KERESZTFALVI TIBOR

ANALÍZIS IV.

Differenciálás

Tartalom

A. SZÁMFÜGGVÉNYEK DIFFERENCIÁLÁSA

B. DIFFERENCIÁLÁS NORMÁLT TEREKBEN

IV. DIFFERENCIÁLHATÓ FÜGGVÉNYEK

V. TÖBBSZÖR DIFFERENCIÁLHATÓ FÜGGVÉNYEK

VI. RÉSZSOKASÁGOK

VII. A FIZIKÁBAN HASZNÁLT NÉHÁNY FOGALOM

.

A. SZÁMFÜGGVÉNYEK DIFFERENCIÁLÁSA

I. DIFFERENCIÁLHATÓ FÜGGVÉNYEK

1. Differenciálhatóság, alapvető tulajdonságok

1.1. A függvények differenciálása a matematikában és alkalmazásaiban lépten-nyomon előforduló fogalom. Sokféleképpen juthatunk el hozzá; említsük meg azokat a klasszikus példákat,amelyekből tulajdonképpen a differenciálszámítás eredt.

Reprezentáljuk fizikai terünket számhármasokkal, az időt pedig számokkal. Ekkor egy tömegpont mozgását, azaz helyzetét mint az idő függvényét egy $r : \mathbb{R} \rightarrowtail \mathbb{R}^3$ függvénnyel írjuk le. A mozgás egyik jellemzője a sebesség, amely megmutatja, egységnyi idő alatt milyen az elmozdulás. Ezt így értelmezzük: az $[s, t]$ idő-intervallumban az átlagsebesség $\dfrac{r(t) - r(s)}{t - s}$; minél kisebb az intervallum, annál jobban megközelíti az átlagsebesség a sebességet, vagyis a sebesség az átlagsebességek határértéke, miközben az idő-intervallum egy pontra zsugorodik.

A másik példa geometriai szemléltetést ad. Az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvény grafikonja jó esetben – például, ha $f = \mathrm{id}_{\mathbb{R}}^2$ – egy görbe $\mathbb{R}^2$-ben (mi a jó eset? mi a görbe? e kérdésekre éppen a differenciálszámítás ad választ). Húzzunk érintőt a grafikon egy kijelölt $(a, f(a))$ pontjához! Ezt a nem is egyszerű feladatot úgy próbáljuk megoldani, hogy szelőkkel közelítjük az érintőt. Az $(x, f(x))$ ponton is áthaladó szelő meredeksége $\dfrac{f(x) - f(a)}{x - a}$. Minél közelebb van x az a ponthoz, annál jobban megközelíti a szelő az érintőt, vagyis az érintő meredeksége a szelőmeredekségek határértéke.

Vizsgálatainknál – a bevezető példákkal ellentétben – nem szorítkozunk valós függvényekre; ebben a részben $\mathbb{K} \rightarrowtail \mathbb{K}^N$ $(N \in \mathbb{N})$ függvények differenciálásáról lesz szó.

1.2. Definíció *(i) Az $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ függvény* ***differenciálható az*** *$a \in \mathbb{K}$* ***pontban****, ha $a \in \overset{\circ}{\mathrm{Dom}}(f)$ és létezik*

$$\lim_{x \to a} \frac{f(x) - f(a)}{x - a} =: f'(a).$$

$f'(a)$-t az f függvény a ***pontbeli deriváltjának*** *nevezzük.*

(ii) Egy függvényt akkor hívunk ***differenciálhatónak egy halmazon****, ha annak minden pontjában differenciálható.*

(iii) Egy függvényt akkor nevezünk ***differenciálhatónak****, ha differenciálható az értelmezési tartományán.*

(iv) Az $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ függvény ***deriváltjának*** *hívjuk azt a $f' := \mathbb{K} \rightarrowtail \mathbb{K}^N$, $x \mapsto f'(x)$ függvényt, amely az $\{x \in \mathrm{Dom}(f) \mid f \text{differenciálható } x\text{-ben}\}$ halmazon van értelmezve.*

Megjegyzések (i) Mint az a definícióból is kitűnik, egy függvény csak olyan halmazon lehet differenciálható, melynek minden pontja belső pontja a függvény értelmezési tartományának. Ezért egy differenciálható függvény értelmezési tartománya szükségképpen nyílt.

(ii) A definíció szerint az üres függvény differenciálható (az értelmezési tartomány nyílt, és nincs olyan pontja, ahol a függvény ne lenne differenciálható).

(iii) Az iménti definíció szerint minden függvénynek van deriváltja, viszont előfordulhat, hogy a derivált értelmezési tartománya üres (pontosan akkor, ha az adott függvény sehol sem diferenciálható).

(iv) Sokszor célszerű a $h := x - a$ jelöléssel élve a deriváltat $\lim\limits_{h \to 0} \dfrac{f(a+h) - f(a)}{h}$ formában előállítani.

(v) Függvények határértékére vonatkozó ismereteinkből közvetlenül adódik, hogy az $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ függvény akkor és csak akkor differenciálható valamely a pontban, ha minden komponense differenciálható a-ban, azaz minden $k = 1, ..., N$ esetén a $\mathrm{pr}_k \circ f : \mathbb{K} \to \mathbb{K}^N$ függvény differenciálható a-ban, és ekkor minden k-ra igaz, hogy

$$\mathrm{pr}_k(f'(a)) = (\mathrm{pr}_k \circ f)'(a).$$

(vi) Mivel $\mathbb{R} \rightarrowtail \mathbb{C}$ függvények differenciálhatóságának értelmezése valós számmal való osztást tartalmaz, és ebből a szempontból $\mathbb{C}$ azonosítható $\mathbb{R}^2$-vel, az előző megjegyzés szerint egy $f : \mathbb{R} \rightarrowtail \mathbb{C}$ függvény akkor és csak akkor differenciálható valamely a pontban, ha mind a valós része, mind a képzetes része differenciálható, és ekkor

$$\mathrm{Re}(f'(a)) = (\mathrm{Re} \circ f)'(a), \quad \mathrm{Im}(f'(a)) = (\mathrm{Im} \circ f)'(a).$$

Figyelem: ez nem igaz $\mathbb{C} \rightarrowtail \mathbb{C}$ függvényekre, hiszen ott a differenciálhatóság értelmezésében komplex számmal való osztás szerepel (lásd az 1.9.9. feladatot is).

(vii) Olykor a differenciálást a vessző helyett a függvény fölé tett ponttal jelöljük, mint például görbék paraméterezésénél (lásd később): $\dot{p} := p'$.

1.3. Állítás *Ha az $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ függvény differenciálható az a pontban, akkor folytonos is a-ban.*

Bizonyítás Ha $x \in \mathrm{Dom}(f)$ és $x \neq a$, akkor

$$f(x) - f(a) = \frac{f(x) - f(a)}{x - a}(x - a).$$

Az f differenciálhatósága miatt az egyenlőség jobb oldalának létezik határértéke $x \to a$ esetén és nullával egyenlő, ezért az egyenlőség bal oldalának is létezik határértéke, amely szintén nulla, vagyis

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

ami pontosan azt jelenti, hogy az f függvény folytonos az a pontban.

1.4. Állítás *Legyenek az $f, g : \mathbb{K} \rightarrowtail \mathbb{K}^N$ és a $\phi : \mathbb{K} \rightarrowtail \mathbb{K}$ függvények differenciálhatók az a pontban. Ekkor*

(i) minden $\lambda \in \mathbb{K}$ esetén λf is differenciálható a-ban és

$$(\lambda f)'(a) = \lambda f'(a),$$

(ii) $f + g$ is differenciálható a-ban és

$$(f + g)'(a) = f'(a) + g'(a),$$

(iii) ϕf is differenciálható a-ban és

$$(\phi f)'(a) = \phi'(a) f(a) + \phi(a) f'(a),$$

(iv) ha $\phi(a) \neq 0$, akkor $\dfrac{f}{\phi}$ is differenciálható a-ban és

$$\left(\frac{f}{\phi}\right)'(a) = \frac{\phi(a) f'(a) - \phi'(a) f(a)}{\phi^2(a)}.$$

Bizonyítás Először is meg kell jegyeznünk a következőket:

$$\mathrm{Dom}(\lambda f) = \mathrm{Dom}(f), \quad \mathrm{Dom}(f+g) = \mathrm{Dom}(f) \cap \mathrm{Dom}(g),$$

$$\mathrm{Dom}(\phi f) = \mathrm{Dom}(\phi) \cap \mathrm{Dom}(f), \quad \mathrm{Dom}\left(\frac{f}{\phi}\right) = \Big(\mathrm{Dom}(\phi) \cap \mathrm{Dom}(f)\Big) \setminus \overset{-1}{\phi}(\{0\}),$$

és a $\mathbb{K}^N$ bármely F és G részhalmazára (Analízis III.A.2.13.) $\mathring{F} \cap \mathring{G} = F \mathring{\cap} G$, tehát ha a belső pontja mind az f mind a g értelmezési tartományának, akkor belső pontja az $f+g$ értelmezési tartományának is stb.

Ezek után az (i) és (ii) pontok egyszerűen a határérték linearitásából következnek. A (iii) pont bizonyításához tekintsük a következő átalakítást:

$$\frac{\phi(x)f(x) - \phi(a)f(a)}{x-a} = \frac{\phi(x)f(x) - \phi(a)f(x)}{x-a} + \frac{\phi(a)f(x) - \phi(a)f(a)}{x-a} =$$
$$= \frac{\phi(x) - \phi(a)}{x-a} f(x) + \phi(a)\frac{f(x) - f(a)}{x-a}.$$

A jobb oldal mindkét tagjának létezik határértéke $x \to a$ esetén, mivel f és ϕ differenciálhatók a-ban (ezért folytonosak is), tehát az egyenlőség bal oldalának is létezik határértéke, és a két határérték megegyezik:

$$\lim_{x\to a} \frac{\phi(x)f(x) - \phi(a)f(a)}{x-a} = \lim_{x\to a} \frac{\phi(x) - \phi(a)}{x-a} f(x) + \phi(a) \lim_{x\to a} \frac{f(x) - f(a)}{x-a} =$$
$$= \lim_{x\to a} \frac{\phi(x) - \phi(a)}{x-a} \lim_{x\to a} f(x) + \phi(a) \lim_{x\to a} \frac{f(x) - f(a)}{x-a} =$$
$$= \phi'(a)f(a) + \phi(a)f'(a),$$

és ezzel a (iii) pontot be is bizonyítottuk.

A (iv) ponttal kapcsolatban vegyük észre, hogy ϕ folytonos a-ban (ezt már láttuk) és $\phi(a) \neq 0$, ezért létezik az $a \in \mathbb{K}$ pontnak egy olyan $G(a) \subset \mathrm{Dom}(\phi)$ környezete, ahol ϕ nem veszi fel a nulla értéket. Legyen $x \in G(a)$, $x \neq a$:

$$\frac{\dfrac{f(x)}{\phi(x)} - \dfrac{f(a)}{\phi(a)}}{x-a} = \frac{\phi(a)f(x) - \phi(x)f(a)}{(x-a)\phi(x)\phi(a)} =$$
$$= \frac{\dfrac{\phi(a)f(x) - \phi(a)f(a)}{x-a} - \dfrac{\phi(x)f(a) - \phi(a)f(a)}{x-a}}{\phi(x)\phi(a)} =$$
$$= \frac{\phi(a)\dfrac{f(x) - f(a)}{x-a} - \dfrac{\phi(x) - \phi(a)}{x-a} f(a)}{\phi(x)\phi(a)}.$$

Az egyenlőség jobb oldalának létezik határértéke $x \to a$ esetén, mivel f és ϕ differenciálható (ezért folytonos is) a-ban, tehát az egyenlőség bal oldalának is létezik határértéke, és a két határérték megegyezik:

$$\lim_{x\to a}\frac{\dfrac{f(x)}{\phi(x)}-\dfrac{f(a)}{\phi(a)}}{x-a}=\lim_{x\to a}\frac{\phi(a)\dfrac{f(x)-f(a)}{x-a}-\dfrac{\phi(x)-\phi(a)}{x-a}f(a)}{\phi(x)\phi(a)}=$$
$$=\frac{\phi(a)f'(a)-\phi'(a)f(a)}{\phi^2(a)},$$

és ezzel a (iv) pontot is bebizonyítottuk.

Megjegyzés Ha f differenciálható a-ban de g nem, viszont $a \in \overset{\circ}{\mathrm{Dom}}(g)$, akkor $f+g$ sem lehet differenciálható a-ban; ugyanis ha az volna, akkor $g \supset (g+f)-f$ is differenciálható volna a-ban.

Természetesen két nem differenciálható függvény összege lehet differenciálható.

1.5. (i) Nyilvánvaló, hogy egy konstans függvény differenciálható az értelmezési tartománya belsejében, és a deriváltja minden pontban nulla.

(ii) Az $\mathrm{id}_{\mathbb{K}}$ függvény szintén differenciálható, és a deriváltja az azonosan 1 függvény: $\mathrm{id}_{\mathbb{K}}' = 1$. Az iménti állítás (iii) pontjából teljes indukcióval könnyen látható, hogy

$$(\mathrm{id}_{\mathbb{K}}^n)' = n\mathrm{id}_{\mathbb{K}}^{n-1} \qquad (n\in\mathbb{N}),$$

és az állitás (iv) pontjából pedig ugyancsak teljes indukcióval

$$\left(\frac{1}{\mathrm{id}_{\mathbb{K}}^n}\right)' = -n\frac{1}{\mathrm{id}_{\mathbb{K}}^{n+1}} \qquad (n\in\mathbb{N})$$

következik. Ha akarjuk, e két formulát és az (i) szerinti $(\mathrm{id}_{\mathbb{K}}^0)' = 0$ összefüggést egyesíthetjük:

$$(\mathrm{id}_{\mathbb{K}}^n)' = n\mathrm{id}_{\mathbb{K}}^{n-1} \qquad (n\in\mathbb{Z}).$$

(iii) Az előző állítás (i) és (ii) pontja alapján most már megállapíthatjuk, hogy minden polinom differenciálható, sőt (iv) miatt minden racionális törtfüggvény is az.

(iv) Az sem igényel bővebb magyarázatot, hogy az $\mathbb{R}\to\mathbb{R}$ abszolútérték-függvény differenciálható az $\mathbb{R}\setminus\{0\}$ halmazon és ott a deriváltja a sign (előjel) függvény.

Ezzel szemben a $\mathbb{C}\to\mathbb{C}$ abszolútérték-függvény sehol sem differenciálható. A nullában azért nem, mert e függvény a valós abszolútérték kiterjesztése, és az sem differenciálható a nullában. Ha viszont $a\neq 0$, akkor az $|a|$ sugarú körvonalon levő z pontokra $\dfrac{|z|-|a|}{z-a}=0$, az a által meghatározott egyenes mentén levő z pontok esetén azonban $\dfrac{|z|-|a|}{z-a}$ nem tart a nullához, miközben z tart a-hoz.

(v) Bármily meglepő, a konjugálás, a Re és az Im, mint $\mathbb{C} \to \mathbb{C}$ függvények sehol sem differenciálhatók.

Legyen $a \in \mathbb{C}$ tetszőleges és $\lambda \in \mathbb{R} \setminus \{0\}$. Ekkor

$$\lim_{\lambda \to 0} \frac{(a+i\lambda)^* - a^*}{(a+i\lambda) - a} = \lim_{\lambda \to 0} \frac{-i\lambda}{i\lambda} = -1,$$

viszont

$$\lim_{\lambda \to 0} \frac{(a+\lambda)^* - a^*}{(a+\lambda) - a} = \lim_{\lambda \to 0} \frac{\lambda}{\lambda} = 1,$$

vagyis nem létezhet a $\lim\limits_{z \to a} \dfrac{z^* - a^*}{z - a}$ határérték, ami azt jelenti, hogy a konjugálás sehol sem differenciálható.

Tekintve, hogy bármilyen $z \in \mathbb{C}$ esetén $\mathrm{Re}(z) = \dfrac{z + z^*}{2}$ és $\mathrm{Im}(z) = \dfrac{z - z^*}{2}$, a Re és az Im függvények függvények sem lehetnek differenciálhatók egyetlen pontban sem.

1.6. Az előbbi példában olyan $\mathbb{C} \to \mathbb{C}$ függvények szerepeltek, amelyek mindenütt folytonosak de sehol sem differenciálhatók. Bár kissé kevésbé egyszerűen, ilyen $\mathbb{R} \to \mathbb{R}$ függvény is megadható.

Jelölje $d(x)$ az x valós számnak a hozzá legközelebbi egész számtól való távolságát, azaz

$$d(x) := \min\{|x - n| \mid n \in \mathbb{Z}\} = d(x, \mathbb{Z}).$$

Tudjuk, hogy $d : \mathbb{R} \to \mathbb{R}$ folytonos (Analízis III.B.8.13.), és $0 \leq d \leq \dfrac{1}{2}$. Ugyancsak folytonos $n \in \mathbb{N}$ esetén az

$$\mathbb{R} \to \mathbb{R}, \quad x \mapsto f_n(x) := d(10^n x) 10^{-n}$$

függvény, és $0 \leq f_n \leq 10^{-n}$. Ezért

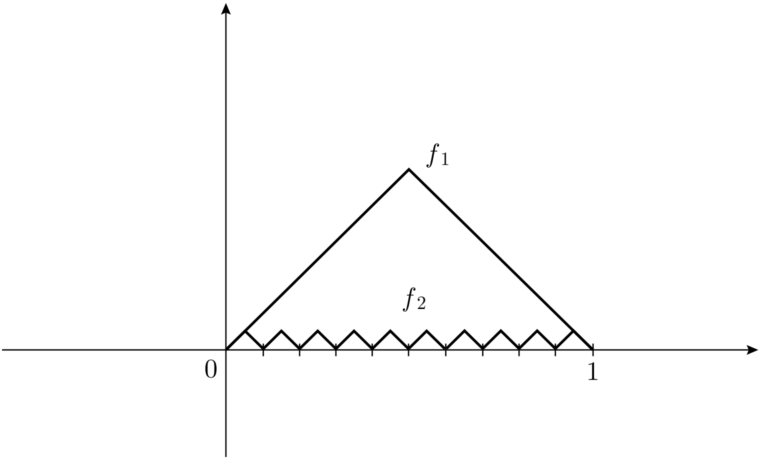
$$f := \sum_{n \in \mathbb{N}} f_n,$$

mint folytonos függvények egyenletesen konvergens sorának összege, szintén folytonos.

Megmutatjuk, hogy f sehol sem differenciálható. Nyilván elég ezt a $[0, 1]$ intervallumon megtenni, hiszen a függvény 1 szerint periodikus.

Legyen $a \in [0, 1]$, és állítsuk elő tizedestört alakban:

$$a = \sum_{k \in \mathbb{N}} \frac{a_k}{10^k}.$$



1. ÁBRA

Ekkor $10^n a$ törtrésze

$$\sum_{k=1}^{\infty} \frac{a_{n+k}}{10^k} =: r_n(a),$$

és az

$$F := \left\{ r_n(a) \leq \frac{1}{2} \right\}$$

jelöléssel

$$f_n(a) = \begin{cases} r_n(a) 10^{-n} = \sum_{k=n+1}^{\infty} \frac{a_{n+k}}{10^k} & \text{ha } n \in F, \\ (1 - r_n(a)) 10^{-n} & \text{ha } n \notin F. \end{cases}$$

Definiáljuk a $(h_m)_{m \in \mathbb{N}}$ sorozatot így:

$$h_m := \begin{cases} -10^{-m} & \text{ha } a_m = 4 \text{ vagy } 9, \\ 10^{-m} & \text{egyébként.} \end{cases}$$

Nyilván $\lim_m h_m = 0$, és $10^n(a + h_m)$ törtrésze pontosan akkor kisebb 1/2-nél, amikor $10^n a$ törtrésze. Ezért

$$f_n(a + h_m) = \begin{cases} f_n(a) & \text{ha } m \leq n, \\ f_n(a) - h_m & \text{ha } m > n \in F, \\ f_n(a) + h_m & \text{ha } m > n \notin F. \end{cases}$$

Tehát

$$\frac{f(a + h_m) - f(a)}{h_m} = \frac{\sum_{n=1}^{m} (\pm h_m)}{h_m} = \sum_{n=1}^{m} \pm 1,$$

ahol $\pm$ itt azt jelenti, hogy valahány $+1$-et és -1-et – mindegy, hogy melyikből mennyit, – kell összegezni. Az utolsó összegnek nincs határértéke $m \to \infty$ esetén, ezért nincs határértéke $m \to \infty$ esetén a bal oldalon álló mennyiségnek sem, azaz f nem differenciálható a-ban.

1.7 Állítás *Legyenek $f : \mathbb{K} \rightarrowtail \mathbb{K}$, $g : \mathbb{K} \rightarrowtail \mathbb{K}^N$ függvények és $a \in \overset{\circ}{\mathrm{Dom}}(g\circ f)$. Ha f differenciálható a-ban, g pedig differenciálható $f(a)$-ban, akkor $g \circ f$ differenciálható a-ban és*

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

Bizonyítás Mindenek előtt megjegyezzük: ha $a \in \overset{\circ}{\mathrm{Dom}}(g\circ f)$, akkor $a \in \overset{\circ}{\mathrm{Dom}}(f)$, viszont $f(a)$ nem szükségképpen belső pontja $\mathrm{Dom}(g)$-nek; ez utóbbit megköveteljük azáltal, hogy feltesszük g differenciálhatóságát.

Definiáljunk két halmazt:

$$H := \{x \in \mathrm{Dom}(g \circ f) \mid f(x) \neq f(a)\}, \quad Q := \{x \in \mathrm{Dom}(g \circ f) \mid f(x) = f(a)\}.$$

Nyilvánvaló, hogy $Q \cap H = \emptyset$ és $Q \cup H = \mathrm{Dom}(g \circ f)$. Az a pont vagy csak az egyik halmaznak, vagy mindkettőnek torlódási pontja.

(i) Ha az a pont csak a Q torlódási pontja, akkor létezik olyan $G(a) \subset Q$ környezete, amelyre leszűkítve az f és a $g \circ f$ függvények állandók. Eszerint $g \circ f$ differenciálható az a pontban és $(g \circ f)'(a) = 0 = f'(a)$, vagyis az állítás igaz.

(ii) Ha az a pont csak a H torlódási pontja, akkor létezik olyan $G(a)$ környezete, hogy $G(a) \setminus \{a\} \subset H$, és f folytonossága miatt az $f(a)$ pont torlódási pontja lesz az $f[H]$ halmaznak. Legyen $x \in G(a) \setminus a$ és tekintsük a következő átalakítást:

$$\frac{(g \circ f)(x) - (g \circ f)(a)}{x - a} = \frac{g(f(x)) - g(f(a))}{f(x) - f(a)} \frac{f(x) - f(a)}{x - a}.$$

Mivel f differenciálható a-ban és g differenciálható $f(a)$-ban; f folytonos is a-ban, ezért $f(x)$ tart $f(a)$-hoz, miközben x tart a-hoz. Így a jobb oldal mindkét tényezőjének létezik határértéke $x \to a$ esetén, és a jobb oldal határértéke $g'(f(a))f'(a)$; tehát a bal oldalnak is van határértéke, ami azt jelenti, hogy $g \circ f$ differenciálható az a pontban, és a deriváltjára az állításban kimondott összefüggés igaz.

(iii) Legyen az a torlódási pontja mind a Q, mind a H halmaznak. Létezik-e a $\lim\limits_{x\to a} \frac{(g\circ f)(x)-(g\circ f)(a)}{x-a}$ határérték? A Q menti határérték nyilvánvalóan létezik és nulla. Az előző pont alapján a H menti határérték is létezik és egyenlő $g'(f(a))f'(a)$-val. Viszont az a torlódási pontja Q-nak, ezért szükségszerűen $f'(a) = 0$. Mindezeket összevetve, létezik a határérték és nulla, és így $0 = (g \circ f)'(a) = g'(f(a))f'(a)$.

Tekintve, hogy az összes lehetséges esetet megvizsgáltuk, az állítást bebizonyítottuk.

1.8. Állítás *Legyen az $f : \mathbb{K} \rightarrowtail \mathbb{K}$ injektív függvény differenciálható az a pontban. Ha*

(i) $f'(a) \neq 0$,
(ii) $f(a)$ *belső pontja* $\mathrm{Ran}(f)$*-nek,*
(iii) f^{-1} *folytonos* $f(a)$*-ban, folytonos,*
akkor f^{-1} *differenciálható* $f(a)$*-ban és*

$$(f^{-1})'(f(a)) = \frac{1}{f'(a)}.$$

Bizonyítás Vegyük egy $G(f(a)) \subset \mathrm{Ran}(f)$ környezetét $f(a)$-nak. Legyen $y \in G(f(a)) \setminus \{a\}$. Ekkor létezik egyetlen olyan $x \in \mathrm{Dom}(f) \setminus \{a\}$, hogy $y = f(x)$. Tekintsük a következő átalakítást:

$$\frac{f^{-1}(y) - f^{-1}(f(a))}{y - f(a)} = \frac{x - a}{f(x) - f(a)} = \frac{1}{\dfrac{f(x) - f(a)}{x - a}}.$$

Ha y tart $f(a)$-hoz, akkor f^{-1} $f(a)$-beli folytonossága miatt $f^{-1}(y) = x$ tart $f^{-1}(f(a)) = a$-hoz, amiből az következik, hogy a jobb oldalnak létezik a határértéke, hiszen a nevezőnek van határértéke és nem nulla. Ezek szerint viszont a bal oldalnak is létezik a határértéke, tehát az f^{-1} függvény differenciálható az $f(a)$ pontban, és deriváltjára az állított egyenlőség teljesül.

Megjegyzések 1. Az állításban lényeges szerepet játszott mind az (ii), mind az (iii) követelmény, amelyeket adott függvény esetén nem mindig könnyű ellenőrizni, teljesülnek-e. Valós függvénynél azonban magának a függvénynek a tulajdonságaiból tudjuk ezeket következtetni: ha $f : \mathbb{R} \rightarrowtail \mathbb{R}$ injektív függvény, és folytonos az értelmezési tartománya egy a belső pontjának egy környezetén, akkor $f(a)$ belső pontja $\mathrm{Ran}(f)$-nek, és f^{-1} folytonos $f(a)$-nak egy környezetén (Analízis II-I.A.30.3.); tehát ekkor f^{-1} differenciálható $f(a)$-ban.

2. Van olyan $f : \mathbb{R} \to \mathbb{R}$ injektív függvény, amely differenciálható az értelmezési tartományának valamely a belső pontjában de az a pont semmilyen környezetén sem folytonos, $f(a)$ nem belső pontja az f értékkészletének, és f^{-1} nem folytonos $f(a)$-ban.

Ilyen például a következő függvény:

$$f :]-1,1[\to \mathbb{R},\ x \mapsto \begin{cases} \frac{1}{n+1}\left(nx - \frac{1}{2^{n-1}}\right) & \text{ha } x \in \left]-\frac{1}{2^{n-1}}, -\frac{1}{2^n}\right] \quad (n \in \mathbb{N}) \\ 0 & \text{ha } x = 0 \\ \frac{1}{n+1}\left(nx + \frac{1}{2^{n-1}}\right) & \text{ha } x \in \left[\frac{1}{2^n}, \frac{1}{2^{n-1}}\right[\quad (n \in \mathbb{N}) \end{cases}.$$

1.9. Feladatok

1. Mi az egészrész-függvény deriváltja?

2. Igazoljuk, hogy $(\text{sign id}_{\mathbb{R}}^2)' = 2|\ \cdot\ |$.

3. Legyen $a, b \in \mathbb{K}^N$. Mi az $\mathbb{R} \to \mathbb{K}^N$, $t \mapsto ta + b$ és $t \mapsto t^2a + b$ függvények deriváltja?

4. Legyen $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ és $\phi : \mathbb{K} \rightarrowtail \mathbb{K}$. Lehet-e ϕf differenciálható, ha
(i) sem ϕ sem f nem differenciálható?
(ii) ϕ és f közül az egyik differenciálható, a másik nem?

5. Legyen $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ és $\phi : \mathbb{K} \rightarrowtail \mathbb{K}$. Tegyük fel, hogy ϕ folytonos a-ban, f differenciálható a-ban és $f(a) = 0$. Bizonyítsuk be, hogy ϕf differenciálható a-ban és $(\phi f)'(a) = \phi(a)f'(a)$.

6. Legyen $f : \mathbb{R} \rightarrowtail \mathbb{C}$ differenciálható függvény. Mutassuk meg, hogy ekkor f^* is differenciálható és $(f^*)' = (f')^*$.

7. (i) Legyenek $f, g : \mathbb{R} \rightarrowtail \mathbb{K}^N$ differenciálható függvények.Igazoljuk, hogy ekkor $< g, f >: \mathbb{R} \rightarrowtail \mathbb{K}$ is differenciálható, és

$$< g, f >' = < g', f > + < g, f' > .$$

(ii) Hozzunk példát arra, hogy $f, g : \mathbb{C} \rightarrowtail \mathbb{C}^N$ differenciálhatók, de $< g, f >$ nem az.

8. Mutassuk meg, hogy a $\chi_{\mathbb{Q}}\text{id}_{\mathbb{R}}^2$ függvény csak egy pontban differenciálható!

9. Igazoljuk, hogy ha $f : \mathbb{C} \rightarrowtail \mathbb{C}$ differenciálható és $\text{Ran}(f) \subset \mathbb{R}$, akkor f konstans függvény.

10. Legyenek $f, g : \mathbb{R} \rightarrowtail \mathbb{R}^3$ differenciálható függvények. Mutassuk meg, hogy a pontonként értelmezett vektoriális szorzatuk, $f \times g$ is differenciálható és

$$(f \times g)' = f' \times g + f \times g'.$$

11. Igazoljuk, hogy az $L : \mathbb{K} \rightarrowtail \mathbb{K}^{M\times N}$ (mátrix értékű függvény) pontosan akkor differenciálható, ha minden $u \in \mathbb{K}^M$ esetén az $Lu : \mathbb{K} \rightarrowtail \mathbb{K}^N$, $x \mapsto L(x)u$ függvény differenciálható.

12. Legyen $L : \mathbb{K} \rightarrowtail \mathbb{K}^{M\times N}$ és $f : \mathbb{K} \rightarrowtail \mathbb{K}^M$ differenciálható. Bizonyítsuk be, hogy $Lf : \mathbb{K} \rightarrowtail \mathbb{K}^N$, $x \mapsto L(x)f(x)$ is differenciálható, és $(Lf)' = L'f + Lf'$.

2. Monotonitás, szélsőérték

2.1. Ebben a fejezetben végig valós változójú és valós értékű függvényekről lesz szó. Emlékeztetünk a következő már ismert fogalmakra:

(i) Az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvény **monoton növekszik (csökken)**, ha bármely $x, y \in \mathrm{Dom}(f)$, $x < y$ esetén

$$f(x) \leq f(y) \qquad \Big(f(x) \geq f(y)\Big).$$

(ii) Az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvény **szigorúan monoton növekszik (csökken)**, ha bármely $x, y \in \mathrm{Dom}(f)$, $x < y$ esetén

$$f(x) < f(y) \qquad \Big(f(x) > f(y)\Big).$$

Továbbá azt mondjuk, hogy f (szigorúan) monoton növekszik (csökken) az értelmezési tartományának egy A részhalmazán, ha $f|_A$ (szigorúan) monoton növekszik (csökken).

(iii) Az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvénynek **maximuma (minimuma)** van az $a \in \mathrm{Dom}(f)$ pontban, ha bármely $x \in \mathrm{Dom}(f)$ esetén

$$f(x) \leq f(a) \qquad \Big(f(x) \geq f(a)\Big).$$

(iv) Az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvénynek **lokális maximuma (minimuma)** van az $a \in \mathrm{Dom}(f)$ pontban, ha létezik olyan $G(a)$ környezete a-nak, hogy bármely $x \in G(a) \cap \mathrm{Dom}(f)$ esetén

$$f(x) \leq f(a) \qquad \Big(f(x) \geq f(a)\Big).$$

A maximumot és a minimumot közösen **szélsőértéknek** (**extrémumnak**) nevezzük. (v) **Szigorú (lokális) szélsőértékről** beszélünk, ha a fenti pontokban a $\leq$ és $\geq$ relációk helyett a szigorúbb $>$ illetve $<$ egyenlőtlenségek teljesülnek, persze $x \neq a$ esetén.

2.2. A derivált függvény definíciójából azonnal adódik a következő fontos eredmény.

Állítás *Legyen $I \subset \mathbb{R}$ intervallum. Ha az $f : I \to \mathbb{R}$ függvény differenciálható, és*

(i) monoton növekszik (csökken), akkor $f' \geq 0 \quad (f' \leq 0)$;
(ii) konstans, akkor $f' = 0$.

2.3. Az előbbi állítás egy kis módosítással fordítva is igaz. Értelem szerint ide kívánkozik, de bizonyításához felhasználjuk a Lagrange-féle középértéktételt, amelyet csak a következő fejezetben bizonyítunk be; kérjük az Olvasót, győződjön meg később arról, hogy a középértéktétel bizonyításához a mostani eredményünket nem használjuk fel, tehát nem követünk el hibát.

Állítás *Legyen $I \subset \mathbb{R}$ nyílt intervallum. Ha az $f : I \to \mathbb{R}$ függvénydifferenciálható és*

(i) $f' > 0 \quad (f' < 0)$, akkor f szigorúan monoton növekszik (csökken) ;
(ii) $f' = 0$, akkor f konstans.

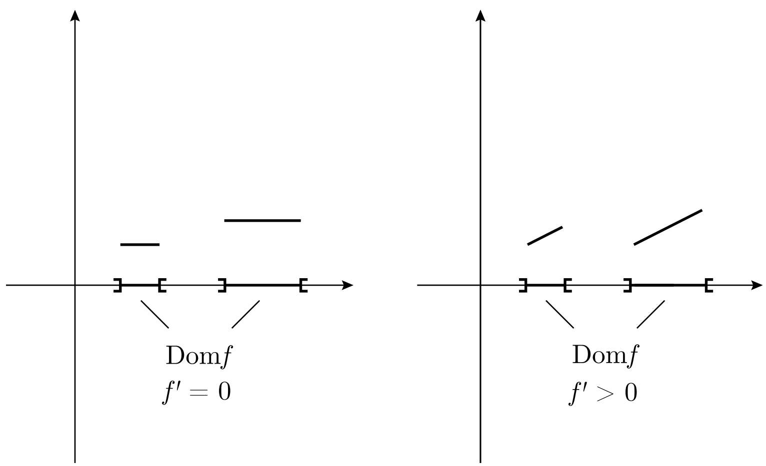
Bizonyítás Legyen x és y két olyan pontja az adott intervallumnak, hogy $x < y$. Az f függvényre alkalmazva a Lagrange-féle középértéktételt az [x,y] intervallumon azt kapjuk, hogy létezik olyan $z \in]x, y[$, amelyre

$$f(y) - f(x) = f'(z)(y - x).$$

(i) A fenti egyenlőséget tekintve nyilvánvaló, hogy $f'(z) > 0 \quad (f'(z) < 0)$ miatt $f(y) > f(x) \quad (f(y) < f(x))$. Az x és y pontokat tetszőlegesen választhattuk I-ből, ezért az f függvény valóban szigorúan monoton növekszik (csökken) ezen az intervallumon.

(ii) $f'(z) = 0$ miatt $f(x) = f(y)$, tehát f konstans függvény, hiszen x és y tetszőleges pontjai voltak I-nek. ■

Az Olvasóra bízzuk annak az egyszerű ténynek az ellenőrzését, hogy ha az iménti állítás (i) pontjában a $>$ illetve a $<$ reláció helyett a gyengébb $\geq$ illetve a $\leq$ relációt tételezzük fel, akkor a szigorú monotonitással szemben csak a monotonitás fog teljesülni (pl. állandó függvény esetében).



2. Ábra

Figyeljünk arra, hogy ezek az állítások csak akkor igazak minden bizonnyal, ha I intervallum (azaz összefüggő halmaz). Ha f értelmezési tartománya nem intervallum, $f' > 0$ illetve $f' = 0$ előfordulhat anélkül, hogy f szigorúan monoton illetve konstans volna. A 2. ábra rajzai jól szemléltetik ezt.

A fenti tételek némiképp általánosabb formában is megfogalmazhatók, mivel $f : I \to \mathbb{R}$ differenciálható függvényre vonatkoznak, és emiatt az I intervallum nyílt kell hogy legyen. Könnyen látható, hogy ezek az állítások igazak maradnak, ha egy $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvényt veszünk, amely differenciálható egy $I \subset \mathrm{Dom}(f)$ intervallumon és $f|_I$-re mondjuk ki őket. Így az I intervallum nem csak nyílt lehet és ezt a tételek alkalmazásakor a későbbiekben (pl. a 7. fejezetben) ki is fogjuk használni.

2.4. Állítás *Ha az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvény differenciálható az a pontban és $f'(a) > 0$ $(f'(a) < 0)$, akkor létezik a-nak olyan $G(a)$ környezete, ahol tetszőleges $x, y \in G(a)$ pontokra $x < a < y$ esetén*

$$f(x) < f(a) < f(y) \qquad \Big(f(x) > f(a) > f(y)\Big).$$

Bizonyítás f differenciálható a-ban, tehát a belső pontja a függvény értelmezési tartományának, valamint létezik a $\lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a}$ határérték és pozitív (negatív). Ezek szerint van olyan $G(a) \subset \mathrm{Dom}(f)$környezete a-nak, hogy bármely $x \in G(a)$, $x \neq a$ esetén $\dfrac{f(x) - f(a)}{x - a} > 0$ (< 0), ami pontosan akkor teljesül, ha a számláló és a nevező azonos (különböző) előjelű. A bizonyítást ezzel be is fejeztük.

Megjegyzés Itt, az előzőtől eltérően csak egyetlen pontban van előírva a derivált létezése és annak előjele, nem egy egész intervallumon. Felhívjuk a figyelmet, hogy az állításban megfogalmazottak nem jelentik azt, hogy f (szigorúan) monoton az a egy $G(a)$ környezetében, ugyanis az eredmény mindig a-t és egy $G(a)$-beli elemet hasonlít össze, nem pedig két akármilyen $G(a)$-belit. Majd később, amikor többet tudunk a differenciálhatóságról, meggyőződhetünk arról, hogy

$$x \mapsto \begin{cases} x + x^2 \sin(1/x) & \text{ha } 0 \neq x, \\ 0 & \text{ha } x = 0 \end{cases}$$

ellenpéldát szolgáltat a nullában.

2.5. Állítás *Ha az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvény differenciálható a-ban és ott lokális szélsőértéke van, akkor $f'(a) = 0$.*

Bizonyítás Ha $f'(a) \neq 0$, tehát $f'(a) > 0$ vagy $f'(a) < 0$, akkor az előző állítás miatt f-nek semmiképpen sem lehet a-ban lokális szélsőértéke. ■

Vigyázzunk arra, hogy az állítás fordítottja nem igaz: $f'(a) = 0$ nem jelenti feltétlenül azt, hogy f-nek a-ban lokális szélsőértéke van. Igen egyszerű az ellenpélda: $\mathrm{id}_{\mathbb{R}}^3$-nek nincs szélsőértéke a nullában, viszont ott a deriváltja nulla.

2.6. Feladatok

1. Adjunk meg olyan I intervallumot és $f : I \to \mathbb{R}$ differenciálható függvényt, amely szigorúan monoton növekszik, mégis van olyan $x \in I$, hogy $f'(x) = 0$.

2. Lehet-e lokális szélsőértéke a következő függvényeknek:

$$\mathrm{id}_{\mathbb{R}}^3 + \mathrm{id}_{\mathbb{R}}; \qquad 2\mathrm{id}_{\mathbb{R}}^4 + 6\mathrm{id}_{\mathbb{R}} + 2; \qquad \frac{\mathrm{id}_{\mathbb{R}}^2 + 3\mathrm{id}_{\mathbb{R}}}{2\mathrm{id}_{\mathbb{R}} + 4}?$$

3. Legyen $a \in \mathbb{R}$ és $f, g : [a, \infty[\to \mathbb{R}$ folytonos függvények, amelyek differenciálhatók az értelmezési tartományuk belsején. Tegyük föl, hogy $f' \le g'$ és $f(a) \le g(a)$. Mutassuk meg, hogy $f \le g$.

4. Nevezzük az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvényt az értelmezési tartományának a belső pontjában lokálisan monoton növőnek (fogyónak), ha létezik a-nak olyan $G(a)$ környezete, amelynek tetszőleges x, y pontjára $x<a<y$ esetén $f(x) \le f(a) \le f(y)$ $(f(x) \ge f(a) \ge f(y))$ teljesül; ha $\le$ helyett $<$ áll, akkor szigorú lokális monotonitásról beszélünk (idézzük fel ezzel kapcsolatban a 2.4. megjegyzését). Bizonyítsuk be, hogy ha f egy intervallum minden pontjában (szigorúan) lokálisan monoton növő (fogyó), akkor (szigorúan) monoton nő (fogy) az intervallumon. (Legyen x és y az intervallum eleme, $x < y$. Az $[x, y]$ kompakt halmazt lefedik a pontjainak azok a környezetei, amelyekben a lokális monotonitás egyenlőtlenségei teljesülnek; ezért véges sok ilyen köryezet is lefedi. Véges sokat "lépegetve" megállapíthatjuk, hogy $f(x) \le f(y)$ vagy $f(x) < f(y)$ stb.)

3. Középértéktételek

3.1. Állítás (***Rolle-tétel***) *Ha $f : [a, b] \to \mathbb{R}$ folytonos, differenciálható $]a, b[$-n és $f(a) = f(b)$, akkor létezik olyan $c \in]a, b[$, hogy*

$$f'(c) = 0.$$

Bizonyítás Az f függvény a Weierstrass-tétel szerint fölveszi az $[a, b]$ intervallumon a maximumát és a minimumát is. Ha mindkettőt a végpontokban, akkor f szükségszerűen konstans függvény és $f' = 0$ az egész intervallumon.

Ha f a minimumot vagy a maximumot az $[a, b]$ valamely c belső pontjában veszi fel, akkor a c pontban f-nek lokális szélsőértéke van, következésképpen $f'(c) = 0$.

3.2. Állítás *(**Cauchy-tétel**) Ha $f, g : [a, b] \to \mathbb{R}$ folytonosak és differenciálhatóak $]a, b[$-n, akkor létezik olyan $c \in]a, b[$, hogy*

$$\big(f(b) - f(a)\big)g'(c) = \big(g(b) - g(a)\big)f'(c).$$

Bizonyítás Az

$$F := \big(g(b) - g(a)\big)f - \big(f(b) - f(a)\big)g$$

függvény eleget tesz az előző állítás feltételeinek $[a, b]$-n, ezért létezik olyan c pontja az $]a, b[$ intervallumnak, hogy

$$F'(c) = \big(g(b) - g(a)\big)f'(c) - \big(f(b) - f(a)\big)g'(c) = 0.$$

A kapott egyenlőséget megfelelően átrendezve pontosan a bizonyítandót kapjuk. ■

Ha g' sehol sem nulla $]a, b[$-n, akkor a Rolle-tétel szerint $g(b) \neq g(a)$, és így a Cauchy-tételt

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}$$

formában is írhatjuk.

3.3. Állítás *(**Lagrange-tétel**) Ha $f : [a, b] \to \mathbb{R}$ folytonos és differenciálható $]a, b[$-n, akkor létezik olyan $c \in]a, b[$, hogy*

$$f(b) - f(a) = f'(c)(b - a).$$

Bizonyítás Ez a tétel az előző állítás következménye, ha a $g := \mathrm{id}_{[a,b]}$ választással élünk. ■

A Lagrange-tételt így szemléltethetjük: az adott feltételek mellett van az $[a, b]$ intervallum belsejében egy pont, ahol az f függvény grafikonjának érintője párhuzamos az $(a, f(a))$ és $(b, f(b))$ pontokon áthaladó szelővel.

A Lagrange-féle középértéktétel igen hasznosnak bizonyul sok alkalmazásban, azonban leginkább egy kicsit más formában jelenik meg, a következőképpen.

Legyen $f : \mathbb{R} \rightarrowtail \mathbb{R}$ differenciálható az x egy környezetében. Ha h olyan valós szám, hogy $x + h$ benne van ebben a környezetben, akkor létezik $\theta_{x,h} \in]0, 1[$ úgy, hogy

$$f(x + h) - f(x) = f'(x + \theta_{x,h}h).$$

Ugyanis, ha $h > 0$, akkor $f|_{[x,x+h]}$ eleget tesz a Lagrange-féle középértéktétel követelményeinek, amelyek biztosítják az $]x, x + h[$ egy c elemét a megfelelő tulajdonságokkal; $\theta_{x,h} := \dfrac{c - x}{h}$. Ha viszont $h < 0$, akkor hasonlóan érvelhetünk az $[x + h, x]$ intervallumra.

Megemlítjük még, hogy a középértéktétel semmit sem mond az $(x,h) \mapsto \theta_{x,h}$ hozzárendelés tulajdonságairól (pl. folytonosság), tehát elképzelhető, hogy $\theta_{x,h}$ teljesen "kaotikusan" változik az x és h függvényében. Érdekes viszont, hogy az $(x,h) \mapsto f'(x+\theta_{x,h})h$ függvény már "szép", hiszen az $(x,h) \mapsto f(x+h) - f(x)$ függvénnyel azonos.

3.4. Mindeddig csak $\mathbb{R} \rightarrowtail \mathbb{R}$ függvényekről volt szó. $N > 1$ esetén $\mathbb{R} \rightarrowtail \mathbb{R}^N$ függvényekre, és akármilyen $N \in \mathbb{N}$ esetén $\mathbb{C} \rightarrowtail \mathbb{C}^N$ függvényekre nem igaz a Lagrange-féle középértéktétel. Megelőlegezve az elemi függvények deriváltjára vonatkozó ismereteket, íme az ellenpéldák:

(i) $f : \mathbb{R} \to \mathbb{R}^2, \quad x \mapsto (\cos 2\pi x, \sin 2\pi x)$.
Erre $f(1) - f(0) = (0,0)$, viszont $f'(x) = 2\pi(-\sin 2\pi x, \cos 2\pi x) \neq (0,0)$ minden $x \in \mathbb{R}$ esetén.

(ii) $f : \mathbb{C} \to \mathbb{C}, \quad z \mapsto \mathrm{Exp}(2\pi i z)$.
Erre $f(1) - f(0) = 0$, viszont $f'(z) = 2\pi i \mathrm{Exp}(2\pi i z) \neq 0$ minden $z \in \mathbb{C}$ esetén. Az általános, $\mathbb{K} \rightarrowtail \mathbb{K}^N$ függvényekre vonatkozó középértéktétel csak egyenlőtlenséget állít, de ez legalább olyan jó szolgálatokat tesz alkalmazásokban.

Állítás (***Általános középértéktétel***) *Legyen $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ és $x, y \in \mathbb{K}$ olyan hogy f differenciálható az $[x,y]$ szakaszon, vagyis az $\{x + t(y-x) \mid t \in [0,1]\}$ halmazon. Ekkor létezik olyan $z \in]x,y[$, hogy*

$$|f(y) - f(x)| \leq |f'(z)|\ |y-x|.$$

Bizonyítás Vegyük észre, hogy $f(x) = f(y)$ esetben az állítás nyilvánvalóan igaz. Ha $f(x) \neq f(y)$, akkor definiáljuk a $\phi : [0,1] \to \mathbb{R}$ folytonos függvényt a következő módon:

$$\phi(t) := \mathrm{Re}\langle f(y) - f(x), f(x + t(y-x)) - f(x)\rangle.$$

Nyilvánvaló, hogy $\phi(1) = |f(y) - f(x)|^2$, $\phi(0) = 0$, valamint ϕ differenciálható]0,1[-en és

$$\phi'(t) = \mathrm{Re}\langle f(y) - f(x), f'(x + t(y-x))(y-x)\rangle.$$

A ϕ függvényre alkalmazhatjuk a Lagrange-féle középértéktételt, miszerint létezik olyan $\tau \in]0,1[$, hogy

$$\phi(1) = \phi(1) - \phi(0) = \phi'(\tau)(1-0) = \phi'(\tau).$$

Legyen $z := x + \tau(y-x)$; világos, hogy $z \in]x,y[$.

A fentieket, valamint a Cauchy-egyenlőtlenséget felhasználva elvégezhetjük a következő átalakítást:

$$|f(y) - f(x)|^2 = \phi(1) = \phi'(\tau) = \mathrm{Re}\langle f(y) - f(x), f'(z)(y-x)\rangle \leq$$
$$\leq \left|\langle f(y) - f(x), f'(z)(y-x)\rangle\right| \leq |f(y) - f(x))\ |f'(z)|\ |y-x|,$$

amiből az $|f(y) - f(x)| > 0$ mennyiséggel osztással megkapjuk a kívánt egyenlőtlenséget.

3.5. Feladatok

1. Ha $\mathbb{R} \to \mathbb{R}^N$ differenciálható, minden komponensére igaz a Lagrange-féle középértéktétel (egyenlőség). Miért nem származtatható ebből egy megfelelő egyenlőség magára a függvényre?

2. Adjunk meg olyan nem konstans $\mathbb{R} \to \mathbb{R}^2$ függvényt, amelyre igaz a Lagrange-féle középértéktétel (egyenlőség).

3. Igazoljuk, hogy a Rolle-féle, a Cauchy-féle és a Lagrange-féle középértéktételek egyenértékűek!

4. Differenciálható függvények sorozata

4.1. Állítás *Legyen $H \subset \mathbb{K}$ konvex, korlátos, nyílt halmaz és $f_n : H \to \mathbb{K}^N$ ($n \in \mathbb{N}$) differenciálható függvények sorozata. Ha van olyan $x_0 \in H$, hogy az $\big(f_n(x_0)\big)_{n\in\mathbb{N}}$ sorozat konvergens, valamint a derivált függvények $(f_n')_{n\in\mathbb{N}}$ sorozata egyenletesen konvergens, akkor az $(f_n)_{n\in\mathbb{N}}$ függvénysorozat is egyenletesen konvergens, a határértéke differenciálható, és fennáll az*

$$\Big((u)\lim_n f_n\Big)' = (u)\lim_n f_n'$$

egyenlőség (ahol az (u) szimbólum az egyenletes konvergenciára utal).

Bizonyítás Mivel H nyílt és korlátos, az átmérője – amelyet D-vel jelölünk – nagyobb nullánál és kisebb végtelennél.

Bármilyen $\epsilon > 0$ számhoz létezik olyan $n_\epsilon \in \mathbb{N}$, hogy minden $m, n \geq n_\epsilon$ természetes szám esetén

$$|f_n(x_0) - f_m(x_0)| < \frac{\epsilon}{2},$$

hiszen az $\big(f_n(x_0)\big)_{n\in\mathbb{N}}$ sorozat konvergens, tehát Cauchy-féle is.

Az $(f_n')_{n\in\mathbb{N}}$ sorozat egyenletes konvergenciája miatt létezik olyan $k_\epsilon \in \mathbb{N}$, hogy minden $m, n \geq k_\epsilon$ természetes szám és minden $z \in H$ esetén

$$|f_n'(z) - f_m'(z)| < \frac{\epsilon}{2D}.$$

Az $(f_n - f_m)$ függvényre az $[x, y]$ szakaszon – ahol $x, y \in H$ tetszőleges – alkalmazhatjuk a középértéktételt, miszerint létezik olyan $z \in]x, y[$, hogy

$$|f_n(x) - f_m(x) - f_n(y) + f_m(y)| \leq |f_n'(z) - f_m'(z)|\ |x - y|.$$

Mivel $|x-y| \le D$, minden $m, n \ge k_\epsilon$ esetén

$$|f_n(x) - f_m(x) - f_n(y) + f_m(y)| < \frac{\epsilon}{2}.$$

A fentieket összevetve, ha m és n nagyobbak a $\max\{n_\epsilon, k_\epsilon\}$ küszöbindexnél, akkor a H minden x elemére

$$\begin{aligned}|f_n(x) - f_m(x)| \le |f_n(x) - f_m(x) - f_n(x_0) + f_m(x_0)| + |f_n(x_0) - f_m(x_0)| < \\ < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,\end{aligned}$$

vagyis az $(f_n)_{n\in\mathbb{N}}$ sorozat egyenletesen konvergens.

Vezessük be az $f := (\mathrm{u})\lim\limits_n f_n$ jelölést. Az f függvény, mint folytonos függvények egyenletes határértéke, folytonos lesz.

Most rögzítsük a H egy x elemét, és definiáljuk a következő függvényeket:

$$\varphi_n : H \setminus \{x\} \to \mathbb{K}^N, \quad y \mapsto \frac{f_n(y) - f_n(x)}{y - x},$$

$$\varphi : H \setminus \{x\} \to \mathbb{K}^N, \quad y \mapsto \frac{f(y) - f(x)}{y - x}.$$

Az látszik, hogy a $(\varphi_n)_{n\in\mathbb{N}}$ sorozat pontonként konvergál a φ függvényhez, de ez nekünk most kevés lenne. Belátjuk, hogy egyenletesen is konvergál. Válasszunk egy $y \in H \setminus \{x\}$ pontot; a következő átalakítások és a középértéktétel szerint

$$\begin{aligned}&|\varphi_n(y) - \varphi_m(x)| = \\ &= \frac{|f_n(y) - f_n(x) - f_m(y) + f_m(x)|}{|y-x|} = \frac{|(f_n - f_m)(y) - (f_n - f_m)(x)|}{|y-x|} \le \\ &\le \frac{|(f_n - f_m)'(z)|\;|y-x|}{y-x} \le \frac{\epsilon}{2D}\end{aligned}$$

valamely $z \in]x, y[$ pontra, ha $m, n \ge k_\epsilon$. Ez pontosan azt jelenti, hogy a $(\varphi_n)_{n\in\mathbb{N}}$ sorozat egyenletesen konvergál a φ függvényhez.

Minden $n \in \mathbb{N}$ esetén az f_n függvény differenciálható, ezért minden n-re létezik $\lim\limits_{y\to x} \varphi(y) = f_n'(x)$. Emiatt létezik (Analízis III.A.31.1.)

$$f'(x) = \lim_{y\to x} \varphi(y) = \lim_{y\to x}(\lim_n \varphi_n(y)) = \lim_n(\lim_{y\to x} \varphi_n(y)) = \lim_n f_n'(x).$$

Az x elemet H-ból tetszőlegesen választottuk, az $(f_n')_{n\in\mathbb{N}}$ sorozat pedig egyenletesen konvergens, tehát

$$f' = \Big((\mathrm{u})\lim_n f_n\Big)' = (\mathrm{u})\lim_n f_n',$$

így a bizonyítást befejeztük.

4.2. Az előbbi eredményünk legtöbbet használt következményét mondjuk ki most.

Állítás *Legyen $H \subset \mathbb{K}$ nyílt halmaz és $f_n : H \to \mathbb{K}^N \quad (n \in \mathbb{N})$ differenciálható függvények sorozata. Ha az $(f_n)_{n\in\mathbb{N}}$ függvénysorozat konvergens H-n, és a derivált függvények $(f_n')_{n\in\mathbb{N}}$ sorozata lokálisan egyenletesen konvergens H-n, akkor az $(f_n)_{n\in\mathbb{N}}$ sorozat határértéke differenciálható, és fennáll a*

$$\left(\lim_n f_n\right)' = \lim_n f_n'$$

egyenlőség.

Bizonyítás Legyen x_0 a H tetszőleges pontja; ekkor az $\big(f_n(x_0)\big)_{n\in\mathbb{N}}$ sorozat konvergens. A H nyíltsága miatt létezik olyan $G(x_0) \subset H$ konvex környezete x_0-nak, hogy ott az $(f_n')_{n\in\mathbb{N}}$ függvénysorozat egyenletesen konvergens, így ebben a környezetben nyilván teljesülnek az előző állítás feltételei; $(f_n)_{n\in\mathbb{N}}$ egyenletesen konvergens $G(x_0)$-on, és a határértékfüggvénye differenciálható $G(x_0)$-on, a differenciálás és a limesz felcserélhető.

Tekintve, hogy x_0 a H tetszőleges eleme, ezzel az állítást bebizonyítottuk.

4.3. Függvénysorokra az előző állítást így lehet megfogalmazni:

Állítás *Legyen $H \subset \mathbb{K}$ nyílt halmaz és $g_n : H \to \mathbb{K}^N \quad (n \in \mathbb{N})$ differenciálható függvények sorozata. Ha a $\sum\limits_{n\in\mathbb{N}} g_n$ függvénysor konvergens H-n, és a derivált függvényekből képezett $\sum\limits_{n\in\mathbb{N}} g_n'$ függvénysor lokálisan egyenletesen konvergens H-n, akkor a $\sum\limits_{n\in\mathbb{N}} g_n$ függvénysor határfüggvénye differenciálható, valamint fennáll a*

$$\left(\sum_{n\in\mathbb{N}} g_n\right)' = \sum_{n\in\mathbb{N}} g_n'$$

egyenlőség.

4.4. A függvénysorokra vonatkozó, előző állítás fontos eredményt szolgáltat hatványsorokra.

Állítás *Legyen a* $\sum_{n=0}^{\infty} c_n(\mathrm{id}_{\mathbb{K}} - a)^n$ *hatványsor konvergenciasugara* $R > 0$. *Ekkor a – $G_R(a)$-n definiált – összegfüggény differenciálható és*

$$\left(\sum_{n=0}^{\infty} c_n(\mathrm{id}_{\mathbb{K}} - a)^n\right)' = \sum_{n=0}^{\infty} c_n n(\mathrm{id}_{\mathbb{K}} - a)^{n-1}.$$

Bizonyítás Mint tudjuk, egy hatványsor lokálisan egyenletesen konvergens a konvergenciakörén. Az egyes tagok

$$(c_n(\mathrm{id}_{\mathbb{K}} - a)^n)' = c_n n(\mathrm{id}_{\mathbb{K}} - a)^{n-1}$$

deriváltjaiból képezett hatványsor konvergenciasugara a gyökkritérium alapján szintén R, hiszen

$$\limsup_n \sqrt[n]{|c_n|n} = \limsup_n \left(\sqrt[n]{|c_n|}\,\sqrt[n]{n}\right) =$$
$$= \left(\limsup_n \sqrt[n]{|c_n|}\right) \lim_n \left(\sqrt[n]{n}\right) = \limsup_n \left(\sqrt[n]{|c_n|}\right).$$

Így a 4.3. állítás minden feltétele teljesül a $G_R(a)$ nyílt halmazon, amely pontosan a szóban forgó összegfüggvények értelmezési tartománya. Ezzel a bizonyítást be is fejeztük.

4.5. Feladatok

1. Fogadjuk el, hogy $\cos' = -\sin$ – ezt később bebizonyítjuk –, és mutassuk meg, hogy

(i) az $f_n(x) := \dfrac{1}{\sqrt{n}}\cos(nx)$ $(x \in \mathbb{R})$ differenciálható függvényekből álló sorozat egyenletesen konvergens – tehát nem a derivált függvények sorozata egyenletesen konvergens! –, és nem áll fenn a $(\lim_n f_n)' = \lim_n f_n'$ egyenlőség;

ii) az $f_n(x) := \sqrt{n}\cos(x/n)$ $(x \in \mathbb{N})$ függvénysorozat nem konvergens, azonban $(f_n')_{n\in\mathbb{N}}$ egyenletesen konvergens (vagyis állításunkból nem hagyható el az $\big(f_n(x_0)\big)_{n\in\mathbb{N}}$ sorozat konvergenciája legalább egy x_0 pontra).

2. A differenciálás és a határérték felcseréléséhez az $(f_n')_{n\in\mathbb{N}}$ lokális egyenletes kovergenciája mellett nyilván nem elég, hogy az $(f_n)_{n\in\mathbb{N}}$ függvénysorozat csak egyetlen pontban legyen konvergens; lehet-e gyengébb feltételt megkövetelni, mint ami a 4.2. állításban szerepel, nevezetesen, hogy a függvénysorozat mindenütt legyen konvergens?

5. Elemi függvények differenciálhatósága

5.1. A 4.4. állítás birtokában megállapíthatjuk, hogy az egész komplex számsíkon konvergens hatványsorokkal definiált

$$\mathrm{Exp} = \sum_{n=0}^{\infty} \frac{\mathrm{id}_{\mathbb{C}}^n}{n!},$$

$$\mathrm{Cos} = \sum_{n=0}^{\infty} (-1)^n \frac{\mathrm{id}_{\mathbb{C}}^{2n}}{(2n)!}, \qquad \mathrm{Sin} = \sum_{n=0}^{\infty} (-1)^n \frac{\mathrm{id}_{\mathbb{C}}^{2n+1}}{(2n+1!)},$$

$$\mathrm{Ch} = \sum_{n=0}^{\infty} \frac{\mathrm{id}_{\mathbb{C}}^{2n}}{(2n)!}, \qquad \mathrm{Sh} = \sum_{n=0}^{\infty} \frac{\mathrm{id}_{\mathbb{C}}^{2n+1}}{(2n+1!)}$$

elemi függvények diferenciálhatók, és

$$\mathrm{Exp}' = \mathrm{Exp}, \quad \mathrm{Cos}' = -\mathrm{Sin}, \quad \mathrm{Sin}' = \mathrm{Cos}, \quad \mathrm{Ch}' = \mathrm{Sh}, \quad \mathrm{Sh}' = \mathrm{Ch}.$$

Ezeknek a függvényeknek az $\mathbb{R}$-re való leszűkítése szintén az egész valós egyenesen konvergens hatványsorral áll elő, tehát az exp, cos, sin, ch és sh függvények is differenciálhatók, és

$$\exp' = \exp, \quad \cos' = -\sin, \quad \sin' = \cos, \quad \mathrm{ch}' = \mathrm{sh}, \quad \mathrm{sh}' = \mathrm{ch}.$$

Ezekből és a függvények hányadosának deriváltjára vonatkozó formulából kaphatjuk, hogy

$$\begin{aligned}
\mathrm{tg}' &= \left(\frac{\sin}{\cos}\right)' = \frac{\cos^2 + \sin^2}{\cos^2} = \frac{1}{\cos^2} = 1 + \mathrm{tg}^2,\\
\mathrm{ctg}' &= \left(\frac{\cos}{\sin}\right)' = \frac{-\sin^2 - \cos^2}{\sin^2} = -\frac{1}{\sin^2} = -(1 + \mathrm{ctg}^2),\\
\mathrm{th}' &= \left(\frac{\mathrm{sh}}{\mathrm{ch}}\right)' = \frac{1}{\mathrm{ch}^2} = 1 - \mathrm{th}^2,\\
\mathrm{cth}' &= \left(\frac{\mathrm{ch}}{\mathrm{sh}}\right)' = -\frac{1}{\mathrm{sh}^2} = 1 - \mathrm{cth}^2.
\end{aligned}$$

5.2. A deriváltakra vonatkozó ismereteink alapján – a 2.3. és az 1.8. állításra támaszkodva – az elemi függvényeknek, illetve bizonyos leszűkítéseiknek inverzét és azok differenciálhatóságát tudjuk tárgyalni.

(i) $\exp' = \exp$ és $\exp > 0$, így a valós exponenciális függvény szigorúan monoton növekszik, tehát injektív (ezt egyébként már korábbról is tudjuk); inverze, a $\log := \exp^{-1} : \mathbb{R}^+ \to \mathbb{R}$ függvény differenciálható és

$$\log' = (\exp^{-1})' = \frac{1}{\exp' \circ \log} = \frac{1}{\exp \circ \log} = \frac{1}{\mathrm{id}_{\mathbb{R}^+}}.$$

(ii) Ha $a \in \mathbb{R}^+$, akkor az $\exp_a := \exp \circ(\log(a)\mathrm{id}_{\mathbb{R}}) : \mathbb{R} \to \mathbb{R}^+$, $x \mapsto a^x$, és a $\log_a := \exp_a^{-1} : \mathbb{R}^+ \to \mathbb{R}$ függvények szintén differenciálhatók:

$$\exp_a' = \big(\exp' \circ(\log(a)\mathrm{id}_{\mathbb{R}})\big) \log(a) = \log(a) \exp_a,$$

$$\log_a' = \frac{1}{\exp_a' \circ \log_a} = \frac{1}{\log(a)} \frac{1}{\exp_a \circ \log_a} = \frac{1}{\log(a)} \frac{1}{\mathrm{id}_{\mathbb{R}^+}}.$$

(iii) Tetszőleges $\alpha \in \mathbb{R}$ esetén értelmezzük az

$$\mathrm{id}_{\mathbb{R}^+}^{\alpha} := \exp \circ(\alpha \log) : \mathbb{R}^+ \to \mathbb{R}^+$$

függvényt. Ez is differenciálható, és

$$(\mathrm{id}_{\mathbb{R}^+}^{\alpha})' = (\exp' \circ(\alpha \log))\, \alpha \frac{1}{\mathrm{id}_{\mathbb{R}^+}} = \alpha\big(\exp \circ(\alpha \log)\big) \frac{1}{\exp \circ \log} = $$
$$= \alpha\big(\exp \circ((\alpha - 1) \log)\big) = \alpha \mathrm{id}_{\mathbb{R}^+}^{\alpha-1}.$$

(iv) A cos és a sin függvénynek a követekező pontban tárgyalt tulajdonságai alapján a

$$]-\pi, \pi] \to \mathbb{S}^1, \quad \alpha \mapsto e^{i\alpha} = \cos \alpha + i \sin \alpha$$

leképezés bijekció. Ezért és az exp függvény tulajdonságai miatt az Exp függvény injektív az $\mathbb{R} + i]-\pi, \pi]$ halmazon; a

$$\mathrm{Log} := \big(\mathrm{Exp}|_{\mathbb{R}+i]-\pi,\pi]}\big)^{-1} : \mathbb{C} \setminus \{0\} \to \mathbb{R} + i]-\pi, \pi],$$
$$z \mapsto \log |z| + i\mathrm{sign}(\mathrm{Im}(z)) \arccos \frac{\mathrm{Re}(z)}{|z|}$$

komplex logaritmusfüggvény folytonos $\mathbb{C} \setminus \mathbb{R}_0^-$-on, és ezért (lévén differenciálható függvény inverze) ott differenciálható is, és

$$\mathrm{Log}' = \frac{1}{\mathrm{Exp}' \circ \mathrm{Log}|_{\mathbb{C}\setminus\mathbb{R}_0^-}} = \frac{1}{\mathrm{id}_{\mathbb{C}\setminus\mathbb{R}_0^-}}.$$

5.3. Nézzük meg a valós trigonometrikus függvények inverzeit!

(i) Tudjuk, hogy $\sin(-\pi/2) = -1$, $\sin(\pi/2) = 1$ és $\sin' x = \cos x > 0$ ha x benne van a $]-\pi/2, \pi/2[$ intervalumben. Ezért a sin függvény szigorúan monoton nő, tehát injektiv a $[-\pi/2, \pi/2]$ intervallumon. Inverze,

$$\arcsin := \left(\sin|_{[-\pi/2,\pi/2]}\right)^{-1} : [-1,1] \to [-\pi/2, \pi/2]$$

folytonos (Analízis III. A.30.3.), és értelmezési tartománya belsejében differenciálható:

$$\arcsin' = \frac{1}{\cos \circ \arcsin} = \frac{1}{\sqrt{1 - \sin^2 \circ \arcsin}} = \frac{1}{\sqrt{1 - \mathrm{id}_{\mathbb{R}}^2}}.$$

Teljesen hasonlóan értelmezhetjük az

$$\arccos = \left(\cos|_{[0,\pi]}\right)^{-1} : [-1,1] \to [0,\pi]$$

függvényt, amelyre

$$\arccos' = \frac{1}{-\sin \circ \arccos} = -\frac{1}{\sqrt{1 - \cos^2 \circ \arccos}} = -\frac{1}{\sqrt{1 - \mathrm{id}_{\mathbb{R}}^2}}.$$

(ii) A tg függvény deriváltja pozitiv, ezért ő is szigorúan monoton nő, azonban nem injektiv, hiszen értelmezési tartománya nem összefüggő; viszont értelmezési tartományának bármely összefüggő részén injektív. Így értelmezzük az

$$\mathrm{arctg} := \left(\mathrm{tg}|_{]-\pi/2,\pi/2[}\right)^{-1} : \mathbb{R} \to]-\pi/2, \pi/2[$$

függvényt, amely differenciálható, és

$$\mathrm{arctg}' = \frac{1}{(1 + \mathrm{tg}^2) \circ \mathrm{arctg}} = \frac{1}{1 + \mathrm{id}_{\mathbb{R}}^2}.$$

Hasonlóan

$$\mathrm{arcctg} := \left(\mathrm{ctg}|_{]0,\pi[}\right)^{-1} : \mathbb{R} \to]0,\pi[,$$

$$\mathrm{arcctg}' = \frac{1}{-(1 + \mathrm{ctg}^2) \circ \mathrm{arcctg}} = -\frac{1}{1 + \mathrm{id}_{\mathbb{R}}^2}.$$

5.4. A valós hiperbolikus függvények inverzeinek értelmezése és a rájuk vonatkozó formulák hasonlók, mint a valós trigonometrikus függvényekre az előbb tárgyaltak.

(i) $\mathrm{sh}' = \mathrm{ch} > 0$, következésképpen a sh függvény szigorúan monoton nő, tehát injektiv; inverze

$$\mathrm{arsh} := \mathrm{sh}^{-1} : \mathbb{R} \to \mathbb{R}$$

is differenciálható, és

$$\operatorname{arsh}' = \frac{1}{\operatorname{ch} \circ \operatorname{arsh}} = \frac{1}{\sqrt{\operatorname{sh}^2 + 1} \circ \operatorname{arsh}} = \frac{1}{\sqrt{\operatorname{id}_{\mathbb{R}}^2 + 1}}.$$

Hasonlóan érvelhetünk megállapítva, hogy

$$\operatorname{arch} := \left(\operatorname{ch}|_{\mathbb{R}_0^+}\right) : [1, \infty[\to \mathbb{R}_0^+$$

differenciálható az értelmezési tartománya belsejében, és

$$\operatorname{arch}' = \frac{1}{\operatorname{sh} \circ \operatorname{arch}} = \frac{1}{\sqrt{\operatorname{ch}^2 - 1} \circ \operatorname{arch}} = \frac{1}{\sqrt{\operatorname{id}_{]1,\infty[}^2 - 1}}.$$

(ii) A th és cth függvények deriváltja pozitív illetve negatív, ezért injektívek; inverzeik

$$\operatorname{arth} := \operatorname{th}^{-1} :]-1, 1[\to \mathbb{R},$$
$$\operatorname{arcth} := \operatorname{cth}^{-1} : \mathbb{R} \setminus [-1, 1] \to \mathbb{R} \setminus \{0\}$$

differenciálhatók, és

$$\operatorname{arth}' = \frac{1}{(1 - \operatorname{th}^2) \circ \operatorname{arth}} = \frac{1}{1 - \operatorname{id}_{]-1,1[}^2},$$
$$\operatorname{arcth}' = \frac{1}{1 - \operatorname{id}_{\mathbb{R}\setminus[-1,1]}^2}.$$

Vigyázzunk: $\operatorname{arth}'(x) = \frac{1}{1-x^2}$, $\operatorname{arcth}'(x) = \frac{1}{1-x^2}$, vagyis ezen függvények hozzárendelési szabálya azonos alakú, de a függvények merőben mások, mert az értelmezési tartományuk különböző; méghozzá annyira, hogy diszjunkt halmazok.

5.5. Tudjuk, hogy az elemi függvények és inverzeik differenciálhatók, ismerjük a deriváltjukat. Ezért az ilyen függvényekből összeadással, szorzással, osztással, kompozícióval készített függvények differenciálóhatósága nem kérdéses, és a deriváltak kiszámítására pontos szabályok állnak a rendelkezésünkre. Vigyáznunk kell azonban az ilyenekből összeillesztett függvényekkel; az illesztési pontokban – ha vannak ilyenek és nem diszjunkt nyílt intervallumokon értelmezett függvényeket illesztünk össze – külön meg kell vizsgálnunk a differenciálhatóságot a definíció alapján, amint azt a következő két példában bemutatjuk.

(i) Az

$$\mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} x^2 \sin(1/x) & \text{ha } x \neq 0, \\ 0 & \text{ha } x = 0 \end{cases}$$

függvény nyilvánvalóan differenciálható a nullán kívűl. A nullában az

$$\frac{x^2\sin(1/x)-0}{x-0}=x\sin(1/x)$$

határértékét kell megvizsgálnunk; mivel a sin függvény korlátos, az identitás pedig a nullához tart a nullában, a szóban forgó határérték létezik és nulla. Ezért az adott függvény differenciálható.

(ii) Az

$$f:\mathbb{R}\to\mathbb{R},\quad x\mapsto\begin{cases} x^2 & \text{ha } x>0,\\ 0 & \text{ha } x\le 0\end{cases}$$

függvény nyilvánvalóan differenciálható a nullán kívűl. A nullában az $\dfrac{f(x)-0}{x-0}$ határértékét kell megvizsgálnunk. Könnyű látni, hogy ennek a törtnek mind a jobb- mind a baloldali határértéke nulla, így nulla maga a határértéke is: a függvény differenciálható.

5.7. Feladatok

1. Számítsuk ki a következő függvények deriváltját:
(i) $\sin^3\cos$, (ii) $\exp^a$ $(=x\mapsto\exp(ax))$, ahol $a\in\mathbb{R}$, (iii) $\dfrac{\cos}{\operatorname{arccos}}$,
(iv) $\mathrm{id}_{\mathbb{R}}^n\exp$, (v) $\mathrm{id}_{\mathbb{R}}\log$, (vi) $\sin\circ\cos$,
(vii) $x\mapsto x^x$ $(x\in\mathbb{R}^+)$, (viii) $x\mapsto(\sin x)^{\cos x}$ $(x\in]0,\pi[)$;
(e két utolsónál vegyük figyelembe, hogy $x^x=\exp(x\log(x))$ és $(\sin x)^{\cos x}=\exp(\cos x\log(\sin x))$).

2. Differenciálható-e az

$$\mathbb{R}\to\mathbb{R},\quad x\mapsto\begin{cases}\exp(-1/x) & \text{ha } x>0,\\ 0 & \text{ha } x\le 0\end{cases}$$

függvény?

3. Van-e olyan $a,b,c,d\in\mathbb{R}$, amelyekre az

$$\mathbb{R}\to\mathbb{R},\quad x\mapsto\begin{cases} x & \text{ha } x<0,\\ ax^3+bx^2+cx+d & \text{ha } 0\le x\le 1,\\ -x^2 & \text{ha } 1<x\end{cases}$$

függvény differenciálható?

4. Igazoljuk, hogy $(\log\circ|\cdot|)'=\frac{1}{\mathrm{id}_{\mathbb{R}}}$!

5. Mutassuk meg, hogy

$$\left(\sqrt[n]{\cdot}\right)'=\frac{1}{n\left(\sqrt[n]{\cdot}\right)^{n-1}},\qquad \left(\frac{1}{\sqrt[n]{\cdot}}\right)'=-\frac{1}{n\left(\sqrt[n]{\cdot}\right)^{n+1}}.$$

6. A L'Hospital-szabály

6.1. A differenciálszámítás sok alkalmazása közül kiemelünk most egyet, amely bizonyos határértékek meghatározására szolgál.

Állítás *Legyenek az $f, g :]a, b[\to \mathbb{R}$ függvények differenciálhatók,és az $]a, b[$ intervallum végpontjában létezzenek a*

$$\lim_{a+0} f = \lim_{a+0} g = 0 \qquad \left(\lim_{b-0} f = \lim_{b-0} g = 0\right)$$

féloldali határértékek. Ha $g'(x) \neq 0$ bármely $x \in]a, b[$ esetén, valamint létezik az $\dfrac{f'}{g'}$ függvénynek a-ban (b-ben) a féloldali határértéke, akkor létezik $\dfrac{f}{g}$-nek is féloldali határértéke a-ban (b-ben), továbbá

$$\lim_{a+0} \frac{f}{g} = \lim_{a+0} \frac{f'}{g'} \qquad \left(\lim_{b-0} \frac{f}{g} = \lim_{b-0} \frac{f'}{g'}\right).$$

Bizonyítás Szorítkozzunk a jobboldali határérték esetére, hiszen a másik eset bizonyítása nyilvánvalóan teljesen hasonló módon végezhető el.

Először is terjesszük ki az f és g függvényeket az a pontra $f(a) := g(a) := 0$ módon, ezáltal azok továbbra is folytonosak maradnak, hiszen jobboldali határértékük éppen nulla a-ban. Az így kapott $[a, b[\to \mathbb{R}$ függvényeket az egyszerűség kedvéért továbbra is az f és g betűkkel jelöljük.

Az $\dfrac{f'}{g'}$ határértékére vonatkozó feltevésünk alapján létezik $A \in \mathbb{R}$ úgy, hogy minden $\epsilon > 0$ számhoz találhatunk olyan $\delta_\epsilon > 0$ számot, hogy tetszőleges $x \in]a, b[$, $|x - a| < \delta_\epsilon$ pontra

$$\left|\frac{f'(x)}{g'(x)} - A\right| < \epsilon$$

teljesül.

Legyen $y \in]a, a + \delta_\epsilon[$. Az $[a, y]$ intervallumon alkalmazhatjuk a Cauchy-féle középértéktételt, miszerint létezik olyan $x \in]a, y[\subset]a, a + \delta_\epsilon[$, hogy

$$\frac{f(y) - f(a)}{g(y) - g(a)} = \frac{f(y)}{g(y)} = \frac{f'(x)}{g'(x)},$$

mivel $f(a) = g(a) = 0$.

Ebből azonban $\left|\dfrac{f(y)}{g(y)} - A\right| < \epsilon$ következik, hiszen $|x-a| < \delta_e$. Tehát $\dfrac{f}{g}$-nek létezik a-ban jobboldali határértéke, és az megegyezik az $\dfrac{f'}{g'}$ ottani határértékével.

Megjegyzések (i) Speciális esete az iménti állításnak, amikor f és g eleve az $[a,b]$ zárt intervallumon van értelmezve, folytonos, és a megfelelő végpontban a nulla értéket veszik föl; ilyenkor a bizonyítás második bekezdését egyszerűen el kell hagyni.

(ii) Az állítás mind jobboldali, mind baloldali határértékekre igaz, következésképpen, ha $c \in]a,b[$ és $f(c) = g(c) = 0$, $g'(x) \neq 0$ bármely $x \in]a,b[$ esetén, továbbá létezik az $\dfrac{f'}{g'}$ függvénynek határértéke c-ben, akkor az $\dfrac{f}{g}$ függvénynek is van határértéke ebben a pontban, és a két határérték megegyezik.

(iii) Az állítás megfordítása nem igaz: ha létezik $\lim\limits_{a+0} \dfrac{f}{g}$, abból nem következik, hogy létezik $\lim\limits_{a+0} \dfrac{f'}{g'}$. Ellenpéldát szolgáltat $g := \mathrm{id}_{\mathbb{R}}$ és az 5.5.(i)-ben adott függvény az f szerepére, az $a := 0$ pontban.

6.2. Állítás *Legyenek $f, g :]a,b[\to \mathbb{R}$ differenciálható függvények, és tegyük fel, hogy mindkét függvény jobbról (balról) végtelenhez tart a-ban (b-ben). Ha $g'(x) \neq 0$ bármely $x \in]a,b[$ esetén, valamint létezik az $\dfrac{f'}{g'}$ függvénynek a-ban (b-ben) a féloldali határértéke, akkor létezik $\dfrac{f}{g}$-nek is féloldali határértéke a-ban (b-ben), továbbá*

$$\lim_{a+0} \frac{f}{g} = \lim_{a+0} \frac{f'}{g'} \qquad \left(\lim_{b-0} \frac{f}{g} = \lim_{b-0} \frac{f'}{g'}\right).$$

Bizonyítás Természetesen most is elegendő a jobboldali határérték esetével foglalkoznunk, hiszen a másik eset szintén hasonló módon látható be.

Létezik $A \in \mathbb{R}$ úgy, hogy minden $\epsilon > 0$ számhoz létezik olyan $\delta_\epsilon > 0$ szám, hogy tetszőleges $x \in]a,b[$, $|x-a| < \delta_\epsilon$ pontra

$$\left|\frac{f'(x)}{g'(x)} - A\right| < \epsilon$$

teljesül.

Tekintve, hogy f és g differenciálható – tehát folytonos is – és mindkét függvény jobbról végtelenhez tart a-ban, van olyan $c \in]a, a+\delta_\epsilon[$, hogy az $]a,c]$ intervallumon már egyik függvény sem veszi föl a nulla értéket. Olyan $d \in]a,c[$ is létezik, hogy $|f(x)| > |f(c)|$, $|g(x)| > |g(c)|$ minden $x \in]a,d]$ esetén.

Rögzítsünk tetszőlegesen egy $x \in]a, d[$ pontot. Az $[x, c]$ intervallumon alkalmazhatjuk a Cauchy-féle középértéktételt, miszerint van olyan $z \in]x, c[$, hogy

$$\big(f(x) - f(c)\big)g'(z) = \big(g(x) - g(c)\big)f'(z).$$

Mivel g' sehol sem nulla $]a, b[$-n, valamint $x \in]a, d[$ miatt $g(x) \neq g(c)$, átírhatjuk a fenti egyenlőséget

$$\frac{f(x) - f(c)}{g(x) - g(c)} = \frac{f(x)}{g(x)} \frac{1 - \frac{f(c)}{f(x)}}{1 - \frac{g(c)}{g(x)}} = \frac{f'(z)}{g'(z)}$$

alakba. Bevezetve a

$$T(x) := \frac{1 - \frac{g(c)}{g(x)}}{1 - \frac{f(c)}{f(x)}}$$

jelölést az

$$\frac{f(x)}{g(x)} = \frac{f'(z)}{g'(z)} T(x) = \frac{f'(z)}{g'(z)} + \frac{f'(z)}{g'(z)}\big(T(x) - 1\big)$$

egyenlőséghez jutunk. Az f és a g függvény jobbról végtelenhez tart a-ban, ezért $\lim\limits_{a+0} T = 1$. Figyelembe véve, hogy az $]a, c[$ interallumon $\dfrac{f'}{g'}$ korlátos, és $z \in]a, d[\subset]a, c[$, létezik olyan $d_1 \in]a, d[$, hogy minden $x \in]a, d_1[$ esetén

$$\left| \frac{f'(z)}{g'(z)}\big(T(x) - 1\big) \right| < \frac{\epsilon}{2}.$$

A fentiek alapján minden $x \in]a, d_1[$ pontra teljesül az

$$\left| \frac{f(x)}{g(x)} - A \right| \leq \left| \frac{f'(z)}{g'(z)} - A \right| + \left| \frac{f'(z)}{g'(z)}\big(T(x) - 1\big) \right| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

egyenlőtlenség. Ez pontosan azt jelenti, hogy az $\dfrac{f}{g}$ függvény jobbról A-hoz konvergál az a pontban, és ezt akartuk bizonyítani.

6.3. Az előző két állításnak megfelelően $\dfrac{0}{0}$ típusú, illetve $\dfrac{\infty}{\infty}$ típusú határértékekről szoktunk beszélni. Előfordulhat azonban 0∞ típusú határérték is: jobbról tartva az a ponthoz f nullához, g viszont végtelenhez tart. Ilyen esetekben az fg függvény a-beli jobboldali határértékének kiszámítására nyújt lehetőséget a L'Hospital-szabály, ha az eredeti szorzatot $\dfrac{0}{0}$ típusú $\dfrac{f}{\frac{1}{g}}$ vagy (ha f nem nulla az a pont egy jobboldali környezetében) $\dfrac{\infty}{\infty}$ típusú $\dfrac{g}{\frac{1}{f}}$ formába írjuk át.

Felhívjuk a figyelmet arra, hogy f-től és g-től függően hol az egyik, hol a másik átírás vezethet célhoz.

6.4. Feladatok

1. Bizonyítsuk be, hogy

(i) $\lim\limits_{0} \dfrac{\sin}{\mathrm{id}_{\mathbb{R}}} = 1$, (ii) $\lim\limits_{0} \dfrac{\exp -1}{\mathrm{id}_{\mathbb{R}}} = 1$, (iii) $\lim\limits_{+0} \mathrm{id}_{\mathbb{R}} \log = 0$.

2. Határozzuk meg a következő határértékeket:

(i) $\lim\limits_{x\to 0} \dfrac{\sqrt{1+3x^2}-1}{x^2}$, (ii) $\lim\limits_{x\to +0} x^x$,

(iii) $\lim\limits_{0} \dfrac{\mathrm{tg} - \sin}{\mathrm{id}_{\mathbb{R}} - \sin}$, (iv) $\lim\limits_{1} \dfrac{\mathrm{id}_{\mathbb{R}} - 1}{\log}$.

3. Bizonyítsuk be a fejezetben szereplő két állítást a következő módosítással: az $]a, b[$ intervallum helyett vegyük a $]-\infty, b[$ illetve az $]a, \infty[$ intervallumot!

7. Határozatlan integrál

Ebben a fejezetben az I és a J betűkkel valós számok nyílt intervallumait jelöljük.

7.1. Definíció *Legyen $f : I \to \mathbb{R}$ függvény. Az $\int f := \{F : I \to \mathbb{R} \mid F' = f\}$ halmazt az f függvény* ***határozatlan integráljának****, elemeit pedig az f* ***primitív függvényeinek*** *nevezzük. Az $\int\limits_a f$ jelölést használjuk arra a primitív függvényre, amelyik az $a \in I$ pontban a nulla értéket veszi föl, vagyis $\int\limits_a f$ az $\int f$ eleme és $\left(\int\limits_a f\right)(a) = 0$.*

Egy függvény határozatlan integrálja lehet üres. Például $\int \chi_{\mathbb{Q}} = 0$. Később, az integrálszámítás során látni fogjuk, hogy folytonos függvénynek mindig létezik primitív függvénye.

Állítás *Ha $F : I \to \mathbb{R}$ az $f : I \to \mathbb{R}$ függvény egy primitív függvénye, akkor $\int f = \{F + C \mid C \in \mathbb{R}\}$.*

Bizonyítás Nyilvánvaló, ha F az f primitív függvénye, akkor tetszőleges C valós szám esetén $F + C$ is az. Ezek után már csak azt kell belátnunk, hogy az f összes primitív függvénye előáll ilyen alakban.

Legyen $G : I \to \mathbb{R}$ szintén az f egy primitív függvénye, tehát $F' = G' = f$. Ekkor $(G - F)' = G' - F' = 0$. Mivel $G - F$ intervallumon értelmezett függvény, a 2.3. állítás miatt $G - F$ konstans, amiből $G - F =: C$, $G = F + C$ következik. ■

Az iménti állítás következménye, hogy amennyiben egy $f : I \to \mathbb{R}$ függvénynek létezik legalább egy primitív függvénye, vagyis $\int f \neq \emptyset$, akkor minden $a \in I$ esetén az $\int\limits_a f$ primitív függvény is biztosan létezik és egyértelmű. (Ha $F' = f$, akkor $\int\limits_a f = F - F(a)$.)

7.2. Állítás *Ha az $f, g : I \to \mathbb{R}$ függvényeknek létezik primitív függvénye, akkor $f + g$-nek és minden $\lambda \in \mathbb{R}$ esetén λf-nek is van primitív függvénye, és minden $a \in I$ esetén*

(i) $\int\limits_a (f + g) = \int\limits_a f + \int\limits_a g$,

(ii) $\int\limits_a (\lambda f) = \lambda \int\limits_a f$ *tetszőleges $\lambda \in \mathbb{R}$ esetén,*

vagyis az $f \mapsto \int\limits_a f$ hozzárendelés lineáris.

Bizonyítás Nyilvánvaló, hogy

$$\left(\int\limits_a f + \int\limits_a g\right)' = f + g \quad \text{és} \quad \left(\int\limits_a f + \int\limits_a g\right)(a) = 0,$$

valamint

$$\left(\lambda \int\limits_a f\right)' = \lambda f \quad \text{és} \quad \left(\lambda \int\limits_a f\right)(a) = 0.$$

7.3. Állítás *(**Parciális integrálás**) Legyenek $f, g : I \to \mathbb{R}$ függvények és $a \in I$. Ha f differenciálható, a g és az $\left(f' \int\limits_a g\right)$ függvényeknek pedig létezik primitív függvénye, akkor az fg függvénynek is van primitív függvénye, valamint*

$$\int\limits_a (fg) = f \int\limits_a g - \int\limits_a \left(f' \int\limits_a g\right).$$

Bizonyítás Nyilvánvaló, hogy a fenti egyenlőség jobb odalán álló függvény az a pontban a nulla értéket veszi fel, továbbá differenciálható, és

$$\left(f \int\limits_a g - \int\limits_a \left(f' \int\limits_a g\right)\right)' = f' \int\limits_a g + fg - f' \int\limits_a g = fg,$$

és ezzel befejeztük a bizonyítást. ■

Ha bevezetjük a $G := \int_a g$ és $F := f$ jelöléseket, akkor a parciális integrálás formulájának egy szokásosabb alakját kapjuk:

$$\int_a (FG') = FG - \int_a (F'G).$$

7.4. Állítás (***Helyettesítéses integrálás***) *Tegyük fel, hogy az $f : I \to \mathbb{R}$ függvénynek van primitív függvénye. Ha $S : J \to I$ differenciálható, akkor az $(f \circ S)S' : J \to \mathbb{R}$ függvénynek is van primitív függvénye, valamint minden $b \in J$ esetén*

$$\int_b (f \circ S)S' = \left(\int_{S(b)} f \right) \circ S.$$

Bizonyítás Csak azt kell tennünk, hogy ellenőrizzük: a fenti képlet jobb oldalán szereplő függvény a b pontban nulla értéket vesz fel – ez nyilvánvaló –, differenciálható – ez is nyilvánvaló –, és a deriváltja

$$\left(\left(\int_{S(b)} f \right) \circ S \right)' = \left(\left(\int_{S(b)} f \right)' \circ S \right) S' = (f \circ S)S'.$$

7.5. Hogy a gyakorlatban is használni tudjuk a határozatlan integrálokról tanultakat, nézzük meg a valós elemi függvények primitív függvényeit, amelyeket alapintegráloknak is szoktak nevezni. Ehhez csak fel kell idéznünk az elemi függvények deriváltjairól tanultakat és kis módosítással felírni az ottani formulákat "jobbról balra olvasva". Előbb azonban említést kell tennünk az irodalomban elfogadott jelölésekről. A 7.1. állításban láttuk, hogy ha $F : I \to \mathbb{R}$ az $f : I \to \mathbb{R}$ egy primitív függvénye, akkor $\int f = \{F + C \mid C \in \mathbb{R}\}$. Ehelyett sokszor csak azt írjuk, hogy $\int f = F + C$, illetve, leginkább konkrét gyakorlati számolások során, használatos az

$$\int f(x)dx = F(x) + C$$

írásmód.

Lássuk akkor az alapintegrálokat, amelyeket itt csak fölsorolunk, bízva abban, hogy ezek a formulák nem okoznak különösebb fejtörést az Olvasónak:

$$\int e^x dx = e^x + C;$$

$$\int \sin(x)dx = -\cos(x) + C; \qquad \int \cos(x)dx = \sin(x) + C;$$

$$\int \mathrm{sh}(x)dx = \mathrm{ch}(x) + C; \qquad \int \mathrm{ch}(x)dx = \mathrm{sh}(x) + C;$$

$$\int \frac{dx}{\cos^2(x)} = \mathrm{tg}(x) + C; \qquad \int \frac{dx}{\sin^2(x)} = -\mathrm{ctg}(x) + C;$$

$$\int \frac{dx}{\mathrm{ch}^2(x)} = \mathrm{th}(x) + C; \qquad \int \frac{dx}{\mathrm{sh}^2(x)} = -\mathrm{cth}(x) + C;$$

$$\int \frac{dx}{x} = \log(x) + C \quad (x \in \mathbb{R}^+);$$

$$\int x^\alpha dx = \frac{1}{\alpha+1}x^{\alpha+1} + C \quad (1 \neq \alpha \in \mathbb{R}^+);$$

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin(x) + C_1 = -\arccos(x) + C_2;$$

$$\int \frac{dx}{x^2+1} = \mathrm{arctg}(x) + C_1 = -\mathrm{arcctg}(x) + C_2;$$

$$\int \frac{dx}{x^2-1} = \frac{1}{2}\log\left|\frac{x-1}{x+1}\right| + C;$$

$$\int \frac{dx}{\sqrt{x^2+1}} = \mathrm{arsh}(x) + C;$$

$$\int \frac{dx}{\sqrt{x^2-1}} = \mathrm{arch}(x) + C;$$

$$\int \frac{dx}{\sqrt{x^2 \pm 1}} = \log\left|x + \sqrt{x^2 \pm 1}\right| + C.$$

A fenti formulák ellenőrzésének egy egyszerű módja, ha igazoljuk, hogy a jobb oldalak deriváltjai pontosan az integrandusok. Ajánljuk az Olvasónak, hogy számoljon utána.

Végül, a gyakorlati integrálási feladatok megoldását segítendő, írjuk fel a parciális és a helyettesítéses integrálás formuláit az alapintegráloknál is használt szokásos (integrálási változós) jelöléssel.

Parciális integrálás:

$$\int u(x)v'(x)dx = u(x)v(x) - \int u'(x)v(x)dx,$$

ahol $u, v : \mathbb{R} \rightarrowtail \mathbb{R}$ megfelelő tulajdonságú függvények.

Helyettesítéses integrálás:

$$\int f(x)dx = \int f(S(t))S'(t)dt, \quad x = S(t);$$

illetve ugyanez "visszafelé", amit új változó bevezetésének hívnak:

$$\int f(S(x))S'(x)dx = \int f(y)dy, \quad S(x) = y;$$

itt $f, S : \mathbb{R} \rightarrowtail \mathbb{R}$ megfelelő tulajdonságú függvények.

7.6. Feladatok

1. Vezessük vissza alapintegrálokra a következő integrálokat:
(i) $\int(1-x)(1-2x)(1-3x)dx$; (ii) $\int\sqrt{1-\sin(2x)}dx$;
(iii) $\displaystyle\int \frac{x+1}{\sqrt{x}}dx$; (iv) $\displaystyle\int \frac{x^2+3}{x^2-1}dx$; (v) $\int \mathrm{tg}^2(x)dx$; (vi) $\int(2^x+3^x)^2dx$.

2. Megfelelő új változó bevezetésével végezzük el az alábbi integrálásokat:
(i) $\displaystyle\int \frac{x^5}{\sqrt{1-x^2}}dx$; (ii) $\displaystyle\int \frac{\sin(x)\cos^3(x)}{1+\cos^2(x)}dx$;
(iii) $\displaystyle\int \frac{\log(x)}{x\sqrt{1+\log(x)}}dx$; (iv) $\displaystyle\int \frac{dx}{\sqrt{1+e^x}}$.

3. Adott $a \in \mathbb{R}$ esetén az $x = a\sin(t)$, $x = a\mathrm{tg}(t)$, $x = a\sin^2(t)$, stb. trigonometrikus helyettesítések alkalmazásával határozzuk meg az alábbi integrálokat:
(i) $\displaystyle\int \frac{dx}{(1-x^2)^{3/2}}$; (ii) $\int\sqrt{a^2-x^2}dx$; (iii) $\displaystyle\int \sqrt{\frac{a+x}{a-x}}dx$.

4. Számoljuk ki a következő integrálokat a parciális integrálás módszerével:
(i) $\int \log(x)dx$; (ii) $\int x^2e^{-2x}dx$; (iii) $\int x^2\sin(2x)dx$;
(iv) $\int x\mathrm{sh}(x)dx$; (v) $\int \log\left(x+\sqrt{1+x^2}\right)dx$; (vi) $\int \sin(x)\log(\mathrm{tg}(x))dx$;
(vii) $\int x^2\arccos(x)dx$.

II. TÖBBSZÖR DIFFERENCIÁLHATÓ FÜGGVÉNYEK

8. Magasabb rendű deriváltak

8.1. Az előző részben tárgyaltuk $\mathbb{K} \rightarrowtail \mathbb{K}^N$ függvények differenciálhatóságát. Láttuk, hogy az ilyen függvények deriváltjai szintén $\mathbb{K} \rightarrowtail \mathbb{K}^N$ függvények. Következésképpen értelmes (és természetes) az a gondolat, hogy a derivált függvények differenciálhatóságát is tárgyaljuk, továbbá értelmezzük $\mathbb{K} \rightarrowtail \mathbb{K}^N$ függvények kétszeres és többszörös differenciálhatóságát.

1. Definíció *Legyen $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ függvény. Értelmezzük rekurzióval az $f^{(n)}$ $n \in \mathbb{N}_0$ függvénysorozatot a következő módon:*

(i) $f^{(0)} := f$,

(ii) $f^{(n)} := \left(f^{(n-1)}\right)'$ $\quad (n \in \mathbb{N})$.

*Az $f^{(n)}$ függvényt az f n-**edik derivált függvényének** (n-**edik deriváltjának**) hívjuk.*

Az iménti függvénysorozat nyilván jól értelmezett, vagyis bármely $\mathbb{K} \rightarrowtail \mathbb{K}^N$ függvénynek létezik akárhányadik deriváltja, viszont az persze előfordulhat, hogy egy bizonyos természetes számra, és akkor már minden annál nagyobb n-re is $f^{(n)}$ az üres függvény (az értelmezési tartománya az üres halmaz). A fenti rekurzióval tehát egy függvény magasabb rendű deriváltjait értelmeztük, de magasabb rendű differenciálhatóságát nem; az n-szer diffrenciálhatóság lényegében azt jelenti, hogy az n-ik derivált nem az üres függvény. A pontos megfogalmazást a következő meghatározásban találjuk.

2. Definíció *Legyen $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ függvény, $n \in \mathbb{N}$. Azt mondjuk, hogy az f függvény*

*(i) n-**szer differenciálható az** $a \in \mathbb{K}$ **pontban**, ha $a \in \mathrm{Dom}(f^{(n)}$;*

*(ii) n-**szer differenciálható a** $H \subset \mathbb{K}^N$ **halmazon**, ha $H \subset \mathrm{Dom}(f^{(n)})$;*

*(iii) n-**szer differenciálható**, ha az értelmezési tartománya nyílt és $\mathrm{Dom}(f^{(n)}) = \mathrm{Dom}(f)$;*

*(iv) **végtelenszer differenciálható egy pontban (egy halmazon)**, ha minden $n \in \mathbb{N}$ esetén n-szer differenciálható abban a pontban (azon a halmazon);*

*(v) **végtelenszer differenciálható**, ha az értelmezési tartománya nyílt, és minden $n \in \mathbb{N}$ esetén n-szer differenciálható.*

Az iménti definícióval kapcsolatban rögtön meg kell jegyeznünk, hogy az "egyszer differenciálhatóság" pontosan az eddig megismert differenciálhatósággal egyezik meg. Említettük már, hogy az első definíció értelmében bármely $\mathbb{K} \rightarrowtail \mathbb{K}^N$ függvénynek létezik akárhányadik deriváltja; világos azonban, ez nem jelenti azt, hogy az adott függvény akár egyszer is differenciálható, hiszen ez a létező derivált a mi értelmezésünk szerint lehet az üres függvény is.

A magasabb rendű deriváltak szintén $\mathbb{K} \rightarrowtail \mathbb{K}^N$ függvények, és a definíció szerint általában az n-edik deriváltat $f^{(n)}$ jelöli; azonban az első egy-két deriváltra többnyire egy-két "vesszővel" utalunk:

$$(f^{(0)} = f) \qquad f^{(1)} = f', \quad f^{(2)} =: f'', \quad f^{(3)} = f'''.$$

Olykor más jelölést is alkalmazunk, amely különösen konkrét formulával megadott függvények esetén előnyös:

$$f'(x) = \frac{\mathrm{d}f(x)}{\mathrm{d}x}, \quad f''(x) = \frac{\mathrm{d}^2 f(x)}{\mathrm{d}x^2}, \quad f^{(n)}(x) = \frac{\mathrm{d}^n f(x)}{\mathrm{d}x^n}.$$

8.2. A magasabbrendű deriváltaknak, mint $\mathbb{K} \rightarrowtail \mathbb{K}^N$ függvényeknek a folytonossága sok alkalmazásban játszik fontos szerepet.

Definíció *Azt mondjuk, hogy az $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ függvény*

(i) folytonosan differenciálható, ha differenciálható és a deriváltja folytonos;

(ii) $n \in \mathbb{N}$ esetén n-szer folytonosan differenciálható, ha n-szer differenciálható és $f^{(n)}$ folytonos.

Nyilvánvalóan egy n-szer differenciálható függvény $(n-1)$-szer folytonosan differenciálható, továbbá a végtelenszer differenciálható függvények akárhányszor folytonosan differenciálhatók is.

Ha $X \subset \mathbb{K}^M$, $Y \subset \mathbb{K}^N$, akkor $C(X,Y)$ az $X \to Y$ folytonos fügvények szokásos jelölése.

Most, ha $X \subset \mathbb{K}$, $n \in \mathbb{N}$ esetére bevezetjük a

$$C^n(X,Y) := \{f : X \to Y \mid f \ n\text{-szer folytonosan differenciálható}\}$$

jelölést, és

$$C^\infty(X,Y) := \{f : X \to Y \mid f \text{ végtelenszer differenciálható}\}.$$

Tehát például $C^1(]a,b[,\mathbb{R})$ az $]a,b[\to \mathbb{R}$ folytonosan differenciálható függvények összességét jelöli. Ugyanúgy, mint a folytonos függvények esetében, ha az érkezési halmaz egyértelmű (pl. a szövegkörnyezetből), akkor azt külön nem jelöljük. Az előző példánál maradva, használatos a $C^1(]a,b[)$ jelölés $C^1(]a,b[,\mathbb{R})$ helyett, de a $C^1(]a,b[,\mathbb{C})$ helyett is.

8.3. (i) Az elemi függvények végtelenszer differenciálhatók.

(ii) Végtelenszer differenciálható az

$$\eta : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} e^{\frac{1}{x}} & \text{ha } x > 0, \\ 0 & \text{ha } x \le 0 \end{cases}$$

függvény is. Ugyanis világos, hogy

– végtelenszer differenciálható $\mathbb{R}^-$-on, ott minden deriváltja 0,

– végtelenszer differenciálható $\mathbb{R}^+$-on is, és teljes indukcióval meggyőződhetünk arról, hogy minden n esetén létezik egy p_n polinom úgy, hogy

$$\eta^{(n)}(x) = p_n\left(\frac{1}{x}\right) e^{-\frac{1}{x}} \qquad (x > 0).$$

Mivel

$$\lim_{x\to -0} \frac{\eta(x)}{x} = \lim_{x\to -0} \frac{0}{x} = 0,$$

$$\lim_{x\to +0} \frac{\eta(x)}{x} = \lim_{x\to +0} \frac{1}{x} e^{-\frac{1}{x}} = \lim_{z\to\infty} z e^{-z} = 0,$$

azt állíthatjuk, hogy

$$\lim_{x\to 0} \frac{\eta(x) - 0}{x - 0} = \lim_{x\to 0} \frac{\eta(x)}{x} = 0,$$

vagyis a függvény differenciálható a nullában is, és ott a deriváltja nulla. Ezután ismét teljes indukcióval beláthatjuk, hogy a függvény akárhányszor differenciálható a nullában, és ott minden rendű deriváltja nulla.

(iii) Az 5.5.(i) példában szereplő függvény differenciálható, a deriváltja

$$x \mapsto \begin{cases} 2x\sin(1/x) - \cos(1/x) & \text{ha } x \neq 0, \\ 0 & \text{ha } x = 0 \end{cases}$$

nem folytonos a nullában; tehát a függvény nem differenciálható folytonosan (és így kétszer sem differenciálható).

(iv) Az 5.5.(ii) példában szereplő függvény differenciálható, a deriváltja

$$x \mapsto \begin{cases} 2x & \text{ha } x > 0, \\ 0 & \text{ha } x \leq 0 \end{cases}$$

folytonos, de nem differenciálható a nullában; tehát a függvény folytonosan differenciálható, de kétszer nem differenciálható.

(v) Későbbi tanulmányaink során megmutatjuk (lásd Analízis VI.), hogy egy $\mathbb{C} \rightarrowtail \mathbb{C}^N$ függvény, ha differenciálható, akkor már végtelenszer is differenciálható. Ez nem áll valós függvényekre, amint azt az előző példák mutatják.

8.4. Megemlítjük a folytonosan differenciálható függvények egy fontos jó tulajdonságát: ha $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ folytonosan differenciálható és $f'(a) \neq 0$, akkor van olyan környezete a-nak, hogy abban levő minden x-re $f'(x) \neq 0$ (Analízis III.A.27.7.). Speciálisan, ha $f : \mathbb{R} \rightarrowtail \mathbb{R}$ folytonosan differenciálható és $f'(a) > 0$, akkor f' pozitív az a-nak egy környezetén. Ez azt jelenti, hogy f szigorúan monoton nő ebben a környezetben. Ajánljuk az Olvasónak, lapozza fel a 2.4. állítást és az utána mondottakat ezzel kapcsolatban.

8.5. Már láttuk, hogy egy differenciálható függvénynek csak akkor lehet szélsőértéke egy pontban, ha ott a deriváltja nulla. Ez azonban csak szükséges feltétel, és arra vonatkozóan sem ad információt, hogy az illető pontban maximuma vagy minimuma van (lehet) az adott függvénynek. A következő állításban a szélsőérték egy elégséges feltételét fogalmazzuk meg és a fenti kérdésre is választ adunk.

1. Állítás *Legyen $f : \mathbb{R} \rightarrowtail \mathbb{R}$ differenciálható az a pontot tartalmazó $I \subset \mathrm{Dom}(f)$ nyílt intervallumon. Ha $f'(a) = 0$ és*

$$f'(x) < 0 < f'(y) \quad \textit{illetve} \quad f'(x) > 0 > f'(y)$$

$$\textit{minden } x, y \in I, \quad x < a < y$$

esetén, vagyis f' a fenti módon előjelet vált az a pontban, akkor az f függvénynek szigorú lokális minimuma illetve maximuma van az a pontban.

Bizonyítás Ha $I =]b, c[$, akkor a 2.3. állítás miatt f a $]b, a[$ intervallumon szigorúan monoton csökken (növekszik), az $]a, c[$ intervallumon pedig szigorúan monoton nő (csökken). Az f függvény folytonos (mert differenciálható) I-n, ezért

$$\lim_{x \to a-0} f(x) = f(a) = \lim_{y \to a+0} f(y),$$

amiből a bizonyítandó állítás már nyilvánvaló. ∎

Gyakran előfordul az az eset, hogy a vizsgált függvénynek létezik az adott pontban második deriváltja is. Ilyenkor az iménti állításnak egy módosított változatát is ki tudjuk mondani.

2. Állítás *Legyen $f : \mathbb{R} \rightarrowtail \mathbb{R}$ az a pontban kétszer differenciálható. Ha $f'(a) = 0$ és $f''(a) > 0$ $(f''(a) < 0)$, akkor az f függvénynek az a-ban szigorú lokális minimuma (maximuma) van.*

Bizonyítás Csak a 2.4. állítást kell alkalmaznunk f'-re, hogy lássuk, az f függvény eleget tesz az előző állítás feltételeinek.

8.6. Megeshet, hogy $f'(a) = 0$ mellett $f''(a) = 0$ is fennáll. Többször differenciálható függvényeknél ilyenkor is tudunk valamit mondani a szélsőértékekre.

Állítás *Legyen $n \in \mathbb{N} \setminus \{1\}$ és $f : \mathbb{R} \rightarrowtail \mathbb{R}$ n-szer differenciálható az a pontban, továbbá*

$$f'(a) = f''(a) = \cdots = f^{(n-1)}(a) = 0, \qquad f^{(n)}(a) \neq 0.$$

Az f függvénynek akkor és csak akkor van szigorú lokális minimuma (maximuma) a-ban, ha n páros és $f^{(n)}(a) > 0$ $(f^{(n)}(a) < 0)$.

Bizonyítás Tekintsük az $f^{(n)}(a) > 0$ esetet. Ekkor 2.4. alapján $f^{(n-1)}$ az a pontban negatívból pozitív értékűbe megy át, hiszen $f^{(n-1)}(a) = 0$. Következésképpen $f^{(n-2)}$-nek a-ban szigorú lokális minimuma van. Mivel $f^{(n-2)}(a) = 0$, van olyan $G(a)$ környezete a-nak, hogy $G(a) \setminus \{a\}$-n $f^{(n-2)}(a) > 0$. Ez azt jelenti, hogy $G(a) \setminus \{a\}$-n $f^{(n-3)}$ szigorúan monoton növekszik, következésképpen $f^{(n-3)}$ is negatív értékről pozitívra vált az a pontban, hiszen $f^{(n-3)} = 0$.

Értelemszerűen folytatva a fenti gondolatmenetet azt kapjuk, hogy az $f^{(n-1)}$, $f^{(n-3)}$, $f^{(n-5)}$, ... függvények az a pontban negatív értékről pozitívra váltanak, az $f^{(n)}$, $f^{(n-2)}$, $f^{(n-4)}$, ... függvényeknek pedig szigorú lokális minimumuk van a-ban.

Innen már nyilvánvaló, hogy páros n esetén f-nek szigorú lokális minimuma van a-ban, páratlan n esetén pedig szigorúan monoton növekszik ebben a pontban, tehát nem lehet szélsőértéke.

Az $f^{(n)} < 0$ eset teljesen hasonló módon tárgyalható, az a különbség, hogy ott lokális maximumot illetve szigorú monoton csökkenést fogunk kapni. ■

Érdekességképpen megemlítjük, hogy f lehet az a pontban végtelenszer differenciálható, és előfordulhat, hogy $f^{(n)}(a) = 0$ minden n-re, vagyis az összes deriváltja "eltűnik" ebben a pontban. A szélsőérték szempontjából ekkor bármi lehetséges, a differenciálszámítás ilyen esetekben alkalmatlan a szélsőérték-vizsgálatra.

8.7. Feladatok

1. Vizsgáljuk meg, hányszor (folytonosan) differenciálható a $(\mathrm{sign})\mathrm{id}_{\mathbb{R}}^2$ függvény!

2. Keressük meg a következő függvények lokális szélsőértékeit:

(i) $(\mathrm{id}_{\mathbb{R}}-1)^2(\mathrm{id}_{\mathbb{R}}+2)^2$; (ii) $\dfrac{1}{1+\mathrm{id}_{\mathbb{R}}^2}$; (iii) $\sin+\cos$,

(iv) $\mathbb{R}\to\mathbb{R}, \quad x\mapsto \mathrm{arctg}(2x)$; (v) $\mathbb{R}^+\to\mathbb{R}^+, \quad x\mapsto x^x$.

3. Bizonyítsuk be: ha $f:\mathbb{R}\rightarrowtail\mathbb{R}$ az a pontban kétszer differenciálható, $f'(a)=0$, $f''(a)=0$, és van az a-nak olyan $G(a)$ környezete, hogy $f''(x)>0$ minden $x\in G(a)\setminus\{a\}$ esetén, akkor az f függvénynek az a-ban szigorú lokális minimuma van.

9. Taylor-formula

9.1. Állítás (***Taylor-formula***) *Legyen az $f:[a,b]\to\mathbb{R}$ folytonos függvény $(n+1)$-szer differenciálható az $]a,b[$ intervallumon. Ha az $f^{(k)}$ $(k=1,\dots,n)$ derivált függvényeknek létezik az $f^k(a)$-val jelölt jobboldali határértéke az a pontban, akkor van olyan $c\in]a,b[$, hogy*

$$f(b)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(b-a)^k+\frac{f^{(n+1)}(c)}{(n+1)!}(b-a)^{n+1}.$$

Bizonyítás Definiáljuk a következő $[a,b]\to\mathbb{R}$ folytonos függvényeket:

$$F(x):=f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k,$$
$$G(x):=(x-a)^{n+1}.$$

Az $]a,b[$ intervallumon az F és a G függvény egyaránt legalább $(n+1)$-szer differenciálható. A Cauchy-féle középértéktétel szerint létezik olyan $c_1\in]a,b[$, hogy

$$\frac{F(b)-F(a)}{G(b)-G(a)}=\frac{F'(c_1)}{G'(c_1)}=\frac{F(b)}{G(b)},$$

mivel $F(a)=G(a)=0$.

Az F és a G függvény deriváltját könnyen megkaphatjuk:

$$F'(x)=f'(x)-\sum_{k=1}^{n}\frac{f^{(k)}(a)}{(k-1)!}(b-a)^{k-1},$$
$$G'(x)=(n+1)(x-a)^n,$$

és látható, hogy $F'(a) = G'(a) = 0$.

Az $[a, c_1]$ intervallumon ismét alkalmazhatjuk a Cauchy-féle középértéktételt, miszerint létezik olyan $c_2 \in]a, c_1[$, hogy

$$\frac{F'(c_1) - F'(a)}{G'(c_1) - G'(a)} = \frac{F''(c_2)}{G''(c_2)} = \frac{F'(c_1)}{G'(c_1)}.$$

Nyilvánvaló, hogy $F^{(k)}(a) = G^{(k)}(a) = 0$ minden $k = 0, \dots, n$ esetén, így még $(n-1)$-szer alkalmazva a Cauchy-féle középértéktételt kapjuk, hogy létezik olyan $c_{n+1} \in]a, b[$, hogy

$$\frac{F(b)}{G(b)} = \frac{F'(c_1)}{G'(c_1)} = \frac{F''(c_2)}{G''(c_2)} = \cdots = \frac{F^{(n+1)}(c_{n+1})}{G_{(n+1)}(c_{n+1})}.$$

Vezessük be a $c := c_{n+1}$ jelölést és írjuk be a kapott egyenlőségbe az F és a G függvény konkrét alakját:

$$\frac{f(b) - \sum\limits_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k}{(b-a)^{n+1}} = \frac{f^{(n+1)}(c)}{(n+1)!}.$$

Ebből egyszerű átrendezéssel megkapjuk azt az egyenlőséget, amelyet bizonyítani akartunk. ■

Teljesen hasonlóan láthatjuk be, hogy ha az $f : [a, b] \to \mathbb{R}$folytonos függvény $(n+1)$-szer differenciálható az $]a, b[$ intervallumon és az $f^{(k)}$ $(k = 1, ..., n)$ derivált függvényeknek létezik az $f^{(k)}(b)$-vel jelölt baloldali határértéke a b pontban, akkor van olyan $c \in]a, b[$, hogy

$$f(a) = \sum_{k=0}^{n} \frac{f^{(k)}(b)}{k!}(a-b)^k + \frac{f^{(n+1)}(c)}{(n+1)!}(a-b)^{n+1}.$$

Azonban a bizonyítást mégsem kell újból végigvinni, mert ezt a formulát egyszerűen úgy is igazolhatjuk, hogy az iménti állítást alkalmazzuk az $f \circ \phi$ függvényre, ahol $\phi : [a, b] \to [a, b]$, $t \mapsto a + b - t$.

9.2. Gyakran, főleg alkalmazásokban a Taylor-formulának egy másik alakjával találkozunk.

Legyen az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvény $(n+1)$-szer differenciálható az $x \in \mathrm{Dom}(f)$ pont egy G környezetében. Ha h olyan valós szám, hogy $x + h$ benne van ebben a környezetben, akkor létezik $\theta_{x,h} \in]0, 1[$ úgy, hogy

$$f(x+h) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}h^k + \frac{f^{(n+1)}(x + \theta_{x,h}h)}{(n+1)!}h^{n+1}.$$

Hogy ezt az alakot megkapjuk, helyettesítsük az előző formulában a-t x-szel, $(b-a)$-t h-val, és $\theta_{x,h} := \dfrac{c-x}{h}$.

9.3. Ha $f:\mathbb{R}\rightarrowtail\mathbb{R}$ $(n+1)$-szer differenciálható az értelmezési tartománya a pontjának egy $G(a)$ környezetében, akkor a 9.1. állításban megjelenő b szerepére vehetjük a $G(a)$ környezet bármely x pontját; így azt mondhatjuk, hogy minden $x\in G(a)$ esetén létezik c_x pont a és x között úgy, hogy

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k+\frac{f^{(n+1)}(c_x)}{(n+1)!}(x-a)^{n+1}.$$

A jobb oldal első tagját, mint x függvényét, az f függvény a-hoz tartozó, **n-ed fokú Taylor-polinomjának** nevezzük.

Ha $f:\mathbb{R}\rightarrowtail\mathbb{R}$ végtelenszer differenciálható a-ban, akkor képezhetjük a

$$\sum_{k=0}^{\infty}\frac{f^{(k)}(a)}{k!}(\mathrm{id}_{\mathbb{R}}-a)^k$$

hatványsort, amelyet az f függvény a-hoz tartozó (a körüli) **Taylor-sorának** hívunk. Ennek n-edik részletösszege az f megfelelő Taylor-polinomja.

Kérdés, hogy

(i) konvergens-e a Taylor-sor az a pont egy környezetében (nagyobb-e a konvergencia-sugara nullánál),

(ii) ha konvergens, mi az összege (egyenlő-e a szóban forgó függvénnyel)?

Azt remélnénk, hogy konvergens Taylor-sor összege megegyezik a függvénnyel. Ez azonban nem feltétlenül áll fenn, amire példát a 8.3.(ii)-beli függvény szolgáltat: a nullában a függvény és minden deriváltja nulla, tehát a nulla körüli Taylor sora az azonosan nulla függvényt állítja elő.

Állítás *Legyen az $f:\mathbb{R}\rightarrowtail\mathbb{R}$ függvény végtelenszer differenciálható az a pont $G\subset\mathrm{Dom}(f)$ környezetében. Ha létezik a derivált függvényeknek közös korlátja G-n, vagyis van olyan $K>0$ szám, hogy $|f^{(n)}(x)|\le K$ minden $n\in\mathbb{N}$ és $x\in G$ esetén, akkor az f függvényt G-n előállítja az a pont körüli Taylor-sora:*

$$f(x)=\sum_{k=0}^{\infty}\frac{f^{(k)}(a)}{k!}(x-a)^k\qquad(x\in G).$$

Bizonyítás Legyen $x\in G$. A korábban már említettek alapján tetszőleges $n\in\mathbb{N}$-re létezik olyan $c_{n,x}\in G$, hogy

$$f(x)-\sum_{k=0}^{n}\frac{f^{(k)}(a)}{k!}(x-a)^k=\frac{f^{(n+1)}(c_{n,x})}{(n+1)!}(x-a)^{n+1}.$$

Tudjuk, hogy $|f^{(n+1)}(c_{n,x})| \leq K$, ezért

$$\left| f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \right| \leq \frac{|x-a|^{n+1}}{(n+1)!} K.$$

Minden x esetén $\lim\limits_{n} \dfrac{(x-a)^{n+1}}{(n+1)!} = 0$, ezért az egész jobb oldal nullához tart, ha n tart a végtelenhez. Ez azt jelenti, hogy a részletösszegek tetszőleges $x \in G$ pontban $f(x)$-hez konvergálnak, és ezt akartuk bizonyítani. ■

Meg kell jegyeznünk, hogy az iménti állításban szereplő feltétel elégséges, de nem szükséges a Taylor-sor konvergenciájához.

Vegyük észre azt is, hogy az állítás pontonkénti konvergenciát biztosít, azonban a Taylor-sor – lévén hatványsor – lokálisan egyenletesen konvergens a G környezeten.

9.4. Ha $f : \mathbb{R} \rightarrowtail \mathbb{R}$ differenciálható a-ban, akkor – éppen a differenciálszámítás alapgondolatához elvezető példának megfelelően – az $(a, f(a))$ ponton átmenő, $f'(a)$ meredekségű egyenest az f függvény a-beli **érintőjének** nevezzük.

Formulával: f-nek a-beli érintője az $\mathbb{R} \to \mathbb{R}$, $x \mapsto f(a) + f'(a)(x-a)$ függvény (vagy ennek a grafikonja). Könnyen felismerhetjük, hogy ez éppen az f függvény a-hoz tartozó elsőfokú Taylor-polinomja. Az f függvény a-beli l érintőjét a következő tulajdonságok határozzák meg:

(i) l elsőfokú polinom,

(ii) $l(a) = f(a)$ és $l'(a) = f'(a)$.

Az n-szer differenciálható függvénynek a-hoz tartozó n-ed fokú Taylor-polimomját pedig az határozza meg, hogy

(i) n-ed fokú polinom,

(ii) k-adik deriváltja a-ban megegyezik a függvény a-beli k-adik deriváltjával, ahol $k = 0, 1, \ldots, n$.

Szokás ezért az n-ed fokú Taylor-polinomot **n-ed rendű érintőnek** is nevezni.

Sőt, általában az $f, g : \mathbb{R} \rightarrowtail \mathbb{R}$ függvényekről azt mondják, hogy az $a \in \mathrm{Dom}(f) \cap \mathrm{Dom}(g)$ pontban n-ed rendben érintkeznek, ha

(i) f és g n-szer differenciálható a-ban,

(ii) $f^{(k)}(a) = g^{(k)}(a), \quad k = 0, 1, \ldots, n.$

9.5. Bár a Taylor-formula a 9.1. formájában csak $\mathbb{R} \rightarrowtail \mathbb{R}$ függvényekre igaz – hiszen $n = 1$ esetén a Taylor-formula a Lagrange-féle középértéktétel –, egy végtelenszer differenciálható $f : \mathbb{K} \rightarrowtail \mathbb{K}$ függvényre is értelmezhetjük az a pont körüli Taylor-sort, és felvethetjük a 9.3.-ban leírt (i) és (ii) kérdést.

Definíció *Egy $\mathbb{K} \rightarrowtail \mathbb{K}$ függvényt **analitikusnak** hívunk, ha végtelenszer differenciálható, és az értelmezési tartományának minden pontja rendelkezik olyan környezettel, hogy a szóban forgó pont körüli Taylor-sor összege egyenlő ebben a környezetben magával a függvénnyel.*

Korábban már említettük, hogy egy $\mathbb{C} \rightarrowtail \mathbb{C}$ differenciálható függvény végtelenszer differenciálható. Későbbi tanulmányaink során (lásd Analízis VI.) azt is látni fogjuk, hogy az ilyen függvény szükségképpen analitikus is; hangsúlyozzuk, valós függvényekre ez nem igaz: a 8.3.(ii)-beli függvény végtelenszer differenciálható, de nem analitikus.

Az analitikus függvények egy alapvető tulajdonsága a következő.

Állítás *Ha egy összefüggő halmazon értelmezett analitikus függvény egy nemüres nyílt halmazon konstans, akkor mindenütt konstans.*

Bizonyítás Vegye fel az $f : \mathbb{K} \rightarrowtail \mathbb{K}$ analitikus függvény a konstans α értéket a $\emptyset \neq H \subset \mathrm{Dom}(f)$ nyílt halmazon; vehetjük H-t összefüggőnek. Ekkor (Analízis III.A.4.4.)

$$D := \bigcup\{N \subset \mathrm{Dom}(f) \mid N \text{ összefögg\H{o}, nyílt }, H \subset N, f|_N = \alpha\}$$

nemüres összefüggő nyílt halmaz, és $f|_D = \alpha$.

Tegyük fel, hogy $D \neq \mathrm{Dom}(f)$. Ekkor f, lévén folytonos, az α konstans a értéket veszi fel a $\overline{D} \cap \mathrm{Dom}(f)$ halmazon is, és ugyanilyen indokok miatt $f^{(n)}|_{\overline{D}} = 0$ minden $n \in \mathbb{N}$ esetén.

A feltételezésünk szerint létezik $a \in (\overline{D} \setminus D) \cap \mathrm{Dom}(f) D$. Az ilyen a-nak létezik $G(a) \in \mathrm{Dom}(f)$ környezete, amelyben a Taylor-sor a függvény állítja elő; azonban az előbb mondottak miatt a Taylor-sor minden együtthatója, a nulladikat kivéve, nulla. Ez azt jelenti, hogy $f(x) = f(a) = \alpha$ minden $x \in G(a)$ esetén, azaz f konstans a D-nél bővebb $D \cup G(a)$ összefüggő nyílt halmazon, ami lehetetlen. Ezek szerint $\mathrm{Dom}(f) = D$, tehát f konstans az egész értelmezési tartományán.

9.6. (i) Az $f := \sum_{k=0}^{\infty} d_k(\mathrm{id}_{\mathbb{K}} - b)^k$ hatványsorral értelmezett $\mathbb{K} \rightarrowtail \mathbb{K}$ függvény – ha a konvergencia-sugár nem nulla – analitikus. Ugyanis végtelenszer differenciálható, amint az a 4.4. állításból következik, és ha a a konvergenciakörön belüli pont, akkor

$$\frac{f^{(n)}(a)}{n} = \frac{1}{n!}\sum_{k=n}^{\infty} c_k \frac{k!}{(k-n)!}(b-a)^{k-n},$$

amiből látjuk (Analízis III.A.24.4.), hogy az a körüli Taylor-sor az adott hatványsornak az átrendezése.

Tehát úgy is átfogalmazhatjuk a definíciót, hogy egy $\mathbb{K} \rightarrowtail \mathbb{K}$ függvény analitikus, ha értelmezési tartományának minden pontjának egy környezetében egy hatványsor összegével egyenlő.

(ii) Az előbbiekből nyilvánvaló, hogy ha $f, g : \mathbb{K} \rightarrowtail \mathbb{K}$ analitikus függvények, akkor $f + g$ és fg is analitikus. Mi több, $\frac{f}{g}$ is analitikus.

Ez utóbbi a $\mathbb{C} \rightarrowtail \mathbb{C}$ esetben azért igaz, mert differenciálható függvény analitikus (lásd Analízis VI.), és differenciálató fügvény reciproka differenciálható.

Egy $\mathbb{R} \rightarrowtail \mathbb{R}$ analitikus függvény pedig kiterjeszthető $\mathbb{C} \rightarrowtail \mathbb{C}$ analitikus függvénnyé: valós hatványsor – $\mathrm{id}_{\mathbb{R}}$ helyett $\mathrm{id}_{\mathbb{C}}$-vel – komplex hatványsornak is tekinthető.

(iii) Ha $f : \mathbb{K} \to \mathbb{K}$ végtelenszer differenciálható és f' analitikus, akkor f is analitikus. Valóban, ha az $a \in \mathrm{Dom}(f)$ egy környezetében levő x-ekre

$$f'(x) = \sum_{n=0}^{\infty} c_n (x-a)^n,$$

akkor

$$f(x) = f(a) + \sum_{n=1}^{\infty} \frac{c_{n-1}}{n}(x-a)^n.$$

Nyilvánvaló az általánosítás arra az esetre, amikor a függvény valahanyadik deriváltja analitikus.

(iv) A $\sqrt{} : \mathbb{R}^+ \to \mathbb{R}$ függvény analitikus (Analízis III.A.24.5.2.).

(v) Az előbb mondottak valamint az 5. pont formulái alapján az elemi függvények és inverzeik analitikus függvények.

(vi) Ha $f : \mathbb{K} \rightarrowtail \mathbb{K}$ analitikus függvény, és $f(a) = 0$, akkor

$$\left(\frac{f}{\mathrm{id}_{\mathbb{K}} - a}\right)_a : \mathbb{K} \rightarrowtail \mathbb{K}, \quad x \mapsto \begin{cases} \frac{f(x)}{x-a} & \text{ha } x \neq a, \\ f'(a) & \text{ha } x = a \end{cases}$$

analitikus függvény, amit az f-nek a körüli Taylor-sorából azonnal láthatunk.

Következésképpen, ha $f, g : \mathbb{K} \rightarrowtail \mathbb{K}$ analitikus függvények, $f(a) = g(a) = 0$, $g'(a) \neq 0$, és minden $x \neq a$ esetén $g(x) \neq 0$, akkor a

$$\mathbb{K} \rightarrowtail \mathbb{K}, \quad x \mapsto \begin{cases} \frac{f(x)}{g(x)} & \text{ha } x \neq a, \\ \frac{f'(a)}{g'(a)} & \text{ha } x = a \end{cases}$$

függvény analitikus, hiszen egyenlő az

$$\left(\frac{f}{\mathrm{id}_{\mathbb{K}} - a}\right)_a \frac{1}{\left(\frac{g}{\mathrm{id}_{\mathbb{K}} - a}\right)_a}$$

függvénnyel.

9.7. Feladatok

1. Bizonyítsuk be a Taylor-formulát más úton: értelmezzük a $g : [a, b] \to \mathbb{R}$ függvényt az

$$x \mapsto f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(b-x)^k - \frac{\lambda}{(n+1)!}(b-x)^{n+1}$$

formulával, ahol

$$\lambda := \frac{(n+1)!}{(b-a)^{n+1}} \left(f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k \right).$$

Ekkor $g(a) = g(b) = 0$, és alkalmazható a Rolle-tétel.

2. Adjuk meg a következő függvényeknek a 0 körüli negyed fokú Taylor-polinomját:

(i) tg, (ii) th, (iii) arctg, (iv) arth.

3. Írjuk föl a következő függvények a-hoz tartozó n-ed fokú Taylor-polinomját:

(i) sin, $a := 1$, $n := 4$;

(ii) $\mathrm{id}_{\mathbb{R}}^2 \mathrm{tg}$, $a := 0$, $n := 4$;

(iii) $\dfrac{1}{\mathrm{id}_{\mathbb{R}}}$, $a := 1$, $n := 5$;

(iv) $\dfrac{1}{1 + \mathrm{id}_{\mathbb{R}}^2}$, $a := 0$, $n := 3$;

(v) $\dfrac{1 - \mathrm{id}_{\mathbb{R}}}{1 + \mathrm{id}_{\mathbb{R}}}$, $a := 0$, $n := 3$;

(vi) $\sin^2$, $a := 0$, $n := 4$.

4. Érintkeznek-e, és ha igen, hányad rendben a következő függvények:

(i) $\dfrac{1}{1 + \mathrm{id}_{\mathbb{R}}^2}$ és $\sqrt{1 - \mathrm{id}_{\mathbb{R}}^2}$,

(ii) $\dfrac{1}{1 + \mathrm{id}_{\mathbb{R}}^2}$ és $\sqrt{\frac{1}{4} - \mathrm{id}_{\mathbb{R}}^2} + \frac{1}{2}$,

(iii) sin és tg,

(iv) cos és $\sqrt{1 - \mathrm{id}_{\mathbb{R}}^2}$?

5. Igazoljuk, hogy $\dfrac{\mathrm{arctg}}{\mathrm{id}_{\mathbb{R}}}$ – amely a nullában nincs értelmezve – kiterjeszthető mindenütt (tehát a nullában is) értelmezett analitikus függvénnyé.

6. Bizonyítsuk be, hogy az

$$L : \mathbb{R} \to \mathbb{R}, \quad x \mapsto \begin{cases} \mathrm{cth} x - \frac{1}{x} & \text{ha } x \neq 0, \\ 0 & \text{ha } x = 0 \end{cases}$$

úgynevezett Langevin-függvény analitikus, és

$$L(x) = \frac{x}{3} - \frac{x^3}{45} + \dots,$$

$$L'(x) = \begin{cases} -\frac{1}{\operatorname{sh}^2 x} + \frac{1}{x^2} > 0 & \text{ha} \quad x \neq 0, \\ \frac{1}{3} & \text{ha} \quad x = 0, \end{cases}$$

$$L''(x) = \begin{cases} 2\left(\frac{\operatorname{ch} x}{\operatorname{sh}^3 x} - \frac{1}{x^3}\right) & \text{ha} \quad x \neq 0, \\ 0 & \text{ha} \quad x = 0. \end{cases}$$

10. Függvénydiszkusszió

10.1. Definíció *(i) Egy $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvényre azt mondjuk, hogy* ***konvex (konkáv)*** *az $I \subset \mathrm{Dom}(f)$ intervallumon, ha bármely két $x_1, x_2 \in I$ pontra és tetszőleges $\lambda_1, \lambda_2 > 0,\ \ \lambda_1 + \lambda_2 = 1$ számra fennáll az*

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2)$$

$$\Big(f(\lambda_1 x_1 + \lambda_2 x_2) \geq \lambda_1 f(x_1) + \lambda_2 f(x_2)\Big)$$

egyenlőtlenség. Ha a $\leq$ $(\geq)$ reláció helyett az erősebb $<$ $(>)$ reláció tejesül, akkor ***szigorúan konvex (konkáv)*** *függvényről beszélünk.*

A határozottság kedvéért a továbbiakban mindig azt vesszük, hogy $x_1 < x_2$. Legyen $x \in]x_1, x_2[$. Ekkor $\lambda_1 := \frac{x_2 - x}{x_2 - x_1} > 0$, $\lambda_2 := \frac{x - x_1}{x_2 - x_1} > 0$ és $\lambda_1 + \lambda_2 = 1$. Ezért a fenti feltétel konvex esetben

$$f(x) \leq \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2)$$

alakban is felírható, ami ekvivalens az

$$f(x) \leq f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1), \tag{$*$}$$

$$f(x) \leq f(x_2) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x_2 - x) \tag{$**$}$$

egyenlőtlenségek bármelyikével. Hasonló ekvivalens feltételeket írhatunk föl konkáv, illetve szigorúan konvex és szigorúan konkáv esetekre is, ha a $\leq$ relációt a megfelelő másik relációval ($\geq$, $<$, $>$) helyettesítjük.

Nyilvánvaló, hogy f pontosan akkor konvex (konkáv) az I intervallumon, ha bármely $x_1, x_2 \in I$ esetén az f grafikonjának $\{(x, f(x)) \mid x \in [x_1, x_2]\}$ része nincs az $(x_1, f(x_1))$ és az $(x_2, f(x_2))$ pontokat összekötő szakasz felett (alatt); illetve pontosan akkor szigorúan konvex (konkáv), ha a grafikon szóban forgó része ezen szakasz alatt (fölött) van.

10.2. Állítás *Legyen $f : \mathbb{R} \rightarrowtail \mathbb{R}$ differenciálható az $I \subset \mathrm{Dom}(f)$ intervallumon. Az f függvény akkor és csak akkor konvex (konkáv) I-n, ha $f'|_I$ monoton növekszik (csökken), és akkor és csak akkor szigorúan konvex (konkáv), ha $f'|_I$ szigorúan monoton növekszik (csökken).*

Bizonyítás (i) Tegyük fel, hogy f konvex I-n, és legyen $x_1, x_2 \in I$, $x_1 < x_2$. Ekkor a 10.1. $(*)$ és $(**)$ képletei alapján minden $x \in]x_1, x_2[$ esetén

$$\frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x) - f(x_2)}{x - x_2}.$$

Ebből először az $x \to x_1 + 0$ majd az $x \to x_2 - 0$ határesetet véve azt kapjuk, hogy

$$f'(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'(x_2),$$

tehát f' monoton növekszik az I intervallumon.

(ii) Most azt tegyük fel, hogy $f'|_I$ monoton növekszik. Legyen $x_1, x_2 \in I$, $x_1 < x_2$; vezessük be az $r : I \to \mathbb{R}$ függvényt az

$$r(x) := f(x) - f(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1)$$

formulával. Nyilvánvaló, hogy r differenciálható $[x_1, x_2]$-n és $r(x_1) = r(x_2) = 0$, így alkalmazhatjuk rá ezen az intervallumon a Rolle-féle középértéktételt: létezik olyan $\xi \in]x_1, x_2[$, hogy

$$r'(\xi) = f'(\xi) - \frac{f(x_2) - f(x_1)}{x_2 - x_1} = 0.$$

Tekintve, hogy r' csak egy konstansban különbözik f'-től és f' monoton növekszik, r' is monoton növő. $r'(\xi) = 0$ miatt ez azt jelenti, hogy r' nem pozitív $[x_1, \xi]$-n és nem negatív $[\xi, x_2]$-n. Ezek szerint r monoton csökken $[x_1, \xi]$-n és monoton növekszik $[\xi, x_2]$-n. Figyelembe véve, hogy $r(x_1) = r(x_2) = 0$, arra következtethetünk, hogy r nem pozitív az egész $[x_1, x_2]$-n, amiből az r értelmezése alapján

$$f(x) \leq f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) \qquad (x \in [x_1, x_2])$$

adódik. Ezzel beláttuk, hogy f konvex az I intervallumon.

(iii) Ha f' szigorúan monoton növekszik I-n, akkor az iménti gondolatmenetet megismételve azt kapjuk, hogy r negatív az egész $]x_1, x_2[$-n, vagyis f szigorúan konvex I-n.

(iv) Ha f szigorúan konvex I-n, akkor f' monoton növekszik, ahogy ezt (i)-ben láttuk. Viszont $f'(x_1) = f'(x_2)$, $x_1 < x_2$ nem teljesülhet, ugyanis akkor f' konstans $[x_1, x_2]$-n, tehát f elsőfokú polinom volna, ami ellentmond annak, hogy f szigorúan konvex. Tehát f' nem csak monoton, hanem szigorúan monoton is I-n.

A konkáv függvényekre vonatkozó állítás teljesen hasonló módon bizonyítható, de érvelhetünk egyszerűbben is: elég azt észrevenni, hogy f pontosan akkor konvex, ha $-f$ konkáv (és viszont), így a konvex függvényekre igazolt állítás rögtön átvihető a konkáv függvények esetére. ■

Nyilvánvaló következménye az iménti állításnak, hogy ha egy $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvény kétszer differenciálható az I intervallumon, akkor ott f pontosan akkor konvex (konkáv), ha $f''|_I \geq 0$ $(f''|_I \leq 0)$. Ha $f''|_I > 0$ $f''|_I < 0$, akkor f szigorúan konvex (konkáv) I-n; fordítva azonban nem igaz. Ugyanis ha f szigorúan konkáv I-n, vagyis f' szigorúan monoton növekszik I-n, attól még előfordulhat olyan $x \in I$, amlyere $f''(x) = 0$.

10.3. Az (intervallumra vonatkozó) konvexitás fogalma mellett differenciálható függvények esetében be szokták vezetni az úgynevezett lokális (vagy pontbeli) konvexitás fogalmát is.

Definíció *Lokálisan konvexnek (konkávnak) hívjuk az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvényt az $a \in \mathrm{Dom}(f)$ pontban, ha f differenciálható a-ban, és létezik a-nak olyan $G(a) \subset \mathrm{Dom}(f)$ környezete, hogy bármely $x \in G(a) \setminus \{a\}$ pontra fennáll az*

$$f(x) \geq f(a) + f'(a)(x-a) \qquad \Big(f(x) \leq f(a) + f'(a)(x-a)\Big)$$

egyenlőtlenség. ***Szigorúan lokálisan konvex (konkáv)*** *függvényről beszélünk, ha a $\geq$ $(\leq)$ reláció helyett az erősebb $>$ $(<)$ reláció teljesül.*

Ez az értelmezés például lokálisan konvex esetben geometriailag azt jelenti, hogy az $f|_{G(a)}$ grafikonja nincs az f a-beli érintője alatt.

Állítás *Ha az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvény az I intervallumon differenciálható és (szigorúan) konvex (konkáv), akkor f az I minden pontjában (szigorúan) lokálisan konvex (konkáv).*

Bizonyítás Ha f konvex I-n, akkor az előző állítás szerint f' monoton növekszik I-n. Legyen a az I intervallum belsejében, és alkalmazzuk a Lagrange-féle középértéktételt az a egy környezetében levő x-re; van olyan c_x az a és az x

között, hogy
$$f(x) = f(a) + f'(c_x)(x - a).$$
A mondottak szerint, ha $x < a$ és ezzel együtt $c_x < a$, akkor $f'(c_x) \leq f'(a)$; ha $x > a$ és ezzel együtt $c_x > a$, akkor $f'(c_x) \geq f'(a)$; ezért mindkét esetben azt kapjuk, hogy
$$f(x) \geq f(a) + f'(a)(x - a),$$
vagyis f lokálisan konvex a-ban.

Az olvasóra bízzuk, hogy a fentiek mintjára végezze el a bizonyítást a szigorúan konvex esetre is.

10.4. Definíció *Az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ differenciálható függvénynek az $a \in \mathrm{Dom}(f)$* ***inflexiós pontja*** *(a-ban az f-nek* ***inflexiója van****), ha létezik olyan $\delta > 0$, hogy az $]a - \delta, a[$ intervallumon f szigorúan konkáv vagy konvex, az $]a, a + \delta[$ intervallumon pedig szigorúan konvex illetve konkáv.*

Felhívjuk a figyelmet, hogy az inflexiós pontot csak differenciálható függvényre értelmeztük, vagyis az inflexiós pontban a függvény "törés nélkül (differenciálhatóan) megy át" konvexből konkávba vagy viszont.

Geometriailag az inflexiót a következőképpen lehet szemléltetni. Tegyük fel, hogy az f függvény differenciálható egy nyílt intervallumon, és az intervallum egy a pontjában inflexiója van. Ekkor az előző állítás szerint az a ponton "áthaladva" az f függvény grafikonja "átmegy" az a-beli érintőegyenes "egyik oldaláról" a "másik oldalára", vagyis az a pontnak van egy olyan környezete, amelyben a ponttól balra a függvény grafikonja az a-beli érintő felett (alatt), a ponttól jobbra pedig ezen érintő alatt (felett) helyezkedik el, más szóval az $x \mapsto f(x) - \big(f(a) - f'(a)(x - a)\big)$ függvény előjelet vált az a pontban.

Állítás *Legyen $f : \mathbb{R} \rightarrowtail \mathbb{R}$ differenciálható függvény.*

(i) Ha f kétszer differenciálható az a pontban és ott inflexiója van, akkor $f''(a) = 0$.

(ii) Ha f háromszor differenciálható az a pontban, $f''(a) = 0$ és $f'''(a) \neq 0$, akkor a az f-nek inflexiós pontja.

Bizonyítás (i) Ha a inflexiós pontja az f függvénynek, akkor létezik olyan $\delta > 0$, hogy f szigorúan konvex (konkáv) $]a - \delta, a[$-n és szigorúan konkáv (konvex) $]a, a + \delta[$-n. Ez azt jelenti, hogy f' szigorúan növekedik (csökken) $]a - \delta, a[$-n és szigorúan csökken (nő) $]a, a + \delta[$-n. Mivel f kétszer differenciálható az a pontban, f' ott folytonos. Mindebből következik, hogy f'-nek lokális maximuma (minimuma) van a-ban, ezért a 2.5. állítás szerint $f''(a) = 0$.

(ii) Az f függvény háromszor differenciálható a-ban, ezért a belső pontja f'' értelmezési tartományának. Tekintve, hogy $f'''(a) \neq 0$, a 2.4. állítás miatt létezik olyan $G(a) \subset \mathrm{Dom}(f'')$ környezete a-nak, amelyben tetszőleges $x, y \in G(a)$

pontokra $x < a < y$ esetén

$$f''(x) < f''(a) < f''(y) \quad \text{vagy} \quad f''(x) > f''a) > f''(y).$$

Mivel $f''(a) = 0$, ez azt jelenti, hogy $G(a)$-n belül a-tól balra az f függvény szigorúan konkáv illetve konvex, a-tól jobbra pedig szigorúan konvex illetve konkáv. Ezek szerint a inflexiós pontja az f függvénynek. ■

Célszerű itt visszautalnunk a 8.6. állításra. Ott az $f : \mathbb{R} \rightarrowtail \mathbb{R}$ függvény n-szer differenciálható volt az a pontban, és fennálltak az

$$f'(a) = f''(a) = \cdots = f^{(n-1)}(a) = 0, \quad f^{(n)}(a) \neq 0$$

egyenlőségek. Azt állítottuk, hogy ha n páros, akkor (és csak akkor) a-ban az f-nek szélsőértéke van. Összevetve az ottani bizonyítást az iménti állítással könnyen láthatjuk, hogy ha n páratlan, akkor (és csak akkor) a-ban az f-nek inflexiója van.

10.5. Az előzőekben szó esett függvények érintőiről. Most a végérintőket értelmezzük, olyan egyeneseket, amelyek a "végtelenben" érintik a függvényt.

Definíció *Azt mondjuk, hogy az $f :]a, +\infty[\to \mathbb{R}$ $\big(f :]-\infty, a[\to \mathbb{R}\big)$ differenciálható függvénynek $+\infty$-ben ($-\infty$-ben)* ***aszimptotája (végérintője)*** *van, ha létezik olyan $l : \mathbb{R} \rightarrowtail \mathbb{R}$ elsőfokú polinom (azaz $l = c\,\mathrm{id}_{\mathbb{R}} + b$ valamely c és b valós számokkal), hogy*

$$\lim_{x \to +\infty} \big(f(x) - l(x)\big) = 0 \qquad \Big(\lim_{x \to -\infty} \big(f(x) - l(x)\big) = 0\Big),$$

és

$$\lim_{x \to +\infty} \big(f'(x) - l'(x)\big) = 0 \qquad \Big(\lim_{x \to -\infty} \big(f'(x) - l'(x)\big) = 0\Big).$$

Ezt az l függvényt hívjuk az f ***aszimptotájának (végérintőjének)****.*

Megjegyezzük, hogy az aszimptota definíciójából a második feltételt olykor elhagyják.

Az alábbi állítás bizonyítása rendkívül egyszerű.

Állítás *Az* $f :]a, +\infty[\to \mathbb{R}$ $\big(f :]-\infty, a[\to \mathbb{R}\big)$ *differenciálható függvénynek a* $c\,\mathrm{id}_{\mathbb{R}} + b$ *elsőfokú polinom akkor és csak akkor aszimptotája* $+\infty$*-ben* $(-\infty$*-ben), ha*

$$\lim_{x\to+\infty}\big(f(x) - cx\big) = b \qquad \left(\lim_{x\to-\infty}\big(f(x) - cx\big) = b\right),$$

és

$$\lim_{x\to+\infty} f'(x) = c \qquad \left(\lim_{x\to-\infty} f'(x) = c\right).$$

Vegyük észre azt is, hogy a

$$\lim_{x\to+\infty}\big(f(x) - cx\big) = b$$

egyenlőségből

$$\frac{f(x)}{x} = \frac{f(x) - cx}{x} + c$$

miatt

$$\lim_{x\to+\infty} \frac{f(x)}{x} = c$$

következik, tehát a fenti egyenlőség szükséges az aszimptota létezéséhez (persze ugyanez a helyzet $-\infty$ esetében is). Ez azonban nem elégséges, amire jó példa a $\sqrt{}$ függvény: $\lim_{x\to+\infty} \frac{\sqrt{x}}{x} = 0$, viszont $(\sqrt{x} - 0x)$-nak nincs határértéke, miközben x tart a $+\infty$-hez.

10.6. Egy $\mathbb{R} \rightarrowtail \mathbb{R}$ függvény "megismeréséhez" nagyban hozzájárul az, ha el tudjuk képzelni (fel tudjuk vázolni) a grafikonját. Ehhez célszerű tisztázni a következőket:

a) mi a függvény értelmezési tartománya, hol folytonos, hol differenciálható, van-e jobb- vagy baloldali határértéke a szakadási pontokban (ott, ahol nem folytonos) illetve az értelmezési tartományának torlódási pontjaiban,

b) metszi-e és hol a függvény grafikonja a koordinátatengelyeket;

c) vannak-e a függvénynek lokális illetve globális szélsőértékei d) vannak-e olyan intervallumok a függvény értelmezési tartományában, ahol a függvény monoton, illetve konvex vagy konkáv;

e) mik a függvény inflexiós pontjai (ha léteznek);

f) hogyan viselkedik a függvény a végtelenben (ha nem korlátos az értelmezési tartománya), van-e aszimptotája.

Egy $\mathbb{R} \rightarrowtail \mathbb{R}$ függvény diszkussziója alatt tulajdonképpen a fenti kérdések megválaszolását értjük. A differenciálszámítás megismert eszközei segítséget nyújthatnak a c)-f) kérdések vizsgálatában.

10.7. Feladatok

1. Fejezzük be a 10.3. állítás bizonyítását.

2. Van-e az $f\ :]a, +\infty[\to \mathbb{R}$ differenciálható függvénynek aszimptotája $+\infty$-ben, ha létezik $\lim\limits_{x\to+\infty} f'(x)$? (Vizsgáljuk meg az $f(x) := cx + \sin(\log x)\quad (x \in \mathbb{R}^+)$ függvényt.)

3. Diszkutáljuk a következő függvényeket:
(i) $\mathrm{id}_{\mathbb{R}} - \log$, (ii) $\sin^3 + \cos^3$, (iii) $\mathrm{ch} + \cos$,
(iv) $6\log - \mathrm{id}_{\mathbb{R}} - 10\mathrm{arctg}$.

4. Mutassuk meg, hogy az L Langevin-függvény (lásd 9.7.6.) páratlan, azaz $L(-x) = -L(x)$, továbbá $L' > 0$, és $L''(x) < 0$ ha $x > 0$, $L''(x) > 0$, ha $x < 0$; tehát L szigorúan monoton nő, a pozitív félegyenesen konkáv, a negatív félegyenesen konvex; ezért $L(x) < \frac{x}{3}$ ha $(x > 0)$.

Végül lássuk be, hogy $\lim\limits_{x\to\infty} L(x) = 1$.

III. GÖRBÉK

11. Féloldali deriváltak

11.1 Definíció *Az $f:\mathbb{R}\rightarrowtail\mathbb{K}^N$ függvényt* ***jobbról (balról) differenciálhatónak*** *hívjuk az a pontban, ha van olyan $\delta>0$, hogy $[a,a+\delta]\subset\mathrm{Dom}(f)$ $\big([a-\delta,a]\subset\mathrm{Dom}(f)\big)$, és létezik a*

$$f'(a+0):=\lim_{x\to a+0}\frac{f(x)-f(a)}{x-a}\qquad\left(f'(a-0):=\lim_{x\to a-0}\frac{f(x)-f(a)}{x-a}\right)$$

határérték, amelyet az f függvény a pontbeli ***jobb oldali (bal oldali)*** *deriváltjának nevezünk.*

A jobb és bal oldali deriváltakat közösen ***féloldali****deriváltaknak szokták nevezni.*

Ugyanúgy, mint 1.3-ban, megmutathatjuk, hogy ha egy függvény balról (jobbról) differenciálható egy pontban, akkor ott balról (jobbról) folytonos is.

Visszaemlékezve a függvények pontbeli határértékének tulajdonságaira, nem szorul különösebb magyarázatra a következő

Állítás *Egy $f:\mathbb{R}\rightarrowtail\mathbb{K}^N$ függvény pontosan akkor differenciálható az értelmezési tartományának egy a belső pontjában, ha mindkét oldalról differenciálható ebben a pontban és $f'(a+0)=f'(a-0)$.*

11.2. Emlékeztetünk arra, hogy eddigi értelmezésünk szerint egy $\mathbb{R}\rightarrowtail\mathbb{K}^N$ differenciálható függvény értelmezési tartománya nyílt kell hogy legyen. Ezért most külön bevezetjük a zárt intervallumon értelmezett differenciálható függvény fogalmát.

Definíció *Az* $f : [a, b] \to \mathbb{K}^N$ *függvényt* ***differenciálhatónak*** *nevezzük, ha* f *differenciálható* $]a, b[$*-n és léteznek az* $f'(a+0)$, $f'(b-0)$ *féloldali deriváltak, amelyeket az egységes írásmód kedvéért* $f'(a)$*-val illetve* $f'(b)$*-vel jelölünk.*

A függvény ***folytonosan differenciálható****, ha differenciálható, és az előbb meghatározott* $f' : [a, b] \to \mathbb{K}^N$ *deriváltfüggvény folytonos.*

Ezek után világos, nem részletezzük, mit értünk azon, hogy egy $[a, b[\to \mathbb{K}^N$ illetve $]a, b] \to \mathbb{K}^N$ függvény differenciálható, ahol az első esetben $b = \infty$, a másik esetben $a = -\infty$ is lehetséges.

A mondottakból az is egyszerűen következik, hogy egy (nem szükségképpen nyílt) intervallumon differenciálható függvény folytonos.

11.3. Definíció *Az* $f : \mathbb{R} \to \mathbb{K}^N$ *függvényt* ***szakaszosan (folytonosan) differenciálhatónak*** *hívjuk, ha*

(i) intervallumon van értelmezve,

(ii) létezik véges sok olyan $a_1 < a_2 < \cdots < a_n$ *pontja az értelmezési tartományának, hogy* f *(folytonosan) differenciálható a* $\mathrm{Dom}(f) \cap] - \infty, a_1]$, $[a_1, a_2]$, $\ldots$, $\mathrm{Dom}(f) \cap [a_n, \infty[$ *intervallumokon.*

Felhívjuk a figyelmet, hogy egy szakaszosan differenciálható függvény folytonos, mert végpontjaikban érintkező intervallumokon folytonos.

Viszont egy szakaszosan akár folytonosan differenciálható függvény nem szükségképpen differenciálható, hiszen általában $f'(a_i - 0) \neq f'(a_i + 0)$, ahol a_i a definícióban szereplő "töréspontok" valamelyike.

Például az $\mathbb{R} \to \mathbb{R}$ abszolútérték-függvény szakaszosan differenciálható, de nem differenciálható.

11.4. Feladatok

1. Igazoljuk, hogy az alábbi függvények szakaszosan differenciálhatók:

(i) $[0, 1] \to \mathbb{R}, \quad x \mapsto \begin{cases} x & \text{ha } 0 \le x \le 1/2, \\ 1 - x & \text{ha } 1/2 < x \le 1; \end{cases}$

(ii) $] - \infty, 1] \to \mathbb{R}, \quad x \mapsto \begin{cases} -x & \text{ha } x \le 0, \\ x^2 & \text{ha } 0 < x \le 1. \end{cases}$

2. Igaz-e, hogy a következő függvények szakaszosan differenciálhatók:

(i) $[-1, 1] \to \mathbb{R}, \quad x \mapsto \begin{cases} 1 - x^2 & \text{ha } -1 \le x \le 0, \\ \cos x, & \text{ha } 0 < x \le 1; \end{cases}$

(ii) $]3, \infty[\to \mathbb{R}, \quad x \mapsto \begin{cases} \sin x & \text{ha } -3 < x \le 0, \\ x + 1 & \text{ha } 0 < x. \end{cases}$

12. Görbék paraméterezése és irányítása

12.1 Definíció *A $\mathbb{K}^N$ egy L részhalmazát* ***görbének*** *hívjuk, ha létezik olyan I nyílt intervallum és olyan $p : I \to \mathbb{K}^N$ függvény, hogy*

(i) $\mathrm{Ran}(p) = L$,
(ii) p folytonosan differenciálható,
(iii) bármely $t \in I$ esetén $\dot{p}(t) \neq 0$,
(iv) p injektív, és az inverze folytonos.

A p függvényt az L görbe ***paraméterezésének*** *nevezzük.*

Vegyük észre, hogy a p paraméterezés is folytonos (nem csak az inverze), mert differenciálható. Azt szokás mondani, hogy p homeomorfizmus ("oda-vissza folytonos") I és L között.

Itt rögtön meg kell jegyeznünk, hogy egy adott görbe paraméterezése egyáltalán nem egyértelmű. Vegyük például a definícióban szereplő p paraméterezést, valamint egy J nyílt intervallumot, és egy olyan $\phi : J \to I$ folytonosan differenciálható bijekciót, hogy ϕ^{-1} is folytonosan differenciálható legyen (persze sok ilyen ϕ létezik). A definíció alapján könnyen látható, hogy a $p \circ \phi : J \to \mathbb{K}^N$ leképezés szintén paraméterezése lesz az L görbének.

12.2. Állítás *Legyen L görbe. Ha $p : I \to L$ és $q : J \to L$ az L paraméterezései, akkor az $S := p^{-1} \circ q : J \to I$ függvény folytonosan differenciálható, és $(\dot{p} \circ S)S' = \dot{q}$.*

Bizonyítás Az nyilvánvaló, hogy S folytonos, injektív és $q = p \circ S$. Ha tehát tudjuk, hogy S differenciálható, akkor fennáll az állított egyenlőség.

Legyen $t, s \in J$ és $t \neq s$. Tekintsük a következő átalakítást:

$$\frac{q(t) - q(s)}{t - s} = \frac{p(S(t)) - p(S(s))}{t - s} = \frac{p(S(t)) - p(S(s))}{S(t) - S(s)} \frac{S(t) - S(s)}{t - s}.$$

Nem követtünk el semmi rosszat, hiszen $S(t) - S(s) \neq 0$, mivel S injektív. Az előbbi egyenlőségből kis átrendezéssel kapjuk, hogy

$$\left| \frac{S(t) - S(s)}{t - s} \right| = \frac{\left| \frac{q(t)-q(s)}{t-s} \right|}{\left| \frac{p(S(t))-p(S(s))}{S(t)-S(s)} \right|},$$

ahol $|\cdot|$ a bal oldalon az abszolútértéket, a jobb oldalon pedig a $\mathbb{K}^N$-beli euklideszi normát jelenti.

Tudjuk, hogy S folytonos, p és q pedig folytonosan differenciálható, következésképpen léteznek a

$$\lim_{t\to s}\frac{q(t)-q(s)}{t-s}=\dot{q}(s)\neq 0,\quad \lim_{t\to s}\frac{p(S(t))-p(S(s))}{S(t)-S(s)}=\dot{p}(S(s))\neq 0$$

határértékek, tehát létezik a

$$\lim_{t\to s}\left|\frac{S(t)-S(s)}{t-s}\right|=\frac{|\dot{q}(s)|}{|\dot{p}(S(s))|}$$

határérték is. S folytonos bijekció két intervallum között, ezért szigorúan monoton, tehát $\frac{S(t)-S(s)}{t-s}$ nem vált előjelet, így abszolútérték nélkül is létezik a határértéke. Ez azt jelenti, hogy az S függvény differenciálható a J intervallum tetszőleges s pontjában. A monotonitás miatt S' sem vált előjelet, tehát

$$S'(s)=\frac{|\dot{q}(s)|}{|\dot{p}(S(s))|}\quad\text{vagy}\quad S'(s)=-\frac{|\dot{q}(s)|}{|\dot{p}(S(s))|},$$

ami mutatja, hogy S' folytonos.

12.3. Egy görbe tetszőleges p és q paraméterezésére $S := p^{-1}\circ q$ a két paraméterezés kapcsolatát tükrözi, hiszen az egyiket a másikba viszi át $q = p\circ S$ és $p = q\circ S^{-1}$ értelemben. Az iménti bizonyítás során láttuk, hogy S szigorúan monoton, és S' sehol sem vált előjelet és a nulla értéket sem veszi föl. Értelmes tehát a következő meghatározás.

1. Definíció *Egy görbe p paraméterezését a q paraméterezéssel* ***azonos (ellentétes) irányításúnak*** *nevezzük, ha a $p^{-1}\circ q$ függvény deriváltja pozitív (negatív).*

Állítás *Az azonos irányításúnak lenni ekvivalencia-reláció a görbe paraméterezéseinek halmazán, és két ekvivalencia-osztály van.*

Bizonyítás A reláció

(i) reflexív: minden paraméterezés azonosan irányított önmagával;

(ii) szimmetrikus: ha p azonos irányítású q-val, akkor q is azonos irányítású p-vel, hiszen $q^{-1}\circ p$ a $p^{-1}\circ q$ inverze, így a deriváltak lényegében egymás reciprokai;

(iii) tranzitív: ha p azonos irányítású q-val és q azonos irányítású r-rel, akkor p azonos irányítású r-rel, hiszen $p^{-1}\circ r = \big(p^{-1}\circ q\big)\circ\big(q^{-1}\circ r\big)$.

Bármely két paraméterezés azonosan vagy ellentétesen irányított (és létezik ilyen is, olyan is), ezért két ekvivalencia-osztály van. ■

2. Definíció ***Irányított görbének*** *nevezünk egy (L, o) párt, ahol L görbe, o pedig, az* ***irányítás****, az L paraméterezéseinek két ekvivalencia-osztálya közül az egyik. Egy irányított görbe valamely paraméterezését* **pozitív (negatív) irányításúnak** *mondjuk, ha ebbe (nem ebbe) a kijelölt ekvivalencia-osztályba tartozik.*

Szokásosan a jelölésből elhagyjuk az irányítást; ha azt mondjuk, L irányított görbe, odaértjük mellé az irányítását is.

12.4. A 12.1. definíció kisebb-nagyobb módosításával más görbefajtákat is értelmezhetünk (és kell is értelmeznünk a gyakorlati alkalmazások miatt).

1. Definíció *Egy $L \subset \mathbb{K}^N$ halmazt* **peremes görbének** *hívunk, ha létezik olyan I kompakt intervallum és olyan $p : I \to \mathbb{K}^N$ függvény, hogy*

(i) $\mathrm{Ran}(p) = L$,
(ii) p folytonosan differenciálható,
(iii) bármely $t \in I$ esetén $\dot{p}(t) \neq 0$;
(iv) p injektív, és az inverze folytonos.

A p függvényt az L peremes görbe **paraméterezésének** *nevezzük. Az I intervallum végpontjainak p általi képét a görbe* **perempontjainak** *hívjuk.*

2. Definíció *Az $L \subset \mathbb{K}^N$ halmazt* **zárt görbének** *nevezzük, ha van olyan $[a, b]$ intervallum és $p : [a, b] \to \mathbb{K}^N$ leképezés, hogy*

(i) $\mathrm{Ran}(p) = L$,
(ii) p folytonosan differenciálható,
(iii) bármely $t \in [a, b]$ esetén $\dot{p}(t) \neq 0$,
(iv) $p|_{]a,b[}$ injektív, és az inverze folytonos,
(v) $p(a) = p(b)$ és $\dot{p}(a) = \dot{p}(b)$.

3. Definíció *Az $L \subset \mathbb{K}^N$ halmazt* **szakaszos görbének hívjuk***, ha van olyan I intervallum és $p : I \to \mathbb{K}^N$ leképezés, hogy*

(i) $\mathrm{Ran}(p) = L$,
(ii) p szakaszosan folytonosan differenciálható,
(iii) $\dot{p}(t) \neq 0$ minden olyan $t \in [a, b]$ esetén, ahol p differenciálható,
(iv) p injektív, és az inverze folytonos.

Be lehet még vezetni a **szakaszos peremes görbét** és a **szakaszos zárt görbét**; kérjük az olvasót, adja meg ezeket a definíciókat az eddigiek alapján.

Felmerül a kérdés: egyértelműek-e a peremes görbe perempontjai, hiszen azokat egy paraméterezzéssel határoztuk meg, és természetesen több paraméterezés lehetséges.

Állítás *A peremes görbe perempontjai függetlenek a paraméterezéstől.*

Bizonyítás Legyen $p : [a, b] \to L$ és $q : [c, d] \to L$ az L peremes görbe két paraméterezése. A $p^{-1} \circ q : [c, d] \to [a, b]$ függvény folytonos bijekció, és mint ilyen, határpontot határpontba képez (Analízis III.A.30.2.), vagyis két eset lehetséges:
– vagy $\big(p^{-1} \circ q\big)(c) = a$ és $\big(p^{-1} \circ q\big)(d) = b$, tehát $q(c) = p(a)$ és $q(d) = p(b)$,
– vagy $\big(p^{-1} \circ q\big)(c) = b$ és $\big(p^{-1} \circ q\big)(d) = a$, tehát $q(c) = p(b)$ és $q(d) = p(a)$.

12.5. Lássunk néhány példát!

(i) Ha I nyílt intervallum és $f : I \to \mathbb{K}^N$ folytonosan differenciálható, akkor az f grafikonja görbe $\mathbb{K}^{1+N}$-ben, amelynek természetes paraméterezése a $p : I \to \mathbb{K}^{1+N}$, $t \mapsto (t, f(t))$ leképezés; erre $\dot{p}(t) = (1, f'(t))$.

(ii) Ha I kompakt intervallum és $f : I \to \mathbb{K}^N$ folytonosan differenciálható, akkor $\operatorname{Graph}(f) \subset \mathbb{K}^{1+N}$ peremes görbe.

(iii) Az $\big\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\big\}$ halmaz zárt görbe, egy lehetséges paraméterezése : $[-\pi, \pi] \to \mathbb{R}^2$, $t \mapsto (\cos t, \sin t)$.

(iv) Az $\big\{(x, y) \in \mathbb{R}^2 \mid |x| + |y| = 1, y \geq 0\big\}$ halmaz szakaszos görbe, amelynek egy paraméterezése a

$$[-1, 1] \to \mathbb{R}^2, \quad t \mapsto \begin{cases} (t, 1 + t) & \text{ha } -1 \leq t < 0, \\ (t, 1 - t) & \text{ha } 0 \leq t \leq 1 \end{cases}$$

függvény.

(v) Sok esetben egy görbe lezártja peremes görbe, de nem mindig. Tekintsük ugyanis az $L := \{(x, \sin(1/x)) \mid x \in \mathbb{R}^+\}$ görbét $\mathbb{R}^2$-ben. Ennek lezártja az $L \cup (\{0\} \times [-1, 1])$ halmaz, amely se nem görbe, se nem peremes görbe.

(vi) Tekintsük a következő $p : \mathbb{R} \rightarrowtail \mathbb{R}^2$ leképezést:

$$p(t) := \left(t\frac{1 + t^2}{1 + t^4},\ t\frac{1 - t^2}{1 + t^4}\right).$$

Ajánljuk az Olvasónak: ellenőrizze, hogy a görbe definíciójában szereplő (i), (ii) és (iii) feltételek teljesülnek az $L := \operatorname{Ran}(p)$ halmazra és p-re mint paraméterezésre, (iv) azonban nem. Ezáltal a szóbanforgó halmaz, amely Bernoulli-lemniszkáta néven ismert, nem lesz görbe a mi értelmezésünk szerint. Gondoljuk végig, mi a probléma a $(0, 0) \in L$ ponttal!

12.6. Feladatok

1. Adjunk meg egy szakaszos peremes görbét és egy szakaszos zárt görbét!
2. Görbék-e valamilyen értelemben (zárt, szakaszos, stb) az alábbi halmazok:
(i) $\big\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = |x|\big\}$, (ii) $\{(x, y) \in \mathbb{R}^2 \mid |x| + |y| = 1\}$,
(iii) $\{(x, y) \in \mathbb{R}^2 \mid x^3 + y^3 = 3axy, x > 0, y > 0\}$ ($a \in \mathbb{R}$ adott),
(iv) $\{(x, y) \in \mathbb{R}^2 \mid y^2(a - x) = x^3\}$ ($a \in \mathbb{R}$ adott),
(v) $\{(x, y) \in \mathbb{R}^2 \mid (x^2 + y^2)^2 = a^2(x^2 - y^2)\}$, ($a \in \mathbb{R}$ adott).

Van-e ezeknek olyan részhalmaza, amely görbe, peremes görbe, zárt görbe, szakaszos görbe? Ha igen, adjuk meg egy paraméterezésüket!

3. Szemléltessük a $(\cos t, \sin t, t) \mid t \in [0, 2\pi]\} \subset \mathbb{R}^3$ peremes görbét!

4. Adjunk meg a 12.5. (i) és (iii) példájában szereplő paraméterezésekkel azonos irányítású más paraméterezést, és velük ellentétes irányítású paraméterezést is!

13. Görbék érintői

13.1. Definíció *Az $L \subset \mathbb{K}^N$ görbe x pontjában a p paraméterezéshez tartozó **érintővektorának** a $\dot{p}\left(p^{-1}(x)\right) \in \mathbb{K}^N$ elemet nevezzük.*

Egy adott pontban különböző paraméterezésekhez más és más érintővektorok tartoznak, azonban ezek az érintővektorok "nem nagyon különböznek" egymástól. Pontosabban arról van szó, hogy egy görbe bármely x pontjában tetszőleges p és q paraméterezéshez tartozó érintővektora párhuzamos egymással. Hogy ezt belássuk, csak vissza kell emlékeznünk a 12.2. állítás képletére, amelyből a $0 \neq \lambda := S'(q^{-1}(x)) \in \mathbb{R}$ jelöléssel

$$\dot{q}\left(q^{-1}(x)\right) = \lambda \dot{p}\left(p^{-1}(x)\right)$$

adódik.

13.2. Az előző fejezetben beszéltünk görbék paraméterezéseinek azonos irányításáról; ez a mondottak szerint az érintővektorokon a szokásos azonos irányítással egyenértékű.

Állítás *Egy görbe p és q paraméterezése pontosan akkor azonos irányítású, ha bármely pontjában a két paraméterezéshez tartozó érintővektorok azonos irányításúak (azaz egymásnak pozitív számszorosai).*

Ennek megfelelően irányított görbék esetében értelemszerűen beszélhetünk egy pontban pozitív illetve negatív irányítású érintővektorokról.

13.3. Definíció *Egy görbe x pontbeli **érintőegyenesének** az*

$$\mathbb{R}\dot{p}\left(p^{-1}(x)\right) := \left\{\alpha\dot{p}\left(p^{-1}(x)\right) \mid \alpha \in \mathbb{R}\right\}$$

halmazt nevezzük, ahol p a görbe egy paraméterezése.

A korábban mondottakból következik, hogy egy pontbeli érintőegyenes nem függ a paraméterezéstől. Speciálisan azt a következtetést is levonhatjuk, hogy egy pontbeli érintőegyenes nem függ a definíciójában szereplő paraméterezés irányításától sem. Viszont a görbe egy irányítása kijelöli az érintőegyenes egy irányítását.

Az így értelmezett érintőegyenes valójában egy $\mathbb{K}^N$-beli, mégpedig az origón keresztül haladó egyenes lesz. Hogy megkapjuk a "geometriai értelemben" vett érintőjét egy görbének egy adott pontban, ahhoz el kell tolnunk az iménti egyenest az adott pontba, vagyis a definíció jelöléseit használva az

$$x + \mathbb{R}\dot{p}\left(p^{-1}(x)\right)$$

egyenest kell vennünk. Az irodalomban azonban (főleg a későbbiekben tárgyalásra kerülő sokaságoknál, melyeknek a görbék speciális esetei) az általunk is bevezetett definíció a megszokott. Ennek például az is az oka, hogy így az érintőegyenes $\mathbb{K}^N$-nek 1 dimenziós valós altere, míg ha eltoljuk a megfelelő pontba (és az nem az origó), akkor már nem lesz vektortér.

13.4. Feladatok

1. Két egymást metsző görbe metszési szögét a metszéspontbeli érintővektoraik szögeként értelmezzük. Bizonyítsuk be, hogy bármely $0 \neq a \in \mathbb{R}$ esetén az

(i) $x \mapsto a\sin(x/a)$, (ii) $x \mapsto a\mathrm{tg}(x/a)$, (iii) $x \mapsto a\log(x/a)$

függvények grafikonjai (mint görbék) ugyanakkora szögben metszik a vízszintes tengelyt, függetlenül a értékétől!

2. Az a és $b \neq 0$ paraméter milyen értékénél metszi az $x \mapsto \dfrac{x^3+ax}{b}$ függvény grafikonja a vízszintes tengelyt 45^o-s szögben?

3. Az origóból kiinduló egyenesek közül melyik metszi az $\{(x,y) \mid xy = a^2\}$ $(a \neq 0)$ hiperbolát derékszögben?

4. Bizonyítsuk be, hogy adott $a > 0$ esetén a $\varphi \mapsto (a\varphi\cos\varphi, a\varphi\sin\varphi)$ paraméterezéssel megadható Archimedész-féle spirálishoz a φ paraméterrel jellemzett ponthoz húzott érintő és az origóból az érintési pontba húzott rádiuszvektor közötti szög 90^o-hoz tart, ha $\varphi \to \infty$.

B. DIFFERENCIÁLÁS NORMÁLT TEREKBEN

I. DIFFERENCIÁLHATÓ FÜGGVÉNYEK

Ebben a részben az U, V, W nagybetűkkel – néha esetleg indexekkel is ellátva – $\mathbb{K}$ test feletti normált vektortereket jelölünk. Az ettől eltérő esetekben arra a figyelmet külön felhívjuk.

1. Differenciálhatóság, alapvető tulajdonságok

1.1. Definíció *Egy $O : V \rightarrowtail W$ függvényt* ***nagy ordó függvénynek*** *hívunk, ha*

(i) létezik olyan $G(0)$ környezete a $0 \in V$ pontnak, hogy $G(0) \setminus \{0\} \subset \mathrm{Dom}(O)$,

(ii) az $x \mapsto \frac{O(x)}{\|x\|}$ $x \in G(0) \setminus \{0\}$ függvény korlátos.

Egy $o : V \rightarrowtail W$ függvényt ***kis ordó függvénynek nevezünk****, ha*

(i) létezik olyan $G(0)$ környezete a $0 \in V$ pontnak, hogy $G(0) \setminus \{0\} \subset \mathrm{Dom}(o)$,

(ii) $\lim\limits_{x \to 0} \frac{o(x)}{\|x\|} = 0.$

Egyszerű tények a következők:

(i) Egy $O : V \rightarrowtail W$ függvény pontosan akkor nagy ordó függvény, ha van olyan $C > 0$, hogy $\|O(x)\| \leq C\|x\|$ a $0 \in V$ valamely környezetéből vett x pontokban; következésképpen $\lim\limits_{x \to 0} O(x) = 0$.

(ii) Ha $O : V \rightarrowtail W$ nagy ordó függvény, akkor az $x \mapsto \|x\|O(x)$ hozzárendeléssel értelmezett függvény kis ordó függvény. Viszont, ha $o : V \rightarrowtail W$ kis ordó függvény, akkor $x \mapsto \frac{o(x)}{\|x\|}$ nem feltétlenül nagy ordó függvény, mert ugyan a határértéke a nullában nulla, de nem biztos, hogy közben $\frac{|o(x)|}{\|x\|} \leq C\|x\|$ is teljesül valamilyen $C > 0$ számra. Tehát nem minden kis ordó függvény $x \mapsto \|x\|O(x)$ alakú, ahol O nagy ordó függvény. Például $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^{4/3}$ kis ordó függvény, de $x \mapsto \dfrac{x^{4/3}}{|x|}$ nem nagy ordó függvény. Viszont nyilvánvaló, hogy egy kis ordó függvény $\| \cdot \|\eta$

alakú, ahol $\lim\limits_{x\to 0} \eta(x) = 0$.

(iii) Egy kis ordó függvény egyben nagy ordó függvény is.

(iv) Nagy (kis) ordó függvények lineáris kombinációja nyilvánvalóan szintén nagy (kis) ordó függvény.

(v) Hogy egy $V \rightarrowtail W$ függvény kis illetve nagy ordó függvény-e, függ attól, milyen norma van megadva V-n és W-n. Ha valamely normákra vonatkozóan adott egy $V \rightarrowtail W$ ordó függvény, és V-n finomabb, W-n pedig durvább normát veszünk, akkor a függvény megmarad ugyanolyan ordó függvénynek. Természetesen ekvivalens normákra nézve ugyanazok az ordó függvények. Speciálisan, ha V és W véges dimenziós vektorterek, akkor a normától függetlenül beszélhetünk $V \rightarrowtail W$ ordó függvényekről, hiszen véges dimenziós vektortereken bármely két norma ekvivalens.

(vi) Nulla dimenziós vektortéren az ordó függvények üresek.

1.2. Definíció *Egy $F : V \rightarrowtail W$ függvényt* ***differenciálhatónak*** *hívunk az $a \in V$ pontban, ha $a \in \overset{\circ}{\mathrm{Dom}}(F)$, és létezik olyan $\mathrm{D}F(a) : V \to W$ folytonos, lineáris leképezés valamint olyan $o : V \rightarrowtail W$ kis ordó függvény, hogy*

$$F(x) - F(a) = DF(a)(x-a) + o(x-a),$$

ahol $x \in \mathrm{Dom}(F)$.

$\mathrm{D}F(a)$-t az F függvény a pontbeli ***deriváltjának*** *nevezzük.*

A fenti egyenlőséget átírhatjuk

$$F(a+h) - F(a) = \mathrm{D}F(a)h + o(h)$$

formába, ahol h a $0 \in V$ egy környezetének tetszőleges eleme.

Még másképpen így fogalmazhatunk: egy $F : V \rightarrowtail W$ függvény differenciálható az $a \in V$ pontban, ha $a \in \overset{\circ}{\mathrm{Dom}}(F)$, és létezik olyan $\mathrm{D}F(a) : V \to W$ folytonos, lineáris leképezés, hogy

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

vagy

$$\lim_{h\to 0} \frac{\|F(a+h) - F(a) - DF(a)h\|}{\|h\|} = 0. \qquad (*)$$

Felhívjuk a figyelmet arra, hogy a definícióban szereplő kis ordó függvény függ az a ponttól, azaz ha egy másik, b pontban is differenciálható a függvény, akkor ott is létezik egy megfelelő kis ordó függvény, de az más lehet, mint ami az a-hoz tartozik.

1.3. (i) A differenciálhatóság függ attól, milyen normát adunk meg V-n és W-n, hiszen mind az $V \to W$ folytonos lineáris leképezések, mind az $V \rightarrowtail W$ ordó függvények függenek a normától. Ha egy függvény differenciálható egy pontban valamely normákra vonatkozóan, és V-n finomabb, W-n pedig durvább normát veszünk, akkor a függvény differenciálható marad ezekre nézve is.

Természetesen ekvivalens normákra nézve ugyanazok a differenciálható függvények. Speciálisan, ha V és W véges dimenziós vektorterek, akkor függetlenül attól, hogy megadunk-e rajtuk normát vagy sem, beszélhetünk $V \rightarrowtail W$ függvények differenciálhatóságáról.

E megjegyzésünknek a gyakorlatban is jó hasznát vehetjük. Meg akarjuk mutatni egy $V \rightarrowtail W$ függvényről, hogy adott normára vonatkozóan differenciálható egy pontban. Ha sikerül belátnunk a differenciálhatóságot V-n durvább, W-n finomabb normában, akkor az igaz lesz az eredeti normákra nézve is.

(ii) Nulla dimenziós vektortéren értelmezett minden függvény differenciálható az egyetlen 0 pontban.

1.4. Az előbbi definícióban csak a létezését követeltük meg a függvény deriváltjának, azaz egy bizonyos tulajdonságokkal rendelkező folytonos, lineáris leképezésnek. Felmerül a kérdés, hány ilyen lehetséges.

Állítás *Ha az $F : V \rightarrowtail W$ függvény differenciálható az $a \in V$ pontban, akkor az a-beli deriváltja egyértelműen meghatározott folytonos, lineáris leképezés.*

Bizonyítás Tegyük föl, hogy $A, B : V \to W$ két olyan folytonos, lineáris leképezés, hogy van olyan $G_r(0)$ környezete az V nulla elemének, és vannak olyan o_A, $o_B : G_r(0) \to W$ kis ordó függvények, hogy

$$F(a+h) - F(a) = Ah + o_A(h) = Bh + o_B(h) \qquad (h \in G_r(0)).$$

Átrendezve az egyenlőséget azt kapjuk, hogy

$$(A - B)h = o_B(h) - o_A(h) =: o(h),$$

ahol $o : G_r(0) \to W$ szintén kis ordó függvény.

Legyen $e \in V$, $\|e\| = 1$. Ekkor minden $t \in]-r, r[$ esetén $te \in G_r(0)$, így

$$(A - B)(te) = o(te),$$

és ha $t > 0$, akkor

$$(A - B)e = \frac{o(te)}{t} = \frac{o(te)}{\|te\|}.$$

A jobb oldal nullához tart, miközben t nullához tart, ezért

$$(A - B)e = 0.$$

Mivel e tetszőleges egységvektor, és minden vektor valamely egységvektor számszorosa, továbbá A és B lineárisak, ez azt jelenti, hogy

$$A - B = 0, \quad \text{vagyis} \quad A = B,$$

és pontosan ezt akartuk bizonyítani. ■

Ha nem követelnénk meg, hogy belső pont legyen, ahol a differenciálhatóságot értelmezzük, akkor általában nem tudnánk bizonyítani a derivált egyértelműségét. Ugyanis, visszautalva az iménti bizonyításra, ha A és B nem az egész egységgömbön egyezik meg, hanem csak annak egy részén, akkor A és B nem feltétlenül egyenlőek.

1.5. Mivel $\mathbb{K}$ és $\mathbb{K}^N$ normált terek, egy $f : \mathbb{K} \rightarrowtail \mathbb{K}^N$ függvény differenciálhatóságáról beszélhetünk a "régi" és az "új" értelemben is. Így egy $a \in \overset{\circ}{\mathrm{Dom}}(f)$ pontban $f'(a) \in \mathbb{K}^N$, viszont $\mathrm{D}f(a) \in \mathrm{Lin}(\mathbb{K}, \mathbb{K}^N)$. Fölmerül a kérdés, hogy ugyanakkor differenciálható-e az "új" és a "régi" értelemben az f függvény, valamint mi a kapcsolat $f'(a)$ és $\mathrm{D}f(a)$ között. Nos, a definíciók alapján világos, hogy $\mathrm{D}f(a)$ pontosan akkor létezik, ha $f'(a)$ is létezik, továbbá tetszőleges $h \in \mathbb{K}$ esetén

$$\mathrm{D}f(a)h = hf'(a),$$

ahol a bal oldalon a lineáris leképezés hat a h-ra, a jobb oldalon a h számmal szorozzuk a vektort. Más szóval, a $h \mapsto hf'(a)$ hozzárendeléssel értelmezett $\mathbb{K} \to \mathbb{K}^N$ lineáris leképezés pontosan a $\mathrm{D}f(a)$ leképezést adja meg. Tulajdonképpen ez a szokásos módja a $\mathbb{K}^N \equiv \mathrm{Lin}(\mathbb{K}, \mathbb{K}^N)$ azonosításnak.

Tekintsünk most egy $F : \mathbb{K} \rightarrowtail W$ függvényt. Úgy tűnik, hogy ilyen függvény differenciálhatóságáról csak az "új" értelemben beszélhetünk, hiszen W tetszőleges normált tér lehet. Vegyük azonban észre, hogy a "régi", $\mathbb{K} \rightarrowtail \mathbb{K}^N$ esetre vonatkozó definíció ebben az esetben is elmondható és értelmes marad. Ezek szerint mind $\mathrm{D}F(a) \in \mathrm{Lin}(\mathbb{K}, W)$, mind

$$F'(a) := \lim_{h \to 0} \frac{F(a+h) - F(a)}{h} \in W$$

értelmezhető, ha a függvény differenciálható a-ban. A kettő között a kapcsolat hasonló, mint ez előbb, vagyis mindkettő ugyanakkor létezik, és bármilyen $h \in \mathbb{K}$ esetén

$$\mathrm{D}F(a)h = hF'(a).$$

Ez szintén a szokásos $W \equiv \mathrm{Lin}(\mathbb{K}, W)$ azonosításnak felel meg. Ilyen függvényeknél a továbbiakban mind a "két" deriváltat fogjuk használni, szem előtt tartva, hogy a fentiek alapján ezek lényegében megegyeznek.

1.6. Állítás *Ha $F : V \rightarrowtail W$ differenciálható az $a \in V$ pontban, akkor folytonos is a-ban.*

BIZONYÍTÁS Mivel egy kis ordó függvény egyben nagy ordó függvény is, és $\mathrm{D}F(a)$ folytonos leképezés, az

$$F(x) - F(a) = \mathrm{D}F(a)(x-a) + o(x-a)$$

kifejezés jobb oldalának határértéke nulla, miközben x tart a-hoz; ezért $\lim_{x\to a} F(x) = F(a)$, amit bizonyítani akartunk.

Megjegyzés Itt látjuk, milyen fontos a differenciálhatóság definíciójában, hogy a derivált *folytonos* lineáris leképezés legyen. Ha ezt nem követelnénk meg, a differenciálhatóságból nem következne a folytonosság.

1.7. Definíció *Egy $F : V \rightarrowtail W$ függvény*

(i) ***differenciálható egy $H \subset V$ halmazon****, ha F differenciálható H minden pontjában,*

(ii) ***differenciálható****, ha differenciálható az értelmezési tartományán,*

(iii) ***derivált függvénye (deriváltja)*** *az $x \mapsto \mathrm{D}F(x)$ hozzárendelési utasítással az $\{x \in \mathrm{Dom}(F) \mid F$ differenciálható x-ben$\}$ halmazon értelmezett $V \rightarrowtail \mathcal{L}in(V, W)$ függvény,*

(iv) ***folytonosan differenciálható*** *egy $H \subset V$ halmazon, ha $H \subset \mathrm{Dom}(\mathrm{D}F)$ és a $\mathrm{D}F$ deriváltfüggvény folytonos H-n,*

(v) ***folytonosan differenciálható****, ha folytonosan differenciálható az értelmezési tartományán.*

Megjegyzések (i) Nyilvánvaló, ha egy függvény differenciálható egy halmazon, akkor annak a halmaznak minden pontja szükségszerűen belső pontja az adott függvény értelmezési tartományának. Ha tehát a függvény differenciálható, akkor az értelmezési tartománya nyílt. (ii) A definíció szerint az üres $V \rightarrowtail W$ függvény differenciálható (ha nem az volna, lenne olyan pont az értelmezési tartományában, amelyben nem differenciálható), sőt folytonosan differenciálható.

(iii) A definíció értelmében minden $V \rightarrowtail W$ függvénynek van deriváltja, azonban az előfordulhat, hogy üres halmazon van értelmezve; pontosan akkor, ha az illető függvény sehol sem differenciálható.

(iv) Emlékeztetünk arra, hogy $\mathcal{L}in(V, W)$, a V normált térből a W normált térbe ható folytonos, lineáris leképezések összessége maga is normált tér. Tehát az F függvény $\mathrm{D}F : V \rightarrowtail \mathcal{L}in(V, W)$, $x \mapsto \mathrm{D}F(x)$ derivált függvénye szintén normált térből normált térbe ható függvény, így beszélhetünk annak folytonosságáról.

(v) Vigyázzunk, hogy ne tévesszük össze a $\mathrm{D}F(x)$ pontbeli deriváltat a $\mathrm{D}F$ derivált függvénnyel. Definíció szerint ugyanis $\mathrm{D}F(x) : V \to W$ mindig folytonos (és lineáris), de $\mathrm{D}F : V \rightarrowtail \mathcal{L}in(V, W)$, $x \mapsto \mathrm{D}F(x)$ nem feltétlenül az.

1.8. (i) Egy normált terek közötti konstans leképezés differenciálható az értelmezési tartománya belsejében, és a deriváltja minden pontban nulla (a nulla folytonos, lineáris leképezés).

(ii) Egy $A : V \to W$ folytonos, lineáris leképezés differenciálható, és $\mathrm{D}A(x) = A$ minden $x \in V$ pontban. Speciálisan az id_V deriváltja minden pontban maga az id_V függvény (folytonos lineáris leképezés).

(iii) Láttuk a könyvünk első részében, hogy a $\mathbb{C} \to \mathbb{R}$ abszolútérték-függvény seholsem differenciálható, az $\mathbb{R} \to \mathbb{R}$ abszolút-értékfüggvény pedig a nullában nem differenciálható. Most egységesen annyit tudunk állítani, hogy a $\|.\| : V \to \mathbb{R}$ norma a nullában nem differenciálható, ha V nem a nulla vektortér. Tegyük fel ugyanis, hogy van olyan $A : V \to \mathbb{R}$ folytonos lineáris leképezés, hogy

$$\|0 + h\| = Ah + o(h)$$

minden $h \in V$ estén. Vegyünk tetszőleges e vektort V-ből, amelyre $\|e\| = 1$ teljesül. Ekkor minden $0 \neq \alpha \in \mathbb{K}$ esetén

$$|\alpha| = \alpha Ae + o(\alpha e), \qquad 1 - \frac{o(\alpha e)}{|\alpha|} = \frac{\alpha}{|\alpha|} Ae.$$

A bal oldalnak van határértéke, miközben α tart a nullához, a jobb oldalnak nincs; ez az ellentmondás cáfolja a norma differenciálhatóságát a nullában.

(iv) Ha V valós vektortér és a norma skaláris szorzatból származik, akkor a skaláris szorzat – mint $V \times V \to \mathbb{R}$ függvény – differenciálható, és (a, b)-beli deriváltja az $V \times V \to \mathbb{R}$, $(h, k) \mapsto < h, b > + < a, k >$ folytonos lineáris leképezés. Következésképpen, ha most r jelöli a normát – azaz $r(x) = \sqrt{\langle x, x\rangle}$ –, akkor

$$\mathrm{D}(r^2)(a) = 2\langle a, .\rangle,$$

és

$$\mathrm{D}r(a) = \frac{\langle a, .\rangle}{r(a)} \qquad (a \neq 0).$$

(v) Legyen V teljes normált tér (Banach-tér). Tudjuk, hogy $\{X \in \mathcal{L}in(V) \mid X \text{ bijekció}, X^{-1} \in \mathcal{L}in(V)\}$ (az invertálható folytonos lineáris leképezések összessége) nyílt halmaz (Analízis III.B.10.10.). Az ezen a halmazon értelmezett $X \mapsto X^{-1}$ leképezés differenciálható, és deriváltja az A pontban a

$$\mathcal{L}in(V) \to \mathcal{L}in(V), \quad H \mapsto -A^{-1}HA^{-1}$$

folytonos lineáris leképezés.

Ugyanis a Neumann-formula szerint, ha $\|H\| < \dfrac{1}{\|A^{-1}\|}$, akkor $(A + H)^{-1}$ is invertálható, és $(A+H)^{-1} = \Big(A(\mathrm{id}_V + A^{-1}H)\Big)^{-1} = (\mathrm{id}_V + A^{-1}H)^{-1}A^{-1}$, ezért

$$(A + H)^{-1} - A^{-1} = \Big((\mathrm{id}_V + A^{-1}H)^{-1} - \mathrm{id}_V\Big)A^{-1} =$$

$$= -A^{-1}HA^{-1} + \sum_{n=2}^{\infty}(-1)^n(A^{-1}H)^nA^{-1}.$$

Könnyű látni, hogy az utolsó tag H-ban kis ordó függvény, és ezzel igazoltuk is, amit akartunk.

(vi) Legyen V teljes normált tér (Banach-tér) és $A \in \mathcal{L}in(V)$. Ekkor értelmes az $\mathbb{R} \to \mathcal{L}in(V)$, $t \mapsto e^{tA}$ leképezés (Analízis III.B.10.8.), amely differenciálható, és a deriváltja az $\mathbb{R} \to \mathcal{L}in(V)$, $t \mapsto Ae^{tA} = e^{tA}A$ leképezés.

Ugyanis

$$e^{(t+h)A} - e^{tA} = \left(e^{hA} - \mathrm{id}_V\right) e^{tA} = e^{tA} \left(e^{hA} - \mathrm{id}_V\right),$$

és

$$e^{hA} - \mathrm{id}_V = hA + \sum_{n=2}^{\infty} \frac{(hA)^n}{n!},$$

tehát

$$\lim_{h \to 0} \frac{e^{hA} - \mathrm{id}_V}{h} = A.$$

1.9. Most következzen egy állítás, amelynek bizonyítása a definíciókra támaszkodva teljesen hasonló módon történik, mint ahogyan azt a $\mathbb{K} \rightarrowtail \mathbb{K}^N$ esetben csináltuk, ezért felkérjük az Olvasót, hogy azt önállóan végezze el.

Állítás *Legyenek az $F, G : V \rightarrowtail W$ és $\phi : V \rightarrowtail \mathbb{K}$ függvények differenciálhatók az $a \in V$ pontban. Ekkor*

(i) minden $\lambda \in \mathbb{K}$ esetén λF is differenciálható a-ban és

$$\mathrm{D}(\lambda F)(a) = \lambda \mathrm{D}F(a),$$

(ii) $F + G$ is differenciálható a-ban és

$$\mathrm{D}(F + G)(a) = \mathrm{D}F(a) + \mathrm{D}G(a),$$

(iii) ϕF is differenciálható a-ban és

$$\mathrm{D}(\phi F)(a) = F(a) \otimes \mathrm{D}\phi(a) + \phi(a)\mathrm{D}F(a),$$

(iv) ha $\phi(a) \neq 0$, akkor $\dfrac{F}{\phi}$ is differenciálható a-ban és

$$\mathrm{D}\left(\frac{F}{\phi}\right)(a) = \frac{\mathrm{D}F(a)}{\phi(a)} - \frac{F(a) \otimes \mathrm{D}\phi(a)}{\phi^2(a)},$$

ahol $F(a) \otimes \mathrm{D}\phi(a) : V \to W$, $h \mapsto (\mathrm{D}\phi(a)h)F(a)$.

1.10. Állítás *Legyen* $F : V \rightarrowtail W$*,* $G : W \rightarrowtail U$ *és* $a \in \overset{\circ}{\mathrm{Dom}}(G \circ F)$*. Ha* F *differenciálható az* a *pontban,* G *pedig differenciálható az* $F(a)$ *pontban, akkor* $G \circ F$ *differenciálható* a*-ban, és*

$$\mathrm{D}(G \circ F)(a) = \mathrm{D}G(F(a)) \circ \mathrm{D}F(a).$$

Bizonyítás Ha $x \in \mathrm{Dom}(G \circ F) \setminus \{a\}$, akkor

$$\begin{aligned}(G \circ F)(x) - (G \circ F)(a) &= G(F(x)) - G(F(a)) = \\ &= \mathrm{D}G(F(a))(F(x) - F(a)) + o_G(F(x) - F(a)) = \\ &= \mathrm{D}G(F(a))\Big(\mathrm{D}F(a)(x-a) + o_F(x-a)\Big) + o_G(F(x) - F(a)) = \\ &= \Big(\mathrm{D}G(F(a)) \circ \mathrm{D}F(a)\Big)(x-a) + \mathrm{D}G(F(a))o_F(x-a) + o_G(F(x) - F(a)), \qquad (*)\end{aligned}$$

ahol $o_G : W \rightarrowtail U$ és $o_F : V \rightarrowtail W$ kis ordó függvények.

Ha a második és a harmadik tagot egy-egy $(x-a)$-ban kis ordó függvénnyel tudnánk helyettesíteni, akkor készen is volnánk. Vizsgáljuk meg őket:

$$\lim_{x \to a} \frac{\mathrm{D}G(F(a))o_F(x-a)}{\|x-a\|} = \lim_{x \to a} \left(\mathrm{D}G(F(a))\frac{o_F(x-a)}{\|x-a\|}\right) = 0, \qquad (1)$$

mivel $\lim\limits_{x \to a} \frac{o_F(x-a)}{\|x-a\|} = 0$ és $\mathrm{D}G(F(a))$ folytonos.

Amint azt az 1.1. pontban láttuk, az $o_G : W \rightarrowtail U$ kis ordó függvény előáll $o_G = \| \,.\, \|\eta$ alakban, ahol $\eta : W \rightarrowtail U$, $\mathrm{Dom}(\eta) = \mathrm{Dom}(o_G)$, és $\lim\limits_{h \to 0} \eta(h) = 0$. Emiatt $x \neq a$ esetén

$$\begin{aligned}\|o_G(F(x) - F(a))\| &= \|F(x) - F(a)\| \; \|\eta(F(x) - F(a))\| = \\ &= \|\mathrm{D}F(a)(x-a) + o_F(x-a)\| \; \|\eta(F(x) - F(a))\| \leq \\ &\leq \|x-a\| \left(\|\mathrm{D}F(a)\| + \frac{\|o_F(x-a)\|}{\|x-a\|}\right) \|\eta(F(x) - F(a))\|.\end{aligned}$$

A jobb oldal középső tényezője korlátos, hiszen $\lim\limits_{x \to a} \dfrac{\|o_F(x-a)\|}{\|x-a\|} = 0$, a harmadik tényezője pedig nullához tart, miközben x tart a-hoz, hiszen η határértéke nulla a nullában, és F folytonos az a pontban, ezért $\lim\limits_{x \to a}(F(x) - F(a)) = 0$.

A fentiek alapján

$$\lim_{x \to a} \frac{o_G(F(x) - F(a))}{\|x-a\|} = 0. \qquad (2)$$

Az (1) és (2) egyenlőségek a $(*)$ összefüggéssel együtt pontosan azt bizonyítják, hogy $G \circ F$ valóban differenciálható az a pontban, és a deriváltjára érvényes az állított egyenlőség. ■

Eredményünk közvetlen következménye, hogy ha $F : V \rightarrowtail W$ injektív, differenciálható a-ban, és az inverze differenciálható $F(a)$-ban, akkor

$$\mathrm{D}F^{-1}(F(a)) \circ \mathrm{D}F(a) = \mathrm{id}_V = \mathrm{D}F(a) \circ \mathrm{D}F^{-1}(F(a)),$$

azaz

$$\mathrm{D}F^{-1}(F(a)) = (\mathrm{D}F(a))^{-1}.$$

Ez jóval kevesebbet mond, mint az A.1.8. állítás, ahol a függvény inverzének a differenciálhatóságánál kevesebbet kellett feltennünk. Többet, bővebbet majd az úgynevezett inverzfüggvény-tétel ad.

1.11. Feladatok

1. Legyen $u : [0,1] \to \mathbb{R}$ adott integrálható függvény. Differenciálható-e a

$$C[0,1] \to \mathbb{R}, \quad f \mapsto \int_0^1 uf$$

leképezés, ha $C[0,1]$-en a $\|.\|_\infty$, az $\|.\|_1$ illetve a $\|.\|_2$ normát vesszük?

2. Differenciálható-e a

$$C[0,1] \to \mathbb{R}, \quad f \mapsto \int_0^1 f^2$$

leképezés az előbbi feladatban szereplő három norma valamelyikében?

3. Differenciálható-e valamely szorzat normára vonatkozóan a

$$C[0,1] \times C[0,1] \to C[0,1], \quad (f,g) \mapsto fg$$

leképezés?

4. Legyen $L : V \rightarrowtail \mathcal{L}in(U,W)$ differenciálható az a pontban. Igazoljuk, hogy minden $u \in U$ esetén $L_u : V \rightarrowtail W$, $x \mapsto L(x)u$ differenciálható a-ban és $\mathrm{D}L_u(a)v = \big(\mathrm{D}L(a)v\big)u$ minden $v \in V$ esetén..

5. Legyen $L : V \rightarrowtail \mathcal{L}in(U,W)$ és $g : V \rightarrowtail U$ differenciálható az a pontban. Bizonyítsuk be, hogy $V \rightarrowtail W$, $x \mapsto L(x)g(x)$ differenciálható a-ban, és adjuk meg az a-beli deriváltját.

6. Mutassuk meg, hogy az $\mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ vektoriális szorzás differenciálható, és adjuk meg a deriváltját.

7. Mutassuk meg, hogy

$$\mathcal{L}in(V) \times V \to V, \quad (A,x) \mapsto Ax,$$

$$\mathcal{L}in(V) \times \mathcal{L}in(V) \to \mathcal{L}in(V), (A, B) \mapsto A \circ B$$

differenciálható leképezések. Adjuk meg deriváltjaikat!

8. Igazoljuk, hogy egy $B : U \times V \to W$ folytonos bilineáris leképezés mindenütt differenciálható, és deriváltja az (a, b) pontban az

$$U \times V \to W, \quad (h, k) \mapsto B(h, b) + B(a, k)$$

folytonos lineáris leképezés.

9. Legyen V és W valós normált tér, és származzon a norma W-n skalárszorzatból. Ha $F, G : V \rightarrowtail W$ differenciálható függvények, akkor a pontonként értelmezett $< F, G >: V \rightarrowtail \mathbb{R}$ függvény differenciálható és a-beli deriváltja a

$$V \rightarrowtail \mathbb{R}, \quad h \mapsto < \mathrm{D}F(a)h, G(a) > + < F(a), \mathrm{D}G(a)h >$$

folytonos lineáris leképezés.

10. Mutassuk meg, hogy komplex skaláris szorzatos téren a skaláris szorzat sehol sem differenciálható.

11. Hol differenciálható a $\mathbb{K}^N$-en szokásosan megadott "egyes", "kettes" és "végtelen" norma? Vizsgáljuk külön a valós és a komplex esetet.

12. Legyen V véges dimenziós vektortér. Mutassuk meg az Analízis II.21.3 és 25.4. formulái alapján, hogy a $\det : \mathrm{Lin}(V) \to \mathbb{K}$ leképezés differenciálható az injektív (tehát invertálható) lineáris leképezések nyílt halmazán, és A-beli deriváltja a

$$\mathrm{Lin}(V) \to \mathbb{K}, \quad H \mapsto (\det A)\mathrm{Tr}(HA^{-1})$$

lineáris leképezés.

13. Emlékeztetünk arra, hogy $A \in \mathcal{L}in(V)$ esetén az A-val való balról illetve jobbról szorzás a

$$L_A : \mathcal{L}in(V) \to \mathcal{L}in(V), \quad X \mapsto AX, \qquad R_A : \mathcal{L}in(V) \to \mathcal{L}in(V), \quad X \mapsto XA$$

folytonos lineáris leképezés.

Értelmezzük $n \in \mathbb{N}_0$ esetén a

$$J_n(A) := \sum_{k=1}^{n} L_A^k R_A^{n-k} : \mathcal{L}in(V) \to \mathcal{L}in(V)$$

folytonos lineáris leképezést.

Mutassuk meg, hogy $n \in \mathbb{N}$ esetén a $\mathcal{L}in(V) \to \mathcal{L}in(V)$, $X \mapsto X^n$ függvény differenciálható, és deriváltja A-ban $J_{n-1}(A)$. Mi ez $A := \mathrm{id}_V$ illetve $A := 0$ esetén?

14. Legyen V teljes normált tér (Banach-tér). Az 1.8. (v), a kompizíció-függvények deriváltjára vonatkozó formula és az előző feladat eredménye alapján mutassuk meg, hogy $X \mapsto X^{-n}$ differenciálható, és deriváltja A-ban $-\sum_{k=0}^{n-1} L_{A^{-1}}^{k+1} R_{A^{-1}}^{n-k}$.

15. Legyen V teljes normált tér (Banach-tér). Mutassuk meg, hogy a $\mathcal{L}in(V) \to \mathcal{L}in(V)$, $X \mapsto e^X$ differenciálható, és A-beli deriváltja a $\displaystyle\sum_{n=0}^{\infty} \frac{J_n(A)}{n!} : \mathcal{L}in(V) \to \mathcal{L}in(V)$ folytonos lineáris leképezés. Mi ez $A := \mathrm{id}_V$ illetve $A := 0$ esetén?

16. Legyen $F : V \rightarrowtail W$ és L lineáris altere V-nek. Ekkor L szintén normált tér, és $F|_L : L \rightarrowtail W$ normált terek közötti leképezés, tehát értelmes a differenciálhatóságáról beszélni. Ha $a \in L$ és $a \in \overset{\circ}{\mathrm{Dom}}(F)$, akkor $a \in \overset{\circ}{\mathrm{Dom}}(F|_L)$; igazoljuk, hogy ha F differenciálható a-ban, akkor $F|_L$ is differenciálható a-ban, és $\mathrm{D}(F|_L)(a) = (\mathrm{D}F(a))|_L$.

2. Középértéktételek

2.1. Állítás *(**Középérték-tétel, valós**) Legyen V valós normált tér, $f : V \rightarrowtail \mathbb{R}$ függvény, és $x, y \in V$. Ha f differenciálható az $[x, y] \subset V$ szakaszon, akkor létezik olyan $z \in]x, y[$, hogy*

$$f(y) - f(x) = \mathrm{D}f(z)(y - x).$$

Bizonyítás Definiáljuk az $\alpha : [0, 1] \to V$ leképezést a $t \mapsto x + t(y - x) =: \alpha(t)$ hozzárendeléssel. Nyilvánvaló, hogy α folytonos függvény, és differenciálható $]0, 1[$-en, valamint $\alpha'(t) = y - x$ minden $t \in]0, 1[$ esetén. Következésképpen az $f \circ \alpha : [0, 1] \to \mathbb{R}$ függvény is folytonos, és differenciálható a $]0, 1]$ intervallumon. Az $f \circ \alpha$ függvényre alkalmazhatjuk a Lagrange-féle középértéktételt, miszerint létezik olyan $s \in]0, 1[$, hogy

$$f(y) - f(x) = (f \circ \alpha)(1) - (f \circ \alpha)(0) = (f \circ \alpha)'(s)(1 - 0) = (f \circ \alpha)'(s).$$

Vezessük be a $z := \alpha(s)$ jelölést. Figyelembe véve az $(f \circ \alpha)'$ és $\mathrm{D}(f \circ \alpha)$ illetve az α' és $\mathrm{D}\alpha$ közötti azonosításokat, valamint a kompozíciófüggvény deriválására vonatkozó tételt azt kapjuk, hogy

$$\begin{aligned} f(y) - f(x) &= \mathrm{D}(f \circ \alpha)(s)1 = \big(\mathrm{D}f(\alpha(s)) \circ \mathrm{D}\alpha(s)\big)1 = \\ &= \mathrm{D}f(z)\big(\mathrm{D}\alpha(s)1\big) = \mathrm{D}f(z)\alpha'(s) = \\ &= \mathrm{D}f(z)(y - x), \end{aligned}$$

és pontosan ezt akartuk bizonyítani.

2.2. Állítás *(**Középérték-tétel, általános**) Legyen $F : V \rightarrowtail W$ függvény és $x, y \in V$. Ha F differenciálható az $[x, y] \subset V$ szakaszon, akkor létezik olyan $z \in]x, y[$, hogy*

$$\|F(y) - F(x)\| \leq \|\mathrm{D}F(z)\| \ \|y - x\|.$$

Bizonyítás Definiáljuk ismét az $\alpha : [0, 1] \to V$ leképezést a $t \mapsto x+t(y-x) =: \alpha(t)$ hozzárendeléssel, így α rendelkezik mindazokkal a tulajdonságokkal, mint az előző állításban.

Előlegezzük meg (későbbi tanulmányaink során látni fogjuk, a Hahn–Banach-tételből következik, lásd Analízis VIII.7.5.), hogy van olyan $p : W \to \mathbb{K}$ folytonos, lineáris leképezés, hogy

$$|p\big(F(y) - F(x)\big)| = \|F(y) - F(x)\| \quad \text{és} \quad \|p\| = 1.$$

A $p \circ F \circ \alpha : [0, 1] \to \mathbb{K}$ függvény folytonos, és differenciálható a $]0, 1[$ intervallumon, így alkalmazhatjuk rá az A.3.4. általános középértéktételt, miszerint létezik olyan $s \in]0, 1[$, hogy

$$|(p \circ F \circ \alpha)(1) - (p \circ F \circ \alpha)(0)| \leq |(p \circ F \circ \alpha)'(s)| \ |1 - 0|.$$

Ezt átalakítva, a deriváltak közötti azonosítások figyelembevételével, a

$$|(p \circ F)(y) - (p \circ F)(x)| \leq |\mathrm{D}p((F \circ \alpha)(s))\mathrm{D}F(\alpha(s))\mathrm{D}\alpha(s)|$$

egyenlőtlenséghez jutunk.

Mivel p folytonos lineáris leképezés, minden pontban a deriváltja önmaga, azaz $\mathrm{D}p[(F \circ \alpha)(s)] = p$, valamint α konkrét alakjából $\mathrm{D}\alpha(s) = y - x$.

Ezek alapján, bevezetve a $z := \alpha(s)$ jelölést arra jutunk, hogy

$$\begin{aligned} \|F(y) - F(x)\| = |p\big(F(y) - F(x)\big)| \leq |p\mathrm{D}F(z)(y - x)| \leq \\ \leq \|p\| \ \|\mathrm{D}F(z)\| \ \|y - x\|, \end{aligned}$$

amit bizonyítani akartunk, hiszen $\|p\| = 1$.

2.3. Eredményünk egyszerű, de fontos következménye:

Állítás *Ha $F : V \rightarrowtail W$ differenciálható, az értelmezési tartománya csillagszerű, és $\mathrm{D}F = 0$, akkor F konstans.*

Egy kicsit általánosabban: ha F differenciálható, $\mathrm{D}F = 0$, és $\mathrm{Dom}(F)$ bármely két pontja összeköthető a $\mathrm{Dom}(F)$-ben levő törött vonallal, akkor F konstans.

3. Függvények együttesének és Descartes-szorzatának differenciálhatósága

3.1. Legyenek V_i $(i = 1, ..., N)$ normált terek. Emlékeztetünk arra, hogy Descartes-szorzatukon, a $\mathop{\times}\limits_{i=1}^{N} V_i$ vektortéren háromféleképpen vezettünk be normát; az "egyes", a "kettes" és a "végtelen" normát:

$$\|(x_1, \ldots, x_N)\|_1 := \sum_{i=1}^{N} \|x_i\|,$$

$$\|(x_1, \ldots, x_N)\|_2 := \sqrt{\sum_{i=1}^{N} \|x_i\|^2},$$

$$\|(x_1, \ldots, x_N)\|_\infty := \max\{\|x_i\| \mid i = 1, \ldots, N\}.$$

E három norma ekvivalens egymással. Amikor tehát normált terek Descartes-szorzatáról beszélünk, úgy tekintjük, hogy az is normált tér, e normák valamelyikével ellátva. Gyakran az "egyes" és a "végtelen" normák kezelhetők jól, viszont szinte mindig a "kettes" normát használják, ha minden V_i-n a norma skalárszorzatból származtatható (gondoljuk meg, miért!).

3.2. Állítás *Az $F_i : V \rightarrowtail W_i$ $(i = 1, \ldots, N)$ függvények akkor és csak akkor differenciálhatók az $a \in V$ pontban, ha az $(F_1, \ldots, F_N) : V \rightarrowtail \mathop{\times}\limits_{i=1}^{N} W_i$ együttes függvény differenciálható a-ban.*

A deriváltakra fennáll a

$$\big(\mathrm{D}(F_1, \ldots, F_N)\big)(a) = \big(\mathrm{D}F_1(a), \ldots, \mathrm{D}F_N(a)\big)$$

összefüggés.

Bizonyítás Vezessük be az $F := (F_1, \ldots, F_N)$ jelölést.

Nyilvánvaló, hogy a pontosan akkor belső pontja az összes F_i értelmezési tartományának, ha belső pontja ezek metszetének is, vagyis az F értelmezési tartományának.

(i) Legyenek az $F_i : V \rightarrowtail W_i$ $(i = 1, ..., N)$ függvények differenciálhatók az a pontban.

Ha $h \in V$ és $a + h \in \mathrm{Dom}(F)$, akkor

$$\begin{aligned} F(a+h) - F(a) &= \big(F_1(a+h) - F_1(a), \ldots, F_N(a+h) - F_N(a)\big) = \\ &= \big(\mathrm{D}F_1(a)h + o_1(h), \ldots, \mathrm{D}F_N(a)h + o_N(h)\big) = \\ &= \big(\mathrm{D}F_1(a)h, \ldots, \mathrm{D}F_N(a)h\big) + \big(o_1(h), \ldots, o_N(h)\big) = \\ &= \big(\mathrm{D}F_1(a), \ldots, \mathrm{D}F_N(a)\big)h + \big(o_1, \ldots, o_N\big)(h). \end{aligned}$$

Az utolsó sor első tagjában $V \to \mathop{\times}\limits_{i=1}^{N} W_i$ folytonos lineáris leképezés szerepel, a második tag pedig kis ordó függvény, hiszen

$$\lim_{h \to 0} \frac{\big(o_1, \ldots, o_N\big)(h)}{\|h\|} = \lim_{h \to 0} \left(\frac{o_1(h)}{\|h\|}, \ldots, \frac{o_N(h)}{\|h\|} \right) = 0,$$

mivel mindegyik komponens nullához tart.

Ezek alapján F differenciálható az a pontban, és deriváltjára fennáll az állított egyenlőség.

(ii) Most azt tegyük fel, hogy $F : V \rightarrowtail \mathop{\times}\limits_{i=1}^{N} W_i$ differenciálható az a pontban. Mivel a pr_i kanonikus projekció folytonos lineáris leképezés, tehát differenciálható, és a deriváltja mindenütt önmaga, a kompozíciófüggvények deriváltjára vonatkozó ismereteink alapján $F_i = \mathrm{pr}_i \circ F$ is differenciálható a-ban, és $\mathrm{D}F_i(a) = \mathrm{pr}_i \circ \mathrm{D}F(a)$. Ezzel a bizonyítást befejeztük.

3.3. Állítás *Az $F_i : V_i \rightarrowtail W_i$ függvények akkor és csak akkor differenciálhatók az $a_i \in V_i$ $(i = 1, \ldots, N)$ pontokban, ha az $\mathop{\times}\limits_{i=1}^{N} F_i : \mathop{\times}\limits_{i=1}^{N} V_i \rightarrowtail \mathop{\times}\limits_{i=1}^{N} W_i$ Descartes-szorzatfüggvény differenciálható az $a := (a_1, \ldots, a_N)$ pontban.*

A deriváltakra fennáll a

$$\left(\mathrm{D} \left(\mathop{\times}_{i=1}^{N} F_i \right) \right)(a) = \mathop{\times}_{i=1}^{N} \mathrm{D}F_i(a_i)$$

összefüggés.

Bizonyítás Vezessük be az $F := \mathop{\times}\limits_{i=1}^{N} F_i$ jelölést.

Nyilvánvaló, hogy a pontosan akkor belső pontja F értelmezési tartományának, ha minden i-re a_i belső pontja F_i értelmezési tartományának.

(i) Legyenek az $F_i : V_i \rightarrowtail W_i$ függvények differenciálhatók az a_i pontokban.

Ha $h := (h_1, \ldots, h_N) \in \bigtimes_{i=1}^{N} V_i$ és $a + h \in \mathrm{Dom}(F)$, akkor

$$\begin{aligned}
F(a+h) - F(a) &= \big(F_1(a_1+h_1) - F_1(a_1), \ldots, F_N(a_N+h_N) - F_N(a_N)\big) = \\
&= \big(\mathrm{D}F_1(a_1)h_1 + o_1(h_1), \ldots, \mathrm{D}F_N(a_N)h_N + o_N(h_N)\big) = \\
&= \left(\bigtimes_{i=1}^{N} \mathrm{D}F_i(a_i)\right) h + \left(\bigtimes_{i=1}^{N} o_i\right)(h).
\end{aligned}$$

Az utolsó sor első tagjában $\bigtimes_{i=1}^{N} V_i \to \bigtimes_{i=1}^{N} W_i$ folytonos lineáris leképezés szerepel, a második tag pedig kis ordó függvény, hiszen

$$\frac{\left\| \left(\bigtimes_{i=1}^{N} o_i\right)(h) \right\|_1}{\|h\|_1} = \frac{\sum_{i=1}^{N} \|o_i(h_i)\|}{\sum_{i=1}^{N} \|h_i\|} \leq \sum_{i=1}^{N} \frac{\|o_i(h_i)\|}{\|h_i\|},$$

és a jobb oldal a nullához tart, miközben h – és ezzel együtt minden h_i – a nullához tart.

Ezek alapján F differenciálható a-ban, és a deriváltjára fennáll az állított egyenlőség.

(ii) Most azt tegyük fel, hogy F differenciálható az a pontban. Az egyszerűség kedvéért tekintsük az $i = 1$ indexet, teljesen hasonlóan érvelhetünk bármelyik másikra. Legyen $h_1 \in V_1$ olyan, hogy $a_1 + h_1 \in \mathrm{Dom}(F_1)$, és $h := (h_1, 0, \ldots, 0)$. Ekkor

$$\begin{aligned}
(F_1(a_1+h_1) - F_1(a_1), 0, \ldots, 0) = F(a+h) - F(a) = \mathrm{D}F(a)h + o(h) = \\
= \mathrm{D}F(a)(h_1, 0, \ldots, 0) + o(h_1, 0, \ldots, 0).
\end{aligned}$$

$\mathrm{D}F(a)$ folytonos és lineáris $\bigtimes_{i=1}^{N} V_i$-n, ezért a $V_1 \times \{0\} \times \ldots \times \{0\}$ altéren is folytonos és lineáris. A második tag viszont $V_1 \rightarrowtail W_1$ kis ordó függvény, hiszen

$$\lim_{h_1 \to 0} \frac{o(h_1, 0, \ldots, 0)}{\|h_1\|} = \lim_{h \to 0} \frac{o(h)}{\|h\|_1} = 0.$$

Tehát F_1 differenciálható az a_1 pontban és $\mathrm{D}F_1(a_1)h_1 = \mathrm{D}F(a)(h_1, 0, \ldots, 0)$, amivel a bizonyítást befejeztük.

4. Iránymenti és parciális differenciálhatóság

4.1. Definíció *Az $F : V \rightarrowtail W$ függvényt* ***differenciálhatónak*** *mondjuk az $a \in \overset{\circ}{\mathrm{Dom}}(F)$ pontban a $0 \neq v \in V$* ***irány mentén****, ha a $\mathbb{K} \rightarrowtail W$, $t \mapsto F(a+tv)$ függvény differenciálható a $0 \in \mathbb{K}$ pontban. Az F függvénynek a v iránymenti deriváltját az a pontban $\mathrm{D}_v F(a)$-val jelöljük:*

$$\mathrm{D}_v F(a) := \lim_{t\to 0} \frac{F(a+tv) - F(a)}{t} \in W.$$

4.2. Állítás *Ha az $F : V \rightarrowtail W$ függvény differenciálható az a pontban, akkor ebben a pontban minden $0 \neq v$ irány mentén is differenciálható, és*

$$\mathrm{D}_v F(a) = \mathrm{D}F(a)v.$$

Bizonyítás Válasszunk tetszőlegesen egy $0 \neq v$ vektort V-ből. Mivel a belső pontja az F értelmezési tartományának, létezik olyan $G(0)$ környezete a $0 \in \mathbb{K}$ pontnak, hogy $a + tv \in \mathrm{Dom}(F)$ minden $t \in G(0)$ esetén. Az összes ilyen t-re

$$F(a+tv) - F(a) = \mathrm{D}F(a)(tv) + o(tv) = t\mathrm{D}F(a)v + o(tv),$$

ahol $o : V \rightarrowtail W$ kis ordó függvény.

Az iménti egyenlőségből már könnyen látható, hogy létezik

$$\lim_{t\to 0} \frac{F(a+tv) - F(a)}{t} = \mathrm{D}F(a)v + \lim_{t\to 0} \frac{o(tv)}{t} = \mathrm{D}F(a)v$$

hiszen $\dfrac{o(tv)}{t} = \dfrac{o(tv)}{\|tv\|}\|v\|$.

Ez pontosan az, amit bizonyítani akartunk.

Megjegyzés Abból, hogy egy függvény egy pontban minden irányban differenciálható, nem következik, hogy differenciálható is ebben a pontban, sőt még az sem, hogy ott folytonos. Íme erre egy példa: az

$$\mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} 1 & \text{ha } 0<y<x^2, \\ 0 & \text{egyébként} \end{cases}$$

függvény nem folytonos a nullában, azonban ott minden irányú deriváltja létezik és nulla.

4.3. Definíció *Legyen $k = 1, \dots, N$. Egy $F : \bigtimes_{i=1}^{N} V_i \rightarrowtail W$ függvény*

(i) ***parciálisan differenciálható*** *az $(a_1, \dots, a_N) \in \bigtimes_{i=1}^{N} V_i$ pontban a k-adik változója szerint, ha a $V_k \rightarrowtail W$, $x_k \mapsto F(a_1, \dots, a_{k-1}, x_k, a_{k+1}, \dots, a_N)$ függvény differenciálható az a_k pontban; ennek deriváltja az a_k pontban az F függvénynek a k-**adik változó szerinti parciális deriváltja** az $(a_1, \dots, a_N)$ pontban, amelyet $\mathrm{D}_k F(a_1, \dots, a_N)$-nel jelölünk;*

(ii) ***parciálisan differenciálható a k-adik változója szerint egy halmazon***, *ha a halmaz minden pontjában parciálisan differenciálható e szerint a változója szerint;*

(iii) ***parciálisan differenciálható a k-adik változója szerint*** *ha parciálisan differenciálható e szerint a változója szerint az értelmezési tartományán;*

*(iv) k-**adik parciális derivált függvénye (deriváltja)** a $\mathrm{D}_k F : \bigtimes_{i=1}^{N} V_i \rightarrowtail \mathcal{L}in(V_k, W)$, $(x_1, \dots, x_N) \mapsto \mathrm{D}_k F(x_1, \dots, x_N)$ hozzárendeléssel megadott függvény, amelynek értelmezési tartománya mindazon pontok összessége, ahol F a k-adik változója szerint parciálisan differenciálható;*

(v) ***folytonosan parciálisan differenciálható a k-adik változója szerint egy halmazon***, *ha a k-ik parciális deriváltja értelmezett és folytonos ezen a halmazon;*

(vi) ***folytonosan parciálisan differenciálható a k-adik változója szerint***, *ha ilyen az értelmezési tartományán.*

Értelmezésünk szerint tehát $\mathrm{D}_k F(a_1, \dots, a_N) \in \mathcal{L}in(V_k, W)$.

A továbbiakban tetszőleges véges számú normált tér $\bigtimes_{i-1}^{N} V_i$ Descartes-szorzata helyett, az egyszerűbb jelölés és a könnyebb érthetőség érdekében, csak két normált tér $V_1 \times V_2$ Descartes-szorzatát fogjuk tekinteni. Természetesen azok az eredmények melyekhez így jutunk, minden különösebb nehézség nélkül átfogalmazhatók akárhány véges számú normált tér Descartes-szorzatára.

4.4. Állítás *Ha az $F : V_1 \times V_2 \rightarrowtail W$ függvény differenciálható az $(a_1, a_2) \in V_1 \times V_2$ pontban, akkor mindkét változójában parciálisan is differenciálható ebben a pontban, és*

$$\mathrm{D}_1 F(a_1, a_2) = \mathrm{D}F(a_1, a_2)|_{V_1 \times \{0\}}, \tag{1}$$

$$\mathrm{D}_2 F(a_1, a_2) = \mathrm{D}F(a_1, a_2)|_{\{0\} \times V_2}. \tag{2}$$

Bizonyítás Szorítkozzunk az első egyenlőségre, hiszen nyilvánvaló, hogy a második teljesen hasonlóan látható be.

Tekintsük az $F(\cdot, a_2) : V_1 \rightarrowtail W$ függvényt. Az a_1 pont belső pontja e függvény értelmezési tartományának, hiszen (a_1, a_2) belső pontja az F értelmezési tartományának, mivel F ott differenciálható. Legyen $h_1 \in V_1$ olyan, hogy $a_1 + h_1 \in \mathrm{Dom} F(\cdot, a_2)$; ekkor

$$F(a_1 + h_1, a_2) - F(a_1, a_2) = \mathrm{D}F(a_1, a_2)(h_1, 0) + o(h_1, 0),$$

ahol $o : V_1 \times V_2 \rightarrowtail W$ kis ordó függvény.

Mivel $\|h_1\| = \|h_1\| + \|0\| = \|(h_1, 0)\|_1$, ahol az 1-es index a szoksásos "egyes" szorzatnormára utal,

$$\lim_{h_1 \to 0} \frac{o(h_1, 0)}{\|h_1\|} = \lim_{(h_1,0) \to (0,0)} \frac{o(h_1, 0)}{\|(h_1, 0)\|} = 0,$$

vagyis $o(\cdot, 0) : V_1 \rightarrowtail W$ szintén kis ordó függvény.

Tehát $F(\cdot, a_2)$ differenciálható a_1-ben és a deriváltja a $h_1 \mapsto \mathrm{D}F(a_1, a_2)(h_1, 0)$ folytonos lineáris leképezés. Ezzel a bizonyítást elvégeztük.

Megjegyzés Felhívjuk a figyelmet arra, hogy az iménti állítás, megfordítva már nem igaz, vagyis az F összes parciális deriváltjának a létezéséből még nem feltétlenül következik F differenciálhatósága. Ellenpéldának az előző pont végén szereplő függvényt hozhatjuk föl, ha $\mathbb{R}^2$-t $\mathbb{R} \times \mathbb{R}$-nek írjuk. A parciális deriváltak lényegében a "tengelyek", vagyis a standard bázisvektorok iránya mentén vett deriváltakkal egyeznek meg.

4.5. Az Olvasó bizonyára észrevette, hogy egy kis "csúsztatást" követtünk el, amikor a 4.4-beli (1) és (2) egyenlőségeket felírtuk, ugyanis $\mathrm{D}_1 F(a_1, a_2) : V_1 \to W$ lineáris leképezés, viszont $\mathrm{D}F(a_1, a_2)|_{V_1 \times \{0\}} : V_1 \times \{0\} \to W$ lineáris leképezés; és hasonló a helyzet a második változóban is.

Amit tettünk, azt a szokásos

$$\mathcal{L}in(V_1, W) \times \mathcal{L}in(V_2, W) \equiv \mathcal{L}in(V_1 \times V_2, W),$$

$$(A_1 \ A_2) \equiv \big((h_1, h_2) \mapsto A_1 h_1 + A_2 h_2\big)$$

azonosítás (Analízis III. B.10.12.) indokolja. Itt a szokásnak megfelelően A_1 és A_2 közé nem tettünk vesszőt, hogy a jelölés ne essen egybe a függvények együttesére már használt jelöléssel. Egyébként $(A_1 \ A_2)$ a lineáris algebrából ismert sormátrixnak felel meg, és a fenti azonosítást

$$(A_1 \ A_2) \begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = A_1 h_1 + A_2 h_2$$

alakban is írhatnánk.

Ezzel a jelöléssel a 4.4-beli (1) és (2) egyenlőségeket az áttekinthetőbb

$$\mathrm{D}F(a_1,a_2) = \big(\mathrm{D}_1F(a_1,a_2)\ \ \mathrm{D}_2F(a_1,a_2)\big)$$

formára hozhatjuk.

Azt is tudjuk, hogy akármelyik szorzatnormát vesszük is $V_1 \times V_2$-n,

$$\|(A_1\ A_2)\| \le \|A_1\| + \|A_2\|,$$

$$\|A_1\| \le \|(A_1\ A_2)\|, \qquad \|A_2\| \le \|(A_1\ A_2)\|.$$

4.6. Az előzőekben láttuk, hogy a parciális deriváltak létezése csak szükséges, de nem elégséges feltétele a differenciálhatóságnak. Folytonos differenciálhatóság esetén ennél többet mondhatunk.

Állítás *Az $F : V_1 \times V_2 \rightarrowtail W$ függvény akkor és csak akkor differenciálható folytonosan, ha folytonosan parciálisan differenciálható mindkét változója szerint.*

Bizonyítás (i) Tegyük fel, hogy F folytonosan differenciálható. Ekkor parciálisan is differenciálható, és a parciális derivált függvények folytonossága az előző pont végén idézett formulák következménye, mivel bármely $x, y \in \mathrm{Dom}(F)$ esetén

$$\begin{aligned}\|\mathrm{D}_1F(y) - \mathrm{D}_1F(x)\| &\le \big\|\big(\mathrm{D}_1F(y) - \mathrm{D}_1F(x)\ \ \mathrm{D}_2F(y) - \mathrm{D}_2F(x)\big)\big\| = \\ &= \big\|\big(\mathrm{D}_1F(y)\ \mathrm{D}_2F(y)\big) - \big(\mathrm{D}_1F(x)\ \mathrm{D}_2F(x)\big)\big\| = \\ &= \|\mathrm{D}F(y) - \mathrm{D}F(x)\|,\end{aligned}$$

és hasonló egyenlőtlenség igaz a második parciális deriváltra is.

(ii) Ha a parciális deriváltak folytonosak és F differenciálható, akkor $\mathrm{D}F$ is folytonos, mert

$$\|\mathrm{D}F(y) - \mathrm{D}F(x)\| \le \|\mathrm{D}_1F(y) - \mathrm{D}_1F(x)\| + \|\mathrm{D}_2F(y) - \mathrm{D}_2F(x)\|.$$

Azt kell tehát csak megmutatnunk, hogy ha F mindkét változójában folytonosan parciálisan differenciálható, akkor differenciálható is.

Legyen $(a_1,a_2) \in \mathrm{Dom}(F)$. A $0 \in V_1$ egy környezetében értelmezett

$$V_1 \to W, \quad h_1 \mapsto F(a_1 + h_1, a_2) - \mathrm{D}_1F(a_1,a_2)h_1$$

függvény differenciálható, ezért alkalmazhatjuk rá a középértéktételt:

$$\begin{aligned}&\|F(a_1+h_1,a_2) - F(a_1,a_2) - \mathrm{D}_1F(a_1,a_2)h_1\| \le \\ &\qquad\qquad \le \sup_{\theta\in[0,1]} \|\mathrm{D}_1F(a_1+\theta h_1, a_2) - \mathrm{D}_1F(a_1,a_2)\|\ \|h_1\|.\end{aligned}$$

Teljesen hasonlóan, rögzített h_1 mellett

$$\begin{aligned}&\|F(a_1+h_1,a_2+h_2)-F(a_1+h_1,a_2)-\mathrm{D}_2F(a_1,a_2)h_2\|\le\\&\qquad\le\sup_{\theta\in[0,1]}\|\mathrm{D}_2F(a_1+h_1,a_2+\theta h_2)-\mathrm{D}_2F(a_1,a_2)\|\,\|h_2\|.\end{aligned}$$

A parciális deriváltak folytonossága miatt minden $\epsilon>0$ számhoz létezik $\delta>0$ úgy, hogy minden $(h_1,h_2)\in G_\delta(0,0)$ esetén

$$\|\mathrm{D}_1F(a_1+h_1,a_2+h_2)-\mathrm{D}_1F(a_1,a_2)\|\le\epsilon,$$

$$\|\mathrm{D}_2F(a_1+h_1,a_2+h_2)-\mathrm{D}_2F(a_1,a_2)\|\le\epsilon.$$

Tehát ilyen (h_1,h_2)-re

$$\begin{aligned}&\|F(a_1+h_1,a_2+h_2)-F(a_1,a_2)-\big(\mathrm{D}_1F(a_1,a_2)h_1+\mathrm{D}_2F(a_1,a_2)h_2\big)\|\le\\&\qquad=\|F(a_1+h_1,a_2,h_2)-F(a_1+h_1,a_2)-\mathrm{D}_2F(a_1,a_2)h_2\|+\\&\qquad\quad+\|F(a_1+h_1,a_2)-F(a_1,a_2)-\mathrm{D}_1F(a_1,a_2)h_1\|\le\\&\qquad\qquad\le\epsilon(\|h_1\|+\|h_2\|)=\epsilon\|(h_1,h_2)\|.\end{aligned}$$

Ez pedig az 1.2. $(*)$ összefüggése szerint azt jelenti, hogy F differenciálható (a_1,a_2)-ben. ■

Felhívjuk a figyelmet arra, hogy az állításban a folytonos (parciális) differenciálhatóság – definíciója szerint – a (parciális) deriváltak nyílt halmazon való folytonosságát jelenti.

Nem elég, ha csak egy pontbeli folytonosságot követelünk meg: abból, hogy egy pontban folytonosak a parciális deriváltak, még az sem következik általában, hogy a függvény az adott pontban differenciálható, nemhogy a derivált folytonos lenne.

Viszont az sem elég, ha egy nyílt halmazon léteznek a parciális deriváltak; ha nem folytonosak, akkor a függvény nem szükségképpen differenciálható az adott halmaz akár egy pontjában is.

4.7. Vegyünk egy $f:\mathbb{K}^N\rightarrowtail\mathbb{K}^M$ függvényt. Legyen $e_k\in\mathbb{K}^N$ a k-adik standard bázisvektor és tegyük fel, hogy f differenciálható a-ban az e_k irány mentén:

$$\mathrm{D}_{e_k}f(a)=\lim_{t\to 0\in\mathbb{K}}\frac{f(a+te_k)-f(a)}{t}.$$

Ugyanakkor, lévén $\mathbb{K}^N$ a $\mathbb{K}$-nak önmagával vett N-szeres Descartes-szorzata, tekinthetjük f-nek a k-adik parciális deriváltját, amelyet

$$f(a+te_k)-f(a)=\mathrm{D}_kf(a)t+o(te_k)\qquad(t\in\mathbb{K},a+te_k\in\mathrm{Dom}(F))$$

jellemez.

Értelmezésünk szerint tehát $\mathrm{D}_{e_k} f(a) \in \mathbb{K}^M$, $\mathrm{D}_k f(a) \in \mathcal{L}in(\mathbb{K}, \mathbb{K}^M)$. Viszont $\mathbb{K}^M$-et és a $\mathbb{K} \to \mathbb{K}^M$ lineáris leképezéseket a szokásos módon azonosítjuk ($\xi \in \mathbb{K}^M$ az a lineáris leképezés, amely az α számhoz $\alpha\xi$-t rendeli). Ebben az azonosításban, ezt jól látjuk, a standard bázisvektorok menti deriváltak és a parciális deriváltak megegyeznek egymással; szokás rájuk a

$$\partial_k f(a) := \mathrm{D}_{e_k} f(a) \equiv \mathrm{D}_k f(a) \qquad (k = 1, \ldots, N)$$

jelölést használni.

Mint tudjuk, az $f : \mathbb{K}^N \rightarrowtail \mathbb{K}^M$ függvény a $\mathbb{K}^N \rightarrowtail \mathbb{K}$ komponenseinek az együttese: $f = (f_1, \ldots, f_M)$. Ha f differenciálható a-ban, akkor minden komponensfüggvény parciálisan differenciálható az összes változója szerint, azaz képezhető a

$$\big(\partial_k f_i(a) \mid i = 1, \ldots, M, k = 1, \ldots, N\big) \in \mathbb{K}^{M \times N}$$

úgynevezett **Jacobi-mátrix**. Vigyázzunk, sajnos a jelölés itt egy kicsit "felfordítja" a rendet: $\partial_k f_i(a)$ az imént definiált mátrix ik-adik tagja és nem ki-edik tagja! Jól jegyezzük meg: a függvény-komponensek adják a mátrix sorait, a változók szerinti parciális deriváltak az oszlopait:

$$\begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \ldots & \partial_N f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \ldots & \partial_N f_2(a) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 f_M(a) & \partial_2 f_M(a) & \ldots & \partial_N f_M(a) \end{pmatrix}$$

Az $f : \mathbb{K}^N \rightarrowtail \mathbb{K}^M$ differenciálható függvény a-beli deriváltja $\mathrm{D}f(a) : \mathbb{K}^N \to \mathbb{K}^M$ lineáris leképezés, vagyis egy $M \times N$-es mátrix. Könnyű meggyőződni arról, hogy ez éppen a Jacobi-mátrix. Valóban, a derivált ik-adik mátrixtagja, a lineáris leképezések mátrixának definíciója szerint

$$\mathrm{pr}_i\big(\mathrm{D}f(a)e_k\big) = \mathrm{pr}_i\left(\lim_{t\to 0} \frac{f(a + te_k) - f(a)}{t}\right) = \lim_{t\to 0} \frac{f_i(a + te_k) - f_i(a)}{t} = \partial_k f_i(a).$$

Itt újra felhívjuk a figyelmet arra, hogy egy függvény minden változója szerinti parciális differenciálhatóságából nem következik a függvény differenciálhatósága. Előfordulhat tehát, hogy minden i és k esetén létezik $\partial_k f_i(a)$, a függvény mégsem differenciálható; vagyis képezhető a Jacobi-mátrix, de az nem a függvény deriváltja.

Mindazonáltal a parciális deriváltak létezése jó szolgálatot tehet annak eldöntésében, differenciálható-e a függvény. Ugyanis a parciális deriváltakat általában könnyű kiszámolni a $\mathbb{K} \rightarrowtail \mathbb{K}$ függvényekre megismert differenciálási szabályok alapján. Ha valamelyik parciális derivált nem létezik, a függvény nem differenciálható. Ha minden parciális derivált létezik, akkor a függvény deriváltja – ha differenciálható – csak a Jacobi-mátrix lehet. Ha most $Jf(a)$ jelöli az f függvény a-beli Jacobi-mátrixát, az a-beli differenciálhatóság eldöntéséhez csak azt kell megvizsgálnunk, hogy a

$$h \mapsto f(a+h) - f(a) - Jf(a)h$$

függvény kis ordó-e vagy sem.

Természetesen igen jó hasznát vehetjük vizsgálatunknál annak a ténynek, hogy ha a parciális deriváltak léteznek és folytonosak egy nyílt halmazon, akkor ott a függvény differenciálható (sőt folytonosan differenciálható).

Megemlítjük, hogy alkalmazásokban használatos és igen jó szolgálatot tesz konkrét formulával adott függvények esetén a

$$\frac{\partial f(x_1, x_2, \dots, x_N)}{\partial x_k} := \partial_k f(x_1, x_2, \dots, x_N) \qquad (k = 1, \dots, N)$$

jelölés.

4.8. Feladatok

1. Vizsgáljuk meg a következő függvényeket parciális differenciálhatóság és folytonos parciális differenciálhatóság szempontjából:

(i) $C[0,1] \times C[0,1] \to C[0,1], \quad (f,g) \mapsto fg,$
(lássuk el $C[0,1]$-et sorban mindhárom ismert normával);

(ii) $\mathcal{L}in(V) \times \mathcal{L}in(V) \to \mathcal{L}in(V), \quad (A,B) \mapsto AB.$

2. Legyen $R : \bigtimes_{i=1}^{n} V_i \to W$ folytonos n-lineáris leképezés. Mutassuk meg, hogy minden változója szerint folytonosan parciálisan differenciálható, és

$$\mathrm{D}_i R(a) : V_i \to W, \quad h_i \mapsto R(a_1, \dots, a_{i-1}, h_i, a_{i+1}, \dots, a_N).$$

Adjuk meg ennek alapján R deriváltját.

3. Legyen $R : V^n \to W$ folytonos n-lineáris leképezés. Mivel $V \to V^n$, $x \mapsto (x, x, \dots x)$ folytonos lineáris leképezés, mutassuk meg az előző feladat alapján, hogy a $V \to W$, $x \mapsto R(x, x, \dots, x)$ leképezés differenciálható, és deriváltja a-ban a

$$h \mapsto R(h, a, \dots, a) + R(a, h, a, \dots, a) + \dots R(a, \dots, a, h, a, \dots, a) + \dots \\ + R(a, \dots, a, h)$$

folytonos lineáris leképezés.

Ha tehát R szimmetrikus, akkor a deriváltja a-ban az $nR(\cdot, a, \dots, a)$ folytonos lineáris leképezés.

4. Vessük össze az előző feladatot az 1.11.13. feladattal!

5. Igazoljuk (az 1.11.12. feladattól függetlenül) parciális deriváltakkal, hogy a $\mathbb{K}^{N \times N} \to \mathbb{K}$, $X \mapsto \det X$ leképezés differenciálható az $\mathrm{id}_{\mathbb{K}^N}$-ben, és ott a deriváltja a $\mathbb{K}^{N \times N} \to \mathbb{K}$, $H \mapsto \mathrm{Tr}(H)$ lineáris leképezés.

6. Adjuk meg a következő függvények derivált függvényeit:
(i) $\mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto (x^2+y^2)\sin x \cos y$;
(ii) $\mathbb{R}^3 \rightarrowtail \mathbb{R}, \quad (x,y,z) \mapsto (\log x)e^{y+z} + \sin y \operatorname{arctg}(z/x)$.
7. Mi az $\mathbb{R}^3 \to \mathbb{R}$, $(x,y,z) \mapsto xyz$ függvény deriváltja az $(1,2,3)$ pontban az $(1,0,1)$ irány mentén?
8. Mely egységvektor irányában a legnagyobb a deriváltja az $\mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto x^2 - xy + y^2$ függvénynek az $(1,1)$ pontban?
9. Differenciálhatók-e az alábbi $\mathbb{R}^2 \to \mathbb{R}$ függvények:
(i) $(x,y) \mapsto \sqrt{|xy|}$, (ii) $(x,y) \mapsto \sqrt[3]{xy}$, (iii) $(x,y) \mapsto \sqrt[3]{x^3+y^3}$,

(iv) $(a,b) \mapsto \begin{cases} \frac{ab}{\sqrt{a^2+b^2}} & \text{ha } a^2+b^2 \neq 0, \\ 0 & \text{ha } a^2+b^2 = 0. \end{cases}$,

(v) $(u,v) \mapsto \begin{cases} (u^2+v^2)\sin\frac{1}{u^2+v^2} & \text{ha } u^2+v^2 \neq 0, \\ 0 & \text{ha } u^2+v^2 = 0. \end{cases}$.

5. A komplex és valós differenciálhatóság összevetése

5.1. Értelmeztük a $\mathbb{K}$ test fölötti normált terek közötti függvények differenciálhatóságát.

Komplex vektorterek egyben valós vektorterek is, és egy $\mathbb{C}$-lineáris leképezés egyben $\mathbb{R}$-lineáris is.

Ha viszont V és W komplex vektortér, az $A : V \to W$ $\mathbb{R}$-lineáris leképezés nem feltétlenül $\mathbb{C}$-lineáris; ismeretes, hogy pontosan akkor $\mathbb{C}$-lineáris, ha $A(ix) = iAx$ minden $x \in V$ esetén, vagyis ha az i imaginárius egység "kiemelhető" A alól.

Legyen V és W komplex (így egyben valós) normált tér, $F : V \rightarrowtail W$ függvény és $a \in \overset{\circ}{\mathrm{Dom}}(F)$. Az 1.2. definíció értelmében • F $\mathbb{C}$-differenciálható a-ban, ha létezik olyan $A : V \to W$ folytonos, $\mathbb{C}$-lineáris leképezés, hogy

$$F(a+h) - F(a) = Ah + o_A(h), \qquad (a+h \in \mathrm{Dom}(F)),$$

ahol $o_A : V \rightarrowtail W$ kis ordó függvény;

• F $\mathbb{R}$-differenciálható a-ban, ha létezik olyan $B : V \to W$ folytonos, $\mathbb{R}$-lineáris leképezés, hogy

$$F(a+h) - F(a) = Bh + o_B(h), \qquad (a+h \in \mathrm{Dom}(F)),$$

ahol $o_B : V \rightarrowtail W$ kis ordó függvény.

A fentiek alapján látható, hogy ha F $\mathbb{C}$-differenciálható a-ban, akkor egyben $\mathbb{R}$-differenciálható is. Fordítva ez nem igaz.

Ha F $\mathbb{R}$-differenciálható a-ban, akkor és csak akkor $\mathbb{C}$-differenciálható is, ha $B(ix) = iBx$ minden $x \in V$ esetén.

5.2. Talán a legjobban látható a különbség a valós és a komplex differenciálhatóság között, ha a $V = W = \mathbb{C}$ esetet vesszük; ekkor az ezeknek megfelelő valós vektortér $\mathbb{R}^2$.

Ha $B : \mathbb{C} \to \mathbb{C}$ $\mathbb{R}$-lineáris, vagyis $B : \mathbb{R}^2 \to \mathbb{R}^2$ lineáris, akkor 2×2-es valós mátrixként írható föl:

$$B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}.$$

Az i imaginárius egységgel való szorzás mátrixa

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

hiszen $i(x + iy) = (-y + ix)$, azaz $\mathbb{R}^2 \to \mathbb{R}^2$ leképezésként az i-vel való szorzás $(x, y) \mapsto (-y, x)$ módon írható.

Az, hogy i kiemelhető a B alól, azt jelenti, hogy a fenti két mátrix felcserélhető, ami pontosan akkor teljesül, ha

$$b_{11} = b_{22}, \quad b_{12} = -b_{21}.$$

Legyen $f : \mathbb{C} \rightarrowtail \mathbb{C}$ mint $f \equiv (f_1, f_2) : \mathbb{R}^2 \rightarrowtail \mathbb{R}^2$ függvény $\mathbb{R}$-differenciálható az a pontban, ahol az azonosításnak megfelelően f_1 és f_2 a valós illetve képzetes része f-nek. A $\mathrm{D}f(a) : \mathbb{R}^2 \rightarrowtail \mathbb{R}^2$ valós lineáris leképezés mátrixa:

$$\begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a). \end{pmatrix}$$

f akkor és csak akkor $\mathbb{C}$-differenciálható is a-ban, ha $\mathrm{D}f(a)$ $\mathbb{C}$-lineáris is, azaz

$$\partial_1 f_1(a) = \partial_2 f_2(a), \quad \partial_1 f_2(a) = -\partial_2 f_1(a).$$

Ezeket az összefüggéseket **Cauchy–Riemann-egyenleteknek** szokták nevezni.

Összefoglalva: a komplex differenciálhatóság sokkal erősebb követelmény, mint a valós differenciálhatóság, hiszen az utóbbi mellett a deriváltnak még plusz feltételeket kell kielégítenie, hogy az előbbi igaz legyen; például a $\mathbb{C} \rightarrowtail \mathbb{C}$ esetben konkrétan a Cauchy–Riemann-egyenleteket.

5.3. Feladatok

1. Mutassuk meg, hogy a $\mathbb{C} \to \mathbb{C}$ komplex konjugálás, valamint a Re és Im függvény $\mathbb{R}$-differenciálhatók (viszont nem $\mathbb{C}$-differenciálhatók, lásd A.1.5.(v)).

2. Adjunk meg az előző példában szereplőktől különböző olyan $\mathbb{C} \to \mathbb{C}$ függvényeket, amelyek $\mathbb{R}$-differenciálhatók, de nem $\mathbb{C}$-differenciálhatók.

6. Az inverzfüggvény-tétel

6.1. Állítás *Legyenek V és W Banach-terek (azaz teljes normált terek), $F : V \rightarrowtail W$ folytonosan differenciálható, $a \in \mathrm{Dom}(F)$, és $\mathrm{D}F(a)$ invertálható (azaz bijektív és az inverze is folytonos). Ekkor létezik olyan $N \subset V$ nyílt halmaz, hogy $a \in N \subset \mathrm{Dom}(F)$, $\mathrm{D}F(x)$ invertálató minden $x \in N$ esetén, $F|_N$ injektív, $F[N]$ nyílt, $(F|_N)^{-1}$ folytonosan differenciálható, és*

$$\mathrm{D}\left(F|_N\right)^{-1}(F(x)) = (\mathrm{D}F(x))^{-1}. \tag{1}$$

Bizonyítás Az egyszerűség kedvéért célszerű bevezetni az $A := \mathrm{D}F(a)$ jelölést. Tetszőleges $y \in W$ esetén definiáljuk a

$$\phi_y : V \rightarrowtail V, \quad x \mapsto x + A^{-1}(y - F(x)) \qquad (x \in \mathrm{Dom}(F))$$

leképezést. Mivel A bijekció, $\phi_y(x) = x$ pontosan akkor teljesül – más szóval, x a ϕ_y fix pontja –, amikor $y = F(x)$. Tehát ϕ_y-nak csak $y \in \mathrm{Ran}(F)$ esetén lehet fix pontja. Továbbá ϕ_y differenciálható, és minden $x \in \mathrm{Dom}(F)$ esetén

$$\mathrm{D}\phi_y(x) = \mathrm{id}_V - A^{-1}\mathrm{D}F(x) = A^{-1}(A - \mathrm{D}F(x)).$$

(i) F folytonosan differenciálható, ezért létezik olyan $N \subset \mathrm{Dom}(F)$ konvex, nyílt halmaz, $a \in N$, hogy

$$\|\mathrm{D}F(x) - A\| \leq \frac{1}{2\|A^{-1}\|} \qquad (x \in N).$$

Ismereteink alapján bármely $x \in N$ pontban $\mathrm{D}F(x)$ invertálható (Analízis II-I.B.10.10.); ha tehát $F|_N$ injektív és az inverze differenciálható, akkor 1.10 szerint fennáll az (1) egyenlőség.

(ii) Megmutatjuk, hogy $F|_N$ injektív. Ugyanis az előbbi két egyenlőségből

$$\|\mathrm{D}\phi_y(x)\| \leq \|A^{-1}\|\ \|A - \mathrm{D}F(x)\| \leq \frac{1}{2} \qquad (x \in N),$$

így az általános középértéktétel miatt

$$\|\phi_y(x_1) - \phi_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\| \qquad (x_1, x_2 \in N). \tag{2}$$

Ezek szerint ϕ_y kontrakció N-en, tehát ott csak egy fix pontja lehet, vagyis bármely $y \in F[N]$ esetén egyetlen olyan $x \in N$ létezhet (és létezik is, hiszen $y \in F[N]$), hogy $y = F(x)$; ez pontosan azt jelenti, hogy $F|_N$ injektív.

(iii) Megmutatjuk, hogy $F[N] \subset W$ nyílt halmaz. Ha $y_0 \in F[N]$, akkor létezik egyetlen olyan $x_0 \in N$, hogy $y_0 = F(x_0)$. Mivel N nyílt, van olyan $r > 0$, hogy az x_0 körüli r sugarú zárt gömb is része N-nek: $B_r(x_0) \subset N$. Ha $x \in B_r(x_0)$ és

$$\|y - y_0\| < \frac{r}{2\|A^{-1}\|} =: s,$$

akkor

$$\begin{aligned}\|\phi_y(x) - x_0\| &\leq \|\phi_y(x) - \phi_y(x_0)\| + \|\phi_y(x_0) - x_0\| \leq \\ &\leq \frac{1}{2}\|x - x_0\| + \|x_0 + A^{-1}(y - F(x_0)) - x_0\| \leq \\ &\leq \frac{r}{2} + \|A^{-1}\| \; \|y - y_0\| \leq r,\end{aligned}$$

vagyis $\phi_y[B_r(x_0)] \subset B_r(x_0)$ ha $\|y - y_0\| \leq s$. Tekintve, hogy $B_r(x_0)$ zárt része egy teljes normált térnek, ő is teljes; ϕ_y pedig kontrakció rajta, ha $y \in G_s(y_0)$. Következésképpen ϕ_y-nak létezik egyetlen fix pontja, azaz (egyértelműen) létezik $x \in B_r(x_0) \subset N$, hogy $y = F(x)$. Tehát ha $y \in G_s(y_0)$, akkor $y \in F(N)$; ez azt jelenti, hogy $F[N]$ nyílt.

Vegyük észre, a gondolatmenetünkkel azt is beláttuk, hogy $F|_N$ nyílt leképezés – és így $(F|_N)^{-1}$ folytonos függvény –, ugyanis a fentiek elmondhatók az N bármely nyílt részhalmazára, vagyis az N nyílt részhalmazainak F általi képe is nyílt.

(iv) Ebben a lépésben azt igazoljuk, hogy a $G := (F|_N)^{-1}$ függvény differenciálható. Ehhez először belátjuk, hogy eleget tesz a Lipschitz-feltételnek (tehát folytonos is) az $F[N]$ halmazon .

Legyen $y, y + k \in F[N]$, és tekintsük a következő átalakítást:

$$\begin{aligned}\|G(y+k) - G(y)\| &\leq \|G(y+k) - G(y) - A^{-1}k\| + \|A^{-1}k\| = \\ &= \|\phi_y(G(y+k)) - \phi_y(G(y))\| + \|A^{-1}k\| \leq \\ &\leq \frac{1}{2}\|G(y+k) - G(y)\| + \|A^{-1}\| \; \|k\|,\end{aligned}$$

ahol felhasználtuk a (2) összefüggést. Egyszerű átrendezéssel

$$\|G(y+k) - G(y)\| \leq 2\|A^{-1}\| \; \|k\| \qquad (y, y + k \in F[N]) \tag{3}$$

adódik.

F folytonosan differenciálható, ezért van olyan $o_F : V \rightarrowtail W$ kis ordó függvény, hogy

$$\begin{aligned}k = (y + k) - y &= F(G(y+k)) - F(G(y)) = \\ &= \mathrm{D}F(G(y))\big(G(y+k) - G(y)\big) + o_F(G(y+k) - G(y)),\end{aligned}$$

azaz

$$\mathrm{D}F(G(y))\big(G(y+k)-G(y)\big)-k=-o_F(G(y+k)-G(y)). \tag{4}$$

Mivel $o_F=\|.\|\eta$, ahol $\eta: V \rightarrowtail W$, $\lim\limits_{0}\eta=0$, a (3) összefüggés miatt

$$\begin{aligned}\|o_F G(y+k)-G(y)\| = \|G(y+k)-G(y)\|\ \|\eta(G(y+k)-G(y))\| \le \\ \le 2\|A^{-1}\|\ \|k\|\ \|\eta(G(y+k)-G(y))\|, \quad (5)\end{aligned}$$

tehát

$$\lim_{k\to 0}\frac{\|o_F G(y+k)-G(y))\|}{\|k\|}\le \lim_{k\to 0}2\|A^{-1}\|\ \|\eta(G(y+k)-G(y)\|=0.$$

Az (i)-ben mondottak szerint G differenciálhatóságához azt kell megmutatnunk, hogy minden $y\in F[N]$ esetén létezik olyan $o_G: W\rightarrowtail V$ kis ordó függvény, hogy

$$G(y+k)-G(y)=(\mathrm{D}F(G(y)))^{-1}\,k+o_G(k) \quad (y+k\in F[N]).$$

Íme: a (4) összefüggés alapján

$$\begin{aligned}G(y+k)-G(y)-(\mathrm{D}F(G(y)))^{-1}\,k= \\ =(\mathrm{D}F(G(y))^{-1}\Big(\mathrm{D}F(G(y))\big(G(y+k)-G(y)\big)-k\Big)= \\ =-(\mathrm{D}F(G(y))^{-1}\Big(o_F(G(y+k)-G(y))\Big).\end{aligned}$$

Mivel $\mathrm{D}F(G(y))^{-1}$ folytonos lineáris leképezés, az (5) összefüggés miatt a

$$W\to V,\quad k\mapsto -(\mathrm{D}F(G(y))^{-1}\Big(o_F(G(y+k)-G(y))\Big)=: o_G(k)$$

lekpezés függvény kis ordó függvény, és ezzel beláttuk, hogy G differenciálható.

(v) Már csak azt kell megnéznünk, hogy az $y\mapsto \mathrm{D}G(y)$ hozzárendelés folytonos-e. Könnyen látható, hogy igen, hiszen

$$y\mapsto G(y)\mapsto \mathrm{D}F(G(y))\mapsto (\mathrm{D}F(G(y))^{-1}=\mathrm{D}G(y)$$

módon írható föl folytonos függvények kompozíciójaként (az invertálás differenciálható, tehát folytonos leképezés).

6.2. (i) A bizonyítás során is utaltunk arra, hogy ha $F: V\rightarrowtail W$ folytonosan differenciálható és $\mathrm{D}F(a)$ invertálható, akkor F nyílt leképezés az a egy környezetében; ha tehát minden $x\in \mathrm{Dom}(F)$ esetén $\mathrm{D}F(x)$ invertálató, akkor F nyílt leképezés.

(ii) Ha $\mathrm{D}F(x)$ invertálható minden $x \in \mathrm{Dom}(F)$ esetén, akkor F lokálisan injektív, vagyis minden pontnak egy környezetén injektív; ez azonban nem jelenti azt, hogy F injektív. Például ha a cos függvény értelmezési tartományából kivesszük a $\mathbb{Z}\pi$ halmazt, lokálisan injektív de nem injektív függvényt kapunk.

(iii) $\mathrm{D}F(a)$ invertálhatósága nem szükséges ahhoz, hogy F injektív legyen az a egy környezetében, de ahhoz igen, hogy az inverz függvény differenciálható legyen. Például az $\mathrm{id}_{\mathbb{R}}^3$ függvény injektív, de a deriváltja a nullában nulla.

(iv) Ha V és W véges dimenziós vektorterek, akkor a $\mathrm{D}F(a) : V \to W$ lineáris leképezés csak úgy lehet bijekció (az inverze eleve folytonos), ha $\dim V = \dim W$, tehát az inverzfüggvény-tétel feltételei csak ebben az esetben teljesülhetnek. Így például $V = \mathbb{R}^M$ és $W = \mathbb{R}^N$ esetben $M = N$ szükséges, és az inverzfüggvény létezése azon fog múlni, hogy az adott pontban a Jacobi-mátrix determinánsa (a Jacobi-determináns) nulla-e vagy sem. Ugyanis a Jacobi-mátrix (amely tulajdonképpen a pontbeli derivált) pontosan akkor injektív, ha a determinánsa nem nulla.

6.3. Feladatok

1. Igazoljuk, hogy egy $\mathbb{R} \rightarrowtail \mathbb{R}$ folytonos függvény, amely összefüggő halmazon (tehát intervallumon) lokálisan injektív, az injektív.

2. Mutassuk meg, hogy az $\{(x,y) \in \mathbb{R}^2 \mid xy > 2\}$ összefüggő halmazon értelmezett $\mathbb{R}^2 \rightarrowtail \mathbb{R}^2$, $(x,y) \mapsto \left(x, \dfrac{x}{y} - \dfrac{1}{y^2}\right)$ függvény lokálisan injektív de nem injektív.

3. Mely pontok környezetében van az alábbi függvényeknek folytonosan differenciálató inverze:

(i) $\mathbb{R}^2 \to \mathbb{R}^2, \quad (x,y) \mapsto (x - y^2, x^2 - y^2)$,

(ii) $\mathbb{R}^2 \to \mathbb{R}^2, \quad (u,v) \mapsto (e^u, u \sin v)$,

(iii) $\mathbb{R}^3 \rightarrowtail \mathbb{R}^3, \quad (x,y,z) \mapsto \left(\sqrt{y^2+z^2}, \dfrac{y}{x+z}, \cos(x - y^2)\right)$.

7. Az implicitfüggvény-tétel

7.1. Idézzük fel, hogy a függvény fogalmának meghatározásában a leglényegesebb szerepet a függvényszerű reláció játszotta: ennek segítségével vezettük be a hozzárendelés fogalmát. A gyakorlatban a függvényeket általában hozzárendelési utasítással adjuk meg. Előfordul azonban, hogy egy reláció áll rendelkezésünkre, és arról kell eldöntenünk, meghatároz-e egy függvényt, azaz függvényszerű-e.

A szokásos szóhasználattal élve gyakran úgy vetik fel a kérdést, hogy adott feltételekből bizonyos számú változó ("mennyiség") függvényeként kifejezhető-e a többi. Íme egy példa:

"Kifejezhető-e, illetve mely pont környezetében fejezhető ki u és v az x és y

függvényében az

$$u^2 + x^2 = vy, \qquad ux = v + y$$

összefüggésekből ?”

Az ilyen kérdések megválaszolására egy speciális, de széles körben alkalmazható esetre a differenciálszámítás jól használható módszert nyújt.

7.2. Állítás *Legyenek V_1, V_2, W Banach-terek, $F : V_1 \times V_2 \rightarrowtail W$ folytonosan differenciálható függvény, $(a_1, a_2) \in \mathrm{Dom}(F)$, $c := F(a_1, a_2)$. Ha a $\mathrm{D}_2 F(a_1, a_2) \in \mathcal{L}in(V_2, W)$ folytonos lineáris leképezés invertálható, akkor van olyan $N_1 \subset V_1$ és $N_2 \subset V_2$ nyílt halmaz, hogy $(a_1, a_2) \in N_1 \times N_2$, és az $(N_1 \times N_2) \cap \overset{-1}{F}\{c\} \subset V_1 \times V_2$ halmaz egy egyértelműen meghatározott folytonosan differenciálható függvény grafikonja; közelebbről, létezik egyetlen olyan $G : N_1 \to N_2$ folytonosan differenciálható függvény, hogy*

$$F(x_1, G(x_1)) = c, \quad \textit{ahol} \quad x_1 \in N_1,$$

továbbá

$$\mathrm{D}G(x_1) = -\left(\mathrm{D}_2 F(x_1, G(x_1))\right)^{-1} \mathrm{D}_1 F(x_1, G(x_1)).$$

Bizonyítás A bizonyítás során az előző pontban tárgyalt inverzfüggvény-tételt fogjuk kihasználni.

Vezessük be az $\hat{F} := (\mathrm{pr}_1, F)$ függvényt, tehát

$$\hat{F} : V_1 \times V_2 \rightarrowtail V_1 \times W, \quad (x_1, x_2) \mapsto (x_1, F(x_1, x_2)).$$

Mint folytonosan differenciálható függvények együttese, $\hat{F}$ is folytonosan differenciálható, és

$$\mathrm{D}\hat{F}(x_1, x_2) = (\mathrm{pr}_1, \mathrm{D}F(x_1, x_2)).$$

Megmutatjuk, hogy $\mathrm{D}\hat{F}(x_1, x_2)$ pontosan akkor invertálható, ha $\mathrm{D}_2 F(x_1, x_2)$ invertálható.

Minden $(h_1, h_2) \in V_1 \times V_2$ esetén

$$\big(\mathrm{D}\hat{F}(x_1, x_2)\big)(h_1, h_2) = (h_1, \mathrm{D}_1 F(x_1, x_2)h_1 + \mathrm{D}_2 F(x_1, x_2)h_2).$$

Ebből látszik, hogy ha $\mathrm{D}_2 F(x_1, x_2)$ invertálható, akkor $\mathrm{D}\hat{F}(x_1, x_2)$ bijekció, és inverze a

$$V_1 \times W \to V_1, \times V_2, \quad (h_1, k) \mapsto (h_1, \mathrm{D}_2 F(x_1, x_2)^{-1} k - \mathrm{D}_2 F(x_1, x_2)^{-1} \mathrm{D}_1 F(x_1, x_2) h_1)$$

leképezés, amely folytonos.

Ha $\mathrm{D}\hat{F}(x_1, x_2)$ invertálható, akkor $\mathrm{D}_2 F(x_1, x_2) \equiv \mathrm{D}\hat{F}(x_1, x_2)|_{\{0\}\times V_2}$ szintén invertálható.

Így az inverzfüggvény-tétel alapján létezik olyan $\hat{N} := \hat{N}_1 \times \hat{N}_2 \subset V_1 \times V_2$ nyílt halmaz, hogy $(a_1, a_2) \in \hat{N}$, $\hat{F}|_{\hat{N}}$ injektív, a deriváltja minden $(x_1, x_2) \in \hat{N}$ esetén invertálható, $\hat{F}[\hat{N}]$ nyílt halmaz, továbbá

$$\hat{G} := \left(\hat{F}|_{\hat{N}}\right)^{-1} : V_1 \times W \rightarrowtail V_1 \times V_2$$

folytonosan differenciálható.

Vizsgáljuk meg a $\hat{G} =: (\hat{G}_1, \hat{G}_2)$ komponenseit: ha $(x_1, y) \in \mathrm{Dom}(\hat{G})$, akkor

$$\begin{aligned}(x_1, y) = \hat{F}(\hat{G}(x_1, y)) = \hat{F}(\hat{G}_1(x_1, y), \hat{G}_2(x_1, y)) = \\ = \Big(\hat{G}_1(x_1, y), F\big(\hat{G}_1(x_1, y), \hat{G}_2(x_1, y)\big)\Big);\end{aligned}$$

következésképpen minden $(x_1, y) \in \mathrm{Dom}(\hat{G})$ esetén

$$x_1 = \hat{G}_1(x_1, y), \qquad y = F\big(x_1, \hat{G}_2(x_1, y)\big). \tag{$*$}$$

Mivel $F(a_1, a_2) = c$, nyilvánvaló, hogy $(a_1, c) \in \mathrm{Dom}(\hat{G}) = \hat{F}[\hat{N}]$. Az $x_1 \mapsto (x_1, c)$ leképezés folytonossága és $\hat{F}[\hat{N}]$ nyíltsága következtében van olyan $N_1 \subset \hat{N}_1$ nyílt halmaz, hogy $N_1 \times \{c\} \subset \hat{F}[\hat{N}]$.

Legyen $N_2 := \hat{N}_2$, és definiáljuk a

$$G : N_1 \to N_2, \quad x_1 \mapsto \hat{G}_2(x_1, c)$$

függvényt. Erre a $(*)$ összefüggés szerint

$$F\big(x_1, G(x_1)\big) = F\big(x_1, \hat{G}_2(x_1, c)\big) = c \qquad (x_1 \in N_1)$$

teljesül, tehát $(N_1 \times N_2) \cap \overset{-1}{F}\{c\}$ a G grafikonja.

A G egyértelműsége az $\hat{F}$ injektivitásából származtatható: ha van egy $H : N_1 \to N_2$ függvény, amelyre $F(x_1, H(x_1)) = c$ azaz $\hat{F}(x_1, H(x_1)) = (x_1, c) = \hat{F}(x_1, G(x_1))$ teljesül, akkor $H(x_1) = G(x_1)$.

G folytononosan differenciálható, hiszen $\hat{G}$ folytonosan differenciálható és emiatt $\hat{G}_2$, továbbá G is az. Az $x_1 \mapsto F\big(x_1, G(x_1)\big) = c$ hozzárendeléssel értelmezett $N_1 \to W$ azonosan nulla függvény deriváltja minden pontban a nulla lineáris leképezés. A közvetett függvények differenciálására vonatkozó szabályokat alkalmazva ezt a következő formában írhatjuk föl:

$$\mathrm{D}_1 F(x_1, G(x_1)) + \mathrm{D}_2 F(x_1, G(x_1)) \mathrm{D}G(x_1) = 0.$$

Mivel $(x_1, G(x_1)) \in \hat{N}_1 \times \hat{N}_2$, az F második parciális deriváltja a fenti egyenlőségben invertálható, ezért azonnal adódik a G deriváltjára állított formula.

7.3. (i) Ha az F függvény eleget tesz az előbbi állítás feltételeinek, akkor létezik olyan N környezete az (a_1, a_2) pontnak, hogy $F|_N$ nyílt leképezés.

Erről könnyen meggyőződhetünk, csak arra kell visszaemlékeznünk, hogy – az $N := \hat{N}$ jelöléssel – $F|_N = \mathrm{pr}_W \circ \hat{F}|_N$, tehát nyílt leképezések kompozíciójáról van szó; ugyanis $\mathrm{pr}_W : V_1 \times W \to W$ nyílt, $\hat{F}|_N$ pedig az inverzfüggvény-tétel után mondottak alapján szintén az.

(ii) Az inverzfüggvény-tétellel kapcsolatban tett megjegyzésünkhöz hasonlóan itt is megemlítjük, hogy ha a szóban forgó vektorterek véges dimenziósak, az implicitfüggvény-tétel feltételei csak akkor teljesülhetnek, ha $\dim V_2 = \dim W$, hiszen a $\mathrm{D}_2F(a_1, a_2) : V_2 \to W$ lineáris leképezés csak ekkor lehet bijektív.

(iii) Ha V véges dimenziós vektortér, $F : V \rightarrowtail W$ folytonosan differenciálható, és $\mathrm{D}F(a) : V \to W$ ráképezés (eszerint W is véges dimenziós), akkor létezik olyan N környezete az a pontnak, hogy $F|_N$ nyílt leképezés.

Ezt az állítást visszavezethetjük az előző esetre. A fenti feltételek mellett ugyanis a $V_1 := \mathrm{Ker}(\mathrm{D}F(a))$ lineáris altérnek egy tetszőleges V_2 kiegészítő alterére $\mathrm{D}F(a)|_{V_2} : V_2 \to W$ bijekció. Ez utóbbi feltétel az ismert $V \equiv V_1 \times V_2$ azonosítással éppen azt jelenti, hogy F második parciális deriváltja invertálható az a pontban. Véges dimenziós – tehát teljes – normált terekről lévén szó, az implicitfüggvény-tétel feltételei teljesülnek.

Ha V végtelen dimenziós Banach-tér, akkor a $V_1 := \mathrm{Ker}(\mathrm{D}F(a))$ zárt lineáris altérnek nem feltétlenül létezik *zárt* kiegészítő altere, ezért nincs olyan $V \equiv V_1 \times V_2$ azonosítás, hogy mind V_1, mind V_2 teljes, ezért általánosan itt nem állíthatunk semmit.

(iv) Az inverzfüggvény-tételből vezettük le az implicitfüggvény-tételt, de ez utóbbiból is követekezik az előbbi, tehát egyenertékűek.

Ha ugyanis $F : V \to W$ folytonosan differenciálható, $\mathrm{D}F(a)$ invertálható, akkor az $\tilde{F} : W \times V \rightarrowtail W$, $(y, x) \mapsto y - F(x)$ függvénynek a második változó szerinti deriváltja az $(F(a), a)$ pontban éppen $\mathrm{D}F(a)$, tehát alkalmazhatjuk az implicitfüggvény-tételt: létezik az a-nak N, az $F(a)$-nak R környezete, $G : R \to N$ folytonosan differenciálható függvény úgy, hogy $\tilde{F}(y, G(y)) = \tilde{F}(F(a), a)$, azaz $y - F(G(y)) = 0$, vagyis $y = F(G(y))$ minden $y \in R$ esetn. Ez azt jelenti, hogy F alkalmas leszűkítése a G inverze, tehát G az F alkalmas leszűkítésének az inverze.

7.4. (i) Formálisan az implicit függvénykapcsolatot így szokás megfogalmazni: az $F(x_1, x_2) = c$ egyenletből kell x_2-t meghatározni az x_1 függvényében. Jegyezzük meg: mindig azon változó szerinti parciális deriváltnak kell invertálhatónak lennie, amelyet ki akarunk fejezni a másik változó függvényében.

(ii) Térjünk vissza a 7.1. pontban említett kérdéshez, amelyet így fogalmazhatunk át: adott az

$$F : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2, \quad \big((x, y), (u, v)\big) \mapsto (u^2 + x^2 - vy, ux - v - y)$$

függvény, és keresendő olyan része az $\overset{-1}{F}\{0\}$ halmaznak, amely egy $(\mathbf{u},\mathbf{v}) : \mathbb{R}^2 \rightarrowtail \mathbb{R}^2$ függvény grafikonja, azaz amelyre

$$\mathbf{u}^2(x,y) + x^2 = \mathbf{v}(x,y)y, \qquad \mathbf{u}(x,y)x = \mathbf{v}(x,y) + y.$$

Az implicitfüggvény-tétel szerint a

$$\begin{pmatrix} \frac{\partial F_1(x,y,u,v)}{\partial u} & \frac{\partial F_1(x,y,u,v)}{\partial v} \\ \frac{\partial F_2(x,y,u,v)}{\partial u} & \frac{\partial F_2(x,y,u,v)}{\partial v} \end{pmatrix} = \begin{pmatrix} 2u & -y \\ x & -1 \end{pmatrix}$$

mátrixnak kell bijektívnek, azaz a determinánsának nem nullának lennie, vagyis az $xy \neq 2u$ feltételnek kell teljesülnie. Egészen pontosan tehát az

$$\{\big((x,y),(u,v)\big) \in \mathbb{R}^2 \times \mathbb{R}^2 \mid u^2 + x^2 - vy = ux - v - y = 0, xy \neq 2u\}$$

halmaz minden pontjának egy környezete, elmetszve ezzel a halmazzal, egy $\mathbb{R}^2 \rightarrowtail \mathbb{R}^2$ folytonosan differenciálható függvény grafikonja.

(iii) Érdemes megemlíteni: ha $M < N$ akkor $\mathbb{K}^N$-et többféleképpen – az elemeinek koordinátáit különféleképpen csoportosítva – azonosíthatjuk $\mathbb{K}^{N-M} \times \mathbb{K}^M$-mel, tehát adott $F : \mathbb{K}^N \rightarrowtail \mathbb{K}^M$ folytonosan differenciálható függvény esetén az $F(x_1, \dots, x_N) = c$ egyenlőségből többféleképpen választott M változó fejezhető ki $N - M$ változó függvényében; a feltétel az, hogy a választott M változó szerinti parciális deriváltakból készített mátrix determinánsa nem nulla.

Például az előbb vizsgált esetben azt is kérdezhetjük, hogy kifejezhető-e u és y a v és x függvényében, stb.

(iv) Igen gyakran talákozunk $F : \mathbb{R}^2 \rightarrowtail \mathbb{R}$ függvénnyel megadott $F(x,y) = c$ implicit kapcsolattal; ebből y az x függvényében olyan pontokban fejezhető ki, ahol $\dfrac{\partial F(x,y)}{\partial y} \neq 0$. Az y-nak mint az x függvényének a deriváltjára a

$$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}$$

jelölést szokás használni, amely egyszerű, kifejező, de félrevezető is lehet; pontosan azt írhatjuk, hogy ha $\mathbf{y}$ jelöli azt a függvényt, amelyre $F(x,\mathbf{y}(x)) = c$ teljesül, akkor

$$\frac{d\mathbf{y}(x)}{dx} = -\frac{\frac{\partial F(x,y)}{\partial x}\Big|_{y=\mathbf{y}(x)}}{\frac{\partial F(x,y)}{\partial y}\Big|_{y=\mathbf{y}(x)}}.$$

7.5. Feladatok

1. Legyen adott $\alpha \in \mathbb{R}$ esetén $y \in \mathbb{R}$ az $x \in \mathbb{R}$ függvénye az $x^3 + y^3 - 3\alpha xy = 0$ implicit kapcsolattal megadva. Határozzuk meg y'-t abban az a pontban, amelyre $y(a) = a$.

Másképp fogalmazva: legyen adott $\alpha \in \mathbb{R}$ esetén

$$F : \mathbb{R}^2 \to \mathbb{R}, \quad x^3 + y^3 - 3\alpha xy.$$

Adjuk meg az $(a, a) \in \overset{-1}{F}\{0\}$ környezetében létező $\mathbf{y} : \mathbb{R} \rightarrowtail \mathbb{R}$, $F(x, \mathbf{y}(x)) = 0$ függvénynek a deriváltját az a pontban.

2. Legyen adott $\alpha \in \mathbb{R}$ esetén $y \in \mathbb{R}$ az $x \in \mathbb{R}$ függvénye az $x = y - \alpha \sin(y)$ implicit kapcsolattal megadva. Írjuk fel y'-t és y''-t!

3. Legyen adott $\alpha, \beta \in \mathbb{R}$ esetén $y \in \mathbb{R}$ az $x \in \mathbb{R}$ függvénye az $(x^2 + y^2 - \beta x)^2 = \alpha^2(x^2 + y^2)$ implicit kapcsolattal megadva. Határozzuk meg a nullában annak a függvénynek a deriváltját, amelyre $y(0) = 0$ teljesül.

4. Az

$$F : \mathbb{R}^3 \to \mathbb{R}^2, \quad (x, y, z) \mapsto (x^2 + y^2 - 2z^2, x^2 + 2y^2 + z^2), \qquad \overset{-1}{F}\{(0, 4)\}$$

formula impliciten megadja x-et és y-t a z függvényeként. Határozzuk meg $\dfrac{dx}{dz}$-t és $\dfrac{d^2y}{dz^2}$-et az $(1, -1, 1)$ pontban.

5. Definiálja adott $\alpha \in \mathbb{R}$ esetén az

$$x + y + z = \alpha, \quad x^3 + y^3 + z^3 = 3xyz,$$

implicit kapcsolat y-t és z-t az x függvényében. Határozzuk meg ezeknek az $\mathbb{R} \rightarrowtail \mathbb{R}$ függvényeknek a deriváltját.

II. TÖBBSZÖR DIFFERENCIÁLHATÓ FÜGGVÉNYEK

8. Magasabb rendű deriváltak, Young-tétel

8.1. Az előző részben tárgyaltuk normált terek közötti $F : V \rightarrowtail W$ függvények differenciálhatóságát. Láttuk, hogy az ilyen függvények deriváltjai minden pontban $V \to W$ folytonos lineáris leképezések, a derivált leképezés $\mathrm{D}F : V \rightarrowtail \mathcal{L}in(V, W)$ függvény. Korábbi tanulmányainkból tudjuk, hogy normált terek közötti, mindenhol értelmezett, folytonos lineáris leképezések, ha ellátjuk őket a szokásos normával, szintén normált teret alkotnak, azaz $\mathcal{L}in(V, W)$ is normált tér (Analízis III.B.10.3.). Ezért beszélhettünk arról az előzőekben, hogy egy függvény folytonosan differenciálható, azaz a deriváltja folytonos. Viszont értelmes a derivált függvények differenciálhatóságának fogalma is, aminek segítségével értelmezhetjük a normált terek közötti függvények kétszeres és többszörös differenciálhatóságát.

Definíció *Egy $F : V \rightarrowtail W$ függvény*

(i) ***kétszer differenciálható*** *egy $a \in V$ pontban, ha $a \in \overset{\circ}{\mathrm{Dom}}(\mathrm{D}F)$, és a $\mathrm{D}F : V \rightarrowtail \mathcal{L}in(V, W)$ függvény differenciálható a-ban, és ekkor a $\mathrm{D}^2F(a) := \mathrm{D}(\mathrm{D}F)(a) \in \mathcal{L}in(V, \mathcal{L}in(V, W))$ deriváltja az F* ***második deriváltja*** *az a pontban,*

(ii) ***kétszer differenciálható*** *egy halmazon, ha $\mathrm{D}F$ differenciálható ezen a halmazon,*

(iii) ***kétszer differenciálható****, ha kétszer differenciálható az értelmezési tartományán.*

(iv) ***második deriváltja*** *az $x \mapsto \mathrm{D}^2F(x)$ hozzárendelési utasítással az $\{x \in \mathrm{Dom}(F) \mid F$ kétszer differenciálható x-ben$\}$ halmazon értelmezett $V \rightarrowtail \mathcal{L}in(V, Lin(V, W))$ függvény.*

8.2. Az előző definícióból kiderül, hogy egy $F : V \rightarrowtail W$ kétszer differenciálható függvény második deriváltja egy a pontban a $\mathcal{L}in(V, \mathcal{L}in(V, W))$ eleme, vagyis olyan folytonos lineáris leképezés, amely minden V-beli elemhez egy $V \to W$ folytonos lineáris leképezést rendel. $\mathcal{L}in(V, \mathcal{L}in(V, W))$ szintén normált tér, így $\mathrm{D}^2F : V \to \mathcal{L}in(V, \mathcal{L}in(V, W))$ differenciálhatóságát is vizsgálhatjuk és nyilvánvaló, hogy a deriváltja egy adott pontban $\mathcal{L}in(V, \mathcal{L}in(V, \mathcal{L}in(V, W)))$ eleme lesz. Látszik, hogy egyre bonyolultabb (legalábbis annak tűnő) lineáris leképezésekhez jutunk, ahogy egyre többször differenciálható függvényeket akarunk értelmezni.

Könnyen áthidalhatjuk ezt a nehézséget korábbi tanulmányainkban megismert azonosításokkal (Analízis III.B.10.12.). Ha $n \in \mathbb{N}$, akkor $\mathcal{L}in^n(V^n, W)$ jelöli a $V^n \to W$ folytonos n-lineáris leképezések normált terét, és van egy

$$\mathcal{L}in^n\left(V^n, W\right) \equiv \mathcal{L}in\left(V, \mathcal{L}in^{n-1}\left(V^{n-1}, W\right)\right),$$

$$R \equiv \big(x_n \mapsto R(\cdot, \ldots, \cdot, x_n)\big)$$

azonosítás (lineáris izometrikus bijekció).

Speciálisan $\mathcal{L}in^2(V^2, W) \equiv \mathcal{L}in(V, \mathcal{L}in(V, W))$, tehát egy $F : V \rightarrowtail W$ függvény második deriváltját egy pontban tekinthetjük úgy, mint egy $V \times V \to W$ folytonos bilineáris leképezést. Közelebbről,

$$\mathrm{D}^2F(a)(h, v) \equiv (\mathrm{D}^2F(a)v)h,$$

ahol a jobb oldalon az eredeti értelmezés szerinti $\mathcal{L}in(V, \mathcal{L}in(V, W))$ leképezés áll. Ha F kétszer diferenciálható a-ban, azaz $\mathrm{D}F : V \rightarrowtail \mathcal{L}in(V, W)$ differenciálható a-ban, akkor az 1.11.4. feladat szerint minden $v \in V$ esetén $V \rightarrowtail W$, $x \mapsto \mathrm{D}F(x)v$ differenciálható a-ban, és deriváltja a $V \to W$, $h \mapsto (\mathrm{D}^2F(a)v)h \equiv \mathrm{D}^2F(a)(h, v)$ lineáris leképezés. Tehát

$$\mathrm{D}F(a + h)v - \mathrm{D}F(a)v = \mathrm{D}^2F(a)(h, v) + o_v(h). \qquad (*)$$

Az F harmadik deriváltja egy pontban (ha létezik) $\mathcal{L}in(V, \mathcal{L}in^2(V^2, W))$ eleme, amelyet tekinthetünk $V \times V \times V \to W$ folytonos trilineáris leképezésnek. Ennek megfelelően az n-ik derivált egy pontban $V^n \to W$ folytonos n-lineáris leképezés lesz.

8.3. Most már könnyen értelmezhetjük (szokás szerint rekurzióval) a normált terek közötti, kettőnél többször differenciálható függvényeket is.

1. Definíció *Legyen $F : V \rightarrowtail W$ függvény. Minden $n \in \mathbb{N}_0$ esetén értelmezzük a $\mathrm{D}^nF : V \rightarrowtail \mathcal{L}in^n(V^n, W)$ függvényt a következő módon:*

(i) $\mathrm{D}^0F := F$,

(ii) $\mathrm{D}^nF := \mathrm{D}(\mathrm{D}^{n-1}F) : V \rightarrowtail \mathcal{L}in(V, \mathcal{L}in^{n-1}(V^{n-1}, W) \equiv \mathcal{L}in^n(V^n, W)$.

A D^nF függvényt az F ***n-edik derivált függvényének*** *(**n-edik deriváltjának**) hívjuk.*

Az iménti függvénysorozat nyilván jól értelmezett, vagyis bármely $V \rightarrowtail W$ függvénynek létezik akárhányadik deriváltja, viszont az persze előfordulhat, hogy egy bizonyos természetes számra, és akkor már minden annál nagyobb n-re is $\mathrm{D}^n F$ az üres függvény (értelmezési tartománya az üres halmaz).

2. Definíció *Legyen $F : V \rightarrowtail W$ függvény, $n \in \mathbb{N}$. Azt mondjuk, hogy az F függvény*

*(i) n-**szer differenciálható az** $a \in V$ **pontban**, ha $a \in \mathrm{Dom}(\mathrm{D}^n F)$;*

*(ii) n-**szer differenciálható a** $H \subset V$ **halmazon**, ha $H \subset \mathrm{Dom}(\mathrm{D}^n F)$;*

*(iii) n-**szer differenciálható**, ha F értelmezési tartománya nyílt és $\mathrm{Dom}(\mathrm{D}^n F) = \mathrm{Dom}(F)$;*

*(iv) n-**szer folytonosan differenciálható a** H **halmazon**, ha $H \subset \mathrm{Dom}(\mathrm{D}^n F)$ és $\mathrm{D}^n F$ folytonos H-n;*

*(v) n-**szer folytonosan differenciálható**, ha n-szer folytonosan differenciálható az értelmezési tartományán;*

*(vi) **végtelenszer differenciálható egy pontban (egy halmazon)**, ha minden $n \in \mathbb{N}$ esetén n-szer differenciálható ebben a pontban (ezen a halmazon);*

*(vii) **végtelenszer differenciálható**, ha az értelmezési tartománya nyílt, és minden $n \in \mathbb{N}$ esetén n-szer differenciálható.*

Rögtön megjegyezzük, hogy az iménti definíció szerinti "egyszer" illetve "kétszer differenciálhatóság" pontosan a korábban bevezetett differenciálhatósággal illetve kétszer differenciálhatósággal egyezik meg. Az első definíció kapcsán már említettük, hogy (abban az értelemben) bármely $V \rightarrowtail W$ függvénynek van akárhányadik deriváltja, világos azonban, hogy ez nem jelenti azt, hogy az adott függvény akár egyszer is differenciálható, hiszen ez a létező derivált az értelmezésünk szerint lehet üres függvény is.

8.4. (i) Egy $B : U \times V \to W$ folytonos bilineáris leképezés differenciálható, deriváltja az

$$U \times V \to \mathcal{L}in(U \times V, W), \quad (x, y) \mapsto B(\mathrm{pr}_U, y) + B(x, \mathrm{pr}_V)$$

leképezés (lásd 1.11.6. feladat). Mivel $(h, k) \in U \times V$ és $(u, v) \in U \times V$ esetén

$$B(h, y + v) + B(x + u, k) - \big(B(h, y) + B(x, k) = B(h, v) + B(u, k),$$

megállapíthatjuk, hogy B kétszer differenciálható, és a deriváltja minden pontban az $(U \times V) \times (U \times V) \to W$, $((h, k)(u, v)) \mapsto B(h, v) + B(u, k)$ bilineáris leképezés (vagyis a második deriváltja konstans).

Tehát B végtelenszer differenciálható, és a 2-nél magasabb rendű deriváltjai nullák. Ezt úgy is beláthattuk volna, hogy megmutatjuk, az első deriváltfüggtvénye lineáris és folytonos.

(ii) Egy $R : \bigtimes_{i=1}^{n} V_i \to W$ folytonos n-lineáris leképezés deriváltja a 4.8.2-beli formula alapján a

$$\bigtimes_{i=1}^{n} V_i \to \mathcal{L}in(\bigtimes_{i=1}^{n} V_i, W), \quad (x_1, \ldots, x_n) \mapsto \sum_{i=1}^{n} R(x_1, \ldots, x_{i-1}, \mathrm{pr}_i, x_{i+1}, \ldots, x_n)$$

leképezés. Ezért R kétszer differenciálható, második deriváltja az $a = (a_1, \ldots, a_n)$ pontban a

$$\begin{aligned} &\left(\bigtimes_{i=1}^{n} V_i\right) \times \left(\bigtimes_{i=1}^{n} V_i\right) \to W, \\ &\big((h_1, \ldots, h_n), (v_1, \ldots, v_n)\big) \mapsto \\ &\qquad \sum_{i=1}^{n} \sum_{k=i+1}^{n} R(a_1, \ldots, a_{i-1}, h_i, a_{i+1}, \ldots, a_{k-1}, v_k, a_{k+1}, \ldots, a_n)+ \\ &\qquad +\sum_{i=1}^{n} \sum_{k=i+1}^{n} R(a_1, \ldots, a_{i-1}, v_i, a_{i+1}, \ldots, a_{k-1}, h_k, a_{k+1}, \ldots, a_n) \end{aligned}$$

bilineáris leképezés.

Ennek alapján látjuk, hogy R végtelenszer differenciálható, és n-nél magasabb rendű deriváltjai nullák.

(iii) Legyen $R : V^n \to W$ szimmetrikus, folytonos n-lineáris leképezés és tekintsük a $V \to W$, $x \to R(x, x, \ldots, x)$ leképezést (lásd 4.8.3.). Ez is végtelenszer differenciálható, második deriváltja az a pontban a

$$V^2 \to W, \quad (h_1, h_2) \to n(n-1)R(h_1, h_2, a, \ldots, a)$$

bilineáris leképezés, a k-adik deriváltja $(1 \leq k \leq n)$ a-ban a

$$V^k \to W, \quad (h_1, \ldots, h_k) \mapsto n(n-1)\ldots(n-k+1)R(h_1, \ldots, h_k, a, \ldots, a)$$

k-lineáris leképezés, és n-nél magasabb rendű deriváltjai nullák.

8.5. Az előbbi példákban szereplő függvények második deriváltja minden pontban szimmetrikus bilineáris leképezés. Azt vizsgáljuk, általában is érvényes-e ez.

Legyen az $F : V \rightarrowtail W$ függvény kétszer differenciálható az a pontban. Ekkor van olyan $r > 0$, hogy $G_r(a) \subset \mathrm{Dom}(F)$ és F differenciálható $G_r(a)$-n.

Ha $v, h \in V \setminus \{0\}$ és $t \in \left]0, \frac{r}{\|v\| + \|h\|}\right[$, akkor F differenciálhatóságának és DF a-beli differenciálhatóságának definíciója és a 8.2. $(*)$ öszefüggés szerint

$$\mathrm{D}^2F(a)(h, v) = \frac{\mathrm{D}F(a + th)v - \mathrm{D}F(a)v}{t} - \frac{o_{\mathrm{D}F}(a, th)v}{t},$$

$$\mathrm{D}F(a+th)v = \frac{F(a+th+tv)-F(a+th)}{t} - \frac{o_F(a+th,tv)}{t},$$

$$\mathrm{D}F(a)v = \frac{F(a+tv)-F(a)}{t} - \frac{o_F(a,tv)}{t},$$

ahol $o_{\mathrm{D}F}(a,\cdot) : V \rightarrowtail \mathcal{L}in(V,W)$ és $o_F(a,\cdot), o_F(a+th,\cdot) : V \rightarrowtail W$ kis ordó függvények.

A két utóbbi egyenlőséget az elsőbe behelyettesítve

$$\mathrm{D}^2F(a)(h,v) = \Delta_{h,v}(t) + \frac{\frac{o_F(a,tv)}{t} - \frac{o_F(a+th,tv)}{t}}{t} - \frac{o_{\mathrm{D}F}(a,th)v}{t} \qquad (*)$$

adódik, ahol

$$\Delta_{h,v} : \left]0, \frac{r}{\|h\|+\|v\|}\right[\to W, \quad t \mapsto \frac{1}{t^2}\Big(F(a+t(h+v))-F(a+th)-F(a+tv)+F(a)\Big).$$

A $(*)$ egyenlőség jobb oldalán a harmadik tag nullához tart $t \to 0$ esetén, és a második tag számlálójának első tagja úgyszintén. Azonban nem látszik, hogy ugyanekkor a szóban forgó számláló második tagja nullához tartana; az pedig végképp nem nyilvánvaló, hogy az egész második tag nullához tartana $t \to 0$ esetén. Ha viszont ez igaz lenne, tehát

$$\lim_{t\to 0} \Delta_{h,v}(t) = \mathrm{D}^2F(a)(h,v) \qquad (**)$$

teljesülne, akkor $\Delta_{h,v} = \Delta_{v,h}$ miatt $\mathrm{D}^2F(a)$ szimmetrikus lenne.

Állítás (***Young-tétel***). *Ha $F : V \rightarrowtail W$ az a pontban kétszer differenciálható, akkor $\mathrm{D}^2F(a) \in \mathcal{L}in^2(V^2,W)$ szimmetrikus.*

Bizonyítás Használjuk az előző jelöléseket.

Rögzítsük tetszőlegesen az $h,v \in V \setminus \{0\}$ vektorokat, és minden $0 < t < \frac{r}{\|h\|+\|v\|}$ esetén legyen $g_t : V \rightarrowtail W$ az a függvény, amely az

$$\left\{x \in V \;\middle|\; a+tx \in \overset{\circ}{\mathrm{Dom}}(F),\ a+t(h+x) \in \overset{\circ}{\mathrm{Dom}}(F)\right\}$$

halmazon van értelmezve az

$$x \mapsto F(a+t(h+x)) - F(a+tx) - t^2\mathrm{D}^2F(a)(h,x)$$

utasítással.

Minden t-re g_t differenciálható függvény és az is könnyen látható, hogy a $[0, v]$ szakasz része g_t értelmezési tartományának, valamint

$$\Delta_{h,v}(t) - \mathrm{D}^2F(a)(h, v) = \frac{1}{t^2}(g_t(v) - g_t(0)). \tag{1}$$

Alkalmazhatjuk a g_t függvényre a $[0, v]$ szakaszon az általános középértéktételt, miszerint létezik olyan $z_t \in]0, v[$, hogy

$$\begin{aligned}\|g_t(v) - g_t(0)\| \leq \|\mathrm{D}g_t(z_t)\|\ \|v\| = \\ \|t\mathrm{D}F(a + t(h + z_t)) - t\mathrm{D}F(a + tz_t) - t^2\mathrm{D}^2F(a)(h, \cdot)\|\ \|v\|.\end{aligned} \tag{2}$$

Mivel $\mathrm{D}F$ az a pontban differenciálható, van olyan $\eta : V \rightarrowtail \mathcal{L}in(V, W)$ függvény, amely a $0 \in V$ egy környezetén értelmezett, folytonos a 0-ban, $\eta(0) = 0$, és minden $x \in \mathrm{Dom}(\eta)$ vektorra $a + x \in \mathrm{Dom}(\mathrm{D}F)$, valamint

$$\mathrm{D}F(a + x) = \mathrm{D}F(a) + \mathrm{D}^2F(a)(x, \cdot) + \|x\|\eta(x) \tag{3}$$

teljesül.

Legyen $\epsilon > 0$ tetszőleges. Ekkor η tulajdonságai szerint van olyan $0 < \delta_\epsilon < r$, hogy $\|x\| < \delta_\epsilon$ esetén $x \in \mathrm{Dom}(\eta)$ és

$$\|\eta(x)\| \leq \frac{\epsilon}{2\|v\|(\|h\| + \|v\|)}.$$

Ekkor $0 < t < \dfrac{\delta_\epsilon}{\|h\| + \|v\|}$ esetén (3)-ba x helyére $t(h + z_t)$-t, illetve tz_t-t behelyettesítve és a kapottakat (2)-be beírva kapjuk, hogy

$$\begin{aligned}
&\|g_t(v) - g_t(0)\| \leq \\
&\quad \leq \|v\|\ \big\|t\mathrm{D}F(a) + t^2\mathrm{D}^2F(a)(h + z_t, \cdot) + t|t|\ \|h + z_t\|\eta(t(h + z_t)) - \\
&\qquad - t\mathrm{D}F(a) - t^2\mathrm{D}^2F(a)(z_t, \cdot) - t|t|\ \|z_t\|\eta(tz_t) - t^2\mathrm{D}^2F(a)(h, \cdot)\big\| = \\
&\quad = \|v\|t^2\big\|\ \|h + z_t\|\eta(t(h + z_t)) - \|z_t\|\eta(tz_t)\big\| \leq \\
&\quad \leq \|v\|t^2\Big(\big(\|h\| + \|z_t\|\big)\|\eta(t(h + z_t))\| + \|z_t\|\ \|\eta(tz_t)\|\Big) \leq \\
&\quad \leq \|v\|t^2\left((\|h\| + \|v\|)\frac{\epsilon}{2\|v\|(\|h\| + \|v\|)} + \|v\|\frac{\epsilon}{2\|v\|(\|h\| + \|v\|)}\right) \leq \\
&\quad \leq t^2\epsilon,
\end{aligned}$$

mert $\|t(h+z_t)\| < \delta_\epsilon$ és $\|tz_t\| < \delta_e$. Ez azt jelenti, hogy minden $0 < t < \dfrac{\delta_\epsilon}{\|h\| + \|v\|}$ esetén (1) miatt

$$\|\Delta_{h,v}(t) - \mathrm{D}^2F(a)(h, v)\| \leq \frac{1}{t^2}\|g_t(v) - g_t(0)\| \leq \epsilon,$$

tehát $\lim_{t\to 0} \Delta_{h,v}(t) = \mathrm{D}^2F(a)(h,v)$, és éppen ezt akartuk belátni.

Megjegyzés Igaz az is, hogy ha $F : V \rightarrowtail W$ n-szer differenciálható az a pontban, akkor a $\mathrm{D}^nF(a) \in \mathcal{L}in^n(V^n, W)$ n-lineáris leképezés szimmetrikus (azaz szimmetrikus bármely két változójában, ha a többit rögzítjük). Ennek bizonyítását itt nem közöljük, csak utalunk arra, hogy n szerinti teljes indukcióval végezhető el.

8.6. Nézzük meg mit jelent az iménti állítás egy $f : \mathbb{K}^N \rightarrowtail \mathbb{K}$ kétszer differenciálható függvényre nézve. A 4.7-ben láttuk, hogy a standard bázisvektorok menti deriváltak gyakorlatilag azonosak a megfelelő parciális deriváltakkal. A $\mathrm{D}^2f(a) \in \mathcal{L}in^2(\mathbb{K}^N, \mathbb{K})$ bilineáris leképezés mátrixa

$$(\partial_i\partial_j f(a) \mid i,j = 1,\dots,N).$$

$\mathrm{D}^2f(a)$ szimmetrikussága azt jelenti, hogy ez a mátrix szimmetrikus, azaz

$$\partial_i\partial_j f = \partial_j\partial_i f \qquad (i,j = 1,\dots,N).$$

Ezt úgy szoktuk kifejezni, hogy az "egyes változók szerinti parciális differenciálás sorrendje fölcserélhető", ha a függvény kétszer differenciálható.

Felhívjuk a figyelmet, hogy előfordulhat, a függvény második parciális deriváltjai léteznek anélkül, hogy a függvény kétszer differenciálható volna; ekkor a különböző sorrendű parciális deriváltak különbözhetnek egymástól.

Például az

$$f : \mathbb{R}^2 \to \mathbb{R}, \quad (x,y) \mapsto \begin{cases} xy\frac{x^2-y^2}{x^2+y^2} & \text{ha } x^2+y^2 \neq 0, \\ 0 & \text{ha } x^2+y^2 = 0 \end{cases}$$

függvényre

$$\frac{\partial f(x,y)}{\partial x} = \begin{cases} \frac{yx^4+4x^2y^3-y^5}{(x^2+y^2)^2} & \text{ha } x^2+y^2 \neq 0, \\ 0 & \text{ha } x^2+y^2 = 0, \end{cases}$$

ezért

$$\left.\frac{\partial^2 f(x,y)}{\partial y \partial x}\right|_{x=0,y=0} = -1,$$

és mivel x és y azonos szerepet játszik egy előjeltől eltekintve,

$$\left.\frac{\partial^2 f(x,y)}{\partial x \partial y}\right|_{x=0,y=0} = 1.$$

8.7. Többször differenciálható függvényekre az inverzfüggvény-tétel és az implicitfüggvény-tétel erősebb formában is igaz. Nevezetesen, ha az inverzfüggvény-tétel (implicitfüggvény-tétel) feltételei teljesülnek, és a szóban forgó függvény n-szer folytonosan differenciálható, akkor az inverz függvény (implicit függvény) is n-szer folytonosan differenciálható. Ezt az inverz függvényre mutatjuk meg, abból könnyen származtatható a másik eset.

Használjuk a 6.1. jelöléseit, és legyen F n-szer folytonosan differenciálható. Ekkor $G := (F|_N)^{-1}$ folytonosan differenciálható, és deriváltja az $y \mapsto \mathrm{D}F(G(y))^{-1}$ függvény, amely folytonosan differenciálható függvények kompozíciójaként írható fel,

$$y \mapsto G(y) \mapsto \mathrm{D}F(G(y)) \mapsto (\mathrm{D}F(G(y))^{-1} = \mathrm{D}G(y),$$

ezért folytonosan differenciálható. Ennek deriváltja – azaz G második deriváltja (lásd 1.8. (v)) – az

$$y \mapsto -\mathrm{D}(F(G(y))^{-1}\Big(\mathrm{D}^2F(G(y))(\mathrm{D}G(y)\cdot, \cdot)\Big)\mathrm{D}F(G(y))^{-1}$$

függvény, amely ismét csak folytonosan differenciálható függvények kompozíciója. Látjuk, hogy G második deriváltjában F első és második deriváltja, valamint G első derviáltja jelenik meg; világos a formulából, hogy G harmadik deriváltjában F első, második és harmadik deriváltja, valamint G első és második deriváltja jelenik meg. Ebből indukcióval könnyen beláthatjuk, hogy G n-szer folytonosan differenciálható.

8.8. Feladatok

1. Bizonyítsuk be, hogy $\mathcal{L}in(V) \to \mathcal{L}in(V)$, $X \mapsto X^n$ végtelenszer differenciálható, és adjuk meg a deriváltjait.

2. Legyen V teljes normált tér (Banach-tér). Igazoljuk, hogy $\mathcal{L}in(V) \rightarrowtail \mathcal{L}in(V)$, $X \mapsto X^{-1}$ végtelenszer differenciálható, és adjuk meg a deriváltjait.

3. Legyen V teljes normált tér és $A \in \mathcal{L}in(V)$. Mutassuk meg, hogy $\mathbb{R} \to \mathcal{L}in(V)$, $t \mapsto e^{tA}$ végtelenszer differenciálható, és adjuk meg a deriváltjait.

4. Határozzuk meg a következő $\mathbb{R}^2 \rightarrowtail \mathbb{R}$ illetve $\mathbb{R}^3 \rightarrowtail \mathbb{R}$ függvények megjelölt magasabb rendű deriváltjait:

(i) $(x, y) \mapsto \sin(x^2 + y^2)$ (harmad rendű),
(ii) $(x, y) \mapsto \cos(x)\sin(y)$ (harmad rendű),
(iii) $(x, y, z) \mapsto xyz$ (negyed rendű),

5. Legyen $f, g : \mathbb{R} \rightarrowtail \mathbb{R}$ végtelenszer differenciálható. Határozzuk meg az alábbi $\mathbb{R}^2 \rightarrowtail \mathbb{R}$ illetve $\mathbb{R}^3 \rightarrowtail \mathbb{R}$ függvények n-ik ($n \in \mathbb{N}$) deriváltját:

(i) $(x, y) \mapsto f(x)g(y)$, (ii) $(x, y) \mapsto f\left(\sqrt{x^2 + y^2}\right)$,
(iii) $(x, y, z) \mapsto f(x + y + z)$, (iv) $(x, y, z) \mapsto f(xyz)$.

6. Legyen $F : V \rightarrowtail W$ kétszer differenciálható a-ban. Mutassuk meg, hogy $\mathrm{Ran}(\mathrm{D}^2F(a)) \subset \overline{\mathrm{Ran}\mathrm{D}F(a)}$. (Útmutatás: vegyünk iránymenti deriváltakat:

$$\mathrm{D}^2F(a)(h, v) = \lim_{t\to 0} \frac{\mathrm{D}F(a + th)v - \mathrm{D}F(a)v}{t}.)$$

9. Taylor-formula

9.1. Állítás *(**Taylor-formula, valós**) Legyen V valós normált tér és $F : V \rightarrowtail \mathbb{R}$ $(n+1)$-szer differenciálható függvény, $a, h \in V$. Ha az $[a, a+h]$ szakasz benne van az F értelmezési tartományában, akkor létezik olyan $\theta \in [0,1]$, hogy*

$$F(a+h) = F(a) + \mathrm{D}F(a)h + \frac{1}{2!}\mathrm{D}^2F(a)(h,h) + \dots + \frac{1}{n!}\mathrm{D}^nF(a)(h,\dots,h) + \\ + \frac{1}{(n+1)!}\mathrm{D}^{n+1}F(a+\theta h)(h,\dots,h,h).$$

Bizonyítás A

$$\phi : [0,1] \to \mathbb{R}, \quad t \mapsto F(a+th)$$

függvényre alkalmazhatjuk az A.9.1. Taylor-formulát a $[0,1]$ intervallumon: létezik olyan $\theta \in]0,1[$, hogy

$$\phi(1) = \phi(0) + \phi'(0) + \frac{1}{2!}\phi''(0) + \dots \frac{1}{n!}\phi^{(n)}(0) + \frac{1}{(n+1)!}\phi^{(n+1)}(\theta). \qquad (*)$$

Nyilvánvaló, hogy

$$\phi(1) = F(a+h), \qquad \phi(0) = F(a),$$

valamint közvetett függvényként egyre többször differenciálva ϕ-t, megkaphatjuk a deriváltjaira a következő összefüggéseket:

$$\begin{aligned} \phi'(0) &= \mathrm{D}F(a)h, \\ \phi''(0) &= \mathrm{D}^2F(a)(h,h), \\ &\cdot \\ &\cdot \\ &\cdot \\ \phi^{(n)}(0) &= \mathrm{D}^nF(a)(h,\dots,h), \\ \phi^{(n+1)}(\theta) &= \mathrm{D}^{n+1}F(a+\theta h)(h,\dots,h,h). \end{aligned}$$

Mindezeket behelyettesítve a $(*)$ összefüggésbe, pontosan a bizonyítandó formulát kapjuk. ∎

Megjegyezzük, hogy a Taylor formulában a jobb oldal utolsó tagját **maradéktagnak** szokás hívni.

9.2. Egy $f : \mathbb{R}^n \rightarrowtail \mathbb{R}$ $(n+1)$-szer differenciálható függvény esetén a Taylor-formulát a parciális deriváltakkal

$$f(a+h) = f(a) + \sum_{i=1}^{n} \partial_i f(a) h_i + \frac{1}{2!} \sum_{i,j=1}^{n} \partial_i \partial_j f(a) h_i h_j + \dots$$
$$+ \frac{1}{n!} \sum_{i,j,\dots,k} \partial_i \partial_j \dots \partial_k f(a) h_i h_j \dots h_k + \text{maradéktag}$$

formába írhatjuk, ahol az utolsó összeg a jobb oldalon n darab index szerinti összegzést jelent, és természetesen h_i stb. a $h \in \mathbb{R}^n$ elem i-edik komponensét jelöli.

A jobb oldal, a maradéktagot leszámítva, h-ban egy legfeljebb n-ed fokú **multipolinom**, vagyis egy olyan $\mathbb{R}^n \to \mathbb{R}$ függvény, amely a változókomponensek hatványainak szorzatából képzett lineáris kombináció.

Ezt a függvény a körüli **n-ed fokú Taylor-multipolinomjának** nevezzük. Ha bevezetjük az $x := a + h$ jelölést, akkor $h = x - a$, és az $\mathbb{R} \rightarrowtail \mathbb{R}$ függvényekre megismert A.9.2. alakhoz hasonlót kapunk.

9.3. A Taylor-formulát egy kicsit más formában felírhatjuk nem csak valós, hanem tetszőleges normált tér értékű függvényekre is. Itt a maradéktag konkrét alakja helyett csak egy becslés jelenik meg.

Állítás (***Taylor formula, általános***) *Legyen $F : V \rightarrowtail W$ n-szer differenciálható az a pontban. Ekkor $h \in V$, $a + h \in \mathrm{Dom}(F)$ esetén*

$$F(a+h) = F(a) + \mathrm{D}F(a)h + \frac{1}{2!}\mathrm{D}^2 F(a)(h,h) + \dots + \frac{1}{n!}\mathrm{D}^n F(a)(h,\dots,h) + o(\|h\|^n),$$

ahol $o : \mathbb{R} \rightarrowtail W$ kis ordó függvény.

Bizonyítás Az egyszerűbb írásmód kedvéért bevezetjük a

$$h^m := (h,\dots,h) \in V^m \qquad (m = 1,\dots,n)$$

jelölést.

A bizonyítást n szerinti teljes indukcióval végezzük. $n = 1$ esetben azt állítjuk, hogy

$$F(a+h) = F(a) + \mathrm{D}F(a)h + o(\|h\|),$$

ez pedig éppen az F függvény a pontbeli differenciálhatóságának a definíciója.

Tegyük fel, hogy az állítás $1 \leq n - 1$ esetén teljesül, és igazoljuk n-re.

Legyen tehát F n-szer differenciálható a-ban; azt kell megmutatnunk, hogy a

$$G : V \rightarrowtail W, \quad h \mapsto F(a+h) - F(a) - \sum_{m=1}^{n} \frac{1}{m!} \mathrm{D}^m F(a) h^m$$

függvény elosztva $\|h\|^n$-nel a nullához tart, miközben h tart a nullához.

Nyilvánvaló, hogy $\mathrm{Dom}(G) = \mathrm{Dom}(F) - a$, tehát $0 \in V$ a $\mathrm{Dom}(G)$ belső pontja. F n-szer differenciálható a-ban, ezért G n-szer differenciálható a 0-ban, és így G $(n-1)$-szer differenciálható a 0 egy környezetében. A $h \to \mathrm{D}^m F(a)h^m$ szimmetrikus m-lineáris leképezés k-ik deriváltjára vonatkozó 8.4.(iii) képlet alapján $h \in \mathrm{Dom}(G)$ és $0 < k \leq n-1$ esetén azt kapjuk, hogy G-nek a k-adik deriváltja h-ban egy $v^k \in V^k$ elemre hatva

$$\mathrm{D}^k G(h)v^k = \mathrm{D}^k F(a+h)v^k - \sum_{m=k}^{n} \frac{1}{m!} m(m-1)\dots(m-k+1)\mathrm{D}^m F(a)(h^{m-k}, v^k),$$

ahol $(h^{m-k}, v^k) \in V^{m-k} \times V^k \equiv V^m$. Ezt tovább alakítva, az $i := m-k$ jelöléssel azt kapjuk, hogy

$$\mathrm{D}^k G(h)v^k = \mathrm{D}^k F(a+h)v^k - \sum_{i=0}^{n-k} \frac{(i+k)(i+k-1)\dots(i+1)}{(i+k)!}\mathrm{D}^{i+k} F(a)(h^i, v^k) =$$
$$= \mathrm{D}^k F(a+h)v^k - \mathrm{D}^k F(a)v^k - \sum_{i=1}^{n-k} \frac{1}{i!}\mathrm{D}^{i+k} F(a)(h^i, v^k),$$

azaz

$$\mathrm{D}^k G(h) = \mathrm{D}^k F(a+h) - \mathrm{D}^k F(a) - \sum_{i=1}^{n-k} \frac{1}{i!}\mathrm{D}^{i+k} F(a)(h^i, \cdot),$$

ahol az utolsó szimbólum azt a $V^k \to W$ szimmetrikus k-lineáris leképezést jelenti, amelyet úgy kapunk, hogy $\mathrm{D}^{i+k}F(a)$ első i változója helyébe beírjuk h^i-t.

Ebből látszik, hogy $0 < k \leq n-1$ esetén $\mathrm{D}^k G(0) = 0$. Továbbá

$$\mathrm{D}^{n-1}G(h) = \mathrm{D}^{n-1}G(h) - \mathrm{D}^{n-1}G(0) = \mathrm{D}^{n-1}F(a+h) - \mathrm{D}^{n-1}F(a) - \mathrm{D}^n F(a)(h, \cdot)$$
$$= o_n(h),$$

ahol $o_n : V \rightarrowtail W$ kis ordó függvény, amely F-nek az a-beli n-szer differenciálhatósága miatt létezik. Ez viszont azt jelenti, hogy $D^n G(0) = 0$ is teljesül.

Az indukciós feltevés DG-re, mint a nullában $(n-1)$-szer differenciálható függvényre azt adja, hogy

$$\mathrm{D}G(h) = \mathrm{D}G(0) + \sum_{j=1}^{n-1} \frac{1}{j!}\big(\mathrm{D}^j(\mathrm{D}G)(0)\big)h^j + o(\|h\|^{n-1}) =$$
$$= o(\|h\|^{n-1}) = \|h\|^{n-1}\eta(\|h\|^{n-1}),$$

ahol $\eta : \mathbb{R} \rightarrowtail \mathcal{L}in(V, W)$ olyan függvény, amelyre $\lim\limits_{t\to 0} \eta(t) = 0$.

Ezért a középértéktétel szerint, ha $h \in V$ olyan, hogy a $[0, h]$ szakaszon G differenciálható, akkor – lévén a G definíciója szerint $G(0) = 0$ –

$$\left\| F(a+h) - F(a) - \sum_{m=1}^{n} \frac{1}{m!} \mathrm{D}^m F(a) h^m \right\| = \|G(h) - G(0)\| \leq$$

$$\leq \|h\| \sup_{\xi \in [0,h]} \|\mathrm{D}G(\xi)\| = \|h\|^n \sup_{\xi \in [0,h]} \left\| \eta(\|\xi\|^{n-1}) \right\|.$$

Mivel $\lim\limits_{h \to 0} \sup\limits_{\xi \in [0,h]} \left\| \eta(\|\xi\|^{n-1}) \right| = 0$, az előzőek alapján

$$\lim_{h \to 0} \frac{1}{\|h\|^n} \left(F(a+h) - F(a) - \sum_{m=1}^{n} \frac{1}{m!} \mathrm{D}^m F(a) h^m \right) = 0,$$

és ezt akartuk bizonyítani.

9.4. Feladatok

1. Írjuk fel a következő $\mathbb{R}^2 \rightarrowtail \mathbb{R}$ függvények harmadrendű Taylor-multipolinomját az adott pont körül:

(i) $(x, y) \mapsto \sqrt{x^2 + y^2}$, $(1, 0)$, (ii) $(x, y) \mapsto \sin x \cos y$, $(0, 0)$,

(iii) $(x, y) \mapsto e^{x+y^2}$, $(0, 1)$, (iv) $(x, y) \mapsto \dfrac{x-y}{x+y}$, $(1, 1)$.

1. Írjuk fel a következő $\mathbb{R}^3 \to \mathbb{R}$ függvények harmadrendű Taylor-multipolinomját az adott pont körül:

(i) $(x, y, z) \mapsto e^z \sqrt{x^2 + y^2}$, $(1, 1, 0)$, (ii) $(x, y) \mapsto \sin(x + zy)$, $(0, 0, 0)$.

10. Szélsőérték, feltételes szélsőérték

Ebben a fejezetben V valós normált teret jelöl. Emlékeztetünk, hogy minden komplex vektortér egyben valós vektortér is, de egy függvény valós differenciálhatósága gyengébb tulajdonság, mint a komplex differenciálhatósága.

10.1. Ahhoz hasonlóan, mint A.2.1-ben, azt mondjuk, hogy az $F : V \rightarrowtail \mathbb{R}$ függvénynek

(i) **maximuma (minimuma)** van az $a \in \mathrm{Dom}(F)$ pontban, ha bármely $x \in \mathrm{Dom}(F)$ esetén

$$F(x) \leq F(a) \qquad \big(F(x) \geq F(a)\big).$$

(ii) **lokális maximuma (minimuma)** van az $a \in \mathrm{Dom}(F)$ pontban, ha létezik olyan $G(a)$ környezete a-nak, hogy bármely $x \in G(a) \cap \mathrm{Dom}(F)$ esetén

$$F(x) \leq F(a) \qquad \Big(F(x) \geq F(a)\Big).$$

Szigorú (lokális) szélsőértékről (más szóval extrémumról) beszélünk, ha a fenti pontokban a $\leq$ és $\geq$ relációk helyett a szigorúbb $>$ illetve $<$ egyenlőtlenségek teljesülnek, persze $x \neq a$ esetén.

A jegyzet első felében már láttuk, hogy egy differenciálható valós-valós függvénynek csak akkor lehet szélsőértéke egy pontban, ha ott a deriváltja nulla. Ezt a tényt általánosíthatjuk normált téren értelmezett függvényekre.

Állítás *Ha az $F : V \rightarrowtail \mathbb{R}$ függvény $0 \neq v \in V$ esetén a v irány mentén differenciálható az a pontban és ott lokális szélsőértéke van, akkor $\mathrm{D}_v F(a) = 0$.*

Bizonyítás Mivel F differenciálható a-ban a v irány mentén, létezik olyan $r > 0$, hogy értelmezhetjük a következő függvényt:

$$\phi :]-r, r[\to \mathbb{R}, \quad t \mapsto F(a + tv).$$

Nyilvánvaló, hogy ϕ differenciálható a nullában, valamint

$$\phi(0) = F(a), \quad \phi'(0) = \mathrm{D}_v F(a),$$

és ϕ-nek lokális szélsőértéke van a nullában. Így az A.2.5. állítás értelmében befejeztük a bizonyítást.

Következmény Ha az $F : V \rightarrowtail \mathbb{R}$ függvény differenciálható az a pontban és ott lokális szélsőértéke van, akkor $\mathrm{D}F(a) = 0 \in \mathcal{L}in(V, \mathbb{R})$, hiszen ekkor bármilyen $0 \neq v \in V$ esetén $\mathrm{D}F(a)v = \mathrm{D}_v F(a) = 0$.

Vigyázzunk arra, hogy az állítás fordítottja nem igaz (hiszen már valós-valós függvények esetében sem volt az), vagyis $\mathrm{D}F(a) = 0$ nem jelenti feltétlenül azt, hogy F-nek a-ban lokális szélsőértéke van.

10.2. Az előző pontban a szélsőértéknek egy szükséges feltételét láttuk be, de ez alapján azt sem tudhatjuk, hogy az illető pontban maximuma vagy minimuma van (lehet) az adott függvénynek. Hogy egy elégséges feltételt fogalmazhassunk meg a szélsőértékre vonatkozóan és a fenti kérdésre is választ adjunk, be kell vezetnünk néhány fogalmat.

Definíció *Egy $B : V \times V \to \mathbb{R}$ bilineáris leképezésről azt mondjuk, hogy*

(i) ***pozitív (negatív)*** *definit, ha minden $0 \neq v \in V$ esetén*

$$B(v, v) > 0 \qquad \big(B(v, v) < 0\big),$$

(ii) ***szigorúan pozitív (negatív) definit****, ha*

$$\inf_{\|v\|=1} B(v, v) > 0 \qquad \big(\sup_{\|v\|=1} B(v, v) < 0\big),$$

(iii) ***pozitív (negatív) szemidefinit****, ha minden $v \in V$ esetén*

$$B(v, v) \geq 0 \qquad \big(B(v, v) \leq 0\big),$$

(iv) ***indefinit****, ha létezik $u, v \in V$, úgy, hogy*

$$B(u, u) > 0 \quad \textit{és} \quad B(v, v) < 0.$$

A szigorú pozitív (negatív) definitség egyenértékű azzal, hogy létezik $\beta > 0$ úgy, hogy minden $v \in V$ esetén

$$B(v, v) \geq \beta\|v\|^2 \qquad \big(B(v, v) \leq -\beta\|v\|^2\big). \tag{$*$}$$

Világos, hogy (ii) erősebb tulajdonság, mint (i): ha a B bilineáris leképezés szigorúan pozitív definit, akkor minden $v \in V$, $\|v\| = 1$ esetén $B(v, v) > 0$, ezért bármilyen $\alpha \neq 0$ estén $B(\alpha v, \alpha v) = \alpha^2 B(v, v) > 0$.

Ha V véges dimenziós, akkor (i) és (ii) megegyezik egymással; ez azon múlik, hogy véges dimenziós normált térben az egységgömbhéj mindig kompakt, a bilineáris leképezések pedig folytonosak. Ezért egy bilineáris leképezés felveszi a minimumát és a maximumát az egységgömbhéjon. Ha B pozitív definit, akkor az egységgömbhéjon a minimuma (infimuma) nem lehet nulla, mert csak a nullában veszi fel a nulla értéket.

Végtelen dimenziós normált térben azonban ez nem igaz, amit a következő példa mutat:

$$l^2(\mathbb{R}) \times l^2(\mathbb{R}) \to \mathbb{R}, \quad (x, y) \mapsto \sum_{n\in\mathbb{N}} \frac{x_n y_n}{n}$$

olyan folytonos bilineáris leképezés, amely pozitív definit, de nem szigorúan pozitív definit, amit könnyen ellenőrizhet az Olvasó is.

10.3. Állítás *Legyen $F : V \rightarrowtail \mathbb{R}$ az a pontban kétszer differenciálható és $\mathrm{D}F(a) = 0$.*

(i) Ha $\mathrm{D}^2 F(a)$ szigorúan pozitív (negatív) definit, akkor F-nek az a pontban szigorú lokális minimuma (maximuma) van;

(ii) ha $\mathrm{D}^2 F(a)$ indefinit, akkor F-nek az a pontban nincs lokális szélsőértéke.

(iii) Ha F-nek az a pontban lokális minimuma (maximuma) van, akkor $\mathrm{D}^2 F(a)$ pozitív (negatív) szemidefinit.

Bizonyítás Nyilvánvaló, hogy a belső pontja az F értelmezési tartományának, ezért létezik olyan $G(a)$ környezete, hogy $G(a) \subset \mathrm{Dom}(F)$. Legyen $0 \neq h \in V$ olyan, hogy $a + h \in G(a)$, és írjuk fel F-re az általános Taylor-formulát:

$$F(a+h) = F(a) + \mathrm{D}F(a)h + \frac{1}{2!}\mathrm{D}^2 F(a)(h,h) + o(\|h\|^2) =$$
$$= F(a) + \|h\|^2 \left(\frac{1}{2}\mathrm{D}^2 F(a) \left(\frac{h}{\|h\|}, \frac{h}{\|h\|} \right) + \frac{o(\|h\|^2)}{\|h\|^2} \right).$$

(i) Ha $\mathrm{D}^2 F(a)$ szigorúan pozitív definit (a másik eset teljesen hasonlóan tárgyalható), akkor a 10.2. $(*)$ összefüggés alapján

$$F(a+h) > F(a) + \|h\|^2 \left(\frac{\beta}{2} + \frac{o(\|h\|^2)}{\|h\|^2} \right)$$

következik.

Mivel $o : \mathbb{R} \rightarrowtail \mathbb{R}$ kis ordó függvény, létezik olyan $\delta > 0$, hogy minden $0 \neq h$, $a + h \in G(a)$, $\|h\| < \delta$ esetén

$$\frac{o(\|h\|^2)}{\|h\|^2} < \frac{\beta}{2}$$

amiből ilyen h-kra $F(a+h) > F(a)$ adódik, vagyis F-nek az a-ban szigorú lokális minimuma van.

(iii) Legyen az a pontban lokális minimuma F-nek (a lokális maximum hasonlóan tárgyalható). Az eddigiek alapján ez azt jelenti, hogy létezik olyan $\delta > 0$, hogy bármely $0 \neq h$, $a + h \in G(a)$, $\|h\| < \delta$ esetén $F(a+h) \geq F(a)$, azaz

$$\frac{1}{2}\mathrm{D}^2 F(a) \left(\frac{h}{\|h\|}, \frac{h}{\|h\|} \right) + \frac{o(\|h\|^2)}{\|h\|^2} \geq 0.$$

Ezt átírhatjuk úgy, hogy bármely $v \in V$, $\|v\| = 1$ és minden $0 < \alpha < \delta$ estén

$$\frac{1}{2}\mathrm{D}^2 F(a)(v,v) + \frac{o(\alpha^2)}{\alpha^2} \geq 0. \tag{$*$}$$

Tekintve, hogy $o : \mathbb{R} \rightarrowtail \mathbb{R}$ kis ordó függvény, ebből

$$\mathrm{D}^2F(a)(v,v) \geq 0 \qquad (\|v\| = 1)$$

szükségszerűen következik, és ez éppen azt jelenti, hogy $\mathrm{D}^2F(a)$ pozitív szemidefinit.

(ii) Ha $\mathrm{D}^2F(a)$ indefinit, akkor éppen az imént mondottak alapján nem lehet lokális szélsőértéke F-nek az a-ban, hiszen akkor vagy pozitív vagy negatív szemidefinitnek kellene lennie.

10.4. Emlékezzünk vissza, hogy A.8.6-ban valós-valós függvényekre tárgyaltuk a lokális szélsőérték elégséges feltételeit. Megvizsgáltuk, hogy mit lehet mondani, amikor az adott függvénynek nem csak az első, hanem a második (vagy az első valahány) deriváltja nulla. Bizonyítás nélkül kimondjuk ennek az állításnak a megfelelőjét normált téren értelmezett függvényekre. Ehhez értelemszerűen általánosítjuk *páros* n esetén n-lineáris leképezésekre a **pozitív (negatív) definit**, **pozitív (negatív) szemidefinit**, **szigorúan pozitív (negatív) definit** és **indefinit** fogalmakat a 10.3. mintájára. Például az $R : V^n \to \mathbb{R}$ n-lineáris leképezés pozitív definit, ha $R(v, v, \ldots, v) > 0$ minden $0 \neq v \in V$ esetén.

Állítás *Legyen $F : V \rightarrowtail \mathbb{R}$ n-szer differenciálható az a pontban és*

$$\mathrm{D}F(a) = 0, \quad \mathrm{D}^2F(a) = 0, \quad \ldots \quad \mathrm{D}^{n-1}F(a) = 0, \quad \mathrm{D}^nF(a) \neq 0.$$

(i) Az F függvénynek csak akkor lehet lokális minimuma (maximuma) az a pontban, ha n páros és $\mathrm{D}^nF(a)$ pozitív (negatív) szemidefinit,

(ii) Ha n páros és $\mathrm{D}^nF(a)$

1) szigorúan pozitív (negatív) definit, akkor F-nek a-ban szigorú lokális minimuma (maximuma) van,

2) indefinit, akkor F-nek nincs az a-ban lokális szélsőértéke.

10.5. Gyakran előfordul, hogy egy függvénynek a szélsőértékét keressük, de csak egy olyan halmazon, amelynek elemei eleget tesznek bizonyos feltételeknek. Például egy többváltozós függvény maximumára vagyunk kiváncsiak, ha néhány változója előre rögzített, vagy mondjuk egy $\mathbb{R}^3 \to \mathbb{R}$ függvény minimumára egy ellipszoid felületén, stb.

Legegyszerűbbnek tűnik leszűkíteni a szóbanforgó függvényt a kérdéses halmazra és ott vizsgálni. Ez a halmaz azonban esetleg nem nyílt, sőt talán a belseje üres (vagyis nem tartalmaz nyílt részhalmazt; ilyen például egy ellipszoid felület $\mathbb{R}^3$-ban), és ekkor a leszűkített függvényünk már differenciálható sem lehet, így az eddig megismert differenciálszámítási módszerek nem alkalmazhatók a szélsőérték meghatározásához.

Matematikailag a problémát következőképpen fogalmazhatjuk meg:

Definíció *Legyen $F : V \rightarrowtail \mathbb{R}$, $S : V \rightarrowtail \mathbb{R}^M$, $0 \in \mathrm{Ran}(S)$. Azt mondjuk, hogy az F függvénynek (lokális)* ***szélsőértéke*** *van az $a \in \mathrm{Dom}(F) \cap \overset{-1}{S}\{0\}$ pontban az S* ***feltétel mellett****, ha az $F|_{\overset{-1}{S}\{0\}}$ függvénynek (lokális) szélsőértéke van a-ban.*

Állítás *Legyen V teljes normált tér (Banach-tér), $F : V \rightarrowtail \mathbb{R}$, $S : V \rightarrowtail \mathbb{R}^M$ folytonosan differenciálható, $a \in \mathrm{Dom}(F) \cap \overset{-1}{S}\{0\}$, és $\mathrm{D}S(a) \in \mathcal{L}in(V, \mathbb{R}^M)$ ráképezés. Ha F-nek lokális szélsőértéke van a-ban az S feltétel mellett, akkor létezik olyan $\lambda : \mathbb{R}^M \to \mathbb{R}$ lineáris leképezés, hogy*

$$\mathrm{D}(F - \lambda \circ S)(a) = 0.$$

Bizonyítás (i) Először legyen V véges dimenziós. A határozottság kedvéért tegyük fel, hogy F-nek lokális maximuma van a-ban az S feltétel mellett; természetesen minimum esetén a bizonyítás teljesen hasonlóan végezhető el. Létezik tehát olyan $N \subset V$ környezete a-nak, hogy $F(x) \leq F(a)$ minden $x \in N \cap \overset{-1}{S}\{0\}$ esetén.

Tekintsük az $(F, S) : V \rightarrowtail \mathbb{R} \times \mathbb{R}^M$ folytonosan differenciálható együttes függvényt. $(F, S)(a) = (F(a), 0) \in (F, S)(N)$ nem lehet belső pontja az $(F, S)(N) \subset \mathbb{R} \times \mathbb{R}^M$ halmaznak, hiszen ha $\alpha > F(a)$, akkor $(\alpha, 0) \notin (F, S)(N)$, mert F-nek lokális maximuma van a-ban az S feltétel mellett.

A mondottak miatt $\mathrm{D}(F, S)(a) = (\mathrm{D}F(a), \mathrm{D}S(a)) : V \to \mathbb{R} \times \mathbb{R}^M$ nem lehet ráképezés, ugyanis ha az volna, akkor a 7.3.(iii) alapján létezne olyan környezete az a pontnak, hogy (F, S) ezen a környezeten nyílt leképezés volna, és ezáltal $(F(a), 0)$ belső pontja lenne $(F, S)(N)$-nek.

Tehát $(\mathrm{D}F(a), \mathrm{D}S(a)) : V \to \mathbb{R} \times \mathbb{R}^M$ nem ráképezés, viszont $\mathrm{D}S(a) : V \to \mathbb{R}^M$ az. Ebből az következik, hogy $(\mathrm{D}F(a), \mathrm{D}S(a))$ értékkészlete, $L := \mathrm{Ran}(\mathrm{D}F(a), \mathrm{D}S(a))$ az $\mathbb{R} \times \mathbb{R}^M$ egy M-dimenziós altere (ugyanis legfeljebb és legalább M-dimenziós). Nyilvánvaló, hogy

$$\mathrm{D}S(a) = \mathrm{pr}_2 \circ (\mathrm{D}F(a), \mathrm{D}S(a)),$$

ahol $\mathrm{pr}_2 : \mathbb{R} \times \mathbb{R}^M \to \mathbb{R}^M$ a szokásos második projekció. A $\mathrm{pr}_2|_L : L \to \mathbb{R}^M$ lineáris leképezés szürjekció, és $\dim L = M$, tehát bijekció is, és így

$$(\mathrm{D}F(a), \mathrm{D}S(a)) = (\mathrm{pr}_2|_L)^{-1} \circ \mathrm{D}S(a).$$

Bevezetve a $\lambda := \mathrm{pr}_1 \circ (\mathrm{pr}_2|_L)^{-1} : \mathbb{R}^M \to \mathbb{R}$ lineáris leképezést (ahol $\mathrm{pr}_1 : \mathbb{R} \times \mathbb{R}^M \to \mathbb{R}$ a szokásos első projekció) ebből rögtön

$$\mathrm{D}F(a) = \lambda \circ \mathrm{D}S(a)$$

adódik, azaz $\mathrm{D}(F - \lambda \circ S)(a) = 0$, és ezzel véges dimenziós esetre a bizonyítást elvégeztük.

(ii) Legyen most V végtelen dimenziós. Mivel $\mathrm{D}S(a) : V \to \mathbb{R}^M$ lineáris ráképezés, van olyan $U \subset V$ lineáris altér, hogy $\dim U = M$ és $\mathrm{D}S(a)|_U$ bijekció.

Vegyünk egy tetszőleges $0 \neq v \in V$ vektort, és jelölje $L \subset V$ az a, v vektorok és az U altér által kifeszített legfeljebb $(M+2)$ dimenziós lineáris alteret. F-nek lokális szélsőértéke van az a pontban az S feltétel mellett, ezért az $F|_L$ függvénynek is lokális szélsőértéke van a-ban az $S|_L$ feltétel mellett. Tekintve, hogy L véges dimenziós, a bizonyítás első része alapján létezik olyan $\lambda_v : \mathbb{R}^M \to \mathbb{R}$ lineáris leképezés, hogy

$$\mathrm{D}\big(F|_L - \lambda_v \circ S|_L\big)(a) = 0.$$

Ebből következik (lásd az 1.11.16. feladatot), hogy

$$\mathrm{D}(F - \lambda_v \circ S(a))|_L = 0,$$

sőt

$$\mathrm{D}F(a)|_U - \lambda_v \circ \mathrm{D}S(a)|_U = \mathrm{D}(F - \lambda_v \circ S)(a)|_U = 0,$$

amiből

$$\lambda_v = \mathrm{D}F(a)|_U \circ \big(\mathrm{D}S(a)|_U\big)^{-1} =: \lambda$$

adódik, azaz λ_v nem függ az v választásától. Ezek szerint

$$\mathrm{D}(F - \lambda \circ S)(a)|_L = 0$$

teljesül, és így készen vagyunk, mert $v \in V$ tetszőlegesen választott elem volt, és L tartalmazza v-t, ezért az utolsó egyenlőségből

$$\mathrm{D}(F - \lambda \circ S)(a) = 0$$

következik.

10.6. A feltételes lokális szélsőértékre vonatkozó szükséges feltétel, amelyet az imént megfogalmaztunk, egy bizonyos tulajdonságú $\lambda \in \mathrm{Lin}(\mathbb{R}^M, \mathbb{R}) \equiv \mathbb{R}^M$ létezését állítja. A gyakorlatban, amikor feltételes szélsőértéket keresünk, ez M darab ismeretlent jelent, viszont maga az $S(a) = 0$ feltétel pontosan M darab egyenletet biztosít számunkra. Ha például (az állításban használt jelöléseknél maradva) $F : \mathbb{R}^N \rightarrowtail \mathbb{R}$ és $S = (S_1, \ldots, S_M) : \mathbb{R}^N \rightarrowtail \mathbb{R}^M$ akkor

$$\begin{aligned} \partial_i \left(F - \sum_{k=1}^{M} \lambda_k S_k \right)(a) &= 0, \qquad (i = 1, ..., N), \\ S_k(a) &= 0, \qquad (k = 1, ..., M) \end{aligned}$$

éppen $N+M$ egyenlet az $a =: (a_1, \dots, a_N)$ és $(\lambda_1, \dots, \lambda_M)$ ismeretlenekre vonatkozóan, ahol $a \in \mathbb{R}^N$ a keresett lehetséges lokális szélsőérték(ek) helye(i), $\lambda_1, \dots, \lambda_M$ pedig az úgynevezett **Lagrange-multiplikátorok**.

10.7. Használjuk a 10.5. jelöléseit, és tegyük fel, hogy F-nek szélsőértéke van a-ban az S feltétel mellett. Ha F és S kétszer differenciálható, akkor olyan h-kra, amelyekre $a + h \in \overset{-1}{S}\{0\}$, az $F - \lambda \circ S$ függvényre felírva az általános Taylor-formulát azt kapjuk, hogy

$$F(a + h) = F(a) + \frac{1}{2}\mathrm{D}^2(F - \lambda \circ S)(a)(h, h) + o(\|h\|^2). \qquad (*)$$

Ezzel elismételhetjük a 10.3. (i) bizonyításában mondottakat: ha $\mathrm{D}^2(F-\lambda\circ S)(a)$ szigorúan pozitív (negatív) definit, akkor F-nek az a pontban szigorú lokális minimuma (maximuma) van az S feltétel mellett. Ez azonban túl erős feltétel, hiszen látjuk, hogy $\mathrm{D}^2(F - \lambda \circ S)(a)$-nak csak bizonyos h-kon felvett értékei szerepelnek. Emiatt viszont a 10.3. (iii) és ezáltal (ii) pontja sem szükségképpen igaz, hiszen az ottani $(*)$ egyenlőtlenséget most nem tudjuk minden v egységvektorra és δ-nál kisebb α-ra, hiszen $\alpha v \in \overset{-1}{S}\{0\}$ fenn kell, hogy álljon.

Ha V véges dimenziós, akkor többet tudunk mondani.

Állítás *Legyen V véges dimenziós, $F : V \rightarrowtail \mathbb{R}$ és $S : V \rightarrowtail \mathbb{R}^M$ az a pontban kétszer differenciálható, és $\lambda \in Lin(\mathbb{R}^M, \mathbb{R})$ olyan, hogy $\mathrm{D}(F - \lambda \circ S)(a) = 0$.*

(i) Ha $\mathrm{D}^2(F - \lambda \circ S)(a)$ pozitív (negatív) definit a $\mathrm{Ker}(\mathrm{D}S)(a)$ altéren, akkor F-nek az a pontban szigorú lokális minimuma (maximuma) van az S feltétel mellett;

(ii) ha $\mathrm{D}^2(F - \lambda \circ S)(a)$ indefinit a $\mathrm{Ker}(\mathrm{D}S)(a)$ altéren, akkor F-nek az a pontban nincs lokális szélsőértéke az S feltétel mellett;

(iii) ha F-nek az a pontban lokális maximuma (minimuma) van, akkor $\mathrm{D}^2(F - \lambda \circ S)(a)$ pozitív (negatív) szemidefinit a $\mathrm{Ker}(\mathrm{D}S)(a)$ altéren.

Bizonyítás Mivel V véges dimenziós, megadhatunk rajta normát skaláris szorzattal (Analízis III.B.6.2.). Legyen U a $\mathrm{D}S(a)$ magjának ortogonális kiegészítő altere. Mivel $\mathrm{D}S(a)$ injektív az U-n, $\mathrm{D}S(a)|_U$ inverze (amely szintén véges dimenziós vektortéren van értelmezve) folytonos, tehát van olyan $\alpha > 0$, hogy

$$\|\mathrm{D}S(a)h_U\| \geq \alpha\|h_U\| \qquad (1)$$

minden $h_U \in U$ esetén. Továbbá minden $h \in V$ egyértelműen előáll $h = h_S + h_U$ alakban, ahol $h_S \in \mathrm{Ker}(\mathrm{D}S(a))$ és $h_U \in U$, és $\|h\|^2 = \|h_S\|^2 + \|h_U\|^2$, amiből

$$\|h_S\| \leq \|h\|. \qquad (2)$$

A Taylor formulából

$$S(a+h) = S(a) + \mathrm{D}S(a)h + o(h) = \mathrm{D}S(a)h_U + o(h);$$

ha $a+h \in \overset{-1}{S}\{0\}$, akkor a bal oldal nulla, tehát (1) alapján

$$\|h_U\| \leq \frac{\|\mathrm{D}S(a)h_U\|}{\alpha} = \frac{\|o(h)\|}{\alpha}. \tag{3}$$

Az egyszerűség kedvéért az $A := \mathrm{D}^2(F - \lambda \circ S)(a)$ jelöléssel a $(*)$ összefüggést átírhatjuk

$$F(a+h) = F(a) + \frac{1}{2}A(h_S, h_S) + A(h_S, h_U) + \frac{1}{2}A(h_U, h_U) + o(\|h\|^2) \tag{4}$$

alakba. A (2) és a (3) összefüggés szerint $\dfrac{|A(h_S, h_U)|}{\|h\|^2} \leq \|A\| \dfrac{\|o(h)\|}{\alpha\|h\|}$, és hasonló becsléssel $\dfrac{|A(h_U, h_U)|}{\|h\|^2}$-ra vonatkozóan megállapíthatjuk, hogy (4)-ben a jobb oldal harmadik és negyedik tagja is kis ordo h-ban, azaz

$$F(a+h) = F(a) + \frac{1}{2}A(h_S, h_S) + o(\|h\|^2).$$

Ezután már ugyanúgy érvelhetünk, mint 10.3-ban.

10.8. Feladatok

1. Az $(a, b, c) \in \mathbb{R}^3$ ponton keresztül fektessünk olyan síkot, amely a koordinátasíkokkal a legkisebb térfogatú tetraédert alkotja!

2. Határozzuk meg a felülről nyitott derékszögű láda méreteit úgy, hogy adott V térfogat és h falvastagság mellett a legkevesebb anyagból legyen elkészíthető!

3. Legyen adva $a_1, \dots, a_m \in \mathbb{R}^N$, és bármely $\mathbb{R}^N$-beli x ponthoz rendeljük hozzá az a_k-tól mért távolságának négyzetösszegét $(k = 1, \dots, m)$. Keressük meg e függvény minimumát!

4. Határozzuk meg az impliciten $x \in \mathbb{R}$ és $y \in \mathbb{R}$ függvényeként adott $z \in \mathbb{R}$ legkisebb és legnagyobb értékeit:
(i) $2x^2 + 2y^2 + z^2 + 8xz - z + 8 = 0$;
(ii) $x^4 + y^4 + z^4 = 2a^2(x^2 + y^2 + z^2)$, ahol $a \in \mathbb{R}$ adott.

5. Határozzuk meg a következő $\mathbb{R}^3 \rightarrowtail \mathbb{R}$ függvényeknek a feltüntetett feltételek melletti szélsőértékeit:
(i) $(x, y, z) \mapsto xyz, \qquad x^2 + y^2 + z^2 = 3$;
(ii) $(x, y, z) \mapsto x^2y^3z^4, \qquad 2x + 3y + 4z = a \quad (a \in \mathbb{R})$;
(iii) $(x, y, z) \mapsto xyz, \qquad x + y + z = 5, \quad xy + yz + xz = 8$;
(iv) $(x, y, z) \mapsto \sin\frac{x}{2}\sin\frac{y}{2}\sin\frac{z}{2}, \qquad x + y + z = \pi, \quad x > 0,\ y > 0,\ z > 0.$

6. Legyen $F : \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \mapsto x^2 + y^2 + z^2$, és $S : \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \mapsto \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1$, ahol $a > b > c > 0$ adott számok. Alkalmazzuk ezekre a függvényekre a 10.7. állítást.

11. A variációszámítás elemei

11.1. A szélsőérték-problémák egy érdekes és fontos fejezete a variációszámítás. Történeti okból azért érdekes, mert a klasszikus analízisen – a néhány valós változós függvények elméletén – kívül eső problémát vet fel és tárgyal; gyakorlati szempontból pedig azért fontos, mert a fizika és a műszaki tudományok megkövetelik olyan függvények szélsőértékeinek megkeresését, amelyeknek értelmezési tartománya nem $\mathbb{R}^N$ részhalmaza, hanem például valamely görbék vagy felületek egy osztálya.

Fogalmazzuk meg példaképpen az egyik ilyen legegyszerűbb feladatot, az úgynevezett brachisztochron-problémát, amelynek nagy szerepe volt a variációszámítás kialakulásában.

Legyen adott két pont a fizikai terünkben, amelyek különböző magasságban, nem egymás alatt helyezkednek el. Vegyük azokat a görbéket, amelyeknek ezek a végpontjai. Melyik az a peremes görbe, amely mentén (mint kényszerpályán) adott kezdősebességgel indított anyagi pont a nehézségi erő hatása alatt az egyik pontból a másikba a legrövidebb idő alatt jut el?

Mi a variációszámításnak arra az esetére korlátozódunk, amikor az előzőekben tanult differenciálási módszerek alkalmazhatók, vagyis a vizsgált függvények értelmezési tartománya egy normált tér részhalmaza. Ennek is csak egy speciális részét érintjük, amelyet a fizikában szokás alkalmazni.

11.2. Tekintsünk egy külső erő hatása alatt mozgó tömegpontot. Az időt valós számokkal, a fizikai teret valós számhármasokkal reprezentáljuk. A mozgás a (t_1, t_2) időtartamban megy végbe, és ez alatt a tömegpont a P_1 pontból a P_2 pontba jut el. A mozgást egy $r : [t_1, t_2] \to \mathbb{R}^3$ kétszer folytonosan differenciálható függvénnyel írjuk le, amelyre $r(t_1) = P_1$, $r(t_2) = P_2$. A mozgást Newton törvénye határozza meg, vagyis egy másodrendű differenciálegyenlet.

Lagrange szerint ehhez a differenciálegyenlethez egy variációs probléma megoldásával juthatunk el. Tekintsük azokat a kétszer folytonosan differenciálható $r : [t_1, t_2] \to \mathbb{R}^3$ függvényeket, amelyekre $r(t_1) = P_1$, $r(t_2) = P_2$. Ezeket a fizikában virtuális (képzeletbeli) mozgásoknak szokás nevezni. A tömegpontra ható erő segítségével megadható egy olyan $L : \mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}^3 \rightarrowtail \mathbb{R}$ úgynevezett **Lagrange-függvény**, hogy a valódi, a létrejövő mozgás az, amelyre az előbbi függvényosztályon adott

$$r \mapsto \int_{t_1}^{t_2} L(t, r(t), \dot{r}(t))dt$$

funkcionál extremális értéket vesz föl.

Ez a módszer alkalmazható arra az esetre is, ha több tömegpont mozgását vizsgáljuk, és még akkor is, ha a tömegpont(ok) nem mozdulhat(nak) el szabadon akárhogyan a térben, hanem kényszer(ek) alá van(nak) vetve. Sőt, gyakorlati szempontból éppen ez utóbbi esetekben nagy jelentőségű a Lagrange-módszer, mert áttekinthetetlennek tűnő problémáknál is többnyire viszonylag egyszerűen írhatjuk fel segítségével a mozgásegyenleteket.

11.3. Fogalmazzuk most meg a pontos matematikai problémát.

Legyen adott $t_1, t_2 \in \mathbb{R}$ és $P_1, P_2 \in \mathbb{R}^N$. A

$$V := \left\{ r \in C^1([t_1, t_2], \mathbb{R}^N) \mid r(t_1) = P_1, r(t_2) = P_2 \right\}$$

függvényosztály affin tér a

$$\delta V := \left\{ \delta r \in C^1([t_1, t_2], \mathbb{R}^N) \mid (\delta r)(t_1) = (\delta r)(t_2) = 0 \right\}$$

vektortér felett, amely a

$$\|\delta r\| := \max_{t \in [t_1, t_2]} \{|(\delta r)(t)| + |(\delta r)^{\cdot}(t)\}$$

normával ellátva Banach-tér, amint azt tudjuk (Analízis III.B.4.3.2.). V tehát normált affin tér, alkalmazhatjuk a differenciálszámítás módszereit (lásd a következő fejezetet).

A kissé szokatlan δ jeleket azért használjuk, mert a fizikában elterjedtek, és így könnyebb összevetni eredményeinket az ottani alkalmazásokkal. Sőt, a további egyszerűsített írásmódot is elfogadjuk: $\delta r(t) := (\delta r)(t)$, $\delta \dot{r} := (\delta r)^{\cdot}$.

Legyen $L : \mathbb{R} \times \mathbb{R}^N \times \mathbb{R}^N \rightarrowtail \mathbb{R}$ folytonos függvény, és keressük az

$$S : V \to \mathbb{R}, \quad r \mapsto \int_{t_1}^{t_2} L(t, r(t), \dot{r}(t))dt$$

leképezésnek – amelyet a fizikában **hatásnak** szokás nevezni – a szélsőértékeit.

A következő fejezetben megmutatjuk, hogy a differenciálszámítás eszközei alkalmazhatók a V normált affin térre. Ha S differenciálható, akkor a szélsőértékeinél a deriváltja nulla, tehát első feladatunk feltételt adni arra, mikor differenciálható S, második feledatunk pedig az, hogy felkutassuk a $\mathrm{D}S$ nulla-helyeit.

Mielőtt erre vonatkozó kijelentést tennénk, állapodjunk meg abban, hogy L változóit rendre nulladiknak, elsőnek és másodiknak nevezzük.

11.4. Állítás *Ha L kétszer folytonosan differenciálható, akkor S differenciálható és*

$$\mathrm{D}S(r)\delta r = \int_{t_1}^{t_2} \Big(\partial_1 L(t, r(t), \dot r(t))\delta r(t) + \partial_2 L(t, r(t), \dot r(t))\delta \dot r(t) \Big) dt,$$

ahol $r \in V$ és $\delta r \in \delta V$.

Bizonyítás Mivel L kétszer folytonosan differenciálható, minden rögzített $t \in [t_1, t_2]$ és $r \in V$, $\delta r \in \delta V$ esetén a Taylor-formulából

$$\begin{aligned} L\big(t, r(t) + \delta r(t), \dot r(t) + \delta \dot r(t)\big) - L(t, r(t), \dot r(t)) = \\ = \partial_1 L(t, r(t), \dot r(t))\delta r(t) + \partial_2 L(t, r(t), \dot r(t))\delta \dot r(t) + \\ + \frac{1}{2}\Big[\partial_1^2 L(t, \rho(t), \xi(t))(\delta r(t))^2 + \\ + 2\partial_1\partial_2 L(t, \rho(t), \xi(t))\delta r(t)\delta \dot r(t) + \\ + \partial_2^2 L(t, \rho(t), \xi(t))(\delta \dot r(t))^2\Big], \end{aligned}$$

ahol $\rho(t)$ és $\xi(t)$ az $r(t)$ és $r(t) + \delta r(t)$ illetve az $\dot r(t)$ és $\dot r(t) + \delta \dot r(t)$ közötti szakasznak eleme $\mathbb{R}^N$-ben. Felhívjuk a figyelmet, hogy itt $\delta r(t), \delta \dot r(t) \in \mathbb{R}^N$, tehát ne tévesszük szem elől, hogy például a fenti jobb oldal első tagjában a parciális derivált mint lineáris leképezés hat $\delta r(t)$-re, az utolsó tagban pedig a második parciális derivált mint bilineáris leképezés hat $(\delta \dot r(t), \delta \dot r(t))$-re.

A fenti formulának megfelelően az

$$S(r + \delta r) - S(r) = \int_{t_1}^{t_2} \Big(L(t, r(t) + \delta r(t), \dot r(t) + \delta \dot r(t)) - L(t, r(t), \dot r(t)) \Big) dt$$

megváltozást két tagra bontjuk.

Az első,

$$\int_{t_1}^{t_2} \Big(\partial_1 L(t, r(t), \dot r(t))\delta r(t) + \partial_2 L(t, r(t), \dot r(t))\delta \dot r(t) \Big) dt$$

nyilván lineáris δr-ben és folytonos is, mert $|\delta r(t)| \le \|\delta r\|$, $|\delta \dot r(t)| \le \|\delta r\|$, és a $\partial_1 L$ illetve $\partial_2 L$ folytonos fügvények korlátosak a $[t_1, t_2] \times \mathrm{Ran}(r) \times \mathrm{Ran}(\dot r)$ kompakt halmazon; ha K egy korlátjuk, akkor az egész kifejezés majorálható a $2K(t_2 - t_1)\|\delta r\|$ értékkel.

A szóban forgó megváltozás második tagjáról pedig kimutatjuk, hogy δr-ben kis ordó függvény. Ugyanis L kétszeres folytonos differenciálhatósága miatt $\partial_1^2 L$, $\partial_1\partial_2 L$ és $\partial_2^2 L$ korlátos függvények a $[t_1, t_2] \times (\mathrm{Ran}(r) + B) \times (\mathrm{Ran}(\dot{r}) + B)$ kompakt halmazon, ahol $B := \{x \in \mathbb{R}^N \mid |x| \leq 1\}$. Így, ha M jelöli e függvények egy korlátját, akkor a $\|\delta r\| \leq 1$ feltétel teljesülése esetén a megváltozás ezen második tagját $2M(t_2 - t_1)\|\delta r\|^2$ felülről becsüli.

Megjegyzés Az iméntinél jóval erősebb állítást is kimondhattunk volna, ugyanis S folytonosan is differenciálható, már akkor is, ha L csak egyszer folytonosan differenciálható. Ez azonban bonylultabb bizonyítást igényelt volna, a fizikában pedig általában elegendő a fentebb igazolt állítás ismerete.

11.5. Állítás *Ha $r \in V$ és r kétszer folytonosan differenciálható, akkor $\mathrm{D}S(r) = 0$ pontosan akkor teljesül, ha r kielégíti az úgynevezett Euler–Lagrange egyenletet:*

$$\partial_1 L(t, r(t), \dot{r}(t)) - \frac{d}{dt}\partial_2 L(t, r(t), \dot{r}(t)) = 0.$$

Bizonyítás r kétszer folytonosan differenciálható, így $t \mapsto \partial_2 L(t, r(t), \dot{r}(t))$ differenciálható, ezért a $\mathrm{D}S(r)$-t megadó formulában parciálisan integrálhatunk és a $\delta r(t_1) = \delta r(t_2) = 0$ feltétel miatt azt kapjuk, hogy

$$\mathrm{D}S(r)\delta r = \int_{t_1}^{t_2} \Big(\partial_1 L(t, r(t), \dot{r}(t)) - \frac{d}{dt}\partial_2 L(t, r(t), \dot{r}(t))\Big)\delta r(t)dt.$$

Ha tehát r kielégíti az Euler–Lagrange-egyenletet, akkor $\mathrm{D}S(r) = 0$.

Ha viszont $\mathrm{D}S(r) = 0$, akkor minden $\delta r \in \delta V$ esetén $\mathrm{D}S(r)\delta r = 0$, amiből a következő úgynevezett Lagrange-lemmával már adódik, hogy r kielégíti az Euler–Lagrange-egyenletet.

11.6. A Lagrange-féle lemma így szól: ha $\phi \in C([t_1, t_2], \mathbb{R}^N)$ és minden $\delta r \in \delta V$ esetén

$$\int_{t_1}^{t_2} \phi(t) \cdot \delta r(t)dt = 0$$

(ahol $\cdot$ jelöli a szokásos skalárszorzást), akkor $\phi = 0$.

Bizonyítását indirekt módon végezzük. Tegyük fel, hogy $\phi \neq 0$. Ekkor van olyan komponense ϕ-nek, amely nem nulla; az általánosság megszorítása nélkül feltehetjük, hogy ez ϕ_1. ϕ_1 folytonossága miatt van olyan $[t_0 - \epsilon, t_0 + \epsilon]$ intervallum $[t_1, t_2)]$-ben, amelyen ϕ_1 azonos előjelű, mondjuk pozitív. A

$$\kappa : [t_1, t_2] \to \mathbb{R}, \quad t \mapsto := \begin{cases} (t - t_0 + \epsilon)^3(t_0 - t + \epsilon)^3 & \text{ha } t \in [t_0 - \epsilon, t_0 + \epsilon], \\ 0 & \text{ha } t \notin [t_0 - \epsilon, t_0 + \epsilon]. \end{cases}$$

függvény folytonosan differenciálható, a $[t_0-\epsilon, t+\epsilon]$ intervallumon pozitív, azon kívül nulla, tehát $\delta r := (\kappa, 0, 0) \in \delta V$, és $\int\limits_{t_1}^{t_2} \phi(t)\cdot \delta r(t) dt = \int\limits_{t_1}^{t_2} \phi_1(t)\kappa(t)dt > 0$, ami ellentmondás.

12. Differenciálás affin terekben

12.1. A differenciálhatóság fogalmát szinte változatlan formában bevezethetjük normált affin terekben is.

Ha V affin tér a $\mathbf{V}$ vektortér fölött (Analízis II.42.), $\mathbf{V}$-n adott egy $\|.\|$ norma, akkor

$$V \times V \to \mathbb{R}_0^+, \qquad (x, y) \mapsto \|x-y\|$$

metrika V-n. Ekkor normált affin térről beszélünk, és szokásosan egyetlen szimbólummal, V-vel jelöljük a továbbiakban. Véges dimenziós affin terek esetén nem is kell konkrét normát megadnunk az alulfekvő vektortéren (hiszen bármely két norma ekvivalens), hogy nyílt halmazokról, konvergenciáról, folytonosságról és differenciálhatóságról beszélhessünk az affin térben.

A differenciálhatóságot formailag ugyanúgy definiálhatjuk normált affin terek közötti függvényekre, mint ahogyan normált vektortereknél tettük.

Legyen V és W normált affin tér. Az $F : V \rightarrowtail W$ függvény differenciálható az a pontban, ha $a \in \overset{\circ}{\mathrm{Dom}}(F)$, és létezik olyan $\mathrm{D}F(a) : \mathbf{V} \to \mathbf{W}$ folytonos lineáris leképezés, hogy

$$F(x) - F(a) = \mathrm{D}F(a)(x-a) + o(x-a),$$

ahol $o : \mathbf{V} \rightarrowtail \mathbf{W}$ kis ordó függvény.

A differenciálhatóságra vonatkozó összefüggések értelemszerűen érvényben maradnak ebben az esetben is. Például, ha $F, G : V \rightarrowtail W$ és $\mathbf{G} : V \rightarrowtail \mathbf{W}$ differenciálhatók a-ban, akkor

– $F - G : V \rightarrowtail \mathbf{W}$ is differenciálható a-ban, és $\mathrm{D}(F-G)(a) = \mathrm{D}F(a) - \mathrm{D}G(a)$,

– $F + \mathbf{G}$ is differenciálható a-ban, és $\mathrm{D}(F+\mathbf{G})(a) = \mathrm{D}F(a) + \mathrm{D}\mathbf{G}(a)$.

A többszörös differenciálhatóság értelmezése is lényegében ugyanúgy megy, mint normált terek esetén. Az $F : V \rightarrowtail W$ első deriváltja $V \rightarrowtail \mathcal{L}in(\mathbf{V}, \mathbf{W})$ függvény, második deriváltja $V \rightarrowtail \mathcal{L}in^2(\mathbf{V}^2, \mathbf{W})$ függvény, stb.

A V affin tér x és y elemei közötti szakasz is értelmezhető a szokásos formulával: $[x, y] := \{x + \alpha(y-x) \mid \alpha \in [0,1]\}$.

Mindezeknek megfelelően, értelemszerű megfogalmazással érvényben marad az általános középértéktétel, a Taylor-formula, a szélsőértékekre vonatkozó feltételek stb.

12.2. Legyen I egy dimenziós (és így normáltnak tekinthető) affin tér, V normált affin tér. Ha $f : I \rightarrowtail V$ differenciálható a $t \in I$ pontban, akkor $\mathrm{D}f(t) \in$

$\mathrm{Lin}(\mathbf{I},\mathbf{V})$. Tudjuk, hogy minden $\mathbf{I}\to\mathbf{V}$ lineáris leképezés folytonos, ezért a differenciálhatóság értelmezésekor a folytonosság megkövetelése elhagyható.

Az is ismert, hogy $\mathrm{Lin}(\mathbf{I},\mathbf{V})$ és $\dfrac{\mathbf{V}}{\mathbf{I}}$ között van egy természetes azonosítás: $\dfrac{\mathbf{x}}{\mathbf{t}}(\mathbf{s})\equiv\dfrac{\mathbf{s}}{\mathbf{t}}\mathbf{x}$. $\dfrac{\mathbf{V}}{\mathbf{I}}$ normált tér a $\left\|\dfrac{\mathbf{x}}{\mathbf{t}}\right\|:=\dfrac{\|\mathbf{x}\|}{|\mathbf{t}|}$ normával, ahol $|.|$ az $\mathbf{I}$-n adott norma. $\mathrm{Lin}(\mathbf{I},\mathbf{V})$ is normált tér, és az említett azonosítás izometrikus:

$$\sup_{|\mathbf{s}|=1}\left\|\frac{\mathbf{x}}{\mathbf{t}}(\mathbf{s})\right\|=\sup_{|\mathbf{s}|=1}\left|\frac{\mathbf{s}}{\mathbf{t}}\right|\|\mathbf{x}\|=\frac{\|\mathbf{x}\|}{|\mathbf{t}|}.$$

Állítás *Legyen I egy dimenziós affin tér, V normált affin tér. Az $f:I\rightarrowtail V$ függvény akkor és csak akkor differenciálható a $t\in\overset{\circ}{\mathrm{Dom}}(f)$ pontban, ha létezik az*

$$f'(t):=\lim_{s\to t}\frac{f(s)-f(t)}{s-t}\in\frac{\mathbf{V}}{\mathbf{I}}$$

határérték, és az említett azonosítás szerint $f'(t)\equiv \mathrm{D}f(t)$.

Bizonyítás Ha f differenciálható t-ben, akkor $s\in\mathrm{Dom}(f)$ esetén

$$f(s)-f(t)=\mathrm{D}f(t)(s-t)+o(s-t),$$

ahol $o:\mathbf{I}\rightarrowtail\mathbf{V}$ kis ordó függvény. Ha $s\neq t$, akkor

$$\frac{f(s)-f(t)}{s-t}=\frac{\mathrm{D}f(t)(s-t)}{s-t}+\frac{o(s-t)}{s-t},$$

és az azonosítás szerint

$$\frac{\mathrm{D}f(t)(s-t)}{s-t}\equiv\mathrm{D}f(t).$$

Továbbá $\left\|\dfrac{o(s-t)}{s-t}\right\|=\dfrac{\|o(s-t)\|}{|s-t|}$, amiből már nyilvánvalóan következik, hogy létezik $f'(t)$ és $f'(t)\equiv\mathrm{D}f(t)$.

Teljesen hasonlóan érvelhetünk, ha azt akarjuk megmutatni, hogy $f'(t)$ létezéséből következik az f függvény t-beli differenciálhatósága; a részleteket az Olvasóra bízzuk. ■

Iménti állításunk az 1.5. pontban mondottak általánosítása.

Ha $f:I\rightarrowtail V$ többször differenciálható, akkor, lévén $f':I\rightarrowtail\dfrac{\mathbf{V}}{\mathbf{I}}$, azt kapjuk, hogy

$$f''(t)\in\frac{\frac{\mathbf{V}}{\mathbf{I}}}{\mathbf{I}}\equiv\frac{\mathbf{V}}{\mathbf{I}\otimes\mathbf{I}}$$

és $f''(t) \equiv D^2 f(t)$; és így tovább a magasabb rendű deriváltakra.

12.3. Ha I egy dimenziós valós irányított affin tér, vagyis az alulfekvő $\mathbf{I}$ vektortér irányított, akkor az irányítás kijelöli az $\mathbf{I}$ pozitív és negatív elemeit, és láncszerű rendezést határoz meg mind $\mathbf{I}$-n, mind I-n: $t < s$ pontosan akkor, ha $\mathbf{0} < s - t$. Ennek alapján értelmes az I intervallumairól beszélni. Ha D is egy dimenziós valós irányított affin tér, akkor értelmes az $I \rightarrowtail D$ monoton függvények fogalma.

Az Olvasóra bízzuk, lássa be azt az igen egyszerű tényt, hogy az A.2. fejezet tételei megfelelő fogalmazással érvényesek $I \rightarrowtail D$ függvényekre is.

III. RÉSZSOKASÁGOK

13. Részsokaságok paraméterezése, érintőterek

Ebben a fejezetben véges dimenziós vektorterek részsokaságainak értelmezésével és alapvető tulajdonságaival foglalkozunk; V mindvégig egy véges N dimenziós valós vektorteret jelöl ($N \in \mathbb{N}$).

13.1. Definíció *Legyen $M \in \mathbb{N}_0$. Egy $R \subset V$ nem üres halmazt M* ***dimenziós elemi részsokaságnak*** *hívunk, ha létezik olyan $p : \mathbb{R}^M \rightarrowtail V$ függvény, hogy*

(i) $\mathrm{Dom}(p)$ *összefüggő és nyílt,* $\mathrm{Ran}(p) = R$*,*
(ii) p folytonosan differenciálható,
(iii) $\mathrm{D}p(\xi) \in \mathrm{Lin}(\mathbb{R}^M, V)$ *injektív bármely $\xi \in \mathrm{Dom}(p)$ esetén,*
(iv) p injektív és az inverze folytonos.

A p függvényt az R ***paraméterezésének*** *nevezzük.*

Megjegyzések (i) A p paraméterezés is folytonos (nem csak az inverze), mert differenciálható. Azt szokás mondani, hogy p homeomorfizmus ("oda-vissza folytonos") $\mathrm{Dom}(p)$ és R között. Mivel összefüggő halmaz folytonos képe összefüggő, az elemi részsokaságok összefüggő halmazok.

(ii) Ha $V = \mathbb{K}^N$ és $M = 1$, visszakapjuk az A.12.1-ben definiált görbe fogalmát. Általában akármilyen V esetén az egy dimenziós elemi részsokaságot görbének hívjuk.

(iii) Egy elemi részsokaság dimenziója nem lehet nagyobb az őt tartalmazó vektortér dimenziójánál, vagyis – a definícióban szereplő jelölésekkel – $M \leq N$, mivel máskülönben a $\mathrm{D}p(\xi) : \mathbb{R}^M \to V$ lineáris leképezés nem lehetne injektív egyetlen $\xi \in \mathrm{Dom}(p)$ pontban sem.

(iv) Az N dimenziós elemi részsokaságok az összefüggő nyílt halmazok. Ugyanis ha $p:\mathbb{R}^N \rightarrowtail V$ a definícióban adott tulajdonságú függvény, akkor az inverzfüggvény-tétel szerint $\mathrm{Ran}(p)$ nyílt. Ha viszont R összefüggő és nyílt, akkor bármely $K:V \to \mathbb{R}^N$ lineáris bijekció (koordinátázás) esetén $K^{-1}|_{K[R]}$ az R paraméterezése.

(v) A 0 dimenziós elemi részsokaságok az 1.3.(ii) szerint az egypont halmazok, ugyanis $\mathbb{R}^0$-t az egyetlen pontból álló vektortérnek tekintjük.

(vi) A görbékhez hasonlóan itt is ki kell emelnünk, hogy egy elemi részsokaság paraméterezése nem egyértelmű. Ugyanis, ha p az R paraméterezése és $\phi : \mathbb{R}^M \rightarrowtail \mathbb{R}^M$ összefüggő halmazon értelmezett folytonosan differenciálható függvény, amely injektív és az inverze is folytonosan differenciálható, továbbá $\mathrm{Ran}(\phi) = \mathrm{Dom}(p)$, akkor a $p \circ \phi : \mathbb{R}^M \rightarrowtail V$ függvény szintén paraméterezése lesz az R elemi részsokaságnak (ϕ deriváltja minden pontban lineáris bijekció).

(vii) Ha U véges dimenziós vektortér és $F : V \rightarrowtail U$ összefüggő halmazon értelmezett folytonosan differenciálható függvény, amely injektív és az inverze is folytonosan differenciálható, továbbá R elemi részsokaság V-ben és $R\cap \subset \mathrm{Dom}F \neq \emptyset$, akkor $F[R]$ is elemi részsokaság; ha p az R paraméterezése, akkor $F \circ p$ az $F[R]$ paraméterezése (F deriváltja minden pontban lineáris bijekció).

13.2. Definíció *Egy $R \subset V$ nem üres halmazt*

*(i) M **dimenziós részsokaságnak** hívunk, ha minden $x \in R$ pontnak létezik olyan $G(x) \subset V$ környezete, hogy $R \cap G(x)$ M dimenziós elemi részsokaság,*

*(ii) M **dimenziós zárt részsokaságnak** hívunk, ha M dimenziós részsokaság és zárt halmaz V-ben,*

*(iii) M **dimenziós peremes részsokaságnak** hívunk ha M dimenziós részsokaság, és a határa (pereme), $\partial R := \overline{R} \setminus R$ $M-1$ dimenziós részsokaság.*

*Az egy dimenziós részsokaságokat **görbéknek**, a két dimenziósokat **felü leteknek**, az $N-1$ dimenziósokat **hiperfelületeknek** hívjuk.*

A peremes görbe itteni definíciója egy kissé eltér a korábbitól; ott ugyanis a perem is hozzátartozott a peremes görbéhez, itt viszont nem. Valójában ez lényegtelen kölönbség, csak $M > 1$ esetén nehézkes volna – de lehetne – úgy definiálni a peremes részsokaságot, hogy a pereme is hozzátartozzon.

Megjegyzések (i) Egy elemi részsokaság egyben részsokaság is, hiszen minden x pontjára $G(x) := V$ teljesíti a definíció feltételét.

(ii) Ha U véges dimenziós vektortér és $F : V \rightarrowtail U$ olyan függvény, amely folytonosan differenciálható, injektív, az inverze is folytonosan differenciálható, továbbá R részsokaság V-ben és $R\cap \subset \mathrm{Dom}F \neq \emptyset$, akkor $F[R]$ is részsokaság.

13.3. Lássunk néhány példát!

(i) Az $r > 0$ sugarú gömbhéj $\mathbb{R}^N$-ben, azaz $\{x \in \mathbb{R}^N \mid |x| = r\}$ zárt hiperfelület. Ugyanis zárt halmaz, és ha x az eleme, akkor valamelyik koordinátája nem nulla, legyen például $x_N > 0$. Ekkor az x egy környezetében

$$\mathbb{R}^{N-1} \rightarrowtail \mathbb{R}^N, \quad (\xi_1, \ldots, \xi_{N-1}) \mapsto \left(\xi_1, \ldots, \xi_{N-1}, \sqrt{r^2 - \sum_{k=1}^{N-1} \xi_k^2} \right)$$

a gömbhéj paraméterezése.

(ii) Az $r > 0$ sugarú fél gömbhéj $\mathbb{R}^N$-ben, azaz $\{x \in \mathbb{R}^N \mid |x| = r, x_N > 0\}$ peremes hiperfelület. Pereme az $\{x \in \mathbb{R}^N \mid |x| = r,\ x_N = 0\}$ halmaz, amely $N-2$ dimenziós részsokaság, hiszen "lényegében" az $\mathbb{R}^{N-1}$-beli r sugarú gömbhéj.

13.4. Állítás *Ha T M dimenziós vektortér, U véges dimenziós vektortér, $u : T \rightarrowtail U$ összefüggő halmazon értelmezett folytonosan differenciálható, akkor u grafikonja M dimenziós elemi részsokaság $T \times U$-ban, amelynek paraméterezését egy $A : \mathbb{R}^M \to T$ lineáris bijekcióval adhatjuk meg $p : A^{-1}(\mathrm{Dom}u) \to T \times U$, $\xi \mapsto (A\xi, u(A\xi))$ alakban.*

Bizonyítás Az állításban szereplő p teljesíti a paraméterezésre kirótt feltételeket, hiszen
– $\mathrm{Dom}(p)$ összefüggő és nyílt,
– p folytonosan differenciálható,
– $\mathrm{D}p(\xi) = (A, \mathrm{D}u(A\xi)A)$ injektív minden $\xi \in \mathrm{Dom}(p)$ esetén;
– p injektív, és $p^{-1} = A^{-1}\mathrm{pr}_T$ folytonos. ■

Természetesen, ha elhagyjuk azt a feltételt, hogy u értelmezési tartománya összefüggő, akkor úgy igaz az állítás, hogy u grafikonja M dimenziós részsokaság.

13.5. Állítás *Legyen W $N-M$ dimenziós valós vektortér, $S : V \rightarrowtail W$ folytonosan differenciálható és $c \in \mathrm{Ran}(S)$. Ha $\mathrm{D}S(x) \in \mathrm{Lin}(V, W)$ ráképezés minden $x \in \overset{-1}{S}\{c\}$ pontban, akkor $\overset{-1}{S}\{c\}$ M dimenziós részsokaság.*

Bizonyítás Legyen $a \in \overset{-1}{S}\{c\}$ tetszőleges. Mivel $\dim(V) = N$ és $\mathrm{D}S(a) \in \mathrm{Lin}(V, W)$ ráképezés, $T := \mathrm{Ker}(\mathrm{D}S(a)) \subset V$ M dimenziós lineáris altér. Ha U ennek egy kiegészítő altere, azaz $T \cap U = \{0\}$, $T + U = V$, akkor $\dim(U) = N - M$, továbbá $\mathrm{D}S(a)|_U$ bijekció U és W között.

A V minden eleme előáll egyértelműen egy T-beli és egy U-beli elem összegeként, azaz $T \times U \to V$, $(t, u) \mapsto t + u$ lineáris bijekció. Ennek segítségével az általánosság megszorítása nélkül vehetjük úgy, hogy $V = T \times U$. Ekkor $\mathrm{D}S(a)|_U$ bijektív volta pontosan azt jelenti, hogy S-nek a második változó szerinti parciális deriváltja a-ban bijektív. Alkalmazhatjuk tehát az implicitfüggvény-tételt, miszerint létezik olyan $G(a) \subset V$ (összefüggő) nyílt környezete az a pontnak, hogy $G(a) \cap \overset{-1}{S}\{c\}$ egy $u_c : T \rightarrowtail U$ folytonosan differenciálható függvény grafikonja, azaz minden $t \in T \cap G(a)$ pontra

$$S(t, u_c(t)) = c.$$

A 13.4. szerint beláttuk, hogy $\overset{-1}{S}\{c\}$ tetszőleges pontjának van olyan környezete V-ben, amelynek $\overset{-1}{S}\{c\}$-val vett metszete részsokaság, tehát ez a szinthalmaz valóban részsokaság.

13.6. Az alábbi állítás egyik eredménye az, hogy az előbbinek a fordítottja is igaz abban az értelemben, hogy lokálisan minden részsokaság előáll egy folytonosan differenciálható függvény szinthalmazaként. A másik fontos, a későbbiekben lényegesen kihasznált eredménye, hogy bármely paraméterezés inverze lokálisan egy folytonosan differenciálható függvény leszűkítse.

Állítás *Ha $R \subset V$ M dimenziós részsokaság, $a \in R$ és p az a pont egy környezetében az R paraméterezése, akkor létezik olyan $G(a) \subset V$ környezete az a pontnak, hogy $R \cap G(a) \subset \mathrm{Ran}(p)$, és*

$$H : G(a) \to \mathbb{R}^M \qquad S : G(a) \to \mathbb{R}^{N-M}$$

folytonosan differenciálható függvények, melyekre

$$H \circ p \subset \mathrm{id}_{\mathbb{R}^M}, \qquad S \circ p \subset 0 : \mathbb{R}^M \to \mathbb{R}^{N-M},$$

(azaz $\overset{-1}{S}\{0\} \subset R$), valamint minden $x \in G(a)$ pontban $\mathrm{D}S(x) : V \to \mathbb{R}^{N-M}$ ráképezés.

Bizonyítás Ha $M = 0$, akkor $H := 0$, $S := L \circ (\mathrm{id}_V - a)$, ahol $L : V \to \mathbb{R}^N$ tetszőleges lineáris bijekció, teljesíti a követelményeket; ha pedig $M = N$, akkor $H := p^{-1}$, $S := 0$. A következőkben tehát $0 < M < N$.

Mivel a részsokaságnak csak a paraméterezett részét tekintjük, az általánosság megszorítása nélkül – az egyszerűbb fogalmazás kedvéért – feltehetjük, hogy $R = \mathrm{Ran}(p)$ (azaz R elemi részsokaság).

Mivel p injektív, egyetlen olyan $\alpha \in \mathrm{Dom}(p)$ létezik, hogy $p(\alpha) = a$. $\mathrm{D}p(\alpha) : \mathbb{R}^M \to V$ injektív lineáris leképezés, tehát $\mathrm{Ran}(\mathrm{D}p(\alpha)) =: T$ M dimenziós lineáris altere V-nek. Legyen U ennek egy kiegészítő altere; ekkor $\dim U = N - M$.

Jelölje $Q : V \to V$ az U mentén a T-ra való vetítést, vagyis Q az a lineáris leképezés, amelyre $Q \circ Q = Q$, $\mathrm{Ker}(Q) = U$, $\mathrm{Ran}(Q) = T$ teljesül. Tekintsük a

$$\Phi = Q \circ p : \mathrm{Dom}(p) \to T$$

leképezést. Ekkor $\mathrm{D}\Phi(\alpha) = Q \circ \mathrm{D}p(\alpha)$, és mivel $\mathrm{Ran}(\mathrm{D}p(\alpha)) = T$, amelyen Q az identitás, $\mathrm{D}\Phi(\alpha) = \mathrm{D}p(\alpha) : \mathbb{R}^M \to T$ lineáris bijekció. Ezért az inverzfüggvény-tétel szerint van olyan $D \subset \mathrm{Dom}(p)$ nyílt halmaz, hogy $\alpha \in D$, $\Phi|_D = Q \circ p|_D$ injektív, $\Phi[D] = Q[p[D]] \subset T$ nyílt halmaz, és $(\Phi|_D)^{-1} = (Q \circ p|_D)^{-1} : \Phi[D] \to \mathbb{R}^M$ folytonosan differenciálható.

$p[D]$ nyílt halmaz R-ben (mert p^{-1} folytonos), továbbá része a $\overset{-1}{Q}(Q[p[D]]) \subset V$ nyílt halmaznak, van tehát olyan $G(a) \subset \overset{-1}{Q}(Q[p[D]])$ nyílt halmaz, hogy $p[D] = R \cap G(a)$.

Legyen $L : V \to \mathbb{R}^{N-M}$ olyan lineáris leképezés, melyre $L|_U$ bijekció, és

$$H := (Q \circ p|_D)^{-1} \circ Q|_{G(a)} : G(a) \to \mathbb{R}^M,$$
$$S := L \circ (\mathrm{id}_V - p \circ H) : G(a) \to \mathbb{R}^{N-M}.$$

Mivel $Q[G(a)] \subset Q[p[D]] = \mathrm{Dom}\left((Q \circ p|_D)^{-1}\right)$, H valóban az egész $G(a)$-n értelmezve van. Továbbá $p[D] \subset G(a)$, ezért $H \circ p = (Q \circ p|_D)^{-1} \circ Q \circ p \subset \mathrm{id}_{\mathbb{R}^M}$ amiből $H|_R \subset p^{-1}$ adódik.

Az utolsó egyenlőség miatt $p \circ H|_R \subset \mathrm{id}_R$ is fennáll. Ezért ha $x \in R \cap G(a)$, akkor $S(x) = 0$; mivel $R \cap G(a) \subset \mathrm{Ran}(p)$, ebből $S \circ p \subset 0$ adódik.

Már csak az maradt hátra, hogy az S deriváltjára vonatkozó kijelentésünket igazoljuk.

Ha $x \in G(a)$, akkor $\mathrm{D}H(x) = \left(\mathrm{D}\left(\Phi|_D^{-1}\right)\right)(Qx)Q$, ezért $\mathrm{D}H(x)Q = \mathrm{D}H(x)$, és így

$$\begin{aligned}
\mathrm{D}S(x)Q &= L\Big(\mathrm{id}_V - \mathrm{D}p(H(x))\mathrm{D}H(x)\Big)Q = L\Big(Q - \mathrm{D}p(H(x))\mathrm{D}H(x)\Big) = \\
&= L\Big(\mathrm{id}_V - \mathrm{D}p(H(x))\mathrm{D}H(x)\Big) - L(\mathrm{id}_V - Q) = \\
&= \mathrm{D}S(x) - L(\mathrm{id}_V - Q),
\end{aligned}$$

tehát

$$\mathrm{D}S(x)(\mathrm{id}_V - Q) = L(\mathrm{id}_V - Q).$$

Mivel $\mathrm{Ran}(\mathrm{id}_V - Q) = U$ és $L|_U : U \to \mathbb{R}^{N-M}$ bijekció, a jobb oldal értékkészlete $\mathbb{R}^{N-M}$, vagyis $\mathrm{Ran}(\mathrm{D}S(x))$ sem lehet szűkebb $\mathbb{R}^{N-M}$-nél, azaz $\mathrm{D}S(x) : V \to \mathbb{R}^{N-M}$ ráképezés.

13.7. Állítás *Ha p és q egy $R \subset V$ M dimenziós részsokaság két paraméterezése, akkor $p^{-1} \circ q : \mathbb{R}^M \rightarrowtail \mathbb{R}^M$ folytonosan differenciálható, és minden $x \in \mathrm{Ran}(p) \cap \mathrm{Ran}(q)$ esetén*

$$D(p^{-1} \circ q)(q^{-1}(x)) = \left(\mathrm{D}p(p^{-1}(x))\right)^{-1} \mathrm{D}q(q^{-1}(x)). \qquad (*)$$

Bizonyítás Nyilván igaz ez, ha $p^{-1} \circ q$ az üres függvény. Tegyük tehát föl, hogy $\mathrm{Ran}(p) \cap \mathrm{Ran}(q) \neq \emptyset$, és legyen $a \in \mathrm{Ran}(p) \cap \mathrm{Ran}(q)$ tetszőleges. A 13.6. állítás szerint létezik olyan $G(a)$ környezete az a pontnak és olyan $H : G(a) \to \mathbb{R}^M$ folytonosan differenciálható függvény, hogy $H \circ p \subset \mathrm{id}_{\mathbb{R}^M}$ és $R \cap G(a) \subset \mathrm{Ran}(p)$.

Nyilvánvaló, hogy $G := q^{-1}(G(a)) \subset \mathrm{Dom}(p^{-1} \circ q)$ nyílt halmaz és $(p^{-1} \circ q)|_G = H \circ p \circ (p^{-1} \circ q)|_G = H \circ q|_G$, tehát $(p^{-1} \circ q)|_G$ is folytonosan differenciálható, mert két ilyen függvény kompozíciójaként állítható elő.

Mivel $a \in R$ tetszőleges pontja volt $\mathrm{Ran}(p) \cap \mathrm{Ran}(q)$-nak, $q^{-1}(a)$ tetszőleges pontja a $q^{-1}\big(\mathrm{Ran}(p) \cap \mathrm{Ran}(q)\big) = \mathrm{Dom}(p^{-1} \circ q)$ halmaznak, ezért $p^{-1} \circ q$ folytonosan differenciálható az értelmezési tartománya minden pontjának egy környezetében, tehát folytonosan differenciálható.

Továbbá legyen az egyszerűség kedvéért $\xi := q^{-1}(x)$. Mivel

$$\begin{aligned}\mathrm{D}q(q^{-1}(x)) &= \mathrm{D}q(\xi) = \mathrm{D}(q|_G)(\xi) = \mathrm{D}(p \circ p^{-1} \circ q|_G)(\xi) = \\ &= \mathrm{D}p\big((p^{-1} \circ q)(\xi)\big)\mathrm{D}(p^{-1} \circ q|_G)(\xi) = \mathrm{D}p(p^{-1}(x))\mathrm{D}(p^{-1} \circ q)(\xi),\end{aligned}$$

amiből már azonnal következik, amit akartunk.

Megjegyzés (i) A $(*)$ összefüggés olyan, hogy a $\mathrm{D}p^{-1}(x) := \big(\mathrm{D}p(p^{-1}(x))\big)^{-1}$ formális definícióval érvényben marad a $p^{-1} \circ q$ kompozíció deriváltjára vonatkozó formula.

(ii) Az iménti állítás feltételei teljesen szimmetrikusak p-ben és q-ban, tehát – a szerepüket megcserélve – a $(p^{-1} \circ q)^{-1} = q^{-1} \circ p$ függvény is folytonosan differenciálható. Ebből következik, hogy $p^{-1} \circ q$ deriváltja minden pontban $\mathbb{R}^M \to \mathbb{R}^M$ lineáris bijekció.

Ezek alapján megállapíthatjuk, hogy a dimenzió egy részsokaság "belső" tulajdonsága, vagyis egy M dimenziós részsokaság nem lehet egyben M' dimenziós is, ha $M \neq M'$. Ugyanis az előbbieket elismételhetjük úgy, hogy $\mathrm{Dom}(p) \subset \mathbb{R}^M$ és $\mathrm{Dom}(q) \subset \mathbb{R}^{M'}$, és azt kapjuk, hogy létezik egy $\mathbb{R}^{M'} \to \mathbb{R}^M$ lineáris bijekció.

13.8. Állítás *Legyen p és q egy $R \subset V$ részsokaság két paraméterezése, $x \in \mathrm{Ran}(p) \cap \mathrm{Ran}(q)$. Ekkor*

$$\mathrm{Ran}\big(\mathrm{D}p(p^{-1}(x))\big) = \mathrm{Ran}\big(\mathrm{D}q(q^{-1}(x))\big).$$

Bizonyítás Az előző pont $(*)$ formulája szerint fennáll a $\mathrm{Ran}\big(\mathrm{D}p(p^{-1}(x))\big) \supset \mathrm{Ran}\big(\mathrm{D}q(q^{-1}(x))\big)$ összefüggés. A p és q szerepének cseréjével a másik irányú tartalmazás is igaz. ■

Az iménti állítás biztosítja számunkra, hogy értelmes a következő meghatározás.

Definíció *Egy $R \subset V$ M dimenziós részsokaság $x \in R$ pontbeli* ***érintőterének*** *a $T_x(R) := \mathrm{Ran}\big(\mathrm{D}p(p^{-1}(x))\big) \subset V$ halmazt nevezzük, ahol p az R egy paraméterezése az x egy környezetében.*

Nyilvánvaló, hogy az imént értelmezett érintőtér a V-nek M dimenziós lineáris altere, tekintve, hogy $\mathrm{D}p(p^{-1}(x)) : \mathbb{R}^M \to V$ injektív lineáris leképezés.

Megadhatjuk az érintőtér egy bázisát egy p paraméterezéssel az $\mathbb{R}^M$ $e_1, \ldots, e_M$ standard bázisából :

$$\partial_k p(p^{-1}(x)) = \mathrm{D}p(p^{-1}(x))e_k \qquad (k = 1, \ldots, M).$$

Még egyszer kiemeljük, hogy az érintőtér független a definíciójában szereplő paraméterezés választásától, a megadott bázis azonban függ tőle.

A részsokaság ,,geometriai értelemben" vett érintőjét egy adott pontban (gondoljunk egy felület érintősíkjára $\mathbb{R}^3$-ban) úgy kapjuk meg, hogy eltoljuk az iménti alteret az adott pontba, vagyis a definíció jelöléseit használva az

$$x + T_x(R) \subset V$$

halmazt vesszük. Ez azonban általában már nem lineáris altere V-nek (kivéve, ha $x \in T_x(R)$).

13.9. Ha a részsokaságot lokálisan szinthalmazként adjuk meg – és ezt mindig megtehetjük a 13.6. állítás értelmében –, akkor az érintőtereit igen kényelmesen tudjuk leírni.

Állítás *Legyen* $S : V \rightarrowtail W$ *folytonosan differenciálható,* $c \in \mathrm{Ran}(S)$, *és* $\mathrm{D}S(x) \in \mathrm{Lin}(V, W)$ *ráképezés minden* $x \in \overset{-1}{S}\{c\}$ *esetén. Ekkor az* $\overset{-1}{S}\{c\}$ *részsokaság* x *pontjában az érintőtér* $\mathrm{Ker}(\mathrm{D}S(x))$.

Bizonyítás A szóban forgó részsokaság bármely p paraméterezésére $S \circ p : \mathbb{R}^M \rightarrowtail W$ folytonosan differenciálható függvény és mindenhol a c értéket veszi föl, tehát konstans, így a deriváltja minden pontban az azonosan nulla lineáris leképezés, azaz a részsokaság minden x elemére

$$\mathrm{D}(S \circ p)(p^{-1}(x)) = \mathrm{D}S(x)\mathrm{D}p(p^{-1}(x)) = 0 \in \mathrm{Lin}(\mathbb{R}^M, W).$$

Ebből rögtön következik, hogy

$$\mathrm{Ran}\big(\mathrm{D}p(p^{-1}(x))\big) \subset \mathrm{Ker}(\mathrm{D}S(x)),$$

ami nekünk elég, hiszen mindkét oldalon M-dimenziós altér áll, így szükségképpen egyenlők.

Megjegyés A feltételes szélsőértéknél a feltételre kirótt tulajdonságok pontosan azt jelentik, hogy a függvénynek a szélsőértékét egy részsokaságon keressük. Továbbá a 10.7. állítás arról beszél, mi a kapcsolat a szélsőérték létezése és aközött, milyen definitási tulajdonsággal rendelkezik $F - \lambda \circ S$ második deriváltja a feltételi részsokaság érintőterén.

13.10. Eddig egyetlen részsokaságról volt szó, mint szinthalmazról. Egy függvénynek sok szinthalmaza van; most az ilyen részsokaságoknak egymáshoz való viszonyát kapjuk meg a 13.5. állítás egyszerű módosításával.

Állítás *Legyen* $S : V \rightarrowtail W$ *folytonosan differenciálható, és* $\mathrm{D}S(x) \in \mathrm{Lin}(V, W)$ *ráképezés minden* $x \in \mathrm{Dom}S$ *esetén. Ekkor* $\mathrm{Dom}S$ *minden* a *pontjának van olyan* $G(a)$ *környezete, és olyan* $P : W \times \mathbb{R}^M$ *folytonosan differenciálható függvény, hogy* $P(c, \cdot)$ *az* $\overset{-1}{S}\{c\} \cap G(a)$ *paraméterezése minden* $c \in S[G(a)]$ *esetén.*

Bizonyítás A 13.5. állítás bizonyítása szerint vehetjük úgy, hogy $V = T \times U$, ahol $T = \mathrm{Ker}\mathrm{D}S(a)$, és U ennek egy kiegészítője. A $W \times T \times U \to W$, $(c, t, x) \mapsto S(t, x) - c$ folytonosan differenciálható fügvényt tekintve látjuk, hogy az implicitfügvény-tétel szerint a 13.5. bizonyításában szereplő $(c, t) \to u_c(t)$ függvény folytonosan differenciálható. Ezért, a 13.4. jelöléseivel, $P(c, \xi) := (A\xi, u_c(A\xi))$.

13.10. Véges dimenziós affin térben a részsokaságot formailag ugyanúgy értelmezzük, mint vektortérben, vagyis az eddigi V lehet affin tér a $\mathbf{V}$ vektortér felett. Értelemszerű módosításokkal minden ugyanúgy érvényben marad. A legfontosabb, hogy a részsokaság érintőtere minden pontban az affin tér alatti vektortér egy lineáris altere.

13.11. Feladatok

1. Igazoljuk, hogy ha $S : V \rightarrowtail \mathbb{R}^{N-M}$ folytonosan differenciálható függvény, $c \in \mathrm{Ran}S$, akkor $\{x \in \mathrm{Dom}(S) \mid S(x) = c,\ \mathrm{D}S(x) \text{ ráképezés}\}$ amennyiben nem üres, M dimenziós részsokaság.

2. Legyen $Z \subset V$ K dimenziós, $R \subset V$ M dimenziós részsokaság, $Z \subset R$. Bizonyítsuk be, hogy a Z minden a pontjának létezik olyan $G(a)$ környezete és $G(a)$-nak olyan $P : R^K \times \mathbb{R}^{M-K} \times \mathbb{R}^{N-M} \rightarrowtail G(a)$ paraméterezése, hogy $P(\cdot, 0, 0)$ a $Z \cap G(a)$ paraméterezése, $P(\cdot, \cdot, 0)$ az $R \cap G(a)$ paraméterezése. (Útmutatás: a 13.5. állítás alapján eleve vehetjük úgy, hogy $V = R^M \times \mathbb{R}^{N-M}$ és $R \subset \mathbb{R}^M \times \{0\}$.)

3. Melyek részsokaságok az $\mathbb{R}^N$ alábbi részhalmazai közül? Van-e olyan részhalmazuk, amely részsokaság? Hány dimenziósak?

(i) $\left\{x \in \mathbb{R}^N \;\middle|\; \sum_{k=1}^{N-1} x_k^2 - x_N^2 = 1\right\}$, (ii) $\left\{x \in \mathbb{R}^N \;\middle|\; \sum_{k=1}^{N-1} x_k^2 - x_N^2 = 0\right\}$,

(iii) $\{x \in \mathbb{R}^N \mid x_1x_2 = 1,\ x_2x_3 = 1\}$, (iv) $\{x \in \mathbb{R}^N \mid x_1x_2 = 0,\ x_2x_3 > 1\}$.

(Emlékeztetünk arra, hogy egy $\mathbb{R}^N \to \mathbb{R}^K$ lineáris leképezés (mátrix) akkor ráképezés, ha a rangja K.)

4. Adjuk meg az előző feladatban szereplő részsokaságok érintőterét az alábbi pontokban:

(i) $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}, 0, \ldots, 0\right)$, (ii) $(1, 0, \ldots, 0, 1)$,

(iii) $\left(\frac{1}{2}, 2, \frac{1}{2}, 0, \ldots, 0\right)$,

(iv) $\left(\frac{1}{3}, 3, \frac{1}{2}, 0, \ldots, 0\right)$.

5. Felületek-e $\mathbb{R}^3$-ban az alábbi halmazok adott $\alpha \in \mathbb{R}$ esetén:

(i) $\{(x, y, z) \in \mathbb{R}^3 \mid xy = \alpha z^2\}$, (ii) $\{(x, y, z) \in \mathbb{R}^3 \mid xyz = \alpha\}$,

(iii) $\left\{(x, y, z) \in \mathbb{R}^3 \mid \sqrt{x^2 + z^2} + \sqrt{y^2 + z^2} = \alpha\right\}$,

(iv) $\{(x, y, z) \in \mathbb{R}^3 \mid y = x\,\mathrm{tg}z\}$.

6. Határozzuk meg a következő $\mathbb{R}^3$-beli felületek érintősíkjait az adott a pontban:

(i) $\{(x, y, z) \in \mathbb{R}^3 \mid xy^2 + z^3 = 12\}, \quad a := (1, 2, 2);$
(ii) $\big\{(x, y, z) \in \mathbb{R}^3 \mid (x^2 + y^2 + z^2)^2 = 4(x^2 - y^2 + z^2)^2\big\}, \quad a := (1, 2, 2).$

7. Írjuk föl az

$$\mathbb{R}^2 \to \mathbb{R}^3, \quad (u, v) \mapsto (u + v, u^2 + v^2, u^3 + v^3)$$

paraméterezéssel adott felülethez a $(2, 2, 2)$ pontban fektetett érintősíkot.

8. Igazoljuk, hogy $\{L \in \mathbb{R}^{3\times 3} \mid L^*L = \mathrm{id}_{\mathbb{R}^3}\}$ három dimenziós részsokaság (a csillag a mátrix transzponáltját jelöli). Mi ennek a részsokaságnak az érintője az L pontban?

9. Írjuk át 13.6. "alaptétel" bizonyítását affin térben levő részsokaságra.

14. Differenciálás részsokaságokon

14.1. Ha $M < N$, akkor egy M dimenziós $R \subset V$ részsokaságnak egyetlen pontja sem belső pont, ezért az eredeti értelmezés szerint nem beszélhetünk arról, hogy egy $f : R \rightarrowtail W$ függvény differenciálható.

Definíció *Legyen $R \subset V$ M dimenziós részsokaság. Azt mondjuk, hogy az $f : R \rightarrowtail W$ függvény* ***differenciálható az*** *$a \in \mathrm{Dom}(f)$* ***pontban****, ha minden $p : \mathbb{R}^M \rightarrowtail R$, $a \in \mathrm{Ran}(p)$ paraméterezés esetén az $f \circ p : \mathbb{R}^M \to W$ függvény differenciálható a $p^{-1}(a)$ pontban.*

Értelemszerűen bevezethetjük azokat a fogalmakat, hogy a függvény differenciálható egy halmazon, folytonosan differenciálható, stb.

Ezen értelmezés szerint a részsokaság bármely paraméterezésének az inverze folytonosan differenciálható: ha p paraméterezés akkor minden q paraméterezés esetén $p^{-1} \circ q$ folytonosan differenciálható.

Fontos az alábbi eredmény, amely szerint a differenciálhatósághoz elég, ha a függvény *egyetlen* paraméterezéssel komponálva differenciálható függvényt ad.

Állítás *Ha $f : R \rightarrowtail W$, és létezik az R-nek olyan p paraméterezése, hogy $f \circ p$ differenciálható $p^{-1}(a)$-ban, akkor f differenciálható a-ban.*

Bizonyítás Azt kell megmutatnunk, hogy ekkor bármely q paraméterezésre is $f \circ q$ differenciálható $q^{-1}(a)$-ban. Ez viszont teljesül, hiszen $p^{-1} \circ q$ folytonosan differenciálható, és $f \circ q = f \circ p \circ p^{-1} \circ q$.

14.2. Azt már értelmeztük, mit jelent egy részsokaságon értelmezett függvény differenciálhatósága, de a deriváltját még nem.

Állítás *Legyen $f : R \rightarrowtail W$ differenciálható az $a \in \mathrm{Dom}(f)$ pontban. Ha p és q olyan paraméterezés, hogy $a \in \mathrm{Ran}(p) \cap \mathrm{Ran}(q)$, akkor*

$$\mathrm{D}(f \circ p)(p^{-1}(a)) \left(\mathrm{D}p(p^{-1}(a))\right)^{-1} = \mathrm{D}(f \circ q)(q^{-1}(a)) \left(\mathrm{D}q(q^{-1}(a))\right)^{-1}.$$

Bizonyítás

$$\mathrm{D}(f\circ q)(q^{-1}(a)) = \mathrm{D}(f\circ p\circ(p^{-1}\circ q))(q^{-1}(a)) = \mathrm{D}(f\circ p)(p^{-1}(a))\mathrm{D}(p^{-1}\circ q)(q^{-1}(a));$$

ezután már csak a 13.6. (∗) összefüggést kell alkalmazni, majd $\mathrm{D}q(q^{-1}(a))$ inverzével balról beszorozni az egyenlőséget.

Definíció *Ha $f : R \rightarrowtail W$ differenciálható az $a \in \mathrm{Dom}(f)$ pontban, akkor az a-beli deriváltján*

$$\mathrm{D}f(a) := \mathrm{D}(f \circ p)(p^{-1}(a)) \left(\mathrm{D}p(p^{-1}(a))\right)^{-1} : T_a(R) \to W$$

lineáris leképezést értjük, ahol p tetszőleges olyan paraméterezés, amelyre $a \in \mathrm{Ran}(p)$.

Ezzel a meghatározással érvényben marad az $f \circ p$ kompozíció deriváltjára vonatkozó formula, és a 13.6-ban tett megjegyzésünkből törölhető a ,,formális definícióval".

14.3. Állítás *Ha $F : V \rightarrowtail W$ differenciálható az a pontban, és a eleme az R részsokaságnak, akkor $F|_R$ is differenciálható a-ban, és*

$$\mathrm{D}(F|_R)(a) = \mathrm{D}F(a)|_{T_a(R)}$$

.

Bizonyítás Az R bármely p paraméterezésére $F|_R \circ p = F \circ p$, ezért $F|_R$ differenciálható a-ban. Tovább a definíció szerint

$$\mathrm{D}(F|_R)(a) = \mathrm{D}(F|_R \circ p)(p^{-1}(a)) \left(\mathrm{D}p(p^{-1}(a))\right)^{-1} =$$
$$= \mathrm{D}F(a)\Big(\mathrm{D}p(p^{-1}(a)) \left(\mathrm{D}p(p^{-1}(a))\right)^{-1}\Big).$$

A nagy zárójelen belül $T_a(R)$ identitása áll, és ezzel be is fejeztük a bizonyítást.

14.4. Egy részsokaság paraméterezésétől azt követeltük meg, hogy legyen folytonosan differenciálható. Előfordulhat, hogy a paraméterezés többször is differenciálható, mint az előbbi fejezet feladataiban szereplő részsokaságok esetén is.

Legyen $n \in \mathbb{N} \cup \infty$. Egy részsokságot C^n-részsokaságnak hívunk, ha van n-szer folytonosan differenciálható paraméterezése. Ebben a megfogalmazásban tehát eddig a C^1-részsokaságokat tekintettük. Természetesen egy C^n-részsokaság egyben C^k-részsokaság is minden $1 \leq k \leq n$ esetén.

Ha a sokaság C^n-sokaság, akkor azt is értelmezhetjük minden $1 \leq k \leq n$ esetén, hogy egy $f : R \rightarrowtail W$ függvény k-szor differenciálható (egy pontban, egy halmazon, folytonosan): ha egy n-szer folytonosan differenciálható p paraméterezésre $f \circ p$ k-szor diffrenciálható stb. Azonban $M < N$ esetén az f magasabb rendű deriváltjait csak igen körülményesen lehet értelmezni, és az eddgi fogalmaink nem is elegek hozzá; de ezek a deriváltak általában nem is merülnek fel alkalmazásokban.

14.5. Lapozzunk vissza a feltételes szélsőérték értelmezéséhez. Véges N dimenziós V esetén az S-re kirótt feltétel azt jelenti, hogy $R := \overset{-1}{S}\{0\}$ $N - M$ dimenziós részsokaság. A $\mathrm{D}F(a) = \lambda \mathrm{D}S(a)$ egyenlőségből $\mathrm{Ker}(\mathrm{D}F(a)) \supset \mathrm{Ker}(\mathrm{D}S(a)) = T_a(R)$, azaz $0 = \mathrm{D}F(a)|_{T_a(R)} = \mathrm{D}(F|_R)(a)$ következik. Ez speciális esete a következőnek.

Állítás *Legyen $R \subset V$ M dimenziós részsokaság, $f : R \rightarrowtail \mathbb{R}$ differenciálható az a pontban, és legyen f-nek lokális szélsőértéke a-ban. Ekkor* $\mathrm{D}f(a) = 0$.

Bizonyítás Vegyünk egy p paraméterezést az a környezetében. Az $f \circ p : \mathbb{R}^M \rightarrowtail \mathbb{R}$ függvénynek lokális szélsőértéke van $p^{-1}(a)$-ban, ezért

$$0 = \mathrm{D}(f \circ p)(p^{-1}(a)) = \mathrm{D}f(a)\mathrm{D}p(p^{-1}(a)).$$

Mivel $\mathrm{D}p(p^{-1}(a)) : \mathbb{R}^M \to T_a(R) = \mathrm{Dom}(\mathrm{D}f(a))$ lineáris bijekció, ebből $\mathrm{D}f(a) = 0$ következik.

14.6. Feladatok

1. Bizonyítsuk be, hogy egy nyílt részsokaságon értelmezett függvény differenciálhatóságának eredeti 1.2. definíciója és itteni 14.1. definíciója egyenértékű.

2. Adjunk meg egy R részsokaságot V-ben (például egy gömbfelületet $\mathbb{R}^3$-ban) és egy olyan $F : V \to \mathbb{R}$ függvényt, amely nem differenciálható az $a \in R$ pontban, de $F|_R$ differenciálható a-ban.

15. Részsokaságok irányítása

15.1. Az egy dimenziós elemi részsokaságok, vagyis a görbék tárgyalásakor már beszéltünk az irányításról és be is vezettük az irányított görbe fogalmát. Ebben a fejezetben, az ottani gondolatmenetetet általánosítva, értelmezzük az irányítást tetszőleges részsokaságokra.

Legyen tehát R M dimenziós elemi részsokaság a véges N dimenziós vektortérben, valamint p és q két tetszőleges paraméterezése R-nek. A 13.6. szerint a

$p^{-1} \circ q : \mathbb{R}^M \rightarrowtail \mathbb{R}^M$ leképezés folytonosan differenciálható és a deriváltja minden pontban lineáris bijekció, azaz a derivált mátrix determinánsa sehol sem nulla. A derivált leképezés és a determináns folytonossága miatt az is világos, hogy a derivált mátrix determinánsa nem is válthat előjelet, értelmes tehát a következő meghatározás.

1. Definíció *Egy részsokaság p paraméterezését a q paraméterezéssel* ***azonos (ellentétes) irányításúnak*** *nevezzük, ha a $p^{-1} \circ q$ függvény deriváltjának determinánsa pozitív (negatív).*

Az R részsokaság paraméterezéseinek egy $(p_i)_{i\in I}$ családját ***irányítónak*** *nevezzük, ha $R = \bigcup_{i\in I} \mathrm{Ran}(p_i)$ és minden $i, j \in I$ esetén p_i és p_j azonos irányítású.*

Két irányító paraméterezés-család ***azonosan (ellentétesen) irányít****, ha egyesítésük is irányító (nem irányító).*

Egy részsokaság ***irányítható****, ha van irányító paraméterezés-családja.*

Ha $\mathrm{Ran}(p) \cap \mathrm{Ran}(q) = \emptyset$, akkor p a q-val azonosan és ellentétesen is irányított.

Ha p és q azonos (ellentétes) irányítású paraméterezés és $L : \mathbb{R}^M \to \mathbb{R}^M$ olyan lineáris bijekció, amelynek determinánsa negatív, akkor p és $q \circ L$ ellentétes (azonos) irányítású; $p \circ L$ és $q \circ L$ viszont ismét azonos (ellentétes) irányítású.

Állítás *Azonosan irányítónak lenni ekvivelencia-reláció az irányító paraméterezés-családok halmazán. Ha az irányítható részsokaság összefüggő, pontosan két ekvivalencia osztály van.*

Bizonyítás Legyen $(p_i)_{i\in I}$ irányító paraméter-család és $L \in \mathrm{Lin}(\mathbb{R}^M)$ olyan, hogy $\det(L) < 0$; ekkor $(p_i)_{i\in I}$ és $(p_i \circ L)_{i\in I}$ nem azonosan irányítanak, tehát legalább két ekvivalencia-osztály van.

Legyen $(p_i)_{i\in I}$ és $(q_k)_{k\in K}$ egy összefüggő részsokaság két nem azonosan irányító paraméter-családja. Ekkor

$$H := \{(i, k) \in I \times K \mid \mathrm{Ran}(p_i) \cap \mathrm{Ran}(q_k) \neq \emptyset,\ p_i \text{ és } q_k \text{ ellentétes irányítású}\}$$

nem üres. Tegyük fel, hogy

$$G := \{(j, l) \in I \times K \mid \mathrm{Ran}(p_j) \cap \mathrm{Ran}(q_l) \neq \emptyset,\ p_j \text{ és } q_l \text{ azonos irányítású}\}$$

sem üres. Könnyű látni, hogy

$$\bigcup_{(i,k)\in H} \big(\mathrm{Ran}(p_i) \cap \mathrm{Ran}(q_k)\big), \qquad \bigcup_{(j,l)\in G} \big(\mathrm{Ran}(p_j) \cap \mathrm{Ran}(q_l)\big)$$

széteső nem üres halmazok, amelyek egyesítése a részsokaság, ami ellentmondás. Következésképpen minden q_k ellentétes irányítású minden p_i-vel, ami épp azt jelenti, hogy két ekvivalencia-osztály van.

2. Definíció ***Irányított részsokaságnak*** *nevezünk egy* (R, o) *párt, ahol* R *irányítható részsokaság,* o *pedig, az* ***irányítás****, az* R *irányító paraméterezés-családok ekvivalencia-osztálya közül az egyik. Egy irányított részsokaság valamely paraméterezését* ***pozitív (negatív) irányításúnak*** *mondjuk, ha ebbe (nem ebbe) a kijelölt ekvivalencia-osztályba tartozik.*

Szokásosan a jelölésből elhagyjuk az irányítást; ha azt mondjuk, R irányított részsokaság, odaértjük mellé az irányítását is.

15.2. Ha p és q az R azonos irányítású paraméterezése, $x \in \mathrm{Ran}(p) \cap \mathrm{Ran}(q) \neq \emptyset$, akkor

$$(\partial_k p(p^{-1}(x)) \mid k = 1, \ldots, M) \quad \text{és} \quad (\partial_k q(q^{-1}(x)) \mid k = 1, \ldots, M)$$

(lásd 13.7.) a $T_x(R)$ azonos irányítású rendezett bázisa (Analízis II.23.1.), hiszen a bázisok közötti áttérés mátrixa éppen $\big(\mathrm{D}p(p^{-1}(x))\big)^{-1}\mathrm{D}q(q^{-1}(x))$, a $p^{-1} \circ q$ deriváltja a $q(x)$ pontban.

Egy részsokaság irányítása tehát meghatározza a részsokaság minden érintőterének egy irányítását.

15.3. Vegyünk $\mathbb{R}^3$-ban egy felületet. Ennek irányíthatósága szemléletesen azt jelenti, hogy meg lehet különböztetni a felület két oldalát (mint egy terítő színét és visszáját): "az egyik oldalon mozogva sosem juthatunk át a másikra". A Möbius-szalag például nem irányítható: egy pontból elindulva "folytonos úton eljuthatunk a ponttal szemközti, ellentétes oldalon lévő pontba". E szemléletes képnek adunk most pontos matematikai értelmet.

Ha az R felület $\mathbb{R}^3$-ben irányítható, akkor megadható egy $n : R \to \mathbb{R}^3$ normálvektor-függvény, azaz olyan folytonos leképezés, hogy minden $x \in R$ esetén $n(x) = 1$ és $n(x)$ merőleges a felület x-beli érintőterére.

Legyen ugyanis $p : \mathbb{R}^2 \rightarrowtail \mathbb{R}^3$ az R egy paraméterezése, $x \in \mathrm{Ran}(p)$. Ekkor $\partial_1 p(p^{-1}(x))$ és $\partial_2 p(p^{-1}(x))$ az x-beli érintőtér bázisát alkotják, ezért

$$\mathrm{Ran}(p) \to \mathbb{R}^3, \quad x \mapsto \frac{\partial_1 p \times \partial_2 p}{|\partial_1 p \times \partial_2 p|}(p^{-1}(x))$$

folytonos leképezés, amely mindenütt merőleges az érintőtérre.

Ha $x \in \mathrm{Ran}(p) \cap \mathrm{Ran}(q)$ és q a p-vel azonos irányítású paraméterezés, akkor

$$\frac{\partial_1 p \times \partial_2 p}{|\partial_1 p \times \partial_2 p|}(p^{-1}(x)) = \frac{\partial_1 q \times \partial_2 q}{|\partial_1 q \times \partial_2 q|}(q^{-1}(x)).$$

Ez abból látható, hogy a $\phi = p^{-1} \circ q$, $\xi := q^{-1}(x)$ jelöléssel 13.6. $(*)$ szerint $\mathrm{D}q(\xi) = \mathrm{D}p(\phi(\xi))\mathrm{D}\phi(\xi)$, azaz

$$\partial_1 q(\xi) = \partial_1 p(\phi(\xi))\partial_1\phi_1(\xi) + \partial_2 p(\phi(\xi))\partial_1\phi_2(\xi),$$

$$\partial_2 q(\xi) = \partial_1 p(\phi(\xi))\partial_2\phi_1(\xi) + \partial_2 p(\phi(\xi))\partial_2\phi_2(\xi),$$

amiből a vektoriális szorzás tulajdonságait kihasználva egyszerű (de hosszadalmas) számolással adódik, hogy

$$(\partial_1 q \times \partial_2 q)(q^{-1}(x)) = \det\left(\mathrm{D}\phi(q^{-1}(x))\right)(\partial_1 p \times \partial_2 p)(p^{-1}(x)).$$

Ezért ha $(p_i)_{i\in I}$ irányító paraméterezés-család, akkor

$$x \mapsto \frac{\partial_1 p_i \times \partial_2 p_i}{|\partial_1 p_i \times \partial_2 p_i|}(p_i^{-1}(x)) \qquad \text{ha } x \in \mathrm{Ran}(p_i)$$

jól definiált, a kívánt tulajdonságú leképezés.

Eszerint szoktak beszélni például arról, hogy egy gömbfelületet “kifelé” illetve “befelé” irányítunk: ha egy irányító paraméterezés-család egy p tagjára $(\partial_1 p \times \partial_2 p)(p^{-1}(x))$ a gömbfelületből kifelé (az x vektorral azonos irányba) illetve befelé (az x vektorral ellentétes irányba) mutat (persze itt most úgy képzeljük, hogy a gömbfelület az origó köré van írva).

15.4. A vektoriális szorzás $\mathbb{R}^3$ – illetve a három dimenziós irányított euklideszi terek – sajátja, amelyet a vektorok külső (antiszimmetrikus) szorzatából származtatunk.

Általában, ha R M dimenziós irányítható részsokaság V-ben, akkor megadható rajta $\bigwedge^M V$ értékű seholsem nulla folytonos leképezés például olyan módon, hogy választunk egy akármilyen normát és ezzel egy paraméterezéshez definiáljuk a

$$\frac{\bigwedge_{k=1}^{M} \partial_k p}{\left|\bigwedge_{k=1}^{M} \partial_k p\right|} \circ p^{-1}$$

leképezést. A nevező seholsem nulla, mert lineárisan független vektorok külső szorzatát vettük.

Ismerjük a külső szorzás következő tulajdonságait: ha $a_1, \ldots, a_M \in V$ és $L : S : \mathbb{R}^M \to \mathbb{R}^M$ lineáris leképezés (azaz $M \times M$-es mátrix), akkor

$$\bigwedge_{k=1}^{M} \sum_{i=1}^{M} L_{ik} a_i = \det(L) \bigwedge_{k=1}^{M} a_k.$$

Ennek alapján azonos irányítású p és q paraméterezésre a fenti formula ugyanazt a leképezést határozza meg p és q értékkészletének közös részén, így ezek a összeilleszthetők az egész irányítható részsokaságon adott folytonos függvénnyé.

15.5. Ha (V, B, h) irányított pszeudo-euklideszi vektortér, akkor, mint tudjuk, megadható egy $J : \bigwedge^{N-1} V \equiv V \otimes (\overset{N-2}{\otimes} B)$ azonosítás a Levi-Civita tenzor segítségével (Analízis II.30.10.).

Ezért, ha R olyan irányítható hiperfelület V-ben, hogy egyetlen érintőterében sincs izotróp vektor, akkor egy p paraméterezésével a $h := J \circ \left(\bigwedge_{k=1}^{N-1} \partial_k p\right) \circ p^{-1}$ jelöléssel

$$\frac{h}{\sqrt{|b \circ (h,h)|}} : R \to \frac{V}{B}$$

folytonos úgynevezett normálvektor-függvény adható meg.

Ez az általánosítása az $\mathbb{R}^3$ esetében tárgyalt normálvektor-függvénynek.

15.6. Feladatok

1. Mutassuk meg, hogy a gömbfelület $\mathbb{R}^3$-ban (általában $\mathbb{R}^N$-ben) irányítható.
2. Minden nem nulla dimenziós elemi részsokaság irányítható.
3. Tudjuk, hogy az

$$\mathbb{R}^{1+N} \times \mathbb{R}^{1+N} \to \mathbb{R}, \quad (\xi, \eta) \mapsto (\xi|\eta) := -\xi_0\eta_0 + \sum_{k=1}^{N} \xi_k\eta_k$$

bilineáris leképezés Minkowski-forma (Lorentz-forma) $\mathbb{R}^{1+N}$-en. Mutassuk meg, hogy

$$\{\xi \in \mathbb{R}^{1+N} \mid (\xi|\xi) = -1, \xi_0 > 0\}$$

összefüggő, N-dimenziós részsokaság (hiperfelület) $\mathbb{R}^{1+N}$-ben. Írjuk le egy ξ pontjához tartozó érintőterét. Adjuk meg egy irányítását.

Tekintsük először az $(\mathbb{R}^{1+N}, \mathbb{R}, (.|.))$ standard módon irányított Minkowski-féle vektortér részhalmazának, majd az $(\mathbb{R}^{1+N}, \mathbb{R}, < .,. >)$ standardul irányított euklideszi vektortér részhalmazának. Mindkét esetben adjuk meg normálvektorfüggvényét. Vizsgáljuk meg külön az $N = 1$ esetet, szemléltessük ábrán.

16. Görbe vonalú koordinátázás

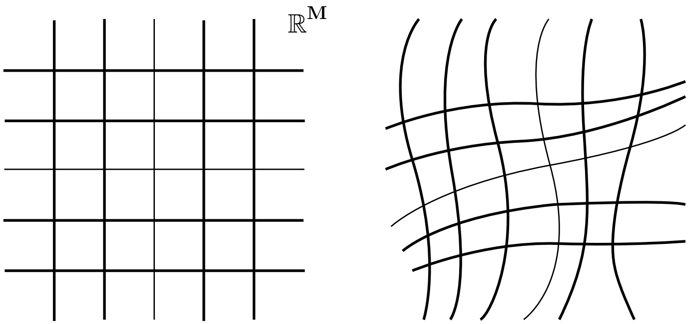
16.1. Legyen R $M > 0$ dimenziós részsokaság a véges N dimenziós valós V vektortérben. Vegyük az R egy p paraméterezését. Ennek inverze, p^{-1} folytonosan differenciálható a 14.1. értelmében; ezt a részsokaság egy **koordinátázásának**

hívjuk. Ha $x \in R$, akkor $p^{-1}(x) \in \mathbb{R}^M$ az x **koordinátái** az adott koordinátázásban. A koordinátázás lényege, hogy a részsokaság elemeit szám M-esekkel reprezentátjuk.

Jelölje, mint szokásosan, $e_1, \ldots, e_M$ az $\mathbb{R}^M$ standard bázisát. Legyen $a \in R$ és $\alpha : p^{-1}(a)$. Minden $k = 1, \ldots, M$ esetén $p(\alpha + \mathbb{R}e_k)$ az a-n átmenő görbe, amelyet az a-n **áthaladó k-adik koordinátavonalnak** hívunk. Ennek egy paraméterezése: $p_k : \mathbb{R} \to V,\quad t \mapsto p(\alpha + te_k)$. A koordinátavonal érintővektora az a pontban $\dot{p}_k(0) = \mathrm{D}p(\alpha + te_k)e_k = \partial_k p(\alpha)$; ez eleme a részsokaság a-beli értintőterének. Az a-n áthaladó különböző koordinátavonalak **transzverzálisak**, ami azt jelenti, hogy érintőjük nem párhuzamos; valóban, az a-n áthaladó koordinátavonalak érintővektorai bázist alkotnak $T_a(R)$-ben (lásd 13.7.), amelyet az a-**beli érintőbázisnak** nevezünk.

Az R minden pontján M darab koordinátavonal halad át, a kordinátavonalak "behálózzák" a részsokaságot, amint azt a 3. ábra szemlélteti.

Megemlítjük, hogy C^n-részsokaság esetén mindig n-szer folytonosan differenciálható paraméterezést (koordinátázást) tekintünk.



3. ÁBRA

16.2. Ha R egy (E, D, b) euklideszi tér részsokasága, akkor az R egy koordinátázását **ortogonálisnak** nevezzük, ha a koordinátavonalak minden pontban ortogonálisan metszik egymást, azaz minden pontban a koordinátázás meghatározta érintőbázis ortogonális. Ekkor szokás az érintőbázisok elemeit "normálni", vagyis elosztani a hosszukkal, és az így kapott $\dfrac{E}{D}$-beli ortonormált bázist használni az eredeti érintőbázis helyett.

Tehát ha p az R ortogonális paraméterezése, $a = p(\alpha)$, és vesszük a

$$T(\alpha) : E \to \frac{E}{D}, \quad \partial_k p(\alpha) \mapsto \frac{\partial_k p(\alpha)}{|\partial_k p(\alpha)|} \quad (k = 1, \dots, M)$$

formulával megadott lineáris bijekciót, akkor

$$Z(\alpha) := T(\alpha)\mathrm{D}p(\alpha) : \mathbb{R}^N \to \frac{E}{D}$$

ortogononális bijekció; tehát $\mathrm{D}p(\alpha)$ a $T(\alpha)^{-1}$ és az ortogonális $Z(\alpha)$ kompozíciója.

16.3. Különleges az N dimenziós részsokaságok esete, amikor tehát R a V nyílt részhalmaza.

Ekkor bármely $K : V \to \mathbb{R}^M$ lineáris bijekció esetén $K|_R$ koordinátázás, azaz $(K|_R)^{-1}$ az R paraméterezése. Ez **egyenes vonalú koordinátázás**, mert a koordinátavonalak egyenesek. Nyílt részsokaságok esetén többnyire egyenesvonalú koordinátázást használunk, de nem mindig.

Mivel a nyílt részhalmazok C^∞-részsokaságok, mindig végtelenszer differenciálható koordinátázásokat tekintünk. Az R nyílt részhalmaz egy $P : \mathbb{R}^N \to R$ paraméterezése a definíció szerint olyan, hogy

(i) $\mathrm{Dom}(P)$ összefüggő és nyílt, $\mathrm{Ran}(P) = R$,

(ii) P végtelenszer differenciálható,

(iii) $\mathrm{D}P$ minden pontban injektív,

(v) P injektív és az inverze folytonos.

Az inverzfüggvény-tétel szerint az (ii) és (iii) feltételből már következik, hogy lokálisan injektív és a lokális inverze is végtelenszer differenciálható (lásd 8.7.), tehát az (v)-ből elhagyható az inverz folytonosságának követelménye. Az R összefüggő nyílt halmaz (görbe vonalú) koordinátázása tehát olyan $K : V \rightarrowtail \mathbb{R}^N$ leképezés, amely

(1) végtelenszer differenciálható,

(2) a deriváltja minden pontban injektív,

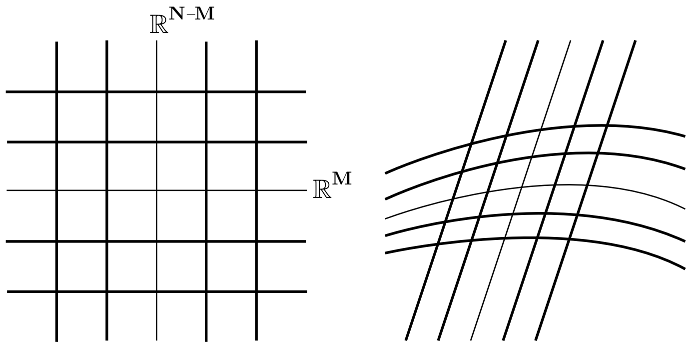
(3) injektív.

A 13.4-ben szereplő $K := (H, S) : V \rightarrowtail \mathbb{R}^M \times \mathbb{R}^{N-M}$ leképezés a $G(a)$ koordinátázása; az inverzét (paraméterezést) 13.9-ben írtuk le. Az ennek megfelelő koordinátavonalakat a 4. ábra szemlélteti.

16.4. A leggyakrabban alkalmazott görbe vonalú koordinátázások a kétdimenziós vektorterek polár-koordinátázása és a háromdimenziós vektorterek henger- valamint gömb-koordinátázása. Ezeket $\mathbb{R}^2 \rightarrowtail \mathbb{R}^2$, illetve $\mathbb{R}^3 \rightarrowtail \mathbb{R}^3$ görbevonalú koordinátázásaként adjuk meg, és általában úgy használjuk, hogy komponáljuk őket egy $V \to \mathbb{R}^2$, illetve $V \to \mathbb{R}^3$ lineáris koordinátázással.

(1) A polárkoordinátázás:

$$K : \mathbb{R}^2 \setminus (\mathbb{R}_0^- \times \{0\}) \to \mathbb{R}^+ \times]-\pi, \pi[,$$



4. ÁBRA

$$x = (x_1, x_2) \mapsto \left(|x|, \operatorname{sign}(x_2) \arccos \frac{x_1}{|x|} \right),$$

amelynek inverze

$$P : \mathbb{R}^+ \times]-\pi, \pi[\to \mathbb{R}^2 \setminus (\mathbb{R}_0^- \times \{0\}),$$
$$(r, \varphi) \mapsto (r \cos \varphi, r \sin \varphi).$$

Könnyen megkapható, hogy

$$\mathrm{D}P(r, \varphi) = \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix},$$

$$\det(\mathrm{D}P(r, \varphi)) = r.$$

A polárkoordinátázás ortogonális, és

$$\mathrm{D}P(r, \varphi) = T^{-1}(r) Z(\varphi),$$

ahol

$$Z(\varphi) := \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix},$$

ortogonális mátrix, és

$$T(r) := \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{r} \end{pmatrix}.$$

(2) A hengerkoordinátázás:

$$K : \mathbb{R}^3 \setminus (\mathbb{R}_0^- \times \{0\} \times \mathbb{R}) \to \mathbb{R}^+ \times]-\pi, \pi[\times \mathbb{R},$$

$$x = (x_1, x_2, x_3) \mapsto \left(\sqrt{x_1^2 + x_2^2}, \operatorname{sign}(x_2 \arccos \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, x_3 \right),$$

amelynek inverze

$$P : \mathbb{R}^+ \times]-\pi, \pi[\times \mathbb{R} \to \mathbb{R}^3 \setminus (\mathbb{R}_0^- \times \{0\} \times \mathbb{R},$$

$$(\rho, \varphi, z) \mapsto (\rho \cos\varphi, \rho \sin\varphi, z).$$

Ennek a deriváltja

$$\mathrm{D}P(\rho, \varphi, z) = \begin{pmatrix} \cos\varphi & -\rho\sin\varphi & 0 \\ \sin\varphi & \rho\cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$\det(\mathrm{D}P(\rho, \varphi, z)) = \rho.$$

A hengerkoordinátázás ortogonális, és

$$\mathrm{D}P(\rho, \varphi, z) = T^{-1}(\rho) Z(\varphi),$$

ahol

$$Z(\varphi) := \begin{pmatrix} \cos\varphi & -\sin\varphi & 0 \\ \sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

ortogonális mátrix, és

$$T(\rho) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{\rho} & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(3) A gömbkoordinátázás:

$$K : \mathbb{R}^3 \setminus (\mathbb{R}^+ \times \{0\} \times \mathbb{R}) \to \mathbb{R}^+ \times]0, \pi[\times]-\pi, \pi[,$$

$$x = (x_1, x_2, x_3) \mapsto \left(|x|, \arccos \frac{x_3}{|x|}, \operatorname{sign}(x_2) \arccos \frac{x_1}{\sqrt{x_1^2 + x_2^2}} \right),$$

amelynek inverze

$$P : \mathbb{R}^+ \times]0, \pi[\times]-\pi, \pi[\to \mathbb{R}^3 \setminus (\mathbb{R}^+ \times \{0\} \times \mathbb{R}),$$

$$(r, \theta, \varphi) \mapsto (r \sin\theta \cos\varphi, r \sin\theta \sin\varphi, r \cos\theta).$$

Ennek a deriváltja

$$\mathrm{D}P(r, \theta, \varphi) = \begin{pmatrix} \sin\theta cos\varphi & r\cos\theta\cos\varphi & -r\sin\theta\sin\varphi \\ \sin\theta\sin\varphi & r\cos\theta\sin\varphi & r\sin\theta\cos\varphi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix},$$

$$\det(\mathrm{D}P(r,\theta,\varphi) = r^2 \sin\theta.$$

A gömbkoordinátázás ortogonális, és

$$\mathrm{D}P(r,\theta,\varphi) = T^{-1}(r,\theta)Z(\theta,\varphi),$$

ahol

$$Z(\theta,\varphi) := \begin{pmatrix} \sin\theta cos\varphi & \cos\theta\cos\varphi & -\sin\varphi \\ \sin\theta\sin\varphi & \cos\theta\sin\varphi & \cos\varphi \\ \cos\theta & -\sin\theta & 0 \end{pmatrix}$$

ortogonális mátrix, és

$$T(r,\theta) := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{r} & 0 \\ 0 & 0 & \frac{1}{r\sin\theta} \end{pmatrix}.$$

16.5. Feladatok

1. Szemléltessük a 16.4-ben adott koordinátázások koordinátavonalait! Adjuk meg e koordinátázosokban tetszőleges $\mathbb{R}^2 \setminus (\mathbb{R}_0^- \times \{0\})$-beli illetve $\mathbb{R}^3 \setminus (\mathbb{R}_0^- \times \{0\} \times \mathbb{R})$-beli ponthoz tartozó érintőbázist!

2. Melyek azok a halmazok, amelyeknek polárkoordinátás alakja
(i) $\{(r,\varphi) \in \mathbb{R}^+ \times]-\pi,\pi[\mid r = \varphi\}$, (ii) $\{(r,\varphi) \in \mathbb{R}^\times]-\pi,\pi[| \; r = \sin\varphi\}$
(iii) $(r,\varphi) \in \{\mathbb{R}^+ \times]-\pi,\pi[\mid r = \cos(\pi/2)\}$?

Pontosan arról van szó, hogy ha K jelöli a polárkoordinátázást, akkor azokat a $H \subset \mathbb{R}^2$ halmazokat keressük, amelyekre $K[H]$ a fent megadott halmazok.

3. Adjuk meg $\mathbb{R}^2 \setminus (\mathbb{R}_0^- \times \{0\})$-nek olyan koordinátázását, amelyben a koordinátavonalak az origóból induló félegyenesek, valamint ellipszisek, amelyek tengelyeinek aránya 1 : 2!

4. Koordinátázás-e az $\mathbb{R}^2 \to \mathbb{R} \times \mathbb{R}^+$, $(x,y) \mapsto (x, e^{x^2+y^2})$ leképezés?. Mely részhalmazra leszűkítve koordinátázás? Mik a koordinátavonalak?

VII. A FIZIKÁBAN HASZNÁLT NÉHÁNY FOGALOM

17. Deriválttenzorok

17.1. Legyen V és W véges $N > 1$ illetve M dimenziós vektortér. Ekkor a

$$W \otimes V^* \equiv \mathrm{Lin}(V, W), \qquad (w \otimes p) \equiv \big(v \mapsto w(p|v)\big)$$

azonosítással élünk (Analízis II.24.6.), és általában $n{\in}\mathbb{N}$ esetére

$$W \otimes (\overset{n}{\otimes} V^*) \equiv \mathrm{Lin}^n(V^n, W), \qquad w \otimes (\bigotimes_{k=1}^{n} p_k) \equiv (v_1, \ldots, v_n) \mapsto w \prod_{k=1}^{n} (p_k|v_k).$$

Ha $F : V \rightarrowtail W$ differenciálható függvény, akkor minden $a \in \mathrm{Dom}(F)$ esetén

$$\mathrm{D}F(a) \in \mathrm{Lin}(V, W) \equiv W \otimes V^*,$$

ha n-szer differenciálható, akkor

$$\mathrm{D}^n F(a) \in \mathrm{Lin}^n(V^n, W) \equiv W \otimes (\overset{n}{\otimes} V^*).$$

Ezért szokás a függvény deriváltjait **deriválttenzoroknak** is hívni.

17.2. Látjuk, hogy a D differenciálás formálisan úgy működik, mint egy V-kovektorral való tenzorszorzás. Egy kis formai hiba, hogy a V^*-gal a tenzorszorzás jobbról történik, a D szimbólumot viszont balról írjuk. Ez kellemetlen mert például $q \in W^*$ és $v \in V$ esetén a $(q|\mathrm{D}F(a)v)$ formulában felborul a megszokott rend, q az F-fel, v pedig a D-vel áll kapcsolatban; különösen zavaró és félreértésre adhat okot ez, amikor $W = V^*$ (ami gyakran előfordul fizikai alkalmazásokban). Magasabb rendű deriváltakra már kisebb a félreértés veszélye, hiszen például $(q|\mathrm{D}^2 F(a)(v_1, v_2))$ már egyértelműen mutatja, hogy a rendzett pár D^2-tel áll kapcsolatban.

A félreérthetőséget elkerülhetjük a transzponáltak használatával. Egy $A \in \mathrm{Lin}(V, W) \equiv W \otimes V^*$ trnaszponáltja $A^* \in \mathrm{Lin}(W^*, V^*) \equiv V^* \otimes W$, amelyet $(q|Av) = (A^*q|v) \quad (q \in W^*, v \in V)$, vagy ami ugyanaz, $(w \otimes p)^* = p \otimes w$ $(w \in W, p \in V^*)$ határoz meg.

A differenciálással mint formális tenzorszorzással összhangban bevezetjük a

$$(\mathrm{D} \otimes F)(a) := (\mathrm{D}F(a))^* \in V^* \otimes W$$

jelölést, amivel kiküszöböltük a sorrendi kellemetlenséget; ekkor tehát

$$\mathrm{D}F(a) = \big((\mathrm{D} \otimes F)(a)\big)^*.$$

17.3. A fizikában a téridőt egy megfigyelő szempontjából $\mathbb{R} \times \mathbb{R}^3$-mel szokás reprezentálni; egy kicsit általánosabban $\mathbf{I} \times \mathbf{E}$-vel, ahol $\mathbf{I}$ egy dimenziós vektortér (idő), $\mathbf{E}$ pedig három dimenziós euklideszi tér (fizikai tér).

Az $F : \mathbf{I} \times \mathbf{E} \rightarrowtail W$ differenciálható függvénynek a második változója (a "térváltozó") szerinti parciális deriváltját ∇F-val szokás jelölni. Ennek is sokszor célszerű inkább a transzponáltját használni, amelyre az előzőekhez hasonlóan a $\nabla \otimes F$ jelet alkalmazzuk.

18. Gradiens, divergencia, rotáció

Ebben a fejezetben V véges $N > 1$ dimenziós valós vektorteret jelöl.

18.1. (i) Ha $f : V \rightarrowtail \mathbb{R}$ differenciálható függvény, akkor $a \in \mathrm{Dom}(f)$ esetén $\mathrm{D}f(a) \in \mathrm{Lin}(V, \mathbb{R}) = V^*$. Az f deriváltfüggvénye tehát $V \rightarrowtail V^*$ leképezés. Szokás ezt az f **gradiensének** is nevezni és a $\mathrm{Grad}f$ szimbólummal jelölni.

Általánosabban, ha Z véges dimenziós valós vektortér és $f : V \rightarrowtail Z$ differenciálható függvény, akkor a $\mathrm{D}f : V \rightarrowtail Z \otimes V^*$ deriváltfüggvényére is használatos a $Gradf$ jelölés.

(ii) Ha $F : V \rightarrowtail V$ differenciálható függvény (**vektormező**), akkor $a \in \mathrm{Dom}(F)$ esetén $\mathrm{D}F(a) \in \mathrm{Lin}(V, V) \equiv V \otimes V^*$. Az F derivált függvénye (deriválttenzora) tehát $V \rightarrowtail V \otimes V^*$ leképezés. Tudjuk, hogy adott a $\mathrm{Tr} : \mathrm{Lin}(V, V) \to \mathbb{R}$ lineáris leképezés, a nyom (Analízis II.24.9.). Képezhetjük ezért $\mathrm{D}F(a)$ nyomát:

$$\mathrm{D} \cdot F(a) := Tr(\mathrm{D}F(a)) \in \mathbb{R}.$$

Szokás $x \mapsto \mathrm{D} \cdot F(x)$ függvényt az F **divergenciájának** nevezni és a $\mathrm{div}F$ szimbólummal jelölni.

Általában, ha Z véges dimenziós valós vektortér és $F : V \rightarrowtail Z \otimes V$ differenciálható függvény, akkor $\mathrm{D}F(a) \in Z \otimes V \otimes V^*$, és képezhetjük a $\mathrm{D} \cdot F(x) :=$

$(\mathrm{id}_Z \otimes \mathrm{Tr})\mathrm{D}F(a) \in Z$ mennyiséget. Most is a $V \rightarrowtail Z$, $x \mapsto \mathrm{D} \cdot F(x)$ az F **divergenciája** .

(iii) Ha $S : V \rightarrowtail V^*$ differenciálható függvény (**kovektormező**), akkor $a \in \mathrm{Dom}(S)$ esetén $\mathrm{D}S(a) \in \mathrm{Lin}(V, V^*) \equiv V^* \otimes V^*$. Az S derivált függvénye (deriválttenzora) tehát $V \rightarrowtail V^* \otimes V^*$ leképezés.

A $\mathrm{D}S(a) \in V^* \otimes V^*$ lineáris leképezés transzponáltja $(\mathrm{D} \otimes S)(a) := \big(\mathrm{D}S(a)\big)^* \in V^* \otimes V^*$. Képezhetjük ezért a

$$(\mathrm{D} \wedge S)(a) := (\mathrm{D} \otimes S)(a) - \big((\mathrm{D} \otimes S)(a)\big)^* \in V^* \wedge V^*$$

mennyiséget (antiszimmetrikus kotenzort), amely $\mathrm{D}S(a)$ kétszeres antiszimmetrikus részének negatívja ($\mathrm{D} \otimes S(a)$ kétszeres antiszimmetrikus része).

A $V \rightarrowtail V^* \wedge V^*$, $x \mapsto (\mathrm{D} \wedge S)(x)$ leképezést az S **külső deriváltjának** vagy **Rotációjának** nevezzük; használatosak még a $dS(x) := (\mathrm{D} \wedge S)(x) =: \mathrm{Rot}S(x)$ jelölések is.

Speciálisan vehetjük S szerepére egy $f : V \rightarrowtail \mathbb{R}$ függvény Gradiensét: $S = \mathrm{D}f = \mathrm{Grad}f$. Minthogy kétszer differenciálható függvény második deriváltja szimmetrikus, $\mathrm{D}^2 f$ antiszimmetrikus része nulla, tehát $\mathrm{D} \wedge (\mathrm{D}f) = 0$, más szóval $\mathrm{Rot\ Grad}f = 0$, ha $f : V \rightarrowtail \mathbb{R}$ kétszer differenciálható függvény.

(iv) Jegyezzük meg jól, hogy

– egy $V \rightarrowtail V$ differenciálható függvénynek (vektormezőnek) van divergenciája, és nincs külső deriváltja,

– egy $V \rightarrowtail V^*$ differenciálható függvénynek nincs divergenciája, és van külső deriváltja.

18.2. Legyen most $V = \mathbb{R}^N$. Ekkor a szokásos azonosítással $(\mathbb{R}^N)^* \equiv \mathbb{R}^N$, így az $f : \mathbb{R}^N \rightarrowtail \mathbb{R}$ differenciálható függvény deriváltja (Gradiense) $\mathbb{R}^N \rightarrowtail \mathbb{R}^N$ leképezésnek tekinthető, amelyet a függvény **gradiensvektorának** nevezünk és a grad szimbólummal jelölünk. Amit $\mathbb{R}^N \rightarrowtail \mathbb{R}$ függvények deriváltjairól tudunk, abból azonnal következik, hogy

$$\mathrm{grad}f = (\partial_1 f, \partial_2 f, \ldots, \partial_N f).$$

Fizikai alkalmazásokban mindig a gradiensvektort használják, és egyszerűen gradiensnek hívják.

Az $F : \mathbb{R}^N \rightarrowtail \mathbb{R}^N$ differenciálható függvény deriválttenzora

$$(\partial_i F_k \mid k, i = 1, \ldots, N),$$

így $\mathrm{div}F = \displaystyle\sum_{k=1}^{N} \partial_k F_k$.

Ismét az $(\mathbb{R}^N)^* \equiv \mathbb{R}^N$ azonosítás miatt az $F : \mathbb{R}^N \rightarrowtail \mathbb{R}^N$ függvénynek értelmezhető a külső deriváltja is, és

$$(\mathrm{D} \wedge F)_{ik} = \partial_i F_k - \partial_k F_i \qquad (i, k = 1, \ldots, N).$$

Ebben az esetben tehát elmosódik a különbség a vektormezők és kovektormezők között; ugyanannak a függvénynek van divergenciája is, külső deriváltja is.

Az $(\mathbb{R}^N)^*$ és $\mathbb{R}^N$ közötti azonosítás fontos következménye, hogy egy $f : \mathbb{R}^N \rightarrowtail \mathbb{R}$ kétszer differenciálható függvény deriváltjának (gradiensének) képezhető a divergenciája, amelyet **Laplace-féle deriváltjának** hívunk:

$$\triangle f := \operatorname{div} \operatorname{grad} f =: \mathrm{D} \cdot \mathrm{D} f = \sum_{k-1}^{N} \partial_k^2 f.$$

18.3. Különösen érdekes $\mathbb{R}^3$ esete a külső derivált szempontjából. Ekkor ugyanis $\mathbb{R}^3$ és $\mathbb{R}^3 \wedge \mathbb{R}^3$ – a 3×3-as valós antiszimmetrikus mátrixok összessége – között van egy természetes lineáris bijekció (azonosítás):

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \equiv \begin{pmatrix} 0 & -\alpha_3 & \alpha_2 \\ \alpha_3 & 0 & \alpha_1 \\ -\alpha_2 & \alpha_1 & 0 \end{pmatrix}.$$

Ha j jelöli ennek az azonosításnak az inverzét, akkor egy $F : \mathbb{R}^3 \rightarrowtail \mathbb{R}^3$ differenciálható függvény esetében képezhető a $\operatorname{rot} F := j \circ \mathrm{D} \wedge F = j \circ \operatorname{Rot} F : \mathbb{R}^3 \rightarrowtail \mathbb{R}^3$ **rotációvektor**. A korábbiak alapján könnyedén megkaphatjuk, hogy

$$\operatorname{rot} F = (\partial_2 F_3 - \partial_3 F_2, \partial_3 F_1 - \partial_1 F_3, \partial_1 F_2 - \partial_2 F_1).$$

Fizikai alkalmazásokban mindig ezt a rotációvektort használják, és egyszerűen rotációnak hívják.

A 18.1-ben mondottakat ezzel az új jelöléssel úgy fejezhetnénk ki, hogy ha $f : \mathbb{R}^3 \rightarrowtail \mathbb{R}$ kétszer differenciálható, akkor $\operatorname{rot} \operatorname{grad} f = 0$. Azt sem nehéz ellenőrizni, hogy ha $F : \mathbb{R}^3 \rightarrowtail \mathbb{R}^3$ kétszer differenciálható, akkor $\operatorname{div} \operatorname{rot} F = 0$ szintén fennáll.

18.4.. Az előbb az $(\mathbb{R}^N)^* \equiv \mathbb{R}^N$ úgynevzett standard azonosítást tekitettük, vagyis azt, amelyet a szokásos $< \cdot, \cdot >$ skaláris szorzattal létesítünk, azaz $\alpha \in \mathbb{R}^N$ a $\xi \mapsto < \alpha, \xi > = \sum_{k=1}^{N} \alpha_k \xi_k$ lineáris leképezésnek felel meg, tehát az $(\mathbb{R}^N, \mathbb{R}, < \cdot, \cdot >)$ euklideszi tér által meghatározott azonosításról van szó.

Egy másik fontos eset, amikor az $\mathbb{R}^{1+N}, \mathbb{R}, (\cdot|\cdot))$ Minkowski-féle térrel van dolgunk. Ekkor az $(\mathbb{R}^{1+N})^* \equiv \mathbb{R}^{1+N}$ azonosításban $\alpha \in \mathbb{R}^{1+N}$ a $\xi \mapsto (\alpha|\xi) := -\alpha_0 \xi_0 + \sum_{k=1}^{N} \alpha_k \xi_k$ lineáris funkcionálnak felel meg.

Ebben az esetben egy $F : \mathbb{R}^{1+N} \rightarrowtail \mathbb{R}^{1+N}$ differenciálható függvény divergenciája $-\partial_0 F_0 + \sum_{k=1}^{N} \partial_k F_k$, és egy $f : \mathbb{R}^{1+N} \rightarrowtail \mathbb{R}$ kétszer differenciálható függvény

Laplace-féle deriváltja $-\partial_0^2 f + \sum_{k=1}^{N} \partial_k^2 f$, amelyet az összetéveszthetőség elkerülése végett D'Alembert-deriváltnak szokás nevezni.

18.5. Az előbbi három pontban mondottak általánosítását fogalmazhatjuk meg (V, B, h) pszeudo-euklideszi vektortér esetére. Ekkor ugyanis van egy $V^* \equiv \frac{V}{B \otimes B}$ azonosításunk.

Ennek az azonosításnak a következményeként

– adódik egy $f : V \rightarrowtail \mathbb{R}$ gradiensvektora, $\mathrm{grad}f : V \rightarrowtail \frac{V}{B \otimes B}$

– egy $F : V \rightarrowtail V \equiv (B \otimes B) \otimes V^*$ differenciálható vektormezőnek a divergenciáján kívül a Rotációja (külső deriváltja) is értelmezető, $\mathrm{Rot}F : V \rightarrowtail \frac{V \wedge V}{B \otimes B}$

– egy $S : V \rightarrowtail V^* \equiv \frac{V}{B \otimes B}$ differenciálható kovektormezőnek a külső deriváltján kívül értelmezhető a divergenciája is, $\mathrm{div}S : V \rightarrowtail \frac{\mathbb{R}}{B \otimes B}$.

Ugyancsak értelmezhető egy $f : V \rightarrowtail \mathbb{R}$ kétszer differenciálható függvény Laplace-féle (D'Alembert-féle) deriváltja, $\mathrm{D} \cdot \mathrm{D}f : V \rightarrowtail \frac{\mathbb{R}}{B \otimes B}$.

Az (E, D, b) három dimenziós, irányított euklideszi vektortéren értelmezhető egy differenciálható $S : E \rightarrowtail E^*$ kovektormező (vagy $F : E \rightarrowtail E$ vektormező) rotációvektora az $E^* \wedge E^* \equiv \frac{E^*}{D}$ azonosítás által: $\mathrm{rot}S : E \rightarrowtail \frac{E^*}{D}$ $\left(\mathrm{rot}F : E \rightarrowtail \frac{E}{D}\right)$.

18.6. Feladatok

1. Legyen $G : V \times V \to \mathbb{R}$ szimmetrikus bilineáris leképezés. Mi a következő $V \to \mathbb{R}$ függvények deriváltja:

(i) $x \mapsto G(x, x)$, (ii) $V \to \mathbb{R}$, $x \mapsto \sqrt{G(x, x)}$.

Milyen következtetést vonhatunk le ebből euklideszi vektortér normájának gradiensére ?

2. Mi az id_V függvény divergenciája, külső deriváltja? (Legyünk figyelmesek!)

3. Legyen (V, B, h) pseudo-euklideszi vektortér. Ekkor id_V felfogható $V \to V^* \otimes (B \otimes B)$ leképezésnek. Mi az id_V külső deriváltja? (Vizsgáljuk meg előbb azt az esetet, amikor $V = \mathbb{R}^N$.)

4. Számoljuk ki az $\frac{1}{|\mathrm{id}_{\mathbb{R}^3}|}$ Laplace-deriváltját!

5. Legyenek $f, g : V \rightarrowtail \mathbb{R}$, $F : V \rightarrowtail V$, $S : V \rightarrowtail V^*$ differenciálható függvények. Bizonyítsuk be a következő azonosságokat:

(i) $\mathrm{Grad}(fg) = g\mathrm{Grad}f + f\mathrm{Grad}g$, másképp ugyanez:

$\mathrm{D}(fg) = g\mathrm{D}f + f\mathrm{D}g$,

(ii) $\mathrm{div}(fF) = f\mathrm{div}F + (\mathrm{Grad}f|F)$,

(iii) $\mathrm{Rot}(fS) = f\mathrm{Rot}S + (\mathrm{Grad}f) \wedge S$, másképp ugyanez:
$\mathrm{D} \wedge (fS) = f\mathrm{D} \wedge S + (\mathrm{D}f) \wedge S$.

6. Legyen (E, D, b) három dimenziós irányított euklideszi vektortér, $f : E \rightarrowtail \mathbb{R}$, $F, G : E \rightarrowtail E$ differenciálható függvények. Lássuk be a következő azonosságokat:
(i) $\mathrm{div}(G \times F) = F \cdot \mathrm{rot}G - G \cdot \mathrm{rot}F$,
(ii) $\mathrm{rot}(fF) = f\mathrm{rot}F - F \times \mathrm{grad}f$,
(iii) $\mathrm{rot}(G \times F) = (F \cdot \mathrm{D})G - (G \cdot \mathrm{D})F + G\mathrm{div}F - F\mathrm{div}G$,
(iv) $\mathrm{grad}\ \mathrm{div}F = \mathrm{rot}\ \mathrm{rot}F + \triangle F$,
ahol a pont a b bilineáris leképezésből származtatott "szorzást", $\times$ pedig a vektoriális szorzást jelöli; az (iii)-ben szereplő első két tagot $F \cdot (\mathrm{D} \otimes G) - G \cdot (\mathrm{D} \otimes F)$ alakba is írhatjuk.

19. Differenciálás és koordinátázás

19.1. Ismételjük át, amit V véges N dimenziós, valós vektortér lineáris koordinátázásáról tudunk (Analízis II.16.).

Vegyük a V egy $(v_1, \ldots, v_N)$ rendezett bázisát, és az ennek megfelelő $K : V \to \mathbb{R}^N$ lineáris bijekciót, azaz koordinátázást, és a $P := K^{-1}$ paraméterezést. Az adott bázis $(p^1, \ldots, p^N)$ duálisa bázisa V^*-nak, amely a $\left(K^{-1}\right)^* = P^* : V^* \to \mathbb{R}^N$ koordinátázást határozza meg.

Egy $x \in V$ koordinátái az adott bázisban

$$K(x) = \Big((p^k|x) \mid k = 1, \ldots, N\Big),$$

egy $p \in V^*$ koordinátái pedig

$$P^*(p) = \Big((p|v_k) \mid k = 1, \ldots, N\Big).$$

Egy $A \in \mathrm{Lin}(V, V) \equiv V \otimes V^*$ mátrixa

$$KAP = \Big((p^i|Av_k) \mid i, k = 1, \ldots, N\Big),$$

egy $R \in \mathrm{Lin}(V, V^*) \equiv V^* \otimes V^*$ mátrixa pedig

$$P^*RP = \Big((Rv_i|v_k) \mid i, k = 1, \ldots, N\Big).$$

19.2. Ezekkel a koordinátázosokkal a V-ben értelmezett valós függvényeket, vektormezőket és kovektormezőket az $\mathbb{R}^N$-ben értelemzett valós illetve $\mathbb{R}^N$ értékű függvényekbe visszük át:

$$\begin{aligned}
&f : V \rightarrowtail \mathbb{R}, && \varphi := f \circ P : \mathbb{R}^N \rightarrowtail \mathbb{R},\\
&F : V \rightarrowtail V, && \mathcal{F} := K \circ F \circ P : \mathbb{R}^N \rightarrowtail \mathbb{R}^N,\\
&S : V \rightarrowtail V^*, && \mathcal{S} := P^* \circ S \circ P : \mathbb{R}^N \rightarrowtail \mathbb{R}^N.
\end{aligned}$$

A táblázat jobb oldali oszlopában álló függvényeket nevezzük a bal oldaliak **koordinátázott alakjának**. Az eredeti függvények deriváltjának koordinátázott alakja,

$$\begin{aligned}
&\mathrm{D}f : V \rightarrowtail V^*, && P^*\mathrm{D}f(\circ P) : \mathbb{R}^N \rightarrowtail \mathbb{R}^N,\\
&\mathrm{D}F : V \rightarrowtail V \otimes V^*, && K\mathrm{D}F(\circ P)P : \mathbb{R}^N \rightarrowtail \mathbb{R}^{N\times N},\\
&\mathrm{D}S : V \rightarrowtail V^* \otimes V^*, && P^*\mathrm{D}S(\circ P)P : \mathbb{R}^N \rightarrowtail \mathbb{R}^{N\times N},
\end{aligned}$$

a függvények koordinátázott alakjának a deriváltja pedig

$$\begin{aligned}
\mathrm{D}\varphi &= ((\mathrm{D}f) \circ P)P = P^*\mathrm{D}f(\circ P),\\
\mathrm{D}\mathcal{F} &= K((\mathrm{D}F) \circ P)P = K\mathrm{D}F(\circ P)P,\\
\mathrm{D}\mathcal{S} &= P^*((\mathrm{D}F) \circ P)P = P^*\mathrm{D}F(\circ P)P,
\end{aligned}$$

ahol a szokásnak megfelelően elhagytuk a lineáris leképezések közül a kompozíció jelét, valamint a kifejező írásmód és a tévedések elkerülése érdekében a

$$\mathrm{D}F(\circ P) : \mathbb{R}^N \rightarrowtail V \otimes V^*, \quad \xi \mapsto \mathrm{D}F(P(\xi)) \tag{$*$}$$

stb. jelölést alkalmaztuk.

Megállapíthatjuk tehát, hogy a lineáris koordinátázás felcserélhető a differenciálással: mindegy, hogy előbb koordinátázunk, és azután differenciálunk, vagy előbb differenciálunk, és azután koordinátázunk. A deriváltak koordinátázott alakja konmponensekben kiírva:

$$\begin{aligned}
&\mathrm{D}f && (\partial_i\varphi \mid i = 1, \dots, N),\\
&\mathrm{D}F && (\partial_i\mathcal{F}^k \mid k, i = 1, \dots, N),\\
&\mathrm{D}S && (\partial_i\mathcal{S}_k \mid k, i = 1, \dots, N),\\
&\mathrm{div}F && \sum_{i=1}^{N} \partial_i\mathcal{F}^i,\\
&\mathrm{D} \wedge S && (\partial_k\mathcal{S}_i - \partial_i\mathcal{S}_k \mid k, i = 1, \dots, N).
\end{aligned}$$

Továbbá, ha (V, B, h) pszeudo-euklideszi vektortér, és h-ortogonális koordinátázást használunk – vagyis a koordinátázást meghatározó bázis h-ortogonális –, akkor (és csak akkor) az $f : V \rightarrowtail \mathbb{R}$ függvény Laplace-féle deriváltjának koordinátás kifejezése:

$$\mathrm{D}\cdot\mathrm{D}f \qquad\qquad \sum_{k=1}^{N} \eta_k \partial_k^2 \phi,$$

ahol $\eta_k = -1$ ha $k = 1, \dots, \mathrm{neg}(h)$ és $\eta_k = 1$ ha $k = \mathrm{neg}(h) + 1, \dots, N$.

Ez visszaadja az euklideszi illetve a Minkowski-féle esetben a 18.2-ben illetve a 18.4-ben megismert formulát.

19.3. A V vektorteret – vagy annak egy nyílt részhalmazát – görbe vonalúan is koordinátázhatjuk.

Ekkor a koordinátázás $K : V \rightarrowtail \mathbb{R}^N$ végtelenszer differenciálható leképezés, amely injektív, és a deriváltja minden pontban bijekció; a $P := K^{-1}$ paraméterezés ugyanilyen tulajdonságú.

Minden $x \in \mathrm{Dom}(K)$ esetén az x-en áthaladó koordináta-vonalak érintői bázist alkotnak V-ben, amelyet a K által meghatározott x-beli érintőbázisnak nevezünk. $\mathrm{D}K(x) : V \to \mathbb{R}^N$ adja azt a lineáris bijekciót, amely ezen bázisvektorokat az $\mathbb{R}^N$ standard bázisvektoraiba képezi. Ennek megfelelően $\big(\mathrm{D}K(K^{-1}(\xi))\big)^{-1} = \mathrm{D}P(\xi) : \mathbb{R}^N \to V$ az a lineáris bijekció, amely az $\mathbb{R}^N$ standard bázisvektorait a $P(\xi)$-beli érintőbázisba képezi.

Az $F : \mathrm{Dom}(K) \rightarrowtail V$ vektormezőt úgy koordinátázzuk, hogy az $F(x) \in V$ vektort az x-beli érintőbázisra vonatkozó koordinátáival adjuk meg minden x esetén, x-et pedig a koordinátázás segítségével rendezett szám N-essel reprezentáljuk. Képletben:

$$F : V \rightarrowtail V \quad \text{koordinátázott alakja} \quad \mathbb{R}^N \rightarrowtail \mathbb{R}^N, \quad \xi \mapsto \mathrm{D}K(P(\xi))F(P(\xi)),$$

és hasonló formulát írhatunk fel egy $S : V \rightarrowtail V^*$ kovektormező koordinátázott alakjára: $S(x)$-et az x-beli érintőbázis duálisára vonatkozó koordintátaival adjuk meg. Értelemszerűen így járunk el $V \rightarrowtail V \otimes V^*$ és $V \rightarrowtail V^* \otimes V^*$ tenzormezőkkel is.

A következő táblázatban foglaljuk össze a különféle függvények koordinátázott alakjait, a 19.2-ben bevezetett jelölés olyan további egyszerűsítésével, hogy a paraméterezéssel való kompozícióból elhagyjuk a kompozíció jelét:

$$\begin{aligned}
f &: V \rightarrowtail \mathbb{R}, & \varphi &:= f(P) : \mathbb{R}^N \rightarrowtail \mathbb{R},\\
F &: V \rightarrowtail V, & \mathcal{F} &:= \mathrm{D}K(P)F(P) : \mathbb{R}^N \rightarrowtail \mathbb{R}^N,\\
S &: V \rightarrowtail V^*, & \mathcal{S} &:= (\mathrm{D}P)^* S(P) : \mathbb{R}^N \rightarrowtail \mathbb{R}^N.
\end{aligned}$$

Az eredeti függvények deriváltjának koordinátázott alakja

$$\begin{aligned} &\mathrm{D}f : V \rightarrowtail V^*, && (\mathrm{D}P)^*\mathrm{D}f(P) : \mathbb{R}^N \rightarrowtail \mathbb{R}^N, \\ &\mathrm{D}F : V \rightarrowtail V \otimes V^*, && \mathrm{D}K(P)\mathrm{D}F(P)\mathrm{D}P : \mathbb{R}^N \rightarrowtail \mathbb{R}^{N\times N}, \\ &\mathrm{D}S : V \rightarrowtail V^* \otimes V^*, && (\mathrm{D}P)^*\mathrm{D}S(P)\mathrm{D}P : \mathbb{R}^N \rightarrowtail \mathbb{R}^{N\times N}, \end{aligned}$$

a függvények koordinátázott alakjának a deriváltja pedig (a jelölést nemsokára megmagyarázzuk)

$$\mathrm{D}\varphi = \mathrm{D}f(P)\mathrm{D}P = (\mathrm{D}P)^*\mathrm{D}f(P), \tag{1}$$

$$\mathrm{D}\mathcal{F} = \mathrm{D}^2K(P)(\mathrm{D}P\cdot, F(P)) + \mathrm{D}K(P)\mathrm{D}F(P)\mathrm{D}P, \tag{2}$$

$$\mathrm{D}\mathcal{S} = S(P)\mathrm{D}^2P(\cdot,\bullet) + (\mathrm{D}P)^*\mathrm{D}S(P)\mathrm{D}P. \tag{3}$$

A valós értékű függvények koordinátázott alakjának a deriváltja megegyezik a derivált koordinátázott alakjával, tehát ilyen függvényekre a koordinátázás és a differenciálás görbe vonalú koordinátázás esetén is felcserélhető.

Ezzel szemben vektormezőkre és kovektormezőkre ez nem igaz. A koordinátázott alakok deriváltjának második tagja megegyezik a deriváltak koordinátázott alakjával, az első tag pedig csak akkor azonosan nulla, ha K (illetve P) második deriváltja azonosan nulla. Fejtsük ki egy kicsit részletesebben, mit is jelentenek ezek a tagok.

$D^2K(x) : V \times V \to \mathbb{R}^N$ bilineáris leképezés, és így (2)-ben az első tag pontos értelme adott $\xi \in \mathrm{Dom}(P)$ esetén az

$$\mathbb{R}^N \to \mathbb{R}^N, \quad \omega \mapsto \mathrm{D}^2K(P(\xi))\Big(\mathrm{D}P(\xi)\omega, F(P(\xi))\Big)$$

lineáris leképezés.

$D^2P(\xi) : \mathbb{R}^N \times \mathbb{R}^N \to V$ bilinéáris leképezés, és így (3)-ban az első tag pontos értelme az

$$\mathbb{R}^N \to \left(\mathbb{R}^N\right)^* \equiv \mathbb{R}^N, \quad \omega \mapsto \big(S(P(\xi)) \mid \mathrm{D}^2P(\xi)(\omega,\bullet)\big)$$

lineáris leképezés, ahol $\bullet$ jelöli a változó helyét.

A $\mathrm{D}K(P)\mathrm{D}P = \mathrm{id}_{\mathbb{R}^N}$ összefüggés differenciálásával azt kapjuk, hogy

$$\mathrm{D}^2K(P)(\mathrm{D}P \times \mathrm{D}P) + \mathrm{D}K(P)\mathrm{D}^2P = 0.$$

Az $F(P) = (\mathrm{D}P)\mathcal{F}$ összefüggés alapján a (2) első tagját átalakítjuk úgy, hogy benne a vektormező koordinátázott alakja jelenjen meg:

$$\mathrm{D}^2K(\mathrm{D}P\cdot, F(P)) = \mathrm{D}^2K(P)(\mathrm{D}P \times \mathrm{D}P)(\cdot, \mathcal{F}),$$

és hasonlóan, $S(P) = \mathcal{S}\mathrm{D}P$ alapján a (3) első tagját is, hogy benne csak a kovektormező koordinátázott alakja jelenjen meg. A

$$\Gamma := -\mathrm{D}^2K(P)(\mathrm{D}P \times \mathrm{D}P) : \mathbb{R}^N \to \mathrm{Lin}^2(\mathbb{R}^N \times \mathbb{R}^N, \mathbb{R}^N)$$

úgynevezett **Christoffel-féle szimbólum** bevezetésével végül a

$$\mathrm{D}K(P)\mathrm{D}F(P)\mathrm{D}P = \mathrm{D}\mathcal{F} + \Gamma(\cdot, \mathcal{F}),$$
$$(\mathrm{D}P)^*\mathrm{D}S(P)\mathrm{D}P = \mathrm{D}\mathcal{S} - \mathcal{S}\Gamma(\cdot, \bullet)$$

eredményt kapjuk.

19.4. Az előző pont formuláiban nagyon oda kellett figyelni, mi milyen lineáris vagy bilineáris leképezés, mit mivel azonosítunk.

Jobban áttekinthető formulákat kapunk, ha a $V = \mathbb{R}^N$ esetet veszük.

$\mathbb{R}^N$ vektorainak standard komponenseit felső indexszel írjuk: $x = (x^1, \ldots, x^N) \in \mathbb{R}^N$; $(\mathbb{R}^N)^* \equiv \mathbb{R}^N$ elemeinek standard komponenseit alsó indexszel látjuk el: $a = (a_1, \ldots, a_N) \in (\mathbb{R}^N)^*$.

A komponensek általános jelölésére az $i, j, k, \ldots$ betűket használjuk: $x = (x^i \mid i = 1, \ldots, N)$, $a = (a_i \mid i = 1, \ldots, N)$.

A görbe vonalú koordinátázás értékkészletében lévő szám N-eseket is felül fogjuk indexelni az $\alpha, \beta, \gamma \ldots$ betűkkel. Ennek megfelelően

– egy vektormező $F = (F^i \mid i = 1, \ldots, N) : \mathbb{R}^N \rightarrowtail \mathbb{R}^N$,

– egy kovektormező $S = (S_i \mid i = 1, \ldots, N) : \mathbb{R}^N \rightarrowtail (\mathbb{R}^N)^* \equiv \mathbb{R}^N$,

– a koordinátázás $K = (K^\alpha \mid \alpha = 1, \ldots, N) : \mathbb{R}^N \rightarrowtail \mathbb{R}^N$,

– és a paraméterezés $K^{-1} = P = (P^i \mid i = 1, \ldots, N) : \mathbb{R}^N \rightarrowtail \mathbb{R}^N$

alakú.

Minthogy minden index az $1, \ldots, N$ értékeket futja végig, elhagyjuk ennek jelölését, valamint megállapodunk az úgynevezett Einstein-féle összegzési szabályban, mely szerint az ellentétes helyzetben (lenn-fönn) levő azonos indexek automatikusan 1-től N-ig történő összegzést jelentenek. Tehát például $a_i x^i := \sum_{i=1}^{N} a_i x^i$.

Ennek megfelelelően a következő táblázatban folglaljuk össze az egyes függvények koordinátázott alakjait (a valós értékű függvények esete érdektelen):

F^i	koordinátázottja	$\mathcal{F}^\alpha = \partial_i K^\alpha(P) F^i(P)$,
S_i	koordinátázottja	$\mathcal{S}_\alpha = \partial_\alpha P^i S_i(P)$,
$\partial_k F^i$	koordinátázottja	$(\mathcal{D}\mathcal{F})^\alpha{}_\beta := \partial_i K^\alpha(P) \partial_k F^i(P) \partial_\beta P^k$,
$\partial_k S_i$	koordinátázottja	$(\mathcal{D}\mathcal{S})_{\alpha\beta} := \partial_\alpha P^i \partial_k S_i(P) \partial_\beta P^k$.

Viszont F és S koordinátázottjának a deriváltja

$$\partial_\beta \mathcal{F}^\alpha = \partial_k \partial_i K^\alpha(P) \partial_\beta P^k F^i(P) + \partial_i K^\alpha(P) \partial_k F^i(P) \partial_\beta P^k,$$
$$\partial_\beta \mathcal{S}_\alpha = \partial_\beta \partial_\alpha P^i S_i(P) + \partial_\alpha P^i \partial_k S_i(P) \partial_\beta P^k.$$

A $\partial_i K^\alpha(P) \partial_\gamma P^i = \delta^\alpha{}_\gamma$ és az ebből adódó

$$\partial_k \partial_i K^\alpha(P) \partial_\beta P^k \partial \gamma P^i + \partial_i K^\alpha(P) \partial_\beta \partial \gamma P^i = 0$$

összefüggésből, a

$$\Gamma^\alpha{}_{\beta\gamma} := -\partial_k \partial_i K^\alpha(P) \partial_\beta P^k \partial_\gamma P^i = \partial_i K^\alpha \partial_\alpha \partial_\beta P^i$$

Christoffel-féle szimbólum bevezetésével végül azt kapjuk, hogy

$$(\mathcal{D}\mathcal{F})^\alpha{}_\beta = \partial_\beta \mathcal{F}^\alpha + \Gamma^\alpha{}_{\beta\gamma} \mathcal{F}^\gamma,$$

$$(\mathcal{D}\mathcal{S})_{\alpha\beta} = \partial_\beta \mathcal{S}_\alpha - \Gamma^\gamma{}_{\alpha\beta} \mathcal{S}_\gamma.$$

19.6. Amint az a fentiekből kiderül, a deriváltak általában elég bonyolultak a görbevonalú koordinátázásokban. Kivételt képeznek azok az esetek, amikor a függvények állandók bizonyos koordinátavonalak mentén; ilyenkor előnyös a görbevonalú koordináták használata. Például, ha olyan függvénnyel van dolgunk, amely állandó $\mathbb{R}^3$-ban a gömbfelületeken, akkor célszerű a gömbkoordinátázás; ha a hengerfelületeken, akkor a hengerkoordinátázás.

Kivétel továbbá a kovektormezők külső deriváltja, amelyre az igaz, hogy koordintázott alakja a vektormező koordinátázott alakjának külső deriváltja, vagyis az eddigi jelölések értelemszerű alkalmazásával

$$(\mathcal{D} \wedge \mathcal{S})_{\alpha\beta} = \partial_\alpha \mathcal{S}_\beta - \partial_\beta \mathcal{S}_\alpha.$$

Ez egyszerűen adódik abból, hogy a koordinátázás második deriváltja szimmetrikus, ezért a Christoffel-féle szimbólum is szimmetrikus az alsó indexeiben: $\Gamma^\alpha{}_{\beta\gamma} = \Gamma^\alpha{}_{\gamma\beta}$.

Itt említjük meg, hogy a fizikában hangsúlyozzák, hogy a Christoffel-féle szimbólum nem harmadrendű tenzor, noha három indexe van. Ez azt jelenti, hogy nincs olyan $V \to \mathrm{Lin}^2(V \times V, V^*)$ leképezés (harmad rendű tenzormező), amelynek a koordinátázott alakja volna a Christoffel-szimbólum.

19.7. Alkalmazásokban legtöbbször a Laplace-féle deriváltat állítjuk elő görbevonalú koordinátákban. A hosszas számítások mellőzésével közöljük a formulákat.

Legyen $f : \mathbb{R}^3 \rightarrowtail \mathbb{R}$ kétszer differenciálható függvény. Ekkor, ha P az $\mathbb{R}^3$ henger-paraméterezése, akkor a $\phi := f \circ P$ valamint a 16.3. jelöléseivel

$$(\triangle f)(P(\rho, \varphi, z)) = \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial \phi(\rho, \varphi, z)}{\partial \rho} \right) + \frac{1}{\rho^2} \frac{\partial^2 \phi(\rho, \varphi, z)}{\partial \varphi^2} + \frac{\partial^2 \phi(\rho, \varphi, z)}{\partial z^2},$$

és ha P az $\mathbb{R}^3$ gömb-paraméterezése, akkor

$$(\triangle f)(P(r,\theta,\varphi)) =$$
$$= \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\phi(r,\theta,\varphi)}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\phi(r,\theta,\varphi)}{\partial\theta}\right) + \frac{1}{r^2\sin\theta}\frac{\partial^2\phi(r,\theta,\varphi)}{\partial\varphi^2}.$$

19.8. Feladatok

1. Mi az $\mathbb{R}^2$-beli α szögű forgatás (mint $\mathbb{R}^2 \to \mathbb{R}^2$ vektormező) polárkoordinátás alakja?

2. Adjuk meg az $A : \mathbb{R}^3 \to \mathbb{R}^3$ lineáris leképezés (mint vektormező) henger- és gömbkoordinátázását! Mikor lesznek ezek különösen egyszerűek?

3. Legyen $0 \neq v \in \mathbb{R}^3$. Adjuk meg az $\mathbb{R}^3 \to \mathbb{R}^3$, $x \mapsto |x|v$ vektormező henger- és gömbkoordinátázását!

4. Adjuk meg $\mathbb{R}^2$ polárkoordinátázásának, valamint $\mathbb{R}^3$ henger- és gömbkoordinátázásának a Christoffel-féle szimbólumát!

5. A fejezet eredményei szerint egy $f : \mathbb{R}^3 \rightarrowtail \mathbb{R}$ függvény gradiense koordinátázott alakjának komponensei a függvény koordinátázot alakjának parciális deriváltjai. Tehát hengerkoordinátákban

$$\frac{\partial f(P(\rho,\varphi,z))}{\partial\rho}, \frac{\partial f(P(\rho,\varphi,z))}{\partial\varphi}, \frac{\partial f(P(\rho,\varphi,z))}{\partial z},$$

gömbkoordinátákban pedig

$$\frac{\partial f(P(r,\theta,\varphi))}{\partial r}, \frac{\partial f(P(r,\theta,\varphi))}{\partial\theta}, \frac{\partial f(P(r,\theta,\varphi))}{\partial\varphi}.$$

Fizikakönyvekben azonban azt találjuk, hogy a gradiens komponensei hengerkoordinátákban

$$\frac{\partial f(P(\rho,\varphi,z))}{\partial\rho}, \frac{1}{\rho}\frac{\partial f(P(\rho,\varphi,z))}{\partial\varphi}, \frac{\partial f(P(\rho,\varphi,z))}{\partial z},$$

gömbkoordinátákban pedig

$$\frac{\partial f(P(r,\theta,\varphi))}{\partial r}, \frac{1}{r}\frac{\partial f(P(r,\theta,\varphi))}{\partial\theta}, \frac{1}{r\sin\theta}\frac{\partial f(P(r,\theta,\varphi))}{\partial\varphi}.$$

Magyarázzuk meg ezt az eltérést!