb{R}^2$. To plyne okamžitě z toho, že $S_u = v$, $S_v = u$, $S_{uv} = S_{vu} = 1$, $S_{uu} = S_{vv} = 0$ a parciální derivace třetího a vyššího řádu jsou nulové.

Z Věty 2.50 okamžitě vyplývá, že složení dvou diferencovatelných zobrazení je opět diferencovatelné. Následující věta říká, že „stabilita vzhledem ke skládání“ platí i pro zobrazení třídy C^p.

2.76 Věta. *Nechť $p \in \mathbb{N} \cup \{\infty\}$, $A \subset \mathbb{R}^n$, $B \subset \mathbb{R}^m$ jsou otevřené množiny, $f: A \to \mathbb{R}^m$, $g: B \to \mathbb{R}^s$ jsou třídy C^p a platí $f(A) \subset B$. Pak zobrazení $h := g \circ f$ je třídy C^p.*

Důkaz. Nechť $f = (f_1, \dots, f_m)$, $g = (g_1, \dots, g_s)$, $h = (h_1, \dots, h_s)$. Zřejmě stačí důkaz provést pro $p \in \mathbb{N}$; budeme postupovat indukcí podle p.

Buď $p = 1$. Zvolme libovolný bod $x \in A$. Podle Poznámky 2.74 existují derivace $f'(x)$, $g'(f(x))$ a tedy podle Věty 2.50 a Důsledku 2.51 existuje $h'(x)$ a pro $1 \leq i \leq s$ a $1 \leq j \leq n$ platí

$$\frac{\partial h_i}{\partial x_j}(x) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}(f(x)) \frac{\partial f_k}{\partial x_j}(x). \tag{2.17}$$

Protože všechny funkce $\dfrac{\partial g_i}{\partial y_k}$ jsou spojité na B, podle Věty 1.38 jsou složené funkce $\dfrac{\partial g_i}{\partial y_k}(f(x))$ spojité na A. Jelikož i všechny funkce $\dfrac{\partial f_k}{\partial x_j}$ jsou spojité na A, z (2.17) tedy dostáváme, že funkce $\dfrac{\partial h_i}{\partial x_j}$ jsou spojité na A. Jsou tedy $h_1, \dots, h_s$, a tudíž i zobrazení h, třídy C^1 na A.

Nyní předpokládejme, že $p > 1$ a věta platí pro hodnotu $p^* = p - 1$. Protože $g_i \in C^p(B)$ pro $1 \leq i \leq s$, máme $\dfrac{\partial g_i}{\partial y_k} \in C^{p-1}(B)$. Podle indukčního předpokladu tedy dostáváme $\dfrac{\partial g_i}{\partial y_k} \circ f \in C^{p-1}(A)$. Zřejmě také $\dfrac{\partial f_k}{\partial x_j} \in C^{p-1}(A)$. Dále si uvědomíme, že z našeho indukčního předpokladu snadno plyne, že součin i součet dvou funkcí z $C^{p-1}(A)$ leží opět v $C^{p-1}(A)$. Skutečně, jsou-li $\varphi, \psi \in C^{p-1}(A)$, pak $\varphi \cdot \psi = S \circ (\varphi, \psi)$, kde $S: \mathbb{R}^2 \to \mathbb{R}$ je zobrazení $S(x,y) = xy$. Protože je snadno vidět (viz Příklad 2.75), že $S \in C^\infty(\mathbb{R}^2)$, indukční předpoklad dává $\varphi \cdot \psi \in C^{p-1}(A)$. Případ součtu je zcela analogický. Ze vzorce (2.17) tedy dostáváme, že každá z funkcí $\dfrac{\partial h_i}{\partial x_j}$ je třídy C^{p-1} na A. Jsou tedy $h_1, \dots, h_s$ a tudíž i zobrazení h třídy C^p na A.

2.77 Poznámka. Z předchozí věty snadno vyplývá, že racionální funkce n proměnných jsou třídy C^∞ na svém definičním oboru (což je otevřená množina). Racionální funkce jsou totiž ty funkce, které lze „vyjádřit pomocí souřadnicových funkcí $x_1, \dots, x_n$, konstant, sčítání, násobení a dělení“, přičemž funkce $s(x,y) = x + y$, $S(x,y) = xy$ a $p(x) = 1/x$ jsou třídy C^∞ na svých definičních oborech.

Běžně se ovšem racionální funkce n proměnných definují jako ty funkce, které jsou podílem dvou polynomů n proměnných. Přitom polynom n proměnných je taková funkce, která je součtem konečně mnoha funkcí tvaru

$$C \cdot x_1^{\alpha_1} \cdots x_n^{\alpha_n},$$

kde $C \in \mathbb{R}$, $\alpha_i = 0, 1, \dots,$ přičemž klademe $x_i^0 = 1$. Jsou-li $P(x)$, $Q(x)$ dva takové polynomy, je tedy racionální funkce $P(x)/Q(x)$ třídy C^∞ na otevřené množině $\{x \in \mathbb{R}^n\colon\ Q(x) \neq 0\}$.

2.6 Záměnnost parciálních derivací

Následující klasický příklad ukazuje, že „záměnnost parciálních derivací“ nenastává vždy. Přesněji: při počítání parciálních derivací vyššího řádu nelze obecně zaměnit pořadí proměnných, podle kterých se derivuje.

2.78 Příklad. Nechť $f(x,y) = xy\,\frac{x^2-y^2}{x^2+y^2}$ pro $(x,y) \neq (0,0)$ a $f(0,0) = 0$. Pak

$$\frac{\partial f}{\partial x}(x,y) = y\left(\frac{x^2-y^2}{x^2+y^2} + \frac{4x^2y^2}{(x^2+y^2)^2}\right) \qquad \text{pro } (x,y) \neq (0,0)$$

a $\dfrac{\partial f}{\partial x}(0,0) = 0$ (parciální funkce $f(\cdot, 0)$ je nulová). Protože $\dfrac{\partial f}{\partial x}(0,y) = -y$ (i pro $y = 0$), je $\dfrac{\partial^2 f}{\partial y \partial x}(0,0) = -1$. Zcela obdobně dostáváme, že $\dfrac{\partial^2 f}{\partial x \partial y}(0,0) = 1$. Snadno lze ověřit, že $\dfrac{\partial^2 f}{\partial y \partial x}, \dfrac{\partial^2 f}{\partial x \partial y}$ existují a jsou si rovny v ostatních bodech roviny, ale přesto jsme dostali, že

$$\frac{\partial^2 f}{\partial y \partial x}(0,0) \neq \frac{\partial^2 f}{\partial x \partial y}(0,0).$$

Chceme tedy odvodit použitelné postačující podmínky pro záměnnost parciálních derivací. Začneme nejjednodušším případem smíšených parciálních derivací druhého řádu funkce dvou proměnných.

Než přistoupíme k formulaci a důkazu tvrzení, provedeme předběžné úvahy, které mají naznačit, *proč* často na pořadí derivování nezáleží. Výsledek těchto úvah pak dále použijeme. Uvažujme funkci dvou proměnných $f(x,y)$, bod $(a,b) \in \mathbb{R}^2$ a předpokládejme, že

(2.18) parciální derivace $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}$ existují na nějakém okolí bodu (a,b).

Podle definice

$$\frac{\partial^2 f}{\partial y \partial x}(a,b) = \lim_{k \to 0} \frac{1}{k} \left(\frac{\partial f}{\partial x}(a,b+k) - \frac{\partial f}{\partial x}(a,b) \right),$$

(má-li jedna strana smysl). Z (2.18) vyplývá existence $\delta > 0$ takového, že $\frac{\partial f}{\partial x}(a,b+k)$ existuje, pokud $|k| < \delta$. Pro taková k platí

$$\frac{\partial f}{\partial x}(a,b+k) - \frac{\partial f}{\partial x}(a,b) = \lim_{h \to 0} \frac{f(a+h,b+k) - f(a,b+k)}{h} - \lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h}$$

$$= \lim_{h \to 0} \frac{1}{h} \big((f(a+h,b+k) - f(a,b+k)) - (f(a+h,b) - f(a,b)) \big),$$

takže rovnost

$$\frac{\partial^2 f}{\partial y \partial x}(a,b) = \lim_{k \to 0} \left(\lim_{h \to 0} \frac{1}{hk} \big((f(a+h,b+k) - f(a,b+k)) - (f(a+h,b) - f(a,b)) \big) \right)$$

platí, má-li její jedna strana smysl. Zcela obdobně dostáváme

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) = \lim_{h \to 0} \left(\lim_{k \to 0} \frac{1}{hk} \big((f(a+h,b+k) - f(a+h,b)) - (f(a,b+k) - f(a,b)) \big) \right).$$

Smíšené parciální derivace jsou tedy dvojnásobné (opakované) limity téže funkce dvou proměnných; označíme-li ji

$$W(h,k) := \frac{f(a+h,b+k) + f(a,b) - f(a+h,b) - f(a,b+k)}{hk},$$

dostáváme toto tvrzení:

2.79 Lemma. *Za předpokladu* (2.18)

(i) *existuje $\delta > 0$ takové, že $\lim_{h \to 0} W(h,k)$ (resp. $\lim_{k \to 0} W(h,k)$) existuje, pokud $|k| < \delta$ (resp. $|h| < \delta$) a*

(ii) *každá z rovností*

$$\frac{\partial^2 f}{\partial y \partial x}(a,b) = \lim_{k\to 0}\left(\lim_{h\to 0} W(h,k)\right), \qquad \frac{\partial^2 f}{\partial x \partial y}(a,b) = \lim_{h\to 0}\left(\lim_{k\to 0} W(h,k)\right)$$

platí, má-li jedna její strana smysl.

2.80 Poznámka. Kdysi se místo o limitách hovořilo o nekonečně malých a nekonečně velkých veličinách. Obě parciální derivace byly v tomto pojetí rovny hodnotě $W(h,k)$ pro nekonečně malé přírůstky h,k a jejich rovnost byla považována za samozřejmou. V dnešní terminologii: tehdy se nerozlišovalo mezi dvojnou a dvojnásobnou limitou a předpokládalo se, že limity vždy komutují. Víme, že pořadí limit jde skutečně „většinou“ zaměnit, ne však vždy. Mnoho hlubokých vět z analýzy spočívá v tom, že podávají postačující podmínky pro možnost záměny dvou „limitních procesů“ (např. dvou limit, limity a derivace, limity a integrálu).

Funkce $C(h,k)$ v čitateli výrazu z definice $W(h,k)$ je tzv. diference druhého řádu funkce f v bodě (a,b) s kroky $v_1 = (h,0), v_2 = (0,k)$. Diference prvního řádu funkce f s krokem v se definuje jako funkce $\Delta_v f\colon x \mapsto f(x+v) - f(x)$. Pak $C(h,k) = \Delta_f^2(a; v_2, v_1) \ := (\Delta_{v_2}(\Delta_{v_1} f))(a)$. Při důkazu toho, že smíšené parciální derivace jsou dvojnásobné limity *téže* funkce, jsme již de facto užili rovnost $\Delta_f^2(a; v_1, v_2) = \Delta_f^2(a; v_2, v_1)$. „Důvodem“ časté záměnnosti parciálních derivací je tedy „záměnnost diferencí“.

Poznamenejme ještě, že symbol $\dfrac{\partial^2 f}{\partial x \partial y}$ původně vyjadřoval, že jde o podíl „nekonečně malé diference druhého řádu“ závisle proměnné a součinu „nekonečně malých diferencí“ nezávisle proměnných (symbol Δ se používal pro konečné diference, zatímco symboly d a ∂ pro „nekonečně malé“ diference).

2.81 Tvrzení. *Nechť $f(x,y)$ je funkce dvou proměnných, $(a,b) \in \mathbb{R}^2$. Nechť dále platí:*

(i) $\displaystyle\lim_{(x,y)\to(a,b)} \frac{\partial^2 f}{\partial y \partial x}(x,y) = A \in \mathbb{R}$.

(ii) *Parciální derivace* $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}$ *existují na nějakém okolí bodu* (a,b).

Pak v bodě (a,b) existují obě smíšené parciální derivace druhého řádu a platí

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) = \frac{\partial^2 f}{\partial y \partial x}(a,b) = A.$$

Důkaz. Označme jako výše

$$W(h,k) \ := \frac{f(a+h,b+k) + f(a,b) - f(a+h,b) - f(a,b+k)}{hk}.$$

Dokážeme-li, že

$$\lim_{(h,k)\to(0,0),\, h\neq 0,\, k\neq 0} W(h,k) = A, \tag{2.19}$$

OBR. 2.7. Znaménka + a − nad body a, $a+h$ se týkají definice $W(h,k)$; nad bodem x definice funkce φ.

budeme hotovi. Pak totiž podle Lemmatu 2.79 (i) a Věty 1.63 o dvojné a dvojnásobné limitě platí

$$\lim_{k\to 0}\left(\lim_{h\to 0} W(h,k)\right) = \lim_{h\to 0}\left(\lim_{k\to 0} W(h,k)\right) = A,$$

takže stačí použít Lemma 2.79 (ii).

Rovnost (2.19) budeme dokazovat z definice; zvolíme tedy libovolné $\varepsilon > 0$. Podle předpokladů (i), (ii) existuje $\delta > 0$ takové, že na $U_\delta^\infty(a,b)$ existují $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$ a

$$\left|\frac{\partial^2 f}{\partial y \partial x}(z) - A\right| < \varepsilon \quad \text{pro} \quad z \in U_\delta^\infty(a,b) \setminus \{(a,b)\}. \tag{2.20}$$

Uvažujme nyní libovolná reálná h, k, $0 < |h| < \delta$, $0 < |k| < \delta$ (viz obr. 2.7 pro h, $k > 0$).

Zavedeme dále pomocnou funkci $\varphi(x) := f(x, b+k) - f(x,b)$ a uvědomíme si, že $W(h,k) = \frac{1}{hk}\left(\varphi(a+h) - \varphi(a)\right)$. Podle volby δ platí pro každé $x \in (a-\delta, a+\delta)$

$$\varphi'(x) = \frac{\partial f}{\partial x}(x, b+k) - \frac{\partial f}{\partial x}(x,b).$$

Podle Lagrangeovy věty o přírůstku funkce existuje tedy číslo $\xi \in (a-\delta, a+\delta)\setminus\{a\}$ (dokonce mezi a a $a+h$) takové, že

$$\varphi(a+h) - \varphi(a) = h\varphi'(\xi) = h\left(\frac{\partial f}{\partial x}(\xi, b+k) - \frac{\partial f}{\partial x}(\xi, b)\right).$$

Dostáváme tedy $W(h,k) = \dfrac{1}{k}\left(\dfrac{\partial f}{\partial x}(\xi, b+k) - \dfrac{\partial f}{\partial x}(\xi, b)\right)$. Zaveďme ještě pomocnou funkci $\psi(y) := \dfrac{\partial f}{\partial x}(\xi, y)$. Podle volby δ pro každé $y \in (b-\delta, b+\delta)$ zřejmě existuje

$\psi'(y) = \dfrac{\partial^2 f}{\partial y \partial x}(\xi, y)$. Podle Lagrangeovy věty tedy existuje $\eta \in (b-\delta, b+\delta)$ (mezi b a $b+k$) takové, že

$$W(h,k) = \frac{\psi(b+k) - \psi(b)}{k} = \psi'(\eta) = \frac{\partial^2 f}{\partial y \partial x}(\xi, \eta).$$

Protože $(\xi, \eta) \in U_\delta^\infty(a,b) \setminus \{(a,b)\}$, podle (2.20) dostáváme $|W(h,k) - A| < \varepsilon$ a důkaz je hotov.

Jiná verze tvrzení o záměnnosti parciálních derivací je tato:

2.82 Tvrzení. *Nechť $f(x,y)$ je funkce dvou proměnných, $(a,b) \in \mathbb{R}^2$ a obě parciální derivace $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y}$ mají totální diferenciál v bodě (a,b). Potom*

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) = \frac{\partial^2 f}{\partial y \partial x}(a,b).$$

Důkaz. (Náznak) Stejně jako v důkazu Tvrzení 2.81 můžeme najít $\delta > 0$ takové, že pro libovolná reálná h, k, $0 < |h| < \delta$, $0 < |k| < \delta$ platí
$W(h,k) = \dfrac{1}{k}\left(\dfrac{\partial f}{\partial x}(\xi, b+k) - \dfrac{\partial f}{\partial x}(\xi, b)\right)$, kde bod $\xi = \xi(h,k)$ leží mezi a a $a+h$.
Protože $\dfrac{\partial f}{\partial x}$ má v (a,b) diferenciál, platí

$$\frac{\partial f}{\partial x}(a+u, b+v) = \frac{\partial f}{\partial x}(a,b) + \frac{\partial^2 f}{\partial x^2}(a,b) \cdot u + \frac{\partial^2 f}{\partial y \partial x}(a,b) \cdot v + r(u,v),$$

kde $r(u,v) = o(\|(u,v)\|)$, $(u,v) \to (0,0)$. Po dosazení máme

$$W(h,k) = \frac{\partial^2 f}{\partial y \partial x}(a,b) + \frac{1}{k}(r(\xi - a, k) - r(\xi - a, 0)).$$

Jestliže $h = k$, pak zřejmě $\|(\xi - a, k)\|_\infty \le |k|$, $\|(\xi - a, 0)\|_\infty \le |k|$, takže platí
$\lim\limits_{k \to 0} W(k,k) = \dfrac{\partial^2 f}{\partial y \partial x}(a,b)$. Postupujeme-li „symetricky“ (nebo již dokázané použijeme na funkci $f^*(x,y) := f(y,x)$), dostaneme $\lim\limits_{k \to 0} W(k,k) = \dfrac{\partial^2 f}{\partial x \partial y}(a,b)$, a tedy i dokazovanou rovnost.

2.83 Poznámka. Tvrzení 2.81 se většinou formuluje v trochu slabším tvaru: podmínka (i) je nahrazena podmínkou

$(i)^*$ funkce $\dfrac{\partial^2 f}{\partial y \partial x}$ je spojitá v bodě (a,b).

Pak ovšem platí podmínka (i) pro $A := \dfrac{\partial^2 f}{\partial y \partial x}(a,b)$.

Obecnou větu již budeme formulovat v příslušné běžnější formě.

2.84 Věta. *Nechť f je funkce n proměnných, $c \in \mathbb{R}^n$, $1 \le i, j \le n$, $i \ne j$ a platí:*

(i) *Funkce* $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ *je spojitá v bodě c.*

(ii) *Funkce* $\dfrac{\partial f}{\partial x_j}$ *je definovaná na nějakém okolí bodu c.*

Pak existuje $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(c)$ a $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(c) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(c)$.

Důkaz. Nejprve si uvědomíme, že z předpokladu (i) okamžitě plyne, že funkce $\frac{\partial f}{\partial x_i}$ je definovaná na nějakém okolí bodu c. Položme $a := c_i$, $b := c_j$ a definujme (parciální) funkci dvou proměnných

$$g(x,y) := f(c + (x - c_i)e_i + (y - c_j)e_j);$$

sk

v případě, že $i < j$, je obvyklé psát

$$g(x,y) = f(c_1, \ldots, c_{i-1}, x, c_{i+1}, \ldots, c_{j-1}, y, c_{j+1}, \ldots, c_n).$$

Označme $z(x,y) := c + (x - c_i)e_i + (y - c_j)e_j$. Je snadné dokázat, že parciální derivace

$$\frac{\partial g}{\partial x}(x,y), \quad \frac{\partial g}{\partial y}(x,y), \quad \frac{\partial^2 g}{\partial x \partial y}(x,y), \quad \frac{\partial^2 g}{\partial y \partial x}(x,y)$$

jsou (co do existence i hodnoty) totéž, co

$$\frac{\partial f}{\partial x_i}(z(x,y)), \quad \frac{\partial f}{\partial x_j}(z(x,y)), \quad \frac{\partial^2 f}{\partial x_i \partial x_j}(z(x,y)), \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(z(x,y)).$$

Z toho pak snadno vyplývá, že $\frac{\partial g}{\partial x}$, $\frac{\partial g}{\partial y}$ existují na nějakém okolí bodu (a,b) a funkce $\dfrac{\partial^2 g}{\partial y \partial x}$ je spojitá v bodě (a,b). Z Tvrzení 2.81 (viz Poznámka 2.83) již okamžitě plyne tvrzení naší věty.

Stejným postupem jako v předchozím důkazu lze z Tvrzení 2.82 snadno odvodit druhou verzi věty o záměnnosti parciálních derivací:

2.85 Věta. *Nechť f je funkce n proměnných, $c \in \mathbb{R}^n$ a $1 \le i, j \le n$. Nechť obě funkce* $\dfrac{\partial f}{\partial x_i}$, $\dfrac{\partial f}{\partial x_j}$ *mají totální diferenciál v bodě c. Pak* $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(c) = \dfrac{\partial^2 f}{\partial x_j \partial x_i}(c)$.

Jako důsledek Věty 2.84 dostáváme snadno tuto větu o záměnnosti parciálních derivací vyšších řádů.

2.86 Věta. *Nechť funkce f je třídy $C^k(G)$ $(k \ge 2)$, kde $G \subset \mathbb{R}^n$ je otevřená množina a nechť $x \in G$. Pak hodnota*

$$\frac{\partial^k f}{\partial x_{i_k} \ldots \partial x_{i_1}}(x) = f_{i_k, \ldots, i_1}(x)$$

nezávisí na pořadí indexů $i_1, \dots, i_k \in \{1, \dots, n\}$.

Důkaz. Zcela přesně řečeno, máme dokázat, že pokud $(j_1, \dots, j_k)$ je taková k-tice, kterou lze dostat z k-tice $(i_1, \dots, i_k)$ přerovnáním, pak

$$f_{j_k,\dots,j_1}(x) = f_{i_k,\dots,i_1}(x). \tag{2.21}$$

Pro $n = 3$ a $k = 5$ má například platit $f_{1,2,3,3,2}(x) = f_{2,2,3,1,3}(x)$. To, že $(j_1, \dots, j_k)$ lze dostat z $(i_1, \dots, i_k)$ přerovnáním, lze formálně vyjádřit takto: existuje permutace (tj. bijekce) $\pi\colon \{1, \dots, k\} \to \{1, \dots, k\}$ taková, že $i_1 = j_{\pi(1)}, \dots, i_k = j_{\pi(k)}$.

Protože libovolná permutace je složením konečně mnoha „sousedních transpozic“ (srov. Poznámka 2.87 (ii)), stačí dokázat, že pro každé $p \in \{1, \dots, k-1\}$ platí

$$f_{i_k,\dots,i_{p+2},i_{p+1},i_p,i_{p-1},\dots,i_1} = f_{i_k,\dots,i_{p+2},i_p,i_{p+1},i_{p-1},\dots,i_1}$$

na G. (Zde jde o běžnou licenci – zápis používáme například i pro $p = 1$, kdy nemá, přesně vzato, smysl. Jde jen o názorný zápis tvrzení, že rovnost (2.21) platí, pokud $1 \le p \le k-1$, $j_p = i_{p+1}$, $j_{p+1} = i_p$ a $j_s = i_s$ pro $s \in \{1, \dots, k\} \setminus \{p, p+1\}$.)

Položme $g := f_{i_{p-1},\dots,i_1}$ pro $p > 1$ a $g := f$ pro $p = 1$. Pak zřejmě $g \in C^2(G)$, a proto podle Věty 2.84 platí $g_{i_p,i_{p+1}}(y) = g_{i_{p+1},i_p}(y)$ pro každé $y \in G$. Na G tedy platí $f_{i_{p+1},i_p,i_{p-1},\dots,1} = f_{i_p,i_{p+1},i_{p-1},\dots,1}$ a tudíž také

$$\begin{aligned} f_{i_k,\dots,i_{p+2},i_{p+1},i_p,i_{p-1},\dots,i_1} &= \left(f_{i_{p+1},i_p,i_{p-1},\dots,i_1}\right)_{i_k,\dots,i_{p+2}} \\ &= \left(f_{i_p,i_{p+1},i_{p-1},\dots,i_1}\right)_{i_k,\dots,i_{p+2}} = f_{i_k,\dots,i_{p+2},i_p,i_{p+1},i_{p-1},\dots,i_1}. \end{aligned}$$

2.87 Poznámka.

(i) V předchozí větě lze předpoklad $f \in C^k(G)$ zeslabit: stačí předpokládat, že f má v bodě x derivaci k-tého řádu (pro definici viz str. 89 a Poznámka 2.94 níže). Důkaz (viz D II, Věta 195) je zcela obdobný důkazu předchozí věty, ale vychází z Věty 2.85.

(ii) Transpozici čísel $\{1, \dots, k\}$, která vyměňuje prvky $i \ne j$, označme $T_{i,j}$; pokud $|j - i| = 1$, nazveme $T_{i,j}$ sousední transpozicí. Každému, kdo chvíli experimentoval se sousedními transpozicemi, je jasné, že každá permutace je složením konečně mnoha sousedních transpozic (a uměl by to asi indukcí dokázat). Víme-li již, že že každá permutace je složením konečně mnoha transpozic ([Be], [Bi]), plyne naše tvrzení ihned z toho, že pro $1 \le i < j \le k$ je zřejmě $T_{i,j} = T_{j-1,j} \circ \cdots \circ T_{i+1,i+2} \circ T_{i,i+1}$.

Jinými slovy můžeme obsah Věty 2.86 vyjádřit takto: Je-li $f \in C^k(G)$, $G \subset \mathbb{R}^n$, pak hodnota parciálních derivací řádu k záleží jen na tom, kolikrát se derivuje podle x_1, kolikrát podle x_2 atd., nikoliv však na pořadí, ve kterém se derivuje.

Každý parciální diferenciální operátor k-tého řádu $f \mapsto \dfrac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}$ chápaný jako zobrazení $C^k(G) \to C(G)$ lze tedy psát ve tvaru

$$f \mapsto \frac{\partial^k f}{(\partial x_1)^{\alpha_1} \dots (\partial x_n)^{\alpha_n}}, \quad 0 \le \alpha_i, \quad \sum_{I=1}^{n} \alpha_i = k. \tag{2.22}$$

(Pokud $\alpha_i = 0$, znamená to, že se podle proměnné x_i nederivuje.)

Uspořádaná n-tice celých nezáporných čísel $\alpha := (\alpha_1, \dots, \alpha_n)$ se nazývá multiindex a diferenciální operátor tvaru (2.22) příslušný multiindexu α se označuje symbolem D^α. Číslo $\sum_{i=1}^n \alpha_i$ se nazývá výška multiindexu α.

2.88 Poznámka. Z kombinatoriky je známo, že počet multiindexů $(\alpha_1, \dots, \alpha_n)$ výšky k je $\binom{n+k-1}{n-1}$. Není těžké ověřit, že $D^\alpha \neq D^\beta$, pokud $\alpha \neq \beta$. Počet různých parciálních operátorů tvaru $f \mapsto \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}}$ na $C^k(G)$, kde $G \subset \mathbb{R}^n$ je neprázdná otevřená množina, je tedy také $\binom{n+k-1}{n-1}$.

2.7 Diferenciály a derivace vyššího řádu a Taylorova věta

Poměrně obtížný pojem derivace (diferenciálu) vyššího řádu je důležitý v řadě teorií.

Pro základní teorii funkcí n proměnných tento pojem již tak důležitý není. Proto jej probereme jen stručně, přičemž se ve větách o derivaci k-tého řádu omezíme na nejdůležitější speciální případ funkcí třídy C^k, který je podstatně snažší. (Podrobný rozbor obecného případu lze nalézt v [D II].)

Derivace a diferenciál prvního řádu reálné funkce jsou jen dva různé názvy pro jeden objekt – lineární formu „dobře aproximující“ přírůstek funkce. Derivace a diferenciál druhého řádu však již pro nás budou dva *různé* objekty.

Druhá *derivace* $f''(a)$ bude mít obvyklý moderní význam – bude to jistá *bilineární forma* na $\mathbb{R}^n$.

Druhý *diferenciál* $\mathrm{d}^2 f(a)$ však zde definujeme klasicky – jako *kvadratickou formu* (příslušnou bilineární formě $f''(a)$).

Pro přesnou definici těchto pojmů použijeme klasický zápis hodnoty diferenciálu: má-li funkce f v bodě $a \in \mathbb{R}^n$ diferenciál a $h = (h_1, \dots, h_n) \in \mathbb{R}^n$, pak hodnotu $\mathrm{d}f(a)(h)$ budeme zapisovat jako $\mathrm{d}_h f(a)$.

Při pevném h bude symbol d_h označovat diferenciální operátor, který funkci f přiřazuje funkci $\mathrm{d}_h f \colon x \mapsto \mathrm{d}_h f(x)$. Funkce $\mathrm{d}_h f$ je ovšem definovaná právě v těch bodech x, ve kterých má f diferenciál („je diferencovatelná“). V těchto bodech se podle Věty 2.28 hodnota $\mathrm{d}_h f(x)$ rovná derivaci $D_h f(x)$ podle vektoru h, která ale může být definovaná i v jiných bodech. Je-li x bodem diferencovatelnosti funkce f, můžeme také psát

$$\mathrm{d}_h f(x) = h_1 \frac{\partial f}{\partial x_1}(x) + \dots + h_n \frac{\partial f}{\partial x_n}(x).$$

Může se ovšem stát, že výraz napravo má smysl, ale výraz nalevo není definován.

Vidíme tedy, že mezi operátorem d_h a parciálními diferenciálními operátory je úzký vztah. Definujeme-li přirozeným způsobem součet operátorů a násobení operátoru reálným číslem, například

$$\left(\frac{\partial}{\partial x} + \frac{\partial}{\partial y}\right) f \; := \frac{\partial}{\partial x} f + \frac{\partial}{\partial y} f,$$

kde pravou stranu chápeme jako součet funkcí, *neplatí* sice obecně rovnost operátorů

$$\mathrm{d}_h = h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n},$$

oba operátory však přiřazují tutéž funkci každé funkci, která je diferencovatelná na svém definičním oboru.

2.89 Příklad. Nechť $f(x,y) = x^2 y$ a $h = (2,3)$. Pak $f_x(x,y) = 2xy$, $f_y(x,y) = x^2$. Obě parciální derivace jsou spojité na $\mathbb{R}^2$, a proto je f všude diferencovatelná. V každém bodě (x,y) platí $\mathrm{d}_h f(x,y) = 2 \cdot 2xy + 3x^2$. Operátor d_h funkci $x^2 y$ přiřazuje funkci $4xy + 3x^2$.

Nyní již můžeme definovat derivaci a diferenciál řádu k funkce n proměnných; pro názornost začneme případem $k = 2$.

(i) Řekneme, že funkce f má v bodě $a \in \mathbb{R}^n$ druhou derivaci $f''(a)$, jestliže funkce $\mathrm{d}_h(\mathrm{d}_k f) = (\mathrm{d}_h \circ \mathrm{d}_k) f$ je definovaná v bodě a pro každou dvojici $h \in \mathbb{R}^n$, $k \in \mathbb{R}^n$. Druhá derivace funkce f v bodě a je pak funkce na $\mathbb{R}^n \times \mathbb{R}^n$ definovaná předpisem

$$f''(a)\colon (h,k) \mapsto (\mathrm{d}_h(\mathrm{d}_k f))(a).$$

(ii) Existuje-li $f''(a)$, pak definujeme druhý diferenciál funkce f v bodě a jako funkci $\mathrm{d}^2 f(a)$ na $\mathbb{R}^n$ definovanou předpisem $\mathrm{d}^2 f(a)\colon\ h \mapsto (\mathrm{d}_h(\mathrm{d}_h f))(a)$, tj.

$$\mathrm{d}^2 f(a)(h) \; := f''(h,h).$$

Místo $\mathrm{d}^2 f(a)(h)$ se často píše $\mathrm{d}_h^2 f(a)$.

(iii) Obdobně pro libovolné přirozené k definujeme k-tou derivaci $f^{(k)}(a)$ funkce f v bodě $a \in \mathbb{R}^n$ jako funkci na $(\mathbb{R}^n)^k$ definovanou předpisem $f^{(k)}(a)\,(v_1,\ldots,v_k) = (\mathrm{d}_{v_1} \circ \mathrm{d}_{v_2} \circ \cdots \circ \mathrm{d}_{v_k}) f\,(a)$ pro ty body a, ve kterých má pravá strana smysl pro všechny k-tice vektorů $(v_1,\ldots,v_k) \in (\mathbb{R}^n)^k$. Někdy se také píše $f^{(k)}(a; v_1,\ldots,v_k)$ místo $f^{(k)}(a)\,(v_1,\ldots,v_k)$.

(iv) Existuje-li $f^{(k)}(a)$, pak definujeme k-tý diferenciál funkce f v bodě a jako funkci $\mathrm{d}^k f(a)$ na $\mathbb{R}^n$ definovanou předpisem

$$\mathrm{d}^k f(a)(h) = f^{(k)}(h,\ldots,h).$$

Místo $\mathrm{d}^k f(a)(h)$ se také píše $\mathrm{d}_h^k f(a)$.

(v) Symboly $f^{(k)}$ (resp. $\mathrm{d}^k f$) budeme rozumět zobrazení, které bodům $x \in \mathbb{R}^n$ přiřazuje funkci $f^{(k)}(x)$ (resp. $\mathrm{d}^k f(x)$), jsou-li ovšem tyto funkce definovány. Definičním oborem $f^{(k)}$ (a také $\mathrm{d}^k f$) je tedy množina těch $x \in \mathbb{R}^n$, ve kterých existuje $f^{(k)}(x)$. (To je modernější pojetí; klasické pojetí bylo formálně jiné.)

2.90 Poznámka. Často se užívá symbolika, při které se k-násobné složení $A \circ \cdots \circ A$ téhož operátoru (zobrazení) A označuje symbolem A^k. Existuje-li tedy $\mathrm{d}^k f(a)$, při použití tohoto „operátorového“ zápisu platí rovnost

$$\mathrm{d}_h^k f(a) = \big((\mathrm{d}_h)^k f\big)\ (a) = \left(\left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n}\right)^k f\right)(a). \tag{2.23}$$

Složitost pojmů derivace a diferenciálu řádu $k > 1$ spočívá podstatně v trochu komplikovaném popisu bodů, ve kterých jsou $f^{(k)}$ a $\mathrm{d}^k f$ definovány. (Tyto body popíšeme bez důkazu v Poznámce 2.94.) Omezíme se proto na nejdůležitější případ funkcí třídy C^k, kdy tyto problémy nenastávají.

V dalším najdeme vyjádření derivace a diferenciálu vyššího řádu pomocí parciálních derivací vyššího řádu. Nejdříve probereme zvlášť nejjednodušší (a také nejdůležitější) případ druhé derivace a diferenciálu. Uvažujme otevřenou množinu $G \subset \mathbb{R}^n$, funkci $f \in C^2(G)$ a vektory $h = (h_1, \ldots, h_n)$, $k = (k_1, \ldots, k_n)$. Pak $\mathrm{d}_k f$ je funkce definovaná na celém G a

$$\mathrm{d}_k f = \left(k_1 \frac{\partial}{\partial x_1} + \cdots + k_n \frac{\partial}{\partial x_n}\right) f = k_1 \frac{\partial f}{\partial x_1} + \cdots + k_n \frac{\partial f}{\partial x_n}.$$

Na pravé straně je tedy funkce třídy C^1 na G, a proto existuje

$$\begin{aligned}\mathrm{d}_h(\mathrm{d}_k f) &= \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n}\right) \circ \left(k_1 \frac{\partial}{\partial x_1} + \cdots + k_n \frac{\partial}{\partial x_n}\right) f \\ &= \sum_{i=1}^n h_i \frac{\partial}{\partial x_i}\left(\sum_{j=1}^n k_j \frac{\partial f}{\partial x_j}\right) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} h_i k_j.\end{aligned}$$

V každém bodě $a \in G$ tedy existuje $f''(a)$ a je to bilineární forma na $\mathbb{R}^n$ daná maticí $\left(\dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)\right)_{i,j=1}^n$; v maticovém zápisu máme

$$f''(a)(h,k) = (h_1, \ldots, h_n) \cdot \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{bmatrix} \cdot (k_1, \ldots, k_n)^T.$$

Existuje tedy i druhý diferenciál – kvadratická forma $\mathrm{d}^2 f(a)$ zadaná předpisem

$$\mathrm{d}^2 f(a)(h) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j. \tag{2.24}$$

Matice bilineární formy $f''(a)$ (a zároveň kvadratické formy $\mathrm{d}^2 f(a)$) se nazývá *Hessova matice* (a její determinant *Hessův determinant* nebo také hessián). Protože $f \in C^2(G)$ a $a \in G$, z Věty 2.84 o záměnnosti parciálních derivací dostáváme, že Hessova matice je symetrická, takže $f''(a)$ je symetrická bilineární forma. Při použití časté dohody o významu symbolu $\mathrm{d}x_i$ (viz Poznámka 2.21; $\mathrm{d}x_i = L_i$) můžeme (2.24) přepsat jako rovnost kvadratických forem:

$$\mathrm{d}^2 f(a) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a) \mathrm{d}x_i \mathrm{d}x_j. \tag{2.25}$$

2.91 Poznámka.

(i) Druhá derivace $f''(a)$ je *vždy* (existuje-li) symetrická bilineární forma. Pokud totiž $f''(a)$ existuje, pak pro $h := e_k$ $(1 \le k \le n)$ má $\mathrm{d}_h f = \dfrac{\partial f}{\partial x_k}$ diferenciál v bodě a, takže podle Věty 2.85 dostáváme, že parciální derivace druhého řádu funkce f v bodě a jsou záměnné.

(ii) Kvůli kolísání zápisu derivací druhého řádu se Hessova matice často uvádí v transponovaném tvaru; vzhledem ke své symetrii je to táž matice.

Výsledek o druhé derivaci nyní snadno zobecníme na případ k-té derivace.

2.92 Věta. *Nechť $k \in \mathbb{N}$, $G \subset \mathbb{R}^n$ je otevřená množina a $f \in C^k(G)$. Pak $f^{(k)}(x)$ a $\mathrm{d}^k f(x)$ existují v každém bodě $x \in G$, $f^{(k)}(x)$ je symetrická k-lineární forma na $\mathbb{R}^n$ a kdykoliv*

$$v^{(1)} = (v_1^{(1)}, \dots, v_n^{(1)}), \dots, v^{(k)} = (v_1^{(k)}, \dots, v_n^{(k)}) \quad \text{a} \quad h = (h_1, \dots, h_n)$$

jsou vektory z $\mathbb{R}^n$, pak

$$f^{(k)}(x)\,(v^{(1)}, \dots, v^{(k)}) = \sum_{i_1, \dots, i_k = 1}^{n} \frac{\partial^k f(x)}{\partial x_{i_1} \dots \partial x_{i_k}} v_{i_1}^{(1)} \cdot v_{i_2}^{(2)} \cdots v_{i_k}^{(k)}, \tag{2.26}$$

$$\mathrm{d}_h^k f(x) = \sum_{i_1, \dots, i_k = 1}^{n} \frac{\partial^k f(x)}{\partial x_{i_1} \dots \partial x_{i_k}} h_{i_1} \cdot h_{i_2} \cdots h_{i_k}. \tag{2.27}$$

Důkaz. Budeme postupovat indukcí podle k. Pro $k = 1$ je tvrzení okamžitým důsledkem Věty 2.20 a Věty 2.10. Dále předpokládejme, že $k > 1$ a věta platí pro hodnotu $k^* = k - 1$. Víme tedy, že pro $x \in G$ platí

$$f^{(k-1)}(x)(v^{(2)}, \dots, v^{(k)}) = (\mathrm{d}_{v^{(2)}} \circ \dots \circ \mathrm{d}_{v^{(k)}}) f(x) = \sum_{i_2,\dots,i_k=1}^{n} \frac{\partial^{k-1} f(x)}{\partial x_{i_2} \dots \partial x_{i_k}} v_{i_2}^{(2)} \cdots v_{i_k}^{(k)},$$

takže s pomocí toho, že $f \in C^k(G)$, Věty 2.20 a Věty 2.10 dostáváme

$$\begin{aligned}(\mathrm{d}_{v^{(1)}} \circ \mathrm{d}_{v^2} \circ \dots \circ \mathrm{d}_{v^{(k)}}) f &= \mathrm{d}_{v^{(1)}} \left(\sum_{i_2,\dots,i_k=1}^{n} \frac{\partial^{k-1} f}{\partial x_{i_2} \dots \partial x_{i_k}} v_{i_2}^{(2)} \cdots v_{i_k}^{(k)} \right) \\ &= \sum_{i_1=1}^{n} v_{i_1}^{(1)} \frac{\partial}{\partial x_{i_1}} \left(\sum_{i_2,\dots,i_k=1}^{n} \frac{\partial^{k-1} f}{\partial x_{i_2} \dots \partial x_{i_k}} v_{i_2}^{(2)} \cdots v_{i_k}^{(k)} \right) \\ &= \sum_{i_1,\dots,i_k=1}^{n} \frac{\partial^{k} f}{\partial x_{i_1} \dots \partial x_{i_k}} v_{i_1}^{(1)} \cdot v_{i_2}^{(2)} \cdots v_{i_k}^{(k)}.\end{aligned}$$

Po dosazení x do obou stran rovnosti dostáváme vzorec pro $f^{(k)}(x)(v^{(1)}, \dots, v^{(k)})$ z tvrzení věty. Z něho pak okamžitě plyne vzorec pro $\mathrm{d}_h^k f(x)$. To, že $f^{(k)}(x)$ je symetrická k-lineární forma na $\mathbb{R}^n$, plyne snadno z dokázaného vzorce pro $f^{(k)}(x)$ a z Věty 2.86 o záměně parciálních derivací vyššího řádu.

Není těžké se přesvědčit, že výpočty z předchozího důkazu ukazují i to, že vzorce (2.26), (2.27) platí, kdykoliv má jejich levá strana smysl.

2.93 Poznámka. Vzorec (2.27) lze upravit. S pomocí Poznámky 2.88 (a textu před ní) snadno vidíme, že každý sčítanec z (2.27) je roven hodnotě ($h_i^0 := 1$)

$$\frac{\partial^k f(x)}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} \cdot h_1^{k_1} \cdots h_n^{k_n} \tag{2.28}$$

pro jistý multiindex $(k_1, \dots, k_n)$ takový, že $k_1 \geq 0, \dots k_n \geq 0$ a $k_1 + \dots k_n = k$. Poměrně snadná kombinatorická úvaha ukazuje, že hodnota (2.28) se rovná $\dfrac{k!}{k_1! \cdots k_n!}$ sčítancům v (2.27), takže (2.27) lze psát ve tvaru

$$\mathrm{d}_h^k f(x) = \sum_{k_1 \geq 0, \dots k_n \geq 0, k_1 + \dots k_n = k} \frac{k!}{k_1! \cdots k_n!} \frac{\partial^k f(a)}{\partial x_1^{k_1} \dots \partial x_n^{k_n}} \cdot h_1^{k_1} \cdots h_n^{k_n},$$

ve kterém je $\binom{n+k-1}{n-1}$ sčítanců (viz Poznámka 2.88). Stejný vzorec ovšem dostaneme, pokud pravou stranu rovnosti (viz Poznámka 2.90)

$$\mathrm{d}_h^k f(x) = \left(h_1 \frac{\partial}{\partial x_1} + \dots + h_n \frac{\partial}{\partial x_n} \right)^k f(x)$$

„formálně umocníme", přičemž s operátory (symboly) $\dfrac{\partial}{\partial x_i}$ pracujeme, jako by šlo o reálná čísla (a součin operátorů interpretujeme jako skládání operátorů). To je intuitivně „jasné"; pro formální důkaz lze užít polynomickou větu (která dává vzorec pro $(a_1+\cdots+a_n)^k$). Obecná teorie, ve které se s operátory počítá jako s reálnými čísly, se nazývá operátorový (dříve symbolický) počet.

2.94 Poznámka. (o bodech existence $f^{(k)}$ a $\mathrm{d}^k f$) Není těžké dokázat, že $f^{(k)}(a)$ existuje, právě když všechny parciální derivace funkce f až do řádu $k-2$ mají totální diferenciál na nějakém okolí bodu a a všechny parciální derivace funkce f řádu $k-1$ mají totální diferenciál v bodě a. (Předpokládá se ovšem, že $k \geq 2$ a parciální derivací řádu 0 se rozumí funkce f.)

Jako důsledek dostáváme, že pokud $f^{(k)}(a)$ existuje, pak f má v bodě a všechny parciální derivace až do řádu k a je třídy C^{k-2} na nějakém okolí bodu a.

Podle *naší definice* diferenciál $\mathrm{d}^k f(a)$ existuje, právě když existuje derivace $f^{(k)}(a)$. To platí i tehdy, pokud diferenciál $\mathrm{d}^k f(a)$ definujeme (přirozeněji) právě tehdy, když funkce $(\mathrm{d}_h \circ \cdots \circ \mathrm{d}_h)f$ (skládá se k-krát) je definována v bodě a pro každý vektor $h \in \mathbb{R}^k$. Důkaz ([D II; Věta 200]) není snadný.

Diferenciály vyššího řádu budeme potřebovat pouze v důkazu Taylorovy věty pro funkce více proměnných a při vyšetřování lokálních extrémů, kde hraje zásadní roli diferenciál druhého řádu.

Přitom budeme pracovat s diferenciály vyššího řádu složené funkce; ovšem jen ve dvou velmi speciálních případech, které nyní vyšetříme. (Pro obecný případ viz Příklad 3.44.)

2.95 Tvrzení. *Nechť $p \in \mathbb{N}$, $G \subset \mathbb{R}^n$ je otevřená množina a nechť f je reálná funkce třídy $C^p(G)$. Nechť dále $a \in \mathbb{R}^n$, $h = (h_1, \ldots, h_n) \in \mathbb{R}^n$ a $\alpha < \beta$ jsou taková reálná čísla, že $a + th \in G$ pro každé $t \in (\alpha, \beta)$. Pak funkce $g(t) := f(a+th)$ má v každém bodě $t \in (\alpha, \beta)$ derivaci řádu p a platí $g^{(p)}(t) = \mathrm{d}_h^p f(a+th)$.*

Důkaz. Důkaz provedeme indukcí podle p. Je-li $p = 1$, můžeme na složenou funkci $g(t) = f(a_1 + th_1, \ldots, a_n + th_n)$ použít Větu 2.54, takže pro všechna $t \in (\alpha, \beta)$ dostáváme

$$g'(t) = \frac{\partial f}{\partial x_1}(a+th) \cdot h_1 + \cdots + \frac{\partial f}{\partial x_n}(a+th) \cdot h_n = \mathrm{d}_h f(a+th).$$

Nyní předpokládejme, že $p > 1$ a tvrzení „platí pro $p^* = p-1$". Položíme-li $f^* := \mathrm{d}_h^{p-1} f$, pak z vyjádření diferenciálu (zde řádu $p-1$) pomocí parciálních derivací (Věta 2.92) vidíme, že $f^* \in C^1(G)$. Označíme-li $g^* := f^*(a+th)$, indukční předpoklad dává $g^*(t) = g^{(p-1)}(t)$ pro $t \in (\alpha, \beta)$). Užijeme-li naše tvrzení na f^* pro (již dokázaný) případ $p = 1$, dostáváme

$$g^{(p)}(t) = (g^*)'(t) = \mathrm{d}_h f^*(a+th) = \mathrm{d}_h(\mathrm{d}_h^{p-1} f)(a+th) = \mathrm{d}_h^p f(a+th).$$

2.96 Tvrzení. *Nechť $H \subset \mathbb{R}^k$, $G \subset \mathbb{R}^n$ jsou otevřené množiny, $f: G \to \mathbb{R}$ a $\varphi = (\varphi_1, \ldots, \varphi_n): H \to \mathbb{R}^n$ jsou třídy C^2 a platí $\varphi(H) \subset G$. Nechť $a \in H$ a předpokládejme, že bod $b := \varphi(a)$ je stacionárním bodem funkce f. Pak pro složenou* *funkci $g := f \circ \varphi$ (tj. $g(t) = f(\varphi_1(t), \ldots, \varphi_n(t))$) platí rovnost kvadratických forem*

$$\mathrm{d}^2 g(a) = \sum_{p,q=1}^{n} \frac{\partial^2 f}{\partial x_p \partial x_q}(b)\, d\varphi_p(a)\, d\varphi_q(a).$$

Důkaz. Podle Věty 2.76 je funkce g třídy C^2 na H, takže podle Věty 2.92 $\mathrm{d}^2 g(a)$ existuje a pro každé $h = (h_1, \ldots, h_k) \in \mathbb{R}^k$ platí

$$\mathrm{d}_h^2 g(a) = \sum_{i,j=1}^{k} \frac{\partial^2 g}{\partial t_i \partial t_j}(a)\, h_i h_j.$$

Podle Věty 2.54 o derivaci složené funkce a vzorce pro parciální derivaci součtu a součinu funkcí (viz Poznámka 2.4) dostáváme pro $i, j = 1, \ldots, k$ a $t \in H$ rovnosti

$$\frac{\partial g}{\partial t_j}(t) = \sum_{q=1}^{n} \frac{\partial f}{\partial x_q}(\varphi_1(t), \ldots, \varphi_n(t))\, \frac{\partial \varphi_q}{\partial t_j}(t),$$

$$\frac{\partial^2 g}{\partial t_i \partial t_j}(t) = \sum_{q=1}^{n} \left(\sum_{p=1}^{n} \frac{\partial^2 f}{\partial x_p \partial x_q}(\varphi(t))\, \frac{\partial \varphi_p}{\partial t_i}(t)\, \frac{\partial \varphi_q}{\partial t_j}(t) \; + \frac{\partial f}{\partial x_q}(\varphi(t))\, \frac{\partial^2 \varphi_q}{\partial t_i \partial t_j}(t) \right).$$

Protože předpokládáme, že $b = \varphi(a)$ je stacionární bod f, dostáváme

$$\frac{\partial^2 g}{\partial t_i \partial t_j}(a) = \sum_{p,q=1}^{n} \frac{\partial^2 f}{\partial x_p \partial x_q}(b)\, \frac{\partial \varphi_p}{\partial t_i}(a)\, \frac{\partial \varphi_q}{\partial t_j}(a),$$

a tedy také

$$\begin{aligned}
\mathrm{d}_h^2 g(a) &= \sum_{i,j=1}^{k} \sum_{p,q=1}^{n} \left(\frac{\partial^2 f}{\partial x_p \partial x_q}(b)\, \frac{\partial \varphi_p}{\partial t_i}(a)\, \frac{\partial \varphi_q}{\partial t_j}(a) \right) h_i h_j \\
&= \sum_{p,q=1}^{n} \frac{\partial^2 f}{\partial x_p \partial x_q}(b) \left(\sum_{i=1}^{k} \frac{\partial \varphi_p}{\partial t_i}(a) h_i \right) \left(\sum_{j=1}^{k} \frac{\partial \varphi_q}{\partial t_j}(a) h_j \right).
\end{aligned}$$

Tím je důkaz proveden.

2.97 Poznámka. Ukázali jsme, že ve velmi speciálním případě platí jakási „invariantnost formy druhého diferenciálu“ (srov. (2.25)). Z důkazu je také patrno, že obecně tato „invariantnost“ neplatí (srov. Příklad 3.44).

2.98 Věta. (Taylorova věta s Lagrangeovým tvarem zbytku) *Nechť jsou dána celá čísla $p \geq 0$, $n \geq 0$, otevřená množina $G \subset \mathbb{R}^p$ a funkce $f \in C^{n+1}(G)$. Nechť*

$a \in G, x \in G, a \neq x$ *a nechť* G *obsahuje uzavřenou úsečku* $\overline{ax}$. *Pak existuje bod* ξ *tvaru* $\xi = a + \theta(x-a)$, $\theta \in (0,1)$, *takový, že*

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{1}{k!} \mathrm{d}^k_{x-a} f(a) \; + \; \frac{1}{(n+1)!} d^{n+1}_{x-a} f(\xi).$$

Položíme-li $h \; := x - a$, *můžeme ekvivalentně psát*

$$\begin{aligned} f(a+h) = f(a) + \frac{1}{1!}\left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_p \frac{\partial}{\partial x_p}\right) f\,(a) + \ldots \\ + \frac{1}{n!}\left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_p \frac{\partial}{\partial x_p}\right)^n f\,(a) + \\ + \frac{1}{(n+1)!}\left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_p \frac{\partial}{\partial x_p}\right)^{n+1} f\,(a+\theta h). \end{aligned}$$

Důkaz. Položme $g(t) \; := f(a+th)$. Protože zobrazení $\varphi\colon t \mapsto a + th$ je spojité, $\varphi(0) = a$, $\varphi(1) = x$ a G je otevřená, zřejmě existuje interval $(\alpha, \beta) \supset [0,1]$ takový, že $a + th = \varphi(t) \in G$ pro každé $t \in (\alpha, \beta)$. Z Tvrzení 2.95 dostáváme, že pro každé $k \in \{1, \ldots, n+1\}$ a $t \in (\alpha, \beta)$ platí

$$g^{(k)}(t) = \mathrm{d}^k_h f(a+th). \tag{2.29}$$

Podle Taylorovy věty s Lagrangeovým tvarem zbytku pro funkce jedné proměnné (viz [D I], Věta 153) existuje číslo $\theta \in (0,1)$, pro které

$$g(1) = g(0) + \sum_{k=1}^{n} \frac{g^{(k)}(0)}{k!} + \frac{g^{(k+1)}(\theta)}{(k+1)!}.$$

Protože $g(0) = f(a)$ a $g(1) = f(x)$, podle (2.29) platí

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{1}{k!} \mathrm{d}^k_{x-a} f(a) \; + \; \frac{1}{(n+1)!} d^{n+1}_{x-a} f(a+\theta h),$$

takže stačí položit $\xi \; := a + \theta h$.

Polynom p proměnných

$$T^{f,a}_n(x) \; := f(a) + \sum_{k=1}^{n} \frac{1}{k!} \mathrm{d}^k_{x-a} f(a)$$

budeme nazývat *Taylorův polynom* n*-tého řádu funkce* f *v bodě* a. Při užití této terminologie má vzorec z Věty 2.98 tvar

$$f(x) = T^{f,a}_n(x) + \frac{1}{(n+1)!} d^{n+1}_{x-a} f(\xi).$$

2.99 Poznámka. Někdy se definuje $\mathrm{d}_h^0 f(a) := f(a)$; při této úmluvě lze psát $T_n^{f,a}(x) = \sum_{k=0}^{n} \frac{1}{k!}\mathrm{d}_{x-a}^k f(a)$.

Následující věta ukazuje, že Taylorův polynom $T_n^{f,a}$ velmi dobře aproximuje funkci f v blízkosti bodu a, je-li f na nějakém okolí bodu a dostatečně hladká.

2.100 Věta. (Taylorova věta s Peanovým tvarem zbytku) *Je-li funkce f třídy C^n $(n \geq 1)$ na nějakém okolí bodu $a \in \mathbb{R}^p$, pak*

$$f(x) = T_n^{f,a}(x) + o\left(\|x-a\|^n\right), \quad x \to a.$$

Důkaz. Zvolme $\delta > 0$ tak malé, aby funkce f byla třídy C^n na $G := U_\delta(a)$. Užijeme-li předchozí větu pro „$n^* = n-1$“, dostáváme, že pro každé $x \in G$ existuje bod $\xi = \xi(x) \in \overline{ax} \setminus \{a, x\}$ takový, že

$$f(x) = T_{n-1}^{f,a}(x) + \frac{1}{n!}\mathrm{d}_{x-a}^n f(\xi).$$

Podle definice Taylorova polynomu

$$T_n^{f,a}(x) = T_{n-1}^{f,a}(x) + \frac{1}{n!}\mathrm{d}_{x-a}^n f(a),$$

takže odečtením rovností dostáváme

$$C(x) := f(x) - T_n^{f,a}(x) = \frac{1}{n!}\left(\mathrm{d}_{x-a}^n f(\xi) - \mathrm{d}_{x-a}^n f(a)\right).$$

Označíme-li ještě $h = (h_1, \ldots, h_p) := x - a$ a použijeme Větu 2.92 o vyjádření n-tého diferenciálu pomocí parciálních derivací, dostáváme

$$C(x) = \frac{1}{n!}\sum_{i_1,\ldots,i_n}^{p} \left(f_{i_1,\ldots,i_n}(\xi) - f_{i_1,\ldots,i_n}(a)\right) h_{i_1} h_{i_2} \cdots h_{i_n},$$

$$\frac{|C(x)|}{\|x-a\|^n} \leq \frac{1}{n!\|h\|^n}\sum_{i_1,\ldots,i_n}^{p} |f_{i_1,\ldots,i_n}(\xi(x)) - f_{i_1,\ldots,i_n}(a)| \cdot |h_{i_1}| \cdot |h_{i_2}| \cdots |h_{i_n}| \leq$$
$$\leq \frac{1}{n!}\sum_{i_1,\ldots,i_n}^{p} |f_{i_1,\ldots,i_n}(\xi(x)) - f_{i_1,\ldots,i_n}(a)|.$$

Protože všechny funkce $f_{i_1,\ldots,i_n}$ jsou spojité v bodě a a $\lim_{x\to a}\xi(x) = a$, podle Věty 1.40 o limitě složené funkce dostáváme $\lim_{x\to a} |C(x)| \cdot \|x-a\|^{-n} = 0$, tj. $f(x) - T_n^{f,a}(x) = o\left(\|x-a\|^n\right), \quad x \to a$, což jsme chtěli dokázat.

2.101 Poznámka. Předpoklad předchozí věty lze zeslabit – stačí předpokládat, že existuje $f^{(n)}(a)$ (viz [D II], Cvičení 2-4 za Větou 203). Pro případ funkcí jedné proměnné je to známo z přednášek pro 1. ročník (srov. [D I], poslední úvaha kap. XI).

2.8 Lokální extrémy

2.102 Definice. (lokální extrém) *Nechť f je funkce n proměnných. Řekneme, že f má v bodě a lokální maximum, existuje-li $\delta > 0$ takové, že pro každý bod $x \in U_\delta(a) \setminus \{a\}$ platí $f(x) \leq f(a)$. Lze-li nerovnost $f(x) \leq f(a)$ nahradit dokonce ostrou nerovností $f(x) < f(a)$, řekneme, že f má v bodě a ostré lokální maximum. Zcela obdobně definujeme pojmy lokálního minima i ostrého lokálního minima funkce.*

Je vhodné již zde definovat obecnější pojem relativního lokálního extrému, kterým se však budeme podrobně zabývat až později.

2.103 Definice. (relativní lokální extrém) *Nechť je dána reálná funkce n proměnných f, množina $M \subset \mathbb{R}^n$ a bod $a \in M \cap D_f$. Řekneme, že f má v bodě a lokální maximum vzhledem k množině M, existuje-li $\delta > 0$ takové, že pro každé $x \in (U_\delta(a) \setminus \{a\}) \cap M$ platí $f(x) \leq f(a)$. Lze-li nerovnost $f(x) \leq f(a)$ nahradit dokonce ostrou nerovností $f(x) < f(a)$, řekneme, že f má v bodě a ostré lokální maximum vzhledem k M. Zcela obdobně definujeme pojmy lokálního minima i ostrého lokálního minima funkce vzhledem ke množině.*

2.104 Poznámka.

(i) „Lokální extrém" je společný název pro lokální maximum a lokální minimum.
(ii) Je-li a vnitřním bodem množiny M, pak pojem lokálního extrému vzhledem k množině M ovšem splývá s pojmem lokálního extrému.

Základní nutná podmínka pro lokální extrémy funkcí více proměnných snadno vyplývá z příslušné nutné podmínky pro funkce jedné proměnné.

2.105 Věta. *Nechť funkce n proměnných f má v bodě $a \in \mathbb{R}^n$ lokální extrém. Potom platí:*

(i) *Je-li $i \in \{1, \ldots, n\}$ a existuje parciální derivace $\frac{\partial f}{\partial x_i}(a)$, pak $\frac{\partial f}{\partial x_i}(a) = 0$.*

(ii) *Má-li funkce f diferenciál v bodě a, pak bod a je stacionárním bodem funkce f, tj.* $\mathrm{d}f(a) = 0$.

Důkaz. Nechť $a = (a_1, \dots, a_n)$ a je dáno $i \in \{1, \dots, n\}$. Snadno vidíme, že parciální funkce $\varphi := f(a_1, \dots, a_{i-1}, \cdot, a_{i+1}, \dots, a_n)$ má v bodě a_i lokální extrém, takže nutně (srov. [D I], Věta 140) $\varphi'(a_i) = \dfrac{\partial f}{\partial x_i}(a) = 0$. Tím je dokázáno tvrzení (i); z něj pak okamžitě plyne tvrzení (ii).

Při hledání lokálních extrémů funkce f tedy stačí vyšetřovat její stacionární body a body, ve kterých f nemá diferenciál.

O tom, zda ve svém stacionárním bodě a má funkce f lokální extrém, můžeme často rozhodnout pomocí druhého diferenciálu $\mathrm{d}^2 f(a)$.

Pro důkaz příslušného výsledku potřebujeme některá fakta o kvadratických formách. Nejdříve připomeňme, že funkce $Q(x)$ na $\mathbb{R}^n$ je kvadratická forma, právě když ji lze psát ve tvaru

$$Q(x) = \sum_{i,j=1}^{n} a_{ij} x_i x_j, \quad \text{kde} \quad a_{ij} \in \mathbb{R} \quad \text{a} \quad x = (x_1, \dots, x_n).$$

Každá kvadratická forma je zřejmě homogenní funkce stupně 2, tj. $Q(tx) = t^2 Q(x)$ kdykoliv $x \in \mathbb{R}^n$ a $t \in \mathbb{R}$.

Je-li $Q(x) > 0$ pro všechna $0 \neq x \in \mathbb{R}^n$, nazývá se kvadratická forma pozitivně definitní. Platí-li $Q(x) \geq 0$ pro všechna x a $Q(x) = 0$ pro nějaké $x \neq 0$, říkáme, že Q je pozitivně semidefinitní.

Obdobně se definují pojmy negativně definitní a negativně semidefinitní kvadratické formy.

Existují-li $x \in \mathbb{R}^n, y \in \mathbb{R}^n$ takové, že $Q(x) > 0$ a $Q(y) < 0$, nazývá se Q indefinitní kvadratická forma.

Pro kvadratické formy na obecném vektorovém prostoru (srov. [Be], [Bi]) se jejich „definitnost“ definuje zcela analogicky.

2.106 Poznámka. *Terminologie kolísá; někdy se pozitivně semidefinitní formou rozumí forma, která je všude nezáporná.*

Lineární algebra zná různé metody, jak určit, kterého z těchto typů je daná kvadratická forma na $\mathbb{R}^n$. Jedna elementární metoda je vyložena v [D II] (za Větou 216). Známé Sylvestrovo pravidlo ([Bi, Důsledek 13.17.]) udává jednoduchou nutnou a postačující podmínku pro to, aby kvadratická forma byla pozitivně definitní (všechny hlavní subdeterminanty její matice jsou kladné).

2.107 Lemma. *Nechť kvadratická forma Q na $\mathbb{R}^n$ je pozitivně definitní. Pak existuje reálné číslo $K > 0$ takové, že $Q(h) \geq K\|h\|^2$ pro každý bod $h \in \mathbb{R}^n$.*

Důkaz. Protože jednotková (eukleidovská) sféra $S := \{x \in \mathbb{R}^n \colon \|x\| = 1\}$ je uzavřená a omezená, a tedy kompaktní, a funkce Q je zřejmě spojitá na $\mathbb{R}^n$, nabývá

Q na S svého minima. Protože Q je pozitivně definitní a $0 \notin S$, je nutně $K := \min\{Q(h)\colon\ h \in S\} > 0$. Pro $h \neq 0$ tedy máme

$$Q(h) = Q\left(||h|| \frac{h}{||h||}\right) = ||h||^2 Q\left(\frac{h}{||h||}\right) \geq K||h||^2.$$

Protože $Q(0) = 0 = K||0||^2$, je důkaz proveden.

Nyní již můžeme dokázat hlavní větu o lokálních extrémech, která mj. udává *postačující podmínky* pro lokální extrémy.

2.108 Věta. (lokální extrémy a druhý diferenciál) *Nechť funkce n proměnných f je třídy C^2 na nějakém otevřeném okolí bodu $a \in \mathbb{R}^n$. Nechť a je stacionární bod funkce f. Položme*

$$Q(h) := \mathrm{d}_h^2 f(a) = \sum_{i,j=1}^{n} \frac{\partial^2 f(a)}{\partial x_i \partial x_j} h_i h_j.$$

Pak platí:

(i) *Je-li kvadratická forma Q pozitivně definitní, má funkce f v bodě a ostré lokální minimum.*
(ii) *Je-li Q negativně definitní, má funkce f v bodě a ostré lokální maximum.*
(iii) *Je-li Q indefinitní, pak f nemá v bodě a lokální extrém.*

Důkaz. Protože a je stacionární bod funkce f, je $\mathrm{d}f(a) = 0$. Podle Taylorovy věty s Peanovým tvarem zbytku (Věta 2.100)

$$f(a+h) - f(a) = \frac{1}{2} Q(h) + o\left(||h||^2\right), \quad h \to 0.$$

Jinými slovy, pro funkci η danou předpisem $\eta(h) := f(a+h) - f(a) - \frac{1}{2}Q(h)$ platí $\lim_{h\to 0} \eta(h)\, ||h||^{-2} = 0$.

Předpokládejme nyní, že Q je pozitivně definitní. Podle Lemmatu 2.107 můžeme zvolit $K > 0$ tak, aby platilo $Q(h) \geq K||h||^2$. Platí tedy

$$f(a+h) - f(a) \geq \frac{1}{2}K||h||^2 + \eta(h) = ||h||^2 \left(\frac{1}{2}K + \eta(h)||h||^{-2}\right).$$

Protože

$$\lim_{h\to 0} \left(\frac{1}{2}K + \eta(h)||h||^{-2}\right) = \frac{1}{2}K > 0,$$

existuje $\delta > 0$ takové, že $f(a+h) - f(a) > 0$, kdykoliv $0 < ||h|| < \delta$. Funkce f má tedy v bodě a ostré lokální minimum.

Je-li Q negativně negativní, pak položíme $f^* := -f$ a $Q^* := \mathrm{d}^2 f^*(a)$. Protože $Q^* = -Q$ je pozitivně definitní, má funkce f^* v bodě a lokální minimum a tedy f má v a lokální maximum.

Je-li Q indefinitní, můžeme zvolit $h_1, h_2 \in \mathbb{R}^n$ tak, že $Q(h_1) > 0$ a $Q(h_2) < 0$. Pak pro $t \neq 0$ máme

$$f(a + th_1) - f(a) = \frac{1}{2} Q(th_1) + \eta(th_1) = t^2 \left(\frac{1}{2} Q(h_1) + \frac{\eta(th_1)}{\|th_1\|^2} \cdot \|h_1\|^2 \right).$$

Protože pro $t \to 0$ má výraz v závorce limitu $\frac{1}{2} Q(h_1) > 0$, existuje $\delta > 0$ takové, že $f(a + th_1) - f(a) > 0$ pro každé $t \in (-\delta, \delta) \setminus \{0\}$. Snadno tedy vidíme, že f má v bodě a ostré relativní lokální minimum vzhledem k přímce $\{a + th_1 \colon t \in \mathbb{R}\}$. Zcela obdobně dostáváme, že f má v bodě a ostré relativní lokální maximum vzhledem k přímce $\{a + th_2 \colon t \in \mathbb{R}\}$. Proto je zřejmé, že f nemá v bodě a lokální extrém.

2.109 Poznámka.

(i) Pokud je Q semidefinitní kvadratická forma, nemůžeme obecně usoudit, zda f má v bodě a lokální extrém; to již známe v případě funkce jedné proměnné (kdy jde o případ $f'(a) = f''(a) = 0$). K dalšímu vyšetřování „semidefinitního případu“ pomocí Taylorovy věty bychom potřebovali znát derivace vyšších řádů.

(ii) Důkaz pro „indefinitní případ“ jsme mohli vést užitím Tvrzení 2.95 na funkce $g_1(t) = f(a + th_1)$, $g_2(t) = f(a + th_2)$ a aplikací věty o souvislosti lokálních extrémů a druhé derivace pro funkce jedné proměnné. Přímočaré užití této metody na „definitní případ“ selhává. To ilustruje následující příklad.

2.110 Příklad. Nechť $f(x, y) = 1$, pokud $y \neq x^2$, $f(0, 0) = 0$ a $f(x, x^2) = -1$ pro každé $x \in \mathbb{R} \setminus \{0\}$. Je snadno vidět, že f má v bodě $(0, 0)$ ostré relativní lokální minimum vzhledem k libovolné přímce procházející bodem $(0, 0)$, ale nemá v bodě $(0, 0)$ lokální extrém.