Analízis 3A

1. Fejezet

1. Emlékeztető.

Az $f \in \mathbf{K} \to \mathbf{K}$ függvények[1] (speciálisan a valós vagy komplex számsorozatok) vizsgálatakor alapvető szerepet játszott az, hogy „mérni" tudtuk két szám „távolságát". Nevezetesen, ha $x, y \in \mathbf{K}$, akkor az x és az y *távolságán* az

$$|x-y|$$

számot értettük.

Példaként idézzük fel a fenti f függvény valamely $a \in \mathcal{D}_f$ pontbeli folytonosságának a definícióját: bármely ε pozitív számhoz van olyan δ pozitív szám, hogy

$$x \in \mathcal{D}_f,\ |x-a| < \delta \Longrightarrow |f(x)-f(a)| < \varepsilon.$$

Ez utóbbi látszólag függ a valós vagy komplex számok algebrai struktúrájától, hogy ti. $x, y \in \mathbf{K}$ esetén létezik az $x-y \in \mathbf{K}$ szám, ill., hogy bármely $z \in \mathbf{K}$ számnak értelmeztük a $|z|$ módon jelölt abszolútértékét (amit a valós számok körében bevezetett rendezés segítségével definiáltunk).

Ugyanakkor mindezen fogalmak nélkül is könnyedén elmondhatjuk az előbbi definíciót, nevezetesen: *az f függvény folytonos az $a \in \mathcal{D}_f$ helyen, ha minden ε pozitív szám esetén létezik olyan δ pozitív szám, amellyel az $f(x)$ és az $f(a)$ távolsága kisebb, mint ε, hacsak az $x \in \mathcal{D}_f$ és az a távolsága kisebb, mint δ.*

Tehát ebből a szempontból igazából csak az alábbi (a „távolságot" illető meghatározó) tulajdonságok játszanak szerepet: bevezetve a

$$\rho(u,v) := |u-v| \qquad (u, v \in \mathbf{K})$$

jelölést, tetszőleges $x, y, z \in \mathbf{K}$ számokra

1^o $\rho(x,y) \geq 0$;

2^o $\rho(x,y) = 0$ akkor és csak akkor teljesül, ha $x = y$;

3^o $\rho(x,y) = \rho(y,x)$;

4^o $\rho(x,y) \leq \rho(x,z) + \rho(z,y)$.

---

[1] A $\mathbf{K}$ szimbólum a továbbiakban a valós számok ($\mathbf{R}$), vagy a komlex számok ($\mathbf{C}$) halmazát jelöli.

(A 4^o ún. *háromszög-egyenlőtlenség* alkalmazása számos bizonyítás kulcsmozzanata volt.) A $\mathbf{K}$-beli számok távolságának az értelmezéséhez tehát tulajdonképpen nincs szükség sem a test-struktúrára, sem (az $\mathbf{R}$-beli) rendezésre, hanem pl. az előbbi $1^o - 4^o$ összefüggéseknek eleget tevő

$$\rho : \mathbf{K} \times \mathbf{K} \to \mathbf{R}$$

függvény segítségével a

$$\rho(x, y) := \rho((x, y)) \qquad (x, y \in \mathbf{K})$$

helyettesítési értéket (nemnegatív) számot nevezhetjük az $x, y \in \mathbf{K}$ számok *távolságának.*

A „távolság" említése nem pusztán formai trükk. Gondoljunk pl. az $\mathbf{R}^2$-ben már a bevezető tanulmányok során megismert euklideszi távolságra, amikor is az $(x, y), (u, v) \in \mathbf{R}^2$ vektorok távolságát a

$$\sqrt{(x-u)^2 + (y-v)^2}$$

nemnegatív számmal mértük. Az itt is megjelenő algebrai struktúra az, ami csupán formai. Az elemi geometriában hamar kidomborodik ui. a tartalmi lényeg, nevezetesen:

i) két vektor távolsága nemnegatív szám;

ii) két vektor távolsága akkor és csak akkor nulla, ha a két vektor azonos;

iii) az egyik vektor távolsága a másiktól ugyanaz, mint a másiknak az előbbitől;

iv) két vektor távolsága nem lehet nagyobb egy harmadik vektortól mért távolságaik összegénél.

Nem nehéz belátni, hogy ezek (az egymástól ugyan nem teljesen független) tulajdonságok azok, amelyek (az előbbiekben a számokkal kapcsolatban megfogalmazottakkal analóg módon) jellemzik az $\mathbf{R}^2$-beli elemek távolságát. Hasonló mondható el akkor is, ha $\mathbf{R}^2$ helyett $\mathbf{R}^3$-at, ill. i) – iv) -ben kétdimenziós vektorok helyett háromdimenziós vektorokat veszünk.

Könnyű meggyőződni arról is, hogy $\mathbf{K}$-ban, ill. $\mathbf{R}^2$-ben, vagy $\mathbf{R}^3$-ban nem egyedül a most említett módon mérhetjük az elemek távolságát. Pl., ha $x, y \in \mathbf{K}$ esetén x és y távolságának az

$$\frac{|x-y|}{1+|x-y|}$$

nemnegatív számot választjuk, akkor az előbbi $1^o - 4^o$ tulajdonságok triviálisan teljesülnek. Hasonlóan, az $(x,y),(u,v) \in \mathbf{R}^2$ vektorok távolságaként vehetnénk pl. a

$$\max\{|x-u|,|y-v|\},$$

vagy az

$$|x-u|+|y-v|$$

számot.

2. Metrikus terek.

Konkrét példák sokasága vezet el a távolság-fogalom absztrakciójához: legyen az $X \neq \emptyset$ egy nem üres halmaz, és tegyük fel, hogy a

$$\rho : X^2 \to [0,+\infty)$$

függvény a következő tulajdonságokkal rendelkezik:[2]

a) minden $x \in X$ esetén $\rho(x,x) = 0$;

b) ha $x,y \in X$ és $\rho(x,y) = 0$, akkor $x = y$;

c) bármely $x,y \in X$ választással $\rho(x,y) = \rho(y,x)$;

d) tetszőleges $x,y,z \in X$ elemekkel $\rho(x,y) \leq \rho(x,z) + \rho(y,z)$.

Azt mondjuk, hogy ekkor a ρ egy *távolságfüggvény* (vagy idegen szóval *metrika*). Ha $x,y \in X$, akkor $\rho(x,y)$ az x,y elemek *távolsága.* Az (X,ρ) rendezett párt *metrikus térnek* nevezzük.

Az X-beli elemek távolsága tehát egy nemnegatív szám. Bármely elem önmagától vett távolsága *nulla* (ld. a)), továbbá két különböző elem távolsága mindig *pozitív* (ld. b)). A távolság *szimmetrikus*, azaz két elem távolsága független az illető elemek sorrendjétől (ld. c)). A d) tulajdonságot *háromszög-egyenlőtlenségként* fogjuk idézni.

Mutassuk meg, hogy a háromszög-egyenlőtlenségből annak az alábbi változata is következik:

$$|\rho(x,z) - \rho(y,z)| \leq \rho(x,y) \qquad (x,y,z \in X).$$

Ugyanis a d) axióma miatt

$$\rho(x,z) \leq \rho(x,y) + \rho(y,z),$$

[2]Valamilyen $(x,y) \in X^2$ esetén a ρ függvény (x,y)-beli helyettesítési értékére a „szabványos" $\rho((x,y))$ szimbólum helyett továbbra is az egyszerűbb $\rho(x,y)$-t használjuk.

tehát
$$\rho(x,z)-\rho(y,z)\le\rho(x,y).$$
Ha itt az x-et és az y-t felcseréljük, akkor a
$$-(\rho(x,z)-\rho(y,z))=\rho(y,z)-\rho(x,z)\le\rho(y,x)=\rho(x,y)$$
egyenlőtlenséghez jutunk. Az utóbbi két becslés egybevetésével kapjuk a jelzett egyenlőtlenséget.

Bármely $X\neq\emptyset$ halmaz esetén megadható
$$\rho:X^2\to[0,+\infty)$$
távolságfüggvény, ui. pl. a
$$\rho(x,y):=\begin{cases}0 & (x=y)\\ \\ 1 & (x\neq y)\end{cases}\qquad \Big((x,y)\in X^2\Big)$$
leképezés nyilván eleget tesz a fenti, a metrikát meghatározó a) – d) axiómáknak. Az így definiált (X,ρ) teret a *diszkrét* jelzővel illetjük.

Megmutatható, hogy az a) – d) axiómák nem függetlenek egymástól, nevezetesen: ha egy
$$\rho:X^2\to\mathbf{R}$$
függvény rendelkezik az a), b), d) tulajdonságokkal, akkor a ρ metrika.[3]

3. Példák.

Soroljunk fel néhány példát, amelyek nem csupán az analízisben játszanak fontos szerepet.

1^o Legyen $1\le n\in\mathbf{N}$, $0<p<+\infty$, és
$$x=(x_1,...,x_n),\ y=(y_1,...,y_n)\in\mathbf{K}^n$$
esetén definiáljuk az (x,y)-ban a
$$\rho_p:\mathbf{K}^n\times\mathbf{K}^n\to[0,+\infty)$$

[3] Ezzel együtt „érdemes" a c)-t is felsorolni az axiómák között, megkönnyítve a rá való hivatkozást a számtalan alkalmazás során.

függvény helyettesítési értékét a következőképpen:

$$\rho_p(x,y) := \begin{cases} \sum\limits_{i=1}^{n} |x_i - y_i|^p & (p \le 1) \\ \Big(\sum\limits_{i=1}^{n} |x_i - y_i|^p\Big)^{1/p} & (p > 1). \end{cases}$$

Terjesszük ki a ρ_p értelmezését $p = \infty := +\infty$-re is az alábbiak szerint:

$$\rho_\infty(x,y) := \max\{|x_i - y_i| : i = 1, ..., n\}.$$

Belátható, hogy $(\mathbf{K}^n, \rho_p)$ metrikus tér.[4] A későbbiekben a ρ_∞ metrika mellett a $\mathbf{K}^n$-beli vektorok távolságának a mérésére többnyire a

$$\rho_2(x,y) = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} \qquad (x, y \in \mathbf{K}^n),$$

$$\rho_1(x,y) = \sum_{i=1}^{n} |x_i - y_i| \qquad (x, y \in \mathbf{K}^n)$$

metrikákat fogjuk használni. Speciálisan az $n = 1$ esetben

$$\rho_p(x,y) = |x - y| \qquad (x, y \in \mathbf{K}, p \ge 1).$$

2^o Tekintsük egy $0 < p < +\infty$ mellett az

$$\ell_p := \Big\{(x_n) : \mathbf{N} \to \mathbf{K} : \sum_{n=0}^{\infty} |x_n|^p < +\infty\Big\}$$

halmazokat. Legyen továbbá $x = (x_n)$, $y = (y_n) \in \ell_p$ esetén

$$\rho_p(x,y) := \begin{cases} \sum\limits_{n=0}^{\infty} |x_n - y_n|^p & (0 < p \le 1) \\ \Big(\sum\limits_{n=0}^{\infty} |x_n - y_n|^p\Big)^{1/p} & (p > 1). \end{cases}$$

Bebizonyítható, hogy az így definiált ρ_p függvény is metrika, azaz (ℓ_p, ρ_p) metrikus tér. A $p = \infty := +\infty$-re való „kiterjesztést" a következőképpen kapjuk:

$$\ell_\infty := \Big\{(x_n) : \mathbf{N} \to \mathbf{K} : \sup\{|x_n| : n \in \mathbf{N}\} < +\infty\Big\}$$

[4] Ennek a bizonyítását nem részletezzük, mindez $p = 1$-re és $p = \infty$-re meglehetősen triviális.

(más szóval az ℓ_∞ szimbólum a korlátos számsorozatok halmazát jelöli), valamint az ℓ_∞-beli $x = (x_n)$, $y = (y_n)$ elemekre (sorozatokra)

$$\rho_\infty(x, y) := \sup\{|x_n - y_n| : n \in \mathbf{N}\}.$$

A ρ_∞ függvény is metrika, tehát $(\ell_\infty, \rho_\infty)$ is metrikus tér.

3^o Valamilyen $[a, b]$ korlátos és zárt intervallum ($a, b \in \mathbf{R}$, $a < b$) esetén jelöljük $C[a, b]$-vel az $[a, b]$-n értelmezett, valós értékű és folytonos függvények halmazát. Ha $0 < p \leq +\infty$, akkor tekintsük az 1^o, 2^o példák alábbi „folytonos" változatait: ha $f, g \in C[a, b]$), akkor[5]

$$\rho_p(f, g) := \begin{cases} \displaystyle\int_a^b |f - g|^p & (0 < p \leq 1) \\ \left(\displaystyle\int_a^b |f - g|^p\right)^{1/p} & (1 < p < +\infty) \\ \max\{|f(x) - g(x)| : x \in [a, b]\} & (p = \infty := +\infty). \end{cases}$$

Az előbbi példákhoz hasonlóan látható be, hogy $(C[a, b], \rho_p)$ is metrikus tér.

Azt mondjuk, hogy valamilyen $X \neq \emptyset$ halmaz és az X^2-en értelmezett

$$\rho, \sigma : X^2 \to [0, +\infty)$$

metrikák esetén a ρ és a σ *ekvivalens*, ha alkalmas c, C pozitív számokkal

$$c\cdot \rho(x, y) \leq \sigma(x, y) \leq C\cdot \rho(x, y) \qquad (x, y \in X).$$

Könnyű belátni, hogy ha $\mathcal{M}$ jelöli az előbb említett metrikák halmazát, és a ρ, $\sigma \in \mathcal{M}$ elemekre $\rho \sim \sigma$ azt jelenti, hogy a ρ és a σ ekvivalens, akkor az így értelmezett ($\mathcal{M}^2$-beli) $\sim$ reláció ekvivalencia.

Pl. a fenti $(\mathbf{K}^n, \rho_p)$ metrikus terekre a ρ_p metrikák közül $p \geq 1$ esetén bármelyik kettő ekvivalens. A továbbiakban az

$$X := \mathbf{K}^n \qquad (1 \leq n \in \mathbf{N})$$

esetben a $\rho_2, \rho_1, \rho_\infty$ metrikák valamelyikét (többnyire a ρ_2 ún. *euklideszi metrikát*) fogjuk használni.

[5]Emlékeztetünk arra, hogy $|f - g|^p$ bármely $0 < p \in \mathbf{R}$ mellett folytonos függvény, így Riemann-szerint integrálható. Továbbá Weierstrass nevezetes tétele alapján az $|f - g|$ folytonos függvénynek van maximuma.

4. Normált terek.

Tegyük fel, hogy a szóban forgó $X \neq \emptyset$ halmaz *lineáris tér* a $\mathbf{K}$ felett.[6] Azt mondjuk, hogy a

$$\varphi : X \to [0, +\infty)$$

leképezés *norma*, ha

a) $\varphi(0) = 0$;

b) ha $x \in X$ és $\varphi(x) = 0$, akkor $x = 0$;

c) bármely $\lambda \in \mathbf{K}$, $x \in X$ esetén $\varphi(\lambda x) = |\lambda| \cdot \varphi(x)$;

d) tetszőleges $x, y \in X$ elemekre $\varphi(x + y) \leq \varphi(x) + \varphi(y)$.

Egy $x \in X$ elemre az

$$\|x\| := \varphi(x)$$

nemnegatív számot az x *hosszának* (vagy *normájának*), az $(X, \|.\|)$ rendezett párt pedig *normált térnek* nevezzük. Következésképpen a fenti a) – d) axiómák alakja ezzel a jelöléssel:

a) $\|0\| = 0$;

b) ha $x \in X$ és $\|x\| = 0$, akkor $x = 0$;

c) bármely $\lambda \in \mathbf{K}$, $x \in X$ esetén $\|\lambda x\| = |\lambda| \cdot \|x\|$;

d) tetszőleges $x, y \in X$ elemekre $\|x + y\| \leq \|x\| + \|y\|$.

A d) axiómát szintén *háromszög-egyenlőtlenségként* említjük a későbbiekben.

Ha pl. X jelöli a

$$\mathbf{K}^n \ (0 < n \in \mathbf{N}), \quad \ell_p \quad (0 < p \in \mathbf{R}), \quad C[a, b] \ (-\infty < a < b < +\infty)$$

halmazok bármelyikét, akkor a vektorok (sorozatok, függvények) szokásos összeadására és számmal való szorzására nézve az X lineáris tér a $\mathbf{K}$, ill. az $\mathbf{R}$ felett. Az említett terekben a nulla-elemet 0-val jelölve azt kapjuk továbbá, hogy $1 \leq p \leq +\infty$ esetén

$$\|x\|_p := \rho_p(x, 0) \quad (x \in X)$$

[6] Nem fog félreértést okozni, ha (a számokhoz hasonlóan) az $x, y \in X$ elemek X-beli „összegét" $x + y$-nal, valamilyen $\lambda \in \mathbf{R}$ esetén az x elem (X-beli) λ-szorosát λx-szel, az x inverzét pedig $-x$-szel, továbbá az X-beli nulla-elemet 0-val jelöljük.

norma, azaz ilyen p-kre

$$(\mathbf{K}^n, \|.\|_p),\ (\ell_p, \|.\|_p),\ (C[a,b], \|.\|_p)$$

normált terek. Tehát

$$\|x\|_p = \begin{cases} \Big(\sum\limits_{i=1}^{n} |x_i|^p\Big)^{1/p} & (1 \le p < +\infty) \\ \max\{|x_i| : i = 1, ..., n\} & (p = +\infty) \end{cases} \qquad \Big(x = (x_1, ..., x_n) \in \mathbf{K}^n\Big),$$

speciálisan az $n = 1$ esetben

$$\|x\|_p = |x| \qquad (x \in \mathbf{K},\ 1 \le p \le +\infty),$$

valamint

$$\|y\|_p = \begin{cases} \Big(\sum\limits_{i=1}^{\infty} |y_i|^p\Big)^{1/p} & (1 \le p < +\infty) \\ \sup\{|y_i| : i \in \mathbf{N}\} & (p = +\infty) \end{cases} \qquad \Big(y = (y_n) \in \ell_p\Big)$$

és

$$\|f\|_p = \begin{cases} \Big(\int_a^b |f|^p\Big)^{1/p} & (1 \le p < +\infty) \\ \max\{|f(x)| : x \in [a,b]\} & (p = +\infty) \end{cases} \qquad \Big(f \in C[a,b]\Big).$$

Világos, hogy a most mondott példákban

$$\rho_p(x,y) = \|x - y\|_p \qquad (x, y \in X).$$

Sőt, ha most $(X, \|.\|)$ egy tetszőleges normált teret jelöl, akkor a

$$\rho(x,y) := \|x - y\| \qquad \Big((x,y) \in X^2\Big)$$

függvény metrika, azaz (X, ρ) metrikus tér:

$$(X, \rho) \equiv (X, \|.\|).$$

Ekkor pl. a

$$|\rho(x,z) - \rho(y,z)| \le \rho(x,y) \qquad (x, y, z \in X)$$

háromszög-egyenlőtlenség az alábbi alakot ölti:

$$\Big|\|x - z\| - \|y - z\|\Big| \le \|x - y\| \qquad (x, y, z \in X).$$

Ha itt $z = 0$, akkor

$$\Big| \|x\| - \|y\| \Big| \leq \|x - y\| \qquad (x, y \in X).$$

Azt mondjuk, hogy az X ($\mathbf{K}$-feletti) vektortéren értelmezett

$$\|.\|,\ \|.\|_* : X \to [0, +\infty)$$

normák *ekvivalensek* (erre is a $\|.\| \sim \|.\|_*$ jelölést fogjuk használni), ha alkalmas c, C pozitív konstansokkal

$$c \cdot \|x\| \leq \|x\|_* \leq C \cdot \|x\| \qquad (x \in X).$$

Világos, hogy mindez azzal ekvivalens, hogy a

$$\rho(x, y) := \|x - y\|,\ \ \rho_*(x, y) := \|x - y\|_* \qquad (x, y \in X)$$

metrikák ekvivalensek.

Az

$$(X, \|.\|) := (\mathbf{K}^n, \|.\|_p) \qquad (n \in \mathbf{N}, 1 \leq p \leq +\infty)$$

terek esetén

$$\|.\|_p \sim \|.\|_q \qquad (1 \leq p, q \leq +\infty).$$

Speciálisan[7]

$$\|x\|_\infty \leq \|x\|_p \leq n^{1/p} \cdot \|x\|_\infty \qquad (x, y \in \mathbf{K}^n,\ 1 \leq p < +\infty),$$

ill.

$$n^{-1/p} \cdot \|x\|_p \leq \|x\|_q \leq n^{1/q} \cdot \|x\|_p \qquad (x, y \in \mathbf{K}^n,\ 1 \leq p,\ q < +\infty).$$

5. Euklideszi terek.

A fent bevezetett $\|.\|_p$ norma a $p = 2$ esetben speciális esete egy tágabb (a lineáris algebrából is jól ismert) normaosztálynak. Legyen ui. X újra egy lineáris tér a $\mathbf{K}$ felett, az

$$s : X^2 \to \mathbf{K}$$

függvényről pedig tegyük fel, hogy

- minden $x, y \in X$ mellett $s(x, y) = \overline{s(y, x)}$ (ahol a $\overline{\xi}$ szimbólum a $\xi \in \mathbf{K}$ szám komplex konjugáltját jelöli);

[7] $1/\infty := 0$.

- bármely $x \in X \setminus \{0\}$ esetén $s(x,x) \in \mathbf{R}$ és $s(x,x) > 0$;
- ha $x, y \in X$ és $\lambda \in \mathbf{K}$, akkor $s(\lambda x, y) = \lambda s(x,y)$;
- tetszőleges $x, y, z \in X$ elemekre fennáll a következő egyenlőség:

$$s(x+y,z) = s(x,z) + s(y,z).$$

Ha $x, y \in X$, akkor az

$$\langle x, y \rangle := s(x,y)$$

számot az x, y elemek *skaláris szorzatának*, az $(X, \langle . \rangle)$ rendezett párt pedig *skaláris szorzat-térnek* (vagy *euklideszi térnek*) nevezzük. A skaláris szorzatra most bevezetett jelöléssel tehát a fenti axiómák a következő alakúak:

- minden $x, y \in X$ mellett $\langle x, y \rangle = \overline{\langle y, x \rangle}$;
- bármely $x \in X \setminus \{0\}$ esetén $\langle x, x \rangle > 0$;
- ha $x, y \in X$ és $\lambda \in \mathbf{K}$, akkor $\langle \lambda x, y \rangle = \lambda \cdot \langle x, y \rangle$;
- tetszőleges $x, y, z \in X$ elemekre fennáll a következő egyenlőség:

$$\langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle.$$

Speciálisan minden $x \in X$ esetén

$$\langle 0, x \rangle = \langle x, 0 \rangle = 0,$$

ill.

$$\langle x, x \rangle = 0 \Longrightarrow x = 0.$$

Tehát

$$\langle x, x \rangle = 0 \iff x = 0.$$

Ha $\mathbf{K} = \mathbf{R}$ (azaz $(X, \langle . \rangle)$ egy ún. *valós euklideszi tér*), akkor

$$\langle x, y \rangle = \langle y, x \rangle \qquad (x, y \in X).$$

Jelentse pl. X a

$$\mathbf{K}^n \;\; (1 \leq n \in \mathbf{N}), \;\; \ell_2, \;\; C[a,b] \;\; (-\infty < a < b < +\infty)$$

halmazok valamelyikét, és

$$\langle x, y\rangle := \begin{cases} \sum_{i=1}^{n} x_i\overline{y}_i & \Big(x = (x_1, ..., x_n),\ y = (y_1, ..., y_n) \in \mathbf{K}^n\Big) \\ \sum_{n=0}^{\infty} x_n\overline{y}_n & \Big(x = (x_n), y = (y_n) \in \ell_2\Big) \\ \int_a^b xy & \Big(x, y \in C[a, b]\Big). \end{cases}$$

Ekkor $(X, \langle . \rangle)$ euklideszi tér, továbbá

$$\|x\|_2 = \sqrt{\langle x, x\rangle} \qquad (x \in X).$$

Ez utóbbi egyenlőségnek sokkal általánosabb háttere van, ui. tetszőleges $(X, \langle . \rangle)$ euklideszi teret véve

$$\|x\| := \sqrt{\langle x, x\rangle} \qquad (x \in X)$$

norma. Itt a háromszög-egyenlőtlenség igazolásában fontos szerep jut az

$$|\langle x, y\rangle| \leq \|x\| \cdot \|y\| \quad (x, y \in X)$$

Cauchy-Bunyakovszkij-egyenlőtlenségnek. Ezt „lefordítva" az előbb említett euklideszi terekre az alábbi egyenlőtlenségeket kapjuk:

$$\left|\sum_{i=1}^{n} x_i\overline{y_i}\right| \leq \sqrt{\sum_{i=1}^{n} |x_i|^2} \cdot \sqrt{\sum_{i=1}^{n} |y_i|^2} \qquad (x, y \in \mathbf{K}^n),$$

$$\left|\sum_{i=1}^{\infty} x_i\overline{y_i}\right| \leq \sqrt{\sum_{i=1}^{\infty} |x_i|^2} \cdot \sqrt{\sum_{i=1}^{\infty} |y_i|^2} \qquad (x, y \in \ell_2),$$

$$\left|\int_a^b fg\right| \leq \sqrt{\int_a^b f^2} \cdot \sqrt{\int_a^b g^2} \qquad (f, g \in C[a, b]).$$

Az $n = 1$ esetben a $\mathbf{K}^n = \mathbf{K}$-ban az előbb értelmezett skaláris szorzás a következő:

$$\langle x, y\rangle = x\overline{y} \qquad (x, y \in \mathbf{K}),$$

ill. ekkor

$$\|x\| = \|x\|_2 = \sqrt{\langle x, x\rangle} = \sqrt{|x|^2} = |x| \qquad (x \in \mathbf{K}).$$

Nem nehéz belátni, hogy a fenti

$$(\mathbf{K}^n, \|.\|_p) \qquad (1 \leq n \in \mathbf{N}, 1 \leq p \leq +\infty)$$

normált terek közül $(\mathbf{K}^n, \|.\|_2)$ az egyetlen, amelyre a $\|.\|_p$ normát skaláris szorzás „generálja". Másképp fogalmazva az a tény, hogy egy alkalmas $\langle . \rangle$ skaláris szorzással

$$\|x\|_p = \sqrt{\langle x, x\rangle} \qquad (x \in \mathbf{K}^n),$$

azzal ekvivalens, hogy $p = 2$. Ha ui. $p = 2$, akkor fentebb láttuk, hogy a $\|.\|_2$ norma skaláris szorzásból származik. Fordítva pedig mindez az ún. *paralelogramma-szabály*[8] következménye: tetszőleges $(X, \langle . \rangle)$ euklideszi tér esetén az

$$\|x\| := \sqrt{\langle x, x\rangle} \qquad (x \in X)$$

normára

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \qquad (x, y \in X).$$

6. Metrikus terek topológiája.

Legyen az (X, ρ) metrikus tér esetén $b \in X$ és $r > 0$, ekkor a

$$K_r(b) := \{x \in X : \rho(x, b) < r\}$$

halmazt a b elem *r-sugarú környezetének* nevezzük. Használni fogjuk a $K(b)$ jelölést is a $K_r(b)$ helyett, ha az adott szituációban a $K_r(b)$ *környezet sugara* (r) nem játszik szerepet.

Tekintsük a

$$(\mathbf{K}^n, \rho_p) \qquad (1 \leq n \in \mathbf{N},\ p = 1, 2, \infty)$$

metrikus tereket. Ekkor $b = (b_1, ..., b_n) \in \mathbf{K}^n$ és $r > 0$ esetén ezekben a terekben a b vektor r-sugarú $K_r(b)$ környezetei (attól függően, hogy $\rho = \rho_1, \rho_2, \rho_\infty$) rendre a következők:

$$K_r^{(1)}(b) := \left\{x = (x_1, ..., x_n) \in \mathbf{K}^n : \sum_{j=1}^{n} |x_j - b_j| < r\right\},$$

$$K_r^{(2)}(b) := \left\{x = (x_1, ..., x_n) \in \mathbf{K}^n : \sqrt{\sum_{j=1}^{n} |x_j - b_j|^2} < r\right\},$$

[8] Neumann–Jordan-tétel: legyen $(X, \|.\|)$ normált tér a $\mathbf{K}$ testre vonatkozóan és tegyük fel, hogy $\|x + y\|^2 + \|x - y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right)$ $(x, y \in X)$. Ekkor megadható olyan $X^2 \ni (x, y) \mapsto \langle x, y\rangle \in \mathbf{K}$ skaláris szorzás, amelyre $\|x\| = \sqrt{\langle x, x\rangle}$ $(x \in X)$. (Neumann János (1903 – 1957) – Ernst Pascual Jordan (1902 – 1980)).

$$K_r^{(\infty)}(b) := \left\{ x = (x_1, ..., x_n) \in \mathbf{K}^n : \max\{|x_j - b_j| : j = 1, ..., n\} < r \right\}.$$

Speciálisan a $\mathbf{K}^n := \mathbf{R}^2$ választással a $b = (b_1, b_2) \in \mathbf{R}^2$ vektor előbbi környezetei geometriailag (az $\mathbf{R}^2$ „síkot" egy derékszögű koordinátarendszerrel reprezentálva) könnyen ellenőrizhetően a következők:

- $K_r^{(1)}(b)$ egy, a

$$(b_1 - r, b_2),\ (b_1, b_2 + r),\ (b_1 + r, b_2),\ (b_1, b_2 - r)$$

  pontok (mint csúcspontok) által meghatározott *rombusz* (csúcsára állított négyzet) belseje,

- $K_r^{(2)}(b)$ egy b középpontú és r sugarú *körlemez* belseje,

- $K_r^{(\infty)}(b)$ pedig egy, a

$$(b_1 - r, b_2 - r),\ (b_1 - r, b_2 + r),\ (b_1 + r, b_2 + r),\ (b_1 + r, b_2 - r)$$

  pontok (mint csúcspontok) által meghatározott *négyzet* belseje.

7. Megjegyzések.

i) Valamely

$$f : [0, +\infty) \to [0, +\infty)$$

függvényről tegyük fel, hogy

α) $f(x) \leq f(y) \quad (x, y \in [0, +\infty),\ x \leq y)$;

β) $f(x) = 0 \iff x = 0 \quad \big(x \in [0, +\infty)\big)$;

γ) $f(x + y) \leq f(x) + f(y) \quad \big(x, y \in [0, +\infty)\big)$.

Ekkor bármelyik (X, ρ) metrikus tér esetén az

$$X^2 \ni (x, y) \mapsto \sigma(x, y) := f(\rho(x, y))$$

leképezés metrika. Pl. tetszőleges $0 < a \in \mathbf{R}$, $r \in (0, 1]$ esetén az

- $f(x) := ax \quad (x \geq 0)$;

- $g(x) := \dfrac{x}{1+x} \quad (x \geq 0)$;
- $h(x) := x^r \quad (x \geq 0)$;
- $s(x) := \lg(1+x) \quad (x \geq 0)$;
- $v(x) := \min\{1, x\} \quad (x \geq 0)$

függvények ilyenek.

Következésképpen, ha (X, ρ) metrikus tér, akkor az

$$a\rho \quad (a > 0), \quad \frac{\rho}{1+\rho}, \quad \sqrt{\rho}, \quad \rho^r,$$

ill. az

$$X^2 \ni (x,y) \mapsto \lg\Big(1 + \rho(x,y)\Big), \; X^2 \ni (x,y) \mapsto \min\{1, \rho(x,y)\}$$

függvények mindegyike metrika.

ii) Vegyük a $\mathbf{K}^n$ $(0 < n \in \mathbf{N})$ vektortéren $1 \leq p \leq +\infty$ esetén az $\|.\|_p$ normát, azaz, amikor

$$\|x\|_p = \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p} \qquad (x \in \mathbf{K}^n).$$

Ekkor az

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p \qquad (x, y \in \mathbf{K}^n)$$

háromszögegyenlőtlenség (az ún. *Minkowski-egyenlőtlenség*) a $p \neq +\infty$ esetben részletesen kiírva a következőt jelenti:

$$\Big(\sum_{i=1}^{n} |x_i + y_i|^p\Big)^{1/p} \leq \Big(\sum_{i=1}^{n} |x_i|^p\Big)^{1/p} + \Big(\sum_{i=1}^{n} |y_i|^p\Big)^{1/p} \qquad (x \in \mathbf{K}^n).$$

Ez a $p = 1$ esetben (maga a Minkowski-egyenlőtlenség $p = +\infty$-re is) triviális. Különben az ún. *Hölder-egyenlőtlenségen* múlik. Ez utóbbi megfogalmazásához legyen

$$q := \begin{cases} \dfrac{p}{p-1} & (1 < p + \infty) \\ 1 & (p = +\infty) \\ +\infty & (p = 1) \end{cases}$$

(tehát $\frac{1}{p}+\frac{1}{q}=1$). Ha $(x_1,...,x_n),\ (y_1,...,y_n)\in \mathbf{K}^n$, akkor

$$\left|\sum_{k=1}^{n} x_k y_k\right| \leq \begin{cases} \Big(\sum\limits_{k=0}^{n}|x_k|^p\Big)^{1/p}\cdot\Big(\sum\limits_{k=0}^{n}|y_k|^q\Big)^{1/q} & (1<p<+\infty) \\ \max\{|y_k| : k=1,...,n\}\cdot\sum\limits_{k=1}^{n}|x_k| & (p=1) \\ \max\{|x_k| : k=1,...,n\}\cdot\sum\limits_{k=1}^{n}|y_k| & (p=+\infty). \end{cases}$$

Ha $p=2$, akkor $q=2$, és a fenti egyenlőtlenség a Cauchy-Bunyakovszkij-egyenlőtlenség speciális esete.

iii) A Hölder-egyenlőtlenség egyszerűen megkapható az alábbi becslésből: legyen $1<p<+\infty$ és $\frac{1}{p}+\frac{1}{q}=1$. Ekkor

$$ab \leq \frac{a^p}{p}+\frac{b^q}{q} \qquad (a,b\geq 0).$$

Ha itt $p=2$, akkor $q=2$, és az elemi

$$ab \leq \frac{a^2+b^2}{2} \qquad (a,b\geq 0)$$

egyenlőtlenséget kapjuk.

2. Fejezet

1. Emlékeztető.

Legyen az $X\neq\emptyset$ egy nem üres halmaz, és tegyük fel, hogy a

$$\rho : X^2 \to [0,+\infty)$$

függvény a következő tulajdonságokkal rendelkezik:

a) minden $x\in X$ esetén $\rho(x,x)=0$;

b) ha $x,y\in X$ és $\rho(x,y)=0$, akkor $x=y$;

c) bármely $x,y\in X$ választással $\rho(x,y)=\rho(y,x)$;

d) tetszőleges $x,y,z\in X$ elemekkel $\rho(x,y)\leq\rho(x,z)+\rho(y,z)$.

Ekkor a ρ egy *távolságfüggvény* (*metrika*). Ha $x, y \in X$, akkor $\rho(x, y)$ az x, y elemek *távolsága*. Az (X, ρ) rendezett párt *metrikus térnek* nevezzük.

Ha $b \in X$ és $r > 0$, akkor a

$$K_r(b) := \{x \in X : \rho(x, b) < r\}$$

halmaz a b elem *r-sugarú környezete*. Használjuk a $K(b)$ jelölést is a $K_r(b)$ helyett, ha az adott szituációban a $K_r(b)$ *környezet sugara* (r) nem játszik szerepet. Nyilvánvaló, hogy $0 < v \leq r$ esetén

$$K_v(b) \subset K_r(b).$$

Tetszőleges $K_r(a)$ környezet és $b \in K_r(a)$ esetén a

$$0 < v < r - \rho(b, a)$$

feltételnek eleget tevő v „sugárral"

$$K_v(b) \subset K_r(a).$$

Ha ui. $x \in K_v(b)$, azaz $\rho(x, b) < v$, akkor a háromszög-egyenlőtlenség szerint

$$\rho(x, a) \leq \rho(x, b) + \rho(b, a) < v + \rho(b, a) < r.$$

Ez azt jelenti, hogy $x \in K_r(a)$, tehát a $K_v(b) \subset K_r(a)$ tartalmazás valóban fennáll.

2. Metrikus terek topológiája.

Nevezzük valamilyen $\emptyset \neq A \subset X$ halmaz esetén az $a \in A$ pontot az A halmaz *belső pontjának,* ha egy alkalmas $K(a)$ környezettel

$$K(a) \subset A$$

teljesül. Az ilyen tulajdonságú pontok által alkotott halmaz az A ún. *belseje,* amit az

$$\operatorname{int} A$$

szimbólummal fogunk jelölni.[9] Nyilván $\operatorname{int} X = X$, míg az

$$X := \mathbf{R},\ \rho(x, y) := |x - y| \quad (x, y \in \mathbf{R})$$

[9] Esetenként a halmaz jelét zárójelbe tesszük, különösen, ha az illető halmazt egy „hosszabb" szimbólum jelöli.

esetben $\operatorname{int}\{a\} = \emptyset \quad (a \in \mathbf{R})$. Állapodjunk meg abban, hogy

$$\operatorname{int}\emptyset := \emptyset.$$

Tehát bármely $A \subset X$ halmazra

$$\operatorname{int} A \subset A.$$

Könnyű meggondolni ugyanakkor (ld. emlékeztető), hogy pl. tetszőleges $K(a)$ környezetre

$$\operatorname{int} K(a) = K(a).$$

Azt mondjuk, hogy az $A \subset X$ halmaz *nyílt,* ha

$$\operatorname{int} A = A.$$

Így pl. az $\emptyset$ (az üreshalmaz) nyílt halmaz, ill. bármely környezet is az. Más megfogalmazásban tehát egy $\emptyset \neq A \subset X$ halmaz akkor és csak akkor nyílt, ha az A minden pontja belső pontja az A-nak:

$$a \in A \implies a \in \operatorname{int} A.$$

Ismételjük el újra, hogy mit is jelent ez: az

$$\emptyset \neq A \subset X$$

halmaz akkor és csak akkor nyílt, ha tetszőleges $a \in A$ elemének létezik olyan $K(a)$ környezete, hogy

$$K(a) \subset A.$$

Bármely (X, ρ) metrikus tér esetén az X „alaphalmaz" nyílt halmaz. Ha pl. (X, ρ) a diszkrét metrikus tér, amikor is

$$\rho(x, y) = \begin{cases} 0 & (x = y) \\ \\ 1 & (x \neq y) \end{cases} \qquad (x, y \in X),$$

akkor az X összes részhalmaza nyílt halmaz. Valóban, ekkor (pl.)

$$K_{1/2}(a) = \{a\} \qquad (a \in X),$$

következésképpen tetszőleges $\emptyset \neq A \subset X$ halmazra és $a \in A$ pontra

$$K_{1/2}(a) = \{a\} \subset A.$$

Tehát $a \in \operatorname{int} A$. Egyúttal minden $x \in X$ pontra az $\{x\}$ halmaz is nyílt.

Ha viszont a ρ metrika olyan, hogy bármelyik $a \in X$ elemhez és tetszőleges $r > 0$ számhoz van olyan $a \neq x \in X$, hogy

$$\rho(x, a) < r,$$

akkor az X egyelemű részhalmazai közül egyik sem nyílt. Ti. ebben az esetben (az előbbi jelölésekkel) $x \in K_r(a)$, ezért $x \neq a$ miatt $K_r(a)$ nem lehet részhalmaza az $\{a\}$ halmaznak. Ez azt jelenti, hogy $\operatorname{int}\{a\} = \emptyset \neq \{a\}$. Ilyen tulajdonságú metrikus terek pl. a

$$(\mathbf{K}^n, \rho_p) \qquad (1 \leq n \in \mathbf{N},\, 1 \leq p \leq +\infty)$$

terek.

Legyen

$$\mathcal{T}_\rho(X) := \mathcal{T}_\rho := \{A \in \mathcal{P}(X) : A \text{ nyílt}\}.$$

Az X nyílt részhalmazai által meghatározott $\mathcal{T}_\rho$ halmazrendszert az (X, ρ) metrikus tér *topológiájának* nevezzük.

1. Tétel. *Tegyük fel, hogy valamilyen $\Gamma \neq \emptyset$ (index)halmaz esetén az $A_\gamma \subset X \quad (\gamma \in \Gamma)$ halmazok valamennyien nyíltak az (X, ρ) metrikus térben. Ekkor*

- *az $\bigcup\limits_{\gamma \in \Gamma} A_\gamma$ egyesítésük is nyílt;*
- *ha a Γ halmaz véges, akkor a $\bigcap\limits_{\gamma \in \Gamma} A_\gamma$ metszetük is nyílt.*

Tehát más megfogalmazásban:

- $\emptyset, X \in \mathcal{T}_\rho$;
- tetszőleges $\Gamma \neq \emptyset$, $A_\gamma \in \mathcal{T}_\rho \quad (\gamma \in \Gamma)$ választással $\bigcup\limits_{\gamma \in \Gamma} A_\gamma \in \mathcal{T}_\rho$;
- ha az előbbi Γ halmaz véges, akkor $\bigcap\limits_{\gamma \in \Gamma} A_\gamma \in \mathcal{T}_\rho$.

Bizonyítás. Legyen $a \in \bigcup\limits_{\gamma \in \Gamma} A_\gamma$. Ekkor egy $\nu \in \Gamma$ indexszel $a \in A_\nu$. Mivel az A_ν halmaz nyílt, ezért egy alkalmas $K(a)$ környezettel $K(a) \subset A_\nu$. Nyilvánvaló, hogy

$$A_\nu \subset \bigcup_{\gamma \in \Gamma} A_\gamma,$$

így egyúttal

$$K(a) \subset \bigcup_{\gamma\in\Gamma} A_\gamma$$

is teljesül. Más szóval $a \in \operatorname{int}\Big(\bigcup\limits_{\gamma\in\Gamma} A_\gamma\Big)$, azaz $\bigcup\limits_{\gamma\in\Gamma} A_\gamma$ nyílt halmaz.

Most tegyük fel, hogy a Γ halmaz véges, és legyen $a \in \bigcap\limits_{\gamma\in\Gamma} A_\gamma$. Ekkor minden $\gamma \in \Gamma$ mellett $a \in A_\gamma$, következésképpen az A_γ-k nyíltsága miatt egy $r_\gamma > 0$ sugárral

$$K_{r_\gamma}(a) \subset A_\gamma.$$

Ha

$$r := \min\{r_\gamma : \gamma \in \Gamma\}$$

(ami egy pozitív szám), akkor $r \le r_\gamma$ $(\gamma \in \Gamma)$ miatt

$$K_r(a) \subset K_{r_\gamma}(a) \subset A_\gamma \qquad (\gamma \in \Gamma),$$

így

$$K_r(a) \subset \bigcap_{\gamma\in\Gamma} A_\gamma.$$

Tehát $a \in \operatorname{int}\Big(\bigcap\limits_{\gamma\in\Gamma} A_\gamma\Big)$, ezért a $\bigcap\limits_{\gamma\in\Gamma} A_\gamma$ metszethalmaz nyílt. ■

Gondoljuk meg, hogy tetszőleges (X, ρ) metrikus térben minden $A \subset X$ halmazra

$$\operatorname{int} A = \bigcup_{T\in\mathcal{A}} T,$$

ahol $\mathcal{A}$ jelöli az X halmaz összes olyan $T \subset X$ nyílt részhalmaza által alkotott halmazrendszert, amelyre $T \subset A$. (Mivel $\emptyset \in \mathcal{A}$, ezért $\mathcal{A} \neq \emptyset$.)

Az 1. Tétel miatt az $\operatorname{int} A = \bigcup\limits_{T\in\mathcal{A}} T$ halmaz nyílt, továbbá

$$\operatorname{int} A = \bigcup_{T\in\mathcal{A}} T \subset A.$$

Az is világos, hogy ha $C \subset X$ olyan nyílt halmaz, amelyre $C \subset A$, akkor $C \subset \operatorname{int} A$. Ezért is szokták az $\operatorname{int} A$ halmazt az A *legbővebb* nyílt részhalmazának nevezni. Speciálisan, ha $A \subset X$ nyílt, akkor $A \in \mathcal{A}$ és $A = \operatorname{int} A$.

Tegyük fel, hogy adottak az (X, ρ), (X, σ) metrikus terek, és $\rho \sim \sigma$ (azaz a ρ metrika ekvivalens a σ-val). Ekkor

$$\mathcal{T}_\rho(X) = \mathcal{T}_\sigma(X),$$

tehát a két tér topológiája egybeesik. Más szóval az X nyílt részhalmazai a két metrika szerint ugyanazok.

Valóban, ha a c, $C > 0$ konstansokkal

$$c\cdot \rho(x,y) \le \sigma(x,y) \le C\cdot \rho(x,y) \qquad (x,y \in X),$$

akkor tetszőleges $a \in X$, $r > 0$ esetén a σ-szerinti

$$K_r^{(\sigma)}(a) := \{x \in X : \sigma(x,a) < r\}$$

környezetre

$$K_{r/C}^{(\rho)}(a) \subset K_r^{(\sigma)}(a),$$

ahol

$$K_{r/C}^{(\rho)}(a) := \{x \in X : \rho(x,a) < r/C\}.$$

Ha ui. $x \in K_{r/C}^{(\rho)}(a)$, azaz $\rho(x,a) < r/C$, akkor

$$\sigma(x,a) \le C\cdot \rho(x,a) < C\cdot \frac{r}{C} = r,$$

más szóval $x \in K_r^{(\sigma)}(a)$. Ugyanígy kapjuk, hogy

$$K_{cr}^{(\sigma)}(a) \subset K_r^{(\rho)}(a).$$

Ha tehát $\emptyset \ne A \subset X$, $a \in A$ és egy alkalmas $r > 0$ sugárral

$$K_r^{(\sigma)}(a) \subset A,$$

akkor az előbbiek szerint

$$K_{r/C}^{(\rho)}(a) \subset A$$

is igaz, ill.

$$K_r^{(\rho)}(a) \subset A$$

esetén

$$K_{cr}^{(\sigma)}(a) \subset A.$$

Következésképpen az a tény, hogy $a \in \operatorname{int} A$, független attól, hogy az (X,ρ), vagy az (X,σ) metrikus térben „vagyunk". Így az $\operatorname{int} A = A$ egyenlőség is pontosan akkor teljesül a ρ metrika értelmében, ha a σ szerint is fennáll. Röviden:

$$A \in \mathcal{T}_\rho(X) \iff A \in \mathcal{T}_\sigma(X).$$

Az (X,ρ) metrikus térben az $A \subset X$ halmazt *zártnak* fogjuk nevezni, ha az $X \setminus A$ (komplementer) halmaz nyílt. Világos, hogy pl. az $\emptyset$, X halmazok zártak, vagy pl. a diszkrét metrikus térben minden halmaz zárt.

Tetszőleges (X, ρ) metrikus térben minden egyelemű halmaz zárt. Legyen ui. $a \in X$, ekkor bármely $b \in X \setminus \{a\}$ elemre $\rho(a, b) > 0$. Ha

$$0 < v \leq \rho(a, b)$$

és $x \in K_v(b)$, akkor

$$\rho(a, x) \geq \rho(a, b) - \rho(x, b) > \rho(a, b) - v \geq \rho(a, b) - \rho(a, b) = 0,$$

azaz $\rho(a, x) > 0$. Ez azt jelenti, hogy $x \neq a$, más szóval $x \in X \setminus \{a\}$. Ezért

$$K_v(b) \subset X \setminus \{a\},$$

röviden: az $X \setminus \{a\}$ halmaz nyílt, tehát $\{a\}$ valóban zárt.

Zártak pl. a

$$G_r(a) := \{x \in X : \rho(x, a) \leq r\} \qquad (a \in X,\ r > 0)$$

halmazok is. Világos, hogy az $A \subset X$ halmaz akkor és csak akkor nyílt, ha az $X \setminus A$ (komplementer) halmaz zárt.

Nyilvánvaló, hogy az ekvivalens metrikákkal kapcsolatban a topológiákra kapott egyenlőség igaz marad az egyes metrikákra nézve zárt halmazok által meghatározott halmazrendszerekre is. Ha tehát (X, ρ), (X, σ) olyan metrikus terek, hogy $\rho \sim \sigma$, akkor

$$\mathcal{C}_\rho(X) = \mathcal{C}_\sigma(X),$$

ahol általában egy (X, δ) metrikus tér esetén

$$\mathcal{C}_\delta(X) := \{A \in \mathcal{P}(X) : A \text{ zárt}\}.$$

Az ismert De Morgan-azonosságokra utalva az 1. Tételből rögtön következik a

2. Tétel. *Tegyük fel, hogy valamilyen $\Gamma \neq \emptyset$ (index)halmaz esetén az $A_\gamma \subset X \quad (\gamma \in \Gamma)$ halmazok valamennyien zártak az (X, ρ) metrikus térben. Ekkor*

- *a $\bigcap_{\gamma \in \Gamma} A_\gamma$ metszetük is zárt;*
- *ha a Γ halmaz véges, akkor az $\bigcup_{\gamma \in \Gamma} A_\gamma$ egyesítésük is zárt.*

Más szóval:

- $\emptyset, X \in \mathcal{C}_\rho$;
- tetszőleges $\Gamma \neq \emptyset$, $A_\gamma \in \mathcal{C}_\rho \quad (\gamma \in \Gamma)$ választással $\bigcap_{\gamma\in\Gamma} A_\gamma \in \mathcal{C}_\rho$;
- ha az előbbi Γ halmaz véges, akkor $\bigcup_{\gamma\in\Gamma} A_\gamma \in \mathcal{C}_\rho$.

Tekintsük az (X, ρ) metrikus térben az $A \subset X$ halmazt, és jelöljük $\mathcal{X}$-szel az összes olyan $B \subset X$ zárt halmaz által alkotott halmazrendszert, amelyre $A \subset B$. Világos, hogy $\mathcal{X} \neq \emptyset$, hiszen nyilván $X \in \mathcal{X}$. Az előző tétel szerint az

$$\overline{A} := \bigcap_{B\in\mathcal{X}} B$$

halmaz zárt, és mivel minden $B \in \mathcal{X}$ esetén $A \subset B$, ezért

$$A \subset \overline{A}.$$

Az is világos, hogy ha a $C \subset X$ halmaz zárt és $A \subset C$, akkor $\overline{A} \subset C$, ui. $C \in \mathcal{X}$. (Ezért is szokták az $\overline{A}$ halmazt az A-t lefedő legszűkebb zárt halmazként említeni.)

Pl.

$$\overline{\emptyset} = \emptyset, \ \overline{X} = X.$$

Ha az $A \subset X$ halmaz zárt, akkor nyilván $A \in \mathcal{X}$, ezért $\overline{A} \subset A$, így $A = \overline{A}$. Következésképpen azt mondhatjuk, hogy az $A \subset X$ halmaz akkor és csak akkor zárt, ha $A = \overline{A}$.

Legyen $A \subset X$, ekkor az előbbiekben definiált $\overline{A}$ halmazt az A *lezárásának* nevezzük. Azt is mondjuk, hogy $\overline{A}$ az A *lezártja.*

Vezessük be a torlódási pont fogalmát. Legyen az (X, ρ) metrikus tér esetén $A \subset X$ és $\alpha \in X$. Azt mondjuk, hogy az α *torlódási pontja* az A halmaznak, ha minden $K(\alpha)$ környezetre

$$(K(\alpha) \setminus \{\alpha\}) \cap A \neq \emptyset.$$

Más megfogalmazásban tehát ez azzal ekvivalens, hogy tetszőleges $r > 0$ számhoz létezik olyan $\alpha \neq x \in A$, amelyre $x \in K_r(\alpha)$, azaz

$$0 < \rho(x, \alpha) < r.$$

Legyen a fenti A halmaz esetén

$$A' := \{\alpha \in X : \text{ az } \alpha \text{ torlódási pontja az } A\text{-nak}\}.$$

Ekkor teljesül a

3. Tétel. *Tetszőleges* (X, ρ) *metrikus tér,* $\emptyset \neq A \subset X$ *halmaz és* $\alpha \in X$ *esetén fennáll a következő ekvivalencia:* $\alpha \in A'$ *akkor és csak akkor igaz, ha az* α *minden* $K(\alpha)$ *környezetére az* $A \cap K(\alpha)$ *metszet végtelen halmaz.*

Bizonyítás. Az állítás egyik „iránya" nyilvánvaló: ha az $A \cap K(\alpha)$ metszet minden $K(\alpha)$ környezetre végtelen halmaz, azaz $K(\alpha)$-ban végtelen sok A-beli elem van, akkor ezek között olyan is van, amelyik különbözik α-tól. Így

$$(K(\alpha) \setminus \{\alpha\}) \cap A \neq \emptyset,$$

következésképpen $\alpha \in A'$.

Most tegyük fel azt, hogy $\alpha \in A'$, és indirekt módon induljunk ki abból, hogy valamilyen $K(\alpha)$ esetén $A \cap K(\alpha)$ véges halmaz. Mivel $\alpha \in A'$, ezért

$$(K(\alpha) \setminus \{\alpha\}) \cap A \neq \emptyset$$

(és ez a halmaz is véges). Legyen

$$0 < r < \min\{\rho(x, \alpha) : \alpha \neq x \in A \cap K(\alpha)\}.$$

Világos, hogy

$$(K_r(\alpha) \setminus \{\alpha\}) \cap A = \emptyset,$$

szemben az $\alpha \in A'$ feltételezésünkkel. ■

Így pl. akármilyen véges $A \subset X$ halmazra $A' = \emptyset$.

4. Tétel. *Tetszőleges* (X, ρ) *metrikus tér és* $A \subset X$ *esetén az* A *halmaz akkor és csak akkor zárt, ha* $A' \subset A$.

Bizonyítás. Tegyük fel először azt, hogy az A halmaz zárt és $\alpha \in A'$. Ha $\alpha \notin A$ teljesülne, akkor nyilván $\alpha \in X \setminus A$ is igaz lenne. Mivel az A zárt, ezért az $X \setminus A$ halmaz nyílt. Így alkalmas $K(\alpha)$ környezettel

$$K(\alpha) \subset X \setminus A.$$

Ezért $K(\alpha) \cap A = \emptyset$, ami ellentmond annak, hogy $\alpha \in A'$.

Most azt tegyük fel azt, hogy $A' \subset A$. Ha az A nem lenne zárt, akkor az $X \setminus A$ halmaz nem lenne nyílt. Ez azt jelenti, hogy egy alkalmas $\alpha \in X \setminus A$ pontra $\alpha \notin \operatorname{int}(X \setminus A)$. Tehát minden $K(\alpha)$ környezetre

$$K(\alpha) \not\subset X \setminus A,$$

más szóval
$$(K(\alpha) \setminus \{\alpha\}) \cap A \neq \emptyset.$$
Ezért $\alpha \in A'$. Tehát az $A' \subset A$ feltételezésünk miatt $\alpha \in A$, ami nyilván ellentmond annak, hogy $\alpha \in X \setminus A$. ■

Speciálisan minden véges $A \subset X$ halmaz zárt, ui. $A' = \emptyset \subset A$.

3. Megjegyzések.

i) A fontosságára való tekintettel mondjuk el újra: tegyük fel, hogy adottak az (X, ρ), (X, σ) metrikus terek, és a ρ metrika ekvivalens a σ-val. Ekkor
$$\mathcal{C}_\rho(X) = \mathcal{C}_\sigma(X),$$
tehát az X zárt részhalmazai a két metrika szerint ugyanazok.

ii) Láttuk, hogy ha az (X, ρ), (X, σ) metrikus terek esetén $\rho \sim \sigma$, akkor tetszőleges $a \in X$ és $r > 0$ esetén megadhatók olyan $u, v > 0$ számok, amelyekkel
$$K_u^{(\sigma)}(a) \subset K_r^{(\rho)}(a) \subset K_v^{(\sigma)}(a).$$
Speciálisan az $(\mathbf{R}^2, \rho_p)$ $(p = 1, 2, \infty)$ választással azt mondhatjuk, hogy geometriailag mindez a következőt jelenti:

az $a \in \mathbf{R}^2$ középpontú $K_r^{(1)}(a) = K_r^{(\rho_1)}(a)$ $(r > 0)$ rombuszok,
az $a \in \mathbf{R}^2$ középpontú $K_u^{(2)}(a) = K_u^{(\rho_2)}(a)$ $(u > 0)$ körlemezek,
és az $a \in \mathbf{R}^2$ középpontú $K_v^{(\infty)}(a) = K_v^{(\rho_\infty)}(a)$ $(v > 0)$ négyzetek

kölcsönösen egymásba illeszthetők.

iii) A
$$(\mathbf{K}^n, \rho_p) \qquad (1 \leq n \in \mathbf{N},\ 1 \leq p \leq +\infty)$$
metrikus terekre azt mondhatjuk, hogy egy $A \subset \mathbf{K}^n$ halmaz nyíltsága ugyanazt jelenti a ρ_p metrika szerint, mint a ρ_q metrika értelmében $(1 \leq p, q \leq +\infty)$. Éppen ezért (is) a $\mathbf{K}^n$ terekkel kapcsolatos vizsgálódásaink során általában a ρ_2, ρ_1, ρ_∞ metrikák valamelyikét (többnyire a ρ_2-t) fogjuk „használni".

iv) Mutassuk meg, hogy minden (X, ρ) metrikus térben
$$\overline{A} = A \cup A' \qquad (A \subset X).$$

Egyrészt tudjuk ui., hogy $A \subset \overline{A}$. Másrészt, ha lenne olyan $a \in A'$, amelyre $a \notin \overline{A}$, akkor nyilván $a \in X \setminus \overline{A}$. Az $X \setminus \overline{A}$ halmaz nyílt, következésképpen egy alkalmas $K(a)$ környezetre

$$K(a) \subset X \setminus \overline{A} \subset X \setminus A,$$

amiből

$$A \subset X \setminus K(a)$$

következik. Így $K(a) \cap A = \emptyset$, ami ellentmond annak, hogy $a \in A'$. Tehát $a \in \overline{A}$, ezért $A' \subset \overline{A}$.

Ezzel beláttuk, hogy $A \cup A' \subset \overline{A}$. Fordítva, ha $b \in \overline{A}$ és $b \notin A \cup A'$, akkor $b \notin A$ és $b \notin A'$, így egy alkalmas $K(b)$ környezettel

$$K(b) \cap A = \emptyset.$$

Ez azt jelenti, hogy $A \subset X \setminus K(b)$. Mivel az $X \setminus K(b)$ halmaz zárt, ezért innen és a halmaz lezárásának az értelmezéséből azt kapjuk, hogy

$$\overline{A} \subset X \setminus K(b).$$

Ez $b \in K(b)$ és a feltételezett $b \in \overline{A}$ tartalmazás miatt nyilván lehetetlenség.

v) Ha az $A \subset X$ halmazra $\overline{A} = X$, akkor az A *mindenütt sűrű.* Ez a iv) megjegyzést figyelembe véve azt jelenti, hogy tetszőleges $x \in X$ és $K(x)$ környezet esetén van olyan $a \in A$, hogy $a \in K(x)$.

Pl. az $\mathbf{R}$-ben (a „szokásos" metrikával) a racionális számok halmaza mindenütt sűrű.

Hasonlóan, a $(C[a,b], \|.\|_\infty)$ tér esetén Weierstrass egyik híres tétele szerint a polinomok $[a,b]$-re való leszűkítéseinek a halmaza mindenütt sűrű. Tehát minden

$$f : [a,b] \to \mathbf{R}$$

folytonos függvényhez és ε pozitív számhoz van olyan P polinom, amellyel

$$|f(x) - P(x)| < \varepsilon \qquad (x \in [a,b]).$$

vi) Legyen (X, ρ) metrikus tér, $A \subset X$, és az $a \in A$ elemről tegyük fel, hogy $a \notin A'$. Ekkor van olyan $K(a)$ környezet, amelyre

$$K(a) \cap A = \{a\}.$$

Ezért azt mondjuk ekkor, hogy az a elem *diszkrét pontja* (vagy más szóval *izolált pontja*) az A halmaznak. Ha viszont $b \in \overline{A}$, akkor tetszőleges $K(b)$ környezetre

$$K(b) \cap A \neq \emptyset.$$

Ui. a iv) megjegyzést figyelembe véve vagy $b \in A$, amikor is

$$b \in K(b) \cap A,$$

vagy pedig $b \in A'$, amikor meg

$$\Big(K(b) \setminus \{b\}\Big) \cap A \neq \emptyset$$

is igaz. Az ilyen b-re ezért azt mondjuk, hogy az A halmaz *érintkezési pontja.*

vii) Számos esetben a metrikus terekkel kapcsolatos meggondolásokban valójában nem is a metrika, hanem csak a $\mathcal{T}_\rho$ topológia tulajdonságai (ld. 1. Tétel) játszanak szerepet. Pl. az

$$f : \mathbf{R} \to \mathbf{R}$$

függvény akkor és csak akkor folytonos, ha minden $A \subset \mathbf{R}$ nyílt halmazra az A-nak az f által létesített ősképe is nyílt.

Mindez a metrikus tereknél általánosabb absztrakt terek bevezetéséhez vezet. Nevezetesen, legyen adott az X halmaz, a $\mathcal{T} \subset \mathcal{P}(X)$ halmazrendszer pedig legyen olyan, hogy

- $\emptyset, X \in \mathcal{T}$;
- bármely $\Gamma \neq \emptyset$, $A_\gamma \in \mathcal{T}$ $\quad (\gamma \in \Gamma)$ esetén $\bigcup_{\gamma \in \Gamma} A_\gamma \in \mathcal{T}$;
- ha az előbbi Γ halmaz véges, akkor $\bigcap_{\gamma \in \Gamma} A_\gamma \in \mathcal{T}$.

Ekkor a $\mathcal{T}$ halmazrendszert *topológiának,* az $(X, \mathcal{T})$ párt pedig *topologikus térnek* nevezzük. Szokás a $\mathcal{T}$ elemeit *nyílt halmazokként* említeni. Világos, hogy pl.

$$(X, \{\emptyset, X\}),\ (X, \mathcal{P}(X))$$

topologikus terek. Speciálisan, ha (X, ρ) metrikus tér, akkor $(X, \mathcal{T}_\rho)$ egyúttal topologikus tér, ahol $\mathcal{T}_\rho$ a ρ *metrika által indukált topológia.*

viii) Van olyan $(X, \mathcal{T})$ topologikus tér, amelyik nem metrizálható, azaz nincs olyan ρ metrika az X-en, hogy a $\mathcal{T} = \mathcal{T}_\rho$ egyenlőség teljesülne.

Hasonlóan, megadható olyan (X, ρ) metrikus tér, ahol az X ugyan vektortér, de nincs rajta olyan $\|.\|$ norma, hogy

$$\rho(x, y) = \|x - y\| \qquad (x, y \in X).$$

ix) Az 1. Tételben szereplő, tetszőleges véges $\Gamma \neq \emptyset$, $A_\gamma \in \mathcal{T}_\rho \quad (\gamma \in \Gamma)$ választással fennálló $\bigcap\limits_{\gamma \in \Gamma} A_\gamma \in \mathcal{T}_\rho$ állításban a „véges" kitétel lényeges. Legyen ui. a számegyenesen (a „szokásos" metrikával)

$$A_n := (-1/n, 1/n) \qquad (1 \leq n \in \mathbf{N}).$$

Ekkor minden itt szereplő intervallum nyílt halmaz, de a

$$\bigcap_{n=1}^{\infty} A_n = \{0\}$$

metszethalmaz nem nyílt.

3. Fejezet

1. Emlékeztető.

Az (X, ρ) metrikus tér esetén az $A \subset X$ halmaz nyílt, ha $A = \emptyset$, vagy tetszőleges $a \in A$ elemnek létezik olyan $K(a)$ környezete, hogy

$$K(a) \subset A.$$

Az A zárt, ha az $X \setminus A$ (komplementer) halmaz nyílt.

A $z \in X$ torlódási pontja az A halmaznak, ha minden $K(z)$ környezetre

$$(K(z) \setminus \{z\}) \cap A \neq \emptyset.$$

Legyen

$$A' := \{z \in X : \text{a } z \text{ torlódási pontja az } A\text{-nak}\}.$$

Ekkor: *$z \in A'$ akkor és csak akkor igaz, ha a z minden $K(z)$ környezetére az $A \cap K(z)$ metszet végtelen halmaz.*

Továbbá: *az* A *akkor és csak akkor zárt, ha* $A' \subset A$.

2. Konvergencia.

Mit értettünk egy

$$x = (x_n) : \mathbf{N} \to \mathbf{K}$$

számsorozat konvergenciáján? Azt mondtuk, hogy az (x_n) sorozat konvergens, ha egy alkalmas $\alpha \in \mathbf{K}$ mellett bármely $\varepsilon > 0$ számhoz megadható az $N \in \mathbf{N}$ „küszöbindex" úgy, hogy

$$(*) \qquad |x_n - \alpha| < \varepsilon \qquad (N < n \in \mathbf{N}).$$

Beláttuk, hogy tetszőleges (x_n) számsorozathoz legfeljebb egyetlen ilyen α létezhet, amit (ha létezik) az (x_n) határértékének neveztünk, és a

$$\lim x, \quad \lim(x_n), \quad \lim_{n\to\infty} x_n$$

szimbólumok valamelyikével jelöltünk. Azt is írtuk ilyenkor, hogy

$$x_n \to \alpha \qquad (n \to \infty).$$

(Szavakban kifejezve: *az* x_n *tart az* α*-hoz, ha az* n *tart a végtelenbe.*)

Vegyük észre, hogy ha a $(*)$-ban (a $\mathbf{K}$ halmaz algebrai és rendezési struktúráját is kihasználó) $|x_n - \alpha|$ helyett az x_n, α számok *távolságáról* beszélünk, akkor máris megkapjuk a konvergencia absztrakt definícióját.

Tekintsünk ehhez egy (X, ρ) metrikus teret, és legyen az

$$(x_n) : \mathbf{N} \to X$$

egy, az X elemeiből álló sorozat. Az (x_n) sorozatot *konvergensnek* nevezzük, ha van olyan $\alpha \in X$, amelyre bármely $\varepsilon > 0$ „hibakorlát" mellett egy alkalmas $N \in \mathbf{N}$ indexszel igaz a

$$\rho(x_n, \alpha) < \varepsilon \qquad (n \in \mathbf{N},\ n > N)$$

becslés.

Ha ilyen α nincs, akkor azt mondjuk, hogy az (x_n) sorozat *divergens.* Világos pl., hogy minden konstans sorozat konvergens, hiszen

$$\alpha \in X,\ x_n = \alpha \qquad (n \in \mathbf{N})$$

esetén bármely $n \in \mathbf{N}$ indexre

$$\rho(x_n, \alpha) = \rho(\alpha, \alpha) = 0.$$

Ugyanakkor a diszkrét metrikus térben valamely (x_n) sorozat akkor és csak akkor konvergens, ha *kvázikonstans*, azaz létezik olyan $M \in \mathbf{N}$ természetes szám, hogy

$$x_n = x_M \qquad (n \in \mathbf{N}, n \geq M).$$

Ha ui. egy sorozat ilyen, akkor a konvergencia definíciójában az α helyébe az x_M-et, tetszőleges $\varepsilon > 0$ mellett pedig N helyébe M-et írva triviálisan fennáll a

$$\rho(x_n, \alpha) = \rho(x_n, x_M) = \rho(x_M, x_M) = 0 < \varepsilon \qquad (n \in \mathbf{N}, n \geq M)$$

egyenlőtlenség.[10] Fordítva, a fenti konvergencia-definícióban $\varepsilon := 1/2$-et választva egy alkalmas $N \in \mathbf{N}$ mellett

$$\rho(x_n, \alpha) < 1/2 \qquad (\mathbf{N} \ni n > N)$$

adódik, amiből (figyelembe véve a „diszkrét" ρ metrika definícióját) $\rho(x_n, \alpha) = 0$, azaz

$$x_n = \alpha \qquad (\mathbf{N} \ni n > N)$$

következik.

1. Tétel. *Legyen valamilyen (X, ρ) metrikus tér esetén az*

$$x = (x_n) : \mathbf{N} \to X$$

sorozat konvergens. Ekkor egyértelműen létezik olyan $\alpha \in X$ elem, amellyel tetszőleges $\varepsilon > 0$ esetén egy alkalmas $N \in \mathbf{N}$ mellett

$$\rho(x_n, \alpha) < \varepsilon \qquad (n \in \mathbf{N},\ n > N).$$

Bizonyítás. Tegyük fel, hogy a tételben említett $X \ni \alpha$-n kívül egy $\beta \in X$ elemre is igaz a konvergencia definíciója: bármilyen $\varepsilon > 0$ számhoz megadható olyan $M \in \mathbf{N}$, hogy

$$\rho(x_n, \beta) < \varepsilon \qquad (n \in \mathbf{N},\ n > M).$$

Ekkor a ρ metrikára vonatkozó háromszög-egyenlőtlenség miatt tetszőlegesen választott

$$n \in \mathbf{N},\ n > \max\{N, M\}$$

[10] Az eddig mondottak nyilván igazak minden metrikus térben is.

indexre
$$\rho(\alpha,\beta) \le \rho(\alpha,x_n) + \rho(x_n,\beta) < 2\varepsilon.$$
Következésképpen
$$0 \le \rho(\alpha,\beta) < 2\varepsilon.$$
Mivel itt az $\varepsilon > 0$ bármilyen (pozitív) szám lehet, ezért csak $\rho(\alpha,\beta) = 0$ lehetséges. A metrika axiómái szerint innen viszont $\alpha = \beta$ adódik. ■

Ha az $x = (x_n)$ sorozat konvergens, akkor a konvergencia definíciójában szereplő, az előző tétel szerint egyértelműen létező $\alpha \in X$ elemet az illető sorozat *határértékének* nevezzük, és (a számsorozatokkal analóg módon) a
$$\lim x, \quad \lim(x_n), \quad \lim_{n\to\infty} x_n$$
szimbólumok valamelyikével jelöljük. Absztrakt esetben is használjuk minderre az
$$x_n \to \alpha \qquad (n \to \infty)$$
jelölést is (amit most is úgy olvasunk, hogy x_n *tart* α*-hoz, ha* n *tart a végtelenbe*).

Egy konvergens
$$(x_n) : \mathbf{N} \to X$$
sorozatra és az $\alpha \in X$ elemre $\alpha = \lim(x_n)$ azt jelenti tehát, hogy bármely $K(\alpha)$ környezethez[11] létezik olyan $N \in \mathbf{N}$, hogy
$$x_n \in K(\alpha) \qquad (n \in \mathbf{N},\, n > N).$$
Röviden:
$$x_n \in K(\alpha) \;\; (\text{m.m.}\, n \in \mathbf{N})$$
(ahol az „m.m." szimbólum a *majdnem minden* kifejezés rövidítése.)[12] Más szóval tetszőleges $K(\alpha)$ környezetre az
$$\{n \in \mathbf{N} : x_n \notin K(\alpha)\}$$
halmaz legfeljebb véges.

A konvergencia definíciója alapján könnyű meggondolni, hogy ha az
$$(x_n),\, (y_n) : \mathbf{N} \to X$$

[11] $z \in K(\alpha) = K_r(\alpha) \iff \rho(z,\alpha) < r$.

[12] angolul: a.a. (*almost all*), németül: f.a. (*fast alle*), oroszul: p.v. (*pacstyi vszje*) .

sorozatokra
$$\{n \in \mathbf{N} : x_n \neq y_n\}$$
legfeljebb véges halmaz,[13] akkor a két sorozat *ekvikonvergens*:
$$(x_n) \text{ konvergens } \iff (y_n) \text{ konvergens.}$$
Ekkor igaz továbbá, hogy ha az (x_n) konvergens, akkor
$$\lim (x_n) = \lim (y_n).$$

Gyakran hivatkozunk a későbbiekben az alábbi, szintén egyszerűen igazolható állításra: ha az
$$(x_n) : \mathbf{N} \to X$$
sorozat konvergens, akkor minden $k \in \mathbf{N}$ esetén az
$$y_n := x_{n+k} \qquad (n \in \mathbf{N})$$
előírással definiált
$$(y_n) : \mathbf{N} \to X$$
(*eltolt*) sorozat is konvergens és
$$\lim (y_n) = \lim (x_n).$$

3. Megjegyzések.

i) Az
$$X := \mathbf{K},\ \rho(x,y) := |x-y| \qquad (x,y \in \mathbf{K})$$
választással a fentiekben a számsorozatokra kapjuk a konvergencia-fogalmat: az
$$(x_n) : \mathbf{N} \to \mathbf{K}$$
sorozat konvergens,
$$\lim (x_n) = \alpha \in \mathbf{K},$$
ha tetszőleges $\varepsilon > 0$ esetén alkalmas $N \in \mathbf{N}$ index mellett
$$|x_n - \alpha| < \varepsilon \qquad (n \in \mathbf{N},\ n > N).$$

Ha itt $\alpha = 0$, akkor az (x_n)-et *nullasorozatnak* nevezzük.

[13] Tehát $x_n = y_n$ (m.m. $n \in \mathbf{N}$).

ii) Világos, hogy egy (X, ρ) metrikus térben az (x_n) sorozat konvergenciája (és $\lim(x_n) = \alpha$) azzal ekvivalens, hogy a $\big(\rho(x_n, \alpha)\big)$ számsorozat nullasorozat.

iii) Valamely $-\infty < a < b < +\infty$ mellett tekintsük az $X := C[a, b]$ halmazt és a ρ_∞ metrikát. Ha az

$$f_n \in C[a, b] \qquad (n \in \mathbf{N})$$

(függvény)sorozat konvergens és

$$f := \lim(f_n) \in C[a, b],$$

akkor tetszőleges $\varepsilon > 0$ számhoz van olyan $N \in \mathbf{N}$, hogy minden $n \in \mathbf{N}$, $n > N$ esetén

$$\rho_\infty(f_n, f) < \varepsilon,$$

azaz

$$\max\{|f_n(x) - f(x)| : a \leq x \leq b\} < \varepsilon.$$

Azt mondjuk, hogy az (f_n) függvénysorozat *egyenletesen konvergens.* Az f az (f_n) sorozat *határfüggvénye.*

Nyilvánvaló, hogy ekkor az előbbi n indexekre bármelyik $x \in [a, b]$ helyen

$$|f_n(x) - f(x)| < \varepsilon$$

igaz. Más szóval tehát ez azt jelenti, hogy az $(f_n(x))$ (szám)sorozat konvergens és a határértéke $f(x)$. Röviden: az (f_n) függvénysorozat *pontonként konvergens.*[14]

Az előbbi „egyenletes" jelző ugyanakkor arra utal, hogy ekkor minden $\varepsilon > 0$ számhoz létezik (csak az ε-tól függő) $N \in \mathbf{N}$, hogy bármely $n \in \mathbf{N}, n > N$ indexre és $x \in [a, b]$ elemre

$$|f_n(x) - f(x)| < \varepsilon.$$

Ezzel szemben a pontonkénti konvergencia esetén az előbbi kijelentésben szereplő N függhet (az ε mellett) az $[a, b] \ni x$-től is. Pl. az

$$f_n(x) := \begin{cases} 2nx & (0 \leq x \leq 1/(2n)) \\ \\ 2 - 2nx & (1/(2n) < x < 1/n) \\ \\ 0 & (1/n \leq x \leq 1) \end{cases} \qquad (0 < n \in \mathbf{N})$$

[14] Tehát: a $(g_n) : \mathbf{N} \to C[a, b]$ függvénysorozat pontonként konvergál a $g \in C[a, b]$ függvényhez, ha minden $x \in [a, b]$ mellett a $(g_n(x))$ számsorozat konvergál a $g(x)$-hez. Ez azt jelenti, hogy tetszőleges $x \in [a, b]$ és $\varepsilon > 0$ esetén megadható olyan (az x-től is és az ε-tól is függő) $N \in \mathbf{N}$, amellyel $|g_n(x) - g(x)| < \varepsilon$ $(N < n \in \mathbf{N})$.

választással az (f_n) függvénysorozat pontonként nyilván konvergál az

$$f(x) := 0 \qquad (x \in [0,1])$$

függvényhez, de a minden pozitív $\mathbf{N} \ni n$-re fennálló

$$\rho_\infty(f_n, f) = \sup\{|f_n(x) - f(x)| : 0 < x < 1\} = 1$$

egyenlőség miatt az (f_n) (a ρ_∞ metrika értelmében, azaz egyenletesen) nem konvergál az f-hez.

iv) Az (X, ρ) metrikus térben legyen adott egy konvergens (x_n) sorozat. Lássuk be, hogy ha

$$\alpha := \lim(x_n),$$

akkor bármely $\beta \in X$ esetén a $\big(\rho(x_n, \beta)\big)$ számsorozat konvergens és

$$\lim\big(\rho(x_n, \beta)\big) = \rho(\alpha, \beta).$$

Sőt, ha az (y_n) is egy konvergens sorozat és

$$\beta := \lim(y_n),$$

akkor a $\big(\rho(x_n, y_n)\big)$ számsorozat is konvergens és

$$\lim\big(\rho(x_n, y_n)\big) = \rho(\alpha, \beta).$$

Valóban, a háromszög-egyenlőtlenség alapján

$$|\rho(x_n, \beta) - \rho(\alpha, \beta)| \le \rho(x_n, \alpha) \qquad (n \in \mathbf{N}).$$

Mivel $\rho(x_n, \alpha) \to 0$ $(n \to \infty)$, ezért

$$\rho(x_n, \beta) \to \rho(\alpha, \beta) \qquad (n \to \infty)$$

is azonnal adódik.

Hasonlóan, ha $\beta = \lim(y_n)$, akkor

$$\rho(\alpha, \beta) \le \rho(\alpha, x_n) + \rho(x_n, y_n) + \rho(y_n, \beta) \qquad (n \in \mathbf{N}),$$

amiből

$$\rho(\alpha, \beta) - \rho(x_n, y_n) \le \rho(\alpha, x_n) + \rho(y_n, \beta) \qquad (n \in \mathbf{N}),$$

ill. (az α-t az x_n-nel, a β-t az y_n-nel felcserélve)

$$\rho(x_n, y_n) - \rho(\alpha, \beta) \leq \rho(\alpha, x_n) + \rho(y_n, \beta) \qquad (n \in \mathbf{N})$$

következik. Ezért

$$|\rho(\alpha, \beta) - \rho(x_n, y_n)| \leq \rho(\alpha, x_n) + \rho(y_n, \beta) \to 0 \qquad (n \to \infty),$$

tehát

$$\rho(x_n, y_n) \to \rho(\alpha, \beta) \quad (n \to \infty).$$

Speciálisan, ha adott az $(X, \|.\|)$ normált tér és

$$\rho(x, y) := \|x - y\| \qquad (x, y \in X),$$

akkor $\lim(x_n) = \alpha$, azaz

$$\|x_n - \alpha\| \to 0 \qquad (n \to \infty)$$

esetén tetszőleges $\beta \in X$ elemre

$$\|x_n - \beta\| \to \|\alpha - \beta\| \qquad (n \to \infty).$$

Sőt, ha $\lim(y_n) = \beta$, akkor

$$\|x_n - y_n\| \to \|\alpha - \beta\| \qquad (n \to \infty).$$

Ha $\beta = 0$, akkor

$$\|x_n\| \to \|\alpha\| \qquad (n \to \infty).$$

A most mondottakat tovább specializálva tekintsük az $(X, \langle . \rangle)$ euklideszi teret. Ha

$$\|x\| := \sqrt{\langle x, x \rangle} \qquad (x \in X),$$

akkor véve az előbbi (x_n), (y_n) konvergens sorozatokat a következőket mondhatjuk: ha

$$\alpha := \lim(x_n),$$

akkor tetszőleges $\beta \in X$ elemre

$$\langle x_n, \beta \rangle \to \langle \alpha, \beta \rangle \qquad (n \to \infty),$$

ill. $\beta := \lim(y_n)$ esetén

$$\langle x_n, y_n \rangle \to \langle \alpha, \beta \rangle \qquad (n \to \infty).$$

v) Ha ρ, σ ekvivalens metrikák (valamilyen $X \neq \emptyset$ „alaphalmazon"), akkor megfelelő pozitív c, C konstansokkal tetszőleges

$$(x_n) : \mathbf{N} \to X$$

sorozatra és $\alpha \in X$ elemre

$$c \cdot \rho(x_n, \alpha) \leq \sigma(x_n, \alpha) \leq C \cdot \rho(x_n, \alpha) \qquad (n \in \mathbf{N}).$$

Ezért

$$\lim\Big(\rho(x_n, \alpha)\Big) = 0 \iff \lim\Big(\sigma(x_n, \alpha)\Big) = 0.$$

Tehát ekvivalens metrikák esetén egy sorozat konvergenciája és a határértéke nem változik, ha az egyik metrikát a másikkal felcseréljük. Éppen ezért a továbbiakban pl. az

$$X := \mathbf{K}^n \qquad (1 \leq n \in \mathbf{N})$$

esetben a $\rho_2, \rho_1, \rho_\infty$ metrikák valamelyikét (többnyire a ρ_2 ún. *euklideszi metrikát*) fogjuk használni.

vi) Világos, hogy tetszőleges (X, ρ) metrikus tér esetén is igaz a számsorozatok körében megismert tény: ha az

$$(x_n) : \mathbf{N} \to X$$

sorozat konvergens, akkor tetszőleges (ν_n) indexsorozat esetén az (x_{ν_n}) részsorozat is konvergens és

$$\lim(x_n) = \lim(x_{\nu_n}).$$

4. Vektorsorozatok.

Legyen most

$$1 \leq s \in \mathbf{N}, \ \ 0 < p \leq +\infty, \ \ X := \mathbf{K}^s, \ \ \rho := \rho_p,$$

és tekintsünk egy

$$x = (x_n) : \mathbf{N} \to \mathbf{K}^s$$

sorozatot. Ha

$$x_n = (x_{n1}, ..., x_{ns}) \ (\in \mathbf{K}^s) \qquad (n \in \mathbf{N}),$$

akkor minden $i = 1, ..., s$ mellett definiálhatjuk az

$$x^{(i)} := (x_{ni})$$

számsorozatot, az x sorozat i-edik *koordináta-sorozatát.*[15] Ekkor az x *vektorsorozat* konvergenciája az alábbiak szerint „kezelhető" a koordináta-sorozatainak a konvergenciája révén.

2. Tétel. *Az*

$$x = (x_n) : \mathbf{N} \to \mathbf{K}^s$$

sorozat akkor és csak akkor konvergens az előbbi $(\mathbf{K}^s, \rho_p)$ *metrikus térben, ha minden* $x^{(i)}$ $(i = 1, ..., s)$ *koordináta-sorozata konvergens. Továbbá*

$$\mathbf{K}^s \ni (\alpha_1, ..., \alpha_s) = \lim x \iff \alpha_i = \lim x^{(i)} \quad (i = 1, ..., s).$$

Bizonyítás. Tegyük fel először, hogy a tételbeli x sorozat konvergens, legyen

$$\alpha = (\alpha_1, ..., \alpha_s) \in \mathbf{K}^s$$

a határértéke. A ρ_p metrika definíciója szerint ez azt jelenti, hogy tetszőleges $\varepsilon > 0$ számot megadva van olyan $N \in \mathbf{N}$, amellyel az $n \in \mathbf{N}$, $n > N$ indexekre

$$\varepsilon > \rho_p(x_n, \alpha) = \begin{cases} \sum_{i=1}^{s} |x_{ni} - \alpha_i|^p & (p < 1) \\ \Big(\sum_{i=1}^{s} |x_{ni} - \alpha_i|^p\Big)^{1/p} & (1 \le p < +\infty) \\ \max\{|x_{ni} - \alpha_i| : i = 1, ..., s\} & (p = \infty := +\infty). \end{cases}$$

Világos, hogy bármely $i = 1, ..., s$ esetén

$$\rho_p(x_n, \alpha) \ge \begin{cases} |x_{ni} - \alpha_i|^p & (0 < p < 1) \\ |x_{ni} - \alpha_i| & (1 \le p \le +\infty) \end{cases} \qquad (n \in \mathbf{N}),$$

ezért

$$|x_{ni} - \alpha_i| < \varepsilon \qquad (n \in \mathbf{N},\ n > N,\ p \ge 1)$$

és

$$|x_{ni} - \alpha_i| < \sqrt[p]{\varepsilon} \qquad (n \in \mathbf{N},\ n > N,\ 0 < p < 1).$$

Mindez pontosan azt jelenti, hogy az $x^{(i)}$ koordináta-sorozat konvergens és

$$\lim x^{(i)} = \alpha_i \qquad (i = 1, ..., s).$$

[15] Az $x^{(i)}$ sorozat n-edik tagja $(n \in \mathbf{N})$ tehát az x sorozat n-edik tagjának az i-edik koordinátája $(i = 1, ..., s)$.

Fordítva, ha minden $x^{(i)}$ koordináta-sorozat konvergens, akkor legyen

$$\alpha_i := \lim x^{(i)} \qquad (i = 1, ..., s)$$

és

$$\alpha := (\alpha_1, ..., \alpha_s) \in \mathbf{K}^s.$$

Ha $\varepsilon > 0$ tetszőleges, akkor minden $i = 1, ..., s$ mellett létezik olyan $N_i \in \mathbf{N}$ küszöbindex, hogy

$$|x_{ni} - \alpha_i| < \varepsilon \qquad (n \in \mathbf{N},\, n > N_i).$$

Legyen $N := \max\{N_1, ..., N_s\}$, ekkor

$$|x_{ni} - \alpha_i| < \varepsilon \qquad (n \in \mathbf{N},\, n > N,\, i = 1, ..., s).$$

Ezért

$$\rho_p(x_n, \alpha) = \sum_{i=1}^{s} |x_{ni} - \alpha_i|^p < s \cdot \varepsilon^p \qquad (n \in \mathbf{N},\, n > N,\, 0 < p < 1),$$

$$\rho_p(x_n, \alpha) = \Big(\sum_{i=1}^{s} |x_{ni} - \alpha_i|^p \Big)^{1/p} < s^{1/p} \cdot \varepsilon \qquad (n \in \mathbf{N},\, n > N,\, 1 \le p < +\infty),$$

$$\rho_p(x_n, \alpha) = \max\{|x_{ni} - \alpha_i| : i = 1, ..., s\} < \varepsilon \quad (n \in \mathbf{N},\, n > N,\, p = +\infty)$$

Így minden $0 < p \le +\infty$ esetén az (x_n) sorozat konvergens a $(\mathbf{K}^s, \rho_p)$ metrikus térben, és $\lim(x_n) = \alpha$. ■

4. Konvergencia – topológia.

Az alábbi tételben halmazok torlódási pontjait sorozatok segítségével fogjuk jellemezni.

3. Tétel. *Tekintsük valamilyen (X, ρ) metrikus térben az $\emptyset \neq A \subset X$ halmazt, és legyen $\alpha \in X$. Az α akkor és csak akkor torlódási pontja az A-nak, ha van olyan*

$$(x_n) : \mathbf{N} \to A \setminus \{\alpha\}$$

sorozat, amelynek létezik határértéke és $\lim(x_n) = \alpha$.

Bizonyítás. Tegyük fel, hogy $\alpha \in A'$. Ekkor tetszőleges $0 < n \in \mathbf{N}$ esetén van olyan $x_n \in A$, hogy

$$x_n \in K_{1/n}(\alpha) \setminus \{\alpha\}.$$

Következésképpen

$$(x_n) : \mathbf{N} \to A \setminus \{\alpha\}$$

és

$$0 < \rho(x_n, \alpha) < \frac{1}{n} \qquad (0 < n \in \mathbf{N}).$$

Mivel $\lim(1/n) = 0$, ezért mindez azt jelenti, hogy $\lim\Big(\rho(x_n, \alpha)\Big) = 0$, azaz az (x_n) sorozat konvergens és $\lim(x_n) = \alpha$.

Induljunk ki most abból, hogy valamilyen

$$(x_n) : \mathbf{N} \to A \setminus \{\alpha\}$$

sorozatra $\lim(x_n) = \alpha$. Ekkor bármilyen $K(\alpha)$ környezetet is véve létezik olyan $k \in \mathbf{N}$ (sőt, egy küszöbindextől kezdve minden $k \in \mathbf{N}$ ilyen), amelyre $x_k \in K(\alpha)$. Mivel

$$\mathcal{R}_{(x_n)} \subset A \setminus \{\alpha\},$$

így $x_k \neq \alpha$, ezért egyúttal

$$x_k \in (K(\alpha) \setminus \{\alpha\}) \cap A$$

is igaz. Következésképpen $\alpha \in A'$. ■

Az előbbi tételben szereplő (x_n) sorozatról az is feltehető, hogy injektív, azaz

$$x_n \neq x_m \qquad (n, m \in \mathbf{N}, n \neq m).$$

Ui. a fenti bizonyításban az $\alpha \in A'$ torlódási pontot „előállító" (x_n) sorozat konstrukciójakor a

$$0 < \rho(x_{n+1}, \alpha) < \min\{\rho(x_n, \alpha), 1/(n+1)\} \qquad (n \in \mathbf{N})$$

egyenlőtlenségnek megfelelően választva a sorozat tagjait, mindez nyilvánvaló.

Az előbbiek alapján már egyszerűen belátható a halmazok zártságát konvergens sorozatokkal jellemző alábbi tétel.

4. Tétel. *Legyen* (X, ρ) *metrikus tér. Az* $\emptyset \neq A \subset X$ *halmaz akkor és csak akkor zárt, ha minden konvergens*

$$(x_n) : \mathbf{N} \to A$$

sorozatra $\lim(x_n) \in A$.

Bizonyítás. Tegyük fel először azt, hogy az A halmaz zárt, de valamilyen

$$(x_n) : \mathbf{N} \to A$$

konvergens sorozatra

$$\alpha := \lim(x_n) \notin A.$$

Ekkor tehát $\alpha \in X \setminus A$, ahol az $X \setminus A$ halmaz nyílt. Így van olyan $K(\alpha)$ környezet, hogy

$$K(\alpha) \subset X \setminus A.$$

Ugyanakkor egy $N \in \mathbf{N}$ indexszel

$$A \ni x_k \in K(\alpha) \subset X \setminus A \qquad (N < k \in \mathbf{N}),$$

ami nyilván nem lehet.

Most tegyük fel azt, hogy tetszőleges konvergens

$$(x_n) : \mathbf{N} \to A$$

sorozat határértékére $\lim(x_n) \in A$, és lássuk be, hogy az A halmaz zárt. Legyen ehhez $\alpha \in A'$, ekkor (ld. 3. Tétel) egy alkalmas

$$(z_n) : \mathbf{N} \to A$$

sorozatra $\lim(z_n) = \alpha$. A kiinduló feltételünk szerint ezért $\alpha \in A$, azaz $A' \subset A$ és (egy korábbi tételre hivatkozva) az A zárt. ∎

5. Megjegyzések.

i) Tegyük fel, hogy az

$$(x_n) : \mathbf{N} \to X$$

sorozat konvergens, α a határértéke, és az $\mathcal{R}_{(x_n)}$ értékkészlet végtelen halmaz. Ekkor az α torlódási pontja az $\mathcal{R}_{(x_n)}$-nek. Ui. tetszőleges $K(\alpha)$ környezetre legfeljebb véges sok $k \in \mathbf{N}$ indexszel $x_k \notin K(\alpha)$. Így a

$$K(\alpha) \cap \mathcal{R}_{(x_n)}$$

metszethalmaz nyilván végtelen.

ii) Ha adott az $\emptyset \neq A \subset X$ halmaz, és a konvergens

$$(x_n) : \mathbf{N} \to A$$

sorozat $\mathcal{R}_{(x_n)}$ értékkészlete véges, akkor

$$\alpha := \lim(x_n) \in \mathcal{R}_{(x_n)}.$$

Ti. az $\alpha \notin \mathcal{R}_{(x_n)}$ esetben az

$$0 < r < \min\{\rho(x_n, \alpha) : n \in \mathbf{N}\}$$

választással

$$K_r(\alpha) \cap \mathcal{R}_{(x_n)} = \emptyset,$$

ami ellentmond annak, hogy az α az (x_n) sorozat határértéke.[16]

Ha viszont az $\mathcal{R}_{(x_n)}$ értékkészlet végtelen halmaz, akkor az i) megjegyzés szerint az α torlódási pontja az $\mathcal{R}_{(x_n)}$-nek,[17] ezért $\mathcal{R}_{(x_n)} \subset A$ miatt nyilván az A-nak is. Következésképpen zárt A halmaz esetén $\alpha \in A$.

4. Fejezet

1. Emlékeztető.

Tekintsük az (X, ρ) metrikus teret és az

$$(x_n) : \mathbf{N} \to X$$

sorozatot. Az (x_n) konvergens, ha van olyan $\alpha \in X$, amelyre bármely $\varepsilon > 0$ számhoz megadható egy $N \in \mathbf{N}$ úgy, hogy

$$\rho(x_n, \alpha) < \varepsilon \qquad (n \in \mathbf{N},\, n > N).$$

Ha az (x_n) sorozat konvergens, akkor egyértelműen létezik a fenti α, az $x = (x_n)$ határértéke, amit a

$$\lim x, \ \ \lim(x_n), \ \ \lim_{n\to\infty} x_n$$

[16] Lássuk be, hogy egy alkalmas $N \in \mathbf{N}$ mellett $x_k = \alpha \quad (N \leq k \in \mathbf{N})$.

[17] Gondoljuk meg, hogy ekkor az α az egyetlen torlódási pontja az $\mathcal{R}_{(x_n)}$-nek.

szimbólumok valamelyikével jelölünk. Használjuk erre az

$$x_n \to \alpha \qquad (n \to \infty)$$

jelölést is.[18]

Mindez azt jelenti tehát, hogy bármely $K(\alpha)$ környezethez létezik olyan $N \in \mathbf{N}$, hogy

$$x_n \in K(\alpha) \qquad (n \in \mathbf{N},\, n > N).$$

Röviden:

$$x_n \in K(\alpha) \;\; (\text{m.m.}\, n \in \mathbf{N})$$

(ahol az „m.m." szimbólum a *majdnem minden* kifejezés rövidítése.)

2. Korlátosság.

Nevezzük az $A \subset X$ halmazt *korlátosnak,* ha megadható olyan $\xi \in X$ elem és annak olyan $K(\xi)$ környezete, hogy

$$A \subset K(\xi).$$

Nyilvánvaló, hogy minden környezet korlátos halmaz. Bármely konvergens

$$(x_n) : \mathbf{N} \to X$$

sorozat esetén az $\mathcal{R}_{(x_n)}$ értékkészlet korlátos halmaz (röviden: az (x_n) *korlátos sorozat*). Ha ui.

$$\alpha := \lim(x_n),$$

akkor egy alkalmas $N \in \mathbf{N}$ küszöbindexszel

$$\rho(x_n, \alpha) < 1 \qquad (n \in \mathbf{N},\, n > N).$$

Legyen

$$v := \max\{1, \rho(x_k, \alpha) : k = 0, ..., N\},$$

akkor minden $r > v$ szám mellett nyilván

$$\rho(x_m, \alpha) < r \qquad (m \in \mathbf{N}),$$

azaz $\mathcal{R}_{(x_n)} \subset K_r(\alpha)$.

Véges sok korlátos halmaz egyesítése is korlátos. Valóban, tegyük fel, hogy $n \in \mathbf{N}$ és az

$$A_i \subset X \qquad (i = 0, ..., n)$$

[18] „x_n tart α-hoz, ha n tart a végtelenbe".

halmazok korlátosak. Tehát alkalmas $a_i \in X,\ r_i > 0$ mellett

$$A_i \subset K_{r_i}(a_i) \qquad (i = 0, ..., n).$$

Ha $\xi := a_0$ és

$$r := \max\{\rho(a_i, \xi) + r_i : i = 0, ..., n\},$$

akkor

$$\bigcup_{i=0}^{n} A_i \subset K_r(\xi).$$

Legyen ui. $x \in \bigcup_{i=0}^{n} A_i$, azaz valamilyen $j = 0, ..., n$ indexre $x \in A_j$. Következésképpen

$$x \in K_{r_j}(a_j),$$

azaz $\rho(x, a_j) < r_j$. Ezért

$$\rho(x, \xi) \le \rho(x, a_j) + \rho(a_j, \xi) < r_j + \rho(a_j, \xi) \le r.$$

Ez azt jelenti, hogy $\rho(x, \xi) < r$, más szóval $x \in K_r(\xi)$, tehát

$$\bigcup_{i=0}^{n} A_i \subset K_r(\xi).$$

Analóg gondolatmenettel láthatók be az alábbi állítások is:

- az $A \subset X$ halmaz akkor és csak akkor korlátos, ha az A alkalmasan választott véges sok környezet egyesítésének a részhalmaza;
- ha az $A \subset X$ halmaz korlátos, akkor minden $\xi \in X$ elemnek van olyan $K(\xi)$ környezete, hogy $A \subset K(\xi)$;
- speciálisan, ha adott az $(X, \|.\|)$ normált tér, és

  $$\rho(x, y) := \|x - y\| \qquad \Big((x, y) \in X^2\Big),$$

  akkor az $A \subset X$ halmaz korlátossága azzal ekvivalens, hogy

  $$\sup\{\|x\| : x \in A\} < +\infty.$$

  Más szóval megadható olyan $r > 0$ szám, amellyel

  $$\|x\| \le r \qquad (x \in A).$$

3. Bolzano–Weierstrass-kiválasztási tétel.

A vektorsorozatokra is igaz a számsorozatok körében megismert Bolzano-Weierstrass-kiválasztási tétel megfelelője, nevezetesen:

1. Tétel. *A* $(\mathbf{K}^s, \rho_p)$ $(1 \leq s \in \mathbf{N},\, 0 < p \leq +\infty)$ *metrikus térben minden korlátos sorozatnak van konvergens részsorozata.*

Bizonyítás. Emlékeztetünk egy korábbi tételre, miszerint az

$$x = (x_n) : \mathbf{N} \to \mathbf{K}^s$$

sorozat akkor és csak akkor konvergens, ha minden $x^{(i)}$ $(i = 1, ..., s)$ koordináta-sorozata konvergens.

A feltételezésünk szerint most az (x_n) sorozat korlátos. Van tehát olyan $r > 0$ szám, amellyel

$$\rho_p(x_n, 0) < r \qquad (n \in \mathbf{N}).$$

A ρ_p metrika definícióját figyelembe véve innen az is rögtön adódik, hogy $1 \leq p \leq +\infty$ esetén

$$|x_{ni}| < r \qquad (n \in \mathbf{N},\, i = 1, ..., s),$$

ill. $0 < p < 1$ mellett

$$|x_{ni}| < r^{1/p} \qquad (n \in \mathbf{N},\, i = 1, ..., s),$$

azaz, hogy minden $x^{(i)}$ $(i = 1, ..., s)$ koordináta-sorozat (mint számsorozat) is korlátos. A számsorozatokra ismert Bolzano-Weierstrass-kiválasztási tétel alapján ezért létezik olyan $\nu^{(1)}$ indexsorozat[19], hogy az

$$x^{(1)} \circ \nu^{(1)}$$

részsorozat konvergens. Világos, hogy az $x^{(2)} \circ \nu^{(1)}$ résszorozat is korlátos, ezért van olyan $\nu^{(2)}$ indexsorozat is, amelyre az

$$\left(x^{(2)} \circ \nu^{(1)}\right) \circ \nu^{(2)} = x^{(2)} \circ \left(\nu^{(1)} \circ \nu^{(2)}\right)$$

résszorozat is konvergens. A konstrukciót folytatva végül olyan

$$\nu^{(i)} \qquad (i = 1, ..., s)$$

[19]Tehát a $\nu^{(1)} : \mathbf{N} \to \mathbf{N}$ sorozat szigorúan monoton növő.

indexsorozatokat kapunk, hogy az

$$x^{(i)} \circ \left(\nu^{(1)} \circ \ldots \circ \nu^{(i)}\right) \qquad (i = 1, ..., s)$$

részsorozatok konvergensek. Legyen

$$\nu := \nu^{(1)} \circ \ldots \circ \nu^{(s)},$$

ekkor a ν sorozat indexsorozat, és minden

$$x^{(i)} \circ \nu \qquad (i = 1, ..., s)$$

sorozat részsorozata a konvergens $x^{(i)} \circ \left(\nu^{(1)} \circ \ldots \circ \nu^{(i)}\right)$ sorozatnak. Így az

$$x^{(i)} \circ \nu \qquad (i = 1, ..., s)$$

számsorozatok mindegyike konvergens. Ez azt jelenti, hogy az $x \circ \nu$ részsorozat is konvergens. ■

4. Cauchy-sorozatok, teljes terek.

Legyen az (X, ρ) metrikus térben az

$$(y_n) : \mathbf{N} \to X$$

sorozat konvergens. Ekkor bármely $\varepsilon > 0$ számhoz van olyan $N \in \mathbf{N}$, amellyel

$$\rho(y_n, y_m) < \varepsilon \qquad (m, n \in \mathbf{N},\ m, n > N).$$

Valóban, ha $\alpha := \lim (y_n)$, akkor az előbbi ε-hoz létezik olyan $N \in \mathbf{N}$, hogy

$$\rho(y_k, \alpha) < \frac{\varepsilon}{2} \qquad (k \in \mathbf{N},\ k > N).$$

A háromszög-egyenlőtlenség miatt minden $m, n \in \mathbf{N}$, $m, n > N$ esetén

$$\rho(y_n, y_m) \leq \rho(y_n \alpha) + \rho(y_m, \alpha) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

A továbbiakban azt mondjuk, hogy az

$$(x_n) : \mathbf{N} \to X$$

sorozat *Cauchy-sorozat*, ha tetszőleges $\varepsilon > 0$ esetén létezik olyan $N \in \mathbf{N}$, hogy

$$\rho(x_n, x_m) < \varepsilon \qquad (m, n \in \mathbf{N},\ m, n > N).$$

A most megfogalmazott feltételt *Cauchy-kritériumként* (vagy *Cauchy-feltételként*) fogjuk idézni.

Tudjuk, hogy az

$$X := \mathbf{K}, \;\; \rho(x, y) := |x - y| \qquad (x, y \in \mathbf{K})$$

speciális esetben – azaz számsorozatokra – a Cauchy-kritérium elégséges is ahhoz, hogy az (x_n) (szám)sorozat konvergens legyen. Ugyanakkor tetszőleges metrikus térben ez már nem mondható el.

Legyen ti.

$$X := \mathbf{Q}, \;\; \rho(x, y) := |x - y| \qquad (x, y \in \mathbf{Q}),$$

valamint

$$x_0 := 2, \;\; x_{n+1} := \frac{x_n}{2} + \frac{1}{x_n} \qquad (n \in \mathbf{N}).$$

Jól ismert, hogy az így definiált (x_n) (racionális)számsorozattal tetszőleges $\varepsilon > 0$ szám mellett egy alkalmas $N \in \mathbf{N}$ küszöbindexszel

$$|x_n - \sqrt{2}| < \frac{\varepsilon}{2} \qquad (n \in \mathbf{N}, n > N),$$

és a $\sqrt{2}$ az egyetlen olyan valós szám, amire ez teljesül. Következésképpen a fentiek szerint

$$|x_n - x_m| < \varepsilon \qquad (m, n \in \mathbf{N},\, m, n > N).$$

Tehát a szóban forgó $(\mathbf{Q}, \rho)$ metrikus térben ez az (x_n) sorozat Cauchy-sorozat. Ha konvergens lenne, akkor egy $\alpha \in \mathbf{Q}$ (racionális) számmal minden $\varepsilon > 0$ mellett valamilyen $M \in \mathbf{N}$ indexszel

$$|x_n - \alpha| < \varepsilon \qquad (n \in \mathbf{N}, n > M)$$

teljesülne. Az előbbiek szerint viszont az α csak $\sqrt{2}$ lehetne, ami $\sqrt{2} \notin \mathbf{Q}$ miatt nem lehetséges.

Különös jelentőséggel bírnak azok a metrikus terek, amelyekben a számsorozatokhoz hasonlóan a Cauchy-kritérium elégséges is az illető sorozat konvergenciájához. Ezekkel kapcsolatos a következő definíció: az (X, ρ) metrikus tér *teljes metrikus tér*, ha a Cauchy-kritériumnak eleget tevő bármely sorozat (azaz minden Cauchy-sorozat) konvergens.

2. Tétel. *Tetszőleges* $1 \leq s \in \mathbf{N}$, $0 < p \leq +\infty$ *esetén a* $(\mathbf{K}^s, \rho_p)$ *metrikus tér teljes.*

Bizonyítás. Legyen adott egy

$$x = (x_n) : \mathbf{N} \to \mathbf{K}^s$$

sorozat, és tekintsük az

$$x^{(i)} = (x_{ni}) : \mathbf{N} \to \mathbf{K} \qquad (i = 1, ..., s)$$

koordináta-sorozatokat. Az 1. Tétel bizonyításában látottakkal analóg módon kapjuk, hogy ha az (x_n) sorozat Cauchy-sorozat, akkor minden koordináta-sorozata, mint számsorozat, Cauchy-sorozat.[20] Ui. a feltételezésünk szerint minden ε pozitív számhoz van olyan $N \in \mathbf{N}$, hogy

$$\rho_p(x_n, x_m) < \varepsilon \qquad (N < n, m \in \mathbf{N}).$$

A ρ_p metrika definíciója szerint innen az előbb mondott $N < n, m \in \mathbf{N}$ indexekre $1 \leq p \leq +\infty$ esetén

$$|x_{ni} - x_{mi}| < \varepsilon \qquad (i = 1, ..., s),$$

ill., ha $0 < p < 1$, akkor

$$|x_{ni} - x_{mi}| < \varepsilon^{1/p} \qquad (i = 1, ..., s).$$

Más szóval bármelyik

$$x^{(i)} : \mathbf{N} \to \mathbf{K} \qquad (i = 1, ..., s)$$

koordináta-(szám)sorozat is Cauchy-sorozat, következésképpen konvergens. Ezért (egy korábbi tételünk miatt) az (x_n) sorozat is konvergens. ■

3. Tétel. *A* $(C[a,b], \rho_\infty)$ *metrikus tér teljes.*

Bizonyítás. Tegyük fel, hogy az

$$(f_n) : \mathbf{N} \to C[a,b]$$

(függvény)sorozat (a ρ_∞ metrika értelmében) Cauchy-sorozat. Ez most azt jelenti, hogy bármilyen $\varepsilon > 0$ számot is adunk meg, ehhez találunk olyan $N \in \mathbf{N}$ indexet, hogy

[20] Sőt, az (x_n) sorozat akkor és csak akkor Cauchy-sorozat, ha minden koordinátasorozata Cauchy-sorozat.

$$\rho_\infty(f_n, f_m) =$$

$$\max\{|f_n(x) - f_m(x)| : a \le x \le b\} < \varepsilon \qquad (n, m \in \mathbf{N},\ n, m > N).$$

Világos, hogy tetszőleges $x \in [a, b]$ esetén egyúttal

$$(*) \qquad |f_n(x) - f_m(x)| < \varepsilon \qquad (n, m \in \mathbf{N},\ n, m > N)$$

is teljesül, más szóval az $\big(f_n(x)\big)$ számsorozat Cauchy-sorozat. Létezik tehát az

$$f(x) := \lim_{n\to\infty} f_n(x) \qquad (x \in [a, b])$$

(„pontonkénti") határérték. Továbbá a $(*)$ miatt

$$|f_n(x) - f(x)| =$$

$$(**) \qquad \lim_{m\to\infty} |f_n(x) - f_m(x)| \le \varepsilon \qquad (x \in [a, b],\ n \in \mathbf{N},\ n > N).$$

Mutassuk meg, hogy az így definiált

$$f : [a, b] \to \mathbf{R}$$

függvény folytonos, azaz $f \in C[a, b]$. Legyen ehhez valamilyen $\xi \in [a, b]$ mellett $\varepsilon > 0$ tetszőleges, ekkor az előbbiek szerint bármilyen (rögzített) $n \in \mathbf{N}$, $n > N$ esetén

$$|f(x) - f(\xi)| \le |f(x) - f_n(x)| + |f_n(x) - f_n(\xi)| + |f_n(\xi) - f(\xi)| \le$$

$$2\varepsilon + |f_n(x) - f_n(\xi)| \qquad (x \in [a, b]).$$

Mivel $f_n \in C[a, b]$, így $f_n \in \mathcal{C}\{\xi\}$ is igaz. Következésképpen létezik olyan $\delta > 0$ szám, amellyel

$$|f_n(x) - f_n(\xi)| < \varepsilon \qquad (x \in [a, b],\ |x - \xi| < \delta).$$

Mindezeket figyelembe véve azt mondhatjuk, hogy

$$|f(x) - f(\xi)| < 2\varepsilon + \varepsilon = 3\varepsilon \qquad (x \in [a, b],\ |x - \xi| < \delta).$$

Ez nem jelent mást, mint azt, hogy $f \in \mathcal{C}\{\xi\}$. Az itt szereplő ξ tetszőleges eleme volt az $[a, b]$ intervallumnak, ezért $f \in C[a, b]$.

Végül, a $(**)$ becslés szerint (az ottani szereplőkkel)

$$\rho_\infty(f_n, f) = \max\{|f_n(x) - f(x)| : a \le x \le b\} \le \varepsilon \qquad (n \in \mathbf{N},\ n > N),$$

azaz

$$\rho_\infty(f_n, f) \to 0 \qquad (n \to \infty).$$

Ez azzal ekvivalens, hogy a $(C[a,b], \rho_\infty)$ metrikus térben az (f_n) sorozat konvergál az f függvényhez.

Ezzel beláttuk, hogy a szóban forgó térben minden Cauchy-sorozat konvergens, azaz a $(C[a,b], \rho_\infty)$ teljes metrikus tér. ■

5. Fixpont-tétel.

A teljes metrikus terekkel kapcsolatban az egyik legfontosabb állításunk a Banach(1892–1945)–Tyihonov(1906–1993)–Cacciopoli(1904–1959)–*fixpont-tétel.*

4. Tétel. *Legyen adott az* (X, ρ) *teljes metrikus tér, valamint az*

$$f : X \to X$$

leképezés. Tegyük fel, hogy valamilyen $0 \le q < 1$ *együtthatóval*

$$\rho\big(f(x), f(y)\big) \le q \cdot \rho(x, y) \qquad (x, y \in X).$$

Ekkor

a) *egyértelműen létezik olyan* $\alpha \in X$, *amire* $f(\alpha) = \alpha$;

b) *bármilyen* $x_0 \in X$ *elemet véve az*

$$x_{n+1} := f(x_n) \qquad (n \in \mathbf{N})$$

rekurzióval definiált sorozat konvergens, és $\lim (x_n) = \alpha$;

c) *igaz a következő becslés*:

$$\rho(x_n, \alpha) \le \frac{q^n}{1-q} \cdot \rho(x_0, x_1) \qquad (n \in \mathbf{N}).$$

Bizonyítás. Teljes indukcióval rögtön adódik, hogy

$$\rho(x_n, x_{n+1}) \le q^n \cdot \rho(x_0, x_1) \qquad (n \in \mathbf{N}).$$

Valóban, $n = 0$-ra mindez triviális, $n \in \mathbf{N}$ esetén pedig

$$\rho(x_{n+1}, x_{n+2}) = \rho\Big(f(x_n), f(x_{n+1})\Big) \leq q \cdot \rho(x_n, x_{n+1}).$$

Ha tehát $n \in \mathbf{N}$ olyan, amelyre a szóban forgó

$$\rho(x_n, x_{n+1}) \leq q^n \cdot \rho(x_0, x_1)$$

egyenlőtlenség teljesül, akkor

$$\rho(x_{n+1}, x_{n+2}) \leq q \cdot q^n \cdot \rho(x_0, x_1) = q^{n+1} \cdot \rho(x_0, x_1)$$

is igaz. Innen – a háromszög-egyenlőtlenség többszöri alkalmazásával – akármilyen $n, s \in \mathbf{N}$ mellett

$$\rho(x_n, x_{n+s}) \leq \sum_{k=n}^{n+s-1} \rho(x_{k+1}, x_k) \leq \rho(x_0, x_1) \cdot \sum_{k=n}^{n+s-1} q^k \leq \rho(x_0, x_1) \cdot \sum_{k=n}^{\infty} q^k,$$

tehát

$$\rho(x_n, x_{n+s}) \leq \rho(x_0, x_1) \cdot \frac{q^n}{1-q}$$

következik. Mivel a (q^n) számsorozat 0-hoz konvergál, ezért az (x_n)-re teljesül a Cauchy-kritérium. Ha ui. a $\delta > 0$ tetszőlegesen adott, akkor egy $N \in \mathbf{N}$ indexszel

$$q^n < \delta \qquad (n \in \mathbf{N}, n > N).$$

Ezért

$$\rho(x_n, x_{n+s}) \leq \rho(x_0, x_1) \cdot \frac{\delta}{1-q} \qquad (n, s \in \mathbf{N}, n > N).$$

Legyen $\varepsilon > 0$, és válasszuk az előbbi δ-t úgy, hogy

$$\rho(x_0, x_1) \cdot \frac{\delta}{1-q} < \varepsilon.$$

Ekkor

$$\rho(x_n, x_{n+s}) < \varepsilon \qquad (n, s \in \mathbf{N}, n > N).$$

Ha tehát $N < k, j \in \mathbf{N}$, és (pl.) $k \leq j$, akkor az $s := j - k \in \mathbf{N}$ jelöléssel

$$\rho(x_k, x_j) = \rho(x_k, x_{k+s}) < \varepsilon.$$

Következésképpen az (x_n) sorozat valóban Cauchy-sorozat, így a tér teljessége miatt konvergens.

Legyen

$$\alpha := \lim{(x_n)}.$$

Ekkor ismét a háromszög-egyenlőtlenség alapján bármilyen $n \in \mathbf{N}$ természetes számmal

$$0 \leq \rho\Big(f(\alpha), \alpha\Big) \leq \rho\Big(f(\alpha), f(x_n)\Big) + \rho\Big(f(x_n), \alpha\Big) =$$

$$\rho\Big(f(\alpha), f(x_n)\Big) + \rho(x_{n+1}, \alpha) \leq$$

$$q \cdot \rho(\alpha, x_n) + \rho(x_{n+1}, \alpha) \to 0 \qquad (n \to \infty).$$

Ezért $\rho\Big(f(\alpha), \alpha\Big) = 0,$ azaz

$$f(\alpha) = \alpha.$$

Ha ez utóbbi egyenlőség valamilyen $\beta \in X$ elemre is fennáll, akkor

$$\rho(\alpha, \beta) = \rho\Big(f(\alpha), f(\beta)\Big) \leq q \cdot \rho(\alpha, \beta),$$

és $0 \leq q < 1$ alapján $\rho(\alpha, \beta) = 0.$ Tehát $\alpha = \beta.$

Továbbá, tetszőleges $n \in \mathbf{N}$ mellett

$$x_{n+s} \to \alpha \qquad (s \to \infty),$$

így az

$$y_s := \rho(x_n, x_{n+s}) \qquad (s \in \mathbf{N})$$

sorozat konvergens, és a határértéke $\rho(x_n, \alpha).$ Innen a fentieket is figyelembe véve

$$\rho(x_n, \alpha) = \lim_{s \to \infty} y_s \leq \frac{q^n}{1-q} \cdot \rho(x_0, x_1) \qquad (n \in \mathbf{N}),$$

azaz a tételünk utolsó állítása is adódik. ∎

6. Megjegyzések.

i) Mutassuk meg, hogy ha az (X, ρ) metrikus térben valamilyen (x_n) Cauchy-sorozatnak van konvergens részsorozata, akkor az (x_n) sorozat konvergens.

Legyen ui. a (ν_n) indexsorozat olyan, hogy az (x_{ν_n}) részsorozat konvergens, ill. legyen

$$\alpha := \lim(x_{\nu_n}).$$

Ekkor minden $\varepsilon > 0$ számhoz van olyan $N \in \mathbf{N},$ hogy

$$\rho(x_{\nu_n}, \alpha) < \varepsilon \qquad (n \in \mathbf{N},\, n > N).$$

Az (x_n) Cauchy-sorozat lévén olyan $M \in \mathbf{N}$ is megadható, amellyel

$$\rho(x_k, x_j) < \varepsilon \qquad (k, j \in \mathbf{N},\, k, j > M).$$

Ekkor

$$\rho(x_n,\alpha)\le\rho(x_n,x_{\nu_n})+\rho(x_{\nu_n},\alpha)<\rho(x_n,x_{\nu_n})+\varepsilon\qquad(N<n\in\mathbf{N}).$$

Vegyük figyelembe, hogy

$$\nu_n\ge n\qquad(n\in\mathbf{N}),$$

ezért

$$\rho(x_n,x_{\nu_n})<\varepsilon\qquad(n\in\mathbf{N},\ n>M).$$

Következésképpen

$$\rho(x_n,\alpha)<2\varepsilon\qquad(n\in\mathbf{N},\ n>\max\{N,M\}).$$

Ezért az (x_n) sorozat konvergens és $\lim(x_n)=\alpha$.

ii) Legyen adott az $(X,\|.\|)$ normált tér. Azt mondjuk, hogy ez a tér *teljes* (vagy *Banach-tér*), ha a $\|.\|$ norma által indukált

$$\rho(x,y):=\|x-y\|\qquad(x,y\in X)$$

metrikával az (X,ρ) metrikus tér teljes. Világos, hogy a

$$(\mathbf{K}^s,\|.\|_p)\qquad(1\le s\in\mathbf{N},\ 1\le p\le+\infty)$$

terek valamennyien Banach-terek. Hasonlóan: a $(C[a,b],\|.\|_\infty)$ tér is Banach-tér.

iii) Azt mondjuk, hogy az $(X,\langle.\rangle)$ euklideszi tér *teljes* (vagy *Hilbert*[21]*-tér*), ha a $\langle.\rangle$ skaláris szorzás által meghatározott

$$\|x\|:=\sqrt{\langle x,x\rangle}\qquad(x\in X)$$

normával $(X,\|.\|)$ Banach-tér. Így pl. a

$$(\mathbf{K}^s,\langle.\rangle)\qquad(1\le s\in\mathbf{N})$$

tér Hilbert-tér, ahol

$$\langle x,y\rangle=\sum_{i=1}^{s}x_i\overline{y}_i\qquad\Big(x=(x_1,...,x_s),\ y=(y_1,...,y_s)\in\mathbf{K}^s\Big).$$

[21] David Hilbert (1862–1943).

iv) A fixpont-tételben szereplő, a valamilyen $0 \le q < 1$ együtthatóval a

$$\rho\Big(f(x), f(y)\Big) \le q\cdot \rho(x,y) \qquad (x, y \in X)$$

feltételnek eleget tevő

$$f : X \to X$$

függvényt *kontrakciónak,* az

$$f(\alpha) = \alpha$$

egyenlőséget kielégítő α elemet pedig az f *fixpontjának* nevezzük.

v) A fixpont-tétel a) és b) pontja egyúttal hatékony algoritmust is kínál a fixpont közelítő meghatározására. Jegyezzük meg, hogy a c) hibabecslő formulából tetszőleges $0 < m \in \mathbf{N}$ esetén

$$\rho(x_m, \alpha) \le \frac{q}{1-q} \cdot \rho(x_m, x_{m-1})$$

következik. Alkalmazzuk ui. az említett becslést az

$$y_n := x_{m-1+n} \qquad (n \in \mathbf{N})$$

sorozatra az $n := 1$ választással.

vi) Felhívjuk a figyelmet arra, hogy a fixpont-tételben az f leképezésre vonatkozó, a $0 \le q < 1$ konstanssal teljesülő

$$\rho\Big(f(x), f(y)\Big) \le q\cdot \rho(x,y) \qquad (x, y \in X)$$

kontrakciós feltétel nem helyettesíthető az alábbival:

$$\rho\Big(f(x), f(y)\Big) < \rho(x,y) \qquad (x, y \in X, x \ne y).$$

Tekintsük ui. az

$$X := \mathbf{R}, \;\; \rho(x,y) := |x-y| \qquad (x, y \in \mathbf{R})$$

választással a (teljes) (X, ρ) metrikus teret és azt az

$$f : X \to X$$

leképezést, amelyre

$$f(x) := \ln\Big(1 + e^x\Big) \qquad (x \in \mathbf{R}).$$

Ekkor az f differenciálható függvény és

$$f'(x) = \frac{e^x}{1+e^x} < 1 \qquad (x \in \mathbf{R}).$$

Alkalmazva a Lagrange-féle középérték-tételt azt kapjuk, hogy tetszőleges $x, y \in \mathbf{R}$, $x < y$ esetén egy alkalmas $\xi \in (x, y)$ számmal

$$|f(x) - f(y)| = |f'(\xi)| \cdot |x - y| < |x - y|.$$

Ugyanakkor az f-nek nincs fixpontja. Ha ui. valamilyen $\alpha \in \mathbf{R}$ fixpontja lenne az f-nek, akkor az $\alpha = f(\alpha)$ egyenlőségből az

$$\ln\Big(1 + e^\alpha\Big) = \alpha,$$

vagy – ami ezzel ekvivalens – az

$$1 + e^\alpha = e^\alpha$$

egyenlőséghez jutnánk. Innen viszont $1 = 0$ következne, ami nem igaz.

5. Fejezet

1. Emlékeztető.

Tétel (fixpont-tétel). *Adott az (X, ρ) teljes metrikus tér, valamint az*

$$f : X \to X$$

leképezés. Tegyük fel, hogy valamilyen $0 \le q < 1$ együtthatóval

$$\rho\Big(f(x), f(y)\Big) \le q \cdot \rho(x, y) \qquad (x, y \in X).$$

Ekkor

i) *egyértelműen létezik olyan $\alpha \in X$, amire $f(\alpha) = \alpha$;*

ii) *bármilyen $x_0 \in X$ elemet véve az*

$$x_{n+1} := f(x_n) \qquad (n \in \mathbf{N})$$

rekurzióval definiált sorozat konvergens, és $\lim\,(x_n) = \alpha$;

iii) *igaz a következő becslés:*

$$\rho(x_n, \alpha) \le \frac{q^n}{1-q} \cdot \rho(x_0, x_1) \qquad (n \in \mathbf{N}).$$

Legyen pl.

$$X := [1, +\infty), \quad \rho(x, y) := |x - y| \qquad (x, y \in X),$$

ill.

$$f(x) := \frac{x}{2} + \frac{1}{x} \qquad (x, y \in X).$$

Ekkor az (X, ρ) teljes metrikus tér, az

$$f : X \to X$$

leképezés kontrakció:

$$\rho(f(x), f(y)) = \left|\frac{x}{2} - \frac{y}{2} + \frac{1}{x} - \frac{1}{y}\right| = \left|\frac{x-y}{2} - \frac{x-y}{xy}\right| =$$

$$|x - y| \cdot \left|\frac{1}{2} - \frac{1}{xy}\right| \le \frac{|x-y|}{2} = \frac{\rho(x, y)}{2} \qquad (x, y \in X)$$

és

$$x_0 := 2, \quad x_{n+1} := f(x_n) = \frac{x_n}{2} + \frac{1}{x_n} \qquad (n \in \mathbf{N}).$$

Az így definiált (x_n) (racionális) számsorozatnak van határértéke és az $\sqrt{2}$. Ezért

$$x_1 = \frac{3}{2}$$

miatt tetszőleges $n \in \mathbf{N}$ indexszel

$$|x_n - \sqrt{2}| \le \frac{(1/2)^n}{1/2} \cdot \frac{1}{2} = \frac{1}{2^n} \qquad (n \in \mathbf{N}),$$

ahol az itteni x_n-ek valamennyien racionális számok. Pl. x_{10}-re igaz az

$$|x_{10} - \sqrt{2}| \le \frac{1}{2^{10}} < \frac{1}{1000}$$

hibabecslés.[22]

[22] Ha $\varepsilon_n := |x_n - \sqrt{2}|/2$, akkor $\varepsilon_{n+1} \le \varepsilon_n^2$, így $\varepsilon_n \le \varepsilon_0^{2^n} < (1/2)^{2^n}$ $(n \in \mathbf{N})$. Ezért valójában már $|x_4 - \sqrt{2}| \le (1/2)^{15} < 10^{-4}$ is fennáll. Sőt, egyszerű számolás után azt kapjuk, hogy $x_4 = 665857/470832 = 1,414213562374...$, ahol (az első 65 tizedesjegyre) $\sqrt{2} = 1,41421356237309504880168872420969807856967187537694807317667973799...$

2. Megjegyzések.

i) Pl. a $(C[-1,1], \|.\|_1)$ normált tér[23] nem teljes: létezik olyan

$$(f_n) : \mathbf{N} \to C[-1,1]$$

Cauchy-sorozat, amelyik nem konvergens. Tehát tetszőleges $\varepsilon > 0$ számhoz van olyan $N \in \mathbf{N}$, amellyel

$$\|f_n - f_m\|_1 = \int_{-1}^{1} |f_n - f_m| < \varepsilon \qquad (n, m \in \mathbf{N},\ n, m > N),$$

de ugyanakkor nem létezik olyan $f \in C[-1,1]$ függvény, hogy

$$\|f_n - f\|_1 = \int_{-1}^{1} |f_n - f| \to 0 \qquad (n \to \infty)$$

teljesülne. Ilyen sorozat pl. a következő:

$$f_n(x) := \begin{cases} 0 & (-1 \le x \le 0) \\ nx & (0 \le x \le 1/n) \\ 1 & (1/n \le x \le 1) \end{cases} \qquad (1 \le n \in \mathbf{N}).$$

Ez Cauchy-sorozat. Ha ui. $0 < n, m \in \mathbf{N}$ és $n < m$, akkor

$$\|f_n - f_m\|_1 = \int_0^{1/m} (m-n)x\,dx + \int_{1/m}^{1/n} (1 - nx)\,dx =$$

$$\frac{m-n}{2m^2} + \frac{1}{n} - \frac{1}{m} - \frac{n}{2} \cdot \left(\frac{1}{n^2} - \frac{1}{m^2}\right) < \frac{1}{n}.$$

Legyen $\varepsilon > 0$ tetszőleges, ekkor az $1/N < \varepsilon$ egyenlőtlenségnek eleget tevő akármelyik $N \in \mathbf{N}$ indexszel

$$\|f_n - f_m\|_1 < \frac{1}{n} < \frac{1}{N} < \varepsilon \qquad (n, m \in \mathbf{N},\ m > n > N).$$

Tegyük fel indirekt módon, hogy van olyan $f \in C[-1,1]$ függvény, amellyel

$$\|f_n - f\|_1 \to 0 \qquad (n \to \infty).$$

[23] Ha $f \in C[-1,1]$, akkor $\|f\|_1 = \int_{-1}^{1} |f|$.

Belátjuk, hogy ekkor szükségszerűen

$$f(\xi) = 0, \quad f(\eta) = 1 \qquad (-1 < \xi < 0 < \eta < 1).$$

Világos, hogy ez a függvény nem folytonos a 0-ban, ezért $f \in C[-1,1]$ nem lehet igaz.[24] Az $f(\xi) = 0$ egyenlőséget indirekt úton igazoljuk. Ha (pl.) $f(\xi) > 0$, akkor $f \in \mathcal{C}\{\xi\}$ miatt van olyan $\delta > 0$ szám, hogy

$$[\xi - \delta, \xi + \delta] \subset (-1, 0)$$

és

$$f(t) > \frac{f(\xi)}{2} \qquad (\xi - \delta \le t \le \xi + \delta).$$

Ezért tetszőleges $1 \le n \in \mathbf{N}$ mellett

$$\|f_n - f\|_1 = \int_{-1}^{1} |f_n - f| \ge \int_{\xi-\delta}^{\xi+\delta} |f_n - f| = \int_{\xi-\delta}^{\xi+\delta} |f| \ge$$

$$\ge \int_{\xi-\delta}^{\xi+\delta} \frac{f(\xi)}{2} = 2\delta \cdot \frac{f(\xi)}{2} = \delta \cdot f(\xi) \ (> 0).$$

Innen nyilvánvaló, hogy

$$\|f_n - f\|_1 \to 0 \qquad (n \to \infty)$$

nem állhat fenn.

Hasonló meggondolással kapjuk az $f(\eta) = 1$ $(0 < \eta < 1)$ egyenlőséget is.

ii) Kiderült, hogy folytonos függvényekből álló *egyenletesen konvergens* függvénysorozat határfüggvénye is folytonos. Az előző megjegyzésben szereplő $f_n \in C[-1,1]$ $(1 \le n \in \mathbf{N})$ függvénysorozatra viszont egyszerűen adódik, hogy létezik a pontonkénti

$$f(x) := \lim_{n\to\infty} f_n(x) = \begin{cases} 0 & (-1 \le x \le 0) \\ \\ 1 & (0 < x \le 1) \end{cases}$$

határfüggvény, viszont $f \notin \mathcal{C}\{0\}$. Ez a példa azt mutatja, hogy a folytonosság minden további nélkül nem „öröklődik" egy pontonként konvergens, folytonos függvényekből álló függvénysorozat határfüggvényére.

[24] Vigyázat: erre az f függvényre (az $f(0) := 0$ kiegészítéssel) triviális módon fennáll a pontonkénti konvergencia: minden $x \in [-1,1]$ helyen $|f_n(x) - f(x)| \to 0$ $(n \to \infty)$. Ez utóbbira azonban nem hivatkozhatunk a várt ellentmondást illetően, ui. csak az (indirekt) $\|f_n - f\|_1 \to 0$ $(n \to \infty)$ (norma-konvergencia) feltevésre támaszkodhatunk.

iii) Egy érdekes tulajdonsága a Cauchy-sorozatoknak a következő állítás. Legyen az (X, ρ) tetszőleges metrikus tér, és jelöljük $\mathbf{C}$-vel az X-beli Cauchy-sorozatok halmazát. Ekkor bármely két

$$x = (x_n), y = (y_n) \in \mathbf{C}$$

sorozat esetén a $(\rho(x_n, y_n))$ számsorozat konvergens.[25]

Valóban, ha $m, n \in \mathbf{N},$ akkor a háromszög-egyenlőtlenség kétszeri alkalmazásával

$$\rho(x_n, y_n) \leq \rho(x_n, x_m) + \rho(x_m, y_n) \leq$$

$$\rho(x_n, x_m) + \rho(x_m, y_m) + \rho(y_m, y_n),$$

azaz

$$\rho(x_n, y_n) - \rho(x_m, y_m) \leq \rho(x_n, x_m) + \rho(y_m, y_n)$$

adódik. Innen (m és n felcserélésével)

$$\rho(x_m, y_m) - \rho(x_n, y_n) \leq \rho(x_n, x_m) + \rho(y_m, y_n),$$

tehát

$$|\rho(x_m, y_m) - \rho(x_n, y_n)| \leq \rho(x_n, x_m) + \rho(y_m, y_n)$$

következik. Tudjuk, hogy bármilyen $\varepsilon > 0$ számhoz megadhatók olyan $N, M \in \mathbf{N}$ indexek, hogy $m, n \in \mathbf{N}, m, n > N,$ ill. $m, n > M$ esetén

$$\rho(x_n, x_m) < \varepsilon/2, \text{ ill. } \rho(y_n, y_m) < \varepsilon/2.$$

Így az

$$m, n \in \mathbf{N}, m, n > \max\{M, N\}$$

választással

$$|\rho(x_m, y_m) - \rho(x_n, y_n)| < \varepsilon.$$

Ez éppen azt jelenti, hogy a

$$\mathbf{N} \ni k \mapsto \rho(x_k, y_k)$$

számsorozat Cauchy-sorozat, azaz konvergens.

---

[25] $x \sim y$, ha $\lim(\rho(x_n, y_n)) = 0$. Ekkor a $\sim$ reláció ekvivalencia, az $\widehat{x} := \{y \in \mathbf{C} : x \sim y\}$ ekvivalenciaosztályok $\widehat{\mathbf{C}}$ halmaza pedig a $\widehat{\rho}(\widehat{u}, \widehat{v}) := \lim(\rho(u_n, v_n))$ $(\widehat{u}, \widehat{v} \in \widehat{\mathbf{C}})$ metrikával teljes metrikus tér. (Ha itt $x \in \widehat{u}$, $y \in \widehat{v}$, akkor $\lim(\rho(x_n, y_n)) = \lim(\rho(u_n, v_n))$.) (Felix Hausdorff (1868-1942) tétele). Az $X := \mathbf{Q}$, $\rho(r, s) := |r - s|$ $(r, s \in \mathbf{Q})$ speciális esetben a $(\widehat{\mathbf{C}}, \widehat{\rho})$ tér a valós számok modelljéül szolgálhat. Így pl. a $\sqrt{2}$-t az emlékeztetőben szereplő (x_n) racionális számsorozat reprezentálja.

iv) Bizonyítsuk be a fixpont-tétel következő változatát: egy (X, ρ) teljes metrikus tér és valamilyen $a \in X$, $0 < r \in \mathbf{R}$ esetén legyen

$$Y := \{x \in X : \rho(x, a) \leq r\}.$$

Tegyük fel, hogy az

$$f \in X \to X$$

leképezésre az alábbi feltételek teljesülnek:

- $Y \subset \mathcal{D}_f$;
- van olyan $0 \leq q < 1$ szám, amellyel minden $x, y \in Y$ mellett

$$\rho\Big(f(x), f(y)\Big) \leq q \cdot \rho(x, y);$$

- $\rho\Big(a, f(a)\Big) \leq (1 - q) r.$

Ekkor egyértelműen létezik egy $\alpha \in Y$ úgy, hogy

a) $f(\alpha) = \alpha$;

b) az

$$x_0 \in Y,\ x_{n+1} := f(x_n) \qquad (n \in \mathbf{N})$$

rekurzióval értelmezett (x_n) sorozat konvergens, $\lim(x_n) = \alpha$, és

$$\rho(x_n, \alpha) \leq \frac{q^n}{1 - q} \cdot \rho\Big(x_0, f(x_0)\Big) \qquad (n \in \mathbf{N}).$$

Tekintsük ui. azt az (Y, σ) metrikus teret, amelyre

$$\sigma(x, y) := \rho(x, y) \qquad (x, y \in Y)$$

(az (X, ρ) (metrikus) „altere"). Legyen

$$F(x) := f(x) \qquad (x \in Y).$$

Lássuk be először is azt, hogy

$$F : Y \to Y,$$

azaz tetszőleges $x \in Y$ esetén $F(x) \in Y$. Ui.

$$\rho\Big(F(x), a\Big) \leq \rho\Big(F(x), F(a)\Big) + \rho\Big(F(a), a\Big) =$$

$$= \rho\Big(f(x), f(a)\Big) + \rho\Big(f(a), a\Big) \leq$$
$$q\cdot \rho(x,a) + (1-q)r \leq qr + (1-q)r = r.$$

Világos, hogy az F leképezés kontrakció, mivel

$$\sigma\Big(F(x), F(y)\Big) = \rho\Big(f(x), f(y)\Big) \leq q\cdot \rho(x,y) = q\cdot \sigma(x,y) \quad (x, y \in Y).$$

Az (Y, σ) tér teljes. Legyen ehhez adott az

$$(x_n) : \mathbf{N} \to Y$$

Cauchy-sorozat. Következésképpen tetszőleges $\varepsilon > 0$ számhoz van olyan $N \in \mathbf{N}$, hogy

$$\sigma(x_n, x_m) = \rho(x_n, x_m) < \varepsilon \qquad (n, m \in \mathbf{N},\ n, m > N).$$

Mindez azt jelenti, hogy az (x_n) sorozat egyúttal mint

$$(x_n) : \mathbf{N} \to X$$

is Cauchy-sorozat (másképp mondva a ρ metrika „szerint" is Cauchy-sorozat). Az (X, ρ) tér feltételezett teljessége miatt tehát létezik olyan $\alpha \in X$, amellyel

$$\rho(x_n, \alpha) \to 0 \qquad (n \to \infty).$$

Megmutatjuk, hogy $\alpha \in Y$ igaz.[26] Tegyük fel ehhez indirekt úton, hogy $\alpha \notin Y$, azaz

$$\rho(a, \alpha) > r.$$

Ekkor $\lim(x_n) = \alpha$ miatt van olyan $n \in \mathbf{N}$, amellyel

$$\rho(x_n, \alpha) < \rho(a, \alpha) - r.$$

Ezért

$$\rho(x_n, a) \geq \rho(a, \alpha) - \rho(x_n, \alpha) > \rho(a, \alpha) - (\rho(a, \alpha) - r) = r,$$

ami nyilván ellentmond annak, hogy $x_n \in Y$.

Tehát $\alpha \in Y$, és ezért

$$\sigma(x_n, \alpha) = \rho(x_n, \alpha) \to 0 \qquad (n \to \infty).$$

Így az (x_n) sorozat a σ metrika értelmében is konvergál α-hoz.

Alkalmazható ezért a fixpont-tétel az (Y, σ) teljes metrikus térre és az

$$F : Y \to Y$$

kontrakcióra, aminek az „eredménye" éppen a bizonyítandó állításunk.

[26] Ez egyébként abból is adódik, hogy az Y halmaz zárt.

v) Legyen $1 \leq n \in \mathbf{N}$ és $1 \leq p \leq +\infty$. Ekkor a $\mathbf{K}^n$-beli $\|.\|_p$ normákra igaz az alábbi határérték-egyenlőség:

$$\|x\|_\infty = \lim_{p \to +\infty} \|x\|_p \qquad (x \in \mathbf{K}^n).$$

Valóban, tetszőleges $1 \leq p < +\infty$ és $x = (x_1, ..., x_n) \in \mathbf{K}^n$ esetén

$$\|x\|_\infty \leq \|x\|_p = \left(\sum_{k=1}^n |x_k|^p \right)^{1/p} \leq \left(\sum_{k=1}^n \|x\|_\infty \right)^{1/p} = n^{1/p} \cdot \|x\|_\infty.$$

Ismert, hogy ha $q > 0$ tetszőlegesen adott pozitív szám, akkor

$$\lim_{p \to +\infty} q^{1/p} = 1,$$

speciálisan

$$\lim_{p \to +\infty} n^{1/p} = 1.$$

Ezért

$$n^{1/p} \cdot \|x\|_\infty \to \|x\|_\infty \qquad (p \to +\infty).$$

Innen a közrefogási elv alapján az állításunk már következik.

2. Kompakt halmazok.

Induljunk ki valamilyen (X, ρ) metrikus térből, és legyen $\emptyset \neq A \subset X$. Azt mondjuk, hogy az A halmaz *kompakt,* ha tetszőleges

$$(x_n) : \mathbf{N} \to A$$

sorozathoz létezik olyan (ν_n) indexsorozat, hogy az (x_{ν_n}) részsorozat konvergens és

$$\lim(x_{\nu_n}) \in A.$$

Világos, hogy ha pl. az $A \subset X$ véges halmaz, akkor az A kompakt. Valóban, ha

$$(x_n) : \mathbf{N} \to A,$$

akkor az A végessége miatt léteznie kell olyan $a \in A$ elemnek és olyan (ν_n) indexsorozatnak, hogy

$$x_{\nu_n} = a \qquad (n \in \mathbf{N}).$$

Következésképpen az (x_{ν_n}) konstanssorozat konvergens és

$$\lim(x_{\nu_n}) = a \in A.$$

Nem nehéz belátni, hogy minden kompakt halmaz egyúttal zárt is. Ui. a kompaktság előbbi definíciójában szereplő tetszőleges

$$(x_n) : \mathbf{N} \to A$$

A-beli sorozatok között ott vannak az A-beli konvergens sorozatok is. Ha tehát az A kompakt, akkor minden A-beli konvergens (x_n) sorozatnak is kell, hogy legyen olyan (x_{ν_n}) konvergens részsorozata, amelyre

$$\lim(x_n) = \lim(x_{\nu_n}) \in A.$$

Tudjuk, hogy ez viszont azt jelenti, hogy az A halmaz zárt.

Az $A \subset X$ halmaz *korlátos,* ha megadható olyan $\xi \in X$ elem és annak olyan $K(\xi)$ környezete, hogy

$$A \subset K(\xi).$$

Mutassuk meg, hogy minden kompakt halmaz korlátos. Indirekt módon gondolkozva ui. tegyük fel, hogy az $\emptyset \neq A \subset X$ halmaz kompakt, de nem korlátos. Legyen $\xi \in X$ tetszőleges, ekkor az indirekt feltételezésünk miatt

$$A \not\subset K_n(\xi) \qquad (1 \leq n \in \mathbf{N}).$$

Ezért létezik olyan

$$(x_n) : \mathbf{N} \to A$$

sorozat, amelyre

$$\rho(x_n, \xi) \geq n \qquad (1 \leq n \in \mathbf{N}).$$

Egyúttal bármilyen (ν_n) indexsorozatra

$$\rho(x_{\nu_n}, \xi) \geq \nu_n \geq n \qquad (1 \leq n \in \mathbf{N})$$

is igaz. Az A kompaktsága miatt viszont van olyan (ν_n) indexsorozat is, amelyre az (x_{ν_n}) konvergens, így korlátos is: alkalmas $\eta \in X$ elem és $r > 0$ szám mellett $\mathcal{R}_{(x_{\nu_n})} \subset K_r(\eta)$, azaz

$$\rho(x_{\nu_n}, \eta) < r \qquad (1 \leq n \in \mathbf{N}).$$

Következésképpen

$$n \leq \rho(x_{\nu_n}, \xi) \leq \rho(x_{\nu_n}, \eta) + \rho(\eta, \xi) < r + \rho(\eta, \xi) \qquad (1 \leq n \in \mathbf{N}),$$

ami nyilván nem lehetséges.

Ezzel beláttuk az alábbi állítást:

1.Tétel. *Bármilyen* (X, ρ) *metrikus tér és* $A \subset X$ *kompakt halmaz esetén az* A *korlátos és zárt.*

Megjegyezzük, hogy az előbbi tétel „megfordítása" nem teljesül: van olyan (X, ρ) metrikus tér és olyan $A \subset X$ halmaz, hogy az A korlátos és zárt, de nem kompakt.[27]

Ugyanakkor igaz a

2. Tétel. *Legyen* $1 \leq s \in \mathbf{N}$, $1 \leq p \leq +\infty$, *és tekintsük a* $(\mathbf{K}^s, \rho_p)$ *metrikus teret. Ekkor tetszőleges korlátos és zárt* $\emptyset \neq A \subset \mathbf{K}^s$ *halmaz kompakt.*

Bizonyítás. Azt kell megmutatnunk, hogy minden

$$x = (x_n) : \mathbf{N} \to A$$

sorozatnak van olyan konvergens (x_{ν_n}) részsorozata, amelyre

$$\lim(x_{\nu_n}) \in A.$$

A feltételezésünk szerint az A halmaz korlátos, ezért az (x_n) sorozat is korlátos. A Bolzano–Weierstrass-kiválasztási tétel alapján van olyan ν indexsorozat, hogy az

$$x \circ \nu : \mathbf{N} \to A$$

részsorozat konvergens. Mivel az A halmaz zárt, ezért

$$\lim(x \circ \nu) \in A,$$

azaz az A halmaz kompakt. ■

3. Folytonosság.

Legyenek adottak az (X, ρ), (Y, σ) metrikus terek, és tekintsük az

$$f \in X \to Y$$

függvényt. Azt mondjuk, hogy az f *folytonos az* $a \in \mathcal{D}_f$ *pontban,* ha minden $\varepsilon > 0$ számhoz létezik olyan $\delta > 0$, hogy

$$\sigma(f(x), f(a)) < \varepsilon \qquad (x \in \mathcal{D}_f,\ \rho(x, a) < \delta).$$

[27] Ha pl. $X := \ell_\infty$, $\rho := \rho_\infty$ és $A := \{(x_n) \in \ell_\infty : x_n \in \{0, 1\} \ \ (n \in \mathbf{N})\}$, akkor az A korlátos és zárt, de nem kompakt. Világos, hogy ha $x, y \in A, x \neq y$, akkor $\rho_\infty(x, y) = 1$.

Mindezt az $f \in \mathcal{C}\{a\}$ szimbólummal fogjuk jelölni.

Ha az előbbi definícióban megkövetelt tulajdonság bármely $a \in \mathcal{D}_f$ helyen igaz, akkor röviden azt mondjuk, hogy az f *folytonos*, és erre az $f \in \mathcal{C}$ jelölést használjuk.

Világos, hogy a leképezések folytonosságának ez a definíciója az elemi folytonosság-fogalom kiterjesztése absztrakt (metrikus terek közötti) függvényekre. Az is nyilvánvaló, hogy a pontbeli folytonosság egy ekvivalens megfogalmazása a környezetek „nyelvén" a következő: $f \in \mathcal{C}\{a\}$ akkor és csak akkor igaz, ha bármely $K(f(a)) \subset Y$ környezethez megadható olyan $K(a) \subset X$ környezet, hogy

$$f[K(a) \cap \mathcal{D}_f] \subset K(f(a)).$$

4. Megjegyzések.

i) Az ekvivalens metrikákkal kapcsolatban mondottakra utalva világos a következő. Ha az

$$(X, \rho),\ (X, \rho_*),\ (Y, \sigma),\ (Y, \sigma_*)$$

metrikus tereket illetően

$$\rho \sim \rho_* \text{ és } \sigma \sim \sigma_*,$$

akkor egy

$$f \in X \to Y$$

függvény valamilyen $a \in \mathcal{D}_f$ pontbeli folytonossága (speciálisan az f folytonossága) független attól, hogy az X-ben, ill. az Y-ban a ρ vagy a ρ_*, ill. a σ vagy a σ_* metrika szerint tekintjük az elemek távolságát.

ii) Tetszőleges (X, ρ), (Y, σ) metrikus terek esetén az

$$f : X \to Y$$

függvény akkor és csak akkor folytonos, ha minden $Z \subset Y$ nyílt halmazra az $f^{-1}[Z]$ őskép is nyílt.[28]

Legyen ui. először az f függvény folytonos, a $Z \subset Y$ halmaz pedig nyílt. Nyilván feltehető, hogy $Z \neq \emptyset$, hiszen

$$f^{-1}[\emptyset] = \emptyset$$

[28]Felhívjuk a figyelmet arra, hogy most $\mathcal{D}_f = X$.

és az $\emptyset$ nyílt. Ha vamilyen $a \in \mathcal{D}_f$ esetén $f(a) \in Z$ (azaz $a \in f^{-1}[Z]$), akkor a Z nyíltsága miatt egy alkalmas $K(f(a))$ környezettel

$$K(f(a)) \subset Z.$$

Mivel $f \in \mathcal{C}\{a\}$, ezért van olyan $K(a) \subset X$ környezet, amellyel

$$f[K(a) \cap \mathcal{D}_f] = f[K(a)] \subset K(f(a)),$$

más szóval

$$K(a) \subset f^{-1}[K(f(a))].$$

A $K(a)$ $(a \in \mathcal{D}_f)$ környezetek nyíltsága alapján azt mondhatjuk, hogy az

$$f^{-1}[Z] = \bigcup_{a \in f^{-1}[Z]} K(a)$$

halmaz nyílt.

Fordítva, ha azt tudjuk, hogy tetszőleges $Z \subset Y$ nyílt halmazra az $f^{-1}[Z]$ nyílt, akkor minden $a \in \mathcal{D}_f$ és $Z := K(f(a))$ esetén is az

$$f^{-1}[K(f(a))]$$

nyílt. Van tehát olyan $K(a)$ környezet, amelyre $K(a) \subset f^{-1}[K(f(a))]$, azaz

$$f[K(a)] = f[K(a) \cap \mathcal{D}_f] \subset K(f(a)).$$

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

iii) Tetszőleges (X, ρ) metrikus tér, vagy $(X, \|.\|)$ normált tér, vagy $(X, \langle . \rangle)$ euklideszi tér esetén minden $a \in X$ rögzített elemmel az

$$f(x) := \rho(x, a),\ g(x) := \|x - a\|,\ h(x) := \langle x, a \rangle \qquad (x \in X)$$

függvények folytonosak. Ugyanez mondható el az $(X, \|.\|)$ normált tér vonatkozásában speciálisan az

$$F(x) := \|x\| \qquad (x \in X)$$

(norma-)függvényről, nevezetesen, hogy az F folytonos.

6. Fejezet

1. Emlékeztető.

Legyenek adottak az (X, ρ), (Y, σ) metrikus terek. Ekkor az

$$f \in X \to Y$$

függvény folytonos az $a \in \mathcal{D}_f$ pontban, ha minden $\varepsilon > 0$ számhoz létezik olyan $\delta > 0$, hogy

$$\sigma\big(f(x), f(a)\big) < \varepsilon \qquad (x \in \mathcal{D}_f,\ \rho(x,a) < \delta).$$

Mindezt az $f \in \mathcal{C}\{a\}$ szimbólummal jelöljük. Ha az előbbi definícióban megkövetelt tulajdonság bármely $a \in \mathcal{D}_f$ helyen igaz, akkor röviden azt mondjuk, hogy az f folytonos, és erre az $f \in \mathcal{C}$ jelölést használjuk.

2. Folytonosság.

A továbbiakban

$$f \in X \to Y$$

függvényekre vonatkozóan fogalmazunk meg a folytonossággal kapcsolatos állításokat. Ezeknek a bizonyítását nem részletezzük, mert azok "egy az egyben" ugyanúgy végezhetők, mint az $f \in \mathbf{R} \to \mathbf{R}$ esetben, csupán az aktuális

$$u, v, f(u), f(v) \in \mathbf{R}$$

számok távolsága (azaz a különbségük abszolút értéke) helyett kell rendre az

$$u, v \in X, f(u), f(v) \in Y$$

elemek távolságát, azaz $\rho(u,v)$-t, ill. $\sigma(f(u), f(v))$-t $(u, v \in \mathcal{D}_f)$ írni.

1. Tétel (átviteli elv). *Legyen $a \in \mathcal{D}_f$. Ekkor $f \in \mathcal{C}\{a\}$ azzal ekvivalens, hogy tetszőleges*

$$(x_n) : \mathbf{N} \to \mathcal{D}_f,\ \lim(x_n) = a$$

sorozatra az $\big(f(x_n)\big)$ sorozat konvergens és

$$\lim\big(f(x_n)\big) = f(a).$$

2. Tétel. *Az eddigi* (X, ρ), (Y, σ) *metrikus terek mellett legyen adott a* (Z, δ) *metrikus tér is, és tegyük fel, hogy a*

$$g \in X \to Y, \ f \in Y \to Z$$

függvényekre a következők teljesülnek:

$$a \in \mathcal{D}_g, \ g(a) \in \mathcal{D}_f$$

és

$$g \in \mathcal{C}\{a\}, \ f \in \mathcal{C}\{g(a)\}.$$

Ekkor az

$$f \circ g \in X \to Z$$

függvény folytonos az a*-ban, azaz* $f \circ g \in \mathcal{C}\{a\}$.

3. Tétel (Weierstrass[29]). *Ha az* $f \in X \to Y$ *függvény folytonos és a* $\mathcal{D}_f$ *értelmezési tartomány kompakt, akkor az* $\mathcal{R}_f$ *értékészlet is kompakt. Speciálisan, ha itt*

$$f \in X \to \mathbf{R},$$

akkor az $\mathcal{R}_f$ *értékészletnek van legkisebb és legnagyobb eleme.*

4. Tétel. *Legyen az* $f \in X \to Y$ *függvény folytonos, invertálható, és az értelmezési tartománya kompakt. Ekkor az* f^{-1} *inverzfüggvény is folytonos.*

A következő tétel megfogalmazásához vezessük be az egyenletes folytonosság fogalmát: az

$$f \in X \to Y$$

függvény *egyenletesen folytonos,* ha bármilyen $\varepsilon > 0$ esetén létezik olyan $\delta > 0$, hogy

$$\sigma\Big(f(x), f(t)\Big) < \varepsilon \qquad (x, t \in \mathcal{D}_f, \ \rho(x, t) < \delta).$$

Világos, hogy ekkor az f folytonos is, de mindez „megfordítva" nem igaz. Ugyanakkor fennáll az alábbi tétel:

[29]Karl Theodor Wilhelm Weierstrass (1815-1897).

5. Tétel (Heine [30]). *Tegyük fel, hogy az* $f \in X \to Y$ *függvény folytonos, és a* $\mathcal{D}_f$ *halmaz kompakt. Ekkor az* f *egyenletesen folytonos.*

3. Többváltozós vektorfüggvények.

A továbbiakban adott $1 \leq s,\, m \in \mathbf{N}$ mellett tekintsünk egy

$$f \in \mathbf{R}^s \to \mathbf{R}^m$$

függvényt. Ha $s > 1$, akkor az f-et *többváltozós függvénynek* is szokás nevezni. Hasonlóan, ha $m > 1$, akkor az f egy ún. *vektorfüggvény.*[31] Így $s, m > 1$ esetén a szóban forgó f függvény egy *többváltozós vektorfüggvény.*

Legyen $x \in \mathcal{D}_f$ és

$$(y_1, ..., y_m) := f(x).$$

Ha $i = 1, ..., m,$ akkor (az előbbi jelölésekkel) értelmezzük az

$$f_i : \mathcal{D}_f \to \mathbf{R}$$

függvényt a következőképpen:

$$f_i(x) := y_i \qquad (x \in \mathcal{D}_f).$$

Az f_i $(i = 1, ..., m)$ függvényt az f függvény i-edik *koordináta-függvényének* nevezzük, és minderre az

$$f = (f_1, ..., f_m)$$

jelölést alkalmazzuk. Tehát

$$f_i \in \mathbf{R}^s \to \mathbf{R} \qquad (i = 1, ..., m),$$

---

[30]Heinrich Eduard Heine (1821-1881).

[31]Speciálisan, az $f \in \mathbf{R} \to \mathbf{R}^m$ egyváltozós vektorfüggvények fontos szerepet játszanak az $\mathbf{R}^m$-beli görbék vizsgálatában: bizonyos feltételek mellett az ilyen f függvények $\mathcal{R}_f$ értékkészlethalmazát nevezzük *görbének* (magát az f függvényt pedig az illető görbe (egy) *paraméterezésének*). Pl. adott $F \in \mathbf{R} \to \mathbf{R}$ egyváltozós valós függvénnyel legyen $f(x) := (x, F(x)) \in \mathbf{R}^2 \quad (x \in \mathcal{D}_F)$. Ekkor (az F-re vonatkozó alkalmas megkötésekkel) az $\mathcal{R}_f$ értékkészlethalmaz (ezt szokás az F grafikonjaként is emlegetni) egy $\mathbf{R}^2$-beli görbe („függvénygörbe"). Az elemi matematikából is jól ismert görbe a körvonal, mint pl. az $f(x) := (\cos x, \sin x) \in \mathbf{R}^2 \quad (x \in [0, 2\pi])$ előírással megadott $f : [0, 2\pi] \to \mathbf{R}^2$ függvény értékkészlete. Hasonlóan, ha tekintjük az $u, v \in \mathbf{R}^m$, $u \neq v$ vektorokat, akkor az $f(x) := u + x(v - u) \quad (x \in [0, 1])$ paraméterezéssel megadott görbe az u, v végpontú *m-dimenziós szakasz.* Ha pl. valamilyen $1 < m, n \in \mathbf{N}$ és a kompakt $I \subset \mathbf{R}$ intervallum mellett adottak a $v_k \in \mathbf{R}^m$ vektorok és a $g_k : I \to \mathbf{R} \quad (k = 1, ..., n)$ függvények, akkor egy (az alkalmazásokban is gyakran alkalmazott) „görbeszerkesztési eljárás" a következő: $f(x) := \sum_{k=1}^{n} g_k(x) v_k \quad (x \in I)$.

azaz $s > 1$ esetén a koordináta-függvények többváltozós valós értékű függvények.

Az $1 \leq p, q \leq +\infty$ paraméterekkel az $\mathbf{R}^s$-ben a ρ_p, az $\mathbf{R}^m$-ben pedig a ρ_q metrikát fogjuk használni, azaz legyen

$$(X, \rho) := (\mathbf{R}^s, \rho_p), \ (Y, \sigma) := (\mathbf{R}^m, \rho_q).$$

Ekkor a fentiek szerint vizsgálhatjuk az

$$f \in \mathbf{R}^s \to \mathbf{R}^m$$

függvények – speciálisan tehát az

$$f_i \in \mathbf{R}^s \to \mathbf{R} \qquad (i = 1, ..., m)$$

koordináta-függvények – folytonosságát. Ebből a szempontból alapvető az alábbi állítás.

6. Tétel. *Az*

$$f = (f_1, ..., f_m) \in \mathbf{R}^s \to \mathbf{R}^m$$

függvény akkor és csak akkor folytonos valamilyen $a \in \mathcal{D}_f$ helyen, ha minden $i = 1, ..., m$ esetén az f_i koordináta-függvény folytonos az a-ban. Speciálisan az f akkor és csak akkor folytonos, ha minden koordináta-függvénye folytonos.

Bizonyítás. Legyen $a \in \mathcal{D}_f$. Azt kell megmutatnunk, hogy az $f \in \mathcal{C}\{a\}$ folytonosság akkor és csak akkor igaz, ha

$$f_i \in \mathcal{C}\{a\} \qquad (i = 1, ..., m).$$

Az átviteli elv szerint mindez a következőkkel egyenértékű: tetszőleges

$$(x_n) : \mathbf{N} \to \mathcal{D}_f, \ \lim(x_n) = a$$

sorozatra

$$\lim\Big(f(x_n)\Big) = f(a)$$

akkor és csak akkor teljesül, ha

$$\lim\Big(f_i(x_n)\Big) = f_i(a) \qquad (i = 1, ..., m).$$

Vegyük észre, hogy az $\Big(f_i(x_n)\Big)$ $(i = 1, ..., m)$ sorozat nem más, mint az $\Big(f(x_n)\Big)$ sorozat i-edik koordináta-sorozata, $f_i(a)$ pedig az $f(a)$ helyettesítési érték i-edik koordinátája. Ezért az állításunk a vektorsorozatok konvergenciájára vonatkozó tétel egyszerű következménye. ■

4. Megjegyzések.

i) Az $f \in \mathbf{R} \to \mathbf{R}$ függvények határértékének a fogalmát könnyen kiterjeszthetjük absztrakt függvényekre. Legyenek ehhez (X, ρ), (Y, σ) metrikus terek, és tekintsünk egy

$$f \in X \to Y$$

függvényt. Tegyük fel, hogy $a \in \mathcal{D}'_f$, azaz az a torlódási pontja a $\mathcal{D}_f$-nek. Ekkor azt mondjuk, hogy *az f függvénynek az a helyen* (vagy *az a-ban*) *van határértéke,* ha létezik olyan $A \in Y$, amelyre tetszőleges $K(A) \subset Y$ környezetet véve megadható az a-nak olyan $K(a) \subset X$ környezete, amellyel

$$f\Big[\Big(K(a) \setminus \{a\}\Big) \cap \mathcal{D}_f\Big] \subset K(A)$$

teljesül. Más szóval tehát

$$f(x) \in K(A) \qquad (a \neq x \in K(a) \cap \mathcal{D}_f).$$

Ekkor egyértelműen létezik az az $A \in Y$, amely eleget tesz a fentieknek. Ezt az $A \in Y$ elemet az *f függvény a helyen vett* (vagy *a-beli*) *határértékének* nevezzük. Az alábbi jelöléseket fogjuk használni:

$$\lim_a f := \lim_{x \to a} f(x) := A.$$

ii) Más megfogalmazásban

$$\lim_a f = A$$

azt jelenti, hogy minden $\varepsilon > 0$ számhoz létezik olyan $\delta > 0$, hogy

$$\sigma\Big(f(x), A\Big) < \varepsilon \qquad (a \neq x \in \mathcal{D}_f,\ \rho(x, a) < \delta).$$

iii) A definícióból eléggé nyilvánvaló, hogy ha

$$f, g \in X \to Y$$

és valamilyen $a \in X$ esetén van olyan $K(a) \subset X$ környezet, hogy

$$\emptyset \neq \mathcal{D} := \Big(K(a) \setminus \{a\}\Big) \cap \mathcal{D}_f = \Big(K(a) \setminus \{a\}\Big) \cap \mathcal{D}_g,$$

valamint

$$f(x) = g(x) \qquad (x \in \mathcal{D}),$$

akkor – feltéve, hogy $a \in \mathcal{D}'_f \cap \mathcal{D}'_g$ – a következő mondható: az f függvénynek akkor és csak akkor van határértéke az a helyen, ha a g függvénynek is van az a-ban határértéke, és az utóbbi esetben

$$\lim_a f = \lim_a g.$$

Röviden azt mondjuk, hogy a $\lim_a f$ létezése vagy nem létezése az illető függvény *lokális tulajdonsága.*

Speciálisan azt kapjuk, hogy a $\lim_a f$ létezése és „milyensége" szempontjából érdektelen az $a \in \mathcal{D}_f$, vagy $a \notin \mathcal{D}_f$ kérdés, ill. az $a \in \mathcal{D}_f$ esetben az $f(a)$ helyettesítési érték. Ugyanakkor világos, hogy

$$a \in \mathcal{D}_f \cap \mathcal{D}'_f$$

esetén az f folytonossága az a-ban azzal ekvivalens, hogy létezik a $\lim_a f$ határérték és

$$\lim_a f = f(a).$$

iv) (*Átviteli elv.*) Bármilyen

$$f \in X \to Y$$

függvény és $a \in \mathcal{D}'_f$ esetén az f-nek az a helyen akkor és csak akkor létezik határértéke, ha tetszőleges

$$(x_n) : \mathbf{N} \to \mathcal{D}_f \setminus \{a\},\ \lim(x_n) = a$$

sorozatra az $\big(f(x_n)\big)$ sorozatnak van határértéke. Ha létezik az

$$A := \lim_a f$$

határérték, akkor az előbbiekben említett minden (x_n) sorozatra

$$\lim\big(f(x_n)\big) = A.$$

v) Legyen pl. $1 \le s,\ m \in \mathbf{N},\ 1 \le p,\ q \le +\infty,$ és vegyük az

$$(\mathbf{R}^s, \rho_p),\ (\mathbf{R}^m, \rho_q)$$

metrikus tereket, továbbá az

$$f = (f_1, ..., f_m) \in \mathbf{R}^s \to \mathbf{R}^m$$

függvényt és az $a \in \mathcal{D}'_f$ pontot. Ekkor az f függvénynek az a helyen akkor és csak akkor létezik határértéke, ha minden $i = 1, ..., m$ esetén létezik az

$$A_i := \lim_a f_i$$

határérték. Az utóbbi esetben

$$\lim_a f = (A_1, ..., A_m).$$

vi) Az approximációelmélet egyik alapkérdése a következő: legyen adott az (X, ρ) metrikus tér, $\emptyset \neq A \subset X$, és tekintsük az

$$f(x) := \inf\{\rho(x, a) : a \in A\} \qquad (x \in X)$$

függvényt. A kérdés az, hogy (adott $x \in X$ elemre) írható-e az előbbi „inf" helyett „min", azaz van-e olyan $a^* \in A$, amelyre

$$f(x) = \rho(x, a^*).$$

Ha igen, akkor az ilyen a^* elemet *extremális elemnek* nevezzük.[32] Mutassuk meg, hogy az f függvény egyenletesen folytonos. Valóban, ha $x, y \in X$, akkor tetszőleges $a \in A$ esetén

$$f(x) \leq \rho(x, a) \leq \rho(x, y) + \rho(y, a),$$

amiből az

$$f(x) \leq \rho(x, y) + f(y)$$

egyenlőtlenség triviálisan következik. Így

$$f(x) - f(y) \leq \rho(x, y),$$

ill. x-et és y-t felcserélve

$$f(y) - f(x) \leq \rho(x, y),$$

más szóval

$$|f(x) - f(y)| \leq \rho(x, y) \qquad (x, y \in X).$$

Innen a mondott állítás már nyilvánvaló.

[32] Ha pl. $(X, \rho) := (C[a, b], \rho_\infty)$, és valamilyen $n \in \mathbf{N}$ mellett A a legfeljebb n-ed fokú polinomok $[a, b]$-re vett leszűkítéseinek a halmaza, akkor tetszőleges $g \in C[a, b]$ függvényhez egyértelműen létezik az előbbi értelemben vett extremális elem: olyan P legfeljebb n-ed fokú polinom, amelyre $\|g - P\|_\infty = \min_Q \|g - Q\|_\infty$, ahol az itteni minimum a legfeljebb n-ed fokú Q polinomokra vonatkozik (Borel-tétel). Ezt a P-t a g-t *egyenletesen legjobban közelítő legfeljebb n-ed fokú polinomnak* nevezik.

5. Lineáris leképezések.

Az $(X, \|.\|_\bullet)$, $(Y, \|.\|_*)$ normált terek közötti folytonos leképezések[33] között különös fontosságúak az ún. korlátos lineáris leképezések. Nevezetesen, az

$$f : X \to Y$$

függvényt *korlátos lineáris leképezésnek* nevezzük, ha

- az f *lineáris,* azaz

$$f(x + \lambda y) = f(x) + \lambda \cdot f(y) \qquad (x, y \in X,\ \lambda \in \mathbf{R})$$

  és

- van olyan $M \geq 0$ szám, hogy

$$\|f(x)\|_* \leq M \cdot \|x\|_\bullet \qquad (x \in X).$$

(Szokás minden ilyen M számot a szóban forgó f korlátos lineáris leképezés *korlátjaként* említeni.) Ekkor

$$\|f(x) - f(y)\|_* = \|f(x - y)\|_* \leq M \cdot \|x - y\|_\bullet \qquad (x, y \in X)$$

miatt minden korlátos lineáris leképezés egyenletesen folytonos.

Jelöljük a fenti korlátos lineáris leképezések halmazát az $L(X, Y)$ szimbólummal. Könnyen meggondolható, hogy a „szokásos" függvények közötti műveletekkel az $L(X, Y)$ halmaz is vektortér (lineáris tér) az $\mathbf{R}$ testre vonatkozóan. Sőt, ha $f \in L(X, Y)$ esetén

$$K_f := \{M \geq 0 : \|f(x)\|_* \leq M \cdot \|x\|_\bullet \ (x \in X)\},$$

akkor az

$$L(X, Y) \ni f \mapsto \|f\| := \inf K_f$$

leképezés norma (a $\|.\|_\bullet$, $\|.\|_*$ normák által *indukált norma*). Így az $(L(X, Y), \|.\|)$ normált tér.

Az is megmutatható, hogy tetszőleges $f \in L(X, Y)$ leképezésre létezik a K_f halmaznak minimuma, azaz

$$\|f\| = \min K_f.$$

[33]Az illető terekben az $X \times X \ni (x, y) \to \|x - y\|_\bullet$ és az $Y \times Y \ni (u, v) \to \|u - v\|_*$ normákat vezetve be.

Ui. lássuk be, hogy $\|f\| \in K_f$, azaz

$$\|f(x)\|_* \leq \|f\| \cdot \|x\|_\bullet \qquad (x \in X).$$

Indirekt gondolkodva tegyük fel, hogy valamilyen $z \in X$ mellett

$$\|f(z)\|_* > \|f\| \cdot \|z\|_\bullet.$$

Legyen ekkor a $q > 0$ szám olyan, hogy még

$$\|f(z)\|_* > (\|f\| + q) \cdot \|z\|_\bullet$$

is igaz legyen. (Ilyen q nyilván létezik.) Az infimum tulajdonságai alapján viszont van olyan

$$0 \leq M < \|f\| + q$$

szám, hogy $M \in K_f$, azaz

$$\|f(x)\|_* \leq M \cdot \|x\|_\bullet \qquad (x \in X).$$

Így

$$(\|f\| + q) \cdot \|z\|_\bullet < \|f(z)\|_* \leq M \cdot \|z\|_\bullet.$$

Ugyanakkor itt $z \neq 0$, hiszen

$$\|f(0)\|_* = \|0\|_* = 0 \not> \|f\| \cdot \|0\|_\bullet = 0.$$

Tehát $\|z\|_\bullet > 0$ és az előzőekből (egyszerűsítés után)

$$\|f\| + q < M < \|f\| + q$$

következik, ami nem igaz.

Ezzel beláttuk, hogy valóban $\|f\| \in K_f$, amiből az állításunk már következik. ■

Tehát

$$\|f(x)\|_* \leq \|f\| \cdot \|x\|_\bullet \qquad (x \in X),$$

az $\|f\|$ norma egyúttal az $f \in L(X, Y)$ korlátos lineáris leképezés legkisebb korlátja.

Legyen pl. valamilyen

$$1 \leq s,\, m \in \mathbf{N},\, 1 \leq p,\, q \leq +\infty$$

esetén

$$(X, \|.\|_\bullet) := (\mathbf{R}^s, \|.\|_p),\ (Y, \|.\|_*) := (\mathbf{R}^m, \|.\|_q),$$

és egy $A \in \mathbf{R}^{m\times s}$ mátrixszal

$$(*) \qquad f_A(x) := Ax \qquad (x \in \mathbf{R}^s).$$

Egyszerű számolással ellenőrizhető, hogy ha $p = q = 1$, akkor

$$f_A \in L(\mathbf{R}^s, \mathbf{R}^m).$$

Jegyezzük meg, hogy ezért, és a $\|.\|_p$, $\|.\|_q$ normák ekvivalenciája miatt az $\mathbf{R}^s$-en, ill. az $\mathbf{R}^m$-en bevezetett bármelyik $\|.\|_p$, $\|.\|_q$ norma esetén a fenti f_A leképezésre

$$f_A \in L(\mathbf{R}^s, \mathbf{R}^m).$$

Sőt, jól ismert a lineáris algebrából, hogy az $L(\mathbf{R}^s, \mathbf{R}^m)$ halmaz minden eleme „ilyen alakú", azaz tetszőleges $f \in L(\mathbf{R}^s, \mathbf{R}^m)$ függvényhez egyértelműen létezik olyan $A \in \mathbf{R}^{m\times s}$ mátrix, amellyel $(*)$ fennáll:

$$f = f_A.$$

Ennek alapján vezessük be az $\|A\|_{(p,q)}$ *mátrixnormát* az alábbiak szerint:

$$\|A\|_{(p,q)} := \|f_A\|.$$

Következésképpen

$$\|Ax\|_q \leq \|A\|_{(p,q)} \cdot \|x\|_p \qquad (x \in \mathbf{R}^s),$$

és egyúttal $M := \|A\|_{(p,q)}$ a legkisebb olyan $M \geq 0$ szám, amellyel

$$\|Ax\|_q \leq M \cdot \|x\|_p \qquad (x \in \mathbf{R}^s).$$

Állapodjunk meg abban, hogy a $p = q$ esetben $\|A\|_{(p,q)}$ helyett $\|A\|_{(p)}$-t írunk. Speciálisan, ha $A = (a_{ik})_{i,k=1}^{m,s}$, akkor jól ismert (de elemi számolással is rövid úton belátható), hogy

$$\|A\|_{(\infty)} = \max\left\{\sum_{k=1}^{s} |a_{ik}| : i = 1, ..., m\right\},$$

$$\|A\|_{(1)} = \max\left\{\sum_{i=1}^{m} |a_{ik}| : k = 1, ..., s\right\}$$

(az A mátrix *sornormája,* ill. *oszlopnormája*), továbbá az $\|A\|_{(2)}$ ún. *spektrálnormára*[34]

$$\|A\|_{(2)} \leq \sqrt{\sum_{i=1}^{m}\sum_{k=1}^{s} |a_{ik}|^2}.$$

[34]Belátható, hogy $\|A\|_{(2)} = \sqrt{\max\{|\lambda| : \lambda \in S_A\}}$, ahol S_A az A^*A mátrix sajátértékeinek a halmaza.

Tegyük fel pl., hogy

$$a, b \in \mathbf{R},\ a < b$$

esetén valamilyen $n \in \mathbf{N}$ mellett adottak az

$$a \leq x_0 < ... < x_n \leq b$$

ún. *alappontok* és a

$$g_0, ..., g_n \in C[a, b]$$

függvények. Világos, hogy az

$$C[a, b] \ni f \mapsto Tf := T(f) := \sum_{k=0}^{n} f(x_k) g_k \in C[a, b]$$

előírással definiált

$$T : C[a, b] \to C[a, b]$$

leképezés lineáris. Mivel

$$|Tf(x)| \leq \sum_{k=0}^{n} |f(x_k)| \cdot |g_k(x)| \leq \|f\|_\infty \cdot \sum_{k=0}^{n} |g_k(x)| \qquad (x \in [a, b])$$

miatt

$$\|Tf\|_\infty \leq \|f\|_\infty \cdot \Big\| \sum_{k=0}^{n} |g_k| \Big\|_\infty,$$

ezért a T egyúttal korlátos lineáris is.[35] Belátható, hogy

$$\|T\| = \Big\| \sum_{k=0}^{n} |g_k| \Big\|_\infty.$$

Ha az $f(x_k)$-knak csak valamilyen y_k közelítéseit ismerjük és egy pozitív ε számmal

$$|f(x_k) - y_k| \leq \varepsilon \qquad (k = 0, ..., n),$$

akkor

$$|Tf(x) - \sum_{k=0}^{n} y_k g_k(x)| \leq \sum_{k=0}^{n} |f(x_k) - y_k| \cdot |g_k(x)| \leq$$

$$\varepsilon \cdot \sum_{k=0}^{n} |g_k(x)| \leq \varepsilon \cdot \Big\| \sum_{k=0}^{n} |g_k| \Big\|_\infty = \|T\| \cdot \varepsilon \qquad (x \in [a, b]).$$

Specálisan, ha itt

$$g_k(x) := \prod_{k \neq j = 0}^{n} \frac{x - x_j}{x_k - x_j} \qquad (x \in [a, b],\ k = 0, ..., n),$$

[35] A $C[a, b]$-n a $\|.\|_\infty$ normát tekintve.

akkor

$$Tf = P_{|_{[a,b]}} \qquad (f \in C[a,b]),$$

ahol a P függvény nem más, mint a fenti alappontokra vonatkozóan az f *Lagrange-interpolációs polinomja*:

$$P(x_k) = f(x_k) \qquad (k = 0, ..., n).$$

Ennek szellemében szokás a fenti T-t *általánosított interpolációs operátornak* is nevezni.

4. Differenciálható leképezések.

A továbbiakban adott $1 \leq n, m \in \mathbf{N}$ mellett definiálni fogjuk az $f \in \mathbf{R}^n \to \mathbf{R}^m$ függvények valamilyen $a \in \operatorname{int} \mathcal{D}_f$ pontbeli differenciálhatóságát. Ezzel kapcsolatban emlékeztetünk az „egyváltozós" analízisben megismert $n = m = 1$ esetre, azaz, az $f \in \mathbf{R} \to \mathbf{R}$ egyváltozós valós függvények differenciálhatóságára. Nevezetesen, az f függvény differenciálható az $a \in \operatorname{int} \mathcal{D}_f$ helyen, ha alkalmas $q \in \mathbf{R}$ számmal és $\varepsilon \in \mathbf{R} \to \mathbf{R}$ függvénnyel az alábbiak teljesülnek:

$$f(a+h) - f(a) = qh + \varepsilon(h) \cdot h \qquad (h \in \mathbf{R},\ a+h \in \mathcal{D}_f),$$

és

$$\varepsilon(h) \to 0 \qquad (h \to 0).$$

A fenti definíciónak eleget tevő $f'(a) := q$ szám (az f függvény a-beli deriváltja) egyértelműen létezik.

Legyen az előbbiekben

$$L(x) := qx \qquad (x \in \mathbf{R}),$$

ekkor (ld. 2.1. vi) megjegyzés) az $L : \mathbf{R} \to \mathbf{R}$ függvény korlátos lineáris leképezés, és a differenciálhatóságra vonatkozó fenti formula így (is) írható:

$$f(a+h) - f(a) = L(h) + \eta(h) \cdot |h| \qquad (h \in \mathbf{R},\ a+h \in \mathcal{D}_f),$$

ahol az $\eta \in \mathbf{R} \to \mathbf{R}$ függvényre

$$\eta(h) \to 0 \qquad (|h| \to 0).$$

Minden „készen áll" ahhoz, hogy a differenciálhatóság utóbbi definícióját kiterjesszük többváltozós vektorfüggvényekre. Legyen tehát

$$1 \leq n,\ m \in \mathbf{N},\ 1 \leq p,\ q \leq +\infty,$$

és tekintsük az

$$(\mathbf{R}^n, \|.\|_p),\ (\mathbf{R}^m, \|.\|_q)$$

normált tereket (ld. 1.2. v) megjegyzés). Azt mondjuk, hogy az f függvény *differenciálható az* $a \in \operatorname{int} \mathcal{D}_f$ *helyen,* ha van olyan $L \in L(\mathbf{R}^n, \mathbf{R}^m)$ korlátos lineáris leképezés és olyan $\eta \in \mathbf{R}^n \to \mathbf{R}^m$ függvény, hogy

$$(*) \qquad f(a+h)-f(a) = L(h)+\eta(h)\cdot \|h\|_p \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f),$$

ahol

$$\eta(h) \to 0 \qquad (\|h\|_p \to 0).$$

Tehát tetszőleges $\varepsilon > 0$ szám esetén létezik olyan $\delta > 0$, hogy

$$\|\eta(h)\|_q < \varepsilon \qquad (0 < \|h\|_p < \delta).$$

Mindezt írhatjuk úgy is, hogy

$$\|\eta(h)\|_q \to 0 \qquad (\|h\|_p \to 0).$$

Más szóval

$$\frac{f(a+h)-f(a)-L(h)}{\|h\|_p} \to 0 \qquad (\|h\|_p \to 0),$$

vagy ami ezzel ekvivalens

$$(*) \qquad \frac{\|f(a+h)-f(a)-L(h)\|_q}{\|h\|_p} \to 0 \qquad (\|h\|_p \to 0).$$

Mindezt az $f \in D\{a\}$ jelöléssel juttatjuk kifejezésre. Ha ez minden $a \in \mathcal{D}_f$ esetén teljesül, akkor röviden azt mondjuk, hogy az f függvény *differenciálható,* és erre az $f \in D$ jelölést használjuk.

A fentiekben – ha kell – feltételezhetjük, hogy $\eta(0) = 0$, azaz $\lim_0 \eta = 0$ miatt az η folytonos a 0-ban: $\eta \in \mathcal{C}\{0\}$.

Az is könnyen átlátható, hogy a $(*)$ összefüggésben a $h \in \mathbf{R}^n$, $a + h \in \mathcal{D}_f$ feltételezés helyett elegendő azt megkövetelni, hogy valamilyen $r > 0$ mellett a $(*)$ egyenlőség minden olyan $h \in K_r(0)$ esetén fennálljon, amelyre $a+h \in \mathcal{D}_f$. Mivel $a \in \operatorname{int} \mathcal{D}_f$, ezért pl. eleve van olyan $r > 0$, hogy $K_r(a) \subset \mathcal{D}_f$. Így minden $h \in K_r(0)$ vektorra $a + h \in \mathcal{D}_f$ „automatikusan" igaz.

Jegyezzük meg továbbá, hogy a pontbeli differenciálhatóság most megfogalmazott definíciója csak látszólag függ a $\|.\|_p$, $\|.\|_q$ normáktól. Ui. a $\|.\|_p$, $\|.\|_q$ normák ekvivalenciája miatt (ld. 1.4. xi) megjegyzés) a $(*)$ tulajdonság vagy minden $1 \le p,\ q \le +\infty$ mellett teljesül, vagy egyik mellett sem. Ezért a továbbiakban többnyire a $\|.\|_\infty$, $\|.\|_1$, $\|.\|_2$ normák valamelyikét fogjuk „használni".

3.1.1. Tétel. *Tegyük fel, hogy az* $f \in \mathbf{R}^n \to \mathbf{R}^m$ *függvény valamilyen* $a \in \operatorname{int}\mathcal{D}_f$ *helyen differenciálható. Akkor az ennek a definíciójában szereplő* $L \in L(\mathbf{R}^n, \mathbf{R}^m)$ *korlátos lineáris leképezés egyértelműen létezik.*

Bizonyítás. Tegyük fel, hogy az

$$L,\ \tilde{L} \in L(\mathbf{R}^n, \mathbf{R}^m)$$

leképezésekkel, és az

$$\eta,\ \tilde{\eta} \in \mathbf{R}^n \to \mathbf{R}^m$$

függvényekkel

$$f(a+h)-f(a) = L(h)+\eta(h)\cdot \|h\|_p = \tilde{L}(h)+\tilde{\eta}(h)\cdot \|h\|_p \quad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f),$$

ahol

$$\eta(h) \to 0\ ,\ \tilde{\eta}(h) \to 0 \qquad (\|h\|_p \to 0).$$

Ekkor az

$$L_* := L - \tilde{L} \in L(\mathbf{R}^n, \mathbf{R}^m)\ \ ,\ \ p = q = \infty := +\infty$$

jelöléssel

$$\frac{\|L_*(h)\|_\infty}{\|h\|_\infty} = \|\tilde{\eta}(h) - \eta(h)\|_\infty \leq \|\tilde{\eta}(h)\|_\infty + \|\eta(h)\|_\infty \to 0 \qquad (\|h\|_\infty \to 0)$$

is igaz. Ha $i = 1, ..., n$ és

$$e_i := (0, ..., 0, 1, 0, ...0) \in \mathbf{R}^n$$

(ahol tehát az e_i vektor i-edik koordinátája 1, a többi 0), akkor $\|te_i\|_\infty = |t|$ $(t \in \mathbf{R})$ miatt

$$\frac{\|L_*(te_i)\|_\infty}{\|te_i\|_\infty} = \frac{\|t\cdot L_*(e_i)\|_\infty}{\|te_i\|_\infty} = |L_*(e_i)| \to 0 \qquad (t \to 0).$$

Ez csak úgy lehetséges, ha $L_*(e_i) = 0\ \ (i = 1, ..., n),$ amiből az

$$x = (x_1, ..., x_n) = \sum_{i=1}^{n} x_i\cdot e_i \in \mathbf{R}^n$$

előállítást figyelembe véve

$$L_*(x) = \sum_{i=1}^{n} x_i\cdot L(e_i) = 0,$$

azaz $L_* \equiv 0$ adódik. Mindez azt jelenti, hogy $L = \tilde{L}$. ■

Az $f \in D\{a\}$ esetben az előbbi tétel szerint egyértelműen létező $L \in L(\mathbf{R}^n, \mathbf{R}^m)$ korlátos lineáris leképezést az f függvény a-beli *deriváltjának* nevezzük, és az $f'(a)$ szimbólummal jelöljük. Az $L \in L(\mathbf{R}^n, \mathbf{R}^m)$ halmaz és az $\mathbf{R}^{m\times n}$ mátrixok kapcsolatára gondolva tehát minden $f \in D\{a\}$ függvény esetén egyértelműen adható meg olyan $A \in \mathbf{R}^{m\times n}$ mátrix, amellyel az $L := f'(a)$ leképezés a következő alakú:

$$L(x) = Ax \qquad (x \in \mathbf{R}^n).$$

A most felidézett $L(\mathbf{R}^n, \mathbf{R}^m) \approx \mathbf{R}^{m\times n}$ megfeleltetés alapján ezért azt is mondjuk (és írjuk), hogy

$$f'(a) := A$$

az f függvény a-beli *deriváltja* vagy *deriváltmátrixa.* Az $f'(a)$ (derivált)mátrixra használatos a *Jacobi-mátrix* elnevezés is. Ha tehát $f \in D\{a\}$, akkor

$$f(a+h) - f(a) = f'(a)h + \eta(h)\cdot \|h\| \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f),$$

ahol $\eta(h) \to 0 \quad (\|h\| \to 0)$, és $\|.\|$ a $\|.\|_p \quad (1 \le p \le +\infty)$ normák bármelyike lehet. Ha itt $m = 1$, azaz $f \in \mathbf{R}^n \to \mathbf{R}$, akkor

$$\mathrm{grad} f(a) := f'(a) \in \mathbf{R}^{1\times n} \approx \mathbf{R}^n,$$

azaz ebben az esetben az $f'(a)$ Jacobi-mátrix tekinthető egy $\mathbf{R}^n$-beli vektornak, amit az f függvény a-beli *gradiensének* nevezünk. Ekkor az előbbi $f'(a)h \quad (h \in \mathbf{R}^n)$ mátrix-vektor-szorzat nem más, mint a $\mathrm{grad} f(a)$ és a h vektor „szokásos" ($\mathbf{R}^n$-beli) skaláris szorzata:

$$\langle u, v\rangle = \sum_{j=1}^{n} u_j v_j \qquad \Big(u = (u_1, ..., u_n),\ v = (v_1, ..., v_n) \in \mathbf{R}^n\Big).$$

Ezért

$$f(a+h) - f(a) = \langle \mathrm{grad} f(a), h\rangle + \eta(h)\cdot \|h\| \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f).$$

Speciálisan, ha $n = 1$ is igaz, akkor $\mathrm{grad} f(a) = f'(a) \in \mathbf{R}$, az $f \in \mathbf{R} \to \mathbf{R}$ egyváltozós valós függvény „klasszikus" a-beli deriváltja.

Ha

$$\mathcal{D} := \{a \in \mathcal{D}_f : f \in D\{a\}\} \ne \emptyset,$$

akkor az

$$\mathcal{D} \ni x \mapsto \mathrm{grad} f(x) \in \mathbf{R}^n$$

függvényt (az f *deriváltfüggvényét*) az f függvény *gradiensének* nevezzük, és a $\text{grad}f$ szimbólummal jelöljük. Tehát

$$\text{grad}f \in \mathbf{R}^n \to \mathbf{R}^n.$$

Ha viszont $n=1$, azaz $f \in \mathbf{R} \to \mathbf{R}^m$, akkor

$$f'(a) \in \mathbf{R}^{m\times 1} \approx \mathbf{R}^m,$$

tehát ebben az esetben is az $f'(a)$ Jacobi-mátrix tekinthető egy, de most $\mathbf{R}^m$-beli vektornak, amit az f függvény a-beli *deriváltvektorának* nevezünk. Ekkor

$$f(a+h) - f(a) = f'(a)h + \eta(h)\cdot |h| \qquad (h \in \mathbf{R},\ a+h \in \mathcal{D}_f).$$

(Az $m=1$ esetben most is visszakapjuk az egyváltozós valós függvények deriváltját.) Tegyük fel, hogy

$$\mathcal{D} := \{a \in \mathcal{D}_f : f \in D\{a\}\} \neq \emptyset.$$

Ekkor az

$$\mathcal{D} \ni x \mapsto f'(x) \in \mathbf{R}^m$$

függvény az f *deriváltfüggvénye,* amit az f' szimbólummal jelölünk:

$$f' \in \mathbf{R} \to \mathbf{R}^m.$$

Legyen továbbra is $1 \leq n, m \in \mathbf{N}$, és állapodjunk meg a következőkben: egy $A \in \mathbf{R}^{m\times n}$ mátrix sorvektorait az $A_i \quad (i = 1, ..., m)$ szimbólummal fogjuk jelölni, magát az A mátrixot pedig az A_i-k szerint particionáljuk. Ha tehát $A = \left(a_{ik}\right)_{i,k=1}^{m,n}$ (azaz a_{ik} az A mátrix i-edik sorának a k-adik eleme $(i = 1, ..., m,$ ill. $k = 1, ..., n)$, akkor

$$A_i := (a_{i1}, ..., a_{in}) \in \mathbf{R}^n \qquad (i = 1, ..., m),$$

és

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}.$$

A mátrix-vektor-szorzás értelmezése szerint

$$Ax = \Big(\langle A_1, x\rangle, \ldots, \langle A_m, x\rangle\Big) \qquad (x \in \mathbf{R}^n).$$

3.1.2. Tétel. *Legyen* $1 \le n, m \in \mathbf{N}$. *Az* $f = (f_1, ..., f_m) \in \mathbf{R}^n \to \mathbf{R}^m$ *függvény akkor és csak akkor differenciálható az* $a \in \operatorname{int} \mathcal{D}_f$ *helyen, ha minden* $i = 1, ..., m$ *esetén az* $f_i \in \mathbf{R}^n \to \mathbf{R}$ *koordináta-függvény differenciálható az* a*-ban. Ha* $f \in D\{a\}$, *akkor az* $f'(a)$ *Jacobi-mátrix a következő alakú:*

$$f'(a) = \begin{bmatrix} \operatorname{grad} f_1(a) \\ \operatorname{grad} f_2(a) \\ \vdots \\ \operatorname{grad} f_m(a) \end{bmatrix}.$$

Bizonyítás. Tegyük fel először, hogy $f \in D\{a\}$, és jelöljük az $f'(a) \in \mathbf{R}^{m \times n}$ Jacobi-mátrix sorvektorait A_i-vel $(i = 1, ..., m)$:

$$f'(a) = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}.$$

Ekkor alkalmas

$$\eta = (\eta_1, ..., \eta_m) \in \mathbf{R}^n \to \mathbf{R}^m \quad , \quad \eta(h) \to 0 \quad (\|h\|_2 \to 0)$$

függvénnyel

$$f(a+h) - f(a) = \Big(f_1(a+h) - f_1(a), \ldots, f_m(a+h) - f_m(a)\Big) =$$

$$f'(a)h + \eta(h) \cdot \|h\|_2 =$$

$$\Big(\langle A_1, h\rangle, \ldots, \langle A_m, h\rangle\Big) + \Big(\eta_1(h) \cdot \|h\|_2, \ldots, \eta_m(h) \cdot \|h\|_2\Big) \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f).$$

Következésképpen minden $i = 1, ..., m$ mellett az η függvény $\eta_i \in \mathbf{R}^n \to \mathbf{R}$ koordináta-függvényeivel

$$(*) \qquad f_i(a+h) - f_i(a) = \langle A_i, h\rangle + \eta_i(h) \cdot \|h\|_2 \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f).$$

Mivel bármely $i = 1, ..., m$ indexre

$$\eta_i(h) \to 0 \qquad (\|h\|_2 \to 0),$$

ezért az előbbi $(*)$ összefüggés azt jelenti, hogy $f_i \in D\{a\}$, és

$$A_i = \operatorname{grad} f_i(a) \qquad (i = 1, ..., m).$$

Most azt tegyük fel, hogy $f_i \in D\{a\}$ $(i = 1, ..., m)$, amikor is valamilyen

$$\eta_i \in \mathbf{R}^n \to \mathbf{R} \ , \ \eta_i \to 0 \ \ (\|h\|_2 \to 0) \qquad (i = 1, ..., m)$$

függvényekkel

$$f_i(a+h)-f_i(a) = \langle \mathrm{grad} f_i(a), h\rangle + \eta_i(h) \cdot \|h\|_2 \quad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f,\ i = 1, ..., m).$$

Ha tehát

$$A := \begin{bmatrix} \mathrm{grad}\, f_1(a) \\ \mathrm{grad}\, f_2(a) \\ \vdots \\ \mathrm{grad}\, f_m(a) \end{bmatrix} \in \mathbf{R}^{m \times n},$$

akkor az

$$\eta := (\eta_1, ..., \eta_m) \in \mathbf{R}^n \to \mathbf{R}^m$$

függvénnyel

$$f(a+h) - f(a) = Ah + \eta(h) \cdot \|h\|_2,$$

ahol (ld. 2.1. xvii) megjegyzés) $\eta(h) \to 0$ $(\|h\|_2 \to 0)$. Ez azt jelenti, hogy $f \in D\{a\}$, és (ld. 3.1.1. Tétel) $f'(a) = A$. ∎

Az előbbi tétel szerint a Jacobi-mátrix „kiszámításához" elegendő a gradiens vektorok „szerkezetét" ismernünk. Ez utóbbihoz induljunk ki (tetszőlegesen adott $1 \leq n \in \mathbf{N}$ mellett) egy $h \in \mathbf{R}^n \to \mathbf{R}$ függvényből. Ha $i = 1, ..., n$ és $a = (a_1, ..., a_n) \in \mathcal{D}_h$, akkor legyen

$$\mathcal{D}_{h,i}^{(a)} := \{t \in \mathbf{R} : (a_1, ..., a_{i-1}, t, a_{i+1}, ..., a_n) \in \mathcal{D}_h\}.$$

Értelemszerűen

$$\mathcal{D}_{h,1}^{(a)} := \{t \in \mathbf{R} : (t, a_2, ..., a_n) \in \mathcal{D}_h\},\ \mathcal{D}_{h,n}^{(a)} := \{t \in \mathbf{R} : (a_1, ..., a_{n-1}, t) \in \mathcal{D}_h\}.$$

Világos, hogy $\mathcal{D}_{h,i}^{(a)} \subset \mathbf{R}$, és $\mathcal{D}_{h,i}^{(a)} \neq \emptyset$, hiszen $a_i \in \mathcal{D}_{h,i}^{(a)}$. Speciálisan, ha $n = 1$, akkor $\mathcal{D}_{h,1}^{(a)} = \mathcal{D}_h$. Legyen ezek után

$$h_{a,i} : \mathcal{D}_{h,i}^{(a)} \to \mathbf{R}$$

az a függvény, amelyre

$$h_{a,i}(t) := h(a_1, ..., a_{i-1}, t, a_{i+1}, ..., a_n) \qquad \left(t \in \mathcal{D}_{h,i}^{(a)}\right),$$

ahol

$$h_{a,1}(t) = h(t, a_2, ..., a_n) \qquad \left(t \in \mathcal{D}_{h,1}^{(a)}\right)$$

és
$$h_{a,n}(t) = h(a_1, ..., a_{n-1}, t) \qquad \left(t \in \mathcal{D}^{(a)}_{h,n}\right).$$
A $h_{a,i}$ $(i = 1, ..., n)$ *parciális függvények* valamennyien egyváltozós valós függvények:
$$h_{a,i} \in \mathbf{R} \to \mathbf{R}.$$
Ha itt $n = 1$, azaz $h \in \mathbf{R} \to \mathbf{R}$, akkor nyilván $h_{a,1} = h$.

A fenti jelöléseket megtartva azt mondjuk, hogy a h függvény az a-ban *az i-edik változó szerint parcálisan deriválható* (vagy differenciálható), ha $h_{a,i} \in D\{a_i\}$. Ekkor a
$$\partial_i h(a) := h'_{a,i}(a_i)$$
valós számot a h függvény a-beli, *az i-edik változó szerinti parcális deriváltjának* nevezzük. Következésképpen
$$\partial_i h(a) = \lim_{x \to a_i} \frac{h_{a,i}(x) - h_{a,i}(a_i)}{x - a_i} = \lim_{x \to a_i} \frac{h(a_1, ..., a_{i-1}, x, a_{i+1}, ..., a_n) - h(a)}{x - a_i} =$$
$$\lim_{t \to 0} \frac{h(a_1, ..., a_{i-1}, a_i + t, a_{i+1}, ..., a_n) - h(a)}{t}.$$
Formálisan szólva tehát az i-edik változó szerinti parciális derivált kiszámításakor a h függvény argumentumában (az a_j $(i \neq j = 1, ..., n)$ komponenseket rögzítve) csak „az i-edik változótól való függést" vesszük figyelembe, és az így kapott egyváltozós valós függvényt deriváljuk az a_i helyen.

Tegyük fel, hogy $i = 1, ..., n$, és a fenti $h \in \mathbf{R}^n \to \mathbf{R}$ függvényre
$$\mathcal{D}_{h,i} := \{a \in \mathcal{D}_h : \text{ létezik a } \partial_i f(a) \text{ parciáis derivált}\} \neq \emptyset.$$
Ekkor a
$$\mathcal{D}_{h,i} \ni x \mapsto \partial_i h(x)$$
függvényt a h függvény *i-edik változó szerinti parciális deriváltfüggvényének* nevezzük, és a $\partial_i h$ szimbólummal jelöljük.

7. Fejezet

1. Differenciálható leképezések.

A továbbiakban adott $1 \leq n, m \in \mathbf{N}$ mellett definiálni fogjuk az
$$f \in \mathbf{R}^n \to \mathbf{R}^m$$
függvények valamilyen $a \in \operatorname{int} \mathcal{D}_f$ pontbeli differenciálhatóságát.

Előzetesen emlékeztetünk az

$$F \in \mathbf{R} \to \mathbf{R}$$

egyváltozós valós függvények differenciálhatóságára. Nevezetesen, az F függvény differenciálható az $a \in \operatorname{int} \mathcal{D}_F$ helyen, ha alkalmas $q \in \mathbf{R}$ számmal és

$$\varphi \in \mathbf{R} \to \mathbf{R}$$

függvénnyel az alábbiak teljesülnek:

$$F(a+h) - F(a) = qh + \varphi(h) \cdot h \qquad (h \in \mathbf{R},\, a+h \in \mathcal{D}_F),$$

és

$$\varphi(h) \to 0 \qquad (h \to 0).$$

A fenti definíciónak eleget tevő $F'(a) := q$ szám (az F függvény a-beli deriváltja) (ha létezik) egyértelműen létezik. Nyilvánvaló, hogy ekkor

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

Legyen az előbbiekben

$$L(x) := qx \qquad (x \in \mathbf{R}),$$

ekkor az

$$L : \mathbf{R} \to \mathbf{R}$$

függvény korlátos lineáris leképezés, és a differenciálhatóságra vonatkozó fenti formula így (is) írható:

$$F(a+h) - F(a) = L(h) + \eta(h) \cdot |h| \qquad (h \in \mathbf{R},\, a+h \in \mathcal{D}_F),$$

ahol az $\eta \in \mathbf{R} \to \mathbf{R}$ függvényre

$$\eta(h) \to 0 \qquad (|h| \to 0).$$

Minden „készen áll" ahhoz, hogy a differenciálhatóság utóbbi definícióját kiterjesszük többváltozós vektorfüggvényekre. Legyen tehát

$$1 \leq n,\, m \in \mathbf{N},\ 1 \leq p,\, q \leq +\infty,$$

és tekintsük az

$$(\mathbf{R}^n, \|.\|_p),\ (\mathbf{R}^m, \|.\|_q)$$

normált tereket. Azt mondjuk, hogy az

$$f \in \mathbf{R}^n \to \mathbf{R}^m$$

függvény *differenciálható az* $a \in \operatorname{int} \mathcal{D}_f$ *helyen,* ha van olyan

$$L \in L(\mathbf{R}^n, \mathbf{R}^m)$$

korlátos lineáris leképezés és olyan

$$\eta \in \mathbf{R}^n \to \mathbf{R}^m$$

függvény, hogy

$$(*) \qquad f(a+h) - f(a) = L(h) + \eta(h) \cdot \|h\|_p \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f),$$

ahol

$$\eta(h) \to 0 \qquad (\|h\|_p \to 0).$$

Tehát tetszőleges $\varepsilon > 0$ szám esetén létezik olyan $\delta > 0$, hogy

$$\|\eta(h)\|_q < \varepsilon \qquad (0 < \|h\|_p < \delta).$$

Mindezt írhatjuk úgy is, hogy

$$\|\eta(h)\|_q \to 0 \qquad (\|h\|_p \to 0).$$

Más szóval

$$\frac{f(a+h) - f(a) - L(h)}{\|h\|_p} \to 0 \qquad (\|h\|_p \to 0),$$

vagy ami ezzel ekvivalens

$$(**) \qquad \frac{\|f(a+h) - f(a) - L(h)\|_q}{\|h\|_p} \to 0 \qquad (\|h\|_p \to 0).$$

Mindezt az $f \in D\{a\}$ jelöléssel juttatjuk kifejezésre. Ha ez minden $a \in \mathcal{D}_f$ esetén teljesül, akkor röviden azt mondjuk, hogy az f függvény *differenciálható,* és erre az $f \in D$ jelölést használjuk.

A fentiekben – ha kell – feltételezhetjük, hogy $\eta(0) = 0$, azaz

$$\lim_0 \eta = 0$$

miatt az η folytonos a 0-ban: $\eta \in \mathcal{C}\{0\}$.

Az is könnyen átlátható, hogy a $(*)$ összefüggésben a

$$h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f$$

feltételezés helyett elegendő azt megkövetelni, hogy valamilyen $r>0$ mellett a $(*)$ egyenlőség minden olyan $h \in K_r(0)$ esetén fennálljon, amelyre

$$a+h \in \mathcal{D}_f.$$

Mivel $a \in \operatorname{int} \mathcal{D}_f$, ezért pl. eleve van olyan $r>0$, hogy $K_r(a) \subset \mathcal{D}_f$. Így minden $h \in K_r(0)$ vektorra $a+h \in \mathcal{D}_f$ „automatikusan" igaz.

A pontbeli differenciálhatóság most megfogalmazott definíciója csak látszólag függ a $\|.\|_p$, $\|.\|_q$ normáktól. Ui. a $\|.\|_p$, $\|.\|_q$ normák ekvivalenciája miatt a $(**)$ tulajdonság vagy minden

$$1 \leq p,\ q \leq +\infty$$

mellett teljesül, vagy egyik mellett sem. Ezért a továbbiakban többnyire a

$$\|\cdot\|_\infty,\ \|\cdot\|_1,\ \|\cdot\|_2$$

normák egyikét fogjuk „használni".

1. Tétel. *Tegyük fel, hogy az $f \in \mathbf{R}^n \to \mathbf{R}^m$ függvény valamilyen $a \in \operatorname{int} \mathcal{D}_f$ helyen differenciálható. Akkor az ennek a definíciójában szereplő $L \in L(\mathbf{R}^n, \mathbf{R}^m)$ korlátos lineáris leképezés egyértelműen létezik.*

Bizonyítás. Tegyük fel, hogy az

$$L,\ \tilde{L} \in L(\mathbf{R}^n, \mathbf{R}^m)$$

leképezésekkel és az

$$\eta,\ \tilde{\eta} \in \mathbf{R}^n \to \mathbf{R}^m$$

függvényekkel

$$f(a+h)-f(a)=$$

$$L(h)+\eta(h)\cdot \|h\|_p = \tilde{L}(h)+\tilde{\eta}(h)\cdot \|h\|_p \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f),$$

ahol

$$\eta(h) \to 0\ ,\ \tilde{\eta}(h) \to 0 \qquad (\|h\|_p \to 0).$$

Ekkor az

$$L_* := L-\tilde{L} \in L(\mathbf{R}^n, \mathbf{R}^m),\ \ p=q=\infty := +\infty$$

jelöléssel

$$L_*(h) = (\tilde{\eta}(h) - \eta(h))\cdot \|h\|_p \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f),$$

így

$$\frac{\|L_*(h)\|_\infty}{\|h\|_\infty} = \|\tilde{\eta}(h) - \eta(h)\|_\infty \leq \|\tilde{\eta}(h)\|_\infty + \|\eta(h)\|_\infty \to 0 \qquad (\|h\|_\infty \to 0)$$

is igaz. Ha $i = 1, ..., n$ és

$$e_i := (0, ..., 0, 1, 0, ...0) \in \mathbf{R}^n$$

(ahol tehát az e_i vektor i-edik koordinátája 1, a többi 0), akkor

$$\|te_i\|_\infty = |t| \qquad (t \in \mathbf{R})$$

miatt

$$\frac{\|L_*(te_i)\|_\infty}{\|te_i\|_\infty} = \frac{\|t\cdot L_*(e_i)\|_\infty}{\|te_i\|_\infty} = \|L_*(e_i)\|_\infty \to 0 \qquad (t \to 0).$$

Ez csak úgy lehetséges, ha

$$L_*(e_i) = 0 \qquad (i = 1, ..., n),$$

amiből az

$$x = (x_1, ..., x_n) = \sum_{i=1}^{n} x_i\cdot e_i \in \mathbf{R}^n$$

előállítást figyelembe véve

$$L_*(x) = \sum_{i=1}^{n} x_i\cdot L_*(e_i) = 0,$$

azaz $L_* \equiv 0$ adódik. Mindez azt jelenti, hogy $L = \tilde{L}$. ∎

Az $f \in D\{a\}$ esetben az előbbi tétel szerint egyértelműen létező

$$L \in L(\mathbf{R}^n, \mathbf{R}^m)$$

korlátos lineáris leképezést az f függvény a-beli *deriváltjának* nevezzük, és az $f'(a)$ szimbólummal jelöljük.

Az $L \in L(\mathbf{R}^n, \mathbf{R}^m)$ halmaz és az $\mathbf{R}^{m\times n}$ mátrixok kapcsolatára gondolva tehát minden $f \in D\{a\}$ függvény esetén egyértelműen adható meg olyan $A \in \mathbf{R}^{m\times n}$ mátrix, amellyel az $L := f'(a)$ leképezés a következő alakú:

$$L(x) = Ax \qquad (x \in \mathbf{R}^n).$$

A most felidézett $L(\mathbf{R}^n, \mathbf{R}^m) \approx \mathbf{R}^{m\times n}$ megfeleltetés alapján ezért azt is mondjuk (és írjuk), hogy

$$f'(a) := A$$

az f függvény a-beli *deriváltja* vagy *deriváltmátrixa.* Az $f'(a)$ (derivált)mátrixra használatos a *Jacobi-mátrix*[36] elnevezés is. Ha tehát $f \in D\{a\}$, akkor

$$f(a+h) - f(a) = f'(a)h + \eta(h)\cdot \|h\| \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f),$$

ahol

$$\eta(h) \to 0 \qquad (\|h\| \to 0),$$

és $\|.\|$ a $\|.\|_p$ $(1 \le p \le +\infty)$ normák bármelyike lehet.

Ha itt $m = 1$, azaz

$$f \in \mathbf{R}^n \to \mathbf{R},$$

akkor

$$\mathrm{grad} f(a) := f'(a) \in \mathbf{R}^{1\times n} \approx \mathbf{R}^n.$$

Ebben az esetben az $f'(a)$ Jacobi-mátrix tekinthető egy $\mathbf{R}^n$-beli vektornak, amit az f függvény a-beli *gradiensének* nevezünk. Ekkor az előbbi

$$f'(a)h \qquad (h \in \mathbf{R}^n)$$

mátrix-vektor-szorzat nem más, mint a $\mathrm{grad} f(a)$ és a h vektor „szokásos" ($\mathbf{R}^n$-beli) skaláris szorzata, ahol

$$\langle u, v\rangle = \sum_{j=1}^{n} u_j v_j \qquad \Big(u = (u_1, ..., u_n),\ v = (v_1, ..., v_n) \in \mathbf{R}^n\Big).$$

Ezért

$$f(a+h) - f(a) = \langle \mathrm{grad} f(a), h\rangle + \eta(h)\cdot \|h\| \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f).$$

Speciálisan, ha $n = 1$ is igaz, akkor

$$\mathrm{grad} f(a) = f'(a) \in \mathbf{R},$$

az

$$f \in \mathbf{R} \to \mathbf{R}$$

egyváltozós valós függvény „klasszikus" a-beli deriváltja.

---

[36] Carl Gustav Jacob Jacobi (1804-1851).

Ha

$$\mathcal{D} := \{a \in \mathcal{D}_f : f \in D\{a\}\} \neq \emptyset,$$

akkor az

$$\mathcal{D} \ni x \mapsto \mathrm{grad} f(x) \in \mathbf{R}^n$$

függvényt (az f *deriváltfüggvényét*) az f függvény *gradiensének* nevezzük, és a $\mathrm{grad} f$ szimbólummal jelöljük. Tehát

$$\mathrm{grad} f \in \mathbf{R}^n \to \mathbf{R}^n.$$

Ha viszont $n = 1$, azaz

$$f \in \mathbf{R} \to \mathbf{R}^m,$$

akkor

$$f'(a) \in \mathbf{R}^{m\times 1} \approx \mathbf{R}^m,$$

tehát ebben az esetben is az $f'(a)$ Jacobi-mátrix tekinthető egy, de most $\mathbf{R}^m$-beli vektornak, amit az f függvény a-beli *deriváltvektorának* nevezünk. Ekkor

$$f(a+h) - f(a) = f'(a)h + \eta(h) \cdot |h| \qquad (h \in \mathbf{R},\ a+h \in \mathcal{D}_f).$$

(Az $m = 1$ esetben most is visszakapjuk az egyváltozós valós függvények deriváltját.) Tegyük fel, hogy

$$\mathcal{D} := \{a \in \mathcal{D}_f : f \in D\{a\}\} \neq \emptyset.$$

Ekkor az

$$\mathcal{D} \ni x \mapsto f'(x) \in \mathbf{R}^m$$

függvény az f *deriváltfüggvénye,* amit az f' szimbólummal jelölünk:[37]

$$f' \in \mathbf{R} \to \mathbf{R}^m.$$

2. A derivált kiszámítása.

Legyen továbbra is $1 \leq n, m \in \mathbf{N}$, és állapodjunk meg a következőkben: egy $A \in \mathbf{R}^{m\times n}$ mátrix sorvektorait az A_i $(i = 1, ..., m)$ szimbólummal fogjuk jelölni, magát az A mátrixot pedig az A_i-k szerint particionáljuk.

[37]Ha $m, n > 1$, akkor nem beszélünk deriváltfüggvényről: az $\mathbf{R}^n \ni x \mapsto f'(x) \in \mathbf{R}^{m\times n}$ leképezés ekkor „nem fér bele" az általunk vizsgált $\mathbf{R}^n \to \mathbf{R}^m$ függvények körébe.

Ha tehát $A = \left(a_{ik}\right)_{i,k=1}^{m,n}$ (azaz a_{ik} az A mátrix i-edik sorának a k-adik eleme $(i = 1, ..., m,$ ill. $k = 1, ..., n)$, akkor

$$A_i := (a_{i1}, ..., a_{in}) \in \mathbf{R}^n \qquad (i = 1, ..., m)$$

és

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}.$$

A mátrix-vektor-szorzás értelmezése szerint

$$Ax = \Big(\langle A_1, x\rangle, \dots, \langle A_m, x\rangle\Big) \qquad (x \in \mathbf{R}^n).$$

2. Tétel. *Legyen* $1 \le n, m \in \mathbf{N}$. *Az*

$$f = (f_1, ..., f_m) \in \mathbf{R}^n \to \mathbf{R}^m$$

függvény akkor és csak akkor differenciálható az $a \in \operatorname{int} \mathcal{D}_f$ *helyen, ha minden* $i = 1, ..., m$ *esetén az*

$$f_i \in \mathbf{R}^n \to \mathbf{R}$$

koordináta-függvény differenciálható az a*-ban. Ha* $f \in D\{a\}$*, akkor az* $f'(a)$ *Jacobi-mátrix a következő alakú:*

$$f'(a) = \begin{bmatrix} \operatorname{grad} f_1(a) \\ \operatorname{grad} f_2(a) \\ \vdots \\ \operatorname{grad} f_m(a) \end{bmatrix}.$$

Bizonyítás. Tegyük fel először is azt, hogy $f \in D\{a\}$, és jelöljük az $f'(a) \in \mathbf{R}^{m \times n}$ Jacobi-mátrix sorvektorait A_i-vel $(i = 1, ..., m)$:

$$f'(a) = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_m \end{bmatrix}.$$

Ekkor alkalmas

$$\eta = (\eta_1, ..., \eta_m) \in \mathbf{R}^n \to \mathbf{R}^m$$

függvénnyel
$$\eta(h) \to 0 \qquad (\|h\| \to 0)$$
és a
$$h \in \mathbf{R}^n \qquad (a+h \in \mathcal{D}_f)$$
helyeken
$$f(a+h)-f(a) = \Big(f_1(a+h)-f_1(a), \ldots, f_m(a+h)-f_m(a)\Big) =$$
$$f'(a)h + \eta(h)\cdot \|h\| =$$
$$\Big(\langle A_1, h\rangle, \ldots, \langle A_m, h\rangle\Big) + \Big(\eta_1(h)\cdot \|h\|, \ldots, \eta_m(h)\cdot \|h\|\Big).$$
Következésképpen minden $i = 1, ..., m$ mellett az η függvény
$$\eta_i \in \mathbf{R}^n \to \mathbf{R}$$
koordináta-függvényeivel
$$(*) \qquad f_i(a+h)-f_i(a) = \langle A_i, h\rangle + \eta_i(h)\cdot \|h\| \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f).$$
Mivel bármely $i = 1, ..., m$ indexre
$$\eta_i(h) \to 0 \qquad (\|h\| \to 0),$$
ezért az előbbi $(*)$ összefüggés azt jelenti, hogy $f_i \in D\{a\}$ és
$$A_i = \mathrm{grad} f_i(a) \qquad (i = 1, ..., m).$$

Most azt tegyük fel, hogy $f_i \in D\{a\} \quad (i = 1, ..., m)$, amikor is valamilyen
$$\eta_i \in \mathbf{R}^n \to \mathbf{R},\ \ \eta_i \to 0 \ \ (\|h\| \to 0) \qquad (i = 1, ..., m)$$
függvényekkel
$$f_i(a+h)-f_i(a) =$$
$$\langle \mathrm{grad} f_i(a), h\rangle + \eta_i(h)\cdot \|h\| \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f,\ i = 1, ..., m).$$
Ha tehát
$$A := \begin{bmatrix} \mathrm{grad}\, f_1(a) \\ \mathrm{grad}\, f_2(a) \\ \vdots \\ \mathrm{grad}\, f_m(a) \end{bmatrix} \in \mathbf{R}^{m\times n},$$
akkor az
$$\eta := (\eta_1, ..., \eta_m) \in \mathbf{R}^n \to \mathbf{R}^m$$

függvénnyel

$$f(a+h)-f(a)=Ah+\eta(h)\cdot\|h\|,$$

ahol $\eta(h)\to 0\quad(\|h\|\to 0)$. Ezért $f\in D\{a\}$ és $f'(a)=A$. ■

3. Parciális derivált.

Az előbbi tétel szerint a Jacobi-mátrix „kiszámításához" elegendő a gradiens vektorok „szerkezetét" ismernünk. Ez utóbbihoz induljunk ki (tetszőlegesen adott $1\le n\in\mathbf{N}$ mellett) egy

$$g\in\mathbf{R}^n\to\mathbf{R}$$

függvényből. Ha $i=1,...,n$ és $a=(a_1,...,a_n)\in\mathcal{D}_g$, akkor legyen

$$\mathcal{D}_{g,i}^{(a)}:=\{t\in\mathbf{R}:(a_1,...,a_{i-1},t,a_{i+1},...,a_n)\in\mathcal{D}_g\}.$$

Értelemszerűen

$$\mathcal{D}_{g,1}^{(a)}:=\{t\in\mathbf{R}:(t,a_2,...,a_n)\in\mathcal{D}_g\}$$

és[38]

$$\mathcal{D}_{g,n}^{(a)}:=\{t\in\mathbf{R}:(a_1,...,a_{n-1},t)\in\mathcal{D}_g\}.$$

Világos, hogy $\mathcal{D}_{g,i}^{(a)}\subset\mathbf{R}$ és $\mathcal{D}_{g,i}^{(a)}\neq\emptyset$, hiszen $a_i\in\mathcal{D}_{g,i}^{(a)}$. Speciálisan, ha $n=1$, akkor $\mathcal{D}_{g,1}^{(a)}=\mathcal{D}_g$. Legyen ezek után

$$g_{a,i}:\mathcal{D}_{g,i}^{(a)}\to\mathbf{R}$$

az a függvény, amire

$$g_{a,i}(t):=g(a_1,...,a_{i-1},t,a_{i+1},...,a_n)\qquad\left(t\in\mathcal{D}_{g,i}^{(a)}\right),$$ [39]

ahol

$$g_{a,1}(t)=g(t,a_2,...,a_n)\qquad\left(t\in\mathcal{D}_{g,1}^{(a)}\right)$$

és

$$g_{a,n}(t)=g(a_1,...,a_{n-1},t)\qquad\left(t\in\mathcal{D}_{g,n}^{(a)}\right).$$

---

[38] Ha pl. $n=2$, akkor a „szokásos" derékszögű koordinátarendszerben a $\mathcal{D}_{g,1}^{(a)}$ halmaz az a_2 „magasságban" az első koordinátatengellyel párhuzamos (vízszintes) egyenesnek a $\mathcal{D}_g$-vel való metszete.

[39] Szemléletesen fogalmazva az a pont koordinátáit az i -edik kivételével rögzítjük, és a g argumentumában csak az i -edik koordinátát változtatjuk. Így pl. $n=2$ esetén $g_{a,1}$ a g függvény leszűkítése az előbb említett vízszintes egyenesre.

A $g_{a,i}$ $(i = 1, ..., n)$ *parciális függvények* valamennyien egyváltozós valós függvények:

$$g_{a,i} \in \mathbf{R} \to \mathbf{R}.$$

Ha itt $n = 1$, azaz $h \in \mathbf{R} \to \mathbf{R}$, akkor nyilván $g_{a,1} = g$.

A fenti jelöléseket megtartva azt mondjuk, hogy a g függvény az a-ban *az i-edik változó szerint parcálisan deriválható* (vagy differenciálható), ha $g_{a,i} \in D\{a_i\}$. Ekkor a

$$\partial_i g(a) := g'_{a,i}(a_i)$$

valós számot a g függvény a-beli, *az i-edik változó szerinti parcális deriváltjának* nevezzük.[40] Következésképpen

$$\partial_i g(a) = \lim_{x \to a_i} \frac{g_{a,i}(x) - g_{a,i}(a_i)}{x - a_i} =$$

$$\lim_{x \to a_i} \frac{g(a_1, ..., a_{i-1}, x, a_{i+1}, ..., a_n) - g(a)}{x - a_i} =$$

$$\lim_{t \to 0} \frac{g(a_1, ..., a_{i-1}, a_i + t, a_{i+1}, ..., a_n) - g(a)}{t}.$$

Formálisan szólva tehát az i-edik változó szerinti parciális derivált kiszámításakor a g függvény argumentumában (az a_j $(i \neq j = 1, ..., n)$ komponenseket rögzítve) csak „az i-edik változótól való függést" vesszük figyelembe, és az így kapott egyváltozós valós függvényt deriváljuk az a_i helyen.

Tegyük fel, hogy $i = 1, ..., n$ és a fenti $g \in \mathbf{R}^n \to \mathbf{R}$ függvényre

$$\mathcal{D}_{g,i} := \{a \in \mathcal{D}_g : \text{ létezik a } \partial_i g(a) \text{ parciáis derivált}\} \neq \emptyset.$$

Ekkor a

$$\mathcal{D}_{g,i} \ni x \mapsto \partial_i g(x)$$

függvényt a g függvény *i-edik változó szerinti parciális deriváltfüggvényének* nevezzük, és a $\partial_i g$ szimbólummal jelöljük.

Legyen pl. $n := 2$ és

$$g(x, y) := x^2 + 2xy^2 + x + 3y + 1 \qquad \big((x, y) \in \mathbf{R}^2\big).$$

---

[40]Legyen pl. $n = 2$, ekkor a $g \in \mathbf{R}^2 \to \mathbf{R}$ kétváltozós függvényről van szó, amikor is az $\mathcal{F} := \{(x, y, z) \in \mathbf{R}^3 : (x, y) \in \mathcal{D}_g, z = g(x, y)\}$ halmaz a g által meghatározott *felület.* Ha itt $a = (u, v) \in \mathcal{D}_g$, és a derékszögű koordináta-rendszerben tekintjük az Y tengelyre az (u, v) pontban állított merőleges S síkot, akkor az $\mathcal{F} \cap S$ halmaz („felületi görbe") a $g_{a,1} \in \mathbf{R} \to \mathbf{R}$ függvény grafikonja. Ha $g_{a,1} \in D\{u\}$, akkor a $g'_{a,1}(u)$ valós szám a g függvény (u, v)-beli, az első változó szerinti $\partial_1 g(u, v)$ parcális deriváltja.

Ekkor tetszőleges $a=(u,v)\in \mathbf{R}^2$ esetén

$$\mathcal{D}_{g,i}^{(a)}=\mathbf{R} \qquad (i=1,2)$$

és

$$g_{a,1}(t)=h(t,v)=t^2+2tv^2+t+3v+1 \qquad (t\in\mathbf{R}),$$
$$g_{a,2}(t)=h(u,t)=u^2+2ut^2+u+3t+1 \qquad (t\in\mathbf{R}).$$

Mivel a

$$g_{a,1},\ g_{a,2}:\mathbf{R}\to\mathbf{R}$$

függvények differenciálhatók, valamint

$$g'_{a,1}(t)=2t+2v^2+1 \qquad (t\in\mathbf{R}),$$

$$g'_{a,2}(t)=4ut+3 \qquad (t\in\mathbf{R}),$$

ezért léteznek a $\partial_1 g(a)$, $\partial_2 g(a)$ parciális deriváltak, és

$$\partial_1 g(a)=g'_{a,1}(u)=2u+2v^2+1,\ \partial_2 g(a)=g'_{a,2}(v)=4uv+3.$$

3. Tétel. *Tegyük fel, hogy valamilyen $1\le n\in\mathbf{N}$, $g\in\mathbf{R}^n\to\mathbf{R}$ és $a\in\mathcal{D}_g$ esetén $g\in D\{a\}$. Ekkor minden $i=1,...,n$ mellett a g függvény az i-edik változó szerint parciálisan deriválható az a-ban, és*

$$\operatorname{grad} g(a)=\Big(\partial_1 g(a),\ldots,\partial_n g(a)\Big).$$

Bizonyítás. A $g\in D\{a\}$ feltételezés miatt $a\in\operatorname{int}\mathcal{D}_g$, így egy alkalmas $r>0$ sugárral $K_r(a)\subset\mathcal{D}_g$. Ha pl. $\mathbf{R}^n$-ben a ρ_∞ metrikát (vagy ami ugyanaz, a $\|.\|_\infty$ normát) tekintjük, akkor

$$(a_i-r,a_i+r)\subset\mathcal{D}_{g,i}^{(a)} \qquad (i=1,...,n),$$

más szóval $a_i\in\operatorname{int}\mathcal{D}_{g,i}^{(a)}$ $(i=1,...,n)$. Továbbá alkalmas $\eta\in\mathbf{R}^n\to\mathbf{R}$ függvénnyel

$$\eta(x)\to 0 \qquad (\|x\|_\infty\to 0),$$

és a

$$(d_1,...,d_n):=\operatorname{grad} g(a)$$

jelöléssel

$$g(a+x)-g(a)=\langle d,x\rangle+\eta(x)\cdot\|x\|_\infty=$$

$$= \sum_{j=1}^{n} d_j x_j + \eta(x) \cdot \|x\|_\infty \qquad (x = (x_1, ..., x_n) \in \mathbf{R}^n,\ a + x \in \mathcal{D}_g).$$

Legyen itt valamilyen $i = 1, ..., n$ esetén

$$x = (0, \dots, 0, t, 0, \dots, 0) \qquad (t \in \mathbf{R},\ a + x \in \mathcal{D}_g)$$

(ahol tehát az x vektor i-edik koordinátája t, a többi nulla), akkor

$$g(a + x) - g(a) = g(a_1, ..., a_{i-1}, a_i + t, a_{i+1}, ..., a_n) - g(a) =$$

$$g_{a,i}(a_i + t) - g_{a,i}(a_i) = d_i t + \eta(x) \cdot |t| \qquad (t \in \mathbf{R},\ |t| < r).$$

Mivel

$$\lim_{t \to 0} \eta(0, \dots, 0, t, 0, \dots, 0) = \lim_{\|x\|_\infty \to 0} \eta(x) = 0,$$

ezért mindez azt jelenti, hogy $g_{a,i} \in D\{a_i\}$ és

$$g'_{a,i}(a_i) = d_i \qquad (i = 1, ..., n).$$

Tehát léteznek a

$$\partial_i g(a) = g'_{a,i}(a_i) = d_i \qquad (i = 1, ..., n)$$

parciális deriváltak, és

$$\operatorname{grad} g(a) = (d_1, ..., d_n) = \big(\partial_1 g(a), \dots, \partial_n g(a)\big).$$

■

Figyelembe véve a 2. Tételt, a Jacobi-mátrix „szerkezetére" az alábbi állítás adódik.

4. Tétel. *Tegyük fel, hogy adott* $1 \le n, m \in \mathbf{N}$ *esetén az*

$$f = (f_1, ..., f_m) \in \mathbf{R}^n \to \mathbf{R}^m$$

függvény differenciálható az $a \in \mathcal{D}_f$ *pontban. Ekkor az* $f'(a)$ *Jacobi-mátrix a következő alakú:*

$$f'(a) = \begin{bmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \dots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \dots & \partial_n f_2(a) \\ \vdots & \vdots & \dots & \vdots \\ \partial_1 f_m(a) & \partial_2 f_m(a) & \dots & \partial_n f_m(a) \end{bmatrix}.$$

Tehát az $f'(a) \in \mathbf{R}^{m\times n}$ mátrix i-edik sorának a k-adik eleme

$$\partial_k f_i(a) \qquad (i = 1, ..., m;\ k = 1, ..., n).$$

Speciálisan az $n = 1$ esetben

$$\partial_1 f_i(a) = f_i'(a) \qquad (i = 1, ..., m),$$

azaz

$$f'(a) = \big(f_1'(a), \ldots, f_m'(a)\big).$$

Például az

$$f(x, y) := (x^2 - y, x + 2y, y^2) \in \mathbf{R}^3 \qquad \big((x, y) \in \mathbf{R}^2\big)$$

függvény koordináta-függvényei

$$f_1(x, y) = x^2 - y,\ f_2(x, y) = x + 2y,\ f_3(x, y) = y^2 \qquad \big((x, y) \in \mathbf{R}^2\big),$$

amelyek – könnyen belátható módon – differenciálható függvények. Ezért $f \in D$ és

$$f'(x, y) = \begin{bmatrix} 2x & -1 \\ 1 & 2 \\ 0 & 2y \end{bmatrix} \qquad \big((x, y) \in \mathbf{R}^2\big).$$

4. Megjegyzések.

i) Az $f \in \mathbf{R}^n \to \mathbf{R}^m$ függvényekre adott differenciálhatósági definíciónk minden további nélkül megfogalmazható normált terek közötti absztrakt függvényekre is. Legyenek ui. adottak az

$$(X, \|.\|_X),\ (Y, \|.\|_Y)$$

normált terek és az

$$f \in X \to Y$$

függvény. Ekkor az f-et differenciálhatónak nevezzük egy $a \in \operatorname{int} \mathcal{D}_f$ pontban, ha alkalmas $L \in L(X, Y)$ korlátos lineáris leképezéssel és

$$\eta \in X \to Y$$

függvénnyel

$$f(a + h) - f(a) = L(h) + \eta(h) \cdot \|h\|_X \qquad (h \in X,\ a + h \in \mathcal{D}_f),$$

ahol
$$\eta(h) \to 0 \qquad (\|h\|_X \to 0).$$
Másképp fogalmazva
$$\frac{f(a+h) - f(a) - L(h)}{\|h\|_X} \to 0 \qquad (\|h\|_X \to 0),$$
vagy ami ezzel ekvivalens
$$\frac{\|f(a+h) - f(a) - L(h)\|_Y}{\|h\|_X} \to 0 \qquad (\|h\|_X \to 0).$$
Megmutatható, hogy ekkor egyértelműen létezik ilyen $L \in L(X, Y)$, amit az f függvény a-beli *deriváltjának* (vagy *Fréchet-deriváltjának*) nevezünk.[41]

ii) Könnyű meggondolni, hogy tetszőleges
$$1 \leq n, m \in \mathbf{N},\ L \in L(\mathbf{R}^n, \mathbf{R}^m)$$
és $a \in \mathbf{R}^n$ esetén $L \in D\{a\}$ és $L'(a) = L$. Ui.
$$L(a+x) - L(a) = L(a) + L(x) - L(a) =$$
$$L(x) + \eta(x) \cdot \|x\| \qquad (x \in \mathbf{R}^n),$$
ahol
$$\eta(t) := 0 \qquad (t \in \mathbf{R}^n).$$
Nyilván
$$\eta(x) \to 0 \qquad (\|x\| \to 0),$$
tehát az L-re teljesül az a-beli differenciálhatóság minden feltétele.

iii) A pontbeli differenciálhatóság a többváltozós vektorfüggvények körében is „erősebb" tulajdonság a pontbeli folytonosságnál: ha $1 \leq n, m \in \mathbf{N}$ és az
$$f \in \mathbf{R}^n \to \mathbf{R}^m$$
függvény differenciálható valamilyen $a \in \mathcal{D}_f$ helyen, akkor az f az a-ban folytonos. Valóban, pl. a $\|.\| := \|.\|_2$ vektornormával
$$f(a+h) - f(a) = f'(a)h + \eta(h) \cdot \|h\| \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f),$$

[41] René Maurice Fréchet (1878-1973).

ahol $\eta(h) \to 0 \quad (\|h\| \to 0)$, ezért

$$\|\eta(h)\| \to 0 \quad (\|h\| \to 0),$$

és

$$\|f(a+h) - f(a)\| \leq$$

$$\|f'(a)\|_{(2)} \cdot \|h\| + \|\eta(h)\| \cdot \|h\| \to 0 \qquad (\|h\| \to 0).$$

Így létezik a

$$\lim_{\|h\| \to 0} f(a+h) = f(a)$$

határérték, más szóval létezik a

$$\lim_{a} f = f(a)$$

határérték, ami a jelen esetben az f függvény a-beli folytonosságát igazolja.

iv) Legyen $1 \leq n, m \in \mathbf{N}$, $f, g \in \mathbf{R}^n \to \mathbf{R}^m$ és tegyük fel, hogy valamilyen $a \in \mathcal{D}_f \cap \mathcal{D}_g$ helyen az f, g függvények differenciálhatók: $f, g \in D\{a\}$. Ekkor tetszőleges $c \in \mathbf{R}$ együtthatóval $f + cg \in D\{a\}$ és

$$(f + cg)'(a) = f'(a) + cg'(a).$$

8. Fejezet

0. Emlékeztető.

I) Legyen

$$1 \leq n,\ m \in \mathbf{N},\ 1 \leq p,\ q \leq +\infty,$$

és tekintsük az

$$(\mathbf{R}^n, \|.\|_p),\ (\mathbf{R}^m, \|.\|_q)$$

normált tereket. Azt mondjuk, hogy az

$$f \in \mathbf{R}^n \to \mathbf{R}^m$$

függvény *differenciálható az* $a \in \operatorname{int} \mathcal{D}_f$ *helyen* $(f \in D\{a\})$, ha van olyan

$$L \in L(\mathbf{R}^n, \mathbf{R}^m)$$

korlátos lineáris leképezés és olyan

$$\eta \in \mathbf{R}^n \to \mathbf{R}^m$$

függvény, hogy

$$f(a+h)-f(a)=L(h)+\eta(h)\cdot\|h\|_p \qquad (h\in\mathbf{R}^n,\ a+h\in\mathcal{D}_f),$$

ahol

$$\eta(h)\to 0 \qquad (\|h\|_p\to 0).$$

Ekkor ez az $L\in L(\mathbf{R}^n,\mathbf{R}^m)$ egyértelműen létezik, amit az f függvény a-beli *deriváltjának* nevezünk, és az $f'(a)$ szimbólummal jelölünk.

Tudjuk, hogy egyértelműen adható meg olyan $A\in\mathbf{R}^{m\times n}$ mátrix, amellyel az $L:=f'(a)$ leképezés a következő alakú:

$$L(x)=Ax \qquad (x\in\mathbf{R}^n).$$

Így azt is mondjuk (és írjuk), hogy

$$f'(a):=A$$

az f függvény a-beli *deriváltja* vagy *deriváltmátrixa* (*Jacobi-mátrixa*). Ha tehát $f\in D\{a\}$, akkor

$$f(a+h)-f(a)=f'(a)h+\eta(h)\cdot\|h\| \qquad (h\in\mathbf{R}^n,\ a+h\in\mathcal{D}_f),$$

ahol

$$\eta(h)\to 0 \qquad (\|h\|\to 0),$$

és $\|.\|$ a $\|.\|_p$ $(1\le p\le+\infty)$ normák bármelyike lehet.

Ha itt $m=1$, azaz

$$f\in\mathbf{R}^n\to\mathbf{R},$$

akkor

$$\mathrm{grad}f(a):=f'(a)\in\mathbf{R}^{1\times n}\approx\mathbf{R}^n.$$

Ebben az esetben az $f'(a)$ Jacobi-mátrix egy $\mathbf{R}^n$-beli vektor, amit az f függvény a-beli *gradiensének* nevezünk. Ekkor

$$f(a+h)-f(a)=\langle\mathrm{grad}f(a),h\rangle+\eta(h)\cdot\|h\| \qquad (h\in\mathbf{R}^n,\ a+h\in\mathcal{D}_f).$$

Ha

$$\mathcal{D}:=\{a\in\mathcal{D}_f : f\in D\{a\}\}\neq\emptyset,$$

akkor az

$$\mathcal{D}\ni x\mapsto \mathrm{grad}f(x)\in\mathbf{R}^n$$

függvényt (az f *deriváltfüggvényét*) az f függvény *gradiensének* nevezzük, és a $\mathrm{grad} f$ szimbólummal jelöljük. Tehát

$$\mathrm{grad} f \in \mathbf{R}^n \to \mathbf{R}^n.$$

Ha viszont $n = 1$, azaz

$$f \in \mathbf{R} \to \mathbf{R}^m,$$

akkor

$$f'(a) \in \mathbf{R}^{m\times 1} \approx \mathbf{R}^m,$$

tehát az $f'(a)$ Jacobi-mátrix egy $\mathbf{R}^m$-beli vektor, az f függvény a-beli *deriváltvektora.*[42] Ekkor

$$f(a+h) - f(a) = f'(a)h + \eta(h)\cdot |h| \qquad (h \in \mathbf{R},\ a+h \in \mathcal{D}_f).$$

Tegyük fel, hogy

$$\mathcal{D} := \{a \in \mathcal{D}_f : f \in D\{a\}\} \neq \emptyset,$$

amikor is az

$$\mathcal{D} \ni x \mapsto f'(x) \in \mathbf{R}^m$$

függvény az f *deriváltfüggvénye,* amit az f' szimbólummal jelölünk:

$$f' \in \mathbf{R} \to \mathbf{R}^m.$$

II) Legyen adott a

$$g \in \mathbf{R}^n \to \mathbf{R}$$

függvény, ill. $i = 1, ..., n$ és $a = (a_1, ..., a_n) \in \mathcal{D}_g$ esetén

$$\mathcal{D}_{g,i}^{(a)} := \{t \in \mathbf{R} : (a_1, ..., a_{i-1}, t, a_{i+1}, ..., a_n) \in \mathcal{D}_g\}.$$

Továbbá legyen

$$g_{a,i} : \mathcal{D}_{g,i}^{(a)} \to \mathbf{R}$$

az az egyváltozós valós függvény, amire

$$g_{a,i}(t) := g(a_1, ..., a_{i-1}, t, a_{i+1}, ..., a_n) \qquad \left(t \in \mathcal{D}_{g,i}^{(a)}\right).$$

[42] Tegyük fel, hogy a szóban forgó f függvény egy görbe paraméterezése az $\mathbf{R}^m$-ben, és $f'(a) \neq 0\,(\in \mathbf{R}^m)$. Ekkor az $\ell := \{f(a) + t\cdot f'(a) \in \mathbf{R}^m : t \in \mathbf{R}\}$ halmaz az illető görbe $f(a)$ pontbeli *érintője,* ez utóbbinak egy *irányvektora* $f'(a)$. Tekintsük mondjuk $m = 2$ esetén az $\{(u, v) \in \mathbf{R}^2 : u \in [0, 1],\ v = u^2\}$ (sík)görbét (parabolaívet), aminek egy paraméterezése az $f(x) := (x, x^2)$ $(x \in [0, 1])$ függvény. Legyen $a := 1/2$, ekkor $f'(a) = (1, 1)$, ezért most az a-beli érintő az $\ell := \{(1/2, 1/4) + t\cdot (1, 1) \in \mathbf{R}^2 : t \in \mathbf{R}\} = \{(1/2 + t, 1/4 + t) \in \mathbf{R}^2 : t \in \mathbf{R}\}$ (vagy ha tetszik: az $y = x - 1/4$ „egyenletű") egyenes. (Mindez elemi matematikai eszközökkel is könnyen „ellenőrizhető".)

A g függvény az a-ban *az i-edik változó szerint parcálisan deriválható* (vagy differenciálható), ha $g_{a,i} \in D\{a_i\}$. Ekkor a

$$\partial_i g(a) := g'_{a,i}(a_i)$$

valós számot a g függvény a-beli, *az i-edik változó szerinti parcális deriváltjának* nevezzük. Következésképpen

$$\partial_i g(a) = \lim_{x \to a_i} \frac{g_{a,i}(x) - g_{a,i}(a_i)}{x - a_i} =$$

$$\lim_{x \to a_i} \frac{g(a_1, ..., a_{i-1}, x, a_{i+1}, ..., a_n) - g(a_1, ..., a_{i-1}, a_i, a_{i+1}, ..., a_n)}{x - a_i}.$$

A $\partial_i g$, ill. a $\partial_i g(a)$ jelölés mellett használatos még a $\partial_{x_i} g$, ill. a $\partial_{x_i} g(a)$ írásmód, esetenként a

$$\partial_x g, \partial_y g, ...,$$

ill. a

$$\partial_x g(a), \partial_y g(a), ...$$

is.

Az egyváltozós esettel analóg módon írjuk az előbbiek helyett gyakran a következőket:

$$\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, ...,$$

stb. Ha pl. $n := 3$ és

$$g(x, y, z) := x^3 + y^2 z + z^3 + xyz \qquad \Big((x, y, z) \in \mathbf{R}^3\Big),$$

akkor

$$\partial_1 g(x, y, z) = \partial_x g(x, y, z) = 3x^2 + yz,$$

$$\partial_2 g(x, y, z) = \partial_y g(x, y, z) = 2yz + xz,$$

$$\partial_3 g(x, y, z) = \partial_z g(x, y, z) = y^2 + 3z^2 + xy \qquad \Big((x, y, z) \in \mathbf{R}^3\Big).$$

III) Tegyük fel, hogy valamilyen $g \in \mathbf{R}^n \to \mathbf{R}$, $a \in \mathcal{D}_g$ esetén $g \in D\{a\}$. Ekkor minden $i = 1, ..., n$ mellett a g függvény az i-edik változó szerint parciálisan deriválható az a-ban, és

$$\operatorname{grad} g(a) = \Big(\partial_1 g(a), \ldots, \partial_n g(a)\Big).$$

IV) Adott $1 \leq n, m \in \mathbf{N}$ mellett legyen az

$$f = (f_1, ..., f_m) \in \mathbf{R}^n \to \mathbf{R}^m$$

függvény differenciálható az $a \in \mathcal{D}_f$ pontban. Ekkor az $f'(a)$ Jacobi-mátrix a következő alakú:

$$f'(a) = \begin{bmatrix} \partial_1 f_1(a) & \partial_2 f_1(a) & \dots & \partial_n f_1(a) \\ \partial_1 f_2(a) & \partial_2 f_2(a) & \dots & \partial_n f_2(a) \\ \vdots & \vdots & \dots & \vdots \\ \partial_1 f_m(a) & \partial_2 f_m(a) & \dots & \partial_n f_m(a) \end{bmatrix}.$$

1. Differenciálhatóság.

Tudjuk, hogy egy többváltozós függvényt illetően a differenciálhatóság maga után vonja a (minden változó szerinti) parciális differenciálhatóságot. A „fordított" következtetés ugyanakkor csak bizonyos egyéb feltételek esetén igaz. Erről szól az

1. Tétel. *Legyen adott valamilyen* $1 \le n \in \mathbf{N}$ *mellett a*

$$h \in \mathbf{R}^n \to \mathbf{R}$$

függvény és $a \in \operatorname{int} \mathcal{D}_h$. *Tegyük fel, hogy egy* $i \in \{1, ..., n\}$ *esetén*

- *alkalmas* $r > 0$ *mellett* $K_r(a) \subset \mathcal{D}_h$, *és tetszőleges* $x \in K_r(a)$ *helyen léteznek a* $\partial_j h(x)$ $(i \neq j = 1, ..., n)$ *parciális deriváltak*;
- *mindegyik* $\partial_j h$ $(i \neq j = 1, ..., n)$ *parciális deriváltfüggvény folytonos az* a*-ban*;
- *létezik a* $\partial_i h(a)$ *parciális derivált.*

Ekkor $h \in D\{a\}$, *azaz a* h *függvény differenciáható az* a*-ban.*

Bizonyítás. Legyen a formai egyszerűség kedvéért $n = 2$ és $i = 1$. Ekkor ($\mathbf{R}^2$-ben a $\|.\|_\infty$ normát vezetve be) minden

$$x = (x_1, x_2) \in \mathbf{R}^2, \ \ \|x\|_\infty < r$$

vektorra a

$$h(a + x) - h(a)$$

különbséget olyan függvényértékek különbségeinek az összegeként fogjuk felírni, ahol az egyes különbségek argumentumaiban csak egyetlen koordinátában van változás: az

$$A := (a_1, a_2), \ B := (a_1 + x_1, a_2 + x_2), \ C := (a_1 + x_1, a_2)$$

jelölésekkel

$$h(a+x) - h(a) = h(B) - h(A) = h(C) - h(A) + h(B) - h(C).$$

Világos, hogy az itteni

$$h(C) - h(A)$$

első különbségben a h argumentumában az a-nak csak az első komponense változik, ezért ezt a különbséget „kezelhetjük" az a-beli, első változó szerinti paricális deriválhatóság alapján. Ugyanakkor a

$$h(B) - h(C)$$

eltérésben a h argumentuma csak a második koordinátában változik, így ezt a különbséget (a $K_r(a)$-ban feltételezett) második változó szerinti parciális deriválhatóságot figyelembe véve számolhatjuk.

Tehát

$$h(a+x) - h(a) = h(a_1 + x_1, a_2 + x_2) - h(a_1, a_2) =$$

$$h(a_1 + x_1, a_2) - h(a_1, a_2) + h(a_1 + x_1, a_2 + x_2) - h(a_1 + x_1, a_2).$$

Mivel feltételeztük, hogy létezik a $\partial_1 h(a)$ parciális derivált, ezért egy alkalmas

$$\omega \in \mathbf{R} \to \mathbf{R},\ \omega(t) \to 0 \qquad (t \to 0)$$

függvénnyel

$$h(a_1 + x_1, a_2) - h(a_1, a_2) =$$

$$\partial_1 h(a) \cdot x_1 + \omega(x_1) \cdot x_1 \qquad (|x_1| < r).$$

A $K_r(a)$ környezet minden $z \in K_r(a)$ pontjában létezik a $\partial_2 h(z)$ parciális derivált. Ez azt jelenti, hogy az $\mathbf{R} \to \mathbf{R}$ függvényekre vonatkozó Lagrange-középérték-tétel szerint egy

$$\xi_x \in [a_2, a_2 + x_2]$$

(vagy $\xi_x \in [a_2 + x_2, a_2]$) helyen

$$h(a_1 + x_1, a_2 + x_2) - h(a_1 + x_1, a_2) = \partial_2 h(a_1 + x_1, \xi_x) \cdot x_2.$$

Feltettük, hogy $\partial_2 h \in \mathcal{C}\{a\}$, ezért minden $\varepsilon > 0$ számhoz megadható olyan $0 < \delta < r$, hogy

$$|\partial_2 h(z) - \partial_2 h(a)| < \varepsilon \qquad (z \in K_\delta(a)),$$

így az $\|x\|_\infty < \delta$ választással

$$|\partial_2 h(a_1 + x_1, \xi_x) - \partial_2 h(a)| < \varepsilon.$$

Az

$$\omega(t) \to 0 \qquad (t \to 0)$$

miatt az is feltehető egyúttal, hogy $|\omega(x_1)| < \varepsilon$. Ha tehát

$$\varphi(x) := \Big(\omega(x_1), \partial_2 h(a_1 + x_1, \xi_x) - \partial_2 h(a)\Big) \qquad \Big(x \in K_r(0)\Big),$$

akkor

$$\|\varphi(x)\|_\infty < \varepsilon \qquad \Big(x \in K_\delta(0)\Big),$$

röviden:

$$\|\varphi(x)\|_\infty \to 0 \qquad (\|x\|_\infty \to 0).$$

Továbbá a

$$v := \Big(\partial_1 h(a), \partial_2 h(a)\Big) \in \mathbf{R}^2$$

vektorral

$$h(a + x) - h(a) = \langle v, x\rangle + \langle \varphi(x), x\rangle \qquad \Big(x \in K_r(0)\Big).$$

Legyen

$$\eta(x) := \langle \varphi(x), x/\|x\|_\infty\rangle \quad \Big(0 \neq x \in K_r(0)\Big), \ \ \eta(0) := 0,$$

ekkor

$$h(a + x) - h(a) = \langle v, x\rangle + \eta(x)\cdot \|x\|_\infty \qquad \Big(x \in K_r(0)\Big),$$

és a Cauchy-Bunyakovszkij-egyenlőtlenséget felhasználva

$$|\eta(x)| \leq \|\varphi(x)\|_2 \cdot \|x\|_2/\|x\|_\infty \leq 2\cdot \|\varphi(x)\|_\infty \to 0 \qquad (\|x\|_\infty \to 0)$$

miatt

$$\eta(x) \to 0 \qquad (\|x\|_\infty \to 0).$$

Összegezve a fentieket azt kaptuk, hogy $h \in D\{a\}$ és

$$h'(a) = \operatorname{grad} h(a) = v.$$

■

2. Megjegyzések.

i) Azt mondjuk, hogy az

$$f = (f_1, ..., f_m) \in \mathbf{R}^n \to \mathbf{R}^m$$

függvény *folytonosan differenciálható az* $a \in \mathcal{D}_f$ *pontban,* ha a következők teljesülnek:

- van az a-nak olyan $K(a) \subset \mathcal{D}_f$ környezete, hogy az $x \in K(a)$ helyeken $f \in D\{x\}$, és
- valamennyi $i = 1, ..., m$ esetén a $\text{grad} f_i$ függvény folytonos az a-ban.

Mindezt az $f \in C^1\{a\}$ szimbólummal fogjuk jelölni. Ha tetszőleges $a \in \mathcal{D}_f$ pontban $f \in C^1\{a\}$, akkor az f-et *folytonosan differenciálhatónak* (vagy *deriválhatónak*) nevezzük, és röviden azt írjuk, hogy $f \in C^1$. Mivel

$$\text{grad} f_i = (\partial_1 f_i, ..., \partial_n f_i) \qquad (i = 1, ..., m),$$

ezért a $\text{grad} f_i$ függvények a-beli folytonossága azzal ekvivalens, hogy

$$\partial_k f_i \in \mathcal{C}\{a\} \qquad (i = 1, ..., m;\ k = 1, ..., n).$$

ii) Adott pontban differenciálható $h \in \mathbf{R}^n \to \mathbf{R}$ függvény ugyanebben a pontban minden változó szerint parciálisan is deriválható. Egy egyszerű példa mutatja azt, hogy mindez fordítva nem igaz. Legyen ui. $n := 2$ és

$$f(x, y) := \begin{cases} 1 & (x = 0 \text{ vagy } y = 0) \\ \\ 0 & (xy \neq 0) \end{cases} \qquad \big((x, y) \in \mathbf{R}^2\big).$$

Világos, hogy ez a függvény nem folytonos az $a := (0, 0)$-ban, ezért $f \notin D\{a\}$. Ugyanakkor

$$\mathcal{D}_{f,i}^{(a)} = \mathbf{R} \qquad (i = 1, 2)$$

és

$$f_{a,i} \equiv 1 \qquad (i = 1, 2).$$

Ezért $f_{a,i} \in D\{0\}$ $(i = 1, 2)$ és

$$f'_{a,i}(0) = \partial_i f(0) = 0 \qquad (i = 1, 2).$$

iii) A parciális derivált fogalma speciális esete az ún. iránymenti deriváltnak. Legyen ennek az értelmezéséhez

$$1 \le n \in \mathbf{N} \ \text{ és } \ h \in \mathbf{R}^n \to \mathbf{R},$$

valamint $a \in \operatorname{int} \mathcal{D}_h$. Tegyük fel, hogy valamilyen

$$e \in \mathbf{R}^n, \ \|e\| = 1$$

vektor („irány") esetén (ahol $\|.\|$ a „szokásos" $\|.\|_p \quad (1 \le p \le +\infty)$ normák egyike) a

$$\mathcal{D}_{h,e}^{(a)} := \{t \in \mathbf{R} : a + te \in \mathcal{D}_h\}$$

jelöléssel a

$$h_e(t) := h(a + te) \qquad \left(t \in \mathcal{D}_{h,e}^{(a)}\right)$$

függvény[43] differenciálható a 0-ban. (Világos, hogy $0 \in \mathcal{D}_{h,e}^{(a)}$ minden h függvényre teljesül, ill. $a \in \operatorname{int} \mathcal{D}_h$ esetén $0 \in \operatorname{int} \mathcal{D}_{h,e}^{(a)}$ is igaz.) Ekkor a

$$\partial_e h(a) := h_e'(0)$$

számot a h függvény a-beli (az e által meghatározott) *iránymenti deriváltjának* nevezzük.

Nyilvánvaló, hogy ha $i = 1, ..., n$ és

$$e := (0, ..., 0, 1, 0, ..., 0) \in \mathbf{R}^n$$

(ahol tehát az e vektor i-edik koordinátája 1, a többi 0), akkor

$$h_e(t) = h_{a,i}(a_i + t) \qquad \left(t \in \mathcal{D}_{h,e}^{(a)}\right).$$

Következésképpen $h_e \in D\{0\}$ azzal ekvivalens, hogy $h_{a,i} \in D\{a_i\}$, azaz azzal, hogy létezik a $\partial_i h(a)$ parciális derivált és

$$\partial_e h(a) = h_e'(0) = h_{a,i}'(a_i) = \partial_i h(a).$$

[43]„Geometriailag" a h-nak az a-n átmenő e irányú egyenesre való leszűkítése.

3. Az összetett függvény differenciálhatósága.

2. Tétel. *Adott* $1 \leq n, m, s \in \mathbf{N}$ *mellett tekintsük a*

$$g \in \mathbf{R}^n \to \mathbf{R}^m, \ f \in \mathbf{R}^m \to \mathbf{R}^s$$

függvényeket. Tegyük fel, hogy az $a \in \mathcal{D}_g$ *pontban* $g(a) \in \mathcal{D}_f$, *valamint* $g \in D\{a\}$ *és* $f \in D\{g(a)\}$. *Ekkor* $f \circ g \in D\{a\}$, *és*

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

A bizonyítás előtt jegyezzük meg, hogy a tétel állításában szereplő $f'(g(a)) \cdot g'(a)$ mátrix az $f'(g(a)) \in \mathbf{R}^{s \times m}$ és a $g'(a) \in \mathbf{R}^{m \times n}$ mátrix szorzata, és mint ilyen, $\mathbf{R}^{s \times n}$-beli. Ezt is „várjuk", hiszen

$$f \circ g \in \mathbf{R}^n \to \mathbf{R}^s$$

miatt (ha $f \circ g \in D\{y\}$ valamilyen $y \in \mathcal{D}_{f \circ g}$ helyen)

$$(f \circ g)'(y) \in \mathbf{R}^{s \times n}.$$

Bizonyítás. Mivel $a \in \operatorname{int} \mathcal{D}_g$, $g(a) \in \operatorname{int} \mathcal{D}_f$, következésképpen megadhatók olyan r, $\delta > 0$ számok, amelyekkel

$$K_r(a) \subset \mathcal{D}_g, \ K_\delta(g(a)) \subset \mathcal{D}_f.$$

A g függvény folytonos is az a-ban, ezért az előbbi $r > 0$ sugárról az is feltehető, hogy

$$g[K_r(a)] \subset K_\delta(g(a)).$$

Tehát egyúttal $g[K_r(a)] \subset \mathcal{D}_f$ is teljesül. Ez azt jelenti, hogy $K_r(a) \subset \mathcal{D}_{f \circ g}$, más szóval $a \in \operatorname{int} \mathcal{D}_{f \circ g}$.

Írjuk fel a g, ill. az f megváltozását az a, ill. a $g(a)$ pont körül a differenciálhatósági feltételeknek megfelelően:

$$g(a+x) - g(a) = g'(a)x + \eta(x) \cdot \|x\| \qquad (x \in \mathbf{R}^n, \ a+x \in \mathcal{D}_g),$$

$$f(g(a)+y) - f((g(a)) = f'(g(a))y + \tilde{\eta}(y) \cdot \|y\| \qquad (y \in \mathbf{R}^m, \ g(a)+y \in \mathcal{D}_f),$$

ahol

$$\eta(x) \to \eta(0) = 0 \quad (\|x\| \to 0)$$

és

$$\tilde{\eta}(y) \to \tilde{\eta}(0) = 0 \quad (\|y\| \to 0).$$

Speciálisan az

$$y := g(a+x) - g(a) \qquad \big(x \in K_r(0)\big)$$

választással

$$(f \circ g)(a+x) - (f \circ g)(a) = f(g(a+x)) - f(g(a)) =$$

$$f'(g(a))\Big(g(a+x) - g(a)\Big) + \tilde{\eta}\Big(g(a+x) - g(a)\Big) \cdot \|g(a+x) - g(a)\| =$$

$$f'(g(a))\Big(g'(a)x + \eta(x) \cdot \|x\|\Big) + \tilde{\eta}\Big(g(a+x) - g(a)\Big) \cdot \|g'(a)x + \|x\| \cdot \eta(x)\| =$$

$$f'(g(a))g'(a)x + f'(g(a))\eta(x) \cdot \|x\| + \tilde{\eta}\Big(g(a+x) - g(a)\Big) \cdot \|g'(a)x + \|x\| \cdot \eta(x)\|.$$

Legyen

$$\varphi(x) :=$$

$$f'(g(a))\eta(x) + \frac{\tilde{\eta}\Big(g(a+x) - g(a)\Big) \cdot \|g'(a)x + \|x\| \cdot \eta(x)\|}{\|x\|} \qquad \Big(0 \neq x \in K_r(0)\Big),$$

ekkor

$$(f \circ g)(a+x) - (f \circ g)(a) = f'(g(a))g'(a)x + \varphi(x) \cdot \|x\| \qquad \Big(0 \neq x \in K_r(0)\Big).$$

Megmutatjuk, hogy $\varphi(x) \to 0 \quad (\|x\| \to 0)$. Ui. $\eta(x) \to 0 \quad (\|x\| \to 0)$ miatt

$$\|f'(g(a))\eta(x)\| \leq q \cdot \|\eta(x)\| \to 0 \qquad (\|x\| \to 0),$$

ahol q jelenti az $f'(g(a))$ mátrix normáját. Továbbá a g folytonos (is) az a-ban, ezért

$$g(a+x) - g(a) \to 0 \qquad (\|x\| \to 0),$$

amiből $\tilde{\eta}(y) \to \tilde{\eta}(0) = 0 \quad (\|y\| \to 0)$ miatt

$$\tilde{\eta}\Big(g(a+x) - g(a)\Big) \to 0 \qquad (\|x\| \to 0)$$

következik. Végül,

$$\frac{\Big\|\tilde{\eta}\Big(g(a+x) - g(a)\Big)\Big\| \cdot \Big\|g'(a)x + \|x\| \cdot \eta(x)\Big\|}{\|x\|} \leq$$

$$\frac{\Big\|\tilde{\eta}\Big(g(a+x) - g(a)\Big)\Big\| \cdot \Big(\Big\|g'(a)x\| + \|x\| \cdot \|\eta(x)\Big\|\Big)}{\|x\|} \leq$$

$$\Big\|\tilde{\eta}\Big(g(a+x) - g(a)\Big)\Big\| \cdot \Big(Q + \|\eta(x)\|\Big) \to 0 \qquad (\|x\| \to 0),$$

ahol Q a $g'(a)$ mátrix normáját jelöli. Mindez azt jelenti, hogy

$$\varphi(x) \to 0 \qquad (\|x\| \to 0)$$

valóban teljesül. Ezért az

$$(f \circ g)(a + x) - (f \circ g)(a)$$

megváltozásra kapott előbbi egyenlőség szerint $f \circ g \in D\{a\}$, és

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

■

Ha pl. a 2. Tételben $n = s = 1$, azaz

$$g \in \mathbf{R} \to \mathbf{R}^m,\ f \in \mathbf{R}^m \to \mathbf{R},$$

akkor

$$f \circ g \in \mathbf{R} \to \mathbf{R},$$

és $g \in D\{a\}$, $f \in D\{g(a)\}$ esetén

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

4. Többször differenciálható függvények.

Az

$$f \in \mathbf{R} \to \mathbf{R}$$

egyváltozós valós függvények körében „nem okozott gondot" az f függvény mondjuk 2-szer, 3-szor, ... való differenciálhatóságának az értelmezése. Pl. a szóban forgó f-re $f \in D^2\{a\}$, ha $f' \in D\{a\}$, és az utóbbi esetben az

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

deriváltat az f függvény a-beli második deriváltjának neveztük. Megjegyezzük, hogy az előbbi $f' \in D\{a\}$ feltételezésben az is „benne volt", hogy egy alkalmas $r > 0$ mellett

$$f \in D\{x\} \qquad (x \in \mathcal{D}_f,\ |x - a| < r).$$

Vegyük észre, hogy a 2-szer való differenciálhatóságnak ez az értelmezése átvihető az

$$f \in \mathbf{R}^n \to \mathbf{R}$$

függvényekre akkor is, ha $1 < n \in \mathbf{N}$. Ti., ha $f \in D$, akkor

$$f' = \operatorname{grad} f = \Big(\partial_1 f, ..., \partial_n f\Big) \in \mathbf{R}^n \to \mathbf{R}^n,$$

és az $\mathbf{R}^n \to \mathbf{R}^n$ típusú vektor-vektorfüggvények körében értelmeztük a deriválhatóságot. Tehát, az $f \in \mathbf{R}^n \to \mathbf{R}$ függvényre minden további nélkül előírhatjuk, hogy $f' \in D\{a\}$ teljesüljön. Más szóval legyen olyan $r > 0$, amellyel

$$K_r(a) \subset \mathcal{D}_f, \quad f \in D\{x\} \qquad \Big(x \in K_r(a)\Big)$$

és

$$f' = \operatorname{grad} f \in D\{a\}.$$

Ekkor

$$f''(a) := (\operatorname{grad} f)'(a) \in \mathbf{R}^{n \times n}.$$

Az is nyilvánvaló azonban, hogy ezen az úton nem tudunk eljutni a 2-nél magasabb rendű deriválhatósághoz, hiszen pl. az $\mathbf{R}^n \to \mathbf{R}^{n \times n}$ típusú függvények differenciálhatóságát nem értelmeztük.

Emlékeztetünk ugyanakkor arra, hogy az előbbi $f' = \operatorname{grad} f \in D\{a\}$ azzal ekvivalens, hogy

$$\partial_i f \in D\{a\} \qquad (i = 1, ..., n).$$

Éppen ezért az alábbiak szerint definiáljuk egy

$$f \in \mathbf{R}^n \to \mathbf{R}$$

függvény magasabb rendű differenciálhatóságát. Az értelmezés teljes indukcióval történik, első lépésként a kétszer deriválhatóságot definiálva.

Tegyük fel ehhez, hogy $a \in \operatorname{int} \mathcal{D}_f$, azaz egy alkalmas $K(a)$ környezettel $K(a) \subset \mathcal{D}_f$. Azt mondjuk, hogy az f függvény *kétszer differenciálható* (vagy *deriválható*) az a-ban (amit az $f \in D^2\{a\}$ szimbólummal jelölünk), ha minden $x \in K(a)$ esetén $f \in D\{x\}$, és

$$(*) \qquad \partial_i f \in D\{a\} \qquad (i = 1, ..., n).$$

A $(*)$ feltétel alapján léteznek a

$$\partial_j(\partial_i f)(a) \qquad (i, j = 1, ..., n)$$

parciális deriváltak. Ehhez persze nem szükséges, hogy a $\partial_i f$ $(i = 1, ..., n)$ függvények deriválhatók legyenek az a helyen: ha a $\partial_i f$ a j-edik változó szerint parciálisan differenciálható az a-ban, akkor az f-et *az i-edik és j-edik* *változó szerint kétszer parciálisan differenciálhatónak* (vagy *deriválhatónak*) nevezzük az a-ban. A $\partial_j(\partial_i f)(a)$ parciális deriváltat (mint a $\partial_i f \in \mathbf{R}^n \to \mathbf{R}$ függvény a-beli, a j-edik változó szerinti parciális deriváltját) a

$$\partial_i \partial_j f(a) := \partial_{ij} f(a) := \partial_j(\partial_i f)(a)$$

szimbólummal jelöljük, és az f függvény *a-beli, i-edik és j-edik változó szerinti másodrendű* (vagy *második*) *parciális deriváltjának* nevezzük. Esetenként használatos $\partial_{ij} f(a)$ helyett a $\partial_{x_i x_j} f(a)$, ill. a $\partial_{xy} f(a), \ldots$ jelölés is.

Tegyük fel, hogy $i, j = 1, \ldots, n$ és

$$\mathcal{D} := \{a \in \mathcal{D}_f : \text{ létezik } \partial_{ij} f(a)\} \neq \emptyset,$$

akkor értelmezhetjük a $\partial_{ij} f$-fel jelölt

$$\mathcal{D} \ni x \mapsto \partial_{ij} f(x)$$

függvényt, az *f függvény i-edik és j-edik változó szerinti másodrendű parciális deriváltfüggvényét.* Világos, hogy

$$\partial_{ij} f = \partial_j(\partial_i f)$$

Például, az

$$f(x, y, z) := x^3 + y^2 z + z^3 + xyz \qquad \Big((x, y, z) \in \mathbf{R}^3\Big)$$

függvényt véve bármelyik $(x, y, z) \in \mathbf{R}^3$ pontban

$$\partial_{11} f(x, y, z) = \partial_{xx} f(x, y, z) = 6x,\ \partial_{12} f(x, y, z) = \partial_{xy} f(x, y, z) = z,$$
$$\partial_{13} f(x, y, z) = \partial_{xz} f(x, y, z) = y,\ \partial_{22} f(x, y, z) = \partial_{yy} f(x, y, z) = 2z,$$
$$\partial_{23} f(x, y, z) = \partial_{yz} f(x, y, z) = 2y + x,\ \partial_{33} f(x, y, z) =$$
$$\partial_{zz} f(x, y, z) = 6z.$$

Ha tehát a fenti

$$f \in \mathbf{R}^n \to \mathbf{R}$$

függvényre $f \in D^2\{a\}$, akkor minden $i, j = 1, \ldots, n$ mellett létezik a $\partial_{ij} f(a)$ másodrendű parciális derivált. Az

$$f''(a) := \Big(\partial_{ij} f(a)\Big)_{i,j=1}^{n} = \begin{bmatrix} \partial_{11} f(a) & \partial_{12} f(a) & \ldots & \partial_{1n} f(a) \\ \partial_{21} f(a) & \partial_{22} f(a) & \ldots & \partial_{2n} f(a) \\ \vdots & \vdots & \ldots & \vdots \\ \partial_{n1} f(a) & \partial_{n2} f(a) & \ldots & \partial_{nn} f(a) \end{bmatrix} \in \mathbf{R}^{n \times n}$$

mátrixot az f függvény a-beli *második deriváltmátrixának* nevezzük.[44] A későbbiekben tárgyalandó Young-tétel miatt ez egy szimmetrikus mátrix. Például, az előbbiekben szereplő

$$f(x,y,z) := x^3 + y^2 z + z^3 + xyz \qquad \big((x,y,z) \in \mathbf{R}^3\big)$$

függvényre

$$f''(1,0,2) = \begin{bmatrix} 6 & 2 & 0 \\ 2 & 4 & 1 \\ 0 & 1 & 12 \end{bmatrix}.$$

4. Megjegyzések

i) Ha az

$$f \in \mathbf{R}^n \to \mathbf{R}^m \qquad (1 \le n, m \in \mathbf{N})$$

függvényre az $a \in \operatorname{int}\mathcal{D}_f$ helyen $f \in D\{a\}$ igaz, akkor tetszőleges $e \in \mathbf{R}^n$, $\|e\| = 1$ vektorral az

$$f_e(t) := f(a + te) \qquad (t \in \mathbf{R},\ a + te \in \mathcal{D}_f)$$

függvény (ld. 2. Tétel) differenciálható a 0-ban (azaz létezik az e által meghatározott iránymenti derivált), és

$$\partial_e f(a) = f_e'(0) = f'(a)e.$$

Az $m = 1$ esetben tehát

$$\partial_e f(a) = \langle \operatorname{grad} f(a), e \rangle.$$

ii) Tekintsük az alábbi függvényt:

$$f(x,y) := \begin{cases} 0 & (y \ne x^2) \\[2ex] 0 & (x = y = 0) \\[2ex] 1 & (y = x^2 \ne 0) \end{cases} \qquad \big((x,y) \in \mathbf{R}^2\big).$$

Ekkor tetszőleges $e \in \mathbf{R}^2$, $\|e\| = 1$ esetén létezik a $\partial_e f(0,0)$ iránymenti derivált, de $f \notin D\{(0,0)\}$. Ti. minden ilyen e-hez nyilván van olyan $0 < r \in \mathbf{R}$, hogy

$$f_e(t) = f(te) = 0 \qquad (|t| < r).$$

[44] Ez nem más, mint a $\operatorname{grad} f$ függvény a-beli Jacobi-mátrixa.

Innen az f_e függvény 0-beli differenciálhatósága, valamint

$$\partial_e f(0,0) = f_e'(0) = 0$$

már nyilvánvaló. Ugyanakkor

$$f(t,t^2) = 1 \qquad (0 \neq t \in \mathbf{R}),$$

így

$$f(t,t^2) \to 1 \neq 0 = f(0,0) \qquad (t \to 0)$$

miatt az f függvény nem folytonos a $(0,0)$-ban, tehát ugyanitt nem is differenciálható.

iv) Határozzuk meg azt az $f : \mathbf{R}^2 \to \mathbf{R}$ differenciálható függvényt, amelyre

$$\partial_x f(x,y) = x^2 y,\ \partial_y f(x,y) = \frac{x^3}{3} + 1 \qquad \big((x,y) \in \mathbf{R}^2\big).$$

Tetszőleges $a = (x,y) \in \mathbf{R}^2$ mellett

$$\mathcal{D}_{f,1}^{(a)} = \{t \in \mathbf{R} : (t,y) \in \mathcal{D}_f\} = \mathcal{D}_{f,2}^{(a)} = \{t \in \mathbf{R} : (x,t) \in \mathcal{D}_f\} = \mathbf{R}$$

és

$$f_{a,1}(t) = f(t,y),\ f_{a,2}(t) = f(x,t) \qquad (t \in \mathbf{R}),$$

ill.

$$f_{a,1}'(t) = \partial_1 f(t,y) = t^2 y,\ f_{a,2}'(t) = \partial_2 f(x,t) = \frac{x^3}{3} + 1 \qquad (t \in \mathbf{R}).$$

Ezért alkalmas

$$\varphi,\ \psi : \mathbf{R} \to \mathbf{R}$$

differenciálható függvényekkel

$$f_{a,1}(t) = f(t,y) = \frac{t^3 y}{3} + \varphi(y) \qquad (t \in \mathbf{R}),$$

$$f_{a,2}(t) = f(x,t) = \frac{x^3 t}{3} + t + \psi(x) \qquad (t \in \mathbf{R}).$$

Tehát

$$f(x,y) = f_{a,1}(x) = \frac{x^3 y}{3} + \varphi(y) = f_{a,2}(y) = \frac{x^3 y}{3} + y + \psi(x).$$

Innen az adódik, hogy

$$\psi(x) = \varphi(y) - y \qquad (x,y \in \mathbf{R}).$$

Világos, hogy ez akkor és csak akkor lehetséges, ha valamilyen $c \in \mathbf{R}$ konstanssal

$$\varphi(y) - y = c = \psi(x) \qquad (x, y \in \mathbf{R}).$$

Más szóval

$$f(x,y) = \frac{x^3 y}{3} + y + c \qquad \Big((x,y) \in \mathbf{R}^2\Big).$$

v) Legyen $a, b \in \mathbf{R}$, $a \neq 0$, és

$$f(x,t) := \frac{1}{2a\sqrt{\pi t}} \cdot \exp\left(-\frac{(x-b)^2}{4a^2 t}\right) \qquad \Big((x,t) \in \mathbf{R} \times (0, +\infty)\Big).$$

A $\partial_t f$, $\partial_{xx} f$ parciális deriváltfüggvények kiszámításával könnyen ellenőrizhető, hogy

$$\partial_t f = a^2 \cdot \partial_{xx} f$$

(ez a fizika egyik „alapegyenlete", az ún. *hővezetési egyenlet*).

9. Fejezet

0. Emlékeztető.

I) Legyen adott az

$$f \in \mathbf{R}^n \to \mathbf{R}$$

függvény. Tegyük fel, hogy $a \in \operatorname{int} \mathcal{D}_f$. Azt mondjuk, hogy az f függvény *kétszer differenciálható* az a-ban ($f \in D^2\{a\}$), ha egy alkalmas $K(a)$ környezettel $K(a) \subset \mathcal{D}_f$, minden $x \in K(a)$ esetén $f \in D\{x\}$, és

$$\partial_i f \in D\{a\} \qquad (i = 1, ..., n).$$

II) Ekkor léteznek a

$$\partial_j(\partial_i f)(a) \qquad (i, j = 1, ..., n)$$

parciális deriváltak. Ehhez nem szükséges, hogy a $\partial_i f$ $(i = 1, ..., n)$ függvények deriválhatók legyenek az a helyen: ha a $\partial_i f$ a j-edik változó szerint parciálisan differenciálható az a-ban, akkor az f-et *az i-edik és j-edik változó szerint kétszer parciálisan differenciálhatónak* nevezzük az a-ban. A $\partial_j(\partial_i f)(a)$ parciális deriváltat (mint a $\partial_i f \in \mathbf{R}^n \to \mathbf{R}$ függvény a-beli, a j-edik változó szerinti parciális deriváltját) a

$$\partial_i \partial_j f(a) := \partial_{ij} f(a) := \partial_j(\partial_i f)(a)$$

szimbólummal jelöljük, és az f függvény *a-beli, i-edik és j-edik változó szerinti másodrendű* (vagy *második*) *parciális deriváltjának* nevezzük. Esetenként használatos $\partial_{ij}f(a)$ helyett a $\partial_{x_ix_j}f(a)$, ill. a $\partial_{xy}f(a), \ldots$ jelölés is.

III) Tegyük fel, hogy $i, j = 1, \ldots, n$ és

$$\mathcal{D} := \{a \in \mathcal{D}_f : \text{ létezik } \partial_{ij}f(a)\} \neq \emptyset,$$

akkor értelmezhetjük a $\partial_{ij}f$-fel jelölt

$$\mathcal{D} \ni x \mapsto \partial_{ij}f(x)$$

függvényt, az *f függvény i-edik és j-edik változó szerinti másodrendű parciális deriváltfüggvényét.* Világos, hogy

$$\partial_{ij}f = \partial_j(\partial_i f)$$

1. Differenciálhatóság.

Értelmezzük most a parciális differenciálhatóságot kettőnél magasabb rendre. Legyen tehát

$$1 \leq n, \ f \in \mathbf{R}^n \to \mathbf{R},$$

és a teljes indukció módszerét alkalmazva tegyük fel, hogy valamilyen

$$1 \leq s \in \mathbf{N}, \ 1 \leq i_1, \ldots, i_s \leq n$$

mellett már definiáltuk a $\partial_{i_1 \ldots i_s}f$ szimbólummal jelölt s-edrendű parciális deriváltfüggvényt. Ha $a \in \mathcal{D}_f$ és egy $j = 1, \ldots, n$ indexre a

$$\partial_{i_1 \ldots i_s}f \in \mathbf{R}^n \to \mathbf{R}$$

függvény parciálisan differenciálható az a-ban a j-edik változó szerint, akkor legyen

$$\partial_{i_1} \ldots \partial_{i_s}\partial_j f(a) := \partial_{i_1 \ldots i_s j}f(a) := \partial_j(\partial_{i_1 \ldots i_s}f)(a)$$

az *f függvény a-beli $(s+1)$-edrendű parciális deriváltja az i_1-edik, $\ldots$ i_s-edik, j-edik változó szerint.* Ha az ilyen tulajdonságú $a \in \mathcal{D}_f$ pontok $\mathcal{X}$ halmaza nem üres, akkor az

$$\mathcal{X} \ni x \mapsto \partial_{i_1 \ldots i_s j}f(x)$$

leképezést az *f függvény* (a jelzett változók szerinti) *$(s+1)$-edrendű parciális deriváltfüggvényének* nevezzük. A változók indexeinek a felsorolása helyett továbbra is használhatjuk magukat a változókat jelölő szimbólumokat. Például, a már többször vizsgált

$$f(x,y,z) := x^3 + y^2 + z^3 + xyz \qquad \big((x,y,z) \in \mathbf{R}^3\big)$$

függvény esetén harmadrendű parciális deriváltak a következők:

$$\partial_{111} f(x,y,z) = \partial_{xxx} f(x,y,z) = 6\ ,\ \partial_{121} f(x,y,z) = \partial_{xyx} f(x,y,z) = 0,$$

$$\partial_{123} f(x,y,z) = \partial_{xyz} f(x,y,z) = 1 \qquad \big((x,y,z) \in \mathbf{R}^3\big).$$

A parciális deriváltak jelölésére is használatosak az alábbi, a *differenciálhányados* szóból a *hányados*ra utaló szimbólumok:

$$\frac{\partial f(a)}{\partial x},\ \frac{\partial^2 f(a)}{\partial x \partial y},\ \frac{\partial^2 f(a)}{\partial x^2},\ \frac{\partial^3 f(a)}{\partial x^2 \partial y},\ \ldots$$

Minden készen áll ahhoz, hogy az

$$f \in \mathbf{R}^n \to \mathbf{R} \qquad (1 \le n \in \mathbf{N})$$

függvények magasabb rendű differenciálhatóságát definiálhassuk. Az előbbi módszert alkalmazva ui. tegyük fel ehhez, hogy

$$a \in \operatorname{int} \mathcal{D}_f,\ 1 \le s \in \mathbf{N},$$

továbbá egy alkalmas $K(a)$ környezettel $K(a) \subset \mathcal{D}_f$, és minden $x \in K(a)$ pontban az f függvény s-szer differenciálható: $f \in D^s\{x\}$. Belátható, hogy ekkor a $K(a)$ pontjaiban az f összes s-edrendű parciális deriváltja létezik. Azt mondjuk, hogy *az f függvény az a-ban $(s+1)$-szer differenciálható,* ha minden s-edrendű parciális deriváltfüggvénye differenciálható az a-ban.

Legyen most $1 \le n, m \in \mathbf{N}$ és

$$f = (f_1, ..., f_m) \in \mathbf{R}^n \to \mathbf{R}^m,\ a \in \operatorname{int} \mathcal{D}_f,$$

ill. $1 \le k \in \mathbf{N}$. Azt mondjuk, hogy az f függvény *k-szor differenciálható az a-ban,* ha

$$f_j \in D^k\{a\} \qquad (j = 1, ..., m).$$

A „szokásos" $f \in D^k\{a\}$ jelölést fogjuk használni. Ha tetszőleges $a \in \mathcal{D}_f$ pontban teljesül, hogy $f \in D^k\{a\}$, akkor az f függvény *k-szor differenciálható*: $f \in D^k$.

Az előbbi

$$f \in \mathbf{R}^n \to \mathbf{R}^m$$

függvény *k-szor folytonosan differenciálható az a-ban,* ha egy alkalmas $K(a) \subset \mathcal{D}_f$ környezettel

$$f \in D^k\{x\} \qquad \big(x \in K(a)\big),$$

és minden $f_j \quad (j = 1, ..., m)$ koordinátafüggvény összes k-adrendű parciális deriváltfüggvénye folytonos az a-ban. Minderre az $f \in \mathcal{C}^k\{a\}$ jelölést alkalmazzuk. Ha $f \in \mathcal{C}^k\{a\}$ minden $a \in \mathcal{D}_f$ helyen igaz, akkor röviden azt mondjuk, hogy az f függvény *k-szor folytonosan differenciálható,* és ezt az $f \in \mathcal{C}^k$ módon jelöljük.

2. Young-tétel.

A magasabb rendű parciális deriváltakkal kapcsolatban joggal merül fel a kérdés, hogy azok kiszámításakor van-e szerepe a változók sorrendjének? Pl. tegyük fel, hogy $2 \le n \in \mathbf{N}$, és egy

$$f \in \mathbf{R}^n \to \mathbf{R}$$

függvény esetén valamilyen $a \in \mathcal{D}_f$ helyen léteznek a $\partial_{12} f(a)$, $\partial_{21} f(a)$ másodrendű parciális deriváltak. Igaz-e, hogy

$$\partial_{12} f(a) = \partial_{21} f(a)?$$

Az alábbi példa azt igazolja, hogy minden további nélkül ez nem teljesül. Legyen ui.

$$f(x, y) := \begin{cases} xy \cdot \dfrac{x^2 - y^2}{x^2 + y^2} & (x^2 + y^2 \neq 0) \\ 0 & (x = y = 0) \end{cases} \qquad \big((x, y) \in \mathbf{R}^2\big).$$

Megmutatjuk, hogy

$$\partial_{12} f(0, 0) = -1, \quad \partial_{21} f(0, 0) = 1.$$

A $\partial_{12} f(0, 0) = -1$ egyenlőség igazolásához induljunk ki a szóban forgó másodrendű parciális derivált definíciójából:

$$\partial_{12} f(0, 0) = \partial_2(\partial_1 f)(0, 0) \,\Big(= \partial_y(\partial_x f)(0, 0)\Big) = \varphi'(0),$$

ahol

$$\varphi(t) := \partial_1 f(0, t) \qquad (t \in \mathbf{R}).$$

A

$$\varphi : \mathbf{R} \to \mathbf{R}$$

függvény kiszámítását hasonlóan végezhetjük: $\varphi(t) = \psi_t'(0)$, ahol $t \in \mathbf{R}$ esetén

$$\psi_t(\tau) := f(\tau, t) \qquad (\tau \in \mathbf{R}).$$

Ha $t = 0$, akkor $\psi_0(\tau) = 0 \quad (\tau \in \mathbf{R})$, ezért $\varphi(0) = 0$. A $t \neq 0$ esetben viszont

$$\varphi(t) = \psi_t'(0) = \lim_{\tau \to 0} \frac{\psi_t(\tau) - \psi_t(0)}{\tau} = \lim_{\tau \to 0} \frac{f(\tau, t)}{\tau} = \lim_{\tau \to 0} \frac{t \cdot (\tau^2 - t^2)}{\tau^2 + t^2} = -t.$$

Tehát

$$\varphi(t) = -t \qquad (t \in \mathbf{R}),$$

következésképpen

$$\partial_{12} f(0,0) = \varphi'(0) = -1.$$

Hasonlóan kapjuk a $\partial_{21} f(0,0) = 1$ egyenlőséget.

Ha viszont az f függvény a szóban forgó a helyen „elég szokszor" differenciálható, akkor a fent említett sorrend elveszti a jelentőségét. Ezzel kapcsolatos az analízis egyik legnevezetesebb állítása, az ún. *Young-tétel.*

1. Tétel (Young[45]). *Legyen* $2 \leq n \in \mathbf{N}$, $f \in \mathbf{R}^n \to \mathbf{R}$, $a \in \operatorname{int} \mathcal{D}_f$, $2 \leq s \in \mathbf{N}$ *és* $f \in D^s\{a\}$. *Ekkor tetszőleges* $k_1, \ldots, k_s \in \{1, ..., n\}$ *indexek esetén ezek bármely* $j_1, \ldots, j_s$ *permutációjára*

$$\partial_{k_1 \ldots k_s} f(a) = \partial_{j_1 \ldots j_s} f(a).$$

Bizonyítás (vázlat). Az s-szerinti teljes indukcióra gondolva elegendő az $s = 2$ esettel foglalkoznunk. Ekkor tehát azt kell belátnunk, hogy ha $f \in D^2\{a\}$, akkor

$$\partial_{ij} f(a) = \partial_{ji} f(a) \qquad (i, j = 1, ..., n).$$

Világos, hogy csak az $i \neq j$ eset az „érdekes". Ezen túl (könnyen meggondolhatóan) azt is feltehetjük, hogy $n = 2$. Más szóval az

$$f \in \mathbf{R}^2 \to \mathbf{R}$$

függvényre

$$a = (a_1, a_2) \in \operatorname{int} \mathcal{D}_f, \ f \in D^2\{a\},$$

[45] William Henry Young (1863–1942).

és ennek alapján azt kell bebizonyítanunk, hogy

$$\partial_{12}f(a) = \partial_{21}f(a).$$

Legyen ehhez $r > 0$ olyan, amellyel ($\mathbf{R}^n$-ben a $\|.\| := \|.\|_\infty$ normát választva)

$$K(a) = \{x \in \mathbf{R}^2 : \|x - a\| < r\} \subset \mathcal{D}_f,$$

és vezessük be az alábbi jelölést: az $u, v \in (-r, r)$ helyeken

$$\Delta(u, v) := f(a_1 + u, a_2 + v) - f(a_1 + u, a_2) + f(a_1, a_2) - f(a_1, a_2 + v).$$

Ha rögzítjük a $v \in (-r, r)$ számot, akkor a

$$\varphi(u) := f(a_1 + u, a_2 + v) - f(a_1 + u, a_2) \qquad \big(u \in (-r, r)\big)$$

függvénnyel

$$\Delta(u, v) = \varphi(u) - \varphi(0) \qquad \big(u \in (-r, r)\big).$$

Az $f \in D^2\{a\}$ feltétel miatt az előbbi $K_r(a)$ környezettől azt is megkövetelhetjük, hogy egyrészt minden $x \in K_r(a)$ helyen $f \in D\{x\}$ (így egyúttal léteznek a $\partial_1 f(x)$, $\partial_2 f(x)$ parciális deriváltak is), másrészt

$$\partial_1 f,\ \partial_2 f \in D\{a\}.$$

Következésképpen a most definiált

$$\varphi : (-r, r) \to \mathbf{R}$$

függvény differenciálható, ezért a Lagrange-középérték-tétel alapján

$$\varphi(u) - \varphi(0) = \varphi'(\xi)\cdot u \qquad \big(u \in (-r, r)\big),$$

ahol $\xi \in (0, u)$ (vagy $\xi \in (u, 0)$). A parciális deriváltak definíciójára gondolva

$$\varphi'(u) = \partial_1 f(a_1 + u, a_2 + v) - \partial_1 f(a_1 + u, a_2) \qquad \big(u \in (-r, r)\big),$$

így

$$\varphi(u) - \varphi(0) = \big(\partial_1 f(a_1 + \xi, a_2 + v) - \partial_1 f(a_1 + \xi, a_2)\big)\cdot u \qquad \big(u \in (-r, r)\big).$$

A $\partial_1 f \in D\{a\}$ differenciálhatósági feltételből

$$\operatorname{grad} \partial_1 f(a) = \big(\partial_{11} f(a), \partial_{12} f(a)\big),$$

és egy alkalmas

$$\eta \in \mathbf{R}^2 \to \mathbf{R},\ \eta(z) \to 0 \qquad (\|z\| \to 0)$$

függvénnyel

$$\partial_1 f(a_1+\xi, a_2+v) - \partial_1 f(a_1+\xi, a_2) =$$
$$\partial_1 f(a_1+\xi, a_2+v) - \partial_1 f(a_1, a_2) - \Big(\partial_1 f(a_1+\xi, a_2) - \partial_1 f(a_1, a_2)\Big) =$$
$$\langle \operatorname{grad} \partial_1 f(a), (\xi, v)\rangle + \eta(\xi, v)\cdot \|(\xi, v)\| - \langle \operatorname{grad} \partial_1 f(a), (\xi, 0)\rangle - \eta(\xi, 0)\cdot \|(\xi, 0)\| =$$
$$\partial_{12} f(a)\cdot v + \eta(\xi, v)\cdot \|(\xi, v)\| - \eta(\xi, 0)\cdot |\xi|.$$

Speciálisan a $0 \neq u = v \in (-r, r)$ választással

$$\Delta(u, u) = \varphi(u) - \varphi(0) = \partial_{12} f(a)\cdot u^2 + \eta(\xi, u)\cdot \|(\xi, u)\|\cdot u - \eta(\xi, 0)\cdot |\xi|\cdot u,$$

amiből

$$\frac{\Delta(u, u)}{u^2} = \partial_{12} f(a) + \eta(\xi, u)\cdot \frac{\|(\xi, u)\|}{u} - \eta(\xi, 0)\cdot \frac{|\xi|}{u}$$

következik. Ezért $|\xi| < |u|$ alapján

$$\left| \frac{\Delta(u, u)}{u^2} - \partial_{12} f(a) \right| \leq |\eta(\xi, u)| + |\eta(\xi, 0)| \to 0 \qquad (u \to 0),$$

hiszen

$$\|(\xi, u)\|,\ \|(\xi, 0)\| \leq |u| \to 0 \qquad (u \to 0).$$

Azt kaptuk ezzel, hogy

$$(*) \qquad \partial_{12} f(a) = \lim_{u \to 0} \frac{\Delta(u, u)}{u^2}.$$

Legyen most rögzített $u \in (-r, r)$ mellett

$$\psi(v) := f(a_1+u, a_2+v) - f(a_1, a_2+v) \qquad \Big(v \in (-r, r)\Big).$$

Ekkor

$$\Delta(u, v) = \psi(v) - \psi(0) \qquad \Big(v \in (-r, r)\Big)$$

és az előbbiekkel analóg módon az adódik , hogy

$$\partial_{21} f(a) = \lim_{v \to 0} \frac{\Delta(v, v)}{v^2}.$$

Itt a jobb oldali limesz ugyanaz, mint a $(*)$-ban. Így $\partial_{21} f(a) = \partial_{12} f(a)$. ∎

3. Taylor-formula.

Emlékeztető: legyen valamilyen

$$g \in \mathbf{R} \to \mathbf{R}$$

függvény, $a \in \operatorname{int} \mathcal{D}_g$ és $r > 0$ szám esetén $(a-r, a+r) \subset \mathcal{D}_g$. Tegyük fel, hogy $s \in \mathbf{N}$ és

$$g \in D^{(s+1)}\{x\} \qquad (x \in (a-r, a+r)).$$

Ekkor bármely $a \neq z \in (a-r, a+r)$ esetén van olyan $\xi \in (a, z)$ (vagy $\xi \in (z, a)$), amellyel

$$g(z) - T_{a,s}g(z) = \frac{g^{(s+1)}(\xi)}{(s+1)!}(z-a)^{s+1},$$

ahol

$$T_{a,s}g(x) := \sum_{k=0}^{s} \frac{g^{(k)}(a)}{k!}(x-a)^k \qquad (x \in \mathbf{R})$$

a g függvény a-beli s-edik Taylor-polinomja.

A továbbiakban az alábbi jelöléseket fogjuk használni (feltételezve egyben az azokban szereplő parciális deriváltak létezését). Legyen

$$1 \leq n,\ k \in \mathbf{N},\ f \in \mathbf{R}^n \to \mathbf{R},\ a \in \operatorname{int} \mathcal{D}_f,\ j = 1, ..., n,$$

és

$$\partial_j^k f(a) := \partial_{j...j} f(a)$$

(ahol a „$j \dots j$" rövidítés k darab j-t jelöl). Speciálisan $\partial_j^1 f(a) = \partial_j f(a)$. Állapodjunk meg abban, hogy

$$\partial_j^0 f(a) := f(a).$$

Ha $i := (i_1, ..., i_n) \in \mathbf{N}^n$, akkor legyen

$$\partial^i f(a) := \partial_1^{i_1} \dots \partial_n^{i_n} f(a).$$

Részletesen kiírva tehát

$$\partial^i f(a) = \partial_{1...1...n...n} f(a),$$

ahol i_1 darab 1-es, ..., i_n darab n-es szerepel. Például, ha

$$n := 3,\ i := (2, 0, 1),$$

valamint

$$f(x,y,z) := x^3 + y^2 z + z^3 + xyz \qquad \Big((x,y,z) \in \mathbf{R}^3\Big),$$

akkor

$$\partial^i f(x,y,z) = \partial_1^2 \partial_3 f(x,y,z) = \partial_{113} f(x,y,z) = 0 \qquad \Big((x,y,z) \in \mathbf{R}^3\Big).$$

Az $i \in \mathbf{N}^n$ *multiindex*[46] esetén az i *hosszát* a következőképpen definiáljuk:

$$|i| := \|i\|_1 = \sum_{j=1}^n i_j.$$

A korábban mondottak szerint pl. $f \in D^{|i|}\{a\}$ elégséges ahhoz, hogy létezzen az $|i|$-edrendű $\partial^i f(a)$ parciális derivált, és ekkor a Young-tétel miatt az $|i|$ hosszúságú $1 \ldots 1 \ldots n \ldots n$ jelsorozat (ld. fent) bármely $\nu_1 \ldots \nu_{|i|}$ permutációjára

$$\partial^i f(a) = \partial_{1 \ldots 1 \ldots n \ldots n} f(a) = \partial_{\nu_1 \ldots \nu_{|i|}} f(a).$$

Legyen az $i = (i_1, ..., i_n) \in \mathbf{N}^n$ multiindex és az $x = (x_1, ..., x_n) \in \mathbf{R}^n$ vektor esetén

$$i! := \prod_{j=1}^n i_j! = \prod_{j=1}^n \prod_{k=1}^{i_j} k$$

és

$$x^i := \prod_{j=1}^n x_j^{i_j}$$

az i *faktoriálisa,* ill. az x vektor i-kitevős *hatványa.* Nyilvánvaló, hogy ha $n = 1$, akkor $i = i_1 \in \mathbf{N}$ és $|i| = i$, ill.

$$i! = \prod_{j=1}^i j,$$

az $x^i \quad (x \in \mathbf{R})$ pedig a „megszokott" hatvány.

Ha $a, b \in \mathbf{R}^n$, akkor az a és b végpontú zárt, ill. $a \neq b$ esetén nyílt *szakaszt* az

$$[a,b] := \{a + t(b-a) \in \mathbf{R}^n : 0 \leq t \leq 1\},$$

ill.

$$(a,b) := \{a + t(b-a) \in \mathbf{R}^n : 0 < t < 1\}$$

módon definiáljuk.

---

[46]Laurent Moise Schwartz (1915 – 2002).

Tekintsük ezek után az

$$f \in \mathbf{R}^n \to \mathbf{R}$$

függvényt, és tegyük fel, hogy valamilyen

$$a \in \operatorname{int} \mathcal{D}_f, \ s \in \mathbf{N}$$

mellett $f \in D^s\{a\}$ (ahol – mint korábban is – a $D^0\{a\} := \mathcal{C}\{a\}$ megállapodással élünk). Ekkor a

$$T_{a,s}f(x) := \sum_{k=0}^{s} \sum_{i \in \mathbf{N}^n, |i|=k} \frac{\partial^i f(a)}{i!} \cdot (x-a)^i \qquad (x \in \mathbf{R}^n)$$

előírással definiált

$$T_{a,s}f : \mathbf{R}^n \to \mathbf{R}$$

függvényt az f függvény a-hoz tartozó s-edrendű *Taylor-polinomjának* nevezzük. Világos, hogy $T_{a,0}f \equiv f(a)$, ill. $T_{a,s}f(a) = f(a)$.

2. Tétel (Taylor[47]-formula). *Tegyük fel, hogy az $f \in \mathbf{R}^n \to \mathbf{R}$ függvényre valamilyen $s \in \mathbf{N}$ mellett $f \in D^{s+1}$ teljesül. Ekkor bármely $[a,b] \subset \mathcal{D}_f$ esetén van olyan $c \in [a,b]$, hogy*

$$f(b) = T_{a,s}f(b) + \sum_{i \in \mathbf{N}^n, |i|=s+1} \frac{\partial^i f(c)}{i!} \cdot (b-a)^i.$$

Bizonyítás (vázlat). Feltehető, hogy $a \neq b$, legyen ekkor $h := b-a$ és

$$F(t) := f(a+th) \qquad (t \in \mathbf{R},\, a+th \in \mathcal{D}_f).$$

Az $[a,b] \subset \mathcal{D}_f$ feltétel miatt $[0,1] \subset \mathcal{D}_F$. Mivel a

$$g(t) := a+th \qquad (t \in \mathbf{R},\, a+th \in \mathcal{D}_f)$$

függvényre $g \in D^{s+1}$ triviális módon igaz, így

$$F = f \circ g \in D^{s+1}$$

következik. Ezért az $F \in \mathbf{R} \to \mathbf{R}$ függvényre alkalmazható az „egyváltozós" Taylor-formula, miszerint egy alkalmas $\xi \in (0,1)$ helyen

$$f(b) = F(1) = \sum_{k=0}^{s} \frac{F^{(k)}(0)}{k!} + \frac{F^{(s+1)}(\xi)}{(s+1)!}.$$

[47]Brook Taylor (1685–1731).

Elég tehát azt megmutatni, hogy

$$(*) \qquad \frac{F^{(k)}(t)}{k!} = \sum_{i\in\mathbf{N}^n, |i|=k} \frac{\partial^i f(a+th)}{i!}\cdot h^i \qquad (t\in\mathcal{D}_F,\ k=0,...,s+1),$$

amiből következően a $c := a+\xi\cdot h$ pont megfelel az állításunknak.

Ha itt $k=0$, akkor

$$\frac{F^{(0)}(t)}{0!} = F(t) = f(a+th) = \sum_{i\in\mathbf{N}^n, |i|=0} \frac{\partial^i f(a+th)}{i!}\cdot h^i \qquad (t\in\mathcal{D}_F)$$

nyilván igaz, ui. $i=(0,\ldots,0)\in\mathbf{N}^n$ az egyetlen olyan multiindex, amelyre $|i|=0$ és

$$\partial^{(0,\ldots,0)} f(a+th) = f(a+th),\ h^i = 1,\ i! = 1.$$

Teljes indukcióval folytatva a bizonyítást tegyük fel, hogy a $(*)$ formula valamilyen $k=0,...,s$ esetén igaz, és lássuk be ugyanezt $(k+1)$-re. Ekkor tehát

$$\frac{F^{(k)}}{k!} = \sum_{i\in\mathbf{N}^n, |i|=k} \frac{\partial^i f\circ g}{i!}\cdot h^i,$$

ahol a feltételeink szerint minden itt szereplő i multiindexre $\partial^i f\circ g\in D$. Ezért

$$\frac{F^{(k+1)}(t)}{(k+1)!} = \frac{1}{k+1}\cdot\left(\frac{F^{(k)}}{k!}\right)'(t) = \frac{1}{k+1}\cdot\sum_{i\in\mathbf{N}^n, |i|=k} \frac{\left(\partial^i f\circ g\right)'(t)}{i!}\cdot h^i =$$

$$\frac{1}{k+1}\cdot\sum_{i\in\mathbf{N}^n, |i|=k} \frac{\langle \operatorname{grad}\partial^i f(g(t)), g'(t)\rangle}{i!}\cdot h^i =$$

$$\frac{1}{k+1}\cdot\sum_{i\in\mathbf{N}^n, |i|=k} \frac{\langle \operatorname{grad}\partial^i f(a+th), h\rangle}{i!}\cdot h^i =$$

$$\frac{1}{k+1}\cdot\sum_{i\in\mathbf{N}^n, |i|=k}\sum_{j=1}^{n} \frac{\partial_j\partial^i f(a+th)}{i!}\cdot h_j h^i$$

Ha $i\in\mathbf{N}^n$, $|i|=k$ és

$$i^{(j)} := (i_1,\ldots,i_{j-1}, i_j+1, i_{j+1},\ldots,i_n)\in\mathbf{N}^n \qquad (j=1,...,n),$$

akkor nyilván

$$\partial_j\partial^i f(a+th) = \partial^{i^{(j)}} f(a+th),$$

$$|i^{(j)}| = k+1, \ \ i^{(j)}! = i! \cdot (i_j+1), \ \ h_j h^i = h^{i^{(j)}}.$$

Összeszámolva, hogy egy

$$l = (l_1, ..., l_n) \in \mathbf{N}^n, \ |l| = k+1$$

multiindexet megadva hányféleképpen lehetséges az

$$l = i^{(j)}$$

előállítás alkalmas $i \in \mathbf{N}^n$, $|i| = k$, valamint $j = 1, ..., n$ esetén, kapjuk:

$$\frac{F^{(k+1)}(t)}{(k+1)!} = \frac{1}{k+1} \cdot \sum_{i \in \mathbf{N}^n, |i|=k} \sum_{j=1}^{n} \frac{\partial_j \partial^i f(a+th)}{i!} \cdot h_j h^i =$$

$$\sum_{l \in \mathbf{N}^n, |l|=k+1} \frac{\partial^l f(a+th)}{l!} \cdot h^l.$$

Ez nem más, mint a $(*)$ egyenlőség k helyett $(k+1)$-re. ■

Alkalmazzuk a 2. Tételt az $s = 0$ esetben:

3. Tétel (Lagrange[48]). *Legyen adott a differenciálható* $f \in \mathbf{R}^n \to \mathbf{R}$ *függvény, és valamilyen* $a, b \in \mathbf{R}^n$, $a \neq b$ *végpontokkal* $[a, b] \subset \mathcal{D}_f$. *Ekkor egy alkalmas* $c \in (a, b)$ *mellett*

$$f(b) - f(a) = \langle \operatorname{grad} f(c), b-a \rangle.$$

Ekkor ui. $T_{a,0} f \equiv f(a)$, és $i \in \mathbf{N}^n, |i| = 1$ nyilván azzal ekvivalens, hogy egy $j = 1, ..., n$ indexszel

$$i_j = 1, \ i_\nu = 0 \qquad (j \neq \nu = 1, ..., n).$$

Ezért a 2. Tételben most

$$\sum_{i \in \mathbf{N}, |i|=1} \frac{\partial^i f(c)}{i!} \cdot (b-a)^i = \sum_{j=1}^{n} \partial_j f(c) \cdot (b_j - a_j) = \langle \operatorname{grad} f(c), b-a \rangle.$$

Megjegyezzük, hogy a 3. Tétel nyilván az egyváltozós valós függvények körében megismert *Lagrange-középérték-tétel* többváltozós megfelelője.

[48] Joseph–Louis Lagrange (1736–1813).

A 2. Tételben szereplő

$$\sum_{i\in\mathbf{N}^n,|i|=s+1} \frac{\partial^i f(c)}{i!}\cdot (b-a)^i$$

összeget *Lagrange-maradéktagnak* nevezzük.

4. Megjegyzések.

i) Valamilyen $1 \le n, m \in \mathbf{N}$ mellett tekintsük az

$$f = (f_1, ..., f_m) \in \mathbf{R}^n \to \mathbf{R}^m$$

differenciálható függvényt. Tegyük fel, hogy az $a, b \in \mathbf{R}^n$, $a \neq b$ végpontokkal meghatározott szakaszra

$$[a, b] \subset \mathcal{D}_f.$$

Ekkor minden $i = 1, ..., m$ mellett egy alkalmas $\xi^{(i)} \in (a, b)$ helyen a

$$h := (h_1, ..., h_n) := b - a$$

jelöléssel

$$f_i(b) - f_i(a) = \langle \operatorname{grad} f_i(\xi^{(i)}), h\rangle = \sum_{j=1}^{n} \partial_j f_i(\xi^{(i)})\cdot h_j,$$

ezért

$$|f_i(b) - f_i(a)| \le \sum_{j=1}^{n} |\partial_j f_i(\xi^{(i)})|\cdot |h_j| \le \sum_{j=1}^{n} |\partial_j f_i(\xi^{(i)})|\cdot \|h\|_\infty \le$$

$$\le \sup\Big\{\sum_{j=1}^{n} |\partial_j f_i(x) : x \in (a, b)\Big\}\cdot \|h\|_\infty.$$

Következésképpen

$$f(b) - f(a) = \Big(\langle \operatorname{grad} f_1(\xi^{(1)}), h\rangle, \ldots, \langle \operatorname{grad} f_m(\xi^{(m)}), h\rangle\Big)$$

miatt

$$\|f(b) - f(a)\|_\infty = \max\{|f_i(b) - f_i(a)| : i = 1, ..., m\} =$$

$$\max\{|\langle \operatorname{grad} f_i(\xi^{(i)}), h\rangle| : i = 1, ..., m\} \le$$

$$\leq \max\Big\{\sup\Big\{\sum_{j=1}^{n}|\partial_j f_i(x) : x\in(a,b)\Big\} : i=1,...,m\Big\}\cdot \|h\|_\infty =$$

$$\sup\Big\{\max\Big\{\sum_{j=1}^{n}|\partial_j f_i(x) : i=1,...,m\Big\} : x\in(a,b)\Big\}\cdot \|h\|_\infty =$$

$$\sup\{\|f'(x)\|_{(\infty)} : x\in(a,b)\}\cdot \|h\|_\infty.$$

Ha tehát az $\mathbf{R}^n$-en is és az $\mathbf{R}^m$-en is a $\|.\|_\infty$ vektornormát vezetjük be, akkor az $f'(x)\quad(x\in\mathcal{D}_f)$ Jacobi-mátrix általuk generált $\|f'(x)\|_{(\infty)}$ (sor)normáját tekintve a

$$q := \sup\{\|f'(x)\|_{(\infty)} : x\in(a,b)\}$$

szimbólummal

$$\|f(b)-f(a)\|_\infty \leq q\cdot \|b-a\|_\infty.$$

(A Lagrange-középérték-tétel „vektorértékű" függvényekre vonatkozó alakja.) Megjegyezzük még, hogy a fentiek szerint az

$$A := \begin{bmatrix} \mathrm{grad} f_1(\xi^{(1)}) \\ \mathrm{grad} f_2(\xi^{(2)}) \\ \vdots \\ \mathrm{grad} f_m(\xi^{(m)}) \end{bmatrix} \in \mathbf{R}^{m\times n}$$

mátrixszal

$$f(b)-f(a) = A(b-a).$$

ii) Tekintsük az

$$f(x,y) := (xy, x^2y^2) \qquad \big((x,y)\in\mathbf{R}^2\big)$$

(könnyen igazolhatóan) differenciálható

$$f:\mathbf{R}^2\to\mathbf{R}^2$$

függvényt. Ekkor

$$f(1,1)-f(0,0) = (1,1),$$

továbbá

$$f'(x,y) = \begin{bmatrix} y & x \\ 2xy^2 & 2x^2y \end{bmatrix} \qquad \big((x,y)\in\mathbf{R}^2\big).$$

Mutassuk meg, hogy a $(0,0)$, $(1,1)$ pontokat össszekötő szakaszon nincs olyan (x,y) pont, amellyel

$$f(1,1)-f(0,0) = (1,1) = f'(x,y)\cdot(1,1) = (x+y, 2xy^2+2x^2y)$$

teljesülne. Valóban, ellenkező esetben

$$x + y = 1,\ 2xy^2 + 2x^2y = 1$$

állna fenn. Innen $2xy(x + y) = 2xy = 1$ és $y = 1 - x$ miatt

$$2x(1 - x) = 1 \iff 2x^2 - 2x + 1 = 0,$$

ami nem lehetséges, hiszen az utóbbi másodfokú egyenlet diszkriminánsa

$$4 - 4 \cdot 2 \cdot 1 < 0.$$

10. Fejezet

0. Emlékeztető.

I) Legyen $1 \leq n,\ k \in \mathbf{N},\ f \in \mathbf{R}^n \to \mathbf{R},\ a \in \operatorname{int} \mathcal{D}_f,\ j = 1, ..., n$ és

$$\partial_j^k f(a) := \partial_{j...j} f(a)$$

(ahol a „$j \ldots j$" rövidítés k darab j-t jelöl). Speciálisan $\partial_j^1 f(a) = \partial_j f(a)$ és

$$\partial_j^0 f(a) := f(a).$$

Ha $i := (i_1, ..., i_n) \in \mathbf{N}^n$, akkor legyen

$$\partial^i f(a) := \partial_1^{i_1} \ldots \partial_n^{i_n} f(a).$$

Részletesen kiírva tehát

$$\partial^i f(a) = \partial_{1...1...n...n} f(a),$$

ahol i_1 darab 1-es, ..., i_n darab n-es szerepel.

II) Az $i \in \mathbf{N}^n$ *multiindex* esetén az i *hossza*:

$$|i| := \|i\|_1 = \sum_{j=1}^n i_j,$$

az i *faktoriálisa*:

$$i! := \prod_{j=1}^n i_j! = \prod_{j=1}^n \prod_{k=1}^{i_j} k.$$

Az $x = (x_1, ..., x_n) \in \mathbf{R}^n$ vektor i-kitevős *hatványa*:

$$x^i := \prod_{j=1}^{n} x_j^{i_j}.$$

III) Az $a, b \in \mathbf{R}^n$ vektorokkal legyen:

$$[a, b] := \{a + t(b - a) \in \mathbf{R}^n : 0 \le t \le 1\},$$

ill.

$$(a, b) := \{a + t(b - a) \in \mathbf{R}^n : 0 < t < 1\}.^{49}$$

IV) Ha $f \in \mathbf{R}^n \to \mathbf{R}$ és $a \in \operatorname{int} \mathcal{D}_f$, $s \in \mathbf{N}$ mellett $f \in D^s\{a\}$, akkor a

$$T_{a,s} f(x) := \sum_{k=0}^{s} \sum_{i \in \mathbf{N}^n, |i| = k} \frac{\partial^i f(a)}{i!} \cdot (x - a)^i \qquad (x \in \mathbf{R}^n)$$

előírással definiált

$$T_{a,s} f : \mathbf{R}^n \to \mathbf{R}$$

függvény az f-nek az a-hoz tartozó s-edrendű *Taylor-polinomja.*

V) Taylor-formula (Lagrange-maradéktaggal): ha az $f \in \mathbf{R}^n \to \mathbf{R}$ függvényre valamilyen $s \in \mathbf{N}$ kitevővel $f \in D^{s+1}$, akkor bármely $[a, b] \subset \mathcal{D}_f$ esetén van olyan $c \in [a, b]$, hogy

$$f(b) = T_{a,s} f(b) + \sum_{i \in \mathbf{N}^n, |i| = s+1} \frac{\partial^i f(c)}{i!} \cdot (b - a)^i.$$

VI) Speciálisan (Lagrange): legyen adott a differenciálható $f \in \mathbf{R}^n \to \mathbf{R}$ függvény, és valamilyen $a, b \in \mathbf{R}^n$, $a \ne b$ végpontokkal $[a, b] \subset \mathcal{D}_f$. Ekkor egy alkalmas $c \in (a, b)$ mellett

$$f(b) - f(a) = \langle \operatorname{grad} f(c), b - a \rangle.^{50}$$

VII) Adott $1 \le n, m \in \mathbf{N}$ paraméterekkel tekintsük a

$$g = (g_1, ..., g_m) \in \mathbf{R}^n \to \mathbf{R}^m$$

[49] Az a és b végpontú zárt, ill. $a \ne b$ esetén nyílt *szakasz.*

[50] Ha tehát $\operatorname{grad} f \equiv 0$, akkor $f(a) = f(b)$, és egyúttal bármely $x, y \in [a, b]$ mellett (ugyanígy) $f(x) = f(y)$. Röviden: az f függvénynek az [a,b] szakaszra való leszűkítése konstansfüggvény.

differenciálható függvényt. Ha az $a, b \in \mathbf{R}^n$, $a \neq b$ végpontokkal meghatározott szakaszra $[a, b] \subset \mathcal{D}_g$, akkor

$$\|g(b) - g(a)\|_\infty \leq \sup\{\|g'(x)\|_{(\infty)} : x \in (a, b)\} \cdot \|b - a\|_\infty.$$

1. Taylor-formula.

Az alkalmazások szempontjából rendkívül hasznos az alábbi, *Peano*[51]-*maradéktagos Taylor-formula* is.

1. Tétel. *Tegyük fel, hogy az* $f \in \mathbf{R}^n \to \mathbf{R}$ *függvényre az* $a \in \operatorname{int} \mathcal{D}_f$ *helyen egy* $1 \leq s \in \mathbf{N}$ *„kitevővel"* $f \in D^s\{a\}$. *Ekkor van olyan*

$$\eta \in \mathbf{R}^n \to \mathbf{R},$$

függvény, amire

$$\eta(x) \to 0 \qquad (\|x\| \to 0)^{52}$$

és

$$f(a+h) - f(a) =$$

$$(*) \quad \sum_{k=1}^{s} \sum_{i \in \mathbf{N}^n, |i| = k} \frac{\partial^i f(a)}{i!} \cdot h^i + \eta(h) \cdot \|h\|^s \qquad (h \in \mathbf{R}^n,\ a + h \in \mathcal{D}_f).$$

Megjegyezzük, hogy a $(*)$ egyenlőség másképp írva a következőképpen szól:

$$\frac{f(x) - T_{a,s} f(x)}{\|x - a\|^s} \to 0 \qquad (\|x - a\| \to 0).$$

Az $s = 1$ esetben

$$T_{a,1} f(x) = f(a) + \sum_{i \in \mathbf{N}, |i| = 1} \frac{\partial^i f(a)}{i!} \cdot h^i =$$

$$f(a) + \sum_{j=1}^{n} \partial_j f(a) \cdot h_j = f(a) + \langle \operatorname{grad} f(a), h \rangle$$

[51] Giuseppe Peano (1858 – 1932).

[52] Ha kell, nyugodtan feltehetjük azt is, hogy $\eta(0) = 0$, azaz, hogy az η folytonos a 0-ban.

miatt[53] $(*)$ pontosan az $f \in D\{a\}$ feltételt jelenti.

Bizonyítás. Tegyük fel: valamilyen $1 \leq s \in \mathbf{N}$ mellett igaz a tétel. Ha $f \in D^{s+1}\{a\}$, akkor egy $K(a) \subset \mathcal{D}_f$ környezettel $f \in D^s\{x\} \quad (x \in K(a))$, és $\partial_j f \in D^s\{a\} \quad (j = 1, ..., n)$. Legyen

$$g(x) := f(x) - f(a) - \sum_{k=1}^{s+1} \sum_{i \in \mathbf{N}^n, |i|=k} \frac{\partial^i f(a)}{i!} \cdot (x-a)^i \qquad \big(x \in K(a)\big).$$

Ekkor $g \in D^{s+1}\{a\}$, és

$$(**) \qquad \partial^j g(a) = 0 \qquad (j \in \mathbf{N}^n,\ |j| = 1, ..., s+1).$$

A $K(a)$ környezetről az is feltehető, hogy $g \in D^s\{x\} \quad (x \in K(a),\ s \geq 1)$, így a $h := x - a \quad (x \in K(a))$ jelöléssel (ld. Lagrange-tétel)

$$g(x) - g(a) = g(x) = f(a+h) - f(a) - \sum_{k=1}^{s+1} \sum_{i \in \mathbf{N}^n, |i|=k} \frac{\partial^i f(a)}{i!} \cdot h^i =$$

$$\langle \operatorname{grad} g(a+th), h \rangle = \sum_{l=1}^{n} \partial_l g(a+th) \cdot h_l.^{54}$$

Az $l = 1, ..., n$ indexekre $\partial_l g \in D^s\{a\}$, ezért (ld. indukciós feltétel)

$$\partial_l g(a+th) = \partial_l g(a+th) - \partial_l g(a) =$$

$$\sum_{k=1}^{s} \sum_{i \in \mathbf{N}^n, |i|=k} \frac{\partial^i \partial_l g(a)}{i!} \cdot (th)^i + \eta_l(th) \cdot \|th\|^s.$$

Itt az $\eta_l \in \mathbf{R}^n \to \mathbf{R} \quad (l = 1, ..., n)$ függvényekre $\eta_l(z) \to 0 \quad (\|z\| \to 0)$, valamint (ld. $(**)$) $\partial^i \partial_l g(a) = 0$. Ez azt jelenti, hogy

$$f(a+h) - f(a) - \sum_{k=1}^{s+1} \sum_{i \in \mathbf{N}^n, |i|=k} \frac{\partial^i f(a)}{i!} \cdot h^i = \sum_{l=1}^{n} \eta_l(th) \cdot \|th\|^s \cdot h_l.$$

Ha $\eta(0) := 0$ és

$$\eta(h) := \frac{\sum_{l=1}^{n} \eta_l(th) \cdot \|th\|^s \cdot h_l}{\|h\|^{s+1}} \qquad \big(0 \neq h \in \mathbf{R}^n,\ a+h \in K(a)\big),$$

[53] $i \in \mathbf{N}^n, |i| = 1 \iff$ egy $j = 1, ..., n$ indexszel $i_j = 1$ és $i_\nu = 0$ $(j \neq \nu = 1, ..., n)$.

[54] Alkalmas $t \in (0,1)$ (a h-tól függő) számmal.

akkor

$$|\eta(h)| \leq \frac{\sum_{l=1}^{n} |\eta_l(th)|\cdot \|h\|^s \cdot \|h\|}{\|h\|^{s+1}} = \sum_{l=1}^{n} |\eta_l(th)| \to 0 \qquad (\|h\| \to 0),$$

így $\eta(h) \to 0 \quad (\|h\| \to 0)$. A $h \in \mathbf{R}^n \ \ (a+h \in K(a))$ vektorokra fennálló

$$f(a+h) - f(a) = \sum_{k=1}^{s+1} \sum_{i \in \mathbf{N}^n, |i|=k} \frac{\partial^i f(a)}{i!} \cdot h^i + \eta(h) \cdot \|h\|^{s+1}$$

egyenlőség szerint az állításunk az s helyett $(s+1)$-re is igaz. ■

Különösen fontos az előző tétel $s = 2$ esete, amikor is $f \in D^2\{a\}$. Ekkor a

$$\sum_{i \in \mathbf{N}^n, |i|=2} \frac{\partial^i f(a)}{i!} \cdot h^i$$

összegben megjelenő $i \in \mathbf{N}^n$, $|i| = 2$ multiindexek mindegyike csak a következők valamelyike lehet:

$$i = (0, \dots, 0, 2, 0, \dots, 0), \ \text{vagy} \ i = (0, \dots, 0, 1, 0, \dots, 0, 1, 0, \dots 0),$$

ahol tehát $r = 1, ..., n$ és

$$i_r = 2, \ i_k = 0 \qquad (r \neq k = 1, ..., n),$$

és ekkor

$$i! = 2, \ h^i = h_r^2,$$

vagy $1 \leq j < l \leq n$ és

$$i_j = i_l = 1, \ i_k = 0 \qquad (j, \, l \neq k = 1, ..., n),$$

így

$$i! = 1, \ h^i = h_j h_l.$$

Ezért (az utóbbi esetben figyelembe véve a $\partial_{jl} f(a) = \partial_{lj} f(a)$ egyenlőséget is) azt kapjuk, hogy az

$$f''(a) := \begin{bmatrix} \partial_{11} f(a) & \partial_{12} f(a) & \dots & \partial_{1n} f(a) \\ \partial_{21} f(a) & \partial_{22} f(a) & \dots & \partial_{2n} f(a) \\ \vdots & \vdots & \dots & \vdots \\ \partial_{n1} f(a) & \partial_{n2} f(a) & \dots & \partial_{nn} f(a) \end{bmatrix} \in \mathbf{R}^{n \times n}$$

második deriváltmátrixszal [55] a $h \in \mathbf{R}^n$ vektorokra

$$\sum_{i \in \mathbf{N}^n, |i|=2} \frac{\partial^i f(a)}{i!} \cdot h^i = \frac{1}{2} \cdot \sum_{j=1}^{n} \sum_{l=1}^{n} \partial_{jl} f(a) \cdot h_j h_l = \frac{1}{2} \cdot \langle f''(a)h, h \rangle.$$

Következésképpen

$$f(a+h) =$$

$$f(a) + \langle \operatorname{grad} f(a), h \rangle + \frac{1}{2} \cdot \langle f''(a)h, h \rangle + \eta(h) \cdot \|h\|^2 \quad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f),$$

ahol

$$\eta(x) \to 0 \qquad (\|x\| \to 0).$$

Az alábbi jelölést fogjuk használni:

$$Q_a^f(x) := \langle f''(a)x, x \rangle \qquad (x \in \mathbf{R}^n).$$

A „szokásos" $x = (x_1, ..., x_n) \in \mathbf{R}^n$ koordinátázást alkalmazva tehát

$$Q_a^f(x) = \sum_{i=1}^{n} \sum_{k=1}^{n} \partial_{ik} f(a) x_i x_k \qquad (x \in \mathbf{R}^n).$$

Így végül azt mondhatjuk, hogy

$$f(a+h) - f(a) =$$

$$\langle \operatorname{grad} f(a), h \rangle + \frac{1}{2} \cdot Q_a^f(h) + \eta(h) \cdot \|h\|^2 \qquad (h \in \mathbf{R}^n,\ a+h \in \mathcal{D}_f).$$

2. Kvadratikus alak.

Legyen az $A = (a_{ik}) \in \mathbf{R}^{n \times n}$ egy szimmetrikus mátrix. Ekkor az

$$\mathbf{R}^n \ni x \mapsto Q(x) := \langle Ax, x \rangle \in \mathbf{R}$$

előírással definiált Q leképezést (az A mátrix által meghatározott) *kvadratikus alaknak* nevezzük. Világos, hogy $Q(0) = 0$ és

$$Q(x) = \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} x_i x_k \qquad \Big(x = (x_1, ..., x_n) \in \mathbf{R}^n\Big).$$

Speciálisan, az $n = 2$ esetben $a, b, c \in \mathbf{R}$ és

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

[55] Ami (ld. Young-tétel) szimmetrikus.

valamint

$$Q(u,v) = \langle A(u,v),(u,v)\rangle =$$

$$\langle (au+bv, bu+cv),(u,v)\rangle = au^2 + 2buv + cv^2 \qquad \Big((u,v)\in \mathbf{R}^2\Big).$$

Legyen pl.

$$Q(u,v) := 2u^2 + 2uv - v^2 \qquad \Big((u,v)\in \mathbf{R}^2\Big),$$

ekkor

$$A = \begin{bmatrix} 2 & 1 \\ 1 & -1 \end{bmatrix}.$$

Ha

$$f \in \mathbf{R}^n \to \mathbf{R}$$

és $f \in D^2\{a\}$, akkor (ld. fentebb) Q_a^f egy kvadratikus alak.

Mutassuk meg, hogy minden

$$Q : \mathbf{R}^n \to \mathbf{R}$$

kvadratikus alak folytonos. Valóban, legyen $y \in \mathbf{R}^n$, ekkor tetszőleges $x \in K_1(y)$ vektorra a $\|.\| := \|.\|_2$ vektornorma által indukált $q := \|A\|_{(2)}$ mátrixnormával (ahol $A \in \mathbf{R}^{n\times n}$ a Q-t a fentiek szerint meghatározó szimmetrikus mátrix) a Cauchy–Bunyakovszkij-egyenlőtlenség szerint

$$|Q(x)-Q(y)| = |\langle Ax,x\rangle - \langle Ay,y\rangle| = |\langle A(x-y),x\rangle + \langle Ay, x-y\rangle| \le$$

$$|\langle A(x-y),x\rangle| + |\langle Ay,x-y\rangle| \le$$

$$q\cdot (\|x-y\|\cdot (\|x\|+\|y\|) \le q(1+2\|y\|)\cdot \|x-y\|,$$

amiből a $Q \in \mathcal{C}\{y\}$ folytonosság már nyilván következik.

Azt mondjuk, hogy az előbbi Q kvadratikus alak

- *pozitív definit*, ha $Q(x) > 0 \quad (x \in \mathbf{R}^n \setminus \{0\})$;
- *negatív definit*, ha $Q(x) < 0 \quad (x \in \mathbf{R}^n \setminus \{0\})$;
- *pozitív szemidefinit*, ha $Q(x) \ge 0 \quad (x \in \mathbf{R}^n)$;
- *negatív szemidefinit*, ha $Q(x) \le 0 \quad (x \in \mathbf{R}^n)$;
- *indefinit*, ha van olyan $x \in \mathbf{R}^n$ és $y \in \mathbf{R}^n$, hogy $Q(y) < 0 < Q(x)$.

Minden definit alak tehát egyúttal szemidefinit is. A szemidefinit Q pontosan akkor nem definit, ha van olyan $x \in \mathbf{R}^n \setminus \{0\}$, amelyre $Q(x) = 0$. Tekintsük pl. a

$$Q(u,v) := 2u^2 + 2uv - v^2 \qquad \Big((u,v) \in \mathbf{R}^2\Big)$$

kvadratikus alakot. A Q indefinit, ui.

$$Q(0,1) = -1 < 0 < Q(1,0) = 2.$$

Speciálisan, ha

$$f \in \mathbf{R}^n \to \mathbf{R}$$

és $f \in D^2\{a\}$, akkor a Q_a^f nem más, mint az $f''(a)$ mátrix által meghatározott kvadratikus alak. Tehát ennek a definitsége, ill. a szemidefinitsége a

$$\partial_{ij} f(a) \qquad (i,j = 1,...,n)$$

másodrendű parciális deriváltak segítségével dönthető el. Pl. az $n = 1$ esetben nyilván

$$Q_a^f(x) = f''(a) \cdot x^2 \qquad (x \in \mathbf{R}),$$

azaz a Q_a^f definitségét az $f''(a)$ előjele határozza meg:

$$Q_a^f \text{ pozitív definit} \iff f''(a) > 0;$$
$$Q_a^f \text{ negatív definit} \iff f''(a) < 0.$$

Ha $f''(a) = 0$, akkor $Q_a^f(x) = 0 \quad (x \in \mathbf{R})$, azaz ekkor a Q_a^f egyszerre pozitív és negatív szemidefinit.

Mivel a Q kvadratikus alakra (továbbra is a $\|.\| := \|.\|_2$ normát használva)

$$Q(x) = \|x\|^2 \cdot Q(x/\|x\|) \qquad (x \in \mathbf{R}^n \setminus \{0\})$$

és itt $\|x\|^2 > 0$, valamint $\Big\|x/\|x\|\Big\| = 1$, ezért a $Q(x)$ helyettesítési értékek előjelének a vizsgálatakor elegendő csak az $\|x\| = 1$ feltételnek eleget tevő x-ekre szorítkozni. A

$$G := \{x \in \mathbf{R}^n : \|x\| = 1\}$$

halmaz korlátos és zárt, azaz kompakt, ezért a Q folytonossága alapján a

$$\{Q(x) : x \in G\}$$

halmaznak van minimuma is meg maximuma is. Legyenek ezek rendre m, M, ekkor az előbbiek szerint

$$m \cdot \|x\|^2 \leq Q(x) \leq M \cdot \|x\|^2 \qquad (x \in \mathbf{R}^n).$$

Ez azt jelenti, hogy

a) Q pozitív definit $\iff m > 0$;

b) Q negatív definit $\iff M < 0$;

c) Q pozitív szemidefinit $\iff m \geq 0$;

d) Q negatív szemidefinit $\iff M \leq 0$;

e) Q indefinit $\iff m \cdot M < 0$.

Jelöljük az A mátrix esetén d_i-vel $(i = 1, ..., n)$ az alábbi („bal felső sarok") determinánst:

$$d_i := \det{(a_{jk})_{j,k=1}^{i}}.$$

Ekkor az algebrából ismert *Sylvester*[56]*-féle kritérium*:

Q pozitív definit $\iff d_i > 0 \quad (i = 1, ..., n)$;

Q negatív definit $\iff (-1)^j \cdot d_j > 0 \quad (j = 1, ..., n)$.

Pl., ha $n = 2$, akkor alkalmas $a, b, c \in \mathbf{R}$ számokkal

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix},$$

valamint

$$Q(u, v) = au^2 + 2buv + cv^2 \qquad \big((u, v) \in \mathbf{R}^2\big).$$

Ekkor

$$d_1 = a, \ d_2 = ac - b^2,$$

és így

$$Q \text{ pozitív definit } \iff a > 0 \text{ és } ac > b^2;$$

$$Q \text{ negatív definit } \iff a < 0 \text{ és } ac > b^2.$$

Sőt, ekkor az is igaz, hogy

$$Q \text{ indefinit } \iff ac < b^2.$$

Mindez egyébként elemi úton is könnyen belátható.

[56]James Joseph Sylvester (1814 – 1897).

3. Lokális szélsőérték.

A matematikai alkalmazások egyik legfontosabb fejezete a függvények szélsőértékeinek a vizsgálata. A címben szereplő lokális szélsőértékek értelmezése az egyváltozós esettel analóg módon történhet. Nevezetesen, legyen valamilyen $1 \leq n \in \mathbf{N}$ mellett adott az

$$f \in \mathbf{R}^n \to \mathbf{R}$$

függvény. Azt mondjuk, hogy az f függvénynek az $a \in \mathcal{D}_f$ helyen *lokális maximuma* van (más szóval az a pont *lokális maximumhely*), ha egy alkalmas $K(a)$ környezettel fennáll a következő becslés:

$$f(x) \leq f(a) \qquad (x \in K(a) \cap \mathcal{D}_f).$$

Ekkor az $f(a)$ függvényértéket az f *lokális maximumának* nevezzük.

Ha a fenti $a \in \mathcal{D}_f$ helyen az előbbinél többet kívánó

$$f(x) \leq f(a) \qquad (x \in \mathcal{D}_f)$$

egyenlőtlenség igaz, akkor az f függvénynek az a-ban *abszolút maximuma* van (vagy másképp fogalmazva az a pont *abszolút maximumhelye* az f-nek). Ekkor maga az $f(a)$ függvényérték az f *abszolút maximuma.* Világos, hogy az utóbbi esetben az $f(a)$ egyúttal az f lokális maximuma is, ill. ekkor az a pont lokális maximumhely is.

Analóg módon kapjuk a minimumhelyek, ill. minimumok fogalmát. Ha ui. $a \in \mathcal{D}_f$ és egy $K(a)$ környezet esetén igaz az

$$f(x) \geq f(a) \qquad (x \in K(a) \cap \mathcal{D}_f)$$

becslés, akkor az $f(a)$ függvényértéket az f *lokális minimumának,* az a-t pedig az f *lokális minimumhelyének* nevezzük. (Más szóval ekkor az f-nek az a-ban *lokális minimuma van.*)

Hasonlóan, ha

$$f(x) \geq f(a) \qquad (x \in \mathcal{D}_f),$$

akkor az f függvénynek az a-ban *abszolút minimuma* van (az a pont *abszolút minimumhelye* az f-nek), ill. az $f(a)$ függvényérték az f *abszolút minimuma.* Ilyenkor az $f(a)$ nyilván az f lokális minimuma is.

Esetenként azt mondjuk röviden, hogy az f függvénynek az a-ban *lokális szélsőértéke* van (az a pont *lokális szélsőértékhelye* az f-nek), ha az f függvénynek az a-ban lokális maximuma, vagy lokális minimuma van. Továbbá az a pont *abszolút szélsőértékhelye* a szóban forgó f-nek (az $f(a)$ függvényérték *abszolút szélsőértéke* az f-nek), ha az a-ban abszolút maximuma, vagy abszolút mimimuma van az f függvénynek.

A többváltozós differenciálható függvényekre vonatkozóan is az egyik legismertebb állítás az alábbi tétel, az ún. *elsőrendű szükséges feltétel lokális szélsőértékre*:

2. Tétel. *Tegyük fel, hogy az $f \in \mathbf{R}^n \to \mathbf{R}$ függvénynek valamilyen $a \in \operatorname{int} \mathcal{D}_f$ helyen lokális szélsőértéke van és $f \in D\{a\}$. Ekkor $f'(a) = 0$.*

Megjegyezzük, hogy

$$f'(a) = \operatorname{grad} f(a) = \Big(\partial_1 f(a), \ldots, \partial_n f(a)\Big) \in \mathbf{R}^n,$$

tehát az $f'(a) = 0$ feltétel azt jelenti, hogy

$$\partial_i f(a_1, ..., a_n) = 0 \qquad (i = 1, ..., n).$$

Ez nem más, mint az a vektor koordinátáira vonatkozó egyenletrendszer.

Bizonyítás. Legyen pl. az a hely lokális maximumhely (minimumhely esetén analóg a bizonyítás). Van tehát olyan $r > 0$, hogy (az $a \in \operatorname{int} \mathcal{D}_f$ feltételt is figyelembe véve) a

$$K(a) := K_r(a) \subset \mathcal{D}_f$$

környezettel

$$f(x) \leq f(a) \qquad \Big(x \in K(a)\Big).$$

Világos, hogy (az $\mathbf{R}^n$-ben a $\|.\| := \|.\|_2$ normát vezetve be) tetszőleges

$$e \in \mathbf{R}^n, \ \|e\| = 1$$

vektorral egyúttal az

$$f_e(t) := f(a + te) \qquad (-r < t < r)$$

függvénynek is lokális maximuma van a 0-ban.[57] Az

$$f_e \in \mathbf{R} \to \mathbf{R}$$

[57]Ekkor ui. $a + te \in K(a)$, és így $f_e(t) = f(a + te) \leq f(a) = f_e(0) \quad (-r < t < r)$.

függvény differenciálható a 0-ban és (az egyváltozós valós függvényekre vonatkozó analóg elsőrendű szükséges feltétel alapján)

$$f_e'(0) = \partial_e f(a) = 0.$$

Más szóval az f függvénynek az a-ban minden $e \in \mathbf{R}^n$, $\|e\| = 1$ irányvektor esetén az a-beli iránymenti deriváltja nulla. Így speciálisan

$$\partial_j f(a) = 0 \qquad (j = 1, ..., n),$$

ami az állításunkat jelenti. ■

Emlékeztetünk arra, hogy már az $n = 1$ esetben láttuk: a 2. Tételbeli $f'(a) = 0$ feltétel csak szükséges, de nem elégséges ahhoz, hogy az f-nek az a-ban lokális szélsőértéke legyen.

A következő állításban elégséges feltételt adunk a lokális szélsőértékekkel kapcsolatban (*másodrendű elégséges feltétel lokális szélsőérték létezésére*):

> **3. Tétel.** *Tegyük fel, hogy az $f \in \mathbf{R}^n \to \mathbf{R}$ függvényre egy adott $a \in \operatorname{int} \mathcal{D}_f$ helyen az alábbi feltételek teljesülnek*:
>
> a) $f \in D^2\{a\}$;
>
> b) $\operatorname{grad} f(a) = 0$;
>
> c) Q_a^f *pozitív (negatív) definit.*
>
> *Ekkor az f-nek az a-ban lokális minimuma (maximuma) van.*

Bizonyítás. Mivel $a \in \operatorname{int} \mathcal{D}_f$, ezért egy alkalmas $r > 0$ mellett a $K_r(a)$ környezetre $K_r(a) \subset \mathcal{D}_f$ teljesül. Továbbá a b) feltételt is figyelembe véve, pl. a $\|.\| := \|.\|_2$ választással (ld. 1. Tétel)

$$f(a+h) - f(a) = \langle \operatorname{grad} f(a), h \rangle + \frac{1}{2} \cdot \langle f''(a)h, h \rangle + \eta(h) \cdot \|h\|^2 =$$

$$\frac{1}{2} \cdot \langle f''(a)h, h \rangle + \eta(h) \cdot \|h\|^2 = \frac{1}{2} \cdot Q_a^f(h) + \eta(h) \cdot \|h\|^2 \qquad \Big(h \in K_r(0)\Big),$$

ahol

$$(*) \qquad \eta(h) \to \eta(0) = 0 \qquad (\|h\| \to 0).$$

Feltéve, hogy a Q_a^f kvadratikus alak pozitív definit (a negatív definit esetben analóg a bizonyítás), egy alkalmas $m > 0$ számmal

$$Q_a^f(\xi) \geq m \cdot \|\xi\|^2 \qquad (\xi \in \mathbf{R}^n).$$

Ezért

$$f(a+h)-f(a)\ge\Big(\frac{m}{2}+\eta(h)\Big)\cdot\|h\|^2\qquad\big(h\in K_r(0)\big).$$

A $(*)$ feltétel miatt a fenti r „sugárról" az is feltehető, hogy

$$|\eta(h)|<\frac{m}{4}\qquad\big(h\in K_r(0)\big),$$

következésképpen

$$f(a+h)-f(a)\ge\Big(\frac{m}{2}-|\eta(h)|\Big)\cdot\|h\|^2\ge$$

$$\Big(\frac{m}{2}-\frac{m}{4}\Big)\cdot\|h\|^2=\frac{m}{4}\cdot\|h\|^2\ge 0\qquad\big(h\in K_r(0)\big).$$

Tehát

$$f(a+h)\ge f(a)\qquad\big(h\in K_r(0)\big),$$

vagy más szóval

$$f(x)\ge f(a)\qquad\big(x\in K_r(a)\big),$$

azaz az f-nek az a-ban lokális minimuma van. ■

Most is elmondhatjuk, hogy az $n=1$ esetben láttuk: a 3. Tételben szereplő feltételek csak elégségesek, de nem szükségesek ahhoz, hogy az f-nek az a-ban lokális szélsőértéke legyen.

A továbbiakban kétszer differenciálható függvények körében adunk szükséges feltételt arra, hogy a szóban forgó függvénynek valamilyen pontban lokális szélsőértéke legyen (*másodrendű szükséges feltétel lokális szélsőérték létezésére*):

4. Tétel. *Tekintsük az $f\in\mathbf{R}^n\to\mathbf{R}$ függvényt, és tételezzük fel, hogy az $a\in\operatorname{int}\mathcal{D}_f$ pontban igazak az alábbiak feltételek:*

a) $f\in D^2\{a\}$;

b) *az f-nek az a-ban lokális minimuma (maximuma) van.*

Ekkor $\operatorname{grad} f(a)=0$, *és a Q_a^f kvadratikus alak pozitív (negatív) szemidefinit.*

Bizonyítás. A $\operatorname{grad} f(a)=0$ egyenlőség nyilván következik a 2. Tételből. Továbbá az $a\in\operatorname{int}\mathcal{D}_f$ feltevés és a b) feltétel miatt egy alkalmas

$r > 0$ mellett a $K_r(a)$ környezettel $K_r(a) \subset \mathcal{D}_f$, és az a-beli lokális minimum létezése mellett (lokális maximum esetén a további gondolatmenet értelemszerűen módosítható)

$$(*) \qquad f(x) \geq f(a) \qquad \Big(x \in K_r(a)\Big).$$

Indirekt módon tegyük fel, hogy a Q_a^f kvadratikus alak nem pozitív szemidefinit, azaz van olyan $h \in \mathbf{R}^n$, hogy

$$Q_a^f(h) < 0.$$

Mivel $Q_a^f(0) = 0$, ezért $h \neq 0$. Világos, hogy minden $0 < t < r/\|h\|$ szám esetén

$$a + th \in K_r(a),$$

ezért a fentebb már idézett formula szerint

$$f(a+th) - f(a) = \frac{1}{2} \cdot Q_a^f(th) + \eta(th) \cdot \|th\|^2 =$$

$$\frac{t^2}{2} \cdot \Big(Q_a^f(h) + 2\eta(th) \cdot \|h\|^2\Big) \qquad (0 < t < r/\|h\|).$$

Itt

$$\eta(\xi) \to \eta(0) = 0 \qquad (\|\xi\| \to 0),$$

és így egyúttal

$$\eta(th) \to 0 \qquad (t \to 0).$$

Létezik tehát olyan $0 < \tau < r/\|h\|$, amellyel

$$2|\eta(\tau h)| \cdot \|h\|^2 < -\frac{Q_a^f(h)}{2}.$$

Innen azt kapjuk, hogy

$$f(a+\tau h) - f(a) < \frac{\tau^2}{2} \cdot \Big(Q_a^f(h) + 2|\eta(\tau h)| \cdot \|h\|^2\Big) <$$

$$\frac{\tau^2}{2} \cdot \left(Q_a^f(h) - \frac{Q_a^f(h)}{2}\right) = \frac{\tau^2}{2} \cdot \frac{Q_a^f(h)}{2} < 0,$$

azaz

$$f(a+\tau h) < f(a).$$

Ez $a + \tau h \in K_r(a)$ miatt nyilván ellentmond a $(*)$ becslésnek. ■

Amint azt már az előző tételekkel kapcsolatban is említettük, a most bebizonyított másodrendű szükséges feltétel sem elégséges ahhoz, hogy az illető függvénynek az adott helyen lokális szélsőértéke legyen.

3. Megjegyzések.

i) Ha $n = 1$, akkor – ahogyan azt korábban is megjegyeztük – a Q_a^f kvadratikus alak definitsége az $f''(a)$ előjelétől függ. Így pl. a másodrendű elégséges feltétel (az egyváltozós valós függvényekkel kapcsolatban tanultak alapján „ismerősen") a következőképpen hangzik: $f'(a) = 0$ és $f''(a) \neq 0$. Az $f''(a) > 0$ esetben lokális minimuma, $f''(a) < 0$ esetén pedig lokális maximuma van az f függvénynek az a helyen.

ii) A másodrendű szükséges feltétel más „olvasatban" azt jelenti, hogy ha a Q_a^f indefinit, akkor az f-nek az a helyen nincs lokális szélsőértéke.

iii) Egy $f \in \mathbf{R}^n \to \mathbf{R}$ differenciálható függvény esetén az

$$\mathcal{S} := \{x \in \mathcal{D}_f : f'(x) = 0\}$$

halmaz elemeit az f *stacionárius pontjainak* nevezzük. Ha $\mathcal{L}$ jelöli az f lokális szélsőértékhelyeinek a halmazát, akkor az elsőrendű szükséges feltétel szerint[58]

$$\mathcal{L} \subset \mathcal{S}.$$

Ha $f \in D^2$ és $a \in \mathcal{S}$, akkor három eset lehetséges:

1^o az a-ban teljesül a másodrendű elégséges feltétel. Ekkor $a \in \mathcal{L}$;

2^o az a-ban nem teljesül a másodrendű szükséges feltétel. Ekkor $a \notin \mathcal{L}$;

3^o az a-ban nem teljesül a másodrendű elégséges feltétel, de teljesül a másodrendű szükséges feltétel. Ekkor akár $a \in \mathcal{L}$, akár pedig $a \notin \mathcal{L}$ is előfordulhat. Tehát az idézett tételek segítségével ebben az esetben nem lehet eldönteni azt, hogy az a lokális szélsőértékhelye-e az f függvénynek, vagy sem.

11. Fejezet

1. Paraméteres integrál.

Valamilyen kompakt $[a, b]$ intervallum $(a, b \in \mathbf{R},\ a < b)$ és $\emptyset \neq U \subset \mathbf{R}^n$ $(1 \leq n \in \mathbf{N})$ nyílt halmaz esetén tekintsük az

$$f : U \times [a, b] \to \mathbf{R}$$

[58]Mivel az f differenciálható, így minden $a \in \mathcal{D}_f$ pontra egyúttal $a \in \operatorname{int} \mathcal{D}_f$ is igaz és alkalmazható a 2. Tétel az a-ra nézve.

függvényt.[59] Ha $x \in U$, akkor legyen $f_x : [a,b] \to \mathbf{R}$ az a függvény, amire

$$f_x(t) := f(x,t) \qquad (t \in [a,b]).$$

Tegyük fel, hogy minden $x \in U$ esetén az f_x függvény Riemann-integrálható: $f_x \in R[a,b]$, legyen ekkor

$$F(x) := \int_a^b f(x,t)\,dt := \int_a^b f_x \qquad (x \in U).$$

Az így definiált $F : U \to \mathbf{R}$ függvényt az f által meghatározott *paraméteres integrálnak* nevezzük.

Világos, hogy

$$\mathcal{D}_f = U \times [a,b] \subset \mathbf{R}^{n+1}.$$

Vezessük be $\mathbf{R}^n$-en is és $\mathbf{R}^{n+1}$-en is a $\|.\| := \|.\|_\infty$ normát. Nem fog félreértést okozni, ha $\xi \in \mathbf{R}^n$ és $\eta \in \mathbf{R}^{n+1}$ esetén egyaránt a $\|\xi\|$, $\|\eta\|$ jelölést használjuk. Így pl. nyilvánvaló, hogy az

$$(x,t) \in U \times [a,b]$$

vektorra

$$\|(x,t)\| = \max\{\|x\|, |t|\}.$$

Továbbá, ha az

$$f : U \times [a,b] \to \mathbf{R}$$

függvény folytonos, akkor tetszőleges $x \in U$ mellett az $f_x : [a,b] \to \mathbf{R}$ függvény is folytonos. Ui., ha $s \in [a,b]$ és $\xi := (x,s)$, akkor $f \in \mathcal{C}\{\xi\}$ miatt bármilyen $\varepsilon > 0$ esetén van olyan $\delta > 0$, amellyel

$$|f(\omega) - f(\xi)| < \varepsilon \qquad (\omega \in U \times [a,b],\ \|\omega - \xi\| < \delta),$$

azaz[60]

$$|f(y,t) - f(x,s)| < \varepsilon \qquad \Big((y,t) \in U \times [a,b],\ \|(y,t) - (x,s)\| < \delta\Big).$$

Ha itt $y := x$, akkor

$$\|(x,t) - (x,s)\| = \|(0, t-s)\| = |t-s|,$$

[59] Valójában egy $f \in \mathbf{R}^{n+1} \to \mathbf{R}$ függvényről van szó, amikor is az $\mathbf{R}^{n+1} \equiv \mathbf{R}^n \times \mathbf{R}$ azonosítás révén az f argumentumában csoportosítottuk a változókat: az első n változó által alkotott vektor egy $\mathbf{R}^n$-beli nyílt halmazból, az utolsó változó pedig egy $\mathbf{R}$-beli kompakt intervallumból van véve.

[60] Az $\omega = (y,t)$ „koordinátázással".

ezért

$$|f(x,t)-f(x,s)| = |f_x(t)-f_x(s)| < \varepsilon \qquad (t \in [a,b],\ |t-s| < \delta),$$

más szóval $f_x \in \mathcal{C}\{s\}$. Következésképpen $f_x \in C[a,b]$, így $f_x \in R[a,b]$, tehát létezik a fenti F paraméteres integrál.

A továbbiakban az F paraméteres integrált vizsgáljuk folytonosság és differenciálhatóság szempontjából. Látni fogjuk, hogy a folytonosság mintegy automatikusan teljesül. A differenciálhatóságot illetően kiderül, hogy bizonyos „természetes" feltételek mellett a $\partial_1 F, ..., \partial_n F$ parciális deriváltak kiszámításakor a szóban forgó $\partial_1, ..., \partial_n$ deriválást elvégezhetjük az „integráljel mögött", azaz a $\partial_1, ..., \partial_n$ parciális deriválás és az $\int_a^b ...$ integrálás felcserélhető.[61]

1. Tétel. *Tegyük fel, hogy adott az* $[a,b]$ $(a,b \in \mathbf{R},\ a<b)$ *kompakt intervallum,* $1 \le n \in \mathbf{N}$, *és* $\emptyset \ne U \subset \mathbf{R}^n$ *nyílt halmaz. Ekkor tetszőleges folytonos*

$$f : U \times [a,b] \to \mathbf{R}$$

függvény esetén az

$$F(x) := \int_a^b f(x,t)\,dt \qquad (x \in U)$$

paraméteres integrálra az alábbiak igazak:

1^o *az* F *függvény folytonos;*

2^o *ha valamilyen* $i = 1, ..., n$ *indexre létezik és folytonos a* $\partial_i f$ *parciális deriváltfüggvény, akkor létezik a* $\partial_i F$ *parciális deriváltfüggvény is, és*

$$\partial_i F(x) = \int_a^b \partial_i f(x,t)\,dt \qquad (x \in U);$$ [62]

3^o *amennyiben az* f *folytonosan differenciálható, azaz* $f \in C^1$, *akkor* $F \in C^1$.

[61] Általában is (mind az elmélet, mind a gyakorlat számára) rendkívül fontos kérdés bizonyos műveletek (operátorok) egymás utáni alkalmazásával kapcsolatban azok sorrendjének a felcserélhetősége.

[62] Ha pl. n=1, akkor formálisan: $\dfrac{d}{dx}\displaystyle\int_a^b f(x,t)\,dt = \int_a^b \frac{\partial f}{\partial x}(x,t)\,dt.$

Bizonyítás. Az F függvény valamely $x \in U$ pontbeli folytonosságához az

$$F(x) - F(y) = \int_a^b (f(x,t) - f(y,t))\, dt \qquad (y \in U)$$

különbséget, azaz az

$$f(x,t) - f(y,t) \qquad (t \in [a,b])$$

megváltozást kell „kezelni". Legyen ehhez tehát adott az $x \in U$ vektor, ekkor az U halmaz nyíltsága miatt egy alkalmas $r > 0$ számmal

$$G_r := \{y \in U : \|x - y\| \leq r\} \subset U.$$

A G_r halmaz könnyen láthatóan zárt, ezért az

$$A := G_r \times [a,b]\, (\subset U \times [a,b] = \mathcal{D}_f)$$

halmaz is zárt. Mivel az A nyilván korlátos is, így kompakt. A Heine[63]-tétel alapján az $f_{|A}$ leszűkítés egyenletesen folytonos, tehát tetszőleges $\varepsilon > 0$ számhoz van olyan $\delta > 0$, hogy

$$|f(\xi) - f(\zeta)| < \varepsilon \qquad (\xi,\, \zeta \in A,\, \|\xi - \zeta\| < \delta).$$

Speciálisan minden $y \in G_r$, $\|x - y\| < \delta$ és $t \in [a,b]$ esetén a

$$\xi := (x,t),\ \zeta := (y,t) \in A$$

pontokban

$$\|\xi - \zeta\| = \|(x,t) - (y,t)\| = \|x - y\| < \delta,$$

így

$$|f(x,t) - f(y,t)| < \varepsilon.$$

Következésképpen

$$|F(y) - F(x)| \leq \int_a^b |f(y,t) - f(x,t)|\, dt \leq (b-a)\cdot \varepsilon,$$

ami azt jelenti, hogy $F \in \mathcal{C}\{x\}$. Mivel az x az U halmaz bármelyik eleme lehetett, ezért az F függvény folytonos.

Ezzel az 1^o állítást beláttuk. A 2^o bizonyítása is hasonlóan történhet. Legyen tehát $i = 1, ..., n$, és tegyük fel, hogy létezik a folytonos $\partial_i f$ parciális

---

[63]Heinrich Eduard Heine (1821 – 1881).

deriváltfüggvény. Ekkor minden $x \in U$ és az 1^o bizonyításában szereplő $r > 0$ számmal

$$F(x_1, ..., x_{i-1}, x_i + h, x_{i+1}, ..., x_n) - F(x_1, ..., x_n) =$$

$$\int_a^b \Big(f(x_1, ..., x_{i-1}, x_i + h, x_{i+1}, ..., x_n, t) - f(x_1, ..., x_n, t)\Big)\, dt \qquad (|h| \leq r).$$

Az egyváltozós valós függvényekre vonatkozó Lagrange-középérték-tétel, ill. a $\partial_i f$ függvény értelmezése alapján egy (h-tól függő) $\theta \in (0, 1)$ számmal

$$f(x_1, ..., x_{i-1}, x_i + h, x_{i+1}, ..., x_n, t) - f(x_1, ..., x_n, t) =$$

$$\partial_i f(x_1, ..., x_{i-1}, x_i + \theta \cdot h, x_{i+1}, ..., x_n, t) \cdot h =: \partial_i f(\tilde{x}, t) \cdot h.$$

Az 1^o állítás bizonyításában bevezetett A halmazra mondottakat az f helyett a $\partial_i f$-re alkalmazva bármilyen $\varepsilon > 0$ mellett kapunk olyan $\delta > 0$ számot, amellyel[64]

$$|\partial_i f(\tilde{x}, t) - \partial_i f(x, t)| < \varepsilon \qquad (t \in [a, b],\ |h| < \delta).$$

Legyen most már $0 < |h| < \delta$, ekkor a fentiek szerint

$$\left| \frac{F(x_1, ..., x_{i-1}, x_i + h, x_{i+1}, ..., x_n) - F(x_1, ..., x_n)}{h} - \int_a^b \partial_i f(x, t)\, dt \right| =$$

$$\left| \int_a^b \Big(\partial_i f(\tilde{x}, t) - \partial_i f(x, t)\Big)\, dt \right| \leq \int_a^b \Big|\partial_i f(\tilde{x}, t) - \partial_i f(x, t)\Big|\, dt \leq (b - a) \cdot \varepsilon.$$

Ezért létezik a

$$\lim_{h \to 0} \frac{F(x_1, ..., x_{i-1}, x_i + h, x_{i+1}, ..., x_n) - F(x_1, ..., x_n)}{h} = \int_a^b \partial_i f(x, t)\, dt$$

határérték, ami a 2^o állítás bizonyítását jelenti.

Ha $f \in C^1$, akkor valamennyi $i = 1, ..., n$ esetén létezik és folytonos a $\partial_i f$ függvény. Ezért (ld. 2^o) léteznek a

$$\partial_i F \qquad (i = 1, ..., n)$$

parciális deriváltfüggvények is, mint a $\partial_i f$-ek paraméteres integráljai:

$$\partial_i F(x) = \int_a^b \partial_i f(x, t)\, dt \qquad (x \in U).$$

[64] $\|(\tilde{x}, t) - (x, t)\| = \|\tilde{x} - x\| = |(1 - \theta)h| < |h|.$

Így 1^o szerint a $\partial_i F \quad (i = 1, ..., n)$ függvények folytonosak. Innen a 3^o állítás már következik.[65] ■

2. Többváltozós függvények Riemann-integrálja.

A továbbiakban legyen $1 \leq n \in \mathbf{N}$ egy adott természetes szám. Ekkor $\mathbf{R}^n$-beli *intervallumon* az

$$I := [a_1, b_1] \times \ldots \times [a_n, b_n]$$

Descartes-szorzatot értjük, ahol

$$a_i,\, b_i \in \mathbf{R},\ a_i < b_i \qquad (i = 1, ..., n).$$

Világos, hogy $n = 1$ esetén a „szokásos" (korlátos és zárt) $\mathbf{R}$-beli intervallumot, $n = 2$-re a koordináta-síkon a koordináta-tengelyekkel párhuzamos oldalú téglalapot, $n = 3$-ra pedig a térbeli derékszögű koordináta-rendszerben a koordináta-síkokkal párhuzamos oldallapú téglatestet kapunk.

Az $\mathbf{R}^n$-ben a $\|.\|_p$ normák közül $(1 \leq p \leq +\infty)$ bármelyiket is vezetve be, az I halmaz korlátos és zárt, azaz kompakt. A

$$d_I := \sup\{\|x - y\|_2 : x, y \in I\} = \sqrt{\sum_{j=1}^{n} (b_j - a_j)^2}$$

szám az I intervallum *átmérője,*

$$|I| := \prod_{j=1}^{n} (b_j - a_j)$$

pedig az I *mértéke.* Nyilván $|I| \leq d_I^n$.

Tehát $n = 1$ esetén

$$|I| = b_1 - a_1$$

az $I \subset \mathbf{R}$ intervallum *hossza,* $n = 2$-re

$$|I| = (b_1 - a_1) \cdot (b_2 - a_2)$$

az

$$I = [a_1, b_1] \times [a_2, b_2] \subset \mathbf{R}^2$$

[65] A $\partial_i F \quad (i = 1, ..., n)$ függvények folytonossága maga után vonja az F függvény differenciálhatóságát.

téglalap *területe,* ha pedig $n = 3$, akkor

$$|I| = (b_1 - a_1)\cdot (b_2 - a_2)\cdot (b_3 - a_3)$$

az

$$I = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3] \subset \mathbf{R}^3$$

téglatest *térfogata.*

Emlékeztetünk arra, hogy egy $[a, b] \subset \mathbf{R}$ kompakt intervallum felosztásán olyan véges $\tau \subset [a, b]$ halmazt értettünk, amelyre $a, b \in \tau$. Továbbá, ha valamilyen $k \in \mathbf{N}$ mellett $\tau = \{x_0, ..., x_k\}$, akkor mindig azzal a formai feltételezéssel éltünk és élünk, hogy

$$a = x_0 < x_1 < ... < x_k = b.$$

Ha az előbbi τ felosztás esetén

$$I_j := [x_j, x_{j+1}] \qquad (j = 0, ..., k-1),$$

akkor az

$$\mathcal{F}(\tau) := \{I_j : j = 0, ..., k-1\}$$

halmaz elemei a τ felosztás által meghatározott ún. osztásintervallumok.

Legyen ezek után

$$\tau_i = \{x_{i0}, ..., x_{ik_i}\} \subset [a_i, b_i] \qquad (k_i \in \mathbf{N},\ i = 1, ..., n)$$

egy-egy felosztása a szóban forgó $[a_i, b_i]$ intervallumoknak. Ekkor a

$$\tau := \tau_1 \times \ldots \times \tau_n \subset I$$

halmazt az I intervallum *felosztásának* nevezzük. Az

$$I_{j_1} \times \ldots \times I_{j_n} \qquad (I_{j_s} \in \mathcal{F}(\tau_s),\ s = 1, ..., n)$$

(a fenti értelemben nyilvánvalóan) intervallumok a τ felosztás által meghatározott *osztásintervallumok.* Ezek halmazára is (az $n = 1$ esettel analóg módon) az $\mathcal{F}(\tau)$ jelölést fogjuk használni. Világos, hogy

$$|I| = \sum_{J \in \mathcal{F}(\tau)} |J|.$$

Azt mondjuk, hogy a $\tau \subset I$ felosztás *finomítása* a $\mu \subset I$ felosztásnak, ha $\mu \subset \tau$.

A fenti $I \subset \mathbf{R}^n$ intervallum és egy

$$f : I \to \mathbf{R}$$

korlátos függvény esetén tetszőleges $\tau \subset I$ felosztást véve definiáljuk az m_J, $M_J \in \mathbf{R}$ $(J \in \mathcal{F}(\tau))$ számokat a következőképpen:

$$m_J := m_J(f) := \inf\{f(x) : x \in J\},$$

$$M_j := M_J(f) := \sup\{f(x) : x \in J\}.$$

Ekkor

$$s(f, \tau) := \sum_{J \in \mathcal{F}(\tau)} m_J \cdot |J|$$

az f függvénynek a τ felosztáshoz tartozó *alsó összege,*

$$S(f, \tau) := \sum_{J \in \mathcal{F}(\tau)} M_J \cdot |J|$$

pedig az f függvénynek a τ felosztáshoz tartozó *felső összege.*[66] Világos, hogy bármely $\tau \subset I$ felosztásra fennáll az

$$s(f, \tau) \leq S(f, \tau)$$

egyenlőtlenség.

2. Tétel. *Legyen adott az $I \subset \mathbf{R}^n$ intervallumon az*

$$f : [a, b] \to \mathbf{R}$$

korlátos függvény. Ekkor tetszőleges τ, $\mu \subset I$ felosztások esetén

1^o $s(f, \tau) \leq S(f, \mu)$;

2^o *ha $\mu \subset \tau$, akkor $s(f, \mu) \leq s(f, \tau)$ és $S(f, \mu) \geq S(f, \tau)$.*

Jelöljük az I intervallum felosztásainak a halmazát $\mathcal{F}_I$-vel. Az előbbi tétel alapján tehát (az abban szereplő f függvényre) az

$$\{s(f, \tau) : \tau \in \mathcal{F}_I\}$$

[66] Ha pl. $n = 2$ és $f \geq 0$, akkor $M_J \cdot |J|$, $m_J \cdot |J|$ a J alaplapú és M_J, m_J magasságú (az f grafikonja fölé és alá beírt) egy-egy hasáb térfogata.

halmaz felülről korlátos, és minden $\mu \in \mathcal{F}_I$ felosztásra az utóbbi halmaznak az $S(f,\mu)$ felső összeg egy felső korlátja. Ezért

$$I_*(f) := \sup\{s(f,\tau) : \tau \in \mathcal{F}_I\} \leq S(f,\mu) < +\infty.$$

Hasonlóan, az

$$\{S(f,\tau) : \tau \in \mathcal{F}_I\}$$

halmaz alulról korlátos, aminek minden $\mu \in \mathcal{F}_I$ felosztásra az $s(f,\mu)$ alsó összeg egy alsó korlátja. Így

$$I^*(f) := \inf\{S(f,\tau) : \tau \in \mathcal{F}_I\} \geq s(f,\mu) > -\infty.$$

Következésképpen $I_*(f)$, $I^*(f) \in \mathbf{R}$ és

$$I_*(f) \leq I^*(f).$$

Az $I_*(f)$-et az f függvény *Darboux*[67]*-féle alsó integráljának,* az $I^*(f)$-et pedig az f függvény *Darboux-féle felső integráljának* nevezzük. A fentiek szerint tetszőleges $\tau, \mu \in \mathcal{F}_I$ felosztásokra

$$s(f,\tau) \leq I_*(f) \leq I^*(f) \leq S(f,\mu).$$

Azt mondjuk, hogy az f függvény *Riemann*[68]*-integrálható,* ha

$$I_*(f) = I^*(f).$$

Ekkor az

$$\int_I f := I_*(f)\,(= I^*(f))$$

számot az f függvény *Riemann-integráljának* (vagy más szóval *határozott integráljának*) nevezzük. Ez utóbbira használatosak még az

$$\int_I f(x)\,dx \text{ vagy az } \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(x_1,...,x_n)\,dx_1 \dots dx_n$$

jelölések is.

Az előbbiekben értelmezett Riemann-integrálható függvények halmazát az $R(I)$ szimbólummal fogjuk jelölni.

Legyen az

$$f : I \to \mathbf{R}$$

[67] Jean Gaston Darboux (1842–1917).
[68] Georg Friedrich Bernhard Riemann (1826–1866).

függvény korlátos, $\tau \in \mathcal{F}_I$. Az

$$\omega(f,\tau) := S(f,\tau) - s(f,\tau)$$

számot az f függvény τ által meghatározott *oszcillációs összegének* nevezzük. Világos, hogy

$$I^*(f) - I_*(f) \le \omega(f,\tau).$$

Az alsó, felső összegek definíciója szerint

$$\omega(f,\tau) = \sum_{J\in\mathcal{F}(\tau)} (M_J - m_J)\cdot |J|,$$

ahol a J intervallumon vett $M_J - m_J$ *oszcillációra*

$$M_J - m_J = \sup\{|f(x) - f(t)| : x,t \in J\} = \sup\{f(x) - f(t) : x,t \in J\}.$$

3. Tétel. *Tegyük fel, hogy az $f : I \to \mathbf{R}$ függvény korlátos. Ekkor a következő ekvivalencia igaz: $f \in R(I)$ akkor és csak akkor, ha tetszőleges $\varepsilon > 0$ számhoz van olyan $\tau \in \mathcal{F}_I$ felosztás, amellyel $\omega(f,\tau) < \varepsilon$.*

A „többváltozós" Riemann-integrálra is teljesül az „egydimenziós" esetben megismert monotonitási tulajdonság. Ezt fejezi ki az alábbi tétel.

4. Tétel. *Legyen $f \in R(I)$. Ekkor:*

a) $|f| \in R(I)$ *és* $\left|\int_I f\right| \le \int_I |f|$;

b) *ha* $g \in R(I)$ *és* $f(x) \le g(x)$ $(x \in I)$, *akkor* $\int_I f \le \int_I g$;

c) *ha* $f \in R(I)$, *és valamilyen* $0 \le c \in \mathbf{R}$ *számmal* $|f(x)| \le c$ $(x \in I)$, *akkor*

$$\left|\int_I f\right| \le c\cdot |I|.$$

A Riemann-integrál a „többváltozós" esetben is rendelekezik azzal a tulajdonsággal, miszerint „érzéketlen" a szóban forgó függvény véges halmazon való „viselkedésére".

5. Tétel. *Tekintsük az* I *intervallumon az*

$$f,\, g : I \to \mathbf{R}$$

korlátos függvényeket. Tegyük fel, hogy az

$$A := \{x \in I : f(x) \neq g(x)\}$$

halmaz véges. Ekkor:

a) $f \in R(I) \iff g \in R(I)$;

b) *ha* $f \in R(I)$, *akkor* $\int_I f = \int_I g$.

A folytonosság „erősebb" tulajdonság a Riemann-integrálhatóságnál. Ezzel kapcsolatban emlékeztetünk arra, hogy pl. az

$$f(x) := \begin{cases} 0 & (0 < x \leq 1) \\ \\ 1 & (x = 0) \end{cases}$$

függvényre $f \in R[0,1] \setminus C[0,1]$. Ugyanakkor igaz a

6. Tétel. *Ha az* $f : I \to \mathbf{R}$ *függvény folytonos, akkor* $f \in R(I)$.

3. Az integrál kiszámítása. Szukcesszív integrálás.

A továbbiakban a „többváltozós" Riemann-integrál egyik leggyakrabban alkalmazott tulajdonságával, az ún. *szukcesszív integrálással* foglalkozunk. Legyen ehhez $2 \leq n \in \mathbf{N}$ és valamilyen $s = 1, ..., n-1$ mellett adottak az

$$1 \leq i_1 < \ldots < i_s \leq n$$

indexek. Ha

$$\xi = (\xi_1, ..., \xi_n) \in \mathbf{R}^n,$$

akkor a $\xi \equiv (x, y)$ szimbólum jelentse a következőt:

$$x := (\xi_{i_1}, ..., \xi_{i_s}) \in \mathbf{R}^s,\ y := (\xi_{j_1}, ..., \xi_{j_{n-s}}) \in \mathbf{R}^{n-s},$$

ahol $j_1 < \ldots < j_{n-s}$ és

$$\{j_1, ..., j_{n-s}\} := \{1, ..., n\} \setminus \{i_1, ..., i_s\}.$$

Ennek megfelelően az $\mathbf{R}^n$ teret az

$$\mathbf{R}^n \equiv \mathbf{R}^s \times \mathbf{R}^{n-s}$$

alakban tekintjük. Továbbá minden

$$f \in \mathbf{R}^n \to \mathbf{R}$$

függvényt

$$f \in \mathbf{R}^s \times \mathbf{R}^{n-s} \to \mathbf{R}$$

„kétváltozós" függvényként fogunk fel, ahol az f függvény

$$f(x,y) \qquad ((x,y) \in \mathcal{D}_f)$$

helyettesítési értékeiben az (x,y) „vektor" első komponense $\mathbf{R}^s$-beli, a második pedig $\mathbf{R}^{n-s}$-beli. Hasonlóan, ha

$$I = [a_1, b_1] \times \ldots \times [a_n, b_n] \subset \mathbf{R}^n$$

intervallum, akkor legyen

$$I \equiv I_1 \times I_2,$$

ahol

$$I_1 := [a_{i_1}, b_{i_1}] \times \ldots \times [a_{i_s}, b_{i_s}]$$

és

$$I_2 := [a_{j_1}, b_{j_1}] \times \ldots \times [a_{j_{n-s}}, b_{j_{n-s}}]$$

$\mathbf{R}^s$-beli, ill. $\mathbf{R}^{n-s}$-beli intervallumok.

Mindezt figyelembe véve legyen

$$f : I_1 \times I_2 \to \mathbf{R}$$

adott „kétváltozós" függvény, és tetszőleges $x \in I_1$, $y \in I_2$ esetén tekintsük az

$$f_x : I_2 \to \mathbf{R},$$

ill. az

$$f^y : I_1 \to \mathbf{R}$$

függvényt, ahol

$$f_x(z) := f(x,z) \qquad (z \in I_2),$$

ill.

$$f^y(t) := f(t,y) \qquad (t \in I_1).$$

Ha az f korlátos, akkor nyilván az f_x, f^y $(x \in I_1,\, y \in I_2)$ függvények is korlátosak. Következésképpen vehetjük azokat az

$$\mathbf{m}_f,\, \mathbf{M}_f : I_1 \to \mathbf{R}, \ \text{ valamint } \ \mathbf{m}^f,\, \mathbf{M}^f : I_2 \to \mathbf{R}$$

függvényeket, amiket az alábbi módon értelmezünk:

$$\mathbf{m}_f(x) := I_*(f_x) \quad (x \in I_1),\ \mathbf{M}_f(x) := I^*(f_x) \quad (x \in I_1),$$

$$\mathbf{m}^f(y) := I_*(f^y) \quad (y \in I_2),\ \mathbf{M}^f(y) := I^*(f^y) \quad (y \in I_2).$$

A most bevezetett jelölésekkel a *szukcesszív integrálás* tétele a következőképpen szól:

7. Tétel (Fubini[69]). *Tegyük fel, hogy* $f \in R(I)$. *Ekkor*

$$\mathbf{m}_f,\ \mathbf{M}_f \in R(I_1) \ \text{ és } \ \mathbf{m}^f,\ \mathbf{M}^f \in R(I_2),$$

továbbá

$$\int_{I_1} \mathbf{m}_f = \int_{I_1} \mathbf{M}_f = \int_{I_2} \mathbf{m}^f = \int_{I_2} \mathbf{M}^f = \int_I f.$$

Bizonyítás. Legyen $J = J_1 \times J_2 \subset I$ egy részintervalluma az I-nek. Ekkor nyilván minden rögzített $x \in J_1$ mellett

$$m_J(f) = m_{J_1 \times J_2}(f) = \inf\{f(t,z) : (t,z) \in J\} \leq$$

$$\inf\{f(x,z) : z \in J_2\} = m_{J_2}(f_x).$$

Ezért tetszőleges

$$\tau = \tau_1 \times \tau_2 \subset I \equiv I_1 \times I_2$$

felosztásra (ahol tehát $\tau_1 \subset I_1$, $\tau_2 \subset I_2$ bármilyen felosztások lehetnek) és $J_1 \in \mathcal{F}(\tau_1)$ osztásintervallumra

$$\mathbf{m}_f(x) \geq s(f_x, \tau_2) = \sum_{J_2 \in \mathcal{F}(\tau_2)} m_{J_2}(f_x) \cdot |J_2| \geq$$

$$\sum_{J_2 \in \mathcal{F}(\tau_2)} m_{J_1 \times J_2}(f) \cdot |J_2| \qquad (x \in J_1).$$

Innen

$$I_*(\mathbf{m}_f) \geq s(\mathbf{m}_f, \tau_1) = \sum_{J_1 \in \mathcal{F}(\tau_1)} \inf\{\mathbf{m}_f(x) : x \in J_1\} \cdot |J_1| \geq$$

[69] Guido Fubini (1879 – 1943).

$$\geq \sum_{J_1\in\mathcal{F}(\tau_1)} \sum_{J_2\in\mathcal{F}(\tau_2)} m_{J_1\times J_2}(f)\cdot |J_2|\cdot |J_1| = s(f,\tau),$$

tehát (a τ-ban szuprémumot véve)

$$I_*(\mathbf{m}_f) \geq I_*(f) = \int_I f$$

következik. Ugyanígy kapjuk az

$$I^*(\mathbf{m}_f) \leq I^*(f) = \int_I f$$

egyenlőtlenséget, amiből

$$\int_I f \leq I_*(\mathbf{m}_f) \leq I^*(\mathbf{m}_f) \leq \int_I f$$

miatt

$$I_*(\mathbf{m}_f) = I^*(\mathbf{m}_f) = \int_I f$$

következik. Ez nem más, mint az $\mathbf{m}_f$-re vonatkozó állításunk.

A tétel többi állítása analóg módon látható be. ∎

4. Megjegyzések.

i) Ha az $f : I \to \mathbf{R}$ függvény folytonos, akkor tetszőleges $x \in I_1$, $y \in I_2$ mellett (ld. **1.**) az f_x, f^y függvények is folytonosak. Következésképpen $f_x \in R(I_2)$ és $f^y \in R(I_1)$. Ezért azt írhatjuk, hogy

$$\mathbf{m}_f(x) = \mathbf{M}_f(x) = \int_{I_2} f_x(y)\,dy = \int_{I_2} f(x,y)\,dy \qquad (x \in I_1),$$

$$\mathbf{m}^f(y) = \mathbf{M}^f(y) = \int_{I_1} f^y(x)\,dx = \int_{I_1} f(x,y)\,dx \qquad (y \in I_2),$$

azaz a 7. Tétel szerint

$$\int_{I_1\times I_2} f(x,y)\,dx\,dy = \int_{I_1} \left(\int_{I_2} f(x,y)\,dy\right) dx = \int_{I_2} \left(\int_{I_1} f(x,y)\,dx\right) dy.$$

Formálisan tehát: egy kétváltozós folytonos függvény integrálját kiszámíthatjuk úgy is, hogy az egyik változót először rögzítjük, és a másik változó szerint integrálunk, majd az így kapott (a rögzített változótól függő) integrált integráljuk. (Innen ered a *szukcesszív* (egymás utáni) jelző.)

ii) Ha $f \in R(I)$ és minden $x \in I_1$, $y \in I_2$ mellett $f_x \in R(I_2)$, valamint $f^y \in R(I_1)$, akkor lehet szukcesszíve integrálni. Ugyanakkor nem minden $f \in R(I)$ függvényre teljesül „automatikusan", hogy $f_x \in R(I_2)$, $f^y \in R(I_1)$. Legyen ui. $I := [0,1] \times [0,1]$ és

$$f(x,y) := \begin{cases} 0 & (0 < x \leq 1 \text{ és } y \in [0,1]) \\ \\ 1 & (x = 0 \text{ és } y \in [0,1] \cap \mathbf{Q}) \\ \\ 0 & (x = 0 \text{ és } y \in [0,1] \setminus \mathbf{Q}). \end{cases}$$

Ekkor $f \in R(I)$, de $f_0 \notin R[0,1]$.

iii) Megmutatható, hogy igaz a következő állítás: ha az $f : I_1 \times I_2 \to \mathbf{R}$ függvény korlátos, és minden $x \in I_1$, $y \in I_2$ esetén léteznek az

$$m_f(x) := \int_{I_2} f(x,z)\,dz,\ m^f(y) := \int_{I_1} f(t,y)\,dt$$

integrálok, akkor $m_f \in R(I_1)$, $m^f \in R(I_2)$, és (dacára annak, hogy az f esetleg *nem integrálható* az $I_1 \times I_2$-n)

$$\int_{I_1} m_f(x)\,dx = \int_{I_2} m^f(y)\,dy,$$

azaz

$$\int_{I_1} \left(\int_{I_2} f(x,y)\,dy \right) dx = \int_{I_2} \left(\int_{I_1} f(x,y)\,dx \right) dy.$$

iv) Vegyük észre, hogy a ii) megjegyzésben szereplő f függvény „majdnem" folytonos. Ti. az $A := \{0\} \times [0,1]$ halmaz („szakasz") pontjaitól eltekintve a $[0,1] \times [0,1]$ „négyzet" minden pontjában az. Az is nyilvánvaló, hogy a ii) megjegyzésbeli u számmal $A \subset J$, ahol $J := [0,u] \times [0,1]$ és $|J| = u$. (Az A *nullamértékű* halmaz (ld. következő előadás).)

12. Fejezet

0. Emlékeztető.

I) Legyen $1 \leq n \in \mathbf{N}$, ekkor $\mathbf{R}^n$-beli *intervallumon* az

$$I := [a_1, b_1] \times \ldots \times [a_n, b_n]$$

Descartes-szorzatot értjük, ahol

$$a_i,\, b_i \in \mathbf{R},\ a_i < b_i \qquad (i = 1, ..., n).$$

Az

$$|I| := \prod_{j=1}^{n} (b_j - a_j)$$

szám az I *mértéke.*

Ha $\tau = \{x_0, ..., x_k\}$ $(k \in \mathbf{N})$ az $[a, b] \subset \mathbf{R}$ kompakt intervallum felosztása, akkor

$$a = x_0 < x_1 < ... < x_k = b.$$

Legyen

$$\mathcal{F}(\tau) := \{I_j := [x_j, x_{j+1}] : j = 0, ..., k-1\}.$$

A $\tau_i \subset [a_i, b_i]$ $(i = 1, ..., n)$ felosztásokkal a

$$\tau := \tau_1 \times \ldots \times \tau_n \subset I$$

halmazt az I intervallum *felosztásának* nevezzük. Az

$$I_{j_1} \times \ldots \times I_{j_n} \qquad (I_{j_s} \in \mathcal{F}(\tau_s),\ s = 1, ..., n)$$

osztásintervallumok halmaza legyen most is $\mathcal{F}(\tau)$.

II) Az $I \subset \mathbf{R}^n$ intervallum és egy $f : I \to \mathbf{R}$ korlátos függvény esetén tetszőleges $\tau \subset I$ felosztásra a $J \in \mathcal{F}(\tau)$ osztásintervallumokkal legyen

$$m_J := m_J(f) := \inf\{f(x) : x \in J\},$$

$$M_j := M_J(f) := \sup\{f(x) : x \in J\},$$

valamint

$$s(f, \tau) := \sum_{J \in \mathcal{F}(\tau)} m_J \cdot |J|,\ S(f, \tau) := \sum_{J \in \mathcal{F}(\tau)} M_J \cdot |J|.$$

Jelöljük az I intervallum felosztásainak a halmazát $\mathcal{F}_I$-vel. Ekkor

$$I_*(f) := \sup\{s(f, \tau) : \tau \in \mathcal{F}_I\} < +\infty,$$

$$I^*(f) := \inf\{S(f, \tau) : \tau \in \mathcal{F}_I\} > -\infty.$$

Továbbá

$$I_*(f) \le I^*(f).$$

Az f *Riemann-integrálható függvény* $(f \in R(I))$, ha $I_*(f) = I^*(f)$. Ekkor az

$$\int_I f := I_*(f)\, (= I^*(f))$$

számot az f függvény *Riemann-integráljának* nevezzük. Ez utóbbira használatosak még az

$$\int_I f(x)\,dx \text{ vagy az } \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(x_1,...,x_n)\,dx_1 \dots dx_n$$

jelölések is.

III) Szukcesszív-integrálás: legyen $2 \le n \in \mathbf{N}$ és $s = 1, ..., n-1$. Minden $f \in \mathbf{R}^n \to \mathbf{R}$ függvényt

$$f \in \mathbf{R}^s \times \mathbf{R}^{n-s} \to \mathbf{R}$$

„kétváltozós" függvényként fogunk fel. Hasonlóan, az $I \subset \mathbf{R}^n$ intervallum $I = I_1 \times I_2$ alakú, ahol $I_1 \subset \mathbf{R}^s$, $I_2 \subset \mathbf{R}^{n-s}$.

Legyen az

$$f : I_1 \times I_2 \to \mathbf{R}$$

függvény korlátos, és $x \in I_1$, $y \in I_2$ esetén

$$f_x(z) := f(x,z) \quad (z \in I_2), \; f^y(t) := f(t,y) \quad (t \in I_1).$$

Továbbá

$$\mathbf{m}_f(x) := I_*(f_x) \quad (x \in I_1), \; \mathbf{M}_f(x) := I^*(f_x) \quad (x \in I_1),$$

$$\mathbf{m}^f(y) := I_*(f^y) \quad (y \in I_2), \; \mathbf{M}^f(y) := I^*(f^y) \quad (y \in I_2).$$

Ekkor igaz a Fubini-tétel: bármely $f \in R(I)$ függvényre

$$\mathbf{m}_f, \; \mathbf{M}_f \in R(I_1) \text{ és } \mathbf{m}^f, \; \mathbf{M}^f \in R(I_2)$$

és

$$\int_{I_1} \mathbf{m}_f = \int_{I_1} \mathbf{M}_f = \int_{I_2} \mathbf{m}^f = \int_{I_2} \mathbf{M}^f = \int_I f.$$

Ha az $f : I \to \mathbf{R}$ függvény folytonos, akkor az f_x, f^y $(x \in I_1,\, y \in I_2)$ függvények is folytonosak, így $f_x \in R(I_2)$ és $f^y \in R(I_1)$. Ezért

$$\mathbf{m}_f(x) = \mathbf{M}_f(x) = \int_{I_2} f_x(y)\,dy = \int_{I_2} f(x,y)\,dy \qquad (x \in I_1),$$

$$\mathbf{m}^f(y) = \mathbf{M}^f(y) = \int_{I_1} f^y(x)\,dx = \int_{I_1} f(x,y)\,dx \qquad (y \in I_2),$$

azaz a Fubini-tétel szerint

$$\int_{I_1 \times I_2} f(x,y)\,dx\,dy = \int_{I_1} \left(\int_{I_2} f(x,y)\,dy \right) dx = \int_{I_2} \left(\int_{I_1} f(x,y)\,dx \right) dy.$$

1. Integrálás normáltartományon.

Legyen $2 \le n \in \mathbf{N}$, és valamilyen $k = 1, ..., n$ mellett tekintsük a

$$\xi = (\xi_1, ..., \xi_n) \in \mathbf{R}^n$$

vektorokat $\xi = (x, y)$ alakban, ahol

$$x := (\xi_1, ..., \xi_{k-1}, \xi_{k+1}, ..., \xi_n) \in \mathbf{R}^{n-1},\ y := \xi_k \in \mathbf{R}.$$

Tegyük fel, hogy valamilyen (kompakt)

$$I \subset \mathbf{R}^{n-1},\ J = [a, b] \subset \mathbf{R}$$

intervallumokkal

$$K := \{\xi \in \mathbf{R}^n : x \in I,\, y \in J\} \equiv I \times J,$$

és adottak a

$$g,\, h : I \to J$$

folytonos függvények, amelyekre

$$g(t) \le h(t) \qquad (t \in I).$$

Ekkor az

$$\mathcal{N} := \{\xi \in K : x \in I,\, g(x) \le y \le h(x)\} \subset \mathbf{R}^n$$

halmaz egy ún. *normáltartomány.* Tételezzük fel továbbá, hogy az

$$f : \mathcal{N} \to \mathbf{R}$$

függvény folytonos, és legyen a fenti K intervallummal

$$F(\xi) := \begin{cases} 0 & (\xi \in K \setminus \mathcal{N}) \\ f(\xi) & (\xi \in \mathcal{N}). \end{cases}$$

Ekkor $F \in R(K)$ és

$$\int_{\mathcal{N}} f := \int_K F = \int_I \left(\int_{g(x)}^{h(x)} f(x, y)\, dy \right) dx.^{70}$$

[70] Az f *integrálja* az $\mathcal{N}$ normáltartományon. „Koordinátás" alakban $\int_{\mathcal{N}} f(\xi)\, d\xi =$ $\int_I \Big(\int_{g(\xi_1, ..., \xi_{k-1}, \xi_{k+1}, ..., \xi_n)}^{h(\xi_1, ..., \xi_{k-1}, \xi_{k+1}, ..., \xi_n)} f(\xi_1, ..., \xi_n)\, d\xi_k \Big)\, d\xi_1 \dots d\xi_{k-1}\, d\xi_{k+1} \dots d\xi_n.$

Ha $n = 2$, akkor $I \subset \mathbf{R}$ és az alábbi normáltartományokról lehet szó:

$$\mathcal{N}_{21} := \{(u,v) \in \mathbf{R}^2 : u \in I,\ g(u) \leq v \leq h(u)\},$$

$$\mathcal{N}_{22} := \{(u,v) \in \mathbf{R}^2 : v \in I,\ g(v) \leq u \leq h(v)\}.$$

Ennek megfelelően

$$\int_{\mathcal{N}_{21}} f = \int_I \left(\int_{g(u)}^{h(u)} f(u,v)\,dv \right) du,$$

$$\int_{\mathcal{N}_{22}} f = \int_I \left(\int_{g(v)}^{h(v)} f(u,v)\,du \right) dv.$$

Hasonlóan, ha $n = 3$, akkor $I \subset \mathbf{R}^2$ és

$$\mathcal{N}_{31} := \{(u,v,z) \in \mathbf{R}^3 : (u,v) \in I,\ g(u,v) \leq z \leq h(u,v)\},$$

$$\mathcal{N}_{32} := \{(u,v,z) \in \mathbf{R}^3 : (u,z) \in I,\ g(u,z) \leq v \leq h(u,z)\},$$

$$\mathcal{N}_{33} := \{(u,v,z) \in \mathbf{R}^3 : (v,z) \in I,\ g(v,z) \leq u \leq h(v,z)\},$$

továbbá

$$\int_{\mathcal{N}_{31}} f = \int_I \left(\int_{g(u,v)}^{h(u,v)} f(u,v,z)\,dz \right) du\,dv,$$

$$\int_{\mathcal{N}_{32}} f = \int_I \left(\int_{g(u,z)}^{h(u,z)} f(u,v,z)\,dv \right) du\,dz,$$

$$\int_{\mathcal{N}_{33}} f = \int_I \left(\int_{g(v,z)}^{h(v,z)} f(u,v,z)\,du \right) dv\,dz.$$

2. Megjegyzések.

i) Legyen $1 \leq s \in \mathbf{N}$, $A \subset \mathbf{R}^s$. Azt mondjuk, hogy az A halmaz *nullamértékű,* ha tetszőleges $\varepsilon > 0$ számhoz megadható $I_k \subset \mathbf{R}^s$ $(k \in \mathbf{N})$ intervallumoknak egy olyan sorozata, hogy

$$A \subset \bigcup_{k=0}^{\infty} I_k \ \text{ és } \ \sum_{k=0}^{\infty} |I_k| < \varepsilon.$$

Pl. egyszerűen belátható: az $\mathbf{R}^s$ minden, legfeljebb megszámlálható részhalmaza nullamértékű. Sőt, ha $X_k \subset \mathbf{R}^s$ $(k \in \mathbf{N})$ nullamértékű, akkor az $\bigcup_{k=0}^{\infty} X_k$ halmaz is nullamértékű. Világos továbbá, hogy egy nullamértékű halmaz minden részhalmaza is nullamértékű.

ii) Tekintsük pl. az $I \subset \mathbf{R}^n$ $(1 \le n \in \mathbf{N})$ intervallum esetén az $f \in R(I)$ függvényt, és legyen

$$\mathrm{graf}\, f := \{(x, f(x)) \in \mathbf{R}^{n+1} : x \in I\}.$$[71]

Ekkor az $\mathrm{graf}\, f \subset \mathbf{R}^{n+1}$ halmaz nullamértékű.

iii) A Riemann-integrálhatóság *Lebesgue*[72]*-kritériuma* a következőképpen szól: *tegyük fel, hogy az $I \subset \mathbf{R}^n$ intervallumon értelmezett $f : I \to \mathbf{R}$ függvény korlátos, és legyen az f szakadási helyeinek a halmaza*

$$\mathcal{A}_f := \Big\{x \in I : f \notin \mathcal{C}\{x\}\Big\}.$$

Ekkor az f Riemann-integrálhatósága, azaz $f \in R(I)$ azzal ekvivalens, hogy az $\mathcal{A}_f$ halmaz nullamértékű.[73]

iv) A fentiekben az $\mathcal{N}$ normáltartományon megadott $f : \mathcal{N} \to \mathbf{R}$ folytonos függvényhez[74] definált F függvény legfeljebb a $\mathrm{graf}\, g \cup \mathrm{graf}\, h$ nullamértékű halmaz pontjaiban nem folytonos. Tehát (ld. iii)) $F \in R(I)$ és

$$\int_K F = \int_I \left(\int_J F(x,y)\,dy\right) dx = \int_I \left(\int_{g(x)}^{h(x)} f(x,y)\,dy\right) dx.$$

v) Ha az $f \in \mathbf{R}^n \to \mathbf{R}$ függvény $\mathcal{D}_f \subset \mathbf{R}^n$ értelmezési tartománya korlátos, akkor van olyan $I \subset \mathbf{R}^n$ intervallum, amellyel $\mathcal{D}_f \subset I$. Legyen ekkor

$$F_I(x) := \begin{cases} f(x) & (x \in \mathcal{D}_f) \\ \\ 0 & (x \in I \setminus \mathcal{D}_f). \end{cases}$$

Nem nehéz meggondolni, hogy ha $F_I \in R(I)$, akkor $F_J \in R(J)$ minden olyan $J \subset \mathbf{R}^n$ intervallumra, amelyre $\mathcal{D}_f \subset J$ és

$$\int_I F_I = \int_J F_J.$$

Ez ad értelmet az alábbi definíciónak: a fenti f függvény *Riemann-integrálható,* ha egy alkalmas $I \subset \mathbf{R}^n$ intervallummal $\mathcal{D}_f \subset I$ és $F_I \in R(I)$. Az utóbbi esetben

$$\int_{\mathbf{R}^n} f := \int_{\mathcal{D}_f} f := \int_I F_I$$

[71] Az f „grafikonja", ami persze valójában maga az f függvény.

[72] Henri Léon Lebesgue (1875–1941).

[73] A „szokásos" szóhasználattal: az f függvény *majdnem mindenütt folytonos* (röviden: $f \in \mathcal{C}\{x\}$ (m.m. $x \in I$)).

[74] Mivel az $\mathcal{N}$ kompakt, ezért az f korlátos.

az f *Riemann-integrálja.* Világos, hogy ha a $\mathcal{D}_f$ intervallum, akkor visszakapjuk az integrálhatóság, ill. az integrál „eredeti" definícióját. Speciálisan, ha a $\mathcal{D}_f$ normáltartomány és az f folytonos, akkor az f integrálható.

vi) Az $f \in \mathbf{R}^n \to \mathbf{R}$ függvény (esetleges) gyökhelyeitől különböző $\mathcal{D}_f$-beli elemek halmazának a lezárását, azaz a

$$\operatorname{supp} f := \overline{\{x \in \mathcal{D}_f : f(x) \neq 0\}}$$

halmazt az f függvény *tartójának* nevezzük (az angol „support" szó után). Világos, hogy a $\operatorname{supp} f$ halmaz zárt, és az $x \in \mathcal{D}_f \setminus \operatorname{supp} f$ helyeken $f(x) = 0$.[75] Ha a $\operatorname{supp} f$ korlátos is, azaz kompakt, akkor egy alkalmas $I \subset \mathbf{R}^n$ intervallummal $\operatorname{supp} f \subset I$. Legyen ebben az esetben most

$$F_I(x) := \begin{cases} f(x) & (x \in \mathcal{D}_f \cap I) \\ \\ 0 & (x \in I \setminus \mathcal{D}_f). \end{cases}$$

Belátható, hogy ha $F_I \in R(I)$, akkor egyúttal $F_J \in R(J)$ minden olyan $J \subset \mathbf{R}^n$ intervallumra, amelyre $\operatorname{supp} f \subset J$, továbbá

$$\int_I F_I = \int_J F_J.$$

Van értelme tehát most is az alábbi definíciónak:

$$\int_{\mathbf{R}^n} f := \int_{\mathcal{D}_f} f := \int_I F_I$$

(az f *(Riemann-integrálja)*, ahol az I egy előbbi tulajdonságú (tetszőleges) intervallum.

3. Integráltranszformáció.

Legyen $f \in R[a,b]$, $\varphi \in C^1[\alpha,\beta]$, ahol $a = \varphi(\alpha)$, $b = \varphi(\beta)$, valamint $\mathcal{R}_\varphi = [a,b]$. Ekkor tetszőleges $f \in C[a,b]$ függvényre

$$\int_a^b f = \int_\alpha^\beta (f \circ \varphi) \cdot \varphi'.$$

Ha $\varphi' > 0$, akkor a φ szigorúan monoton növő, következésképpen az

$$a = \varphi(\alpha) < b = \varphi(\beta)$$

[75] Ugyanakkor előfordulhat, hogy bizonyos $x \in \operatorname{supp} f$ helyeken is $f(x) = 0$.

és az $\mathcal{R}_\varphi = [a, b]$ feltételek „automatikusan" teljesülnek. Ha viszont $\varphi' < 0$, akkor a φ szigorúan monoton fogyó, $\varphi(\alpha) > \varphi(\beta)$ és $\mathcal{R}_\varphi = [b, a]$. Ezért $f \in C[b, a]$ esetén

$$\int_b^a f = \int_\alpha^\beta (f \circ \varphi) \cdot \varphi',$$

azaz

$$\int_a^b f := -\int_b^a f = -\int_\alpha^\beta (f \circ \varphi) \cdot \varphi',$$

amiből formálisan

$$\int_a^b f = \int_\alpha^\beta (f \circ \varphi) \cdot (-\varphi') = \int_\alpha^\beta (f \circ \varphi) \cdot |\varphi'|$$

adódik.

Összefoglalva azt mondhatjuk, hogy ha

$$\varphi'(x) \neq 0 \qquad (x \in [\alpha, \beta])$$

(azaz a φ' Darboux-tulajdonsága miatt $\varphi' > 0$, vagy pedig $\varphi' < 0$), akkor

$$\int_a^b f = \int_\alpha^\beta (f \circ \varphi) \cdot |\varphi'|.$$

Az integráltranszformációval kapcsolatos következő, eléggé általános feltételek melletti állítás igazolható:

1. Tétel. *Tekintsük a nyílt halmazon értelmezett és folytonosan differenciálható*

$$g \in \mathbf{R}^n \to \mathbf{R}^n$$

függvényt. Tegyük fel, hogy az $I \subset \mathcal{D}_g$ halmaz kompakt intervallum, továbbá az I belsejére való $g_{|\mathrm{int}\, I}$ leszűkítés injektív függvény. Ekkor az

$$f : g[I] \to \mathbf{R}$$

korlátos függvény akkor és csak akkor integrálható, ha az

$$I \ni x \mapsto f(g(x)) \cdot |\det g'(x)|$$

függvény is integrálható.[76] *Az utóbbi esetben*

$$\int_I f(g(x)) \cdot |\det g'(x)| \, dx = \int_{g[I]} f.$$

[76] Emlékeztetünk arra, hogy $g'(x) \in \mathbf{R}^{n \times n}$ a g függvény Jacobi-mátrixa az x pontban, $\det g'(x)$ ennek a mátrixnak a determinánsa.

4. Megjegyzések.

i) Legyen az $A \subset \mathbf{R}^n$ halmaz korlátos,

$$\chi_A(x) := \begin{cases} 1 & (x \in A) \\ \\ 0 & (x \in \mathbf{R}^n \setminus A) \end{cases}$$

az A *karakterisztikus függvénye.* Azt mondjuk, hogy az A halmaz *Jordan*[77]*-mérhető,* ha a χ_A függvény integrálható, amikor is a

$$\mu(A) := \int_{\mathbf{R}^n} \chi_A$$

nemnegatív szám az A halmaz *Jordan-mértéke.*[78] Nyilvánvaló, hogy $\mu(\emptyset) = 0$, valamint minden $J \subset \mathbf{R}^n$ intervallum Jordan-mérhető és $\mu(J) = |J|$. Ha pl. az $A := \mathcal{N}$ egy normáltartomány, akkor az $\mathcal{N}$ Jordan-mérhető és (ld. **1.**)

$$\mu(\mathcal{N}) = \int_{\mathbf{R}^n} \chi_{\mathcal{N}} = \int_I \Big(\int_{g(x)}^{h(x)} 1\, dy \Big)\, dx = \int_I \Big(h(x) - g(x) \Big) dx.$$[79]

Pl. az $n = 2$ esetben az

$$\mathcal{N} := \{(x, y) \in \mathbf{R}^2 : x \in [-1, 1],\ 0 \le y \le \sqrt{1 - x^2}\}$$

halmaz (félkörlemez) normáltartomány, ahol

$$g(x) := 0,\ h(x) := \sqrt{1 - x^2} \qquad (-1 \le x \le 1).$$

Ezért

$$\mu(\mathcal{N}) = \int_{-1}^{1} \sqrt{1 - x^2}\, dx = 2 \cdot \int_0^1 \sqrt{1 - x^2}\, dx = 2 \cdot \int_0^{\pi/2} \cos^2 t\, dt =$$

$$\int_0^{\pi/2} (1 + \cos(2t))\, dt = \frac{\pi}{2}$$

(a félkörlemez „jól ismert" területe). Ugyanakkor az

$$A := \{x \in [0, 1] : x \in \mathbf{Q}\}$$

---

[77]Marie Ennemond Camille Jordan (1838–1922).

[78]Gondoljuk meg, hogy ha az $A, B \subset \mathbf{R}^n$ halmazok Jordan-mérhetők, akkor az $A \cup B$ egyesítésük is az. Ha még diszjunktak is, akkor $\mu(A \cup B) = \mu(A) + \mu(B)$.

[79]Speciálisan, ha itt $n = 2$ és $g \equiv 0$, akkor $\mu(\mathcal{N}) = \int_I h$ a h grafikonja alatti terület nagysága.

halmaz (a $[0,1]$-beli racionális számok halmaza) nem Jordan-mérhető, ui. (pl.) az $I := [0,1]$ választással az

$$F_I(x) := \begin{cases} \chi_A(x) = 1 & (x \in A) \\ \\ 0 & (x \in I \setminus A) \end{cases}$$

függvény a Dirichlet[80]-függvény, ami nem Riemann-integrálható.

ii) Az 1. Tételben az I intervallum helyett tetszőleges $\emptyset \neq A \subset \mathcal{D}_g$ kompakt, Jordan-mérhető halmaz is vehető. Ekkor az is igaz, hogy a $g[A]$ képhalmaz is kompakt, Jordan-mérhető halmaz.

iii) Legyenek az $\emptyset \neq U, V \subset \mathbf{R}^n$ halmazok nyíltak, a $\varphi : U \to V$ egy folytonosan differenciálható bijekció, amire

$$\det \varphi'(x) \neq 0 \qquad (x \in U).$$

Jelöljük az ilyen φ-k halmazát $\mathcal{C}(U,V)$-vel. Tekintsünk egy kompakt tartójú, folytonos $f : V \to \mathbf{R}$ függvényt, amiről feltesszük, hogy

$$\operatorname{supp} f \subset V,$$

amikor is $\mathcal{K}(V)$ az ilyen f függvények összessége. Ha pl. $\varphi \in \mathcal{C}(U,V)$, akkor

$$(f \circ \varphi) \cdot |\det \varphi'| \in \mathcal{K}(U).$$

Igaz az alábbi tétel: *a* $\varphi \in \mathcal{C}(U,V)$, $f \in \mathcal{K}(V)$ *függvényekkel*

$$\int_V f = \int_U (f \circ \varphi) \cdot |\det \varphi'|.$$

5. Speciális esetek.

a) Legyen $n = 2$, és tekintsük azt a $g : \mathbf{R}^2 \to \mathbf{R}^2$ leképezést, amire

$$g(u,v) := (u \cdot \cos v, u \cdot \sin v) \qquad \big((u,v) \in \mathbf{R}^2\big)$$

(*síkbeli polárkoordináta-transzformáció*). Világos, hogy $g \in C^1$, továbbá

$$g'(u,v) = \begin{bmatrix} \cos v & -u \cdot \sin v \\ \sin v & u \cdot \cos v \end{bmatrix} \text{ és } \det g'(u,v) = u \qquad \big((u,v) \in \mathbf{R}^2\big).$$

[80] Johann Peter Gustav Lejeune Dirichlet(1805–1859).

Az is könnyen belátható, hogy bármilyen

$$\emptyset \neq A \subset [0, +\infty) \times [0, 2\pi]$$

kompakt, Jordan-mérhető halmaz esetén a g függvény $g_{|\mathrm{int}\, A}$ leszűkítése injektív. Ha tehát az

$$f : g[A] \to \mathbf{R}$$

függvény integrálható, akkor a $g[A]$ képhalmaz is kompakt, Jordan-mérhető, és

$$\int_A |u| \cdot f(u \cdot \cos v, u \cdot \sin v)\, du\, dv = \int_{g[A]} f(x, y)\, dx\, dy.$$

Ez a helyzet pl. akkor, ha az f függvény folytonos. Ha mondjuk

$$A := [0, 1] \times [0, 2\pi],$$

akkor

$$g[A] = K_1 := \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 \leq 1\},$$

ezért

$$\int_{K_1} f(x, y)\, dx\, dy = \int_0^1 \int_0^{2\pi} u \cdot f(u \cdot \cos v, u \cdot \sin v)\, du\, dv.$$

Speciálisan az $f \equiv 1$ függvényre

$$\int_{K_1} f(x, y)\, dx\, dy = \mu(K_1) = \int_0^1 u \left(\int_0^{2\pi} dv \right) du = 2\pi \cdot \int_0^1 u\, du = \pi$$

(a K_1 körlemez területe).

b) Az előbbiek „térbeli" megfelelője a $g : \mathbf{R}^3 \to \mathbf{R}^3$ transzformáció, ahol

$$g(u, v, w) := (u \cdot \sin w \cdot \cos v, u \cdot \sin w \cdot \sin v, u \cdot \cos w) \qquad \big((u, v, w) \in \mathbf{R}^3\big)$$

(*térbeli polárkoordináta-transzformáció*). Nyilván $g \in C^1$ és

$$g'(u, v) = \begin{bmatrix} \sin w \cdot \cos v & -u \cdot \sin w \cdot \sin v & u \cdot \cos w \cdot \cos v \\ \sin w \cdot \sin v & u \cdot \sin w \cdot \cos v & u \cdot \cos w \cdot \sin v \\ \cos w & 0 & -u \cdot \sin w \end{bmatrix} \qquad \big((u, v) \in \mathbf{R}^3\big),$$

valamint

$$\det g'(u, v, w) =$$

$$\cos w \cdot \Big(-u^2 \cdot \sin^2 v \cdot \sin w \cdot \cos w - u^2 \cdot \cos^2 v \cdot \sin w \cdot \cos w \Big) -$$

$$u \cdot \sin w \Big(u \cdot \sin^2 w \cdot \cos^2 v + u \cdot \sin^2 w \cdot \sin^2 v \Big) =$$

$$= -u^2\cdot \sin w\cdot \cos^2 w - u^2\cdot \sin^3 w = -u^2\cdot \sin w \qquad \Big((u,v,w)\in \mathbf{R}^3\Big).$$

Ha az

$$\emptyset \neq A \subset [0,+\infty)\times[0,2\pi]\times[0,\pi]$$

kompakt, Jordan-mérhető halmaz, akkor a $g_{|\mathrm{int}\, A}$ leszűkítés injektív. Ezért a $g[A]$ képhalmaz is kompakt, Jordan-mérhető, és az integrálható

$$f : g[A] \to \mathbf{R}$$

függvény esetén (pl., ha az f folytonos)

$$\int_A u^2\cdot f(u\cdot \sin w\cdot \cos v, u\cdot \sin w\cdot \sin v, u\cdot \cos w)\cdot \sin w\, du\, dv\, dw =$$

$$\int_{g[A]} f(x,y,z)\, dx\, dy\, dz.$$

Legyen

$$A := [0,1]\times[0,2\pi]\times[0,\pi],$$

ekkor

$$g[A] = G_1 := \{(x,y,z)\in \mathbf{R}^3 : x^2+y^2+z^2 \le 1\},$$

és

$$\int_{G_1} f(x,y,z)\, dx\, dy\, dz =$$

$$\int_0^1 \int_0^{2\pi} \int_0^{\pi} u^2\cdot f(u\cdot \sin w\cdot \cos v, u\cdot \sin w\cdot \sin v, u\cdot \cos w)\cdot \sin w\, du\, dv\, dw.$$

Az $f \equiv 1$ esetben

$$\int_{G_1} f(x,y,z)\, dx\, dy\, dz = \mu(G_1) = \int_0^1 u^2 \left(\int_0^{2\pi}\left(\int_0^{\pi} \sin w\, dw\right) dv\right) du =$$

$$= 2\pi\cdot \int_0^1 u^2 \left(\int_0^{\pi} \sin w\, dw\right) du = 4\pi\cdot \int_0^1 u^2\, du = \frac{4\pi}{3}$$

(a G_1 gömb térfogata).

c) Legyen most a $g \in \mathbf{R}^3 \to \mathbf{R}^3$ olyan, hogy

$$g(u,v,w) := (u\cdot \cos v, u\cdot \sin v, w) \qquad \Big((u,v,w)\in \mathbf{R}^3\Big)$$

(*hengerkoordináta-transzformáció*). Ekkor nyilván $g \in C^1$ és

$$g'(u,v) = \begin{bmatrix} \cos v & -u\cdot \sin v & 0 \\ \sin v & u\cdot \cos v & 0 \\ 0 & 0 & 1 \end{bmatrix} \qquad \Big((u,v)\in \mathbf{R}^3\Big),$$

továbbá

$$\det g'(u,v,w) = u \qquad \Big((u,v) \in \mathbf{R}^3\Big).$$

Ha az

$$\emptyset \neq A \subset [0,+\infty) \times [0,2\pi] \times \mathbf{R}$$

tetszőleges kompakt, Jordan-mérhető halmaz, akkor egyszerűen ellenőrizhető, hogy a $g_{|\mathrm{int}\, A}$ leszűkítés injektív. Így a $g[A]$ képhalmaz is kompakt, Jordan-mérhető, és az integrálható

$$f : g[A] \to \mathbf{R}$$

függvényre (ilyen pl. minden folytonos függvény)

$$\int_A u \cdot f(u \cdot \cos v, u \cdot \sin v, w)\, du\, dv\, dw = \int_{g[A]} f(x,y,z)\, dx\, dy\, dz.$$

Pl. az

$$A := [0,1] \times [0,2\pi] \times [0,1]$$

halmazzal

$$g[A] = H_1 := \{(x,y,z) \in \mathbf{R}^3 : x^2 + y^2 \leq 1\,; 0 \leq z \leq 1\},$$

és

$$\int_{H_1} f(x,y,z)\, dx\, dy\, dz = \int_0^1 \int_0^{2\pi} \int_0^1 u \cdot f(u \cdot \cos v, u \cdot \sin v, w)\, du\, dv\, dw.$$

Az $f \equiv 1$ esetben

$$\int_{H_1} f(x,y,z)\, dx\, dy\, dz = \mu(H_1) = \int_0^1 u \left(\int_0^{2\pi} \left(\int_0^1 dw \right) dv \right) du =$$

$$= 2\pi \cdot \int_0^1 u\, du = \pi$$

(a H_1 henger térfogata).

6. Megjegyzések.

i) Tekintsük az $I \subset \mathbf{R}^n$ kompakt intervallumon adott korlátos $f : I \to \mathbf{R}$ függvényt és az I egy $\tau \in \mathcal{F}_I$ felosztását. Ekkor tetszőleges $J \in \mathcal{F}(\tau)$ osztásintervallum és $y_J \in J$ pont választással legyen

$$\sigma(f,\tau,y) := \sum_{J \in \mathcal{F}(\tau)} f(y_J) \cdot |J|,$$

ahol az y az alábbi „vektor":

$$y := (y_J, J \in \mathcal{F}(\tau)).$$

Ez utóbbi $\sigma(f,\tau,y)$ összeget az f függvény (integrál) *közelítő összegének* nevezzük. Legyen továbbá

$$\widehat{\tau} := \{(y_J, J \in \mathcal{F}(\tau)) : y_J \in J \quad (J \in \mathcal{F}(\tau))\}.$$

Nyilvánvaló, hogy bármelyik $\tau \in \mathcal{F}_I$ felosztásra

$$s(f,\tau) \leq \sigma(f,\tau,y) \leq S(f,\tau) \qquad (y \in \widehat{\tau}).$$

ii) Belátható, hogy (az i)-beli f függvénnyel) igaz az alábbi tétel:

a) *tetszőleges* $\varepsilon > 0$ *számhoz megadható olyan* $\delta > 0$, *hogy*

$$|s(f,\tau) - I_*(f)| < \varepsilon, \ |S(f,\tau) - I^*(f)| < \varepsilon \quad (\tau \in \mathcal{F}_I, \ \delta_\tau < \delta);$$

b) *az* $f \in R(I)$, $\int_I f = q \, (\in \mathbf{R})$ *kijelentés azzal ekvivalens, hogy bármilyen* $\varepsilon > 0$ *számhoz létezik olyan* $\delta > 0$, *amellyel*

$$|\sigma(f,\tau,y) - q| < \varepsilon \qquad (\tau \in \mathcal{F}_I, \ \delta_\tau < \delta, \ y \in \widehat{\tau})$$

(ahol $d_J := \sup\{\|x-y\|_2 : x,y \in J\}$ a J osztásintervallum „átmérője", $\delta_\tau := \sup\{d_J : J \in \mathcal{F}(\tau)\}$ a τ felosztás „finomsága).

iii) Legyen adott a $\Phi : \mathcal{F}_I \to \mathbf{R}$ függvény. Azt mondjuk, hogy $\delta_\tau \to 0$ esetén a Φ-nek *van határértéke,* ha létezik olyan $\alpha \in \mathbf{R}$ szám, amelyre minden $\varepsilon > 0$ esetén egy alkalmas $0 < \delta$-val

$$|\Phi(\tau) - \alpha| < \varepsilon \qquad \big(\tau \in \mathcal{F}_I, \ \delta_\tau < \delta\big).$$

Ekkor egyetlen ilyen α van, amit a Φ *határértékének* (vagy *limeszének*) nevezünk, és a $\lim \Phi$ $\big($vagy a $\lim_{\delta_\tau \to 0} \Phi(\tau)\big)$ szimbólummal jelölünk.

iv) Ha például az $f : I \to \mathbf{R}$ egy korlátos függvény, akkor léteznek az alábbi határértékek:

$$\lim_{\delta_\tau \to 0} s(f,\tau) = I_*(f), \ \lim_{\delta_\tau \to 0} S(f,\tau) = I^*(f).$$

Az $f \in R(I)$ esetben (tetszőleges $y \in \hat{\tau}$ $(\tau \in \mathcal{F}_I)$ választással) létezik a $\lim_{\delta_\tau \to 0} \sigma(f,\tau,y)$ határérték is és

$$\lim_{\delta_\tau \to 0} \sigma(f,\tau,y) = \lim_{\delta_\tau \to 0} s(f,\tau) = \lim_{\delta_\tau \to 0} S(f,\tau) = \int_I f.$$

v) Az ún. *középérték-tétel* számos gyakorlati alkalmazással bír:

> *tegyük fel, hogy a fenti $I \subset \mathbf{R}^n$ intervallummal $f,\, g \in R(I)$, ahol $g \geq 0$. Ekkor:*
>
> a) *az $m := \inf \mathcal{R}_f$, $M := \sup \mathcal{R}_f$ jelölésekkel*
>
> $$m\cdot \int_I g \leq \int_I fg \leq M\cdot \int_I g;$$
>
> b) *ha az f függvény folytonos, akkor létezik olyan $\xi \in I$, hogy*
>
> $$\int_I fg = f(\xi)\cdot \int_I g.$$

13. Fejezet

1. Alkalmazások.

a) **Tömeg.** Képzeljük el, hogy a

$$G_1 := \{(x, y, z) \in \mathbf{R}^3 : x^2 + y^2 + z^2 \leq 1\}$$

gömböt valamilyen anyag tölti ki. Legyen ez utóbbinak a

$$\rho : G_1 \to [0, +\infty)$$

függvény a sűrűségfüggvénye, ahol a ρ-ról feltételezzük, hogy folytonos.

Emlékeztetünk arra, hogy ha valamilyen $q > 0$ számmal $\rho \equiv q$, akkor tetszőleges $V \subset G_1$ téglatest esetén a V-ben lévő anyag T_V tömege $|V|\cdot q$, ahol $|V|$ a szóban forgó téglatest térfogata.[81]

Ha a ρ függvény nem állandó, akkor legyen $I := [-1, 1]^3$, amikor is az I olyan intervallum (kocka), hogy $G_1 \subset I$. Terjesszük ki a ρ függvényt az I-re a „szokásos" módon:

$$\widetilde{\rho}(\xi) := \begin{cases} \rho(\xi) & (\xi \in G_1) \\ 0 & (\xi \in I \setminus G_1). \end{cases}$$

[81] Nem foglalkozunk a „sűrűség", ill. a „tömeg" egzakt fizikai értelmezésével. Elfogadjuk, hogy a G_1-ben, ill. a V-ben lévő anyagnak van tömege, és az utóbbi $|V|\cdot q$, továbbá azonos sűrűség esetén a nagyobb térfogatú test tömege nagyobb, mint egy kisebb térfogatú test tömege, valamint azonos térfogatú testek esetén a nagyobb sűrűségű test tömege a nagyobb.

Vegyük az I valamilyen $\tau \in \mathcal{F}_I$ felosztását, ekkor az

$$m_V := \inf\{\widetilde{\rho}(\xi) : \xi \in V \cap G_1\} \qquad (V \in \mathcal{F}(\tau)),$$

$$M_V := \sup\{\widetilde{\rho}(\xi) : \xi \in V \cap G_1\} \qquad (V \in \mathcal{F}(\tau))$$

jelölésekkel és az

$$s(\widetilde{\rho}, \tau) = \sum_{V \in \mathcal{F}(\tau),\, V \cap G_1 \neq \emptyset} m_V \cdot |V|,$$

$$S(\widetilde{\rho}, \tau) = \sum_{V \in \mathcal{F}(\tau),\, V \cap G_1 \neq \emptyset} M_V \cdot |V|$$

alsó, felső összegekkel a G_1 gömb $\mathcal{M}_{G_1}$ tömegére

$$s(\widetilde{\rho}, \tau) \leq \mathcal{M}_{G_1} \leq S(\widetilde{\rho}, \tau).$$

A ρ függvény integrálható, és az $\int_{G_1} \rho = \int_I \widetilde{\rho}$ szám az egyetlen olyan, amelyik minden $\tau \in \mathcal{F}_I$ esetén az $s(\widetilde{\rho}, \tau)$, $S(\widetilde{\rho}, \tau)$ összegek közé esik. Ezért az előbbieket figyelembe véve a „fizikai" elvárásoknak megfelelőnek tűnik az a matematikai definíció, hogy

$$\mathcal{M}_{G_1} := \int_{G_1} \rho$$

a G_1 gömb (pontosabban a benne lévő anyag) *tömege.* Speciálisan (ld. fent), ha $\rho \equiv q$, akkor

$$\mathcal{M}_{G_1} = q \cdot \int_{G_1} 1 = q \cdot |G_1|.$$

Ennek megfelelően általában is azt mondjuk, hogy ha adott egy

$$\emptyset \neq A \subset \mathbf{R}^3$$

„test", amit egy

$$\rho : A \to [0, +\infty)$$

sűrűségű anyag tölt ki, továbbá $\rho \in R(A)$, akkor a szóban forgó testnek van tömege ($\mathcal{M}_A$), ami az

$$\mathcal{M}_A := \int_A \rho$$

integrál.[82]

[82]Hangsúlyozni kell, hogy nem „kiszámoltuk" az illető test tömegét, hanem – a gyakorlati tapasztalatokkal összhangban – matematikailag „deklaráltuk" azt.

Legyen pl. a fenti G_1 esetén az őt kitöltő anyag olyan, amelynek a sűrűsége egyenesen arányos a gömb középpontjától vett távolsággal. Ez matematikailag azt jelenti, hogy valamilyen $\lambda > 0$ számmal

$$\rho(\xi) = \lambda \cdot \|\xi\|_2 \qquad (\xi \in G_1).$$

Világos, hogy a ρ függvény folytonos, továbbá (a térbeli polárkoordinátatranszformációt is alkalmazva)

$$\mathcal{M}_{G_1} = \int_{G_1} \rho(\xi)\, d\xi = \int_{G_1} \rho(x,y,z)\, dx\, dy\, dz =$$

$$\lambda \cdot \int_{G_1} \sqrt{x^2+y^2+z^2}\, dx\, dy\, dz =$$

$$= \lambda \cdot \int_0^1 \int_0^{2\pi} \int_0^{\pi} u^2 \cdot \rho(u \cdot \sin w \cdot \cos v, u \cdot \sin w \cdot \sin v, u \cdot \cos w) \cdot \sin w\, du\, dv\, dw,$$

ahol az $(u,v,w) \in [0,1] \times [0,2\pi] \times [0,\pi]$ helyeken

$$\rho(u \cdot \sin w \cdot \cos v, u \cdot \sin w \cdot \sin v, u \cdot \cos w) =$$

$$u \cdot \sqrt{\sin^2 w \cdot \cos^2 v + \sin^2 w \cdot \sin^2 v + \cos^2 w} = u.$$

Innen azt kapjuk, hogy

$$\mathcal{M}_{G_1} = \lambda \cdot \int_0^1 \int_0^{2\pi} \int_0^{\pi} u^3 \cdot \sin w\, du\, dv\, dw =$$

$$\lambda \cdot \int_0^1 u^3 \left(\int_0^{2\pi} \left(\int_0^{\pi} \sin w\, dw \right) dv \right) du =$$

$$= \lambda \cdot 2\pi \cdot \left(\int_0^1 u^3 du \right) \cdot \left(\int_0^{\pi} \sin w\, dw \right) = \lambda \cdot \pi.$$

b) **Súlypont.** Valamilyen $1 \leq m \in \mathbf{N}$ esetén a

$$P_i := (x_i, y_i, z_i) \in \mathbf{R}^3 \qquad (i = 1, ..., m)$$

helyek mindegyikében elhelyezünk egy

$$m_i > 0 \qquad (i = 1, ..., m)$$

tömegű pontot. Ekkor az így kapott $(P_i, i = 1, ..., m)$ (diszkrét) pontrendszer

$$S = (x_s, y_s, z_s) \in \mathbf{R}^3$$

súlypontján azt a pontot értettük, amire

$$x_s := \frac{1}{m} \cdot \sum_{i=1}^{s} m_i \cdot x_i,$$

$$y_s := \frac{1}{m} \cdot \sum_{i=1}^{s} m_i \cdot y_i,$$

$$z_s := \frac{1}{m} \cdot \sum_{i=1}^{s} m_i \cdot z_i,$$

ahol

$$m := \sum_{i=1}^{s} m_i$$

a pontrendszer össztömege.

Ennek megfelelően egy $\emptyset \neq A \subset \mathbf{R}^3$ test és az azt kitöltő

$$\rho : A \to [0, +\infty)$$

sűrűségű anyag esetén a test $S = (x_s, y_s, z_s) \in \mathbf{R}^3$ *súlypontja* az a pont, amelyre a test

$$\mathcal{M}_A = \int_A \rho \ \ (> 0)$$

tömegével

$$x_s := \frac{1}{\mathcal{M}_A} \cdot \int_A x \cdot \rho(x, y, z)\, dx\, dy\, dz,$$

$$y_s := \frac{1}{\mathcal{M}_A} \cdot \int_A y \cdot \rho(x, y, z)\, dx\, dy\, dz,$$

$$z_s := \frac{1}{\mathcal{M}_A} \cdot \int_A z \cdot \rho(x, y, z)\, dx\, dy\, dz$$

(feltéve, hogy az összes itt szereplő integrál létezik).

Hasonlóan definiáljuk a súlypont fogalmát „lemezek" esetén is: ha $\emptyset \neq B \subset \mathbf{R}^2$, és a B-t (lemezt) egy

$$\rho : B \to [0, +\infty)$$

sűrűségű anyag tölti ki, akkor a B lemez

$$S = (x_s, y_s) \in \mathbf{R}^2$$

súlypontja:

$$x_s := \frac{1}{\mathcal{M}_B} \cdot \int_B x \cdot \rho(x, y)\, dx\, dy,$$

$$y_s := \frac{1}{\mathcal{M}_B} \cdot \int_B y \cdot \rho(x,y)\, dx\, dy,$$

ahol

$$\mathcal{M}_B = \int_B \rho \ \ (> 0)$$

a lemez tömege. (Most is feltételezzük, hogy az összes itteni integrál létezik.)

Legyen pl.

$$B := \{(x,y) \in \mathbf{R} \times [0, +\infty) : x^2 + y^2 \le 1\}$$

(félkörlemez), és valamilyen $\lambda > 0$ együtthatóval

$$\rho(\xi) = \lambda \cdot \|\xi\|_2 \qquad (\xi \in B).$$

Ekkor (síkbeli polárkoordináta-transzformációval)

$$\mathcal{M}_B = \lambda \cdot \int_B \rho(x,y)\, dx\, dy = \lambda \cdot \int_B \sqrt{x^2 + y^2}\, dx\, dy =$$

$$\lambda \cdot \int_0^1 \int_0^\pi u^2 \cdot \sqrt{\cos^2 v + \sin^2 v}\, du\, dv = \lambda \cdot \int_0^1 u^2 \left(\int_0^\pi dv \right) du =$$

$$\pi \cdot \lambda \cdot \int_0^1 u^2 du = \frac{\pi \cdot \lambda}{3}.$$

Továbbá

$$x_s = \frac{3}{\pi \cdot \lambda} \cdot \int_B x \cdot \rho(x,y)\, dx\, dy = \frac{3}{\pi} \cdot \int_B x \cdot \sqrt{x^2 + y^2}\, dx\, dy =$$

$$= \frac{3}{\pi} \cdot \int_0^1 \int_0^\pi u^3 \cdot \cos v\, du\, dv = \frac{3}{\pi} \cdot \int_0^1 u^3 \left(\int_0^\pi \cos v\, dv \right) du = \frac{3}{\pi} \cdot \frac{1}{4} \cdot 0 = 0$$

(ami a szimmetriaviszonyok alapján nem „meglepő"). Ugyanakkor

$$y_s = \frac{3}{\pi \cdot \lambda} \cdot \int_B y \cdot \rho(x,y)\, dx\, dy = \frac{3}{\pi} \cdot \int_B y \cdot \sqrt{x^2 + y^2}\, dx\, dy =$$

$$\frac{3}{\pi} \cdot \int_0^1 \int_0^\pi u^3 \cdot \sin v\, du\, dv = \frac{3}{\pi} \cdot \int_0^1 u^3 \left(\int_0^\pi \sin v\, dv \right) du = \frac{3}{\pi} \cdot \frac{1}{4} \cdot 2 = \frac{3}{2\pi}.$$

c) **Tehetetlenségi nyomaték.** A mechanikában tanultak alapján egy m tömegű tömegpontnak a tőle r távolságban lévő tengelyre vonatkozóan az ún. tehetetlenségi nyomatéka $m \cdot r^2$.

Ebből kiindulva vezessük be a tehetetlenségi nyomaték általános fogalmát az alábbiak szerint. Tegyük fel ehhez, hogy adott az

$$\emptyset \neq A \subset \mathbf{R}^3$$

test, amit a

$$\rho : A \to [0, +\infty)$$

sűrűségű anyag tölt ki. Legyen adott továbbá valamilyen $0 \neq e \in \mathbf{R}^3$ irányvektor és egy $a \in \mathbf{R}^3$ pont esetén a

$$\mathbf{R} \ni t \mapsto a + t\cdot e \in \mathbf{R}^3$$

hozzárendeléssel definiált ℓ egyenes („tengely"). Egy $\xi \in \mathbf{R}^3$ pontnak az ℓ tengelytől mért távolságát jelöljük $r_\ell(\xi)$-vel. Tehát

$$r_\ell(\xi) := \inf\{\|\xi - (a + t\cdot e)\|_2 : t \in \mathbf{R}\}.$$

Ekkor a szóban forgó testnek az ℓ tengelyre vonatkozó *tehetetlenségi nyomatékán* a

$$\Theta_\ell := \int_A r_\ell^2(\xi)\cdot \rho(\xi)\, d\xi$$

integrált értjük (feltéve, hogy ez az integrál létezik).

Speciálisan, az ℓ_i-vel jelölt $(i = 1, 2, 3)$ koordináta-tengelyekre vonatkozó tehetetlenségi nyomatékokat a következőképpen számíthatjuk ki. Most tehát $a := (0, 0, 0)$, az e vektor pedig rendre

$$e_1 := (1, 0, 0),\ e_2 := (0, 1, 0),\ e_3 := (0, 0, 1).$$

Világos, hogy tetszőleges $\xi = (x, y, z) \in \mathbf{R}^3$ mellett

$$r_{\ell_1}^2(x, y, z) = y^2 + z^2,$$

$$r_{\ell_2}^2(x, y, z) = x^2 + z^2,$$

$$r_{\ell_3}^2(x, y, z) = x^2 + y^2,$$

azaz

$$\Theta_{\ell_1} = \int_A (y^2 + z^2)\cdot \rho(x, y, z)\, dx\, dy\, dz,$$

$$\Theta_{\ell_2} = \int_A (x^2 + z^2)\cdot \rho(x, y, z)\, dx\, dy\, dz,$$

$$\Theta_{\ell_3} = \int_A (x^2 + y^2)\cdot \rho(x, y, z)\, dx\, dy\, dz.$$

Legyen pl.

$$A := G_1 := \{(x,y,z) \in \mathbf{R}^3 : x^2+y^2+z^2 \le 1\}$$

(gömb), és valamilyen $\rho_0 > 0$ konstanssal $\rho \equiv \rho_0$. Ekkor a nyilvánvaló szimmetriaviszonyok alapján

$$\Theta_{\ell_1} = \Theta_{\ell_2} = \Theta_{\ell_3},$$

és pl. (ismét a térbeli polárkoordináta-transzformáció segítségével)

$$\Theta_{\ell_3} = \int_{G_1} (x^2+y^2)\cdot \rho(x,y,z)\,dx\,dy\,dz =$$

$$\rho_0\cdot \int_0^1 \int_0^{2\pi} \int_0^{\pi} u^2\cdot (u^2\cdot \sin^2 w\cdot \cos^2 v + u^2\cdot \sin^2 w\cdot \sin^2 v)\cdot \sin w\,du\,dv\,dw =$$

$$\rho_0\cdot \int_0^1 \int_0^{2\pi} \int_0^{\pi} u^4\cdot \sin^3 w\,du\,dv\,dw = 2\pi\rho_0\cdot \left(\int_0^1 u^4\,du\right)\cdot\left(\int_0^{\pi} \sin^3 w\,dw\right) =$$

$$\frac{2\pi\rho_0}{5}\cdot \int_0^{\pi} \sin^3 w\,dw = \frac{2\pi\rho_0}{5}\cdot \int_0^{\pi} (\sin w)\cdot (1-\cos^2 w)\,dw =$$

$$= \frac{2\pi\rho_0}{5}\cdot (2-2/3) = \frac{8\pi\rho_0}{15} = \frac{2M}{5},$$

ahol a G_1 gömb $4\pi/3$ térfogatával

$$M := \frac{4\pi}{3}\cdot \rho_0$$

a gömb tömege.

d) **Improprius integrál.** Mutassuk meg, hogy létezik az

$$\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

improprius integrál. Legyen ehhez

$$I_r := \int_0^r e^{-x^2} dx \qquad (r > 0).$$

Azt kell tehát megmutatnunk, hogy létezik a

$$\lim_{r\to+\infty} I_r = \frac{\sqrt{\pi}}{2}$$

határérték. Világos, hogy az $N_r := [0,r]^2$ jelöléssel („négyzettel") a szukcesszív integrálás szabálya alapján

$$I_r^2 = \left(\int_0^r e^{-x^2} dx\right)\cdot\left(\int_0^r e^{-y^2} dy\right) = \int_{N_r} e^{-(x^2+y^2)}\,dx\,dy.$$

Ha

$$K_r := \{(u,v) \in [0,+\infty)^2 : u^2+v^2 \le r^2\}$$

(„negyedkörlemez"), akkor $K_r \subset N_r$, így[83]

$$\int_{K_r} e^{-(x^2+y^2)}\,dx\,dy \le \int_{N_r} e^{-(x^2+y^2)}\,dx\,dy.$$

Továbbá

$$x^2+y^2 > r^2 \qquad \big((x,y) \in N_r \setminus K_r\big).$$

Következésképpen

$$0 \le \int_{N_r} e^{-(x^2+y^2)}\,dx\,dy - \int_{K_r} e^{-(x^2+y^2)}\,dx\,dy =$$

$$= \int_{N_r\setminus K_r} e^{-(x^2+y^2)}\,dx\,dy \le e^{-r^2}\cdot |N_r \setminus K_r| \le e^{-r^2}\cdot r^2 \to 0 \quad (r \to +\infty).$$

Ha tehát létezik a

$$\lim_{r\to+\infty} \int_{K_r} e^{-(x^2+y^2)}\,dx\,dy$$

határérték, akkor létezik

$$\lim_{r\to+\infty} \int_{N_r} e^{-(x^2+y^2)}\,dx\,dy$$

is, és

$$\lim_{r\to+\infty} \int_{N_r} e^{-(x^2+y^2)}\,dx\,dy = \lim_{r\to+\infty} \int_{K_r} e^{-(x^2+y^2)}\,dx\,dy.$$

Az

$$\int_{K_r} e^{-(x^2+y^2)}\,dx\,dy \qquad (r>0)$$

integrálok kiszámítására alkalmazzuk a síkbeli polárkoordináta-transzformációt:

$$\int_{K_r} e^{-(x^2+y^2)}\,dx\,dy = \int_0^r \int_0^{\pi/2} u\cdot e^{-(u^2\cdot\cos^2 v + u^2\cdot\sin^2 v)}\,du\,dv =$$

$$\int_0^r u\cdot e^{-u^2}\left(\int_0^{\pi/2} dv\right) du = \frac{\pi}{4}\cdot \int_0^r 2u\cdot e^{-u^2}\,du = \frac{\pi}{4}\cdot(1-e^{-r^2}),$$

amiből

$$\lim_{r\to+\infty} \int_{K_r} e^{-(x^2+y^2)}\,dx\,dy = \frac{\pi}{4}$$

---

[83] A többszörös Riemann-integrálokra is igaz az „integrálási tartomány" szerinti additivitás, jelen esetben: $\int_{N_r} e^{-(x^2+y^2)}\,dx\,dy = \int_{K_r} e^{-(x^2+y^2)}\,dx\,dy + \int_{N_r\setminus K_r} e^{-(x^2+y^2)}\,dx\,dy$, ahol $\int_{N_r\setminus K_r} e^{-(x^2+y^2)}\,dx\,dy \ge 0$.

adódik. Az előzményeket is figyelembe véve tehát létezik az

$$\int_0^{+\infty} e^{-x^2}\,dx = \lim_{r\to+\infty} I_r = \lim_{r\to+\infty}\sqrt{\int_{N_r} e^{-(x^2+y^2)}\,dx\,dy} = \frac{\sqrt{\pi}}{2}$$

határérték.

e) **Guldin[84]-tétel.** Legyen adott az

$$[a,b]\subset \mathbf{R} \qquad (a,b\in\mathbf{R},\ a<b)$$

kompakt intervallum, továbbá egy

$$0\le f:[a,b]\to\mathbf{R}$$

folytonos függvény, és tekintsük az

$$A := \{(x,y)\in\mathbf{R}^2 : x\in[a,b],\ 0\le y\le f(x)\}$$

halmazt (a függvény(grafikon) alatti síkidomot). A

$$\rho: A\to[0,+\infty),\ \rho\equiv 1$$

sűrűségű homogén anyaggal kitöltött lemez

$$S := (x_s,y_s)\in\mathbf{R}^2$$

súlypontja:

$$x_s = \frac{1}{|A|}\cdot\int_A x\,dx\,dy,$$

$$y_s = \frac{1}{|A|}\cdot\int_A y\,dx\,dy,$$

ahol

$$|A| = \mu(A) = \int_A 1\,dx\,dy$$

az A területe („tömege"). Feltesszük, hogy $|A|>0$. Ekkor (szukcesszív integrálással)

$$y_s = \frac{1}{|A|}\cdot\int_a^b\left(\int_0^{f(x)} y\,dy\right)dx =$$

$$\frac{1}{2\cdot|A|}\cdot\int_a^b f^2(x)\,dx = \frac{1}{2\pi\cdot|A|}\cdot\pi\cdot\int_a^b f^2(x)\,dx.$$

[84]Paul Guldin (1577–1643).

Jól ismert, hogy az f függvénynek az X-tengely körüli „megforgatásával" létrejövő

$$V := \{(x,y,z) \in \mathbf{R}^3 : x \in [a,b],\ y^2 + z^2 \le f^2(x)\}$$

forgástest $|V|$ térfogata:

$$|V| = \pi\cdot \int_a^b f^2(x)\,dx.$$

Következésképpen

$$|V| = |A|\cdot (2\pi\cdot y_s).$$

Itt $2\pi\cdot y_s$ annak a körnek a kerülete, amelyet az S súlypontnak az X-tengely körüli megforgatásával kapunk.[85]

f) **Felszín.** Tegyük fel, hogy a $\Phi : [0,1]^2 \to \mathbf{R}^3$ függvény folytonosan differenciálható, és a $\partial_1\Phi(\xi)$, $\partial_2\Phi(\xi)$ parciális deriváltvektorok egyetlen egy $\xi \in [0,1]^2$ esetén sem párhuzamosak.[86] Tekintsük a $[0,1]^2$ „négyzet" valamilyen $\tau \subset [0,1]^2$ felosztását, legyen az $ABCD$ téglalap ennek egy osztásintervalluma. Ekkor a $\Phi(A)$, $\Phi(B)$, $\Phi(C)$, $\Phi(D)$ felületi pontok által „kijelölt"

$$\{\Phi(\xi) \in \mathbf{R}^3 : \xi \in ABCD\}$$

felületdarabot a $\Phi(D) - \Phi(A)$, $\Phi(B) - \Phi(A)$ vektorok által kifeszített paralelogrammával helyettesítjük.[87] Ez utóbbinak a t_{ABCD} területe az

$$a \otimes b = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}$$

[85]Ugyanez adódik akkor is, ha valamilyen $0 \le g \le f : [a,b] \to \mathbf{R}$ folytonos függvénnyel az $A := \{(x,y) \in \mathbf{R}^2 : x \in [a,b],\ g(x) \le y \le f(x)\}$ normáltartományt forgatjuk meg az X-tengely körül. Legyen pl. $0 < r < R$ és $h(x) := \sqrt{r^2 - x^2}$ $(x \in [-r,r])$, valamint $f := R + h$ és $g := R - h$. Ekkor az A normáltartomány egy r sugarú körlemez, a V forgástest pedig egy ***tórusz***. Nyilván $y_s = R$, ezért $|V| = |A|\cdot (2\pi\cdot y_s) = \pi r^2\cdot (2\pi\cdot R) = 2\pi^2 r^2 R$. Mindezt könnyen „ellenőrizhetjük is" a térfogat „direkt" kiszámításával: $|V| = \pi\cdot \int_{-r}^r (f^2(x) - g^2(x))\,dx = 8\pi R \int_0^r \sqrt{r^2 - x^2}\,dx = 8\pi r R \int_0^r \sqrt{1 - (x/r)^2}\,dx$, ahol az $x = r\sin t$ $(0 \le t \le \pi/2)$ helyettesítéssel $\int_0^r \sqrt{1 - (x/r)^2}\,dx = \pi r/4$, és kapjuk az előző eredményt.

[86]Geometriailag a Φ értékkészlete egy „felület", a Φ ennek a (Gauss-féle) „paraméterezése". Ha mondjuk a $\varphi : [a,b] \to [0,1]^2$ függvény folytonosan differenciálható (a $[0,1]^2$ egységnégyzetbeli valamely görbe paraméterezése), akkor a $\psi := \Phi \circ \varphi : [a,b] \to \mathcal{R}_\Phi$ egy ún. ***felületi görbe*** (paraméterezése). Legyen $u \in [a,b]$ és $\xi := \varphi(u)$, amikor is a szóban forgó felületi görbe $\Phi(\xi)$-beli irányvektora: $\psi'(u) = \varphi_1'(u)\partial_1\Phi(\xi) + \varphi_2'(u)\partial_2\Phi(\xi)$.

[87]Feltéve egy pillanatra, hogy ezek a vektorok nem párhuzamosak.

vektoriális szorzat[88] és az $\|a\| := \|a\|_2$ $(a, b \in \mathbf{R}^3)$ norma alapján a következő:

$$t_{ABCD} = \|(\Phi(D) - \Phi(A)) \otimes (\Phi(B) - \Phi(A))\|.$$

Itt a parciális deriváltak definíciójára gondolva közelítőleg

$$t_{ABCD} \approx \|(\partial_2\Phi(A)\cdot|AD|) \otimes (\partial_1\Phi(A)\cdot|AB|)\| =$$

$$\|\partial_1\Phi(A) \otimes \partial_2\Phi(A)\|\cdot|ABCD|,$$

ahol $|AD|$, $|AB|$ az AD, AB szakaszok hosszát, $|ABCD| := |AD|\cdot|AB|$ pedig a jelzett osztásintervallum(téglalap) területét jelöli. A

$$\sum_{ABCD} t_{ABCD} \approx \sum_{ABCD} \|\partial_1\Phi(A) \otimes \partial_2\Phi(A)\|\cdot|ABCD|$$

közelítéssel az utóbbi összeg nem más, mint a

$$\|\partial_1\Phi \otimes \partial_2\Phi\| : [0,1]^2 \to \mathbf{R}$$

folytonos függvénynek a τ felosztáshoz tartozó (Riemann-féle) integrálközelítő összege. Ezért „logikusnak" látszik a szóban forgó felület *felszínét* a

$$\Upsilon_\Phi := \int_{[0,1]^2} \|\partial_1\Phi \otimes \partial_2\Phi\|$$

kettős integrálként definiálni.[89]

Legyen pl. $R > 0$ és az $(u,v) \in [0,1]^2$ helyeken

$$\Phi(u,v) := R\cdot\Big(\sin(\pi v)\cdot\cos(2\pi u), \sin(\pi v)\cdot\sin(2\pi u), \cos(\pi v)\Big)$$

(amikor is $\mathcal{R}_\Phi \subset \mathbf{R}^3$ egy R sugarú gömb). Ekkor

$$\partial_1\Phi(u,v) = 2\pi R\cdot\sin(\pi v)\cdot\Big(-\sin(2\pi u), \cos(2\pi u), 0\Big),$$

$$\partial_2\Phi(u,v) = \pi R\cdot\Big(\cos(\pi v)\cdot\cos(2\pi u), \cos(\pi v)\cdot\sin(2\pi u), -\sin(\pi v)\Big),$$

---

[88] Amit az itt szereplő determinánsnak az első sor szerinti formális kifejtése révén kapunk az $\mathbf{i} := (1,0,0)$, $\mathbf{j} := (0,1,0)$, $\mathbf{k} := (0,0,1) \in \mathbf{R}^3$ vektorokkal.

[89] A $\partial_1\Phi(A)$, $\partial_2\Phi(A)$ vektorok egy síkot határoznak meg, a szóban forgó felület $\Phi(A)$ pontbeli ***érintősíkját,*** amit (ld. előbb) a $\Phi(A)$-n átmenő felületi görbék érintővektorai feszítenek ki. Az érintősík (egy) ***normálvektora*** a $\partial_1\Phi(A) \otimes \partial_2\Phi(A)$ vektor. A fent említett $\Phi(A)$, $\Phi(B)$, $\Phi(C)$, $\Phi(D)$ „csúcspontú" felületdarabot végül ennek a síknak egy darabjával, a $\partial_1\Phi(A)\cdot|AB|$, $\partial_2\Phi(A)\cdot|AD|$ vektorokkal megadott paralelogrammával helyettesítettük. Szemléletesen szólva, ha a τ felosztás elég „finom", akkor ez a paralelogramma „elég közel" van a mondott felületdarabhoz.

valamint (könnyen ellenőrizhetően)

$$\partial_1\Phi(u,v)\otimes\partial_2\Phi(u,v) =$$

$$2\pi^2R^2\cdot\sin(\pi v)\cdot\Big(-\sin(\pi v)\cdot\cos(2\pi u), -\sin(\pi v)\cdot\sin(2\pi u), -\cos(\pi v)\Big).$$

Ezért

$$\|\partial_1\Phi(u,v)\otimes\partial_2\Phi(u,v)\| = 2\pi^2R^2\cdot\sin(\pi v) \qquad \Big((u,v)\in[0,1]^2\Big),$$

amiből

$$\Upsilon_\Phi = 2\pi^2R^2\cdot\int_{[0,1]^2}\sin(\pi v)\,du\,dv = 2\pi^2R^2\cdot\int_0^1\sin(\pi v)\,dv = 4\pi R^2.$$

Ha a fenti Gauss[90]-paraméterezés helyett (explicit) Euler[91]–Monge[92]-módon van megadva a felület, azaz valamilyen folytonosan differenciálható

$$f:[0,1]^2\to\mathbf{R}$$

függvénnyel

$$\Phi(u,v) := (u,v,f(u,v)) \qquad \Big((u,v)\in[0,1]^2\Big),^{93}$$

akkor

$$\partial_1\Phi = (1,0,\partial_1 f),\ \partial_2\Phi = (0,1,\partial_2 f),$$

és így

$$\partial_1\Phi\otimes\partial_2\Phi = (1,0,\partial_1 f)\otimes(0,1,\partial_2 f) = (-\partial_1 f, -\partial_2 f, 1),$$

továbbá az illető felület felszíne:

$$\Upsilon_\Phi = \int_{[0,1]^2}\sqrt{1+(\partial_1 f)^2+(\partial_2 f)^2}.$$

Legyen pl. $0\leq h\in C^1[a,b]$, és tekintsük a h „megforgatásával" létrejövő

$$\mathcal{F} := \{(x,y,z)\in\mathbf{R}^3 : a\leq x\leq b,\ y^2+z^2 = h^2(x)\ \ (y,z\in\mathbf{R})\}$$

forgásfelületet. Világos, hogy az utóbbi negyedrészét az

$$f(u,v) := \sqrt{h^2(u)-v^2} \qquad \Big(a\leq u\leq b;\ 0\leq v\leq h(u)\Big)$$

[90]Johann Carl Friedrich Gauss (1777–1855).

[91]Leonhard Euler (1707–1783).

[92]Gaspard Monge (1746–1818).

[93]Ekkor a szóban forgó felület nem más, mint az f függvény grafikonja. A $[0,1]^2$ „paramétertartomány" helyett vehetnénk más „szerkezetű" halmazokat is (pl. kompakt tartományokat).

függvénnyel paraméterezhetjük explicit Euler-Monge-módon. Ekkor az előbbi (u,v) helyeken

$$\partial_1 f(u,v) = \frac{h(u)\cdot h'(u)}{\sqrt{h^2(u)-v^2}},\ \partial_2 f(u,v) = -\frac{v}{\sqrt{h^2(u)-v^2}},$$

következésképpen (a paraméteres integrállal kapcsolatos tételt és a normáltartományon való integrálásra vonatkozó formulát is felhasználva) a jól ismert

$$\Upsilon_\Phi = 4\cdot \lim_{\varepsilon\to 0}\int_a^b \int_0^{h(u)-\varepsilon} \sqrt{1+\frac{h^2(u)\cdot (h'(u))^2+v^2}{h^2(u)-v^2}}\,du\,dv =$$

$$4\cdot \lim_{\varepsilon\to 0}\int_a^b h(u)\cdot \sqrt{1+(h'(u))^2}\cdot \left(\int_0^{h(u)-\varepsilon} \frac{dv}{\sqrt{h^2(u)-v^2}}\,dv\right)du =$$

$$4\cdot \lim_{\varepsilon\to 0}\int_a^b h(u)\cdot \sqrt{1+(h'(u))^2}\cdot \arcsin\frac{h(u)-\varepsilon}{h(u)}\,du =$$

$$4\cdot \int_a^b h(u)\cdot \sqrt{1+(h'(u))^2}\cdot \arcsin 1\,du = 2\pi\cdot \int_a^b h(u)\cdot \sqrt{1+(h'(u))^2}\,du$$

„képlet" adódik.[94]

Érdemes megjegyezni, hogy az előbbi $\mathcal{F}$ forgásfelület egy másik lehetséges paraméterezése a

$$\Phi(u,v) := (u, h(u)\cdot \cos v, h(u)\cdot \sin v) \qquad \big((u,v)\in [a,b]\times[0,2\pi]\big)$$

függvény is. Ekkor

$$\partial_1\Phi(u,v) = (1, h'(u)\cdot \cos v, h'(u)\cdot \sin v),$$

$$\partial_2\Phi(u,v) = (0, -h(u)\cdot \sin v, h(u)\cdot \cos v) \qquad \big((u,v)\in [a,b]\times[0,2\pi]\big),$$

amiből

$$\partial_1\Phi(u,v)\otimes\partial_2\Phi(u,v) =$$

$$h(u)\cdot (h'(u), -\cos v, -\sin v) \qquad \big((u,v)\in [a,b]\times[0,2\pi]\big),$$

[94]Ha pl. $0<r<R$ és $h(x) := R+\sqrt{r^2-x^2}$ $(x\in[-r,r])$, akkor az $\mathcal{F}$ egy tórusz „külső" felülete. Továbbá (némi számolás után) $\alpha := \int_{-r}^r h(u)\cdot\sqrt{1+(h'(u))^2}\,du = 2r^2+\pi rR$. Hasonlóan, ha most $h(x) := R-\sqrt{r^2-x^2}$ $(x\in[-r,r])$, akkor az $\mathcal{F}$ a tórusz „belső" felülete, és $\beta := \int_{-r}^r h(u)\cdot\sqrt{1+(h'(u))^2}\,du = \pi rR-2r^2$. Következésképpen a szóban forgó tórusz felszíne $2\pi(\alpha+\beta) = 4\pi^2 rR$.

és így

$$\|\partial_1\Phi(u,v)\otimes\partial_2\Phi(u,v)\| = h(u)\cdot\sqrt{(h'(u))^2+\cos^2 v+\sin^2 v} =$$

$$= h(u)\cdot\sqrt{1+(h'(u))^2} \qquad \Big((u,v)\in[a,b]\times[0,2\pi]\Big)$$

következik. Tehát

$$\Upsilon_\Phi = \int_{[a,b]\times[0,2\pi]} \|\partial_1\Phi(u,v)\otimes\partial_2\Phi(u,v)\|\,du\,dv =$$

$$\int_{[a,b]\times[0,2\pi]} h(u)\cdot\sqrt{1+(h'(u))^2}\,du\,dv = \int_a^b h(u)\cdot\sqrt{1+(h'(u))^2}\left(\int_0^{2\pi} dv\right)du =$$

$$2\pi\cdot\int_a^b h(u)\cdot\sqrt{1+(h'(u))^2}\,du.$$

Itt a h függvény grafikonjának (függvénygörbének) az ívhossza a már korábban tanult formula szerint

$$\ell := \int_a^b \sqrt{1+(h'(u))^2}\,du,^{95}$$

az ℓ súlypontjának az ordinátája pedig (a b) pontban mondottak analógiájára)

$$y := \frac{1}{\ell}\int_a^b h(u)\cdot\sqrt{1+(h'(u))^2}\,du.$$

Más szóval $\Upsilon_\Phi = 2\pi y\cdot\ell,$ azaz az illető forgásfelület felszíne nem más, mint az azt meghatározó függvény(görbe) ívhosszának és a súlypontja által a forgatás során leírt kör kerületének a szorzata. Ez az állítás szintén *Guldin-tételként* ismert.[96]

g) **Felületi integrál.** *Képzeljük el, hogy áramló folyadék „útjában" elhelyezünk egy felületdarabot. Arra vagyunk kiváncsiak, hogy a szóban forgó felületdarabon (mondjuk időegység alatt) mennyi folyadék áramlik át?* (Röviden: mennyi az áramló folyadéknak az illető felületdarabra vonatkozó *fluxusa*?)

[95] Általában, ha valamilyen $n>1$ esetén adott egy $\Gamma\subset\mathbf{R}^n$ görbe a (kompakt $[a,b]$ intervallumon értelmezett) folytonosan differenciálható $g:[a,b]\to\Gamma$ paraméterezéssel, akkor a Γ ívhossza az $\int_a^b\|g'\|_2$ integrállal egyenlő. Speciálisan, ha az $n=2$ esetben $g(x)=(x,h(x))$ $(x\in[a,b]$, akkor $\|g'\|_2=\sqrt{1+(h'(u))^2}$. (Emlékeztetünk arra, hogy egy görbét ***rektifikálhatónak*** nevezünk (azaz, hogy van ívhossza), ha a görbébe írt poligonok hosszai által alkotott P halmaz felülről korlátos. Ez utóbbi esetben $\sup P$ a görbe *ívhossza.*)

[96] Vegyük észre, hogy a fentiekben a tórusz felszínére kapott $4\pi^2 rR=(2\pi r)(2\pi R)$ képletben $2\pi r$ a megforgatott körvonal kerülete (ívhossza), $2\pi R$ pedig annak a körnek a kerülete, amit a megforgatott körvonal középpontja (súlypontja) ír le.

A most mondott (kissé „szabadon" megfogalmazott) feladat matematikai modellezéséhez induljunk ki abból, hogy az itt említett „folyadékáramlás" egy

$$f \in \mathbf{R}^3 \to \mathbf{R}^3$$

(sebesség)függvénnyel adható meg a következő értelemben: ha $x \in \mathcal{D}_f$, akkor az $f(x) \in \mathbf{R}^3$ vektor a folyadék x pontbeli sebességét jelenti. A *felület* matematikai leírásához a

$$\Phi \in \mathbf{R}^2 \to \mathbf{R}^3$$

Gauss-paraméterezést használhatjuk (ld. f)), ahol a $\mathcal{D}_\Phi$ értelmezési tartomány egy kompakt intervallum (téglalap):

$$\mathcal{D}_\Phi = [a, b] \times [c, d].$$

Feltesszük, hogy az f függvény folytonos, a Φ pedig folytonosan differenciálható, és teljesül az alábbi feltétel:

$$n_\Phi(t) := \partial_1\Phi(t) \otimes \partial_2\Phi(t) \neq (0, 0, 0) \qquad (t \in [a, b] \times [c, d]).^{97}$$

Az itt szereplő

$$n_\Phi(t) \in \mathbf{R}^3$$

vektor az adott felület $\Phi(t)$-beli *normálvektora,* a

$$\partial_1\Phi \otimes \partial_2\Phi := n_\Phi : [a, b] \times [c, d] \to \mathbf{R}^3$$

függvény pedig a *normálvektorfüggvénye.*

Vegyük a fenti $[a, b]$ intervallumnak egy

$$x_0 = a < x_1 < ... < x_m = b \qquad (0 < m \in \mathbf{N})$$

felosztását és legyen

$$\mathcal{F}(x_0, ..., x_m) := \{[x_j, x_{j+1}] : j = 0, ..., m - 1\}.$$

Hasonlóan, ha

$$y_0 = c < y_1 < ... < y_p = d \qquad (0 < p \in \mathbf{N})$$

egy felosztása a $[c, d]$-nek, akkor legyen

$$\mathcal{F}(y_0, ..., y_p) := \{[y_i, y_{i+1}] : i = 0, ..., p - 1\}$$

[97] Az $\mathbf{R}^3$-beli A, B vektorok vektoriális szorzatát továbbra is az $A \otimes B$ szimbólummal jelölve.

és tekintsük az $[a,b]\times[c,d]$ téglalapnak a

$$J\times K \qquad (J\in\mathcal{F}(x_0,...,x_m),\ K\in\mathcal{F}(y_0,...,y_p))$$

„osztásintervallumait". Ha itt

$$J=[u,v],\ K=[w,z]$$

és

$$A:=\Phi(u,w),B:=\Phi(v,w),C:=\Phi(v,z),D:=\Phi(u,z),$$

akkor az A,B,C,D „sarokpontú"

$$\Phi_{ABCD}:=\{\Phi(x,y)\in\mathbf{R}^3:(x,y)\in J\times K\}$$

felületdarabot helyettesítsük a

$$B-A,D-A$$

vektorok által meghatározott (A csúcspontú)

$$P_{ABCD}$$

paralelogrammával. Ez utóbbinak a területe:

$$|P_{ABCD}|=|(B-A)\otimes(D-A)|=$$

$$|(\Phi(v,w)-\Phi(u,w))\otimes(\Phi(u,z)-\Phi(u,w))|,$$

ahol a Lagrange-középértéktétel szerint

$$\Phi(v,w)-\Phi(u,w)=\partial_1\Phi(\xi,w)\cdot(v-u)\approx\partial_1\Phi(u,w)\cdot(v-u)$$

és

$$\Phi(u,z)-\Phi(u,w)=\partial_2\Phi(u,\eta)\cdot(z-w)\approx\partial_2\Phi(u,w)\cdot(z-w)$$

alkalmas

$$\xi\in(u,v),\ \eta\in(w,z)$$

választással. Ezért

$$|P_{ABCD}|\approx|n_\Phi(u,w)|\cdot|J\times K|.$$

Így a Φ_{ABCD} felületdarabon időegység alatt átáramló folyadékmennyiség közelítőleg megegyezik annak a paralelepipedonnak a térfogatával, aminek az alapja a P_{ABCD} paralelogramma, a magassága pedig az $f(\Phi(u,w))$ vektor $n_\Phi(u,w)$ irányú vetületének a hossza:

$$\frac{|\langle f(\Phi(u,w), n_\Phi(u,w)\rangle|}{|n_\Phi(u,w)|}.$$

Más szóval a Φ_{ABCD}-n időegység alatt átáramló előjeles folyadékmennyiség megközelítőleg:

$$\langle f(\Phi(u,w), n_\Phi(u,w)\rangle \cdot |J \times K|.$$

Ha mindezt a szóba jövő összes $J \times K$ osztásintervallumra összegezzük, akkor a

$$\sum_{J \times K} \langle f(\Phi(u,w), n_\Phi(u,w)\rangle \cdot |J \times K|$$

összeg a keresett folyadékmennyiség egy közelítése. Ez nem más, mint az

$$\langle f \circ \Phi, n_\Phi \rangle : [a,b] \times [c,d] \to \mathbf{R}$$

függvénynek egy (Riemann-)integrál közelítő összege.

Összefoglalva az eddigieket azt mondjuk, hogy a keresett fluxus az

$$\int_\Phi f := \int_a^b \int_c^d \langle f \circ \Phi, n_\Phi \rangle$$

kettős integrál, az f függvénynek a Φ felületre vonatkozó *felületi integrálja.*