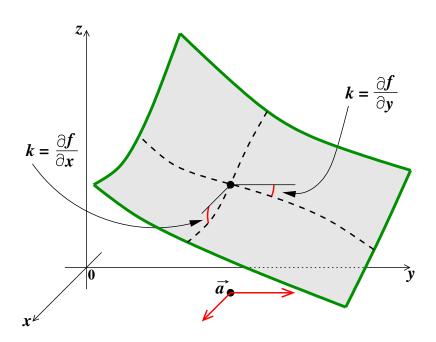


Ilustrovaný úvod do funkcí více proměnných

Petr Habala

 $\begin{array}{c} \texttt{habala@fel.cvut.cz} \\ \text{FEL, \ \ } \check{\text{C}}\text{VUT Praha} \\ 2022 \end{array}$



Úvod

Cílem této brožurky je představit důležité pojmy teorie funkcí více proměnných tak, aby je čtenář dokázal chápat intuitivně a dokázal si propojit matematické vzorce s geometrickými představami. Slouží tak jako doplněk, nikoliv náhrada klasických knih o analýze více proměnných, které vykládají teorii a obsahují důkazy či cvičné příklady, což zde naopak zcela chybí.

Značení je standardní, pro jistotu přidáváme na konci Dodatky se shrnutím základních topologických pojmů používaných při práci s vícerozměrnými prostory. Přidali jsme i přehled základních analytických objektů v rovině a prostoru, protože se s nimi hodně pracuje.

Také formátování je vcelku standardní. Máme tu jednu zvláštnůstku, všechny Příklady a Poznámky jsou zakončeny trojúhelníčkem, který vidíte níže.

 \triangle

Obsah 1a. Definiční obor 1b. Vizualizace funkcí (náčrt grafu) 2. Představujeme limitu a spojitost......11 3a. Derivace ve vyšší dimenzi 3b. Význam parciálních derivací 3c. Parciální derivace vyššího řádu 3d. Význam derivací vyššího řádu 4a. Lokální extrémy 4b. Globální extrémy 5a. Parametrické křivky 5b. Základní analytické pojmy 5c. Diferenciální operátory (div. rot) 6a. Dvoudimenzionální integrál 6b. Třídimenzionální integrál 6c. Křivkové a povrchové integrály (včetně divergenční a Stokesovy věty) 6d. Parametrizace, substituce 7a. Derivace a diferenciální operátory 7b. Skládání funkcí, transformace (včetně změny souřadnic) 7c. Diferenciál 7d. Taylorův polynom 7e. Implicitní křivky a funkce 8a. Množiny 8b. Integrál 9a. Geometrické objekty v \mathbb{R}^n 9b. Topologické pojmy v \mathbb{R}^n

1. Představujeme funkce více proměnných

Funkce více proměnných jsou přirozeným zobecněním funkcí jedné proměnné. Funkce jedné proměnné je zobrazení (jakési posílátko) z jedné kopie reálných čísel (typicky její podmnožiny) do druhé, funkci více proměnných dostaneme, když výchozí jednorozměrnou množinu nahradíme množinou vícerozměrnou.



Lidé si obvykle s pojmem funkce spojují vzorec (což není vždy možné, ale typicky to tak potkáváme), i pak je funkce více proměnným zcela logickým krokem od vzorců typu $f(x) = x^2$ ke vzorcům typu $f(w,x,y,z) = (e^w + y)^x \cdot \sin(x - \pi z)$. Toto značení je velice pohodlné pro práci s konkrétními funkcemi, kdy známe počet proměnných a jejich jména, ale nehodí se pro obecné úvahy, kdy pracujeme s neznámým počtem proměnných. V takovém případě se vyplatí vidět situaci tak, že pořád máme jednu proměnnou, jen se tentokráte bere z množiny vektorů. Namísto f(x,y), f(x,y,z), f(u,v,w,x,y) atd. prostě píšeme $f(\vec{x})$.

V tomto značení je pak také práce s funkcemi více proměnných v mnohém obdobná práci s jednou proměnnou, hlavní myšlenky zůstávají, jen jsou výpočty a postupy občas komplikovanější. Vše se ale zvládá výrazně snáze, pokud si to člověk umí propojit s geometrickou představivostí, zde se hodí zkušenost se základními objekty ve vícedimenzionálním prostoru (přímky, roviny atd.)

To je také hlavním cílem těchto kapitolek, ukázat na geometrický význam pojmů, výpočtů a postupů. Nejprve si dáme definici funkce více proměnných, ať víme, s čím budeme pracovat.

Definice.

Funkcí více proměnných rozumíme libovolné zobrazení $f: D \mapsto \mathbb{R}$, kde D = D(f) je nějaká podmnožina \mathbb{R}^n .

Není-li D explicitně dáno, pak jako definiční obor D(f) bereme množinu všech $\vec{x} \in \mathbb{R}^n$, pro kterou má $f(\vec{x})$ smysl.

Což nás přivádí k prvnímu tématu.

1a. Definiční obor

U funkcí zadaných vzorcem se definiční obor hledá obdobně jako u funkce jedné proměnné, tedy položíme si otázku, jaká omezení na proměnné vyplývají z onoho daného vzorce.

Na rozdíl od případu jedné proměnné se nemusíme snažit vyjádřit výslednou množinu v nějakém standardizovaném tvaru (sjednocení intervalů), protože to ve více rozměrech prostě nelze, množiny jsou příliš rozmanité na to, aby se daly zapsat pomocí jednoho základního typu (či několika typů). Ušetříme tedy práci, stačí výsledek napsat ve tvaru

$$D(f) = {\vec{x} \in \mathbb{R}^n; \text{ podmínky pro } \vec{x}}.$$

Naopak nová věc je, že u množin ve vícerozměrném prostoru někdy můžeme rozeznat, jaký je to vlastně objekt.

Příklad: Určíme definiční obory těchto funkcí:

- a) $f(x,y) = x^2 \sin(x+y)$: evidentně $D(f) = \mathbb{R}^2$.
- b) $f(x,y) = \sqrt{9 x^2 y^2}$: $D(f) = \{(x,y) \in \mathbb{R}^2; \ x^2 + y^2 \le 3^2\},$

je to kruh se středem v počátku o poloměru 3.

c) $f(x,y,z) = \sqrt{9 - x^2 - y^2 - z^2}$: $D(f) = \{(x,y,z) \in \mathbb{R}^3; \ x^2 + y^2 + z^2 \le 3^2\},$

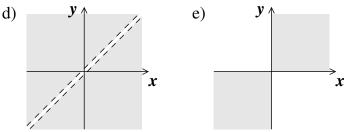
je to koule se středem v počátku o poloměru 3.

d)
$$f(x,y) = \frac{e^{x+y}}{x-y}$$
: $D(f) = \{(x,y) \in \mathbb{R}^2; y \neq x\},\$

d) $f(x,y) = \frac{e^{x+y}}{x-y}$: $D(f) = \{(x,y) \in \mathbb{R}^2; \ y \neq x\}$, je to rovina s vynechanou přímkou—hlavní diagonálou (viz obrázek níže).

e)
$$f(x,y) = \sqrt{x \cdot y}$$
: $D(f) = \{(x,y) \in \mathbb{R}^2; \ y \cdot x \ge 0\},\$

je to uzavřený první a třetí kvadrant v rovině.

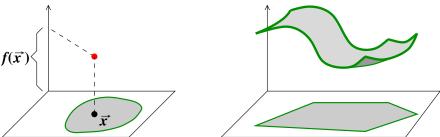


 \triangle

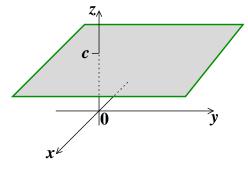
1b. Vizualizace funkcí (náčrt grafu)

Funkce jedné proměnné si nejčastěji vizualizujeme pomocí grafu. Zobecnění na více proměnných je vcelku snadné. V jedné proměnné je grafem funkce množina všech bodů ve tvaru (x, f(x)) pro $x \in D(f)$, potřebuje tedy dvě dimenze. Pokud se proměnná funkce $f(\vec{x})$ bere odněkud z prostoru \mathbb{R}^n , pak na graf budeme potřebovat o jednu dimenzi víc a vznikne jako množina bodů ve tvaru $(x_1, x_2, \dots, x_n, f(\vec{x})), \text{ kde } \vec{x} = (x_1, \dots, x_n).$

V případě $n \geq 2$ tedy potřebujeme alespoň tři dimenze, čímž vyskočíme z papíru a hned vidíme, že graf vlastně nelze nakreslit. Přesto je dobré si zachovat alespoň pocitovou představu grafu. Na "vodorovné" reprezentaci definičního oboru (přesné či symbolické) najdeme výchozí hodnotu vícerozměrné proměnné, nad ní zakreslíme bod grafu ve výšce odpovídající funkční hodnotě v daném bodě.



Pokud to uděláme ve všech bodech definičního oboru, body grafu vytvoří jakýsi útvar, který si můžeme představit jako zvlněný povrch. V zásadě takto to bude fungovat v případě dvou proměnných, typickým grafem je nějaký tenký (v principu dvourozměrný) objekt coby součást třírozměrného prostoru, při troše štěstí si jej dokážeme perspektivou, vhodným stínováním a podobně alespoň přiblížit. Například grafem konstantní funkce f(x,y)=c bude vodorovná rovina vznášející se nad rovinou xy v patřičné výšce c.



Zkušenost by měla naznačit, že jsou i funkce dvou proměnných, u kterých to tak pěkné nebude, grafem je obecně nějaká množina v \mathbb{R}^3 , která může být i velmi divoká, ale takové funkce nejspíše v běžných aplikacích nepotkáme. S funkcemi dvou proměnných pracujeme rádi právě proto, že si je obvykle umíme načrtnout.

3

Pokud máme tři a více proměnných, tak už si graf neznázorníme ani přibližně, čtyřrozměrný prostor je nad naši představivost. Proto se hledaly jiné způsoby, jak si chování funkcí více proměnných vizuálně přiblížit. Zde si ukážeme ty nejpoužívanější.

Budeme je ilustrovat na funkcích dvou proměnných, protože si tu situaci jsme schopni nakreslit, ale myšlenky budou platné i ve vyšších dimezních.

Obrázky grafů funkcí dvou proměnných lze získat také pomocí rozličných programových prostředků (Maple, Mathematica atd.), což je pohodlné a příjemné, nicméně je dobré věcem rozumět, ostatně občas tyto výpočetní prostředky selhávají a zůstává starý dobrý mozek a znalosti.

Hladiny konstantnosti

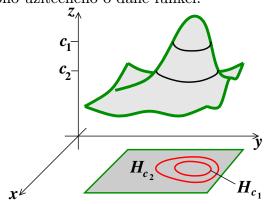
Toto je jedna z nejsilnějších metod pro vizualizaci chování funkce, ale tento pojem se dá využít i šířeji, například ke zkoumání implicitně zadaných útvarů.

Definice.

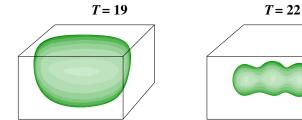
Nechť $f: D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$. Pro $c \in \mathbb{R}$ definujeme příslušnou hladinu konstantnosti $H_c = \{\vec{x} \in D(f); \ f(\vec{x}) = c\}$.

Z definice vidíme, že H_c leží v D(f), takže na znázornění hladin nám stačí jen n rozměrů, což je pokrok oproti grafům.

Jak to funguje? Nejprve graf řízneme na úrovni c (zeptáme se, kde všude jsou hodnoty funkce rovny c) a pak se podíváme, ze kterých hodnot proměnných se tam dostaneme, neboli si tento řez promítneme do definičního oboru. Pokud si graf představíme jako reliéf krajiny, tak hladina konstantnosti jsou místa daná pomocí zeměpisných souřadnic (tedy v rovině mapy) taková, že je v nich nadmořská výška přesně c. Jinými slovy, hladiny konstantnosti jsou analogické vrstevnicím na mapě. Zkušený turista dokáže z vrstevnic odhadnout tvarování krajiny, stejně tak lze z hladin konstantnosti odhadnout mnoho užitečného o dané funkci.



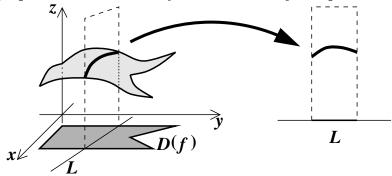
Je to dobrý způsob, jak analyzovat funkce tří proměnných. Pokud máme například funkci T(x,y,z) popisující teplotu v různých místech jisté místnosti, pak pro rozličné volby teploty c dostáváme příslušné hladiny konstantnosti H_c jako jakési "obláčky" ukazující, kde v místnosti je zrovna dotyčná teplota. Tyto obláčky jsou třírozměrné, je tedy možné si je pomocí perspektivy znázornit ve 2D (na papíře). Při porovnání těchto nákresů pro různé hodnoty c získáme velice dobrou představu o chování teploty, je také možné tyto nákresy spojit do animace a podobně. Obrázek níže ukazuje, jak by také mohly vypadat dva teplotní "snímky" učebny v únoru, kdy ji vytápějí tři studenti potící se u zkoušek.



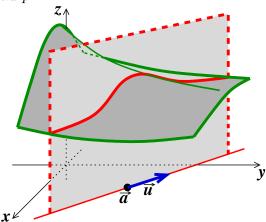
Řezy

Řezy jsou jednoduchým, ale velice užitečným nástrojem pro zkoumání vícerozměrných situací. Aby byly opravdu jednoduché, musíme jim dobře rozumět a vnímat vztahy mezi vzorci a geometrickou představou. Základní myšlenkou je, že si z mnoharozměrné situace vyřízneme pomocí roviny dvourozměrný "průřez", který již umíme (doufejme) prozkoumat pomocí nástrojů pro zkoumání (grafu) funkce jedné proměnné.

Myšlenka je jednoduchá. Máme graf funkce n proměnných, tedy útvar v n+1 dimenzích, kdy se proměnné berou ze světa \mathbb{R}^n , který si v obrázku symbolicky představíme jako vodorovný útvar. Když si vybereme v definičním oboru přímku a omezíme se čistě na ni, pak vlastně dostáváme plochý útvar, který po přenesení do \mathbb{R}^2 lze vystihnout funkcí jedné proměnné.



Formálně to obvykle děláme takto. Vybereme si nějaký počáteční bod \vec{a} a uvažujeme přímku p v \mathbb{R}^n procházející tímto bodem ve směru \vec{u} . Body této přímky jsou dány (pokud se pohybujeme rovnoměrně) parametrickou rovnicí $\vec{x} = \vec{a} + t\vec{u}$, můžeme si představit, že je to popis cesty, kterou vykonáváme v definičním oboru D(f). V čase t jsme v jistém bodě z D(f) a tomu odpovídá hodnota funkce $f(\vec{a} + t\vec{u})$. Vzniklo zobrazení $t \mapsto f(\vec{a} + t\vec{u})$, které popisuje tvar grafu, který nad sebou "vidíme" při pohybu po přímce p. Přímka p tedy určila řez, který získáme jako průnik grafu f a "svislé" roviny vztyčené nad p.



Tento řez má tvar, jenž zjevně souvisí s funkcí $\varphi(t)=f(\vec{a}+t\vec{u})$, která má jen jednu proměnnou a umíme ji hravě vyšetřit. Bohužel tato souvislost není tak jednoduchá, jak bychom doufali, tedy že když si nakreslíme graf funkce φ , tak rovnou uvidíme tvar řezu. Problém je v měřítku, na našem dvourozměrném obrázku řezu bude značka pro "1" (pozice v čase t=1) v místě $\vec{a}+\vec{u}$, což ovšem nemusí být ve vzdálenosti 1 od bodu \vec{a} v původním mnoharozměrném obrázku grafu f. Míra

zkreslení evidentně závisí na velikosti směrového vektoru \vec{u} , což by nemělo překvapit, čím rychleji jedeme, tím více se nám subjektivně zkreslí krajina kolem ("zmáčkne se", míjíme věci rychlejí).

Řešením je omezit se pouze na směrové vektory velikosti 1, pak nám bude geometrická informace souhlasit. Budeme se o to snažit, ale ne vždy je to možné, například v aplikacích (fyzika a spol.), je tedy dobré umět si poradit i s obecným případem. Narazíme na to později u derivací.

Příklad: Uvažujme funkci $f(x,y)=\frac{3x^2+10x+2y^2}{4x^2+y^2}$. Podíváme se, jak vypadá řez jejím grafem svislou rovinou nad přímkou (x,y)=(2,3)+t(-1,1). Interpretace: Funkce f dává nadmořskou výšku, my stojíme na místě s GPS souřadnicemi (2,3) a vydáme se směrem (-1,1), přitom nás zajímá, jakým způsobem stoupáme a klesáme.

Předpis pro přímku nám dává $x=2-t,\,y=3+t,$ po dosazení do funkce f dostáváme pomocnou funkci

$$\varphi:\ t\mapsto f(2-t,3+t)=\frac{5t^2-10t+50}{2t^2-10t+25}=1+\frac{5}{t^2-2+5}=1+\frac{5}{(t-1)^2+4}.$$

Poslední tvar ukazuje, že graf funkce φ má tvar kopečku, jehož vrchol je v čase t=1.

To znamená, že když se podíváme na odpovídající řez grafem funkce f, tak tento řez má tvar kopečku, ale ne přesně stejně tvarovaného, protože $\|\vec{u}\| = \sqrt{2}$, není to jednotkový vektor. Ovšem mění se jen vodorovné měřítko obrázku, nikoliv tvar řezu. Můžeme tedy konstatovat, že řez má nějaký tvar kopečku, také platí, že jeho vrchol je v bodě odpovídajícím času t=1, tedy v bodě (1,4).

Situace je tedy hodně podobná tomu obrázku výše, kde sice graf funkce nesouhlasí, ale bod \vec{a} , vektor \vec{u} a přímka p jsou na správných místech a řez má víceméně správný tvar. Mimochodem, obrázek ukazuje, že se snadno může stát, že sice na řezu vidíme kopeček (tedy lokální maximum), ale funkce jako taková tam extrém nemá, řežeme skrz úbočí.

Poznamenejme, že pokud bychom chtěli dostat přesný tvar řezu, tak bychom museli uvažovat směrový vektor $\frac{(-1,1)}{\|(-1,1)\|} = \frac{1}{\sqrt{2}}(-1,1)$. Výpočty jsou pak obdobné, jen drobet méně příjemné.

Nejčastěji používáme přímky rovnoběžné s osami, protože to zjednoduší analytickou práci a výsledky se dobře interpretují. Například pokud jsme v \mathbb{R}^3 v bodě (x_0, y_0, z_0) a chceme se pohybovat ve směru osy y, pak ostatní proměnné zůstávají konstantní a popis pohybu se tedy velmi zjednoduší. Vzhledem k tomu, že preferujeme směrový vektor o velikosti 1, vychází velmi příjemný popis přímky (a pohybu) ve tvaru $t \mapsto (x_0, y_0 + t, z_0)$. Podobně pomocí $t \mapsto (x_0 + t, y_0, z_0)$ popíšeme pohyb ve směru osy x a tak dále.

Abychom mohli takto pracovat obecně, použijeme souřadnicové vektory $\vec{e}_1 = (1, 0, \dots, 0, 0)$, $\vec{e}_2 = (0, 1, \dots, 0, 0)$ až $\vec{e}_n = (0, 0, \dots, 0, 1)$, jde tedy o obvyklou kanonickou bázi v \mathbb{R}^n . Pohyb z bodu \vec{x}_0 ve směru i-té souřadnice pak lze vyjádřit parametrickým vzorcem $t \mapsto \vec{x}_0 + t\vec{e}_i$. Protože $\|\vec{e}_i\| = 1$, tvary se nedeformují, takže například vyšetřením grafu funkce $t \mapsto f(x_0 + t, y_0, z_0)$ získáme přesný tvar řezu grafem funkce f nad přímkou, která je rovnoběžná s osou x a prochází bodem (x_0, y_0, z_0) .

V situaci, kdy pracujeme se známým a rozumně malým počtem proměnných (což je skoro vždy kromě obecných tvrzení), se často vzorce dále zjednodušší tím, že se vzdáme jinak rozumné představy, že pro t=0 nás parametrický popis umístí do výchozího bodu. Místo toho používáme jako parametr přímo proměnnou vybraného směru, tedy prostě pracujeme se vzorci $x\mapsto (x,y_0,z_0),$ $y\mapsto (x_0,y,z_0)$ atd., přičemž do výchozího bodu se dostaneme volbou $x=x_0,$ $y=y_0$ atd. Je to přirozené a zároveň funkční, získáváme tak informaci o vlivu jednotlivých proměnných na chování funkce.

Příklad: Uvažujme funkci $f(x,y)=\frac{3x^2+10x+2y^2}{4x^2+y^2}$. Jaký je vliv jednotlivých proměnných

[1b] 6

okolo nějakého bodu (x_0, y_0) ? Odpověď nám dají dvě funkce jedné proměnné,

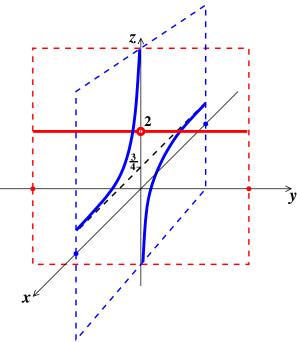
$$x \mapsto f(x, y_0) = \frac{3x^2 + 10x + 2y_0^2}{4x^2 + y_0^2}$$
$$y \mapsto f(x_0, y) = \frac{3x_0^2 + 10x_0 + 2y^2}{4x_0^2 + y^2}.$$

Pokud bychom například chtěli vědět, co se děje na okolí počátku (ve kterém funkce neexistuje, ale existuje všude jinde), pak bychom pracovali s výchozím bodem $x_0 = 0$, $y_0 = 0$ a tedy s funkcemi

$$x \mapsto f(x,0) = \frac{3x^2 + 10x}{4x^2} = \frac{3}{4} + \frac{5}{2x}$$

 $y \mapsto f(0,y) = \frac{2y^2}{y^2} = 2.$

První vzorec popisuje tvar grafu f nad osou x, druhý nad osou y. Naše zjištění lze zachytit následujícím náčrtem:



Úplně přesnou představu o grafu funkce f z toho ještě nezískáme, ale už tušíme, že to okolo počátku bude zajímavé.

 \triangle

Tento pohled se dá zobecnit tak, že si zafixujeme určitý počet proměnných a hýbeme s ostatními, čímž si snižujeme počet dimenzí, se kterými pracujeme, podle toho, co nás zrovna zajímá. Představme si na chvíli funkci T(x,y,z) tří proměnných, například popisující teplotu v různých místech v přednáškárně, a určitý bod $\vec{a} = (x_0, y_0, z_0)$.

Pokud si zafixujeme hodnoty $y = y_0$ a $z = z_0$, dostáváme jednorozměrný objekt neboli přímku, kdy se z bodu \vec{a} pohybujeme jen ve směru osy x a sledujeme teploty. Jde o jednorozměrnou situaci $x \mapsto T(x, y_0, z_0)$, kterou snadno prozkoumáme a znázorníme.

Pokud si zafixujeme pouze proměnnou $z=z_0$, pak máme dva stupně volnosti, pohybujeme se tedy na (vodorovné) rovině procházející bodem \vec{a} , která je kolmá na osu z. Vzniká tedy funkce dvou proměnných $(x,y)\mapsto T(x,y,z_0)$, jejíž graf si pořád ještě dokážeme při troše štěstí přiblížit v perspektivě (třeba stínováním).

Dostáváme tak další nástroj ke zkoumání funkcí. Nejčastěji ovšem pracujeme právě s pohybem po přímce a dvourozměrnými řezy.

Standardní tvary

V případě funkce jedné proměnné často nahlížíme na graf jako na analytický objekt v rovině daný rovnicí y = f(x), mnohé z těchto objektů umíme poznat. Například graf funkce f(x) = 1 - 2x je útvar popsaný rovnicí y = 1 - 2x neboli 2x + y = 1 a my víme, že jde o přímku. Toto často výtečně funguje i v případě více proměnných.

U funkcí n proměnných dostáváme objekt v \mathbb{R}^{n+1} daný rovicí $y=f(\vec{x})$. I zde lze doufat, že takový útvar poznáme, například funkce f(x,y)=2x+y-5 vede na rovnici z=f(x,y) neboli 2x+y-z=5, což definuje rovinu v \mathbb{R}^3 .

Někdy si vzniklou rovnici musíme upravit, pak ale musíme dát pozor, že tím může vzniknout větší množina. Na to už jsme ostatně museli dávat pozor i u funkcí jedné proměnné. Víme například, jak vypadá tvar grafu funkce $f(x) = \sqrt{x}$. Pokud si rovnici $y = \sqrt{x}$ přepíšeme jako $x = y^2$, hned rozpoznáme parabolu, jen má prohozené role proměnných, tudíž to bude parabola, která jde směrem vpravo, nad a pod osou x. Graf funkce $f(x) = \sqrt{x}$ je ovšem jen horní polovina tohoto útvaru, při umocňování rovnice jsme zároveň přišli o jednu z podmínek omezujících y.

Taková změna množiny nastává, pokud při změně rovnice provedeme neekvivalentní úpravu, což je například ono umocnění. Zkušenost napoví, kdy je třeba dát si pozor.

Aby mělo rozpoznávání tvarů úspěch, musíme znát základní geometrické objekty. Jde zejména o ploché útvary (přímky, roviny) a útvary s kvadratickými rovnicemi, což je ve dvou rozměrech známá rodinka kuželoseček (parabola, kružnice a obecněji elipsa, hyperbola), ve více rozměrech se to dá zkombinovat bohatěji. Je samozřejmě jasné, že by člověk musel mít opravdu velké štěstí, aby se při výběru z nekonečně mnoha možných funkcí trefil zrovna do jedné z těch pár rovnic, které zná, ale ono to nastává častěji, než by se čistě na základě pravděpodobnosti dalo čekat, a když to vyjde, tak to velmi potěší.

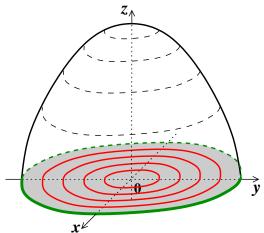
Příklad: Zde si ukážeme jednotlivé metody na funkci $f(x,y) = \sqrt{9 - x^2 - y^2}$.

Definiční obor: $D(f) = \{(x,y) \in \mathbb{R}^2; \ x^2 + y^2 \le 3^2\}$. Je to kruh o poloměru 3 se středem (0,0). Graf tedy bude nad ním, viz $f(x,y) \ge 0$. Můžeme si také všimnout, že $f(x,y) \le 3$ a do všech hodnot mezi 0 a 3 se dostat umíme, takže obor hodnot je $R(f) = \langle 0, 3 \rangle$.

Začneme hladinami konstantnosti. Jak jsme již poznamenali, hladiny konstantnosti budou zajímavé jen pro hodnoty mezi 0 a 3, ostatní jsou prázdné. Jak pro $c \in \langle 0, 3 \rangle$ hladina konstantnosti H_c vypadá?

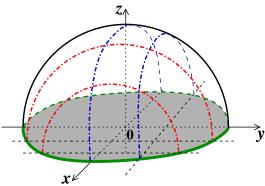
Je to množina daná vztahem f(x,y)=c neboli $x^2+y^2=9-c^2$, toto specifikuje kružnici o poloměru $\sqrt{9-c^2}$ ležící v definičním oboru. Pro větší hodnoty c (tedy větší hodnoty funkce) jsou kružnice menší, blíže ke středu, a naopak, pro hodnotu c=0 dostáváme jako hladinu konstantnosti přímo kružnici o poloměru 3, tedy hranici definičního oboru.

Závěr: Funkce je nulová na krajích definičního oboru, největší v jeho středu, tak vypadají kopečky. Na obrázku je čárkovaně několik úrovní hodnot funkce, v definičním oboru odpovídající hladiny konstantnosti.



Protože jsou hladiny coby kružnice rotačně symetrické vůči středu, můžeme usoudit, že dotyčný graf je symetrický vzhledem k rotaci okolo osy z. Lze to i potvrdit. Funkci si lze totiž přepsat jako $f(x,y) = \sqrt{9 - \sqrt{x^2 + y^2}}^2$, tedy její hodnota záleží jen na vzdálenosti $\sqrt{x^2 + y^2}$ bodu (x,y) od počátku, nikoliv na konkrétní pozici na odpovídající kružnici.

Nyní se podíváme na řezy. Zvolíme-li x=0, pak se ptáme, jak vypadá graf nad osou y. Dostáváme $f(0,y)=\sqrt{9-y^2}$, grafem je horní půlkružnice. Pro jiná pevně zvolená x=a jsou grafem menší půlkružnice $f(a,y)=\sqrt{(9-a^2)-y^2}$. Takto tedy vypadají řezy svislými rovinami rovnoběžnými s osou y.



Funguje to i symetricky, takže také svislé řezy grafem rovnoběžné s osou x jsou horní půlkružnice. Nejde tedy patrně o kopeček ledajaký, ale o kopeček sférického tvaru.

Zkusíme to potvrdit rozpoznáním křivky. Graf je dán rovnicí $z=\sqrt{9-(x^2+y^2)}$, což si hravě přepíšeme jako $x^2+y^2+z^2=3^2$ a dostáváme rovnici sféry. Graf funkce určitě nebude celá sféra (ta nabízí dvě různé hodnoty pro jeden bod (x,y)), ale nějaká její podmnožina. Protože D(f) je kruh v rovině xy o poloměru 3, bude graf zabírat celou rozlohu oné sféry, a ze vztahu $f(x,y)=\sqrt{**}\geq 0$ vyplývá, že jde jmenovitě o její horní polovinu, což ostatně naznačily již hladiny konstantnosti a řezy.

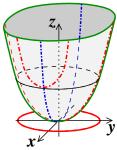
Závěr: Grafem je horní polosféra (dóm). \triangle

Příklad: Uvažujme $f(x,y)=x^2+y^2$. Evidentně $D(f)=\mathbb{R}^2$.

Hladiny konstantnosti: $x^2 + y^2 = c$, jsou to kružnice, jejichž poloměry se zvětšují se zvětšující se hodnotou funkce. Tedy čím dál jsme od středu, tím větší je funkce, to vypadá na jámu. Protože jsou všechny hladiny konstantnosti rotačně invariantní (po otočení okolo svislé osy vypadají pořád stejně), bude takový i graf.

Řezy: Zvolíme-li x=0, dostáváme $f(0,y)=y^2$, grafem je parabola. Pro jiná x jsou grafem paraboly posunuté nahoru tím více, čím dále jsme od počátku, $f(x_0,y)=x_0^2+y^2$, vrcholy těchto parabol jsou samy na parabole. Funguje to i symetricky.

Když k tomu přidáme pozorování o symetrii, závěr je jasný. Graf je rotační paraboloid.



 \triangle

Příklad: Uvažujme $f(x,y) = y^2 - x^2$. Zjevně $D(f) = \mathbb{R}^2$.

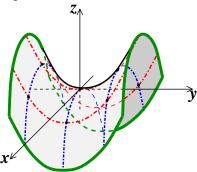
Hladiny konstantnosti: $y^2 - x^2 = c$ neboli hyperboly se zvětšujícím se poloměrem.

Řezy: Zvolíme-li x = 0 (řez nad osou y), dostáváme $f(0,y) = y^2$, grafem je standardní parabola. Pro jiná x jsou grafem paraboly posunuté dolů tím více, čím dále jsme od počátku, $f(x_0,y) = y^2 - x_0^2$. Vrcholy těchto parabol jsou na parabole otočené dolů $z = -x^2$.

Zvolíme-li y = 0 (řez nad osou x), dostáváme $f(x,0) = -x^2$, grafem je parabola otočená dolů. Pro jiná y jsou grafem paraboly posunuté nahoru tím více, čím dále jsme od počátku, $f(x,y_0) = -x^2 + y_0^2$, vrcholy těchto parabol jsou na parabole $z = y^2$.

Pomocné metody: Rotační symetrii nemáme. Rovnice $z=x^2-y^2$ nepopisuje nějaký všeobecně známý útvar. Nezbývá než si o tvaru udělat představu z řezů.

Závěr: Je to zvláštní graf, který stojí za nakreslení.



Vlastně je to dost známý tvar zvaný sedlo.

 \triangle

Na závěr si dáme ještě jedno zobecnění užitečného pojmu.

Definice.

Nechť $f: D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$.

Řekneme, že f je **omezená**, jestliže existuje K > 0 takové, že $|f(\vec{x})| \leq K$ pro všechna $\vec{x} \in D(f)$.

Nechť $M \subseteq D(f)$. Řekneme, že f je **omezená na** M, jestliže existuje K > 0 takové, že $|f(\vec{x})| \le K$ pro všechna $\vec{x} \in M$.

Grafická interpretace je jednoduchá. Pro funkci dvou proměnných to znamená, že se její graf dá sevřít mezi dvě vodorovné roviny. Ve vyšších dimezních je to obdobné, jen si to moc neumíme představit.

2. Představujeme limitu a spojitost

Pojem limity má ve více dimezních prakticky identickou definici jako v jedné proměnné:

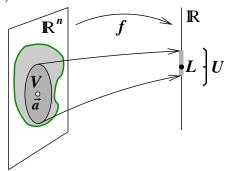
Definice.

Nechť $f: D(f) \to \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$.

Nechť \vec{a} je bod takový, že f je definováno na nějakém jeho prstencovém okolí, nechť $L \in \mathbb{R}^*$.

Řekneme, že L je limita f pro \vec{x} jdoucí k \vec{a} , značeno $\lim_{\vec{x}\to\vec{a}} (f(\vec{x})) = L$, jestliže pro každé okolí U hodnoty L existuje prstencové okolí V bodu \vec{a} takové, že $f[V] \subseteq U$.

Pro ilustraci tohoto pojmu je lepší si funkce představit způsobem, který se používá pro obecné zobrazení, tedy šipkový graf $D(f) \mapsto \mathbb{R}$.



Co se změnilo přechodem do více proměnných? Evidentní modifikací je, že okolí v definičním oboru jsou teď vícerozměrná. Pořád ale jde o okolí prstencová, tedy hodnota funkce v samotném bodě \vec{a} je irelevantní a dokonce ani nemusí existovat.

Přechod do více dimenzí se odrazí i v klasické epsilon-deltové formulaci podmínky:

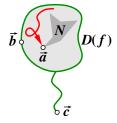
• Pro každé $\varepsilon > 0$ existuje $\delta > 0$ takové, že $|f(\vec{x}) - L| < \varepsilon$ kdykoliv $0 < ||\vec{x} - \vec{a}|| < \delta$.

V definičním oboru jsme museli použít vícerozměrnou vzdálenost mezi vektory, je to klasická Euklidovská vzdálenost.

Podstatný rozdíl oproti světu jedné proměnné je, že existují dva názory na to, jak má vypadat limita pro funkce více proměnných. Zde jsme zvolili variantu, která je přímým zobecněním jednorozměrného případu, vystačí ve většině případů a je jednoduchá. Její nevýhodou je, že nám neumožní spočítat některé zajímavé limity. Například funkce $f(x,y) = \sqrt{x} + y$ existuje na polorovině $\{(x,y); x \geq 0\}$ a jistě by bylo zajímavé spočítat si limitu v bodě (0,13), ale bohužel leží na hranici definičního oboru, takže funkce neexistuje na žádném jeho okolí, ať už prstencovém či plném. Definice nám tedy nedovolí se na limitu v tomto bodě zeptat. Některé lidi to trápí a tudíž někteří autoři volí obecnější a také komplikovanější definici, která by výpočet této limity umožnila.

Náš přístup zde vychází z představy, že je lepší mít jednoduchou základní definici limity a pro ty komplikovanější případy jsme si v jedné dimenzi vymysleli pojem jednostranné limity. Uděláme to i tady, ale ve více dimenzích se to podstatně zkomplikuje. Pokud máme bod \vec{a} ve vícerozměrném prostoru, pak nemá smysl mluvit o tom, že se k němu blížíme "zleva" či "zprava". My se totiž k němu můžeme blížit z nekonečně mnoha směrů, pohybujíce se po mnoha různých křivkách, dokonce můžeme uvažovat situaci, kdy se k němu blížíme v rámci výsečí, kuželů a po dalších zajímavých útvarech. Podívejme se na následující situaci.

Na obrázku vidíme definiční obor jisté funkce (asi nebude zcela obvyklá) a možná by nás zajímala limita v bodech \vec{b} a \vec{c} . Podle definice to nepůjde (nejsou obklopeny definičním oborem), ale smysl to dává. Například k bodu \vec{c} vede v rámci definičního oboru "ocásek", takže je možné se po něm pohybovat. Obrázek naznačuje, že se po něm umíme dostat libovolně blízko k \vec{c} , a celou tu dobu máme k dispozici hodnoty funkce f, takže má smysl se zeptat: Jdou ty hodnoty k nějakému číslu (k limitě)?



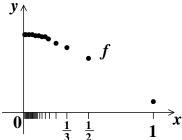
Obdobně se dokážeme dostat k bodu \vec{b} , dokonce má hned vedle sebe "plnou" množinu nabízející spoustu cest, jak se k \vec{b} dostat, ale to nestačí na to, aby šla použít naše definice limity, i zde budeme potřebovat ten nový pojem. Naopak u bodu \vec{a} definici limity aplikovat můžeme, ale i zde se můžeme rozhodnout, že bude z nějakého důvodu zajímavé zkusit jít k \vec{a} třeba po nakreslené křivce, a nebo možná v rámci naznačené množiny N. V takovém případě vlastně pracujeme s restrikcí funkce na nějakou podmnožinu definičního oboru (tu křivku či množinu N).

Když se hodláme přibližovat k nějakému bodu \vec{a} pomocí nějaké množiny, na které je f definované, tak musíme zajistit, že nám ta množina dovolí dojít libovolně blízko k \vec{a} , aby mělo smysl ptát se na limitu v \vec{a} . Matematicky tento požadavek zachycuje pojem hromadného bodu (viz základní topologické pojmy v apendixu).

Uvažujme zvláštní funkci f(x), která má definiční obor

$$\left\{\frac{1}{n}; \ n \in \mathbb{N}\right\} = \left\{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\right\}.$$

K bodu 0 se umíme pomocí této množiny dostat libovolně blízko, tudíž je to hromadný bod a má smysl se tam ptát na limitu.



Obdobně to funguje i ve více dimenzích.

Definice.

Nechť $f: D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$.

Nechť $\vec{a} \in \mathbb{R}^n$ a $N \subseteq D(f)$ jsou takové, že \vec{a} je hromadný bod N. Nechť $L \in \mathbb{R}^*$. Řekneme, že L je limita f pro \vec{x} jdoucí k \vec{a} vzhledem k N, značeno

$$\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in N}} (f(\vec{x})) = L,$$

jestliže pro každé okolí U hodnoty L existuje prstencové okolí V bodu \vec{a} takové, že $f[V\cap N]\subseteq U.$

Množina N může být nějaká oblast, může to být křivka, nebo prostě jen množina nějakých bodů. Vzhledem k omezením diskutovaným u definice limity se jako zajímavé jeví brát přímo N=D(f), protože pak nám tato definice umožňuje hledat limity na hranici definičního oboru. Vrátíme-li se k tomu úvodnímu příkladu, teď má smysl ptát se na limitu $\lim_{\substack{(x,y)\to(0,13)\\x>0}} (\sqrt{x}+y)$.

Pro funkce jedné proměnné platí, že pokud limita v nějakém bodě existuje, pak tam jednostranné limity musí dát stejnou odpověď. Totéž platí pro naše nové pojmy. Měli jsme také dvě tvrzení, která vedou v jistém smyslu v opačném směru. Obojí má vícedimenzionální verze.

Věta.

Nechť $f: D(f) \to \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$.

Nechť \vec{a} je bod takový, že f je definováno na nějakém jeho prstencovém okolí, nechť $L \in \mathbb{R}^*$.

(i)
$$\lim_{\vec{x}\to\vec{a}} (f(\vec{x})) = L$$
 právě tehdy, když $\lim_{\substack{\vec{x}\to\vec{a}\\\vec{x}\in N}} (f(\vec{x})) = L$ pro všechny $N\subseteq D(f)$

takové, že \vec{a} je hromadný bod N.

(ii) (Heineho věta) $\lim_{\vec{x}\to\vec{a}}(f(\vec{x}))=L$ právě tehdy, když $\lim_{k\to\infty}(f(\vec{x}(k)))=L$ pro všechny posloupnosti $\{\vec{x}(k)\}\subseteq D(f)-\{\vec{a}\}$ takové, že $\vec{x}(k)\to\vec{a}$.

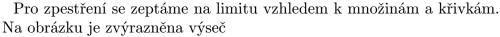
Tato věta je silným nástrojem při zkoumání limity.

Příklad: Uvažujme funkci

$$f(x,y) = \begin{cases} 1, & y = x^2, x > 0; \\ 0, & \text{jinde.} \end{cases}$$

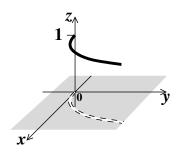
Její graf je vlastně planina na úrovni 0, nad kterou se tyčí strmý útes ve tvaru velice úzké lavice nad parabolickou stopou v definičním oboru.

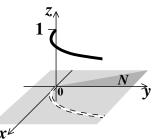
Jak to vypadá s limitou v počátku? Případná limita L by musela aproximovat hodnoty funkce f na okolí počátku s libovolnou přesností. Z obrázku je zjevné, že žádná taková hodnota existovat nemůže, protože ať už si vezmeme jakkoliv malé okolí počátku, tak tam vždycky najdeme v definičním oboru body s funkční hodnotou 0 a také body s funkční hodnotou 1.



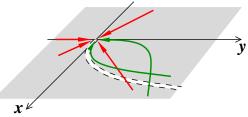
$$N = \{(x, y) \in \mathbb{R}^2; \ x < 0, y > -x\}.$$

Pro tuto množinu je $\vec{0}$ hromadným bodem, tudíž má smysl ptát se na limitu. Protože funkce f je na této množině vždy nulová, pak je i limita v $\vec{0}$ vzhledem k N rovna nule.





Teď vyzkoušíme množiny N ve tvaru křivky, jmenovitě se podíváme na přímky jdoucí do počátku. Takovéto "paprsky" lze kódovat pomocí směrnice, takže nás zajímají množiny bodů typu (x,kx) pro x>0, popřípadě pro x<0. To ale nezahrne dva speciální paprsky, jmenovitě když se k počátku blížíme po ose y, takže ještě přidáme množinu bodů typu (0,y) pro y>0, popřípadě y<0. Na obrázku několik takovýchto množin vidíme, šipkou naznačujeme, že se v rámci těchto množin blížíme do počátku. Jaké hodnoty funkce přitom potkáváme?



Je zjevné, že šipky ze tří kvadrantů, včetně těch jdoucích podél os x a y, se s tou parabolou v definičním oboru vůbec nepotkají. Takže i na těchto množinách je f vždy rovna nule a vychází limita nulová. Zbývají přímky z prvního kvadrantu. Ty se sice s dotyčnou parabolou protnou, ale jen jednou. To znamená, že u každého z těchto paprsků existuje nějaké okolí počátku, na kterém už se s parabolou nepotká, "vidí" jen hodnoty nula a takovou také určí limitu. Závěr tedy je, že všechny limity po přímkách vycházejí nulové.

Zajímavé je, že také skoro všechny limity po parabolických křivkách vyjdou nulové, protože i paraboly různé od té naší se s ní neprotnou buď vůbec, nebo jen v jednom bodě, to pokud bychom

připustili obecné paraboly typu y = ax(x - b) a x = ay(y - b), viz zelené křivky. Někdo by si mohl začít myslet, že funkce má limitu nula v počátku.

Je jen jediná "jednodušší" cesta, jak zjistit, že je s limitou v počátku problém, a to pustit se do něj po klíčové parabole $y=x^2,\,x>0$. Pak vidíme jen hodnoty funkce jedna a taková bude i limita vzhledem k této množině N. Dokázali jsme najít dvě množiny, vůči kterým dává limita v počátku dva různé výsledky, což dokazuje, že limita jako taková (dle první definice) neexistuje. Všimněte si nicméně, že pokud by nás nenapadla ta speciální cesta, tak bychom o problému s limitou nevěděli.

Ve skutečnosti je těch možností víc, postačila by libovolná křivka, která se s naší parabolou protne libovolně blízko k počátku, třeba $y = x^2 + \sin(\frac{1}{x})$, ale to by asi běžného vyšetřovatele limit nenapadlo.

 \triangle

Toto "chození po křivkách" je nejsilnější a nejčastěji používaný nástroj na zkoumání limit a také oblíbený typ důkazu, že zkoumaná limita neexistuje. Bohužel má svá omezení. Je zjevné, že není problém najít si namísto paraboly nějakou vhodnou opravdu nečekanou příšerku jako základ pro naši funkci, a dostaneme funkci dvou proměnných, která se v počátku tváří, jako že má limitu nula, ať už se blížíme po snad jakékoliv křivce přívětivého tvaru, ale ve skutečnosti limitu nemá. Je to tedy nástroj, který může snadno selhat, ale nic lepšího nemáme.

U funkcí jedné proměnné jsme výsledek často dostali prostým dosazením. Na to jsme potřebovali pojem spojitosti, který se také snadno zobecní.

Definice.

Nechť $f: D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$.

Nechť $\vec{a} \in D(f)$. Řekneme, že funkce f je spojitá v \vec{a} , jestliže

- \bullet \vec{a} je hromadný bod D(f) a $\lim_{\substack{\vec{x}\to \vec{a}\\ \vec{x}\in D(f)}} \left(f(\vec{x})\right) = f(\vec{a})$
- nebo \vec{a} je izolovaný bod D(f).

Druhá podmínka je opět něco, co ne všichni autoři dělají. My jsme zde vyhlásili automatickou spojitost, protože to usnadňuje formulaci některých tvrzení a má to smysl i logicky. Spojitost se vyjadřuje k souladu mezi hodnotou funkce v bodě a hodnotami na jeho bezprostředním okolí. Pokud funkce existuje v izolovaném bodě, pak na bezprostředním okolí neexistuje a tudíž se soulad nemůže pokazit. Někteří autoři to ovšem vidí tak, že tam tím pádem soulad není, a taky jim teorie fungují. Mi ten jejich přístup přijde o trošku lepší, ale pro tento text jsem zvolil variantu souhlasící s učebnicemi používanými na naší škole. Naštěstí v aplikacích funkce s izolovanými body definičního oboru v zásadě nepotkáváme, takže nás pak nemusí trápit, kterou teorii vlastně používáme.

Od definice spojitosti v bodě je už jen krůček ke spojitosti globální.

Definice.

Nechť $f: D(f) \to \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$.

Řekneme, že funkce f je spojitá na množině $M \subseteq D(f)$, jestliže je restrikce f na množinu M spojitá ve všech bodech M.

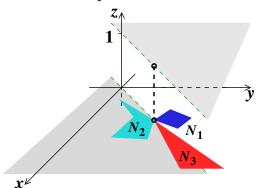
Řekneme, že funkce je spojitá, jestliže je spojitá ve všech bodech svého definičního oboru.

Pojem spojitosti ve více rozměrech odpovídá svými vlastnostmi své jednorozměrné inspiraci. Představujeme si je zase jako funkce, jejichž grafy (například ty plachty ve třírozměrném prostoru) nejsou "potrhané".

Příklad: Uvažujme funkci

$$f(x,y) = \begin{cases} 1, & y > x; \\ 0, & y \le x. \end{cases}$$

Definiční obor je zjevně \mathbb{R}^2 . Množina daná podmínkou y>x je "nekonečný trojúhelník" nad diagonálou v definičním oboru, tam je grafem funkce vodorovná "plachta" na úrovni 1. Na zbývající části definičního oboru je grafem vodorovná planina na základní úrovni 0.



Z obrázku se zdá zjevné, že tato funkce je spojitá na množinách $\{(x,y) \in \mathbb{R}^2; \ y > x\}$ a $\{(x,y) \in \mathbb{R}^2; \ y < x\}$ a nespojitá v bodech na přímce y = x.

Zastavme se ještě u bodu (1,1). Na obrázku jsme vyznačili tři množiny N, vůči kterým se zamyslíme nad limitou.

Je doufejme zjevné, že limita vzhledem k modré množině N_1 je 1, zatímco limita vzhledem k tyrkysové N_2 je 0. Protože tyto dvě relativní limity nesouhlasí, podle jedné z vět výše nemůže existovat ani limita v (1,1) jako taková, proto tam také nemáme spojitost. Obdobně usoudíme, že limita vzhledem k červené množině N_3 neexistuje.

Δ

Pro spojitost platí klíčové analytické věty. Za prvé, platí, že elementární funkce jsou spojité na svých definičních oborech, což zejména zahrnuje polynomy (i více proměnných). Dále pak platí, že když vytvoříme ze spojitých funkcí novou funkci pomocí algebraických operací, popřípadě skládání, tak je na svém definičním oboru spojitá.

Díky tomu získáme velkou zásobu spojitých funkcí. To je klíčové, protože tím také získáme snadný způsob, jak hledat celou řadu limit: Prostým dosazením do vzorce.

Příklad: Protože je exponenciála spojitá a polynomy také, je složená funkce $f(x,y) = e^{x^2y+y^2}$ spojitá na svém definičním oboru, což je evidentně \mathbb{R}^2 . Můžeme tedy snadno počítat limity ve všech bodech \mathbb{R}^2 , například

$$\lim_{(x,y)\to(2,-4)} \left(e^{x^2y+y^2}\right) = e^{-16+16} = 1.$$

 \triangle

Samozřejmě se nedá čekat, že by to takto krásně vyšlo vždy, konec konců, seznam neurčitých výrazů je platný i pro tyto limitní výpočty. A zde se věci silně komplikují, protože neexistuje vícerozměrná náhrada za l'Hospitalovo pravidlo, tedy nejmocnější nástroj na neurčité limity v jedné proměnné. Z praktického pohledu to znamená, že i když je pojem limity pro více proměnných teoreticky obdobný připadu jedné proměnné, tak při výpočtech to je úplně jiná (a horší) situace.

Hlavním problémem je dokázat, že nějaké L je opravdu limitou. Podíváme se na několik zajímavých slepých uliček, které nicméně nabídnou alespoň něco. Nabízí se třeba myšlenka, že kdybychom chodili k \vec{a} po řezech (křivkách), tak už jde o jednorozměrné situace, kde l'Hospitalovo pravidlo lze použít. Bohužel, existenci limity takto potvrdit neumíme, protože podle Heineho věty bychom museli prozkoumat úplně všechny možné křivky jdoucí do \vec{a} , což není prakticky možné. Jak už jsme viděli, pomocí limit po křivkách lze dokázat (při troše štěstí) jedině neexistenci limity.

Existuje ještě jeden zajímavý nápad, jak se dostat k práci s jednou proměnnou, jmenovitě tzv. "opakované limity". Tím se myslí proces, kdy se u vícerozměrné limity neblížíme s body \vec{x} k \vec{a} ve smyslu prostorovém, ale postupně po jednotlivých souřadnicích. Možných pořadí je více, jak ukazuje příklad dvou proměnných.

$$\lim_{(x,y)\to(a,b)} (f(x,y)) \implies \begin{cases} \lim_{y\to a} \left(\lim_{x\to a} (f(x,y))\right) \\ \lim_{x\to a} \left(\lim_{y\to b} (f(x,y))\right) \end{cases}$$

Podívejme se na tu první opakovanou limitu. Dá se ukázat, že pokud limita $\lim_{(x,y)\to(a,b)} (f(x,y)) \text{ (ta}$ standardní, v této souvislosti se o ní někdy mluví jako o "dvojné limitě") existuje a je rovna La pro všechna y_0 z nějakého okolí b existují také limity $\lim_{x\to a} (f(x,y_0))$, pak už nutně existuje také opakovaná limita $\lim_{y\to a} \left(\lim_{x\to a} \left(f(x,y)\right)\right)$ a je rovna L. Obdobné tvrzení platí i pro druhou opakovanou limitu. To znamená, že pokud ty dvě opakované

limity vyjdou různé, pak to dokazuje, že zkoumaná dvojná limita neexistuje.

Příklad: Vyšetříme limitu funkce
$$f(x,y) = \frac{x^2}{x^2 + y^2}$$
 v bodě $(0,0)$.

Tato funkce existuje na množině $\mathbb{R}^2 \setminus \{(0,0)\}$, hlavně tedy na prstencovém okolí bodu (0,0) a počítání limity má tedy smysl. Zkusíme opakované limity.

$$\lim_{y \to 0} \left(\lim_{x \to 0} \left(\frac{x^2}{x^2 + y^2} \right) \right) = \lim_{y \to 0} \left(\frac{0}{0 + y^2} \right) = \lim_{y \to 0} (0) = 0,$$

zatímco

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

Protože to vyšlo různě, $\lim_{(x,y)\to(0,0)} \left(\frac{x^2}{x^2+y^2}\right)$ neexistuje. \triangle

Bohužel, ta implikace spojující dvojné a opakované limity neplatí naopak. Pokud opakované limity vyjdou stejně, tak to nepomůže s dvojnou limitou.

Určování limit pro funkce více proměnných je evidentně adrenalinový sport, ne nepodobný integrování. Spolehlivý algoritmus neexistuje, ale jsou vcelku účinné postupy.

Vyšetřování limity

- Můžeme zkusit dosadit limitní bod \vec{a} do dané funkce f a pokud to vede k výsledku, pak máme i hodnotu limity. U zkouškových příkladů je pravděpodobnost blízká nule. V typickém případě dostaneme neurčitý výraz, což bohužel na rozdíl od funkcí jedné proměnné nevede na vcelku přímočarý postup.
- Pokud máme podezření, že limita neexistuje, můžeme to potvrdit tím, že dokážeme najít dva rozdílné výsledky, když se k \vec{a} blížíme po různých křivkách. Často nejprve zkoušíme jít po přímkách (paprscích), někdy dokonce jen po souřadnicových osách, protože je to jednodušší. Když to pořád vychází stejně, můžeme v zoufalství chodit i po komplikovanějších útvarech. Ve dvourozměrném případě je také zajímavou možností zkusit opakované limity. Pokud všechny pokusy vycházejí stejně, tak začneme mít podezření, že jsme možná našli limitu a bylo by třeba její existenci dokázat.
- Pokud máme podezření, že limita existuje, pak je potřeba to potvrdit. Bohužel, nějaká často fungující metoda neexistuje, je třeba přemýšlet nad různými přístupy. Někdy se podaří zkrátit algebraicky. Někdy pomůže srovnávací kritérium, které ve více dimenzích funguje obdobně. Někdy

zase pomůže substituce, zejména když se vydaří taková, která vyrobí jednorozměrnou limitu. Obvykle docela pomůže, pokud se podaří do situace nějak zahrnout vzdálenost mezi \vec{x} a \vec{a} , protože ta pak jde během limitění k nule. No a když všechno selže, je vždy možné ještě zkusit dokázat výsledek z definice.

Protože dokazování správnosti odpovědi na limitu je náročný úkol, tak se v přednáškách limity často spíše jen proletí a mnohdy vůbec nejsou zkoušeny.

Pojďme se blíže podívat na jednotlivé nástroje potvrzování existence limity. Co se týče srovnávacího kritéria, uvedeme si verzi vzhledem ke množině, ze které pak volbou $N = \mathbb{R}^n$ vyplyne i verze pro standardní limitu.

Věta. (o sevření)

Nechť $f: D(f) \to \mathbb{R}$, $g: D(g) \to \mathbb{R}$ a $h: D(h) \to \mathbb{R}$ jsou funkce,

kde $D(f), D(g), D(h) \subseteq \mathbb{R}^n$.

Nechť $\vec{a} \in \mathbb{R}^n$ a $N \subseteq D(g)$ jsou takové, že \vec{a} je hromadný bod N. Nechť $L \in \mathbb{R}^*$.

Předpokládejme, že $N \subseteq D(f) \cap D(h)$ a $f \leq g \leq h$ na N.

Jestliže
$$\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in N}} (f(\vec{x})) = \lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in N}} (h(\vec{x})) = L$$
, pak také $\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in N}} (g(\vec{x})) = L$.

Co se týče substituce, funkce $f: \mathbb{R}^n \to \mathbb{R}$ lze napsat jako složenou funkci vícero způsoby, takže je i vícero možných vět. My se zde podíváme na situaci, která se objevuje často, a to když máme f = h(g) pro nějaké $g: \mathbb{R}^n \to \mathbb{R}$ a $h: \mathbb{R} \to \mathbb{R}$. Pak můžeme počítat limitu jako

$$\lim_{\vec{x} \to \vec{a}} (h(g(\vec{x}))) = \begin{vmatrix} y = g(\vec{x}) \\ b = \lim_{\vec{x} \to \vec{a}} (g(\vec{x})) \end{vmatrix} = \lim_{y \to b} (h(y)),$$

ovšem vyžaduje to jisté předpoklady. Populární je podmínka, že h je spojitá v bodě b. Další možnost je ověřit, že hodnoty $g(\vec{x})$ nejsou rovny b poblíž \vec{a} . Viz tvrzení níže v sekci Fakta o limitách.

Jako ukázku postupů a situací se podíváme na několik limit.

Příklad: Vyšetříme limitu funkce $f(x,y) = \frac{x^2y^2}{x^4 + y^4}$ v bodě (0,0).

Funkce existuje na $\mathbb{R}^2 \setminus \{(0,0)\}$, tedy limita má smysl.

Pokusíme se o dosazení zadaného bodu do funkce a zjistíme, že není v definičním oboru, dále vidíme, že dosazení vedlo na oblíbený neurčitý výraz $\frac{0}{0}$.

Nejprve zkusíme jít do zadaného bodu po souřadnicových osách. Po ose x se pohybujeme volbou y = 0 a $x \to 0$, jde tedy o limitu pro $(x, 0) \to (0, 0)$.

$$\lim_{(x,0)\to(0,0)} \left(\frac{x^2y^2}{x^4+y^4}\right) = \lim_{x\to 0} \left(\frac{0}{x^4+0}\right) = \lim_{x\to 0} (0) = 0.$$

Obdobně po ose y:

$$\lim_{(0,y)\to(0,0)} \left(\frac{x^2y^2}{x^4+y^4}\right) = \lim_{y\to 0} \left(\frac{0}{0+y^4}\right) = \lim_{y\to 0} (0) = 0.$$

Mohla by být limita rovna nule? Zkusíme dojít do počátku po přímkách. Typická přímka skrz počátek je dána rovnicí y = kx, takže nás zajímá limitní proces pro $(x, kx) \to (0, 0)$.

$$\lim_{(x,kx)\to(0,0)} \left(\frac{x^2y^2}{x^4+y^4}\right) = \lim_{x\to 0} \left(\frac{k^2x^4}{x^4+k^4x^4}\right) = \lim_{x\to 0} \left(\frac{k^2}{1+k^4}\right) = \frac{k^2}{1+k^4}.$$

Vidíme, že výsledek záleží na k, tedy na tom, z které strany se po přímce blížíme k počátku. Závěr: Limita f v (0,0) neexistuje.

Mimo jiné z toho vyplývá, že funkce definovaná vzorcem

$$f(x) = \begin{cases} \frac{x^2 y^2}{x^4 + y^4}, & (x, y) \neq (0, 0); \\ 0, & (x, y) = (0, 0) \end{cases}$$

není spojitá v počátku.

Pro zajímavost si ještě vyzkoušíme opakované limity.

$$\lim_{y \to 0} \left(\lim_{x \to 0} \left(\frac{x^2 y^2}{x^4 + y^4} \right) \right) = \lim_{y \to 0} \left(\frac{0}{0 + y^4} \right) = \lim_{y \to 0} (0) = 0,$$

také

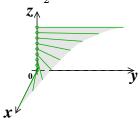
$$\lim_{x \to 0} \left(\lim_{y \to 0} \left(\frac{x^2 y^2}{x^4 + y^4} \right) \right) = \lim_{x \to 0} \left(\frac{0}{x^4 + 0} \right) = \lim_{x \to 0} (0) = 0.$$

Takže z nich bychom se o problémech s limitou nedozvěděli.

 \triangle

Jak si máme představit situaci, kdy výsledek limity záleží na směru, ze kterého se blížíme? Asi nejjednodušší je podívat se na funkci $f(x,y)=\frac{y}{x}$ na kvadrantu $D=\{(x,y)\in\mathbb{R}^2;\ x,y>0\}$. Podíváme-li se na vzorec, vidíme, že hodnota funkce záleží čistě na tom, jaká je směrnice přímky, na které se bod (x,y) nachází. Pro všechny body ležící na přímce y=kx je tedy f(x,y)=k. Jak vypadá graf? Tak trochu jako kousek točitého schodiště, kde se jednotlivé schody postupně otáčejí, ten dolní je ve směru osy x a horní ve směru osy y. Pak ty schody ještě zbrousíme, aby se nám lépe klouzalo dolů.

Protože směrnice jdou do nekonečna pro body velmi blízké ose y, nakreslím raději graf funkce $f(x,y) = \arctan\left(\frac{y}{x}\right)$. Pro body velmi blízké ose x se pak hodnoty f blíží nule, zatímco pro body velmi blízké ose y má funkce hodnoty blízké $\frac{\pi}{2}$.



Příklad: Vyšetříme limitu funkce $f(x,y) = \sqrt{2x^2 + y^4}$ v bodě (0,0).

Protože je (0,0) v definičním oboru této zjevně spojité funkce, hravě dostaneme výsledek dosazením, je to 0.

To byla nuda, zkusíme něco mírně komplikovanějšího, ale pořád se šťastným koncem:

$$\lim_{(x,y)\to(0,0)} \left(\frac{x^4-y^4}{x^2+y^2}\right) = \lim_{(x,y)\to(0,0)} \left(\frac{(x^2+y^2)(x^2-y^2)}{x^2+y^2}\right) = \lim_{(x,y)\to(0,0)} (x^2-y^2) = 0.$$

Zadaná funkce existuje na $\mathbb{R}^2 \setminus \{(0,0)\}$, tedy na jistém prstencovém okolí počátku a tudíž je možné tam o limitě uvažovat.

 \triangle

Příklad: Vyšetříme limitu funkce $f(x,y) = \frac{x^2 + y^4}{\sqrt{4x^2 + y^2}}$ v bodě (0,0).

Pokusíme se o dosazení zadaného bodu do funkce a zjistíme, že není v definičním oboru, dále vidíme, že dosazení vedlo na oblíbený neurčitý výraz $\frac{0}{0}$.

Pokus o přístup do (0,0) po souřadnicových osách:

$$\lim_{(x,0)\to(0,0)} \left(\frac{x^2+y^4}{\sqrt{4x^2+y^2}}\right) = \lim_{x\to 0} \left(\frac{x^2}{\sqrt{4x^2}}\right) = \lim_{x\to 0} \left(\frac{x^2}{2|x|}\right) = \lim_{x\to 0} \left(\frac{|x|}{2}\right) = 0.$$

Obdobně

$$\lim_{(y,0)\to(0,0)} \left(\frac{x^2+y^4}{\sqrt{4x^2+y^2}}\right) = \lim_{y\to 0} \left(\frac{y^4}{\sqrt{y^2}}\right) = \lim_{y\to 0} (|y|\cdot y^2) = 0.$$

Zkusíme to po přímkách:

$$\lim_{(x,kx)\to(0,0)}\Bigl(\frac{x^2+y^4}{\sqrt{4x^2+y^2}}\Bigr) = \lim_{x\to 0}\Bigl(\frac{x^2+k^4x^4}{\sqrt{4x^2+k^2x^2}}\Bigr) = \lim_{x\to 0}\Bigl(\frac{x^2(1+k^4x^2)}{|x|\sqrt{4+k^2}}\Bigr) = \lim_{x\to 0}\Bigl(\frac{|x|(1+k^4x^2)}{\sqrt{4+k^2}}\Bigr) = 0.$$

Co kdybychom zkusili jít po parabolách? Pro zjednodušení jen po těch snadných, $y = kx^2$.

$$\lim_{(x,kx^2)\to(0,0)} \left(\frac{x^2+y^4}{\sqrt{4x^2+y^2}}\right) = \lim_{x\to 0} \left(\frac{x^2+k^4x^8}{\sqrt{4x^2+k^2x^4}}\right) = \lim_{x\to 0} \left(\frac{|x|(1+k^4x^6)}{\sqrt{4+k^2x^2}}\right) = 0.$$

Což otočit jejich orentaci, tedy použít křivky $(ky^2, y) \rightarrow (0, 0)$?

$$\lim_{(ky^2,y)\to(0,0)} \left(\frac{x^2+y^4}{\sqrt{4x^2+y^2}}\right) = \lim_{y\to 0} \left(\frac{k^2y^4+y^4}{\sqrt{4k^2y^4+y^2}}\right) = \lim_{y\to 0} \left(\frac{y^4(1+k^2)}{|y|\sqrt{4k^2y^2+1}}\right)$$
$$= \lim_{y\to 0} \left(\frac{|y|y^2(1+k^2)}{\sqrt{4k^2y^2+1}}\right) = 0.$$

Koho to baví, zkusí si ještě dojít do počátku po sinusoidách $(x,k\sin(x))$ či dalších křivkách, ale možná je na čase smířit se s tím, že ta limita by opravdu mohla být nula, a začít přemýšlet nad tím, jak by se to dokázalo. Nezdá se, že by šlo nějak algebraicky zkrátit výrazy, které vytvářejí ty neurčité nuly. Substituce se také nenabízí, protože výrazy v čitateli a jmenovateli nemají zpolečný základ. Co zbývá? Srovnání. Dobré by bylo najít nějaké, ktere pracuje se vzdáleností bodu od počátku, protože pro $(x,y) \to (0,0)$ nutně platí $\sqrt{x^2 + y^2} \to 0$, možná by to šlo využít.

Jaké odhady potřebujeme? Abychom vynutili $f \to 0$, stačí vynutit $|f| \to 0$, tedy použít srovnávací větu ve verzi s absolutní hodnotou. Hledáme tedy **horní** odhad pro funkci |f|, ovšem vzhledem k tomu, že evidentně $f \geq 0$, tak nám vlastně stačí najít vhodný horní odhad pro samotné f. To znamená, že bychom potřebovali dolní odhad pro jmenovatel, pokud možno aby nějak zahrnul vzdálenost od počátku. Vyzkoušíme:

$$\sqrt{4x^2 + y^2} \ge \sqrt{x^2 + y^2} \implies \frac{1}{\sqrt{4x^2 + y^2}} \le \frac{1}{\sqrt{x^2 + y^2}}$$

To vlastně bylo lehké. S čitatelem to až tak snadné nebude, zlobí tam to y^4 . Klíčem je uvědomit si, že když $(x,y) \to (0,0)$, tak jsou x,y malé, zejména pak nakonec musí platit |y| < 1. Pak ovšem $y^4 < y^2$ a odhadujeme

$$x^2 + y^4 < x^2 + y^2 = \sqrt{x^2 + y^2}^2$$
.

Dáme to dohromady:

$$|f(x,y)| = \frac{x^2 + y^4}{\sqrt{4x^2 + y^2}} \le \frac{\sqrt{x^2 + y^2}^2}{\sqrt{x^2 + y^2}} = \sqrt{x^2 + y^2} = h(x,y).$$

Protože $h(x,y)=\sqrt{x^2+y^2}\to 0$ v počátku, podle srovnávacího kritéria platí, že $f(x,y)\to 0$ v (0,0).

 \triangle

Jen mírně pozměněným postupem bychom ukázali, že

$$\frac{\sqrt{4x^2 + y^2}}{x^2 + y^4} \ge \frac{1}{\sqrt{x^2 + y^2}} \to \infty \qquad \text{pro } (x, y) \to (0, 0).$$

Je evidentní, že ty odhady nám vyšly jen díky pečlivě připravenému příkladu, jinak bychom se zapotili.

[2]

Příklad: Vyšetříme limitu funkce $f(x,y) = \frac{x^2y}{x^4 + y^2}$ v bodě (0,0).

I tato otázka má smysl a vede na limitu typu $\frac{0}{0}$. Rovnou zkusíme jít po přímkách.

$$\lim_{(x,kx)\to(0,0)} \left(\frac{x^2y}{x^4+y^2}\right) = \lim_{x\to 0} \left(\frac{kx^3}{x^4+k^2x^2}\right) = \lim_{x\to 0} \left(\frac{kx}{x^2+k^2}\right) = \frac{0}{0+k} = \frac{0}{k} = 0.$$

Opravdu? Co když k=0? Pak ten výpočet v posledním kroku ztroskotá. My ale při zkoumání případu k=0 tu informaci máme už od začátku, takže ji můžeme využít ještě dříve, než na ten osudový poslední krok dojde.

$$\lim_{(x,0 \cdot x) \to (0,0)} \left(\frac{x^2 y}{x^4 + y^2} \right) = \lim_{(x,0) \to (0,0)} \left(\frac{x^2 y}{x^4 + y^2} \right) = \lim_{x \to 0} \left(\frac{0}{x^4 + 0} \right) = \lim_{x \to 0} (0) = 0.$$

Takže to ta nula byla, ale museli jsme to správně odůvodnit.

Že by limita vycházela nulová? Pro jistotu se ještě podíváme na paraboly, třeba tu nejjednodušší $y = x^2$.

$$\lim_{(x,x^2)\to(0,0)} \left(\frac{x^2y}{x^4+y^2}\right) = \lim_{x\to 0} \left(\frac{x^4}{x^4+x^4}\right) = \lim_{x\to 0} \left(\frac{1}{2}\right) = \frac{1}{2}.$$

Vyšlo to jinak, takže skvělé, máme hotovo.

Závěr: f nemá limitu v (0,0).

 \triangle

Příklad: Vyšetříme limitu funkce $f(x,y) = \frac{e^{x^2+y^2}-1}{x^2+y^2}$ v bodě (0,0).

Překvapivě, i tato otázka má smysl a vede na limitu typu $\frac{0}{0}$. Teď tam nemáme jen polynomy, tak začneme opatrně po osách.

$$\lim_{(x,0)\to(0,0)} \left(\frac{e^{x^2+y^2}-1}{x^2+y^2}\right) = \lim_{x\to 0} \left(\frac{e^{x^2}-1}{x^2}\right) \frac{\frac{0}{0}}{\frac{1}{1}H} \lim_{x\to 0} \left(\frac{2x\,e^{x^2}}{2x}\right) = \lim_{x\to 0} (e^{x^2}) = 1.$$

Tu jednorozměrnou limitu šlo také počítat přes Taylorův rozvoj, substituci $t=x^2$ a někteří si tento typ dokonce pamatují.

Díky symetrii by limita vzhledem k ose y dopadla stejně. Jak dopadnou přímky?

$$\lim_{(x,kx)\to(0,0)} \left(\frac{e^{x^2+y^2}-1}{x^2+y^2}\right) = \lim_{x\to 0} \left(\frac{e^{x^2(1+k^2)}-1}{x^2(1+k^2)}\right) = 1$$

obdobným výpočtem.

Kdo chce, vyzkouší si třeba pohyb po parabolách, ale asi je načase zamyslet se, jestli zkoumaná limita náhodou není rovna 1. Jak bychom to dokázali? Jedna možnost je použít Taylorův rozvoj, ale není jasné, zda jsme na to pro více proměnných připraveni. Co se týče srovnání, tak se tam nějaké užitečné nenabízí. Zato si všimneme, že proměnné se vyskytují vždy ve stejném výrazu, což napovídá substituci:

$$\lim_{(x,y)\to(0,0)} \left(\frac{e^{x^2+y^2}-1}{x^2+y^2}\right) = \left| \begin{array}{c} r = x^2+y^2 \\ (x,y)\to(0,0) \implies r\to 0 \end{array} \right| = \lim_{r\to 0} \left(\frac{e^r-1}{r}\right) \frac{\frac{0}{0}}{\Pr } \lim_{r\to 0} (e^r) = 1.$$

 \triangle

Fakta o limitě a spojitosti

Pro úplnost zde stručně připomeneme některá klíčová tvrzení.

(o limitě a operacích) Věta.

Nechť $f: D(f) \mapsto \mathbb{R}$ a $g: D(g) \mapsto \mathbb{R}$ jsou funkce, kde $D(f), D(g) \subseteq \mathbb{R}^n$.

Předpokládejme, že f, g jsou definovány na nějakém prstencovém okolí jistého bodu \vec{a} . Pak platí:

(i)
$$\lim_{\vec{x} \to \vec{a}} \left((f+g)(\vec{x}) \right) = \lim_{\vec{x} \to \vec{a}} \left(f(\vec{x}) \right) + \lim_{\vec{x} \to \vec{a}} \left(g(\vec{x}) \right),$$

(ii)
$$\lim_{\vec{x} \to \vec{q}} ((f - g)(\vec{x})) = \lim_{\vec{x} \to \vec{q}} (f(\vec{x})) - \lim_{\vec{x} \to \vec{q}} (g(\vec{x})),$$

(iii)
$$\lim_{\vec{x} \to \vec{a}} ((f \cdot g)(\vec{x})) = \lim_{\vec{x} \to \vec{a}} (f(\vec{x})) \cdot \lim_{\vec{x} \to \vec{a}} (g(\vec{x})),$$

$$(i) \lim_{\vec{x} \to \vec{a}} \left((f+g)(\vec{x}) \right) = \lim_{\vec{x} \to \vec{a}} \left(f(\vec{x}) \right) + \lim_{\vec{x} \to \vec{a}} \left(g(\vec{x}) \right),$$

$$(ii) \lim_{\vec{x} \to \vec{a}} \left((f-g)(\vec{x}) \right) = \lim_{\vec{x} \to \vec{a}} \left(f(\vec{x}) \right) - \lim_{\vec{x} \to \vec{a}} \left(g(\vec{x}) \right),$$

$$(iii) \lim_{\vec{x} \to \vec{a}} \left((f \cdot g)(\vec{x}) \right) = \lim_{\vec{x} \to \vec{a}} \left(f(\vec{x}) \right) \cdot \lim_{\vec{x} \to \vec{a}} \left(g(\vec{x}) \right),$$

$$(iv) \lim_{\vec{x} \to \vec{a}} \left(\frac{f}{g}(\vec{x}) \right) = \frac{\lim_{\vec{x} \to \vec{a}} \left(f(\vec{x}) \right)}{\lim_{\vec{x} \to \vec{a}} \left(g(\vec{x}) \right)},$$

$$(v) \lim_{\vec{x} \to \vec{a}} ((f^g)(\vec{x})) = \lim_{\vec{x} \to \vec{a}} (f(\vec{x}))^{\lim_{\vec{x} \to \vec{a}} (g(\vec{x}))},$$
pokud mají pravé strany smysl

pokud mají pravé strany smysl.

Věta. (o limitě složené funkce)

Nechť $g: D(g) \mapsto \mathbb{R}$ je funkce, kde $D(g) \subseteq \mathbb{R}^n$.

Předpokládejme, že g je definována na nějakém prstencovém okolí jistého bodu $\vec{a},$ že existuje $b=\lim_{\vec{x}\to\vec{a}} \bigl(g(\vec{x})\bigr)$ a že existuje prstencové okolí $P\subseteq D(g)$ bodu \vec{a} takové, že $g(\vec{x}) \neq b$ pro $\vec{x} \in P$.

Předpokládejme dále, že h je reálná funkce definovaná na nějakém prstencovém okolí b. Pak

$$\lim_{\vec{x}\to\vec{a}} ((h\circ g)(\vec{x})) = \lim_{t\to b} (h(t)).$$

Někdy je také praktické formulovat závěr jako $\lim_{\vec{x} \to \vec{a}} \left(h(g(\vec{x})) \right) = h \left[\lim_{\vec{x} \to \vec{a}} \left(g(\vec{x}) \right) \right]$, což platí, pokud má výraz na pravé straně smysl. Ve výpočtech se často stane, že je h spojitá v b, které není izolovaným bodem D(h), pak už se nemusí kontrolovat podmínka $g(\vec{x}) \neq b$.

Heineho věta a obecně myšlenka zkoumat limitu přechodem k podmnožinám platí také obecně pro situaci, kdy zkoumáme limitu vzhledem k jisté množině.

Věta.

Nechť $f: D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$.

Nechť $\vec{a} \in \mathbb{R}^n$ a $N \subseteq D(f)$ jsou takové, že \vec{a} je hromadný bod N. Nechť $L \in \mathbb{R}^*$.

(i) $\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in N}} (f(\vec{x})) = L$ právě tehdy, když $\lim_{\substack{\vec{x} \to \vec{a} \\ \vec{x} \in M}} (f(\vec{x})) = L$ pro všechny $M \subseteq N$ $\vec{x} \rightarrow \vec{a} \\ \vec{x} \in N$

takové, že \vec{a} je hromadný bod M.

(ii) $\lim_{\vec{x} \to \vec{a}} (f(\vec{x})) = L$ právě tehdy, když $\lim_{k \to \infty} (f(\vec{x}(k))) = L$ pro všechny posloup-

nosti $\{\vec{x}(k)\} \subseteq N - \{\vec{a}\}$ takové, že $\vec{x}(k) \to \vec{a}$.

Ve více dimenzích platí i věta o souvislosti konvergence a omezenosti.

Věta.

Nechť $f: D(f) \to \mathbb{R}$ je funkce, kde $D(f) \subset \mathbb{R}^n$. Nechť $\vec{a} \in \mathbb{R}^n$ a $N \subset D(f)$ jsou takové, že \vec{a} je hromadný bod N.

Jestliže limita f v \vec{a} vzhledem k N existuje a je konečná, tak existuje okolí Ubodu \vec{a} takové, že f je omezená na $U \cap N$.

Teď něco o spojitosti.

Fakt.

Funkce vytvořené pomocí elementárních funkcí, algebraických operací a skládání jsou spojité na svých definičních oborech.

Věta. (o extrémní hodnotě)

Nechť $f: D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$. Jestliže je M omezená uzavřená podmnožina D(f) a f je spojitá na M, pak f nabývá svého minima a maxima na M, tedy existují $\vec{x}_{\min}, \vec{x}_{\max} \in M$ takové, že $f(\vec{x}_{\min}) = \inf\{f(\vec{x}); \vec{x} \in M\}$ a $f(\vec{x}_{\max}) = \sup\{f(\vec{x}); \vec{x} \in M\}$.

Věta. (o mezihodnotě)

Nechť $f: D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$. Jestliže je f spojitá, pak pro každou souvislou $M \subseteq D(f)$ je i f[M] souvislá.

Koutek pro pokročilé V mnoha aplikacích se nám na spojitých funkcích líbí, že když jsme v nějakém bodě \vec{a} a o kousek se pohneme, tak se f také změní jen o kousek. Někdy ale potřebujeme vědět, jak malý ten kousek je, a mohou se objevit problémy, dokonce v jedné proměnné. Víme například, že funkce $\operatorname{arctg}(x)$, $\frac{1}{x}$ a x^2 jsou spojité. Teď si představme, že jsme v bodě a a posuneme se o $\Delta x = 0.1$.

U funkce $\operatorname{arctg}(x)$ je úplně jedno, kde je a, hodnoty funkce se prostě při tom malém posunu nemůžou změnit o víc než o 0.1. Toto ale neplatí o dalších dvou funkcí. Při pohledu na graf $y=\frac{1}{x}$ vidíme, že když bereme a blízko počátku, tak posun o Δx může znamenat velkou změnu hodnoty funkce, přičemž pro tuto změnu není horní mez. Čím blíže k 0 se a dostane, tím větší změnu funkčních hodnot vyvolá posun o 0.1, přičemž můžeme dosáhnout libovolně velké změny. U funkce $y=x^2$ je podobný problém u nekonečna.

Dá se nicméně ukázat, že například na intervalu (0,13) už se x^2 chová rozumně, tam posun o 0.1 nikdy nevyvolá změnu větší než 2.6. Podobně je $\frac{1}{x}$ pěkná třeba na intervalu $(1,\infty)$.

Funkce, jejíž změny lze mít pod kontrolou, si zaslouží speciální jméno.

Definice.

Nechť $f \colon D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$. Nechť $M \subseteq D(f)$. Řekneme, že f je **stejnoměrně spojitá** na M, jestliže pro každé $\varepsilon > 0$ existuje $\delta > 0$ takové, že $||f(\vec{x}) - f(\vec{y})|| < \varepsilon$ pro všechna $\vec{x}, \vec{y} \in M$ splňující $||\vec{x} - \vec{y}|| < \delta$.

Některé funkce jsou stejnoměrně spojité na celém definičním oboru, třeba onen výše zmíněný $\operatorname{arctg}(x)$, také třeba $\sin(x)$ či $\frac{1}{x^2+1}$. Nemá to ale nic společného s omezeností, například $\sin(x^2)$ je omezený, ale není to stejnoměrně spojitá funkce na \mathbb{R} . Naopak funkce $\sqrt{x}\sin(\sqrt{x})$ nebo třeba 13x nejsou na $(0,\infty)$ omezené, ale stejnoměrně spojité tam jsou.

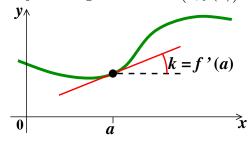
Velice mnoho užitečných funkcí není stejnoměrně spojitých na \mathbb{R} (třeba e^x či mocniny), ale podle následující věty se dokážeme ke stejnoměrné spojitosti dostat přechodem na podmnožiny, což často stačí.

Věta.

Nechť $f: D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$. Jestliže je M omezená a uzavřená podmnožina D(f) a f je spojitá na M, pak je f stejnoměrně spojitá na M.

3. Představujeme derivace

Nejprve si připomeňme hlavní významy derivace u funkce jedné proměnné. V aplikacích významně používáme fakt, že derivace f'(a) udává rychlost změny hodnoty funkce f, pokud procházíme proměnnou přes zvolený bod a. Geometrická interpretace říká, že f'(a) udává směrnici tečny ke grafu f v bodě a, což charakterizuje sklon grafu v bodě (a, f(a)).

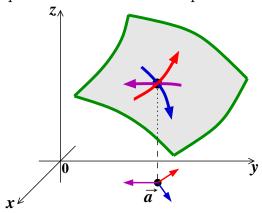


Třetí užitečný pohled je, že derivace umožňuje aproximovat hodnoty funkce na okolí bodu a pomocí tečny, vzorec je (ve dvou podobách)

$$f(x) \approx f(a) + f'(a) \cdot (x - a), \qquad f(a + h) \approx f(a) + f'(a)h.$$

3a. Derivace ve vyšší dimenzi

Je zjevné, že pokud budeme chtít tyto myšlenky aplikovat na funkce více proměnných, tak bude potřeba je výrazně modifikovat. Vidíme to na obrázku funkce dvou proměnných: Pokud zvolíme bod $\vec{a} \in D(f)$, pak rychlost klesání či růstu grafu při průchodu přes \vec{a} záleží na tom, kudy půjdeme, a na rozdíl od situace s jednou proměnnou se teď nabízí spousta směrů a tedy i odpovědí.



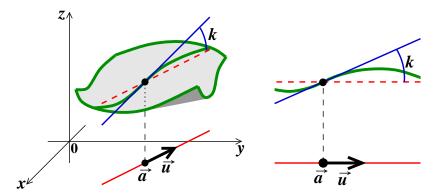
Touto úvahou jsme se nicméně dostali k jednomu z klíčových pojmů. Pokud nám někdo řekne, kterým směrem se z \vec{a} vydat, pak už má otázka na rychlost růstu grafu smysl, stejně tak pak má smysl i přiložit ke grafu přímku coby tečnu a tou pak hodnoty funkce aproximovat (v daném směru).

Příklad: Uvažujme funkci $f(x,y) = x^2 + y^2$, jsme v bodě $\vec{a} = (1,2)$. Co se stane, když se vydáme v určitém směru $\vec{u} = (h,k)$?

Pohybujeme se na přímce dané parametrickou rovnicí (x, y) = (1, 2) + t(h, k), přitom potkáváme funkční hodnoty

$$\varphi(t) = f(1+th, 2+tk) = (1+th)^2 + (2+tk)^2 = (h^2+k^2)t^2 + (2h+4k)t + 5.$$

Jde samozřejmě o situaci s řezem, kterou jsme již viděli v úvodní kapitole. Graf funkce f jsme proťali svislou rovinou nad přímkou danou vzorcem $t \mapsto \vec{a} + t\vec{u}$ a očekáváme, že dostaneme snadno zvládnutelnou dvourozměrnou situaci popsanou pomocí jedné proměnné.



Přesně takový popis nám poskytuje naše funkce φ . Můžeme ji tedy zderivovat, v bodě \vec{a} jsme v čase t=0 a hodnoty φ se v té chvíli mění rychlostí

$$\varphi'(0) = 2t(h^2 + k^2) + (2h + 4k)\big|_{t=0} = 2h + 4k.$$

 \triangle

Jak tento výsledek můžeme interpretovat? Pokud se například vydáme z bodu \vec{a} ve směru $\vec{v} = (-1,1)$, tak nám výsledek derivování funkce $\varphi(t) = 2t^2 + 2t + 5$ v čase t=0 dává číslo 2. Jaký je jeho význam? Je to rychlost, s jakou se pro nás (subjektivně) mění hodnota funkce při průchodu bodem \vec{a} . Tato pozorovaná rychlost změny ovšem závisí na tom, jak rychle se po přímce pohybujeme, tedy na velikosti směrového vektoru \vec{v} . Ono číslo 2 tedy nenese geometrickou informaci, kterou bychom rádi měli pro hrátky s tečnami či aproximacemi.

Aby se pozorované a geometrické vlastnosti shodovaly, je potřeba použít směrový vektor velikosti 1. V našem případě bychom použili vektor

$$\vec{u} = \frac{(-1,1)}{\|(-1,1)\|} = \frac{1}{\sqrt{2}}(-1,1).$$

Podle výpočtu výše dostáváme, že v tomto směru se graf funkce mění rychlostí $\frac{1}{\sqrt{2}}(-2+4) = \sqrt{2}$. Toto už je také geometricky významná informace, například to opravdu souhlasí se směrnicí

"tečny ve směru", tedy tečny, kterou bychom v bodě \vec{a} sestrojili na skutečném řezu grafem. Pokud bychom ze dřeva vyřezali dřevěný model grafu, tak bychom mohli měřením potvrdit, že v daném směru graf svírá s vodorovnou rovinou úhel $\arctan(\sqrt{2})$.

Budeme pochopitelně v matematice preferovat směrové vektory velikosti 1. Teorii ale vybudujeme obecně, protože směrové vektory libovolných délek mají své místo v aplikacích, například ve fyzice (vektor rychlosti).

Poznámka: Pomocí tečny bychom dostali i možnost aproximace hodnot funkce ve směru \vec{u} . Pro funkci φ platí $\varphi(t) \approx \varphi(0) + \varphi'(0)t$. Pokud toto použijeme s normalizovaným vektorem \vec{u} a přejdeme zpět k funkci f, dostáváme vzorec

$$f((1,2) + \frac{1}{\sqrt{2}}(-1,1)t) \approx 5 + \sqrt{2}t$$

neboli

$$f\left(1 - \frac{1}{\sqrt{2}}t, 2 + \frac{1}{\sqrt{2}}t\right) \approx 5 + \sqrt{2}t.$$

Substitucí $s=\frac{1}{\sqrt{2}}t$ dostáváme rovnocenný, ale mnohem příjemnější vzorec

$$f(1-s, 2+s) \approx 5+2s$$
.

Zajímavé je, že přesně totéž bychom totiž dostali přímo, pokud bychom použili φ s původním směrovým vektorem \vec{v} . Takže pokud nás nezajímá přesný tvar grafu, ale jen hledáme aproximaci, tak vlastně nemusíme normalizovat. Každopádně závěr je, že pokud se chceme z bodu \vec{a} pohybovat jen v dotyčném směru, pak lze hodnoty funkce aproximovat (pro malá t) vzorcem

$$f(1-t, 2+t) \approx f(1, 2) + 2t$$
.

 \triangle

Rychlost změny funkce f v bodě \vec{a} jsme našli jako $\varphi'(0)$. Co jsme to vlastně počítali?

$$\varphi'(0) = \lim_{t \to 0} \left(\frac{\varphi(t) - \varphi(0)}{t} \right) = \lim_{t \to 0} \left(\frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t} \right).$$

K informaci se tedy umíme dostat přímo, bez pomocné funkce φ . Ta poslední limita je doufejme čtenáři povědomá a nepřekvapí nás v následující definici.

Definice.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Nechť \vec{u} je vektor

Řekneme, že funkce f je **diferencovatelná** v bodě \vec{a} ve směru \vec{u} , jestliže limita $\lim_{t\to 0} \left(\frac{f(\vec{a}+t\vec{u})-f(\vec{a})}{t}\right) \text{ existuje konečná.}$ Pak definujeme **(směrovou) derivaci** f v bodě \vec{a} ve směru \vec{u} jako

$$D_{\vec{u}}f(\vec{a}) = \lim_{t \to 0} \left(\frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t} \right).$$

Mimochodem, při udávání směru evidentně potřebujeme nenulový vektor \vec{u} . V definici jsme toto omezení neuvedli, ale nevadí to. Pro $\vec{u} = \vec{0}$ vychází z definice $D_{\vec{0}}f = 0$, což v praxi nevyužijeme, ale nepokazí to různé věty.

Jeden z výsledků onoho příkladu výše je nyní možné zachytit vzorcem $D_{(-1,1)}f(1,2)=2$. Tento příklad také naznačil, že dost možná nezáleží na tom, ve které fázi výpočtu provedeme normalizaci směrového vektoru. Následující tvrzení to potvrzuje.

Fakt.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Nechť \vec{u} je vektor z \mathbb{R}^n a předpokládejme, že f je diferencovatelná v bodě \vec{a} ve směru \vec{u} . Pak je také f pro libovolné $\lambda \in \mathbb{R}$ diferencovatelná ve směru $\lambda \vec{u}$ a platí

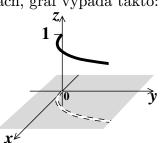
$$D_{\lambda \vec{n}} f(\vec{a}) = \lambda D_{\vec{n}} f(\vec{a}).$$

Dá se ukázat mnohem více. Když si pevně zvolíme nějaký směr \vec{u} , tak se příslušná směrová derivace chová jako běžná derivace funkce jedné proměnné, zejména platí běžná výpočetní pravidla typu součinového vzorce. Dokonce platí i věta o střední hodnotě, viz kapitola Více o derivaci.

Ovšem zkoumat funkci jen v jednom směru asi nedá dostatečné informace o jejím chování. Selský rozum napovídá, že pokud chceme rozumět chování funkce v nějakém bodě, pak bychom se asi měli podívat na derivace ve všech směrech. Kupodivu se ukáže, že to ještě nestačí. Ukážeme si jeden zajímavý zádrhel.

Příklad: Uvažujme funkci $f(x,y) = \begin{cases} 1, & y = x^2, x > 0; \\ 0, & \text{jinde.} \end{cases}$

Již jsme ji potkali v kapitole o limitách, graf vypadá takto:



Zjistili jsme, že v počátku (0,0) neexistuje limita, teď se tam podíváme na směrové derivace. Vezměme libovolný vektor $\vec{u} \neq \vec{0}$. Jak už jsme diskutovali při minulém setkání s touto funkcí, přímka procházející počátkem ve směru \vec{u} se s parabolou $y=x^2$ buď neprotne vůbec, nebo jen jednou, každopádně na nějakém okolí počátku platí, že když jdeme po naší přímce, tak vidíme hodnoty funkce nulové. To znamená, že vnímáme f jako konstantní, tudíž $D_{\vec{u}}f(0,0)=0$.

My se ale hlavně zaměříme na to, že v bodě (0,0) existují ve všech směrech směrové derivace, dokonce jsou stejné, ale f není spojitá v (0,0).

Pro směrové derivace tedy neplatí obdoba věty "diferencovatelná, pak spojitá", tudíž to není to správné zobecnění pojmu derivace z jedné proměnné. To ale neznamená, že by směrové derivace nebyly užitečné, jak záhy uvidíme.

V úvodní části jsme uvedli, že nejjednodušší řezy jsou v souřadnicových směrech, protože nemusíme zavádět nový parametr, pracujeme s funkcemi $x \mapsto (x, y_0, z_0, \dots), y \mapsto (x_0, y, z_0, \dots)$ atd, kde navíc máme směrový vektor o normě 1, čili vše je v nejlepším možném stavu.

Jak by vypadala derivace ve směru \vec{e}_1 , tedy ve směru osy x? Pracujeme s parametrickým vyjádřením $x \mapsto (x, y_0, z_0, \dots)$, máme tedy funkci $\varphi(x) = f(x, y_0, z_0, \dots)$, kterou chceme zderivovat podle x. Vidíme, že vlastně ani žádnou novou funkci nepotřebujeme, stačí si ve funkci f zafixovat ostatní souřadnice a derivovat podle x běžným způsobem.

Příklad: Vratme se k funkci $f(x,y) = x^2 + y^2$, zajímá nás derivace ve směru osy y v bodě (1,2). Nejprve to uděláme podle definice. Pohybujeme se po parametrické přímce $t\mapsto (1,2)+t(0,1)$ se směrovým vektorem $\vec{u}=(0,1)$, vzniká funkce $\varphi(t)=1^2+(2+t)^2=t^2+4t+5$. Pak $D_{\vec{u}}f(1,2)=$ $\varphi'(0)=4.$

Alternativní přístup: Do funkce f(x,y) dosadíme za x číslo 1 a výsledný vzorec $f(1,y) = 1^2 + y^2$ zderivujeme "podle y" běžnými pravidly: $[1+y^2]'=2y$. Nakonec dosadíme y=2 a máme stejný výsledek.

 \triangle

Je vidět, že takto snadno získáme derivace obecně v libovolném bodě $\vec{a}=(x_0,y_0)$, například ve směru x ta derivace vychází derivováním funkce $x^2+y_0^2$, kde teď y_0 je konstanta (neznámé, ale pevně zvolené číslo), a protože derivace konstanty y_0^2 (ať už je jakákoliv) je nula, vychází $[x^2 + y_0^2]' = 2x$, tedy derivace v bodě (x_0, y_0) ve směru osy x je $2x_0$.

V praxi se ty nulky nepíšou, prostě se prohlásí, že derivace $f(x,y) = x^2 + y^2$ ve směru x je 2x a myslí se tím v libovolném bodě (x,y), podobně derivace ve směru osy y je 2y. A to je celé tajemství. Protože se tyto derivace počítají tak snadno a zároveň jde o klíčové směry, není divu, že tato myšlenka má speciální jméno.

Definice.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Uvažujme jednotkové vektory \vec{e}_i ve směru souřadnicových os, $\vec{e}_1 = (1, 0, 0, \dots, 0)$, $\vec{e}_2=(0,1,0,\ldots,0),\ldots,\vec{e}_n=(0,0,0,\ldots,1).$ Pro $i=1,\ldots,n$ definujeme **parciální derivaci** f vzhledem k x_i jako $\frac{\partial f}{\partial x_i}(\vec{a}) = D_{\vec{e}_i} f(\vec{a})$, pokud tato existuje.

Alternativní značení: $\frac{\partial f}{\partial x_i}(\vec{a}) = \frac{\partial}{\partial x_i}[f](\vec{a}) = D_i f(\vec{a}) = f'_{x_i}(\vec{a}) = f_{x_i}(\vec{a})$. V praxi se parciálně derivuje podlé dané proměnné tak, že si představujeme, že všechny ostatní proměnné jsou konstanty (a tudíž i všechny výrazy, které vytvoří), a derivuje se podle dané proměnné běžnými pravidly.

Příklad: Najdeme všechny parciální derivace funkce $f(x, y, z) = x^2y + \sin(y^3 + 2z)$. Začneme pomalu a s detaily.

Parciální derivaci podle x získáme tak, že si představujeme, že y a z jsou nějaká konkrétní čísla, třeba $\sqrt{2}$ a π , pak je i $\sin(\sqrt{2}^3+2\pi)$ číslo neboli konstanta. Derivováním tudíž vyjde

$$[x^2\sqrt{2} + \sin(\sqrt{2}^3 + 2\pi)]' = 2x \cdot \sqrt{2} + 0.$$

Tohle jsme samozřejmě nechtěli, ale tento pomocný výpočet nám ukazuje, jaký osud postihl proměnné y a z, když se s nimi zachází jako s konstantami. Všimněte si, že v prvním členu ta $\sqrt{2}$ reprezentující nějakou zvolenou hodnotu y zůstala, šlo o násobící konstantu. Přesně stejnou roli bude v součinu x^2y hrát y, když předstíráme, že je konstantní. Podobně pokud bereme "y" a "z" jako konstanty, bude konstantní i výraz $\sin(y^3+2z)$ a stejnou úvahou jako před chvílí odvodíme

$$f_x(x,y) = \frac{\partial f}{\partial x}(x,y) = 2x \cdot y + 0 = 2xy.$$

Obdobně si pro parciální derivaci podle y můžeme představit, že namísto x a z jsou konstanty, třeba $\sqrt{5}$ a π , a počítáme

$$[\sqrt{5}^{2}y + \sin(y^{3} + 2\pi)]' = \sqrt{5}^{2} + \cos(y^{3} + 2\pi) \cdot 3y^{2}$$

neboli

$$f_y(x,y) = \frac{\partial f}{\partial y}(x,y) = x^2 + \cos(y^3 + 2z) \cdot 3y^2.$$

Samozřejmě zkušený borec si čísla nepíše, jen se naučí dobře předstírat, že někde je konstanta, a trochu si to rozmyslí. Ještě chybí derivace podle z, na to si bereme x a y jako konstanty, pak je ovšem celý člen x^2y konstantní. Proto

$$f_z = \frac{\partial f}{\partial z} = 0 + \cos(y^3 + 2z) \cdot 2.$$

Teď jsme u derivací vynechali to (x, y), lidé to často dělají.

 \triangle

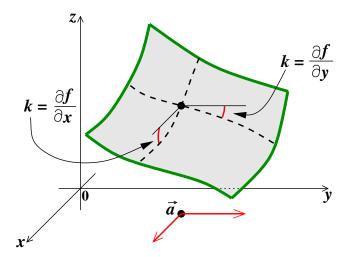
Asi bychom preferovali to krásné stručné značení indexem f_x , ale není tak flexibilní jako to zavedené v definici s krouceníčkem, které kupodivu nemá nějaké jméno kromě "značky pro parciální derivaci". Značka $\frac{\partial}{\partial x_i}$ se dá používat stejně příjemně jako derivační čárka k indikaci derivování konkrétních výrazů, jen se píše zleva. Například u té poslední derivace můžeme použití pravidla pro derivování složené funkce vyznačit zápisem

$$\frac{\partial}{\partial z}[\cos(y^3 + 2z)] = \cos(y^3 + 2z) \cdot \frac{\partial}{\partial z}[y^3 + 2z] = 2\cos(y^3 + 2z).$$

Budeme se tedy tohoto značení držet.

3b. Význam parciálních derivací

Již víme, že parciální derivace dává informace o tom, jak rychle se funkce mění (roste, klesá) v klíčových směrech.



Zdálo by se, že v ostatních směrech si funkce může dělat, co chce, a je to pravda.

Příklad: Uvažujme
$$f(x,y) = \frac{x^2y^2}{x^4 + y^4}$$
 pro $(x,y) \neq (0,0)$ a $f(0,0) = 0$.

Nejprve si jen tak cvičně spočítáme derivaci v bodech $(x, y) \neq (0, 0)$. Tam je vždy funkce definovaná na nějakém okolí konkrétním vzorcem, tudíž stačí použít klasické vzorce.

$$\frac{\partial f}{\partial x} = \frac{2xy^2(x^4 + y^4) - x^2y^2 \cdot 4x^3}{(x^4 + y^4)^2} = \frac{2xy^6 - 2x^5y^2}{(x^4 + y^4)^2},$$
$$\frac{\partial f}{\partial y} = \frac{2x^2y(x^4 + y^4) - x^2y^2 \cdot 4y^3}{(x^4 + y^4)^2} = \frac{2x^6y - 2x^2y^5}{(x^4 + y^4)^2}.$$

To podstatné se bude dít v bodě (0,0). Tam funkce není definována oním algebraickým vzorcem, proto musíme počítat podle definice:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{x \to 0} \left(\frac{f(x,0) - f(0,0)}{x} \right) = \lim_{x \to 0} \left(\frac{0-0}{x} \right) = 0.$$

Podobně $\frac{\partial f}{\partial y}(0,0) = 0.$

Parciální derivace tedy v počátku existují, ale derivace v ostatních směrech překvapivě ne. Abychom se přesvědčili, zvolíme $k \neq 0$ a posuneme se z počátku po přímce y = kx pro x > 0, třeba se směrovým vektorem $\vec{u} = (1, k)$. Potkáváme pak funkční hodnoty $f(x, kx) = \frac{k^4}{1+k^4}$ pro $x \neq 0$, zatímco f(0, 0) = 0. Směrová derivace proto vychází

$$D_{\vec{u}}f(0,0) = \lim_{h \to 0} \left(\frac{f(h,kh) - f(0,0)}{h} \right) = \lim_{h \to 0} \left(\frac{k^4}{1 + k^4} \cdot \frac{1}{h} \right) = \frac{1}{0}.$$

Tato směrová derivace tedy neexistuje

 \triangle

To ukazuje, že obecně nestačí znát parciální derivace k určení derivací směrových, funkce může zlobit. Pokud ovšem po funkci f požadujeme, aby se "nelámala", pak tuto svobodu ztrácí. Poněkud převapivě stačí docela málo, aby byl růst či klesání funkce ve všech směrech jednoznačně určen tím, jak se chová v souřadnicových směrech.

Věta.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Jestliže existuje nějaké okolí bodu \vec{a} , na kterém pro všechna $i=1,\ldots,n$ existují $\frac{\partial f}{\partial x_i}(\vec{x})$ a jsou spojité v \vec{a} , pak má f v \vec{a} derivace ve všech směrech a pro každé \vec{u} platí

$$D_{\vec{u}}f(\vec{a}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a}) \cdot u_i.$$

Požadavky na spojitost derivací jsou splněny často, v zásadě každá funkce zadaná vzorcem toto splňuje, a pro takovéto funkce tedy derivaci v libovolném směru určíme čistě ze znalosti parciálních derivací. To ukazuje, že podmínka spojitých derivací má docela velké dopady. Proto se zavádí užitečné značení, pro oblast G v \mathbb{R}^n je $C^1(G)$ množina všech funkcí na G, které mají na G spojité první parciální derivace. S těmi pracujeme nejraději, například jsou už automaticky spojité.

Pro lepší manipulaci bývá zvykem schovat parciální derivace do jednoho balíčku.

Definice.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Jestliže existují všechny parciální derivace $\frac{\partial f}{\partial x_i}(\vec{a})$ pro = 1,..., n, pak definujeme **gradient** f v \vec{a} jako vektor

$$\nabla f(\vec{a}) = \left(\frac{\partial f}{\partial x_1}(\vec{a}), \dots, \frac{\partial f}{\partial x_n}(\vec{a})\right).$$

Alternativní značení: $\nabla f(\vec{a}) = \vec{\nabla} f(\vec{a}) = \operatorname{grad}(f)(\vec{a}).$

Značení "nabla" ∇ je univerzálně používané, varianta $\vec{\nabla}$ se někdy čte "del". Značka grad(f) je dobrá v tom, že čtenáři napoví, ale je to dlouhé. Budu proto tady nablovat, ale občas připomenu tu alternativu.

Je dobré si uvědomit, že gradient je vektor z \mathbb{R}^n , tedy vnímáme jej jako objekt ze světa definičního oboru funkce, na symbolickém grafu jej vidíme v rámci vodorovného znázornění D(f) (viz obrázek níže).

Pro funkci se spojitými derivacemi (tedy pro naprostou většinu běžně potkávaných funkcí) se závěr věty dá elegantně napsat pomocí skalárního součinu jako

$$D_{\vec{u}}f(\vec{a}) = \nabla f(\vec{a}) \bullet \vec{u}.$$

Příklad: Uvažujme zase $f(x,y)=x^2+y^2$. Parciální derivace jsme již počítali, tudíž hravě dostáváme gradient $\nabla f(x,y)=(2x,2y)$.

V bodě $\vec{a} = (1,2)$ pak $\nabla f(1,2) = (2,4)$ a nový vzorec dává

$$D_{(h,k)}f(1,2) = \nabla f(1,2) \bullet (h,k) = 2h + 4k,$$

přesně jak jsme to měli výše výpočtem podle definice.

Δ

Gradient má zajímavé teoretické vlastnosti, například splňuje povědomá počítací pravidla či pro něj platí povědomé tvrzení, že pokud má dostatečné hladká funkce na nějaké oblasti nulový gradient, pak musí být konstantní. To ale necháme do kapitoly Více o derivaci, zde nás zajímá, jakou užitečnou informaci gradient umí poskytnout. Ukáže se, že je to jeden z klíčových pojmů při zkoumání funkcí.

Gradient a spád

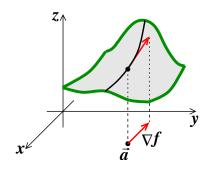
Představme si, že jsme v bodě \vec{a} , sedíme na grafu a rozhlížíme se, jak ten povrch vypadá. Podle toho, kterým směrem se díváme, tak stoupá či klesá, rychlost stoupání či klesání je dána směrovou derivací. Jinak řečeno, je dána výrazem $\nabla f(\vec{a}) \bullet \vec{u}$, kde \vec{u} jsou jednotkové vektory.

Podle známého vzorce platí

$$\nabla f(\vec{a}) \bullet \vec{u} = \|\nabla f(\vec{a})\| \cdot \|\vec{u}\| \cos(\alpha) = \|\nabla f(\vec{a})\| \cos(\alpha),$$

kde α je úhel mezi vektory $\nabla f(\vec{a})$ a \vec{u} .

Zajímají nás extrémní hodnoty této směrové derivace. Vidíme, že stoupáme nejrychleji, pokud se vydáme tak, aby $\cos(\alpha) = 1$, což je pro $\alpha = 0$ neboli v směru gradientu. Naopak největší klesání nastane v případě $\cos(\alpha) = -1$, tedy přesně v opačném směru.



Fakt.

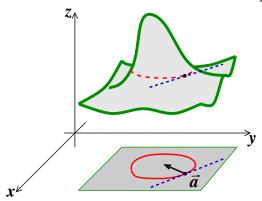
Nechť f je funkce, která má na okolí bodu \vec{a} spojité první parciální derivace. Pak gradient $\nabla f(\vec{a})$ udává směr největšího růstu funkce v bodě \vec{a} , funkce tam roste rychlostí $\|\nabla f(\vec{a})\|$.

Vektor $-\nabla f(\vec{a})$ udává směr největšího spádu v bodě \vec{a} .

Gradient a hladiny konstantnosti

Jsme stále v bodě \vec{a} a sedíme na grafu. Toto místo na grafu má jistě nějakou úroveň (nadmořskou výšku), jmenovitě úroveň $c=f(\vec{a})$, a bod \vec{a} pak nutně patří do příslušné hladiny konstantnosti (která leží v definičním oboru). Pokud chodíme po této hladině, pak se naše výška nebude měnit. Předpokládejme, že se v bodě \vec{a} hladina konstantnosti neláme (jako křivka je diferencovatelná, tedy hladká), takže má smysl mluvit o tečně k této hladině v \vec{a} . Uvažujme směrový vektor \vec{u} této tečny.

Pokud se vydáme z bodu \vec{a} v tomto směrem \vec{u} neboli po tečně, pak se na malou chvíli od naší hladiny konstantnosti téměř neodchýlíme, a proto také hodnota funkce zůstane v zásadě stejná. To znamená, že $D_{\vec{u}}f(\vec{a}) = 0$ neboli $\|\nabla f(\vec{a})\|\cos(\alpha) = 0$ neboli $\alpha = \frac{\pi}{2}$.



Právě jsme nahlédli, že směr, kterým příslušná hladina konstantnosti vychází z bodu \vec{a} , je kolmý ke směru, který tam má gradient. Jinak řečeno, směr, ve kterém bychom šli, abychom zůstali ve stejné výšce, je vždy kolmý na směr, kterým v tom místě svah roste nepříkřeji. Potvrdíme si to.

Fakt.

Má-li funkce f na okolí bodu \vec{a} spojité první parciální derivace, pak je gradient $\nabla f(\vec{a})$ kolmý na hladinu konstantnosti procházející bodem \vec{a} .

Toto je velice užitečné. Mnohé útvary lze reprezentovat jako hladiny konstantnosti vhodné funkce, gradient pak umožňuje snadno získat kolmice k dotyčnému útvaru, což se hodí v řadě aplikací.

Příklad: Uvažujme elipsu danou rovnicí $\frac{x^2}{6} + \frac{y^2}{3} = 1$, chceme najít její tečnu v bodě (2,1). Jedna možnost je použít přístup přes grafy. Daný bod leží na horní polovině elipsy, kde ji lze vnímat (pokud vzorec řešíme pro y) jako graf funkce $f(x) = \sqrt{3 - \frac{x^2}{2}}$. K nalezení tečny pro x = 2

potřebujeme derivaci $f'(x)=\frac{-x}{2\sqrt{3-\frac{x^2}{2}}}$, směrnice tečny je proto k=f'(2)=-1. Dostáváme přímku y-1=-(x-2) neboli x+y=3.

Alternativní přístup: Danou rovnici upravíme na příjemnější tvar $x^2 + 2y^2 = 6$ a rozhodneme se, že je to vlastně hladina konstantnosti funkce $F(x,y) = x^2 + 2y^2$ odpovídající c = 6. Najdeme gradient v bodě (2,1): $\nabla F = (2x,4y)$, proto $\nabla F(2,1) = (4,4)$.

Tento vektor je kolmý na hladinu konstantnosti procházející bodem (1,2) neboli naši elipsu, proto je i kolmý na její tečnu. Rovnice přímky kolmé na (4,4) a procházející bodem (2,1) je 4(x-2)+4(y-1)=0 neboli x+y=3.

Gradient a tečny, aproximace

Intuitivně vzato, tečna ke grafu funkce f(x) je objekt stejného typu (graf je zakřivená čára, tečna je přímá čára), který se dotýká a splývá s grafem v daném bodě. To naznačuje, jakým způsobem budeme tuto myšlenku adaptovat pro případ více proměnných. Představíme si graf funkce dvou proměnných, to je jakási dvourozměrná plachta vznášející se ve třírozměrném prostoru, a je přirozené přiložit k ní rovinu, tedy plochý dvourozměrný objekt.

Podobně bychom pro funkci tří proměnných hledali tečné třírozměrné prostory (které vypadají "ploše", když jsou zasazeny do čtyřrozměrného prostoru, ve kterém žije graf) atd. Obecně nás zajímají **nadroviny** v prostoru \mathbb{R}^{n+1} , což jsou afinní lineární prostory o dimenzi n. My hledáme jisté speciální.

Jak je najdeme? Tím, že se na chvíli oprostíme od geometrie a podíváme se na tečny analyticky. Víme, že u funkcí jedné proměnné je tečna v bodě a taková přímka, která nejlépe ze všech přímek aproximuje chování f okolo bodu a,

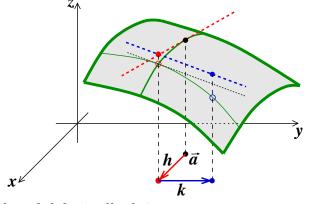
$$f(a+h) \approx f(a) + f'(a)h.$$

Jak bychom co nejlépe aproximovali hodnoty funkce f(x,y) okolo bodu $\vec{a} = (a,b)$? Představme si, že se z něj o kousíček posuneme, jmenovitě o vektor $\vec{u} = (h,k)$. Jak se změní funkce?

Namísto pohybu jakoby diagonálního se do stejného místa (a+h,b+k) dostaneme, pokud nejprve popojdeme o h podél osy x a pak o k podél osy y. Ten první pohyb je jednorozměrná záležitost, měníme jednu proměnnou, a tam už odhadnout změnu funkce umíme, použijeme derivování v příslušném směru:

$$f(a+h,b) \approx f(\vec{a}) + \frac{\partial f}{\partial x}(\vec{a})h.$$

Vidíme to na obrázku červenou barvou, namísto hodnoty v kroužku bereme hodnotu v plném kolečku.



Z bodu (a+h,b) se teď posuneme ve směru osy y o k a obdobně odhadujeme

$$f(a+h,b+k) \approx f(a+h,b) + \frac{\partial f}{\partial y}(a+h,b)k.$$

Dáme to dohromady:

$$f(a+h,b+k) \approx f(\vec{a}) + \frac{\partial f}{\partial x}(\vec{a})h + \frac{\partial f}{\partial y}(a+h,b)k.$$

Nepříjemné je, že v posledním členu nepoužíváme derivaci v \vec{a} . Pokud ale budeme předpokládat, že tato derivace existuje na nějakém okolí \vec{a} a je v \vec{a} spojitá, pak je možné v případě velmi malého h ten posun o h zanedbat v naději, že se jím hodnota $\frac{\partial f}{\partial y}$ příliš nezmění. Dostaneme

$$f(a+h,b+k) \approx f(\vec{a}) + \frac{\partial f}{\partial x}(\vec{a})h + \frac{\partial f}{\partial y}(\vec{a})k.$$

Dostali jsme aproximační vzorec, který je lineární, a tedy popisuje rovinu ve třídimenzionálním prostoru. Abychom to lépe viděli, použijeme obvyklou substituci $x=a+h,\,y=b+k$ a obdržíme druhou populární podobu aproximačního vzorce

$$f(x,y) \approx f(\vec{a}) + \frac{\partial f}{\partial x}(\vec{a})(x-a) + \frac{\partial f}{\partial y}(\vec{a})(y-b).$$

Výraz napravo lze definuje funkci, jejíž graf lze popsat vzorcem

$$z = f(\vec{a}) + \frac{\partial f}{\partial x}(\vec{a})(x - a) + \frac{\partial f}{\partial y}(\vec{a})(y - b),$$

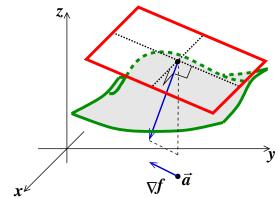
ve kterém již rozpoznáme rovnici roviny, naší hledané tečné roviny. Můžeme ji zjednodušit na tvar

$$\frac{\partial f}{\partial x}(\vec{a})x + \frac{\partial f}{\partial y}(\vec{a})y - z = -f(\vec{a}) + \frac{\partial f}{\partial x}(\vec{a})a + \frac{\partial f}{\partial y}(\vec{a})b$$

neboli

$$\frac{\partial f}{\partial x}(\vec{a})x + \frac{\partial f}{\partial y}(\vec{a})y - z = d$$

pro vhodné d a vidíme, že tato rovina je dána normálovým vektorem $\left(\frac{\partial f}{\partial x}(\vec{a}), \frac{\partial f}{\partial x}(\vec{a}), -1\right)$



Principiálně to dává smysl i na intuitivní úrovni. Představme si tečnou rovinu umístěnou na vrcholku onoho kopce. Pak by měla být vodorovná a její normálový směr svislý, představme si ho směrem dolů a s přirozenou velikostí 1, jde tedy o vektor (0,0,-1). Pokud chceme přiložit tečnou rovinu k bodu na svahu, pak tu rovinu musíme z vodorovné polohy vychýlit, a zdá se zjevné, že rovinu a tedy i její normálový vektor musíme vychýlit směrem, kterým onen svah roste. A to je právě směr gradientu (který vidíme v rovině definičního oboru). Čím je svah příkřejší, tím více musíme normálový vektor vychýlit, ale příkřejší svah také znamená větší gradient, to souhlasí.

Obdobná úvaha platí i ve více rozměrech, máme aproximační vzorec

$$f(\vec{x} + \vec{h}) \approx f(\vec{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a})h_i = f(\vec{a}) + \nabla f(\vec{a}) \bullet \vec{h},$$

popřípadě verzi

$$f(\vec{x}) \approx f(\vec{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a})(x_i - a_i) = f(\vec{a}) + \nabla f(\vec{a}) \bullet (\vec{x} - \vec{a}),$$

která vede k rovnici tečné nadroviny

$$z = f(\vec{a}) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a})(x_i - a_i).$$

Potvrdíme si to oficiálně.

Fakt.

Nechť funkce f má spojité první derivace na okolí bodu \vec{a} . Jestliže vektor $\nabla f(\vec{a})$ rozšíříme o jednu souřadnici navíc, jmenovitě přidáme -1 jako souřadnici (n+1), tak dostaneme vektor v \mathbb{R}^{n+1} kolmý na tečnou nadrovinu ke grafu v bodě \vec{a} .

Příklad: Uvažujme $f(x,y)=x^2+y^2$ a bod (1,2). Najdeme tečnou nadrovinu ke grafu f v odpovídajícím bodě.

Spočítali jsme výše, že $\nabla f(1,2)=(2,4)$. Normálový vektor ke grafu je tedy například $\vec{n}=(2,4,-1)$.

Skrz jaký bod má tečná rovina jít? Protože f(1,2) = 5, je to bod (1,2,5). Máme bod a normálový vektor, z toho napíšeme rovnici roviny snadno:

$$0 = \vec{n} \bullet ((x, y, z) - (1, 2, 5)) = 2(x - 1) + 4(y - 2) - (z - 5) \implies 2x + 4y - z = 5.$$

Alternativa: Rovina kolmá na vektor (2,4,-1) má rovnici 2x+4y-z+d=0. Dosazením bodu (1,2,5) dostaneme d=-5, odtud 2x+4y-z-5=0.

Další alternativa: Graf je dán rovnicí $z=x^2+y^2$. Po přepisu na tvar $x^2+y^2-z=0$ toto lze vnímat jako hladinu konstantnosti funkce $F(x,y,z)=x^2+y^2-z$ odpovídající hodnotě c=0. Snadno najdeme $\nabla F=(2x,2y,-1)$ a víme, že vektor $\nabla F(1,2,5)=(2,4,-1)$ je kolmý na tuto hladinu, tedy i na náš graf a potažmo na hledanou tečnou rovinu. Její rovnice je proto

$$2(x-1) + 4(y-2) - (z-5) = 0$$

a jsme hotovi.

Závěr: Tečná rovina ke grafu f v bodě daném $\vec{a} = (1, 2)$ má rovnici 2x + 4y - z = 5.

Ještě si na tomto příkladě zopakujeme další poznatky o gradientu.

Funkce poroste nejrychleji, pokud se z bodu (1,2) vydáme ve směru $\nabla f(1,2) = (2,4)$ neboli ve směru (1,2) (každý kladný násobek má stejný směr), rychlost růstu funkce pak bude $\|\nabla f(1,2)\| = \|(2,4)\| = \sqrt{20} = 2\sqrt{5}$.

Bod $(1,2) \in D(f)$ leží na hladině konstantnosti f(1,2)=5 neboli na kružnici dané rovnicí $y^2+x^2=5$. V bodě (1,2) je vektor $\nabla f(1,2)=(2,4)$ kolmý na tuto křivku, díky čemuž hravě získáme rovnici tečny k této křivce:

$$0 = \nabla f(1,2) \bullet ((x,y) - (1,2)) = 2(x-1) + 4(y-2) \implies 2x + 4y = 10.$$

Z normálového směru (1,2) dostaneme známým trikem k němu kolmý vektor v \mathbb{R}^2 , třeba (2,-1) coby směr tečný k oné hladině konstantnosti. \triangle

3c. Parciální derivace vyššího řádu

Podobně jako u funkcí jedné proměnné, i funkce více proměnných lze derivovat vícekrát, pokud nám to dovolí. Představte si například, že máme funkci f, směrový vektor \vec{u} a zjistíme, že $D_{\vec{u}}f$ existuje na nějakém okolí bodu \vec{a} . Pak tato směrová derivace vlastně vytváří funkci $\vec{x} \mapsto D_{\vec{u}}f(\vec{x})$ na okolí \vec{a} a tu lze zase derivovat v \vec{a} , například ve směru \vec{v} . Dostaneme $D_{\vec{v}}[D_{\vec{u}}f] = D_{\vec{v}}D_{\vec{u}}f$. Všimněte si pořadí diferenciálních operátorů ve výrazu napravo. Aplikují se zprava doleva, takže nejprve derivujeme ve směru \vec{u} .

Dala by se takto rozvinout obecná teorie, ale my se přirozeně zaměříme na parciální derivace. Dají se aplikovat opakovaně (přinejmenším pro dostatečně příjemné funkce), a je jasné, že máme na výběr, jaké derivace chceme aplikovat. U funkce dvou proměnných máme $\frac{\partial f}{\partial x}$ a $\frac{\partial f}{\partial y}$ a obě tyto derivace můžeme následně derivovat podle x či podle y, čímž vzniknou celkem čtyři různé parciální derivace druhého řádu, například následující dvě. Ukážeme nejprve podrobný zápis postupu a pak standardní kondenzované značení:

$$\frac{\partial}{\partial x} \left[\frac{\partial}{\partial x} [f] \right] = \frac{\partial}{\partial x} \left[\frac{\partial f}{\partial x} \right] = \frac{\partial^2 f}{\partial x^2}, \qquad \frac{\partial}{\partial x} \left[\frac{\partial}{\partial y} [f] \right] = \frac{\partial}{\partial x} \left[\frac{\partial f}{\partial y} \right] = \frac{\partial^2 f}{\partial x \partial y}.$$

[3c] 33

Všimněte si pořadí derivování, symboly ve jmenovateli se berou zprava doleva, tedy začneme proměnnou nejvíce vpravo. Například parciální derivaci třetího řádu $\frac{\partial^3 f}{\partial x \partial y \partial x}$ bychom získali tak, že nejdříve derivujeme f podle x, výsledek pak podle y a to zase dle x, zatímco pro získání $\frac{\partial^3 f}{\partial x^2 \partial y}$ bychom nejprve derivovali podle y a pak dvakrát dle x.

Definice.

Uvažujme funkci f definovanou na okolí bodu $\vec{a} \in \mathbb{R}^n$.

Nechť $i_1, i_2, \ldots, i_m \in \{1, 2, \ldots, n\}$ jsou indexy souřadnic. Definujeme odpovídající **parciální drivaci řádu** m funkce f indukcí jako

$$\frac{\partial^m f}{\partial x_{i_m} \partial x_{i_{m-1}} \cdots \partial x_{i_2} \partial x_{i_1}} = \frac{\partial}{\partial x_{i_m}} \left[\frac{\partial^{m-1} f}{\partial x_{i_{m-1}} \cdots \partial x_{i_2} \partial x_{i_1}} \right]$$

za předpokladu, že všechny potřebné derivace existují.

Pokud nejsou všechny indexy souřadnic i_k stejné, pak se taková derivace nazývá smíšená.

Aby to bylo zajímavější, i pro opakované derivování existuje velmi příjemné alternativní značení dolním indexem, u kterého je ale naopak přirozené psát pořadí zleva doprava: $[f_x]_y = f_{xy} = \frac{\partial^2 f}{\partial y \partial x}$. Pokud si nejste jisti pořadím, obvykle pomůže rozepsat si kompaktní zápis do podoby opakovaného derivování a to pak naznačí postup, například

$$f_{x^2zy} = [[[f_x]_x]_z]_y = \frac{\partial}{\partial y} \left[\frac{\partial}{\partial z} \left[\frac{\partial}{\partial x} \left[\frac{\partial}{\partial x} [f] \right] \right] \right] = \frac{\partial^4 f}{\partial y \partial z \partial x^2}.$$

V obou případech tedy nejprve derivujeme dvakrát podle x, pak podle z a nakonec podle y.

Obecně má funkce o n proměnných celkem n^m parciálních derivací řádu m. Je snadné najít příklady, kde na pořadí opravdu záleží, ale pokud tak čtenář učiní, tak uvidí, že obvykle nejde o zrovna přátelské funkce. Jinak řečeno, pro pěkné funkce na pořadí derivování nezáleží. To nám samozřejmě silně zjednodušší život.

Věta.

Jestliže má funkce f všechny parciální derivace řádu m na nějakém okolí bodu \vec{a} a jsou všechny spojité v \vec{a} , pak při výpočtu konkrétní parciální derivace až po řád m nezáleží na pořadí derivování.

Co to znamená? Parciálních derivací je obecně docela dost, například pokud pracujeme s funkcí tří proměnných, tak je teoreticky celkem $3^4=81$ parciálních derivací čtvrtého řádu. Pokud jsou ale derivace čtvrtého řádu spojité (například když je naše funkce zadaná vzorcem poskládaným z elementárních funkcí), tak stačí najít jen 10 derivací třetího řádu namísto 27 a 15 derivací čtvrtého řádu namísto 81. Popravdě řečeno to je spíše úspora teoretická, protože v praxi si většinou vystačíme s prvními dvěma řády, ale potěší.

Podobně jako u derivací prvního řádu, i vyšší derivace se dají sebrat do jednoho celku.

Definice.

Předpokládejme, že funkce f má všechny parciální derivace druhého řádu v bodě \vec{a} . Pak definujeme její **Hessovu matici** v bodě \vec{a} jako

$$H(\vec{a}) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}\right)_{i,j=1,\dots,n} = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

Prakticky řečeno, zderivujeme funkci f dle první proměnné, tuto derivaci pak zderivujeme znovu postupně podle všech proměnných a z výsledků vytvoříme první řádek matice. Tento první řádek je tedy vlastně $\nabla \left(\frac{\partial f}{\partial x_1}\right)$, obdobně pak další. Na diagonále vidíme nesmíšené derivace $\frac{\partial^2 f}{\partial x_i^2}$, mimo ni pak tv smíšené.

Pokud je daná funkce trochu rozumná, aby na pořadí derivování nezáviselo, pak je Hessova matice symetrická, takže to zabere jen zhruba polovinu práce. Pro funkce dvou proměnných je nicméně dobrý nápad počítat smíšené derivace nezávisle, protože to není zase tolik práce navíc a jejich rovnost pak slouží jako ověření správnosti.

Pro shrnutí derivací třetího řádu bychom potřebovali třírozměrnou matici, jakousi krychli, čímž se dostáváme k tenzorům, které tady rozhodně nechceme probírat. V mnoha (většině?) aplikacích si vystačíme s prvními dvěma derivacemi, spokojíme se s nimi i zde.

Protože se funkce se spojitými parciálními derivacemi dobře chovají, často se v teorii omezujeme jen na ně. Abychom si usnadnili zápis, zavedeme si užitečné značení.

Definice.

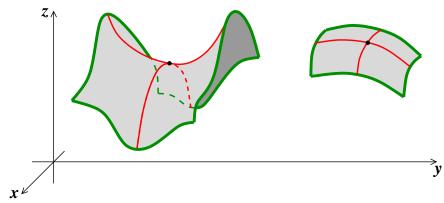
Nechť G je otevřená množina v \mathbb{R}^n . Pro $m \in \mathbb{N}$ značíme symbolem $C^m(G)$ množinu všech funkcí definovaných (alespoň) na G, které mají ve všech bodech G všechny parciální derivace řádu m a ty jsou na G spojité.

Jako u funkcí jedné proměnné i zde platí, že pokud má funkce spojité všechny parciální derivace řádu m, pak musí být také parciální derivace všech nižších řádů spojité.

3d. Význam derivací vyššího řádu

U funkcí jedné proměnné obvykle nabízíme interpretace prvních dvou derivací, ta první ukazuje směrnici tečny či rychlost změny funkce, ta druhá intenzitu zakřivení grafu schovanou pod jménem konvexita a konkávnost. O parciálních derivacích prvního řádu jsme již hovořili, takže se podíváme na význam parciálních derivací druhého řádu.

Nejjednodušší je to s nesmíšenými derivacemi, protože pokud opakovaně derivujeme podle stejné proměnné, pak ty ostatní směry lze ignorovat. Pracujeme pak vlastně s řezem naší situace, se kterým lze zacházet jako s funkcí jedné proměnné. Pokud například graf funkce řízneme ve směru osy x, tak nám $\frac{\partial^2 f}{\partial x^2}$ určuje konvexitu řezu, přesně jak jsme zvyklí, obdobnou informaci podávají $\frac{\partial^2 f}{\partial y^2}$, $\frac{\partial^2 f}{\partial z^2}$ atd. To nám říká něco o tvaru grafu.



Nalevo vidíme případ, kdy mají nesmíšené druhé parciální derivace v jistém bodě rozdílná znaménka, tedy funkce je tam v jednom směru konvexní a v druhém konkávní. Dokážeme si to docela dobře představit. Napravo je pak situace, kdy je v obou směrech konkávnost.

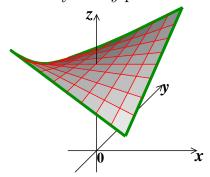
Tento druhý příklad inspiruje otázku obdobnou té, kterou jsme již položili ohledně prvních derivací. Tam nám vyšlo, že pokud u hladké funkce známe gradient, tedy rychlost růstu v souřadnicových směrech, tak to již jednoznačně určuje růst ve všech směrech. Platí podobně, že když u

hladké funkce známe nesmíšené druhé derivace, tedy konkávnost v souřadnicových směrech, tak už je tím určena konkávnost i v dalších řezech? Odpověď zní, že ne, pro posouzení vypouklosti grafu je potřeba znát všechny parciální derivace druhého řádu. Čímž se dostáváme k významu smíšených derivací.

Nejprve se podívejme na derivaci $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y} \left[\frac{\partial f}{\partial x} \right]$ a představme si situaci, že je kladná. Máme funkci $\frac{\partial f}{\partial x}$ a derivujeme ji podle y, takže se pohybujeme ve směru y. Předpokládali jsme, že tato druhá derivace je kladná, což nám říká, že funkce $\frac{\partial f}{\partial x}$ roste, když se pohybujeme ve směru osy y. Geometricky, pohybujeme se ve směru osy y a díváme se na směrnice tečen ve směru x, tedy v

Umíte si představit situaci, kdy se pohybujete ve směru osy y a přitom se tečny ve směry x otáčejí směrem k většímu růstu? Takový graf je zkroucený. Ukážeme to na obrázku, na kterém se podíváme, co se děje okolo počátku. Abychom to viděli lépe, otočili jsme osy a x jde doprava, jak jsme zvyklí při znázorňování směrnice tečny. Osa y pak nutně musí jít směrem od nás.

kolmém směru. Zjišťujeme, že tyto směrnice rostou, tedy tečny se stáčejí směrem nahoru.



Všimněte si, že řezy grafem ve směrech x a y jsou přímky. To je možné (například jednoduchá funkce f(x,y)=xy to splňuje) a to důležité na tom je, že pro funkci na obrázku platí $\frac{\partial^2 f}{\partial x^2}=\frac{\partial^2 f}{\partial y^2}=0$. Jinak řečeno, parciální derivace $\frac{\partial^2 f}{\partial x^2}$ a $\frac{\partial^2 f}{\partial y^2}$ pak neovlivňují tvar grafu, takže můžeme přímo pozorovat vliv smíšené druhé derivace.

Pokud je funkce dostatečně hladká, tak by mělo platit $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$, měli bychom tedy dostat stejný obrázek i při interpretaci výrazu $\frac{\partial}{\partial x} \left[\frac{\partial f}{\partial y} \right] > 0$: Pokud se pohybujeme ve směru osy x, tak se směrnice tečen ve směru y stáčejí nahoru. Obrázek toto odráží, větší směrnice "ypsilonové tečny" znamená příkřejší růst ve směru osy y, tedy směrem od nás.

Viděli jsme, že význam smíšené druhé derivace je směr a míra zkroucení grafu. Právě takovéto zkroucení grafu může způsobit, že konvexita v souřadnicových směrech ještě neznamená tvar dolíku, při zkoumání zakřivení grafu je třeba (matematickým způsobem) posoudit vzájemné působení konvexity v souřadnicových směrech a zkroucení grafu indikované smíšenou derivací, na což samozřejmě narazíme v části o lokálních extrémech. Z této krátké sekce si odneseme poznání, že informaci, kterou u funkcí jedné proměnné hledáme pomocí druhé derivace, najdeme v případě více proměnných v Hessově matici, kde všechny složky (všechny derivace druhého řádu) hrají stejně důležitou roli.

4. Představujeme lokální a globální extrémy

Lokální i globální extrémy se přímo zobecní pro více dimenzí. Ačkoliv definice vypadají stejně, při vyšetřování extrémů narazíme na podstatné rozdíly.

4a. Lokální extrémy

Definice lokálního extrému prostě zkopírujeme a nahradíme body značkami vektorů.

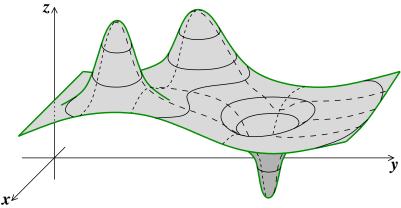
Definice.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$.

Řekneme, že f má v \vec{a} lokální maximum nebo že $f(\vec{a})$ je lokální maximum, jestliže existuje okolí U bodu \vec{a} takové, že $f(\vec{a}) \geq f(\vec{x})$ pro všechna $x \in U$.

Řekneme, že f má v \vec{a} lokální minimum nebo že $f(\vec{a})$ je lokální minimum, jestliže existuje okolí U bodu \vec{a} takové, že $f(\vec{a}) \leq f(\vec{x})$ pro všechna $x \in U$.

Obrázek níže pro případ dvou proměnných ukazuje vlevo dvě lokální maxima a vpravo jedno lokální minimum.



Podobně si to představujeme ve více dimenzích. Lokální maximum má vlastnost, že když skrze něj řízneme graf v libovolném směru (čímž se z toho stane situace funkce jedné proměnné), tak na tom řezu máme lokální maximum v klasickém významu. Obdobná vlastnost platí pro lokální minimum.

Existuje ještě jeden zajímavý typ chování, který na obrázku vidíme mezi oněmi dvěma kopci. Pokud tam graf řízneme svislou rovinou jdoucí od vrcholu k vrcholu, pak v tom údolí mezi nimi vidíme na řezu lokální minimum. Když ale řízneme graf ve směru kolmém jdoucím mezi vrchy, tak v tom údolí uvidíme na řezu lokální maximum. Takovýmto bodům říkáme sedlo či sedlový bod a při zkoumání extrémů se také připlétají do hry, takže se běžně berou jako součást zkoumání extrémů.

Jak lokální extrémy najdeme? Postup je obdobný vyšetřování lokálních extrémů pro funkce jedné proměnné. Zhruba řečeno, nejprve najdeme kandidáty pomocí první derivace, pak je posoudíme pomocí druhé.

Když graf řízneme libovolnou svislou rovinou skrz nějaký lokální extrém, bude extrém i na řezu, tedy derivace v tomto směru musí být nulová. Zejména to platí pro řezy rovnoběžné s osami, tedy parciální derivace musí být nulové. Jinak řečeno, gradient musí být nulový (jako vektor).

Jiná úvaha: V lokálním extrému je tečná nadrovina vodorovná, musí tedy mít normálový vektor svislý. Normálový vektor je ovšem $\left(\frac{\partial f}{\partial x_1}(\vec{a}), \ldots, \frac{\partial f}{\partial x_n}(\vec{a}), -1\right)$ a ten je svislý, pokud $\frac{\partial f}{\partial x_i}(\vec{a}) = 0$ pro všechna i neboli $\nabla f(\vec{a}) = \vec{0}$.

Věta.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Jestliže má f v \vec{a} lokální extrém, pak $\nabla f(\vec{a}) = \vec{0}$, pokud gradient v \vec{a} existuje.

Bodům, kde $\nabla f(\vec{x}) = \vec{0}$, říkáme **stacionární body**. Najdeme je při troše štěstí řešením soustavy n rovnic $\frac{\partial f}{\partial x_i}(\vec{x}) = 0$ o n neznámých x_1, \ldots, x_n .

Jako obvykle neplatí naopak, že stacionární body dávají lokální extrémy. Například typický sed-

Jako obvýkle neplatí naopak, že stacionární body dávají lokální extrémy. Například typický sedlový bod u funkce dvou proměnných je také stacionární bod. Význam věty je následující: Pokud budeme hledat lokální extémy jen mezi stacionárními body (popřípadě těmi bez derivace), tak nám určitě žádný neunikne.

Když tedy shromáždíme kandidáty (body s nulovým či žádným gradientem), je potřeba je ještě klasifikovat. To není snadné pro body bez gradientu, takže si usnadníme život a od teď v této kapitole pracujeme jen s hladkými funkcemi. Pak jsou jediným zdrojem kandidátů stacionární body, pro které existuje vcelku spolehlivý test zvaný Sylvestrovo kritérium. Je mnohem snažší si jej pamatovat, pokud nám dává smysl intuitivně, takže se podíváme, odkud přichází.

Představme si, že funkce f má v bodě \vec{a} lokální maximum (kopeček). Pokud graf řízneme ve směru osy x, pak v řezu opět vidíme kopeček, tedy funkce $x \mapsto f(x, a_2, a_3, \dots)$ má v bodě a_1 také lokální maximum. Její druhá derivace, což je vlastně $\frac{\partial^2 f}{\partial x^2}(\vec{a})$, tedy musí být záporná (nebo nulová). Protože souřadnice x není ničím speciální, usoudíme, že pokud má funkce f lokální maximum v bodě \vec{a} , pak nutně $\frac{\partial^2 f}{\partial x_i^2}(\vec{a}) \leq 0$ pro všechna i. Jinak řečeno, a pokud vynecháme velmi řídký případ nul, dá se očekávat, že v bodě lokálního maxima bude mít Hessova matice záporná čísla na diagonále.

Podobně u lokálního minima zase očekáváme, že platí $\frac{\partial^2 f}{\partial x_i^2}(\vec{a}) > 0$ (popřípadě zřídka i nula) pro všechna i, tedy máme kladná čísla na diagonále Hessovy matice. Když se podíváme na oba případy najednou, vidíme, že existence lokálního extrému znamená, že nesmíšené druhé parciální derivace (neboli diagonální členy Hessovy matice) musejí mít stejná znaménka. Je zajímavé se zamyslet, jaká situace by vedla na případ, kdy máme na diagonále Hessovy matice čísla s rozdílnými znaménky. To by znamenalo, že na některých řezech grafu máme konvexitu a na jiných konkávnost, neboli jde o sedlo.

Zdá se, že nám tu vzniká pěkný test na rozeznávání lokálních extrémů a sedel založený na diagonále Hessovy matice, ale bohužel to tak snadné nebude. Je potřeba vzít také v úvahu kroutivost grafu, jinak řečeno smíšené derivace také mají vliv. Abychom to měli snadnější, začneme případem dvou proměnných.

Tam lze shodu či neshodu znamének vyjádřit elegantně součinem. Víme tedy, že lokální extrémy mají splňovat $\frac{\partial^2 f}{\partial x^2}(\vec{a}) \frac{\partial^2 f}{\partial y^2}(\vec{a}) \geq 0$, zatímco sedla mají splňovat $\frac{\partial^2 f}{\partial x^2}(\vec{a}) \frac{\partial^2 f}{\partial y^2}(\vec{a}) \leq 0$. Mimochodem, vlastně jsme se dívali na součin diagonály Hessovy matice

$$H(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial y \partial x}(x,y) \\ \frac{\partial^2 f}{\partial x \partial y}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{pmatrix}.$$

Pokud chceme, aby toto fungovalo i v opačném směru, tak tento výraz musíme obohatit i o zbývající členy této matice. Známe nějaký matematický vzorec pro matice 2×2 , který zahrnuje všechny členy, přičemž diagonála se násobí? Ano, determinant, a kupodivu jsme se trefili. Dá se dokázat, že pro stacionární body funkcí dvou proměnných dokáže znaménko $\det(H)$ poznat lokální extrémy a sedla.

Když už máme lokální extrém, potřebujeme rozpoznat minima a maxima. Ale to je snadné, stačí se podívat na tvar nějakého řezu, například pohledem na znaménko $\frac{\partial^2 f}{\partial x^2}(\vec{a})$. Dostáváme tak následující algoritmus.

Algoritmus. (Vyšetřování lokálních extrémů pro f(x,y))

1. Řešením rovnice $\nabla f(x,y) = \vec{0}$ neboli soustavy

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

najdeme stacionární body.

- 2. Pro každý stacionární bod (a,b) najdeme odpovídající Hessovu matici H=H(a,b).
- Jestliže det(H) < 0, je v (a, b) sedlový bod.
- Jestliže $\det(H) > 0$, je v (a,b) lokální extrém. Je to lokální maximum pokud $\frac{\partial^2 f}{\partial x^2}(a,b) < 0$, a je to lokální minimum pokud $\frac{\partial^2 f}{\partial x^2}(a,b) > 0$.

Tento specifický algoritmus není jen inspirací pro to, co záhy přijde, ale může být pro studenty užitečný sám o sobě, protože mnozí z nich se potkají pouze s funkcemi o dvou proměnných, pak je tato speciální verze příjemná. Poznamenejme, že v případech, kdy vychází v klíčových okamžicích nuly, rozhodovací algoritmus selhává, nic nevíme a je třeba použít pokročilejší metody, které jsou za rámec tohoto úvodu.

Pokud budeme chtít tento proces zobecnit pro více proměnných, musíme oba testy sjednotit. Nejprve jsme se dívali na determinant celé matice, v případě lokálního extrému jsme se pak rozhodovali podle $\frac{\partial^2 f}{\partial x^2}(a, b)$, ale to je vlastně také determinant, jmenovitě determinant levé horní submatice 1×1 . Pracujeme tedy se stejnými objekty, subdeterminanty Hessovy matice, což se dá snadno zobecnit na větší matice.

Aby nám podmínky, které dostaneme, dávaly smysl, představme si zatím jednodušší situaci, kdy smíšené parciální derivace jsou nula, tedy Hessova matice je diagonální a determinanty se počítají jednoduše.

$$\Delta_{1} = \begin{pmatrix} \frac{\partial^{2}f}{\partial x_{1}^{2}} & 0 & 0 & 0 & \cdots \\ 0 & \frac{\partial^{2}f}{\partial x_{2}^{2}} & 0 & 0 & \cdots \\ \Delta_{2} & 0 & 0 & \frac{\partial^{2}f}{\partial x_{3}^{2}} & 0 & \cdots \\ \Delta_{3} & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

Jak se taková matice bude chovat v bodě lokálního maxima? Očekáváme $\frac{\partial^2 f}{\partial x_i^2}(\vec{a}) < 0$ pro všechna i. Budeme teď postupně počítat levé horní subdeterminanty $i \times i$ tradičně značené Δ_i .

- První subdeterminant je dán levým horním rohem H, tedy $\Delta_1 = \frac{\partial^2 f}{\partial x_1^2}(\vec{a}) < 0$.
- \bullet Druhý subdeterminant je dán maticí 2 × 2 v levém horním rohu H,tedy $\Delta_2 = \frac{\partial^2 f}{\partial x_1^2}(\vec{a}) \cdot \frac{\partial^2 f}{\partial x_2^2}(\vec{a}) > 0.$ • Třetí subdeterminant je dán maticí 3×3 v levém horním rohu H,

tedy
$$\Delta_2 = \frac{\partial^2 f}{\partial x_1^2}(\vec{a}) \cdot \frac{\partial^2 f}{\partial x_2^2}(\vec{a}) > 0.$$

tedy
$$\Delta_3 = \frac{\partial^2 f}{\partial x_1^2}(\vec{a}) \cdot \frac{\partial^2 f}{\partial x_2^2}(\vec{a}) \cdot \frac{\partial^2 f}{\partial x_3^2}(\vec{a}) < 0.$$

Čtenář si jistě hravě domyslí, jak to funguje dál. Pro lokální maxima se stále střídají znaménka, obecný vzorec je sign $(\Delta_i) = (-1)^i$. Dá se to také vyjádřit jako $(-1)^i \Delta_i > 0$.

Pro lokální minima jsou všechny parciální derivace na diagonále kladné, a proto také všechny subdetetminanty vyjdou kladné.

Pokud nastane jiná konstelace znamének, pak maximum ani minimum nenastává. Pokud někde vyskočí nuly, pak tento rozhodovací proces neuspěl a nevíme, co se v \vec{a} děje.

To, co jsme právě odvodili pro případ diagonální H neboli nulových smíšených derivací, platí i obecně.

Věta. (Sylvestrovo kritérium)

Nechť f je definováno a má spojité derivace druhého řádu na nějakém okolí bodu \vec{a} , který je pro f stacionární, tedy $\nabla f(\vec{a}) = 0$. Nechť H je Hessova matice f v \vec{a} , nechť Δ_i jsou její levé horní subdeterminanty, $i = 1, \ldots, n$.

Jestliže $\Delta_i > 0$ pro všechna i, pak je $f(\vec{a})$ lokální minimum.

Jestliže $\Delta_1<0,\,\Delta_2>0,\,\Delta_3<0$ atd. až $(-1)^n\Delta_n>0,$ pak je $f(\vec{a})$ lokální maximum.

Vydatnější vysvětlení podáme po zavedení vhodných nástrojů v sekci 7d.

Algoritmus. (hledání a klasifikace lokálních extrémů pro $f(\vec{x})$)

1. Řešením rovnice $\nabla f(\vec{x}) = \vec{0}$ neboli soustavy

$$\frac{\partial f}{\partial x_1}(x_1, \dots, x_n) = 0$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n}(x_1, \dots, x_n) = 0$$

najdeme stacionární body.

- 2. Pro každý stacionární bod \vec{a} najdeme odpovídající Hessovu matici $H = H(\vec{a})$ a spočítáme subdeterminanty Δ_i neboli determinanty levých horních submatic velikosti $i \times i$.
- Pokud $\Delta_i > 0$ pro všechna i, je v bodě \vec{a} lokální minimum.
- \bullet Pokud se znaménka střídají $\Delta_1<0,\,\Delta_2>0,\,\Delta_3<0,\,\dots$, je v bodě \vec{a} lokální maximum. \triangle

Příklad: Najdeme a klasifikujeme lokální extrémy funkce $f(x, y, z) = 2xy^2 - 4xy + x^2 + z^2 - 2z$. Nejprve najdeme stacionární body.

$$\frac{\partial f}{\partial x} = 2y^2 - 4y + 2x = 0$$
$$\frac{\partial f}{\partial y} = 4xy - 4x = 0$$
$$\frac{\partial f}{\partial z} = 2z - 2 = 0.$$

Je to soustava tří rovnic o třech neznámých, to zní nadějně, ale rovnice nejsou lineární, tudíž celá ta krásná teorie nepomůže. Jak řešíme obecné soustavy?

My si hned všimneme, že třetí rovnice je nezávislá, tudíž jistě z=1. Co dál? Nejspolehlivější metoda je eliminační, kdy si postupně vyjadřujeme proměnné a snižujeme počet rovnic. Šlo by například vyjádřit z první rovnice $x=2y-y^2$ a dosadit do rovnice druhé, vznikla by tak rovnice třetího stupně s y jako neznámou a to se při troše štěstí a chytrém vytýkání dá umlátit (zkuste to). Vypadá to nicméně dobrodružně a je dobré znát alternativy.

Zaměříme se na druhou rovnici, kterou si přepíšeme do tvaru 4x(y-1)=0. Pokud se nám podaří vytvořit součin na straně jedné a nulu na straně druhé, bývá to velký bonus. V tomto případě například vidíme, že jsou dvě možnosti, x=0 a y=1.

V případě y=1 dává první rovnice -2+2x=0 neboli x=1 a máme první stacionární bod, (1,1,1).

V případě x = 0 první rovnice zní $2y^2 - 4y = 0$ neboli y(y - 2) = 0 a máme dvě řešení, y = 0 a y = 2. Vzniknou další dva stacionární body, (0,0,1) a (0,2,1).

Teď je třeba všechny tři stacionární body prozkoumat, tedy potřebujeme Hessovu matici. Připravíme si druhé parciální derivace, díky symetrii u smíšených stačí šest:

$$\frac{\partial^2 f}{\partial x^2} = 2, \qquad \frac{\partial^2 f}{\partial y \partial x} = 4y - 4, \qquad \frac{\partial^2 f}{\partial z \partial x} = 0,$$

$$\frac{\partial^2 f}{\partial y^2} = 4x, \qquad \frac{\partial^2 f}{\partial z \partial y} = 0, \qquad \frac{\partial^2 f}{\partial z^2} = 2,$$

Hessova matice je

$$H(x,y,z) = \begin{pmatrix} 2 & 4y-4 & 0 \\ 4y-4 & 4x & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

A jdeme na to:

Bod
$$(1,1,1)$$
: $H = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, tedy $\Delta_1 = 2$, $\Delta_2 = \det \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} = 8$ a $\Delta_3 = \det(H) = 16$.

Znaménka jsou postupně +, +, +, v bodě
$$f(1,1,1) = -2$$
 je proto lokální minimum. Bod $(0,0,1)$: $H = \begin{pmatrix} 2 & -4 & 0 \\ -4 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, tedy $\Delta_1 = 2$, $\Delta_2 = \det \begin{pmatrix} 2 & -4 \\ -4 & 0 \end{pmatrix} = -16$ a $\Delta_3 = \det(H) = -16$

-32. Znaménka jsou postupně +, -, -, v bodě f(0,0,1) = -1 proto není lokální extrém. Podle znamének to vypadá, že v některých směrech je minimum a v některých maximum v řezu, u funkce dvou proměnných bychom hovořili o sedlu.

Bod
$$(0,2,1)$$
: $H = \begin{pmatrix} 2 & 4 & 0 \\ 4 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}$, tedy $\Delta_1 = 2$, $\Delta_2 = \det \begin{pmatrix} 2 & 4 \\ 4 & 0 \end{pmatrix} = -16$ a $\Delta_3 = \det(H) = -32$.

Znaménka jsou postupně +, -, -, v bodě f(0, 2, 1) = 3 proto není lokální extrém, viz předchozí bod.

 \triangle

Příklad: Najdeme a klasifikujeme lokální extrémy funkce $f(x,y) = xy e^{x-y^2/2}$ Nejprve najdeme stacionární body.

$$\frac{\partial f}{\partial x} = y e^{x-y^2/2} + xy e^{x-y^2/2} \cdot 1 = (y+xy)e^{x-y^2/2} = 0$$

$$\frac{\partial f}{\partial y} = x e^{x-y^2/2} + xy e^{x-y^2/2} \cdot (-y) = (x-xy^2)e^{x-y^2/2} = 0.$$

Exponenciála je vždy kladná, mohou se jí rovnice vydělit a stačí řešit soustavu (1+x)y=0 a $x(1-y^2)=0$. Rovnice jsme upravili do příznivého tvaru součinu a první dává dvě možnosti. Jestliže y=0, pak z druhé rovnice x=0. Jestliže x=-1, pak z druhé rovnice $y=\pm 1$. Máme stacionární body (0,0), (-1,-1), (-1,1).

Připravíme si druhé parciální derivace:

$$\frac{\partial^2 f}{\partial x^2} = y e^{x - y^2/2} + (y + xy)e^{x - y^2/2} \cdot 1 = (x + 2)y e^{x - y^2/2},$$

$$\frac{\partial^2 f}{\partial x \partial y} = (1 + x)e^{x - y^2/2} + (y + xy)e^{x - y^2/2}(-y) = (x + 1)(1 - y^2)e^{x - y^2/2},$$

$$\frac{\partial^2 f}{\partial y^2} = -2xy e^{x - y^2/2} + (x - xy^2)e^{x - y^2/2}(-y) = xy(y^2 - 3)e^{x - y^2/2}.$$

Hessova matice je

$$H(x,y) = \begin{pmatrix} (x+2)y e^{x-y^2/2} & (x+1)(1-y^2)e^{x-y^2/2} \\ (x+1)(1-y^2)e^{x-y^2/2} & xy(y^2-3)e^{x-y^2/2} \end{pmatrix}.$$

Člen $e^{x-y^2/2}$ je vždy kladný, proto jeho vytknutí ze všech členů neovlivní znaménka determinantů, zajímá nás tedy spíše matice

$$\widehat{H}(x,y) = \begin{pmatrix} (x+2)y & (x+1)(1-y^2) \\ (x+1)(1-y^2) & xy(y^2-3) \end{pmatrix}.$$

Protože jde o funkci dvou proměnných, použijeme první algoritmus, kdy nejprve posoudíme Δ_2 .

Bod
$$(0,0)$$
: $H = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, tedy $\Delta_2 = -1 < 0$ a $f(0,0) = 0$ je sedlo.

Bod (-1,1): $H = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}$, tedy $\Delta_2 = 2 > 0$ a máme lokální extrém. Protože $\Delta_1 = 1 > 0$, je $f(-1,1) = -e^{-3/2}$ lokální minimum.

Bod (-1,-1): $H=\begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}$, tedy $\Delta_2=2>0$ a máme lokální extrém. Protože $\Delta_1=-1<0$, je $f(-1,-1)=e^{-3/2}$ lokální maximum.

4b. Globální extrémy

Definice.

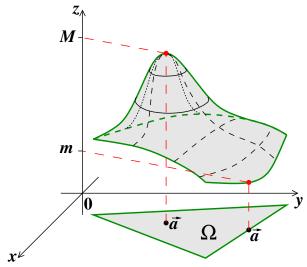
Nechť $f: D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$. Nechť $\Omega \subseteq D(f)$. Řekneme, že M je **(globální) maximum** f **na** Ω , značeno $M = \max_{\vec{x} \in \Omega} (f(\vec{x}))$, jestliže platí:

- $f(\vec{x}) \leq M$ pro všechna $\vec{x} \in \Omega$ a
- $f(\vec{a}) = M$ pro nějaké $\vec{a} \in \Omega$.

Řekneme, že m je (globální) minimum f na Ω , značeno $m=\min_{\vec{x}\in\Omega}\bigl(f(\vec{x})\bigr),$ jestliže platí:

- $f(\vec{x}) \geq m$ pro všechna $\vec{x} \in \Omega$ a
- $f(\vec{a}) = m$ pro nějaké $\vec{a} \in \Omega$.

Někdy také píšeme jen $\max_{\Omega}(f)$ a $\min_{\Omega}(f)$. Pokud nějaký bod $\vec{a} \in \Omega$ splňuje $f(\vec{a}) = \max_{\Omega}(f)$, pak říkáme, že f nabývá své maximum v \vec{a} , obdobně se nabývá minimum. Obrázek ukazuje typickou situaci.



Vysvětluje také, proč by mělo platit následující tvrzení.

Věta.

Nechť $f \colon D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$. Nechť $\Omega \subseteq D(f)$. Jestliže f nabývá své maximum nebo minimum na Ω v \vec{a} , pak f má lokální extrém v \vec{a} nebo \vec{a} leží na hranici $\partial \Omega$ množiny Ω .

Jde o zcela stejnou formulaci jako u funkcí jedné proměnné, a stejně jako u nich je to i zde východisko pro algoritmus určování globálních extrémů.

Hlavní kroky:

- 1. Najdeme kandidáty.
- 2. Porovnáme hodnoty f v těchto kandidátech a vybereme největší a nejmenší.

Druhý krok by měl být pohoda, takže se podíváme na první. Podle věty připadají v úvahu lokální extrémy a body na hranici. Lokální extrémy určovat umíme, teď nás navíc nezajímá klasifikace. Pokud totiž přidáme mezi kandidáty bod navíc, tak jej stejně vyloučíme v druhém (porovnávacím) kroku. Takže mezi kandidáty nepřidáváme přesně body lokálních extrémů, ale body, které jsou z toho podezřelé. Stačí tedy přidat všechny body z množiny Ω , kde je gradient nulový nebo neexistuje. Formálně bychom měli brát jen body z vnitřku Ω , ale zase to nemusíme řešit, přidáním pár bodů z hranice nic nezkazíme (jen si přidáme práci v druhém kroku).

S hranicí je to komplikovanější. Když jsme hledali globální extrémy pro funkce jedné proměnné, tak byla množina Ω obvykle interval a jeho hranice byly krajní body, které jsme přidali do množiny kandidátů (pokud nebyly nekonečno). U funkcí více proměnných je to úplně jinak, protože typická množina v \mathbb{R}^n má nekonečnou hranici. Například hranice čtverce v \mathbb{R}^2 je jeho obvod, hranice koule v \mathbb{R}^3 je její povrch, tedy sféra.

Rozhodně nechceme ke kandidátům přidávat nekonečně mnoho bodů, ale naštěstí nemusíme. Vezměme si nějakou hodnotu, kterou lze dosáhnout na hranici. Pokud není v rámci hodnot z hranice největší nebo nejmenší, tak také nemůže být největší nebo nejmenší vzhledem k celé množině, tudíž je zbytečné ji uvažovat. Proto když hledáme kandidáty na globální extrémy z hranice, tak se stačí zabývat pouze těmi, které jsou extrémní vzhledem k této hranici.

To znamená, že když zkoumáme hranici, tak vlastně řešíme nový podproblém, kdy hledáme globální extrémy (či přesněji body, které jsou z toho podezřelé) dané funkce, ale pouze na hranici zadané množiny. V typickém případě má tato hranice o dimenzi méně než původní množina a coby množina nižší dimenze má svůj vnitřek a hranici, takže na tento podproblém aplikujeme obecný algoritmus: Shromáždíme stacionární body z vnitřku a prozkoumáme hranici.

Například hranice čtverce se skládá ze čtyř stran. každá z nich je v zásadě jednorozměrná množina, segment, a jako taková má vnitřek a hranici, jmenovitě koncové body. Může se stát, že tato hranice hranice je zase nekonečná, pak bychom na ni vyšetřovali globální extrémy neboli přešli na podproblém podproblému. Protože pokaždé snížíme dimenzi, dříve či později musí tento proces skončit.

Příklad: Představme si nějakou funkci tří proměnných f(x, y, z). Hledáme extrémy na množině Ω , což je horní polokoule o poloměru 2, matematicky

$$\Omega = \{(x, y, z) \in \mathbb{R}^3; \ z \ge 0 \text{ a } x^2 + y^2 + z^2 \le 2^2\}.$$

Při hledání kandidátů nejprve najdeme body podezřelé z lokálního extrému a vybereme jen ty, které jsou v Ω . Pak se podíváme na hranici, což je povrch polokoule.

Ten se skládá ze dvou částí, základny a horní polosféry. Oba dva objekty jsou v zásadě dvoudimenzionální, a my u nich budeme hledat globální extrémy dané funkce, tedy přejdeme na dva podproblémy.

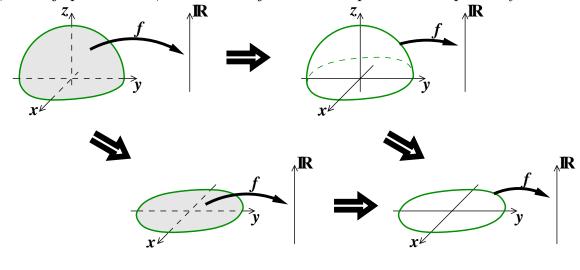
Co se týče základny, je to kruh o poloměru 2 v rovině xy, tedy opravdu dvoudimenzionální objekt. Má vnitřek, na kterém bychom hledali lokální extrémy, a hranici, jmenovitě obvod. Ten je nekonečný, takže přejdeme na další podporoblém: Hledáme globální extrémy dané funkce na obvodu základy. Je to kružnice, což je svou podstatou jednorozměrný objekt (lze jej popsat jedním

 \triangle

parametrem), a je zajímavý tím, že coby takový nemá hranici (přesněji řečeno je prázdná), a všechny body kružnice jsou jejími vnitřními body. Stačí tedy hledat lokální extrémy na kružnici.

Pokud čtenáři není zcela jasná ta záležitost s dimenzemi a prázdnou hranicí, tak se nemusí bát, uvidíme to přesně, až se později k tomuto příkladu vrátíme a jednotlivé objekty popíšeme matematicky.

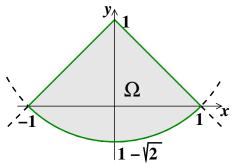
Druhý hlavní podproblém, zkoumání funkce na vrchlíku, je obdobné. Má vnitřek, jmenovitě otevřenou horní polosféru, na které budeme hledat lokální extrémy. Pak je potřeba podívat se na hranici, ale to je přesně tatáž, na kterou už jsme se dívali při zkoumání podstavy.



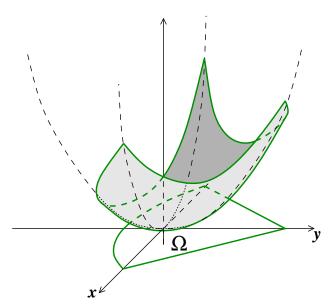
Jak vlastně vyšetřujeme funkce na množinách nižší dimenze? My ty množiny parametrizujeme (popíšeme je matematicky) a počet parametrů, které na to potřebujeme, určí dimenzi. Tyto parametry se pak také objeví v naší funkci jako nové proměnné, takže vlastně pracujeme s pomocnými funkcemi s nižším počtem proměnných. Ukážeme, jak to funguje.

Příklad: Najdeme globální extrémy funkce $f(x,y) = x^2 + y^2$ na množině Ω , což je množina vymezená křivkami y = 1 - |x| a $x^2 + (y - 1)^2 = 2$ obsahující počátek.

Nejprve identifikujeme oblast Ω . Nakreslíme křivky, první je zalomená čára, ta druhá je kružnice se středem v (0,1) o poloměru $\sqrt{2}$. Rozdělí rovinu na čtyři oblasti, my vybereme tu, která obsahuje počátek.



Nyní najdeme globální extrémy f vůči této množině. Teď budeme trochu podvádět a ukážeme, jak vlastně situace vypadá. Zajímá nás kus rotačního paraboloidu nad naší oblastí.



To bychom normálně nevěděli, budeme tedy dále předstírat, že tento obrázek nevidíme, a vyřešíme vše čistě matematicky, nicméně pro čtenáře může být užitečné si jednotlivé kroky ukazovat na tomto náčrtu.

- 1. Nejprve shromáždíme kandidáty.
- a) Lokální extrémy: Rovnice $\nabla f = \vec{0}$ znamená (2x, 2y) = (0, 0), tedy x = 0 a y = 0. Bod (0, 0) leží v oblasti Ω a je to tedy kandidát.
 - b) Extrémy na hranici: Potřebujeme prozkoumat tři části hranice.

Začneme úsečkou v kvadrantu napravo nahoře. Je to část přímky dané vzorcem y=1-x a je tedy přirozené ji popsat parametricky jako $x\mapsto (x,1-x)$ pro $x\in \langle 0,1\rangle$. Tato parametrizace potvrzuje, že tento segment je svou podstatou jednorozměrný. Pomocí vzorce y=1-x eliminujeme y v f a dostaneme pomocnou funkci ϕ jedné proměnné popisující chování f na tomto úseku hranice.

$$\phi(x) = f(x, 1 - x) = x^2 + (1 - x)^2 = 2x^2 - 2x + 1.$$

Tímto způsobem jsme převedli otázku hledání globálních extrémů funkce f na segmentu na problém hledání globálních extrémů funkce $\phi(x)$ na intervalu (0,1). To už umíme.

Nejprve najdeme kritické body z dotyčného intervalu:

$$\phi'(x) = 0 \implies 4x - 2 = 0 \implies x = \frac{1}{2}.$$

Toto x leží ve zkoumaném intervalu, je tedy platné, takže do hlavního seznamu kandidátů přidáme bod z hraničního úseku (návrat do dvourozměrné situace), který tomuto x odpovídá: $y = 1 - \frac{1}{2} = \frac{1}{2}$, tedy přidáme bod $\left(\frac{1}{2}, \frac{1}{2}\right)$.

Musíme se také podívat na hranici. Pro interval je to jednoduché, hranice jsou krajní body x=0 a x=1, takže přidáme odpovídající dvourozměrné body (0,1) a (1,0) do seznamu kandidátů. Jsou to vlastně koncové body segmentu hranice množiny Ω , který zde zkoumáme, takže to do sebe zapadá.

Všimněte si, že bod $\left(\frac{1}{2},\frac{1}{2}\right)$ není lokálním extrémem funkce f samotné. Při zkoumání úseku hranice jsme ale tuto funkci omezili pouze na dotyčnou množinu a vzhledem k této množině už je ten bod lokálním minimem, viz obrázek.

Nyní obdobně zpracujeme část hranice Ω danou vzorcem y=1+x pro $x\in\langle -1,0\rangle$. Vyšetřením funkce

$$\phi(x) = f(x, 1+x) = 2x^2 + 2x + 1$$

najdeme $x = -\frac{1}{2}$, kde by mohl být lokální extrém, a přidáme bod $\left(-\frac{1}{2}, \frac{1}{2}\right)$ do seznamu. Pohled na krajní body poskytne (-1,0) a (0,1), první do seznamu přidáme, druhý už tam je.

Zbývá třetí část hranice, ten oblouk. Vzorec $x^2 + (y-1)^2 = 2$ nabízí dvě možnosti, jak se zbavit

jedné proměnné, zkusíme $y=1-\sqrt{2-x^2}$ pro $x\in\langle -1,1\rangle$. Zajímá nás pomocná funkce

$$\phi(x) = f(x, 1 - \sqrt{2 - x^2}) = x^2 + 1 + (2 - x^2) - 2\sqrt{2 - x^2} = 3 - 2\sqrt{2 - x^2}.$$

Kde může mít extrémy na intervalu $\langle -1, 1 \rangle$? Krajní body x = -1, 1 vedou na body, které už máme v seznamu. Zbývá tedy podívat se na lokální extrémy.

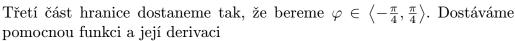
$$\phi'(x) = 0 \implies \frac{2x}{\sqrt{2 - x^2}} = 0 \implies x = 0.$$

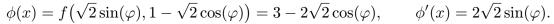
Do seznamu přidáváme poslední bod $(0, 1 - \sqrt{2})$.

Alternativa pro oblouk: Protože je to část kružnice, nabízí se polární souřadnice. Použijeme speciální verzi adaptovanou na naši situaci.

$$x = \sqrt{2}\sin(\varphi),$$

$$y = 1 - \sqrt{2}\cos(\varphi).$$





Kandidáti na extrém jsou koncové body $\varphi=\pm\frac{\pi}{4}$, dávající $(\pm 1,0)$, které už máme, a nulový bod $\varphi=0$ derivace, který dodá bod $\left(0,1-\sqrt{2}\right)$, což potvrzuje naši předchozí práci.

2. Nyní porovnáme hodnoty.

$$f(0,0) = 0, \ f(0,1) = 1, \ f(1,0) = 1, \ f(-1,0) = 1, \ f\left(\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{2}, \ f\left(-\frac{1}{2}, \frac{1}{2}\right) = \frac{1}{2}, \ f\left(0,1-\sqrt{2}\right) = 3-2\sqrt{2} \approx 0.17.$$
 Zjistili jsme, že $\min_{\Omega}(f) = 0 = f(0,0)$ a $\max_{\Omega}(f) = 1 = f(0,1) = f(\pm 1,0).$ \triangle

Co vlastně míníme pojmem "lokální extrém vzhledem k množině"? Uvažujme funkci $f(\vec{x})$ s $D(f) \subseteq \mathbb{R}^n$ a množinu $M \subseteq D(f)$. Bod $\vec{a} \in M$ je "lokální maximum vzhledem k M", jestliže existuje okolí U bodu \vec{a} takové, že $f(\vec{x}) \le f(\vec{a})$ pro všechna $\vec{x} \in U \cap M$. Relativní lokální minimum se definuje obdobně. Nedělali jsme oficiální definici, protože lidé se obvykle soustředí na speciálnější situaci.

Jmenovitě nás bude zajímat situce, kdy je množina M popsána rovnicí, popřípadě několika rovnicemi. Můžeme si představit třeba kružnici v rovině nebo rovinu v třírozměrném prostoru. Rovnicím, které množinu M definují, se říká "vazby". Abychom sjednotili značení, připomeneme, že každou (algebraickou) rovnici lze přepsat do tvaru $g(x,y,\dots)=0$. Množina M je pak jednoznačně určena vazební funkcí g, popřípadě vazebními funkcemi g_j .

$$M = {\vec{x} \in D; \ g_1(\vec{x}) = 0, \dots, g_p(\vec{x}) = 0},$$

zde D je množina, na které jsou všechny vazební funkce definovány. Velmi často je vazba jen jedna:

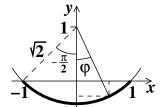
$$M = \{ \vec{x} \in D; \ g(\vec{x}) = 0 \}.$$

Co se dá od takové množiny čekat? Obecně mohou být takové množiny velmi divoké, ale v aplikacích to obvykle funguje tak, že jsme svým požadavkem na platnost rovnice g=0 vzali množině jeden stupeň volnosti, tudíž se dá očekávat, že svou podstatou půjde o množinu o dimenzi n-1, což mimo jiné znamená, že bychom ji měli být schopni popsat za pomocí n-1 parametrů.

Například vazební funkce $g(x,y)=169-x^2-y^2$ definovaná na dvoudimenzionálním prostoru \mathbb{R}^2 vede na vazbu $169-x^2-y^2=0$ neboli $x^2+y^2=169$, takže výsledná množina je kružnice o poloměru 13 se středem v počátku, což je svou podstatou jednodimenzionální objekt. Poznáme to třeba podle toho, že k určení pozice bodu na kružnici stačí jeden parametr (například úhel).

Nemusí to tak ale vyjít vždy. Třeba funkce $g(x) = x - \sqrt{x^2}$ je definována na \mathbb{R} , což je jednorozměrný svět, a definuje množinu bodů splňujících x = |x|, takže $M = \langle 0, \infty \rangle$. Tato množina je také jednodimenzionální, takže nám žádná dimenze neubyla.

46



Naopak funkce $g(x,y)=(x-1)^2+(y-13)^2$ existuje na \mathbb{R}^2 a definuje množinu bodů splňujících vazbu $(x-1)^2+(y-13)^2=0$, což ale zvládne jen jeden bod, jmenovitě (1,13). Tato množina tedy není jednorozměrná, ale má dimenzi nula, takže jsme svým požadavkem ztratili dokonce dvě dimenze.

Ovšem takové příklady si vymýšlejí zlomyslní matematici, v aplikacích obvykle funguje to ztrácení jedné dimenze.

Zajímavé to začne být, když se objeví více vazebních funkcí g_j . Pokud požadujeme současnou platnost všech rovností $g_j = 0$, pak je výsledná množina M vlastně průnik $\bigcap M_j$ množin M_j definovaných jednotlivými vazebními funkcemi g_j . V typickém případě má každá z množin M_j dimenzi n-1 (pokud předpokládáme, že se to vše děje v prostoru \mathbb{R}^n), takže člověk by čekal, že při každém novém průniku by mohlo dojít k odebrání jedné dimenze a při p vazbách by odpovídající průnik neboli množina M mohl mít dimenzi n-p. Čtenář by mohl něco takového znát z lineární algebry. Lineární funkce $g = a_1x_1 + \cdots + a_nx_n + d$ vede na rovnici, která definuje nadrovinu v \mathbb{R}^n , tedy útvar dimenze n-1. Pronikáním různoběžných rovin pak postupně ubíráme dimenze, díky čemuž jsme schopni dostat přímky, tedy jednorozměrné útvary. Například přímka v \mathbb{R}^3 je dána dvěma rovnicemi.

Ale i v té lineární algebře víme, že se toto může pokazit, například průnik dvou rovnoběžných a různých rovin je prázdný, u obecných útvarů se dají čekat i komplikovanější situace, což je pro teorii nepříjemné. My chceme pracovat se situací, kdy má výsledná množina dimenzi n-p, a proto ve větě, která brzy přijde, budeme muset přidat technický předpoklad o vazebních funkcích g_j .

Než se tam dostaneme, popíšeme problém, který chceme řešit. Máme funkci $f(\vec{x})$ a množinu M určenou nějakými vazbami. Hledáme lokální extrémy vzhledem k této množině. Takovýmto lokálním extrémům se říká **vázané extrémy**. V příkladě výše jsme tento problém řešili pomocí eliminace proměnných (či parametrického popisu množiny), ale to u nespolupracujících vazeb může vést ke komplikovaným výpočtům. Následující věta nabízí alternativní přístup.

Věta. (Lagrangeovy multiplikátory)

Nechť D je otevřená množina v \mathbb{R}^n a $f \in C^1(D)$.

Uvažujme vazební funkce $g_1, \ldots, g_p \in C^1(D)$, označme

$$M = \{ \vec{x} \in D; \ g_1(\vec{x}) = 0, \dots, g_p(\vec{x}) = 0 \}$$

a předpokládejme, že vektory $\nabla g_1(\vec{a}), \ldots, \nabla g_p(\vec{a})$ jsou lineárné nezávislé pro všechny body $\vec{a} \in M$.

Jestliže je \vec{x}_0 lokální extrém funkce f vzhledem k množině M, pak existují čísla $\lambda_1, \ldots, \lambda_p$ taková, že

$$\nabla f(\vec{x}_0) = \sum_{j=1}^p \lambda_j \nabla g_j(\vec{x}_0).$$

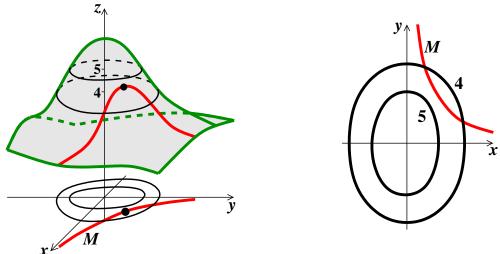
Podmínky na gradienty g_j zajistí, že tyto vazby vytvoří rozumnou množinu. Číslům λ_j se říká Lagrangeovy multiplikátory. Z hlediska hledání matematických extrémů vystupují jako čistě pracovní parametry a obvykle se jich hned zbavujeme eliminací. Stojí nicméně za zmínku, že mají praktický význam ve fyzice. Tam se pracuje s tzv. Lagrangeovou funkcí

$$L = f - \sum_{j=1}^{p} \lambda_j g_j$$

a závěr věty vlastně říká, že $\nabla L = \vec{0}$.

Proč by něco takového mělo platit? Pro případ jedné vazební funkce se to dá docela pěkně nakreslit.

Na obrázku vidíme graf nějaké funkce f, nás ovšem zajímají pouze hodnoty, kterých se nabývá na množině M vyznačené v rovině xy neboli v definičním oboru f. Tyto hodnoty jsme také vyznačili červeně v grafu funkce. Odhad největší možné hodnoty, kterou lze v rámci M nabýt, jsme zaznačili jak na grafu, tak v samotné množině M, ale to je spíše jen pro ilustraci. Mnohem důležitější jsou černé křivky, které ukazují na grafu f hodnoty f a f z nich vyčteme, že hodnota hledaného maxima na f je někde mezi f a f z není to správné místo, kde máme hledat informace, protože s grafy se nepracuje snadno. Přenesme proto pozornost do definičního oboru, kde zcela přirozeně přejdeme k hladinám konstantnosti, které odpovídají hodnotám f a f pro funkci f. Dokážeme něco rozpoznat čistě z pohledu na situaci v definičním oboru?



Na první pohled vidíme, že funkce v rámci množiny M nikdy nedosáhne hodnoty 5, protože množina M má nulový průnik s množinou těch \vec{x} , pomocí kterých se k hodnotě 5 dostaneme. Vidíme také, že hodnotu 4 získat umíme, dokonce pomocí dvou různých \vec{x} . Klíčová otázka ovšem je, zda z obrázku poznáme, že 4 není největší možná hodnota, tedy není to to, co hledáme. Zdálo by se, že důležitou roli bude hrát fakt, že těch průniků je více, ale to není ono. Podstatné je, že v bodě průniku se hladina konstantnosti H_4 a množina M protínají pod nenulovým úhlem (měřeno jako úhel jejich tečen).

Souvisí to s faktem, že hladiny konstantnosti H_c se chovají spojitě vzhledem k hodnotě c, pokud je zkoumaná funkce f spojitá. Znamená to, že když změníme c jen nepatrně, tak se také množina H_c v prostoru \mathbb{R}^n změní jen málo. Dobrá představa je, že když spojitým pohybem otáčíme ovládací knoflík c z hodnoty 4 na hodnotu 5, tak se odpovídající hladiny konstantnosti plynule přesouvají od H_4 k H_5 , asi jako populární morphing. Co to znamená pro nás? Jestliže se množina M a hladina konstantnosti H_4 protínají pod nenulovým úhlem, pak při malé změně konstanty c budou nové hladiny konstantnosti stále protínat množinu M. Z toho pak vyplývá, že hodnota c=4 nemůže být ani maximální, ani minimální.



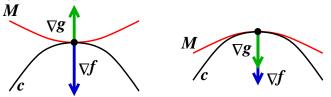
Jak tedy poznáme, že jsme našli maximum či minimum? Příslušná hladina konstantnosti se musí s množinou M protínat tak, aby měli stejné tečny. Protože se blížíme k cíli, je potřeba zkusit toto pozorování naformulovat tak, aby fungovalo i ve více dimenzích. Pokud máme funkci definovanou na \mathbb{R}^3 , pak hladiny konstantnosti jsou plochy ve 3D, stejně tak i množina M. Aby se neprotínaly šikmo, musí být tečně roviny totožné. Podobně přejdeme i k vyšším dimenzím.

Jak tedy obecně poznáme, že tečné nadroviny dvou (n-1)-rozměrných ploch jsou totožné? Nejsnadněji pomocí normálových vektorů, které v tom případě musí být rovnoběžné. Jsme skoro v cíli. Jak najdeme normálové vektory k takové ploše? Co se týče hladiny konstantnosti, tam máme k

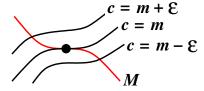
 \triangle

dispozici tvrzení z kapitoly 3, normálovým vektorem je přímo gradient $\nabla f.$ Jak je na tom množina M?

Je to množina bodů, které splňují $g(\vec{x}) = 0$, takže je to také hladina konstantnosti, tentokráte pro funkci g. Normálový vektor množiny M je tedy ∇g . Došli jsme k následujícímu závěru. Aby měla funkce f maximum či minimum s hodnotou c vzhledem k množině M v nějakém bodě \vec{a} , tak se v bodě \vec{a} musí množina M protínat s hladinou konstantnosti H_c a vektory $\nabla f(\vec{a})$, $\nabla g(\vec{a})$ musejí být rovnoběžné, což namená, že pro nějakou nenulovou konstantu λ platí $\nabla f(\vec{a}) = \lambda \nabla g(\vec{a})$. Přesně to tvrdí věta o Lagrangeových multiplikátorech.



Pro úplnost poznamenejme, že i tato věta je jen implikace, obdobně jako věta o kritických bodech pro lokální extrémy. Může se stát, že v nějakém bodě \vec{a} jsou $\nabla f(\vec{a})$ a $\nabla g(\vec{a})$ rovnoběžné, ale vázaný lokální extrém tam není.



V praxi nás to netrápí, protože stejně jen shromažďujeme kandidáty, a pokud si přidáme nějakého kandidáta navíc, tak jej v druhém, porovnávacím kole stejně vyřadíme. Hlavní význam věty je, že když podle ní hledáme kandidáty, tak žádného nevynecháme.

Doufejme, že nám teď věta dává smysl z hlediska matematického, další dobrá otázka zní, zda dává smysl z hlediska praktického. Podívejme se na příklad.

Příklad: Vrátíme se k příkladu, kde jsme hledali extrémy funkce $f(x,y) = x^2 + y^2$. Jmenovitě se podíváme na fázi, kdy jsme hledali extrémy f vzhledem k množině dané vzorcem $x^2 + (y-1)^2 = 2$. V prvním řešení jsme tuto vazbu použili k eliminaci jedné proměnné a vyšetřovali jednorozměrnou situaci.

Teď zkusíme Lagrangeovy multiplikátory. Vazbu si upravíme jako $x^2 + (y-1)^2 - 2 = 0$. Máme tedy vazební funkci $g(x,y) = x^2 + (y-1)^2 - 2$ a sepíšeme si rovnice, které vyplývají z Lagrangeovy podmínky a vazby.

$$\nabla f(x,y) = \lambda \nabla g(x,y) \implies 2x = 2x\lambda$$

$$g(x,y) = 0 \implies 2y = 2(y-1)\lambda$$

$$x^2 + (y-1)^2 = 2.$$

Dostali jsme tři rovnice pro tři neznámé, což zní nadějně, ale bohužel nejsou lineární. Naštěstí se nabízejí zajímavé možnosti, například první rovnice vyloženě volá po zkrácení. Pak dostaneme $1=\lambda$. Druhá rovnice pak říká y=y-1, což nemá řešení. Toto je tedy slepá ulička, ale není to ještě konec. Krátit totiž lze pouze za předpokladu, že $x\neq 0$, je tedy třeba ještě prozkoumat tu druhou možnost.

Pokud x=0, pak je první rovnice automaticky splněna. Dosazením do vazební rovnice dostáváme $(y-1)^2=2$. Pak $y=1\pm\sqrt{2}$, ovšem ze zadání je jasné, že nás zajímají jen hodnoty $y\leq 0$. Proto nás zajímá jen $y=1-\sqrt{2}$. Vzhledem k tomu, že λ nepotřebujeme, tak by se zdálo, že máme hotovo, ale ještě musíme ověřit, že druhou rovnici lze splnit, což se povede pro $\lambda=\frac{\sqrt{2}-1}{\sqrt{2}}$.

Dostali jsme podezřelý bod $(0,1-\sqrt{2})$, přesně jako v předchozím výpočtu. Bylo to snažší? To musí posoudit čtenář.

Jsme připravení porovnat strategie. Uvažujme funkci f, která má n proměnných, a p nezávislých vazebních funkcí q taktéž o n proměnných, kde p < n. Naše předchozí (intuitivní) strategie spočí-

[4b] 49

vala v tom, že při troše štěstí je možné si pomocí vazebních rovnic vyjádřit některé proměnné v závislosti na jiných, což po dosazení do f vede na vznik pracovní funkce ϕ o n-p proměnných. U ní pak hledáme extrémy, což mimo jiné vede na podmínku $\nabla \phi = \vec{0}$, tedy soustavu n-p rovnic o n-p neznámých.

Pokud bychom následovali postup z naší nové věty, tak bychom si uchovali původní proměnné a ještě navrch přidali λ_j , což dává celkem n+p proměnných. Kolik máme rovnic? Rovnost $\nabla f = \sum \lambda_j \nabla g_j$ porovnává vektory, jde tedy o rovnost všech složek obou vektorů. Tím vzniká n rovnic, plus ještě pořád máme vazby $g_j = 0$, tedy dostáváme celkem n+p rovnic

$$\frac{\partial f}{\partial x_1}(x_1, \dots, x_n) = \lambda_1 \frac{\partial g_1}{\partial x_1}(x_1, \dots, x_n) + \dots + \lambda_p \frac{\partial g_p}{\partial x_1}(x_1, \dots, x_n)$$

$$\vdots$$

$$\frac{\partial f}{\partial x_n}(x_1, \dots, x_n) = \lambda_1 \frac{\partial g_1}{\partial x_n}(x_1, \dots, x_n) + \dots + \lambda_p \frac{\partial g_p}{\partial x_n}(x_1, \dots, x_n)$$

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

$$\vdots$$

$$g_p(x_1, \dots, x_n) = 0.$$

Dobrá zpráva je, že počet rovnic souhlasí s počtem neznámých, špatná zpráva je, že máme o 2p rovnic a 2p neznámých víc než při předchozím postupu. Proč bychom tedy měli chtít jít touto novou cestou? Záleží na tom, jak komplikovaný vzorec jsme předtím dostali při vyjadřování proměnných z vazeb $g_j=0$ a jak pěkně (či škaredě) zapadly do funkce f. Pokud vznikly nepříjemné výrazy, pak by Lagrangeův přístup mohl pomoci a často také pomůže. Tento přístup je velmi oblíbený v aplikacích a dokonce nám poskytne některé obecné výsledky a důkazy vět.

Příklad: Najdeme vzdálenost bodu $P = (x_0, y_0, z_0)$ od roviny q: ax + by + cz + d = 0.

Tato vzdálenost se najde jako vzdálenost mezi P a bodem Q=(x,y,z) z roviny q, který je k P nejblíže. Hledáme tedy minimum funkce

$$(x, y, z) \mapsto \operatorname{dist}(P, (x, y, z)) = \sqrt{(x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2}$$

pro (x, y, z) z dané roviny. Protože odmocnina je rostoucí funkce, stačí hledat minimum funkce

$$f(x, y, z) = (x - x_0)^2 + (y - y_0)^2 + (z - z_0)^2,$$

a to na množině dané podmínkou ax + by + cz + d = 0. Budeme tedy pracovat s vazební funkcí g(x, y, z) = ax + by + cz + d.

Podle věty se hledaný bod skrývá mezi řešeními soustavy vzniklé z Lagrangeovy podmínky a dané vazby

$$\nabla f = \lambda \nabla g$$
$$g = 0$$

neboli

$$2(x - x_0) = \lambda \cdot a$$
$$2(y - y_0) = \lambda \cdot b$$
$$2(z - z_0) = \lambda \cdot c$$
$$ax + by + cz + d = 0.$$

Rovnice jsou lineární, ale standardní přístup Gaussovou eliminací či Kramerovým pravidlem necháme zvědavému čtenáři a zaměříme se na postupy obvyklé u Lagrangeových multiplikátorů.

Vzhledem k tomu, že λ nepotřebujeme, můžeme se jej zbavit eliminací. Předpokládejme, že $a \neq 0$ (alespoň jeden z koeficientů a, b, c není nula, jinak nemáme rovnici roviny). Pak z první rovnice

 $\lambda = \frac{2}{a}(x-x_0)$ a první tři rovnice se redukují na

$$2(y - y_0) = b\frac{2}{a}(x - x_0) \implies y = y_0 + \frac{b}{a}(x - x_0)$$
$$2(z - z_0) = c\frac{2}{a}(x - x_0) \implies z = z_0 + \frac{c}{a}(x - x_0).$$

Dosadíme do vazby:

$$ax + b(y_0 + \frac{b}{a}(x - x_0)) + c(z_0 + \frac{c}{a}(x - x_0)) + d = 0$$

$$\implies x = \frac{-ad + (a^2 + b^2 + c^2)x_0 - a(ax_0 + by_0 + cz_0)}{a^2 + b^2 + c^2}$$

$$\implies x = x_0 - a\frac{ax_0 + by_0 + cz_0 + d}{a^2 + b^2 + c^2}$$

$$\implies y = y_0 - b\frac{ax_0 + by_0 + cz_0 + d}{a^2 + b^2 + c^2}, z = z_0 - c\frac{ax_0 + by_0 + cz_0 + d}{a^2 + b^2 + c^2}.$$

Toto je nejbližší bod Q roviny q k bodu P. Než spočítáme žádanou vzdálenost, podíváme se na další zajímavou možnost, jak vyřešit soustavu. Z prvních tří rovnic si vyjádříme x, y, z jako funkce Lagrangeova multiplikátoru:

$$x = x_0 + \frac{a}{2}\lambda$$
$$y = y_0 + \frac{b}{2}\lambda$$
$$z = z_0 + \frac{c}{2}\lambda.$$

Toto dosadíme do vazby a dostaneme

$$a(x_0 + \frac{a}{2}\lambda) + b(y_0 + \frac{b}{2}\lambda) + c(z_0 + \frac{c}{2}\lambda) + d = 0$$

$$\implies \lambda = -2\frac{ax_0 + by_0 + cz_0 + d}{a^2 + b^2 + c^2}.$$

Takže

$$x = x_0 - a \frac{ax_0 + by_0 + cz_0 + d}{a^2 + b^2 + c^2}$$
$$y = y_0 - b \frac{ax_0 + by_0 + cz_0 + d}{a^2 + b^2 + c^2}$$
$$z = z_0 - c \frac{ax_0 + by_0 + cz_0 + d}{a^2 + b^2 + c^2}.$$

Takto to možná bylo snažší.

Oba postupy byly úspěšné a musely vést ke stejnému bodu Q. Při výpočtu vzdálenosti se hodí nejprve používat λ ,

$$dist(P,Q) = \sqrt{(x-x_0)^2 + (y-y_0)^2 + (z-z_0)^2}$$

$$= \sqrt{\left(\frac{a}{2}\lambda\right)^2 + \left(\frac{b}{2}\lambda\right)^2 + \left(\frac{c}{2}\lambda\right)^2} = \sqrt{a^2 + b^2 + c^2} \frac{1}{2}|\lambda|$$

$$= \frac{|ax_0 + by_0 + cz_0 + d|}{\sqrt{a^2 + b^2 + c^2}}.$$

Toto je standardní vzorec pro vzdálenost bodu od roviny: Bod dosadíme do rovnice roviny, aplikujeme absolutní hodnotu a vydělíme velikostí normálového vektoru spojeného s rovnicí roviny.

Šťouravá otázka: Jak víme, že to je minimum a ne maximum? Sofistikovaná odpověď by poukázala na to, že f je konvexní funkce. Na této úrovni ale nejsme, zkusíme to jinak. Pokud bereme body Q roviny utíkající někam do nekonečna, tak se i vzdálenost od P bude zvětšovat do nekonečna, tudíž číslo nalezené výše nemůže být maximem. Mohlo by to ale být lokálním maximem. jenže naše funkce je spojitá, omezená zdola a "na okrajích" jde do nekonečna, takže musí mít lokální minimum. My jsme našli jen jeden podezřelý bod, takže to je on. Není to matematicky přesné, ale

úplná odpopvěď by byla dost dlouhá.

 \triangle

Tento příklad potvrdil, že Lagrangeovy multiplikátory umí nabídnout obecné výsledky, a také ukázal dvě nejčastější strategie pro řešení rovnic, které dostaneme aplikací Lagrangeova multiplikátoru.

- \bullet Velmi oblíbený přístup je použít rovnice vzniklé z gradientní rovnosti k eliminaci parametrů λ_j (stejně nás nezajímají) a nalezení dalších vazeb mezi proměnnými. Pomocí těchto a také původních vazeb pak zjistíme podezřelé body.
- Někdy se podaří z prvních rovnic (vzniklých z gradientu) vyjádřit proměnné pomocí parametru λ . Po dosazení do vazebné rovnice získáme hodnotu λ , následně pak hodnoty proměnných. To už ale musíme mít štěstí.

Pokud ani jeden z těchto přístupů nevyjde, tak prostě musíme zkusit něco jiného, podle tvaru rovnic, kterým čelíme.

Už jsme potkali všechny klíčové nástroje, můžeme tedy zformulovat algoritmus pro hledání globálních extrémů, který bude spíš rámcový a rekurentní.

Algoritmus. (Vyšetřování globálních extrémů $f(\vec{x})$ na Ω)

- 1. Připravíme seznam kandidátů $\{x_1, \ldots, x_N\}$.
- a) Najdeme body z Ω , ve kterých $\nabla f = \vec{0}$ nebo ∇f neexistuje.
- b) Najdeme body z hranice $\partial\Omega$, ve kterých může f nabývat globálních extrémů vůči $\partial\Omega$. Přitom opět aplikujeme základní algoritmus, tedy kroky a) a b), na množinu $\partial\Omega$.

K hledání lokálních extrémů na vnitřku $\partial\Omega$ lze použít Lagrangeovy multiplikátory, pokud je tato množina zadána pomocí vazebních podmínek, druhou obvyklou možností je použití parametrizace. Pokud je hranice množiny $\partial\Omega$ nekonečná, tak k jejímu zkoumání použijeme rekurzivně krok b). Tato redukce by měla v typickém případě skončit zkoumáním funkce jedné proměnné na uzavřeném omezeném intervalu, kde hranici tvoří jen krajní body, které zahrneme mezi kandidáty.

2. Porovnáme hodnoty $f(x_1), \ldots, f(x_N)$, vybereme mezi nimi největší a nejmenší.

 \triangle

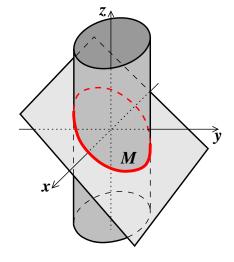
Teď už je čas podívat se na pořádný příklad.

Příklad: Najdeme globální extrémy funkce f(x, y, z) = xyz vzhledem k podmínkám $x^2 + y^2 = 3$ a y + z = 0.

Nejprve si jen ze zvědavosti rozmyslíme, na jaké množině M vlastně hledáme naše extrémy. Rovnice $x^2+y^2=3$ by v \mathbb{R}^2 definovala kružnici o poloměru $\sqrt{3}$, ale my zde pracujeme v \mathbb{R}^3 . Podmínka nijak neomezuje z, takže se ta kružnice může posouvat ve směru osy z dle libosti a vytvoří nekonečný válec.

Rovnice y+z=0 definuje přímku v rovině yz. Prochází počátkem a klesá, je to vlastně vedlejší diagonála z=-y. Tato přímka se pak volně posouvá ve směru osy x a vytvoří rovinu. Dobrá vizuální představa je, že vezmeme rovinu xz a otočíme ji kolem osy x o 45 stupňů.

Když válec a rovinu protneme, dostaneme elipsu, což je svou podstatou jednodimenzionální útvar. To souhlasí, dvěma podmínkami jsme odebrali dva rozměry z původně třírozměrného prostoru. Uvedli jsme si, že to tak nemusí vyjít vždy, zde máme štěstí.



Tato elipsa coby jednodimenzionální objekt nemá hranici, takže globální extrémy na ní musí nastat v bodech lokálních extrémů. Najdeme je pomocí věty o Lagrangeových multiplikátorech. Nejprve ověříme, že jsou splněny její předpoklady. Máme dvě vazební funkce, $g(x, y, z) = x^2 + y^2 - 3$

a h(x, y, z) = y + z, které definují množinu M. Věta chce, abychom ověřili, že na této množině jsou gradienty g a h lineárně nezávislé. Tak schválně:

$$\nabla g = (2x, 2y, 0),$$
$$\nabla h = (0, 1, 1).$$

Podle třetí souřadnice vidíme, že vektor ∇h nelze získat jako násobek vektoru ∇g , takže máme nezávislost. Ještě si všimneme, že všechny zůčastněné funkce mají spojité derivace prvního řádu, a víme, že máme právo větu použít. Zavedeme dva Largangeovy multiplikátory, λ a μ (nechce se mi psát indexy), a píšeme rovnice.

$$yz = \lambda \cdot 2x + \mu \cdot 0$$

$$\nabla f = \lambda \nabla g + \mu \nabla h$$

$$yz = \lambda \cdot 2x + \mu \cdot 0$$

$$xz = \lambda \cdot 2y + \mu \cdot 1$$

$$yz = \lambda \cdot 2x + \mu \cdot 0$$

$$xz = \lambda \cdot 2y + \mu \cdot 1$$

$$xy = \lambda \cdot 0 + \mu \cdot 1$$

$$x^2 + y^2 = 3$$

$$y + z = 0$$

Zkusíme doporučený postup a použijeme první tři rovnice k eliminaci multiplikátorů.

$$yz = 2\lambda x$$

$$xz = 2\lambda y + \mu \implies yz = 2\lambda x$$

$$xy = \mu$$

$$xz = 2\lambda y + xy.$$

Teď bychom rádi vyjádřili λ z první rovnice, ale nejprve se musíme zeptat: Co když x=0? Pak z této rovnice dostáváme yz=0, ale druhá vazebná podmínka y+z=0 dává y=-z, což vede na y=z=0. Dostáváem bod (0,0,0), ale ten nesplňuje první vazebnou podmínku, takže toto není platný kandidát.

Pokračujeme tedy za předpokladu $x \neq 0$. Pak z první rovnice vyjádříme 2λ a dosadíme do druhé:

Eliminací parametrů jsme získali vazbu mezi proměnnými, teď ji spojíme s vazbami ze zadání. Z druhé z nich dosadíme z=-y:

$$-x^2y = -y^3 + x^2y \implies y^3 = 2x^2y.$$

Je možné, aby y=0? Pak i z=0 a z první vazebné podmínky vyjde $x=\pm\sqrt{3}$. Máme první dva kandidáty.

Pokud $y \neq 0$, pak zkrátíme a dostaneme $y^2 = 2x^2$. Dosadíme do první vazební podmínky a dostáváme

$$x^2 + 2x^2 = 3 \implies x = \pm 1, \ y = \pm \sqrt{2}.$$

Dostáváme další čtyři body, kde třetí souřadnici doplníme pomocí z = -y.

Víc odboček k prozkoumání už není, seznam je hotov. Nyní porovnáme hodnoty pro kandidáty. $f(\pm\sqrt{3},0,0)=0$,

$$\begin{split} f\left(1,\sqrt{2},-\sqrt{2}\right) &= -2, \ f\left(1,-\sqrt{2},\sqrt{2}\right) = -2, \ f\left(-1,\sqrt{2},-\sqrt{2}\right) = 2, \ f\left(-1,-\sqrt{2},\sqrt{2}\right) = 2. \\ \text{Závěr: } \max_{\Omega}(f) &= 2 = f\left(-1,\sqrt{2},-\sqrt{2}\right) = f\left(-1,-\sqrt{2},\sqrt{2}\right) \\ \text{a } \min_{\Omega}(f) &= -2 = f\left(1,\sqrt{2},-\sqrt{2}\right) = f\left(1,-\sqrt{2},\sqrt{2}\right). \end{split}$$

Alternativa: Co kdybychom větu o Lagrangeových multiplikátorech neznali? Pak bychom mohli použít vazby k redukci počtu proměnných. Pokud se rozhodneme použít x jako parametr, pak $y=\pm\sqrt{3-x^2}$, zjevně $-\sqrt{3} \le x \le \sqrt{3}$.

Začneme verzí $y = \sqrt{3-x^2}$. Pak $z = -y = -\sqrt{3-x^2}$ a potřebujeme vyšetřit pomocnou funkci $\phi(x) = f\left(x, \sqrt{3-x^2}, -\sqrt{3-x^2}\right) = x\sqrt{3-x^2} \cdot \left(-\sqrt{3-x^2}\right) = -x(3-x^2) = x^3 - 3x$.

Máme najít globální extrémy na intervalu $\langle -\sqrt{3}, \sqrt{3} \rangle$. Rovnou přidáme do seznamu kandidátů krajní body $x = \pm \sqrt{3}$, vzhledem k původnímu zadání tedy vlastně body $(\pm \sqrt{3}, 0, 0)$.

Zbývá najít lokální extrémy:

$$\phi'(x) = 0 \implies 3x^2 - 3 = 0 \implies x = \pm 1.$$

Dostáváme body $(1, \sqrt{2}, -\sqrt{2})$ a $(-1, \sqrt{2}, -\sqrt{2})$.

Obdobně pak vyšetříme verzi $y=-\sqrt{3-x^2},\ z=-y=\sqrt{3-x^2},\ \text{což dá body } \left(1,-\sqrt{2},\sqrt{2}\right)$ a $\left(-1,-\sqrt{2},\sqrt{2}\right)$.

Tento postup je tedy možný a docela příjemný.

 \triangle

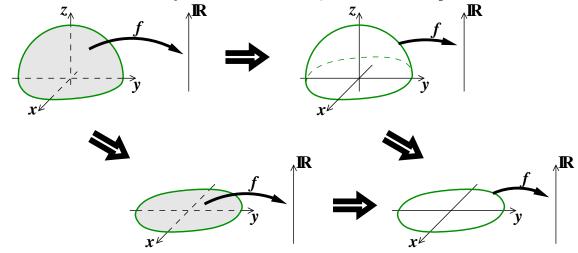
Tento příklad pěkně ukázal, jak pracujeme s dvěma vazbami, ale možná to není nejlepší reklama na Lagrangeovy multiplikátory. Kapitolu tedy uzavřeme pořádným příkladem.

Příklad: Najdeme globální extrémy funkce $f(x, y, z) = xy + (z - 1)^2$ na množině

$$\Omega = \{(x,y,z) \in \mathbb{R}^3; \ z \geq 0 \ {\rm a} \ x^2 + y^2 + z^2 \leq 2^2 \}$$

neboli na horní polokouli o průměru 2.

Již jsme si situaci trochu rozmysleli včetně obrázku, teď to uděláme pořádně.



- 1. Sběr kandidátů.
- a) Lokální extrémy: Množina Ω má neprázdný vnitřek, tedy jde o plně n-dimenzionální množinu, proto lokální extrémy vůči této množině odpovídají klasickým lokálním extrémům. Rovnice $\nabla f = \vec{0}$ vede na soustavu

$$y = 0$$

$$x = 0 \implies x = 0, y = 0, z = 1.$$

$$2(z - 1) = 0$$

Bod (0,0,1) leží v naší polokouli, je to tedy validní kandidát.

b) Extrémy na hranici.

Hranice polokoule je její povrch a skládá se ze dvou částí, vrchlíku neboli horní polosféry a kruhové základy.

b1) Hledáme extrémy funkce f na vrchlíku neboli horní polosféře. Jde vlastně o novou podúlohu, kterou vyřešíme standardním postupem, tedy prozkoumáme chování f uvnitř této množiny a také na její hranici.

Lokální extrémy: Hledáme extrémy funkce $f(x, y, z) = xy + (z-1)^2$ vzhledem k vazební podmínkce $x^2 + y^2 + z^2 - 4 = 0$, tedy s vazební funkcí $g(x, y, z) = x^2 + y^2 + z^2 - 4$. Použijeme Lagrangeovy

multiplikátory. Sestavíme příslušné rovnice:

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) \implies \begin{cases} y = 2x\lambda \\ x = 2y\lambda \\ 2(z - 1) = 2z\lambda \end{cases}$$

$$x^{2} + y^{2} + z^{2} = 4.$$

Je možné použít oblíbenou strategii a rovnou eliminovat parametr λ , když jej stejně nepotřebujeme? Rádi bychom jej vyjádřili z první rovnice, na to ale potřebujeme $x \neq 0$. Pojďme tedy nejprve prozkoumat případ x = 0.

Z první rovnice pak máme y=0, vazba pak poskytne $z=\pm 2$, nás ale zajímá jen $z\geq 0$. Ještě ověříme, že pro $\lambda=\frac{1}{2}$ platí i třetí rovnice, a dostáváme podezřelý bod (0,0,2), první do seznamu.

Pokud $x \neq 0$, pak z první rovnice $\lambda = \frac{y}{2x}$. Dosadíme do druhé rovnice a dostáváme $x = \frac{2y^2}{2x}$ neboli $x^2 = y^2$. Takže $y = \pm x$ a proto $\lambda = \pm \frac{1}{2}$.

Případ y=x a $\lambda=\frac{1}{2}$: Pak třetí rovnice říká 2(z-1)=z neboli z=2. Pak z vazby $x^2+x^2+2^2=4$ vyjde x=y=0, tento bod už máme v seznamu.

Případ y=-x a $\lambda=-\frac{1}{2}$: Pak třetí rovnice říká 2(z-1)=-z neboli $z=\frac{2}{3}$. Dosazením tohoto a y=-x do vazební rovnice máme $2x^2+\frac{4}{9}=4$ neboli $x=\pm\frac{4}{3}=-y$. Dostáváme podezřelé body $\left(\frac{4}{3},-\frac{4}{3},\frac{2}{3}\right)$ a $\left(-\frac{4}{3},\frac{4}{3},\frac{2}{3}\right)$.

Teď se musíme podívat na hranici vrchlíku. Je to vlastně kružnice v rovině xy o poloměru 2. Pokud bychom to chtěli dělat formálně, mohli bychom to vnímat jako křivku v \mathbb{R}^3 , která je dána vazbami $x^2 + y^2 = 4$ a z = 0 neboli vazebními funkcemi $g(x, y, z) = x^2 + y^2 - 4$ a h(x, y, z) = z. Sestavíme Lagrangeovské rovnice:

$$y = 2x\lambda + 0 \cdot \mu \qquad y = 2x\lambda$$

$$\nabla f(x, y, z) = \lambda \nabla g(x, y, z) + \mu \nabla h(x, y, z) \qquad x = 2y\lambda + 0 \cdot \mu \qquad x = 2y\lambda$$

$$g(x, y, z) = 0 \qquad \Longrightarrow 2(z - 1) = 0 \cdot \lambda + 1 \cdot \mu \implies 2(z - 1) = \mu$$

$$h(x, y, z) = 0 \qquad x^2 + y^2 = 4 \qquad z = 0$$

Z toho z=0 hned máme $\mu=2$, což nám k ničemu není, důležité je, že třetí rovnice je splněna. Obvykle rádi eliminujeme parametry, dá se vyjádřit λ z první rovnice? Co když x=0? Pak z té rovnice také y=0, ale tyto hodnoty nesplňují vazební rovnici kružnice, tudíž toto nemůže v naší soustavě nastat. Pak z první rovnice $\lambda=\frac{y}{2x}$. Dosazením do druhé rovnice získáme $x=\frac{2y^2}{2x}$ neboli $x^2=y^2$ neboli $x=\pm y$. Dosadíme do vazební podmínky a dostaneme $2x^2=4$, tedy $x=\pm\sqrt{2}$. Dostáváme čtyři podezřelé body $(\pm\sqrt{2},\pm\sqrt{2},0)$.

Čtenáře možná napadlo, že jsme souřadnici z vůbec nepoužili, takže šlo v podstatě o rovinný problém. Ano, mohli jsme zavést pomocnou funkci

$$\phi(x,y) = f(x,y,0) = xy + 1$$

a hledat její extrémy vzhledem k vazbě $x^2+y^2=4$. Pak bychom pracovali s vazební funkcí $g(x,y)=x^2+y^2-4$ a se soustavou

$$\nabla \phi(x,y) = \lambda \nabla g(x,y) \implies \begin{aligned} y &= 2x\lambda \\ g(x,y) &= 0 \end{aligned} \implies \begin{aligned} x &= 2y\lambda \\ x^2 + y^2 &= 4, \end{aligned}$$

což jsou přesně ty rovnice, které nám vyprodukovaly body $(\pm\sqrt{2},\pm\sqrt{2},0)$.

b2) Nyní hledáme extrémy funkce f na základně neboli na kruhu v rovině xy daném podmínkou $x^2 + y^2 \le 4$. Jde o v podstatě dvourozměrnou úlohu, proto zavedeme pracovní funkci

$$\phi(x, y) = f(x, y, 0) = xy + 1.$$

Hledáme její extrémy na kruhu. Opět jde o podproblém, který řešíme jako samostatnou úlohu, tedy zkoumáme dva zdroje kandidátů, uvnitř a na hranici.

Hranice našeho kruhu je kružnice, kterou jsme již zkoumali, tedy to už máme hotovo. Zbývá najít lokální extrémy funkce ϕ v kruhu. Zde jde o klasické lokální extrémy, protože už díky zavedení pracovní funkce žádnou vazbu nemáme, takže stačí hledat stacionární body:

$$\nabla \phi(x,y) = \vec{0} \implies \begin{cases} y = 0 \\ x = 0 \end{cases} \implies x = 0, \ y = 0.$$

Dostali jsme další (a poslední) podezřelý bod (0,0,0).

Poznamenejme, že zde jsme nemohli použít Lagrangeovy multiplikátory, protože naše množina byla dána nerovností, nikoliv rovnicí.

2. Porovnání kandidátů.

Nalezli jsme celkem devět podezřelých bodů. Hodnoty funkce v těchto bodech jsou:

$$f(0,0,1) = 0, \ f(0,0,0) = 1, \ f(\sqrt{2},\sqrt{2},0) = 3, \ f(-\sqrt{2},\sqrt{2},0) = -1, \ f(\sqrt{2},-\sqrt{2},0) = -1,$$

Porovnáním zjišťujeme, že

$$\min_{\Omega}(f) = -\frac{5}{3} = f\left(\frac{4}{3}, -\frac{4}{3}, \frac{2}{3}\right) = f\left(-\frac{4}{3}, \frac{4}{3}, \frac{2}{3}\right),$$

$$\max_{\Omega}(f) = 3 = f(\sqrt{2}, \sqrt{2}, 0) = f(-\sqrt{2}, -\sqrt{2}, 0).$$

To byl masakr.

 \triangle

Kapitolu uzavřeme zajímavou otázkou: Jak by to dopadlo, kdybychom namísto Lagrangeových multiplikátorů použili selskorozumový přístup, tedy redukci počtu proměnných a podobné triky? Začneme základnou, protože tam pracujeme s příjemnou dvourozměrnou funkcí

$$\phi(x, y) = f(x, y, 0) = xy + 1.$$

Lokální extrém jsme našli přímo, ale co ta hranice? Je dána podmínkou $x^2 + y^2 = 4$, která nabízí dva možné způsoby redukce počtu proměnných.

Jedna možnost je rozdělit si kružnici na dva úseky, jeden daný rovnicí $y=\sqrt{4-x^2}$ a druhý daný rovnicí $y=-\sqrt{4-x^2}$ pro $x\in\langle -2,2\rangle$. Chování ϕ a tedy i f na první polovině kružnice popisuje pomocná funkce

$$\psi(x) = f(x, \sqrt{4 - x^2}, 0) = x\sqrt{4 - x^2} + 1.$$

Hledáme extrémy ψ na množině $\langle -2, 2 \rangle$, jde tedy o novou optimalizační úlohu.

Nejprve najdeme lokální extrémy.

$$\psi'(x) = 0 \implies \sqrt{4 - x^2} + \frac{-x^2}{\sqrt{4 - x^2}} = 0 \implies x = \pm \sqrt{2}.$$

Dostáváme body $(\sqrt{2}, \sqrt{2}, 0)$ a $(-\sqrt{2}, \sqrt{2}, 0)$ do seznamu podezřelých.

Zbývá se podívat na hranici množiny $\langle -2,2\rangle$. Tato hranice se skládá ze dvou hodnot, jmenovitě $x=\pm 2$, a přidáme odpovídající prostorové body do seznamu: (2,0,0) a (-2,0,0).

Podobně zpracujeme druhou polovinu kružnice danou vzorcem $y = -\sqrt{4 - x^2}$, krajní body $x = \pm 2$ se budou shodovat, přibudou dva kandidáti $(\sqrt{2}, -\sqrt{2}, 0)$ a $(-\sqrt{2}, -\sqrt{2}, 0)$.

Vidíme, že jsme našli ty čtyři body jako u Lagrangeova postupu, ale také dva body navíc $(\pm 2, 0, 0)$. To jsou artefakty parametrizace a znamenají práci navíc v porovnávací fázi.

Je také možné použít jiné dvě poloviny kružnice dané vzorci $x = \pm \sqrt{4 - y^2}$, což by opět vedlo na kandidáty $(\pm \sqrt{2}, \pm \sqrt{2}, 0)$ a k tomu krajní body intervalů $(0, \pm 2, 0)$. I zde by tedy zvnikly dva body navíc.

Nepříjemná na těchto přístupech byla nutnost vždy pracovat dvakrát, protože jsme neuměli vyjádřit kružnici jedním vzorcem. To nás ponouká, abychom zkusili lepší parametrizaci, jmenovitě

polární souřadnice. Připomeňme, že hledáme extrémy funkce $\phi(x,y)=xy+1$ na jisté kružnici o poloměru 2. Zkusíme parametrizaci $x=2\cos(\varphi),\,y=2\sin(\varphi)$ pro $\varphi\in\langle 0,2\pi\rangle$. Zkoumáme hodnoty pomocné funkce

$$\psi(\varphi) = f(2\cos(\varphi), 2\sin(\varphi), 0) = 4\cos(\varphi)\sin(\varphi) + 1 = 2\sin(2\varphi) + 1.$$

Podezřelými body jsou samozřejmě ty krajní pro φ , tedy 0 a 2π . Obě hodnoty nám do seznamu přidají stejný bod, jmenovitě (2,0,0). Zase jde o artefakt parametrizace.

Zajímavější je otázka lokálních extrémů.

$$\psi'(\varphi) = 0 \implies 4\cos(2\varphi) = 0 \implies \varphi = \frac{\pi}{4} + k\frac{\pi}{2}, \ k \in \mathbb{Z}.$$

Vznikají tak čtyři hodnoty, $\varphi = \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}$, po dosazení do vzorců dostáváme ony čtyři podezřelé body $(\pm\sqrt{2}, \pm\sqrt{2}, 0)$.

Čtenář si jistě závěr udělá sám, mi z toho vycházejí ty Lagrangeovy multiplikátory lépe, hlavně protože nevyráběly body navíc.

Teď už je čas podívat se na hledání lokálních extrémů na vrchlíku. Ukážeme tři přístupy.

1. Při pohledu na rovnici vrchlíku $x^2 + y^2 + z^2 = 4$ nás třeba napadne, že by bylo zajímavé si vyjádřit $z = \sqrt{4 - x^2 + y^2}$. Pak bychom měli hledat lokální extrémy pomocné funkce

$$\phi(x,y) = f(x,y,\sqrt{4-x^2-y^2}) = xy + (\sqrt{4-x^2-y^2}-1)^2$$

na množině bodů (x,y) splňujících $x^2 + y^2 \le 2$. Jdeme na to:

$$\nabla \phi = \vec{0} \implies \begin{cases} y - 2x \frac{\sqrt{4 - x^2 - y^2} - 1}{\sqrt{4 - x^2 - y^2}} = 0 \\ x - 2y \frac{\sqrt{4 - x^2 - y^2} - 1}{\sqrt{4 - x^2 - y^2}} = 0 \end{cases} \implies \begin{cases} y = 2x \frac{\sqrt{4 - x^2 - y^2} - 1}{\sqrt{4 - x^2 - y^2}} \\ x = 2y \frac{\sqrt{4 - x^2 - y^2} - 1}{\sqrt{4 - x^2 - y^2}} \end{cases}$$

Dosadíme-li z druhé rovnice do první, dostáváme

$$y = 4y \left(\frac{\sqrt{4 - x^2 - y^2} - 1}{\sqrt{4 - x^2 - y^2}} \right)^2.$$

Pokud je y=0, pak vychází i x=0 a dostáváme další bod (0,0,2) do seznamu kandidátů.

Pokud $y \neq 0$, můžeme zkrátit a dostáváme

$$1 = 4\left(\frac{\sqrt{4 - x^2 - y^2} - 1}{\sqrt{4 - x^2 - y^2}}\right)^2 \implies \frac{\sqrt{4 - x^2 - y^2} - 1}{\sqrt{4 - x^2 - y^2}} = \pm \frac{1}{2}.$$

Případ, kdy to vyjde $+\frac{1}{2}$: Pak z první rovnice gradientové podmínky máme y=x. Máme také

$$\frac{\sqrt{4-x^2-y^2}-1}{\sqrt{4-x^2-y^2}} = \frac{1}{2} \implies \sqrt{4-x^2-y^2} = 2 \implies x^2+y^2 = 0.$$

To má jediné řešení, x = y = 0. Pak také z = 2 a tento bod už máme v seznamu.

Případ, kdy to vyjde $-\frac{1}{2}$: Pak z první rovnice gradientové podmínky máme y=-x. Máme také

$$\frac{\sqrt{4-x^2-y^2}-1}{\sqrt{4-x^2-y^2}} = -\frac{1}{2} \implies \sqrt{4-x^2-y^2} = \frac{2}{3}.$$

To nám říká, že $z=\frac{2}{3}$. Když tam dosadíme y=-x, dostaneme

$$\sqrt{4-2x^2} = \frac{2}{3} \implies x = \pm \sqrt{2-\frac{2}{9}} = \pm \frac{4}{3}.$$

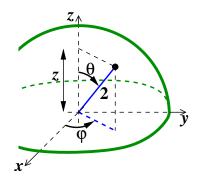
Dostáváme tedy dva nové podezřelé body, jmenovitě $\left(\frac{4}{3}, -\frac{4}{3}, \frac{2}{3}\right)$ a $\left(-\frac{4}{3}, \frac{4}{3}, \frac{2}{3}\right)$.

Vyšlo to stejně jako při prvním přístupu, z mého pohledu ten Lagrange vedl na jednodušší výpočty.

[4b]

2. Zkušenější řešitel by si mohl připomenout, že kulová plocha se dobře vyjadřuje sférickými souřadnicemi. Pro daný bod na sféře máme polární úhel θ vystihující, jak se tento bod odklání od osy z. Tím se určí souřadnice z a také kružnice, na které bod leží. Jeho polohu na kružnici pak určíme pomocí azimutu φ .

$$x = 2\cos(\varphi)\sin(\theta)$$
$$y = 2\sin(\varphi)\sin(\theta)$$
$$z = 2\cos(\theta).$$



Pracujeme pak s pomocnou funkcí

$$\phi(\varphi,\theta) = 4\cos(\varphi)\sin(\varphi)\sin^2(\theta) + (2\cos(\theta) - 1)^2 = 2\sin(2\varphi)\sin^2(\theta) + (2\cos(\theta) - 1)^2.$$

Hledáme lokální extrémy pro $0 \le \varphi \le 2\pi$ a $0 \le \theta \le \frac{\pi}{2}$. Podíváme se zde na lokální extrémy, podmínka je

$$\nabla \phi = \vec{0} \implies \frac{4\cos(2\varphi)\sin^2(\theta) = 0}{4\sin(2\varphi)\sin(\theta)\cos(\theta) - 4(2\cos(\theta) - 1)\sin(\theta) = 0.}$$

První rovnice nabízí dvě možnosti. Jedna je, že $\sin(\theta) = 0$. Vzhledem k rozsahu proměnné to nabízí jen hodnotu $\theta = 0$ neboli z = 2. Pak x = y = 0, tedy máme bod (0, 0, 2).

Druhá možnost je, že $\sin(\theta) \neq 0$, takže jej lze zkrátit v druhé rovnici a první dává $\cos(2\varphi) = 0$. Pak $2\varphi = \frac{\pi}{2} + 2k\pi$ neboli $\varphi = \frac{\pi}{4} + k\pi$.

Případ
$$\varphi = \frac{\pi}{4}$$
 nebo $\varphi = \frac{5\pi}{4}$: Pak $\sin(2\varphi) = 1$, druhá rovnice tedy říká, že $\cos(\theta) - (2\cos(\theta) - 1) = 0 \implies \cos(\theta) = 1$.

Vychází z=2, to už tu bylo.

Případ $\varphi=\frac{3\pi}{4}$ nebo $\varphi=\frac{7\pi}{4}$: Pak $\sin(2\varphi)=-1,$ druhá rovnice tedy říká, že

$$-\cos(\theta) - (2\cos(\theta) - 1) = 0 \implies \cos(\theta) = \frac{1}{3}.$$

Vychází $z = \frac{2}{3}$. Co ještě víme? $Z \cos(\theta) = \frac{1}{3}$ dostaneme $\sin(\theta) = \sqrt{\frac{8}{9}} = \frac{2}{3}\sqrt{2}$. Opět dva podpřípady:

Pokud $\varphi = \frac{3\pi}{4}$, pak

$$x = 2 \cdot \frac{2}{3}\sqrt{2} \cdot \left(-\frac{\sqrt{2}}{2}\right) = -\frac{4}{3},$$
$$y = 2 \cdot \frac{2}{3}\sqrt{2} \cdot \frac{\sqrt{2}}{2} = \frac{4}{3}.$$

Máme tedy podezřelý bod $\left(-\frac{4}{3}, \frac{4}{3}, \frac{2}{3}\right)$.

Pokud $\varphi = \frac{7\pi}{4}$, pak

$$x = 2 \cdot \frac{2}{3}\sqrt{2} \cdot \frac{\sqrt{2}}{2} = \frac{4}{3},$$
$$y = 2 \cdot \frac{2}{3}\sqrt{2} \cdot \left(-\frac{\sqrt{2}}{2}\right) = -\frac{4}{3}.$$

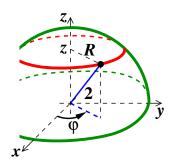
Máme tedy podezřelý bod $(\frac{4}{3}, -\frac{4}{3}, \frac{2}{3})$.

Opět stejný výsledek, tentokráte to možná bylo o chlup náročnější.

3. Zdá se, že výpočty nám komplikovala složená funkce ve výrazu $(z-1)^2$. Dalo by se tomu vyhnout, pokud bychom použili z jako jeden z parametrů. Takto to dělají válcové souřadnice, ale ty potřebují třírozměrný objekt. Budeme to tedy muset vymyslet sami.

Když si zvolíme nějakou hodnotu z, tak to již na vrchlíku jednoznačně určí kružnici v této výšce. Bude mít poloměr $R = \sqrt{4-z^2}$. Bod na ní pak jednoznačně určíme například úhlem, což se hodí ke kruhové situaci. Dostáváme tak nestandardní parametrizaci

$$x = \sqrt{4 - z^2} \cos(\varphi)$$
$$y = \sqrt{4 - z^2} \sin(\varphi)$$
$$z = z.$$



kde $z \in \langle 0, 2 \rangle$ a $\varphi \in \langle 0, 2\pi \rangle$. Pak hledáme lokální extrémy pomocné funkce

$$\phi(\varphi, z) = (4 - z^2)\cos(\varphi)\sin(\varphi) + (z - 1)^2 = \frac{1}{2}(4 - z^2)\sin(2\varphi) + (z - 1)^2.$$

Gradient poskytne rovnice:

$$\nabla \phi = \vec{0} \implies \frac{(4 - z^2)\cos(2\varphi) = 0,}{-z\sin(2\phi) + 2(z - 1) = 0.}$$

V první rovnici můžeme mít $4-z^2=0$, pak z parametrických vzorců dostaneme x=y=0 a máme kandidáta (0,0,2).

Jinak musí platit $cos(2\varphi) = 0$. Podobně jako u předchozího řešení rozebereme případy.

Případ $\varphi = \frac{\pi}{4}$ nebo $\varphi = \frac{5\pi}{4}$: Pak $\sin(2\varphi) = 1$, druhá rovnice tedy říká, že

$$-z + 2(z - 1) = 0 \implies z = 2.$$

To už tu bylo.

Případ $\varphi = \frac{3\pi}{4}$ nebo $\varphi = \frac{7\pi}{4}$: Pak $\sin(2\varphi) = -1$, druhá rovnice tedy říká, že

$$z + 2(z - 1) = 0 \implies z = \frac{2}{3}$$
.

Dopočítáme ostatní proměnné. Pro $\varphi = \frac{3\pi}{4}$ máme

$$x = \sqrt{4 - \frac{4}{9}} \cdot \left(-\frac{\sqrt{2}}{2}\right) = -\frac{4}{3},$$
$$y = \sqrt{4 - \frac{4}{9}} \cdot \frac{\sqrt{2}}{2} = \frac{4}{3}.$$

Máme tedy podezřelý bod $\left(-\frac{4}{3}, \frac{4}{3}, \frac{2}{3}\right)$. Pro $\varphi = \frac{7\pi}{4}$ pak

$$x = \sqrt{4 - \frac{4}{9}} \cdot \frac{\sqrt{2}}{2} = \frac{4}{3},$$
$$y = \sqrt{4 - \frac{4}{9}} \cdot \left(-\frac{\sqrt{2}}{2}\right) = -\frac{4}{3}.$$

Máme tedy podezřelý bod $(\frac{4}{3}, -\frac{4}{3}, \frac{2}{3})$.

Vyšlo to stejně jako u polárních souřadnic, ale výpočty byly možná o trochu příjemnější.

 \triangle

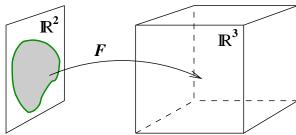
Tím jsme probrali všechny základní přístupy ke zkoumání globálních a vázaných extrémů. Poznamenejme, že k parametrickým plochám a křivkám se vrátíme z obecného pohledu v sekci 5a a prakticky v sekci 6d, zatímco k souřadnicím se vrátíme v sekci 7b.

59

5. Představujeme vektorové funkce

Funkce více proměnných jsme si uvedli jako přirozené zobecnění, kdy se jednorozměrný definiční obor nahradí vícerozměrným. Stejně přirozené je provést totéž i s oborem hodnot.

Vektorová funkce je libovolné zobrazení $F: D \mapsto \mathbb{R}^m$, kde $D \subseteq \mathbb{R}^n$. Někdy se také používá název vektorové pole, motivaci uvidíme v sekci 6c.

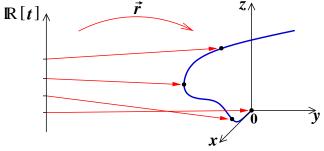


Jeden příklad vektorové funkce jsme už potkali. Je-li dána funkce více proměnných $f(\vec{x})$, pak její gradient také používá proměnnou \vec{x} a jeho hodnoty jsou vektory, je to tedy vektorová funkce. Například pro funkci $f(x,y) = xy^2$ dostáváme gradient $\nabla f(x,y) = (y^2, 2xy)$, je to funkce $\mathbb{R}^2 \to \mathbb{R}^2$.

Protože hodnoty funkce F jsou vektory, docela by se hodilo vektorové funkce značit \vec{F} , ať se to čtenáři připomíná. Matematikům se ale nikdy nechtělo, tak i zde budeme psát jen F, abychom to měli jednotné s literaturou. Další důležitá úmluva je, že hodnoty takové vektorové funkce se dají napsat po složkách, tedy píšeme $F(\vec{x}) = (F_1(\vec{x}), \dots, F_m(\vec{x}))$. Standardně si představujeme v cílovém prostoru kartézské souřadnice, ale není to nutné. Podstatné je, že chování funkce F je úzce spojeno s chováním komponent F_j . V řadě situací je přirozné pracovat jen s těmito komponentami (viz transformace souřadnic níže), takže pak vlastně pojem vektorové funkce nepotřebujeme. Jsou ale situace, kdy vektorové funkce poskytují přirozený a efektivní pohled na situaci.

5a. Parametrické křivky

V rámci matematiky se lidé asi nejčastěji setkávají s vektorovými funkcemi v podobě parametrických křivek a ploch. Představme si letadlo poletující ve vzduchu, zavedeme si nějakou přirozenou kartézskou soustavu, třeba s počátkem na letišti a osami vedoucími na východ a sever. Aktuální pozice letadla je pak dána třírozměrným vektorem tradičně značeným \vec{r} závisejícím na čase, jde tedy o funkci $\vec{r} = \vec{r}(t)$ neboli vektorovou funkci $\mathbb{R} \mapsto \mathbb{R}^3$.



Dráze, kterou letadlo letělo, přesněji řečeno množině bodů, kterými proletělo, ještě přesněji řečeno množině hodnot vektorové funkce $\vec{r}(t)$ říkáme parametrická křivka. V tomto případě jde o útvar, který je svou podstatou jednorozměrný (lze jím procházet jen v jednom směru, tam nebo zpět po křivce), ale je umístěn do třírozměrného prostoru. V mnoha případech tento útvar dokážeme rozeznat a pojmenovat matematicky. Pokud by například letadlo kroužilo, pak by výsledná křivka byla kružnice v \mathbb{R}^3 .

Funkci $\vec{r}(t)$ říkáme parametrizace této křivky. Nese informaci o jejím tvaru, ale také o fyzikálním ději, který ji vytvořil, například rychlosti pohybu po této křivce a podobně, takže parametrické křivky jsou velmi užitečné v aplikacích. Čtenáře snad napadne, že tatáž křivka může být parametrizována více způsoby, například každé z aut snažící se dojet z Prahy do Brna vytvoří svou vlastní parametrizaci křivky zvané D1 (pokud bychom dálnici s přimhouřením oka považovali za křivku o šířce jednoho bodu, její propustnost by tomu někdy ostatně napovídala).

Obecně vzato, všechny vektorové funkce typu $\mathbb{R} \to \mathbb{R}^n$ lze interpretovat jako parametrické křivky, matematické popisy svou podstatou jednorozměrných útvarů umístěných v prostoru \mathbb{R}^n . Protože se v této kapitole soustředíme na vektorové funkce, budeme místo $\vec{r}(t)$ používat obecné značení F(t).

Z matematického pohledu je zajímavá otázka odvodit, jakou křivku jistá vektorová funkce vlastně určuje. Jindy naopak máme matematický útvar a rádi bychom jej popsali. Někdy to lze pomocí rovnic (např. přímka nebo kružnice v rovině), ale někdy se zase hodí popis prostřednictvím parametrické křivky, pak říkáme, že jsme ten objekt parametrizovali.

Příklad: Uvažujme vektorovou funkci $F(t) = (a_1t+b_1, a_2t+b_2)$, kde a_i, b_i jsou zvolené parametry. Jde o funkci $\mathbb{R} \mapsto \mathbb{R}^2$. Budeme-li předpokládat standardní značení os v cílovém prostoru, tak vlastně máme popis pohybu ve formě

$$x = a_1 t + b_1$$
$$y = a_2 t + b_2.$$

S touto podobou se pracuje docela často, odvoláváme se přímo na souřadnice bez využití pojmu vektorové funkce.

Můžeme si představit, že jde o zápis pohybu brouka po stole, přičemž t je čas. Pokud bychom mu namočili bříško do inkoustu, zanechal by po sobě trasu, nějaký matematický objekt v rovině. Jaký objekt to bude? Klasický trik spočívá v eliminaci parametru.

$$t = \frac{1}{a_1}(x - b_1) \implies y = a_2 \frac{1}{a_1}(x - b_1) + b_2 \implies y = \frac{a_2}{a_1}x + \left(b_2 - \frac{a_2}{a_1}b_1\right).$$

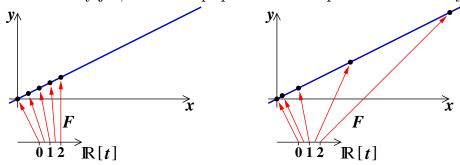
Právě jsme zjistili, že tento brouček lezl po přímce. Matematicky řečeno, daná funkce je parametrizací přímky. Rovnou vidíme jednu výhodu parametrického popisu. Vzorcem y = ax + b svislou přímku popsat neumíme (není to tedy univerzální rovnice pro přímky v rovině), v parametrickém popisu to půjde snadno volbou $a_1 = 0$.

Stejnou přímku půjde parametricky popsat i jinými způsoby, někdy je to jedno, jindy může správná volba parametrizace výrazně ušetřit práci. Například přímku o konkrétní rovnici y = ax + b lze parametrizovat jako F(t) = (t, at + b) neboli

$$x = t$$
$$y = at + b,$$

ale také jako F(t) = (13t + 7, 13at + 7a + b).

Lze si hrát i zajímavěji. Například teď už umíme spočítat, že F(t)=(2t,t) je přímka o rovnici $y=\frac{1}{2}x$, ale tutéž přímku dokážeme získat i jako $F(t)=(2t^3,t^3)$. Geometricky jde o stejný objekt, ale fyzikálné o velmi rozdílný jev, v druhém případě brouček opravdu masivně zrychluje.



Teď si ovšem představme, že takto zrychleně jedeme po této přímkové silnici, ale ona není postavena na zemi, je kvůli močálům vedena na pylonech dva metry nad zemí. Pak bychom pozici popsali jako $F(t) = (2t^3, t^3, 2)$, tedy je to jednodimenzionální parametrická křivka v 3D prostoru.

 \triangle

Příklad: Křivka daná vektorovou funkcí $F(t) = (\cos(t), \sin(t))$ neboli

$$x = \cos(t)$$

$$y = \sin(t)$$

se také dá rozpoznat eliminací parametru t, zde je asi nejsnažší obě rovnice umocnit na druhou a sečíst. Zjistíme, že jde o jednotkovou kružnici $x^2 + y^2 = 1$, parametrický popis ovšem říká více. Můžeme si třeba představit pejska s vodítkem připevněným na kolíku, jak obíhá konstantní rychlostí v kladném směru (proti směru hodinových ručiček). Pokud by pejsek obíhal v opačném směru a dvakrát rychleji, vytvořil by parametrizaci $F(t) = (\cos(-2t), \sin(-2t))$ stejné kružnice.

Parametrická křivka $F(t) = (\cos(t^2), \sin(t^2))$ také popisuje jednotkovou kružnici, v tomto případě se ovšem rychlost obíhání stále zvětšuje. Teď se podíváme na několik zajímavých modifikací.

 $F(t)=(t\cos(t),t\sin(t))$ pro $t\in\langle 1,\infty\rangle$ dává spirálu, obíhá se dokola, přičemž se zvětšuje i poloměr.

 $F(t) = (\frac{1}{t}\cos(t), \frac{1}{t}\sin(t))$ pro $t \in (1, \infty)$ je pro změnu spirála zavinutá směrem k počátku.

 $F(t) = (\cos(t), \sin(t), t)$ je pohyb ve třech rozměrech. Průmět na rovinu xy je naše stará známá kružnice, teď se ovšem také postupně zvětšuje výška. Výsledná křivka je tedy stoupající spirála, můžeme si třeba představit větron, který našel termiku a tak si krouží a stoupá.

Na závěr něco trochu jiného, uvažujme $F(t) = (t^2, 2t, \sin(t))$. První dvě souřadnice nám dávají průmět dráhy brouka do roviny xy. Z $x = t^2$, y = 2t eliminujeme t a dostaneme $x = \frac{1}{4}y^2$. Brouk se tedy pohybuje nad parabolou otevřenou doprava, do směru osy x. Výška brouka nad rovinou xy je dána třetí souřadnicí, což je $\sin(t)$. Brouk tedy letí střídavě nad a pod onou parabolou. Na otázku proč už matematika odpovědět neumí.

 \triangle

Pro úplnost připomeneme klasické parametrické křivky:

Kružnice v \mathbb{R}^2 : $x = R\cos(t)$, $y = R\sin(t)$ pro $t \in (0, 2\pi)$ či nějaký jiný interval dostatečné délky, klidně i pro $t \in \mathbb{R}$, kružnici je možné oběhnout vícekrát.

Elipsa v \mathbb{R}^2 : $x = A\cos(t)$, $y = B\sin(t)$ pro $t \in (0, 2\pi)$ či nějaký jiný interval dostatečné délky, klidně i pro $t \in \mathbb{R}$.

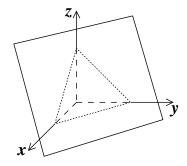
Přímka v \mathbb{R}^n procházející bodem \vec{a} ve směru \vec{u} je dána funkcí $F(t) = \vec{a} + t\vec{u}$ pro $t \in \mathbb{R}$. Zde vidíme jednu z výhod parametrického zápisu. Přímku v třírozměrném prostoru nelze popsat jednou algebraickou rovnicí, musíme použít dvě (průnik dvou rovin), zatímco parametrický popis stačí jeden.

Dalším zajímavým objektem jsou **parametrické plochy**. Jsou to objekty, které jsou svou podstatou dvourozměrné, ale vyskytují se v prostoru vyšší dimenze. Budou tedy popsány pomocí vektorových funkcí typu $\mathbb{R}^2 \mapsto \mathbb{R}^n$ pro $n \geq 3$. Můžeme si představit prostěradlo vznášející se ve vzduchu.

Z fyzikálního pohledu je užitečné to vnímat jako popis chování nějakého zařízení, které se pohybuje v třírozměrném prostoru, ale nikoliv libovolně, takže jsme díky nějakým mechanickým omezením schopni jej ovládat pouze dvěma ovladači, má tedy dva stupně volnosti. Představme si například portálový jeřáb, tedy vozíček jezdící po přímém kusu kolejnice, často namontované pod stropem tovární haly, ze kterého je na kabelu zavěšen hák. Operátor může jezdit vozíčkem po kolejnici a spouštět či vytahovat hák, tedy má dva ovladače. Množina všech míst, na kterých se hák může nacházet, je svislý obdélník, tedy v podstatě dvourozměrný objekt situovaný ve třírozměrném světě. Pokud je ta kolejnice zatočená, pak množina všech možných pozic háku už není plochá, ale je to "ohnutý" obdélník, nicméně stále v podstatě dvourozměrný objekt.

Příklad: Uvažujme rovinu v \mathbb{R}^3 , která protíná všechny tři osy v hodnotě 1, tedy prochází body (1,0,0), (0,1,0) a (0,0,1), je dána rovnicí x+y+z=1.

Pokud si vyjádříme z=1-x-y, můžeme to vnímat jako popis toho, jak vysoko nad (pod) sebou vidíme dotyčnou rovinu, když stojíme v rovině xy na souřadnici (x,y). Z toho snadno vyrobíme parametrický popis dotyčné plochy:



$$x = x$$

$$y = y$$

$$z = 1 - x - y.$$

Bude to vypadat lépe, pokud použijeme jiné názvy parametrů:

$$x = s$$

$$y = t$$

$$z = 1 - s - t.$$

Jinak řečeno, je to vektorová funkce F(s,t)=(s,t,1-s-t). Jako obvykle jsou i jiné parametrizace, například F(u,v)=(u+v,u-v,1-2u). Shodu bychom povrdili tak, že by po eliminaci parametrů u,v z rovnic

$$x = u + v$$
$$y = u - v$$
$$z = 1 - 2u$$

vyšla rovnice x + y + z = 1.

 \triangle

S parametrizací ploch a křivek jsme se velice zblízka seznámili v sekci 4b, zejména ke konci, a hodně to využijeme v sekci 6d.

Obecně řečeno, každou vektorovou funkci $F: \mathbb{R}^n \mapsto \mathbb{R}^m$, kde n < m, si můžeme interpretovat jako popis nějakého n-rozměrného útvaru umístěného v m-rozměrném prostoru, jen si to pro m > 3 neumíme představit.

Případ m=n má zase jinou užitečnou interpretaci, jmenovitě změnu souřadnic. Například pokud chceme v dvourozměrném prostoru přejít mezi polárními a kartézskými souřadnicemi, máme v jednom směru příjemné vzorce $x=r\cos(\varphi), y=r\sin(\varphi)$. Jde vlastně o vektorovou funkci $F(r,\varphi)=\left(r\cos(\varphi),r\sin(\varphi)\right)$ z \mathbb{R}^2 do \mathbb{R}^2 .

Lidé v praxi často pracují spíš s jednotlivými vzorci pro nové souřadnice $x = x(r, \varphi), y = y(r, \varphi)$, pohled skrz vektorové funkce ale někdy situaci zjednoduší. Teoreticky se o změně souřadnic mluví v sekci 7b, používáme ji v sekci 6d.

5b. Základní analytické pojmy

Limita a spojitost se dají velice snadno zobecnit, stačí uvažovat vícerozměrná okolí v cílovém prostoru.

Definice.

Nechť F je vektorová funkce s hodnotami v \mathbb{R}^m definovaná na nějakém prstencovém okolí bodu $\vec{a} \in \mathbb{R}^n$, nechť $\vec{L} \in \mathbb{R}^m$.

Řekneme, že \vec{L} je limita F pro \vec{x} jdoucí k \vec{a} , značeno $\lim_{\vec{x} \to \vec{a}} \left(F(\vec{x}) \right) = \vec{L}$, jestliže pro každé okolí U hodnoty \vec{L} existuje prstencové okolí V bodu \vec{a} takové, že $F[V] \subseteq U$.

Vektorovou funkci můžeme psát jako $F=(F_1,\ldots,F_m)$. Snadno se dokáže, že $\lim_{\vec{x}\to\vec{a}} \left(F(\vec{x})\right) = \vec{L}$ právě tehdy když $\lim_{\vec{x}\to\vec{a}} \left(F_j(\vec{x})\right) = L_j$ pro každé $j=1,\ldots,m$. Bylo by tedy možné také definovat

limitu takto "po souřadnicích".

Teď ten nápad použijeme a prostě prohlásíme, že vektorová funkce F je spojitá v jistém bodě, na jisté množině či prostě spojitá (tedy na svém definičním oboru), pokud toto platí pro všechny komponenty F_j . Jak je zvykem, spojité vektorové funkce jsou příjemné, například grafy nejsou "roztrhané" a platí zobecnění některých vět o spojitých funkcích. Udělejme si stručný přehled známých tvrzení.

Věta. (Heineho věta)

Nechť F je vektorová funkce s hodnotami v \mathbb{R}^m definovaná na nějakém prstencovém okolí D bodu $\vec{a} \in \mathbb{R}^n$, nechť $\vec{L} \in \mathbb{R}^m$.

Pak $\lim_{\vec{x} \to \vec{a}} (F(\vec{x})) = \vec{L}$ právě tehdy, když $\lim_{k \to \infty} (F(\vec{x}(k))) = \vec{L}$ pro všechny posloupnosti $\{\vec{x}(k)\} \subseteq D - \{\vec{a}\}$ splňující $\vec{x}(k) \to \vec{a}$.

Věta.

Nechť F je vektorová funkce s hodnotami v \mathbb{R}^m definovaná na nějakém prstencovém okolí bodu $\vec{a} \in \mathbb{R}^n$, nechť $\vec{L} \in \mathbb{R}^m$.

Jestliže limita F v \vec{a} existuje a je konečná, tak je F omezená na nějakém prstencovém okolí bodu \vec{a} .

Věta.

Nechť $F \colon D \mapsto \mathbb{R}^m$ je vektorová funkce, kde $D \subseteq \mathbb{R}^n$ je omezená uzavřená množina.

- (i) Jestliže je F spojitá na D, pak je F[D] omezená uzavřená množina.
- (ii) Jestliže je F spojitá a prostá na D, pak je i inverzní funkce F_{-1} : $F[D] \mapsto D$ spojitá.

Nesmíme se ale nechat unést, například věta o mezihodnotě se již zobecnit nedá.

Příklad: Uvažujme parametrickou křivku $F(t) = (2t - t^2, t)$ pro $t \in (0, 2)$.

Jde o kopeček v rovině xy. Snadno spočítáme F(0)=(0,0) a F(2)=(0,2), krajní body křivky tedy leží na ose y. Rozhodně ale neplatí, že by funkce také nabývala hodnot mezi nimi. V tomto případě by šlo o úsečku mezi cílovými body (0,0) a (0,2), ale funkce se tam nedostane, i když je spojitá.

 $\mathbb{R}[t]$ $\downarrow p$ $\downarrow p$

Platí ale, že definiční obor $\langle 0,2\rangle$ coby souvislá množina se zobrazí na křivku, která je zase souvislá.

 \triangle

Věta.

Nechť $F: D(F) \to \mathbb{R}^m$ je vektorová funkce, kde $D(F) \subseteq \mathbb{R}^n$. Jestliže je F spojitá, pak pro každou souvislou $M \subseteq D(F)$ je i F[M] souvislá.

Pro vektorové funkce můžeme zavést i stejnoměrnou spojitost.

Definice.

Nechť $F: D(F) \to \mathbb{R}^m$ je vektorová funkce, kde $D(F) \subseteq \mathbb{R}^n$. Nechť M je podmnožina D(F).

Řekneme, že F je **stejnoměrně spojitá** na M, jestliže

pro každé $\varepsilon > 0$ existuje $\delta > 0$ takové, že $||F(\vec{x}) - F(\vec{y})|| < \varepsilon$ pro všechna $\vec{x}, \vec{y} \in M$ splňující $||\vec{x} - \vec{y}|| < \delta$.

Věta.

Nechť $F: D \mapsto \mathbb{R}^m$ je vektorová funkce, kde $D \subseteq \mathbb{R}^n$.

Jestliže je F je spojitá na D a D je omezená a uzavřená, pak je F stejnoměrně spojitá na D.

Co se týče derivací, podobně jako pro funkce více proměnných, i u vektorových funkcí začneme směrovými derivacemi.

Definice.

Nechť F je vektorová funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Nechť \vec{u} je vektor z \mathbb{R}^n .

Řekneme, že funkce F je **diferencovatelná** v bodě \vec{a} ve směru \vec{u} , jestliže limita $\lim_{t\to 0} \left(\frac{F(\vec{a}+t\vec{u})-F(\vec{a})}{t}\right)$ existuje konečná.

Pak definujeme (směrovou) derivaci F v bodě \vec{a} ve směru \vec{u} jako

$$D_{\vec{u}}F(\vec{a}) = \lim_{t \to 0} \left(\frac{F(\vec{a} + t\vec{u}) - F(\vec{a})}{t} \right).$$

Vlastně jsme jen v definici směrové derivace zaměnili f za F. Teď se ovšem v limitě pracuje s vektory, a proto je i směrová derivace F vektorem z cílového prostoru \mathbb{R}^m funkce F.

Protože všechny operace ve vzorci v definici—odčítání, dělení skalárem, limitní proces—fungují po souřadnicích, vidíme, že vlastně aplikujeme proces pro nalezení směrové derivace zvlášť na každou složku funkce $F = (F_1, \ldots, F_m)$. Z toho plyne, že směrová derivace $D_{\vec{u}}F(\vec{a})$ existuje právě tehdy, když existují směrové derivace $D_{\vec{u}}F_j(\vec{a})$ pro všechna $j=1,\ldots,m$, a v takovém případě můžeme psát

$$D_{\vec{u}}F(\vec{a}) = (D_{\vec{u}}F_1(\vec{a}), \dots, D_{\vec{u}}F_m(\vec{a})).$$

Jako obvykle zaujímají speciální místo mezi směrovými derivacemi derivace v souřadnicových směrech. Zajímají nás tedy $\frac{\partial F_j}{\partial x_i}$, kde $i=1,\ldots,n$ a $j=1,\ldots,m$.

U funkcí více proměnných jsme parciální derivace zabalili do objektu zvaného gradient, který se následně docela hodil. U vektorových funkcí máme něco podobného.

Definice.

Nechť $F = (F_1, \ldots, F_m)$: $D(F) \mapsto \mathbb{R}^m$ je vektorová funkce, kde $D(F) \subseteq \mathbb{R}^n$. Definujeme její **Jacobiho matici**

$$J_F(\vec{a}) = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(\vec{a}) & \dots & \frac{\partial F_1}{\partial x_n}(\vec{a}) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1}(\vec{a}) & \dots & \frac{\partial F_m}{\partial x_n}(\vec{a}) \end{pmatrix}$$

v bodech, kde příslušné parciální derivace existují.

Pro případ m = n pak definujeme také **jakobián** jako

$$\Delta_F(\vec{a}) = \det(J_F(\vec{a})).$$

Jacobiho matice vznikne tak, že do prvního řádku umístíme všechny parciální derivace F_1 , tedy vlastně jej vytvoříme z gradientu F_1 . Stejně plníme další řádky, matice má tedy m řádků a n sloupců. Symbolicky,

$$J_F = \begin{pmatrix} \nabla F_1 \\ \vdots \\ \nabla F_m \end{pmatrix}.$$

Jak se dá čekat, souvislost mezi gradientem a směrovou derivací se také přenáší do světa vektorových funkcí. Ve vzorci budeme potřebovat vynásobit Jacobiho matici vektorem, který proto musí být sloupcový, takže použijeme transpozice.

$$D_{\vec{u}}F(\vec{a})^T = J_F(\vec{a})\vec{u}^T.$$

Jacobiho matice a jakobián mají velký význam při studiu parametrických povrchů nebo třeba při substituci v integrálu.

Protože s funkcemi majícími derivace budeme pracovat často, zavedeme praktické prostory hladkých funkcí také pro vektorové funkce. Pro oblast G značí $[C(G)]^m$ množinu všech vektorových funkcí $G \mapsto \mathbb{R}^m$ spojitých na G a $[C^1(G)]^m$ je množinou všech vektorových funkcí $G \mapsto \mathbb{R}^m$, které mají všechny parciální derivace prvního řádu (tedy existuje Jacobiho matice) spojité na G.

Značení je vcelku intuitivní, protože pro obecnou množinu M představuje M^m množinu vektorů dimenze m se složkami z množiny M. Zde v množině $[C(G)]^m$ najdeme vektorové funkce, tedy funkce zapsatelné jako (G_1, \ldots, G_m) , a ty leží v $[C(G)]^m$ právě tehdy, když jednotlivé složky G_j jsou z C(G). Obdobně to funguje i pro $[C^1(G)]^m$.

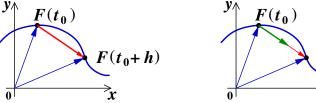
Podívejme se na některé zajímavé interpretace derivací vektorových funkcí. Začneme parametrickými křivkami. Uvažujme parametrickou křivku F(t): $\mathbb{R} \mapsto \mathbb{R}^m$. Pak n=1, tedy namísto parciálních derivací derivujeme standardním způsobem. Jacobiho matice je pak sloupcový vektor

$$J_F(t) = \begin{pmatrix} F_1'(t) \\ \vdots \\ F_n'(t) \end{pmatrix}.$$

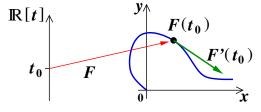
Lidé v aplikacích často preferují řádkový vektor $F'(t) = (F'_1(t), \dots, F'_m(t))$ a říkají tomu derivace F. Je snadné ukázat, že

$$F'(t_0) = \lim_{h \to 0} \left(\frac{F(t_0 + h) - F(t_0)}{h} \right).$$

Toto již trochu napovídá, jakou informaci nám derivace poskytuje. Podívejme se na zlomek v limitě. Funkci F(t) si představujeme jako vektor pozice v čase t. Pohybujeme se tedy po křivce, rozdíl $F(t_0+h)-F(t_0)$ ukazuje, jak se změnila naše pozice mezi časy t_0 a t_0+h . Tato změna pozice je také vektor, na obrázku vlevo jej vidíme červeně. Z obrázku je pěkně vidět, že při pohybu po křivce tato změna polohy nemusí odpovídat dráze, kterou jsme urazili. Pokud je ale h malé, pak jde o velmi dobrou aproximaci. Ve zlomku pak dělíme časem, čímž dostaneme průměrnou rychlost (zelená šipka v obrázku napravo). Tato průměrná rychlost má stejný směr jako změna pozice, což by pro malé h mělo souhlasit velmi dobře se směrem, kterým zrovna vede křivka.



Protože se dá očekávat, že pro rozumné pohyby bude zlomek aproximovat okamžitou rychlost v t_0 tím lépe, čím je h menší, dostáváme následující interpretaci: Derivace $F'(t_0)$ (pokud existuje) udává okamžitou rychlost pohybu po křivce v čase t_0 . Tato rychlost je vektor, který je tečný k dané křivce v bodě $F(t_0)$.



V této souvislosti jsou důležitou kategorií tzv. **hladké křivky**, což jsou parametrické křivky F(t) definované na intervalu, na kterém také mají spojitou derivaci a $F'(t) \neq 0$ na vnitřku tohoto intervalu.

Často se také hodí mít k dispozici normálový vektor ke křivce (například kdybychom se chtěli

bavit o odstředivé síle). Najdeme jej vzorcem

$$\left[\frac{F'(t)}{\|F'(t)\|}\right]'.$$

Vyplývá to ze zajímavého faktu, který pak aplikujeme na vektorovou funkci T = F':

Fakt.

Jestliže je vektorová funkce T(t) diferencovatelná na okolí t_0 a velikost ||T(t)|| je konstantní na tomto okolí, pak je $T'(t_0)$ kolmý vektor k $T(t_0)$.

Ukažme si ještě jedno zajímavé použití derivace. Uvažujme diferencovatelnou vektorovou funkci jedné proměnné F(t) a zvolme bod a. Zajímá nás aproximace vektoru F(a+h). Protože je vektorová funkce F diferencovatelná, můžeme derivovat i její jednotlivé složky $F_j(t)$, což jsou vlastně obyčejné funkce. Můžeme proto udělat následující:

$$F(a+h) = (F_1(a+h), \dots, F_m(a+h))$$

$$\approx (F_1(a) + F'_1(a)h, \dots, F_m(a) + F'_m(a)h)$$

$$= (F_1(a), \dots, F_m(a)) + (F'_1(a), \dots, F'_m(a))h$$

$$= F(a) + F'(a)h.$$

Není to zase tak překvapující. Zajímá nás, jak se změní naše pozice F v prostoru \mathbb{R}^m , pokud se proměnná posune o kousek z bodu t=a. Pro hladké funkce F neboli hladké křivky by ten pohyb mělo jít velmi dobře aproximovat pohybem po tečně, a my už víme, že směr tečny ke křivce dané funkcí F je právě F'(a).

Tento výsledek naznačuje, že by i pro vektorové funkce mohlo fungovat něco jako aproximace Taylorovým polynomem. Pro vektorové funkce jedné proměnné se to dá snadno odvodit prací po souřadnicích, jak jsme to právě viděli, pro funkce více proměnných už je Taylorův polynom dobrodružnější a necháme si to do kapitoly 7. Významné aplikace derivací se najdou také v sekci 6d.

5c. Diferenciální operátory

Operátor je zobrazení, u kterého očekáváme nějaké příjemné algebraické vlastnosti, například oblíbenou linearitu. Diferenciální operátor je takový, který funguje pomocí derivování. Zatím jsme potkali dva důležité diferenciální operátory. Pro funkce více proměnných je to gradient, který vyrábí z funkcí vektorové funkce, a pro vektorové funkce pak Jacobiho matice, což lze vnímat jako proces, který vyrábí z vektorových funkcí maticové funkce. Pro oba operátory platí řada pravidel, na která ovšem v textu zvaném Ilustrovaný úvod není místo a odkážeme čtenáře na kapitolu Více o derivaci; zde jen zmiňme, že oba jsou lineární, což je velmi užitečné při praktických výpočtech.

V této sekci představíme dva další diferenciální operátory, oba hrají důležitou roli v aplikacích, zejména ve fyzice, a také se objeví v několika klíčových větách integrálního počtu. Oba operátory vyžadují u vektorové funkce m=n, tedy omezíme se na transformace, i když zde bude výhodnější zaměřit se na jinou interpretaci.

Správná představa pro tuto sekci je, že vektorová funkce zachycuje proudění kapaliny. V případě n=3 si můžeme představit třeba koryto řeky, pak definiční obor D je množina pozic, ve kterých se nachází voda, a $F(\vec{a})$ uvádí směr a rychlost částice vody, která se nachází v lokaci \vec{a} . Vzhledem k tomu, že ve funkci nevidíme jako proměnnou čas, pak jde o ustálené proudění.

V případě n=2 si třeba představíme vodu tekoucí po stole ve velmi tenké vrstvě (opět ustáleně, tedy teče pořád stejně). Na tento případ se zaměříme, protože si jej umíme nakreslit, s tím že naše pozorování budou platit také pro vyšší dimenze, kde je obvykle budeme používat.

Nemusí ovšem proudit jen voda, proudit může třeba vzduch a další podobná média, dokonce i teplo nebo může jít o tok magnetického pole. Právě v kontextu těchto představ se používá název vektorové pole namísto vektorová funkce.

Pro úplnost ještě poznamenejme, že existuje úzká spojitost mezi funkcemi více proměnných a vektorovými poli. Představme si funkci tří proměnných zaznamenávající teplotu v jednotlivých bodech místnosti. Fyzika (či selský rozum) nám říkají, že pokud jsou vedle sebe místa s rozdílnou teplotou, pak dojde k přenosu tepla neboli ke vzniku proudění. Intenzita souvisí s tím, jak moc se teploty liší. Přesně tuto informaci nabízí gradient. Ten ale dává směr a velikost největšího růstu, zatímco teplo proudí směrem největšího spádu (od teplého ke studentému), tedy přesně opačně. Vychází nám tak zajímavý nápad, že pokud funkce T více proměnných popisuje rozložení teploty, pak vektorová funkce $-\nabla T$ popisuje tok tepla.

Tento princip lze nalézt v mnoha oblastech fyziky. Ne každé vektorové pole vzniká pomocí gradientu, ale mnohé ano a pak mají speciální vlastnosti, říká se jim konzervativní pole. To už se ale zase dostáváme někam jinam, proto zpět k tématu a uvedeme první diferenciální operátor této sekce.

Definice.

Nechť $F = (F_1, \ldots, F_n)$: $D(F) \mapsto \mathbb{R}^n$ je vektorová funkce, kde $D(F) \subseteq \mathbb{R}^n$. Definujeme její **divergenci** jako

$$\operatorname{div}(F)(\vec{a}) = \sum_{i=1}^{n} \frac{\partial F_i}{\partial x_i}(\vec{a}) = \frac{\partial F_1}{\partial x_1}(\vec{a}) + \frac{\partial F_2}{\partial x_2}(\vec{a}) + \dots + \frac{\partial F_n}{\partial x_n}(\vec{a})$$

v bodech $\vec{a} \in D$, kde příslušné parciální derivace existují.

Vlastně sčítáme čísla z diagonály Jacobiho matice, což je známá matematická operace trace neboli stopa. Příznivci lineární algebry by tedy mohli psát $\operatorname{div}(F) = \operatorname{tr}(J_F)$, ale obvykle se preferuje ještě jiné značení, viz kapitola Více o derivaci.

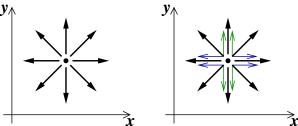
Co nám divergence vektorového pole říká? Standardní interpretace je, že když div $(F)(\vec{a}) > 0$, tak v tom místě přibývá média, je tam takzvané **zřídlo**. Pokud například F popisuje šíření tepla, tak by v bodě \vec{a} mohl být zdroj vytápění. Naopak div $(F)(\vec{a}) < 0$ znamená, že tam média ubývá, je tam tzv. **odtok** či také propad.

Je možné to prostě akceptovat, ale zvídavému čtenáři nejspíše vrtá hlavou několik otázek: Jak tu zřídelnost pozná zrovna tento vzoreček? A jak to, že tam zcela chybí ostatní derivace, to se opravdu obejdou třeba bez znalosti $\frac{\partial F_1}{\partial x_2}$? (To je moje oblíbená otázka.)

Vše je samozřejmě možné korektně odvodit středně komplikovaných výpočtem pomocí limity

Vše je samozřejmě možné korektně odvodit středně komplikovaných výpočtem pomocí limity integrálu a pak bychom tomu věřili, ale cílem tohoto ilustrovaného úvodu je, abychom si význam pojmů uměli intuitivně představit, aby nám to dávalo smysl. Dá se to udělat pro divergenci?

Rada autorů předvádí dvě výrazné situace pro dvourozměrný tok, aby čtenáře přesvědčili, že definice poskytuje správné odpovědi.



Na levém obrázku vidíme typický případ zřídla, proudění zachycují černé šipky. Třeba nám na ten stůl crčí pramínek vody ze stropu, protože soused nad námi vytopil byt. Z pohledu tenké, dvoudimenzionální vody se na stole jakoby odnikud neustále magicky objevuje nová voda.

Voda dopadá na bod \vec{a} a pak se rozlévá do stran, takže vektory jejího proudění v obrázku nás jistě nepřekvapí. Podíváme-li se na složky jednotlivých vektorů ve směru osy x (u těch šikmých jsme je vyznačili modře v obrázku napravo, jde vlastně o složku F_1 vektorové funkce F), pak si můžeme všimnout, že nalevo od bodu \vec{a} ukazují doleva, tedy mají záporné hodnoty, zatímco napravo od bodu \vec{a} mají hodnoty kladné. Jinak řečeno, když se pohybujeme zleva doprava přes bod \vec{a} , tak se

ty vodorovné komponenty zvětšují (jde ze záporných hodnot na kladné), dokonce tomu tak je, i když se pohybujeme vodorovně kousek nad a pod bodem \vec{a} . Z toho usoudíme, že první složka F_1 naší vektorové funkce je okolo \vec{a} rostoucí při posunu ve směru osy x, tedy $\frac{\partial F_1}{\partial x} > 0$.

Podobně když se pohybujeme ve svislém směru vzhůru, tak se zelené složky F_2 vektorů F ve směru y zvětšují (ze záporných se stanou kladnými), tedy $\frac{\partial F_2}{\partial y} > 0$. Dostáváme

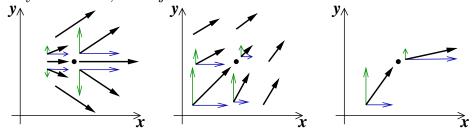
$$\operatorname{div}(F) = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} > 0.$$

Teď na levém obrázku vidíme typický odtok, například když jsme měli vodu na stole a rozhodli jsme se jí zbavit navrtáním dírky v bodě \vec{a} . Směry proudění nás opět nepřekvapují a tentokráte se směry složek překlápějí z kladných na záporné při pohybu zleva doprava, tedy první složka F_1 vektorové funkce F se při pohledu zleva doprava zmenšuje, obdobně pro druhou složku F_2 při posunu zdola nahoru. Závěr: Ony dvě parciální derivace jsou záporné a máme

$$\operatorname{div}(F) = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} < 0.$$

Toto vysvětlení je možné nalézt na spoustě míst a pokud vám stačí, tak máte splněno, ale mě to moc nepřesvědčilo. Ukázalo to, že divergence dává u zdroje a odtoku správné znaménko, ale nevysvětlilo mi to, proč ten zdroj a odtok musím rozeznávat zrovna tímto vzorcem. V obou případech se skokově mění směr vektoru F při průchodu bodem \vec{a} a není jasné, proč by se to mělo testovat zrovna parciálními derivacemi, jejichž cílem je poskytovat jinou informaci (rychlost změny).

Podívejme se tedy na situace, které jsou méně extrémní.



V první nám teče řeka, koukáme na ni shora, a vidíme také místo, kde do ní nějaká továrna vypouští odpadní vodu (patřičně vyčištěnou, samozřejmě). Je to tedy příklad zřídla. Když sledujeme změnu modré první složky F_1 funkce F při pohybu zleva doprava, popřípadě změnu zelené druhé složky F_2 při pohybu zdola nahoru, pak to vypadá, že se zvětšují, takže bychom asi věřili, že v tomto případě také $\operatorname{div}(F)(\vec{a}) > 0$.

Na druhém obrázku teče řeka (diagonálně, ať máme opravdu pořádný příklad, ne něco speciálního), ale někdo umístil na dno v bodě \vec{a} jeden konec mimoprostorového tunelu, kterým část vody odčerpává, ale ne moc, aby něco zbylo (zdá se, že emzákům už nestačí unášet lidi). Tentokráte se zdá, že se modrá složka F_1 při pohybu zleva doprava zmenšuje, stejně jako zelená složka F_2 při pohybu zdola nahoru, takže příslušné parciální derivace $\frac{\partial F_1}{\partial x}$ a $\frac{\partial F_2}{\partial y}$ vycházejí záporné a $\operatorname{div}(F)(\vec{a}) < 0$.

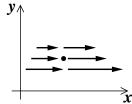
V obou případech nešlo o změnu orientace, ale jen velikosti, čili začíná to vypadat, že ony parciální derivace opravdu mohou být tím správným rozeznávacím nástrojem.

Pro úplnost se podíváme na třetí obrázek. Tam do řeky ani nic nevypouštíme, ani neodčerpáváme, jen se tok stočí. Mělo by pak platit $\operatorname{div}(F) = 0$. Na obrázku se modrá složka F_1 při pohybu zleva doprava zvětšuje, zatímco zelená složka F_2 se při pohledu zdola nahoru zmenšuje. Příslušné

parciální derivace tedy mají opačná znaménka. Že by se opravdu pokrátily na nulu? To z hrubého náčrtu nepoznáme, tohle by potřebovalo přesnější výpočet.

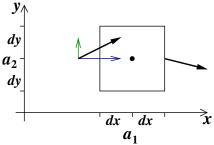
To už bylo lepší vysvětlení, ale pořád nevíme, proč se ve vzorci pro n=2 vůbec nevyskytují další dvě parciální derivace $\frac{\partial F_1}{\partial y}$ a $\frac{\partial F_2}{\partial x}$. V této souvislosti je zajímavé vrátit se ke všem situacím výše a všimnout si, že u složky F_1 je znaménko $\frac{\partial F_1}{\partial y}$ vždycky stejné jako znaménko $\frac{\partial F_1}{\partial x}$, obdobně to platí i pro složku F_2 . Možná je tedy zbytečné ty dvě derivace uvažovat, protože poskytují stejnou informaci, kterou už máme. Ale co si pak máme myslet o tomto obrázku?

Pro zjednodušení je složka F_2 vždy nulová, všechno se děje jen ve směru x. Pohledem zleva doprava snadno nahlédneme, že $\frac{\partial F_1}{\partial x} > 0$ okolo bodu \vec{a} , ale také $\frac{\partial F_1}{\partial y} < 0$ okolo \vec{a} . Co teď s tím? Opravdu můžeme ten druhý údaj ignorovat a prostě prohlásit, že v bodě \vec{a} je zdroj? Nebo že by takový obrázek nebyl možný?



Pro mě je z toho všeho závěr, že ty obrázky ukázaly, že definice divergence, jak ji máme zavedenou, dokáže v mnoha situacích správně odpovědět (což může někomu stačit). Pokud jí ale chceme rozumět lépe, tak to chce ještě nějaký jiný přístup. Proto provedeme takzvané intuitivní odvození, což znamená, že budeme podvádět, ale jen přiměřeně, a náš příběh bude mít správné poučení.

Vezmeme bod \vec{a} a budeme sledovat, kolik média vyprodukuje. Podle toho zjistíme, zda je to zdroj, odtok či třetí možnost, kdy se nic neztrácí ani nepřidává. Aby se nám lépe počítalo, uděláme si kolem něj malinkatý čtereček přesahující ve vodorovném směru nalevo i napravo o dx, zatímco ve svislém směru o dy.



Spočítáme, zda tento čtvereček produkuje či požírá médium. Nejprve si uděláme bilanci vodorovných stran. Kolik média proteče levou stranou za čas dt? To záleží na rychlosti, kterou médium vtéká, a ta nemusí být všude stejná. Naštěstí máme opravdu malý čtvereček, dokonce nekonečně malý, takže můžeme předpokládat, že rychlost vtékání je na celé levé straně stejná, řekněme daná hodnotou uprostřed, tedy v bodě $(a_1 - dx, a_2)$. Zároveň si všimneme, že svislá složka vektoru rychlosti nevede ke vtékání, o tom rozhoduje pouze první složka. Rychlost vtékání je tedy $F_1(a_1 - dx, a_2)$. Celkový objem vtečeného média pak je rychlost krát velikost strany krát čas neboli $F_1(a_1 - dx, a_2)(2dy)dt$. Stejně spočítáme, kolik vyteče pravou stranou ven, vyjde $F_1(a_1 + dx, a_2)(2dy)dt$. Odečtením získáme celkovou bilanci, a protože chceme kladné číslo v případě, že čtvereček k médiu přidává, budeme počítat

$$F_1(a_1 + dx, a_2)(2dy)dt - F_1(a_1 - dx, a_2)(2dy)dt.$$

Jsou tady ovšem dvě dobré otázky. Za prvé, co když proudění teče zprava doleva? Pak jsou složky F_1 záporné a čtenář si jistě snadno rozmyslí, že pokud vtéká méně než vytéká, tak z našeho vzorce zase vyjde kladné číslo, což přesně chceme. Obdobně dostaneme správné znaménko, pokud by do obou stran médium vtékalo (pak je $F_1(a_1-dx,a_2)>0$ a $F_1(a_1+dx,a_2)<0$) nebo z obou vytékalo. Tento vzorec je tedy univerzální.

Poslední dva případy nás ovšem přivádějí k druhé otázce, která je principiální: Co když část média, která vtéká levou stranou, opustí čtverec některou z vodorovných stran? Odpověď zní, že nás to zde nemusí trápit, protože pak se prostě započítá při zkoumání vodorovných stran.

Než se k němu dostaneme, zjednodušíme si vzorec, který jsme dostali. Protože je dx hodně malé, můžeme si dovolit nahradit obě funkční hodnoty pomocí lineárního Taylorova rozvoje. Všimněte

si, že v obou případech měníme jen první souřadnici.

$$F_{1}(a_{1} + dx, a_{2})(2dy)dt - F_{1}(a_{1} - dx, a_{2})(2dy)dt$$

$$= \left(F_{1}(a_{1}, a_{2}) + \frac{\partial F_{1}}{\partial x}(a_{1}, a_{2}) \cdot dx\right)(2dy)dt - \left(F_{1}(a_{1}, a_{2}) + \frac{\partial F_{1}}{\partial x}(a_{1}, a_{2}) \cdot (-dx)\right)(2dy)dt$$

$$= \frac{\partial F_{1}}{\partial x}(\vec{a})(2dx)(2dy)dt.$$

Tyto úvahy konečně vysvětlily, proč nás u složky F_1 nezajímá, jak se mění ve svislém směru neboli $\frac{\partial F_1}{\partial u}$, jen vodorovném.

Nyní naprosto stejným způsobem rozebereme bilanci toku skrz vodorovné strany, tentokráte nás pro změnu nebude zajímat vodorovná složka pohybu, jen svislá F_2 .

$$F_{2}(a_{1}, a_{2} + dy)(2dx)dt - F_{2}(a_{1}, a_{2} - dy)(2dx)dt$$

$$= \left(F_{2}(a_{1}, a_{2}) + \frac{\partial F_{2}}{\partial y}(a_{1}, a_{2}) \cdot dy\right)(2d2)dt - \left(F_{2}(a_{1}, a_{2}) + \frac{\partial F_{2}}{\partial y}(a_{1}, a_{2}) \cdot (-dy)\right)(2dx)dt$$

$$= \frac{\partial F_{2}}{\partial y}(\vec{a})(2dy)(2dx)dt.$$

Když naše dvě částečné bilance sečtme, dostaneme číslo vyjadřující, kolik za čas dt náš čtverec přidal, popřípadě ubral proudícího média.

$$\frac{\partial F_1}{\partial x}(\vec{a})(2dx)(2dy)dt + \frac{\partial F_2}{\partial y}(\vec{a})(2dy)(2dx)dt = \left(\frac{\partial F_1}{\partial x}(\vec{a}) + \frac{\partial F_2}{\partial y}(\vec{a})\right)(2dy)(2dx)dt$$
$$= \operatorname{div}(F)(\vec{a}) \cdot dA \cdot dt.$$

V závěrečném kroku jsme součin (2dx)(2dy) napsali jako obsah zkoumaného čtverce. Pokud bychom chtěli přínos bodu \vec{a} jako takového, pak bychom museli zprůměrovat vydělením oním obsahem dA a časem dt a dostáváme divergenci F. Tento výpočet ukázal, že definice divergence byla udělána správně vzhledem k informaci, kterou chceme získat.

Stojí za zmínku, že lokální přibývání či ubývání média nemusí být způsobeno vnějším zásahem (topidlo, vodovodní kohoutek, magnet), ale může také být způsobeno změnou hustoty média. Například lokální zahřátí plynu znamená, že se začne rozpínat, tedy vnikne tam zřídlo, ale také se sníží hustota.

Naopak pokud je někde médium hustší, tak to znamená, že do tohoto místa proudilo víc, než odplynulo. Vyplývá z toho intuitivní pravidlo pro médium bez vnějšího přidávání/odebírání: $\operatorname{div}(F)>0$ znamená odtok neboli snížení hustoty, naopak $\operatorname{div}(F)<0$ znamená hromadění neboli nárust hustoty.

Pokud proudí dokonale nestlačitelná kapalina, pak hustotu měnit nesmí, v případě bez vnějšího ovlivnění pak máme $\operatorname{div}(F) = 0$. Tato podmínka tedy hraje v mechanice proudění kapalin významnou úlohu.

Nyní se podíváme na druhý operátor užitečný v aplikacích, a to je rotace vektorového pole (neboli vektorové funkce). Existuje univerzální definice, která je ale pro praktické výpočty v zásadě nepoužitelná a všichni pracují s jiným vzorcem, která má bohužel speciální verze v závislosti na dimenzi.

Definice.

Nechť $F = (F_1, F_2)$: $D(F) \mapsto \mathbb{R}^2$ je vektorová funkce, kde $D(F) \subseteq \mathbb{R}^2$. Definujeme její **rotaci** jako

$$rot(F)(\vec{a}) = \frac{\partial F_2}{\partial x}(\vec{a}) - \frac{\partial F_1}{\partial y}(\vec{a})$$

v bodech $\vec{a} \in D(F)$, kde příslušné parciální derivace existují.

Nechť $F = (F_1, F_2, F_3)$: $D(F) \mapsto \mathbb{R}^3$ je vektorová funkce, kde $D(F) \subseteq \mathbb{R}^3$. Definujeme její **rotaci** jako

$$rot(F)(\vec{a}) = \left(\frac{\partial F_3}{\partial y}(\vec{a}) - \frac{\partial F_2}{\partial z}(\vec{a}), \frac{\partial F_1}{\partial z}(\vec{a}) - \frac{\partial F_3}{\partial x}(\vec{a}), \frac{\partial F_2}{\partial x}(\vec{a}) - \frac{\partial F_1}{\partial y}(\vec{a})\right)$$

v bodech $\vec{a} \in D(F)$, kde příslušné parciální derivace existují.

V anglické literatuře se tomu říká **curl** a víc než značení rot(F) se používá curl(F).

Jako informaci nám tento operátor poskytuje? Standardní odpověď je, že rotace popisuje tendenci vektorového pole se stáčet. Co se tím myslí?

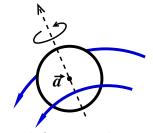
Uvažujme proudící řeku. Abychom posoudili její točivý vliv v bodě \vec{a} , tak tam vložíme virtuální kuličku, která je udržována na této pozici (třeba magneticky), ale může se volně otáčet. Pokud v tom místě médium proudí rovně, tak působí třením na povrch kuličky symetricky z pohledu proudění (nahoře i dole, vlevo i vpravo) a tudíž kulička nemá důvod se otáčet. Pokud se ovšem proudící médium v místě \vec{a} stáčí, pak mezi silami na různých stranách kuličky vznikne nerovnováha. Ostatně jistě všichni intuitivně tušíme, že když se proudění stáčí, tak na vnější straně proudí rychleji, tedy tato strana kuličky bude více ponoukána k pohybu.

Teď už tedy víme, co chceme poznat, ale ještě si musíme vyjasnit, jak se takové otáčení kuličky zapíše. Ve dvou dimenzích si představujeme rotující kotouč umístěný v bodě \vec{a} . Jaká data potřebujeme znát, abychom jeho rotaci popsali? Jednak směr rotace a druhak buď okamžitou úhlovou rychlost, nebo úhlové zrychlení, podle toho, co zrovna zkoumáme. Nás zajímá působení proudění na kuličku, což znamená sílu, což znamená zrychlení. Chceme si tedy schovat velikost úhlového zrychlení a jeho směr, na což stačí jedno číslo. Shodou okolností i rotace dvoudimenzionálního vektorového pole je výše definována jako číslo, takže přinejmenším v tomto to dává smysl.

Poznamenejme ještě, že ve fyzice se kladným směrem rotace myslí rotace vedoucí od osy x k ose y, tedy proti směru hodinových ručiček, pokud bychom přidali ještě osu z a dívali se na rovinu xy z její "špičky".

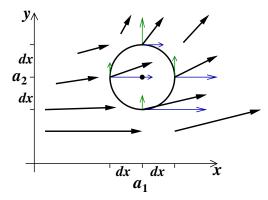
Teď si představme kuličku ve třírozměrném prostoru se středem v bodě \vec{a} . Jak bychom zachytili její rotaci? Důležitou informací je jistě osa rotace. Jak víme, že nějaká existuje? To je jednoduché. Pokud by na kuličce neexistovaly dva protilehlé body, které zůstávají na místě, pak by nešlo o rotaci, ale o nějaký jiný pohyb.

K uložení směru potřebujeme vektor. Jak zvolíme jeho orientaci? Tak, abychom tím zachytili směr rotace. Úmluva je orientovat vektor tak, abychom při pohledu z jeho špičky viděli kuličku rotovat v kladném směru, tedy proti směru hodinových ručiček. Ještě zbývá vybrat velikost tohoto vektoru, čímž hravě zakódujeme úhlové zrychlení a vše potřebné máme uloženo v jednom vektoru. Podíváme se, že rotace pro třírozměrné vektorové pole je definovaná zrovna jako třírozměrný vektor, takže tady nám to taky hraje.



Teď už víme, co hledáme, a rádi bychom viděli, proč zrovna ty vzorečky z definice poskytují správnou informaci.

Začneme ve dvourozměrném případě. Položíme kuličku o poloměru dx do bodu \vec{a} uvnitř dvourozměrného vektorového pole, takže je to vlastně kotouček, a prozkoumáme situaci. Na obrázku se proudění stáčí doleva neboli v kladném směru, tudíž bychom čekali snahu o rotaci v kladném směru. Podíváme se, jakou intenzitou na kotouček působí síly proudění na různých místech.



Začneme bodem dole. Nejprve si na něm zopakujeme klíčové fyzikální poznatky. Intenzita rotačního působení se jmenuje točivý moment a spočítá se jako moment síly. Ten se nejlépe počítá v případě, kdy je síla kolmá na rameno síly, tedy na spojnici středu rotace s působištěm síly. Pak stačí velikost síly vynásobit délkou ramene a máme točivý moment.

Pokud síla kolmá není (což je obvyklý případ a vidíme to i na dolním bodu kotoučku), pak je třeba sílu rozložit na dvě složky. Jedna je radiální, vede ke středu či od středu kuličky (kotoučku) a k rotaci nepřispívá. Druhá složka je kolmá k ramenu síly a tu pak použijeme při výpočtu momentu.

Na obrázku je příslušný rozklad označen zeleně a modře a shodou okolností jsou to právě složky F_2 a F_1 naší vektorové funkce, proto jsme také začali tím příjemným bodem dole, aby to tak pěkně vyšlo. Snadno spočítáme točivý moment v tomto bodě, je to $F_1(a_1, a_2 - dx) \cdot dx$. Znaménko souhlasí, síla působící vpravo má kladnou složku F_1 a tlakem dole chce roztáčet kotouček v kladném směru.

V horním bodě opět uvažujeme jen složku F_1 . Tentokráte ovšem kladná hodnota F_1 vede na rotaci v záporném směru, proto započítáme točivý moment $-F_1(a_1, a_2 + dx)dx$. Celková bilance horního a dolního bodu je tedy

$$F_1(a_1, a_2 - dx)dx - F_1(a_1, a_2 + dx)dx.$$

Dává znaménko smysl? V našem případě je síla dole větší než nahoře, čili bychom čekali, že celkově bude tendence roztáčet kuličku proti směru hodinových ručiček, a onen rozdíl také vyjde kladný, to souhlasí. Můžete si rozmyslet, že i pro jiné směry vektorů F bude tento rozdíl vždy dávat správné znaménko.

Výraz zase zjednodušíme lineární aproximací a dostaneme

$$F_{1}(a_{1}, a_{2} - dx)dx - F_{1}(a_{1}, a_{2} + dx)dx$$

$$= \left(F_{1}(a_{1}, a_{2}) + \frac{\partial F_{1}}{\partial y}(a_{1}, a_{2}) \cdot (-dx)\right)dx - \left(F_{1}(a_{1}, a_{2}) + \frac{\partial F_{1}}{\partial y}(a_{1}, a_{2}) \cdot dx\right)dx$$

$$= -2\frac{\partial F_{1}}{\partial y}(\vec{a})(dx)^{2}.$$

Nyní obdobným způsobem uděláme bilanci točivého momentu pro bod vlevo a vpravo. Vychází to

$$F_{2}(a_{1} + dx, a_{2})dx - F_{2}(a_{1} - dx, a_{2})dx$$

$$= \left(\left[F_{2}(a_{1}, a_{2}) + \frac{\partial F_{2}}{\partial x}(a_{1}, a_{2}) \cdot dx \right) dx - \left(F_{2}(a_{1}, a_{2}) + \frac{\partial F_{2}}{\partial x}(a_{1}, a_{2}) \cdot (-dx) \right) dx$$

$$= 2 \frac{\partial F_{2}}{\partial x} (\vec{a})(dx)^{2}.$$

Celkové působení sil je pak

$$-2\frac{\partial F_1}{\partial y}(\vec{a})(dx)^2 + 2\frac{\partial F_2}{\partial x}(\vec{a})(dx)^2 = \operatorname{rot}(F) \cdot 2(dx)^2.$$

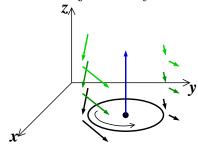
Pro zjištění rotačního zrychlení musíme vydělit rotačním momentem setrvačnosti, což vede ke zkrácení toho $(dx)^2$ a vidíme, že rot(F) opravdu ukazuje, jaké rotační zrychlení bude tekoucím médiem způsobeno.

Podobně jako u divergence, i zde toto zjednodušené odvození vypadá důvěryhodně, ale při bližším pohledu se tam najdou zásadní otazníky. Jedním z nich je tradiční otázka, proč nejsou třeba derivace $\frac{\partial F_1}{\partial x}$ a $\frac{\partial F_2}{\partial y}$. V našem výpočtu sice nefigurují, ale to je tím, že jsme se omezili jen na působení toku na čtyři pečlivě vybrané body. Kdybychom se podívali na rovnováhu momentů vůči jiným bodům, například tomu diagonálnímu vpravo dole, tak už by složky F_1 a F_2 nedávaly kolmý a radiální průmět, čímž by se celý výpočet značně zkomplikoval. Museli bychom počítat správné průměty F a ve formulkách by se objevily ony dvě chybějící derivace.

Nakonec se dá dojít ke vzorci z definice, ale na to je třeba udělat podrobnou fyzikální analýzu, což je něco, co v ilustrovaném úvodu nechceme vidět. Bohužel, na rozdíl od divergence mi není známo žádné relativně snadné obrázkové zdůvodnění významu rotace, které by zcela fungovalo, ale ten příběh výše nám dal alespoň nějaký vhled.

Přijměme tedy dvourozměrné vysvětlení a zkusme se na jeho základě podívat, nakolik bude dávat smysl třírozměrná definice.

Začneme tím, že se vrátíme k našemu dvourozměrnému příběhu, ale posadíme jej do tří rozměrů. Na stole, po kterém něco teklo v tenké vrstvě, nám rotoval virtuální papírový kotouček připíchnutý špendlíkem, ale najednou jsme si všimli, že se to vlastně odehrává ve třírozměrném prostoru. Formálně můžeme z naší dvourozměrné vektorové funkce $F(x,y) = (F_1(x,y), F_2(x,y))$ vyrobit třírozměrnou předpisem $F(x,y,z) = (F_1(x,y), F_2(x,y), 0)$, kde do funkce F formálně připíšeme i třetí proměnnou, ale není využita. Vznikne tak situace, kdy se dění na stole zkopíruje směrem vzhůru a dolů, i kotouček si pak představíme jako svislý válec.



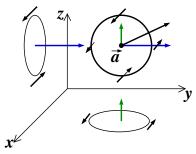
Rotace kotoučku neboli válce pak stále probíhá v rovině xy, tedy osa rotace je rovnoběžná s osou z. Na našem obrázku se kotouč otáčí od osy x k ose y, tedy kladně při pohledu od osy z. Vektor rotace tedy musí vést nahoru, jeho velikost je určena točivým momentem, který jsme spočítali jako dvourozměrnou rotaci. Hledaný vektor je tedy

$$(0,0,\mathrm{rot}(F)) = (0,0,\frac{\partial F_2}{\partial x}(\vec{a}) - \frac{\partial F_1}{\partial y}(\vec{a})).$$

Čtenář si může všimnout, že třetí souřadnice souhlasí s tím, jak je rotace definována pro třírozměrná vektorová pole.

Důležité je, že tento přechod od dvou dimenzí ke třem se dá udělat i naopak. Pokud nám rotuje kulička ve třech dimenzích, tak tuto situaci můžeme říznout skrz střed rovinou, nejlépe rovnoběžnou z některých souřadnicových rovin, a na řezu se zachovají fyzikální jevy. Zejména bude pravda, že točivý moment, který na tom řezu spočítáme, odpovídá příslušné složce třírozměrného vektoru točivého zrychlení.

Podívejte se na obrázek. Máme rotující kuličku ve třírozměrném prostoru, kde ji popohání vektorové pole $F(x,y,z)=(F_1,F_2,F_3)$. Abychom si to zjednodušili, představíme si situaci, kdy je vyznačená osa rotace a zároveň vektor točivého zrychlení rovnoběžná s rovinou yz, tedy nemusíme řešit první složku ve směru x. Znamená to, že se kulička točí, jako by se chtěla kutálet z obrázku směrem k nám. To může být způsobeno například vodním vírem, kdy se voda točí tak, že nad kuličkou teče z obrázku proti nám, zatímco pod kuličkou teče od nás, což jsme zkusili naznačit šipkami.



Přímo před kuličkou jakoby teče směrem dolů, ale ne zcela, protože pak by osa rotace ukazovala vodorovně doprava. Ona je ale trochu vytočená vzhůru, což znamená, že proudění před kuličkou vede dolů a také trochu doprava. To pak znamená, že proudění má složku i nalevo a napravo od kuličky, ale poněkud slabší než nad a pod.

Když třírozměrnou situaci řízneme rovinou skrz $\vec{a} = (a_1, a_2, a_3)$ rovnoběžnou s rovinou xy, získáme dvourozměrnou situaci, kterou jsme pro přehlednost zakreslili dolů do skutečné roviny xy. Přenesli jsme tam i dvě typické síly, které jsou ve skutečném obrázku vidět nalevo a napravo od kuličky. V tomto řezu vidíme jen to, nakolik se kulička točí vzhledem ke svislé ose, proto nás zajímaly jen díly nalevo a napravo, zatímco síly nad a pod kuličkou k rotaci viditelné v rovině xy nepřispívají.

Vlastně pracujeme na řezu s vektorovou funkcí

$$F(x,y) = (F_1(x,y,a_3), F_2(x,y,a_3)).$$

Spočítáme její točivý vektor okolo svislé osy přesně tak, jak jsme to udělali výše.

$$\left(0,0,\frac{\partial F_2}{\partial x}(\vec{a}) - \frac{\partial F_1}{\partial y}(\vec{a})\right).$$

Vidíme jej nakreslený zeleně a jeho velikost přesně souhlasí s třetí složkou skutečného třírozměrného vektoru točivého momentu.

Pokud si situaci řízneme rovinou rovnoběžnou s rovinou xz, kterou jsme si zase pro přehlednost nakreslili přímo do roviny xz včetně ukázek podstatných sil nad a pod koulí, pak vlastně pracujeme s pomocnou funkcí

$$F(x,z) = (F_1(x,a_2,z), F_3(x,a_2,z)).$$

Hrátky s dvourozměrným kotoučkem skončí u modře označeného vektoru ukazujícího ve směru osy y a s točivým momentem $\frac{\partial F_3}{\partial x}(\vec{a}) - \frac{\partial F_1}{\partial z}(\vec{a})$. Zde je ovšem třeba být opatrný. Moment vede od první pracovní osy ke druhé. V předchozím výpočtu se tedy šlo od x k y, což je kladná rotace. Teď ale jdeme od x k z, což je z pohledu osv y rotace v záporném směru. Abychom získali směr kladný, je třeba u naší dvourozměrné rotace změnit znaménko. Bereme tedy točivý moment okolo osy y jako $\frac{\partial F_1}{\partial z}(\vec{a}) - \frac{\partial F_3}{\partial x}(\vec{a})$. Vznikne vektor

$$\left(0, \frac{\partial F_1}{\partial z}(\vec{a}) - \frac{\partial F_3}{\partial x}(\vec{a}), 0\right).$$

Pokud modrý a zelený vektor složíme, měli bychom na našem speciálním obrázku, kde složka xnefiguruje, dostat skutečný vektor točivého momentu původní třírozměrné situace. V obecném případě bychom ještě kuličku řízli rovinou rovnoběžnou s rovinou yz. Pomocná funkce

$$F(y,z) = (F_2(a_1, y, z), F_3(a_1, y, z))$$

 $F(y,z)=(F_2(a_1,y,z),F_3(a_1,y,z))$ by nám dala točivý moment $\frac{\partial F_3}{\partial y}(\vec{a})-\frac{\partial F_2}{\partial z}(\vec{a})$. Má správné znaménko? Plus znamená rotaci od osy y k ose z, což je z pohledu osy x v kladném směru, souhlasí to. Toto číslo přijde jako první souřadnice vektoru točivého momentu ukazující směrem osy x:

$$\Big(\frac{\partial F_3}{\partial u}(\vec{a}) - \frac{\partial F_2}{\partial z}(\vec{a}), 0, 0\Big).$$

Pokud získané tři složky rotace poskládáme do jednoho vektoru, dostaneme přesně ten vzorec, který vidíme v definici třírozměrné rotace:

$$\Big(\frac{\partial F_3}{\partial y}(\vec{a}) - \frac{\partial F_2}{\partial z}(\vec{a}), \frac{\partial F_1}{\partial z}(\vec{a}) - \frac{\partial F_3}{\partial x}(\vec{a}), \frac{\partial F_2}{\partial x}(\vec{a}) - \frac{\partial F_1}{\partial y}(\vec{a})\Big).$$

Tolik k významu rotace vektorového pole.

Pro vzorce platné pro diferenciální operátory se podívejte do sekce 7a. Integrální věty s divergencí a rotací jsou v sekci 6c.

6. Představujeme integrál

Začneme návratem k určitému integrálu pro funkci jedné proměnné. Výraz $\int_a^b f(x) dx$ je definován jako "obsah pod grafem". Formálně nesprávná, ale výstižná představa je, že jakoby "sčítáme" hodnoty funkce v bodech x, přičemž je sčítáme ve formě obsahů teňounkých sloupečků, obdélníčků. Klíčovým pojmem je zde dx. Formálně jej lze interpretovat jako diferenciál proměnné x, což někteří autoři zdůrazňují značením dx.

Nicméně zde nás bude zajímat spíš původní význam; v předchozích staletích byl dx vnímán jako symbol pro nekonečně malý kousek osy x, který je také nazýván infinitesimála. Zjevně teď opouštíme svět rigorózní matematiky. Abychom to zamotali ještě víc, symbolem dx se také značí délka takového nekonečně malého kousku. Naštěstí je vhodná interpretace symbolu dx obvykle zjevná z kontextu.

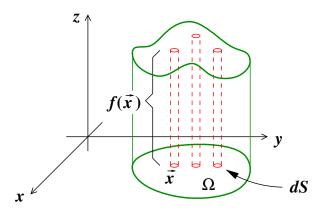
Integrování na intervalu I si běžně představujeme tak, že jsme si interval I rozdělili na malinké úseky délky dx a obrazec pod grafem následně aproximovali jako spojení tenkých obdélníkových sloupečků. Každý z nich má obsah $f(x) \cdot dx$ a my je sčítáme integrálem. Nám se zde bude hodit ještě trochu jiný pohled. Při integrování postupně procházíme všemi body intervalu I. Pro každý takový bod x započítáme příslušnou hodnotu funkce f(x) tak, že si kolem x vezmeme onen nekonečně malý kousek osy x s délkou dx a nad ním vytvoříme sloupeček s výškou f(x). Jeho obsah $f(x) \cdot dx$ pak přidáme do celkového součtu.

Z toho vychází představa pro funkce více proměnných. Máme-li nějakou množinu Ω v \mathbb{R}^n a na ní funkci f, pak také můžeme "sčítat" hodnoty. V prostoru \mathbb{R}^n máme infinitesimály, které budeme značit $d[\vec{x}]$ pokud je proměnná \vec{x} , ale typicky používáme specifické značení podle dimenze. V \mathbb{R}^2 máme rovinné infinitesimály, můžeme si představovat třeba nekonečně malé čtverečky či kolečka, a značíme je dS. Můžeme také psát $dS[\vec{x}]$, dS[x,y] a podobně, když chceme zdůraznit proměnné. Tyto infinitesimály mají obsah, který značíme dS.

Podobně máme v \mathbb{R}^3 příslušné 3D infinitesimály, můžeme si představit krychličky nebo kuličky, a značíme je dV nebo dV[x,y,z] atd. I zde používáme dV také pro jejich objem. Obecně pak používáme n-rozměrný objem pro infinitesimály $d[\vec{x}]$ v \mathbb{R}^n .

Snadno si představíme situaci v \mathbb{R}^2 , tedy integrování funkce dvou proměnných. Inspirováni integrálem jedné proměnné, chceme teď najít objem tělesa, které má jako základnu Ω a horní strana je dána grafem funkce f. Najdeme to následovně. Kolem každého bodu (x,y) z Ω si vezmeme vhodnou infinitesimálu, můžeme si představit malé kolečko. To bude základna pro svislý sloup (v tomto případě válec) o výšce f(x,y). Jeho objem je pak $f(x,y) \cdot dS$. Když tyto objemy "sečteme",

dostaneme integrál funkce f na Ω , značený $\iint_{\Omega} f(x,y) dS$.



Pro funkci tří proměnných a množinu Ω bychom pro každý bod $\vec{x} \in \Omega$ vzali 3D-infinitesimálu dV a nad ní vztyčili čtyřdimenzionální "sloup" o výšce $f(\vec{x})$. Sečtením všech odpovídajících 4D-

objemů $f(\vec{x}) \cdot dV$ pak dostaneme $\iiint\limits_{\Omega} f(\vec{x}) \, dV$, což je čtyřrozměrný objem objektu mezi základnou

 Ω a grafem f,což je chvíle, kdy to naše představivost vzdá. Nicméně myšlenka je snad jasná.

Tímto způsobem dojdeme k obecnému pojmu integrálu $\int \cdots \int f(\vec{x})d[\vec{x}]$. Poznamenejme, že kon-

krétní tvar diferenciálu není podstatný, jen musejí být natolik jednoduché, abychom mohli spočítat jejich obsah, objem atd. Ovšem schopnost rozložit množinu Ω na vhodné elementární částečky není zaručena. Když jsme integrovali funkci jedné proměnné na intervalu, tak nás zajímala jen kvalita funkce (některé jsou integrabilní, třeba ty spojité, jiné ne). Když přejdeme k funkcím více proměnných, tak možnost integrovat závisí na tom, zda jsou jak funkce, tak množina Ω rozumné. Přesná specifikace (zejména specifikace rozumnosti pro množinu) jsou komplikované a rozhodně mimo obzor ilustrovaného úvodu, podrobnější výklad se najde v kapitole 8. Jak záhy uvidíme, v praxi integrujeme jen přes množiny speciálních typů.

Co se stane, když se přes nějakou množinu integruje konstantní funkce f=1? Pak mají i všechny "sloupečky" výšku 1, tedy jejich n+1-dimenzionální objem je roven n-dimenzionálnímu objemu podstav $d[\vec{x}]$. Jejich "sčítání" integrováním pak odpovídá sčítání objemů podstav, na které jsme rozdělili množinu Ω , a tedy integrál poskytne n-rozměrný objem množiny Ω . To je zjevné v jednom rozměru, kde "objem" (neboli velikost) odpovídá délce. Integrál $\int\limits_a^b 1\,dx\,$ dává $b-a\,$ neboli délku

úsečky $\langle a,b\rangle$. Ve dvou rozměrech pak integrál dá plošnou výměru, ve třech běžný objem a pro vyšší dimenze to funguje obdobně.

To se hodí například v situaci, kdy chceme najít průměr funkce vzhledem k množině Ω . Vzorec pro funkce jedné proměnné je znám a hravě se zobecní i pro funkce více proměnných:

• Je-li funkce $f(\vec{x})$ integrovatelná na množině Ω , pak se průměr spočítá vzorcem

$$\operatorname{Ave}_{\Omega}(f) = \frac{\int \cdots \int f(\vec{x}) \, d[\vec{x}]}{\int \cdots \int 1 \, d[\vec{x}]}.$$

Hlavní myšlenka vícerozměrného integrálu je jasná, teď se podíváme, jak se počítá.

Než s tím začneme, je třeba vyjasnit terminologii. Standardní význam slova "oblast" v analýze je, že označuje omezenou otevřenou souvislou množinu. Bohužel, ve světě integrování lidé běžně používají slovo "oblast" pro množinu, přes kterou integrujeme, bez ohledu na její vlastnosti. Popravdě řečeno, při integrování se to spíš cítí tak, že integrujeme přes uzavřené množiny. Abych se tomu problému vyhnul, snažil jsem se mluvit spíš o množinách, ale nebyl jsem zcela úspěšný, protože mi občas přišlo, že situace vyloženě volá po slově oblast, a já jsem se s tím pokušením nepral. Naštěstí je obvykle jedno, jestli integrujeme přes množinu, její uzávěr nebo její vnitřek, takže nás to nemusí moc trápit.

6a. Dvoudimenzionální integrál

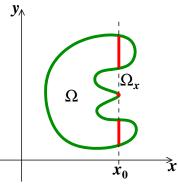
Integrálu, ve kterém integrujeme funkci dvou proměnných na množině $\Omega \subseteq \mathbb{R}^2$, říkáme **dvojný** integrál,

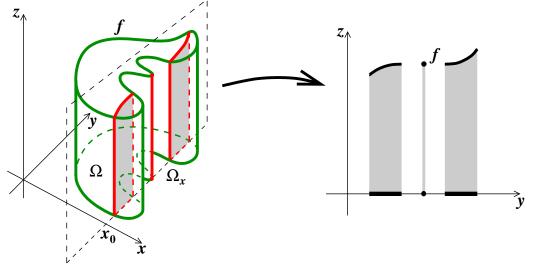
$$\iint\limits_{\Omega} f(x,y)dS.$$

Chceme najít proceduru pro jeho výpočet. Musíme sečíst objemy "sloupů" nad všemi body množiny Ω . Klíčovou myšlenkou vícedimenzionálního integrování je, že abychom prošli všemi body množiny Ω , zorganizujeme si je do skupin a zpracováváme každou zvlášť. Finta je, že tyto skupiny vytváříme tak, aby měly nižší dimenzi.

Jako to děláme? Zvolíme si základní směr, řekněme směr osy y, což v rovině xy znamená svislý směr. Konkrétní přímka v rovině v tomto směru je dána tím, že si pro x pevně zvolíme jistou hodnotu $x=x_0$. Předpokládejme, že tato přímka protne množinu Ω , a nazvěme tento průnik Ω_{x_0} . Budeme tomu říkat řez, a tyto řezy mohou být podle tvaru množiny Ω docela divoké, což nás později donutí omezit se jen na rozumnější množiny.

Rovnici $x = x_0$ lze uvažovat i v \mathbb{R}^3 , kde je to rovnice roviny kolmé na osu x neboli rovnoběžně s rovinou yz. Tato rovina protne těleso pod grafem f a vytvoří jeho řez.



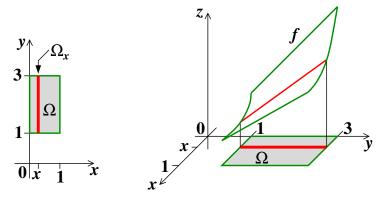


Poznamenejme, že teď už máme dva druhy řezů, řezy Ω_x skrz Ω a také řezy skrz těleso pod grafem. Snad se nám je podaří vždy rozlišit, naštěstí při praktickém počítání pracujeme jen s Ω , a tak nás to brzy přestane trápit.

Když se na řez tělesem pod grafem podíváme, tak vlastně vidíme graf funkce jedné proměnné $y\mapsto f(x_0,y)$, kde se y berou z bodů řezu Ω_{x_0} . Pokud je množina Ω_{x_0} rozumná (ideálně interval), tak tam můžeme funkci $f(x_0,y)$ integrovat (samozřejmě podle y) a získáme obsah toho řezu tělesem pod grafem. Všimneme si, že jde o určitý integrál vzhledem k y, takže při tom y zmizí a ve výsledku bude jen parametr x_0 , což se dá čekat vzhledem k tomu, že obsahy řezů budou nejspíše různé pro různé x_0 .

Příklad: Uvažujme funkci $f(x,y)=y\,e^{-x}$ na 2D intervalu $\langle 0,1\rangle \times \langle 1,3\rangle$ neboli na množině $\Omega=\{(x,y)\in\mathbb{R}^2;\ 0\leq x\leq 1\ \text{a}\ 1\leq y\leq 3\}.$

Zvolíme $x_0 = 0.3$. Tím vznikne řez $\Omega_{0.3}$ množinou Ω , který má tvar úsečky. Ukážeme obrázek. Obvykle pracujeme jen s obrázkem množiny Ω , zde pro dodání kontextu výjimečně ukážeme i pohled na třídimenzionální situaci s grafem funkce.



Na řezu se hodnoty y mění mezi 1 a 3 a máme tam funkci

$$y \mapsto f(x_0, y) = f(0.3, y) = y e^{-0.3},$$

kterou můžeme integrovat.

Při integrování je $e^{-0.3}$ konstanta, proto snadno získáme

$$\int_{1}^{3} y e^{-0.3} dy = e^{-0.3} \int_{1}^{3} y dy = e^{-0.3} \left[\frac{1}{2} y^{2} \right]_{1}^{3} = e^{-0.3} \left[\frac{9}{2} - \frac{1}{2} \right] = 4e^{-0.3}.$$

Toto je tedy obsah příslušného řezu tělesa pod grafem, podle očekávání se proměnná y již nevyskytuje. Výsledek ovšem záleží na hodnotě x, pro $x_0 = 0$ bychom dostali

$$\int_{1}^{3} y e^{0} dy = e^{0} \int_{1}^{3} y dy = 1 \cdot \left[\frac{1}{2} y^{2} \right]_{1}^{3} = 4.$$

Je užitečné si rozmyslet, že pro obecné x mezi 0 a 1 příslušný integrál vyjde

$$\int_{1}^{3} y e^{-x} dy = e^{-x} \int_{1}^{3} y dy = e^{-x} \left[\frac{1}{2} y^{2} \right]_{1}^{3} = e^{-x} \left[\frac{9}{2} - \frac{1}{2} \right] = 4e^{-x}.$$

Je to trochu jako parciální integrování. Získáme tak informaci o obsazích všech řezů.

Ne vždy je možné nepoužívanou proměnnou úplně vytknout. Neměl by to být problém, stačí si pořád představovat, že jisté členy jsou konstantní, a pracovat s nimi jako obvykle. Náš integrál se tedy dal počítat i jako

$$\int_{1}^{3} y e^{-x} dy = \left[\frac{1}{2} y^{2} e^{-x} \right]_{y=1}^{y=3} = \frac{9}{2} e^{x} - \frac{1}{2} e^{-x} = 4e^{-x}.$$

Protože v primitivní funkci jsou dvě proměnné, raději jsme si v dosazovací fázi připomněli, za co máme meze dosadit.

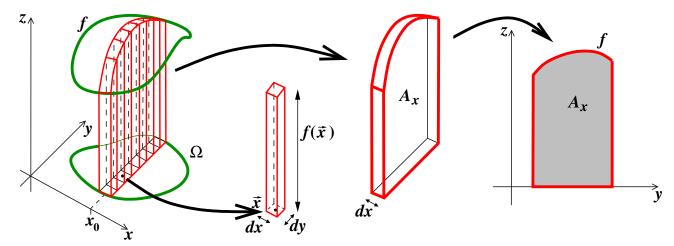
 \triangle

Počítat jednotlivé obsahy řezů tělesem pomocí integrace tedy dává smysl. Zavedeme pro takový integrál symbolické značení

$$\int_{\Omega_{\pi}} f(x,y) \, dy.$$

Je jen symbolické, protože jej nelze brát doslovně, na to by Ω_x musela být nějaká množina z \mathbb{R} . Ve skutečnosti je Ω_{x_0} množina z \mathbb{R}^2 , ve které všechny body mají společnou první souřadnici x_0 . Tomu integračnímu značení rozumíme tak, že dotyčné x_0 dosadíme do f za x a pak integrujeme přes všechna y takové, že body (x_0, y) jsou v Ω_x .

Čím je pro nás takové integrování na řezu Ω_x užitečné? Připomeňme, že naším konečným cílem je posčítat příspěvky $f(x,y)\cdot dS$ z celé množiny Ω , ale teď se chceme omezit jen na příspěvky pro (x,y) z jistého řezu Ω_{x_0} množiny Ω . Bude pro nás výhodné si představovat, že dS jsou nekonečně malé čtverečky $dx \times dy$. Při integraci přes řez vlastně spojujeme objemy kvádříků v tělese odpovídající bodům z Ω_{x_0} . Stojí těsně vedle sebe, mají shodnou šířku dx a vytvářejí proto jakousi desku s vykousaným horním okrajem (ten je daný funkcí f). Objem této desky získáme, když obsah její boční plochy vynásobíme tloušťkou dx.



Obsah takového řezu jsme vlastně počítali v příkladě výše. Obsah určitého řezu zjevně závisí na volbě x, má tedy smysl označit jej A_x , v příkladě to takto vyšlo. Selský rozum nám říká, že objem tělesa by měl být součtem objemů oněch desek, tedy chceme "sečíst" objemy $dV = A_x \cdot dx$. Infinitesimály se sčítají integrací, takže objem tělesa (neboli dvojný integrál) se dá přepsat jako

$$\iint\limits_{\Omega} f \, dS = \int A_x \, dx.$$

To vyjadřuje užitečný neformální princip:

• Objem tělesa lze získat sečtením obsahů všech jeho rovnoběžných řezů.

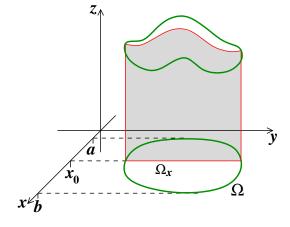
Platí to i v dalších dimenzích, třeba obsah rovinného obrazce lze získat součtem délek jeho rovnoběžných řezů, nebo čtyřrozměrný objem čtyřrozměrné množiny lze získat sečtením objemů jeho řezů, což jsou třírozměrná tělesa.

V integrálu napravo potřebujeme doplnit meze. Pokud je množina Ω omezená, tak velká část jejích řezů Ω_x bude prázdných, pak jsou i obsahy A_x řezů tělesem nulové a není důvod je zahrnovat do integrálu. Mělo by tedy být možné najít rozmezí $\langle a,b\rangle$ pro proměnnou x takové, že neprázdné Ω_x jsou jen pro x z tohoto rozmezí, tam pak má smysl integrovat.

Když si ještě připomeneme, že obsah A_x lze získat jako integrál na řezu

$$A_x = \int_{\Omega_x} f(x, y) \, dy,$$

dojdeme ke klíčové identitě.



$$\iint\limits_{\Omega} f(x,y) \, dS = \int\limits_{a}^{b} \int\limits_{\Omega} f(x,y) \, dy \, dx.$$

Integrálu napravo se říká **dvojnásobný integrál**. Procesu, kdy pro daný dvojný integrál najdeme nějaký přepis tohoto druhu, říkáme **sestavení dvojnásobného integrálu**.

Jak se takový integrál počítá? Je to vlastně integrál z integrálu, které jsou do sebe "vnořeny", což znamená, že jej interpetujeme zvenčí dovnitř: integrály na začátku bereme zleva a diferenciály na konci zprava. Dalo by se to vyjádřit závorkou, ale obvykle se to z lenosti nedělá.

$$\iint_{\Omega} f(x,y) dS = \int_{a}^{b} \left(\int_{\Omega_{x}} f(x,y) dy \right) dx.$$

Vidíme tedy **vnější integrál**, ten pracuje s proměnnou x a patří k ní integrační meze a, b. Aplikuje

se na výraz mezi, což je zase integrál. Je jasné, že nemá smysl zkoušet to integrování sx, dokud nezjistíme, co vlastně máme integrovat. Výpočet opakovaného integrálu tedy vždy začínáme tím, že vypočítáme **vnitřní integrál** $\int\limits_{\Omega} f(x,y)\,dy$.

Jak jsme viděli, výsledkem takového integrálu je nějaký výraz sx, pak máme šanci spočítat vnější integrál $\int \cdots dx$.

Ty třídimenzionální obrázky s grafem f nám pomohly pochopit motivaci za tímto postupem, ale pro praktické výpočty je lepší pracovat jen s množinou Ω . Vzorec výše pak čteme následovně: Integrál na Ω lze nahradit integrováním na řezech Ω_x (nižší dimenze) a ten vnější integrál pak zajistí, že započítáme všechny řezy, které přispívají.

Tato redukce dimenze je klíčový princip pro počítání vícerozměrných integrálů:

 \bullet Integrál na množině Ω převedeme na integrování po jejích řezech, které mají dimenzi (o jednu) menší. Pokud je tato nová dimenze stále příliš vysoká, proces řezání zopakujeme.

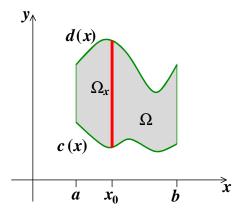
Tento obecný princip naznačuje, že řezat můžeme i jinými směry, než jsme zatím ukazovali. Šikmé řezy množinou Ω jsou principiálně možné, ale obtížně se zpracovávají. Nabízí se ale alternativa stejně příjemná jako řezy vzniklé volbou x. Můžeme si vytvářet řezy rovnoběžné s osou x, které vznikají volbou hodnoty proměnné y. Pokud si odpovídající řez množinou Ω označíme jako Ω_y , pak máme vzorec

$$\iint_{\Omega} f(x,y) dS = \int_{c}^{d} \int_{\Omega_{T}} f(x,y) dx dy.$$

Opět se počítá zevnitř, tedy tentokráte začínáme integrováním vzhledem k x.

Tyto abstraktní úvahy teď převedeme v konkrétní výpočty, což by čtenáři mělo pomoci při jejich vstřebání. Budeme se přitom muset začít zajímat o tvar integrační množiny Ω . Když je Ω velmi divoká, pak mohou být divoké i řezy Ω_x či Ω_y a může se stát, že integrování selže. Proto se v praxi omezujeme na integrační domény rozumného tvaru (neboli typu).

Jako první se podíváme na situaci, kdy řežeme ve směru y, tedy konkrétní řez zvolíme volbou $x=x_0$. V ideálním případě bude mít řez Ω_x množinou Ω tvar úsečky, což znamená, že abychom prošli všemi body množiny Ω_x , budeme pohybovat proměnnou y mezi dvěma čísly. Tato čísla, tedy integrační meze pro vnitřní integrál vzhledem k y, obecně závisí na zvoleném řezu neboli na x. Označme je c(x) a d(x). Jinak řečeno, zajímá nás případ, kdy existují funkce c(x) a d(x) takové, že řez příslušný hodnotě x je úsečka $\langle c(x), d(x) \rangle$. Množiny Ω , které takto fungují, mají specifický tvar: Jsou to oblasti uzavřené mezi dvěma grafy funkcí nad intervalem:



Někteří lidé jim říkají množiny typu I. Matematicky řečeno, zajímají nás množiny, které lze

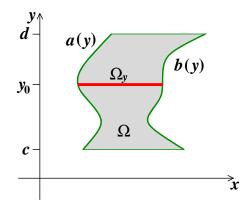
vyjádřit takto:

$$\Omega = \{(x,y) \in \mathbb{R}^2; \ a \le x \le b \ \text{a} \ c(x) \le y \le d(x)\}.$$

Pro takovéto množiny je integrování přes řezy Ω_x přímočaré a dostáváme následující dvojnásobný integrál:

$$\iint\limits_{\Omega} f(x,y)\,dS = \int\limits_{a}^{b} \int\limits_{c(x)}^{d(x)} f(x,y)\,dy\,dx.$$

Druhý směr řezání pracuje s volbou y a řezy Ω_y určující integrační množinu pro vnitřní integrál s proměnnou x. Zde opět budeme chtít, aby byly řezy Ω_y intervaly s rozumnými mezemi, což vede na typ množiny, kterým někteří říkají množiny typu II.



Formálně jde o množiny ve tvaru

$$\Omega = \{(x,y) \in \mathbb{R}^2; \ c \le y \le d \ \text{a} \ a(y) \le x \le b(y)\}.$$

Integrování přes takovouto množinu lze pak převést na dvojnásobný integrál

$$\iint\limits_{\Omega} f(x,y) \, dS = \int\limits_{c}^{d} \int\limits_{a(y)}^{b(y)} f(x,y) \, dx \, dy.$$

Všimněte si u obou přepisů důležitého prvku, který platí obecně. Správně sestavený vícenásobný integrál mimo jiné splňuje následující: Vnější integrál má jako meze pouze čísla. Vnitřní integrál může mít jako meze čísla nebo výrazy s proměnnou, v tom případě to ale musí být proměnná, která je v seznamu diferenciálů vně, tedy musí jít o proměnnou, vůči které se bude integrovat později.

Některé množiny vyhovují oběma specifikacím, pak máme na výběr, kterým směrem budeme řezat. V takových případech se obvykle rozhoduje podle toho, které řezy vedou na příjemnější integraci.

Příklad: Budeme integrovat funkci $f(x,y) = y e^{-x}$ přes interval $\langle 0,1 \rangle \times \langle 1,3 \rangle$ neboli na množině $\Omega = \{(x,y) \in \mathbb{R}^2; \ 0 \le x \le 1 \ \text{a} \ 1 \le y \le 3\}.$

Máme na výběr, pomocí které proměnné integrovat nejdříve, protože v obdélníku budou obojí směry řezů fungovat stejně dobře.

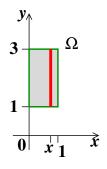
Zkusíme fixovat x a nejprve integrovat podle y (tedy řezat množinu rovnoběžně s osou y), protože už jsme si to v předchozím příkladě připravili.

Již jsme spočítali integrál přes takový řez,

$$\int_{1}^{3} y e^{-x} dy = \left[e^{-x} \frac{1}{2} y^{2} \right]_{y=1}^{y=3} = 4e^{-x}.$$

Tyto integrály přes řezy pak "posčítáme" pomocí x, dostáváme

$$\iint\limits_{\Omega} y \, e^{-x} \, dS = \int\limits_{0}^{1} \int\limits_{1}^{3} y \, e^{-x} \, dy \, dx = \int\limits_{0}^{1} 4e^{-x} \, dx = \left[-4e^{-x} \right]_{0}^{1} = 4 - 4e^{-1}.$$



Teď zkusíme ten druhý možný způsob řezání. Jak to vypadá při vodorovných řezech, tedy ve směru osy x?

Řezy jsou neprázdné pro y mezi 1 a 3, dostáváme tak rozklad

$$\iint_{\Omega} y e^{-x} dS = \int_{1}^{3} \int_{\Omega_{y}} y e^{-x} dx dy.$$

Z obrázku vyčteme, že na řezu se proměnná x vždy pohybuje mezi 0 a 1, což určí vnitřní integrál.

$$\iint_{\Omega} y \, e^{-x} \, dS = \int_{1}^{3} \int_{0}^{1} y \, e^{-x} \, dx \, dy.$$

Počítáme zevnitř:
$$\iint_{\Omega} y \, e^{-x} \, dS = \int_{1}^{3} \int_{0}^{1} y \, e^{-x} \, dx \, dy = \int_{1}^{3} \left[-y \, e^{-x} \right]_{x=0}^{x=1} dy = \int_{1}^{3} y - e^{-1} y \, dy$$
$$= \int_{1}^{3} (1 - e^{-1}) y \, dy = \left[(1 - e^{-1}) \frac{1}{2} y^{2} \right]_{1}^{3} = 4(1 - e^{-1}).$$

Při dosazování do vnitřního integrálu jsme si pro jistotu připomněli, která proměnná byla v té chvíli pracovní, tedy za kterou máme dosazovat.

 \triangle

Porovnejme oba způsoby přepisu daného dvojného integrálu v příkladě výše:

$$\iint_{\Omega} y e^{-x} dS = \int_{0}^{1} \int_{1}^{3} y e^{-x} dy dx = \int_{1}^{3} \int_{0}^{1} y e^{-x} dx dy.$$

Vypadá to, že jsme prostě jen prohodili pořadí integrálů a odpovídajících diferenciálů. Takto snadno to ovšem funguje pouze u integrace přes obdélníky, které jsou tak nejoblíbenějšími integračními množinami. Jinak je tzv. "záměna pořadí integrace" poněkud složitější.

Příklad: Budeme integrovat funkci $f(x,y) = e^{x^2}$ přes omezenou oblast Ω vymezenou křivkami y = 2x, x = 3 a y = 0.

Vždy se vyplatí nakreslit obrázek. Dané tři přímky rozdělí rovinu na sedm částí, ale jen jedna z nich je omezená: trojúhelník s vrcholy (0,0), (3,0) a (3,6).

Svislé řezy (ve směru osy y) budou určitě fungovat. Vidíme, že pro určení polohy řezu má smysl volit x jen z rozmezí 0 až 3, výsledný dvojnásobný integrál tedy bude mít vnější integrál $\int_{0}^{3} \cdots dx$.

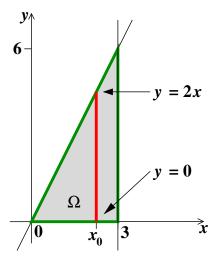
Na jednom konkrétním řezu se pak proměnná y pohybuje mezi hodnotami y=0 a y=2x. To ostatně odpovídá formálnímu přepisu množiny ve tvaru

$$\Omega = \{(x, y) \in \mathbb{R}^2; \ 0 \le x \le 3 \text{ a } 0 \le y \le 2x\}.$$

Tím je jasné, jak bude vypadat vnitřní integrál po typickém řezu. Můžeme počítat.

$$\iint_{\Omega} e^{x^2} dS = \int_{0}^{3} \int_{0}^{2x} e^{x^2} dy dx = \int_{0}^{3} \left[y e^{x^2} \right]_{y=0}^{y=2x} dx = \int_{0}^{3} 2x e^{x^2} dx$$

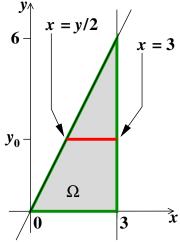
$$= \begin{vmatrix} w = x^2 \\ dw = 2x dx \\ x = 0 \implies w = 0 \\ x = 3 \implies w = 9 \end{vmatrix} = \int_{0}^{9} e^w dw = \left[e^w \right]_{0}^{9} = e^9 - 1.$$



Nyní je čas na řezy ve vodorovném směru.

Poloha řezu je dána volbou y v rozmezí 0 až 6, tím je jasný vnější integrál. Řez (na kterém je pracovní proměnná x) je úsečka s pravým koncem na úrovni x=3, levý konec leží na křivce y=2x a nás zajímá, kolik je tam x. Vidíme tedy, že při pohybu podél takového řezu se x mění od $\frac{y}{2}$ do 3. Dostáváme

$$\iint_{\Omega} e^{x^2} dS = \int_{0}^{6} \int_{y/2}^{3} e^{x^2} dx dy.$$



Začneme vnitřním integrálem $\int_{y/2}^3 e^{x^2} dx$. A hned máme problém, primitivní funkci k e^{x^2} nelze vyjádřit uzavřeným algebraickým vzorcem, takže jsme skončili.

 \triangle

Vidíme, že někdy volba směru řezu neboli volba pořadí při integrování může mít vliv na to, jak to dopadne. V méně extrémním případě může ovlivnit složitost výpočtu.

Ačkoliv jeden výpočet nevyšel, přesto to bylo užitečné cvičení. Oba dva přepisy do dvojnásobného integrálu byly zcela v pořádku a ukazují, že se při změně pořadí integrace meze mohou také změnit.

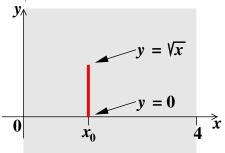
$$\iint\limits_{\Omega} e^{x^2} dS = \int_0^3 \int_0^{2x} e^{x^2} dy dx = \int_0^6 \int_{y/2}^3 e^{x^2} dx dy.$$

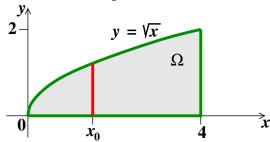
To je typické, při změně pořadí se musí meze znovu přepočítat. Někdy se potkáme s úlohou, že je nám zadán dvojnásobný integrál a máme změnit pořadí integrace (například proto, že v daném pořadí nelze najít primitivní funkci). V takovém případě je třeba nejdříve z tvaru mezí odvodit, jak vlastně vypadá integrační doména Ω (jakýsi "reverse engineering"), a na ní pak aplikovat řezy druhým směrem.

Příklad: Zaměňte pořadí integrace v integálu

$$\int_{0}^{4} \int_{0}^{\sqrt{x}} f \, dy \, dx.$$

Nejprve musíme identifikovat integrační doménu Ω . Vnější integrál ukazuje meze pro x, což nám říká dvě věci: Bereme řezy kolmé na osu x, tedy svislé, a ty jsou relevantní pouze pro x mezi 0 a 4. To znamená, že množina Ω musí ležet v nekonečném svislém pruhu daném $0 \le x \le 4$.

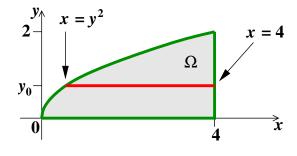




Teď vezmeme nějaké $x \in \langle 0,4 \rangle$ a podíváme se na rozsah, který je pro y na příslušném řezu povolen. Vidíme, že hodnoty začínají na y=0 a končí na $y=\sqrt{x}$. Jinými slovy, spodní hrana množiny je dána vzorcem y=0 a horní hrana je dána $y=\sqrt{x}$.

Jsme připravení a řezy v kolmém směru. Vodorovný řez vybereme volbou y z intervalu $\langle 0,2\rangle$. Na tomto konkrétním řezu je hodnota pro x nejvíce vlevo dána vzorcem $y=\sqrt{x}$, takže se x mění mezi $x=y^2$ a 4. Můžeme napsat žádaný integrál.

$$\int_{0}^{4} \int_{0}^{\sqrt{x}} f \, dy \, dx = \int_{0}^{2} \int_{x^2}^{4} f \, dx \, dy.$$



 \triangle

Poznamenejme, že takováto přímočará změna není vždy možná, protože pro některé množiny nevedou řezy jistým směrem na jeden výsledný integrál. Uvidíme to v následujícím příkladě.

Zvídavého čtenáře jistě napadlo, že ne všechny množiny Ω spadají do dvou probraných typů. V takovém případě obvykle není problém dotyčnou oblast rozdělit na podmnožiny, které jsou již vhodného typu, a pro každou z nich integrujeme zvlášť. Toto je možné díky aditivitě integrálu vzhledem k integrační množině. Čtenář zná vzorec

$$\int_{a}^{c} f(x) dx = \int_{a}^{b} f(x) dx + \int_{b}^{c} f(x) dx,$$

který pro a < b < c říká, že si integrační interval můžeme rozdělit na dvě části. Ty by se neměly překrývat, na druhou stranu rozhodně neplatí, že by intervaly $\langle a,b \rangle$ a $\langle b,c \rangle$ měly prázdný průnik. Potřebujeme povolit, aby dvě množiny mohly sdílet hranici, což se dá matematicky zařídit tak, že z podmínky disjunktnosti tu hranici vynecháme. Dostáváme následující princip.

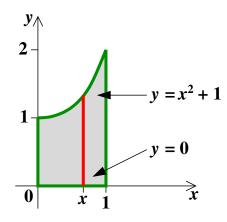
• Uvažujme množinu Ω , kterou si rozdělíme na podmnožiny $\Omega_1, \ldots, \Omega_m$ tak, že $\Omega = \Omega_1 \cup \cdots \cup \Omega_m$ a pro libovolné $j \neq k$ jsou vnitřky množin Ω_j, Ω_k disjunktní. Pak

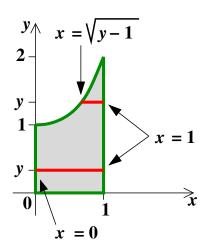
$$\iint_{\Omega} f(\vec{x}) d[\vec{x}] = \sum_{k=1}^{m} \iint_{\Omega_k} f(\vec{x}) d[\vec{x}].$$

85

Příklad: Nechť Ω je oblast mezi $y = x^2 + 1$ a osou x na (0,1). Chceme najít $\iint_{\Omega} 8x(x^2 - y + 1)^3 dS$.

Nejprve si nakreslíme Ω .





Tvar oblasti přímo vyzývá ke svislým řezům, kdy nejprve integrujeme podle y. Kdybychom se totiž pokusili o řezy ve směru x, tak bychom neměli jednotné zadání pro levé konce řezů.

Raději tedy volíme svislé řezy určené volbou x, zjevně z intervalu (0,1), a vevnitř budeme integrovat na řezu pomocí y s mezemi, které vidíme z obrázku. Dostáváme dvojnásobný integrál

$$\iint\limits_{\Omega} 8x(x^2 - y + 1)^3 dS = \int\limits_{0}^{1} \int\limits_{0}^{x^2 + 1} 8x(x^2 - y + 1)^3 dy dx.$$

Jak zpracujeme vnitřní integrál? Když integrujeme funkci $8x(x^2-y+1)^3$ podle y, tak je 8x vlastně konstanta a dá se z integrálu vytknout. Při integrování výrazu $(x^2-y+1)^3$ podle y je i x^2+1 jakoby konstanta, čili vlastně integrujeme výraz ve tvaru $(-y+a)^3$, který zkušení integrátoři zvládnou zpaměti, ti opatrní či méně zkušení si pomohou substitucí:

$$8x \int (x^2 - y + 1)^3 dy = \begin{vmatrix} w = -y + x^2 + 1 \\ dw = \frac{\partial}{\partial y} [-y + x^2 + 1] dy = -dy \end{vmatrix}$$
$$= -8x \int w^3 dw = -2xw^4 + C = -2x(x^2 - y + 1)^4 + C.$$

Dostáváme

$$\iint_{\Omega} 8x(x^2 - y + 1)^3 dS = \int_0^1 \int_0^{x^2 + 1} 8x(x^2 - y + 1)^3 dy dx = \int_0^1 \left[-2x(x^2 - y + 1)^4 \right]_{y=0}^{y=x^2 + 1} dx$$

$$= \int_0^1 0 - (-2x(x^2 + 1)^4) dx = \int_0^1 2x(x^2 + 1)^4 dx = \left| \frac{w = x^2 + 1}{dw = 2x dx} \right|$$

$$= \left[\frac{1}{5} (x^2 + 1)^5 \right]_0^1 = \frac{2^5}{5} - \frac{1}{5} = \frac{31}{5}.$$

Co kdybychom z nějakého důvodu mermomocí chtěli řezat ve směru x? Zatímco pravý konec řezů se stále řídí stejným předpisem x=1, levý konec záleží na tom, kde řežeme. Pro řezy dané $y \in \langle 0,1 \rangle$ bude levý konec x=0, budeme tedy tuto část množiny Ω (spodní polovinu) integrovat zvlášť. V horní polovině, tedy pro řezy dané $y \in \langle 1,2 \rangle$, leží levé konce řezů na křivce $y=x^2+1$. Proto má x na takovém řezu počáteční hodnotu $\sqrt{y-1}$. Sestavíme tedy naše integrování jako

součet dvou integrálů.

$$\iint_{\Omega} 8x(x^2 - y + 1)^3 dS = \int_{0}^{1} \int_{0}^{1} 8x(x^2 - y + 1)^3 dx dy + \int_{1}^{2} \int_{\sqrt{y - 1}}^{1} 8x(x^2 - y + 1)^3 dx dy$$

$$= \begin{vmatrix} w = x^2 - y + 1 \\ dw = \frac{\partial}{\partial x} [x^2 - y + 1] dx = 2x dx \end{vmatrix}$$

$$= \int_{0}^{1} \left[(x^2 - y + 1)^4 \right]_{x = 0}^{x = 1} dy + \int_{1}^{2} \left[(x^2 - y + 1)^4 \right]_{x = \sqrt{y - 1}}^{x = 1} dy$$

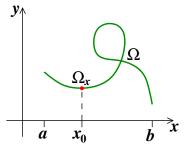
$$= \int_{0}^{1} (2 - y)^4 - (1 - y)^4 dy + \int_{1}^{2} (2 - y)^4 - 0 dy$$

$$= \left[-\frac{1}{5} (2 - y)^5 + \frac{1}{5} (1 - y)^5 \right]_{0}^{1} + \left[-\frac{1}{5} (2 - y)^5 \right]_{1}^{2} = \frac{31}{5}.$$

Uf, vyšlo to stejně.

Poznámka: Tuto sekci uzavřeme pohledem na jednu zajímavou situaci. Všichni známe vzorec $\int_a^a f(x) \, dx = 0$. Co nám to říká? Počítáme jednorozměrný integrál na intervalu, který je jen jeden bod, takže je to množina dimenze nula. Integrační interval tvoří základnu pro obrazec pod grafem f, jehož obsah chceme najít, a pokud má tato základna délku nula, pak ten obsah je samozřejmě také nula.

Stejný princip platí také pro dvourozměrný integrál. Může se stát, že množina Ω sice leží v \mathbb{R}^2 , ale její skutečná dimenze je nižší. Typický příkladem je úsečka. Na ní se lze pohybovat jen tam a zpět v jednom směru, tedy má jen jednu dimenzi. Dá se také říct, že pro pohyb nabízí jen jeden stupeň volnosti. Další významný indikátor je, že úsečka má jako množina v \mathbb{R}^2 prázdný vnitřek. My jsme zvyklí, že vnitřkem intervalu $\langle a,b\rangle$ je interval (a,b), ale to platí jen v jednom rozměru. Ve dvou rozměrech je nějaký bod vnitřním bodem množiny, pokud v ní leží také s nějakým svým okolím, což je ve dvou dimenzích kruh. Na úsečce v rovině takové body nenajdeme. Stejný argument platí pro všechny rozumné křivky v rovině obecně.



Pokud máme takovou množinu s prázdným vnitřkem, pak jsou řezy Ω_x (nebo Ω_y) typicky jednobodové nebo několik bodů. Pak jsou jednorozměrné integrály \int_{Ω_x} automaticky nula a tedy i celý dvojný integrál vyjde nulový.

Takže když se pokusíme integrovat dvojným integrálem přes množinu, která není dostatečně "masivní", například když má dvourozměrný vnitřek nulový, tak integrál automaticky vyjde nula. Typické množiny tohoto typu jsou křivky, třeba úsečky, oblouky, kružnice a tak podobně.

 \triangle

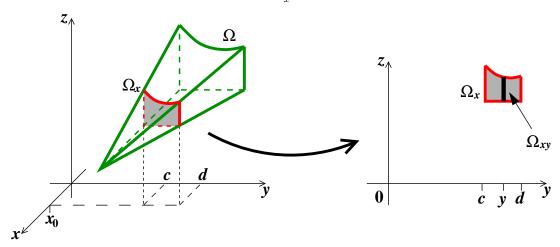
6b. Třídimenzionální integrál

Trojný integrál $\iiint\limits_{\Omega} f(x,y,z)\,dV$ počítáme obdobně jako dvojný. Potřebujeme projít všechny body

tělesa Ω , což si usnadníme organizací do řezů. Jinak řečeno, postupně po jedné zmenšujeme dimenzi množiny Ω .

Když si zvolíme pevně hodnotu jedné proměnné, například x jako x_0 , tak to ve třírozměrném prostoru vybere rovinu kolmou na osu x, která řízne těleso Ω a vytvoří dvourozměrný řez, nazvěme ho Ω_x . Pro rozumné těleso Ω jsou tyto řezy neprázdné pouze pro x z nějakého rozmezí $\langle a,b\rangle$, což určí první etapu sestavení integrálu. Vnějším integrálem zajistíme, že projdeme všechny relevantní řezy. Vnitřním integrálem se pak postaráme o konkrétní řez, což je dvourozměrná množina, tedy vnitřní integrál bude dvourozměrný. Po řezu se pohybujeme změnou y a z. Můžeme psát

$$\iiint\limits_{\Omega} f(x,y,z) \, dV = \int_a^b \iint\limits_{\Omega_x} f(x,y,z) \, dS[y,z] \, dx.$$



Dalším krokem je vypracování vnitřního integrálu přes dvourozměrnou množinu. Použijeme obvyklý postup, tedy budeme řezat řezy Ω_x . Například při troše štěstí dokážeme najít rozumné řezy množinou Ω_x ve směru z, tedy určené volbou y. Můžeme je značit Ω_{xy} . Rozsah pro y teď závisí na tvaru konkrétního řezu Ω_x , takže pro ten další integrál budeme uvažovat y mezi jistými hodnotami c(x) a d(x). Máme

$$\iint_{\Omega_x} f(x, y, z) dS[y, z] = \int_{c(x)}^{d(x)} \int_{\Omega_{xy}} f(x, y, z) dz dy.$$

To zredukuje dvojný integrál na integrál přes jednorozměrný řez, s trochou štěstí to bude interval, jehož koncové body závisí na y a to zase na x. Dostaneme tak

$$\iiint_{\Omega} f(x, y, z) \, dV = \int_{a}^{b} \int_{c(x)}^{d(x)} \int_{g(x, y)}^{h(x, y)} f(x, y, z) \, dz \, dy \, dx.$$

Samozřejmě nevíme, zda zrovna toto pořadí řezů bude fungovat pro náš tvar tělesa Ω . Je celkem šest možností pořadí, alespoň jedna by mohla zabrat.

Je zjevné, že pro zdárné sestavení takového integrálu je třeba umět se dobře orientovat v prostoru a mít zvládnutou analytickou práci s geometrickými útvary.

Příklad: Spočítáme průměrnou hodnotu funkce f(x, y, z) = 8xy(4 - z) na množině

$$\Omega = \{(x, y, z) \in \mathbb{R}^3; \ x, y \ge 0, \ 0 \le z \le 4 - 4\sqrt{x^2 + y^2}\}.$$

Nejprve je třeba si rozmyslet, jak tato množina vypadá. Jako první si všimneme, že máme omezení

 $x,y,z\geq 0$, takže půjde o útvar v hlavním kvadrantu. Podmínka na z nám říká, že Ω je těleso mezi dvěma plochami. Ta spodní je vodorovná rovina z=0, bude tvořit základnu tělesa. Horní povrch je dán rovnicí $z=4-4\sqrt{x^2+y^2}$. Co je to za povrch?

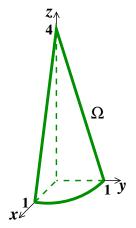
Je to vlastně transformace jedné ze základních rovnic. Rovnice $z=\sqrt{x^2+y^2}$ dává kužel (výška v bodě (x,y) je dána jeho vzdáleností od počátku) se špičkou v počátku a rozevírající se směrem vzhůru, jeho osa souhlasí s osou z.

Když zavedeme násobící konstantu, vznikne rovnice $z=4\sqrt{x^2+y^2}$, která čtyřikrát zrychlí růst kužele. Je to tedy kužel výrazně špičatější, jeho strana už nemá stoupání 1:1, ale 4:1. Změna znaménka $z=-4\sqrt{x^2+y^2}$ tento kužel obrátí směrem dolů a přičtení čtyřky posune nahoru.

Dostáváme tedy kužel, který má špičku nahoře v bodě (0,0,4) a stěny mu docela příkře padají dolů. Tento kužel vymezuje naše těleso shora, máme také omezení rovinou z=0 zdola. Kde se protnou?

$$0 = 4 - 4\sqrt{x^2 + y^2} \implies x^2 + y^2 = 1.$$

Z toho plyne, že podmínka $0 \le z \le 4 - 4\sqrt{x^2 + y^2}$ určuje kužel, který má vrchol v bodě (0,0,4) a jehož podstava je dána jednotkovou kružnicí v rovině xy se středem (0,0,0). Z tohoto kužele vyřízneme čtvrtinu, která se nachází v hlavním kvadrantu.

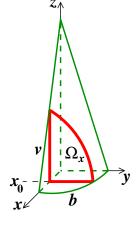


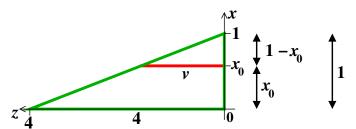
Přes tuto množinu bude třeba integrovat, takže si rozmyslíme řezy. V obecném úvodu jsme začali s x, takže to vyzkoušíme. Je zjevné, že má smysl brát jen x z rozmezí $\langle 0,1 \rangle$. Pro jisté takové x pak vznikne svislý řez Ω_x rovnoběžný s rovinou yz. Dostáváme první redukci pro integrál,

$$\iiint_{\Omega} f(x, y, z) dV = \int_{0}^{1} \iint_{\Omega_{x}} f(x, y, z) dS dx.$$

Řez Ω_x je dvourozměrný, takže potřebujeme integrál přes tuto množinu dále redukovat. Jak tato množina vypadá?

Co vidíme hned je, že tento útvar má svislou výšku a vodorovnou základnu, které jsou přímé, jde tedy o úsečky. Jak jsou dlouhé? Výšku v odvodíme bližším pohledem na svislou trojúhelníkovou stěnu čtvrtkužele, která leží v rovině xz, pro úsporu místa jsme ji položili na bok.





Podobnost trojúhelníků dává

$$\frac{v}{1-x} = \frac{4}{1} \implies v = 4(1-x).$$

Velikost základny b zase odvodíme z toho, co se děje v základně kužele, ta je dána kružnicí a proto $b = \sqrt{1 - x^2}$. Tyto dva vzorce nám také udávají rozumné hodnoty pro proměnné z, popřípadě y pro další krok, kdy budeme chtít tento řez dále krájet.

Tyto další řezy už budou úsečky. Pro jejich přesné určení potřebujeme znát přesný tvar "přepony", jenže ona to žádná přepona není, svislý řez kuželem mimo osu bohužel nedává trojúhelníky. Tato třetí strana je dána rovnicí kužele, ve kterém vezmeme x jako parametr (pro vytvoření řezu jsme

zvolili pevnou hodnotu $x = x_0$). Proměnné y, z jsou tedy na oné třetí straně řezu svázány vzorcem

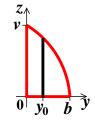
$$z = 4 - 4\sqrt{x_0^2 + y^2}$$

neboli

$$(4-z)^2 - y^2 = x_0^2.$$

Třetí strana řezu má tedy tvar hyperboly.

Jsme připravení sestavit integrál pro integraci na Ω_x , zkusíme svislé řezy. To znamená, že volíme $y=y_0$ z rozmezí 0 až $b=\sqrt{1-x_0^2}$ a vznikne svislá úsečka, na které se z mění mezi 0 a $4-4\sqrt{x_0^2+y_0^2}$. Jinak řečeno, náš trojný integrál se změní v trojnásobný integrál



$$\iiint_{\Omega} f(x, y, z) dV = \int_{0}^{1} \int_{0}^{\sqrt{1 - x^{2}}} \int_{0}^{4 - 4\sqrt{x^{2} + y^{2}}} f(x, y, z) dz dy dx.$$

Ke spočítání průměru potřebujeme integrovat danou funkci. Začneme vnitřním integrálem a postupně se propracujeme ven.

$$\iiint_{\Omega} 8xy(4-z) \, dV = \int_{0}^{1} \int_{0}^{\sqrt{1-x^2}} \int_{0}^{4-4\sqrt{x^2+y^2}} 8xy(4-z) \, dz \, dy \, dx = \left| \begin{array}{c} w = 4-z \\ dw = -dz \end{array} \right| \\
= \int_{0}^{1} \int_{0}^{\sqrt{1-x^2}} \left[-4xy(4-z)^2 \right]_{z=0}^{z=4-4\sqrt{x^2+y^2}} \, dy \, dx \\
= \int_{0}^{1} \int_{0}^{\sqrt{1-x^2}} 4^3xy - 4^3xy(x^2+y^2) \, dy \, dx = \left| \begin{array}{c} w = x^2+y^2 \\ dw = 2y \, dy \end{array} \right| \\
= \int_{0}^{1} \left[2 \cdot 4^2xy^2 - 4^2x(x^2+y^2)^2 \right]_{y=0}^{y=\sqrt{1-x^2}} \, dx \\
= \int_{0}^{1} 2 \cdot 4^2x(1-x^2) - 4^2x + 4^2x^5 \, dx = 4^2 \int_{0}^{1} x - 2x^3 + x^5 \, dx \\
= 4^2 \left[\frac{1}{2}x^2 - \frac{1}{2}x^4 + \frac{1}{6}x^6 \right]_{0}^{1} = 4^2 \frac{1}{6} = \frac{8}{3}.$$

Alternativa: Ze symetrie je docela jasné, že vést první řez kolmo na osu y, rovnoběžně s rovinou xz, by vedlo na stejné tvary řezu, x a y jsou zaměnitelné i ve funkci f, takže by následovaly stejné výpočty, jen s prohozenou rolí x a y. Pokud chceme vidět něco jiného, zkusíme vodorovné řezy.

Zvolme tedy nějakou hodnotu $z=z_0$, to má smysl pro rozmezí $\langle 0,4\rangle$. Vznikne tak řez, který by měl mít podobu čtvrtiny kruhu.

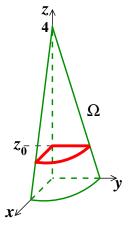
Potvrdíme to tak, že v rovnici pláště

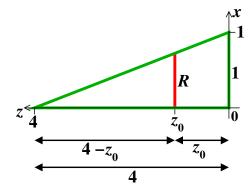
$$z = 4 - 4\sqrt{x^2 + y^2}$$

považujeme $z=z_0$ za parametr. Dostáváme pak

$$x^2 + y^2 = \left(1 - \frac{1}{4}z_0\right)^2.$$

Toto je tvar vypouklé strany onoho řezu a je to opravdu kružnice, má poloměr $R=1-\frac{1}{4}z_0$. To druhé lze nezávisle potvrdit i z podobnosti trojúhelníků na levé svislé stěně našeho čtvrtkužele.

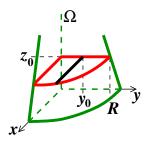




Dostáváme

$$\frac{R}{4-z} = \frac{1}{4} \implies R = \frac{1}{4}(4-z).$$

Uvažujme teď jeden takový čtvrtkruh. Rozložíme jej na řezy tvaru úsečky. Pokud zvolíme hodnotu $y=y_0$ z rozmezí 0 až $R=\frac{1}{4}(4-z)$, vznikne řez neboli úsečka rovnoběžná s osou x a možné hodnoty pro x jsou v rozmezí 0 až $\sqrt{R^2-y_0^2}$ neboli $\sqrt{\frac{1}{4^2}(4-z_0)^2-y_0^2}$. Vznikne tak převod trojného integrálu na opakovaný integrál



$$\iiint_{\Omega} f(x,y,z) \, dV = \int_{0}^{4} \int_{0}^{(z-4)/4} \int_{0}^{\sqrt{(4-z)^2/16-y^2}} f(x,y,z) \, dx \, dy \, dz.$$

Zase budeme integrovat danou funkci.

$$\iiint_{\Omega} 8xy(z-4) \, dV = \int_{0}^{4} \int_{0}^{(z-4)/4} \int_{0}^{\sqrt{(4-z)^{2}/16-y^{2}}} 8xy(z-4) \, dx \, dy \, dz$$

$$= \int_{0}^{4} \int_{0}^{(z-4)/4} \left[4x^{2}y(z-4) \right]_{x=0}^{x=\sqrt{(4-z)^{2}/16-y^{2}}} dy \, dz$$

$$= \int_{0}^{4} \int_{0}^{(z-4)/4} 4\left(\frac{1}{16}(4-z)^{2} - y^{2}\right) y(4-z) \, dy \, dz.$$

Zde se nabízejí dvě možnosti. Lze roznásobit první závorku a pak integrovat zvlášť y a y^3 . Zajímavější myšlenka je použít substituci, protože celý člen $\frac{1}{16}(4-z)^2$ je v této chvíli vlastně konstanta. Položíme-li $w=\frac{1}{16}(4-z)^2-y^2$, dostáváme $-2y\,dy=dw$, takže počítáme

$$\int 4\left(\frac{1}{16}(4-z)^2 - y^2\right)y(4-z)\,dy = \int -2w(4-z)dw = -w^2(4-z) = -\left(\frac{1}{16}(4-z)^2 - y^2\right)^2(4-z).$$

[6c]

Máme tedy

$$\iiint_{\Omega} 8xy(z-4) \, dV = \int_{0}^{4} \int_{0}^{(z-4)/4} 4\left(\frac{1}{16}(4-z)^{2} - y^{2}\right)y(4-z) \, dy \, dz$$

$$= \int_{0}^{4} \left[-\left(\frac{1}{16}(4-z)^{2} - y^{2}\right)^{2}(4-z) \right]_{y=0}^{y=(z-4)/4} dz$$

$$= \int_{0}^{4} 0 - \left[-\left(\frac{1}{16}(4-z)^{2} - 0\right)^{2}(4-z) \right] dz$$

$$= \int_{0}^{4} \frac{1}{4^{4}}(4-z)^{5} \, dz = \left[\frac{1}{-6 \cdot 4^{4}}(4-z)^{6} \right]_{0}^{4} = \frac{16}{6} = \frac{8}{3}.$$

Potvrdili jsme výsledek.

Ke spočítání průměru potřebujeme ještě znát objem tělesa Ω . Obecně bychom mohli integrovat funkci 1 přes Ω a máme teď na to dva přístupy. Je tu ale zajímavá alternativa.

Vychází z interpretace, že objem je součet obsahů řezů. To nás navádí na vodorovné řezy, které dávaly čtvrtkruhy. Pokud si označíme obsah takového řezu A_z , můžeme formálně psát

$$\operatorname{vol}(\Omega) = \iiint_{\Omega} 1 \, dV = \int_{0}^{4} \iint_{\Omega_{z}} 1 \, dS \, dz = \int_{0}^{4} A_{z} \, dz.$$

Spočítali jsme, že čtvrtkruh odpovídající volbě $z=z_0$ má poloměr $R=\frac{1}{4}(4-z_0)$, takže $A_z=\frac{1}{4}\pi\left(\frac{1}{4}(4-z)\right)^2$. Odtud

$$\operatorname{vol}(\Omega) = \int_{0}^{4} \pi \frac{1}{4^{3}} (4 - z)^{2} dz = \left[-\pi \frac{1}{3 \cdot 4^{3}} (4 - z)^{3} \right]_{0}^{4} = \frac{\pi}{3}.$$

Popravdě řečeno, vlastně stačilo vzít standardní vzorec pro objem kužele a vydělit čtyřmi, vyšlo by to stejně. Ale to bychom se zase nenaučili nic nového.

Nyní již spočítáme průměr:

$$\operatorname{Ave}_{\Omega}(8xy(4-z)) = \frac{\iint 8xy(4-z) \, dV}{\operatorname{vol}(\Omega)} = \frac{8}{\pi}.$$

 \triangle

Tento příklad ukazuje, že samotná myšlenka integrování ve více rozměrech není zase tak obtížná, ale rozebrat zadané těleso na řezy nemusí být snadné, pokud si řešitel opravdu dobře netyká s geometrií ve 3D.

Poznámka: Sekci uzavřeme zopakováním poznámky, kterou jsme uzavřeli předchozí sekci. Pokud se pokusíme integrovat funkci f(x, y, z) přes množinu z \mathbb{R}^3 , která ale není opravdu třírozměrná, například když má prázdný třídimenzionální vnitřek, tak to automaticky vyjde nula.

 \triangle

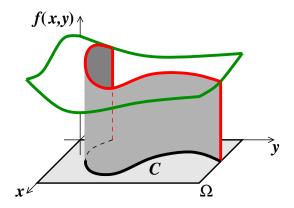
6c. Křivkové a plošné integrály

Představme si plotnu sporáku, na které je nějak rozložená stabilní teplota f(x,y). Pokud si reprezentujeme tuto plotnu množinou Ω , tak snadno najdeme průměrnou teplotu plotny pomocí dvojného integrálu přes množinu Ω .

Teď si ale představme, že po té plotně přelezl brouk (s azbestovýma bačkůrkama) a cesta, kterou šel, vytváří rovinnou křivku C. Brouk je přirozeně zvědavý, jakou průměrnou teplotu během své procházky zažil. Protože C nejspíše nebude přímka, ale nějak zatáčí a kroutí se, tak brouk nemůže použít běžný jednorozměrný integrál. Na druhou stranu je tato křivka svou podstatou jednorozměrná, takže pokud se brouk rozhodne spočítat dvojný integrál teploty přes tuto křivku C, vyjde automaticky nula.

Broukova otázka je ovšem zcela legitimní a ukazuje nám, že potřebujeme vymyslet nějaký nový typ integrálu. Jak by měl fungovat? Situace je vlastně následující:

Máme funkci f(x,y) definovanou na plošné množině Ω . Nás ale zajímají jen hodnoty f na nějaké křivce C. Máme v úmyslu je "sečíst" integrací, takže vlastně hledáme obsah obsah řezu, který křivka C vytvoří v tělese pod grafem f. Tento řez je svou podstatou dvourozměrný, takže použití obsahu je správné, a není plochý, ale zakřivený. Obojí je třeba vzít v úvahu, když vytváříme nový typ integrálu.



Když vezmeme nějaký bod \vec{x} na křivce C, tak nemá smysl zkoušet kolem něj umístit rovinnou infinitesimálu, protože to by vedlo na 3D sloupec s objemem, ale my potřebujeme plochý objekt s obsahem. Proto si místo toho vezmeme "délkovou infinitesimálu" ve směru křivky, tedy nekonečně krátký úsek dané křivky okolo \vec{x} o délce ds. Toto bude základna, na které vztyčíme obdélník o výšce $f(\vec{x})$. Coby plochý objekt má obsah, jmenovitě $f(\vec{x}) \cdot ds$. Když tyto příspěvky posčítáme integrálem, dostaneme číslo, které zajímalo toho brouka, plochu onoho zakřiveného řezu.

Takovému číslu se říká **křivkový integrál** a značíme ho

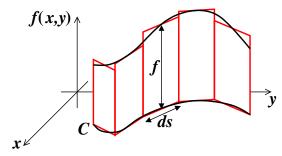
$$\int_{C} f(\vec{x}) \, ds.$$

Existuje speciální značení

$$\oint\limits_C f(\vec{x})\,ds$$

pro křivkové integrály na uzavřených křivkách (konec se napojuje na začátek)

Abychom lépe ocenili myšlenku za křivkovým integrálem, přidáme ještě příběh, který se obvykle vypráví u běžného určitého integrálu. Můžeme si představit, že jsme křivku C rozdělili na velmi malé úseky délky ds. Protože je každý takový úsek velmi malý, tak můžeme předpokládat, že se na něm křivka nestihne zatočit, takže je to vlastně úsečka, a také funkce f nestihne příliš změnit hodnotu, takže je možné ji aproximovat třeba hodnotou uprostřed úsečky. Tímto způsobem vlastně onen zakřivený povrch aproximujeme plochými obdélníkovými panely.



Když obsahy panelů posčítáme, dostaneme aproximaci obsahu řezu, přičemž je možné tuto aproximaci dle libosti zpřesnit braním velmi úzkých obdélníčků.

Všimneme si, že jsme výše ve značení křivkového integrálu použili vektor \vec{x} namísto souřadnic (x,y). Má to svůj důvod. Při naší úvaze nebylo vůbec podstatné, že křivka byla zrovna na plotně, to hlavní bylo rozdělit si ji na úseky, u kterých jsme znali délku. Myšlenka křivkového integrálu je tedy aplikovatelná na křivku v libovolném prostoru \mathbb{R}^n a funkci o n proměnných.

Představme si například mouchu poletující učebnou, kde známe teplotu f(x, y, z) ve všech bodech místnosti. Pokud by ta moucha chtěla znát průměrnou teplotu na dráze, kterou letěla (což je jednodimenzionální křivka), tak by ji zase rozdělila na kousíčky délky ds a každý by přispěl hodnotou $f(x, y, z) \cdot ds$.

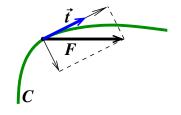
Jakmile máme zaveden křivkový integrál, dokážeme najít **délku křivky** integrováním jedničky coby konstantní funkce přes křivku, protože sčítáním čísel $1 \cdot ds$ vlastně sčítáme délky jednotlivých úseků. Průměr funkce na křivce pak získáme tak, že integrál funkce na křivce vydělíme její délkou.

Důležitá aplikace křivkového integrálu je spojena s vektorovými funkcemi. Operace spojené s derivováním a integrováním se obvykle aplikují na vektorové funkce po jednotlivých souřadnicích. Takto bychom také interpretovali křivkový integrál $\int_C F(\vec{x}) \, ds$, výsledkem by pak byl zase vektor.

V praxi se ale spíše potřebuje něco jiného.

Představme si, že vektorová funkce F popisuje působící sílu, kvůli které se nějaké těleso pohybuje po dráze C. Jaká práce se pak vykonala? Pokud je síla konstantní a ve směru přímočarého pohybu, tak stačí pracovat s její velikostí ||F|| a vzorcem ze základní školy, že práce rovná se síla krát dráha: $W = ||F|| \cdot s$. Pokud se velikost síly mění, ale pořád působí ve směru pohybu po přímé dráze, pak si dráhu rozkouskujeme na malé úseky ds, na každém se síla nestačí moc změnit, a proto integrací spojujeme lokální příspěvky k práci ||F(x)||ds. Tento přístup bude fungovat pomocí křivkového integrálu i tehdy, pokud dráha bude nějak zakřivená v rovině či v prostoru, za předpokladu, že síla působí správným směrem.

To ale obecně není zaručeno. Pak je třeba sílu rozložit na dvě složky, jedna jde ve směru pohybu a druhá je kolmá na směr pohybu. Práci vykonává jen ta tečná složka. Jak ji najdeme? Potřebujeme jednotkový tečný vektor \vec{t} , který navíc ukazuje ve směru pohybu (pak je ještě jednotkový tečný vektor ukazující v opačném směru, ten nechceme). Lineární algebra nám říká, že velikost tečné složky vektoru F získáme skalárním součinem $F \bullet \vec{t}$. Odpovídající přírůstek k práci je tedy $F \bullet \vec{t} \, ds$.



Pokud označíme jako $\vec{t}(\vec{x})$ jednotkový tečný vektor ke křivce C v bodě \vec{x} vedoucí ve směru pohybu, pak se celková práce získá křivkovým integrálem

$$\int_C F(\vec{x}) \bullet \vec{t}(\vec{x}) \, ds.$$

Takto se tedy v aplikacích obvykle integrují vektorové funkce po křivkách. Člen $\bar{t}(\vec{x})ds$ v sobě zahrnuje informaci o délce nekonečně malého kousku křivky a také o jeho směru, takže je to možné považovat za "orientovanou infinitesimálu". Někteří autoři ji značí $d\vec{s} = \bar{t}(\vec{x})ds$. Při správném zave-

dení je možné mluvit o "orientovaném diferenciálu". Tento typ křivkového integrálu pro vektorové funkce se pak značí

$$\int_C F(\vec{x}) \bullet d\vec{s}.$$

Je třeba poznamenat, že když máme křivku C, tak to ještě nestačí k tomu, aby měl integrál tohoto typu smysl, protože křivka jako taková nemá přiřazen směr pohybu. Tento integrál tedy zjevně používáme pouze v situacích, kdy máme ještě další informace, zejména směr pohybu po křivce.

Oblíbená intepretace vektorových funkcí je jako vektorové pole. Představme si tedy, že F popisuje lokální rychlost pohybu nějaké kapaliny či plynu. Podíváme-li se do bodu \vec{x} křivky C, pak rozklad F na tečnou a normálovou složku nám říká, kolik média proudí podél křivky v tomto bodě a kolik přechází přes křivku. Když sečteme lokální informace o proudění podél křivky C křivkovým integrálem, dostáváme další užitečnou interpretaci: Integrál

$$\int_C F(\vec{x}) \bullet d\vec{s}$$

udává celkový **tok podél křivky**, zvaný také cirkulace toku (zejména pokud je křivka C uzavřená).

Někdy se tok po křivce pro případ $F = (F_1, F_2): \mathbb{R}^2 \mapsto \mathbb{R}^2$ udává také ve tvaru

$$\int_{C} F_2 dx + F_2 dy.$$

Abychom to rozklíčovali, potřebujeme se bavit o parametrizaci, takže si počkáme do příští sekce.

Tok podél křivky udává tu složku proudění, která neovlivňuje, kolik média je nalevo nebo napravo od křivky. Někdy nás ale zajímá právě toto. Pak potřebujeme pracovat s onou druhou složkou z rozkladu funkce F, která je kolmá na křivku. K výpočtu její velkosti potřebujeme znát jednotkový normálový vektor $\vec{n}(\vec{x})$ ke křivce. Opět existuje více kandidátů a z kontextu je třeba získat informaci, která orientace je správná pro naši situaci. Velikost toku napříč křivkou v bodě \vec{x} se pak získá jako $F(\vec{x}) \bullet \vec{n}(\vec{x})$. Integrací distaneme

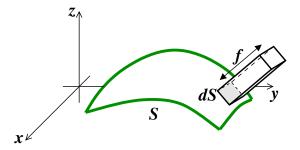
$$\int\limits_C F(\vec{x}) \bullet \vec{n}(\vec{x}) \, ds.$$

Tento **tok přes křivku** nám říká, kolik celkem média se přesunulo z jedné strany křivky na druhou. Pro element $\vec{n}(\vec{x}) ds$ není speciální značení.

Myšlenka křivkového integrálu se snadno zobecní. Pro danou funkci o n proměnných nás zajímá její chování na množině, jejíž skutečná dimenze je nižší. Taková obecná teorie je možná, ale v běžných aplikacích ji nepotřebujeme. Proto se lidé obvykle zaměřují na dva důležité případy, když je množina v podstatě jednorozměrná (to jsme právě probrali) a když je dvourozměrná. Takovým množinám říkáme povrchy.

Představme si místnost se zaznamenanou lokální teplotou f(x, y, z). Mohli bychom tam umístit balónek a zeptat se, jaká je průměrná teplota na jeho povrchu. To je taky rozumná otázka, a evidentně nám s ní křivkový integrál nepomůže. Použijeme obvyklý postup.

Uvažujme povrch S v \mathbb{R}^n pro $n \geq 3$. Pro bod \vec{x} na tomto povrchu vezmeme nekonečně malý čtvereček kolem něj s obsahem rovným dS. Nad touto čtvercovou základnou vztyčíme "sloup" o výšce $f(\vec{x})$, čímž vznikne třírozměrný objekt, jehož objem snadno spočítáme.



Objemy těchto "sloupečků" pak integrací "sečteme". Dostáváme tak **povrchový integrál**, značíme jej

$$\iint\limits_{S} f(\vec{x}) \, dS.$$

Pokud je povrch uzavřený (třeba ten povrch balónku), pak píšeme

$$\iint_{S} f(\vec{x}) \, dS.$$

Jako obvykle nám tento vzorec snadno poskytne **obsahu povrchu**, stačí přes něj integrovat jedničku.

Jaká je typická aplikace povrchových integrálů na vektorové funkce? I zde lze definici interpretovat po souřadnicích, ale v aplikacích se preferuje jiný přístup značený

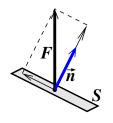
$$\iint\limits_{S} F(\vec{x}) \bullet d\vec{S}.$$

Co se tím myslí? Uvažujme nějaký nekonečně malý kousek povrchu S okolo bodu \vec{x} . Pak dS uchovává jeho obsah a \vec{x} dává lokaci, ale na přesnou informaci o poloze bychom ještě potřebovali znát naklonění, které se u zakřivených povrchů určuje jako sklon tečné roviny v \vec{x} . Problém zachycení sklonu jsme už řešili u délkových infinitesimál, kde jsme zachytili směr pomocí tečného vektoru. To ale funguje jen pro jednodimenzionální objekty. Pokud bychom chtěli zachytit sklon roviny v \mathbb{R}^3 tečnými vektory, tak bychom na to potřebovali dva, což by se nám zde vůbec nehodilo. Pokud chceme sklon zachytit jen jedním vektorem, tak musíme použít normálový vektor, tedy vektor kolmý k povrchu. Ten sloupek v obrázku výše ukazuje správný směr.

Takže pro nekonečně malý kousek povrchu okolo bodu \vec{x} najdeme jednotkový normálový vektor a nazveme jej $\vec{n}(\vec{x})$, pak $d\vec{S} = \vec{n}(\vec{x}) \cdot dS$. Můžeme to považovat za "orientovanou plošnou infinite-simálu". Všimneme si, že pro normálový vektor jsou vždy dva kandidáti, mířící opačnými směry. Předpokládá se, že když používáme $d\vec{S}$, pak jsou k dispozici další informace, které umožní určit správnou orientaci \vec{n} . Aby platila klíčová tvrzení, musí být tato orientace konzistentní, jmenovitě by $\vec{n}(\vec{x})$ mělo být spojité jako funkce \vec{x} .

Je zajímavé, že to není vždy možné. Například není možné najít spojitou orientaci pro známý Mobiův pruh, což je problém, protože spojitost patří mezi základní předpoklady integrovatelnosti. V praktických aplikacích ale takové komplikace nečekáme.

Představme si tedy, že díky informaci máme určeny orientované jednotkové normálové vektory $\vec{n}(\vec{x})$. Pak skalární součin $F \bullet \vec{n}$ poskytuje velikost normálové komponenty F. Pokud by F byla vektorová funkce udávající proudění kapaliny či plynu, pak tečná komponenta F popisuje, kolik částic se pohybuje rovnoběžně s povrchem S, zatímco normálová složka nám říká, jaká část toku opravdu překročí onen povrch.



Typická aplikace povrchového integrálu na vektorovou funkci, které říkáme tok plochou, má

tedy podobu

$$\iint\limits_{S} F(\vec{x}) \bullet d\vec{S} = \iint\limits_{S} F(\vec{x}) \bullet \vec{n}(\vec{x}) \, dS.$$

Zde možná čtenáře napadne, že není úplně příjemné, že délkový diferenciál $d\vec{s}$ pracuje s tečnou, zatímco plošný diferenciál $d\vec{S}$ s normálou, takže skalární součin s vektorovou funkcí dává rozdílné informace (tok podél a tok napříč). Dá se na to zvyknout a má to vnitřní logiku v tom, že v obou případech ukládáme směr ds či dS jedním vektorem.

Oba integrály, křivkový i povrchový, lze zavést jen pro křivky a povrchy rozumné, v praxi to obvykle znamená, že jsme schopni je popsat parametricky pomocí diferencovatelných vektorových funkcí. Blíže se na to podíváme v následující sekci o praktickém výpočtu těchto integrálů, zde se ještě podíváme na dva zajímavé vztahy spojující běžné vícedimenzionální integrály a křivkové/povrchové integrály, které jsou užitečné v aplikacích.

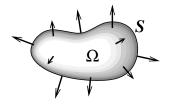
Uvažujme nějakou vektorovu funkci F(x,y,z), kterou si interpretujeme jako vektorové pole, tedy je to například záznam o rychlosti proudění vody v jednotlivých místech koryta řeky, která plyne ustáleně, tedy její proudění se nemění v čase. Uvažujme nějakou oblast, tedy imaginární těleso Ω umístěné v korytě řeky. Protože je imaginární, voda přes něj volně proudí. Zajímá nás, zda tato oblast jako celek přidává nebo ubírá z proudící vody.

Existují dva možné přístupy. Víme, že v každém bodě řeky nám divergence dává lokální bilanci (zdroj, odtok). Pokud tyto lokální bilance posčítáme přes celou oblast, tak najdeme její celkovou bilanci:

$$\iiint_{\Omega} \operatorname{div}(f) \, dV.$$

Stejnou informaci bychom ale měli být schopni získat jinak, jmenovitě tím, že se podíváme, kolik vody proudí dovnitř a ven přes hranici S tohoto tělesa. Hranice tělesa je dvourozměrný objekt neboli povrch. Pokud v každém bodě určíme, kolik vody vtéká či vytéká, tak tato čísla budeme schopni sečíst povrchovým integrálem. To ale bude vyžadovat trochu práce.

Jak jsme před chvílí viděli, tok skrz nějaký kousek plochy okolo bodu \vec{x} lze získat vzorcem $F \bullet \vec{n}(\vec{x})$, ale musíme vyjasnit problém orientace. Pokud je povrch S hranicí třírozměrného tělesa rozumného tvaru, pak existují dva rozeznatelné směry, dovnitř a ven. My zde očekáváme, že kladné znaménko naší celkové bilance znamená produkci média, tedy tok ven.



To dostaneme, pokud budeme brát normálové vektory směřující ven z oblasti. Těmto vektorům se říká **vnější normály**. Uzavřeme tedy úmluvu, že při těchto úvahách vždy bereme tento typ normálových vektorů.

Když jsme tedy vyřešili problém orientace, připomeneme si předchozí rozjímání o "toku plochou" a konstatujeme, že množství média, které oblast Ω vyprodukuje, je dáno také integrálem.

$$\iint\limits_{S} F(\vec{x}) \bullet d\vec{S}.$$

Oba dva přístupy dávají stejnou informaci, takže se musejí rovnat ony dva integrály. Dostáváme tak významné tvrzení zvané **věta o divergenci** či také **Gaussova-Ostrogradského věta**:

$$\iiint\limits_{\Omega} \operatorname{div}(F)(\vec{x}) \, dV = \iint\limits_{S} F(\vec{x}) \bullet d\vec{S}$$

nebo

$$\iiint\limits_{\Omega} \operatorname{div}(F)(\vec{x}) \, dV = \iint\limits_{S} F(\vec{x}) \bullet \vec{n}(\vec{x}) \, dS.$$

Jak značení naznačuje, pro rozumné oblasti Ω bude povrch S uzavřený.

Jak se dá čekat, tato rovnost platí, pouze pokud jsou oblast Ω a její povrch dostatečně rozumné a vektorová funkce F je spojitě diferencovatelná.

Věta o divergenci přirozeně platí i v dalších dimenzích. Můžeme si představit tenkou plotnu, která je odizolovaná od okolí, takže se v ní teplo šíří jen dvourozměrně. Je tam nějaké ohřívátko a jinde zase chladítko a po nějaké době by měl vzniknout ustálený stav, kdy teplo proudí beze změny v čase, což zachytíme vektorovou funkcí F(x,y). Když si na této plotně nakreslíme tužkou bramboroid (uzavřenou křivku C) a zeptáme se, jaká je tepelná bilance útvaru Ω , který jsme tak vymezili, tak se zase můžeme podívat na všechny body uvnitř a sečíst divergence. Nebo se podíváme, co prostupuje hranicí, která je teď jednorozměrná neboli je to (uzavřená) křivka C. I ke křivce umístěné v rovině lze najít vnější normálový vektor, takže předchozí úvahy vedou na rovnost

$$\iint\limits_{\Omega} \operatorname{div}(F)(\vec{x}) \, dS = \oint\limits_{C} F(\vec{x}) \bullet \vec{n}(\vec{x}) \, ds.$$

Co je zajímavé, má smysl udělat takovouto analýzu také pro jednorozměrný případ $\Omega = \langle a, b \rangle$. Hranice $\partial \Omega$ se skládá ze dvou bodů a, b a úvahy o bilanci nás zavedou ke známému vzorci

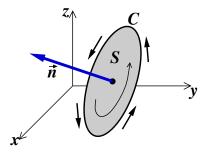
$$\int_{[a,b]} f'(x) \, dx = \left[f(x) \right]_a^b = f(b) - f(a).$$

Druhá důležitá rovnost v tomto kontextu je **Stokesova věta**, obvykle formulovaná pro tři dimenze:

$$\iint\limits_{S} \operatorname{rot}(F) \bullet d\vec{S} = \oint\limits_{C} F \bullet d\vec{s}.$$

Zde S je nějaká plocha v \mathbb{R}^3 a C je její hranice, tedy uzavřená křivka. Dále se předpokládá, že orientace diferenciálů $d\vec{S}$ a $d\vec{s}$ jsou sladěny tak, aby tvořily "pravoruký pár". Můžeme si to představit tak, že nejprve určíme směr oběhu okolo křivky C, čímž jsou dány orientace $d\vec{s}$, a pak se mezi dvěma možnostmi volby normál $\vec{n}(\vec{x})$ k S zvolí ta orientace, aby při pohledu dolů na plochu S ze špičky vektoru \vec{n} bylo obíhání hranice ve směru d \vec{s} v kladném směru, tedy proti směru hodinových ručiček.

Co nám Stokesova věta říká? Představme si plochý disk S v třírozměrném vektorovém poli F neboli proudícím médiu, který je držen ve stejné pozici, ale může se otáčet kolem své osy. Ta je pak zároveň směrem pro normály \vec{n} . Tím se následně určí směr pohybu okolo obvodu.



Jak už jsme viděli výše, integrál

$$\oint\limits_C F \bullet d\vec{s} = \oint\limits_C F \bullet \vec{t} \, ds$$

nám říká, jak velká část média proudí podél hranice C. Díky tření se snaží diskem otočit, integrál tyto účinky sčítá po celém obvodu a tudíž nám dává celkový rotační moment okolo osy \vec{n} , kterým médium působí na obvod našeho disku. Je samozřejmě možné, že médium působí na hranici disku i jinak, například někde dokonce může plynout kolmo na směr hranice, takže zde může být i snaha disk naklonit, ale ten skalární součin v integrálu tuto část působení vyřadí z úvah.

Všimněte si, že integrál vyjde kladný, pokud médium proudí ve zvoleném směru oběhu, a záporný, pokud proudí v opačném směru.

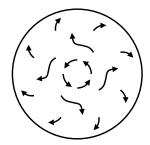
Integrál

$$\iint\limits_{S} \operatorname{rot}(F) \bullet d\vec{S} = \iint\limits_{S} \operatorname{rot}(F) \bullet \vec{n}(\vec{x}) dS$$

sčítá lokální tendence otáčet diskem (rozpoznatelné operátorem rotace) přes celý disk. Rotační působení poskytnuté $\operatorname{rot}(F)$ má formu vektoru a skalární součin s \vec{n} z něj extrahuje složku, která se snaží o rotaci okolo osy \vec{n} . Tato informace je tedy konzistentní s tím, co jsme se snažili vysledovat na obvodu pomocí křivkového integrálu. Všimneme si, že znaménka souhlasí s křivkovým integrálem, tato lokální působení jsou kladná, pokud se točí ve směru oběhu kolem obvodu.

Stokesova věta říká, že vyjde nastejno, jestli určíme rotační působení kolem osy \vec{n} sečtením lokálních účinků na desku nebo sečtením účinků na obvodu. Na rozdíl od věty o divergenci, kde šlo čistě o přidávání či ubírání média, zde již není tak jasné, proč by ony dva součty měly vyjít stejně. Umím si třeba představit médium, které by bylo na obvodu disku v klidu, ale uprostřed se točilo. Pak by věta logicky nemohla platit.

Jenže takové médium by nebylo reálné, protože aby se kapalina uprostřed točila a na obvodu zůstávala v klidu, tak by někde mezi muselo být mezikruži, kde zase proudí zpět. Jinak řečeno, pokud je uprostřed rotace kladná, tak blízko obvodu bude zase záporná. Stokesova věta nám říká, že tato představa je správná, kladné a záporné rotace v oblasti by se měly navzájem zkrátit, pokud chceme nehybné médium na obvodu.

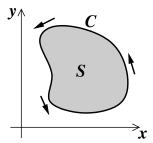


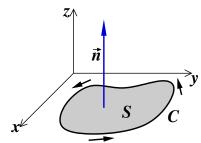
Pracovali jsme zde s plochým diskem, ale Stokesova věta je mnohem obecnější. Rovnost rotačního působení na obvodu objektu a na jeho ploše platí i pro jiné tvary, tedy mohou být různě vypouklé a nepravidelné a podobně. Je to opravdu silná věta.

Jak se Stokesova věta převede do dvou dimenzí? Tok okolo hranice je dán stejným křivkovým integrálem, ale musíme změnit ten povrchový. U dvourozměrné funkce F dává rotace jen číslo, protože už se nemusíme zabývat směrem rotace, jen velikostí rotačního momentu. Celkový vliv tak získáme přímým integrováním těchto lokálních příspěvků, bez nutnosti skalárního součinu.

$$\iint\limits_{S} \operatorname{rot}(F) \, dS = \oint\limits_{C} F \bullet d\vec{s}.$$

Jen si musíme pohlídat, že orientace cestování okolo C je správně zvolena. Pravidlo zní, že když cestujeme po C, tak bychom měli vidět vnitřek S po naší levici. Tomuto vzorci se také říká **Greenova věta** a je zajímavé, že je ekvivalentní dvoudimenzionální verzi věty o divergenci. Jinak řečeno, Stokesova věta a věta o divergenci dávají pro funkce dvou proměnných stejnou informaci, jen jinak formulovanou.





Dvoudimenzionální verzi lze také odvodit z původního tvrzení tak, že \mathbb{R}^2 vnoříme do \mathbb{R}^3 . Přesně řečeno, máme-li vektorovou funkci $F = (F_1(x, y), F_2(x, y))$ a oblast S v \mathbb{R}^2 , pak definujeme pomocnou funkci

$$G(x, y, z) = (F_1(x, y), F_2(x, y), 0)$$

a povrch $S_3 = \{(x, y, 0); (x, y) \in S\}$. Pak můžeme aplikovat třídimenzionální Stokesovu větu na G a S_3 . Snadno zjistíme, že

$$rot(G)(x,y,z) = \left(0,0,\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right) = \left(0,0,rot(F)(x,y)\right).$$

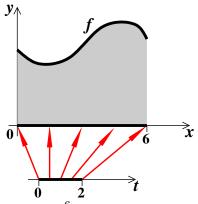
Protože normálový vektor k plochému objektu v rovině xy je $\vec{n} = (0,0,1)$, dostáváme

$$rot(G)(x, y, z) \bullet \vec{n} = rot(F)(x, y).$$

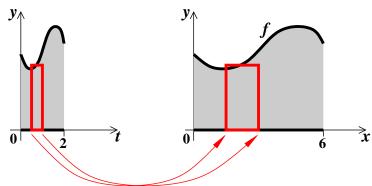
6d. Parametrizace, substituce

Začneme návratem k funkci jedné proměnné. Chceme počítat integrál z funkce f(x) na intervalu (0,6), jinými slovy zajímá nás obsah jistého obrazce.

My ovšem z nějakého důvodu nechceme přistupovat k ose x přímo, ale prostřednictvím nějakého parametru t, například tak, že pro $t \in \langle 0, 2 \rangle$ budeme pracovat s x = 3t.



Co by se stalo, kdybychom místo integrálu $\int\limits_0^6 f(x)\,dx$ počítali integrál $\int\limits_0^2 f(3t)\,dt$? Procházíme body intervalu $\langle 0,2\rangle$ a "sčítáme" hodnoty f(3t), což vlastně znamená, že zahrnujeme všechny hodnoty f(x) pro $x\in\langle 0,6\rangle$. Jenže my je projdeme, zatímco se pohybujeme na menším intervalu $\langle 0,2\rangle$. To znamená, že pracujeme s původním tvarem oblasti pod grafem funkce f, ale zmáčknutým třikrát ze strany.



Dá se čekat, že výsledný obsah a tedy i integrál budou třikrát menší, než by měly.

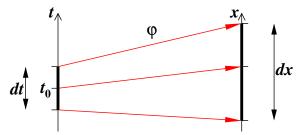
Obrázek ukazuje ještě jiný argument, kterým se to dá odůvodnit. Riemannovská integrace nám říká, že bychom měli $\int\limits_0^6 f(x)\,dx$ aproximovat rozdělením intervalu $\langle 0,6\rangle$ na části a pracovat s obsahy $f(x)\cdot dx$ obdélníků nad nimi. Zobrazení x=3t nám umožní přesunout tuto situaci do světa t, ale přitom zkrátí velikosti základen třikrát, takže obsahy $f(3t)\cdot dt$ jsou třikrát menší, než by měly.

Pokud chceme obsahy zachovat, musíme infinitesimály dx nahradit nikoliv kratšími dt, ale jejich trojnásobkem, tedy dx = 3dt. Jinak řečeno, máme počítat integrál $\int\limits_0^2 f(3t)\,3dt$. Shodou okolností nám přesně totéž radí věta o substituci, pro vzorec x=3t nám poradí nahrazovat diferenciály pomocí $dx=[3t]'dt=3\,dt$.

Náš příklad byl samozřejmě velmi snadný, protože měnil osu t na osu x rovnoměrně, ale je možné to dělat i jinak. Třeba substituce $x=t^2$ infinitesimály dt někdy roztahuje a někdy zkracuje. Substituční pravidlo říká, že nám pak pomůže derivace, a bude pro nás užitečné pochopit, proč by to tak mělo být.

Uvažujme tedy převod souřadnic $x = \varphi(t)$. Podíváme se, jak se deformuje infinitesimála dt.

Uvažujme na ose t nějaký bod t_0 a kolem něj malý úsek šířky dt, takže vlastně jde o úsečku mezi body $t_0 - \frac{1}{2}dt$ a $t_0 + \frac{1}{2}dt$. Uvažujme pro jednoduchost rostoucí φ , pak se tato úsečka zobrazí na úsečku mezi body $\varphi(t_0 - \frac{1}{2}dt)$ a $\varphi(t_0 + \frac{1}{2}dt)$.



To znamená, že výsledná délka dx je dána jako

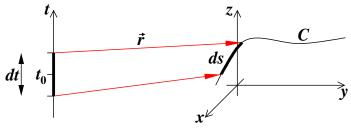
$$dx = \varphi(t_0 + \frac{1}{2}dt) - \varphi(t_0 + \frac{1}{2}dt).$$

Pokud je dt opravdu malé, lze hodnoty funkce aproximovat lineárními Taylorovými polynomy a dostáváme

$$dx = \left[\varphi(t_0) + \varphi'(t_0)\frac{1}{2}dt\right] - \left[\varphi(t_0) - \varphi'(t_0)\frac{1}{2}dt\right] = \varphi'(t_0)dt.$$

Nyní povedeme stejnou úvahu pro **křivkový integrál**. Uvažujme nějakou křivku C v \mathbb{R}^n . Musí být nějak zadána, a pro křivky je přirozené mít ji zachycenou jako parametrickou křivku, tedy křivka C je obraz nějaké vektorové funkce tradičně značené $\vec{r}(t)$ pro $t \in \langle a, b \rangle$.

Naším cílem je projít postupně body křivky C a "sčítat" přírůstky $f(\vec{x}) \cdot ds$. Parametrizace nám umožňuje dostávat se k bodům křivky pomocí \vec{r} , tedy máme $\vec{x} = \vec{r}(t)$, a potřebujeme ještě zjistit, jak se při této transformaci deformuje infinitesimála dt. Budeme postupovat podobně, tedy zvolíme si t_0 a podíváme se na úsečku o celkové délce dt kolem něj.



Výpočet bude obdobný, protože jsme v sekci 5b ukázali, že lineární aproximace funguje také pro vektorové funkce jedné proměnné. Jedna podstatná změna ale ve výpočtu bude. Protože se naše elementární úsečka z osy t přenáší do vícerozměrného světa \mathbb{R}^n , vznikne vektor a my ke zjištění jeho délky budeme potřebovat normu.

$$ds = \|\vec{r}(t_0 + \frac{1}{2}dt) - \vec{r}(t_0 - \frac{1}{2}dt)\|$$

= $\|[\vec{r}(t_0) + \vec{r}'(t_0)\frac{1}{2}dt] - [\vec{r}(t_0) - \vec{r}'(t_0)\frac{1}{2}dt]\| = \|\vec{r}'(t_0)dt\|$
= $\|\vec{r}'(t_0)\|dt$.

Takže pokud chceme přičíst obsah jistého elementárního obdélníka $f(\vec{x}) \cdot ds$ pro bod $\vec{x} = \vec{r}(t)$ na křivce, tak to lze nahradit přičtením obsahu $f(\vec{r}(t)) \cdot ||\vec{r}'(t)|| dt$. Sečtením takovýchto obdélníků pak

získáme hodnotu integrálu po křivce C. Protože je to důležitý výsledek, dáme tomu nějakou formu a přidáme nutné předpoklady.

Fakt.

Nechť f je spojitá funkce na nějaké množině $\Omega \subset \mathbb{R}^n$.

Nechť C je křivka v Ω daná parametrizací $\vec{r}(t)$ pro $t \in \langle a, b \rangle$, kde \vec{r} je vektorová funkce $\langle a, b \rangle \mapsto \Omega$, taková, že $\vec{r}[\langle a, b \rangle] = \Omega$, \vec{r} je na $\langle a, b \rangle$ prostá a má spojitou nenulovou derivaci na (a, b). Pak

$$\int_{C} f(\vec{x}) ds = \int_{a}^{b} f(\vec{r}(t)) ||\vec{r}'(t)|| dt.$$

Snadno se to pamatuje, protože to je velmi podobné klasické substituci

$$\begin{vmatrix} \vec{x} = \vec{r}(t) \\ ds = ||\vec{r}'(t)|| dt \end{vmatrix}.$$

Existuje speciální a vcelku častá situace, pro kterou se tento vzorec zjednoduší. Uvažujme křivku C v \mathbb{R}^2 , která je grafem nějaké diferencovatelné funkce g(x) na intervalu $\langle a,b\rangle$. Pak existuje přirozený parametrický popis $\vec{r}(x)=(x,g(x))$ a vzorec pro křivkový integrál funkce f(x,y) na C se dá upravit na

$$\oint_C f(x,y) ds = \int_a^b f(x,g(x)) \sqrt{1 + [g'(x)]^2} dx.$$

Pokud by nás zajímala délka grafu funkce g, tak bychom použili f=1 a vyjde vzorec, který známe z úvodního kurzu analýzy.

Existuje alternativní pohled na tuto substituci v křivkovém integrálu. Druhé rozšířené značení pro parametrický popis je (ve dvou dimenzích)

$$x = x(t),$$

$$y = y(t).$$

Zde x(t) a y(t) odpovídají komponentám $r_1(t)$ a $r_2(t)$ parametrického popisu $\vec{r}(t)$. Jednorozměrná substituce nám říká, že

$$dx = x'(t)dt,$$

 $dy = y'(t)dt.$

Toto nám říká, jak se změní jednotlivé souřadnice bodu $(x_0, y_0) = (x(t_0), y(t_0))$, pokud parametr změníme o dt. Z bodu (x_0, y_0) se tak posuneme do bodu $(x_0 + dx, y_0 + dy)$ a vznikne úsečka, která je obrazem infinitesimály dt. Jak je dlouhá? Zajímá nás délka vektoru

$$(x_0 + dx, y_0 + dy) - (x_0, y_0) = (dx, dy),$$

takže

$$ds = \sqrt{dx^2 + dy^2}.$$

Tento vzorec nabízí alternativní přepis křivkového integrálu, který ale odpovídá tomu původnímu, jak se snadno přesvědčíme:

$$ds = \sqrt{dx^2 + dy^2} = \sqrt{[x'(t)dt]^2 + [y'(t)dt]^2}$$

$$= \sqrt{([x'(t)]^2 + [y'(t)]^2)dt^2} = \sqrt{[x'(t)]^2 + [y'(t)]^2} dt$$

$$= \sqrt{[r'_1(t)]^2 + [r'_2(t)]^2} dt = ||(r'_1(t), r'_2(t))|| dt = ||\vec{r}'(t)|| dt.$$

Obdobně se lze ve třech rozměrech setkat se zápisem $ds = \sqrt{dx^2 + dy^2 + dz^2}$. Pro čtenáře, který

ještě nepracoval s transformacemi proměnných ve více dimenzích, může být užitečné podívat se na část Transformace souřadnic v sekci 7b.

Obecný substituční vzorec nabývá specifické podoby, když integrujeme vektorovou funkci. Nejprve se podíváme na tok ve směru křivky, tedy $\int_C F(\vec{x}) \cdot d\vec{s}$. Potřebujeme jednotkový tečný vektor \vec{t} v bodě $\vec{x} = \vec{r}(t)$, který snadno najdeme jako $\frac{\vec{r}'(t)}{\|\vec{r}'(t)\|}$, protože ta derivace podle našich předpokladů není nulová. Můžeme tedy psát

$$\int_{C} F(\vec{x}) \bullet d\vec{s} = \int_{C} F(\vec{x}) \bullet \vec{t}(\vec{x}) ds = \int_{a}^{b} F(\vec{r}(t)) \bullet \frac{\vec{r}'(t)}{\|\vec{r}'(t)\|} \|\vec{r}'(t)\| dt = \int_{a}^{b} F(\vec{r}(t)) \bullet \vec{r}'(t) dt.$$

Tento vzorec se hodí, když počítáme integrály ze Stokesova vzorce. Podívejme se na něj ve dvou dimenzích.

$$\int_{C} F(\vec{x}) \bullet d\vec{s} = \int_{a}^{b} F(\vec{r}(t)) \bullet \vec{r}'(t) dt = \int_{a}^{b} (F_{1}, F_{2}) \bullet (r'_{1}, r'_{2}) dt$$

$$= \int_{a}^{b} F_{1}r'_{1} + F_{2}r'_{2} dt = \int_{a}^{b} F_{1}(\vec{r}(t))r'_{1}(t) dt + \int_{a}^{b} F_{2}(\vec{r}(t))r'_{2}(t) dt.$$

Ovšem vzpomeneme-li si na alternativní značení substituce,

$$\begin{array}{l} x = x(t) = r_1(t) \\ y = y(t) = r_2(t) \end{array} \implies \begin{array}{l} r'_1(t)dt = x'(t)dt = dx \\ r'_2(t)dt = y'(t)dt = dy \end{array}$$

vidíme, že můžeme provést zpětnou substituci

$$\int_{C} F(\vec{x}) \bullet d\vec{s} = \int_{C} F_1 dx + \int_{C} F_2 dy = \int_{C} F_1 dx + F_2 dy.$$

Někteří autoři preferují tento tvar ve vzorci Greenovy věty.

Jak by vypadal tok přes hranici, tedy v situaci věty o divergenci? Tam je potřeba najít vnější jednotkový normálový vektor ke křivce. Máme tečný vektor $\vec{r}' = (r'_1, r'_2)$ a z něj máme dva přirozené kandidáty na kolmý vektor, $(r'_2, -r'_1)$ a $(-r'_2, r'_1)$. Není těžké si rozmyslet, že pokud se pohybujeme ve směru \vec{r}' v rovině, tak ten první kolmý ukazuje doprava a ten druhý doleva. My pro správnou orientaci potřebujeme vidět vnitřek množiny po své levé ruce, takže ten první vektor ukazuje ven a je to ten pravý. Pak jej ještě musíme vydělit normou, aby byl jednotkový.

$$\int_{C} F(\vec{x}) \bullet \vec{n}(\vec{x}) ds = \int_{a}^{b} F \bullet \frac{(r'_{2}, -r'_{1})}{\|(r'_{2}, -r'_{1})\|} \|\vec{r}'(t)\| dt = \int_{a}^{b} (F_{1}, F_{2}) \bullet (r'_{2}, -r'_{1}) \frac{\|(r'_{1}, r'_{2})\|}{\|(r'_{2}, -r'_{1})\|} dt$$

$$= \int_{a}^{b} F_{1}r'_{2} - F_{2}r'_{1} \frac{\sqrt{[r'_{1}]^{2} + [r'_{2}]^{2}}}{\sqrt{[r'_{2}]^{2} + [-r'_{1}]^{2}}} dt = \int_{a}^{b} F_{1}r'_{2} - F_{2}r'_{1} dt$$

$$= \int_{a}^{b} F_{1}r'_{2} dt - \int_{a}^{b} F_{2}r'_{1} dt = \int_{C} F_{1} dy - \int_{C} F_{2} dx = \int_{C} F_{1} dy - F_{2} dx.$$

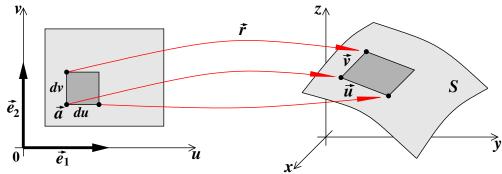
Tento výraz se někdy objevuje v dvoudimenzionálních formulacích věty o divergenci.

Nyní se podíváme na **povrchový integrál**. Budeme postupovat obdobně. V povrchovém integrálu "sčítáme" objemy $f(\vec{x}) \cdot dS$. Rozhodneme se, že se k povrchu S budeme dostávat pomocí parametrizace, tedy vektorové funkce $\vec{r}(u,v)$. Jak se přenese plošná infinitesimála ze světa u,v,

tedy z \mathbb{R}^2 , do \mathbb{R}^n ? Přesněji řečeno nás zajímá, jak se změní její obsah, protože nás zajímá velikost podstavy onoho objemu $f(\vec{x}) \cdot dS$.

Ve výchozím prostoru $\mathbb{R}^2[u,v]$ si teď budeme představovat infinitesimály jako nekonečně malé čtverečky. Jejich strany jsou sice stejné, ale odpovídají různým souřadnicovým osám, takže je budeme značit du a dv. Obsah je tedy $du \cdot dv$. Tento čtverec přeneseme pomocí \vec{r} do cílového prostoru \mathbb{R}^m a podíváme se, jak se změnil jeho obsah.

Zavedeme si přesnější značení. Základem je bod \vec{a} , který je zároveň vrcholem sledovaného čtverce. Jeho vodorovná strana o délce du se dá vnímat jako úsečka s krajními body \vec{a} a $\vec{a} + du \cdot \vec{e_1}$. Obdobně vidíme svislou stranu jako spojnici bodů \vec{a} a $\vec{a} + dv \cdot \vec{e_2}$. Transformací těchto bodů zjistíme, co se děje se čtvercem.



Dá se čekat, že se kolmé vektory nemusí nutně převést zase na kolmé vektory, ale tím to nemusí skončit. Jak naznačuje obrázek, obrazy jednotlivých stran se také mohou zakřivit. To by bylo nepříjemné. Když je ale výchozí čtvereček extrémně malý, pak se strany nemohly příliš zdeformovat a obraz elementárního čtverce lze považovat za rovnoběžník. Jaké jsou jeho strany?

Z vodorovné strany infinitesimálního čtverce vznikl vektor, který vede z $\vec{r}(\vec{a})$ do $\vec{r}(\vec{a} + du \cdot \vec{e}_1)$. Máme tedy

$$\vec{u} = \vec{r}(\vec{a} + du \cdot \vec{e}_1) - \vec{r}(\vec{a}).$$

Protože je du velmi malé, použijeme pro první člen lineární aproximaci využívající směrovou derivaci. Zde jde ovšem o derivaci ve směru $\vec{e_1}$, tedy parciální derivaci vůči u.

$$\vec{r}(\vec{a} + du \cdot \vec{e}_1) \approx \vec{r}(\vec{a}) + D_{\vec{e}_1}\vec{r}(\vec{a}) \cdot du = \vec{r}(\vec{a}) + \frac{\partial \vec{r}}{\partial u}(\vec{a}) \cdot du.$$

Pokud je to pro čtenáře moc rychlé, můžeme se na to podívat po složkách funkce $\vec{r} = (r_1, \dots, r_n)$.

$$r_j(\vec{a} + du \cdot \vec{e}_1) = r_j(a_1 + du, a_2) \approx r_j(a_1, a_2) + \frac{\partial r_j}{\partial u}(a_1, a_2) du = r_j(\vec{a}) + \frac{\partial r_j}{\partial u}(\vec{a}) du.$$

Teď zpětně poskládáme vektor.

$$\vec{r}(\vec{a} + du \cdot \vec{e}_1) \approx \left(r_1(\vec{a}) + \frac{\partial r_1}{\partial u}(\vec{a})du, r_2(\vec{a}) + \frac{\partial r_2}{\partial u}(\vec{a})du, \dots\right)$$

$$= \left(r_1(\vec{a}), r_2(\vec{a}), \dots\right) + \left(\frac{\partial r_1}{\partial u}(\vec{a}), \frac{\partial r_2}{\partial u}(\vec{a}), \dots\right)du = \vec{r}(\vec{a}) + \frac{\partial \vec{r}}{\partial u}(\vec{a}) \cdot du.$$

Bylo to užitečné už proto, že jsme si ujasnili, že $\frac{\partial \vec{r}}{\partial u}(\vec{a})$ je vlastně vektor. To podstatné je, že jedna strana našeho rovnoběžníku je

$$\vec{u} = \vec{r}(\vec{a} + du \cdot \vec{e}_1) - \vec{r}(\vec{a}) \approx \frac{\partial \vec{r}}{\partial u}(\vec{a}) \cdot du.$$

Obdobně zjistíme, že druhá strana je

$$\vec{v} = \vec{r}(\vec{a} + dv \cdot \vec{e}_2) - \vec{r}(\vec{a}) \approx \frac{\partial \vec{r}}{\partial v}(\vec{a}) \cdot dv.$$

Vektory \vec{u} a \vec{v} tvoří rovnoběžník v \mathbb{R}^n . Jak najdeme jeho obsah? To není obecně snadné. Proto se omezíme na případ povrchů v \mathbb{R}^3 . Tam existuje vzorec, který nachází obsah dotyčného rovnoběžníka

jako velikost kartézského součinu sousedících stran.

$$dS = \|\vec{u} \times \vec{v}\| = \left\| \left(\frac{\partial \vec{r}}{\partial u} (\vec{a}) du \right) \times \left(\frac{\partial \vec{r}}{\partial v} (\vec{a}) dv \right) \right\| = \left\| \frac{\partial \vec{r}}{\partial u} (\vec{a}) \times \frac{\partial \vec{r}}{\partial v} (\vec{a}) \right\| du dv.$$

Poznamenejme, že to $du\,dv$ vyšlo z intuitivních výpočtů, ale neměli bychom se vázat na přesnou podobu, například že bychom pak museli integrovat nejprve vzhledem ku a pak vzhledem kv. Spíše nám to říká, že tam máme očekávat d[u,v] neboli infinitesimálu z $\mathbb{R}^2[u,v]$, která může mít různé podoby, například tu nahoře nebo také $dv\,du$.

Teď tedy umíme nahradit v objemu $f(\vec{x})dS$ diferenciál. Hodnota funkce se přirozeně nahradí jako $f(\vec{x}) = f(\vec{r}(u,v))$ a jsme připraveni to dát dohromady a dojít ke kýženému vzorci. Bohužel platí jen pro n=3, ale zrovna tato dimenze je potřeba v aplikacích, takže toto omezení nevadí.

Fakt.

Nechť f je spojitá funkce na nějaké množině $\Omega \subset \mathbb{R}^3$.

Nechť S je povrch v Ω daný parametrizací $\vec{r}(u,v)$ pro $(u,v) \in D$, kde D je nějaká množina v \mathbb{R}^2 a \vec{r} je vektorová funkce $D \mapsto \Omega$ taková, že $\vec{r}[D] = S$, \vec{r} je na D prostá a má nenulové spojité parciální derivace na vnitřku D. Pak

$$\iint_{S} f(\vec{x}) dS = \iint_{D} f(\vec{r}(u, v)) \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| d[u, v].$$

Opět to lze vnímat jako jistou formu substituce

$$\begin{vmatrix} \vec{x} = \vec{r}(u, v) \\ dS = \left\| \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right\| du \, dv \end{vmatrix}.$$

I zde máme zajímavý speciální případ. Uvažujme povrch S v \mathbb{R}^3 , který je grafem nějaké diferencovatelné funkce g(x,y) na D. Přirozená parametrizace $\vec{r}(x,y) = (x,y,g(x,y))$ vede na vzorec

$$\iint\limits_{S} f(x,y,z) dS = \iint\limits_{D} f(x,y,g(x,y)) \sqrt{\left(\frac{\partial g}{\partial x}\right)^{2} + \left(\frac{\partial g}{\partial y}\right)^{2} + 1} d[x,y].$$

Jak parametrizace funguje pro vektorové funkce? Zejména nás zajímá integrál $\iint_S F(\vec{x}) \bullet d\vec{S}$.

Uvažujme nějakou parametrizaci \vec{r} . Pak vektorový součin $\vec{N} = \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v}$ dává vektor kolmý na povrch v daném bodě. Podle našich předpokladů je nenulový, tudíž pomocí něj lze získat jednotkový normálový vektor $\vec{n} = \frac{\vec{N}}{\|\vec{N}\|}$. Pak dostáváme

$$\begin{split} \iint\limits_{S} F(\vec{x}) \bullet d\vec{S} &= \iint\limits_{S} F(\vec{x}) \bullet \vec{n} \, dS = \iint\limits_{D} F(\vec{r}(u,v)) \bullet \frac{\vec{N}}{\|\vec{N}\|} \|\vec{N}\| \, d[u,v] \\ &= \iint\limits_{D} F(\vec{r}(u,v)) \bullet \vec{N} \, d[u,v] = \iint\limits_{D} F(\vec{r}(u,v)) \bullet \left(\frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} \right) d[u,v]. \end{split}$$

To se hodí, když počítáme integrály ve Stokesově vzorci. Podívejme se na ten vektorový součin blíže.

$$\begin{split} \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} &= \left(\frac{\partial r_1}{\partial u}, \frac{\partial r_2}{\partial u}, \frac{\partial r_3}{\partial u}\right) \times \left(\frac{\partial r_1}{\partial v}, \frac{\partial r_2}{\partial v}, \frac{\partial r_3}{\partial v}\right) \\ &= \left(\left|\frac{\partial r_2}{\partial u}, \frac{\partial r_3}{\partial v}\right|, \left|\frac{\partial r_3}{\partial u}, \frac{\partial r_1}{\partial v}\right|, \left|\frac{\partial r_1}{\partial v}, \frac{\partial r_2}{\partial v}\right|, \left|\frac{\partial r_1}{\partial v}, \frac{\partial r_2}{\partial v}\right|\right). \end{split}$$

Proto

$$\iint_{S} F(\vec{x}) \bullet d\vec{S} = \iint_{D} F_{1}(\vec{r}(u,v)) \begin{vmatrix} \frac{\partial r_{2}}{\partial u} & \frac{\partial r_{3}}{\partial u} \\ \frac{\partial r_{2}}{\partial v} & \frac{\partial r_{3}}{\partial v} \end{vmatrix} d[u,v] + \iint_{D} F_{2}(\vec{r}(u,v)) \begin{vmatrix} \frac{\partial r_{3}}{\partial u} & \frac{\partial r_{1}}{\partial u} \\ \frac{\partial r_{3}}{\partial v} & \frac{\partial r_{1}}{\partial v} \end{vmatrix} d[u,v] + \iint_{D} F_{3}(\vec{r}(u,v)) \begin{vmatrix} \frac{\partial r_{1}}{\partial u} & \frac{\partial r_{2}}{\partial u} \\ \frac{\partial r_{1}}{\partial v} & \frac{\partial r_{2}}{\partial v} \end{vmatrix} d[u,v].$$

Teď to přijde. Parametrizace $(u, v) \mapsto \vec{r}(u, v) = (r_1(u, v), r_2(u, v), r_3(u, v))$ také vytváří tři přidružené transformace $\mathbb{R}^2 \mapsto \mathbb{R}^2$. Jedna z nich funguje takto:

$$x = r_1(u, v),$$

$$y = r_2(u, v).$$

Jak uvidíme dále, při této substituci se diferenciál transformuje podle vzorce

$$dx \, dy = \begin{vmatrix} \frac{\partial r_1}{\partial u} & \frac{\partial r_2}{\partial u} \\ \frac{\partial r_1}{\partial v} & \frac{\partial r_2}{\partial v} \end{vmatrix} du \, dv,$$

což souhlasí s výrazem ve třetím integrálu. V prvním integrálu se podobně použije substituce $y = r_2(u, v)$, $z = r_3(u, v)$ a v tom prostředním $z = r_3(u, v)$, $x = r_1(u, v)$ (pozor na pořadí!). Dostáváme tak pomocí zpětné substituce přepis

$$\iint\limits_{S} F(\vec{x}) \bullet d\vec{S} = \iint\limits_{S} F_1(\vec{x}) dy dz + \iint\limits_{S} F_2(\vec{x}) dz dx + \iint\limits_{S} F_3(\vec{x}) dx dy.$$

Když toto použijeme s rot(F) namísto F a spojíme s předchozím přepisem křivkového integrálu, dostaneme alternativní zápis Stokesovy věty:

$$\iint\limits_{S} \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z} \right) dy \, dz + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x} \right) dz \, dx + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx \, dy = \oint\limits_{C} F_1 dx + F_2 dy + F_3 dz.$$

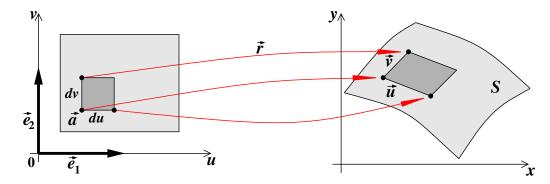
Není to můj oblíbený tvar, ale stálo za to vysvětlit zde jeho smysl, kdybyste na to někdy narazili.

Je načase podívat se na **substituci** pro běžný vícerozměrný integrál, kterým jsme tuto kapitolu začali. Máme tedy funkci f definovanou na množině Ω v \mathbb{R}^n , o které předpokládáme, že má opravdu dimenzi n, což si můžeme představit tak, že má neprázný vnitřek v \mathbb{R}^n , popřípadě že má nenulový n-rozměrný objem, takže n-rozměrný integrál f na Ω nevyjde automaticky nula.

Substituce předpokládá, že jsme si tuto množinu popsali parametricky, tedy existuje nějaká vektorová funkce \vec{r} na jisté množině D taková, že $\Omega = \vec{r}[D]$. Protože je množina Ω n-rozměrná a my chceme, aby byla funkce \vec{r} hladká, musí být i množina D n-rozměrná. Odpovídá to naší intuici, že na popis množiny, která má opravdu n rozměrů, potřebujeme n parametrů.

Máme tedy vektorovou funkci \vec{r} : $\mathbb{R}^n \to \mathbb{R}^n$. Naším cílem je nahradit výrazy typu $f(\vec{x})d[\vec{x}]$, kde $d[\vec{x}]$ je n-rozměrná infinitesimála prostoru \mathbb{R}^n . Potřebujeme proto zjistit, jak se n-dimenzionální infinitesimála z výchozí množiny D pro funkci \vec{r} přenese do cílového prostoru; přesněji řečeno nás zajímá změna n-dimenzionálního objemu při této transformaci.

Začněme případem n=2, tedy máme vektorovou funkci $\vec{r}(u,v)$, jejíž obrazy jsou také dvourozměrné. Chceme přenést čtvereček o stranách ds, dt, a to je něco, co jsme již dělali před chvílí. Teď je cílový prostor dvoudimenzionální.



To znamená, že v závěru postupu musíme zvolit jiný přístup pro výpočet obsahu rovnoběžníka. Ve dvou dimenzích máme z vektorů stran vytvořit matici a najít absolutní hodnotu jejího determinantu. Vektory $\vec{u} = \frac{\partial \vec{r}}{\partial u} du$ a $\vec{v} = \frac{\partial \vec{r}}{\partial v} dv$ lze dávat jako řádky i jako sloupce, my zvolíme druhou variantu. Dostáváme tedy

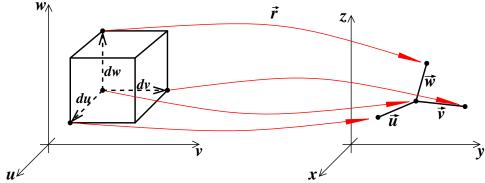
$$dS = \left| \det \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \end{pmatrix} \right| = \left| \det \begin{pmatrix} \frac{\partial r_1}{\partial u} du & \frac{\partial r_1}{\partial v} dv \\ \frac{\partial r_2}{\partial u} du & \frac{\partial r_2}{\partial v} dv \end{pmatrix} \right| = \left| \det \begin{pmatrix} \frac{\partial r_1}{\partial u} & \frac{\partial r_1}{\partial v} \\ \frac{\partial r_2}{\partial u} & \frac{\partial r_2}{\partial v} \end{pmatrix} \right| \cdot du \cdot dv$$

Zajímavou shodou okolností tam vidíme Jacobiho matici funkce \vec{r} . Pro její determinant jsme si zavedli speciální značení a jméno jakobián, takže můžeme psát

$$dS = |\Delta_{\vec{r}}(\vec{a})| \, du \, dv.$$

Podívejme se na případ n=3. Takže nás zajímá trojný integrál funkce f na množině $\Omega \subseteq \mathbb{R}^3$ rozumného tvaru, který "sčítá" čtyřrozměrné objemy $f(\vec{x}) \cdot dV$, kde dV je objem elementární třírozměrné krychličky.

Uvažujeme parametrizaci $\vec{r}(u,v,w)$: $D\mapsto\mathbb{R}^3$, kde $D\subset\mathbb{R}^3$ a $\Omega=\vec{r}[D]$. Jak se změní objem elementární krychličky se stranami $du,\,dv,\,dw$, když ji pomocí \vec{r} přeneseme do Ω ?



Opět si jednotlivé strany napíšeme pomocí standardních vektorů $\vec{e}_1,\,\vec{e}_2,\,\vec{e}_3$ a stejné výpočty jako dříve ukážou, že se klíčové strany krychle přenesou na vektory

$$\vec{u} \approx \frac{\partial \vec{r}}{\partial u}(\vec{a}) \cdot du, \quad \vec{v} \approx \frac{\partial \vec{r}}{\partial v}(\vec{a}) \cdot dv, \quad \vec{w} \approx \frac{\partial \vec{r}}{\partial w}(\vec{a}) \cdot dw.$$

Tyto vektory vytvářejí rovnoběžnostěn v \mathbb{R}^3 . I jeho objem lze nalézt pomocí determinantu matice z vektorů, dostáváme tak

$$dV = |\Delta_{\vec{r}}(\vec{a})| du dv dw.$$

Vzhledem k tomu, že tento postup pro výpočet objemu n-rozměrného rovnoběžníka v \mathbb{R}^n je univerzální, dá se čekat, že náš výsledek platí ve všech rozměrech, tedy n-dimenzionální infinitesimála $d[\vec{x}]$ se bude nahrazovat členem $|\Delta_{\vec{r}}(\vec{a})| d[\vec{u}]$, kde $d[\vec{u}]$ značí infinitesimálu v definičním oboru transformace \vec{r} a bude ve výpočtech nahrazena příslušnými infinitesimálami $du \, dv$, $dt \, du \, dv \, dw$ a podobně.

Věta.

Nechť f je funkce spojitá na omezené množině Ω , pro kterou existuje integrál funkce f.

Uvažujme transformaci (vektorovou funkci) \vec{r} : $D[\vec{u}] \mapsto \Omega[\vec{x}]$ takovou, že $\vec{r}[D] = \Omega$, \vec{r} je spojitá na D, prostá a spojitě diferencovatelná na vnitřku D a její jakobián $\Delta_{\vec{r}}$ je tam nenulový.

Pak

$$\int_{\Omega} \cdots \int_{\Omega} f(\vec{x}) d[\vec{x}] = \int_{\Omega} \cdots \int_{\Omega} f(\vec{r}(\vec{u})) |\Delta_{\vec{r}}(\vec{u})| d[\vec{u}].$$

Tím končí náš ilustrovaný úvod do integrálů.

Poznámka: Většina odvození v této kapitole (i těch předchozích) byla intuitivní, což znamená, že díky nim máme pocit, že to všechno dává smysl, ale zároveň jim nemůžeme úplně věřit. Ten pocit porozumění je samozřejmě užitečný, ale kdo chce mít čisté svědomí, podívá se na korektní důkazy v nějaké pořádné učebnici, třeba doporučených skriptech.

 \triangle

Teď si všechny integrály vyzkoušíme v monstrózním příkladě.

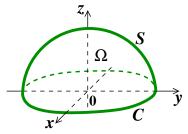
Příklad: Uvažujme těleso

$$\Omega = \{(x, y, z) \in \mathbb{R}^3; \ 0 \le z \le \sqrt{1 - (x^2 + x^2)}\}.$$

Podmínka $0 \leq z$ ukazuje, že tato množina leží nad rovinou xy. Shora je vymezena plochou

$$z = \sqrt{1 - (x^2 + y^2)} \implies 1 = x^2 + y^2 + z^2.$$

Toto je rovnice jednotkové sféry v \mathbb{R}^3 , takže Ω je vlastně horní polokoule o poloměru jedna, budeme tomu říkat dóm.



Bude nás také zajímat plocha S, jmenovitě "střecha" tohoto dómu, tedy horní polosféra, a křivka C, jmenovitě hranice S neboli obvod základny, je to vlastně jednotková kružnice v rovině xy.

Na těchto třech množinách budeme integrovat funkci

$$f(x, y, z) = 4xy + z.$$

Než začneme, tak si všimneme, že všechny tři množiny Ω , S, C jsou zrcadlově souměrné vzhledem k rovinám xz a yz. Také obě složky funkce f jsou symetrické vzhledem k x a y, což znamená, že by mělo stačit najít integrál na jedné čtvrtině těchto množin, například na výseči splňující $x \geq 0$ a $y \geq 0$, říkejme tomu hlavní čtvrtina. Integrály přes zbývající čtvrtiny se pak z toho odvodí.

Konkrétně výraz 4xy je lichý vzhledem kx a také vzhledem ky. Graficky to znamená, že graf 4xy nalevo od osy y (pro x < 0) má stejný tvar jako pravá polovina, ale zrcadlený okolo roviny xy, tedy s opačnými znaménky. Integrování na nějaké množině nalevo pak samozřejmě musí dát opačné číslo ve srovnání se symetrickou integrací napravo.

Máme stejnou symetrii vůči levé a pravé straně osy x. Tyto dvě symetrie se spojí a dovedou nás k následujícímu závěru: Výsledek integrování 4xy na hlavní čtvrtině bude také platit pro protější čtvrtinu, zatímco u sousedních čtvrtin bude mít hodnota integrálu opačné znaménko. Takže ať už

integrujeme 4xy přes celou množinu Ω nebo S či C, vždycky to musí vyjít nula. Jinými slovy, až to za chvíli začneme pořádně počítat, tak se v průběhu výpočtu první člen funkce f vytratí.

Na druhou stranu, výraz z je sudý vzhledem kx a y, takže integrály na všech čtyřech čtvrtinách souhlasí. Z toho všeho vyplývá, že pokud chceme znát hodnoty oněch tří integrálů funkce f (na Ω , S, C), tak stačí integrovat jen výraz z na hlavní čtvrtině a pak výsledné číslo vynásobit čtyřmi. Nás ale výsledky zas tak nezajímají, smyslem příkladu je ukázat metody, tak uděláme plné výpočty. Jdeme na to.

1) **Křivkový integrál** na *C*.

Křivka C je dána rovnicemi $x^2 + y^2 = 1$, z = 0, což napovídá standardní parametrizaci

$$x = x,$$

$$y = \sqrt{1 - x^2},$$

$$z = 0$$

pro jednu její polovinu a parametrizaci se vzorcem $y=-\sqrt{1-x^2}$ pro tu druhou polovinu, u obou se bere $x\in\langle -1,1\rangle$. Přístup s tímto značením je docela populární, ale zde budeme raději následovat značení použité v teoretických úvahách. Proto zavedeme parametrizaci $\vec{r}(t)=(t,\sqrt{1-t^2},0)$ pro $t\in\langle -1,1\rangle$.

Měli bychom najít náhradu za diferenciál ds.

$$\vec{r}'(t) = \left(1, \frac{-t}{\sqrt{1-t^2}}, 0\right) \implies ||\vec{r}'|| = \sqrt{1^2 + \left(\frac{-t}{\sqrt{1-t^2}}\right)^2 + 0^2} = \frac{1}{\sqrt{1-t^2}}$$

$$\implies ds = \frac{1}{\sqrt{1-t^2}} dt.$$

Podobně zpracujeme druhou půlku kružnice parametrizovanou pomocí $\vec{r}(t) = (t, -\sqrt{1-t^2}, 0)$ a dostaneme

$$\oint_C 4xy + z \, ds = \int_{-1}^1 \left(4t\sqrt{1-t^2} + 0\right) \frac{dt}{\sqrt{1-t^2}} + \int_{-1}^1 \left(4t\left(-\sqrt{1-t^2}\right) + 0\right) \frac{dt}{\sqrt{1-t^2}}.$$

Vypočteme:

$$\oint_C 4xy + z \, ds = \int_{-1}^1 4t \, dt + \int_{-1}^1 -4t \, dt = \left[2t^2\right]_{-1}^1 + \left[-2t^2\right]_{-1}^1 = 0 + 0 = 0.$$

Alternativa: V mnoha situacích je nejlepší parametrizace pro kružnici dána polárními souřadnicemi. Jak by to fungovalo zde?

$$x = \cos(\varphi),$$

$$y = \sin(\varphi),$$

$$z = 0.$$

Pracujeme tedy s vektorovou funkcí $\vec{r}(\varphi) = (\cos(\varphi), \sin(\varphi), 0)$ pro $t \in (0, 2\pi)$. Vidíme jednu z výhod tohoto přístupu, jedním vzorcem dostáváme celou kružnici.

Spočítáme vztah mezi diferenciály:

$$\vec{r}'(\varphi) = (-\sin(\varphi), \cos(\varphi), 0) \implies ||\vec{r}'(\varphi)|| = \sqrt{\sin^2(\varphi) + \cos^2(\varphi) + 0^2} = 1$$

 $\implies ds = d\varphi.$

Proto

$$\oint_C 4xy + z \, ds = \int_0^{2\pi} 4\cos(\varphi)\sin(\varphi) + 0 \, d\varphi = \int_0^{2\pi} 2\sin(2\varphi) \, d\varphi = \left[-\cos(2\varphi)\right]_0^{2\pi} = -1 + 1 = 0.$$

2) Povrchový integrál na S.

Povrch S je popsán podmínkami $x^2+y^2+z^2=1, z\geq 0$, což lze ekvivalentně nahradit podmínkou $z=\sqrt{1-x^2-y^2}$. Můžeme tedy vidět S jako graf funkce $g(x,y)=\sqrt{1-x^2-y^2}$. Výše jsme ukázali speciální vzorec pro tento typ povrchu, spočítáme to zde podrobně, abychom viděli, jak se k tomu dojde.

Formálně zavedeme parametrický popis $\vec{r}(u,v) = \left(u,v,\sqrt{1-u^2-v^2}\right)$, kde se (u,v) bere z jednotkového kruhu v \mathbb{R}^2 , říkejme mu D. Jinak řečeno, jako D bereme množinu parametrů (u,v) splňujících $u^2+v^2\leq 1$.

Potřebujeme najít normálový vektor:

$$\begin{split} \vec{N} &= \frac{\partial \vec{r}}{\partial u} \times \frac{\partial \vec{r}}{\partial v} = \left(1, 0, \frac{-u}{\sqrt{1 - u^2 - v^2}}\right) \times \left(0, 1, \frac{-v}{\sqrt{1 - u^2 - v^2}}\right) \\ &= \left(\frac{u}{\sqrt{1 - u^2 - v^2}}, \frac{v}{\sqrt{1 - u^2 - v^2}}, 1\right). \end{split}$$

Proto

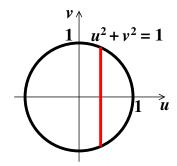
$$dS = \|\vec{N}\| dS[u, v] = \sqrt{\left(\frac{u}{\sqrt{1 - u^2 - v^2}}\right)^2 + \left(\frac{v}{\sqrt{1 - u^2 - v^2}}\right)^2 + 1} dS[u, v]$$
$$= \frac{1}{\sqrt{1 - u^2 - v^2}} dS[u, v].$$

Značku dS[u,v] jsme zapsali rovinný diferenciál v \mathbb{R}^2 s proměnnými u,v.

Můžeme sestavit integrál:

$$\iint\limits_{S} 4xy + z \, dS = \iint\limits_{D} 4uv + \sqrt{1 - u^2 - v^2} \frac{dS[u,v]}{\sqrt{1 - u^2 - v^2}} = \iint\limits_{D} \frac{4uv}{\sqrt{1 - u^2 - v^2}} + 1 \, dS[u,v].$$

Všimněte si, že jsme dostali integrál funkce dvou proměnných na plnohodnotné dvoudimenzionální množině, je to tedy standardní dvojný integrál. Ten počítáme redukcí dimenze pomocí řezů množinou D. Tato množina je jednotkový kruh, tedy je to rutinní situace. Budeme řezat volbou u mezi -1 a 1, pak se v pohybuje mezi $-\sqrt{1-u^2}$ a $\sqrt{1-u^2}$.



$$\iint_{S} 4xy + z \, dS = \int_{-1}^{1} \int_{-1}^{\sqrt{1-u^2}} \frac{4uv}{\sqrt{1-u^2-v^2}} + 1 \, dv \, du.$$

Vnitřní integrál (jeho první člen) se dá zvládnout substitucí:

$$\int \frac{4uv}{\sqrt{1-u^2-v^2}} \, dv = \begin{vmatrix} w = \sqrt{1-u^2-v^2} \\ dw = \frac{-v}{\sqrt{1-u^2-v^2}} \, dv \end{vmatrix} = \int -4u \, dw = -4uw = -4u\sqrt{1-u^2-v^2}.$$

Proto

$$\iint_{S} 4xy + z \, dS = \int_{-1}^{1} \int_{-\sqrt{1-u^{2}}}^{\sqrt{1-u^{2}}} \frac{4uv}{\sqrt{1-u^{2}-v^{2}}} + 1 \, dv \, du$$

$$= \int_{-1}^{1} \left[-4u\sqrt{1-u^{2}-v^{2}} + v \right]_{v=-\sqrt{1-u^{2}}}^{v=\sqrt{1-u^{2}}} du = \int_{-1}^{1} 0 + 2\sqrt{1-u^{2}} \, du.$$

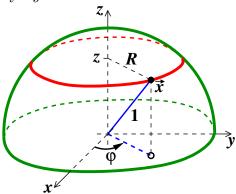
Standardní přístup k tomuto integrálu je přes nepřímou substituci $t = \sin(w)$. Při úpravě budeme díky tomu, že $\cos(w) \ge 0$ na $\left\langle -\frac{\pi}{2}, \frac{\pi}{2} \right\rangle$, moci zjednodušit

$$\sqrt{\cos^2(w)} = |\cos(w)| = \cos(w).$$

Jdeme na to.

$$\iint_{S} 4xy + z \, dS = \int_{-1}^{1} 2\sqrt{1 - u^2} \, du = \int_{-\pi/2}^{\pi/2} 2\sqrt{\cos^2(w)} \cos(w) \, dw = \int_{-\pi/2}^{\pi/2} 2\cos^2(w) \, dw$$
$$= \int_{-\pi/2}^{\pi/2} 1 + \sin(2w) \, dw = \left[w - \frac{1}{2}\cos(2w)\right]_{-\pi/2}^{\pi/2} = \pi.$$

Alternativa: Potřebujeme projít všechny body sféry a další možný přístup je zorganizovat je do kružnic, tedy pomocí vodorovných řezů. Volba nějakého $z \in \langle 0,1 \rangle$ určí na naší polosféře kružnici, jejíž poloměr je $R = \sqrt{1-z^2}$. Pak už jen zbývá projít všemi body této kružnice, což snadno zařídíme pomocí vhodného úhlu φ . Formálně je φ úhel mezi osou x a paprskem spojujícím počátek s projekcí bodu kružnice do roviny xy.



Vzorce pro parametrický popis jsou

$$x = \sqrt{1 - z^2} \cos(\varphi),$$

$$y = \sqrt{1 - z^2} \sin(\varphi),$$

$$z = z.$$

Formálně pracujeme s vektorovou funkcí

$$\vec{r}(\varphi,t) = (\sqrt{1-t^2}\cos(\varphi), \sqrt{1-t^2}\sin(\varphi), t).$$

Zde $\varphi \in \langle 0, 2\pi \rangle$ a $t \in \langle 0, 1 \rangle$. Přesně řečeno, horní polosféru dostaneme tak, že bereme parametry (φ, t) z množiny $D = \langle 0, 2\pi \rangle \times \langle 0, 1 \rangle$. Teď zpracujeme povrchovou substituci:

$$\begin{split} \vec{N} &= \frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial t} = \left(-\sqrt{1-t^2} \sin(\varphi), \sqrt{1-t^2} \cos(\varphi), 0 \right) \times \left(\frac{-t\cos(\varphi)}{\sqrt{1-t^2}}, \frac{-t\sin(\varphi)}{\sqrt{1-t^2}}, 1 \right) \\ &= \left(\sqrt{1-t^2} \cos(\varphi), \sqrt{1-t^2} \sin(\varphi), t\sin^2(\varphi) + t\cos^2(\varphi) \right) = \left(\sqrt{1-t^2} \cos(\varphi), \sqrt{1-t^2} \sin(\varphi), t \right). \end{split}$$

Proto

$$dS = \|\vec{N}\| dS[\varphi, t] = \sqrt{(1 - t^2)\cos^2(\varphi) + (1 - t^2)\sin^2(\varphi) + t^2} dS[\varphi, t]$$

= $\sqrt{1 - t^2 + t^2} d[\varphi, t] = dS[\varphi, t].$

Můžeme sestavit příslušný integrál:

$$\iint\limits_{S} 4xy + z \, dS = \iint\limits_{D} 4\sqrt{1 - t^2} \cos(\varphi) \sqrt{1 - t^2} \sin(\varphi) + t \, dS[\varphi, t].$$

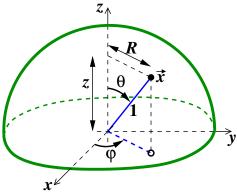
Zase máme napravo standardní dvojný integrál. Tentokráte je D obdélník, takže si můžeme volně

vybrat pořadí integrace a integrační meze jsou zjevné.

$$\iint_{S} 4xy + z \, dS = \int_{0}^{1} \int_{0}^{2\pi} 2(1 - t^{2}) \sin(2\varphi) + t \, d\varphi \, dt = \int_{0}^{1} \left[-(1 - t^{2}) \cos(2\varphi) + t\varphi \right]_{\varphi=0}^{\varphi=2\pi} dt$$
$$= \int_{0}^{1} 0 + 2\pi t \, dt = \left[\pi t^{2} \right]_{0}^{1} dt = \pi.$$

Tento přístup se zdá příjemnější.

Alternativa: Velmi příjemný popis sféry pochází z třírozměrných polárních souřadnic, tedy ze sférických souřadnic. Pro daný bod na sféře máme polární úhel θ vystihující, jak se tento bod odklání od osy z. Když máme θ , tak už víme, jakou má bod souřadnici z, a také známe poloměr kružnice, na které leží. Jeho polohu na kružnici pak určíme pomocí azimutu φ jako u předchozího přístupu.



Dostáváme

$$x = \cos(\varphi)\sin(\theta),$$

$$y = \sin(\varphi)\sin(\theta),$$

$$z = \cos(\theta),$$

tedy pracujeme s funkcí

$$\vec{r}(\varphi, \theta) = (\cos(\varphi)\sin(\theta), \sin(\varphi)\sin(\theta), \cos(\theta)).$$

Horní polosféru dostaneme, když bereme $\varphi \in \langle 0, 2\pi \rangle$ a $\theta \in \langle 0, \frac{\pi}{2} \rangle$, neboli máme $D = \langle 0, 2\pi \rangle \times \langle 0, \frac{\pi}{2} \rangle$. Jdeme na to:

$$\begin{split} \vec{N} &= \frac{\partial \vec{r}}{\partial \varphi} \times \frac{\partial \vec{r}}{\partial \theta} = \left(-\sin(\varphi)\sin(\theta), \cos(\varphi)\sin(\theta), 0 \right) \times \left(\cos(\varphi)\cos(\theta), \sin(\varphi)\cos(\theta), -\sin(\theta) \right) \\ &= \left(-\cos(\varphi)\sin^2(\theta), -\sin(\varphi)\sin^2(\theta), -\sin^2(\varphi)\sin(\theta)\cos(\theta) - \cos^2(\varphi)\sin(\theta)\cos(\theta) \right) \\ &= \left(-\cos(\varphi)\sin^2(\theta), -\sin(\varphi)\sin^2(\theta), -\sin(\theta)\cos(\theta) \right). \end{split}$$

Proto

$$dS = \|\vec{N}\| dS[\varphi, \theta] = \sqrt{\cos^2(\varphi) \sin^4(\theta) + \sin^2(\varphi) \sin^4(\theta) + \sin^2(\theta) \cos^2(\theta)} dS[\varphi, \theta]$$
$$= \sqrt{\sin^4(\theta) + \sin^2(\theta) \cos^2(\theta)} dS[\varphi, \theta] = \sqrt{\sin^2(\theta) [\sin^2(\theta) + \cos^2(\theta)]} dS[\varphi, \theta]$$
$$= |\sin(\theta)| dS[\varphi, \theta] = \sin(\theta) dS[\varphi, \theta].$$

Absolutní hodnotu jsme mohli vynechat, protože sinus není nikdy záporný pro $\theta \in \langle 0, \frac{\pi}{2} \rangle$. Integrování na obdélníku je zase snadné.

$$\iint_{S} 4xy + z \, dS = \iint_{D} \left(4\cos(\varphi)\sin(\theta)\sin(\varphi)\sin(\theta) + \cos(\theta) \right) \sin(\theta) \, dS[\varphi, \theta]$$

$$= \int_{0}^{\pi/2} \int_{0}^{2\pi} 2\sin(2\varphi)\sin^{3}(\theta) + \frac{1}{2}\sin(2\theta)\,d\varphi\,d\theta$$

$$= \int_{0}^{\pi/2} \left[-\cos(2\varphi)\sin^{3}(\theta) + \frac{1}{2}\sin(2\theta)\varphi\right]_{\varphi=0}^{\varphi=2\pi}d\theta$$

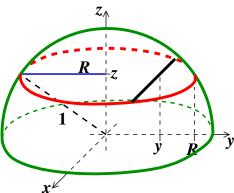
$$= \int_{0}^{\pi/2} 0 + \pi\sin(2\theta)\,d\theta = \left[-\frac{1}{2}\pi\cos(2\theta)\right]_{0}^{\pi/2} = \pi.$$

Tohle taky docela šlo.

3) Integrál na Ω .

Protože Ω neboli náš dóm má neprázdný vnitřek, je to opravdu třídimenzionální množina. Integrál přes Ω je tedy standardní trojný integrál. Použijeme běžný postup, tedy redukci dimenze pomocí řezů.

Měli bychom vybrat první směr, a jako dobrý nápad se zdá řezat zvolením $z \in \langle 0, 1 \rangle$ neboli rovnoběžně s rovinou xy, protože takový řez bude vlastně kruh. Snadno najdeme jeho poloměr coby $R = \sqrt{1-z^2}$.



Potřebujeme projít body takového kruhu, proto jej rozřežeme na proužky. Zvolíme nějaké y, přirozeně mezi -R a R, to určí pruh, na kterém se x může pohybovat mezi $-\sqrt{R^2-y^2}$ a $\sqrt{R^2-y^2}$. Dostáváme následující trojnásobný integrál:

$$\iiint\limits_{\Omega} 4xy + z\, dV = \int\limits_{0}^{1} \int\limits_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}} \int\limits_{-\sqrt{1-z^2-y^2}}^{\sqrt{1-z^2-y^2}} 4xy + z\, dx\, dy\, dz.$$

Jako obvykle začneme zevnitř.

$$\begin{split} \iiint_{\Omega} 4xy + z \, dV &= \int_{0}^{1} \int_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}} \left[2x^2y + zx \right]_{x=-\sqrt{1-z^2-y^2}}^{x=\sqrt{1-z^2-y^2}} dy \, dz \\ &= \int_{0}^{1} \int_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}} 0 + 2z\sqrt{1-z^2-y^2} \, dy \, dz. \end{split}$$

Zde je standardním postupem substituce $y=\sqrt{1-z^2}\sin(w),\ dy=\sqrt{1-z^2}\cos(w)\,dw,$ kde se z

bere jako parametr, takže pokračujeme následovně.

$$\iiint_{\Omega} 4xy + z \, dV = \int_{0}^{1} \int_{-\pi/2}^{\pi/2} 2z \sqrt{(1 - z^2) - (1 - z^2) \sin^2(w)} \sqrt{1 - z^2} \cos(w) \, dw \, dz$$

$$= \int_{0}^{1} \int_{-\pi/2}^{\pi/2} 2z \sqrt{1 - z^2} \sqrt{1 - \sin^2(w)} \sqrt{1 - z^2} \cos(w) \, dw \, dz$$

$$= \int_{0}^{1} \int_{-\pi/2}^{\pi/2} 2z (1 - z^2) \cos^2(w) \, dw \, dz = \int_{0}^{1} \int_{-\pi/2}^{\pi/2} z (1 - z^2) (1 + \cos(2w)) \, dw \, dz$$

$$= \int_{0}^{1} \left[z (1 - z^2) (w - \frac{1}{2} \sin(2w)) \right]_{w = -\pi/2}^{w = \pi/2} dz = \int_{0}^{1} \pi z (1 - z^2) \, dz$$

$$= \left| \int_{0}^{\omega} \frac{1 - z^2}{d\omega} \right| = \int_{0}^{0} -\frac{1}{2} \pi \omega \, d\omega = \left[-\frac{1}{2} \pi \frac{1}{2} \omega^2 \right]_{1}^{0} = \frac{\pi}{4}.$$

Jde to i jinak? Mohli bychom teď zkusit svislé řezy dómem, ty budou ve tvaru půlkruhu, ale vede to na obdobné integrály. Bude zajímavější zkusit jiné přístupy.

Alternativa: Použijeme sférické souřadnice

$$x = r \cos(\varphi) \sin(\theta),$$

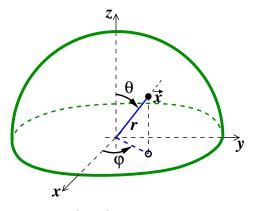
$$y = r \sin(\varphi) \sin(\theta),$$

$$z = r \cos(\theta).$$

Protože počet výchozích i cílových proměnných souhlasí, je to substituce. Formálně pracujeme s funkcí

$$\vec{r}(r,\varphi,\theta) = \left(r\cos(\varphi)\sin(\theta),r\sin(\varphi)\sin(\theta),r\cos(\theta)\right).$$

Celou horní polokouli dostaneme, když vezmeme $r \in \langle 0, 1 \rangle$, $\varphi \in \langle 0, 2\pi \rangle$ a $\theta \in \langle 0, \frac{\pi}{2} \rangle$.



Formálně bychom brali \vec{r} jako zobrazení z množiny $D=\langle 0,1\rangle \times \langle 0,2\pi\rangle \times \langle 0,\frac{\pi}{2}\rangle$ na Ω , obě množiny jsou třírozměrné.

Potřebujeme najít jakobián.

$$\Delta_{\vec{r}} = \det \begin{pmatrix} \frac{\partial r_1}{\partial r} & \frac{\partial r_1}{\partial \varphi} & \frac{\partial r_1}{\partial \theta} \\ \frac{\partial r_2}{\partial r} & \frac{\partial r_2}{\partial \varphi} & \frac{\partial r_2}{\partial \theta} \\ \frac{\partial r_3}{\partial r} & \frac{\partial r_3}{\partial \varphi} & \frac{\partial r_3}{\partial \theta} \end{pmatrix} = \det \begin{pmatrix} \cos(\varphi)\sin(\theta) & -r\sin(\varphi)\sin(\theta) & r\cos(\varphi)\cos(\theta) \\ \sin(\varphi)\sin(\theta) & r\cos(\varphi)\sin(\theta) & r\sin(\varphi)\cos(\theta) \\ \cos(\theta) & 0 & -r\sin(\theta) \end{pmatrix}$$

$$= -r^2\cos^2(\varphi)\sin^3(\theta) - r^2\sin^2(\varphi)\sin(\theta)\cos^2(\theta)$$

$$- r^2\cos^2(\varphi)\sin(\theta)\cos^2(\theta) - r^2\sin^2(\varphi)\sin^3(\theta)$$

$$= -r^2\left(\left[\cos^2(\varphi) + \sin^2(\varphi)\right]\sin^3(\theta) + \left[\sin^2(\varphi) + \cos^2(\varphi)\right]\sin(\theta)\cos^2(\theta)\right)$$

$$= -r^2\left(\sin^3(\theta) + \sin(\theta)\cos^2(\theta)\right) = -r^2\sin(\theta)\left(\sin^2(\theta) + \cos^2(\theta)\right)$$

$$= -r^2\sin(\theta).$$

Aplikujeme absolutní hodnotu a vezmeme v úvahu, že $\theta \in \left<0, \frac{\pi}{2}\right>$ a tam je sinus kladný či nula, takže dostáváme

$$dV = |-r^2 \sin(\theta)| dV[r, \varphi, \theta] = r^2 \sin(\theta) dV[r, \varphi, \theta].$$

Jsme připravení aplikovat substituci. Nový integrál bude trojný integrál na množině D, což je

třídimenzionální hranol. Máme tedy svobodnou volbu pořadí integrování a integrační meze jsou zjevné.

$$\begin{split} \iint\limits_{\Omega} 4xy + z \, dV &= \iiint\limits_{D} \left(4r \cos(\varphi) \sin(\theta) r \sin(\varphi) \sin(\theta) + r \cos(\theta) \right) r^2 \sin(\theta) \, dV[r, \varphi, \theta] \\ &= \int\limits_{0}^{1} \int\limits_{0}^{\pi/2} \int\limits_{0}^{2\pi} \left(2r^2 \sin(2\varphi) \sin^2(\theta) + r \cos(\theta) \right) r^2 \sin(\theta) \, d\varphi \, d\theta \, dr \\ &= \int\limits_{0}^{1} \int\limits_{0}^{\pi/2} \int\limits_{0}^{2\pi} 2r^4 \sin(2\varphi) \sin^3(\theta) + r^3 \cos(\theta) \sin(\theta) \, d\varphi \, d\theta \, dr \\ &= \int\limits_{0}^{1} \int\limits_{0}^{\pi/2} \left[-r^4 \cos(2\varphi) \sin^3(\theta) + r^3 \cos(\theta) \sin(\theta) \varphi \right]_{\varphi=0}^{\varphi=2\pi} \, d\theta \, dr \\ &= \int\limits_{0}^{1} \int\limits_{0}^{\pi/2} 0 + 2\pi r^3 \cos(\theta) \sin(\theta) \, d\theta \, dr = \int\limits_{0}^{1} \int\limits_{0}^{\pi/2} \pi r^3 \sin(2\theta) \, d\theta \, dr \\ &= \int\limits_{0}^{1} \left[-\frac{1}{2}\pi r^3 \cos(2\theta) \right]_{\theta=0}^{\theta=\pi/2} dr = \int\limits_{0}^{1} \pi r^3 \, dr = \left[\pi \frac{1}{4} r^4 \right]_{0}^{1} = \frac{\pi}{4}. \end{split}$$

Alternativa: Další zajímavý přístup je použít válcové souřadnice. Standardní převod je

$$x = r\cos(\varphi),$$

$$y = r\sin(\varphi),$$

$$z = z.$$

To odpovídá vektorové funkci

$$\vec{r}(z, r, \varphi) = (r\cos(\varphi), r\sin(\varphi), z).$$

Potřebujeme najít správný definiční obor D pro \vec{r} tak, aby obrazem byl náš dóm. Zjevná omezení jsou $z \in \langle 0, 1 \rangle$, $r \in \langle 0, 1 \rangle$ a $\varphi \in \langle 0, 2\pi \rangle$. Kdybychom ale jako D vzali hranol $\langle 0, 1 \rangle \times \langle 0, 1 \rangle \times \langle 0, 2\pi \rangle$, tak by obrazem byl válec. Potřebujeme omezit poloměr r v závislosti na výšce neboli na z, a snadno odvodíme, že tím správným omezením je $0 \le r \le \sqrt{1-z^2}$. Použijeme proto množinu

$$D = \{(z, r, \varphi) \in \mathbb{R}^3; \ 0 \le \varphi \le 2\pi, \ 0 \le z \le 1, 0 \le r \le \sqrt{1 - z^2}\}.$$

Protože omezení pro r závisí na z, bude záležet na pořadí integrace, r musí být v integrálu "hlouběji" než z, aby se přes něj integrovalo dříve.

$$\Delta_{\vec{r}} = \det \begin{pmatrix} \frac{\partial r_1}{\partial z} & \frac{\partial r_1}{\partial r} & \frac{\partial r_1}{\partial \varphi} \\ \frac{\partial r_2}{\partial z} & \frac{\partial r_2}{\partial r} & \frac{\partial r_2}{\partial \varphi} \\ \frac{\partial r_3}{\partial z} & \frac{\partial r_3}{\partial r} & \frac{\partial r_3}{\partial \varphi} \end{pmatrix} = \det \begin{pmatrix} 0 & \cos(\varphi) & -r\sin(\varphi) \\ 0 & \sin(\varphi) & r\cos(\varphi) \\ 1 & 0 & 0 \end{pmatrix} = r\cos^2(\varphi) + r\sin^2(\varphi) = r$$

$$\implies dV = r \, dV[z, r, \varphi].$$

Proto

$$\iiint\limits_{\Omega} 4xy + z \, dV = \iiint\limits_{D} \left(4r\cos(\varphi)r\sin(\varphi) + z\right)r \, dV[z, r, \varphi]$$

$$= \int_{0}^{1} \int_{0}^{\sqrt{1-z^2}} \int_{0}^{2\pi} 2r^3 \sin(2\varphi) + zr \, d\varphi \, dr \, dz$$

$$= \int_{0}^{1} \int_{0}^{\sqrt{1-z^2}} \left[-r^3 \cos(2\varphi) + zr\varphi \right]_{\varphi=0}^{\varphi=2\pi} dr \, dz = \int_{0}^{1} \int_{0}^{\sqrt{1-z^2}} 0 + 2\pi zr \, dr \, dz$$

$$= \int_{0}^{1} \left[\pi zr^2 \right]_{r=0}^{r=\sqrt{1-z^2}} dz = \int_{0}^{1} \pi z(1-z^2) \, dz = \left| \frac{w = 1-z^2}{dw = -2z \, dz} \right|$$

$$= \int_{1}^{0} -\frac{1}{2}\pi w \, dw = \left[-\frac{1}{4}\pi w^2 \right]_{1}^{0} = \frac{\pi}{4}.$$

Alternativa: To, že jsme si nemohli vybrat pořadí integrace, je trochu nemilé. Dá se to obejít následujícím zajímavým trikem. Budeme vždy brát $r \ge \langle 0, 1 \rangle$, ale skutečný poloměr pak při výpočtu x a y příslušně upravíme. Uvažujeme tedy následující transformaci:

$$x = r\sqrt{1 - z^2}\cos(\varphi),$$

$$y = r\sqrt{1 - z^2}\sin(\varphi),$$

$$z = z.$$

To odpovídá vektorové funkci

$$\vec{r}(z, r, \varphi) = \left(r\sqrt{1 - z^2}\cos(\varphi), r\sqrt{1 - z^2}\sin(\varphi), z\right)$$

a teď v \vec{r} bereme $z \in \langle 0, 1 \rangle$, $r \in \langle 0, 1 \rangle$ a $\varphi \in \langle 0, 2\pi \rangle$ neboli $D = \langle 0, 1 \rangle \times \langle 0, 1 \rangle \times \langle 0, 2\pi \rangle$.

$$\Delta_{\vec{r}} = \det \begin{pmatrix} \frac{\partial r_1}{\partial z} & \frac{\partial r_1}{\partial r} & \frac{\partial r_1}{\partial \varphi} \\ \frac{\partial r_2}{\partial z} & \frac{\partial r_2}{\partial r} & \frac{\partial r_2}{\partial \varphi} \\ \frac{\partial r_3}{\partial z} & \frac{\partial r_3}{\partial r} & \frac{\partial r_3}{\partial \varphi} \end{pmatrix} = \det \begin{pmatrix} r \frac{-z}{\sqrt{1-z^2}} \cos(\varphi) & \sqrt{1-z^2} \cos(\varphi) & -r\sqrt{1-z^2} \sin(\varphi) \\ r \frac{-z}{\sqrt{1-z^2}} \sin(\varphi) & \sqrt{1-z^2} \sin(\varphi) & r\sqrt{1-z^2} \cos(\varphi) \end{pmatrix} \\
= r(1-z^2)\cos^2(\varphi) + r(1-z^2)\sin^2(\varphi) = r(1-z^2) \\
\implies dV = r(1-z^2) dV[z, r, \varphi].$$

Proto

$$\iint_{\Omega} 4xy + z \, dV = \iiint_{D} \left(4r\sqrt{1 - z^2} \cos(\varphi)r\sqrt{1 - z^2} \sin(\varphi) + z \right) r(1 - z^2) \, dV[z, r, \varphi]$$

$$= \int_{0}^{1} \int_{0}^{1} \int_{0}^{2\pi} 2r^3 (1 - z^2)^2 \sin(2\varphi) + z(1 - z^2) r \, d\varphi \, dr \, dz$$

$$= \int_{0}^{1} \int_{0}^{1} \left[-r^3 (1 - z^2)^2 \cos(2\varphi) + z(1 - z^2) r \varphi \right]_{\varphi=0}^{\varphi=2\pi} dr \, dz$$

$$= \int_{0}^{1} \int_{0}^{1} 0 + 2\pi z (1 - z^2) r \, dr \, dz = \int_{0}^{1} \left[\pi z (1 - z^2) r^2 \right]_{r=0}^{r=1} dz$$

$$= \int_{0}^{1} \pi z (1 - z^2) \, dz = \frac{\pi}{4}.$$

Došli jsme ke stejnému integrálu jako v předchozím výpočtu, tak jsme jen zkopírovali výsledek.

U přímé aplikace válcových souřadnic se snáze hledal jakobián, u této alternativy to bylo trochu delší, ale zato jsme neměli starosti s novou integrační doménou D a mohli si vybrat pořadí integrace. Já bych to viděl na remízu.

 \triangle

Když se na ty výpočty podíváme s odstupem, můžeme si všimnout, že všechny probíhají stejným způsobem. Danou integrační množinu si popíšeme tak, že původní souřadnice nahradíme nějakými novými (tj. parametry). Integrál přes danou množinu se pak nahradí integrálem přes novou množinu D, ke kterému přistupujeme jako k nové úloze. Rozdíl mezi substitucí a speciálními typy integrálů se projeví v kroku, kdy nahrazujeme diferenciál.

- Pokud počet původních a nových proměnných souhlasí, jde o substituci a při převodu diferenciálů použijeme jakobián.
- Pokud počet proměnných nesouhlasí a je jen jedna nová proměnná (parametr), jde o přepis křivkového integrálu a na to máme vzorec pro diferenciály.
- Pokud počet proměnných nesouhlasí a jsou dvě nové proměnné (parametry), jde o přepis povrchového integrálu. Na to máme vzorec pro diferenciály, pokud byly původní proměnné tři.

Většina substitucí v tomto příkladě byla standardní, což se týká zejména polárních, sférických a válcových souřadnic.

$$x = r \cos(\varphi),$$

$$y = r \sin(\varphi) \implies dS = r dS[r, \varphi]$$

$$x = \cos(\varphi) \sin(\theta),$$

$$y = \sin(\varphi) \sin(\theta), \implies dS = \sin(\theta) dS[\varphi, \theta]$$

$$z = \cos(\theta)$$

$$x = r \cos(\varphi) \sin(\theta),$$

$$y = r \sin(\varphi) \sin(\theta), \implies dV = r^2 \sin(\theta) dV[r, \varphi, \theta]$$

$$z = r \cos(\theta)$$

$$x = r \cos(\varphi),$$

$$y = r \sin(\varphi), \implies dV = r dV[z, r, \varphi]$$

Ještě v nedávné minulosti se studenti učili tyto převodní vzorce pro diferenciály nazpaměť. Není pokrok úžasný?

7. Více o derivaci

Zde opouštíme ilustrovaný úvod a jednak spíše pro úplnost zpřesníme naše předchozí často jen geometrické úvahy, ale hlavně představíme další důležitá témata.

7a. Derivace a diferenciální operátory

V kapitole 3 jsme představili směrové derivace.

Definice.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Nechť \vec{u} je vektor z \mathbb{R}^n .

Řekneme, že funkce f je **diferencovatelná** v bodě \vec{a} ve směru \vec{u} , jestliže limita $\lim_{t\to 0} \left(\frac{f(\vec{a}+t\vec{u})-f(\vec{a})}{t}\right)$ existuje konečná.

Pak definujeme (směrovou) derivaci f v bodě \vec{a} ve směru \vec{u} jako

$$D_{\vec{u}}f(\vec{a}) = \lim_{t \to 0} \left(\frac{f(\vec{a} + t\vec{u}) - f(\vec{a})}{t} \right).$$

Napsali jsme, že směrová derivace v jednom konkrétním směru se chová jako běžná derivace. Potvrdíme to formálními tvrzeními.

Věta.

Nechť f,g jsou funkce definované na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$ a $c \in \mathbb{R}$. Nechť \vec{u} je vektor z \mathbb{R}^n .

Předpokládejme, že f a g jsou diferencovatelné v \vec{a} ve směru \vec{u} . Pak jsou v \vec{a} ve směru \vec{u} diferencovatelné i funkce cf, f+g, f-g, $f\cdot g$, $\frac{f}{g}$ (pokud $g(\vec{a})\neq 0$) a platí

$$D_{\vec{u}}(cf)(\vec{a}) = cD_{\vec{u}}f(\vec{a}),$$

$$D_{\vec{u}}(f+g)(\vec{a}) = D_{\vec{u}}f(\vec{a}) + D_{\vec{u}}g(\vec{a}),$$

$$D_{\vec{u}}(f-g)(\vec{a}) = D_{\vec{u}}f(\vec{a}) - D_{\vec{u}}g(\vec{a}),$$

$$D_{\vec{u}}(f \cdot g)(\vec{a}) = D_{\vec{u}}f(\vec{a}) \cdot g(\vec{a}) + f(\vec{a}) \cdot D_{\vec{u}}g(\vec{a}),$$

$$D_{\vec{u}}(\frac{f}{g})(\vec{a}) = \frac{D_{\vec{u}}f(\vec{a}) \cdot g(\vec{a}) - f(\vec{a}) \cdot D_{\vec{u}}g(\vec{a})}{g^2(\vec{a})}.$$

Dokonce máme i následující. Připomeňme, že Věta o střední hodnotě pro funkci jedné proměnné vede na rovnost f(y) - f(x) = f'(c)(y - x).

Věta. (o střední hodnotě)

Nechť $f: D(f) \mapsto \mathbb{R}$ je funkce, kde $D(f) \subseteq \mathbb{R}^n$. Nechť $\vec{x}, \vec{y} \in D(f)$.

Předpokládejme, že úsečka $\langle \vec{x}, \vec{y} \rangle$ leží v D(f) a f má na této úsečce derivaci ve směru $\vec{u} = \frac{\vec{y} - \vec{x}}{\|\vec{y} - \vec{x}\|}$. Pak existuje $\vec{c} \in \langle \vec{x}, \vec{y} \rangle$ takové, že

$$f(\vec{y}) - f(\vec{x}) = D_{\vec{u}}f(\vec{c}) \cdot ||\vec{y} - \vec{x}|| = D_{\vec{y} - \vec{x}}f(\vec{c}).$$

Úsečkou $\langle \vec{a}, \vec{x} \rangle$ v \mathbb{R}^n míníme množinu bodů ve tvaru $\vec{a} + t(\vec{x} - \vec{a})$ pro $t \in \langle 0, 1 \rangle$. Věta tedy říká, že existuje $t \in (0, 1)$ takové, že

$$f(\vec{y}) - f(\vec{x}) = D_{\vec{u}}f(t\vec{x} + (1-t)\vec{y}) \cdot ||\vec{y} - \vec{x}|| = D_{\vec{u} - \vec{x}}f(t\vec{x} + (1-t)\vec{y}).$$

Protože parciální derivace jsou jen speciálním případem směrových derivací, věta o operacích nám potvrzuje, že při počítání parciálních derivací můžeme používat běžné vzorce. Když už mluvíme o parciálních derivacích, přidejme i formální definici prostorů \mathbb{C}^k .

Definice.

Nechť D je oblast v \mathbb{R}^n .

Definujeme $C^k(D)$ jako množinu všech funkcí $f: D \mapsto \mathbb{R}$, které mají všechny parciální derivace až po řád k a ty jsou spojité na D.

Při počítání parciálních derivací vyššího řádu je příjemné, když se nemusíme starat o pořadí derivování. Ukážeme obecnější tvrzení.

Věta.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$.

Předpokládejme, že pro $\vec{u}, \vec{v} \in \mathbb{R}^n$ existují $D_{\vec{u}}(D_{\vec{v}}f)$ a $D_{\vec{v}}(D_{\vec{u}}f)$ na nějakém okolí \vec{a} a jsou spojité v \vec{a} . Pak $D_{\vec{u}}(D_{\vec{v}}f)(\vec{a}) = D_{\vec{v}}(D_{\vec{u}}f)(\vec{a})$.

Ještě nám chybí derivování složené funkce. Ve více dimenzích to je docela zajímavé, takže si to necháme na samostatnou sekci 7b.

Gradient jsou vlastně jen poskládané parciální derivace. Není těžké si rozmyslet, že vzorce ve větě o operacích lze poskládat i do vektorů. Dostaneme tak pravidla pro gradient.

Věta.

Nechť f,g jsou funkce definované na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$, nechť $c \in \mathbb{R}$. Předpokládejme, že pro obě funkce existuje gradient $\nabla f(\vec{a})$ a $\nabla g(\vec{a})$. Pak mají gradient i funkce cf, f+g, $f\cdot g$, $\frac{f}{g}$ (pokud $g(\vec{a})\neq 0$) a platí

$$\begin{split} &\nabla(cf)(\vec{a}) = c\nabla f(\vec{a}),\\ &\nabla(f+g)(\vec{a}) = \nabla f(\vec{a}) + \nabla g(\vec{a}),\\ &\nabla(f\cdot g)(\vec{a}) = \nabla f(\vec{a})g(\vec{a}) + f(\vec{a})\nabla g(\vec{a}),\\ &\nabla\Big(\frac{f}{g}\Big)(\vec{a}) = \frac{\nabla f(\vec{a})g(\vec{a}) - f(\vec{a})\nabla g(\vec{a})}{[g(\vec{a})]^2}. \end{split}$$

Jako obvykle je možné tyto vzorce interpretovat jako rovnost funkcí na množině, v tomto případě ovšem vektorových funkcí.

$$\nabla(cf + g) = c\nabla f + \nabla g,$$

$$\nabla(f \cdot g) = \nabla f \cdot g + f \cdot \nabla g,$$

$$\nabla\left(\frac{f}{g}\right) = \frac{\nabla f \cdot g - f \cdot \nabla g}{g^2}.$$

Dávají ty vzorce smysl? Víme, že vektory je možné sčítat a odčítat, popřípadě násobit či dělit skalárem. Zde je užitečné si uvědomit, že skalár klidně může být proměnný, tedy vektor můžeme násobit funkcí.

Představme si například funkce $f(x,y) = xy^2$ a $g(x,y) = e^{x+y}$. Pak je ∇f vektorová funkce a můžeme ji vynásobit skalární funkcí g:

$$\nabla f \cdot g = (y^2, 2xy) \cdot e^{x+y} = (y^2 e^{x+y}, 2xy e^{x+y}).$$

Vidíme, že výsledkem je vektorová funkce. To vysvětluje, proč mají vzorce v součinovém a podílovém pravidle smysl.

Užitečnost gradientu jsme už viděli v kapitole 3, zde si přidáme jednu povědomou větu.

Věta.

Nechť D je oblast v \mathbb{R}^n a $f \in C^1(D)$. Jestliže $\nabla f = \vec{0}$ na D, pak je f konstantní na D.

U vektorových funkcí je to zajímavější. Připomeňme, že pro otevřenou množinu D označujeme symbolem $[C(D)]^m$ množinu všech vektorových funkcí $F: D \mapsto \mathbb{R}^m$, které jsou spojité, tedy jejichž složky F_j jsou z C(D). Podobně symbolem $[C^k(D)]^m$ označujeme množinu všech vektorových funkcí $F: D \mapsto \mathbb{R}^m$, jejichž složky F_j jsou z $C^k(D)$, tedy mají parciální derivace až do řádu k, které jsou navíc na D spojité. Následující věty aplikujeme obvykle na funkce z $[C^1(D)]^m$.

Uvedli jsme tři diferenciální operátory, jmenovitě Jacobiho matici coby zobecnění gradientu, divergenci a rotaci. Začneme tím, že potvrdíme linearitu, což je tím hlavním, co od operátorů očekáváme. Napíšeme globální verze.

Věta.

Nechť $F, G: D \mapsto \mathbb{R}^m$ jsou vektorové funkce, kde $D \subseteq \mathbb{R}^n$. Nechť $c \in \mathbb{R}$. Pak

$$J_{cF+G} = cJ_F + J_G,$$

$$\operatorname{div}(cF+G) = c\operatorname{div}(F) + \operatorname{div}(G),$$

$$\operatorname{rot}(cF+G) = c\operatorname{rot}(F) + \operatorname{rot}(G)$$

všude, kde tyto výrazy existují.

Jak už jsme si rozmysleli dříve, vektorovou funkci je možno vynásobit skalární funkcí. Dvě vektorové funkce F, G má smysl sčítat či odčítat, můžeme je také násobit skalárním či vektorovým součinem.

Věta.

Nechť $F,G:D\mapsto\mathbb{R}^m$ jsou vektorové funkce, kde $D\subseteq\mathbb{R}^n$. Nechť f,g jsou funkce $D\mapsto\mathbb{R}.$ Pak

(i)
$$J_{fG} = (\nabla f)^T G + f J_G$$
,

(ii)
$$\operatorname{div}(fG) = f \operatorname{div}(G) + \nabla f \bullet G$$
,

(iii)
$$\operatorname{div}\left(\frac{G}{f}\right) = \frac{f\operatorname{div}(G) - \nabla f \bullet G}{f^2},$$

(iv)
$$\operatorname{rot}(fG) = f \operatorname{rot}(G) + \nabla f \times G$$
,

(v)
$$\operatorname{rot}\left(\frac{G}{f}\right) = \frac{f \operatorname{rot}(G) - \nabla f \times G}{f^2},$$

(vi)
$$\nabla(F \bullet G) = F \cdot J_G + G \cdot J_F$$
,

(vii)
$$\operatorname{div}(F \times G) = G \bullet \operatorname{rot}(F) - F \bullet \operatorname{rot}(G)$$
,

(viii)
$$\operatorname{rot}(F \times G) = F \cdot \operatorname{div}(G) - G \cdot \operatorname{div}(F) + G \cdot (J_F)^T - F \cdot (J_G)^T$$
,

(ix)
$$rot(f\nabla g) = \nabla f \times \nabla g$$
,

(x)
$$\operatorname{div}(\operatorname{rot}(F)) = 0$$
,

(xi)
$$rot(\nabla f) = 0$$

všude, kde tyto výrazy existují.

Při interpretaci je potřeba si uvědomit, že gradient coby operátor má přednost před dalšími operacemi, takže třeba vzorec (ix) napravo interpretujeme jako $(\nabla f) \times (\nabla g)$.

Všechny výrazy dávají smysl. Například zrovna v (ix) se gradient g (tedy vektor) násobí skalární funkcí, vznikne tak vektorová funkce. Z ní je možné udělat rotaci, výsledkem je zase vektor. Na pravé straně vidíme vektorový součin dvou vektorů, což také dává vektor. Na levé a pravé straně tedy porovnáváme vektory (a stejného rozměru), což dává smysl.

Podívejme se na pravou stranu (i). Gradient se bere jako řádkový vektor, takže $(\nabla f)^T$ je sloupcový vektor neboli matice o rozměru $n \times 1$. Vektorová funkce $G = (G_1, \ldots, G_n)$ je řádkový vektor neboli matice o rozměru $1 \times n$. Když tyto dvě matice běžným maticovým způsobem vynásobíme,

dostaneme matici $n \times n$, kterou pak lze sečíst s Jacobiho maticí J_G stejného rozměru vynásobenou funkcí f, která zde násobí matici jako skalár.

Ve vzorci (vi) máme nalevo nejprve skalární součin dvou vektorů, výsledkem je tedy číslo neboli skalární funkce více proměnných. Aplikací gradientu vznikne vektor (řádkový). Na pravé straně vidíme vektorovou funkci neboli řádkový vektor, jak násobí matici. V tomto případě jde o běžné maticové násobení matice $1 \times n$ s maticí $n \times n$, vznikne tedy matice $1 \times n$ neboli řádkový vektor. Typy výrazů nalevo i napravo souhlasí.

Na pravé straně (ii) pak na konci vidíme skalární součin dvou řádkových vektorů, výsledkem je číslo neboli skalární funkce, což lze zakomponovat do zlomku.

Gradient je možné vnímat jako diferenciální operátor, který lze symbolicky značit jako

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n}\right).$$

Na funkci se pak aplikuje takto:

$$\left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\right) f = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}\right).$$

Úžasná věc je, že se s ním dá zacházet jako se skutečným vektorem. Například víme, že derivaci ve směru lze počítat pomocí gradientu a skalárního součinu: $D_{\vec{u}}f = \nabla f \bullet \vec{u}$. Podívejme se na to blíže:

$$\nabla f \bullet \vec{u} = \sum \frac{\partial f}{\partial x_i} u_i = \sum u_i \frac{\partial f}{\partial x_i} = \left(\sum u_i \frac{\partial}{\partial x_i}\right) f = (\vec{u} \bullet \nabla) f.$$

Aplikací skalárního součinu na vektor (u_1, \ldots, u_n) a abstraktní vektor $(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n})$ vznikne diferenciální operátor $\sum u_i \frac{\partial}{\partial x_i}$, který realizuje směrovou derivaci. Symbolicky: $D_{\vec{u}} = \vec{u} \cdot \nabla$.

Zde se uplatňuje konvence, že když máme $\frac{\partial}{\partial x_i}f$, tak se derivuje a je hotovo: $\frac{\partial f}{\partial x_i}$. V opačném pořadí $f\frac{\partial}{\partial x_i}$ se nederivuje, ale modifikuje diferenciální operátor násobením. Neplatí tedy, že $\nabla \bullet F$ by bylo totéž co $F \bullet \nabla$ pro vektorovou funkci F. První výraz totiž dává výsledek:

$$\left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\right) \bullet (F_1, \dots, F_n) = \frac{\partial}{\partial x_1} F_1 + \dots + \frac{\partial}{\partial x_n} F_n = \frac{\partial F_1}{\partial x_1} + \dots + \frac{\partial F_n}{\partial x_n}.$$

Je to vlastně divergence.

Naopak $F \bullet \nabla$ vytváří nový diferenciální operátor, který lze následně aplikovat na nějakou funkci či vektorovou funkci (pak se aplikuje na jednotlivé složky):

$$(F_1,\ldots,F_n) \bullet \left(\frac{\partial}{\partial x_1},\ldots,\frac{\partial}{\partial x_n}\right) = F_1 \frac{\partial}{\partial x_1} + \cdots + F_n \frac{\partial}{\partial x_n}.$$

Je to tedy vlastně směrová derivace, zde ovšem i směr záleží na proměnných.

Podobnými triky lze vytvořit všechny populární operátory. Udělejme si seznam:

- $\bullet \ D_{\vec{u}} = \vec{u} \bullet \nabla,$
- $\operatorname{div}(F) = \nabla \bullet F$ pro vektorové pole F,
- $\operatorname{rot}(F) = \nabla \times F$ pro třírozměrné vektorové pole F.

Tento jazyk je docela populární. Například poslední dvě pozorování z předchozí věty se často zapisují jako $\nabla \bullet (\nabla \times F) = 0$ a $\nabla \times (\nabla f) = 0$.

U pravidla (viii) je dokonce obvyklé jej psát jako

$$\bullet \ \mathrm{rot}(F \times G) = F(\nabla \bullet G) - G(\nabla \bullet F) + (G \bullet \nabla)F - (F \bullet \nabla)G.$$

Jak výrazy napravo interpretujeme? Mějme funkce $F = (F_1, \ldots, F_n)$ a $G = (G_1, \ldots, G_n)$. Podívejme se na první výraz napravo. $\nabla \bullet G$ je skalární součin gradientu $\left(\frac{\partial}{\partial x_1}, \ldots, \frac{\partial}{\partial x_n}\right)$ s G, je to tedy výraz $\sum \frac{\partial}{\partial x_i} G_i$ neboli $\sum \frac{\partial G_i}{\partial x_i}$, ty derivace se rovnou aplikují. Vznikne skalární funkce, kterou je vynásobena vektorová funkce, první výraz tedy je

$$\left(F_1 \sum \frac{\partial G_i}{\partial x_i}, \dots, F_n \sum \frac{\partial G_i}{\partial x_i}\right) = \left(\sum F_1 \frac{\partial G_i}{\partial x_i}, \dots, \sum F_n \frac{\partial G_i}{\partial x_i}\right).$$

Protože jsme zvyklí na komutativitu, mohli bychom si myslet, že třetí výraz by měl vyjít stejně, ale tam se nenásobí. V závorce se skalárně násobí vektor G s vektorem gradientu, ale v pořadí $G \bullet \nabla$, takže se nederivuje, ale vznikne diferenciální operátor $\sum G_i \frac{\partial}{\partial x_i}$. Napravo od něj pak vidíme funkci, na kterou se aplikuje, jmenovitě na vektor, takže se aplikuje na každou složku zvlášť:

$$\left(\sum G_i \frac{\partial}{\partial x_i} F_1, \dots, \sum G_i \frac{\partial}{\partial x_i} F_n\right) = \left(\sum G_i \frac{\partial F_1}{\partial x_i}, \dots, \sum G_i \frac{\partial F_n}{\partial x_i}\right).$$

Porovnáním vidíme, že první a třetí člen napravo v (viii) nejsou stejné.

Podobně si lze rozmyslet ostatní dva výrazy.

Tuto sekci uzavřeme pohledem na problematiku inverzních funkcí. Víme, že když má dostatečně rozumná funkce f inverzi f_{-1} , tak lze její derivaci získat z derivace funkce f prostřednictvím vzorce

$$[f_{-1}]'(y) = \frac{1}{f'(x)}.$$

Lokální existenci inverzní funkce lze rozpoznat pomocí podmínky $f'(x) \neq 0$.

Přirozené zobecnění tohoto výsledku pracuje s vektorovými funkcemi mezi prostory o stejné dimenzi, například s transformacemi souřadnic. Začneme poněkud obecnějším výsledkem, který je vcelku snadné dokázat.

Věta.

Nechť $F: D \to \mathbb{R}^n$ je vektorová funkce, kde $D \subseteq \mathbb{R}^n$. Předpokládejme, že na nějakém okolí $\vec{a} \in D$ existuje inverzní funkce F^{-1} k F. Označme $\vec{b} = F(\vec{a})$. Jestliže je F diferencovatelná v \vec{a} a F^{-1} je diferencovatelná v \vec{b} , pak

$$J_{F^{-1}}(\vec{b}) = J_F(\vec{a})^{-1}$$

Součástí závěru je, že Jacobiho matice $J_F(\vec{a})$ je invertibilní. Můžeme také psát $J_{F^{-1}}(\vec{b}) = J_F(F^{-1}(\vec{b}))^{-1}$. Toto značení předpokládá, že nejprve dosadíme za \vec{x} v J_F a pak najdeme inverzní matici, ale jde to také dělat v opačném pořadí $J_F^{-1}(F^{-1}(\vec{b}))$.

Nyní si ukážeme náročnou verzi (s obtížným důkazem), kde se existence inverzní funkce odvodí z analogie nenulovosti derivace.

Věta.

Nechť $F\colon D\mapsto \mathbb{R}^n$ je vektorová funkce, kde $D\subseteq \mathbb{R}^n$, a $\vec{a}\in D$. Předpokládejme, že $F\in C^1(U)$ pro nějaké okolí U bodu \vec{a} .

Jestliže $J_F(\vec{a})$ není singulární, pak existuje okolí B bodu \vec{a} takové, že F je na B prostá funkce s obrazem C=F[B] a inverzní funkce $F^{-1}\colon C\mapsto B$ je z $C^1(C)$, přičemž

$$J_{F^{-1}} = J_F^{-1}(F^{-1})$$
 na C .

7b. Skládání funkcí, transformace

Začneme elegantním obecným výsledkem o derivaci pro složenou funkci.

Věta.

Nechť $F \colon D \mapsto \mathbb{R}^m$ je vektorová funkce, kde $D \subseteq \mathbb{R}^n$ je otevřená množina. Nechť $G \colon M \mapsto \mathbb{R}^p$ je vektorová funkce, kde $M \subseteq \mathbb{R}^m$ je otevřená množina, přičemž $F[D] \subseteq M$.

Jestliže $F \in [C^1(D)]^m$ a $G \in [C^1(M)]^p$, pak $G \circ F \in [C^1(D)]^p$ a pro $\vec{a} \in D$ platí $J_{G \circ F}(\vec{a}) = J_G(F(\vec{a}))J_F(\vec{a})$.

Stručně: $J_{G \circ F} = J_G(F)J_F$ na D.

Tento výsledek čtenáři jistě připomene známé řetízkové pravidlo g'(f)f'. Stojí za to si připomenout jeho podstatu, která je pěkně vidět na případě skládání tří funkcí.

$$\mathbb{R}[x] \xrightarrow{f} \mathbb{R}[y] \xrightarrow{g} \mathbb{R}[z] \xrightarrow{h} \mathbb{R}.$$

Složenou funkci $h \circ g \circ f$ si představujeme tak, že do funkce h(z) dosadíme z = g(y), načež za y dosadíme y = f(x), vznikne h(g(f(x))).

Tato funkce má derivaci $[h \circ g \circ f]' = h'(g(f)) \cdot g'(f) \cdot f'$, ale často píšeme jen $[h \circ g \circ f]' = h' \cdot g' \cdot f'$. Tento vzoreček dobře vystihuje podstatu pravidla: Chceme-li derivovat složenou funkci h(g(f(x))), tak prostě začneme vně a postupně krok za krokem přecházíme dovnitř k x, přičemž vše, co potkáme cestou, zderivujeme. Tyto derivace pak spojujeme násobením. Ovšem hotový výraz musí mít jedinou proměnnou, jmenovitě x, takže nemůžeme vzít jako odpověď h'(z)g'(y)f'(x), ale musíme dosadit za y a z, čímž vznikne ten přesný a správný tvar.

Jak se toto odrazí ve verzi pro vektorové funkce? Nejprve se přesvědčíme, že výraz ve větě vůbec dává smysl. Násobí se dvě Jacobiho matice běžným maticovým způsobem. Aby šlo funkce skládat v pořadí F pak G, musí být dimenze cílového prostoru F stejná jako dimenze definičního oboru G. Jinak řečeno, počet proměnných G musí odpovídat počtu složek vektorové funkce F. To znamená, že matice J_G má stejný počet sloupců, jako má matice J_F řádků, a proto jdou násobit v pořadí $J_G \cdot J_F$.

Co to násobení ale konkrétně znamená? Nejprve si ujasníme situaci a zavedeme názvy proměnných v jednotlivých prostorech, se kterými pracujeme.

$$\mathbb{R}^n[\vec{x}] \xrightarrow{F} \mathbb{R}^m[\vec{y}] \xrightarrow{G} \mathbb{R}^p.$$

Máme tedy funkci $G(\vec{y})$ a do ní dosazujeme $\vec{y} = F(\vec{x})$. Výsledkem je vektorová funkce, jejíž i tá složka je

$$G_i(F(\vec{x})) = G_i(F_1(x_1, \dots, x_n), F_2(x_1, \dots, x_n), \dots, F_m(x_1, \dots, x_n)).$$

Jacobiho matice $J_{G \circ F}$ této složené funkce vznikne tak, že na řádek s indexem i umístíme parciální derivace této složky podle všech proměnných. Na pozici (i, j) je tedy číslo

$$\frac{\partial G_i(F(\vec{x}))}{\partial x_j} = \frac{\partial}{\partial x_j} G_i(F_1(x_1, \dots, x_n), F_2(x_1, \dots, x_n), \dots, F_m(x_1, \dots, x_n)).$$

Podle vzorce z věty se k tomuto číslu umíme dostat i tak, že vezmeme matici J_G , za proměnnou \vec{y} dosadíme $F(\vec{x})$, a pak její řádek i pronásobíme oním maticovým způsobem se sloupcem j matice J_F . Ono násobení je vlastně skalární součin onoho řádku a sloupce coby vektorů, takže lze napsat

$$\frac{\partial G_i(F)}{\partial x_j} = \nabla G_i(F) \bullet \frac{\partial}{\partial x_j} F.$$

Toto elegantní vyjádření má svá použití, ale my teď chceme vědět, co to opravdu říká:

$$\frac{\partial G_i(F)}{\partial x_j} = \left(\frac{\partial G_i}{\partial y_1}(F), \dots, \frac{\partial G_i}{\partial y_m}(F)\right) \bullet \left(\frac{\partial F_1}{\partial x_j}, \dots, \frac{\partial F_m}{\partial x_j}\right)
= \frac{\partial G_i}{\partial y_1}(F) \frac{\partial F_1}{\partial x_j} + \dots + \frac{\partial G_i}{\partial y_m}(F) \frac{\partial F_m}{\partial x_j}
= \sum_{k=1}^m \frac{\partial G_i}{\partial y_k}(F) \frac{\partial F_k}{\partial x_j}.$$

Toto je klíčový vzorec. Abychom lépe ocenili jeho význam, vynecháme teď to dosazování F:

$$\frac{\partial}{\partial x_j} G_i(F_1(x_1, \dots, x_n), F_2(x_1, \dots, x_n), \dots, F_m(x_1, \dots, x_n))$$

$$= \frac{\partial G_i}{\partial y_1} \frac{\partial F_1}{\partial x_j} + \frac{\partial G_i}{\partial y_2} \frac{\partial F_2}{\partial x_j} + \dots + \frac{\partial G_i}{\partial y_m} \frac{\partial F_m}{\partial x_j}.$$

Tento vzorec zachovává podstatu řetízkového pravidla. Jsme s derivací vně a chceme se dostat k

proměnné x_j . Teď je ale novinka v tom, že se proměnná x_j nachází na více místech, například vidíme m výskytů proměnné x_1 . Vzorec nám říká, že se máme dostat ke všem výskytům proměnné x_j , což je možné skrz F_1 , skrz F_2 atd. Pokaždé máme v souladu s podstatou řetízkového pravidla zderivovat vše, co potkáme cestou, a derivace násobit; to, jakou cestu ke konkrétnímu výskytu x_j volíme, určuje, podle které proměnné máme derivovat G_i , takže například při cestě k x_j schovanému v F_2 jdeme skrz druhou proměnnou funkce G_i , takže budeme derivovat G_i podle y_2 . Součiny derivací těchto možných cest pak spojujeme sčítáním.

Jedna zajímavá verze řetízkového pravidla si pomáhá označením $y_k = F_k(\vec{x})$. Můžeme pak symbolicky psát

$$\frac{\partial G_i}{\partial x_j} = \frac{\partial G_i}{\partial y_1} \frac{\partial F_1}{\partial x_j} + \frac{\partial G_i}{\partial y_2} \frac{\partial F_2}{\partial x_j} + \dots + \frac{\partial G_i}{\partial y_m} \frac{\partial F_m}{\partial x_j}$$
$$= \frac{\partial G_i}{\partial y_1} \frac{\partial y_1}{\partial x_j} + \frac{\partial G_i}{\partial y_2} \frac{\partial y_2}{\partial x_j} + \dots + \frac{\partial G_i}{\partial y_m} \frac{\partial y_m}{\partial x_j}.$$

Příklad: Uvažujme funkce $F: \mathbb{R}^2 \to \mathbb{R}^3$ a $G: \mathbb{R}^3 \to \mathbb{R}^2$ definované takto:

$$F(x,y) = (x^2y^2, x^2 + y^2, xy)$$
$$G(u, v, w) = (uvw, u^2 + v^2 + w^2).$$

Najdeme parciální derivace pro funkci $H=G\circ F=G(F)$. Můžeme to vidět tak, že do funkce G(u,v,w) dosazujeme

$$u = x^{2}y^{2},$$

$$v = x^{2} + y^{2},$$

$$w = xy.$$

Nejprve derivace najdeme formálně podle věty. Začneme Jacobi maticemi pro funkce F a G. Jednotlivé složky jsou $F_1(x,y) = x^2y^2$, $F_2(x,y) = x^2 + y^2$, $F_3(x,y) = xy$ a $G_1(u,v,w) = uvw$, $G_2(u,v,w) = u^2 + v^2 + w^2$. Pak

$$J_F(x,y) = \begin{pmatrix} 2xy^2 & 2x^2y \\ 2x & 2y \\ y & x \end{pmatrix}, \qquad J_G(u,v,w) = \begin{pmatrix} vw & uw & uv \\ 2u & 2v & 2w \end{pmatrix}.$$

Potřebujeme ještě dosadit $u = F_1(x, y), v = F_2(x, y), w = F_3(x, y)$ do J_G , dostaneme

$$J_G(F(x,y)) = \begin{pmatrix} (x^2 + y^2)xy & x^2y^2xy & x^2y^2(x^2 + y^2) \\ 2x^2y^2 & 2(x^2 + y^2) & 2xy \end{pmatrix}.$$

Máme tedy

$$J_{H}(x,y) = J_{G}(F(x,y)) \cdot J_{F}(x,y) = \begin{pmatrix} (x^{2} + y^{2})xy & x^{3}y^{3} & x^{2}y^{2}(x^{2} + y^{2}) \\ 2x^{2}y^{2} & 2(x^{2} + y^{2}) & 2xy \end{pmatrix} \begin{pmatrix} 2xy^{2} & 2x^{2}y \\ 2x & 2y \\ y & x \end{pmatrix}$$

$$= \begin{pmatrix} (x^{2} + y^{2})2x^{2}y^{3} + 2x^{4}y^{3} + x^{2}y^{3}(x^{2} + y^{2}) & (x^{2} + y^{2})2x^{3}y^{2} + 2x^{3}y^{4} + x^{3}y^{2}(x^{2} + y^{2}) \\ 4x^{3}y^{4} + 4x(x^{2} + y^{2}) + 2xy^{2} & 4x^{4}y^{3} + 4y(x^{2} + y^{2}) + 2x^{2}y \end{pmatrix}$$

$$= \begin{pmatrix} 5x^{4}y^{3} + 3x^{2}y^{5} & 3x^{5}y^{2} + 5x^{3}y^{4} \\ 4x^{3} + 4x^{3}y^{4} + 6xy^{2} & 4x^{4}y^{3} + 6x^{2}y + 4y^{3} \end{pmatrix}.$$
The second of the first standard of the following standard of th

Tento postup je výhodný v teoretických úvahách, v praxi častěji hledáme jednotlivé parciální derivace složené funkce $H(x,y)=G(F_1(x,y),F_2(x,y),F_3(x,y))$ aplikací řetízkového pravidla. Ukážeme si $\frac{\partial H_1}{\partial x}$ neboli levý horní roh v J_H .

První složka H je $H_1(x,y) = G_1(F_1(x,y), F_2(x,y), F_3(x,y))$. Pokud chceme najít derivaci vzhledem k x, pak se k tomu x musíme zvenčí dostat. Vidíme jej na třech místech (ve třech proměnných

 G_1), takže všechny tyto cesty musíme projít: k x se dostaneme skrz F_1 , F_2 a F_3 .

$$\begin{split} \frac{\partial H_1}{\partial x} &= \frac{\partial G_1}{\partial u} \cdot \frac{\partial F_1}{\partial x} + \frac{\partial G_1}{\partial v} \cdot \frac{\partial F_2}{\partial x} + \frac{\partial G_1}{\partial w} \cdot \frac{\partial F_3}{\partial x} \\ &= vw \cdot 2xy^2 + uw \cdot 2x + uv \cdot y = (x^2 + y^2)xy2xy^2 + (x^2y^2)xy2x + x^2y^2(x^2 + y^2)y \\ &= 5x^4y^3 + 3x^2y^5. \end{split}$$

Dobře to dopadlo, dostali jsme stejný výsledek.

Nyní ještě zkusíme najít parciální derivace přímo. Nejprve zjistíme, jak vypadá H:

$$H(x,y) = G(x^2y^2, x^2 + y^2, xy) = (x^2y^2(x^2 + y^2)xy, (x^2y^2)^2 + (x^2 + y^2)^2 + (xy)^2)$$
$$= (x^5y^3 + x^3y^5, x^4 + y^4 + x^4y^4 + 3x^2y^2).$$

Nyní najdeme parciální derivace složek $H_1(x,y)=x^5y^3+x^3y^5$ a $H_2(x,y)=x^4+y^4+x^4y^4+3x^2y^2$ přímo:

$$\frac{\partial H_1}{\partial x} = 5x^4y^3 + 3x^2y^5$$

$$\frac{\partial H_1}{\partial y} = 3x^5y^2 + 5x^3y^4$$

$$\frac{\partial H_2}{\partial x} = 4x^3 + 4x^3y^4 + 6xy^2$$

$$\frac{\partial H_2}{\partial y} = 4x^4y^3 + 6x^2y + 4y^3.$$

Tím jsme potvrdili výsledek získaný vzorcem pro derivaci složené funkce. Toto bylo samozřejmě nejsnažší, ale ne vždy je to možné, například pokud některé zúčastněné funkce neznáme. Je tedy důležité umět aplikovat řetízkové pravidlo i na funkce více proměnných.

 \triangle

V aplikacích se často stává, že jeden z prostorů, které u skládání máme, je jednorozměrný. Pak se hodí zjednodušené verze pravidla pro derivaci složené funkce. Nejprve si jako ilustraci ukážeme případ, kdy m=p=1. Máme tedy funkci $f\colon \mathbb{R}^n \mapsto \mathbb{R}$ více proměnných a my ji "prodloužíme" další funkcí $g\colon \mathbb{R} \mapsto \mathbb{R}$. Po složení vznikne funkce více proměnných $\mathbb{R}^n \mapsto \mathbb{R}$. Z Jacobiho matice se pro funkci více proměnných stane gradient, pro jednoduchou funkci pak derivace. Dostáváme následující tvrzení, ve kterém jsme se pro jednoduchost omezili na hladké funkce.

Věta.

Nechť $D \subseteq \mathbb{R}^n$ je otevřená množina a $f \in C^1(D)$ je funkce. Nechť M je otevřený interval takový, že $f[D] \subseteq M$, uvažujme funkci $g \in C^1(M)$.

Pak $g \circ f \in C^1(D)$ a na D platí

$$\nabla(g \circ f) = g'(f)\nabla f.$$

Ve vzorci je napravo gradient, tedy vektor, násobený skalárem, což dává smysl. O jednotlivých parciálních derivacích nám vzorec říká toto:

$$\frac{\partial}{\partial x_i}g(f(x_1,\ldots,x_n))=g'(f(x_1,\ldots,x_n))\cdot\frac{\partial f}{\partial x_i}(x_1,\ldots,x_n).$$

Příklad: Uvažujme $f(x,y) = \sin(xy)$ a $g(u) = u^2$, zajímá nás derivace složené funkce $g \circ f$.

$$\mathbb{R}^2[x,y] \xrightarrow{u=\sin(xy)} \mathbb{R}[u] \xrightarrow{u^2} \mathbb{R}$$

Použijeme větu. g'(u) = 2u, proto

$$\frac{\partial}{\partial x}g(f(x,y)) = 2u\big|_{u=f(x,y)} \cdot \frac{\partial}{\partial x}[\sin(xy)] = 2f(x,y) \cdot \cos(xy) \cdot y = 2y\sin(xy)\cos(xy).$$

Ověříme: Máme $(g \circ f)(x,y) = [\sin(xy)]^2$, tedy opravdu derivace je

$$\frac{\partial}{\partial x}g(f(x,y)) = 2\sin(xy)\frac{\partial}{\partial x}[\sin(xy)] = 2\sin(xy)\cos(xy)y.$$

Derivace podle y se dělá obdobně.

 \triangle

Pojďme se teď podívat na dvě velmi užitečné situace, které se pojí s příběhem.

Pohyb po křivce.

Představme si následující situaci. Funkce g více proměnných popisuje třeba teplotu v různých místech učebny, má tedy tři proměnné. Vektorová funkce G může popisovat třeba proudění vzduchu v dotyčné místnosti, je to tedy funkce $\mathbb{R}^3 \mapsto \mathbb{R}^3$. My touto místností pocházíme, pohybujeme se tedy po parametrické křivce s časem jako proměnnou.

Tradiční označení naší polohy (vůči nějaké souřadnicové soustavě) by ve fyzice bylo $\vec{r}(t)$. Je to tedy funkce $\mathbb{R} \mapsto \mathbb{R}^3$. Pro parametrické křivky $\vec{r}(t) = (r_1(t), r_2(t), r_3(t))$ jsme si zavedli užitečné značení $\vec{r}'(t) = (r'_1(t), r'_2(t), r'_3(t))$. Ovšem pokud se chceme odvolat na obecnou větu, pak bychom to měli vnímat jako speciální případ Jacobiho matice, což je pro takovouto funkci sloupec; budeme tedy ve vzorci transponovat, ať se z $\vec{r}'(t)$ stane sloupcový vektor.

Složená funkce $(g \circ \vec{r})(t) = g(\vec{r}(t))$ popisuje, jakou teplotu při své procházce registrujeme. Je to vlastně obyčejná funkce $\mathbb{R} \mapsto \mathbb{R}$. Pokud ji zderivujeme, zjistíme, jak rychle se při našem pohybu teplota mění subjektivně, z našeho pohledu. To jistě nějak souvisí s tím, jak se mění teplota v místnosti jako takové, tedy s prostorovým gradientem funkce g.

Složená funkce $(G \circ \vec{r})(t) = G(\vec{r}(t))$ popisuje, jaký vítr cítíme při své procházce. Je to vektorová funkce $\mathbb{R} \mapsto \mathbb{R}^3$, tedy formálně parametrická křivka, budeme hledat její derivaci $[G(\vec{r})]'$, kterou budeme v obecném vzorci transponovat na sloupcový vektor, aby zapadl do obecného rámce.

Souvislosti prostorových derivací (Jacobiho matice) s derivacemi, které se vážou na naši cestu místností, nabídne speciální verze věty o derivaci složené funkce.

Věta.

Nechť D je otevřený interval v \mathbb{R} a $\vec{r} \in [C^1(D)]^m$ je vektorová funkce.

Nechť M je otevřená množina v \mathbb{R}^m taková, že $\vec{r}[D] \subseteq M$.

(i) Pro funkci $g\in C^1(M)$ pak $g(\vec{r}\,)\in C^1(D)$ a platí

$$[g(\vec{r})]' = \nabla g(\vec{r}) \cdot \vec{r}'^T.$$

(i) Pro vektorovou funkci $G \in [C^1(M)]^p$ pak $G(\vec{r}\,) \in [C^1(D)]^p$ a platí

$$[G(\vec{r})]^{\prime T} = J_G(\vec{r}) \cdot \vec{r}^{\prime T}.$$

V prvním vzorci napravo máme

$$\nabla g \cdot \vec{r}'^T = \left(\frac{\partial g}{\partial x_1}, \dots, \frac{\partial g}{\partial x_m}\right) \cdot \left(r'_1, \dots, r'_m\right)^T = \left(\frac{\partial g}{\partial x_1}, \dots, \frac{\partial g}{\partial x_m}\right) \bullet \left(r'_1, \dots, r'_m\right)$$
$$= \frac{\partial g}{\partial x_1} r'_1 + \dots + \frac{\partial g}{\partial x_m} r'_m.$$

Protože zkoušíme derivovat funkci $g(r_1(t), \ldots, r_m(t))$ podle t, odpovídá to intuici řetízkového pravidla, že máme vyzkoušet všechny možnosti, jak se dostat k t (tyto možnosti sčítáme), pokaždé derivujeme vše, co cestou potkáme, a násobíme ty derivace. V tomto vzorci do bylo pěkně vidět, ale správně bychom měli ještě kvůli správné proměnné dosadit do parciálních derivací:

$$[g(\vec{r})]'(t) = \nabla g(\vec{r}(t)) \cdot \vec{r}'(t)^T = \sum_{k=1}^m \frac{\partial g}{\partial x_k} (\vec{r}(t)) r_k'(t).$$

Tento vzorec lze přepsat ještě jinak:

$$[g(\vec{r})]'(t) = D_{\vec{r}'(t)}g(\vec{r}(t)).$$

To má přirozenou interpretaci. Pohybujeme se po jisté křivce. V čase t jsme v místě $\vec{r}(t)$ a chceme vědět, jak rychlou změnu pozorujeme u funkce g. Vzorec nám říká, ať se podíváme na směrovou derivaci g ve směru $\vec{r}'(t)$, tedy ve směru ve kterém se zrovna pohybujeme.

U vektorové verze máme funkci $G(\vec{r})$: $\mathbb{R} \mapsto \mathbb{R}^p$, jmenovitě funkci $(G_1(\vec{r}(t)), \dots, G_p(\vec{r}(t)))$. Hledáme tedy sloupcovou Jacobiho matici $p \times 1$ se složkami $[G_j(\vec{r})]'$. Na pravé straně vzorce vidíme, jak se $p \times m$ matice násobí sloupcovým vektorem $m \times 1$, což rozměrově souhlasí. Maticové násobení skrývá vzorce

$$[G_j(r_1(t),\ldots,r_m(t))]' = \sum_{k=1}^m \frac{\partial G_j}{\partial x_k}(\vec{r}(t))r'_k(t),$$

což opět odpovídá základní filosofii řetízkového pravidla. Je to vlastně stejný vzorec jako pro g výše, což dává smysl, protože prostě aplikujeme derivaci na jednotlivé složky vektorové funkce $G = (G_1, \ldots, G_m)$.

Transformace souřadnic

Zde uvažujeme případ n=m a p=1. Máme tedy vektorovou funkci $F: \mathbb{R}^n \to \mathbb{R}^n$ a na ni navazující funkci $g: \mathbb{R}^n \to \mathbb{R}$. Věta o derivaci složené funkce pak má tuto podobu:

Věta.

Nechť D je otevřená množina v \mathbb{R}^n a $F \in [C^1(D)]^n$ je vektorová funkce. Nechť M je otevřená množina v \mathbb{R}^n taková, že $F[D] \subseteq M$, uvažujme funkci $g \in C^1(M)$. Pak $g(F) \in C^1(D)$ a platí

$$\nabla(g(F)) = \nabla g(F)J_F.$$

Jako obvykle pro praktické počítání potřebujeme rozklíčovat tento kompaktní výsledek. Začneme tím, že zavedeme názvy pro proměnné v jednotlivých prostorech vyskytujících se v této větě.

$$\mathbb{R}^n[\vec{x}] \xrightarrow{F} \mathbb{R}^n[\vec{y}] \xrightarrow{g} \mathbb{R} \qquad \mathbb{R}^n[\vec{x}] \xrightarrow{g(F)} \mathbb{R}.$$

Jinak řečeno, máme funkci $g(\vec{y})$, do které dosazujeme $\vec{y} = F(\vec{x})$ a dostaneme $g(F)(\vec{x})$.

Nalevo ve vzorci teď vidíme vektor, jehož složky jsou $\frac{\partial g(F)}{\partial x_i}$. Napravo násobíme (řádkový) vektor

se složkami $\frac{\partial g}{\partial y_i}$ maticí se složkami $\frac{\partial F_i}{\partial y_j}$. Když v tomto násobení spočítáme jednu složku, dostáváme vzorec

$$\frac{\partial g(F)}{\partial x_i} = \sum_{j=1}^n \frac{\partial g}{\partial y_j} \frac{\partial F_j}{\partial x_i}.$$

Jako obvykle je to v souladu se základní myšlenkou. Když chceme derivovat funkci

$$g(y_1, y_2, \dots, y_n) = g(F_1(x_1, \dots, x_n), F_2(x_1, \dots, x_n), \dots, F_n(x_1, \dots, x_n))$$

podle x_i , tak se musíme dostat ke všem výskytům x_i ve vzorci a cestou zderivujeme všechno, co potkáme:

$$\frac{\partial g}{\partial y_1} \frac{\partial F_1}{\partial x_i} + \frac{\partial g}{\partial y_2} \frac{\partial F_2}{\partial x_i} + \dots + \frac{\partial g}{\partial y_n} \frac{\partial F_n}{\partial x_i}.$$

Teď si k tomu postavíme praktickou interpretaci. Představme si, že máme funkci $g(\vec{y}): \mathbb{R}^n \to \mathbb{R}$, která popisuje nějakou fyzikální veličinu (třeba lokální teplotu), přičemž pozice se udává vzhledem k nějaké souřadnicové soustavě \vec{y} , můžeme si představit třeba kartézské souřadnice. My ale zjistíme, že je pro nás výhodnější použít nějaký jiný souřadnicový systém se souřadnicemi \vec{x} . Vzniká tak nová funkce $\hat{g}(\vec{x})$, která popisuje sice stejnou kvantitu (stejný jev), ale vůči jiným souřadnicím, tedy bude to téměř jistě formálně jiná funkce (jiný vzorec).

Protože funkce $g(\vec{y})$ a $\hat{g}(\vec{x})$ popisují stejný jev, dá se čekat, že mezi nimi existuje nějaký vztah. Ten je dán převodem mezi souřadnicemi. Předpokládejme, že existuje vektorová funkce $F: \mathbb{R}^n \to \mathbb{R}^n$, která převádí souřadnice \vec{x} na souřadnice \vec{y} , formálně $\vec{y} = F(\vec{x})$. Pokud sedíme v bodě s novými souřadnicemi \vec{x} , tak hodnotu dotyčné fyzikální veličiny můžeme získat dosazením do \hat{g} . Alternativní možnost je nejprve zjistit, že vzhledem ke starým souřadnicím máme polohu $\vec{y} = F(\vec{x})$, a to dosadit do g. Protože se ptáme na tutéž hodnotu na stejném místě, musí platit

$$\hat{q}(\vec{x}) = q(F(\vec{x})).$$

Zdá se to zjevné, prostě do g dosadíme za \vec{y} , ale je dobré vědět, že to dává i smysl v aplikacích. Protože je F transformací souřadnic, tak bychom čekali, že je invertibilní, tedy máme k dispozici i funkci $\vec{x} = F_{-1}(\vec{y})$. Pak také $g(\vec{y}) = \hat{g}(F_{-1}(\vec{y}))$.

Příklad: Uvažujme funkci $g(x,y) = x^2y$. Vyjadřují, jak hodnoty jistého jevu (třeba teplotu) vidí pozorovatel, který svou pozici měří v kartézských souřadnicích. Druhý pozorovatel vidí stejný jev, ale používá polární souřadnice, tedy vidí závislost $\hat{g}(r,\varphi)$.

Předpokládejme, že oba pozorovatelé používají stejný počátek a osu x. Pak mají k dispozici pohodlný převod souřadnic. Od polárních souřadnic se k těm kartézským dostaneme vzorci

$$x = r\cos(\varphi),$$

$$y = r\sin(\varphi).$$

Zde se bere $r \geq 0$ a $0 \leq \varphi < 2\pi$. Formálně vidíme tuto transformaci coby vektorovou funkci

$$F(r,\varphi) = (r\cos(\varphi), r\sin(\varphi)).$$

Druhý pozorovatel tak vidí, že jev závisí na jeho pozici dle vzorce

$$\hat{g}(r,\varphi) = g(r\cos(\varphi), r\sin(\varphi)) = r^3\cos^2(\varphi)\sin(\varphi).$$

Vidíme, že je to opravdu jiná funkce než ta, kterou pozoruje první pozorovatel. Funguje to ale tak, že když oba stojí na stejném místě a dosadí své souřadnice do své funkce, tak dostanou stejnou hodnotu.

Polární souřadnice jsou populární, ačkoliv nejsou zcela příjemné. Jsou dva problémy. Za prvé, je třeba rozhodnout, jak kódovat počátek, jinak by F nebylo prosté. Zatímco totiž pro počátek máme evidentně r=0, s úhlem je to horší, všechny by fungovaly. Je tedy nutno vybrat.

Když to uděláme, tak kF bude existovat inverzní zobrazení F_{-1} , které dává převod od kartézských souřadnic k polárním. Tam ovšem narazíme na druhou nepříjemnou komplikaci: pro tento převod nemáme algebraický vzorec pro φ .

 \triangle

V praxi, zejména ve fyzice, si zjednodušují život tím, že nepoužívají přesné značení. Protože funkce g a \hat{g} popisují stejný jev, značí je obě jako g a rozlišují je označením proměnné. Psali by tedy $g(\vec{y})$ a také $g(\vec{x})$ namísto $\hat{g}(\vec{x})$, ačkoliv jde vzorečkově o dvě různé funkce. Další rozšířený zvyk je nepoužívat pro transformaci jméno F, ale nazývat ji jménem cílové proměnné. Fyzici by tedy psali $\vec{y} = \vec{y}(\vec{x})$, popřípadě $\vec{x} = \vec{x}(\vec{y})$ pro inverzní transformaci. Často se ani nepoužívá aparát vektorové funkce, ale pracuje se s transformačními funkcemi pro jednotlivé proměnné, tedy psali by $y_j = y_j(\vec{x})$ nebo $y_j = y_j(x_1, \dots, x_n)$.

Příklad: Vrátíme-li se k předchozímu příkladu, tak praktici by viděli dvě verze funkce g:

$$g(x, y) = x^{2}y,$$

 $g(r, \varphi) = r^{3} \cos^{2}(\varphi) \sin(\varphi).$

Transformaci by pak značili takto:

$$x = x(r, \varphi) = r \cos(\varphi),$$

 $y = y(r, \varphi) = r \sin(\varphi).$

 \triangle

V praktickém značení lze pravidlo pro derivaci zachytit takto:

$$\frac{\partial g(\vec{x})}{\partial x_i} = \sum_{k=1}^n \frac{\partial g(\vec{y})}{\partial y_k} \frac{\partial y_k}{\partial x_i}.$$

Příklad: Uvažujme opět funkci $g(x,y)=x^2y$ a převod od polárních na kartézské souřadnice. Pokud chceme znát parciální derivace funkce g vůči polárním souřadnicím, můžeme prostě zderivovat funkci $g(r,\varphi)=r^3\cos^2(\varphi)\sin(\varphi)$:

$$\begin{split} \frac{\partial g}{\partial r} &= 3r^2 \cos^2(\varphi) \sin(\varphi), \\ \frac{\partial g}{\partial \varphi} &= 2r^3 \cos(\varphi) (-\sin(\varphi)) \sin(\varphi) + r^3 \cos^2(\varphi) \cos(\varphi) \\ &= r^3 \cos(\varphi) [\cos^2(\varphi) - 2\sin^2(\varphi)]. \end{split}$$

Jak to bude fungovat podle věty?

$$\frac{\partial g}{\partial r}(r,\varphi) = \frac{\partial g}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial r} = 2xy\cos(\varphi) + x^2\sin(\varphi)
= 2r\cos(\varphi)r\sin(\varphi)\cos(\varphi) + (r\cos(\varphi))^2\sin(\varphi) = 3r^2\cos^2(\varphi)\sin(\varphi),
\frac{\partial g}{\partial \varphi}(r,\varphi) = \frac{\partial g}{\partial x}\frac{\partial y}{\partial \varphi} + \frac{\partial g}{\partial y}\frac{\partial y}{\partial \varphi} = -2xyr\sin(\varphi) + x^2r\cos(\varphi)
= -2r\cos(\varphi)r\sin(\varphi)r\sin(\varphi) + (r\cos(\varphi))^2r\cos(\varphi) = -2r^3\cos(\varphi)\sin^2(\varphi) + r^3\cos^3(\varphi).$$

Všimněte si, že derivace $\frac{\partial g}{\partial x}$, $\frac{\partial g}{\partial y}$ pracují s proměnnými x, y, takže aby měl výsledek smysl (chceme derivaci vzhledem k r či φ), museli jsme ještě dosadit.

Tento druhý výpočet byl delší, nicméně jsou situace, kdy to jinak nejde. Představme si například, že primární informace jsou ty od polárního pozorovatele, tedy máme k dispozici vzorec $g(r,\varphi) = r^3 \cos^2(\varphi) \sin(\varphi)$ a také informace o tom, jak rychle se g mění, tedy známe gradient čili parciální derivace vůči r a φ . Tuto informaci by rád získal i kartézský pozorovatel.

Přirozený přístup by bylo vzít inverzní transformační vzorce r = r(x, y) a $\varphi = \varphi(x, y)$ a následovat předchozí výpočet. Bohužel, my sice máme $r = \sqrt{x^2 + y^2}$, ale neexistuje diferencovatelný vzorec pro φ v závislosti na x, y. Je tedy třeba zkusit něco jiného.

Nejprve se vrátíme k předchozímu procesu, kdy jsme přecházeli od kartézských na polární souřadnice a hledali souvislost mezi derivacemi.

$$\frac{\partial g}{\partial r} = \frac{\partial g}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial g}{\partial y} \frac{\partial y}{\partial r}$$
$$\frac{\partial g}{\partial \varphi} = \frac{\partial g}{\partial x} \frac{\partial y}{\partial \varphi} + \frac{\partial g}{\partial y} \frac{\partial y}{\partial \varphi}.$$

Dosadíme známé údaje:

$$3r^{2}\cos^{2}(\varphi)\sin(\varphi) = \frac{\partial g}{\partial x}\cdot\cos(\varphi) + \frac{\partial g}{\partial y}\cdot\sin(\varphi)$$
$$r^{3}\cos(\varphi)[\cos^{2}(\varphi) - 2\sin^{2}(\varphi)] = -\frac{\partial g}{\partial x}\cdot r\sin(\varphi) + \frac{\partial g}{\partial y}\cdot r\cos(\varphi).$$

Vznikly dvě lineární rovnice, ve kterých máme hledané parciální derivace jako dvě neznámé, takže by neměl být problém je najít (alespoň teoreticky). V tomto případě vypadá nadějně jinak nepříliš populární Kramerovo pravidlo, které přímým výpočtem poskytne

$$\frac{\partial g}{\partial x} = \frac{2r^3 \cos(\varphi) \sin(\varphi)}{r} = 2r^2 \cos(\varphi) \sin(\varphi),$$
$$\frac{\partial g}{\partial y} = \frac{r^3 \cos^2(\varphi)}{r} = r^2 \cos^2(\varphi).$$

Všimněte si, že nám to úplně nesedí. Našli jsme parciální derivace vzhledem kx a y, ale ve vzorcích vystupují ty druhé proměnné. Zajímavé je, že za chvíli uvidíme aplikaci, kde nám to přesně takto vyhovuje, takže tento postup má něco do sebe. Pro kartézského pozorovatele je to nicméně nepříjemné, protože ten vidí svět očima proměnných x a y. Dá se s tím něco dělat?

Pokud máme k dispozici převodní vzorce z kartézských souřadnic na polární, tedy r = r(x, y) a $\varphi = \varphi(x, y)$, tak stačí dosadit do výrazů napravo a dostaneme ty správné vzorce pro parciální derivace. Zde ale ty převodní vzorce nemáme, ostatně proto jsme volili tento pracnější postup, ale zato máme štěstí (což je také velmi důležité). Výsledky lze totiž upravit tak, že v nich kartézské souřadnice poznáme.

$$\frac{\partial g}{\partial x} = 2r\cos(\varphi)r\sin(\varphi) = 2xy,$$
$$\frac{\partial g}{\partial y} = [r\cos(\varphi)]^2 = x^2.$$

Kartézský pozorovatel je tedy nakonec spokojen.

 \triangle

Shrňme si situaci. Máme funkci dvou proměnných, která je zadána vůči souřadnicím x, y a také vůči souřadnicím u, v. Pro převod souřadnic máme vzorce x = x(u, v), y = y(u, v). Pak máme vzorce

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u}$$
$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial u} \frac{\partial y}{\partial v}$$

Tyto vzorce je možné použít přímo, pokud chceme najít parciální derivace vzhledem ku a v. Má smysl pak očekávat odpovědi v řeči proměnných u,v, což znamená, že do $\frac{\partial f}{\partial x}$ a $\frac{\partial f}{\partial y}$ ještě musíme dosadit x=x(u,v) a y=y(u,v), což u řetízkového pravidla děláme standardně.

Druhá možnost je využít tyto rovnice jako soustavu a jejím vyřešením získat parciální derivace vzhledem k x a y. Ty by zase měly být v řeči proměnných x, y, což může být problém, pokud se ve výrazech neumíme zbavit nechtěných proměnných dosazením r = r(x, y) a $\varphi = \varphi(x, y)$.

Vše funguje analogicky i pro více proměnných.

Příklad: Na obdélníkovém hřišti dva správci měří hustotu trávníku f. Dohodli se na společném počátku souřadnic v západním rohu, oba hodlají používat kartézské souřadnice.

Jeden k určení pozice používá GPS, takže orientuje osy na východ a sever a měří v metrech. Vznikla tak souřadnicová soustava [xy], vyznačená na obrázku černou čtvercovou sítí. Pozoruje pak funkci f(x,y).

Ten druhý vzdálenosti krokuje, a to vzhledem ke stranám hrací plochy. Délka kroku definuje jednotkové směrové vektory \vec{u} a \vec{v} , jejich souřadnice vzhledem k soustavě [xy] jsou $\vec{u}=(1,-1)$ a $\vec{v}=(1,1)$; ten tedy má dlouhé nohy. Vzniká tak souřadnicová soustava [uv], vyznačená modrou čtvercovou sítí, která je diagonální vůči soustavě [xy].

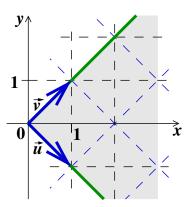
Když [uv]-pozorovatel stojí v místě, na které se dostal u kroky ve směru \vec{u} a v kroky ve směru \vec{v} , tak má vzhledem ke své souřadné soustavě souřadnice (u, v). Ovšem [xy]-pozorovatel jej vidí v lokaci

$$u\vec{u} + v\vec{v} = u(1, -1) + v(1, 1) = (u + v, -u + v).$$

Máme tedy transformaci

$$x = u + v,$$

$$y = -u + v.$$



Ta umožňuje následující přímý přepočet parciálních derivací:

$$\begin{split} \frac{\partial f}{\partial u} &= \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u} = \frac{\partial f}{\partial x} \cdot 1 + \frac{\partial f}{\partial y} \cdot (-1), \\ \frac{\partial f}{\partial v} &= \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v} = \frac{\partial f}{\partial x} \cdot 1 + \frac{\partial f}{\partial y} \cdot 1. \end{split}$$

Takže

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x} - \frac{\partial f}{\partial y},$$
$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y},$$

přesněji

$$\frac{\partial f}{\partial u}(u, v) = \frac{\partial f}{\partial x}(x, y) - \frac{\partial f}{\partial y}(x, y),$$
$$\frac{\partial f}{\partial v}(u, v) = \frac{\partial f}{\partial x}(x, y) + \frac{\partial f}{\partial y}(x, y).$$

Pak napravo dosadíme x = u + v, y = v - u.

Pokud by nás zajímal opačný převod, mohli bychom tuto soustavu vyřešit pro parciální derivace napravo, což se snadno udělá například sečtením a odečtením rovnic:

$$\frac{\partial f}{\partial x} = \frac{1}{2} \frac{\partial f}{\partial u} + \frac{1}{2} \frac{\partial f}{\partial v},$$
$$\frac{\partial f}{\partial y} = -\frac{1}{2} \frac{\partial f}{\partial u} + \frac{1}{2} \frac{\partial f}{\partial v}.$$

Také bychom mohli najít inverzní transformaci souřadnic tím, že původní transformační rovnice vyřešíme pro u, v:

$$u = \frac{1}{2}x - \frac{1}{2}y,$$

$$v = \frac{1}{2}x + \frac{1}{2}y.$$

Teď lze ty druhé převodní vzorce získat přímou aplikací řetízkového pravidla:

$$\begin{split} \frac{\partial f}{\partial x} &= \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x} = \frac{\partial f}{\partial u} \cdot \frac{1}{2} + \frac{\partial f}{\partial v} \cdot \frac{1}{2}, \\ \frac{\partial f}{\partial y} &= \frac{\partial f}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial y} = \frac{\partial f}{\partial u} \cdot \left(-\frac{1}{2}\right) + \frac{\partial f}{\partial v} \cdot \frac{1}{2}. \end{split}$$

 \triangle

Transformace diferenciálních výrazů

Fyzikální zákony jsou typicky vyjádřeny diferenciálními rovnicemi, což jsou rovnice, ve kterých se vyskytují výrazy s derivacemi. Fyzika není jediná, s diferenciálními rovnicemi pracují i další přírodní vědy nebo třeba inženýři. Standardně se přírodní zákonitosti formulují vzhledem ke kartézským souřadnicím. Někdy ale potřebujeme zjistit, jak dotyčný zákon vypadá, když na svět nahlížíme pohledem jiných souřadnic. V takovém případě musíme převést zúčastněné diferenciální výrazy do nových souřadnic.

Příklad: Máme funkci f(x,y) v kartézských souřadnicích, pro kterou má platit diferenciální rovnice $\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} = 0$.

Přejdeme k polárním souřadnicím $x = r\cos(\varphi)$, $y = r\sin(\varphi)$. Po transformaci vznikne nová funkce $(r, \varphi) \mapsto f(r\cos(\varphi), r\sin(\varphi))$ a my chceme vědět, co pro ni znamená ta zadaná rovnice.

Potřebujeme v dotyčné rovnici nahradit prostorové derivace $\frac{\partial f}{\partial x}$ a $\frac{\partial f}{\partial y}$ výrazy, ve kterých budou parciální derivace vzhledem k r a φ . To by se snadno udělalo, pokud bychom měli povědomé vzorce

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r} \frac{\partial r}{\partial x} + \frac{\partial f}{\partial \varphi} \frac{\partial \varphi}{\partial x}$$
$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial r} \frac{\partial r}{\partial y} + \frac{\partial f}{\partial \varphi} \frac{\partial \varphi}{\partial y}$$

Bohužel nemáme potřebný vzorec $\varphi = \varphi(x, y)$, takže musíme použít přístup jako v předchozím příkladě a začít rovnicemi pro opačnou transformaci:

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial r} = \frac{\partial f}{\partial x} \cdot \cos(\varphi) + \frac{\partial f}{\partial y} \cdot \sin(\varphi),$$

$$\frac{\partial f}{\partial \varphi} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial \varphi} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial \varphi} = -\frac{\partial f}{\partial x} \cdot r \sin(\varphi) + \frac{\partial f}{\partial y} \cdot r \cos(\varphi).$$

Stručným zápisem: $f_r = \cos(\varphi) f_x + \sin(\varphi) f_y$ a $f_{\varphi} = -r \sin(\varphi) f_x + r \cos(\varphi) f_y$. Když tyto rovnice vyřešíme pro f_x a f_y , dostaneme

$$f_x = f_r \cos(\varphi) - \frac{1}{r} f_\varphi \sin(\varphi),$$

$$f_y = f_r \sin(\varphi) + \frac{1}{r} f_\varphi \cos(\varphi).$$

Můžeme je dosadit do dané rovnice a dostáváme

$$f_r \cdot (\cos(\varphi) + \sin(\varphi)) + \frac{1}{r} f_{\varphi} \cdot (\cos(\varphi) - \sin(\varphi)) = 0$$

neboli

$$\frac{\partial f}{\partial r} \cdot (\cos(\varphi) + \sin(\varphi)) + \frac{1}{r} \frac{\partial f}{\partial \varphi} \cdot (\cos(\varphi) - \sin(\varphi)) = 0.$$

Takto tedy zkoumaný (imaginární) přírodní zákon vypadá v polárních souřadnicích.

Všimněte si, že jsme ve vzorcích pro f_x a f_y měli výrazy závisející na nových proměnných r, φ , ale nevadilo nám to, protože cílem tentokráte nebylo najít informaci z pohledu x, y, ale naopak z pohledu r, φ .

 \triangle

Rovnici v příkladě výše jsme si vycucali z prstu, ale postup je užitečný. U řady skutečných rovnic o přírodě je transformace do vhodné soustavy v zásadě jediným rozumným způsobem, jak je řešit. Shrňme si postup:

Je dán výraz s derivacemi nějaké funkce f vzhledem k proměnným x, y. Chceme jej transformovat do řeči proměnných u, v. Uděláme to tak, že odvodíme rovnice ve tvaru

$$\frac{\partial f}{\partial x} = a_{1,1}(u,v)\frac{\partial f}{\partial u} + a_{1,2}(u,v)\frac{\partial f}{\partial v},$$
$$\frac{\partial f}{\partial u} = a_{2,1}(u,v)\frac{\partial f}{\partial u} + a_{2,2}(u,v)\frac{\partial f}{\partial v}.$$

Pokud máme k dispozici převodní vzorce u = u(x, y), v = v(x, y), pak tyto rovnice dostaneme přímou aplikací řetízkového pravidla:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial x},$$
$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial u} \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \frac{\partial v}{\partial y},$$

přičemž na pravé straně také výskyty proměnných x,y nahradíme pomocí transformačních vzorců tak, aby se vyskytovaly jen proměnné u,v.

Pokud tyto převodní vzorce nemáme, ale máme transformační vzorce $x=x(u,v),\,y=y(u,v),$

pak potřebné vzorce získáme pomocí soustavy rovnic

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u},$$
$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial v} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial v}.$$

Vyřešíme pro parciální derivace vzhledem k x, y a ty pak dosadíme do dané rovnice.

Pokud se ve výrazu vyskytují vyšší derivace, tak opakovaně aplikujeme derivaci na rovnice z řetízkového pravidla. Pak je ovšem třeba pečlivě sledovat, které členy používají původní proměnné x,y a které už pracují s novými proměnnými u,v.

Příklad: Uvažujme tzv. Laplaceovu rovnici

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0.$$

Popisuje například rovnováhu sil na bublině nebo rozložení tepla na plotně v ustáleném stavu. Jak vypadá v polárních souřadnicích?

Už víme, že přímý výpočet potřebných vztahů není možný, začneme tedy vztahem v opačném směru. Pro první derivace již máme

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}(x, y) \cdot \cos(\varphi) + \frac{\partial f}{\partial y}(x, y) \cdot \sin(\varphi),$$

$$\frac{\partial f}{\partial \varphi} = -\frac{\partial f}{\partial x}(x, y) \cdot r \sin(\varphi) + \frac{\partial f}{\partial y}(x, y) \cdot r \cos(\varphi).$$

Připomněli jsme si, že prostorové derivace mají jako proměnné původní souřadnice x,y. To je velmi důležité, protože my teď tyto rovnice budeme ještě jednou derivovat, a to podle r a φ , takže bude třeba vzít vúvahu, že tam vlastně máme $\frac{\partial f}{\partial x} \big(x(r,\varphi), y(r,\varphi) \big)$ a $\frac{\partial f}{\partial y} \big(x(r,\varphi), y(r,\varphi) \big)$. Proto budeme při derivování muset použít řetízkové pravidlo. Naštěstí nemusíme počítat všechny čtyři derivace druhého řádu, stačí jen dvě.

$$\begin{split} \frac{\partial^2 f}{\partial r^2} &= \left(\frac{\partial^2 f}{\partial x^2}(x,y) \cdot \frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y \partial x}(x,y) \cdot \frac{\partial y}{\partial r}\right) \cos(\varphi) \\ &\quad + \left(\frac{\partial^2 f}{\partial x \partial y}(x,y) \cdot \frac{\partial x}{\partial r} + \frac{\partial^2 f}{\partial y^2}(x,y) \cdot \frac{\partial y}{\partial r}\right) \cdot \sin(\varphi) \\ &= \frac{\partial^2 f}{\partial x^2}(x,y) \cdot \cos^2(\varphi) + \frac{\partial^2 f}{\partial y \partial x}(x,y) \cdot \sin(\varphi) \cos(\varphi) \\ &\quad + \frac{\partial^2 f}{\partial x \partial y}(x,y) \cdot \cos(\varphi) \sin(\varphi) + \frac{\partial^2 f}{\partial y^2}(x,y) \cdot \sin^2(\varphi), \\ \frac{\partial^2 f}{\partial \varphi^2} &= -\left(\frac{\partial^2 f}{\partial x^2}(x,y) \frac{\partial x}{\partial \varphi} + \frac{\partial^2 f}{\partial y \partial x}(x,y) \frac{\partial y}{\partial \varphi}\right) \cdot r \sin(\varphi) - \frac{\partial f}{\partial x}(x,y) \cdot r \cos(\varphi) \\ &\quad + \left(\frac{\partial^2 f}{\partial x \partial y}(x,y) \frac{\partial x}{\partial \varphi} + \frac{\partial^2 f}{\partial y^2}(x,y) \frac{\partial y}{\partial \varphi}\right) \cdot r \cos(\varphi) - \frac{\partial f}{\partial y}(x,y) \cdot r \sin(\varphi) \\ &= \frac{\partial^2 f}{\partial x^2}(x,y) r^2 \sin^2(\varphi) - \frac{\partial^2 f}{\partial y \partial x}(x,y) r^2 \cos(\varphi) \sin(\varphi) - \frac{\partial f}{\partial x}(x,y) \cdot r \cos(\varphi) \\ &\quad - \frac{\partial^2 f}{\partial x \partial y}(x,y) r^2 \cos(\varphi) \sin(\varphi) + \frac{\partial^2 f}{\partial y^2}(x,y) r^2 \cos^2(\varphi) - \frac{\partial f}{\partial y}(x,y) \cdot r \sin(\varphi). \end{split}$$

Aby se nám s tím lépe manipulovalo, přepíšeme si to pomocí alternativního značení.

$$f_{rr} = f_{xx}\cos^{2}(\varphi) + f_{xy}\cos(\varphi)\sin(\varphi) + f_{yx}\cos(\varphi)\sin(\varphi) + f_{yy}\sin^{2}(\varphi),$$

$$f_{\varphi\varphi} = f_{xx}r^{2}\sin^{2}(\varphi) - f_{xy}r^{2}\cos(\varphi)\sin(\varphi) - f_{yx}r^{2}\cos(\varphi)\sin(\varphi) + f_{yy}r^{2}\cos^{2}(\varphi)$$

$$- r(f_{x}\cos(\varphi) + f_{y}\sin(\varphi)).$$

Rovnice si ještě připravíme. Tu první vynásobíme členem r^2 , u té druhé si všimneme, že se výraz v závorce shoduje se vzorcem odvozeným před chvílí pro $\frac{\partial f}{\partial r}$, můžeme tedy nahradit. Takto připravené rovnice pak sečteme.

$$r^{2}f_{rr} + f_{\varphi\varphi} = r^{2}f_{xx}[\cos^{2}(\varphi) + \sin^{2}(\varphi)] + r^{2}f_{yy}[\cos^{2}(\varphi) + \sin^{2}(\varphi)] - rf_{r}$$

$$\implies r^{2}f_{rr} + f_{\varphi\varphi} = r^{2}f_{xx} + r^{2}f_{yy} - rf_{r}.$$

Dostáváme

$$r^2 f_{xx} + r^2 f_{yy} = r^2 f_{rr} + f_{\varphi\varphi} + r f_r$$

neboli

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

Toto je kýžený přepis Laplaceova operátoru do polárních souřadnic. Laplaceova rovnice tedy v polárních souřadnicích zní

$$\frac{\partial^2 f}{\partial r^2} + \frac{1}{r} \frac{\partial f}{\partial r} + \frac{1}{r^2} \frac{\partial^2 f}{\partial \varphi^2} = 0.$$

 \triangle

7c. Diferenciál

V kapitole 3 jsme poznamenali, že směrové (a tedy i parciální) derivace nejsou tím správným zobecněním derivace. My potřebujeme, aby z diferencovatelnosti již vyplynula spojitost. Pokud chceme najít nějaký pojem pro funkce více proměnných, který takto funguje, tak se musíme podívat na obyčejnou derivaci jinak. Obvyklý pohled, tedy rychlost změny funkce, totiž zobecnit rozumně nejde.

Ten správný pohled získáme, když se vrátíme k otázce aproximace. Pro diferencovatelnou funkci máme velice dobrou lineární aproximaci tečnou

$$f(a+h) \approx f(a) + f'(a)h$$
.

Tuto aproximaci si přepíšeme jako

$$f(a+h) - f(a) \approx f'(a)h$$
.

Výraz napravo definuje lineární zobrazení $h \mapsto f'(a)h$. Když hledáme tečnu, tak vlastně hledáme lineární zobrazení L[h], které nejlépe aproximuje f(a+h)-f(a), jmenovitě chceme, aby L[h] vystihlo lineární složku v rozdílu f(a+h)-f(a). Jak to poznáme?

Od toho rozdílu odečteme L[h] a co zbyde, rozvineme v řadu,

$$f(a+h) - f(a) - L[h] = c_0 + c_1h + c_2h^2 + c_3h^2 + \cdots$$

Dosazením h=0 a připomenutím, že L[0]=0 pro všechna lineární zobrazení L, dostaneme $c_0=f(a)-f(a)-L[0]=0$. Doufáme, že také $c_1=0$, tedy že po odebrání L[h] z f(a+h)-f(a) už nezůstala žádná lineární složka. Vyjádříme c_1 jako

$$c_1 = \frac{f(a+h) - f(a) - L[h]}{h} - c_2 h - c_3 h^2 - \cdots$$

Abychom se zbavili toho rozvoje, přejdeme k limitě $h \to 0$:

$$c_1 = \lim_{h \to 0} (c_1) = \lim_{h \to 0} \left(\frac{f(a+h) - f(a) - L[h]}{h} \right) - \lim_{h \to 0} (c_2 h + c_3 h^2 + \cdots)$$
$$= \lim_{h \to 0} \left(\frac{f(a+h) - f(a) - L[h]}{h} \right) - 0 = \lim_{h \to 0} \left(\frac{f(a+h) - f(a) - L[h]}{h} \right).$$

Takže zobrazení L[h] vyčerpalo lineární složku rozdílu f(a+h) - f(a) právě tehdy, když je ta poslední limita nulová. Takovéto lineární zobrazení je speciální a dáme mu jméno.

Definice.

Nechť f je funkce definována na nějakém okolí bodu \vec{a} . Řekneme, že lineární zobrazení $L: \mathbb{R} \to \mathbb{R}$ je (totální) diferenciál $f \vee \vec{a}$, jestliže

$$\lim_{h\to 0} \left(\frac{f(a+h)-f(a)-L[h]}{h}\right) = 0.$$

Pokud takovéto lineární zobrazení existuje, pak jej značíme df(a), popř. Df(a).

Každé lineární zobrazení lze vyjádřit jako L[h] = Ah, takže hledání diferenciálu je vlastně hledáním A. Už jsme si připomněli, že vlastně víme, jak vypadá nejlepší lineární aproximace, jmenovitě by to mělo být f'(a)h neboli A = f'(a). Snadno to potvrdíme:

$$f'(a) = \lim_{h \to 0} \left(\frac{f(a+h) - f(a)}{h} \right)$$

$$\implies 0 = \lim_{h \to 0} \left(\frac{f(a+h) - f(a)}{h} \right) - f'(a) = \lim_{h \to 0} \left(\frac{f(a+h) - f(a)}{h} - f'(a) \right)$$

$$= \lim_{h \to 0} \left(\frac{f(a+h) - f(a) - f'(a)h}{h} \right).$$

Naopak, pokud máme v a diferenciál, tak máme nějaké A splňující

$$\lim_{h \to 0} \left(\frac{f(a+h) - f(a) - Ah}{h} \right) = 0.$$

Snadno to přeorganizujeme na tvar

$$A = \lim_{h \to 0} \left(\frac{f(a+h) - f(a)}{h} \right),$$

což ukazuje, že máme derivaci v a a A = f'(a).

Dokázali jsme, že u funkcí jedné proměnné je existence diferenciálu ekvivalentní existenci derivace a obojí nese stejnou informaci, jen trochu jinak vyjádřenou.

Ukáže se, že pokud se na derivaci díváme očima diferenciálu, tak je již možné ji zobecnit i pro funkce více proměnných.

V kapitole 3 jsme již nahlédli, že pokud je funkce dvou proměnných f(x,y) rozumná na okolí nějakého bodu (a,b), tak ji můžeme aproximovat tečnou rovinou danou gradientem, což lze napsat jako

$$f(a+h,b+k) - f(a,b) \approx \frac{\partial f}{\partial x}h + \frac{\partial f}{\partial y}k.$$

Na pravé straně tentokráte vidíme zobrazení ve tvaru

$$L[(h,k)] = Ah + Bk,$$

což je lineární zobrazení $\mathbb{R}^2 \mapsto \mathbb{R}$. Nalevo vlastně máme f((a,b) + (h,k)). Vidíme tedy přímou analogii jednorozměrného případu. Toto snadno zobecníme a budeme chtít, aby

$$f(\vec{a} + \vec{h}) - f(\vec{a}) \approx L[\vec{h}]$$

pro nějaké lineární zobrazení $L: \mathbb{R}^n \to \mathbb{R}$. Lineární algebra nás učí, že každé lineární zobrazení $L: \mathbb{R}^n \to \mathbb{R}$ je dáno určitým vektorem (A_1, \ldots, A_n) dle vzorce $L[\vec{h}] = \sum A_i h_i$, takže tento obecný pohled dobře zapadá do motivačních příkladů. Test, zda je nějaká lineární aproximace optimální, je obdobný jako u jednorozměrného případu.

Definice.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Řekneme, že lineární zobrazení $L: \mathbb{R}^n \to \mathbb{R}$ je (totální) diferenciál f v \vec{a} , jestliže

$$\lim_{\vec{h}\to\vec{0}}\Bigl(\frac{f(\vec{a}+\vec{h})-f(\vec{a})-L[\vec{h}]}{||\vec{h}||}\Bigr)=0.$$

Pokud takovéto lineární zobrazení existuje, pak jej značíme d $f(\vec{a})$, popř. $Df(\vec{a})$ a řekneme, že f je diferencovatelná v \vec{a} .

Diferenciál je to pravé zobecnění derivace. Platí například následující.

Věta.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Jestliže existuje d $f(\vec{a})$, pak je f spojitá v \vec{a} .

Máme i tuto větu:

Věta.

Nechť D je oblast v \mathbb{R}^n , $f: D \mapsto \mathbb{R}$. Jestliže je f diferencovatelná na D a df = 0 na D, pak je f konstantní na D.

Poznamenejme, že df=0 znamená, že je to triviální zobrazení, tedy d $f[\vec{h}]=0$ pro všechna \vec{h} . Jak existenci diferenciálu rozpoznáme a jak jej najdeme? Vzpomeneme-li si na to, jak se hledá tečná nadrovina pomocí gradientu, tak bychom si tipli, že onen vektor (A_1,\ldots,A_n) by měl odpovídat gradientu ∇f , tedy $L[\vec{h}]=\sum \frac{\partial f}{\partial x_i}h_i=\nabla f\bullet \vec{u}$. Jak už jsme napsali, pro jednorozměrné funkce jsou existence derivace a diferenciálu ekvivalentní, ale u vícerozměrného případu je ta souvislost poněkud komplikovanější

Začneme výsledkem, který je trochu obecnější.

Věta.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Jestliže existuje $\mathrm{d}f(\vec{a})$, pak pro každé $\vec{u} \in \mathbb{R}^n$ existuje $D_{\vec{u}}f(\vec{a})$ a platí

$$D_{\vec{u}}f(\vec{a}) = \mathrm{d}f(\vec{a})[\vec{u}].$$

Co dostaneme, když toto aplikujeme na $\vec{u} = \vec{e_i}$? Jestliže $\mathrm{d}f(\vec{a})[\vec{h}] = \sum A_i h_i$, pak

$$A_i = \mathrm{d}f(\vec{a})[\vec{e}_i] = D_{\vec{e}_i}f(\vec{a}) = \frac{\partial f}{\partial x_i}(\vec{a}).$$

Důsledek.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Jestliže existuje $\mathrm{d}f(\vec{a})$, pak pro všechny $i=1,\ldots,n$ existuje $\frac{\partial f}{\partial x_i}(\vec{a})$ a platí

$$\mathrm{d}f(\vec{a})[\vec{h}] = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a}) \cdot h_i = \nabla f(\vec{a}) \bullet \vec{h}.$$

Takže ∇f je opravdu ten správný vektor pro diferenciál, je to jediný možný kandidát. Všimněte si, že věta říká, že pokud máme diferenciál, tak máme i gradient. Bohužel, pro funkce více proměnných to už nemusí platit naopak. Někdy to vyjde.

Příklad: Pro $f(x,y) = x^2 + y^2$ jsme už v kapitole 3 našli $\nabla(f)(1,2) = (2,4)$. Podle věty je pak $(2,4) \bullet (h,k) = 2h + 4k$ jediný kandidát na diferenciál df(1,2). Ověříme podle definice, že to opravdu vyjde:

$$\lim_{\vec{h}\to\vec{0}} \left(\frac{f(\vec{a}+\vec{h}) - f(\vec{a}) - L[\vec{h}]}{\|\vec{h}\|} \right) = \lim_{(h,k)\to(0,0)} \left(\frac{[(1+h)^2 + (2+k)^2] - (1^2 + 2^2) - (2h + 4k)}{\sqrt{h^2 + k^2}} \right)$$

$$= \lim_{(h,k)\to(0,0)} \left(\frac{h^2 + k^2}{\sqrt{h^2 + k^2}} \right) = \lim_{(h,k)\to(0,0)} \left(\sqrt{h^2 + k^2} \right) = 0.$$

Potvrdili jsem, že $\mathrm{d}f(1,2)[h,k] = \nabla(f)(1,2) \bullet (h,k) = 2h+4k.$ \triangle

V kapitole 3 jsme ale ukázali, že funkce $\frac{x^2y^2}{x^4+y^4}$ má gradient v (0,0), ale nemá tam všechny směrové derivace, tudíž ani nemůže mít totální diferenciál. Ještě jeden takový případ:

Příklad: Uvažujme $f(x,y) = \sqrt{|xy|}$. Pak v jednotlivých kvadrantech, kde nám znalost znamének umožní zbavit se absolutní hodnoty, snadno derivujeme. Výsledky se dají sjednodit do vzorce $\nabla f(x,y) = \left(\frac{1}{2\sqrt{|xy|}}\operatorname{sign}(xy)y, \frac{1}{2\sqrt{|xy|}}\operatorname{sign}(xy)x\right)$, který platí pro $(x,y) \neq (0,0)$.

V počátku se pomocí definice dá dokázat, že $\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0$. Intuitivně je to zjevné, protože f(x,0) = f(0,y) = 0, tedy f je na řezech nad souřadnicovými osami konstantní a musí mít nulové příslušné parciální derivace.

Takže tato funkce má gradient dokonce na celém \mathbb{R}^2 , ale není diferencovatelná v (0,0) (je tam jen spojitá). To je vidět například tak, že se omezíme na přímku y=x, tam je funkce rovna f(x,x)=|x| a nemá směrovou derivaci.

Abychom z existence gradientu získali diferenciál, musíme ještě přidat spojitost.

Věta.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Jestliže existuje nějaké okolí U bodu \vec{a} takové, že $f \in C^1(U)$, pak je f diferencovatelná v \vec{a} a

$$\mathrm{d}f(\vec{a})[\vec{h}] = \nabla f(\vec{a}) \bullet \vec{h}.$$

Důsledek.

Nechť D je otevřená množina a $f \in C^1(D)$. Pak je f diferencovatelná na D a $\mathrm{d} f[\vec{h}] = \nabla f \bullet \vec{h}.$

Opět to neplatí naopak, existence diferenciálu sice dává existenci parciálních derivací, ale ne nutně jejich spojitost. Máme tedy hierarchii: Mnohé funkce mají gradient. Některé z nich mají i diferenciál, z nich zase některé mají dokonce spojitý gradient. Vlastnost spojitého gradientu je tedy nejsilnější, proto máme rádi funkce z C^1 .

Shrňme si používané zápisy diferenciálu, každý má své místo podle toho, co zrovna chceme s diferenciálem dělat.

$$df(\vec{a})[\vec{h}] = \frac{\partial f}{\partial x_1}(\vec{a})h_1 + \dots + \frac{\partial f}{\partial x_n}(\vec{a})h_n = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\vec{a})h_i = \nabla f(\vec{a}) \bullet \vec{h} = \nabla f(\vec{a})\vec{h}^T = D_{\vec{h}}f(\vec{a}).$$

Existuje ještě jeden pohled na diferenciál, který je docela užitečný v aplikacích. Když máme prostor vektorů, tak máme automaticky také funkce, které z vektorů izolují určitou souřadnici,

například $(x_1, x_2, \dots) \mapsto x_1$. Říká se tomu projekce a víme, že je lineární. Protože funkce $\vec{x} \mapsto x_i$ extrahuje z vektoru jeho x_i -tou souřadnici, je zvykem takovou projekci značit jako $x_i(\vec{x}) = x_i$.

Protože je to funkce, má smysl ptát se na její diferenciál. Jakožto lineární funkce má za graf nadrovinu, tudíž tečná rovina bude zase tato funkce. Jinak řečeno, tato funkce je sama sobě diferenciálem, $dx_i[\vec{x}] = x_i$.

Diferenciál nějaké funkce f pak můžeme napsat takto:

$$df(\vec{a}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\vec{a}) dx_i.$$

Například výsledek obdržený výše lze zapsat jako df(1,2) = 2dx + 4dy. Formálně se zde pracuje s diferenciály, ale častá praktická interpretace vidí spíše dx a dy neboli nekonečně malé infinitesimály jednotlivých proměnných. Spočítaný diferenciál by se pak interpretoval tak, že když se z bodu (1,2) posuneme o nekonečně malý vektor (dx, dy), pak se hodnota funkce změní o df = 2dx + 4dy. Takovéto úvahy se používají při intuitivním odvozování vzorců.

Mimochodem, pro funkci jedné proměnné pak máme df = f'(a)dx, což je vzorec známý ze substituce, který lze dále upravit na $\frac{df}{dx} = f'(a)$. Vidíme spoustu zajímavých souvislostí.

Jaká pravidla máme pro diferenciál? Protože je diferenciál daný gradientem, je to lineární operace, tedy platí, že d(cf+g) = c df + dg. U ostatních algebraických operací se pravidla příliš nepoužívají a proto toto téma přeskočíme. Naopak skládání je velice důležitá situace, podíváme se jmenovitě na transformace souřadnic.

Věta.

Nechť D je otevřená množina v \mathbb{R}^m , mějme funkci $f(y_1, \ldots, y_m): D \mapsto \mathbb{R}$ diferencovatelnou na D.

Uvažujme následující transformaci souřadnic: Nechť M je otevřená množina v \mathbb{R}^n , pro $j=1,\ldots,m$ nechť $y_j=y_j(x_1,\ldots,x_n)\in C^1(M)$ a platí $\vec{y}=(y_1,\ldots,y_m)$: $M\mapsto D$. Pak je $f\circ\vec{y}=f(y_1(x_1,\ldots,x_n),\ldots,y_m(x_1,\ldots,x_n))$ diferencovatelná na M a platí

$$d(f(\vec{y}))(\vec{x}) = \sum_{j=1}^{m} \frac{\partial f}{\partial y_j}(\vec{y}(\vec{x}))dy_j.$$

Totální diferenciál lze snadno zobecnit pro vektorové funkce. Máme-li funkci $F: \mathbb{R}^n \mapsto \mathbb{R}^m$, pak rozdíl $F(\vec{a} + \vec{h}) - F(\vec{a})$ je vektor z \mathbb{R}^m . Pokud jej budeme chtít aproximovat lineárním zaobrazením závisejícím na \vec{h} , tak toto zobrazení musí mít definiční obor v \mathbb{R}^n a dávat vektory z \mathbb{R}^m , tedy musí to být vektorová funkce $\mathbb{R}^n \mapsto \mathbb{R}^m$.

Definice.

Nechť F je vektorová funkce s hodnotami v \mathbb{R}^m definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Řekneme, že lineární zobrazení $L: \mathbb{R}^n \mapsto \mathbb{R}^m$ je (totální) diferenciál F v \vec{a} , jestliže

$$\lim_{\vec{h} \to \vec{0}} \Bigl(\frac{\|F(\vec{a} + \vec{h}) - F(\vec{a}) - L[\vec{h}]\|}{\|\vec{h}\|} \Bigr) = 0.$$

Pokud takovéto lineární zobrazení existuje, pak jej značíme d $F(\vec{a})$, popř. $DF(\vec{a})$, a řekneme, že F je diferencovatelná v \vec{a} .

I pro vektorové funkce platí věta, kterou chceme.

Věta.

Nechť F je vektorová funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Jestliže existuje d $F(\vec{a})$, pak je F spojitá v \vec{a} .

Každé lineární zobrazení $\mathbb{R}^n \to \mathbb{R}^m$ je dáno nějakou $m \times n$ maticí A podle vzorce $L[\vec{h}] = A\vec{h}^T$. Jaká matice dává totální diferenciál? Je to Jacobiho matice funkce F, tedy

$$dF(\vec{a})[\vec{h}] = J_F(\vec{a})\vec{h}^T.$$

I pro vektorové funkce platí analogická tvrzení spojující existenci diferenciálu a gradientu. Máme tedy hierarchii tří kvalit, kde nejvýše stojí funkce z $[C^1(D)]^m$, tedy funkce se spojitými parciálními derivacemi. Protože je existence diferenciálu spojena s Jacobiho maticí, lze řadu poznatků z předchozích sekcí přepsat do jazyka diferenciálu. Získáme tak například větu o diferenciálu složené funkce.

Věta.

Nechť $F: D(F) \mapsto \mathbb{R}^m$ je vektorová funkce, kde $D(F) \subseteq \mathbb{R}^n$.

Nechť $G: D(G) \to \mathbb{R}^p$ je vektorová funkce, kde $D(G) \subseteq \mathbb{R}^m$.

Nechť \vec{a} je vnitřní bod D(F) takový, že $\vec{b}=F(\vec{a})$ je vnitřní bod D(G). Předpokládejme, že F je diferencovatelná v \vec{a} a G je diferencovatelná v \vec{b} . Pak $G\circ F$ je diferencovatelná v \vec{a} a platí

$$d(G \circ F)(\vec{a}) = dG(F(\vec{a})) \circ dF(\vec{a}).$$

Zvedli jsme diferenciál coby zobecnění derivace a přirozená otázka zní, jak se zobecní vyšší derivace. Pro jednu proměnnou je najdeme tak, že se derivuje funkce, která ukazuje závislost derivace na pozici, tedy funkce $x\mapsto f'(x)$. Diferenciál $\mathrm{d}f(\vec{a})[\vec{h}]$ má dvě proměnné, což vyřešíme tak, že si zvolíme nějaký směrový vektor $\vec{k}\in\mathbb{R}^n$ a podíváme se na funkci $\vec{x}\mapsto \mathrm{d}f(\vec{x})[\vec{k}]$. To je funkce více proměnných, tudíž se můžeme zeptat na její diferenciál. Výsledek ale bude záležet na \vec{k} , takže se v odpovědi objeví jak \vec{h} , tak \vec{k} . U obou složek očekáváme (a vyžadujeme) lineární chování, takže nejprve definice.

Definice.

Zobrazení $L: \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}^m$ se nazývá **bilineární**, jestliže

$$L(c\vec{x}_1 + \vec{x}_2, \vec{y}) = cL(\vec{x}_1, \vec{y}) + L(\vec{x}_2, \vec{y})$$

$$L(\vec{x}, c\vec{y}_1 + \vec{y}_2) = cL(\vec{x}, \vec{y}_1) + L(\vec{x}, \vec{y}_2)$$

pro všechny vektory $\vec{x}, \vec{y} \in \mathbb{R}^n$ a skaláry $c \in \mathbb{R}$.

Teď si mezi bilineárními zobrazeními jedno vybereme.

Definice.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$. Řekneme, že bilineární zobrazení $L: \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ je $d^2 f(\vec{a})$, jestliže pro všechna $\vec{k} \in \mathbb{R}^n$ je

$$\lim_{\vec{h}\to\vec{0}}\Bigl(\frac{\mathrm{d}f(\vec{a}+\vec{h})[\vec{k}]-\mathrm{d}f(\vec{a})[\vec{k}]-L[\vec{h},\vec{k}]}{\|\vec{h}\|}\Bigr)=0.$$

Tato definice opravdu dělá to, co očekáváme od druhé derivace:

Fakt.
$$d^2 f(\vec{a})[\vec{h}, \vec{k}] = d(df[\vec{k}])[\vec{h}].$$

Co se myslí výrazem napravo? Ten vnější diferenciál se dělá z funkce $\vec{x} \mapsto \mathrm{d}f(\vec{x})[\vec{k}]$, přesně jak bychom chtěli.

Druhý diferenciál je tedy opravdu diferenciál z diferenciálu, podobně jako je druhá derivace vlastně derivace z derivace. Dává to smysl, ale jak se to spočítá?

Lineární algebra říká, že každé bilineární zobrazení lze vyjádřit pomocí jisté matice $A=(a_{ij})_{i,j=1}^n$ tak, že $L[\vec{h},\vec{k}]=\sum_{i,j=1}^n a_{ij}h_ik_j=\vec{h}A\vec{k}^T$. Nás samozřejmě zajímá, jaká matice A nám dá d 2f . V tom nám pomůže onen fakt výše. My vlastně pro pevně zvolené \vec{k} hledáme totální diferenciál funkce $\vec{x}\mapsto\sum_{i=1}^n\frac{\partial f}{\partial x_j}(\vec{x})k_j$, díky linearitě se stačí zamyslet, že

$$d\left(k_j \frac{\partial f}{\partial x_j}(\vec{x})\right) = k_j d\left(\frac{\partial f}{\partial x_j}(\vec{x})\right) = k_j \sum_{i=1}^n \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j} \cdot h_i = k_j \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j} h_i,$$

a pak to posčítat.

Věta

Nechť D je otevřená množina v \mathbb{R}^n . Jestliže $f \in C^2(D)$, pak pro $\vec{a} \in D$ máme

$$d^{2}f(\vec{a})[\vec{h}, \vec{k}] = \sum_{i,j=1}^{n} \frac{\partial^{2}f}{\partial x_{i}\partial x_{j}}(\vec{a})h_{i}k_{j}.$$

Vidíme, že matice, která udává naše bilineární zobrazení d 2f , je matice $H=\left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$ neboli Hessova matice známá z kapitoly 3. Druhý diferenciál tedy najdeme jako

$$d^2 f(\vec{a})[\vec{h}, \vec{k}] = \vec{h} H(\vec{a}) \vec{k}^T.$$

Jak by fungovaly diferenciály vyšších řádů? Používali bychom n-lineární zobrazení, která lze vyjádřit pomocí "mnoharozměrných matic". Například d $^3f[\vec{u},\vec{v},\vec{w}] = \sum_{i,j,k=1}^n a_{ijk}u_iv_jw_k$, přičemž

se dá dokázat, že $a_{ijk} = \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}$, namísto matice bychom si představili něco jako krychli s parciálními derivacemi třetího řádu. Čtenář si už sám představí vyšší diferenciály.

Takovéto sumy jsou v praxi dosti nepříjemné, jak záhy uvidíme. Je tam také docela dost parciálních derivací. Tady pomůže, když pracujeme s funkcemi se spojitými parciálními derivacemi, kde na pořadí derivování nezáleží a počet nutných výpočtů se revolučně zmenší. Všimneme si, že pak při výběru proměnných, podle kterých derivujeme, nezáleží na jménech směrů \vec{u}, \vec{v} atd., ale jen na indexech u nich. Například v diferenciálu čtvrtého řádu d $f(\vec{a})[\vec{t}, \vec{u}, \vec{v}, \vec{w}]$ má typický člen tvar

$$\frac{\partial^4 f}{\partial x_i \partial x_j \partial x_k \partial x_l}(\vec{a}) t_i u_j v_k w_l.$$

Pracuje se tam s i-tou souřadnicí \vec{t} , takže derivujeme podle i-té proměnné, obdobně pro ostatní členy. Takže když vybíráme správnou parciální derivaci, tak se prostě v členu $t_i u_j v_k w_l$ podíváme na indexy. Například když chceme najít koeficient u $t_1 u_4 v_2 w_1$, tak projdeme indexy a vidíme, že máme derivovat dvakrát vzhledem k x_1 a jednou vzhledem k x_2 a x_4 . Pokud je funkce f dostatečně hladká, tak na pořadí nezáleží a můžeme použít třeba $\frac{\partial^2 f}{\partial x_4 \partial x_2 \partial x_1^2}$, ale jiná pořadí budou také správně.

Příklad: Najdeme diferenciály prvních tří řádů pro $f(x,y) = x^2y + y^3$ v bodě $\vec{a} = (1,2)$. Protože má tato funkce spojité parciální derivace všech řádů, ušetříme spoustu výpočtů díky symetrii smíšených derivací.

Protože zde nebudeme používat x_1, x_2 jako ve větě, ale pracujeme se jmény x, y, vyjasníme si, jak se určují parciální derivace pro koeficienty. Když vidíme součin typu $h_i k_j$ či $u_i v_j w_k$, tak projdeme indexy a vidíme, podle kterých proměnných máme derivovat. Ve větě proměnné poznáváme podle jejich čísla, zde podle abecedního pořadí. Index 1 indikuje derivaci podle x, index 2 indikuje derivaci podle y. Například abychom našli koeficient u $u_1 v_2 w_1$, budeme derivovat dvakrát podle x a jednou podle y.

Nejprve diferenciál prvního řádu: Potřebujeme $\frac{\partial f}{\partial x} = 2xy$ a $\frac{\partial f}{\partial y} = x^2 + 3y^2$.

Podle součtového vzorce máme

$$df(1,2)[(u_1,u_2)] = \frac{\partial f}{\partial x}(1,2)u_1 + \frac{\partial f}{\partial y}(1,2)u_2 = 4u_1 + 13u_2.$$

Nebo najdeme $\nabla f(1,2) = (4,13)$ a

$$df(1,2)[(u_1,u_2)] = (4,13) \bullet (u_1,u_2) = 4u_1 + 13u_2.$$

Diferenciál druhého řádu: Potřebujeme $\frac{\partial^2 f}{\partial x^2}=2y, \ \frac{\partial^2 f}{\partial x \partial y}=2x$ a $\frac{\partial^2 f}{\partial y^2}=6y$.

Podle součtového vzorce máme

$$df(1,2)[(u_1, u_2)] = \frac{\partial^2 f}{\partial x^2}(1,2)u_1v_1 + \frac{\partial^2 f}{\partial x \partial y}(1,2)u_1v_2 + \frac{\partial^2 f}{\partial y \partial x}(1,2)u_2v_1 + \frac{\partial^2 f}{\partial y^2}(1,2)u_2v_2$$

$$= 4u_1v_1 + 2u_1v_2 + 2u_2v_1 + 12u_2v_2.$$

Nebo použijeme $H(1,2)=\begin{pmatrix} 4 & 2 \\ 2 & 12 \end{pmatrix}$ a

$$d^2 f(1,2)[(u_1, u_2), (v_1, v_2)] = (u_1, u_2) \begin{pmatrix} 4 & 2 \\ 2 & 12 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Diferenciál třetího řádu: Potřebujeme $\frac{\partial^3 f}{\partial x^3} = 0$, $\frac{\partial^3 f}{\partial x^2 \partial y} = 2$, $\frac{\partial^3 f}{\partial x \partial y^2} = 0$ a $\frac{\partial^3 f}{\partial y^3} = 6$.

Dosadíme bod (1,2) (to je snadné) a počítáme, musíme si v tom udělat pořádek, ať nezapomeneme na žádný z 2^3 členů trojité sumy. Musí se tam objevit kombinace indexů 111, 112, 121, 122, 211, 212, 221, 222, u každého takového součinu $u_i v_j w_k$ bude jistá parciální derivace, což jsme si před chvílí rozmysleli.

$$d^{3}f(1,2)[(u_{1},u_{2}),(v_{1},v_{2}),(w_{1},w_{2})]$$

$$= 0u_{1}v_{1}w_{1} + 2u_{1}v_{1}w_{2} + 2u_{1}v_{2}w_{1} + 0u_{1}v_{2}w_{2} + 2u_{2}v_{1}w_{1} + 0u_{2}v_{1}w_{2} + 0u_{2}v_{2}w_{1} + 6u_{2}v_{2}w_{2}$$

$$= 2u_{1}v_{1}w_{2} + 2u_{1}v_{2}w_{1} + 2u_{2}v_{1}w_{1} + 6u_{2}v_{2}w_{2}.$$

Docela maso.

 \triangle

Nešel by výpočet diferenciálů vyšších řádů nešel nějak zjednodušit? Obecně bohužel ne. Nicméně diferenciály se často používají pro stejný směr, tedy $d^2f(\vec{a})[\vec{h},\vec{h}],\ d^3f(\vec{a})[\vec{h},\vec{h},\vec{h}]$ atd. Jak by pak dopadly diferenciály v příkladě výše? Mnohem lépe, protože pak lze spojit členy se smíšenými indexy. Použijeme-li výpočty z příkladu, dostáváme

$$df(1,2)[(h_1,h_2)] = 4h_1 + 13h_2,$$

$$d^2f(1,2)[(h_1,h_2),(h_1,h_2)] = 4h_1^2 + 2h_1h_2 + 2h_2h_1 + 12h_2^2$$

$$= 4h_1^2 + 4h_1h_2 + 12h_2^2,$$

$$d^3f(1,2)[(h_1,h_2),(h_1,h_2),(h_1,h_2)] = 2h_1h_1h_2 + 2h_1h_2h_1 + 2h_2h_1h_1 + 6h_2h_2h_2$$

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

Uměli bychom se k těmto výsledkům dostat efektivněji? Zde odpověď zní, že ano. Připomeneme si, že gradient je možné vnímat jako diferenciální operátor $\nabla = \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n}\right)$, který se také může stát součástí výrazů. To nám umožnilo najít alternativní vyjádření směrové derivace, tedy také

prvního diferenciálu:

$$\mathrm{d}f[\vec{h}] = \sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i} f = \left(\sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i}\right) f = (\vec{h} \bullet \nabla) f.$$

Teď se podíváme na druhý diferenciál. Tam máme

$$d^{2}f(\vec{a})[\vec{h}, \vec{h}] = \sum_{i,j=1}^{n} \frac{\partial^{2}f}{\partial x_{i}\partial x_{j}} h_{i}h_{j} = \left(\sum_{i,j=1}^{n} h_{i}h_{j} \frac{\partial^{2}}{\partial x_{i}\partial x_{j}}\right) f = \left(\sum_{i,j=1}^{n} h_{i}h_{j} \frac{\partial}{\partial x_{i}} \frac{\partial}{\partial x_{j}}\right) f$$
$$= \left(\sum_{i=1}^{n} h_{i} \frac{\partial}{\partial x_{i}}\right) \left(\sum_{j=1}^{n} h_{j} \frac{\partial}{\partial x_{j}}\right) f = (\vec{h} \cdot \nabla)(\vec{h} \cdot \nabla) f = (\vec{h} \cdot \nabla)^{2} f.$$

Takže druhý diferenciál je prostě jen dvakrát aplikovaný operátor $(\vec{h} \bullet \nabla)$. Pro vyšší diferenciály to funguje zcela analogicky. Nejlepší na tom je, že není třeba opakovaně aplikovat tento operátor, jako když vyšší derivaci hledáme opakovaným derivováním, ale můžeme si nejprve zjistit, jak vlastně celý operátor funguje, roznásobením mocniny pomocí obvyklých vzorců. Například takto se najde třetí diferenciál.

$$(\vec{h} \bullet \nabla)^3 = \left(h_1 \frac{\partial}{\partial x_1} + h_2 \frac{\partial}{\partial x_2}\right)^3$$

$$= \left(h_1 \frac{\partial}{\partial x_1}\right)^3 + 3\left(h_1 \frac{\partial}{\partial x_1}\right)^2 h_2 \frac{\partial}{\partial x_2} + 3h_1 \frac{\partial}{\partial x_1} \left(h_2 \frac{\partial}{\partial x_2}\right)^2 + \left(h_2 \frac{\partial}{\partial x_2}\right)^3$$

$$= h_1^3 \frac{\partial^3}{\partial x_1^3} + 3h_1^2 h_2 \frac{\partial^3}{\partial x_1^2 \partial x_2} + 3h_1 h_2^2 \frac{\partial^3}{\partial x_1 \partial x_2^2} + h_2^3 \frac{\partial^3}{\partial x_2^3}.$$

Teď tento operátor aplikujeme na f a dostaneme

$$d^{3}f[(h_{1}, h_{2}), (h_{1}, h_{2}), (h_{1}, h_{2})] = \frac{\partial^{3}f}{\partial x_{1}^{3}}h_{1}^{3} + \frac{\partial^{3}f}{\partial x_{1}^{2}\partial x_{2}}3h_{1}^{2}h_{2} + \frac{\partial^{3}f}{\partial x_{1}\partial x_{2}^{2}}3h_{1}h_{2}^{2} + \frac{\partial^{3}f}{\partial x_{2}^{3}}h_{2}^{3}.$$

Pokud toto aplikujeme na funkci f a bod (1,2) z příkladu, tak dostaneme shodný vzorec

$$d^{3}f(1,2)[(h_{1},h_{2}),(h_{1},h_{2}),(h_{1},h_{2})] = 6h_{1}^{2}h_{2} + 6h_{2}^{3}$$

Tento postup je vysoce efektivní, bude se nám velice hodit v následující sekci.

7d. Taylorův polynom

Taylorův polynom je inspirován následující otázkou: Máme dobré informace o funkci f v bodě \vec{a} , a chtěli bychom z toho odvodit co nejkvalitnější informaci o hodnotě f v bodě \vec{x} , který je typicky blízko k \vec{a} .

Pro funkce jedné proměnné nám tuto službu poskytuje právě Taylorův polynom. Pro funkci více proměnných se k této situaci dostaneme oblíbeným trikem, tedy přechodem k řezu. Od bodu \vec{a} se k bodu \vec{x} dostaneme nejsnáze po přímce, ta vede směrem $\vec{x} - \vec{a}$. My ale preferujeme jednotkový směrový vektor, proto vezmeme $\vec{u} = \frac{\vec{x} - \vec{a}}{\|\vec{x} - \vec{a}\|}$. Pak $\vec{x} - \vec{a} = \|\vec{x} - \vec{a}\|\vec{u}$. Naše přímka má tedy parametrické vyjádření $\vec{a} + t\vec{u}$, přičemž nás zajímá zejména pro $t_x = \|\vec{x} - \vec{a}\|$, protože pak $\vec{a} + t_x\vec{u} = \vec{x}$.

Na této přímce můžeme funkci f sledovat prostřednictvím pomocné funkce $\varphi(t) = f(\vec{a} + t\vec{u})$. Zajímavé jsou zejména hodnoty $\varphi(0) = f(\vec{a})$ a $\varphi(t_x) = f(\vec{x})$. Pokud má funkce f dost derivací v bodě \vec{a} , bude mít také funkce φ dost derivací v odpovídající hodnotě parametru t = 0 a můžeme použít Taylorův polynom k odhadu její hodnoty:

$$f(\vec{x}) = \varphi(t_x) \approx \varphi(0) + \varphi'(0)t_x + \frac{1}{2!}\varphi''(0)t_x^2 + \frac{1}{3!}\varphi'''(0)t_x^3 + \ldots + \frac{1}{N!}\varphi^{(N)}(0)t_x^N.$$

My ovšem víme, že derivace φ souvisejí se směrovými derivacemi funkce f ve směru \vec{u} . V kapitole 3 jsme viděli, že $\varphi'(0) = D_{\vec{u}}f(\vec{a})$, snadno se ukáže, že také $\varphi''(0) = D_{\vec{u}}D_{\vec{u}}f(\vec{a}) = D_{\vec{u}}^2f(\vec{a})$ a tak dále. Pokud si ještě připomeneme, že $t_x = ||\vec{x} - \vec{a}||$, můžeme psát

$$f(\vec{x}) \approx f(\vec{a}) + D_{\vec{u}}f(\vec{a}) \|\vec{x} - \vec{a}\| + \frac{1}{2}D_{\vec{u}}^2 f(\vec{a}) \|\vec{x} - \vec{a}\|^2 + \dots + \frac{1}{N!}D_{\vec{u}}^N f(\vec{a}) \|\vec{x} - \vec{a}\|^N.$$

To už je výraz, který přímo pracuje s funkcí f, takže bychom jej mohli vzít jako inspiraci pro definici Taylorova polynomu pro funkce více proměnných. Směrová derivace má ale z teoretického pohledu jisté nedostatky, takže na tom výrazu ještě zapracujeme. Nejprve si připomeneme, že $D_{c\vec{u}}f = cD_{\vec{u}}f$. Pak

$$D_{c\vec{u}}^2 f = D_{c\vec{u}}[D_{c\vec{u}}f] = D_{c\vec{u}}[cD_{\vec{u}}f] = cD_{c\vec{u}}[D_{\vec{u}}f] = c^2D_{\vec{u}}[D_{\vec{u}}f] = c^2D_{\vec{u}}^2 f.$$

Obdobně se ukáže, že obecně $D^k_{\vec{u}}f\cdot c^k=D^k_{c\vec{u}}f,$ vzorec tak lze díky $\|\vec{x}-\vec{a}\|\vec{u}=\vec{x}-\vec{a}$ přepsat jako

$$f(\vec{x}) \approx f(\vec{a}) + D_{\vec{x}-\vec{a}}f(\vec{a}) + \frac{1}{2}D_{\vec{x}-\vec{a}}^2f(\vec{a}) + \frac{1}{3!}D_{\vec{x}-\vec{a}}^3f(\vec{a}) + \dots + \frac{1}{N!}D_{\vec{x}-\vec{a}}^Nf(\vec{a}).$$

Viděli jsme, že $D_{\vec{h}}f$ odpovídá diferenciálu d $f[\vec{h}]$. Víme také, že vyšší diferenciály vzniknou opakovanou aplikací diferenciálu, takže

$$d^2 f[\vec{h}, \vec{h}] = df (df[\vec{h}])[\vec{h}] = D_{\vec{h}} (df[\vec{h}]) = D_{\vec{h}} D_{\vec{h}} f = D_{\vec{h}}^2 f.$$

Obdobně to platí i pro vyšší řády, což nás přivádí ke tvaru

$$f(\vec{x}) \approx f(\vec{a}) + df(\vec{a})[\vec{x} - \vec{a}] + \frac{1}{2}d^2f(\vec{a})[\vec{x} - \vec{a}, \vec{x} - \vec{a}] + \dots + \frac{1}{N!}d^Nf(\vec{a})[\vec{x} - \vec{a}, \dots, \vec{x} - \vec{a}].$$

Pro dostatečně rozumné funkce f by měl výraz napravo nabízet dobrou aproximaci, navíc vidíme zjevnou podobnost s klasickým Taylorovým polynomem. Následující definici jde tedy čekat.

Definice.

Nechť f je funkce definovaná na nějakém okolí bodu $\vec{a} \in \mathbb{R}^n$, která je v \vec{a} diferencovatelná až do řádu $N \in \mathbb{N}$.

Pak definujeme její **Taylorův polynom** v \vec{a} stupně N jako

$$T_{N}(\vec{x}) = f(\vec{a}) + df(\vec{a})[\vec{x} - \vec{a}] + \frac{1}{2}d^{2}f(\vec{a})[\vec{x} - \vec{a}, \vec{x} - \vec{a}] + \cdots$$
$$\cdots + \frac{1}{N!}d^{N}f(\vec{a})[\vec{x} - \vec{a}, \dots, \vec{x} - \vec{a}]$$
$$= \sum_{k=0}^{N} \frac{1}{k!}d^{k}f(\vec{a})[\vec{x} - \vec{a}, \dots, \vec{x} - \vec{a}].$$

V této podobě je Taylorův polynom přirozeným zobecněním definice pro jednu proměnnou, protože už víme, že diferenciály jsou tím správným zobecněním běžných derivací. Jak brzy uvidíme, nabízí také efektivní postup výpočtu. Přepis do směrových derivací má ale také své výhody, například nabízí vhodný jazyk pro zápis obecné verze jednoho z klíčových poznatků o Taylorově polynomu, jmenovitě větu o Lagrangeově tvaru zbytku.

Věta.

Nechť $\vec{a} \in \mathbb{R}^n$. Nechť $f \in C^{n+1}(U)$ pro nějaké okolí U bodu \vec{a} . Pak pro každé $\vec{x} \in U$ existuje $\vec{c} \in \langle \vec{a}, \vec{x} \rangle$ takové, že při označení $\vec{u} = \frac{\vec{x} - \vec{a}}{\|\vec{x} - \vec{a}\|}$ máme

$$f(\vec{x}) - T_N(\vec{x}) = \frac{1}{(N+1)!} D_{\vec{x}-\vec{a}}^{N+1} f(\vec{c}) = \frac{1}{(N+1)!} D_{\vec{u}}^{N+1} f(\vec{c}) ||\vec{x} - \vec{a}||^{N+1}.$$

Zde $\langle \vec{a}, \vec{x} \rangle$ označuje úsečku v \mathbb{R}^n , jinak řečeno množinu bodů ve tvaru $\vec{a} + t(\vec{x} - \vec{a})$ pro $t \in \langle 0, 1 \rangle$. V aplikacích často preferujeme variantu, kde nepřecházíme od \vec{a} do \vec{x} , ale od \vec{a} do $\vec{a} + \vec{h}$. Odpovídající verzi snadno vytvoříme přímo s $\vec{x} - \vec{a} = \vec{h}$ či s použitím jednotkového směrového vektoru $\vec{u} = \frac{\vec{h}}{\|\vec{h}\|}$. Nabídneme tři verze.

$$f(\vec{a} + \vec{h}) \approx f(\vec{a}) + df(\vec{a})[\vec{h}] + \frac{1}{2}d^{2}f(\vec{a})[\vec{h}, \vec{h}] + \dots + \frac{1}{N!}d^{N}f(\vec{a})[\vec{h}, \dots, \vec{h}]$$

$$\approx f(\vec{a}) + D_{\vec{h}}f(\vec{a}) + \frac{1}{2}D_{\vec{h}}^{2}f(\vec{a}) + \dots + \frac{1}{N!}D_{\vec{h}}^{N}f(\vec{a})$$

$$\approx f(\vec{a}) + D_{\vec{u}}f(\vec{a})||\vec{h}|| + \frac{1}{2}D_{\vec{u}}^{2}f(\vec{a})||\vec{h}||^{2} + \dots + \frac{1}{N!}D_{\vec{u}}^{N}f(\vec{a})||\vec{h}||^{N}.$$

Taylorův polynom tedy opravdu i ve více dimenzích funguje dle očekávání. Jak jej spočítáme efektivně? Na to si připomeneme praktický přepis diferenciálu.

$$T_N(\vec{x}) = f(\vec{a}) + ((\vec{x} - \vec{a}) \bullet \nabla) f(\vec{a}) + \frac{1}{2} ((\vec{x} - \vec{a}) \bullet \nabla)^2 f(\vec{a}) + \dots + \frac{1}{N!} ((\vec{x} - \vec{a}) \bullet \nabla)^N f(\vec{a})$$

$$= \sum_{k=0}^{N} \frac{1}{k!} ((\vec{x} - \vec{a}) \bullet \nabla)^k f(\vec{a}).$$

S těmito výrazy již umíme pracovat. I zde se nabízí alternativní přepis

$$T_N(\vec{a} + \vec{h}) = \sum_{k=0}^N \frac{1}{k!} (\vec{h} \bullet \nabla)^k f(\vec{a}).$$

Příklad: Uvažujme funkci $f(x,y) = \cos(xy) + xy$. Najdeme pro ni T_2 se středem v bodě $(0,\pi)$.

Nejprve spočítáme parciální derivace, pro úsporu místa použijeme oblíbené fyzikální značení:

$$f_x = -\sin(xy) \cdot y + y, f_y = -\sin(xy) \cdot x + x;$$

$$f_{xx} = -\cos(xy) \cdot y^2$$
, $f_{yx} = f_{xy} = -\cos(xy) \cdot xy - \sin(xy) + 1$, $f_{yy} = -\cos(xy) \cdot x^2$.

Dosadíme: $f_x(0,\pi) = \pi, f_y(0,\pi) = 0;$

$$f_{xx}(0,\pi) = -\pi^2$$
, $f_{yx}(0,\pi) = f_{xy}(0,\pi) = 1$, $f_{yy}(0,\pi) = 0$.

Ještě budeme potřebovat $f(0,\pi) = 1$.

Protože jde o polynom stupně dva a pro tyto diferenciály máme příjemné vzorce

$$\mathrm{d}f(\vec{a})[\vec{h}] = \nabla f \bullet \vec{h}, \quad \mathrm{d}^2 f(\vec{a})[\vec{h}, \vec{h}] = \vec{h}H(\vec{a})\vec{h}^T,$$

mohli bychom najít hledaný polynom přímo podle definice.

$$T(x,y) = f(0,\pi) + df(0,\pi)[(x,y) - (0,\pi)] + \frac{1}{2}d^2f(0,\pi)[(x,y) - (0,\pi), (x,y) - (0,\pi)]$$

$$= 1 + (\pi,0) \bullet ((x,y) - (0,\pi)) + \frac{1}{2}((x,y) - (0,\pi)) \begin{pmatrix} -\pi^2 & 1\\ 1 & 0 \end{pmatrix} ((x,y) - (0,\pi))^T$$

$$= 1 + (\pi,0) \bullet (x,y-\pi) + \frac{1}{2}(x,y-\pi) \begin{pmatrix} -\pi^2 & 1\\ 1 & 0 \end{pmatrix} \begin{pmatrix} x\\ y-\pi \end{pmatrix}$$

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

U polynomů vyššího stupně už bychom museli použít $(\vec{h} \bullet \nabla)^k f$, tak si to tady natrénujeme. Nejprve si zjistíme, jak vypadají potřebné diferenciály:

$$(((x,y) - (0,\pi)) \bullet \nabla) = ((x - 0, y - \pi) \bullet \nabla) = x \frac{\partial}{\partial x} + (y - \pi) \frac{\partial}{\partial y},$$
$$(((x,y) - (0,\pi)) \bullet \nabla)^2 = \left(x \frac{\partial}{\partial x} + (y - \pi) \frac{\partial}{\partial y}\right)^2 = x^2 \frac{\partial^2}{\partial x^2} + 2x(y - \pi) \frac{\partial^2}{\partial x \partial y} + (y - \pi)^2 \frac{\partial^2}{\partial y^2}.$$

Nyní tyto diferenciály aplikujeme na f a dosadíme $(0, \pi)$:

$$((x,y-\pi) \bullet \nabla) f(0,\pi) = x \frac{\partial f}{\partial x}(0,\pi) + (y-\pi) \frac{\partial f}{\partial y}(0,\pi) = \pi x,$$

$$((x,y-\pi) \bullet \nabla)^2 f(0,\pi) = x^2 \frac{\partial^2 f}{\partial x^2}(0,\pi) + 2x(y-\pi) \frac{\partial^2 f}{\partial x \partial y}(0,\pi) + (y-\pi)^2 \frac{\partial^2 f}{\partial y^2}(0,\pi)$$

$$= -\pi^2 x^2 + 2x(y-\pi).$$

Takže

$$T_2(x,y) = 1 + \pi x + \frac{1}{2} \left(-\pi^2 x^2 + 2x(y-\pi) \right),$$

což dává stejný polynom jako předtím.

Další možný přístup je spočítat $T_N(0+h,\pi+k)$. Našli bychom

$$\begin{split} \left((h,k) \bullet \nabla \right) &= h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y}, \\ \left((h,k) \bullet \nabla \right)^2 &= h^2 \frac{\partial^2}{\partial x^2} + 2hk \frac{\partial^2}{\partial x \partial y} + k^2 \frac{\partial^2}{\partial y^2}, \end{split}$$

pak

$$((h,k) \bullet \nabla) f(0,\pi) = h \frac{\partial f}{\partial x}(0,\pi) + k \frac{\partial f}{\partial y}(0,\pi) = \pi h,$$

$$((h,k) \bullet \nabla)^2 f(0,\pi) = h^2 \frac{\partial^2 f}{\partial x^2}(0,\pi) + 2hk \frac{\partial^2 f}{\partial x \partial y}(0,\pi) + k^2 \frac{\partial^2 f}{\partial y^2}(0,\pi) = -\pi^2 h^2 + 2hk.$$

Konečně

$$T(0+h,\pi+k) = f(0,\pi) + df(0,\pi)[(h,k)] + \frac{1}{2}d^2f(0,\pi)[(h,k),(h,k)]$$
$$= f(0,\pi) + ((h,k) \bullet \nabla)f(0,\pi) + \frac{1}{2}((h,k) \bullet \nabla)^2f(0,\pi)$$
$$= 1 + \pi h - \frac{1}{2}\pi^2h^2 + hk.$$

Protože f je rozumná funkce, očekáváme, že pro malá h,k je

$$f(0+h,\pi+k) \approx 1 + \pi h - \frac{1}{2}\pi^2 h^2 + hk.$$

 \triangle

Na závěr se vrátíme do kapitoly 4, kde jsme si uvedli Sylvestrovo kritérium pro klasifikaci lokálních extrémů. Odkud se vzalo?

Uvažujme funkci f definovanou na okolí stacionárního bodu \vec{a} . Máme tedy podezření, že je tam lokální extrém. Pokud je tato funkce dostatečně hladká (dvakrát spojitě diferencovatelná), pak bychom ji mohli na okolí bodu \vec{a} aproximovat Taylorovým polynomem stupně dva. Vezmeme-li v úvahu, že $\nabla f(\vec{a}) = \vec{0}$, pak

$$f(\vec{a} + \vec{h}) \approx f(\vec{a}) + \frac{1}{2}\vec{h}H\vec{h}^T,$$

kde $H = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$ je Hessova matice. Potřebujeme vědět, zda jsou hodnoty f okolo \vec{a} větší či menší než $f(\vec{a})$.

(Ostré) lokální minimum vyžaduje $f(\vec{a}+\vec{h}) > f(\vec{a})$, tedy rozpoznáme jej podle toho, že $\vec{h}H\vec{h}^T > 0$ pro všechna nenulová \vec{h} . Obdobně (ostré) lokální maximum nastane, pokud $\vec{h}H\vec{h}^T < 0$ pro všechna nenulová \vec{h} . Zajímavou shodou okolností přesně tuto vlastnost znají v lineární algebře.

Definice.

Nechť A je symetrická matice $n \times n$.

Řekneme, že A je **positivně definitní**, jestliže

 $\vec{x}A\vec{x}^T > 0$ pro všechny $\vec{x} \in \mathbb{R}^n - \{\vec{0}\}.$

Řekneme, že A je **negativně definitní**, jestliže

 $\vec{x}A\vec{x}^T < 0$ pro všechny $\vec{x} \in \mathbb{R}^n - \{\vec{0}\}.$

Řekneme, že A je **indefinitní**, jestliže

existují $\vec{x}, \vec{y} \in \mathbb{R}^n$ takové, že $\vec{x} A \vec{x}^T > 0$ a $\vec{y} A \vec{y}^T < 0$.

Tato vlastnost je velice užitečná v některých aplikacích, zde je dokonce evidentně klíčová. Vidíme, že pokud je Hessova matice positivně definitní, tak máme lokální minimum, negativní definitnost dá lokální maximum a indefinitní Hessova matice signalizuje stacionární bod, který není lokálním extrémem. Jak se definitnost u matice pozná?

Pro diagonální matice je to snadné, pak totiž $a_{ij}=0$ pro $i\neq j$ a tedy

$$\vec{h}A\vec{h}^T = \sum_{i,j=1}^n a_{ij}h_ih_j = \sum_{i=1}^n a_{ii}h_i^2 + \sum_{\substack{i,j=1\\i\neq j}}^n a_{ij}h_ih_j = \sum_{i=1}^n a_{ii}h_i^2.$$

Je zjevné, že znaménko $\vec{h}A\vec{h}^T$ poznáme podle znamének a_{ii} . Pokud jsou všechna kladná, pak $\vec{h}A\vec{h}^T>0$ pro libovolný (nenulový) vektor. Pokud jsou všechna záporná, pak vždy $\vec{h}A\vec{h}^T<0$. Pokud by byla mezi a_{ii} znaménka kladná i záporná, tak máme indefinitní matici.

Pokud matice A není diagonální, ale je symetrická, pak se dle teorie matic dá v součinu $\vec{h}A\vec{h}^T$ nahradit diagonální maticí D, která má na diagonále vlastní čísla matice A. Dostáváme tak následující tvrzení.

Věta.

Nechť A je symetrická matice $n \times n$ s vlastními čísly $\lambda_1, \ldots, \lambda_n$. Jestliže $\lambda_i > 0$ pro všechna $i = 1, \dots, n$, pak je A positivně definitní. Jestliže $\lambda_i < 0$ pro všechna $i = 1, \dots, n$, pak je A negativně definitní.

Hledání vlastních čísel je ale obtížné. Mnohem snadnější je následující kritérium.

Věta. (Sylvestrovo kritérium)

Nechť
$$A = (a_{ij})$$
 je symetrická matice $n \times n$. Uvažujme její subdeterminanty $\Delta_1 = a_{11}, \ \Delta_2 = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, atd. až $\Delta_n = \det(A)$.

Jestliže $\Delta_i > 0$ pro všechna i, pak je A positivně definitní.

Jestliže $\Delta_1 < 0, \, \Delta_2 > 0, \, \Delta_3 < 0$ atd. až $(-1)^n \Delta_n > 0$, pak je A negativně definitní.

Jestliže $\Delta_2 < 0$, pak existují $\vec{x}, \vec{y} \in \mathbb{R}^n$ takové, že $\vec{x} A \vec{x}^T > 0$ a $\vec{y} A \vec{y}^T < 0$.

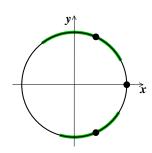
Protože je Hessova matice hladkých funkcí symetrická, můžeme tento poznatek na ni aplikovat a dostaneme Sylvestrovo kritérium, které jsme poznali v kapitole 4.

7e. Implicitní křivky a funkce

Mnohé útvary jsou dány rovnicemi, například jistá kružnice v rovině je dána rovnicí $x^2+y^2=169$. Jak bychom se o takových útvarech něco dozvěděli?

Jedna možnost je vnímat útvar daný rovnicí f(x, y, ...) = c jako hladinu konstantnosti funkce f. Získáme tak užitečné informace. Pokud například vezmeme libovolný bod na této křivce, tak získáme vektor kolmý k dané křivce v tomto bodě jednoduše jako ∇f . Pomocí tohoto vektoru se pak dají snadno napsat tečná přímka, rovina atd. (podle počtu dimenzí). Jsou ovšem informace, které se takto nezískají snadno. Proto se používá ještě jiný pohled na situaci.

Často se stává, že určité části takového útvaru lze interpretovat i jako grafy vhodných funkcí. Například pokud uvažujeme tu kružnici $x^2 + y^2 = 169$, tak okolo bodu (5,12) odpovídá ona kružnice grafu funkce $y = \sqrt{169 - x^2}$, prostě jsme danou rovnici vyřešili pro y. Takové funkci říkáme **implicitní** funkce. Výhodou je, že pak můžeme na příslušnou část křivky aplikovat standardní metody analýzy, např. hledání lokálních extrémů, konvexita atd. Nevýhodou tohoto přístupu je, že ne vždy je možné takto popsat celou danou křivku. Vidíme to právě v tomto příkladě, celou kružnici rozhodně nedokážeme vidět jako graf jedné funkce.



V takovém případě se obvykle postupuje lokálně, například na okolí bodu (5, -12) je už kružnice totožná s grafem funkce $y=-\sqrt{169-x^2}$, což je jiná funkce než předtím. Proto se od začátku smíříme s tím, že budeme dané křivky převádět na grafy funkcí pouze lokálně.

Bohužel, ani tuto ambici nelze zcela naplnit. Jak vidíme z obrázku, nelze najít okolí bodu (13,0), na kterém by šlo část dotyčné kružnice vyjádřit jako graf funkce. Na libovolném okolí jsou totiž dva body kružnice se stejnou souřadnicí x, ale žádná funkce nemůže mít v jednom x dvě různé hodnoty.

Přesto je tento přístup přínosný. Pokud jej chceme aplikovat, tak potřebujeme vyloučit patologické situace. Příklad s kružnicí ukazuje, že může nastat problém, když má křivka v bodě svislou tečnu (nebo také ne, záleží na tom, jestli se v takovém bodě křivka začne vracet nad sebe). Bohužel mohou nastat i horší věci, množina řešení rovnice f(x, y) = c vůbec nemusí být křivka.

Příklad: Implicitní rovnice $xy - \sqrt{x^2y^2} = 0$ dává množinu všech bodů prvního a třetího kvadrantu.

Implicitní rovnice $(x^2-4)^2+(y-1)^2=0$ dává množinu dvou bodů, jmenovitě (2,1) a (-2,1). Implicitní rovnice $(|x|-1)^2+y^2=1$ dává "osmičku", dvě kružnice dotýkající se v jednom bodě. Konečně něco rozumného.

 \triangle

Soustředíme se teď na rovinné objekty, tedy rovnice o dvou proměnných x, y. Abychom si situaci formálně zjednodušili, budeme pracovat s rovnicemi typu f(x, y) = 0, na které se ty ostatní snadno převedou.

Dá se ukázat, že pokud vyloučíme případy se svislou (nebo neexistující) tečnou, tak už vyloučíme všechny patalogické jevy. Bohužel tím také vyloučíme některé slušně vychované křivky, například $x - y^3 = 0$, což je graf třetí odmocniny, ale jemnější vybírací nástroj nemáme.

Než se dostaneme k větě, ještě klíčová poznámka: jak poznáme, že nějaká funkce y = y(x) popisuje lokálně stejný objekt jako rovnice f(x, y) = 0?

Vraťme se k příkladu $x^2 + y^2 = 169$. To, že jsme si dokázali vyjádřit y, vlastně znamená, že jsme našli řešení této rovnice, pokud považujeme y za neznámou a x za parametr. Správnost řešení se standardně ověřuje zpětným dosazením:

$$x^{2} + \sqrt{169 - x^{2}}^{2} = x^{2} + 169 - x^{2} = 169.$$

Obecně tedy vzorec y = y(x) dosadíme do dané rovnice za y a očekáváme, že to dobře dopadne, tedy že f(x, y(x)) = 0 pro všechna x, se kterými pracujeme, tedy obvykle z nějakého okolí daného bodu. Protože máme zájem o křivku procházející daným bodem (x_0, y_0) , budeme také chtít $y(x_0) = y_0$.

Věta. (o implicitní funkci)

Nechť D je otevřená množina v \mathbb{R}^2 , $f \in C^1(D)$. Nechť $\vec{a} = (x_0, y_0) \in D$ splňuje $f(\vec{a}) = 0$.

Jestliže $\frac{\partial f}{\partial y}(\vec{a}) \neq 0$, pak existuje okolí U bodu x_0 a funkce y = y(x) na U taková, že $y(x_0) = y_0$ a f(x, y(x)) = 0 pro $x \in U$.

Navíc má y(x) spojitou derivaci na U a

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

Jinak řečeno, rovnice f(x,y)=0 se dá na okolí bodu \vec{a} vyřešit pro y v závislosti na x. Věta nezaručuje, že výslednou závislost y=y(x), tedy vzniklou funkci, půjde vyjádřit algebraickým vzorcem, ale už jen vědomost, že existuje, je někdy přínosná. Vzorec z věty pak nabízí možnost dostat se k derivaci této funkce, i když si ji třeba neumíme vyjádřit formulkou.

Nemusíme ten vzorec aplikovat formálně, protože se odvozuje způsobem, kterým můžeme derivaci rovnou hledat pro konkrétní funkci. Začneme tím, že si v rovnici f(x,y) = 0 představíme, že y je ve skutečnosti funkce s proměnnou x. Je to pak rovnice porovnávající dvě funkce s neznámou

x, přičemž na pravé straně je to konstantní funkce. Můžeme tedy obě strany zderivovat, na levé straně bude třeba použít řetízkové pravidlo:

$$f(x, y(x)) = 0$$
$$[f(x, y(x))]' = [0]'$$
$$\frac{\partial f}{\partial x}(x, y(x)) \cdot 1 + \frac{\partial f}{\partial y}(x, y(x)) \cdot y'(x) = 0.$$

Když si z této rovnice vyjádříme y'(x), dostáváme vzorec z věty.

Když se tento postup provádí s konkrétní funkcí, říkáme tomu implicitní derivování.

Příklad: Uvažujme implicitní křivku danou rovnicí $x^4 - 4xy + y^4 = 1$ (je to mimochodem šišoid s dlouhou osou na hlavní diagonále, ale většinou tvar křivky neznáme, tak se bez této informace zkusíme i zde obejít). Protože bod x = 0, y = 1 vyhovuje dané implicitní rovnici, leží na této křivce. Je daná množina na okolí tohoto bodu pěkná (je to vůbec křivka?) a můžeme ji vyjádřit jako graf funkce?

Potřebujeme ještě upravit rovnici, aby měla napravo nulu, takže použijeme větu s funkcí $f(x,y) = x^4 - 4xy + y^4 - 1$. Máme

$$\frac{\partial f}{\partial y}(0,1) = (-4x + 4y^3)\big|_{x=0,y=1} = 4 \neq 0,$$

takže podmínka je splněna. Proto na nějakém okolí bodu (0,1) máme křivku, která odpovídá grafu jisté neznámé funkce y(x). Bohužel, vzorec pro tuto funkci y(x) nejsme z dané rovnice schopni odvodit.

Jak vypadá tečna k jejímu grafu (neboli křivce) v daném bodě? Protože pro y(x) nemáme vzorec, nezbývá než najít potřebnou derivaci pomocí implicitního derivování. V dané rovnici (klidně té původní) tedy považujeme y za funkci proměnné x a derivujeme rovnici $x^4 - 4xy(x) + (y(x))^4 = 1$. Obvykle si lidé tu proměnnou u y jen představují, ale nepíšou.

$$[x^{4} - 4xy + y^{4}]' = [1]'$$

$$4x^{3} - 4y - 4xy' + 4y^{3}y' = 0$$

$$y' = \frac{4y - 4x^{3}}{4y^{3} - 4x}.$$

Nás zajímá situace v bodě (0,1), tedy pro x=0 a y=1, po dosazení máme $y'(0)=\frac{4}{4}=1$. Známe tedy bod i směrnici, tečna v daném bodě je $y-1=1\cdot(x-0)$ neboli y=x+1.

Alternativa: Jde o hladinu konstantnosti funkce $f(x,y) = x^4 - 4xy + y^4$, máme

$$\nabla f = (4x^3 - 4y, -4x + 4y^3) \implies \nabla f(0,1) = (-4,4).$$

Víme, že $\vec{u} = \nabla f(0,1) = (-4,4)$ je vektor kolmý na tuto hladinu, proto je kolmý i na tečnu. Přímka, která prochází bodem (0,1) a je kolmá na vektor \vec{u} je dána rovnicí

$$\vec{u} \bullet ((x,y) - (0,1)) = 0 \implies (-4,4) \bullet (x,y-1) = 0 \implies -4x + 4(y-1) = 0 \implies y = x+1.$$

Implicitní derivování pomůže často i v případě, kdy bychom byli schopni funkci y z rovnice vyjádřit, ale byla by komplikovaná.

Všimněte si, že vzorec pro derivaci, který jsme dostali implicitním derivováním,

$$y' = \frac{4y - 4x^3}{4y^3 - 4x},$$

zahrnuje x i y, tedy jak volnou proměnnou, tak funkci, kterou se snažíme zkoumat. To je typické. Pokud bychom znali vzorec pro y = y(x), tak bychom jej mohli dosadit a dostali obvyklý vzoreček pro derivaci používající x (ovšem kdybychom znali vzorec pro y(x), tak derivujeme rovnou). My

ale vzorec pro y(x) neznáme, takže nám je náš derivační vzorec k užitku pouze ve známých bodech (x,y) implicitní křivky.

Jak by toto vypadalo ve více dimenzích? V dané implicitní rovnici pak máme více proměnných, například rovnice 2x+3y+z=5 definuje rovinu v \mathbb{R}^3 . V tomto případě z rovnice snadno vyjádříme třeba z, takže celou rovinu vyjádříme jako graf funkce z(x,y)=5-2x-3y.

Obecně, pokud máme jednu rovnici s více proměnnými, tak můžeme doufat, že se nám podaří (přinejmenším lokálně) jednu proměnnou vyjádřit. To odpovídá situaci, kdy jsme si kousek útvaru vytvořeného tou rovnicí vyjádřili jako graf funkce, jako jsme to udělali s rovnicí roviny.

Formálně se hodí si tu jednu speciální proměnnou oddělit značením. Pracujeme tedy s rovnicí ve tvaru $f(x_1, \ldots, x_n, y) = 0$, přičemž doufáme, že se nám z ní podaří vyjádřit y v závislosti na ostatních proměnných. Jinak řečeno, doufáme, že vznikne funkce $y = y(x_1, \ldots, x_n)$, jejíž graf odpovídá (lokálně) tomu, co vytvořila rovnice. Opět to lze ověřit dosazením, tedy podmínka je

$$f(x_1,\ldots,x_n,y(x_1,\ldots,x_n))=0.$$

Jak to vypadá v případě, že máme rovnic více? Představme si typickou sadu rovnic popisujících přímku v \mathbb{R}^3 , třeba

$$4x + 2y + z = 6$$
$$x + y + z = 4.$$

Z takovéto soustavy dokážeme eliminovat jednu proměnnou, takže dostaneme například (odečtením druhé od první) y=2-3x. Ale můžeme také odečíst první rovnici od dvojnásobku druhé a dostaneme z=2+2x. Získali jsme vzorce, které dávají y,z v závislosti na x. Není těžké zjistit, že množina bodů (x,2-3x,2+2x) je totožná s přímkou danou původními rovnicemi. Teď potřebujeme naše pozorování správně interpretovat.

Na odvozené závislosti lze nahlížet jako na vektorovou funkci

$$x \mapsto (2 - 3x, 2 + 2x) = (y(x), z(x)).$$

Množina $\{(x, y(x), z(x))\}$ pak představuje graf této funkce. Ukázali jsme, že z původních rovnic jsme dokázali odvodit vektorovou funkci, jejíž graf souhlasí s objektem definovaným původními rovnicemi.

Toto je správný pohled na zobecnění implicitních křivek a implicitního derivování. Pokud máme m rovnic, tak je při troše štěstí možné z nich odvodit vzorce pro m vybraných proměnných (značených y_j), které už budou záviset jen na těch ostatních proměnných (značených x_i). Lze to zapsat i vektorově, pak pracujeme s \vec{y} a \vec{x} .

Dokonce i těch m rovnic lze zapsat jako jednu rovnici, ale s vektorovou funkcí nalevo a vektorem funkcí napravo. Například ty dvě rovnice výše lze přepsat jako

$$\begin{pmatrix} 4x + 2y + z - 6 \\ x + y + z - 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Vznikne tak velmi kompaktní formulace.

Věta. (o implicitní funkci)

Nechť $F(\vec{x}, \vec{y}): D \mapsto \mathbb{R}^m$ je vektorová funkce, kde $D \subseteq \mathbb{R}^n \times \mathbb{R}^m$. Předpokládejme, že pro jistý bod $(\vec{a}, \vec{b}) \in D$ platí $F(\vec{a}, \vec{b}) = \vec{0}_m$ a F má spojité parciální derivace na nějakém okolí (\vec{a}, \vec{b}) .

Uvažujme matice $m \times m$ a $m \times n$

$$J_{F,\vec{y}} = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}, \qquad J_{F,\vec{x}} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \cdots & \frac{\partial F_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1} & \cdots & \frac{\partial F_m}{\partial x_n} \end{pmatrix}.$$

Jestliže je $J_{F,\vec{y}}(\vec{a},\vec{b})$ regulární, tak existuje okolí U bodu \vec{a} a funkce $Y(\vec{x}) \in [C^1(U)]^m$ taková, že $Y(\vec{a}) = \vec{b}$ a $F(\vec{x},Y(\vec{x})) = \vec{0}_m$ na U. Tato funkce je jednoznačně určena a její Jacobiho matice je dána jako

$$J_Y(\vec{x}) = -[J_{F,\vec{y}}(\vec{x}, Y(\vec{x}))]^{-1} \cdot J_{F,\vec{x}}(\vec{x}, Y(\vec{x})).$$

Přeloženo do méně formálního jazyka, sadu rovnic

$$F_1(x_1, \dots, x_n, y_1, \dots, y_m) = 0$$

$$\vdots$$

$$F_m(x_1, \dots, x_n, y_1, \dots, y_m) = 0$$

dokážeme lokálně upravit na vzorce

$$y_1 = y_1(x_1, \dots, x_n)$$

$$\vdots$$

$$y_m = y_m(x_1, \dots, x_n),$$

které popisují stejný objekt jako je množina řešení původních rovnic. Jako obvykle není zaručeno, že ty vzorce musejí mít formu vzorců algebraických, mohou to být abstraktní funkce.

Jejich parciální derivace lze získat implicitním derivováním. Když chceme parciální derivaci vzhledem k nějakému x_i , tak prostě vezmeme původní rovnice, představíme si, že y_1 až y_m jsou funkce x_i , a aplikujeme parciální derivaci vzhledem k x_i na levé a pravé strany rovnic. Výslednou soustavu pak vyřešíme pro $\frac{\partial y_j}{\partial x_i}$.

Příklad: Uvažujme rovnice

$$s - e^{t}v^{2} - u + 2v^{2} = -1$$
$$3s - e^{t}v^{2} + u - 2v^{3} = -3$$
$$3vs + e^{t}v - u + v^{2} - 2v^{3} = 0.$$

Definují nějaký objekt v \mathbb{R}^4 , tak se rovnou vzdáme naděje, že bychom si to uměli představit. Zato si všimneme, že pokud jsou ty rovnice nezávislé (a vypadají na to), tak každá odebere jeden stupeň volnosti z původních čtyř, takže by ten objekt měl být jednodimenzionální (pokud to není jeden z těch patalogických) a můžeme mu říkat křivka.

Snadno ověříme, že bod (3,0,8,2) tyto rovnice splňuje, takže na té křivce leží. Je kolem něj křivka rozumná? Věta nás vyzývá ke zkoumání jisté matice, na to si nejprve zavedeme příslušné značení:

$$\begin{pmatrix} s - e^t v^2 - u + 2v^2 \\ 3s - e^t v^2 + u - 2v^3 \\ 3vs + e^t v - u + v^2 - 2v^3 \end{pmatrix} = \begin{pmatrix} -1 \\ -3 \\ 0 \end{pmatrix}$$

$$\implies F(s, t, u, v) = (s - e^t v^2 - u + 2v^2, 3s - e^t v^2 + u - 2v^3, 3vs + e^t v - u + v^2 - 2v^3).$$

Cílový prostor je třídimenzionální, takže pokud budou splněny předpoklady, měli bychom být

schopni (teoreticky) vyjádřit tři z proměnných jako funkce té čtvrté. Při pohledu na rovnice se mi nechce hledat vzorec pro v, protože se tam vyskytuje v různých mocninách, takže jsem se rozhodl z něj udělat volnou proměnnou. Jinými slovy, soustředíme se teď na na možnost mít funkce s=s(v), t=t(v), u=u(v). Pokud se vám má volba nelíbí, tak si udělejte vlastní a proveďte příslušné výpočty, já na to nemám dost odvahy.

Věta po nás chce zkoumat jistou matici.

$$J_{F,(s,t,u)} = \begin{pmatrix} \frac{\partial F_1}{\partial s} & \frac{\partial F_1}{\partial t} & \frac{\partial F_1}{\partial u} \\ \frac{\partial F_2}{\partial s} & \frac{\partial F_2}{\partial t} & \frac{\partial F_2}{\partial u} \\ \frac{\partial F_3}{\partial s} & \frac{\partial F_3}{\partial t} & \frac{\partial F_3}{\partial u} \end{pmatrix} = \begin{pmatrix} 1 & -e^t v^2 & -1 \\ 3 & -e^t v^2 & 1 \\ 3v & e^t v & -1 \end{pmatrix}.$$

Potřebujeme najít její determinant v daném bodě.

$$\det(J_{F,(s,t,u)}(3,0,8,2)) = \det\begin{pmatrix} 1 & -4 & -1 \\ 3 & -4 & 1 \\ 6 & 2 & -1 \end{pmatrix} = -64.$$

Determinant není nula, takže v bodě (3,0,8,2) není křivka patologická a musí existovat nějaké okolí tohoto bodu, kde se tato křivka dá vnímat jako graf jisté vektorové funkce. Nás spíše zajímá interpretace, že je tedy (teoreticky) možné z těch tří rovnic získat vztahy $s=s(v),\ t=t(v),\ u=u(v)$. Dokážeme je najít?

Budeme brát v jako parametr. Odečtením první rovnice od druhé a přičtením v-násobku třetí k druhé rovnici dostaneme

$$2s + 2u - 2v^{3} - 2v^{2} = -2 \implies s + u = v^{3} + v^{2} - 1$$
$$3s(1+v^{2}) + u(1-v) - v^{3} - 2v^{4} = -3 \implies s(3v^{2} + 3) + u(1-v) = 2v^{4} + v^{3} - 3.$$

Pokud tu první vynásobíme členem 1-v a odečteme od druhé, získáme

$$s(3v^2 + v + 2) = 3v^4 + v^3 - v^2 - v - 2.$$

Proto

$$s = \frac{3v^4 + v^3 - v^2 - v - 2}{3v^2 + v + 2} = v^2 - 1.$$

Dosazením do první rovnice hned nad námi dostaneme

$$u = (v^3 + v^2 - 1) - (v^2 - 1) = v^3.$$

Nakonec dosadíme s a u do první původní rovnice a vyjde

$$e^{t}v^{2} = 1 + (v^{2} - 1) - v^{3} + 2v^{2} = 3v^{2} - v^{3} \implies t = \ln(3 - v).$$

Dostali jsme vzorce

$$s(v) = v^2 - 1,$$

 $t(v) = \ln(3 - v),$
 $u(v) = v^3,$

které popisují danou křivku na nějakém okolí bodu (3,0,8,2). Vektorová funkce z věty je

$$G(v) = (v^2 - 1, \ln(3 - v), v^3).$$

Měli jsme štěstí a vzorce jsme našli, takže snadno najdeme derivace

$$s'(v) = 2v,$$

 $t'(v) = \frac{-1}{3-v},$
 $u'(v) = 3v^2,$

což má platit na nějakém okolí bodu (3,0,8,2) na křivce, tedy pro v blízko k 2. Například s'(2)=4.

Kdybychom toto štěstí neměli, tak by se pro nalezení derivace nabízely dvě možnosti. Jedna z

nich je použít vzorec z věty. Na to potřebujeme najít invezní matici k $J_{F,(s,t,u)}$, což je

$$J_{F,(s,t,u)}^{-1} = \begin{pmatrix} -\frac{1}{2} \frac{v-1}{3v^2+v+2} & \frac{1}{2} \frac{v+1}{3v^2+v+2} & \frac{v}{3v^2+v+2} \\ -\frac{3}{2} \frac{1}{e^t v} \frac{v+1}{3v^2+v+2} & -\frac{1}{2} \frac{1}{e^t v} \frac{3v-1}{3v^2+v+2} & \frac{1}{e^t v} \frac{2}{3v^2+v+2} \\ -\frac{3}{2} \frac{v^2+1}{3v^2+v+2} & \frac{1}{2} \frac{3v^2+1}{3v^2+v+2} & -\frac{v}{3v^2+v+2} \end{pmatrix}.$$

Našli jsme ji pomocí Gaussovy eliminace a triku $(A|E_n) \mapsto (E_n|A^{-1})$ a už z výsledku se dá usoudit, že to nebylo moc příjemné. Tuto matici pak usíme vynásobit maticí

$$J_{F,v} = \begin{pmatrix} \frac{\partial F_1}{\partial v} \\ \frac{\partial F_2}{\partial v} \\ \frac{\partial F_3}{\partial v} \end{pmatrix} = \begin{pmatrix} -2e^t v + 4v \\ -2e^t v - 6v^2 \\ 3s + e^t + 2v - 6v^2 \end{pmatrix}.$$

Podle vzorce pak u výsledku změníme znaménko a dostaneme

$$\begin{pmatrix} s' \\ t' \\ u' \end{pmatrix} = \begin{pmatrix} v \frac{9v^2 + 3v - 2 + e^t - 3s}{3v^2 + v + 2} \\ -\frac{1}{e^t v} \frac{9v^3 - 21v^2 - 2v + 6e^t v^2 + 2e^t v + 2e^t + 6s}{3v^2 + v + 2} \\ -v \frac{-9v^3 - 5v - 6 + e^t - 3s}{3v^2 + v + 2} \end{pmatrix}$$

Obecně bychom tento vzorec mohli používat pouze ve známých bodech křivky. My známe bod (3,0,8,2), takže dosazením $s=3,\,t=0,\,u=8,\,v=2$ snadno získáme třeba

$$s'(2) = 2\frac{36+6-2+1-9}{12+2+2} = 4.$$

To se shoduje s výsledkem přímého výpočtu výše. Protože my zde vlastně známe vzorce pro s(v) a t(v), můžeme je do vzorce pro s' dosadit a dostaneme

$$s' = v \frac{9v^2 + 3v - 2 + e^{\ln(3-v)} - 3(v^2 - 1)}{3v^2 + v + 2} = v \frac{6v^2 + 2v + 4}{3v^2 + v + 2} = 2v.$$

Vypadá to tedy, že ten vzorec s Jacobiho maticemi opravdu dává správnou informaci.

Ovšem ruční výpočet inverzní matice byl poněkud výživný, takže stojí za to zkusit ten druhý přístup, jmenovitě implicitní derivování. Nejprve derivujeme dané rovnice, maje přitom na paměti, že s,t,u jsou teď vlastně funkce s proměnnou v.

$$s' - e^{t}t'v^{2} - e^{t}2v - u' + 4v = 0$$
$$3s' - e^{t}t'v^{2} - e^{t}2v + u' - 6v^{2} = 0$$
$$3s + 3vs' + e^{t}t'v + e^{t} - u' + 2v - 6v^{2} = 0.$$

Toto se dá vyřešit pro s', t', u'. Všimněte si, že teď bereme i s, t, u jako parametry, takže taková soustava bude vždy lineární.

$$s' - e^{t}v^{2}t' - u' = e^{t}2v - 4v$$
$$3s' - e^{t}v^{2}t' + u' = e^{t}2v + 6v^{2}$$
$$3vs' + e^{t}vt' - u' + 2v = -3s - e^{t} + 6v^{2}.$$

Můžeme například použít Kramerovo pravidlo a dostaneme

$$s' = \frac{-2e^{t}v^{2}(e^{t} + 9v^{2} + 3v - 2 - 3s)}{-2e^{t}v(3v^{2} + v + 2)} = v\frac{e^{t} + 9v^{2} + 3v - 2 - 3s}{3v^{2} + v + 2}.$$

Je to stejný vzorec, jaký jsme dostali formálním výpočtem přes Jacobiho matice, a zdálo se to o něco příjemnější. Čtenář si může podobně ověřit, že také vzorce pro t' a u' získané tímto způsobem souhlasí.

 \triangle

8. Více o integrálu

V této kapitole zavedeme integrál pro funkce více proměnných. Možných přístupů je více, my zde budeme následovat definici Riemannova integrálu pro funkce jedné proměnné na intervalech. Její podstatou je dělení intervalu, kdy si daný interval rozdělíme na úsečky a pak sčítáme obsahy obdélníků, které jsme nad nimi vztyčili pomocí funkce f. To je zabaleno do mechanismu, který nás motivuje brát jemnější dělení a tak získávat přesnější aproximace integrálu.

Pro úspěch tohoto přístupu je podstatné, že úsečky umíme změřit (pomocí jejich délky pak počítáme obsahy). Relativní jednoduchost je pak dána tím, že intervaly umíme pohodlně rozřezat na menší úsečky.

8a. Množiny

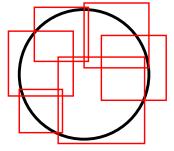
Nyní tyto myšlenky aplikujeme na integrál funkce n proměnných přes množinu Ω z \mathbb{R}^n pro $n \in \mathbb{N}$. Rádi bychom integrační množiny dělili na nějaké "základní tvary" v \mathbb{R}^n , tak jako jsme v jednom rozměru dělili integrační interval na úsečky. Už pro n=2 ale narazíme na problém. Určitě bychom rádi integrovali přes kruhy, ale kruh nedokážeme rozřezat na kousky, jejichž tvary by byly jednoduše zpracovatelné (čtverce, kolečka, trojúhelníky, . . .).

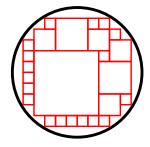
Je dobré si uvědomit, že problém není ve vyšší dimenzi, ale ve vyšší ambici. V jednom rozměru jsme se spokojili s integrováním přes intervaly. Pokud bychom se pro funkce dvou proměnných spokojili s integrováním přes obdélníky, tak vesele řežeme na menší obdélníčky. Ale my jsme neskromní a chceme integrovat i přes jiné tvary, pak se nemůžeme divit, že máme problém.

Pokud jej chceme vyřešit, musíme se blíže podívat na problematiku dělení množin na menší kousky. Vlastně nás v té chvíli nezajímá funkce, kterou integrujeme, je to otázka práce s množinami v \mathbb{R}^n . Jak záhy uvidíme, vlastně řešíme otázku, jak množiny změřit, tedy jak najít jejich n-rozměrnou analogii obsahu a objemu.

Klíčová myšlenka je, že už nebudeme množiny řezat, ale budeme je aproximovat pomocí množin příznivého typu. Z praktického pohledu jsou dva rozumné přístupy.

První možnost je, že pro nás bude hlavní, abychom z množiny nic nevynechali. Toto se velmi snadno zapíše matematicky. Na obrázku níže vlevo vidíme, jak pomocí čtverců aproximujeme kruh. Je zjevné, že plýtváme. To nevadí, protože podobně jako u Riemannova integrálu vznikne mechanismus, který nás bude motivovat k co nejmenšímu plýtvání.





Na obrázku vpravo vidíme druhý přístup k tomu, jak pomocí čtverců aproximovat kruh. Zde je naopak motivací zabránit plýtvání a nějak nás nemrzí, že jsme množinu nepokryli. Jak zabráníme plýtání? Pokud bychom požadovali, aby byly zúčastněné množiny disjunktní, tak by nás to v mnoha případech zbytečně omezovalo. Praktičtější je dovolit, aby se základní tvary mohly překrývat svými hranicemi. To se dá elegantně zachytit požadavkem, že by se množinám neměly překrývat vnitřky.

Definice.

Uvažujme množinu $\Omega \subseteq \mathbb{R}^n$ a konečnou kolekci množin $\Omega_1, \ldots, \Omega_m \subseteq \mathbb{R}^n$.

Řekneme, že $\{\Omega_k\}$ je **pokrytí** množiny Ω , jestliže $\Omega\subseteq\bigcup\limits_{k=1}^m\Omega_k$.

Řekneme, že $\{\Omega_k\}$ je **vyplnění** množiny Ω , jestliže $\Omega_k \subseteq \Omega$ pro všechna k a $\Omega_k^O \cap \Omega_l^O = \emptyset$ pro všechna $k \neq l \in \{1, \ldots, m\}$.

Zde bychom měli čtenáře varovat, že zatímco "pokrytí" je oficiální název, "vyplnění" jsme si tu vymysleli, abychom se o tom mohli bavit intuitivně. V teorii, jak se běžně zavádí, se nutnost tu situaci nějak pojmenovat obchází různými triky.

Z toho, jak definujeme pokrytí a vyplnění, by mělo platit následující: Jestliže bychom uměli všechny zúčastněné množiny "změřit", tedy měli bychom něco jako n-rozměrnou obdobu objemu, tak by pro všechna pokrytí $\{\Omega_k\}$ mělo platit

$$\operatorname{vol}_n(\Omega) \le \sum_{k=1}^m \operatorname{vol}_n(\Omega_k)$$

a pro všechna vyplnění $\{\Omega_k\}$ pak

$$\operatorname{vol}_n(\Omega) \ge \sum_{k=1}^m \operatorname{vol}_n(\Omega_k).$$

Jinak řečeno, zatím tyto definice dávají smysl, budeme se pomocí kvalitnějších aproximací množiny snažit shora a zdola přiblížit k jejímu "objemu".

Teď přichází okamžik, kdy bychom se měli rozhodnout, jaký základní tvar budeme používat. Hlavní požadavek je, že by měl být přirozeně zobecnitelný pro všechny dimenze a měli bychom umět spočítat jeho objem způsobem, který je natolik intuitivní, že by nám měl přijít rozumný i pro vyšší dimenze, kde vlastně zatím nevíme, co je to "objem".

To ukazuje na vícerozměrné intervaly, tedy obdélníky, kvádry atd. Při jednorozměrné integraci jsme pracovali s úsečkami $\langle a,b\rangle$, délka je b-a. Je pak přirozené pracovat ve dvou dimenzích s obdélníky $\langle a_1,b_1\rangle \times \langle a_2,b_2\rangle$ s obsahem $(b_1-a_1)\cdot (b_2-a_2)$. Ve třech rozměrech přejdeme ke kvádrům $\langle a_1,b_1\rangle \times \langle a_2,b_2\rangle \times \langle a_3,b_3\rangle$ s objemem $(b_1-a_1)\cdot (b_2-a_2)\cdot (b_3-a_3)$ a je snadné se nechat přemluvit, že má smysl tímto způsobem pokračovat.

Definice.

Nechť $n \in \mathbb{N}$. Pojem n-obdélník označuje libovolnou množinu typu

$$\langle a_1, b_1 \rangle \times \langle a_2, b_2 \rangle \times \cdots \times \langle a_n, b_n \rangle$$

pro nějaká reálná čísla $a_k \leq b_k$ pro $k = 1, \ldots, n$.

Pro daný n-obdélník $\Omega = \langle a_1, b_1 \rangle \times \langle a_2, b_2 \rangle \times \cdots \times \langle a_n, b_n \rangle$ definujeme jeho **míru** jako

$$m(\Omega) = (b_1 - a_1) \cdot (b_2 - a_2) \cdot \cdot \cdot (b_n - a_n) = \prod_{i=1}^n (b_i - a_i).$$

Míra je matematický způsob, jak neříct "n-rozměrný objem", ale myslet si to.

Intuice nám říká, že délka intervalu nezáleží na tom, jestli je otevřený, polouzavřený nebo uzavřený. Podobně je to u n-obdélníků. Někteří autoři preferují v definici otevřené intervaly, mnozí používají intervaly $\langle a_i, b_i \rangle$, snad kvůli tomu, že na sebe pěkně navazují. Dá se ukázat, že to nakonec vyjde nastejno. My jsme zde zvolili uzavřené intervaly, protože při integrování se množiny obvykle berou jako uzavřené.

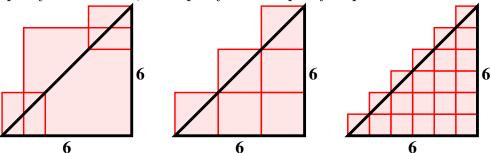
Budeme tedy množiny pokrývat a vyplňovat n-obdélníky, pro zkrácení tomu někdy budeme říkat obdélníkové pokrytí a obdélníkové vyplnění. Nejprve si rozmyslíme, že abychom mohli množinu pokrýt obdélníky, tak musí být omezená, protože obdélníky jsou omezené a definice pokrytí neumožňuje vzít jich nekonečně mnoho. V této kapitole se tedy omezíme na omezené množiny Ω a ty už pak jdou pokrýt všechny. Každá omezená množina je totiž z definice obsažena v n-rozměrné kouli o jistém poloměru r a ta je zase obsažena v n-rozměrné krychli o straně 2r.

Jak je to s vyplněním? Tady nám pomůže, že v definici n-obdélníka se připouští možnost, že $a_i = b_i$. To znamená, že i jeden bod je považován za n-obdélník, jeho míra je zjevně nula. Každá

neprázdná množina tudíž v sobě obsahuje nějaký n-obdélník, tedy má nějaké obdélníkové vyplnění. V této kapitole proto také budou všechny množiny neprázdné. Shrnuto, každá neprázdná omezená množina má nějaké obdélníkové pokrytí i vyplnění.

Nyní jsme připraveni pokusit se co nejlépe aproximovat tvar dané množiny pomocí n-obdélníků. Zavedeme mechanismus inspirovaný Darbouxovým přístupem k Riemannovu integrálu, který nás bude motivovat k co nejlepším aproximacím množiny. Máme dvě možnosti, jak se k tvaru množiny Ω přiblížit.

Můžeme zkoušet různá pokrytí pomocí n-obdélníků. Jak poznáme, která pokrytí jsou lepší a která horší? Každé takové pokrytí $\{\Omega_k\}$ poskytuje jistý odhad n-rozměrného objemu množiny Ω , v našem případě ve tvaru $\sum m(\Omega_k)$. Jak jsme již poznamenali výše, tento odhad (pro rozumně zavedený pojem objemu) nemůže být menší než skutečný objem množiny Ω (pokud něco takového existuje), ale zato může být větší, a to tím větší, čím víc se plýtvá. Neboli naopak, čí menší odhad pomocí obdélníkového pokrytí dostáváme, tím lépe by mělo to pokrytí aproximovat danou množinu.



Na obrázku vidíme několik pokrytí trojúhelníka o obsahu 18 pomocí 2-obdélníků. V levém pokrytí jsme použili velký čtverec o straně 5 a dva menší o straně 2, takže jsme dostali horní odhad

$$5^2 + 2 \cdot 2^2 = 33 = 18 + 15.$$

To jsme docela plýtvali. V dalším pokrytí už jsme se poučili, nepřekrývali jsme čtverce o straně 2 a horní odhad je

$$6 \cdot 2^2 = 24 = 18 + 6$$
.

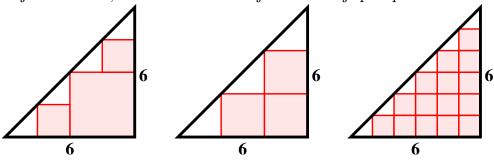
Jsme blíže. U třetího pokrytí jsme použili čtverce o straně 1 a dostali odhad

$$21 \cdot 1^2 = 21 = 18 + 3.$$

Vidíme, že se součty opravdu přibližují ke správné hodnotě 18 s tím, jak pokrytí stále lépe vystihují tvar trojúhelníka. Zdá se také, že bychom se mohli přiblížit libovolně blízko volbou stále menších čtverců.

Pro obecnou množinu Ω tedy má smysl hledat nejmenší možné číslo, ke kterému je možné se pomocí pokrytí dostat. U některých množin Ω přímo existuje konkrétní pokrytí, které dá jeho "objem", ale pro jiné (třeba ten trojúhelník) se dokážeme jen libovolně blízko přiblížit, takže je potřeba místo minima vzít infimum možných odhadů.

Druhý možný přístup je naopak množinu Ω vyplňovat. Pro každé vyplnění pomocí n-obdélníků pak $\sum m(\Omega_k)$ představuje dolní odhad n-dimenzionálního objemu. Když povzbudíme vyplnění k poskytnutí co největšího čísla, tak tím zase motivujeme k co nejlepší aproximaci tvaru množiny Ω .



Na obrázku vlevo je nepravidelné vyplnění čtverci, které dává dolní odhad

$$3^2 + 2 \cdot 1.5^2 = 13.5 = 18 - 4.5$$
,

zatímco další pravidelná vyplnění dávají dolní odhady 18-6 a 18-3. Opět by si člověk tipnul, že zjemňováním čtverců by se uměl dostat libovolně blízko k 18.

To nás přivádí k následující definici.

Definice.

Nechť Ω je neprázdná omezená množina v \mathbb{R}^n .

Definujeme pro ni vnější míru

$$\mathbf{m}^*(\Omega) = \inf\Bigl\{\sum_k \mathbf{m}(\Omega_k); \ \{\Omega_k\} \text{ obdélníkové pokrytí } \Omega\Bigr\}$$

a vnitřní míru

miru
$$\mathrm{m}_*(\Omega) = \sup\Bigl\{\sum_k \mathrm{m}(\Omega_k); \ \{\Omega_k\} \text{ obdélníkové vyplnění } \Omega\Bigr\}.$$

Řekneme, že množina Ω je Jordanovsky měřitelná, jestliže $m^*(\Omega) = m_*(\Omega)$. Pak definujeme její (Peanovu-Jordanovu) míru jako $m(\Omega) = m^*(\Omega)$.

Myšlenka je podobná jako u definice Riemannova integrálu. Tam jsme také odhadovali shora a zdola a pokud se odhady sešly, tak jsme prohlásili, že funkce je dostatečně rozumná. Teď to máme podobně s tvarem množin. Klíčová dobrá zpráva je, že ve dvou a třech dimenzích Jordanova míra odpovídá obvyklému obsahu a objemu.

Které množiny jsou měřitelné? Všechny běžné množiny. Například pokud vezmeme spojitou funkci na intervalu, tak oblast pod grafem (coby množina v \mathbb{R}^2) bude určitě měřitelná. Stejně tak jsou měřitelné množiny "mezi grafy", jak jsme o nich mluvili v kapitole 6. Přejdeme-li do více dimenzí, tak pokud máme měřitelnou množinu D a na ní dvě spojité funkce, tak je množina "mezi grafy" těchto dvou funkcí (která má o jednu dimenzi víc než D) také měřitelná.

Neměřitelné množiny musí mít velmi divokou hranici, obvykle všelijak potrhanou, takže je v běžných aplikacích nepotkáváme.

Příklad: Ukážeme neměřitelnou množinu v \mathbb{R} . Připomeňme že pak n-obdélníky odpovídají úsečkám.

Uvažujme množinu $\Omega = \mathbb{Q} \cap \langle 0, 1 \rangle$. Jde tedy o množinu všech racionálních čísel mezi 0 a 1. Racionální čísla mají dvě důležité topologické vlastnosti.

Tou první je, že racionální čísla jsou hustá, tedy v každém úseku reálné osy kladné délky je nějaké racionální číslo. To znamená, že pokud množinu Ω pokryjeme úsečkami, tak ty úsečky na sebe musejí navazovat bez mezer, protože kdyby se mezi nimi objevila mezera (a byla mezi 0 a 1), tak už je v té mezeře racionální číslo a vlastně to není pokrytí. Sjednocení takového pokrytí proto musí obsahovat interval $\langle 0, 1 \rangle$. Pokud posčítáme délky úseček z takového pokrytí, tak musí být nejméně 1. Infimum získané z pokrytí tedy nemůže být menší než jedna. A nemůže být ani větší, protože interval $\langle 0, 1 \rangle$ o délce 1 je sám o sobě je pokrytí Ω . Proto m*(Ω) = 1.

Druhou významnou vlastností je, že i iracionální čísla jsou hustá, jinak řečeno, neexistuje žádný souvislý úsek reálné osy, který by se skládal čistě z racionálních čísel. To znamená, že Ω nemůže obsahovat žádnou podmnožinu ve tvaru intervalu kladné délky. Každé obdélníkové vyplnění se tedy nutně musí skládat z jednobodových intervalů, jejichž míra je nulová a nulový je pak i součet měr. Právě jsme ukázali, že všechny odhady pomocí vyplnění jsou nula, takže i jejich supremum musí být nula. Matematicky: $m_*(\Omega) = 0$.

Ukázali jsme, že $m_*(\Omega) < m^*(\Omega)$, takže množina Ω není měřitelná.

Tento příklad se hravě zobecní i pro vyšší dimenze, například pro \mathbb{R}^2 můžeme vzít všechny body s racionálními souřadnicemi ze čtverce $\langle 0, 1 \rangle \times \langle 0, 1 \rangle$.

Δ

Odborná poznámka: Existují i jiné typy měr než Jordanova. V matematické analýze je asi nejpoužívanější míra Lebesgueova, která používá nekonečná pokrytí. Je výrazně mocnější než Jordanova a umí změřit i dosti divoké množiny (například \mathbb{Q}), ale její zavedení dá dost práce.

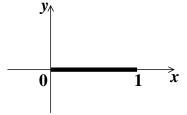
Popravdě řečeno, i naše zavedení Jordanovy míry by ještě potřebovalo podepřít, například vůbec není jasné, jestli to infimum a supremum v definici existují konečná. Zde je klíčové lemma, které je velmi podobné klíčovému lemmatu při zavádění Riemannova integrálu: Nemůže se stát, že by nějaké pokrytí množiny poskytlo menší odhad její míry než její vyplnění. I důkaz je podobný jako pro Riemannův integrál, jmenovitě přes zjemnění.

Poznámka: V kapitole 6 jsme se mimo jiné zmínili o zajímavé kategorii množin, které formálně patří do n-rozměrného prostoru, ale jejich podstata je méně dimenzionální. Pokud jsou měřitelné, tak by měly mít míru nula.

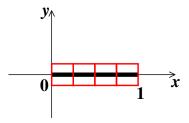
Podívejme se na případ úsečky v \mathbb{R}^2 . Pro jednoduchost si vybereme interval (0,1) na ose x, který ale teď vidíme jako množinu

$$\Omega = \{(x,0); \ 0 \le x \le 1\}.$$

Je jasné, že jediné 2-obdélníky (neboli obdélníky), které se do Ω vejdou, jsou obdélníky typu $\langle 0,0\rangle \times \langle a_2,b_2\rangle$. Protože 0-0=0, mají všechny míru nula. Proto pro libovolné obdélníkové vyplnění $\{\Omega_k\}$ musí platit $\sum m(\Omega_k) = \sum 0 = 0$. To dokazuje, že $m_*(\Omega) = 0$.



Teď budeme pokrývat, a aby se nám to lépe dělalo, budeme pokrývat pomocí čtverců o straně $\frac{1}{m}$. Zvolme si tedy $m \in \mathbb{N}$ a uvažujme čtverce o straně $\frac{1}{m}$, jejichž středy leží na ose x. Protože jsme naše n-obdélníky definovali jako uzavřené, je možné je na sebe nalepit a hned vidíme, že k pokrytí naší množiny Ω stačí m takových čtverců.



Pro takovéto pokrytí pak dostáváme

$$\sum_{k=1}^{m} m(\Omega_k) = \sum_{k=1}^{m} \left(\frac{1}{m}\right)^2 = m \cdot \frac{1}{m^2} = \frac{1}{m}.$$

Pokud teď uvažujeme všechna pokrytí tohoto typu a vezmeme infimum odpovídajících sum, dostáváme inf $\left(\frac{1}{m}\right)=0$. Přibráním dalších pokrytí se infimum nemůže zvětšit a zmenšit se nemá kam, takže vidíme, že také $m^*(\Omega)=0$. Protože $m^*(\Omega)=m_*(\Omega)$, je naše úsečka měřitelná, ale hlavně vidíme, že opravdu $m(\Omega)=0$.

Podobně se dá ukázat, že všechny (rozumné) křivky v rovině mají míru nula, obdobný argument pak funguje i pro povrchy v \mathbb{R}^3 a podobně.

 \wedge

8b. Integrál

Pokud množina není měřitelná, tak na ní nemá smysl ani integrovat. Pokud měřitelná je, tak se dá dobře aproximovat pomocí *n*-obdélníků, pak má smysl nad nimi stavět "sloupce" a sčítat jejich příspěvky k integrálu. U klasického Riemannova integrálu jsme při počítání takového příspěvku jednoho "sloupce" násobili délku základny výškou, zde přirozeně pracujeme s mírou základny.

Nejprve zavedeme pojmy. Připomeňme, že v této kapitole pracujeme jen s neprázdnými a omezenými množinami $\Omega.$

157

Definice.

Nechť f je omezená funkce na měřitelné množině Ω .

Pro pokrytí $\{\Omega_k\}$ množiny Ω pomocí n-obdélníků definujeme

$$U(f, {\Omega_k}) = \sum_k \sup_{\Omega_k} (f) \cdot m(\Omega_k).$$

Dále definujeme **horní integrál** f na Ω jako

$$\overline{S}(f,\Omega) = \inf\{U(f,\{\Omega_k\}); \{\Omega_k\} \text{ obdélníkové pokrytí } \Omega\}.$$

Pro vyplnění $\{\Omega_k\}$ množiny Ω pomocí n-obdélníků definujeme

$$L(f, \{\Omega_k\}) = \sum_{k} \inf_{\Omega_k} (f) \cdot m(\Omega_k).$$

Dále definujeme **dolní integrál** f na Ω jako

$$\underline{S}(f,\Omega) = \sup\{L(f,\{\Omega_k\}); \{\Omega_k\} \text{ obdélníkové vyplnění } \Omega\}.$$

Pro smysluplnost definice zase potřebujeme klíčové lemma, které říká, že když pro množinu Ω máme nějaké pokrytí $\{\Omega_k\}$ a nějaké vyplnění $\{\widetilde{\Omega}_l\}$, tak pro f>0 musí platit

$$L(f, {\widetilde{\Omega}_l}) \le U(f, {\Omega_k}).$$

Jen to dá o něco víc práce než v jedné dimenzi. Z toho pak vyplyne, že definice horního a dolního integrálu má smysl, výsledkem jsou reálná čísla a vždycky platí

$$\underline{S}(f,\Omega) \leq \overline{S}(f,\Omega).$$

Pak už následuje povědomá definice.

Definice.

Nechť f je omezená funkce na měřitelné množině Ω .

Řekneme, že f je **Riemannovsky integrovatelná** na Ω , jestliže

$$\underline{S}(f,\Omega) = \overline{S}(f,\Omega).$$

V takovém případě definujeme

$$\int \dots \int f(\vec{x}) d[\vec{x}] = \overline{S}(f, \Omega).$$

Například všechny spojité funkce jsou integrovatelné, což pro běžné aplikace stačí. Takto definovaný integrál splňuje všechny obvyklé vlastnosti, třeba linearitu a aditivitu vůči integrační množině.

Věta.

Nechť f, g jsou funkce integrovatelné na měřitelné množině $\Omega \subseteq \mathbb{R}^n$, nechť $c \in \mathbb{R}$. Pak je také funkce cf + g integrovatelná na Ω a

$$\int_{\Omega} \cdots \int cf + g \, d[\vec{x}] = c \int_{\Omega} \cdots \int f \, d[\vec{x}] + \int_{\Omega} \cdots \int_{\Omega} g \, d[\vec{x}].$$

Věta.

Nechť f je funkce integrovatelná na měřitelné množině Ω .

Předpokládejme, že množiny $\Omega_1, \ldots, \Omega_m$ jsou měřitelné, jejich vnitřky jsou po dvou disjunktní a $\Omega = \bigcup_{k=1}^m \Omega_k$.

Pak

$$\int_{\Omega} \cdots \int f d[\vec{x}] = \sum_{k=1}^{m} \int_{\Omega_k} \cdots \int f d[\vec{x}].$$

Kdyby toto byla pořádná učebnice a ne ilustrovaný úvod, tak by teď přišly důkazy těchto vět a také věty další, například že tento Riemannův integrál je možné počítat pomocí řezů, jak jsme se to naučili v kapitole 6. Ale ona není. Věty o substituci už jsme také probrali v příslušné kapitole, takže si na rozloučenou dáme jeden integrál.

Příklad: Uvažujme funkci $f(x,y)=(x^2+y^2+1)^3$ na množině Ω , kde Ω je kruh v \mathbb{R}^2 se středem v počátku a poloměrem 3. Chceme integrál, přímý výpočet vede na dvojný integrál

$$\int_{-3}^{3} \int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}} (x^2 + y^2 + 1)^3 dy \, dx.$$

Zkušený integrátor si hned všimne, že preferovaný přístup, tedy substituce $w = x^2 + 1 + y^2$, zde není průchozí. Tento integrál se musí počítat roznásobením třetí mocniny a integrací všech výsledných členů, což není moc přitažlivé.

Kruh ale vyzývá k přechodu na polární souřadnice $x = r\cos(\varphi), y = r\sin(\varphi)$, které zobrazují $D = \langle 0, 3 \rangle \times \langle 0, 2\pi \rangle$ na Ω . Najdeme jakobián:

$$\Delta = \begin{vmatrix} \cos(\varphi) & \sin(\varphi) \\ -r\sin(\varphi) & r\cos(\varphi) \end{vmatrix} = r,$$

proto $dS = r dS[r, \varphi]$. Dostáváme

$$\iint_{\Omega} (x^2 + y^2 + 1)^3 dx dy = \int_{0}^{3} \int_{0}^{2\pi} (r^2 \cos^2(\varphi) + r^2 \sin^2(\varphi) + 1)^3 r \, d\varphi \, dr = \int_{0}^{3} \int_{0}^{2\pi} (r^2 + 1)^3 r \, d\varphi \, dr$$

$$= \int_{0}^{3} \left[r(r^2 + 1)^3 \varphi \right]_{\varphi=0}^{\varphi=2\pi} dr = \int_{0}^{3} 2\pi r (r^2 + 1)^3 \, dr = \left| \frac{w = r^2 + 1}{dw = 2r \, dr} \right|$$

$$= \int_{0}^{10} \pi w^3 \, dw = \left[\pi \frac{1}{4} w^4 \right]_{0}^{10} = \frac{\pi}{4} \cdot 9999.$$

 \triangle

9. Dodatky

9a. Geometrické objekty v \mathbb{R}^n

Zde si připomeneme některé populární objekty v \mathbb{R}^n .

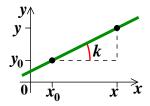
Lineární objekty

Začneme s přímkou v \mathbb{R}^2 procházející daným bodem (x_0, y_0) . Je několik způsobů, jak určit její směr.

ullet Jeden populární způsob je zadat její směrnici k. Rovnice přímky pak je

$$y = k(x - x_0) + y_0$$
 nebo $y - y_0 = k(x - x_0)$.

Druhá podoba nám připomíná skutečný význam čísla k. Je to poměr výšky k základně v libovolném trojúhelníku přilepeném k přímce. Někteří lidé dávají přednost napsat y=kx+q a dopočítat q z dat.

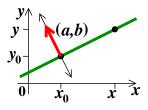


Hlavní nevýhoda tohoto přístupu je, že nám nedovolí zapsat svislé přímky.

• Další populární popis přímky je rovnicí ve tvaru

$$ax + by = c$$
 nebo $ax + by + c = 0$.

Koeficienty a,b tvoří vektor (a,b), který je na dotyčnou přímku kolmý a jednoznačně určuje její směr. Říká se mu normálový vektor a rozhodně není jediný. Libovolný nenulový násobek normálového vektoru zase poslouží jako normálový vektor.



Je-li dán normálový vektor $\vec{n}=(a,b)$ a bod (x_0,y_0) , tak lze odpovídající přímku popsat rovnicí

$$a(x - x_0) + b(y - y_0) = 0.$$

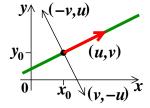
Dá se to také zapsat pomocí skalárního součinu jako $(a,b) \bullet (x-x_0,y-y_0) = 0$. To určuje přímku jako množinu všech bodů (x,y) takových, že vektor $(x,y) - (x_0,y_0)$ je kolmý k \vec{n} .

Přepsáním y=kx+q do tvaru kx+(-1)y+q=0 vidíme, že pro danou směrnici k lze použít (k,-1) coby normálový vektor naší přímky. Je-li naopak dána přímka specifikací ax+by+c=0 s $b\neq 0$, pak $y=-\frac{a}{b}x-\frac{c}{b}$, tedy jde o přímku se směrnicí $k=-\frac{a}{b}$.

 \bullet Směr přímky lze specifikovat směrovým vektorem (u,v). Pak lze použít parametrický popis bodů dotyčné přímky

$$(x,y) = (x_0,y_0) + t(u,v)$$
 neboli $x = a + tu,$ $y = b + tv, t \in \mathbb{R}.$

Eliminací t dostaneme $y = \frac{v}{u}x - \frac{v}{u}a + b$, následně vx - uy - va + ub = 0. Takováto přímka má tedy směrnici $k = \frac{v}{u}$ a normálový vektor $\vec{n} = (v, -u)$ nebo $\vec{n} = (-v, u)$ nebo jejich libovolný nenulový násobek.



Naopak, přímka daná rovnicí y = kx + q má směrový vektor (1, k) a přímka daná ax + by + c = 0 má směrový vektor (b, -a) nebo (-b, a) nebo jejich libovolný nenulový násobek.

Tyto tři způsoby popisu přímky v \mathbb{R}^2 mají rozličný zobecňovací potenciál.

Parametrický popis lze přímo použít také pro přímky v \mathbb{R}^n pro libovolné $n \geq 2$. Je-li dán bod \vec{a} a směrový vektor \vec{u} , pak jsou body přímky procházející bodem \vec{a} ve směru \vec{u} dány jako

$$\vec{x} = \vec{a} + t\vec{u}, \ t \in \mathbb{R}.$$

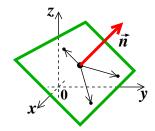
Je to náš preferovaný způsob zápisu přímek ve vyšší dimenzi.

Směrnicová verze žádné zobecnění nenabízí. Verze s normálovým vektorem zobecnění nabízí, ale jiným směrem. Už v prostoru \mathbb{R}^3 snadno nahlédneme, že když daný vektor \vec{n} umístíme v nějakém bodě \vec{a} , tak existuje nekonečně mnoho přímek, které daným bodem procházejí a jsou k tomuto vektoru kolmé. To znamená, že ve vyšších dimenzích nelze specifikovat přímky pomocí normálového vektoru.

Na druhou stranu není těžké nahlédnout, že body $\vec{x} \in \mathbb{R}^3$ takové, že vektor $\vec{x} - \vec{x}_0$ je kolmý na $\vec{n} = (a, b, c)$, tvoří v \mathbb{R}^3 plochou rovinu. Vlastně sjednocujeme všechny ty kolmé přímky. Výsledná rovnice je

$$a(x - x_0) + b(y - y_0) + c(z - z_0) = 0.$$

Obecně, každá rovnice ve tvaru ax + by + cz = d nebo ax + by + cz + d = 0 definuje rovinu v \mathbb{R}^3 , jejíž normálový vektor je $\vec{n} = (a, b, c)$.



Sklon takovéto roviny je svým normálovým vektorem jednoznačně určen (zase je nekonečně mnoho možných normálových vektorů). To mimo jiné znamená, že když chceme vědět, zda jsou dvě roviny rovnoběžné, tak položíme stejnou otázku o jejich normálových vektorech.

Je nemožné (pro většínu lidí) nahlédnout, že množina všech bodů $\vec{x} \in \mathbb{R}^4$, pro které je $\vec{x} - \vec{x}_0$ kolmý k \vec{n} , je vlastně objekt, jehož skutečný rozměr je 3 a který je také plochý. Je to ale tak. Obecně, libovolný nenulový vektor $\vec{n} \in \mathbb{R}^n$ umístěný v nějakém bodě \vec{a} určuje specifický objekt, "plochou" množinu, která je svou podstatou n-1 rozměrná a kolmá k \vec{n} . Takovým objektům říkáme nadroviny. Matematicky, nadrovinou v \mathbb{R}^n rozumíme libovolnou množinu ve tvaru $H + \vec{a}$, kde H je nějaký lineární podprostor \mathbb{R}^n dimenze n-1.

Nadrovina daná bodem \vec{a} a normálovým vektorem \vec{n} je popsána rovnicí

$$\vec{n} \bullet (\vec{x} - \vec{a}) = 0$$

neboli

$$\sum_{i=1}^{n} n_i (x_i - a_i) = 0.$$

Průsečíky nadrovin vedou k množinám nižšího rozměru, což nabízí další možnost, jak zachytit přímky v \mathbb{R}^n , jmenovitě coby průsečíky n-1 nerovnoběžných nadrovin.

Kvadriky (kvadratické objekty)

Intuitivně jde o objekty popsatelné rovnicemi, ve kterých jsou kvadratické polynomy.

 $V \mathbb{R}^2$ jsou tyto objekty známé coby kuželosečky. Jsou dva základní typy:

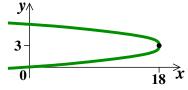
• Rovnice, kde jedna proměnná není přítomna na druhou: V této rodině je jen jeden tvar. Rovnice $y = ax^2 + bx + c$ a $x = ay^2 + by + c$ definují **paraboly**. Ta první se otevírá nahoru (pro a > 0) nebo dolů (pro a < 0), zatímco ta druhá se otevírá doprava či doleva.

Důležité prvky, zejména vrchol, lze identifikovat doplněním čtverce.

Například rovnici $x = 12y - 2y^2$ lze přepsat jako

$$x = -2(y^2 - 6y) = -2(y^2 - 6y + 9 - 9) = -2((y - 3)^2 - 9) = 18 - 2(y - 3)^2.$$

Vidíme, že popisuje parabolu otevřenou doleva a s vrcholem (18,3).



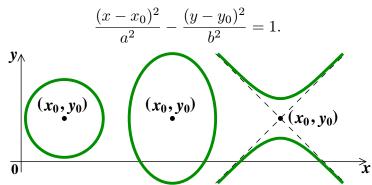
- Rovnice, kde obě proměnné mají druhou mocninu: Typický reprezentant je rovnice $ax^2 + by^2 = c$ s $a, b \neq 0$. Jou dva základní tvary, jeden z nich má podtyp, který stojí za vypíchnutí.
 - Kružnice se středem (x_0, y_0) a poloměrem $r \ge 0$ je dána rovnicí

$$(x - x_0)^2 + (y - y_0)^2 = r^2.$$

- Elipsa se středem (x_0,y_0) a poloosami a,b>0 je dána rovnicí

$$\frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} = 1.$$

- **Hyperbola** je dána rovnicí



Rovnice s linárními členy se zpracovávají doplněním čtverců. Například

$$x^{2} + y^{2} - 4x + 2y = 4 \implies (x - 2)^{2} + (y + 1)^{2} = 3^{2}$$
.

Toto popisuje kružnici o poloměru 3 se středem v (2, -1).

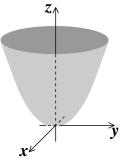
Katalog kvadrických povrchů v \mathbb{R}^3 je podstatně bohatší. Proto zde uvedeme jen některé populární typy a se středem v počátku, což se snadno upraví na jiné středy.

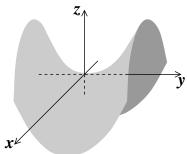
- Začneme rovnicemi, kde jedna proměnná není v mocnině:
- $-x^2+y^2-z=0$ neboli $z=x^2+y^2$ popisuje **paraboloid**. Viz příklad v sekci 1b.

Tento paraboloid lze zploštit ze strany, čímž vznikne eliptický paraboloid daný rovnicí

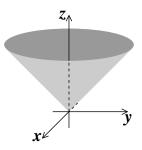
$$z = \frac{x^2}{a^2} + \frac{y^2}{b^2}.$$

 $-x^2-y^2-z=0$ neboli $z=x^2-y^2$ popisuje **hyperbolický paraboloid**. Jde o zajímavý objekt ve tvaru sedla, viz příklad v sekci 1b.





- A teď rovnice, ve kterých jsou všechny proměnné v druhé mocnině:
- $-x^2+y^2+z^2=r^2$ definuje **sféru** o poloměru r. Jde o populární speciální případ následujícího typu.
- $-\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ definuje **elipsoid**.
- $a^2+b^2+c^2$ $-x^2+y^2-z^2=0$ neboli $z=\sqrt{x^2+y^2}$ popisuje **kužel** s vrcholem v počátku. Všimneme si, že pro bod (x,y) v rovině xy udává $r=\sqrt{x^2+y^2}$ jeho vzdálenost od počátku. Vzorec $z=\sqrt{x^2+y^2}$ říká, že v bodě (x,y) je hodnota z rovna vzdálenosti od počátku. Graf se tedy zvedá stejnou rychlostí, jakou se vzdalujeme od počátku, jinak řečeno, strana kužele má sklon 1:1.



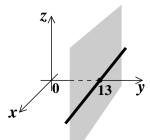
Úplný seznam kvadrik v \mathbb{R}^3 se snadno najde.

Degenerované křivky a povrchy

Jde o rovnice, ve kterých chybí některá proměnná. Základní pravidlo je, že degenerované objekty nejprve nakreslíme jen v prostoru, jehož souřadnice jsou v rovnici přítomny, a pak jej protáhneme do nekonečna ve směru proměnných, které v rovnici chybí.

Uvažujme rovnici y=13. V $\mathbb R$ definuje bod na přímce. Ve vyšších dimenzích se stane degenerovanou. V $\mathbb R^2$ definuje vodorovnou přímku. V $\mathbb R^3$ definuje rovinu, která je rovnoběžná s rovinou xz. Obdobně to platí pro rovnice $x=x_0$ a $z=z_0$.

Podobně, rovnice x+2y=13 definuje přímku v \mathbb{R}^2 . V prostoru \mathbb{R}^3 se stane degenerovanou a definuje svislou rovinu procházející onou přímkou v rovině xy.



Rovnice $x^2 + y^2 = r^2$ v \mathbb{R}^3 definuje svislý **válec** o poloměru r, zatímco $y^2 + z^2 = r^2$ tam definuje válec rovnoběžný s osou x.

Obecně rovnice

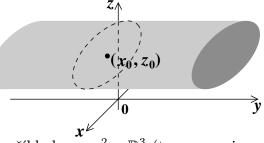
$$\frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} = 1$$

definuje **eliptický válec**, tedy válec rovnoběžný s osou z, který je ze stran zploštělý a jeho osa prochází bodem $(x_0, y_0, 0)$.

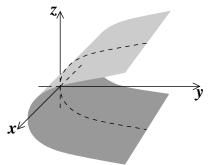
I zde jsou proměnné zaměnitelné, takže například

$$\frac{(x-x_0)^2}{a^2} + \frac{(z-z_0)^2}{c^2} = 1$$

definuje zploštělý válec rovnoběžný s osou y jako na obrázku.



Čtenář si jistě dokáže vymyslet řadu takových příkladů, například $y = z^2$ v \mathbb{R}^3 (tomu se mimochodem říká **parabolický válec**).



Naštěstí skoro nikdy nepracujeme v ještě vyšších dimenzích, protože tam už je počet tvarů kvadrik masivní a stejně si to neumíme představit. Za zmínku stojí, že jeden tvar se škáluje dobře, jmenovitě sféra, která je ve všech dimenzích reprezentovaná rovnicí

$$\sum (x_i - a_i)^2 = r^2.$$

Obecně lze uvažovat elipsoidy

$$\sum \frac{(x_i - a_i)^2}{r_i^2} = 1.$$

9b. Topologické pojmy v \mathbb{R}^n

Značení:

Nechť $n \in \mathbb{N}$. Symbol \mathbb{R}^n značí prostor všech n-rozměrných vektorů s reálnými souřednicemi značených $\vec{x} = (x_1, \dots, x_n)$, také \bar{x} nebo třeba x. Čísla $x_i \in \mathbb{R}$ jsou **souřadnice** či **složky**.

Pro vektory máme základní operace.

Definice.

Pro vektor $\vec{x} \in \mathbb{R}^n$ a skalár $c \in \mathbb{R}$ definujeme **násobek** jako

$$c\vec{x} = (cx_1, \dots, cx_n).$$

Pro vektory $\vec{x}, \vec{y} \in \mathbb{R}^n$ definujeme jejich

• součet jako

$$\vec{x} + \vec{y} = (x_1 + y_1, \dots, x_n + y_n);$$

• skalární součin jako

$$\vec{x} \bullet \vec{y} = x_1 y_1 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i.$$

Pro vektory \vec{x}, \vec{y} z \mathbb{R}^3 také máme **vektorový součin**

$$\vec{x} \times \vec{y} = \begin{pmatrix} x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1 \end{pmatrix}$$

$$= \begin{pmatrix} \det \begin{pmatrix} x_2 & x_3 \\ y_2 & y_3 \end{pmatrix}, \det \begin{pmatrix} x_3 & x_1 \\ y_3 & y_1 \end{pmatrix}, \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} \end{pmatrix}$$

$$= \det \begin{pmatrix} \vec{e}_1 & \vec{e}_2 & \vec{e}_3 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} = \det \begin{pmatrix} \vec{i} & \vec{j} & \vec{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}.$$

Pro skalární součin existují i alternativní značení či alternativní vzorce, například

$$\vec{x} \bullet \vec{y} = \langle \vec{x}, \vec{y} \rangle = \vec{x} \cdot \vec{y}^T.$$

V tom posledním vzorečku se řádkový vektor násobí sloupcovým vektorem, jako by to byly dvě matice.

Pojem vzdálenosti neboli metriku nám určuje norma.

Definice.

Pro vektor $\vec{x} \in \mathbb{R}^n$ definujeme jeho (Euklidovskou) normu jako

$$\|\vec{x}\| = \sqrt{x_1^2 + \dots + x_n^2}.$$

Pro vektory $\vec{x}, \vec{y} \in \mathbb{R}^n$ definujeme jejich (Euklidovskou) vzdálenost jako

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

Vektorovému prostoru \mathbb{R}^n s Euklidovskou normou říkáme **Euklidovský prostor** dimenze n, někdy se značí $(\mathbb{R}^n, \|\cdot\|_2)$ či E_n . My budeme psát jen \mathbb{R}^n , protože jiné normy neuvažujeme.

Každá norma musí splňovat určité vlastnosti už proto, že je to norma, viz (i)–(iii) níže. Přidáme i nějaká další užitečná pozorování o Euklidovské normě.

Věta.

Nechť $n \in \mathbb{N}$. Euklidovská norma na \mathbb{R}^n splňuje následující vlastnosti:

- (i) Pro všechna $\vec{x} \in \mathbb{R}^n$ platí: $\|\vec{x}\| = 0$ právě tehdy, když $\vec{x} = \vec{0}$;
- (ii) Pro všechna $\vec{x} \in \mathbb{R}^n$ a $c \in \mathbb{R}$ platí $||c\vec{x}|| = |c| \cdot ||\vec{x}||$;
- (iii) Pro všechna $\vec{x}, \vec{y} \in \mathbb{R}^n$ platí $||\vec{x} + \vec{y}|| \le ||\vec{x}|| + ||\vec{y}||$; (trojúhelníková nerovnost)
- (iv) Pro všechna $\vec{x} \in \mathbb{R}^n$ platí $||\vec{x}|| = \sqrt{\vec{x} \cdot \vec{x}};$
- (v) Pro všechna $\vec{x}, \vec{y} \in \mathbb{R}^n$ platí $|\vec{x} \cdot \vec{y}| \le ||\vec{x}|| \cdot ||\vec{y}||$;
- (Cauchy-Schwarzova nerovnost) (vi) Pro všechna $\vec{x} \in \mathbb{R}^n$ platí $\max |x_i| \le ||\vec{x}|| \le \sqrt{n} \max |x_i|$.

Skalární součin se dá spočítat pomocí normy a informaci o pozici vektorů.

Fakt.

Nechť $n \in \mathbb{N}$. Pro $\vec{x}, \vec{y} \in \mathbb{R}^n$ platí $\vec{x} \cdot \vec{y} = ||\vec{x}|| \cdot ||\vec{y}|| \cdot \cos(\alpha)$, kde α je úhel mezi vektory \vec{x} a \vec{y} .

Skalární součin je míra kolmosti. Platí $\vec{x} \bullet \vec{y} = 0$ právě tehdy, když $\vec{x} \perp \vec{y}$. Naopak pro \vec{x}, \vec{y} rovnoběžné nabývá součin $\vec{x} \bullet \vec{y}$ maximální možné hodnoty $\pm ||\vec{x}|| \, ||\vec{y}||$ (viz Cauchy-Schwarz výše), znaménko záleží na tom, zda je orientace shodná nebo opačná.

Definice.

Nechť M je podmnožina \mathbb{R}^n .

Řekneme, že je **omezená**, jestliže existuje K>0 tak, aby $\|\vec{x}\| \leq K$ pro všechna $\vec{x} \in M$.

Definujeme **diametr** M jako diam $(M) = \sup\{\|\vec{x} - \vec{y}\|; \ \vec{x}, \vec{y} \in M\}.$

Evidentně je množina M omezená právě tehdy, když diam $(M) < \infty$.

Definice.

Nechť $\vec{a} \in \mathbb{R}^n$ a $\varepsilon > 0$. Definujeme

- ε -okolí bodu \vec{a} jako $U_{\varepsilon}(\vec{a}) = \{\vec{x} \in \mathbb{R}^n; ||\vec{x} \vec{a}|| < \varepsilon\};$
- prstencové ε -okolí bodu \vec{a} jako $P_{\varepsilon}(\vec{a}) = \{\vec{x} \in \mathbb{R}^n; \ 0 < ||\vec{x} \vec{a}|| < \varepsilon\}.$

 $U_R(\vec{a})$ je otevřená "koule" o poloměru R se středem v \vec{a} . Množina M je omezená právě tehdy, pokud je podmnožinou nějakého $U_R(\vec{a})$; pak diam $(M) \leq 2R$.

Definice.

Nechť M je podmnožina \mathbb{R}^n . Řekneme, že $\vec{x} \in \mathbb{R}^n$ je

- vnitřní bod M, jestliže existuje $U = U_{\varepsilon}(\vec{x})$ takové, že $U \subseteq M$;
- vnější bod M, jestliže existuje $U = U_{\varepsilon}(\vec{x})$ takové, že $U \cap M = \emptyset$;
- hraniční bod M, jestliže pro každé $U = U_{\varepsilon}(\vec{x})$ platí $U \cap M \neq \emptyset$ a $U M \neq \emptyset$;
- izolovaný bod M, jestliže existuje $U = U_{\varepsilon}(\vec{x})$ takové, že $U \cap M = \{\vec{x}\}$;
- hromadný bod M, jestliže pro každé $P = P_{\varepsilon}(\vec{x})$ platí $P \cap M \neq \emptyset$.

Ekvivalentně, \vec{x} je hromadný bod M přesně tehdy, když existuje posloupnost $\{\vec{x}(k)\}\subseteq M-\{\vec{x}\}$ taková, že $\vec{x}(k)\to \vec{x}$.

Fakt.

Jestliže má množina $M \subseteq \mathbb{R}^n$ hromadný bod, pak musí být nekonečná.

Neplatí naopak, například \mathbb{N} je nekonečná podmnožina \mathbb{R}^1 , ale nemá hromadný bod.

Věta.

Každá nekonečná omezená podmnožina \mathbb{R}^n má alespoň jeden hromadný bod.

Definice.

Nechť M je podmnožina \mathbb{R}^n . Definujeme její

- vnitřek M^O jako množinu všech vnitřních bodů M;
- hranici ∂M jako množinu všech hraničních bodů M;
- uzávěr \overline{M} jako $M \cup \partial M$.

Pro každou množinu M platí $M^O \subseteq M \subseteq \overline{M}$ a $M^O \cap \partial M = \emptyset$.

Uzávěr je množina všech bodů, ke kterým se dá libovolně blízko dojít pomocí bodů z M.

$$\vec{x} \in \overline{M} \iff \text{existuje posloupnost } \{\vec{x}(k)\} \subseteq M \text{ aby } \vec{x}(k) \to \vec{x}.$$

To samozřejmě zahrnuje i body z M, k takovému bodu se dá dojít libovolně blízko pomocí jeho samého.

Definice.

Nechť M je podmnožina \mathbb{R}^n .

Řekneme, že M je **otevřená**, jestliže $M^O = M$.

Řekneme, že M je **uzavřená**, jestliže $\overline{M} = M$.

Množina je otevřená, je-li každý její bod vnitřní.

Množina je uzavřená, jestliže je každý její hromadný bod částí množiny. Jinými slovy, pro libovolnou posloupnost $\{\vec{x}(k)\}\subseteq M$, která konverguje k nějakému \vec{x} , musí platit $\vec{x}\in M$.

Definice.

Nechť M je podmnožina \mathbb{R}^n . Řekneme, že je **souvislá**, jestliže neexistují otevřené množiny $G_1, G_2 \subseteq \mathbb{R}^n$ takové, že $M \subseteq G_1 \cup G_2$, $G_1 \cap M \neq \emptyset$, $G_2 \cap M \neq \emptyset$ a $G_1 \cap G_2 = \emptyset$.

Intuitivně řečeno, množina $M \subseteq \mathbb{R}^n$ je souvislá právě tehdy, jestliže pro libovolné dva body $\vec{x}, \vec{y} \in M$ existuje nepřerušovaná čára (viz spojité parametrické křivky) z \vec{x} do \vec{y} , která celá leží v M.

Věta.

Jediné souvislé podmnožiny \mathbb{R} jsou intervaly.

Když už jsme u intervalů: Vícerozměrné intervaly v \mathbb{R}^n jsou množiny typu $I_1 \times I_2 \times \cdots \times I_n$, kde I_i jsou intervaly v \mathbb{R} . Představujeme si obdélník, hranol, atd. Ve více dimenzích jsou ale i jiné souvislé množiny, třeba kruh v \mathbb{R}^2 .

Definice.

Množina $M \subseteq \mathbb{R}^n$ se nazývá **oblast**, jestliže je otevřená a souvislá.

Definice.

Nechť $\vec{x}, \vec{y} \in \mathbb{R}^n$. Definujeme **úsečku** $\langle \vec{x}, \vec{y} \rangle$ jako množinu

$$\{t\vec{x} + (1-t)\vec{y}; \ t \in \langle 0, 1 \rangle\}.$$

Definice.

Nechť M je podmnožina \mathbb{R}^n . Řekneme, že je **konvexní**, jestliže pro libovolné $\vec{x}, \vec{y} \in M$ platí $\langle \vec{x}, \vec{y} \rangle \subseteq M$.

S topologií úzce souvisí pojem konvergence.

Definice.

Nechť $n \in \mathbb{N}$, uvažujme vektor $\vec{x} \in \mathbb{R}^n$ a posloupnost vektorů $\{\vec{x}(k)\} \subseteq \mathbb{R}^n$. Řekneme, že $\{\vec{x}(k)\}$ konverguje k \vec{x} nebo že \vec{x} je limita posloupnosti $\{\vec{x}(k)\}$, značeno $\lim_{k \to \infty} (\vec{x}(k)) = \vec{x}$ či jen $\vec{x}(k) \to \vec{x}$, jestliže $\lim_{k \to \infty} (\|\vec{x}(k) - \vec{x}\|) = 0$. Ekvivalentně,

$$\forall \varepsilon > 0 \ \exists N \in \mathbb{N} \ \forall k \ge N : \ \|\vec{x}(k) - \vec{x}\| < \varepsilon$$

neboli

$$\forall U = U(\vec{x}) \ \exists N \in \mathbb{N} \ \forall k \ge N : \ \vec{x}(k) \in U.$$

Fakt.

Nechť $n \in \mathbb{N}$, uvažujme vektor $\vec{x} \in \mathbb{R}^n$ a posloupnost vektorů $\{\vec{x}(k)\}\subseteq \mathbb{R}^n$, kde $\vec{x}(k) = (x(k)_1, \dots, x(k)_n)$.

Platí $\vec{x}(k) \to \vec{x}$ právě tehdy, když pro všechna $i = 1, \ldots, n$ platí $x(k)_i \to x_i$.

Konvergence tedy funguje po souřadnicích.