1. előadás

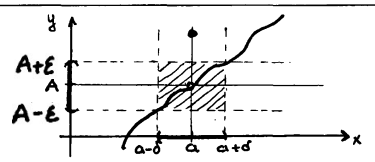
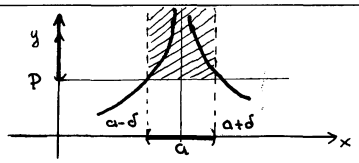
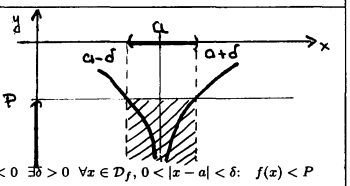
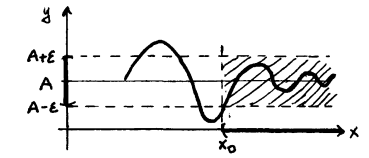
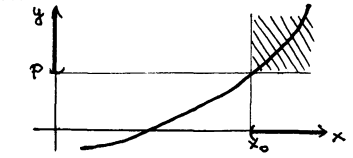
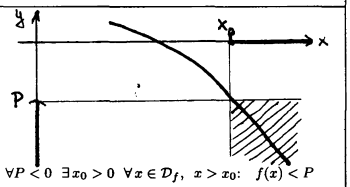
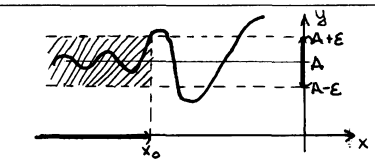
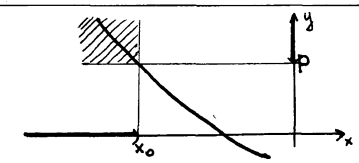
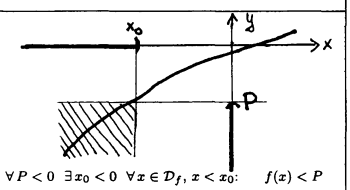
DIFFERENCIÁLSZÁMÍTÁS 1.

A differenciálszámítás a matematikai analízis egyik legfontosabb eszköze, amellyel meg tudjuk vizsgálni milyen gyorsan változik a függvény értéke az egyes helyeken, és ez egy jóval részletesebb függvénymegismerésre ad lehetőséget. Segítségével fontos gyakorlati alkalmazhatósággal bíró tulajdonságokat is meghatározhatunk, mint például a függvény monotonitását, szélsőértékhelyeit, a függvényhez húzott érintőegyenes meredekségét, stb. Ezek olyan fontos alkalmazási területek, melyeknek már az ókorban is komoly hagyományai voltak, igaz akkor még végtelen kicsiny mennyiségekkel, ú.n. infinitezimálisokkal próbáltak számolni. Sőt, még a XVIII. században is olyan módszereket alkalmaztak, melyek mai szemmel nézve nem voltak elég precízek. Csak a határérték fogalmának tisztázása tette lehetővé a pontos számítások elvégzését.

Emlékeztető a függvény határértékéről

Az Analízis I. kurzusban megismertük a matematikai analízis egyik legfontosabb fogalmát, nevezetesen a valós–valós függvények pontbeli határértékének definícióját. Emlékeztetünk arra, hogy ezzel a fogalommal egy függvénynek azt a – szemléletünk alapján eléggé világos – tulajdonságát fogalmaztuk meg matematikai szempontból pontos formában, hogy egy adott a ponthoz „közeli” helyeken a függvényértékek „közel” vannak valamely A értékhez.

Azonban, egy ennél általánosabb fogalmat adtunk meg, ahol a és A nem csak valós értékek, hanem $\pm\infty$ is lehetnek. Környezetek segítségével sikerült egy egységes fogalmat alkotni, ami egyenlőtlenségekkel kilenc speciális esetre át lehetett fogalmazni. Ezek a következő ábrán szemléltethetők.

	$A \in \mathbb{R}$	$A = +\infty$	$A = -\infty$
$a \in \mathbb{R}$	 $\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathcal{D}_f, 0 < \lvert x-a \rvert < \delta\colon\ \lvert f(x) - A \rvert < \varepsilon$	 $\forall P > 0\ \exists \delta > 0\ \forall x \in \mathcal{D}_f, 0 < \lvert x-a \rvert < \delta\colon\ f(x) > P$	 $\forall P < 0\ \exists \delta > 0\ \forall x \in \mathcal{D}_f, 0 < \lvert x-a \rvert < \delta\colon\ f(x) < P$
$a = +\infty$	 $\forall \varepsilon > 0\ \exists x_0 > 0\ \forall x \in \mathcal{D}_f, x > x_0\colon\ \lvert f(x) - A \rvert < \varepsilon$	 $\forall P > 0\ \exists x_0 > 0\ \forall x \in \mathcal{D}_f,\ x > x_0\colon\ f(x) > P$	 $\forall P < 0\ \exists x_0 > 0\ \forall x \in \mathcal{D}_f,\ x > x_0\colon\ f(x) < P$
$a = -\infty$	 $\forall \varepsilon > 0\ \exists x_0 < 0\ \forall x \in \mathcal{D}_f,\ x < x_0\colon\ \lvert f(x) - A \rvert < \varepsilon$	 $\forall P > 0\ \exists x_0 < 0\ \forall x \in \mathcal{D}_f, x < x_0\colon\ f(x) > P$	 $\forall P < 0\ \exists x_0 < 0\ \forall x \in \mathcal{D}_f, x < x_0\colon\ f(x) < P$

A fenti esetek közül nézzük meg közelebbről azt az esetet, amikor a és A valós számok. Legyen $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \mathcal{D}'_f$. Azt mondjuk, hogy ***az f függvénynek az a pontban vett határértéke az A szám***, ha

$$\forall \varepsilon > 0\text{-hoz } \exists \delta > 0, \forall x \in \mathcal{D}_f,\ 0 < |x - a| < \delta \colon |f(x) - A| < \varepsilon.$$

Ezt az alábbi szimbólumok valamelyikével jelöljük:

$$\lim_a f = A, \qquad \lim_{x \to a} f(x) = A, \qquad f(x) \to A, \text{ ha } x \to a.$$

A függvény határértékének definíciójának van egy vele ekvivalens állítás, amelyet ***átviteli elvnek*** nevezzük:

$$\lim_a f = A \quad \Longleftrightarrow \quad \forall (x_n) \colon \mathbb{N} \to \mathcal{D}_f \setminus \{a\},\ \lim_{n \to +\infty} x_n = a \text{ esetén } \lim_{n \to +\infty} f(x_n) = A.$$

Gyakorlaton konkrét határértékek kiszámításában nagy segítséget jelentet ***a határérték és a műveletek kapcsolata***, amely lényegében azt mondja ki, hogy az algebrai műveletek és a határérték képzési sorrendje felcserélhető, ha a műveletek elvégezhetők az $\overline{\mathbb{R}}$ kibővített valós számok halmazán. Így többek között nem tudjuk alkalmazni az előbbi kapcsolatot, ha olyan hányados határértékéről van szó, amelynek számlálójában és nevezőjében szereplő függvények határértéke az adott pontban egyszerre nulla. Ekkor $0/0$ típusú kritikus határértékről beszélünk.

Több nevezetes határértéket is tanultunk az Analízis I. kurzusban. Ezek közül szeretnék kiemelne a következő

$$\lim_{x \to 0} \frac{\sin x}{x} = 1, \qquad \lim_{x \to 0} \frac{e^x - 1}{x} = 1$$

határértékeket. Nem nehéz igazolni ezeket, ha kiindulunk a szinusz és exponenciális függvény hatványsoros definíciójából, és alkalmazzuk ***a hatványsor összegfüggvényének a határértékéről*** szóló tételt. Ez azt mondja ki, hogy egy hatványsor összegfüggvénynek van határértéke a konvergenciahalmazának minden pontjában, és ez megegyezik az összegfüggvény pontbeli értékével, ha a hatványsor konvergenciasugara nem nulla.

Az előbbi tulajdonság a ***függvény pontbeli folytonossághoz*** vezet. Akkor mondjuk, hogy az $f \in \mathbb{R} \to \mathbb{R}$ függvény folytonos az $a \in \mathcal{D}_f$ pontban, ha

$$\forall \varepsilon > 0\text{-hoz } \exists \delta > 0, \forall x \in \mathcal{D}_f,\ |x - a| < \delta \colon |f(x) - f(a)| < \varepsilon.$$

Jelölés: $f \in C\{a\}$. Két lehetőség adódik:

- ha $a \in \mathcal{D}_f$, de $a \notin \mathcal{D}'_f$, azaz izolált pont, akkor $f \in C\{a\}$.
- ha $a \in \mathcal{D}_f \cap \mathcal{D}'_f$, akkor $f \in C\{a\} \quad \Longleftrightarrow \quad \exists \lim_a f$ és $\lim_a f = f(a)$.

Ezért, ha egy függvény értelmezési tartománya intervallum, akkor a függvény akkor és csak akkor folytonos valamely pontjában, ha ott létezik határértéke, és ez megegyezik a függvény pontbeli értékével. Az előbbiek szerint egy hatványsor összegfüggvénye folytonos a konvergenciahalmazának minden pontjában. Így az exponenciális, a szinusz és a koszinusz függvények folytonosak a valós számok halmazán.

Egyoldali határértékeket is értelmeztünk. Lényegében akkor mondjuk, hogy egy f függvénynek van jobb oldali határértéke az a pontban, ha az f a $\mathcal{D}_f \cap (a, +\infty)$ halmazra való leszűkítésének van határértéke az a pontban. Bal oldali határérték esetén a $\mathcal{D}_f \cap (-\infty, a)$ halmazra való leszűkítést nézzük. A határértékre vonatkozó alaptételek az egyoldali határértékekre is érvényesek. Továbbá, ha az a pont bal és jobb oldali torlódási pontja $\mathcal{D}_f$-nek, akkor

$$\exists \lim_a f \quad \Longleftrightarrow \quad \exists \lim_{a-0} f,\ \exists \lim_{a+0} f \text{ és } \lim_{a-0} f = \lim_{a+0} f\ (= \lim_a f).$$

Két szemléletes probléma

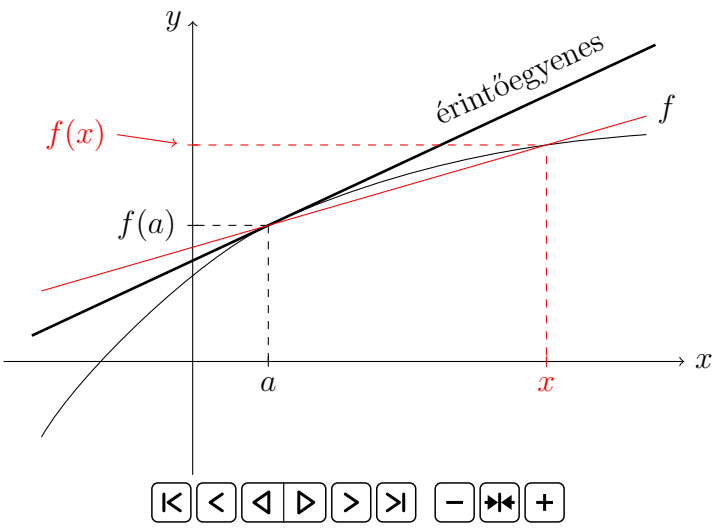
A derivált definíciójának értelmezése előtt két olyan problémát vázolunk, amelyek jól megvilágítják a derivált értelmezésének szükségességét, sőt a definíció célszerű módját is sugallják.

Az ún. ***érintő problémával*** fogjuk elkezdeni. Legyen I egy nyílt intervallum, illetve $f\colon I \to \mathbb{R}$ és $a \in I$. A kérdés az, hogy hogyan határozhatjuk meg az f függvény görbéjéhez az a abszcisszájú pontban húzott érintőegyenes (vagy egyszerűen érintő) egyenletét. Az előző kérdés egy másikat vet fel, mégpedig az érintő fogalmát az analízisben. A geometriából olyan intuitív fogalom alakul ki bennünk az érintőről, amely igazán csak olyan speciális görbéknél alkalmazható, mint a kör és további másodrendű görbék. Olyan fogalomra van szükségünk, amivel el tudjuk dönteni, hogy létezik-e az érintő a megadott a abszcisszájú pontban, és ha igen, akkor meg tudjuk vele határozni az érintő egyenletét.

Induljunk ki abból, hogy az (x_0, y_0) ponton átmenő m meredekségű egyenes egyenlete

$$y = m(x - x_0) + y_0.$$

Az érintő az f függvény görbéjét az $(a, f(a))$ pontban metszi, azaz $x_0 = a$ és $y_0 = f(a)$. Ha még az m meredekséget is ismernénk, akkor fel tudnánk írni az érintő egyenletét. Ehhez tekintsük meg a következő animációt.



Ha x az I intervallum egy a-tól különböző pontja, akkor tekintsük azt a szelőegyenest, amely az $\big(a, f(a)\big)$ és az $\big(x, f(x)\big)$ pontokon megy át (piros színnel ábrázolva). Mivel ismerjük a szelőegyenes két pontját, ezért ki tudjuk számolni a meredekségét. Ez természetesen függ az x érték megválasztásától, így értékét a következő függvény adja:

$$F(x) := \frac{f(x) - f(a)}{x - a} \qquad (x \in I,\ x \neq a).$$

Az animációban látható, hogy ha az x érték elég közel van az a értékéhez, akkor a szelőegyenes egy „hátáregyenes” közelébe kerül. Ezt fogjuk majd érintőnek nevezni. Meredeksége a szelőegyenesek meredekségének határértéke, ha $x \to a$. Ez precízen kifejezhető az F függvény a pontban vett határértékkel:

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

Célszerűnek tűnik tehát ezzel a határértékkel értelmezni az érintő meredekségét, ha a határérték létezik.

Tekintsük most a következő fizikai problémát. Egy pontszerű test akkor végez egyenletes mozgást egy egyenes pályán, hogy ha bármely megtett szakasz hosszát elosztjuk a szakasz megtételére szükséges idő nagyságával, akkor ugyanazt az értéket kapjuk. Ezt a hányadost a test ***sebességének*** nevezzük. Ha egy test nem egyenletesen mozog, akkor is beszélhetünk a test ***átlagsebességéről*** egy adott szakaszon, ami nem más, mint a szakasz hossza és az ennek megtételéhez szükséges idő hányadosa. Hogyan tudnánk egy test sebességét értelmezni egy adott időpontban?

Ha egy egészen kicsi időintervallumot nézünk, akkor azt várjuk, hogy ez idő alatt a sebesség nem változik jelentősen, így a mozgás közelítőleg egyenletes lesz. Ezért egy test sebességét egy adott időpontban úgy kellene értelmezni, mint egy kellően kis időtartamra vett átlagsebességet, de ez így nem elég precíz. A pontos értelmezéshez ismernünk kell a test mozgását az egyenes pályán a vizsgált I időintervallumban, azaz tudnunk kell milyen távolságban van a test egy, a pályán található referenciaponttól a t időpillanatban, ahol $t \in I$. Jelölje $s(t)$ ezt a távolságot, de mivel két irány van, meg kell őket különböztetni. Az egyik irányban legyen az s értéke pozitív, a másikban pedig negatív. Ezt a függvényt a test ***elmozdulásfüggvényének*** hívjuk.

Vegyünk egy t_0 időpontot az I intervallum belsejéből! Ha $t \in I$ egy másik időpont, akkor $\Delta t := t - t_0$ a két időpont között eltelt idő előjelesen, azaz pozitív, ha $t > t_0$, és negatív, ha $t < t_0$. Ez idő alatt a test egy $\Delta s := s(t) - s(t_0)$ hosszúságú szakaszt tesz meg, ami szintén előjeles mennyiség. Ezért a

$$\frac{\Delta s}{\Delta t} = \frac{s(t) - s(t_0)}{t - t_0}$$

különbséghányados a két időpont között eltelt időintervallumban megtett szakaszra vonatkozó átlagsebesség. Ez pozitív, ha a test pozitív irányba halad, és negatív, ha az ellenkező irányba. Az elvárás szerint a test sebessége a t_0 időpontban egy kellően kis Δt időtartamra vett átlagsebesség. Ezért célszerű lenne egy egyenes pályán mozgó test ***pillanatnyi sebességét*** egy t_0 időpontban a

$$\lim_{t \to t_0} \frac{s(t) - s(t_0)}{t - t_0}$$

határértékkel értelmezni, ami „figyel” a test haladási irányára, azaz pozitív, ha a test pozitív irányba halad, és negatív, ha az ellenkező irányba.

Vegyük észre, hogy a fenti határérték azonos az érintő problémában kapott határértékkel, ezért a továbbiakban kiemelten fogunk ilyen határértékekkel foglalkozni.

A pontbeli derivált fogalma

Egy adott függvény pontbeli deriváltját a függvény értelmezési tartományának olyan pontjaiban értelmezzük, amelyek valamely környezete is az értelmezési tartományhoz tartozik. Az ilyen pontokat belső pontoknak fogjuk nevezni.

1. Definíció. *Tegyük fel, hogy $\emptyset \neq A \subseteq \mathbb{R}$. Az $a \in A$ pont az A halmaz* **belső pontja**, *ha*

$$\exists K(a), \text{ hogy } K(a) \subseteq A.$$

Az **int A** *szimbólummal jelöljük* **az A halmaz belső pontjainak a halmazát**.

Példák: a) Ha $A = [0, 1]$, akkor $\operatorname{int} A = (0, 1)$.

b) Ha $A = (5, 6] \cup \{7\}$, akkor $\operatorname{int} A = (5, 6)$.

c) Ha $A = \{2, 3, 4\}$, akkor $\operatorname{int} A = \emptyset$.

2. Definíció. *Az $f \in \mathbb{R} \to \mathbb{R}$ függvény az $a \in \operatorname{int} \mathcal{D}_f$ pontban* **differenciálható** *(vagy* **deriválható***), ha*

$$\exists \quad \text{és véges a} \quad \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \quad \text{határérték.}$$

Ezt a határértéket az $f'(a)$ szimbólummal jelöljük, és **az f függvény a pontbeli deriváltjának** *(vagy* **differenciálhányadosának***) nevezzük, azaz*

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

Ezt a tényt a következőképpen fogjuk jelölni: $\boldsymbol{f \in D\{a\}}$.

Megjegyzések.

1. A fenti definícióban szereplő határértéket az $h = x - a$ helyettesítéssel gyakran így írjuk fel:
$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}.$$
2. Ha $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$, akkor az
$$F(x) := \frac{f(x) - f(a)}{x - a} \qquad \Big(x \in \mathcal{D}_f \setminus \{a\}\Big)$$
függvényt ***az f függvény az a ponthoz tartozó különbséghányados-függvényének*** vagy ***differenciahányados-függvényének*** nevezzük.
3. A derivált definíciójában 0/0-típusú kritikus határértékről van szó.

A differenciálhatóság erősebb megkötés, mint a folytonosság.

1. Tétel (A folytonosság és a deriválhatóság kapcsolata). *Tegyük fel, hogy $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$. Ekkor*

a) $f \in D\{a\} \quad \Longrightarrow \quad f \in C\{a\}$,

b) Az állítás megfordítása nem igaz.

Bizonyítás.

a) Ha $f \in D\{a\}$, akkor

$$\lim_{x \to a} \Big(f(x) - f(a)\Big) = \lim_{x \to a} \left(\frac{f(x) - f(a)}{x - a} \cdot (x - a) \right) =$$
$$= \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \cdot \lim_{x \to a} (x - a) = f'(a) \cdot 0 = 0.$$

Ez éppen azt jelenti, hogy $f \in C\{a\}$.

b) Van olyan folytonos függvény, amelyik egy pontban folytonos, de ott nem deriválható. Például az
$$f \colon \mathbb{R} \to \mathbb{R}, \qquad f(x) := |x|$$
abszolút érték függvény az $a = 0$ pontban.

Valóban $f \in C\{0\}$, de $f \notin D\{0\}$, hiszen minden $x \neq 0$ esetén

$$\frac{f(x)-f(0)}{x-0} = \frac{|x|-|0|}{x} = \frac{|x|}{x} = \begin{cases} 1, & \text{ha } x > 0, \\ -1, & \text{ha } x < 0, \end{cases}$$

és így

$$\lim_{x\to 0-0} \frac{f(x)-f(0)}{x-0} = -1 \qquad \text{és} \qquad \lim_{x\to 0+0} \frac{f(x)-f(0)}{x-0} = 1.$$

Ezért az f függvény 0 pontbeli differenciálhányadosa nem létezik.

Az érintő fogalma

Az előzőek alapján, ha $f \in D\{a\}$, akkor az $\big(a, f(a)\big)$ és az $\big(x, f(x)\big)$ pontokon átmenő szelőegyeneseknek van „határegyenese”, ha $x \to a$. Függvény grafikonjának (mint síkbeli halmaznak) az érintőjén éppen ezt az egyenest célszerű érteni. Az $f \in D\{a\}$ függvény esetén a szóban forgó egyenes átmegy az $\big(a, f(a)\big)$ pontban és a meredeksége $f'(a)$.

3. Definíció. *Az $f \in \mathbb{R} \to \mathbb{R}$ függvény grafikonjának az $\big(a, f(a)\big)$ pontban* **van érintője**, *ha $f \in D\{a\}$. Az f függvény grafikonjának $\big(a, f(a)\big)$ pontbeli* **érintőjén** *az*

$$y = f'(a) \cdot (x-a) + f(a)$$

egyenletű egyenest értjük.

Vegyük észre, hogy a $f'(a)$ definíciójában szereplő határérték véges, ezért az érintő nem lehet párhuzamos az y-tengellyel.

Megjegyzés. Érdemes meggondolni, hogy a kör és a parabola érintőjének a fenti definíciója ekvivalens a középiskolában geometriai úton megadott definícióval.

1. Feladat. *A differenciálhányados fogalmának segítségével határozzuk meg az $f(x) = x^2$ függvény görbéjéhez az $x_0 = 1$ abszcisszájú pontban húzott érintőegyenes egyenletét!*

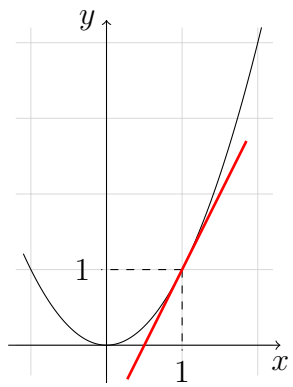
Megoldás. A differenciálhányados fogalma szerint

$$f'(1) = \lim_{h\to 0} \frac{(1+h)^2 - 1^2}{h} = \lim_{h\to 0} \frac{1+2h+h^2-1}{h} =$$

$$= \lim_{h\to 0} \frac{2h+h^2}{h} = \lim_{h\to 0}(2+h) = 2.$$

Másrészt $f(1) = 1^2 = 1$, így behelyettesíthetjük ezeket az érintő egyenes egyenletébe:

$$y = f'(1)(x-1) + f(1) \qquad \Longrightarrow \qquad y = 2x - 1.$$



$f(x) = x^2$

Deriválási szabályok

A definíció alapján annak eldöntése, hogy egy adott függvény deriválható-e valamilyen pontban, esetenként szinte reménytelen feladat. Éppen ezért különös jelentőséggel bírnak azok az állítások, amelyek ezt megkönnyítik. Ezek az ún. ***deriválási szabályok***, amelyek segítségével bizonyos függvények differenciálhatóságából és deriváltjának ismeretéből következtetni tudunk további függvények differenciálhatóságára és deriváltjára.

2. Tétel (Algebrai műveletekre vonatkozó deriválási szabályok). *Tegyük fel, hogy $f, g \in \mathbb{R} \to \mathbb{R}$ és $f, g \in D\{a\}$ valamilyen $a \in \operatorname{int}(\mathcal{D}_f \cap \mathcal{D}_g)$ pontban. Ekkor*

1. *a szorzó konstansokat ki tudjuk emelni a deriválásból, azaz*
$$cf \in D\{a\} \quad \text{és} \quad (cf)'(a) = cf'(a) \qquad (c \in \mathbb{R})$$
2. *tagokból álló függvényeket tagonként deriválhatjuk, azaz*
$$f + g \in D\{a\} \quad \text{és} \quad (f+g)'(a) = f'(a) + g'(a),$$
3. *egy szorzat deriváltja az az összeg, amelynek tagjai az egyik tényező deriváltja megszorozva a másik tényezővel, azaz*
$$f \cdot g \in D\{a\} \quad \text{és} \quad (f \cdot g)'(a) = f'(a)g(a) + f(a)g'(a),$$
4. *ha még a $g(a) \neq 0$ feltétel is teljesül, akkor*
$$\frac{f}{g} \in D\{a\} \quad \text{és} \quad \left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g^2(a)}.$$

Bizonyítás. A bizonyításban többször alkalmazzuk az algebrai műveletek és a határérték felcserélhetőségéről szóló tételt, illetve azt a tényt, hogy ha $g \in D\{a\}$, akkor $g \in C\{a\}$, tehát
$$\lim_{x \to a} g(x) = g(a).$$

1. Nyilván $a \in \operatorname{int} \mathcal{D}_{cf}$, másrészt
$$(cf)'(a) = \lim_{x \to a} \frac{cf(x) - cf(a)}{x - a} = \lim_{x \to a} c \cdot \frac{f(x) - f(a)}{x - a} = c \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = cf'(a).$$
2. Nyilván $a \in \operatorname{int} \mathcal{D}_{f+g}$, másrészt
$$\begin{aligned}(f+g)'(a) &= \lim_{x \to a} \frac{(f+g)(x) - (f+g)(a)}{x - a} = \lim_{x \to a} \frac{f(x) + g(x) - (f(a) + g(a))}{x - a} = \\ &= \lim_{x \to a} \frac{f(x) - f(a) + g(x) - g(a)}{x - a} = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} + \lim_{x \to a} \frac{g(x) - g(a)}{x - a} = \\ &= f'(a) + g'(a).\end{aligned}$$

3. Nyilván $a \in \operatorname{int} \mathcal{D}_{fg}$, másrészt

$$\begin{aligned}
(fg)'(a) &= \lim_{x\to a} \frac{(fg)(x)-(fg)(a)}{x-a} = \lim_{x\to a} \frac{f(x)g(x)-f(a)g(a)}{x-a} = \\
&= \lim_{x\to a} \frac{f(x)g(x)-f(a)g(x)+f(a)g(x)-f(a)g(a)}{x-a} = \\
&= \lim_{x\to a} \frac{\big(f(x)-f(a)\big)g(x)+f(a)\big(g(x)-g(a)\big)}{x-a} = \\
&= \lim_{x\to a} \frac{f(x)-f(a)}{x-a} \lim_{x\to a} g(x) + f(a) \lim_{x\to a} \frac{g(x)-g(a)}{x-a} = f'(a)g(a)+f(a)g'(a).
\end{aligned}$$

4. Ha $g \in D\{a\}$, akkor $g \in C\{a\}$, ezért a $g(a) \neq 0$ feltétel miatt

$$\exists K(a) \subseteq \mathcal{D}_f,\ \forall x \in K(a)\colon g(x) \neq 0 \qquad \Longrightarrow \qquad a \in \operatorname{int} \mathcal{D}_{f/g},$$

hiszen a folytonos függvények előjeltartóak. Másrészt

$$\begin{aligned}
\left(\frac{f}{g}\right)'(a) &= \lim_{x\to a} \frac{\left(\frac{f}{g}\right)(x) - \left(\frac{f}{g}\right)(a)}{x-a} = \lim_{x\to a} \frac{\frac{f(x)}{g(x)} - \frac{f(a)}{g(a)}}{x-a} = \lim_{x\to a} \frac{f(x)g(a)-f(a)g(x)}{g(a)g(x)(x-a)} = \\
&= \lim_{x\to a} \frac{f(x)g(a)-f(a)g(a)+f(a)g(a)-f(a)g(x)}{g(a)g(x)(x-a)} = \\
&= \lim_{x\to a} \left(\frac{1}{g(a)g(x)} \cdot \frac{\big(f(x)-f(a)\big)g(a)-f(a)\big(g(x)-g(a)\big)}{x-a} \right) = \\
&= \frac{1}{g(a)\lim\limits_{x\to a} g(x)} \left(\lim_{x\to a} \frac{f(x)-f(a)}{x-a} g(a) - f(a) \lim_{x\to a} \frac{g(x)-g(a)}{x-a} \right) = \\
&= \frac{1}{g^2(a)} \left(f'(a)g(a)-f(a)g'(a)\right).
\end{aligned}$$

Az összetett függvényre vonatkozó szabály azt mondja ki, hogy egy összetett függvény deriváltjához először a külső függvényt deriváljuk, miközben belseje érintetlen marad. Ezt utána még megszorozzuk a belső függvény deriváltjával.

3. Tétel (Összetett függvény deriváltja). *Tegyük fel, hogy $f, g \in \mathbb{R} \to \mathbb{R}$ és valamilyen $a \in \operatorname{int} \mathcal{D}_g$ pontban $g \in D\{a\}$, továbbá $f \in D\{g(a)\}$. Ekkor $f \circ g \in D\{a\}$, és*

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

Bizonyítás. A tétel feltételei alapján

$$f \in D\{g(a)\} \qquad \Longrightarrow \qquad g(a) \in \operatorname{int} \mathcal{D}_f \qquad \Longrightarrow \qquad \exists \varepsilon > 0\colon K_\varepsilon\Big(g(a)\Big) \subseteq \mathcal{D}_f,$$

illetve az $a \in \operatorname{int} \mathcal{D}_g$ feltétel miatt

$$g \in D\{a\} \qquad \Longrightarrow \qquad g \in C\{a\} \qquad \Longrightarrow \qquad \varepsilon\text{-hoz } \exists \delta > 0, \forall x \in K_\delta(a)\colon g(x) \in K_\varepsilon\Big(g(a)\Big) \subseteq \mathcal{D}_f.$$

Ebből következik, hogy $K_\delta(a) \subseteq \mathcal{D}_{f\circ g}$, azaz $a \in \operatorname{int} \mathcal{D}_{f\circ g}$.

Másrészt, jelölje $b := g(a)$ és $\varepsilon_g : \mathcal{D}_g \to \mathbb{R}$, $\varepsilon_f : \mathcal{D}_f \to \mathbb{R}$ a következő függvényeket:

$$\varepsilon_g(x) := \begin{cases} \dfrac{g(x)-g(a)}{x-a} - g'(a), & \text{ha } x \neq a, \\ 0, & \text{ha } x = a, \end{cases} \qquad \varepsilon_f(y) := \begin{cases} \dfrac{f(y)-f(b)}{y-b} - f'(b), & \text{ha } y \neq b, \\ 0, & \text{ha } y = b. \end{cases}$$

Mivel $g \in D\{a\}$ és $f \in D\{b\}$, így $\varepsilon_g \in C\{a\}$ és $\varepsilon_f \in C\{b\}$. Továbbá a fentiekből

$$g(x) - g(a) = g'(a)(x-a) + \varepsilon_g(x)(x-a) \qquad (x \in \mathcal{D}_g),$$
$$f(y) - f(b) = f'(b)(y-b) + \varepsilon_f(y)(y-b) \qquad (y \in \mathcal{D}_f).$$

Ezért, ha $x \in \mathcal{D}_{f\circ g}$ és $y = g(x)$, akkor $y - b = g(x) - g(a)$ és

$$\begin{aligned} f\Big(g(x)\Big) - f\Big(g(a)\Big) &= f(y) - f(b) = f'(b)(y-b) + \varepsilon_f(y)(y-b) = \\ &= f'\Big(g(a)\Big)\Big(g(x)-g(a)\Big) + \varepsilon_f\Big(g(x)\Big)\Big(g(x)-g(a)\Big) = \\ &= f'\Big(g(a)\Big)\Big(g'(a)(x-a) + \varepsilon_g(x)(x-a)\Big) + \varepsilon_f\Big(g(x)\Big)\Big(g'(a)(x-a) + \varepsilon_g(x)(x-a)\Big) = \\ &= f'\Big(g(a)\Big)g'(a)(x-a) + \Big(f'(b)\varepsilon_g(x) + \varepsilon_f\Big(g(x)\Big)g'(a) + \varepsilon_f\Big(g(x)\Big)\varepsilon_g(x)\Big)(x-a). \end{aligned}$$

Tudjuk, hogy $\varepsilon_g \in C\{a\} \implies \lim\limits_{x\to a} \varepsilon_g(x) = 0$. Mivel $\varepsilon_f \in C\{b\}$, így alkalmazható az összetett függvény határértékére vonatkozó tételt: $\lim\limits_{x\to a} \varepsilon_f\Big(g(x)\Big) = \lim\limits_{y\to b} \varepsilon_f(y) = 0$. Így

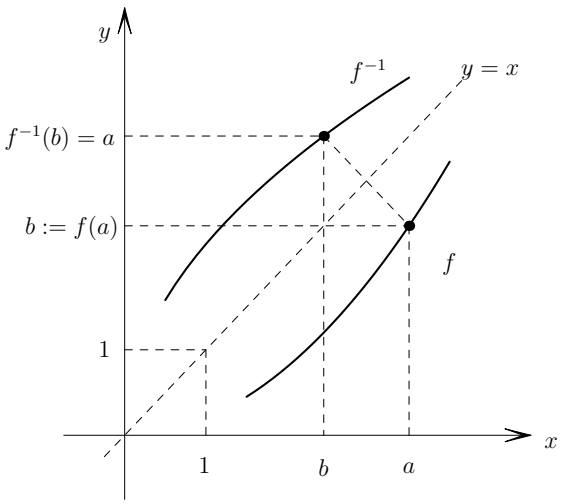
$$\begin{aligned} (f\circ g)'(a) &= \lim_{x\to a} \frac{f\Big(g(x)\Big) - f\Big(g(a)\Big)}{x-a} = \\ &= f'\Big(g(a)\Big)g'(a) + \lim_{x\to a}\Big(f'(b)\varepsilon_g(x) + \varepsilon_f\Big(g(x)\Big)g'(a) + \varepsilon_f\Big(g(x)\Big)\varepsilon_g(x)\Big) = \\ &= f'\Big(g(a)\Big)g'(a) + f'(b)\cdot 0 + 0\cdot g'(a) + 0\cdot 0 = f'\Big(g(a)\Big)g'(a). \end{aligned}$$

Tegyük fel, hogy az f függvény invertálható. Emlékeztetünk arra, hogy az f és az f^{-1} függvények grafikonjai egymásnak az $y = x$ egyenletű egyenesre vonatkozó tükörképei.

Tekintsük az f függvény grafikonjának egy $\Big(a, f(a)\Big) = (a, b)$ pontját. Ennek tükörképe az $y = x$ egyenletű szögfelező egyenesre a (b, a) pont. Mivel $a = f^{-1}(b)$, ezért a (b, a) pont rajta van az f^{-1} függvény grafikonján.

Az f függvény grafikonjának $\Big(a, f(a)\Big) = (a, b)$ pontbeli érintőegyenesének tükörképe az f^{-1} függvény grafikonjának az $\Big(f(a), a\Big) = (b, a)$ pontbeli érintője. Ha az f-hez húzott érintő nem párhuzamos az x-tengellyel (vagyis $f^{-1}(a) \neq 0$), akkor a tükörképe nem párhuzamos az y-tengellyel. Ekkor a meredekségeik egymás reciprokai, vagyis

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



4. Tétel (Inverz függvény deriváltja). *Legyen I egy nyílt intervallum, és $f : I \to \mathbb{R}$. Tegyük fel, hogy*

a) f szigorúan monoton és folytonos az I intervallumon,

b) valamilyen $a \in I$ pontban $f \in D\{a\}$ és $f'(a) \neq 0$.

Ekkor az f^{-1} függvény deriválható a $b = f(a)$ pontban és

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

Bizonyítás. Tudjuk, hogy ha f szigorúan monoton, akkor invertálható és inverze is szigorúan monoton. Mivel $\mathcal{D}_f = I$ egy nyílt intervallum, illetve f szigorúan monoton és folytonos függvény, így $\mathcal{D}_{f^{-1}} = \mathcal{R}_f$ is egy nyílt intervallum, ezért $b \in \operatorname{int} \mathcal{D}_{f^{-1}}$.

Legyen $y \in \mathcal{D}_{f^{-1}}$. Tudjuk, hogy $b = f(a)$ és $a = f^{-1}(b)$. Mivel f^{-1} szigorúan monoton, így egy-egyértelmű, ezért alkalmazhatjuk az $x = f^{-1}(y)$ helyettesítést az alábbi határértékben, ahol $x \to f^{-1}(b) = a$, ha $y \to b$, és $y = f(x)$.

$$(f^{-1})'(b) = \lim_{y \to b} \frac{f^{-1}(y) - f^{-1}(b)}{y - b} = \lim_{x \to a} \frac{x - a}{f(x) - f(a)} = \frac{1}{\lim\limits_{x \to a} \frac{f(x) - f(a)}{x - a}} = \frac{1}{f'(a)},$$

hiszen $f \in D\{a\}$ és $f'(a) \neq 0$.

Hatványsor összegfüggvénye a konvergenciaintervallum belsejében differenciálható és a hatványsor deriválását szabad tagonként végezni.

5. Tétel (Hatványsor összegfüggvényének deriváltja). *Legyen $a \in \mathbb{R}$ és $\alpha_n \in \mathbb{R}$ $(n = 0, 1, \ldots)$. Tegyük fel, hogy a*

$$\sum_{n=0} \alpha_n (x - a)^n \qquad (x \in \mathbb{R})$$

hatványsor R konvergenciasugara pozitív, és jelölje f az összegfüggvényét:

$$f(x) := \sum_{n=0}^{+\infty} \alpha_n (x - a)^n \qquad \big(x \in K_R(a)\big).$$

Ekkor minden $x \in K_R(a)$ pontban $f \in D\{x\}$ és

$$f'(x) = \sum_{n=1}^{+\infty} n\alpha_n (x - a)^{n-1} \qquad \big(x \in K_R(a)\big).$$

Bizonyítás. Nem bizonyítjuk.

A deriváltfüggvény

Látni fogjuk, hogy a derivált a leghatékonyabb segédeszköz egy függvény tulajdonságainak vizsgálatára. Ez lokálisan és globálisan is igaz. Az $f'(a)$ derivált létezése és értéke az f függvény a-beli (lokális) viselkedésére jellemző: $f'(a)$ értékéből az f függvény a pont körüli viselkedésére vonhatunk le következtetéseket. (Egy ilyen kapcsolatot már találtunk, amikor beláttuk, hogy a differenciálhatóságból következik a folytonosság.)

Ha viszont f egy A halmaz (például intervallum) minden pontjában deriválható, akkor az $f'(x)$ értékekből az f függvény globális viselkedésére következtethetünk. Erre azt mondjuk, hogy ***f* az *A* halmazon deriválható** vagy ***differenciálható***, és az $\boldsymbol{f \in D(A)}$ jelölést alkalmazzuk (vagy egyszerűen $\boldsymbol{f \in D(a,b)}$, ha $A = (a,b)$ egy intervallum).

Az alkalmazásokban legtöbbször olyan f függvények szerepelnek, amelyek valamely A halmazon (például intervallumon) deriválhatók. Célszerű tehát a deriválást olyan operációként felfogni, amely függvényekhez rendel függvényeket.

4. Definíció. *Ha $f \in \mathbb{R} \to \mathbb{R}$ és $\Big\{x \in \operatorname{int} \mathcal{D}_f \mid f \in D\{x\}\Big\} \neq \emptyset$, akkor az*

$$\Big\{x \in \operatorname{int} \mathcal{D}_f \mid f \in D\{x\}\Big\} \ni x \mapsto f'(x)$$

függvényt az f **deriváltfüggvényének** *(vagy* **differenciálhányados-függvényének***) nevezzük, és az f' szimbólummal jelöljük.*

Néhány speciális függvény deriváltja

Ebben a táblázatban felsoroltuk azokat a speciális függvényeket, amelyek deriváltjait meg kell jegyezni. Ezek közül néhánynak a bizonyítását most megmutatjuk.

$\boxed{\textbf{1.}}$ ***Konstans függvények.*** Tetszőleges $c \in \mathbb{R}$ esetén az

$$f(x) := c \qquad (x \in \mathbb{R})$$

konstans függvény minden $x \in \mathbb{R}$ pontban deriválható és deriváltja

$$\boxed{f'(x) = 0 \quad (x \in \mathbb{R})} \qquad \text{vagyis} \qquad \boxed{\big(c\big)' = 0 \quad (x \in \mathbb{R})}.$$

Bizonyítás. Minden $x \in \mathbb{R}$ esetén $x \in \operatorname{int} \mathcal{D}_f = \mathbb{R}$, és

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{c - c}{h} = 0. \quad \blacksquare$$

$\boxed{\textbf{2.}}$ ***Hatványfüggvények.*** Tetszőleges n természetes szám esetén az

$$f(x) := x^n \qquad (x \in \mathbb{R})$$

hatványfüggvény minden $x \in \mathbb{R}$ pontban deriválható, és deriváltja

$$\boxed{\big(x^n\big)' = nx^{n-1} \quad (x \in \mathbb{R})}.$$

Bizonyítás. Minden $x \in \mathbb{R}$ esetén $x \in \operatorname{int} \mathcal{D}_f = \mathbb{R}$, és

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \lim_{h\to 0} \frac{(x+h)^n - x^n}{h} =$$

$$= \lim_{h\to 0} \frac{h\Big[(x+h)^{n-1} + (x+h)^{n-2}x + \cdots + (x+h)x^{n-2} + x^{n-1}\Big]}{h} = nx^{n-1}$$

a hatványfüggvény folytonossága alapján. Az előző határérték kiszámításában a következő azonosságot alkalmaztuk:

$$a^n - b^n = (a-b)\Big(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1}\Big). \quad \blacksquare$$

$\boxed{\textbf{3.}}$ ***A természetes alapú exponenciális függvény*** minden $x \in \mathbb{R}$ pontban deriválható, és

$$\boxed{\exp'(x) = (e^x)' = e^x \qquad \Big(x \in \mathbb{R}\Big)}.$$

Bizonyítás. Az exp függvényt az egész $\mathbb{R}$-en konvergens hatványsor összegfüggvényeként értelmeztük:

$$\exp(x) := e^x := \sum_{n=0}^{+\infty} \frac{1}{n!}\, x^n \qquad (x \in \mathbb{R}).$$

A hatványsor összegfüggvényének deriválására vonatkozó tételt alkalmazva azt kapjuk, hogy minden $x \in \mathbb{R}$ pontban $\exp \in D\{x\}$, és

$$\exp'(x) = \sum_{n=1}^{+\infty} n \cdot \frac{1}{n!}\, x^{n-1} = \sum_{n=1}^{+\infty} \frac{1}{(n-1)!}\, x^{n-1} = (k = n-1) = \sum_{k=0}^{+\infty} \frac{1}{k!}\, x^k = \exp(x) = e^x. \quad \blacksquare$$

Megjegyzés. Az állítás úgy is megfogalmazható, hogy az **exp** ***függvény deriváltfüggvénye önmaga***. Tulajdonképpen ez az összefüggés indokolja, hogy az e számot tekintjük az analízis (és általában a matematika) egyik legfontosabb állandójának.

$\boxed{\textbf{4.}}$ ***A természetes alapú logaritmusfüggvény*** minden $x \in (0, +\infty)$ pontban deriválható, és

$$\boxed{\ln' x = \Big(\ln x\Big)' = \frac{1}{x} \qquad \Big(x \in (0, +\infty)\Big)}.$$

Bizonyítás. Emlékeztetünk arra, hogy az ln függvényt az exp függvény inverzeként értelmeztük:

$$\ln := \exp^{-1}.$$

Az exp függvényre az inverz függvény deriválására vonatkozó tétel valamennyi feltétele teljesül. Ezért minden $x \in (0, +\infty)\ (= \mathcal{D}_{\ln})$ pontban $\ln \in D\{x\}$, és

$$\ln' x = \Big(\ln x\Big)' = \frac{1}{\exp'\Big(\ln x\Big)} = \frac{1}{\exp\Big(\ln x\Big)} = \frac{1}{x}. \quad \blacksquare$$

$\boxed{5.}$ ***Exponenciális függvények.*** Ha $a > 0$ valós szám, akkor minden $x \in \mathbb{R}$ pontban $\exp_a \in D\{x\}$, és

$$\boxed{\exp_a'(x) = (a^x)' = a^x \ln a \qquad \Big(x \in \mathbb{R}\Big)}.$$

Bizonyítás. Emlékeztetünk arra, hogy $a > 0$ valós szám esetén az a alapú exponenciális függvényt így értelmeztük:

$$a^x := \exp_a(x) := \exp(x \cdot \ln a) \quad (x \in \mathbb{R}).$$

Ha $f := \exp$ és $g(x) := x \ln a$ $(x \in \mathbb{R})$, akkor

$$\exp_a = f \circ g.$$

Az exp és az összetett függvény deriválására vonatkozó állítások alapján azt kapjuk, hogy az $\exp_a$ függvény minden $x \in \mathbb{R} = \mathcal{D}_{\exp_a}$ pontban deriválható, és

$$\exp_a'(x) = f'\Big(g(x)\Big) \cdot g'(x) = \exp(x \ln a) \cdot \ln a = a^x \ln a. \quad \blacksquare$$

$\boxed{6.}$ ***Logaritmusfüggvények.*** Ha $a > 0$ valós szám és $a \neq 1$, akkor $\forall\, x \in (0, +\infty)\, (= \mathcal{D}_{\log_a})$ pontban $\log_a \in D\{x\}$, és

$$\boxed{\log_a' x = \Big(\log_a x\Big)' = \frac{1}{x \ln a} \qquad \Big(x \in (0, +\infty)\Big)}.$$

Bizonyítás. Emlékeztetünk arra, hogy $a \in (0, +\infty) \setminus \{1\}$ esetén az a alapú logaritmusfüggvényt így értelmeztük:

$$\log_a := \Big(\exp_a\Big)^{-1}.$$

Az $\exp_a$ függvényre az inverz függvény deriválására vonatkozó tétel valamennyi feltétele teljesül. Ezért minden $x \in (0, +\infty)\, (= \mathcal{D}_{\log_a})$ pontban $\log_a \in D\{x\}$, és

$$\log_a' x = \Big(\log_a x\Big)' = \frac{1}{\exp_a'\Big(\log_a x\Big)} = \frac{1}{(\ln a) \cdot \exp_a\Big(\log_a x\Big)} = \frac{1}{x \ln a}. \quad \blacksquare$$

$\boxed{7.}$ ***Általánosított hatványfüggvények.*** Ha α tetszőleges valós szám, akkor a

$$h_\alpha : (0, +\infty) \ni x \mapsto x^\alpha$$

általánosított hatványfüggvény minden $x \in (0, +\infty)\, (= \mathcal{D}_{h_\alpha})$ pontban deriválható, és

$$\boxed{(x^\alpha)' = \alpha x^{\alpha - 1} \qquad \Big(x \in (0, +\infty)\Big)}.$$

Bizonyítás. Rögzítsük az $\alpha \in \mathbb{R}$ számot. Írjuk fel az $x > 0$ alapot e-hatványként: $x = e^{\ln x}$. Ekkor

$$x^\alpha = \Big(e^{\ln x}\Big)^\alpha = e^{\alpha \ln x} \qquad \Big(x > 0\Big),$$

majd alkalmazzuk az összetett függvény deriválására vonatkozó állításunkat. Azt kapjuk, hogy minden $x > 0$ esetén $h_\alpha \in D\{x\}$ és

$$h_\alpha'(x) = (x^\alpha)' = \Big(e^{\alpha \ln x}\Big)' = e^{\alpha \ln x} \cdot (\alpha \ln x)' = e^{\alpha \ln x} \cdot \alpha \cdot \frac{1}{x} = x^\alpha \cdot \alpha \cdot \frac{1}{x} = \alpha\, x^{\alpha - 1}. \quad \blacksquare$$

$\boxed{\textbf{8.}}$ ***A szinuszfüggvény*** minden $x \in \mathbb{R}$ pontban deriválható, és

$$\boxed{\sin' x = \cos x \qquad \Big(x \in \mathbb{R}\Big)}.$$

Bizonyítás. A szinuszfüggvényt az egész $\mathbb{R}$-en konvergens hatványsor összegfüggvényeként értelmeztük:

$$\sin(x) := \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \qquad (x \in \mathbb{R}).$$

A hatványsor összegfüggvényének deriválására vonatkozó tételt alkalmazva azt kapjuk, hogy minden $x \in \mathbb{R}$ pontban $\sin \in D\{x\}$, és

$$\sin'(x) = \sum_{n=0}^{+\infty} (-1)^n \cdot (2n+1) \frac{x^{2n}}{(2n+1)!} = \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \cos(x). \quad \blacksquare$$

$\boxed{\textbf{9.}}$ ***A koszinuszfüggvény*** minden $x \in \mathbb{R}$ pontban deriválható, és

$$\boxed{\cos' x = -\sin x \qquad \Big(x \in \mathbb{R}\Big)}.$$

Bizonyítás. A koszinuszfüggvényt az egész $\mathbb{R}$-en konvergens hatványsor összegfüggvényeként értelmeztük:

$$\cos(x) := \sum_{n=0}^{+\infty} (-1)^n \frac{x^{2n}}{(2n)!} \qquad (x \in \mathbb{R}).$$

A hatványsor összegfüggvényének deriválására vonatkozó tételt alkalmazva azt kapjuk, hogy minden $x \in \mathbb{R}$ pontban $\cos \in D\{x\}$, és

$$\cos'(x) = \sum_{n=0}^{+\infty} (-1)^n \cdot 2n \frac{x^{2n-1}}{(2n)!} = \sum_{n=1}^{+\infty} (-1)^n \cdot 2n \frac{x^{2n-1}}{(2n)!} = \sum_{n=1}^{+\infty} (-1)^n \frac{x^{2n-1}}{(2n-1)!} =$$

$$= -\sum_{n=1}^{+\infty} (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!} = (k = n-1) = -\sum_{k=0}^{+\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} = -\sin(x). \quad \blacksquare$$

2. előadás

DIFFERENCIÁLSZÁMÍTÁS 2.

Egyoldali pontbeli deriváltak

Vannak olyan esetek, amikor a differenciálhányados létezése a bal és a jobb oldali határértékek egyezésével szükséges vizsgálni. Például, az

$$f\colon \mathbb{R} \to \mathbb{R}, \qquad f(x) := |x| = \begin{cases} x, & \text{ha } x \geq 0, \\ -x, & \text{ha } x < 0 \end{cases}$$

abszolút érték függvény az $a = 0$ pontban nem differenciálható, mert

$$\lim_{x\to 0-0} \frac{f(x)-f(0)}{x-0} = \lim_{x\to 0-0} \frac{-x}{x} = -1 \qquad \text{és} \qquad \lim_{x\to 0+0} \frac{f(x)-f(0)}{x-0} = \lim_{x\to 0+0} \frac{x}{x} = 1.$$

Érdemes tehát az egyoldali pontbeli deriváltakat értelmezni.

1. Definíció. *Legyen $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \mathcal{D}_f$ olyan pont, hogy $\exists \delta > 0\colon [a, a+\delta) \subseteq \mathcal{D}_f$. Azt mondjuk, hogy* **az f függvény jobbról differenciálható** *(vagy* **jobbról deriválható***)* **az a pontban**, *ha*

$$\exists \ \text{ és véges a } \lim_{x\to a+0} \frac{f(x)-f(a)}{x-a} \ \text{ határérték.}$$

Ezt a határértéket az $f'_+(a)$ szimbólummal jelöljük, és **az f függvény a pontbeli jobb oldali deriváltjának** *(vagy* **differenciálhányadosának***) nevezzük, azaz*

$$f'_+(a) := \lim_{x\to a+0} \frac{f(x)-f(a)}{x-a} \in \mathbb{R}.$$

Az f függvény bal oldali differenciálhatóságát az a pontban *hasonlóan értelmezzük, és az $f'_-(a)$ szimbólummal jelöljük* **az a pontbeli bal oldali deriváltját**.

Az előző jelölés értelmében, ha $f(x) = |x| \ \ (x \in \mathbb{R})$, akkor $f'_-(0) = -1$ és $f'_+(0) = 1$.

Az egyoldali függvényhatárértékeknél tanultak szerint világos, hogy

$$f \in D\{a\} \quad \Longleftrightarrow \quad \exists f'_-(a),\ \exists f'_+(a) \ \text{ és } \ f'_-(a) = f'_+(a) \ (= f'(a)).$$

A bal és jobb oldali deriváltak segítséget nyújtanak a következő általános probléma megoldásában. Legyen $b, j \in \mathbb{R} \to \mathbb{R}$, $a \in \mathbb{R}$ olyan pont, hogy $\exists \delta > 0\colon (a-\delta, a) \subseteq \mathcal{D}_b$, $(a, a+\delta) \subseteq \mathcal{D}_j$, és $A \in \mathbb{R}$. Milyen legyenek a b és j függvények ahhoz, hogy az

$$f(x) := \begin{cases} b(x), & \text{ha } x < a \ (x \in \mathcal{D}_b), \\ A, & \text{ha } x = a, \\ j(x), & \text{ha } x > a \ (x \in \mathcal{D}_j) \end{cases}$$

függvény differenciálható legyen az a pontban?

Világos, hogy $a \in \operatorname{int} \mathcal{D}_f$. Ha $f \in C\{a\}$, akkor $\exists \lim\limits_{a-0} f$, $\exists \lim\limits_{a+0} f$ és $\lim\limits_{a-0} f = \lim\limits_{a+0} f = f(a) = A$. Ezért

$$\exists \lim_{a-0} b = \lim_{a-0} f = A \qquad \text{és} \qquad \exists \lim_{a+0} j = \lim_{a+0} f = A,$$

azaz

$$\boxed{\text{I.} \quad \exists \lim_{a-0} b,\ \exists \lim_{a+0} j \quad \text{és} \quad \lim_{a-0} b = A = \lim_{a+0} j}\ .$$

Az I. feltétel azt jelenti, hogy $f \in C\{a\}$, ami szükséges ahhoz, hogy $f \in D\{a\}$ teljesüljön.

Gyakori eset, amikor b balról, j jobbról folytonos a-ban és $b(a) = j(a) = A$. Világos, hogy ekkor az I. feltétel teljesül. Fordítva, ha az I. feltétel teljesül, akkor a

$$b(a) := \lim_{a-0} b \qquad \text{és} \qquad j(a) := \lim_{a+0} j.$$

értékadással (módosítással) elérhetjük, hogy b balról és j jobbról folytonos legyen a-ban, valamint $b(a) = j(a) = A$. Ez a módosítás nem befolyásolja az f függvény értékét, hiszen $f(a) = A$.

Ha az I. feltétel teljesül, és $b(a) = j(a) = A$, akkor $f \in D\{a\} \iff f'_-(a) = f'_+(a)$, ahol

$$f'_-(a) = \lim_{x \to a-0} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a-0} \frac{b(x) - A}{x - a} = \lim_{x \to a-0} \frac{b(x) - b(a)}{x - a} = b'_-(a)$$

$$f'_+(a) = \lim_{x \to a+0} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a+0} \frac{j(x) - A}{x - a} = \lim_{x \to a+0} \frac{j(x) - j(a)}{x - a} = j'_+(a),$$

azaz

$$\boxed{\text{II.} \quad b'_+(a) = j'_+(a)}\ .$$

Ekkor $f'(a) = b'_+(a) = j'_+(a)$. Ha $b \in D\{a\}$ és $j \in D\{a\}$, akkor a II. feltétel ekvivalens, azzal, hogy $b'(a) = j'(a)$.

1. Feladat. *Állapítsa meg, hogy differenciálható-e az alábbi függvény az $a = 0$ pontban!*

$$f(x) = \begin{cases} 1 - x, & \text{ha } x \le 0, \\ e^{-x} & \text{ha } x > 0. \end{cases}$$

Megoldás. Alkalmazzuk a fenti jelöléseket! Ekkor $A := f(0) = 1 - 0 = 1$, illetve legyen

$$b(x) := 1 - x \quad (x \in \mathbb{R}) \qquad \text{és} \qquad j(x) := e^{-x} \quad (x \in \mathbb{R}).$$

A deriválási szabályok alapján $b, j \in D(\mathbb{R})$ és

$$b'(x) = -1, \qquad j'(x) = -e^{-x} \qquad (x \in \mathbb{R}).$$

Ekkor

- I. teljesül, hiszen $b, j \in C\{0\}$, és $b(0) = j(0) = 1 = A$.
- II. teljesül, hiszen $b, j \in D\{0\}$ és $b'(0) = -1 = j'(0)$.

Ezért $f \in D\{0\}$ és $f'(0) = -1$.

Magasabb rendű deriváltak

Ha valamely valós-valós függvénynek létezik a deriváltfüggvénye, akkor természetes módon felvethetjük annak újbóli deriválhatóságát, és így eljuthatunk a ***többször deriválható függvények*** és a ***magasabb rendű deriváltak*** fogalmához. A rekurzió módszerét alkalmazzuk. Először a kétszer deriválhatóság fogalmát definiáljuk.

2. Definíció. *Legyen $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$. Azt mondjuk, hogy* ***f*** **kétszer deriválható az *a* pontban** *(jelölése: $f \in D^2\{a\}$), ha*

- $\exists r > 0\colon f \in D\Big(K_r(a)\Big)$, *és*
- *az f' deriváltfüggvény deriválható a-ban, azaz $f' \in D\{a\}$.*

Legyen ekkor

$$f''(a) := \Big(f'\Big)'(a)$$

az f függvény ***a*-beli második deriváltja**.

Ha $H := \Big\{x \in \operatorname{int} \mathcal{D}_f \mid f \in D^2\{x\}\Big\} \neq \emptyset$, akkor $H \ni x \mapsto f''(x)$ az f függvény ***második deriváltfüggvénye***, amit röviden az f'' szimbólummal jelölünk.

Jelölések. A deriváltakra és a deriváltfüggvényekre a következő jelöléseket is fogjuk használni:

$$f^{(1)}(a) := f'(a) \quad \text{és} \quad f^{(1)} := f',$$
$$f^{(2)}(a) := f''(a) \quad \text{és} \quad f^{(2)} := f''.$$

Megállapodunk abban is, hogy $\quad f^{(0)}(a) := f(a) \quad$ és $\quad f^{(0)} := f$.

Indukcióval értelmezzük az ***n*-szeri deriválhatóságot** és az ***n*-edik deriváltat**. Tegyük fel, hogy valamely $n \in \mathbb{N}$ esetében már értelmeztük azt, hogy valamely $f \in \mathbb{R} \to \mathbb{R}$ függvény mikor deriválható $(n-1)$-szer egy $a \in \operatorname{int} \mathcal{D}_f$ pontban (jelölése: $f \in D^{n-1}\{a\}$), továbbá azt is, hogy mikor létezik és mi az $(n-1)$-edik deriváltfüggvénye. Ha ez utóbbi létezik, akkor jelöljük azt az $f^{(n-1)}$ szimbólummal.

3. Definíció. *Legyen $f \in \mathbb{R} \to \mathbb{R}$, $a \in \operatorname{int} \mathcal{D}_f$, és tegyük fel, hogy valamely $n = 2, 3, \ldots$ esetén létezik az $f^{(n-1)}$-gyel jelölt $(n-1)$-edik deriváltfüggvény. Azt mondjuk, hogy* ***f*** ***n*-szer deriválható az *a* pontban** *(jelölése: $f \in D^n\{a\}$), ha*

- $\exists r > 0\colon f \in D^{n-1}\Big(K_r(a)\Big)$, *és*
- *az $f^{(n-1)}$ deriváltfüggvény deriválható a-ban, azaz $f^{(n-1)} \in D\{a\}$.*

Legyen ekkor

$$f^{(n)}(a) := \Big(f^{(n-1)}\Big)'(a)$$

az f függvény ***a*-beli *n*-edik deriváltja.**

Ha $H := \Big\{x \in \operatorname{int} \mathcal{D}_f \mid f \in D^n\{x\}\Big\} \neq \emptyset$, akkor $H \ni x \mapsto f^{(n)}(x)$ az f függvény ***n*-edik deriváltfüggvénye***, amit röviden az $f^{(n)}$ szimbólummal jelölünk.

Példa. Ha $f(x) := x^4$ $(x \in \mathbb{R})$, akkor

$$f'(x) = 4x^3 \ (x \in \mathbb{R}), \quad f''(x) = 12x^2 \ (x \in \mathbb{R}), \quad f^{(3)}(x) = 24x \ (x \in \mathbb{R}), \quad f^{(4)}(x) = 24 \ (x \in \mathbb{R}).$$

Ha egy f függvényre valamilyen $a \in \operatorname{int} \mathcal{D}_f$ pontban minden $n \in \mathbb{N}$ mellett teljesül, hogy $f \in D^n\{a\}$, akkor azt mondjuk, hogy az f függvény ***a*-ban végtelen sokszor** (vagy **akárhányszor**) deriválható. Ennek jelölésére az $f \in D^\infty\{a\}$ szimbólumot használjuk. Ha ez minden $a \in \operatorname{int} \mathcal{D}_f$ pontban igaz, akkor az ***f* függvény végtelen sokszor** (vagy **akárhányszor**) deriválható, amit röviden így jelölünk: $f \in D^\infty$.

Könnyen igazolható, hogy ha a p függvény egy polinom, akkor $p \in D^\infty$. Másrészt $\exp \in D^\infty$, $\sin \in D^\infty$, $\cos \in D^\infty$.

A deriválási szabályok némelyike magasabb rendű deriváltakra is átvihető.

1. Tétel. *Ha valamilyen $n \in \mathbb{N}$ esetén $f, g \in D^n\{a\}$, akkor*

a) $f + g \in D^n\{a\}$ *és* $(f+g)^{(n)}(a) = f^{(n)}(a) + g^{(n)}(a)$,

b) $f \cdot g \in D^n\{a\}$ *és* $(f \cdot g)^{(n)}(a) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(a) g^{(n-k)}(a)$ (**Leibniz-szabály**).

Bizonyítás. Mindkét állítás teljes indukcióval igazolható.

- Az a) bizonyítása szinte triviális.
- A b) belátása némi számolgatást igényel.

FÜGGVÉNYTULAJDONSÁGOK KAPCSOLATA A DERIVÁLTTAL 1.

Lokális szélsőértékek

Korábban már értelmeztük az abszolút szélsőértékek fogalmát. Célszerű bevezetni ezek lokális változatait.

4. Definíció. *Az $f \in \mathbb{R} \to \mathbb{R}$ függvénynek az $a \in \operatorname{int} \mathcal{D}_f$ pontban* **lokális maximuma van**, *ha*

$$\exists K(a) \subseteq \mathcal{D}_f : \forall\, x \in K(a) \subseteq \mathcal{D}_f \text{ esetén } f(x) \leq f(a).$$

Ekkor az a pontot ***f* lokális maximumhelyének** *nevezzük, az $f(a)$ függvényérték pedig a függvény* **lokális maximuma**.

Hasonlóan értelmezzük a ***lokális minimum*** fogalmát. A lokális maximumot, illetve minimumot közösen ***lokális szélsőértéknek***, a lokális maximumhelyet, illetve lokális minimumhelyet ***lokális szélsőértékhelynek*** nevezzük.

Megjegyzés. Az abszolút szélsőértékhely és a lokális szélsőértékhely fogalmai között a következő kapcsolat áll fenn.

- Egy abszolút szélsőértékhely nem szükségképpen lokális szélsőértékhely, mert a lokális szélsőértékhelynek feltétele, hogy a függvény értelmezve legyen a pont egy környezetében. Így például az $f(x) = x$ ($x \in [0, 1]$) függvénynek a 0 pontban abszolút minimuma van, de ez nem lokális minimumhely. Azonban, ha az $f : A \to \mathbb{R}$ függvénynek az $a \in A$ pontban abszolút szélsőértéke van, és A tartalmazza az a pont egy környezetét, akkor a lokális szélsőértékhely.

- Egy lokális szélsőértékhely nem szükségképpen abszolút szélsőértékhely, hiszen attól, hogy az f függvénynek az a pont egy környezetében nincs $f(a)$-nál nagyobb értéke, a környezeten kívül f felvehet $f(a)$-nál nagyobb értéket.

2. Tétel (A lokális szélsőértékre vonatkozó elsőrendű szükséges feltétel). *Tegyük fel, hogy $f \in \mathbb{R} \to \mathbb{R}$ és*

- *$f \in D\{a\}$ valamilyen $a \in \operatorname{int} \mathcal{D}_f$-ben,*
- *f-nek a-ban lokális szélsőértéke van.*

Ekkor $f'(a) = 0$.

Bizonyítás. Tegyük fel, hogy az $a \in \operatorname{int} \mathcal{D}_f$ pont lokális maximumhelye az $f \in D\{a\}$ függvénynek. Ekkor

$$\exists r > 0 \colon \forall x \in (a - r, a + r) \text{ esetén } f(x) \leq f(a).$$

Tekintsük az f függvény a-hoz tartozó különbséghányados-függvényét:

$$\frac{f(x) - f(a)}{x - a} \qquad \Big(x \in \mathcal{D}_f \setminus \{a\}\Big). \tag{1}$$

Ha $a < x < a + r$, azaz $x - a > 0$, akkor $f(x) \leq f(a)$ (vagyis $f(x) - f(a) \leq 0$) miatt (1)-ben szereplő különbséghányados nem pozitív. Mivel $f \in D\{a\}$, ezért

$$\lim_{x \to a+0} \frac{f(x) - f(a)}{x - a} = f'_+(a) = f'(a) \leq 0.$$

Ha $a - r < x < a$, azaz $x - a < 0$, akkor $f(x) \leq f(a)$ (vagyis $f(x) - f(a) \leq 0$) miatt (1)-ben szereplő különbséghányados nem negatív. Mivel $f \in D\{a\}$, ezért

$$\lim_{x \to a-0} \frac{f(x) - f(a)}{x - a} = f'_-(a) = f'(a) \geq 0.$$

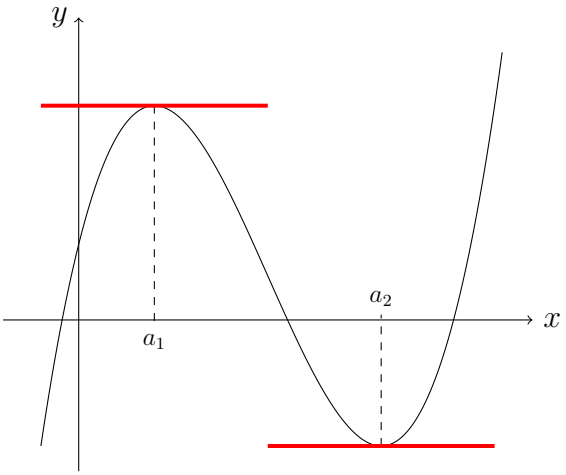
Azt kaptuk tehát, hogy $f'(a) \leq 0$ és $f'(a) \geq 0$, ami csak úgy lehetséges, ha $f'(a) = 0$.

A bizonyítás hasonló akkor is, ha a lokális minimumhelye az f függvénynek.

Megjegyzések.

1. A deriválható f függvénynek csak olyan pontban lehet lokális szélsőértéke, ahol a függvény deriváltja nulla. A lokális szélsőértékek meghatározásához tehát az $f'(x) = 0$ egyenletet kell megoldani.

2. A tételnek az a geometriai jelentése, hogy ha a függvény grafikonjának létezik érintője egy lokális szélsőértékpontban, akkor az érintő vízszintes, párhuzamos az x tengellyel.



3. Abból, hogy $f'(a) = 0$, nem következik, hogy az f függvénynek az $a \in \operatorname{int} \mathcal{D}_f$ pont lokális szélsőértékhelye. Például az

$$f(x) := x^3 \quad (x \in \mathbb{R}) \qquad \text{függvényre} \qquad f'(x) = 3x^2 \quad (x \in \mathbb{R}) \qquad \text{miatt} \qquad f'(0) = 0,$$

de a függvénynek nincs 0-ban lokális szélsőértéke (hiszen a függvény az egész számegyenesen szigorúan monoton növekedő). Ezt úgy is megfogalmazhatjuk, hogy ha f differenciálható a-ban, akkor az $f'(a) = 0$ csak szükséges, de nem elégséges feltétele annak, hogy az f függvénynek a-ban lokális szélsőértéke legyen.

A fenti példa motiválja a következő fogalom bevezetését.

5. Definíció. *Az $f \in \mathbb{R} \to \mathbb{R}$ függvénynek az $a \in \operatorname{int} \mathcal{D}_f$* **stacionárius pontja**, *ha $f \in D\{a\}$ és $f'(a) = 0$.*

A lokális szélsőértékre vonatkozó elsőrendű szükséges feltétel azt állítja, hogy deriválható függvénynek lokális szélsőértékhelyei a függvény stacionárius pontjaiban lehetnek. A fenti példa azonban azt mutatja, hogy lehetnek olyan stacionárius pontok, amelyek nem lokális szélsőértékhelyek. Fontos feladat tehát annak eldöntése, hogy egy stacionárius pont vajon lokális szélsőértékhely-e. Erre hamarosan jól használható eredményeket fogunk mutatni.

A differenciálszámítás középértéktételei

A középértéktételek olyan állítások, amelyek egy intervallumon értelmezett függvényről garantálják, hogy mindig van olyan pont az intervallum belsejében, amelyre bizonyos tulajdonság teljesül. Ilyen típusú állítás például a Bolzano tétele, miszerint ha $f \in C[a,b]$ és $f(a) \cdot f(b) < 0$, akkor $\exists \xi \in (a,b)\colon f(\xi) = 0$. Emlékezzük arról, hogy $f \in C[a,b]$ azt jelenti, hogy $[a,b] \in \mathcal{D}_f$, $f \in C\{x\} \quad \big(x \in (a,b)\big)$, illetve f jobbról folytonos a-ban és balról folytonos b-ben

Differenciálszámításnál a középértéktételek egy olyan pont létezését garantálják, ahol a függvény deriváltja olyan értéket vesz fel, ami a függvényhez kapcsolódik. Három ilyen tételt mutatjuk meg: a Rolle-, a Lagrange- és a Cauchy-féle középértéktételt. A középértéktételekkel fontos tulajdonságokat tudunk igazolni a függvény viselkedéséről.

3. Tétel (Rolle-féle középértéktétel). *Legyen $a, b \in \mathbb{R}$ és $a < b$. Ekkor*

$$\left.\begin{array}{l} \bullet\ f \in C[a,b], \\ \bullet\ f \in D(a,b), \\ \bullet\ f(a) = f(b) \end{array}\right\} \quad \Longrightarrow \quad \exists \xi \in (a,b)\colon f'(\xi) = 0.$$

Bizonyítás. A Weierstrass-tétel szerint, ha $f \in C[a,b]$, akkor f felveszi a maximumát és a minimumát az $[a,b]$ intervallumon, azaz

$$\exists \xi_1, \xi_2 \in [a,b]\colon f(\xi_1) = \min\{f(x) \mid x \in [a,b]\},\ f(\xi_2) = \max\{f(x) \mid x \in [a,b]\}$$

Ekkor két eset lehetséges

a) $\underline{f(\xi_1) = f(\xi_2)}$. Ekkor a maximum és a minimum megegyezik, azaz f konstans az $[a,b]$ intervallumon. Ezért $f'(\xi) = 0$ minden $\xi \in (a,b)$ esetén.

b) $\underline{f(\xi_1) \neq f(\xi_2)}$. Ekkor az $f(a) = f(b)$ feltétel miatt nem lehetséges, hogy ξ_1 és ξ_2 legyen az $[a,b]$ intervallum két végpontja. Tehát egyikük az intervallum belsejében van, amelyet ξ-vel fogjuk jelölni. Így már ξ lokális szélsőértékhelye lesz f-nek. Ezért a lokális szélsőértékre vonatkozó elsőrendű szükséges feltétel miatt $f'(\xi) = 0$.

Megjegyzés. A Rolle-féle középértéktételből következik, hogy ha $f \in C[a,b]$, $f \in D(a,b)$ és $\forall x \in (a,b)\colon f'(x) \neq 0$, akkor $f(a) \neq f(b)$.

2. Feladat. *Igazoljuk, hogy az*

$$e^x = x + 1$$

egyenletnek egyetlen egy megoldása van!

Megoldás. Vegyük észre, hogy az egyenlet megoldása egyben az

$$f(x) := e^x - x - 1 \qquad (x \in \mathbb{R})$$

függvény zérushelye is. Világos, hogy $a = 0$ megoldása az egyenletnek, azaz $f(a) = 0$. Azonban az

$$f'(x) = e^x - 1$$

deriváltfüggvény csak az $x = a = 0$ helyen lehet nulla az exponenciális függvény szigorú monotonitása miatt.

Ezért, ha $b > a$ tetszőleges szám, akkor $f \in C[a,b]$, $f \in D(a,b)$ és $\forall x \in (a,b)\colon f'(x) \neq 0$, amiből a Rolle-tétel alapján következik, hogy $f(a) \neq f(b)$, azaz $f(b) \neq 0$. Ez azt jelenti, hogy az egyenletnek nincs pozitív megoldása.

Hasonlóan igazolható, hogy az egyenletnek nincs negatív megoldása.

A Rolle-féle középértéktétel a következő módon általánosítható.

4. Tétel (Cauchy-féle középértéktétel). *Legyen $a, b \in \mathbb{R}$ és $a < b$. Ekkor*

$$\left.\begin{array}{l} \bullet\ f, g \in C[a,b], \\ \bullet\ f, g \in D(a,b), \\ \bullet\ \forall x \in (a,b)\colon g'(x) \neq 0 \end{array}\right\} \quad \Longrightarrow \quad \exists \xi \in (a,b)\colon \ \frac{f'(\xi)}{g'(\xi)} = \frac{f(b)-f(a)}{g(b)-g(a)}.$$

Bizonyítás. Először vegyük észere, hogy a Rolle-tételből következik, hogy $g(a) \neq g(b)$. Tekintsük az

$$F(x) := f(x) - \lambda g(x) \quad (x \in [a,b]), \qquad \text{ahol} \qquad \lambda := \frac{f(b)-f(a)}{g(b)-g(a)}$$

függvényt! Világos, hogy

$$f, g \in C[a,b] \implies F \in C[a,b] \qquad \text{és} \qquad f, g \in D(a,b) \implies F \in D(a,b).$$

Másrészt

$$F(a) = f(a) - \lambda g(a) = \ (\lambda \text{ definiciója szerint}) \ = f(b) - \lambda g(b) = F(b).$$

Ez azt jelenti, hogy F eleget tesz a Rolle-tétel feltételeinek, és így $\exists \xi \in (a,b)\colon F'(\xi) = 0$. Azonban

$$F'(x) = f'(x) - \lambda g'(x) \quad (x \in (a,b)),$$

tehát

$$f'(\xi) - \lambda g'(\xi) = F'(\xi) = 0 \qquad \Longrightarrow \qquad \frac{f'(\xi)}{g'(\xi)} = \lambda = \frac{f(b)-f(a)}{g(b)-g(a)}.$$

A Cauchy-féle középértéktételnek van egy fontos speciális esete.

5. Tétel (Lagrange-féle középértéktétel). *Legyen $a, b \in \mathbb{R}$ és $a < b$. Ekkor*

$$\left.\begin{array}{l} \bullet\ f \in C[a,b], \\ \bullet\ f \in D(a,b) \end{array}\right\} \quad \Longrightarrow \quad \exists \xi \in (a,b)\colon \ f'(\xi) = \frac{f(b)-f(a)}{b-a}.$$

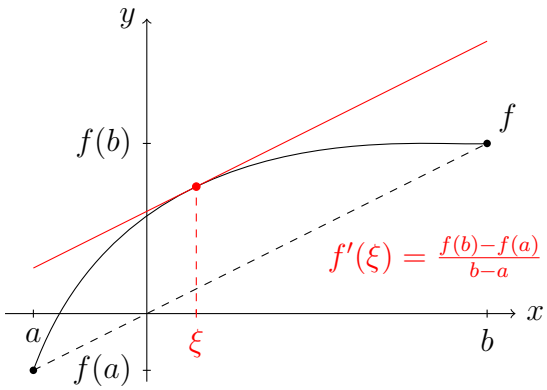
Bizonyítás. Legyen $g(x) := x$ a Cauchy-féle középértéktételben.

A Rolle-féle középértéktétel a Lagrange-féle középértéktétel speciális esete.

A Lagrange-féle középértéktétel geometriai jelentése: az $\big(a, f(a)\big)$ és $\big(b, f(b)\big)$ pontokon átmenő húr meredeksége

$$\frac{f(b)-f(a)}{b-a}.$$

Tehát f grafikonjának van olyan pontja, ahol az érintő párhuzamos ezzel a húrral.



A Lagrange-féle középértéktétel egyszerű, de fontos következményei a következő állítások:

6. Tétel (A deriváltak egyenlősége). *Legyen $a, b \in \mathbb{R}$ és $a < b$.*

1. *Tegyük fel, hogy $f : (a, b) \to \mathbb{R}$ és $f \in D(a, b)$. Ekkor*

$$f' \equiv 0 \ (a,b)\text{-}n \qquad \Longleftrightarrow \qquad f \equiv \text{állandó } (a,b)\text{-}n.$$

2. *Tegyük fel, hogy $f, g : (a, b) \to \mathbb{R}$ és $f, g \in D(a, b)$. Ekkor*

$$f' \equiv g' \ (a,b)\text{-}n \qquad \Longleftrightarrow \qquad \exists c \in \mathbb{R}, \forall x \in (a,b) \colon f(x) = g(x) + c.$$

Bizonyítás.

1. Az állítás $\boxed{\Leftarrow}$ irányát már ismerjük, ugyanis a konstansfüggvény deriváltja 0.

   A $\boxed{\Rightarrow}$ irány a Lagrange-féle középértéktétel következménye. Valóban, $\forall x, y \in (a, b)$, $x < y$ esetén $f \in C[x, y]$, $f \in D(x, y)$, és így

$$\exists \xi \in (x,y) \colon \frac{f(y) - f(x)}{y - x} = f'(\xi) = 0 \qquad \Longrightarrow \qquad f(x) = f(y).$$

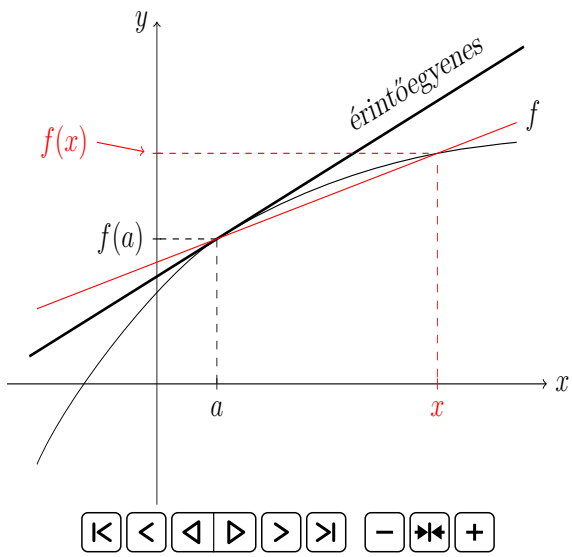
2. Alkalmazzuk az 1. állítást az $F := f - g$ függvényre.

Monotonitás

Az egyszerűség kedvéért csak intervallumon vizsgáljuk a monotonitást. Az $(a, b) \subseteq \mathbb{R}$ szimbólummal jelölünk egy korlátos vagy nem korlátos nyílt intervallumot, tehát $a = -\infty$ vagy $b = +\infty$ is lehetséges.

Egy függvény esetén a „monoton növekedő”, a „monoton csökkenő”, a „szigorúan monoton növekedő”, illetve a „szigorúan monoton csökkenő” kifejezések helyett gyakran a „$\nearrow$”, a „$\searrow$”, a „$\uparrow$”, illetve a „$\downarrow$” jeleket használjuk.

Az első fontos észrevétel az, hogy az első derivált előjeléből következtethetünk a függvény monotonitására. A következő animációban „látjuk”, hogy a lokális szélsőértékeken a függvényhez húzott érintő párhuzamos az x tengellyel, továbbá a szigorúan monoton növekvő szakaszokon az érintő meredeksége pozitív, illetve a szigorúan monoton csökkenő szakaszokon az érintő meredeksége negatív.



7. Tétel (A monotonitás és a derivált kapcsolata). *Legyen $(a,b) \subset \mathbb{R}$ egy nyílt intervallum. Tegyük fel, hogy $f \in D(a,b)$. Ekkor*

1. $f \nearrow$ [*illetve* $\searrow$] (a,b)-*n* $\iff$ $f' \geq 0$ [*illetve* $f' \leq 0$] (a,b)-*n;*
2. *ha* $f' > 0$ [*illetve* $f' < 0$] (a,b)-*n* $\implies$ $f \uparrow$ [*illetve* $\downarrow$] (a,b)-*n.*

Bizonyítás.

1. $\boxed{\Rightarrow}$ Ha $f \nearrow$ (a,b)-n és $t \in (a,b)$ egy tetszőleges pont, akkor

$$\frac{f(x)-f(t)}{x-t} \geq 0 \qquad (t < x < b),$$

hiszen $x - t > 0$ és a monotonitás miatt $f(x) - f(t) \geq 0$. Mivel $f \in D\{t\}$, így

$$f'(t) = f'_+(t) = \lim_{x \to t+0} \frac{f(x)-f(t)}{x-t} \geq 0.$$

$\boxed{\Leftarrow}$ Ha $\forall x \in (a,b)\colon f'(x) \geq 0$, akkor legyen $x, y \in (a,b)$, $x < y$ két tetszőleges pont. Ekkor $f \in C[x,y]$, $f \in D(x,y)$, és így a Lagrange-féle középértéktétel szerint

$$\exists \xi \in (x,y)\colon \frac{f(y)-f(x)}{y-x} = f'(\xi) \geq 0 \qquad \implies \qquad f(x) \leq f(y).$$

Ezért $f \nearrow$ (a,b)-n.

Az állítás hasonlóan igazolható monoton csökkenő függvények esetében is.

2. Alkalmazzuk „éles” egyenlőtlenségeket az 1. pont $\boxed{\Leftarrow}$ irányában.

Megjegyzések.

1. A tételben lényeges feltétel, hogy intervallumon értelmezett függvényről van szó. Például, ha

$$f(x) := \frac{1}{x} \quad \Big(x \in \mathbb{R} \setminus \{0\}\Big), \qquad \text{akkor} \qquad \forall x \in \mathcal{D}_f\colon f'(x) = -\frac{1}{x^2} < 0,$$

de az f függvény nem szigorúan monoton csökkenő a $\mathcal{D}_f = \mathbb{R} \setminus \{0\}$ halmazon, ami ***nem intervallum***.

2. Az 1. pont $\boxed{\Rightarrow}$ irányában nem tudunk „éles” egyenlőtlenségeket alkalmazni. ***A szigorú monotonitásra vonatkozó elégséges feltételek nem is fordíthatók meg.*** Például az

$$f(x) := x^3 \qquad (x \in \mathbb{R})$$

függvény szigorúan monoton növekedő az egész valós számok halmazán, de $f'(0) = 0$.

3. Nem nehéz igazolni (pl. átviteli elvvel), hogy ha f monoton az (a,b) intervallumon, és folytonos az intervallum egyik végpontján, akkor a monotonitás kiterjeszthető a végponttal bővített intervallumra. Ez az állítás szigorú monotonitás esetén is igaz. Így tudjuk az előbbi tételt kiterjeszteni tetszőleges intervallumra. Például, $f \in C[a,b]$ és $f \in D(a,b)$ esetén

$$f \nearrow [a,b]-n \quad \iff \quad f' \geq 0 \;\; (a,b)-n.$$

A lokális szélsőértékre vonatkozó elégséges feltételek

Az eddigiek alapján könnyen kaphatunk elégséges feltételeket arra, hogy egy függvénynek valamilyen pontban lokális szélsőértéke legyen.

Az egyik ahhoz kapcsolódik, hogy a deriváltfüggvény ***előjelet vált***. Azt mondjuk, hogy a $g \in \mathbb{R} \to \mathbb{R}$ függvény az $t \in \operatorname{int} \mathcal{D}_g$ pontban negatívból pozitívba megy át (röviden: $(-,+)$ előjelváltása van), ha $g(t) = 0$ és $\exists \delta > 0$ úgy, hogy $g < 0$ $(t-\delta, t)$-n és $g > 0$ $(t, t+\delta)$-n. A $(+,-)$ előjelváltást hasonló módon értelmezzük.

8. Tétel (A lokális szélsőértékre vonatkozó elsőrendű elégséges feltétel). *Legyen $a, b \in \mathbb{R}$, $a < b$ és $f : (a,b) \to \mathbb{R}$. Tegyük fel, hogy*

- *$f \in D(a,b)$,*
- *egy $c \in (a,b)$ pontban $f'(c) = 0$ és*
- *az f' deriváltfüggvény előjelet vált c-ben.*

Ekkor,

1. *ha az f' függvény a c pontban negatív értékből pozitív értékbe megy át, akkor c az f függvénynek lokális minimumhelye,*
2. *ha az f' függvény a c pontban pozitív értékből negatív értékbe megy át, akkor a c pont az f függvénynek lokális maximumhelye.*

Bizonyítás. Az állítás azonnal következik a monotonitás és a derivált kapcsolatárról szóló tételből, hiszen ha f-nek $(-,+)$ előjelváltása van a c pontban, akkor $\exists \delta > 0$ úgy, hogy $f' < 0$ $(c-\delta, c)$-n és $f' > 0$ $(c, c+\delta)$-n. Ezért $f \downarrow$ $(c-\delta, c]$-n és $f \uparrow$ $[c, c+\delta)$-n. Emiatt $\forall x \in (c-\delta, c+\delta)\colon f(x) > f(c)$, tehát c az f függvénynek lokális minimumhelye.

Az állítás hasonlóan igazolható $(+,-)$ előjelváltás esetén.

3. Feladat. *Határozzuk meg az*

$$f(x) = \frac{x+1}{x^2} \qquad (x \neq 0)$$

függvény lokális szélsőértékhelyeit és monotonitási intervallumait!

Megoldás. A tanult tételek értelmében előjelvizsgálatot fogunk végezni az

$$f'(x) = \frac{1 \cdot x^2 - (x+1) \cdot 2x}{x^4} = -\frac{x+2}{x^3} \qquad (x \neq 0)$$

deriváltfüggvény esetében. Mivel f' folytonos minden $x \neq 0$ pontban, így csak az $x = 0$ és az f' zérushelyei olyan pontok, ahol eltérhet f' előjele a pont bal és jobb oldali környezetén. A példában f'-nek egyetlen zérushelye van:

$$f'(x) = -\frac{x+2}{x^3} = 0 \qquad \Longrightarrow \qquad x = -2.$$

Ezzel három részintervallumot kapunk, ahol egységes f' előjele: $(-\infty, -2)$, $(-2, 0)$ és $(0, +\infty)$. Nem nehéz meghatározni f' előjeleit ezeken az intervallumokon, hiszen ehhez elegendő kiszámolni f' értékét az intervallumok egyik pontján. A következő táblázat tartalmazza az előjelvizsgálatot és ennek következményeit az f függvényre vonatkozóan.

	$x < -2$	-2	$-2 < x < 0$	$x > 0$
f'	$-$	0	$+$	$-$
f	$\downarrow$	$-1/4$	$\uparrow$	$\downarrow$
lok.		min		

A táblázatból rögtön leolvashatók f monotonitási szakaszait, illetve azt, hogy f-nek lokális minimumhelye van az $x = -2$ pontban, és lokális minimuma $f(-2) = -1/4$.

Kétszer differenciálható függvények esetében van egy alternatív módszer, amivel nem szükséges előjelvizsgálatot végezni, ha a függvényt kétszer deriváljuk.

9. Tétel (A lokális szélsőértékre vonatkozó másodrendű elégséges feltétel). *Legyen $a, b \in \mathbb{R}$, $a < b$ és $f : (a, b) \to \mathbb{R}$. Tegyük fel, hogy*

- *$f \in D^2\{c\}$, azaz kétszer deriválható egy $c \in (a, b)$ pontban,*
- *$f'(c) = 0$ és*
- *$f''(c) \neq 0$.*

Ekkor c lokális szélsőértékhelye az f függvénynek, és

1. *ha $f''(c) > 0$, akkor f-nek c-ben lokális minimuma van,*
2. *ha $f''(c) < 0$, akkor f-nek c-ben lokális maximuma van.*

Bizonyítás. Tegyük fel, hogy $f''(c) > 0$. Mivel

$$0 < f''(c) = \lim_{x \to c} \frac{f'(x) - f'(c)}{x - c} = \lim_{x \to c} \frac{f'(x) - 0}{x - c} = \lim_{x \to c} \frac{f'(x)}{x - c},$$

így a c pontnak van olyan bal oldali környezete, ahol $f' < 0$, van olyan jobb oldali környezete, ahol $f' > 0$. Tehát f'-nek $(-, +)$ előjelváltása van, ami azt jelenti, hogy f'-nek c-ben lokális minimuma van.

Az állítás hasonlóan igazolható $f''(c) < 0$ esetén.

Megjegyzések.

1. A 3. Feladatban $f \in D^2\{x\}$ minden $x \neq 0$ esetén, és

$$f''(x) = \left(-\frac{x+2}{x^3}\right)' = -\frac{1 \cdot x^3 - (x+2) \cdot 3x^2}{x^6} = \frac{2x+6}{x^4} \qquad \Longrightarrow \qquad f(-2) > 0.$$

Ezért $x = -2$ valóban lokális minimumhely.

2. Ha egyszerre $f'(c) = 0$ és $f''(c) = 0$, akkor c lehet lokális szélsőértékhely, de lehet, hogy nem. A különböző lehetőségeket mutatják például az $f(x) := x^3$, $f(x) := x^4$ és az $f(x) := -x^4$ $(x \in \mathbb{R})$ függvények a $c = 0$ helyen. Ebben az esetben további vizsgálatok kellenek.

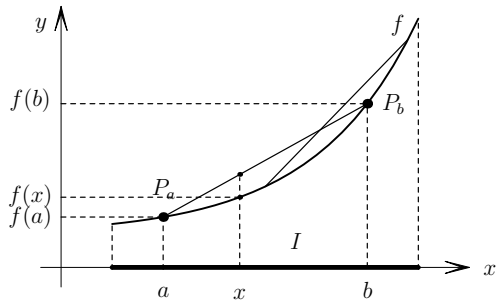
3. előadás

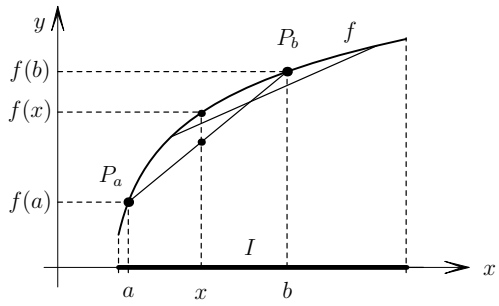
FÜGGVÉNYTULAJDONSÁGOK KAPCSOLATA A DERIVÁLTTAL 2.

Konvex és konkáv függvények

A konvex és konkáv függvények fogalmát már az Analízis I. kurzuson bevezettük. Ezzel az volt a célunk, hogy néhány alapfüggvény konvexitási tulajdonságát vizsgáljuk, és így teljes képet kapjunk ezekről a függvényekről. Nevezetesen a hatvány-, a reciprok és a gyökfüggvények esetében sikerült ezt „ügyes” számításokkal megvalósítani a definíció alapján. Azonban az exponenciális függvénynél azt mondtuk, hogy szigorú konvexitása elemi úton igazolható ugyan, de a differenciálszámítás eszköztárával ezt az állítást jóval egyszerűbben tudjuk bizonyítani.

Emlékezzünk, hogy egy függvény konvexitása bizonyos „alaki” tulajdonságaival van összefüggésben. Az mondtuk, hogy egy függvény szigorúan konvex (konkáv) egy I intervallumon, ha tetszőleges $a, b \in I$, $a < b$ pontpár esetén a függvény (a, b) intervallumhoz tartozó része az $(a, f(a))$ és $(b, f(b))$ pontokat összekötő húr alatt (felett) van. Az alábbi ábrák ezt illusztrálják. Az f függvény szigorúan konvex, és a g függvény szigorúan konkáv.





A szóban forgó húr egyenesének egyenlete:

$$y = \frac{f(b) - f(a)}{b - a}(x - a) + f(a), \qquad \text{vagy} \qquad y = \frac{f(b) - f(a)}{b - a}(x - b) + f(b).$$

A konvex és konkáv függvények pontos fogalma a következő:

1. Definíció. *Legyen $f \in \mathbb{R} \to \mathbb{R}$ és $I \subseteq \mathcal{D}_f$ egy intervallum. Ha $\forall a, b \in I$, $a < b$ esetén igaz az, hogy*

- *ha $\forall x \in (a, b)$: $f(x) \leq \dfrac{f(b) - f(a)}{b - a}(x - a) + f(a)$, akkor azt mondjuk, hogy az f függvény* **konvex az I intervallumon**,
- *ha $\forall x \in (a, b)$: $f(x) \geq \dfrac{f(b) - f(a)}{b - a}(x - a) + f(a)$, akkor azt mondjuk, hogy az f függvény* **konkáv az I intervallumon**,

Szigorú egyenlőtlenségek esetén **szigorúan konvex**, *illetve* **szigorúan konkáv** *függvényekről beszélünk.*

Megjegyzés. Mivel a húr egyenesének egyenlete kétféle módon írható fel, így a fenti definícióban szereplő

$$\frac{f(b)-f(a)}{b-a}(x-a)+f(a) \quad \text{kifejezés helyett az} \quad \frac{f(b)-f(a)}{b-a}(x-b)+f(b)$$

kifejezés is írható.

Most tegyük fel, hogy az f függvény konvex az I nyílt intervallumon, és $x \in I$ egy tetszőleges pont. Legyen tovább $u, v \in I$, $u < x < v$ két tetszőleges pont. Ekkor a fenti megjegyzés miatt egyszerre igaz, hogy

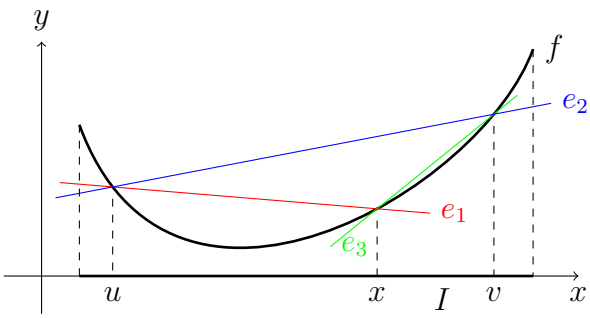
$$f(x) \leq \frac{f(v)-f(u)}{v-u}(x-u)+f(u) \qquad \text{és} \qquad f(x) \leq \frac{f(v)-f(u)}{v-u}(x-v)+f(v),$$

ami egyszerű ekvivalens átalakításokkal a következő módon írható át

$$(*) \qquad \frac{f(x)-f(u)}{x-u} \leq \frac{f(v)-f(u)}{v-u} \leq \frac{f(x)-f(v)}{x-v}.$$

Ha a függvény konkáv, szigorúan konvex vagy szigorúan konkáv, akkor a $(*)$-ban szereplő relációk irányán és élességén kell megfelelően változtatni.

Az ábra szemlélteti $(*)$ geometriai jelentését. Látható, hogy az e_1, az e_2 és az e_3 húrok meredekségei rendre a $(*)$-ban szereplő különbséghányadosok, és ezek nagysága megfelelnek a $(*)$-ban szereplő egyenlőtlenségeknek.



Megjegyzés. Legyen $a \in I$ és

$$F_a(x) := \frac{f(x)-f(a)}{x-a} \qquad (x \in I \setminus \{a\}).$$

Mivel $u < x < v$ tetszőleges pontok, így $(*)$-ból következik, hogy f akkor és csak akkor (szigorúan) konvex, ha az $F_a(x)$ különbséghányados-függvény (szigorúan) monoton növekvő. Konkáv függvények esetén csökkenő monotonitás lép fel.

1. Tétel (A konvexitás és a deriválhatóság kapcsolata). *Legyen $(a,b) \subseteq \mathbb{R}$ egy nyílt intervallum. Tegyük fel, hogy $f \in D^2(a,b)$. Ekkor*

1. *f konvex [illetve f konkáv] (a,b)-n $\iff$ $f'' \geq 0$ [illetve $f'' \leq 0$] (a,b)-n;*
2. *ha $f'' > 0$ [illetve $f'' < 0$] (a,b)-n $\implies$ f szigorúan konvex [illetve f szigorúan konkáv] (a,b)-n.*

Bizonyítás. Igazolható, hogy

- f konvex (a,b)-n $\iff$ $f' \nearrow (a,b)$-n,
- f szigorúan konvex (a,b)-n $\iff$ $f' \uparrow (a,b)$-n,
- f konkáv (a,b)-n $\iff$ $f' \searrow (a,b)$-n,
- f szigorúan konkáv (a,b)-n $\iff$ $f' \downarrow (a,b)$-n.

Ekkor a tétel állítása a monotonitás és a derivált kapcsolatáról szóló tételből következik, ha a tételt alkalmazzuk az f' függvényre, hiszen $f'' = (f')'$.

A négy állításból csak az elsőt fogjuk igazolni. A többi hasonlóan igazolható. Tegyük fel, hogy f konvex (a,b)-n, $u,v \in (a,b)$, $u<v$ két tetszőleges pont és $u<x<v$ egy tetszőleges pont. Ekkor $(*)$ miatt határátmenettel:

$$f'(u) = \lim_{x\to u} \frac{f(x)-f(u)}{x-u} \le \frac{f(v)-f(u)}{v-u} \quad \text{és} \quad f'(v) = \lim_{x\to v} \frac{f(x)-f(v)}{x-v} \ge \frac{f(v)-f(u)}{v-u}$$

amiből $f'(u) \le f'(v)$ következik, azaz f' monoton növekvő az (a,b) intervallumon.

Fordítva, tegyük fel, hogy az f' derivált függvény monoton növekvő az (a,b) intervallumon és vegyük az $a<u<x<v<b$ tetszőleges pontokat. Ekkor a Lagrange-féle középértéktétel szerint létezik egy $u<\xi_1<x$ és egy $x<\xi_2<v$ szám, hogy

$$f'(\xi_1) = \frac{f(x)-f(u)}{x-u}, \qquad f'(\xi_2) = \frac{f(v)-f(x)}{v-x}.$$

A monotonitás miatt $f'(\xi_1) \le f'(\xi_2)$, azaz

$$\frac{f(x)-f(u)}{x-u} \le \frac{f(v)-f(x)}{v-x},$$

ami a következő módon alakítható át:

$$\begin{aligned}
(f(x)-f(u))(v-x) &\le (f(v)-f(x))(x-u)\\
f(x)\Big((v-x)+(x-u)\Big) &\le f(u)(v-x)+f(v)(x-u)\\
f(x)(v-u) &\le f(u)\Big((v-u)-(x-u)\Big)+f(v)(x-u)\\
f(x)(v-u) &\le (f(v)-f(u))(x-u)+f(u)(v-u)\\
f(x) &\le \frac{f(v)-f(u)}{v-u}(x-u)+f(u).
\end{aligned}$$

Ez azt jelenti, hogy f konvex függvény.

Megjegyzés. Az előző tétel értelmében az exponenciális függvény szigorúan konvex $\mathbb{R}$-n azért, mert $\forall x \in \mathbb{R}\colon (\exp)''(x) = \exp(x) > 0$.

2. Definíció. *Legyen $(a,b) \subseteq \mathbb{R}$ egy nyílt intervallum, és tegyük fel, hogy $f \in D(a,b)$. Ekkor azt mondjuk, hogy a $c \in (a,b)$ pont az f függvénynek* **inflexiós pontja**, *ha*

$$\exists \delta > 0\colon f \text{ konvex } (c-\delta, c]\text{-n} \quad \text{és} \quad \text{konkáv } [c, c+\delta)\text{-n}, \quad \text{vagy fordítva.}$$

Megjegyzések.

1. A konvexitás és a deriválhatóság kapcsolatáról szóló tétel értelmében, ha $f \in D^2(a,b)$ és $c \in (a,b)$ inflexiós pont, akkor f'' előjele különbözik c egy bal és egy jobb oldali környezetében. Ha még $f'' \in C\{c\}$, akkor ez csak akkor lehetséges, ha $f''(c) = 0$. Ezt hívjuk ***az inflexiós pontra vonatkozó másodrendű szükséges feltételnek***.

2. Igazolható, hogy a függvény inflexiós pontjain az érintő „átszeli” a függvény grafikonját szigorú konvexitási váltás mellet. Pontosabban, ha e jelöli az f függvény érintőjét a c pontban, akkor a $\varphi := f - e$ függvénynek előjelváltása van a c pontban. Ez a jelenség könnyen megfigyelhető az $f(x) := x^3$ $(x \in \mathbb{R})$ függvény esetén a $c = 0$ pontban, ahol a függvénynek inflexiós pontja van, és ott az érintő egyenlete $e(x) = 0$ $(x \in \mathbb{R})$.

1. Feladat. *Vizsgáljuk meg konvexitás szempontjából a következő függvényt*

$$f(x) = \frac{x+1}{x^2} \qquad (x \neq 0),$$

és határozzuk megy az inflexiós pontjait!

Megoldás. Előjelvizsgálatot fogunk végezni a függvény második deriváltjával, amelyet már kiszámoltunk egy korábbi előadásban:

$$f''(x) = \frac{2x+6}{x^4}, \qquad (x \neq 0).$$

Ennek egyetlen zérushelye az $x = -3$. Ekkor a monotonitáshoz hasonló táblázatot készítünk, de ennek első sorában f'' szerepel. A $\smile$ szimbólummal jelöljük azt, hogy a függvény konvex, a $\frown$ szimbólummal pedig azt, hogy a függvény konkáv az adott intervallumon.

	$x < -3$	-3	$-3 < x < 0$	$x > 0$
f''	$-$	0	$+$	$+$
f	$\frown$	$-2/9$	$\smile$	$\smile$
		infl.		

A táblázatból rögtön leolvasható, mely intervallumokon lesz konvex vagy konkáv a függvény, illetve az, hogy f-nek inflexiós pontja van az $x = -3$ helyen, ahol a függvény értéke $f(-3) = -2/9$.

Aszimptoták

Ha egy függvény grafikonja nem korlátos, akkor úgy kell ábrázolni, hogy érzékelhető legyen az a tendencia, amit a függvény követ, amikor „elhagyja” a rajzterületet. Ehhez a függvény határérték nagyon fontos segítséget nyújt. Előfordul, hogy a grafikon pontjai tetszőleges közelségbe kerülnek egy adott egyeneshez, ún. ***aszimptotához***. A legegyszerűbb ilyen eset, amikor a függvénynek egy a pontban van bal vagy jobb oldali határértéke, és ez $-\infty$ vagy $+\infty$. Ekkor az $x = a$ egyenletű egyenes egy függőleges aszimptotája lesz a függvénynek.

Ha a függvény értelmezési tartománya nem korlátos, akkor lehet vízszintes aszimptotája is. Ez akkor fordul elő, amikor létezik a függvény határértéke a $-\infty$-ben vagy a $+\infty$-ben, és ez egy A számmal egyenlő. Ekkor $y = A$ az aszimptota egyenlete. De előfordulnak más esetek is.

3. Definíció. *Legyen $a \in \mathbb{R}$ és $f : (a, +\infty) \to \mathbb{R}$. Azt mondjuk, hogy az f függvénynek* **van aszimptotája** $(+\infty)$**-ben**, *ha*

$$\exists l(x) = Ax + B \quad (x \in \mathbb{R})$$

elsőfokú függvény, amelyre

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

Ekkor az $y = Ax + B$ egyenletű egyenes az f függvény **aszimptotája** $(+\infty)$**-ben**.

A függvény $(-\infty)$**-beli aszimptotáját** *is hasonló módon értelmezzük.*

2. Tétel. *Legyen $a \in \mathbb{R}$. Az $f : (a, +\infty) \to \mathbb{R}$ függvénynek akkor és csak akkor van aszimptotája $(+\infty)$-ben, ha léteznek és végesek a következő határértékek:*

$$\lim_{x\to+\infty} \frac{f(x)}{x} =: A \in \mathbb{R} \qquad \text{és} \qquad \lim_{x\to+\infty} \Big(f(x) - Ax\Big) =: B \in \mathbb{R}.$$

Ekkor az

$$l(x) = Ax + B \quad (x \in \mathbb{R})$$

egyenes az f függvény aszimptotája $(+\infty)$-ben.

Bizonyítás. A tétel könnyen igazolható a határérték alaptulajdonságai alapján.

Megjegyzés. Hasonló állítás érvényes a $(-\infty)$-beli aszimptoták meghatározására.

A L'Hospital-szabály

Ebben a fejezetben láttuk, hogy a differenciálszámítás eszköztárával milyen hatékonyan tudunk meghatározni több függvénytulajdonságot. Az aszimptoták kivételnek tűnnek, mert ehhez határértékszámítás szükséges. Most megmutatjuk, hogy a differenciálszámítás is remek módon alkalmazható függvények határértékének kiszámításában.

Először lássuk hogyan számítjuk a következő határértéket az eddigi ismereteink alapján:

$$\lim_{x\to1} \frac{x^8 + 3x^5 - 2x^2 - 2}{x^4 - 6x^2 + 5x} = \lim_{x\to1} \frac{\cancel{(x-1)}(x^7 + x^6 + x^5 + 4x^4 + 4x^3 + 4x^2 + 2x + 2)}{\cancel{(x-1)}(x^3 + x^2 - 5x)} = -\frac{19}{3}.$$

Egy 0/0 típusú kritikus határértékről van szó. A számlalóból és a nevezőből kiemeltük az $x-1$ tényezőt, és egyszerűsítés után kiértékeltük $x = 1$-re a fennmaradó kifejezést. Itt a problémát a kiemelés jelenti, amire több módszer ismert, de ezek általában sok számítással járnak. Most megpróbálkozunk valami mással. Legyen $a = 1$, illetve

$$f(x) := x^8 + 3x^5 - 2x^2 - 2 \quad (x \in \mathbb{R}) \qquad \text{és} \qquad g(x) := x^4 - 6x^2 + 5x \quad (x \in \mathbb{R}).$$

Tudjuk, hogy f és g differenciálható függvények,

$$f'(x) = 8x^7 + 15x^4 - 4x \quad (x \in \mathbb{R}) \qquad \text{és} \qquad g'(x) = 4x^3 - 12x + 5 \quad (x \in \mathbb{R}).$$

Mivel $f, g \in D\{a\}$, $f(a) = g(a) = 0$ és $g'(a) \neq 0$, így

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f(x) - f(a)}{g(x) - g(a)} = \lim_{x\to a} \frac{\frac{f(x)-f(a)}{x-a}}{\frac{g(x)-g(a)}{x-a}} = \frac{\lim\limits_{x\to a} \dfrac{f(x) - f(a)}{x - a}}{\lim\limits_{x\to a} \dfrac{g(x) - g(a)}{x - a}} = \frac{f'(a)}{g'(a)}.$$

Ha az $f'(a) = f'(1) = 19$ és a $g'(a) = g'(1) = -3$ értékeket egymással elosztjuk, akkor megkapjuk a már ismert végeredményt. A határérték kiszámításához „csak" külön kellett deriválni a számlálót és a nevezőt, és az így kapott hányadost kiértékelni az $x = a$ pontban.

A most bemutatott módszer csak a fent megadott feltételek mellett alkalmazható. Több olyan eset van, amikor az $\overline{\mathbb{R}}$-beli műveletek és a határérték kapcsolatára vonatkozó tételek nem alkalmazhatók, ezek az ún. ***kritikus határértékek***. Ilyenek például a következők:

$$(+\infty) + (-\infty), \quad 0 \cdot (\pm\infty), \quad \frac{0}{0}, \quad \frac{\pm\infty}{\pm\infty}, \quad 0^0.$$

Szerencsére a bemutatott módszer csak egy speciális esete egy jóval általánosabb állításnak, amit L'Hospital-szabálynak nevezünk.

3. Tétel (L'Hospital-szabály a $0/0$ esetben). *Legyen* $(-\infty \le a < b < +\infty)$, *illetve* $f, g \in \mathbb{R} \to \mathbb{R}$. *Ekkor*

$$\left.\begin{array}{l} \bullet\ f, g \in D(a,b), \\ \bullet\ \forall x \in (a,b)\colon g'(x) \neq 0, \\ \bullet\ \lim\limits_{a+0} f = \lim\limits_{a+0} g = 0, \\ \bullet\ \exists \lim\limits_{a+0} \dfrac{f'}{g'} \in \overline{\mathbb{R}}. \end{array}\right\} \quad \Longrightarrow \quad \exists \lim_{a+0} \frac{f}{g} \quad \textit{és} \quad \lim_{a+0} \frac{f}{g} = \lim_{a+0} \frac{f'}{g'}.$$

Bizonyítás. Először vegyük észre, hogy a Rolle-tétel alapján, ha $g' \neq 0$ (a,b)-n, akkor g-nek legfeljebb egy zérushelye van (a,b)-n. Ekkor választhatunk olyan $b > a$ számot, hogy $g \neq 0$ (a,b)-n.

Két esetet fogunk megkülönböztetni

1. $\boxed{a > -\infty \text{ (véges).}}$ Legyen $A := \lim\limits_{a+0} \dfrac{f'}{g'} \in \overline{\mathbb{R}}$. Azt kell igazolni, hogy $\lim\limits_{a+0} \dfrac{f}{g} = A$, azaz

$$(\#) \qquad \forall \varepsilon > 0\text{-hoz } \exists \delta > 0,\ \forall x \in (a, a+\delta) \subseteq (a,b)\colon \frac{f(x)}{g(x)} \in K_\varepsilon(A).$$

Az $A = \lim\limits_{a+0} \dfrac{f'}{g'}$ feltétel azt jelenti, hogy

$$(\#\#) \qquad \forall \varepsilon > 0\text{-hoz } \exists \delta > 0,\ \forall y \in (a, a+\delta) \subseteq (a,b)\colon \frac{f'(y)}{g'(y)} \in K_\varepsilon(A).$$

Értelmezzük az f és a g függvényt az a pontban úgy, hogy

$$f(a) := 0 \quad \text{és} \quad g(a) := 0.$$

A $\lim\limits_{a+0} f = \lim\limits_{a+0} g = 0$ feltételből következik, hogy ekkor $f, g \in C[a, a+\delta)$.

Legyen most $x \in (a, a+\delta)$ tetszőleges pont. A Cauchy-féle középértéktétel feltételei az f és a g függvényre az $[a,x]$ intervallumon teljesülnek. Így $\exists\, \xi_x \in (a,x)$, amelyre

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(\xi_x)}{g'(\xi_x)} \text{ (és ez } (\#\#) \text{ miatt) } \in K_\varepsilon(A).$$

A $(\#)$ állítást tehát bebizonyítottuk. A $\lim\limits_{a+0} \dfrac{f}{g}$ határérték létezik, és $\lim\limits_{a+0} \dfrac{f}{g} = A$.

2. $\boxed{a = -\infty.}$ Legyen $-\infty < c < \min\{0, b\}$, $d := -1/c$ és

$$F(x) := f(-1/x), \quad G(x) := g(-1/x) \qquad \big(x \in (0,d)\big).$$

Ekkor F és G kielégítik a $0/0$ esetre vonatkozó L'Hospital-szabály feltételeit $(0,d)$-n, illetve az összetett függvény deriváltja szerint

$$F'(x) := f'(-1/x) \cdot \frac{1}{x^2}, \quad G(x) := g'(-1/x) \cdot \frac{1}{x^2} \qquad \big(x \in (0,d)\big).$$

Így az $x = -1/y$ helyettesítéssel

$$\lim_{y\to-\infty} \frac{f(y)}{g(y)} = \lim_{x\to 0+0} \frac{F(x)}{G(x)} = \text{(L'Hospital)} = \lim_{x\to 0+0} \frac{F'(x)}{G'(x)} = \lim_{x\to 0+0} \frac{\frac{1}{x^2} f'(-1/x)}{\frac{1}{x^2} g'(-1/x)} =$$

$$= \lim_{x\to 0+0} \frac{f'(-1/x)}{g'(-1/x)} = \lim_{y\to-\infty} \frac{f'(y)}{g'(y)}.$$

Most megfogalmazzuk a $(+\infty)/(+\infty)$ kritikus határértékre vonatkozó L'Hospital-szabályt.

4. Tétel (L'Hospital-szabály a $(+\infty)/(+\infty)$ esetben). *Legyen $(-\infty \le a < b < +\infty)$, illetve $f, g \in \mathbb{R} \to \mathbb{R}$. Ekkor*

$$\left.\begin{array}{l} \bullet\ f, g \in D(a,b), \\ \bullet\ \forall x \in (a,b)\colon g'(x) \neq 0, \\ \bullet\ \lim\limits_{a+0} f = \lim\limits_{a+0} g = +\infty, \\ \bullet\ \exists \lim\limits_{a+0} \dfrac{f'}{g'} \in \overline{\mathbb{R}}. \end{array}\right\} \quad \Longrightarrow \quad \exists \lim_{a+0} \frac{f}{g} \quad \text{és} \quad \lim_{a+0} \frac{f}{g} = \lim_{a+0} \frac{f'}{g'}.$$

Bizonyítás. Nem bizonyítjuk.

Megjegyzések.

1. Ha $(-\infty < a < b \le +\infty)$, akkor a L'Hospital-szabály minkét esete átfogalmazható a ***b* pontbeli bal oldali határértékre** (így tehát $+\infty$-ben vett határértékekre is). Ehhez elegendő az $y = -x$ helyettesítést elvégezni. Így a L'Hospital-szabály könnyen kétoldali határértékekre is átfogalmazható.

2. A $(+\infty)/(+\infty)$ esetre vonatkozó L'Hospital-szabály alkalmazható a $(+\infty)/(-\infty)$, a $(-\infty)/(+\infty)$ és a $(-\infty)/(-\infty)$ típusú esetekre is.

3. Mivel az $a < b$ értékeket úgy választhatjuk, hogy egymáshoz tetszőleges közelségben legyenek, így a $\forall x \in (a,b)\colon g'(x) \neq 0$ feltétel szinte mindig előfordul a gyakorlatban.

4. A L'Hospital-szabály alkalmazása előtt győződjünk meg, hogy a határérték teljesíti a szabály alkalmazásához szükséges feltételeket. Például, ha

$$f(x) := \cos x, \quad \text{és} \quad g(x) := x + 1 \qquad (x \in \mathbb{R}),$$

akkor

$$\lim_0 \frac{f}{g} = \lim_{x\to 0} \frac{\cos x}{x+1} = \frac{1}{0+1} = 1, \quad \text{de} \quad \lim_0 \frac{f'}{g'} = \lim_{x\to 0} \frac{-\sin x}{1} = 0 \neq \lim_0 \frac{f}{g}.$$

5. A L'Hospital-szabály többször egymás után is alkalmazható, ha szükséges. Például

$$\lim_{x\to 0} \frac{1-\cos x}{x^2} = \left(\frac{0}{0}\right) \overset{\text{L'Hospital}}{=} \lim_{x\to 0} \frac{\sin x}{2x} = \left(\frac{0}{0}\right) \overset{\text{L'Hospital}}{=} \lim_{x\to 0} \frac{\cos x}{2} = \frac{1}{2}.$$

Azonban előfordul, hogy ez soha nem vezet eredményhez. Például

$$\lim_{x\to+\infty} \frac{x}{\sqrt{1+x^2}} = \left(\frac{+\infty}{+\infty}\right) \overset{\text{L'Hospital}}{=} \lim_{x\to+\infty} \frac{1}{\frac{1}{2\sqrt{1+x^2}} 2x} = \lim_{x\to+\infty} \frac{\sqrt{1+x^2}}{x},$$

és ha még egyszer alkalmazzuk a L'Hospital-szabályt, akkor az induló kifejezést kapjuk.

6. A L'Hospital-szabály nem fordítható meg abban az értelemben, hogy ha elvégzése után a kapott határérték nem létezik, akkor attól még lehet, hogy a keresett határérték létezik. Például

$$\lim_{x\to+\infty}\frac{x+\sin x}{x}=\left(\frac{+\infty}{+\infty}\right)\overset{\text{L'Hospital}}{=}\lim_{x\to+\infty}\frac{1+\cos x}{1}\ \nexists,\ \text{de}\ \lim_{x\to+\infty}\frac{x+\sin x}{x}=\lim_{x\to+\infty}\left(1+\frac{\sin x}{x}\right)=1.$$

Ha $a>1$ és $1\le n\in\mathbb{N}$, akkor a L'Hospital-szabály n-szeri alkalmazásával azt kapjuk, hogy

$$\lim_{x\to+\infty}\frac{a^x}{x^n}=\left(\frac{+\infty}{+\infty}\right)\overset{\text{L'Hospital}}{=}\lim_{x\to+\infty}\frac{(\ln a)\cdot a^x}{n\cdot x^{n-1}}=\left(\frac{+\infty}{+\infty}\right)\overset{\text{L'Hospital}}{=}\lim_{x\to+\infty}\frac{(\ln a)^2\cdot a^x}{n\cdot(n-1)\cdot x^{n-2}}=\cdots=$$

$$=\left(\frac{+\infty}{+\infty}\right)\overset{\text{L'Hospital}}{=}\lim_{x\to+\infty}\frac{(\ln a)^{n-1}\cdot a^x}{n(n-1)\cdot\ldots\cdot 2\cdot x}=\left(\frac{+\infty}{+\infty}\right)\overset{\text{L'Hospital}}{=}\lim_{x\to+\infty}\frac{(\ln a)^n\cdot a^x}{n!}=+\infty.$$

Ezt úgy is kifejezhetjük, hogy ha $a>1$, akkor $x\to+\infty$ esetén az a^x ($x\in\mathbb{R}$) függvény gyorsabban tart $(+\infty)$-hez, mint x bármelyik pozitív kitevőjű hatványa, és ezt szokás így is jelölni:

$$x^n\ll a^x,\quad \text{ha}\ \ x\ \ \text{elég nagy.}$$

Hasonlóan, ha $1\le m,n\in\mathbb{N}$, akkor a L'Hospital-szabály n-szeri alkalmazásával azt kapjuk, hogy

$$\lim_{x\to+\infty}\frac{\ln^n x}{x^m}=\left(\frac{+\infty}{+\infty}\right)\overset{\text{L'Hospital}}{=}\lim_{x\to+\infty}\frac{n\cdot\ln^{n-1}x}{m\cdot x^m}=\cdots=\left(\frac{+\infty}{+\infty}\right)\overset{\text{L'Hospital}}{=}\lim_{x\to+\infty}\frac{n!}{m^n\cdot x^m}=0,$$

azaz x bármely pozitív kitevőjű hatványa gyorsabban tart $(+\infty)$-hez $x\to+\infty$ esetén, mint $\ln x$ bármely pozitív kitevőjű hatványa. Röviden: minden $1\le n,m\in\mathbb{N}$ esetén

$$\Big(\ln x\Big)^n\ll x^m,\quad \text{ha}\ \ x\ \ \text{elég nagy.}$$

A többi kritikus határértéktípust gyakran vissza lehet vezetni $0/0$ vagy $(+\infty)/(+\infty)$ típusú határértékre, és így megpróbálhatjuk alkalmazni a L'Hospital-szabályt.

Példák.

1. $$\lim_{x\to0+0}x\ln x=\Big(0\cdot(-\infty)\Big)=\lim_{x\to0+0}\frac{\ln x}{\frac{1}{x}}=\left(\frac{-\infty}{+\infty}\right)\overset{\text{L'Hospital}}{=}\lim_{x\to0+0}\frac{\frac{1}{x}}{-\frac{1}{x^2}}=\lim_{x\to0+0}(-x)=0.$$

2. $$\lim_{x\to0}\left(\frac{1}{\sin x}-\frac{1}{x}\right)=(?)=\lim_{x\to0}\frac{x-\sin x}{x\ \sin x}=\left(\frac{0}{0}\right)\overset{\text{L'Hospital}}{=}\lim_{x\to0}\frac{1-\cos x}{\sin x+x\ \cos x}=$$
$$=\left(\frac{0}{0}\right)\overset{\text{L'Hospital}}{=}\lim_{x\to0}\frac{\sin x}{\cos x+\cos x-x\ \sin x}=\frac{0}{1+1-0}=0.$$

3. $$\lim_{x\to0+0}x^x=\Big(0^0\Big)=\lim_{x\to0+0}\Big(e^{\ln x}\Big)^x=\lim_{x\to0+0}e^{x\ln x}=\exp\Big(\lim_{x\to0+0}x\ln x\Big)=\exp(0)=1,$$

   hiszen exp folytonos függvény, és már igazoltuk, hogy $\lim\limits_{x\to0+0}x\ln x=0$.

Teljes függvényvizsgálat

Adott f valós-valós függvény ***teljes függvényvizsgálatán*** f analitikus és geometriai tulajdonságainak a megállapítását értjük. Ennek során a következőket kell meghatározni:

1. Kezdeti vizsgálatok. (Deriválhatóság, zérushelyek, előjelvizsgálat, paritás, periodicitás megállapítása.)
2. Lokális szélsőértékek és monotonitási intervallumok.
3. Konvexitási intervallumok és inflexiós pontok.
4. Határértékek és aszimptoták.
5. A függvény grafikonjának felrajzolása.

2. Feladat. *Teljes függvényvizsgálat végzése után vázolja az*

$$f(x) := x - 1 - \frac{4x}{1+x^2} \qquad (x \in \mathbb{R})$$

függvény grafikonját!

Megoldás.

1. ***Kezdeti vizsgálatok***. A deriválási szabályok alapján az f függvény minden $x \in \mathbb{R}$ pontban akárhányszor deriválható. Mivel

$$f(x) = x - 1 - \frac{4x}{x^2+1} = \frac{x^3 - x^2 - 3x - 1}{x^2+1} = \frac{(x+1)(x^2-2x-1)}{x^2+1},$$

így $f(x) = 0 \iff x = -1$ vagy $x^2 - 2x - 1 = 0$, azaz

$$x = -1, \qquad x = x_1 := 1 - \sqrt{2} \approx -0,414 \qquad \text{vagy} \qquad x = x_2 := 1 + \sqrt{2} \approx 2,414.$$

Előjelvizsgálat

	$x < -1$	-1	$-1 < x < x_1$	x_1	$x_1 < x < x_2$	x_2	$x > x_2$
f	$-$	0	$+$	0	$-$	0	$+$

A függvény nem páros, páratlan vagy periodikus.

2. ***Monotonitás***. Minden $x \in \mathbb{R}$ esetén

$$f'(x) = 1 - \frac{4 \cdot (x^2+1) - 4x \cdot 2x}{(x^2+1)^2} = \frac{x^4 + 6x^2 - 3}{(x^2+1)^2},$$

ezért

$$f'(x) = 0 \qquad \iff \qquad x^4 + 6x^2 - 3 = 0.$$

A $t = x^2$ helyettesítéssel

$$t^2 + 6t - 3 = 0 \qquad \iff \qquad t = -3 - 2\sqrt{3} < 0, \quad \text{vagy} \quad t = -3 + 2\sqrt{3} > 0.$$

Ezért $x^2 = 2\sqrt{3} - 3 \qquad \Longleftrightarrow$

$$x = x_3 := \sqrt{2\sqrt{3} - 3} \approx 0,681 \quad \text{vagy} \quad x = -x_3 := -\sqrt{2\sqrt{3} - 3} \approx -0,681$$

	$x < -x_3$	$-x_3$	$-x_3 < x < x_3$	x_3	$x > x_3$
f'	$+$	0	$-$	0	$+$
f	$\uparrow$	$0,18$	$\downarrow$	$-2,18$	$\uparrow$
lok.		max		min	

$\boxed{3.}$ ***Konvexitás***. Minden $x \in \mathbb{R}$ esetén

$$f''(x) = \frac{(4x^3 + 12x)(x^2+1)^2 - (x^4 + 6x^2 - 3) \cdot 2(x^2+1) \cdot 2x}{(x^2+1)^4} = \frac{8x(3 - x^2)}{(x^2+1)^3},$$

ezért $f''(x) = 0 \quad \Longleftrightarrow \quad x = 0, \quad x = -\sqrt{3} \approx -1,73, \quad$ vagy $\quad x = \sqrt{3} \approx 1,73.$

	$x < -\sqrt{3}$	$-\sqrt{3}$	$-\sqrt{3} < x < 0$	0	$0 < x\sqrt{3}$	$\sqrt{3}$	$x > \sqrt{3}$
f''	$+$	0	$-$	0	$+$	0	$-$
f	$\smile$	-1	$\frown$	-1	$\smile$	-1	$\frown$
		infl.		infl.		infl.	

$\boxed{4.}$ ***Határértékek és aszimptoták***. Mivel

$$\lim_{x\to\pm\infty} \frac{4x}{1+x^2} = \left(\frac{\pm\infty}{+\infty}\right) \overset{\text{L'Hospital}}{=} \lim_{x\to\pm\infty} \frac{4}{2x} = \lim_{x\to\pm\infty} \frac{2}{x} = 0,$$

ezért

$$\lim_{x\to\pm\infty} f(x) = \lim_{x\to\pm\infty} \left(x - 1 - \frac{4x}{1+x^2}\right) = \pm\infty - 1 - 0 = \pm\infty.$$

Továbbá

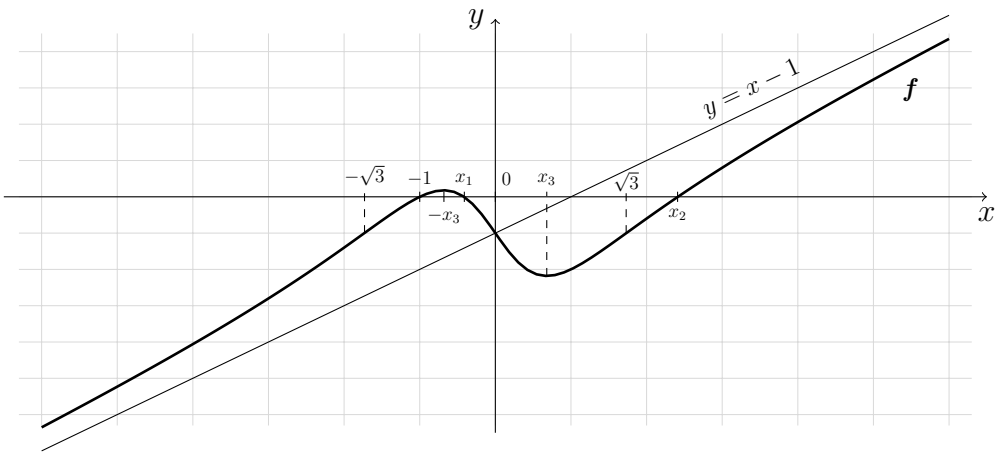
$$\lim_{x\to\pm\infty} \frac{f(x)}{x} = \lim_{x\to\pm\infty} \left(1 - \frac{1}{x} - \frac{4}{1+x^2}\right) = 1 =: A$$

és

$$\lim_{x\to\pm\infty} (f(x) - 1 \cdot x) = \lim_{x\to\pm\infty} \left(-1 - \frac{4x}{1+x^2}\right) = -1 =: B.$$

Ez azt jelenti, hogy az $y = Ax + B$, azaz az $y = x - 1$ egyenletű egyenes az f függvény aszimptotája $(+\infty)$-ben és $(+\infty)$-ben is.

$\boxed{5.}$ ***Ábrázolás***.



4. előadás

SPECIÁLIS FÜGGVÉNYEK

Már elemzett speciális függvények

Az Analízis I. kurzuson több speciális függvénnyel foglalkoztunk, megadtuk és igazoltuk tulajdonságaikat, ábrázoltuk a grafikonjukat. Első körben sikerült teljesen elemezni a

- ***hatványfüggvényeket:*** $f(x) := x^n \quad (x \in \mathbb{R})$, ahol $n = 0, 1, 2, \dots,$
- ***reciprokfüggvényeket:*** $f(x) := 1/x^n \quad \big(x \in \mathbb{R} \setminus \{0\}\big)$, ahol $n = 1, 2, 3, \dots,$
- ***gyökfüggvényeket:*** $f(x) := \sqrt[n]{x} \quad \big(x \in [0, +\infty)\big)$, ahol $n = 2, 3, 4, \dots.$

A következő körben az exponenciális függvényre és logaritmusfüggvényre került sor. Velük kapcsolatban a következő függvényeket is elemeztük:

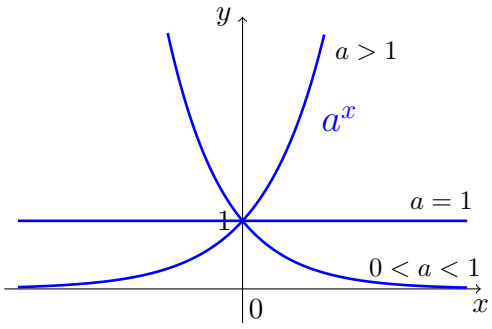
- ***a alapú exponenciális függvény:*** $f(x) := \exp_a(x) = a^x \quad (x \in \mathbb{R})$, ahol $a > 0$,
- ***a alapú logaritmusfüggvény:*** $f(x) := \log_a x \quad \big(x \in (0, +\infty)\big)$, ahol $a > 0$ és $a \neq 1$,
- ***általános hatványfüggvény:*** $f(x) := x^\alpha \quad \big(x \in (0, +\infty)\big)$, ahol $\alpha \in \mathbb{R}$.

Ez utóbbi körben még nem vizsgáltuk a függvények konvexitását, hiszen ehhez a differenciálszámítás eszközeit szerettük volna alkalmazni. Most fogjuk ezt pótolni. A konvexitáshoz a függvény második deriváltjának előjeleit kell meghatározni.

Az $\exp_a$ függvény szigorúan konvex $\mathbb{R}$-n, ha $a > 0$ és $a \neq 1$. Ez abból következik, hogy minden $x \in \mathbb{R}$ esetén

$$(\exp_a(x))'' = (\exp_a(x) \ln a)' = \exp_a(x) \ln^2 a > 0.$$

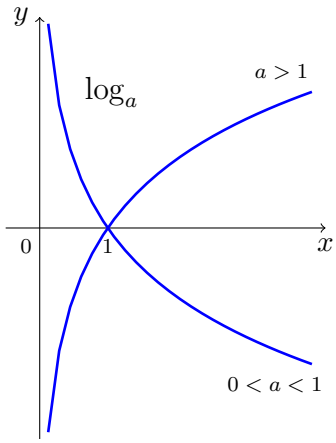
Ha $a = 1$, akkor az $\exp_a \equiv 1$ lineáris függvény egyszerre konvex és konkáv $\mathbb{R}$-n, de nem szigorú értelemben.



A $\log_a$ függvény szigorúan konkáv $(0, +\infty)$-n, ha $a > 1$, és szigorúan konvex $(0, +\infty)$-n, ha $0 < a < 1$, hiszen minden $x > 0$ esetén

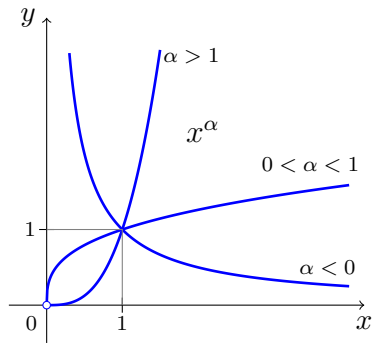
$$(\log_a(x))'' = \left(\frac{1}{x \ln a}\right)' = -\frac{1}{x^2 \ln a} \begin{cases} < 0, \text{ ha } a > 1, \\ > 0, \text{ ha } 0 < a < 1, \end{cases}$$

mivel $\ln a > 0$, ha $a > 1$, illetve $\ln a < 0$, ha $0 < a < 1$.



Az x^α függvény konvexitása is függ az α értéktől. Valóban, minden $x > 0$ esetén

$$(x^\alpha)'' = (\alpha x^{\alpha-1})' = \alpha(\alpha-1)x^{\alpha-2}.$$



Tehát az $\alpha(\alpha-1)$ előjelén múlik a függvény konvexitása. Ha $\alpha > 1$ vagy $\alpha < 0$, akkor $(x^\alpha)'' > 0$, és így az x^α függvény konvex $(0, +\infty)$-n. Ha $0 < \alpha < 1$, akkor $(x^\alpha)'' < 0$, és így az x^α függvény konkáv $(0, +\infty)$-n. Ha $\alpha = 0$ vagy $\alpha = 1$, akkor x^α lineáris függvény, azaz egyszerre konvex és konkáv $(0, +\infty)$-n, de nem szigorú értelemben.

Az Analízis I. kurzuson szintén foglalkoztunk a szinusz- és a koszinuszfüggvénnyel, de nem fejeztük be az elemzésüket. A következő részben ezt fogjuk elvégezni.

Trigonometrikus függvények

A középiskolában már megismerkedtünk tetszőleges $x \in \mathbb{R}$ esetén a $\sin x$, a $\cos x$, valamint alkalmas $x \in \mathbb{R}$ esetén a $\operatorname{tg} x$ és a $\operatorname{ctg} x$ számok szemléletes definícióival, amiket érdemes felidézni és megjegyezni. Ezekből kiindulva értelmeztük a trigonometrikus függvényeket, és megállapítottuk számos érdekes és fontos tulajdonságaikat. A szóban forgó értelmezésekhez a következő megjegyzéseket fűzzük: Egyrészt ezek a definíciók még utalást sem adnak a függvényértékek (akárcsak közelítő) kiszámolására. Másrészt az egyszerű geometriai fogalmakon túl szerepelnek viszonylag bonyolult és definiálatlan fogalmak is, így a valós számoknak a kör kerületére való „felmérése” vagy a körív hossza. A **π számot** az egységsugarú kör kerületének a felével definiáltuk, amelyről megtudtuk, hogy az egy *irracionális szám*, század pontossággal $3,14$.

Az Analízis I. kurzuson a **szinusz-** és a **koszinuszfüggvényt** hatványsor összegfüggvényeként értelmeztük. Egyáltalán nem nyilvánvaló, hogy ezek ekvivalensek a középiskolai definíciókkal. Néhány már megismert összefüggés azonban jelzi a hasonlóságot. A kétféle bevezetés ekvivalenciájának az igazolását majd az integrálszámítás alkalmazásainak a tárgyalásánál fejezzük be, amikor is értelmezzük a körív hosszát, és meghatározzuk a kör kerületét. A hatványsoros definíció alapján kapott szinusz- és koszinuszfüggvény jelölésére a jelzett ekvivalencia miatt használtuk a „szokásos” sin és cos szimbólumokat.

• *A* sin *és a* cos *függvény*

A ***szinusz-*** és a ***koszinuszfüggvénnyt*** az egész $\mathbb{R}$-en konvergens hatványsor összegfüggvényeként értelmeztük:

$$\sin x := \sin(x) := x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots = \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} \qquad (x \in \mathbb{R}),$$

$$\cos x := \cos(x) := 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots = \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!} x^{2k} \qquad (x \in \mathbb{R}).$$

Most bizonyítás nélkül összefoglaljuk azokat az állításokat, amelyeket az Analízis I. kurzuson már megismertünk.

1. ***Paritás:*** a sin függvény páratlan, és a cos függvény páros, azaz

$$\sin(-x) = -\sin x, \quad \cos(-x) = \cos x \qquad (x \in \mathbb{R}).$$

$\boxed{\textbf{2.}}$ ***Addíciós képletek:*** minden $x, y \in \mathbb{R}$ esetén

$$\sin(x+y) = \sin x \cos y + \cos x \sin y,$$
$$\cos(x+y) = \cos x \cos y - \sin x \sin y.$$

$\boxed{\textbf{3.}}$ Érdemes megjegyezni azt a tényt, hogy ***két szinusz, illetve koszinusz összege és különbsége szorzattá alakítható***. A következő azonosságok az addíciós képletek egyszerű következményei. Tetszőleges $x, y \in \mathbb{R}$ esetén

$$\begin{aligned} &\sin x + \sin y = 2 \cdot \sin \frac{x+y}{2} \cdot \cos \frac{x-y}{2}, \qquad && \sin x - \sin y = 2 \cdot \sin \frac{x-y}{2} \cdot \cos \frac{x+y}{2}, \\ &\cos x + \cos y = 2 \cdot \cos \frac{x+y}{2} \cdot \cos \frac{x-y}{2}, \qquad && \cos x - \cos y = -2 \cdot \sin \frac{x+y}{2} \cdot \sin \frac{x-y}{2}. \end{aligned} \tag{1}$$

$\boxed{\textbf{4.}}$ Minden $x \in \mathbb{R}$ esetén

$$\sin(2x) = 2 \sin x \cos x, \qquad \cos(2x) = \cos^2 x - \sin^2 x.$$

$\boxed{\textbf{5.}}$ ***Négyzetes összefüggés:***

$$\sin^2 x + \cos^2 x = 1 \quad (x \in \mathbb{R}).$$

$\boxed{\textbf{6.}}$ ***Folytonosság:*** A sin és a cos függvény folytonos $\mathbb{R}$-en.

$\boxed{\textbf{7.}}$ ***A π szám értelmezése:*** A cos függvénynek a $[0, 2]$ intervallumban pontosan egy zérushelye van, azaz $[0, 2]$-nek pontosan egy ξ pontjában áll fenn a $\cos \xi = 0$ egyenlőség. Ennek a ξ számnak a kétszereseként értelmezzük a π számot:

$$\pi := 2\xi.$$

Megjegyzések.

1. Az előző állítás bizonyításából (Analízis I.-ben) kiderült, hogy $\forall x \in (0, 2)\colon\ \sin x > 0$. Ezt a tulajdonságot fogjuk a továbbiakban is alkalmazni.

2. Az addíciós képletek, valamint a négyzetes összefüggés felhasználásával a sin és a cos függvény számos helyen vett helyettesítési értékeit pontosan ki tudjuk számolni. Például:

   - $\sin(\pi/2) = 1$, mert a négyzetes összefüggés miatt:

     $$\sin^2 \frac{\pi}{2} = 1 - \cos^2 \frac{\pi}{2} = 1 - 0 = 1 \qquad \Longrightarrow \qquad \sin \frac{\pi}{2} = 1 \quad \text{vagy} \quad \sin \frac{\pi}{2} = -1$$

     de $\pi/2 \in (0, 2) \implies \sin(\pi/2) > 0$, azaz $\sin(\pi/2) = 1$.

   - $\sin(\pi/6) = 1/2$. Az addíciós képletek miatt

     $$\begin{aligned} \cos 3x = \cos(2x + x) &= \cos 2x \cos x - \sin 2x \sin x = \\ &= (\cos^2 x - \sin^2 x) \cos x - 2 \sin x \cos x \sin x = \\ &= (1 - \sin^2 x - \sin^2 x) \cos x - 2 \sin^2 x \cos x = (1 - 4 \sin^2 x) \cos x. \end{aligned}$$

     Ha $x = \pi/6$, akkor $\cos 3x = \cos(\pi/2) = 0$ és $\cos x = \cos(\pi/6) \neq 0$, mert cos-nak nincs $\pi/2$-től különböző zérushelye $(0, 2)$-n. Ezért a fenti egyenlőségből következik, hogy $1 - 4 \sin^2(\pi/6) = 0$, és így $\sin(\pi/6) = 1/2$, mert $\sin(\pi/6) > 0$.

$\boxed{\textbf{8.}}$ Az addíciós képletekből a következő kapcsolat adódik:

$$(2)\quad \sin x = \cos\left(x - \frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2} - x\right), \qquad \cos x = \sin\left(\frac{\pi}{2} - x\right) = \sin\left(x + \frac{\pi}{2}\right) \quad (x \in \mathbb{R}).$$

$\boxed{\textbf{9.}}$ ***Periodicitás:*** A sin és a cos függvény 2π szerint periodikus, azaz

$$\sin x = \sin(x + 2k\pi), \qquad \cos x = \cos(x + 2k\pi) \quad (x \in \mathbb{R},\ k \in \mathbb{Z}).$$

és 2π mindegyik függvénynek a legkisebb periódusa.

Az eddig felsorolt tulajdonságokat tanultunk az Analízis I. kurzuson. Ezentúl a differenciálszámítás eszköztárát felhasználva fogjuk a sin és a cos függvény „alaki” tulajdonságait tanulmányozni. A deriváltjukat már ismerjük.

$\boxed{\textbf{10.}}$ ***Differenciálhatóság:*** A sin és a cos függvény differenciálható $\mathbb{R}$-en, és

$$\sin'(x) = \cos(x), \quad \cos'(x) = -\sin(x) \qquad (x \in \mathbb{R}).$$

A periodicitást, valamint a paritást figyelembe véve, a vizsgálatokat elég a $[0, \pi]$ intervallumon elvégezni.

1. Tétel.

1. *$\sin \uparrow [0, \frac{\pi}{2}]$-en, $\downarrow [\frac{\pi}{2}, \pi]$-n és szigorúan konkáv $[0, \pi]$-n.*
2. *$\cos \downarrow [0, \pi]$-n, szigorúan konkáv $[0, \frac{\pi}{2}]$-en és szigorúan konvex $[\frac{\pi}{2}, \pi]$-n.*

Bizonyítás. A (2) azonosságok alapján egyszerűen igazolható a

$$\cos x > 0 \ \left(x \in (0, \tfrac{\pi}{2})\right), \quad \cos x < 0 \ \left(x \in \left(\tfrac{\pi}{2}, \pi\right)\right) \quad \text{és} \quad \sin x > 0 \ \left(x \in (0, \pi)\right)$$

egyenlőtlenségeket, a deriváltakra vonatkozó már ismert

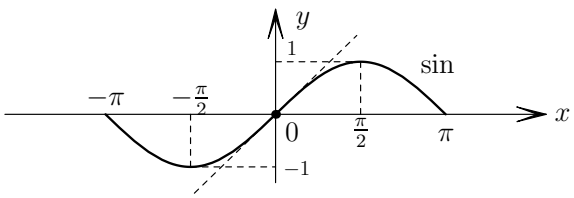
$$\sin' x = \cos x, \quad \cos' x = -\sin x, \quad \sin'' x = -\sin x, \quad \cos'' x = -\cos x \qquad (x \in \mathbb{R})$$

képleteket, továbbá a monotonitás, illetve a konvexitás és a derivált kapcsolatára vonatkozó eredményeket felhasználva kapjuk a tétel állítását.

A sin függvény páratlan, ezért a grafikonja szimmetrikus az origóra. A következő ábrán szemléltetjük a sin függvény grafikonját a $[-\pi, \pi]$ intervallumon, majd felsoroljuk a függvény néhány fontos tulajdonságát:

A* sin *függvény

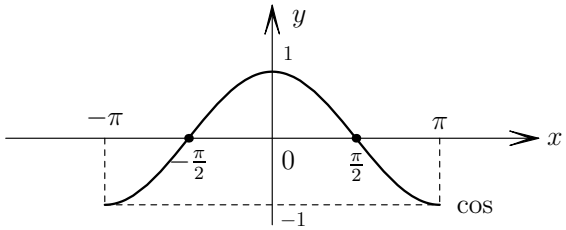
- $\downarrow [-\pi, -\frac{\pi}{2}]$-en, $\uparrow [-\frac{\pi}{2}, \frac{\pi}{2}]$-en és $\downarrow [\frac{\pi}{2}, \pi]$-n;
- szigorúan konvex $[-\pi, 0]$-n és
- 0 inflexiós pont.



A cos függvény páros, ezért a grafikonja szimmetrikus az y-tengelyre. A következő ábrán a cos függvény grafikonját szemléltetjük a $[-\pi, \pi]$ intervallumon, majd felsoroljuk a függvény néhány fontos tulajdonságát:

A* cos *függvény

- $\uparrow$ $[-\pi, 0]$-n és $\downarrow$ $[0, \pi]$-n;
- szigorúan konvex $[-\pi, -\frac{\pi}{2}]$-en,
  szigorúan konkáv $[-\frac{\pi}{2}, \frac{\pi}{2}]$-en,
  szigorúan konvex $[\frac{\pi}{2}, \pi]$-n;
- $\pm\frac{\pi}{2}$ inflexiós pontok.



A periodicitást figyelembe véve az előzőekből már következnek a sin és a cos függvény ***zérushelyeire*** vonatkozó alábbi állítások:

$$\boxed{\begin{array}{lcl} \sin x = 0 & \Longleftrightarrow & x = k\pi \quad (k \in \mathbb{Z}), \\ \cos x = 0 & \Longleftrightarrow & x = \frac{\pi}{2} + k\pi \quad (k \in \mathbb{Z}). \end{array}}$$

Az (1) azonosságokat és a fenti egyenlőségeket felhasználva egyszerűen bebizonyíthatjuk, hogy

$$\boxed{\begin{array}{lcl} \sin x = \sin y & \Longleftrightarrow & x - y = 2k\pi \ \text{ vagy } \ x + y = (2l+1)\pi \quad (k, l \in \mathbb{Z}), \\ \cos x = \cos y & \Longleftrightarrow & x - y = 2k\pi \ \text{ vagy } \ x + y = 2l\pi \quad (k, l \in \mathbb{Z}). \end{array}}$$

• *A* tg *és a* ctg *függvény*

A ***tangensfüggvényt*** így értelmezzük:

$$\operatorname{tg} x := \operatorname{tg}(x) := \frac{\sin x}{\cos x} \qquad \left(x \in \mathbb{R} \setminus \{\tfrac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\}\right).$$

A sin és a cos függvények tulajdonságait felhasználva adódnak a következő állítások:

a) tg *páratlan* függvény, azaz $\operatorname{tg}(-x) = -\operatorname{tg} x \quad (x \in \mathcal{D}_{\operatorname{tg}})$;

b) a tg függvény π *szerint periodikus*, azaz $\operatorname{tg}(x + k\pi) = \operatorname{tg} x \quad (x \in \mathcal{D}_{\operatorname{tg}},\ k \in \mathbb{Z})$;

c) a tg függvény *zérushelyei*:

$$\boxed{\operatorname{tg} x = 0 \quad \Longleftrightarrow \quad \sin x = 0 \quad \Longleftrightarrow \quad x = k\pi \qquad (k \in \mathbb{Z});}$$

d) Mivel

$$\operatorname{tg} x - \operatorname{tg} y = \frac{\sin x}{\cos x} - \frac{\sin y}{\cos y} = \frac{\sin x \cdot \cos y - \cos x \cdot \sin y}{\cos x \cdot \cos y} = \frac{\sin(x-y)}{\cos x \cdot \cos y},$$

ezért

$$\boxed{\operatorname{tg} x = \operatorname{tg} y \quad \Longleftrightarrow \quad \sin(x-y) = 0 \quad \Longleftrightarrow \quad x = y + k\pi \qquad (k \in \mathbb{Z}).}$$

Ezek alapján a tg függvény „alaki tulajdonságait” elég a $(-\pi/2, \pi/2)$ intervallumon megállapítani. Mivel $\operatorname{tg} \in D^\infty(-\pi/2, \pi/2)$, és például

$$\operatorname{tg}' x = \frac{1}{\cos^2 x}, \qquad \operatorname{tg}'' x = 2\frac{\sin x}{\cos^3 x} \qquad \left(x \in \left(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right)\right),$$

ezért a monotonitási, illetve a konvexitási tulajdonságok egyszerűen megállapíthatók.

Meg kell azonban vizsgálnunk a tg függvény $\pm\frac{\pi}{2}$ pont körüli viselkedését. Mivel

$$\lim_{x\to\pm\frac{\pi}{2}} \sin x = \sin\left(\pm\tfrac{\pi}{2}\right) = \pm 1, \quad \text{és}$$

$$\lim_{x\to\pm\frac{\pi}{2}} \cos x = \cos\left(\pm\tfrac{\pi}{2}\right) = 0, \quad \text{továbbá} \quad \cos x > 0, \text{ ha } x \in \left(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right),$$

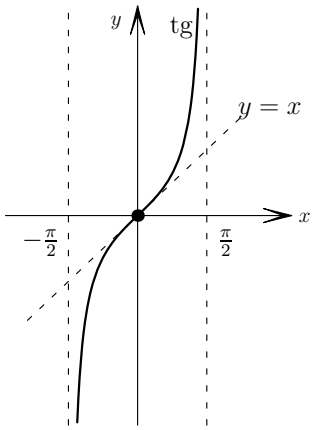
ezért

$$\lim_{x\to-\frac{\pi}{2}+0} \operatorname{tg} x = -\infty \quad \text{és} \quad \lim_{x\to\frac{\pi}{2}-0} \operatorname{tg} x = +\infty\,.$$

A tg függvény grafikonja a $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ intervallumon, és alapvető tulajdonságai:

A* tg *függvény

- páratlan,
- π szerint periodikus,
- $\uparrow$ $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$-en,
- szigorúan konkáv $\left(-\frac{\pi}{2}, 0\right)$-n,
  szigorúan konvex $\left(0, \frac{\pi}{2}\right)$-en,
- 0 inflexiós pont,
- $\lim\limits_{x\to-\frac{\pi}{2}+0} \operatorname{tg} x = -\infty$ és $\lim\limits_{x\to\frac{\pi}{2}-0} \operatorname{tg} x = +\infty$.



A **kotangensfüggvényt** így értelmezzük:

$$\operatorname{ctg} x := \operatorname{ctg}(x) := \frac{\cos x}{\sin x} \qquad (x \in \mathbb{R} \setminus \{k\pi \mid k \in \mathbb{Z}\})\,.$$

A ctg függvény *páratlan* és π *szerint periodikus*, továbbá

$$\boxed{\operatorname{ctg} x = 0 \quad \Longleftrightarrow \quad \cos x = 0 \quad \Longleftrightarrow \quad x = \tfrac{\pi}{2} + k\pi \qquad (k \in \mathbb{Z}),}$$

$$\boxed{\operatorname{ctg} x = \operatorname{ctg} y \quad \Longleftrightarrow \quad \sin(x-y) = 0 \quad \Longleftrightarrow \quad x = y + k\pi \qquad (k \in \mathbb{Z}).}$$

A tg és a ctg függvények között a következő kapcsolat áll fenn:

$$\operatorname{ctg} x = \frac{1}{\operatorname{tg} x}, \qquad \operatorname{ctg} x = -\operatorname{tg}\left(x - \frac{\pi}{2}\right), \qquad \operatorname{tg} x = -\operatorname{ctg}\left(x - \frac{\pi}{2}\right) \tag{3}$$

$$\left(x \in \mathcal{D}_{\operatorname{tg}} \cap \mathcal{D}_{\operatorname{ctg}} = \mathbb{R} \setminus \left\{k\,\tfrac{\pi}{2} \mid k \in \mathbb{Z}\right\}\right).$$

A ctg függvény tulajdonságait elég egy π hosszú intervallumon, mondjuk $(0,\pi)$-n megvizsgálni. Mivel $\operatorname{ctg} \in D^\infty(0,\pi)$, és

$$\operatorname{ctg}' x = -\frac{1}{\sin^2 x}, \qquad \operatorname{ctg}'' x = 2 \cdot \frac{\cos x}{\sin^3 x} \qquad \big(x \in (0,\pi)\big),$$

ezért a monotonitási, illetve a konvexitási tulajdonságok egyszerűen megállapíthatók.

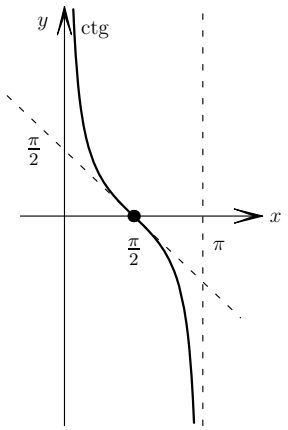
A ctg függvény 0 és π pont körüli viselkedésére a következők teljesülnek:

$$\lim_{x\to 0+0} \operatorname{ctg} x = +\infty, \qquad \lim_{x\to\pi-0} \operatorname{ctg} x = -\infty\,.$$

A ctg függvény grafikonja a $(0,\pi)$ intervallumon, és néhány tulajdonsága:

A* ctg *függvény

- páratlan,
- π szerint periodikus,
- $\downarrow$ $(0,\pi)$-n,
- szigorúan konvex $\left(0,\frac{\pi}{2}\right)$-en,
  szigorúan konkáv $\left(\frac{\pi}{2},\pi\right)$-n,
- $\frac{\pi}{2}$ inflexiós pont,
- $\lim\limits_{x\to 0+0} \operatorname{ctg} x = +\infty$ és $\lim\limits_{x\to \pi-0} \operatorname{ctg} x = -\infty$.



Trigonometrikus függvények inverzei (arkuszfüggvények)

A trigonometrikus függvények mindegyike periodikus, ezért egyikük sem invertálható. Mind a négy függvénynek vannak azonban olyan alkalmas intervallumra vonatkozó leszűkítéseik, amelyeken a függvények szigorúan monotonok, ezért invertálhatók. Ki fogunk választani egy-egy ilyen intervallumot, és ezekre vonatkozó leszűkítéseket invertáljuk. Az így kapott függvényeket ***arkuszfüggvényeknek*** nevezzük. (Az *arcus* szó — ívet jelent — azt jelzi, hogy a függvények helyettesítési értékei bizonyos körív hosszával hozható kapcsolatba.)

1. Definíció. *A szigorúan monoton*

$$\sin_{|\,[-\frac{\pi}{2},\frac{\pi}{2}]}, \quad \cos_{|\,[0,\pi]}, \quad \operatorname{tg}_{|\,(-\frac{\pi}{2},\frac{\pi}{2})}, \quad \operatorname{ctg}_{|\,(0,\pi)}$$

függvények inverzeit rendre **arkuszszinusz-, arkuszkoszinusz, arkusztangens-, arkuszkotangens-függvényeknek** *nevezzük és így jelöljük:*

$$\arcsin := \left(\sin_{|\,[-\frac{\pi}{2},\frac{\pi}{2}]}\right)^{-1}, \qquad \arccos := \left(\cos_{|\,[0,\pi]}\right)^{-1},$$
$$\operatorname{arc\,tg} := \left(\operatorname{tg}_{|\,(-\frac{\pi}{2},\frac{\pi}{2})}\right)^{-1}, \qquad \operatorname{arc\,ctg} := \left(\operatorname{ctg}_{|\,(0,\pi)}\right)^{-1}.$$

Emlékeztetünk arra, hogy ha egy invertálható függvényt és annak inverzét egy koordinátarendszerben ábrázoljuk (feltéve azt, hogy a tengelyeken az egységek egyenlő hosszúak), akkor a szóban forgó függvények grafikonjai egymás tükörképei az $y = x$ egyenletű szögfelező egyenesre vonatkozóan. Következésképpen mindegyik arkuszfüggvény grafikonját a neki megfelelő trigonometrikus függvény (már ismert) grafikonjából kapjuk meg.

Az* arc sin *függvény definíciójából következik, hogy tetszőleges $x \in [-1,1]$ esetén $\arcsin x$ az a $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$ intervallumba eső y szög, amelynek a szinusza x-szel egyenlő, azaz

$$\boxed{\begin{array}{ccccc} \arcsin x & = & y & \Longleftrightarrow & \sin y = x. \\ \left(x \in [-1,1]\right) & & \left(y \in \left[-\frac{\pi}{2},\frac{\pi}{2}\right]\right) & & \end{array}}$$

Az arc sin függvény *folytonos* $[-1,1]$*-en* (lásd az „inverz függvény folytonosságára” vonatkozó tételt), a függvény *deriválhatósága* pedig egyszerűen adódik az inverz függvény deriválási szabályából: Legyen $y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ és $\sin y = x \in [-1,1]$, azaz $y = \arcsin x$. Mivel $\sin' y = \cos y > 0$, ha $y \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, ezért $\arcsin \in D(-1,1)$ és

$$\arcsin' x = \frac{1}{\sin' y} = \frac{1}{\cos y} = \frac{1}{\sqrt{1-\sin^2 y}} = \frac{1}{\sqrt{1-x^2}},$$

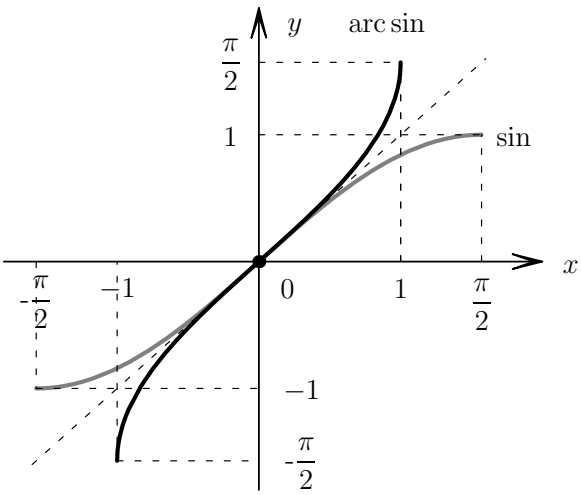
azaz

$$\boxed{\arcsin' x = \frac{1}{\sqrt{1-x^2}} \quad \Big(x \in (-1,1)\Big)}. \tag{4}$$

A következő ábrán szemléltetjük az arc sin függvény grafikonját, és felsoroljuk néhány fontos tulajdonságát:

Az* arc sin *függvény

- $\mathcal{D}_{\arcsin} = [-1,1]$, $\mathcal{R}_{\arcsin} = \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$,
- folytonos $[-1,1]$-en,,
- deriválható $(-1,1)$-en, és
  $$\arcsin' x = \frac{1}{\sqrt{1-x^2}} \quad \Big(x \in (-1,1)\Big)$$
- $\uparrow$ $[-1,1]$-en,
- szigorúan konkáv $[-1,0]$-en,
  szigorúan konvex $[0,1]$-n,
- 0 inflexiós pont.



Az* arc cos *függvény definíciójából következik, hogy

$$\boxed{\begin{array}{ccccc} \arccos x & = & y & \Longleftrightarrow & \cos y = x. \\ \Big(x \in [-1,1]\Big) & & (y \in [0,\pi]) & & \end{array}}$$

Az (2) azonosságok felhasználásával igazolható, hogy

$$\boxed{\arcsin x + \arccos x = \frac{\pi}{2} \quad \Big(x \in [-1,1]\Big)}. \tag{5}$$

A bizonyításához vegyünk egy tetszőleges $x \in [-1,1]$ elemet. Ekkor

$$\arcsin x =: y_1 \in \left[-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right] \iff x = \sin y_1 \iff x = \cos\left(\tfrac{\pi}{2} - y_1\right), \text{ és}$$

$$\arccos x =: y_2 \in [0,\pi] \iff x = \cos y_2.$$

Mivel $\frac{\pi}{2} - y_1 \in [0,\pi]$, és $\cos \downarrow [0,\pi]$-n, így

$$x = \cos\left(\tfrac{\pi}{2} - y_1\right) = \cos y_2 \iff \tfrac{\pi}{2} - y_1 = y_2 \iff y_1 + y_2 = \tfrac{\pi}{2} \iff \arcsin x + \arccos x = \frac{\pi}{2}.$$

Az inverz függvény folytonosságára vonatkozó tétel alapján az arc cos függvény *folytonos* $[-1,1]$-*en*. A (5) és a (4) képletekből pedig az következik, hogy $\arccos \in D(-1,1)$ és

$$\boxed{\arccos' x = -\frac{1}{\sqrt{1-x^2}} \quad \big(x \in (-1,1)\big)}.$$

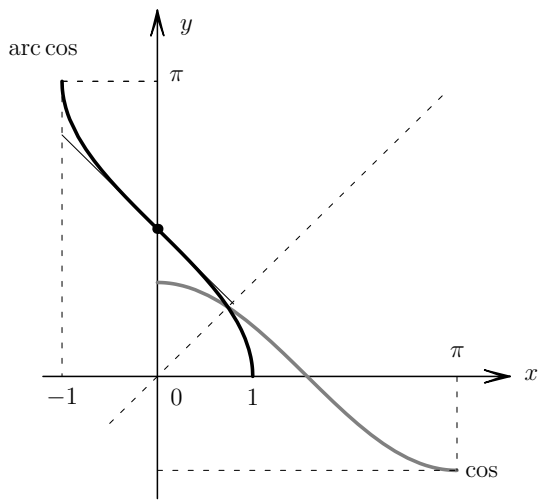
A következő ábrán szemléltetjük az arc cos függvény grafikonját, és felsoroljuk néhány fontos tulajdonságát:

Az* arc cos *függvény

- $\mathcal{D}_{\arccos} = [-1,1]$, $\mathcal{R}_{\arccos} = [0,\pi]$,
- folytonos $[-1,1]$-en,,
- deriválható $(-1,1)$-en, és

  $$\arccos' x = -\frac{1}{\sqrt{1-x^2}} \quad \big(x \in (-1,1)\big)$$

- $\downarrow$ $[-1,1]$-en,
- szigorúan konvex $[-1,0]$-en,
  szigorúan konkáv $[0,1]$-n,
- 0 inflexiós pont.



Az* arc tg *függvény definíciójából és a korábbi eredményeinkből azonnal következik, hogy

$$\mathcal{D}_{\operatorname{arc\,tg}} = \mathbb{R}, \quad \mathcal{R}_{\operatorname{arc\,tg}} = \left(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right), \quad \operatorname{arc\,tg} \uparrow \text{ és folytonos } \mathbb{R}\text{-en},$$

$$\boxed{\begin{array}{lll} \operatorname{arc\,tg} x = y & \Longleftrightarrow & \operatorname{tg} y = x, \\ \big(x \in \mathbb{R}\big) \quad \big(y \in \left(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right)\big) & & \end{array}}$$

továbbá

$$\lim_{x\to-\infty} \operatorname{arc\,tg} x = -\frac{\pi}{2} \quad \text{és} \quad \lim_{x\to+\infty} \operatorname{arc\,tg} x = +\frac{\pi}{2}.$$

Az utóbbi két állítás azt jelenti, hogy az $y = -\frac{\pi}{2}$ (illetve az $y = \frac{\pi}{2}$) egyenletű egyenes az arc tg függvény aszimptotája a $(-\infty)$-ben (illetve a $(+\infty)$-ben).

Mivel minden $y \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ esetén $\operatorname{tg}' y = \frac{1}{\cos^2 y} > 0$, ezért minden $x = \operatorname{tg} y \in \mathbb{R}$ pontban az arc tg függvény deriválható és az inverz függvény deriválási szabálya alapján

$$\operatorname{arc\,tg}' x = \frac{1}{\operatorname{tg}' y} = \cos^2 y = \frac{1}{1+\operatorname{tg}^2 y} = \frac{1}{1+x^2},$$

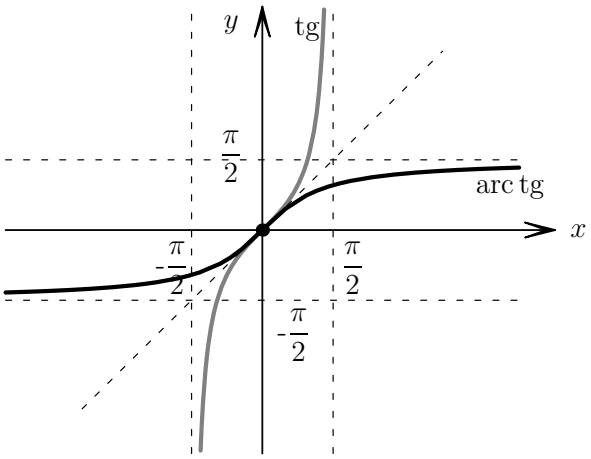
azaz

$$\boxed{\operatorname{arc\,tg}' x = \frac{1}{1+x^2} \quad (x \in \mathbb{R})}.$$

A következő ábrán szemléltetjük az arc tg függvény grafikonját, és felsoroljuk néhány fontos tulajdonságát:

Az* arc tg *függvény

- $\mathcal{D}_{\operatorname{arc\,tg}} = \mathbb{R}$, $\mathcal{R}_{\operatorname{arc\,tg}} = \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$,
- folytonos és deriválható $\mathbb{R}$-en, ill.
$$\operatorname{arc\,tg}' x = \frac{1}{1+x^2} \quad (x \in \mathbb{R})$$
- $\uparrow$ $\mathbb{R}$-en,
- szigorúan konvex $(-\infty, 0]$-n, szigorúan konkáv $[0, +\infty)$-en,
- 0 inflexiós pont,
- $y = \pm\frac{\pi}{2}$ aszimptota a $(\pm\infty)$-ben.



Az* arc ctg *függvény definíciójából és a korábbi eredményeinkből azonnal következik, hogy

$$\mathcal{D}_{\operatorname{arc\,ctg}} = \mathbb{R}, \quad \mathcal{R}_{\operatorname{arc\,ctg}} = (0, \pi), \quad \operatorname{arc\,ctg} \downarrow \text{ és folytonos } \mathbb{R}\text{-en},$$

$$\boxed{\begin{array}{lcl} \operatorname{arc\,ctg} x = y & \Longleftrightarrow & \operatorname{ctg} y = x, \\ \left(x \in \mathbb{R}\right) \quad (y \in (0, \pi)) & & \end{array}}$$

továbbá

$$\lim_{x \to -\infty} \operatorname{arc\,ctg} x = \pi \quad \text{és} \quad \lim_{x \to +\infty} \operatorname{arc\,ctg} x = 0.$$

Az utóbbi két állítás azt jelenti, hogy az $y = \pi$ (illetve az $y = 0$) egyenletű egyenes az arc ctg függvény aszimptotája $(-\infty)$-ben (illetve $(+\infty)$-ben).

A (3) azonosságokból következik, hogy az arc tg és az arc ctg függvény között a következő összefüggés áll fenn:

$$\boxed{\operatorname{arc\,tg} x + \operatorname{arc\,ctg} x = \frac{\pi}{2} \quad (x \in \mathbb{R})},$$

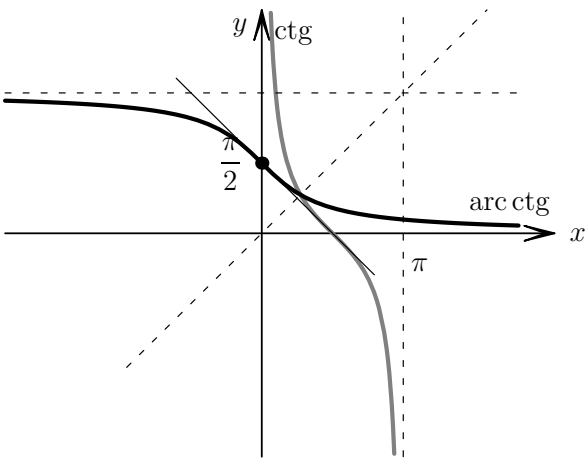
ezért $\operatorname{arc\,ctg} \in D(\mathbb{R})$, és

$$\boxed{\operatorname{arc\,ctg}' x = -\frac{1}{1+x^2} \quad (x \in \mathbb{R})}.$$

Most szemléltetjük az arc ctg függvény grafikonját, és felsoroljuk néhány fontos tulajdonságát:

Az* arc ctg *függvény

- $\mathcal{D}_{\operatorname{arc\,ctg}} = \mathbb{R}$ $\mathcal{R}_{\operatorname{arc\,ctg}} = (0, \pi)$,
- folytonos és deriválható $\mathbb{R}$-en, ill.
$$\operatorname{arc\,ctg}' x = -\frac{1}{1+x^2} \quad (x \in \mathbb{R})$$
- $\downarrow$ $\mathbb{R}$-en,
- szigorúan konkáv $(-\infty, 0]$-n, szigorúan konvex $[0, +\infty)$-en,
- π inflexiós pont,
- $y = \pi$ aszimptota a $(-\infty)$-ben és $y = 0$ aszimptota a $(+\infty)$-ben.



Hiperbolikus függvények és inverzeik

• Hiperbolikus függvények

A ***szinuszhiberbolikusz-*** és a ***koszinuszhiberbolikusz-függvényt*** az egész $\mathbb{R}$-en konvergens hatványsor összegfüggvényeként értelmezzük:

$$\operatorname{sh} x := \operatorname{sh}(x) := x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots = \sum_{k=0}^{+\infty} \frac{x^{2k+1}}{(2k+1)!} \quad (x \in \mathbb{R}),$$

$$\operatorname{ch} x := \operatorname{ch}(x) := 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots = \sum_{k=0}^{+\infty} \frac{x^{2k}}{(2k)!} \quad (x \in \mathbb{R}).$$

Az exp függvény

$$e^x := \exp(x) := 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \sum_{k=0}^{+\infty} \frac{x^k}{k!} \quad (x \in \mathbb{R})$$

definíciójából közvetlenül következnek az alábbi fontos formulák:

$$\operatorname{sh} x = \frac{e^x - e^{-x}}{2}, \qquad \operatorname{ch} x = \frac{e^x + e^{-x}}{2} \qquad (x \in \mathbb{R}). \tag{6}$$

A hiperbolikus függvények sok rokon vonást mutatnak a trigonometrikus függvényekkel, ezekre utalnak az elnevezések. Az (6) formulák, valamint az exp függvény alaptulajdonságai felhasználásával egyszerűen bizonyíthatók a trigonometrikus függvényekhez hasonló alábbi állítások:

1. ***Paritás:*** A sh *páratlan*, a ch pedig *páros* függvény.

2. ***Addíciós képletek:*** minden $x, y \in \mathbb{R}$ esetén

$$\operatorname{sh}(x+y) = \operatorname{sh} x \operatorname{ch} y + \operatorname{ch} x \operatorname{sh} y,$$

$$\operatorname{ch}(x+y) = \operatorname{ch} x \operatorname{ch} y + \operatorname{sh} x \operatorname{sh} y.$$

3. ***Négyzetes összefüggés:***

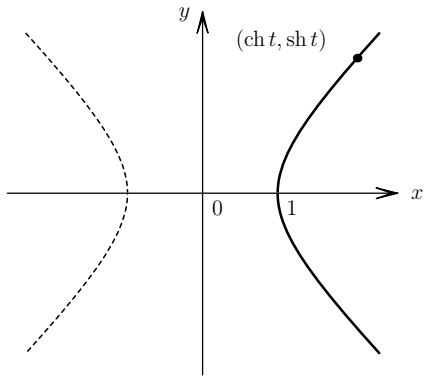
$$\operatorname{ch}^2 x - \operatorname{sh}^2 x = 1 \qquad (x \in \mathbb{R}).$$

Megjegyzés. A négyzetes összefüggés geometriai tartalma a következő:

Minden $t \in \mathbb{R}$ valós szám esetén az

$$(x, y) = (\operatorname{ch} t, \operatorname{sh} t)$$

síkbeli pont rajta van az $x^2 - y^2 = 1$ $(x > 0)$ egyenletű hiperbolaágon (innen származik a szóban forgó függvények nevében szereplő „hiperbolikus” jelző).

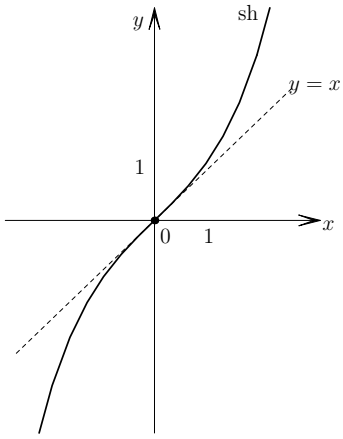


4. ***Differenciálhatóság:*** A sh és a ch függvény differenciálható (tehát folytonos is) $\mathbb{R}$-en, és

$$\operatorname{sh}'(x) = \operatorname{ch}(x), \quad \operatorname{ch}'(x) = \operatorname{sh}(x) \qquad (x \in \mathbb{R}).$$

A differenciálszámítás eszköztárának a felhasználásával vizsgálhatjuk a sh és a ch függvény analitikus és „alaki” tulajdonságait. Most a részletek mellőzésével felsoroljuk ezeknek a függvényeknek a tulajdonságait, majd azok felhasználásával ábrázoljuk a grafikonjaikat.

A sh *függvény*

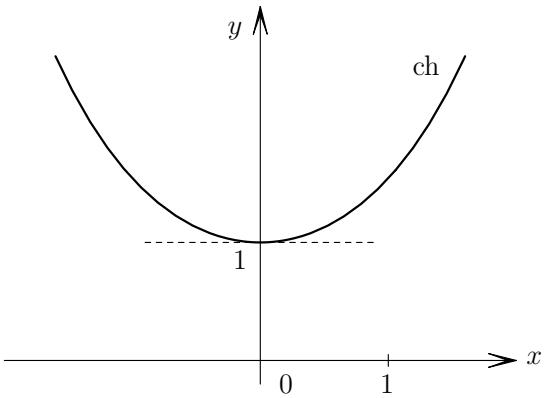


- $\mathcal{D}_{\text{sh}} = \mathbb{R},\ \mathcal{R}_{\text{sh}} = \mathbb{R}$,
- páratlan függvény,
- folytonos és deriválható $\mathbb{R}$-en, ill.

$$\operatorname{sh}' x = \operatorname{ch} x \quad (x \in \mathbb{R})$$

- $\uparrow$ $\mathbb{R}$-en,
- szigorúan konkáv $(-\infty, 0]$-n, szigorúan konvex $[0, +\infty)$-en,
- 0 inflexiós pont.

A ch *függvény*



- $\mathcal{D}_{\text{ch}} = \mathbb{R},\ \mathcal{R}_{\text{ch}} = [1, +\infty)$,
- páros függvény,
- folytonos és deriválható $\mathbb{R}$-en, ill.

$$\operatorname{ch}' x = \operatorname{sh} x \quad (x \in \mathbb{R})$$

- $\downarrow$ $(-\infty, 0)$-en, és $\uparrow$ $(0, +\infty)$-en,
- szigorúan konvex $\mathbb{R}$-en,
- 0 abszolút minimumhely.

Megjegyzés. A ch függvény képét ***láncgörbének*** is nevezik, mert egy homogén, hajlékony, nyúlásmentes, két végén felfüggesztett fonal (lánc) ilyen alakot vesz fel.

A ***tangenshiperbolikusz-*** és a ***kotengenshiperbolikusz-függvényeket*** a tg és a ctg függvények mintájára értelmezzük:

$$\operatorname{th} x := \operatorname{th}(x) := \frac{\operatorname{sh} x}{\operatorname{ch} x}\ \Big(x \in \mathbb{R}\Big), \qquad \operatorname{cth} x := \operatorname{cth}(x) := \frac{\operatorname{ch} x}{\operatorname{sh} x}\ \Big(x \in \mathbb{R} \setminus \{0\}\Big).$$

(A definícióknál figyelembe vettük azt, hogy minden $x \in \mathbb{R}$-re $\operatorname{ch} x \neq 0$, és $\operatorname{sh} x = 0 \iff x = 0$.)

Mindkét függvény páratlan, ezért a tulajdonságaikat elég a $(0, +\infty)$ intervallumon megállapítani. Meg kell még vizsgálni a $\lim\limits_{+\infty} \operatorname{th}$, a $\lim\limits_{+\infty} \operatorname{cth}$ és a $\lim\limits_{0+0} \operatorname{cth}$ határértékeket. Mivel

$$\operatorname{th} x = \frac{\operatorname{sh} x}{\operatorname{ch} x} \overset{(6)}{=} \frac{e^x - e^{-x}}{e^x + e^{-x}} = \frac{1 - e^{-2x}}{1 + e^{-2x}}, \quad \lim_{x \to +\infty} e^{-2x} = 0, \quad \operatorname{cth} x = \frac{1}{\operatorname{th} x} \quad \text{és} \quad \lim_{x \to 0+0} \operatorname{sh} x = 0,$$

ezért

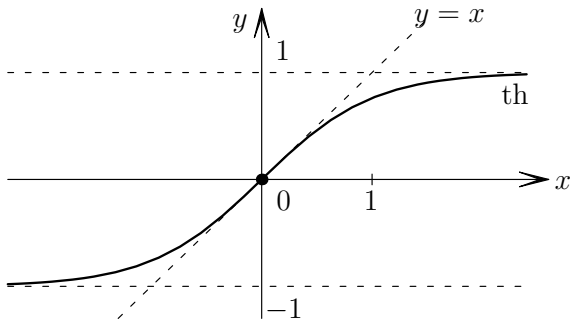
$$\lim_{x \to +\infty} \operatorname{th} x = 1, \quad \lim_{x \to +\infty} \operatorname{cth} x = 1 \quad \text{és} \quad \lim_{x \to 0+0} \operatorname{cth} x = +\infty.$$

A th *függvény*

- $\mathcal{D}_{\text{th}} = \mathbb{R}$, $\mathcal{R}_{\text{th}} = (-1, 1)$,
- páratlan függvény,
- folytonos és deriválható $\mathbb{R}$-en, ill.

$$\text{th}'\, x = \frac{1}{\text{ch}^2\, x} \quad (x \in \mathbb{R})$$

- $\uparrow$ $\mathbb{R}$-en,
- szigorúan konvex $(-\infty, 0]$-n,
  szigorúan konkáv $[0, +\infty)$-en,
- 0 inflexiós pont,
- $y = \pm 1$ aszimptota $(\pm\infty)$-ben.

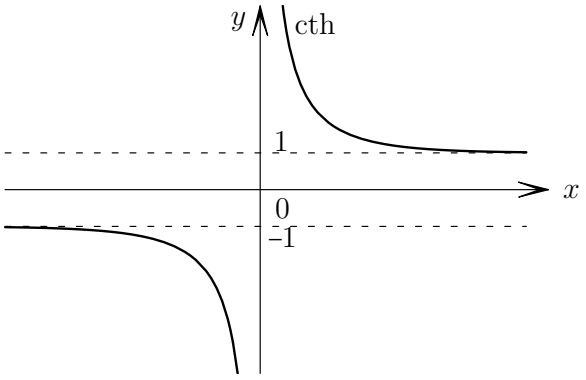


A cth *függvény*

- $\mathcal{D}_{\text{cth}} = \mathbb{R} \setminus \{0\}$, $\mathcal{R}_{\text{cth}} = \mathbb{R} \setminus [-1, 1]$,
- páratlan függvény,
- folytonos és deriválható $\mathbb{R}$-en, ill.

$$\text{cth}'\, x = -\frac{1}{\text{sh}^2\, x} \quad \big(x \in \mathbb{R} \setminus \{0\}\big)$$

- $\downarrow$ $(-\infty, 0)$-en, és $\downarrow$ $(0, +\infty)$-en,,
- szigorúan konkáv $(-\infty, 0]$-n,
  szigorúan konvex $[0, +\infty)$-en,
- $y = \pm 1$ aszimptota $(\pm\infty)$-ben.

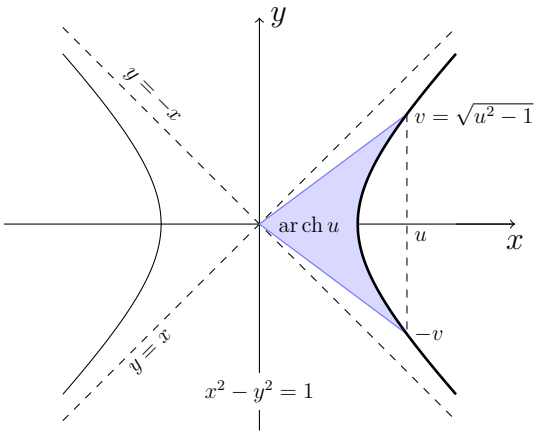


Hiperbolikus függvények inverzei (areafüggvények)

2. Definíció. *A* sh, $\text{ch}_{|\,[0,+\infty)}$, th, cth *függvények invertálhatók. Az inverzeiket rendre* **areaszinuszhiperbolikusz-, areakoszinuszhiperbolikusz-, areatangenshiperbolikusz-, areakotangenshiperbolikusz-függvényeknek** *nevezzük és így jelöljük:*

$$\text{ar sh} := \text{sh}^{-1}, \quad \text{ar ch} := \left(\text{ch}_{|\,[0,+\infty)}\right)^{-1}, \quad \text{ar th} := \text{th}^{-1}, \quad \text{ar cth} := \text{cth}^{-1}.$$

Megjegyzés. Az inverz hiperbolikus függvények nevében megjelenő *area*=terület szó azt jelzi, hogy az $\text{ar ch}\, u$ mennyiség egy bizonyos síkidom területével egyenlő. Pontosabban: Legyen $u \geq 1$ és $v = \sqrt{u^2 - 1}$. Jelölje A_u azt a tartományt, amelyet az origót az (u, v) és $(u, -v)$ pontokkal összekötő két szakasz, valamint az $x^2 - y^2 = 1$ hiperbolának az (u, v) és $(u, -v)$ pontok közötti íve határol. Meg lehet mutatni, hogy az A_u tartomány területe éppen $\text{ar ch}\, u$-val egyenlő. Ezt szemlélteti a következő ábra.

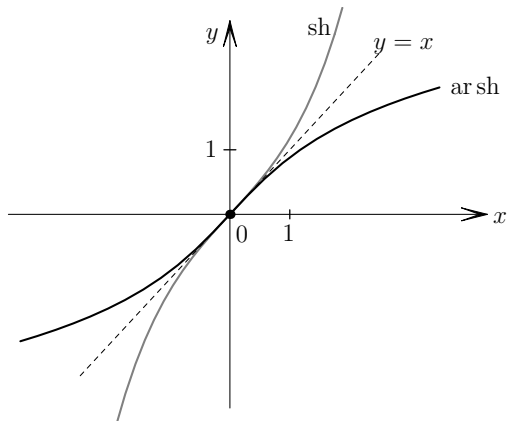


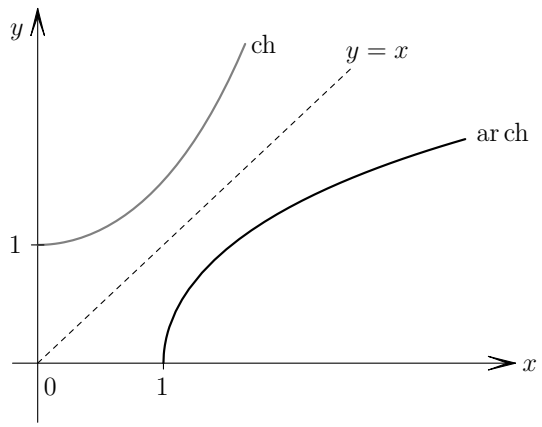
Az inverz függvény deriválási szabályából következik, hogy mindegyik areafüggvény az értelmezési tartományának minden belső pontjában deriválható, és

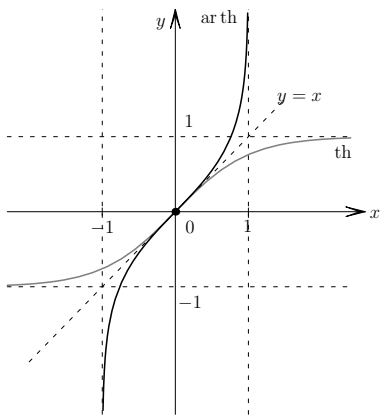
$$\operatorname{ar\,sh}' x = \frac{1}{\sqrt{x^2+1}} \quad (x \in \mathbb{R}), \qquad \operatorname{ar\,ch}' x = \frac{1}{\sqrt{x^2-1}} \quad \big(x \in (1, +\infty)\big),$$

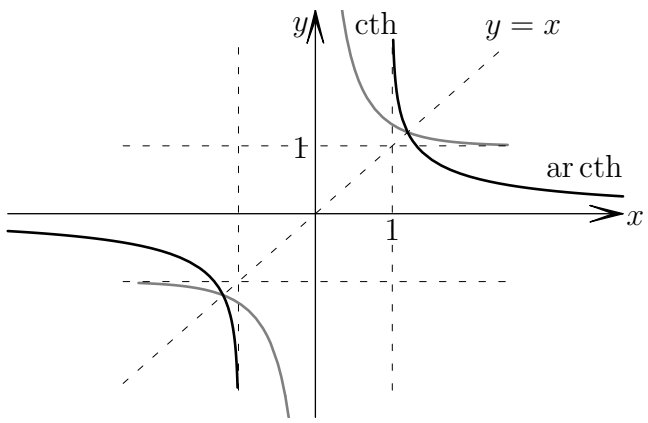
$$\operatorname{ar\,th}' x = \frac{1}{1-x^2} \quad \big(x \in (-1, 1)\big), \qquad \operatorname{ar\,cth}' x = \frac{1}{1-x^2} \quad \big(|x| > 1\big).$$

Az areafüggvények alábbi ábrákon szemléltetett analitikus és geometriai tulajdonságai a korábbiakhoz hasonlóan állapíthatók meg.









A hiperbolikus függvényeket ki lehet fejezni az exp függvénnyel. Az exp függvény inverze az ln függvény, ezért nem meglepő, hogy az areafüggvényeket az ln segítségével is fel tudjuk írni.

2. Tétel. *A következő azonosságok teljesülnek:*

$$\operatorname{ar\,sh} x = \ln\left(x + \sqrt{x^2+1}\right) \qquad (x \in \mathbb{R}),$$

$$\operatorname{ar\,ch} x = \ln\left(x + \sqrt{x^2-1}\right) \qquad (x \in [1, +\infty)),$$

$$\operatorname{ar\,th} x = \tfrac{1}{2} \cdot \ln\left(\frac{1+x}{1-x}\right) \qquad (x \in (-1, 1)),$$

$$\operatorname{ar\,cth} x = \tfrac{1}{2} \cdot \ln\left(\frac{x+1}{x-1}\right) \qquad (|x| > 1).$$

Bizonyítás. A bizonyítások hasonlók, ezért csak az első azonosság igazolását részletezzük. Legyen $x \in \mathbb{R}$ és $y := \operatorname{ar\,sh} x$, azaz $x = \operatorname{sh} y = \dfrac{e^y - e^{-y}}{2}$. Bevezetve a $t := e^y$ jelölést, t-re a $t^2 - 2tx - 1 = 0$ másodfokú egyenletet kapjuk. Mivel $t > 0$, ezért $t = e^y = x + \sqrt{x^2+1}$, azaz

$$y = \operatorname{ar\,sh} x = \ln\left(x + \sqrt{x^2+1}\right) \qquad (x \in \mathbb{R}).$$

Megjegyzés. A fenti képletek jelentősége a következő: Ha az ln függvény helyettesítési értékeit ki tudjuk számolni, akkor az areafüggvények helyettesítési értékei is számolhatók.

Az elemi függvények helyettesítési értékeinek a kiszámolása

Az analízisben (és általában a matematikában) a leggyakrabban előforduló függvények a polinomok, a racionális függvények, az exponenciális-, hatvány- és logaritmusfüggvények, a trigonometrikus függvények, a hiperbolikus függvények és az inverzeik. ***Elemi függvényeknek*** nevezzük azokat a függvényeket, amelyeket a fentiekből kaphatjuk meg a négy alapművelet, a kompozíció és valamely nyílt intervallumra való leszűkítés véges számú alkalmazásával.

Fontos kérdés az elemi függvények helyettesítési értékeinek „tetszőleges pontossággal” való kiszámíthatósága. Polinomok és racionális függvények helyettesítési értékei a négy alapművelet véges sokszori alkalmazásával egyszerűen számolhatók. Hatványsor összegfüggvényei polinomok sorozatának határértékei, ezért azok helyettesítési értékeit általában nem tudjuk pontosan kiszámolni. A közelítő értékeit azonban (elvileg) tetszőleges pontossággal meg tudjuk határozni a négy alapművelet véges sokszori felhasználásával.

Érdekes és fontos tény az, hogy mindegyik elemi függvényt *ki lehet fejezni* néhány „alapfüggvény” segítségével. Pontosabban igaz a következő állítás: *Mindegyik elemi függvény kifejezhető az*

- $\mathbb{R} \ni x \mapsto 1$,
- $\mathbb{R} \ni x \mapsto x$,
- $\mathbb{R} \ni x \mapsto e^x$,
- $\mathbb{R} \ni x \mapsto \sin x$,
- $(0, +\infty) \ni x \mapsto \ln x$,
- $\mathbb{R} \ni x \mapsto \operatorname{arc\,tg} x$

függvényekkel a négy alapművelet, a kompozíció és valamely nyílt intervallumra való leszűkítés véges számú alkalmazásával.

Ez azt is jelenti, hogy az *összes* elemi függvény tetszőleges pontossággal való kiszámolhatóságához elég előállítani csak az utolsó négy függvényt hatványsor összegfüggvényeként. Az exp és a sin függvényekre ilyen előállítást már megismertünk. Az ln és az arc tg függvények alkalmas leszűkítéseit hamarosan elő fogjuk állítani hatványsor összegfüggvényeként.

Most felsoroljuk a fenti állítás bizonyításához alkalmazható formulákat:

- $x^\alpha = e^{\alpha \ln x}$ $\quad \Big(x \in (0, +\infty),\ \alpha \in \mathbb{R}\Big)$
- $a^x = e^{x \ln a}$ $\quad \Big(x \in \mathbb{R},\ a \in (0, +\infty)\Big)$
- $\log_a x = \dfrac{\ln x}{\ln a}$ $\quad \Big(x \in (0, +\infty),\ a \in \mathbb{R}^+ \setminus \{1\}\Big)$
- $\arcsin x = \operatorname{arc\,tg} \dfrac{x}{\sqrt{1 - x^2}}$ $\quad \Big(x \in (-1, 1)\Big)$
- $\cos x = \sin\Big(x + \dfrac{\pi}{2}\Big)$ $\quad \Big(x \in \mathbb{R}\Big)$
- $\arccos x = \dfrac{\pi}{2} - \arcsin x$ $\quad \Big(x \in [-1, 1]\Big)$
- $\operatorname{tg} x = \dfrac{\sin x}{\cos x}$ $\quad \Big(x \in \mathbb{R} \setminus \Big\{\frac{\pi}{2} + k\pi \mid k \in Z\Big\}\Big)$
- $\operatorname{ctg} x = \dfrac{\cos x}{\sin x}$ $\quad \Big(x \in \mathbb{R} \setminus \Big\{k\pi \mid k \in Z\Big\}\Big)$
- $\operatorname{arc\,ctg} x = \dfrac{\pi}{2} - \operatorname{arc\,tg} x$ $\quad \Big(x \in \mathbb{R}\Big)$
- $\operatorname{sh} x = \dfrac{e^x - e^{-x}}{2}$ $\quad \Big(x \in \mathbb{R}\Big)$
- $\operatorname{ar\,sh} x = \ln\Big(x + \sqrt{x^2 + 1}\Big)$ $\quad \Big(x \in \mathbb{R}\Big)$
- $\operatorname{ch} x = \dfrac{e^x + e^{-x}}{2}$ $\quad \Big(x \in \mathbb{R}\Big)$
- $\operatorname{ar\,ch} x = \ln\Big(x + \sqrt{x^2 - 1}\Big)$ $\quad \Big(x \in [1, +\infty)\Big)$
- $\operatorname{th} x = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$ $\quad \Big(x \in \mathbb{R}\Big)$
- $\operatorname{ar\,th} x = \dfrac{1}{2} \cdot \ln\left(\dfrac{1 + x}{1 - x}\right)$ $\quad \Big(x \in (-1, 1)\Big)$
- $\operatorname{cth} x = \dfrac{e^x + e^{-x}}{e^x - e^{-x}}$ $\quad \Big(x \in \mathbb{R} \setminus \{0\}\Big)$
- $\operatorname{ar\,cth} x = \dfrac{1}{2} \cdot \ln\left(\dfrac{1 + x}{x - 1}\right)$ $\quad \Big(x \in (-\infty, -1) \cup (1, +\infty)\Big)$

5. előadás

TAYLOR POLINOMOK ÉS TAYLOR-SOROK

Lineáris közelítés

Gyakori jelenség, hogy valamely problémánál fellépő függvénnyel dolgozva egyszerűbb és áttekinthetőbb eredményhez juthatunk, ha a függvény helyett egy másik, az eredetit „jól közelítő”, de egyszerűbb típusú függvényt tekintünk. Az egyik legegyszerűbb függvénytípus az elsőfokú polinom (vagyis az $mx + b$ $(x \in \mathbb{R})$ függvény). Megmutatjuk, hogy *egy f függvény deriválhatósága az $a \in \operatorname{int} \mathcal{D}_f$ pontban éppen azt jelenti, hogy a függvény bizonyos értelemben „jól közelíthető” elsőfokú polinommal.*

1. Tétel (Lineáris közelítés). *Legyen $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \mathcal{D}_f$. Ekkor*

$$f \in D\{a\} \quad \Longleftrightarrow \quad \begin{cases} \exists A \in \mathbb{R} \ \text{ és } \ \exists \varepsilon : \mathcal{D}_f \to \mathbb{R}, \ \lim\limits_a \varepsilon = 0 : \\ f(x) - f(a) = A(x-a) + \varepsilon(x)(x-a) \quad \big(x \in \mathcal{D}_f\big). \end{cases}$$

Az A szám az f függvény $a \in \mathcal{D}_f$ pontbeli deriváltja, vagyis $A = f'(a)$.

Bizonyítás. $\boxed{\Rightarrow}$

$$f \in D\{a\} \implies \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = f'(a) \in \mathbb{R} \iff \lim_{x \to a} \left(\frac{f(x) - f(a)}{x - a} - f'(a) \right) = 0.$$

Ha $\varepsilon(a) := 0$ és

$$\varepsilon(x) := \frac{f(x) - f(a)}{x - a} - f'(a) \qquad \big(x \in \mathcal{D}_f \setminus \{a\}\big),$$

akkor $\lim\limits_a \varepsilon = 0$ és

$$f(x) - f(a) = f'(a)(x - a) + \varepsilon(x)(x - a) \qquad \big(x \in \mathcal{D}_f\big),$$

ezért a feltétel az $A = f'(a)$ választással teljesül.

$\boxed{\Leftarrow}$ Most tegyük fel, hogy $\exists A \in \mathbb{R}$ és $\exists \varepsilon : \mathcal{D}_f \to \mathbb{R}$, $\lim\limits_a \varepsilon = 0$, hogy

$$f(x) - f(a) = A(x - a) + \varepsilon(x)(x - a) \qquad \big(x \in \mathcal{D}_f\big).$$

Ebből

$$\frac{f(x) - f(a)}{x - a} = A + \varepsilon(x) \ \longrightarrow A + 0 = A, \quad \text{ha} \ \ x \longrightarrow a$$

adódik, ami azt jelenti, hogy $f \in D\{a\}$ és $f'(a) = A$.

Megjegyzés. Az

$$f(x) = f(a) + f'(a)(x-a) + \varepsilon(x)(x-a) \qquad \big(x \in \mathcal{D}_f\big)$$

felírásban a jobb oldal az $f(a)+f'(a)(x-a)$ lineáris függvény és az $\varepsilon(x)(x-a)$ tag összegéből áll, amely tag $\lim\limits_{a} \varepsilon = 0$ és $\lim\limits_{x\to a}(x-a) = 0$ miatt „meglehetősen gyorsan" tart nullához, ha $x \to a$. Az f függvény ***a pontbeli deriválhatósága tehát azt jelenti, hogy a függvény az a pont környezetében „jól" közelíthető lineáris függvénnyel.*** Ezt a közelítést gyakran az

$$f(x) \sim f(a) + f'(a)(x-a) \quad (\text{ha } x \sim a)$$

jelöléssel fejezzük ki. A szóban forgó lineáris függvény grafikonja az

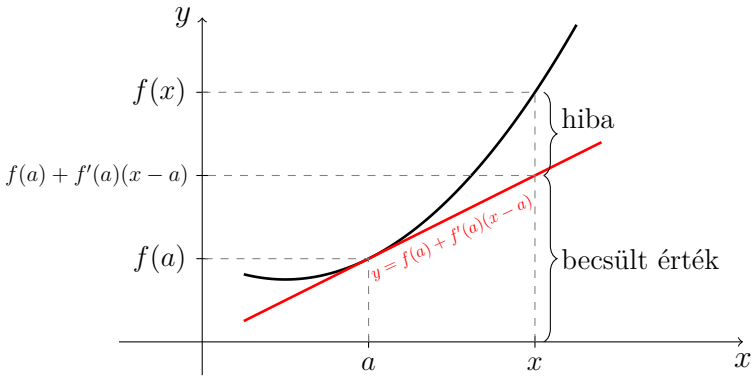
$$y = f(a) + f'(a)(x-a)$$

egyenletű egyenes, ami az f függvény grafikonjának $\big(a, f(a)\big)$ pontbeli *érintője*.

A lineáris közelítés alkalmazható értékbecslésben. Ha ismerjük egy differenciálható f függvény és f' deriváltja pontos értékeit egy „nevezetes" a pontban, akkor az

$$f(a) + f'(a)(x-a)$$

értékkel becsülhetjük az f függvény értékét minden olyan x pontban, amely az a pont „kis sugarú" környezetében van.



1. Feladat. *Milyen lineáris becslést tudunk adni az $f(x) = \sqrt{1+x}$ függvényre az $a = 0$ pont környezetében? Becsüljük vele a $\sqrt{1,1}$ értéket!*

Megoldás. $\mathcal{D}_f = [-1, +\infty)$, $a = 0 \in \operatorname{int} \mathcal{D}_f$, $f \in D(-1, +\infty)$ és

$$f'(x) = \frac{1}{2\sqrt{1+x}}.$$

Mivel $f(0) = 1$ és a $f'(0) = 1/2$, így a keresett becslés

$$\sqrt{1+x} \sim f(a) + f'(a)(x-a) = 1 + \frac{x}{2}.$$

Ha $x = 0,1$, akkor

$$\sqrt{1,1} = \sqrt{1+0,1} \approx 1 + \frac{0,1}{2} = 1,05.$$

Megjegyzés. A kapott $1,05$ eredmény nem olyan rossz közelítés, hiszen a valódi érték első 10 tizedesjegye $1,0488088480$.

Taylor-polinomok

Ha a lineáris közelítés nem elég pontos, akkor magasabb fokszámú polinomokkal is próbálkoztunk. Jobbnak tűnik az

$$(*) \qquad f(x) = P_n(x) + \varepsilon(x)(x-a)^n \qquad \big(x \in \mathcal{D}_f\big),$$

típusú közelítés, ahol $\varepsilon : \mathcal{D}_f \to \mathbb{R}$, $\lim\limits_a \varepsilon = 0$ és P_n egy legfeljebb n-ed fokú polinom, hiszen az a pont közelében $(x-a)^n$ „jobban tart" nullához, mint $x-a$. Ez persze függ az új ε függvénytől is, ezért fogjuk közelebbről megvizsgálni a fenti közelítést. Mindenesetre, a polinomokkal történő közelítés azért jó választás, mert a polinomok értékeinek kiszámításához csak az alapműveletek kellenek.

A vizsgálathoz feltételezzük, hogy az f függvény n-szer differenciálható az a pontban. Ekkor értelmezhető az alábbi polinom.

1. Definíció. *Ha $n \in \mathbb{N}$, $a \in \mathcal{D}_f$ és $f \in D^n\{a\}$, akkor a*

$$T_{n,a}f(x) := f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n =$$
$$= \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k \qquad (x \in \mathbb{R})$$

polinomot az **f függvény a ponthoz tartozó n-edik Taylor-polinomjának** *nevezzük.*

Megjegyzések.

1. Ne felejtsük el, hogy az f^0 jelölés maga az f függvényt jelenti.
2. A Taylor-polinom úgy van ügyesen kialakítva, hogy i-dik deriváltja megegyezzen a függvény i-dik deriváltjával az a pontban minden $i = 0, 1, \ldots, n$ esetén, azaz

   $$(\#) \qquad \big(T_{n,a}f\big)^{(i)}(a) = f^{(i)}(a) \qquad (i = 0, 1, \ldots, n).$$

   Valóban, ha $x = a$, akkor $T_{n,a}f(x)$ mindegyik tagja nulla a nulladik tag ($k = 0$) kivételével, azaz $T_{n,a}f(a) = f^{(0)}(a) = f(a)$. Mivel

   $$\big(T_{n,a}f\big)'(x) = \sum_{k=0}^{n} \frac{k \cdot f^{(k)}(a)}{k!}(x-a)^{k-1} = \sum_{k=1}^{n} \frac{f^{(k)}(a)}{(k-1)!}(x-a)^{k-1} = \sum_{k=0}^{n-1} \frac{f^{(k+1)}(a)}{k!}(x-a)^k,$$

   így ha $x = a$, akkor $\big(T_{n,a}f\big)'(x)$ mindegyik tagja nulla a nulladik tag ($k = 0$) kivételével, azaz $\big(T_{n,a}f\big)'(a) = f^{(0+1)}(a) = f'(a)$. Ha egymás után deriválunk i-szer, akkor

   $$\big(T_{n,a}f\big)^{(i)}(x) = \sum_{k=0}^{n-i} \frac{f^{(k+i)}(a)}{k!}(x-a)^k \qquad (i = 0, 1, \ldots, n).$$

   Ha $x = a$, akkor $\big(T_{n,a}f\big)^{(i)}(x)$ mindegyik tagja nulla a nulladik tag ($k = 0$) kivételével, azaz $\big(T_{n,a}f\big)^{(i)}(a) = f^{(0+i)}(a) = f^{(i)}(a)$.
3. Igazolható, hogy ha $f \in D^n\{a\}$, akkor $T_{n,a}f$ az egyetlen legfeljebb n-edfokú polinom, amire $(*)$ teljesül.

Ahhoz, hogy becsülni tudjuk a Taylor-polinommal történő közelítést, szükségünk lenne a következő állításra.

2. Tétel (Taylor-formula). *Legyen $n \in \mathbb{N}$, és tegyük fel, hogy $f \in D^{n+1}\big(K(a)\big)$. Ekkor*

$$\forall\, x \in K(a) \quad \textit{ponthoz} \quad \exists\, \xi \quad \textit{a és x között}:$$

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

A fenti képlet jobb oldalán álló függvényt **Lagrange-féle maradéktagnak** *nevezzük.*

Bizonyítás. A Cauchy-féle középértéktételt fogjuk felhasználni. Legyen

$$F(x) := f(x) - T_{n,a}f(x) \qquad \big(x \in K(a)\big).$$

(#)-ből következik, hogy

$$F^{(i)}(a) = f^{(i)}(a) - \big(T_{n,a}f\big)^{(i)}(a) = 0 \qquad (i = 0, 1, \ldots, n).$$

Továbbá, $F^{(n+1)}(x) = f^{(n+1)}(x)$, hiszen $\big(T_{n,a}f\big)^{(n+1)} \equiv 0$, mert $T_{n,a}f$ egy legfeljebb n-edfokú polinom.

Másrészt, legyen $G(x) := (x-a)^{n+1} \quad \big(x \in K(a)\big)$. Ekkor minden $x \in K(a)$ esetén

$$G'(x) = (n+1)(x-a)^n, \quad G''(x) = n(n+1)(x-a)^{n-1}, \quad \ldots \quad , G^{(n)}(x) = (n+1)!(x-a),$$

amiből következik, hogy $G^{(i)}(a) = 0$ $(i = 0, 1, \ldots, n)$, és $G^{(n+1)}(x) = (n+1)!$.

Tegyük fel, hogy $x \in K(a)$ és például $x > a$. (Az $x < a$ eset hasonlóan vizsgálható.) Az F és a G függvényekre az $[a, x]$ intervallumon alkalmazható a Cauchy-féle középértéktétel, következésképpen

$$\exists\, \xi_1 \in (a, x)\colon \frac{f(x) - T_{n,a}f(x)}{(x-a)^{n+1}} = \frac{F(x)}{G(x)} = \frac{F(x) - F(a)}{G(x) - G(a)} = \frac{F'(\xi_1)}{G'(\xi_1)}.$$

A Cauchy-féle középértéktételt most az F' és a G' függvényekre az $[a, \xi_1]$ intervallumon alkalmazzuk:

$$\exists\, \xi_2 \in (a, \xi_1) \subseteq (a, x)\colon \frac{F'(\xi_1)}{G'(\xi_1)} = \frac{F'(\xi_1) - F'(a)}{G'(\xi_1) - G'(a)} = \frac{F''(\xi_2)}{G''(\xi_2)}.$$

A fenti gondolatmenetet n-szer megismételve adódik, hogy

$$\exists\, \xi_{n+1} \in (a, \xi_n) \subseteq (a, x)\colon \frac{F^{(n)}(\xi_n)}{G^{(n)}(\xi_n)} = \frac{F^{(n)}(\xi_n) - F^{(n)}(a)}{G^{(n)}(\xi_n) - G^{(n)}(a)} = \frac{F^{(n+1)}(\xi_{n+1})}{G^{(n+1)}(\xi_{n+1})}.$$

A bizonyítás során kapott egyenlőségeket egybevetve azt kapjuk, hogy

$$\frac{f(x) - T_{a,n}(f, x)}{(x-a)^{n+1}} = \frac{F(x)}{G(x)} = \frac{F'(\xi_1)}{G'(\xi_1)} = \cdots = \frac{F^{(n)}(\xi_n)}{G^{(n)}(\xi_n)} = \frac{F^{(n+1)}(\xi_{n+1})}{G^{(n+1)}(\xi_{n+1})} = \frac{f^{(n+1)}(\xi_{n+1})}{(n+1)!},$$

hiszen minden $x \in K(a)$ esetén $F^{(n+1)}(x) = f^{(n+1)}(x)$ és $G^{(n+1)}(x) = (n+1)!$. A konstrukcióból látható, hogy ξ_{n+1} az a pont és x között van, ezért a $\xi := \xi_{n+1}$ választással a bizonyítandó állítást kapjuk.

2. Feladat. *Írjuk fel az*

$$f(x) := \sqrt{1+x} \qquad (x > -1)$$

függvény 0 pont körüli másodfokú Taylor-polinomját! Becsüljük vele a $\sqrt{1,1}$ *értéket és a közelítés hibáját!*

Megoldás. A másodfokú Taylor-polinom

$$T_{2,a}f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 \qquad (x \in \mathbb{R}),$$

ahol $a = 0$. Minden $x > -1$ esetén

$$f(x) = \sqrt{1+x}, \qquad f(0) = 1, \qquad f(a) = 1,$$

$$f'(x) = \frac{1}{2\sqrt{1+x}}, \qquad f'(0) = \frac{1}{2}, \qquad f'(a) = \frac{1}{2},$$

$$f''(x) = -\frac{1}{4\sqrt{(1+x)^3}}, \qquad f''(0) = -\frac{1}{4}, \qquad \frac{f''(a)}{2} = -\frac{1}{8}.$$

Ebből

$$T_{2,0}f(x) = 1 + \frac{1}{2}x - \frac{1}{8}x^2 \qquad (x \in \mathbb{R}).$$

A keresett érték becslése:

$$\sqrt{1,1} = f(0,1) \approx T_{2,0}f(0,1) = 1 + \frac{1}{2}\cdot 0,1 - \frac{1}{8}\cdot(0,1)^2 = 1,04875.$$

A hibabecsléshez legyen $x = 0,1$. A Taylor-formula szerint $\exists\ 0 < \xi < 0,1$, hogy

$$f(x) - T_{2,0}f(x) = \frac{f'''(\xi)}{3!}x^3.$$

Mivel

$$f'''(x) = \frac{3}{8\sqrt{(1+x)^5}} \qquad (x > -1),$$

így

$$|f(x) - T_{2,0}f(x)| = \frac{|f'''(\xi)|}{3!}|x|^3 = \frac{1}{6}\cdot\frac{3}{8\sqrt{(1+\xi)^5}}\cdot(0,1)^3 <$$

$$< \frac{1}{6}\cdot\frac{3}{8\sqrt{(1+0)^5}}\cdot(0,1)^3 = \frac{1}{6}\cdot\frac{3}{8}\cdot\frac{1}{1000} = 0,000375$$

Megjegyzés. A $\sqrt{1,1}$ érték becslésében alkalmazott másodfokú közelítés jobb, mint a lineáris közelítés. Felmerül a kérdés, hogy az n érték növelésével az n-edik Taylor-polinommal egyre kisebb hibakorlátot kapunk-e. Ez sajnos nem minden esetben igaz, ahogy az sem, hogy minden rögzített a pont esetén a Lagrange-féle maradéktag tart nullához a $K(a)$ környezet minden pontjában, ha $n \to +\infty$.

Taylor-sorok

A $\sum\limits_{k=0} \alpha_k(x-a)^k$ hatványsor konvergenciahalmazán értelmezett

$$f(x) := \sum_{k=0}^{+\infty} \alpha_k(x-a)^k$$

összegfüggvénynek nagyon jó tulajdonságai vannak. Ez a függvény többek között folytonos a konvergenciahalmaz minden pontjában, és tagonként differenciálható a belsejében. Helyettesítési értékeit (elvben) tetszőleges pontossággal meg lehet határozni csupán a négy alapművelet felhasználásával. Nem véletlen, hogy az exp, a sin, a cos, a sh, valamint a ch függvényeket hatványsorok összegfüggvényként értelmeztük. Ezért (is) fontos a következő kérdésfelvetés.

Probléma. *Egy adott függvényt vajon elő lehet-e állítani hatványsor összegfüggvényeként? Ha igen, akkor a függvény ismeretében hogyan lehet az együtthatókat meghatározni?*

A hatványsor összegfüggvényének deriváltjára vonatkozó tétel azt mondja ki, hogy ha a hatványsor R konvergenciasugara pozitív, akkor összegfüggvénye differenciálható minden $x \in K_R(a)$ pontban, és

$$f'(x) = \sum_{k=1}^{+\infty} k\alpha_k(x-a)^{k-1} \qquad \Big(x \in K_R(a)\Big).$$

Így f' egy újabb hatványsor összegfüggvénye, ami a fenti tétel alapján szintén differenciálható $K_R(a)$-n, vagyis $f \in D^2\Big(K_R(a)\Big)$ és

$$f''(x) = \sum_{k=2}^{+\infty} k(k-1)\alpha_k(x-a)^{k-2} \qquad \Big(x \in K_R(a)\Big).$$

Világos, hogy f''-re mindaz elmondható, ami f'-re. Ebben az esetben $f \in D^3\Big(K_R(a)\Big)$ és

$$f'''(x) = \sum_{k=3}^{+\infty} k(k-1)(k-2)\alpha_k(x-a)^{k-3} \qquad \Big(x \in K_R(a)\Big).$$

Ezt a gondolatmenetet folytatva azt kapjuk, hogy minden $n = 1, 2, \ldots$ esetén $f \in D^n\Big(K_R(a)\Big)$ és

$$(\star) \qquad f^{(n)}(x) = \sum_{k=n}^{+\infty} k(k-1)(k-2)\cdots(k-n+1)\alpha_k(x-a)^{k-n} \qquad \Big(x \in K_R(a)\Big).$$

Azt a tényt, hogy az f függvény n-szer deriválható minden $n \in \mathbb{N}$ esetén úgy fejeztünk ki, hogy f végtelen sokszor (vagy akárhányszor) deriválható. Ennek jelölésére az $f \in D^\infty$ szimbólumot vezettük be.

Ha $(\star)$-ban $x = a$, akkor a sor minden tagja nulla a $k = n$-re vonatkozó tag kivételivel. Ebből következik, hogy $f^{(n)}(a) = n! \cdot \alpha_n$, azaz

$$\boxed{\alpha_n = \frac{f^{(n)}(a)}{n!} \qquad (n \in \mathbb{N}).}$$

Ezzel sikerült megoldanunk a kérdésfelvetés második részét, azaz az összegfüggvény ismeretében meghatározni a hatványsor együtthatóit. De ilyen együtthatókat már láttuk ezelőtt, ezek éppen a Taylor-polinom együtthatói. Olyan sorokkal van tehát dolgunk, amelynek részletösszegei a függvény Taylor-polinomjai.

2. Definíció. *Ha $a \in \mathcal{D}_f$ és $f \in D^\infty\{a\}$, akkor a*

$$T_a f(x) := f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(k)}(a)}{k!}(x-a)^k + \cdots =$$
$$= \sum_{k=0} \frac{f^{(k)}(a)}{k!}(x-a)^k \qquad (x \in \mathbb{R})$$

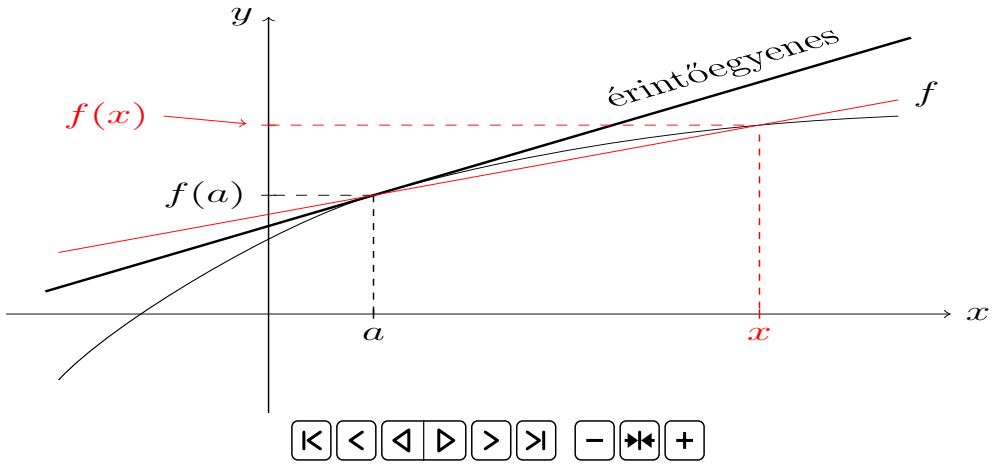
hatványsort az ***f* függvény *a* ponthoz tartozó Taylor-sorának** *nevezzük.*

Az $a = 0$ esetben használatos a **Maclaurin-sor** *elnevezés is.*

Megjegyzés. Az exp, sin, cos, sh, ch függvények definícióiban megadott hatványsorok a szóban forgó függvények $a = 0$ ponthoz tartozó Taylor-sorai. Az animációban példaként látjuk, hogy a szinusz függvény 0 ponthoz tartozó Taylor-polinomjai milyen közel kerülnek a Taylor-sorhoz, azaz az $f(x) := \sin x \quad (x \in \mathbb{R})$ függvényhez a

$$\sin x = T_0 f(x) = \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = \lim_{n\to\infty} \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = \lim_{n\to\infty} T_{n,0} f(x) \qquad (x \in \mathbb{R})$$

határérték értelmében.



A kérdésfelvetés második részére a következő választ adjuk.

3. Tétel. *Minden pozitív konvergenciasugárral rendelkező hatványsor összegfüggvénye a Taylor-sorával egyenlő a konvergenciahalmaz belsejében.*

Bizonyítás. Azonnal következik a fenti fogalmakból és eredményekből.

Azt mondjuk, hogy egy függvény ***hatványsorba fejthető*** egy intervallumon, ha megegyezik egy hatványsor összegfüggvényével ezen az intervallumon. A tétel tehát azt is állítja, hogy ha f hatványsorba fejthető egy nyílt intervallumon, akkor a szóban forgó hatványsor szükségképpen az f függvény Taylor-sora. Ez azt jelenti, hogy a kérdésfelvetés első részének megválaszolásához vizsgálni kell a Taylor-sorok tulajdonságait.

A sorfejtés problémája. Legyen $f \in D^\infty\{a\}$ egy adott függvény.

1. ***A konvergencia problémája:*** Konvergens-e a $T_a f$ Taylor-sor az a pont egyik környezetébe?

2. ***Az előállítás problémája:*** Ha a Taylor-sor konvergens egy $a \in I \subseteq \mathbb{R}$ nyílt intervallumon, akkor vajon fennáll-e az

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

egyenlőség? Ha ez igaz, akkor azt mondjuk, hogy a Taylor-sor előállítja az f függvényt az I intervallumon.

Sajnálatos módon a sorfejtés problémájára nem adhatunk pozitív válasz minden akárhányszor deriválható függvény esetében.

- Bonyolult konstrukcióval lehet példát adni olyan akárhányszor deriválható f függvényre, amelynek *bármely* a ponthoz tartozó Taylor-sora az a ponton kívül sehol sem konvergens.

- Van olyan függvény, amelynek Taylor-sora konvergenciasugara végtelen, de csak egy pontban (a középpontban) állítja elő a függvényt. Legyen

$$f(x) := \begin{cases} e^{-\frac{1}{x^2}}, & \text{ha } x \in \mathbb{R} \setminus \{0\}, \\ 0, & \text{ha } x = 0. \end{cases}$$

Bebizonyítható, hogy $f \in D^\infty(\mathbb{R})$ és $f^{(n)}(0) = 0 \quad (n \in \mathbb{N})$. Ebből következik, hogy f Taylor-sorának minden együtthatója 0, ezért az mindenütt konvergens, és az összegfüggvénye az azonosan 0 függvény, ami nyilván csak a 0 pontban egyenlő f-fel.

Ettől függetlenül kereshetünk olyan függvényosztályokat, ahol pozitív válasz tudunk adni a sorfejtés problémájára. A következő tételben adunk erre egy jól kezelhető feltételt.

4. Tétel (Elégséges feltétel függvények Taylor-sorral történő előállítására). *Legyen $f \in D^\infty\big(K(a)\big)$, és tegyük fel, hogy $\exists M > 0$ valós szám, amire*

$$\forall x \in K(a),\ \forall n \in \mathbb{N}\colon\ \left|f^{(n)}(x)\right| \le M$$

teljesül. Ekkor f-nek az a ponthoz tartozó Taylor-sora a $K(a)$ halmazon előállítja az f függvényt, vagyis fennáll a következő egyenlőség

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

Bizonyítás. Legyen $x \in K(a)$ egy tetszőleges pont. Azt kell igazolni, hogy

$$f(x) = \lim_{n\to+\infty} T_{n,a}(x) \qquad \big(x \in K(a)\big).$$

Ez a Taylor-formula szerint igaz, hiszen létezik olyan ξ pont a és x között, hogy

$$\left|f(x) - T_{n,a}(x)\right| = \left|\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}\right| \le M \cdot \frac{|x-a|^{n+1}}{(n+1)!} \to 0, \qquad \text{ha } n \to +\infty.$$

Megjegyzések.

1. Az Analízis I. kurzuson tanultuk a
$$\lim_{n\to+\infty} \frac{a^n}{n!} = 0 \qquad (a \in \mathbb{R})$$
nevezetes határértéket. Ebből következik a tétel bizonyításában alkalmazott
$$\lim_{n\to+\infty} \frac{|x-a|^{n+1}}{(n+1)!} = 0$$
határérték.

2. A sin és a cos függvény teljesíti az előző tétel feltételeit minden $a \in \mathbb{R}$, $K(a) = \mathbb{R}$ és $M = 1$ esetén, hiszen
$$|\sin x| \le 1, \qquad |\cos x| \le 1 \qquad (x \in \mathbb{R}).$$
Ezért is tudjuk mindkét függvényt előállítani egy tetszőleges a ponthoz tartozó Taylor-sorral. Pl. ha $a = \pi/2$, akkor
$$\sin\Big(\frac{\pi}{2}\Big) = 1, \quad \sin'\Big(\frac{\pi}{2}\Big) = 0, \quad \sin''\Big(\frac{\pi}{2}\Big) = -1, \quad \sin'''\Big(\frac{\pi}{2}\Big) = 0, \quad \dots,$$
és így minden $x \in \mathbb{R}$ esetén
$$\sin(x) = 1 - \frac{1}{2!}\Big(x - \frac{\pi}{2!}\Big)^2 + \frac{1}{4!}\Big(x - \frac{\pi}{2}\Big)^4 - \frac{1}{6!}\Big(x - \frac{\pi}{2}\Big)^6 + \dots = \sum_{k=0}^{+\infty} \frac{(-1)^k}{(2k)!}\Big(x - \frac{\pi}{2}\Big)^{2k}.$$
Ez az előállítás nem annyira meglepő, mert rögtön következik a
$$\sin(x) = \cos\Big(x - \frac{\pi}{2}\Big) \qquad (x \in \mathbb{R})$$
azonosságból, és így a 3. Tételből következik, hogy ez a sin függvény $\pi/2$ ponthoz tartozó Taylor-sora.

3. Az exp függvény nem korlátos $\mathbb{R}$-n, ezért nem teljesíti az előző tétel feltételeit $K(a) = \mathbb{R}$-en. Azonban minden $a \in \mathbb{R}$ és $R > 0$ valós szám esetén az exp függvény korlátos a $K_R(a) = (a - R, a + R)$ intervallumon. Így $\exists M > 0$ valós szám, amire
$$\forall x \in K_R(a),\ \forall n \in \mathbb{N}\colon\ \Big|\exp^{(n)}(x)\Big| = \Big|\exp(x)\Big| \le M$$
teljesül. Tehát exp teljesíti az előző tétel feltételeit, és így előállítható egy tetszőleges a ponthoz tartozó Taylor-sorral minden $(a - R, a + R)$ intervallumon. Ez azt jelenti, hogy előállítható ezzel a Taylor-sorral a teljes való számok halmazán.

Egy f függvény a-hoz tartozó Taylor-sorának a felírásához kétféle módon járhatunk el.

a) A definíció értelmében meghatározzuk a függvény *összes* magasabb rendű deriváltját az a pontban, azaz minden $n \in \mathbb{N}$ számra az $f^{(n)}(a)$ függvényértékeket. Ezek meghatározása általában nem egyszerű feladat, de néhány esetben ki lehet következtetni az első néhány deriválás után. Ezzel fel tudjuk írni a keresett Taylor-sort, de még meg kell határozni a konvergenciahalmazát, illetve megvizsgálni milyen pontokban állítja elő a függvényt.

b) Egy ismert függvény hatványsoros előállításából kiindulva helyettesítéssel, tagonkénti deriválással vagy egyéb módon megkapjuk a függvény hatványsorba fejtését egy $K(a)$ környezetben, ami a 3. Tétel szerint a függvény Taylor-sora lesz. Ezzel rögtön kapunk egy konvergenciahalmazt, ahol a Taylor-sor előállítja a függvényt.

Hasonlítsuk össze a két módszert egyazon feladat megoldásában.

3. Feladat. *Legyen*

$$f(x) := \frac{1}{1+x} \qquad \Big(x \in (-1, +\infty)\Big).$$

Állítsuk elő az f függvény $a = 0$ ponthoz tartozó Taylor-sorát, és vizsgáljuk meg az előállítás problémáját!

Megoldás.

a) Világos, hogy $f \in D^\infty\Big(-1, +\infty\Big)$, és

$$f(x) = (1+x)^{-1}, \quad f'(x) = -(1+x)^{-2}, \quad f''(x) = 2(1+x)^{-3}, \quad f'''(x) = -6(1+x)^{-4}, \dots$$

Sejthető, hogy minden $n = 0, 1, 2, \dots$ esetén

$$f^{(n)}(x) = (-1)^n n!(1+x)^{-n-1} \quad (x > -1), \qquad \text{és így} \qquad f^{(n)}(0) = (-1)^n n!,$$

ami teljes indukcióval könnyen igazolható. A keresett Taylor-sor tehát

$$T_0(f, x) = \sum_{n=0} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0} (-1)^n x^n = 1 - x + x^2 - x^3 + x^4 - \cdots,$$

ami konvergens a $(-1, 1)$ intervallumon (az intervallum határait nem vesszük figyelembe), hiszen a Cauchy–Hadamard-tétel szerint konvergenciasugara

$$R = \frac{1}{\lim\limits_{n\to+\infty} \sqrt[n]{|(-1)^n|}} = 1.$$

Az előállítás problémájához alkalmazzuk a Taylor-formulát. Minden $|x| < 1$ számhoz létezik olyan ξ pont 0 és x között, hogy

$$\Big|f(x) - T_{n,0}(x)\Big| = \left|\frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1}\right| = \left|\frac{(-1)^{n+1}(n+1)!(1+\xi)^{-n-2}}{(n+1)!} x^{n+1}\right| = \frac{|x|^{n+1}}{(1+\xi)^{n+2}}.$$

Ekkor a mértani sorozat határértéke miatt ($q^n \to 0$, ha $|q| < 1$):

- ha $0 < \xi < x < 1$, akkor $\dfrac{1}{1+\xi} < 1$. Ezért

$$|f(x) - T_{n,0}(x)| < x^{n+1} \to 0,$$

- ha $-\dfrac{1}{2} < x < \xi < 0$, akkor $\dfrac{|x|}{1+\xi} < \dfrac{|x|}{1+x} = \dfrac{|x|}{1-|x|} < 1$. Ezért

$$|f(x) - T_{n,0}(x)| = \frac{1}{|x|} \cdot \left(\frac{|x|}{1+\xi}\right)^{n+2} < \frac{1}{|x|} \cdot \left(\frac{|x|}{1-|x|}\right)^{n+2} \to 0.$$

Ezért a Taylor-sor előállítja a függvényt a $(-1/2, 1)$ intervallumon.

Vegyük észre, hogy a Taylor-formula segítségével nem tudjuk igazolni az előállítást a teljes $(-1, 1)$ intervallumon.

b) Ismerjük a mértani sor összegfüggvényét:

$$\frac{1}{1-x} = \sum_{n=0}^{+\infty} x^n \qquad (-1 < x < 1)$$

Ha ebben x helyett $-x$-et írunk, akkor

$$f(x) = \frac{1}{1+x} = \frac{1}{1-(-x)} = \sum_{n=0}^{+\infty} (-x)^n = \sum_{n=0}^{+\infty} (-1)^n x^n \qquad (\underbrace{-1 < -x < 1}_{-1<x<1})$$

Ezzel sikerült az f függvényt hatványsorba fejteni. Ekkor a 3. Tétel szerint ez a hatványsor a függvény 0 ponthoz tartozó Taylor-sora, ami előállítja a függvényt a $(-1, 1)$ intervallumon.

Megjegyzés. A helyettesítéses módszert már alkalmaztuk az Analízis I. kurzuson hatványsorok előállítására. Ehhez most hozzájön a hatványsor összegfüggvényének deriváltjára vonatkozó tétel, amivel nagyon érdekes eredményekhez juthatunk.

4. Feladat. *Legyen*

$$f(x) := \ln(1+x) \qquad \Big(x \in (-1, +\infty)\Big).$$

Állítsuk elő az f függvény $a = 0$ ponthoz tartozó Taylor-sorát, és vizsgáljuk meg az előállítás problémáját!

Megoldás. Először vegyük észre, hogy $f \in D(-1, 1)$ és

$$f'(x) = \frac{1}{1+x} = \sum_{n=0}^{+\infty} (-1)^n x^n = 1 - x + x^2 - x^3 + \cdots \qquad \Big(x \in (-1, 1)\Big).$$

Legyen

$$g(x) := \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} x^n = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots \qquad \Big(x \in (-1, 1]\Big).$$

A fenti hatványsor konvergenciahalmaza valóban a $(-1, 1]$ intervallum, hiszen a Cauchy–Hadamard-tétel szerint konvergenciasugara

$$R = \frac{1}{\lim\limits_{n\to+\infty} \sqrt[n]{|(-1)^{n+1}/n|}} = \lim_{n\to+\infty} \sqrt[n]{n} = 1,$$

az $x = -1$ pontban divergens (harmonikus sor), és az $x = 1$ pontban konvergens (Leibniz-sor). A hatványsorok összegfüggvényének deriváltjára vonatkozó tétel szerint $g \in D(-1, 1)$ és

$$g'(x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} n x^{n-1} = \sum_{n=1}^{+\infty} (-1)^{n-1} x^{n-1} = \sum_{n=0}^{+\infty} (-1)^n x^n \qquad \Big(x \in (-1, 1)\Big).$$

Ezért $f'(x) = g'(x)$ minden $x \in (-1,1)$ pontban. A deriváltak egyenlőségére vonatkozó tételből következik, hogy van olyan $c \in \mathbb{R}$ állandó, hogy

$$f(x) - g(x) = c \quad \Big(x \in (-1,1)\Big).$$

Ugyanakkor $f(0) - g(0) = 0$, így $c = 0$. Bebizonyítottuk tehát azt, hogy

$$f(x) = \ln(1+x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} x^n \qquad \text{ha} \quad x \in (-1,1).$$

Ezzel sikerült az f függvényt hatványsorba fejteni. Ekkor a 3. Tétel szerint ez a hatványsor a függvény 0 ponthoz tartozó Taylor-sora, ami előállítja a függvényt a $(-1,1)$ intervallumon.

Megjegyzés. Az előző feladat megoldásában szereplő

$$f(x) = \ln(1+x) \quad (-1 < x < +\infty) \quad \text{és} \quad g(x) = \sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} x^n \quad (-1 < x \le 1)$$

függvények folytonosak az $x = 1$ pontban. A g függvény folytonossága abból következik, hogy minden hatványsor folytonos a konvergenciahalmazának bármely pontjában. Mivel $f(x) = g(x)$ minden $x \in (-1,1)$ esetén, így a folytonosság miatt

$$\sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} = g(1) = \lim_{x \to 1-0} g(x) = \lim_{x \to 1-0} f(x) = f(1) = \ln 2.$$

Tehát

$$\boxed{\sum_{n=1}^{+\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \ln 2\,.}$$

Az Analízis I. kurzuson ennek a sornak a konvergenciáját már beláttuk (a Leibniz-sorról van szó), és most már a sor összegét is megismertük.

5. Feladat. *Legyen*

$$f(x) := \operatorname{arc\,tg}(x) \qquad (x \in \mathbb{R}).$$

Állítsuk elő az f függvény $a = 0$ ponthoz tartozó Taylor-sorát, és vizsgáljuk meg az előállítás problémáját!

Megoldás. Az arc tg függvény 0 pont körüli Taylor-sorának előállítása a definíció alapján nem egyszerű feladat, mert a magasabb rendű deriváltak (az előző két példával ellentétben) kiszámolása jóval bonyolultabb.

Először vegyük észre, hogy $f \in D(-1,1)$ és ha x helyett x^2-tet írunk az

$$\frac{1}{1+x} = \sum_{n=0}^{+\infty} (-1)^n x^n \qquad (-1 < x < 1)$$

egyenlőségbe, akkor azt kapjuk, hogy

$$f'(x) = \frac{1}{1+x^2} = \sum_{n=0}^{+\infty} (-1)^n (x^2)^n = \sum_{n=0}^{+\infty} (-1)^n x^{2n} = 1 - x^2 + x^4 - x^6 + \cdots \quad (\underbrace{-1 < x^2 < 1}_{-1<x<1}).$$

Legyen

$$g(x) := \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1} = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots \qquad \Big(x \in [-1,1]\Big).$$

A fenti hatványsor konvergenciahalmaza valóban a $[-1,1]$ intervallum, hiszen ez hányadoskritériummal és a végpontokban Leibniz-kritériummal könnyen igazolható. A hatványsorok összegfüggvényének deriváltjára vonatkozó tétel szerint $g \in D(-1,1)$ és

$$g'(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} (2n+1)x^{2n} = \sum_{n=0}^{+\infty} (-1)^n x^{2n} \qquad \Big(x \in (-1,1)\Big).$$

Ezért $f'(x) = g'(x)$ minden $x \in (-1,1)$ pontban. A deriváltak egyenlőségére vonatkozó tételből következik, hogy van olyan $c \in \mathbb{R}$ állandó, hogy

$$f(x) - g(x) = c \quad \Big(x \in (-1,1)\Big).$$

Ugyanakkor $f(0) - g(0) = 0$, így $c = 0$. Bebizonyítottuk tehát azt, hogy

$$f(x) = \operatorname{arc\,tg} x = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \qquad \text{ha} \quad x \in (-1,1).$$

Ezzel sikerült az f függvényt hatványsorba fejteni. Ekkor a 3. Tétel szerint ez a hatványsor a függvény 0 ponthoz tartozó Taylor-sora, ami előállítja a függvényt a $(-1,1)$ intervallumon.

Megjegyzés. Az előző feladat megoldásában szereplő

$$f(x) = \operatorname{arc\,tg} \quad (x \in \mathbb{R}) \quad \text{és} \quad g(x) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} x^{2n+1} \quad (-1 \le x \le 1)$$

függvények folytonosak az $x = -1$ és az $x = 1$ pontokban. Mivel $f(x) = g(x)$ minden $x \in (-1,1)$ esetén, így a folytonosság miatt

$$\sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} = g(1) = \lim_{x \to 1-0} g(x) = \lim_{x \to 1-0} f(x) = f(1) = \operatorname{arc\,tg} 1 = \frac{\pi}{4}.$$

Tehát

$$\boxed{\sum_{n=0}^{+\infty} \frac{(-1)^n}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots = \frac{\pi}{4}.}$$

A fenti képletek alapján a π értéke tetszőleges pontossággal számolható.

6. előadás

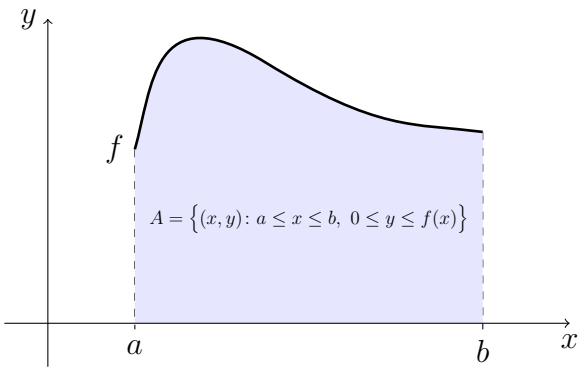
INTEGRÁLSZÁMÍTÁS 1.

Differenciálszámítással számos fontos gyakorlati problémát tudtunk megoldani, ilyenek többek között a görbékhez húzott érintőegyenes felállítása, szélsőértékfeladatok megoldása, kritikus határértékek kiszámítása, egy függvényt jól megközelítő polinom meghatározása, stb. Azonban vannak olyan problémák, amelyeknek megoldásához a deriválással ellentétes folyamat szükséges, vagyis a megoldásban olyan függvény játszi szerepet, melyeknek deriváltja az adott függvény. Ezt az ellentétes folyamatot egyes szerzők „antideriválásnak” nevezik, bár ezt az elnevezést csak most, ismereteink bevezetésében fogjuk alkalmazni. Az antideriválás olyan jelentős gyakorlati problémák megoldásában alkalmazható, mint az ívhossz-, a terület-, a térfogat-, valamint a súlypont- és a nyomatékszámítás.

A területmérő függvény

A következőben azt fogjuk illusztrálni, hogy az antideriválás hogyan játszik szerepet a területszámításban. Fontos megjegyezni, hogy most nem foglalkozunk a terület pontos értelmezésével, hiszen az illusztrációhoz elegendő a területnek az az intuitív megközelítése, amelyet a mindennapokban alkalmazunk.

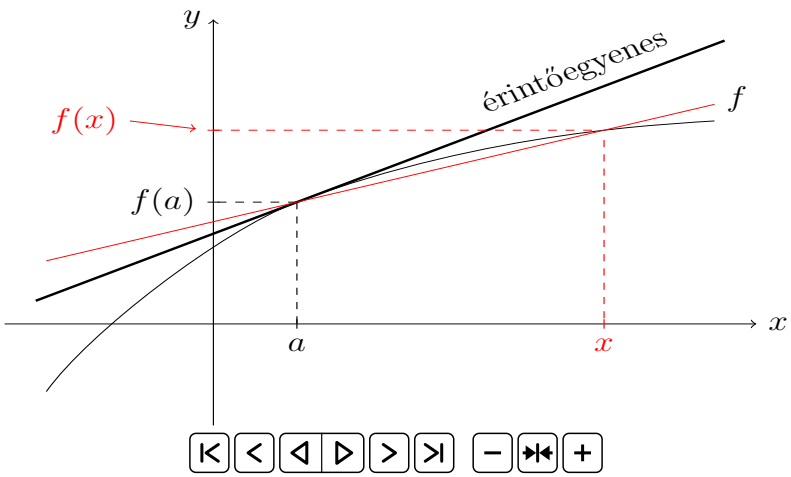
Legyen f egy folytonos függvény, amely egy adott $[a, b]$ zárt intervallumon pozitív értékeket vesz fel. Az a célunk, hogy kiszámítjuk a függvény görbe alatti területét, azaz a függvény grafikonja, az $x = a$ és az $x = b$ egyenesek, valamint az x tengely által közrezárt korlátos síkidom területét. Nevezetesen, az ábrán látható A síkidom területét szeretnénk meghatározni.



A probléma megoldásához vezessünk be egy újabb függvényt! Minden egyes $x \in [a, b]$ értékhez rendeljük az f függvény a-tól x-ig terjedő görbe alatti területét, amit $T(x)$-szel jelölünk. Ezt a T függvényt az f függvény a ponttól induló ***területmérő függvényének*** nevezzük. A területmérő függvény ismeretében egyszerűen megállapíthatjuk a keresett síkidom területét, hiszen ehhez elegendő kiszámolni a területmérő függvény értékét az $x = b$ pontban. Ezért ezentúl a figyelmünk a területmérő függvény meghatározására fog irányulni.

Rögzítsünk egy $x_0 \in (a, b)$ értéket! A következő ábrán zöld színnel jelölt síkidom területe nem más, mint az a ponttól induló területmérő függvény értéke az x_0 pontban. Vegyünk egy olyan $h > 0$ elegendően kicsi számot, hogy $x_0 + h < b$ teljesüljön, és tekintsük azt a síkidomot, amely az f függvény x_0-tól $x_0 + h$-ig terjedő görbéje alatt van. Ezt az ábrán kék színnel jelöltük. A T területmérő függvény ismeretében nem nehéz kiszámolni a kék síkidom területét, hiszen elég kivonni az a-tól $x_0 + h$-ig terjedő görbe alatti területből az a-tól x_0-ig terjedő görbe alatti területet, azaz a kék síkidom területének mértéke:

$$T(x_0 + h) - T(x_0).$$



Másrészt ha egyesítjük az ábrán szereplő kék és piros területrészt, akkor egy olyan téglalapot kapunk, melynek egyik oldala h-val, a másik oldala $f(x_0)$-val egyenlő, így területe $f(x_0)h$. Ezért, ha $\varepsilon(h)$-val jelöljük a piros síkidom területét, akkor fennáll a

$$T(x_0+h) - T(x_0) = f(x_0)h - \varepsilon(h) \tag{1}$$

összefüggés. A piros síkidom része egy olyan téglalapnak, melynek egyik oldala h-val, a másik oldala az $f(x_0) - f(x_0+h)$ értékkel egyenlő, így területe $(f(x_0) - f(x_0+h))h$. Ezért

$$\varepsilon(h) \le \Big(f(x_0) - f(x_0+h)\Big)h.$$

Ebből következik, hogy

$$0 \le \lim_{h\to 0+0} \frac{\varepsilon(h)}{h} \le \lim_{h\to 0+0}\Big(f(x_0) - f(x_0+h)\Big) = 0 \qquad \Longleftrightarrow \qquad \lim_{h\to 0+0} \frac{\varepsilon(h)}{h} = 0,$$

hiszen az f függvény folytonossága azt jelenti, hogy $f(x_0) - f(x_0+h) \to 0$, ha $h \to 0$. Így, ha az (1) egyenlet két oldalát h-val osztjuk és tartunk vele a nullához, akkor azt kapjuk, hogy

$$T'_+(x_0) = \lim_{h\to 0+} \frac{T(x_0+h) - T(x_0)}{h} = f(x_0).$$

Hasonló eredményt kapunk, ha egy kis $h < 0$ számmal az előbbihez hasonló számításokat végzünk, nevezetesen $T'_-(x_0) = f(x_0)$, és így $T'(x_0) = f(x_0)$. Ez, a területmérő függvény egy nagyon érdekes tulajdonságára enged következtetni, miszerint T olyan függvény, amelynek deriváltja az f függvény az (a, b) intervallum minden pontjában.

Fontos hangsúlyozni, hogy az előbbi gondolatmenet csak szemlélteti, de nem igazolja kellő precizitással az előbbi következtetést. Mégis mutatja, hogy miért szükséges az antideriválással foglalkozni. A téma kidolgozását a primitív függvény fogalmával kezdjük.

A határozatlan integrál fogalma

A deriválás egy differenciálható függvényhez hozzárendeli a deriváltfüggvényét. Az előző részben láttuk, hogy célszerű közelebbről megvizsgálni ennek a műveletnek a „megfordítását”: egy adott függvényhez keresünk olyan differenciálható függvényt, hogy ez utóbbinak a deriváltja a kiindulási függvény legyen. A keresendő függvényre érdemes külön elnevezést bevezetni. A következő fontos fogalmat csak az $\mathbb{R}$ egy adott $I \subset \mathbb{R}$ ***nyílt intervallumán*** (ez lehet korlátos és nem korlátos is) értelmezett függvényekre fogjuk bevezetni.

1. Definíció. *Legyen adott az $I \subset \mathbb{R}$ nyílt intervallumon értelmezett $f : I \to \mathbb{R}$ függvény. Azt mondjuk, hogy a $F : I \to \mathbb{R}$ függvény f* **primitív függvénye**, *ha $F \in D(I)$ és $F'(x) = f(x)$ $(x \in I)$.*

Világos, hogy ha a F függvény primitív függvénye f-nek, akkor $\forall\, c \in \mathbb{R}$ esetén $F + c$ is primitív függvénye f-nek.

Példa. Az $f(x) := x^2 \quad (x \in \mathbb{R})$ függvény két különböző primitív függvénye:

$$F_1(x) := \frac{x^3}{3} \quad (x \in \mathbb{R}) \qquad \text{és} \qquad F_2(x) := \frac{x^3}{3} + 1 \quad (x \in \mathbb{R}).$$

Az `elemi függvények deriváltjaira` gondolva több függvény primitív függvényét meg tudjuk már határozni. Például a $(\cos x)' = -\sin x$ deriváltból nem nehéz következtetni, hogy az $f(x) := \sin x \quad (x \in \mathbb{R})$ függvény egyik primitív függvénye $F(x) := -\cos x \quad (x \in \mathbb{R})$.

Kérdések:

1. Milyen függvényeknek van primitív függvénye?
2. Ha f-nek van primitív függvénye, akkor hogyan lehet azt meghatározni?
3. Ha az f egyik primitív függvényét ismerjük, akkor hogyan tudjuk meghatározni az összes többit?

$\boxed{\textbf{1.}}$ Jelenleg nem ismert olyan egyszerűen megfogalmazható, a függvény belső tulajdonságain alapuló feltétel, amelyik *szükséges és elégséges* feltételt ad arra, hogy az adott függvénynek legyen primitív függvénye, de külön tudunk ilyen feltételeket mondani.

Elégséges feltétel primitív függvény létezésére. *Ha $I \subset \mathbb{R}$ nyílt intervallum és $f : I \to \mathbb{R}$ folytonos függvény, akkor f-nek van primitív függvénye.*

A folytonosság mint elégséges feltétel azért nem meglepő, mert a bevezetőből sejteni lehet, hogy egy f folytonos függvény területmérő függvénye az f primitív függvénye.

Szükséges feltétel primitív függvény létezésére. *Ha $I \subset \mathbb{R}$ nyílt intervallum, valamint az $f : I \to \mathbb{R}$ függvénynek van primitív függvénye, akkor f* **Darboux-tulajdonságú** *az I intervallumon, azaz tetszőleges $a, b \in I$, $a < b$, $f(a) \neq f(b)$ esetén az f függvény minden $f(a)$ és $f(b)$ közötti értéket felvesz $[a, b]$-ben.*

Példa. Az előjelfüggvény (vagyis a sgn függvény) a $(-1, 1)$ intervallumon nem Darboux-tulajdonságú, ezért ezen az intervallumon *nincs primitív függvénye.*

$\boxed{\textbf{2.}}$ Egy primitív függvény keresése a deriválás műveletének a „megfordítása”. Deriválni általában nem volt bonyolult feladat, mert a deriválási szabályok összhangban vannak a műveletekkel, de az összeadás és a konstanssal való szorzás kivitelével a deriválási szabályok eredménye összetett, így megfordításukkal nem tudunk egyszerű műveleti szabályokat alkotni a primitív függvények keresésére.

A helyzet ennél súlyosabb. Könnyű meggondolni azt, hogy egy elemi függvény deriváltja mindig elemi függvény. Talán első hallásra meglepőnek tűnhet, de *vannak olyan elemi függvények, amelyeknek a primitív függvénye nem elemi függvény*, így semmilyen szabállyal vagy „ügyes” átalakítással nem tudjuk felírni. Ez jelentős különbség a deriválás és az integrálás között. *Joseph Liouville* (1809–1882) francia matematikus volt az első, aki megmutatta, hogy léteznek ilyen

elemi függvények. Bebizonyítható, hogy pl. az alábbi (folytonos) elemi függvények primitív függvényei nem elemi függvények:

$$e^{-x^2} \quad (x \in \mathbb{R}), \quad \sin x^2 \quad (x \in \mathbb{R}), \quad \frac{\sin x}{x} \quad \big(x \in (0, +\infty)\big), \quad \frac{e^x}{x} \quad \big(x \in (0, +\infty)\big),$$

$$\frac{1}{\ln x} \quad \big(x \in (0, +\infty)\big), \quad \sqrt{x^3 + 1} \quad \big(x \in (0, +\infty)\big).$$

Ennek ellenére, amikor egy függvény primitív függvényeit keressük, akkor ugyanazt a módszert próbáljuk követni, amelyet a határérték és a deriváltak kiszámításánál alkalmaztunk. Először is ismerni kell a legegyszerűbb függvények primitív függvényeit. Ezután szabályokat alkotunk, fogásokat, és egyéb módszereket sajátítunk el, amivel „bonyolultabb" kifejezéseket is tudunk kezelni.

3. Egy primitív függvény ismeretében már könnyedén meghatározható az összes többi, konstansok hozzáadásával. A következő tétel állításait egy korábbi előadáson igazoltuk.

1. Tétel. *Legyen $I \subset \mathbb{R}$ nyílt intervallum és $f : I \to \mathbb{R}$ adott függvény.*

1. *Ha $F : I \to \mathbb{R}$ a f függvény egy primitív függvénye, akkor minden $c \in \mathbb{R}$ esetén az $F + c$ függvény is primitív függvénye f-nek.*
2. *Ha $F_1, F_2 : I \to \mathbb{R}$ primitív függvényei a f függvénynek, akkor*

$$\exists\, c \in \mathbb{R} : \quad F_1(x) = F_2(x) + c \qquad (x \in I),$$

azaz a primitív függvények csak konstansban különböznek egymástól.

Megjegyzés. A tétel 2. állításában lényeges, hogy f **intervallumon** értelmezett függvény. Tekintsük például az

$$f(x) := \begin{cases} 2x, & \text{ha } x \in (0, 1) \\ 0, & \text{ha } x \in (2, 3) \end{cases}$$

függvényt. Legyen

$$F_1(x) := \begin{cases} x^2, & \text{ha } x \in (0, 1) \\ 1, & \text{ha } x \in (2, 3) \end{cases} \qquad \text{és} \qquad F_2(x) := \begin{cases} x^2, & \text{ha } x \in (0, 1) \\ 0, & \text{ha } x \in (2, 3). \end{cases}$$

Ekkor $F_1' = f = F_2'$, de F_1 és F_2 nem csak egy konstansban különböznek egymástól.

2. Definíció. *Az $I \subset \mathbb{R}$ nyílt intervallumon értelmezett f függvény primitív függvényeinek a halmazát f* **határozatlan integráljának** *nevezzük, és így jelöljük:*

$$\int f := \int f(x)\, dx := \Big\{ F : I \to \mathbb{R} \mid F \in D \text{ és } F' = f \Big\}.$$

Ilyenkor f-re az **integrandus**, *illetve az* **integrálandó függvény** *elnevezéseket is használjuk.*

Ha $F \in \int f$, akkor az előző tétel 2° állításából következő $\int f = \Big\{ F + c \mid c \in \mathbb{R} \Big\}$ egyenlőséget rövidebben (és kevésbé precízen) az alábbi formában fogjuk jelölni:

$$\int f(x)\, dx = F(x) + c \quad \big(x \in I\big).$$

Az adott f függvény értelmezési tartományát – vagyis az I intervallumot – mindig feltüntetjük, a $c \in \mathbb{R}$ feltételt a képletbe „beleértjük”, de azt nem írjuk ki. Így például

$$\int \frac{1}{1+x^2}\,dx = \operatorname{arc\,tg} x + c \quad \big(x \in \mathbb{R}\big), \qquad \text{vagy} \qquad \int \frac{1}{\cos^2 x}\,dx = \operatorname{tg} x + c \quad \Big(x \in \big(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\big)\Big).$$

Előfordul, hogy különböző módszerekkel határozunk meg egy határozatlan integrált, de különböző képleteket kapunk. Ez azért lehetséges, mert a két képlet valójában csak egy konstanssal tér el egymástól, ami több esetben csak megfelelő átalakítások után látható.

Határozatlan integrálra vonatkozó szabályok

1. *Alapintegrálok, a határozatlan integrál linearitása*

Az alapintegrálok az alapvető elemi függvények deriváltjának megfordításával kapott határozatlan integrálok. Ezeket `ebben a táblázatban` soroltuk fel. Például

- $\displaystyle\int \frac{1}{x}\,dx = \ln x + c \quad \big(x \in (0, +\infty)\big)$,
- $\displaystyle\int \frac{1}{x}\,dx = \ln(-x) + c \quad \big(x \in (-\infty, 0)\big)$,
- $\displaystyle\int x^\alpha\,dx = \frac{x^{\alpha+1}}{\alpha+1} + c \quad \big(x \in (0, +\infty),\ \alpha \in \mathbb{R} \setminus \{-1\}\big)$.

Itt hívjuk fel a figyelmet arra, hogy a „határozzuk meg az $\int f$-et” feladatokban természetesen elengedhetetlen az f függvény, és így (többek között) az $I = \mathcal{D}_f$ értelmezési tartomány ismerete. Ez utóbbi időnként „elsikkad” a feladat kitűzésekor, mondván, hogy az „magától értetődik”. Ez gyakran vezethet félreértésekre, ezért kerülni kell ezt a pongyolaságot. A továbbiakban mindig megadjuk, hogy melyik $I = \mathcal{D}_f$ *intervallumon* tekintjük az f függvény primitív függvényeit.

Az összeadásra és a konstanssal való szorzásra vonatkozó deriválási szabály megfordítása alapján kapjuk a következő szabályt.

2. Tétel (A határozatlan integrál linearitása). *Legyen $I \subset \mathbb{R}$ nyílt intervallum. Ha az $f, g : I \to \mathbb{R}$ függvényeknek létezik primitív függvénye, akkor tetszőleges $\alpha, \beta \in \mathbb{R}$ mellett $(\alpha f + \beta g)$-nek is létezik primitív függvénye és*

$$\int \big(\alpha f(x) + \beta g(x)\big)\,dx = \alpha \int f(x)\,dx + \beta \int g(x)\,dx \quad (x \in I).$$

Bizonyítás. Legyen $F \in \int f$ és $G \in \int g$. Ekkor $F, G \in D(I)$, illetve $F' = f$ és $G' = g$. Ekkor az említett deriválási szabályok miatt

$$(\alpha F + \beta G)' = \alpha F' + \beta G' = \alpha f + \beta g.$$

Ez azt jelenti, hogy $\alpha F + \beta G \in \int(\alpha f + \beta g)$.

Példa. $\displaystyle\int \big(6x^2 - 8x + 3\big)\,dx = 6 \cdot \frac{x^3}{3} - 8 \cdot \frac{x^2}{2} + 3x + c \qquad (x \in \mathbb{R}).$

Az ebbe a körbe tartozó „elemi fogásokkal” megoldható feladatok sokszor nem egyszerűek, mert át kell alakítani az integrandust úgy, hogy fel tudjuk írni alapintegrálok lineáris kombinációjaként. Például

$$\int \frac{x^2}{1+x^2}\,dx = \int \frac{1+x^2-1}{1+x^2}\,dx = \int \left(1 - \frac{1}{1+x^2}\right)\,dx = x - \operatorname{arc\,tg} x + c \qquad (x \in \mathbb{R}).$$

2. *Az első helyettesítési szabály*

Az összetett függvény deriválására vonatkozó tételnek a „megfordításával” kapcsolatban két állítást fogunk megmutatni. Az egyiket most, a másikat pedig a következő órán ismertetjük.

3. Tétel (Az első helyettesítési szabály). *Legyenek $I, J \subset \mathbb{R}$ nyílt intervallumok és $g : I \to \mathbb{R}$, $f : J \to \mathbb{R}$ függvények. Tegyük fel, hogy $g \in D(I)$, $\mathcal{R}_g \subset J$ és az f függvénynek van primitív függvénye. Ekkor az $(f \circ g) \cdot g'$ függvénynek is van primitív függvénye és*

$$\int f\Big(g(x)\Big) \cdot g'(x)\, dx = F\Big(g(x)\Big) + c \qquad (x \in I),$$

ahol F az f függvény egy primitív függvénye.

Bizonyítás. Legyen $F \in \int f$. Ekkor $F \in D(J)$ és $F' = f$. Az összetett függvény deriválására vonatkozó tétel szerint ekkor $F \circ g \in D(I)$ és

$$(F \circ g)' = \Big(F' \circ g\Big) \cdot g' = (f \circ g) \cdot g'.$$

Ez azt jelenti, hogy $F \circ g \in \int (f \circ g) \cdot g'$.

Példa. $\displaystyle\int xe^{x^2}\, dx = \frac{1}{2}\int 2xe^{x^2}\, dx = \frac{1}{2}\int \exp(x^2) \cdot (x^2)'\, dx = \frac{1}{2}\exp(x^2) + c = \frac{1}{2}e^{x^2} + c \quad (x \in \mathbb{R}).$

Az első helyettesítési szabály speciális eseteit gyakorlaton fogjuk megismerni.

3. *A parciális integrálás*

A szorzatfüggvény deriválására vonatkozó tétel „megfordítását” fejezi ki a következő állítás.

4. Tétel (A parciális integrálás szabálya). *Legyen $I \subset \mathbb{R}$ nyílt intervallum. Tegyük fel, hogy $f, g \in D(I)$ és az $f'g$ függvénynek létezik primitív függvénye I-n. Ekkor az fg' függvénynek is van primitív függvénye és*

$$\int f(x)g'(x)\, dx = f(x)g(x) - \int f'(x)g(x)\, dx \quad (x \in I).$$

Bizonyítás. Ha $F \in \int f'g$, akkor $F \in D(I)$ és $F' = f'g$. Mivel $fg \in D(I)$, illetve $(fg)' = f'g + fg'$, ezért $(fg - F) \in D(I)$ és

$$(fg - F)' = f'g + fg' - f'g = fg'.$$

Így $(fg - F) \in \int fg'$ valóban fennáll.

Példa. $\displaystyle\int xe^x\, dx = \int \underbrace{x}_{f} \cdot \underbrace{e^x}_{g'}\, dx = \int \underbrace{x}_{f} \cdot (\underbrace{e^x}_{g})'\, dx = \underbrace{x}_{f} \cdot \underbrace{e^x}_{g} - \int \underbrace{1}_{f'} \cdot \underbrace{e^x}_{g}\, dx =$

$$= xe^x - \int e^x\, dx = xe^x - e^x + c \qquad (x \in \mathbb{R}).$$

Gyakorlaton olyan integráltípusokat fogunk tanulni, ahol érdemes a parciális integrálás szabályát alkalmazni.

4. *A második helyettesítési szabály*

Megemlítettük már, hogy az összetett függvény deriválására vonatkozó tételnek a „megfordításával” kapcsolatban két állítást fogalmazunk meg. Az elsőt már megismertük. A másodikhoz szükséges, hogy a g belső függvény ***invertálható*** legyen.

5. Tétel (A második helyettesítési szabály). *Legyenek $I, J \subset \mathbb{R}$ nyílt intervallumok. Tegyük fel, hogy $f : I \to \mathbb{R}$, $g : J \to I$, $\mathcal{R}_g = I$, $g \in D(J)$, $g' > 0$ J-n (vagy $g' < 0$ J-n) és az $(f \circ g) \cdot g' : J \to \mathbb{R}$ függvénynek van primitív függvénye. Ekkor az f függvénynek is van primitív függvénye és*

$$\int f(x)\,dx \underset{x=g(t)}{=} \int f\big(g(t)\big) \cdot g'(t)\,dt\Big|_{t=g^{-1}(x)} \qquad (x \in I).$$

Bizonyítás. Ha $g' > 0$ J-n, akkor g szigorúan monoton növekvő, és ha $g' < 0$ J-n, akkor g szigorúan monoton csökkenő. Mindkét esetben g invertálható, és mivel g folytonos és $g' \neq 0$ J-n, így az inverz függvény deriváltjára vonatkozó tétel szerint a $g^{-1} \colon I \to J$ függvény differenciálható I minden pontjában, és

$$\big(g^{-1}\big)' = \frac{1}{g' \circ g^{-1}}.$$

Legyen $H \in \int (f \circ g) \cdot g'$, akkor $H \in D(J)$ és $H' = (f \circ g) \cdot g'$. Ekkor az összetett függvény deriválására vonatkozó tétel szerint $H \circ g^{-1} \in D(I)$ és

$$\begin{aligned}\big(H \circ g^{-1}\big)' &= \big(H' \circ g^{-1}\big) \cdot \big(g^{-1}\big)' = \Big(\big((f \circ g) \cdot g'\big) \circ g^{-1}\Big) \cdot \big(g^{-1}\big)' = \\ &= \big((f \circ g \circ g^{-1}) \cdot (g' \circ g^{-1})\big) \cdot \big(g^{-1}\big)' = \big(f \cdot (g' \circ g^{-1})\big) \cdot \big(g^{-1}\big)' = \\ &= f \cdot (g' \circ g^{-1}) \cdot \frac{1}{g' \circ g^{-1}} = f.\end{aligned}$$

Ez azt jelenti, hogy $H \circ g^{-1} \in \int f$, amiből a tétel állítása következik.

A második helyettesítési szabályt a következő módon szoktuk alkalmazni. Tegyük fel, hogy egy $\int f(x)\,dx$ határozatlan integrált szeretnénk kiszámítani. Egy alkalmas, a szabály feltételeit teljesítő g függvénnyel a „régi” x változó helyett vezessük be az $x = g(t)$ egyenlőségből adódó $t = g^{-1}(x)$ „új” változót. Ezzel az x változót t-re helyettesítjük az f függvényben és ezt „megszorozzuk” $g'(t)\,dt$-vel. Az így kapott integrált kiszámítjuk (***ha tudjuk***) és utána a t változót x-re visszahelyettesítjük.

Példa. Az

$$\int \sqrt{1 - x^2}\,dx \quad \big(x \in (-1, 1)\big)$$

határozatlan integrál kiszámításához az

$$x = \sin t =: g(t) \quad \Big(t \in \big(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\big)\Big)$$

helyettesítést alkalmazzuk. Mivel $g'(t) = \cos t > 0$ $\Big(t \in \big(-\frac{\pi}{2}, \frac{\pi}{2}\big)\Big)$, így $g \uparrow$, azaz g invertálható. Másrészt $\mathcal{R}_g = (-1, 1)$, így

$$t = g^{-1}(x) = \arcsin x \ \big(x \in (-1, 1)\big).$$

A második helyettesítési szabályra vonatkozó tétel tehát alkalmazható. Így

$$\int \sqrt{1-x^2}\,dx = \int \sqrt{1-\sin^2 t}\cdot\cos t\,dt = \Big(\cos t > 0 \text{ miatt } \sqrt{1-\sin^2 t} = \cos t\Big) =$$

$$= \int \cos^2 t\,dt = \int \frac{1+\cos 2t}{2}\,dt = \frac{t}{2} + \frac{\sin 2t}{4} + c\Big|_{t=\arcsin x} = \frac{\arcsin x}{2} + \frac{\sin(2\arcsin x)}{4} + c.$$

Itt a második tag egyszerűbb alakra hozható:

$$\sin 2t = 2\sin t\cos t = 2\sin t\sqrt{1-\sin^2 t} = 2x\sqrt{1-x^2}.$$

Végül azt kapjuk, hogy

$$\int \sqrt{1-x^2}\,dx = \frac{1}{2}\cdot\arcsin x + \frac{1}{2}\cdot x\sqrt{1-x^2} + c \quad \Big(x\in(-1,1)\Big).$$

A határozott integrál motivációja

Az előadás bevezetőjében folytonos függvények görbe alatti területét vizsgáltuk. Ott még nem foglalkoztunk a terület pontos értelmezésével, csak a területről alkotott intuitív megközelítést alkalmaztunk. Differenciálszámítással több geometria fogalomnak (pl. az érintőnek) adtunk analitikus értelmezést, amivel már számolni tudunk. Jogos kérdés, hogy nem tudnánk-e hasonlóan eljárni a függvények görbe alatti területével is.

A pontos feladat a következő: Legyen f nemnegatív korlátos (***nem feltétlenül folytonos***) függvény a korlátos és zárt $[a,b]$ intervallumon, és tekintsük az f grafikonja alatti

$$A := \Big\{(x,y) \mid x\in[a,b],\ 0\le y\le f(x)\Big\}$$

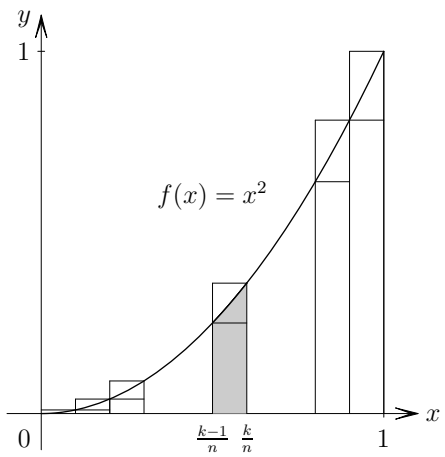
síkidomot. Hogyan érdemes A területét ***értelmezni***, milyen f esetén beszélhetünk az A halmaz területéről, és hogyan lehet a $T(A)$-val jelölt területet ***kiszámolni***?

Abból a görög matematikusok által már alkalmazott elég „természetes” ötletből indulunk ki, hogy a szóban forgó síkidom területét téglalapok területeinek az összegével közelítjük. Lássuk, hogyan határozta meg Arkhimédész a parabola alatti területet! A modern jelöléseket és a már megismert fogalmakat fogjuk használni.

Tekintsük az

$$f(x) := x^2 \quad \Big(x\in[0,1]\Big)$$

függvényt! Rögzítsünk egy $1\le n\in\mathbb{N}$ számot, és osszuk fel a $[0,1]$ intervallumot az $x_k := \frac{k}{n}$ $(k=0,1,2,\dots,n)$ osztópontokkal. Közelítsük az A halmazt az alábbi ábrán szemléltetett „beírt” és „körülírt” téglalapokkal.



Jelölje s_n, illetve S_n a szóban forgó téglalapok területeinek az összegét. Az

$$1^2+2^2+3^2+\dots+m^2 = \frac{m(m+1)(2m+1)}{6}$$

$(m\in\mathbb{N}^+)$ képletből következik, hogy

$$s_n = \frac{1}{n}\left(\left(\frac{0}{n}\right)^2 + \left(\frac{1}{n}\right)^2 + \left(\frac{2}{n}\right)^2 + \cdots \left(\frac{n-1}{n}\right)^2\right) = \frac{(n-1)n(2n-1)}{6n^3} \quad \text{és}$$

$$S_n = \frac{1}{n}\left(\left(\frac{1}{n}\right)^2 + \left(\frac{2}{n}\right)^2 + \cdots \left(\frac{n}{n}\right)^2\right) = \frac{n(n+1)(2n+1)}{6n^3}.$$

Ha növeljük az osztópontok n számát, akkor a „lépcsősidomok” egyre jobban közelítenek az A halmazhoz. Világos, hogy minden $1 \leq n \in \mathbb{N}$ esetén

$$s_n \leq T(A) \leq S_n.$$

Mivel

$$\lim_{n\to+\infty} s_n = \lim_{n\to+\infty} \frac{1}{6} \cdot \left(1 - \frac{1}{n}\right) \cdot \left(2 - \frac{1}{n}\right) = \frac{1}{3} \quad \text{és}$$
$$\lim_{n\to+\infty} S_n = \lim_{n\to+\infty} \frac{1}{6} \cdot \left(1 + \frac{1}{n}\right) \cdot \left(2 + \frac{1}{n}\right) = \frac{1}{3},$$

ezért a fenti egyenlőtlenségeket csak az $1/3$ szám elégíti ki. Kézenfekvő tehát azt mondani, hogy az A halmaznak van területe, és az legyen $1/3$.

Megjegyzés. A fenti eredmény összhangban van azzal, amit a területmérő függvényről mondtunk, de még nem igazoltunk. Ha $T(x)$ a 0-tól x-ig terjedő parabola alatti terület, akkor a $T(1)$ értékre vagyunk kíváncsiak. Azt mondtuk, hogy T olyan függvény, amelynek deriváltja az $f(x) = x^2$ függvényt adja a $(0, 1)$ intervallum minden pontjában. Ezért $\exists c \in \mathbb{R}$ úgy, hogy

$$T(x) = \frac{x^3}{3} + c \qquad (0 < x < 1),$$

de mivel logikus, hogy területmérő függvény folytonos, ezért a fenti egyenlőség kiterjeszthető a $[0, 1]$ intervallumra. 0-tól 0-ig még 0 a görbe alatti terület, azaz $T(0) = 0$. Tehát

$$T(0) = \frac{0^3}{3} + c = 0 \quad \Longrightarrow \quad c = 0, \qquad \text{így} \qquad T(1) = \frac{1^3}{3} + 0 = \frac{1}{3}.$$

A határozott integrál értelmezése

A határozott integrált a korlátos és zárt $[a, b] \subset \mathbb{R}$ intervallumon értelmezett korlátos függvények körében értelmezzük. A $\boldsymbol{K[a, b]}$ szimbólummal fogjuk jelölni az ilyen függvények halmazát:

$$K[a, b] := \Big\{ f : [a, b] \to \mathbb{R} \mid f \text{ korlátos } [a, b]\text{-n} \Big\}.$$

Legyen $a, b \in \mathbb{R}$, $a < b$ és $n \in \mathbb{N}$. Az $[a, b]$ intervallum egy ***felosztásán*** a

$$\tau := \Big\{ a = x_0 < x_1 < x_2 < \cdots < x_n = b \Big\}$$

halmazt értjük. $\boldsymbol{\mathcal{F}[a, b]}$ jelöli az $[a, b]$ intervallum felosztásainak a halmazát. A

$$\|\tau\| := \max\Big\{ x_{k+1} - x_k \mid k = 0, 1, \ldots, n-1 \Big\}$$

számot a τ felosztás ***finomságának*** nevezzük. Legyen $f \in K[a, b]$, $\tau \in \mathcal{F}[a, b]$, továbbá

$$m_k := \inf\{ f(x) \mid x_k \leq x \leq x_{k+1} \}, \qquad \text{illetve} \qquad M_k := \sup\{ f(x) \mid x_k \leq x \leq x_{k+1} \}$$

minden $k = 0, ..., n-1$ indexre. A

$$s(f, \tau) := \sum_{k=0}^{n-1} m_k (x_{k+1} - x_k), \qquad \text{illetve} \qquad S(f, \tau) := \sum_{k=0}^{n-1} M_k (x_{k+1} - x_k)$$

összegeket az f függvény τ felosztáshoz tartozó ***alsó***, illetve ***felső közelítő összegének*** nevezzük.

Mivel tetszőleges $f \in K[a,b]$ függvény és $\tau \in \mathcal{F}[a,b]$ felosztás esetén

$$-\infty < \inf_{x\in[a,b]} f(x)\cdot(b-a) \le s(f,\tau) \le S(f,\tau) \le \sup_{x\in[a,b]} f(x)\cdot(b-a) < +\infty,$$

ezért az

$$\Big\{s(f,\tau) \mid \tau \in \mathcal{F}[a,b]\Big\} \qquad \text{és az} \qquad \Big\{S(f,\tau) \mid \tau \in \mathcal{F}[a,b]\Big\}$$

halmazok korlátosak. Következésképpen mindegyik halmaz infimuma és szuprémuma véges. Az

$$I_*(f) := \sup\Big\{s(f,\tau) \;\Big|\; \tau \in \mathcal{F}[a,b]\Big\}, \quad \text{illetve az} \quad I^*(f) := \inf\Big\{S(f,\tau) \;\Big|\; \tau \in \mathcal{F}[a,b]\Big\}$$

valós számot az f függvény ***Darboux-féle alsó integráljának***, illetve ***Darboux-féle felső integráljának*** nevezzük. Világos, hogy

$$I_*(f) \le I^*(f) \quad \Big(f \in K[a,b]\Big).$$

3. Definíció. *Tegyük fel, hogy $-\infty < a < b < +\infty$ és $f : [a,b] \to \mathbb{R}$ egy korlátos függvény. Azt mondjuk, hogy f* ***Riemann-integrálható*** *az $[a,b]$ intervallumon (röviden* **integrálható** *$[a,b]$-n), ha $I_*(f) = I^*(f)$. Ekkor az*

$$\int_a^b f := \int_a^b f(x)\,dx := I_*(f) = I^*(f)$$

számot az f függvény $[a,b]$ intervallumon vett **Riemann-integráljának** *(vagy más szóval* **határozott integráljának***) nevezzük. Az $[a,b]$ intervallumon Riemann-integrálható függvények halmazát az $\boldsymbol{R[a,b]}$ szimbólummal fogjuk jelölni.*

Az előzőekből következik, hogy bizonyos síkidomok területét a következőképpen célszerű *értelmezni.*

4. Definíció. *Ha a korlátos $f : [a,b] \to \mathbb{R}$ függvény Riemann-integrálható az $[a,b]$ intervallumon és $f(x) \ge 0$ $(x \in [a,b])$, akkor* **az $\boldsymbol{f}$ grafikonja alatti**

$$A := \{(x,y) \in \mathbb{R}^2 \mid a \le x \le b,\ 0 \le y \le f(x)\}$$

síkidomnak *van* **területe***, és a területét így értelmezzük:*

$$T(A) := \int_a^b f(x)\,dx.$$

Egyszerű példát lehet megadni olyan korlátos f függvényre, amelyre $I_*(f) < I^*(f)$, ami azt jelenti, hogy a függvény *nem integrálható.* Például legyen

$$f(x) := \begin{cases} 0, & \text{ha } x \in [0,1] \cap \mathbb{Q} \\ 1, & \text{ha } x \in [0,1] \cap (\mathbb{R} \setminus \mathbb{Q}). \end{cases}$$

Ekkor $I_*(f) = 0$ és $I^*(f) = 1$, ezért $f \notin R[0,1]$.

A határozott integrál tulajdonságai

Az ebben a részben szereplő tételeket bizonyítás nélkül mondjuk ki.

Először azt fogalmazzuk meg, hogy a Riemann-integrál „érzéketlen” a függvény *véges* halmazon való „viselkedésére”. Más szóval, ha egy Riemann-integrálható függvényt egy véges halmazon (tetszőlegesen) megváltoztatunk, akkor az így kapott „új” függvény is Riemann-integrálható lesz, és a (Riemann-)integrálja ugyanaz marad, mint a kiindulási függvényé.

6. Tétel. *Legyen $f, g \in K[a,b]$ és tegyük fel, hogy az $A := \left\{x \in [a,b] \;\middle|\; f(x) \neq g(x)\right\}$ halmaz véges. Ekkor*

a) $f \in R[a,b] \iff g \in R[a,b]$,

b) ha $f \in R[a,b]$, akkor $\int\limits_a^b f = \int\limits_a^b g$.

A következő tételben kiderül, hogy a folytonosság „erősebb” tulajdonság a Riemann-integrálhatóságnál.

7. Tétel. *Legyen $a, b \in \mathbb{R}$, $a < b$. Tegyük fel, hogy az $f : [a,b] \to \mathbb{R}$ függvény folytonos. Ekkor $f \in R[a,b]$, azaz* **minden folytonos függvény integrálható**. *(jelekkel $C[a,b] \subset R[a,b]$).*

Az állítás megfordítása nem igaz. Legyen például

$$f(x) := \begin{cases} 1, & \text{ha } 0 < x \leq 1 \\ 0, & \text{ha } x = 0. \end{cases}$$

Ekkor $f \in R[0,1]$, de $f \notin C[0,1]$.

Megjegyzés. A 6. és 7. Tételből következik, hogy a véges sok szakadási hellyel rendelkező függvények integrálhatók. Kérdés, hogy a szakadási helyek számát valamilyen értelemben lehet-e növelni úgy, hogy a függvény továbbra is integrálható maradjon. Kiderül, hogy egy függvény Riemann-integrálhatósága lényegében azon múlik, hogy a függvény szakadási helyeinek a halmaza mennyire „kicsi”.

Precízen: Azt mondjuk, hogy az $A \subset \mathbb{R}$ halmaz ***nullmértékű***, ha tetszőleges $\varepsilon > 0$ számhoz megadható $I_k \subset \mathbb{R}$ $(k \in \mathbb{N})$ intervallumoknak egy olyan sorozata, hogy

$$A \subset \bigcup_{k=1}^{+\infty} I_k \qquad \text{és} \qquad \sum_{k=1}^{+\infty} |I_k| < \varepsilon.$$

A Riemann-integrálhatóság Lebesgue-kritériuma: *Tegyük fel, hogy $f \in K[a,b]$ és legyen az f szakadási helyeinek a halmaza $\mathcal{A} := \left\{x \in [a,b] \;\middle|\; f \notin C\{x\}\right\}$. Ekkor $f \in R[a,b]$ azzal ekvivalens, hogy az $\mathcal{A}$ halmaz nullmértékű.*

Az integrálás és a függvényműveletek kapcsolatára vonatkoznak az alábbi állítások.

8. Tétel. *Tegyük fel, hogy f, $g \in R[a,b]$. Ekkor*

a) minden $\alpha, \beta \in \mathbb{R}$ esetén

$$\alpha f + \beta g \in R[a,b] \qquad \text{és} \qquad \int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g,$$

b) $f \cdot g \in R[a,b]$,

c) ha valamilyen $m > 0$ állandóval fennáll az

$$|g(x)| \geq m > 0 \quad (x \in [a,b])$$

egyenlőtlenség, akkor az $\dfrac{f}{g}$ függvény is integrálható az $[a,b]$ intervallumon.

A Riemann-integrál monotonitását fejezi ki az alábbi állítás.

9. Tétel. *Tegyük fel, hogy $f \in R[a,b]$. Ekkor*

a) ha $g \in R[a,b]$, és $f(x) \leq g(x)$ $(x \in [a,b])$, akkor $\int_a^b f \leq \int_a^b g$,

b) $|f| \in R[a,b]$, és $\left|\int_a^b f\right| \leq \int_a^b |f|$.

Az $\int_a^b f$ jelölés használatánál eddig feltettük, hogy $a < b$. Az $\int_a^b f$ szimbólumnak $a = b$ és $a > b$ esetén is célszerű értelmet tulajdonítani. Megállapodunk abban, hogy

$$\int_a^a f := 0, \qquad \text{és} \qquad \int_a^b f := -\int_b^a f, \quad \text{ha} \quad a > b.$$

A Riemann-integrál egy további fontos tulajdonságát fejezi ki a következő állítás.

10. Tétel (Intervallum szerinti additivitás). *Ha $f \in R[A,B]$, akkor igaz, hogy minden $a, b, c \in [A,B]$ esetén*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$

Történeti utalások

A matematikai analízis alapgondolatának a felfedezése – vagyis hogy a keresett mennyiséget tetszőleges pontossággal való megközelítések segítségével határozzuk meg – *Eudoxos* (i.e. 408–355) nevéhez fűződik, aki megalkotta a *kimerítés módszerét*.

Arkhimédész (i.e. 287–212) minden idők egyik legnagyobb, de az ókornak minden bizonnyal legnagyobb matematikusa volt. A kimerítés módszerét továbbfejlesztve kiszámította különböző

görbevonalú idomok (pl. parabolaszelet) területét, meghatározta a gömb térfogatát és felszínét, bizonyos spirálok ívhosszúságát, vizsgálta a forgási paraboloidokat és hiperboloidokat. Munkásságának nagyobb része elveszett, de így is hatalmas művet hagyott hátra. Meg kell jegyezni azonban azt is, hogy gondolatai nagyon sokáig nem találtak méltó folytatásra.

Az analízis mint széles körben alkalmazható általános módszer, mint tudományág csak akkor született meg, amikor XVII. századi európai matematikusok kidolgoztak egy elméletet, az ún. *kalkulust* vagy a mai szóval *differenciálszámítást*. Ezt nagy matematikusok sora (*Barrow, Cavalieri, Fermat, Kepler* és sokan mások) fejlesztették ki, majd *Isaac Newton* (1643–1727) és *Gottfried Wilhelm Leibniz* (1646–1716) foglalták össze. Így a század végére már megérett az idő egy nagyszabású monográfia megírására. Ez *L'Hospital* (1661–1704) *Infinitézimál-számítás* (azaz végtelen kicsiny mennyiségekkel való számolás) című műve volt (1696), amely csaknem 100 évig a téma legfontosabb tankönyve maradt.

A kalkulust kezdettől fogva sok kritika és támadás érte – tegyük hozzá teljes joggal. A módszer logikai tisztasága nagyon is vitatható volt, mert homályos fogalmakkal dolgozott, és a gondolatmenetei néha zavarosak voltak. Pl. mit jelent az, hogy végtelenül kicsiny mennyiség? A kalkulus körüli vita egészen a XIX. század végéig zajlott, és nemegyszer filozófiai síkra terelődött (Berkeley, Hegel).

Ezeket a belső problémákat végül mégis a matematikusok oldották meg a XIX. században, amikor a kalkulus intuitív, de homályos és ellentmondásos fogalmait precízen definiált matematikai fogalmakkal helyettesítették. A változó mennyiség fogalmát a függvény fogalmával, a differenciált a határértékkel, a differenciálhányadost pedig a deriválttal váltották fel. Ennek a tisztázási folyamatnak az eredményeképpen – amelyben *Augustin Cauchy* (1789–1857), *Karl Weierstrass* (1815–1897) és *Richard Dedekind* (1831–1916) vállaltak úttörő szerepet – a XIX. század végére a ***differeciál- és integrálszámítás*** (röviden ***analízis***) elérte a logikai tisztaságnak azt a fokát, amelyet a matematika megkövetel.

Az analízis precíz elméletének kidolgozása az újkori nyugati kultúra egyik legnagyobb szellemi teljesítménye volt. Ne csodálkozzunk hát, ha ezt az elméletet – főleg az alapjait, mindenekelőtt pedig annak centrális fogalmát, a határértéket – nehéznek találjuk.

7. előadás

INTEGRÁLSZÁMÍTÁS 2.

Az integrálfüggvény

A határozott integrál értelmezésével és tulajdonságai ismeretében már érdemben tudunk foglalkozni a területmérő függvénnyel, két apró megjegyzéssel. Az egyik, hogy az $\int_a^b f$ integrált érdemes az $a \geq b$ esetekre is értelmezni a következő módon:

$$\int_a^a f := 0, \qquad \text{és} \qquad \int_a^b f := -\int_b^a f, \quad \text{ha} \quad a > b.$$

Ezzel értelmeztük a területmérést „visszafelé” is. A másik, hogy a területmérő elnevezést már nem fogjuk többet használni.

1. Definíció. *Tegyük fel, hogy $f \in R[a,b]$ és $x_0 \in [a,b]$. Az*

$$F(x) := \int_{x_0}^x f(t)\,dt \qquad \Big(x \in [a,b]\Big)$$

függvényt **az f függvény x_0 pontban eltűnő integrálfüggvényének** *nevezzük.*

Megjegyzések.

1. F jól értelmezett, mert azt tanultuk, hogy $f \in R[a,b] \implies f \in R[x_0,x]$.
2. az „x_0 pontban eltűnő” kifejezés arra utal, hogy $F(x_0) = 0$.

1. Tétel (Az integrálfüggvény folytonossága). *Legyen $f \in R[a,b]$, $x_0 \in [a,b]$ és F az f függvény x_0 pontban eltűnő integrálfüggvénye. Ekkor $F \in C[a,b]$.*

Bizonyítás. Mivel f korlátos függvény, hiszen $f \in R[a,b]$, így $\exists K > 0$, hogy

$$|f(x)| \leq K \qquad \Big(x \in [a,b]\Big).$$

Legyen $x \in [a,b]$ egy tetszőleges pont és $(x_n)\colon \mathbb{N} \to [a,b]$ olyan sorozat, hogy $\lim(x_n) = x$. Tegyük fel, hogy $x_n \geq x \quad (n \in \mathbb{N})$. Ekkor a határozott integrál tulajdonságai alapján

$$\Big|F(x_n) - F(x)\Big| = \left|\int_{x_0}^{x_n} f(t)\,dt - \int_{x_0}^{x} f(t)\,dt\right| = \left|\int_{x}^{x_n} f(t)\,dt\right| \leq \int_{x}^{x_n} |f(t)|\,dt \leq \int_{x}^{x_n} K\,dt =$$

$$= K(x_n - x) \to 0, \quad \text{azaz} \quad F(x_n) \to F(x), \quad \text{ha} \quad n \to +\infty.$$

Ezért az átviteli elv szerint F jobbról folytonos az x pontban.

Hasonlóan igazolható, hogy F balról folytonos az x pontban. Ezzel a tétel állítását igazoltuk.

2. Tétel (Az integrálfüggvény deriválthatósága). *Legyen $f \in R[a,b]$, $x_0 \in [a,b]$ és F az f függvény x_0 pontban eltűnő integrálfüggvénye. Tegyük fel, hogy $x \in (a,b)$ olyan pont, amire $f \in C\{x\}$ teljesül. Ekkor $F \in D\{x\}$ és $F'(x) = f(x)$.*

Bizonyítás. Legyen $h > 0$ olyan szám, amire $x + h < b$ teljesül. Ekkor a határozott integrál tulajdonságai alapján

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\left(\int_{x_0}^{x+h} f(t)\,dt - \int_{x_0}^{x} f(t)\,dt\right) = \frac{1}{h}\int_{x}^{x+h} f(t)\,dt$$

Mivel x egy rögzített szám, így

$$\frac{1}{h}\int_{x}^{x+h} f(x)\,dt = f(x)\cdot\frac{1}{h}\cdot\int_{x}^{x+h} 1\,dt = f(x)\cdot\frac{1}{h}\cdot(x+h-x) = f(x).$$

Ezért az integrál linearitása alapján

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h}\int_{x}^{x+h} f(t)\,dt - f(x) + f(x) = \frac{1}{h}\int_{x}^{x+h} f(t)\,dt - \frac{1}{h}\int_{x}^{x+h} f(x)\,dt + f(x) =$$

$$= \frac{1}{h}\int_{x}^{x+h} \Big(f(t) - f(x)\Big)\,dt + f(x).$$

Mivel $f \in C\{x\}$, így a definíció szerint

$$\forall\varepsilon > 0\text{-hoz } \exists\delta > 0,\ \text{hogy } \forall t \in [a,b],\ |t - x| < \delta\colon \Big|f(t) - f(x)\Big| < \varepsilon.$$

Ekkor a fentiek alapján, ha $h < \delta$, akkor minden $x \le t < x + h$ szám esetén

$$\left|\frac{F(x+h) - F(x)}{h} - f(x)\right| = \left|\frac{1}{h}\int_{x}^{x+h} \Big(f(t) - f(x)\Big)\,dt\right| \le \frac{1}{h}\int_{x}^{x+h} \Big|f(t) - f(x)\Big|\,dt <$$

$$< \frac{1}{h}\int_{x}^{x+h} \varepsilon\,dt = \frac{1}{h}\cdot\varepsilon\cdot(x+h-x) = \varepsilon \qquad \Longrightarrow \qquad \exists F'_{+}(x) = f(x).$$

Hasonlóan igazolható, hogy létezik F bal oldali deriváltja az x pontban, és $F'_{-}(x) = f(x)$. Ezzel a tétel állítását igazoltuk.

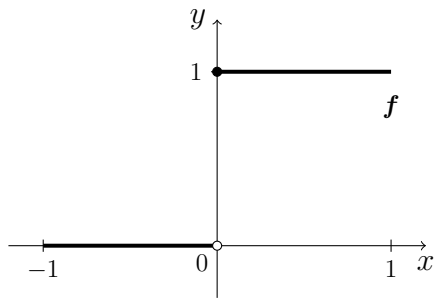
Megjegyzések.

1. Az előző állításokat így foglalhatjuk össze: az integrálfüggvény minden pontban folytonos, és olyan pontokban differenciálható, ahol az eredeti függvény folytonos. Ezekben a pontokban az integrálfüggvény deriváltja az eredeti függvény értékével egyenlő.

2. Következmény: minden nyílt intervallumon értelmezett folytonos függvénynek van primitív függvénye.

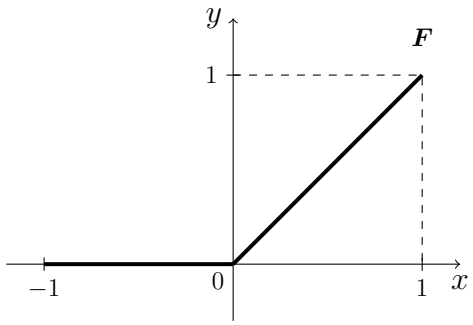
Példa. Ha $f\colon [-1,1] \to \mathbb{R}$,

$$f(x) := \begin{cases} 0, & \text{ha } -1 \le x < 0, \\ 1, & \text{ha } 0 \le x \le 1, \end{cases}$$



akkor az $x_0 = 0$ pontban eltűnő integrálfüggvény

$$F(x) := \begin{cases} 0, & \text{ha } -1 \le x < 0, \\ x, & \text{ha } 0 \le x \le 1. \end{cases}$$



A határozott integrál kiszámítása

Az előzőek szerint, ha $f \in C[a,b]$, akkor tetszőleges $x_0 \in [a,b]$ pontban eltűnő F integrálfüggvényre igaz, hogy $F \in C[a,b]$, $F \in D(a,b)$ és $F'(x) = f(x)$ minden $x \in (a,b)$ esetén. Az ilyen függvényekre érdemes külön elnevezést bevezetni.

2. Definíció. *Legyen $[a,b] \subset \mathbb{R}$ korlátos és zárt intervallum. A $F : [a,b] \to \mathbb{R}$ függvény az $f : [a,b] \to \mathbb{R}$ függvény egy* **primitív függvénye az $[a,b]$ intervallumon**, *ha*

- $F \in C[a,b]$,
- $F \in D(a,b)$ *és* $F'(x) = f(x) \quad \big(x \in (a,b)\big)$.

A következő rendkívül fontos tételt a kalkulus alaptételének is nevezik.

3. Tétel (Newton–Leibniz-formula). *Tegyük fel, hogy*

- $f \in R[a,b]$ *és*
- *az f függvénynek van primitív függvénye az $[a,b]$ intervallumon.*

Ekkor

$$\int_a^b f(x)\,dx = F(b) - F(a) =: \Big[F(x)\Big]_a^b,$$

ahol F az f függvény egy primitív függvénye.

Bizonyítás. Legyen $n \in \mathbb{N}^+$, és tekintsük az $[a,b]$ intervallum egy tetszőlegesen megválasztott $\tau = \Big\{a = x_0 < x_1 < \cdots < x_n = b\Big\}$ felosztását. A Lagrange-féle középértéktétel szerint minden $i = 0, 1, \ldots, n-1$ indexre van olyan $\xi_i \in (x_i, x_{i+1})$ pont, amelyre

$$F(x_{i+1}) - F(x_i) = F'(\xi_i)\,(x_{i+1} - x_i) = f(\xi_i)\,(x_{i+1} - x_i)$$

teljesül. Ha ezeket az egyenlőtlenségeket összeadjuk minden $i = 0, 1, \ldots, n-1$ indexre, akkor a bal oldalon minden tag kiesik, kivéve az $F(x_n) = F(b)$ és az $F(0) = F(a)$ tagokat. Így azt kapjuk, hogy

$$F(b) - F(a) = \sum_{i=0}^{n-1} f(\xi_i)\,(x_{i+1} - x_i).$$

Mivel $\inf\limits_{x\in[x_i,x_{i+1}]} f(x) \le f(\xi_i) \le \sup\limits_{x\in[x_i,x_{i+1}]} f(x)$ $(i = 0, 1, \ldots, n-1)$, ezért a

$$s(f,\tau) \le F(b) - F(a) \le S(f,\tau)$$

egyenlőtlenség minden $\tau \in \mathcal{F}[a,b]$ felosztásra teljesül. Következésképpen

$$I_*(f) = \sup_{\tau\in\mathcal{F}[a,b]} s(f,\tau) \le F(b) - F(a) \le \inf_{\tau\in\mathcal{F}[a,b]} S(f,\tau) = I^*(f).$$

Az $f \in R[a,b]$ (azaz az $I_*(f) = I^*(f)$) feltételünkből így az következik, hogy

$$F(b) - F(a) = I_*(f) = I^*(f) = \int_a^b f(x)\,dx.$$

Megjegyzés. A Newton–Leinbiz-formula feltételei közül egyik sem hagyható el. Belátható, hogy ezek egymástól függetlenek (egyikből sem következik a másik): Létezik ui. olyan integrálható függvény, amelynek nincs primitív függvénye (ilyen pl. a szignum függvény a $[-1,1]$ intervallumon). Jóval nehezebb annak a megmutatása, hogy van olyan nem integrálható függvény, amelynek van primitív függvénye.

Példa. Számítsuk ki a

$$\int_0^\pi \sin x\,dx$$

határozott integrált! Világos, hogy a $\sin x$ $\Big(x \in [0,\pi]\Big)$ függvényre teljesülnek a Newton–Leibniz-formula feltételei, és $F(x) = -\cos x$ $\Big(x \in [0,\pi]\Big)$ a sin függvény egy primitív függvénye $[0,\pi]$-n. Így

$$\int_0^\pi \sin x\,dx = \Big[-\cos x\Big]_0^\pi = \Big(-\cos\pi\Big) - \Big(-\cos 0\Big) = 1 - (-1) = 2.$$

Ezzel megkaptuk a $\sin_{|\,[0,\pi]}$, függvény grafikonja alatti síkidom területét.

A HATÁROZOTT INTEGRÁL NÉHÁNY ALKALMAZÁSA

Ebben a pontban a határozott integrál néhány geometriai alkalmazását mutatjuk be.

Síkidom területe

Emlékeztetünk arra, hogy ha a korlátos $f : [a, b] \to \mathbb{R}$ függvény Riemann-integrálható az $[a, b]$ intervallumon és $f(x) \geq 0$ $(x \in [a, b])$, akkor az f grafikonja alatti

$$A := \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b,\ 0 \leq y \leq f(x)\}$$

síkidom területét így *értelmeztük:*

$$T(A) := \int_a^b f(x)\,dx.$$

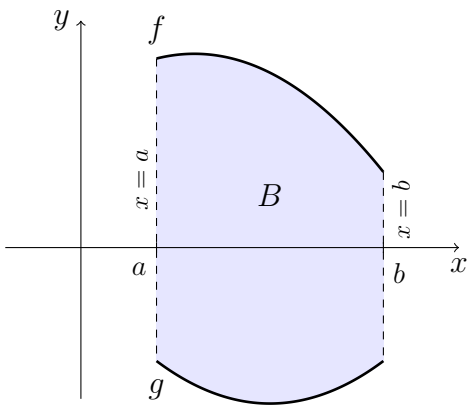
Két $f, g\colon [a, b] \to \mathbb{R}$ korlátos és Riemann-integrálható függvény esetében, ha $g(x) \leq f(x)$ minden $x \in [a, b]$ esetén, akkor a függvények az $x = a$ és $x = b$ egyenesekkel által közrezárt

$$B = \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b,\ g(x) \leq y \leq f(x)\}$$

síkidom területét a

$$\boxed{T(B) = \int_a^b \Big(f(x) - g(x)\Big)\,dx}$$

határozott integrállal célszerű értelmezni.



Ez könnyen látható, ha $g \geq 0$, hiszen ekkor az f függvény grafikonja alatti A_f síkidom tartalmazza a g függvény grafikonja alatti A_g síkidomot, azaz $A_g \subseteq A_f$, és így

$$T(B) = T(A_f \setminus A_g) = T(A_f) - T(A_g) = \int_a^b f(x)\,dx - \int_a^b g(x)\,dx = \int_a^b \Big(f(x) - g(x)\Big)\,dx.$$

Ha a $g \geq 0$ feltétel nem teljesül, akkor a függvény korlátossága miatt $\exists c > 0$ szám, hogy $f(x) + c \geq g(x) + c \geq 0$ minden $x \in [a, b]$ esetén. Ezzel feltoltuk a B síkidomot az x tengely fölé, amit az $f + c, g + c$ függvények az $x = a$ és $x = b$ egyenesekkel zárnak közre. Ezért területe

$$T(B) = \int_a^b f(x) + c\,dx - \int_a^b g(x) + c\,dx = \int_a^b \Big(f(x) + c - (g(x) + c)\Big)\,dx = \int_a^b \Big(f(x) - g(x)\Big)\,dx.$$

Példa (A kör területe). Helyezzük el az $R > 0$ sugarú körlapot a koordináta-rendszerben úgy, hogy az origó legyen a körlap középpontja. Ekkor a körvonal egyenlete $x^2 + y^2 = R^2$. Világos, hogy elég a felső félsíkba eső félkörlap területét meghatározni. A körvonal felső félsíkba eső része az

$$f(x) := \sqrt{R^2 - x^2} \quad \Big(x \in [-R, R]\Big)$$

függvény grafikonja.

Mivel az f függvény folytonos, ezért Riemann-integrálható az $[a, b]$ intervallumon, következésképpen a szóban forgó félkörlapnak *van* területe, és az egyenlő a következő határozott integrállal:

$$\int_{-R}^{R} \sqrt{R^2 - x^2}\, dx.$$

A Newton–Leibniz-tétel szerint először az integrandus egy primitív függvényét kell meghatározni. Ez létezik, mert a szóban forgó függvény folytonos.

Mivel

$$\int \sqrt{1 - x^2}\, dx = \frac{1}{2} \cdot \arcsin x + \frac{1}{2} \cdot x\sqrt{1 - x^2} + c \quad \Big(x \in (-1, 1)\Big),$$

ezért lineáris helyettesítéssel

$$\int \sqrt{R^2 - x^2}\, dx = R \int \sqrt{1 - \left(\frac{x}{R}\right)^2}\, dx = R \cdot \frac{1}{\frac{1}{R}} \left(\frac{1}{2} \arcsin \frac{x}{R} + \frac{1}{2} \cdot \frac{x}{R}\sqrt{1 - \left(\frac{x}{R}\right)^2}\right) + c =$$

$$= \frac{R^2}{2} \cdot \arcsin \frac{x}{R} + \frac{R\,x}{2}\sqrt{1 - \left(\frac{x}{R}\right)^2} + c \qquad \Big(x \in (-R, R)\Big).$$

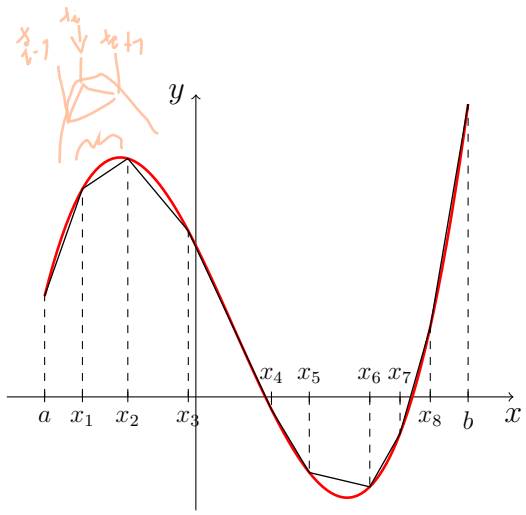
Így a Newton–Leibniz-tétel szerint

$$\int_{-R}^{R} \sqrt{R^2 - x^2}\, dx = \left[\frac{R^2}{2} \cdot \arcsin \frac{x}{R} + \frac{R\,x}{2}\sqrt{1 - \left(\frac{x}{R}\right)^2}\right]_{-R}^{R} =$$

$$= \frac{R^2}{2} \cdot \arcsin 1 - \frac{R^2}{2} \cdot \arcsin(-1) = \frac{R^2}{2} \cdot \frac{\pi}{2} - \frac{R^2}{2} \cdot \left(-\frac{\pi}{2}\right) = \frac{R^2\pi}{2},$$

vagyis az R sugarú félkörlap területe $R^2\pi/2$. Ezzel megkaptuk az R sugarú körlap területének ismert képletét: $R^2\pi$.

Síkbeli görbe ívhossza



Az ívhossz problémájánál a terület definíciója során megismert gondolatmenetet követjük. A görbét egy felosztáshoz tartozó törtvonallal közelítjük. A szemléletre hivatkozva azt várjuk, hogy egy „elég finom” beírt törtvonal annyira megközelíti a görbét, hogy a hosszúsága is közel lesz a görbe hosszához. Mindezekből azt szűrhetjük le, hogy a görbe ívhossza egyenlő a beírt törtvonal hosszainak a szuprémumával. Ezt a megállapítást fogjuk definícióként elfogadni.

A definíciót csak függvénygrafikonokra fogalmazzuk meg. Legyen valamilyen korlátos és zárt $[a, b]$ intervallum esetén adott az $f : [a, b] \to \mathbb{R}$ függvény. Emlékeztetünk arra, hogy a

$$\Gamma_f := \Big\{\Big(x, f(x)\Big) \mid x \in [a, b]\Big\}$$

síkbeli halmazt (görbét) neveztük f grafikonjának.

Tetszőleges $\tau = \{a = x_0 < x_1 < \cdots < x_n = b\} \in \mathcal{F}[a,b]$ felosztás esetén (alkalmas $n \in \mathbb{N}^+$ mellett) tekintsük az

$$\big(x_0, f(x_0)\big),\ \big(x_1, f(x_1)\big),\ \ldots,\ \big(x_n, f(x_n)\big)$$

pontokat összekötő szakaszokat; ezt nevezzük az f függvénygrafikon τ felosztáshoz tartozó ***beírt törtvonalának***. Világos, hogy ennek hossza a következő összeg:

$$\ell_f(\tau) = \sum_{i=0}^{n-1} \sqrt{\big(x_{i+1} - x_i\big)^2 + \big(f(x_{i+1}) - f(x_i)\big)^2}.$$

Az előzőek alapján elég természetes a következő definíció.

3. Definíció. *Legyen $a, b \in \mathbb{R}$, $a < b$ és $f : [a,b] \to \mathbb{R}$ adott függvény. Azt mondjuk, hogy a*

$$\Gamma_f := \Big\{\big(x, f(x)\big) \mid x \in [a,b]\Big\}$$

függvénygrafikon ***rektifikálható*** (*vagy más szóval* **van ívhossza**), *ha*

$$\ell\big(\Gamma_f\big) := \sup\Big\{\ell_f(\tau) \mid \tau \in \mathcal{F}[a,b]\Big\} < +\infty.$$

Ezt a valós számot a szóban forgó függvénygrafikon **ívhosszának** *nevezzük.*

A terület problémájához hasonlóan a következő kérdéseket is felvetjük:

1. Milyen Γ_f görbének van ívhossza?
2. Hogyan lehet ℓ_f-et kiszámolni?

A válaszok motiválásához az egyszerűség kedvéért tegyük fel azt, hogy az f *függvény folytonosan deriválható* az $[a,b]$ intervallumon (röviden $f \in C^1[a,b]$), és tekintsük az $\ell_f(\tau)$ összeg i-edik tagját:

$$\ell_i = \sqrt{\big(x_{i+1} - x_i\big)^2 + \big(f(x_{i+1}) - f(x_i)\big)^2} = (x_{i+1} - x_i)\sqrt{1 + \left(\frac{f(x_{i+1}) - f(x_i)}{x_{i+1} - x_i}\right)^2}.$$

A Lagrange-féle középértéktétel szerint van olyan $\xi_i \in (x_i, x_{i+1})$ pont, amelyre

$$f(x_{i+1}) - f(x_i) = f'(\xi_i)(x_{i+1} - x_i)$$

teljesül, ezért

$$\ell_i = (x_{i+1} - x_i)\sqrt{1 + [f'(\xi_i)]^2}.$$

Így az f függvény grafikonjához közel eső törtvonal hossza

$$\sum_{i=0}^{n-1} \ell_i = \sum_{i=0}^{n-1} \sqrt{1 + [f'(\xi_i)]^2} \cdot (x_{i+1} - x_i).$$

Vegyük észre, hogy a jobb oldal a $\varphi(x) := \sqrt{1 + [f'(x)]^2}$ $(x \in [a,b])$ függvény egy integrálközelítő összege. Várható tehát az, hogy a szóban forgó görbének van ívhossza, és az egyenlő az $\int_a^b \sqrt{1 + [f'(x)]^2}\, dx$ határozott integrállal.

A részletek mellőzésével itt csak azt jegyezzük meg, hogy az imént vázolt gondolatmenet precíz formában is „végigvihető”, ezért igaz a következő állítás.

4. Tétel. *Legyen $a, b \in \mathbb{R}$ $a < b$ és tegyük fel, hogy $f \in C^1[a,b]$. Ekkor az f függvény*

$$\Gamma_f := \Big\{\big(x, f(x)\big) \mid x \in [a,b]\Big\}$$

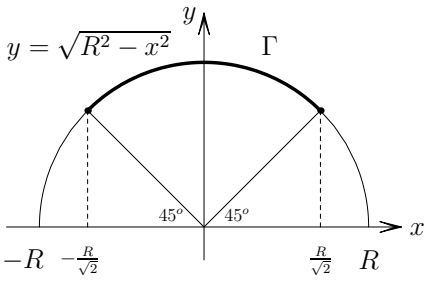
grafikonjának van ívhossza, és az egyenlő az

$$\ell\big(\Gamma_f\big) = \int_a^b \sqrt{1 + \big[f'(x)\big]^2}\, dx$$

határozott integrállal.

Példa (<u>A kör kerülete</u>). Az alábbi ábrán jelzett negyedkör ívhosszát fogjuk kiszámolni. Legyen

$$f(x) := \sqrt{R^2 - x^2} \qquad \Big(|x| \le \tfrac{R}{\sqrt{2}}\Big)$$



Világos, hogy $f \in C^1\Big[-\frac{R}{\sqrt{2}}, \frac{R}{\sqrt{2}}\Big]$ és

$$f'(x) = \frac{1}{2\sqrt{R^2 - x^2}} \cdot (-2x) \qquad \Big(|x| < R\Big)$$

Így

$$\sqrt{1 + \big[f'(x)\big]^2} = \sqrt{1 + \frac{x^2}{R^2 - x^2}} = \frac{R}{\sqrt{R^2 - x^2}} = \frac{1}{\sqrt{1 - \left(\frac{x}{R}\right)^2}}.$$

Az előző tétel szerint a Γ görbének van ívhossza és

$$\ell(\Gamma) = \int_{-R/\sqrt{2}}^{R/\sqrt{2}} \sqrt{1 + \big[f'(x)\big]^2}\, dx = \int_{-R/\sqrt{2}}^{R/\sqrt{2}} \frac{1}{\sqrt{1 - \left(\frac{x}{R}\right)^2}} = \left[R \arcsin \frac{x}{R}\right]_{-R/\sqrt{2}}^{R/\sqrt{2}} = \tag{1}$$

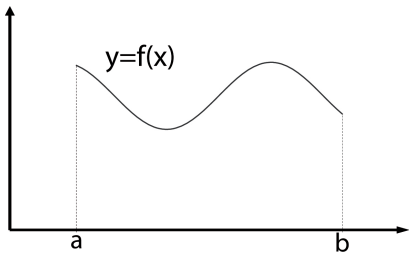
$$= R\left(\arcsin \frac{1}{\sqrt{2}} - \arcsin\left(-\frac{1}{\sqrt{2}}\right)\right) = R\left(\frac{\pi}{4} - \left(-\frac{\pi}{4}\right)\right) = R \cdot \frac{\pi}{2}. \tag{2}$$

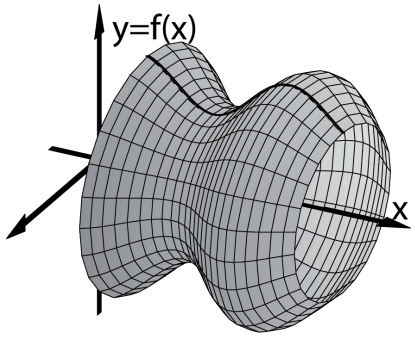
Az R sugarú kör kerülete tehát $4 \cdot R \cdot \dfrac{\pi}{2} = 2R\pi$.

Megjegyzés. A középiskolában a π számot az egységsugarú kör kerületének a felével értelmeztük. A 4. előadáson a (hatványsor összegfüggvényeként bevezetett) cos függvény első pozitív zérushelyének a kétszeresével definiáltuk a π számot. A fentiek alapján a két definíció ekvivalens. Ebből az is következik, hogy a középiskolában bevezetett sin, illetve cos függvény valóban egyenlő az Analízis I. tantárgyban definiált sin, illetve cos függvénnyel.

Forgástest térfogata

A Riemann-integrál eszköztárát a térfogat problémájának a vizsgálatánál is felhasználhatjuk. A továbbiakban csak *forgástesteket* (vagyis olyan térrészt amelyet egy függvénygrafikon alatti tartomány x tengely körüli megforgatásával kapunk) fogunk vizsgálni. A terület és az ívhossz problémájánál alkalmazott gondolatmenetet követjük: a forgástestet beírt és körülírt hengerekkel (ezek térfogatát ismertnek tekintjük) közelítjük.





Tekintsünk egy nemnegatív f függvényt a korlátos és zárt $[a,b]$ intervallumon. Ekkor az

$$A_f := \Big\{(x,y,z) \in \mathbb{R}^3 \mid a \le x \le b,\ y^2+z^2 \le f^2(x)\Big\}$$

halmazt az f függvény által meghatározott **forgástestnek** nevezzük.

Ha $n \in \mathbb{N}$ és $\tau = \Big\{a = x_0 < x_1 < \cdots < x_n = b\Big\}$ az $[a,b]$ intervallum egy felosztása, akkor legyen

$$m_i := \inf_{x\in[x_i,x_{i+1}]} f(x) \quad \text{és} \quad M_i := \sup_{x\in[x_i,x_{i+1}]} f(x) \qquad (i = 0,1,\dots,n-1).$$

A

$$h_i = \Big\{(x,y,z) \in \mathbb{R}^3 \mid x_i \le x \le x_{i+1},\ y^2+z^2 \le m_i^2\Big\}$$

beírt hengerekre, illetve a

$$H_i = \Big\{(x,y,z) \in \mathbb{R}^3 \mid x_i \le x \le x_{i+1},\ y^2+z^2 \le M_i^2\Big\}$$

körülírt hengerekre nyilvánvalóan fennáll, hogy

$$\bigcup_{i=0}^{n-1} h_i \subset A_f \subset \bigcup_{i=0}^{n-1} H_i.$$

Jelöljük $V(B)$-vel a $B \subset \mathbb{R}^3$ test térfogatát. Az r alapsugarú és m magasságú henger térfogata $r^2\pi \cdot m$, ezért

$$V\left(\bigcup_{i=0}^{n-1} h_i\right) = \pi \sum_{i=0}^{n-1} m_i^2 \cdot (x_{i+1}-x_i) = \pi s\Big(f^2,\tau\Big),$$

illetve

$$V\left(\bigcup_{i=0}^{n-1} H_i\right) = \pi \sum_{i=0}^{n-1} M_i^2 \cdot (x_{i+1}-x_i) = \pi S\Big(f^2,\tau\Big).$$

Ha tehát az A_f forgástestnek is akarunk a $V(A_f)$-fel jelölt térfogatot tulajdonítani, akkor fenn kell állnia az alábbi egyenlőtlenségeknek:

$$\pi \cdot s\Big(f^2,\tau\Big) \le V(A_f) \le \pi \cdot S\Big(f^2,\tau\Big) \quad (\tau \in \mathcal{F}[a,b])$$

Abban az esetben, ha f Riemann-integrálható az $[a,b]$ intervallumon, akkor $f^2 \in R[a,b]$ is teljesül, azaz

$$I_*(f^2) = \sup_{\tau\in\mathcal{F}[a,b]} s\Big(f^2,\tau\Big) = \inf_{\tau\in\mathcal{F}[a,b]} S\Big(f^2,\tau\Big) = \int_a^b f^2.$$

Az előzőek alapján kézenfekvő az alábbi értelmezés.

4. Definíció. *Legyen $0 \le f \in R[a,b]$. Ekkor f grafikonjának az x tengely körüli megforgatásával adódó*

$$A_f := \Big\{(x,y,z) \in \mathbb{R}^3 \mid a \le x \le b,\ y^2 + z^2 \le f^2(x)\Big\}$$

forgástestnek van térfogata, és az egyenlő a

$$\pi \int_a^b f^2(x)\,dx$$

integrállal.

Példa (A gömb térfogata). Tekintsük valamilyen $R > 0$ mellett az

$$f(x) := \sqrt{R^2 - x^2} \quad \Big(x \in [-R, R]\Big)$$

függvényt. Mivel f folytonos, ezért Riemann-integrálható $[-R, R]$-en. Az A_f forgástest egy R sugarú gömb. A fenti definíció szerint ennek van térfogata, és a Newton–Leibniz-formula felhasználásával a térfogata

$$\pi \int_a^b f^2(x)\,dx = \pi \int_{-R}^{R} \Big(R^2 - x^2\Big)\,dx = \pi \left[R^2 x - \frac{x^3}{3}\right]_{-R}^{R} =$$

$$= \pi \left(\left(R^2 \cdot R - \frac{R^3}{3}\right) - \left(R^2 \cdot (-R) - \frac{(-R)^3}{3}\right)\right) = \pi \left(\frac{2R^3}{3} - \left(-\frac{2R^3}{3}\right)\right) = \frac{4R^3\pi}{3}.$$

Az R sugarú gömb térfogata tehát $4R^3\pi/3$, és ez megegyezik a korábbi tanulmányokban gyakran használt képlettel.

Forgástest felszíne

Felületek felszínének a problémája (még forgásfelület esetén is) jóval bonyolultabb, mint a terület vagy a térfogat problémája. A továbbiakban azonban az alkalmazások szempontjából jól használható eredményt fogunk ismertetni.

Legyen f a korlátos és zárt $[a,b]$ intervallumon értelmezett nemnegatív függvény. Jelöljük $\mathcal{A}_f$-fel f grafikonjának az x tengely körüli megfogatásával kapott halmazt:

$$\mathcal{A}_f := \Big\{(x,y,z) \in \mathbb{R}^3 \mid a \le x \le b,\ y^2 + z^2 = f^2(x)\ (y, z \in \mathbb{R})\Big\},$$

amit az f függvény által meghatározott *forgásfelületnek* nevezünk.

Kézenfekvő az a feltételezés, hogy $\mathcal{A}_f$ felszínét megközelítik az f grafikonjába beírt törtvonalak megforgatásával kapott halmazok (ezek *csonkakúp palástok* egyesítései) felszínei.

Tekintsük az $[a,b]$ intervallum egy τ felosztását, és jelöljük Φ_τ-val a szóban forgó csonkakúp palástok felszíneinek (ezt ismertnek tekintjük) az összegét. Mivel f grafikonjának ívhossza (ha létezik) egyenlő kell hogy legyen a beírt törtvonalak ívhosszai halmazának a szuprémumával, ezért első gondolatunk az lehetne, hogy az $\mathcal{A}_f$ halmaz felszíne egyenlő kell hogy legyen a Φ_τ értékek ($\tau \in \mathcal{F}[a,b]$) szuprémumával. Ez azonban már egészen egyszerű függvények esetében sem igaz. Tekintsük például az $|x|$ függvényt a $[-1,1]$ intervallumon. Ekkor $\mathcal{A}_f$ két egybevágó kúppalást egyesítése, ezért a felszíne $2 \cdot (2\,\pi \cdot \sqrt{2}/2) = 2\sqrt{2}\,\pi$. Ha azonban a τ felosztás csupán a -1 és 1 osztópontokból áll, akkor $\Phi_\tau = 2\,\pi \cdot 2 = 4\,\pi$, ami nagyobb, mint $\mathcal{A}_f$ felszíne.

Ez a példa a helyes definíciót is sugallja. A technikai nehézségeket elkerülendő azt az egyszerűbb utat választjuk, hogy az imént jelzett okoskodás végeredményeképpen kapott integrállal definiáljuk a felszínt.

5. Definíció. *Legyen $a, b \in \mathbb{R}$, $a < b$ és tegyük fel, hogy $0 \le f \in C^1[a,b]$. Ekkor f grafikonjának az x-tengely körüli forgatásával adódó*

$$\mathcal{A}_f = \Big\{(x,y,z) \in \mathbb{R}^3 \mid a \le x \le b, \quad y^2 + z^2 = f^2(x) \ (y,z \in \mathbb{R})\Big\}$$

forgásfelületnek van felszíne, és értéke

$$2\,\pi \int\limits_a^b f(x) \cdot \sqrt{1 + \Big[f'(x)\Big]^2}\, dx.$$

Példa (A gömb felszíne). Az origó középpontú és R sugarú gömbfelületet az

$$f(x) = \sqrt{R^2 - x^2} \quad \Big(x \in [-R, R]\Big)$$

függvény grafikonjának az x tengely körüli megforgatásával kapjuk. Legyen $0 < r < R$ és tekintsük az

$$f_r(x) = \sqrt{R^2 - x^2} \quad \Big(x \in [-r, r]\Big).$$

Ekkor $f_r \in C^1[-r, r]$ és

$$f_r'(x) = -\frac{x}{\sqrt{R^2 - x^2}} \quad \Big(x \in [-r, r]\Big),$$

továbbá

$$1 + \Big[f'(x)\Big]^2 = 1 + \frac{x^2}{R^2 - x^2} = \frac{R^2}{R^2 - x^2} \quad \Big(x \in [-r, r]\Big).$$

Ezért az f_r által a fentiekben meghatározott és $\mathcal{A}_r$-rel jelölt forgásfelület felszíne

$$F_r := 2\,\pi \int\limits_{-r}^{r} \sqrt{R^2 - x^2} \cdot \sqrt{\frac{R^2}{R^2 - x^2}}\, dx = 2\,\pi\, R \int\limits_{-r}^{r} 1\, dx = 4\, R\, r\, \pi.$$

A „szemlélet alapján” könnyen elfogadható (az egzakt meggondolásokat nem részletezve), hogy az $r \to R$ határátmenettel $\mathcal{A}_r \to \mathcal{A}_R$, ahol $\mathcal{A}_R$ az R sugarú gömb felszíne. Ugyanakkor

$$\lim_{r \to R} F_r = \lim_{r \to R} 4\, R\, r\, \pi = 4\, R^2\, \pi,$$

ami valóban nem más, mint az R sugarú gömb felszíne.

Improprius integrálok

A Riemann-integrál értelmezésénél a kiindulópontunk az volt, hogy csak olyan f függvényeket tekintettünk, amelyekre a következő két feltétel teljesül:

a) f értelmezési tartománya egy ***korlátos és zárt*** $[\boldsymbol{a}, \boldsymbol{b}]$ ***intervallum***,

b) az $\boldsymbol{f}$ ***függvény korlátos*** $[a, b]$-n.

Felvethető az a ***probléma***, hogy ezeket a feltételeket nem kielégítő függvényekre vajon értelmezhető-e az integrál fogalma. Egyfajta kiterjesztést tesznek lehetővé az ún. **improprius integrálok**.

6. Definíció. *Legyen* $-\infty \leq a < b < +\infty$ *és* $f\colon (a,b] \to \mathbb{R}$. *Tegyük fel, hogy* $f \in R[x,b]$ *minden* $x \in (a,b)$ *esetén. Vezessük be a*

$$G(x) := \int_x^b f(t)\,dt \qquad \big(x \in (a,b)\big)$$

függvényt. Azt mondjuk, hogy az f *függvény* **impropriusan integrálható**, *ha* $\exists \lim\limits_a G \in \mathbb{R}$ *véges határérték. Ekkor az*

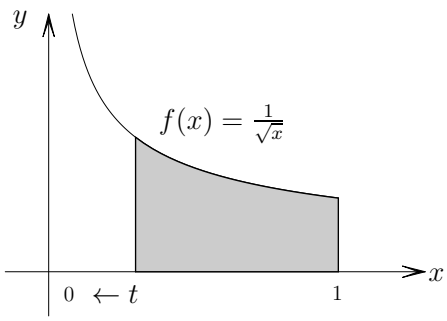
$$\int\limits_a^b f := \lim_{x\to a} G(x)$$

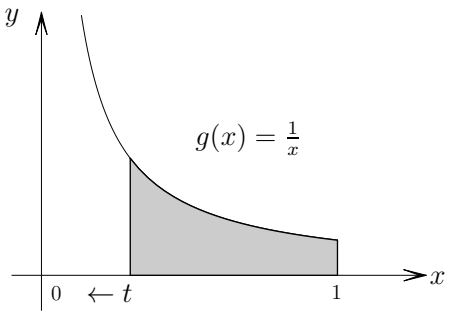
számot az f **improprius integráljának** *nevezzük.*

Példa. Legyen

$$f(x) := \frac{1}{\sqrt{x}} \quad \big(x \in (0,1]\big) \qquad \text{és} \qquad g(x) := \frac{1}{x} \quad \big(x \in (0,1]\big).$$

Tekintsük a következő ábrákat:





Ekkor

$$\int\limits_0^1 \frac{1}{\sqrt{x}}\,dx = \lim_{t\to 0+0} \int\limits_t^1 \frac{1}{\sqrt{x}}\,dx = \lim_{t\to 0+0} \Big[2\sqrt{x}\Big]_t^1 = \lim_{t\to 0+0} \Big(2 - 2\sqrt{t}\Big) = 2,$$

$$\int\limits_0^1 \frac{1}{x}\,dx = \lim_{t\to 0+0} \int\limits_t^1 \frac{1}{x}\,dx = \lim_{t\to 0+0} \Big[\ln x\Big]_t^1 = \lim_{t\to 0+0} \Big(\ln 1 - \ln t\Big) = +\infty.$$

Megjegyzések.

1. Vigyázat! Az improprius integrál jelöléseiből nem derül ki, hogy nem a szokásos határozott integrálról van szó.

2. Nem nehéz meggondolni, hogy ha $f \in R[a,b]$ akkor az improprius integrál megegyezik a szokásos határozott integrállal. Ennek oka, hogy ekkor az integrálfüggvény tulajdonságai szerint G folytonos lesz az a pontban.

3. Az improprius integrállal bizonyos nem korlátos síkidomok területét is értelmezhetjük.

Analóg módon értelmezhető $-\infty < a < b \leq < +\infty$ esetén az $f\colon [a,b) \to \mathbb{R}$ függvény improprius integrálja az

$$\int\limits_a^b f := \lim_{x\to b} G(x), \qquad G(x) := \int_a^x f(t)\,dt \qquad \big(x \in (a,b)\big)$$

összefüggéssel.

Példa. Legyen

$$f(x) := \frac{1}{x^2} \qquad \Big(x \in [1, +\infty)\Big).$$

Ekkor

$$\int_1^{+\infty} \frac{1}{x^2}\,dx = \lim_{t\to+\infty} \int_1^t \frac{1}{x^2}\,dx = \lim_{t\to+\infty} \left[-\frac{1}{x}\right]_1^t = \lim_{t\to+\infty} \left(-\frac{1}{t} - (-1)\right) = 1.$$

Még egy harmadik eset marad, ahol szükséges értelmezni az improprius integrált.

7. Definíció. *Legyen $-\infty \leq a < b \leq +\infty$ és $f\colon (a,b) \to \mathbb{R}$. Tegyük fel, hogy $f \in R[x,y]$ minden $a < x < y < b$ esetén. Azt mondjuk, hogy az f függvény* **impropriusan integrálható***, ha minden $c \in (a,b)$ esetén $f_{|(a,c]}$ és $f_{|[c,b)]}$ impropriusan integrálható. Ekkor*

$$\int_a^b f := \int_a^c f + \int_c^b f.$$

Megjegyzés. Nem nehéz meggondolni, hogy a c értéke nem befolyásolja az $\int_a^b f$ eredményét.

Példa. Legyen

$$f(x) := \frac{1}{1+x^2} \qquad \Big(x \in (-\infty, +\infty)\Big).$$

Igaz, hogy

$$\int_0^{+\infty} \frac{1}{1+x^2}\,dx = \lim_{t\to+\infty} \int_0^t \frac{1}{1+x^2}\,dx = \lim_{t\to+\infty} \Big[\operatorname{arc\,tg} x\Big]_0^t = \lim_{t\to+\infty} (\operatorname{arc\,tg} t - 0) = \frac{\pi}{2}.$$

Hasonlóan

$$\int_{-\infty}^0 \frac{1}{1+x^2}\,dx = \lim_{t\to-\infty} \int_t^0 \frac{1}{1+x^2}\,dx = \lim_{t\to-\infty} \Big[\operatorname{arc\,tg} x\Big]_t^0 = \lim_{t\to-\infty} (0 - \operatorname{arc\,tg} t) = \frac{\pi}{2}.$$

Ezért

$$\int_{-\infty}^{+\infty} \frac{1}{1+x^2}\,dx = \int_{-\infty}^0 \frac{1}{1+x^2}\,dx + \int_0^{+\infty} \frac{1}{1+x^2}\,dx = \frac{\pi}{2} + \frac{\pi}{2} = \pi.$$

Az egyik legfontosabb eredmény az improprius integrálok körében az, hogy

$$\int_{-\infty}^{+\infty} e^{-x^2}\,dx = \sqrt{\pi},$$

ami a valószínűségszámításban játszik fontos szerepet. Mivel $\int e^{-x^2}\,dx$ nem elemi függvényekből áll, így nem tudjuk a Newton–Leibniz-formulával kiszámítani a fenti improprius integrál értékét. Egy későbbi előadáson, teljesen más eszközökkel megmutatjuk ennek az állításnak a bizonyítását.

8. előadás

TÖBBVÁLTOZÓS FÜGGVÉNYEK.

Eddigi tanulmányaink során ***egyváltozós analízissel***, vagyis $\mathbb{R} \to \mathbb{R}$ típusú (vagy másképpen fogalmazva valós-valós) függvényekkel foglalkoztunk. Láttuk, hogy az alapvető fogalmak a szóban forgó függvényeknek a *határértéke, folytonossága, deriváltja és integrálja.* A továbbiakban a ***többváltozós analízis*** alapjaival fogunk megismerkedni. Az egyváltozós analízis alapvető fogalmainak és eredményeinek az $\mathbb{R}^n \to \mathbb{R}^m$ $(1 \leq n, m \in \mathbb{N})$ típusú (az ún. vektor-vektor) függvényekre való kiterjesztéséről lesz szó.

De mi az alapja ennek a kiterjesztésnek? Elevenítsük fel újra a *függvény pontbeli határértékének* fogalmát! Legyen $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \mathcal{D}'_f$. Azt mondjuk, hogy az f függvénynek az a pontban vett határértéke az A szám, ha

$$\forall \varepsilon > 0\text{-hoz } \exists \delta > 0, \forall x \in \mathcal{D}_f,\ 0 < |x - a| < \delta \colon |f(x) - A| < \varepsilon.$$

Ugyanez ***környezetekkel*** kifejezve:

$$\forall \varepsilon > 0\text{-hoz } \exists \delta > 0, \forall x \in \Big(K_\delta(a) \setminus \{a\}\Big) \cap \mathcal{D}_f \colon f(x) \in K_\varepsilon(A),$$

ahol

$$K_r(x) := \{y \in \mathbb{R} \colon |x - y| < r\} = (x - r, x + r)$$

az x valós szám $r > 0$ sugarú környezetét jelenti. Látható, hogy a fenti fogalom (még az $a \in \mathcal{D}'_f$ torlódási pont fogalma is) teljesen leírható környezetekkel. Ez azt jelenti, hogy ha bevezetnénk a környezet fogalmát $\mathbb{R}^n$-ben, akkor változtatás nélkül általánosítani tudnánk a függvény pontbeli határértékének fogalmát.

Metrikus és normált terek

A környezet fogalma nem lehet akármilyen, mert akkor nem fogjuk tudni a határértéktől „elvárt” tulajdonságokat igazolni. A valós számok halmazán a $K_r(x)$ környezet nem más, mint azon pontok halmaza, amelyeknek távolsága az x ponttól kisebb, mint r. Ha ezen az úton maradunk, akkor csak egy *távolságfüggvényre* vagy más néven *metrikára* lesz szükségünk.

Mit jelent az, hogy metrika? Gondolhatunk arra, hogy a valós térben, ahol élünk, két pont távolsága mindig a két pontot összekötő egyenes szakasz hossza. Sajnos nem mindig tudunk egyenes úton eljutni az egyik ponttól a másikig, ezért sokszor szükséges egy ettől eltérő metrikát értelmezni. A matematikai analízis különböző metrikákat enged alkalmazni, azonban ezekkel szemben megkövetel néhány tulajdonságot.

Legyen $M \neq \emptyset$ egy adott halmaz és $d : M \times M \to \mathbb{R}$ olyan függvény, amelyre teljesül

$$\begin{array}{lll} i. & d(x,y) \geq 0 \quad \text{és} \quad d(x,y) = 0 \iff x = y, & \\ ii. & d(x,y) = d(y,x) & \text{(szimmetrikus)}, \\ iii. & d(x,y) \leq d(x,z) + d(z,y) & \text{(háromszög egyenlőtlenség)} \end{array}$$

minden $x, y, z \in M$ esetén. Ekkor az (M, d) együttest ***metrikus térnek*** nevezzük. Ugyanakkor azt mondjuk, hogy d ***metrika*** vagy ***távolságfüggvény*** M-en.

A $d(x,y) := |x-y|$ függvény metrika $\mathbb{R}$-en, és ***természetes távolságnak*** fogjuk nevezni, de ettől lényegesen eltérő metrikák is értelmeztetők. Pl. igazolható, hogy a

$$d_1(x,y) := \begin{cases} 0 & \text{ha } x = y, \\ 1 & \text{ha } x \neq y \end{cases}$$

függvény metrika (ún. diszkrét metrika) $\mathbb{R}$-en, illetve a

$$d_2(x,y) := \frac{|x-y|}{1+|x-y|} \qquad \text{és} \qquad d_3(x,y) := |e^x - e^y|$$

függvények szintén metrikák $\mathbb{R}$-en.

Ha az (M,d) metrikus tér egyben lineáris tér (vektortér), akkor minden x elem (vektor) nagyságát (hosszát) úgy értelmezzük, mint az elem nullától való távolságát. Erre a $\|x\|$ jelölést alkalmazzuk, azaz $\|x\| := d(x,0)$. Tudjuk, hogy $\mathbb{R}$ egy 1 dimenziós vektortérnek tekinthető, így ha d a természetes metrika, akkor $\|x\| = |x|$. Csak hogy az abszolút értéknek a következő tulajdonságai vannak:

a) $|x| \geq 0, \qquad \text{és} \qquad |x| = 0 \iff x = 0,$

b) $|xy| = |x||y|,$

c) $|x+y| \leq |x| + |y|$

minden $x, y \in \mathbf{R}$ esetén. Ezeket az Analízis I. kurzuson igazoltuk, és számos állítás bizonyításában alkalmaztuk. Nem okoz meglepetést tehát a következő fogalom bevezetése.

Legyen $X \neq \emptyset$ egy lineáris tér $\mathbb{R}$-felett és $\|.\| : X \to \mathbb{R}$ olyan függvény (ún. ***norma***), amelyre teljesül

$$\begin{array}{lcl} i) & \|x\| \geq 0 \quad \text{és} \quad \|x\| = 0 \iff x = 0, & \\ ii) & \|\lambda x\| = |\lambda|\|x\| & \text{(abszolút homogén)}, \\ iii) & \|x+y\| \leq \|x\| + \|y\| & \text{(szubadditív)} \end{array}$$

minden $x, y \in X$ és $\lambda \in \mathbb{R}$ esetén. Ekkor az $(X, \|.\|)$ együttest ***normált térnek*** nevezzük. Ugyanakkor azt mondjuk, hogy $\boldsymbol{\|x\|}$ ***az*** $\boldsymbol{x \in X}$ ***elem normája***. Könnyen igazolható, hogy (X,d) metrikus tér, ha a d távolságot a norma segítségével értelmezzük:

$$d(x,y) := \|x-y\| \qquad (x \in X).$$

Igazolható, hogy egy normából származó metrika abszolút homogén és eltolás invariáns, azaz

$$d(\lambda x, \lambda y) = |\lambda| d(x,y) \qquad \text{és} \qquad d(x+z, y+z) = d(x,y)$$

minden $x, y, z \in X$ és $\lambda \in \mathbb{R}$ esetén. Az abszolút homogenitás nem érvényes a fenti példákban szereplő d_1, d_2 és d_3 metrikákra, de nyilván érvényes a természetes metrikára. A normákra még igazolható az

$$\Big| \|x\| - \|y\| \Big| \leq \|x-y\| \qquad (x, y \in X)$$

egyenlőtlenséget.

A fentiek értelmében normákat kell keresünk $\mathbb{R}^n$-en, és rögtön hármat meg is tudunk adni:

$$\|x\|_2 := \left(\sum_{k=1}^n x_k^2 \right)^{1/2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \qquad (\textbf{\textit{euklideszi norma}}),$$

$$\|x\|_1 := \sum_{k=1}^n |x_k| = |x_1| + |x_2| + \cdots + |x_n|,$$

$$\|x\|_\infty := \max\{|x_k| \colon k = 1, 2, \ldots, n\} = \max\{|x_1|, |x_2|, \ldots, |x_n|\},$$

ahol $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

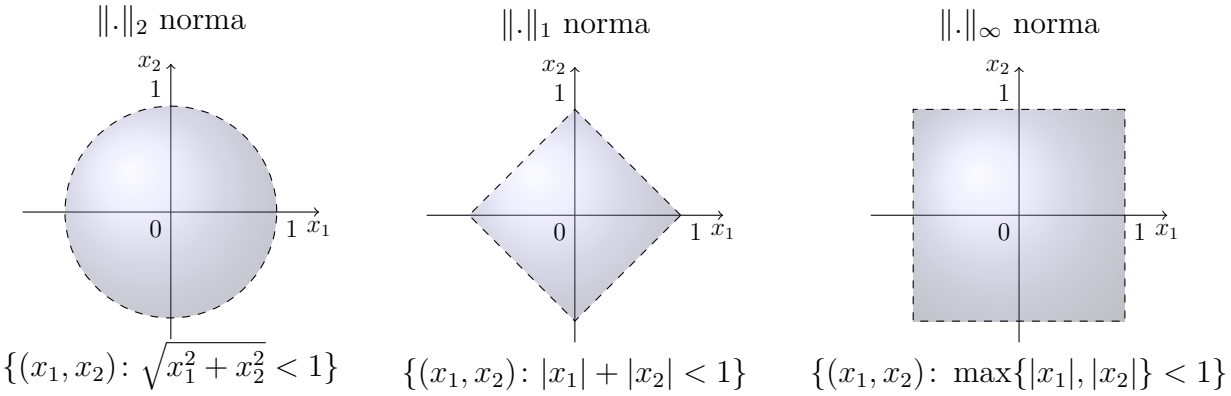
Mindhárom norma jó abban az értelemben, hogy visszaadják a szám abszolút értékét $n = 1$ esetén. Tehát mindhárom az abszolút érték általánosításának tekinthető. Akkor melyiket kellene használni? Ezek a normák egymással ekvivalensek, azaz bármelyik a másik konstansszorosával felülről becsülhető. Ez valóban így van, hiszen igazolható, hogy

$$\frac{1}{n}\,\|x\|_1 \le \|x\|_\infty \le \|x\|_2 \le \|x\|_1 \qquad (x \in \mathbb{R}^n).$$

De az is igaz, hogy véges dimenziós térben bármely két norma egymással ekvivalens. Ez azt jelenti hogy mindegy melyiket használunk, de mi az $\|x\|_2$-t fogjuk preferálni. A normaekvivalencia nem azt jelenti, hogy hasonló környezeteket kapunk. A következő ábra mutatja a

$$K_1(0) = \{x \in \mathbb{R}^2 \colon \|x - 0\| < 1\}$$

környezetet $\mathbb{R}^2$-ben $\big(0 = (0,0)\big)$.



$\mathbb{R}^n$ mint euklideszi tér

A `Matematikai alapok` tantárgyban az $\mathbb{R}^n$ tér számos tulajdonságáról volt szó. Most felsoroljuk azokat az ismerteket, amelyekre a továbbiakban szükségünk lesz.

Legyen $n \in \mathbb{N}^+$ egy adott pozitív természetes szám. Az $\mathbb{R}^n$ szimbólummal jelöljük a rendezett valós szám n-esek halmazát:

$$\mathbb{R}^n := \Big\{x = (x_1, x_2, \ldots, x_n) \colon x_k \in \mathbb{R},\ k = 1, 2, \ldots, n\Big\}.$$

Az $x_1, x_2, \ldots, x_n$ számokat az $x = (x_1, x_2, \ldots, x_n)$ pont (vektor) ***koordinátáinak*** vagy ***komponenseinek*** nevezzük.

$\mathbb{R}^1$-et azonosítjuk $\mathbb{R}$-rel. A sík pontjai rendezett valós számpárokkal (vagyis az $\mathbb{R}^2$ halmaz elemeivel), a tér pontjai pedig rendezett valós számhármasokkal (vagyis $\mathbb{R}^3$ elemeivel) azonosíthatók. Az $\mathbb{R}^n$ halmaz tehát ezek „természetes” általánosításaként fogható fel. Az $n > 3$ esetben $\mathbb{R}^n$-nek nincs szemléletes jelentése, de a fogalom mégis nélkülönözhetetlen mind az elmélet, mind pedig az alkalmazások szempontjából.

A középiskolában a sík és a tér vektoraival több műveletet is értelmeztünk. Vektorok ***összeadásának***, valamint **vektor** (valós) ***számmal való szorzásának*** a mintájára vezetjük be az $\mathbb{R}^n$ halmazon az alábbi komponensenkénti műveleteket: ha $x = (x_1, x_2, \ldots, x_n)$, $y = (y_1, y_2, \ldots, y_n)$ és $\lambda \in \mathbb{R}$, akkor

$$x + y := (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n), \qquad \lambda \cdot x := (\lambda\, x_1, \lambda\, x_2, \ldots, \lambda\, x_n).$$

Ezek az $\mathbb{R}^n$-beli műveletek rendelkeznek a sík és a tér vektorainak a középiskolában megismert 10 alapvető tulajdonságával. Röviden ezt úgy fejezzük ki, hogy $\mathbb{R}^n$ ***ezekkel a műveletekkel lineáris tér*** (vagy ***vektortér*** $\mathbb{R}$ ***felett***). Ennek a vektortérnek a dimenziója pontosan n, azaz rendelkezik egy n darab tagból álló bázissal.

(Kiemeljük azt fontos tényt is, hogy rögzített $n, m \in \mathbb{N}^+$ esetén az $n \times m$-es valós elemű mátrixok $\mathbb{R}^{n\times m}$ szimbólummal jelölt halmazában is értelmezzük az összeadás és a számmal való szorzás műveleteket, és $\mathbb{R}^{n\times m}$ ezekkel a műveletekkel $\mathbb{R}$ feletti lineáris tér.)

A középiskolában a sík és tér vektorainak az összeadásán és a számmal való szorzásán kívül megismerkedtünk még egy fontos művelettel, vektorok *skaláris szorzatával.* Ezt a fogalmat is fogjuk az $\mathbb{R}^n$ lineáris térre is kiterjeszteni: Az $x = (x_1, x_2, \dots, x_n)$, $y = (y_1, y_2, \dots, y_n)$ vektorok ***skaláris szorzatát*** az

$$\langle x, y\rangle := \sum_{k=1}^{n} x_k y_k = x_1y_1 + x_2y_2 + \cdots + x_ny_n$$

valós számmal definiáljuk. A skaláris szorzat rendelkezik az euklideszi tér fogalmában szereplő 5 axiómával, ezért $\mathbb{R}^n$ egy n-dimenziós euklideszi tér $\mathbb{R}$ felett.

A skaláris szorzat segítségével értelmezhetjük $\mathbb{R}^n$-beli vektorok szögét, merőlegességét, illetve a hosszát (normát) és a távolságot. Ezekre a geometriában megszokott tulajdonságok jelentős része megmarad. Az $x = (x_1, x_2, \dots, x_n)$ vektor ***normáját*** (***hosszát*** vagy ***abszolút értékét***) az

$$\|x\| := \|x\|_2 := \sqrt{x_1^2 + x_2^2 + \cdot + x_n^2} = \sqrt{\langle x, x\rangle}$$

képlettel definiáljuk. Az $x, y \in \mathbb{R}^n$ vektorok ***távolságán*** az $\|x - y\|$ számot értjük. Ha a továbbiakban az $\mathbb{R}^n$ euklideszi térről beszélünk, akkor mindig az $\mathbb{R}^n$ lineáris térre és az azon értelmezett, a fenti skaláris szorzatból származó euklideszi normára gondolunk.

Egy $a \in \mathbb{R}^n$ pont r (> 0) sugarú ***környezetén*** a

$$K_r(a) := \Big\{x \in \mathbb{R}^n \colon \|x - a\| < r\Big\}$$

halmazt értjük. $n = 1$ esetén $K_r(a)$ az a pontra szimmetrikus $(a - r, a + r)$ nyílt intervallum. Ha $n = 2$, akkor $K_r(a)$ az a pont körüli r sugarú nyílt körlap, $n = 3$ esetén pedig az a pont körüli r sugarú nyílt gömb. A „nyílt gömb” elnevezést használjuk akkor is, ha $n > 3$.

A $\emptyset \neq A \subset \mathbb{R}^n$ halmazt ***korlátosnak*** nevezzük, ha $\exists\, r > 0$ úgy, hogy $A \subseteq K_r(0)$, vagyis A benne van egy 0 középpontú, alkalmas sugarú nyílt gömbben.

Környezetek segítségével (hasonlóan mint $\mathbb{R}$-ben) értelmezhetjük $\mathbb{R}^n$-ben is a következő „topológiai” fogalmakat. Tegyük fel, hogy A az $\mathbb{R}^n$ euklideszi térnek egy nem üres részhalmaza. Ekkor

- $a \in \mathbb{R}^n$ az A halmaz ***torlódási pontja*** (jelekkel $a \in A'$), ha $\forall K(a) \colon K(a) \cap A$ végtelen halmaz, azaz az a pont minden környezete végtelen sok A-beli pontot tartalmaz,
- $a \in A$ az A halmaz ***belső pontja*** (jelekkel $a \in \operatorname{int} A$), ha $\exists K(a) \colon K(a) \subset A$,
- az A halmaz ***nyílt halmaz***, ha minden pontja belső pont,
- az A halmaz ***zárt halmaz***, ha $\mathbb{R}^n \setminus A$ nyílt halmaz.

Látható, hogy a környezettel kapcsolatos fogalmak és jelölés módja nem változtak a már ismert $\mathbb{R}$-beli fogalmakhoz és jelöléhez képes, de tulajdonságai különbözhetnek. A nyílt és zárt halmazok struktúrája jóval gazdagabb. Pl. $\mathbb{R}$-ben egy nyílt halmaz mindig előáll megszámlálhatóan sok nyílt intervallum uniójaként.

Többdimenziós térben nem fogunk rendezést értelmezni, így nem beszélhetünk alsó, felső korlátokról, maximum, minimumról, ill. szuprémum-, infimumról. A teret nem fogjuk bővíteni olyan ideális elemekkel, mint a $+\infty$ és a $-\infty$ szimbólumokkal tettük a valós számok halmazán.

Konvergencia az $\mathbb{R}^n$ euklideszi térben

1. Definíció. *Az $x : \mathbb{N} \to \mathbb{R}^n$ függvényt* **$\mathbb{R}^n$-beli sorozatnak** *nevezzük. Az*

$$x(k) =: x_k \quad (k \in \mathbb{N})$$

helyettesítési érték a sorozat ***k*-adik** *vagy* ***k*-indexű tagja***, a tag sorszámát jelző szám a tag* **indexe***. Lehetséges jelölései: x, (x_k) vagy $(x_0, x_1, x_2, \dots)$.*

Mivel egy $\mathbb{R}^n$-beli pont koordinátainak jelölésére szintén alsó indexet használunk, így a félreértések elkerülésére az (x_k) sorozat k-adik tagjának i-edik koordinátájára az $x_k^{(i)}$ jelölést alkalmazzuk. Adott (x_k) $\mathbb{R}^n$-beli sorozat és fix $i = 1, 2, \dots, n$ esetén beszélhetünk az $(x_k^{(i)})$ koordinátasorozatról, ami már valós sorozat lesz. A koordinátasorozatok fontos szerepet játszanak az $\mathbb{R}^n$-beli sorozatok vizsgálatában.

Azt mondjuk, hogy az (x_k) $\mathbb{R}^n$-beli sorozat ***korlátos***, ha a sorozat értékkészlete korlátos, azaz ha az

$$\mathcal{R}_x = \{x_k \in \mathbb{R}^n : k \in \mathbb{N}\}$$

halmaz korlátos. Rendezés híján ***egy $\mathbb{R}^n$-beli sorozat monotonitása nem értelmezhető***, de a koordinátasorozatok esetében van értelme a monotonitásnak.

Emlékeztetünk arra, hogy az $(x_k) : \mathbb{N} \to \mathbb{R}$ valós sorozatot akkor neveztük *konvergensnek*, ha

$$\exists A \in \mathbb{R} \text{ úgy, hogy } \forall \varepsilon > 0 \text{ számhoz } \exists k_0 \in \mathbb{N},\ \forall\, k > k_0 \text{ indexre } |x_k - A| < \varepsilon.$$

Látható, hogy a fogalom lényegében az $\mathbb{R}$-en értelmezett $d(x, y) = |x - y|$ természetes távolságon múlik. Ha ehelyett az $\mathbb{R}^n$ euklideszi téren értelmezett $d(x, y) = \|x - y\|$ távolságfüggvényt használjuk, akkor általánosíthatjuk a sorozatok konvergenciájának fogalmát $\mathbb{R}^n$-re.

2. Definíció. *Legyen $1 \le n \in \mathbb{N}$. Azt mondjuk, hogy az $\mathbb{R}^n$ euklideszi tér $(x_k) : \mathbb{N} \to \mathbb{R}^n$ sorozata* **konvergens***, ha*

$$\exists A \in \mathbb{R}^n \text{ úgy, hogy } \forall \varepsilon > 0 \text{ számhoz } \exists k_0 \in \mathbb{N},\ \forall k > k_0 \text{ indexre } \|x_k - A\| < \varepsilon.$$

Ha A létezik, akkor az egyértelmű, és A-t az (x_k) sorozat **határértékének** *nevezzük, és az alábbi szimbólumok valamelyikével jelöljük:*

$$\lim (x_k) = A, \qquad \lim_{k \to +\infty} x_k = A, \qquad x_k \to A, \text{ ha } k \to +\infty.$$

Az (x_k) sorozat **divergens***, ha nem konvergens.*

Figyeljük meg, hogy az (x_k) vektorsorozat pontosan akkor tart az A vektorhoz, ha az $\|x_k - A\|$ $(k \in \mathbb{N})$ normák sorozata $\mathbb{R}$-beli nullsorozat, azaz

$$\lim_{k \to +\infty} x_k = A \quad \Longleftrightarrow \quad \lim_{k \to +\infty} \|x_k - A\| = 0.$$

A következő tétel szerint egy vektorsorozat konvergenciája ekvivalens a koordináták sorozatainak a konvergenciájával.

1. Tétel. *Legyen $n \in \mathbb{N}^+$. Egy $\mathbb{R}^n$-beli sorozat akkor és csak akkor konvergens, ha a sorozat minden koordinátasorozata konvergens, és a határértéke a határvektor megfelelő koordinátája, azaz*

$$\mathbb{R}^n \ni x_k = \left(x_k^{(1)}, x_k^{(2)}, \dots, x_k^{(n)}\right) \to A = \left(A^{(1)}, A^{(2)}, \dots, A^{(n)}\right), \quad ha \quad k \to +\infty$$

pontosan akkor igaz, ha minden $i = 1, 2, \dots, n$ koordinátára

$$x_k^{(i)} \to A^{(i)}, \quad ha \quad k \to +\infty.$$

Bizonyítás. $\boxed{\Rightarrow}$ Tegyük fel, hogy $\lim\limits_{k\to+\infty} x_k = A$, azaz $\lim\limits_{k\to+\infty} \|x_k - A\| = 0$. Rögzítsük az $i = 1, 2, \dots, n$ indexet. Mivel

$$0 \le \left|x_k^{(i)} - A^{(i)}\right| \le \sqrt{\sum_{j=1}^{n} \left|x_k^{(j)} - A^{(j)}\right|^2} = \|x_k - A\| \to 0, \text{ ha } k \to +\infty,$$

ezért a közrefogási elv szerint $\lim\limits_{k\to+\infty} \left|x_k^{(i)} - A^{(i)}\right| = 0$, azaz $\lim\limits_{k\to+\infty} x_k^{(i)} = A^{(i)}$.

$\boxed{\Leftarrow}$ Tegyük fel, hogy minden $i = 1, 2, \dots, n$ indexre $\lim\limits_{k\to+\infty} x_k^{(i)} = A^{(i)}$, azaz $\lim\limits_{k\to+\infty} |x_k^{(i)} - A^{(i)}| = 0$. Ekkor az

$$0 \le \|x_k - A\| = \|x_k - A\|_2 \le \|x_k - A\|_1 = \sum_{i=1}^{n} \left|x_k^{(i)} - A^{(i)}\right| \to 0, \text{ ha } k \to +\infty,$$

egyenlőtlenség és ismét a közrefogási elv alkalmazásával azt kapjuk, hogy $\|x_k - A\| \to 0$, ha $k \to +\infty$, azaz $\lim\limits_{k\to+\infty} x_k = A$.

Példák:

- $\lim\limits_{k\to+\infty} \left(\frac{1}{k}, \left(1 + \frac{1}{k}\right)^k\right) = (0, e)$, hiszen $\lim\limits_{k\to+\infty} \frac{1}{k} = 0$ és $\lim\limits_{k\to+\infty} \left(1 + \frac{1}{k}\right)^k = e$.
- Az $x_k := \left(\frac{1}{k^2}, \frac{\sin k}{k}, k\right) \quad (k \in \mathbb{N}^+)$ sorozat divergens, mert $\lim\limits_{k\to+\infty} k = +\infty$.

A tétel segítségével a legtöbb számsorozatra vonatkozó állítást általánosíthatjuk $\mathbb{R}^n$-beli sorozatokra. A bizonyítás többnyire abból áll, hogy a koordináták sorozataira alkalmazzuk a megfelelő számsorozatokra vonatkozó tételt. Ezért $\mathbb{R}^n$-beli sorozatokra is igaz a határérték egyértelműségére vonatkozó tétel, az összegsorozat és a számszoros sorozat határértékére vonatkozó tétel, illetve a konvergens sorozat részsorozataira vonatkozó tétel.

A következő két állításban azt fogalmazzuk meg, hogy az $\mathbb{R}$-beli sorozatok konvergenciájára vonatkozó alapvető jelentőségű tételek az $\mathbb{R}^n$ euklideszi térben is érvényesek.

2. Tétel (Cauchy-féle konvergenciakritérium). *Legyen $n \in \mathbb{N}^+$. Az $\mathbb{R}^n$ euklideszi tér (x_k) sorozata akkor és csak akkor konvergens, ha (x_k) Cauchy-sorozat, azaz*

$$\forall \varepsilon > 0 \text{ számhoz } \exists k_0 \in \mathbb{N}, \text{ hogy } \forall k, l > k_0 \text{ indexre } \|x_k - x_l\| < \varepsilon.$$

3. Tétel (Bolzano–Weierstrass-féle kiválasztási tétel). *Az $\mathbb{R}^n$ $(n \in \mathbb{N}^+)$ euklideszi térben minden korlátos sorozatnak van konvergens részsorozata.*

Az $\mathbb{R}^n \to \mathbb{R}^m$ típusú függvények

Olyan függvényekkel fogunk foglalkozni, amelyeknek értelmezési tartományuk része az $\mathbb{R}^n$ halmaznak, és értékkészletük része az $\mathbb{R}^m$ halmaznak, ahol n és m pozitív egész számok. Tehát

$$f\colon \mathbb{R}^n \ni (x_1, x_2, \ldots, x_n) \mapsto (y_1, y_2, \ldots, y_m) \in \mathbb{R}^m.$$

Ha $n = 1$ vagy $m = 1$, akkor ezek a függvények leegyszerűsödnek, és speciális értelmezésekkel kerülünk szembe.

$\boxed{\mathbf{1.}}$ $f \in \mathbb{R} \to \mathbb{R}$ ✓ A valós-valós függvényekkel már részletesen foglalkoztunk.

$\boxed{\mathbf{2.}}$ $f \in \mathbb{R} \to \mathbb{R}^m$ $(m > 1)$. Ekkor ***valós (egy)változós vektor értékű függvényekről*** beszélünk. Az ilyen függvények felírhatók

$$f(t) = \Big(x_1(t), x_2(t), \ldots, x_m(t)\Big) \qquad (t \in \mathcal{D}_f \subseteq \mathbb{R})$$

alakban, ahol az $x_i\colon \mathcal{D}_f \to \mathbb{R}$ valós-valós függvényeket ***koordinátafüggvényeknek*** nevezzük $(i = 1, 2, \ldots, m)$. Ha $m = 2$, akkor úgy tudjuk szemléltetni egy ilyen függvény értékészletét, hogy a koordinátasíkon ábrázoljuk az $\big(x_1(t), x_2(t)\big)$ koordinátájú pontokat, ahol $t \in \mathcal{D}_f$. Ha egy síkbeli görbe pontjait ilyen módon előállítjuk, akkor ***paraméteres görbéről*** beszélünk, ahol t a paraméter. Hasonlóan járunk el térbeli görbék esetén, ebben az esetben $m = 3$.

Matlab programmal a következő egyszerű kóddal tudunk paraméteres görbéket előállítani:

Síkbeli görbék

```
syms t
x1 = ...;              %x1(t) fv.
x2 = ...;              %x2(t) fv.
fplot(x1,x2,[a b])     %a<=t<=b
```

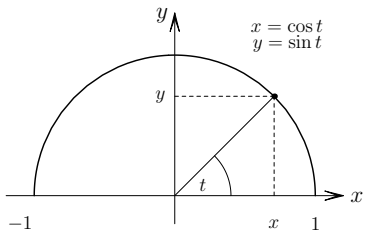
Térbeli görbék

```
syms t
x1 = ...;                  %x1(t) fv.
x2 = ...;                  %x2(t) fv.
x3 = ...;                  %x3(t) fv.
fplot3(x1,x2,x3,[a b])     %a<=t<=b
```

Példák:

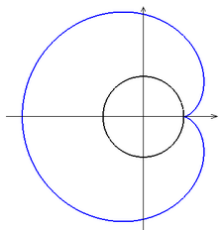
a) ***a félkörív:***

$$f(t) := (\cos t,\ \sin t) \quad \big(t \in [0, \pi]\big)$$



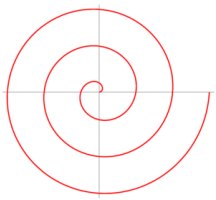
b) ***a kardiodid:***

$$f(t) := (2\cos t - \cos 2t,\ 2\sin t - \sin 2t)$$
$$\big(t \in [0, 2\pi]\big)$$



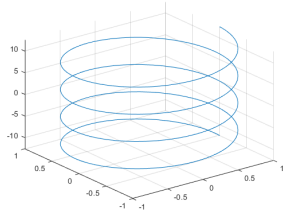
c) ***az arkhimédészi spirális:***

$$f(t) := (t\cos t,\ t\sin t) \quad (t \geq 0)$$



d) ***a hengerre írható csavarvonal:***
(a sugarú, m menetemelkedésű csavarvonal)

$$f(t) := \left(a\cos t,\ a\sin t,\ \frac{m}{2\pi}\,t\right) \qquad (t \in \mathbb{R})$$



Megjegyzések.

1. Itt hívjuk fel ismét a figyelmüket a `MacTutor` honlapra. Ezen – többek között – matematikusok (Arkhimédésztől napjainkig) életrajzáról és munkásságáról találhatnak részletes információkat. Ugyanezen az oldalon a „CURVES" menüpont alatt számos klasszikus görbe leírását találhatják meg.

2. A Néhány nevezetes síkgörbe című segédanyagban pedig bizonyos görbék származtatásáról olvashatnak.

$\boxed{\textbf{3.}}$ $f \in \mathbb{R}^n \to \mathbb{R}$ $(n > 1)$. Ekkor ***n változós valós értékű függvényekről*** beszélünk. Pl.

$$f \in \mathbb{R}^3 \to \mathbb{R}, \qquad f(x) := f(x_1, x_2, x_3) := \sqrt{1 - x_1^2 - x_2^2 - x_3^2} = \sqrt{1 - \|x\|^2} \qquad (\|x\| \leq 1)$$

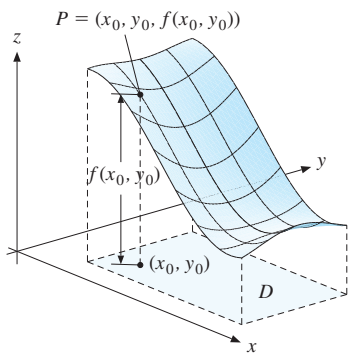
vagy

$$f \in \mathbb{R}^2 \to \mathbb{R}, \qquad f(x,y) = x^3 y - e^{xy} \qquad \big((x,y) \in \mathbb{R}^2\big).$$

Ha $n = 2$, akkor ***kétváltozós valós értékű függvényekről*** beszélünk. Az ilyen függvényeket a

$$\mathrm{Gr}_f := \Big\{(x, y, f(x,y)) \in \mathbb{R}^3 \mid (x,y) \in \mathcal{D}_f\Big\}$$

térbeli halmazzal, az ún. **függvény grafikonjával** tudjuk ábrázolni, ami egy térbeli felületet határoz meg. Ennek alakját úgy tudjuk szemléltetni, hogy a felületen olyan görbesereget rajzolunk fel, amelynek tagjai a felület és olyan sík metszete, amely az xy síkra merőleges, de az x vagy az y tengellyel párhuzamos. Egy másik módszer olyan görbesereget felrajzolni, amelynek tagjai a felület és olyan sík metszete, amely párhuzamos az xy síkkal. Adott $c \in \mathcal{R}_f$ az



$$\Big\{(x, y, c) \in \mathbb{R}^3 \mid (x,y) \in \mathcal{D}_f,\ f(x,y) = c\Big\}$$

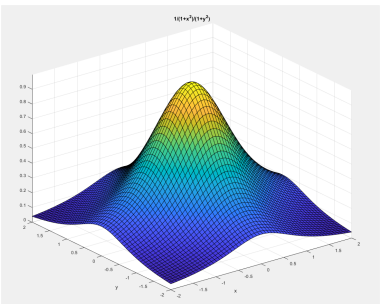
halmazt a grafikon c paraméterhez tartozó ***szintvonalának*** nevezzük. Ilyen ábrázolást a térképészetben használnak.

Matlab programmal az `ezsurf` függvénnyel tudunk kétváltozós valós értékű függvények grafikonját ábrázolni. Pl. az

$$f(x,y) = \frac{1}{(1+x^2)(1+y^2)} \qquad \big((x,y) \in \mathbb{R}^2\big)$$

függvény kódja:

```
ezsurf('1/(1+x^2)/(1+y^2)',[-2,2,-2,2])
```



Előfordul, hogy a kétváltozós f függvény értéke minden $(x,y) \in \mathcal{D}_f$ pontban csak az x^2+y^2 értéktől függ, azaz a $\|(x,y)\|$ értéktől, ami a pont nullától (origótól) való távolsága. Ekkor a függvény szintvonalai olyan körök (vagy egy pont), amiknek középpontja a z tengelyen található. Elég lenne megtartani mindegyikből egyetlen egy pontot, és ezeket megforgatni a z tengely körül, hogy előállítsuk a felületet. Legyen ez a pont az, amire $x \geq 0$ és $y = 0$ teljesül. Így az f függvény grafikonját a

$$g(x) := f(x,0) \qquad \Big((x,0) \in \mathcal{D}_f,\ x \geq 0\Big)$$

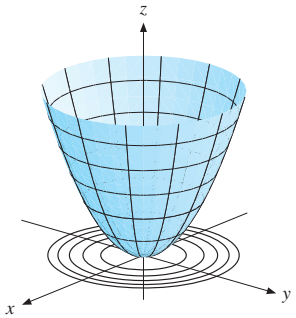
függvény a z tengely körüli megforgatásával kapjuk.

Példák:

a) ***a forgásparaboloid:***

$$f(x,y) := x^2 + y^2 \qquad \Big((x,y) \in \mathbb{R}^2\Big)$$

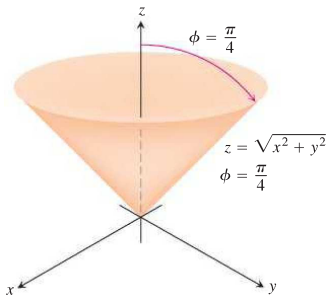
a $g(x) := x^2 \ \ (x \geq 0)$ parabolaág megforgatásával kapott forgásfelület.



b) ***a forgáskúp:***

$$f(x,y) := \sqrt{x^2+y^2} \qquad \Big((x,y) \in \mathbb{R}^2\Big)$$

a $g(x) := \sqrt{x^2} = x \ \ (x \geq 0)$ félegyenes megforgatásával kapott forgásfelület.



$\boxed{\textbf{4.}}$ $f \in \mathbb{R}^n \to \mathbb{R}^m \ \ (n,m > 1)$. Ekkor ***n változós m dimenziós vektor értékű függvényekről*** beszélünk röviden ***vektor-vektor függvényekről***. Az ilyen függvények felírhatók

$$f(x) = \Big(f_1(x), f_2(x), \ldots, f_m(x)\Big) \qquad (x \in \mathcal{D}_f \subseteq \mathbb{R}^n)$$

alakban, ahol az $f_i \colon \mathcal{D}_f \to \mathbb{R}$ n változós valós értékű függvényeket ***koordinátafüggvényeknek*** nevezzük $(i = 1, 2, \ldots, m)$. Pl.

$$f \in \mathbb{R}^2 \to \mathbb{R}^3, \qquad f(x) := f(x_1,x_2) := (x_1^2,\ x_1+x_2, x_1x_2 - 3) \qquad \Big((x_1,x_2) \in \mathbb{R}^2\Big)$$

Az $\mathbb{R}^n \to \mathbb{R}^m$ típusú függvények folytonossága

A sorozatok határértékéhez hasonlóan a többváltozós függvények folytonosságát is a valós-valós függvények folytonosságából nyerjük az euklideszi norma alkalmazásával.

3. Definíció. *Azt mondjuk, hogy az $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n,m \in \mathbb{N}^+)$ függvény* **folytonos az $a \in \mathcal{D}_f$ pontban**, *(jelben $f \in C\{a\}$), ha*

$$\forall \varepsilon > 0 \text{ számhoz } \exists \delta > 0 \text{ úgy, hogy } \forall x \in \mathcal{D}_f,\ \|x-a\| < \delta \text{ pontban } \|f(x) - f(a)\| < \varepsilon.$$

Megjegyzések.

1. Az euklideszi normára mindig a $\|\cdot\|$ jelölést alkalmazzuk függetlenül attól, hogy hány dimenziós a benne szereplő vektor.

2. A folytonosság fogalma leírható környezetekkel. $f \in C\{a\}$, ha $a \in \mathcal{D}_f$ és

$$\forall \varepsilon > 0 \text{ számhoz } \exists \delta > 0 \text{ úgy, hogy } \forall x \in K_\delta(a) \cap \mathcal{D}_f \text{ pontban } f(x) \in K_\varepsilon\Big(f(a)\Big).$$

   Ez pontosan megegyezik a valós-valós függvényeknél tanult fogalmával.

3. A folytonosság most is az f függvénynek azt a szemléletes tulajdonságát fejezi ki, hogy „ha x közel van az a ponthoz, akkor az $f(x)$ függvényérték közel van $f(a)$-hoz". De a közelség most azt jelenti, hogy a két pont euklideszi távolsága kicsi.

Nem nehéz igazolni, hogy ***a norma folytonos függvény***. Ez azért igaz, mert ha $f \in \mathbb{R}^n \to \mathbb{R}$, $f(x) := \|x\|$, akkor $\forall \varepsilon > 0$-hoz $\exists \delta := \varepsilon > 0$ úgy, hogy $\forall x \in \mathbb{R}^n$ esetén ha $\|x - a\| < \delta$, akkor

$$\|f(x) - f(a)\| = \Big|\|x\| - \|a\|\Big| \leq \|x - a\| < \delta = \varepsilon.$$

Másrészt, a $\mathrm{pr}_i \in \mathbb{R}^n \to \mathbb{R}$, $\mathrm{pr}_i(x) := x_i \quad (i = 1, 2, \ldots, n)$ ún. ***projekciók*** (a ponthoz rendeli az i-dik koordinátáját) szintén folytonos függvények. Valóban, $\forall \varepsilon > 0$-hoz $\exists \delta := \varepsilon > 0$ úgy, hogy $\forall x \in \mathbb{R}^n$ esetén ha $\|x - a\| < \delta$, akkor

$$\|\mathrm{pr}_i(x) - \mathrm{pr}_i(a)\| = |x_i - a_i| \leq \sqrt{\sum_{k=1}^{n} (x_k - a_k)^2} = \|x - a\| < \delta = \varepsilon.$$

Azt a tényt, hogy $\forall a \in \mathcal{D}_f \colon a \in C\{a\}$, azaz f folytonos minden értelmezési tartománybeli pontjában, az $f \in C$ jelöléssel fogjuk rövidíteni. Az előzőek szerint $\|.\| \in C$ és $\mathrm{pr}_i \in C$ $(i = 1, 2, \ldots, n)$.

4. Tétel (A folytonosságra vonatkozó átviteli elv). *Legyen $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ és $a \in \mathcal{D}_f$. Ekkor*

$$f \in C\{a\} \quad \Longleftrightarrow \quad \forall (x_k) \colon \mathbb{N} \to \mathcal{D}_f, \ \lim_{k \to +\infty} x_k = a \text{ esetén } \lim_{k \to +\infty} f(x_k) = f(a).$$

Bizonyítás. Hasonlóan igazolható, mint valós-valós függvények esetén, a környezetekkel leírt folytonosság fogalmából kiindulva.

Megjegyzés. Az átviteli elvből következik, hogy ha $\exists (x_k) : \mathbb{N} \to \mathcal{D}_f$ sorozat, amely az a ponthoz konvergál, de

$$\lim_{k \to +\infty} f(x_k) \neq f(a),$$

akkor az f függvény nem folytonos a pontban.

Az átviteli elvvel és a sorozatok határértékére vonatkozó műveleti tételekkel nem nehéz igazolni a következő állításokat.

<u>Műveletek folytonos függvényekkel:</u>

1. Ha $f, g \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ és $f, g \in C\{a\}$, akkor

   a) $f + g \in C\{a\}$ és $\lambda f \in C\{a\}$ $(\lambda \in \mathbb{R})$

   b) az $m = 1$ esetben $f \cdot g \in C\{a\}$ és $g(a) \neq 0$ esetén $\dfrac{f}{g} \in C\{a\}$.

2. Ha $g \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$, $g \in C\{a\}$ és $f \in \mathbb{R}^m \to \mathbb{R}^p$ $(m, p \in \mathbb{N}^+)$, $f \in C\{g(a)\}$, akkor $f \circ g \in C\{a\}$.

Példa: A fenti állításokból igazolható, hogy

$$f \in \mathbb{R}^2 \to \mathbb{R}, \quad f(x,y) := \frac{x+y}{x^2+1} - \sin(xe^{x-y^3} + \pi) \in C$$

hiszen $\sin, \exp \in C$, illetve ha $\mathrm{pr}_1(x,y) := x$ és $\mathrm{pr}_2(x,y) := y$, akkor $\mathrm{pr}_1, \mathrm{pr}_2 \in C$.

Példa: Az

$$f(x,y) := \begin{cases} \dfrac{xy}{x^2+y^2}, & \text{ha } (x,y) \neq (0,0) \\ 0, & \text{ha } (x,y) = (0,0) \end{cases}$$

függvény nem folytonos a $(0,0)$ pontban. Valóban az átviteli elv szerint $f \notin C\{(0,0)\}$, hiszen $(x_k, y_k) := \left(\frac{1}{k}, \frac{1}{k}\right) \to (0,0)$ ha $k \to +\infty$, de

$$f(x_k, y_k) = f\left(\tfrac{1}{k}, \tfrac{1}{k}\right) = \frac{\frac{1}{k} \cdot \frac{1}{k}}{\frac{1}{k^2} + \frac{1}{k^2}} = \frac{1}{2} \not\to 0 = f(0,0).$$

A következő tétel azt mondja ki, hogy az n változós m dimenziós vektor értékű függvények folytonossága visszavezethető n változós valós értékű függvények folytonosságára.

5. Tétel. *Legyen $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ és $a \in \mathcal{D}_f$. Ekkor*

$$f \in C\{a\} \qquad \Longleftrightarrow \qquad f_i \in C\{a\} \quad (i = 1, 2, \ldots, m),$$

ahol $f_i \colon \mathcal{D}_f \to \mathbb{R}$ az f függvény koordinátafüggvényei.

Bizonyítás. A sorozatok konvergenciája visszavezethető a koordinátasorozatok konvergenciájára. Ez azt jelenti, hogy az átviteli elv a következő alakban írható:

$$f \in C\{a\} \quad \Longleftrightarrow \quad \forall (x_k) \colon \mathbb{N} \to \mathcal{D}_f, \ \lim_{k \to +\infty} x_k^{(i)} = a_i \text{ esetén } \lim_{k \to +\infty} \Big(f(x_k)\Big)^{(i)} = \Big(f(a)\Big)_i.$$

minden $i = 1, 2, \ldots, m$ esetén. Ekkor a tétel állítása az $\Big(f(x_k)\Big)^{(i)} = f_i(x_k)$ egyenlőségből következik.

Példa: A fenti állításból igazolható, hogy

$$f \in \mathbb{R}^2 \to \mathbb{R}^2, \quad f(x,y) := \left(\frac{x+y}{x^2+1}, \sin(xe^{x-y^3} + \pi)\right) \in C$$

hiszen az

$$f_1(x,y) := \frac{x+y}{x^2+1}, \quad f_2(x,y) := \sin(xe^{x-y^3} + \pi) \qquad \Big((x,y) \in \mathbb{R}^2\Big)$$

koordinátafüggvényeiről a műveleti tétel alapján igazolhatjuk, hogy $f_1, f_2 \in C$.

Felmerül a kérdés, hogy az n változós valós értékű függvények folytonosságát vissza tudjuk-e vezetni valós-valós függvények folytonosságára. Arra gondolnánk, hogy ha lerögzítjük az $a \in \mathcal{D}_f$ pont koordinátait az i-edik koordináta kivételével, akkor elegendő lenne megvizsgálni a

$$g_i^{(a)}(x) := f(a_1, a_2, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n) \qquad \Big(x \in \mathbb{R} \colon (a_1, a_2, \ldots, a_{i-1}, x, a_{i+1}, \ldots, a_n) \in \mathcal{D}_f\Big)$$

függvényt minden $i = 1, 2, \ldots, n$ esetén. Ez sajnos ***nem igaz***. Pl. már láttuk, hogy az

$$f(x,y) := \begin{cases} \dfrac{xy}{x^2+y^2}, & \text{ha } (x,y) \neq (0,0) \\ 0, & \text{ha } (x,y) = (0,0) \end{cases}$$

függvény nem folytonos az $a = (0,0)$ pontban. Azonban a $g_1^{(a)}(x) = f(x,0) = 0 \ \ (x \in \mathbb{R})$ és a $g_2^{(a)}(y) = f(0,y) = 0 \ \ (y \in \mathbb{R})$ függvények folytonosak a 0 pontban.

6. Tétel (Weierstrass tétele.). *Legyen $n \in \mathbb{N}^+$ és tegyük fel, hogy*

a) $f \in \mathbb{R}^n \to \mathbb{R}$,

b) $\mathcal{D}_f$ *korlátos és zárt halmaz az* $\mathbb{R}^n$ *euklideszi térben,*

c) $f \in C$.

Ekkor az f függvénynek vannak abszolút szélsőértékhelyei, azaz

$$\exists\, x_1 \in \mathcal{D}_f\colon \quad f(x) \le f(x_1) \quad (\forall\, x \in \mathcal{D}_f) \quad (x_1 \text{ abszolút maximumhely}),$$

$$\exists\, x_2 \in \mathcal{D}_f\colon \quad f(x_2) \le f(x) \quad (\forall\, x \in \mathcal{D}_f) \quad (x_2 \text{ abszolút minimumhely}).$$

Bizonyítás. Az egyváltozós esetben bemutatott bizonyítás adaptálható többváltozós esetre.

Az $\mathbb{R}^n \to \mathbb{R}^m$ típusú függvények határértéke

A függvény határértéke szintén a valós-valós eset általánosításaként kerül bevezetésre.

4. Definíció. *Az $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ függvénynek az $a \in \mathcal{D}'_f$ pontban* **van határértéke**, *ha $\exists A \in \mathbb{R}^m$, hogy*

$$\forall \varepsilon > 0 \text{ számhoz } \exists \delta > 0 \text{ úgy, hogy } \forall x \in \mathcal{D}_f,\ 0 < \|x - a\| < \delta \text{ pontban } \|f(x) - A\| < \varepsilon.$$

Ekkor $A-$t a **függvény *a* pontbeli határértékének** *nevezzük, és az alábbi szimbólumok valamelyikével jelöljük:*

$$\lim_a f = A, \qquad \lim_{x \to a} f(x) = A, \qquad f(x) \to A, \text{ ha } x \to a.$$

A határérték fogalma is megadható környezetekkel, ami pontosan megegyezik a valós-valós függvényeknél leírtakkal. A határérték egyértelműsége környezetekkel ugyanúgy igazolható, mint valós-valós függvények esetében. Ugyanez mondható az átviteli elvre.

7. Tétel (A határértékre vonatkozó átviteli elv). *Legyen $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ és $a \in \mathcal{D}'_f$. Ekkor*

$$\lim_a f = A \in \mathbb{R}^m \quad \Longleftrightarrow \quad \forall (x_k) : \mathbb{N} \to \mathcal{D}_f \backslash \{a\},\ \lim_{k \to +\infty} (x_k) = a \text{ esetén } \lim_{k \to +\infty} f(x_k) = A.$$

Megjegyzés. Az átviteli elvből következik, hogy ha van két olyan $(x_k), (y_k) : \mathbb{N} \to \mathcal{D}_f$ sorozat, amely az a ponthoz konvergál, de

$$\lim_{k \to +\infty} f(x_k) \neq \lim_{k \to +\infty} f(y_k),$$

akkor az f függvénynek nincs határértéke az a pontban.

Példa: Az

$$f(x, y) := \begin{cases} \dfrac{xy}{x^2 + y^2}, & \text{ha } (x, y) \neq (0, 0) \\ 0, & \text{ha } (x, y) = (0, 0) \end{cases}$$

függvénynek nincs határértéke a $(0,0)$ pontban, hiszen ha $k \to +\infty$, akkor

$$\left(\tfrac{1}{k}, \tfrac{1}{k}\right) \to (0,0) \qquad \text{és} \qquad f\left(\tfrac{1}{k}, \tfrac{1}{k}\right) = \frac{\frac{1}{k} \cdot \frac{1}{k}}{\frac{1}{k^2} + \frac{1}{k^2}} = \frac{1}{2},$$

$$\left(0, \tfrac{1}{k}\right) \to (0,0) \qquad \text{és} \qquad f\left(0, \tfrac{1}{k}\right) = \frac{0 \cdot \frac{1}{k}}{0 + \frac{1}{k^2}} = 0.$$

Tehát két, a $(0,0)$ ponthoz tartó sorozat képsorozatának határtétéke különbözik.

A folytonosság és a határérték kapcsolatát fejezi ki a következő állítás.

8. Tétel. *Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ és $a \in \mathcal{D}_f \cap \mathcal{D}'_f$. Ekkor*

$$f \in C\{a\} \qquad \Longleftrightarrow \qquad \exists \lim_a f \quad \text{és} \quad \lim_a f = f(a).$$

9. előadás

TÖBBVÁLTOZÓS FÜGGVÉNYEK DIFFERENCIÁLSZÁMÍTÁSA 1.

Többváltozós függvények esetén is megpróbálhatjuk átívelni a valós-valós függvényeknél alkalmazott differenciálhatóság fogalmát úgy, ahogy a folytonossággal és a határértékkel tettük. Azonban ez nem megy olyan egyszerűen. A problémát az okozza, hogy a differenciálhányados meghatározásához, mint a nevéből is kitűnik, elemek hányadosát kellene képezni, amit többdimenziós térben nem értelmezünk. A probléma elkerülésére először megpróbálunk olyan „metszeteket” kinyerni a többváltozós függvényből, amelyek már valós-valós függvényekkel leírhatók, és ezeket így a differenciálszámítás eszközeivel már tudjuk elemezni.

Parciális deriváltak $\mathbb{R}^n \to \mathbb{R}$ típusú függvényekre

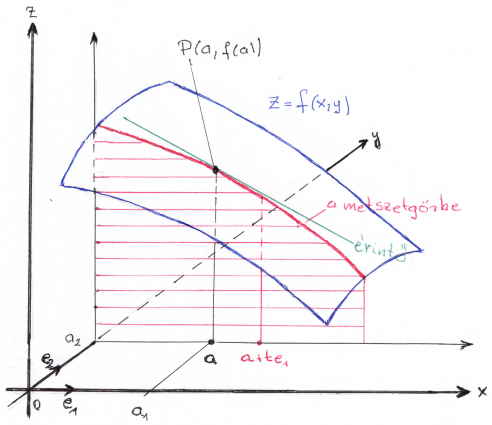
A parciális deriváltakat úgy kapjuk, hogy egy híján minden változót rögzítünk, és az így kapott egyváltozós függvényt deriváljuk.

A kétváltozós esetben a fogalomnak szemléletes jelentés adható. Legyen

$$f \in \mathbb{R}^2 \to \mathbb{R} \quad \text{és} \quad a = (a_1, a_2) \in \operatorname{int} \mathcal{D}_f.$$

A függvény grafikonja a térben a $z = f(x, y)$ egyenletű felület. Fektessünk az a ponton át az x tengellyel párhuzamos egyenest. Ennek pontjai az xy síkban

$$(a_1 + t, a_2) \qquad (t \in \mathbb{R})$$



Vegyük a függvény értékeit ezekben a pontokban, és képezzük velük az

$$F_x(t) := f(a_1 + t, a_2)$$

valós-valós függvényt. Ez a függvény értelmezhető a $t = 0$ pontnak egy $K(0)$ környezetében, mert $a \in \operatorname{int} \mathcal{D}_f$. Az F_x függvény képe egy, a felületen futó (metszet)görbe, vagyis a $z = f(x, y)$ egyenletű felület és az $y = a_2$ (x, z tetszőleges) egyenletű sík metszésvonala. ***Az f függvény x változó szerinti parciális deriváltját az a pontban*** (jele: $\boldsymbol{\partial_x f(a)}$) úgy értelmezzük, mint ***az F_x függvény deriváltja a 0 pontban***, feltéve, hogy a derivált létezik, azaz

$$\partial_x f(a) := F_x'(0) = \lim_{t \to 0} \frac{F_x(t) - F_x(0)}{t} = \lim_{t \to 0} \frac{f(a_1 + t, a_2) - f(a_1, a_2)}{t}.$$

$F_x'(0)$ a metszetgörbe $P\big(a, f(a)\big)$ pontbeli érintőjének a meredeksége.

Az y változó szerinti parciális deriváltat hasonló módon értelmezzük. Ebben az esetben az a ponton átmenő, az y tengellyel párhuzamos egyenest kell figyelembe venni, amelynek pontjai $(a_1, a_2 + t)$ $(t \in \mathbb{R})$. Ez pedig az alábbi értelmezéshez vezet.

$$\partial_y f(a) := \lim_{t \to 0} \frac{f(a_1, a_2 + t) - f(a_1, a_2)}{t}.$$

Egy $f \in \mathbb{R}^n \to \mathbb{R}$ függvény mindegyik változója szerint képezhetjük a parciális deriváltat valamely $a = (a_1, a_2, \dots, a_n) \in \operatorname{int} \mathcal{D}_f$ pontban. Rögzítsük az $a = (a_1, a_2, \dots, a_n)$ pont koordinátáit az i-edik kivételével. Az

$$F_i \colon K(0) \ni t \mapsto f(a_1, a_2, \dots, a_{i-1}, a_i + t, a_{i+1}, \dots, a_n) = f(a + te_i)$$

valós-valós függvény $t = 0$ pontban vett deriváltját (amennyiben létezik) nevezzük az f függvény a-ban vett i-edik parciális deriváltjának. $e_1, e_2, \dots, e_n$ jelenti a kanonikus bázist $\mathbb{R}^n$-ben, azaz

$$e_i := (0, 0, \dots, 0, \underbrace{1}_{i\text{-edik}}, 0, \dots, 0) \qquad (i = 1, 2, \dots, n).$$

1. Definíció. *Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$, $a \in \operatorname{int} \mathcal{D}_f$ és $e_1, \dots, e_n$ a kanonikus bázis $\mathbb{R}^n$-ben. Az f függvénynek az a pontban létezik az i-edik $(i = 1, 2, \dots, n)$ változó szerinti* **parciális deriváltja**, *ha az*

$$F_i \colon K(0) \ni t \mapsto f(a + te_i)$$

valós-valós függvény deriválható a 0 pontban. A $F_i'(0)$ valós számot az ***f* függvény *a* pontbeli, *i*-edik változó szerinti parciális deriváltjának** *nevezzük, és az alábbi szimbólumok valamelyikével jelöljük:*

$$\partial_i f(a), \quad \partial_{x_i} f(a), \quad \frac{\partial f}{\partial x_i}(a), \quad f'_{x_i}(a), \quad D_i f(a).$$

A parciális deriváltakat a következő módon számíthatjuk ki. Tekintsük meg a

$$G_i \colon K(a_i) \ni x_i \mapsto f(a_1, a_2, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$$

ún. ***parciális függvényt***! Mivel

$$F_i(t) = G_i(a_i + t) \quad \Longrightarrow \quad F_i'(t) = G_i'(a_i + t) \cdot (a_i + t)' = G_i'(a_i + t) \quad \Longrightarrow \quad F_i'(0) = G_i'(a_i),$$

így $\partial_i f(a) = G_i'(a_i)$. Más szavakkal, G_i az a függvény, amit úgy kapunk az f függvényből, hogy csak az i-edik változója marad változónak, a többi az a pont megfelelő koordinátájával egyenlő. Ezért $G_i'(a_i)$ az f függvény deriváltja az i-edik változója szerint úgy, hogy a többi változóját konstansnak kell tekinteni, és a végeredményben az a pont koordinátait kell behelyettesíteni. Ha $n = 1$, azaz f valós-valós függvény, akkor a parciális derivált megegyezik a „rendes” deriválttal.

Példa: Ha ki akarjuk számítani az

$$f(x, y) := xe^{x^2+y^2} - 3y \qquad \left((x, y) \in \mathbb{R}^2\right)$$

függvény x változója szerinti parciális deriváltját az $a = (a_1, a_2)$ pontban, akkor az y változót konstansnak tekintjük, és az x változó szerint deriválunk:

$$\begin{aligned}
\partial_x f(a_1, a_2) &= (x)'_x \cdot e^{x^2+y^2} + x \cdot (e^{x^2+y^2})'_x - (3y)'_x \Big|_{x=a_1,\, y=a_2} = \\
&= 1 \cdot e^{x^2+y^2} + x \cdot e^{x^2+y^2} \cdot (x^2 + y^2)'_x - 0 \Big|_{x=a_1,\, y=a_2} = \\
&= e^{x^2+y^2} + x \cdot e^{x^2+y^2} \cdot 2x \Big|_{x=a_1,\, y=a_2} = \\
&= e^{a_1^2+a_2^2} + 2a_1^2\, e^{a_1^2+a_2^2}.
\end{aligned}$$

Ugyanígy az y változó szerinti parciális deriválás során az x változót tekintjük konstansnak, és az y változó szerint deriválunk:

$$\begin{aligned}\partial_y f(a_1,a_2) &= x\cdot (e^{x^2+y^2})'_y - (3y)'_y\Big|_{x=a_1,\,y=a_2} = \\ &= x\cdot e^{x^2+y^2}\cdot (x^2+y^2)'_y - 3\Big|_{x=a_1,\,y=a_2} = \\ &= x\cdot e^{x^2+y^2}\cdot 2y - 3\Big|_{x=a_1,\,y=a_2} = \\ &= 2a_1a_2\,e^{a_1^2+a_2^2} - 3.\end{aligned}$$

A fenti számításaink szerint például

$$\partial_x f(0,0) = 1 \qquad \text{és} \qquad \partial_y f(0,0) = -3.$$

Legyen az f függvény értelmezve $\mathbb{R}^n$ egy részhalmazán. Az f függvény i-edik ***parciális deriváltfüggvényén*** azt a $\partial_i f$ függvényt értjük, amely azokban a pontokban van értelmezve, ahol az f függvény i-edik parciális deriváltja létezik és véges, és ott az értéke $\partial_i f(a)$.

Magasabb rendű parciális deriváltak

Ha egy n-változós függvénynek az $a \in \mathbb{R}^n$ pont egy környezetében létezik a függvény valamely változó szerinti parciális deriváltja, akkor a parciális deriváltfüggvény szintén n-változós valós értékű függvény. Így ennek bármelyik másik változó szerinti parciális deriválhatóságát vizsgálhatjuk.

Legyen f értelmezve az $a \in \mathbb{R}^n$ pont egy környezetében. Ha rögzített $i = 1, 2, \ldots, n$ esetén a $\partial_i f$ parciális derivált létezik az a pont egy környezetében és a $\partial_i f$ parciális deriváltfüggvénynek létezik a j-edik ($j = 1, 2, \ldots, n$) változó szerinti parciális deriváltja az a pontban, akkor a $\partial_j\Big(\partial_i f\Big)(a)$ számot (mint $\partial_i f \in \mathbb{R}^n \to \mathbb{R}$ függvény a-beli j-edik változó szerinti parciális deriváltját) a függvény a-beli ij-edik **másodrendű parciális deriváltjának** nevezzük és így jelöljük:

$$\partial_{ij} f(a) := \partial_i\partial_j f(a) := \partial_j\Big(\partial_i f\Big)(a)$$

$\partial_{ij} f(a)$ helyett használatosak még a következő jelölések is:

$$\partial_{x_ix_j} f(a), \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(a), \quad f''_{x_ix_j}(a), \quad D_jD_i f(a), \quad D_{ji} f(a).$$

A $\partial_{ij} f(a)$ deriváltat $i = j$ esetén másodrendű ***tiszta parciális deriváltnak*** nevezzük, és a

$$\partial_i^2 f(a), \quad \frac{\partial^2 f}{\partial x_i^2}(a), \quad \ldots .$$

szimbólumok valamelyikével jelöljük. Ha $i \neq j$, akkor $\partial_{ij} f(a)$-t másodrendű ***vegyes parciális deriváltnak*** is szokás nevezni.

Kettőnél magasabb rendre az s-edrendű ($2 < s \in \mathbb{N}$) parciális deriválhatóságot s szerinti indukcióval definiálhatjuk. Ha $1 \le i_1, i_2, \ldots, i_s \le n$ tetszőleges indexek, akkor az f függvény a-beli s-edrendű i_1-edik, …, i_s-edik változó szerinti parciális deriváltját az alábbi szimbólumok valamelyikével jelöljük:

$$\partial_{i_1}\partial_{i_2}\ldots\partial_{i_s} f(a), \quad \frac{\partial^s f}{\partial x_{i_s}\ldots\partial x_{i_1}}(a), \quad \ldots .$$

Iránymenti deriváltak $\mathbb{R}^n \to \mathbb{R}$ típusú függvényekre

A parciális deriváltaknál az e_i kanonikus vektorokkal párhuzamos „irányokban” deriváltuk az $f\colon \mathbb{R}^n \to \mathbb{R}$ függvény értékeiből keletkezett valós-valós függvényt az a pontban. Ezt úgy fogjuk általánosítani, hogy egy tetszőleges irányban csináljuk ugyanezt.

Egy $f \in \mathbb{R}^n \to \mathbb{R}$ $(2 \le n \in \mathbb{N})$ függvény minden $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}_n$ egységvektor szerint képezhetjük a v irányú iránymenti deriváltat valamely $a = (a_1, a_2, \dots, a_n) \in \operatorname{int} \mathcal{D}_f$ pontban. Mivel v egységvektor, így

$$\|v\|^2 = v_1^2 + v_2^2 + \dots + v_n^2 = 1.$$

Az

$$F_v\colon K(0) \ni t \mapsto f(a_1 + tv_1, a_2 + tv_2, \dots, a_n + tv_n) = f(a + tv)$$

valós-valós függvény $t = 0$ pontban vett deriváltját (amennyiben létezik) nevezzük az f függvény v irányú iránymenti deriváltjának az a pontban.

2. Definíció. *Legyen $n \in \mathbb{N}^+$, $f \in \mathbb{R}^n \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$. Tegyük fel, hogy $v \in \mathbb{R}^n$ egy egységvektor: $\|v\| = 1$. A f függvénynek az a pontban létezik a v irányú* **iránymenti deriváltja**, *ha a*

$$F_v : K(0) \ni t \mapsto f(a + tv)$$

valós-valós függvény deriválható a 0 pontban. A $F_v'(0)$ valós számot az $\boldsymbol{f}$ **függvény** $\boldsymbol{a}$ **pontbeli** $\boldsymbol{v}$ **irányú iránymenti deriváltjának** *nevezzük, és a* $\boldsymbol{\partial_v f(a)}$ *vagy a* $\boldsymbol{f_v'(a)}$ *szimbólummal jelöljük.*

Ha $i = 1, 2, \dots, n$ egy rögzített index és $v = e_i = (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{R}^n$ (ahol tehát a v vektor i-edik koordinátája 1, a többi 0), akkor $F_v = F_{e_i} = F_i$, és így

$$\partial_v f(a) = \partial_{e_i} f(a) = F_{e_i}'(0) = F_i'(0) = \partial_i f(a).$$

Az iránymenti derivált tehát a parciális derivált általánosítása.

Példa: Számítsuk ki az

$$f(x, y) := xe^{x+y} + \sin xy \qquad \big((x, y) \in \mathbb{R}^2\big)$$

függvény $v := \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ irányú iránymenti deriváltját az $a := (0, 0)$ pontban! Mivel

$$F_v(t) := f(a + tv) = f\left((0,0) + t\left(\tfrac{1}{2}, \tfrac{\sqrt{3}}{2}\right)\right) = f\left(\tfrac{1}{2}t, \tfrac{\sqrt{3}}{2}t\right) =$$

$$= \frac{1}{2}t\, e^{\frac{1}{2}t + \frac{\sqrt{3}}{2}t} + \sin\left(\frac{1}{2}t \cdot \frac{\sqrt{3}}{2}t\right) = \frac{1}{2}t\, e^{\frac{1+\sqrt{3}}{2}t} + \sin\left(\frac{\sqrt{3}}{4}t^2\right),$$

így

$$F_v'(t) := \frac{1}{2}e^{\frac{1+\sqrt{3}}{2}t} + \frac{1}{2}t\, e^{\frac{1+\sqrt{3}}{2}t} \cdot \frac{1+\sqrt{3}}{2} + \cos\left(\frac{\sqrt{3}}{4}t^2\right) \cdot \frac{\sqrt{3}}{2}t.$$

Ekkor

$$\partial_v f(a) = F_v'(0) = \frac{1}{2}.$$

Az esetek többségében az iránymenti derivált kiszámítása (vagyis F_v deriválhatóságának a vizsgálata) hosszadalmas feladat (ti. a parciális deriválttal ellentétben most f minden komponense változik). Bizonyos feltételek mellett gyorsabban tudunk eljutni a végeredményhez. Már láttuk, hogy a parciális deriváltak speciális iránymenti deriváltak, hiszen $\partial_{e_i} f(a) = \partial_i f(a)$ minden $i = 1, 2, \dots, n$ esetén. A következő állítás egy fordított kapcsolatot mutat.

1. Tétel. *Tegyük fel, hogy $f \in \mathbb{R}^n \to \mathbb{R}$ ($n \in \mathbb{N}^+$), $a \in \operatorname{int} \mathcal{D}_f$, illetve az f függvénynek léteznek a parciális deriváltjai egy $K(a) \subset \mathcal{D}_f$ környezetben, és ezek folytonosak az a pontban. Ekkor az f függvénynek az a pontból induló tetszőleges $v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n$ egységvektor irányban létezik az iránymenti deriváltja, és*

$$\partial_v f(a) = \sum_{i=1}^{n} \partial_i f(a) \cdot v_i = \partial_1 f(a) \cdot v_1 + \partial_2 f(a) \cdot v_2 + \dots + \partial_n f(a) \cdot v_n.$$

Bizonyítás. Később.

A tétel segítségével lerövidülnek a számítások. A fenti példánál:

$$f(x,y) := xe^{x+y} + \sin xy \qquad \big((x,y) \in \mathbb{R}^2\big), \qquad v := \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right), \quad a := (0,0).$$

Mivel

$$\begin{aligned} \partial_x f(x,y) &= e^{x+y} + xe^{x+y} + y \cos xy, & \partial_x f(0,0) &= 1, \\ \partial_y f(x,y) &= xe^{x+y} + x \cos xy, & \partial_y f(0,0) &= 0, \end{aligned}$$

ezért

$$\partial_v f(a) = \partial_x f(0,0) \cdot v_1 + \partial_y f(0,0) \cdot v_2 = 1 \cdot \frac{1}{2} + 0 \cdot \frac{\sqrt{3}}{2} = \frac{1}{2}.$$

A totális derivált

A totális deriválhatóság fogalma ugyan pontosan megfelel az egyváltozós deriválhatóság definíciójának, de mégis bonyolultabb, mint egyváltozós esetben. Idézzünk fel a valós-valós függvények deriválhatóságával kapcsolatos néhány fontos ismeretet.

Az $f \in \mathbb{R} \to \mathbb{R}$ függvény *differenciálható* vagy *deriválható* az $a \in \operatorname{int} \mathcal{D}_f$ pontban (jelben: $f \in D\{a\}$), ha létezik és véges a

$$\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} \quad \text{határérték.}$$

Ezt a határértéket az $f'(a)$ szimbólummal jelöljük, és az f függvény a pontbeli *deriváltjának* vagy *differenciálhányadosának* nevezzük.

A többváltozós függvények esetében a differenciahányadosnak nincs közvetlen megfelelője (hiszen vektorok körében nem tudunk osztani), ezért a deriválhatóságot nem tudjuk differenciahányadosok határértékeként értelmezni. Az egyváltozós analízisben azonban láttuk azt, hogy az elsőfokú polinomokkal való lokális közelíthetőség ekvivalens a differenciálhatósággal. Nevezetesen: ha $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$, akkor

$$f \in D\{a\} \quad \Longleftrightarrow \quad \begin{cases} \exists A \in \mathbb{R} \ \text{ és } \ \exists \varepsilon : \mathcal{D}_f \to \mathbb{R}, \ \lim\limits_{0} \varepsilon = 0 : \\ f(a+h) - f(a) = A \cdot h + \varepsilon(h) \cdot h \qquad (a + h \in \mathcal{D}_f). \end{cases}$$

Ekkor az A szám az f függvény a pontbeli deriváltja, vagyis $A = f'(a)$. Az f függvény a *pontbeli deriválhatósága tehát azt jelenti, hogy a függvény megváltozása az a pont környezetében „jól” közelíthető lineáris függvénnyel.*

Vegyük észre, hogy a fentiekben az ε függvény szerepeltetése „kiküszöbölhető”. Pontosabban: ha $f \in \mathbb{R} \to \mathbb{R}$ és $a \in \operatorname{int} \mathcal{D}_f$, akkor

$$f \in D\{a\} \quad \Longleftrightarrow \quad \exists A \in \mathbb{R}\colon \quad \lim_{h \to 0} \frac{f(a+h) - f(a) - A \cdot h}{h} = \lim_{0} \varepsilon = 0.$$

Ha még azt is figyelembe vesszük, hogy $\lim\limits_{0} \varepsilon = 0 \iff \lim\limits_{0} |\varepsilon| = 0$, akkor végül azt kapjuk, hogy

$$(*) \qquad \boxed{f \in D\{a\} \quad \Longleftrightarrow \quad \exists\, A \in \mathbb{R}: \quad \lim_{h \to 0} \frac{|f(a+h) - f(a) - A \cdot h|}{|h|} = 0.}$$

Valós-valós függvény deriválhatóságára vonatkozó $(*)$ ekvivalens átfogalmazás már kiterjeszthető vektor-vektor függvényre is. Ehhez vegyük észre, hogy a $(*)$-ban szereplő $L(h) := A \cdot h$ tag felfogható, mint egy $L : \mathbb{R} \to \mathbb{R}$ *lineáris transzformáció*, ami azt jelenti, hogy

$$(\#) \qquad L(\alpha x + \beta y) = \alpha L(x) + \beta L(y)$$

teljesül minden $x, y \in \mathbb{R}$ és $\alpha, \beta \in \mathbb{R}$ esetén. A lineáris transzformáció fogalma is hasonlóan megadható $L : \mathbb{R}^n \to \mathbb{R}^m$ típusú leképezések esetén, nevezetesen úgy, hogy $(\#)$ teljesüljön minden $x, y \in \mathbb{R}^n$ és $\alpha, \beta \in \mathbb{R}$ esetén. Tudjuk, hogy ekkor van egy $m \times n$-es A mátrix, amire $L(h) = A \cdot h$ teljesül. Itt $h \in \mathbb{R}^n$-t oszlopvektorként kell tekinteni, és a szorzás a mátrixszorzatot jelenti. ***Így $(*)$-ban az abszolút értéket a megfelelő normákkal, az A valós számot pedig egy $m \times n$-es mátrixszal helyettesítjük***.

3. Definíció. *Az $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ függvény* **totálisan deriválható** *az $a \in \operatorname{int} \mathcal{D}_f$ pontban (jelben: $\boldsymbol{f \in D\{a\}}$), ha*

$$\exists A \in \mathbb{R}^{m \times n}\colon \quad \lim_{h \to 0} \frac{\|f(a+h) - f(a) - A \cdot h\|}{\|h\|} = 0.$$

Ekkor $f'(a) := A$ az f függvény **deriváltmátrixa** *az a pontban.*

Megjegyzések.

1. Az euklideszi normára mindig a $\|\cdot\|$ jelölést alkalmazzuk függetlenül attól, hogy hány dimenziós a benne szereplő vektor. $h \to 0$ azt jelenti, hogy a h vektor tart a 0 vektorhoz $\mathbb{R}^n$-ben, és így a határértékben lévő valós kifejezésnek tartania kell a 0 számhoz.

2. A fenti definíciót úgy lehet átfogalmazni, hogy

$$f \in D\{a\} \quad \Longleftrightarrow \quad \begin{cases} \exists A \in \mathbb{R}^{m \times n} \ \text{ és } \ \exists \varepsilon : \mathcal{D}_f \to \mathbb{R}^m, \ \lim\limits_{0} \varepsilon = 0 : \\ f(a+h) - f(a) = A \cdot h + \varepsilon(h) \cdot \|h\| \qquad (a + h \in \mathcal{D}_f). \end{cases}$$

2. Tétel. *Ha $f \in D\{a\}$, akkor az $f'(a)$ deriváltmátrix egyértelműen meghatározott.*

Bizonyítás. Tegyük fel, hogy az A_1 és A_2 mátrixok kielégítik a totálisan deriválhatóság definiciójában szereplő feltételeket. Ekkor

$$0 \le \frac{\|(A_1 - A_2)h\|}{\|h\|} = \frac{\|(f(a+h) - f(a) - A_2 h) - (f(a+h) - f(a) - A_1 h)\|}{\|h\|} \le$$

$$= \frac{\|f(a+h) - f(a) - A_2 h\|}{\|h\|} + \frac{\|f(a+h) - f(a) - A_1 h\|}{\|h\|} \to 0 + 0 = 0 \quad (h \to 0).$$

Ezért

$$\lim_{h\to 0}\frac{\|(A_1-A_2)h\|}{\|h\|}=0 \quad \overset{h=te_i}{\Longrightarrow} \quad 0=\lim_{t\to 0}\frac{\|(A_1-A_2)(te_i)\|}{\|te_i\|}=\lim_{t\to 0}\frac{|t|\cdot\|(A_1-A_2)e_i\|}{|t|\cdot\|e_i\|}=$$

$$=\|(A_1-A_2)e_i\|.$$

Ezért $(A_1-A_2)e_i=0$, azaz $A_1e_i=A_2e_i$ minden $i=1,2,\dots,n$ esetén. Tehát a mátrixok mindegyik i-edik oszlopa megegyezik, és így $A_1=A_2$.

Példa: Az $f(x,y)=xy \quad \big((x,y)\in\mathbb{R}\big)$ függvény differenciálható az $a=(1,2)$ pontban és ebben a pontban vett deriváltmátrixa: $f'(1,2)=\begin{pmatrix}2 & 1\end{pmatrix}$. Valóban, ha $h=(h_1,h_2)\to(0,0)$ és $A=\begin{pmatrix}2 & 1\end{pmatrix}$, akkor

$$0\le\frac{\|f(a+h)-f(a)-A\cdot h\|}{\|h\|}=\frac{\left|(1+h_1)(2+h_2)-1\cdot 2-\begin{pmatrix}2 & 1\end{pmatrix}\cdot\begin{pmatrix}h_1\\ h_2\end{pmatrix}\right|}{\|h\|}=$$

$$=\frac{|2+2h_1+h_2+h_1h_2-2-(2h_1+h_2)|}{\|h\|}=\frac{|h_1h_2|}{\|h\|}\le\left(|h_1h_2|\le\frac{h_1^2+h_2^2}{2}\right)\le\frac{\frac{\|h\|^2}{2}}{\|h\|}=\frac{\|h\|}{2}\to 0,$$

és így a definícióban szereplő határérték tart nullához.

3. Tétel. *Ha $f\in D\{a\}$, akkor $f\in C\{a\}$.*

Bizonyítás. A $\|.\|$ euklideszi és a $\|x\|_\infty:=\max\{|x_i|\colon i=1,2,\dots,n\}$ normák ekvivalenciája miatt

$$f\in C\{a\} \quad\Longleftrightarrow\quad \lim_{x\to a}f(x)=f(a) \quad\Longleftrightarrow\quad \lim_{x\to a}\|f(x)-f(a)\|_\infty=0.$$

Ha $A\in\mathbb{R}^{m\times n}$ és $x\in\mathbb{R}^n$, akkor az $A\cdot x\in\mathbb{R}^m$ oszlopvektor i-edik koordinátára igaz, hogy

$$|(A\cdot x)_i|=\left|\sum_{j=1}^n a_{ij}x_j\right|\le\sum_{j=1}^n(|a_{ij}|\cdot|x_j|)\le\sum_{j=1}^n(|a_{ij}|\cdot\|x\|_\infty)=\|x\|_\infty\sum_{j=1}^n|a_{ij}|,$$

ahol a_{ij} az A mátrix i-edik sorában és j-edik oszlopában lévő eleme. Ezért

$$\|A\cdot x\|_\infty=\max\{|(A\cdot x)_i|\colon i=1,2,\dots,m\}\le\|x\|_\infty\cdot\max\left\{\sum_{j=1}^n|a_{ij}|\colon i=1,2,\dots,m\right\}.$$

Tehát

$$\alpha:=\max\left\{\sum_{j=1}^n|a_{ij}|\colon i=1,2,\dots,m\right\} \quad\Longrightarrow\quad \|A\cdot x\|_\infty\le\alpha\|x\|_\infty.$$

Legyen $f\in D\{a\}$. Ekkor a $h=x-a\to 0$ helyettesítéssel

$$0\le\|f(x)-f(a)\|_\infty=\|f(a+h)-f(a)\|_\infty=\Big\|A\cdot h+\varepsilon(h)\cdot\|h\|\Big\|_\infty\le$$

$$\le\|A\cdot h\|_\infty+\Big\|\varepsilon(h)\cdot\|h\|\Big\|_\infty\le\alpha\|h\|_\infty+\|h\|\cdot\|\varepsilon(h)\|_\infty\to 0+0=0,$$

amiből a tétel állítása következik.

Megjegyzés. Az előző tétel nem fordítható meg. Nem nehéz igazolni, hogy az $f(x)=\|x\|$ $(x\in\mathbb{R}^n)$ függvény folytonos, de nem differenciálható az $a=0$ pontban.

4. Tétel. *Legyen az $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ függvény koordinátafüggvényei $f_j \colon \mathbb{R}^n \supseteq \mathcal{D}_f \to \mathbb{R}$, ahol $j = 1, 2, \ldots, m$, illetve $a \in \operatorname{int} \mathcal{D}_f$. Ekkor*

$$f \in D\{a\} \qquad \Longleftrightarrow \qquad f_j \in D\{a\} \quad \text{minden } j = 1, 2, \ldots, m \text{ esetén.}$$

Továbbá $f_j'(a) = e_j^\top \cdot f'(a)$, azaz $f_j'(a)$ az $f'(a)$ mátrix j-edik sora.

Bizonyítás. Alkalmazzuk az

$$\|y\|_\infty \le \|y\| \le \sqrt{m}\, \|y\|_\infty \qquad (y \in \mathbb{R}^m).$$

normák ekvivalenciáját, ahol $\|x\|_\infty := \max\{|y_j| \colon j = 1, 2, \ldots, m\}$. Vegyük észre, hogy

$$(f(a+h) - f(a) - A \cdot h)_j = f_j(a+h) - f_j(a) - e_j^\top \cdot A \cdot h$$

minden $j = 1, 2, \ldots, m$. Így a normák ekvivalenciájából:

$$0 \le \frac{\max\limits_{1 \le j \le m} |f_j(a+h) - f_j(a) - e_j^\top \cdot A \cdot h|}{\|h\|} \le \frac{\|f(a+h) - f(a) - A \cdot h\|}{\|h\|}$$
$$= \sqrt{m}\, \frac{\max\limits_{1 \le j \le m} |f_j(a+h) - f_j(a) - e_j^\top \cdot A \cdot h|}{\|h\|}.$$

Ebből következik a tétel állítása, hiszen ha az egyik hányados nullához tart $h \to 0$ esetén, akkor a másik is nullához tart.

Megjegyzés. Az előző tétel szerint elég lenne valós értékű függvényekkel foglalkozni.

A totális és a parciális derivált kapcsolata

A totális derivált az erősebb.

5. Tétel. *Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ és $a \in \operatorname{int} \mathcal{D}_f$. Ha $f \in D\{a\}$, akkor $\forall v \in \mathbb{R}^n$ egységvektor esetén az f függvénynek van v irányú iránymenti deriváltja az a pontban, és*

$$\partial_v f(a) = f'(a) \cdot v$$

Bizonyítás. Ha $f \in D\{a\}$, akkor $\exists \varepsilon : \mathcal{D}_f \to \mathbb{R}$, $\lim\limits_0 \varepsilon = 0$ úgy, hogy

$$f(a+h) - f(a) = f'(a) \cdot h + \varepsilon(h) \cdot \|h\| \qquad (a + h \in \mathcal{D}_f).$$

Legyen $h = tv$ $(t \in \mathbb{R})$. Ekkor $\|h\| = |t| \cdot \|v\| = |t|$, hiszen $\|v\| = 1$, és így

$$f(a+tv) - f(a) = f'(a) \cdot tv + \varepsilon(tv) \cdot |t|.$$

Ezért

$$\partial_v f(a) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t} = f'(a) \cdot v + \lim_{t \to 0} \Big(\varepsilon(tv) \operatorname{sgn}(t)\Big) = f'(a) \cdot v + 0 = f'(a) \cdot v,$$

hiszen $\operatorname{sgn}(t)$ korlátos függvény és $\lim\limits_{t \to 0} \varepsilon(tv) = \lim\limits_{h \to 0} \varepsilon(h) = 0$.

A totális derivált definíciója nem nyújt információt a deriváltmátrix előállításáról, de az előző tétel alapján igazolni tudjuk, hogy az $f \in \mathbb{R}^n \to \mathbb{R}^m$ függvény a pontbeli deriváltmátrixa egyszerűen felírható a koordinátafüggvények a pontban vett parciális deriváltjai segítségével.

6. Tétel (A deriváltmátrix előállítása). *Legyen az $f \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ függvény koordinátafüggvényei $f_j \colon \mathbb{R}^n \supseteq \mathcal{D}_f \to \mathbb{R}$, ahol $j = 1, 2, \ldots, m$, illetve $a \in \operatorname{int} \mathcal{D}_f$. Ha $f \in D\{a\}$, akkor*

$$\exists\, \partial_i f_j(a) \quad (\forall\, i = 1, \ldots, n;\ j = 1, \ldots, m) \quad \textit{és}$$

$$f'(a) = \begin{pmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \cdots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \cdots & \partial_n f_2(a) \\ \vdots & \vdots & \vdots & \vdots \\ \partial_1 f_m(a) & \partial_2 f_m(a) & \cdots & \partial_n f_m(a) \end{pmatrix} \in \mathbb{R}^{m \times n}$$

az ún. **Jacobi-mátrix**.

Bizonyítás. A 4. Tétel szerint, ha $f \in D\{a\}$, akkor $f_j \in D\{a\}$ minden $j = 1, 2, \ldots, m$ esetén, és az $f'(a)$ mátrix j-edik sora az $f_j'(a)$ sormátrix. Alkalmazzuk az 5. tételt az f_j koordinátafüggvényekre a $v = e_i$ irányok esetén! Ekkor

$$\exists\, \partial_i f_j(a) = \partial_{e_i} f_j(a) = f'(a) \cdot e_i = a_i \qquad (i = 1, 2, \ldots, n).$$

Ez azt jelenti, hogy

$$f_j'(a) = \begin{pmatrix} \partial_1 f_j(a) & \partial_2 f_j(a) & \ldots & \partial_n f_j(a) \end{pmatrix},$$

amiből a tétel állítása következik.

A parciális deriváltak létezéséből *nem következik* a totális deriválhatóság. Például gyakorlaton fogjuk igazolni, hogy az

$$f(x, y) := \sqrt{xy} \qquad \big((x, y) \in \mathbb{R}^2\big)$$

függvény folytonos az $a = (0, 0)$ pontban, itt léteznek a parciális deriváltak, de f nem totálisan deriválható ebben a pontban.

Azonban, ha a parciális deriváltak létezésénél valamivel többet feltételezünk, akkor már tudjuk garantálni a totális deriválhatóságot. A következő tétel egy ilyen gyakran alkalmazható *elégséges feltételt* ad a függvény totális deriválhatóságára.

7. Tétel (Elégséges feltétel a totális deriválhatóságra). *Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ és $a \in \operatorname{int} \mathcal{D}_f$. Tegyük fel, hogy az a pontnak létezik olyan $K(a) \subset \mathcal{D}_f$ környezete, amelyre minden $i = 1, \ldots, n$ index esetén a következők teljesülnek:*

a) $\exists\, \partial_i f(x)$ *minden $x \in K(a)$ pontban,*

b) a $\partial_i f : K(a) \to \mathbb{R}$ parciális deriváltfüggvény folytonos az a pontban.

Ekkor az f függvény totálisan deriválható az a pontban.

Bizonyítás. A tételt nem bizonyítjuk.

Felület érintősíkja

Az egyváltozós analízisben láttuk, hogy ha $f \in D\{a\}$, akkor az f függvényt az a pont környezetében jól közelítő elsőfokú polinom nem más, mint a függvénygrafikon $\big(a, f(a)\big)$ pontbeli érintője. Most megvizsgáljuk, hogy mi felel meg ennek az állításnak az $\mathbb{R}^2 \to \mathbb{R}$ függvények körében.

Tudjuk, hogy az a pontbeli totális deriválthatóságot át lehet fogalmazni a következő módon:

$$\exists A \in \mathbb{R}^{m\times n} \quad \text{és} \quad \exists \varepsilon : \mathcal{D}_f \to \mathbb{R}^m,\ \lim_0 \varepsilon = 0 : f(a+h) - f(a) = A \cdot h + \varepsilon(h) \cdot \|h\| \quad (a+h \in \mathcal{D}_f),$$

ahol $A = f'(a)$. Ezért

$$f(a+h) - f(a) \approx f'(a) \cdot h \qquad \text{ha } h \approx 0.$$

Az $\approx$ jelölés azt jelenti, hogy két vektor távolsága (különbségük normája) kicsi.

Legyen $f \in \mathbb{R}^2 \to \mathbb{R}$, $a = (x_0, y_0) \in \operatorname{int} \mathcal{D}_f$ és $f \in D\{a\}$. Ha egy a-hoz közeli $(x, y) \in \mathcal{D}_f$ pontot felírunk $(x, y) = a + h$ alakban, akkor a fentiek szerint

$$(*) \qquad \begin{aligned} f(x,y) - f(x_0,y_0) &\approx \big(\partial_x f(x_0,y_0) \quad \partial_y f(x_0,y_0)\big) \cdot \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix} = \\ &= \partial_x f(x_0,y_0)(x - x_0) + \partial_y f(x_0,y_0)(y - y_0), \qquad \text{ha } (x,y) \approx (x_0,y_0). \end{aligned}$$

Legyen $z_0 := f(x_0, y_0)$ és tekintsük a

$$z - z_0 = \partial_x f(x_0,y_0)(x - x_0) + \partial_y f(x_0,y_0)(y - y_0)$$

egyenletű síkot. Ez egy olyan sík, ami átmegy az $\big(x_0, y_0, f(x_0, y_0)\big)$ ponton és egyik normálvektora

$$\vec{n}\big(\partial_x f(x_0,y_0), \partial_y f(x_0,y_0), -1\big).$$

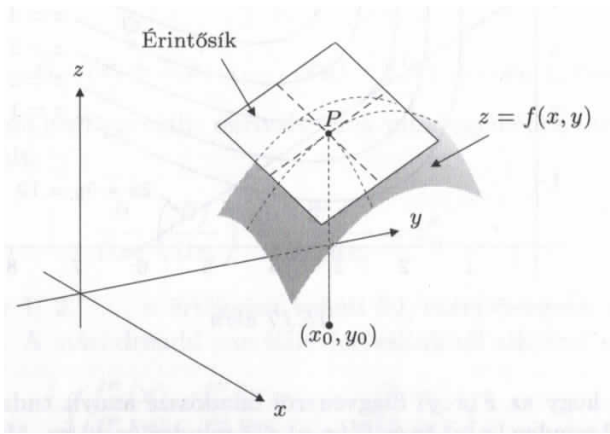
Mivel $(*)$ miatt $f(x, y) \approx z$, ha $(x, y) \approx (x_0, y_0)$, ezért érdemes a felület érintősíkját az előbbi síkkal értelmezni.

4. Definíció. *Az $f \in \mathbb{R}^2 \to \mathbb{R}$ függvény grafikonjának az $\big(x_0, y_0, f(x_0, y_0)\big)$ pontban van érintősíkja, ha $f \in D\big\{(x_0, y_0)\big\}$.* **Az érintősík egyenlete***:*

$$z - f(x_0,y_0) = \partial_x f(x_0,y_0) \cdot (x - x_0) + \partial_y f(x_0,y_0) \cdot (y - y_0),$$

amelynek egyik **normálvektora***:* $\vec{n}\big(\partial_x f(x_0,y_0), \partial_y f(x_0,y_0), -1\big)$.

Az érintősíkot az alábbi ábrán szemléltetjük:



Megjegyzés. A háromdimenziós térben a sík általános egyenlete

$$Ax + By + Cz = D,$$

ahol az A, B, C együtthatók legalább egyike nem nulla. Ekkor a

$$\vec{n}(A, B, C) \in \mathbb{R}^3$$

egy olyan vektor, amelyik a sík minden vektorára merőleges (ez a sík egyik normálvektora). Ez azért igaz, mert ha $P_0 = (x_0, y_0, z_0)$ a sík egyik pontja, akkor $P = (x, y, z)$ akkor és csak akkor van rajta a síkon, ha merőleges az $\vec{n}$ vektorra, azaz $\langle \overrightarrow{P_0P}, \vec{n} \rangle = 0$. Ebből

$$\langle \overrightarrow{P_0P}, n \rangle = A(x - x_0) + B(y - y_0) + C(z - z_0) = 0,$$

amiből átrendezéssel adódik az állítás, ha $D = Ax_0 + By_0 + Cz_0$.

Példa. Írjuk fel a $z = xy$ egyenletű felület $P_0(1, 2)$ pontjához tartozó érintősíkjának az egyenletét! Legyen $f(x, y) = xy \quad \big((x, y) \in \mathbb{R}^2\big)$ és $a = (x_0, y_0)$, ahol $x_0 = 1$ és $y_0 = 2$. Ekkor

$$\partial_x f(x, y) = y \quad \text{és} \quad \partial_y f(x, y) = x \qquad \big((x, y) \in \mathbb{R}^2\big).$$

Mivel ezek folytonos függvények, ezért f differenciálható minden pontban, azaz $f \in D\{a\}$. Másrészt

$$\partial_x f(1, 2) = 2, \quad \partial_y f(1, 2) = 1, \quad f(1, 2) = 2.$$

Ezért az érintősík egyenlete:

$$z - f(1, 2) = \partial_x f(1, 2) \cdot (x - 1) + \partial_y f(1, 2) \cdot (y - 2) \quad \Longrightarrow \quad z - 2 = 2(x - 1) + (y - 2),$$

azaz $2x + y - z = 2$.

Deriválási szabályok

$\mathbb{R}^n \to \mathbb{R}^m$ típusú vektor-vektor függvényekre hasonló deriválási szabályok érvényesek, mint valós-valós függvények esetén.

8. Tétel (Algebrai műveletekre vonatkozó deriválási szabályok).

- *Ha $f, g \in \mathbb{R}^n \to \mathbb{R}^m$ $(n, m \in \mathbb{N}^+)$ és $f, g \in D\{a\}$, akkor $\forall \alpha, \beta \in \mathbb{R}$ esetén*

$$\big(\alpha f + \beta g\big) \in D\{a\} \qquad \text{és} \qquad \big(\alpha f + \beta g\big)'(a) = \alpha f'(a) + \beta g'(a).$$

- *Ha $m = 1$, akkor az $f \cdot g$ és az f/g függvényekre az egyváltozós esethez hasonló deriválási szabályok teljesülnek.*

Bizonyítás. A tétel állításai igazolhatók követlenül a totális derivált definíciójából a valós-valós esethez hasonló átalakításokkal.

Az összetett függvény deriválhatóságára vonatkozó állítás meglepő hasonlóságot mutat a valós-valós függvényekre jól ismert állítással. De vigyázzunk, mert a többváltozós képletben mátrixszorzás szerepel.

9. Tétel (Az összetett függvény deriválhatósága). *Legyen $n, m, s \in \mathbb{N}^+$. Ha $g \in \mathbb{R}^n \to \mathbb{R}^m$ és $g \in D\{a\}$, továbbá $f \in \mathbb{R}^m \to \mathbb{R}^s$ és $f \in D\{g(a)\}$, akkor $f \circ g \in D\{a\}$ és*

$$(\#) \qquad \Big(f \circ g\Big)'(a) = f'\Big(g(a)\Big) \cdot g'(a),$$

ahol $\cdot$ a mátrixok közötti szorzás műveletét jelöli.

Bizonyítás. Valós-valós esetben a bizonyítás a derivált lineáris közelítéssel történő átfogalmazásán alapszik. $\mathbb{R}^n \to \mathbb{R}^m$ típusú függvények esetén is megvan ennek a megfelelője, így a bizonyítás lényegében hasonló, de az egyes lépéseket adaptálni kell többváltozós függvények esetére.

Megjegyzések.

1. Figyeljük meg, hogy a (#) képletben szereplő mátrixszorzat elvégezhető és az eredmény olyan típusú mátrix, mint az $\Big(f \circ g\Big)'(a)$ deriváltmátrix, hiszen

$$\Big(f \circ g\Big)'(a) \in \mathbb{R}^{s\times n}, \qquad f'\Big(g(a)\Big) \in \mathbb{R}^{s\times m}, \qquad g'(a) \in \mathbb{R}^{m\times n}.$$
$$(f \circ g \in \mathbb{R}^n \to \mathbb{R}^s) \qquad (f \in \mathbb{R}^m \to \mathbb{R}^s) \qquad (g \in \mathbb{R}^n \to \mathbb{R}^m)$$

2. Koordinátafüggvényekkel felírva az összetett függvény általános alakja:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \xrightarrow{\ g\ } y = \begin{pmatrix} y_1 := g_1(x_1, x_2, \dots, x_n) \\ y_2 := g_2(x_1, x_2, \dots, x_n) \\ \vdots \\ y_m := g_m(x_1, x_2, \dots, x_n) \end{pmatrix} \xrightarrow{\ f\ } z = \begin{pmatrix} z_1 := f_1(y_1, y_2, \dots, y_m) \\ z_2 := f_2(y_1, y_2, \dots, y_m) \\ \vdots \\ z_s := f_s(y_1, y_2, \dots, y_m) \end{pmatrix}.$$

Az összetett függvény deriválhatóságáról szóló tétel szerint, ha $g \in D\{x\}$ és $f \in D\{g(x)\}$, akkor $f \circ g \in D\{x\}$, és így léteznek az összetett függvény parciális deriváltjai az x pontban. Ekkor (#) alapján minden $i = 1, 2, \dots, n$ és $j = 1, 2, \dots, s$ rögzített indexpár esetén

$$\frac{\partial z_j}{\partial x_i}(x) = \sum_{k=1}^{m} \frac{\partial z_j}{\partial y_k}\Big(y(x)\Big) \cdot \frac{\partial y_k}{\partial x_i}(x).$$

Példa. Legyen $f\colon \mathbb{R}^3 \to \mathbb{R}$, $f(x, y, z) = x + yz$, és

$$g\colon \mathbb{R}^3 \to \mathbb{R}^3, \qquad g(x, y, z) = \begin{pmatrix} x^2 + y^2 + z^2 \\ xyz \\ x \end{pmatrix}.$$

Az f és g függvények parciális deriváltjai folytonosak minden $\mathbb{R}^3$-beli pontban, ezért $f \in D(\mathbb{R}^3)$ és $g \in D(\mathbb{R}^3)$, és így $f \circ g \in D(\mathbb{R}^3)$. Vezessük be a következő jelöléseket:

$$\begin{aligned} y_1 &:= g_1(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 \\ y_2 &:= g_2(x_1, x_2, x_3) = x_1 x_2 x_3 \qquad \text{és} \qquad z = f(y_1, y_2, y_3) = y_1 + y_2 y_3. \\ y_3 &:= g_3(x_1, x_2, x_3) = x_1 \end{aligned}$$

Legyen $F := f \circ g$. Ekkor az összetett függvény deriválhatóságáról szóló tétel szerint

$$\frac{\partial F}{\partial x_1} = \frac{\partial z}{\partial x_1} = \frac{\partial z}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_1} + \frac{\partial z}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_1} + \frac{\partial z}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_1} =$$

$$= 1 \cdot 2x_1 + y_3 \cdot x_2x_3 + y_2 \cdot 1 = 2x_1 + x_1 \cdot x_2x_3 + x_1x_2x_3 = 2x_1 + 2x_1x_2x_3,$$

$$\frac{\partial F}{\partial x_2} = \frac{\partial z}{\partial x_2} = \frac{\partial z}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_2} + \frac{\partial z}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_2} + \frac{\partial z}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_2} =$$

$$= 1 \cdot 2x_2 + y_3 \cdot x_1x_3 + y_2 \cdot 0 = 2x_2 + x_1 \cdot x_1x_3 = 2x_2 + x_1^2x_3,$$

$$\frac{\partial F}{\partial x_3} = \frac{\partial z}{\partial x_3} = \frac{\partial z}{\partial y_1} \cdot \frac{\partial y_1}{\partial x_3} + \frac{\partial z}{\partial y_2} \cdot \frac{\partial y_2}{\partial x_3} + \frac{\partial z}{\partial y_3} \cdot \frac{\partial y_3}{\partial x_3} =$$

$$= 1 \cdot 2x_3 + y_3 \cdot x_1x_2 + y_2 \cdot 0 = 2x_3 + +x_1 \cdot x_1x_2 = 2x_3 + x_1^2x_2.$$

Nem nehéz ellenőrizni a kapott eredmények helyességét, hiszen

$$f(g(x_1, x_2, x_3)) = z = y_1 + y_2y_3 = (x_1^2 + x_2^2 + x_3^2) + (x_1x_2x_3) \cdot x_1 = x_1^2 + x_2^2 + x_3^2 + x_1^2x_2x_3.$$

Az eredeti jelölés alkalmazásával:

$$\frac{\partial(f \circ g)}{\partial x} = 2x + 2xyz, \quad \frac{\partial(f \circ g)}{\partial y} = 2y + x^2z, \quad \frac{\partial(f \circ g)}{\partial z} = 2z + x^2y,$$

Gradiens vektor

A 3. előadáson megismertük az $\mathbb{R}^n \to \mathbb{R}$ típusú függvények iránymenti deriváltját. Egy olyan tételt mondtunk ki, amely szerint bizonyos feltételek mellett a $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$ egységvektor irányú iránymenti derivált az a pontban kiszámolható a parciális deriváltak segítségével a következő módon:

$$(\star) \qquad \partial_v f(a) = \sum_{i=1}^{n} \partial_i f(a) \cdot v_i.$$

Ehhez az kellet, hogy létezzenek az f függvény parciális deriváltjai egy $K(a) \subset \mathcal{D}_f$ környezetben, és folytonosak legyenek az a pontban. Ez éppen az elégséges feltétel a totális deriválhatóságra, tehát $f \in D\{a\}$. De ekkor $(\star)$ következik a 4. előadás 4. tételéből, ahol azt igazoltuk, hogy ha f totális deriválható az a pontban, akkor ott bármilyen v irányban van iránymenti deriváltja, és

$$\partial_v f(a) = f'(a)v.$$

Ez azért igaz, mert valós értékű függvények esetében

$$f'(a) = \begin{pmatrix} \partial_1 f(a) & \partial_1 f(a) & \ldots & \partial_n f(a) \end{pmatrix},$$

és így $(\star)$ adódik a fenti sormátrix és v koordinátaiból álló oszlopmátrix szorzatából. Érdekes, hogy $(\star)$ is igazolható az összetett függvény deriválhatóságáról szóló tétel alapján. Valóban, a definíció szerint $\partial_v f(a)$ nem más, mint az $F_v(t) = f(a+tv)$ függvény deriváltja a $t = 0$ pontban. De $F_v = f \circ g$, ahol

$$g \colon \mathbb{R} \supseteq K(0) \ni t \mapsto a + tv \in \mathbb{R}^n.$$

Látható, hogy $g(0) = a$, illetve nem nehéz igazolni, hogy $g'(0) = v$ (g konstans + lineáris függvény). Ezért

$$\partial_v f(a) = F_v'(0) = \Big(f \circ g\Big)'(0) = f'\Big(g(0)\Big) \cdot g'(0) = f'(a)v.$$

A ($\star$) képletből következik, hogy az összes v egységvektor közül az, amelynek iránya olyan, mint a

$$\operatorname{grad} f(a) := \Big(\partial_1 f(a), \partial_2 f(a), \dots, \partial_n f(a)\Big)$$

(ún. ***gradiens vektor***) vektor iránya, azzal a tulajdonsággal rendelkezik, hogy az ebbe az irányba mutató iránymenti derivált a legnagyobb. A szóban forgó egységvektor:

$$w := \frac{\operatorname{grad} f(a)}{\|\operatorname{grad} f(a)\|} = \left(\frac{\partial_1 f(a)}{\|\operatorname{grad} f(a)\|}, \frac{\partial_2 f(a)}{\|\operatorname{grad} f(a)\|}, \dots, \frac{\partial_n f(a)}{\|\operatorname{grad} f(a)\|}\right).$$

Ezért ($\star$) szerint

$$\max\{\partial_v f(a) \colon v \in \mathbb{R}^n, \|v\| = 1\} = \partial_w f(a) = \sum_{i=1}^{n} \left(\partial_i f(a) \cdot \frac{\partial_i f(a)}{\|\operatorname{grad} f(a)\|}\right) =$$

$$= \frac{1}{\|\operatorname{grad} f(a)\|} \sum_{i=1}^{n} \Big(\partial_i f(a)\Big)^2 = \frac{1}{\|\operatorname{grad} f(a)\|} \cdot \|\operatorname{grad} f(a)\|^2 = \|\operatorname{grad} f(a)\|.$$

Megjegyzés. A gradiens irányában a leggyorsabb a függvény növekedése, ellentétes irányban a leggyorsabb a csökkenése. A gradiensre merőleges irányban az iránymenti derivált nulla.

10. előadás

TÖBBVÁLTOZÓS FÜGGVÉNYEK DIFFERENCIÁLSZÁMÍTÁSA 2.

Magasabb rendű deriváltak

Az $f \in \mathbb{R} \to \mathbb{R}$ függvények körében „nem okozott gondot” a 2-szer, 3-szor, ... való differenciálhatóság teljes indukcióval történő értelmezése. Először a kétszer deriválhatóság fogalmát definiáltuk. Azt mondtuk, hogy a szóban forgó függvény *kétszer differenciálható az* $a \in \operatorname{int} \mathcal{D}_f$ *pontban* (röviden $f \in D^2\{a\}$), ha f az a pont egy $K(a) \subseteq \mathcal{D}_f$ környezetében deriválható (f' értelmezve van a $K(a)$ halmazon) és f' differenciálható az a pontban, azaz $f' \in D\{a\}$. Ekkor az $f''(a) := \big(f'\big)'(a)$ számot az f függvény *a pontbeli második deriváltjának neveztük.* Teljes indukcióval hasonlóan értelmeztük a 2-nél magasabb rendű deriválhatóság fogalmát.

A kétszeri differenciálhatóságnak ez az értelmezése átvihető az $f \in \mathbb{R}^n \to \mathbb{R}$ függvényekre akkor is, ha $n > 1$. Ti., ha $a \in \operatorname{int} \mathcal{D}_f$ és $\exists\, K(a) \subseteq \mathcal{D}_f \colon f \in D\big(K(a)\big)$, akkor az

$$f'(x) = \begin{pmatrix} \partial_1 f(x) & \partial_2 f(x) & \dots & \partial_n f(x) \end{pmatrix} \qquad \big(x \in K(a)\big),$$

deriváltmátrixok és a $\operatorname{grad} f(x) = \big(\partial_1 f(x), \partial_2 f(x), \dots, \partial_n f(x)\big)$ gradiens vektorok segítségével értelmezhető az

$$f' \colon \mathbb{R}^n \supseteq K(a) \ni x \mapsto \operatorname{grad} f(x) \in \mathbb{R}^n$$

deriváltfüggvény. Továbbá az $\mathbb{R}^n \to \mathbb{R}^n$ vektor-vektor függvények körében értelmeztük a deriválhatóságot. Tehát az $f \in \mathbb{R}^n \to \mathbb{R}$ függvényre minden további nélkül előírhatjuk, hogy $f' = \operatorname{grad} f \in D\{a\}$ teljesüljön. Ekkor

$$f''(a) := \big(f'\big)'(a) = \big(\operatorname{grad} f\big)'(a)$$

az f függvény második deriváltja az a pontban. Világos, hogy egy $f \in \mathbb{R}^n \to \mathbb{R}$ függvény esetén $f''(a)$ egy $(n \times n)$-es mátrix. Az is nyilvánvaló azonban, hogy ezen az úton nem tudunk eljutni a 2-nél magasabb rendű deriválhatósághoz, hiszen pl. $\mathbb{R}^n \to \mathbb{R}^{n \times n}$ típusú függvények differenciálhatóságát nem értelmeztük.

Emlékeztetünk ugyanakkor arra, hogy egy függvény akkor és csak akkor differenciálható egy pontban, ha a koordinátafüggvényei differenciálhatók ebben a pontban. Mivel a $\operatorname{grad} f$ függvény koordinátafüggvényei az f függvény parciális deriváltjai, így

$$f' = \operatorname{grad} f = \big(\partial_1 f, \partial_2 f, \dots, \partial_n f\big) \in D\{a\} \qquad \Longleftrightarrow \qquad \partial_i f \in D\{a\} \quad (i = 1, 2, \dots, n).$$

Ez az észrevétel lehetőséget ad arra, hogy értelmezni tudjuk egy $f \in \mathbb{R}^n \to \mathbb{R}$ függvény magasabb rendű deriválhatóságát. Az értelmezés teljes indukcióval történik, első lépésként a kétszer deriválhatóságot definiálva.

1. Definíció. *Az* $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ *függvény* **kétszer deriválható** (*vagy* **differenciálható**) *az* $a \in \operatorname{int} \mathcal{D}_f$ *pontban* (*jelben:* $f \in D^2\{a\}$), *ha*

a) $\exists K(a) \subset \mathcal{D}_f$, *hogy* $f \in D\{x\}$ *minden* $x \in K(a)$ *pontban, és*

b) $\forall i = 1, 2, \dots, n$ *indexre* $\partial_i f \in D\{a\}$.

Az a) feltételt röviden úgy is írhatjuk, hogy a $\exists K(a) \subseteq \mathcal{D}_f$, hogy $f \in D\Big(K(a)\Big)$. Ebből következik, hogy $K(a)$ környezetben létezik az

$$f' = \Big(\partial_1 f, \partial_2 f, \dots, \partial_n f\Big) = \operatorname{grad} f \in \mathbb{R}^n \to \mathbb{R}^n$$

deriváltfüggvény, ami tehát már egy vektor-vektor függvény.

A b) feltétel pedig azzal ekvivalens, hogy $f' \in D\{a\}$. Így minden $i = 1, 2, \dots, n$ esetén a $\partial_i f \in \mathbb{R}^n \to \mathbb{R}$ parciális deriváltfüggvényeknek léteznek az a pontban mindegyik változó szerinti parciális deriváltjai:

$$\partial_j\Big(\partial_i f\Big)(a) \quad (j = 1, 2, \dots, n).$$

Ezeket a számokat az ***f* függvény *a* pontbeli, *i*-edik és *j*-edik változó szerinti másodrendű** (vagy ***második***) ***parciális deriváltjának*** nevezzük.

Az f függvény a ***pontbeli második deriváltját*** így értelmezzük:

$$f''(a) := \Big(\operatorname{grad} f\Big)'(a).$$

Világos, hogy egy $f \in \mathbb{R}^n \to \mathbb{R}$ függvény esetén $f''(a)$ egy $(n \times n)$-es mátrix.

2. Definíció. *Ha az $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ függvény kétszer deriválható az $a \in \operatorname{int} \mathcal{D}_f$ pontban, akkor*

$$f''(a) = \begin{pmatrix} \partial_{11} f(a) & \partial_{12} f(a) & \dots & \partial_{1n} f(a) \\ \partial_{21} f(a) & \partial_{22} f(a) & \dots & \partial_{2n} f(a) \\ \vdots & \vdots & \vdots & \vdots \\ \partial_{n1} f(a) & \partial_{n2} f(a) & \dots & \partial_{nn} f(a) \end{pmatrix} \in \mathbb{R}^{n \times n}$$

az ***f* függvény *a* pontbeli Hesse-féle mátrixa**, *ahol*

$$\partial_{ij} f(a) = \partial_j\Big(\partial_i f\Big)(a) \quad \Big(i = 1, 2, \dots, n;\ j = 1, 2, \dots, n\Big).$$

A 2-nél magasabb rendű deriválhatóságot teljes indukcióval így értelmezzük: az $f \in \mathbb{R}^n \to \mathbb{R}$ $(2 \le n \in \mathbb{N})$ függvény ***s*-szer $(2 \le s \in \mathbb{N})$ *deriválható*** az $a \in \operatorname{int} \mathcal{D}_f$ pontban (jelben $f \in D^s\{a\}$), ha

a) $\exists K(a) \subseteq \mathcal{D}_f$ környezet, hogy $f \in D^{s-1}\Big(K(a)\Big)$ és

b) minden $(s-1)$-edrendű $\partial_{i_1} \partial_{i_2} \dots \partial_{i_{s-1}} f$ $\Big(1 \le i_1, i_2, \dots, i_{s-1} \le n\Big)$ parciális deriváltfüggvény deriválható az a pontban.

Az előbbihez kapcsolódik egy fontos fogalom: azt mondjuk, hogy az $f \in \mathbb{R}^n \to \mathbb{R}$ függvény ***s*-szer $(s \in \mathbb{N}^+)$ *folytonosan deriválható*** az $a \in \operatorname{int} \mathcal{D}_f$ pontban (jelben $f \in C^s\{a\}$), ha

a) $\exists K(a) \subseteq \mathcal{D}_f$ környezet, hogy $f \in D^s\Big(K(a)\Big)$ és

b) minden s-edrendű $\partial_{i_1} \partial_{i_2} \dots \partial_{i_s} f$ $\Big(1 \le i_1, i_2, \dots, i_s \le n\Big)$ parciális deriváltfüggvény folytonos az a pontban.

Megjegyzés. A totális deriválhatóságra vonatkozó elégséges feltétel azt jelenti, hogy f (egyszer) folytonosan deriválható az a pontban.

A magasabb rendű parciális deriváltakkal kapcsolatban joggal merül fel a *kérdés*, hogy azok kiszámításakor van-e szerepe a változók sorrendjének? A következő tétel azt állítja, hogy ha az f függvény a szóban forgó a helyen „elég sokszor” deriválható, akkor a sorrend elveszti a jelentőségét.

1. Tétel (Young-tétel). *Ha $f \in \mathbb{R}^n \to \mathbb{R}$ $(2 \le n \in \mathbb{N})$ és $f \in D^2\{a\}$, akkor*

$$\partial_{ij} f(a) = \partial_{ji} f(a) \quad \forall\, i, j = 1, \dots, n \quad \textit{indexre.}$$

Bizonyítás. A tételt nem bizonyítjuk.

Példa. Az $f \in D^2\{a\}$ feltétel hiánya esetén a parciális deriváltak képzésének a sorrendje általában nem cserélhető fel. Ha például

$$f(x,y) := \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2}, & \text{ha } (x,y) \in \mathbb{R}^2 \setminus \{(0,0)\} \\ 0, & \text{ha } (x,y) = (0,0), \end{cases}$$

akkor

$$\partial_{12} f(0,0) = -1, \qquad \partial_{21} f(0,0) = 1.$$

Érdemes meggondolni azt, hogy $f \notin D^2\{a\}$ miért igaz.

Megjegyzések.

1. A Young-tételből következik, hogy ha $f \in D^2\{a\}$, akkor az f függvény a pontbeli Hesse-féle mátrixa *szimmetrikus* mátrix.

2. Teljes indukcióval igazolható a Young-tétel következő általánosítása: *Tegyük fel, hogy az $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N})$ függvény s-szer $(s \in \mathbb{N})$ differenciálható az $a \in \operatorname{int} \mathcal{D}_f$ pontban. Legyen $i_k \in \mathbb{N}$, $1 \le i_k \le n$ $(k = 1, 2, \dots, s)$ és $j_1, j_2, \dots, j_s$ az $i_1, i_2, \dots, i_s$ indexek egy tetszőleges permutációja. Ekkor*

$$\partial_{i_1} \partial_{i_2} \cdots \partial_{i_s} f(a) = \partial_{j_1} \partial_{j_2} \cdots \partial_{j_s} f(a).$$

Taylor-polinomok

Induljunk ki az egyváltozós ismereteinkből. *Motivációként* akkor azt a fontos *problémát* vetettük fel, hogy egy adott „bonyolult” függvényt vajon meg lehet-e közelíteni egyszerű szerkezetű függvényekkel, például a jól kezelhető és könnyen számolható polinomokkal.

A valós-valós esetben beláttuk, hogy egy függvény akkor és csak akkor differenciálható egy pontban, ha annak a pontnak egy környezetében a függvény lokálisan jól közelíthető elsőfokú polinommal. Ennek általánosításaként jutottunk el ahhoz a fontos eredményhez, hogy ha egy függvény m-szer deriválható a szóban forgó pont egy környezetében, akkor ebben a környezetben a függvény jól közelíthető egy alkalmasan megválasztott $(m-1)$-edfokú polinommal, nevezetesen a függvény adott ponthoz tartozó *Taylor-polinomjával.* Ezzel kapcsolatos alapvető eredményünk volt a *Taylor-formulára* vonatkozó alábbi állítás:

Ha $f \in \mathbb{R} \to \mathbb{R}$, $m \in \mathbb{N}$ és egy $K(a) \subseteq \mathcal{D}_f$ környezetben $f \in D^m\big(K(a)\big)$, akkor minden $h > 0$ $(a + h \in K(a))$ számhoz létezik olyan $\nu \in (0,1)$ szám, amelyre az

$$f(a+h) = \sum_{k=0}^{m-1} \frac{f^{(k)}(a)}{k!} h^k + \frac{f^{(m)}(a+\nu h)}{m!} h^m \tag{1}$$

egyenlőség teljesül.

A legfeljebb $(m-1)$-edfokú

$$T_{m-1,a}f(h) := \sum_{k=0}^{m-1} \frac{f^{(k)}(a)}{k!} h^k \qquad (h \in \mathbb{R})$$

polinom az f függvény *a ponthoz tartozó* $(m-1)$*-edik Taylor-polinomja*, az (1) egyenlőség jobb oldalán álló összeg utolsó tagja pedig a *Taylor-formula Lagrange-féle maradéktagja*. Az elnevezést az a tény indokolja, hogy ez a tag $T_{m-1,a}f(h)$-hoz képest kicsi, ha h közel van 0-hoz.

Figyeljük meg, hogy a fenti egyváltozós eredmény $f \in \mathbb{R}^n \to \mathbb{R}$ függvényekre vonatkozó általánosításához *egyrészt* értelmezni kellene $h \in \mathbb{R}^n$ esetén a h^k hatványokat, *másrészt* az $f^{(k)}(a)$ deriváltakat, ha $k = 1, 2, \ldots, m$.

A továbbiakban ezt a kiterjesztést csak az $n = 2$ és az $m = 2$ speciális esetben igazoljuk. Legyen tehát $f \in \mathbb{R}^2 \to \mathbb{R}$, $a = (a_1, a_2) \in \operatorname{int} \mathcal{D}_f$ és $h = (h_1, h_2) \in \mathbb{R}^2$. Tegyük fel, hogy $f \in C^2\{a\}$, azaz egy $K(a) \subseteq \mathcal{D}_f$ környezetben az f függvény kétszer deriválható, azaz $f \in D^2\Big(K(a)\Big)$, illetve a másodrendű parciális deriváltak folytonosak az a pontban. Vegyük a síkon az a és az $a + h \in K(a)$ pontokat összekötő egyenes $a + th \;\; (t \in \mathbb{R})$ pontjait. Tekintsük a

$$F(t) := f(a + th) \qquad (t \in \mathbb{R})$$

valós-valós függvényt. Mivel $f \in D^2\Big(K(a)\Big)$, ezért az összetett függvény deriválására vonatkozó tételből következik, hogy a F és a F' függvény folytonos a $[0, 1]$ intervallumon és F' differenciálható $(0, 1)$-en.

Alkalmazzuk a F függvényre az (1) alatti Taylor-formulát a $[0, 1]$ intervallumon: létezik tehát olyan $\nu \in (0, 1)$, hogy

$$F(1) = F(0) + F'(0) + \frac{1}{2!} F''(\nu).$$

Most kiszámítjuk az $F(0)$, az $F'(0)$ és az $F''(\nu)$ értékeket. Egyrészt

$$F(0) = f(a + 0 \cdot h) = f(a).$$

Másrészt az összetett függvény deriválására vonatkozó tétel alapján

$$F'(t) = f'(a + th) \cdot h = \partial_1 f(a + th)h_1 + \partial_2 f(a + th)h_2 = \langle f'(a + th), h \rangle,$$

következésképpen

$$F'(0) = \langle f'(a), h \rangle.$$

A F függvény második deriváltja:

$$\begin{aligned} F''(t) &= (\partial_1 f(a + th)h_1 + \partial_2 f(a + th)h_2)' = \Big[(\partial_1 f)'(a + th) \cdot h\Big]h_1 + \Big[(\partial_2 f)'(a + th) \cdot h\Big]h_2 = \\ &= \Big(\partial_{11} f(a + th)h_1 + \partial_{12} f(a + th)h_2\Big)h_1 + \Big(\partial_{21} f(a + th)h_1 + \partial_{22} f(a + th)h_2\Big)h_2 = \\ &= \partial_{11} f(a + th)h_1^2 + \partial_{12} f(a + th)h_1 h_2 + \partial_{21} f(a + th)h_1 h_2 + \partial_{22} f(a + th)h_2^2 = \\ &= \langle f''(a + th) \cdot h, h \rangle. \end{aligned}$$

(Az utolsó egyenlőséget gondoljuk végig a mátrixszorzás és a skaláris szorzat definícióinak birtokában. Itt $f''(a + th) = \begin{pmatrix} \partial_{11} f(a + th) & \partial_{12} f(a + th) \\ \partial_{21} f(a + th) & \partial_{22} f(a + th) \end{pmatrix}$ egy (2×2)-es mátrix, $h = \begin{pmatrix} h_1 \\ h_2 \end{pmatrix}$ oszlopvektor, tehát $f''(a + th) \cdot h \in \mathbb{R}^{2\times 1} \approx \mathbb{R}^2$ egy oszlopvektor.)

Így $F''(\nu) = \langle f''(a + \nu h) \cdot h, h \rangle$.

Mivel $f \in C^2\{a\}$, ezért a másodrendű parciális deriváltak folytonos az a pontban, és így

$$f''(a+\nu h) = \begin{pmatrix} \partial_{11} f(a+\nu h) & \partial_{12} f(a+\nu h) \\ \partial_{21} f(a+\nu h) & \partial_{22} f(a+\nu h) \end{pmatrix} = \begin{pmatrix} \partial_{11} f(a) + \eta_{11}(h) & \partial_{12} f(a) + \eta_{12}(h) \\ \partial_{21} f(a) + \eta_{21}(h) & \partial_{22} f(a) + \eta_{22}(h) \end{pmatrix} =$$

$$= f''(a) + \eta(h),$$

ahol $\eta = \left[\eta_{ij}\right] \in \mathbb{R}^2 \to \mathbb{R}^{2\times 2}$ egy olyan függvény, amelyre $\lim\limits_{0} \eta = \mathbf{0}$.

A fentiek alapján tehát

$$F(1) = f(a+h) = f(a) + \langle f'(a), h\rangle + \tfrac{1}{2!}\langle f''(a)\cdot h, h\rangle + \tfrac{1}{2!}\langle \eta(h)\cdot h, h\rangle.$$

Az utolsó tagot még tovább alakítjuk:

$$\langle \eta(h)\cdot h, h\rangle = \sum_{i,j=1}^{2} \eta_{ij}(h)\cdot h_i h_j = \|h\|^2 \cdot \sum_{i,j=1}^{2} \eta_{ij}(h)\cdot \frac{h_i h_j}{\|h\|^2}.$$

Mivel $\dfrac{|h_i h_j|}{\|h\|^2} \le 1$ $(i,j=1,2)$ és $\lim\limits_{0} \eta_{ij} = 0$, ezért

$$\tfrac{1}{2!}\langle \eta(h)\cdot h, h\rangle = \varepsilon(h)\cdot \|h\|^2,$$

ahol $\varepsilon \in \mathbb{R}^2 \to \mathbb{R}$ olyan függvény, amelyre $\lim\limits_{0} \varepsilon = 0$ teljesül.

Az eddigieket összefoglalva azt kaptuk, hogy van olyan $\varepsilon \in \mathbb{R}^2 \to \mathbb{R}$, a $\lim\limits_{0} \varepsilon = 0$ feltételt kielégítő függvény, amelyre

$$f(a+h) = f(a) + \langle f'(a), h\rangle + \tfrac{1}{2}\langle f''(a)h, h\rangle + \varepsilon(h)\|h\|^2. \tag{2}$$

A fenti gondolatmenetet követve viszonylag egyszerűen belátható, hogy a (2) képlet tetszőleges $f \in \mathbb{R}^n \to \mathbb{R}$ $(2 < n \in \mathbb{N})$ függvényre is teljesül.

2. Tétel (A Taylor-formula a Peano-féle maradéktaggal). *Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ és tegyük fel, hogy $a \in \operatorname{int} \mathcal{D}_f$ és $f \in C^2\{a\}$. Ekkor van olyan $\varepsilon \in \mathbb{R}^n \to \mathbb{R}$, a $\lim\limits_{0} \varepsilon = 0$ feltételnek eleget tevő függvény, hogy*

$$f(a+h) = f(a) + \langle f'(a), h\rangle + \tfrac{1}{2}\langle f''(a)h, h\rangle + \varepsilon(h)\cdot\|h\|^2 \quad \Big(h \in \mathbb{R}^n,\ a+h \in \mathcal{D}_f\Big).$$

A formula jobb oldala első három tagjának az összege egy n-változós legfeljebb másodfokú polinom, amit az ***f függvény a ponthoz tartozó második Taylor-polinomjának*** nevezünk:

$$T_{2,a}f(h) := f(a) + \langle f'(a), h\rangle + \tfrac{1}{2}\langle f''(a)h, h\rangle = f(a) + \sum_{i=1}^{n} \partial_i f(a) h_i + \frac{1}{2}\sum_{i,j=1}^{n} \partial_{ij} f(a) h_i h_j.$$

A formulában szereplő $\varepsilon(h)\cdot\|h\|^2$ tagot a ***Taylor-formula Peano-féle maradéktagjának*** nevezzük.

Megjegyzések.

1. A 2. Tétel jelentősége egyrészt abban van, hogy a felhasználásával „bonyolult” n-változós valós értékű függvények helyettesítési értékeire lehet „jó” közelítő értékeket adni. Másrészt, azt a következő órán látni fogjuk, hogy a 2. Tétel fontos szerepet játszik $\mathbb{R}^n \to \mathbb{R}$ típusú függvények szélsőérték-problémáinak a vizsgálatánál.

2. A 2. Tétel általánosítható magasabb rendű Taylor-polinomokra is.

Példa. Számítsuk ki az

$$f(x,y) = e^x \cos y \qquad \big((x,y) \in \mathbb{R}^2\big)$$

függvény $a = (0,0)$ pont körüli második Taylor polinomját! A

$$T_{2,a}f(h) = f(a) + \sum_{i=1}^{n} \partial_i f(a) h_i + \frac{1}{2} \sum_{i,j=1}^{n} \partial_{ij} f(a) h_i h_j$$

képletben legyen $a = (0,0)$ és $(x,y) = a + h = (h_1, h_2)$, azaz $h_1 = x$ és $h_2 = y$. Mivel

$$\begin{aligned}
&f(x,y) = e^x \cos y && \to && f(0,0) = 1,\\
&\partial_1 f(x,y) = e^x \cos y, \quad \partial_2 f(x,y) = -e^x \sin y && \to && \partial_1 f(0,0) = 1, \quad \partial_2 f(0,0) = 0,\\
&\partial_{11} f(x,y) = e^x \cos y, \quad \partial_{12} f(x,y) = -e^x \sin y && \to && \partial_{11} f(0,0) = 1, \quad \partial_{22} f(0,0) = 0,\\
&\partial_{21} f(x,y) = -e^x \sin y, \quad \partial_{22} f(x,y) = -e^x \cos y && \to && \partial_{21} f(0,0) = 0, \quad \partial_{22} f(0,0) = -1,
\end{aligned}$$

így a keresett Taylor polinom:

$$T_{2,a}f(x,y) = 1 + 1 \cdot x + 0 \cdot y + \frac{1}{2} \cdot \left(1 \cdot x^2 + 2 \cdot 0 \cdot xy - 1 \cdot y^2\right) = \frac{x^2}{2} - \frac{y^2}{2} + x + 1.$$

$\mathbb{R}^n \to \mathbb{R}$ típusú függvények szélsőértékei

Amint azt már az „egyváltozós analízisben” is hangsúlyoztuk, a matematikai alkalmazások egyik legfontosabb fejezete a függvények szélsőértékeinek a vizsgálata. Valós-valós függvényeknél megismerkedtünk az abszolút- és a lokális szélsőértékek fogalmával, a lokális szélsőértékekre vonatkozó szükséges feltétellel, valamint több elégséges feltétellel. Most ezeket az ismereteket terjesztjük ki $\mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ típusú függvényekre. Az egyváltozós esetben bevezetett fogalmak minden további nehézség nélkül átvihetők a többváltozós függvényekre.

3. Definíció. *Azt mondjuk, hogy az $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ függvénynek az $a \in \mathcal{D}_f$ pontban* **abszolút maximuma van** *(vagy másképp fogalmazva az a pont az f függvénynek* **abszolút maximumhelye***), ha*

$$\forall x \in \mathcal{D}_f \colon f(x) \le f(a).$$

Ekkor az $f(a)$ függvényértéket az f függvény **abszolút maximumának** *nevezzük.*

Analóg módon értelmezzük az ***abszolút minimumhely*** és az ***abszolút minimum*** fogalmát.

Az abszolút maximumhelyet, illetve az abszolút minimumhelyet közösen ***abszolút szélsőértékhelynek***, az abszolút maximumot, illetve az abszolút minimumot közösen ***abszolút szélsőértéknek*** nevezzük. Minden további nehézség nélkül definiálhatjuk ezeknek a fogalmaknak a *lokális* változatait.

4. Definíció. *Azt mondjuk, hogy az $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ függvénynek az $a \in \operatorname{int} \mathcal{D}_f$ pontban* **lokális maximuma van** *(vagy másképp fogalmazva az a pont az f függvénynek* **lokális maximumhelye***), ha van olyan $K(a) \subseteq \mathcal{D}_f$ környezet, hogy*

$$\forall x \in K(a) \colon f(x) \le f(a).$$

Ekkor az $f(a)$ függvényértéket az f függvény **lokális maximumának** *nevezzük.*

Analóg módon értelmezzük az ***lokális minimumhely*** és az ***lokális minimum*** fogalmát.

A lokális maximumhelyet, illetve a lokális minimumhelyet közösen ***lokális szélsőértékhelynek***, a lokális maximumot, illetve a lokális abszolút minimumot közösen ***lokális szélsőértéknek*** nevezzük.

Elsőrendű szükséges a lokális szélsőértékre

A valós-valós függvények lokális szélsőértékeire vonatkozó szükséges feltétel lényeges nehézség nélkül átvihető az $\mathbb{R}^n \to \mathbb{R}$ típusú függvényekre.

3. Tétel (Elsőrendű szükséges feltétel a lokális szélsőértékre). *Tegyük fel, hogy* $f \in \mathbb{R}^n \to \mathbb{R} \quad (n \in \mathbb{N}^+)$ *és* $a \in \operatorname{int} \mathcal{D}_f$. *Továbbá*

- $f \in D\{a\}$ *és*
- *az* f *függvénynek az* a *pontban lokális szélsőértéke van.*

Ekkor $f'(a) = 0$, *azaz*

$$f'(a) = \begin{pmatrix} \partial_1 f(a) & \partial_2 f(a) & \dots & \partial_n f(a) \end{pmatrix} = \begin{pmatrix} 0 & 0 & \dots & 0 \end{pmatrix}.$$

Bizonyítás. Legyen $i = 1, 2, \dots, n$ rögzített és tekintsük meg a

$$G_i(t) = f\Big(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n\Big) \qquad \Big(t \in K(a_i)\Big)$$

parciális függvényt! Tudjuk, hogy $\exists \partial_i f(a) = G_i'(a_i)$, hiszen $f \in D\{a\}$. Minden $K(a) \in \mathcal{D}_f$ környezethez van olyan $r > 0$ szám, amire igaz, hogy

$$\Big(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_n\Big) \in K(a) \qquad \Big(t \in (a_i - r, a_i + r)\Big).$$

Ezért, ha az f függvénynek az a pontban lokális szélsőértéke van, akkor a G_i függvénynek az a_i pontban is lokális szélsőértéke van, mivel $G_i(a_i) = f(a)$. Ekkor a valós-valós függvényeknél tanult, a lokális szélsőértékre vonatkozó elsőrendű szükséges feltétel szerint $G_i'(a_i) = 0$, ami éppen azt jelenti, hogy $\partial_i f(a) = 0$.

5. Definíció. *Az* $a \in \operatorname{int} \mathcal{D}_f$ *pont az* $f \in \mathbb{R}^n \to \mathbb{R}$ *függvény* **stacionárius pontja**, *ha* $f \in D\{a\}$ *és* $f'(a) = 0$.

Megjegyzések.

1. A tétel tehát azt állítja, hogy a lokális szélsőértékhelyek szükségképpen a függvény stacionárius pontjai. Az $f'(a) = 0$ azonban csak *szükséges*, de *nem elégséges* feltétel a lokális szélsőértékre. Az $n = 1$ esetben például már ismerjük az $f(x) := x^3 \quad (x \in \mathbb{R})$ függvényt, amelynek az $a = 0$ stacionárius pontjában nincs lokális szélsőértéke. Az $n = 2$ esetben az $f(x) := x^2 - y^2 \quad \big((x, y) \in \mathbb{R}^2)\big)$ függvénynek az $a = (0, 0)$ pont stacionárius pontja, de itt sincs lokális szélsőérték.

2. A tétel értelmében egy $f \in \mathbb{R}^n \to \mathbb{R}$ differenciálható függvény stacionárius pontjainak megkeresésére szükséges megoldani a

$$\begin{aligned} \partial_1 f(x_1, \dots, x_n) &= 0, \\ \partial_2 f(x_1, \dots, x_n) &= 0, \\ &\vdots \\ \partial_n f(x_1, \dots, x_n) &= 0 \end{aligned}$$

egyenletrendszert. Csak az így kapott $(x_1, \dots, x_n)$ pontok *lehetnek* az f függvény lokális szélsőértékhelyei.

Másodrendű elégséges feltétel a lokális szélsőértékre

A fentiek alapján a stacionárius pontok között lehetnek olyanok, amelyekben a függvénynek nincs lokális szélsőértéke. Fontos kérdés tehát annak eldöntése, hogy egy stacionárius hely vajon lokális szélsőértékhely-e. Ennek eldöntéséhez a valós-valós esetben az elsőrendű- vagy a másodrendű elégséges feltételt használtuk. Az elsőrendű elégséges feltételhez nem tudunk megfelelő állítást kimondani többváltozós függvények esetén. A másodrendű elégséges feltétel azt mondja ki, hogy ha $f \in D^2\{a\}$, $f'(a) = 0$ és $f''(a) > 0$ (illetve $f''(a) < 0$), akkor az f függvénynek az a pontban lokális minimuma (illetve lokális maximuma) van. Ezt fogjuk általánosítani $\mathbb{R}^n \to \mathbb{R}$ típusú függvényekre.

A többváltozós esetben a kiindulópontunk *alapötlete* a Peano-féle maradéktagos Taylor-formula alkalmazása. Idézzük fel ezt az állítást: *Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$ és tegyük fel, hogy $a \in \operatorname{int} \mathcal{D}_f$ és $f \in C^2\{a\}$. Ekkor van olyan $\varepsilon \in \mathbb{R}^n \to \mathbb{R}$, a $\lim\limits_{0} \varepsilon = 0$ feltételnek eleget tevő függvény, hogy*

$$f(a+h) = f(a) + \langle f'(a), h\rangle + \tfrac{1}{2}\langle f''(a)h, h\rangle + \varepsilon(h) \cdot \|h\|^2 \qquad \Big(h \in \mathbb{R}^n,\ a+h \in \mathcal{D}_f\Big).$$

Ha az a pont az f függvény stacionárius pontja, azaz $f'(a) = 0$, akkor $\langle f'(a), h\rangle = 0$ minden $h \in \mathbb{R}^n$ esetén. Vizsgájuk meg a

$$\mathbb{R}^n \ni h \mapsto \langle f''(a)h, h\rangle \in \mathbb{R}$$

függvényt. Emlékezzük arra, hogy ha $f \in D^2\{a\}$, akkor a Young-tétel miatt az $f''(a) \in \mathbb{R}^{n\times n}$ másodrendű parciális deriváltakat tartalmazó Hesse-féle mátrix szimmetrikus.

6. Definíció. *Legyen $n \in \mathbb{N}^+$ és $A = [a_{ij}] \in \mathbb{R}^{n\times n}$ egy szimmetrikus mátrix. Ekkor a*

$$Q(h) := \langle A \cdot h, h\rangle = \sum_{i,j=1}^{n} a_{ij} h_i h_j \in \mathbb{R} \qquad \Big(h = (h_1, h_2, \dots, h_n) \in \mathbb{R}^n\Big)$$

függvényt az A mátrix által meghatározott **kvadratikus alaknak** *nevezzük.*

A $Q\colon \mathbb{R}^n \to \mathbb{R}$ kvadratikus alakok többváltozós polinomfüggvények, ezért minden pontban folytonosak és differenciálhatóak. Másrészt minden $\lambda \in \mathbb{R}$ esetén

$$Q(\lambda h) = \sum_{i,j=1}^{n} a_{ij} \cdot (\lambda h_i) \cdot (\lambda h_j) = \lambda^2 \sum_{i,j=1}^{n} a_{ij} h_i h_j = \lambda^2 Q(h) \qquad (h \in \mathbb{R}^n).$$

Mivel a $Q\colon \mathbb{R}^n \to \mathbb{R}$ kvadratikus alakok folytonosak a $\{h \in \mathbb{R}^n\colon \|h\| = 1\}$ korlátos és zárt halmazon, így a Weierstrass-tétel szerint

$$\exists m_Q := \min\{Q(h)\colon h \in \mathbb{R}^n,\ \|h\| = 1\} \in \mathbb{R} \quad \text{és} \quad \exists M_Q := \max\{Q(h)\colon h \in \mathbb{R}^n,\ \|h\| = 1\} \in \mathbb{R}.$$

Ha $h \neq 0$, akkor

$$Q(h) = Q\left(\|h\| \cdot \frac{h}{\|h\|}\right) = \|h\|^2 \cdot Q\left(\frac{h}{\|h\|}\right),$$

és mivel $\frac{h}{\|h\|}$ egységvektor, így

$$(*) \qquad m_Q\|h\|^2 \leq Q(h) \leq M_Q\|h\|^2 \qquad (h \in \mathbb{R}^n).$$

7. Definíció. *Azt mondjuk, hogy az $A \in \mathbb{R}^{n\times n}$ szimmetrikus mátrix, illetve a hozzá tartozó $Q(h) = \langle A \cdot h, h\rangle$ $(h \in \mathbb{R}^n)$ kvadratikus alak*

- **pozitív definit**, *ha $\forall\, h \in \mathbb{R}^n \setminus \{0\}$ esetén $Q(h) > 0$,*
- **negatív definit**, *ha $\forall\, h \in \mathbb{R}^n \setminus \{0\}$ esetén $Q(h) < 0$,*
- **indefinit**, *ha Q pozitív és negatív értéket is felvesz.*

A $(*)$ egyenlőtlenségből és az m_Q, M_Q számok értelmezéséből következik, hogy Q pontosan pozitív definit, ha $m_Q > 0$; negatív definit, ha $M_Q < 0$; és indefinit, ha $m_Q < 0 < M_Q$.

Most térjünk vissza a Taylor-formulához. Ha $f'(a) = 0$ és $Q(h) := \langle f''(a)h, h\rangle$ akkor

$$(\#) \qquad f(a+h) - f(a) = \tfrac{1}{2}Q(h) + \varepsilon(h) \cdot \|h\|^2 \qquad \Big(h \in \mathbb{R}^n,\ a + h \in \mathcal{D}_f\Big).$$

Ha Q pozitív definit, akkor $(*)$ miatt

$$f(a+h) - f(a) \geq \frac{m_Q}{2} \cdot \|h\|^2 + \varepsilon(h) \cdot \|h\|^2 = \left(\frac{m_Q}{2} + \varepsilon(h)\right) \cdot \|h\|^2.$$

Mivel $\lim\limits_{0} \varepsilon = 0$, így van olyan $K_\delta(0)$ környezet, ahol $\frac{m_Q}{2} + \varepsilon(h) > 0$ teljesül. Ezért, ha $h \in K_\delta(0)$, akkor $f(a+h) - f(a) \geq 0$. Az $x = a + h$ helyettesítéssel azt kapjuk, hogy

$$\forall x \in K_\delta(a)\colon f(x) \geq f(a),$$

ami azt jelenti, hogy az f függvénynek lokális minimuma van az a pontban.

Hasonlóan igazolható, hogy ha Q negatív definit, akkor az f függvénynek lokális maximuma van az a pontban. Az eddigi eredményeket a következő állításban tudjuk összefoglalni.

4. Tétel (Másodrendű elégséges feltétel a lokális szélsőértékre). *Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$, $a \in \operatorname{int} \mathcal{D}_f$ és $f \in C^2\{a\}$. Tegyük fel, hogy*

- *$f'(a) = 0$,*
- *az $f''(a)$ Hesse-féle mátrix pozitív (negatív) definit.*

Ekkor az f függvénynek az a pontban lokális minimuma (maximuma) van.

Másodrendű szükséges feltétel a lokális szélsőértékre

8. Definíció. *Azt mondjuk, hogy az $A \in \mathbb{R}^{n\times n}$ szimmetrikus mátrix, illetve a hozzá tartozó $Q(h) = \langle A \cdot h, h\rangle$ $(h \in \mathbb{R}^n)$ kvadratikus alak*

- **pozitív szemidefinit**, *ha $\forall\, h \in \mathbb{R}^n$ esetén $Q(h) \geq 0$,*
- **negatív szemidefinit**, *ha $\forall\, h \in \mathbb{R}^n$ esetén $Q(h) \leq 0$.*

A $(*)$ egyenlőtlenségből és az m_Q, M_Q számok értelmezéséből következik, hogy Q pontosan pozitív szemidefinit, ha $m_Q \geq 0$, és negatív szemidefinit, ha $M_Q \leq 0$. Ha Q nem (pozitív vagy negatív) szemidefinit, akkor indefinit.

5. Tétel (Másodrendű szükséges feltétel a lokális szélsőértékre). *Legyen $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$, $a \in \operatorname{int} \mathcal{D}_f$ és $f \in C^2\{a\}$. Ha az f függvénynek az a pontban lokális minimuma (maximuma) van, akkor*

- *$f'(a) = 0$,*
- *az $f''(a)$ Hesse-féle mátrix pozitív (negatív) szemidefinit.*

Bizonyítás. Az elsőrendű szükséges feltétel miatt $f'(a) = 0$. A Hesse-féle mátrixszal kapcsolatos állítás igazolásához tegyük fel indirekt módon, hogy

$$Q(h) := \langle f''(a)h, h\rangle$$

nem pozitív szemidefinit, azaz $m_Q < 0$. Ekkor $\exists u \in \mathbb{R}^n$ úgy, hogy $\|u\| = 1$ és $Q(u) < 0$. Legyen $h = tu$. Ekkor $(\#)$ miatt

$$f(a + tu) - f(a) = \tfrac{1}{2}Q(tu) + \varepsilon(tu) \cdot \|tu\|^2 = \frac{t^2}{2}Q(u) + \varepsilon(tu) \cdot t^2 = \left(\frac{1}{2}Q(u) + \varepsilon(tu)\right) \cdot t^2.$$

Mivel $Q(u) < 0$ és $\lim\limits_{0} \varepsilon = 0$, így van olyan $\delta > 0$, hogy $\frac{1}{2}Q(u) + \varepsilon(tu) < 0$ minden $|t| < \delta$ esetén, és így $f(a + tu) - f(a) < 0$. Azonban minden $K(a)$ környezet tartalmaz egy olyan $x = a + tu$ pontot, amire $|t| < \delta$ teljesül. Ez azt jelenti, hogy a nem lokális minimumhelye az f függvénynek.

Hasonlóan igazolható, hogy a Hesse-féle mátrix negatív szemidefinit, ha az f függvénynek az a pontban lokális maximuma van.

A tétel fontos következménye: *Ha $f \in \mathbb{R}^n \to \mathbb{R}$ $(n \in \mathbb{N}^+)$, $a \in \operatorname{int} \mathcal{D}_f$ és $f \in C^2\{a\}$, továbbá $f'(a) = 0$ és az $f''(a)$* **Hesse-féle mátrix indefinit**, *akkor az f függvénynek az a pontban* **nincs lokális szélsőértéke**.

Egy mátrix, illetve kvadratikus alak definitségének az eldöntése nem egyszerű feladat. A következő állításban a gyakorlatban jól használható eredményt fogalmazunk meg.

6. Tétel (Sylvester-kritérium). *Legyen $n \in \mathbb{N}^+$, $A \in \mathbb{R}^{n\times n}$ egy szimmetrikus mátrix és $Q(h) = \langle A \cdot h, h\rangle \quad (h \in \mathbb{R}^n)$ az A által meghatározott kvadratikus alak. Jelölje*

$$D_k := \det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \dots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{pmatrix} \qquad (k = 1, 2, \dots, n)$$

az A mátrix „bal felső sarokmátrixainak" a determinánsát. Ekkor az A mátrix, illetve a Q kvadratikus alak

- *pozitív definit* $\iff$ *ha $\forall k = 1, 2, \dots, n$ esetén $D_k > 0$,*
- *negatív definit* $\iff$ *ha $\forall k = 1, 2, \dots, n$ esetén $(-1)^k D_k > 0$.*

Bizonyítás. A tételt nem bizonyítjuk.

Megjegyzés. A Sylvester-kritériumból nem lehet megtudni mikor indefinit egy szimmetrikus mátrix. Azonban $n = 2$ esetén van egy egyszerű elégséges feltétel erre az esetre is, nevezetesen ha $\det A < 0$. Ez elemi úton, a másodfokú függvény tulajdonságai alapján könnyen igazolható.

Foglaljuk össze a kapott eredményeket kétváltozós függvények esetében:

Legyen $f \in \mathbb{R}^2 \to \mathbb{R}$, $a \in \operatorname{int} \mathcal{D}_f$ és $f \in C^2\{a\}$. Tegyük fel, hogy

$$\partial_1 f(a) = 0 \quad \text{és} \quad \partial_2 f(a) = 0.$$

Jelölje

$$D(a) := \det \begin{pmatrix} \partial_{11} f(a) & \partial_{12} f(a) \\ \partial_{21} f(a) & \partial_{22} f(a) \end{pmatrix}.$$

Ekkor

1. *ha $D(a) > 0$ és $\partial_{11} f(a) > 0$ [illetve $\partial_{11} f(a) < 0$], akkor az f függvénynek a-ban lokális minimuma [illetve maximuma] van.*
2. *ha $D(a) < 0$, akkor f-nek a-ban nincs lokális szélsőértéke (ezt nevezzük nyeregpontnak)*
3. *ha $D(a) = 0$, akkor így nem tudjuk megállapítani, hogy az a pont vajon lokális szélsőértékhely-e vagy sem.*

Példa. Legyen

$$f(x, y) = x^2 + xy + 2y^2 \qquad \big((x, y) \in \mathbb{R}^2\big).$$

Ekkor $f \in C^2(\mathbb{R}^2)$. Továbbá minden $(x, y) \in \mathbb{R}^2$ esetén

$$\partial_x f(x, y) = 2x + y = 0, \quad \partial_y f(x, y) = x + 4y = 0 \qquad \Longrightarrow \qquad x = y = 0.$$

Ez azt jelenti, hogy f-nek csak a $P(0, 0)$ pontban lehet lokális szélsőértéke. Másrészt

$$\partial_{xx} f(x, y) = 2, \quad \partial_{xy} f(x, y) = \partial_{yx} f(x, y) = 1, \quad \partial_{yy} f(x, y) = 4 \qquad \big((x, y) \in \mathbb{R}^2\big).$$

Mivel

$$D(0, 0) = \det \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix} = 7 > 0 \qquad \text{és} \qquad \partial_{xx} f(x, y) = 2 > 0,$$

így f-nek a $P(0, 0)$ pontban lokális minimuma van.

Példa. Legyen

$$f(x,y) = x^2 - y^2 \qquad \big((x,y) \in \mathbb{R}^2\big).$$

Ekkor $f \in C^2(\mathbb{R}^2)$. Továbbá minden $(x,y) \in \mathbb{R}^2$ esetén

$$\partial_x f(x,y) = 2x = 0, \quad \partial_y f(x,y) = -2y = 0 \qquad \Longrightarrow \qquad x = y = 0.$$

Ez azt jelenti, hogy f-nek csak a $P(0,0)$ pontban lehet lokális szélsőértéke. Másrészt

$$\partial_{xx} f(x,y) = 2, \quad \partial_{xy} f(x,y) = \partial_{yx} f(x,y) = 0, \quad \partial_{yy} f(x,y) = -2 \qquad \big((x,y) \in \mathbb{R}^2\big).$$

Mivel

$$D(0,0) = \det \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} = -4 < 0$$

így f-nek nincs a $P(0,0)$ pontban lokális szélsőértéke.

Abszolút szélsőértékek

Az a megfigyelés, hogy a lokális szélsőértékhelyeken a függvény deriváltja nulla (feltéve, hogy létezik), lehetővé tette olyan f egyváltozós függvény abszolút szélsőértékeinek meghatározását, amelyik *folytonos egy korlátos és zárt* $[a,b]$ *intervallumban, és differenciálható annak* (a,b) *belsejében.* Ekkor ui. f-nek van legnagyobb és legkisebb értéke a *Weierstrass-tétel* szerint. Ha f ezek valamelyikét egy c pontban veszi fel, akkor vagy $c = a$, vagy $c = b$, vagy pedig $c \in (a,b)$. Ez utóbbi esetben lokális szélsőértékről van szó, és így stacionárius pont, azaz $f'(c) = 0$.

Ha tehát megkeressük f összes $c \in (a,b)$ stacionárius pontját, akkor biztos, hogy az abszolút szélsőértékhelyek ezek közül, valamint az a és a b végpontok közül kerülnek ki. Például az abszolút maximumhelyet úgy határozzuk meg, hogy kiszámítjuk f értékeit ezekben a pontokban (nem feledkezve meg az a és b végpontokról sem), és kiválasztjuk azokat, amelyekben f értéke a legnagyobb.

A fenti gondolatmenetet könnyen általánosíthatjuk többváltozós függvényekre.

7. Tétel. *Legyen $H \subset \mathbb{R}^n$ korlátos és zárt halmaz. Tegyük fel, hogy az $f : H \to \mathbb{R}$ függvény folytonos és léteznek a parciális deriváltjai H belsejének minden pontjában. Ekkor f a legnagyobb (legkisebb) értékét vagy a H halmaz határán veszi fel, vagy pedig egy olyan $a \in \operatorname{int} \mathcal{D}_f$ belső pontban, ahol $\partial_i f(a) = 0$ teljesül minden $i = 1, 2, \ldots, n$ indexre.*

11. előadás

TÖBBVÁLTOZÓS FÜGGVÉNYEK DIFFERENCIÁLSZÁMÍTÁSA 3.

Az inverzfüggvény-tétel

A valós-valós függvények inverzére vonatkozó deriválási szabály azt mondja ki, hogy ha az I nyílt intervallumon értelmezett, és ott szigorúan monoton és folytonos f függvény egy $a \in I$ pontban differenciálható, és $f'(a) \neq 0$, akkor a létező f^{-1} függvény differenciálható a $b = f(a)$ pontban és

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

A fenti állítás lokálisan kiterjeszthető az a pont egyik környezetére, ha $f : I \to \mathbb{R}$ folytonosan deriválható az I nyílt intervallumon, és $f'(a) \neq 0$. Ti. a folytonosság miatt $\exists U := K(a)$ környezet, hogy $f'(u) > 0$ (vagy $f'(u) < 0$) minden $u \in U$ esetén, ezért f szigorúan monoton és folytonos U-n, továbbá a $V := f[U]$ képhalmaz olyan nyílt intervallum, amely tartalmazza az $f(a)$ pontot, és

1. f lokálisan invertálható, azaz $f_{|U} : U \to V$ függvény bijekció,
2. az f^{-1} inverz függvény folytonosan deriválható V-n és

$$\left(f^{-1}\right)'(y) = \frac{1}{f'\left(f^{-1}(y)\right)} \qquad (y \in V).$$

A többváltozós esetben hasonló állítás érvényes.

1. Tétel (Inverzfüggvény-tétel). *Legyen $\Omega \subseteq \mathbb{R}^n$ nyílt halmaz és $f : \Omega \to \mathbb{R}^n$. Tegyük fel, hogy,*

a) f folytonosan deriválható Ω-n,

b) az $a \in \Omega$ pontban $\det f'(a) \neq 0$.

Ekkor

1. *f lokálisan invertálható, azaz van olyan $a \in U$ és $f(a) \in V$ nyílt halmazok, hogy az $f_{|U} : U \to V$ függvény bijekció (következésképpen invertálható),*
2. *az f^{-1} inverz függvény folytonosan deriválható V-n és*

$$(*) \qquad \left(f^{-1}\right)'(y) = \left[f'(f^{-1}(y))\right]^{-1} \qquad (y \in V).$$

Bizonyítás. A tételt nem bizonyítjuk.

Megjegyzések.

1. Az inverz függvény létezése a többváltozós esetben *minőségileg bonyolultabb* az egyváltozós esetnél; ez tehát egy olyan pont, ahol az egyváltozós analógia létezik ugyan, az immár nem elegendő.

2. Az f függvény explicit alakjának az ismeretében f^{-1} helyettesítési értékeire általában nincs explicit képlet, viszont $(*)$ alapján a derivált helyettesítési értékei az f' helyettesítési értékeinek felhasználásával már kiszámíthatók.

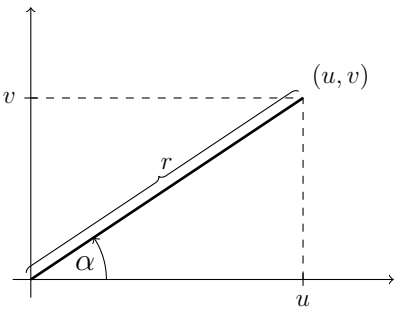
Példa. Legyen

$$f(x,y) := \Big(e^x \cos y, e^x \sin y\Big) \qquad \Big((x,y) \in \mathbb{R}^2\Big).$$

Nem nehéz igazolni, hogy $\mathcal{R}_f = \mathbb{R}^2 \setminus \{(0,0)\}$. Valóban $(0,0) \notin \mathcal{R}_f$, hiszen a sin és a cos függvény zérushelyei különbözőek. Másrészt, minden $(u,v) \neq (0,0)$ pont egyértelműen felírható

$$(u,v) = (r\cos\alpha, r\sin\alpha) \qquad \Big(r > 0,\ \alpha \in [0, 2\pi)\Big)$$

alakban, az ún. ***polárkoordinátákkal***, az ábrán szereplő jelölésekkel. Ekkor $x = \ln r$, azaz $e^x = r$, és $y = \alpha$ esetén $f(x,y) = (u,v)$.



A sin és a cos függvény periodicitása miatt f *globálisan* nem invertálható. Azonban f *lokálisan* invertálható az $\mathbb{R}^2$ minden pontjában. Ez utóbbi állítást az inverzfüggvény-tétellel tudjuk a legegyszerűbben igazolni. Világos, hogy $f \in C^1(\mathbb{R}^2)$. Az $f : \mathbb{R}^2 \to \mathbb{R}^2$ függvény Jacobi-mátrixa egy tetszőleges $(x,y) \in \mathbb{R}^2$ pontban

$$(\#) \qquad f'(x,y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}.$$

Mivel

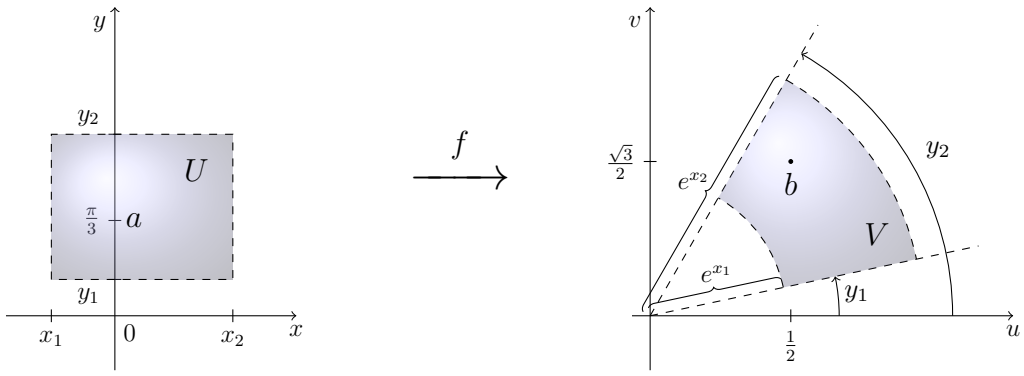
$$\det f'(x,y) = e^{2x} \cdot \cos^2 y + e^{2x} \cdot \sin^2 y = e^{2x} \neq 0 \qquad \Big((x,y) \in \mathbb{R}^2\Big),$$

ezért az inverzfüggvény-tétel feltételei teljesülnek minden $(x,y) \in \mathbb{R}^2$ pontban. Tehát f lokálisan invertálható.

Legyen $a := (0, \pi/3)$ és $b := f(a) = f(0, \pi/3) = (1/2, \sqrt{3}/2)$. Tetszőleges $x_1 < 0 < x_2$ és $0 < y_1 < \pi/3 < y_2 < \pi/2$ valós számok esetén

$$U := \{(x,y) \in \mathbb{R}^2 \colon x_1 < x < x_2,\ y_1 < y < y_2\} \ni a$$

nyílt halmaz, és $V := f[U] \ni b$ olyan nyílt halmaz, amelynek minden pontja pozitív koordinátákkal rendelkezik.



Ezért, ha $(x,y) \in U$, akkor

$$\begin{cases} e^x \cos y = u > 0 \\ e^x \sin y = v > 0 \end{cases} \implies \begin{cases} e^{2x} = u^2 + v^2 \\ \operatorname{tg} y = \dfrac{v}{u} \end{cases} \implies \begin{cases} x = \dfrac{1}{2}\ln(u^2+v^2) \\ y = \operatorname{arc\,tg} \dfrac{v}{u} \end{cases}.$$

Ekkor minden $(u,v) \in V$ esetén

$$f^{-1}(u,v) = \begin{pmatrix} \dfrac{1}{2}\ln\left(u^2+v^2\right) \\ \operatorname{arc\,tg} \dfrac{v}{u} \end{pmatrix} \implies \left(f^{-1}\right)'(u,v) = \begin{pmatrix} \dfrac{u}{u^2+v^2} & \dfrac{v}{u^2+v^2} \\ -\dfrac{v}{u^2+v^2} & \dfrac{u}{u^2+v^2} \end{pmatrix}.$$

Így

$$\left(f^{-1}\right)'(b) = \left(f^{-1}\right)'\left(\tfrac{1}{2}, \tfrac{\sqrt{3}}{2}\right) = \begin{pmatrix} \dfrac{1}{2} & \dfrac{\sqrt{3}}{2} \\ -\dfrac{\sqrt{3}}{2} & \dfrac{1}{2} \end{pmatrix}.$$

Az előző eredmény az inverzfüggvény-tételben szereplő

$$\left(f^{-1}\right)'(b) = \left[f'\left(f^{-1}(b)\right)\right]^{-1} = [f'(a)]^{-1}$$

összefüggéssel is megkaphatjuk. Valóban (#)-ból

$$f'(a) = f'\left(0, \tfrac{\pi}{3}\right) = \begin{pmatrix} \dfrac{1}{2} & -\dfrac{\sqrt{3}}{2} \\ \dfrac{\sqrt{3}}{2} & \dfrac{1}{2} \end{pmatrix} \implies \left(f^{-1}\right)'(b) = \begin{pmatrix} \dfrac{1}{2} & -\dfrac{\sqrt{3}}{2} \\ \dfrac{\sqrt{3}}{2} & \dfrac{1}{2} \end{pmatrix}^{-1} = \begin{pmatrix} \dfrac{1}{2} & \dfrac{\sqrt{3}}{2} \\ -\dfrac{\sqrt{3}}{2} & \dfrac{1}{2} \end{pmatrix}.$$

Megjegyzés. Érdemes megjegyezni a (2×2)-es mátrixok inverzére vonatkozó alábbi képletet: ha $ad - bc \neq 0$, akkor

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-cb} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Az inverzfüggvény-tételnek egyenletrendszerek megoldásával kapcsolatos értelmezés is adható. Legyen $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ és $y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$. Jelölje

$$f_i \in \mathbb{R}^n \to \mathbb{R} \qquad (i = 1, 2, \ldots, n)$$

az f függvény koordinátafüggvényeit: $f = (f_1, f_2, \ldots, f_n) \in \mathbb{R}^n \to \mathbb{R}^n$. Tekintsük az

$$f(x) = y$$

egyenletet. A komponensekre bontott alakba írva kapjuk az n egyenletből álló

$$\begin{aligned} f_1(x_1, x_2, \ldots, x_n) &= y_1, \\ f_2(x_1, x_2, \ldots, x_n) &= y_2, \\ &\vdots \\ f_n(x_1, x_2, \ldots, x_n) &= y_n \end{aligned}$$

egyenletrendszert, ahol az $y_1, y_2, \ldots, y_n$ számokat paramétereknek tekintjük, és $x_1, x_2, \ldots, x_n$ az ismeretlenek.

Legyen $a = (a_1, a_2, \ldots, a_n) \in \mathcal{D}_f$ és $b = (b_1, b_2, \ldots, b_n) := f(a)$. Tegyük fel, hogy f folytonosan deriválható az a pont egyik Ω környezetében, továbbá teljesül a $\det f'(a) \neq 0$ feltétel. Ekkor az inverzfüggvény-tétel szerint megadható olyan $b \in V$ paramétertartomány, hogy az egyenletrendszer egyértelműen megoldható az a pont egy U környezetében.

Implicit függvények (egyenletek megoldása)

Legyen $f \in \mathbb{R}^2 \to \mathbb{R}$ olyan függvény, amelyre

$$H := \Big\{(x,y) \in \mathbb{R}^2 \;\Big|\; f(x,y) = 0\Big\} \neq \emptyset$$

teljesül. Az lenne a kérdés, hogy ki tudjuk-e fejezni az y változót az x változó függvényeként, azaz van-e olyan $\varphi \in \mathbb{R} \to \mathbb{R}$ függvény, hogy $y = \varphi(x)$ ekvivalens legyen az $f(x,y) = 0$ egyenlettel.

Tudjuk, hogy az $x^2 + y^2 - 1 = 0$ egyenlet megoldásai az origó középpontú egység sugarú kör pontjai, és nincs olyan valós-valós függvény, amelynek grafikonja kört alkot. Ezért az eredeti kérdésnek egy „lokális” változatával foglalkozunk, nevezetesen olyan $\varphi : I \to \mathbb{R}$ nyílt intervallum értelmezett függvényt keresünk, amire

$$f\Big(x, \varphi(x)\Big) = 0 \qquad (x \in I)$$

teljesül. Ekkor azt mondjuk, hogy φ az $f(x,y)$ egyenletnek egy ***implicit megoldása***.

Nézzük újra az

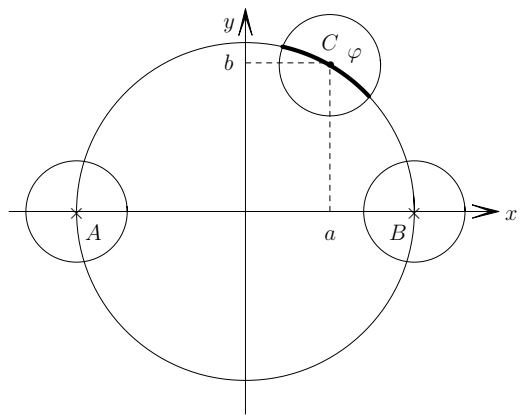
$$f(x,y) := x^2 + y^2 - 1 \quad \Big((x,y) \in \mathbb{R}^2\Big)$$

függvényből származó $x^2 + y^2 - 1 = 0$ egyenletet! Ha $C(a,b)$ olyan pont, hogy $f(a,b) = 0$ és $b > 0$, akkor $\exists I \ni a$ nyílt halmaz, hogy

$$\varphi(x) = \sqrt{1 - x^2} \qquad (x \in I)$$

implicit megoldása lesz az egyenletnek, illetve ha $b < 0$, akkor

$$\varphi(x) = -\sqrt{1 - x^2} \qquad (x \in I)$$



implicit megoldása lesz az egyenletnek. Azonban nincs olyan implicit megoldás, amely az $A(-1,0)$ vagy a $B(1,0)$ ponton menne át, azaz ha $b = 0$.

Vegyük észre, hogy $\partial_2 f(x,y) = 2y \implies \partial_2 f(A) = \partial_2 f(B) = 0$, de a többi C pontban (ahol $\exists\, \varphi$) igaz, hogy $\partial_2 f(C) \neq 0$.

2. Tétel (Egyváltozós implicitfüggvény-tétel). *Legyen $\Omega \subseteq \mathbb{R}^2$ nyílt halmaz és $f : \Omega \to \mathbb{R}$. Tegyük fel, hogy*

a) f folytonosan deriválható Ω-n,

b) az $(a,b) \in \Omega$ pontban $f(a,b) = 0$ és $\partial_2 f(a,b) \neq 0$.

Ekkor

1. *van olyan $U := K(a)$ környezet és $b \in V$ nyílt halmaz $\mathbb{R}$-ben, hogy minden $x \in U$ ponthoz létezik egyetlen $\varphi(x) \in V$, amelyre $f(x, \varphi(x)) = 0$ teljesül,*

2. *az így definiált $\varphi : U \to V$ függvény folytonosan deriválható U-n és*

$$(**) \qquad \varphi'(x) = -\frac{\partial_1 f(x, \varphi(x))}{\partial_2 f(x, \varphi(x))} \qquad (x \in U).$$

Bizonyítás. A tételt nem bizonyítjuk.

Megjegyzések.

1. Világos, hogy $\varphi(a) = b$. A φ függvényt az $f(x, \varphi(x)) = 0$ $(x \in U)$ egyenlőség „implicit" (= nem kifejtett, burkolt, rejtett) módon definiálja. Innen származik a tétel neve.

2. A tétel igazolható az inverzfüggvény-tétel segítségével. Ennek vázlata a következő: legyen

$$F : \Omega \to \mathbb{R}^2, \qquad F(x, y) := (x, f(x, y)).$$

Ekkor F folytonosan deriválható Ω-n, és

$$F'(x, y) = \begin{pmatrix} 1 & 0 \\ \partial_1 f(x, y) & \partial_2 f(x, y) \end{pmatrix} \qquad \Longrightarrow \qquad \det F'(a, b) = \partial_2 f(a, b) \neq 0.$$

Így F kielégíti az inverzfüggvény-tétel feltételeit az (a, b) pontban, ahol $F(a, b) = (a, 0)$. A létező folytonosan differenciálható $F^{-1} \colon V^* \to U^*$ inverz függvény F_1^{-1} és F_2^{-1} koordinátafüggvényei szintén folytonosan differenciálhatók, és minden $(x, y) \in V^*$ esetén

$$(x, y) = F\Big(F^{-1}(x, y)\Big) = F\Big(F_1^{-1}(x, y), F_2^{-1}(x, y)\Big) = \Big(F_1^{-1}(x, y), f\Big(F_1^{-1}(x, y), F_2^{-1}(x, y)\Big)\Big).$$

Ebből $F_1^{-1}(x, y) = x$, és így $f\Big(x, F_2^{-1}(x, y)\Big) = y$. Mivel $(a, 0) \in V^*$ nyílt halmaz, ezért $\exists K(a), K(0)$ környezetek, hogy $K(a) \times K(0) \in V^*$. Ekkor

$$f\Big(x, F_2^{-1}(x, 0)\Big) = 0 \quad \Big(x \in K(a)\Big) \qquad \Longrightarrow \qquad \varphi(x) := F_2^{-1}(x, 0) \quad \Big(x \in K(a)\Big)$$

a keresett implicit függvény, amiről nem nehéz igazolni, hogy ez az egyedüli ilyen függvény, ami értelmezett a $K(a)$ környezetben.

A $(**)$ összefüggésből abból következik, hogy

$$h(x) := f(x, \varphi(x)) = 0 \qquad \Big(x \in K(a)\Big).$$

Ezért az összetett függvényre vonatkozó deriválási szabály alapján

$$0 = h'(x) = \partial_1 f(x, \varphi(x)) \cdot 1 + \partial_2 f(x, \varphi(x)) \cdot \varphi'(x) \qquad \Big(x \in K(a)\Big).$$

3. Tétel (Implicitfüggvény-tétel az általános esetben). *Legyenek $\Omega_1 \subset \mathbb{R}^{n_1}$ és $\Omega_2 \subset \mathbb{R}^{n_2}$ nyílt halmazok $(n_1, n_2 \in \mathbb{N}^+)$, illetve $f : \Omega_1 \times \Omega_2 \to \mathbb{R}^{n_2}$. Tegyük fel, hogy,*

a) f folytonosan deriválható az $\Omega_1 \times \Omega_2$ halmazon,

b) az $(a, b) \in \Omega_1 \times \Omega_2$ pontban $f(a, b) = 0$ és $\det \partial_2 f(a, b) \neq 0$.

Ekkor

1. *létezik a-nak olyan $U := K(a) \subseteq \Omega_1$ környezet és $b \in V \subseteq \Omega_2$ nyílt halmaz, hogy minden $x \in U$ ponthoz létezik egyetlen $\varphi(x) \in V$, amelyre $f(x, \varphi(x)) = 0 \in \mathbb{R}^{n_2}$,*

2. *az így definiált $\varphi : U \to V$ függvény folytonosan deriválható U-n és*

$$\varphi'(x) = -\left[\partial_2 f(x, \varphi(x))\right]^{-1} \cdot \partial_1 f(x, \varphi(x)) \qquad (x \in U).$$

Megjegyzések.

1. A tételben $\partial_2 f(a,b)$ jelöli az f függvény *második változócsoport szerinti parciális deriváltját* az (a,b) pontban. Ez az alábbi módon definiált $n_2 \times n_2$-típusú mátrix:

$$\partial_2 f(a,b) := (\mathbb{R}^{n_2} \supset \Omega_2 \ni y \mapsto f(a,y) \in \mathbb{R}^{n_2})'_{y=b} \in \mathbb{R}^{n_2 \times n_2}.$$

   A $\partial_1 f(a,b)$ derivált definíciója hasonló.

2. A tételnek egyenletrendszerek *megoldhatóságával* kapcsolatos értelmezés is adható.

   Legyen $n_1, n_2 \in \mathbb{N}$, $x = (x_1, x_2, \ldots, x_{n_1}) \in \mathbb{R}^{n_1}$, $y = (y_1, y_2, \ldots, y_{n_2}) \in \mathbb{R}^{n_2}$ és $f = (f_1, f_2, \ldots, f_{n_2}) \in \mathbb{R}^{n_1} \times \mathbb{R}^{n_2} \to \mathbb{R}^{n_2}$.

   Tekintsük az $f(x,y) = 0$ egyenletrendszert, amelyet komponensekre bontott alakban így írhatunk fel:

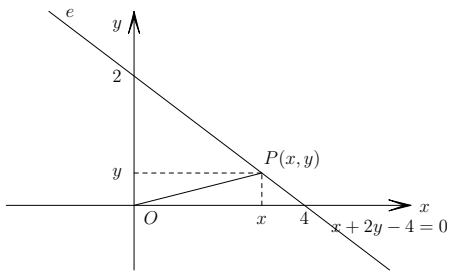
$$\begin{aligned} f_1(x_1, x_2, \ldots, x_{n_1}; y_1, y_2, \ldots, y_{n_2}) &= 0, \\ f_2(x_1, x_2, \ldots, x_{n_1}; y_1, y_2, \ldots, y_{n_2}) &= 0, \\ &\vdots \\ f_{n_2}(x_1, x_2, \ldots, x_{n_1}; y_1, y_2, \ldots, y_{n_2}) &= 0. \end{aligned}$$

   Itt az $y_1, y_2, \ldots, y_{n_2}$ számok az ismeretlenek és $x_1, x_2, \ldots, x_{n_1}$ a paraméterek. Feltesszük, hogy *ismerjük* ennek egy megoldását, azaz tudjuk, hogy az $a = (a_1, a_2, \ldots, a_{n_1})$ paraméter esetén $b = (b_1, b_2, \ldots, b_{n_2})$ egy megoldás, vagyis $f(a,b) = 0$. A fenti egyenletrendszerből szeretnénk kifejezni az $y_1, y_2, \ldots, y_{n_2}$ ismeretleneket az $x_1, x_2, \ldots, x_{n_1}$ paraméterek függvényében. A 2. Tétel szerint ez minden a-hoz közeli x esetén megtehető, ha f folytonosan deriválható és $\partial_2 f(a,b) \neq 0$; a megoldások egyértelműek és x-nek folytonosan deriválható függvényei.

$\mathbb{R}^2 \to \mathbb{R}$ típusú függvények feltételes szélsőértékei

Vannak olyan problémák, ahol egy $f \in \mathbb{R}^2 \to \mathbb{R}$ függvény szélsőértékét kell keresni, de csak bizonyos egyenletet kielégítő pontok jöhetnek számításba.

1. Példa: Keressük meg az $x + 2y - 4 = 0$ egyenletű egyenesnek azt a pontját, amely legközelebb van az origótól!



A probléma a következő módon modellezhető:

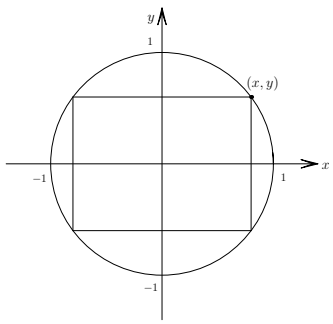
Keressük meg az

$$f(x,y) := x^2 + y^2 \quad \big((x,y) \in \mathbb{R}^2\big)$$

függvény minimumát a

$$H_g := \Big\{(x,y) \in \mathbb{R}^2 \;\Big|\; g(x,y) = 0\Big\} \quad \text{halmazon, ahol} \quad g(x,y) := x + 2y - 4 \quad \big((x,y) \in \mathbb{R}^2\big).$$

2. Példa: Határozzuk meg az egységsugarú körbe írt téglalapok között a maximális területű téglalapot!



A probléma a következő módon modellezhető:

Keressük meg az

$$f(x,y) := 4xy \quad \big((x,y) \in \mathbb{R}^2\big)$$

függvény maximumát a

$$H_g := \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; g(x,y) = 0\right\}$$

halmazon, ahol

$$g(x,y) := x^2 + y^2 - 1 \quad \left((x,y) \in \mathbb{R}^2\right).$$

Általános feladat: Adott

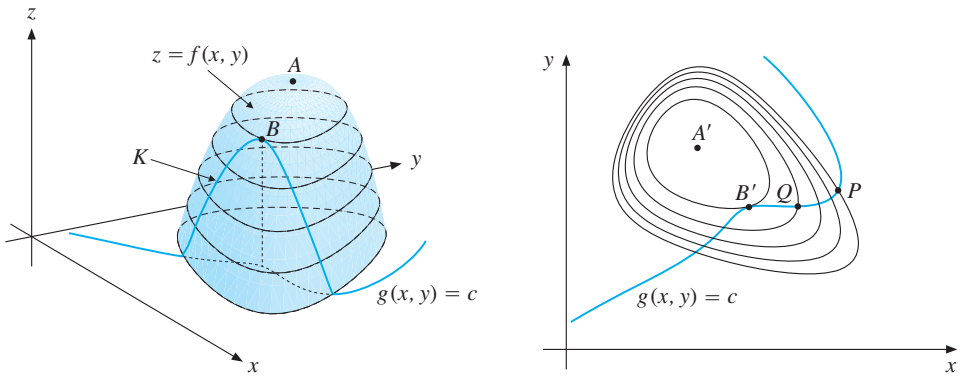
- $U \subseteq \mathbb{R}^2$ nyílt halmaz,
- $f : U \to \mathbb{R}$ (célfüggvény) és
- $g : U \to \mathbb{R}$ (feltételfüggvény).

Keressük az f függvény szélsőértékeit a

$$H_g := \left\{(x,y) \in U \;\middle|\; g(x,y) = 0\right\}$$

halmazon, azaz határozzuk meg az $f_{|H_g}$ függvény szélsőértékeit!

A problémát az alábbi ábrákon szemléltetjük:



1. Definíció. *Legyen $U \subseteq \mathbb{R}^2$ nyílt halmaz. Tegyük fel, hogy $f, g : U \to \mathbb{R}$ adott függvények és*

$$a \in H_g := \left\{z \in U \;\middle|\; g(z) = 0\right\} \neq \emptyset.$$

Azt mondjuk, hogy az f függvénynek a $g = 0$ feltétel mellett az a pontban

- **feltételes abszolút maximuma van**, *ha*

$$\forall x \in H_g \colon f(x) \leq f(a),$$

- **feltételes lokális maximuma van**, *ha*

$$\exists K(a) \subseteq U, \forall x \in K(a) \cap H_g \colon f(x) \leq f(a).$$

A ***minimummal*** kapcsolatban hasonló fogalmakat kapunk, ha a fentiekben a $\leq$ egyenlőtlenség helyett $\geq$-t írunk. A korábbiakkal összhangban használjuk $f(a)$-ra a ***feltételes abszolút (lokális) maximum (minimum)***, illetve ***szélsőérték***, továbbá a-ra a ***feltételes abszolút (lokális) maximumhely (minimumhely)***, illetve ***szélsőértékhely*** elnevezést is.

Megjegyzés. Az $f_{|H_g} \in \mathbb{R}^2 \to \mathbb{R}$ függvény lokális szélsőértékeire *nem alkalmazhatók* az előző előadáson megfogalmazott tételek, hiszen a $H_g \subseteq \mathbb{R}^2$ halmaznak nincsenek belső pontjai. Ezért úgy értelmeztük a feltételes lokális szélsőértékhelyeket, hogy minden feltételes abszolút szélsőértékhely egyben lokális szélsőértékhely is legyen az f függvénynek.

4. Tétel (Szükséges feltétel a feltételes lokális szélsőértékre). *Tegyük fel, hogy*

a) $U \subseteq \mathbb{R}^2$ *nyílt halmaz és az* $f, g : U \to \mathbb{R}$ *függvényeknek léteznek a parciális deriváltjaik és ezek folytonosak az* U *halmazon,*

b) az $(x_0, y_0) \in U$ *pontban az* f *függvénynek a* $g = 0$ *feltételre vonatkozóan feltételes lokális szélsőértéke van,*

c) $g'(x_0, y_0) = \big(\partial_1 g(x_0, y_0), \partial_2 g(x_0, y_0)\big) \neq \big(0, 0\big)$.

Ekkor van olyan $\lambda \in \mathbb{R}$ *valós szám (ezt* ***Lagrange-szorzónak*** *szokás nevezni), hogy az*

$$\mathcal{L}(x, y) := f(x, y) + \lambda g(x, y) \qquad \big((x, y) \in U\big)$$

Lagrange-függvénynek (x_0, y_0) *stacionárius pontja, azaz*

$$\mathcal{L}'(x_0, y_0) = \big(\partial_x \mathcal{L}(x_0, y_0), \partial_y \mathcal{L}(x_0, y_0)\big) = (0, 0).$$

Bizonyítás. A c) feltétel alapján feltételezhető, hogy $\partial_2 g(x_0, y_0) \neq 0$, ellenkező esetben $\partial_1 g(x_0, y_0) \neq 0$, és így a bizonyítás további részében x és y szerepe felcserélhető. Ha $\partial_2 g(x_0, y_0) \neq 0$, akkor a g függvény az (x_0, y_0) pontra vonatkozóan teljesíti az implicitfüggvény-tétel feltételeit, hiszen $g(x_0, y_0) = 0$ és az a) feltétel alapján g folytonosan deriválható U-n. Ezért van olyan $\varphi\colon U^* := K(x_0) \to V$ folytonosan deriválható függvény, hogy $g(x, \varphi(x)) = 0$ teljesül, és

$$(*) \qquad \varphi'(x) = -\frac{\partial_1 g\big(x, \varphi(x)\big)}{\partial_2 g\big(x, \varphi(x)\big)} \qquad (x \in U^*).$$

A b) feltétel alapján a

$$h(x) := f\big(x, \varphi(x)\big) \qquad (x \in U^*)$$

függvénynek lokális szélsőértéke van az x_0 pontban. Ezért az összetett függvényre vonatkozó deriválási szabály szerint

$$0 = h'(x_0) = \partial_1 f\big(x_0, \varphi(x_0)\big) + \partial_2 f\big(x_0, \varphi(x_0)\big)\varphi'(x_0) = \partial_1 f(x_0, y_0) + \partial_2 f(x_0, y_0)\varphi'(x_0),$$

hiszen $\varphi(x_0) = y_0$. De $(*)$ miatt $\varphi'(x_0) = -\partial_1 g(x_0, y_0)/\partial_2 g(x_0, y_0)$. Ha ezt behelyettesítjük a fenti egyenletbe, akkor átrendezés után azt kapjuk, hogy

$$\partial_1 f(x_0, y_0)\partial_2 g(x_0, y_0) - \partial_2 f(x_0, y_0)\partial_1 g(x_0, y_0) = 0.$$

Mivel $\partial_2 g(x_0, y_0) \neq 0$, így $\exists \lambda \in \mathbb{R}\colon \partial_y \mathcal{L}(x_0, y_0) = \partial_2 f(x_0, y_0) + \lambda \partial_2 g(x_0, y_0) = 0$. Ezért

$$\partial_2 f(x_0, y_0) = -\lambda \partial_2 g(x_0, y_0) \quad \Longrightarrow \quad \partial_1 f(x_0, y_0)\partial_2 g(x_0, y_0) - \big(-\lambda \partial_2 g(x_0, y_0)\big)\partial_1 g(x_0, y_0) = 0.$$

$\partial_2 g(x_0, y_0)$-vel való egyszerűsítés után

$$\partial_x \mathcal{L}(x_0, y_0) = \partial_1 f(x_0, y_0) + \lambda \partial_1 g(x_0, y_0) = 0.$$

Megjegyzés. A feltételes szélsőértékek vizsgálatára alkalmazható módszer kitalálója *Joseph Louis Lagrange* (1736–1813) francia matematikus. Ezért a szóban forgó módszert ***Lagrange-szorzók*** (vagy ***Lagrange-féle multiplikátorok***) ***módszerének*** nevezzük.

A módszer alkalmazása:

1. Képezzük az
$$\mathcal{L}(x,y) := f(x,y) + \lambda g(x,y) \qquad \big((x,y) \in U\big)$$
Lagrange függvényt.

2. Az x, y, λ ismeretlenekre megoldjuk a következő egyenletrendszert:
$$\begin{aligned} \partial_x \mathcal{L}(x,y) &= \partial_x f(x,y) + \lambda \partial_x g(x,y) = 0, \\ \partial_y \mathcal{L}(x,y) &= \partial_y f(x,y) + \lambda \partial_y g(x,y) = 0, \\ g(x,y) &= 0. \end{aligned}$$
Az így kapott (x_0, y_0) pont(ok)ban *lehet(nek)* a feltételes lokális szélsőértékhelyek.

Azonban az $\mathcal{L}'(x_0, y_0) = (0,0)$ csak *szükséges*, de *nem elégséges* feltétel a feltételes lokális szélsőértékre.

5. Tétel (A feltételes lokális szélsőértékre vonatkozó másodrendű elégséges feltétel). *Tegyük fel, hogy*

a) $U \subseteq \mathbb{R}^2$ nyílt halmaz és az $f, g : U \to \mathbb{R}$ függvényeknek léteznek a másodrendű parciális deriváltjaik és ezek folytonosak az U halmazon,

b) az $(x_0, y_0) \in U$ pontban a $\lambda_0 \in \mathbb{R}$ számmal teljesül a szükséges feltétel.

Tekintsük ezzel a λ_0 számmal az
$$\mathcal{L}(x,y) := f(x,y) + \lambda_0 g(x,y) \qquad \big((x,y) \in U\big)$$
Lagrange-függvényt. Legyen
$$D(x_0, y_0; \lambda_0) := \det \begin{pmatrix} 0 & \partial_1 g(x_0,y_0) & \partial_2 g(x_0,y_0) \\ \partial_1 g(x_0,y_0) & \partial_{11}\mathcal{L}(x_0,y_0) & \partial_{12}\mathcal{L}(x_0,y_0) \\ \partial_2 g(x_0,y_0) & \partial_{21}\mathcal{L}(x_0,y_0) & \partial_{22}\mathcal{L}(x_0,y_0) \end{pmatrix}.$$
Ekkor,

- $D(x_0, y_0; \lambda_0) > 0 \quad \Longrightarrow \quad (x_0, y_0)$ *feltételes lokális* **maximumhely**,
- $D(x_0, y_0; \lambda_0) < 0 \quad \Longrightarrow \quad (x_0, y_0)$ *feltételes lokális* **minimumhely**.

Bizonyítás. A tételt nem bizonyítjuk.

Megjegyzések.

1. A fentiekben két változó és egy egyenlőségi feltétel mellett vizsgáltuk a feltételes szélsőértékproblémát. Az eredmények kiterjeszthetők arra az esetre is, amikor az f célfüggvény n-változós ($2 < n \in \mathbb{N}$), és ekkor az egyetlen $g = 0$ feltétel helyett akár több $g_1 = 0$, $g_2 = 0, \ldots, g_m = 0$ egyenlőségi feltételt is előírhatunk, ahol $1 \le m < n$. Ekkor a Lagrange-függvény
$$\mathcal{L}(x,y) := f(x,y) + \sum_{k=1}^{m} \lambda_k g_k(x,y) \qquad \big((x,y) \in U\big).$$

Ha $m > 1$, akkor több λ szorzó szerepel a Lagrange-függvényben, ami igazolja a „Lagrange-szorzók” elnevezésben szereplő többes számot.

2. A gyakorlat felvet számos olyan szélsőérték-problémát, amelyekben a változókra tett korlátozó feltételek *nem egyenlőségekkel*, hanem *egyenlőtlenségekkel* adottak. Az ilyen típusú problémákat *(lineáris) programozási* feladatoknak hívják. Vizsgálatukhoz nem csak az *analízis*, hanem a *lineáris algebra* eszköztárát is fel kell használni.

3. Ha szükséges feltétel bizonyításában szereplő $\varphi\colon U^* \to \mathbb{R}$ implicit függvény ki tudjuk fejezni, például az y változó kifejezhető az x változó függvényeként, akkor az f függvénynek a H_g halmaz pontjaiban felvett értékeit a

$$h(x) := f\Big(x, \varphi(x)\Big) \qquad (x \in U^*)$$

valós-valós függvénnyel lehet kifejezni. Ekkor a kétváltozós függvényekre vonatkozó feltételes szélsőérték-problémát a h egyváltozós függvény szélsőérték-problémájára lehet visszavezetni.

Példa: Tekintsük az

$$f(x,y) := xy, \quad g(x,y) := \frac{x^2}{8} + \frac{y^2}{2} - 1 \qquad \Big((x,y) \in \mathbb{R}^2\Big)$$

függvényeket, és határozzuk meg az f feltételes lokális szélsőértékeit a $g = 0$ feltétel mellett!

A szükséges feltételre vonatkozó tétel feltételei teljesülnek, mert $f, g \in C^1\Big(\mathbb{R}^2\Big)$ és

$$g'(x,y) = \Big(\partial_1 g(x,y), \partial_2 g(x,y)\Big) = \Big(\tfrac{x}{4}, y\Big) \neq (0,0)$$

minden $H_g := \Big\{(x,y) \in \mathbb{R}^2 \;\Big|\; g(x,y) = 0\Big\}$-beli pontban, hiszen ha $\Big(\frac{x}{4}, y\Big) = (0,0)$, akkor $x = 0$ és $y = 0$, de $g(x,y) = -1 \neq 0$.

A feladat Lagrange-függvénye:

$$\mathcal{L}(x,y) := f(x,y) + \lambda g(x,y) = xy + \lambda\left(\frac{x^2}{8} + \frac{y^2}{2} - 1\right) \qquad \Big((x,y) \in \mathbb{R}^2\Big).$$

A feltételes lokális szélsőértékre vonatkozó *szükséges* feltétel az x, y, λ ismeretlenekre az alábbi egyenletrendszert adja:

$$\begin{aligned} \partial_1\mathcal{L}(x,y) &= y + \lambda\frac{x}{4} &&= 0,\\ \partial_2\mathcal{L}(x,y) &= x + \lambda y &&= 0,\\ g(x,y) &= \frac{x^2}{8} + \frac{y^2}{2} - 1 &&= 0. \end{aligned}$$

A második egyenletből $x = -\lambda y$ adódik. Ezt beírjuk az első egyenletbe:

$$0 = y + \lambda\frac{-\lambda y}{4} \quad\Longrightarrow\quad 0 = y\left(1 - \frac{\lambda^2}{4}\right) \quad\Longrightarrow\quad \lambda^2 = 4 \quad\Longrightarrow\quad \lambda = \pm 2,$$

hiszen $y \neq 0$ (ha $y = 0$, akkor $x = 0$ és így $g(x,y) = g(0,0) = -1 \neq 0$).

i) Ha $\lambda = 2$, akkor $x = -2y$. Ekkor a harmadik egyenletből:

$$\frac{(-2y)^2}{8} + \frac{y^2}{2} - 1 = 0 \quad\Longrightarrow\quad y^2 = 1 \quad\Longrightarrow\quad \begin{aligned} &y = 1,\ x = -2, \quad P_1(-2,1),\\ &y = -1,\ x = 2, \quad P_2(2,-1). \end{aligned}$$

ii) Ha $\lambda = -2$, akkor $x = 2y$. Ekkor a harmadik egyenletből:

$$\frac{(2y)^2}{8} + \frac{y^2}{2} - 1 = 0 \quad \Longrightarrow \quad y^2 = 1 \quad \Longrightarrow \quad \begin{array}{l} y = 1,\ x = 2, \quad P_3(2,1), \\ y = -1,\ x = -2, \quad P_4(-2,-1). \end{array}$$

Az elégséges feltétel. Minden $(x, y) \in \mathbb{R}^2$ pontban

$$\partial_1 g(x,y) = \frac{x}{4}, \qquad \partial_2 g(x,y) = y;$$

$$\partial_{11}\mathcal{L}(x,y) = \frac{\lambda}{4}, \quad \partial_{12}\mathcal{L}(x,y) = 1 = \partial_{21}\mathcal{L}(x,y), \quad \partial_{22}\mathcal{L}(x,y) = \lambda.$$

Mivel $x = -\lambda y$, így

$$D(x,y;\lambda) = \det\begin{pmatrix} 0 & x/4 & y \\ x/4 & \lambda/4 & 1 \\ y & 1 & \lambda \end{pmatrix} = \det\begin{pmatrix} 0 & -\lambda y/4 & y \\ -\lambda y/4 & \lambda/4 & 1 \\ y & 1 & \lambda \end{pmatrix} = \frac{1}{16}\cdot\det\begin{pmatrix} 0 & -\lambda y & 4y \\ -\lambda y & \lambda & 4 \\ y & 1 & \lambda \end{pmatrix} =$$

$$= \frac{y^2}{16}\cdot\det\begin{pmatrix} 0 & -\lambda & 4 \\ -\lambda & \lambda & 4 \\ 1 & 1 & \lambda \end{pmatrix} = \frac{y^2}{16}\cdot\det\begin{pmatrix} 0 & -\lambda & 4 \\ 0 & 2\lambda & 4+\lambda^2 \\ 1 & 1 & \lambda \end{pmatrix} = \frac{y^2}{16}\cdot\det\begin{pmatrix} -\lambda & 4 \\ 2\lambda & 4+\lambda^2 \end{pmatrix} =$$

$$= \frac{\lambda y^2}{16}\cdot\det\begin{pmatrix} -1 & 4 \\ 2 & 4+\lambda^2 \end{pmatrix} = \frac{\lambda y^2}{16}\cdot\Big(-(4+\lambda^2) - 2\cdot 4\Big) = -\lambda\cdot\frac{y^2(\lambda^2+12)}{16}.$$

$\underline{P_1(-2,1),\ \lambda = 2}$: $D(-2,1;-2) < 0$, ezért a $P_1(-2,1)$ pont *feltételes lokális minimumhely.*

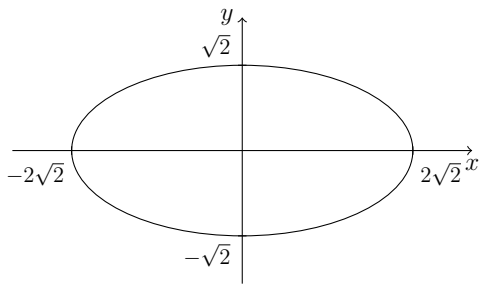
$\underline{P_2(2,-1),\ \lambda = 2}$: $D(2,-1;-2) < 0$, ezért a $P_2(2,-1)$ pont *feltételes lokális minimumhely.*

$\underline{P_3(2,1),\ \lambda = -2}$: $D(2,1;-2) > 0$, ezért a $P_3(2,1)$ pont *feltételes lokális maximumhely.*

$\underline{P_4(-2,-1),\ \lambda = -2}$: $D(-2,-1;-2) > 0$, ezért a $P_4(-2,-1)$ pont *feltételes lokális maximumhely.*

Megjegyzés. Az

$$\frac{x^2}{8} + \frac{y^2}{2} = 1$$



egyenletű görbe pontjai az ábrán látható ellipszist alkotják. Ezekből a pontokból álló H_g halmaz korlátos és zárt az $\mathbb{R}^2$ térben. Mivel f folytonos, így a Weierstrass-tétel szerint felveszi a maximumát és a minimumát ezen a halmazon. A feltételes abszolút maximum- és minimumhely csak a fenti négy pontból kerülhet ki, továbbá

$$f(P_1) = f(P_2) = -2 \qquad \text{és} \qquad f(P_3) = f(P_4) = 2.$$

Ezért

$\underline{P_1(-2,1), \text{ és } P_2(2,-1)}$: *feltételes abszolút minimumhely.*

$\underline{P_3(2,1), \text{ és } P_4(-2,-1)}$: *feltételes abszolút maximumhely.*

12. előadás

TÖBBSZÖRÖS INTEGRÁLOK 1.

Az egyváltozós analízisben hangsúlyoztuk, hogy a matematikai alkalmazások igen fontos fejezete az *integrálszámítás*. Bevezettük a határozott integrál (vagy Riemann-integrál) fogalmát, megismertük a legfontosabb tulajdonságait, és bemutattuk számos gyakorlati alkalmazását. A továbbiakban a Riemann-integrál többváltozós függvényekre való kiterjesztéséről lesz szó.

Fontos megjegyezni, hogy a valós-valós függvények körében megismert integrálfogalmat többféle módon lehet általánosítani. Sőt: különböző (pl. geometriai, fizikai és egyéb) problémák vizsgálata szükségessé is tette több integrálfogalom bevezetését. Ilyen pl. egy kétváltozós függvény grafikonja alatti térrész térfogatának a problémája, ami az $\mathbb{R}^n \to \mathbb{R}$ függvények **többszörös integráljának** a fogalmához vezet. A továbbiakban csak a többszörös integrálokról lesz szó.

A többszörös integrálok értelmezése n-dimenziós intervallumokon

Szeretnénk ara emlékeztetni, hogy a Riemann-integrál bevezetését az motiválta, hogy ki akartunk számítani függvények görbe alatti területét. Abból az Arkhimédész-óta ismert, egyébként elég természetes ötletből indultunk ki, hogy a szóban forgó (görbe vonallal határolt) síkidom területét téglalapok területeinek az összegével közelítsük. Ebből kiindulva jutottunk el a *Riemann-integrálhatóság* fogalmához.

A többszörös integrál bevezetését hasonló geometriai, illetve fizikai problémák motiválják. Példaként tekintsünk egy kétváltozós, valós értékű pozitív függvényt, amelyik az egyszerűség végett például egy, a koordináta-tengelyekkel párhuzamos oldalú téglalapon van értelmezve. A függvény grafikonja alatti térrész *térfogatát* téglatestek térfogatainak az összegével lehet közelíteni.

Látni fogjuk, hogy az egyváltozós Riemann-integrál fogalmának bevezetésénél követett út szó szerint átvihető $\mathbb{R}^n \to \mathbb{R}$ függvényekre, ezért a többszörös integrál értelmezése az egyváltozós Riemann-integrál definíciójának *közvetlen* általánosításaként adódik.

Induljunk ki a legegyszerűbb $\mathbb{R}^n$-beli halmazokból, az ún. n-dimenziós intervallumokból: Egy ***n-dimenziós intervallum*** (vagy más szóval ***n-dimenziós tégla***) az

$$I := [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n \qquad (n \in \mathbb{N}^+) \tag{1}$$

Descartes-szorzattal értelmezett halmaz, ahol $a_k, b_k \in \mathbb{R}$, $a_k < b_k$ $(k = 1, 2, \ldots, n)$. Az

$$|I| := \mu(I) := \prod_{k=1}^{n} (b_k - a_k)$$

számot az I intervallum ***mértékének*** nevezzük.

- Ha $n = 1$, akkor a „szokásos” (korlátos és zárt) $\mathbb{R}$-beli intervallumot kapunk, amelynek mértéke az $|I| = b_1 - a_1$ intervallum *hossza*.
- Ha $n = 2$, akkor a koordináta-síkon a koordináta-tengelyekkel párhuzamos oldalú téglalapot kapunk, amelynek mértéke az $|I| = (b_1 - a_1) \cdot (b_2 - a_2)$ téglalap *területe*.

- Ha $n = 3$, akkor a térbeli derékszögű koordináta-rendszerben a koordináta-síkokkal párhuzamos oldallapú téglatestet kapunk, amelynek mértéke az $|I| = (b_1 - a_1)\cdot(b_2 - a_2)\cdot(b_3 - a_3)$ téglatest *térfogata.*

Emlékeztetünk arra, hogy egy korlátos és zárt $[a, b] \subset \mathbb{R}$ intervallum felosztásán olyan $\tau \subset [a, b]$ *véges* halmazt értettünk, amelyre $a, b \in \tau$, azaz

$$\tau := \Big\{a = x_0 < x_1 < x_2 < \cdots < x_m = b\Big\},$$

ahol m egy adott természetes szám. Az intervallum felosztásainak a halmazát az $\mathcal{F}[a, b]$ szimbólummal jelöltük. A többdimenziós intervallum felosztásának az értelmezéséhez vegyük észre, hogy a fenti osztópontokkal megadott felosztást az $I_j := [x_j, x_{j+1}]$ intervallumok ($j = 0, 1, 2, \ldots, m-1$) halmazaként is értelmezhetjük:

$$\tau = \Big\{I_j = [x_j, x_{j+1}] \;\Big|\; j = 0, 1, 2, \ldots, m-1\Big\}.$$

Legyen ezek után egy $k = 1, 2, \ldots, n$ index esetén az $[a_k, b_k]$ intervallum egy felosztása

$$\tau_k = \Big\{x_{k,0} = a_k < x_{k,1} < x_{k,2} < \cdots < x_{k,m_k} = b_k\Big\} =$$
$$= \Big\{I_{k,j} = [x_{k,j}, x_{k,j+1}] \;\Big|\; j = 0, 1, \ldots, m_k - 1\Big\}.$$

(A felosztás tehát $m_k + 1$ osztópontot, illetve m_k intervallumot tartalmaz.) Ekkor az (1) n-dimenziós I intervallum egy ***felosztásán*** a

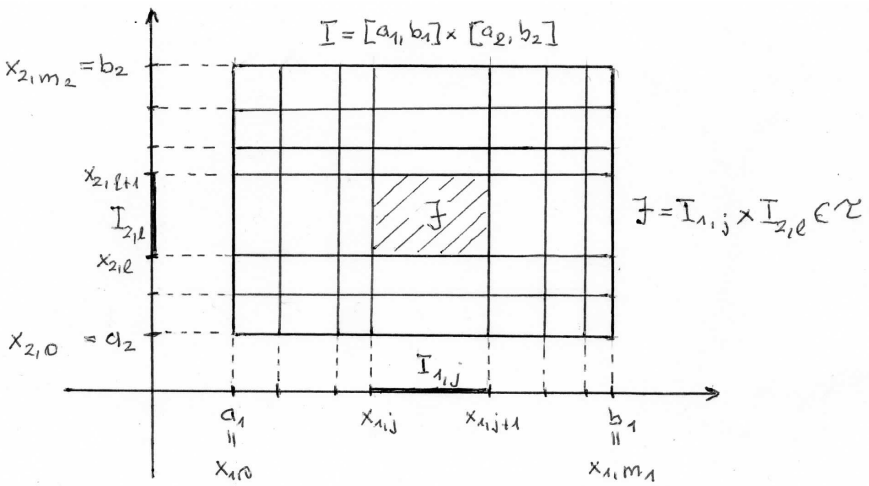
$$\tau := \tau_1 \times \tau_2 \times \cdots \times \tau_n \subset I$$

halmazt értjük, a felosztások halmazát a $\mathcal{F}(I)$ szimbólummal jelöljük. A τ halmaz elemei tehát az

$$I_{1,j_1} \times I_{2,j_2} \times \cdots \times I_{n,j_n}$$

n-dimenziós intervallumok, ahol $0 \le j_l \le m_l - 1$ ($l = 1, 2, \ldots, n$).

A fentieket az $n = 2$ esetben az alábbi ábra szemlélteti.



Egyszerűen igazolható, hogy

$$I = \bigcup_{J \in \tau} J, \qquad \mu(I) = \sum_{J \in \tau} \mu(J).$$

Az egyváltozós esethez hasonlóan értelmezzük az alsó-, illetve a felső közelítő összeg fogalmát. Legyen τ az n-dimenziós I intervallum egy felosztása és $f : I \to \mathbb{R}$ korlátos függvény. Ekkor

$$s(f,\tau) := \sum_{J\in\tau} \inf_{x\in J} f(x)\cdot\mu(J)$$

az f függvény τ felosztáshoz tartozó ***alsó közelítő összege***,

$$S(f,\tau) := \sum_{J\in\tau} \sup_{x\in J} f(x)\cdot\mu(J)$$

az f függvény τ felosztáshoz tartozó ***felső közelítő összege***. Mivel tetszőleges $\tau \in \mathcal{F}(I)$ felosztás esetén

$$\inf_{x\in I} f(x)\cdot\mu(I) \le s(f,\tau) \le S(f,\tau) \le \sup_{x\in I} f(x)\cdot\mu(I),$$

ezért minden korlátos f függvényre az

$$\Big\{s(f,\tau)\ \Big|\ \tau\in\mathcal{F}(I)\Big\} \quad \text{és az} \quad \Big\{S(f,\tau)\ \Big|\ \tau\in\mathcal{F}(I)\Big\}$$

halmazok korlátosak. Az

$$I_*(f) := \sup\Big\{s(f,\tau)\ \Big|\ \tau\in\mathcal{F}(I)\Big\}$$

valós számot az f függvény ***Darboux-féle alsó integráljának***, az

$$I^*(f) := \inf\Big\{S(f,\tau)\ \Big|\ \tau\in\mathcal{F}(I)\Big\}$$

valós számot pedig az f függvény ***Darboux-féle felső integráljának*** nevezzük.

1. Definíció. *Akkor mondjuk, hogy a korlátos $f : I \to \mathbb{R}$ függvény* **Riemann-integrálható** (*röviden* **integrálható**) *az I intervallumon (jelben $f \in R(I)$), ha $I_*(f) = I^*(f)$. A közös $I_*(f) = I^*(f)$ számot az f függvény I intervallumon vett* **Riemann-integráljának** (*röviden* **integráljának**) *nevezzük, és az alábbi szimbólumok valamelyikével jelöljük:*

$$\int_I f, \quad \int_I f(x)\,dx, \quad \int\cdots\int_I f(x_1,\dots,x_n)\,dx_1\dots dx_n.$$

Jegyezzük meg, hogy tetszőleges $f : I \to \mathbb{R}$ korlátos függvényre $I_*(f)$ és $I^*(f)$ létezik, mindegyik véges, továbbá bármely két $\tau,\sigma \in \mathcal{F}(I)$ felosztásra $s(f,\tau) \le S(f,\sigma)$, következésképpen

$$I_*(f) \le I^*(f).$$

Az egyváltozós esethez hasonlóan egyszerű példát lehet megadni olyan korlátos f függvényre, amelyre az $I_*(f) < I^*(f)$, ami azt jelenti, hogy a függvény *nem integrálható.*

Példa. Legyen $f : [0,1]\times[0,1] \to \mathbb{R}$,

$$f(x,y) := \begin{cases} 0, & \text{ha } x \text{ és } y \text{ racionális,} \\ 1, & \text{különben.} \end{cases}$$

Ekkor $I_*(f) = 0$ és $I^*(f) = 1$, ezért az f függvény *nem integrálható* a $[0,1]\times[0,1] \subset \mathbb{R}^2$ intervallumon.

A többszörös integrál tulajdonságai

A továbbiakban felsorolt állítások azt fejezik ki, hogy a Riemann-integrálhatóság, illetve maga a Riemann-integrál a többváltozós esetben is rendelkezik az egyváltozós esetben megismert tulajdonságokkal. Az állításokat nem fogjuk bizonyítani.

Először azt fogalmazzuk meg, hogy az említett fogalmak „érzéketlenek” a függvény véges halmazon való „viselkedésére”. Más szóval, ha egy Riemann-integrálható függvényt egy véges halmazon (tetszőlegesen) megváltoztatunk, akkor az így kapott „új” függvény is Riemann-integrálható lesz, és a (Riemann-)integrálja ugyanaz marad, mint a kiindulási függvényé. Tehát, ha egy intervallumon értelmezett két (valós) értékű függvény legfeljebb véges sok helyen különbözik egymástól, akkor vagy mindkettő integrálható (és ekkor az integráljuk megegyezik), vagy egyikük sem integrálható.

1. Tétel. *Tekintsük az $I \subset \mathbb{R}^n$ $(n \in \mathbb{N}^+)$ intervallumon értelmezett $f, g : I \to \mathbb{R}$ korlátos függvényeket. Tegyük fel, hogy az $A := \left\{x \in I \;\middle|\; f(x) \neq g(x)\right\}$ halmaz véges. Ekkor*

a) $f \in R(I) \iff g \in R(I)$,

b) ha $f \in R(I)$, akkor $\int\limits_I f = \int\limits_I g$

A következő tételben kiderül, hogy a folytonosság „erősebb” tulajdonság a Riemann-integrálhatóságnál.

2. Tétel. *Tegyük fel, hogy az $I \subset \mathbb{R}^n$ $(n \in \mathbb{N}^+)$ intervallumon értelmezett $f : I \to \mathbb{R}$ függvény folytonos. Ekkor $f \in R(I)$, azaz $C(I) \subseteq R(I)$.*

Az állítás megfordítása nem igaz. Az $n = 1$ esetben például az

$$f(x) := \begin{cases} 1, & \text{ha } 0 < x \leq 1 \\ 0, & \text{ha } x = 0. \end{cases}$$

függvényre $f \in R[0,1]$, de $f \notin C[0,1]$.

Megjegyzés. Az előző tételekből következik, hogy a véges sok szakadási hellyel rendelkező korlátos függvények integrálhatók. Kérdés, hogy a szakadási helyek számát valamilyen értelemben lehet-e növelni úgy, hogy a függvény továbbra is integrálható maradjon. Kiderül, hogy egy függvény Riemann-integrálhatósága lényegében azon múlik, hogy a függvény szakadási helyeinek a halmaza mennyire „kicsi”.

Precízen: Azt mondjuk, hogy az $A \subset \mathbb{R}^n$ halmaz ***nullmértékű***, ha tetszőleges $\varepsilon > 0$ számhoz megadható $I_k \subset \mathbb{R}^n$ $(k \in \mathbb{N}^+)$ n-dimenziós intervallumoknak egy olyan sorozata, hogy

$$A \subseteq \bigcup_{k=1}^{+\infty} I_k \qquad \text{és} \qquad \sum_{k=1}^{+\infty} |I_k| < \varepsilon.$$

A Riemann-integrálhatóság Lebesgue-kritériuma: *Tegyük fel, hogy az $I \subset \mathbb{R}^n$ n-dimenziós intervallumon értelmezett $f \colon I \to \mathbb{R}$ függvény korlátos, és legyen az f szakadási helyeinek a halmaza*

$$\mathcal{A} := \left\{x \in I \;\middle|\; f \notin C\{x\}\right\}.$$

Ekkor $f \in R(I)$ azzal ekvivalens, hogy az $\mathcal{A}$ halmaz nullmértékű.

Az integrálás és bizonyos függvényműveletek kapcsolatára vonatkozik az alábbi állítás.

3. Tétel. *Legyen $I \subset \mathbb{R}^n$ $(n \in \mathbb{N}^+)$ egy intervallum, és tegyük fel, hogy $f, g \in R(I)$. Ekkor*

a) minden $\alpha, \beta \in \mathbb{R}$ esetén

$$\alpha f + \beta g \in R(I) \quad \text{és} \quad \int_I (\alpha f + \beta g) = \alpha \int_I f + \beta \int_I g,$$

b) $f \cdot g \in R(I)$,

c) ha valamilyen $m > 0$ állandóval fennáll az

$$|g(x)| \geq m > 0 \quad (x \in I)$$

egyenlőtlenség, akkor az $\dfrac{f}{g}$ függvény is integrálható az I intervallumon.

A többváltozós Riemann-integrál is rendelkezik az egydimenziós esetben megismert monotonitási tulajdonsággal. Ezt fejezi ki az alábbi állítás.

4. Tétel. *Tegyük fel, hogy $f \in R(I)$, ahol I egy intervallum. Ekkor*

a) ha $g \in R(I)$, és $f(x) \leq g(x)$ $(x \in I)$, akkor $\int_I f \leq \int_I g$,

b) $|f| \in R(I)$, és $\left|\int_I f\right| \leq \int_I |f|$.

Szukcesszív integrálás

Egy n-dimenziós intervallumon értelmezett függvény integráljának a kiszámítását vissza lehet vezetni valós-valós függvények egymásra következő (szukcesszív) integráljának a kiszámolására. A tételt $n = 2$-re fogjuk kimondani, az ún. ***kettős integrálokra***, de hasonlóan általánosítható $n > 2$-re is.

A továbbiakban feltesszük, hogy adott egy

$$I := I_1 \times I_2 := [a, b] \times [c, d] \subset \mathbb{R}^2$$

kétdimenziós intervallum és egy $f : I \to \mathbb{R}$ *korlátos* függvény.

Kétváltozós függvény viselkedésének az áttekintését megkönnyítheti, ha az egyik változóját rögzítjük, és a függvényt a másik változó függvényének fogjuk fel. Az így kapott függvények az eredeti függvény ún. **szekciófüggvényei**.

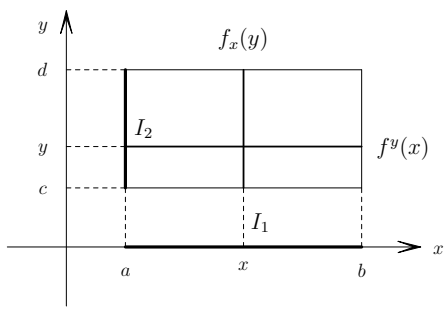
Ha $f : I_1 \times I_2 \to \mathbb{R}$ adott kétváltozós függvény, akkor tetszőlegesen rögzített $x \in I_1$ esetén az

$$f_x : I_2 \to \mathbb{R}, \quad f_x(y) := f(x, y) \qquad (y \in I_2);$$

tetszőlegesen rögzített $y \in I_2$ esetén az

$$f^y : I_1 \to \mathbb{R}, \quad f^y(x) := f(x, y) \qquad (x \in I_1)$$

az f függvény szekciófüggvényei.



5. Tétel (Fubini-tétel). *Legyen $I = [a,b] \times [c,d]$ és $f : I \to \mathbb{R}$. Tegyük fel, hogy*

a) $f \in R(I)$,

b) $\forall x \in [a,b]\colon f_x \in R[c,d]$,

c) $\forall y \in [c,d]\colon f^y \in R[a,b]$.

Ekkor

$$\iint\limits_I f(x,y)\,dx\,dy = \int\limits_a^b \left(\int\limits_c^d f(x,y)\,dy \right) dx = \int\limits_c^d \left(\int\limits_a^b f(x,y)\,dx \right) dy. \tag{2}$$

Bizonyítás. A tételt nem bizonyítjuk.

Megjegyzés.

1. Ha az f függvény *folytonos* az I téglalapon, akkor $f \in R(I)$, illetve az f_x ($x \in [a,b]$) és az f^y ($y \in [c,d]$) szekciófüggvények is folytonosak, következésképpen Riemann-integrálhatóak. Így a tétel feltételei teljesülnek. Ebben az esetben az állítás *egyszerűen* bebizonyítható. Ennek a fontos speciális esetnek a felfedezése Leonhard Euler (1707–1783) érdeme. Euler eredményét *Guido Fubini* (1879–1943) általánosította *integrálható függvényekre*. Az „igazi" Fubini-tétel ennél sokkal általánosabb érvényű az ún. Lebesgue-integrálható, ill. az absztrakt integrálható függvények elméletében.

2. Formálisan megfogalmazva tehát a fenti feltételek teljesülése esetén egy kétváltozós függvény integrálját kiszámíthatjuk úgy is, hogy az egyik változót először (tetszőlegesen) rögzítjük, és a másik változó szerint integrálunk, majd az így kapott (a rögzített változótól függő) integrált integráljuk. (Innen ered a *szukcesszív* (egymás utáni) jelző.) Az (2) egyenlőség azt is állítja, hogy az integrálást bármelyik változóval kezdhetjük, tehát az ***integrálás sorrendje felcserélhető***.

Példa. Számítsuk ki a következő kettős integrált!

$$\iint\limits_I \left(x^3y + xy^2 + 1 \right) dx\,dy \qquad \Big(I := [0,1] \times [1,2] \Big).$$

Az integrálandó $f(x,y) := x^3y + xy^2 + 1$ $\Big((x,y) \in \mathbb{R}^2\Big)$ függvény folytonos I-n, ezért $f \in R(I)$. A Fubini-tétele alapján mindegy, hogy milyen sorrendben integrálunk, az eredmény ugyanaz lesz. Ezzel kétféle módon tudjuk kiszámítani az integrált. Ha először az $x \in [0,1]$ változót (tetszőlegesen) rögzítjük, és az y változó szerint integrálunk, akkor

$$\iint\limits_I (x^3y + xy^2 + 1)\,dx\,dy = \int\limits_0^1 \left(\int\limits_1^2 (x^3y + xy^2 + 1)\,dy \right) dx = \int\limits_0^1 \left[\frac{x^3y^2}{2} + \frac{xy^3}{3} + y \right]_{y=1}^{y=2} dx =$$

$$= \int\limits_0^1 \left(\left(\frac{x^3 \cdot 2^2}{2} + \frac{x \cdot 2^3}{3} + 2 \right) - \left(\frac{x^3 \cdot 1^2}{2} + \frac{x \cdot 1^3}{3} + 1 \right) \right) dx = \int\limits_0^1 \left(\frac{3x^3}{2} + \frac{7x}{3} + 1 \right) dx =$$

$$= \left[\frac{3x^4}{8} + \frac{7x^2}{6} + x \right]_{x=0}^{x=1} = \left(\frac{3}{8} + \frac{7}{6} + 1 \right) - 0 = \frac{61}{24}.$$

Másképpen, ha először az $y \in [1,2]$ változót (tetszőlegesen) rögzítjük, és az x változó szerint integrálunk, akkor

$$\iint\limits_I (x^3y + xy^2 + 1)\,dx\,dy = \int\limits_1^2 \left(\int\limits_0^1 (x^3y + xy^2 + 1)\,dx \right) dy = \int\limits_1^2 \left[\frac{x^4y}{4} + \frac{x^2y^2}{2} + x \right]_{x=0}^{x=1} dy =$$

$$= \int\limits_1^2 \left(\left(\frac{1^4 \cdot y}{4} + \frac{1^2 \cdot y^2}{2} + 1 \right) - 0 \right) dy = \int\limits_1^2 \left(\frac{y}{4} + \frac{y^2}{2} + 1 \right) dy = \left[\frac{y^2}{8} + \frac{y^3}{6} + y \right]_{y=1}^{y=2} =$$

$$= \left(\frac{2^2}{8} + \frac{2^3}{6} + 2 \right) - \left(\frac{1^2}{8} + \frac{1^3}{6} + 1 \right) = \frac{61}{24}.$$

Megjegyzés. A szukcesszív integrálás tétele azt állítja, hogy (a tétel feltételeinek a teljesülése esetén) mindegy, hogy melyik sorrendben integrálunk, az eredmény ugyanaz lesz. Ez azonban *nem* jelenti azt, hogy a kétféle sorrendben történő kiszámolás során ugyanolyan technikai jellegű nehézségek lépnek fel.

Példa. Számítsuk ki a következő kettős integrált!

$$\iint\limits_I x \cdot \sin(xy)\,dx\,dy \qquad \left(I := [1,3] \times [0, \tfrac{\pi}{2}] \right).$$

Az integrálandó $f(x,y) := x \cdot \sin(xy)$ $\left((x,y) \in \mathbb{R}^2 \right)$ függvény folytonos I-n, ezért $f \in R(I)$. A Fubini-tétele alapján mindegy, hogy milyen sorrendben integrálunk, az eredmény ugyanaz lesz.

Ha először az $x \in [1,3]$ változót (tetszőlegesen) rögzítjük, és az y változó szerint integrálunk, akkor az

$$(*) \qquad \int\limits_1^3 \left(\int\limits_0^{\pi/2} x \cdot \sin(xy)\,dy \right) dx$$

egyváltozós integrálokat kell egymás után kiszámolni.

Ha először az $y \in \left[0, \frac{\pi}{2}\right]$ változót (tetszőlegesen) rögzítjük, és az x változó szerint integrálunk, akkor pedig azt kapjuk, hogy

$$(**) \qquad \int\limits_0^{\pi/2} \left(\int\limits_1^3 x \cdot \sin(xy)\,dx \right) dy$$

Vegyük észre azonban azt, hogy a $(**)$ esetben először parciálisan kell integrálni, a $(*)$ esetben pedig a belső integrált rögtön kiszámíthatjuk. Ezért a $(*)$ alatti sorrendben integrálunk:

$$\iint\limits_I x \cdot \sin(xy)\,dx\,dy = \int\limits_1^3 \left(\int\limits_0^{\pi/2} x \cdot \sin(xy)\,dy \right) dx = \int\limits_1^3 \Big[-\cos(xy) \Big]_{y=0}^{y=\pi/2} dx = \int\limits_1^3 \left(-\cos\frac{\pi x}{2} + \cos 0 \right) dx =$$

$$= \int\limits_1^3 \left(-\cos\frac{\pi x}{2} + 1 \right) dx = \left[-\frac{\sin\frac{\pi x}{2}}{\frac{\pi}{2}} + x \right]_1^3 = \left(-\frac{2}{\pi}\sin\frac{3\pi}{2} + 3 \right) - \left(-\frac{2}{\pi}\sin\frac{\pi}{2} + 1 \right) = 2 + \frac{4}{\pi}.$$

A többszörös integrálok értelmezése $\mathbb{R}^n$-beli korlátos halmazokon

Az integrálhatóság fogalma egyszerűen kiterjeszthető *tetszőleges* korlátos $H \subset \mathbb{R}^n$-beli halmazokon értelmezett *korlátos* függvényekre. Legyen H egy ilyen halmaz és $f : H \to \mathbb{R}$ egy adott korlátos függvény. Ekkor van olyan n-dimenziós I intervallum, amelyre $H \subseteq I$. Terjesszük ki az f függvény értelmezését az I intervallumra a következőképpen:

$$\tilde{f}(x) := \begin{cases} f(x), & \text{ha } x \in H, \\ 0, & \text{ha } x \in I \setminus H. \end{cases}$$

Azt mondjuk, hogy az $f : H \to \mathbb{R}$ függvény ***(Riemann)-integrálható a H halmazon*** (jelben $f \in R(H)$), ha az $\tilde{f} : I \to \mathbb{R}$ függvény integrálható az I intervallumon. Ekkor legyen

$$\int\limits_H f := \int\limits_I \tilde{f}.$$

Egyszerűen belátható, hogy ez az értelmezés *független* a H-t tartalmazó intervallum megválasztásától.

A kettős integrálok geometriai jelentései

Ha az $f \in \mathbb{R}^2 \to \mathbb{R}$ korlátos függvény integrálható a $H \subset \mathbb{R}^2$ korlátos halmazon, akkor az integrálját a korábban alkalmazott jelölésekhez hasonlóan a

$$\iint\limits_H f \quad \text{vagy az} \quad \iint\limits_H f(x,y)\,dx\,dy$$

szimbólumok valamelyikével fogjuk jelölni.

Az egyváltozós esetben már definiáltuk egy korlátos, nem negatív függvény *grafikonja alatti síkidom területét*. Kettős integrálokkal bármely *korlátos síkidom területét* is értelmezhetjük.

2. Definíció. *Legyen $H \subset \mathbb{R}^2$ egy korlátos halmaz és $f(x,y) := 1 \ \big((x,y) \in H\big)$. Azt mondjuk, hogy a H síkidomnak* ***van területe****, ha $f \in R(H)$, és ekkor H* ***területét*** *a*

$$t(H) := \iint\limits_H f = \iint\limits_H 1\,dx\,dy$$

kettős integrállal értelmezzük.

Megjegyzés. Igazolható, hogy függvények grafikonja alatti síkidomok esetén a két definíció ekvivalens.

Most arra emlékeztetünk, hogy az egyváltozós analízisben már értelmeztük *forgástestek térfogatát*, és annak kiszámolására egy képletet is megismertünk.

Kettős integrálok alsó- és a felső közelítő összegeinek geometriai jelentése alapján kézenfekvő bizonyos „térrészek” (pl. kétváltozós függvény grafikonja alatti tartományok) térfogatának az alábbi értelmezése.

3. Definíció. *Legyen $H \subset \mathbb{R}^2$ korlátos halmaz és $f : H \to \mathbb{R}$ nemnegatív korlátos függvény. Azt mondjuk, hogy a*

$$T := \Big\{(x,y,z) \;\Big|\; (x,y) \in H, \quad 0 \le z \le f(x,y)\Big\} \subset \mathbb{R}^3$$

„térrésznek" (hengerszerű testnek) **van térfogata**, *ha $f \in R(H)$. Ekkor a*

$$V(T) := \iint\limits_H f = \iint\limits_H f(x,y)\,dx\,dy$$

kettős integrált a T test **térfogatának** *nevezzük.*

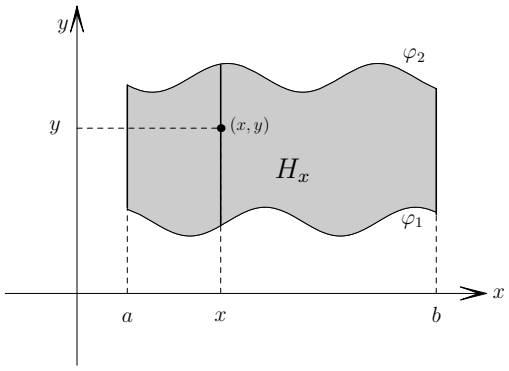
Kettős integrál kiszámítása normáltartományon

Gyakran előfordul, hogy két függvény által határolt tartományon kell egy integrált kiszámítani.

Legyenek $\varphi_1, \varphi_2 : [a,b] \to \mathbb{R}$ folytonos függvények, és tegyük fel, hogy $\varphi_1(x) \le \varphi_2(x)$ minden $x \in [a,b]$ esetén. A

$$H_x := \Big\{(x,y) \in \mathbb{R}^2 \;\Big|\; a \le x \le b,\ \varphi_1(x) \le y \le \varphi_2(x)\Big\}$$

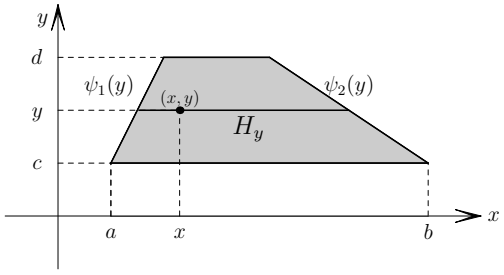
halmazt a *x tengelyre nézve normáltartománynak* nevezzük.



Legyenek $\psi_1, \psi_2 : [c,d] \to \mathbb{R}$ folytonos függvények, és tegyük fel, hogy $\psi_1(y) \le \psi_2(y)$ minden $x \in [a,b]$ esetén. A

$$H_y := \Big\{(x,y) \in \mathbb{R}^2 \;\Big|\; c \le y \le d,\ \psi_1(y) \le x \le \psi_2(y)\Big\}$$

halmazt a *y tengelyre nézve normáltartománynak* nevezzük.



Az eddigiekből egyszerűen adódnak az alábbi fontos állítások, amelyeket bizonyítás nélkül kimondunk.

6. Tétel (Integrálás H_x normáltartományon). *Legyen H_x az x tengelyre nézve normáltartomány, és tegyük fel, hogy az $f : H_x \to \mathbb{R}$ függvény folytonos. Ekkor $f \in R(H_x)$ és*

$$\iint\limits_{H_x} f(x,y)\,dx\,dy = \int\limits_a^b \left(\int\limits_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)\,dy \right) dx.$$

7. Tétel (Integrálás H_y normáltartományon). *Legyen H_y az y tengelyre nézve normáltartomány, és tegyük fel, hogy az $f : H_y \to \mathbb{R}$ függvény folytonos. Ekkor $f \in R(H_y)$ és*

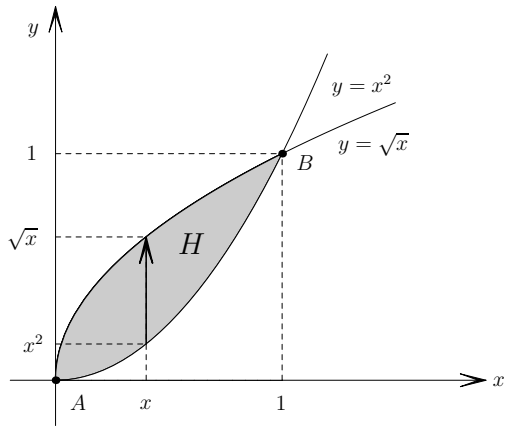
$$\iint\limits_{H_y} f(x,y)\,dx\,dy = \int\limits_c^d \left(\int\limits_{\psi_1(y)}^{\psi_2(y)} f(x,y)\,dx \right) dy.$$

Példa. Számítsuk ki a következő kettős integrált:

$$\iint\limits_H xy^2\,dx\,dy,$$

ahol H az $y = x^2$ és az $y = \sqrt{x}$ egyenletű görbék által közrezárt korlátos síkrész!

Ábrázoljuk a H halmazt!



Az integrandus folytonos az egész $\mathbb{R}^2$-őn, tehát a korlátos H halmazon is. Következésképpen $f \in R(H)$.

Az integrál kiszámításához először a görbék metszéspontjainak a koordinátáit határozzuk meg:

$$\left.\begin{array}{l} y = x^2 \\ y = \sqrt{x} \end{array}\right\} \iff \sqrt{x} = x^2 \iff x = 0 \text{ v. } x = 1.$$

A metszéspontok tehát $A(0,0)$ és $B(1,1)$.

A H halmaz a x tengelyre és az y tengelyre nézve is normáltartomány, ezért mindegyik megismert képletet használhatjuk. (érdemes arra is figyelni, hogy mindegyik esetben a „belső” integrálokat könnyen kiszámolhatjuk, ezért bármelyik változó szerinti integrálással kezdhetünk.)

Tekintsük a H halmazt az x tengelyre nézve normáltartománynak:

$$0 \le x \le 1, \qquad x^2 \le y \le \sqrt{x}.$$

Ekkor először y szerint kell integrálnunk. (A nyíl jelzi a „belső” integrál irányát.) Így

$$\iint\limits_H xy^2\,dx\,dy = \int\limits_0^1 \left(\int\limits_{x^2}^{\sqrt{x}} xy^2\,dy \right) dx = \int\limits_0^1 \left[x \cdot \frac{y^3}{3} \right]_{y=x^2}^{y=\sqrt{x}} dx = \frac{1}{3} \int\limits_0^1 x \cdot \left(x^{3/2} - x^6 \right) dx =$$

$$= \frac{1}{3} \int\limits_0^1 \left(x^{5/2} - x^7 \right) dx = \frac{1}{3} \cdot \left[\frac{x^{7/2}}{7/2} - \frac{x^8}{8} \right]_0^1 = \frac{1}{3} \cdot \left(\frac{1}{7/2} - \frac{1}{8} \right) = \frac{3}{56}.$$

Tegyük fel, hogy a H integrálási tartomány az x tengelyre és az y tengelyre nézve is normáltartomány és az f függvény folytonos H-n. Ekkor a fenti tétel szerint az $\iint_H f$ kettős integrált kétféle sorrendben is kiszámíthatjuk. Az integrálás sorrendjének felcserélésénél azonban körültekintően kell eljárnunk.

Példa. Tegyük fel, hogy $f : \mathbb{R}^2 \to \mathbb{R}$ folytonos függvény. Tekintsük a következő integrált:

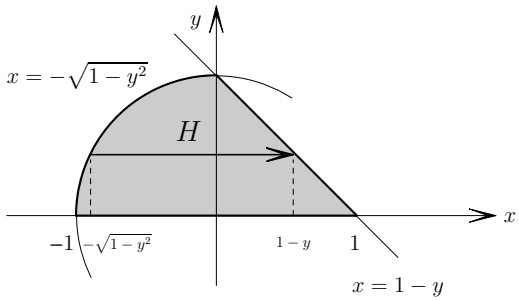
$$\int\limits_0^1 \left(\int\limits_{-\sqrt{1-y^2}}^{1-y} f(x,y)\,dx \right) dy\,.$$

Szemléltessük az integrálási tartományt! Cseréljük fel az integrálás sorrendjét!

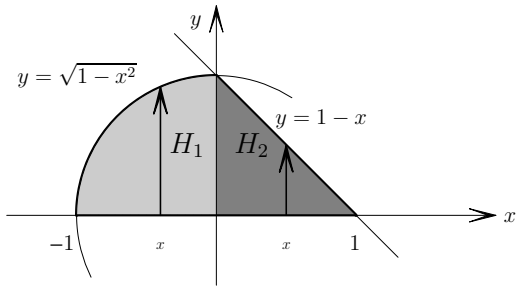
A H-val jelölt integrálási tartomány a

$$0 \le y \le 1, \quad -\sqrt{1-y^2} \le x \le 1-y$$

egyenlőtlenségeknek eleget tevő $(x,y) \in \mathbb{R}^2$ pontok halmaza. Mivel H az y tengelyre nézve normáltartomány, ezért először az x változó, utána pedig az y változó szerint integrálunk. (A nyíl jelzi a belső integrálban az integrálás irányát.)



Ha a H halmazt az x tengelyre nézve normáltartománynak tekintjük, akkor először az y, utána pedig az x változó szerint kell integrálnunk. Az f függvény folytonos $\mathbb{R}^2$-őn (tehát a H halmazon is), ezért az integrálás sorrendje felcserélhető. A H tartományt ebben az esetben az alábbi módon két részre bontjuk.



A tartományokat a következő egyenlőtlenségrendszerek határozzák meg:

$$H_1: \quad -1 \le x \le 0, \quad 0 \le y \le \sqrt{1-x^2},$$

$$H_2: \quad 0 \le x \le 1, \quad 0 \le y \le 1-x.$$

Így

$$\iint\limits_{H_1} f = \int\limits_{-1}^{0} \left(\int\limits_{0}^{\sqrt{1-x^2}} f(x,y)\,dy \right) dx, \qquad \iint\limits_{H_2} f = \int\limits_{0}^{1} \left(\int\limits_{0}^{1-x} f(x,y)\,dy \right) dx.$$

Mivel

$$\iint\limits_{H} f = \iint\limits_{H_1} f + \iint\limits_{H_2} f,$$

ezért

$$\int\limits_{0}^{1} \left(\int\limits_{-\sqrt{1-y^2}}^{1-y} f(x,y)\,dx \right) dy = \int\limits_{-1}^{0} \left(\int\limits_{0}^{\sqrt{1-x^2}} f(x,y)\,dy \right) dx + \int\limits_{0}^{1} \left(\int\limits_{0}^{1-x} f(x,y)\,dy \right) dx.$$

Itt emlékeztetünk arra, hogy egyváltozós határozott integrálok kiszámításához bizonyos esetekben a Newton–Leibniz-formula nem használható. Ez a helyzet például akkor, amikor az integrálandó függvénynek van primitív függvénye (mert pl. folytonos), azonban a primitív függvény nem elemi függvény. Bebizonyítható, hogy ilyenek a következő függvények:

$$e^{\pm x^2} \ \ (x \in \mathbb{R}), \quad \sin x^2 \ \ (x \in \mathbb{R}), \quad \cos x^2 \ \ (x \in \mathbb{R}),$$

$$\frac{\sin x}{x} \ \ \big(x \in (0,+\infty)\big), \quad \frac{\cos x}{x} \ \ \big(x \in (0,+\infty)\big), \quad \frac{e^x}{x} \ \ \big(x \in (0,+\infty)\big),$$

$$\frac{1}{\ln x} \ \ \big(x \in (0,+\infty)\big), \sqrt{x^3+1} \ \ \big(x \in (0,+\infty)\big).$$

Ha egy kettős integrál kiszámolásánál ilyen függvények adódnak, akkor bizonyos esetekben az integrálás sorrendjének a felcserélése segíthet.

Példa. Számítsuk ki a

$$\iint\limits_H y \sin x^2 \, dx \, dy \qquad H := \{(x,y) \in \mathbb{R}^2 \colon 0 \le y \le 1,\ y^2 \le x \le 1\}$$

integrált!

A H-val jelölt integrálási halmaz az

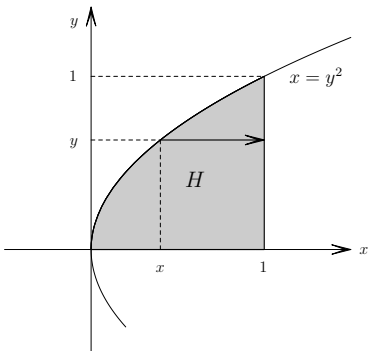
$$y^2 \le x \le 1, \quad 0 \le y \le 1$$

egyenlőtlenségekkel meghatározott y tengelyre nézve normáltartomány (l. az (a) ábrát).

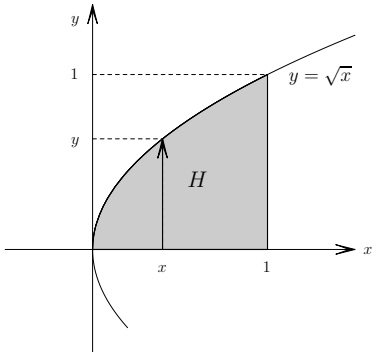
Ezért

$$\iint\limits_H y \sin x^2 \, dx \, dy = \int\limits_0^1 \left(\int\limits_{y^2}^1 y \sin x^2 \, dx \right) dy.$$

Ha a fenti képlet szerint először x szerint integrálunk, akkor a következő problémába ütközünk: A $\sin x^2$ ($x \in \mathbb{R}$) függvénynek *van* primitív függvénye (hiszen folytonos), de az *nem elemi függvény*, így a belső (egyváltozós) integrál kiszámítására a Newton–Leibniz-tétel *nem alkalmazható*. Próbáljuk meg az integrálás sorrendjét felcserélni, azaz először y szerint integrálni. Ezt megtehetjük, mert a szóban forgó halmaz az x tengelyre nézve is normáltartomány, amelyet a (b) ábra alatti egyenlőtlenségek határoznak meg.



(a) ábra
H az y-ra normáltartomány
$0 \le y \le 1, \quad y^2 \le x \le 1$



(b) ábra
H az x-re normáltartomány
$0 \le x \le 1, \quad 0 \le y \le \sqrt{x}$

(A nyíl jelzi az eredeti felírásban, illetve az integrálok felcserélése után a belső integrálok irányát.)

Így

$$\iint\limits_H y \sin x^2 \, dx \, dy = \int\limits_0^1 \left(\int\limits_0^{\sqrt{x}} y \sin x^2 \, dy \right) dx = \int\limits_0^1 \left(\sin x^2\right) \cdot \left[\frac{y^2}{2}\right]_{y=0}^{y=\sqrt{x}} dx =$$

$$= \frac{1}{2} \int\limits_0^1 x \cdot \sin x^2 \, dx = \frac{1}{2} \cdot \frac{1}{2} \int\limits_0^1 2x \cdot \sin x^2 \, dx = \frac{1}{4}\Big[-\cos x^2\Big]_0^1 = \frac{1}{4}\Big(1 - \cos 1\Big).$$

13. előadás

TÖBBSZÖRÖS INTEGRÁLOK 2.

Kettős integrálok kiszámítása integráltranszformációval

A helyettesítéssel való integrálást illetően idézzük fel a valós-valós függvényekre vonatkozó állításokat. Először a határozatlan integrálokkal kapcsolatos ***második helyettesítési szabályra*** emlékeztetünk:

Legyenek $I, J \subset \mathbb{R}$ nyílt intervallumok. Tegyük fel, hogy $f : I \to \mathbb{R}$, $g : J \to I$, $\mathcal{R}_g = I$, $g \in D(J)$, $g' > 0$ J-n (vagy $g' < 0$ J-n) és az $(f \circ g) \cdot g' : J \to \mathbb{R}$ függvénynek van primitív függvénye. Ekkor az f függvénynek is van primitív függvénye és

$$\int f(x)\,dx \underset{x=g(t)}{=} \int f\Big(g(t)\Big) \cdot g'(t)\,dt\Big|_{t=g^{-1}(x)} \qquad (x \in I).$$

Tegyük fel, hogy egy $\int f(x)\,dx$ alakú határozatlan integrált akarunk kiszámítani. Olyan g-t keresünk, amelyre az $\int f\Big(g(t)\Big) \cdot g'(t)\,dt$ integrált ki tudjuk számítani. E cél érdekében általában olyan g függvényt próbálunk választani, amelyre $f \circ g \cdot g'$ egyszerűbb, mint f.

A Newton–Leibniz-formulából következik a helyettesítéssel való integrálás (vagyis az integráltranszformációs formula) határozott integrálokra vonatkozó alábbi változata: Tegyük fel, hogy az $f \in R[a,b]$ függvénynek van primitív függvénye $[a,b]$-n, és $g : [\alpha, \beta] \to [a,b]$ olyan folytonos bijekció, ami differenciálható (α, β)-n. Ekkor

$$(*) \qquad \int_\alpha^\beta f \circ g \cdot g' = \int_{g(\alpha)}^{g(\beta)} f.$$

A feltételek szerint g szigorúan monoton $[\alpha, \beta]$-n.

Ha $g \uparrow [\alpha, \beta]$-n, akkor $g(\alpha) = a$ és $g(\beta) = b$, ezért $(*)$-ból következik, hogy

$$\int_a^b f = \int_\alpha^\beta f \circ g \cdot g'.$$

Ha viszont $g \downarrow [\alpha, \beta]$-n, akkor $g(\alpha) = b$ és $g(\beta) = a$, ezért $(*)$-ból azt kapjuk, hogy

$$\int_\alpha^\beta f \circ g \cdot g' = \int_b^a f = -\int_a^b f, \quad \Longrightarrow \quad \int_a^b f = \int_\alpha^\beta f \circ g \cdot \Big(-g'\Big).$$

Összefoglalva a következő állítás igaz:

Tegyük fel, hogy $f \in R[a,b]$, $g : [\alpha, \beta] \to [a,b]$ folytonosan deriválható bijekció és $g'(t) \neq 0$ $\Big(t \in [\alpha, \beta]\Big)$. Ekkor

$$\int_a^b f = \int_\alpha^\beta f \circ g \cdot \Big|g'\Big|.$$

Az előző állítás általánosítása többszörös integrálokra már jóval bonyolultabb. Valós-valós esetben az f és a g függvények intervallumokon értelmeztük. Többdimenziós esetben olyan H halmazok kerülnek szóba, amelyeknek „van mértékük”, és így érdemes ezeken az integrált értelmezni. Emlékezzünk arra, hogy egy $H \subset \mathbb{R}^2$ korlátos síkidomnak, akkor van területe, ha a konstans 1 függvény Riemann-integrálható H-n, és ekkor a területét a

$$t(H) := \iint\limits_H 1 \, dx \, dy$$

kettős integrállal értelmezzük. Hasonlóan járunk el egy $H \subset \mathbb{R}^3$ téridom térfogatának értelmezésekor hármas integrállal. Általánosan, akkor mondjuk, hogy a $H \subset \mathbb{R}^n \quad (n \in \mathbb{N}^+)$ korlátos halmaz Jordan-mérhető, ha a konstans 1 függvény Riemann-integrálható H-n, és ekkor a H halmaz Jordan-mértéke a

$$\mu(H) := \int\limits_H 1$$

integrállal értelmezzük.

1. Tétel (Integráltranszformáció). *Legyen $U \subseteq \mathbb{R}^n \quad (n \in \mathbb{N}^+)$ egy nem üres nyílt halmaz, és $H \subset U$ egy nem üres, Jordan-mérhető és zárt halmaz. Tegyük fel, hogy*

a) $g : U \to \mathbb{R}^n$ egy folytonosan differenciálható függvény,

b) a g függvény injektív a H halmaz belsejében, azaz $g_{|\operatorname{int} H}$ invertálható.

Ekkor a $g[H]$ halmaz is Jordan-mérhető, illetve az $f \colon g[H] \to \mathbb{R}$ korlátos függvény akkor és csak akkor integrálható, ha a

$$H \ni t \mapsto f\Big(g(t)\Big) \cdot \Big|\det g'(t)\Big|$$

függvény is integrálható, és

$$\int\limits_{g[H]} f(x) \, dx = \int\limits_H f\Big(g(t)\Big) \cdot \Big|\det g'(t)\Big| \, dt.$$

Bizonyítás. A tételt nem bizonyítjuk.

Megjegyzés. Az alkalmazások szempontjából az integráltranszformációra két okból is szükség lehet. Egyrészt, ha H olyan tartomány, amelyen az integrált csak „körülményesen” lehet kiszámolni, akkor *kereshetünk* olyan g-t, amely már egy „egyszerűbb” halmazon van értelmezve (pl. téglalapon), ezért a jobb oldali integrált könnyebb kiszámolni. Másrészt előfordulhat az is, hogy sikerül olyan g függvényt *találni*, amelyre $f \circ g \cdot |\det g'|$ egyszerűbb, mint f.

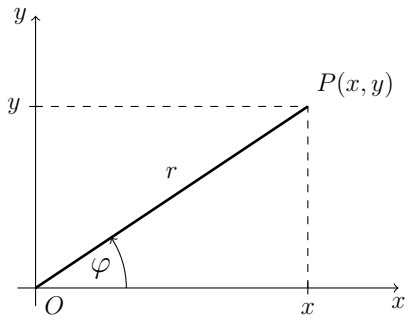
A következőben néhány nevezetes integráltranszformációt fogjuk bemutatni.

Síkbeli polárkoordináta-transzformáció:

Sok esetben a sík Descates-féle derékszögű koordinátarendszere helyett/mellett célszerű ***polárkoordináta-rendszert*** bevezetni a következő módon. Kiválasztunk a síkon egy rögzített O pontot (pólus) és egy ebből kiinduló félegyenest (polártengely). A pólustól különböző P pont *polárkoordinátáin* az (r, φ) számpárt értjük, ahol $r = \overline{OP}$ és φ az $\overrightarrow{OP}$ félegyenesnek a polártengellyel bezárt szöge.

Világos, hogy r és φ egyértelműen meghatározza a P pont helyzetét, ezzel szemben a P pont csak r-et határozza meg egyértelműen, a φ szöget csak 2π egész számú többszörösétől eltekintve. Az O pont polárszöge határozatlan.

A vizsgálataink során gyakran egymás mellett használjuk a Descates-féle derékszögű és a polárkoordináta-rendszert. Ha a kétféle koordinátarendszer kezdőpontja, valamint a polártengely és az x tengely pozitív fele egybeesik, akkor a következő összefüggések állnak fenn az (x, y) derékszögű és az (r, φ) polárkoordináták között:



$$\begin{aligned} x &= r\cos\varphi \\ y &= r\sin\varphi \end{aligned} \qquad \longleftrightarrow \qquad r = \sqrt{x^2+y^2}, \quad \varphi = \begin{cases} \operatorname{arc\,tg}\frac{y}{x} & \text{ha } x>0,\ y\geq 0 \\ 2\pi + \operatorname{arc\,tg}\frac{y}{x} & \text{ha } x>0,\ y<0 \\ \pi + \operatorname{arc\,tg}\frac{y}{x} & \text{ha } x<0, \\ \pi/2 & \text{ha } x=0,\ y>0 \\ 3\pi/2 & \text{ha } x=0,\ y<0. \end{cases}$$

Síkbeli polárkoordináta-transzformációról beszélünk, ha a

$$g(r,\varphi) := (r\cos\varphi,\ r\sin\varphi) \qquad \big((r,\varphi)\in\mathbb{R}^2\big)$$

leképezést alkalmazzuk. Világos, hogy $g \in C^1(\mathbb{R}^2)$, illetve

$$g'(r,\varphi) = \begin{pmatrix} \partial_r(r\cos\varphi) & \partial_\varphi(r\cos\varphi) \\ \partial_r(r\sin\varphi) & \partial_\varphi(r\sin\varphi) \end{pmatrix} = \begin{pmatrix} \cos\varphi & -r\ \sin\varphi \\ \sin\varphi & r\ \cos\varphi \end{pmatrix} \qquad \big((r,\varphi)\in\mathbb{R}^2\big).$$

Ezért

$$\det g'(r,\varphi) = r\cos^2\varphi + r\sin^2\varphi = r(\cos^2\varphi+\sin^2\varphi) = r.$$

Adott $R>0$, legyen

$$H \subseteq [0,R]\times[0,2\pi]$$

egy nem üres, Jordan-mérhető, zárt halmaz. $g_{|\operatorname{int} H}$ invertálható, azaz

$$\begin{aligned} &(r_1,\varphi_1),\ (r_2,\varphi_2) \in \operatorname{int} H \\ &g(r_1,\varphi_1) = g(r_2,\varphi_2) \end{aligned} \qquad \Longrightarrow \qquad (r_1,\varphi_1) = (r_2,\varphi_2),$$

hiszen nem nehéz igazolni, hogy

$$\begin{aligned} &0 < r_1, r_2 < R,\ 0<\varphi_1,\varphi_2<2\pi \\ &r_1\cos\varphi_1 = r_2\cos\varphi_2,\ r_1\sin\varphi_1 = r_2\sin\varphi_2 \end{aligned} \qquad \Longrightarrow \qquad r_1 = r_2,\ \varphi_1=\varphi_2.$$

Ekkor az integráltranszformációról szóló tétel feltételei teljesülnek, és így az

$$x = r\cos\varphi, \quad y = r\sin\varphi$$

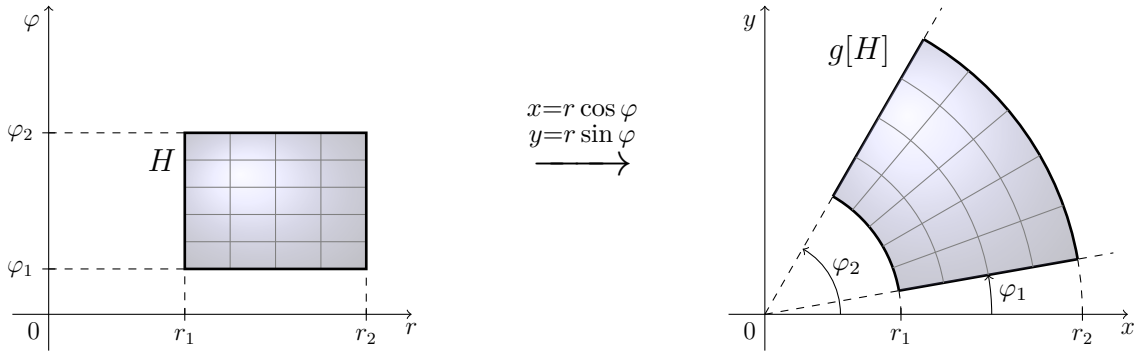
transzformációval

$$(*) \qquad \iint\limits_{g[H]} f(x,y)\,dx\,dy = \iint\limits_{H} f(r\cos\varphi, r\sin\varphi)\cdot r\,dr\,d\varphi,$$

hiszen $\det g'(r,\varphi) = r \geq 0$, ha $H \subseteq [0,R]\times[0,2\pi]$.

Polárkoordináta-transzformációval egy téglalapot körgyűrűcikkbe képezhetünk. Ezt szemlélteti az alábbi ábra, ahol a téglalap:

$$H := \{(r,\varphi) \in \mathbb{R}^2 \colon 0 \le r_1 < r_2,\ 0 \le \varphi_1 < \varphi_2 \le 2\pi\}.$$



Példa. Számítsuk ki a

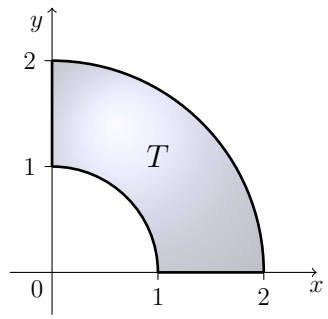
$$\iint\limits_T x^2\, y\, dx\, dy$$

kettős integrált, ahol T az

$$1 \le x^2 + y^2 \le 4, \quad y \ge 0, \quad x \ge 0$$

egyenlőtlenségekkel meghatározott korlátos síkrész!

Megoldás. Az ábra a T-vel jelölt integrálási tartományt szemlélteti.



Az integrandus folytonos, következésképpen integrálható a T halmazon. Az integrál kiszámításához az

$$x = r\cos\varphi, \quad y = r\sin\varphi$$
$$\Big(1 \le r \le 2,\ 0 \le \varphi \le \pi/2\Big)$$

polárkoordinátás helyettesítést alkalmazzuk. A $(*)$ integráltranszformáció alapján $T = g\Big([1,2]\times[0,\pi/2]\Big)$ és

$$\iint\limits_T x^2\, y\, dx\, dy = \iint\limits_{[1,2]\times[0,\pi/2]} (r\cos\varphi)^2 \cdot (r\sin\varphi) \cdot r\, dr\, d\varphi = \iint\limits_{[1,2]\times[0,\pi/2]} r^4 \cdot (\sin\varphi) \cdot \cos^2\varphi\, dr\, d\varphi =$$

$$= \left(\int\limits_1^2 r^4\, dr\right) \cdot \left(\int\limits_0^{\pi/2} \sin\varphi \cos^2\varphi\, d\varphi\right) = \left[\frac{r^5}{5}\right]_{r=1}^{r=2} \cdot \left[-\frac{\cos^3\varphi}{3}\right]_{\varphi=0}^{\varphi=\pi/2} =$$

$$= \left(\frac{2^5}{5} - \frac{1}{5}\right) \cdot \left(-\frac{cos^3(\pi/2)}{3} + \frac{cos^3(0)}{3}\right) = \frac{31}{5} \cdot \frac{1}{3} = \frac{31}{15}.$$

Példa. Kettős integrállal határozzuk meg az R sugarú kör területét!

Megoldás. Jelölje H_R az origó középpontú R sugarú zárt körlapot. Legyen

$$f(x,y) := 1 \quad \Big((x,y) \in H_R\Big).$$

Mivel $f \in R\Big(H_R\Big)$, ezért a H_R halmaznak van területe, és az egyenlő a

$$\iint\limits_{H_R} 1\, dx\, dy$$

kettős integrállal. Ezt az

$$x = r\cos\varphi, \quad y = r\sin\varphi$$
$$\Big(0 \le r \le R,\ 0 \le \varphi \le 2\pi\Big)$$

polárkoordinátás helyettesítéssel számoljuk ki. A megismert képlet alapján

$$\iint\limits_{H_R} 1\,dx\,dy = \iint\limits_{[0,R]\times[0,2\pi]} r\,dr\,d\varphi,$$

és ez utóbbi integrált szukcesszív integrálással könnyű kiszámítani:

$$\iint\limits_{[0,R]\times[0,2\pi]} r\,dr\,d\varphi = \left(\int\limits_0^R r\,dr\right)\cdot\left(\int\limits_0^{2\pi} 1\,d\varphi\right) = \left[\frac{r^2}{2}\right]_0^R \cdot 2\pi = R^2\pi.$$

Az R sugarú kör területére tehát így is megkaphatjuk a jól ismert $R^2\pi$ képletet.

Megjegyzés. Emlékeztetünk arra, hogy az egyváltozós analízisben a félkör területét a

$$\int\limits_{-R}^{R} \sqrt{R^2 - x^2}\,dx$$

határozott integrállal számoltuk ki.

Példa. Kettős integrállal határozzuk meg az R sugarú gömb térfogatát!

Megoldás. Legyen $R > 0$ adott valós szám és

$$f(x,y) := \sqrt{R^2 - x^2 - y^2} \quad \left(H_R := \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; \sqrt{x^2+y^2} \le R\right\}\right).$$

Az f függvény grafikonja az origó középpontú R sugarú gömb felső féltérbe eső felülete; az ez alatti térrész pedig a félgömb. Mivel $f \in C\Big(H_R\Big)$, ezért $f \in R\Big(H_R\Big)$. Így a félgömbnek van térfogata, és az egyenlő az alábbi kettős integrállal:

$$\iint\limits_{H_R} f = \iint\limits_{H_R} f(x,y)\,dx\,dy = \iint\limits_{H_R} \sqrt{R^2 - x^2 - y^2}\,dx\,dy.$$

Ezt az

$$x = r\cos\varphi, \quad y = r\sin\varphi$$
$$\Big(0 \le r \le R,\ 0 \le \varphi \le 2\pi\Big)$$

polárkoordinátás helyettesítéssel számoljuk ki. A megismert képlet alapján

$$\iint\limits_{H_R} f(x,y)\,dx\,dy = \iint\limits_{[0,R]\times[0,2\pi]} f\Big(r\,\cos\varphi,\, r\,\sin\varphi\Big)\cdot r\,dr\,d\varphi,$$

és ez utóbbi integrált szukcesszív integrálással könnyű kiszámítani:

$$\iint\limits_{[0,R]\times[0,2\pi]} f\Big(r\,\cos\varphi,\, r\,\sin\varphi\Big)\cdot r\,dr\,d\varphi = \iint\limits_{[0,R]\times[0,2\pi]} \sqrt{R^2 - r^2}\cdot r\,dr\,d\varphi =$$

$$\left(\int\limits_0^R \sqrt{R^2 - r^2}\cdot r\,dr\right)\cdot\left(\int\limits_0^{2\pi} 1\,d\varphi\right) = \left[-\frac{1}{3}(R^2 - r^2)^{3/2}\right]_0^R \cdot 2\pi = \frac{2R^3\pi}{3}.$$

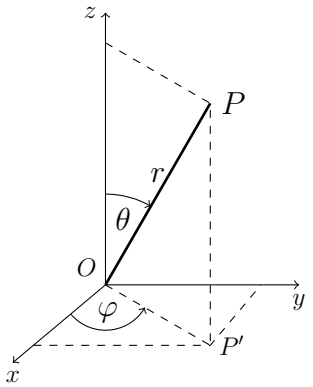
Az R sugarú gömb térfogata tehát $4R^3\pi/3$.

Megjegyzés. Emlékeztetünk arra, hogy az egyváltozós analízisben a gömb (forgástest) térfogatát a

$$\pi \cdot \int_{-R}^{R} \left(\sqrt{R^2 - x^2}\right)^2 \, dx$$

határozott integrállal számoltuk ki.

Térbeli polárkoordináta-transzformáció:



A síkbeli polárkoordináta-rendszert „térbeli” megfelelője az ábrán található jelölésekkel tudjuk megvalósítani. Ha $\overrightarrow{OP}(x, y, z)$ vektor hossza

$$r = \sqrt{x^2 + y^2 + z^2} > 0,$$

akkor a síkbeli polárkoordináták alapján

$$z = r \cos\theta$$

és $\overrightarrow{OP'}$ hossza $r \sin\theta$, ahol $0 \le \theta \le \pi$. Ha felírjuk a P' pont síkbeli polárkoordinátait, akkor

$$x = r \sin\theta \cos\varphi \qquad \text{és} \qquad y = r \sin\theta \sin\varphi,$$

ahol $0 \le \varphi < 2\pi$.

Térbeli polárkoordináta-transzformációról beszélünk, ha a

$$g(r, \varphi, \theta) := (r \sin\theta \cos\varphi,\ r \sin\theta \sin\varphi,\ r \cos\theta) \qquad \left((r, \varphi, \theta) \in \mathbb{R}^3\right)$$

leképezést alkalmazzuk. Világos, hogy $g \in C^1(\mathbb{R}^3)$, illetve

$$g'(r, \varphi, \theta) = \begin{pmatrix} \partial_r(r \sin\theta \cos\varphi) & \partial_\varphi(r \sin\theta \cos\varphi) & \partial_\theta(r \sin\theta \cos\varphi) \\ \partial_r(r \sin\theta \sin\varphi) & \partial_\varphi(r \sin\theta \sin\varphi) & \partial_\theta(r \sin\theta \sin\varphi) \\ \partial_r(r \cos\theta) & \partial_\varphi(r \cos\theta) & \partial_\theta(r \cos\theta) \end{pmatrix} =$$

$$= \begin{pmatrix} \sin\theta \cos\varphi & -r \sin\theta \sin\varphi & r \cos\theta \cos\varphi \\ \sin\theta \sin\varphi & r \sin\theta \cos\varphi & r \cos\theta \sin\varphi \\ \cos\theta & 0 & -r \sin\theta \end{pmatrix} \qquad \left((r, \varphi, \theta) \in \mathbb{R}^3\right).$$

Ezért

$$\det g'(r, \varphi, \theta) = \cos\theta \cdot \det \begin{pmatrix} -r \sin\theta \sin\varphi & r \cos\theta \cos\varphi \\ r \sin\theta \cos\varphi & r \cos\theta \sin\varphi \end{pmatrix} - r \sin\theta \cdot \det \begin{pmatrix} \sin\theta \cos\varphi & -r \sin\theta \sin\varphi \\ \sin\theta \sin\varphi & r \sin\theta \cos\varphi \end{pmatrix} =$$

$$= \cos\theta \cdot (r^2 \sin\theta \cos\theta) \cdot \underbrace{\det \begin{pmatrix} -\sin\varphi & \cos\varphi \\ \cos\varphi & \sin\varphi \end{pmatrix}}_{=-1} - r \sin\theta \cdot (r \sin^2\theta) \cdot \underbrace{\det \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}}_{=1} =$$

$$= -r^2 \sin\theta (\cos^2\theta + \sin^2\theta) = -r^2 \sin\theta$$

Adott $R > 0$, legyen

$$H \subseteq [0, R] \times [0, 2\pi] \times [0, \pi]$$

egy nem üres, Jordan-mérhető, zárt halmaz. Nem nehéz igazolni, hogy $g_{|\operatorname{int} H}$ invertálható.

Ekkor az integráltranszformációról szóló tétel feltételei teljesülnek, és így az

$$x = r\sin\theta\cos\varphi, \qquad y = r\sin\theta\sin\varphi, \qquad \text{és} \qquad z = r\cos\theta$$

transzformációval

$$(**) \qquad \iiint\limits_{g[H]} f(x,y,z)\,dx\,dy\,dz = \iiint\limits_{H} f(r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)\cdot r^2\sin\theta\,dr\,d\varphi\,d\theta,$$

hiszen $\det g'(r,\varphi,\theta) = -r^2\sin\theta \le 0$, ha $H \subseteq [0,R]\times[0,2\pi]\times[0,\pi]$.

Példa. Számítsuk ki az R sugarú gömb térfogatát!

Megoldás. A gömb térfogatát kiszámíthatjuk az alábbi hármas integrállal:

$$V = \iiint\limits_{G} 1\,dx\,dy\,dz, \qquad \text{ahol} \qquad G := \{x^2+y^2+z^2 \le R^2\}.$$

Térbeli polárkoordináta-transzformációval

$$H := [0,R]\times[0,2\pi]\times[0,\pi] \qquad \Longrightarrow \qquad G = g[H],$$

ezért $(**)$ alapján

$$V = \iiint\limits_{G} 1\,dx\,dy\,dz = \iiint\limits_{H} 1\cdot r^2\sin\theta\,dr\,d\varphi\,d\theta = \left(\int\limits_0^R r^2\,dr\right)\cdot\left(\int\limits_0^{2\pi} 1\,d\varphi\right)\cdot\left(\int\limits_0^{\pi}\sin\theta\,d\theta\right) =$$

$$= \left[\frac{r^3}{3}\right]_{r=0}^{r=R}\cdot\Big[\varphi\Big]_{\varphi=0}^{\varphi=2\pi}\cdot\Big[-cos\theta\Big]_{\theta=0}^{\theta=\pi} = \frac{R^3}{3}\cdot 2\pi\cdot 2 = \frac{4\pi R^3}{3}.$$

A normális eloszlás sűrűségfüggvénye

A valószínűségszámításban nagyon fontos szerepet játszanak azok az $f\colon\mathbb{R}\to\mathbb{R}$ függvények, amikre

$$f(x)\ge 0 \quad (x\in\mathbb{R}) \qquad \text{és} \qquad \int\limits_{-\infty}^{+\infty} f(x)\,dx = 1$$

teljesül, az ún. ***sűrűségfüggvények***. A folytonos eloszlások közül a normális eloszlás központi szerepet tölt be. Ennek sűrűségfüggvénye

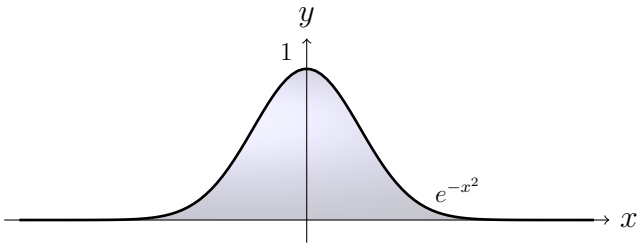
$$f(x) := \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-m)^2}{2\sigma^2}} \qquad (m\in\mathbb{R},\ \sigma>0,\ x\in\mathbb{R}).$$

De miért lesz f sűrűségfüggvény? Az $f(x)\ge 0$ tulajdonság nyilván teljesül, de a teljes $\mathbb{R}$ halmazon vett improprius integrálját ki kellene számítani. Azonban ezt nem tudjuk az improprius integrál fogalma alapján kiszámolni, mert ehhez ismerni kellene f primitív függvényeit, de ezek nem elemi függvények. Más szavakkal, nincs képlet, amivel tudnánk a Newton–Leibniz-formulát alkalmazni, és majd határértéket venni.

Érdekes módon, a többszörös integrálásnál megismert módszerek segítenek megoldani a feladatot. Először számítsuk ki az

$$\int\limits_{-\infty}^{+\infty} e^{-x^2}\,dx$$

ún. ***Gauss-integrált***. Ez úgy értelmezhető, mint az e^{-x^2} függvény alatti terület, amely az alábbi ábrán látható.



Az improprius integrál fogalma szerint

$$\int_{-\infty}^{+\infty} e^{-x^2}\,dx = \int_{-\infty}^{0} e^{-x^2}\,dx + \int_{0}^{+\infty} e^{-x^2}\,dx, \quad \text{ahol} \quad \int_{-\infty}^{0} e^{-x^2}\,dx = \int_{0}^{+\infty} e^{-x^2}\,dx = \lim_{t\to+\infty}\int_{0}^{t} e^{-x^2}\,dx,$$

hiszen az e^{-x^2} függvény páros. A fenti határérték létezik és véges, mert az

$$F(t) = \int_{0}^{t} e^{-x^2}\,dx \qquad (t > 0)$$

integrálfüggvény szigorúan monoton növekvő és korlátos. Valóban, F létezik, mert az e^{-x^2} függvény folytonos, tehát integrálható minden $[0,t]$ intervallumon. F szigorúan monoton növekvő, mert az e^{-x^2} függvény pozitív. F korlátos a $[0,1]$ intervallumon, mert folytonos, illetve minden $t > 1$ esetén

$$F(t) = \int_{0}^{1} e^{-x^2}\,dx + \int_{1}^{t} e^{-x^2}\,dx < \int_{0}^{1} 1\,dx + \int_{1}^{t} e^{-x}\,dx = 1 + \left[\frac{e^{-x}}{-1}\right]_1^t = 1 - \frac{1}{e^t} + \frac{1}{e} < 2.$$

A fentiekből következik, hogy a Gauss-integrál konvergens, azaz

$$\int_{-\infty}^{+\infty} e^{-x^2}\,dx = \int_{-\infty}^{0} e^{-x^2}\,dx + \int_{0}^{+\infty} e^{-x^2}\,dx = \lim_{t\to+\infty}\int_{-t}^{0} e^{-x^2}\,dx + \lim_{t\to+\infty}\int_{0}^{t} e^{-x^2}\,dx = \lim_{t\to+\infty}\int_{-t}^{t} e^{-x^2}\,dx$$

egy véges szám.

Most nézzük meg az

$$I_R := \iint_{H_R} e^{-x^2-y^2}\,dx\,dy$$

kettős integrált, ahol $H_R := \left\{(x,y) \in \mathbb{R}^2 \colon x^2 + y^2 \le R^2\right\}$ az origó középpontú R sugarú zárt körlap, és természetesen $R > 0$. Mivel $f \in C(\mathbb{R}^2)$, ezért a fenti integrál létezik és véges. Az integrál kiszámításához az

$$x = r\cos\varphi, \quad y = r\sin\varphi$$
$$\big(0 \le r \le R,\ 0 \le \varphi \le 2\pi\big)$$

polárkoordinátás helyettesítést alkalmazzuk. Azt kapjuk, hogy

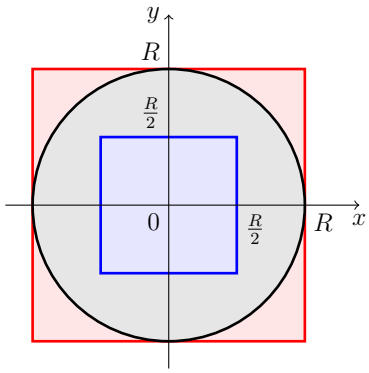
$$I_R = \iint_{[0,R]\times[0,2\pi]} e^{-r^2}\cdot r\,dr\,d\varphi = \left(\int_0^R e^{-r^2}\cdot r\,dr\right)\cdot\left(\int_0^{2\pi} 1\,d\varphi\right) = \left(-\frac{1}{2}\int_0^R e^{-r^2}\cdot(-2r)\,dr\right)\cdot 2\pi =$$

$$= -\pi\cdot\left[e^{-r^2}\right]_{r=0}^{r=R} = -\pi\left(e^{-R^2} - e^0\right) = \pi\cdot\left(1 - e^{-R^2}\right).$$

A következő lépésben tekintsük a

$$T_R := [-R, R] \times [-R, R] \qquad (R > 0)$$

téglalapokat. Nem nehéz igazolni, hogy $T_{R/2} \subset H_R \subset T_R$ (lásd a jobb oldali ábrát). Ezért



$$\iint\limits_{T_{R/2}} e^{-x^2-y^2}\,dx\,dy \le \iint\limits_{H_R} e^{-x^2-y^2}\,dx\,dy \le \iint\limits_{T_R} e^{-x^2-y^2}\,dx\,dy.$$

Azonban minden $R > 0$ esetén

$$\iint\limits_{T_R} e^{-x^2-y^2}\,dx\,dy = \iint\limits_{T_R} e^{-x^2}\cdot e^{-y^2}\,dx\,dy = \left(\int\limits_{-R}^{R} e^{-x^2}\,dx\right)\cdot\left(\int\limits_{-R}^{R} e^{-y^2}\,dy\right) = \left(\int\limits_{-R}^{R} e^{-x^2}\,dx\right)^2.$$

Ennek következtében

$$\left(\int\limits_{-R/2}^{R/2} e^{-x^2}\,dx\right)^2 \le \pi\cdot\left(1-e^{-R^2}\right) \le \left(\int\limits_{-R}^{R} e^{-x^2}\,dx\right)^2.$$

adódik minden $R > 0$-ra. Tudjuk, hogy a Gauss-integrál konvergens. Ezért a fenti egyenlőtlenségben R-rel plusz végtelenhez tartva azt kapjuk, hogy

$$\left(\int\limits_{-\infty}^{+\infty} e^{-x^2}\,dx\right)^2 \le \pi \le \left(\int\limits_{-\infty}^{+\infty} e^{-x^2}\,dx\right)^2,$$

tehát

$$\int\limits_{-\infty}^{+\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$

Most térünk vissza az eredeti feladathoz, azaz a normális eloszlás sűrűségfüggvényéhez:

$$f(x) := \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}} \qquad (m \in \mathbb{R},\ \sigma > 0,\ x \in \mathbb{R}).$$

Minden $t > m$ esetén a következő helyettesítéssel:

$$u = \frac{x-m}{\sqrt{2}\sigma} \qquad \Longrightarrow \qquad x = \sqrt{2}\sigma u + m := g(u) \quad (u > 0) \qquad \Longrightarrow \qquad g'(u) = \sqrt{2}\sigma > 0,$$

azt kapjuk, hogy

$$\int\limits_m^t \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}\,dx = \frac{1}{\sqrt{2\pi}\sigma}\int\limits_m^t e^{-\left(\frac{x-m}{\sqrt{2}\sigma}\right)^2}\,dx = \frac{1}{\sqrt{2\pi}\sigma}\int\limits_0^{\frac{t-m}{\sqrt{2}\sigma}} e^{-u^2}\cdot\sqrt{2}\sigma\,du = \frac{1}{\sqrt{\pi}}\int\limits_0^{\frac{t-m}{\sqrt{2}\sigma}} e^{-u^2}\,du.$$

Így ha t tart a plusz végtelenhez, akkor

$$\int\limits_m^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}\,du = \frac{1}{\sqrt{\pi}}\int\limits_0^{+\infty} e^{-u^2}\,dx.$$

Hasonlóan igazolható, hogy

$$\int_{-\infty}^{m} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}\, du = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{0} e^{-u^2}\, dx.$$

Ennek következtében

$$\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}\, dx = \int_{-\infty}^{m} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}\, dx + \int_{m}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-m)^2}{2\sigma^2}}\, dx =$$

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{0} e^{-u^2}\, du + \frac{1}{\sqrt{\pi}} \int_{0}^{+\infty} e^{-u^2}\, du = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{+\infty} e^{-u^2}\, du = 1.$$