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 $f(x_1, \ldots, x_n)$ 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 $(x_1, \ldots, x_n)$ a nazýváme ji bodem prostoru $\mathbb{R}^n$; reálná čísla $x_1, \ldots, x_n$ 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 $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ a $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ 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:

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

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

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

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

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

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

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

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 $x = (x_1,\dots,x_n) \in \mathbb{R}^n$ a $1 \le p < \infty$ klademe

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

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 $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ s nezápornými složkami se nazývá Minkowského nerovnost (viz [D II], Věta 105), což je jedna z poměrně hlubokých klasických nerovností. Pro obecné $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$ trojúhelníkovou nerovnost ihned dostaneme, aplikujeme-li Minkowského nerovnost na vektory $x^* = (|x_1|, \dots, |x_n|)$, $y^* = (|y_1|, \dots, |y_n|)$. 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

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

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

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

$$U^1_\varepsilon(a) \subset U^2_\varepsilon(a) \subset U^\infty_\varepsilon(a) \subset U^1_{n\varepsilon}(a),$$

(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

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

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

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

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

$$\langle f, g \rangle \ := \int_a^b f(x) g(x) \, \mathrm{d}x$$

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

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

$$\|f\|_1 \ := \int_a^b |f(x)| \, \mathrm{d}x.$$

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 $\langle f, g \rangle \ := \int_a^b f(x) \overline{g(x)} \, \mathrm{d}x$.

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 $f\restriction_A\colon (A,\rho) \to (Y,\sigma)$ 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 $f\colon (X,\rho) \to (Y,\sigma)$, $a \in X$. Pak následující tvrzení jsou po dvou ekvivalentní.*

(i) *f je spojité v bodě a.*

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

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

(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 $f\colon (X,\rho) \to (Y,\sigma)$. 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 $f\colon (X,\rho) \to (Y,\sigma)$ 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 $f\colon X \to Y$, $g\colon Y \to Z$ jsou spojitá zobrazení. Pak zobrazení $g \circ f\colon X \to Z$ 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 $f\colon (X,\rho) \to (\mathbb{R},\rho_2)$, 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 $f\colon (X,\rho) \to (\mathbb{R}^*,\sigma)$, 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 $\lim_{x \to a} \frac{f(x)}{g(x)} = 0$. (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 $\lim_{x \to a} \frac{f(x)}{g(x)} = 1$. 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*

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

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 $K\|x\|_1 \le \|x\|_2 \le C\|x\|_1$ 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

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

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

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

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, $x_1, x_2, \ldots$ 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 $X := X_1 \times \dots \times X_n$. Položme si otázku, jak na množině X „nejpřirozeněji" definovat metriku. Vedeni analogií se speciálním případem $\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R}$, 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

$$\sigma_2(x,y) := \sqrt{(\rho_1(x_1,y_1))^2 + \dots + (\rho_n(x_n,y_n))^2},$$

$$\sigma_\infty(x,y) := \max\{\rho_1(x_1,y_1), \dots, \rho_n(x_n,y_n)\},$$

$$\sigma_1(x,y) := \rho_1(x_1,y_1) + \dots + \rho_n(x_n,y_n).$$

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* $X := X_1 \times \dots \times X_n$ *opatřený metrikou* σ_∞. *Tento metrický prostor označujeme symbolem* $(X_1, \rho_1) \times \dots \times (X_n, \rho_n)$.

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 $X = X_1 \times \dots \times X_n$ 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 $X = X_1 \times \dots \times X_n$. 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:*

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

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 $X_1, \dots X_n$ 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 $\lim_{n\to\infty} \operatorname{diam} F_n = 0$. Pak $\bigcap_{n=1}^{\infty} F_n$ 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 $\{x \in X\colon x \text{ má vlastnost } V\}$ je reziduální; tj. $\{x \in X\colon x \text{ nemá vlastnost } V\}$ 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 $X := X_1 \times \dots \times X_n$ 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 $f\colon X \to Y$ je f spojité, právě když obě zobrazení $f\restriction_A\colon (A, \rho) \to Y$, $f\restriction_B\colon (B, \rho) \to Y$ jsou spojitá.*
(viii) *Charakteristická funkce $C_A\colon X \to R$ 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 $\bigcap\limits_{\omega\in\Omega} A_\omega \neq \emptyset$, pak množina $\bigcup\limits_{\omega\in\Omega} A_\omega$ je také souvislá.*

1.118 Věta. *Nechť $f\colon (X,\rho) \to (Y,\sigma)$ 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 $\varphi\colon [0,1] \to X$ 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 $L\colon\mathbb{R}^n\to\mathbb{R}^k$ 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 $L\colon X\to Y$ 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 $(e) \implies (d)$ a $(d) \implies (c)$ 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

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

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*

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

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

$$\|\alpha \cdot L\| = \sup\{\|\alpha \cdot L(x)\|: \|x\| \le 1\} = |\alpha| \cdot \sup\{\|L(x)\|: \|x\| \le 1\} = |\alpha| \cdot \|L\|,$$

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

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 $F\colon C[0,1]\to C[0,1]$ 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 $f\colon X\to Y$ je izomorfismus. Pro $x\in X$ položme $\|x\|_3 := \|f(x)\|_2$. Snadno je vidět, že $\|\cdot\|_3$ je norma na X a $f\colon(X,\|\cdot\|_3)\to(Y,\|\cdot\|_2)$ je lineární izometrie. Z toho vyplývá, že $(X,\|\cdot\|_3)$ je také úplný a identita $I\colon(X,\|\cdot\|_1)\to(X,\|\cdot\|_3)$ 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 $f\colon(Y,\rho)\to X$ je spojité v bodě y, právě když všechny funkce*

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

jsou spojité v bodě y.

Důkaz. Zobrazení $\varphi=(\varphi_1,\dots,\varphi_n)\colon X\to\mathbb{R}^n$ 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 $(\varphi\circ f)_i=\varphi_i\circ f\colon(Y,\rho)\to\mathbb{R}$, $i=1,\dots,n$, 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.

$$\begin{aligned}(x_1, \dots, x_n) + (y_1, \dots, y_n) &:= (x_1 + y_1, \dots, x_n + y_n),\\ \alpha \cdot (x_1, \dots, x_n) &:= (\alpha x_1, \dots, \alpha x_n),\end{aligned}$$

pak X je zřejmě lineární prostor. Položíme-li

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

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 $F\colon X_1 \times \cdots \times X_n \to Y$ 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 $F\colon X \to Y$ 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 $\|F(x_1, \dots, x_n)\| \leq C \cdot \|x_1\| \cdot \|x_2\| \cdots \|x_n\|$.*
(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

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

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 $x = (x_1, \dots, x_n)$, $h = (h_1, \dots, h_n)$ 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 $F(x + h) = F(x_1 + h_1, \dots, x_n + h_n)$, není těžké nahlédnout, že diference $D := F(x_1 + h_1, \dots, x_n + h_n) - F(x_1, \dots, x_n)$ 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 $X_1, \dots, X_n$ a Y normované lineární prostory, $X_1, \dots, X_n$ 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 $F(a_1, \dots, a_{i-1}, \cdot, a_{i+1}, \dots, a_n)$ je spojité lineární.

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

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

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

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

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

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 $f(v_1,\dots,v_n) := \det[v_1,\dots,v_n]$ 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 $f(e_1,\dots,e_n) = 1$.

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 $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. 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“

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

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

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

Tato limita však není nic jiného, než derivace funkce

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

v bodě a_i. Skutečně,

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

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

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

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 $K \leq \|h\|_1/\|h\|_2 \leq C$ 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*

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

Důkaz. Protože L je lineární forma, je dána předpisem

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

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

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

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 $\dfrac{\partial f}{\partial x}(0,0)$, $\dfrac{\partial f}{\partial y}(0,0)$, 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

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

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

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

takže $\|z - c\| \le 2\|z^* - c\|$.

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

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

takže $\|z - c\| \le 2\sqrt{1 + K^2}\,\|x - a\|$. 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 $\delta\, a$, 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 $L\colon \mathbb{R}^n \to \mathbb{R}^k$ 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 $\dfrac{\partial F}{\partial x_1}(a)$.) 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 $L_w(t)=t\,w$. 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 $i = 1, \dots, n$ 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 $F = (F_1, \dots, F_k)$ 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*

$$J_F(a), \quad \frac{D(F_1, \dots, F_n)}{D(x_1, \dots, x_n)}(a).$$

2.43 Poznámka.

(i) Jacobián se často nazývá také *funkční determinant* funkcí $F_1, \dots, F_n$.

(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

$$\frac{\mathcal{D}(F_1, \dots, F_k)}{\mathcal{D}(x_1, \dots, x_n)}(a).$$

Souvislost derivace $F'(a)$ s parciálními derivacemi složek $F_1, \dots, F_k$ udává následující věta.

2.44 Věta. *Nechť $F = (F_1, \dots, F_k)$ 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*

$$[F'(a)] = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(a) & \dots & \frac{\partial F_1}{\partial x_n}(a) \\ \frac{\partial F_2}{\partial x_1}(a) & \dots & \frac{\partial F_2}{\partial x_n}(a) \\ \vdots & & \vdots \\ \frac{\partial F_k}{\partial x_1}(a) & \dots & \frac{\partial F_k}{\partial x_n}(a) \end{pmatrix}, \tag{2.10}$$

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, $1 \leq j \leq n$. 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

$$\Delta z = g(f(a + \Delta x)) - g(f(a)) = g(b + \Delta y) - g(b)$$

$$\approx g'(b)(\Delta y) \approx g'(b)\big(f'(a)(\Delta x)\big) = g'(b) \circ f'(a)(\Delta x).$$

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)

$$\|f(a + h) - f(a)\| = \|f'(a)h + r(h)\| \le \|f'(a)\| \cdot \|h\| + \|r(h)\| \le (\|f'(a)\| + 1)\|h\|.$$

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 $f = (f_1, \ldots, f_m)$ je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^m$ a $g = (g_1, \ldots, g_s)$ 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 $h = (h_1, \ldots, h_s) := g \circ f$. Pak platí*

$$[h'(a)] = [g'(b)]\cdot[f'(a)], \quad \text{tj.}$$

$$\frac{\partial h_i}{\partial x_j}(a) = \sum_{k=1}^{m}\frac{\partial g_i}{\partial y_k}(b)\,\frac{\partial f_k}{\partial x_j}(a), \quad 1 \le i \le s,\ 1 \le j \le n. \tag{2.11}$$

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

$$\begin{pmatrix}\frac{\partial h_1}{\partial x_1}(a) & \ldots & \frac{\partial h_1}{\partial x_n}(a)\\ \vdots & & \vdots\\ \frac{\partial h_s}{\partial x_1}(a) & \ldots & \frac{\partial h_s}{\partial x_n}(a)\end{pmatrix} = \begin{pmatrix}\frac{\partial g_1}{\partial y_1}(b) & \ldots & \frac{\partial g_1}{\partial y_m}(b)\\ \vdots & & \vdots\\ \frac{\partial g_s}{\partial y_1}(a) & \ldots & \frac{\partial g_s}{\partial y_m}(b)\end{pmatrix}\begin{pmatrix}\frac{\partial f_1}{\partial x_1}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a)\\ \vdots & & \vdots\\ \frac{\partial f_m}{\partial x_1}(a) & \ldots & \frac{\partial f_m}{\partial x_n}(a)\end{pmatrix},$$

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 $L\colon \mathbb{R}^m \to \mathbb{R}^s$ 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

$$dz^*(a) = \frac{\partial z}{\partial y_1}(b)\,\mathrm{d}y_1^*(a) + \cdots + \frac{\partial z}{\partial y_m}(b)\,\mathrm{d}y_m^*(a),$$

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

$$h'(a) = \frac{\partial g}{\partial y_1}(b)\,f_1'(a) + \cdots + \frac{\partial g}{\partial y_m}(b)\,f_m'(a).$$

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 $\varphi\colon t \mapsto a + th$ 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.

$$\|F(b) - F(a)\| \le K\,\|b - a\|, \qquad \textit{kdykoliv}\ \ a, b \in C.$$

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 $\|F(b) - F(a)\| \le \sqrt{k}K\|b-a\|$. 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 $\|D_{b-a}F(x)\| \le K$ pro každé $x \in \overline{ab} \setminus \{a, b\}$, pak $\|F(b) - F(a)\| \le K$. 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 $i \ne j$, 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 $\dfrac{\partial f}{\partial x}$ má definiční obor menší než f; definičním oborem funkce $\dfrac{\partial f}{\partial x}$ je x-ová osa. Proto zřejmě $\dfrac{\partial^2 f}{\partial y \partial x}$ neexistuje v žádném bodě, takže operátor $\dfrac{\partial^2}{\partial y \partial x}$ zobrazuje funkci f na prázdnou funkci (s definičním oborem $\emptyset$). Naproti tomu zřejmě $\dfrac{\partial^2 f}{\partial x^2} = \dfrac{\partial f}{\partial x}$.

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:

$$\frac{\partial x^y}{\partial x} = x^y \, \frac{y}{x} = x^{y-1}\, y, \quad \frac{\partial^2 (x^y)}{\partial y \partial x} = x^{y-1} y \ln x + x^{y-1},$$

$$\frac{\partial x^y}{\partial y} = x^y \ln x, \qquad \frac{\partial^2 (x^y)}{\partial x \partial y} = x^y \, \frac{y}{x} \ln x + x^y \, \frac{1}{x}.$$

Vidíme tedy, že $\dfrac{\partial^2 f}{\partial y \partial x} = \dfrac{\partial^2 f}{\partial x \partial y}$ 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 $f\colon G \to \mathbb{R}^k$. 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, $f\colon A \to \mathbb{R}^m$, $g\colon B \to \mathbb{R}^s$ 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 $S\colon \mathbb{R}^2 \to \mathbb{R}$ 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 $h_1, \ldots, h_s$ 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 $x_1, \ldots, x_n$, 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 $C \in \mathbb{R}$, $\alpha_i = 0, 1, \ldots,$ 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

$$\text{(2.18)} \qquad \text{parciální derivace } \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \text{ existují na nějakém okolí bodu } (a,b).$$

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, $c \in \mathbb{R}^n$, $1 \leq i,j \leq n$, $i \neq j$ 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 $1 \leq i,j \leq n$. 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 $C^k(G)$ $(k \geq 2)$, 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 $T_{i,j} = T_{j-1,j} \circ \dots \circ T_{i+1,i+2} \circ T_{i,i+1}$.

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 $f(x,y)=x^2y$ 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^2y 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 $f^{(k)}(a)\,(v_1,\dots,v_k) = (\mathrm{d}_{v_1}\circ\mathrm{d}_{v_2}\circ\cdots\circ\mathrm{d}_{v_k})f\,(a)$ pro ty body a, ve kterých má pravá strana smysl pro všechny k-tice vektorů $(v_1,\dots,v_k)\in(\mathbb{R}^n)^k$. Někdy se také píše $f^{(k)}(a;v_1,\dots,v_k)$ místo $f^{(k)}(a)\,(v_1,\dots,v_k)$.

(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

$$\mathrm{d}^k f(a)(h) = f^{(k)}(h,\dots,h).$$

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

$$f''(a)(h,k) = (h_1, \ldots, h_n) \cdot \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \ldots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2^2}(a) & \ldots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \ldots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{bmatrix} \cdot (k_1, \ldots, k_n)^T.$$

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

$$(\mathrm{d}_{v^{(1)}} \circ \mathrm{d}_{v^2} \circ \dots \circ \mathrm{d}_{v^{(k)}}) f = \mathrm{d}_{v^{(1)}} \left(\sum_{i_2,\dots,i_k=1}^{n} \frac{\partial^{k-1} f}{\partial x_{i_2} \dots \partial x_{i_k}} v_{i_2}^{(2)} \cdots v_{i_k}^{(k)} \right)$$
$$= \sum_{i_1=1}^{n} v_{i_1}^{(1)} \frac{\partial}{\partial x_{i_1}} \left(\sum_{i_2,\dots,i_k=1}^{n} \frac{\partial^{k-1} f}{\partial x_{i_2} \dots \partial x_{i_k}} v_{i_2}^{(2)} \cdots v_{i_k}^{(k)} \right)$$
$$= \sum_{i_1,\dots,i_k=1}^{n} \frac{\partial^{k} f}{\partial x_{i_1} \dots \partial x_{i_k}} v_{i_1}^{(1)} \cdot v_{i_2}^{(2)} \cdots v_{i_k}^{(k)}.$$

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.* $g(t) = f(\varphi_1(t), \dots, \varphi_n(t))$*) 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 $h = (h_1, \dots, h_k) \in \mathbb{R}^k$ 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 $i, j = 1, \dots, k$ a $t \in H$ rovnosti

$$\frac{\partial g}{\partial t_j}(t) = \sum_{q=1}^{n} \frac{\partial f}{\partial x_q}(\varphi_1(t), \dots, \varphi_n(t))\, \frac{\partial \varphi_q}{\partial t_j}(t),$$

$$\frac{\partial^2 g}{\partial t_i \partial t_j}(t) = \sum_{q=1}^{n} \left(\sum_{p=1}^{n} \frac{\partial^2 f}{\partial x_p \partial x_q}(\varphi(t))\, \frac{\partial \varphi_p}{\partial t_i}(t)\, \frac{\partial \varphi_q}{\partial t_j}(t) \; + \frac{\partial f}{\partial x_q}(\varphi(t))\, \frac{\partial^2 \varphi_q}{\partial t_i \partial t_j}(t) \right).$$

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

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{1}{k!} \mathrm{d}_{x-a}^k f(a) \; + \; \frac{1}{(n+1)!} d_{x-a}^{n+1} f(\xi).$$

Položíme-li $h := x - a$, můžeme ekvivalentně psát

$$\begin{aligned} f(a+h) = f(a) &+ \frac{1}{1!} \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_p \frac{\partial}{\partial x_p} \right) f(a) + \dots \\ &+ \frac{1}{n!} \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_p \frac{\partial}{\partial x_p} \right)^n f(a) + \\ &+ \frac{1}{(n+1)!} \left(h_1 \frac{\partial}{\partial x_1} + \cdots + h_p \frac{\partial}{\partial x_p} \right)^{n+1} f(a + \theta h). \end{aligned}$$

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 $k \in \{1, \dots, n+1\}$ a $t \in (\alpha, \beta)$ platí

$$g^{(k)}(t) = \mathrm{d}_h^k f(a + th). \tag{2.29}$$

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

$$f(x) = f(a) + \sum_{k=1}^{n} \frac{1}{k!} \mathrm{d}_{x-a}^k f(a) \; + \; \frac{1}{(n+1)!} d_{x-a}^{n+1} f(a + \theta h),$$

takže stačí položit $\xi := a + \theta h$.

Polynom p proměnných

$$T_n^{f,a}(x) := f(a) + \sum_{k=1}^{n} \frac{1}{k!} \mathrm{d}_{x-a}^k f(a)$$

budeme nazývat *Taylorův polynom n-tého řádu funkce f v bodě a*. Při užití této terminologie má vzorec z Věty 2.98 tvar

$$f(x) = T_n^{f,a}(x) + \frac{1}{(n+1)!} d_{x-a}^{n+1} f(\xi).$$

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. $f(x) - T_n^{f,a}(x) = o\left(\|x-a\|^n\right), \ x \to a$, 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 $\dfrac{\partial f}{\partial x_i}(a)$, pak $\dfrac{\partial f}{\partial x_i}(a) = 0$.*

(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.

Obr. 2.8.

2.9 Věta o implicitních funkcích

Nechť $g(x, y)$ je funkce dvou proměnných a uvažujme rovnici

$$g(x, y) = 0. \tag{2.30}$$

Ptejme se, zda z rovnice (2.30) lze „jednoznačně vypočítat y pomocí x“. Je-li tomu tak, říkáme, že takto zadaná funkce $y(x)$ je *implicitně zadána* rovnicí (2.30). Jinými slovy a přesněji: (2.30) považujeme za rovnici s parametrem x a neznámou y; implicitně zadaná („implicitní“) funkce $y(x)$ pak vyjadřuje závislost (jediného) kořene y na parametru x. Jejím definičním oborem je ovšem množina parametrů x, pro které má (2.30) řešení.

Je-li například $g(x, y) = 4x^2 + 2y$, potom pro každý parametr $x \in \mathbb{R}$ existuje právě jeden kořen $y(x) = -2x^2$; rovnicí (2.30) je tedy implicitně zadána funkce $x \mapsto -2x^2, \ \ x \in \mathbb{R}$.

Je-li však $g(x, y) = x^2 + y^2 - 1$, pak rovnicí (2.30) není implicitně zadána žádná funkce: pro každý parametr $x \in (-1, 1)$ má totiž rovnice (2.30) dva kořeny $y = \pm\sqrt{1 - x^2}$.

Podíváme-li se na problém geometricky, vidíme, že rovnice (2.30) jednoznačně určuje „implicitní funkci“ právě tehdy, když množina

$$N_g \ := \{(x, y)\colon \ g(x, y) = 0\}$$

je grafem funkce. V posledním případě N_g grafem funkce není; ptejme se tedy, zda N_g je aspoň „lokálně grafem funkce“. Přesněji: pro které body $c \in N_g$ existuje okolí $U = U_\delta(c)$ takové, že $N_g \cap U$ je grafem funkce? Na obr. 2.8 vlevo vidíme, že to je pravda pro $c = e$, ale ne pro $c = d$. Po krátké úvaze vidíme, že to platí pro všechny body $c \in N_g$ s výjimkou bodů $(-1, 0)$ a $(1, 0)$.

Zdá se tedy, že pro „většinu pěkných funkcí“ g by odpověď na naši otázku mohla být kladná pro „většinu“ bodů $c \in N_g$. Uveďme pro to ještě „algebraický důvod“:

Je-li $c = (x_0, y_0) \in N_g$, má rovnice $g(x_0, y) = 0$ kořen y_0. Ze zkušenosti víme, že „většina obvyklých rovnic“ má jen konečně mnoho kořenů, takže je „dosti věrohodné“, že existuje $\Delta > 0$ takové, že v intervalu $(y_0 - \Delta, y_0 + \Delta)$ tato rovnice pouze kořen y_0. Pokud platí, že $\frac{\partial g(c)}{\partial y} = (g(x_0,\cdot))'(y_0) \neq 0$, pak existenci takového $\Delta > 0$ snadno umíme dokázat: parciální funkce $g(x_0,\cdot)$ je pak totiž v bodě y_0 rostoucí nebo klesající. Zdá se také, že pro „pěkné“ funkce g bude mít velmi malá změna parametru jen malý vliv na řešení rovnice, takže pro parametry x velmi blízké parametru x_0 bude mít rovnice (2.30) v nějakém okolí bodu y_0 také právě jeden kořen.

Všimněme si ještě výjimečných bodů $(-1,0)$, $(0,1)$ z předchozího příkladu: jsou to jediné dva body $(x,y) \in N_g$, ve kterých je $\frac{\partial g(x,y)}{\partial y} = 2y$ nulová. Zdá se tedy, že již uvažovaná podmínka $\frac{\partial g(c)}{\partial y} \neq 0$ míří k podstatě problému.

Uveďme ještě heuristickou úvahu naznačující, že hladká „vazbová funkce“ $g(x,y)$ by měla „zpravidla“ rovnicí (2.30) „lokálně“ definovat hladké „implicitní funkce“: Ztotožníme-li body (x,y) a $(x,y,0)$, pak N_g je průnikem hladké plochy o rovnici $z = g(x,y)$ a roviny o rovnici $z = 0$. Je tedy přirozené očekávat, že N_g bude hladká křivka. Hladká křivka pak bývá obvykle „lokálně“ grafem hladké funkce.

2.111 Poznámka. Tradiční termín „implicitní funkce“ neoznačuje ovšem vlastnost této funkce, ale hovoří o způsobu, jakým byla funkce zadána (nikoliv „explicitně vzorcem“, ale „implicitně rovnicí“).

Pokud uvažujeme rovnici $g(x_1,\dots,x_n,y) = 0$ s neznámou y a s n parametry, je situace zcela analogická. Naše nepřesné úvahy již tedy částečně předjímají následující větu.

V ní uvažujeme na všech eukleidovských prostorech *maximovou* metriku; speciálně $U_\delta(x) := U_\delta^\infty(x)$. Pro $x = (x_1,\dots,x_n) \in \mathbb{R}^n$ a $y \in \mathbb{R}$ ztotožňujeme bod $(x,y) \in \mathbb{R}^n \times \mathbb{R}$ s bodem $(x_1,\dots,x_n,y) \in \mathbb{R}^{n+1}$. Platí tedy

$$\begin{aligned} U_\delta(x,y) &= U_\delta(x) \times U_\delta(y) = \\ &= (x_1-\delta,x_1+\delta) \times \cdots \times (x_n-\delta,x_n+\delta) \times (y-\delta,y+\delta). \end{aligned}$$

Tvrzení věty je znázorněno na obr. 2.8 vpravo.

2.112 Věta. (o implicitní funkci) *Nechť n,p jsou přirozená čísla, $a \in \mathbb{R}^n$, $b \in \mathbb{R}$ a $g(x_1,\dots,x_n,y) = g(x,y)$ je funkce $(n+1)$ proměnných. Nechť dále platí:*

(i) *Rovnice $g(x,y) = 0$ je splněna pro $x = a \in \mathbb{R}^n$ a $y = b \in \mathbb{R}$ (tj. $g(a,b) = 0$).*
(ii) *Funkce g je třídy C^p na nějakém otevřeném okolí V bodu $c := (a,b) \in \mathbb{R}^{n+1}$.*
(iii) $\dfrac{\partial g}{\partial y}(c) \neq 0$.

Pak existují čísla $\delta > 0$ a $\Delta > 0$ taková, že platí:

(a) $U_\delta(a) \times U_\Delta(b) \subset V$.
(b) *Pro každý bod $x = (x_1,\dots,x_n) \in U_\delta(a) = (a_1-\delta,a_1+\delta)\times\cdots\times(a_n-\delta,a_n+\delta)$ existuje v intervalu $(b-\Delta,b+\Delta)$ právě jedno číslo $y =: f(x)$, pro které je splněna rovnice $g(x,y) = 0$.*
(c) *Takto definovaná funkce f je třídy C^p na svém definičním oboru $U_\delta(a)$.*

Obr. 2.9.

2.113 Poznámka. Tvrzení věty o implicitní funkci lze formulovat různými způsoby. Například místo dvou tvrzení (b) a (c) lze psát stručně pouze

(d) Množina $\{(x,y) \in U_\delta(a) \times U_\Delta(b)\colon\ g(x,y) = 0\}$ je grafem funkce, která je třídy C^p na $U_\delta(a)$.

Tvrzení (d) lze ekvivalentně formulovat také takto:

(d^*) Existuje funkce $f\colon U_\delta(a) \to \mathbb{R}$ třídy C^p taková, že (pro každý bod $(x,y) \in \mathbb{R}^{n+1}$) platí ekvivalence

$$(x \in U_\delta(a)\ \wedge\ y \in (b-\Delta, b+\Delta)\ \wedge\ g(x,y) = 0) \quad\Longleftrightarrow\quad y = f(x).$$

Někdy se také vynechává tvrzení (a); tím se dostane „formálně slabší" věta, z které lze však snadno odvodit větu ve výše uvedené formě.

Důkaz Věty 2.112 provedeme ve 4 krocích.

V 1. kroku dokážeme, že z předpokladů věty plyne její tvrzení bez bodu (c). Bez újmy na obecnosti můžeme předpokládat $\dfrac{\partial g}{\partial y}(c) > 0$ (v opačném případě bychom mohli položit $g^* := -g$ a řešit rovnici $g^*(x,y) = 0$, která je ekvivalentní s rovnicí $g(x,y) = 0$). Protože funkce $\dfrac{\partial g}{\partial y}$ je spojitá na V, můžeme najít $\Delta > 0$ takové, že $U_{2\Delta}(c) = U_{2\Delta}(a) \times (b-2\Delta, b+2\Delta) \subset V$ a $\dfrac{\partial g}{\partial y}(z) > 0$ pro každý bod $z \in U_{2\Delta}(c)$ (srov. obr. 2.9).

Z věty o souvislosti znaménka derivace a monotonie funkce jedné proměnné pak vyplývá, že pro každé $x \in U_\Delta(a)$ je parciální funkce $g(x,\cdot)$ rostoucí na intervalu $(b-2\Delta, b+2\Delta)$. Speciálně platí, že $g(a, b+\Delta) > 0$ a $g(a, b-\Delta) < 0$. Ze spojitosti funkce g v bodech $(a, b+\Delta) \in V$, $(a, b-\Delta) \in V$ plyne existence čísla $0 < \delta < \Delta$ takového, že $g(x, b+\Delta) > 0$ a $g(x, b-\Delta) < 0$ pro každé $x \in U_\delta(a)$.

Zvolme nyní libovolně $x \in U_\delta(a)$. Protože funkce $g_x := g(x,\cdot)$ je spojitá a rostoucí na intervalu $[b-\Delta, b+\Delta]$, přičemž $g_x(b-\Delta) < 0$, $g_x(b+\Delta) > 0$, existuje v intervalu $(b-\Delta, b+\Delta)$ právě jedno číslo $y =: f(x)$, pro které platí rovnost $g_x(y) = g(x,y) = 0$. Dokázali jsme tedy, že z předpokladů věty plyne existence čísel $\delta > 0$, $\Delta > 0$, pro která platí (a) a (b).

Ve 2. kroku použijeme právě dokázané tvrzení k důkazu, že výše nalezená funkce $f(x)$ je spojitá na $U_\delta(a)$. Zvolíme tedy libovolné $a^* \in U_\delta(a)$ a $\varepsilon > 0$. (Je vhodné si kreslit obrázek, na kterém jsou znázorněny body a okolí, *nikoliv však funkce.*) Potřebujeme nalézt okolí W bodu a^* takové, že $|f(x)-f(a^*)| < \varepsilon$ pro každé $x \in W$. Označme $b^* := f(a^*)$, zvolme $0 < \eta < \varepsilon$ takové, že $U_\eta(a^*, b^*) \subset U_\delta(a) \times U_\Delta(b)$ a položme
$V^* := U_\eta(a^*, b^*)$. Pak zřejmě $g(a^*, b^*) = 0$, g je třídy C^p na okolí V^* bodu
$c^* := (a^*, b^*)$ a $\dfrac{\partial g}{\partial y}(c^*) \neq 0$. Podle 1. kroku existují čísla $\delta^* > 0$, $\Delta^* > 0$ taková, že
$U_{\delta^*}(a^*) \times U_{\Delta^*}(b^*) \subset V^*$ (tj. $\delta^* < \eta$, $\Delta^* < \eta$) a pro každé $x \in U_{\delta^*}(a^*)$ existuje v $U_{\Delta^*}(b^*)$ právě jeden bod y, pro který platí $g(x,y) = 0$. Protože $x \in U_{\delta^*}(a^*) \subset U_\delta(a)$ a $y \in U_{\Delta^*}(b^*) \subset U_\Delta(b)$, je podle definice funkce f nutně $f(x) = y$. Protože $y \in U_{\Delta^*}(b^*) \subset U_\varepsilon(b^*)$, dostáváme $|f(x) - f(a^*)| < \varepsilon$. Stačí tedy položit $W := U_{\delta^*}(a^*)$.

Ve 3. kroku dokážeme i tvrzení (c) v případě, že $p = 1$. Předpokládáme tedy, že g je třídy C^1 na V a chceme dokázat, že f je třídy C^1 na $U_\delta(a)$. Zvolme libovolně $x \in U_\delta(a)$, $1 \le i \le n$ a počítejme $\dfrac{\partial f}{\partial x_i}(x)$, tj. limitu

$$\lim_{h\to 0} \frac{f(x+he_i) - f(x)}{h}.$$

Zřejmě existuje $h_0 > 0$ takové, že $x + he_i \in U_\delta(a)$, kdykoliv $0 < |h| < h_0$; jedno takové h uvažujme. Položme $u := (x, f(x))$, $v = v(h) := (x + he_i, f(x+he_i))$. Podle definice funkce f platí $g(u) = g(v) = 0$. Protože oba body u, v leží v otevřené konvexní množině $U_\delta(a) \times U_\Delta(b)$ a g je třídy C^1 na této množině, podle věty o přírůstku funkce (Věta 2.60) existuje bod $\xi = \xi(h) \in \overline{uv} \setminus \{u, v\}$, pro který platí

$$\begin{aligned} 0 = g(v) - g(u) &= (\operatorname{grad} g(\xi(h)), (he_i, f(x+he_i) - f(x))) = \\ &= \frac{\partial g}{\partial x_i}(\xi(h))h + \frac{\partial g}{\partial y}(\xi(h))(f(x+he_i) - f(x)). \end{aligned}$$

Po úpravě dostáváme

$$\frac{f(x+he_i) - f(x)}{h} = -\frac{\frac{\partial g}{\partial x_i}(\xi(h))}{\frac{\partial g}{\partial y}(\xi(h))}.$$

Protože f je spojitá v bodě x, platí $\lim_{h\to 0} v(h) = u$ a tedy i $\lim_{h\to 0} \xi(h) = u$. Ze spojitosti funkcí $\dfrac{\partial g}{\partial x_i}$, $\dfrac{\partial g}{\partial y}$ v bodě $u = (x, f(x))$, věty o limitě složeného zobrazení (Věta 1.40) a věty o limitě podílu funkcí dostáváme

$$\frac{\partial f}{\partial x_i}(x) = \lim_{h\to 0} \frac{f(x+he_i) - f(x)}{h} = -\frac{\frac{\partial g}{\partial x_i}(x, f(x))}{\frac{\partial g}{\partial y}(x, f(x))}. \tag{2.31}$$

Protože funkce $\dfrac{\partial g}{\partial x_i}$, $\dfrac{\partial g}{\partial y}$ jsou spojité na $U_\delta(a)\times U_\Delta(b)$ a zobrazení $h(x) := (x, f(x))$, $h\colon U_\delta(a) \to U_\delta(a) \times U_\Delta(b)$ je spojité zobrazení, dostáváme, že každá z funkcí $\dfrac{\partial f}{\partial x_i}$, $i = 1, \dots, n$, je spojitá na $U_\delta(a)$, takže f je třídy C^1 na $U_\delta(a)$.

Ve čtvrtém, posledním kroku dokážeme platnost (c) v případě $p > 1$. Nechť $V(k)$ je výrok, že f je třídy C^k na $U_\delta(a)$. Ve 3. kroku jsme dokázali, že platí $V(1)$. Předpokládejme nyní, že $V(k)$ platí pro nějaké přirozené číslo $k < p$. Pak je i zobrazení $h(x) := (x, f(x))$ třídy C^k na $U_\delta(a)$. Protože funkce $\dfrac{\partial g}{\partial x_i}$, $\dfrac{\partial g}{\partial y}$ jsou třídy C^{p-1} (a tedy i třídy C^k) na $U_\delta(a) \times U_\Delta(b)$, rovnost (2.31) a Věta 2.76 (o skládání zobrazení třídy C^k) dává, že funkce $\dfrac{\partial f}{\partial x_i}$, $i = 1, \dots, n$, jsou třídy C^k na $U_\delta(a)$. Platí tedy $V(k+1)$. Indukcí pak snadno dostáváme, že platí $V(p)$.

Nyní uvažujme obecnější případ, kdy místo jedné rovnice je dáno několik rovnic

$$
\begin{aligned}
g_1(x_1, \dots, x_n, y_1, \dots, y_k) &= 0 \\
&\cdots \\
g_k(x_1, \dots, x_n, y_1, \dots, y_k) &= 0.
\end{aligned}
\tag{2.32}
$$

Považujeme-li $y_1, \dots, y_k$ za neznámé a $x_1, \dots, x_n$ za parametry, jde o systém k rovnic s k neznámými a n parametry. Naše zkušenost s rovnicemi nám říká, že „poměrně často“ nastává případ, že k rovnic o k neznámých má právě jedno řešení. (Z lineární algebry je např. dobře známo, kdy tento případ nastává pro systém lineárních rovnic.) Pokud tento případ jednoznačné řešitelnosti nastává, je ke každé n-tici parametrů $x_1, \dots, x_n$ přiřazena k-tice kořenů $y_1, \dots, y_k$; tímto předpisem je pak určeno k funkcí o n proměnných $y_1(x_1, \dots, x_n), \dots, y_k(x_1, \dots, x_n)$. Stručněji: z rovnic (2.32) jsme schopni jednoznačně vypočítat $y_1, \dots, y_k$ pomocí $x_1, \dots, x_k$.

Jednou z nejdůležitějších vět matematické analýzy je následující „věta o implicitních funkcích“, která ukazuje, že výsledek o jednoznačné řešitelnosti systémů lineárních rovnic se dá do jisté míry zobecnit na nelineární případ, pokud jsou všechny funkce $g_1, \dots, g_k$ hladké. „Důvodem“ je to, že hladké funkce jsou „lokálně téměř afinní“. Příslušná věta má ovšem nutně také lokální charakter. Základní podmínkou bude - stejně jako v teorii lineárních rovnic - „podmínka regularity“, která požaduje, aby jistá matice byla regulární.

2.114 Věta. (klasický složkový zápis věty o implicitních funkcích) *Nechť n, k, p jsou přirozená čísla. Předpokládejme, že:*

(i) *Je dán bod $c = (a_1, \dots, a_n, b_1, \dots, b_k) \in \mathbb{R}^{n+k}$ a k funkcí $n+k$ proměnných*

$$g_1(x_1, \dots, x_n, y_1, \dots, y_k), \dots, g_k(x_1, \dots, x_n, y_1, \dots, y_k),$$

které jsou třídy C^p na nějakém otevřeném okolí bodu c.

(ii) *Rovnice*

$$
\begin{aligned}
&g_1(x_1,\dots,x_n,y_1,\dots,y_k) = 0\\
&\qquad\qquad\dots\\
&g_k(x_1,\dots,x_n,y_1,\dots,y_k) = 0
\end{aligned}
\tag{2.33}
$$

jsou splněny pro $(x_1,\dots,x_n,y_1,\dots,y_k) = (a_1,\dots,a_n,b_1,\dots,b_k) = c$.

(iii)

$$
\frac{D(g_1,\dots,g_k)}{D(y_1,\dots,y_k)}(c) = \begin{vmatrix} \dfrac{\partial g_1}{\partial y_1}(c) & \dots & \dfrac{\partial g_1}{\partial y_k}(c) \\ \dfrac{\partial g_2}{\partial y_1}(c) & \dots & \dfrac{\partial g_2}{\partial y_k}(c) \\ \vdots & & \vdots \\ \dfrac{\partial g_k}{\partial y_1}(c) & \dots & \dfrac{\partial g_k}{\partial y_k}(c) \end{vmatrix} \neq 0.
$$

Pak existují čísla $\delta_1 > 0$, $\delta_2 > 0$ *s těmito vlastnostmi:*

(a) *Pro každý bod* $(x_1,\dots,x_n)$ *takový, že*

$$|x_1 - a_1| < \delta_1, \dots, |x_n - a_n| < \delta_1$$

existuje právě jeden bod

$$(y_1,\dots,y_k) =: (f_1(x_1,\dots,x_n),\dots,f_k(x_1,\dots,x_n)),$$

pro který platí nerovnosti $|y_1 - b_1| < \delta_2, \dots, |y_k - b_k| < \delta_2$ *a rovnice* (2.33).

(b) *Takto definované funkce* $f_1(x_1,\dots,x_n),\dots,f_k(x_1,\dots,x_n)$ *jsou třídy* C^p *na intervalu* $(a_1 - \delta_1, a_1 + \delta_1) \times \dots \times (a_n - \delta_1, a_n + \delta_1)$.

Dříve než větu o implicitních funkcích dokážeme, uvedeme ji ještě v jiném, „bezsložkovém" tvaru.

2.115 Věta. (moderní zápis věty o implicitních funkcích) *Nechť* n, k, p *jsou přirozená čísla; na prostorech* $X = \mathbb{R}^n, Y = \mathbb{R}^k, X \times Y = \mathbb{R}^{n+k}$ *uvažujme maximovou normu* $\|\cdot\|_\infty$. *Předpokládejme toto:*

$(i)^*$ *Je dáno zobrazení* g *z* $X \times Y$ *do* Y, *které je třídy* C^p *na nějakém otevřeném okolí bodu* $c = (a,b) \in X \times Y$.

$(ii)^*$ $g(a,b) = 0$.

$(iii)^*$ *Parciální diferenciál* $g'_y(c) = (g(a,\cdot))'(b)$ *je lineární bijekce* Y *na* Y.

Pak existují čísla $\delta_1 > 0$, $\delta_2 > 0$ *s těmito vlastnostmi:*

$(a)^*$ *Existuje zobrazení* $f \colon U_{\delta_1}(a) \to U_{\delta_2}(b)$, *pro které platí ekvivalence*

$$(g(x,y) = 0 \ \wedge\ (x,y) \in U_{\delta_1}(a) \times U_{\delta_2}(b)) \iff y = f(x).$$

$(b)^*$ *Toto zobrazení* f *je třídy* C^p *na* $U_{\delta_1}(a)$.

Je snadno vidět, že pokud položíme $g := (g_1, \dots, g_k)$, $f := (f_1, \dots, f_k)$, $a := (a_1, \dots, a_k)$, $b := (b_1, \dots, b_k)$, jsou podmínky $(i)^*, (ii)^*, (iii)^*, (a)^*, (b)^*$ popořadě ekvivalentní s podmínkami (i), (ii), (iii), (a), (b).

V této modernější formulaci (při vhodném zobecnění pojmů derivace a zobrazení třídy C^p) věta platí i v případě, že X, Y jsou obecné (nekonečně dimenzionální) Banachovy prostory (viz Věta 3.52).

Důkaz. Idea klasického důkazu je přirozená a elementární; postupujeme známou dosazovací metodou. Z poslední rovnice systému (2.33) vypočítáme („lokálně") neznámou y_k pomocí ostatních neznámých $y_1, \dots, y_{k-1}$ a parametrů $x_1, \dots, x_n$; k tomu podle Věty 2.112 potřebujeme vědět, že $\dfrac{\partial g_k}{\partial y_k}(c) \neq 0$, což lze zařídit případným přečíslováním funkcí $g_1, \dots, g_k$. Když toto vyjádření dosadíme do předchozích $k-1$ rovnic, dostaneme systém $k-1$ rovnic o $k-1$ neznámých a n parametrech. Abychom mohli postupovat stejným způsobem dále, potřebujeme dokázat, že vzniklá soustava rovnic opět splňuje podmínku regularity (iii). Tento krok si podstatně usnadníme, uvědomíme-li si, že tvrzení věty stačí dokázat, zaměníme-li podmínku regularity (iii) silnější podmínkou:

$(iii)_s$ Matice $\left(\dfrac{\partial g_i}{\partial y_j}(c)\right)_{i,j=1}^{k}$ je jednotková; tj. $\dfrac{\partial g_i}{\partial y_j}(c) = \delta_{ij}$,

(kde δ_{ij} je Kroneckerův symbol). To je vidět z této jednoduché úvahy: Označme $g := (g_1, \dots, g_k)$. Víme, že $L := (g(a, \cdot))'(b)$ je lineární bijekce $\mathbb{R}^k$ na $\mathbb{R}^k$. Položme nyní $g^* := L^{-1} \circ g$. Pak g^* je zřejmě třídy C^p na V, $g^*(c) = 0$ a

$$(g^*(a, \cdot))'(b) = (L^{-1} \circ g(a, \cdot))'(b) = L^{-1} \circ (g(a, \cdot))'(b) = I,$$

kde $I: \mathbb{R}^k \to \mathbb{R}^k$ je identické zobrazení. Je tedy $\dfrac{\partial g_i^*}{\partial y_j}(c) = \delta_{ij}$. Nyní si stačí uvědomit, že $g(x, y) = 0$ platí, právě když $g^*(x, y) = 0$.

Zmíněnou dosazovací metodu uplatníme tak, že důkaz provedeme indukcí vzhledem ke k. Nejprve pevně zvolíme libovolná přirozená čísla p a n. Dále pro každé přirozené číslo k označíme symbolem $V(k)$ tvrzení, že z předpokladů (i), (ii), $(iii)_s$ vyplývá existence čísel $\delta_1 > 0$, $\delta_2 > 0$ s vlastnostmi (a) a (b). Víme, že stačí dokázat platnost $V(k)$ pro všechna přirozená k.

Je-li $k = 1$, pak (2.33) je pouze jedna rovnice $g_1(x_1, \dots, x_n, y_1) = 0$. Z podmínky $(iii)_s$ plyne, že $\dfrac{\partial g_1}{\partial y_1}(c) = 1$. Věta 2.112 o implicitní funkci pak okamžitě dává platnost $V(1)$.

Nyní předpokládejme, že je dáno přirozené číslo $k > 1$ a tvrzení $V(k-1)$ platí. Abychom dokázali $V(k)$, uvažujme systém rovnic (2.33) a předpokládejme platnost podmínek (i), (ii), $(iii)_s$. Protože $\dfrac{\partial g_k}{\partial y_k}(c) = 1 \neq 0$, podle Věty 2.112 existuje funkce $n + k - 1$ proměnných ψ třídy C^p definovaná na nějakém otevřeném okolí bodu

$\tilde{c} := (a_1, \ldots, a_n, b_1, \ldots, b_{k-1})$ a otevřené okolí $\widetilde{V}$ bodu c takové, že platí tato ekvivalence:

$$\text{(2.34)} \qquad \Big(g_k(x_1, \ldots, x_n, y_1, \ldots, y_k) = 0 \quad \wedge \quad (x,y) \in \widetilde{V} \Big) \iff$$
$$\iff \quad y_k = \psi(x_1, \ldots, x_n, y_1, \ldots, y_{k-1}).$$

Jsou tedy ekvivalentní následující dvě podmínky:

(A) $(x,y) \in \widetilde{V}$ a jsou splněny rovnice (2.33).

(B) $y_k = \psi(x_1, \ldots, x_n, y_1, \ldots, y_{k-1})$ a je splněna tato soustava $k-1$ rovnic:

$$\text{(2.35)} \qquad h_1(x_1, \ldots, x_n, y_1, \ldots, y_{k-1}) :=$$
$$= g_1(x_1, \ldots, x_n, y_1, \ldots, y_{k-1}, \psi(x_1, \ldots, x_n, y_1, \ldots, y_{k-1})) = 0$$
$$\cdots$$
$$h_{k-1}(x_1, \ldots, x_n, y_1, \ldots, y_{k-1}) :=$$
$$= g_{k-1}(x_1, \ldots, x_n, y_1, \ldots, y_{k-1}, \psi(x_1, \ldots, x_n, y_1, \ldots, y_{k-1})) = 0.$$

Protože pro $(x,y) = c = (a,b)$ je splněna levá strana ekvivalence (2.34), její pravá strana dává $\psi(\tilde{c}) = b_k$. Z Věty 2.76 snadno dostáváme, že

$(i)'$ funkce $h_1, \ldots, h_{k-1}$ jsou třídy C^p na nějakém otevřeném okolí bodu $\tilde{c}$

a podle (ii) platí

$(ii)'$ $h_1(\tilde{c}) = 0, \ldots, h_{k-1}(\tilde{c}) = 0.$

Podle Věty 2.54 dostáváme pro každé $1 \le i,j \le k-1$

$$\frac{\partial h_i}{\partial y_j}(\tilde{c}) = \frac{\partial g_i}{\partial y_j}(c) + \frac{\partial g_i}{\partial y_k}(c) \frac{\partial \psi}{\partial y_j}(\tilde{c}).$$

Protože podle $(iii)_s$ je $\dfrac{\partial g_i}{\partial y_k}(c) = 0$ a $\dfrac{\partial g_i}{\partial y_j}(c) = \delta_{ij}$, máme

$(iii)'$ $\quad \dfrac{\partial h_i}{\partial y_j}(\tilde{c}) = \delta_{ij}, \quad 1 \le i,j \le k-1.$

Na soustavu rovnic (2.35) můžeme tedy použít indukční předpoklad (platnost $V(k-1)$) a dostaneme okolí W bodu $\tilde{c}$ a funkce $\varphi_1(x_1, \ldots, x_n), \ldots, \varphi_{k-1}(x_1, \ldots, x_n)$, které jsou třídy C^p na nějakém otevřeném okolí bodu a takové, že následující podmínky jsou ekvivalentní:

(C) $(x_1, \ldots, x_n, y_1, \ldots, y_{k-1}) \in W$ a platí (2.35),

(D) $y_1 = \varphi_1(x_1, \ldots, x_n), \ \ldots, \ y_{k-1} = \varphi_{k-1}(x_1, \ldots, x_n)$.

Uvědomíme si ještě, že pro každé dva body $x \in \mathbb{R}^n$ a $y \in \mathbb{R}^k$ je podmínka $(x_1, \ldots, x_n, y_1, \ldots, y_{k-1}) \in W$ ekvivalentní podmínce $(x,y) \in W \times \mathbb{R}$. Pak již snadno dostáváme, že (pro $(x,y) \in \mathbb{R}^{n+k}$) platí tato ekvivalence:

$$\text{(2.36)} \qquad \Big((2.33) \quad \wedge \quad (x,y) \in \widetilde{V} \cap (W \times \mathbb{R}) \Big) \iff$$
$$\iff \quad ((D) \quad \wedge \quad y_k = \psi(x_1, \ldots, x_n, \varphi_1(x_1, \ldots, x_n), \ldots, \varphi_{k-1}(x_1, \ldots, x_n)))\,.$$

Podrobné zdůvodnění: Platí-li podmínka nalevo, platí podmínka (A) a tedy i (B). Proto platí také podmínky (C) a (D). Dosadíme-li nyní rovnosti (D) do rovnosti $y_k = \psi(x_1, \dots, x_n, y_1, \dots, y_{k-1})$, dostáváme podmínku napravo. Dále předpokládejme, že platí podmínka napravo. Po dosazení (D) do druhé části této podmínky máme $y_k = \psi(x_1, \dots, x_n, y_1, \dots, y_{k-1})$. Z (D) dostáváme (C) a tedy také $(x, y) \in W \times \mathbb{R}$ a (2.35). Vzhledem k rovnosti $y_k = \psi(x_1, \dots, x_n, y_1, \dots, y_{k-1})$ dostáváme (B) a tedy i (A). Platí tudíž podmínka nalevo.

Z ekvivalence (2.36) vidíme, že „v blízkosti bodu c“ lze z rovnic (2.33) vypočítat $y_1, \dots, y_k$ pomocí $x_1, \dots, x_n$. Abychom však dostali přesné tvrzení věty, musíme učinit další snadné, ale trochu nepřehledné úvahy:

Položme $U := \widetilde{V} \cap (W \times \mathbb{R})$; vidíme, že U je otevřené okolí bodu c v $\mathbb{R}^{n+k}$. Dále položme

$$\varphi_k(x_1, \dots, x_n) := \psi(x_1, \dots, x_n, \varphi_1(x_1, \dots, x_n), \dots, \varphi_{k-1}(x_1, \dots, x_n)).$$

Protože pro bod $(x_1, \dots, x_n, y_1, \dots, y_{k-1}) = (a_1, \dots, a_k, b_1, \dots, b_{k-1}) = \tilde{c}$ je splněna podmínka (C), je pro něj (podle $(ii)^*$) splněna i podmínka (D), takže platí $\varphi_1(a) = b_1, \dots, \varphi_{k-1}(a) = b_{k-1}$. Z Věty 2.76 o skládání funkcí třídy C^p pak snadno dostáváme, že φ_k je třídy C^p na nějakém okolí bodu a. Položíme-li nyní $\varphi := (\varphi_1, \dots, \varphi_k)$, je $\varphi(a) = b$ a φ je třídy C^p na nějakém otevřeném okolí $Z \subset \mathbb{R}^n$ bodu a. Ekvivalenci (2.36) můžeme nyní přepsat ve tvaru

$$((2.33) \;\wedge\; (x, y) \in U) \iff y = \varphi(x). \tag{2.37}$$

Zvolme nyní $\delta_2 > 0$ tak, aby

$$U_{\delta_2}(c) = U_{\delta_2}(a) \times U_{\delta_2}(b) \subset U.$$

Protože zobrazení φ je spojité v bodě a, lze zvolit $\delta_1 > 0$ tak malé, aby platilo $U_{\delta_1}(a) \subset Z \cap U_{\delta_2}(a)$ a $\varphi(U_{\delta_1}(a)) \subset U_{\delta_2}(b)$. Restrikce $f := \varphi \restriction_{U_{\delta_1}(a)}$ je pak zobrazení $f: U_{\delta_1}(a) \to U_{\delta_2}(b)$, pro které zřejmě platí podmínka $(b)^*$. Protože platí $U_{\delta_1}(a) \times U_{\delta_2}(b) \subset U$, z (2.37) okamžitě vyplývá platnost ekvivalence z $(a)^*$.

2.116 Poznámka. Intuitivně je jasné, že ve větě o implicitních funkcích „uspořádání souřadnic nehraje roli“, takže platí následující tvrzení:

Nechť $1 \le k < n$ jsou přirozená čísla a zobrazení $g = (g_1, \dots, g_{n-k})$ je třídy C^1 na nějakém otevřeném okolí V bodu $c \in \mathbb{R}^n$. Nechť $(i_1, \dots, i_k, j_1, \dots, j_{n-k})$ je přerovnání čísel $(1, \dots, n)$ takové, že $\dfrac{\mathcal{D}(g_1, \dots, g_{n-k})}{\mathcal{D}(x_{j_1}, \dots, x_{j_{n-k}})}(c) \neq 0$. Pak existuje otevřené okolí U bodu $(c_{i_1}, \dots, c_{i_k})$, otevřené okolí W bodu $(c_{j_1}, \dots, c_{j_{n-k}})$ a zobrazení $f: U \to W$ třídy C^1 takové, že pro $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ platí ekvivalence

$$(g(x) = 0 \;\wedge\; (x_{i_1}, \dots, x_{i_k}) \in U \;\wedge\; (x_{j_1}, \dots, x_{j_{n-k}}) \in W) \iff (x_{j_1}, \dots, x_{j_{n-k}}) = f(x_{i_1}, \dots, x_{i_k}).$$

Pro důkaz stačí „přečíslovat souřadnice “. Formálně: definujeme zobrazení $s: \mathbb{R}^n \to \mathbb{R}^n$ předpisem $s(x_1, \dots, x_n) = (x_{i_1}, \dots, x_{i_k}, x_{j_1}, \dots, x_{j_{n-k}})$, položíme $g^* := g \circ s^{-1}$ a použijeme Větu 2.115 na řešení rovnice $g^*(x) = 0$.

2.117 Poznámka. Závěr Věty 2.115 o implicitních funkcích, tj. tvrzení, že existují $\delta_1 > 0$, $\delta_2 > 0$, pro která platí (a*) a (b*), je ekvivalentní následujícímu „formálně slabšímu“ tvrzení:

(S) Existují otevřené okolí W bodu c a zobrazení h z $\mathbb{R}^n$ do $\mathbb{R}^k$ takové, že h je třídy C^p na nějakém otevřeném okolí bodu a a platí ekvivalence:

$$(g(x,y) = 0 \ \wedge \ (x,y) \in W) \iff y = h(x).$$

Platí-li totiž (S), lze snadno nalézt $\delta_2 > 0$ tak malé, aby h bylo třídy C^p na $U_{\delta_2}(a)$ a platilo $U_{\delta_2}(a) \times U_{\delta_2}(b) \subset W$. Protože h je spojité v a, lze zvolit $0 < \delta_1 < \delta_2$, pro které $h(U_{\delta_1}(a)) \subset U_{\delta_2}(b)$. Je snadno vidět, že pro tato δ_1, δ_2 platí (a*) i (b*).

Tuto úvahu, kterou jsme již vlastně provedli v předchozím důkazu, budeme ještě potřebovat.

Někdy se ve formulaci věty o implicitních funkcích uvádí také to, jak se počítá diferenciál implicitně zadaného zobrazení f pomocí parciálních diferenciálů „vazbového" zobrazení g. Tento vzorec nyní stručně odvodíme, přičemž budeme užívat značení z předešlé věty.

Nejprve si uvědomíme, že jacobián $J(x) := \dfrac{D(g_1,\dots,g_k)}{D(y_1,\dots,y_k)}(x)$ je funkce spojitá v bodě c; to vyplývá z vyjádření determinantu pomocí složek matice. Protože $J(c) \neq 0$, je jacobián J nenulový na nějakém okolí Z bodu c. Můžeme tedy zvolit otevřené okolí $U \subset U_{\delta_1}(a)$ bodu a takové, že $\omega(x) := (x, f(x)) \in Z$ pro $x \in U$. Protože $g(x, f(x)) = 0$ pro $x \in U$, platí pro tato x podle Poznámky 2.58, (2.15)

$$(g \circ \omega)'(x) = g'_x(x, f(x)) + g'_y(x, f(x)) \circ f'(x) = 0.$$

Protože parciální diferenciál $g'_y(x, f(x))$ je pro $x \in U$ lineární bijekce $\mathbb{R}^n$ na $\mathbb{R}^n$, dostáváme pro $x \in U$ vzorec

$$f'(x) = -\left(g'_y(x, f(x))\right)^{-1} \circ g'_x(x, f(x)). \tag{2.38}$$

Derivaci $f'(x)$ tedy umíme spočítat pomocí $f(x)$ (tuto hodnotu jsme ovšem schopni spočítat často jen pro $x = a$). Vzorec (2.38), který jsme odvodili nepříliš podrobně, si však není nutno pamatovat. V případech potřeby jej snadno odvodíme a v konkrétních početních příkladech se neužívá: stačí si pamatovat následující přirozenou metodu výpočtu.

Pro implicitně zadané zobrazení $f\colon U_{\delta_1}(a) \to U_{\delta_2}(b)$, $f = (f_1,\dots,f_k)$, z Věty 2.114 je na $U_{\delta_1}(a)$ splněna rovnice $g(x, f(x)) = 0$, tj. je splněn systém k rovnic

$$g_1(x_1,\dots,x_n, f_1(x_1,\dots,x_n),\dots, f_k(x_1,\dots,x_n)) = 0$$

$$\dots$$

$$g_k(x_1,\dots,x_n, f_1(x_1,\dots,x_n),\dots, f_k(x_1,\dots,x_n)) = 0.$$

Funkce n proměnných stojící na levých stranách těchto rovností *jsou tedy nulové na $U_{\delta_1}(a)$, takže mají na této množině nulové všechny parciální derivace.*

Chceme-li vypočítat například $\dfrac{\partial f_i}{\partial x_1}(x)$ pro $i = 1,\dots,k$, derivujeme podle řetízkového pravidla levé strany rovnic podle x_1, čímž dostaneme soustavu k rovnic o k neznámých (pro $x \in U$ nutně s regulární maticí soustavy), kterou vyřešíme.

Pro výpočet derivací druhého řádu funkcí f_i pak stačí derivovat nalezené vzorce pro derivace prvního řádu atd. Nejlépe je osvětlit věc na konkrétním příkladu.

2.118 Příklad. Uvažujme systém tří rovnic

$$x = u + \ln v \qquad y = v - \ln u \qquad z = 2u + 2v.$$

Tyto rovnice („vazbové podmínky“) jsou splněny pro $(x, y, z, u, v) = (1, 1, 4, 1, 1)$. Chápejme nyní x, y jako parametry a z, u, v jako neznámé. Pro hodnotu parametrů $(x, y) = (1, 1)$ má naše soustava rovnic aspoň jedno řešení $(z, u, v) = (4, 1, 1)$.

Ptejme se nyní, zda pro parametry (x, y) blízké bodu $(1, 1)$ bude mít naše soustava opět řešení (blízké bodu $(4, 1, 1)$), a pokud ano, co lze říci o závislosti řešení (z, u, v) na parametrech (x, y); tj. jaké vlastnosti mají funkce $z(x, y), u(x, y), v(x, y)$.

Odpovědi na tyto otázky nám dává Věta 2.114 o implicitních funkcích: vazbové podmínky můžeme napsat ve tvaru

$$\begin{aligned} g_1(x, y, z, u, v) &= x - u - \ln v = 0 \\ g_2(x, y, z, u, v) &= y - v + \ln u = 0 \\ g_3(x, y, z, u, v) &= z - 2u - 2v = 0. \end{aligned}$$

Je tedy $k = 3$, $n = 2$ a píšeme zde (x, y) místo (x_1, x_2) a (z, u, v) místo (y_1, y_2, y_3).

Podmínku (ii) z Věty 2.114 jsme již (pro $a = (1, 1)$, $b = (4, 1, 1), c = (1, 1, 4, 1, 1)$) ověřili a podmínka (i) je splněna pro $p = \infty$: není těžké ověřit, že vazbové funkce g_1, g_2, g_3 jsou třídy C^∞ na $\mathbb{R} \times \mathbb{R} \times \mathbb{R} \times (0, \infty) \times (0, \infty)$.

„Podmínka regularity“ (iii) také platí:

$$\frac{D(g_1, g_2, g_3)}{D(z, u, v)}(c) = \begin{vmatrix} 0 & -1 & -1 \\ 0 & 1 & -1 \\ 1 & -2 & -2 \end{vmatrix} = 2 \neq 0.$$

Existují tedy čísla $\delta_1 > 0, \delta_2 > 0$ taková, že pro parametry $(x, y) \in U_{\delta_1}(1, 1)$ má v $U_{\delta_2}(4, 1, 1)$ systém rovnic právě jedno řešení (z, u, v), přičemž takto definivané funkce $z(x, y)$, $u(x, y)$, $v(x, y)$ jsou na $U_{\delta_1}(1, 1)$ třídy C^∞. Pro každý bod $(x, y) \in U_{\delta_1}(1, 1)$ tedy platí

$$\begin{aligned} x - u(x, y) - \ln v(x, y) &= 0 \\ y - v(x, y) + \ln u(x, y) &= 0 \\ z(x, y) - 2u(x, y) - 2v(x, y) &= 0 \end{aligned}$$

Parciální derivace funkcí na levých stranách jsou tedy na $U_{\delta_1}(1, 1)$ nulové, takže na tomto okolí platí tyto systémy rovnic (*) :

$$\begin{aligned} 1 - u_x - \frac{v_x}{v} &= 0 \qquad & -u_y - \frac{v_y}{v} &= 0 \\ 0 - v_x + \frac{u_x}{u} &= 0 \qquad & 1 - v_y + \frac{u_y}{u} &= 0 \\ z_x - 2u_x - 2v_x &= 0 \qquad & z_y - 2u_y - 2v_y &= 0 \end{aligned}$$

Dosadíme-li $(x, y) = (1, 1)$ a použijeme $u(1, 1) = v(1, 1) = 1$, snadno spočteme

$$u_x(1, 1) = v_x(1, 1) = 1/2,\ z_x(1, 1) = 2,\ u_y(1, 1) = -1/2,\ v_y(1, 1) = 1/2,\ z_y(1, 1) = 0.$$

Můžeme tedy například psát (viz Poznámka 2.21)

$$u'(1, 1) = \mathrm{d}u(1, 1) = \frac{1}{2}\mathrm{d}x - \frac{1}{2}\mathrm{d}y. \tag{2.39}$$

Z rovnic (*) můžeme vypočítat parciální derivace prvního řádu funkcí u, v, z na nějakém otevřeném okolí bodu $(1,1)$ (případně menším, než je $U_{\delta_1}(1,1)$) pomocí funkcí u, v, z. Derivováním těchto vzorců pak postupně můžeme (na nějakém otevřeném okolí bodu $(1,1)$) vyjádřit parciální derivace všech řádů funkcí u, v, z pomocí funkcí u, v, z; speciálně tedy můžeme najít hodnoty všech parciálních derivací těchto funkcí v bodě $(1,1)$.

Z rovnic (*) například dostáváme, že $u_x = \dfrac{uv}{1+uv}$ na nějakém otevřeném okolí bodu $(1,1)$. (Takové okolí existuje, protože funkce $1+uv$ je v bodě $(1,1)$ nenulová a je v tomto bodě spojitá.) Na tomto okolí tedy platí

$$u_{yx} = \left(1 - \frac{1}{1+uv}\right)_y = \frac{u_y v + u v_y}{(1+uv)^2}.$$

Vypočteme-li z rovnic (*) u_y a v_y (případně na ještě menším otevřeném okolí bodu $(1,1)$), po dosazení dostáváme na tomto okolí vyjádření funkce u_{yx} pomocí funkcí u, v, z.

Ná základě znalosti diferenciálů $dz(1,1)$, $du(1,1)$, $dv(1,1)$ jsme schopni vyslovit „věrohodnou domněnku“ o přibližné hodnotě řešení z^*, u^*, v^* pro parametry (x^*, y^*) „velmi blízké“ $(1,1)$, například pro hodnoty $x^* = 1 + 4\cdot 10^{-6}, y^* = 1 - 2\cdot 10^{-6}$. Protože čísla $4\cdot 10^{-6}, -2\cdot 10^{-6}$ se zdají být „dostatečně malá“, jsme například vedeni k přibližnému výpočtu podle (2.39) (ve kterém nahrazujeme diferenci příslušnou hodnotou diferenciálu)

$$\begin{aligned} u^* &:= u(1+4\cdot 10^{-6}, 1-2\cdot 10^{-6}) \approx u(1,1) + du(1,1)(4\cdot 10^{-6}, -2\cdot 10^{-6}) \\ &= 1 + \frac{1}{2}4\cdot 10^{-6} - \frac{1}{2}(-2\cdot 10^{-6}) = 1 + 3\cdot 10^{-6}. \end{aligned}$$

To je ovšem pouze „věrohodná domněnka“; ve skutečnosti (bez dodatečných úvah) nevíme, zda $\|(4\cdot 10^{-6}, -2\cdot 10^{-6})\| < \delta_1$, takže ani nevíme, zda je vůbec hodnota $u(1+4\cdot 10^{-6}, 1-2\cdot 10^{-6})$ definovaná, natož abychom byli schopni udělat odhad chyby.

Pokud spočteme i druhé diferenciály funkcí z, u, v v bodě $(1,1)$, jsme schopni použít Taylorovu formuli pro ještě lepší (ale stále pouze „pravděpodobný“) odhad hodnot z^*, u^*, v^*.

2.10 Difeomorfismus a regulární zobrazení

2.119 Definice. *Řekneme, že zobrazení F z $\mathbb{R}^n$ do $\mathbb{R}^n$ je difeomorfismus, jestliže F je taková bijekce otevřené množiny $U \subset \mathbb{R}^n$ na otevřenou množinu $V \subset \mathbb{R}^n$, že $F \in C^1(U)$ a $F^{-1} \in C^1(V)$.*

Obecněji, jestliže platí, že $F \in C^p(U)$ a $F^{-1} \in C^p(V)$, (kde $p \in \mathbb{N} \cup \{\infty\}$), řekneme, že F je difeomorfismus třídy C^p.

2.120 Poznámka.

(i) Pojmy „difeomorfismus třídy C^1“ a „difeomorfismus“ jsou totožné.

(ii) Z Věty 2.126 níže ihned vyplývá, že F je difeomorfismus třídy C^p, právě když zobrazení F je difeomorfismus a zároveň je třídy C^p. Zavedená terminologie tedy nemůže vést k nedorozumění.

(iii) Zatímco homeomorfismus (otevřené množiny $U \subset \mathbb{R}^n$ na otevřenou množinu $V \subset \mathbb{R}^n$) lze intuitivně chápat jako „spojitou deformaci“ části prostoru, u difeomorfismu jde o jakousi „hladkou“ deformaci. Speciálně „hladké křivky“ se difeomorfismem převádějí na „hladké křivky“ a „hladké plochy dimenze k“ na „hladké plochy dimenze k“ (srov. Poznámka 2.155 (ii)).

2.121 Věta. (o inverzním zobrazení, o lokálním difeomorfismu) *Nechť F je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^n$, $a \in \mathbb{R}^n$, $p \in \mathbb{N} \cup \{0\}$ a platí:*

(i) *F je třídy C^p na nějakém otevřeném okolí bodu a.*

(ii) *$F'(a)$ je lineární bijekce $\mathbb{R}^n$ na $\mathbb{R}^n$ (tj. $J_F(a) \neq 0$).*

Pak existují otevřené okolí U bodu a, na kterém je F difeomorfismus třídy C^p (tj. $F \restriction_U$ je difeomorfismus třídy C^p).

Důkaz. Hrubá myšlenka důkazu je jednoduchá: inverzní zobrazení $F^{-1}(y) = x$ počítáme z rovnice $y = F(x)$, která je ekvivalentní rovnici $h(y,x) := y - F(x) = 0$. To, že z této rovnice lze (lokálně) skutečně počítat x pomocí y, nám umožní Věta 2.112 o implicitních funkcích. Provedení přesného důkazu je početně triviální, ale myšlenkově poměrně náročné.

Na $\mathbb{R}^n$ budeme opět pracovat s maximovou metrikou. Zvolme otevřené okolí $\widetilde{U}$ bodu a, na kterém je F třídy C^p. Pro $(y,x) \in \mathbb{R}^n \times \widetilde{U}$ položme $h(y,x) := y - F(x)$. Protože

$(i)^*$ $h(b,a) = 0$,

$(ii)^*$ h je třídy C^p na $\mathbb{R}^n \times \widetilde{U}$,

$(iii)^*$ $(h(b,\cdot))'(a) = (b - F(\cdot))'(a) = -F'(a)$ je lineární bijekce $\mathbb{R}^n$ na $\mathbb{R}^n$,

lze aplikovat Větu 2.112 o implicitních funkcích. Existují tedy čísla $\delta_1 > 0, \delta_2 > 0$ a zobrazení $G: U_{\delta_1}(b) \to U_{\delta_2}(a)$ třídy C^p takové, že $U_{\delta_1}(b) \times U_{\delta_2}(a) \subset \mathbb{R}^n \times \widetilde{U}$ (takže $U_{\delta_2}(a) \subset \widetilde{U}$) a

$$(h(y,x) = 0 \quad \wedge \quad (y,x) \in U_{\delta_1}(b) \times U_{\delta_2}(a)) \iff x = G(y). \tag{2.40}$$

Položme $V := U_{\delta_1}(b)$ a $U := G(V)$; zřejmě $U \subset U_{\delta_2}(a)$. Je-li nyní $y \in V$ a $x = G(y)$, podle (2.40) máme $h(y,x) = 0$, takže $F(x) = y$. Zobrazení $G: V \to U$ je surjektivní a $F(G(y)) = y$ pro každé $y \in V$. Je tedy G prosté a $G^{-1} = F \restriction_U$, takže $G = (F \restriction_U)^{-1}$; dále zřejmě $V = F(U)$ a $a \in U$. Protože G je třídy C^p na V, $F \in C^p(\widetilde{U})$ a $U \subset \widetilde{U}$, zbývá jen ověřit, že U je otevřená množina. K tomu ukážeme, že

$$U = \{x \in U_{\delta_2}(a)\colon\ F(x) \in V\}. \tag{2.41}$$

Protože $F(U) = V$ a $U \subset U_{\delta_2}(a)$, je U částí množiny napravo. Naopak, leží-li bod x v množině napravo a položíme $y := F(x)$, platí podle (2.40) $x = G(y)$, a tedy

$x \in U$. Protože zobrazení F je spojité na $U_{\delta_2}(a)$, z rovnosti (2.41) plyne, že U je relativně otevřená množina v otevřené množině $U_{\delta_2}(a)$, a proto je otevřená v $\mathbb{R}^n$ (srov. Tvrzení 1.22).

2.122 Poznámka. „Z lineární algebry víme, že podmínka regularity“ (ii) z Věty 2.121 je ekvivalentní regularitě Jacobiho matice $[F'(a)]$ a tedy také nenulovosti jacobiánu $J_F(a)$.

Důležitým zobecněním difeomorfismů jsou regulární zobrazení.

2.123 Definice. (regulární zobrazení) *Nechť $G \subset \mathbb{R}^n$ je otevřená množina a je dáno zobrazení $F: G \to \mathbb{R}^n$. Řekneme, že F je regulární zobrazení, jestliže*

(i) $F \in C^1(G)$ *a*
(ii) $J_F(x) \neq 0$ *pro každý bod* $x \in G$.

2.124 Tvrzení. *Nechť F je regulární zobrazení (třídy C^p) na otevřené množině $G \subset \mathbb{R}^n$. Pak platí:*

(i) *Pro každý bod $a \in G$ existuje jeho otevřené okolí $U \subset G$ takové, že restrikce $F\restriction_U$ je difeomorfismus (třídy C^p).*
(ii) *$F(G)$ je otevřená množina.*

Důkaz. Tvrzení (i) je okamžitým důsledkem Věty 2.121 (srov. Poznámka 2.122).

Zvolme nyní libovolný bod $b \in F(G)$. Nalezněme $a \in G$, pro který $F(a) = b$ a jeho otevřené okolí U podle (i). Pak $F(U) = (F\restriction_U)(U)$ je otevřená množina a $b \in F(U) \subset F(G)$, takže b je vnitřním bodem množiny $F(G)$.

2.125 Poznámka. Někteří autoři formulují větu o inverzním zobrazení jako Tvrzení 2.124. Z tohoto tvrzení snadno plyne i zde uvedená verze (Věta 2.121). Z toho, že $J_F(a) \neq 0$ a F je třídy C^1 na okolí bodu a totiž snadno vyplývá (ze spojitosti funkce J_F), že jacobián J_F je nenulový i na nějakém okolí bodu a.

Následující snadný důsledek Věty 2.121 dává velmi užitečnou charakterizaci difeomorfismů.

2.126 Věta. *Nechť je dána otevřená množina $G \subset \mathbb{R}^n$, zobrazení $F: G \to \mathbb{R}^n$ a $p \in \mathbb{N} \cup \{\infty\}$. Potom platí:*

(i) *Zobrazení F je difeomorfismus, právě když je regulární a prosté.*
(ii) *Zobrazení F je difeomorfismus třídy C^p, právě když je regulární, prosté a třídy C^p.*

Důkaz. Předpokládejme, že F je difeomorfismus. Protože $F^{-1} \circ F$ je identické zobrazení $I: G \to G$, pro každý bod $x \in G$ podle Důsledku 2.51 platí

$$1 = J_I(x) = J_{F^{-1}}(F(x)) \cdot J_F(x),$$

takže $J_F(x) \neq 0$. Každý difeomorfismus je tedy regulární zobrazení.

Nyní předpokládejme, že F je regulární, prosté a třídy C^p. Podle Tvrzení 2.124 je $H := F(G)$ otevřená množina. Pro libovolný bod $b \in H$ označme $a := F^{-1}(b)$ a najděme otevřenou množinu U, $a \in U \subset G$ podle Tvrzení 2.124 (i). Pak zobrazení $(F \restriction_U)^{-1} = F^{-1} \restriction_{F(U)}$ je třídy C^p na otevřené množině $F(U)$, která obsahuje bod b. Je tedy F^{-1} lokálně třídy C^p, a proto je také („globálně") třídy C^p. Tím jsou obě tvrzení (i), (ii) dokázána.

V aplikacích se často vyskytují zobrazení, která jsou regulární, ale nejsou prostá. Typickým příkladem je zobrazení $F\colon (r, \varphi) \mapsto (r\cos\varphi, r\sin\varphi)$, které se používá při zavádění polárních souřadnic (viz Příklad 2.130). Toto zobrazení je na množině $(0, \infty) \times \mathbb{R}$ regulární, ale není na ní prosté.

Větu o inverzním zobrazení jsme dostali jako poměrně snadný důsledek věty o implicitních funkcích. Je však možný i opačný postup: lze nejprve dokázat větu o inverzním zobrazení a z ní pak poměrně snadno odvodit větu o implicitních funkcích.

V Kapitole 3 zobecníme větu o inverzním zobrazení na případ zobrazení mezi Banachovými prostory. Její důkaz je možná snažší než (detailní) důkaz Věty 2.114 o implicitních funkcích, i když je méně elementární (je založen na použití Banachovy věty o kontrakci). Přitom lze tento důkaz číst (v podstatě bez jakýchkoliv znalostí teorie Banachových prostorů; viz Poznámka 3.51) jako důkaz Věty 2.121 o inverzním zobrazení mezi eukleidovskými prostory.

Proto nyní ukážeme, jak lze Větu 2.114 snadno dokázat pomocí Věty 2.121; čtenář pak bude mít k dispozici také moderní důkaz obou vět.

Důkaz věty o implicitních funkcích pomocí věty o inverzním zobrazení.

Nechť jsou splněny předpoklady věty 2.114. Zvolme otevřené okolí V bodu $c = (a, b)$, na kterém je zobrazení $g = (g_1, \ldots, g_k)$ třídy C^p a uvažujme zobrazení $r\colon V \to \mathbb{R}^n \times \mathbb{R}^k$, $r(x, y) = (x, g(x, y))$, tj.

$$\begin{aligned} &r(x_1, \ldots, x_n, y_1, \ldots, y_k) = \\ &\qquad = (x_1, \ldots, x_n, g_1(x_1, \ldots, x_n, y_1, \ldots, y_k), \ldots, g_k(x_1, \ldots, x_n, y_1, \ldots, y_k)). \end{aligned}$$

Podle Věty 2.76 je zobrazení r třídy C^p na V a jeho Jacobiho matice (v jejímž zápisu hvězdičky označují zcela libovolná reálná čísla)

$$[r'(c)] = \begin{bmatrix} 1 & 0 & \ldots & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 1 & 0 & \ldots & 0 \\ * & \ldots & \ldots & * & \frac{\partial g_1}{\partial y_1}(c) & \ldots & \frac{\partial g_1}{\partial y_k}(c) \\ \vdots & & & \vdots & \vdots & & \vdots \\ * & \ldots & \ldots & * & \frac{\partial g_k}{\partial y_1}(c) & \ldots & \frac{\partial g_k}{\partial y_k}(c) \end{bmatrix}$$

je zřejmě regulární. (Snadno např. ověříme, že její řádky jsou lineárně nezávislé.) Podle Věty 2.121 o inverzním zobrazení existuje otevřené okolí U bodu c, na kterém je r difeomorfismus třídy C^p.

Tedy $s := (r\restriction_U)^{-1}$ je zobrazení třídy C^p definované na nějakém otevřeném okolí W bodu $d := r(c) = (a, g(c)) = (a, 0)$. Je-li $(x,y) \in W$ a $(u,v) = s(x,y)$, pak $(x,y) = r(u,v) = (u, g(u,v))$, takže $u = x$. Zobrazení s je tedy tvaru

$$s(x,y) = (x, h(x,y)), \quad (x,y) \in W,$$

kde $h\colon W \to \mathbb{R}^k$ je třídy C^p. Protože $s = (r\restriction_U)^{-1}$, jsou (pro $x \in \mathbb{R}^n$, $y \in \mathbb{R}^k$) následující tvrzení ekvivalentní.

(i) $\quad (x,y) \in U, \quad g(x,y) = 0,$
(ii) $\quad (x,y) \in U, \quad r(x,y) = (x,0),$
(iii) $\quad (x,0) \in W, \quad s(x,0) = (x,y),$
(iv) $\quad (x,0) \in W, \quad h(x,0) = y.$

Položme $\quad Z := \{x \in \mathbb{R}^n\colon (x,0) \in W\}; \quad f(x) := h(x,0), \quad x \in Z.$
Zřejmě $a \in Z$, Z je otevřená množina a $f \in C^p(Z)$. Z ekvivalence $(i) \iff (iv)$ ihned dostáváme, že

$$((x,y) \in U \ \wedge \ g(x,y) = 0) \quad \iff \quad y = f(x).$$

Z této ekvivalence a Poznámky 2.117 vyplývá závěr Věty 2.114 o implicitních funkcích (platnost tvrzení (a), (b), resp. $(a)^*$, $(b)^*$).

2.11 Křivočaré souřadnice

Řekneme, že na množině M je zadán *souřadnicový systém* (případně souřadný systém, souřadný systém dimenze n, apod.), je-li zadáno prosté zobrazení $s\colon M \to \mathbb{R}^n$. Píšeme-li $s(x) = (s_1(x), \dots, s_n(x))$ pro $x \in M$, říkáme číslům $s_1(x), \dots, s_n(x)$ souřadnice bodu x a funkce $s_1, \dots, s_n$ se nazývají souřadnicové funkce (nebo také krátce „souřadnice").

Nás zde bude zajímat jen případ, kdy $M \subset \mathbb{R}^n$ je otevřená množina a zobrazení s je difeomorfismus.

2.127 Definice. *Nechť $G \subset \mathbb{R}^n$ je otevřená množina. Řekneme, že jsme na G zavedli křivočaré souřadnice, jestliže jsme na G zadali souřadný systém $s = (s_1, \dots, s_n)$, $s\colon G \to \mathbb{R}^n$, který je difeomorfismus. Souřadnice $s_1(x), \dots, s_n(x)$ pak nazýváme křivočarými souřadnicemi bodu $x \in G$.*

2.128 Poznámka.

(i) Název „křivočaré souřadnice“ odpovídá tomu, že množiny bodů, na kterých se mění jen jedna souřadnice a ostatní jsou konstantní, tedy (neprázdné) množiny tvaru

$$K = \{x \in G\colon\ s_k(x) = c_k\} \quad \text{pro všechna} \quad k \in \{1, \dots, n\} \setminus \{i\},$$

jsou pro $1 \leq i \leq n$ a $c_k \in \mathbb{R}$ „křivé čáry“(v terminologii oddílu 2.13 to jsou 1-rozměrné C^1 plochy). Je-li s kartézský (nebo afinní) souřadný systém, pak jsou množiny tohoto tvaru „rovné čáry“ tj. části přímek. (To se ale může přihodit i pro jiné difeomorfismy s, např. $s(x,y) := (x^2, y^2)$, $(x,y) \in (0,\infty)^2$.

(ii) Již nyní poznamenejme, že v aplikacích bývají křivočaré souřadnice často určeny explicitním zadáním inverzního zobrazení s^{-1}, přičemž s výpočtem vzorců pro zobrazení s bývají problémy a při aplikaci je často vůbec nepotřebujeme. Tak je tomu při aplikaci důležitých křivočarých souřadnic, kterými jsou polární souřadnice (v $\mathbb{R}^2$), válcové souřadnice (v $\mathbb{R}^3$) a sférické souřadnice (v $\mathbb{R}^3$); viz Příklady 2.130, 2.131, 2.132.

Zavádění vhodných křivočarých souřadnic je velmi důležité: bez jejich zavedení bychom často nebyli schopni spočítat vícerozměrný integrál či vyřešit parciální diferenciální rovnici. V těchto aplikacích používáme křivočaré souřadnice jako „nové nezávislé proměnné pro reálné funkce“.

O co jde a proč to bývá výhodné, osvětlí snad následující „fyzikální“ příklad.

2.129 Příklad. Zaveďme v prostoru kartézské souřadnice a do počátku umístěme zdroj tepla (bodový s konstantní intenzitou). Předpokládejme, že zbytek prostoru je vyplněn homogenním prostředím (např. vzduchem) a teplota se již ustálila. Pak můžeme uvažovat funkci $T(x,y,z)$ na množině $G := \mathbb{R}^3 \setminus \{(0,0,0)\}$, která má význam teploty v bodě o souřadnicích (x,y,z). Intuitivně je jasné, že funkce T je „rotačně symetrická“ tj. závisí jen na vzdálenosti od počátku: je tvaru $T(x,y,z) = f(\sqrt{x^2+y^2+z^2})$. Zdá se být tedy výhodné zavést na G křivočaré souřadnice s_1, s_2, s_3 tak aby jedna z nich měla význam vzdálenosti od počátku; například $s_1(x,y,z) = \sqrt{x^2+y^2+z^2}$. V tomto a podobných „rotačně symetrických případech“ se často používají sférické souřadnice (r,σ,δ), pro které platí
$r(x,y,z) = \sqrt{x^2+y^2+z^2}$. Funkce $T^*(r,\sigma,\delta) := T(s^{-1}(r,\sigma,\delta))$ (která se často a ne zcela korektně zapisuje jako $T(r,\sigma,\delta)$; popisuje totiž závislost „teploty T“ v bodě na jeho sférických souřadnicích) pak závisí pouze na jedné proměnné: $T^*(r,\sigma,\delta) = f(r)$. Není tedy divu, že často bývá jednodušší pracovat s funkcí T^* než s funkcí T.

Pro studium funkcí $H(x,y,z)$, které závisí pouze na vzdálenosti od osy z
(tj. $H(x,y,z) = h\left(\sqrt{x^2+y^2}\right)$) je zase vhodné použít tzv. válcové souřadnice, jedna z válcových souřadnic je totiž $r(x,y,z) = \sqrt{x^2+y^2}$.

(V předchozí úvaze jsme situaci úmyslně trochu zjednodušili — sférické souřadnice bohužel nelze zavést na celé množině G; srov. Příklad 2.132.)

V následujících příkladech pojednáme o důležitých křivočarých souřadnicích: polárních, válcových, sférických a polárních v $\mathbb{R}^n$.

2.130 Příklad. (polární souřadnice) Uvažujme zobrazení $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ dané předpisem $F(r,\varphi) = (x,y)$, kde $x = r\cos\varphi$, $y = r\sin\varphi$. Je snadno vidět, že F je třídy C^1 (dokonce třídy

C^∞) a

$$J_F(r,\varphi) = \frac{D(x,y)}{D(r,\varphi)}(r,\varphi) = \begin{vmatrix} \cos\varphi & -r\sin\varphi \\ \sin\varphi & r\cos\varphi \end{vmatrix} = r\cos^2\varphi + r\sin^2\varphi = r.$$

Zobrazení F tedy není regulární na $\mathbb{R}^2$; největší otevřená množina, na které je regulární, je $\mathbb{R}^2 \setminus (\{0\} \times \mathbb{R})$. Při zavádění polárních souřadnic však F uvažujeme pouze na otevřené polorovině $H := (0,\infty) \times \mathbb{R}$. Pro $(r,\varphi) \in H$ zřejmě platí $r = \sqrt{x^2+y^2}$, tj. r je vzdálenost bodu $(x,y) := F(r,\varphi)$ od počátku.

Vidíme tedy, že počátek $(0,0)$ nepatří do množiny $F(H)$; kterýkoliv jiný bod $z = (x,y) \in \mathbb{R}^2 \setminus\{(0,0)\}$ již ale v $F(H)$ leží. Ztotožníme-li z s komplexním číslem $x+iy$ a zapíšeme-li je v goniometrickém tvaru, máme $z = r(\cos\varphi + i\sin\varphi)$, tj. $(x,y) = F(r,\varphi)$. Víme, že $r = \sqrt{x^2+y^2}$, ale argument (amplituda) φ není určen jednoznačně; jednoznačně (podmínkou $\varphi \in (-\pi,\pi]$) je však určena jeho „hlavní hodnota" $\operatorname{Arg} z$.

Každý bod $z = (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$ má nekonečně mnoho vzorů:

$$F^{-1}(\{z\}) = \{(|z|, \operatorname{Arg} z + 2k\pi) \colon\ k \in \mathbb{Z}\}.$$

Regulární zobrazení $F\restriction_H$ tedy není prosté. Jestliže zúžíme F na množinu $P := (0,\infty)\times[0,2\pi)$, je $F\restriction_P$ prosté zobrazení, které má obor hodnot $\mathbb{R}^2 \setminus \{(0,0)\}$; jeho definiční obor však není otevřená množina. Dokonce lze dokázat, že neexistuje otevřená množina $G \subset H$, na které je F prosté a $F(G) = \mathbb{R}^2 \setminus \{(0,0)\}$.

Jestliže zúžíme F na otevřenou množinu $P^* = (0,\infty) \times (0,2\pi)$, je výsledné zobrazení $F^* = F\restriction_{P^*}$ difeomorfismus, ale jeho obor hodnot je pouze $\mathbb{R}^2 \setminus A$, kde $A = [0,\infty) \times \{0\}$. Při aplikacích v torii integrálu stačí pracovat se zúžením F^*, protože polopřímka A je z hlediska teorie míry zanedbatelná: její Lebesgueova míra je nulová.

Pro aplikace v diferenciálním počtu nám však vadí, že zobrazení F^* neurčuje křivočaré (polární) souřadnice na okolí bodů množiny A. V těchto aplikacích však pracujeme s lokálními pojmy, a proto nám stačí zavést polární souřadnice v okolí každého bodu $z \in \mathbb{R}^2 \setminus \{(0,0)\}$. Tím se rozumí toto: zvolíme některý z bodů $u = (r,\varphi) \in F^{-1}(\{z\}) \cap H$, podle Věty 2.121 zvolíme okolí U bodu u, pro který je $F\restriction_U$ difeomorfismus, a na okolí $V := F(U)$ bodu z pak zavedeme křivočaré „polární souřadnice" pomocí souřadného systému $s := (F\restriction_U)^{-1}$. Konkrétní volba u a U (a následně s) nemá na výpočty vliv. Označíme-li $s(x,y) =: (r(x,y), \varphi(x,y))$, máme ovšem vždy $r(x,y) = \sqrt{x^2+y^2}$, ale explicitní vzorec pro $\varphi(x,y)$ většinou nepotřebujeme. V případě potřeby můžeme užít následující vzorce:

$$\varphi(x,y) = \operatorname{arctg}\tfrac{y}{x} = \arcsin\tfrac{y}{\sqrt{x^2+y^2}} \quad \text{pro} \quad x>0 \quad (\text{tj. pro } (x,y) \in (0,\infty)\times\mathbb{R}),$$
$$\varphi(x,y) = \operatorname{arctg}\tfrac{y}{x} + \pi = \pi - \arcsin\tfrac{y}{\sqrt{x^2+y^2}} \quad \text{pro} \quad x<0,$$
$$\varphi(x,y) = \operatorname{arccotg}\tfrac{x}{y} = \arccos\tfrac{x}{\sqrt{x^2+y^2}} \quad \text{pro} \quad y>0$$
$$\varphi(x,y) = \operatorname{arccotg}\tfrac{x}{y} - \pi = -\arccos\tfrac{x}{\sqrt{x^2+y^2}} \quad \text{pro} \quad y<0.$$

2.131 Příklad. (válcové souřadnice) Válcové souřadnice (r,φ,Z) se zavádějí rovnicemi

$$x = r\cos\varphi,\ y = r\sin\varphi,\ z = Z.$$

Názorně řečeno: souřadnici z neměníme (jen z formálních důvodů pro „novou" souřadnici používáme písmeno Z) a souřadnice x,y „měníme na příslušné polární souřadnice".

Je snadno vidět, že zobrazení $F(r,\varphi,Z) = (r\cos\varphi, r\sin\varphi, Z)$ je třídy C^∞. Jeho jacobián je pochopitelně „stejný" jako u polárních souřadnic:

$$J_F(r,\varphi,Z) = \begin{vmatrix} \cos\varphi & -r\sin\varphi & 0 \\ \sin\varphi & r\cos\varphi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r.$$

Obr. 2.10. Sférické souřadnice.

Zúžení zobrazení F na množinu $P := (0,\infty)\times\mathbb{R}\times\mathbb{R}$ je tedy regulární; z předchozího příkladu (o polárních souřadnicích) vidíme, že $F(P)=\mathbb{R}^3\setminus L$, kde $L=\{0\}\times\{0\}\times\mathbb{R}$ je osa z. Na okolí každého bodu $a\in\mathbb{R}^3\setminus L$ lze zavést válcové souřadnice tímto způsobem: zvolíme některý z bodů $u=(r,\varphi,Z)\in F^{-1}(\{a\})$, pro který $r>0$, podle Věty 2.121 zvolíme okolí U bodu u, pro který je $F\restriction_U$ difeomorfismus, a na otevřeném okolí $V := F(U)$ bodu a pak zavedeme válcové souřadnice pomocí souřadného systému $s := (F\restriction_U)^{-1}$. Pro aplikace v integrálním počtu stačí pracovat se zúžením zobrazení F na množinu $G := (0,\infty)\times(0,2\pi)\times\mathbb{R}$, které je difeomorfismem množiny G na $\mathbb{R}^3\setminus([0,\infty)\times\{0\}\times\mathbb{R})$.

2.132 Příklad. (sférické souřadnice) Tyto křivočaré souřadnice jsou v podstatě všeobecně známy ze zeměpisu. Nechť je dán bod $a=(x,y,z)\in\mathbb{R}^3$, pro který $(x,y)\neq(0,0)$.

Jeho první sférická souřadnice r má význam jeho vzdálenosti od počátku, tedy $r := \sqrt{x^2+y^2+z^2}$. Nyní uvažujeme sféru $S_r=\{w\in\mathbb{R}^3\colon \|w\|=r\}$, kterou si představíme jako „povrch zeměkoule“ (viz obr. 2.10), a to tak, že

(i) rovina xy je rovina rovníku, $(0,0,r)$ je severní pól a $(0,0,-r)$ je pól jižní,
(ii) uzavřená polorovina $P := [0,\infty)\times\{0\}\times\mathbb{R}$ je „polorovina nultého poledníku“ a
(iii) východní zeměpisná délka bodu $(0,r,0)$ (v obloukové míře) je $\pi/2$.

Při této interpretaci je druhá sférická souřadnice $\sigma\in[-\pi/2,\pi/2]$ bodu a dána jeho „zeměpisnou šířkou“, přičemž znaménko $+$ (resp. $-$) přísluší severní (resp. jižní) šířce.

Podobně je třetí sférická souřadnice $\delta\in(-\pi,\pi]$ bodu a dána jeho zeměpisnou délkou, přičemž znaménko $+$ přísluší východní délce.

Kolmá projekce a^* bodu a do roviny xy má pak vzdálenost od počátku $r\cos\sigma$, takže x-ová souřadnice bodu a^* i bodu a je $x=r\cos\sigma\cos\delta$ a jejich y-ová souřadnice je $y=r\cos\sigma\sin\delta$. Přitom z-ová souřadnice bodu a je zřejmě $z=r\sin\sigma$.

Místo úhlu $\sigma\in[-\pi/2,\pi/2]$ se ale častěji používá úhel $\theta := \pi/2-\sigma\in[0,\pi]$; takže $\sin\theta=\cos\sigma$ a $\cos\theta=\sin\sigma$. Navíc úhel δ je zvykem označovat písemenem φ a uvažovat jej většinou v intervalu $[0,2\pi)$.

Při zavádění sférických souřadnic (neboli polárních souřadnic v prostoru) tedy uvažujeme

sobrazení
$$S^*\colon (r,\sigma,\delta)\mapsto (r\cos\sigma\cos\delta, r\cos\sigma\sin\delta, r\sin\sigma),$$
nebo častěji
$$S\colon (r,\theta,\varphi)\mapsto (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta).$$
Z geometrického názoru je vidět (a lze snadno dokázat), že $S([0,\infty)\times[0,\pi]\times[0,2\pi))=\mathbb{R}^3$. Snadný výpočet dává
$$J_{S^*}(r,\sigma,\delta)=-r^2\cos\sigma,\quad J_S(r,\theta,\varphi)=r^2\sin\theta.$$

(Toto vyjádření jacobiánu lze také snadno odvodit ze znalosti „jacobiánu pro polární souřadnice“; viz následující poznámka.)

Pomocí zobrazení S lze tedy zavést sférické souřadnice na okolí každého bodu, který neleží na ose z. Pro aplikace v integrálním počtu stačí pracovat se zúžením zobrazení S na množinu $G := (0,\infty)\times(0,\pi)\times(0,2\pi)$, které je difeomorfismem množiny G na $\mathbb{R}^3\setminus([0,\infty)\times\{0\}\times\mathbb{R})$.

2.133 Poznámka. Vzorec pro jacobián zobrazení S z předchozího příkladu lze jednoduše získat, uvědomíme-li si, že zobrazení S je v podstatě složením dvou zobrazení, které jsme užili v Příkladu 2.131 při zavádění válcových souřadnic. Položíme-li totiž $P(x,y,z) := (y,z,x)$,
$$F(r,\theta,\varphi) := (r\cos\theta, r\sin\theta,\varphi),\quad F^*(u,v,w) := (u, v\cos w, v\sin w),$$
pak snadno vidíme, že $S=P\circ F^*\circ F$, $J_F(r,\theta,\varphi)=r$, $J_{F^*}(u,v,w)=v$ a $J_P(u,v,w)=1$, takže podle Důsledku 2.51
$$J_S(r,\theta,\varphi)=1\cdot J_{F^*}(F(r,\theta,\varphi))\cdot J_F(r,\theta,\varphi)=(r\sin\theta)\cdot r=r^2\sin\theta.$$

2.134 Příklad. (polární souřadnice v $\mathbb{R}^n$) Na pozorování z předchozí poznámky je založena definice polárních (sférických) souřadnic v $\mathbb{R}^n$, $(n\ge 3)$ při jejichž zavádění uvažujeme zobrazení
$$S\colon (r,\varphi_1,\dots,\varphi_{n-1})\mapsto (x_1,\dots,x_n)$$
zadané takto:
$$\begin{aligned}
x_1 &= r\,\cos\varphi_1\\
x_2 &= r\,\sin\varphi_1\,\cos\varphi_2\\
x_3 &= r\,\sin\varphi_1\,\sin\varphi_2\,\cos\varphi_3\\
&\dots\dots\dots\dots\dots\dots\\
x_{n-2} &= r\,\sin\varphi_1\cdots\sin\varphi_{n-3}\,\cos\varphi_{n-2}\\
x_{n-1} &= r\,\sin\varphi_1\cdots\sin\varphi_{n-3}\,\sin\varphi_{n-2}\,\cos\varphi_{n-1}\\
x_n &= r\sin\varphi_1\cdots\sin\varphi_{n-3}\,\sin\varphi_{n-2}\,\sin\varphi_{n-1}.
\end{aligned}$$
Je snadno vidět, že
$$S=V_{n-1,n}\circ V_{n-2,n-1}\circ\dots\circ V_{2,3}\circ V_{1,2},$$
kde $V_{i,i+1}$ je zobrazení, při kterém se na dvojici souřadnic (t_i,t_{i+1}) aplikuje zobrazení F, kterým se zavádějí polární souřadnice, a ostatní souřadnice se nemění, tj.
$$V_{i,i+1}(t_1,\dots,t_i,t_{i+1},\dots,t_n) := (t_1,\dots,t_i\cos t_{i+1}, t_i\sin t_{i+1},\dots,t_n).$$
Z tohoto vyjádření pak dostáváme
$$\begin{aligned}
J_S(r,\varphi_1,\dots,\varphi_{n-1}) &= r\cdot(r\,\sin\varphi_1)\cdot(r\,\sin\varphi_1\,\sin\varphi_2)\cdot(r\,\sin\varphi_1\cdots\sin\varphi_{n-3}\,\sin\varphi_{n-2})\\
&= r^{n-1}\,\sin^{n-2}\varphi_1\,\sin^{n-3}\varphi_2\cdots\sin\varphi_{n-2}.
\end{aligned}$$

Pro aplikace v integrálním počtu stačí uvažovat (difeomorfní) zúžení zobrazení S na množinu $(0,\infty)\times(0,\pi)^{n-2}\times(0,2\pi)$. Pro aplikace v diferenciálním počtu lze pomocí zobrazení S zavést „n-rozměrné“ polární souřadnice v okolí každého bodu $x\in\mathbb{R}^n$, který je tvaru $x=S(r,\varphi_1,\ldots,\varphi_{n-1})$, kde $J_S(r,\varphi_1,\ldots,\varphi_{n-1})\neq 0$.

2.12 Záměna proměnných

Předpokládejme, že na otevřené množině $G\subset\mathbb{R}^n$ je zadán souřadný systém (tj. difeomorfismus) $S=(S_1,\ldots,S_n)$, $S\colon G\to\mathbb{R}^n$. Pak $H\ :=S(G)$ je otevřená množina. Difeomorfismus S^{-1} udává závislost „starých“ kartézských souřadnic (které značíme $x_1,\ldots,x_n$) na „nových“ křivočarých souřadnnicích $s_1,\ldots,s_n$; jeho složky je proto přirozené označovat $X_1(s_1,\ldots,s_n),\ldots,X_n(s_1,\ldots,s_n)$. (Pro přesnost výkladu zde souřadnicové funkce značíme velkými písmeny, abychom je v dalším, poměrně složitém výkladu, odlišili od příslušných souřadnic konkrétních bodů.
V početní praxi souřadnicové funkce a souřadnice zpravidla označujeme stejnými písmeny.)

Kdykoliv je nyní $z\colon G\to\mathbb{R}$ reálná funkce, budeme symbolem z^* označovat funkci $z^*\ :=z\circ S^{-1}$. Funkce $z^*\colon H\to\mathbb{R}$ je tedy zadána předpisem

$$z^*(s_1,\ldots,s_n)=z\left(X_1(s_1,\ldots,s_n),\ldots,X_n(s_1,\ldots,s_n)\right).$$

Pokud $n=3$ a $z(x_1,\ldots,x_n)$ má význam velikosti nějaké fyzikální veličiny (třeba teploty) v bodě $(x_1,\ldots,x_n)$, vyjadřuje funkce z^* závislost této veličiny na křivočarých souřadnicích $s_1,\ldots,s_n$ tohoto bodu. Fyzici, a někdy i matematici, proto (ne zcela korektně) píší $z(s_1,\ldots,s_n)$ místo $z^*(s_1,\ldots,s_n)$.

Fyzikální zákony lze velmi často formulovat ve tvaru parciálních diferenciálních rovnic. V obecné parciální diferenciální rovnici k-tého řádu pro neznámou funkci $z(x)=z(x_1,\ldots,x_n)$ se kromě hodnoty $z(x)$ a čísel $x_1,\ldots,x_n$ mohou vyskytovat i hodnoty parciálních derivací funkce z v bodě x až do řádu k.

Obecné parciální diferenciální rovnice prvního a k-tého řádu mají tvar

$$f\left(x_1,\ldots,x_n,z(x),\frac{\partial z}{\partial x_1}(x),\ldots,\frac{\partial z}{\partial x_n}(x)\right)=0$$

a

$$F\left(x_1,\ldots,x_n,z(x),\frac{\partial z}{\partial x_1}(x),\ldots,\frac{\partial z}{\partial x_n}(x),\frac{\partial^2 z}{\partial x_1^2}(x),\ldots,\frac{\partial^k z}{\partial x_n^k}(x)\right)=0 \tag{2.42}$$

Má-li například $z(x)$ význam ustálené teploty v bodě x, lze ukázat, že funkce z musí splňovat (Laplaceovu) parciální diferenciání rovnici druhého řádu

$$\frac{\partial^2 z}{\partial x_1^2}(x)+\frac{\partial^2 z}{\partial x_2^2}(x)+\frac{\partial^2 z}{\partial x_3^2}(x)=0$$

v každém bodě množiny G.

Je přirozené očekávat, že parciální diferenciální rovnici (2.42) pro funkci z lze „ekvivalentně převést“ na parciální diferenciální rovnici pro funkci z^*. Abychom příslušnou rovnici pro z^* napsali, potřebujeme vypočítat souřadnice $x_1, \dots, x_n$ bodu x, hodnotu $z(x)$ a hodnoty parciálních derivací (až do řádu k) funkce z v bodě x pomocí křivočarých souřadnic $S_1(x), \dots, S_n(x)$ bodu x, hodnoty $z^*(S(x))$ a hodnot parciálních derivací funkce z^* v bodě $S(x)$. Pak již stačí tato vyjádření dosadit do (2.42), místo $S_i(x)$ psát s_i a dostaneme příslušnou rovnici pro z^*.

Dále budeme předpokládat, že máme v „explicitním tvaru“ zadáno zobrazení S^{-1}, tj. jsou dány vzorce pro funkce $X_1(s_1, \dots, s_n), \dots, X_n(s_1, \dots, s_n)$. (Pak ovšem můžeme „explicitně vyjádřit“ také všechny parciální derivace funkcí $X_1, \dots, X_n$.) Tak tomu bývá v nejdůležitějších případech (užití polárních, válcových a sférických souřadnic) a navíc: pokud se v rovnici (2.42) vyskytují složky x (tj. funkce F závisí na $x_1, \dots, x_n$), pak vzorce pro S^{-1} nutně potřebujeme.

Předpokládejme nyní, že funkce z (a tedy i z^*, viz Věta 2.76) je třídy C^1. Pak můžeme vypočítat parciální derivace prvního řádu funkce z pomocí parciálních derivací prvního řádu funkce z^*, tj. získat „explicitní vyjádření“

$$\frac{\partial z}{\partial x_i}(x) = h_i\left(S_1(x), \dots, S_n(x), \frac{\partial z^*}{\partial s_1}(S(x)), \dots, \frac{\partial z^*}{\partial s_n}(S(x))\right) \tag{2.43}$$

následujícím způsobem.

Derivováním rovnice $z^*(s_1, \dots, s_n) = z(X_1(s_1, \dots, s_n), \dots, X_n(s_1, \dots, s_n))$ podle proměnných $s_1, \dots, s_n$ získáme systém n rovnic (ve kterých značíme $s = (s_1, \dots, s_n)$ a $x = (X_1(s), \dots, X_n(s))$)

$$\begin{gathered}
\frac{\partial z^*}{\partial s_1}(s) = \frac{\partial z}{\partial x_1}(x)\frac{\partial X_1}{\partial s_1}(s) + \dots + \frac{\partial z}{\partial x_n}(x)\frac{\partial X_n}{\partial s_1}(s) \\
\dots \\
\frac{\partial z^*}{\partial s_n}(s) = \frac{\partial z}{\partial x_1}(x)\frac{\partial X_1}{\partial s_n}(s) + \dots + \frac{\partial z}{\partial x_n}(x)\frac{\partial X_n}{\partial s_n}(s),
\end{gathered}$$

z nichž vypočítáme (například pomocí Cramerova pravidla; matice soustavy je zřejmě regulární) „neznámé“ $\frac{\partial z}{\partial x_1}(x), \dots, \frac{\partial z}{\partial x_n}(x)$.

Za předpokladu, že souřadný systém S i funkce z (a tedy i z^*) jsou třídy C^2, bychom mohli každou z těchto n rovnic derivovat podle $s_1, \dots, s_n$ a z obdržených n^2 rovnic nalézt vhodné vyjádření pro parciální derivace druhého řádu funkce z.

Naštěstí je věc podstatně jednodušší, protože můžeme použít následující „trik“, který spočívá v tom, že použijeme vzorec (2.43) na funkci $\frac{\partial z}{\partial x_j}(x)$ místo na funkci $z(x)$. (Na tomto místě lze doporučit čtenáři, aby se před dalším čtením seznámil s použitím tohoto triku na konkrétním příkladě 2.135.)

Pak pro $i = 1, \dots, n$ dostáváme

(2.44)
$$\frac{\partial^2 z}{\partial x_i \partial x_j}(x) = h_i \left(S_1(x), \dots, S_n(x), \frac{\partial \left(\left(\frac{\partial z}{\partial x_j} \right)^* \right)}{\partial s_1}(S(x)), \dots, \frac{\partial \left(\left(\frac{\partial z}{\partial x_j} \right)^* \right)}{\partial s_n}(S(x)) \right).$$

Užijeme-li (2.43) pro $i = j$ a $x = S^{-1}(s)$, dostaneme pro funkci $\left(\frac{\partial z}{\partial x_j} \right)^*$ následující vzorec:

$$\left(\frac{\partial z}{\partial x_j} \right)^* (s) = \frac{\partial z}{\partial x_j} \left(S^{-1}(s) \right) = h_j \left(\frac{\partial z^*}{\partial s_1}(s), \dots, \frac{\partial z^*}{\partial s_n}(s), s_1, \dots, s_n \right).$$

Po zderivování posledního výrazu a dosazení do (2.44) zřejmě dostaneme vyjádření tvaru

$$\begin{aligned} \frac{\partial^2 z}{\partial x_i \partial x_j}(x) = h_{ij} \Big(& S_1(x), \dots, S_n(x), \frac{\partial z^*}{\partial s_1}(S(x)), \dots, \frac{\partial z^*}{\partial s_n}(S(x)), \\ & \frac{\partial^2 z^*}{\partial s_1^2}(S(x)), \frac{\partial^2 z^*}{\partial s_2 \partial s_1}(S(x)), \dots \frac{\partial^2 z^*}{\partial s_n^2}(S(x)), \Big), \end{aligned} \tag{2.45}$$

kde funkce h_{ij} dovedeme explicitně spočítat.

Postupujeme-li stejnou metodou dále, dostaneme (za předpokladu, že S je třídy C^k) pro funkce $z \in C^k(G)$ hledané vyjádření parciálních derivací až do řádu k. Pokud je například z třídy C^3, použijeme „trik“ pro výpočet $\dfrac{\partial^3 z}{\partial x_k \partial x_i \partial x_j}$ tak, že aplikujeme vzorec (2.43) na funkci $\dfrac{\partial^2 z}{\partial x_i \partial x_j}$ místo na funkci z, přičemž pro vyjádření funkce $\left(\dfrac{\partial^2 z}{\partial x_i \partial x_j} \right)^*$ použijeme vzorec (2.45), ve kterém položíme $x = S^{-1}(s)$.

Právě provedená odvozování si není nutno pamatovat, stačí si osvojit metodu výpočtu na následujícím typickém příkladě, ve kterém již používáme obvyklé „licence“: píšeme z místo z^* a nepíšeme body, ve kterých se parciální derivace počítají.

2.135 Příklad. Nechť je dána funkce $z(x, y)$, která je třídy C^1 na nějaké otevřené množině $M \subset \mathbb{R}^2 \setminus \{(0, 0)\}$. Je-li $c \in M$, existuje jeho okolí $G \subset M$, na kterém můžeme zavést polární souřadnice difeomorfismem S. Na $H := S(G)$ pak platí $S^{-1}(r, \varphi) = (r \cos \varphi, r \sin \varphi)$. Derivujeme-li rovnici

$$z(r, \varphi) := z^*(r, \varphi) := z(r \cos \varphi, r \sin \varphi)$$

podle r a také podle φ, dostáváme

$$\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x} \cos \varphi + \frac{\partial z}{\partial y} \sin \varphi, \qquad \frac{\partial z}{\partial \varphi} = -\frac{\partial z}{\partial x} r \sin \varphi + \frac{\partial z}{\partial y} r \cos \varphi. \tag{2.46}$$

(Poznamenejme, že přesný zápis 1. rovnice z (2.46) je

$$\frac{\partial z^*}{\partial r}(r,\varphi) = \frac{\partial z}{\partial x}(r\cos\varphi, r\sin\varphi)\cos\varphi + \frac{\partial z}{\partial y}(r\cos\varphi, r\sin\varphi)\sin\varphi, \quad (r,\varphi)\in H.)$$

Řešením soustavy (2.46) vzhledem k neznámým $\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y}$ dostáváme

$$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial r}\cos\varphi - \frac{1}{r}\frac{\partial z}{\partial \varphi}\sin\varphi, \qquad \frac{\partial z}{\partial y} = \frac{\partial z}{\partial r}\sin\varphi + \frac{1}{r}\frac{\partial z}{\partial \varphi}\cos\varphi. \tag{2.47}$$

Dále předpokládáme, že z je třídy C^2 na G; pak můžeme použít výše popsaný „trik“: první vzorec z (2.47) použijeme na funkci $\frac{\partial z}{\partial x}$ místo na funkci z. Přitom funkci $\frac{\partial z}{\partial x}$ vyjádříme pomocí prvního vzorce z (2.47). Dostáváme

$$\frac{\partial^2 z}{\partial x^2} = \frac{\partial\left(\frac{\partial z}{\partial x}\right)}{\partial x} = \frac{\partial\left(\frac{\partial z}{\partial x}\right)}{\partial r}\cos\varphi - \frac{1}{r}\frac{\partial\left(\frac{\partial z}{\partial x}\right)}{\partial \varphi}\sin\varphi =$$

$$= \frac{\partial}{\partial r}\left(\frac{\partial z}{\partial r}\cos\varphi - \frac{1}{r}\frac{\partial z}{\partial \varphi}\sin\varphi\right)\cos\varphi - \frac{1}{r}\frac{\partial}{\partial \varphi}\left(\frac{\partial z}{\partial r}\cos\varphi - \frac{1}{r}\frac{\partial z}{\partial \varphi}\sin\varphi\right)\sin\varphi =$$

$$= \frac{\partial^2 z}{\partial r^2}\cos^2\varphi - \frac{2}{r}\frac{\partial^2 z}{\partial \varphi\,\partial r}\cos\varphi\sin\varphi + \frac{1}{r^2}\frac{\partial^2 z}{\partial \varphi^2}\sin^2\varphi +$$

$$+\frac{1}{r}\frac{\partial z}{\partial r}\sin^2\varphi + \frac{2}{r^2}\frac{\partial z}{\partial \varphi}\sin\varphi\cos\varphi.$$

Zcela obdobně lze vypočítat $\frac{\partial^2 z}{\partial x\,\partial y}$ a $\frac{\partial^2 z}{\partial y^2}$ a odvodit například vzorec

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = \frac{\partial^2 z}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 z}{\partial \varphi^2} + \frac{1}{r}\frac{\partial z}{\partial r}.$$

Z něho vyplývá, že funkce z spňuje na G důležitou Laplaceovu rovnici

$$\frac{\partial^2 z}{\partial x^2} + \frac{\partial^2 z}{\partial y^2} = 0, \tag{2.48}$$

právě když funkce z^* splňuje na H rovnici

$$\frac{\partial^2 z^*}{\partial r^2} + \frac{1}{r^2}\frac{\partial^2 z^*}{\partial \varphi^2} + \frac{1}{r}\frac{\partial z^*}{\partial r} = 0. \tag{2.49}$$

Z toho můžeme také usoudit, že funkce $z \in C^2(M)$ splňuje na M Laplaceovu rovnici, právě když funkce $z^*(r,\varphi) := z(r\cos\varphi, r\sin\varphi)$ splňuje na $(0,\infty)\times\mathbb{R}$ rovnici (2.49).

2.13 Hladké k-rozměrné plochy v $\mathbb{R}^n$

2.13.1 Úvodní úvahy

Vyložíme pouze základní fakta teorie k-rozměrných ploch, která umožňují přirozený a elegantní výklad teorie „vázaných extrémů“.

Začněme „orientační“ diskuzí následující otázky: Které podmnožiny $\mathbb{R}^2$ je přirozené považovat za hladké 1-rozměrné plochy („křivé čáry“) v $\mathbb{R}^2$?

Naší intuitivní představě jistě vyhovuje libovolná přímka; dále také asi graf „hladké“ funkce f (tedy množina $\{(x,y)\colon\ y = f(x)\}$ a také „graf“ hladké funkce $x = g(y)$ (přesněji: množina $\{(x,y)\colon\ x = g(y)\}$). Naší představě však také asi vyhovuje kružnice $\{(x,y)\colon\ x^2 + y^2 = 1\}$. Ta však není ani grafem funkce $y = f(x)$ ani „grafem“ funkce $x = g(y)$, je však „lokálně“ jednoho z těchto typů.

Jako přirozená se tedy jeví „lokální definice“: hladkou 1-rozměrnou plochu budeme definovat jako takovou množinu $M \subset \mathbb{R}^2$, pro jejíž každý bod $x \in M$ existuje otevřené okolí U bodu x takové, že $U \cap M$ je „speciální kus hladké 1-rozměrné plochy“: „hladce zdeformovaná otevřená úsečka“. Pod tímto pojmem můžeme rozumět množinu, která je grafem funkce $y = f(x), x \in (a,b)$ nebo $x = g(y)$, $y \in (c,d)$; jsou však možné i jiné přirozené definice (například obraz úsečky při „hladké deformaci roviny“ nebo obraz intervalu (a,b) při vhodném hladkém zobrazení $\varphi\colon\ (a,b) \to \mathbb{R}^2$).

Tyto přístupy v obecném případě vedou k různým pojmům „kusu k-rozměrné C^p plochy“ v $\mathbb{R}^n$. Je důležitou a zajímavou skutečností, že tyto *různé přístupy* vedou ke *stejnému pojmu* k-rozměrné C^p plochy v $\mathbb{R}^n$.

Úmluva Použijeme-li kdekoliv v dalším výkladu symbol C^p, rozumí se, že $p \in \mathbb{N} \cup \{\infty\}$ je libovolné pevně zvolené číslo. Není-li řečeno jinak, jsou v dalším k, n přirozená čísla, pro která $k < n$.

2.13.2 Různé typy „kusů ploch“

Explicitně zadaný kus plochy

Zhruba řečeno, jde o množinu, která je zadaná jako „graf funkce (zobrazení)“. Jinak a trochu přesněji: pro určení bodu v explicitně zadaném kusu k-rozměrné plochy stačí znát jeho jistých k souřadnic; ostatních $n - k$ souřadnic pak můžeme jednoznačně dopočítat.

2.136 Definice. *Množina $M \subset \mathbb{R}^n$ je explicitně zadaný kus k-rozměrné C^p plochy v $\mathbb{R}^n$ $(1 \leq k < n)$, jestliže existují přirozená čísla $1 \leq i_1 < i_2 < \cdots < i_k \leq n$, neprázdná otevřená množina $G \subset \mathbb{R}^k$ a zobrazení $f\colon\ G \to \mathbb{R}^{n-k}$ třídy C^p takové, že*

$$M = \{(x_1, \ldots, x_n)\colon\ (x_{j_1}, \ldots, x_{j_{n-k}}) = f(x_{i_1}, \ldots, x_{i_k})\},$$

Obr. 2.11.

kde $x_{j_1}, \dots, x_{j_{n-k}}$ *jsou „zbylé souřadnice“, přesněji:* $j_1 < j_2 < \dots < j_{n-k}$ *a* $\{j_1, \dots, j_{n-k}\} = \{1, \dots, n\} \setminus \{i_1, \dots, i_k\}$.

2.137 Příklad.

a) Horní polosféra $\{(x,y,z)\colon\ x^2 + y^2 < 1,\ z = \sqrt{1 - x^2 - y^2}\}$ znázorněná na obr. 2.11 vlevo je explicitně zadaný kus 2-rozměrné C^∞ plochy v $\mathbb{R}^3$.

b) Šroubovice $\{(x,y,z)\colon\ x = r\cos cz,\ y = r\sin cz\}$ $(r > 0, c \neq 0)$ (srov. obr. 2.11 vpravo) je explicitně zadaný kus 1-rozměrné C^∞ plochy v $\mathbb{R}^3$.

c) Nechť G je libovolná otevřená neprázdná podmnožina $\mathbb{R}^2$. Pak množina

$$M = \{(x,y,z,t)\colon\ (y,t) \in G, x = y^2 t, z = \sin(y+t)\}$$

je explicitně zadaný kus 2-rozměrné C^∞ plochy v $\mathbb{R}^4$.

d) Množina $\bigcup_{n=1}^{\infty} (n, n + 1/2) \times \{1\}$ je podle naší definice explicitně zadaný kus 1-rozměrné C^∞ plochy v $\mathbb{R}^2$, i když není souvislá – souvislost kusu plochy jsme v definici nepožadovali.

c) Kružnice $\{(x,y)\colon\ x^2 + y^2 = 1\}$ není explicitně zadaný kus 1-rozměrné C^1 plochy v $\mathbb{R}^2$, ale je sjednocením 4 takových kusů:

$$\{(x,y)\colon\ x \in (-1,1), y = \sqrt{1-x^2}\},\ \{(x,y)\colon\ x \in (-1,1), y = -\sqrt{1-x^2}\},$$

$$\{(x,y)\colon\ y \in (-1,1), x = \sqrt{1-y^2}\},\ \{(x,y)\colon\ y \in (-1,1), x = -\sqrt{1-y^2}\}.$$

Parametricky zadaný kus plochy

Myšlenka parametrického zadání plochy je analogická obvyklé definici křivky: bod k-rozměrné plochy $P \subset \mathbb{R}^n$ by měl být určen zadáním k reálných parametrů, tj. $P = \varphi(G)$, kde $G \subset \mathbb{R}^k$ a φ je „rozumné zobrazení“.

Pro „parametrické zadání“ kusu k-rozměrné C^p plochy v $\mathbb{R}^n$ $(1 \le k < n)$, je přirozené uvažovat otevřenou množinu $G \subset \mathbb{R}^k$ a prosté zobrazení $\varphi\colon\ G \to \mathbb{R}^n$ třídy

C^p. Hlubší rozbor ukazuje, že je vhodné připouštět pouze zobrazení φ, která splňují další dodatečné požadavky (srov. Příklad 2.143).

2.138 Definice. *Nechť $1 \le k < n$ jsou přirozená čísla, $G \subset \mathbb{R}^k$ je neprázdná otevřená množina a $\varphi\colon G \to \mathbb{R}^n$ je zobrazení. Řekneme, že $\varphi\colon G \to \mathbb{R}^n$ je regulární homeomorfismus z $\mathbb{R}^k$ do $\mathbb{R}^n$, jestliže platí:*

(i) *Zobrazení φ je třídy C^1 na G.*
(ii) *Derivace $\varphi'(a)\colon \mathbb{R}^k \to \mathbb{R}^n$ je prosté zobrazení pro každý bod $a \in G$.*
(iii) *Zobrazení $\varphi\colon G \to \varphi(G)$ je homeomorfismus.*

2.139 Definice. *Nechť $1 \le k < n$. Řekneme, že $M \subset \mathbb{R}^n$ je parametricky zadaný kus k-rozměrné C^p plochy v $\mathbb{R}^n$, jestliže existuje neprázdná otevřená množina $G \subset \mathbb{R}^k$ a regulární homeomorfismus $\varphi\colon G \to \mathbb{R}^n$ třídy C^p takový, že $M = \varphi(G)$.*

2.140 Poznámka.

(a) Pokud G je navíc souvislá množina (případně množina homeomorfní s otevřenou koulí v $\mathbb{R}^k$), používá se v literatuře pro $\varphi(G)$ někdy název „jednoduchá plocha“ nebo „elementární plocha“.

(b) Podmínka (ii) z Definice 2.138 je *podmínka regularity*. Z lineární algebry je známo, že tuto podmínku lze ekvivalentně psát i v jiných formách:

$(ii)^*$ $\dim\big(\varphi'(a)(\mathbb{R}^k)\big) = k$ pro každý bod $a \in G$.

$(ii)^{**}$ Hodnost Jacobiho matice zobrazení φ je v každém bodě $a \in G$ rovna k.

Regulárními zobrazeními z $\mathbb{R}^k$ do $\mathbb{R}^n$ $(k < n)$ se často rozumí zobrazení s vlastnostmi (i) a (ii). My tento pojem nebudeme zavádět; „regulární homeomorfismus“je pro nás „nedělitelný termín“.

(c) Vzhledem k tomu, že z podmínky (i) vyplývá spojitost φ, lze podmínku (iii) (ekvivalentně) nahradit kteroukoliv z následujících dvou podmínek, které jsou ekvivalentní spojitosti φ^{-1}:

$(iii)^*$ Zobrazení φ je prosté a pro každou otevřenou množinu $H \subset G$ je $\varphi(H)$ množina (relativně) otevřená ve $\varphi(G)$.

$(iii)^{**}$ Zobrazení φ je prosté a pro každou posloupnost (t_n) v G a $t \in G$ platí implikace $\varphi(t_n) \to \varphi(t) \Longrightarrow t_n \to t$.

(d) Pokud $k = n$ a $G \subset \mathbb{R}^n$, pak zobrazení $\varphi\colon G \to \mathbb{R}^n$ splňuje podmínky (i), (ii), (iii), právě když φ je difeomorfismus. To plyne okamžitě z Tvrzení 2.124.

Následující příklad (a také Příklad (?)) ukazuje, že kdybychom v Definici 2.138 vynechali podmínku regularity (ii), plocha $\varphi(G)$ by nemusela být „hladká“.

2.141 Příklad. Nechť $\varphi\colon \mathbb{R} \to \mathbb{R}^2$, $\varphi(x) = (x^3, |x^3|)$. Pak $\varphi(\mathbb{R})$ je grafem funkce $f(x) = \|x\|$, takže to není „hladká“ 1-rozměrná plocha. Předpoklady (i) a (iii) se snadno ověří; podmínka (ii) však není splněna pro $a = 0$.

2.142 Příklad. Předpokládejme, že $\varphi\colon \mathbb{R}^k \to \mathbb{R}^n, (1 \le k < n)$ je lineární zobrazení. Pak $\varphi'(a) = \varphi$ pro každý bod $a \in \mathbb{R}^k$. Podmínka regularity (ii) je tedy splněna právě

Obr. 2.12.

tehdy, když φ je prosté. V tom případě je však vždy splněna i podmínka (iii). To můžeme zdůvodnit tak, že $\varphi(\mathbb{R}^k)$ je konečně rozměrný normovaný lineární podprostor $\mathbb{R}^n$, a protože zobrazení $\varphi^{-1}\colon\ \varphi(\mathbb{R}^k) \to \mathbb{R}^k$ je lineární, je nutně spojité (viz Věta 1.130). Zcela analogický je případ, kdy φ je zúžení lineárního (resp. afinního) zobrazení na neprázdnou otevřenou množinu $G \subset \mathbb{R}^n$.

Následující příklad však ukazuje, že obecně nelze podmínku (iii) z definice regulárního difeomorfismu nahradit předpokladem prostoty φ.

2.143 Příklad. (Srov. obr. 2.12).
Uvažujme množinu

$$M := ((-2,0) \times \{-1\}) \cup \{(x,y)\colon\ x^2 + y^2 = 1\}.$$

Položíme nyní $G := (-2, 2\pi)$ a budeme uvažovat prostou parametrizaci $\varphi\colon\ G \to M$, která popisuje pohyb bodu, který v „časovém intervalu“ $(-2,0)$ (jednotkovou rychlostí) probíhá úsečku $(-2,0) \times \{-1\}$ „zleva doprava“ a pak v „časovém intervalu“ $[0, 2\pi]$ probíhá kružnici $\{(x,y)\colon\ x^2+y^2 = 1\}$ v kladném smyslu. Přesněji:

$$\varphi(t) = (t,-1), t \in (-2,0), \varphi(t) = (\cos(-\pi/2+t), \sin(-\pi/2+t)), t \in [0, 2\pi).$$

Elementární výpočet ukazuje, že φ splňuje podmínky (i) a (ii). Navíc je φ zřejmě prosté; není však regulárním homeomorfismem z $\mathbb{R}$ do $\mathbb{R}^2$, protože $\varphi^{-1}\colon\ M \to G$ *není spojité*. To ukážeme pomocí podmínky $(iii)^{**}$ z Poznámky 2.140: klademe-li $t_k = 2\pi - 1/k$ a $t = 0$, platí $\varphi(t_k) \to \varphi(t) = (0,-1)$, ale neplatí $t_k \to t$. Lze ukázat, že M vůbec není parametricky zadaný kus 1-rozměrné C^p plochy v $\mathbb{R}^2$. „Vadí“ však pouze bod $(0,-1)$: zúžení φ na otevřenou množinu $(-2,0) \cup (0, 2\pi)$ je regulární homeomorfismus z $\mathbb{R}$ do $\mathbb{R}^2$, takže $M^*\colon = M \setminus \{(0,-1)\}$ je parametricky zadaný kus 1-rozměrné C^1 plochy v $\mathbb{R}^2$.

S přesným ověřováním spojitosti φ^{-1} mohou být potíže i tehdy, když se věc zdá být z geometrického názoru zcela jasná. V případě, že G je omezená množina, lze často užít následující tvrzení.

2.144 Tvrzení. *Nechť $1 \le k < n$, $G \subset \mathbb{R}^k$ je neprázdná omezená otevřená množina a $\varphi\colon\ G \to \mathbb{R}^n$ je prosté spojité zobrazení. Předpokládejme, že existuje* *spojité zobrazení* $\widetilde{\varphi}\colon \overline{G} \to \mathbb{R}^n$, *které rozšiřuje zobrazení* φ. *Pak* $\varphi\colon G \to \varphi(G)$ *je homeomorfismus, právě když platí podmínka*

$$\widetilde{\varphi}(\partial G) \cap \varphi(G) = \emptyset. \tag{2.50}$$

Důkaz. Předpokládejme nejprve, že platí podmínka (2.50). Víme, že zobrazení $\varphi^{-1}\colon \varphi(G) \to G$ je spojité, právě když vzor $(\varphi^{-1})^{-1}(A) = \varphi(A)$ je (relativně) uzavřená množina ve $\varphi(G)$ pro každou (relativně) uzavřenou množinu A v G. Nechť tedy A je relativně uzavřená v G, tj. $A = \overline{A} \cap G$. Protože G je omezená, je podle Věty 1.104 $\overline{A}$ kompaktní. Podle Věty 1.106 je tedy množina $\widetilde{\varphi}(\overline{A})$ kompaktní, a proto je uzavřená v $\mathbb{R}^n$. Protože zřejmě $\overline{A} = A \cup (\overline{A} \cap \partial G)$, podle podmínky (2.50) platí

$$\widetilde{\varphi}(\overline{A}) \cap \varphi(G) = \varphi(A) \cup \big(\widetilde{\varphi}(\overline{A} \cap \partial G) \cap \varphi(G))\big) = \varphi(A),$$

takže dostáváme, že $\varphi(A)$ je (relativně) uzavřená množina ve $\varphi(G)$, což jsme měli dokázat.
Nyní předpokládejme, že podmínka (2.50) neplatí. Existují tedy body $t \in G$ a $t^* \in \partial G$, pro které $\varphi(t) = \widetilde{\varphi}(t^*)$. Zvolíme-li nyní posloupnost (t_n) bodů z G konvergující k t^*, dostáváme, že $\varphi(t_n) \to \varphi(t)$, takže neplatí podmínka $(iii)^{**}$ z Poznámky 2.140 a $\varphi\colon G \to \varphi(G)$ není homeomorfismus.

2.145 Příklad. Nechť $S\colon \mathbb{R}^3 \to \mathbb{R}^3$ je zobrazení „zavádějící sférické souřadnice“ z Příkladu 2.132. Uvažujme zobrazení

$$\Psi\colon \mathbb{R}^2 \to \mathbb{R}^3, \quad \Psi(\theta,\varphi) = S(1,\theta,\varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$$

Z vlastností S vyplývá, že $\Psi(\mathbb{R}^2) = \{(x,y,z)\colon x^2+y^2+z^2=1\}$. Zřejmě Ψ je třídy C^1 na $\mathbb{R}^2$ a úvahou nebo výpočtem snadno dostáváme, že platí rovnost $\Psi'(\theta,\varphi)\,(h,k) = S'(1,\theta,\varphi)\,(0,h,k)$. Protože $J_S(1,\theta,\varphi) = \sin\theta$, je lineární zobrazení $\Psi'(\theta,\varphi)$ prosté, pokud $\sin\theta \neq 0$. Položíme-li $G\colon = (0,\pi)\times(0,2\pi)$ a $\Phi\colon = \Psi\restriction_G$, splňuje tedy Φ podmínky (i) a (ii) z Definice 2.138. Z geometrického názoru „je vidět“ a snadno se přesně dokáže, že Φ je prosté a $\Phi(G)$ je jednotková sféra bez „nultého poledníku“, tj.

$$\Phi(G) = P\colon = \{(x,y,z)\colon x^2+y^2+z^2=1\} \setminus \{(x,y,z)\colon y=0, x\ge 0, x^2+z^2=1\}.$$

Zobrazení $\widetilde{\Phi}\colon = \Psi\restriction_{\overline{G}}$ je zřejmě spojité rozšíření Φ a je snadné ověřit platnost podmínky (2.50). Zobrazení Φ je tedy regulární homeomorfismus a M je parametricky zadaný kus dvourozměrné C^1 plochy.

Kus k-rozměrné C^p plochy zadaný difeomorfismem.

2.146 Definice. *Nechť* $1 \le k < n$. *Řekneme, že* $\emptyset \neq M \subset \mathbb{R}^n$ *je kus* k*-rozměrné* C^p *plochy zadaný difeomorfismem, jestliže existuje otevřená množina* $H \subset \mathbb{R}^n$ *a difeomorfismus* $\psi\colon H \to \mathbb{R}^n$ *třídy* C^p *takový, že*

$$M = \psi(H \cap \{(x_1,\dots,x_n)\colon x_{k+1} = x_{k+2} = \dots = x_n = 0\}).$$

2.147 Poznámka. Kdybychom v definici místo lineárního k-rozměrného prostoru $\{(x_1,\dots,x_n)\colon x_{k+1} = x_{k+2} = \dots = x_n = 0\}$ připouštěli libovolný k-rozměrný afinní podprostor $\mathbb{R}^n$, dostali bychom stejný pojem. To je snadno vidět z toho, že libovolné dva k-rozměrné afinní podprostory $\mathbb{R}^n$ lze na sebe převést (afinním) difeomorfismem, který zobrazuje $\mathbb{R}^n$ na $\mathbb{R}^n$.

Implicitně zadaný kus k-rozměrné C^p plochy.

2.148 Definice. *Nechť $1 \le k < n$. Řekneme, že $\emptyset \ne M \subset \mathbb{R}^n$ je implicitně zadaný kus k-rozměrné C^p plochy v $\mathbb{R}^n$, existuje-li otevřená množina $G \subset \mathbb{R}^n$ a zobrazení $g\colon G \to \mathbb{R}^{n-k}$ takové, že g je třídy C^p v G, Jacobiho matice $[g'(x)]$ zobrazení g má v každém bodě $x \in G$ hodnost $n-k$ a*

$$M = \{x \in G\colon\ g(x) = 0\}.$$

Dále budeme potřebovat následující snadné vztahy mezi právě definovanými pojmy.

2.149 Tvrzení. *Nechť $1 \le k < n$ a $M \subset \mathbb{R}^n$. Uvažujme následující výroky.*

(i) *M je explicitně zadaný kus k-rozměrné C^p plochy.*
(ii) *M je parametricky zadaný kus k-rozměrné C^p plochy.*
(iii) *M je implicitně zadaný kus k-rozměrné C^p plochy.*
(iv) *M je kus k-rozměrné C^p plochy zadaný difeomorfismem.*

Pak platí následující implikace:

$$(i) \Rightarrow (ii),\ (i) \Rightarrow (iii),\ (i) \Rightarrow (iv),\ (iv) \Rightarrow (ii).$$

Důkaz. „$(i) \Rightarrow (ii)$“: Předpokládejme, že M je explicitně zadaný kus k-rozměrné C^p plochy. Pro větší přehlednost předpokládejme, že pro indexy $i_1, \dots, i_k$ z Definice 2.136 platí $i_1 = 1, \dots, i_k = k$. Pak existuje neprázdná otevřená množina $G \subset \mathbb{R}^k$ a zobrazení $f\colon G \to \mathbb{R}^{n-k}$ třídy C^p takové, že

$$M = \{(x_1, \dots, x_n)\colon\ (x_{k+1}, \dots, x_n) = f(x_1, \dots, x_k)\}.$$

Je snadno vidět, že zobrazení $\varphi\colon G \to \mathbb{R}^n$ dané předpisem $\varphi(t) = (t, f(t))$ je regulární homeomorfismus z $\mathbb{R}^k$ do $\mathbb{R}^n$ a $\varphi(G) = M$. Množina M je tedy parametricky zadaný kus k-rozměrné C^p plochy. Je zřejmé, že případ obecných indexů $i_1, \dots, i_k$ je zcela obdobný, jen formální zápis je méně přehledný.

„$(i) \Rightarrow (iii)$“: Opět předpokládáme, že $i_1 = 1, \dots, i_k = k$, a G, f mají stejný smysl jako výše. Položme $G^* := G \times \mathbb{R}^{n-k}$ a definujme zobrazení $g\colon G^* \to \mathbb{R}^{n-k}$ předpisem $g(t,u)\colon = u - f(t)$. Snadno vidíme, že $g \in C^1(G^*)$, Jacobiho matice zobrazení g má v každém bodě $(t,u) \in G^*$ hodnost $n-k$ a

$$g(t,u) = 0 \iff u = f(t) \iff (t,u) \in M,$$

takže M je skutečně implicitně zadaný kus k-rozměrné C^p plochy.

„$(i) \Rightarrow (iv)$“: Stále předpokládáme, že $i_1 = 1, \dots, i_k = k$ a G, f jsou jako výše. Položme $H\colon = G \times \mathbb{R}^{n-k}$ a $\psi(t,u) := (t, u + f(t))$ pro $(t,u) \in H$. Zřejmě $\psi^{-1}(\tau, v) = (\tau, v - f(\tau))$, $(\tau, v) \in H$, takže ψ je difeomorfismus H na H. Tvrzení (iv) vyplývá z rovností

$$\psi(H \cap \{(x_1, \dots, x_n)\colon x_{k+1} = x_{k+2} = \cdots = x_n = 0\}) = \{(t, f(t)\colon t \in G\} = M.$$

„$(iv) \Rightarrow (ii)$“: Nechť M je kus k-rozměrné C^p plochy zadaný difeomorfismem. Existuje tedy otevřená množina $H \subset \mathbb{R}^n$ a difeomorfismus $\psi\colon H \to \mathbb{R}^n$ takový, že

$$M = \psi\left(H \cap \{(x_1, \dots, x_n)\colon x_{k+1} = x_{k+2} = \cdots = x_n = 0\}\right).$$

Nechť $G := \{(t_1, \dots, t_k) \in \mathbb{R}^k\colon (t_1, \dots, t_k, 0, \dots, 0) \in H\}$ a pro $t = (t_1, \dots, t_k) \in G$ položme $\varphi(t) = \psi(t_1, \dots, t_k, 0, \dots, 0)$. Zřejmě $\varphi(G) = M$ a snadno lze ověřit, že φ je regulární homeomorfismus z $\mathbb{R}^k$ do $\mathbb{R}^n$. Je tedy M parametricky zadaný kus k-rozměrné C^p plochy.

Obtížnější je následující důležité tvrzení, a proto je dokážeme podrobněji.

2.150 Tvrzení. *Nechť $1 \le k < n, M \subset \mathbb{R}^n$ je parametricky zadaný kus k-rozměrné C^p plochy a $a \in M$. Pak existuje otevřené okolí U bodu a takové, že $U \cap M$ je explicitně zadaný kus k-rozměrné C^p plochy.*

Důkaz. Podle definice existuje otevřená množina $G \subset \mathbb{R}^k$ a regulární homeomorfismus $\varphi = (\varphi_1, \dots, \varphi_n)\colon G \to \mathbb{R}^n$ takový, že $M = \varphi(G)$. Označme $t_0 := \varphi^{-1}(a)$. Protože hodnost Jacobiho matice zobrazení φ v bodě t_0 je podle předpokladu k a řádky této matice jsou

$$\operatorname{grad} \varphi_1(t_0), \dots, \operatorname{grad} \varphi_n(t_0),$$

existují indexy $1 \le i_1 < i_2 < \cdots < i_k \le n$ takové, že vektory

$$\operatorname{grad} \varphi_{i_1}(t_0), \dots, \operatorname{grad} \varphi_{i_k}(t_0)$$

jsou lineárně nezávislé. Položíme-li

$$\omega(t)\colon = (\varphi_{i_1}(t), \dots, \varphi_{i_k}(t)),\ t \in G,$$

dostáváme, že $\omega\colon G \to \mathbb{R}^k$ je zobrazení třídy C^p na G, jehož Jacobiho matice v bodě t_0 je regulární. Podle Věty 2.121 o inverzním zobrazení existuje otevřené okolí V bodu t_0 takové, že zúžení $\tau\colon = \omega\restriction_V$ je difeomorfismus. Je tedy $W := \tau(V)$ otevřená podmnožina $\mathbb{R}^k$ a τ^{-1} je difeomorfismus W na V. Definujme ještě

$$\psi(t) := (\varphi_{j_1}(t), \dots, \varphi_{j_{n-k}}(t)),\quad t \in G,$$

kde $j_1 < \cdots < j_{n-k}$ jsou „zbylé indexy“, $(\{j_1, \dots, j_{n-k}\} = \{1, \dots, n\} \setminus \{i_1, \dots, i_k\})$. Položíme-li $f := \psi \circ \tau^{-1}$, je $f\colon W \to \mathbb{R}^{n-k}$ třídy C^p a snadno vidíme, že

$$\varphi(V) = \{(x_1, \dots, x_n)\colon (x_{j_1}, \dots, x_{j_{n-k}}) = f(x_{i_1}, \dots, x_{i_k})\}. \tag{2.51}$$

Skutečně, je-li $(x_1, \dots, x_n) = \varphi(t)$ pro nějaký bod $t \in V$, pak podle definice zobrazení τ, ψ je $(x_{i_1}, \dots, x_{i_k}) = \tau(t) \in W$ a $(x_{j_1}, \dots, x_{j_{n-k}}) = \psi(t)$, takže

$$(x_{j_1},\dots,x_{j_{n-k}}) = \psi(\tau^{-1}(x_{i_1},\dots,x_{i_k})) = f(x_{i_1},\dots,x_{i_k}).$$

Jestliže naopak pro bod $(x_1,\dots,x_n)$ platí $(x_{j_1},\dots,x_{j_{n-k}}) = f(x_{i_1},\dots,x_{i_k})$, pak pro $t := \tau^{-1}(x_{i_1},\dots,x_{i_k})$ platí $(x_{i_1},\dots,x_{i_k}) = \tau(t)$ a $(x_{j_1},\dots,x_{j_{n-k}}) = \psi(\tau^{-1}(x_{i_1},\dots,x_{i_k})) = \psi(t)$, takže $(x_1,\dots,x_n) = \varphi(t) \in \varphi(V)$.

Protože $\varphi\colon G \to \mathbb{R}^n$ je regulární homeomorfismus z $\mathbb{R}^k$ do $\mathbb{R}^n$, je obraz $\varphi(V)$ otevřené množiny V množina relativně otevřená v $M = \varphi(G)$. Existuje tudíž otevřená množina $U \subset \mathbb{R}^n$ taková, že $U \cap M = \varphi(V)$. Podle (2.51) je tedy $U \cap M$ explicitně zadaný kus k-rozměrné C^p plochy.

Analogické tvrzení pro implicitně zadaný kus k-rozměrné C^p plochy je vlastně jen jinou formulací věty o implicitních funkcích.

2.151 Tvrzení. *Nechť $1 \le k < n$, $M \subset \mathbb{R}^n$ je implicitně zadaný kus k-rozměrné C^p plochy a $a \in M$. Pak existuje otevřené okolí U bodu a takové, že $U \cap M$ je explicitně zadaný kus k-rozměrné C^p plochy.*

Důkaz. Podle Definice 2.148 můžeme najít otevřenou množinu $G \subset \mathbb{R}^n$ a zobrazení $g = (g_1,\dots,g_{n-k})\colon G \to \mathbb{R}^{n-k}$ takové, že g je třídy C^p v G, Jacobiho matice $[g'(a)]$ má v každém bodě $x \in G$ hodnost $n-k$ a $M = \{x \in G\colon\ g(x) = 0\}$. Protože matice $[g'(a)]$ má hodnost $n-k$, existují indexy $1 \le j_1 < \cdots < j_{n-k} \le n$ takové, že

$$\frac{D(g_1,\dots,g_{n-k})}{D(x_{j_1},\dots,x_{j_{n-k}})}(a) \neq 0.$$

Označme „zbylé indexy“ $i_1 < \cdots < i_k$ ($\{i_1,\dots,i_k\} \cup \{j_1,\dots,j_{n-k}\} = \{1,\dots,n\}$). Podle věty o implicitních funkcích dostáváme (viz. Poznámka 2.116), že existují otevřené okolí $U \subset G$ bodu a a zobrazení f z $\mathbb{R}^k$ do $\mathbb{R}^{n-k}$, které je definované a třídy C^p na nějakém otevřeném okolí bodu $(a_{i_1},\dots,a_{i_k})$, takové, že

$$M \cap U = \{x \in U\colon\ g(x) = 0\} = \{(x_1,\dots,x_n)\colon\ (x_{j_1},\dots,x_{j_{n-k}}) = f(x_{i_1},\dots,x_{i_k})\}.$$

Je tedy $U \cap M$ explicitně zadaný kus k-rozměrné C^p plochy.

2.152 Poznámka. Lze snadno dokázat, že pokud M je kus k-rozměrné C^p plochy v $\mathbb{R}^n$ zadaný parametricky (implicitně, difeomorfismem) a $\psi\colon \mathbb{R}^n \to \mathbb{R}^n$ je difeomorfismus, pak $\psi(M)$ je opět kus plochy stejného druhu. Speciálně tyto pojmy „kusů ploch“ jsou „geometrické “ — nezávisí na zvoleném systému kartézských (dokonce i afinních) souřadnic. Můžeme tedy tyto pojmy definovat i v obecném eukleidovském či afinním prostoru (kde žádný souřadnicový systém není předem zadán). Jinak je tomu s pojmem explicitně zadaného kusu plochy, který není invariantní ani vůči změně kartézských souřadnic. Uvažujme například „horní půlkružnici“
$M = \{(x,y)\colon\ x^2 + y^2 = 1, y > 0\}$, která je zřejmě explicitně zadaný kus 1-rozměrné C^∞ plochy v $\mathbb{R}^2$. Otočíme-li ale množinu M kolem počátku v kladném smyslu o $\pi/4$, výsledná množina již takovým explicitně zadaným kusem plochy nebude.

2.153 Věta. *Nechť $1 \le k < n$ a $P \subset \mathbb{R}^n$ je neprázdná množina. Pak následující vlastnosti množiny P jsou ekvivalentní.*

(i) *Pro každý bod $a \in P$ existuje otevřené okolí U bodu a takové, že $U \cap P$ je explicitně zadaný kus k-rozměrné C^p plochy.*
(ii) *Pro každý bod $a \in P$ existuje otevřené okolí U bodu a takové, že $U \cap P$ je parametricky zadaný kus k-rozměrné C^p plochy.*
(iii) *Pro každý bod $a \in P$ existuje otevřené okolí U bodu a takové, že $U \cap P$ je implicitně zadaný kus k-rozměrné C^p plochy.*
(iv) *Pro každý bod $a \in P$ existuje otevřené okolí U bodu a takové, že $U \cap P$ je kus k-rozměrné C^p plochy zadaný difeomorfismem.*

Důkaz. Implikace $(i) \implies (ii), (i) \implies (iii), (i) \implies (iv)$ a $(iv) \Rightarrow (ii)$ plynou okamžitě z Tvrzení 2.149. Stačí tedy dokázat implikace $(ii) \Rightarrow (i)$ a $(iii) \Rightarrow (i)$.

K důkazu prvé z nich zvolme libovolný bod $a \in P$. Podle (ii) existuje otevřené okolí V bodu a, pro které je $V \cap P$ parametricky zadaný kus k-rozměrné C^p plochy. Podle Tvrzení 2.150 existuje otevřené okolí V^* bodu a takové, že $V^* \cap (V \cap P)$ je explicitně zadaný kus k-rozměrné C^p plochy. Položíme-li $U := V \cap V^*$, dostáváme platnost (i).

Důkaz implikace $(iii) \Rightarrow (i)$ dostáváme stejným způsobem z Tvrzení 2.151.

Nyní můžeme vyslovit *základní definici* tohoto oddílu.

2.154 Definice. *Nechť $1 \le k < n$. Neprázdnou množinu $P \subset \mathbb{R}^n$ budeme nazývát k-rozměrnou C^p plochou, jsou-li splněny (ekvivalentní) podmínky (i)-(iv) z Věty 2.153.*

2.155 Poznámka.

(i) Je snadné ověřit, že každý z výše definovaných typů kusů k-rozměrných C^p ploch je k-rozměrná C^p plocha. Speciálně (viz Příklad 2.137 d)) vidíme, že k-rozměrná C^p plocha nemusí být souvislá množina.
(ii) Je-li ψ difeomorfismus $\mathbb{R}^n$ do $\mathbb{R}^n$ třídy C^p a $M \subset \mathbb{R}^n$ je k-rozměrná C^p plocha, pak (srov. Poznámka 2.152) $\psi(M)$ je opět k-rozměrná C^p plocha.
(iii) Je-li $M \subset \mathbb{R}^n$ k-rozměrná C^p plocha a $G \subset \mathbb{R}^n$ je otevřená množina protínající M, pak $M \cap G$ je zřejmě opět k-rozměrná C^p plocha.

Autoři učebnic berou za základ při definici k-rozměrné C^p plochy v $\mathbb{R}^n$ různé podmínky (většinou některou z podmínek (i)-(iv) z Věty 2.153). Někdy se v definici předpokládá souvislost plochy; někdy se místo o k-rozměrné C^p ploše hovoří o k-rozměrné varietě třídy C^p v $\mathbb{R}^n$. Poznamenejme, že pojem variety, který je základem moderní diferenciální geometrie, je velmi obecný — prvky (body) variety mohou být libovolné objekty. Pojem k-rozměrné C^p plochy lze přirozeným způsobem ztotožnit (srov. [Kow; Věta 2.27]) s pojmem k-rozměrné variety třídy C^p, která je *vlastní podvarietou* prostoru $\mathbb{R}^n$.

Na přirozenou otázku po velikosti k-rozměrných ploch odpovídá následující tvrzení.

2.156 Tvrzení. *Nechť $1 \le k < n$ a $P \subset \mathbb{R}^n$ je k-rozměrná C^1 plocha. Pak P je řídká, lebesgueovsky nulová a borelovská množina.*

Důkaz. Zvolme libovolný bod $a \in P$. Pak existuje jeho otevřené okolí U_a takové, že $P^* := P \cap U$ je kus k-rozměrné plochy zadaný difeomorfismem.

Existuje tedy otevřená množina $H \subset \mathbb{R}^n$ a difeomorfismus ψ: $H \to \mathbb{R}^n$ třídy C^p takový, že $P^* = \psi(H \cap A)$, kde $A = \{(x_1, \dots, x_n)\colon x_{k+1} = x_{k+2} = \cdots = x_n = 0\}$. Množina $H \cap A$ je zřejmě řídká v H a $\lambda_n(H \cap A) = 0$.

Protože $\psi H \to \psi(H)$ je homeomorfismus, je P^* řídká v otevřené množině $\psi(H)$, takže je řídká (v $\mathbb{R}^n$). Je tedy P lokálně řídká, z čehož snadno vyplývá (např. pomocí podmínky (iv) z Tvrzení 1.90), že je řídká.

Z věty o substituci (Věta 6.34) okamžitě vyplývá, že $\lambda_n(P^*) = 0$. Pomocí Poznámky 1.69 snadno dostáváme $\lambda_n(P) = 0$.

Protože množina $H \cap A$ je množina uzavřená v H, je P^* uzavřená v otevřené množině $\psi(H)$, a proto je také uzavřená v otevřené množině $V := U \cap \psi(H) \supset P^* = P \cap V$, tj. $\overline{P \cap V} \cap V = P \cap V$. Protože zřejmě $\overline{P \cap V} \cap V = \overline{P} \cap V$, snadno tedy vidíme, že a je vnitřním bodem množiny P v prostoru $\overline{P}$. Tím jsme dokázali, že P je otevřená podmnožina $\overline{P}$. Podle Tvrzení 1.22 platí $P = \overline{P} \cap G$ pro nějakou otevřenou množinu $G \subset \mathbb{R}^n$; P je tedy borelovská množina.

2.157 Poznámka. Je zřejmé, že k-rozměrná C^∞ plocha v $\mathbb{R}^n$ nemusí (ale může) být uzavřená podmnožina $\mathbb{R}^n$. Otevřená však být nemůže (to by nebyla řídká).

??? Dat do wint? Následující tvrzení okamžitě plyne z definice k-plochy a Poznámky 1.69.

2.158 Tvrzení. *Nechť $P \subset \mathbb{R}^n$ je k-plocha. Pak P je spočetným sjednocením explicitně zadaných kusů k-rozměrných C^1 ploch.*

2.13.3 Tečný prostor ke k-rozměrné C^p ploše

Velmi důležitý je pojem tečného prostoru k ploše P v bodě $a \in P$. *Afinním tečným prostorem* $T_a^{\mathrm{af}}(P)$ ke ploše P v bodě a rozumíme, zhruba řečeno, jediný afinní k-rozměrný prostor

$A \subset \mathbb{R}^n$, který obsahuje bod a a v blízkosti bodu a se ke ploše P „velmi těsně přimyká“. Tento smysl je možno různými způsoby přesně definovat (viz Poznámka ?? níže). Velmi zhruba lze říci, že v dostatečně malém okolí bodu a již „od sebe plochu P a afinní prostor A nerozeznáme“.

Pojem afinního tečného prostoru bude ovšem zobecňovat již definovaný pojem tečné nadroviny ke grafu funkce: Je-li f: $\mathbb{R}^{n-1} \to \mathbb{R}$ funkce třídy C^1, pak graf $P = \{(x_1, \dots, x_n)\colon x_n = f(x_1, \dots, x_{n-1})\}$ funkce f je $(n-1)$-rozměrná C^1 plocha

Obr. 2.13. Tečný vektor v ke ploše P.

v $\mathbb{R}^n$; afinním tečným prostorem k P v bodě $a \in P$ pak bude tečná nadrovina k P v tomto bodě.

Základním geometrickým pojmem v teorii ploch však není pojem afinního tečného prostoru, ale pojem (vektorového) tečného prostoru $T_a(P)$, který je zaměřením afinního tečného prostoru. Jinými slovy, $T_a(P)$ je ten (k-rozměrný) vektorový podprostor $\mathbb{R}^n$, pro který

$$T_a^{\mathrm{af}}(P) = a + T_a(P).$$

Probíhá-li v vektorový tečný prostor, vyplní body $a + v$ celý afinní tečný prostor.

Tečný vektorový prostor ke ploše se často definuje následujícím způsobem (srov. obr. 2.13) pomocí derivací zobrazení z $\mathbb{R}$ do $\mathbb{R}^n$ („cest“).

2.159 Definice. *Nechť $P \subset \mathbb{R}^n$ je k-rozměrná C^p plocha a je dán bod $a \in P$. Pak vektor $v \in \mathbb{R}^n$ nazveme tečným vektorem ke ploše P v bodě a, existuje-li $\delta > 0$ a spojité zobrazení $\gamma\colon (-\delta, \delta) \to P$ takové, že $\gamma(0) = a$ a $\gamma'(0) = v$. Množinu všech tečných vektorů ke ploše P v bodě a označujeme $T_a(P)$ a nazýváme ji (vektorovým) tečným prostorem ke ploše P v bodě a.*

2.160 Poznámka.

(i) Je vhodné si uvědomit toto: Je-li U otevřené okolí bodu a, pak $P^* = P \cap U$ je opět k-rozměrná C^p plocha a $T_a(P^*) = T_a(P)$. To plyne okamžitě z toho, že pro každou spojitou „cestu“ $\gamma\colon (-\delta, \delta) \to P$, pro kterou $\gamma(0) = a$, zřejmě existuje $\delta^* > 0$ takové, že $\gamma((-\delta^*, \delta^*)) \subset P^*$.

(ii) I kdybychom v Definici 2.159 nepožadovali spojitost „cesty“ γ, dostali bychom stejný pojem (to je vidět z *důkazu* následující věty).

V následující větě dokážeme, že $T_a(P)$ je skutečně vektorový prostor (dimenze k), a ukážeme, jak jej v konkrétních případech můžeme snadno určit.

2.161 Věta. *Nechť $P \subset \mathbb{R}^n$ je k-rozměrná C^p plocha a je dán bod $a \in P$. Pak platí následující tvrzení.*

(i) *$T_a(P)$ je k-rozměrný vektorový prostor.*

(ii) *Je-li P parametricky zadaný kus k-rozměrné C^p plochy a $\varphi\colon H \to \mathbb{R}^n$ je příslušný regulární homeomorfismus z $\mathbb{R}^k$ do $\mathbb{R}^n$, pro který $\varphi(H) = P$ a $a = \varphi(u_0)$, pak*

$$T_a(P) = \operatorname{Im}(\varphi'(u_0)).$$

(iii) *Nechť P je implicitně zadaný kus k-rozměrné C^p plochy, tj.*

$$P = \{x \in G\colon\ g_1(x) = 0, \dots, g_{n-k}(x) = 0\},$$

kde g_i, $i = 1, \dots, n-k$, jsou funkce třídy C^p na nějaké otevřené množině $G \subset \mathbb{R}^n$ takové, že gradienty $\operatorname{grad} g_1(x), \dots, \operatorname{grad} g_{n-k}(x)$ jsou lineárně nezávislé pro každé $x \in G$. Pak

$$T_a(P) = \left(\operatorname{Lin}\{\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_{n-k}(a)\}\right)^{\perp}.$$

Důkaz. Důkaz provedeme ve čtyřech krocích. (Na obr. 2.14 je zobrazen případ, kdy $k = 2$, $n = 3$ a jsou splněny předpoklady (ii), (iii).)

1. krok: Nejdříve dokážeme inkluzi $\operatorname{Im}(\varphi'(u_0)) \subset T_a(P)$ z (ii). Zvolme tedy libovolný vektor $w \in \operatorname{Im}(\varphi'(u_0))$; nechť $v \in H$ je ten vektor, pro který $\varphi'(u_0)(v) = w$. Uvažujme funkci $\gamma(t) := \varphi(u_0 + tv)$. Zřejmě můžeme zvolit $\delta > 0$ tak malé, aby funkce γ byla definována na celém intervalu $(-\delta, \delta)$; pak ovšem $\gamma((-\delta, \delta)) \subset P$. Protože

$$\gamma'(0) = \lim_{t \to 0} \frac{\varphi(u_0 + tv) - \varphi(u_0)}{t} = D_v\varphi(u_0) = \varphi'(u_0)(v) = w,$$

dokázali jsme, že $w \in T_a(P)$.

2. krok: Nyní dokážeme inkluzi $T_a(P) \subset \left(\operatorname{Lin}\{\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_{n-k}(a)\}\right)^{\perp}$ z podmínky (iii). Pro libovolný vektor $w \in T_a(P)$ najděme $\delta > 0$ a spojitou cestu $\gamma\colon (-\delta, \delta) \to P$ takovou, že $\gamma(0) = a$ a $\gamma'(0) = w$. Pro každé $t \in (-\delta, \delta)$ platí

$$g_1(\gamma(t)) = 0, \dots, g_{n-k}(\gamma(t)) = 0.$$

Pro každé $i \in \{1, \dots, n-k\}$ tedy podle věty o derivaci složené funkce (srov. Poznámka 2.56) dostáváme

$$(g_i \circ \gamma)'(0) = \langle \operatorname{grad} g_i(a), \gamma'(0) \rangle = 0,$$

takže skutečně $w \in \left(\operatorname{Lin}\{\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_{n-k}(a)\}\right)^{\perp}$.

3. krok: Uvažujme obecnou k-rozměrnou C^p plochu P. Zvolme otevřené okolí U bodu a, pro které je $U \cap P$ parametricky zadaný kus k-rozměrné C^p plochy. Podle 1. kroku důkazu a Poznámky 2.160(i) dostáváme, že $T_a(P)$ obsahuje k-rozměrný vektorový podprostor $\mathbb{R}^n$, totiž $V := \operatorname{Im}(\varphi'(u_0))$. Existuje však také otevřené okolí U^* bodu a, pro které je $U^* \cap P$ implicitně zadaný kus k-rozměrné C^p plochy. Podle

OBR. 2.14.

2. kroku důkazu a Poznámky 2.160(i) dostáváme, že $T_a(P)$ je obsažen v k-rozměrném vektorovém podprostoru $\mathbb{R}^n$; totiž v ortogonálním doplňku

$$W := (\operatorname{Lin}\{\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_{n-k}(a)\})^{\perp}$$

$(n-k)$-rozměrného prostoru $\operatorname{Lin}\{\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_{n-k}(a)\}$. Platí tedy $V \subset T_a(P) \subset W$. Protože $\dim V = \dim W$, snadno vidíme, že $V = W$, takže $T_a(P) = V = W$ je skutečně k-rozměrný vektorový podprostor $\mathbb{R}^n$.

4. krok: Teď již můžeme dokončit důkaz (ii) a (iii). Protože už víme, že $T_a(P)$ je v obou případech k-rozměrný vektorový prostor, z již dokázaných inkluzí mezi k-rozměrnými vektorovými prostory vyplývá, že je lze nahradit rovnostmi.

2.162 Definice. *Nechť $P \subset \mathbb{R}^n$ je k-rozměrná C^p plocha a je dán bod $a \in P$. Pak ortogonální doplněk $(T_a(P))^{\perp}$ k tečnému prostoru ke ploše P v bodě a se nazývá normálový prostor k ploše P v bodě a. Jeho prvky se nazývají normálové vektory k ploše P v bodě a.*

2.163 Poznámka.

(a) Je-li P implicitně zadaný kus k-rozměrné C^p plochy jako v (iii), pak normálový prostor k ploše P v bodě a je $\operatorname{Lin}\{\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_{n-k}(a)\}$.

(b) V situaci na obr. 2.14 je normálový prostor k ploše P v bodě a jednorozměrný. Tečný prostor $T_a(P) = \operatorname{Im}(\varphi'(u_0))$ je lineárním obalem vektorů $\frac{\partial \varphi}{\partial t_1}(u_0)$, $\frac{\partial \varphi}{\partial t_2}(u_0)$.

Nakonec ještě dokážeme důležité tvrzení, které budeme potřebovat v teorii plošného integrálu.

2.164 Tvrzení. *Nechť P je parametricky zadaný kus k-rozměrné C^1 plochy v $\mathbb{R}^n$, $A \subset \mathbb{R}^k$, $B \subset \mathbb{R}^k$ jsou otevřené množiny a $\alpha\colon A \to \mathbb{R}^n$, $\beta\colon B \to \mathbb{R}^n$ jsou* *regulární homeomorfismy takové, že $\alpha(A) = \beta(B) = P$. Pak existuje (bijektivní) difeomorfismus ω: $A \to B$ takový, že $\alpha = \beta \circ \omega$.*

Důkaz. Stačí zřejmě dokázat, že $\omega := \beta^{-1} \circ \alpha$ je difeomorfismus. Zvolme libovolný bod $a \in A$ a položme $p := \alpha(a)$, $b := \beta^{-1}(p)$. Podle Věty 2.153 existuje otevřené okolí U bodu p takové, že $P \cap U$ je kus k-rozměrné C^1 plochy zadaný difeomorfismem. Existuje tedy otevřená množina $H \subset \mathbb{R}^n$ a difeomorfismus ψ: $H \to \mathbb{R}^n$ třídy C^1 takový, že

$$(2.52) \qquad P \cap U = \psi(H \cap \{(x_1, \ldots, x_n):\ x_{k+1} = x_{k+2} = \cdots = x_n = 0\}).$$

Označme $A^* := \alpha^{-1}(P \cap U)$, $B^* := \beta^{-1}(P \cap U)$ a

$$\alpha^* := \alpha \restriction_{A^*}, \quad \beta^* := \beta \restriction_{B^*}, \quad \mu := \psi^{-1} \circ \alpha^*, \quad \nu := \psi^{-1} \circ \beta^*.$$

Zřejmě A^*, B^* jsou pořadě otevřená okolí bodů a, b a μ, ν jsou prostá zobrazení třídy C^1. Z (2.52) vyplývá, že $\mu = (\mu_1, \ldots, \mu_k, 0, \ldots, 0)$, takže také $\widetilde{\mu} := (\mu_1, \ldots, \mu_k)$ je prosté a třídy C^1. Pro každé $t \in A^*$ platí $\mu'(t) = (\psi^{-1})'(\alpha(t)) \circ \alpha'(t)$. Protože $(\psi^{-1})'(\alpha(t))$ je lineární bijekce $\mathbb{R}^n$ na $\mathbb{R}^n$, platí

$$\dim(\mathrm{Lin}\{\mathrm{grad}(\mu_1(t)), dots, \mathrm{grad}(\mu_k(t)))\}) = h([\mu'(t)]) = \dim(\mathrm{Im}\,\mu'(t)) = \dim(\mathrm{Im}\,\alpha'(t)) = k,$$

takže $[(\widetilde{\mu})'(t)]$ je regulární matice. Zobrazení $\widetilde{\mu}$: $A^* \to \mathbb{R}^k$ je tedy prosté regulární zobrazení, takže je podle Věty 2.126 difeomorfismus.

Zcela obdobně dostáváme, že $\widetilde{\nu} := (\nu_1, \ldots, \nu_k)$ je difeomorfismus. Na základě (2.52) snadno nahlédneme, že $\omega \restriction_{A^*} = (\beta^*)^{-1} \circ \alpha^* = (\widetilde{\nu})^{-1} \circ \widetilde{\mu}$, takže $\omega \restriction_{A^*}$ je regulární zobrazení. Z toho ihned vyplývá, že ω je regulární, a protože je prosté, je difeomorfismus.

2.14 Vázané extrémy

Problém hledání bodů relativního lokálního extrému (viz. Definice 2.103) funkce n proměnných f (vzhledem k množině $M \subset \mathbb{R}^n$) je důležitý pro matematiku i její aplikace (např. fyziku).

V matematice tento problém přirozeně vzniká při vyšetřování *globálních extrémů* funkce. Ve fyzice je pojem lokálního (vázaného) extrému důležitý například při vyšetřování otázek stability.

Budeme se zabývat pro aplikace důležitým případem, kdy množina M je zadána pomocí „vazbových podmínek“ (často je M implicitně zadaný kus k-rozměrné C^1 plochy). V tom případě se o lokálních relativních extrémech (vzhledem k M) hovoří jako o „vázaných extrémech“. Pokud lze M jednoduchým způsobem parametrizovat, ke zjišťování relativních lokálních extrémů funkce f vzhledem k množině M zpravidla žádnou speciální teorii nepotřebujeme; problém se snadno převádí na hledání lokálních extrémů jisté funkce k proměnných. Můžeme totiž vyjít z následujícího téměř zřejmého tvrzení.

2.165 Tvrzení. *Nechť je dána reálná funkce n proměnných f, množina $M \subset \mathbb{R}^n$ a bod $a \in M \cap D_f$. Nechť $H \subset \mathbb{R}^k$ je otevřená množina, $c \in H$ a φ: $H \to M$ je zobrazení, pro které $\varphi(c) = a$. Pak platí:*

(i) *Je-li φ spojité v bodě c a f má v bodě a lokální maximum vzhledem k množině M, pak funkce $f \circ \varphi$ má v bodě c lokální maximum.*

(ii) *Existuje-li okolí U bodu a takové, že φ je homeomorfismus mezi množinami H a $M \cap U$, pak f má lokální maximum (resp. ostré lokální maximum) vzhledem k M v bodě a, právě když $f \circ \varphi$ má v bodě c lokální maximum (resp. ostré lokální maximum).*

2.166 Příklad. Nechť $f(x, y) = x^2 + xy$ a $M = \{(x, y)\colon x^2 + y^2 = 1\}$. Zkoumejme, ve kterých bodech má funkce f lokální relativní extrémy vzhledem k množině M. Máme-li užít předchozí tvrzení, musíme hledat vhodné parametrizace M (resp. částí M). Nabízejí se dva přirozené přístupy:

a) Můžeme použít dvě parametrizace $\varphi(x) = (x, \pm\sqrt{1 - x^2})$ otevřených půlkružnic, které náš problém převádějí (podle Tvrzení 2.165 (ii)) na vyšetřování lokálních extrémů funkcí $f(\varphi(x)) = x^2 \pm x\sqrt{1 - x^2}$.

Zbývají ovšem body $(1, 0), (-1, 0)$; pro vyšetření, zda v nich má f lokální extrém, můžeme použít další parametrizace $\psi(y) = (y, \pm\sqrt{1 - y^2})$.

b) Najednou můžeme parametrizovat M pomocí zobrazení $\varphi(\alpha) := (\cos \alpha, \sin \alpha)$, $\alpha \in \mathbb{R}$, které ovšem není prosté. Má-li však funkce f v bodě $\varphi(t)$ lokální extrém vzhledem k M, má podle Tvrzení 2.165 (i) funkce $g(t) := f(\varphi(t))$ lokální extrém v bodě t, takže $g'(t) = -2 \cos t \sin t - \sin^2 t + \cos^2 t = \cos 2t - \sin 2t = 0$, a proto $\sin 2t = \cos 2t$, $2t = \pi/4 + k\pi$, $t = \pi/8 + k\pi/2$ $(k \in \mathbb{Z})$. Protože $g(t) = 1/2 \cdot (1 + \cos 2t + \sin 2t)$, je buď (pro sudá k) $g(t) = 1/2 \cdot (1 + \sqrt{2})$ nebo (pro lichá k) $g(t) = 1/2 \cdot (1 - \sqrt{2})$. V této konkrétní speciální úloze již učiněné výpočty stačí k úplnému řešení. Protože funkce f je spojitá na kompaktní množině M, nutně na ní nabývá v některých bodech svého maxima a minima; v těchto bodech má ovšem f také lokální extrém vzhledem k množině M. Z toho již vyplývá, že funkce f nabývá na M svého maxima (a tedy i ostrého relativního lokálního maxima) $1/2 \cdot (1 + \sqrt{2})$ ve dvou bodech $(\cos \pi/8, \sin \pi/8)$, $(\cos(\pi/8 + \pi), \sin(\pi/8 + \pi))$ a svého minima (a tedy i ostrého relativního lokálního minima) $1/2 \cdot (1 - \sqrt{2})$ v bodech $(\cos(\pi/8 + \pi/2), \sin(\pi/8 + \pi/2))$, $(\cos(\pi/8 + 3\pi/2), \sin(\pi/8 + 3\pi/2))$.

V předchozím příkladu jsme hledali „vázané extrémy“ tak, že jsme se snažili vypočítat parametrizaci množiny M a použít Tvrzení 2.165. Následující důležitá věta mj. ukazuje, jak se při hledání „vázaných extrémů“ obejít bez výpočtu jakékoliv parametrizace. Tím se někdy vyhneme pracnému rozlišování případů (jako v postupu a) Příkladu 2.166) a výpočty se často podstatně zjednoduší. V některých složitějších případech nelze vhodnou parametrizaci vůbec vypočítat (nebo to není snadné).

2.167 Věta. (nutné podmínky pro vázaný extrém, 1. část věty o Lagrangeových multiplikátorech) *Nechť $G \subset \mathbb{R}^n$ je otevřená množina, f a $g_1, \dots, g_s$ $(1 \leq s < n)$ jsou funkce třídy C^1 na G a nechť v každém bodě $x \in G$ má (Jacobiho) matice*

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \dots & \frac{\partial g_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial g_s}{\partial x_1}(x) & \dots & \frac{\partial g_s}{\partial x_n}(x) \end{pmatrix}$$

hodnost s. Nechť

$$M = \{z \in G\colon\ g_1(z) = 0, \dots, g_s(z) = 0\}$$

a funkce f má v bodě $a \in M$ lokální extrém vzhledem k M. Pak existuje právě jedna s-tice reálných čísel $(\lambda_1, \dots, \lambda_s)$ taková, že a je stacionárním bodem funkce

$$L(x) = f(x) + \sum_{i=1}^{s} \lambda_i g_i(x).$$

Než přistoupíme k důkazu věty, je vhodné udělat několik poznámek a snadných úvah.

2.168 Poznámka.

(i) Rovnicím $g_1(z) = 0, \dots, g_s(z) = 0$ se často říká „vazbové podmínky“ a funkcím $g_1, \dots, g_s$ „vazbové funkce“.

(ii) Číslům $\lambda_1, \dots, \lambda_s$ se říká „Lagrangeovy multiplikátory“ a funkce L se nazývá „Lagrangeova funkce“. Někteří autoři říkají „Lagrangeovy multiplikátory“ číslům $\lambda_i^* := -\lambda_i$; Lagrangeova funkce má pak tvar $L(x) = f(x) - \sum_{i=1}^{s} \lambda_i^* g_i(x)$.

(iii) Bod a je stacionárním bodem funkce L, právě když $\operatorname{grad} L(a) = 0$, což je totéž jako

$$-\operatorname{grad} f(a) = \sum_{i=1}^{s} \lambda_i \operatorname{grad} g_i(a).$$

Z „podmínky regularity“ hovořící o hodnosti Jacobiho matice vyplývá, že vektory

$$\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_s(a)$$

jsou lineárně nezávislé. Z toho vidíme, že pokud čísla $\lambda_1, \dots, \lambda_s$ (Lagrangeovy multiplikátory) existují, jsou určena jednoznačně. Dále: existence čísel $\lambda_1, \dots, \lambda_s$ je ekvivalentní tomu, že vektor $\operatorname{grad} f(a)$ je lineární kombinací vektorů $\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_s(a)$. Závěr věty o vázaných extrémech lze tedy ekvivalentně přepsat podstatně jednodušeji: vektor $\operatorname{grad} f(a)$ je lineární kombinací vektorů $\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_s(a)$.

Zdá se tedy divné, proč se pojem Lagrangeovy funkce zavádí. To má ale své dobré důvody. Ve Větě 2.171 například uvidíme, jak elegantně se *postačující podmínky* pro vázané extrémy formulují pomocí Lagrangeovy funkce.

Důkaz. Podstatně užijeme teorii k-rozměrných ploch a jejich tečných prostorů. Předně si uvědomíme, že M je implicitně zadaný kus $(n-s)$-rozměrné C^1 plochy. Uvažujme nyní libovolný tečný vektor $v \in T_a(M)$. Podle definice tedy existuje spojité zobrazení $\gamma\colon (-\delta,\delta) \to M$ takové, že $\gamma(0) = a$ a $\gamma'(0) = v$. Protože f má v bodě a lokální extrém vzhledem k M, má reálná funkce $f \circ \gamma\colon (-\delta,\delta) \to \mathbb{R}$ lokální extrém v bodě 0 (Tvrzení 2.165 (i)), takže (viz Poznámka 2.56)

$$(f \circ \gamma)'(0) = \langle \operatorname{grad} f(a), \gamma'(0)\rangle = \langle \operatorname{grad} f(a), v\rangle = 0.$$

Dokázali jsme tedy, že $\operatorname{grad} f(a)$ je normálový vektor ke ploše M v bodě a. Z Věty 2.161 (iii) (srov. Poznámka 2.163 (a)) dostáváme, že

$$\operatorname{grad} f(a) \in \operatorname{Lin}\{\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_s(a)\}.$$

Existují tudíž (jednoznačně určená) čísla $\lambda_1, \dots, \lambda_s$, pro která nastává rovnost $-\operatorname{grad} f(a) = \sum_{i=1}^s \lambda_i \operatorname{grad} g_i(a)$, která je ekvivalentní s tím, že a je stacionárním bodem Lagrangeovy funkce $L = f + \sum_{i=1}^s \lambda_i g_i$ (srov. Poznámka 2.168 (iii)).

Následující verze věty o Lagrangeových multiplikátorech (která snadno plyne z předchozí verze) se hodí v početní praxi, kdy vazbové funkce někdy nesplňují podmínku regularity.

2.169 Věta. *Nechť $H \subset \mathbb{R}^n$ je otevřená množina a $f, g_1, \dots, g_s$ $(s < n)$ jsou funkce třídy C^1 na H. Nechť*

$$M = \{z \in H\colon\ g_1(z) = 0, \dots, g_s(z) = 0\}$$

a funkce f má v bodě $a \in M$ lokální extrém vzhledem k M. Pak jsou vektory $\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_s(a)$ lineárně závislé nebo existují čísla $(\lambda_1, \dots, \lambda_s)$ taková, že a je stacionárním bodem Lagrangeovy funkce $L(x) = f(x) + \sum_{i=1}^s \lambda_i g_i(x)$.

Důkaz. Jsou-li vektory $\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_s(a)$ lineárně závislé, je tvrzení věty splněno. V opačném případě jsou vektory $\operatorname{grad} g_1(x), \dots, \operatorname{grad} g_s(x)$ lineárně nezávislé i na nějakém otevřeném okolí $G \subset H$ bodu a. To nahlédneme například takto: Protože řádky $\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_s(a)$ Jacobiho matice zobrazení $g = (g_1, \dots, g_s)$ v bodě a jsou nezávislé, má tato matice nenulový subdeterminant řádu s. Jinými slovy, existují indexy $1 \le i_1 < i_2 < \dots < i_s \le n$ takové, že Jacobiho determinant $J(x) := \dfrac{D(g_1, \dots, g_s)}{D(x_{i_1}, \dots, x_{i_s})}(x)$ je nenulový v bodě a. Protože determinant matice lze vyjádřit pomocí vzorce, který používá pouze prvky této matice a operace násobení, sčítání a odčítání, je funkce J spojitá na H, a proto je nenulová na nějakém otevřeném okolí $G \subset H$ bodu a. Jsou tedy vektory $\operatorname{grad} g_1(x), \dots, \operatorname{grad} g_s(x)$ lineárně nezávislé v každém bodě $x \in G$. Nyní již můžeme na G a $M^* := M \cap G$ použít Větu 2.167, takže dostáváme existenci hledaných čísel $\lambda_1, \dots, \lambda_s$.

2.170 Poznámka. Závěr předchozí věty lze ovšem formulovat také takto:
Pak jsou vektory $\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_s(a), \operatorname{grad} f(a)$ *lineárně závislé.*

V praxi je však často výhodnější postupovat podle komplikovanějšího znění Věty theTheoremvaex2. Jedním z důvodů je to, pokud chceme použít následující větu, Lagrangeovu funkci nutně potřebujeme.

Pomocí Věty 2.167 lze často najít konečnou množinu všech bodů „podezřelých“ z toho, že v nich má funkce relativní lokální extrém. Na otázku, které z těchto bodů jsou skutečně body relativního lokálního extrému, nám často dá odpověď následující věta, která poskytuje mj. účinné *postačující podmínky pro vázané extrémy.*

2.171 Věta. (2. část věty o Lagrangeových multiplikátorech) *Nechť $G \subset \mathbb{R}^n$ je otevřená množina, f, $g_1, \dots, g_s$ $(s < n)$ jsou funkce třídy C^2 na G, nechť v každém bodě $x \in G$ má (Jacobiho) matice*

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \dots & \frac{\partial g_1}{\partial x_n}(x) \\ \vdots & & \vdots \\ \frac{\partial g_s}{\partial x_1}(x) & \dots & \frac{\partial g_s}{\partial x_n}(x) \end{pmatrix}$$

hodnost s a nechť

$$M = \{z \in G\colon\ g_1(z) = 0, \dots, g_s(z) = 0\}.$$

Nechť $a \in M$, $(\lambda_1, \dots, \lambda_s) \in \mathbb{R}^s$ a nechť a je stacionárním bodem funkce

$$L(x) = f(x) + \sum_{i=1}^{s} \lambda_i g_i(x).$$

Pak platí:

(i) *Je-li kvadratická forma $Q := \mathrm{d}^2 L(a)$ positivně definitní na $T_a(M)$ (tj. je-li kvadratická forma $Q \restriction_{T_a(M)}$ pozitivně definitní), pak f má v bodě a ostré lokální minimum vzhledem k množině M.*

(ii) *Je-li $Q := \mathrm{d}^2 L(a)$ negativně definitní na $T_a(M)$, pak f má v bodě a ostré lokální maximum vzhledem k množině M.*

(iii) *Je-li $Q := \mathrm{d}^2 L(a)$ indefinitní na $T_a(M)$, pak f nemá v bodě a lokální extrém vzhledem k množině M.*

Důkaz. Protože M je $(n-s)$-rozměrná C^2 plocha, podle definice existuje otevřené okolí U bodu a takové, že $U \cap M$ je parametricky zadaný kus $(n-s)$-rozměrné C^2 plochy. Nechť $\varphi\colon H \to \mathbb{R}^n$, $\varphi = (\varphi_1, \dots, \varphi_n)$ je příslušná parametrizace, tj. $H \subset \mathbb{R}^{n-s}$ je otevřená množina, φ je regulární homeomorfismus a $\varphi(H) = U \cap M$. Zřejmě f má v bodě a (ostrý) lokální extrém vzhledem k M, právě když $g := f \circ \varphi$ má v bodě $c := \varphi^{-1}(a)$ (ostrý) lokální extrém stejného typu (Tvrzení 2.165 (ii)). Protože $f = L$ na M, platí $g = L \circ \varphi$. Jelikož $g'(c) = L'(a) \circ \varphi'(c)$ a $L'(a) = 0$, je c stacionárním bodem funkce g. Funkce g je třídy C^2 na H (Věta 2.76), takže na ni můžeme aplikovat Větu 2.108. Budeme tedy vyšetřovat kvadratickou formu $Q^* := \mathrm{d}^2 g(c)$. Protože a je stacionárním bodem L, podle Tvrzení 2.96 platí

$$\mathrm{d}^2 g(c) = \sum_{p,q=1}^{n} \frac{\partial^2 L}{\partial x_p \partial x_q}(a)\, d\varphi_p(c)\, d\varphi_q(c) = \mathrm{d}^2 L(a) \circ \varphi'(c).$$

Pro $h \in \mathbb{R}^{n-s}$ je tedy $Q^*(h) = Q(\varphi'(c)h)$. Přitom víme (Věta 2.161(ii)), že vektory $\varphi'(c)h$ vyplní tečný prostor $T_a(M)$, když h probíhá prostor $\mathbb{R}^{n-s}$, a rovnost $\varphi'(c)h = 0$ platí pouze pro $h = 0$. Je tedy Q^* positivně definitní (resp. negativně definitní, resp. indefinitní), právě když Q je positivně definitní (resp. negativně definitní, resp. indefinitní) na $T_a(M)$. Z Věty 2.108 pak již okamžitě plyne tvrzení dokazované věty.

2.172 Poznámka.

(i) Pokud je $Q = \mathrm{d}^2 L(a)$ pozitivně (negativně) semidefinitní na $T_a(M)$, nemůžeme obecně udělat žádný závěr.

(ii) „Definitnost“ kvadratické formy Q na $T_a(M)$ lze zjišťovat takto: Parametrizujeme $T_a(M)$ lineární bijekcí $A\colon \mathbb{R}^{n-s} \to T_a(M)$ a zjistíme, jakého druhu je kvadratická forma $Q^* := Q \circ A$ na $\mathbb{R}^{n-s}$.

Tečný prostor $T_a(M)$ lze vždy parametrizovat pomocí vhodných $(n-s)$ kartézských souřadnic $h_{i_1}, \ldots, h_{i_{n-s}}$ tak, že z rovnic

$$\langle \operatorname{grad} g_1(a), (h_1, \ldots, h_n)\rangle = 0, \ldots, \langle \operatorname{grad} g_s(a), (h_1, \ldots, h_n)\rangle = 0$$

vypočítáme zbylé souřadnice.

2.15 Věta o hodnosti; funkce závislé a nezávislé

Výsledky tohoto odstavce odpovídají na tři přirozené otázky, které spolu velmi úzce souvisí.

Nechť f je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^m$, které je třídy C^1 na nějakém okolí bodu $a \in \mathbb{R}^n$. *První ze zmíněných otázek se ptá, co lze říci o povaze množiny $f(U)$, kde U je „dostatečně malé“ otevřené okolí bodu a.* Na základě předchozích výsledků můžeme již nyní dát uspokojivou odpověď na tuto otázku ve dvou speciálních případech.

(i) Pokud $n = m$ a $J_f(a) \neq 0$, z Věty 2.121 vyplývá, že pro všechna dostatečně malá otevřená okolí U bodu a je $f(U)$ otevřená podmnožina $\mathbb{R}^m$.

(ii) Předpokládejme, že $n < m$ a hodnost Jacobiho matice $[f'(a)]$ je rovna n (je tedy „maximální možná“). Z důkazu Tvrzení 2.150 je možno nahlédnout, že pro všechna dostatečně malá otevřená okolí U bodu a je $f(U)$ explicitně zadaný kus n-rozměrné C^1 plochy.

V těchto případech je tedy „dimenze“ $f(U)$ (pro dostatečně malá otevřená okolí U bodu a) rovna hodnosti Jacobiho matice $[f'(a)]$ (tj. dimenzi oboru hodnot lineárního zobrazení $f'(a)$). Ukážeme, že totéž platí vždy, kdy hodnost $\mathrm{h}([f'(x)])$ *je konstantní na nějakém okolí bodu a.*

Bez tohoto předpokladu to však obecně pravda být nemusí: již pro $n = m = 1$ a $f(x) = x^2 \sin x$ je $\mathrm{h}([f'(0)]) = 0$, ale obraz $f(U)$ každého otevřeného okolí U bodu a má zřejmě neprázdný vnitřek, takže má dimenzi 1 (při každé „rozumné“ definici dimenze).

Ještě uveďme jeden konkrétní netriviální příklad, kdy hodnost $\mathrm{h}([f'(x)])$ je konstantní. Nechť $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ má složky $f_1(x,y) = x+2y$, $f_2(x,y) = x^2+4y^2+4xy-x-2y$. Pak

$$[f'(x,y)] = \begin{bmatrix} 1 & 2 \\ 2x+4y-1 & 4x+8y-2 \end{bmatrix},$$

takže $\mathrm{h}[f'(x,y)] = 1$ v každém bodě $(x,y) \in \mathbb{R}^2$. Protože $f_2(x,y) = (f_1(x,y))^2 - f_1(x,y)$, snadno vidíme, že $f(\mathbb{R}^2)$ je parabola $\{(u,v)\colon v = u^2 - u\}$; navíc není těžké dokázat, že $f(U)$ je explicitně zadaný kus 1-rozměrné C^1 plochy pro každou neprázdnou otevřenou množinu $U \subset \mathbb{R}^2$.

Skutečnost, že obor hodnot tohoto zobrazení $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ je „jednorozměrný“, souvisí s tím, že jeho složky f_1, f_2 nejsou „nezávislé“: hodnotu $f_2(x,y)$ lze vypočítat, známe-li hodnotu $f_1(x,y)$; tj. f_2 „je funkcí“ f_1. Tak jsme vedeni k následující definici.

2.173 Definice. *Nechť $f_1, f_2, \dots, f_k$, g jsou reálné funkce na množině M. Řekneme, že funkce g je (funkčně) závislá na funkcích $f_1, \dots, f_k$ (nebo také, že g je funkcí $f_1, \dots, f_k$), existuje-li funkce k proměnných $h(y_1, \dots, y_k)$ taková, že*

$$g(x) = h(f_1(x), \dots, f_k(x)), \quad x \in M.$$

Lze-li h zvolit tak, že je třídy C^p $(p \in \mathbb{N} \cup \{\infty\})$ na svém definičním oboru, řekneme, že funkce g je C^p-závislá na funkcích $f_1, \dots, f_k$.

2.174 Poznámka. Terminologie ohledně „funkční závislosti a nezávislosti“ v literatuře velmi kolísá a často se zavádějí i další pojmy. Někdy se například říká, že funkce $f_1, \dots, f_k$ na otevřené podmnožině G eukleidovského prostoru jsou funkčně závislé, jestliže některá z nich je závislá (případně C^1-závislá) na ostatních $k-1$ funkcích. Pak je tato definice zřejmým rozšířením pojmu lineární závislosti funkcí (kde za funkce h připouštíme jen funkce lineární). O funkcích $f_1, \dots, f_k$ se také často říká, že jsou funkčně nezávislé, jestliže nejsou funkčně závislé na žádné neprázdné otevřené množině $G^* \subset G$.

Častý je případ, kdy funkce $g(x_1, \dots, x_n)$ na $M \subset \mathbb{R}^n$ je závislá na několika souřadnicových funkcích, např. na $f_1(x_1, \dots, x_n) = x_1, \dots, f_k(x_1, \dots, x_n) = x_k$. Pak ovšem používáme frázi „funkce g závisí pouze na souřadnicích $x_1, \dots, x_k$“. Budeme potřebovat následující snadné tvrzení.

2.175 Tvrzení. *Nechť $p \in \mathbb{N} \cup \{\infty\}$, $1 \le k < n$ a g je funkce na otevřeném intervalu $I = I_1 \times \cdots \times I_n \subset \mathbb{R}^n$. Pak následující tvrzení jsou ekvivalentní:*

(i) *g je C^p-závislá na funkcích $x_1, \dots, x_k$.*

(ii) *Existuje funkce $h \in C^p(I_1 \times \cdots \times I_k)$ taková, že $g(x_1, \dots, x_n) = h(x_1, \dots, x_k)$ pro $(x_1, \dots, x_n) \in I$.*

(iii) *$\dfrac{\partial g}{\partial x_j}(x) = 0$ pro $k < j \le n$ a $x \in I$.*

Důkaz. Podmínky (i) a (ii) jsou zřejmě ekvivalentní a implikace $(ii) \Longrightarrow (iii)$ je zřejmá. Jestliže platí (iii), pak pro každý bod $(x_1, \dots, x_k) \in I_1 \times \cdots \times I_k$ je podle Věty 2.62 parciální funkce $(x_{k+1}, \dots, x_n) \mapsto g(x_1, \dots, x_n)$ konstantní na intervalu $I_{k+1} \times \cdots \times I_n$. Zvolíme-li tedy bod $a_{k+1} \in I_{k+1}, \dots, a_n \in I_n$ a položíme-li $h(x_1, \dots, x_k) := g(x_1, \dots, x_k, a_{k+1}, \dots, a_n)$, potom zřejmě $h \in C^p(I_1 \times \cdots \times I_k)$ a $g(x_1, \dots, x_n) = g(x_1, \dots, x_k, a_{k+1}, \dots, a_n) = h(x_1, \dots, x_k)$ pro $(x_1, \dots, x_n) \in I$.

2.176 Poznámka. Není těžké dokázat $(i) \Longleftrightarrow (iii)$ i v případě, že $I \subset \mathbb{R}^n$ je obecná konvexní otevřená množina; předpoklad otevřenosti a souvislosti I však nestačí.

Předpokládejme nyní, že funkce $f_1, \dots, f_k$ jsou třídy C^1 na nějakém otevřeném okolí bodu $a \in \mathbb{R}^n$. Naskýtá se přirozená *otázka, jak zjistit, zda na nějakém otevřeném okolí bodu a je některá z těchto funkcí závislá na ostatních funkcích.* Tato *druhá otázka* ovšem velmi úzce souvisí s naší první otázkou. Existuje-li například pro f z $\mathbb{R}^n$ do $\mathbb{R}^m$ číslo $1 \le k < m$ a otevřené okolí U bodu a, pro které je $f(U)$ explicitně zadaný k-rozměrný kus C^1 plochy, pak zřejmě existují indexy $1 \le i_1 < i_2 < \cdots < i_k \le m$ takové, že každá složka f_j, $j \in \{1, \dots, m\} \setminus \{i_1, \dots, i_k\}$, je na U C^1-závislá na složkách $f_{i_1}, \dots, f_{i_k}$.

Pro formulaci *třetí otázky* opět uvažujme zobrazení f z $\mathbb{R}^n$ do $\mathbb{R}^m$, které je třídy C^1 na nějakém otevřeném okolí bodu a. Ptejme se nyní *zda lze zvolit na nějakém otevřeném okolí bodu a křivočaré souřadnice $s_1, \dots, s_n$ a na nějakém otevřeném okolí bodu $f(a)$ křivočaré souřadnice $q_1, \dots, q_m$ tak, aby vyjádření (zúžení) zobrazení f pomocí těchto souřadnic mělo velmi jednoduchý „kanonický“ tvar.*

Ukážeme, že pokud hodnost $\mathrm{h}([f'(x)])$ je konstantní (rovna k) na nějakém okolí bodu a, pak to možné je: f má lokálně vzhledem ke vhodným křivočarým souřadnicím tvar

$$(s_1, \dots, s_n) \mapsto (q_1, \dots, q_m) = (s_1, \dots, s_k, 0, \dots, 0).$$

Přesná formulace odpovědí na všechny tři otázky je obsažena v následující větě.

2.177 Věta. (věta o hodnosti) *Nechť $f = (f_1, \dots, f_m)$ je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}^m$, které je třídy C^p ($p \in \mathbb{N} \cup \{\infty\}$) na nějakém otevřeném okolí bodu $a \in \mathbb{R}^n$. Nechť hodnost $\mathrm{h}([f'(x)])$ je rovna $k \ge 1$ na nějakém okolí bodu a. Potom:*

(i) *Existují indexy $1 \le j_1 < \cdots < j_k \le m$ a okolí U bodu a takové, že*

(a) *zobrazení $x \mapsto (f_{j_1}(x), \dots, f_{j_k}(x))$, $x \in U$ zobrazuje každou otevřenou množinu $G \subset U$ na otevřenou podmnožinu $\mathbb{R}^k$ a*

Obr. 2.15.

(b) každá z funkcí f_j, $j \in \{1, \dots, m\} \setminus \{j_1, \dots, j_k\}$ je na U C^p-závislá na funkcích $f_{j_1}, \dots, f_{j_k}$.

(ii) *Existuje otevřené okolí U bodu a takové, že pro každou otevřenou množinu $\emptyset \neq G \subset U$ platí, že*

(c) $f(G)$ je otevřená, jestliže $k = m$;

(d) $f(G)$ je explicitně zadaný kus k-rozměrné C^p-plochy, jestliže $k < m$.

(iii) *Existují otevřené okolí $U \subset D_f$ bodu a, otevřená množina $V \supset f(U)$ a difeomorfismy třídy C^p $S\colon U \to \mathbb{R}^n$, $Q\colon V \to \mathbb{R}^m$ takové, že zobrazení*

$$f^* := Q \circ f \circ S^{-1}, \quad f^*\colon S(U) \to \mathbb{R}^m$$

je dáno předpisem

$$f^*(s_1, \dots, s_n) = (s_1, \dots, s_k, 0, \dots, 0).$$

Důkaz. (Pro částečné znázornění tvrzení (iii) viz obr. 2.15.)

Podle předpokladu má Jacobiho matice $[f'(a)]$ aspoň jeden nenulový subdeterminant řádu k, tj. existují indexy $1 \leq i_1 < \dots < i_k \leq n$, $1 \leq j_1 < \dots < j_k \leq m$ takové, že $\dfrac{D(f_{j_1}, \dots, f_{j_k})}{D(x_{i_1}, \dots, x_{i_k})}(a) \neq 0$. Aby zápisy byly přehlednější a důkaz pochopitelnější, budeme dále bez újmy na obecnosti předpokládat, že $i_p = j_p = p$, tj. $\dfrac{D(f_1, \dots, f_k)}{D(x_1, \dots, x_k)}(a) \neq 0$. (Obecný případ lze na tento speciální zřejmě převést „přečíslováním" souřadnic bodu x a složek zobrazení f.)

Klíčem k důkazu věty je zavedení zobrazení $\widetilde{S}$ z $\mathbb{R}^n$ do $\mathbb{R}^n$ předpisem

$$\widetilde{S}(x_1, \dots, x_n) = (f_1(x), \dots, f_k(x), x_{k+1}, \dots, x_n), \quad x \in D_f.$$

Zobrazení $\widetilde{S}$ je zřejmě třídy C^p na nějakém otevřeném okolí bodu a a Jacobiho matice zobrazení $\widetilde{S}$ v bodě a má tvar (ve kterém hvězdičky označují libovolná reálná čísla)

$$[\widetilde{S}'(a)] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_k}(a) & * & \dots & \dots & * \\ \vdots & & \vdots & \vdots & & & \vdots \\ \frac{\partial f_k}{\partial x_1}(a) & \dots & \frac{\partial f_k}{\partial x_k}(a) & * & \dots & \dots & * \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & 0 & 0 & 1 & \dots & 0 \\ \vdots & & \vdots & \vdots & & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$

a je tedy regulární. (Snadno např. ověříme, že její řádky jsou lineárně nezávislé.) Podle Věty 2.121 o inverzním zobrazení existuje otevřené okolí U bodu a takové že $S := \widetilde{S} \restriction_U$ je difeomorfismus třídy C^p. Případným vhodným zmenšením množiny U můžeme zřejmě dosáhnout toho, že množina $U^* := S(U)$ je otevřený interval a $\mathrm{h}([f'(x)]) = k$ pro každý bod $x \in U$.

Protože S je difeomorfismus, pro každou otevřenou množinu $G \subset U$ je množina $S(G)$ otevřená, takže je otevřená i projekce množiny $S(G)$ na prostor prvních k souřadnic. Tím je dokázáno tvrzení (a).

Zobrazení $g := f \circ S^{-1}$, $g\colon U^* \to \mathbb{R}^m$ je zřejmě třídy C^p na U^*. Položíme-li $x = (x_1, \dots, x_n) := S^{-1}(s)$ pro libovolný bod $s = (s_1, \dots, s_n) \in U^*$, zřejmě platí

$$g_j(s) = f_j(x), \quad j = 1, \dots, m; \qquad g_j(s) = f_j(x) = s_j, \quad j = 1, \dots, k. \tag{2.53}$$

Jacobiho matice $[g'(s)]$ zobrazení g v bodě s má tedy tvar

$$[g'(s)] = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ * & \dots & \dots & * & \frac{\partial g_{k+1}(s)}{\partial s_{k+1}} & \dots & \frac{\partial g_{k+1}(s)}{\partial s_n} \\ \vdots & & & \vdots & \vdots & & \vdots \\ * & \dots & \dots & * & \frac{\partial g_m(s)}{\partial s_{k+1}} & \dots & \frac{\partial g_m(s)}{\partial s_n} \end{bmatrix} \tag{2.54}$$

Pomocí souřadnicového systému S jsme zavedli na U křivočaré souřadnice; nyní pomocí Tvrzení 2.175 dokážeme, že „zobrazení f závisí jen na prvních k z nich“. Přesněji: ukážeme, že (složky) $g(s_1, \dots, s_n)$ závisí pouze na $s_1, \dots, s_k$. Protože $(S^{-1})'(s)$ je lineární bijekce $\mathbb{R}^n$ na $\mathbb{R}^n$ a $g'(s) = f'(x) \circ (S^{-1})'(s)$, dostáváme

$$\mathrm{h}([g'(s)]) = \dim\left(\mathrm{Im}(g'(s))\right) = \dim\left(\mathrm{Im}(f'(x))\right) = \mathrm{h}([f'(x)]) = k.$$

Z tvaru (2.54) matice $[g'(s)]$ tedy okamžitě vidíme, že

$$\frac{\partial g_j(s)}{\partial s_i} = 0, \quad \text{pro} \quad k+1 \le i \le n, \ k+1 \le j \le m \ \ s \in U^*.$$

Protože U^* je interval, je tvaru $U^* = I_1 \times \cdots \times I_n$, kde $I_j \subset \mathbb{R}$ jsou otevřené intervaly. Z Tvrzení 2.175 vyplývá, že existují funkce φ_j, $j = k+1, \ldots, m$ třídy C^p na $I_1 \times \cdots \times I_k$ takové, že

$$g_j(s) = \varphi_j(s_1, \ldots, s_k), \quad s \in U^*, \qquad j = k+1, \ldots, m.$$

Podle (2.53) tedy pro každé $k+1 \leq j \leq m$ platí

$$f_j(x) = \varphi_j(f_1(x), \ldots, f_k(x)), \quad x \in U.$$

Tím je dokázáno i tvrzení (b).

Z tvrzení (a), (b) již okamžitě vyplývá, že pro výše uvedené U platí také tvrzení (c) a (d).

Zbývá dokázat (iii). Víme, že zobrazení $g = f \circ S^{-1}$ je tvaru

$$g(s_1, \ldots, s_n) = (s_1, \ldots, s_k, \varphi_{k+1}(s_1, \ldots, s_k), \ldots, \varphi_m(s_1, \ldots, s_k)).$$

Otevřená množina $V := I_1 \times \cdots \times I_k \times \mathbb{R}^{m-k}$ zřejmě obsahuje množinu $f(U)$. Položíme-li pro $y = (y_1, \ldots, y_m) \in V$

$$Q(y_1, \ldots, y_m) := (y_1, \ldots, y_k, y_{k+1} - \varphi_{k+1}(y_1, \ldots, y_k), \ldots, y_m - \varphi_m(y_1, \ldots, y_k)),$$

je snadno vidět, že $Q\colon V \to V$ je difeomorfismus třídy C^p a zobrazení

$$f^* := Q \circ f \circ S^{-1} = Q \circ g, \qquad f^*\colon S(U) = U^* \to \mathbb{R}^m$$

má žádaný („kanonický“) tvar

$$f^*(s_1, \ldots, s_n) = (s_1, \ldots, s_k, 0, \ldots, 0).$$

2.178 Poznámka.

(α) Je-li $\mathrm{h}([f'(x)]) = 0$ na nějakém okolí bodu a, pak z Věty 2.62 snadno vyplývá, že f je konstantní na nějakém okolí bodu a.

(β) Z (a) snadno vyplývá, že žádná z funkcí $f_{j_1}, \ldots, f_{j_k}$ není funkčně závislá na ostatních na žádné otevřené množině $\emptyset \neq G \subset U$. (V tomto smyslu jsou tedy $f_{j_1}, \ldots, f_{j_k}$ „nezávislé na G“.)

(γ) Z (d) vyplývá (např. podle Tvrzení 2.156), že $f(G)$ je řídká a lebesgueovsky nulová v $\mathbb{R}^m$.

3. Úvod do diferenciálního počtu v Banachových prostorech

Smyslem tohoto krátkého úvodu je seznámit čtenáře s některými základními pojmy, metodami a výsledky této teorie (někdy bez důkazu nebo pouze s náznakem důkazu). Konkrétní cíle jsou čtyři:

(i) Podat úplný důkaz věty o inverzním zobrazení (pro C^1 zobrazení mezi Banachovými prostory). V Kapitole 2 je přesně ukázáno, jak lze odtud snadno vyvodit klasickou větu o implicitních funkcích. Důkaz věty o implicitním zobrazení mezi Banachovými prostory je pouze naznačen.

(ii) Ukázat moderní definici derivací zobrazení vyššího řádu, při které například chápeme druhou derivaci $x \mapsto F''(x)$ jako derivaci zobrazení $x \mapsto F'(x)$. Druhá derivace je tedy opět (stejně jako v případě funkcí jedné proměnné) definovaná jako derivace derivace. Po jistém „přirozeném ztotožnění" je při této definici druhá derivace bilineární zobrazení, které je pro případ zobrazení $F\colon \mathbb{R}^n \to \mathbb{R}$ druhou derivací ve smyslu Kapitoly 2.

(iii) Ukázat, že toto moderní pojetí vyšších derivací může být užitečné i pro diferenciální počet v eukleidovských prostorech. Jako příklad uvádíme „bezsouřadnicový" výpočet derivací druhého a třetího řádu složené funkce, při kterém si ušetříme psaní poměrně dlouhých vzorců – ovšem za cenu (myšlenkově náročné) práce v abstraktních prostorech.

(iv) Na konkrétních příkladech ukázat, jak v „abstraktním kalkulu" přebírá roli součinu obecné bilineární (resp. multilineární) zobrazení a roli reálné funkce $x \mapsto 1/x$ zobrazení $S \mapsto S^{-1}$ (kde $S\colon X \to X$ je izomorfismus normovaného lineárního prostoru na sebe).

V této kapitole budeme někdy pro přehlednost zápisu (zmenšení počtu závorek) používat znak $\mathbf{.}$ pro „operaci evaluace". Například místo $f(x)$ budeme někdy psát $f\mathbf{.}x$. V jednoduchých případech však vynecháváme závorku i znak $\mathbf{.}$; například jako obvykle píšeme $f'(x)h$.

3.1 Fréchetova a Gâteauxova derivace; derivace složeného zobrazení

V tomto oddíle budeme předpokládat, že X ($\dim X \geq 1$) a Y jsou normované lineární prostory a f je zobrazení z X do Y. V každé větě jsou buď *všechny* normované lineární prostory reálné nebo jsou komplexní. Čtenáři lze doporučit, aby si vždy představoval „reálný případ“, který je pochopitelnější.

Často se teorie buduje pouze pro případ Banachových prostorů, který má nejdůležitější aplikace, a ve kterém platí hlubší věty; pro základní definice a nejjednodušší definice však předpoklad úplnosti není nutný.

Pojem derivace podle vektoru se ve variačním počtu fakticky používal podstatně dříve, než vznikla teorie normovaných lineárních prostorů. Jde o přímočaré zobecnění definice v klasickém případě.

3.1 Definice. *Derivací zobrazení f v bodě $a \in X$ podle vektoru $v \in X$ rozumíme limitu*

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

3.2 Poznámka.

(i) Pro zápis derivace podle vektoru se někdy používají také symboly

$$\partial_v f(a), f'(a,v), \delta_v f(a).$$

(ii) Pro derivaci podle vektoru se používá i název směrová derivace nebo (při užití zápisu $\delta_v f(a)$) Gâteauxova (nebo Lagrangeova) variace. Často se místo „derivace podle vektoru v“ užívá fráze „derivace ve směru v“

(iii) V některých aplikacích má velký význam i pojem jednostranné derivace podle vektoru, definované předpisem

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

Pojem parciální derivace zobrazení f nelze obecně rozumně definovat. Přirozené zobecnění pojmu parciální derivace pro zobrazení z kartézského součinu normovaných lineárních prostorů je podáno níže (Definice 3.28).

Základním pojmem diferenciálního počtu v normovaných lineárních prostorech je pojem *Fréchetovy derivace*, který přirozeně zobecňuje klasický pojem totálního diferenciálu.

3.3 Definice. *Nechť f je zobrazení z normovaného lineárního prostoru X do normovaného lineárního prostoru Y a nechť $L\colon X \to Y$ je spojité lineární zobrazení. Řekneme, že L je Fréchetova derivace zobrazení f v bodě a, jestliže*

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - L(h)}{\|h\|} = 0. \tag{3.1}$$

Fréchetovu derivaci zobrazení f v bodě a budeme značit symbolem $f'(a)$ a někdy ji budeme nazývat pouze derivace.

Fréchetově derivaci se také říká Fréchetův diferenciál a místo $f'(a)$ se často píše $Df(a)$. Existuje-li $f'(a)$ (resp. $f'(x)$ pro každé x z otevřené množiny $G \subset X$), říkáme, že f je diferencovatelné v a (resp. na G).

Pro korektnost definice je ovšem třeba vědět, že derivace zobrazení f v bodě a existuje nejvýše jedna; to plyne ihned z Tvrzení 3.6.

3.4 Poznámka.

(i) Aby měl pojem Fréchetovy derivace užitečné vlastnosti, je třeba požadovat nejen linearitu zobrazení $f'(a)$, ale také jeho *spojitost.*

Pro případ $X = \mathbb{R}^n, Y = \mathbb{R}^m$ je ovšem (viz oddíl 1.11) každé lineární zobrazení L: $X \to Y$ spojité, takže Definice 3.3 je skutečně zobecněním definice z Kapitoly 2.

(ii) Snadno je vidět, že pokud f je spojité v a a pro lineární zobrazení L: $X \to Y$ platí (3.1), je L spojité.

(iii) Přímo z definice okamžitě vyplývá, že pokud f: $X \to Y$ je spojité lineární zobrazení, pak $f'(a) = f$ pro každý bod $a \in X$.

(iv) Je-li f: $X \to Y$ konstantní na otevřené množině $G \subset X$, pak zřejmě platí $f'(a) = 0 \in \mathcal{L}(X, Y)$ pro každý bod $a \in G$.

(iv) Snadno je vidět, že pokud v X a Y přejdeme k ekvivalentním normám, pojem Fréchetovy derivace se nezmění.

Zcela stejným způsobem, jakým jsme dokázali v klasickém případě Tvrzení 2.8, Větu 2.14 a rovnost $D_v f(a) = \mathrm{d}f(a)(v)$ (z Věty 2.28), lze snadno dokázat tato jejich zobecnění.

3.5 Tvrzení. *Nechť L: $X \to Y$ je spojité lineární zobrazení a $r(h)$ je zobrazení určené rovnicí*

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

Pak následující tvrzení jsou ekvivalentní.

(i) $f'(a) = L.$

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

(iii) *Existuje zobrazení ω z X do Y, které je spojité v $0 \in X$, a pro které platí $\omega(0) = 0$ a $r(h) = \|h\|\,\omega(h)$.*

3.6 Tvrzení. *Nechť existuje Fréchetova derivace $f'(a)$. Pak f je spojité v bodě a a pro každý vektor $v \in X$ platí*

$$D_v f(a) = f'(a)v.$$

Často se užívá i následující (méně důležitý) pojem Gâteauxovy derivace.

3.7 Definice. *Řekneme, že spojité lineární zobrazení $L\colon X \to Y$ je Gâteauxova derivace zobrazení f v bodě $a \in X$, jestliže pro každý vektor $v \in X$ platí rovnost $L(v) = D_v f(a)$. Gâteauxovu derivaci budeme dále označovat symbolem $f'_G(a)$.*

Zobrazení f má tedy Gâteauxovu derivaci v bodě a, právě když existují všechny směrové derivace $D_v f(a)$ a zobrazení $v \mapsto D_v f(a)$ je lineární a spojité; toto zobrazení je pak $f'_G(a)$. Jednoznačnost Gâteauxovy derivace je tedy zřejmá.

Z Tvrzení 3.6 okamžitě vyplývá, že pojem Gâteauxovy derivace je slabší než pojem Fréchetovy derivace:

Existuje-li $f'(a)$, existuje i $f'_G(a)$ a platí $f'(a) = f'_G(a)$.

Někdy se proto Gâteauxova derivace nazývá „slabá derivace" a Fréchetova derivace „silná derivace".

3.8 Poznámka. V případě, že f je reálná funkce na reálném Hilbertově prostoru X, $a \in X$ a existuje $f'(a)$, pak je přirozené definovat gradient $\operatorname{grad} f(a) \in X$ jako jediný prvek X, pro který $f'(a) = \langle \cdot, \operatorname{grad} f(a) \rangle$ (srov. Věta 1.137). Při této definici platí zobecnění Věty 2.28, speciálně $\operatorname{grad} f(a)$ určuje směr největšího růstu funkce. Poznamenejme ještě, že někteří autoři při definici gradientu pracují s $f'_G(a)$ místo $f'(a)$ a že i pak platí zobecnění Věty 2.28.

Užitečné je následující snadné tvrzení, které více objasňuje vztah mezi Gâteauxovou a Fréchetovou derivací, a které lze také chápat jako alternativní definici Fréchetovy derivace.

3.9 Tvrzení. *Fréchetova derivace $f'(a)$ existuje právě tehdy, existuje-li Gâteauxova derivace $f'_G(a)$ a limita*

$$\lim_{t \to 0} \frac{f(a+tv) - f(a)}{t} = D_v f(a) = f'_G(a)v \tag{3.2}$$

je stejnoměrná vzhledem k vektorům v z jednotkové sféry $S := \{v \in X\colon \|v\| = 1\}$.

Důkaz. Předpokládejme, že existuje $L := f'_G(a)$. Pak zmíněná stejnoměrnost limity (3.2) ovšem znamená toto (viz Poznámka 1.34):

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall v \in S\, \forall t \in \mathbb{R}\colon \quad (0 < |t| < \delta) \implies \left\| \frac{f(a+tv) - f(a)}{t} - L(v) \right\| < \varepsilon. \tag{3.3}$$

Podle definice platí $L = f'(a)$, právě když

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall h \in X\colon \quad (0 < \|h\| < \delta) \implies \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|} < \varepsilon. \tag{3.4}$$

Položíme-li $h := tv$ pro $v \in S$ a $t \neq 0$, pak $\|h\| = |t|$ a $L(h) = t\,L(v)$, takže

$$\left\| \frac{f(a+tv) - f(a)}{t} - L(v) \right\| = \frac{\|f(a+h) - f(a) - L(h)\|}{\|h\|}$$

Uvážíme-li ještě, že každý vektor $0 \neq h \in X$ lze psát ve tvaru $h = t\,v$, kde $t = \|h\|$ a $v = h/\|h\| \in S$, pak snadno vidíme, že tvrzení (3.3) a (3.4) jsou ekvivalentní.

3.10 Poznámka.

(i) Z předchozího důkazu je zřejmé, že pokud v (3.2) uvažujeme jednostrannou limitu ($\lim_{t\to 0+}$), tvrzení opět platí.

(ii) Je snadné ukázat, že Gâteauxova derivace $f'_G(a$ je Fréchetovou derivací, právě když je splněna některá ze dvou podmínek

(a) Existuje okolí U bodu $0 \in X$ takové, že limita (3.2) je stejnoměrná vzhledem k $v \in U$.

(b) Pro každou omezenou množinu $M \subset X$ je limita (3.2) stejnoměrná vzhledem k $v \in M$.

Někdy je vyšetření existence Gâteauxovy derivace snazší než vyšetření Fréchetovy derivace. V těchto případech je důležitá následující věta, kterou uvedeme bez důkazu. Poznamenejme jen, že v klasickém případě ($X = \mathbb{R}^n$, $Y = \mathbb{R}^m$) snadno vyplývá z Věty 2.20 a v obecném případě se snadno dokáže z Věty 3.19 o přírůstku vektorové funkce.

3.11 Věta. *Nechť X, Y jsou normované lineární prostory, $a \in X$ a f je zobrazení z X do Y. Jestliže f má Gâteauxovu derivaci v každém bodě nějakého okolí bodu a a Gâteauxova derivace $x \mapsto f'_G(x)$ (zobrazení z X do $\mathcal{L}(X,Y)$) je spojitá v bodě a, pak existuje Fréchetova derivace $f'(a)$.*

Pro Fréchetovu derivaci platí také přirozené zobecnění Tvrzení 2.48:

Existuje-li Fréchetova derivace $f'(a)$, pak $\|f(a+h) - f(a)\| = O(\|h\|)$, $h \to 0$.

Použijeme-li toto tvrzení a Tvrzení 3.5, můžeme postupovat zcela stejně jako při důkazu Věty 2.50 o derivaci složeného zobrazení a dokázat její přirozené zobecnění.

3.12 Věta. (věta o Fréchetově derivaci složeného zobrazení) *Nechť X, Y, Z jsou normované lineární prostory, f je zobrazení z X do Y a g je zobrazení z Y do Z. Nechť f má Fréchetovu derivaci v bodě $a \in X$ a nechť g má Fréchetovu derivaci v bodě $b := f(a)$. Pak zobrazení $g \circ f$ má Fréchetovu derivaci v bodě a a platí*

$$(g \circ f)'(a) = g'(b) \circ f'(a).$$

Důležitým speciálním případem je následující tvrzení.

3.13 Důsledek. *Nechť X, Y, Z jsou normované lineární prostory, f je zobrazení z X do Y a $L \in \mathcal{L}(Y, Z)$. Má-li f Fréchetovu derivaci v bodě $a \in X$, pak zobrazení $L \circ f$ má Fréchetovu derivaci v bodě a a platí $(L \circ f)'(a) = L \circ f'(a)$.*

Z něj pak snadno plyne:

3.14 Důsledek. *Nechť X, Y_1, Y_2 jsou normované lineární prostory, Y_1 a Y_2 jsou izomorfní a φ: $Y_1 \to Y_2$ je izomorfismus. Nechť f je zobrazení z X do Y_1, a $a \in X$. Položme $f^* := \varphi \circ f$. Pak platí rovnost $(f^*)'(a) = \varphi \circ f'(a)$, jakmile má jedna její strana smysl.*

Snadným důsledkem věty o derivaci složeného zobrazení je následující tvrzení.

3.15 Důsledek. *Nechť X, Y, Z jsou normované lineární prostory, f je zobrazení z X do Y a g je zobrazení z Y do Z. Pak platí následující tvrzení.*

(i) *Jestliže $a, v \in X$ a existují $D_v f(a)$ a $g'(f(a))$, pak existuje $D_v (g \circ f)(a)$ a platí*

$$D_v (g \circ f)(a) = g'(f(a)) \cdot D_v f(a).$$

(ii) *Nechť f má Gâteauxovu derivaci v bodě $a \in X$ a nechť g má Fréchetovu derivaci v bodě $b := f(a)$. Pak zobrazení $g \circ f$ má Gâteauxovu derivaci v bodě a a platí*

$$(g \circ f)'_G(a) = g'(b) \circ f'_G(a).$$

Důkaz. (Náznak.) Tvrzení (i) snadno dostáváme, aplikujeme-li Větu 3.12 na zobrazení f^* z $\mathbb{R}$ do Y definované předpisem $f^*(t) := f(a + tv)$, zobrazení g a na bod $a^* := 0 \in R$. Tvrzení (ii) je pak okamžitým důsledkem tvrzení (i) a toho, že složení $g'(b) \circ f'_G(a)$ dvou spojitých lineárních zobrazení je spojité lineární zobrazení.

3.16 Poznámka. Nechť X, Y, Z, f, g jsou jako ve Větě 3.12. Existuje-li $f'(a)$ a $g'_G(f(a))$, nemusí obecně existovat $(g \circ f)'_G(a)$, a ani v případě, že existuje, nemusí platit $(g \circ f)'_G(a) = g'_G(f(a)) \circ f'_G(a)$ (Stačí volit $X = \mathbb{R}$, $Y = \mathbb{R}^2$, $Z = \mathbb{R}$, $f(x) = (x, x^2)$, $a = 0$ a za g vzít po řadě funkce g^*, g z Příkladu 2.33). Předpokládáme-li však, že zobrazení g je *lipschitzovské* na nějakém okolí bodu $f(a)$, pak z existence *Gâteauxových* derivací $f'_G(a)$, $g'_G(f(a))$ lze (poměrně snadno) dokázat existenci Gâteauxovy derivace $(g \circ f)'_G(a)$ a rovnost $(g \circ f)'_G(a) = g'_G(f(a)) \circ f'_G(a)$.

Zcela analogicky jako bylo dokázáno Tvrzení 2.38, lze dokázat jeho následující zobecnění.

3.17 Věta. *Nechť $f = (f_1, \ldots, f_n)$ je zobrazení z normovaného lineárního prostoru Y do součinu $X := X_1 \times \cdots \times X_n$ normovaných lineárních prostorů a nechť $a \in Y$. Pak*

$$f'(a) = (f'_1(a), \ldots, f'_n(a)),$$

má-li jedna strana rovnosti smysl.

Velmi užitečné je také následující tvrzení.

3.18 Tvrzení. *Nechť X je n-rozměrný normovaný lineární prostor, Z je normovaný lineární prostor a nechť $(\varphi_1, \ldots, \varphi_n)$ je báze duálního prostoru X^*. Nechť f je* *zobrazení ze Z do X. Pak f má derivaci v bodě $a \in Z$, právě když všechny reálné (resp. komplexní) funkce $\varphi_i \circ f$ mají derivaci v bodě a.*

Důkaz. (Stručný.) Nechť X a Z jsou reálné (komplexní případ je zcela obdobný). Pak zobrazení $\varphi = (\varphi_1, \ldots, \varphi_n)\colon X \to \mathbb{R}^n$ je lineární bijekce, takže je to podle Věty 1.130 izomorfismus (na $\mathbb{R}^n$ uvažujeme maximovou normu). Podle Důsledku 3.14 má f v bodě a derivaci, právě když $\varphi \circ f$ má v tomto bodě derivaci. A to podle Věty 3.17 nastane právě tehdy, když všechny funkce $\varphi_i \circ f$ mají derivaci v bodě a.

3.19 Věta. *Nechť X, Y jsou normované lineární prostory, $C \subset X$ je otevřená konvexní množina, $F\colon C \to Y$ je zobrazení diferencovatelné na množině C a nechť*

$$\sup\{\|F'(x)\|\colon x \in C\} \le K.$$

Pak zobrazení F je na C lipschitzovské s konstantou K, tj.

$$\|F(b) - F(a)\| \le K\,\|b - a\|, \quad \textit{kdykoliv}\ \ a, b \in C. \tag{3.5}$$

Důkaz. (Náznak.) Naznačíme dvě možnosti důkazu.

První možnost (kterou ukážeme jen v případě reálných prostorů X, Y) je založena na užití Lagrangeovy věty podobně jako v důkazu Věty 2.64. Zvolíme libovolně a, $b \in C$ a použijeme Větu 1.136 (kterou jsme uvedli bez důkazu) k nalezení funkcionálu $\psi \in Y^*$ takového, že platí $\|\psi\| = 1$ a $\psi(F(b) - F(a)) = \|F(b) - F(a)\|$. Použijeme-li Lagrangeovu větu na reálnou funkci $f(t) = \psi(F(a+t(b-a)))$ na $[0,1]$, dostáváme

$$\|F(b) - F(a)\| = \psi(F(b)) - \psi(F(a)) = f(1) - f(0) = f'(\xi),$$

kde $\xi \in (0,1)$. Lagrangeovu větu lze použít, protože podle Věty 3.12 má f v každém bodě $t \in [0,1]$ derivaci $f'(t) \in \mathcal{L}(\mathbb{R},\mathbb{R})$ a $f'(t) = \psi \circ F'(a + t(b-a)) \circ L$, kde $L \in \mathcal{L}(\mathbb{R}, X)$, $L(h) = h(b-a)$. Platí tedy nerovnosti (srov. Tvrzení 1.135)

$$\|f'(\xi)\| \le \|\psi\| \cdot \|F'(a + \xi(b-a))\| \cdot \|b - a\| \le K\|b - a\|.$$

Chápeme-li $f'(\xi)$ klasicky jako reálné číslo, dostáváme nerovnost (srov. Poznámka 2.37 (iii)) $|f'(\xi)| \le K\|b-a\|$, takže (3.5) platí.

Druhá možnost je zcela přímočará. Pro body $a, b \in C$ a $\varepsilon > 0$ položme

$$M := \{t \in [0,1]\colon \|F(a + t(b-a)) - F(a)\| \le (K+\varepsilon)t\|b-a\|\}.$$

Zřejmě $0 \in M$ a ze spojitosti F dostáváme, že $\tau := \sup M \in M$. Pokud $\tau < 1$, z definice derivace snadno vyplývá existence čísla $t \in (\tau, 1)$, pro které platí $\|F(a + t(b-a)) - F(a + \tau(b-a))\| \le (K+\varepsilon)(t-\tau)\|b-a\|$. Pak ale

$$\begin{aligned}\|F(a+t(b-a))-F(a)\| &\le \|F(a+\tau(b-a))-F(a)\|+\|F(a+t(b-a))-F(a+\tau(b-a))\|\\ &\le (K+\varepsilon)\tau\|b-a\| + (K+\varepsilon)(t-\tau)\|b-a\| = (K+\varepsilon)t\|b-a\|,\end{aligned}$$

takže $t \in M$, což je spor. Je tedy $\tau = 1 \in M$, tj. $\|F(b) - F(a)\| \le (K + \varepsilon)\|b - a\|$. Protože ε bylo libovolné, platí (3.5).

3.20 Poznámka. Pro odhad $\|F(b) - F(a)\|$ jsme využili pouze odhad velikosti derivace na úsečce $\overline{ab}$. Platí zobecnění tvrzení z Poznámky 2.65 (ii).

Někdy se používá také pojem striktní (ostré, silné) derivace, který je *silnější* než pojem Fréchetovy derivace.

Nechť X, Y jsou normované lineární prostory a f je zobrazení z X do Y. Řekneme, že $L \in \mathcal{L}(X, Y)$ je striktní derivace zobrazení f v bodě a, jestliže pro každé $\varepsilon > 0$ existuje okolí U bodu a takové, že zobrazení $f - L$ je na U lipschitzovské s konstantou ε.

Je snadno vidět, že L je striktní derivace f v a, právě když $L = f'(a)$ a navíc

$$\lim_{(x,y)\to(a,a),x\neq y} \frac{f(x) - f(y) - f'(a)(x - y)}{\|x - y\|} = 0.$$

Tento pojem derivace pro funkce z $\mathbb{R}$ do $\mathbb{R}$ definoval již koncem 19. století Peano. Poznamenejme ještě toto: a) Z předpokladů Věty 3.11 vyplývá, že $f'(a)$ je striktní derivace. b) Je-li X normovaný lineární prostor, $f\colon X \to \mathbb{R}$ je spojitá konvexní funkce, $a \in X$ a existuje $f'(a)$, pak $f'(a)$ je striktní derivace. c) Jsou-li ve Větě 3.12 o derivaci složeného zobrazení derivace $f'(a)$, $g'(b)$ striktní, pak je také $(g \circ f)'(a)$ striktní derivace.

3.2 Zobrazení třídy C^1

3.21 Definice. *Nechť X, Y jsou normované lineární prostory a $G \subset X$ je otevřená množina. Řekneme, že zobrazení $f\colon G \to Y$ je třídy C^1 (je spojitě diferencovatelné), jestliže má Fréchetovu derivaci v každém bodě a zobrazení $f'\colon x \mapsto f'(x)$ je spojité zobrazení množiny G do prostoru $\mathcal{L}(X, Y)$.*

Množinu všech takových zobrazení označujeme symbolem $C^1(G; Y)$; v případě $Y = \mathbb{R}$ pouze $C^1(G)$.

3.22 Poznámka.

(i) Později budeme definovat i pro $p > 1$ ($p \in \mathbb{N}$) množinu $C^p(G; Y)$ zobrazení G do Y, která jsou p-krát spojitě diferencovatelná.

(ii) Z Věty 3.11 ihned vyplývá, že $f \in C^1(G; Y)$, právě když Gâteauxova derivace $f'_G : x \mapsto f'_G(x)$ je spojitá na G.

Pojem zobrazení třídy C^1 jsme již definovali pro zobrazení $f\colon \mathbb{R}^n \to \mathbb{R}^m$. Musíme tedy ukázat, že nová definice není ve sporu se starou. To okamžitě vyplývá z následujícího snadného tvrzení.

3.23 Tvrzení. *Nechť $G \subset \mathbb{R}^n$ je otevřená množina a nechť je dáno zobrazení $f = (f_1, \dots, f_m)\colon G \to \mathbb{R}^m$. Pak následující tvrzení jsou ekvivalentní.*

(i) *Zobrazení $f'\colon x \mapsto f'(x)$ je spojité na G.*

(ii) *Všechny funkce $\dfrac{\partial f_j}{\partial x_i}$, $i = 1, \dots, n$, $j = 1, \dots, m$, jsou spojité na G.*

Důkaz. Nejprve si uvědomíme (pro názorný důkaz lze „ztotožnit“ $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ s prostorem matic typu $m \times n$ pomocí lineární bijekce $L \mapsto [L]$), že lineární formy $\varphi_{i,j} \in (\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))^*$, $i = 1, \dots, m$, $j = 1, \dots, n$, dané předpisem

$$\varphi_{i,j}(L) = \langle L(e_j), e_i \rangle, \quad L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$$

tvoří bázi prostoru $(\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))^*$.

Platí-li (i), víme (viz (2.10)), že $\frac{\partial f_i}{\partial x_j}(x) = \langle f'(x)e_j, e_i \rangle = \varphi_{i,j}(f'(x))$ pro $x \in G$, takže všechny funkce $\frac{\partial f_i}{\partial x_j} = \varphi_{i,j} \circ f'$ jsou spojité.

Nechť platí (ii). Pak podle Věty 2.44 existuje $f'(x)$ pro každé $x \in G$. Protože funkce $\varphi_{i,j} \circ f'$, $i = 1, \dots, n$, $j = 1, \dots, m$, jsou spojité, je podle Tvrzení 1.142 spojitá i derivace $f'\colon G \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Zcela analogicky jako v klasickém případě se definuje pojem difeomorfismu.

3.24 Definice. *Nechť X, Y jsou normované lineární prostory a f je zobrazení z X do Y. Řekneme, že f je difeomorfismus, jestliže f je bijekce otevřené množiny $G \subset X$ na otevřenou množinu $H \subset Y$ a zobrazení f, f^{-1} jsou třídy C^1.*

3.3 Zobrazení ze součinu prostorů

V diferenciálním počtu v normovaných lineárních prostorech hraje pojem spojitého bilineárního zobrazení roli „zobecněného součinu“. To uvidíme zejména ve Větě 3.26.

3.25 Věta. *Nechť X, Y, Z jsou normované lineární prostory a $B\colon X \times Y \to Z$ je spojité bilineární zobrazení. Pak B má Fréchetovu derivaci v každém bodě $(a, b) \in X \times Y$ a platí*

$$B'(a,b)\,.\,(h,k) = B(h,b) + B(a,k).$$

Důkaz. Platí

$$B(a+h, b+k) - B(a,b) = B(h,b) + B(a,k) + B(h,k),$$

a zobrazení $(h,k) \mapsto B(h,b) + B(a,k)$ je prvkem prostoru $\mathcal{L}(X \times Y, Z)$ (srov. Poznámka 1.146). Stačí tedy dokázat, že

$$\|B(h,k)\| = o(\|(h,k)\|), \quad (h,k) \to 0.$$

Tento vztah však okamžitě plyne z nerovností

$$\|B(h,k)\| \le \|B\| \cdot \|h\| \cdot \|k\| \le \|B\| \cdot \|(h,k)\|^2.$$

Jako důsledek ovšem dostáváme větu o derivaci „zobecněného součinu" dvou zobrazení.

3.26 Věta. *Nechť X, Y_1, Y_2 a Z jsou normované lineární prostory, $a \in X$ a B: $Y_1 \times Y_2 \to Z$ je spojité bilineární zobrazení. Nechť zobrazení f_1 z X do Y_1 a f_2 z X do Y_2 jsou diferencovatelná v bodě $a \in X$. Pak zobrazení $g := B \circ (f_1, f_2)$ má derivaci v bodě a, přičemž*

$$g'(a)h = B(f_1'(a)h, f_2(a)) + B(f_1(a), f_2'(a)h), \quad h \in X. \tag{3.6}$$

Důkaz. Stačí si uvědomit, že $g = B \circ f$, kde $f := (f_1, f_2)$ je zobrazení z X do $Y_1 \times Y_2$, použít Větu 3.12 o derivaci složeného zobrazení a předchozí Větu 3.25. Podle Věty 3.17 $f'(a) = (f_1'(a), f_2'(a))$, takže existuje $g'(a)$ a

$$\begin{aligned} g'(a)h = (B'(f_1(a), f_2(a)) \circ f'(a))\, h &= B'(f_1(a), f_2(a)) \,.\, (f_1'(a)h, f_2'(a)h) \\ &= B(f_1'(a)h, f_2(a)) + B(f_1(a), f_2'(a)h). \end{aligned}$$

Zcela analogicky jako Větu 3.25 a Větu 3.26 lze dokázat jejich zobecnění pro multilineární zobrazení:

3.27 Věta. *Nechť $Y_1, Y_2, \dots, Y_n$, Z a X jsou normované lineární prostory, $b = (b_1, \dots, b_n) \in Y := Y_1 \times \dots \times Y_n$ a M: $Y_1 \times \dots \times Y_n \to Z$ je spojité multilineární zobrazení. Pak existuje $M'(b)$ a pro $h = (h_1, \dots, h_n) \in Y$ platí*

$$M'(b) \,.\, h = M(h_1, b_2, \dots, b_n) + M(b_1, h_2, b_3, \dots, b_n) + \dots + M(b_1, \dots, b_{n-1}, h_n).$$

Jestliže navíc zobrazení f_i z X do Y_i, $i = 1, \dots, n$, jsou diferencovatelná v bodě $a \in X$, pak zobrazení $g := M \circ (f_1, \dots, f_n)$ má derivaci v bodě a, přičemž pro $h \in X$ platí

$$g'(a)h = M(f_1'(a)h, f_2(a), \dots, f_n(a)) + \dots + M(f_1(a), \dots, f_{n-1}(a), f_n'(a)h). \tag{3.7}$$

Pro zobrazení definovaná na podmnožině součinu normovaných lineárních prostorů lze zavést pojem parciálních derivací zcela analogicky jako pro funkce více proměnných.

3.28 Definice. *Nechť $X_1, X_2, \dots, X_n$ a Y jsou normované lineární prostory, f je zobrazení z $X := X_1 \times \dots \times X_n$ do Y, $a = (a_1, \dots, a_n) \in X$ a $1 \le i \le n$. Pak klademe*

$$f_i'(a) := (f(a_1, \dots, a_{i-1}, \cdot, a_{i+1}, \dots, a_n)'(a_i)),$$

má-li pravá strana rovnosti smysl. Existuje-li $f_i'(a)$, nazýváme ji parciální derivací (podle i-té proměnné) zobrazení f v bodě a; pak zřejmě $f_i'(a) \in \mathcal{L}(X_i, Y)$.

Pro označení parciální derivace $f_i'(a)$ se také používají symboly

$$D_i f(a),\ f'_{x_i},\ \frac{\partial f}{\partial x_i}(a).$$

Analogicky lze definovat i parciální derivace vyšších řádů. Pro takto zobecněný pojem parciálních derivací platí analogie většiny klasických vět o parciálních derivacích. Speciálně platí tyto dvě základní věty (ve kterých užíváme označení z definice 3.28).

3.29 Věta. *Existuje-li $f'(a)$, pak existují všechny parciální derivace $f_i'(a)$, $i = 1, \dots, n$ a platí*

$$f'(a) \centerdot (h_1, \dots, h_n) = f_1'(a)h_1 + \dots + f_n'(a)h_n, \qquad (h_1, \dots, h_n) \in X.$$

3.30 Věta. *Nechť všechny parciální derivace f_i' $(i = 1, \dots, n)$ existují na nějakém okolí U bodu a a všechny parciální derivace (tj. zobrazení $f_i'\colon U \to \mathcal{L}(X_i, Y)$) jsou spojité v bodě a. Pak existuje $f'(a)$.*

3.4 Příklady výpočtu Fréchetovy derivace

3.31 Příklad. (derivace normy v Hilbertově prostoru) Nechť X je reálný Hilbertův prostor ($\dim X \ge 1$). Vyšetřujme fréchetovskou diferencovatelnost normy, tj. funkce $n(x) = \|x\| = \sqrt{\langle x, x\rangle}$. Protože skalární součin $(x, y) \mapsto \langle x, y\rangle$ je spojité bilineární zobrazení (viz. Příklad 1.149), můžeme použít Větu 3.26 (pro $Y_1 = Y_2 = X$, $Z = \mathbb{R}$ a $f_1(x) = f_2(x) = x$) a dostáváme, že funkce $g\colon x \mapsto \langle x, x\rangle$ má v každém bodě $a \in X$ derivaci a platí

$$g'(a)\,h = \langle h, a\rangle + \langle a, h\rangle = 2\langle a, h\rangle, \qquad h \in X.$$

Zobrazení $\omega(y) = \sqrt{y}$ z $\mathbb{R}$ do $\mathbb{R}$ má ovšem v každém bodě $y > 0$ derivaci $\omega'(y)\colon h \mapsto 1/(2\sqrt{y}) \cdot h$, $h \in \mathbb{R}$.

Protože pro $0 \ne x \in X$ platí $g(x) > 0$, podle Věty 3.12 o derivaci složeného zobrazení dostáváme, že funkce $n = \omega \circ g$ má v bodě $x \ne 0$ derivaci a

$$n'(x)h = 1/(2\sqrt{\langle x,x\rangle}) \cdot 2\langle x,h\rangle = \langle x/\|x\|, h\rangle, \qquad h \in X.$$

Podle Poznámky 3.8 můžeme také psát $\operatorname{grad} n(x) = x/\|x\|$.

Derivace $n'(0)$ však neexistuje. Skutečně, pro $0 \neq v \in X$ neexistuje

$$D_v n(0) = \lim_{t\to 0} \frac{\|tv\|}{t} = \lim_{t\to 0} \ \operatorname{sgn} t \cdot \|v\|,$$

takže podle Tvrzení 3.6 neexistuje ani $n'(0)$.

3.32 Příklad. Nechť $\varphi\colon \mathbb{R} \to \mathbb{R}$ je funkce třídy C^2. Uvažujme zobrazení $F\colon C[0,1] \to C[0,1]$, které každé funkci $x \in C[0,1]$ přiřazuje funkci $\varphi \circ x$. Zkoumejme, zda F má Fréchetovu derivaci v bodě $x_0 \in C[0,1]$. Zvolme tedy $h \in C[0,1]$ a zkoumejme diferenci $F(x_0+h)-F(x_0)$. Pro každé $t \in [0,1]$ použijeme pro výpočet hodnoty $F(x_0+h)(t) = \varphi(x_0(t)+h(t))$ Taylorovu větu s Lagrangeovým tvarem zbytku na funkci φ. Dostáváme

$$\varphi(x_0(t) + h(t)) = \varphi(x_0(t)) + \varphi'(x_0(t)) \cdot h(t) + \frac{1}{2}\varphi''(\xi(t)) \cdot h^2(t),$$

kde $\xi(t)$ leží (neostře) mezi $x_0(t)$ a $x_0(t) + h(t)$ (jestliže $h(t) = 0$, rovnost platí pro $\xi(t) = x_0(t)$). Platí tedy

$$(F(x_0 + h) - F(x_0))(t) = (\varphi' \circ x_0)(t) \cdot h(t) + \frac{1}{2}\varphi''(\xi(t)) \cdot h^2(t).$$

Položíme-li $L(h) := (\varphi' \circ x_0)\cdot h$ (tj. $L(h)\colon t \mapsto \varphi'(x_0(t))\cdot h(t)$), je L zřejmě lineární zobrazení $C[0,1]$ do $C[0,1]$. Podle Příkladu 1.138 je L také spojité. Dokážeme, že $F'(x_0) = L$. K tomu potřebujeme dokázat, že

$$\lim_{h\to 0} \frac{F(x_0 + h) - F(x_0) - L(h)}{\|h\|} = 0. \tag{3.8}$$

Protože φ'' je spojitá, můžeme zvolit $K > 0$ takové, že $|\varphi''(\xi)| \leq K$, kdykoliv $|\xi| \leq \|x_0\| + 1$. Je-li nyní $h \in C[0,1]$ s $\|h\| < 1$, pak pro každé $t \in [0,1]$ platí

$$|(F(x_0 + h) - F(x_0) - L(h))(t)| = \left|\frac{1}{2}\varphi''(\xi(t)) \cdot h^2(t)\right| \leq K \cdot h^2(t).$$

Dostáváme tedy

$$\begin{aligned}\|F(x_0 + h) - F(x_0) - L(h)\| &= \sup\,\{|(F(x_0 + h) - F(x_0) - L(h))(t)|\colon\ t \in [0,1]\}\\ &\leq K \sup_{t\in[0,1]} h^2(t) = K\|h\|^2,\end{aligned}$$

takže (3.8) zřejmě platí. V každém bodě $x \in C[0,1]$ tedy existuje $F'(x)$ a platí $F'(x)h = (\varphi' \circ x) \cdot h$. (Poznamenejme, že stejný výsledek platí i za slabšího předpokladu, že φ je třídy C^1.)

3.33 Příklad. Uvažujme funkci $f\colon C[0,1] \to \mathbb{R}$ danou předpisem

$$f(x) = \int_0^1 \sin(x(t))\,\mathrm{d}t.$$

Zřejmě platí $f = g \circ F$, kde

$$F\colon\ C[0,1] \to C[0,1], \quad F(x) = \sin \circ x \qquad \text{a}$$

$$g\colon\ C[0,1] \to \mathbb{R}, \quad g(y) = \int_0^1 y(t)\,\mathrm{d}t.$$

Podle předchozího příkladu (zde $\varphi(x) = \sin x$) pro každou funkci $x \in C[0,1]$ dostáváme $F'(x)h = (\cos \circ x) \cdot h$. Protože g je podle Příkladu 1.139 spojitý lineární funkcionál na $C[0,1]$, podle Důsledku 3.13 platí $f'(x) = g \circ F'(x)$, a tedy

$$f'(x) \mathbin{.} h = \int_0^1 \cos x(t) \cdot h(t)\,\mathrm{d}t, \qquad h \in C[0,1].$$

3.5 Derivace vyšších řádů

V tomto oddíle stále předpokládáme, že X, Y jsou normované lineární prostory a f je zobrazení z X do Y.

První derivací f' rozumíme zobrazení z X do normovaného prostoru $\mathcal{L}(X,Y)$, které je definováno předpisem

$$f'\colon\ x \mapsto f'(x), \quad x \in D_{f'},$$

kde $D_{f'}$ je množina všech bodů $x \in X$, ve kterých existuje Fréchetova derivace $f'(x)$. Protože f' je zobrazení z normovaného prostoru do normovaného prostoru, má smysl hovořit o jeho derivaci a můžeme vyslovit následující definici.

3.34 Definice. *Řekneme, že zobrazení f má v bodě $a \in X$ druhou derivaci $f''(a)$, má-li zobrazení f' v bodě a Fréchetovu derivaci; pak klademe*

$$f''(a) := (f')'(a).$$

Při této přirozené definici je $f''(a) \in \mathcal{L}(X, \mathcal{L}(X,Y))$; je to spojité lineární zobrazení prostoru X do prostoru všech spojitých lineárních zobrazení prostoru X do prostoru Y. To však není ve shodě s definicí z Kapitoly 2, podle které v případě $X = \mathbb{R}^n$ a $Y = \mathbb{R}$ je $f''(a)$ bilineární forma na prostoru $X = \mathbb{R}^n$.

Mezi prostorem $\mathcal{L}(X, \mathcal{L}(X,Y))$ a prostorem $\mathcal{L}_2(X,Y)$ všech spojitých bilineárních zobrazení prostoru $X \times X$ do Y však existuje kanonická lineární izometrie (viz Tvrzení 1.147) a je běžné „ztotožňovat“ prvek prostoru $\mathcal{L}(X, \mathcal{L}(X,Y))$ s jeho obrazem (který je prvkem $\mathcal{L}_2(X,Y)$) při této izometrii. Jak za chvíli ukážeme v Tvrzení 3.35, při této běžné úmluvě je Definice 3.34 zobecněním definice z Kapitoly 2.

Druhou derivaci $f''(a)$ tedy ztotožňujeme se spojitým bilineárním zobrazením prostoru $X \times X$ do Y definovaným předpisem

$$(h,k) \mapsto (f''(a) \,.\, h) \,.\, k, \qquad h,k \in X.$$

Při tomto ztotožnění lze tedy psát

$$f''(a) \,.\, (h,k) = (f''(a) \,.\, h) \,.\, k, \qquad h,k \in X. \tag{3.9}$$

3.35 Tvrzení. *Nechť f je zobrazení z $\mathbb{R}^n$ do $\mathbb{R}$, $a \in \mathbb{R}^n$ a B je bilineární forma na $\mathbb{R}^n$. Pak rovnost $B = f''(a)$ má ve smyslu staré definice (viz str. 89) tentýž význam, jako ve smyslu Definice 3.34 (a jí následující úmluvy o ztotožnění).*

Důkaz. Nejdříve připomeňme (srov. Poznámka 3.4 (i)), že $f'(x)$ podle Definice 3.3 je přesně totéž, co $\mathrm{d}f(x)$ podle klasické definice. Je tedy $f'(x) \,.\, h$ totéž (co do existence i hodnoty), co $d_h f(x)$ podle klasické definice. Předpokládejme nejprve, že $f''(a)$ existuje ve smyslu Definice 3.34 (a dále má přesně tento význam; tj. zatím „neztotožňujeme“). Zvolme libovolně $h,k \in \mathbb{R}^n$. Pak zřejmě $d_k f(x) = f'(x) \,.\, k$ existuje pro všechna x z nějakého okolí bodu a. Nechť zobrazení φ_k: $\mathcal{L}(\mathbb{R}^n,\mathbb{R}) \to \mathbb{R}$ je dáno předpisem $\varphi_k(L) = L(k)$. Podle Příkladu 1.150 je φ_k spojité lineární zobrazení, takže podle Důsledku 3.13 existuje

$$(d_k f)'(a) = (f'(\cdot) \,.\, k)'(a) = (\varphi_k \circ f')'(a) = \varphi_k \circ f''(a).$$

Existuje tedy

$$d_h(d_k f)(a) = (\varphi_k \circ f''(a)) \,.\, h = \varphi_k \,.\, (f''(a) \,.\, h) = (f''(a) \,.\, h) \,.\, k.$$

Nechť nyní $f''(a)$ existuje ve smyslu klasické definice, tj. $d_h(d_k f)(a)$ existuje pro všechny $h,k \in \mathbb{R}^n$. Je snadno vidět, že $\psi_i = \varphi_{e_i}$, $i = 1,\ldots,n$, tvoří bázi prostoru $(\mathcal{L}(\mathbb{R}^n,\mathbb{R}))^*$, takže podle Tvrzení 3.18 $f''(a) = (f')'(a)$ existuje, právě když existují všechny derivace $(\psi_i \circ f')'(a)$, $i = 1,\ldots,n$. To je však podle předpokladu pravda, protože $\psi_i \circ f' = f'(\cdot) \,.\, e_i = d_{e_i} f$.

Dále klademe

$$f'''(a) := (f'')'(a) \in \mathcal{L}(X, \mathcal{L}(X, \mathcal{L}(X,Y)))$$

a n-tou derivaci definujeme indukcí pomocí rekurentního vzorce

$$f^{(n)}(a) := (f^{(n-1)})'(a).$$

Třetí derivaci $f'''(a)$ ztotožňujeme přirozeným způsobem (viz 1.147) s prvkem prostoru $\mathcal{L}(X, \mathcal{L}_2(X,Y))$ a ten s prvkem prostoru $\mathcal{L}_3(X,Y)$; tj. s multilineárním zobrazením prostoru $X \times X \times X$ do prostoru Y. Po tomto ztotožnění můžeme psát

$$f'''(a) \,.\, (v_1,v_2,v_3) = ((f'''(a) \,.\, v_1) \,.\, v_2) \,.\, v_3.$$

Podobně n-tou derivaci $f^{(n)}$ chápeme jako prvek prostoru $\mathcal{L}_n(X,Y)$: n-lineární zobrazení prostoru X^n do Y dané předpisem

$$f^{(n)}(a)(v_1,\ldots,v_n) := (\ldots((f^{(n)}(a) \,.\, v_1) \,.\, v_2) \,.\, \ldots) \,.\, v_n.$$

Je jen formálně obtížné indukcí zobecnit Tvrzení 3.35; dokázat, že pro případ $X = \mathbb{R}^n$ a $Y = \mathbb{R}$ je právě uvedená definice n-té derivace ve shodě s klasickou definicí.

3.36 Definice. *Nechť X, Y jsou normované lineární prostory, $G \subset X$ je otevřená množina a $p \in \mathbb{N}$. Řekneme, že zobrazení $f\colon G \to Y$ je třídy C^p (náleží do $C^p(G,Y)$), jestliže p-tá derivace $f^{(p)}\colon G \to \mathcal{L}_p(X,Y)$ je spojité zobrazení.*

Následující zobecnění Věty 3.17 se snadno dokáže indukcí.

3.37 Věta. *Nechť $p \in \mathbb{N}$, $f = (f_1, \dots, f_n)$ je zobrazení z normovaného lineárního prostoru Y do součinu $X := X_1 \times \dots \times X_n$ normovaných lineárních prostorů a nechť $a \in Y$. Pak*

$$f^{(p)}(a) = (f_1^{(p)}(a), \dots, f_n^{(p)}(a)),$$

má-li jedna strana rovnosti smysl.

Pokud víme, že existuje $f^{(n)}(a)$, můžeme pro výpočet $f^{(n)}(a)(v_1, \dots, v_n)$ s výhodou použít následující tvrzení (ve kterém operátory d_v a D_v používáme ve zcela analogickém významu jako na str. 88).

3.38 Tvrzení. *Nechť existuje $f^{(n)}(a)$ a $v_1, \dots, v_n \in X$. Pak platí rovnosti*

$$f^{(n)}(a)(v_1, \dots, v_n) = \Big(f^{(n-1)}(\cdot)(v_2, \dots, v_n)\Big)'(a) \,\boldsymbol{.}\, v_1, \tag{3.10}$$

$$f^{(n)}(a)(v_1, \dots, v_n) = (d_{v_1} \circ \dots \circ d_{v_n} f)(a) = (D_{v_1} \circ \dots \circ D_{v_n} f)(a). \tag{3.11}$$

Důkaz. (Náznak.) Uvažujme spojitý lineární funkcionál $\Psi \in (\mathcal{L}_{n-1}(X,Y))^*$ daný předpisem $\Psi(\omega) := \omega(v_2, \dots, v_n)$. Podle definice n-té derivace máme

$$f^{(n)}(a)(v_1, \dots, v_n) = \Big((f^{(n-1)})'(a) \,\boldsymbol{.}\, v_1\Big)(v_2, \dots, v_n) = \Psi\Big((f^{(n-1)})'(a) \,\boldsymbol{.}\, v_1\Big).$$

Podle Důsledku 3.13 platí

$$\Big(f^{(n-1)}(\cdot)(v_2, \dots, v_n)\Big)'(a) \,\boldsymbol{.}\, v_1 = \Big(\Psi \circ f^{(n-1)}\Big)'(a) \,\boldsymbol{.}\, v_1 = \Big(\Psi \circ (f^{(n-1)})'(a)\Big) \,\boldsymbol{.}\, v_1.$$

Protože zřejmě $\Psi\big((f^{(n-1)})'(a) \,\boldsymbol{.}\, v_1\big) = \big(\Psi \circ (f^{(n-1)})'(a)\big) \,\boldsymbol{.}\, v_1$, dostáváme (3.10).

Vzorec (3.11) je zřejmý pro $n = 1$. Předpokládáme-li jeho (obecnou) platnost pro $n - 1$, vidíme, že pro x z nějakého okolí bodu a platí $f^{(n-1)}(x)(v_2, \dots, v_n) = (d_{v_2} \circ \dots \circ d_{v_n} f)(x)$, takže z rovnosti (3.10) vyplývá (3.11).

3.39 Poznámka. Z (3.11) okamžitě vyplývá, že pokud funkce z $\mathbb{R}^n$ do $\mathbb{R}$ má n-tou derivaci $f^{(n)}(a)$ podle moderní definice, je to n-tá derivace i podle klasické definice.

3.40 Tvrzení. (symetrie n-té derivace) *Existuje-li $f^{(n)}(a)$, je tato n-tá derivace symetrické n-lineární zobrazení; hodnota $f^{(n)}(a)(v_1, \dots, v_n)$ se nezmění, permutujeme-li libovolně vektory $v_1, \dots, v_n$.*

Důkaz. (Náznak možného postupu v případě reálných prostorů.) Symetrii n-té derivace pro případ $Y = \mathbb{R}$ lze odvodit z její symetrie v klasickém případě. Zvolme libovolně vektory $v_1, \dots, v_n$ a pro $t = (t_1, \dots, t_n) \in \mathbb{R}^n$ položme

$$\varphi(t_1, \dots, t_n) := t_1 v_1 + \dots + t_n v_n, \quad g(t_1, \dots, t_n) := f(a + \varphi(t)).$$

Není obtížné dokázat, že existuje $g^{(n)}(0)$ a pro $u_1, \dots, u_n \in \mathbb{R}^n$ platí

$$g^{(n)}(0)(u_1, \dots, u_n) = f^{(n)}(a)(\varphi(u_1), \dots, \varphi(u_n)).$$

(Pro případy $n = 2, 3$ to okamžitě vyplývá ze vzorců pro druhou a třetí derivaci složeného zobrazení, které jsou odvozeny v Příkladu 3.44 níže. Tam použitý postup také dává návod, jak důkaz indukcí provést.) Z Poznámky 3.39 a Poznámky 2.87 (i) vyplývá symetrie $g^{(n)}(0)$ a odtud i symetrie $f^{(n)}(a)$.

Pro případ obecného reálného Y je možno nejprve indukcí dokázat, že pro každé $\varphi \in Y^*$ platí

$$(\varphi \circ f)^{(n)}(a) = \varphi \circ f^{(n)}(a)$$

a ze symetrie $\varphi \circ f^{(n)}(a)$ pro všechna $\varphi \in Y^*$ pak pomocí Věty 1.136 snadno dokázat symetrii $f^{(n)}(a)$.

(Pro přímý důkaz, který je založen na stejné myšlence jako důkaz Tvrzení 2.82, ale je technicky náročnější, viz [Cart] nebo [FM].)

V dalším budeme potřebovat následující snadné důležité tvrzení.

3.41 Tvrzení. *Nechť x, Y, Z jsou normované lineární prostory a $B \in \mathcal{L}(X, Y; Z)$ je bilineární zobrazení. Pak B: $X \times Y \to Z$ je třídy C^∞. Navíc platí, že druhá derivace B'' je konstantní zobrazení a derivace $B^{(n)}$ pro $n > 2$ jsou nulové.*

Důkaz. Podle Věty 3.25 víme, že B je diferencovatelné na $X \times Y$ a pro $(x, y) \in X \times Y$ platí

$$B'(x, y) \,.\, (h, k) = B(h, y) + B(x, k), \qquad (h, k) \in X \times Y.$$

Snadno tedy vidíme, že zobrazení

$$B' \colon\ X \times Y \to \mathcal{L}(X \times Y, Z)$$

je lineární. Protože

$$\begin{aligned} \|B'(x, y)\| &= \sup\{\|B(h, y) + B(x, k)\| \colon\ \|(h, k)\| \le 1\} \\ &\le \|B\| \cdot \|y\| + \|B\| \cdot \|x\| \le 2\|B\|\|(x, y)\|, \end{aligned}$$

je zobrazení B' také spojité, takže $B' \in \mathcal{L}(X \times Y, \mathcal{L}(X \times Y, Z))$. Je tedy (srov. Poznámka 3.4 (iii)) $B'' = (B')'$ konstantní zobrazení, které každému $(x, y) \in X \times Y$ přiřazuje B', tj. prvek $\mathcal{L}(X \times Y, \mathcal{L}(X \times Y, Z))$ daný předpisem

$$(h_1, k_1) \mapsto ((h_2, k_2) \mapsto\ B(h_2, k_1) + B(h_1, k_2))\,,$$

které ztotožňujeme s bilineárním zobrazením (prvkem $\mathcal{L}_2(X \times Y, Z)$) daným přespisem

$$((h_1, k_1), (h_2, k_2)) \ \mapsto\ B(h_2, k_1) + B(h_1, k_2).$$

Nyní již stačí použít Poznámku 3.4 (iv).

3.6 Derivace vyššího řádu složeného zobrazení

Připomeňme, že v klasickém případě jsme pro jednoduchost dokazovali věty o n-té derivaci pouze pro zobrazení třídy C^p, takže následující výsledek je nový i v klasickém případě.

3.42 Věta. *Nechť $n \in \mathbb{N}$, X, Y, Z jsou normované lineární prostory, f je zobrazení z X do Y a g je zobrazení z Y do Z. Nechť $a \in X$ a nechť existují n-té derivace $f^{(n)}(a)$ a $g^{(n)}(f(a))$. Pak složené zobrazení $h := g \circ f$ má n-tou derivaci $h^{(n)}(a)$ v bodě a.*

Důkaz. Víme, že věta platí pro $n = 1$. Předpokládejme tedy, že $n \geq 2$ a že věta platí pro $n - 1$. Zřejmě $f'(x)$ existuje pro všechna x z nějakého okolí bodu a a $g'(y)$ existuje pro všechna y z nějakého okolí bodu $f(a)$. Existuje tedy $(g \circ f)'(x) = g'(f(x)) \circ f'(x)$ pro všechna x z nějakého okolí U bodu a. Jinými slovy, na U platí $h' = B \circ \Psi$, kde

$$\Psi : U \to \mathcal{L}(X,Y) \times \mathcal{L}(Y,Z), \quad \Psi(x) = (f'(x), g'(f(x)) \text{ a}$$
$$B : \mathcal{L}(X,Y) \times \mathcal{L}(Y,Z) \to \mathcal{L}(X,Z), \quad B(p,q) = q \circ p.$$

Abychom dokázali, že vnitřní zobrazení Ψ má $(n-1)$-tou derivaci v bodě a, stačí dokázat (Věta 3.37), že jeho složky mají $(n-1)$-tou derivaci v a. To je zřejmě pravda pro první složku a pro druhou to plyne z indukčního předpokladu, protože existují $(g')^{(n-1)}(f(a))$ a $f^{(n-1)}(a)$. Vnější zobrazení B je spojité bilineární zobrazení, takže je podle Tvrzení 3.41 dokonce třídy C^∞. Podle indukčního předpokladu tedy existuje $h^{(n)}(a) = (h')^{(n-1)}(a)$.

Zcela analogicky lze dokázat zobecnění Věty 2.76 o skládání zobrazení třídy C^p.

3.43 Věta. *Nechť $p \in \mathbb{N} \cup \{\infty\}$, X, Y, Z jsou normované lineární prostory, $G \subset X$, $H \subset Y$ jsou otevřené množiny, $f\colon G \to Y$, $g\colon H \to Z$ jsou třídy C^p a platí $f(G) \subset H$. Pak zobrazení $h := g \circ f$ je třídy C^p.*

3.44 Příklad. Za předpokladů Věty 3.42 odvodíme vzorec pro $(g \circ f)^{(n)}(a)$ v případech $n = 2$ a $n = 3$. V případě $n = 2$ s pomocí Tvrzení 3.38 a vzorce (3.6) použitého na spojité bilineární zobrazení (srov. Příklad 1.150) $B(p,q) = p \centerdot q$, $p \in$

$\mathcal{L}(Y,Z)$, $q \in Y$ dostáváme

$$\begin{aligned}(g\circ f)''(a)\,.\,(h,k) &= ((g\circ f)'(\cdot)\,.\,k)'(a)\,.\,h = (g'(f(\cdot))\,.\,(f'(\cdot)\,.\,k))'(a)\,.\,h = \\ &= (g''(f(a))\,.\,(f'(a)\,.\,h))\,.\,(f'(a)\,.\,k) + g'(f(a))\,.\,((f''(a)\,.\,h)\,.\,k) = \\ &= g''(f(a))\,.\,(f'(a)\,.\,h, f'(a)\,.\,k) + g'(f(a))\,.\,(f''(a)\,.\,(h,k))\,.\end{aligned}$$

Na třech řádcích jsme odvodili vzorec, který jsme klasicky vypočetli v důkazu Tvrzení 2.96 ve velmi speciálním případě. Cenou za zkrácení výpočtu je ovšem myšlenkově náročná práce v abstraktních prostorech.

Pro zajímavost ještě ukážeme výpočet pro $n=3$. Podle Tvrzení 3.38 a výsledku pro $n=2$ máme

$$\begin{aligned}(g\circ f)'''(a)\,.\,(l,h,k) &= \big((g\circ f)''(\cdot)\,.\,(h,k)\big)'(a)\,.\,l \\ &= \big(g''(f(\cdot))\,.\,(f'(\cdot)\,.\,h, f'(\cdot)\,.\,k) + g'(f(\cdot))\,.\,((f''(\cdot)\,.\,(h,k))\big)'(a)\,.\,l.\end{aligned}$$

Nyní postupujeme jako pro $n=2$ a navíc při derivování prvního sčítance užijeme vzorec (3.7) pro spojité 3-lineární zobrazení $M : \mathcal{L}_2(Y,Z)\times Y\times Y \to Z$, $M(B,y_1,y_y) = B(y_1,y_2)$. Dostáváme

$$\begin{aligned}(g\circ f)'''(a)(l,h,k) &= \big(g'''(f(a))\,.\,(f'(a)\,.\,l)\big)\,.\,(f'(a)\,.\,h, f'(a)\,.\,k) + \\ &+ g''(f(a)\,.\,((f''(a)\,.\,l)\,.\,h, f'(a)\,.\,k) + g''(f(a))\,.\,(f'(a)\,.\,h,(f''(a)\,.\,l)\,.\,k) + \\ &+ \big(g''(f(a))\,.\,(f'(a)\,.\,l)\big)\,.\,(f''(a)\,.\,(h,k)) + g'(f(a))\,.\,\big((f'''(a)\,.\,l)\,.\,(h,k)\big) = \\ &= g'''(f(a))\,.\,(f'(a)\,.\,l, f'(a)\,.\,h, f'(a)\,.\,k) + \\ &+ g''(f(a))\,.\,(f''(a)\,.\,(l,h), f'(a)\,.\,k)) + g''(f(a)\,.\,(f'(a)\,.\,h, f''(a)\,.\,(l,k)) + \\ &\quad g''(f(a))\,.\,(f'(a)\,.\,l, f''(a)\,.\,(h,k)) + g'(f(a))\,.\,\big(f'''(a)\,.\,(l,h,k)\big)\,.\end{aligned}$$

Nakonec ještě vysvětlíme, jak se získá vzorec pro $(g\circ f)^{(n)}(a)(v_1,\dots,v_n)$. Tuto hodnotu lze vypočíst jako součet hodnot, z nichž každá odpovídá právě jednomu rozkladu R množiny $\{1,\dots,n\}$ na neprázdné podmnožiny následujícím způsobem:

Jde-li o rozklad na s množin, uspořádáme si tyto množiny do konečné posloupnosti $M_1,\dots,M_s$ a prvky každé z množin M_i ($1\le i\le s$) si uspořádáme (třeba podle velikosti) do konečné posloupnosti $(k_1^i, k_2^i, \dots, k_{p_i}^i)$. Rozkladu R pak přiřadíme hodnotu

$$g^{(s)}(f(a))\,.\,\Big(f^{(p_1)}(a)\,.\,(v_{k_1^1},\dots,v_{k_{p_1}^1}),\dots,f^{(p_s)}(a)\,.\,(v_{k_1^s},\dots,v_{k_{p_s}^s})\Big)\,.$$

Vzhledem k symetrii derivací vyššího řádu vůbec nezáleží na tom, jak jsme uspořádávali. Toto tvrzení lze dokázat indukcí stejným postupem, jakým jsme ze vzorce pro $n=2$ odvodili vzorec pro $n=3$; problémy dělá pouze nepřehlednost formálních zápisů.

3.45 Příklad. Pro $n=3$ existuje 5 rozkladů množiny $1,2,3$:

$$\{\{1,2,3\}\},\ \{\{1,2\},\{3\}\},\ \{\{1,3\},\{2\}\},\ \{\{2,3\},\{1\}\},\ \{\{1\},\{2\},\{3\}\}.$$

Těmto pěti rozkladům zřejmě odpovídá pět sčítanců ve vzorci, který jsme odvodili výše.

Poznamenejme ještě, že příslušný vzorec pro obecné n lze také odvodit pomocí Taylorových rozvojů (srov. [Cart] a [Fe]).

3.7 Vlastnosti zobrazení L 7! L^{-1}, L 2 Izom(X,Y)

V tomto oddíle budeme zkoumat zobrazení z, které každému izomorfismu $L \in \mathrm{Izom}(X,Y)$ přiřazuje jeho inverzi $L^{-1} \in \mathrm{Izom}(Y,X)$. Výsledky přitom platí jen pro úplné (tj. Banachovy) prostory X, Y.

Ztotožníme-li pro $X = Y = \mathbb{R}^n$ množinu $\mathrm{Izom}(\mathbb{R}^n, \mathbb{R}^n)$ kanonicky s množinou všech regulárních matic typu $n \times n$, jde o zobrazení $z\colon\ A \mapsto A^{-1}$, které matici přiřazuje matici inverzní. Toto zobrazení umíme explicitně spočítat a na základě příslušného vzorce není těžké usoudit, že z je třídy C^∞. Pro případ $X = Y = \mathbb{R}$ má zobrazení z tvar (matice 1×1 ztotožňujeme s reálnými čísly) $z\colon\ x \mapsto 1/x, \quad x \neq 0$. Lze říci, že zobrazení z hraje v „abstraktním diferenciálním počtu“ podobnou roli, jako funkce $1/x$ v klasickém diferenciálním počtu. Jeho vlastnosti podstatně užijeme v důkazu Věty 3.50 o inverzním zobrazení.

Připomeňme ještě, že symbolem I_X označujeme identické zobrazení na množině X.

3.46 Lemma. *Nechť X je Banachův prostor , $u \in \mathcal{L}(X,X)$ a $\|u\| < 1$. Pak*

$$v := I_X - u \in \mathrm{Izom}(X,X) \quad \text{a} \quad \|I_X - v^{-1}\| \leq \frac{\|u\|}{1 - \|u\|}.$$

Důkaz. Pro $x \in X$ platí

$$\|v(x)\| = \|x - u(x)\| \geq \|x\| - \|u(x)\| \geq \|x\| - \|u\| \cdot \|x\| = (1 - \|u\|)\|x\|. \tag{3.12}$$

Vidíme, že $\mathrm{Ker}\, v = \{x \in X\colon\ v(x) = 0\} = \{0\}$. Z lineární algebry tedy vyplývá, že v je prosté zobrazení a v^{-1} je lineární zobrazení definované na vektorovém prostoru $\mathrm{Im}\, v = v(X)$. K důkazu, že $\mathrm{Im}\, v = X$, je třeba dokázat, že rovnice $v(x) = y$, která je ekvivalentní s rovnicí $y + u(x) = x$, má řešení pro každý bod $y \in X$. To však plyne okamžitě z Banachovy věty o kontrakci (Věty 1.87), protože zobrazení

$$g\colon\ X \to X, \quad g(x) = y + u(x)$$

je zřejmě kontrakce, a má tedy pevný bod.

Uvažujme libovolný bod $y \in X$ a položme $x := v^{-1}(y)$. Pak $y = v(x)$, takže podle (3.12) platí

$$\|y\| \geq (1 - \|u\|)\|x\|, \quad \text{tj.} \quad \|v^{-1}(y)\| \leq (1 - \|u\|)^{-1}\, \|y\|. \tag{3.13}$$

Z této nerovnosti vyplývá, že $v \in \mathrm{Izom}(X,X)$. Dále platí

$$(I_X - v^{-1})(y) = y - x = v(x) - x = -u(x),$$

takže s pomocí (3.13) dostáváme

$$\|(I_X - v^{-1})(y)\| = \|u(x)\| \leq \|u\| \cdot \|x\| \leq \frac{\|u\|}{1 - \|u\|}\, \|y\|,$$

z čehož vyplývá hledaný odhad pro $\|I_X - v^{-1}\|$.

3.47 Poznámka. Běžně se při důkazu předchozího lemmatu (a následujícího tvrzení) používá přirozenější postup, který však vyžaduje řadu předběžných úvah. Vychází se přitom z explicitního vzorce

$$v^{-1} = I_X + u + u^2 + u^3 + \dots$$

(ve kterém u^n je ovšem n-násobné složení $u \circ \dots \circ u$), který je zobecněním vzorce pro součet geometrické řady čísel.

3.48 Tvrzení. *Nechť X, Y jsou izomorfní Banachovy prostory. Pak platí:*

(i) $\operatorname{Izom}(X,Y)$ *je otevřená podmnožina prostoru* $\mathcal{L}(X,Y)$.

(ii) *Zobrazení* $z\colon \operatorname{Izom}(X,Y) \to \mathcal{L}(Y,X)$, $z\colon u \mapsto u^{-1}$ *je spojité.*

(iii) *Zobrazení z má derivaci* $z'(u) \in \mathcal{L}\left(\mathcal{L}(X,Y), \mathcal{L}(Y,X)\right)$ *v každém bodě* $u \in \operatorname{Izom}(X,Y)$ *a platí*

$$z'(u)\colon\ h \mapsto -u^{-1} \circ h \circ u^{-1}, \quad h \in \mathcal{L}(X,Y).$$

(iv) *Zobrazení z je třídy C^∞ na* $\operatorname{Izom}(X,Y)$.

Důkaz. Uvažujme libovolný prvek $u \in \operatorname{Izom}(X,Y)$ a zvolme $h \in \mathcal{L}(X,Y)$ s normou $\|h\| < 1/\|u^{-1}\|$. Protože $\|u^{-1} \circ h\| \le \|u^{-1}\|\,\|h\| < 1$, podle Lemmatu 3.46 dostáváme

$$u^{-1} \circ (u+h) = I_X + u^{-1} \circ h \in \operatorname{Izom}(X,X).$$

Z toho snadno vyplývá, že $u + h = u \circ (I_X + u^{-1} \circ h)$ je prvek $\operatorname{Izom}(X,Y)$, takže jsme dokázali (i). Dále platí

$$(u+h)^{-1} = (I_X + u^{-1} \circ h)^{-1} \circ u^{-1}, \quad (u+h)^{-1} - u^{-1} = \left((I_X + u^{-1} \circ h)^{-1} - I_X\right) \circ u^{-1}.$$

S pomocí Lemmatu 3.46 tedy dostáváme

$$\|(u+h)^{-1} - u^{-1}\| \le \frac{\|-u^{-1} \circ h\|}{1 - \|-u^{-1} \circ h\|}\|u^{-1}\| \le \frac{\|u^{-1}\|^2\,\|h\|}{1 - \|u^{-1}\|\|h\|}.$$

Zobrazení z je tedy spojité v bodě u; tím je dokázáno tvrzení (ii).

Abychom dokázali (iii), vyšetřujme pro každý prvek $u \in \operatorname{Izom}(X,Y)$ diferenci $z(u+h) - z(u)$, $h \in \mathcal{L}(X,Y)$, která je podle předchozího definovaná, je-li $\|h\|$ dostatečně malé. Zřejmě platí

$$\begin{aligned} z(u+h) - z(u) &= (u+h)^{-1} - u^{-1} = \\ &= (u+h)^{-1} \circ u \circ u^{-1} - (u+h)^{-1} \circ (u+h) \circ u^{-1} = \\ &= -(u+h)^{-1} \circ h \circ u^{-1}. \end{aligned}$$

Zobrazení $L\colon h \mapsto -u^{-1} \circ h \circ u^{-1}$, $h \in \mathcal{L}(X,Y)$, je zřejmě lineární; chceme dokázat že pro zobrazení $r(h) := z(u+h) - z(u) - L(h)$ platí $\|r(h)\| = o(\|h\|), h \to 0$. To vyplývá z nerovnosti

$$\begin{aligned} \|r(h)\| &= \|(u+h)^{-1} \circ h \circ u^{-1} - u^{-1} \circ h \circ u^{-1}\| = \\ &\quad \|\left((u+h)^{-1} - u^{-1}\right) \circ h \circ u^{-1}\| \le \|(u+h)^{-1} - u^{-1}\|\,\|h\|\,\|u^{-1}\| \end{aligned}$$

a toho, že podle (ii) platí $\lim_{h\to 0} \|(u+h)^{-1} - u^{-1}\| = 0$. Spojitost lineárního zobrazení L vyplývá z nerovnosti $\|L(h)\| \le \|u^{-1}\|\,\|h\|\,\|u^{-1}\|$ (nebo ze spojitosti zobrazení z, viz Poznámka 3.4 (ii)).

Nakonec stručně dokážeme (iv). Pro $v \in \mathcal{L}(Y,X)$ a $w \in \mathcal{L}(Y,X)$ je předpisem

$$B(v,w)(h) := -v \circ h \circ w, \qquad h \in \mathcal{L}(X,Y)$$

zřejmě definováno zobrazení $B(v,w) \in \mathcal{L}\left(\mathcal{L}(X,Y), \mathcal{L}(Y,X)\right)$. Snadno se ověří, že zobrazení

$$B := (v,w) \mapsto B(v,w), \qquad B\colon\ \mathcal{L}(Y,X) \times \mathcal{L}(Y,X) \to \mathcal{L}\left(\mathcal{L}(X,Y), \mathcal{L}(Y,X)\right)$$

je spojité lineární zobrazení. Podle (iii) platí

$$z'(u) = B(z(u), z(u)), \qquad u \in \operatorname{Izom}(X,Y),$$

tj. $z' = B \circ \Psi$, kde

$$\Psi\colon\ u \mapsto (z(u), z(u)), \qquad \Psi\colon\ \operatorname{Izom}(X,Y) \to \mathcal{L}(Y,X) \times \mathcal{L}(Y,X).$$

Podle (ii) je Ψ spojité zobrazení, takže také $z' = B \circ \Psi$ je spojité zobrazení, takže z je třídy C^1 na $\operatorname{Izom}(X,Y)$.

Předpokládejme nyní, že z je třídy C^p ($p \in \mathbb{N}$). Pak z Věty 3.37 snadno vyplývá, že Ψ je třídy C^p, takže podle Tvrzení 3.41 a Věty 3.43 je $z' = B \circ \Psi$ třídy C^p, takže z je třídy C^{p+1}. Indukcí jsme dokázali (iv).

3.8 Věta o inverzním zobrazení a věta o implicitních funkcích

Velmi důležité věty o implicitních funkcích a o inverzním zobrazení lze zcela přirozeně zobecnit na případ zobrazení mezi Banachovými prostory. V Kapitole 2 jsme při důkazu těchto vět použili „konečně dimenzionální“ metodu (řešení soustavy rovnic dosazovací metodou). V Banachových prostorech se nejčastěji důkaz vede tak, že se rovnice řeší vhodným převedením na úlohu o hledání pevného bodu a užitím Banachovy věty o kontrakci. Tento postup vyznívá poněkud trikově; je vhodné proto poznamenat, že užití věty o kontrakci je pouze abstraktní verzí metody postupných aproximací, která je v tomto případě naprosto přirozená (jde v podstatě o řešení rovnice modifikovanou Newtonovou metodou tečen).

Budeme postupovat tak, že nejprve dokážeme větu o inverzním zobrazení. Pak pouze naznačíme, jak z ní lze odvodit větu o implicitním zobrazení. Příslušný *úplný důkaz* v konečně rozměrném případě je proveden na str. 115.

Nejdříve dokážeme následující verzi věty o derivaci inverzního zobrazení.

3.49 Tvrzení. *Nechť X, Y jsou Banachovy prostory, $G \subset X$, $H \subset Y$ jsou otevřené množiny a je dán bod $a \in G$. Nechť $f\colon\ G \to H$ je homeomorfní zobrazení a $f'(a)$*

je izomorfismus X na Y. Pak zobrazení f^{-1} má derivaci v bodě $b := f(a)$ a $(f^{-1})'(b) = (f'(a))^{-1}$.

Důkaz. (i) Nejprve provedeme důkaz v případě, že $Y = X, a = f(a) = b = 0$ a $f'(a) = I_X$. Pak ukážeme, jak se obecný případ na tento speciální případ snadno redukuje. Z definice derivace dokážeme, že $(f^{-1})'(0) = I_X$. Protože $f^{-1}(0) = 0$, stačí pro libovolné $1 > \varepsilon > 0$ najít okolí V bodu 0 takové, že

$$\|f^{-1}(y) - y\| < \varepsilon\|y\|, \quad \text{kdykoliv} \quad y \in V \setminus \{0\}. \tag{3.14}$$

Protože $f'(0) = I_X$ a $f(0) = 0$, lze najít otevřené okolí $U \subset G$ bodu 0 tak, že

$$\|f(x) - x\| < \frac{\varepsilon}{2}\|x\|, \quad \text{kdykoliv} \quad x \in U \setminus \{0\}. \tag{3.15}$$

Ukážeme, že stačí položit $V := f(U)$. Protože f je homeomorfismus a množina H je otevřená, je zřejmě otevřená i množina V. Zvolme libovolný bod $y \in V \setminus \{0\}$ a položme $x := f^{-1}(y)$. Pak zřejmě $x \in U \setminus \{0\}$ a $f(x) = y$, takže podle (3.15) máme

$$\|y - x\| < \frac{\varepsilon}{2}\|x\|, \qquad \|y\| \geq \|x\| - \|y - x\| > \|x\| - \frac{\varepsilon}{2}\|x\| > \frac{1}{2}\|x\|.$$

Z těchto dvou nerovností dostáváme

$$\|f^{-1}(y) - y\| = \|y - x\| < \frac{\varepsilon}{2}\|x\| < \varepsilon\|y\|,$$

takže jsme dokázali (3.14).

(ii) V obecném případě položíme

$$\widetilde{f}(x) := (f'(a))^{-1}\,(f(a + x) - f(a)) \qquad \text{a} \qquad \widetilde{G} := G - a.$$

Pak zobrazení $\widetilde{f}$ homeomorfně zobrazuje otevřenou množinu $\widetilde{G} \subset X$ na otevřenou množinu $\widetilde{H} := (f'(a))^{-1}(H - b) \subset X$, přičemž platí $0 \in \widetilde{G}$, $0 \in \widetilde{H}$, $\widetilde{f}(0) = 0$ a $(\widetilde{f})'(0) = I_X$. Podle první části důkazu tedy dostáváme $\left((\widetilde{f})^{-1}\right)'(0) = I_X$. Uvážíme-li, že $f(z) = f'(a)(\widetilde{f}(z - a)) + f(a)$, $z \in G$, a tedy

$$f^{-1}(w) = \widetilde{f}^{-1}\left((f'(a))^{-1}(w - b)\right) + a, \quad w \in H,$$

z Věty 3.12 o derivaci složeného zobrazení dostáváme $(f^{-1})'(b) = (f'(a))^{-1}$.

3.50 Věta. (věta o inverzním zobrazení v Banachových prostorech) *Nechť X, Y jsou Banachovy prostory, $p \in \mathbb{N} \cup \{\infty\}$ a V je otevřené okolí bodu $a \in X$. Nechť je dáno zobrazení $f\colon V \to Y$ třídy C^p takové, že $f'(a)$ je izomorfismus prostoru X na prostor Y. Pak existuje otevřené okolí U bodu a takové, že $f \restriction_U$ je difeomorfismus a zobrazení $(f \restriction_U)^{-1}$ je třídy C^p.*

Důkaz. Nejprve provedeme důkaz v případě, že $Y = X$, $a = f(a) = 0$ a platí $f'(a) = I_X$. Pak ukážeme, jak se obecný případ na tento speciální případ snadno redukuje.

Položme $g := I_X - f$; zřejmě g je třídy C^1 na V, $g(0) = 0$ a $g'(0) = 0$. Ze spojitosti f' a g' v bodě 0 a Tvrzení 3.48 (i) vyplývá existence $\delta > 0$ takového, že

$$\|g'(x)\| \le \frac{1}{2}, \quad \text{kdykoliv } x \in U_{2\delta}(0) \text{ a} \tag{3.16}$$

$$f'(x) \text{ je izomorfismus, kdykoliv } x \in U_{2\delta}(0). \tag{3.17}$$

Podle (3.16) a Věty 3.19 dostáváme, že zobrazení g je lipschitzovské s konstantou $1/2$ na (konvexní) množině $U_{2\delta}(0)$. Pro každé dva body $U_{2\delta}(0)$ tedy platí

$$\begin{aligned} \|f(v) - f(u)\| &= \|v - u - (g(v) - g(u))\| \ge \\ &\ge \|v - u\| - \|g(v) - g(u)\| \ge \frac{1}{2}\|v - u\|. \end{aligned} \tag{3.18}$$

Odtud snadno dostáváme, že zobrazení $f \restriction_{U_{2\delta}(0)}$ je prosté a jeho inverze $(f \restriction_{U_{2\delta}(0)})^{-1}$ je lipschitzovská s konstantou 2.

Nyní ukážeme, že $U_{\delta/2}(0) \subset f(U_\delta(0))$. Zvolme proto libovolný bod $y \in U_{\delta/2}(0)$ a uvažme, že rovnice $f(x) = y$ je ekvivalentní s rovnicí $y + g(x) = x$. Položíme-li tedy $h(x) := y + g(x)$, chceme dokázat, že zobrazení h má v množině $U_\delta(0)$ pevný bod. K důkazu použijeme Banachovu větu o kontrakci (1.92) na restrikci $h^* := h \restriction_{\overline{U_\delta(0)}}$. Pro $x \in \overline{U_\delta(0)}$ platí $\|g(x)\| = \|g(x) - g(0)\| \le (1/2)\,\|x - 0\| \le \delta/2$, a tedy

$$\|h^*(x)\| = \|y + g(x)\| \le \|y\| + \|g(x)\| < \frac{\delta}{2} + \frac{\delta}{2} = \delta.$$

Zobrazení $h^*\colon \overline{U_\delta(0)} \to U_\delta(0)$ je lipschitzovské s konstantou $1/2$, protože zobrazení g je na $\overline{U_\delta(0)}$ takové. Je tedy h^* kontrakce na úplném metrickém prostoru $\overline{U_\delta(0)}$, takže h^* má podle Věty 1.92 pevný bod; tj. existuje $x \in \overline{U_\delta(0)}$, pro který $h^*(x) = x$, a tedy $f(x) = y$. Víme také, že $x = h^*(x) \in U_\delta(0)$.

Množina $U := f^{-1}\left(U_{\delta/2}(0)\right) \cap U_\delta(0)$ je zřejmě otevřená podmnožina $U_\delta(0)$. Položíme-li $\varphi := f \restriction_U$, je $\varphi\colon U \to U_{\delta/2}(0)$ bijekce třídy C^1. Zobrazení φ^{-1} je lipschitzovské (protože $(f \restriction_{U_{2\delta}(0)})^{-1}$ je lipschitzovské s konstantou 2), takže φ je homeomorfismus. Z (3.17) a Tvrzení 3.49 okamžitě dostáváme, že pro každý bod $y \in U_{\delta/2}(0)$ platí $\left((\varphi)^{-1}\right)'(y) = \left(\varphi'(\varphi^{-1}(y))\right)^{-1}$. Z Tvrzení 3.48 (ii) nyní snadno vyplývá, že zobrazení $y \mapsto (\varphi^{-1})'(y)$ je spojité na $U_{\delta/2}(0)$, protože je složením tří spojitých zobrazení:

$$(\varphi^{-1})' = z \circ \varphi' \circ \varphi^{-1}, \quad \text{kde} \quad z\colon L \mapsto L^{-1},\ L \in \mathrm{Izom}(X, Y). \tag{3.19}$$

Zobrazení φ^{-1} je tedy třídy C^1.

Nyní indukcí podle k dokážeme, že φ^{-1} je třídy C^k pro $1 \le k \le p$. Pro $k = 1$ to již víme. Předpokládejme, že $1 < k \le p$ a φ^{-1} je třídy C^{k-1}. Protože φ' je třídy C^{p-1} a z je třídy C^∞ (Tvrzení 3.48 (iv)), z rovnosti (3.19) a Věty 3.43 dostáváme, že $(\varphi^{-1})'$ je třídy C^{k-1}. Je tedy φ^{-1} třídy C^k a indukční krok je proveden. Dokázali jsme tedy, že φ^{-1} třídy C^p; důkaz v našem speciálním případě je dokončen.

V obecném případě položíme (podobně jako v důkazu Tvrzení 3.49)

$$\widetilde{f}(x) := (f'(a))^{-1}\,(f(a+x) - f(a)) \qquad \text{a} \qquad \widetilde{V} := V - a.$$

Snadno se ověří, že zobrazení $\widetilde{f}\colon \widetilde{V} \to X$ je třídy C^p, $\widetilde{f}(0) = 0$ a $(\widetilde{f})'(0) = I_X$. Podle první části důkazu tedy dostáváme, že existuje otevřené okolí $\widetilde{U}$ bodu 0 takové, že $\widetilde{f} \restriction_{\widetilde{U}}$ je difeomorfismus a $\left(\widetilde{f} \restriction_{\widetilde{U}}\right)^{-1}$ je třídy C^p. Uvážíme-li, že $f(z) = f'(a)(\widetilde{f}(z - a)) + f(a)$ pro $z \in U := \widetilde{U} + a$, snadno dostáváme, že $f \restriction_U$ je difeomorfismus. Vyjádříme-li ještě $(f \restriction_U)^{-1}$ pomocí $\left(\widetilde{f} \restriction_{\widetilde{U}}\right)^{-1}$ podobně jako v důkazu Tvrzení 3.49, snadno vidíme, že $(f \restriction_U)^{-1}$ je třídy C^p.

3.51 Poznámka.

" (i)" V případě $X = Y = \mathbb{R}^n$ k odvození (3.17) nepotřebujeme Tvrzení 3.48; stačí použít spojitost funkce $\det([f'(x)])$. Abychom dokázali, že φ je třídy C^p, potřebujeme vědět, že z je třídy C^p. To však lze v tomto případě dokázat zcela elementárně, srov. text před Lemmatem 3.46.

" (ii)" Ve větě stačí předpokládat, že $f'(a)\colon X \to Y$ je lineární bijekce (srov. Poznámka (?)izo2¿).

Následující věta o implicitně zadaném zobrazení se tradičně nazývá věta o implicitních funkcích (o implicitní funkci). Jedná se o její nejjednodušší formu; existuje mnoho jejích verzí (ve kterých jsou často předpoklady různým způsobem oslabeny).

3.52 Věta. (věta o implicitních funkcích) *Nechť X, Y, Z jsou Banachovy prostory a $p \in \mathbb{N} \cup \{\infty\}$. Předpokládejme toto:*

(i) Je dáno zobrazení g z $X \times Y$ do Z, které je třídy C^p na nějakém otevřeném okolí bodu $c = (a, b) \in X \times Y$.

(ii) $g(a, b) = 0$.

(iii) Parciální derivace $g_2'(c) = (g(a, \cdot))'(b)$ je izomorfismus prostoru Y na prostor Z.

Pak existují čísla $\delta_1 > 0$, $\delta_2 > 0$ a zobrazení $f\colon U_{\delta_1}(a) \to U_{\delta_2}(b)$ třídy C^p, pro které platí ekvivalence

$$(g(x, y) = 0 \ \wedge\ (x, y) \in U_{\delta_1}(a) \times U_{\delta_2}(b)) \iff y = f(x).$$

Větu lze dokázat zcela analogicky, jako je to provedeno na str. 115 pro případ eukleidovských prostorů (srov. [Ca]). Věta 3.50 se aplikuje na zobrazení r z $X \times Y$

do $X \times Z$ zadané předpisem $r(x,y) := (x, g(x,y))$. To, že $r'(c)$ je izomorfismus $X \times Y$ na $X \times Z$ je ovšem nutno ověřit jinak; inverzi $(r'(c))^{-1}$ není těžké explicitně spočítat.

4. Fourierovy řady a Fourierova transformace

4.1 Fourierovy řady periodických funkcí

Nejdůležitějším typem funkčních řad jsou bezesporu mocninné řady, pomocí kterých lze (lokálně) vyjádřit všechny elementární funkce a řadu dalších důležitých funkcí.

Velmi důležité jsou také trigonometrické řady, pomocí kterých lze vyjádřit „téměř všechny“ zajímavé *periodické* funkce. V základních aplikacích se pracuje pouze s nejdůležitějšími trigonometrickými řadami, které se nazývají Fourierovy řady.

Bez újmy na obecnosti (srov. Poznámka 4.16) se v dalším omezíme na případ 2π-periodických funkcí. Víme, že funkce $\cos x$, $\sin x$ jsou 2π-periodické funkce na $\mathbb{R}$. Dalšími 2π-periodickými funkcemi jsou ovšem také funkce

$$\cos nx, \ \sin nx, \quad n = 2, 3, \dots$$

(které mají minimální periodu $2\pi/n$) a také funkce konstantní.

Pokud řada tvaru $c + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)$ (neboli trigonometrická řada) konverguje v každém bodě $x \in \mathbb{R}$, je tedy funkce zadaná součtem této řady 2π-periodická.

Funkce $\cos nx$, $\sin nx$ ve fyzice popisují tzv. „harmonické kmitání“. Obecný „harmonický kmit“, který má minimální periodu $2\pi/n$, je popsán nenulovou funkcí tvaru $y(x) = a_n \cos nx + b_n \sin x$ (viz Poznámka 4.15).

Přirozeně (zvláště ve fyzice) vzniká otázka, zda lze *každou* „rozumnou“ 2π-periodickou funkci $f(x)$ vyjádřit ve tvaru

$$f(x) = c + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx),$$

tedy jako součet „harmonických kmitů“(a konstanty). Jinak řečeno: ptáme se, zda lze funkci f vyjádřit jako součet trigonometrické řady. To byl historicky první problém rozsáhlé teorie zvané *harmonická analýza*.

4.1 Poznámka. Tato otázka byla diskutována r. 1753 v souvislosti s popisem kmitání struny. D. Bernoulli (na rozdíl např. od D'Alemberta) se na základě fyzikálně motivovaných úvah domníval, že každou 2π-periodickou funkci (jejíž graf „lze nakreslit jedním tahem“) takto vyjádřit lze. Tento názor však tenkrát nepřevládl, protože D. Bernoulli nedal návod, jak čísla c, a_n, b_n pro danou f vypočítat. Takové vzorce nalezl (a aplikoval) mj. Fourier (r. 1822), který je právem považován za zakladatele harmonické analýzy.

Ukazuje se (první zcela přesný výsledek dokázal Dirichlet r. 1829), že jako součet trigonometrické řady se skutečně dají napsat „téměř všechny“ 2π-periodické spojité funkce – i ty, které v některých bodech nemají derivaci, a dokonce i některé funkce nespojité.

4.2 Poznámka. Do té doby se soudilo, že skutečný význam v matematice mají pouze funkce, které jsou „velmi hladké“ – k tomu přispívala známá skutečnost, že funkce, která je součtem mocninné řady, je třídy C^∞. To, že funkce, které jsou součtem i velmi jednoduchých trigonometrických řad (a jsou tedy zadány „jednoduchou formulí“), nemusí být hladké, bylo velmi překvapivé a zmíněný názor se začal měnit. To vedlo na začátku 20. století ke vzniku teorie reálných funkcí, která pracuje s velmi obecnými (i nespojitými) funkcemi, a je nyní nezbytná i v řadě důležitých aplikací.

Po těchto důležitých motivačních poznámkách začneme s přesnou teorií.

4.3 Definice. *Nechť a_i, $i = 0, 1, \dots$ a b_i, $i = 1, 2, \dots$ jsou reálná čísla.*

(1) Pak řadu funkcí

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx), \qquad x \in \mathbb{R},$$

nazýváme trigonometrickou řadou.

(2) Je-li dáno ještě $n \in \mathbb{N} \cup \{0\}$, funkce tvaru

$$p(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx), \qquad x \in \mathbb{R},$$

se nazývá trigonometrický polynom. Pokud navíc $a_n \neq 0$ nebo $b_n \neq 0$, řekneme, že tento trigonometrický polynom má stupeň n. (Stupeň nulového polynomu nedefinujeme.)

Pokud a_i, b_i jsou komplexní čísla, hovoříme o komplexní trigonometrické řadě a o komplexním trigonometrickém polynomu.

4.4 Poznámka.

(i) Důvod, proč se „konstantní člen“ tradičně píše ve tvaru $\frac{a_0}{2}$, je čistě formální – některé vzorce se zjednoduší.

(ii) Trigonometrický polynom (podobně jako obyčejný polynom) lze chápat jako „formální výraz“ $\frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$. Za chvíli uvidíme (viz Poznámka 4.11), že stupeň n i koeficienty $a_0, a_1, \dots, a_n$, $b_1, \dots, b_n$ jsou funkcí $p(x)$ jednoznačně určeny, takže je jedno (podobně jako u obyčejných polynomů), zda trigonometrický polynom definujeme jako funkci nebo jako výraz. Definice stupně trigonometrického polynomu je tedy korektní.

Systém funkcí

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x, \dots$$

se nazývá *reálný trigonometrický systém.*

Funkce je tedy trigonometrickým polynomem (resp. komplexním trigonometrickým polynomem), právě když je reálnou (resp. komplexní) lineární kombinací funkcí z reálného trigonometrického systému. (Trigonometrickou řadu lze *nepřesně* interpretovat jako „formální nekonečnou lineární kombinaci" funkcí z reálného trigonometrického systému.)

Reálnému trigonometrickému systému je velmi blízký *komplexní trigonometrický systém*, což je systém

$$\{e^{inx}\}_{n=-\infty}^{\infty} := \{e^{inx}\}_{n\in\mathbb{Z}}$$

komplexních funkcí reálné proměnné.

Z Eulerových vzorců snadno vyplývá, že komplexní trigonometrické polynomy jsou právě všechny komplexní lineární kombinace funkcí z komplexního trigonometrického systému. Přitom je funkce $p(x)$ komplexním trigonometrickým polynom stupně n, právě když lze psát ve tvaru

$$p(x) = \sum_{k=-n}^{n} c_k e^{ikx},$$

kde $c_k \in \mathbb{C}$ a $c_n \neq 0$ nebo $c_{-n} \neq 0$.

Nyní ještě ukážeme, že trigonometrické řady úzce souvisí s teorií (komplexních) mocninných řad. Uvažujme komplexní mocninnou řadu $\sum_{n=0}^{\infty} c_n z^n$ pro z ležící na jednotkové kružnici. Je tedy $c_n = a_n + ib_n$, kde $a_n, b_n \in \mathbb{R}$ a $z = e^{ix} = \cos x + i\sin x$, kde $x \in \mathbb{R}$. Pak platí

$$\begin{aligned}\sum_{n=0}^{\infty} c_n z^n &= \sum_{n=0}^{\infty} (a_n + ib_n)(\cos nx + i\sin nx)\\ &= \sum_{n=0}^{\infty}(a_n\cos nx - b_n \sin nx) + i\sum_{n=0}^{\infty}(b_n\cos nx + a_n\sin nx),\end{aligned}$$

má-li jeden ze tří výrazů smysl.

Speciálně: mocninnou řadu $\sum_{n=0}^{\infty} c_n z^n$ umíme sečíst na jednotkové kružnici, právě když umíme sečíst trigonometrické řady

$$\sum_{n=0}^{\infty}(a_n\cos nx - b_n\sin nx), \quad \sum_{n=0}^{\infty}(b_n\cos nx + a_n\sin nx)$$

na $\mathbb{R}$. (Poznamenejme, že tyto trigonometrické řady se nazývají vzájemně konjugované.)

4.5 Příklad. Mocninnou řadu $\sum_{n=0}^{\infty}\frac{z^n}{n!}$ dovedeme sečíst v celé komplexní rovině; tedy i na jednotkové kružnici. V našem případě $c_n = a_n = 1/n!$; dostáváme tedy

$$e^{\cos x + i\sin x} = \sum_{n=0}^{\infty}\frac{1}{n!}\cos nx + i\sum_{n=0}^{\infty}\frac{1}{n!}\sin nx,$$

$$\sum_{n=0}^{\infty}\frac{1}{n!}\cos nx = e^{\cos x}\cos(\sin x), \qquad \sum_{n=0}^{\infty}\frac{1}{n!}\sin nx = e^{\cos x}\sin(\sin x).$$

Dále budeme potřebovat základy teorie Lebesgueova integrálu; nebudeme pracovat s obecnými 2π-periodickými funkcemi, ale pouze s těmi, které jsou navíc lokálně Lebesgueovsky integrovatelné.

Úmluva Dále v této kapitole (není-li řečeno jinak) *jsou všechny integrály Lebesgueovy a všechny funkce jsou komplexní funkce reálné proměnné.* Také (Lebesgueovy) prostory L^p chápeme jako „komplexní prostory“ tvořené komplexními funkcemi (hovoříme-li o funkci jako o prvku Lebesgueova prostoru, myslíme tím ovšem příslušnou třídu ekvivalence). Důkazy jsou často prováděny jen pro reálné funkce, pokud důkaz pro komplexní funkci se ihned dostane po aplikaci výsledku na reálnou a imaginární část.

4.6 Definice. *Symbolem $\mathscr{P}(2\pi)$ budeme označovat množinu všech 2π-periodických funkcí které jsou (lebesgueovsky) integrovatelné (tj. mají konečný Lebesgueův integrál) na intervalu $[0, 2\pi]$.*

Je snadno vidět, že 2π-periodická funkce f leží v $\mathscr{P}(2\pi)$ právě tehdy, když je integrovatelná na každém omezeném intervalu, a to nastane právě tehdy, je-li f lokálně integrovatelná (užití Borelovy věty o pokrytí).

V dalším budeme funkce z $\mathscr{P}(2\pi)$ často integrovat přes různé intervaly délky 2π; ukážeme, že vždy dostaneme stejný výsledek.

4.7 Lemma. *Nechť $f \in \mathscr{P}(2\pi)$ a $a \in \mathbb{R}$. Pak*

$$\int_0^{2\pi} f(x)\,\mathrm{d}x = \int_a^{a+2\pi} f(x)\,\mathrm{d}x.$$

Důkaz. Nejdříve si uvědomíme, že pomocí substituce $x = y + 2k\pi$ pro libovolná $\alpha < \beta$ a $k \in \mathbb{Z}$ dostáváme

$$\int_{\alpha+2k\pi}^{\beta+2k\pi} f(x)\,\mathrm{d}x = \int_\alpha^\beta f(y+2k\pi)\,\mathrm{d}y = \int_\alpha^\beta f(y)\,\mathrm{d}y. \tag{4.1}$$

Nyní najděme $k \in Z$ takové, že $0 \le a + 2k\pi < 2\pi$. Podle (4.1) platí

$$\int_a^{a+2\pi} f(x)\,\mathrm{d}x = \int_{a+2k\pi}^{a+2k\pi+2\pi} f(x)\,\mathrm{d}x, \qquad \int_0^{a+2k\pi} f(x)\,\mathrm{d}x = \int_{2\pi}^{a+2k\pi+2\pi} f(x)\,\mathrm{d}x.$$

Protože $0 \le a + 2k\pi < 2\pi \le a + 2k\pi + 2\pi$, dostáváme

$$\begin{aligned}\int_a^{a+2\pi} f(x)\,\mathrm{d}x &= \int_{a+2k\pi}^{a+2k\pi+2\pi} f(x)\,\mathrm{d}x = \int_{a+2k\pi}^{2\pi} f(x)\,\mathrm{d}x + \int_{2\pi}^{a+2k\pi+2\pi} f(x)\,\mathrm{d}x \\ &= \int_{a+2k\pi}^{2\pi} f(x)\,\mathrm{d}x + \int_0^{a+2k\pi} f(x)\,\mathrm{d}x = \int_0^{2\pi} f(x)\,\mathrm{d}x.\end{aligned}$$

Jednou ze základních vlastností trigonometrického systému

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x, \dots$$

je jeho „ortogonalita". Pokud tento sytém chápeme jako podmnožinu (komplexního nebo reálného) prostoru $L^2(0,2\pi)$, můžeme uvozovky vynechat, protože $L^2(0,2\pi)$ je unitární prostor, ve kterém je pojem kolmosti dvou prvků definován. (Připomeňme, že skalární součin na komplexním $L^2(0,2\pi)$ je definován rovností $\langle f,g\rangle = \int_0^{2\pi} f\overline{g}$.) Nejdříve však dokážeme ortogonalitu komplexního trigonometrického systému, což je početně snazší.

4.8 Lemma.

(i) Jsou-li $f,g \in \{e^{inx}\colon n\in\mathbb{Z}\}$ *a* $f\neq g$, *pak* $\displaystyle\int_0^{2\pi} f\overline{g} = 0.$

(ii) $\displaystyle\int_0^{2\pi} e^{inx}\cdot\overline{e^{inx}}\,\mathrm{d}x = 2\pi.$

Důkaz. Tvrzení (i) plyne z toho, že $\int_0^{2\pi} e^{ikx}\,\mathrm{d}x = \left[e^{ikx}/ik\right]_0^{2\pi} = 0$ pro $k\in\mathbb{Z}$, $k\neq 0$, $\overline{e^{ikx}} = e^{-ikx}$ $(k\in\mathbb{Z})$ a $e^{inx}\cdot e^{-imx} = e^{i(n-m)x}$. Tvrzení (ii) je zřejmé.

4.9 Lemma.

(i) Jsou-li $f,g\in\{1,\cos x,\sin x,\cos 2x,\sin 2x,\dots\}$ *a* $f\neq g$, *pak* $\displaystyle\int_0^{2\pi} fg = 0.$

(ii) $\displaystyle\int_0^{2\pi} 1\cdot 1\,\mathrm{d}x = 2\pi,\quad \int_0^{2\pi}\sin kx\,\sin kx\,\mathrm{d}x = \int_0^{2\pi}\cos kx\,\cos kx\,\mathrm{d}x = \pi.$

Důkaz. (Náznak.) Jde jen o elementární výpočty. Jejich počet můžeme zmenšit, uvědomíme-li si, že pokud $m\in\mathbb{N}\cup\{0\}$, $n\in\mathbb{N}$ a $m\neq n$, pak rovnosti

$$\int_0^{2\pi}\cos mx\,\cos nx\,\mathrm{d}x = 0,\quad \int_0^{2\pi}\cos mx\,\sin nx\,\mathrm{d}x = 0,\quad \int_0^{2\pi}\sin mx\,\sin nx\,\mathrm{d}x = 0$$

snadno plynou z Eulerových vzorců a Lemmatu 4.8. Npapříklad zpaměti vidíme, že

$$\int_0^{2\pi}\cos mx\,\sin nx\,\mathrm{d}x = \langle\cos mx\ \sin nx\rangle = \langle\frac{e^{imx}+e^{-imx}}{2},\frac{e^{inx}-e^{-inx}}{2i}\rangle = 0.$$

Následující tvrzení dává přirozenou motivaci pro definici Fourierovy řady funkce $f\in\mathscr{P}(2\pi)$.

4.10 Tvrzení. *Nechť* $f\in\mathscr{P}(2\pi)$ *a platí*

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty}(a_k\cos kx + b_k\sin kx),\quad x\in\mathbb{R},$$

přičemž řada napravo je stejnoměrně konvergentní.

Pak platí (Fourierovy vzorce)

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \,\mathrm{d}x, \qquad k = 0, 1, 2, \ldots,$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \,\mathrm{d}x, \qquad k = 1, 2, \ldots.$$

Důkaz. Důkaz stačí provést v případě, že f je reálná funkce a čísla a_k, b_k jsou reálná čísla.

Nechť $n \in \mathbb{N} \cup \{0\}$. Pak pro každé $x \in \mathbb{R}$ platí

$$f(x) \cos nx = \frac{a_0}{2} \cos nx + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \cos nx. \tag{4.2}$$

Protože $\sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$ je stejnoměrně konvergentní řada omezených funkcí, snadno dostáváme, že tato řada, a tudíž i řada z (4.2), mají stejně omezené posloupnosti částečných součtů. Podle „Lebesgueovy věty pro řady“ (srov. [LM; 8. 14] tedy platí

$$\int_0^{2\pi} f(x) \cos nx \,\mathrm{d}x =$$

$$\int_0^{2\pi} \frac{a_0}{2} \cos nx \,\mathrm{d}x + \sum_{k=1}^{\infty} \int_0^{2\pi} (a_k \cos kx \cos nx + b_k \sin kx \cos nx) \,\mathrm{d}x.$$

Podle Lemmatu 4.9 snadno dostáváme pro $m \geq n$ rovnost

$$\int_0^{2\pi} s_m(x) \cos nx \,\mathrm{d}x =$$

$$= \int_0^{2\pi} \frac{a_0}{2} \cos nx + \sum_{k=1}^{m} \left(a_k \int_0^{2\pi} \cos kx \cos nx \,\mathrm{d}x + b_k \int_0^{2\pi} \sin kx \cos nx \,\mathrm{d}x \right) =$$

$$= a_n \int_0^{2\pi} \cos^2 nx \,\mathrm{d}x = \pi a_n,$$

takže

$$\int_0^{2\pi} f(x) \cos nx \,\mathrm{d}x = \pi a_n.$$

Zcela obdobně dostáváme

$$\int_0^{2\pi} f(x) \sin nx \,\mathrm{d}x = b_n \int_0^{2\pi} \sin^2 nx \mathrm{d}x = \pi b_n \qquad \text{a}$$

$$\int_0^{2\pi} f(x)\,\cos 0x\,\mathrm{d}x = \int_0^{2\pi} f(x)\cdot 1\,\mathrm{d}x = \frac{a_0}{2}\int_0^{2\pi} 1\,\mathrm{d}x = \pi a_0.$$

Z těchto rovností již okamžitě plynou Fourierovy vzorce.

4.11 Poznámka.

(i) Nyní je vidět, proč konstantní člen zapisujeme ve tvaru $a_0/2$; pro a_k pak dostáváme pro $k = 0$ vzorec stejného tvaru jako pro $k \geq 1$.

(ii) Nechť je dán trigonometrický polynom

$$p(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx).$$

Definujeme-li $a_k = b_k = 0$ pro $k > n$, máme funkci p vyjádřenu jako součet stejnoměrně konvergentní trigonometrické řady, takže a_k, b_k jsou určeny funkcí $p(x)$ pomocí Fourierových vzorců.

(i) Pro zdůvodnění záměny sumace a integrálu v předchozím důkazu jsme mohli použít i větu pro Riemannův (nebo Newtonův) integrál, protože je snadno vidět, že řada v (4.2) konverguje stejnoměrně. Také by místo stejnoměrné konvergence řady $\sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx)$ stačilo předpokládat, že její částečné součty mají na $[0, 2\pi]$ společnou integrabilní majorantu. Pak totiž řada v (4.2) má stejnou vlastnost.

Tvrzení 4.10 naznačuje, že pokud chceme funkci $f \in \mathscr{P}(2\pi)$ vyjádřit jako součet trigonometrické řady, měli bychom především zkoumat řadu, jejíž koeficienty jsou dány Fourierovými vzorci.

4.12 Definice. *Nechť $f \in \mathscr{P}(2\pi)$. Pak definujeme Fourierovy koeficienty funkce f takto:*

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\,\cos nx\,\mathrm{d}x, \qquad n = 0, 1, 2, \ldots,$$

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\,\sin nx\,\mathrm{d}x, \qquad n = 1, 2, \ldots.$$

Trigonometrická řada

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

se pak nazývá Fourierova řada funkce f a budeme ji označovat symbolem F_f. Zápis

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

vyjadřuje, že řada napravo je Fourierovou řadou funkce f.

Fourierova řada je řadou

$$a_0/2 + (a_1 \cos x + b_1 \sin x) + (a_2 \cos 2x + b_2 \sin 2x) + \dots,$$

přičemž se její částečné součty tradičně „indexují od 0", tj.

$$s_0(x) := \frac{a_0}{2}, \dots, s_n(x) := \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx), \dots.$$

V případě, že f je reálná, je Fourierova řada funkce f reálná trigonometrická řada.

4.13 Poznámka.

(i) Někdy se říká, že F_f je *Lebesgue – Fourierova řada* funkce f, aby se zdůraznila skutečnost, že při definici Fourierových koeficientů se používá Lebesgueův integrál (před jeho zavedením se ovšem pracovalo hlavně s Riemannovým integrálem a tedy s Riemann–Lebesgueovými řadami).

(ii) Zápis (zcela běžný)

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

nic neříká o konvergenci (Fourierovy) řady stojící napravo. Ta dokonce může divergovat ve všech bodech.

(iii) Symbol F_f pro Fourierovu řadu není běžný a nebudeme jej zde proto příliš používat. Někteří autoři používají symbol $S[f]$.

(iv) Podle Lemmatu 4.7 pro Fourierovy koeficienty platí

$$a_n = \frac{1}{\pi} \int_a^{a+2\pi} f(x) \cos nx \,\mathrm{d}x, \qquad n = 0, 1, 2, \dots,$$

$$b_n = \frac{1}{\pi} \int_a^{a+2\pi} f(x) \sin nx \,\mathrm{d}x, \qquad n = 1, 2, \dots$$

pro libovolné $a \in \mathbb{R}$. Pro $a = -\pi$, jde o integrály od $-\pi$ do π.

Fourierovu řadu lze ekvivalentně zapsat v tzv. *komplexním tvaru*, ve kterém se nepracuje s reálným, ale s komplexním trigonometrickým systémem. Tento intuitivně méně pochopitelný, ale z teoretického i početního hlediska často výhodnější tvar, berou dnes někteří autoři přímo za definici Fourierovy řady.

Uvažujme $f \in \mathscr{P}(2\pi)$ a upravme vzorec pro částečný součet Fourierovy řady funkce f

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

podle Eulerových vzorců

$$\cos kx = \frac{1}{2}(e^{ikx} + e^{-ikx}), \quad \sin kx = \frac{1}{2i}(e^{ikx} - e^{-ikx}) = \frac{i}{2}(e^{-ikx} - e^{ikx}).$$

Dostáváme

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(\frac{1}{2}(a_k - ib_k)e^{ikx} + \frac{1}{2}(a_k + ib_k)e^{-ikx} \right).$$

Položíme-li

$$c_0 := \frac{1}{2}a_0 \quad \text{a} \quad c_k := \frac{1}{2}(a_k - ib_k), \quad c_{-k} := \frac{1}{2}(a_k + ib_k), \quad k \in \mathbb{N}, \tag{4.3}$$

máme

$$s_n(x) = \sum_{k=-n}^{n} c_k e^{ikx}. \tag{4.4}$$

Dosadíme-li do (4.3) z definice a_n, b_n a užijeme Eulerovy vzorce, snadno dostáváme

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\, e^{-ikx}, \qquad k \in Z.$$

Je zvykem psát

$$f(x) \sim \sum_{k=-\infty}^{\infty} c_k e^{ikx}$$

a výraz napravo nazývat *komplexním tvarem Fourierovy řady funkce* f. Čísla c_k se nazývají *komplexní Fourierovy koeficienty* funkce f.

Přitom se ale řada napravo *nechápe* jako zobecněná řada. Její částečné součty $s_n(x)$ jsou *podle definice* dány vzorcem (4.4) a

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx} = s \quad \Longleftrightarrow \lim_{n\to\infty} s_n(x) = s.$$

4.14 Poznámka. Fourierova řada funkce $f \in \mathscr{P}(2\pi)$ nemusí konvergovat absolutně, v tom případě pak při jejím sčítání záleží na uspořádání jejích členů. Vzhledem k tomu, že podle (4.3) a Věty 4.23 platí $\lim_{n\to\pm\infty} c_n = 0$, číslo s je součtem Fourierovy řady funkce f v bodě x, právě když

$$c_o + c_1 e^{ix} + c_{-1}e^{-ix} + c_2 e^{2ix} + c_{-2}e^{-2ix} + \ldots \;=\; s.$$

4.15 Poznámka. Obecná reálná funkce popisující tzv. „harmonický kmit“ má tvar $x(t) = A\,\sin(\omega t + \varphi)$, kde $A > 0$ je tzv. amplituda, $\varphi \in \mathbb{R}$ počáteční fáze a $\omega > 0$ je kruhová frekvence. Nejmenší perioda této funkce je $T = 2\pi/\omega$. Harmonické kmitání v přírodě vzniká přirozeným způsobem (je řešením „rovnice harmonického oscilátoru“) a je to z fyzikálního hlediska nejpřirozenější kmitání (periodický děj). Použijeme-li součtové vzorce, vidíme, že každou reálnou funkci (proměnné t) tvaru $x(t) = A\,\sin(\omega t + \varphi)$ lze psát ve tvaru $a\,\cos\,\omega t + b\,\sin\,\omega t$, kde

$$a = A\,\sin\varphi,\ b = A\,\cos\varphi, \quad \text{tj.}\ \ b + ai = A\,e^{i\varphi}. \tag{4.5}$$

Naopak, jsou-li dána čísla a, b, $(a, b) \neq (0, 0)$, můžeme z (4.5) určit čísla A, φ (číslo $b+ai$ vyjádříme v trigonometrickém tvaru; jednoznačnost „argumentu“ φ lze dosáhnout požadavkem $\varphi \in (-\pi, \pi]$), takže nenulové funkce tvaru $a\,\cos\,\omega t + b\,\sin\,\omega t$ popisují obecný harmonický kmit s kruhovou frekvencí ω. Z toho ihned plyne, že složení (součet) konečně mnoha harmonických („stejnosměrných“) kmitů stejné kruhové frekvence je opět harmonický kmit téže kruhové frekvence. Vidíme také, že každou reálnou trigonometrickou řadu (tedy i Fourierovu řadu) lze psát v (tzv. fázovém) tvaru

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} A_n\,\sin(nx + \varphi_n),$$

kde $A_n \geq 0$ a $\varphi_n \in (-\pi, \pi]$. (Poznamenejne, že často se ve fázovém tvaru užívá kosinus místo sinu; vše vychází zcela analogicky.)

Základní otázku teorie Fourierových řad (viz Poznámka 4.1) lze tedy „fyzikálně" formulovat takto: Je každý „rozumný" 2π-periodický kmit složením harmonických periodických kmitů (a konstantního kmitu) ?

4.16 Poznámka. Jestliže vyšetřujeme periodické funkce s obecnou periodou $l > 0$, máme dvě možnosti:

a) Lze postupovat zcela analogicky jako v případě $l = 2\pi$. Příslušný reálný trigonometrický systém pak je

$$1, \cos \tfrac{2\pi}{l} x,\ \sin \tfrac{2\pi}{l} x, \ldots, \cos \tfrac{2k\pi}{l} x,\ \sin \tfrac{2k\pi}{l} x, \ldots;$$

příslušná Fourierova řada má tvar

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos \frac{2k\pi}{l} x + b_k \sin \frac{2k\pi}{l} x),$$

kde Fourierovy koeficienty a_k, b_k jsou definovány vzorci

$$a_k = \frac{2}{l} \int_0^l f(x) \cos \frac{2k\pi}{l} x \,\mathrm{d}x, \qquad b_k = \frac{2}{l} \int_0^l f(x) \sin \frac{2k\pi}{l} x \,\mathrm{d}x$$

a celá teorie je naprosto analogická.

b) Teorii lze vybudovat jen pro případ $l = 2\pi$ a výsledky o l-periodických funkcích $f(x)$ snadno získat z výsledků o 2π-periodických funkcích $f^*(y) := f(ly/2\pi)$.

(Obvyklý je druhý postup.)

Pokud $f \in \mathscr{P}(2\pi)$ a $a_k = 0$ pro $k = 0, 1, \ldots$, máme

$$f \sim 0 + (0 + b_1 \sin x) + (0 + b_2 \sin 2x) + \ldots;$$

je ale zvykem nulové členy vynechávat a psát

$$f \sim b_1 \sin x + b_2 \sin 2x + \ldots.$$

V tom případě říkáme, že Fourierova řada F_f je sinová řada.

Pokud $b_k = 0$ pro $k = 1, 2, \ldots$, říkáme, že F_f je kosinová řada a píšeme

$$f \sim \tfrac{a_0}{2} + a_1 \cos x + a_2 \cos 2x + \ldots.$$

4.17 Tvrzení. *Je-li $f \in \mathscr{P}(2\pi)$ sudá funkce, pak F_f je kosinová řada. Jestliže $f \in \mathscr{P}(2\pi)$ je lichá funkce, pak F_f je sinová řada.*

Důkaz. Pokud g je integrovatelná a lichá na intervalu $(-\pi, \pi)$, pak

$$\int_{-\pi}^{\pi} g(x) \,\mathrm{d}x = \int_{-\pi}^{0} g(x) \,\mathrm{d}x + \int_0^{\pi} g(x) \,\mathrm{d}x =$$
$$= \int_0^{\pi} g(-x) \,\mathrm{d}x + \int_0^{\pi} g(x) \,\mathrm{d}x = \int_0^{\pi} (g(-x) + g(x)) \,\mathrm{d}x = 0.$$

Nyní si stačí uvědomit, že funkce $f(x) \cos kx$ je lichá, pokud f je lichá a funkce $f(x) \sin kx$ je lichá, pokud f je sudá.

Obr. 4.16.

4.18 Poznámka. Tvrzení lze „obrátit“ (viz Poznámka 4.42 níže); například platí: Je-li F_f kosinová řada, pak f je s.v. rovna sudé funkci.

4.19 Příklad. Snažme se funkci x^2 napsat na intervalu $(-\pi, \pi]$ jako součet trigonometrické řady. Funkci $g(x) = x^2, x \in (-\pi, \pi]$ můžeme zřejmě rozšířit právě jedním způsobem na celé $\mathbb{R}$ tak, aby výsledná funkce byla 2π-periodická. Toto rozšíření g^* je zřejmě na $\mathbb{R}$ spojité a sudé, proto patří do $\mathscr{P}(2\pi)$ a má kosinovou Fourierovu řadu. Integrací per partes dostáváme

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2 \cos kx \,\mathrm{d}x = (-1)^k \frac{4}{k^2}, \quad k \geq 1, \qquad a_0 = \frac{1}{\pi}\int_{-\pi}^{\pi} x^2 = \frac{2\pi^2}{3}.$$

Platí tedy (nuly vynecháváme)

$$f(x) \sim \frac{\pi^2}{3} - 4\frac{\cos x}{1^2} + 4\frac{\cos 2x}{2^2} - 4\frac{\cos 3x}{3^2} + \dots .$$

Zatím ovšem *nevíme*, zda (v některých bodech) platí

$$f(x) = \frac{\pi^2}{3} - 4\frac{\cos x}{1^2} + 4\frac{\cos 2x}{2^2} - 4\frac{\cos 3x}{3^2} + \dots \tag{4.6}$$

V dalším ukážeme několika způsoby, že rovnost platí pro každé $x \in \mathbb{R}$. Pro $x = 0$ a $x = \pi$ pak budeme mít dokázány zajímavé vzorce

$$\frac{1}{1^2} - \frac{1}{2^2} + \frac{1}{3^2} - \dots = \frac{\pi^2}{12}, \qquad \frac{1}{1^2} + \frac{1}{2^2} + \frac{1}{3^2} + \dots = \frac{\pi^2}{6}.$$

Na obr. 4.16 je znázorněna funkce $g^*(x)$ a částečný součet $s_2(x)$ její Fourierovy řady.

Dále se snažíme najít obecné věty, které pro „pěkné“ funkce $f \in \mathscr{P}(2\pi)$ udávají vztah mezi součtem Fourierovy řady a funkcí f. Pro důkazy takových vět je vhodné vyjádřit částečný součet Fourierovy řady v jednoduchém tvaru. Takové vyjádření nalezl již Dirichlet (r. 1829) pomocí tzv. Dirichletova jádra.

4.20 Lemma. (o Dirichletově jádře) *Pro každé $n \in \mathbb{N} \cup \{0\}$ položme*

$$D_n(x) := \frac{1}{2} + \cos x + \cos 2x + \dots + \cos nx.$$

OBR. 4.17.

Pro tuto funkci, kterou nazýváme Dirichletovo jádro, platí:

(i) $$D_n(x) = \frac{\sin{(n+\frac{1}{2})x}}{2\sin\frac{x}{2}} \qquad \textit{pro} \qquad x \neq 2k\pi \ (k \in \mathbb{Z});$$

(ii) D_n *je funkce sudá, spojitá,* 2π*-periodická a* $D_n(0) = n + \frac{1}{2}$.

(iii) $$\int_{-\pi}^{\pi} D_n = \pi.$$

Důkaz. Pomocí Eulerových vzorců

$$\cos kx = \frac{1}{2}\left(e^{ikx} + e^{-ikx}\right), \qquad \sin kx = \tfrac{1}{2i}\left(e^{ikx} - e^{-ikx}\right)$$

a vzorce pro částečný součet geometrické řady dostáváme, že

$$\begin{aligned} D_n(x) &= \frac{1}{2}\left(e^{-inx} + \cdots + e^{-ix} + 1 + e^{ix} + \cdots + e^{inx}\right) = \frac{1}{2}e^{-inx}\frac{e^{ix(2n+1)}-1}{e^{ix}-1} \\ &= \frac{1}{2}e^{-inx}\frac{e^{ix(2n+1)}-1}{e^{ix}-1}\,\frac{e^{-\frac{ix}{2}}}{e^{-\frac{ix}{2}}} = \frac{1}{2}\,\frac{e^{ix(n+\frac{1}{2})} - e^{-ix(n+\frac{1}{2})}}{e^{\frac{ix}{2}} - e^{\frac{-ix}{2}}} = \frac{\sin{(n+\frac{1}{2})x}}{2\sin\frac{x}{2}}, \end{aligned}$$

má-li výraz napravo smysl; tím je dokázáno (i). Tvrzení (ii) je zřejmé a (iii) dostáváme okamžitě z rovností

$$\int_{-\pi}^{\pi}\frac{1}{2} = \pi, \qquad \int_{-\pi}^{\pi}\cos kx\,\mathrm{d}x = 0, \quad k \in \mathbb{N}.$$

Dirichletovo jádro je (pro $n = 5$) znázorněno na obr. 4.17 vlevo.

4.21 Lemma. (vyjádření částečných součtů Fourierovy řady) *Nechť* $f \in \mathscr{P}(2\pi)$

a $f \sim \dfrac{a_0}{2} + \displaystyle\sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$. *Pak*

$$\begin{aligned} s_n(x) &:= \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x+y) D_n(y)\,\mathrm{d}y = \frac{1}{\pi} \int_{0}^{\pi} (f(x+y) + f(x-y)) D_n(y)\,\mathrm{d}y. \end{aligned}$$

Důkaz. Po dosazení za Fourierovy koeficienty a snadné úpravě máme

$$\begin{aligned} s_n(x) &= \frac{1}{2}\left(\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\mathrm{d}t\right) + \\ &+ \sum_{k=1}^{n}\left[\left(\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos kt\,\mathrm{d}t\right)\cos kx + \left(\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\sin kt\,\mathrm{d}t\right)\sin kx\right] \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left[\frac{1}{2} + \sum_{k=1}^{n}(\cos kt \cos kx + \sin kt \sin kx)\right]\mathrm{d}t. \end{aligned}$$

Užijeme-li vzorec

$$\cos kt \cos kx + \sin kt \sin kx = \cos(k(t-x)),$$

definici $D_n(t-x)$ a Lemma 4.7, dostáváme

$$\begin{aligned} s_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t) D_n(t-x)\,\mathrm{d}t &= \frac{1}{\pi}\int_{-\pi-x}^{\pi-x} f(x+y) D_n(y)\,\mathrm{d}y \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+y) D_n(y)\,\mathrm{d}y. \end{aligned}$$

Pomocí substituce $z = -y$ s využitím sudosti Dirichletova jádra dostáváme:

$$\begin{aligned} \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+y) D_n(y)\,\mathrm{d}y &= \frac{1}{\pi}\int_{-\pi}^{0} f(x+y) D_n(y)\,\mathrm{d}y + \frac{1}{\pi}\int_{0}^{\pi} f(x+y) D_n(y)\,\mathrm{d}y \\ &= \frac{1}{\pi}\int_{0}^{\pi} f(x-z) D_n(z)\,\mathrm{d}z + \frac{1}{\pi}\int_{0}^{\pi} f(x+y) D_n(y)\,\mathrm{d}y \\ &= \frac{1}{\pi}\int_{0}^{\pi} (f(x+y) + f(x-y)) D_n(y)\,\mathrm{d}y. \end{aligned}$$

4.22 Poznámka. Integrálu $\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+y) D_n(y)\,\mathrm{d}y$ se říká Dirichletův integrál. Protože Dirichletovo jádro je (i co do absolutní hodnoty) největší v 0, zdá se být pravděpodobné, že hodnota Dirichletova integrálu (tj. $s_n(x)$) bude nejvíce záviset na hodnotách funkce f v blízkosti bodu x. Tato domněnka skutečně platí ve velmi silném smyslu, srov. Poznámka 4.28. Klíčem k přesnému důkazu je následující důležitá věta.

4.23 Věta. (Riemann–Lebesgueovo lemma) *Nechť $a, b \in \mathbb{R}^*$ a funkce f je lebesgueovsky integrovatelná na (a, b). Pak*

$$\text{(4.7)} \qquad \lim_{\lambda\to\infty} \int_a^b f(x) \cos \lambda x \, \mathrm{d}x = 0 \quad \text{a} \quad \lim_{\lambda\to\infty} \int_a^b f(x) \sin \lambda x \, \mathrm{d}x = 0.$$

Speciálně: Posloupnosti Fourierových koeficientů (a_n), (b_n) *libovolné funkce z* $\mathscr{P}(2\pi)$ *mají limitu* 0.

Důkaz. Větu stačí dokázat pro reálnou funkci f. Dokážeme první část (4.7); druhou část lze dokázat zcela obdobně. Pro reálný prostor $L^1(a, b)$ položme

$$Q := \{g \in L^1(a,b) : \lim_{\lambda\to\infty} \int_a^b g(x) \cos \lambda x \, \mathrm{d}x = 0\};$$

chceme dokázat, že $Q = L^1(a, b)$. Nejdříve ukážeme, že

$$\text{(4.8)} \qquad Q \quad \text{je uzavřený lineární podprostor} \quad L^1(a, b).$$

To, že Q je lineární prostor, plyne okamžitě z rovnosti

$$\int_a^b (c_1 g_1(x) + c_2 g_2(x)) \cos \lambda x \, \mathrm{d}x = c_1 \int_a^b g_1(x) \cos \lambda x \, \mathrm{d}x + c_2 \int_a^b g_2(x) \cos \lambda x \, \mathrm{d}x.$$

Nyní uvažujme funkci $h \in \overline{Q}$. Máme dokázat, že $h \in Q$, tj.

$$\lim_{\lambda\to\infty} \int_a^b h(x) \cos \lambda x \, \mathrm{d}x = 0;$$

to provedeme z definice limity. Zvolme tedy libovolné $\varepsilon > 0$ a nalezněme funkci $g \in Q$ takovou, že

$$\|h - g\|_1 = \int_a^b |h(x) - g(x)| \, \mathrm{d}x < \frac{\varepsilon}{2}.$$

Protože $g \in Q$, existuje $\lambda_0 \in \mathbb{R}$ takové, že

$$\left| \int_a^b g(x) \cos \lambda x \, \mathrm{d}x \right| < \frac{\varepsilon}{2} \qquad \text{pro} \qquad \lambda > \lambda_0.$$

Pro $\lambda > \lambda_0$ tedy platí

$$\begin{aligned}
\left| \int_a^b h(x) \cos \lambda x \, \mathrm{d}x \right| &= \left| \int_a^b (h(x) - g(x)) \cos \lambda x \, \mathrm{d}x + \int_a^b g(x) \cos \lambda x \, \mathrm{d}x \right| \le \\
&\le \left| \int_a^b (h(x) - g(x)) \cos \lambda x \, \mathrm{d}x \right| + \left| \int_a^b g(x) \cos \lambda x \, \mathrm{d}x \right| \le \\
&\le \int_a^b |h(x) - g(x)| \, \mathrm{d}x + \frac{\varepsilon}{2} < \varepsilon.
\end{aligned}$$

Pomocí (4.8) budeme nyní postupně pro stále obecnější druhy funkcí $f \in L^1(a, b)$ dokazovat, že patří do Q, až nakonec dostaneme kýženou rovnost $Q = L^1(a, b)$. Všechny funkce tedy uvažujeme pouze na (a, b) a vždy předpokládáme $f \in L^1(a, b)$.

a) Nechť $f = C_I$ (charakteristická funkce I), přičemž $I = (c, d) \subset (a, b)$ je otevřený (nutně omezený) interval. Pak pro $\lambda > 0$ máme

$$\left|\int_a^b f(x)\,\cos\lambda x\,\mathrm{d}x\right| = \left|\int_c^d \cos\lambda x\,\mathrm{d}x\right| = \left|\frac{1}{\lambda}(\sin\lambda d - \sin\lambda c)\right| \le \frac{2}{\lambda},$$

takže zřejmě $f \in Q$.

b) Nechť $f = C_G$, kde $G = \bigcup_{k=1}^n I_k$, $I_1, \dots, I_n$ jsou po dvou disjunktní otevřené intervaly a $G \subset (a, b)$. Pak $f = \sum_{k=1}^n C_{I_k}$, takže podle (4.8) a kroku a) máme $f \in Q$.

c) Nechť $f = C_G$, kde $G = \bigcup_{k=1}^\infty I_k$, $I_1, I_2, \dots$ jsou po dvou disjunktní otevřené intervaly a $G \subset (a, b)$. Pak zřejmě

$$\|f - f_n\|_1 = \lambda\left(\bigcup_{k=n+1}^{\infty} I_k\right) \to 0, \quad n \to \infty, \qquad \text{kde} \qquad f_n := C_{\bigcup_{k=1}^n I_k}.$$

Protože $f_n \in Q$ podle kroku b), z (4.8) dostáváme $f \in Q$.

d) Nechť $f = C_M$, kde $M \subset (a, b)$ je lebesgueovsky měřitelná množina; pak $\lambda(M) < \infty$. Z regularity Lebesgueovy míry snadno dostaneme otevřené množiny G_n, $n = 1, 2, \dots$, takové, že $M \subset G_n \subset (a, b)$ a $\lambda(G_n \setminus M) < 1/n$. Zřejmě $C_{G_n} \to f$ v $L^1(a, b)$. Protože podle kroků b), c) máme $C_{G_n} \in Q$, z (4.8) dostáváme $f \in Q$.

e) Nechť f je jednoduchá funkce. Pak lze psát ve tvaru $\sum_{k=1}^n c_i C_{M_i}$, kde $c_i \in \mathbb{R}$ a $M_i \subset (a, b)$ mají konečnou Lebesgueovu míru. Podle (4.8) a kroku d) máme $f \in Q$.

f) Protože jednoduché měřitelné funkce jsou husté v $L^1(a, b)$ (viz např. [Ru2; Věta 3.13]), je podle kroku e) i množina Q hustá v $L^1(a, b)$; z uzavřenosti Q tedy máme konečně $Q = L^1(a, b)$.

4.24 Poznámka. V důkazu jsme podstatně využili, že množina jednoduchých funkcí je hustá v $L^1(a, b)$. Na základě hustoty jiných podmnožin $L^1(a, b)$ lze podat alternativní důkazy:

a) Velmi krátký důkaz dostáváme, pokud víme, že v $L^1(a, b)$ je hustá množina M těch funkcí $f \in C^1(a, b)$, pro které $\overline{\{x\colon f(x) \neq 0\}} \subset (a, b)$. Pomocí integrace per partes totiž snadno dostáváme, že $M \subset Q$.

b) Víme-li, že množina C spojitých funkcí f na (a, b), pro které $\overline{\{x\colon f(x) \neq 0\}} \subset (a, b)$, je hustá v $L^1(a, b)$, lze z kroku a) předchozího důkazu usoudit, že množina $A \subset L^1(a, b)$ po částech konstantních funkcí je částí Q, a pak (ze stejnoměrné spojitosti funkcí z C) dokázat, že $C \subset \overline{A}$, takže $\overline{A} = L^1(a, b)$.

4.25 Poznámka. (o Fourierových koeficientech) Nechť $a_n(f)$, $b_n(f)$ jsou Fourierovy koeficienty funkce $f \in \mathscr{P}(2\pi)$. Podle Riemann-Lebesgueova lemmatu $a_n(f) \to 0$, $b_n(f) \to 0$. Existují však posloupnosti $a_n \to 0, b_n \to 0$, které nejsou Fourierovy koeficienty (srov. [Ru2; Věta 5.15]). Pokud však (a_n), (b_n) jdou k nule dostatečně rychle (stačí, aby řada $\sum(a_n^2 + b_n^2)$ konvergovala; viz Věta 4.76), pak $(a_n)_0^\infty$, $(b_n)_0^\infty$ jsou Fourierovy koeficienty.

Dále lze zhruba říci toto: čím hladší je funkce $f \in \mathscr{P}(2\pi)$, tím rychleji jdou její koeficienty k nule. Je-li například $f \in C^k(\mathbb{R})$, snadno lze dokázat, že $a_n(f) = o(n^k)$, $n \to \infty$, $b_n(f) = o(n^k)$, $n \to \infty$; (srov. [J II; Věta 189]). Pro $k \ge$ pak vidíme, že f lze „velmi dobře aproximovat“ trigonometrickými polynomy s_n stupně nejvýše n.

4.26 Příklad. Jako příklad na užití předchozí věty dokážeme důležitou rovnost

$$\int_0^\infty \frac{\sin x}{x}\,\mathrm{d}x = \frac{\pi}{2}. \tag{4.9}$$

Jak známo, tento integrál v Lebesgueově smyslu neexistuje. Jako nevlastní integrál (zobecněný Riemannův, Newtonův nebo zobecněný Lebesgueův) existuje (konverguje). (To plyne z Dirichletova kritéria, ale také snadno integrací per partes.) K jeho výpočtu použijeme vlastnosti Dirichletova jádra. Vzhledem k sudosti D_n a Lemmatu 4.20 (iii) platí

$$\int_0^\pi D_n(x)\,\mathrm{d}x = \int_0^\pi \frac{\sin\,(n+\frac{1}{2})x}{2\sin\frac{x}{2}}\,\mathrm{d}x = \frac{\pi}{2}. \tag{4.10}$$

Ukážeme, že (pro $n \to \infty$)

$$\int_0^\pi \frac{\sin\,(n+\frac{1}{2})x}{2\sin\frac{x}{2}}\,\mathrm{d}x - \int_0^\pi \frac{\sin(n+\frac{1}{2})x}{x}\,\mathrm{d}x = \int_0^\pi \left(\frac{1}{2\sin\frac{x}{2}} - \frac{1}{x}\right)\sin(n+\frac{1}{2})x \;\to\; 0.$$

K tomu stačí podle předešlé věty (a snadné části Heineho věty) ukázat, že funkce $f(x) := (2\sin(x/2))^{-1} - 1/x$ leží v $L(0,\pi)$. To však snadno plyne z toho, že f lze spojitě rozšířit na $[0,\pi]$, protože elementární výpočet dává $\lim_{x\to 0+} f(x) = 0$. Dostáváme tedy

$$\lim_{n\to\infty}\int_0^\pi \frac{\sin(n+\frac{1}{2})x}{x}\,\mathrm{d}x = \frac{\pi}{2}.$$

Pomocí substituce $(n+1/2)x = y$ a konvergence integrálu z (4.9) máme

$$\lim_{n\to\infty}\int_0^\pi \frac{\sin(n+\frac{1}{2})x}{x}\,\mathrm{d}x = \lim_{n\to\infty}\int_0^{\pi(n+1/2)} \frac{\sin y}{y}\,\mathrm{d}y = \int_0^\infty \frac{\sin y}{y}\,\mathrm{d}y,$$

takže jsme dokázali (4.9).

4.27 Věta. (O lokalizaci) *Nechť $f \in \mathscr{P}(2\pi)$, $x \in \mathbb{R}$, $0 < \delta < \pi$, $s \in \mathbb{R}$ a nechť $s_n(x)$ je částečný součet Fourierovy řady funkce f v bodě x. Pak*

$$\lim_{n\to\infty} s_n(x) = s \iff \lim_{n\to\infty}\int_0^\delta (f(x+v)+f(x-v)-2s)\,\frac{\sin\,(n+\frac{1}{2})v}{\sin\frac{v}{2}}\,\mathrm{d}v \;=\; 0.$$

Důkaz. Protože podle Lemmatu 4.21 a Lemmatu 4.20 (ii), (iii)

$$s_n(x) = \frac{1}{\pi}\int_0^\pi (f(x+v)+f(x-v))\,D_n(v)\,\mathrm{d}v \quad \text{a} \quad s = \frac{1}{\pi}\int_0^\pi 2sD_n(v)\,\mathrm{d}v,$$

dostáváme

$$\begin{aligned} s_n(x) - s &= \frac{1}{\pi}\int_0^\pi (f(x+v)+f(x-v)-2s)\,D_n(v)\,\mathrm{d}v = \\ &= \frac{1}{2\pi}\int_0^\delta (f(x+v)+f(x-v)-2s)\,\frac{\sin\,(n+\frac{1}{2})v}{\sin\frac{v}{2}}\,\mathrm{d}v + \\ &+ \frac{1}{2\pi}\int_\delta^\pi \frac{f(x+v)+f(x-v)-2s}{\sin\frac{v}{2}}\,\sin\,(n+\frac{1}{2})v\,\mathrm{d}v =: A_n + B_n. \end{aligned}$$

Položíme-li pro $v \in (\delta, \pi)$

$$g(v) := \frac{f(x+v) + f(x-v) - 2s}{\sin \frac{v}{2}}, \quad \text{máme} \quad |g(v)| \le \frac{|f(x+v)| + |f(x-v)| + 2|s|}{\sin \frac{\delta}{2}},$$

takže zřejmě $g \in L^1(\delta, \pi)$ (substituce $t = x \pm v$). Z Riemann-Lebesgueova lemmatu dosáváme ($(a,b) := (\delta, \pi)$, $\lambda_n := n + \frac{1}{2}$), že $\lim_{n\to\infty} B_n = 0$, z čehož již tvrzení věty snadno plyne.

4.28 Poznámka. Protože integrál z tvrzení věty závisí jen na hodnotách funkce f v intervalu $(x-\delta, x+\delta)$, vidíme, že *konvergence Fourierovy řady funkce f v bodě x je „lokální vlastnost“*. Přesněji: pokud $f \in \mathscr{P}(2\pi)$, $g \in \mathscr{P}(2\pi)$ se rovnají na nějakém okolí bodu x, pak Fourierova řada funkce f konverguje v bodě x, právě když v něm konverguje Fourierova řada funkce g, a to ke stejnému součtu.

4.29 Věta. (Diniho kritérium) *Nechť $f \in \mathscr{P}(2\pi)$, $x \in \mathbb{R}$, $s \in \mathbb{R}$ a nechť existuje $\delta > 0$ takové, že Lebesgueův integrál*

$$\int_0^\delta \frac{f(x+v) + f(x-v) - 2s}{v} \, \mathrm{d}v \tag{4.11}$$

konverguje. Pak s je součtem Fourierovy řady funkce f v bodě x.

Důkaz. Nechť je δ z předpokladů věty dáno; můžeme předpokládat, že $0 < \delta < \pi$. Podle Věty 4.27 stačí dokázat, že

$$\lim_{n\to\infty} \int_0^\delta \frac{f(x+v) + f(x-v) - 2s}{v} \cdot \frac{v}{\sin \frac{v}{2}} \sin{(n+1/2)v} \, \mathrm{d}v = 0. \tag{4.12}$$

To však okamžitě plyne z Riemann-Lebesgueova lemmatu, protože funkce $\frac{v}{\sin \frac{v}{2}}$ je omezená na $(0,\delta)$ (lze zřejmě spojitě rozšířit na $[0,\delta]$), takže z konvergence (4.11) vyplývá, že funkce

$$g(v) := \frac{f(x+v) + f(x-v) - 2s}{v} \cdot \frac{v}{\sin \frac{v}{2}}$$

je integrovatelná na $(0,\delta)$.

Z Diniho kritéria lze snadno odvodit následující tvrzení, které pro „běžné funkce“ umožňuje určit součet její Fourierovy řady.

4.30 Důsledek. (důsledky Diniho kritéria)

(i) Nechť $f \in \mathscr{P}(2\pi)$ je reálná funkce, $x \in \mathbb{R}$, existují vlastní jednostranné limity $f(x+)$, $f(x_-)$ a také limity

$$\lim_{t\to x+} \frac{f(t) - f(x_+)}{t - x} \in \mathbb{R}, \qquad \lim_{t\to x-} \frac{f(t) - f(x_-)}{t - x} \in \mathbb{R}.$$

Pak Fourierova řada funkce f *v bodě* a *konverguje a její součet je roven* $\frac{1}{2}(f(x_+)+f(x_-))$.

Speciálně, pokud f *má konečné jednostranné derivace v bodě* x*, pak součet Fourierovy řady funkce* f *v bodě* x *je roven* $f(x)$.

(ii)

Nechť $f \in \mathscr{P}(2\pi)$, $x \in \mathbb{R}$ *a existují čísla* $\alpha > 0$, $\delta > 0$ *a* $K > 0$ *taková, že*

$$|f(x+v) - f(x)| \le K|v|^{\alpha}, \qquad \textit{pokud} \quad |v| < \delta.$$

Pak Fourierova řada funkce f *v bodě* a *konverguje a má součet* $f(x)$.

Důkaz. (i) Provedeme-li v limitách z (i) pořadě substituce $t = x+v$, $t = x - v$, snadno vidíme, že existuje číslo $\delta > 0$ takové, že funkce $(1/v)(f(x+v) - f(x+))$, $(1/v)(f(x-v) - f(x-))$ jsou omezené na $(0,\delta)$. Pro $s := (1/2)(f(x_+) + f(x_-))$ pak dostáváme, že (zřejmě měřitelná) funkce

$$\frac{f(x+v)+f(x-v)-2s}{v} = \frac{f(x+v)-f(x_+)}{v} + \frac{f(x-v)-f(x_-)}{v}$$

je omezená na $(0,\delta)$, takže platí (4.11).

(ii) Zvolme příslušná $\alpha > 0, \delta > 0\, K > 0$ a položme $s := f(x)$. Pak pro $0 < v < \delta$ platí

$$\frac{|f(x+v)+f(x-v)-2s|}{v} \le \frac{|f(x+v)-f(x)|}{v} + \frac{|f(x-v)-f(x)|}{v} \le 2Kv^{\alpha-1},$$

z čehož ihned plyne (4.11).

4.31 Příklad. Nechť g^* je funkce z Příkladu 4.19. Protože g^* má zřejmě konečné jednostranné derivace v každém bodě, podle Důsledku 4.30 (i) víme, že g^* je součtem své Fourierovy řady, takže platí

$$x^2 \;=\; \frac{\pi^2}{3} - 4\frac{\cos x}{1^2} + 4\frac{\cos 2x}{2^2} - 4\frac{\cos 3x}{3^2} + \dots, \qquad x \in [-\pi,\pi].$$

Diniho kritérium (se svými důsledky) se snadno dokáže a stačí pro obvyklé aplikace týkající se bodové konvergence. Následující Jordan–Dirichletovo kritérium je k důkazu podstatně obtížnější (i část týkající se bodové konvergence, srov. [J II]), takže je uvedeme bez důkazu.

4.32 Věta. (Jordan-Dirichletovo kritérium) *Nechť* $f \in \mathscr{P}(2\pi)$ *má konečnou variaci v intervalu* $[a,b]$*. Pak platí:*

(i) *V každém bodě* $x \in (a,b)$ *Fourierova řada funkce* f *konverguje a má součet* $\frac{1}{2}(f(x_+)+f(x_-))$.

(ii) *Je-li* f *navíc spojitá v* (a,b)*, konverguje Fourierova řada funkce* f *lokálně stejnoměrně na* (a,b) *(k funkci* f*).*

4.33 Poznámka.

(i) Funkce f má na $[a, b]$ konečnou variaci právě tehdy, když je rozdílem dvou neklesajících funkcí na $[a, b]$. Z toho ihned vyplývá, že v $x \in (a, b)$ existují konečné limity $f(x+), f(x-)$.

(ii) V Příkladu 4.19 je ovšem možno aplikovat i Jordan-Dirichletovo kritérium, ze kterého také snadno vyplývá stejnoměrná konvergence příslušné Fourierovy řady na $\mathbb{R}$. Dovedeme-li ovšem (jako v tomto příkladě) jednoduše spočítat Fourierovy koeficienty, není těžké stejnoměrnou konvergenci Fourierovy řady vyšetřit bez užití jakékoliv obecné věty o Fourierových řadách.

(iii) Existuje řada jiných postačujících podmínek pro bodovou a (lokálně) stejnoměrnou konvergenci Fourierovy řady.

4.34 Poznámka. (o bodové konvergenci) Existuje $f \in \mathscr{P}(2\pi)$, jejíž Fourierova řada diverguje v každém bodě. Fourierova řada spojité $f \in \mathscr{P}(2\pi)$ může divergovat v bodech nespočetné husté podmnožiny $\mathbb{R}$. Hluboký výsledek L. Carlesona (1966) však říká, že každá $f \in L^2(-\pi, \pi)$ (a tedy speciálně každá spojitá $f \in \mathscr{P}(2\pi)$) je součtem své Fourierovy řady ve skoro všech bodech.

4.2 Fejérova věta a její důsledky

Značný význam v analýze má teorie „sčítatelnosti divergentních řad“. Ukazuje se totiž, že pro některé divergentní řady lze přirozeným způsobem (pomocí některé „zobecněné sčítací metody“) definovat jejich „zobecněný součet“ a počítání s těmito zobecněnými součty přináší někdy zajímavé výsledky.

V teorii Fourierových řad prvně použil tuto myšlenku r. 1900 Fejér, když se zabýval sčítatelností Fourierových řad metodou aritmetických průměrů (které se také říká Cesarova sčítací metoda).

Připomeňme, že řada (komplexních) čísel $\sum_{i=0}^{\infty} a_i$ (indexujeme od 0, protože máme na mysli aplikaci na Fourierovy řady) je sčítatelná metodou aritmetických průměrů (Cesarovou metodou) k číslu σ, jestliže platí

$$\lim_{n\to\infty} \sigma_n = \sigma, \qquad \text{kde} \qquad s_n = a_0 + \cdots + a_n, \qquad \sigma_n = \frac{s_0 + \cdots + s_n}{n+1}.$$

Pak píšeme $\quad (C) \displaystyle\sum_{i=0}^{\infty} a_i = \sigma.$

Má-li řada součet s, víme, že je sčítatelná metodou aritmetických průměrů k s (viz [DII]), naopak to však neplatí (snadno se například ověří, že divergentní řada $\sum_{n=0}^{\infty} (-1)^n$ je sčítatelná metodou aritmetických průměrů k $1/2$).

Při vyšetřování sčítatelnosti Fourierových řad přebírá roli Dirichletova jádra tzv. Fejérovo jádro. Dříve než je definujeme a odvodíme jeho základní vlastnosti, potřebujeme dokázat, že pro $n \in \mathbb{N}$ a $v \neq k\pi$ $(k \in \mathbb{Z})$ platí rovnost

$$\sin v + \sin 3v + \cdots + \sin (2n+1)v = \frac{\sin^2 (n+1)v}{\sin v}. \tag{4.13}$$

Pro důkaz si uvědomíme, že výraz vlevo je součtem imaginárních částí konečné geometrické řady $e^{iv} + e^{i3v} + \cdots + e^{i(2n+1)v}$, jejíž součet je

$$e^{iv}\,\frac{1-e^{2i(n+1)v}}{1-e^{2iv}} = \frac{1-e^{2i(n+1)v}}{e^{-iv}-e^{iv}}$$

a jeho imaginární část je rovna

$$\frac{1-\cos 2(n+1)v}{2\sin v} = \frac{\sin^2 (n+1)v}{\sin v}.$$

4.35 Lemma. *Pro každé $n \in \mathbb{N} \cup \{0\}$ položme*

$$K_n(v) := \frac{1}{n+1}\sum_{j=0}^{n} D_j(v),$$

kde D_j je Dirichletovo jádro. Pro funkci K_n, kterou nazýváme Fejérovo jádro, platí:

(i) $$K_n(v) = \frac{1}{2(n+1)}\left(\frac{\sin (n+1)\frac{v}{2}}{\sin\frac{v}{2}}\right)^2, \quad v \neq 2k\pi \ (k \in \mathbb{Z}).$$

(ii) K_n *je funkce sudá, spojitá, 2π-periodická a* $K_n(0) = \dfrac{n+1}{2}$.

(iii) $$\int_{-\pi}^{\pi} K_n = \pi.$$

Důkaz. Platí

$$D_j(v) = \frac{1}{2}\,\frac{\sin (j+\frac{1}{2})v}{\sin\frac{v}{2}} = \frac{1}{2}\,\frac{\sin (2j+1)\frac{v}{2}}{\sin\frac{v}{2}}, \quad v \neq 2k\pi \ (k \in \mathbb{Z}).$$

Pomocí (4.13) tedy dostáváme

$$K_n(v) = \frac{1}{(n+1)2\sin\frac{v}{2}}\left(\sin\frac{v}{2} + \sin 3\frac{v}{2} + \cdots + \sin (2n+1)\frac{v}{2}\right) =$$
$$= \frac{1}{(n+1)2\sin\frac{v}{2}} \cdot \frac{\sin^2 (n+1)\frac{v}{2}}{\sin\frac{v}{2}} = \frac{1}{2(n+1)}\left(\frac{\sin (n+1)\frac{v}{2}}{\sin\frac{v}{2}}\right)^2, \quad v \neq 2k\pi \, (k \in \mathbb{Z}).$$

Z definice K_n a (i) snadno plyne (ii). Vlastnost (iii) ihned vyplývá z Lemmatu 4.20 (iii).

4.36 Poznámka. Fejérovo jádro má „lepší“ vlastnosti než Dirichletovo jádro. Jednou z nich je jeho nezápornost. Také snadno vidíme, že $K_n \rightrightarrows 0$ na $(-\pi,\pi) \setminus (-\delta,\delta)$ pro každé $0 < \delta < \pi$. Pro Dirichletovo jádro analogické tvrzení zřejmě neplatí ($D_n(\pi) = \pm 1/2$). Fejérovo jádro je (pro $n = 5$) znázorněno na obr. 4.17 vpravo.

Bude-li v dalším dána funkce $f \in \mathscr{P}(2\pi)$ a $x \in \mathbb{R}$, budeme symbolem $s_n(x)$, ($n = 0, 1, \dots$) rozumět n-tý částečný součet Fourierovy řady funkce f v bodě x a položíme

$$\sigma_n(x) := \frac{1}{n+1}(s_0(x) + \cdots + s_n(x)). \tag{4.14}$$

Pro zkoumání konvergence posloupnosti $(\sigma_n(x))$, tj. sčítatelnosti Fourierovy řady funkce f v bodě x metodou aritmetických průměrů, budeme potřebovat následující vyjádření $\sigma_n(x)$.

4.37 Lemma. *Nechť $f \in \mathscr{P}(2\pi)$. Pak*

$$\sigma_n(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+y)K_n(y)\,\mathrm{d}y = \frac{1}{\pi}\int_0^{\pi} (f(x+v) + f(x-v))\,K_n(v)\,\mathrm{d}v.$$

Důkaz. Dosadíme-li do (4.14) podle Lemmatu 4.21 a užijeme definici K_j, dostáváme

$$\sigma_n(x) = \frac{1}{n+1}\left(\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+y)D_0(y)\,\mathrm{d}y + \cdots + \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+y)D_n(y)\,\mathrm{d}y\right) =$$
$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)K_n(y)\,\mathrm{d}y.$$

Druhá rovnost se dokáže analogicky z Lemmatu 4.21 nebo užitím sudosti K_n.

4.38 Věta. (Fejérova) *Nechť $f \in \mathscr{P}(2\pi)$ a $x \in \mathbb{R}$. Potom platí:*

(i) Má-li f v bodě x konečné jednostranné limity $f(x+)$, $f(x_-)$, pak

$$\lim_{n\to\infty} \sigma_n(x) = \frac{1}{2}(f(x+) + f(x_-)).$$

(ii) Je-li f spojitá na intervalu $(a,b) \subset \mathbb{R}$, pak $\sigma_n(x) \to f(x)$, $x \in (a,b)$, a konvergence je lokálně stejnoměrná na (a,b).

Důkaz. (i) Položme $s := \frac{1}{2}(f(x+) + f(x_-))$. Pak z Lemmatu 4.35 (iii) máme

$$s = \frac{1}{\pi}\int_{-\pi}^{\pi} sK_n(v)\,\mathrm{d}v = \frac{1}{\pi}\int_0^{\pi} 2sK_n(v)\,\mathrm{d}v,$$

takže podle Lemmatu 4.37 dostáváme

$$\sigma_n(x) - s = \frac{1}{\pi} \int_0^\pi (f(x+v) + f(x-v) - 2s)\, K_n(v)\, \mathrm{d}v.$$

Nechť je dáno $\varepsilon > 0$. Protože zřejmě $\lim_{v \to 0+} (f(x+v) + f(x-v) - 2s) = 0$, můžeme zvolit $0 < \delta < \pi$ takové, že $|f(x+v) + f(x-v) - 2s| < \varepsilon/2$ pro každé $v \in (0, \delta)$. S pomocí Lemmatu 4.35 (i), (iii) dostáváme

$$\begin{aligned}|\sigma_n(x) - s| &\le \frac{1}{\pi} \int_0^\delta |f(x+v) + f(x-v) - 2s| \cdot K_n(v)\, \mathrm{d}v \\ &+ \frac{1}{\pi} \int_\delta^\pi |f(x+v) + f(x-v) - 2s| \cdot \frac{1}{2(n+1)} \left(\frac{\sin (n+1) \frac{v}{2}}{\sin \frac{v}{2}} \right)^2 \mathrm{d}v \le \\ &\le \frac{1}{\pi} \frac{\varepsilon}{2} \pi + \frac{1}{\pi} \frac{1}{2(n+1)} \frac{1}{\sin^2 \frac{\delta}{2}} \left(\int_0^\pi |f(x+v) + f(x-v) - 2s|\, \mathrm{d}v \right).\end{aligned}$$

Existuje tedy $n_0 \in \mathbb{N}$ takové, že $|\sigma_n(x) - s| < \varepsilon$ pro každé $n > n_0$.

(ii) Chceme dokázat, že $\sigma_n \rightrightarrows f$ na libovolném intervalu $[A, B] \subset (a, b)$. Zvolme $\omega > 0$, pro které $[A - \omega, A + \omega] \subset (a, b)$. Nechť je dáno $\varepsilon > 0$. Protože f je stejnoměrně spojitá na $[A - \omega, A + \omega]$, existuje $0 < \delta < \omega$ takové, že $|f(p) - f(q)| < \varepsilon/2$, kdykoliv $p, q \in [A - \omega, A + \omega]$. Pro každé $x \in [A, B]$ a každé $0 < v < \delta$ tedy platí

$$|f(x+v) + f(x-v) - 2f(x)| \le |f(x+v) - f(x)| + |f(x-v) - f(x)| < \varepsilon.$$

Pro libovolné $x \in [A, B]$ tedy dostáváme stejně jako v (i) (nyní ovšem $s = f(x)$):

$$\begin{aligned}|\sigma_n(x) - f(x)| &\le \frac{1}{\pi} \int_0^\delta |f(x+v) + f(x-v) - 2f(x)|\, K_n(v)\, \mathrm{d}v + \\ &+ \frac{1}{\pi} \int_\delta^\pi |f(x+v) + f(x-v) - 2f(x)| \frac{1}{2(n+1)} \left(\frac{\sin (n+1) \frac{v}{2}}{\sin \frac{v}{2}} \right)^2 \mathrm{d}v \le \\ &\le \frac{1}{\pi} \frac{\varepsilon}{2} \pi + \frac{1}{\pi} \frac{1}{2(n+1)} \frac{1}{\sin^2 \frac{\delta}{2}} \left(\int_0^\pi |f(x+v) + f(x-v) - 2f(x)|\, \mathrm{d}v \right).\end{aligned}$$

Abychom dostali odhad nezávislý na x, položíme $M := \max\{|f(x)| : x \in [A, B]\}$ a uvážíme, že podle Lemmatu 4.7 platí

$$\int_0^\pi |f(x+v)|\, \mathrm{d}v = \int_x^{x+\pi} |f(t)|\, \mathrm{d}t \le \int_x^{x+2\pi} |f(t)|\, \mathrm{d}t = \int_0^{2\pi} |f(t)|\, \mathrm{d}t$$

a (z podobného důvodu) $\int_0^\pi |f(x-v)|\, \mathrm{d}v \le \int_0^{2\pi} |f(t)|\, \mathrm{d}t$. Platí tedy

$$\begin{aligned}&\int_0^\pi |f(x+v) + f(x-v) - 2f(x)|\, \mathrm{d}v \le \\ &\le \int_0^{2\pi} |f(x+v)|\, \mathrm{d}v + \int_0^{2\pi} |f(x-v)|\, \mathrm{d}v + 2\pi M \le 2 \int_0^{2\pi} |f(t)|\, \mathrm{d}t + 2\pi M.\end{aligned}$$

Dostáváme

$$|\sigma_n(x) - f(x)| \leq \frac{\varepsilon}{2} + \frac{1}{\pi}\frac{1}{2(n+1)}\frac{1}{\sin^2\frac{\delta}{2}}\left(2\int_0^{2\pi}|f(t)|\,\mathrm{d}t + 2\pi M\right).$$

Existuje tedy $n_0 \in \mathbb{N}$ takové, že $|\sigma_n(x) - f(x)| < \varepsilon$ pro každé $x \in [A, B]$ a $n > n_0$.

Má-li funkce $f \in \mathscr{P}(2\pi)$ v bodě x konečné obě jednostranné limity, nemusí Fourierova řada funkce f v bodě x konvergovat (a to ani v případě, že f je spojitá na $\mathbb{R}$). Fejérova věta však ukazuje, že *„jediným kandidátem na součet Fourierovy řady v bodě x“ je číslo* $1/2\,(f(x_+) + f(x_-))$ (které je rovno $f(x)$, je-li f v x spojitá).

V tom spočívá jeden z významů Fejérovy věty.

4.39 Příklad. Protože funkce g^* z Příkladu 4.19 je spojitá na $\mathbb{R}$ a její Fourierova řada zřejmě konverguje ve všech bodech, dostáváme podle Fejérovy věty, že g^* je ve všech bodech součtem své Fourierovy řady. Při použití Fejérovy věty jsme ovšem museli nejdříve spočítat Fourierovy koeficienty (abychom zdůvodnili konvergenci Fourierovy řady), což při aplikaci Diniho kritéria (viz Příklad 4.31) není nutné.

Z Fejérovy věty také snadno dostáváme následující důležitou větu.

4.40 Věta. (Weirstrassovy věty)

(i) *Nechť $f \in \mathscr{P}(2\pi)$ je spojitá reálná funkce. Pak pro každé $\varepsilon > 0$ existuje reálný trigonometrický polynom T takový, že*

$$|f(x) - T(x)| < \varepsilon \qquad \text{pro všechna } x \in \mathbb{R}.$$

(ii) *Nechť f je reálná spojitá funkce v intervalu $[a, b]$. Pak pro každé $\varepsilon > 0$ existuje reálný polynom P takový, že*

$$|f(x) - P(x)| < \varepsilon \qquad \text{pro všechna } x \in [a, b].$$

Důkaz. (i) Aplikujeme-li (Fejérovu) Větu 4.38 (ii) na funkci f a interval $(a, b) := (-2\pi, 2\pi)$, dostáváme, že $\sigma_n \rightrightarrows f$ na $[-\pi, \pi]$, a (z periodicity) tedy i na $\mathbb{R}$. Protože σ_n jsou zřejmě trigonometrické polynomy, dostáváme (i).

(ii) Nejprve uvažujme případ, kdy $[a, b] \subset (-\pi, \pi)$. Nechť $\varepsilon > 0$ je dáno. Zřejmě existuje spojitá reálná funkce $g \in \mathscr{P}(2\pi)$, která rozšiřuje f. Aplikujeme-li na takovou funkci (i), vidíme, že existuje trigonometrický polynom T takový, že

$$|f(x) - T(x)| < \varepsilon/2 \qquad \text{pro všechna } x \in [a, b].$$

Protože funkce $\cos x$, $\sin x$ jsou na $\mathbb{R}$ součtem své Taylorovy řady v bodě 0, snadno dostáváme, že tuto vlastnost má i trigonometrický polynom T. Jeho Taylorova řada tedy konverguje stejnoměrně na $[a, b]$, takže existuje její částečný součet P_n takový, že

$$|T(x) - P_n(x)| < \varepsilon/2 \qquad \text{pro všechna } x \in [a, b].$$

Protože P_n je polynom, stačí položit $P:= P_n$.

Je-li $[a,b]$ obecný interval, zvolíme $c > 0$ tak malé, aby $[ca, cb] \subset (-\pi, \pi)$. Funkce $f^*(y) := f(y/c)$ je spojitá na $[ca, cb]$, takže podle předchozího existuje polynom P^* takový, že

$$|f(y/c) - P^*(y)| < \varepsilon \qquad \text{pro všechna } y \in [ca, cb].$$

Funkce $P(x) := P^*(cx)$ je zřejmě polynom a použijeme-li předchozí nerovnost pro $y = cx$, dostáváme

$$|f(x) - P(x)| < \varepsilon \qquad \text{pro všechna } x \in [a,b].$$

Věta ovšem platí i pro komplexní funkce reálné proměnné (aproximujeme zvlášť reálnou a imaginární část funkce). Obě tvrzení předchozí věty dokázal (zcela jiným způsobem) Weirstrass r. 1885. *Weirstrassova věta* (o aproximaci) je všeobecně přijatý název pro tvrzení (ii).

Pomocí Fejérovy věty dokážeme i následující větu.

4.41 Věta. (Fourierovy koeficienty určují funkci.) *Jestliže funkce g, $h \in \mathscr{P}(2\pi)$ mají stejné všechny Fourierovy koeficienty, pak $g(x) = h(x)$ pro skoro všechna $x \in \mathbb{R}$. Speciálně, má-li $f \in \mathscr{P}(2\pi)$ nulové všechny Fourierovy koeficienty, pak f je skoro všude nulová.*

Důkaz. (Náznak.) Důkaz stačí provést pro reálné funkce. Nechť g, $h \in \mathscr{P}(2\pi)$ mají stejné všechny Fourierovy koeficienty. Pak funkce $f:= g - h \in \mathscr{P}(2\pi)$ má všechny Fourierovy koeficienty nulové. Položme

$$D:= \{\varphi \in L^\infty(-\pi,\pi)\colon \int_{-\pi}^{\pi} f\cdot\varphi = 0\}.$$

Snadno vidíme, že

(A) D je lineární podprostor prostoru $L^\infty(-\pi,\pi)$

a z Lebesgueovy věty „o konvergenci s majorantou“ snadno vyplývá:

(B) Jestliže (φ_n) je stejně omezená posloupnost funkcí z D a $\varphi_n(x) \to \varphi(x)$ pro skoro všechna $x \in (-\pi,\pi)$, pak $\varphi \in D$.

Pomocí (A) a (B) není obtížné (s užitím Fejérovy věty pro (a) $\Longrightarrow$ (b)) postupně dokázat, že do množiny D patří následující reálné funkce na $(-\pi,\pi)$:

(a) Všechny trigonometrické polynomy.
(b) Spojité funkce f na $(-\pi,\pi)$, pro které $\overline{\{x\colon f(x) \neq 0\}} \subset (-\pi,\pi)$.
(c) Charakteristické funkce otevřených intervalů $(a,b) \subset (-\pi,\pi)$.
(d) Charakteristické funkce otevřených množin $G \subset (-\pi,\pi)$.
(e) Charakteristické funkce lebesgueovsky měřitelných množin $M \subset (-\pi,\pi)$.

Položme $M^+:= \{x \in (-\pi,\pi)\colon f(x) \geq 0\}$ a $M^-:= \{x \in (-\pi,\pi)\colon f(x) < 0\}$. Podle (e) platí

$$\int_{-\pi}^{\pi} |f| = \int_{-\pi}^{\pi} f\, C_{M^+} - \int_{-\pi}^{\pi} f\, C_{M^-} = 0,$$

takže $f = 0$ skoro všude na $(-\pi,\pi)$.

4.42 Poznámka. Nechť funkce $g \in \mathscr{P}(2\pi)$ má kosinovou Fourierovu řadu. Aplikujeme-li předchozí větu na funkci $f(x) := g(x) - g(-x)$, snadno dostáváme, že g se skoro všude rovná nějaké sudé funkci.

4.43 Příklad. Nechť g^* je funkce z Příkladu 4.19. Její Fourierovu řadu jsme spočítali; snadno vidíme, že stejnoměrně konverguje k nějaké (spojité) funkci $h \in \mathscr{P}(2\pi)$. Podle Tvrzení 4.10 mají g^* a h stejné Fourierovy koeficienty, takže se podle Věty 4.41 rovnají ve skoro všech bodech. Protože jsou obě spojité, rovnají se všude. Opět jsme dokázali, že g^* je součtem své Fourierovy řady.

4.3 Fourierovy řady v Hilbertově prostoru

Velmi důležitou vlastností trigonometrického systému je jeho ortogonalita; z ní byly odvozeny důležité klasické výsledky jako je Besselova nerovnost a Parsevalova rovnost a později, po zavedení Lebesgueova integrálu, tzv. L^2-teorie Fourierových řad. V této teorii se uvažují funkce $f \in L^2(-\pi, \pi)$ (ty se dají rozšířit – až na množinu míry nula jednoznačně – na funkci z $\mathscr{P}(2\pi)$) a *nezkoumá se bodová* konvergence její Fourierovy řady, ale konvergence této řady *v prostoru* $L^2(-\pi, \pi)$. Ukazuje se, že jejím „L^2 - součtem“ je vždy funkce f: pro její částečné součty platí

$$s_n \to f \quad \text{v} \quad L^2(-\pi, \pi), \quad \text{tj.} \quad \lim_{n\to\infty} \int_{-\pi}^{\pi} |f(x) - s_n(x)|^2 \, \mathrm{d}x = 0.$$

Tato L^2-teorie klasických Fourierových řad byla přirozeným způsobem zobecněna na případ „abstraktních Fourierových řad“ vzhledem k ortogonálním systémům v obecném separabilním Hilbertově prostoru. Přitom důkazy jsou stejně obtížné jako v klasickém případě a teorii lze použít i na jiné ortogonální systémy funkcí, než je trigonometrický systém. V [J II] je „L^2-teorie“ vyložena ve speciálním případě ortogonálních systémů funkcí v $L^2(\mu)$ pro některé míry μ na $\mathbb{R}^n$. Tam lze také najít důležité klasické příklady ortogonálních systémů, které tvoří polynomy (Legendrovy, Čebyševovy, Jacobiho, Hermiteovy a Laguerreovy).

Teorii lze zobecnit i na případ neseparabilního Hilbertova prostoru; my se zde však omezíme jen na jednodušší (a nejdůležitější) separabilní případ.

Protože úplnost není v řadě úvah podstatná, budeme uvažovat obecný unitární prostor X nad tělesem $\mathbb{T}$, kde $\mathbb{T} = \mathbb{R}$ nebo $\mathbb{T} = \mathbb{C}$.

Z Cauchyovy nerovnosti snadno vyplývá spojitost skalárního součinu:

4.44 Tvrzení. *Nechť X je unitární prostor. Pak skalární součin $(x, y) \mapsto \langle x, y \rangle$ je spojité zobrazení prostoru $X \times X$ do T.*

Důkaz. Předpokládejme, že je dána posloupnost $\{(x_n, y_n)\}_{n=1}^{\infty}$ prvků $X \times X$, která konverguje k $(x, y) \in X \times X$. Chceme dokázat, že $\langle x_n, y_n \rangle \to \langle x, y \rangle$. Označíme-li $h_n := x_n - x$, $k_n := y_n - y$, máme $\|h_n\| \to 0, \|k_n\| \to 0$. Protože

$$|\langle x_n, y_n \rangle - \langle x, y \rangle| = |\langle x + h_n, y + k_n \rangle - \langle x, y \rangle| = |\langle x, k_n \rangle + \langle h_n, y \rangle + \langle h_n, k_n \rangle| \leq$$

$$\leq |\langle x, k_n \rangle| + |\langle h_n, y \rangle| + |\langle h_n, k_n \rangle| \leq \|x\| \cdot \|k_n\| + \|h_n\| \cdot \|y\| + \|h_n\| \cdot \|k_n\|,$$

dostáváme $|\langle x_n, y_n \rangle - \langle x, y \rangle| \to 0$.

Je-li X reálný prostor, pak skalární součin je bilineární forma na X, takže z Věty 1.145 snadno vyplývá, že skalární součin je dokonce lipschitzovský na každé omezené podmnožině $X \times X$. To lze snadno dokázat i v případě komplexního prostoru, kdy skalární součin není bilineární forma.

Jestliže pro prvky $x \in X, y \in X$ platí $\langle x, y \rangle = 0$, říkáme, že x, y jsou na sebe kolmé (ortogonální) a píšeme $x \perp y$. Zřejmě $x \perp x$ pouze pro $x = 0$. Ze spojitosti skalárního součinu okamžitě vyplývá následující tvrzení (o prvcích X):

$$(x_n \perp y_n, \quad x_n \to x, \quad y_n \to y) \quad \Longrightarrow \quad x \perp y. \tag{4.15}$$

Pro libovolnou množinu $M \subset X$ definujeme $M^{\perp} := \{x \in X : y \in M \implies x \perp y\}$. Je zřejmé, že $M^{\perp}$ je lineární podprostor X; použijeme-li (4.15), snadno dostáváme, že

$$M^{\perp} \text{ je uzavřený lineární podprostor } X \text{ a } M^{\perp} = \overline{\operatorname{Lin} M}^{\perp}. \tag{4.16}$$

4.45 Definice. *Nechť X je unitární prostor a $(x_\alpha)_{\alpha \in A}$ je systém (s libovolnou indexovou množinou A) prvků prostoru X. Řekneme, že systém $(x_\alpha)_{\alpha \in A}$ je ortogonální, jestliže $x_\alpha \perp x_\beta$, kdykoliv $\alpha \neq \beta$. Jestliže navíc $\|x_\alpha\| = 1$ pro každé $\alpha \in A$, říkáme, že systém $(x_\alpha)_{\alpha \in A}$ je ortonormální.*

4.46 Poznámka. Je-li $(x_\alpha)_{\alpha \in A}$ ortogonální systém nenulových prvků unitárního prostoru X, pak $(x_\alpha / \|x_\alpha\|)_{\alpha \in A}$ je zřejmě ortonormální systém.

Následující tvrzení ukazuje, že v separabilních prostorech stačí zkoumat ortogonální posloupnosti (konečné nebo spočetné).

4.47 Tvrzení. *Nechť X je separabilní unitární prostor a $(x_\alpha)_{\alpha \in A}$ je ortogonální systém nenulových prvků prostoru X. Pak A je spočetná množina.*

Důkaz. Podle Poznámky 4.46 můžeme předpokládat, že jde o ortonormální systém. Pro α, $\beta \in A$, $\alpha \neq \beta$, pak platí

$$\|x_\alpha - x_\beta\| = \sqrt{\langle x_\alpha - x_\beta, x_\alpha - x_\beta \rangle} = \sqrt{\|x_\alpha\|^2 + \|x_\beta\|^2} = \sqrt{2}.$$

Podle Tvrzení 1.73 je $\{x_\alpha : \alpha \in A\}$, a tedy i A, spočetná množina.

4.48 Příklad. Nechť $\alpha \in \mathbb{R}$ a $X := L^2(\alpha, \alpha + 2\pi)$ je komplexní (nebo reálný) unitární prostor. Pak reálný trigonometrický systém

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x, \ldots$$

je podle Lemmatu 4.9 (i) a Lemmatu 4.7 ortogonální systém v X. (Přesněji: jde o systém $v_1 = 1 \restriction_{(\alpha,\alpha+2\pi)}$, $v_2 = \cos \restriction_{(\alpha,\alpha+2\pi)}, \ldots,$ kde ovšem funkce obvyklým způsobem chápeme jako příslušné třídy funkcí.)

Podle Lemmatu 4.9 (ii) je („normovaný reálný trigonometrický systém")

$$\frac{1}{\sqrt{2\pi}},\ \frac{\cos x}{\sqrt{\pi}},\ \frac{\sin x}{\sqrt{\pi}},\ \frac{\cos 2x}{\sqrt{\pi}},\ \frac{\sin 2x}{\sqrt{\pi}}, \ldots$$

ortonormální systém v X.

V komplexním X je komplexní trigonometrický systém $\left(e^{inx}\right)_{n=-\infty}^{\infty} := \left(e^{inx}\right)_{n\in\mathbb{Z}}$ podle Lemmatu 4.8 ortogonální systém a $\left(e^{inx}/\sqrt{2\pi}\right)_{n\in\mathbb{Z}}$ je ortonormální systém v X.

V normovaných lineárních prostorech definujeme přirozeným způsobem pojem součtu nekonečné řady a pomocí něj pojem (Schauderovy) báze, který má však *odlišný význam* než pojem báze v *lineární algebře*, protože při definici (Schauderovy) báze se připouštějí i *„nekonečné lineární kombinace"*.

Nás zde budou zajímat pouze *ortogonální báze* v unitárním prostoru.

4.49 Definice. *Nechť X je normovaný lineární prostor. Pak (nekonečnou) řadou v X rozumíme symbol*

$$\sum_{n=1}^{\infty} x_n, \qquad \text{kde } x_n \in X,\ n = 1, 2, \ldots .$$

Říkáme, že součtem této řady je $s \in X$ a píšeme $\displaystyle\sum_{n=1}^{\infty} x_n = s,$ *jestliže* $\displaystyle\lim_{n\to\infty} \sum_{k=1}^{n} x_k = s.$

Má-li řada součet, říkáme, že je konvergentní; jinak je divergentní.

4.50 Definice. *Nechť X je normovaný lineární prostor (nad $\mathbb{T}$). Posloupnost $(x_n)_{n=1}^{\infty}$ prvků z X se nazývá (Schauderovou) bazí prostoru X, jestliže pro každý bod $x \in X$ existuje právě jedna posloupnost $(c_n)_{n=1}^{\infty}$ prvků $\mathbb{T}$, pro kterou platí*

$$x = \sum_{n=1}^{\infty} c_n x_n .$$

Číslům c_n říkáme souřadnice bodu x vzhledem k bázi (x_n).

4.51 Poznámka.

(i) Každý prvek prostoru X se dá tedy právě jedním způsobem vyjádřit jako „nekonečná lineární kombinace" prvků ze Schauderovy báze. Pojem Schauderovy báze se většinou definuje pouze v Banachových prostorech.

(ii) Je zřejmé, že báze $(x_n)_{n=1}^{\infty}$ je lineárně nezávislý systém (tj. $x_i \neq x_j$ pro $i \neq j$ a množina $\{x_1, x_2, \ldots\}$ je lineárně nezávislá). Obrácená implikace však neplatí.

(iii) Je snadné dokázat, že pokud normovaný lineární prostor má Schauderovu bázi, je nutně separabilní a nemá konečnou dimenzi.
(iv) Je zřejmé, že pokud (α_n) je posloupnost nenulových prvků z $\mathbb{T}$, pak $(x_n)_{n=1}^{\infty}$ je báze, právě když $(\alpha_n\, x_n)_{n=1}^{\infty}$ je báze.

Dále budeme potřebovat následující dva pojmy.

4.52 Definice. *Nechť $(x_\alpha)_{\alpha\in A}$ je ortogonální systém v unitárním prostoru X.*

(i) Řekneme, že systém $(x_\alpha)_{\alpha\in A}$ je úplný, jestliže jeho lineární obal je hustý v X, tj.

$$\overline{\operatorname{Lin}\{x_\alpha : \alpha\in A\}} = X.$$

(ii) Řekneme, že ortogonální systém $(x_\alpha)_{\alpha\in A}$ je maximální, jestliže neexistuje $0\neq u\in X$, který by byl kolmý na všechny vektory x_α, $\alpha\in A$.

4.53 Poznámka.

(a) Podmínka (ii) říká, že k systému $(x_\alpha)_{\alpha\in A}$ již nelze přidat nenulový prvek tak, aby výsledkem byl opět ortogonální systém.
(b) V literatuře se používá také pojem „totální ortogonální systém“, a to někdy pro úplný systém a jindy pro maximální systém. Také „úplný systém“ někdy označuje maximální systém (a pro úplný systém se užívá termín „uzavřený systém“). Toto kolísání terminologie bylo zřejmě způsobeno tím, že v (klasickém) případě Hilbertova prostoru jsou podmínky (i) a (ii) ekvivalentní (což zde dokážeme pouze v separabilním případě).

První část následujícího tvrzení lze chápat jako zobecnění Pythagorovy věty.

4.54 Tvrzení. (O součtu ortogonální řady.) *Nechť $(x_n)_{n=1}^{\infty}$ je ortogonální posloupnost v unitárním prostoru X. Potom platí:*

(i) Je-li řada $\sum_{n=1}^{\infty} x_n$ konvergentní, pak konverguje i řada $\sum_{n=1}^{\infty}\|x_n\|^2$ a

$$\left\|\sum_{n=1}^{\infty} x_n\right\|^2 = \sum_{n=1}^{\infty}\|x_n\|^2.$$

(ii) Je-li X Hilbertův prostor, pak

$$\sum_{n=1}^{\infty} x_n \ \text{ konverguje } \iff \sum_{n=1}^{\infty}\|x_n\|^2<\infty.$$

Důkaz. (i) Označme $s_n := x_1+\cdots+x_n$. Z ortogonality posloupnosti (x_n) plyne

$$\|s_n\|^2 = \langle x_1+\cdots+x_n, x_1+\cdots+x_n\rangle = \|x_1\|^2+\cdots+\|x_n\|^2.$$

Protože $s_n\to\sum_{k=1}^{\infty} x_k$, ze spojitosti funkce $x\mapsto\|x\|^2$ vyplývá

$$\|x_1\|^2+\cdots+\|x_n\|^2 = \|s_n\|^2 \to \left\|\sum_{k=1}^{\infty} x_k\right\|^2,$$

takže tvrzení (i) je dokázáno.

(ii) Předpokládejme, že $\sum_{n=1}^{\infty} \|x_n\|^2 < \infty$. Pro libovolné $\varepsilon > 0$ pak můžeme najít $n_0 \in \mathbb{N}$, pro které $\sum_{k=n_0+1}^{\infty} \|x_k\|^2 < \varepsilon^2$. Pak pro $m > n \geq n_0$ platí

$$\|s_m - s_n\|^2 = \|x_{n+1} + \cdots + x_m\|^2 = \|x_{n+1}\|^2 + \cdots + \|x_m\|^2 < \varepsilon^2,$$

takže $\|s_n - s_m\| < \varepsilon$. Dokázali jsme, že posloupnost (s_n) je caychyovská; protože X je úplný, je také konvergentní. Obrácená implikace již byla dokázána v (i).

4.55 Poznámka.

(a) Nechť (x_n) je ortogonální a řada $\sum_{n=1}^{\infty} x_n$ konverguje. Tato řada nemusí konvergovat absolutně, tj. nemusí platit $\sum_{n=1}^{\infty} |x_n| < \infty$. To vidíme na jednoduchém příkladě řady $\sum (1/n) e_n$ v reálném prostoru ℓ^2, kde $e_n = (0, \dots, 0, 1, 0, \dots)$ (1 stojí na n-tém místě). Tato řada konverguje podle (ii) a jejím součtem je zřejmě $(1, 1/2, 1/3, \dots)$.

(b) Modifikací důkazu (ii) není obtížné ukázat, že řada $\sum_{n=1}^{\infty} x_n$ konverguje bezpodmínečně, tj. konverguje po libovolném přerovnání ke stejnému součtu.

Následující tvrzení je zcela analogické (i s důkazem) Tvrzení 4.10, které motivovalo zavedení „klasických Fourierových koeficientů“.

4.56 Tvrzení. *Nechť $(v_n)_{n=1}^{\infty}$ je ortogonální posloupnost nenulových prvků v unitárním prostoru X. Pak (pro $x \in X$ a $c_n \in \mathbb{T}$) platí implikace*

$$x = \sum_{n=1}^{\infty} c_n v_n \quad \Longrightarrow \quad c_n = \frac{\langle x, v_n \rangle}{\|v_n\|^2}, \; n \in \mathbb{N}. \tag{4.17}$$

Důkaz. Označme $s_p := \sum_{k=1}^{p} c_k v_k$ a zvolme $n \in \mathbb{N}$. Pak pro $p \geq n$ z ortogonality (v_n) plyne

$$\langle s_p, v_n \rangle = \langle c_1 v_1 + \cdots + c_p v_p, v_n \rangle = \langle c_n v_n, v_n \rangle = c_n \|v_n\|^2.$$

Ze spojitosti skalárního součinu dostáváme

$$\langle x, v_n \rangle = \lim_{p \to \infty} \langle s_p, v_n \rangle = c_n \|v_n\|^2,$$

z čehož již okamžitě vyplývá (4.17).

Předcházející tvrzení motivuje následující definici, která rozšiřuje známou definici z lineární algebry (kde se definují Fourierovy koeficienty vzhledem k ortogonální bázi konečně rozměrného unitárního prostoru).

4.57 Definice. *Nechť X je unitární prostor, $(v_\alpha)_{\alpha \in A}$ je ortogonální systém nenulových prvků prostoru X a $x \in X$. Pak definujeme Fourierovy koeficienty vektoru x vzhledem k systému $(v_\alpha)_{\alpha \in A}$ předpisem*

$$c_\alpha := \frac{\langle x, v_\alpha \rangle}{\|v_\alpha\|^2}, \qquad \alpha \in A. \tag{4.18}$$

V případě $A = \mathbb{N}$ řadu $\sum_{n=1}^{\infty} c_n v_n$ nazýváme Fourierovou řadou bodu x vzhledem k systému $(v_n)_{n=1}^{\infty}$.

Pokud posloupnost $(v_n)_{n=1}^\infty$ je ortonormální, platí ovšem $c_n = \langle x, v_n\rangle$.

Z Tvrzení 4.56 okamžitě plyne následující tvrzení.

4.58 Důsledek. *Souřadnice bodu vzhledem k ortogonální bázi (v_n) jsou jeho Fourierovy koeficienty vzhledem k (v_n).*

4.59 Poznámka. Nechť $(v_n)_{n=1}^\infty$ je ortogonální posloupnost nenulových prvků unitárního prostoru X a $u_n := v_n/\|v_n\|$, $n \in \mathbb{N}$ je příslušná ortonormální posloupnost. Pak Fourierovy řady libovolného $x \in X$ vzhledem k $(v_n)_{n=1}^\infty$ a $(u_n)_{n=1}^\infty$ se rovnají. To je vidět ihned z rovností

$$\langle x, u_n\rangle\, u_n = \left\langle x, \frac{v_n}{\|v_n\|}\right\rangle \frac{v_n}{\|v_n\|} = \frac{\langle x, v_n\rangle}{\|v_n\|^2}\cdot v_n .$$

4.60 Poznámka. Z Lemmatu 4.9 snadno vidíme, že „abstraktní“ Definice 4.57 Fourierových koeficientů v případě $X = L^2(\alpha, \alpha+2\pi)$, $f \in X$ a trigonometrického (reálného nebo komplexního) systému splývá s klasickou definicí Fourierových koeficientů. Přesněji; je-li $f \in \mathscr{P}(2\pi)$ a funkce $|f|^2$ je lokálně integrovatelná na $\mathbb{R}$, pak posloupnost „klasických“ Fourierových koeficientů

$$a_0,\, a_1,\, b_1,\, a_2,\, b_2,\, \ldots$$

je posloupností Fourierových koeficientů ve smyslu Definice 4.57 funkce $f\restriction_{(\alpha,\alpha+2\pi)} \in X$ vzhledem k (nepatrně změněnému) reálnému trigonometrickému systému

$$1/2,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x, \ldots$$

a systém $(c_n)_{n\in\mathbb{Z}}$ „klasických“ komplexních Fourierových koeficientů (viz str. 183) je systémem Fourierových koeficientů ve smyslu Definice 4.57 funkce $f\restriction_{(\alpha,\alpha+2\pi)} \in X$ vzhledem ke komplexnímu systému $(e^{inx})_{n\in Z}$.

4.61 Věta. *Nechť X je unitární prostor, $(u_n)_{n=1}^\infty$ je ortonormální systém v X, $x \in X$ a $(c_n)_{n=1}^\infty$ jsou jeho Fourierovy koeficienty vzhledem k (u_n). Pak platí následující tvrzení:*

(i)
$$\sum_{n=1}^\infty |c_n|^2 \le \|x\|^2 .$$

(ii)
$$x = \sum_{n=1}^\infty c_n u_n \iff \|x\|^2 = \sum_{n=1}^\infty |c_n|^2 .$$

(iii) Je-li X Hilbertův, pak řada $\sum_{k=1}^\infty c_k u_k$ konverguje a

$$x - \sum_{k=1}^\infty c_k u_k \perp u_n, \quad n \in \mathbb{N}.$$

Důkaz. (i) Položme $z_n := x - \sum_{k=1}^n c_k u_k$. Pak pro $1 \le j \le n$ platí

$$\langle z_n, u_j\rangle = \langle x - \sum_{k=1}^n c_k u_k, u_j\rangle = \langle x, u_j\rangle - c_j\langle u_j, u_j\rangle = 0, \tag{4.19}$$

takže $z_n \perp u_j$. Konečná posloupnost $c_1u_1, \dots, c_nu_n, z_n$ je tedy ortogonální systém, takže (lze užít Tvrzení 4.54 (i))

$$\|x\|^2 = \|c_1u_1 + \dots + c_nu_n + z_n\|^2 = \sum_{k=1}^{n} |c_k|^2 + \|z_n\|^2, \tag{4.20}$$

z čehož snadno vyplývá (i).

(ii) Protože $\|c_nu_n\|^2 = |c_n|^2$, implikace „$\Longrightarrow$“ plyne ihned z Tvrzení 4.54 (i).

Platí-li $\|x\|^2 = \sum_{n=1}^{\infty} |c_n|^2$, pak podle (4.20) dostáváme

$$\|z_n\|^2 = \|x\|^2 - \sum_{k=1}^{n} |c_k|^2 \to 0,$$

takže $z_n \to 0$, a tedy $x = \sum_{n=1}^{\infty} c_nu_n$.

(iii) Podle (i) a Tvrzení 4.54 (ii) řada $\sum_{k=1}^{\infty} c_ku_k$ konverguje. Položíme-li $z := x - \sum_{k=1}^{\infty} c_ku_k$, zřejmě $z_n \to z$. Nechť $j \in \mathbb{N}$. Podle (4.19) platí $z_n \perp u_j$ pro každé $n \geq j$; takže také $z \perp u_j$ (viz (4.15)).

Nerovnosti z (i) se říká *Besselova nerovnost* a rovnost na pravé straně (ii) se nazývá *Parsevalova rovnost*.

4.62 Poznámka. Nechť $(v_n)_{n=1}^{\infty}$ je ortogonální posloupnost nenulových prvků unitárního prostoru X, $u_n := v_n/\|v_n\|$ je příslušná ortonormální posloupnost, $x \in X$ a d_n (resp. c_n) jsou Fourierovy koeficienty x vzhledem k (v_n) (resp. (u_n)). Podle Poznámky 4.59 platí $d_n\, v_n = c_n\, u_n$, takže

$$|c_n|^2 = \|c_n\, u_n\|^2 = \|d_n\, v_n\|^2 = d_n^2\, \|v_n\|^2.$$

Platí tedy (Besselova nerovnost)

$$\sum_{n=1}^{\infty} |d_n|^2\, \|v_n\|^2 \leq \|x\|^2$$

a Parsevalova rovnost

$$\sum_{n=1}^{\infty} |d_n|^2\, \|v_n\|^2 = \|x\|^2$$

platí, právě když $x = \sum_{n=1}^{\infty} d_n\, v_n$.

Tvrzení (iii) z Věty 4.61 platí (podle Poznámky 4.59) zřejmě i v případě, kdy (u_n) je otogonální posloupnost nenulových prvků.

4.63 Věta. (o ortogonální bázi) *Nechť X je Hilbertův prostor a $(v_n)_{n=1}^{\infty}$ je ortogonální systém nenulových prvků prostoru X. Pak následující tvrzení jsou ekvivalentní.*

(i) $(v_n)_{n=1}^{\infty}$ je ortogonální (Schauderova) báze v X.
(ii) $(v_n)_{n=1}^{\infty}$ je úplný otogonální systém.
(iii) $(v_n)_{n=1}^{\infty}$ je maximální ortogonální systém.

Důkaz. Bez újmy na obecnosti můžeme předpokládat, že $(v_n)_{n=1}^{\infty}$ je ortonormální. (Položíme-li totiž $v_n^* := v_n/\|v_n\|$, pak každá z podmínek (i), (ii), (iii) platí právě tehdy, když platí příslušná podmínka pro ortonormální systém (v_n^*).)

Nechť platí (i). Každý prvek $x \in X$ pak lze psát ve tvaru $x = \sum_{n=1}^{\infty} c_n v_n$. Protože $s_n := c_1 v_1 + \dots c_n v_n \in \operatorname{Lin}\{v_1,\ v_2, \dots\}$ a $s_n \to x$, dostáváme (ii).

Nechť platí (ii) a nechť $u \perp v_n$, $n \in \mathbb{N}$. Množina $M := \{u\}^{\perp}$ je podle (4.16) uzavřený lineární podprostor X a $v_n \in M$, $n \in \mathbb{N}$. Z (ii) tedy ihned plyne $M = X$, takže $u \perp u$, a tudíž $u = 0$. Dokázali jsme (ii).

Nechť platí (iii), $x \in X$ a c_n jsou Fourierovy koeficienty bodu x vzhledem k (v_n). Podle Věty 4.61 (iii) dostáváme, že řada $\sum_{n=1}^{\infty} c_n v_n$ konverguje a $z := x - \sum_{n=1}^{\infty} c_n v_n$ je kolmý na všechny v_n, $n \in \mathbb{N}$. Z (iii) tedy vyplývá $z = 0$, tj. $x = \sum_{n=1}^{\infty} c_n v_n$. Použijeme-li ještě Tvrzení 4.56, dostáváme (i).

4.64 Poznámka. Nepředpokládáme-li úplnost unitárního prostoru X, jsou podmínky (i) a (ii) ekvivalentní a implikují (iii) (srov. Poznámka 4.73). Podmínka (iii) však obecně neimplikuje (i) a (ii).

4.65 Tvrzení. *Nechť $(u_n)_{n=1}^{\infty}$ je ortonormální systém v unitárním prostoru X, $x, y \in X$ a platí*

$$x = \sum_{n=1}^{\infty} c_n u_n, \quad y = \sum_{n=1}^{\infty} d_n u_n.$$

Pak

$$\langle x, y\rangle = \sum_{n=1}^{\infty} c_n \overline{d_n}.$$

Důkaz. Pomocí spojitosti skalárního součinu dostáváme

$$\begin{aligned} \langle x, y\rangle = \left\langle \sum_{n=1}^{\infty} c_n u_n, \sum_{n=1}^{\infty} d_n u_n \right\rangle &= \lim_{n\to\infty} \langle \sum_{k=1}^{n} c_k u_k, \sum_{k=1}^{n} d_k u_k \rangle \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} \langle c_k u_k, d_k u_k\rangle = \sum_{n=1}^{\infty} c_k \overline{d_k}. \end{aligned}$$

4.66 Věta. *Nechť X je reálný (resp. komplexní) Hilbertův prostor a $(u_n)_{n=1}^{\infty}$ je jeho ortonormální báze. Pak zobrazení*

$$F : x \mapsto (c_1, c_2, \dots), \quad c_n = \langle x, u_n\rangle,$$

je unitární bijekce prostoru X na reálný (resp. komplexní) Hilbertův prostor ℓ^2, Inverzní zobrazení je dáno vzorcem

$$F^{-1}(c_1, c_2, \dots) = \sum_{n=1}^{\infty} c_n u_n. \tag{4.21}$$

Důkaz. Z Besselovy nerovnosti (Věta 4.61 (i)) máme $F(x) \in \ell^2$. Linearita zobrazení F je zřejmá. Je-li $F(x) = (c_1, c_2, \dots)$, pak podle Důsledku 4.58 platí $x = \sum_{n=1}^{\infty} c_n u_n$, takže zobrazení F je prosté.

Je-li $(c_n) \in \ell^2$, pak ortogonální řada $\sum_{n=1}^{\infty} c_n u_n$ konverguje podle Tvrzení 4.54 (ii). Pro $x := \sum_{n=1}^{\infty} c_n u_n$ podle Tvrzení 4.56 platí $F(x) = (c_n)$. Zobrazení F je tedy bijekce a inverzní zobrazení F^{-1} má tvar (4.21).

Jestliže $x, y \in X$, $F(x) = (c_n)$ a $F(y) = (d_n)$, pak $x = \sum_{n=1}^{\infty} c_n u_n$, $y = \sum_{n=1}^{\infty} d_n u_n$, takže podle Tvrzení 4.65 $\langle F(x), F(y)\rangle = \sum_{n=1}^{\infty} c_n \overline{d_n} = \langle x, y\rangle$.

4.67 Tvrzení. *Nechť X je nekonečně dimenzionální separabilní Hilbertův prostor. Pak v X existuje ortonormální báze.*

Důkaz. (Náznak.) Protože X je separabilní, můžeme najít posloupnost $(x_n)_1^{\infty}$, která je hustá v X. Indukcí snadno sestrojíme vybranou lineárně nezávislou posloupnost $(x_{n_k})_{k=1}^{\infty}$ takovou, že $\mathrm{Lin}\{x_{n_1}, \dots, x_{n_k}\} = \mathrm{Lin}\{x_1, x_2, \dots, x_{n_k}\}$. Použijeme-li na posloupnost $w_k := x_{n_k}$ známý Gram–Schmidtův ortogonalizační proces (srov. [Be], str. 376), dostaneme ortogonální posloupnost $(v_k)_{k=1}^{\infty}$, přičemž všechny uvažované posloupnosti mají stejný lineární obal, který je hustý v X. Podle Věty 4.63 je tedy $(v_k)_{k=1}^{\infty}$ ortogonální báze. „Znormováním“ dostaneme ortonormální bázi.

Z předchozího tvrzení a Věty 4.66 okamžitě dostáváme následující důsledek.

4.68 Důsledek. *Nechť X_1, X_2 jsou nekonečně dimenzionální separabilní Hilbertovy prostory. Pak X_1 a X_2 jsou izometricky izomorfní.*

Nakonec se zmíníme o významu Fourierových řad z hlediska teorie aproximace (který je částečně znám již z lineární algebry, srov. [Be, str. 374]).

4.69 Definice. *Nechť (X, ρ) je metrický prostor, $Y \subset X$ a $x \in X$. Řekneme, že $y \in Y$ je nejlepší aproximace bodu x v množině Y (nejbližší bod k x v množině Y), jestliže $\rho(x, y) = \mathrm{dist}(x, Y)$.*

4.70 Poznámka. Poznamenejme, že množina (bodů nejlepší aproximace, nejbližších bodů) $P_Y(x) := \{y \in Y : \rho(x, y) = \mathrm{dist}(x, Y)\}$ se někdy nazývá *metrická projekce* prvku x na množinu Y.

4.71 Definice. *Nechť X je unitární prostor, $Y \subset X$ je jeho lineární podprostor a $x \in X$. Řekneme, že $y \in Y$ je ortogonální projekce bodu x na prostor Y, jestliže platí $x - y \in Y^{\perp}$.*

Snadno je vidět, že pokud y je ortogonální projekce bodu x na prostor Y, pak y je *jediná* nejlepší aproximace bodu x v prostoru Y. Pokud je totiž $y^* \in Y$ a $y \neq y^*$, pak $y - y^* \in Y$, takže $y - y^* \perp x - y$ a z Pythagorovy věty dostáváme

$$\|x - y^*\|^2 = \|(x - y) + (y - y^*)\|^2 = \|x - y\|^2 + \|y - y^*\|^2,$$

a tedy $\|x-y^*\| > \|x-y\|$.

V důkazu Věty 4.61 jsme fakticky dokázali (viz (4.19) a (iii)) následující tvrzení. (Tam jsme sice pracovali s ortonormální posloupností, to však ale podle Poznámky 4.59 nevadí.)

4.72 Tvrzení. *Nechť X je unitární prostor, $(v_k)_{k=1}^{\infty}$ je ortogonální posloupnost nenulových prvků X, $x \in X$ a $\sum_{k=1}^{\infty} c_k\, v_k$ je jeho Fourierova řada vzhledem k (v_k). Potom platí:*

(a) *Pro každé $n \in \mathbb{N}$ je $\sum_{k=1}^{n} c_k\, v_k$ ortogonální projekce bodu x na prostor $\operatorname{Lin}\{v_1,\dots,v_n\}$ a tedy také nejlepší aproximace bodu x v $\operatorname{Lin}\{v_1,\dots,v_n\}$.*

(b) *Je-li X úplný, pak $\sum_{k=1}^{\infty} c_k\, v_k$ je ortogonální projekce bodu x na prostor $\overline{\operatorname{Lin}\{v_1,v_2,\dots\}}$ a tedy také nejlepší aproximace bodu x v $\overline{\operatorname{Lin}\{v_1,v_2,\dots\}}$.*

4.73 Poznámka.

(i) Podle Důsledku 1.141 a Tvrzení 1.82 je $\operatorname{Lin}\{v_1,\dots,v_n\}$ uzavřený lineární podprostor X. Je snadné dokázat, že uzávěr lineárního podprostoru V normovaného lineárního prostoru X je opět lineární prostor; takže $\overline{\operatorname{Lin}\{v_1,v_2,\dots\}}$ je skutečně lineární podprostor X.

(ii) Tvrzení z (a) ovšem platí (se „stejným“ důkazem) pro konečný ortogonální systém nenulových prvků $(v_k)_{k=1}^{n}$.

(iii) V Tvrzení (b) stačí místo úplnosti předpokládat, že řada $\sum_{k=1}^{\infty} c_k\, v_k$ konverguje.

(iv) Použijeme-li (a) a postup z důkazu Věty 4.61, snadno dostaneme tvrzení z Poznámky 4.64. Je-li totiž ortonormální systém (u_n) úplný, pak z (a) snadno vyplývá, že $\|x-\sum_{k=1}^{n} c_k u_k\| \to 0$.

Je-li X Hilbertův prostor a $Y \subset X$ jeho separabilní nekonečně dimenzionální uzavřený podprostor, pak Y je Hilbertův prostor, takže podle Tvrzení 4.67 v něm existuje ortonormální báze $(u_n)_{n=1}^{\infty}$. Pak $Y = \overline{\operatorname{Lin}\{u_1,u_2,\dots\}}$, takže podle Tvrzení 4.72 (b) má každý prvek $x \in X$ ortogonální projekci na Y. Z toho okamžitě vyplývá, že každý prvek $x \in X$ lze jednoznačně zapsat ve tvaru

$$x = y + y', \quad \text{kde } \ y \in Y \ \text{ a } \ y' \in Y^{\perp};$$

toto tvrzení se zapisuje ve formě $X = Y \oplus Y^{\perp}$. (Pro konečně dimenzionální Y platí totéž, srov. Poznámka 4.73.)

4.4 Aplikace na klasické Fourierovy řady

V dalším textu používáme následující úmluvu (která nemůže vést k nedorozumění): pokud o funkci $f \in \mathscr{P}(2\pi)$ hovoříme jako o prvku $L^2(\alpha, \alpha + 2\pi)$, máme na mysli prvek $L^2(\alpha, \alpha + 2\pi)$ (tj. třídu funkcí), který obsahuje restrikci $f \restriction_{(\alpha,\alpha+2\pi)}$.

4.74 Věta. *Reálný trigonometrický systém je ortogonální báze reálného (nebo komplexního) prostoru* $L^2(\alpha, \alpha + 2\pi)$, $(\alpha \in \mathbb{R})$.

Důkaz. Zřejmě stačí dokázat „reálný případ“. Podáme dva důkazy.

a) Protože prostor $L^2(\alpha, \alpha + 2\pi)$ je úplný, podle Věty 4.63 stačí dokázat, že ortogonální systém

$$1,\ \cos x,\ \sin x,\ \cos 2x,\ \sin 2x, \dots$$

je maximální. Předpokládejme tedy, že $f \in L^2(\alpha, \alpha + 2\pi)$ je kolmá na všechny prvky tohoto systému. Nalezněme $f^* \in \mathscr{P}(2\pi)$, která „je rozšířením funkce f“. Pak funkce f^* má podle Lemmatu 4.7 nulové všechny Fourierovy koeficienty, takže je podle Věty 4.41 skoro všude nulová.

b) Podle Věty 4.63 stačí dokázat úplnost trigonometrického systému, tj. tvrzení, že množina trigonometrických polynomů $\mathcal{T} = \mathrm{Lin}\{1, \cos x, \sin x, \dots\}$ je hustá v $L^2(\alpha, \alpha+2\pi)$. K tomu použijeme skutečnost (srov. [LM], Cvičení 15.17 a 18.5), že množina $C := C_C(\alpha, \alpha+2\pi)$ spojitých funkcí na $(\alpha, \alpha+2\pi)$ s kompaktním nosičem (tj. těch, pro které $\overline{\{x\colon f(x) \neq 0\}}$) je hustá v $L^2(\alpha, \alpha + 2\pi)$. Je-li $g \in C$, zřejmě lze najít její spojité rozšíření $g^* \in \mathscr{P}(2\pi)$. Podle (Fejérovy) Věty 4.38 (ii) existuje posloupnost trigonometrických polynomů T_n, která konverguje stejnoměrně ke g^* na $\mathbb{R}$. Z toho snadno vyplývá, že $T_n \to g$ v $L^2(\alpha, \alpha + 2\pi)$, takže $C \subset \overline{\mathcal{T}}$. Platí tedy $\overline{\mathcal{T}} = L^2(\alpha, \alpha + 2\pi)$.

4.75 Věta. *Nechť* $\alpha \in \mathbb{R}$, $f \in \mathscr{P}(2\pi)$ *a funkce* $|f|^2$ *je lokálně integrovatelná, a* $a_0, a_1, \dots,\ b_1, b_2, \dots$ *jsou její Fourierovy koeficienty. Pak*

$$f = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \quad \text{v} \quad L^2(\alpha, \alpha + 2\pi). \tag{4.22}$$

Dále platí (klasická Parsevalova) rovnost

$$\frac{1}{\pi} \int_{\alpha}^{\alpha+2\pi} |f|^2 = \frac{|a_0|^2}{2} + \sum_{n=1}^{\infty} \left(|a_n|^2 + |b_n|^2\right). \tag{4.23}$$

Navíc platí, že v prostoru $L^2(\alpha, \alpha + 2\pi)$ *je (pro* $n = 0, 1, \dots$*) částečný součet Fourierovy řady funkce* f

$$s_n(x) = \frac{a_0}{2} + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx)$$

nejlepší aproximací funkce f v množině všech trigonometrických polynomů stupně nejvýše n (mezi které počítáme i nulovou funkci).

Důkaz. Z Věty 4.74 ihned vyplývá, že také $1/2, \cos x, \sin x$ je ortogonální báze $L^2(\alpha, \alpha + 2\pi)$. Protože $a_0, a_1, b_1, \dots$ jsou Fourierovy koeficienty f vzhledem k této bázi, podle Důsledku 4.58 v $L^2(\alpha, \alpha + 2\pi)$ platí

$$f = a_0/2 + a_1 \cos x + b_1 \sin x + \dots, \tag{4.24}$$

a tedy i (4.22).

Klasická Parsevalova rovnost (4.23) je speciálním případem abstraktní Parsevalovy rovnosti (viz Poznámka 4.62) pro ortogonální systém $1/2, \cos x, \sin x$. Je možné ovšem také použít zobecněnou Pythagorovu větu na ((?))pocts¿ a dostat

$$\|f\|^2 = \|a_0/2\|^2 + \|a_1 \cos x\|^2 + \|a_1 \sin x\|^2 + \dots .$$

Protože podle Lemmatu 4.9 máme $\|1\|^2 = 2\pi$, $\|\cos nx\|^2 = \|\sin nx\|^2 = \pi$, platí

$$\int_{\alpha}^{\alpha+2\pi} |f|^2 = \frac{\pi |a_0|^2}{2} + \pi \sum_{n=1}^{\infty} \left(|a_n|^2 + |b_n|^2\right),$$

z čehož ihned plyne (4.23).

Poslední tvrzení věty je snadným důsledkem Tvrzení 4.72 (a).

4.76 Věta. (Riesz–Fischerova věta) *Nechť a_0, a_1, b_1, a_2, $b_2, \dots$ jsou reálná (resp. komplexní) čísla, pro která řada $\sum(|a_n|^2 + |b_n|^2)$ konverguje. Pak existuje reálná (resp. komplexní) funkce $f \in \mathscr{P}(2\pi)$, jejíž Fourierovy koeficienty jsou tato čísla. Funkce f je určena jednoznačně (až na množinu míry nula) a leží v $L^2(0, 2\pi)$.*

Důkaz. Uvažujme v prostoru $L^2(0, 2\pi)$ řadu

$$a_0/2 + a_1 \cos x + b_1 \sin x + \dots . \tag{4.25}$$

Protože $\|1\|^2 = 2\pi$ a $\|\cos nx\|^2 = \|\sin nx\|^2 = \pi$, řada

$$\|a_0/2\|^2 + \|a_1 \cos x\|^2 + \|a_1 \sin x\|^2 + \dots$$

konverguje, takže podle Tvrzení 4.54 (ii) řada (4.25) konverguje v $L^2(0, 2\pi)$ k nějaké funkci g; zvolme její 2π-periodické rozšíření (které zřejmě leží v $\mathscr{P}(2\pi)$) a označme je f. Podle Tvrzení 4.56 jsou čísla a_0, a_1, b_1, a_2, $b_2, \dots$ „abstraktními“ Fourierovými koeficienty funkce g vzhledem k ortogonální bázi $1/2, \cos x, \sin x, \dots$ a jsou tedy „klasickými“ Fourierovými koeficienty funkce f (srov. Poznámka 4.60). Jednoznačnost f vyplývá z Věty 4.41.

4.77 Poznámka. a) Z uvedeného důkazu je jasně vidět jednoduchá myšlenka důkazu Riesz-Fischerovy věty. Kratší však bylo užít Větu 4.66 (jejíž důkaz jsme částečně zopakovali) na znormovaný trigonometrický systém.

b) Je zřejmé, jak se odvodí analogické výsledky pro klasické Fourierovy řady v komplexní formě. Uveďme pouze, že Parsevalova rovnost má tvar

$$\frac{1}{2\pi}\int_{\alpha}^{\alpha+2\pi} |f|^2 = \sum_{n=-\infty}^{\infty} |c_n|^2.$$

4.5 Fourierův integrál a Fourierova transformace

Dokázali jsme, že dosti obecné 2π-periodické funkce jsou součtem své Fourierovy řady; lze je tedy vyjádřit jako součet konstanty a harmonických 2π-periodických funkcí. Totéž ovšem platí pro periodické funkce s obecnou periodou $l > 0$.

Fourier ukázal, že analogicky lze vyjádřit i funkce, které *nejsou periodické*, je však nutno použít harmonické funkce *všech period* a místo součtu použít *integrál*. Tuto svou teorii „Fourierova integrálu" přitom úspěšně užil ve fyzice při řešení rovnice vedení tepla.

Připomeňme (viz. Poznámka 4.15), že obecnou harmonickou funkci („harmonický kmit") lze psát ve tvaru

$$h(x) = a\cos\omega x + b\sin\omega x,$$

sk

kde $\omega > 0$; jde o harmonickou funkci s minimální periodou $l := 2\pi/\omega$.

Položme si otázku, zda lze danou funkci $f : \mathbb{R} \to \mathbb{R}$ vyjádřit pomocí harmonických funkcí ve tvaru

$$f(x) = \int_0^{\infty} (a(\omega)\cos\omega x + b(\omega)\sin\omega x)\, d\omega. \tag{4.26}$$

Fourier objevil, že toto vyjádření pro mnoho funkcí f platí, definujeme-li koeficienty $a(\omega), b(\omega)$ pomocí vzorců analogických vzorcům pro Fourierovy koeficienty pro řady:

$$a(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(y)\,\cos\omega y\,\,\mathrm{d}y, \qquad b(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(y)\,\sin\omega y\,\,\mathrm{d}y. \tag{4.27}$$

Přijít na tyto vzorce je obtížnější než na vzorce pro Fourierovy koeficienty a_n, b_n pro řady (srov. Tvrzení 4.10). Nepřesné (heuristické) úvahy, které vedou k objevení vzorců (4.27) lze najít např. v [J II].

O reálné funkci f budeme předpokládat, že má konvergentní Lebesgueův integrál, tj. $f \in L(\mathbb{R})$. Pak oba Lebesgueovy integrály v (4.27) zřejmě konvergují.

Navíc z věty o spojitosti Lebesgueova integrálu závislého na parametru snadno vyplývá, že funkce $a(\omega)$, $b(\omega)$ jsou spojité na $[0,\infty)$, takže Lebesgueův integrál $\displaystyle\int_0^{\eta} (a(\omega)\cos\omega x + b(\omega)\sin\omega x)\, d\omega$ konverguje pro každé $\eta > 0$.

Integrál stojící napravo v (4.26) (kde $a(\omega)$, $b(\omega)$ jsou vypočteny podle (4.27)) se nazývá *Fourierův integrál* funkce f. Tento integrál chápeme jako *„nevlastní integrál"*. Jinými slovy, Fourierův integrál má v bodě x hodnotu $I(x) \in \mathbb{R}$, právě když

$$\lim_{\eta\to\infty} I_\eta(x) = I(x),$$

kde $I_\eta(x)$ je „částečný Fourierův integrál", což je „vlastní" integrál (ze spojité funkce)

$$I_\eta(x) := \int_0^\eta (a(\omega)\cos\omega x + b(\omega)\sin\omega x)\, d\omega. \tag{4.28}$$

To je zcela analogické případu Fourierových řad: Fourierova řada také nemusí konvergovat absolutně a její součet je limita částečných součtů.

Dosadíme-li vzorce (4.27) do integrandu ve Fourierově integrálu a použijeme vzorec $\cos\omega y\,\cos\omega x + \sin\omega y\,\sin\omega x = \cos(\omega(y-x))$, dostáváme

$$\begin{aligned}
&a(\omega)\cos\omega x + b(\omega)\sin\omega x = \\
&= \left(\frac{1}{\pi}\int_{-\infty}^{\infty} f(y)\,\cos\omega y\,\mathrm{d}y\right)\cos\omega x + \left(\frac{1}{\pi}\int_{-\infty}^{\infty} f(y)\sin\omega y\,\mathrm{d}y\right)\sin\omega x = \\
&= \frac{1}{\pi}\int_{-\infty}^{\infty} f(y)\,\cos(\omega(y-x))\,\mathrm{d}y.
\end{aligned}$$

Fourierův integrál lze tedy psát ve tvaru

$$\frac{1}{\pi}\int_0^\infty \left(\int_{-\infty}^{\infty} f(y)\,\cos(\omega(y-x))\,\mathrm{d}y\right)\mathrm{d}\omega, \tag{4.29}$$

ve kterém se také uvádí nejčastěji (z kterého však není ihned vidět analogie s Fourierovou řadou).

Všechny předchozí úvahy jsou správné i pro funkce f z komplexního prostoru $L(\mathbb{R})$; funkce $a(\omega)$, $b(\omega)$ jsou pak ovšem také komplexní. Dále již uvažujeme *komplexní* f, není-li řečeno jinak.

Rovnosti (4.26) se říká *Fourierův integrální vzorec.* Pro platnost Fourierova integrálního vzorce platí zcela analogická kritéria jako pro konvergenci Fourierovy řady (srov. [J II; Věta 193]). My zde dokážeme pouze Diniho kritérium pro Fourierův integrál. Jeho důkaz je analogický důkazu pro Fourierovu řadu; navíc však budeme potřebovat vzorec, který jsme odvodili v Příkladu 4.26:

$$\int_0^\infty \frac{\sin\omega}{\omega}\,\mathrm{d}\omega = \frac{\pi}{2}. \tag{4.30}$$

Podobně jako jsme v Lemmatu 4.21 odvodili vyjádření pro částečný součet $s_n(x)$ Fourierovy řady, odvodíme nyní analogické vyjádření pro „částečný Fourierův integrál" $I_\eta(x)$.

4.78 Lemma. *Nechť $f \in L(\mathbb{R})$ a $x \in \mathbb{R}$. Pak pro $\eta > 0$ platí*

$$I_\eta(x) = \frac{1}{\pi} \int_{-\infty}^{\infty} f(x+z)\, \frac{\sin \eta z}{z}\, \mathrm{d}z. \tag{4.31}$$

Důkaz. Stačí uvažovat reálné funkce f. Podle (4.29) platí

$$I_\eta(x) = \frac{1}{\pi} \int_0^\eta \left(\int_{-\infty}^{\infty} f(y)\, \cos(\omega(y-x))\, \mathrm{d}y \right) \mathrm{d}\omega.$$

Po substituci $y = x + z$ vychází

$$I_\eta(x) = \frac{1}{\pi} \int_0^\eta \left(\int_{-\infty}^{\infty} f(x+z)\, \cos(\omega z)\, \mathrm{d}z \right) \mathrm{d}\omega.$$

Protože funkce $g(\omega, z) := f(x+z)\, \cos(\omega z)$ je zřejmě měřitelná na $P := (0, \eta) \times \mathbb{R}$, $|g(\omega, z)| \leq |f(x+z)|$ a zřejmě $h(\omega, z) := |f(x+z)| \in L(P)$, je také $g \in L(P)$. Lze tedy použít Fubiniovu větu, podle které dostáváme

$$I_\eta(x) = \frac{1}{\pi} \left(\int_{-\infty}^{\infty} \int_0^\eta f(x+z)\, \cos(\omega z)\, \mathrm{d}\omega \right) \mathrm{d}z = \frac{1}{\pi} \int_{-\infty}^{\infty} f(x+z)\, \frac{\sin \eta z}{z}\, \mathrm{d}z.$$

4.79 Poznámka. Funkce $F_\eta(z) = \dfrac{\sin \eta z}{z}$ ($F_\eta(0) = \eta$), kde $\eta > 0$, se někdy nazývá *Fourierovo jádro*. Hraje v teorii Fourierova integrálu analogickou roli, jako Dirichletovo jádro v teorii Fourierových řad. Použijeme-li substituci $\omega = \eta z$ a (4.30), okamžitě dostáváme

$$\int_0^\infty \frac{\sin \eta z}{z}\, \mathrm{d}z = \frac{\pi}{2}. \tag{4.32}$$

4.80 Věta. (Diniho kritérium pro Fourierův integrál.) *Nechť $f \in L(\mathbb{R})$, $x \in \mathbb{R}$, $s \in \mathbb{C}$ a nechť existuje $\delta > 0$ takové, že Lebesgueův integrál*

$$\int_0^\delta \frac{f(x+v) + f(x-v) - 2s}{v}\, \mathrm{d}v \tag{4.33}$$

konverguje. Pak s je hodnotou Fourierova integrálu funkce f v bodě x, tj. platí rovnost $\lim_{\eta\to\infty} I_\eta(x) = s$.

Důkaz. Bez újmy na obecnosti můžeme předpokládat, že f je reálná funkce a $c \in \mathbb{R}$. Nechť je dáno $\delta > 0$ z předpokladů věty dáno. Z (4.32) okamžitě dostáváme rovnost $s = \dfrac{1}{\pi} \displaystyle\int_0^\infty 2s\, \frac{\sin \eta z}{z}\, \mathrm{d}z$.

Lemma 4.78 a sudost Fourierova jádra dávají

$$I_\eta(x) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(x+z)\,\frac{\sin\eta z}{z}\,\mathrm{d}z = \frac{1}{\pi}\int_0^{\infty} \big(f(x+z)+f(x-z)\big)\,\frac{\sin\eta z}{z}\,\mathrm{d}z.$$

Platí tedy (1. a 4. integrál jsou „nevlastní“zobecněné Lebesgueovy integrály)

$$\begin{aligned}
I_\eta(x) - s &= \frac{1}{\pi}\int_0^{\infty} \big(f(x+z)+f(x-z)-2s\big)\,\frac{\sin\eta z}{z}\,\mathrm{d}z = \\
&= \frac{1}{\pi}\int_0^{\delta} \frac{f(x+z)+f(x-z)-2s}{z}\,\sin\eta z\,\mathrm{d}z + \\
&+ \frac{1}{\pi}\int_\delta^{\infty} \frac{f(x+z)+f(x-z)}{z}\,\sin\eta z\,\mathrm{d}z - \frac{2s}{\pi}\int_\delta^{\infty}\frac{\sin\eta z}{z}\,\mathrm{d}z.
\end{aligned}$$

Z Riemann–Lebesgueova lemmatu (Věta 4.23) snadno dostáváme, že první dva sčítanci v posledním výrazu mají (pro $\eta\to\infty$) nulovou limitu. Protože (substituce $\eta z = t$)

$$\int_\delta^{\infty}\frac{\sin\eta z}{z}\,\mathrm{d}z = \int_{\eta\delta}^{\infty}\frac{\sin t}{t}\,\mathrm{d}t,$$

z konvergence integrálu (4.30) dostáváme, že i třetí sčítanec má nulovou limitu.

Z Diniho kritéria vyplývá stejně jako v případě Fourierových řad následující tvrzení, které pro „běžné funkce“ umožňuje určit hodnotu jejich Fourierova integrálu.

4.81 Důsledek. (důsledky Diniho kritéria)

(i) Nechť $f\in L(\mathbb{R})$, $x\in\mathbb{R}$, existují (komplexní) jednostranné limity $f(x_+)$, $f(x_-)$ a také (komplexní) limity

$$\lim_{t\to x+}\frac{f(t)-f(x_+)}{t-x},\qquad \lim_{t\to x-}\frac{f(t)-f(x_-)}{t-x}.$$

Pak Fourierův integrál funkce f v bodě x má hodnotu $\frac12\big(f(x_+)+f(x_-)\big)$.

Speciálně, pokud $f\in L(\mathbb{R})$ má konečné jednostranné derivace v každém bodě, pak pro funkci f v každém bodě $x\in\mathbb{R}$ platí Fourierův integrální vzorec.

(ii) Nechť $f\in L(\mathbb{R})$, $x\in\mathbb{R}$ a existují čísla $\alpha>0$, $\delta>0$ a $K>0$ taková, že

$$|f(x+v)-f(x)| \le K|v|^\alpha \qquad \text{pro} \quad |v|<\delta.$$

Pak pro funkci f v bodě x platí Fourierův integrální vzorec.

4.82 Příklad. Nechť $f(x) = \max(1-|x|,0)$. Zřejmě $f\in L(\mathbb{R})$ a má v každém bodě konečné jednostranné derivace; podle Důsledku 4.81 pro každé $x\in\mathbb{R}$ platí Fourierův vzorec (4.26). Protože f je sudá, ihned vidíme, že $b(\omega)=0$, $\omega>0$. Snadný výpočet dává (pro $\omega>0$)

$$a(\omega) = \frac{1}{\pi}\int_{-\infty}^{\infty} f(y)\,\cos\omega y\,\mathrm{d}y = \frac{2}{\pi}\int_0^1 (1-y)\,\cos\omega y\,\mathrm{d}y = \frac{2}{\pi}\,\frac{1-\cos\omega}{\omega^2}.$$

Pro každé $x \in \mathbb{R}$ tedy platí

$$\frac{2}{\pi}\int_0^\infty \frac{1-\cos\omega}{\omega^2}\cos\omega x \,\mathrm{d}\omega = f(x).$$

Použijeme-li tento vzorec pro $x=0$ a v integrálu provedeme substituci $\omega = 2z$, dostáváme rovnost

$$\int_0^\infty \frac{\sin^2 z}{z^2}\,\mathrm{d}z = \frac{\pi}{2}.$$

Zcela analogicky jako jsme odvodili komplexní tvar (4.4) částečného součtu Fourierovy řady lze odvodit komplexní tvar „částečného Fourierova integrálu“; platí

$$I_\eta(x) = \int_{-\eta}^{\eta} c(\omega)\, e^{i\omega x}\,\mathrm{d}\omega, \qquad \text{kde} \quad c(\omega) := \frac{1}{2\pi}\int_{-\infty}^{\infty} f(y)e^{-i\omega y}\,\mathrm{d}y. \tag{4.34}$$

Fourierův integrál funkce $f \in L(\mathbb{R})$ má tedy v bodě $x \in \mathbb{R}$ hodnotu s, právě když

$$s = \lim_{\eta\to\infty}\int_{-\eta}^{\eta} c(\omega)\, e^{i\omega x}\,\mathrm{d}\omega =: (v.p.)\int_{-\infty}^{\infty} c(\omega)\, e^{i\omega x}\,\mathrm{d}\omega. \tag{4.35}$$

Výraz napravo je tzv. *hlavní hodnota integrálu v Caychyově smyslu* definovaná rovností $(v.p.)\int_{-\infty}^{\infty} g := \lim_{\eta\to\infty}\int_{-\eta}^{\eta} g$ (kde integrály napravo jsou Lebesgueovy).

Funkci $c(\omega)$ se někdy říká Fourierova transformace funkce f; častější je však následující nepatrně odlišná, početně výhodnější definice.

4.83 Definice. *Pro (komplexní) funkci $f \in L(\mathbb{R})$ položme*

$$\widehat{f}(\omega) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(x)e^{-i\omega x}\,\mathrm{d}x. \tag{4.36}$$

Funkci $\widehat{f}$ nazýváme Fourierovým obrazem funkce f (někdy také Fourierovou transformací funkce f). Zobrazení $F\colon f \mapsto \widehat{f}$ nazýváme Fourierovou transformací (takže $\widehat{f} = F(f)$).

Inverzní (konjugovanou) Fourierovu transformaci definujeme jako zobrazení F^*, které funkci $f \in L(\mathbb{R})$ přiřazuje její inverzní (konjugovaný) Fourierův obraz

$$f^\vee(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(\omega)e^{i\omega x}\,\mathrm{d}\omega. \tag{4.37}$$

Pomocí věty o spojitosti integrálu závislého na parametru snadno dostáváme, že funkce $\widehat{f} = F(f)$, $f^\vee = F^*(f)$ jsou spojité, a z Riemann–Lebesgueova lemmatu vyplývá, že

$$\lim_{\omega\to\pm\infty}\widehat{f}(\omega) = 0, \qquad \lim_{x\to\pm\infty} f^\vee(x) = 0.$$

Víme-li, že pro funkci $f \in L(\mathbb{R})$ platí v bodě x Fourierův integrální vzorec (což často snadno vyplývá z Důsledku 4.81), můžeme jej podle (4.35) psát ve formě

$$f(x) = \frac{1}{\sqrt{2\pi}}(v.p.)\int_{-\infty}^{\infty} \widehat{f}(\omega)e^{i\omega x}\, \mathrm{d}\omega.$$

Platí-li ještě navíc, že $\widehat{f} \in L(\mathbb{R})$, lze Fourierův integrální vzorec psát ve tvaru

$$F^*(F(f))(x) = f(x);$$

této rovnosti se někdy říká (Fourierova) *inverzní formule* (v bodě x).

4.84 Poznámka.

(i) Platí věta (viz [Ru2]), že již podmínka $\widehat{f} \in L(\mathbb{R})$ stačí k tomu, aby inverzní formule platila pro skoro všechna $x \in \mathbb{R}$.

(ii) Vzorce pro F a F^* jsou „téměř stejné“; liší se jen znaménkem v exponentu. Odtud jsou vidět vzorce

$$F^*(f) = \overline{F(\overline{f})} \quad \text{a} \quad F^*(f) = F(f_s),$$

kde $f_s(x) := f(-x)$. Inverzní formuli lze tedy psát ve tvarech

$$\overline{F\left(\overline{F(f)}\right)}(x) = f(x), \qquad F\left((F(f))_s\right)(x) = f(x).$$

(iii) Pojmy Fourierova integrálu i Fourierovy transformace se zobecňují různými způsoby, přičemž se dokazují rozmanité formy Fourierova integrálního vzorce (inverzní formule).

???? Poznámka o vlastnostech F.tr. ?

5. Úvod do teorie plošného a křivkového integrálu

5.1 Úvod

K nejdůležitějším klasickým výsledkům diferenciálního a integrálního počtu patří nepochybně Greenova, Gaussova a Stokesova věta, které mají důležité aplikace ve fyzice a v řadě matematických teorií. Velmi zhruba lze říci, že tyto věty říkají, že jistý integrál přes množinu $M \subset \mathbb{R}^n$ je roven jistému příslušnému integrálu přes hranici ∂M této množiny a že jde o zobecnění Newton-Leibnizovy formule $f(b) - f(a) = \int_a^b f'(x)\,\mathrm{d}x$. Pro formulaci zmíněných vět je třeba zavést tzv. plošné a křivkové integrály.

Již přesná formulace těchto „integrálních vět“ je dosti obtížná. Ve starších klasických učebnicích je podán názorný výklad teorie plošného integrálu a integrálních vět, který čtenáře seznamuje se základními myšlenkami a umožňuje mu zvládnout početní techniku, nesnese však současné nároky na matematickou přesnost. Dokonce ani základní pojmy nejsou přesně definovány, natož aby byly přesně dokázány základní věty. V. Jarník do svých učebnic teorii plošného integrálu nezařadil; v předmluvě (z r. 1955) k [J II] poznamenává, že nezná v literatuře výklad, který by vyhovoval současně z vědeckého i pedagogického hlediska. Nyní existuje několik přesných česky psaných výkladů (na různých stupních obecnosti) této teorie (např. [ČM], [Si], [Kow], [Kop], [LM], [KST]).

Cílem této krátké kapitoly je co nejstručněji přesně vyložit Gaussovu a Greenovu větu, ale tak, aby všechny pojmy byly motivovány a aby *standardní konkrétní příklady bylo možno pomocí těchto vět co nejpohodlněji počítat.* Motivace definic základních pojmů je vedena obvyklým klasickým způsobem. Také výklad (hodně podobný výkladu z [Kop]) se od klasického liší v podstatě jen v tom, že užívá standardní výsledky abstraktní teorie míry a integrálu. Důkaz (trochu zdlouhavý, ale ne obtížný) věty, která „axiomaticky“ definuje k-rozměrnou míru (na „minimální“ σ-algebře $\mathcal{P}_k$), je proveden v Dodatku 6.4. Standardní snadný důkaz Gaussovy věty pro velmi speciální množiny je proveden v základním textu, poměrně obtížný důkaz v obecném případě (jen málo se lišící od důkazu z [Kop]) lze nalézt v Dodatku 6.5.

Klasická Stokesova věta je uvedena bez důkazu spolu s poznámkami o diferenciálních formách a obecné Stokesově větě (jejíž jedna elementární verze je také bez důkazu uvedena).

Důkazy a poznámky, které nejsou pro pochopení textu zásadní, jsou psány petitem.

Protože terminologie z teorie ploch zavedená v Kapitole 2 je příliš těžkopádná, dohodneme se pro účely této kapitoly na jejím zjednodušení.

Zjednodušení terminologie Není-li řečeno jinak, jsou v této kapitole n, k vždy přirozená čísla, pro která $1 \leq k < n$. Místo k-rozměrná plocha třídy C^1 budeme psát pouze *k-plocha* a místo parametricky zadaný kus k-rozměrné C^1 plochy budeme psát *jednoduchá k-plocha. Parametrizací jednoduché k-plochy* P v $\mathbb{R}^n$ budeme rozumět libovolný regulární homeomorfismus $f\colon G \to \mathbb{R}^n$ ($G \subset \mathbb{R}^k$), pro který $f(G) = P$. Dále pod pojmem „jednoduchá 0-plocha v $\mathbb{R}^n$“ rozumíme jednobodovou podmnožinu $\mathbb{R}^n$ a pod pojmem „0-plocha v $\mathbb{R}^n$“ izolovanou podmnožinu $\mathbb{R}^n$ (tj. množinu, jejíž každý bod je izolovaný).

5.2 Starší a novější přístup k plošnému integrálu 1. druhu

Nejprve provedeme (nepřesnou) heuristickou úvahu, která vysvětluje *starší přístup* k plošnému integrálu 1. druhu. Nechť $P \subset \mathbb{R}^n$ je k-plocha, jejíž k-rozměrný obsah je konečný a nechť f je stejnoměrně spojitá funkce na P. Proveďme rozklad plochy P na konečně mnoho plošek $P_1, \dots P_r$, v každé z nich zvolme bod $x_i \in P_i$ a utvořme součet

$$\sum_{i=1}^{r} f(x_i)\,(\Delta S)_i, \tag{5.1}$$

kde $(\Delta S)_i$ je k-rozměrný obsah plošky P_i. Vzhledem ke stejnoměrné spojitosti f se zdá být věrohodné, že pokud diametr plošek P_i neomezeně zmenšujeme (a jejich počet zvětšujeme), blíží se „integrální součty“ (5.1) k jistému číslu, které nazýváme plošným integrálem 1. druhu a značíme (vedeni analogií s výrazem (5.1)) $\int_P f \,\mathrm{dS}$.

Pro výpočet integrálních součtů (5.1) ovšem potřebujeme znát definici velikosti plošek P_i. Tento problém lze překonat například tak, že $\int_P f \,\mathrm{dS}$ se nedefinuje jako limita součtů (5.1), ale jako limita součtů $\sum_{i=1}^{r} f(x_i)\,\sigma_i$, kde σ_i je k-rozměrný obsah kolmého průmětu P_i^* plošky P_i na afinní tečný prostor A k ploše P v bodě x_i. Přitom se vychází z předpokladu, že pokud je ploška P_i velmi malá, P_i a P_i^* od sebe „nerozeznáme“, proto platí $(\Delta S)_i \approx \sigma_i$, a tedy se také domníváme, že $\sum_{i=1}^{r} f(x_i)\,(\Delta S)_i \approx \sum_{i=1}^{r} f(x_i)\,\sigma_i$.

Plošný integrál 1. druhu má řadu důležitých fyzikálních aplikací. Předpokládejme například, že dvourozměrná plocha P v $\mathbb{R}^3$ je „zhotovena z nehomogenního materiálu“ a $f(x)$ má význam „plošné hustoty“ této „hmotné plochy“ v bodě $x \in P$. Pokud plošky P_i jsou dostatečně malé, zdá se být zřejmé, že číslo $f(x_i)(\Delta S)_i$ velmi dobře aproximuje hmotnost m_i plošky P_i. Součet (5.1) proto velmi dobře aproximuje hmotnost plochy P a integrál $\int_P f\,\mathrm{dS}$ má význam hmotnosti plochy P. Také pro výpočet těžiště plochy P (srov. Příklad 5.95) nebo gravitační síly, kterou P přitahuje hmotný bod (srov. Příklad 5.27), nutně potřebujeme plošný integrál 1. druhu.

Současný přístup k integrálu 1. druhu však není založen na limitě integrálních součtů. Známe-li totiž teorii abstraktního Lebesgueova integrálu, je přirozené nejprve na ploše P zavést k-rozměrnou míru μ_k (která množinám $B \subset P$ přiřazuje jejich k-rozměrný plošný obsah $\mu_k(B)$) a plošný integrál 1. druhu pak definovat rovností

$$\int_P f\,\mathrm{dS} := \int_P f\ \mathrm{d}\mu_k. \tag{5.2}$$

V následujících oddílech se proto budeme věnovat zavedení k-rozměrné míry. Zdůrazněme, že symbolem μ_k budeme vždy označovat k-rozměrnou míru definovanou na jisté σ-algebře borelovských množin. Přirozenější a obecnější by bylo pracovat se zúplněnou mírou $\widehat{\mu}_k$; tím by se ale výklad komplikoval a pro formulaci Gaussovy věty zúplněnou míru nepotřebujeme.

Dále poznamenejme toto:

a) V klasických aplikacích jsou nejčastější případy $k = 1$ a $k = 2$; 1-rozměrné míře $\mu_1(A)$ (množiny $A \subset \mathbb{R}^2$ nebo $A \subset \mathbb{R}^3$) se často říká délka množiny A a 2-rozměrné míře $\mu_2(A)$ množiny $A \subset \mathbb{R}^3$ se říká plošný obsah množiny A.

b) Rovnost (5.2) definuje plošný integrál 1. druhu i pro velmi obecné nespojité funkce (např. pro každou nezápornou borelovsky měřitelnou funkci na P), pro které klasická definice pomocí integrálních součtů selhává.

c) Je-li $\mu_k(P) < \infty$ a f je stejnoměrně spojitá (jak jsme předpokládali výše), lze snadno dokázat (srov. Tvrzení 6.39), že $\int_P f\ \mathrm{d}\mu_k$ je limitou integrálních součtů (5.1).

5.1 Poznámka. Pro geometrii je jedním ze základních pojmů integrál diferenciální formy přes orientovanou plochu (resp. varietu). Pomocí tohoto pojmu lze integrál prvního druhu také definovat, obtížný pojem orientace plochy však pro definici integrálu 1. druhu není nutný. Jinak je tomu s integrálem 2. druhu, kde je pojem orientace plochy podstatný a při důkladném výkladu je velmi elegantní (ale při přesném výkladu dosti náročná) teorie diferenciálních forem nezbytná.

5.3 Definice k-rozměrné míry na k-rozměrném afinním podprostoru $\mathbb{R}^n$

Uvažujme nejprve nejjednodušší případ plochy, která „není křivá“, totiž afinního prostoru. Některá základní fakta o afinních podprostorech eukleidovských prostorů a o afinních zobrazeních mezi nimi, která budeme používat, jsou shrnuta v Dodatku 6.7.

Jestliže A je k-rozměrný afinní podprostor $\mathbb{R}^n$ $(1 \leq k < n)$, existuje (srov. Dodatek 6.7 c)) izometrie $\varphi: \mathbb{R}^k \to A$. Lze tedy A (touto izometrií) „ztotožnit“ s $\mathbb{R}^k$. Protože na $\mathbb{R}^k$ existuje kanonická k-rozměrná míra (totiž Lebesgueova míra); zdá se být zřejmé, že i na A lze definovat „kanonickou k-rozměrnou míru“ μ_k^A (která je obrazem Lebesgueovy míry při každé takové izometrii).

Pro jednoduchost budeme na A definovat pouze „kanonickou“ míru μ_k^A na σ-algebře $\mathcal{B}(A)$ borelovských podmnožin A.

Dále λ_k je Lebesgueova míra v $\mathbb{R}^k$ a λ_k^b je její restrikce na σ-algebru $\mathcal{B}(\mathbb{R}^k)$ borelovských podmnožin $\mathbb{R}^k$.

Je-li $\varphi: \mathbb{R}^k \to A$ izometrie, je $\varphi: (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k)) \to (A, \mathcal{B}(A))$ zřejmě měřitelné zobrazení, takže na $\mathcal{B}(A)$ je definován obraz $\mu_\varphi := \varphi(\lambda_k^b)$ míry λ_k^b při zobrazení φ; podle definice platí

$$\mu_\varphi(B) := \lambda_k(\varphi^{-1}(B)), \qquad B \in \mathcal{B}(A).$$

Snadno vidíme, že μ_φ nezávisí na výběru izometrie φ. Je-li totiž $\psi: \mathbb{R}^k \to A$ jiná izometrie a $B \in \mathcal{B}(A)$, pak zřejmě $\psi^{-1}(B) = (\psi^{-1} \circ \varphi)(\varphi^{-1}(B))$, a protože $\psi^{-1} \circ \varphi$ je izometrie $\mathbb{R}^k$ na $\mathbb{R}^k$, platí (viz Tvrzení 6.36)

$$\mu_\varphi(B) = \lambda_k(\varphi^{-1}(B)) = \lambda_k(\psi^{-1}(B)) = \mu_\psi(B).$$

Následující definice je tedy korektní.

5.2 Definice. *Nechť $A \subset \mathbb{R}^n$ je k-rozměrný $(1 \leq k \leq n)$ afinní podprostor $\mathbb{R}^n$ a $\varphi: \mathbb{R}^k \to A$ je izometrie. Pak klademe $\mu_k^A := \varphi(\lambda_k^b)$.*

Je-li $k = n$, zřejmě $\mu_k^A = \lambda_k^b$. Je-li A pevně dáno, budeme psát místo μ_k^A jen μ_k.

5.3 Poznámka. Míra μ_k^A byla definována v závislosti na afinním prostoru A. Pokud ale B je borelovská podmnožina dvou různých k-rozměrných afinních prostorů $A_1 \subset \mathbb{R}^n$, $A_2 \subset \mathbb{R}^n$, pak $A_1 \cap A_2$ je afinní prostor dimenze menší než k. Z toho snadno vyplývá, že $\mu_k^{A_1}(A_1 \cap A_2) = \mu_k^{A_2}(A_1 \cap A_2) = 0$, takže také $\mu_k^{A_1}(B) = \mu_k^{A_2}(B) = 0$. Pro každou borelovskou množinu $B \subset \mathbb{R}^n$, která je podmnožinou nějakého k-rozměrného afinního podprostoru $\mathbb{R}^n$ (tyto množiny ovšem netvoří σ-algebru!) máme tedy korektně definovaný její k-rozměrný objem $\mu_k(B)$.

Z vlastností λ_k nyní odvodíme některé vlastnosti μ_k^A. Předně je zřejmé, že $\mu_k^A(K) < \infty$, kdykoliv $K \subset A$ je kompaktní (takže μ_k^A je Radonova míra) a také to, že míra μ_k^A je σ-konečná.

Dále ukážeme, jak se jednoduše spočte k-rozměrná míra k-rozměrného rovnoběžnostěnu v A. Pojem rovnoběžnostěnu lze přirozeně definovat v libovolném vektorovém prostoru:

5.4 Definice. *Nechť je dán vektorový prostor V a jeho prvky $v_1, \dots, v_k$, a. Pak klademe*

$$\mathcal{R}_a(v_1, \dots, v_k) := \{a + t_1 v_1 + \dots + t_k v_k : \ 0 \le t_1 \le 1, \dots, 0 \le t_k \le 1\}.$$

Jsou-li $v_1, \dots, v_k$ lineárně nezávislé vektory, nazýváme $\mathcal{R}_a(v_1, \dots, v_k)$ k-rozměrným rovnoběžnostěnem.

Ještě zopakujme definici Gramovy matice a Gramova determinantu (gramiánu).

5.5 Definice. *Nechť V je unitární prostor a $v_1, \dots, v_k$ jsou jeho prvky. Pak definujeme Gramovu matici a Gramův determinant (gramián) vektorů $v_1, \dots, v_k$ takto:*

$$G(v_1, \dots, v_k) := (\langle v_i, v_j \rangle)_{i,j=1}^k, \qquad \Gamma(v_1, \dots, v_k) := \det G(v_1, \dots, v_k).$$

5.6 Tvrzení. *Nechť $v_1, \dots, v_k \in \mathbb{R}^n$, $k \le n$. Pak*

$$\Gamma(v_1, \dots, v_k) = \det\left([v_1, \dots, v_k]^T \cdot [v_1, \dots, v_k]\right).$$

Pokud navíc $k = n$, pak

$$\Gamma(v_1, \dots, v_k) = (\det[v_1, \dots, v_k])^2.$$

Vektory $v_1, \dots, v_k$ jsou lineárně závislé, právě když $\Gamma(v_1, \dots, v_k) = 0$.

Důkaz. Obě rovnosti plynou z toho, že $G(v_1, \dots, v_k) = [v_1, \dots, v_k]^T \cdot [v_1, \dots, v_k]$ a základů teorie determinantů. Poslední tvrzení je Věta 28.4. (i) z [Be].

Nejdříve odvodíme (z věty o substituci) vzorec pro Lebesgueovu míru rovnoběžnostěnu. Čtenáři, který nezná důkaz věty o substituci a chce vzorec pochopit, lze doporučit [Ru2; Věta 8.28] nebo [LM; Lemma 34.7].

5.7 Tvrzení. *Nechť $v_1, \dots, v_k, a \in \mathbb{R}^k$. Pak*

$$\lambda_k\left(\mathcal{R}_a(v_1, \dots, v_k)\right) = |\det[v_1, \dots, v_k]| = \sqrt{\Gamma(v_1, \dots, v_k)}. \tag{5.3}$$

Důkaz. Jsou-li vektory $v_1, \dots, v_k$ lineárně nezávislé, pak $\mathcal{R}_a(v_1, \dots, v_k)$ je k-rozměrný rovnoběžnostěn a zřejmě $\mathcal{R}_a(v_1, \dots, v_k) = f(I)$, kde $I := [0,1]^k$ a $f\colon \mathbb{R}^k \to \mathbb{R}^k$ je prosté afinní zobrazení dané předpisem

$$f(t_1, \dots, t_k) := a + t_1 v_1 + \dots + t_k v_k.$$

Pro Jacobiho matici difeomorfismu f v libovolném bodě $t \in \mathbb{R}^k$ máme zřejmě $[f'(t)] = [v_1, \dots, v_k]$, takže podle věty o substituci pro Lebesgueův integrál a Tvrzení 5.6 dostáváme

$$\lambda_k(\mathcal{R}_a(v_1, \dots, v_k)) = \int_{f(I)} 1 \, \mathrm{d}\lambda_k = \int_I |\det[v_1, \dots, v_k]| \, \mathrm{d}\lambda_k =$$
$$= |\det[v_1, \dots, v_k]| = \sqrt{\Gamma(v_1, \dots, v_k)}.$$

Jsou-li vektory $v_1, \dots, v_k$ lineárně závislé, pak $\mathcal{R}_a(v_1, \dots, v_k)$ je zřejmě podmnožina afinního prostoru dimenze menší než k, takže má nulovou Lebesgueovu míru. Všechny členy (5.3) jsou tedy nulové.

5.8 Věta. *Nechť $A \subset \mathbb{R}^n$ je k-rozměrný afinní prostor ($1 \le k < n$), $a \in A$, vektorový prostor $V \subset \mathbb{R}^n$ je zaměření A (tj. $A = a + V$) a $v_1, \dots, v_k \in V$. Pak*

$$\mu_k\left(\mathcal{R}_a(v_1, \dots, v_k)\right) = \sqrt{\Gamma(v_1, \dots, v_k)}. \tag{5.4}$$

Důkaz. Zvolme izometrii $F: \mathbb{R}^k \to A$. Podle Dodatku 6.7 c) existuje unitární bijekce $L: \mathbb{R}^k \to V$ taková, že $F(x) = F(0) + L(x)$, $x \in \mathbb{R}^k$. Položme $c := F^{-1}(a)$, $w_1 := L^{-1}(v_1), \dots, w_k := L^{-1}(v_k)$. Pak zřejmě $\mathcal{R}_a(v_1, \dots, v_k) = F\left(\mathcal{R}_c(w_1, \dots, w_k)\right)$, takže

$$\mu_k^A\left(\mathcal{R}_a(v_1, \dots, v_k)\right) = \lambda_k\left(\mathcal{R}_c(w_1, \dots, w_k)\right) = \sqrt{\Gamma(w_1, \dots, w_k)}.$$

Rovnost (5.4) pak vyplývá z toho, že $L: \mathbb{R}^k \to V$ je unitární zobrazení, takže $\langle v_i, v_j \rangle = \langle w_i, w_j \rangle$, $1 \le i, j \le k$, a tedy $\Gamma(w_1, \dots, w_k) = \Gamma(v_1, \dots, v_k)$.

V předchozím výkladu jsme mohli předpokládat, že A je k-rozměrný podprostor libovolného unitárního prostoru; takovou obecnost však nepotřebujeme. Vzorec (5.4) motivuje následující definici.

5.9 Definice. *Nechť V je unitární prostor. Pak pro vektory $v_1, \dots, v_k$ z V klademe*

$$\operatorname{vol}(v_1, \dots, v_k) := \sqrt{\Gamma(v_1, \dots, v_k)}.$$

Za předpokladů Věty 5.8 lze pak rovnost (5.4) zapsat ve tvaru

$$\mu_k\left(\mathcal{R}_a(v_1, \dots, v_k)\right) = \operatorname{vol}(v_1, \dots, v_k). \tag{5.5}$$

Jsou-li $v_1, \dots, v_k$ prvky $\mathbb{R}^k$, pak podle Tvrzení 5.7 platí

$$\operatorname{vol}(v_1, \dots, v_k) = |\det[v_1, \dots, v_k]|. \tag{5.6}$$

Důležitou skutečností je existence „koeficientu změny k-rozměrné míry“ při afinním zobrazení. Nejdříve uvažujme afinní bijekci $F: \mathbb{R}^k \to \mathbb{R}^k$. Lze psát $F(x) = F(0) + L(x)$, kde $L: \mathbb{R}^k \to \mathbb{R}^k$ je lineární bijekce. Protože zřejmě $J_F(x) = \det[L]$, $x \in \mathbb{R}^k$, z věty o substituci pro Lebesgueův integrál vyplývá, že pro každou borelovskou množinu $B \subset \mathbb{R}^n$ platí

$$\lambda_n\left(F(B)\right) = |\det[L]| \cdot \lambda_n(B).$$

Zobrazení F tedy buď vůbec nemění míru borelovské množiny (je-li $|\det[L]| = 1$) nebo ji „násobí“ jistým konstantním kladným koeficientem, totiž číslem $\kappa(F) := |\det[L]| > 0$. (Pro změnu vzdáleností bodů nebo změnu úhlů taková zákonitost při $k \ge 2$ neplatí!)

5.10 Věta. *Nechť $T \subset \mathbb{R}^n$, $Z \subset \mathbb{R}^m$ jsou k-rozměrné afinní prostory a V_T, V_Z jsou jejich zaměření. Nechť $f: T \to Z$ je afinní zobrazení tvaru $f(t) = z_0 + L(t - t_0)$, kde $t_0 \in T$, $z_0 \in Z$ a $L: V_T \to V_Z$ je lineární zobrazení. Pak platí následující tvrzení:*

(i) *Existuje jediné číslo $\kappa(f) \geq 0$ takové, že pro každou kompaktní množinu $E \subset T$ platí*

$$\mu_k^Z(f(E)) = \kappa(f) \cdot \mu_k^T(E). \tag{5.7}$$

(ii) *Zobrazení f je prosté právě když $\kappa(f) > 0$. V tom případě platí rovnost (5.7) pro každou borelovskou množinu $E \subset T$.*

(iii) *Je-li $(v_1, \ldots, v_k)$ báze prostoru V_T, pak*

$$\kappa(f) = \frac{\operatorname{vol}(L(v_1), \ldots, L(v_k))}{\operatorname{vol}(v_1, \ldots, v_k)}, \tag{5.8}$$

(iv)
$$\kappa(f) = \kappa(L).$$

Důkaz. Nejprve předpokládejme, že f je prosté; pak f je homeomorfismus. Uvažujme izometrie $\varphi: \mathbb{R}^k \to T$, $\psi: \mathbb{R}^k \to Z$ a borelovskou množinu $E \subset T$. Pak $\mu_k^T(E) = \lambda_k(\varphi^{-1}(E))$ a $\mu_k^Z(f(E)) = \lambda_k(\psi^{-1}(f(E)))$. Zobrazení $F := \psi^{-1} \circ f \circ \varphi$ je zřejmě afinní bijekce $\mathbb{R}^k$ na $\mathbb{R}^k$, takže podle úvahy před větou existuje číslo $\kappa(F) > 0$ takové, že $\lambda_k(F(B)) = \kappa(F)\lambda_k(B)$ pro každou $B \in \mathcal{B}(\mathbb{R}^k)$. Pro $B := \varphi^{-1}(E)$ dostáváme $\lambda_k(\varphi^{-1}(E)) = \kappa(F)\lambda_k(\psi^{-1}(f(E)))$. Položíme-li tedy $\kappa(f) := \kappa(F)$, vidíme, že (5.7) platí pro každou $B \in \mathcal{B}(\mathbb{R}^k)$.

Pokud f není bijekce, je $f(T)$ afinní podprostor prostoru Z dimenze menší než k, takže se snadno ukáže, že $\mu_k^Z(f(T)) = 0$, a vzorec (5.7) tedy platí pro každou kompaktní $E \subset T$ pro $\kappa(f) := 0$ ($f(E)$ je zřejmě kompaktní, a tedy i borelovská). Dokázali jsme tedy (i) a (ii) (kromě jednoznačnosti $\kappa(f)$).

Je-li $(v_1, \ldots, v_k)$ báze prostoru V_T, pak zřejmě platí rovnost $f(\mathcal{R}_{t_0}(v_1, \ldots, v_k)) = \mathcal{R}_{z_0}(L(v_1), \ldots, L(v_k))$, takže

$$\begin{aligned}\operatorname{vol}(L(v_1), \ldots, L(v_k)) &= \mu_k^Z(\mathcal{R}_{z_0}(L(v_1), \ldots, L(v_k))) \\ &= \kappa(f) \cdot \mu_k^T(\mathcal{R}_{t_0}(v_1, \ldots, v_k)) = \kappa(f) \cdot \operatorname{vol}(v_1, \ldots, v_k).\end{aligned}$$

Okamžitě tedy dostáváme (iii) a tudíž i jednoznačnost čísla $\kappa(f)$. Protože L je také afinní zobrazení mezi k-rozměrnými afinními prostory, podle (iii) platí rovnost

$$\operatorname{vol}(L(v_1), \ldots, L(v_k)) = \kappa(L) \cdot \operatorname{vol}(v_1, \ldots, v_k),$$

ze které pak ihned vyplývá (iv).

5.11 Poznámka.

(i) Je-li zobrazení L z předchozí věty unitární, z (5.8) a Definice 5.9 vyplývá, že $\kappa(f) = \kappa(L) = 1$.

(ii) Pokud f z předchozí věty není prosté, může se stát, že obraz $f(E)$ borelovské množiny není borelovský. Kdybychom však místo μ_k^T a μ_k^Z uvažovali jejich zúplnění, mohli bychom psát (5.7) pro libovolnou borelovskou (dokonce μ_k^T-měřitelnou) množinu $E \subset T$ i pro f, které není prosté.

Nyní určíme koeficient změny míry κ_f v jednom důležitém speciálním případě.

5.12 Tvrzení. *Nechť $f: \mathbb{R}^k \to \mathbb{R}^n$ je prosté afinní zobrazení tvaru $f = a + L$, kde $L: \mathbb{R}^k \to \mathbb{R}^n$ je lineární zobrazení a $a \in \mathbb{R}^n$. Pak $Z := f(\mathbb{R}^k)$ je k-rozměrný afinní prostor a platí*

$$\kappa(f) = \kappa(L) = \operatorname{vol}(L(e_1), \ldots, L(e_k)) = \sqrt{\Gamma(L(e_1), \ldots, L(e_k))}. \tag{5.9}$$

V případě $k = n$ *platí* $\kappa(f) = |\det[L]|$.

Důkaz. Stačí užít vzorec (5.8) na kanonickou bázi, tj. pro $v_i = e_i$, $i = 1, \dots, k$. Případ $k = n$ jsme již diskutovali (lze též užít Tvrzení 5.6).

Následující tvrzení, které je téměř zřejmé, nám v dalším umožní zkrátit a vyjasnit některé výpočty.

5.13 Tvrzení. *Nechť* $T \subset \mathbb{R}^n$, $Z \subset \mathbb{R}^m$, $S \subset \mathbb{R}^p$ *jsou* k*-rozměrné afinní prostory a nechť* $f: T \to Z$ *a* $g: Z \to S$ *jsou afinní zobrazení. Pak*

$$\kappa(g \circ f) = \kappa(g) \cdot \kappa(f). \tag{5.10}$$

Důkaz. Nechť $K \subset T$ je kompaktní množina s kladnou mírou $\mu_k^T(K)$. Množiny $f(K)$ i $g(f(K))$ jsou kompaktní, takže platí $\mu_k^S(g(f(K))) = \kappa(g \circ f)\mu_k^T(K)$ a $\mu_k^S(g(f(K))) = \kappa(g) \cdot \mu_k^Z(f(K)) = \kappa(g) \cdot \kappa(f) \cdot \mu_k^T(K)$. Z těchto rovností (5.10) ihned plyne.

5.4 Definice k-rozměrné míry na jednoduché k-ploše

Je-li P křivá k-plocha, je přesná definice k-rozměrné míry μ_k^P na P podstatně obtížnější. Každému je jasné, jak se měří obsah křivé plochy v praxi:

Chceme-li změřit velikost plochy velké parcely v hornatém terénu, rozdělíme si ji na malé části, které již nerozeznáme od rovinných útvarů, jejich plochy změříme a výsledky sečteme.

Tento praktický návod odpovídá základní ideji diferenciálního počtu, která se někdy vyjadřuje heslem, že hladká funkce (resp. plocha) je „v malém“ afinní.

Podat přesnou konstruktivní definici na základě tohoto návodu však není snadné. Proto tento postup pouze použijeme k heuristickému „odvození“ vzorce pro výpočet k-rozměrné míry $\mu_k(P)$ jednoduché k-plochy P. Nechť $\varphi: G \to \mathbb{R}^n$ je parametrizace P. Rozdělme nyní otevřenou množinu $G \subset \mathbb{R}^k$ na konečně (nebo spočetně) mnoho (borelovských, po dvou disjunktních) částí $G_1, G_2, \dots$ tak, že jejich diametry i diametry jejich obrazů $P_j := \varphi(G_j)$ jsou velmi malé. V každé množině G_j zvolme bod $\xi_j \in G_j$ a uvažujme afinní zobrazení $\alpha_j: \mathbb{R}^k \to \mathbb{R}^n$ dané předpisem $\alpha_j(t) := \varphi(\xi_j) + \varphi'(\xi_j)(t - \xi_j)$. (Připomeňme, že afinní zobrazení $\alpha_j(t)$ „velmi dobře“ aproximuje φ v blízkosti bodu ξ_j a jeho oborem hodnot je afinní tečný prostor $T^{\mathrm{af}}_{\varphi(\xi_j)}(P)$). Za našich předpokladů je tedy přirozené předpokládat, že $\varphi(G_j)$ a $\alpha_j(G_j)$ si jsou „k nerozeznání podobné“ (viz obr. 5.18, kde $k = 2, n = 3, G_j$ je

OBR. 5.18.

uzavřený čtverec a ξ_j je jeho vrchol), takže se zdá být jasné, že

$$\mu_k(\varphi(G_j)) \approx \mu_k(\alpha_j(G_j)) = \kappa(\alpha_j) \cdot \lambda_k(G_j) = \kappa(\varphi'(\xi_j)) \cdot \lambda_k(G_j). \tag{5.11}$$

(Poslední rovnost vyplývá z Věty 5.10 (iv).) Také se zdá, že sečtením přes j dostaneme

$$\mu_k(P) = \sum \mu_k(P_j) = \sum \mu_k(\varphi(G_j)) \approx \sum \kappa(\varphi'(\xi_j)) \cdot \lambda_k(G_j).$$

Je tedy přirozené se domnívat, že

$$\mu_k(P) = \int_G \kappa(\varphi'(t))\,\mathrm{d}t = \int\limits_G \sqrt{\Gamma\left(\frac{\partial\varphi(t)}{\partial t_1}, \ldots, \frac{\partial\varphi(t)}{\partial t_k}\right)}\,\mathrm{d}t. \tag{5.12}$$

(Poslední rovnost vyplývá z (5.9), protože $\frac{\partial\varphi(t)}{\partial t_i} = \varphi'(t)(e_i)$.)

Přesně definovat, co je to k-rozměrná míra $\mu_k(B)$ pro dostatečně obecné (borelovské) podmnožiny $\mathbb{R}^n$ a opravdu *dokázat*, že platí vzorec (5.12) (který je pro konkrétní výpočty nezbytný) je značně obtížný úkol (srov. Dodatek 6.3.). Nabízí se tedy myšlenka přijmout vzorec (5.12) *za definici* čísla $\mu_k(P)$.

My však chceme definovat nejen $\mu_k(P)$, ale i borelovskou míru μ_k^P na P. Vzorec (5.12) nás vede k tomu, že každé borelovské množině $B \subset P$ přiřadíme číslo

$$\mu_k^P(B) := \int_{\varphi^{-1}(B)} \kappa(\varphi'(t))\,\mathrm{d}t.$$

Nejprve ovšem musíme dokázat nezávislost na parametrizaci.

5.14 Tvrzení. *Nechť P je jednoduchá k-plocha v $\mathbb{R}^n$ a $\varphi: G \to \mathbb{R}^n, \psi: H \to \mathbb{R}^n$ jsou její dvě parametrizace. Pak pro každou borelovskou množinu $B \subset P$ platí*

$$\int_{\varphi^{-1}(B)} \kappa(\varphi'(t))\,\mathrm{d}t = \int_{\psi^{-1}(B)} \kappa(\psi'(t))\,\mathrm{d}t.$$

Důkaz. Ze spojitosti φ vyplývá, že $\varphi^{-1}(B)$ je borelovská podmnožina $\mathbb{R}^k$, takže integrál nalevo existuje: je to integrál ze spojité (a tedy borelovsky měřitelné) nezáporné funkce přes měřitelnou množinu. Podle Tvrzení 2.164 existuje difeomorfismus $\omega: G \to H$ takový, že $\varphi = \psi \circ \omega$. Pak pro každý bod $t \in G$ platí $\varphi'(t) = \psi'(\omega(t)) \circ \omega'(t)$, takže podle Tvrzení 5.13 a Tvrzení 5.12 dostáváme

$$\kappa(\varphi'(t)) = \kappa(\psi'(\omega(t))) \cdot \kappa(\omega'(t)) = \kappa(\psi'(\omega(t))) \cdot |J_\omega(t)|.$$

Podle věty o substituci pro Lebesgueův integrál tudíž platí

$$\int_{\varphi^{-1}(B)} \kappa(\varphi'(t))\,\mathrm{d}t = \int_{\omega^{-1}(\psi^{-1}(B))} \kappa(\psi'(\omega(t))) \cdot |J_\omega(t)|\,\mathrm{d}t = \int_{\psi^{-1}(B)} \kappa(\psi'(t))\,\mathrm{d}t.$$

Následující definice je tedy korektní.

5.15 Definice. *Nechť $P \subset \mathbb{R}^n$ je jednoduchá k-rozměrná plocha a $\varphi: G \to P$ je její parametrizace. Pak borelovskou k-rozměrnou míru μ_k^P na ploše P definujeme rovností*

$$\mu_k^P(B) := \int\limits_{\varphi^{-1}(B)} \kappa(\varphi'(t))\,\mathrm{d}t = \int_{\varphi^{-1}(B)} \sqrt{\Gamma\left(\frac{\partial\varphi(t)}{\partial t_1}, \dots, \frac{\partial\varphi(t)}{\partial t_k}\right)}\,\mathrm{d}t \tag{5.13}$$

pro každou borelovskou množinu $B \subset P$.

Musíme ovšem ověřit, že množinová funkce, kterou jsme na $\mathcal{B}(P)$ definovali vzorcem (5.13), je skutečně míra. To snadno vyplývá ze známých vlastností Lebesgueova integrálu. (Můžeme se však také odvolat na to, že μ_k^P je zřejmě obrazem $\varphi\left(\kappa(\varphi') \cdot \lambda_k^b\right)$ borelovské míry $\kappa(\varphi') \cdot \lambda_k^b$ při zobrazení φ; srov. Věty 6.32 a 6.31.)

Z Tvrzení 5.12 také snadno vyplývá, že právě definovaný pojem k-rozměrné míry na jednoduché k-ploše je zobecněním pojmu k-rozměrné míry na k-rozměrném afinním prostoru.

5.16 Poznámka. Nechť P, G, φ jsou jako v Definici 5.15. Protože funkce $\kappa(\varphi'(t))$ je spojitá na G a φ je homeomorfismus, lze zřejmě ke každému bodu $x \in P$ najít jeho otevřené okolí U_x takové, že množina $\varphi^{-1}(U_x)$ je omezená a $\kappa(\varphi'(t))$ je na ní omezená, takže $\mu_k^P(U_x \cap P) < \infty$. Míra μ_k^P je tedy lokálně konečná a pomocí Poznámky 1.69 snadno dostáváme, že míra μ_k^P je σ-konečná.

5.5 Definice k-rozměrné míry na „minimální“ σ-algebře $\mathcal{P}_k$

V předchozím oddíle jsme definovali míru μ_k^P na každé jednoduché ploše $P \subset \mathbb{R}^n$. Již pro nejjednodušší klasické aplikace je však důležité umět integrovat i přes k-plochy, které nejsou jednoduché (jako je například jednotková sféra v $\mathbb{R}^3$), a také přes „po částech hladké plochy“ (jako je například povrch krychle nebo kužele). Pro klasický přístup k tomuto problému viz např. [Zo] nebo [Kop] (srov. Poznámka 5.26 (ii)). Obecnější přístup (srov. Dodatek 6.3) pracuje s k-rozměrnými mírami definovanými na σ-algebře všech borelovských podmnožin $\mathbb{R}^n$ (nejčastěji s Hausdorffovou mírou). My zde budeme pracovat s mírou μ_k na nejmenší σ-algebře $\mathcal{P}_k^n$, která stačí pro klasické aplikace.

5.17 Definice. *Nechť $1 \le k < n$. Pak symbolem $\mathcal{P}_k^n$ označme nejmenší σ-algebru podmnožin $\mathbb{R}^n$, která obsahuje každou borelovskou podmnožinu každé jednoduché k-rozměrné plochy v $\mathbb{R}^n$. Je-li zřejmá hodnota n, budeme místo $\mathcal{P}_k^n$ psát pouze $\mathcal{P}_k$.*

Není těžké ukázat, že $\mathcal{P}_k$ je nejmenší σ-algebra, která obsahuje každou jednoduchou k-plochu v $\mathbb{R}^n$; menší σ-algebru proto jistě nemá smysl uvažovat.

Následující větu lze chápat jako „axiomatickou definici“ k-rozměrné míry na σ-algebře $\mathcal{P}_k$. Tato definice je sice heuristickým odvozením vzorce (5.12) dobře motivovaná, ale pro svou složitost není zcela uspokojivá. (Elegantnější axiomatická definice, s kterou je ale obtížné pracovat, je obsažena ve Větě 6.18.)

5.18 Věta. *Nechť $1 \le k < n$. Na σ-algebře $\mathcal{P}_k$ existuje právě jedna míra μ_k, pro kterou platí:*

Je-li P jednoduchá k-plocha v $\mathbb{R}^n$, φ je parametrizace P a B je borelovská podmnožina P, pak

$$\mu_k(B) := \int_{\varphi^{-1}(B)} \kappa(\varphi'(t))\ \mathrm{d}t. \tag{5.14}$$

Důkaz věty, který je elementární, ale trochu zdlouhavý, je v Dodatku 6.4.

5.19 Poznámka. Platnost vzorce (5.14) je ovšem ekvivalentní požadavku, aby μ_k rozšiřovala každou míru μ_k^P (kde P je jednoduchá k-plocha).

V Dodatku 6.4 je plně popsána struktura σ-algebry $\mathcal{P}_k$, pro aplikace však stačí znát pouze Větu 5.18. V obvyklých aplikacích totiž vždy integrujeme přes „po částech hladké plochy“ P, které lze psát ve tvaru

$$P = \bigcup_{i=1}^{u} P_i \ \cup\ \bigcup_{i=1}^{v} Q_i, \tag{5.15}$$

kde $P_1, \dots, P_u$ jsou po dvou disjunktní jednoduché k-plochy a pro každé $i \in \{1, \dots, v\}$ je Q_i k_i-plocha, kde $0 \le k_i < k$. Z Tvrzení 2.156 a Tvrzení 5.20 pak snadno vyplývá, že každá borelovská množina $B \subset P$ patří do $\mathcal{P}_k$ a její k-rozměrnou míru lze počítat pomocí vzorce

$$\mu_k(B) = \mu_k(P_1 \cap B) + \cdots + \mu_k(P_u \cap B).$$

5.20 Tvrzení. *Nechť $0 \le s < k < n$, a nechť $Q \subset \mathbb{R}^n$ je s-plocha. Pak $Q \in \mathcal{P}_k$ a $\mu_k(Q) = 0$.*

Důkaz. (Stručný.) Zřejmě můžeme předpokládat $s \neq 0$. Podle Tvrzení 2.156 je Q borelovská množina. Nejprve předpokládejme, že Q je explicitně zadaný kus s-rozměrné C^1 plochy. Bez újmy na obecnosti můžeme předpokládat, že

$$Q = \{(x_1, \dots, x_n)\colon\ x_{s+1} = f_{s+1}(x_1, \dots, x_s), \dots, x_n = f_n(x_1, \dots, x_s)\},$$

kde $f_{s+1}, \dots, f_n$ jsou funkce třídy C^1 na otevřené množině $G \subset \mathbb{R}^s$. Položme

$$P := \{(x_1, \dots, x_n)\colon\ x_{k+1} = f_{k+1}(x_1, \dots, x_s), \dots, x_n = f_n(x_1, \dots, x_s)\}.$$

Pak P je zřejmě explicitně zadaný kus k-rozměrné C^1 plochy s „přirozenou parametrizací“

$$\varphi(t_1, \dots, t_k) := (t_1, \dots, t_k, f_{k+1}(t_1, \dots, t_s), \dots, f_n(t_1, \dots, t_s)), \qquad t \in G \times \mathbb{R}^{k-s}.$$

Přitom zřejmě $Q = \varphi(T)$, kde

$$T := \{(t_1, \dots, t_k) \in G \times \mathbb{R}^{k-s}\colon\ t_{s+1} = f_{s+1}(t_1, \dots, t_s), \dots, t_k = f_k(t_1, \dots, t_s)\}$$

je explicitně zadaný kus s-rozměrné C^1 plochy v $\mathbb{R}^k$. Podle Tvrzení 2.156 máme $\lambda_k(T) = 0$, takže podle (5.14) $\mu_k(Q) = \mu_k^P(Q) = 0$.

Obecný případ pak snadno dostáváme pomocí Tvrzení 2.158.

5.21 Příklad. Spočtěme plošný obsah $\mu_2(P)$ jednotkové sféry $P = \{(x, y, z)\colon\ x^2 + y^2 + z^2 = 1\}$. Položíme-li (viz Příklad 2.145)

$$G := (0, \pi) \times (0, 2\pi), \quad \Phi(\theta, \varphi) := (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta), \quad P_1 := \Phi(G),$$

je Φ parametrizace jednoduché 2-plochy P_1 a $P = P_1 \cup Q$, kde $Q = \{(x, y, z)\colon\ x^2 + z^2 = 1,\ y = 0\}$. Je snadné ověřit, že Q je 1-plocha. Platí tedy

$$\mu_2(P) = \mu_2(P_1) = \int_G \kappa\left(\Phi'(\theta, \varphi)\right)\ \mathrm{d}\theta\mathrm{d}\varphi.$$

Snadný výpočet dává

$$\begin{aligned}\kappa\left(\Phi'(\theta, \varphi)\right) &= \sqrt{\Gamma\left(\frac{\partial\Phi}{\partial\theta}(\theta, \varphi), \frac{\partial\Phi}{\partial\varphi}(\theta, \varphi)\right)} \\ &= \sqrt{\Gamma\left((\cos\theta\cos\varphi, \cos\theta\sin\varphi, -\sin\theta), (-\sin\theta\sin\varphi, \sin\theta\cos\varphi, 0)\right)} = \sin\theta,\end{aligned}$$

takže $\mu_2(P) = \int_G \sin\theta\ \mathrm{d}\theta\mathrm{d}\varphi = 2\pi \int_0^\pi \sin\theta\ \mathrm{d}\theta = 4\pi.$

5.6 Definice a výpočet plošného integrálu 1. druhu

5.22 Definice. *Nechť $1 \leq k < n$, $M \in \mathcal{P}_k$ a f je funkce z $\mathbb{R}^n$ do $\mathbb{R}^*$. Potom (k-rozměrný) plošný integrál prvního druhu $\int_M f \, \mathrm{dS}_k$ funkce f přes množinu M definujeme rovností*

$$\int_M f \, \mathrm{dS}_k := \int_M f \, \mathrm{d}\mu_k, \tag{5.16}$$

konverguje-li integrál napravo. Je-li zřejmé, jaká je hodnota k, píšeme místo $\int_M f \, \mathrm{dS}_k$ pouze $\int_M f \, \mathrm{dS}$.

5.23 Poznámka.

(a) Pravá strana rovnosti (5.16) je definovaná také v některých případech, kdy f není definovaná na celé množině M (srov. Dodatek 6.6).

(b) Někdy budeme místo $\int_M f \, \mathrm{dS}$ psát $\int_M f(x) \, \mathrm{dS}(x)$.

V konkrétních případech počítáme plošný integrál prvního druhu tak, že jej převádíme na integrál podle Lebesgueovy míry, který pak počítáme běžnými způsoby (užitím Fubiniho věty a věty o substituci). K tomu nám slouží následující věta.

5.24 Věta. *Nechť $1 \leq k < n$. Nechť $P \subset \mathbb{R}^n$ je jednoduchá k-plocha, $\varphi\colon G \to \mathbb{R}^n$ je její parametrizace a $f\colon P \to \mathbb{R}^*$ je borelovsky měřitelná funkce. Pak*

$$\int_P f \, \mathrm{dS} = \int_G (f \circ \varphi)(t) \cdot \kappa(\varphi'(t)) \, \mathrm{d}t = \int\limits_G (f \circ \varphi)(t) \cdot \sqrt{\Gamma\left(\frac{\partial \varphi(t)}{\partial t_1}, \dots, \frac{\partial \varphi(t)}{\partial t_k}\right)} \, \mathrm{d}t,$$

konverguje-li jeden ze tří integrálů.

Důkaz. Podle Poznámky 5.19 platí $\displaystyle\int_P f \, \mathrm{d}S = \int_P f \, \mathrm{d}\mu_k^P$. Protože platí rovnost $\mu_k^P = \varphi(\kappa(\varphi') \cdot \lambda_k^b))$ (kde $\kappa(\varphi')\colon x \mapsto \kappa(\varphi'(x))$), podle věty o integraci podle obrazu míry (Věta 6.31) platí

$$\int_P f \, \mathrm{d}S = \int_G (f \circ \varphi) \, \mathrm{d}(\kappa(\varphi') \cdot \lambda_k^b),$$

konverguje-li jeden z integrálů. Podle věty o integraci podle neurčitého integrálu (Věta 6.32) platí

$$\int_G (f \circ \varphi) \, \mathrm{d}(\kappa(\varphi') \cdot \lambda_k^b) = \int_G (f \circ \varphi) \cdot \kappa(\varphi') \, \mathrm{d}\lambda_k,$$

má-li jedna strana smysl. (Zde jsme použili to, že funkce $(f \circ \varphi) \cdot \kappa(\varphi')$ je borelovsky měřitelná na G.)

Pro základní aplikace plošného integrálu 1. druhu si stačí pamatovat Větu 5.24 a následující tvrzení (které je snadným důsledkem Tvrzení 5.20).

5.25 Tvrzení. *Nechť P je „po částech hladká plocha“ tvaru (5.15) a f je reálná funkce definovaná aspoň na $P_1 \cup \dots \cup P_u$, pak*

$$\int_P f\,\mathrm{dS} = \int_{P_1} f\,\mathrm{dS} + \dots + \int_{P_u} f\,\mathrm{dS} \tag{5.17}$$

má-li jedna strana smysl.

Integrály napravo již totiž můžeme počítat podle Věty 5.24.

5.26 Poznámka.

(i) V aplikacích funkce f není někdy definovaná na celé množině P (ale pouze μ_k-skoro všude na P). Vyjádření (5.15) lze však zpravidla zvolit tak, aby f byla definována na $P_1 \cup \dots \cup P_u$.

(ii) Klasicky se definuje plošný integrál 1. druhu přes „po částech hladkou plochu“ $P \subset \mathbb{R}^n$ pomocí vzorce (5.17). Pak je ovšem nutno dokázat nezávislost na vyjádření (5.15). Pokud se připouštějí jen vyjádření speciálního typu (srov. definice zobecněné plochy v [Kop]), je tento důkaz jednodušší než důkaz Věty 5.18 (není nutno dokazovat Lemma 6.21).

5.27 Příklad. Uvažujme jednotkovou sféru $P = \{(x,y,z)\colon x^2+y^2+z^2 = 1\}$ jako hmotnou plochu s plošnou hustotou 1 a zkoumejme jí vytvořené gravitační pole F v bodě $a = (0,0,v), 0 < v < 1$. Rozložení uvažované hmoty je zřejmě popsáno mírou $\mu = \mu^P = C_P \cdot \mu_2$, takže složka $F_z(a)$ vektoru $F(a) = (F_x(a), F_y(a), F_z(a))$ je podle (6.12) (za předpokladu, že gravitační konstanta je 1) dána vzorcem

$$F_z(a) = \int_P \frac{z-v}{(x^2+y^2+(z-v)^2)^{3/2}}\,\mathrm{d}S(x,y,z).$$

Ze symetrie „je vidět“ (a snadno se spočte), že $F_x(a) = F_y(a) = 0$. Při vyjádření $P = P_1 \cup Q$ a parametrizaci Φ plochy P_1 jako v Příkladu 5.21 dostáváme

$$\begin{aligned} F_z(a) &= \int_{P_1} \frac{z-v}{(x^2+y^2+(z-v)^2)^{3/2}}\,\mathrm{d}S(x,y,z) = \int_G \frac{\cos\theta - v}{\left(\sin^2\theta + (\cos\theta - v)^2\right)^{3/2} \cdot \sin\theta}\,\mathrm{d}\theta\mathrm{d}\varphi \\ &= 2\pi \int_0^\pi \frac{\cos\theta - v}{(1+v^2-2v\cos\theta)^{3/2}} \cdot \sin\theta\,\mathrm{d}\theta. \end{aligned}$$

Pomocí substituce $\sqrt{1+v^2-2v\cos\theta} = y$ dostáváme $F_z(a) = 0$. Ze symetrie se zdá být zřejmé (a lze přesně dokázat), že gravitační pole uvnitř jednotkové koule je nulové. Pro $v > 1$ stejný postup dává, že gravitační pole v bodě a je stejné, jako kdyby všechna hmotnost plochy P byla soustředěna v počátku. (Je zřejmé, jak formulovat analogický výsledek pro elektrostatické pole.)

5.7 Vektorový součin

5.28 Definice. *Nechť $u^1, \dots, u^{n-1}$ jsou prvky $\mathbb{R}^n$ $(n > 1)$. Pak definujeme vektorový součin $u^1 \times \cdots \times u^{n-1}$ takto:*

$$u^1 \times \cdots \times u^{n-1} := (\det[e_1, u^1, \dots, u^{n-1}], \dots, \det[e_n, u^1, \dots, u^{n-1}]). \tag{5.18}$$

Vektorovému součinu se také někdy říká „vnější součin“.

5.29 Poznámka.

(i) Vektorový součin chápaný jako operace je tedy zobrazení $\times\colon (\mathbb{R}^n)^{n-1} \to \mathbb{R}^n$. Pokud $n = 2$, budeme vektorový součin jediného prvku u^1 zapisovat symbolem $\times(u_1)$. (V literatuře se nejčastěji používá zápis $u^1 \times u^2$ pro vektorový součin dvou vektorů; jinak jsou častější zápisy $[u^1, \dots, u^{n-1}]$ nebo $\langle u^1, \dots, u^{n-1}\rangle$, které však kolidují s naší symbolikou.)

(ii) Definice vektorového součinu se někdy zapisuje (a dobře pamatuje) pomocí „symbolického determinantu“ (kde $u^i = (u^i_1, \dots, u^i_n)$):

$$u^1 \times \cdots \times u^{n-1} = \begin{vmatrix} e_1 & u^1_1 & \dots & u^{n-1}_1 \\ \vdots & \vdots & \ddots & \vdots \\ e_n & u^1_n & \dots & u^{n-1}_n \end{vmatrix}$$

Jeho „formální rozvoj“ podle prvního sloupce dává (5.18), například

$$\times(1,3) = \begin{vmatrix} e_1 & 1 \\ e_2 & 3 \end{vmatrix} = 3 \cdot e_1 - 1 \cdot e_2 = (3, -1),$$

$$(2,1,1) \times (1,0,3) = \begin{vmatrix} e_1 & 2 & 1 \\ e_2 & 1 & 0 \\ e_3 & 1 & 3 \end{vmatrix} = (3, -5, -1).$$

(iii) Vidíme, že i-tou souřadnici vektorového součinu $u^1 \times \cdots \times u^{n-1}$ vypočteme tak, že číslo $(-1)^{i+1}$ znásobíme determinantem matice, kterou dostaneme, vyškrtneme-li i-tý řádek z matice $[u^1, \dots, u^{n-1}]$. Z toho snadno dostáváme, že vektorový součin $\times\colon (\mathbb{R}^n)^{n-1} \to \mathbb{R}^n$ je spojité zobrazení. Z vlastností determinantů snadno plyne, že je to $(n-1)$-lineární antisymetrické zobrazení.

(iv) V literatuře se vyskytuje i odlišná definice vektorového součinu, ve které se vektory e_i „kladou nakonec“, takže například jeho první složka je $\det[u^1, \dots, u^{n-1}, e_1]$. V nejběžnějším případě $n = 3$ pak obě definice dávají stejný pojem, pro $n = 2$ se však liší a dávají k sobě opačné vektory.

Následující dvě tvrzení dávají alternativní definice vektorového součinu.

5.30 Tvrzení. *Nechť $u^1, \dots, u^{n-1}, w \in \mathbb{R}^n$ $(n > 1)$. Pak $w = u^1 \times \cdots \times u^{n-1}$, právě když pro každý vektor $v \in \mathbb{R}^n$ platí*

$$\langle v, w\rangle = \det[v, u^1, \dots, u^{n-1}]. \tag{5.19}$$

Důkaz. Protože na levé i pravé straně (5.19) jsou lineární formy (proměnné v), (5.19) platí pro všechna $v \in \mathbb{R}^n$, právě když platí pro $v = e_i$, $i = 1, \dots, n$, což je ekvivalentní s (5.18).

5.31 Poznámka. Přímo z definice nebo z předchozího tvrzení snadno vidíme, že $u^1 \times \cdots \times u^{n-1} = 0$, právě když jsou vektory $u^1, \dots, u^{n-1}$ lineárně závislé.

5.32 Tvrzení. *Nechť $u^1, \dots, u^{n-1}, w \in \mathbb{R}^n$ $(n > 1)$. Pak $w = u^1 \times \cdots \times u^{n-1}$, právě když platí tyto podmínky:*

(i) *w je kolmý na každý z vektorů $u^1, \dots, u^{n-1}$.*
(ii) *$\|w\| = \operatorname{vol}(u^1, \dots, u^{n-1})$.*
(iii) *Jsou-li vektory $u^1, \dots, u^{n-1}$ lineárně nezávislé, pak $\det[w, u^1, \dots, u^{n-1}] > 0$.*

Důkaz. Jsou-li $u^1, \dots, u^{n-1}$ lineárně závislé, pak (i) – (iii) splňuje zřejmě pouze vektor $w = u^1 \times \cdots \times u^{n-1} = 0$.

Předpokládejme tedy, že $u^1, \dots, u^{n-1}$ jsou lineárně nezávislé. Pak vlastnosti (*i*) a (*ii*) mají zřejmě právě dva k sobě opačné nenulové vektory z, $-z$, z nichž právě jeden splňuje i podmínku (iii). Stačí tedy dokázat, že $w := u^1 \times \cdots \times u^{n-1}$ splňuje (i), (ii), (iii). Užijeme-li (5.19) na $v = u^i$, $i = 1, \dots, n-1$, dostáváme platnost (i). Pro důkaz (ii) vyjádříme $\operatorname{vol}(w, u^1, \dots, u^{n-1})$ dvěma způsoby. Podle (5.19) a (5.6) máme

$$\operatorname{vol}(w, u^1, \dots, u^{n-1}) = \langle w, w \rangle = \|w\|^2.$$

Podle Definice 5.9 platí

$$\operatorname{vol}^2(w, u^1, \dots, u^{n-1}) = \begin{vmatrix} \langle w, w \rangle & 0 & \dots & 0 \\ 0 & \langle u^1, u^1 \rangle & \dots & \langle u^1, u^{n-1} \rangle \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \langle u^{n-1}, u^1 \rangle & \dots & \langle u^{n-1}, u^{n-1} \rangle \end{vmatrix}$$

$$= \|w\|^2 \cdot \Gamma(u^1, \dots, u^{n-1}) = \|w\|^2 \cdot \operatorname{vol}^2(u^1, \dots, u^{n-1}),$$

takže $\operatorname{vol}(w, u^1, \dots, u^{n-1}) = \|w\| \cdot \operatorname{vol}(u^1, \dots, u^{n-1})$. Protože $w \neq 0$ (viz Poznámka 5.31), z těchto dvou vyjádření dostáváme (ii). Tvrzení (iii) vyplývá z rovnosti $\langle w, w \rangle = \det[w, u^1, \dots, u^{k-1}]$ (viz (5.19)).

5.33 Poznámka.

(a) Geometrický smysl nerovnosti $\det[w, u^1, \dots, u^{n-1}] > 0$ je ten, že báze $w, u^1, \dots, u^{n-1}$ je orientovaná souhlasně s kanonickou bazí $e_1, \dots, e_n$, srov. Dodatek 6.2.
(b) Rovnost $\operatorname{vol}(w, u^1, \dots, u^{n-1}) = \|w\| \cdot \operatorname{vol}(u^1, \dots, u^{n-1})$ je „geometricky zřejmá“; $\operatorname{vol}(u^1, \dots, u^{n-1})$ je $(n-1)$-rozměrná míra „podstavy“ rovnoběžnostěnu $\mathcal{R}_a(w, u^1, \dots, u^{n-1})$ a $\|w\|$ je velikost jeho „výšky“ (srov. obr. 5.20).

Jistou názornou představu o vektorovém součinu v $\mathbb{R}^3$ a $\mathbb{R}^2$ lze získat z obr. 5.19.

5.34 Poznámka. Názorná geometrická interpretace vektorového součinu v $\mathbb{R}^2$ je zcela jednoduchá: vektor $\times(u)$ dostaneme, „otočíme-li vektor u o úhel $\pi/2$ ve směru

OBR. 5.19.

OBR. 5.20.

hodinových ručiček". Měla by tedy platit rovnost

$$\langle u, v\rangle = \langle \times(u), \times(v)\rangle. \tag{5.20}$$

Analytický důkaz je triviální, protože pro $u = (u_1, u_2)$, $v = (v_1, v_2)$ platí $\times(u) = (u_2, -u_1)$, $\times(v) = (v_2, -v_1)$.

Uvažujme nyní v $\mathbb{R}^n$ nadrovinu A. Nechť $\alpha_i \in [0, \pi/2]$ je úhel, který svírá normála k A s osou x_i a $c_i = \cos\alpha_i$. Úhel α_i je tedy také úhel (srov. [Bi; Def. 20.20]), který svírá nadrovina A a nadrovina $S_i{:}= \{x \in \mathbb{R}^n{:}\, x_i = 0\}$. Je-li $0 \neq w = (w_1, \ldots, w_n)$ normálový vektor k A, pak zřejmě $c_i = |w_i|/\|w\|$, takže

$$(c_1)^2 + \cdots + (c_n)^2 = 1. \tag{5.21}$$

Uvažujme dále „projekci" $\pi^{(i)}{:}\, \mathbb{R}^n \to \mathbb{R}^{n-1}$ danou předpisem

$$\pi^{(i)}(x_1, \ldots, x_n) := (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$$

a projekci $P^{(i)}{:}\, \mathbb{R}^n \to S_i$, kde

$$P^{(i)}(x_1, \ldots, x_n) := (x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n).$$

Zobrazení $\varphi_i{:}= \pi^{(i)} \restriction_{S_i}$, $\varphi{:}\, S_i \to \mathbb{R}^{n-1}$ je zřejmě unitární bijekce (pomocí níž se obvykle S_i ztotožňuje s $\mathbb{R}^{n-1}$). Podle Poznámky 5.11 platí $\kappa(\varphi_i) = 1$. Z rovnosti $\pi^{(i)} = \varphi_i \circ P^{(i)}$ a Tvrzení 5.13 tedy máme $\kappa(\pi^{(i)} \restriction_A) = \kappa(P^{(i)} \restriction_A)$. Pomocí vektorového součinu nyní dokážeme tvrzení, které je pro $n = 2$ a $n = 3$ „geometricky zřejmé".

5.35 Tvrzení. *Nechť A je nadrovina v $\mathbb{R}^n$ a nechť c_i, $\pi^{(i)}$, $P^{(i)}$ jsou definovány jako výše. Pak*

$$\kappa(\pi^{(i)} \restriction_A) = \kappa(P^{(i)} \restriction_A) = c_i. \tag{5.22}$$

Důkaz. Je-li $u^1, \dots, u^{n-1}$ báze zaměření A a $w := u^1 \times \cdots \times u^{n-1}$, podle (5.8), (5.3) a Tvrzení 5.32 (ii) platí

$$\kappa(\pi^{(i)} \restriction_A) = \frac{\operatorname{vol}(\pi^{(i)}(u^1), \dots, \pi^{(i)}(u^{n-1}))}{\operatorname{vol}(u^1, \dots, u^{n-1})} = \frac{|\det[\pi^{(i)}(u^1), \dots, \pi^{(i)}(u^{n-1})]|}{\|w\|}. \tag{5.23}$$

Protože $c_i = |w_i|/\|w\|$ a podle definice w máme

$$w_i = \det[e_i, u^1, \dots, u^{n-1}] = (-1)^{i+1} \det[\pi^{(i)}(u^1), \dots, \pi^{(i)}(u^{n-1})] \tag{5.24}$$

po dosazení do (5.23) dostáváme dokazovanou rovnost.

5.36 Poznámka. Z rovnosti (5.24) je vidět, jaký je geometrický význam znaménka $\operatorname{sgn}(w_i)$. Máme-li totiž na zaměření V nadroviny A zadánu orientaci tak, že $(u^1, \dots, u^{n-1})$ je kladná báze V (srov. Dodatek 6.2), pak z (5.24) vyplývá, že $(\pi^{(i)}(u^1), \dots, \pi^{(i)}(u^{n-1}))$ je kladná (resp. záporná) báze $\mathbb{R}^{n-1}$ (tj. $\pi^{(i)} \restriction_A : A \to \mathbb{R}^{n-1}$ zachovává (resp. mění) orientaci), právě když $(-1)^{i+1} w_i > 0$ (resp. $(-1)^{i+1} w_i < 0$). Poznamenejme ještě, že číslo $w_i/\|w\|$ je kosinus úhlu z $[0, \pi]$, který svírájí vektory w, e_i, a tedy orientované prostory A a $S_i = \{x \in \mathbb{R}^n \colon x_i = 0\}$, orientujeme-li A normálovým vektorem w. Číslo $(-1)^{i+1} w_i$ je kosinus úhlu, který svírá A s S_i, pokud v S_i za kladnou bázi zvolíme $e_1, \dots, e_{i-1}, e_{i+1}, \dots, e_n$.

Okamžitým důsledkem Tvrzení 5.35 je následující „zobecnění Pythagorovy věty“.

5.37 Věta. *Nechť $A \subset \mathbb{R}^n$ je afinní prostor dimenze $n-1$ a nechť $F \subset A$ je uzavřená množina, pro kterou $\mu_{n-1}(F) < \infty$. Pak*

$$(\mu_{n-1}(F))^2 = \sum_{i=1}^{n} \Big(\lambda_{n-1}(\pi^{(i)}(F))\Big)^2 = \sum_{i=1}^{n} \Big(\mu_{n-1}(P^{(i)}(F))\Big)^2. \tag{5.25}$$

Důkaz. Podle (5.22) platí $\lambda_{n-1}\left(\pi^{(i)}(F)\right) = c_i \mu_{n-1}(F)$, takže (5.25) plyne okamžitě z (5.21).

Je-li $k = 1$ a F je úsečka $\overline{ab}$, pak (5.25) je rovnost $\|b - a\|^2 = \sum_{i=1}^n (b_i - a_i)^2$, což je běžné zobecnění Pythagorovy věty. Projekce množiny F v případech $k = 1$, $n = 2$ a $k = 2$, $n = 3$ jsou znázorněny na obr. 5.21. Je-li F trojúhelník XYZ na obr. 5.21, vzorec (5.25) přejde v rovnost

$$(\mu_2(XYZ))^2 = (\mu_2(X0Y))^2 + (\mu_2(X0Z))^2 + (\mu_2(Y0Z))^2.$$

OBR. 5.21. Ke „zobecnění Pythagorovy věty“; množina F leží v rovině trojúhelníku XYZ.

5.38 Poznámka. Přirozená analogie předchozí věty platí pro afinní prostory A libovolné dimenze $1 \leq k < n$. Na pravé straně příslušného vzorce je pak součet čtverců k-rozměrných měr projekcí na všechny „k-rozměrné souřadnicové roviny“. Při důkazu je nutno užít Binet-Cauchyův vzorec o determinantu součinu obdélníkových matic (srov. [Kop], [LM]).

5.8 Lokální koeficient změny k-rozměrné míry a jeho výpočet

Nechť $G \subset \mathbb{R}^k$ je otevřená množina a $\varphi: G \to \mathbb{R}^n$ je parametrizace jednoduché k-plochy. Jestliže φ je afinní, víme, že existuje „koeficient změny k-rozměrné míry“ $\kappa = \kappa(\varphi)$, pro který platí $\mu_k(\varphi(B)) = \kappa \cdot \lambda_k(B)$ pro každou borelovskou $B \subset G$. Pro φ, která není afinní, již takový koeficient existovat nemusí. Víme však, že pro $t_0 \in G$ je $\kappa(\varphi'(t_0))$ „lokální koeficient změny k-rozměrné míry“: je-li $B \subset G$ borelovská množina, která je „hodně blízko“ bodu t_0, pak $\mu_k(\varphi(B)) \approx \kappa(\varphi'(t_0)) \cdot \lambda_k(B)$. To jsme použili při heuristickém odvozování vzorce pro obsah plochy. Nyní, když jsme tento vzorec použili v definici k-rozměrné míry, můžeme ovšem tuto „přibližnou rovnost“ po upřesnění dokázat:

5.39 Tvrzení. *Nechť G, φ jsou jako výše a $t_0 \in G$. Pak pro každé $\varepsilon > 0$ existuje otevřené okolí $U \subset G$ bodu t_0 takové, že pokud $B \subset U$ je borelovská množina a $\lambda_k(B) > 0$, pak*

$$\left| \frac{\mu_k(\varphi(B))}{\lambda_k(B)} - \kappa(\varphi'(t_0)) \right| < \varepsilon .$$

Důkaz. Protože funkce

$$\kappa(\varphi'(t)) = \sqrt{\Gamma\left(\frac{\partial\varphi(t)}{\partial t_1}, \dots, \frac{\partial\varphi(t)}{\partial t_k}\right)}$$

je zřejmě spojitá, lze zvolit okolí $U \subset G$ bodu t_0 tak, aby pro každé $t \in U$ platily nerovnosti $\kappa(\varphi'(t_0)) - \varepsilon < \kappa(\varphi'(t)) < \kappa(\varphi'(t_0)) + \varepsilon$. Je-li $B \subset U$ borelovská množina a $\lambda_k(B) > 0$, pak

$$\lambda_k(B) \cdot (\kappa(\varphi'(t_0)) - \varepsilon) \le \int_B \kappa(\varphi'(t))\ \mathrm{d}t \le \lambda_k(B) \cdot (\kappa(\varphi'(t_0)) + \varepsilon),$$

takže

$$\kappa(\varphi'(t_0)) - \varepsilon \le \frac{\mu_k(\varphi(B))}{\lambda_k(B)} \le \kappa(\varphi'(t_0)) + \varepsilon.$$

Umíme-li vypočítat „lokální koeficient“ $\kappa(\varphi'(t))$ parametrizace $\varphi\colon G \to \mathbb{R}^n$ jednoduché k-plochy P, výpočet plošného integrálu 1. druhu $\int_P f\ \mathrm{dS}$ se redukuje na výpočet Lebesgueova integrálu $\int_G f(\varphi(t)) \cdot \kappa(\varphi'(t))\ \mathrm{d}t$.

5.40 Poznámka.

(i) Někdy se píše $\mathrm{dS} = \kappa(\varphi'(t))\,\mathrm{d}t$ a tento výraz se nazývá „element plochy“. Místo o výpočtu „lokálního koeficientu“ se pak hovoří o výpočtu „elementu plochy“.

(ii) Pro zobrazení $\varphi\colon \mathbb{R}^k \to \mathbb{R}^k$ má význam „lokálního koeficientu změny k-rozměrné míry“ absolutní hodnota jacobiánu: $\kappa(\varphi'(t_0)) = |J_\varphi(t_0)|$. Proto někteří autoři číslu $\kappa(\varphi'(t_0))$ pro $\varphi\colon \mathbb{R}^k \to \mathbb{R}^n$ říkají „k-rozměrný jacobián“.

Protože $\varphi'(t)\,e_i = \dfrac{\partial\varphi(t)}{\partial t_i}$, platí podle Tvrzení 5.12

$$\kappa(\varphi'(t)) = \operatorname{vol}\left(\frac{\partial\varphi(t)}{\partial t_1}, \dots, \frac{\partial\varphi(t)}{\partial t_k}\right) = \sqrt{\Gamma\left(\frac{\partial\varphi(t)}{\partial t_1}, \dots, \frac{\partial\varphi(t)}{\partial t_k}\right)}. \tag{5.26}$$

Tento vzorec (který jsme již použili) je univerzální. V následujících speciálních případech ukážeme, jak jej lze převést na vzorec jiný (často jednodušší). Příslušné vzorce pro $\int_P f\ \mathrm{dS}$ však psát nebudeme.

Případ $k = 1$.

V tomto případě $\dfrac{\partial\varphi(t)}{\partial t_1} = \varphi'(t) \in \mathbb{R}^n$, takže

$$\kappa(\varphi'(t)) = \sqrt{\Gamma(\varphi'(t))} = \sqrt{\langle\varphi'(t), \varphi'(t)\rangle} = \|\varphi'(t)\|. \tag{5.27}$$

Případ $k = n - 1$.

V tomto případě lze $\kappa(\varphi'(t))$ vyjádřit pomocí vektorového součinu. Podle Tvrzení 5.32 (ii) totiž máme

$$\kappa(\varphi'(t)) = \operatorname{vol}\left(\frac{\partial\varphi(t)}{\partial t_1}, \dots, \frac{\partial\varphi(t)}{\partial t_{n-1}}\right) = \left\|\frac{\partial\varphi(t)}{\partial t_1} \times \dots \times \frac{\partial\varphi(t)}{\partial t_{n-1}}\right\|. \tag{5.28}$$

Pro zkrácení zápisů zavádíme označení

$$w_\varphi(t) := \frac{\partial\varphi(t)}{\partial t_1} \times \dots \times \frac{\partial\varphi(t)}{\partial t_{n-1}}, \tag{5.29}$$

při kterém tedy platí

$$\kappa(\varphi'(t)) = \|w_\varphi(t)\|. \tag{5.30}$$

Podle definice vektorového součinu můžeme souřadnice $w_\varphi(t)$ jednoduše vyjádřit pomocí jacobiánů:

$$(w_\varphi(t))_i = \det\left[e_i, \frac{\partial\varphi(t)}{\partial t_1}, \dots, \frac{\partial\varphi(t)}{\partial t_{n-1}}\right] =$$

$$= \begin{vmatrix} 0 & \frac{\partial\varphi_1(t)}{\partial t_1} & \dots & \frac{\partial\varphi_1(t)}{\partial t_{n-1}} \\ \vdots & \vdots & & \vdots \\ 1 & \frac{\partial\varphi_i(t)}{\partial t_1} & \dots & \frac{\partial\varphi_i(t)}{\partial t_{n-1}} \\ \vdots & \vdots & & \vdots \\ 0 & \frac{\partial\varphi_n(t)}{\partial t_1} & \dots & \frac{\partial\varphi_n(t)}{\partial t_{n-1}} \end{vmatrix} = (-1)^{i+1} \frac{D(\varphi_1, \dots, \varphi_{i-1}, \varphi_{i+1}, \dots, \varphi_n)}{D(t_1, \dots, t_{n-1})}(t).$$

Platí tedy

$$\kappa(\varphi'(t)) = \|w_\varphi(t)\| = \sqrt{\sum_{i=1}^n \left(\frac{D(\varphi_1, \dots, \varphi_{i-1}, \varphi_{i+1}, \dots, \varphi_n)}{D(t_1, \dots, t_{n-1})}(t)\right)^2}. \tag{5.31}$$

Je užitečné si rozmyslet, jakým způsobem odpovídá vzorec (5.31) Větě 5.37 a jaký vzorec pro výpočet $\kappa(\varphi'(t))$ opovídá v případě obecného $1 \le k < n$ dalšímu zobecnění Pythagorovy věty zmíněnému v Poznámce 5.38.

Případ $n = 3$, $k = 2$.
V klasické diferenciální geometrii se pro parametrizaci φ jednoduché 2-plochy v $\mathbb{R}^3$ často používá zápis

$$\varphi\colon (u, v) \mapsto (x(u, v), y(u, v), z(u, v)).$$

Pak na G podle (5.31) platí

$$\kappa(\varphi') = \sqrt{\left(\frac{D(y,z)}{D(u,v)}\right)^2 + \left(\frac{D(x,z)}{D(u,v)}\right)^2 + \left(\frac{D(x,y)}{D(u,v)}\right)^2}. \tag{5.32}$$

Použijeme-li (5.26), dostáváme

$$\kappa(\varphi') = \sqrt{\Gamma\left(\frac{\partial\varphi}{\partial u}, \frac{\partial\varphi}{\partial v}\right)} = \sqrt{\left\|\frac{\partial\varphi}{\partial u}\right\|^2 \left\|\frac{\partial\varphi}{\partial v}\right\|^2 - \left\langle\frac{\partial\varphi}{\partial u}, \frac{\partial\varphi}{\partial v}\right\rangle^2}. \tag{5.33}$$

Často se používá označení

$$E = g_{11} := \left\|\frac{\partial\varphi}{\partial u}\right\|^2 = \left(\frac{\partial x}{\partial u}\right)^2 + \left(\frac{\partial y}{\partial u}\right)^2 + \left(\frac{\partial z}{\partial u}\right)^2,$$

$$F = g_{12} = g_{21} := \frac{\partial x}{\partial u}\cdot\frac{\partial x}{\partial v} + \frac{\partial y}{\partial u}\cdot\frac{\partial y}{\partial v} + \frac{\partial z}{\partial u}\cdot\frac{\partial z}{\partial v},$$

$$G = g_{22} := \left\| \frac{\partial \varphi}{\partial v} \right\|^2 = \left(\frac{\partial x}{\partial v} \right)^2 + \left(\frac{\partial y}{\partial v} \right)^2 + \left(\frac{\partial z}{\partial v} \right)^2 .$$

Zde E, F, G jsou tzv. Gaussovy koeficienty (koeficienty první základní kvadratické formy plochy). Při tomto označení má (5.33) tvar

$$\kappa(\varphi') = \sqrt{EG - F^2} = \sqrt{g_{11} \cdot g_{22} - g_{12}^2},$$

formálně:

$$\mathrm{dS} = \sqrt{EG - F^2}\,\mathrm{d}u\,\mathrm{d}v = \sqrt{g_{11} \cdot g_{22} - g_{12}^2}\,\mathrm{d}u\,\mathrm{d}v.$$

Případ explicitně zadané $(n-1)$-plochy.

Nyní předpokládejme, že P je explicitně zadaný kus $(n-1)$-dimenzionální C^1 plochy v $\mathbb{R}^n$. Pak existuje otevřená množina $G \subset \mathbb{R}^{n-1}$, $1 \le i \le n$ a funkce f třídy C^1 na G, pro kterou

$$P = \{(x_1, \dots, x_n) \colon x_i = f(x^{(i)})\},$$

zzz

kde jsme použili označení $x^{(i)} := (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$. Ukážeme, jak počítat plošný integrál 1. druhu přes plochu P pomocí funkce f. Plochu P přirozeně parametrizujeme regulárním homeomorfismem $\varphi \colon G \to \mathbb{R}^n$,

$$\varphi(t) = \varphi(t_1, \dots, t_{n-1}) := (t_1, \dots, t_{i-1}, f(t), t_{i+1}, \dots, t_{n-1}).$$

Jacobiho matice φ má pak tvar

$$[\varphi'(t)] = \left[\frac{\partial \varphi(t)}{\partial t_1}, \dots, \frac{\partial \varphi(t)}{\partial t_{n-1}} \right] = \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 1 & 0 & \dots & 0 \\ \frac{\partial f(t)}{\partial t_1} & \frac{\partial f(t)}{\partial t_2} & \dots & \frac{\partial f(t)}{\partial t_{i-1}} & \frac{\partial f(t)}{\partial t_i} & \dots & \frac{\partial f(t)}{\partial t_{n-1}} \\ 0 & 0 & \dots & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 1 \end{pmatrix}$$

a platí

$$w_\varphi(t) = (-1)^{i+1} \left(-\frac{\partial f(t)}{\partial t_1}, \dots, -\frac{\partial f(t)}{\partial t_{i-1}}, 1, -\frac{\partial f(t)}{\partial t_i}, \dots, -\frac{\partial f(t)}{\partial t_{n-1}} \right). \tag{5.34}$$

Tento vzorec lze přímočaře odvodit z definice vektorového součinu. Přehlednější je však tato úvaha: $[\varphi'(t)]$ má hodnost $n-1$ a vektor $w_\varphi(t)$ je kolmý na všechny její sloupce. Krátký elementární výpočet okamžitě tedy dává, že $w_\varphi(t)$ je násobkem vektoru $\left(-\frac{\partial f(t)}{\partial t_1}, \dots, -\frac{\partial f(t)}{\partial t_{i-1}}, 1, -\frac{\partial f(t)}{\partial t_i}, \dots, -\frac{\partial f(t)}{\partial t_{n-1}} \right)$. Vyškrtneme-li z $[\varphi'(t)]$ i-tý řádek, dostáváme jednotkovou matici, takže platí $(w_\varphi(t))_i = (-1)^{i+1}$, z čehož pak okamžitě vyplývá (5.34). Dostáváme tedy

$$\kappa(\varphi'(t)) = \|w_\varphi(t)\| = \sqrt{1 + \sum_{j \ne i} \left(\frac{\partial f(t)}{\partial t_j} \right)^2}. \tag{5.35}$$

5.9 Orientace $(n \,\Gamma\, 1)$-rozměrných ploch pomocí normálového pole

Definovat pojem orientace obecné k-plochy není snadné. (srov. ???). Pojem orientace $(n-1)$-plochy však lze snadno, přesně a názorně definovat pomocí pojmu *spojitého jednotkového normálového pole.*

5.41 Poznámka. Pojem „vektorové pole v na množině $A \subset \mathbb{R}^n$“ je v klasické analýze jen jiný název pro zobrazení $v: A \to \mathbb{R}^n$. (Tato terminologie se používá zejména tehdy, kdy je užitečné si představovat „pole vázaných vektorů“ $\{\overrightarrow{x, x+v(x)}\colon\ x \in A\}$.)

Nechť P je $(n-1)$-plocha v $\mathbb{R}^n$. Řekneme, že ν je spojité jednotkové normálové pole na P, jestliže $\nu: P \to \mathbb{R}^n$ je spojité zobrazení takové, že pro každé $x \in P$ je $\nu(x)$ jednotkový normálový vektor k ploše P v bodě x. Jedno takové normálové pole ν k ploše P je znázorněno na obr. 5.22 vlevo.

Obr. 5.22.

Nechť P je $(n-1)$-plocha v $\mathbb{R}^n$. Řekneme, že P je *orientovatelná plocha*, jestliže na ní existuje spojité jednotkové normálové pole. Pokud jsme na ní jedno takové pole zvolili, řekneme, že P je *orientovaná plocha.* (Chceme-li o orientaci plochy hovořit jako o matematickém objektu, můžeme ji ztotožnit s tímto zvoleným spojitým jednotkovým normálovým polem.)

5.42 Tvrzení. *Nechť P je souvislá orientovatelná $(n-1)$-plocha v $\mathbb{R}^n$. Pak ji lze orientovat právě dvěma způsoby. Je-li ν jedno spojité jednotkové normálové pole na P, pak to druhé je $-\nu$.*

Důkaz. Nechť ν je jedno spojité jednotkové normálové pole na P. Pak zřejmě $\nu^* := -\nu$ je také takové pole. Uvažujme nyní libovolné spojité jednotkové normálové pole $\widetilde{\nu}$ na P, bod $a \in A$ a funkce $d(x) = \|\nu(x) - \widetilde{\nu}(x)\|$, $d^*(x) = \|\nu^*(x) - \widetilde{\nu}(x)\|$.

Funkce d, d^* jsou zřejmě spojité na souvislé množině P a nabývají pouze hodnoty 0 a 2; jsou tedy konstantní. Protože zřejmě buď $\widetilde{\nu}(a) = \nu(a)$ nebo $\widetilde{\nu}(a) = \nu^*(a)$, platí $\widetilde{\nu} = \nu$ nebo $\widetilde{\nu} = \nu^*$.

5.43 Poznámka. Je snadno vidět, že *nesouvislou* orientovatelnou $(n-1)$-plochu v $\mathbb{R}^n$ můžeme orientovat aspoň čtyřmi způsoby.

Ve starší literatuře se také říká, že souvislá orientovatelná plocha „má dvě strany" a orientace je výběrem jedné z těchto dvou stran. To odpovídá názorné představě, že jednotkový normálový vektor $\nu(x)$ „chodí" po jedné straně plochy P, na druhou stranu (po které „chodí" $\nu^*(x) = -\nu(x)$) se však nikdy nedostane.

Nejznámější neorientovatelnou (jednostrannou) souvislou plochou je Möbiův list zobrazený vpravo na obr. 5.22. Zde je znázorněno, jak se jednotkový normálový vektor n v bodě a „spojitě přemístil" do opačné polohy $n^* = -n$; z toho snadno vyplývá (pomocí modifikace úvahy z předchozího důkazu), že spojité jednotkové normálové pole na Möbiově listu neexistuje.

Snadno je vidět, že každá *jednoduchá* $(n-1)$-plocha je orientovatelná:

5.44 Tvrzení. *Nechť P je jednoduchá $(n-1)$-plocha v $\mathbb{R}^n$ a $\varphi: G \to \mathbb{R}^n$ je její parametrizace. Pro $t \in G$ a $x \in P$ položme*

$$w_\varphi(t) := \frac{\partial\varphi(t)}{\partial t_1} \times \cdots \times \frac{\partial\varphi(t)}{\partial t_{n-1}} \qquad \text{a} \qquad \nu_\varphi(x) := \frac{w_\varphi(\varphi^{-1}(x))}{\|w_\varphi(\varphi^{-1}(x))\|}.$$

Pak ν_φ je spojité jednotkové normálové pole na P.

Důkaz. Protože φ je třídy C^1, z Poznámky 5.29 (iii) ihned vyplývá, že zobrazení $w_\varphi: G \to \mathbb{R}^n$ je spojité. Protože $\left(\dfrac{\partial\varphi(t)}{\partial t_1}, \dots, \dfrac{\partial\varphi(t)}{\partial t_{n-1}}\right)$ je báze tečného prostoru $T_x(P)$, z Tvrzení 5.32 vyplývá, že pro každé $t \in G$ je $w_\varphi(t)$ nenulový normálový vektor k ploše P v bodě $\varphi(t)$. Závěr tvrzení nyní vyplývá z toho, že φ je homeomorfismus G na $\varphi(G)$.

Orientaci určenou polem ν_φ nazýváme *orientací určenou parametrizací* φ. Je-li P orientovaná a její orientace je táž, jako orientace určená parametrizací φ, řekneme, že parametrizace φ je *kladná*. Je-li orientace P táž, jako orientace určená polem $-\nu_\varphi$, říkáme, že parametrizace φ je *záporná*.

5.45 Poznámka. Nechť P je souvislá jednoduchá $(n-1)$-plocha v $\mathbb{R}^n$ orientovaná spojitým jednotkovým normálovým polem ν, $\varphi: G \to \mathbb{R}^n$ je její parametrizace a $t_0 \in G$. Pak z Tvrzení 5.42 a Tvrzení 5.44 okamžitě vyplývá, že φ je kladná (resp. záporná) parametrizace, právě když $w_\varphi(t_0) = \nu(\varphi(t_0))$ (resp. $w_\varphi(t_0) = -\nu(\varphi(t_0))$).

Budeme potřebovat také následující snadné tvrzení.

5.46 Tvrzení. *Nechť P je $(n-1)$-plocha v $\mathbb{R}^n$, ν je spojité jednotkové normálové pole na P a $Q \subset P$ je také $(n-1)$-plocha. Pak $\nu\restriction_Q$ je spojité jednotkové normálové* *pole na Q. (Orientace P tedy na Q kanonicky indukuje orientaci.)*

Důkaz. Zřejmě stačí dokázat, že pro každý bod $a \in Q$ platí $\nu(a) \perp T_a(Q)$. To však platí, protože podle definice tečného prostoru zřejmě $T_a(Q) \subset T_a(P)$.

(Protože oba tečné prostory mají stejné dimenze, nutně se rovnají. Také není těžké ukázat, že Q je vždy relativně otevřená v P; to však potřebovat nebudeme.)

5.10 Regulární hranice otevřené podmnožiny $\mathbb{R}^n$

Většina otevřených podmnožin $G \subset \mathbb{R}^n$, které jsou důležité pro aplikace, má hranici, která je „skoro hladká“: je sjednocením $(n-1)$-plochy a množiny N, která má nulovou $(n-1)$-rozměrnou míru. (Tuto vlastnost mají G_1 i G_2 z obr. 5.23).

OBR. 5.23.

Ve formulaci Gaussovy věty se však připouštějí jen některé takové množiny. Abychom je popsali, potřebujeme následující definici.

5.47 Definice. *Nechť $G \subset \mathbb{R}^n$ je otevřená množina a $x \in \partial G$. Řekneme, že x je regulární hraniční bod, jestliže platí:*

(i) *Existuje $\delta > 0$ takové, že $U_\delta(x) \cap \partial G$ je jednoduchá $(n-1)$-rozměrná plocha.*
(ii) $x \notin \operatorname{int}(\overline{G})$.

Množinu všech regulárních hraničních bodů budeme značit $\partial_* G$ a nazývat ji regulární hranicí množiny G.

5.48 Příklad. Nechť $G_1 \subset \mathbb{R}^2$, $G_2 \subset \mathbb{R}^2$ jsou množiny z obr. 5.23. Pak zřejmě platí $\partial G_1 = \partial_* G_1$. Pro bod $c \in \partial G_2$ neplatí (i), ale platí (ii); pro $b \in \partial G_2$ platí (i), ale neplatí (ii); pro $a \in \partial G_2$ neplatí ani (i), ani (ii).

Obr. 5.24.

Je snadno vidět, že *regulární hranice* $\partial_* G$ *je buď prázdná množina nebo* $(n-1)$-*plocha.* Situace v okolí regulárního hraničního bodu a vypadá velmi jednoduše. Pro případ $n = 2$ jsou možné právě 4 typy chování (které jsou znázorněny na obr. 5.24 vlevo), pro obecné n je $2n$ analogických typů. Toto tvrzení nebudeme potřebovat, a proto ani precizovat; bude nám stačit Tvrzení 5.51.

5.49 Definice. *Nechť* $G \subset \mathbb{R}^n$ *je otevřená množina a* $x \in \partial G$. *Řekneme, že vektor* $0 \neq v \in \mathbb{R}^n$

(i) *míří v bodě* x *ven z* G, *jestliže existuje* $\delta > 0$ *takové, že* $x + tv \notin \overline{G}$ *pro každé* $t \in (0, \delta)$;
(ii) *míří v bodě* x *do* G, *jestliže existuje* $\delta > 0$ *takové, že* $x + tv \in G$ *pro každé* $t \in (0, \delta)$.

5.50 Definice. *Nechť* $G \subset \mathbb{R}^n$ *je otevřená množina,* $x \in \partial_* G$ *a* $\nu \in \mathbb{R}^n$. *Řekneme, že* ν *je vnější (resp. vnitřní) jednotkový normálový vektor ke* G *v bodě* x, *jestliže*

(i) ν *je jednotkový normálový vektor k* $\partial_* G$ *v bodě* x *a*
(ii) ν *míří ven z* G *(resp.* ν *míří do* G*).*

Ilustrace k následujícímu tvrzení je na obr. 5.24 vpravo.

5.51 Tvrzení. *Nechť* $G \subset \mathbb{R}^n$ *je otevřená množina a* $x \in \partial_* G$. *Pak platí:*

(i) *Existuje právě jeden vnější jednotkový normálový vektor ke* G *v bodě* x *(který budeme označovat symbolem* $\nu_G^e(x)$*) a právě jeden vnitřní jednotkový normálový vektor ke* G *v bodě* x $(\nu_G^i(x) = -\nu_G^e(x))$.
(ii) *Jestliže* $v \in \mathbb{R}^n$ *a* $\langle v, \nu_G^e(x) \rangle > 0$, *pak* v *míří ven z* G. *Jestliže* $v \in \mathbb{R}^n$ *a* $\langle v, \nu_G^e(x) \rangle < 0$, *pak* v *míří do* G.
(iii) *Jednotkové normálové pole* ν_G^e *je spojité na ploše* $\partial_* G$.

Důkaz. (Stručný.) Nejdříve dokážeme, že existuje otevřené okolí U bodu x a na U „rozhraničující funkce“ h, tj. C^1 funkce na U taková, že $\operatorname{grad} h(x) \neq 0$ pro $x \in U$ a

$$G \cap U = \{t\colon\ h(t) < 0\}, \qquad \partial G \cap U = \{t\colon\ h(t) = 0\}. \tag{5.36}$$

Zvolme otevřené okolí V bodu x takové, že $\partial G \cap V$ je kus $(n-1)$-rozměrné C^1 plochy zadaný difeomorfismem, tj. existuje otevřená množina $H \subset \mathbb{R}^n$ a difeomorfismus $\psi\colon H \to \mathbb{R}^n$ takový, že

$$\partial G \cap V = \psi(H \cap \{(x_1, \ldots, x_n)\colon\ x_n = 0\}). \tag{5.37}$$

Nechť $y := \psi^{-1}(x)$ a $\delta > 0$ je tak malé, že $U^* := U_\delta^\infty(y) \subset H$. Položme

$$U := \psi(U^*), \quad M_1 := G \cap U, \quad M_2 := U \setminus \overline{G};$$

z Definice 5.47 (ii) plyne, že $M_2 \neq 0$. Dále označme

$$h^*(y_1, \ldots, y_n) := y_n, \quad I_1 := \{y \in U^* : h^*(y) < 0\}, \quad I_2 := \{y \in U^*\colon\ h^*(y) > 0\}.$$

Zobrazení $\psi \restriction_{I_1 \cup I_2}$ je homeomorfismus prostoru $I_1 \cup I_2$ na $M_1 \cup M_2$, intervaly I_1, I_2 jsou souvislé a M_1, M_2 jsou neprázdné obojetné podmnožiny prostoru $M_1 \cup M_2$. Z toho snadno vyplývá, že buď $M_1 = \psi(I_1)$ nebo $M_2 = \psi(I_2)$. V prvním případě stačí zřejmě položit $h := h^* \circ \psi^{-1} \restriction_U$ a v druhém případě $h := -h^* \circ \psi^{-1} \restriction_U$.

Protože $\partial G \cap U$ je kus $(n-1)$-rozměrné C^1 plochy zadaný implicitně vazbovou funkcí h, je $\operatorname{grad} h(x)$ normálový vektor k $\partial_* G$ v bodě x (srov. Poznámka 2.163 (a)). Pro $v \in \mathbb{R}^n$ platí

$$\lim_{t \to 0} \frac{h(x+tv) - h(x)}{t} = D_v h(x) = \langle v, \operatorname{grad} h(x) \rangle.$$

Z (5.36) a definice limity tedy vyplývá, že vektor v míří v bodě x ven z G (resp. do G), jestliže $\langle v, \operatorname{grad} h(x) \rangle > 0$ (resp. $\langle v, \operatorname{grad} h(x) \rangle < 0$). Vektory

$$\nu_G^e(x) := \frac{\operatorname{grad} h(x)}{\|\operatorname{grad} h(x)\|} \quad \text{a} \quad \nu_G^i(x) := -\frac{\operatorname{grad} h(x)}{\|\operatorname{grad} h(x)\|} \tag{5.38}$$

jsou tedy (jediné) vnější a vnitřní jednotkové normálové vektory ke G v bodě x a platí také (ii). Z (5.38) vyplývá, že ν_G^e je lokálně spojité na $\partial_* G$, takže platí též (iii).

Na $\partial_* G$ budeme většinou volit orientaci zadanou spojitým jednotkovým normálovým polem ν_G^e; v tom případě budeme říkat, že plocha $\partial_* G$ je *orientovaná vnější normálou.*

Obr. 5.25.

5.11 Tok vektorového pole orientovanou plochou a Gaussova věta

5.52 Definice. *Nechť $P \subset G$ je $(n-1)$-plocha v $\mathbb{R}^n$ orientovaná spojitým jednotkovým normálovým polem ν. Nechť $F\colon P \to \mathbb{R}^n$ je vektorové pole. Existuje-li reálné číslo*

$$\int_P \overrightarrow{F}\,\mathrm{d}\overrightarrow{S} := \int_P \langle F(x), \nu(x)\rangle\ \mathrm{d}S(x),$$

nazýváme je tokem vektorového pole F orientovanou plochou P.

Tok vektorového pole orientovanou plochou se často nazývá *plošným integrálem 2. druhu.* Je-li tento plošný integrál definován, říkáme, že konverguje.

Právě zavedený pojem toku má pro $n = 3$ přirozenou hydrodynamickou interpretaci (je však důležitý i v jiných fyzikálních oborech).

Předpokládejme, že otevřená množina $G \subset \mathbb{R}^3$ je zaplněna pohybující se nestlačitelnou kapalinou a $F(x)$ má význam vektorové rychlosti částice tekutiny v bodě x. (Předpokládáme tedy, že proudění je stacionární, tj. ustálené; rychlost částic v bodě x nezávisí na čase.) Tok pole F plochou P má pak význam objemu kapaliny, který za jednotkový čas proteče plochou P. Přitom objem kapaliny, který proteče plochou „na stranu, do které míří ν" se počítá se znaménkem $+$ a objem, který proteče plochou v opačném směru, se znaménkem $-$. Celkový tok může tedy být kladný, ale také záporný nebo nulový. Uvažujme nyní na ploše P velmi malou plošku ΔP s obsahem ΔS a zvolme $x \in \Delta P$ (viz obr. 5.25).

Částice kapaliny, které za velmi krátký čas Δt protekly ploškou ΔP, vyplní množinu, která je (za předpokladu hladkosti vektorového pole F) „k nerozeznání" od „válcového tělesa" s podstavou o obsahu ΔS a výškou $(\Delta t) \cdot \|F(x)\| \cdot |\cos\alpha|$, kde $\alpha \in [0, \pi]$ je úhel, který svírají vektory $\nu(x)$ a $F(x)$. Ploškou ΔP tedy za čas Δt proteče objem kapaliny ΔV, který má absolutní hodnotu $|\Delta V| \approx (\Delta S)\cdot(\Delta t)\cdot\|F(x)\|\cdot|\cos\alpha|$, přičemž se tento objem počítá se znaménkem $+$, pokud $\alpha\in[0,\pi/2]$ a se znaménkem $-$, pokud $\alpha\in(\pi/2,\pi]$, takže

$$\Delta V \approx (\Delta S)\cdot(\Delta t)\cdot\|F(x)\|\cdot\cos\alpha = (\Delta S)\cdot(\Delta t)\cdot\langle F(x),\nu(x)\rangle.$$

Za jednotkový čas tedy ploškou ΔP proteče objem kapaliny $\Delta T\approx(\Delta S)\cdot\langle F(x),\nu(x)\rangle$ a plošný integrál 1. druhu $\int_P\langle\nu(x),F(x)\rangle\,\mathrm{d}S(x)$ má význam celkového toku.

5.53 Poznámka. Celkový tok je (za vhodných předpokladů) limitou „integrálních součtů“ $\sum\langle F(x),\Delta S\cdot\nu(x)\rangle$. Z toho vychází označení toku, kde $\vec{F}\,\mathrm{d}\vec{S}$ je skalární součin a $\mathrm{d}\vec{S}$ se „heuristicky interpretuje“ jako nekonečně malý vektor $\Delta S\cdot\nu(x)$.

Výpočet toku přes jednoduchou orientovanou $(n-1)$-plochu v $\mathbb{R}^n$ zpravidla provádíme podle následujícího tvrzení (ve kterém $w_\varphi(t)$ má význam z Tvrzení 5.44, tj. $w_\varphi(t)=\frac{\partial\varphi}{\partial t_1}(t)\times\cdots\times\frac{\partial\varphi}{\partial t_{n-1}}(t)$).

5.54 Tvrzení. *Nechť P je jednoduchá orientovaná $(n-1)$-plocha v $\mathbb{R}^n$ a nechť $\varphi\colon G\to P$ je její kladná (nebo záporná) parametrizace. Nechť $F\colon P\to\mathbb{R}^n$ je borelovsky měřitelné vektorové pole. Pak*

$$\int_P \vec{F}\,\mathrm{d}\vec{S} = \pm\int_G\langle F(\varphi(t)),w_\varphi(t)\rangle\,\mathrm{d}t$$
$$= \pm\int_G \det\left[F(\varphi(t)),\frac{\partial\varphi}{\partial t_1}(t),\ldots,\frac{\partial\varphi}{\partial t_{n-1}}(t)\right]\,\mathrm{d}t,$$

konverguje-li jeden z integrálů. Přitom znaménko $+$ (resp. $-$) se bere, je-li parametrizace φ kladná (resp. záporná).

Důkaz. Podle Tvrzení 5.44 je plocha P orientovaná jednotkovým vektorovým polem $\nu_\varphi(x) := \pm\dfrac{w_\varphi(\varphi^{-1}(x))}{\|w_\varphi(\varphi^{-1}(x))\|}$. Pomocí Věty 5.24 a (5.30) dostáváme

$$\int_P\vec{F}\,\mathrm{d}\vec{S} := \int_P\langle F(x),\nu_\varphi(x)\rangle\,\mathrm{d}S(x) = \int_G\left\langle F(\varphi(t)),\pm\frac{w_\varphi(t)}{\|w_\varphi(t)\|}\right\rangle\cdot\|w_\varphi(t)\|\,\mathrm{d}t,$$

z čehož ihned plyne první z dokazovaných rovností. Druhá rovnost okamžitě vyplývá z Tvrzení 5.30.

5.55 Poznámka. V konkrétních příkladech počítáme tok vektorového pole F obecnou orientovanou $(n-1)$-plochou P (např. regulární hranicí ∂_*G) tak, že nalezneme vyjádření

$$P = P_1\cup\cdots\cup P_u\cup N,$$

kde $P_1,\ldots,P_u$ jsou jednoduché souvislé po dvou disjunktní plochy a $\mu_{n-1}(N)=0$. Na ploše P_i uvažujme orientaci indukovanou orientací plochy P (viz Tvrzení 5.46). Pak zřejmě

$$\int_P \vec{F}\,\mathrm{d}\vec{S} = \int_{P_1} \vec{F}\,\mathrm{d}\vec{S} + \cdots + \int_{P_u} \vec{F}\,\mathrm{d}\vec{S}.$$

Je-li nyní $\varphi_i\colon G_i \to \mathbb{R}^n$ parametrizace P_i, pak φ_i je buď kladná nebo záporná parametrizace (viz Poznámka 5.45). Toky F přes plochy $P_1, \ldots, P_u$ můžeme tedy počítat podle Tvrzení 5.54.

5.56 Příklad. Nechť $G = \{(x,y,z)\colon\ x^2+y^2 < (1-z)^2,\ 0 < z < 1\}$. Pak G je vnitřek kužele s vrcholem $(0,0,1)$, jehož podstava je kruh v rovině xy o poloměru 1 a středu v počátku. Položme $D := \{(x,y)\colon\ x^2+y^2 < 1\}$ a definujme $\varphi\colon D \to \mathbb{R}^3$ a $\omega\colon D \to \mathbb{R}^3$ předpisy

$$\varphi(x,y) = 1 - \sqrt{x^2+y^2}, \qquad \omega(x,y) = (x,y,0).$$

Je snadné ověřit, že φ, ω jsou parametrizace disjunktních jednoduchých ploch $P_1 := \varphi(D)$, $P_2 := \omega(D)$ a není těžké dokázat, že $\partial_* G = P_1 \cup P_2$.

Počítejme tok vektorového pole $F(x,y,z) = (x,y,x+z)$ plochou $\partial_* G$, která je orientovaná vnější normálou.

Protože pro $(x,y) \in D$ vektor $w_\omega(x,y) = (1,0,0)\times(0,1,0) = (0,0,1)$ v bodě $\omega(x,y) = (x,y,0)$ zřejmě míří do G, je ω záporná parametrizace P_2. Platí tedy

$$\int_{P_2} \vec{F}\,\mathrm{d}\vec{S} = -\int_D \langle F((\omega(x,y))), w_\omega(x,y)\rangle\,\mathrm{d}x\,\mathrm{d}y =$$
$$= -\int_D \langle (x,y,x),(0,0,1)\rangle\,\mathrm{d}x\,\mathrm{d}y = -\int_D x\mathrm{d}x\,\mathrm{d}y = 0.$$

Abychom se vyhnuli odmocninám, položme $\eta(r,\alpha) = (r\cos\alpha, r\sin\alpha)$ pro $(r,\alpha) \in \Omega := (0,1) \times (0,2\pi)$. Pak $\eta(\Omega) = D \setminus ([0,1]\times\{0\})$, takže $\psi := \varphi\circ\omega$ je parametrizace plochy $P_3 = P_1 \setminus A$, kde $A = \{(x,y,z)\colon\ y=0,\ 0 \le x \le 1,\ z = 1-x\}$. Zřejmě $\mu_2(A) = 0$, takže $\int_{P_1} \vec{F}\,\mathrm{d}\vec{S} \int_{P_3} \vec{F}\,\mathrm{d}\vec{S}$. Platí

$$\psi(r,\alpha) = (r\,\cos\alpha, r\,\sin\alpha, 1-r), \quad w_\psi(r,\alpha) = (\cos\alpha,\sin\alpha,-1)\times(-r\sin\alpha, r\cos\alpha, 0) = (r\,\cos\alpha, r\,\sin\alpha, r).$$

Protože v každém bodě P_3 zřejmě vektor $v = (0,0,1)$ míří ven z G a $\langle w_\psi(r,\alpha), v\rangle = r > 0$ pro $(r,\alpha)\in\Omega$, dostáváme, že ψ je kladná parametrizace P_3. Platí tedy

$$\int_{P_3} \vec{F}\,\mathrm{d}\vec{S} = \int_\Omega \langle F((\psi(r,\alpha))), w_\psi(r,\alpha)\rangle\,\mathrm{d}r\,\mathrm{d}\alpha =$$
$$= \int_\Omega \langle (r\,\cos\alpha, r\,\sin\alpha, 1-r+r\,\cos\alpha),(r\,\cos\alpha, r\,\sin\alpha, r)\rangle\,\mathrm{d}r\,\mathrm{d}\alpha = \int_\Omega (r + r^2\,\cos\alpha)\,\mathrm{d}r\,\mathrm{d}\alpha = \pi.$$

Je tedy $\int_{\partial_* G} \vec{F}\,\mathrm{d}\vec{S} = \pi$.

Abychom mohli zformulovat základní (Gaussovu) větu o toku vektorového pole plochou, potřebujeme definovat pojem divergence vektorového pole.

5.57 Definice. *Nechť $F\colon G \to \mathbb{R}^n$, $F = (F_1, \ldots F_n)$ je vektorové pole třídy C^1 na otevřené množině $G \subset \mathbb{R}^n$ a $x \in G$. Pak číslo*

$$\operatorname{div} F(x) := \frac{\partial F_1}{\partial x_1}(x) + \cdots + \frac{\partial F_n}{\partial x_n}(x)$$

nazýváme divergencí vektorového pole F v bodě x.

Nyní již můžeme zformulovat Gaussovu větu, které se také někdy říká věta Gauss–Ostrogradského nebo věta o divergenci.

5.58 Věta. (Gaussova věta) *Nechť $G \subset \mathbb{R}^n$ je omezená otevřená množina, pro kterou*

$$\partial G \in \mathcal{P}_{n-1}, \quad \mu_{n-1}(\partial G) < \infty \quad \text{a} \quad \mu_{n-1}(\partial G \setminus \partial_* G) = 0. \tag{5.39}$$

Nechť F je vektorové pole, které je třídy C^1 na nějaké otevřené množině obsahující $\overline{G}$. Orientujeme-li plochu $\partial_* G$ vnější normálou, pak

$$\int_{\partial_* G} \overrightarrow{F}\,\mathrm{d}\overrightarrow{S} = \int_{\partial G} \langle F(x), \nu(x)\rangle\,\mathrm{d}S(x) = \int_G \operatorname{div} F(x)\,\mathrm{d}x. \tag{5.40}$$

Poznamenejme, že první rovnost v (5.40) je zřejmá z definice toku a předpokladu $\mu_{n-1}(\partial G \setminus \partial_* G) = 0$.

Gaussova věta se někdy vyslovuje v následující „skalární formě“.

5.59 Věta. *Nechť G je jako v předchozí větě a $\nu = (\nu_1, \dots, \nu_n)$ je pole jednotkových vnějších normálových vektorů ke G na $\partial_* G$. Nechť f je reálná funkce, která je třídy C^1 na nějaké otevřené množině obsahující $\overline{G}$. Pak pro každé $i \in \{1, \dots, n\}$ platí*

$$\int_{\partial G} f \cdot \nu_i\,\mathrm{d}S = \int_G \frac{\partial f(x)}{\partial x_i}\,\mathrm{d}x. \tag{5.41}$$

Je snadno vidět, že Věta 5.59 je v podstatě pouze přeformulací Věty 5.58. Víme-li totiž, že platí Věta 5.58, a aplikujeme ji na vektorové pole $F := f \cdot e_i$, dostaneme rovnost (5.41). Naopak, je-li F pole z Věty 5.58 pak rovnost (5.40) dostaneme, aplikujeme-li pro každé $i \in \{1, \dots, n\}$ Větu 5.59 na $f := F_i$ a příslušné rovnosti (5.41) sečteme.

Následující dvě tvrzení ukazují, jak počítat $\int_P f \cdot \nu_i\,\mathrm{dS}$ pro jednoduché a explicitně zadané plochy.

5.60 Tvrzení. *Nechť P je jednoduchá $(n-1)$-plocha v $\mathbb{R}^n$ orientovaná jednotkovým normálovým polem ν a $\varphi\colon G \to P$ je její kladná (nebo záporná) parametrizace. Pak pro každou borelovsky měřitelnou funkci $f\colon P \to \mathbb{R}$ a $i \in \{1, \dots, n\}$ platí*

$$\int_P f \cdot \nu_i\,\mathrm{dS} = \pm(-1)^{i+1} \int_G f(\varphi(t)) \cdot \frac{D(\varphi_1, \dots, \varphi_{i-1}, \varphi_{i+1}, \dots, \varphi_n)}{D(t_1, \dots, t_{n-1})}(t)\,\mathrm{d}t, \tag{5.42}$$

konverguje-li jeden z integrálů. Přitom znaménko $+$ (resp. $-$) se bere, je-li parametrizace φ kladná (resp. záporná).

Důkaz. Protože podle Tvrzení 5.44 platí $\nu_i(\varphi(t)) = \pm \dfrac{w_\varphi(t)}{\|w_\varphi(t)\|}$, dostáváme

$$\int_P f \cdot \nu_i \; \mathrm{dS} = \pm \int_G f(\varphi(t)) \cdot \frac{(w_\varphi(t))_i}{\|w_\varphi(t)\|} \cdot \|w_\varphi(t)\| \; \mathrm{d}t.$$

Nyní stačí použít vyjádření $(w_\varphi(t))_i$ ze vzorce za (5.30).

5.61 Tvrzení. *Nechť P je explicitně zadaný kus $(n-1)$-rozměrné C^1 plochy tvaru*

$$P = \{(x_1, \dots, x_n) : x_i = h(x^{(i)}),\ x^{(i)} \in G\},$$

kde jsme použili označení $x^{(i)} := (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$. Pak platí:

(i) *Je-li ν spojité jednotkové normálové pole na P, pak buď $\nu_i(x) > 0,\ x \in P$ nebo $\nu_i(x) < 0,\ x \in P$.*

(ii) *Pro každou borelovsky měřitelnou reálnou funkci f na P platí*

$$\int_P f \cdot \nu_i \; \mathrm{dS} = \pm \int_G f(x_1, \dots, x_{i-1}, h(x^{(i)}), x_{i+1}, \dots, x_n) \; \mathrm{d}x^{(i)}, \tag{5.43}$$

(konverguje-li jeden z integrálů), přičemž znaménko + (resp. −) se bere, pokud $\nu_i(x) > 0,\ x \in P$ (resp. $\nu_i(x) < 0,\ x \in P$) a symbol $\mathrm{d}x^{(i)}$ nahrazuje $\mathrm{d}x_1 \dots \mathrm{d}x_{i-1}\, \mathrm{d}x_{i+1} \dots \mathrm{d}x_n$.

Důkaz. Uvažujme parametrizaci φ plochy P danou předpisem

$$\varphi(t) = \varphi(t_1, \dots, t_{n-1}) := (t_1, \dots, t_{i-1}, h(t), t_i, \dots, t_{n-1}), \quad t \in G.$$

Pak podle (5.34) $(w_\varphi(t))_i = (-1)^{i+1}$. Položíme-li $\nu_i^*(x) := (-1)^{i+1} \dfrac{w_\varphi(\varphi^{-1}(x))}{\|w_\varphi(\varphi^{-1}(x))\|}$, je tedy ν^* spojité jednotkové normálové pole na P a $\nu_i^*(x) = 1/\|w_\varphi(\varphi^{-1}(x))\| > 0$ pro $x \in G$. (Pro druhé spojité jednotkové normálové pole $-\nu^*$ ovšem platí $(-\nu_i^*)(x) < 0,\ x \in G$.) Vzorec (5.43) jistě stačí ověřit pro $\nu = \nu^*$. Platí

$$\int_P f \cdot \nu_i^* \; \mathrm{dS} = \int_G f(\varphi(t)) \cdot \frac{1}{\|w_\varphi(t)\|} \cdot \|w_\varphi(t)\| \; \mathrm{d}t = \int_G f(\varphi(t)) \; \mathrm{d}t.$$

Označíme-li proměnné $t_1, \dots, t_{n-1}$ pořadě symboly $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n$, dostáváme vzorec (5.43).

Gaussovu větu lze vyslovit také v řeči diferenciálních forem. Teorie diferenciálních forem umožňuje dokázat velmi elegantní zobecnění Gaussovy věty – obecnou Stokesovu větu. Nyní budeme uvažovat jen nejjednodušší diferenciální formy (řádu $n-1$) v $\mathbb{R}^n$ tvaru

$$f(x) \; \mathrm{d}x_1 \dots \mathrm{d}x_{i-1} \; \mathrm{d}x_{i+1} \dots \mathrm{d}x_n \;=\; f(x) \; \mathrm{d}x_1 \wedge \dots \wedge \mathrm{d}x_{i-1} \wedge \mathrm{d}x_{i+1} \wedge \dots \wedge \mathrm{d}x_n,$$

kde f je spojitá funkce na otevřené množině $\Omega \subset \mathbb{R}^n$.

Při použití staršího zápisu vlevo se diferenciální forma chápe pouze jako „formální symbol“. V novějším pojetí je diferenciální forma přesně definovaný objekt; je to tenzorová funkce (a $\wedge$ je symbol pro tzv. vnější součin tenzorů).

Integrál diferenciální formy $\omega = f(x)\ \mathrm{d}x_1 \dots \mathrm{d}x_{i-1}\ \mathrm{d}x_{i+1} \dots \mathrm{d}x_n$ přes orientovanou $(n-1)$-plochu lze definovat pomocí integrálu 1. druhu takto.

5.62 Definice. *Nechť $1 \le i \le n$, f je spojitá funkce na otevřené množině $\Omega \subset \mathbb{R}^n$ a $P \subset \Omega$ je $(n-1)$-plocha orientovaná jednotkovým normálovým polem ν. Pak klademe*

$$\int_P f(x)\ \mathrm{d}x_1 \dots \mathrm{d}x_{i-1}\ \mathrm{d}x_{i+1} \dots \mathrm{d}x_n := (-1)^{i+1} \int_P f \cdot \nu_i\ \mathrm{d}S, \tag{5.44}$$

má-li pravá strana smysl.

Integrálu z diferenciální formy přes plochu se také často říká *plošný integrál 2. druhu.*

Motivace Definice 5.62 je (aspoň částečně) dána vzorcem (5.43). Koeficient $(-1)^{i+1}$ úzce souvisí se vzorcem (5.42) (z kterého ihned plyne následující tvrzení). Viz též Poznámka (?).

5.63 Tvrzení. *Nechť $1 \le i \le n$, f je spojitá funkce na otevřené množině $\Omega \subset \mathbb{R}^n$, $P \subset \Omega$ je jednoduchá orientovaná $(n-1)$-plocha a $\varphi\colon G \to P$ je její kladná parametrizace. Nechť $\omega = f(x)\ \mathrm{d}x_1 \dots \mathrm{d}x_{i-1}\ \mathrm{d}x_{i+1} \dots \mathrm{d}x_n$. Pak*

$$\int_P \omega = \int_G f(\varphi(t)) \cdot \frac{D(\varphi_1, \dots, \varphi_{i-1}, \varphi_{i+1}, \dots, \varphi_n)}{D(t_1, \dots, t_{n-1})}(t)\ \mathrm{d}t, \tag{5.45}$$

konverguje-li jeden z integrálů.

V „řeči diferenciálních forem“ má Věta 5.59 následující tvar.

5.64 Věta. *Nechť otevřená množina $G \subset \mathbb{R}^n$ je jako ve Větě 5.58. Plochu $\partial_* G$ orientujme vnější normálou. Nechť f je reálná funkce, která je třídy C^1 na nějaké otevřené množině obsahující $\overline{G}$. Pak pro každé $i \in \{1, \dots, n\}$ platí*

$$\int_{\partial_* G} f(x)\ \mathrm{d}x_1 \dots \mathrm{d}x_{i-1}\ \mathrm{d}x_{i+1} \dots \mathrm{d}x_n = (-1)^{i+1} \int_G \frac{\partial f(x)}{\partial x_i}\ \mathrm{d}x_1 \dots \mathrm{d}x_n. \tag{5.46}$$

5.65 Poznámka. geometrický smysl integrálu z diferenciální formy Popíšeme názorný geometrický význam integrálu $\int_P f(x)\ \mathrm{d}x_1 \dots \mathrm{d}x_{i-1}\ \mathrm{d}x_{i+1} \dots \mathrm{d}x_n$. (Za jistých předpokladů kladených na P a f lze heuristickým úvahám níže dát přesný smysl.)

Uvažujme „projekci“ $\pi^{(i)}\colon \mathbb{R}^n \to \mathbb{R}^{n-1}$ danou předpisem

$$\pi^{(i)}(x_1, \dots, x_n) := (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$$

.

Rozložme P na konečně (nebo spočetně) mnoho borelovských množin $P_1, P_2, \dots$ s velmi malým diametrem a v každé z nich zvolme bod $\xi_k \in P_k$. Pak $\int_P f(x)\,\mathrm{d}x_1 \dots \mathrm{d}x_{i-1}\,\mathrm{d}x_{i+1} \dots \mathrm{d}x_n$ je limitou „integrálních součtů“

$$\sum_k f(\xi_k) \cdot (\pm\lambda_{n-1}(\pi(P_k))),$$

kde $\pm\lambda_{n-1}(\pi(P_k))$ je „orientovaný $(n-1)$-rozměrný objem „orientovaného průmětu $\pi^{(i)}(P_k)$“. To lze říci trochu srozumitelněji tak, že znaménko + (resp. -) se bere, pokud zobrazení $\pi^{(i)} \restriction_{P_k} : P_k \to \mathbb{R}^{n-1}$ „zachovává orientaci“ (resp. „mění orientaci“).

Zkusíme vysvětlit, jak tato názorná představa vede ke vzorci (5.44) (pro heuristické zdůvodnění vzorce (5.45) srov. ???).

Protože P_k „nerozeznáme“ od množiny ležící v afinním prostoru $A := A_P^{af}(\xi_k)$ a podle Tvrzení 5.35 $\kappa(\pi^{(i)} \restriction_A) = c_i = |\nu_i(\xi_k)|$, usuzujeme, že

$$\lambda_{n-1}(\pi^{(i)}(P_k)) \approx \mu_{n-1}(P_k)|\nu_i(\xi_k)|.$$

Orientujme A pomocí normálového vektoru $\nu(\xi_k)$ (srov. Dotatek 6.2.). Nahraďme nyní frázi „$\pi^{(i)} \restriction_{P_k}$ zachovává orientaci“ tvrzením $\pi^{(i)} \restriction_A$ zachovává orientaci, které podle Poznámky 5.36 platí, právě když $(-1)^{i+1}\nu_i(\xi_k) > 0$. Pak usuzujeme, že

$$\sum_k f(\xi_k) \cdot (\pm\lambda_{n-1}(\pi(P_k))) \approx \sum_k f(\xi_k) \cdot (-1)^{i+1}\nu_i(\xi_k)\,\mu_{n-1}(P_k) \approx (-1)^{i+1} \int_P f \cdot \nu_i \,\mathrm{d}S.$$

5.66 Poznámka. V teorii diferenciálních forem se definuje (vnější) diferenciál diferenciální formy

$$\omega \;=\; f(x)\,\mathrm{d}x_1 \wedge \dots \wedge \mathrm{d}x_{i-1} \wedge \mathrm{d}x_{i+1} \wedge \dots \wedge \mathrm{d}x_n$$

jako diferenciální forma

$$\mathrm{d}\omega := \mathrm{d}f(x) \wedge \mathrm{d}x_1 \wedge \dots \wedge \mathrm{d}x_{i-1} \wedge \mathrm{d}x_{i+1} \wedge \dots \wedge \mathrm{d}x_n.$$

Neznáme-li teorii diferenciálních forem, nemá pro nás ovšem výraz napravo smysl. Zde poznamenejme pouze to, že

$$\mathrm{d}f(x) = \frac{\partial f(x)}{\partial x_1}\,\mathrm{d}x_1 + \dots + \frac{\partial f(x)}{\partial x_n}\,\mathrm{d}x_n,$$

pro vnější součin $\wedge$ platí distributivní zákon a platí

$$\mathrm{d}x_i \wedge \dots \wedge \mathrm{d}x_j \wedge \dots \wedge \mathrm{d}x_k \wedge \dots \wedge \mathrm{d}x_l = -\mathrm{d}x_i \wedge \dots \wedge \mathrm{d}x_k \wedge \dots \wedge \mathrm{d}x_j \wedge \dots \wedge \mathrm{d}x_l$$

(takže $\mathrm{d}x_i \wedge \dots \wedge \mathrm{d}x_j \wedge \dots \wedge \mathrm{d}x_j \wedge \dots \wedge \mathrm{d}x_l = 0$). Počítáme-li zcela formálně podle těchto pravidel, snadno dostáváme, že

$$\mathrm{d}\omega = \frac{\partial f(x)}{\partial x_i}\,\mathrm{d}x_i \wedge \mathrm{d}x_1 \wedge \dots \wedge \mathrm{d}x_{i-1} \wedge \mathrm{d}x_{i+1} \wedge \dots \wedge \mathrm{d}x_n = (-1)^{i+1}\frac{\partial f(x)}{\partial x_i}\,\mathrm{d}x_1 \wedge \dots \wedge \mathrm{d}x_n.$$

Rovnost (5.46) lze tedy psát ve tvaru

$$\int_{\partial_* G} \omega = \int_G \mathrm{d}\omega.$$

Rovnost z obecné Stokesovy věty v $\mathbb{R}^n$ má tentýž tvar, ale G je v ní orientovaná plocha dimenze k v $\mathbb{R}^n$ $(n \ge k)$ a $\partial_* G$ je její „okraj“, což je jistá orientovaná plocha dimenze $k-1$ v $\mathbb{R}^n$. Přitom ω je „obecná hladká diferenciální forma řádu $(k-1)$ v $\mathbb{R}^n$“. Pro další informace o diferenciálních formách viz oddíl 5.15.

Uveďme ještě vyjádření toku vektorového pole pomocí integrálů z diferenciálních forem.

Nechť $F = (F_1, \dots, F_n)$ je spojité vektorové pole na otevřené množině $\Omega \subset \mathbb{R}^n$ a $P \subset \Omega$ je orientovaná $(n-1)$-plocha. Z Definice 5.52 a Definice 5.62 okamžitě dostáváme rovnost

(5.47)
$$\int_P \vec{F}\,\mathrm{d}\vec{S} = \int_P F_1(x)\,\mathrm{d}x_2\dots\mathrm{d}x_n - \int_P F_2(x)\,\mathrm{d}x_1\mathrm{d}x_3\dots\mathrm{d}x_n + \dots$$
$$+ (-1)^{n+1}\int_P F_n(x)\,\mathrm{d}x_1\dots\mathrm{d}x_{n-1}.$$

Gaussovu větu nyní dokážeme pouze pro velmi speciální množiny G. Přitom však uvidíme základní pricip důkazu Gaussovy věty spočívající v použití Newton-Leibnizovy formule a Fubiniovy věty. Úplný důkaz lze nalézt v Dodatku 6.5. Pro zjednodušení vyjadřování zavedeme (jen pro potřebu následujícího výkladu) technický pojem i-speciální množiny.

5.67 Definice. *Nechť $1 \le i \le n$. Řekneme, že $G \subset \mathbb{R}^n$ je i-speciální množina, jestliže platí:*

(i) *G je omezená otevřená množina.*
(ii) $\partial G \in \mathcal{P}_{n-1}$, $\quad \mu_{n-1}(\partial G) < \infty \quad$ a $\quad \mu_{n-1}(\partial G \setminus \partial_* G) = 0.$
(iii) *Existuje otevřená množina $\Omega \subset \mathbb{R}^{n-1}$ a dvě funkce $d(t) < h(t)$ třídy C^1 na Ω takové, že*
$$G = \{(x_1,\dots,x_n) : x^{(i)} \in \Omega,\ d(x^{(i)}) < x_i < h(x^{(i)})\},$$
kde jsme položili $x^{(i)} := (x_1,\dots,x_{i-1},x_{i+1},\dots,x_n)$.

5.68 Tvrzení. *Nechť $1 \le i \le n$ a $G \subset \mathbb{R}^n$ je i-speciální množina. Nechť ν je pole vnějších jednotkových normálových vektorů na $\partial_* G$ a nechť f je reálná funkce, která je třídy C^1 na nějaké otevřené množině obsahující $\overline{G}$. Pak*

$$\int_{\partial G} f\cdot\nu_i\,\mathrm{d}S = \int_G \frac{\partial f(x)}{\partial x_i}\,\mathrm{d}x. \tag{5.48}$$

Důkaz. Pro zjednodušení zápisu budeme předpokládat, že $i = n$, důkaz v obecném případě je však zcela analogický. Kvůli zkrácení zápisu dále budeme místo $(x_1,\dots,x_{n-1})$ psát $x^{(n)}$, takže např. $(x_1,\dots,x_n) = (x^{(n)}, x_n)$.

Je snadné ověřit, že $\partial G = H \cup D \cup B$, kde $B = \{x \in \partial G : x^{(n)} \in \partial\Omega\}$,
$H = \{x \in \mathbb{R}^n : x^{(n)} \in \Omega,\ x_n = h(x^{(n)})\}$, $\quad D = \{x \in \mathbb{R}^n : x^{(n)} \in \Omega,\ x_n = d(x^{(n)})\}$,
a platí inkluze $H \subset \partial_* G$, $D \subset \partial_* G$ (srov. obr. 5.26).

Je zřejmé, že v $x \in H$ míří vektor e_n ven z G a v $x \in D$ míří ven z G vektor $-e_n$. Z toho podle Tvrzení 5.51 vyplývá, že $\nu_i(x) = \langle \nu, e_n\rangle > 0$ pro $x \in H$ a $\nu_i(x) < 0$ pro $x \in D$. V bodě $x \in B \cap \partial_* G$ zřejmě nemíří do G žádný z vektorů e_n, $-e_n$, z čehož podle Tvrzení 5.51 vyplývá, že $\nu_i(x) = 0$. Platí tedy $\int_B f\cdot\nu_i\,\mathrm{dS} = 0$ a podle Tvrzení 5.61 dostáváme

Obr. 5.26.

$$\int_H f \cdot \nu_i \,\mathrm{dS} = \int_\Omega f(x^{(n)}, h(x^{(n)}))\,\mathrm{d}x_1 \ldots \mathrm{d}x_{n-1},$$
$$\int_D f \cdot \nu_i \,\mathrm{dS} = -\int_\Omega f(x^{(n)}, d(x^{(n)}))\,\mathrm{d}x_1 \ldots \mathrm{d}x_{n-1}.$$

(Lebesgueovy integrály vpravo totiž zřejmě konvergují.) Platí tedy

$$\int_{\partial G} f \cdot \nu_i \,\mathrm{dS} = \int_\Omega \Big(f(x^{(n)}, h(x^{(n)})) - f(x^{(n)}, d(x^{(n)})) \Big)\,\mathrm{d}x_1 \ldots \mathrm{d}x_{n-1}. \tag{5.49}$$

Protože funkce $\dfrac{\partial f}{\partial x_n}$ je spojitá na kompaktní množině $\overline{G}$, je na ní integrovatelná, takže podle Fubiniovy věty máme

$$\int_G \frac{\partial f(x)}{\partial x_n}\,\mathrm{d}x = \int_\Omega \left(\int_{d(x^{(n)})}^{h(x^{(n)})} \frac{\partial f}{\partial x_n}(x^{(n)}, x_n)\,\mathrm{d}x_n \right) \mathrm{d}x_1 \ldots \mathrm{d}x_{n-1}. \tag{5.50}$$

Protože podle Newton-Leibnizovy formule platí

$$\int_{d(x^{(n)})}^{h(x^{(n)})} \frac{\partial f}{\partial x_n}(x^{(n)}, x_n)\,\mathrm{d}x_n = f(x^{(n)}, h(x^{(n)})) - f(x^{(n)}, d(x^{(n)})),$$

ze vzorců (5.49) a (5.50) dostáváme dokazovanou rovnost (5.48).

5.69 Poznámka. Z předchozího tvrzení ihned vyplývá platnost Gaussovy věty pro množiny, které jsou i-speciální pro každé $i = 1, \ldots, n$. Tuto velmi speciální vlastnost mají například otevřené intervaly a otevřené koule. Většinu „obvyklých otevřených množin“ G pak lze pro každé $i = 1, \ldots, n$ „rozložit“ na konečně mnoho i-speciálních množin takovým způsobem, že pokud na každou z nich použijeme

Obr. 5.27.

rovnost (5.48) a obdržené rovnosti sečteme, dostaneme vzorec (5.48) pro G (viz obr. 5.27, kde vlevo je znázorněn „rozklad“ mezikruží G na 2-speciální množiny a napravo na 1-speciální množiny). Pro takové množiny G platí tedy i Gaussova věta. Tuto úvahu lze v obvyklých příkladech přesně provést, ověřování předpokladů obecné Gaussovy věty je však nesrovnatelně pohodlnější.

Nechť $G \subset \mathbb{R}^n$ je omezená otevřená množina, pro kterou platí (5.39). Pak lze Gaussovu větu použít k *výpočtu $\lambda_n(G)$ pomocí plošného integrálu přes ∂G.*

Položíme-li například $f(x) := x_i \quad (1 \leq i \leq n)$, pak $\frac{\partial f}{\partial x_i} = 1$, takže rovnost (5.41) má tvar

$$\lambda_n(G) = \int_{\partial G} x_i \, \nu_i(x) \, \mathrm{d}S(x). \tag{5.51}$$

Pro vektorové pole $v(x) := \frac{x}{n}$ zase platí $\operatorname{div} v(x) = 1$, takže vzorec (5.40) má tvar

$$\lambda_n(G) = \frac{1}{n} \int_{\partial_* G} \overrightarrow{x} \, \mathrm{d}\overrightarrow{S(x)} = \frac{1}{n} \int_{\partial G} \langle x, \nu(x) \rangle \, \mathrm{d}S(x). \tag{5.52}$$

5.70 Příklad. Nechť $G := B(0,1)$ je jednotková koule v $\mathbb{R}^n$. Je snadné ověřit, že

$$\partial G = \partial_* G = \{x : \|x\| = 1\}$$

a $\nu(x) := x$ je jednotková vnější normála ke G v bodě $x \in \partial G$. Vzorec (5.52) má tedy tvar

$$\lambda_n\left(B(0,1)\right) = \frac{1}{n} \int_{\partial G} \langle x, x \rangle \, \mathrm{dS}(x) = \frac{1}{n} \, \mu_{n-1}\left(\partial \, B(0,1)\right).$$

Dostali jsme jednoduchý vztah mezi objemem jednotkové koule a plošným obsahem jednotkové sféry v $\mathbb{R}^n$.

5.71 Příklad. (obsah smyčky Descartova listu) Spočtěme $\lambda_2(G)$, kde $G = \{(x,y) \colon x^3 < 2xy, x > 0, y > 0\}$. Zřejmě G je otevřená a $\overline{G} \subset M := \{(x,y) \colon x^3 \leq 2xy, x \geq 0, y \geq 0\}$.

Obr. 5.28.

Protože M má se sjednocením os společný pouze počátek, platí $\partial G = \overline{G} \setminus G \subset P \cup \{(0,0)\}$, kde $P = \{(x,y)\colon\ x^3 = 2xy, x > 0, y > 0\}$. Položíme-li $g(x,y) := x^3 + y^3 - 2xy$, snadno se spočte, že $\operatorname{grad} g(x,y) = (3x^2 - 2y, 3y^2 - 2x) \neq (0,0)$ pro $(x,y) \in P$, z čehož snadno plyne $P \subset \partial_* G$. Jestliže pro $(x,y) \in P$ položíme $t = \psi(x,y) = y/x$, snadno dostáváme, že ψ je bijekce P na $(0,\infty)$ a pro $\varphi := \psi^{-1}$ platí $\varphi(t) = (2t\,(1+t^3)^{-1}, 2t^2\,(1+t^3)^{-1}),\ t \in (0,\infty)$. Protože $\lim_{t\to 0+} \varphi(t) = (0,0)$ a ∂G je uzavřená, vidíme, že $\partial G = P \cup \{(0,0)\}$. (Není těžké dokázat, že $(0,0) \notin \partial_* G$, ale to nepotřebujeme.) Snadný výpočet dává, že $\varphi'(t) \neq 0,\ t > 0$. Protože $\psi = \varphi^{-1}$ je spojitá na P, je φ parametrizace jednoduché jednorozměrné plochy P. Platí $\varphi(1) = (1,1)$, $\varphi'(1) = (-1,1)$, $\quad w_\varphi(1) = \times\,\varphi'(1) = (1,1)$, a protože $\operatorname{grad} g(1,1) = (1,1)$, vidíme, že $w_\varphi(1)$ směřuje v bodě $(1,1)$ ven z G, takže parametrizace φ plochy P (orientované vnější normálou) je kladná (srov. Poznámka 5.45). Platí tedy

$$\lambda_2(G) = -\int_P y\,\mathrm{d}x = -\int_0^\infty \frac{2t^2}{1+t^3} \cdot \left(\frac{2t}{1+t^3}\right)' \,\mathrm{d}t = \frac{2}{3}.$$

5.12 Křivkové integrály 1. a 2. druhu

Pojem křivky a cesty

Pojem křivky se v matematice používá v několika různých významech. Nejdůležitější tři z nich se pokusíme objasnit na názorném příkladu. Představme si, že jsme na tabuli ($\mathbb{R}^2$) nakreslili křivou čáru (která může „sama sebe protínat“). Křivkou se někdy rozumí:

(a) Výsledek naší činnosti, tedy množina (křivá čára) $K \subset \mathbb{R}^2$, kterou jsme nakreslili.

(b) Přesný popis naší činnosti, tedy přesný popis pohybu (*cesty*) hmotného bodu (konce křídy) v časovém intervalu $[a,b]$. Tento popis je dán zobrazením $\varphi\colon [a,b] \to \mathbb{R}^2$, kde $\varphi(t)$ je poloha hmotného bodu v čase t.

(c) Křivá čára $K \subset \mathbb{R}^2$ a informace, „jakým způsobem“ jsme tuto čáru namalovali, přičemž nás nezajímá, kdy a jakou rychlostí jsme čáru malovali, ale pouze „směr, kterým bod (konec křídy) probíhal množinu K“. Na obr. 5.28 jsou znázorněny pomocí šipek dva různé způsoby probíhání téže množiny K. Formalizací této představy dostáváme třetí význam pojmu křivky; někdy se hovoří o „orientované křivce“; srov. Poznámka 5.77.

5.72 Definice.

(i) *Cestou v* $\mathbb{R}^n$ *rozumíme spojité zobrazení* $\varphi\colon[a,b]\to\mathbb{R}^n$, *kde* $-\infty<a<b<\infty$. *Obor hodnot* $\varphi([a,b])$ *nazýváme geometrickým obrazem cesty a označujeme jej symbolem* $\langle\varphi\rangle$.

(ii) *Řekneme, že cesta* $\varphi\colon[a,b]\to\mathbb{R}^n$ *je třídy* C^1, *má-li spojitou derivaci na* $[a,b]$. *Podrobněji řečeno: pro každý bod* $t_0\in[a,b]$ *existuje limita* $\varphi'(t_0):=\lim\limits_{t\to t_0,t\in[a,b]}\dfrac{\varphi(t)-\varphi(t_0)}{t-t_0}$ *a vektorová funkce* $t\mapsto\varphi'(t)$ *je spojitá na intervalu* $[a,b]$.

(iii) *Řekneme, že cesta* $\varphi\colon[a,b]\to\mathbb{R}^n$ *je skoro regulární, jestliže existuje dělení* $\{t_0=a<t_1<\cdots<t_k=b\}$ *intervalu* $[a,b]$ *takové, že* φ *je třídy* C^1 *na každém intervalu* $[t_{i-1},t_i]$, $i=1,\ldots,k$ *(tj.* $\varphi\restriction_{[t_{i-1},t_i]}$ *je cesta třídy* C^1*) a pro* $t\in[a,b]\setminus\{t_0,t_1,\ldots,t_k\}$ *platí* $\varphi'(t)\neq 0$.

(iv) *Cesta* φ *se nazývá jednoduchá, je-li* φ *prostá;* φ *je uzavřená cesta, jestliže* $\varphi(a)=\varphi(b)$. *Je-li* φ *uzavřená a* $\varphi\restriction_{[a,b)}$ *je prosté, říkáme, že* φ *je jednoduchá uzavřená cesta.*

5.73 Poznámka. Popisuje-li cesta $\varphi\colon[a,b]\to\mathbb{R}^3$ pohyb hmotného bodu v $\mathbb{R}^3$ a $t\in[a,b]$, pak $\varphi'(t)$ má význam vektorové okamžité rychlosti bodu v čase t a $\|\varphi'(t)\|$ má význam velikosti rychlosti (skalární rychlosti).

5.74 Definice. *Řekneme, že* $K\subset\mathbb{R}^n$ *je křivka (skoro regulární křivka, jednoduchá křivka, uzavřená křivka, jednoduchá uzavřená křivka), jestliže existuje cesta (skoro regulární cesta, ..., jednoduchá uzavřená cesta)* $\varphi\colon[a,b]\to\mathbb{R}^n$ *jejímž geometrickým obrazem* $\langle\varphi\rangle$ *je* K. *Jestliže* $\langle\varphi\rangle=K$, *říkáme, že* φ *je parametrizací křivky* K. *Jednoduché křivce se říká oblouk.*

5.75 Poznámka.

(i) Podle Věty 1.106 a Věty 1.118 je každá křivka v $\mathbb{R}^n$ kompaktní souvislá množina. Jak ukázal Peano, čtverec $[0,1]^2$ je křivka; existuje cesta („Peanova křivka“) $\varphi\colon[0,1]\to\mathbb{R}^2$, jejímž geometrickým obrazem je tento čtverec. Není těžké dokázat, že každá jednoduchá a také každá skoro regulární křivka v $\mathbb{R}^n$, $n>1$, je řídká.

(ii) Terminologie kolísá; někdy se cestě říká „křivka“ a křivce „geometrický obraz křivky“.

(iii) Z Věty 1.106 vyplývá, že množina $K\subset\mathbb{R}^n$ je oblouk, právě když je homeomorfní s intervalem $[0,1]$. Není obtížné dokázat, že množina $K\subset\mathbb{R}^n$ je jednoduchá uzavřená křivka právě tehdy, když je homeomorfní s jednotkovou kružnicí v $\mathbb{R}^2$.

(iv) *Pojem skoro regulární křivky není běžný.* Skoro regulární cesta φ může mít v některých bodech t nulovou derivaci; v bodech $\varphi(t)$ pak nemusí mít křivka $\langle\varphi\rangle$ „polotečnu“ (srov. Příklad 5.76). Důvod pro práci se skoro regulárními křivkami je ten, že obvyklé parametrizace některých klasických křivek (srov. Příklad 5.94) mají v některých bodech nulovou derivaci. Kdybychom pracovali s obvyklejšími „po částech regulárními“ křivkami, museli bychom zbytečně měnit parametrizaci.

5.76 Příklad. Nechť $K\subset\mathbb{R}^2$ je graf funkce $f(x)=x\sin(\ln x)$, $x\in(0,1]$, $f(0)=0$. Pak K „zřejmě“ nemá v bodě $(0,0)$ „polotečnu“ (tento pojem jsme nedefinovali). Přesto je K skoro regulární křivka; je snadné ověřit, že její parametrizace $\varphi(t) := (t^2, f(t^2))$, $t \in [0,1]$ je skoro regulární cesta.

5.77 Poznámka. (O „orientované křivce“) Řekneme, že cesty $\varphi\colon [a,b] \to \mathbb{R}^n$ a $\psi\colon [c,d] \to \mathbb{R}^n$ se liší nepodstatně (a píšeme $\varphi \sim \psi$), existuje-li spojitá rostoucí funkce ω zobrazující interval $[c,d]$ na $[a,b]$ taková, že $\psi = \varphi \circ \omega$. Je snadné ukázat, že $\sim$ je vztah ekvivalence; třídy této ekvivalence se někdy nazývají „orientovanými křivkami“ (nepodstatně se lišící cesty tedy určují jednu orientovanou křivku). Tím je upřesněn výklad (c) o třetím pojetí pojmu křivky.

Křivkový integrál 1. druhu

Nechť $\varphi\colon [a,b] \to \mathbb{R}^n$ je cesta a f je reálná funkce na $\langle\varphi\rangle$. Myšlenka klasické definice křivkového integrálu 1. druhu $\int_\varphi f \,\mathrm{d}s$ je tato:

Uvažujme dělení $D = \{t_0 = a < t_1 < \cdots < t_k = b\}$ intervalu $[a,b]$, body $\xi_i \in [t_{i-1}, t_i]$ a integrální součet

$$\sum_{i=1}^{k} f(\varphi(\xi_i)) \cdot (\Delta s)_i, \tag{5.53}$$

kde $(\Delta s)_i$ je délka cesty $\varphi \restriction_{[t_{i-1},t_i]}$. Limita, ke které se blíží integrální součty (5.53), když norma dělení $\nu(D)$ jde k nule, je pak $\int_\varphi f \,\mathrm{d}s$.

Aby tato definice byla korektní, je třeba nejprve definovat pojem délky cesty a předpokládat, že φ má konečnou délku. Je však možno uvážit, že pro dostatečně jemná dělení je asi $(\Delta s)_i \approx \|\varphi(t_i) - \varphi(t_{i-1})\|$ a $\int_\varphi f \,\mathrm{d}s$ definovat jako limitu integrálních součtů

$$\sum_{i=1}^{k} f(\varphi(\xi_i)) \cdot \|\varphi(t_i) - \varphi(t_{i-1})\|.$$

Je-li navíc φ cesta třídy C^1, pro velmi jemná dělení se zdá (a lze dokázat), že $\|\varphi(t_i) - \varphi(t_{i-1})\| \approx \|\varphi'(\xi_i)\| \cdot (t_i - t_{i-1})$, takže se domníváme, že

$$\sum_{i=1}^{k} f(\varphi(\xi_i)) \cdot \|\varphi(t_i) - \varphi(t_{i-1})\| \approx \sum_{i=1}^{k} f(\varphi(\xi_i)) \cdot \|\varphi'(\xi_i)\| \cdot (t_i - t_{i-1}).$$

Jsme tedy vedeni k následující definici.

5.78 Definice. *Nechť $\varphi\colon [a,b] \to \mathbb{R}^n$ je skoro regulární cesta a f je reálná funkce definovaná na nějaké podmnožině $\mathbb{R}^n$. Pak definujeme křivkový integrál 1. druhu $\int_\varphi f \,\mathrm{d}s$ rovností*

$$\int_\varphi f \,\mathrm{d}s := \int_a^b f(\varphi(t)) \cdot \|\varphi'(t)\| \,\mathrm{d}t,$$

konverguje-li Lebesgueův integrál napravo.

5.79 Poznámka. Nechť $\varphi\colon [a,b] \to \mathbb{R}^n$ je skoro regulární cesta. Poznamenejme toto:

(i) Pro existenci $\int_\varphi f \,\mathrm{d}s$ je zřejmě nutné, aby hodnota $f(\varphi(t))$ byla definovaná pro λ_1-skoro všechna $t \in [a,b]$.

(ii) Je-li f spojitá reálná funkce na $\langle\varphi\rangle$, pak $\int_\varphi f \,\mathrm{d}s$ existuje. Jsou-li totiž $\{t_0, \ldots, t_k\}$ body z Definice 5.72 (iii), pak konverguje každý z integrálů $\int_{t_{i-1}}^{t_i} f(\varphi(t)) \cdot \|\varphi'(t)\| \,\mathrm{d}t$, protože vektorovou funkci $\varphi'(t)$ a tedy i funkci $f(\varphi(t)) \cdot \|\varphi'(t)\|$ lze spojitě rozšířit na $[t_{i-1}, t_i]$, $i = 1, \ldots, k$.

(iii) Křivkový integrál $\int_\varphi 1 \text{ ds} = \int_a^b \|\varphi'(t)\| \text{ d}t$ nazýváme délkou cesty φ. Lze ukázat, že tato definice splývá s obvyklou (mnohem obecnější) definicí délky cesty pomocí suprema délek „vepsaných lomených čar“.

Následující snadné tvrzení ukazuje vztah mezi křivkovým integrálem 1. druhu a plošným 1-rozměrným integrálem 1. druhu pro případ prosté (a „téměř prosté“) skoro regulární cesty.

5.80 Tvrzení. *Nechť $\varphi: [a, b] \to \mathbb{R}^n$ je skoro regulární cesta, pro kterou existuje konečná množina $F \subset [a, b]$ taková, že φ je prostá na $[a, b] \setminus F$. Pak $\langle\varphi\rangle \in \mathcal{P}_1$ a pro každou reálnou funkci f (definovanou na podmnožině $\mathbb{R}^n$) platí*

$$\int_\varphi f \text{ ds} = \int_{\langle\varphi\rangle} f \text{ d}S_1,$$

má-li jedna strana rovnosti smysl.

Důkaz. (Stručný.) Z předpokladů snadno vyplývá existence bodů $t_0 = a < t_1 < \cdots < t_k = b$ takových, že $\varphi \restriction_{[t_{i-1}, t_i]}$ je prostá cesta třídy C^1, $\varphi'(t) \neq 0$ pro $t \in (t_{i-1}, t_i)$ a množiny $P_i := \varphi((t_{i-1}, t_i))$, $i = 1, \ldots, k$ jsou po dvou disjunktní. Podle Tvrzení 2.144 je každá P_i jednoduchá 1-plocha a podle (5.27) platí

$$\int_{P_i} f \text{ d}S_1 = \int_{t_{i-1}}^{t_i} f(\varphi(t)) \cdot \|\varphi'(t)\| \text{ d}t,$$

konverguje-li jeden z integrálů. Protože konečná množina je nulová pro míru μ_1, je $\langle\varphi\rangle \in \mathcal{P}_1$ a

$$\int_{\langle\varphi\rangle} f \text{ d}S_1 = \sum_{i=1}^k \int_{P_i} f \text{ d}S_1 = \sum_{i=1}^k \int_{t_{i-1}}^{t_i} f(\varphi(t)) \cdot \|\varphi'(t)\| \text{ d}t =$$

$$= \int_a^b f(\varphi(t)) \cdot \|\varphi'(t)\| \text{ d}t = \int_\varphi f \text{ ds},$$

je-li jeden z pěti výrazů konečný.

5.81 Důsledek. *Nechť φ_1, φ_2 jsou skoro regulární cesty, které jsou buď jednoduché nebo uzavřené jednoduché a $\langle\varphi_1\rangle = \langle\varphi_2\rangle$. Nechť f je funkce n proměnných. Pak $\int_{\varphi_1} f \text{ ds} = \int_{\varphi_2} f \text{ ds}$, má-li jedna strana rovnosti smysl.*

Křivkový integrál 2. druhu

5.82 Definice. *Nechť $\varphi: [a, b] \to \mathbb{R}^n$ je skoro regulární cesta a $F: \langle\varphi\rangle \to \mathbb{R}^n$ je vektorové pole. Pak integrál (druhého druhu) vektorového pole F podél cesty φ definujeme rovností*

$$\int_\varphi F \text{ d}\varphi := \int_a^b \langle F(\varphi(t)), \varphi'(t)\rangle \text{ d}t,$$

Obr. 5.29.

pokud Lebesgueův integrál napravo konverguje.

Křivkový integrál druhého druhu má důležité fyzikální interpretace, z nichž nejdůležitější je (pro $n = 3$) interpretace $\int_\varphi F\ d\varphi$ jako práce vykonaná silovým polem F po cestě φ.

Nechť φ a F jsou jako v Definici 5.82 a navíc předpokládejme, že vektorové pole F je spojité. Nechť $\varphi(t) \in \mathbb{R}^3$ má význam polohy hmotného bodu B v čase $t \in [a, b]$ a F má význam silového pole. Na hmotný bod B tedy v čase t silové pole F působí silou $F(\varphi(t))$. Ptáme se, jak velkou práci W vykoná při tomto pohybu bodu B silové pole F. K heuristickému odvození uvažujme velmi krátký časový interval $[t_1, t_2] \subset [a, b]$ délky $\Delta t = t_2 - t_1$, na kterém je φ třídy C^1 (viz obr. 5.29).

Spojitá vektorová funkce $F(\varphi(t))$ se na $[t_1, t_2]$ téměř nemění, takže ji nahradíme konstantní silou $F(\varphi(t_1))$. Dále pohyb bodu v časovém intervalu $[t_1, t_2]$ „nerozeznáme“ od rovnoměrného přímočarého pohybu, za který můžeme vzít například pohyb (popsaný cestou $\psi(t) = \varphi(t_1) + (t - t_1)\varphi'(t_1),\ t \in [t_1, t_2]$), při kterém je bod v čase t_1 v bodě $\varphi(t_1)$ a v časovém intervalu $[t_1, t_2]$ se pohybuje rychlostí $\varphi'(t_1)$. Práce, kterou vykoná síla $F(\varphi(t_1))$ při tomto rovnoměrném přímočarém pohybu, je podle středoškolské fyziky rovna

$$\|F(\varphi(t_1))\| \cdot \|\varphi'(t_1) \cdot \Delta t\| \cdot \cos\alpha = \Delta t \cdot \langle F(\varphi(t_1)), \varphi'(t_1)\rangle,$$

kde $\alpha \in [0, \pi]$ je úhel sevřený vektory $F(\varphi(t_1))$ a $\varphi'(t_1)$. Domníváme se proto, že pro práci ΔW vykonanou silovým polem F v časovém intervalu $[t_1, t_2]$ platí „přibližná rovnost“ $\Delta W \approx \Delta t \cdot \langle F(\varphi(t_1)), \varphi'(t_1)\rangle$. Je-li tedy $a = t_0 < t_1 < \cdots < t_k = b$ velmi jemné dělení intervalu $[a, b]$ takové, že φ je na každém $[t_{i-1}, t_i]$ třídy C^1, domníváme se, že

$$W \approx \sum_{i=1}^{k} (t_i - t_{i-1}) \cdot \langle F(\varphi(t_{i-1})), \varphi'(t_{i-1})\rangle,$$

takže je přirozené předpokládat, že $W = \int_\varphi F\ d\varphi$.

5.83 Poznámka. Aproximujeme-li pohyb v intervalu $[t_1, t_2]$ rovnoměrným přímočarým pohybem, při kterém je bod v čase t_1 v bodě $\varphi(t_1)$ a v čase t_2 v bodě $\varphi(t_2)$, a silové pole konstantní silou $F(\varphi(\xi))$, kde $\xi \in [t_1, t_2]$ (srov. obr. 5.29), dostáváme

$$\Delta W \approx \langle F(\varphi(\xi)), \varphi(t_2) - \varphi(t_1)\rangle$$

a (pro dělení jako výše a body $\xi_i \in [t_{i-1}, t_i]$)

$$W \approx \sum_{i=1}^{k} \langle F(\varphi(\xi_i), \varphi(t_i) - \varphi(t_{i-1})\rangle. \tag{5.54}$$

Práce W by tedy měla být limitou součtů z (5.54). Takto, tj. jako Stieltjesův integrál se $\int_\varphi F\, d\varphi$ definuje (srov. [Kr]) pro případ obecnějších cest. Této definici také zřejmým způsobem odpovídá značení integrálu. Pro případ spojitého F a skoro regulární φ obě definice splývají, takže jsme zvolili tu elementárnější.

5.84 Poznámka. Je-li vektorové pole F spojité na $\langle\varphi\rangle$, snadno vidíme (srov. Poznámka 5.79(ii)), že integrál $\int_\varphi F\, \mathrm{d}\varphi$ existuje.

Nyní budeme vyšetřovat vztah mezi křivkovým integrálem 2. druhu a plošným 1-rozměrným integrálem 2. druhu pro případ jednoduché a jednoduché uzavřené cesty v $\mathbb{R}^n$; tedy vztah mezi „prací“ a „tokem“. Napřed zavedeme některá označení.

5.85 Definice. *Nechť $\varphi\colon [a,b] \to \mathbb{R}^n$ je skoro regulární cesta, která je jednoduchá nebo jednoduchá uzavřená. Pak položíme*

$$D_{\varphi_*} := \{x \in (a,b)\colon\ \varphi'(x) \neq 0\}, \qquad \varphi_* := \varphi\restriction_{D_{\varphi_*}}, \qquad H_{\varphi_*} := \varphi(D_{\varphi_*}).$$

Jsou tedy D_{φ_*} a H_{φ_*} po řadě definiční obor a obor hodnot zobrazení φ_*. Je snadno vidět, že $[a,b]\setminus D_{\varphi_*}$ je konečná množina, $\varphi_* \in C^1(D_{\varphi_*})$ a $\varphi_*'(x) \neq 0$, $x \in D_{\varphi_*}$. Protože φ je jednoduchá nebo jednoduchá uzavřená, platí

$$H_{\varphi_*} \cap \varphi(\partial D_{\varphi_*}) = H_{\varphi_*} \cap \varphi([a,b]\setminus D_{\varphi_*}) = \emptyset,$$

takže podle Tvrzení 2.144 je φ_* homeomorfismus. *Je tedy φ_* regulární homeomorfismus a H_{φ_*} je jednoduchá 1-plocha.*

5.86 Poznámka. Jednoduchá 1-plocha H_{φ_*} není jednoznačně určena křivkou $\langle\varphi\rangle$. Je-li například $\varphi^1(t) = (t^3, 0)$, $\varphi^2(t) = (t, 0)$, $t \in [-1,1]$, pak $\langle\varphi^1\rangle = \langle\varphi^2\rangle$, ale $H_{\varphi^1_*} = H_{\varphi^2_*} \setminus \{(0,0)\}$.

5.87 Tvrzení. *Nechť $\varphi\colon [a,b] \to \mathbb{R}^2$ je skoro regulární cesta, která je jednoduchá nebo jednoduchá uzavřená. Na jednoduché 1-ploše H_{φ_*} uvažujme orientaci určenou parametrizací φ_*. Nechť F je spojité vektorové pole na $\langle\varphi\rangle$. Položíme-li*

$$\Phi := \times(F) = (F_2, -F_1),$$

pak platí

$$\int_\varphi F\, \mathrm{d}\varphi = \int_{H_{\varphi_*}} \overrightarrow{\Phi}\, \mathrm{d}\overrightarrow{S}. \tag{5.55}$$

Důkaz. Podle Definice 5.82 platí

$$\int_\varphi F\, \mathrm{d}\varphi := \int_a^b \langle F(\varphi(t)), \varphi'(t)\rangle\, \mathrm{d}t$$

a podle Tvrzení 5.54

$$\int_{H_{\varphi_*}} \overrightarrow{\Phi}\, \mathrm{d}\overrightarrow{S} = \int_{D_{\varphi_*}} \langle \Phi(\varphi_*(t)), w_{\varphi_*}(t)\rangle\, \mathrm{d}t.$$

Protože pro $t \in D_{\varphi_*}$ podle Poznámky 5.34 platí

$$\langle \Phi(\varphi_*(t)), w_{\varphi_*}(t) \rangle = \langle \times\left(F(\varphi_*(t))\right), \times\left(\varphi_*'(t)\right)\rangle = \langle F(\varphi_*(t)), \varphi_*'(t)\rangle$$

a $[a,b] \setminus D_{\varphi_*}$ je konečná množina, dostáváme (5.55).

Pod pojmem „křivkový integrál 2. druhu“ se často rozumí integrál z diferenciální formy přes cestu.

Budeme uvažovat diferenciální formy na otevřené množině $\Omega \subset \mathbb{R}^n$, které jsou tvaru

$$f_1 \,\mathrm{d}x_1 + \dots + f_n \,\mathrm{d}x_n, \tag{5.56}$$

kde $f_1, \dots, f_n$ jsou spojité reálné funkce na Ω.

Je-li φ skoro regulární cesta v Ω (tj. $\langle\varphi\rangle \subset \Omega$), pak klademe

$$\int_\varphi f_1 \,\mathrm{d}x_1 + \dots + f_n \,\mathrm{d}x_n \ := \int_\varphi f \,\mathrm{d}\varphi, \tag{5.57}$$

kde $f := (f_1, \dots, f_n)$.

Diferenciální formu (5.56) zde opět chápeme pouze jako „formální výraz“. Pokud $f_j = 0$ na Ω, pak člen $f_j \,\mathrm{d}x_j$ ve výrazu (5.56) nepíšeme. Podle této úmluvy výrazem $f_i \,\mathrm{d}x_i$ $(1 \le i \le n)$ tedy rozumíme diferenciální formu (5.56), ve které $f_j = 0$ na Ω pro $j \neq i$.

Je-li tedy g spojitá funkce na Ω, platí podle této úmluvy a Definice 5.82

$$\int_\varphi g \,\mathrm{d}x_i = \int_a^b g(t) \cdot \varphi_i'(t) \,\mathrm{d}t. \tag{5.58}$$

5.88 Poznámka. Nechť φ a φ_* jsou jako v Tvrzení 5.87 a f je spojitá funkce na otevřené množině $\Omega \subset \mathbb{R}^2$, která obsahuje $\langle\varphi\rangle$. Pak z (5.55) a (5.47) vyplývá, že

$$\int_\varphi f \,\mathrm{d}x_1 = \int_{H_{\varphi_*}} f \,\mathrm{d}x_1, \qquad \int_\varphi f \,\mathrm{d}x_2 = \int_{H_{\varphi_*}} f \,\mathrm{d}x_2,$$

takže integrál z diferenciální formy přes cestu a příslušnou 1-rozměrnou orientovanou plochu se rovnají.

5.13 Greenova věta

Klasická formulace Greenovy věty je založena na dosti hlubokých poznatcích o topologii roviny. Proto začneme nepřesným heuristickým výkladem:

Pro „nepříliš složité“ jednoduché uzavřené křivky $K \subset \mathbb{R}^2$, které „umíme nakreslit jedním tahem“ se zdá být zřejmé, že:

(i) Křivka K rozděluje rovinu na dvě souvislé otevřené množiny, „vnitřní“ a „vnější“.
(ii) Jestliže φ je jednoduchá uzavřená cesta a $\langle\varphi\rangle = K$, pak jsou jen dvě možnosti: cesta φ je kladně orientovaná (tj. probíhá křivku K v kladném smyslu; proti směru hodinových ručiček) nebo je záporně orientovaná (tj. probíhá K v záporném smyslu; po směru hodinových ručiček).

Tvrzení (i) platí i pro zcela obecnou jednoduchou uzavřenou křivku K; to říká slavná Jordanova věta:

5.89 Věta. (Jordanova) *Nechť $K \subset \mathbb{R}^2$ je jednoduchá uzavřená křivka. Pak otevřená množina $\mathbb{R}^2 \setminus K$ má dvě komponenty, z nichž jedna (které říkáme vnitřek křivky K a značíme ji $\operatorname{Int} K$) je omezená a druhá (které říkáme vnějšek křivky K a značíme ji $\operatorname{Ext} K$) neomezená. Přitom platí $\partial(\operatorname{Int} K) = \partial(\operatorname{Ext} K) = K$.*

Poznamenejme, že platnost Jordanovy věty není vůbec samozřejmá; připouštíme totiž i takové křivky, které nemají tečnu v žádném bodě, takže se jistě vymykají naší geometrické intuici. Přesný důkaz Jordanovy věty (viz např. [Čer]) je obtížný. (Ani v případě, že K je skoro regulární, není důkaz snadný.)

Také tvrzení (ii) lze přesně formulovat a dokázat pro zcela obecné jednoduché uzavřené křivky K a cesty φ, ale již jen přesná definice kladně (záporně) orientované jednoduché uzavřené cesty φ vyžaduje nesnadné předběžné úvahy (srov. [Kr], [Čer]). Je-li však cesta φ skoro regulární, lze tyto pojmy ekvivalentně definovat podstatně elementárněji na základě představy, že „jdeme-li po $\langle\varphi\rangle$ v kladném smyslu, je $\operatorname{Int}\langle\varphi\rangle$ po naší levé ruce“(srov. obr. 5.30 vlevo). Platí totiž tvrzení: ZZZ

5.90 Tvrzení. *Nechť $\varphi\colon [a,b] \to \mathbb{R}^2$ je jednoduchá uzavřená skoro regulární cesta a $G := \operatorname{Int}\langle\varphi\rangle$. Pak platí právě jedno z následujících tvrzení:*

(K) *Je-li $t \in (a,b)$ a existuje $\varphi'(t) \neq 0$, pak vektorový součin $\times(\varphi'(t))$ míří v bodě $\varphi(t)$ ven z G.*
(Z) *Je-li $t \in (a,b)$ a existuje $\varphi'(t) \neq 0$, pak vektorový součin $\times(\varphi'(t))$ míří v bodě $\varphi(t)$ do G.*

Řekneme, že cesta φ je kladně (resp. záporně) orientovaná, platí-li (K) (resp. (Z)).

Na základě Věty 5.89 a Tvrzení 5.90 již můžeme Greenovu větu formulovat:

5.91 Věta. *Nechť φ je jednoduchá uzavřená skoro regulární cesta v $\mathbb{R}^2$. Nechť $F = (F_1, F_2)$ je vektorové pole, které je třídy C^1 na nějaké otevřené množině* *obsahující* $\overline{\operatorname{Int}\varphi}$*. Pak*

$$\int_\varphi F = \pm \int_{\operatorname{Int}\varphi} \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \, \mathrm{d}x. \tag{5.59}$$

Přitom znaménko + (resp. −) se bere, je-li φ *kladně (resp. záporně) orientovaná.*

Tato věta je snadným důsledkem Gaussovy věty. Chceme-li se ale obejít bez Věty 5.89 a Tvrzení 5.90, musíme předpoklady věty formulovat jinak:

5.92 Věta. *Nechť* φ *je jednoduchá uzavřená skoro regulární cesta v* $\mathbb{R}^2$ *a* $G \subset \mathbb{R}^n$ *je taková otevřená omezená množina, že* $\langle\varphi\rangle = \partial G$ *a platí jedna z podmínek (K), (Z) (z Tvrzení 5.90). Nechť* $F = (F_1, F_2)$ *je vektorové pole, které je třídy* C^1 *na nějaké otevřené množině obsahující* $\overline{G}$*. Pak*

$$\int_\varphi F = \pm \int_G \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \, \mathrm{d}x. \tag{5.60}$$

Přitom znaménko + (resp. −) se bere, platí-li podmínka (K) (resp. (Z)).

Důkaz. Nechť φ_* a H_{φ_*} jsou jako v Definici 5.85. Předpokládejme, že platí (K). Nejdříve dokážeme, že $H_{\varphi_*} \subset \partial_* G$ Protože $\langle\varphi\rangle \setminus H_{\varphi_*} = \partial G \setminus H_{\varphi_*}$ je konečná množina, pro každý bod $x \in H_{\varphi_*}$ existuje jeho otevřené okolí U, pro které $U \cap \partial G = U \cap H_{\varphi_*}$. Protože H_{φ_*} je jednoduchá 1-plocha, je také $U \cap H_{\varphi_*}$ jednoduchá 1-plocha. Podmínka (ii) z Definice 5.47 je také splněna: Je-li $x = \varphi_*(t) \in H_{\varphi_*}$ a U je jeho okolí, pak množina $U \setminus \overline{G}$ je neprázdná, protože obsahuje všechny body tvaru $x + h w_{\varphi_*}(t)$ pro dostatečně malá $h > 0$. Z podmínky (K) vyplývá, že φ_* je kladná parametrizace jednoduché 1-plochy H_{φ_*}, kterou orientujeme vnější normálou.

Protože množiny $\partial G \setminus \partial_* G$, $\partial_* G \setminus H_{\varphi_*}$ jsou konečné (a tudíž μ_1-nulové), můžeme použít Gaussovu větu pro vektorové pole

$$\Phi := \times(F) = (F_2, -F_1)$$

a dostáváme

$$\int_{H_{\varphi_*}} \overrightarrow{\Phi} \, \mathrm{d}\overrightarrow{S} = \int_{\partial_* G} \overrightarrow{\Phi} \, \mathrm{d}\overrightarrow{S} = \int_G \operatorname{div} \Phi \, \mathrm{d}x = \int_G \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \, \mathrm{d}x.$$

Protože podle (5.55) platí

$$\int_{H_{\varphi_*}} \overrightarrow{\Phi} \, \mathrm{d}\overrightarrow{S} = \int_\varphi F \, \mathrm{d}\varphi,$$

je důkaz proveden. Platí-li podmínka (Z), je důkaz analogický.

Větu 5.91 zřejmě dostáváme jako okamžitý důsledkem Věty 5.92 (použijeme-li ovšem také Větu 5.89 a Tvrzení 5.90).

Obr. 5.30.

Věta 5.91 (kterou jsme však plně nedokázali) je zřejmě výhodnější v konkrétních aplikacích, protože se podstatně snadněji ověřují její předpoklady (srov. Příklad 5.94).

Za použití symboliky diferenciálních forem (viz (5.57)) můžeme vzorec (5.60) zapsat ve tvaru

$$\int_\varphi F_1 \mathrm{d}x_1 + F_2 \mathrm{d}x_2 = \pm \int_G \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) \mathrm{d}x. \tag{5.61}$$

Užijeme-li tento vzorec pro vektorová pole $F := (f, 0)$ a $\tilde{F} := (0, f)$ (kde f je funkce třídy C^1 na nějaké otevřené množině obsahující $\overline{G}$), dostáváme rovnosti

$$\int_\varphi f \,\mathrm{d}x_1 = \pm \int_G \left(-\frac{\partial f}{\partial x_2} \right) \mathrm{d}x, \qquad \int_\varphi f \,\mathrm{d}x_2 = \pm \int_G \frac{\partial f}{\partial x_1} \,\mathrm{d}x. \tag{5.62}$$

(Tyto rovnosti odpovídají vzorcům (5.46) a nejlépe se pamatují pomocí Poznámky 5.66.)

Pro funkce $f(x_1, x_2) = x_2$ a $f(x_1, x_2) = x_1$ dostáváme klasické vzorce

$$\pm \lambda_2(G) = - \int_\varphi x_2 \,\mathrm{d}x_1 \int_\varphi x_1 \,\mathrm{d}x_2 == \frac{1}{2} \int_\varphi x_1 \,\mathrm{d}x_2 - x_2 \,\mathrm{d}x_1. \tag{5.63}$$

5.93 Poznámka. Při ověřování předpokladů Věty 5.92 musíme nejen dokázat rovnost $\langle \varphi \rangle = \partial G$, ale také podmínku (K) (resp. (Z)). Místo přímého ověřování podmínky (K) můžeme ale postupovat také tak, že nejprve ověříme podmínku

(R) $\quad H_{\varphi_*} \cap \operatorname{int}(\overline{G}) = \emptyset$.

Pak (jako v důkazu Věty 5.92) dostáváme, že $H_{\varphi_*} \subset \partial_* G$, takže podmínku (K) (resp. (Z)) lze formulovat takto:

φ_* je kladná (resp. záporná) parametrizace 1-plochy H_{φ_*}, (kterou uvažujeme orientovanou vnější normálou).

K ověření této podmínky stačí napsat D_{φ_*} jako disjunktní sjednocení otevřených intervalů $I_1, \dots, I_s$ (což jistě lze) a dokázat podmínku

(K^*) (resp. (Z^*)) Existují body $t_k \in I_k$ a vektory $v_k \in \mathbb{R}^2$, $i = 1, \dots, s$ takové, že v_k míří v bodě $\varphi(t_k)$ ven z G (resp. do G) a $\langle v_k, w_\varphi(t_k)\rangle > 0$.

Místo (K) tedy stačí ověřit (A) a (K^*), což může být ve složitějších případech o dost snazší.

5.94 Příklad. (Steinerova hypocykloida) Uvažujme cestu

$$\varphi(t) := (2\cos t + \cos 2t, 2\sin t - \sin 2t), \quad t \in [0, 2\pi],$$

a počítejme obsah množiny ohraničené křivkou $\langle\varphi\rangle$.

a) Okamžitě je vidět, že φ je uzavřená cesta třídy C^1 a snadný výpočet dává, že φ je skoro regulární ($\varphi'(t) = 0$ právě pro $t = 0$, $2\pi/3$, $4\pi/3, 2\pi$). Ukažme, že $\varphi = (\varphi_1, \varphi_2)$ je jednoduchá uzavřená:

Ztotožníme-li $\mathbb{R}^2$ a $\mathbb{C}$, lze psát $\varphi(t) = 2e^{it} + e^{-2it}$. Předpokládejme, že $t_1, t_2 \in [0, 2\pi)$, $t_1 \neq t_2$ a $\varphi(t_1) = \varphi(t_2)$. Položíme-li $v := e^{it_1}$, $w := e^{it_2}$, dostáváme

$$2v + \frac{1}{v^2} = 2w + +\frac{1}{w^2}, \qquad \frac{v + w}{v^2 w^2} = 2,$$

takže $|v + w| = 2$, což zřejmě dává spor.

b) Podle klasické formy Greenovy věty (Věta 5.91) použité na vektorové pole $F(x_1, x_2) = (0, x_1)$ dostáváme

$$\lambda_2(\operatorname{Int}\varphi) = \int_{\operatorname{Int}\varphi} \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right) \,\mathrm{d}x = \pm \int_\varphi F \,\mathrm{d}\varphi = \int_0^{2\pi} \varphi_1'(t)\, \varphi_2'(t) = \pm 2\pi.$$

Tento výpočet okamžitě dává, že φ je kladně orientovaná a $\lambda_2(\operatorname{Int}\varphi) = 2\pi$, protože apriori víme, že $\lambda_2(\operatorname{Int}\varphi) > 0$.

c) Křivka $\langle\varphi\rangle$ (tzv. Steinerova hypocykloida) je zobrazena na obr. 5.30 vpravo. (Geometrický názor jsme zatím ale vůbec nepoužili.) Uveďme dvě interpretace této křivky:

(1) Bod $H := (0, 0)$ je hvězda. Planeta P se pohybuje jednotkovou úhlovou rychlostí kolem H po kružnici s poloměrem 2 v kladném smyslu, přičemž v čase 0 je v bodě $(2, 0)$. Měsíc M se pohybuje kolem P úhlovou rychlostí 2 po kružnici s poloměrem 1 v záporném smyslu, přičemž v čase 0 je v bodě $(3, 0)$. Pak měsíc M se pohybuje po křivce $\langle\varphi\rangle$. (Pohyb popsaný cestou φ vzniká složením dvou kruhových pohybů popsaných cestami $2(\cos t, \sin t)$ a $(\cos 2t, -\sin 2t)$.)

(2) Název „hypocykloida“ odpovídá následující interpretaci: Uvažujme kružnici k o poloměru 1, která má v čase 0 střed v bodě $(0, 2)$ a její bod B (který je v čase 0 v bodě $(0, 3)$. Nechť K je kružnice o středu $(0, 0)$ a poloměru 3. Kutálí-li se nyní kružnice k po kružnici K (je jedno, v jakém smyslu), opíše bod B křivku $\langle\varphi\rangle$.

Nakonec ještě poznamenejme, že

$$G = \operatorname{Int}\varphi = \{(x, y) : (x^2 + y^2)^2 + 8x(3y^2 - x^2) + 18(x^2 + y^2) - 27 < 0\}.$$

d) Ověřit předpoklady Věty 5.92 není vůbec snadné.

5.95 Příklad. Nechť $G \subset \mathbb{R}^2$ je množina omezená křivkou $\langle\varphi\rangle$, kde $\varphi(t) = (t - \sin t, 1 - \cos t)$, $t \in [0, 2\pi]$ (částí cykloidy) a osu x. Představme si G jako hmotnou dvourozměrnou desku s plošnou hustotou 1 a počítejme její těžiště. (Srov. Dodatek 6.6; zde je $\mu = C_G \cdot \lambda_2$.) Pro souřadnice x_t, y_t těžiště tedy platí

$$x_t = (1/\lambda_2(G)) \int_G x \,\mathrm{d}x\mathrm{d}y, \qquad y_t = (1/\lambda_2(G)) \int_G y \,\mathrm{d}x\mathrm{d}y.$$

Ze symetrie cykloidy podle přímky $x = \pi$ lze usoudit, že $x_t = \pi$ (a lze to snadno potvrdit výpočtem). Spočítejme y_t pomocí Greenovy věty. Zde (na rozdíl od předchozího příkladu) nedá příliš práce přesně ověřit předpoklady Věty 5.92. protože φ_1 je rostoucí na $[0, 2\pi]$, snadno vidíme, že $\langle\varphi\rangle$ je graf spojité funkce $g := \varphi_2 \circ (\varphi_1)^{-1}$, která je kladná na $(0, 2\pi)$. Snadno je tedy vidět, že $G = \{(x, y) \colon x \in (0, 2\pi), 0 < y < g(x)\}$ je otevřená množina. Dodefinujeme-li φ na $(2\pi, 4\pi]$ předpisem $\varphi(t) = (4\pi - t, 0]$, snadno se ověří rovnost $\partial G = \langle\varphi\rangle$. Dále platí podmínka (Z) (k ověření je možno použít Poznámku 5.93; snadno vidíme, že v bodech $\varphi((0,\pi))$ míří do G vektor $-e_2$ a v bodech $\varphi((\pi,2\pi))$ vektor e_2). Můžeme tedy podle (5.63) a (5.62) počítat

$$\lambda_2(G) = \int_\varphi y \,\mathrm{d}x = \int_0^{2\pi} (1-\cos t)^2 \,\mathrm{d}t = 3\pi,$$

$$\int_G y \,\mathrm{d}x\mathrm{d}y = \int_\varphi \frac{1}{2}y^2 \,\mathrm{d}x = \frac{1}{2}\int_0^{2\pi} (1-\cos t)^3 \,\mathrm{d}t = \frac{5\pi}{2}.$$

Dostáváme $y_t = 5/6$.

5.96 Poznámka.

(i) Greenova věta se někdy formuluje také pro souvislé omezené otevřené množiny $G \subset \mathbb{R}^2$, jejichž hranice je konečným sjednocením jednoduchých uzavřených křivek. Přesná formulace předpokladů pak opět závisí na tom, užíváme-li Větu 5.89, Tvrzení 5.90 (a jejich zobecnění) nebo ne. V jednoduchých konkrétních příkladech můžeme pro takové množiny užít Gaussovu věty místo Greenovy věty.

(i) Zhruba lze říci, že Gaussova věta v $\mathbb{R}^2$ „ekvivalentní“ Greenově větě. Není to však zcela pravda. V Příkladu 5.94 bychom však při výpočtu podle Gaussovy věty měli problémy s přesným ověřením předpokladů. V Příkladu 5.71 bychom zase museli měnit nejpřirozenější parametrizaci φ regulární hranice (protože φ je definovaná na $(0,\infty)$, takže nelze rozšířit na uzavřenou cestu).

5.14 O Stokesově větě

Název oddílu odpovídá tomu, že v něm podáme jen úvodní nesystematické informace o Stokesově větě a o jejím elegantním a důležitém zobecnění – obecné Stokesově větě. K formulaci klasické Stokesovy věty potřebujeme pojem rotace vektorového pole.

5.97 Definice. *Nechť $F = (F_1, F_2, F_3)$ je vektorové pole třídy C^1 na otevřené množině $G \subset \mathbb{R}^3$. Pak klademe*

$$\operatorname{rot} F(x) = \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}, \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}, \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right), \qquad x \in G.$$

5.98 Poznámka. Je zřejmé, že $\operatorname{rot} F$ je spojité vektorové pole na G. Místo rotace se někdy říká „vír“; pak se píše $\operatorname{curl} F$ místo $\operatorname{rot} F$. (Pokud $F(x)$ má smysl rychlosti proudící tekutiny v bodě x, vektor $\operatorname{rot} F$ dává odpověď na otázku, zda v bodě x je „vír“, a jaké má tento vír vlastnosti.)

5.99 Věta. (klasická Stokesova věta) *Nechť $\varphi\colon [a,b] \to \mathbb{R}^2$ je skoro regulární jednoduchá uzavřená cesta, která je kladně orientovaná. Položme $G := \operatorname{Int}\varphi$. Nechť*

Obr. 5.31.

Ψ je regulární difeomorfismus z $\mathbb{R}^2$ do $\mathbb{R}^3$, který je definován na otevřené množině obsahující $\overline{G}$. Položme

$$\varphi^* := \Psi \circ \varphi, \qquad P := \Psi(G)$$

a na 2-ploše P uvažujme orientaci určenou parametrizací $\Psi \restriction_G$.

Nechť dále F je vektorové pole, které je třídy C^1 na nějaké otevřené množině obsahující $\overline{P}$. Pak

$$\int_{\varphi^*} F \, \mathrm{d}\varphi^* = \int_P \overrightarrow{\operatorname{rot} F} \, \mathrm{d}\overrightarrow{S}\,. \tag{5.64}$$

V této situaci je přirozené nazývat křivku $\langle\varphi^*\rangle$ (orientovanou parametrizací φ^*) „okrajem" plochy P. Na obr. 5.31 „je vidět", že křivka $\langle\varphi^*\rangle$ je „kladně orientovaná vzhledem k orientované ploše" P. To se někdy populárně vysvětluje takto: pokud $\mathbb{R}^3$ ztotožníme s „naším fyzikálním prostorem tak, že přirozená orientace $\mathbb{R}^3$ splývá s pravotočivou orientací fyzikálního prostoru (srov. Dodatek 6.2) a jdeme-li po $\langle\varphi^*\rangle$ ve směru daném parametrizací φ^* tak, že směr od nohou k hlavě odpovídá směru normálového pole orientujícího P, pak plocha P je po naší levé ruce. Pro přesnou definici okraje plochy ve velmi obecných situacích viz [ČM], [LM], [Kow], [KST]).

V případě, že Ψ je třídy C^2, není obtížné Větu 5.99 (pomocí přímočarých, ale trochu nepřehledných výpočtů) dokázat pomocí Greenovy věty (viz [Ko; Věta 15.8.] nebo [KST; Věta 1.9.]).

Zobecněním klasické Stokesovy věty je elegantní a důležitá *obecná Stokesova věta*, která hovoří o integraci diferenciálních forem (viz. [Kow], [LM], [KST]).

Následující nesystematické poznámky o diferenciálních formách nám pouze umožní vyslovit (bez důkazu) jistou formu obecné Stokesovy věty a naznačit, jak z ní plyne klasická Stokesova věta.

Začněme heuristickou poznámkou o názorném geometrickém smyslu integrálu z diferenciální formy (o kterém se v moderních učebnicích ne vždy hovoří).

Půjde o zobecnění Poznámku 5.65

Již dříve (Definice 5.62) jsme definovali integrál diferenciální formy tvaru

$$\omega \;=\; f(x)\,\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_{i-1} \wedge \mathrm{d}x_{i+1} \wedge \cdots \wedge \mathrm{d}x_n$$

(kterou jsme zapisovali bez symbolů $\wedge$) přes orientovanou $(n-1)$-plochu a v Poznámce 5.65 jsme se zmínili o názorné geometrické interpretaci tohoto integrálu.

Nyní budeme uvažovat $1 \leq k < n$ a diferenciální formu (kterou zatím chápeme jako formální výraz)

$$\omega = f(x)\,ddx_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}$$

sk (kde $1 \leq< i_1 < \cdots < i_k \leq n$ a f je reálná funkce na otevřené množině $G \subset \mathbb{R}^n$). Položme si otázku, jak definovat $\int_P \omega$, jestliže $P \subset G$ je orientovaná k-plocha (srov. Dodatek 6.2). Pro jednoduchost uvažujme případ, kdy P je jednoduchá k-plocha orientovaná parametrizací $\varphi\colon D \to P$ $(D \subset \mathbb{R}^k)$. (Pro každé $t \in D$ je tedy $(\frac{\partial\varphi}{\partial t_1}(t), \ldots, \frac{\partial\varphi}{\partial t_k}(t))$ kladná báze tečného prostoru $T_{\varphi(t)}(P)$.)

Názorný geometrický smysl integrálu $\int_P \omega$ je zcela analogický případu $k = n-1$ (srov. Poznámka 5.65):

Uvažujme projekci $\pi\colon (x_1, \ldots, x_n) \mapsto (x_{i_1}, \ldots, x_{i_k})$. Rozložme P na konečně (nebo spočetně) mnoho borelovských množin $P_1, P_2, \ldots$ s velmi malým diametrem a v každé z nich zvolme bod $\xi_j \in P_j$. Pak $\int_P \omega$ je limitou „integrálních součtů“

$$\sum_j f(\xi_j) \cdot (\pm\lambda_k(\pi(P_j))),$$

kde znaménko + se bere, pokud zobrazení $\pi \restriction_{P_j}$ „zachovává orientaci“. Označme $D_j := \varphi^{-1}(P_j)$, $\eta_j := \varphi^{-1}(\xi_j)$ a $\psi := \pi \circ \varphi = (\varphi_{i_1}, \ldots, \varphi_{i_k})$. Pak $\pi(P_j) = \psi(D_j)$, takže lze očekávat, že

$$\lambda_k(\pi(P_j)) = \lambda_k(\psi(D_j)) \approx |J_\psi(\eta_j)| \cdot \lambda_k(D_j) = \left| \frac{D(\varphi_{i_1}, \ldots, \varphi_{i_k})}{D(t_1, \ldots, t_k)}(\eta_j) \right| \cdot \lambda_k(D_j)$$

Přitom lze očekávat, že znaménko + bereme, pokud zúžení π na T_{ξ_j} zachovává orientaci, tj. pokud $\left(\pi\left(\frac{\partial\varphi}{\partial t_1}(\eta_j)\right), \ldots, \pi\left(\frac{\partial\varphi}{\partial t_k}(\eta_j)\right)\right)$ je kladná báze prostoru $\mathbb{R}^k$. To však platí právě tehdy, když $\frac{D(\varphi_{i_1}, \ldots, \varphi_{i_k})}{D(t_1, \ldots, t_k)}(\eta_j) > 0$. Proto lze očekávat, že

$$\int_P \omega \approx \sum_j f(\xi_j) \cdot (\pm\lambda_k(\pi(P_j))) \approx \sum_j f(\varphi(\eta_i))\, \frac{D(\varphi_{i_1}, \ldots, \varphi_{i_k})}{D(t_1, \ldots, t_k)}(\eta_j) \cdot \lambda_k(D_j).$$

Je tedy přirozené definovat

$$\text{(5.65ikfdef)} \qquad \int_P f(x)\,ddx_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k} := \int_D f(\varphi(t)) \cdot \frac{D(\varphi_{i_1}, \ldots, \varphi_{i_k})}{D(t_1, \ldots, t_k)}(t)\;\mathrm{d}t,$$

konverguje-li Lebesgueův integrál napravo.

Jde o přímé zobecnění vzorce (5.45). Tato skutečnost a výše provedené heuristické úvahy napovídají (a není těžké to přesně dokázat), že definice je korektní, tj. nezávisí na výběru kladné orientace φ orientované jednoduché k-plochy P.

Libovolnou diferenciální formu řádu $1 \le k \le n$ na otevřené množině $G \subset \mathbb{R}^n$ lze psát právě jedním způsobem v *kanonickém tvaru*

$$\omega = \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} a_{i_1,\dots,i_k}(x)\, \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}, \tag{5.66}$$

kde $a_{i_1,\dots,i_k}(x)$ jsou funkce na G. Jsou-li tyto funkce třídy C^p $(0 \le p \le \infty)$ na G, říkáme, že ω je třídy C^p na G. Formy tvaru (5.66) se sčítají a násobí reálným číslem přirozeným způsobem a klademe

$$\int_P \omega := \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} \int_P a_{i_1,\dots,i_k}(x)\, \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}.$$

Je však nutno uvažovat diferenciální formy zapsané i v jiných tvarech, například ve tvaru

$$\omega = f(x)\, \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k},$$

sk

kde nyní $1 \le< i_1 \le n, \dots, 1 \le i_k \le n$ jsou nyní *zcela libovolné indexy.* Integrál i takové diferenciální formy definujeme rovností ((?)). V případě, že v zápisu ω „jsou dva stejné diferenciály", tj. $i_r = i_s$ pro některou dvojici $1 \le r < s \le k$, je $\frac{D(\varphi_{i_1},\dots,\varphi_{i_k})}{D(t_1,\dots,t_k)}(t) = 0$ na D, takže $\int_P \omega = 0$ pro každou orientovanou jednoduchou k-plochu $P \subset G$. Je proto přirozené ω považovat za nulovou diferenciální formu.

Jsou-li dále $i_1, \dots, i_k$ indexy z $\{1, \dots, n\}$ a pro indexy $j_1, \dots, j_k$ platí $j_k = i_{\pi(k)}$, kde $\pi\colon \{1, \dots, k\} \to \{1, \dots, k\}$ je permutace, pak podle známých vlastností determinantů na G platí

$$\frac{D(\varphi_{i_1},\dots,\varphi_{i_k})}{D(t_1,\dots,t_k)} = (-1)^{\operatorname{sgn}(\pi)} \frac{D(\varphi_{j_1},\dots,\varphi_{j_k})}{D(t_1,\dots,t_k)}.$$

Platí tedy (pro libovolnou funkci f na G a libovolnou P)

$$\int_P f(x) \cdot \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_k} = (-1)^{\operatorname{sgn}(\pi)} \int_P f(x) \cdot \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}, \tag{5.67}$$

jakmile má jedna strana rovnosti smysl. Proto je přirozené postulovat rovnost

$$f(x) \cdot \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_k} = (-1)^{\operatorname{sgn}(\pi)} f(x) \cdot \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}.$$

Je tedy zřejmé, jak se libovolná diferenciální forma tvaru $\omega = f(x)\, \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}$ převede na kanonický tvar (například $f(x) \cdot \mathrm{d}x_1 \wedge \mathrm{d}x_2 = -f(x) \cdot \mathrm{d}x_2 \wedge \mathrm{d}x_1$).

K formulaci jakékoliv verze obecné Stokesovy věty je nutno ještě definovat („vnější") diferenciál k-formy třídy C^1. Pro případ k-formy třídy C^1 na G, jejíž „kanonický" tvar je

$$\omega = f(x)\, \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}$$

je jejím diferenciálem $(k+1)$-forma tvaru

$$\mathrm{d}\omega := \mathrm{d}f \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k},$$

což je ovšem rovnost, jejíž pravou stranu jsme ještě nedefinovali. Položíme-li však

$$\mathrm{d}f := \frac{\partial f}{\partial x_1}(x)\,\mathrm{d}x_1 + \cdots + \frac{\partial f}{\partial x_n}(x)\,\mathrm{d}x_n \tag{5.68}$$

a předpokládáme „asociativitu vnějšího součinu“, máme

$$\mathrm{d}\omega = \frac{\partial f}{\partial x_1}(x)\,\mathrm{d}x_1 \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k} + \cdots + \frac{\partial f}{\partial x_n}(x)\,\mathrm{d}x_n \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}, \tag{5.69}$$

přičemž nyní má výraz napravo již zřejmý smysl (formy se převedou na kanonický tvar a sečtou). Má-li tedy obecná k-forma třídy C^1 na $G \subset \mathbb{R}^n$ kanonický tvar (5.66), můžeme definovat

$$\mathrm{d}\omega := \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} \mathrm{d}\left(a_{i_1,\ldots,i_k}(x)\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}\right). \tag{5.70}$$

5.100 Příklad. Na otevřené množině $G \subset \mathbb{R}^3$ uvažujme diferenciální formu třídy C_1 tvaru

$$\omega := F_1 \mathrm{d}x_1 + F_2 \mathrm{d}x_2 + F_3 \mathrm{d}x_3.$$

Pak

$$\begin{aligned}
\mathrm{d}\omega &= \left(\frac{\partial F_1}{\partial x_1}\mathrm{d}x_1 + \frac{\partial F_1}{\partial x_2}\mathrm{d}x_2 + \frac{\partial F_1}{\partial x_3}\mathrm{d}x_3\right) \wedge \mathrm{d}x_1 + \left(\frac{\partial F_2}{\partial x_1}\mathrm{d}x_1 + \frac{\partial F_2}{\partial x_2}\mathrm{d}x_2 + \frac{\partial F_2}{\partial x_3}\mathrm{d}x_3\right) \wedge \mathrm{d}x_2 \\
&\quad + \left(\frac{\partial F_3}{\partial x_1}\mathrm{d}x_1 + \frac{\partial F_3}{\partial x_2}\mathrm{d}x_2 + \frac{\partial F_3}{\partial x_3}\mathrm{d}x_3\right) \wedge \mathrm{d}x_3 = \\
&= \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}\right)\mathrm{d}x_2 \wedge \mathrm{d}x_3 - \left(\frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}\right)\mathrm{d}x_1 \wedge \mathrm{d}x_3 + \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2}\right)\mathrm{d}x_1 \wedge \mathrm{d}x_2.
\end{aligned}$$

Při označeních z Věty 5.99 lze tedy psát (viz ???? a ????)

$$\int_{\varphi^*} F\,\mathrm{d}\varphi^* = \int_{\varphi^*} \omega, \qquad \int_P \overrightarrow{\operatorname{rot} F}\,\mathrm{d}\overrightarrow{S} = \int_P \mathrm{d}\omega,$$

takže rovnost ???? z klasické Stokesovy věty lze přepsat do tvaru

$$\int_{\varphi^*} \omega = \int_P \mathrm{d}\omega.$$

Nyní bez důkazu uvedeme jednu (dosti speciální) verzi obecné Stokesovy věty v $\mathbb{R}^n$.

5.101 Věta. *Nechť $1 < k < n$ a otevřená omezená množina $G \subset \mathbb{R}^k$ splňuje předpoklady (Gaussovy) Věty (?) a nechť $P_1 \subset \partial_* G, \ldots, P_s \subset \partial_* G$ jsou jednoduché $(k-1)$-plochy (orientované vnější normálou) takové, že $\mu_{k-1}\left(\partial G \setminus \bigcup_{i=1}^s P_i\right) = 0$. Nechť φ_i je kladná parametrizace plochy P_i, $i = 1, \ldots, s$.*

Nechť $H \supset \overline{G}$ je otevřená množina a $\Psi\colon H \to \mathbb{R}^n$, je regulární homeomorfismus. Na jednoduché k-ploše $\Psi(G)$ uvažujme orientaci indukovanou parametrizací $\Psi \restriction_G$ a na jednoduché $(k-1)$-ploše $\Psi(P_i)$ uvažujme orientaci indukovanou parametrizací $\Psi \circ \varphi_i$ $(i = 1, \dots, s)$.

Nechť ω je diferenciální forma řádu k, která je třídy C^1 na otevřené množině obsahující $\Psi(\overline{G})$. Pak

$$\sum_{i=1}^{s} \int_{\Psi(P_i)} \omega = \int_{\Psi(G)} \mathrm{d}\omega. \tag{5.71}$$

Na základě výpočtů z Příkladu 5.100 není těžké ukázat, že klasická Stokesova věta je důsledkem Věty 5.101.

V předchozím textu jsme již vysvětlili, jak se s diferenciálními formami v $\mathbb{R}^n$ počítá. Výklad však byl dosti vzdálený přesné matematické teorii; dokonce jsme ani přesně nedefinovali, co to diferenciální forma je. Proto se pokusíme krátce vysvětlit obvyklý moderní přístup.

Nechť $(f_1, \dots, f_n)$ je kanonická duální báze v $(\mathbb{R}^n)^*$, $1 \le k \le n$ a jsou dány indexy $1 \le i_1 \le n, \dots, 1 \le i_k \le n$. Nechť $\pi := (f_{i_1}, \dots, f_{i_k})$, takže $\pi\colon (x_1, \dots, x_n) \mapsto (x_{i_1}, \dots, x_{i_k})$.

Nyní definujme funkci

$$f_{i_1} \wedge \dots \wedge f_{i_k}\colon (\mathbb{R}^n)^k \to \mathbb{R}$$

předpisem

$$(f_{i_1} \wedge \dots \wedge f_{i_k})(v_1, \dots, v_k) = \det[\pi(v_1), \dots, \pi(v_k)] = \det(f_{i_r}(v_s))_{r,s=1}^{k}.$$

Z vlastností determinantu je snadno vidět, že $f_{i_1} \wedge \dots \wedge f_{i_k} \in \mathcal{L}_k(\mathbb{R}^n, \mathbb{R})$ (srov. ???): je to k-lineární forma na $\mathbb{R}^n$. Navíc je vidět, že $f_{i_1} \wedge \dots \wedge f_{i_k}$ je *antisymetrická* k-lineární forma. Množina všech antisymetrických k-lineárních forem na konečně dimenzionálním vektorovém prostoru V se obvykle označuje symbolem $\Lambda^k(V^*)$ a jeho prvky se nazývají k-kovektory nebo k-krát kovariantní antisymetrické tenzory (antisymetrické tenzory typu $(0, k)$). Platí tedy $f_{i_1} \wedge \dots \wedge f_{i_k} \in \Lambda^k((\mathbb{R}^n)^*)$. Tento k-kovektor $f_{i_1} \wedge \dots \wedge f_{i_k}$ má jednoduchou geometrickou interpretaci:

Uvažujme projekci $\pi = (f_{i_1}, \dots, f_{i_k})$, libovolný bod $a \in \mathbb{R}^n$ a vektory $v_1, \dots v_k \in \mathbb{R}^n$. Pak $(f_{i_1} \wedge \dots \wedge f_{i_k})(v_1, \dots, v_k)$ má význam „orientovaného objemu rovnoběžnostěnu

$$\pi\left(\mathcal{R}_a(v_1, \dots, v_k)\right) = \mathcal{R}_{\pi(a)}(\pi(v_1), \dots, \pi(v_k)),$$

tj. číslu $\pm \operatorname{vol}(\pi(v_1), \dots, \pi(v_k))$, kde znaménko $+$ (resp. $-$) se bere, jestliže $(\pi(v_1), \dots, \pi(v_k))$ je kladná (resp. záporná) báze $\mathbb{R}^k$ (pokud nejde o bázi, a tedy ani o rovnoběžnostěn, je zkoumaná hodnota nulová).

Ukazuje se, že každý prvek $w \in \Lambda^k((\mathbb{R}^n)^*)$ lze psát jednoznačně ve tvaru (kde $a_{i_1,\dots,i_k}$)

$$w = \sum_{1 \le i_1 < \dots < i_k \le n} a_{i_1,\dots,i_k} \cdot f_{i_1} \wedge \dots \wedge f_{i_k}.$$

Diferenciální k-forma na otevřené množině $G \subset \mathbb{R}^n$ se obvykle definuje jako zobrazení $\omega\colon G \to \Lambda^k((\mathbb{R}^n)^*)$ a diferenciální forma $\omega = \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}$ je konstantní zobrazení, pro které $\omega(x) = f_{i_1} \wedge \cdots \wedge f_{i_k}, \quad x \in G$.

Při tomto obvyklém pojetí opět dostáváme, že každou k-formu na G lze jednoznačně psát v (kanonickém tvaru) (5.66).

Pro diferenciální formu v tomto tvaru můžeme vzorec ((?)) přepsat ve tvaru

$$\int_\varphi \omega = \int_D \omega(\varphi(u))(\varphi'(u)(e_1), \ldots, \varphi'(u)(e_k))\,\mathrm{d}u \tag{5.72}$$

$$\int_\varphi \omega = \int_D \omega(\varphi(t))(\frac{\partial\varphi}{\partial t_1}(t), \ldots, \frac{\partial\varphi}{\partial t_k}(t))\,\mathrm{d}t \tag{5.73.}$$

Tento vzorec lze tedy přijmout za definici integrálu z diferenciální formy, pokud ji chápeme jako tezorovou funkci.

Z něho také vychází následující heuristická „infinitezimální“ interpretace:

Orientovanou plochu P napíšeme jako sjednocení nekonečně malých nepřekrývajících se rovnoběžnostěnů $\mathcal{R}_{b_i}(v_1^i, \ldots, v_k^i)$, $i = 1, 2, \ldots$ tak, aby $(v_1^i, \ldots, v_k^i)$ byla kladně orientovaná báze tečného prostoru $T_{a_i}(P)$. Pak

$$\int_P \omega = \sum_i \omega(b_i)((v_1^i, \ldots, v_k^i)).$$

Vztah k (5.73) ukazuje následující úvaha:

Nechť $\varphi\colon D \to P$ je kladná parametrizace P. Napišme D jako sjednocení konečně (nebo spočetně) mnoha uzavřených velmi malých nepřekrývajících se intervalů $I_1, I_2, \ldots$. nechť $I_i = \mathcal{R}_{a_i}(h_1^i e_1, \ldots, h_k^i e_k)$ $(h_j^i > 0)$. Jeho obraz $\varphi(I_i)$ je „křivočarý rovnoběžnostěn“, který je velmi dobře aproximován rovnoběžnostěnem

$$\mathcal{R}_{\varphi(a_i)}(h_1^i \frac{\partial\varphi}{\partial t_1}(a_i), \ldots, h_k^i \frac{\partial\varphi}{\partial t_k}(a_i)) =: \mathcal{R}_{b_i}(v_1^i, \ldots, v_k^i).$$

Protože $\omega(b_j)$ je multilineární forma, je

$$\sum_i \omega(b_i)((v_1^i, \ldots, v_k^i)) == \sum_i \omega(\varphi(a_i))(\frac{\partial\varphi}{\partial t_1}(a_i), \ldots, \frac{\partial\varphi}{\partial t_k}(a_i)) \cdot h_1^i \cdot h_2^i \cdots h_k^i) \approx \int_\varphi \omega = \int_D \omega(\varphi(t))(\frac{\partial\varphi}{\partial t_1}(t), \ldots, \frac{\partial\varphi}{\partial t_k}(t))\,\mathrm{d}t = \int_P \omega = \int_\varphi$$

6. Dodatky

6.1 Limita vzhledem k bázi filtru

Definice limit posloupnosti reálných čísel a limity (resp. jednostranné limity) reálné funkce reálné proměnné jsou zřejmě založeny na stejném schématu, a proto také důkazy odpovídajících si vět (např. o limitě součinu) lze vést zcela analogicky. Pouze tam, kde u posloupností volíme $n_0 := \max(n_1, n_2)$, v případě funkcí klademe $\delta_0 := \min(\delta_1, \delta_2)$. Proto nepřekvapuje, že existuje přirozený abstraktní pojem limity, jehož speciálním případem oba zmíněné typy limit jsou.

Nechť X je množina a je zadán neprázdný systém $\mathcal{B}$ neprázdných podmnožin množiny X. Pro reálnou funkci f definovanou na nějaké podmnožině X a $A \in \mathbb{R}$ definujme rovnost $\lim_{\mathcal{B}} f = A$ takto:

$$(6.1) \qquad \lim_{\mathcal{B}} f = A \quad \Longleftrightarrow \quad \forall \varepsilon > 0\ \exists B \in \mathcal{B}\ \forall x \in B: \quad f(x) \in U_\varepsilon(A).$$

Je-li $X = \mathbb{R}$ a $a \in \mathbb{R}$, pak zřejmě

$$\lim_{x \to a} f(x) = A \iff \lim_{\mathcal{B}} f = A \ \text{ pro } \ \mathcal{B} := \{(a-\delta, a+\delta) : \delta > 0\}$$

a

$$\lim_{x \to a+} f(x) = A \iff \lim_{\mathcal{B}} f = A \ \text{ pro } \ \mathcal{B} := \{(a, a+\delta) : \delta > 0\}.$$

Pro případ $X = \mathbb{N}$ a $f(n) = a_n$ platí

$$\lim_{n \to \infty} a_n = A \iff \lim_{\mathcal{B}} f = A \ \text{ pro } \ \mathcal{B} := \{\{n_0, n_0+1, \dots\} : n_0 \in \mathbb{N}\}.$$

Takto definovaná limita vzhledem k libovolnému systému $\mathcal{B}$ však nemusí mít vlastnost jednoznačnosti (není těžké si uvědomit, že tuto vlasnost má, pokud $B_1 \cap B_2 \neq \emptyset$ pro každou dvojici $B_1 \in \mathcal{B}$, $B_2 \in \mathcal{B}$). Podstatné je ovšem také to, aby pro limitu vzhledem k systému $\mathcal{B}$ platily obvyklé věty o limitách (např. věty o limitě součtu, součinu a podílu funkcí). Není těžké si rozmyslet, že to, co pro důkazy těchto vět potřebujeme, je tato podmínka:

(*) Pro každé $B_1 \in \mathcal{B}$ a $B_2 \in \mathcal{B}$ existuje $B_3 \in \mathcal{B}$ taková, že $B_3 \subset B_1 \cap B_2$.

Ve všech třech výše uvedených případech je podmínka (*) zřejmě splněna (dokonce můžeme vzít $B_3 := B_1 \cap B_2$). Systémy splňující (*) mají svůj název:

6.1 Definice. *Řekneme, že neprázdný systém $\mathcal{B}$ neprázdných podmnožin množiny X je báze filtru na množině X, jestliže má vlastnost (*).*

Limitu vzhledem k bázi filtry je přirozené definovat i pro zobrazení do obecného metrického prostoru:

6.2 Definice. *Nechť X je množina a $\mathcal{B} \subset \exp X$ je báze filtru. Nechť (Y,ρ) je metrický prostor, $A \in Y$ a f je zobrazení z X do Y. Pak rovnost $\lim_{\mathcal{B}} f = A$ definujeme ekvivalencí* (6.1).

6.3 Poznámka. Říkáme, že neprázdný sytém $\mathcal{F}$ neprázdných podmnožin množiny X je filtr, jestliže pro každé $F_1 \subset X, F_2 \subset X$ platí implikace

(i) $F_1 \in \mathcal{F},\ F_1 \subset F_2 \implies F_2 \in \mathcal{F}$,
(ii) $F_1 \in \mathcal{F},\ F_2 \in \mathcal{F} \implies F_1 \cap F_2 \in \mathcal{F}$

Každý filtr je zřejmě bazí filtru, ale ne naopak. Je-li $\mathcal{B}$ báze filtru a položíme $\mathcal{F} := \{F \subset X\colon \exists B \in \mathcal{B}\colon\ B \subset F$, je zřejmě $\mathcal{F}$ nejmenší filtr, který obsahuje $\mathcal{B}$. V tom případě říkáme, že B je báze filtru $\mathcal{F}$.

Někdy se definuje limita pouze vzhledem k filtru, pro elementární aplikace je však jednodušší pracovat s limitou podle báze filtru

Dále budeme v tomto oddíle pro jednoduchost místo báze filtru psát pouze *„báze“*.

Jak už jsme řekli, pro limitu reálných funkcí vzhledem k bázi platí všechny základní věty o limitách a důkazy jsou zcela analogické důkazům příslušných vět o limitách posloupností (nebo funkcí). Pokud tedy nějaký „limitní proces“ lze chápat jako limitu vzhledem k bázi, víme okamžitě, že pro něj tyto věty platí. Ještě užitečnější je aplikace Věty 6.8, jejíž (ne zcela triviální) důkaz pak nemusíme v konkrétních případech opakovat.

6.4 Příklad. V literatuře se občas vyskytují limity tvaru ($x,\ y \in \mathbb{R}$)

$$\lim_{x\to\infty,\ y\to\infty} f(x,y), \quad \lim_{g(x,y)\to\infty} f(x,y)$$

apod., jejichž smysl je intuitivně jasný, nedají se ale chápat jako limita zobrazení mezi metrickými prostory. Zde jde ovšem (pořadě) o limity vzhledem k bázím

$$\mathcal{B}_1 := \{\{(x,y) \in \mathbb{R}^2\colon\ x > a, y > b\}\colon\ a,b \in \mathbb{R}\};$$
$$\mathcal{B}_2 := \{\{(x,y) \in \mathbb{R}^2\colon\ g(x,y) > c\}\colon\ c \in \mathbb{R}\}.$$

(Množina $\mathcal{B}_2$ je ovšem báze pouze tehdy, když funkce g z $\mathbb{R}^2$ do $\mathbb{R}$ není shora omezená.)

6.5 Příklad. (Riemannův integrál) Nechť $a,b \in \mathbb{R}$, $a < b$. Řekneme, že dvojice konečných posloupností $D^* = ((x_i)_{i=0}^n, (\xi_i)_{i=1}^n)$ je „dělení s body“ (intervalu $[a,b]$), jestliže

$$a = x_0 < x_1 < \cdots < x_n = b \ \text{ a } \ x_{i-1} \le \xi_i \le x_i,\ i = 1,\ldots,n.$$

Normou D^* rozumíme číslo $\nu(D^*) := \max\{x_i - x_{i-1}\colon\ i = 1,\ldots,n\}$ a pro každou funkci f na $[a,b]$ definujeme Riemannův součet

$$\sigma(f, D^*) := \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}).$$

Riemann definoval Riemannův integrál jako limitu:

$$(R)\int_a^b f := \lim_{\nu(D^*)\to 0} \sigma(f, D^*). \tag{6.2}$$

(Ekvivalentní definici pomocí dolních a horních součtů podal o něco později Darboux.) Limitu (6.2) lze přirozeně chápat jako limitu podle báze, podobně jako v Příkladu 6.4. Zde za X bereme množinu všech dělení s body intervalu $[a, b]$, zobrazení $F: X \to \mathbb{R}$ je dáno předpisem $F(D^*) := \sigma(f, D^*)$ a na X uvažujeme bázi

$$\mathcal{B} := \{\{D^* \in X: \nu(D^*) < \delta\}: \ \delta > 0\}.$$

Limitu z (6.2) pak lze chápat jako $\lim_{\mathcal{B}} F$.

6.6 Příklad. (Kurzweil–Henstockův integrál, srov. [LM; 25.2.]) Nechť X a F jsou jako v předchozím příkladu. Položme

$$\mathcal{B}^* := \{\{D^* \in X: x_i - x_{i-1} < \delta(\xi_i),\ i = 1, \dots, n\}: \ \delta: [a, b] \to (0, \infty)\}.$$

Lze (poměrně snadno) dokázat, že $\mathcal{B}^*$ je báze (zřejmé není pouze to, že $\mathcal{B}^*$ neobsahuje prázdnou množinu). Kurzweil–Henstockův integrál reálné funkce f na $[a, b]$ lze definovat rovností

$$(KH)\int_a^b f := \lim_{\mathcal{B}^*} F,$$

je-li limita napravo konečná.

6.7 Definice. (Bolzano-Cauchyova podmínka pro limitu vzhledem k bázi) *Nechť $\mathcal{B}$ je báze na množině X, (Y, ρ) je metrický prostor a f je zobrazení z X do Y. Řekneme, že f splňuje Bolzano-Cauchyovu podmínku vzhledem k bázi $\mathcal{B}$, jestliže platí podmínka*

$$\forall \varepsilon > 0\ \exists B \in \mathcal{B}\ \forall x, y \in B: \ \ \rho(f(x), f(y)) < \varepsilon.$$

Snadno vidíme, že tato Bolzano-Cauchyova podmínka je nutná pro existenci $\lim_{\mathcal{B}} f$. Je-li Y úplný, jde o nutnou a postačující podmínku:

6.8 Věta. *Nechť $\mathcal{B}$ je báze na množině X, (Y, ρ) je úplný metrický prostor a f je zobrazení z X do Y, které splňuje Bolzano-Cauchyovu podmínku vzhledem k bázi $\mathcal{B}$. Pak existuje $\lim_{\mathcal{B}} f$.*

Důkaz. Pro každé $n \in \mathbb{N}$ najděme $B_n \in \mathcal{B}$ tak, aby $\operatorname{diam} f(B_n) < 1/n$ a zvolme $x_n \in B_1 \cap B_2 \cap \dots \cap B_n$. Je snadno vidět, že $(f(x_n))$ je cauchyovská posloupnost, takže $f(x_n) \to y \in Y$. Pro libovolné $\varepsilon > 0$ lze zvolit $n > 2/\varepsilon$ tak, aby $f(x_n) \in U_{\varepsilon/2}(y)$. Protože $x_n \in B_n$ a $\operatorname{diam} f(B_n) < 1/n$, platí $f(B_n) \subset U_\varepsilon(y)$, c.b.d.

Hromadné hodnoty vzhledem k bázi filtru

6.9 Definice. *Nechť $\mathcal{B}$ je báze na množině X, (Y,ρ) je metrický prostor a f je zobrazení z X do Y. Předpokládejme, že existuje množina $B \in \mathcal{B}$, která je částí definičního oboru f.*

Řekneme, že $y \in Y$ je hromadnou hodnotou zobrazení f vzhledem k bázi $\mathcal{B}$, jestliže pro každé $\varepsilon > 0$ a $B \in \mathcal{B}$ existuje bod $x \in B$, pro který $f(x) \in U_\varepsilon(y)$.

Množinu všech hromadných hodnot f vzhledem k $\mathcal{B}$ budeme značit $C(f,\mathcal{B})$.

Přímo z definice je okamžitě vidět, že $Y \setminus C(f,\mathcal{B})$ je otevřená množina, takže platí:

Množina hromadných hodnot $C(f,\mathcal{B})$ je vždy uzavřená.

6.10 Lemma. *Nechť $\mathcal{B}$, X, Y, f jsou jako v Definici 6.9 a $K \subset Y$ je taková kompaktní množina, že $f^{-1}(K) \cap B \neq \emptyset$ pro každou množinu $B \in \mathcal{B}$. Pak $C(f,\mathcal{B}) \cap K \neq \emptyset$.*

Důkaz. Nechť závěr neplatí. Pak ke každému $x \in K$ můžeme zvolit $\varepsilon_x > 0$ a $B_x \in \mathcal{B}$ tak, aby $f^{-1}(U_{\varepsilon_x}(x)) \cap B_x = \emptyset$. Protože K je kompaktní, existují body $x_1, \dots, x_n$, pro které $K \subset U_{\varepsilon_{x_1}}(x_1) \cup \cdots \cup U_{\varepsilon_{x_n}}(x_n)$. Protože $\mathcal{B}$ je báze, zřejmě existuje $B \in \mathcal{B}$ taková, že $B \subset B_{x_1} \cap \cdots \cap B_{x_n}$. Dostáváme $B \cap f^{-1}(K) = \emptyset$, což je spor.

6.11 Věta. *Nechť $\mathcal{B}$, X, Y, f jsou jako v Definici 6.9. Nechť Y je kompaktní a $y \in Y$. Pak platí:*

(i) $C(f,\mathcal{B}) \neq \emptyset$.
(ii) $\lim_{\mathcal{B}} f = y \iff C(f,\mathcal{B}) = \{y\}$.

Důkaz. Tvrzení (i) je speciálním případem předchozího lemmatu. Dále předpokládejme, že $C(f,\mathcal{B}) = \{y\}$ a je dáno libovolné $\varepsilon > 0$. Protože $K := Y \setminus U_\varepsilon(y)$ je kompaktní, podle předchozího lemmatu existuje $B_1 \in \mathcal{B}$, pro kterou $B_1 \cap f^{-1}(K) = \emptyset$. Nechť $B_2 \in \mathcal{B}$ je podmnožina D_f a $B \in \mathcal{B}$, $B \subset B_1 \cap B_2$. Pak $f(B) \subset U_\varepsilon(y)$, takže $\lim_{\mathcal{B}} f = y$. Tím je dokázána obtížnější implikace z (ii), snadný důkaz obrácené implikace vynecháváme.

Dále stále předpokládejme, že $\mathcal{B}$, X, Y, f jsou jako v Definici 6.9 a navíc $Y = \mathbb{R}^*$ a ρ je redukovaná metrika na $\mathbb{R}^*$. Pak $Y = \mathbb{R}^*$ je kompaktní, takže $C(f,\mathcal{B})$ je podle Věty 6.11 neprázdná uzavřená podmnožina $\mathbb{R}^*$. Snadno je vidět, že $C(f,\mathcal{B})$ má největší a nejmenší prvek (pokud $\infty \notin C(f,\mathcal{B})$ a $C(f,\mathcal{B}) \neq \{-\infty\}$, zřejmě $\max C(f,\mathcal{B}) = \sup\,(\mathbb{R} \cap C(f,\mathcal{B}))$).

Můžeme tedy definovat limes superior a limes inferior funkce f vzhledem k bázi $\mathcal{B}$ rovnostmi

$$\overline{\lim}_{\mathcal{B}} f := \max C(f,\mathcal{B}), \quad \underline{\lim}_{\mathcal{B}} f := \min C(f,\mathcal{B}).$$

Pomocí Lemmatu 6.10 je snadné dokázat, že $H := \overline{\lim}_{\mathcal{B}} f$ je jediný prvek $\mathbb{R}^*$ s těmito dvěma vlastnostmi:

Jestliže $-\infty < c < H$, pak pro každou $B \in \mathcal{B}$ existuje $x \in B$, pro který $f(x) > c$.
Jestliže $H < c < \infty$, pak existuje $B \in \mathcal{B}$ taková, že pro každý bod $x \in B$ platí $f(x) \leq c$.

Z Věty 6.11 (ii) okamžitě vyplývá toto tvrzení:

Limita $\lim_{\mathcal{B}} f$ existuje, právě když $\overline{\lim}_{\mathcal{B}} f = \underline{\lim}_{\mathcal{B}} f$.

6.2 Orientace vektorového prostoru

Již na střední škole se užívá pojem orientované úsečky a pojem orientovaného úhlu. V obou případech platí, že daný objekt lze orientovat právě dvěma způsoby, a že orientace nějak souvisí s „udáním směru“. V matematice zavádíme pojem orientace řady jiných objektů. Intuitivně je jasný pojem orientované jednoduché uzavřené křivky – orientace nám udává volbu jednoho ze dvou možných „smyslů (směrů) jejího probíhání“. Totéž platí o orientaci jednorozměrného prostoru $\mathbb{R}$. O něco obtížnější je pojem orientace roviny $\mathbb{R}^2$. Asi nejnázornějším způsobem zadání orientace v rovině je zadání orientace (tj. smyslu probíhání) kružnice nebo jiné jednoduché uzavřené křivky. Tato metoda má však své nevýhody: její přesná formulace není snadná a navíc ji nelze použít při definici orientace prostoru $\mathbb{R}^n$ pro $n \geq 3$.

Tyto nevýhody nemá nejčastěji používaný způsob zadání orientace pomocí volby báze (vektorového prostoru $\mathbb{R}^2$). Musíme ovšem určit, kdy dvě báze (v_1, v_2), (w_1, w_2) prostoru $\mathbb{R}^2$ určují stejnou orientaci. Intuitivní odpověď je přirozená: je to tehdy, když smysl probíhání trojúhelníku o vrcholech $0, v_1, v_2$ (daný pořadím vrcholů) je stejný jako smysl probíhání trojúhelníku o vrcholech $0, w_1, w_2$. Tuto odpověď lze formalizovat, nelze ji ale použít pro vyšší dimenze.

Než vyslovíme obvyklou definici souhlasnosti dvou bazí (tj. toho, že určují stejnou orientaci prostoru), provedeme následující úvahu. Z elementární geometrie víme, že dva shodné geometrické útvary U_1, U_2 v rovině nebo v prostoru mohou být „přímo shodné“ nebo „nepřímo shodné“. Nepřímo shodné jsou například dvě boty ze stejného páru. Jsou shodné, ale „jinak orientované“. Jsou-li U_1, U_2 přímo shodné, můžeme U_1 přemístit na U_2 spojitým pohybem (stačí provést několik otočení), pokud jsou však nepřímo shodné, možné to není (nevystačíme s otáčeními, ale potřebujeme alespoň jedno „zrcadlení“).

Pro dvě ortonormální báze v $\mathbb{R}^2$ již idea „spojitého pohybu“ dává odpověď, kterou lze použít i pro vyšší dimenze: dvě (uspořádané) ortonormální báze (v_1, v_2), (w_1, w_2) jsou souhlasné, jestliže lze jednu spojitým pohybem převést na druhou. Pro obecné báze je nutno ideu spojitého pohybu nahradit ideou „spojité deformace“ (homotopie). Definice homotopičnosti dvou bazí (viz Tvrzení 6.14) je velmi přirozená. Pokud s ní však chceme pracovat, neobejdeme se bez teorie determinantů, kterou také užívá následující (hůře motivovaná) standardní definice.

6.12 Definice. *Nechť V je n-dimenzionální vektorový prostor. Řekneme, že jeho báze $(v_1, \dots, v_n)$ a $(w_1, \dots, w_n)$ jsou souhlasné (souhlasně orientované, ekvivalentní), jestliže determinant matice přechodu od báze $(v_1, \dots, v_n)$ k bázi $(w_1, \dots, w_n)$ je kladný. Pro souhlasné báze budeme psát $(v_1, \dots, v_n) \sim (w_1, \dots, w_n)$. Nejsou-li báze souhlasně orientované, říkáme, že jsou opačně orientované.*

Připomeňme, že matice přechodu $P = (p_{ij})_{i,j=1}^n$ od báze $(v_1, \dots, v_n)$ k bázi $(w_1, \dots, w_n)$ je určena vztahy $w_i = \displaystyle\sum_{j=1}^n p_{ij} v_j,\ i = 1, \dots, n$ a je vždy regulární. Dále je známo, že inverzní matice P^{-1} je maticí přechodu od báze $(w_1, \dots, w_n)$ k bázi $(v_1, \dots, v_n)$. Protože $\det P = \det P^{-1}$, relace $\sim$ je symetrická.

6.13 Poznámka. Někteří autoři nazývají maticí přechodu inverzní matici P^{-1}. To však zřejmě nemá vliv na definici souhlasnosti bazí.

Velmi jednoduchá algebraická Definice 6.12 je ekvivalentní intuitivně pochopitelnější geometrické („homotopické“) definici:

6.14 Tvrzení. *Dvě báze $(v_1, \dots, v_n)$, $(w_1, \dots, w_n)$ prostoru $\mathbb{R}^n$ jsou souhlasné, právě když jsou homotopické, tj. existují spojitá zobrazení*

$$u_1\colon [0,1] \to \mathbb{R}^n, \dots, u_n\colon [0,1] \to \mathbb{R}^n$$

taková, že

(i) pro každé $t \in [0,1]$ je $(u_1(t), \dots, u_n(t))$ báze $\mathbb{R}^n$ a
(ii) $(u_1(0), \dots, u_n(0)) = (v_1, \dots, v_n)$, $(u_1(1), \dots, u_n(1)) = (w_1, \dots, w_n)$.

Důkaz toho, že dvě homotopické báze jsou souhlasné, je snadný. Důkaz obrácené implikace je pracnější, ale také není obtížný.

Nechť dále V je konečně dimenzionální vektorový prostor a $B(V)$ množina všech (uspořádaných) bází prostoru V. Jsou-li $B_1, B_2, B_3 \in B(V)$ tři báze a P_{ij} je matice přechodu od báze B_i k bázi B_j, je známo, že $P_{13} = P_{12} \cdot P_{23}$ (viz [Bi; Věta 10.15]), takže $\det P_{13} = \det P_{12} \cdot \det P_{23}$. Relace $\sim$ je tedy také tranzitivní; protože je zřejmě reflexivní, je to relace ekvivalence na $B(V)$.

Přímým výpočtem determinantu matice přechodu se snadno ověří následující tvrzení:

6.15 Tvrzení. *Nechť $(v_1, \dots, v_n)$ je báze vektorového prostoru V a $\pi : \{1, \dots, n\} \to \{1, \dots, n\}$ je permutace. Pak platí:*

(i) Báze $(v_1, \dots, v_n)$ a $(v_{\pi(1)}, \dots v_{\pi(n)})$ jsou souhlasné právě když $\operatorname{sgn}(\pi) = 1$.
(ii) Báze $(v_1, \dots, v_n)$ a $(-v_1, \dots, v_n)$ jsou opačně orientované.

Snadno vidíme, že (pokud V není triviální) vztah ekvivalence $\sim$ má právě dvě třídy ekvivalence. (Je-li $(v_1, \dots, v_n)$ báze V, pak není ekvivalentní s $(-v_1, \dots, v_n)$ a každá třetí báze je zřejmě ekvivalentní právě s jednou z nich.) Vybereme-li jednu z těchto tříd, říkáme, že jsme prostor V orientovali. Chceme-li o orientaci mluvit jako o matematickém objektu, ztotožňujeme ji s touto třídou $\mathcal{O} \in B(V)/\!\sim$ a orientovaným prostorem V rozumíme dvojici $(V, \mathcal{O})$. Orientaci $\mathcal{O}$ pak nazýváme kladnou orientací a tu druhou (opačnou) nazýváme zápornou orientací. Báze patřící do $\mathcal{O}$ nazýváme kladné (kladně orientované) a ty druhé záporné.

Orientaci na V zadáváme nejčastěji výběrem jedné báze, kterou prohlásíme za kladnou. V $\mathbb{R}^n$ vybíráme zpravidla orientaci která je určena (obsahuje) kanonickou bázi; tuto orientaci nazýváme *přirozená (kanonická, standardní) orientace* prostoru $\mathbb{R}^n$.

Je ihned vidět, že báze $(v_1, \dots, v_n)$ prostoru $\mathbb{R}^n$ s přirozenou orientací je kladná, právě když $\det[v_1, \dots, v_n] > 0$.

6.16 Poznámka. (orientace v reálném světě) Tato poznámka je ovšem nematematická: hovoří o našem „reálném světě“. Vybereme-li ve vesmíru jeden bod (například severní pól), který prohlásíme za počátek, a vzdálenost jistých dvou bodů prohlásíme za jednotkovou, můžeme (na nejjednodušší intuitivní úrovni) náš „fyzikální prostor, ve kterém žijeme“ interpretovat jako unitární prostor. Pro fyzikální teorie, ale také pro běžné lidské dorozumívání, je důležité tento unitární prostor orientovat. Za kladnou orientaci se bere většinou „pravá (pravotočivá)“ orientace, záporná orientace se pak nazývá „levá (levotočivá)". Abychom někomu sdělili, kterou orientaci jsme prohlásili za kladnou, musíme se odvolat na nějakou „fyzikální skutečnost", kterou zná. Člověku (který ví co je levá a pravá ruka) stačí říci, že kladnou (pravotočivou) bází je trojice (v_1, v_2, v_3), pokud v_1 má směr upažené ukazující pravé ruky, v_2 má směr ukazující levé předpažené ruky a v_3 má směr trupu od nohou k hlavě.

6.3 O pojmu k-rozměrné míry v $\mathbb{R}^n$

Pro řadu aplikací je „minimální σ-algebra $\mathcal{P}_k$, na které jsme zde k-rozměrnou míru μ_k v $\mathbb{R}^n$ definovali, příliš malá.

Například v teorii konvexních množin je nutno definovat $\mu_2(A)$, jestliže $A = \partial K$, kde $K \subset \mathbb{R}^3$ je omezená otevřená konvexní množina. Není těžké dokázat, že taková A obecně neleží v $\mathcal{P}_2$; její plošný obsah však lze definovat (pomocí Lebesgueovy míry λ_3) rovností

$$\mu_2(A) := \lim_{\varepsilon \to 0+} \frac{1}{2\varepsilon} \lambda_3(U_\varepsilon(A)), \tag{6.3}$$

kde $U_\varepsilon(A)$ je ε-ové okolí množiny A, tj. $U_\varepsilon(A) := \{x \in \mathbb{R}^3 \colon \operatorname{dist}(x, A) < \varepsilon\}$.

Intuitivně pochopitelný vzorec (6.3) (na základě něhož je definován tzv. Minkowského objem) však nelze použít pro výpočet plošného obsahu dostatečně obecných množin.

V teorii diferenciálních rovnic a ve variačním počtu se nejčastěji pracuje s pojmem Hausdorffovy k-rozměrné míry, jejíž teorie je však dosti obtížná a definice nepříliš názorná.

Pojem k-rozměrné míry značně osvětlila Kolmogorovova práce z r. 1932, ve které jsou formulovány (v trochu jiné formě) následující jednoduché axiomy pro k-rozměrnou míru.

6.17 Definice. *Řekneme, že míra μ definová na σ-algebře $\mathcal{A}$ podmnožin $\mathbb{R}^n$ splňuje Kolmogorovovy axiomy (pro k-rozměrnou míru), jestliže platí:*

(A_1) Je-li $M \in \mathcal{A}$, $f : M \to \mathbb{R}^n$ je c-lipschitzovské zobrazení a $f(M) \in \mathcal{A}$, pak

$$\mu((f(M)) \le c^k \mu(M).$$

(A_2) $\quad \mu\left([0,1]^k \times \{0\}^{n-k}\right) = 1.$

Tyto dva axiomy ovšem neříkají nic o tom, jak velká je σ-algebra $\mathcal{A}$; ani triviální případ $\mathcal{A} = \{\emptyset, \mathbb{R}^n\}$ není vyloučen. Většina měr studovaných v geometrické teorii míry splňuje axiomy (A_1), (A_2) a navíc také třetí axiom, hovořící o tom, že $\mathcal{A}$ je velmi bohatá:

(A_3) Míra μ je zúplněním míry definované na σ-algebře všech borelovských podmnožin $\mathbb{R}^n$,

Míry, které splňují axiomy (A_1), (A_2) i (A_3) je přirozené nazývat k-rozměrnými mírami v Kolmogorovově smyslu; mezi ně patří i nejznámější k-rozměrná Hausdorffova míra. Všechny takové míry jsou rozšířením „minimální“ k-rozměrné míry μ_k. To okamžitě plyne z následující věty, která podává mnohem jednodušší „axiomatickou definici“ míry μ_k na $\mathcal{P}_k$.

6.18 Věta. *Na σ-algebře $\mathcal{P}$ existuje právě jedna míra, která splňuje Kolmogorovovy axiomy (A_1), (A_2). Touto mírou je μ_k.*

Důkaz této věty není snadný. Jeho základní myšlenka spočívá v pozorování, že Kolmogorovovy axiomy (A_1), (A_2) „vysvětlují“ nejméně jasný krok v heuristickém odvození základního vzorce (5.12) pro výpočet jednoduché plochy, totiž „důvod“, proč platí přibližná rovnost (5.11).

Tam jsme se omezili na nejasné konstatování, že je to proto, že množiny $\varphi(G_j)$ a $P_j = \alpha_j(G_j)$ si jsou „k nerozeznání podobné“, což bylo znázorněno na obr. 5.18, na kterém je vidět, že rovnoběžník $\alpha_j(G_j)$ „velmi dobře aproximuje“ „křivočarý rovnoběžník“ $\varphi(G_j)$.

Zde je ovšem podstatné, že plocha $\alpha_j(G_j)$ dobře aproximuje plochu P_j nejenom „ve smyslu polohy“, ale také „ve smyslu směru“.

To můžeme částečně vysvětlit na příkladu délky (tj. 1-rozměrné míry) grafů funkcí na $(0,1)$. Pro velká n graf funkce $f_n(x) = n^{-1} \cdot \sin nx$ jistě velmi dobře aproximuje graf funkce $f(x) = 0$ „ve smyslu polohy“, ale ne „ve smyslu směru“ (protože $f_n'(x) = \cos nx$ a $f'(x) = 0$). Není proto důvod si myslet, že by se při $n \to \infty$ délky grafů funkcí f_n blížily k délce grafu f a také tomu tak není.

Pokud položíme $g_n(x) = n^{-2} \cdot \sin nx$, pak $g_n'(x) = n^{-1} \cdot \cos nx$, takže pro velká n graf funkce g_n dobře aproximuje graf f nejen „ve smyslu polohy“, ale také „ve smyslu směru“, takže lze očekávat, že délky grafů g_n konvergují k délce grafu f (a také tomu tak je).

Zcela upřesnit předchozí ideu o aproximaci „ve smyslu směru“ však není snadné.

Je diskutabilní, nakolik je Kolmogorovův axiom (A1) „intuitivně zřejmý“. Pokud jej však přijmeme za správný, dává nám přesný prostředek, jak o dvou „k nerozeznání podobných množinách“ A, $B \in \mathcal{A}$ přesně dokázat, že $\mu_k(A) \approx \mu_k(B)$. Najdeme-li totiž pro velmi malá ε bijekci $f\colon A \to B$, která je $(1+\varepsilon)$-lipschitzovská a má také $(1+\varepsilon)$-lipschitzovskou inverzi $f^{-1}\colon B \to A$, podle axiomu (A1) platí

$$\mu_k(B) \leq (1+\varepsilon)^k \mu_k(A), \qquad \mu_k(A) \leq (1+\varepsilon)^k \mu_k(B).$$

Víme-li ještě, že $0 < \mu_k(A) < \infty$, můžeme psát $\mu_k(A) \approx \mu_k(B)$.

Pro případ $A := \alpha_j(G_j)$ a $B := \varphi(G_j)$ máme k dispozici přirozenou bijekci $f\colon A \to B$, $f(x) := \varphi(\alpha_j^{-1}(x))$. Dokážeme-li nyní, že má-li G_j velmi malý diametr, jsou f i f^{-1} $(1+\varepsilon)$-lipschitzovská s velmi malým $\varepsilon > 0$, dostáváme přibližnou rovnost $\mu_k(\alpha_j(G_j)) \approx \mu_k(\varphi(G_j))$. Na základě této myšlenky lze dokázat Větu 6.18.

Poznemenejme ještě, že pokud hladkou plochu $P \subset \mathbb{R}^3$ aproximujeme vepsanou po částech afinní („mnohostěnnou“) plochou M velmi dobře „ve smyslu polohy“ (což zřejmě nastane, pokud „stěny“ této plochy jsou velmi malé), neplyne z toho automaticky, že by M také dobře aproximoval P „ve smyslu směru“, a proto může mít M podstatně větší plošný obsah než P. Na této skutečnosti je založen slavný *Schwartzův příklad* (srov. [Kop], [Fi]), který ukazuje, že plošný obsah hladké plochy nelze jednoduše definovat jako limitu plošných obsahů vepsaných po částech afinních ploch. (Pro křivky analogický příklad neexistuje.)

Nakonec ještě poznamenejme, že všechny k-rozměrné míry v $\mathbb{R}^n$ v Kolmogorovově smyslu mají stejnou hodnotu pro všechny borelovské množiny B, které jsou k-rektifikovatelné, tj. pro které existují lipschitzovská zobrazení $f_i\colon [0,1]^k \to \mathbb{R}^n$, $i \in \mathbb{N}$, taková, že $B \subset \bigcup_{i=1}^{\infty} f_i([0,1]^k)$. Na některých „fraktálních" množinách se však různé k-rozměrné míry mohou lišit.

6.4 Důkaz věty o existenci a jednoznačnosti μ_k

6.19 Tvrzení. *Nechť P, Q jsou jednoduché k-plochy v $\mathbb{R}^n$ $(1 \le k \le n)$, které se nedotýkají v bodě $a \in P \cap Q$ (tj. $T_a(P) \neq T_a(Q)$). Pak existuje otevřené okolí U bodu a takové, že $U \cap (P \cap Q)$ je podmnožinou nějaké $(k-1)$-plochy.*

Důkaz. (Stručný.) Je snadno vidět, že existuje otevřené okolí $\widetilde{V}$ bodu a takové, že $\widetilde{V} \cap P$ a $\widetilde{V} \cap Q$ jsou implicitně zadané kusy k-rozměrných C^1 ploch tvaru

$$\widetilde{V} \cap P = \{x \in \widetilde{V} : g_1(x) = 0, \dots, g_{n-k}(x) = 0\},$$
$$\widetilde{V} \cap Q = \{x \in \widetilde{V} : h_1(x) = 0, \dots, h_{n-k}(x) = 0\},$$

kde všechny g_i a h_i jsou třídy $C^1(\widetilde{V})$ a pro každé $x \in \widetilde{V}$ platí

$$\mathrm{h}([\operatorname{grad} g_1(x), \dots, \operatorname{grad} g_{n-k}(x)]) = \mathrm{h}([\operatorname{grad} h_1(x), \dots, \operatorname{grad} h_{n-k}(x)]) = n-k.$$

Protože $T_a(P) \neq T_a(Q)$, platí

$$\operatorname{Lin}\{\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_{n-k}(a)\} \neq \operatorname{Lin}\{\operatorname{grad} h_1(a), \dots, \operatorname{grad} h_{n-k}(a)\},$$

takže existuje $j \in \{1, \dots, n-k\}$ takové, že

$$\mathrm{h}([\operatorname{grad} g_1(a), \dots, \operatorname{grad} g_{n-k}(a), \operatorname{grad} h_j(a)]) = n-k+1.$$

Existuje tedy otevřené okolí $U \subset \widetilde{V}$ bodu a takové, že

$$\mathrm{h}([\operatorname{grad} g_1(x), \dots, \operatorname{grad} g_{n-k}(x), \operatorname{grad} h_j(x)]) = n-k+1 \quad \text{pro} \quad x \in U,$$

takže pro $k > 1$ je

$$Z := \{x \in U : g_1(x) = 0, \dots, g_{n-k}(x) = 0,\ h_j(x) = 0\}$$

implicitně zadaný kus $(k-1)$-rozměrné plochy obsahující $P \cap Q \cap U$.

Je-li $k = 1$, z Věty 2.121 snadno vyplývá, že U lze zvolit tak, že $U \cap Z = \{a\}$.

6.20 Lemma. *Nechť P, Q jsou jednoduché k-plochy v $\mathbb{R}^n$ $(1 \le k \le n)$ a D je množina všech bodů $x \in P \cap Q$, ve kterých se P a Q dotýkají (tj. $T_x(P) = T_x(Q)$). Pak pro každý bod $d \in D$ existuje jeho otevřené okolí U takové, že pro každou borelovskou množinu $B \subset D \cap U$ platí $\mu_k^P(B) = \mu_k^Q(B)$.*

Důkaz. (Stručný.) Nechť φ a ψ jsou pořadě parametrizace P a Q a $d \in D$. Zvolme indexy $1 \le i_1 < i_2 < \cdots < i_k \le n$ takové, že $\dfrac{D(\varphi_{i_1}, \dots, \varphi_{i_k})}{D(t_1, \dots, t_k)}(d) \neq 0$. Pro jednoduchost značení (bez újmy na obecnosti) předpokládejme, že $i_1 = 1, \dots, i_k = k$. Projekce $\pi : (x_1, \dots, x_n) \mapsto (x_1, \dots, x_k)$ tedy zobrazuje $T_d(P)$ na $\mathbb{R}^k$; protože platí

$T_d(P) = T_d(Q)$, snadno dostáváme, že také $\dfrac{D(\psi_{i_1}, \dots, \psi_{i_k})}{D(t_1, \dots, t_k)}(d) \neq 0$. Stejně jako v důkazu Tvrzení 2.150 dostáváme, že existují otevřená okolí U_1, U_2 bodu d taková, že $U_1 \cap P$, $U_2 \cap Q$ jsou explicitně zadané kusy k-rozměrných C^1-ploch (parametrizované souřadnicemi $x_1, \dots, x_k$). Odtud lze snadno usoudit, že existují otevřené okolí U bodu d, otevřené okolí W bodu $(d_1, \dots, d_k)$ a C^1 zobrazení $f : W \to \mathbb{R}^{n-k}$, $g : W \to \mathbb{R}^{n-k}$ taková, že pro $F(w) := (w, f(w))$ a $G(w) := (w, g(w))$ platí $F(W) = P \cap U$ a $G(W) = Q \cap U$. Nechť je dána borelovská množina $B \subset D \cap U$. Množina $C := \pi(B) = F^{-1}(B) = G^{-1}(B)$ je zřejmě také borelovská. Protože pro každé $c \in C$ platí

$$F'(c)(\mathbb{R}^k) = T_{F(c)}(P) = T_{G(c)}(Q) = G'(c)(\mathbb{R}^k)$$

a

$$F'(c)t = (t, f'(c)t), \quad G'(c)t = (t, g'(c)t),$$

snadno vidíme, že $f'(c) = g'(c)$, a tedy i $F'(c) = G'(c)$. Protože zřejmě platí $\mu_k^{P\cap U}(B) = \mu_k^P(B)$ a $\mu_k^{Q\cap U}(B) = \mu_k^Q(B)$, dostáváme

$$\mu_k^P(B) = \int_{F^{-1}(B)} \kappa(F'(t))\, \mathrm{d}\lambda_k(t) = \int_{G^{-1}(B)} \kappa(G'(t))\, \mathrm{d}\lambda_k(t) = \mu_k^Q(B).$$

6.21 Lemma. *Nechť P, Q jsou jednoduché k-plochy v $\mathbb{R}^n$ a $A \subset P \cap Q$ je borelovská množina. Pak*

$$\mu_k^P(A) = \mu_k^Q(A). \tag{6.4}$$

Důkaz. (Stručný.) Nechť D je množina těch bodů $x \in A$, ve kterých se P a Q dotýkají, tj. $T_x(P) = T_x(Q)$. Položme $N := A \setminus D$. Nechť φ a ψ jsou pořadě parametrizace P a Q. Je snadno vidět, že

$$N = \{x \in A :\ h\left[\frac{\partial\varphi(\varphi^{-1}(x))}{\partial t_1}, \dots, \frac{\partial\varphi(\varphi^{-1}(x))}{\partial t_k}, \frac{\partial\psi(\psi^{-1}(x))}{\partial t_1}, \dots, \frac{\partial\psi(\psi^{-1}(x))}{\partial t_k}\right] > k\}$$

je množina otevřená v A. Platí tedy $N = A \cap G$, kde $G \subset \mathbb{R}^n$ je otevřená, takže N i D jsou borelovské.

Uvažujme nyní borelovské míry ω a ω^* na metrickém prostoru N, které jsou definovány rovnostmi

$$\omega(B) := \mu_k^P(B),\ \omega^*(B) := \mu_k^Q(B), \qquad B \in \mathcal{B}(N) = \mathcal{B}(\mathbb{R}^n) \cap \exp(N).$$

a borelovské míry ν a ν^* na D definované rovnostmi

$$\nu(B) := \mu_k^P(B),\ \nu^*(B) := \mu_k^Q(B), \qquad B \in \mathcal{B}(D) = \mathcal{B}(\mathbb{R}^n) \cap \exp(D).$$

Podle Tvrzení 6.19 a Tvrzení 5.20 dostáváme, že ω a ω^* jsou lokálně nulové; podle Tvrzení 6.30 jsou tedy nulové. Míry ν a ν^* se podle Lemmatu 6.20 lokálně rovnají, takže se podle Tvrzení 6.30 rovnají. Platí tedy

$$\mu_k^P(A) = \omega(N) + \nu(D) = \omega^*(N) + \nu^*(D) = \mu_k^Q(A).$$

Označme symbolem $\mathcal{P}_k^*$ systém všech borelovských množin $B \subset \mathbb{R}^n$, které lze pokrýt spočetně mnoha jednoduchými k-plochami. Je zřejmé, že $\mathcal{P}_k^*$ je σ-okruh, tj. obsahuje $\emptyset$ a je uzavřený na množinové rozdíly a spočetná sjednocení. Dále je vidět, že $\mathcal{P}_k^*$ je nejmenší σ-okruh, který obsahuje každou borelovskou podmnožinu každé jednoduché k-plochy, takže $\mathcal{P}_k^* \subset \mathcal{P}_k$. Ovšem $\mathcal{P}_k$ obsahuje také doplňky prvků z $\mathcal{P}_k^*$; množiny jiných typů již neobsahuje; to m.j. ukazuje následující lemma.

6.22 Lemma. *Označme $(\mathcal{P}_k^*)^c := \{\mathbb{R}^n \setminus M : M \in \mathcal{P}_k^*\}$. Pak platí následující tvrzení.*

(i) $\mathcal{P}_k = \mathcal{P}_k^* \cup (\mathcal{P}_k^*)^c$ a $\mathcal{P}_k^* \cap (\mathcal{P}_k^*)^c = \emptyset$.

(ii) Je-li $B \in \mathcal{P}_k^*$, pak existuje její rozklad $B = \bigcup_{i=1}^\infty B_i$ na disjunktní borelovské množiny B_i takové, že každá B_i je podmnožinou nějaké jednoduché k-plochy P_i a $\mu_k^{P_i}(B_i) < \infty$.

(iii) Splňuje-li μ na $\mathcal{P}_k$ vzorec (5.14), pak pro každou $D \in (\mathcal{P}_k^*)^c$ platí $\mu(D) = \infty$.

Důkaz. (Stručný.) (i) Pouze z toho, že $\mathcal{P}_k^*$ je σ-okruh, snadno plyne, že $\mathcal{P}_k^* \cup (\mathcal{P}_k^*)^c$ je σ-algebra. Platí tedy $\mathcal{P}_k = \mathcal{P}_k^* \cup (\mathcal{P}_k^*)^c$. Z Tvrzení 2.156 vyplývá, že $\lambda_n(B) = 0$ pro každou množinu $B \in \mathcal{P}_k^*$, takže $\lambda_n(D) = \infty$ pro každou $D \in (\mathcal{P}_k^*)^c$. Platí tedy $\mathcal{P}_k^* \cap (\mathcal{P}_k^*)^c = \emptyset$.

(ii) Z definice $\mathcal{P}_k^*$ a σ-konečnosti měr $\mu_k^{P_i}$ (viz str. 220) vyplývá, že existují borelovské množiny A_i takové, že každá A_i je podmnožinou nějaké jednoduché k-plochy P_i a $\mu_k^{P_i}(A_i) < \infty$. Nyní stačí položit $B_1 := A_1$ a $B_i = A_i \setminus A_{i-1}$ pro $i > 1$.

(iii) Podle (ii) platí $\lambda_n(\mathbb{R}^n \setminus D) = 0$. Z Fubiniovy věty (v prostoru $\mathbb{R}^n = \mathbb{R}^{n-k} \times \mathbb{R}^k$ s λ_n) ihned vyplývá, že pro λ_{n-k} skoro všechny body $u \in \mathbb{R}^{n-k}$ platí $\lambda_k((\mathbb{R}^n \setminus D)^{u,*}) = 0$. Zvolme jeden takový bod u a položme $P := \{u\} \times \mathbb{R}^k$ a $B := \{u\} \times D^{u,*}$. Zřejmě $B \subset D$ je borelovská množina a $\mu_k^P(B) = \infty$. Tudíž $\mu(D) \geq \mu(B) = \mu_k^P(B) = \infty$.

Nyní již můžeme dokázat Větu 5.18 o existenci a jednoznačnosti míry μ_k na $\mathcal{P}_k^n$.

Důkaz Věty 5.18 (Stručný.) Z Lemmatu 6.22 okamžitě vidíme, že pokud hledaná míra μ_k na $\mathcal{P}_k$ existuje, je určena jednoznačně; je totiž definována nutně takto:

(i) Pro $B \in (\mathcal{P}_k^*)^c$ platí $\mu_k(B) = \infty$.

(ii) Pro $B \in \mathcal{P}_k^*$ platí

$$\mu_k(B) = \sum_{i=1}^\infty \mu_k^{P_i}(B_i), \tag{6.5}$$

kde každá B_i je borelovská podmnožina nějaké jednoduché k-plochy P_i, $\mu_k^{P_i}(B_i) < \infty$ a $B_1, B_2, \ldots$ tvoří disjunktní rozklad množiny B.

Budeme tedy pro každou borelovskou množinu $B \in \mathcal{P}_k$ definovat číslo $\mu_k(B)$ podle (i) a (ii). Aby však tato definice byla korektní, musíme dokázat, že součet v (6.5) nezávisí ani na rozkladu B (tj. na volbě množin B_i), ani na volbě P_i. Nechť je tedy dán jiný takový rozklad množiny B tvořený množinami $B_i^* \subset P_i^*$. Podle Lemmatu 6.21 platí pro $i, j \in \mathbb{N}$ rovnost $\mu_k^{P_i}(B_i \cap B_j^*) = \mu_k^{P_j^*}(B_i \cap B_j^*)$, takže

$$\sum_{i=1}^\infty \mu_k^{P_i}(B_i) = \sum_{i=1}^\infty \sum_{j=1}^\infty \mu_k^{P_i}(B_i \cap B_j^*) =$$

$$= \sum_{j=1}^\infty \sum_{i=1}^\infty \mu_k^{P_j^*}(B_i \cap B_j^*) = \sum_{j=1}^\infty \mu_k^{P_j^*}(B_j^*).$$

Máme tedy na $\mathcal{P}_k$ korektně definovanou množinovou funkci μ_k. Důkaz toho, že μ_k je skutečně míra, je nyní (s použitím základních poznatků o zobecněných řadách čísel) zcela snadné.

6.5 Důkaz Gaussovy věty

Nejprve si uvědomíme to, že při důkazu Tvrzení 5.68 jsme dokázali následující tvrzení.

6.23 Lemma. *Nechť $\Omega \subset \mathbb{R}^{n-1}$ je otevřená množina, funkce $d(x) < h(x)$ jsou třídy C^1 na Ω a (zřejmě otevřená) množina*

$$G := \{(x, y) \in \mathbb{R}^n : d(x) < y < h(x)\}$$

je omezená. Označíme-li

$$D := \{(x, y) :\ y = d(x)\}, \qquad H := \{(x, y) :\ y = h(x)\},$$

pak $D \subset \partial_* G$, $H \subset \partial_* G$ a pro každou funkci f, která je třídy C^1 na nějaké otevřené množině obsahující $\overline{G}$, platí

$$\int_{D \cup H} f \cdot \nu_i \,\mathrm{d}S = \int_G \frac{\partial f(x)}{\partial x_i}\,\mathrm{d}x, \tag{6.6}$$

kde ν je pole vnějších jednotkových normálových vektorů na $\partial_* G$.

Stačí si uvědomit, že předpoklady $\partial G \in \mathcal{P}_{n-1}$ a $\mu_k(\partial G \setminus \partial_* G) = 0$ jsme v důkazu Tvrzení 5.68 potřebovali jen k tomu, abychom mohli napsat $\int_B f \cdot \nu_i \,\mathrm{d}S = 0$. Předpoklad $\mu_k(\partial G) < \infty$ jsme nepotřebovali vůbec, pro platnost *obecné* Gaussovy věty je však nezbytný.

Důkaz (skalární formy) obecné Gaussovy věty povedeme tak, že množinu G „rozložíme“ na spočetně mnoho otevřených množin, na které lze použít předchozí lemma; sečtením rovností (6.6) pak dostaneme dokazovanou rovnost.

Přesný důkaz vyžaduje řadu úvah, proto některé zformulujeme ve formě lemmat. Nejdříve však uděláme několik úmluv:

Budeme jako obvykle psát $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$; pro $x \in \mathbb{R}^{n-1}$ a $y \in \mathbb{R}$ je tedy $(x, y) \in \mathbb{R}^n$. Na všech eukleidovských prostorech budeme uvažovat maximovou normu. Řekneme-li, že $M \subset \mathbb{R}^n$ je graf C^1 funkce f, myslíme tím, že existuje otevřená množina $G \subset \mathbb{R}^{n-1}$ a funkce $f \in C^1(G)$, pro které $M = \{(x, y) : y = f(x)\}$.

Symbolem π budeme značit projekci $\pi : \mathbb{R}^n \to \mathbb{R}^{n-1}$, $\pi(x, y) = x$.

6.24 Lemma. *Nechť P je $(n-1)$-plocha v $\mathbb{R}^n$ a pro bod $c = (a, b) \in P$ platí, že $e_n \notin T_c(P)$. Pak existují čísla $\delta > 0$, $\Delta > 0$, pro která $P \cap (U_\delta(a) \times U_\Delta(b))$ je graf nějaké funkce $f \in C^1\,(U_\delta(a))$.*

Důkaz. (Stručný.) Bez újmy na obecnosti můžeme zřejmě předpokládat, že $P = \{z :\ g(z) = 0\}$, kde $G \subset \mathbb{R}^n$ je otevřená, $g \in C^1(G)$ a $\operatorname{grad} g(z) \neq 0$, $z \in G$. Protože $e_n \notin T_c(P)$, platí $\frac{\partial g(c)}{\partial x_n} \neq 0$, takže závěr lemmatu plyne z věty o implicitních funkcích.

6.25 Lemma. *Nechť $P \subset \mathbb{R}^n$ je $(n-1)$-plocha a $M \subset P$ je borelovská množina taková, že*

(i) $e_n \notin T_z(P)$ *pro každý bod* $z \in M$ *a*
(ii) $\lambda_{n-1}(\pi(M)) = 0$.

Pak $\mu_{n-1}(M) = 0$.

Důkaz. (Stručný.) Podle Lemmatu 6.24 ke každému $x \in M$ existuje jeho otevřené okolí U_x, pro které je $U_x \cap P$ grafem C^1 funkce. S pomocí Poznámky 1.69 snadno vidíme, že lemma stačí dokázat v případě, když P je graf nějaké funkce $f \in C^1(G)$ (kde $G \subset \mathbb{R}^{n-1}$) je otevřená. Položíme-li $\varphi(x) := (x, f(x))$, $x \in G$, je φ zřejmě parametrizace P a $\varphi(\pi(M)) = M$. Platí tedy

$$\mu_{n-1}(M) = \mu_{n-1}(\varphi(\pi(M)) = \int_{\pi(M)} \kappa(\varphi'(x))\,\mathrm{d}x = 0.$$

6.26 Lemma. *Nechť $P \subset \mathbb{R}^n$ je $(n-1)$-plocha a $S := \{z \in P: e_n \in T_z(P)\}$. Pak $\lambda_{n-1}(\pi(S)) = 0$.*

Důkaz. Použijeme-li Poznámku 1.69 analogicky jako v předchozím důkazu, ihned vidíme, že lemma stačí dokázat v případě, že P je jednoduchá $(n-1)$-plocha. Nechť $\varphi = (\varphi_1, \dots, \varphi_n): G \to \mathbb{R}^n$ je parametrizace P a $T := \varphi^{-1}(S)$. Označíme-li $\psi := \pi \circ \varphi$, pak zřejmě $\psi \in C^1(G)$, $\psi(T) = \pi(S)$ a pro každé $t \in T$ platí $J_\psi(t) = 0$. Poslední tvrzení plyne například z toho, že $\psi'(t) = \pi \circ \varphi'(t)$, takže $\psi'(t)(v) = 0$ pro ten (nutně nenulový) vektor $v \in \mathbb{R}^{n-1}$, pro který $\varphi'(t)(v) = e_n$. Podle Tvrzení 6.37 (verze Sardovy věty) platí $\lambda_{n-1}(\pi(S)) = \lambda_{n-1}(\psi(T)) = 0$.

6.27 Lemma. *Nechť $N \subset \mathbb{R}^n$, $N \in \mathcal{P}_{n-1}$ a $\mu_{n-1}(N) = 0$. Pak $\lambda_{n-1}(\pi(N)) = 0$.*

Důkaz. (Stručný.) Podle Lemmatu 6.22 $N \in \mathcal{P}^*_{n-1}$, takže ji lze pokrýt jednoduchými $(n-1)$-plochami $P_1, P_2, \dots$. Stačí tedy dokázat $\lambda_{n-1}(\pi(N \cap P_i)) = 0$, $i \in \mathbb{N}$. Zvolme tedy $i \in \mathbb{N}$ a parametrizaci $\varphi: G \to \mathbb{R}^n$ plochy P_i. Protože $\mu_{n-1}(N \cap P_i) = \int_{\varphi^{-1}(N \cap P_i)} \kappa(\varphi'(t))\,\mathrm{d}t = 0$, nutně $\lambda_{n-1}(T) = 0$, kde $T := \varphi^{-1}(N \cap P_i)$. Zobrazení $\psi := \pi \circ \varphi$ je zřejmě třídy C^1 na G a $\psi(T) = \pi(N \cap P_i)$. Podle Tvrzení 6.38 platí $\lambda_{n-1}(\pi(N \cap P_i)) = \lambda_{n-1}(\psi(T)) = 0$.

6.28 Lemma. *Nechť $F \subset \mathbb{R}^n$ je uzavřená omezená množina, $a \in \mathbb{R}^{n-1}$ a nechť $F^{a,*} = \{y \in \mathbb{R}: (x, y) \in F\}$ je neprázdná konečná množina, která je tvořena body $b_1 < b_2 < \dots < b_r$. Pak pro každou r-tici čísel $\Delta_1 > 0, \dots, \Delta_r > 0$ existuje $\delta > 0$ takové, že pro každé $x \in U_\delta(a)$ platí*

$$F^{x,*} \subset (b_1 - \Delta_1, b_1 + \Delta_1) \cup \dots \cup (b_r - \Delta_r, b_r + \Delta_r).$$

Důkaz. Předpokládejme, že že pro nějaká čísla $\Delta_1 > 0, \dots, \Delta_r > 0$ příslušné $\delta > 0$ neexistuje. Pak lze najít posloupnost bodů $x_n \to a$ a posloupnost čísel $y_n \notin (b_1 - \Delta_1, b_1 + \Delta_1) \cup \dots \cup (b_r - \Delta_r, b_r + \Delta_r)$ tak, že $(x_n, y_n) \in F$. Protože F je kompaktní, existuje vybraná posloupnost $(x_{n_k}, y_{n_k}) \to (a, y) \in F$. Zřejmě platí $y \notin (b_1 - \Delta_1, b_1 + \Delta_1) \cup \dots \cup (b_r - \Delta_r, b_r + \Delta_r)$, což je ale spor s tím, že $y \in F^{a,*} = \{b_1, \dots, b_r\}$.

Víme, že stačí dokázat „skalární formu“ Gaussovy věty (Větu 5.59), jejíž znění zopakujeme:

6.29 Věta. *Nechť $G \subset \mathbb{R}^n$ je omezená otevřená množina, pro kterou*

$$\partial G \in \mathcal{P}_{n-1}, \quad \mu_k(\partial_* G) < \infty, \quad \mu_k(\partial G \setminus \partial_* G) = 0$$

a $\nu = (\nu_1, \dots, \nu_n)$ je pole jednotkových vnějších normálových vektorů na $\partial_* G$. Nechť f je funkce třídy C^1 na nějaké otevřené množině obsahující $\overline{G}$. Pak pro každé $i \in \{1, \dots, n\}$

platí

$$\int_{\partial G} f \cdot \nu_i \, \mathrm{d}S = \int_G \frac{\partial f(x)}{\partial x_i} \, \mathrm{d}x. \tag{6.7}$$

Důkaz. Pro jednoduchost zápisu předpokládejme, že $i = n$; důkaz obecného případu je zcela analogický. Zopakujme, že klademe $\pi(x_1, \ldots, x_n) := (x_1, \ldots, x_{n-1})$. Položme

$$S := \{x \in \partial_* G : \nu_n(x) = 0\} \quad \text{a} \quad N := \partial G \setminus \partial_* G.$$

Množina S je zřejmě uzavřená v borelovské (viz Tvrzení 2.156) množině $\partial_* G$, takže je také borelovská; množina N je zřejmě kompaktní. Z Lemmatu 6.26 snadno vyplývá, že $\lambda_{n-1}(\pi(S)) = 0$ a z Lemmatu 6.27 ihned plyne, že $\lambda_{n-1}(\pi(N)) = 0$. Z definice otevřené množiny snadno vyplývá, že $\pi(G)$ je otevřená množina. Snadno je vidět, že množina $N \cup S$ je uzavřená v množině ∂G, takže je kompaktní. Je tedy kompaktní i množina $\pi(N \cup S) = \pi(N) \cup \pi(S)$, takže množina $C := \pi(G) \setminus (\pi(N) \cup \pi(S))$ je otevřená. Nechť $\{C_\alpha : \alpha \in A\}$ je (prostý) systém všech komponent množiny C; A je zřejmě spočetná. Označme $V := C \times \mathbb{R}$ a $V_\alpha := C_\alpha \times \mathbb{R}$.

Protože $\lambda_{n-1}(\pi(N) \cup \pi(S)) = 0$, platí podle Fubiniovy věty $\lambda_n(G \setminus V) = 0$, takže

$$\int_G \frac{\partial f(x)}{\partial x_i} \, \mathrm{d}x = \int_V \frac{\partial f(x)}{\partial x_i} \, \mathrm{d}x = \sum_{\alpha \in A} \int_{G \cap V_\alpha} \frac{\partial f(x)}{\partial x_i} \, \mathrm{d}x. \tag{6.8}$$

(Integrál vlevo zřejmě konverguje, protože jde o integrál ze spojité omezené funkce přes omezenou otevřenou množinu.)

Nyní ukážeme, že

$$\int_{\partial G} f \cdot \nu_n \, \mathrm{d}S = \int_{\partial G \cap V} f \cdot \nu_n \, \mathrm{d}S = \sum_{\alpha \in A} \int_{\partial G \cap V_\alpha} f \cdot \nu_n \, \mathrm{d}S. \tag{6.9}$$

K tomu si předně uvědomíme, že integrál zcela vlevo konverguje, protože funkce $f \cdot \nu_n$ je omezená a spojitá (a tedy borelovsky měřitelná) na množině $\partial_* G \in \mathcal{P}_{n-1}$, $\mu_{n-1}(\partial_* G) < \infty$ a $\mu_{n-1}(\partial G \setminus \partial_* G) = 0$. Dále snadno vidíme, že

$$\partial G \setminus V = N \cup S \cup M, \quad \text{kde} \quad M := \{z \in \partial_* G \setminus S \colon \pi(z) \in \pi(N) \cup \pi(S)\}.$$

Protože podle předpokladu $\mu_{n-1}(N) = 0$, podle Lemmatu 6.25 platí $\mu_{n-1}(M) = 0$ a $\nu_n(z) = 0$ pro $z \in S$, dostáváme

$$\int_{\partial G \setminus V} f \cdot \nu_n \, \mathrm{d}S = \int_N f \cdot \nu_n \, \mathrm{d}S + \int_S f \cdot \nu_n \, \mathrm{d}S + \int_M f \cdot \nu_n \, \mathrm{d}S = 0,$$

takže platí první rovnost v (6.9). Druhá rovnost je zřejmá.

Stačí tedy pro každé $\alpha \in A$ dokázat, že

$$\int_{\partial G \cap V_\alpha} f \cdot \nu_i \, \mathrm{d}S = \int_{G \cap V_\alpha} \frac{\partial f(x)}{\partial x_i} \, \mathrm{d}x. \tag{6.10}$$

Zvolme tedy $\alpha \in A$ a pro každý bod $x \in C_\alpha$ zkoumejme řez $(\partial G)^{x,*}$. Je-li $y \in (\partial G)^{x,*}$, zřejmě platí $(x,y) \in \partial G_* \setminus S$, takže $\nu_n(x,y) = \langle e_n, \nu(x,y) \rangle \neq 0$. Podle Tvrzení 5.51 (iii) tedy jeden z vektorů e_n, $-e_n$ míří v bodě (x,y) ven z G a druhý do G. Z toho snadno vyplývá, že každý bod množiny $(\partial G)^{x,*}$ je izolovaný. Protože $(\partial G)^{x,*}$ je zřejmě kompaktní, je nutně konečná. Označme její prvky

$$p_1(x) < p_2(x) < .... < p_{r(x)}(x),$$

položme (závislost množin T_i na x pro přehlednost zápisu nevyznačujeme)

$$T_0 := \{x\} \times (-\infty, p_1(x)),\ T_{r(x)} := \{x\} \times (p_{r(x)}(x), \infty) \ \text{ a}$$
$$T_i := \{x\} \times (p_i(x), p_{i+1}(x)) \ \text{ pro } \ i = 1, \dots, r(x) - 1.$$

Protože $T_i \cap \partial G = \emptyset$, zřejmě buď $T_i \subset G$ nebo $T_i \cap \overline{G} = \emptyset$. Protože G je omezená, je nutně $T_0 \cap \overline{G} = \emptyset$, takže v bodě $(x, p_1(x))$ vektor $-e_n$ míří ven z G a e_n do G. Postupně dostáváme

$$T_1 \subset G,\ T_2 \cap \overline{G} = \emptyset,\ T_3 \subset G,\ T_4 \cap \overline{G} = \emptyset, \dots .$$

Protože zřejmě $T_{r(x)} \cap \overline{G} = \emptyset$, je nutně $r(x)$ sudé číslo. Označíme-li $s(x) = r(x)/2$ a

$$p_1(x) =: d_1(x),\ p_2(x) =: h_1(x),\ p_3(x) =: d_2(x),\ p_4(x) =: h_2(x), \dots$$
$$\dots, p_{r(x)-1}(x) =: d_{s(x)}(x),\ p_{r(x)}(x) =: h_{s(x)}(x),$$

platí zřejmě $G^{x,*} = \bigcup_{i=1}^{s(x)} (d_i(x), h_i(x))$.

Nyní ukažme, že pro každé $a \in C_\alpha$ existuje $\omega > 0$ takové, že $U_\omega(a) \subset C_\alpha$, pro každé $x \in U_\omega(a)$ platí $r(x) = r(a)$ a funkce $p_1, \dots, p_{r(a)}$ jsou třídy C^1 na $U_\omega(a)$.

K tomu nejdříve pomocí Lemmatu 6.24 zvolme čísla $\Delta_1 > 0, \dots, \Delta_{r(a)} > 0$ a číslo $\eta > 0$ takové, že $U_\eta(a) \subset C_\alpha$, intervaly $(p_i(a) - \Delta_i, p_i(a) + \Delta_i)$, $i = 1, \dots, r(a)$ jsou po dvou disjunktní a každá z množin

$$\partial G \cap (U_\eta(a) \times (p_i(a) - \Delta_i, p_i(a) + \Delta_i)), \qquad i = 1, \dots, r(a),$$

je grafem funkce z $C^1(U_\eta(a))$. Užijeme-li Lemma 6.28 pro $F := \partial G$, $r := r(a)$, $b^i := p_i(a)$, dostaneme jisté $\delta > 0$. Je snadno vidět, že pak stačí položit $\omega := \min(\delta, \eta)$.

Ukázali jsme tedy, že funkce $r(x)$ je lokálně konstantní na souvislé množině C_α, je tedy na C_α konstantní s hodnotou $r_\alpha =: 2s_\alpha$.

Pro $i = 1, \dots, s_\alpha$ položme $G_i := \{(x, y) \in \mathbb{R}^n :\ d_i(x) < y < h_i(x)\}$,

$$D_i := \{(x, y) :\ y = d_i(x)\}, \quad H_i := \{(x, y) :\ y = h_i(x)\}.$$

Z předchozích úvah je zřejmé, že

$$G \cap V_\alpha = \bigcup_{i=1}^{s_\alpha} G_i, \qquad \partial G \cap V_\alpha = \partial G_* \cap V_\alpha = \bigcup_{i=1}^{s_\alpha} (D_i \cup H_i)$$

a pro $z \in D_i \cup H_i = \partial_* G_i \cap V_\alpha$ je $\nu(z)$ jednotková vnější normála ke G_i v bodě z. Použijeme-li tedy Lemma 6.23 na množiny $G_1, \dots, G_{s_\alpha}$ a příslušné rovnosti (6.6) sečteme, dostáváme (6.10), čímž je důkaz ukončen.

6.6 Některé věty z teorie míry a integrálu

Nechť $(X, \mathcal{A}, \mu)$ je prostor s mírou, tj. $\mathcal{A} \subset \exp X$ je σ-algebra podmnožin množiny X a μ je (nezáporná) míra na $\mathcal{A}$. Pak je pro některé funkce $f: X \to \mathbb{R}^*$ definován abstraktní Lebesgueův integrál $\int_X f \, \mathrm{d}\mu \in \mathbb{R}^*$ (viz [LM] nebo [Ru2]); pak říkáme, že tento integrál existuje. Je-li tento integrál konečný, říkáme, že konverguje.

Integrál $\int_X f \, \mathrm{d}\mu$ se však definuje i pro některé funkce f, které nejsou definovány na celém X; stačí, aby existovala množina $A \subset D_f$, která je měřitelná (tj. patří do $\mathcal{A}$), a pro kterou $\mu(X \setminus A) = 0$ (V případě, že μ je úplná míra, jde o funkce, které jsou definovány μ-skoro všude na X, tj. ty, pro které $\mu(X \setminus D_f) = 0$.) Položíme-li pak $\int_X f \, \mathrm{d}\mu := \int_A f \, \mathrm{d}\mu$, lze snadno dokázat, že definice je korektní: nezáleží na výběru množiny A. (Lze ukázat, že při této definici $\int_X f \, \mathrm{d}\mu = \int_X f \, \mathrm{d}\widehat{\mu}$, má-li jeden z integrálů smysl, přičemž $\widehat{\mu}$ je zúplnění míry μ.)

Je-li X metrický prostor, pak σ-algebru borelovských podmnožin X budeme značit $\mathcal{B}(X)$. Míry na $\mathcal{B}(X)$ budeme nazývat *borelovskými mírami* na X. (Terminologie kolísá, někteří autoři nazývají borelovskými mírami míry definované aspoň na $\mathcal{B}(X)$.)

Symboly λ_n a λ_n^b rozumíme Lebesgueovu míru a borelovskou Lebesgueovu míru na $\mathbb{R}^n$. Pro integrál funkce f vzhledem k λ_n používáme většinou klasické značení $\int_M f(x) \, \mathrm{d}x$ (nebo také $\int_M f(x, y) \, \mathrm{d}x\mathrm{d}y$ pro $n = 2$).

6.30 Tvrzení. *Nechť μ, ν jsou borelovské míry na separabilním metrickém prostoru X. Nechť μ a ν se lokálně rovnají (tj. pro každý bod $x \in X$ existuje jeho otevřené okolí U_x takové, že $\mu(B) = \nu(B)$ pro každou borelovskou množinu $B \subset U_x$). Pak $\mu = \nu$.*

Speciálně, je-li μ lokálně nulová, pak je nulová.

Důkaz. Podle Věty 1.68 existuje posloupnost (x_n) taková, že $X = \bigcup_{n=1}^{\infty} U_{x_n}$. Pro každou $B \in \mathcal{B}(X)$ nyní z rozkladu

$$B = (B \cap U_{x_1}) \cup ((B \cap U_{x_2}) \setminus U_{x_1}) \cup ((B \cap U_{x_3}) \setminus (U_{x_1} \cup U_{x_2})) \cup \ldots$$

ihned vyplývá $\mu(B) = \nu(B)$.

6.31 Věta. (o obrazu míry) *Nechť $(X, \mathcal{A}, \mu)$ je prostor s mírou, $(Y, \mathcal{T})$ je měřitelný prostor (tj. $\mathcal{T}$ je σ-algebra) a $f: X \to Y$ je měřitelné zobrazení (tj. $f^{-1}(E) \in \mathcal{A}$, kdykoliv $E \in \mathcal{T}$). Položíme-li*

$$\nu(E) := \mu(f^{-1}(E)), \qquad E \in \mathcal{T},$$

pak ν je míra na $\mathcal{T}$; tuto míru označujeme $f(\mu)$ a říkáme ji obraz míry μ při zobrazení f.

Je-li $g: Y \to \mathbb{R}^*$ $\mathcal{T}$-měřitelná funkce, pak

$$\int_Y g \, \mathrm{d}f(\mu) = \int_X g \circ f \, \mathrm{d}\mu,$$

jakmile má jedna strana rovnosti smysl.

6.32 Věta. (o neurčitém integrálu) *Nechť $(X, \mathcal{A}, \mu)$ je prostor s mírou a $f \geq 0$ je nezáporná $\mathcal{A}$-měřitelná funkce. Položíme-li*

$$\nu(A) := \int_A f \, \mathrm{d}\mu, \qquad A \in \mathcal{A},$$

pak ν je míra na $\mathcal{A}$; tuto míru označujeme $f \cdot \mu$ a říkáme ji neurčitý integrál f podle μ.

Je-li $g: X \to \mathbb{R}^*$ $\mathcal{A}$-měřitelná funkce, pak

$$\int_X g \, \mathrm{d}(f \cdot \mu) = \int_X g \cdot f \, \mathrm{d}\mu,$$

jakmile má jedna strana rovnosti smysl.

6.33 Poznámka.

(i) Míře $f \cdot \mu$ je také přirozené říkat „míra s hustotou f“ (srov. [LM; 8.19]; je to jediná míra na $\mathcal{A}$, jejíž Radon-Nikodýmova hustota vzhledem k μ je f (tj. $\frac{\mathrm{d}(f \cdot \mu)}{\mathrm{d}\mu} = f$). O míře $f \cdot \mu$ se také hovoří jako o součinu funkce f a míry μ. Někdy (srov. [Ru2; str. 36] se používá zápis $\mathrm{d}(f \cdot \mu) = f \, \mathrm{d}\mu$.

(ii) Důkazy předchozích vět jsou snadné a vedou se stejným standardním postupem; tvrzení se postupně dokáže pro g, která je: a) charakteristická funkce, b) jednoduchá nezáporná funkce, c) nezáporná funkce, d) obecná měřitelná funkce. (Srov. Věta 1.29 z [Ru2] a [LM], 8.23).)

V Kapitole 5 se podstatně využívá následující klasická věta o substituci pro Lebesgueův integrál. Její důkaz je dosti obtížný; lze jej nalézt např. v [J II] nebo [Ru2]. Podstatně obecnější větu lze nalézt v [LM] (Věta 34.18) nebo ve [Fe].

6.34 Věta. (věta o substituci) *Nechť $G \subset \mathbb{R}^k$ je otevřená množina a $\varphi: G \to \mathbb{R}^k$ je prosté regulární zobrazení. Nechť f je funkce z $\varphi(G)$ do $\mathbb{R}^*$. Pak*

$$\int_{\varphi(G)} f(x) \, \mathrm{d}x = \int_G f(\varphi(t)) \cdot |J_\varphi(t)| \, \mathrm{d}t,$$

jakmile jeden z integrálů existuje.

6.35 Poznámka. Zobrazení φ je difeomorfizmus (viz Věta 2.126), takže množina $\varphi(G)$ je otevřená. Aby integrály mohly existovat, je ovšem nutné, aby f byla definována ve skoro všech bodech množiny $\varphi(G)$.

Z Věty 6.34, Tvrzení 6.40 a Tvrzení 5.7 snadno vyplývá následující tvrzení.

6.36 Tvrzení. *Nechť $\varphi: \mathbb{R}^k \to \mathbb{R}^k$ je izometrie a $B \in \mathbb{R}^k$ je borelovská množina. Pak $\lambda_k(\varphi(B)) = \lambda_k(B)$.*

Následující tvrzení je nejjednodušší speciální případ tzv. Sardovy věty.

6.37 Tvrzení. *Nechť $G \subset \mathbb{R}^k$ je otevřená množina a $\varphi: G \to \mathbb{R}^k$ je zobrazení třídy C^1. Nechť $S := \{x \in G\colon J_\varphi(x) = 0\}$. Pak $\lambda_k(\varphi(S)) = 0$.*

Toto tvrzení okamžitě vyplývá z [LM; Věta 10.4] (případ $k = n$ a $\sigma = \lambda_k$). Pro přímý důkaz (který je podstatně jednodušší než důkaz věty o substituci) viz [ČM; Věta 10.4] nebo [FM; Věta 5.35].

Dalším tvrzením, které potřebujeme pro důkaz Gaussovy věty je toto:

6.38 Tvrzení. *Nechť $G \subset \mathbb{R}^k$ je otevřená množina, $\varphi: G \to \mathbb{R}^k$ je zobrazení třídy C^1, $M \subset G$ a $\lambda_k(M) = 0$. Pak $\lambda_k(\varphi(M)) = 0$.*

Důkaz. (Náznak.) Nechť S je jako v Tvrzení 6.37. Snadno vidíme, že S je uzavřená množina, φ je na otevřené množině $H := G \setminus S$ regulární. Pro každý bod $x \in H$ existuje podle Věty 2.121 jeho otevřené okolí U_x, na kterém je φ regulární a prosté. Podle Věty 6.34 (kde za f vezmeme charakteristickou funkci množiny $\varphi(U_x \cap M)$) dostáváme $\lambda_k(\varphi(U_x \cap M)) = 0$. Pomocí Poznámky 1.69 snadno dostáváme $\lambda_k(\varphi(M \setminus S)) = 0$. Podle Tvrzení 6.37 však také $\lambda_k(\varphi(S)) = 0$, takže jsme hotovi. Poznamenejme, že důkaz lze vést snadno přímo bez dosti hluboké věty o substituci; stačí využít toho, že φ je lokálně lipschitzovské.

V této publikaci při různých heuristických úvahách tvrdíme, že jisté „integrální součty" konvergují k některému integrálu. Protože používáme Lebesgueův integrál, který není definován jako limita integrálních součtů, pro lepší pochopení těchto úvah uvádíme následující snadné tvrzení.

6.39 Tvrzení. *Nechť $G \subset \mathbb{R}^n$ je otevřená množina, μ je konečná borelovská míra v $\mathbb{R}^n$ soustředěná v G (tj. $\mu(\mathbb{R}^n \setminus G) = 0$) a nechť f je stejnoměrně spojitá funkce na G. Pak pro každé $\varepsilon > 0$ existuje $\delta > 0$, že kdykoliv zvolíme rozklad G na borelovské množiny $B_1, \dots, B_s$ a body $\xi_i \in B_i$, $i = 1, \dots, s$ tak, že $\operatorname{diam} B_i < \delta$, $i = 1, \dots, s$, pak*

$$\left| \int_{\mathbb{R}^n} f \, \mathrm{d}\mu - \sum_{i=1}^{s} f(\xi_i)\, \mu(B_i) \right| < \varepsilon \tag{6.11}$$

Důkaz. (Náznak.) Nechť $\varepsilon > 0$; k číslu $\varepsilon^* := \varepsilon(\mu(G) + 1)^{-1}$ zvolme $\delta > 0$ z definice stejnoměrné spojitosti. Nechť nyní $B_1, \dots, B_s$ je libovolný příslušný rozklad G. Protože zřejmě

$$\sum_{i=1}^{s} \mu(B_i) \inf_{x \in B_i} f(x) \le \int_{\mathbb{R}^n} f \, \mathrm{d}\mu \le \sum_{i=1}^{s} \mu(B_i) \sup_{x \in B_i} f(x),$$

závěr tvrzení plyne z toho, že podle definice δ se „horní součet" liší od „dolního součtu" nejvýše o ε a číslo $\sum_{i=1}^{s} f(\xi_i)\, \mu(B_i)$ leží mezi nimi.

Každá míra μ na σ-algebře borelovských podmnožin $\mathbb{R}^3$ modeluje rozložení hmoty (nebo kladného elektrického náboje) v prostoru. Ke klasických fyzikálním aplikacím matematiky (které lze pochopit na základě středoškolské fyziky) patří výpočet těžiště a gravitačního (resp. elektrostatického) pole. Níže uvedeme (a částečně motivujeme) příslušné definice, které jsou užity v Příkladech ???? a ?????.

Gravitační (elektrostatické) pole Je-li rozložení hmoty v prostoru dáno konečnou borelovskou mírou μ v $\mathbb{R}^3$, pak příslušné gravitační pole je vektorové pole $F(x)$, kde $F(x)$ je síla, kterou uvažovaná hmota působí na hmotný bod jednotkové hmotnosti umístěný v bodě x. Zkoumejme $F(a)$, kde $a = (a_1, a_2, a_3)$ a pro jednoduchost předpokládejme, že existují čísla $\delta > 0, K > 0$ takové, že μ je soustředěná v množině $M := B(a, K) \setminus B(a, \delta)$ (tj. $\mu(\mathbb{R}^3 \setminus M) = 0$) a gravitační konstanta je 1. Rozdělíme-li nyní M na konečně mnoho borelovských množin $M^1, \dots, M^s$ s velmi malým diametrem a zvolíme body $\xi^i \in M^i$, pak (známe-li Newtonův gravitační zákon) jistě očekáváme, že pro gravitační sílu $\Delta F^i(a)$, kterou působí hmota obsažená v množině M^i na jednotkový hmotný bod umístěný v a, platí

$$\Delta F^i(a) \approx \frac{\mu(M^i)}{\|\xi^i - a\|^2} \cdot \frac{\xi^i - a}{\|\xi^i - a\|} = \frac{\xi^i - a}{\|\xi^i - a\|^3} \mu(M^i).$$

Je tedy přirozené očekávat (srov. Tvrzení 6.39), že pro vektor $F(a)$ a jeho složky $F_i, i = 1, 2, 3$ platí rovnosti

$$F(a) = \int_{\mathbb{R}^3} \frac{x-a}{\|x-a\|^3} \,\mathrm{d}\mu(x), \qquad F_i(a) = \int_{\mathbb{R}^3} \frac{x_i - a}{\|x-a\|^3} \,\mathrm{d}\mu(x). \tag{6.12}$$

Stejným vzorcem definujeme $F(a)$, kdykoliv má příslušný integrál smysl.

Těžiště. Ve fyzice má v různých (statických i kinematických) úvahách základní význam pojem těžiště tělesa (soustavy těles). Je-li příslušné rozložení hmoty dáno konečnou mírou μ, je její těžiště $T(\mu)$ definováno rovností

$$T(\mu) := \frac{1}{\mu(\mathbb{R}^3)} \int_{\mathbb{R}^3} x \,\mathrm{d}\mu(x),$$

konverguje-li integrál vpravo. Vektorová funkce $f(x) = x = (x_1, x_2, x_3)$ se ovšem „integruje po složkách“, takže pro souřadnice T_1, T_2, T_3 těžiště platí

$$T_i = \frac{1}{\mu(\mathbb{R}^3)} \int_{\mathbb{R}^3} x_i \,\mathrm{d}\mu(x), \qquad 1 \le i \le 3. \tag{6.13}$$

Zcela analogicky se těžiště míry definuje (a platí zřejmé zobecnění (6.13)) v obecném $\mathbb{R}^n$.

6.7 O pojmech z lineární algebry

V této publikaci používáme řadu pojmů a výsledků z lineární algebry. Zde nejdříve zavedeme pro některé základní pojmy označení *obvyklá v analýze*. Pak zavedeme některá (nepříliš běžná) označení pro potřeby této publikace a nakonec zopakujeme potřebná fakta týkající se afinních prostorů a afinních zobrazení.

a) Označení obvyklá v analýze

Pod pojmem *lineární prostor* se v analýze rozumí vektorový prostor (nad tělesem $\mathbb{T}$, kde $\mathbb{T} = \mathbb{R}$ nebo $\mathbb{T} = \mathbb{C}$).

Lineární zobrazení je obvyklé označení pro homomorfizmus vektorového (lineárního) prostoru do vektorového (lineárního) prostoru. Místo monomorfismu, epimorfismu a izomorfizmu budeme hovořit o prostém, surjektivním a bijektivním lineárním zobrazení.

Lineární obal podmnožiny M lineárního prostoru (který se v algebře většinou označuje $\langle M \rangle$ nebo $[M]$) se v analýze zpravidla označuje symbolem Lin M.

Značení skalárního součinu i v analýze hodně kolísá, zde používáme symbol $\langle x, y \rangle$.

Nechť U, V jsou dva unitární prostory nad týmž tělesem. Řekneme, že $f: U \to V$ je *unitární zobrazení*, jestliže zachovává skalární součin, tj. $\langle f(x), f(y) \rangle = \langle x, y \rangle$, $x, y \in U$. Každé unitární zobrazení je lineární a prosté (viz [Be], [Bi]).

6.40 Tvrzení. *Nechť U a V jsou reálné unitární prostory. Pak bijekce $g: U \to V$ je izometrie (ve smyslu teorie metrických prostorů), právě když je tvaru $g(x) = f(x) + a$, kde $f: U \to V$ je unitární bijekce a $a \in V$.*

Důkaz lze nalézt v [D II] (par. 3 za Větou 112) pro případ $U = V = \mathbb{R}^n$. Pak stačí použít Větu 14.18 z [Bi].

Množina $\mathbb{R}^n$ (s obvyklou strukturou) se nazývá eukleidovským prostorem, místo o $\mathbb{R}^2$ a $\mathbb{R}$ někdy hovoříme o rovině a reálné přímce.

Kanonickou bázi v $\mathbb{R}^n$ značíme vždy $(e_1, \ldots, e_n)$.

b) Několik speciálních označení

Nulový vektor lineárního prostoru X budeme vždy značit symbolem 0 (případně 0_X, pokud by mohlo dojít k omylu).

Jsou-li $v_1, \ldots, v_k$ prvky prostoru $\mathbb{R}^n$, rozumíme symbolem $[v_1, \ldots, v_k]$ matici typu $n \times k$, jejíž i-tý sloupcový vektor je v_i.

Je-li $L: \mathbb{R}^n \to \mathbb{R}^m$ lineární zobrazení, označujeme jeho matici (typu $m \times n$) symbolem $[L]$.

c) Afinní podprostory eukleidovských prostorů a afinní zobrazení

V této publikaci nepotřebujeme pojem abstraktního afinního prostoru (viz [Bi; str. 128]), stačí nám pojem afinního podprostoru $\mathbb{R}^n$.

Řekneme, že množina $A \subset \mathbb{R}^n$ je afinní podprostor $\mathbb{R}^n$, je-li tvaru $A = a + V := \{a + v\colon\ v \in V\}$, kde V je lineární podprostor $\mathbb{R}^n$.

Lineární prostor V, který je afinním prostorem A jednoznačně určen, se nazývá *zaměření* afinního prostoru A. Je-li V zaměření A, platí $A = b + V$ pro každý bod $b \in A$. Dimenzí afinního prostoru rozumíme dimenzi jeho zaměření, jednorozměrný afinní prostor nazýváme přímka, dvourozměrný rovina a $(n-1)$-rozměrný podprostor $\mathbb{R}^n$ se nazývá nadrovina.

Nechť $A_1 \subset \mathbb{R}^n$, $A_2 \subset \mathbb{R}^m$ jsou afinní prostory a V_1, V_2 jsou jejich zaměření. Řekneme, že zobrazení $f: A_1 \to A_2$ je *afinní zobrazení*, existují-li lineární zobrazení $f_z: V_1 \to V_2$ a body a_1, a_2 takové, že

$$(*) \qquad f(x) = a_2 + f_z(x - a_1).$$

Zobrazení f_z je jednoznačně určeno zobrazením f a vzorec (*) platí, právě když $a_2 = f(a_1)$.

Z Tvrzení 6.40 snadno vyplývá, že zobrazení $f: A_1 \to A_2$ je izometrie, právě když je afinní a zobrazení f_z je unitární bijekce unitárního prostoru V_1 na unitární prostor V_2.

Jsou-li $A_1 \subset \mathbb{R}^n$, $A_2 \subset \mathbb{R}^m$, $A_3 \subset \mathbb{R}^k$ afinní prostory a $f: A_1 \to A_2$, $g: A_2 \to A_3$ jsou afinní zobrazení, pak i jejich složení $g \circ f$ je afinní zobrazení.

6.8 Úmluvy a terminologie

Některé obvyklé úmluvy

a) Při definování nového pojmu má tradičně „implikace význam ekvivalence“. Například fráze „Řekneme, že platí V, jestliže platí W“ vyjadřuje to, že podle definice $V \iff W$.

b) Mluvíme-li o výrokové funkci (predikátu) $V(x)$ jako o pravdivém výroku, myslíme tím, že výroková funkce platí obecně (pro x z dohodnutého oboru). Řekneme-li například, že platí výrok $x > 3 \implies x^2 > 9$; z kontextu je zřejmé, že dohodnutý obor je $\mathbb{R}$.

c) Často místo „$V(x)$ platí pro všechna $x \in M$“píšeme pouze „$V(x),\ x \in M$“.

d) Oba symboly $A := B$, $B =: A$ znamenají to, že A se podle definice rovná B.

e) V případě, že nepřesnost nevede k omylům, se často *úmyslně používají nepřesné zápisy*, které jsou ale přehlednější než zápisy zcela přesné. Například se velmi často vynechávají závorky: pro $f : \mathbb{R}^2 \to \mathbb{R}$ a (x, y) se píše $f(x, y)$ místo $f((x, y))$; místo $f'(x)(v)$ se píše pouze $f'(x)v$ apod. Také bez častého „ztotožňování“ (např. $\mathbb{R}^n \times \mathbb{R}^m$ s $\mathbb{R}^{n+m}$) by (zcela přesné) zápisy byly velmi nepřehledné. Dalším formálně nepřesným značením je klasický zápis funkcí. Hovoříme např. o funkcích $f(x)$, $\sin x$, x^2 (místo o funkcích $f, \sin x \mapsto x^2$). Toto klasické značení často ulehčuje čtení; ve složitějších situacích, kdy může dojít k omylu (jako v Kapitole 3), je nutno používat značení formálně přesné.

Některá obvyklá označení

Symbol $[a, b]$ označuje uzavřený interval v $\mathbb{R}$ s koncovými body $a < b$.

Klademe $\mathbb{R}^* := \mathbb{R} \cup \{\infty, -\infty\}$ a na $\mathbb{R}^*$ definujeme obvyklé operace a uspořádání.

Uzavřeným (resp. otevřeným) intervalem v $\mathbb{R}^n$ rozumíme množinu $I_1 \times \dots I_n$, kde všechny intervaly I_i jsou otevřené (resp. uzavřené). (Uzavřenost a otevřenost intervalů v $\mathbb{R}$ bereme ve smyslu metrických prostorů.) Je-li $M \subset X$, pak $C_M : X \to \mathbb{R}$ je charakteristická funkce množiny M.

Slovo „systém“ se užívá ve dvojím smyslu (z kontextu je vždy patrno, ve kterém). Například systém množin je nejčastěji jen přehlednější označení množiny množin; může však také jít o „indexovaný“ systém množin (tj. soubor množin) $(M_\alpha)_{\alpha \in A}$ (který zapisujeme také $M_\alpha (\alpha \in A)$).

Je-li X vektorový prostor a $a, b \in X$, pak uzavřenou úsečkou $\overline{ab}$ rozumíme množinu $\{a + t(b - a) :\ t \in [0, 1]\}$. Množina $C \subset X$ se nazývá konvexní, jestliže $\overline{ab} \subset C$, kdykoliv $a, b \in C$.

Terminologie týkající se zobrazení není v literatuře jednotná; zde používáme následující přístup.

Zobrazení f je dáno, když známe jeho definiční obor (který značíme D_f) a víme, jaké obrazy $f(x)$ mají prvky $x \in D_f$. Užíváme běžnou symboliku pro zadávání zobrazení; například oba zápisy

$$f(x) := x^2, \quad x \in (0, 1); \qquad\qquad f : x \mapsto x^2, \quad x \in (0, 1)$$

zadávají zobrazení f s definičním oborem $(0, 1)$, které číslu $x \in (0, 1)$ přiřazuje číslo x^2.

Občas používáme značení, kdy nezávisle proměnnou neznačíme písmenem, ale tečkou. Například místo $x \mapsto f'(x)$ píšeme $f'(\cdot)$ apod.

Zobrazení f je dáno, právě když je dán jeho graf $G_f := \{(x, f(x) :\ x \in D_f\}$ a někdy se f a G_f ztotožňují; my to však (jak je v analýze z tradičních terminologických důvodů obvyklé) dělat nebudeme.

Pokud $A = D_f$ a pro každé $x \in A$ platí, že $f(x) \in B$, říkáme, že f je zobrazení *množiny A do množiny B*, a píšeme $f\colon A \to B$. Množina B však není zobrazením f jednoznačně určena. Je-li dáno $f\colon A \to B$ a $A \subset X$, říkáme, že f je zobrazení *z množiny X do B*. Tvrdíme-li, že pro každé $x \in M$ má $f(x)$ nějakou vlastnost, je v tom implicitně obsaženo, že $M \subset D_f$. Jen tehdy, když $M \subset A$, definujeme její obraz

$$f(M) = \{f(x)\colon\ x \in M\} := \{y \in B\colon\ \exists x \in M\colon\ y = f(x)\}.$$

Vzor $f^{-1}(M) = \{x \in A\colon f(x) \in M\}$ definujeme pro libovolnou množinu M. Pro prosté zobrazení f je definováno obvyklým způsobem inverzní zobrazení $f\colon H_f \to D_f$, kde $H_f := f(D_f)$ je obor hodnot zobrazení f. Připouštíme také *prázdné zobrazení*, pro které $D_f = G_f = \emptyset$.

Jsou-li f, g dvě zcela libovolná zobrazení, definujeme *vždy* jejich složení $g \circ f$ předpisem $g \circ f(x) := g(f(x)),\quad x \in f^{-1}(D_g)$. Tímto složením může být ovšem i prázdné zobrazení. To je definice obecnější, než se obvykle v učebnicích uvádí, Je však zcela přirozená a užitečná: složené zobrazení je dáno předpisem $x \mapsto g(f(x))$ pro všechna x, pro které má výraz $g(f(x))$ smysl. I při této obecné definici zřejmě platí „asociativní zákon“: $h \circ (g \circ f) = (h \circ g) \circ f$.

Podobně vždy definujeme součin dvou reálných (resp. komplexních) funkcí. Také pro zobrazení $f_1\colon A_1 \to B_1, f_2\colon A_2 \to B_2$ definujeme zobrazení $(f,g)\colon (A_1 \cap A_2) \to B_1 \times B_2$ předpisem $(f,g)(x) := (f(x), g(x))$.

Funkcí se někdy rozumí zobrazení do čísel nebo do vektorového prostoru (vektorová funkce), často je však funkce pouze jiný název pro zobrazení.

Nakonec ještě zdůrazněme, že *často říkáme, že zobrazení f má nějakou vlastnost, i když jde o vlastnost f vzhledem k dalším objektům.* Řekneme-li, že $f\colon A \to B$ je surjektivní (surjekce) nebo bijektivní (bijekce), není to vlastnost zobrazení f, ale dvojice (f, B).

Řekneme-li, že zobrazení z X do B má nějakou vlastnost, může jít o vlastnost trojice (X, f, B), případně struktur (například struktury metrického nebo vektorového prostoru) zadananých na X a B.

Literatura

[Be] J. Bečvář, *Lineární algebra*, matfyzpress, Praha, 2000.

[Bi] L. Bican, *Lineární algebra a geometrie*, Academia, Praha, 2000.

[Ca] H. Cartan, *Calcul différentiel, Formes différentielles*, Hermann, Paris, 1967. (Francouzsky, existuje ruský překlad)

[Čech] E. Čech, *Bodové množiny*, NČSAV, Praha, 1966.

[Čer] I. Černý, *Základy analysy v komplexním oboru*, Academia, Praha, 1967.

[ČM] I. Černý, J. Mařík, *Integrální počet II*, SPN, Praha, 1961.

[D I] V. Jarník, *Diferenciální počet I*, Academia, Praha, 1984.

[D II] V. Jarník, *Diferenciální počet II*, Nakladatelství ČSAV, Praha, 1956.

[Fe] H. Federer, *Geometric measure theory*, Springer-Verlag, 1969. (Anglicky)

[Fi] G. M. Fichtengol'c, *Kurs differencial'nogo i integral'nogo isčislenija, I-III*, Fizmatgiz, Moskva, 1963. (Rusky)

[FM] S. Fučík, J. Milota, *Matematická analýza II. Diferenciální počet funkcí více proměnných*, SPN, Praha, 1975.

[J II] V. Jarník, *Integrální počet II*, Nakladatelství ČSAV, Praha, 1955.

[Kop] J. Kopáček, *Matematická analýza pro fyziky (III)*, matfyzpress, Praha, 1999.

[Kow] O. Kowalski, *Základy matematické analýzy na varietách*, Universita Karlova, Praha, 1975.

[Kr] J. Král, *Teorie potenciálu I*, SPN, Praha, 1965.

[KF] A.N. Kolmogorov, S.V. Fomin, *Základy teorie funkcí a funkcionální analýzy*, SNTL, Praha, 1975.

[KST] L. Krump, V. Souček, J. Těšínský, *Matematická analýza na varietách*, Karolinum, Praha, 1998.

[LM] J. Lukeš, J. Malý, *Míra a integrál*, Karolinum, Praha, 1993.

[Mi] L. Mišík, *Funkcionálna analýza*, Alfa, Bratislava, 1989.

[Ru1] W. Rudin, *Principles of mathematical analysis*, McGraw-Hill, New York, 1964. (Anglicky, existuje ruský překlad)

[Ru2] W. Rudin, *Reálná a komplexní analýza*, Academia, Praha, 1977.

[Si] R. Sikorski, *Diferenciální a integrální počet. Funkce více proměnných.*, Academia, Praha, 1973.

[] .

[Zo] V. A. Zorič, *Matematičeskij analiz, I, II*, Nauka, Moskva, 1984. (Rusky)

Rejstřík