

**Dancs István–Magyarkúti Gyula–Medvegyev Péter–Puskás Csaba–Tallos Péter**

# **Bevezetés a matematikai analízisbe**

Budapesti Közgazdaságtudományi Egyetem.  
Budapest: Aula, 1996.



# Tartalomjegyzék

Tartalomjegyzék	i
<b>1 Halmazelmélet</b>	<b>1</b>
1.1 A matematika módszere és szaknyelve . . . . .	1
1.2 Bevezetés, alapfogalmak . . . . .	6
1.2.1 Bevezetés . . . . .	6
1.2.2 A halmazelmélet alapfogalmai . . . . .	7
1.3 Halmazok közötti műveletek . . . . .	10
1.3.1 Halmazok uniója, metszete és komplementere . . . . .	11
1.3.2 Halmazok szimmetrikus differenciája . . . . .	14
1.4 Halmazok szorzata, relációk . . . . .	15
1.4.1 Halmazok szorzata . . . . .	15
1.4.2 Relációk . . . . .	16
1.5 Függvények . . . . .	22
1.5.1 A függvény fogalma . . . . .	22
1.5.2 Függvények kompozíciója, inverze . . . . .	25
1.5.3 A direkt- és inverz-kép leképezés . . . . .	28
1.6 A számosságok . . . . .	31
<b>2 A számfogalom és valós függvények</b>	<b>37</b>
2.1 A számfogalom felépítése . . . . .	37
2.1.1 Természetes, egész és racionális számok . . . . .	37
2.1.2 Valós számok . . . . .	38
2.1.3 Nevezetes azonosságok és egyenlőtlenségek . . . . .	43
2.1.4 Komplex számok . . . . .	46
2.2 Valós vektorok, az $\mathbb{R}^p$ tér . . . . .	64
2.3 Algebrai struktúrák . . . . .	69
2.4 Valós függvények . . . . .	76
2.4.1 Valós függvények bevezetése . . . . .	77
2.4.2 Polinomok és racionális törtfüggvények . . . . .	83
2.4.3 A hatvány, exponenciális és logaritmus függvény. . . . .	85
2.4.4 Trigonometrikus függvények és inverzeik . . . . .	89

<b>3</b>	<b>Metrikus terek és leképezéseik</b>	<b>95</b>
3.1	Metrikus terek . . . . .	95
3.1.1	Példák metrikus terekre . . . . .	95
3.1.2	Gömb-környezetek metrikus térben . . . . .	97
3.1.3	Nyílt és zárt halmazok metrikus térben . . . . .	103
3.2	Folytonos függvények metrikus téren . . . . .	107
3.2.1	Definíciók . . . . .	107
3.2.2	Az alpműveletek folytonossága . . . . .	111
3.2.3	Formális szabályok, elemi függvények . . . . .	114
3.2.4	Folytonos függvények alaptulajdonságai . . . . .	117
3.3	Határérték . . . . .	120
3.3.1	Véges határérték végesben . . . . .	120
3.3.2	A végtelen szerepe a határértékeknel . . . . .	127
<b>4</b>	<b>Sorozatok és sorok</b>	<b>135</b>
4.1	Sorozatok metrikus terekben . . . . .	135
4.2	Számsorozatok . . . . .	138
4.2.1	A Bolzano-Weierstrass tétel . . . . .	138
4.2.2	Valós Cauchy-sorozatok . . . . .	139
4.2.3	Limesz superior és limesz inferior . . . . .	140
4.2.4	Határérték és műveletek . . . . .	142
4.2.5	Sorozatok végtelen határértéke . . . . .	142
4.2.6	$\mathbb{R}^2$ -beli sorozatok . . . . .	143
4.3	Numerikus sorok . . . . .	147
4.3.1	Alapfogalmak . . . . .	147
4.3.2	Konvergencia kritériumok . . . . .	152
4.4	Sorozatok és függvények határértéke . . . . .	156
<b>5</b>	<b>Differenciálszámítás</b>	<b>161</b>
5.1	Definíciók, értelmezések . . . . .	161
5.1.1	Bevezetés, definíciók . . . . .	161
5.1.2	Érintő és érintőapproximáció . . . . .	174
5.1.3	Sebesség . . . . .	179
5.1.4	Relatív sebesség, elaszticitás . . . . .	181
5.2	Differenciálás, kalkulus . . . . .	183
5.2.1	Formális szabályok . . . . .	183
5.2.2	Speciális függvények differenciálása . . . . .	188
5.2.3	A szélsőérték szükséges feltételei . . . . .	190
5.3	Középértéktételek . . . . .	192
5.3.1	A differenciálszámítás középértéktétele . . . . .	192
5.3.2	Magasabbrendű approximációk, Taylor-formula . . . . .	194
5.3.3	Általánosított középértéktétel, L'Hospital-szabály . . . . .	199

<b>6</b>	<b>Monoton és konvex függvények</b>	<b>207</b>
6.1	Alaptulajdonságok . . . . .	207
6.1.1	Monoton függvények . . . . .	207
6.1.2	Konvex és konkáv függvények . . . . .	210
6.2	Differenciálható függvények vizsgálata . . . . .	217
6.2.1	A monotonitásra vonatkozó feltételek . . . . .	217
6.2.2	A szélsőérték elégséges feltételei . . . . .	220
6.2.3	Differenciálható konvex függvények . . . . .	221
6.2.4	Az Euler-szám és az exponenciális függvény . . . . .	226
6.2.5	Konvexitási egyenlőtlenségek . . . . .	233
6.2.6	Az elaszticitás és a logaritmikus derivált . . . . .	235
6.2.7	Kalkulus-összefoglaló . . . . .	237
6.2.8	Függvények diszkussziója . . . . .	238
<b>7</b>	<b>Antiderivált és differenciálegyenletek</b>	<b>243</b>
7.1	Az antiderivált, határozatlan integrál . . . . .	243
7.1.1	Definíciók, elemi tulajdonságok . . . . .	243
7.1.2	Parciális integrálás . . . . .	251
7.1.3	Helyettesítéssel való integrálás . . . . .	255
7.2	Differenciálegyenletek . . . . .	259
<b>8</b>	<b>Határozott integrál</b>	<b>267</b>
8.1	A Riemann integrál definíciója . . . . .	267
8.2	Integrálhatósági tételek . . . . .	275
8.3	A határozott integrál és a differenciálás . . . . .	281
8.4	Integrálszámítási szabályok, példák . . . . .	285
8.5	Improprius integrálok . . . . .	287
8.6	Hatványsorok . . . . .	296
8.6.1	Alapvető tulajdonságok . . . . .	296
8.6.2	Példák . . . . .	299
	<b>Tárgymutató</b>	<b>301</b>



# 1.

## Halmazelmélet

### 1.1 A matematika módszere és szaknyelve

Ebben a pontban a matematika tárgyalási módszeréről és a matematikai szaknyelvről fogunk röviden beszélni.

#### A matematika tárgyalási módszere

A matematika deduktív tudomány. Ez a megnevezés már meg is mutatja tárgyalásának a módszerét. A “deduktív” (következtető, levezető) szó azt mondja, hogy a matematikában igaz állításokból igaz állításokat vezetünk le, a logika felhasználásával. Emiatt a logikai következtetési módok helyes használata nyilvánvalóan elengedhetetlen.

A “logikus” gondolkodásnak, szerencsére, a legtöbb ember birtokában van anélkül, hogy a szabályokat tudatosította volna. Így nem szükséges feltétlenül, hogy logikai tanulmányokkal kezdjük a tárgyalásunkat. Ez azonban nem jelenti azt, hogy ne lenne komoly jelentősége egy olyan tárgynak, ami a helyes gondolkodás szabályait tudatosítja. A gondolkodás tiszta voltához nagyban hozzájárulhat az, ha megismerjük, hogy mit is jelent definiálni, tételt állítani, következtetni, mik ennek a logikai szabályai, stb. Mi nem adhatunk itt ilyen irányban részletes bevezetést, csak néhány fontos kérdésre hívjuk fel a figyelmet.

Az állítások lényegében véve kijelentő mondatokból állnak (amelyek vagy igazak vagy nem). A kiemelt matematikai állításokra a “tétel”, “állítás”, “lemma” (segéd-tétel) megjelölések használatosak. Mi általában az “állítás” szót fogjuk használni, és csak a kiemelten jelentős állításokat nevezzük tételnek, de nem leszünk makacsul következtetések, és a leghelyesebb, ha az említett szavakat azonos jelentésűeknek vesszük.

A közönséges nyelvben úgy ismerkedünk meg egy fogalommal, hogy gyakorlás útján megtanuljuk a fogalmat jelző névnek a fogalom körébe tartozó objektumokra

való alkalmazását. A tárgyak hasonló tulajdonságainak a tapasztalása során kialakul bennünk — többé-kevésbé világosan — egy fogalom tartalma. Az így kialakult fogalmak a hétköznapi eligazodás számára megfelelőek, de tudományos célra használhatatlanok, hiszen ki tudná például pontosan megmondani, hogy mi is az “asztal”.

A matematikai fogalom-alkotás módja a *definíció* (meghatározás), amikor már meglévő fogalmakkal új fogalmat adunk meg. A definíciónál nem az igazság a megkövetelés, hanem a korrektség, amin azt értjük, hogy pontosan mondja meg azt, hogy mit értünk a definiált fogalmon. Nem szabad például még nem definiált fogalmat felhasználni a meghatározáshoz. Az átlagos logikai készségünk alapján általában el tudjuk dönteni egy definíció korrektségét, de némelykor kifejezetten indokolnunk kell, hogy helyes a meghatározás.

A definíció fogalma már egyértelműen elvezet a deduktív tárgyalás legalapvetőbb problémájához. Mivel minden definiálás feltételez már ismert fogalmakat, ezért felvethető a kérdés, hogy mi van a “kezdő”, “kiinduló” fogalmakkal, amelyek definiálhatatlanok, mivel nem vezethetők vissza már meghatározott fogalmakra.

Gondoljunk például a geometria alapfogalmaira. Az alapvető objektumokat: “sík”, “egyenes”, “pont”; és a kiinduló relációkat: “illeszkedés” (pont rajta van az egyenesen), “metszés” (két egyenesnek van közös pontja), stb., nem tudjuk definiálni. Megfogalmazunk viszont az alapfogalmak és relációk neveivel állításokat, amiket igaznak fogadunk el. Ezeket az állításokat axiómáknak nevezzük. Például két ilyen állítás:

*Két különböző pont pontosan egy egyenesre illeszkedik.*

Az axiomatikus kiinduló állítások voltaképpen tetszőlegesek, de a megfogalmazásukban azért alapvető szerepe van a tapasztalásnak. Az egyenes geometriai fogalma a természetben tapasztalt, vagy általunk rajzolt “egyenes darabok” elvonatkoztatásából ered.

Az axiómák *nem bizonyított* állítások, *nem definiált* fogalmakkal, ezért jogosan megkérdozhető, hogy mi értelme van ilyeneket kimondani. Mondhatnak ezek egyáltalán valamit? Erre, meglepő módon könnyű válaszolni. Feltétlenül adnak valamit, mert ezek után már nem mondhatunk bármit a fogalmakra, például a geometriai esetben nem állíthatjuk azt, hogy két különböző pontra két különböző egyenes illeszkedik.

Az a helyzet, hogy ha a nem meghatározott nevekkkel kimondott állítások egy rendszerét igaznak követeljük meg, akkor az már egy határozott megkötést — mondhatnánk: valamilyen szokatlan formaájú definíciót — ad a szóbanforgó szavaknak. Az axiómarendszer megadása után a kiinduló nevekkkel új fogalmakat tudunk meghatározni, meg tudunk fogalmazni állításokat, amelyekről megkérdozhetjük, hogy igazak-e vagy hamisak-e. Így pontosan felépíthető egy deduktív tudományos terület, például a geometria.

Az axiomatikus módszer felfedezése a tudomány történetének talán a legfontosabb felfedezése. Az alapvető tudományos felfedezések a történelem során

a különböző kultúrákban rendszerint párhuzamosan keletkeztek vagy legalábbis vetődtek fel. Érdekes, hogy az axiomatikus módszer születése egyetlen kultúrkörhöz kapcsolható. Szinte csaknem teljesen “kész” állapotban jelenik meg Krisztus előtt 300 körül a klasszikus görög tudományban. A már említett logika tudományának a felfedezése sem előzi meg. Tudomány-történetek szerint a logika az axiomatikus módszer kielemezésével született és a tárgyalás is az axiomatikus módszert követi.

Meg kell fontolnunk, hogy az axióma milyen viszonyban van a valóság megismerésével. Azt hihetnénk, hogy az axiómák a “valóságra” vonatkozó legáltalánosabb, legelemibb igazságok. Az nem tagadható, hogy a valóság szemlélete szerepet játszik a megfogalmazásukban, de nem helyes, ha valamilyen kiinduló abszolút igazságoknak tartjuk őket.

Gondoljunk csak a geometriára, amiben először vetődött fel az, hogy egyes axiómákat kétségbe vonjanak — például az előzőekben felírt párhuzamossági axiómát — és különféle, egymásnak ellentmondó geometriát építsenek fel. Az a kérdés, hogy melyik geometria a valóságos tér geometriája, nem matematikai kérdésfeltevés.

Egy matematikai elmélet korrekt, ha ellentmondástól mentes axiómarendszer a kiinduló pontja. Az “igazság” a matematika (és más axiomatikus elmélet) állításainak a tulajdonsága, nem a valóságra vonatkozó állításoké. A tudomány-elméletben a valóságra vonatkozó állításokra az “érvényesség” (alkalmazhatóság) kategóriáját használják az “igazság” kategóriája helyett. A matematika nem kíván a valóság egyetlen szeletéről sem érvényes állításokat mondani, csak ellentmondástól mentes elméleteket épít fel, amelyek a szaktudományok számára egzakt nyelveket adnak.

Az axiomatikus módszer nem a matematika privilégiuma. A tizenhetedik században széles körben megindult a klasszikus görög tudomány újra való felfedezése, és ekkor az axiomatikus módszer a tudományos tárgyalás egyetlen mintája lett. Filozófusok próbáltak axiomatikus filozófiai elméleteket felépíteni, és a fizika nagy elindulása már ilyen elvek szerint történt.

Ma már sok tudomány, ilyen vagy olyan fokban, axiomatikus tárgyalási módot követ. Ez többé-kevésbé megegyezik a matematika felhasználásának a fokával. Egy tudomány egzaktsága az axiomatikus szemlélettel azonosítható. Az egyes tudományok egzaktsági foka különböző, és ami számunkra lényeges megállapítás: a közgazdaságtan már axiomatikus tudománynak tekinthető.

Maga az axiomatikus módszer önmagában is vizsgálatra szorul. Ezzel a megfelelő matematika alapjait vizsgáló diszciplína foglalkozik.

Az axiomatikus módszer nagyon pontos tárgyalást tesz lehetővé, de ez nem teszi feleslegessé, sőt abszolút igényli a “szemléletes” látásmód felhasználását. Enélkül olyan száraz és megjegyezhetetlen lenne a tárgyalás, hogy tanulhatatlan lenne. A kutató matematikus munkájában is döntő szerepe van a “szemléletes fantáziálásnak”. Ezt a praktikus elvet mi is igyekszünk hasznosítani.

Fontos dolog még, hogy — noha a matematika önmagukban zárt elméleteket épít fel — nagyon sokat segít a fogalmak és tételek elsajátításában az, ha a szak-

tudományos alkalmazásokban való szerepüket megismerjük. Erre hangsúlyozott figyelmet fordítunk, elsősorban a közgazdaságtanra koncentrálván.

### A matematikai szaknyelv

Minden tudomány szakmai nyelve különbözik valamilyen mértékben a megszokott hétköznapi nyelvtől. Nem a speciális jelentésű szakmai szavakra gondolunk itt, ami természetes, hanem arra, hogy a mondatok szerkesztése is eltérő, illetve bizonyos speciális mondat szerkezetek gyakorta szerepelnek. Ezeket folyamatosan elsajátítjuk, hiszen jórészüket már a középiskolában megtanultuk.

Az egyes jelölésekről a megfelelő helyeken szólunk. Az általános elvünk az lesz, hogy nem ragaszkodunk mereven egyetlen jelöléshez még akkor sem, ha a legjobbnak tartjuk. A matematikai jelölésekben, mint minden tudomány szaknyelvében, nagy szerepe van a hagyománynak, ami nem mindig eléggé helyes, de úgy célszerű használni a jelöléseket, ahogyan az általában, a jelenlegi történelmi időpontban általános szokás.

A matematika nyelve megalkotható lenne úgy is, hogy szigorúan egyértelmű legyen és ne használjon hétköznapi fordulatokat. Ez a módszer azonban teljesen alkalmatlan lenne mind az oktatáshoz mind a kutatáshoz. Ennek elsősorban az az oka, hogy a matematika sem nélkülözheti a fogalmak és tételek körüli „magyarázkodást”, az olyan értelmezéseket, amelyek szemléletileg közelebb hozzák az olvasóhoz a dolgokat. Egy sikeres mondat sokszor nagyon sokat segíthet egy fogalom vagy állítás megértésében. Egy tankönyv (és minden más) írása küzdelem a helyes és célszerű nyelvi kifejezésekért.

Van azonban egy problémakör, ami nagyon szigorúan pontos leírását kívánja meg a matematikának: a matematika alapjainak a kérdéseit kutató *matematikai logika*. Ezzel most nem kívánunk foglalkozni, de el kell mondanunk, hogy bizonyos logikai jelölések és leírási formák — kizárólag csak a tömörebb írásmód kedvéért — már bekerültek a közönséges matematikai szaknyelvbe is.

Jelöljenek a továbbiakban az  $A$  és  $B$  betűk állításokat, amelyek vagy igazak vagy nem igazak (hamisak). Az állításokból újabb állításokat lehet készíteni a logikai műveletekkel. Négy műveletet fogunk használni némelykor a leírásokban:

**AZ “és” MŰVELET.** Az “ $A$  és  $B$ ” állítás pontosan akkor igaz, ha mind az  $A$  mind a  $B$  állítások igazak. Mi a köznyelvi kapcsolatnak megfelelő “és” műveleti jelet részesítjük előnyben, de szokás például az “ $\vee$ ” műveleti jel is.

**AZ “vagy” MŰVELETI JEL.** Az “ $A$  vagy  $B$ ” állítás pontosan akkor igaz, ha az  $A$  és  $B$  közül legalább az egyik igaz. Itt is van eltérő jelölés, az “ $\wedge$ ” műveleti jel. Szigorúan különböztessük meg ezt a “vagy” műveletet a “kizárólagos vagy”-tól: akkor igaz, ha pontosan az egyik igaz az  $A$  és  $B$  állítás közül.

**AZ  $\Rightarrow$  MŰVELET.** Az “ $A \Rightarrow B$ ” állítás pontosan akkor igaz, ha az  $A$  állításból következik a  $B$  állítás.

AZ  $\Leftrightarrow$  MŰVELET. Az “ $A \Leftrightarrow B$ ” állítás akkor igaz, ha az  $A$  és  $B$  állítások ekvivalensek, azaz egyik következik a másikból.

Két logikai szimbólumot használunk még a “minden” jelentésű “ $\forall$ ” és a “létezik” jelentésű “ $\exists$ ” szimbólumokat. Ezeket mindig egy állítás elején szerepeltetjük, és az állítások változóira vonatkoznak, például a halmazelméletből véve egy példát a

$$\forall x \in A \quad \exists y \in B \quad x = y$$

állítás akkor igaz, ha  $A \subseteq B$ . Fontos megjegyeznünk, hogy egy ilyen állítás tagadása — amire a “ $\neg$ ” jel használatos — úgy történik, hogy a minden és létezik jelek megfordulnak, és tagadjuk a mögöttük álló állítást:

$$(\forall x \in A \quad \exists y \in B \quad x = y) = \exists x \in A \quad \forall y \in B \quad (x \neq y).$$

Végezetül — mivel a matematika gyakran használja a görög betűket — ezért leírjuk a legfontosabbakat:

#### KIS GÖRÖG BETŰK

$\alpha$	alfa	$\beta$	béta	$\gamma$	gamma
$\delta$	delta	$\epsilon \quad \varepsilon$	epszilon	$\zeta$	dzeta
$\eta$	éta	$\theta \quad \vartheta$	théta	$\kappa$	kappa
$\lambda$	lambda	$\mu$	mű	$\nu$	nű
$\xi$	kszi	$\pi$	pi	$\rho$	ró
$\sigma$	szigma	$\tau$	tau	$\phi \quad \varphi$	fi
$\chi$	khi	$\psi$	pszi	$\omega$	omega

A görög nagybetűk többsége erősen hasonlít a latin nagybetűkre, ezért azok közül kevés használatos:

$\Gamma$	gamma	$\Delta$	delta	$\Theta$	theta
$\Lambda$	lambda	$\Xi$	kszi	$\Pi$	pi
$\Sigma$	szigma	$\Phi$	fi	$\Psi$	pszi

Régebben a nagy gót betűk is használatosak voltak, de ma már csak a matematika speciális fejezeteiben lehet velük találkozni. A halmazelméletben használatos még a megszámlálhatóan végtelen számosság jelölésére a héber alef betű nullával indexelve:  $\aleph_0$ .

## 1.2 Bevezetés, alapfogalmak

### 1.2.1 Bevezetés

Ami a halmazok hétköznapi használatát illeti, látszólag nagyon könnyen tudunk példát mondani halmazokra: ezen teremben lévő hallgatók összessége; az európai országok halmaza, stb. . Ezek a példák azonban éppen úgy nem matematikai halmazok, mint ahogyan a vonalzó éle nem egy geometriai egyenes.

Megkísérelhetnénk definiálni is a halmaz fogalmát, ez azonban nem lehetséges, mivel a matematikai *halmaz* fogalma a matematika tudományának a “legelső” alapfogalma, éppen ezért más matematikai fogalmakra nem vezethető vissza, és most az elején, még matematikai példa sem mondható rá. Az axiomatikus tárgyalás számára azonban a mondottak nem jelentenek gondot. A halmazelmélet axiómái pontosan olyan állítások, amelyek megmondják, hogy miként lehet új halmazt alkotni már meglévő halmazokból, és így például definiálható a természetes számok halmaza, valós számok halmaza, és minden egyéb olyan objektum, ami matematikai vizsgálat tárgya.

Nagyon fáradságos — és egy nem specifikusan az axiomatikus halmazelmélet iránt érdeklődő számára nem sokat nyújtó — munka lenne az axiomatikus halmazelmélet felépítése, amire nem is vállalkozunk. Ahogyan az előszóban is mondtuk, “naív” halmazelméletet fogunk adni, de törekszünk a lehetséges pontosságra. Kezdjük először is azzal, hogy egy halmazt közönséges módon a következőképpen adhatunk meg:

Jelöljön a  $T$  egy tulajdonságot, és a megfelelő  $T(x)$  állító mondat legyen igaz, ha az  $x$  rendelkezik a  $T$  tulajdonsággal, és hamis, ha nem. Ekkor a  $T$  tulajdonsággal rendelkező dolgok összességét az

$$\{x : T(x)\} \quad (1.1)$$

módon jelöljük. Olvasva:

A  $T$  tulajdonsággal rendelkező dolgok összessége.

vagy:

Azon  $x$  elemek összessége, amelyek rendelkeznek a  $T$  tulajdonsággal.

Ezekután azt, hogy egy  $y$  objektum rendelkezik a  $T$  tulajdonsággal így is megmondhatjuk: benne van az  $\{x : T(x)\}$  összességben, amit formálisan

$$y \in \{x : T(x)\}$$

módon írunk, és így olvassuk: az  $y$  eleme a  $\{x : T(x)\}$  összességnek.

A  $T$  tulajdonság helyett azt is mondhatnánk, hogy van egy  $T(x)$  kijelentő mondat, ami bizonyos  $x$  alanyokra igaz, bizonyosakra pedig nem, és a  $\{x : T(x)\}$  halmaz azon  $x$  elemekből áll, amelyekre a  $T(x)$  mondat igaz. Ezzel elkerülhetnénk a

“tulajdonság” fogalom használatát. Észre kell venni, hogy a leírt halmaz-megadási mód nagyvonalú, hiszen sem a “tulajdonság”, sem a “mondat” nem pontosan értelmezett fogalmak. Ennek ellenére az előző naív gondolatmenetben szerepel a halmazelmélet minden alapfogalma. Ezekután már csak pontosan ki kellene mondani a halmazelmélet axiómáit, és azután deduktív módon felépíteni a matematika épületét. A következő alpontban pontosítjuk az alapfogalmakat és a kiindulást, de az axiomatikus tárgyalást nem erőltetjük.

### 1.2.2 A halmazelmélet alapfogalmai

A halmazelméletnek csak két alapfogalma van, amiket egy definícióban rögzítünk. Ez azonban nem lesz igazi definíció, hiszen olyan fogalmakról van szó, amelyek kiindulóak, ezért nem vezethetők vissza más fogalmakra, ahogyan egy definíciótól elvárnánk. Tekintsük ezért ezt csupán a kiinduló elnevezések felsorolásának.

**Definíció 1 (A halmazelmélet kiinduló fogalmai)** *A halmazelmélet alapvető objektuma a halmaz, amire ezen kívül nagyon sokféle azonos értelmű szó használatos: elem, összesség, pont, tér, rendszer, stb. .*

*A másik alapfogalom egy, halmazok közötti reláció, az elemének lenni, amit az “ $\in$ ” jellel jelölünk. A “ $\in$ ” reláció tagadásának a jelölésére az “ $\notin$ ” használatos.*

A halmazra azért szükséges a változatos szóhasználat, mert olyan sokszor kerül elő valamilyen formában a fogalom (voltaképpen minden matematikai fogalom halmaz), hogy egyetlen elnevezéshez való ragaszkodás rendkívüli mértékben áttekinthetatlenné tenné a stílust.

Talán meglepetést okoz az, hogy az “elemet” is a “halmazzal” azonos jelentésűnek vesszük. Ez azonban valóban így helyes, mert a halmazelméletben csak egyetlen objektum van: a halmaz. Az a hétköznapi ismeretben általános szokás, hogy az elemet és halmazt különböző objektumnak fogják fel, abból ered, hogy az elemének lenni reláció nem szimmetrikus azaz az  $x \in a$  és  $a \in x$  relációk nem ekvivalensek. Emiatt az “elemének lenni” reláció baloldalán lévő halmazt elemnek szokás mondani. A “pont” szinonima is akkor használatos, ha az “elemének lenni” reláció baloldalán szerepel.

Hasznos jelölési szokás, hogy az “ $\in$ ” reláció baloldalára “kisebb rangú” betűt írunk:  $a \in X$ ,  $D \in \mathcal{H}$ . Ez jól mutatja azt a hierarchiát, ami a halmazok és elemeik között van.

A halmazokat néha az u. n. *Venn-diagramokon* szokás szemléltetni, ahogyan azt a későbbiekben néhányszor mi is tenni fogjuk. Ez nagyon hasznos lehet, de olyan vizuális, hogy elfedi a sokkal absztraktabb hátteret, ezért bizonyításra semmi esetre se használjuk, csak illusztrálásra.

Most pedig, habár nem kívánunk axiomatikusan pontosak lenni, de az első két axiómát — definícióban rögzítve — pontosan kimondjuk:

**Definíció 2 (Azonossági axióma)** *Két halmaz pontosan akkor egyenlő (azonos), ha az elemeik megegyeznek.*

A definíció röviden szólva azt mondja, hogy *egy halmazt az elemei pontosan meghatározzák*. Formálisan is megfogalmazva az egyenlőség feltételét:

$$A = B \iff [(x \in A \Rightarrow x \in B) \text{ és } (x \in B \Rightarrow x \in A)].$$

**Definíció 3 (Halmazmegadási axióma)** *Egy  $X$  halmaz és  $T(x)$ ,  $(x \in X)$  tulajdonság esetén van olyan  $A$  halmaz, amelyhez pontosan azon elemei tartoznak az  $X$ -nek, amelyek kielégítik a  $T(x)$  tulajdonságot. Az  $A$  halmaz formális jelölése:*

$$\{x \in X : T(x)\} \quad \text{vagy} \quad \{x \in X \mid T(x)\}.$$

Vegyük észre, hogy a bevezetésben mondott (1.1) kevésbé pontos halmaz-megadási módtól csak annyiban különbözik a definíció, hogy egy adott — de tetszőleges — halmaz elemeinek a  $T$  tulajdonságú elemeit vesszük. Ez a különbség azonban nagyon lényeges, mert kizárja azt, hogy ellentmondás keletkezzék, ami a bevezetés naív megadási módja mellett nem kerülhető el. Ezt a kiegészítésben részletezzük.

Egy halmazt esetleg felsorolással is megadhatunk, például:  $\{1, 2, 3\}$ . Belátható lenne, hogy ez visszavezethető a tulajdonsággal való megadási módra.

A  $T(x)$  tulajdonság egy  $x$ -et tartalmazó állítás, ami az  $x$  bizonyos értékeire igaz, bizonyosakra pedig hamis. Az állítást a logikai jelekkel és a halmazelmélet alapfogalmaival lehet megfogalmazni, és ha már vannak — végül is a halmazelmélet alapfogalmaiból felépített — matematikai foglmaink, akkor azok is szerepelhetnek. A matematikai foglmainak a halmazelmélet foglmainból való felépítését nem fogjuk elvégezni, de a kiegészítő részben majd szerepel a természetes számok bevezetésének a vázolása. További érdekes példát nyújt erre a számfoglmaink (egész, racionális, valós) felépítése, amit egy más anyag tartalmaz.

A mondottak ellenére a tárgyalás során olyan példákat is fogunk mondani, amelyek olyan foglmainkat használnak, amik csak a későbbiekben kerülnek bevezetésre (valós számok, stb.). Ezt azért célszerű tennünk, mert a halmazelmélet foglmai és tételei annyira elvontak, hogy feltétlenül szükséges konkrétabb példákkal segíteni az elsajátítását. Olyan tárgyalás, amely szigorúan lineáris menetben adja az ismereteket, csak elméletileg létezik még a matematikában is, és az oktatásban követhetetlen.

A definiált halmaz megadási módnak két fontos következménye van.

**Állítás 4 (Nincs univerzum)** *Tetszőleges  $A$  halmazhoz létezik olyan  $B$  halmaz, amelyik nem eleme az  $A$  halmaznak.*

Másként fogalmazva:

*Az az objektum, ami minden halmazt tartalmaz nem lehet halmaz.*

vagy még más módon:

*Nincs olyan halmaz, amelyik minden halmazt tartalmazna.*

Eleinte azt gondolták, hogy ilyen van, és *univerzumnak* nevezték el, ami aztán ellentmondáshoz is vezetett. A 3. definíció (axióma) kizárja az univerzum létezését, és megszüntet bizonyos ellentmondásokat.

**Bizonyítás.** A halmazmegadási axióma alapján vehetjük a

$$B \doteq \{x \in A : x \notin x\} \quad (1.2)$$

halmazt. Megmutatjuk, hogy a  $B$  halmaz nem eleme az  $A$  halmaznak, amivel az állítást nyilvánvalóan belátjuk.

Tegyük fel — állításunkkal ellentétben — hogy

$$B \in A, \quad (1.3)$$

és megmutatjuk, hogy ez ellentmondáshoz vezet. Nyilvánvaló, hogy a

$$B \in B, \quad B \notin B$$

állításoknak pontosan az egyike igaz, és így — ha az (1.3) alatti állítás igaz — a következő két állítás egyikének teljesülnie kellene:

$$1) \quad B \in A \text{ és } B \in B \qquad 2) \quad B \in A \text{ és } B \notin B.$$

1) A  $B \in B$  reláció a (1.2) definíció szerint azt jelenti, hogy  $B \in A$  és  $B \notin B$ , ami ellentétben van a  $B \in B$  állítással, tehát az 1) nem teljesülhet.

2) A 2) alatti relációk (1.2) definíció szerint pontosan azt jelentik, hogy  $B \in B$ , ami ellentétben van a 2)-vel, tehát a 2) sem teljesülhet.

Mivel sem az 1) sem a 2) nem igaz, ezért a (1.3) állítás nem állhat fenn.  $\square$

**Állítás 5 (Az üres halmaz létezése)** *Ha létezik halmaz, akkor létezik olyan halmaz, amelynek nincs eleme. Ezt üres halmaznak nevezzük, és  $\emptyset$ -vel jelöljük.*

Ne lepődjünk meg azon, hogy fel kellett tenni az állításban azt, hogy létezik halmaz, de látni fogjuk a bizonyításban, hogy ez szükséges. Egyébként a halmazelméletben halmaz létezését vagy külön fel kell tenni, vagy más axiómával kell biztosítani, például lehetne: Létezik az üres halmaz. Erre most nem térünk ki.

**Bizonyítás.** Jelölje  $A$  a feltevés szerint létező halmazt. A halmazmegadási axiómában vegyük a  $T(x) = (x \neq x)$  állítást. Így az  $\{x \in A : x \neq x\}$  objektum helyesen megadott halmaz, és nyilvánvalóan nincs eleme, hiszen minden  $x$ -re  $x = x$ .  $\square$

Ezzel befejeztük a halmazelmélet alapfogalmai körüli vizsgálatainkat, és a következőkben célratörően — kevésbé törődve az elvi alapokkal — úgy tárgyaljuk a halmazelméletet, ahogyan azt egy — a matematikai analízissel foglalkozó — matematikusnak, közgazdásznak tudnia kell.

### 1.3 Halmazok közötti műveletek

A bevezetésben szemléletesen behozott két alapfogalom — a *halmaz*, az *elemének lenni* — fogalmakból kiindulva, a *halmaz megadási módjának* a segítségével most további fogalmakat vezetünk be. Kezdjük a “Részének lenni” viszonyt, és a hozzá kapcsolódó elnevezéseknek a definiálásával:

**Definíció 6 (Tartalmazás, részhalmaz)** *Ha egy  $A$  halmaz minden eleme eleme a  $B$  halmaznak is, akkor azt mondjuk, hogy az  $A$  halmaz részhalmaza (része) a  $B$  halmaznak — jelölésben:  $A \subseteq B$  —. Formálisan:*

$$A \subseteq B \iff (x \in A \Rightarrow x \in B).$$

*Valamely  $X$  halmaz részhalmazainak az összességét — amit szintén halmaznak tekintünk — a továbbiakban  $\mathcal{P}(X)$ -szel fogjuk jelölni, de szokásos még a  $2^X$  jelölés is.*

Az  $A \subseteq B$  viszonyra ezt is szokás mondani: Az  $A$  *szűkebb* mint a  $B$ , illetve a  $B$  *bővebb*, mint az  $A$ .

Ha az  $A \subseteq B$  és  $A \neq B$ , azaz az  $A$  minden eleme eleme a  $B$ -nek és a  $B$ -nek van olyan eleme, ami nem eleme az  $A$ -nak, akkor azt mondjuk, hogy az  $A$  *valódi része* a  $B$  halmaznak. Használatos még ebben az esetben: Az  $A$  *szigorúan szűkebb* a  $B$ -nél, illetve a  $B$  *szigorúan bővebb* az  $A$ -nál. A szigorú tartalmazás jelölése:

$$A \subset B \text{ illetve } B \supset A.$$

Azonnal látható, hogy a “ $\subseteq$ ” tartalmazás viszony rendelkezik a következő állításban felsorolt tulajdonságokkal.

**Állítás 7 (Tartalmazás tulajdonságai)** *A halmaz tartalmazás rendelkezik a következő tulajdonságokkal.*

- (1)  $A \subseteq A$ , (reflexivitás),
- (2)  $(A \subseteq B \text{ és } B \subseteq A) \implies A = B$  (antiszimmetria),
- (3)  $(A \subseteq B \text{ és } B \subseteq C) \implies A \subseteq C$ , (transzitivitás).

A (2) tulajdonságot különösen sokszor használjuk, mert rendszerint ezen a módon látjuk be két halmaz egyenlőségét: megmutatjuk, hogy az egyik része a másiknak, és a másik része az egyiknek.

**Bizonyítás.** Bizonyításra alig szorulnak, de azért leírjuk a rövid indoklásokat:

Az (1) a halmazok egyenlőségének az axiómájából (2. definíció) nyilvánvalóan adódik, a (2) pedig megegyezik az említett definícióval.

A (3) igazolása: Ha  $x \in A$ , akkor az  $A \subseteq B$  definíciója szerint  $x \in B$ , és  $B \subseteq C$  miatt ugyanezen okból kapjuk:  $x \in C$ , és így ismét a tartalmazás definíciója szerint:  $A \subseteq C$ . □

### 1.3.1 Halmazok uniója, metszete és komplementere

**Definíció 8 (Halmazok uniója, metszete, kivonása)** Tetszőleges  $A, B$  halmazból képezzük az

$$A \cup B \stackrel{\text{def}}{=} \{x \mid x \in A \text{ vagy } x \in B\}, \quad (1.4)$$

$$A \cap B \stackrel{\text{def}}{=} \{x \mid x \in A \text{ és } x \in B\}, \quad (1.5)$$

$$A \setminus B \stackrel{\text{def}}{=} \{x \mid x \in A \text{ és } x \notin B\}. \quad (1.6)$$

halmazokat.

Az  $A \cup B$  halmazt az  $A$  és  $B$  halmazok uniójának (egyesítésének), az  $A \cap B$  halmazt pedig az  $A$  és  $B$  metszetének (közös részének) szokás nevezni. Az  $A \setminus B$  halmaz elnevezése: az  $A$  és  $B$  halmazok különbsége.

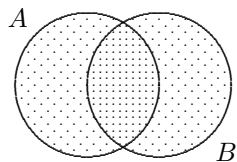
Ha az  $A \cap B$  halmaz üres, akkor az  $A$  és  $B$  halmazokat diszjunktak mondjuk. Halmazok egy rendszerét páronként diszjunktak nevezzük, ha bármely két benne lévő halmaz diszjunkt.

A műveletek a bevezetésben is említett Venn-diagramokon jól szemléltethetők: 1.1.–1.2. ábrák. Általában hasznos, ha a definíciókat szavakban is megfogalmazzuk, mert ez segíti a megértést és a memorizálást.

Ha egy rögzített  $X$  halmaz részhalmazait tekintjük, akkor az  $X$  és  $A \subseteq X$  halmazok különbségére külön elnevezést vezetünk be:

**Definíció 9 (Részhalmaz komplementere)** Legyen  $A \subseteq X$ . Ekkor az  $X \setminus A$ , halmaz jelölésére az  $A^c$  szimbólumot is használjuk, és az  $A$  részhalmaz  $X$ -re vonatkozó komplementerének, röviden: komplementerének mondjuk.

A komplementer képzés ezek szerint egy  $X$  rögzített de tetszőleges halmaz részhalmazaira definiált operáció, ami egy részhalmazhoz egy másik részhalmazt rendel. Az  $A^c$  jelölés csak akkor egyértelmű, ha közben tudjuk, hogy van egy  $X$  kiinduló halmaz, aminek az  $A$  halmazt részhalmazának tekintjük.



Az  $A$  és  $B$  uniója: a valamilyen módon pontozott rész.

Az  $A$  és  $B$  metszete: a duplán pontozott rész.

1.1. ábra: Két halmaz uniójának és metszetének a szemléltetése.

Most először megnézzük, hogy milyen tulajdonságokkal rendelkeznek a metszet, unió műveletek és a komplementer képzés. A tulajdonságok rendszerezésénél az

algebrai struktúráknál használatos fogalmak vezethetnek bennünket. A műveletek tulajdonságait egy adott halmaz részhalmazainak a  $\mathcal{P}(X)$  összességére fogalmazzuk meg, de ez nem jelent megszorítást, mert az  $X$  halmaznak mindig vehetünk valamilyen halmazt, például az azonosságban szereplő halmazok unióját.

**Állítás 10 (Metszet, unió és komplementer tulajdonságai)** *Legyenek  $A, B, C$  az  $X$  halmaz részhalmazai. Ekkor a metszet, unió és komplementer műveletek teljesítik a következő azonosságokat.*

$$(1) \text{ Kommutativitás: } A \cup B = B \cup A \quad \text{és} \quad A \cap B = B \cap A.$$

$$(2) \text{ Asszociativitás:}$$

$$A \cup (B \cup C) = (A \cup B) \cup C \quad \text{és} \quad A \cap (B \cap C) = (A \cap B) \cap C.$$

$$(3) \text{ Idempotencia: } A \cup A = A \quad \text{és} \quad A \cap A = A.$$

$$(4) \text{ Disztributivitás:}$$

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \quad \text{és} \quad (A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

$$(5) \quad A \cup \emptyset = A \quad \text{és} \quad A \cap \emptyset = \emptyset.$$

$$(6) \quad A \cup X = X \quad \text{és} \quad A \cap X = A.$$

$$(7) \quad A \cup A^c = X \quad \text{és} \quad A \cap A^c = \emptyset.$$

$$(8) \text{ deMorgan azonosságok: } (A \cup B)^c = A^c \cap B^c \quad \text{és} \quad (A \cap B)^c = A^c \cup B^c.$$

$$(9) \quad A \subseteq B \implies B^c \subseteq A^c.$$

$$(10) \quad (A^c)^c = A.$$

**Bizonyítás.** Az (1), (3), (5), (6), (7), (9) és (10) tulajdonságok nyilvánvalóak vagy nagyon könnyen indokolhatók. Az asszociativitás és disztributivitás igazolása egyszerű, de hosszadalmas. Azt a már említett utat követjük, hogy belátjuk: a megfelelő azonosság baloldala része a jobboldalnak, és a jobboldal része a baloldalnak. Hasznos, ha a bizonyítások folyamatát a Venn-diagramokon is követjük.

**(2):** Lássuk először az unió asszociativitását. A következő ekvivalencia sorozat igazolja az azonosságot:  $x \in (A \cup B) \cup C \Leftrightarrow x \in A \cup B \text{ vagy } x \in C \Leftrightarrow (x \in A \text{ vagy } x \in B) \text{ vagy } x \in C \Leftrightarrow x \in A \text{ vagy } x \in B \text{ vagy } x \in C \Leftrightarrow x \in A \text{ vagy } (x \in B \text{ vagy } x \in C)$

Ha az előző ekvivalencia sorozatban a “vagy” helyére “és” műveletet helyettesítünk, akkor adódik a metszet asszociativitása.

**(4):** Az első disztributivitást látjuk be, a másik hasonlóan igazolható. A következő implikáció sorozat adja az igazolást:  $x \in (A \cup B) \cap C \Leftrightarrow (x \in A \text{ vagy } x \in B) \text{ és } x \in C \Leftrightarrow (x \in A \text{ és } x \in C) \text{ vagy } (x \in B \text{ és } x \in C) \Leftrightarrow x \in (A \cap C) \cup (B \cap C).$

(8): A következő ekvivalencia sorozatból:  $x \in (A \cup B)^c \Leftrightarrow x \notin A \cup B \Leftrightarrow x \notin A \text{ és } x \notin B \Leftrightarrow x \in A^c \text{ és } x \in B^c \Leftrightarrow x \in A^c \cap B^c$ .  $\square$

Az asszociativitás különösen fontos tulajdonság, mert az teszi lehetővé a művelet több elemre való kiterjesztését. Aszerint ugyanis az  $A \cap B \cap C$  három tagú művelet kiszámolása akár az  $A \cap (B \cap C)$ , akár pedig az  $(A \cap B) \cap C$  módon megtörténhet, ezért egyértelműen definiáltnak tekinthető. Így tetszőleges véges számú tagra kiterjeszthetők a műveletek teljes indukció segítségével.

Szembeötlő az, hogy az azonosságok “párban” teljesülnek. A párok képzési szabálya:

*Cseréld fel az unió és metszet jeleket, és vedd mindegyik halmaz komplementerét.*

Ez első pillantásra nem látszik teljesülni az első öt azonosságban, de tartalmilag mindegyikre helyes az elv, amit az (1) kommutativitásra meg is mutatunk.

Az unió kommutativitásából az elvet követve azt kapjuk, hogy

$$A^c \cap B^c = B^c \cap A^c. \quad (1.7)$$

Vegyük figyelembe, hogy ha az  $A$  és  $B$  halmazok tetszőleges részhalmazai az  $X$  halmaznak, akkor az  $A^c$  és  $B^c$  komplementerek is azok, és így a (1.7) azonosság pontosan azt jelenti, hogy a metszet művelet kommutatív, tehát igaz az (1) azonosság-pár második azonossága is, ha az első igaz.

Arra a tényre, hogy az  $X$  halmaz részhalmazainak a  $\mathcal{P}(X)$  rendszere kielégíti az előző tételben szereplő tulajdonságokat, azt mondjuk, hogy a  $\mathcal{P}(X)$  a metszet és unió műveletekkel *Boole-algebrát* alkot.

A műveletek tulajdonságai alapján “számolni” tudunk a halmazokkal, de meg kell mondani, hogy ez a számolás meglehetősen sajátos és szokatlan. Ha ebben a struktúrában be kell látni egy azonosságot, akkor kétféle út kínálkozik:

- 1) Számolunk a halmazalgebra 10. tételben felsorolt szabályai szerint.
- 2) Közvetlenül a halmazok azonosságának a definícióját alkalmazva, megmutatjuk: Ha egy  $x$  eleme a bizonyítandó azonosság egyik oldalának, akkor eleme a másiknak is.

A tétel bizonyításában természetesen a második módszert alkalmaztuk, most pedig egy példát mutatunk halmaz algebrában való számolásra, de általában felesleges ezt erőltetni, mert a 2) alatti módszer rendszerint célravezetőbb.

**Példa 1.1** *Mutassuk meg, hogy*

$$(A \cap B^c) \cup (A^c \cap B) = (A \cup B) \cap (A \cap B)^c.$$

Megoldás. A disztributivitás alapján

$$(A \cap B^c) \cup (A^c \cap B) = (A \cup (A^c \cap B)) \cap (B^c \cup (A^c \cap B)).$$

A jobboldal első tagja a disztributivitás és a tétel (7) és (5) azonossága felhasználásával

$$A \cup (A^c \cap B) = (A \cup A^c) \cap (A \cup B) = X \cap (A \cup B) = A \cup B.$$

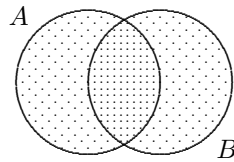
A jobboldal második tagja hasonló átalakítással, a deMorgan azonosságot is alkalmazva:

$$\begin{aligned} B^c \cup (A^c \cap B) &= (B^c \cup A^c) \cap (B^c \cup B) = (B^c \cup A^c) \cap X = \\ &= B^c \cup A^c = (A \cap B)^c. \end{aligned}$$

Összefoglalva az előző eredményeket, adódik, hogy

$$(A \cap B^c) \cup (A^c \cap B) = (A \cup B) \cap (A \cap B)^c.$$

□



Az  $A \setminus B$  különbség: az  $A$  egyszeresen pontozott része.

Az  $A \triangle B$  szimmetrikus differencia: az egyszeresen pontozott rész.

1.2. ábra: Két halmaz különbsége és szimmetrikus differenciája.

### 1.3.2 Halmazok szimmetrikus differenciája

Ebben az alponban egy további — kevésbé közismert, de igen fontos — halmazok közötti műveletet vezetünk be:

**Definíció 11 (Szimmetrikus differencia)** Két  $A$  és  $B$  halmaz szimmetrikus differenciája azon elemekből áll, amelyek elemei az  $A$  vagy  $B$  halmazoknak, de nem elemei mindkettőnek. Formálisan — a “ $\triangle$ ” műveleti jelet használva — :

$$A \triangle B \stackrel{\text{def}}{=} (A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A).$$

A definíció helyességéhez igazolni kell, hogy

$$(A \cup B) \setminus (A \cap B) = (A \setminus B) \cup (B \setminus A).$$

Az  $A \setminus B = A \cap B^c$  azonosság alapján kiindulva, majd pedig az 1.1. példa azonosságát használva azt kapjuk, hogy

$$\begin{aligned} (A \setminus B) \cup (B \setminus A) &= (A \cap B^c) \cup (A^c \cap B) = (A \cup B) \cap (A \cap B)^c = \\ &= (A \cup B) \setminus (A \cap B). \end{aligned}$$

A szimmetrikus differencia műveletet az 1.2. ábrán illusztráljuk, s a tulajdonságainak a vizsgálatát a gyakorlatokra hagyjuk.

## 1.4 Halmazok szorzata, relációk

### 1.4.1 Halmazok szorzata

Az előzőekben láttuk, hogy a halmazműveletek segítségével halmazokból újabb halmazokat lehet megadni. Most is azt tesszük, hogy halmazokból újabb halmazokat állítunk elő.

Illusztrálva a bevezetendő fogalom fontosságát, emlékeztetünk arra, hogy Descartes nevéhez fűződik a következő nevezetes felfedezés. Ha a sík pontjaihoz koordinátákat — valós számpárokat — rendelünk, akkor a geometriai alakzatok vizsgálata számpárok algebrai, analízisbeli tanulmányozására vezet. Ez az észrevétel azért jelentős, mert így az algebra és analízis jól kidolgozott eszköztára a geometriai kutatások szolgálatába állítható. Erre a történeti érdemre való tekintettel halmazok Descartes-szorzatának nevezték el a következő fogalmat:

**Definíció 12 (Két Halmaz szorzata)** Az  $A$  és  $B$  halmazok szorzatának (Descartes-szorzatának) nevezzük —  $A \times B$  módon jelöljük, és “ $A$  kereszt  $B$ ”-nek mondjuk — a halmazok elemeiből alkotott

$$(a, b) \quad a \in A \quad \text{és} \quad b \in B$$

rendezett párok halmazát, formálisan:

$$A \times B \stackrel{\text{def}}{=} \{(a, b) \mid a \in A \quad \text{és} \quad b \in B\}.$$

Ha két azonos halmazról van szó, akkor az  $A \times A$  halmazra az  $A^2$  jelölés is szokásos, aminek az olvasása: “ $A$  kettő”.

A halmazszorzat helyett Descartes-szorzatot is szokás mondani, de mi hacsak nem feltétlenül szükséges nem használjuk a matematikusok neveit. A definíció kiterjeszthető tetszőleges véges számú halmazra is és később majd végtelen sok halmazra is kiterjesztjük:

**Definíció 13 (Halmazok szorzata)** A 12. definíciót általánosítva az  $A_1, \dots, A_n$  halmazok szorzatának (Descartes-szorzatának) nevezzük az

$$A_1 \times A_2 \times \dots \times A_n \stackrel{\text{def}}{=} \{(a_1, \dots, a_n) : a_1 \in A_1, \dots, a_n \in A_n\}.$$

halmazt. Ha  $A_i = A$  minden  $i$ -re, akkor az  $A^n$  jelölést is használhatjuk.

A legfontosabb példát akkor kapjuk, amikor a szorzat mindegyik tényezője a valós számok  $\mathbb{R}$  halmaza, és a valós számokból alkotott rendezett  $n$ -esek összességét kapjuk meg, amit  $\mathbb{R}^n$  jelöl. Különösen fontos az  $\mathbb{R}^2$  “sík”, amely — szemléletes volta miatt — a legtöbb illusztratív példára nyújt lehetőséget.

Halmazok (Descartes-)szorzata természetesen újabb halmaz, és semmiféle külön összefüggés, struktúra sincsen az elemei között. Ha azonban valamilyen struktúra van a halmazokon, akkor annak a segítségével a szorzatukon is megadható

valamilyen struktúra. Például a valós számok az összeadásra nézve kommutatív, asszociatív, nulla elemes és inverz (negatív) elemes struktúra. Ha az  $\mathbb{R}^2$  szorzaton egy szintén  $+$ -szal jelölt műveletet az

$$(a, b) + (c, d) \stackrel{\text{def}}{=} (a + c, c + d)$$

módon értelmeziünk, akkor ezzel a művelettel az  $\mathbb{R}^2$  halmazon az egész számok struktúrájával azonos tulajdonságú struktúrát adtunk meg. Ez a gondolat igen sokszor előfordul a későbbiekben, mint hasznos struktúraépítő módszer. Általános elvnek is tekinthetjük: egy struktúra vizsgálatában mindig sorra kell kerülnie a *szorzatstruktúra* vizsgálatának.

### 1.4.2 Relációk

A reláció fogalmának a bevezetéséhez vegyünk egy előzetes példát: Ha az ebben a teremben lévő emberek halmaza az  $A$ , a  $B$  pedig egy édességeket áruló bolt áruinak az összessége, akkor felvethető az a kérdés, hogy ki szereti vagy nem szereti valamelyik édességet.

Mit is jelent ez formálisabban? Ha az  $a \in A$  egyed szereti  $b \in B$  édességet, akkor ezt azzal jellemezhetjük, hogy az  $(a, b)$  párost kivesszük az összes lehetséges párok  $A \times B$  halmazából. Az ilyen módon kivett párosok halmazát — az egyedek és az általuk szeretett édességek párosait — jelöljük  $\mathcal{S}$ -sel. Így azt mondhatjuk, hogy az  $\mathcal{S}$  egyszerűen egy részhalmaza az  $A \times B$  szorzatnak. Mivel a “szeretem ezt az édességet” viszonyt (relációt) az  $\mathcal{S}$  egyértelműen jellemzi, ezért nem meglepő, hogy egyszerűen ezt a halmazt nevezzük a megfelelő relációnak. Lássuk ezek után a pontos definíciót.

**Definíció 14 (Bináris reláció)** *Legyenek  $X$  és  $Y$  tetszőleges halmazok. Az  $(X, Y)$  halmazpáron értelmezett (bináris) relációnak mondjuk az  $X \times Y$  szorzat egy tetszőleges  $\mathcal{R}$  részhalmazát. Azt, hogy egy  $(x, y)$  elempáros eleme az*

$$\mathcal{R} \subseteq X \times Y$$

*relációnak úgy is mondjuk, hogy az  $x$  az  $\mathcal{R}$  relációban van az  $y$  elemmel, és ezt*

$$x \mathcal{R} y$$

*módon is írjuk.*

*Ha  $Y = X$ , akkor az  $\mathcal{R} \subseteq X \times X = X^2$  relációt az  $X$  halmazon lévő (bináris) relációnak mondjuk.*

Jegyezzük meg, hogy egy reláció *halmaz*, és az  $x \mathcal{R} y$  írásmód ne zavarjon meg bennünket. Ezen mindig csak azt értsük, hogy  $(x, y) \in \mathcal{R}$ . A megelőző jelölés onnan ered, hogy az a szokásos a valós számok két nevezetes relációjának az egyenlőségnek és az egyenlőtlenségnek az írásánál.

Felhívjuk a figyelmet arra, hogy egy reláció egy halmaz ugyan, de nem pusztán az elemei határozzák meg, hanem az is, hogy milyen halmazok szorzatának a részhalmaza. Ezért is mondjuk így: az  $\mathcal{R}$  egy  $(X \times Y)$ -ből való — vagy rövidebben:  $(X \times Y)$  — reláció.

A reláció fogalom általánosabban is megfogalmazható, de ezt nem gyakran fogjuk használni:

**Definíció 15 (Reláció)** *Legyenek az  $X_1, \dots, X_n$  tetszőleges halmazok. Az  $A_1 \times \dots \times A_n$  szorzat egy  $\mathcal{R}$  részhalmazát ( $n$ -áris) relációnak nevezzük.*

*Ha  $X = X_1 = \dots = X_n$ , akkor és  $\mathcal{R} \subseteq X^n$ , akkor azt mondjuk, hogy az  $\mathcal{R}$   $n$ -áris reláció az  $X$  halmazon.*

Lássunk most néhány példát. Először — mint legklasszikusabb példát — a valós számok egyenlőségét és egyenlőtlenségét tekintjük.

**Példa 1.2 (Valós számok egyenlősége)** *Az  $\mathbb{R}$  valós számokon vegyük az egyenlőség “=” relációját, azaz*

$$= \doteq \{(x, y) \in \mathbb{R}^2 : x = y\}.$$

Ne lepődjünk meg azon, hogy az “=” egyenlőség jel halmazt jelöl, hiszen reláció, tehát halmaz. Ennek a következő közismert tulajdonságai vannak: 1)  $(x, x) \in \mathbb{R}^2$ , azaz minden elem relációban van önmagával. 2) Ha  $x = y$ , akkor  $y = x$ . 3) Ha  $x = y$  és  $y = z$ , akkor  $x = z$ . Az egyenlőség relációját — mint az  $\mathbb{R}^2$  sík egy részhalmazát — az 1.3. ábrán szemléltetjük. A felrajzolásra gondolva a relációt megadó halmazt — a jelen esetben: az  $\{(x, x) : x \in \mathbb{R}\}$  egyenest — a reláció gráfjának is szokás mondani, ami egy kicsit szószaporítás, hiszen az maga a reláció (1.3. ábra).

**Példa 1.3 (Valós számok egyenlőtlensége)** *A valós számokon tekintsük az*

$$\leq \doteq \{(x, y) \in \mathbb{R}^2 : x \leq y\}$$

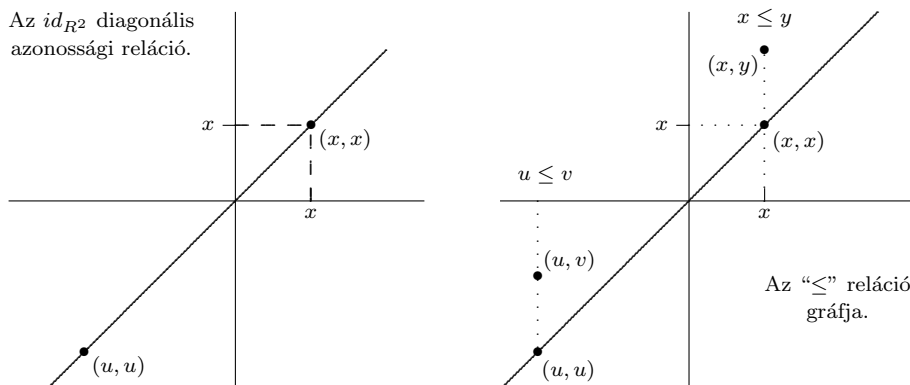
*bináris relációt.*

A “nem nagyobb, mint” (“kisebb egyenlő, mint” relációnak a következő tulajdonságai vannak: 1) Minden  $x$ -re  $x \leq x$ . 2) Ha  $x \leq y$  és  $y \leq x$ , akkor  $x = y$ . 3) Ha  $x \leq y$  és  $y \leq z$ , akkor  $x \leq z$ . Az egyenlőtlenség relációját — mint az  $\mathbb{R}^2$  sík egy részhalmazát — az 1.3. ábrán szemléltetjük.

Mutassunk egy példát hármas relációra is:

**Példa 1.4** *Vegyük az  $\mathbb{R}^3$  szorzat következő relációját:*

$$\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3 : z = x + y\}.$$

1.3. ábra: Az " $=$ " és " $\leq$ " relációk gráfjai.

Az  $\mathcal{S}$  relációt ésszerű lett volna " $+$ " jellel jelölni, mivel egy hármas pontosan akkor tartozik bele, ha a harmadik tag az első kettő összege.

Az előző példákban azt találtuk, hogy a fontos relációknak vannak bizonyos tulajdonságaik. A következő definícióban felsoroljuk a leggyakrabban használt reláció tulajdonságokat:

**Definíció 16 (Reláció tulajdonságok)** Egy  $\mathcal{R} \subseteq X \times X = X^2$  relációra a következő tulajdonságok elnevezését vezetjük be.

**Reflexivitás.**  $x \mathcal{R} x$ ,  $((x, x) \in X^2)$  minden  $x \in X$  elemre.

**Irreflexivitás** Az  $x \mathcal{R} x$  semmilyen  $x \in X$  elemre sem teljesül.

**Szimmetricitás** Az  $x \mathcal{R} y$  teljesülése maga után vonja az  $y \mathcal{R} x$  teljesülését.

**Antiszimmetria** Az  $x \mathcal{R} y$  és  $y \mathcal{R} x$  relációk fennállásából következik az  $x$  és  $y$  elemek egyenlősége, az  $x = y$ ;

**Szigorú antiszimmetria, aszimmetria** Ha az  $x \mathcal{R} y$  teljesül, akkor az  $y \mathcal{R} x$  nem teljesülhet.

**Tranzitivitás** Az  $x \mathcal{R} y$  és  $y \mathcal{R} z$  relációk fennállásából következik az  $x \mathcal{R} z$  teljesülése.

**Teljesség** Minden  $x, y \in X$  elempárra az  $x \mathcal{R} y$  és  $y \mathcal{R} x$  relációk közül legalább az egyik teljesül.

A következő definícióban néhány fontos, speciális reláció elnevezését adjuk meg.

**Definíció 17 (Komplett, üres, diagonális relációk)** Az  $X$  halmazon értelmezett  $\emptyset \subseteq X^2$  relációt üres, az  $X^2$  relációt komplett, az  $\{(x, x) : x \in X\}$  relációt pedig diagonális relációnak vagy azonosságnak fogjuk nevezni, és az  $\Delta_X$  vagy  $\text{id}_X$  jelölést használjuk.

Nyilvánvalóak a következők: Az üres reláció az 16. definíció valamennyi tulajdonságát teljesíti, kivéve a reflexivitást és a teljességet. A komplett reláció az 16. definíció valamennyi tulajdonságával rendelkezik, kivéve az irreflexivitást, antiszimetriát és szigorú antiszimetriát. Az azonosság relációja reflexív, szimmetrikus és tranzitív és antiszimetrikus.

A relációk között műveleteket is tudunk definiálni a következők szerint:

**Definíció 18 (Relációk kompozíciója, inverze, komplementere)** Legyenek  $X, Y$  és  $Z$  tetszőleges halmazok. Az  $\mathcal{R} \subseteq X \times Y$  és  $\mathcal{S} \subseteq Y \times Z$  relációk kompozíciója — amit  $\mathcal{S} \circ \mathcal{R}$  jelöl — az alábbi  $X \times Z$ -ből vett reláció:

$$\mathcal{S} \circ \mathcal{R} \stackrel{\text{def}}{=} \{(x, z) \in X \times Z : \exists y \in Y \quad (x, y) \in \mathcal{R} \text{ és } (y, z) \in \mathcal{S}\}.$$

Az  $\mathcal{R} \subseteq X \times Y$  reláció inverzének mondjuk, és  $\mathcal{R}^{-1}$ -vel jelöljük az

$$\mathcal{R}^{-1} \stackrel{\text{def}}{=} \{(y, x) \in Y \times X : (x, y) \in \mathcal{R}\}$$

$(Y \times X)$ -ből való relációt.

Az  $\mathcal{R} \subseteq X \times Y$  reláció komplementerének (tagadásának) mondjuk, és  $\mathcal{R}^c$ -vel jelöljük az

$$\mathcal{R}^c \stackrel{\text{def}}{=} \{(x, y) \in X \times Y : (x, y) \notin \mathcal{R}\}$$

$(X \times Y)$ -beli relációt.

A relációk szorzatának a következő fontos tulajdonságai vannak:

**Állítás 19 (Relációk kompozíciójának a tulajdonságai)** Legyenek az  $X, Y, Z, W$  tetszőleges halmazok, és

$$\mathcal{R}_1 \subseteq X \times Y, \quad \mathcal{R}_2 \subseteq Y \times Z, \quad \mathcal{R}_3 \subseteq Z \times W.$$

(1) A relációk kompozíciója asszociatív:

$$\mathcal{R}_3 \circ (\mathcal{R}_2 \circ \mathcal{R}_1) = (\mathcal{R}_3 \circ \mathcal{R}_2) \circ \mathcal{R}_1.$$

(2) Legyen  $\mathcal{R} \subseteq X \times Y$ . Az azonosság relációja bizonyos reprodukáló tulajdonsággal bír: (a)  $\text{id}_Y \circ \mathcal{R} = \mathcal{R}$  és (b)  $\mathcal{R} \circ \text{id}_X = \mathcal{R}$ .

(3) Ha  $\mathcal{R} \subseteq X^2$ , akkor  $\text{id}_X \circ \mathcal{R} = \mathcal{R} \circ \text{id}_X = \mathcal{R}$ .

Bizonyítás. (1): Ha

$$(x, w) \in \mathcal{R}_3 \circ (\mathcal{R}_2 \circ \mathcal{R}_1), \quad (1.8)$$

akkor van olyan  $z \in Z$ , amelyre  $(x, z) \in (\mathcal{R}_2 \circ \mathcal{R}_1)$ , és  $(z, w) \in \mathcal{R}_3$ . Az első “ $\in$ ” pontosan akkor áll fenn, ha van olyan  $y \in Y$ , amelyre  $(x, y) \in \mathcal{R}_1$  és  $(y, z) \in \mathcal{R}_2$ . Összefoglalva: az (1.8) pontosan akkor áll fenn, ha

$$\exists z \in Z \ y \in Y, \quad (x, y) \in \mathcal{R}_1, \quad (y, z) \in \mathcal{R}_2 \quad \text{és} \quad (z, w) \in \mathcal{R}_3. \quad (1.9)$$

A másik irányú tartalmazás pontosan így igazolható: Ha

$$(x, w) \in (\mathcal{R}_3 \circ \mathcal{R}_2) \circ \mathcal{R}_1, \quad (1.10)$$

akkor van olyan  $y \in Y$ , amelyre  $(y, w) \in \mathcal{R}_3 \circ \mathcal{R}_2$  és  $(x, y) \in \mathcal{R}_1$ . Az első “ $\in$ ” pontosan akkor áll fenn, ha van olyan  $z \in Z$ , amelyre  $(y, z) \in \mathcal{R}_2$  és  $(z, w) \in \mathcal{R}_3$ . Összefoglalva: az (1.10) pontosan akkor áll fenn, ha az (1.9) teljesül.

(2): (a): A kompozíció képzésének a szabálya szerint:

$$\text{id}_Y \circ \mathcal{R} = \{(x, y) : \exists y \in Y \quad (x, y) \in \mathcal{R} \text{ és } (y, y) \in \text{id}_Y\} = \mathcal{R}.$$

(b) Az előzőhöz hasonlóan:

$$\mathcal{R} \circ \text{id}_X = \{(x, y) : \exists x \in X \quad (x, x) \in \text{id}_X \text{ és } (x, y) \in \mathcal{R}\} = \mathcal{R}.$$

(3): Az (2)(a) és (2)(b) azonosságokból evidens.  $\square$

A 17. definícióban elnevezett speciális relációkat és az 18. definícióban meghatározott reláció műveleteket hasznosan alkalmazhatjuk a 16. definícióban felsorolt reláció tulajdonságok ekvivalens megadására:

**Állítás 20 (Reláció tulajdonságok)** *Legyen az  $\mathcal{R} \subseteq X \times X = X^2$  egy reláció. A 16. definícióban felsorolt reláció tulajdonságok a következő módokon is meghatározhatók.*

**Reflexivitás.**  $\text{id}_X \subseteq \mathcal{R}$ .

**Irreflexivitás.**  $\mathcal{R} \cap \text{id}_X = \emptyset$ .

**Szimmetritás.**  $\mathcal{R} = \mathcal{R}^{-1}$ .

**Antiszimmetria.**  $\mathcal{R} \cap \mathcal{R}^{-1} \subseteq \text{id}_X$ .

**Szigorú antiszimmetria, aszimmetria.**  $\mathcal{R} \cap \mathcal{R}^{-1} = \emptyset$ .

**Tranzitivitás.**  $\mathcal{R} \circ \mathcal{R} \subseteq \mathcal{R}$ .

**Teljesség.**  $\mathcal{R} \cup \mathcal{R}^{-1} = X \times X$ .

**Bizonyítás.** Az indoklások nagyon rövidek, az olvasóra is hagyhatnánk az igazolást.  
*Reflexivitás:* A reflexivitás pontosan azt jelenti, hogy  $(x, x) \in \mathcal{R}$  minden  $x \in X$  esetében, azaz  $\text{id}_X \subseteq \mathcal{R}$ .

*Irreflexivitás:* Az irreflexivitás azt jelenti, hogy az  $(x, x)$  nem eleme az  $\mathcal{R}$  relációnak semmilyen  $x \in X$  mellett sem:  $\text{id}_X \cap \mathcal{R} = \emptyset$ .

*Szimmetritás:* A szimmetritás és az inverz definíciójából evidens.

*Antiszimmetria:* Ha  $(x, y) \in \mathcal{R} \cap \mathcal{R}^{-1} \subseteq \text{id}_X$ , akkor  $(x, y) \in \mathcal{R}$  és  $(x, y) \in \mathcal{R}^{-1}$ , amiből  $(x, y), (y, x) \in \mathcal{R}$ , és így az antiszimmetria miatt:  $x = y$ , tehát  $(x, y) \in \text{id}_X$ .

*Aszimmetria:* Az  $\mathcal{R} \cap \mathcal{R}^{-1} = \emptyset$  baloldalának pontosan akkor nincs eleme, ha nincs olyan  $(x, y)$ , amelyre  $(x, y) \in \mathcal{R}$  és  $(x, y) \in \mathcal{R}^{-1}$ , ami pontosan az aszimmetria.

*Tranzitivitás:* Az  $\mathcal{R} \circ \mathcal{R} \subseteq \mathcal{R}$  egyenlőség pontosan akkor áll fenn, ha  $(x, y) \in \mathcal{R}$  és  $(y, z) \in \mathcal{R}$  maga után vonja, hogy  $(x, z) \in \mathcal{R}$ , ami éppen a tranzitivitás.

*Teljesség:* Az  $\mathcal{R} \cup \mathcal{R}^{-1} = X \times X$  egyenlőség pontosan azt mondja, hogy tetszőleges  $(x, y \in X)$  esetében vagy  $(x, y) \in \mathcal{R}$  vagy  $(x, y) \in \mathcal{R}^{-1}$ . Mivel az utóbbi azzal ekvivalens, hogy  $(y, x) \in \mathcal{R}$ . Emiatt az  $\mathcal{R} \cup \mathcal{R}^{-1} = X \times X$  egyenlőség azzal ekvivalens, hogy  $(x, y) \in \mathcal{R}$  vagy  $(y, x) \in \mathcal{R}$ , ami a teljesség definíciója.  $\square$

Két fontos relációtípus tárgyalunk még. Az egyenlőség relációjának az absztrakciójából ered a következő:

**Definíció 21 (Ekvivalencia reláció)** *Ha az  $\mathcal{R} \subseteq Y$  egy reflexív, szimmetrikus és tranzitív reláció, akkor ekvivalencia relációnak fogjuk mondani.*

Egy ekvivalencia reláció mindig megadja a halmaznak egy diszjunkt részhalmazokra való felbontását, amelyhez szükséges fogalmat el is nevezzük:

**Definíció 22** *Legyen az “ $\equiv$ ” egy ekvivalencia reláció az  $X$  halmazon. Valamely  $x \in X$  elemmel “ $\equiv$ ” relációban lévő elemek*

$$H_x = \{y \in X : y \mathcal{R} x\} \quad (1.11)$$

*halmazát az  $x$  elemhez tartozó  $\equiv$ -ekvivalencia osztálynak fogjuk nevezni.*

Ha egy fix relációval foglalkozunk, akkor nem okoz félreértést, ha  $\equiv$ -ekvivalencia osztály helyett egyszerűen csak ekvivalencia osztályt mondunk.

Az ekvivalencia osztályok halmazok, és az “osztály” elnevezéssel kapcsolatban meg kell említeni, hogy ettől az egy esettől eltekintve ne használjuk a halmaz szinonimájaként, mert másra van fenntartva.

Az ekvivalencia osztályok tulajdonságait fogalmazzuk meg a következő állításban:

**Állítás 23 (Ekvivalencia osztályok tulajdonságai)** *Legyen az  $\equiv$  egy ekvivalencia reláció az  $X$  halmazon. A  $H_x, x \in X$  ekvivalencia osztályoknak a következő tulajdonságaik vannak.*

- (1)  $x \in H_x$ , minden  $x$ -re.

(2) Ha  $y \in H_x$ , akkor  $H_x = H_y$ , amit így szoktunk mondani: egy ekvivalencia osztály bármelyik elemével megadható, reprezentálható.

(3) A különböző ekvivalencia osztályok diszjunktak, vagyis

$$H_x \cap H_y \neq \emptyset \iff H_x = H_y.$$

Bizonyítás. (1): A reláció reflexivitása szerint  $x \equiv x$ , tehát  $x \in H_x$ .

(2): A szokásos módon kétirányú tartalmazást látunk be:

*Első lépés:* Ha  $y \in H_x$ , akkor  $H_y \subseteq H_x$ . Legyen az  $u$  egy tetszőleges eleme a  $H_y$  halmaznak, azaz  $u \equiv y$ . Hozzávéve az utóbbihoz az  $y \in H_x$  alapján fennálló,  $y \equiv x$  összefüggést, a tranzitivitás miatt  $u \equiv x$  adódik, tehát  $u \in H_x$ , ezért  $H_y \subseteq H_x$ .

*Második lépés:* Ha  $y \in H_x$ , akkor  $H_x \subseteq H_y$ . Legyen az  $u$  egy tetszőleges eleme a  $H_x$  halmaznak, azaz  $u \equiv x$ . Hozzávéve az utóbbihoz az  $y \in H_x$  alapján fennálló,  $y \equiv x$  összefüggést, a tranzitivitás miatt  $u \equiv y$  adódik, tehát  $u \in H_y$ , ezért  $H_x \subseteq H_y$ .

(3): Ha  $u \in H_x \cap H_y$ , akkor az ekvivalencia osztályok definíciója szerint:  $u \equiv x$  és  $u \equiv y$ , amit a reláció szimmetritása miatt így is írhatunk:  $x \equiv u$  és  $u \equiv y$ . Ebből pedig a tranzitivitás alapján az  $x \equiv y$  adódik, ami az ekvivalencia osztály definíciója szerint azt jelenti, hogy  $x \in H_y$ . Ez utóbbiból pedig a már belátott (2) állítás alapján azt kapjuk, hogy  $H_x = H_y$ , ami bizonyítandó volt.  $\square$

A következő reláció típus a valós számok “kisebb-egyenlő” relációja általánosításának tekinthető:

**Definíció 24 (Parciális rendezés)** Ha az  $\mathcal{R} \subseteq Y$  egy reflexív, antiszimmetrikus és tranzitív reláció, akkor parciális rendezésnek fogjuk mondani.

Parciális rendezésre példa még: Egy  $X$  halmaz részhalmazainak a  $\mathcal{P}(X)$  rendszerén a tartalmazás relációja. Ez a reláció parciális rendezés, de nem teljes. Az olyan parciális rendezést, amely teljes is *rendezésnek* szokás mondani. A közgazdaságtan legfontosabb relációja a reflexív, tranzitív és teljes reláció, amit preferenciának nevezünk.

## 1.5 Függvények

### 1.5.1 A függvény fogalma

Ebben a pontban talán a matematika legfontosabb fogalmát, a *függvényt* és a vele kapcsolatos legáltalánosabb fogalmakat vezetjük be. A függvényt, mint speciális relációt definiáljuk:

**Definíció 25 (Függvény, leképezés)** Legyenek az  $X$  és  $Y$  tetszőleges halmazok. Egy  $f \subseteq X \times Y$  relációt az  $X$  halmazból az  $Y$  halmazba menő függvénynek (leképezésnek) mondunk, ha teljesíti a következő két feltételt.

(i) Minden  $x \in X$  elemre van olyan  $y \in Y$  elem, hogy  $(x, y) \in X \times Y$ .

(ii) Ha  $(x, y_1) \in f$  és  $(x, y_2) \in f$ , akkor  $y_1 = y_2$ .

A függvény objektum leggyakrabban használatos jelölése:  $f : X \rightarrow Y$ . Az  $x \in X$  elemhez az  $Y$  halmazból párként egyértelműen hozzá tartozó  $y$  elemet  $f(x)$ -szel szokás jelölni, amit helyettesítési értéknek is mondunk.

Az  $X$  halmazt az  $f$  értelmezési tartományának is szokás mondani, és a jelölése:  $\text{dom}(f)$ . Az  $Y$  azon  $\{y \in Y : \exists x \in X \ y = f(x)\}$  elemeinek az összességét pedig, amelyekre az  $f$  képez valamilyen elemet, az  $f$  értékkészletének mondjuk, és  $\text{ran}(f)$  a szokásos jelölés.

Az  $X$  halmazból  $Y$ -ba menő függvények összességét az  $Y^X$  módon fogjuk jelölni.

Más szavakkal megismételve: Az  $X$  halmazról az  $Y$  halmazba menő  $f$  leképezés olyan  $(x, y) \in X \times Y$  elempárokból áll, amelyeknél az  $X$  minden  $x$  eleméhez egyetlen  $y \in Y$  elem tartozik, amit  $f(x)$ -szel jelölünk. A “leképezés” elnevezés szemléletére épülnek azok az ábrák, ahol nyíl mutatja a leképezés irányát.

Nyilvánvaló, hogy nem minden  $\mathcal{R} \subseteq X \times Y$  reláció függvény, mivel relációnál nem kívánalom az, hogy az értelmezési tartománya megegyezze az  $X$ -szel, továbbá nem feltétlenül egyetlen elem tartozik egy  $x$  elemhez. Példaképpen: Nem függvények a valós számokon vett “ $\leq$ ” és “ $<$ ” relációk.

Hangsúlyozni kell, hogy az  $f$  függvény egyetlen objektum, ami az  $X \times Y$  része, és nem szabad összekeverni az  $f$  függvényt az  $x$  helyen felvett  $f(x)$  helyettesítési értékével, ami az  $Y$  eleme.

A függvény és leképezés elnevezéseken kívül szokásosak még: *megfeleltetés*, *transzformáció*, *operátor*, *operáció*, *funkcionál* elnevezések, amelyek közül némelyek speciális esetekre vannak fenntartva.

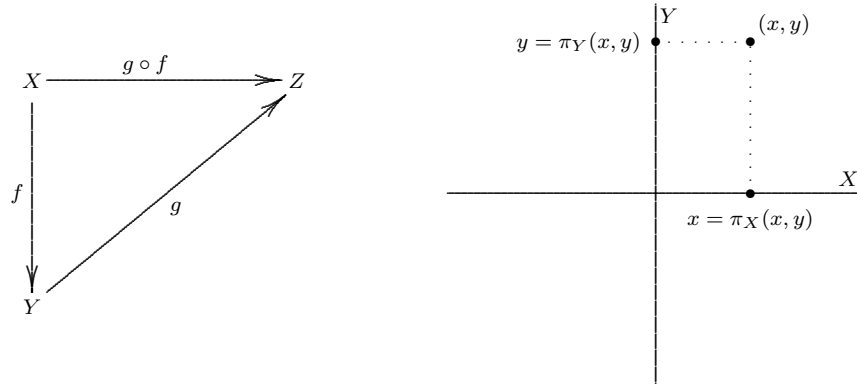
Az  $f : X \rightarrow Y$  jelölésen kívül — mások mellett — használatosak még:

$$x \mapsto f(x), \quad y = f(x), \quad \text{és} \quad x \overset{f}{\mapsto} y = f(x).$$

Ezek azonban csak akkor nyújtanak teljes információt, ha külön megmondjuk, hogy az  $X$  halmaz elemei jönnek az  $x$  helyére és az  $f(x)$  az  $Y$  halmazba tartozik.

Szokásos az is, hogy az  $x \in X$  elemet *független változónak*, az  $y = f(x)$  elemet pedig *függő változónak* nevezik. Ez az elnevezés onnan ered, hogy az  $x$  megadásával meg tudjuk mondani a hozzá tartozó  $y = f(x)$  értéket. A megadott függvény fogalom “statikus”, abban az értelemben, hogy a függvény “egyszerre” adva van, mint egyetlen objektum. Ezzel szemben — elsősorban a fizikai, időtől függő függvény elképzelése alapján — sokszor úgy képzeljük el a függvényt, mint egy “változás” leírását. Ehhez olyan elképzelés társul, hogy a függvény “fokozatosan” áll elő, ahogyan az  $x$  (az idő) halad. Ez a szemlélet — annak ellenére, hogy szellemében nem egyezik meg definícióinkkal — nem rossz.

Az  $\{(x, f(x)) : x \in X\}$  halmazt az  $f$  leképezés *gráfjának* (*grafikonjának*) is szokás nevezni. Ez az elnevezés onnan ered, hogy ha valós függvényről van



1.4. ábra: Függvények kompozíciója, vetítés.

szó, akkor a szóbanforgó pontok összessége a függvény gráfját adja az  $\mathbb{R}^2$  síkon. Hangsúlyozni kell azonban, hogy ez csak egy más neve az  $f$  függvényt megadó relációnak, hiszen a definíció szerint:  $f = \{(x, f(x)) : x \in X\}$ .

Két  $f$  és  $g$  függvény akkor egyenlő, ha mint reláció egyenlő, azaz ha mindkettő az  $X$  halmazból az  $Y$  halmazba képez, és minden  $x \in X$  elemre  $f(x) = g(x)$ . Ez nyilván többet jelent, mint az  $f = \{(x, f(x)) : x \in X\}$  és  $g = \{(x, g(x)) : x \in X\}$  halmazok megegyezése, hiszen ezek nem mondanak semmit sem az  $Y$  halmazról.

Három fontos leképezés elnevezése szerepel a következő definícióban:

**Definíció 26 (Identitás, beágyazás, projekció)** Az  $\text{id}_X$  relációt *identitás-, azonosság-leképezésnek* is nevezzük.

Legyen  $A \subseteq X$ . Azt a  $j_A : A \rightarrow X$  leképezést, amelyre

$$\forall x \in A \quad j_A(x) = x,$$

az  $A$  halmaz *beágyazási leképezésének* mondjuk.

Azt az  $\pi_X : (X \times Y) \rightarrow X$  leképezést, amely egy  $(x, y) \in X \times Y$  elempárhoz az  $x$  elemet rendeli, azaz  $\pi_X((x, y)) = x$ , az  $X$ -re való *vetítésnek* mondjuk. Hasonló az  $Y$ -ra való *vetítés definíciója*.

A vetítés (projekció) elnevezés nyilvánvalóan az  $\mathbb{R}^2$  sík szemlélete alapján született fogalom (1.4. ábra). A beágyazó függvény voltaképpen “beteszi” az  $A$  részhalmazt az  $X$  halmazba, a “beágyazás” szó az angol terminológia szokásos fordítása.

A következő két fogalom — meglepő egyszerűsége ellenére — nagyon fontos szerepet fog betölteni.

**Definíció 27 (Leszűkítés, kiterjesztés)** Az  $f : X \rightarrow Y$  leképezésnek a  $B \subseteq X$  részhalmazra vett *leszűkítése* az az  $f|_B : B \rightarrow Y$  függvény, amelyik a  $B$  halmazon megegyezik az  $f$  leképezéssel:  $\forall x \in B \quad f|_B(x) = f(x)$ .

Legyen  $f : X \rightarrow Y$ ,  $A \subseteq X$  és  $g : A \rightarrow Y$ . Ha az  $f$  és  $g$  függvények megegyeznek az  $A$  halmazon, akkor a  $f$ -et a  $g$  leképezés ( $A$ -ról  $X$ -re való) kiterjesztésének mondjuk.

Példaképpen lássuk még azt, hogy véges  $X$  értelmezési tartomány esetében hogyan adható meg egy függvény:

**Példa 1.5 (Rendezett  $n$ -esek)** Legyen az  $X$  halmaz az  $1, 2, \dots, n$  szimbólumokat tartalmazó  $\{1, 2, \dots, n\}$  halmaz, az  $Y$  pedig tetszőleges halmaz. Ekkor egy  $a : X \rightarrow Y$  függvény az

$$(a(1), a(2), \dots, a(n)) \in Y^n$$

rendezett  $n$ -essel azonosítható, ahol az  $a(n)$  helyett az  $a_n$  írásmód a szokásosabb.

Más szavakkal: Egy  $n$ -elemű halmazon definált,  $Y$ -ba képező függvény úgy fogható fel, mint az  $Y^n$  egy eleme.

Különösen fontos az az eset, amikor az  $Y$  a valós számokkal egyezik meg. Ekkor egy rendezett szám  $n$ -esekről beszélünk.

Az azonosításon azt értjük, hogy a függvénynek a definíció szerinti

$$\{(1, a(1)), \dots, (n, a(n))\}$$

megadását pontosan meg tudjuk mondani az

$$(a(1), a(2), \dots, a(n)) = (a_1, a_2, \dots, a_n)$$

rendezett  $n$ -es segítségével, és megfordítva. Ez voltaképpen a példa megoldása is.  $\square$

### 1.5.2 Függvények kompozíciója, inverze

Az  $f = \{(x, f(x)) : x \in X, f(x) \in Y\}$  függvényt mint relációt definiáltuk, és így — mint relációnak — van inverze:

$$f^{-1} = \{(f(x), x) : x \in X, f(x) \in Y\} \subseteq Y \times X.$$

Kérdés azonban:  $(Y \rightarrow X)$  függvény-e az inverz reláció. Ehhez — a 25. definíció szerint — két feltétel teljesülése szükséges:

- 1) Az  $f^{-1}$ -ben lévő elempárok első eleme ki kell hogy adja az  $Y$  halmazt, ami nyilván azt jelenti, hogy az  $f$  értékkészete a teljes  $Y$ :  $\text{Ran}(f) = Y$ .
- 2) Egy  $f(x) \in Y$  elemhez meghatározott (pontosan egy)  $x$  elem tartozzék:

$$f(x) = f(u) \implies x = u.$$

Rövidebben: Az  $f(x) \in Y$  elem párja az  $f^{-1}$  relációban csak az  $x$  lehet.

Az 1) feltétel elérhető azzal, hogy az  $Y$  halmazt megváltoztatjuk, és az  $f : X \rightarrow \text{Ran}(f)$  függvényt tekintjük, ami persze, nem egy  $(X \rightarrow Y)$  leképezés, ha a függvény nem teljesíti az 1)-et.

A 2) feltételt más nézőpontból közelítve: Egy  $y \in Y$  elemre vagy képez az  $f$  leképezés elemet az  $X$ -ből, vagy nem. Kérdés az, hogy ha képez elemet, akkor hányat. Ha csak egyet, akkor azt mondhatjuk, hogy a leképezés “egyrétű”, nincs “ráfényképezés”.

Az elmondott bevezetésből kiindulva rögzítsük a fogalmakat:

**Definíció 28 (Szurjektív, injektív, bijektív leképezés)** Egy  $f : X \rightarrow Y$  leképezés

**szurjektív**, ha az  $Y$  minden elemére képez elemet, azaz  $\text{Ran}(f) = Y$ ;

**injektív**, ha az  $Y$  halmaz egy elemére legfeljebb egy  $X$ -beli elemet képez,

**bijektív**, ha szurjektív és injektív.

A szurjektív, injektív és bijektív leképezésekre rövid elnevezések: szurjekció, injekció, bijekció.

A “bijektív” elnevezés helyett a “*kölcsönösen egyértelmű*” is használatos. A bijektivitás más szavakkal: Egy  $f : X \rightarrow Y$  leképezés pontosan akkor bijektív, ha minden  $y \in Y$  elemre pontosan egy  $x \in X$  elemet képez.

Tegyük fel, hogy van egy  $f$ , az  $X$  halmazból az  $Y$  halmazba menő, és egy  $g$ , az  $Y$  halmazból a  $Z$  halmazba menő leképezésünk. Ekkor egy  $x \in X$  pontot az  $f$  leképezéssel átviszünk az  $Y$  halmaz  $f(x)$  elemébe, majd pedig onnan a  $g$  leképezéssel tovább visszük a  $Z$  halmaz  $g(f(x))$  pontjába. Ennek az eljárásnak a definícióját adja meg a következő:

**Definíció 29 (Függvények kompozíciója)** Az  $f : X \rightarrow Y$  és  $g : Y \rightarrow Z$  leképezések kompozíciójának mondjuk — és a  $g \circ f$  szimbólummal jelöljük — a következő módon megadott  $(X \rightarrow Z)$  függvényt:

$$x \mapsto g(f(x)), \quad x \in X$$

Szokásos még az “összetett függvény” elnevezés is. Az  $f \circ g$  kompozíciót a 1.4. ábra első fele szemlélteti.

A relációkra már definiáltunk egy kompozíciónak nevezett műveletet, és függvény is reláció, ezért a következetes szóhasználat miatt meg kell néznünk, hogy a definíció kompozíciója megegyezik-e a relációkra definiált kompozícióval. A “\*” jelölje most átmenetileg a relációk kompozícióját. Ekkor

$$\begin{aligned} g * f &= \{(x, z) \in X \times Z : \exists y \in Y \ (x, y) \in f \text{ és } (y, z) \in g\} = \\ &= \{(x, z) : \exists y \in Y \ y = f(x) \text{ és } z = g(y)\}. \end{aligned}$$

Az utolsó tagban az  $\exists y \in Y$  előtag felesleges, hiszen létezik az  $y = f(x)$ , ezért végül is:

$$g * f = \{(x, z) : y = f(x) \text{ és } z = g(y)\} = \{(x, z) : z = g(f(x))\},$$

tehát  $g * f = g \circ f$ , tehát a relációk és függvények kompozíciója azonos fogalom.

A relációk kompozíciójára már kimondtuk a következő állítást, tehát voltaképpen ismételnünk:

**Állítás 30 (Függvény kompozíció tulajdonságai)** *Legyenek  $f : X \rightarrow Y$ ,  $g : Y \rightarrow Z$ . Ekkor*

- (1) *a kompozíció asszociatív:  $h \circ (g \circ f) = (h \circ g) \circ f$ ;*
- (2)  *$f \circ \text{id}_X = f$ ;*
- (3)  *$\text{id}_Y \circ f = f$ .*

**Állítás 31 (Függvény inverze)** *Egy  $f : X \rightarrow Y$  függvény — mint reláció —  $f^{-1}$  inverze pontosan akkor függvény, ha az  $f$  bijektív, és ekkor*

$$(f^{-1})^{-1} = f, \quad (1.12)$$

$$f^{-1} \circ f = \text{id}_X, \quad (1.13)$$

$$f \circ f^{-1} = \text{id}_Y. \quad (1.14)$$

Ha egy  $f : X \rightarrow Y$  függvény inverzéről, inverz függvényéről vagy invertálhatóságáról beszélünk, akkor — hacsak külön mást nem mondunk — mindig arra gondolunk, hogy bijektív leképezésről van szó, és ekkor az  $f^{-1} : Y \rightarrow X$  az inverzfüggvényt jelöli, ami szintén bijekció.

**Bizonyítás.** Az  $f$  reláció

$$f^{-1} = \{(f(x), x) : x \in X\} \quad (1.15)$$

inverze pontosan akkor egy  $Y$ -ből  $X$ -be képező függvény — ha a függvény definíciójának megfelelően — két feltételt teljesít:

1) Az értelmezési tartománya az  $Y$  halmaz. Ez pedig az (1.15) szerint azzal ekvivalens, hogy  $\{f(x) : x \in X\} = Y$ , azaz hogy az  $f$  szurjektív.

2) Az  $f^{-1}$  ben lévő elempárok első eleméhez egyértelműen van hozzárendelve a második elem: az  $y = f(x)$  elemnek egyetlen párja van, azaz csak az  $x$ , tehát:  $f^{-1}(f(x)) = x$ . Ez pedig pontosan az  $f$  injektivitása.

Tehát: az  $f$  szurjektív és injektív, azaz bijektív.

(1.12): A definícióból nyilvánvaló.

(1.13): Az előbb láttuk, hogy  $f^{-1}(f(x)) = x$ , ezért

$$(f^{-1} \circ f)(x) = f^{-1}(f(x)) = x = \text{id}_X(x).$$

(1.14): Legyen az  $y = f(x)$ . Ekkor

$$(f \circ f^{-1})(y) = f(f^{-1}(y)) = f(x) = y = \text{id}_Y.$$

□

**Állítás 32 (Az  $(X \rightarrow X)$  bijekciók struktúrája)** Az  $(X \rightarrow X)$  bijekciók összesége a leképezések kompozíció műveletével a következőket teljesíti.

- (1)  $(X \rightarrow X)$  bijekciók kompozíciója is  $(X \rightarrow X)$  bijekció.
- (2) A kompozíció asszociatív.
- (3) Van reprodukáló elem:  $f \circ \text{id}_X = \text{id}_X \circ f = f$ .
- (4) Minden elemnek van inverze:  $f \circ f^{-1} = \text{id}_X$ , ahol az  $f^{-1}$  az  $f$  inverz függvénye.

**Bizonyítás.** (1): Legyen az  $f, g : X \rightarrow X$  bijekció. Ekkor a  $g \circ f$  szurjektív: Tetszőleges  $x \in X$  elemhez van olyan  $u \in X$ , hogy  $x = g(u)$ , és az  $u$  elemhez van olyan  $v$ , hogy  $u = f(v)$ , ezért

$$(g \circ f)(v) = g(f(v)) = g(u) = x.$$

Vegyük észre, hogy azt láttuk be, hogy szurjektív leképezések szorzata is szurjektív. A  $g \circ f$  leképezés injektivitása: Ha  $(g \circ f)(u) = (g \circ f)(v)$ , azaz  $g(f(u)) = g(f(v))$ , akkor a  $g$  injektivitásából:  $f(u) = f(v)$ , az  $f$  injektivitásából pedig:  $u = v$ .

- (2): A függvények kompozíciója a 30. tétel szerint asszociatív.
- (3): A 31. tételből adódik.
- (4): A 19. tételből adódik.

□

A bijekciók kompozíciója általában nem kommutatív, ahogyan azt a következő példa is mutatja:

Legyen  $X = \mathbb{R}_+$ , és

$$f(x) = 2x, \quad g(x) = x^2, \quad x \in \mathbb{R}_+.$$

Azonnal adódik, hogy  $(g \circ f)(x) = g(f(x)) = 4x^2$  és  $(f \circ g)(x) = f(g(x)) = 2x^2$ .

### 1.5.3 A direkt- és inverz-kép leképezés

A következő definícióban egy adott leképezéshez két leképezést vezetünk be, amelyeknek döntő szerepe lesz a függvények vizsgálatában.

**Definíció 33 (Direkt- és inverz-kép)** Legyen az  $f : X \rightarrow Y$  egy tetszőleges leképezés.

- (1) Az  $f$  leképezéshez tartozó direkt-kép leképezésnek mondjuk az

$$f(A) \doteq \{y \in Y : x \in A \text{ és } y = f(x)\}, \quad \text{ahol } A \subseteq X$$

módon definiált  $\mathcal{P}(X) \rightarrow \mathcal{P}(Y)$  függvényt.

- (2) Az  $f$  leképezéshez tartozó inverz-kép leképezésnek nevezzük az

$$f^{-1}(B) \doteq \{x \in X : f(x) \in B\}, \quad \text{ahol } B \subseteq Y$$

módon definiált  $\mathcal{P}(Y) \rightarrow \mathcal{P}(X)$  függvényt.

Jelölési megállapodás, hogy az  $f^{-1}(\{x\})$  helyett  $f^{-1}(x)$  is írható.

Az  $X$  halmaz  $f(X) \subseteq Y$  direkt képe már előkerült, és az  $f$  leképezés érték-készletének neveztük (25. definíció)

Azonnal észrevehetjük, hogy a jelölés meglehetősen következtelen, hiszen az “ $f$ ” szimbólum elsődlegesen jelöli az  $f : X \rightarrow Y$  függvényt, most pedig a direkt-kép leképezést jelöltük így, ami pedig a részhalmazok rendszeréről képez:  $f : \mathcal{P}(X) \rightarrow \mathcal{P}(Y)$ . Ez kétségtelenül helytelen jelölési szokás, de a szöveg-összefüggésből mindig kiderül, hogy mit is jelöl az “ $f$ ”, és így nem okoz félreértést. Ez a probléma pontosan ugyanígy fennáll az  $f^{-1}$  jelöléssel kapcsolatban is.

A direkt-kép és inverz-kép leképezések részhalmazokhoz részhalmazokat rendelnek. A részhalmazok összességét a halmazelméleti műveletek “strukturálják”, ezért természetes megnézni azt: hogyan viselkednek direkt- és inverz-kép leképezések a halmazelméleti műveletekkel szemben. Ezt fogalmazzuk meg a következő két tételben.

**Állítás 34 (Direkt-kép és a halmazműveletek)** Legyen az  $f : X \rightarrow Y$  egy tetszőleges leképezés, és  $A, B \subseteq X$ . Ekkor igazak a következő állítások.

$$A \subseteq B \implies f(A) \subseteq f(B). \quad (1.16)$$

$$f(A \cup B) = f(A) \cup f(B). \quad (1.17)$$

$$f(A \cap B) \subseteq f(A) \cap f(B), \quad \text{egyenlőség van, ha az } f \text{ injektív} \quad (1.18)$$

A direktkép leképezés általában nem őrzi meg a metszetet, amint azt a következő példa is mutatja: Legyen  $X = \{x_1, x_2\}$ ,  $Y = \{y\}$ ,  $f(x_1) = f(x_2) = y$ ,  $A = \{x_1\}$  és  $B = \{x_2\}$ . Ekkor

$$f(A \cap B) = f(\emptyset) = \emptyset \subset f(A) \cap f(B) = \{y\}.$$

**Bizonyítás.** (1.16): Nyilvánvaló.

(1.17): A következő ekvivalencia sorozatból adódik:

$$\begin{aligned} y \in f(A \cup B) &\Leftrightarrow \exists x \in A \cup B \quad y = f(x) \Leftrightarrow \exists x \in A \text{ vagy } x \in B, \quad y = f(x) \Leftrightarrow \\ &\Leftrightarrow y \in f(A) \text{ vagy } y \in f(B) \Leftrightarrow y \in f(A) \cup f(B). \end{aligned}$$

(1.18): A következő implikáció sorozatból:

$$\begin{aligned} y \in f(A \cap B) &\Leftrightarrow \exists x \quad x \in A \cap B, y = f(x) \Leftrightarrow \exists x \quad x \in A \text{ és } x \in B, y = f(x) \Rightarrow \\ &\Rightarrow y \in f(A) \text{ és } y \in f(B) \Leftrightarrow y \in f(A) \cap f(B). \end{aligned}$$

Ha az  $f$  injektív, akkor fennáll a másik irányú tartalmazás is: Ha  $y \in f(A) \cap f(B)$ , akkor

$$\exists a \in A \quad y = f(a) \quad \text{és} \quad \exists b \in B \quad y = f(b).$$

Mivel az  $f$  injektív, ezért  $a = b \in A \cap B$ , és így  $y = f(a) = f(b) \in f(A \cap B)$ .  $\square$

**Állítás 35 (Inverzkép és a halmazműveletek)** *Legyen az  $f : X \rightarrow Y$  egy tetszőleges leképezés. Ekkor az  $Y$  tetszőleges  $A$  és  $B$  részhalmazaira igazak a következő állítások.*

$$A \subseteq B \implies f^{-1}(A) \subseteq f^{-1}(B). \quad (1.19)$$

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B). \quad (1.20)$$

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B). \quad (1.21)$$

$$f^{-1}(A^c) = (f^{-1}(A))^c. \quad (1.22)$$

$$f^{-1}(Y) = X \quad \text{és} \quad f^{-1}(\emptyset) = \emptyset. \quad (1.23)$$

$$f(f^{-1}(A)) \subseteq A. \quad (1.24)$$

$$f^{-1}(f(C)) \supseteq C, \quad C \subseteq X. \quad (1.25)$$

Az (1.19) implikáció szavakban: az inverzkép leképezés monoton a halmazok tartalmazására nézve.

Az (1.23)–(1.22) azonosságok szerint az inverzkép leképezés őrzi a metszet, unió műveleteket és a komplementer képzést, abban az értelemben, hogy például metszet képe megegyezik a képek metszetével. Ilyen esetben azt szokás mondani, hogy az  $f^{-1} : \mathcal{P}(Y) \rightarrow \mathcal{P}(X)$  leképezés *morfizmus* a  $\mathcal{P}(Y)$  Boole-algebráról a  $\mathcal{P}(X)$  Boole-algebrába.

Az (1.24) és (1.25) szerint a direktkép és inverzkép leképezések nem inverzei ugyan egymásnak, de fennállnak a leírt tartalmazások.

**Bizonyítás.** Az (1.19) és (1.23) állítások nyilvánvalóak.

(1.20): A következő ekvivalencia sorozatból:

$$\begin{aligned} x \in f^{-1}(A \cup B) &\Leftrightarrow f(x) \in A \cup B \Leftrightarrow f(x) \in A \text{ vagy } f(x) \in B \Leftrightarrow \\ &\Leftrightarrow x \in f^{-1}(A) \text{ vagy } x \in f^{-1}(B) \Leftrightarrow x \in f^{-1}(A) \cup f^{-1}(B). \end{aligned}$$

(1.21): A következő ekvivalencia sorozatból:

$$\begin{aligned} x \in f^{-1}(A \cap B) &\Leftrightarrow f(x) \in A \cap B \Leftrightarrow f(x) \in A \text{ és } f(x) \in B \\ &\Leftrightarrow x \in f^{-1}(A) \text{ és } x \in f^{-1}(B) \Leftrightarrow x \in f^{-1}(A) \cap f^{-1}(B). \end{aligned}$$

(1.22): Legyen  $A \in \mathcal{P}(Y)$ . Ekkor

$$\begin{aligned} x \in f^{-1}(A^c) &\Leftrightarrow f(x) \in A^c \Leftrightarrow f(x) \notin A \\ &\Leftrightarrow x \notin f^{-1}(A) \Leftrightarrow x \in (f^{-1}(A))^c. \end{aligned}$$

Más indoklás: Az (1.23)–(1.21) felhasználásával:

$$X = f^{-1}(Y) = f^{-1}(A \cup A^c) = f^{-1}(A) \cup f^{-1}(A^c)$$

$$\emptyset = f^{-1}(\emptyset) = f^{-1}(A \cap A^c) = f^{-1}(A) \cap f^{-1}(A^c),$$

amelyekből adódik:  $(f^{-1}(A))^c = f^{-1}(A^c)$ .

(1.24): Legyen  $A \in \mathcal{P}(Y)$ . Ekkor

$$y \in f(f^{-1}(A)) \Leftrightarrow \exists x \in f^{-1}(A) \ y = f(x) \Leftrightarrow f(x) \in A \text{ és } y = f(x) \Rightarrow y \in A.$$

(1.25): Legyen  $C \in \mathcal{P}(X)$ . Ekkor  $x \in C \Rightarrow f(x) \in f(C) \Leftrightarrow x \in f^{-1}(f(C))$ .  $\square$

Az inverzkép művelet-örző tulajdonságai tetszőleges halmaz összesség mellett is fennmaradnak:

**Állítás 36 (Inverzkép tulajdonságok)** *Legyen  $f : X \rightarrow Y$  és az  $A_\gamma$ ,  $\gamma \in \Gamma$  az  $Y$  részhalmazainak egy családja. Ekkor fennállnak az alábbiak.*

$$\begin{aligned} f^{-1}\left(\bigcup_{\gamma \in \Gamma} A_\gamma\right) &= \bigcup_{\gamma \in \Gamma} f^{-1}(A_\gamma). \\ f^{-1}\left(\bigcap_{\gamma \in \Gamma} A_\gamma\right) &= \bigcap_{\gamma \in \Gamma} f^{-1}(A_\gamma). \end{aligned}$$

Az igazolás ugyanaz, mint két halmaz esetében. Az unió és metszet műveletek is ugyanígy kiterjeszthetők tetszőleges halmazrendszerekre.

## 1.6 A számosságok

A következőkben a halmazok “elemszámának” a megadására szeretnénk fogalmat alkotni. Ha két  $X$  és  $Y$  halmaz között egy  $h$  bijekció van, akkor úgy képzelhetjük el, hogy a két halmaz elemei “párba vannak állítva”, abban az értelemben, hogy egy  $x$  elemnek a párja az  $y = h(x)$ , az  $y$  párja pedig az  $h^{-1}(y)$ . Emiatt természetes elvárásunk: egymással bijektív kapcsolatban lévő halmazok elemeinek a “számát” vegyük azonosnak. A következő definícióban ennek megfelelő fogalmakat vezetünk be:

**Definíció 37 (Ekvipotencia, számosság)** *Ha két  $X$  és  $Y$  halmaz között létezik bijektív leképezés, akkor a két halmazt ekvipotensnek fogjuk nevezni. Ha az  $X$  és  $Y$  ekvipotensek, akkor azt is szokás mondani: azonos számosságúak. Ezt a tényt formálisan a  $\text{card}(X) = \text{card}(Y)$  módon fejezzük ki.*

Felhívjuk a figyelmet arra, hogy az “elemszámot” nem definiáltuk, csak azt mondtuk meg, hogy két halmazt mikor tekintünk elemszám szempontjából azonosnak.

Az elemszám fogalomtól elvárjuk: Az  $X$  számossága ugyanaz legyen, mint az  $X$  számossága. Ez valóban igaz, mivel az  $X$  és  $X$  között van bijekció: az  $\text{id}_X : X \rightarrow X$  identitás leképezés.

Szintén elvárás: Ha az  $A$  ugyanolyan számosságú, mint  $B$ , a  $B$  pedig ugyanolyan számosságú, mint a  $C$ , akkor az  $A$  is ugyanolyan számosságú, mint a  $C$ . Ez is teljesül, mert ha az  $f : A \rightarrow B$  és  $g : B \rightarrow C$  bijekciók, akkor — mivel a 32. tétel szerint két bijekció kompozíciója is bijekció — az  $g \circ f$  leképezés bijekció az  $A$  és  $C$  között.

A “számosság” szót azért használtuk a “szám” helyett, mert nemcsak véges halmazokkal kívánunk foglalkozni. A definíció alapján most azt tesszük, hogy kiszemelünk bizonyos rögzített halmazokat, és az azokkal ekvipotens halmazok “számosságára” megfelelő elnevezéseket vezetünk be:

**Definíció 38 (Véges halmazok)** *Egy  $A$  halmaz  $n$  elemszámú, vagy  $n$ -elemű halmaz, ha ekvipotens az első  $n$  egész számból álló  $\{1, 2, \dots, n\}$  halmazzal. Az üres halmaz számosságát nullának vesszük.*

*Egy halmazt végesnek mondunk, ha az elemszáma nemnegatív egész szám. Végtelennek nevezünk egy halmazt, ha nem véges.*

**Definíció 39 (Megszámlálhatóan végtelen számosság)** *Egy  $A$  halmazt megszámlálhatóan végtelen számosságúnak mondunk, ha ekvipotens a természetes számok halmazával.*

Ha az  $\mathbb{N}$  és  $A$  halmazok között létezik egy bijektív leképezés, akkor — egymás alá írva az egymásnak megfelelő elemeket — így szemléltethetjük:

$$\begin{array}{cccccc} 1 & 2 & 3 & \dots & n & \dots \\ \downarrow \uparrow & \downarrow \uparrow & \downarrow \uparrow & \dots & \downarrow \uparrow & \dots \\ a_1 & a_2 & a_3 & \dots & a_n & \dots \end{array}$$

Ennek az alapján ezt mondhatjuk: *Egy  $A$  halmaz pontosan akkor megszámlálhatóan végtelen számosságú, ha az elemei egy végtelen sorozatba rendezhetőek.*

Nézzünk most egy — első pillantásra meglepő — példát. Vegyük a pozitív egészek  $\{1, 2, \dots, n, \dots\}$  és a páros számok  $\{2, 4, \dots, 2k, \dots\}$  halmazát. Az utóbbi halmaz része az előzőnek, szemléletesen szólva a páros számokat fele annyinak képzeljük, mint az összes egész számot. Ennek ellenére a két halmaz számossága

azonos. Könnyen meg tudunk adni ugyanis egy bijektív leképezést a két halmaz között: egy  $k$  pozitív egész számot a  $2k$  páros számnak feleltetünk meg, szemléltetve:

$$\begin{array}{cccccc} 1 & 2 & 3 & \dots & k & \dots \\ \downarrow \uparrow & \downarrow \uparrow & \downarrow \uparrow & \dots & \downarrow \uparrow & \dots \\ 2 & 4 & 6 & \dots & 2k & \dots \end{array}$$

Ez az észrevétel azért okoz meglepetést, mert azt várnánk, hogy “a valódi rész elemeinek a száma kisebb az egész elmeinek a számánál”, ami itt nyilvánvalóan nem teljesül. Véges elemszámú halmaznál viszont igaz az, hogy egy valódi részhalmaz elemszáma kisebb, ezért azt is mondhatnánk: *Egy halmaz pontosan akkor végtelen, ha van olyan valódi részhalmaza, amelyik vele azonos számosságú, vele ekvipotens.* Egy egyszerű szóhasználatot rögzítünk a következő definícióban:

**Definíció 40 (Megszámlálható halmaz)** *Egy halmazt megszámlálhatónak fogunk mondani, ha véges vagy megszámlálhatóan végtelen számosságú.*

Az elnevezés oka az, hogy egy véges vagy megszámlálhatóan végtelen halmaz “megszámlálható”, abban az értelemben, hogy az elemei egymás után “sorbaszédhetők” úgy, hogy minden eleméhez eljutunk, azaz felírható egy véges vagy végtelen sorozatba rendezve. Azt gondolhatnánk, hogy minden halmaz ilyen tulajdonságú, de a következő tétel szerint ez a vélekedés nem helyes.

**Állítás 41 (A valós számok nem megszámlálhatóak)** *A valós számok halmaza nem megszámlálható.*

Ezek szerint létezik “nagyobb” számosság, mint a természetes számok által meghatározott megszámlálhatóan végtelen, és megemlítjük, hogy a valós számok halmazával ekvipotens halmazokat *kontinuum számosságúnak* nevezik. Ez általános tudományos szempontból is nagyon érdekes állítás, de tanulmányaink során voltaképpen csak a megszámlálható számosságokat fogjuk használni.

**Bizonyítás.** Ha az  $\mathbb{R}$  halmaz sorozatba rendezhető lenne, akkor minden  $A$  részhalmaza is ugyanilyen lenne: a sorozatba rendezett  $\mathbb{R}$ -ből el kell hagyni azokat, amelyek nem szerepelnek az  $A$  részhalmazban. Emiatt elégséges azt megmutatnunk, hogy az  $\mathbb{R}$  valamilyen részhalmaza nem megszámlálható. Azt bizonyítjuk be, hogy a  $(0, 1]$  intervallumban lévő valós számok nem rendezhetők sorozatba. Ehhez szükségünk lesz a következő állításra:

*A  $(0, 1]$  intervallum minden  $x$  pontja egyértelműen írható fel*

$$0, x_1 x_2 \dots x_n \dots, \quad x_n \in \{0, 1, \dots, 9\} \quad n = 1, 2, \dots$$

*tizedestört alakba, ahol a számjegyek egy indextől kezdve nem mind nullák.*

Hangsúlyozni kell, hogy a valós számok és a tizedestörtek közötti kapcsolat nem bijektív, ha elhagyjuk azt a feltevést, hogy nem lehet a tizedes tört minden számjegye nulla egy bizonyos indextől kezdve. Például a

$$0.50000\dots \quad \text{és} \quad 0.49999\dots$$

tizedestörtek mindketten az  $1/2$ , valós számot reprezentálják. Az idézett állítás a valós számok tárgyalásában szerepel.

Most pedig belátjuk, hogy a  $(0, 1]$  intervallumban lévő valós számokat megadó tizedestörtek halmaza nem megszámlálható. Tegyük fel ezzel ellentétben, hogy a tizedes törtet sikerült egy sorozatba felírni:

$$\begin{array}{cccccc} a_{11} & a_{12} & a_{13} & \dots & a_{1n} & \dots \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

Megmutatjuk, hogy ebben a sorozatban nem szerepelhet az összes szóban forgó tizedes tört. Vegyük ehhez azokat a  $b_1, b_2, \dots, b_n, \dots$  elemeket, amelyekre

$$b_i = \begin{cases} 1 & \text{ha } a_{ii} \neq 1 \\ 2 & \text{ha } a_{ii} = 1 \end{cases}$$

A  $0, b_1 b_2 \dots b_n \dots$  tizedestört az általunk vizsgált összességben van, hiszen egyetlen számjegye sem nulla, de nincs benne a felírt sorozatban: a sorozat első tagjával az  $a_{11} \neq b_1$  miatt nem lehet azonos, a másodikkal  $a_{22} \neq b_2$  miatt, és így tovább.  $\square$

A következő tételbe a megszámlálhatóan végtelen halmazokkal kapcsolatos legfontosabb tudnivalókat gyűjtöttük össze:

**Állítás 42 (A megszámlálhatóan végtelen számosság tulajdonságai)**

- (1) *Megszámlálhatóan végtelen halmaznak minden végtelen részhalmaza is megszámlálhatóan végtelen számosságú.*
- (2) *Minden végtelen halmaznak van megszámlálhatóan végtelen részhalmaza.*
- (3) *Két megszámlálhatóan végtelen halmaz szorzata is megszámlálhatóan végtelen.*
- (4) *Megszámlálhatóan végtelen sok megszámlálhatóan végtelen halmaz uniója is megszámlálhatóan végtelen.*

**Bizonyítás.** (1): Legyen az  $X$  megszámlálhatóan végtelen, és a  $C$  egy végtelen részhalmaza. Ekkor az  $X$  egy végtelen sorozatba rendezhető, és ha elhagyjuk ebből azokat, amelyek nem elemei a  $C$  halmaznak, akkor a  $C$  egy végtelen sorozatba rendezve marad vissza.

(2): Legyen az  $Y$  egy végtelen halmaz. Teljes indukció segítségével veszünk ki belőle egy megszámlálhatóan végtelen részhalmazt. Vegyük először az  $Y$  egy tetszőleges  $y_1$  elemét, amit megtehetünk, mivel az  $Y$  nem üres. Ha már kivettünk egy  $n$ -elemű  $\{y_1, \dots, y_n\}$  részhalmazt, akkor az  $Y$  halmaz nem véges volta miatt az  $Y \setminus \{y_1, \dots, y_n\}$  halmaz nem üres, ezért vehetünk belőle egy  $y_{n+1}$  elemet. Így kivettük az  $\{y_1, \dots, y_n, y_{n+1}\}$  részhalmazt. Ezzel az eljárással egy

$$\{y_1, \dots, y_n, \dots\}$$

végtelen sorozatba rendezett — azaz megszámlálhatóan végtelen — részhalmazát sikerült vennünk az  $Y$ -nak.

(3): Elégséges azt belátnunk, hogy a természetes számok esetében igaz az állítás. Az  $\mathbb{N} \times \mathbb{N}$  szorzat a természetes számok számpárjaiból áll, és ezeket kell egy végtelen sorozatba rendeznünk. Ezt elérhetjük, ha egy négyzetes sémába írjuk fel a számpárokat, és megmondjuk, hogyan kell végigmenni rajtuk a sorozatba rendezéshez:

$$\begin{array}{cccccccc}
 (1, 1) & & (1, 2) & & (1, 3) & & (1, 4) & \dots & (1, n) & \dots \\
 & \swarrow & & \swarrow & & \swarrow & & & & \\
 (2, 1) & & (2, 2) & & (2, 3) & & (2, 4) & \dots & (2, n) & \dots \\
 & \swarrow & & \swarrow & & & & & & \\
 (3, 1) & & (3, 2) & & (3, 3) & & (3, 4) & \dots & (3, n) & \dots \\
 & \swarrow & & & & & & & & \\
 (4, 1) & & (4, 2) & & (4, 3) & & (4, 4) & \dots & (4, n) & \dots \\
 \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 (n, 1) & & (n, 2) & & (n, 3) & & (n, 4) & \dots & (n, n) & \dots \\
 \vdots & & \vdots & & \vdots & & \vdots & \vdots & \vdots & \vdots
 \end{array}$$

A berajzolt nyilaknak megfelelően megyünk végig az “átlókon”, és egymásután vesszük az átlókat:

$$(1, 1), \quad (1, 2), (2, 1), \quad (1, 3), (2, 2), (3, 1), \dots$$

Igy egy sorozatba tudjuk felírni a számpárokat.

Ha valaki idegenkedik a szemléletes elmondástól — ami teljesen jogos lehet — akkor pontossá tehető az eljárás a következőképpen: Vegyük az

$$f : (n, m) \mapsto \frac{(n+m)(n+m+1)}{2} + n$$

leképezést. Belátható, hogy az  $f$  bijekció az  $\mathbb{N}^2$  és az  $\mathbb{N}$  között.

(4): Legyenek a  $B_i = \{b_{i,1}, b_{i,2}, b_{i,3}, \dots, b_{i,n}, \dots\}$ ,  $i = 1, 2, \dots, m, \dots$  a megszámlálhatóan végtelen halmazok. Megmutatjuk, hogy az uniójuk is megszámlálhatóan végtelen. A  $B_i$  halmazokat írjuk fel a következő táblázatba:

$b_{1,1}$		$b_{1,2}$		$b_{1,3}$		$b_{1,4}$	$\dots$	$b_{1,n}$	$\dots$
	$\swarrow$		$\swarrow$		$\swarrow$				
$b_{2,1}$		$b_{2,2}$		$b_{2,3}$		$b_{2,4}$	$\dots$	$b_{2,n}$	$\dots$
	$\swarrow$		$\swarrow$						
$b_{3,1}$		$b_{3,2}$		$b_{3,3}$		$b_{3,4}$	$\dots$	$b_{3,n}$	$\dots$
	$\swarrow$								
$b_{4,1}$		$b_{4,2}$		$b_{4,3}$		$b_{4,4}$	$\dots$	$b_{4,n}$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		
$b_{m,1}$		$b_{m,2}$		$b_{m,3}$		$b_{m,4}$	$\dots$	$b_{m,n}$	$\dots$
$\vdots$		$\vdots$		$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$

Ezen a táblázaton ugyanúgy lehet végigmenni a sorbarendehez, ahogyan a (3) bizonyításában.  $\square$

## 2.

# A számfogalom és valós függvények

Az előszóban mondottak szerint ez a fejezet is elsősorban ismétlő és összefoglaló jellegű. Nem elégszünk meg egyszerű ismétléssel, hanem magasabb szempontú visszatekintést szeretnénk adni, anélkül azonban, hogy a vizsgálatok részleteiben elmerülnénk. A fogalmi rendszerünk lesz voltaképpen újszerűbb, és azt lehetne mondani, hogy már meglévő ismereteinket töltjük át egy új, általánosabb keretbe.

## 2.1 A számfogalom felépítése

Ebben a pontban a számfogalmat szeretnénk olyan formában elismételni, hogy közben rövid kitekintést nyújtsunk a modern strukturális szemlélet irányába.

Amint látni fogjuk, a számfogalom bővítését az egyre bonyolultabb problémák matematikai modellezésének igénye tette szükségessé.

### 2.1.1 Természetes, egész és racionális számok

A legnyilvánvalóbb fogalmaink közé tartoznak az  $1, 2, \dots, n, \dots$  pozitív egész számok, ezért szokás ezeket “természetes számoknak” is nevezni. A fogalom kialakulását nyilvánvalóan a számlálás igénye segítette. Megjegyezzük, hogy a természetes számok  $\mathbb{N} = \{1, 2, \dots, n, \dots\}$  halmazát is röviden csak “természetes számoknak”

fogjuk mondani, az általános szokásnak megfelelően. A természetes számok halmazán két művelet is értelmezett, az összeadás és a szorzás. Mindkét művelet kommutatív és asszociatív, a szorzásra nézve az 1 természetes szám a reprodukáló elem szerepét játsza, amin azt értjük, hogy bármely  $n \in \mathbb{N}$ -re  $1 \cdot n = n \cdot 1 = n$ . A szorzás az összeadásra nézve disztributív, azaz minden  $m, n, p \in \mathbb{N}$ -re  $m \cdot (n + p) = m \cdot n + m \cdot p$ . A természetes számok halmaza szűknek bizonyul, amint felmerül a természetes számok kivonásának igénye. Ezért a természetes számok halmazát ki kellett bővítenünk, hozzávéve azokhoz a 0 és a negatív egész számokat. Így kaptuk az egész számok  $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$  halmazát. Az egészek halmazán is értelmezett az összeadás és szorzás művelet, azok megtartják kommutatív és asszociatív tulajdonságukat, de már az összeadásra nézve is van reprodukáló elem: a 0, és bármely két egész szám különbsége is az egészek halmazában van. Szeretnénk felhívni az olvasó figyelmét arra, hogy a kivonás az összeadás úgynevezett inverz művelete, amin azt értjük, hogy bármely két  $m, n$  egész számra  $m - n$  az az egész szám, melyet az  $n$ -hez adva  $m$ -et kapunk. Az egészek szorzása is disztributív azok összeadására nézve. El tudjuk képzelni, hogy az egészek halmaza is hamarosan szűknek bizonyulhatott, hiszen már a következő egyszerű probléma megválaszolása sem lehetséges kizárólag egész számokat használva: ha a piacon 2 almáért 3 szilvát lehet cserélni, akkor mennyi almát ér 1 szilva. Kevésbé gyermeken fogalmazva, az egészek hányadosa nem fejezhető ki mindig egész számmal. A probléma megoldása az egészek halmazának további bővítése, hozzávéve ahhoz az összes olyan számot, amely előáll egészek hányadosaként. Így jutottunk a racionális számok  $\mathbb{Q}$  halmazához. Mivel bármely egész szám is megkapható egészek hányadosaként, ezért a racionális számok halmazát a következőképpen is megadhatjuk:

$$\mathbb{Q} = \left\{ \frac{m}{n} \mid m, n \in \mathbb{Z} \right\}.$$

A racionális számok halmazára kiterjesztett összeadás és szorzás művelet kommutatív, asszociatív, mindkét műveletre nézve létezik reprodukáló elem, nevezetesen a 0 és az 1 és mindkét művelet invertálható, abban az értelemben, hogy bármely  $r \in \mathbb{Q}$  racionális számhoz létezik olyan  $-r$ -rel jelölt racionális szám, hogy  $r + (-r) = 0$  és  $r \neq 0$  esetén olyan  $r^{-1}$ -gyel jelölt racionális szám is, hogy  $r \cdot r^{-1} = 1$  teljesül. A szorzás disztributív az összeadásra nézve.

Ha valamely halmazon két művelet: egy összeadás és egy szorzás van értelmezve és a műveletek rendelkeznek mindazokkal a tulajdonságokkal, mint a racionális számok összeadás és szorzás műveletei, akkor az így kapott algebrai struktúrát (halmaz és azon értelmezett műveletek) *testnek* nevezzük.

### 2.1.2 Valós számok

A racionális számok összessége sok szempontból kielégítő. Az összeadás és annak inverz művelete a kivonás korlátlanul elvégezhető. A szorzás szintén minden korlátozás nélkül elvégezhető, és az inverz műveleténél az osztásnál csak arra kell

ügyelni, hogy nullával nem lehet osztani. Mindezek ellenére már régen felvetődtek olyan problémák, amelyek túlmutatnak a racionális számok fogalmán.

Már a görögök is jól ismerték a következő, geometriai eredetű, szakaszhosszúság méréssel kapcsolatos nehézséget.

Ha egy olyan egyenlőszárú derékszögű háromszöget veszünk, amelyiknek mindkét befogója 1, akkor a Pithagorasz-tétel alapján, a  $c$  átfogójának a négyzete

$$c^2 = 1^2 + 1^2 = 2.$$

A nagy gondot az okozza (már abban az időben belátták), hogy ilyen racionális  $c$  szám nem létezik. Azt a pozitív  $c$  számot, aminek a négyzete kettő, középiskolai ismereteink szerint  $\sqrt{2}$ -vel jelöljük. A számfogalom eddig vázolt bővítési módszere azt sugallja, hogy nem kell mást tennünk, mint a racionális számok halmazát kiegészíteni olyan szimbólumokkal, amelyek a racionális számok gyökeit jelölik. Sajnos bonyolultabb a probléma, a racionális számok gyökeinek birtokában sem lehet minden távolságot az így kapott számokkal kifejezni. A nevezetes  $\pi$  szám például nem kapható meg racionális számból gyökvonás útján, és így a kör területe és kerülete nem válik mérhetővé a racionális számok és azok gyökeinek birtokában sem. A valós számok halmazának megkonstruálása a racionális számokból lényegében úgy történhet, hogy képezzük a racionális számok sorozatait és azok határértékeivel bővítjük a racionális számok testét.

Középiskolai tanulmányainkból jól ismert, hogy a valós számok halmazán értelmezett összeadás és szorzás műveletek is kommutatívak, asszociatívak, mindkét műveletre nézve van reprodukáló elem és mindkét művelet invertálható. A valós számok szorzása is disztributív azok összeadására nézve, azaz a valós számok struktúrája is test, csakúgy mint a racionális számok teste. A valós számok halmazának "gazdagabb" volta tehát másból ered. A valós számok halmazán értelmezett jólismert ( $\leq$  szimbólummal jelölt) rendezési reláció az, amely új tulajdonsággal rendelkezik a racionális számok halmazán értelmezett rendezési relációhoz képest. A különbség megvilágítása érdekében néhány fogalomra van szükségünk.

**Definíció 43** Egy  $T$  testet rendezettnek mondunk, ha értelmezve van rajta egy olyan " $\leq$ " módon jelölt rendezés (reflexív, tranzitív, antiszimmetrikus és teljes reláció), amely az algebrai test tulajdonságokkal a következő kapcsolatban van.

- (1) Az  $x \leq y$  relációból minden  $z$  esetében következik az  $x + z \leq y + z$  reláció. (Az összeadással konform a rendezés)
- (2) Az  $0 \leq x$  és  $0 \leq y$  relációkból következik a  $0 \leq xy$  reláció. (A szorzással konform a rendezés)

Ahhoz, hogy a felsőhatár tulajdonságú rendezett test fogalmát megad hassuk, szükségünk van az alábbi definíciókban rögzített elnevezésekre.

**Definíció 44** Legyen az  $S$  egy  $T$  rendezett testnek egy nem üres részhalmaza.

Egy  $k \in T$  felső korlátja az  $S$  halmaznak, ha az  $S$  minden  $s$  elemére  $s \leq k$ . Az  $S$  halmazt felülről korlátosnak nevezzük, ha van felső korlátja.

Egy  $h \in T$  elem felső határa az  $S$  halmaznak, ha

- (1) a  $h$  felső korlát,
- (2) a  $h$  a legkisebb felső korlát, azaz ha a  $k$  egy tetszőleges felső korlátja az  $S$  halmaznak, akkor  $h \leq k$ .

A felső határ jelölésére a  $\sup S$  (az  $S$  szuprémuma) szimbólumot fogjuk használni.

Lássuk be, hogy a felső határ, ha van, egyetlen, amit a definícióban már fel is tételeztünk. Ha a  $h_1$  és  $h_2$  felső határa az  $S$  halmaznak, akkor mivel mindkettő legkisebb felső korlát, ezért teljesülnek a

$$h_1 \leq h_2 \quad \text{és} \quad h_2 \leq h_1$$

egyenlőtlenségek, ezért  $\leq$  antiszimmetrikus volta miatt  $h_1 = h_2$ .

A következő definíció az előző megismétlése a felső helyett az alsó korlátokra.

**Definíció 45** Legyen az  $S$  egy  $T$  rendezett testnek egy nemüres részhalmaza.

Egy  $k \in T$  alsó korlátja az  $S$  halmaznak, ha az  $S$  minden  $s$  elemére  $s \geq k$ . Az  $S$  halmazt alulról korlátosnak nevezzük, ha van alsó korlátja.

Ha az  $S$  halmaz alulról is és felülről is korlátos, akkor korlátosnak mondjuk.

Egy  $h \in T$  elem alsó határa az  $S$  halmaznak, ha

- (1) a  $h$  alsó korlát,
- (2) a  $h$  a legnagyobb alsó korlát, azaz ha a  $k$  egy tetszőleges alsó korlátja az  $S$  halmaznak, akkor  $h \geq k$ .

A alsó határ jelölésére a  $\inf S$  (az  $S$  infimuma) szimbólumot fogjuk használni.

Természetesen, ha létezik egy halmaznak alsó határa, akkor az is egyértelmű.

Ha egy  $S$  halmaznak van felső határa és az eleme a halmaznak, akkor azt a halmaz maximális elemének mondjuk és a  $\max S$  módon jelöljük, azaz ezen esetben  $\sup S = \max S$ . Hasonló a halmaz minimumának a definíciója.

A bevezetett fogalmakat a racionális számok rendezett halmazán illusztráljuk. Az egész számokat szépen szemléltethetjük egy egyenlőközű kétirányban végtelen skálán, és közbe berajzolva képzelhetjük a többi racionális számot is, nagyságuknak megfelelően. Így kapjuk meg a racionális számok "száamegyenesét".

**Példa 2.1** Vizsgáljuk meg a negatív egészek  $\{-1, -2, \dots, -n, \dots\}$  halmazát.

Felülről korlátos, felső korlát például a nulla. Alulról viszont nem korlátos, mivel nincs olyan racionális szám, ami kisebb lenne mindegyik negatív egész számnál. Ez a számegyeneset szemlélve eléggé nyilvánvaló.

Nézzük most a felső korlátok összességét. Ha egy szám felső korlát, akkor minden nála nem kisebb is az, ezért a felső korlátok halmaza egy félegyenes, a halmaztól jobbra a számegyenesen. A jelen példánkban azonnal láthatjuk, hogy a felső korlátok halmaza az

$$\{x : x \text{ racionális} \mid -1 \leq x\}$$

félegyenes, aminek van legkisebb pontja, a  $(-1)$ , tehát van felső határ is, ami — eleme lévén a halmaznak — egyben maximum értéke is.  $\square$

**Példa 2.2** *Nézzük meg a racionális számok  $\{x : 0 < x < 1\}$  és  $\{x : 0 < x \leq 1\}$  részhalmazait.*

Az első halmaznál a felső korlátok összessége az  $\{x : 1 \leq x\}$  halmaz, aminek van legkisebb eleme, azaz felső határ, mégpedig az 1 racionális szám. Ez a szám azonban nem eleme a halmaznak, tehát nem maximum.

Ha a második,  $\{x : 0 < x \leq 1\}$  halmazt vesszük, akkor már maximum is lesz a felső határ, hiszen

$$\sup\{x : 0 < x \leq 1\} = \max\{x : 0 < x \leq 1\} = 1.$$

$\square$

**Példa 2.3** *Vegyük azoknak a nemnegatív racionális számoknak az  $K$  halmazát, amelyek négyzete kisebb, mint 2. Van-e a  $K$  halmaznak felső határa a racionális számok körében?*

Mivel a nemnegatív számok körében kisebb szám négyzete kisebb és nagyobbé nagyobb, ezért a  $K$  halmaz a 0 és 2 között helyezkedik el, hiszen  $2^2 = 4 > 2$ . A felső korlátok halmaza — az előző példában mondottak szerint — egy félegyenes jobbra a racionális számegyenesen, ami azokból a pozitív racionális számokból áll, amelyeknek a négyzete nem kisebb a 2-nél.

A  $K$  halmaznak nincs felső határa a racionális számok körében. Ugyanis minden olyan nemnegatív racionális számhoz, amelynek a négyzete nem kisebb mint 2 lehet találni olyan nálánál kisebb nemnegatív racionális számot, amelynek négyzete nem kisebb mint 2.

Geometriailag azonban tudjuk, hogy hol helyezkedik el a számegyenesen az a “szám”, ami felső határ lehetne, de nem az, mert a pont elején mondottak szerint nem racionális szám. Szemléletesen szólva azt tapasztaltuk most, hogy a racionális számegyenes “lyukas”.  $\square$

Az utóbbi példánk jól érzékelteti, hogy milyen hiányossága van a racionális számoknak. Ezt a hézagot fogjuk egy újabb követelmény megfogalmazásával

megszüntetni, ami a valós számok fogalmához vezet. Az előzőekben bevezetett fogalmakkal könnyen meg tudjuk mondani, hogy mit követelünk meg a valós számoknak nevezett bővebb számfogalomtól.

**Definíció 46** Egy  $T$  rendezett testet a rendezésre nézve teljesnek nevezzük, ha teljesül az alábbi u. n. "felső határ tulajdonság"

*Minden, nem-üres felülről korlátos halmaznak van felső határa.*

Bebizonyítható, hogy lényegében egyetlen olyan rendezett test létezik, amely a rendezésre nézve teljes.

Ezekután a valós számok axiomatikus bevezetése a következő:

*Ha egy rendezett test a rendezésre nézve teljes, akkor a valós számok testének nevezzük.*

A valós számok halmazát  $\mathbb{R}$ -rel fogjuk jelölni. Időnként szükségünk lesz a nemnegatív valós számok halmazának  $\mathbb{R}_+$  és a pozitív valós számok halmazának  $\mathbb{R}_+^\circ$  jelölésére is.

Belátható, hogy tetszőleges pozitív valós számnak létezik akármilyen gyöke, és ez megoldja a pont elején felvetett szakasz-mérési problémát, de ennek a bizonyítását most mellőzzük.

A racionális számoknak van egy igen fontos tulajdonsága az  $\mathbb{R}$  valós számok között:

*Tetszőleges  $x < y$  valós számhoz van olyan  $r$  racionális szám, hogy*

$$x < r < y.$$

Másképpen fogalmazva: két tetszőleges (különböző) valós szám között van racionális szám. Ezt a tulajdonságot úgy is szokás mondani, hogy "a racionális számok sűrűn vannak a valós számok között".

A valós számok halmazának azon tulajdonsága, hogy a rendezésre nézve teljes két ugyancsak szemléletes tulajdonsággal ekvivalens:

**Archimédeszi tulajdonság** A valós számok  $\mathbb{R}$  halmaza archimédeszi módon rendezett, azaz tetszőleges  $x$  és  $y$  pozitív valós számokhoz van olyan  $n$  egész szám, hogy  $x < ny$ .

**Cantor-féle tulajdonság** A valós számok  $\mathbb{R}$  halmazában tetszőleges olyan

$$[a_n, b_n] \doteq \{x : a \leq x \leq b\}, \quad a_n \leq b_n \quad n = 1, 2, \dots$$

intervallum rendszer közös része nem üres, amelyek egymásba vannak skatulyázva, amin azt értjük, hogy

$$a_n \leq a_{n+1} \quad \text{és} \quad b_{n+1} \leq b_n, \quad n = 1, 2, \dots$$

Az archimédeszi tulajdonság igen szemléletes: Egy adott  $x$  pozitív számon túljuthatunk a számegyenesen, ha egy másik adott  $y$  pozitív számmal eléggé sok lépést teszünk meg. Természetesen, ha az  $y$  pozitív szám kicsi, az  $x$  pedig nagy, akkor sok  $y$  hosszúságú lépést kell megtennünk ahhoz, hogy túljussunk az  $x$ -en, azaz nagy a megkívánt  $n$  egész szám.

A Cantor-féle tulajdonság is eléggé elfogadható, mert azt mondja, hogy ha rendre egymásba rakunk intervallumokat, akkor a metszetük nem lesz üres. Ez valami olyasmit állít, hogy a valós számokat szemléltető számegyenes sehol sem "lyukas" .

### 2.1.3 Nevezetes azonosságok és egyenlőtlenségek

Ebben az alponthan először közöljük a teljes indukcióval való bizonyítási módszert, majd annak birtokában igazoljuk a számtani és mértani közép közötti egyenlőtlenséget és végül az úgynevezett Bernoulli-féle egyenlőtlenséget.

Az alábbi állítások a természetes számok halmazának azon egyszerű tulajdonságából következnek, hogy minden természetes számot felsorolunk, amennyiben az 1-gyel kezdjük a felsorolást és minden már megnevezett természetes szám rákövetkezőjét is felsoroljuk.

**Állítás 47 (Teljes indukció. I.)** *Ha egy olyan  $n$ -től függő  $T_n$  állításunk van (az  $n$  természetes szám), amelyik*

- (1) *igaz az 1 természetes számra,*
  - (2) *ha igaz az  $n$  természetes számra, akkor igaz a rákövetkező  $(n + 1)$  számra is,*
- akkor a  $T_n$  állítás minden természetes számra igaz.*

Az állításra alapozva, a teljes indukciós bizonyítás menete a következő:

**Kezdő lépés.** Ellenőrizzük azt, hogy az állítás igaz az 1 számra, azaz hogy a  $T_1$  igaz.

**Indukciós lépés.** Feltesszük, hogy igaz az állítás az  $n$  természetes szám esetében azaz a  $T_n$  igaz (indukciós feltevés), és ebből belátjuk, hogy igaz az  $(n + 1)$  esetében is, azaz a  $T_{n+1}$  is igaz.

Ezek után azt állíthatjuk, hogy igaz az állítás minden természetes számra.

Megjegyezzük, hogy némelykor nem az 1 a kezdő eset. Ha mondjuk egy  $k$  szám a kezdő szám, akkor első lépésként be kell látni az állítást a  $k$  esetére. Az indukciós lépés ugyanaz, csak az ott szereplő  $n$  szám nem lehet kisebb a kezdő  $k$  számnál.

A teljes indukció gyakorta a következő, az előzőből könnyen adódó tétel szerint történik.

**Állítás 48 (Teljes indukció. II.)** *Ha egy olyan  $n$ -től függő  $T_n$  állításunk van, amelyik*

- (1) igaz az 1 természetes számra,
- (2) ha igaz az  $n$  természetes számnál nem nagyobb minden természetes számra, akkor igaz a rákövetkező  $n + 1$  számra is,

akkor a szóbanforgó állítás igaz minden természetes számra.

**Példa 2.4** Bizonyítsuk be, hogy minden  $n$  pozitív egész számra igaz az

$$n < 2^n$$

egyenlőtlenség.

Az első lépés minden teljes indukciós bizonyításban az, hogy találunk kell egy olyan  $k$  természetes számot, amire igaz az állítás. A jelen esetben ez az 1 szám lehet, hiszen nyilvánvalóan  $1 < 2^1 = 2$ .

A második lépés az, hogy belássuk: Ha igaz az állításunk egy  $n$ , ( $\geq 1$ ) számra, akkor igaz az  $n + 1$  számra is. Tegyük fel tehát, hogy

$$n < 2^n.$$

Adjunk hozzá mindkét oldalhoz az 1 számot:

$$n + 1 < 2^n + 1.$$

Mivel  $1 < 2^n$  kisebb, ezért egyenlőtlenséget a következőképpen folytathatjuk

$$n + 1 < 2^n + 2^n = 2 \cdot 2^n = 2^{n+1},$$

és ezzel beláttuk az állítást az  $n + 1$  számra is. □

A következő példában azt mutatjuk meg, hogy a teljes indukció jó szolgálatot tesz a definíciókban is.

**Példa 2.5** Defináljuk egy  $x$  változó  $n$ -edik hatványát (az  $n$  természetes szám).

Az  $x$   $n$ -edik hatványának a pontos definíciója teljes indukcióval:

Az első hatvány legyen  $x^1 = x$ .

Az  $(n + 1)$ -edik hatványt az  $n$ -edik segítségével definiáljuk a következő módon:  
 $x^{n+1} \stackrel{\text{def}}{=} x \cdot x^n$ .

Indokolnunk kell, hogy ezzel minden  $n$  egész számra definiált a hatvány, de ez most csak két rövid mondat:

Az  $n = 1$  esetében definiált. Ha  $n$ -re definiált, akkor  $(n + 1)$ -re is definiált, ezért az első indukciós tétel miatt definiált minden természetes számra. □

A bemutatott definiálási módszert *rekurzív definíciónak* nevezik, és nagyon gyakran használjuk, legtöbbször anélkül hogy részleteznénk az indukciós bizonyítást.

Ugyancsak teljes indukcióval bizonyítjuk a számtani és mértani közép közötti egyenlőtlenséget:

**Tétel 49** *Legyenek az  $x_1, x_2, \dots, x_n$  nem-negatív számok. Ekkor igaz az alábbi egyenlőtlenség.*

$$x_1 x_2 \cdots x_n \leq \left( \frac{x_1 + x_2 + \cdots + x_n}{n} \right)^n.$$

*Egyenlőség pontosan akkor áll fenn, ha valamennyi  $x_i$  azonos.*

**Bizonyítás.** A teljes indukcióval folyó bizonyítás előtt néhány egyszerű észrevételt teszünk.

Az első: Ha

$$0 < a < x < b, \quad (2.1)$$

akkor

$$\frac{ab}{x} < a + b - x. \quad (2.2)$$

Indoklasképpen tekintsük az alábbi egyenlőséget

$$(b - x)(x - a) = (b - x)x - (b - x)a = (b - x)x - ba + xa = (b - x + a)x - ba$$

A baloldal (2.1) miatt pozitív, így

$$0 < (a + b - x)x - ab$$

ami éppen (2.2)-at jelenti.

A második észrevétel: A számtani közép nem nagyobb (nem kisebb), mint azon számok maximuma (minimuma), amiknek a számtani közepét vesszük, azaz

$$\min_{1 \leq i \leq n} x_i \leq \frac{x_1 + x_2 + \cdots + x_n}{n} \leq \max_{1 \leq i \leq n} x_i. \quad (2.3)$$

Az indoklás: Ha minden  $x_i$  helyett a nála nem kisebb (nem nagyobb)  $\max_{1 \leq i \leq n} x_i$  ( $\min_{1 \leq i \leq n} x_i$ ) mennyiséget tesszük, akkor azonnal adódik az egyenlőtlenség. Az is világos, hogy ha nem minden  $x_i$  ugyanaz a szám, akkor határozott egyenlőtlenség is írható.

Lássuk az indukció első lépését. Az  $n = 1$  esetben igaz az állítás, hiszen ekkor az  $x_1 \leq x_1$  egyenlőtlenségre egyszerűsödik.

Tegyük fel, hogy igaz az állítás  $n$ -re és lássuk be  $(n + 1)$ -re. Legyen az  $n + 1$  számú szám, nagyság szerint rendezve  $x_1 \leq x_2 \leq \dots \leq x_n \leq x_{n+1}$ , és tegyük fel, hogy nem mind azonosak. Ekkor ha bevezetjük az

$$A_n \doteq \frac{x_1 + x_2 + \cdots + x_n}{n}$$

jelölést, akkor (2.3) szerint,  $x_1 < A_{n+1} < x_{n+1}$ . Alkalmazhatjuk tehát az indukciós feltevést a következő  $n$  darab nemnegatív számra

$$(x_1 + x_{n+1} - A_{n+1}), x_2, \dots, x_n.$$

Így

$$\begin{aligned} (x_1 + x_{n+1} - A_{n+1}) x_2 \cdots x_n &\leq \left( \frac{x_1 + x_2 + \cdots + x_n + x_{n+1} - A_{n+1}}{n} \right)^n = \\ &= \left( \frac{(n+1) A_{n+1} - A_{n+1}}{n} \right)^n = A_{n+1}^n. \end{aligned}$$

De alkalmazva a (2.2) észrevételünket, azt kapjuk, hogy

$$\frac{x_1 x_{n+1}}{A_{n+1}} x_2 \cdots x_n < (x_1 + x_{n+1} - A_{n+1}) x_2 \cdots x_n \leq A_{n+1}^n$$

amiből már következik a bizonyítandó

$$x_1 x_2 \cdots x_n x_{n+1} < A_{n+1}^{n+1} = \left( \frac{x_1 + x_2 + \cdots + x_n + x_{n+1}}{n+1} \right)^{n+1}.$$

egyenlőtlenség. □

Végül lássuk be a Bernoulli-féle egyenlőtlenséget.

**Állítás 50 (Bernoulli-egyenlőtlenség)** Minden  $n$  pozitív egészre és bármely  $h > -1$  valós számra igaz az alábbi egyenlőtlenség.

$$(1+h)^n \geq 1 + n \cdot h.$$

**Bizonyítás.**  $n = 1$  esetén nyilvánvalóan egyenlőség formájában teljesül az egyenlőtlenség. Lássuk ezért be, hogy amennyiben  $n (\geq 1)$ -re igaz, akkor abból már következik, hogy  $n+1$ -re is igaz. Figyeljük meg, hogy a  $h > -1$  feltétel biztosítja, hogy  $1+h$ -val szorozva nem változik az egyenlőtlenség iránya.

$$(1+h)^{n+1} = (1+h)^n (1+h) \geq (1+n \cdot h)(1+h) = 1 + n \cdot h + h + n \cdot h^2 \geq 1 + (n+1) \cdot h,$$

ahol az utolsó lépésben a nem-negatív  $n \cdot h^2$  tag elhagyása nyilván csak csökkentheti a jobboldalt. Ezzel a bizonyítást befejeztük. □

### 2.1.4 Komplex számok

A kindulás itt is egy hiányérzet. Már a másodfokú egyenleteknél beleütköztünk abba problémába, hogy egy olyan egyszerű egyenletnek, mint az  $x^2 + 1 = 0$  nincs valós gyöke, hiszen a baloldala mindig legalább egy; vagy másképpen indokolva: nincs olyan valós szám aminek a négyzete a negatív,  $-1$  szám lenne.

Az algebrai egyenletek megoldhatóságának a lehetősége volt a közvetlen ösztönzője annak, hogy egy — a valós számok testénél bővebb — testet, a komplex

számokat bevezessék. Az új, bővebb számfogalom egyrészt harmonikus megoldást ad a felvetődött problémára, másrészt erre a számfogalomra alapozva nagy elméletek épültek fel (komplex függvénytan).

### A komplex számtest, normál alak

A valós számok bevezetésénél az indított el bennünket, hogy mérni akartuk az egyenes szakaszokat, számokat akartunk rendelni a számegyenes minden pontjához. Ehhez hasonlóan most azt szeretnénk elérni, hogy az  $\mathbb{R}^2$  sík pontjaival tudjunk úgy számolni, ahogyan egy testben lehet, azaz testet akarunk definiálni a valós számok rendezett párpárjainak az  $\mathbb{R}^2$  összességén.

**Állítás 51** *A valós számok  $\mathbb{R}^2$ , rendezett számpárjai testet adnak a következőképpen definiált összeadás és szorzás műveletekkel.*

$$(a, b) + (c, d) \stackrel{\text{def}}{=} (a + c, b + d), \quad (2.4)$$

$$(a, b) \cdot (c, d) \stackrel{\text{def}}{=} (ac - bd, ad + bc). \quad (2.5)$$

A test összeadásra nézve reprodukáló eleme (nulla eleme) a  $(0, 0)$ , a szorzás reprodukáló eleme (egység eleme) pedig az  $(1, 0)$ .

Egy  $(a, b)$  elem összeadási (additív) inverze (ellentettje) a  $(-a, -b)$ , és az  $(a, b) \neq (0, 0)$  esetben a szorzási (multiplikatív) inverze (reciproka):

$$(a, b)^{-1} = \left( \frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right). \quad (2.6)$$

Ezt a testet komplex számoknak fogjuk nevezni, és  $\mathbb{C}$ -vel jelöljük.

Hangsúlyoznunk kell, hogy az állításban lévő műveleti jelek mást jelölnek a számpárok közé téve, és mást a koordináták között. A számpárok között a komplex számok most definiált műveleteit jelölik, a koordináták között pedig a valós számok műveleteit.

**Bizonyítás.** Az könnyen látható, hogy az (2.4) összeadás kommutatív, asszociatív, reprodukáló elemes és invertálható, hiszen a párok komponensenként adódnak össze a valós számok összeadásának megfelelően. A (2.5) szorzás tulajdonságainak a belátása egyszerű, de részletezzük:

1) Kommutativitás:

$$(a, b)(c, d) = (ac - bd, ad + bc) \quad \text{és} \quad (c, d)(a, b) = (ca - db, cb + ad).$$

A baloldalokon lévő párok a valós számok összeadásának és szorzásának a kommutativitása miatt azonosak.

2) Asszociativitás:

$$\begin{aligned} [(a, b)(c, d)](e, f) &= (ac - bd, ad + bc)(e, f) = \\ &= ([ac - bd]e - [ad + bc]f, [ac - bd]f + [ad + bc]e) = \\ &= (ace - bde - adf - bcf, acf - bdf + ade + bce), \end{aligned}$$

és

$$\begin{aligned} (a, b)[(c, d)(e, f)] &= (a, b)(ce - df, cf + de) = \\ &= (a[ce - df] - b[cf + de], a[cf + de] + b[ce - df]) = \\ &= (ace - adf - bcf - bde, acf + ade + bce - bdf). \end{aligned}$$

3) Reprodukáló elem az  $(1, 0)$ :

$$(a, b)(1, 0) = (a \cdot 1 - b \cdot 0, a \cdot 0 + b \cdot 1) = (a, b).$$

4) Egy  $(a, b)$ , nem nulla pár inverze: Kétféle módon is eljárhatunk, vagy ellenőrizzük, hogy az (2.6) párral szorozva az  $(a, b)$  párt az  $(1, 0)$ , egységelemet kapjuk, vagy kiszámoljuk az inverzet. Az utóbbi utat választjuk. Olyan  $(x, y)$  elemet keresünk, amelyre

$$(a, b)(x, y) = (1, 0),$$

azaz

$$(ax - by, ay + bx) = (1, 0),$$

amiből

$$ax - by = 1 \quad \text{és} \quad ay + bx = 0.$$

Ezt az lineáris egyenletrendszert kell csak megoldanunk. Egyszerű számolással adódik, hogy

$$(a^2 + b^2)x = a \quad \text{és} \quad (a^2 + b^2)y = -b.$$

5) Disztributivitás:

$$\begin{aligned} (a, b)[(x, y) + (u, v)] &= (a, b)(x + u, y + v) = \\ &= (ax - by, ay + bx) + (au - bv, av + bu) = \\ &= (ax - yb + au - bv, ay + bx + av + bu) = \\ &= (a[x + u] - b[y + v], a[y + v] + b[x + u]) = (a, b)(x + u, y + v), \end{aligned}$$

és ezzel minden szükséges tulajdonságot beláttunk.  $\square$

**Példa 2.6** Számoljuk ki az  $(a, b)$  és  $(c, d) \neq (0, 0)$ , komplex számok hányadosát.

A reciprokra vonatkozó (2.6) formula szerint

$$\begin{aligned}\frac{(a,b)}{(c,d)} &= (a,b)(c,d)^{-1} = (a,b) \left( \frac{c}{c^2+d^2}, \frac{-d}{c^2+d^2} \right) = \\ &= \left( \frac{ac+bd}{a^2+b^2}, \frac{bc-ad}{c^2+d^2} \right).\end{aligned}\quad \square$$

A számfogalom kialakítása során az a vezérlő elv, hogy az általánosabb számkör valamilyen értelemben tartalmazza a kevésbé általános számokat. A komplex számokat a valós számok kiterjesztéseként szeretnénk felfogni, ezért meg kell vizsgálnunk, hogy milyen értelemben “része” a valós számok a komplex számoknak. Az  $\mathbb{R}^2$  síknak, amit alkalmas műveletekkel a komplex testté tettünk, úgy része a valós egyenes, hogy az  $(x, 0)$ ,  $x \in \mathbb{R}$  számpárokat tekinthetjük a valós egyenesnek, vagy helyesebben: a valós egyenes képének, a síkba való “behelyezésének” (“beágyazásának”).

Nézzük meg, hogy mit adnak a komplex számok közötti műveletek az  $(x, 0)$  alakú számpárok esetében:

$$(x, 0) + (u, 0) = (x + u, 0)$$

és

$$(x, 0)(y, 0) = (xy - 0 \cdot 0, x \cdot 0 + 0 \cdot y) = (xy, 0).$$

Ezek alapján megállapíthatjuk, hogy az  $(x, 0)$ ,  $x \in \mathbb{R}$  egyenesen úgy mennek végbe a műveletek, hogy az első koordinátával pontosan a valós számok közötti műveletek szerint járunk el, a másik koordináta pedig nulla. Szemléletesen szólva, ha letörölnénk a zárójeleket a nullát és az elválasztó vesszőt, akkor a valós számok műveleteihez jutnánk. Ezt a nagyon fontos gondolatot a következőképpen fogalmazzuk meg pontosan.

**Állítás 52** Az

$$x \mapsto (x, 0), \quad x \in \mathbb{R} \tag{2.7}$$

módon definiált  $\phi : \mathbb{R} \rightarrow \mathbb{C}$  leképezés injektív és művelet-tartó, abban az értelemben, hogy

$$\phi(u+v) = \phi(u) + \phi(v), \tag{2.8}$$

$$\phi(uv) = \phi(u) \cdot \phi(v). \tag{2.9}$$

A  $\phi$  leképezés az  $\mathbb{R}$  valós számokat a számpárok  $\{(x, 0) : x \in \mathbb{R}\}$  halmazára képezi, és az  $\mathbb{R}$  és a kép között a leképezés bijektív, és művelettartó. Ez azt jelenti, hogy az  $\mathbb{R}$  és  $\{(x, 0) : x \in \mathbb{R}\}$  halmazok a szóban forgó műveleteket tekintve, mint két algebrai struktúra, azonosak abban az értelemben, hogy ahogyan az egyikben mennek végbe a műveletek, ugyanúgy mennek végbe a másikon. Az ilyen  $\phi$  leképezést *izomorfíának* nevezik. Ilyen elnevezéssel azt mondhatjuk, hogy az  $\mathbb{R}$

izomorf képe része a  $\mathbb{C}$ -nek. Lazább szóhasználattal azt is szokás mondani, hogy az  $\mathbb{R}$  *izomorfia értelemben része (részhalmaza)* a  $\mathbb{C}$ -nek. Ha az elmondottaknak a tudatában vagyunk, akkor mondhatjuk azt, hogy egy  $r$  valós szám komplex szám is, de közben az  $(r, 0)$  párra kell gondolnunk.

Minden  $(a, b)$  komplex szám felírható a következőképpen

$$(a, b) = (a, 0) + (0, 1)(b, 0). \quad (2.10)$$

A  $(0, 1)$  elemmel egy gyökeresen új tulajdonság jelenik meg a komplex számtestben:

$$(0, 1)^2 = (0, 1)(0, 1) = (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = -(1, 0),$$

amit — ha figyelembe vesszük az izomorfjáról mondottakat — (egy kicsit lazábban) így is írhatunk

$$(0, 1)^2 = -1.$$

A  $(0, 1)$  elem tehát egy megoldása a

$$z^2 + (1, 0) = (0, 0)$$

(vagy lazábban:  $z^2 + 1 = 0$ ) egyenletnek.

Másik felépítését kapjuk a komplex számoknak a következő konstrukcióval. Vegyük az

$$a + ib$$

alakú szimbólumokat, és a műveleteket úgy definiáljuk, hogy az  $i$  szimbólumra feltesszük, hogy  $i^2 = -1$ , és szem előtt tartjuk a formális műveleti szabályokat:

$$(a + ib) + (c + id) \doteq a + c + i(b + d) \quad (2.11)$$

és

$$(a + ib)(c + id) \doteq ac + aid + ibc + i^2bd = ac - bd + i(ad + bc). \quad (2.12)$$

Azonnal látható, hogy az  $(a, b)$  és  $(c, d)$  számpárok és  $a + ib$  és  $c + id$  szimbólumok között mennyire azonosan történnek a műveletek. Az elmondottakat egy tételben foglaljuk össze.

**Állítás 53** Az  $a + ib$ ,  $a, b \in \mathbb{R}$  alakú szimbólumok összessége a (2.11) és (2.12) műveletekkel testet ad, amelyik izomorf a komplex számok  $\mathbb{C}$  testével az

$$(a, b) \mapsto a + ib.$$

izomorf leképezés mellett.

Az  $a + ib$  szimbólumot az  $(a, b)$  komplex szám normál-alakjának nevezzük.

A normál-alakokkal előnyösebb számolni, mint a számpárokkal, mert — a formális számolási szabályok betartása mellett — csak arra kell ügyelni, hogy  $i^2 = -1$ . Aggályosan pontos gondolkodás mellett, a normál-alakok teste csak izomorfia értelemben azonos a bevezetett komplex számokkal, de ettől a fölösleges fontoskodástól eltekintünk. A következő példák szerint a normál-alakkal való számolás több esetben is előnyös.

**Példa 2.7** Írjuk fel az

$$a + ib \quad \text{és} \quad c + id \neq 0 + i0$$

*hányados normál alakját.*

A 53. állítás szerint a normál-alakok összessége test, ezért úgy számolhatunk, ahogyan tesben szabad:

Az

$$\frac{a + ib}{c + id}$$

tört számlálóját és nevezőjét szorozzuk meg a  $c - id$  komplex számmal:

$$\begin{aligned} \frac{a + ib}{c + id} &= \frac{(a + ib)(c - id)}{(c + id)(c - id)} = \\ &= \frac{ac + bd + i(bc - ad)}{c^2 - (id)^2} = \frac{ac + bd + i(bc - ad)}{c^2 + d^2}. \end{aligned}$$

A számolásnál felhasználtuk a testekben fennálló  $(x - y)(x + y) = x^2 - y^2$  azonosságot.  $\square$

**Példa 2.8** Keressük meg azokat a komplex számokat, amelyeknek a négyzete  $i$ .

Olyan  $x + iy$  számot keresünk, amelyre  $(x + iy)^2 = i$ , azaz

$$x^2 - y^2 + 2xyi = i.$$

Ebből

$$x^2 - y^2 = 0 \quad \text{és} \quad 2xy = 1.$$

Az elsőből  $x^2 = y^2$  és ezt második négyzetébe téve  $4x^2x^2 = 4x^4 = 1$  adódik, amiből  $x^2 = 1/2$ , ezért az

$$x = \pm \frac{1}{\sqrt{2}} \quad \text{és} \quad y = \pm \frac{1}{\sqrt{2}}$$

értékek jöhetnek szóba. Az  $x$  és  $y$  a  $2xy = 1$  miatt egyező előjelű, és így a következő két komplex szám lehet megoldás:

$$\frac{1}{\sqrt{2}} + i \frac{1}{\sqrt{2}} \quad \text{és} \quad -\frac{1}{\sqrt{2}} - i \frac{1}{\sqrt{2}}.$$

Négyzetre emeléssel ellenőrizhetjük, hogy valóban megoldások.  $\square$

Talán úgy tűnhet most, hogy a normál-alakokkal könnyebb bevezetni a komplex számokat. Azért választottunk mégis eltérő módot, mert azt tisztább elindulásnak éreztük. Egy kicsit zavaró lehet ugyanis a normál-alakos bevezetésnél egy olyan objektumnak a kezdeti használata, aminek a négyzete  $-1$ . Az  $i$  számnak ez a különös viselkedése volt az oka annak, hogy a történelem során “lehetetlen” számnak is nevezték, és ma is imaginárius (képzetes, képzelt) egység a neve. Azt hitték, hogy a többi szám valahogyan valóságosabban létezik, a helyes szemlélet szerint azonban minden struktúra egyformán “valóságos”.

Eddig még nem is hasznosítottuk geometriailag azt, hogy az  $\mathbb{R}^2$  sík elemei között vezettünk be műveleteket, és így jó szemléltetésre van lehetőségünk. Ha az  $\mathbb{R}^2$  síkról ilyen értelemben beszélünk, akkor *komplex (szám)síknak* fogjuk mondani. Ez ugyanolyan szerepet tölt be a komplex számoknál, mint a valós egyenes a valós számoknál. Az  $(x, 0)$  számpárok összességét valós, az  $(0, y)$  alakú számpárok összességét pedig képzetes egyenesnek szokás nevezni.

A következő definícióban bevezetett elnevezéseket a 2.1. ábrán szemléltetjük.

**Definíció 54** Legyen az  $z = x + iy$  egy tetszőleges komplex szám. A  $z$  szám valós részének mondjuk az  $x$ , imaginárius részének pedig az  $y$  valós számot, jelölésben:

$$\operatorname{Re}(z) = \operatorname{Re}(x + iy) = x \quad \text{és} \quad \operatorname{Im}(z) = \operatorname{Im}(x + iy) = y.$$

A  $z$  konjugáltjának nevezzük — jelölésben:  $\bar{z}$  — az  $(x - iy)$  komplex számot.

**Állítás 55** A konjugálásnak a következő tulajdonságai vannak.

$$(1) \quad \overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2.$$

$$(2) \quad \overline{z_1 z_2} = \bar{z}_1 \cdot \bar{z}_2.$$

$$(3) \quad \text{Ha } z_2 \neq 0, \text{ akkor}$$

$$\overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}.$$

$$(4) \quad \bar{\bar{z}} = z.$$

$$(5) \quad z + \bar{z} = 2 \cdot \operatorname{Re}(z).$$

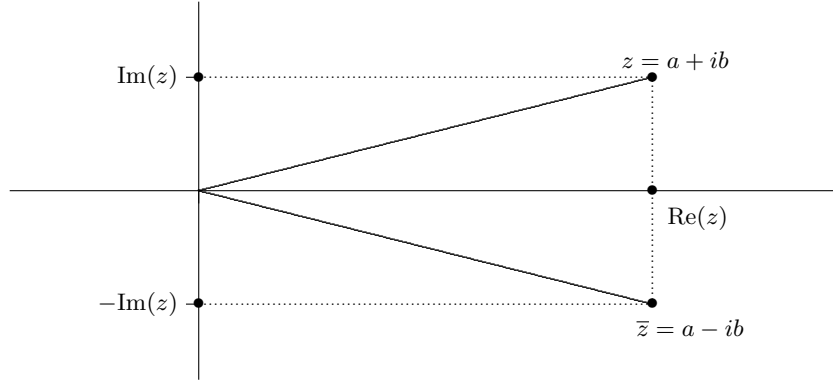
$$(6) \quad z = \bar{z} \iff z \in \mathbb{R}.$$

$$(7) \quad z - \bar{z} = 2i \cdot \operatorname{Im}(z).$$

**Bizonyítás.** Legyen a bizonyítások során  $z_1 = x_1 + iy_1$ ,  $z_2 = x_2 + iy_2$  és  $z = x + iy$ .

$$(1): \quad \overline{z_1 + z_2} = \overline{x_1 + x_2 + i(y_1 + y_2)} = (x_1 + x_2 - i(y_1 + y_2)) = (x_1 - iy_1) + (x_2 - iy_2) = \bar{z}_1 + \bar{z}_2.$$

$$(2): \quad \overline{z_1 z_2} = \overline{x_1 x_2 - y_1 y_2 + i(x_1 y_2 + x_2 y_1)} = \\ = x_1 x_2 - y_1 y_2 - i(x_1 y_2 + x_2 y_1) = (x_1 - iy_1)(x_2 - iy_2) = \bar{z}_1 \cdot \bar{z}_2.$$



2.1. ábra: Valós és képzetes részek, konjugált.

(3): Először azt lássuk be, hogy reciprokok konjugáltja a konjugált reciproka. Ez a következő számolásokból adódik:

$$\frac{1}{z} = \frac{1}{x + iy} = \frac{x - iy}{(x + iy)(x - iy)} = \frac{x - iy}{x^2 + y^2},$$

$$\frac{1}{\bar{z}} = \frac{1}{x - iy} = \frac{x + iy}{(x - iy)(x + iy)} = \frac{x + iy}{x^2 + y^2},$$

és ezekből már látszik, hogy

$$\overline{\left(\frac{1}{z}\right)} = \frac{1}{\bar{z}}.$$

A hányados konjugálása a szorzat és reciprokok konjugálási szabálya alapján:

$$\overline{\left(\frac{z_1}{z_2}\right)} = \overline{\left(z_1 \frac{1}{z_2}\right)} = \overline{z_1} \overline{\left(\frac{1}{z_2}\right)} = \overline{z_1} \frac{1}{\bar{z}_2} = \frac{\bar{z}_1}{\bar{z}_2}.$$

(4):  $\bar{\bar{z}} = \overline{x - iy} = \overline{x - iy} = x + iy = z.$

(5):  $z + \bar{z} = (x + iy) + (x - iy) = 2x = 2 \cdot \operatorname{Re}(z).$

(6): A  $\bar{z} = z$  azaz az  $x - iy = x + iy$  pontosan akkor áll fenn, ha  $2iy = 0$ , tehát ha a  $z$  valós.

(7):  $z - \bar{z} = (x + iy) - (x - iy) = 2iy = 2i \cdot \operatorname{Im}(z).$  □

**Definíció 56** Egy  $z = x + iy$  komplex szám abszolút értékének vagy hosszának mondjuk az

$$|z| \stackrel{\text{def}}{=} \sqrt{z\bar{z}} = \sqrt{(x + iy)(x - iy)} = \sqrt{x^2 + y^2}.$$

számot.

Az  $(x, y)$  komplex szám abszolút értéke megegyezik az  $(x, y)$  síkbeli pont origótól való távolságával. Ha egy  $z = (x, 0)$  valós komplex számot veszünk, akkor  $|z| = \sqrt{x^2} = |x|$ , ezért a komplex szám abszolútértéke a valós abszolút érték kiterjesztésének tekinthető. Az abszolút érték fontos tulajdonságait a következő állításban soroljuk fel.

**Állítás 57** Legyenek a  $z = x + iy$ ,  $z_1 = x_1 + iy_1$  és  $z_2 = x_2 + iy_2$  tetszőleges komplex számok.

$$(1) |z_1 z_2| = |z_1| |z_2|.$$

$$(2) |z| = 0 \iff z = 0.$$

$$(3) \operatorname{Re}(z_1 \bar{z}_2) \leq |z_1| \cdot |z_2|.$$

$$(4) |z_1 + z_2| \leq |z_1| + |z_2|.$$

A bizonyítás előtt két megjegyzés az állításokhoz:  
Figyelembe véve azt, hogy

$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + y_1 x_2),$$

az elsőből kapható,  $|z_1 z_2|^2 = |z_1|^2 |z_2|^2$  egyenlőség részletesen kiírva:

$$(x_1 x_2 - y_1 y_2)^2 + (x_1 y_2 + y_1 x_2)^2 = (x_1^2 + y_1^2)(x_2^2 + y_2^2).$$

A (3) egyenlőtlenség részletesen írva:

$$x_1 x_2 + y_1 y_2 \leq \sqrt{x_1^2 + x_2^2} \sqrt{y_1^2 + y_2^2}.$$

Ez az u. n. Cauchy-egyenlőtlenség. Az állítás negyedik egyenlőtlenségét háromszög egyenlőtlenségnek is szokás nevezni, mivel a 2.2. ábra szerint azt fejezi ki, hogy egy háromszög egyik oldalának a hossza nem lehet nagyobb, mint a másik két oldal hosszának az összege.

**Bizonyítás.**

(1): A szorzat konjugálási szabályának a felhasználásával:

$$|z_1 z_2|^2 = (z_1 z_2) \overline{z_1 z_2} = z_1 z_2 \bar{z}_1 \bar{z}_2 = (z_1 \bar{z}_1)(z_2 \bar{z}_2) = |z_1|^2 \cdot |z_2|^2.$$

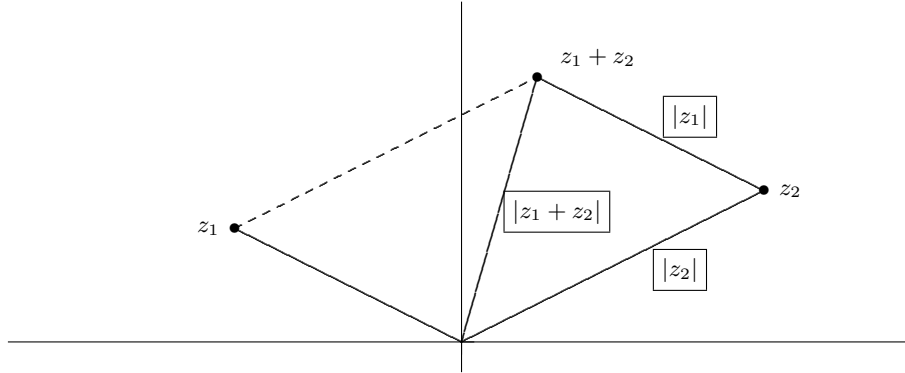
(2): Az  $|z|^2 = x^2 + y^2 = 0$  egyenlőség azzal ekvivalens, hogy  $x = 0$  és  $y = 0$ , azaz  $z = 0$ .

(3): Figyelembe véve azt, hogy

$$z_1 \bar{z}_2 = (x_1 + iy_1)(x_2 - iy_2) = (x_1 x_2 + y_1 y_2) + i(-x_1 y_2 + y_1 x_2),$$

a  $|z_1 \bar{z}_2|^2 = |z_1|^2 |z_2|^2$  egyenlőség (amely könnyen láthatóan igaz) részletesen kiírva:

$$(x_1 x_2 + y_1 y_2)^2 + (-x_1 y_2 + y_1 x_2)^2 = (x_1^2 + y_1^2)(x_2^2 + y_2^2).$$



2.2. ábra: Háromszög egyenlőtlenség.

A baloldali második, nemnegatív tagját elhagyva kapjuk:

$$(x_1 x_2 + y_1 y_2)^2 \leq (x_1^2 + y_1^2)(x_2^2 + y_2^2).$$

(4): A konjugálás szabályait alkalmazva:

$$\begin{aligned} |z_1 + z_2|^2 &= (z_1 + z_2)(\overline{z_1 + z_2}) = \\ &= z_1 \overline{z_1} + z_1 \overline{z_2} + \overline{z_1} z_2 + \overline{z_2} z_2 = |z_1|^2 + 2\operatorname{Re}(z_1 \overline{z_2}) + |z_2|^2. \end{aligned}$$

A jobboldali középső tagra a (3) egyenlőtlenséget alkalmazva:

$$|z_1 + z_2|^2 \leq |z_1|^2 + 2|z_1||z_2| + |z_2|^2 = (|z_1| + |z_2|)^2,$$

amiből azonnal adódik a háromszög egyenlőtlenség.  $\square$

Végezetül egy történelmi megjegyzés. Alig lehet elhinni, hogy az első elektronikus számítógép, amelyik elektromos jelfogókból épült fel, komplex számokkal számolt (1938–40, Bell Telephone Laboratories), és elektromos hálózatok vizsgálatára készült. Ennek az az oka, hogy az elektromosságtan elmélete és gyakorlata intenzíven támaszkodik a komplex számokra. Ez a számítógép azonban még nem volt, a mai értelemben programozható.

### Trigonometrikus alak, egységgyökök

A komplex számok számpárokkal való bevezetése és normál alakja szorosan kapcsolódik az  $\mathbb{R}^2$  sík pontjainak a Descartes-féle (derékszögű) koordinátákkal való megadásához. A sík pontjai azonban megadhatók u. n. polár koordinátákkal is. Egy  $(a, b) \neq (0, 0)$  pontot egyértelműen megad az origótól vett

$$r \stackrel{\text{def}}{=} \sqrt{a^2 + b^2} \quad (2.13)$$

távolsága és az  $(a, b)$  ponthoz mutató szakasznak az  $x$  tengely pozitív irányú részével bezárt  $\alpha$  szöge, amit a

$$\cos \alpha = \frac{a}{\sqrt{a^2 + b^2}} \quad (2.14)$$

egyenlőségből határozhatunk meg. Az origót már a távolság megadja, és szögről itt nem beszélhetünk, ezért ezt a pontot a továbbiakban kizárjuk a vizsgálatokból.

Megfordítva az  $(r, \alpha)$ ,  $r > 0$  polár koordináták alapján az  $(a, b)$  derékszögű koordinátákra való áttérés egyenletei:

$$a = r \cos \alpha \quad \text{és} \quad b = r \sin \alpha. \quad (2.15)$$

Az elmondottakat 2.3. ábrán szemléltetjük, és egy definícióban is rögzítjük:

**Definíció 58** Egy  $a + ib$  nem nulla komplex szám trigonometrikus alakjának mondjuk az

$$r(\cos \alpha + i \sin \alpha)$$

alakot, ahol az  $(r, \alpha)$  számok és az  $(a, b)$  számok közötti kapcsolatot az (2.13)–(2.15) formulák mutatják.

Ne feledjük, hogy — a szög mérésének a szabálya szerint — két szög pontosan akkor azonos, ha a mérőszámaik különbsége a  $2\pi$  egész számú többszöröse, ezért az  $(r_1, \alpha_1)$  és  $(r_2, \alpha_2)$  polár koordinátákkal megadott komplex számok pontosan akkor azonosak, ha

$$r_1 = r_2 \quad \text{és} \quad \alpha_1 - \alpha_2 = (\text{egész}) \cdot 2\pi.$$

Az  $\alpha$  szöget általában célszerű a  $[0, 2\pi)$  intervallumba esőnek választani.

A trigonometrikus alaknál természetesen nem a szögmérés egysége, hanem maga a “szög” a lényeges, ezért a szöget fokban is mérhetjük, és célszerűség szerint felváltva használhatjuk a két mértékegységet.

Most pedig megvizsgáljuk, hogy a trigonometrikus alak esetében hogyan végezhetők el a komplex számok közötti műveletek. Az összeadás esetében nem tudunk érdekeset mondani, de annál inkább a szorzás esetében:

**Állítás 59** Legyenek  $a$

$$z = r(\cos \alpha + i \sin \alpha)$$

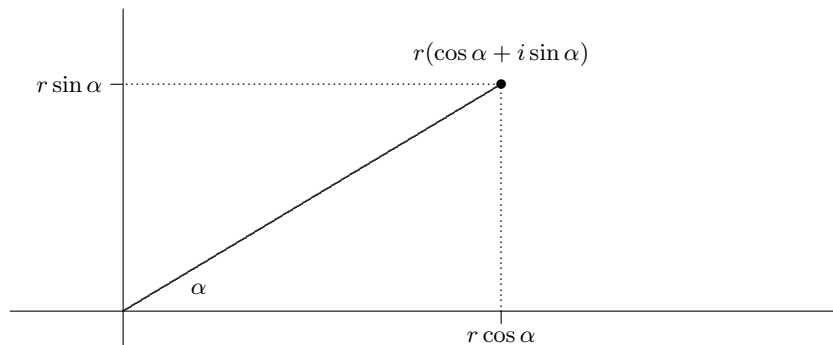
és

$$z_1 = r_1(\cos \alpha_1 + i \sin \alpha_1), \quad z_2 = r_2(\cos \alpha_2 + i \sin \alpha_2),$$

trigonometrikus alakban adott komplex számok. Ekkor

$$(1) \quad z_1 z_2 = r_1 r_2 (\cos(\alpha_1 + \alpha_2) + i(\sin(\alpha_1 + \alpha_2))),$$

$$(2) \quad z^n = r^n (\cos n\alpha + i \sin n\alpha), \quad \text{az } n \text{ egész,}$$



2.3. ábra: Komplex számok trigonometrikus alakja.

$$(3) \quad \frac{z_1}{z_2} = \frac{r_1}{r_2} (\cos(\alpha_1 - \alpha_2) + i \sin(\alpha_1 - \alpha_2)) \quad (z_2 \neq 0).$$

A szabályokat röviden, szavakban is érdemes megjegyezni:

**Szorzás:** Az abszolút értékeket összeszorozzuk, a szögeket összeadjuk.

**Hatványozás:** Az abszolút értéket hatványozzuk, a szögeket megszorozzuk a kitevővel.

**Osztás:** Az abszolút értékeket elosztjuk, a szögeket kivonjuk.

**Bizonyítás.** (1): A normál-alaknak megfelelően számolva

$$\begin{aligned} z_1 z_2 &= r_1 (\cos \alpha_1 + i \sin \alpha_1) r_2 (\cos \alpha_2 + i \sin \alpha_2) = \\ &= r_1 r_2 (\cos \alpha_1 + i \sin \alpha_1) (\cos \alpha_2 + i \sin \alpha_2) = \\ &= r_1 r_2 ((\cos \alpha_1 \cos \alpha_2 - \sin \alpha_1 \sin \alpha_2) + i(\cos \alpha_1 \sin \alpha_2 + \sin \alpha_1 \cos \alpha_2)). \end{aligned}$$

Ebből pedig a koszinusz és szinusz függvények addíciós képlete alapján:

$$z_1 z_2 = r_1 r_2 (\cos(\alpha_1 + \alpha_2) + i \sin(\alpha_1 + \alpha_2)).$$

(2): Ha az  $\alpha_1 = \alpha_2$  és  $r_1 = r_2$ , akkor az (1) azonosságból

$$z^2 = r^2 (\cos 2\alpha + i \sin 2\alpha).$$

Teljes indukcióval tetszőleges  $n$  pozitív számra: Ha az  $(n-1)$ -re már igaz, akkor

$$z^{n-1} z = r^{n-1} (\cos(n-1)\alpha + i \sin(n-1)\alpha) r (\cos \alpha + i \sin \alpha)$$

és ez az (1) alkalmazásával:

$$= r^n (\cos n\alpha + i \sin n\alpha),$$

amivel az állítást pozitív egész  $n$ -re be is láttuk. Ha az  $n$  negatív egész, akkor a  $z^{-n} = 1/z^n$  megállapodás alapján — feltéve, hogy a  $z$  nem nulla —

$$\begin{aligned} z^n &= \frac{1}{z^{-n}} = \frac{1}{r^{-n}(\cos(-n\alpha) + i \sin(-n\alpha))} = \\ &= r^n \frac{\cos(-n\alpha) - i \sin(-n\alpha)}{\cos^2(-n\alpha) + \sin^2(-n\alpha)} = r^n(\cos n\alpha + i \sin n\alpha), \end{aligned}$$

amivel a negatív  $n$  esetét is beláttuk. A  $z^0 = 1$  megállapodás a  $z^0 = r^0(\cos 0 + i \sin 0) = 1$  egyenlőség szerint teljesül.  $\square$

**Példa 2.9** Számoljuk ki az  $(1 + i)$  századik hatványát.

Mivel az  $(1 + i)$  szám abszolút értéke  $\sqrt{2}$ , a szöge pedig a

$$\cos \alpha = \frac{1}{\sqrt{2}}$$

egyenlőség szerint  $\pi/4$ , ezért

$$1 + i = \sqrt{2}(\cos(\pi/4) + i \sin(\pi/4)),$$

és így

$$\begin{aligned} (1 + i)^{100} &= 2^{50}(\cos(100(\pi/4)) + i \sin(100(\pi/4))) = \\ &= 2^{50}(\cos 25\pi + i \sin 25\pi) = -2^{50} \end{aligned}$$

$\square$

Egy  $z$  szám  $n$ -edik gyökének ( $n$  természetes szám) nevezzük a  $w$  számot, ha  $w^n = z$ . A trigonometrikus alak igen jól alkalmazható az  $n$ -edik gyök kiszámolásánál, amit állításban is megfogalmazunk.

**Állítás 60** Legyen a  $z = r(\cos \alpha + i \sin \alpha)$  egy nemnulla komplex szám, és az  $n$  pozitív egész. Ekkor a  $z$  komplex számnak  $n$  számú, különböző  $n$ -edik gyöke van:

$$\sqrt[n]{r} \left[ \cos \left( \frac{\alpha}{n} + \frac{2k\pi}{n} \right) + i \sin \left( \frac{\alpha}{n} + \frac{2k\pi}{n} \right) \right], \quad (2.16)$$

ahol a  $k$  az

$$0, 1, 2, \dots, (n-1) \quad (2.17)$$

értékeket veszi fel.

**Bizonyítás.** A  $w = h(\cos \phi + i \sin \phi)$  komplex szám pontosan akkor  $n$ -edik gyöke a  $z = r(\cos \alpha + i \sin \alpha)$  komplex számnak, ha  $w^n = z$ , azaz ha

$$h^n (\cos n\phi + i \sin n\phi) = r (\cos \alpha + i \sin \alpha).$$

Ez az egyenlőség pontosan akkor áll fenn, ha egyrészt

$$h^n = r, \quad \text{azaz} \quad h = \sqrt[n]{r},$$

másrészt

$$n\phi - \alpha = 2k\pi,$$

ahol a  $k$  egész szám, azaz

$$\phi = \frac{\alpha}{n} + \frac{2k\pi}{n}, k = 0, \pm 1, \pm 2, \dots$$

Ezek szerint az állításban szereplő (2.16) komplex számok  $n$ -edik hatványai valóban a  $z$  számot adják. Ha megmutatjuk, hogy mind különbözőek, akkor készen is leszünk, mert  $n$ -nél több gyöke nem lehet a  $z^n = w$  egyenletnek. Ehhez elégséges belátnunk azt, hogy az

$$\frac{\alpha}{n} + \frac{2k\pi}{n}, \quad k = 0, 1, 2, \dots, n-1$$

szögmértékek különböző szögeket határoznak meg. Ez pedig azonnal következik abból, hogy az

$$\left(\frac{\alpha}{n} + \frac{2k\pi}{n}\right) - \left(\frac{\alpha}{n} + \frac{2j\pi}{n}\right) = \frac{2k\pi}{n} - \frac{2j\pi}{n} = \frac{2(k-j)\pi}{n}$$

különbség kisebb mint  $2\pi$ , hiszen  $|k-j| \leq |n-1|$ . □

A  $z^n = 1$  egyenletnek a 60. állítás szerint  $n$  számú különböző gyöke van. Ezeket a következő definícióban el is nevezzük.

**Definíció 61** Azt az  $n$  darab

$$\cos\left(\frac{2k\pi}{n}\right) + i \sin\left(\frac{2k\pi}{n}\right), \quad k = 0, 1, \dots, n-1$$

(különböző) komplex számot — amelyeknek az  $n$ -edik hatványa 1 —  $n$ -edik egységgyököknek nevezzük.

Határozzunk meg néhány egységgyököt.

1) A második egységgyökök:

$$\cos 0 + i \sin 0 \quad \text{és} \quad \cos\left(\frac{2\pi}{2}\right) + i \sin\left(\frac{2\pi}{2}\right),$$

azaz az 1 és  $(-1)$ .

2) A harmadik egységgyökök: Figyelembe véve, hogy  $(360^\circ)/3 = 120^\circ$ ,

$$\begin{aligned}\cos 0^\circ + i \sin 0^\circ &= 1, \\ \cos 120^\circ + i \sin(120)^\circ &= -\frac{1}{2} + i\frac{\sqrt{3}}{2}, \\ \cos 240^\circ + i \sin 240^\circ &= -\frac{1}{2} - i\frac{\sqrt{3}}{2}.\end{aligned}$$

3) A negyedik egységgyökök: Figyelembe véve, hogy  $(2\pi)/4 = \pi/2$ ,

$$\begin{aligned}\cos 0 + i \sin 0 &= 1, \\ \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} &= i, \\ \cos \pi + i \sin \pi &= -1, \\ \cos \frac{3\pi}{4} + i \sin \frac{3\pi}{4} &= -i.\end{aligned}$$

4) A hatodik egységgyökök: Figyelembe véve, hogy  $360^\circ/6 = 60^\circ$ ,

$$\begin{aligned}\cos 0^\circ + i \sin 0^\circ &= 1 \\ \cos 60^\circ + i \sin 60^\circ &= \frac{1}{2} + i\frac{\sqrt{3}}{2}, \\ \cos 120^\circ + i \sin 120^\circ &= -\frac{1}{2} + i\frac{\sqrt{3}}{2}, \\ \cos 180^\circ + i \sin 180^\circ &= -1, \\ \cos 240^\circ + i \sin 240^\circ &= -\frac{1}{2} - i\frac{\sqrt{3}}{2}, \\ \cos 300^\circ + i \sin 300^\circ &= \frac{1}{2} - i\frac{\sqrt{3}}{2}.\end{aligned}$$

Az  $n$ -edik egységgyökök olyan — az egységkörbe írt — szabályos  $n$  szög csúcspontjaiban helyezkednek el, amelynek az 1 csúcspontja. (a 2.4.–2.6. ábrák).

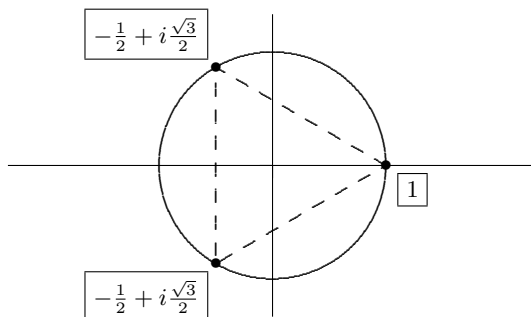
Az egységgyökök segítségével a 60. tételt a következőképpen fogalmazhatjuk át:

**Állítás 62** Legyen  $a z = r(\cos \alpha + i \sin \alpha)$  egy nemnulla komplex szám, és az  $n$  pozitív egész. Ekkor a  $z$  komplex számnak  $n$  darab  $n$ -edik gyöke van:

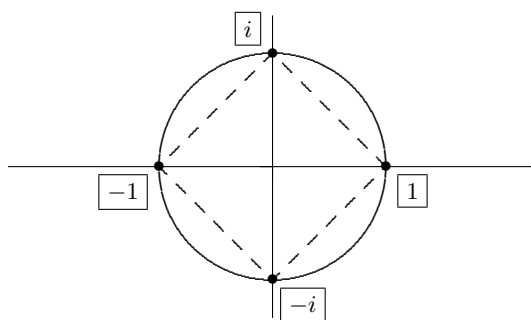
$$\sqrt[n]{r} \left[ \cos \left( \frac{\alpha}{n} \right) + i \sin \left( \frac{\alpha}{n} \right) \right] \cdot \epsilon_k, \quad k = 0, 1, \dots, n-1, \quad (2.18)$$

ahol az  $\epsilon_k$  az  $n$ -edik egységgyökök közül a  $k$ -adik, azaz

$$\epsilon_k = \cos \left( \frac{2k\pi}{n} \right) + i \sin \left( \frac{2k\pi}{n} \right).$$



2.4. ábra: A harmadik egységgyökök..



2.5. ábra: A negyedik egységgyökök..

A (2.18) szerint a  $z$  komplex szám  $n$ -edik gyökeit úgy kapjuk meg, hogy vesszük az egyik

$$\sqrt[n]{r} \left[ \cos \left( \frac{\alpha}{n} \right) + i \sin \left( \frac{\alpha}{n} \right) \right]$$

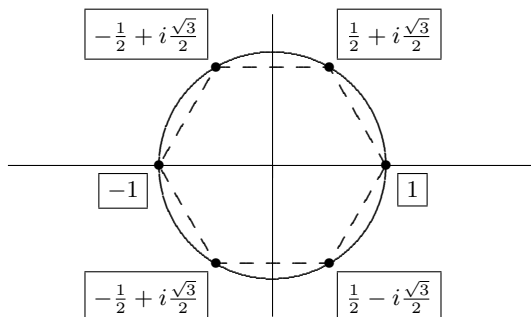
$n$ -edik gyökét, és rendre megszorozzuk az  $n$ -edik egységgyökökkel.

### Algebrai (polinom) egyenletek gyökei

A komplex számok testének egyik nagy előnye, hogy lehetővé teszi az algebrai (polinom) egyenletek gyökeinek a harmonikus vizsgálatát. A tanulmányozás kiinduló állítása az u. n. algebra alaptétele. Ezt a nevezetes tételt mondjuk ki először. Igazolni, a jelenleg rendelkezésre álló eszközeinkkel nem tudjuk. Későbbi tanulmányaink során—alkalmas elméleti ismeretek birtokában— majd adunk rá egy bizonyást.

**Állítás 63 (Az algebra alaptétele)** Minden komplex együtthatós

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0 = 0$$



2.6. ábra: A hatodik egységgyökök..

*algebrai (polinom) egyenletnek van gyöke a komplex számok körében.*

A komplex számok bevezetésével az  $x^2 + 1 = 0$  egyenlet megoldhatatlanságát igyekeztünk megszüntetni, ezért eléggé meglepő, hogy ez a célkitűzés olyan nagy eredményt hozott, hogy most már minden algebrai egyenletnek van gyöke.

Az alaptételnek fontos következménye a gyöktényezős alakra vonatkozó tétel.

**Állítás 64** Legyen  $a$

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = 0, \quad a_n \neq 0$$

egy  $n$ -ed fokú, ( $n \geq 1$ ), algebrai egyenlet. Ekkor vannak olyan

$$z_1, z_2, \dots, z_k, \quad 1 \leq k \leq n$$

különböző komplex és

$$m_1, m_2, \dots, m_k, \quad m_1 + m_2 + \dots + m_k = n$$

pozitív egész számok, hogy a  $p$  polinom a tényezőik sorrendjétől eltekintve egyértelműen írható fel

$$p(z) = a_n (z - z_1)^{m_1} (z - z_2)^{m_2} \dots (z - z_k)^{m_k}, \quad (2.19)$$

alakban.

A bizonyítás előtt néhány elnevezés:

**Definíció 65** A 64. állításban szereplő  $m_1, \dots, m_k$  egész számokat a  $z_1, \dots, z_k$  gyökökhöz tartozó multiplicitásoknak mondjuk a

$$(z - z_1), \dots, (z - z_k)$$

elsőfokú polinomokat pedig gyöktényezőknak. Az (2.19) formát a  $p$  gyöktényezős alakjának mondjuk.

A tétel szerint egy  $n$ -ed fokú algebrai egyenletnek pontosan  $n$  számú gyöke van (a komplex számok körében), ha a gyököket multiplicitással vesszük figyelembe.

**Bizonyítás.** Először azt mutatjuk meg teljes indukcióval, hogy minden polinom felírható gyöktényezős alakba, majd a felbontás egyértelműségét igazoljuk.

Ha  $n = 1$ , akkor nyilvánvalóan igaz az állítás, hiszen

$$a_1 z + a_0 = a_1 \left( z - \left( \frac{a_0}{a_1} \right) \right).$$

Tegyük fel most, hogy minden  $(n - 1)$ -ed fokú polinom felírható gyöktényezős alakba, és legyen a

$$p(z) = a_n z^n + \dots + a_1 z + a_0$$

egy tetszőleges  $n$ -ed fokú polinom. Az algebra alaptétele szerint  $p(z)$ -nek van  $\alpha$  gyöke. Az ismert

$$z^i - \alpha^i = (z - \alpha)(z^{i-1} + z^{i-2}\alpha + \dots + z\alpha^{i-2} + \alpha^{i-1})$$

azonosság felhasználásával, kapjuk, hogy

$$\begin{aligned} p(z) &= p(z) - p(\alpha) = \\ &= a_n(z^n - \alpha^n) + a_{n-1}(z^{n-1} - \alpha^{n-1}) + \dots + a_1(z - \alpha) = a_n(z - \alpha)g(z), \end{aligned}$$

ahol a  $g(z)$  1 főegyütthatójú  $(n - 1)$ -ed fokú polinom. Az indukciós feltevés szerint a  $g$  felírható

$$g(z) = (z - z_1)^{n_1} \dots (z - z_s)^{n_s},$$

$$1 \leq s \leq n - 1, \quad z_1, \dots, z_s \text{ különbözőek}, \quad n_1 + \dots + n_s = n - 1$$

alakba, ezért ha az  $\alpha$  gyök a  $z_1, \dots, z_s$  gyökök valamelyikével egyezik meg, akkor multiplicitást növeljük a megfelelő helyen eggyel, ha pedig nem, akkor a  $(z - \alpha)$  új gyöktényezőként lép fel a  $p$  előállításában. Igazolnunk kell még a gyöktényezős alak egyértelműségét. Tegyük fel ezért, hogy

$$(1) \quad p(z) = a_n(z - z_1)^{m_1}(z - z_2)^{m_2} \dots (z - z_k)^{m_k}$$

és

$$(2), \quad p(z) = a_n(z - z_1)^{\ell_1}(z - z_2)^{\ell_2} \dots (z - z_k)^{\ell_k}$$

továbbá létezik olyan  $i (= 1, \dots, k)$ , hogy  $m_i \neq \ell_i$  mondjuk  $m_i < \ell_i$ . De ez lehetetlen, mert akkor a

$$\frac{p(z)}{(z - z_i)^{m_i}}$$

polinomnak (1) szerint nem gyöke (2) alapján viszont zérushelye a  $z_i$  komplex szám.  $\square$

Valós együtthatós polinom esetében a gyöktényezős előállítást — a valós számok körében maradván — a következőképpen lehet megadni:

**Állítás 66** Legyen  $a$   $p$  egy valós együtthatós  $n$ -ed fokú polinom,  $a_n$  főegyütthatóval. Ekkor  $a$   $p$  egyértelműen írható fel

$$a_n(z - \alpha_1)^{k_1} \cdots (z - \alpha_s)^{k_s} (z^2 + \beta_1 z + \gamma_1)^{l_1} \cdots (z^2 + \beta_r z + \gamma_r)^{l_r}$$

formában, ahol az

$$\alpha_1, \dots, \alpha_s, \quad \beta_1, \dots, \beta_r, \quad \gamma_1, \dots, \gamma_r$$

olyan valós számok, hogy az első és másodfokú gyöktényezők különbözőek és a multiplicitásokra:

$$k_1 + \cdots + k_s + 2(l_1 + \cdots + l_r) = n.$$

Bizonyítás. A 55. állítás szerint minden valós együtthatós

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$$

polinomra

$$\begin{aligned} \overline{p(z)} &= \overline{a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0} = \\ &= a_n \bar{z}^n + a_{n-1} \bar{z}^{n-1} + \cdots + a_1 \bar{z} + a_0 = p(\bar{z}). \end{aligned}$$

Ebből viszont az következik, hogy a valós együtthatós polinomnak  $z_j$  gyökei vagy valósak, vagy a konjugáltjukkal együtt (párban) szerepelnek. Jelölve a valós gyököket  $\alpha_i$ -vel, ( $i = 1, \dots, s$ ) megkapjuk a bebizonyítandó előállítás első felét. Egy  $(a + ib)$  és  $(a - ib)$  konjugált párra vonatkozó gyöktényezők szorzata

$$(z - (a + ib))(z - (a - ib)) = z^2 - 2az + (a^2 + b^2),$$

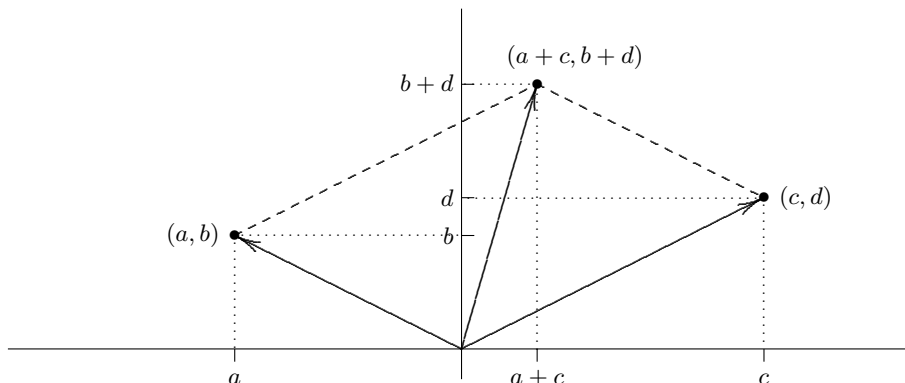
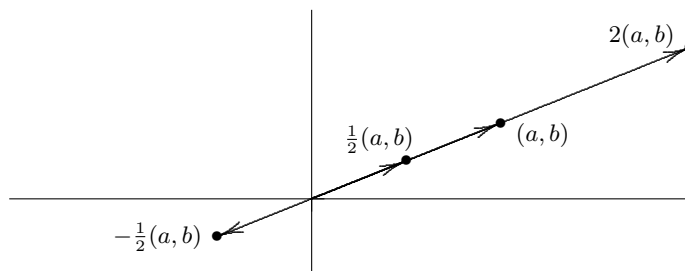
alakú. A konjugált párokban lévő komplex gyökökből így adódnak az előállításban szereplő másodfokú tényezők. Innen már a 64. állítás felhasználásával azonnal adódik a tétel.  $\square$

## 2.2 Valós vektorok, az $\mathbb{R}^p$ tér

Már a középiskolai tanulmányaink során találkoztunk olyan mennyiségekkel, amelyek nem jellemezhetők egyetlen számadattal, ilyenek például az erő, elmozdulás, sebesség stb. Ezek reprezentálására vektorokat használtunk. Kiderült, hogy a síkbeli helyvektorok és a rendezett valós számpárok között létezik egy bijektív megfeleltetés. Teljesen hasonlóan lehet kölcsönösen egyértelmű hozzárendelést létesíteni a térbeli helyvektorok halmaza és a rendezett valós számhármasok halmaza között.

Azt is láttuk, hogy ha a rendezett valós számpárok  $\mathbb{R}^2$  halmazán a

$$(a, b) + (c, d) \stackrel{\text{def}}{=} (a + c, b + d), \quad (a, b), (c, d) \in \mathbb{R}^2, \quad (2.20)$$

2.7. ábra: Az  $\mathbb{R}^2$  sík két elemének az összeadása.2.8. ábra: A számmal való szorzás az  $\mathbb{R}^2$  síkon.

definíció szerinti összeadást értelmezünk, akkor az a helyvektorok "parallelogramma szabály" szerinti összeadásának felel meg, amint azt a 2.7. ábrán szemléltetjük.

A helyvektorok valós számmal (skalárral) való szorzásának megfelelő operáció az  $\mathbb{R}^2$  halmazon az

$$\alpha \cdot (a, b) \stackrel{\text{def}}{=} (\alpha a, \alpha b) \quad (2.21)$$

egyenlőséggel értelmezett ú.n. *skalárral való szorzás* ahol az  $\alpha$  tetszőleges valós szám (skalár).

A skalárral való szorzást a 2.8. ábra szemlélteti.

Az is jól ismert, hogy a fenti összeadás kommutatív, asszociatív, reprodukáló elemes (a zero hosszúságú vektornak megfelelő  $(0,0)$  pár a reprodukáló elem) és invertálható (tetszőleges  $(a,b)$  pár inverze a  $(-a,-b)$  pár).

A skalárral való szorzásnak pedig a következő könnyen ellenőrizhető tulajdonságai vannak:

- (1)  $(\alpha + \beta) \cdot (a, b) = \alpha(a, b) + \beta(a, b)$ .
- (2)  $\alpha((a, b) + (c, d)) = \alpha(a, b) + \alpha(c, d)$ .

$$(3) \alpha(\beta(a, b)) = (\alpha\beta)(a, b).$$

$$(4) 1 \cdot (a, b) = (a, b).$$

A továbbiakban minden olyan algebrai struktúrát, amelyben a fenti négy tulajdonsággal rendelkező összeadás és ugyancsak a fenti négy tulajdonsággal bíró skalárral való szorzás van értelmezve — ahol a skalárok valamely számtest elemei — *vektortérnek* fogunk nevezni.

Az  $\mathbb{R}^2$ -beli összeadás és skalárral való szorzás definíciókat véve mintául további vektorterek konstruálhatók.

Jelölje  $\mathbb{R}^p$  ( $p$  pozitív egész) az összes rendezett valós szám- $p$ -esek halmazát, azaz legyen

$$\mathbb{R}^p = \{x = (x_1, \dots, x_p) : x_i \in \mathbb{R}, i = 1, \dots, p\}.$$

A síkbeli vektorok összeadásának mintájára értelmezzük két  $p$ -komponensű vektor összegét a következőképpen:

$$(x_1, \dots, x_p) + (y_1, \dots, y_p) \stackrel{\text{def}}{=} (x_1 + y_1, \dots, x_p + y_p),$$

azaz legyen az összegvektor minden komponense egyenlő az összeadandók megfelelő komponenseinek összege. Könnyen ellenőrizhető, hogy a definiált összeadás kommutatív és asszociatív, reprodukáló elem a csupa zéró komponensből álló  $p$ -es, és tetszőleges  $x = (x_1, \dots, x_p)$  vektornak az inverze a  $-x = (-x_1, \dots, -x_p)$  vektor.

Ugyancsak az  $\mathbb{R}^2$ -beli vektorok skalárokkal való szorzásának analógiájára, definiáljuk az  $\mathbb{R}^p$ -beli vektorok tetszőleges  $\alpha$  valós skalárral való szorzatát az

$$\alpha \cdot (x_1, \dots, x_p) \stackrel{\text{def}}{=} (\alpha x_1, \dots, \alpha x_p)$$

egyenlőséggel. Nem nehéz ellenőrizni, hogy a skalárokkal való szorzás rendelkezik az 1–4 tulajdonságokkal,

Igy  $\mathbb{R}^p$ -t az azon értelmezett összeadással és valós skalárral való szorzással szintén vektortérre tettük.

A középiskolában értelmeztük a sík vektorainak skaláris szorzatát, hozzárendelve tetszőleges két síkbeli  $x$  és  $y$  vektorhoz az  $x \cdot y \stackrel{\text{def}}{=} \|x\| \|y\| \cos \phi$  valós számot, ahol  $\|x\|$  illetve  $\|y\|$  az  $x$  illetve  $y$  vektorok hosszát jelentette, míg  $\phi$  a vektorok által bezárt szöget jelölte. Később azt is beláttuk, hogy ha  $x = (x_1, x_2)$  és  $y = (y_1, y_2)$ , akkor a két vektor skaláris szorzatát a koordináták szorzatösszegeként is megkaphatjuk, azaz

$$i \quad x \cdot y = x_1 y_1 + x_2 y_2.$$

A skaláris szorzatra támaszkodva kiszámítható a vektor hossza a

$$\|x\| = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2}$$

képlet alapján, hiszen az origó kezdőpontú és az  $x = (x_1, x_2)$  pontba mutató vektor hossza az  $x_1$  és  $x_2$  befogójú derékszögű háromszög átfogójának hosszával egyenlő. Emiatt a Pithagorasz tétel szerint

$$\|x\|^2 = x_1^2 + x_2^2,$$

amiből

$$\|x\| = \sqrt{x_1^2 + x_2^2}.$$

Kézenfekvőnek látszik, hogy az  $\mathbb{R}^p$  vektortérben is értelmezzük vektorok skaláris szorzatát, és annak definiálásakor az (i) képletet használjuk az általánosításhoz.

**Definíció 67** Legyen  $x = (x_1, \dots, x_p)$  és  $y = (y_1, \dots, y_p)$  az  $\mathbb{R}^p$  vektortér tetszőleges két vektora. Ekkor az  $x$  és  $y$  vektorok skaláris szorzatán az

$$x \cdot y = \sum_{i=1}^p x_i y_i$$

valós számot értjük.

Az  $\mathbb{R}^p$ -beli vektorok skaláris szorzata is rendelkezik azokkal a tulajdonságokkal, amelyet a síkbeli vektorok skaláris szorzatáról már beláttunk középiskolai tanulmányaink során. Ezek a következők: minden  $x = (x_1, \dots, x_p)$ ,  $y = (y_1, \dots, y_p)$  és  $z = (z_1, \dots, z_p)$  vektorokra és minden  $\lambda$  valós skalárra

$$(a) \quad x \cdot y = y \cdot x,$$

$$(b) \quad (\lambda x) \cdot y = \lambda(x \cdot y),$$

$$(c) \quad (x + y) \cdot z = x \cdot z + y \cdot z,$$

$$(d) \quad x \cdot x \geq 0 \text{ és } x \cdot x = 0 \text{ akkor és csak akkor ha } x = (0, \dots, 0).$$

A fenti tulajdonságok igazolása nagyon egyszerű, ezért azzal most nem is foglalkozunk. Igazoljuk viszont, hogy tetszőleges  $x, y \in \mathbb{R}^p$  vektorra és  $\lambda$  valós számra az

$$x \cdot (\lambda y) = \lambda(x \cdot y)$$

tulajdonság is teljesül. Valóban, az (a), majd a (b), végül ismét az (a) tulajdonság kihasználásával kapjuk, hogy

$$x \cdot (\lambda y) = (\lambda y) \cdot x = \lambda(y \cdot x) = \lambda(x \cdot y).$$

Az (a) és (c) tulajdonságokból az is következik, hogy

$$x \cdot (y + z) = (y + z) \cdot x = y \cdot x + z \cdot x = x \cdot y + x \cdot z.$$

Megmutatjuk, hogy a skaláris szorzatra támaszkodva  $\mathbb{R}^p$ -ben is definiálható vektorhosszúság függvény.

Egy ilyen valós értékű függvénytől azt várjuk, hogy rendelkezzen azokkal a tulajdonságokkal, amit a síkbeli, illetve térbeli vektorok hossza teljesít, azaz minden  $x, y$  vektorra és  $\lambda$  valós számra

1.  $\|x\| \geq 0$  és  $\|x\| = 0$  pontosan akkor, ha  $x = 0$ ,
2.  $\|\lambda x\| = |\lambda| \cdot \|x\|$ ,
3. és teljesül a háromszög egyenlőtlenség:  $\|x + y\| \leq \|x\| + \|y\|$ .

**Állítás 68** Legyen  $x = (x_1, \dots, x_p) \in \mathbb{R}^p$ . Az

$$\|x\| \stackrel{\text{def}}{=} \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^p x_i^2}$$

módon definiált függvény vektorhosszúság vagy más néven norma az  $\mathbb{R}^p$  vektortérben. Az  $\|\cdot\|$  normát euklideszi normának nevezzük.

Az euklideszi norma a hosszúság megszokott fogalmának általánosítása, hiszen  $p = 2$ -re vagy  $p = 3$ -ra a Pithagorasz tétel felhasználásával kiszámítható vektorhosszúsággal azonos.

**Bizonyítás.** 1) A skaláris szorzat negyedik tulajdonságából azonnal kapjuk, hogy  $\|x\| \geq 0$  és pontosan akkor nulla, ha  $x = (0, \dots, 0)$ .

2) A második norma tulajdonság egyszerű számolással adódik

$$\|\alpha x\| = \sqrt{\alpha x \cdot \alpha x} = \sqrt{\alpha^2 (x \cdot x)} = |\alpha| \sqrt{x \cdot x} = |\alpha| \|x\|.$$

3) A háromszög egyenlőtlenség bizonyításának magja a következő u.n. *Cauchy — Schwarz egyenlőtlenség*:

minden  $x, y \in \mathbb{R}^p$  vektorpárra

$$|x \cdot y| \leq \|x\| \cdot \|y\|.$$

A skaláris szorzat (d) tulajdonsága miatt bármely valós  $\lambda$  skalárra

$$(x + \lambda y) \cdot (x + \lambda y) \geq 0.$$

A skaláris szorzat tulajdonságai alapján a fenti egyenlőtlenség

$$\|x\|^2 + 2\lambda x \cdot y + \lambda^2 \|y\|^2 \geq 0$$

alakban írható. Az egyenlőtlenség bal oldalán szereplő másodfokú polinom a  $\lambda$  bármely értéke mellett csak úgy lehet nemnegatív, ha diszkriminánsa nempozitív, azaz

$$4(x \cdot y)^2 - 4\|y\|^2 \|x\|^2 \leq 0.$$

Ezt az egyenlőtlenséget átrendezve 4-el egyszerűsítve, majd mindkét oldalból négyzetgyököt vonva kapjuk a kívánt Cauchy — Schwarz egyenlőtlenséget.

Ezekután a háromszög egyenlőtlenség igazolása: tekintve, hogy bármely valós szám kisebb vagy egyenlő mint az abszolút értéke, a Cauchy — Schwarz egyenlőtlenségből következik, hogy

$$x \cdot y \leq \|x\| \|y\|,$$

ezért

$$\begin{aligned} \|x + y\|^2 &= (x + y) \cdot (x + y) = \\ &= \|x\|^2 + 2x \cdot y + \|y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2. \end{aligned}$$

A számolás sorozat első és utolsó tagjából gyököt vonva kapjuk a kívánt egyenlőtlenséget.  $\square$

Megjegyezzük, hogy az euklideszi normára vonatkozó háromszög egyenlőtlenséget Minkowski-egyenlőtlenségnek is nevezik.

Meg kell említenünk, hogy  $\mathbb{R}^p$ -ben másképpen is definiálható norma. Például az  $x = (x_1, \dots, x_p)$  vektorhoz az

$$\|x\| \stackrel{\text{def}}{=} \max(|x_1|, \dots, |x_p|)$$

egyenlőséggel értelmezett függvény is rendelkezik mindhárom norma tulajdonsággal. Ennek igazolását az olvasóra bizzuk.

## 2.3 Algebrai struktúrák

Először néhány nagyon általános fogalmat vezetünk be, amelyek — a tárgyalás jelenlegi szintjén — elsősorban mint elnevezések érdekesek. A bevezetett fogalmak az algebra tárgykörébe tartoznak, aminek a részletes tanulmányozása jelenleg nem elsődleges feladatunk.

Találkoztunk már nevezetes struktúrákkal, például a Boole algebrával, a rendezett halmazokkal (rendezett terekkel).

Most olyan struktúrákat fogunk megismerni, amelyek a szám fogalmával kapcsolatban vetődnek fel. A bevezetett absztrakciók eredeténél tartsuk szem előtt a megszokott egész számokat, racionális számokat és valós számokat. Ahogyan már mondtuk is, úgy tekintsük az itt elmondottakat, hogy a már meglévő ismereteinket rendezzük el, egy új és célszerű fogalom-rendszerbe.

Az első absztrakció azt a tapasztalást általánosítja, hogy a közönséges számolásban az összeadás, szorzás és egyéb műveletek két számból egy újabb számot “csinálnak”.

**Definíció 69** Legyen az  $X$  egy nem-üres halmaz. Egy  $f : X \times X \rightarrow X$  leképezést műveletnek is szokás nevezni, és olyan írásmód is lehetséges, hogy egy műveleti jelet — a jelen esetben legyen ez “ $\bullet$ ” — használunk a függvény értékének a megadására:

$$f(x, y) \stackrel{\text{def}}{=} x \bullet y. \quad (2.22)$$

Ha egy halmazon valamilyen művelet van definiálva, akkor algebrai struktúrának mondjuk.

Hétköznapiiban fogalmazva: Ha egy halmaz tetszőleges két eleméből alkotott pároshoz hozzá van rendelve a halmaznak egy eleme, akkor ezt a hozzárendelést műveletnek is nevezhetjük.

**Definíció 70** Az előző definícióban szereplő (2.22) művelet

- (1) *asszociatív, ha minden  $x, y, z \in X$  elemre*

$$x \bullet (y \bullet z) = (x \bullet y) \bullet z;$$

- (2) *kommutatív, ha tetszőleges két  $x, y \in X$  elemre*

$$x \bullet y = y \bullet x;$$

- (3) *reprodukáló elemes, ha létezik olyan  $e$  elem az  $X$  halmazban, hogy minden  $x \in X$  elemre*

$$e \bullet x = x \bullet e = x;$$

*Az  $e$  elemet reprodukáló elemnek fogjuk nevezni.*

- (4) *inverz elemes, ha minden  $x \in X$  elemhez van olyan  $x' \in X$  inverz elemnek nevezett elem, hogy*

$$x \bullet x' = x' \bullet x = e,$$

*ahol az  $e$  a reprodukáló elem.*

A definiált tulajdonságok természetesen a függvényes jelölési móddal is felírhatóak. Vegyük például az asszociativitást:

$$f(x, f(y, z)) = f(f(x, y), z).$$

Szembetűnő, hogy a műveleti jel áttekinthetőbb, legalábbis számunkra megszokottabb írásmódot ad.

Azonnal látható, hogy az asszociativitás arra szolgál, hogy három elemre is kiterjeszthessük a műveletünket, hiszen az  $x \bullet y \bullet z$  felírás az asszociativitás teljesülése esetében már értelmes, hiszen éppen azt mondja, hogy akárhogyan zárójelezve is kiszámolható a három elemhez rendelt érték. Teljes indukcióval, az asszociativitás segítségével tetszőleges véges számú elemre is kiterjeszthető a műveletünk.

Most pedig lássunk néhány példát, véve először a legklasszikusabb eseteket.

**Példa 2.10** Az egész számok  $\mathbb{Z}$  összességénél az összeadás művelete

- *kommutatív,*
- *asszociatív,*
- *reprodukáló elemes az 0 elemmel,*

- inverz elemes, egy  $x$  elem inverze az ellentettje, azaz a  $-x$  elem.

**Példa 2.11** Az egész számok összességénél a szorzás művelete:

- kommutatív és asszociatív,
- reprodukáló elemes, a reprodukáló elem az 1,
- nem inverz-elemes, mert inverze csak az 1 és  $-1$  elemnek van.

**Példa 2.12** A racionális számok  $\mathbb{Q}$  halmazát véve,

- az összeadás kommutatív, asszociatív, reprodukáló elem a 0, inverz elemes, egy  $x$  szám (additív) inverze a  $-x$ ;
- a szorzás kommutatív, asszociatív, reprodukáló elem az 1, (multiplikatív) inverze minden nem nulla  $x$  számnak van, az  $1/x$  (vagyis a  $\mathbb{Q} \setminus \{0\}$  inverz-elemes).

**Példa 2.13** Vegyük a következő öt szimbólumból álló halmazt

$$\{0, 1, 2, 3, 4\}.$$

Ezen a halmazon két műveletet is bevezetünk. Az egyiket “szorzásnak” a másikat pedig “összeadásnak” fogjuk nevezni. Megmaradunk a racionális számoknál megszokott műveleti jeleknél is, persze új jelentéssel. A műveleti szabályokat a következőképpen definiáljuk:

**ÖSSZEADÁS:** Összeadjuk a két szimbólumot jelölő számjegyet, és az összeget az 5 számmal leosztva a maradékot vesszük. Például:  $2 + 4 = 6$ , ezért a 2 és 4 szimbólumok összege az 1 szimbólum.

**SZORZÁS:** Összeszorozzuk a két szimbólumot jelölő számot, és a szorzatot az 5 számmal leosztva vesszük a maradékot. Például:  $2 \cdot 4 = 8$  ezért a 2 és 4 szimbólumok szorzata a 3 szimbólum.

Állítsunk elő egy áttekinthető műveleti táblázatot, és vizsgáljuk meg, hogy milyen tulajdonságai vannak a definiált műveleteknek.

### Összeadási és szorzási táblázatok.

+	0	1	2	3	4
0	0	1	2	3	4
1	1	2	3	4	0
2	2	3	4	0	1
3	3	4	0	1	2
4	4	0	1	2	3

·	0	1	2	3	4
0	0	0	0	0	0
1	0	1	2	3	4
2	0	2	4	1	3
3	0	3	1	4	2
4	0	4	3	2	1

A műveletek tulajdonságai akár a műveleti táblázatok, akár közvetlenül a definíciók alapján megvizsgálhatóak. Nézzük például a kommutativitást. Egy táblázaton ez a tulajdonság úgy jelentkezik, hogy a táblázat szimmetrikus, a főátlóra, ami azonnal látható mindkét műveletnél.

A reprodukáló elem léte úgy olvasható le a táblázatból, hogy valamely elem alatti oszlop megegyezik a táblázat bal szélén lévő oszloppal. Az összeadás műveleténél az első oszlop ilyen, tehát a reprodukáló elem a 0. A szorzás műveleti táblázatában a második, az 1 alatti, oszlop ilyen tulajdonságú, ezért ott a reprodukáló elem az 1.

Vizsgáljuk meg az inverz elem létezését. Nyilvánvalóan azt kell néznünk, hogy egy elem sorában szerepel-e a reprodukáló elem, mert ha igen, akkor a neki megfelelő oszlop fejlécében lévő elem az inverze.

Például:

Az összeadásnál a 2 sorának és a 3 oszlopa találkozásánál van a nulla, tehát a 2 elem inverze a 3 elem. A közösleges összeadástól kölcsönözve az elnevezést azt mondhatnánk: a 2 elentettje a 3 elem. Könnyen látható, hogy minden elemnek van inverze.

A szorzásnál a 0 elemnek nincs inverze, mert a sorában nem szerepel az 1 elem. Az összes többi elemnek van inverze. Például a 4 elem inverze a szorzásra nézve a 4 elem önmaga. A közösleges szorzásból véve az elnevezést azt is mondhatnánk, hogy a 4 elem reciproka maga a 4 elem. Ez persze eléggé meglepő jelenségnek tűnhet.

Gondoljuk meg azt is, hogy miért asszociatívak a műveletek. Ennek a belátása a táblázatok ismeretében véges sok eset megvizsgálását jelenti. Ez persze az esetek nagy száma miatt hosszú ideig tarthat. Egy kis módszeresség segíthet az összes eset végignézésében.

Erre a — talán különösnek érzett — példára egyébként még visszatérünk, ezért foglaljuk össze a műveletek tulajdonságait:

Az összeadás: kommutatív, asszociatív, a reprodukáló elem a 0, inverz elemes.

A szorzás: kommutatív, asszociatív, a reprodukáló elem az 1, a 0 elemen kívül minden elemnek van inverz eleme.  $\square$

**Példa 2.14** Az  $X$   $n$ -elemű halmazt nevezzük ábécének, az elemeit betűknek, véges sok egymás mellé írt  $"x_1x_2\dots x_k"$  betűt (a sorrend lényeges) pedig egy szónak. A szavak összességét jelöljük  $\langle X \rangle$ -szel.

Definiáljunk egy  $"\bullet"$  műveletet a szavak között az egymás után való írással:

$$x_1x_2\dots x_k \bullet y_1y_2\dots y_i = x_1x_2\dots x_ky_1y_2\dots y_i.$$

Az egyetlen betűt sem tartalmazó, üres szót jelölje az  $\emptyset$  szimbólum.

Milyen tulajdonságokat teljesít ez a művelet?

Könnyen látható, hogy a  $"\bullet"$  művelet asszociatív és a reprodukáló elem az  $\emptyset$  üres szó:  $x_1x_2\dots x_n\emptyset = x_1x_2\dots x_n$ . Nem inverz-elemes.  $\square$

A példáinkban több olyan művelet szerepelt, amelyik kommutatív, asszociatív, reprodukáló és inverz elemes volt, ezért célszerű elnevezést is bevezetni az ilyen tulajdonságú struktúrákra, amit a következő definícióban teszünk meg.

**Definíció 71** Ha egy  $X$  halmazon olyan művelet van értelmezve, amelyik asszociatív, reprodukáló és inverz elemes, akkor az algebrai struktúrát csoportnak nevezzük. Ha még a kommutativitás is teljesül, akkor kommutatív csoportot mondunk.

Hangsúlyozzuk, hogy a csoport fogalom semmi más, mint rövid, egy szóval való kifejezése ennek: "asszociatív, reprodukáló és inverz elemes algebrai struktúra".

Az előző példáinkban már többször megjelent a csoport struktúra; nézzük végig őket ebből a szempontból:

- Az egész számok halmaza az összeadás műveletével kommutatív csoport,
- A racionális számok halmaza az összeadás műveletével kommutatív csoport.
- A racionális számok összességéből elvéve a nulla elemet  $(\mathbb{Q} \setminus \{0\})$ , a szorzás műveletével kommutatív csoportot kapunk.
- A 2.13. példánk műveletei ugyanazokat a tulajdonságokat elégítik ki, mint a racionális számok műveletei, ezért az előzőek rá is érvényesek.
- A valós számok halmaza az összeadás műveletével kommutatív csoport.
- A valós számok összességéből elvéve a nulla elemet  $(\mathbb{R} \setminus \{0\})$ , a szorzás műveletével kommutatív csoport.
- A komplex számok halmaza az összeadás műveletével kommutatív csoport.
- A komplex számok összességéből elvéve a nulla elemet  $(\mathbb{C} \setminus \{0\})$ , a szorzás műveletével kommutatív csoport.

Most elérkeztünk ahhoz a ponthoz, hogy a számunkra legfontosabb algebrai struktúrát a testet precízen definiáljuk.

**Definíció 72** Legyen egy  $X$  halmazon két művelet definiálva, amelyek közül az egyiket összeadásnak nevezzük, és a “+” műveleti jelet használjuk, a másikat pedig szorzásnak fogjuk mondani, és a műveleti jele a “.”. Ha az  $X$  struktúra két műveletére teljesülnek az alábbiak, akkor testnek nevezzük.

- (a) Az  $X$  kommutatív csoport az összeadás műveletére nézve.
- (b) Az  $X \setminus \{0\}$  halmaz, ahol a  $0$  az összeadás reprodukáló eleme, a szorzásra nézve kommutatív csoport.
- (c) A két műveletet összekapcsolja az a követelmény, hogy tetszőleges három  $x, y, z \in X$  elemre teljesül az

$$x \cdot (y + z) = x \cdot y + x \cdot z,$$

azonosság (disztributivitás).

A “.” műveleti jelet általában el is szokás hagyni — ha nem vezet félreértésre — és ekkor az egyszerű egymás mellé írás jelöli a szorzás műveletét.

Mielőtt részleteznénk a test tulajdonságait a tömör definíciót aprólékosan megismétljük:

A két, “+” és “.” művelettel ellátott  $X$  halmazon bevezetett struktúrát testnek nevezzük, ha teljesülnek a következőkben leírt azonosságok illetve állítások.

#### AZ ÖSSZEADÁS MŰVELETE

- (1) kommutatív, azaz tetszőleges két  $x, y \in X$  elemre

$$x + y = y + x;$$

- (2) asszociatív, azaz minden  $x, y, z \in X$  elemre

$$x + (y + z) = (x + y) + z;$$

- (3) van reprodukáló elem, amit  $0$ -val fogunk jelölni, amelyre minden  $x \in X$  elemre

$$0 + x = x + 0 = x;$$

- (4) minden  $x \in X$  elemnek van inverze, amit  $-x$ -szel fogunk jelölni, amelyre

$$x + (-x) = (-x) + x = 0.$$

## A SZORZÁS MŰVELETE

- (5) kommutatív, azaz tetszőleges két
- $x, y \in X$
- elemre

$$x \cdot y = y \cdot x.$$

- (6) asszociatív, azaz minden
- $x, y, z \in X$
- elemre

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z.$$

- (7) van reprodukáló elem, amit 1-gyel fogunk jelölni, amelyre minden
- $x \in X$
- elemre

$$1 \cdot x = x \cdot 1 = x.$$

- (8) minden
- $x \in X$
- nem-nulla elemnek van inverze, amit
- $x^{-1}$
- gyel fogunk jelölni, amelyre

$$x \cdot x^{-1} = x^{-1} \cdot x = 1.$$

## A KÉT MŰVELETET ÖSSZEKAPCSOLÓ AZONOSSÁG

- (9) minden
- $x, y, z \in X$
- elemre teljesül a disztributív tulajdonság:

$$x \cdot (y + z) = x \cdot y + x \cdot z.$$

Most pedig lássunk néhány példát. A racionális számok  $\mathbb{Q}$  összessége az előző évek során szerzett ismereteink szerint pontosan ilyen struktúra azaz test. Ugyancsak testet alkot a valós számok halmaza a jól ismert összeadás és szorzás műveletekkel. A komplex számok is testet alkotnak, a műveletek definiálásakor éppen arra ügyeltünk, hogy a kapott struktúra test legyen.

Meglepő, hogy a 2.13. példában leírt struktúra is test, és ehhez az ott modottak után csak a disztributivitás szorul belátásra, amit az olvasó könnyen ellenőrizhet.

A testek műveleteivel a leírt műveleti tulajdonságok alapján nagyon eredményesen lehet számolni. A számolásnál használt apró szabályokat (például az előjel-szabályokat) nem igazoljuk a tárgyalásban, de az olvasónak ajánljuk azok elvégzését.

Emlékeztetőül, hogy helyesen helyezzük el megszokott fogalmainkat a mostani rendszerben, szóljunk még néhány szót a kivonás és az osztás műveletekről.

Könnnyen beláthatjuk, hogy egy testben minden  $a$  és  $b$  elemre van megoldása az

$$x + a = b$$

egyenletnek. Adjuk ugyanis hozzá az egyenlet mindkét oldalához az  $a$  elem  $-a$  el-lentettjét, és alkalmazva a testben teljesülő azonosságokat, számoljunk az alábbiak szerint

$$(x + a) + (-a) = b + (-a), \quad x + (a + (-a)) = b + (-a),$$

$$x + 0 = b + (-a), \quad x = b + (-a).$$

Az egyenlet megoldására adódott  $(b + (-a))$  alakot  $b - a$  módon is szokás írni, de ezen utóbbi esetben a “ $-$ ” jel már nem az  $a$  szám negatívját jelenti, hanem a következő utasítást adja: Vedd az  $a$  szám negatívját és add hozzá a  $b$  számhoz. Ezt az utasítást, mint az  $(a, b) \mapsto b + (-a)$  leképezést, a  $b$  és  $a$  számok *különbségének* nevezzük.

A műveletre adott definíciónk szerint a kivonás is művelet, azt mondhatnánk, hogy az összeadás inverz művelete, de a műveletekre adott speciális tulajdonságok közül egyikkel sem rendelkezik, ezért másodlagosnak is kezeljük az összeadással összevetve.

Az elmondottakhoz teljesen azonos gondolatokat lehetne leírni a szorzás inverz műveletére, az *osztásra*, amire az  $a/b = \frac{a}{b}$ , ( $b \neq 0$ ) írásmód a szokásos, amit törtnek nevezünk. Ez természetesen a megoldása az

$$a \cdot x = b, \quad a \neq 0$$

egyenletnek.

## 2.4 Valós függvények

A jelen pont a valós függvények bevezetésével foglalkozik. Az első alpontban a legfontosabb általánosságokat mondjuk el. A második és harmadik pontban az olyan függvényekkel foglalkozunk, amelyek a test műveletei segítségével építhetők fel, a polinomokkal és a racionális törtfüggvényekkel. Ezekhez elégséges lett volna a racionális számok ismerete is, mivel csak a test-műveletek az építkezés alapeszközei.

A negyedik és ötödik alpontban a logaritmus, exponenciális és trigonometrikus függvényeket vezetjük be. Ezek nem építhetők fel a test műveletek véges alkalmazásával, ezek csak a valós számok rendezésre nézve teljes test struktúrájának a birtokában definiálhatók kielégítően.

### 2.4.1 Valós függvények bevezetése

A függvény fogalmát már az első fejezetben definiáltuk, és most az első definícióban nem teszünk mást, csak elnevezzük azokat a függvényeket, amelyeknek az értelmezési tartománya és az értékkészlete a valós számok valamilyen részhalmaza.

**Definíció 73** *Legyen az  $A$  egy tetszőleges halmaz. Egy, az  $A$  halmazon értelmezett, valós értékű — szimbólikusan:  $A \rightarrow \mathbb{R}$  — leképezést valós értékű függvénynek fogunk nevezni.*

*Amennyiben az  $A$  értelmezési tartomány is a valós számok halmazának egy részhalmaza, akkor valós változós, valós értékű függvényünk van, amit röviden valós függvénynek fogunk mondani.*

Bár a függvény halmazelméleti definíciója után minden fontos gondolatot elmondtunk általában a leképezésekről, a legfontosabbakat újra felidézzük.

Legyen az  $A$  tetszőleges halmaz, a  $B$  pedig a valós számok egy részhalmaza, és tekintsük az  $f : A \rightarrow B$  függvényt. Az  $f$  értelmezési tartománya az egész  $A$  halmaz, de az  $f(A)$  értékkészlete nem feltétlenül a teljes  $B$  halmaz, hanem annak csak egy részhalmaza.

A függvény definíciójára visszaemlékezve az  $f$  függvény az  $A$  halmaz minden  $x \in A$  eleméhez hozzárendel egy jól definiált  $f(x) \in B$  valós számot. Hangsúlyozzuk, hogy az  $f$  egyetlen objektum, és nem keverendő össze az  $f(x)$  helyettesítési értékével.

A függvény definíciója szerint az  $f$  az  $A \times B$  szorzat halmaz egy megfelelő részhalmaza, tehát — valós függvény esetében ( $A \subseteq \mathbb{R}$ ) — részhalmaza az  $\mathbb{R}^2$  síknak. Formálisan:

$$f = \{(x, f(x)) : x \in A\} \subseteq A \times B \subseteq \mathbb{R} \times \mathbb{R} = \mathbb{R}^2.$$

Minden valós függvény az  $\mathbb{R}^2$  részhalmaza, és az  $\mathbb{R}^2$  sík szokásos szemléletes geometriai modelljén fel tudjuk rajzolni a függvényt. Emiatt az  $\mathbb{R}^2$  síknak az

$$\{(x, f(x)) : x \in A\}$$

részhalmazát az  $f$  függvény gráfjának (grafikonjának) is szokás mondani, de — ahogyan ezt már a függvény halmazelméleti definíciójánál is mondtuk — voltaképpen ez a halmaz pontosan az  $f$  függvény maga, és a gráf szó csak a grafikus szemléltre utaló szószaporítás, amit ennek ellenére nem fogunk elvetni.

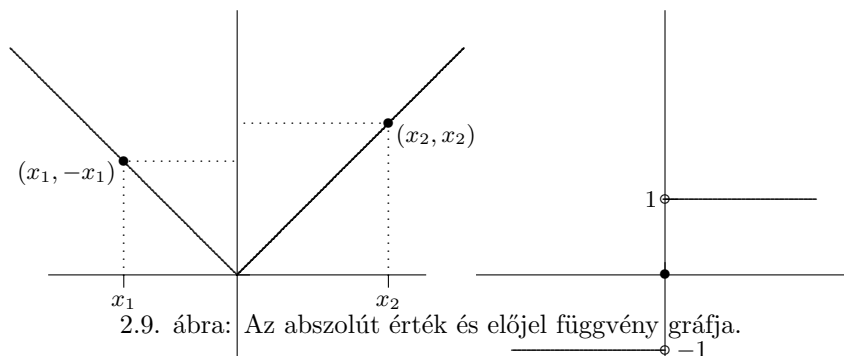
Ha az  $A$  halmaz helyett annak egy részhalmazát vesszük csak, akkor a függvény is megváltozik, például az alábbi két függvény nyilvánvalóan különböző:

$$x \mapsto x^2, \quad x \in \mathbb{R}_+,$$

$$x \mapsto x^2, \quad x \in \mathbb{R}.$$

Az első egy  $\mathbb{R}_+ \rightarrow \mathbb{R}$  függvény, az  $\{(x, x^2) : x \in \mathbb{R}_+\}$  halmaz; a második pedig egy  $\mathbb{R} \rightarrow \mathbb{R}$  függvény, az  $\{(x, x^2) : x \in \mathbb{R}\}$  halmaz.

A  $B$  halmaz megváltoztatása viszont látszólag nem változtatja meg feltétlenül a függvényt, hiszen annak a megadásánál csak arra kell tekintettel lenni, hogy tartalmaznia kell az  $f(A)$  értékkészletet. Ez a megállapítás azonban nem helyes, mert lényeges az  $f(A)$  és  $B$  viszonya is, hiszen ha például az  $f(A)$  halmazt vennénk  $B$ -nek, akkor szurjektív a függvény.



2.9. ábra: Az abszolút érték és előjel függvény gráfja.

Néhány fontos függvényt külön el is nevezünk a következő definícióban.

**Definíció 74** (1) *Előjel függvény.*  $\text{sgn} : \mathbb{R} \rightarrow \mathbb{R}$ :

$$\text{sgn}(x) \stackrel{\text{def}}{=} \begin{cases} -1 & \text{ha } x < 0 \\ 0 & \text{ha } x = 0 \\ 1 & \text{ha } x > 0 \end{cases}.$$

(2) *Abszolút érték függvény.*  $|\cdot| : \mathbb{R} \rightarrow \mathbb{R}_+$ :

$$|x| \stackrel{\text{def}}{=} \begin{cases} -x & \text{ha } 0 \geq x \\ x & \text{ha } x < 0 \end{cases}.$$

A két definiált függvényt a 2.9. ábrán rajzoltuk fel.

A következő definícióban definiáljuk a valós függvények közötti legfontosabb műveleteket.

**Definíció 75** Egy nemüres  $A \subseteq \mathbb{R}$  halmazon értelmezett valós értékű, azaz az  $A \rightarrow \mathbb{R}$  függvények között négy műveletet definiálunk.

1. *Összeadás:*

$$(f + g)(x) \stackrel{\text{def}}{=} f(x) + g(x), \quad \text{minden } x \in A \text{ elemre.}$$

2. *Szorzás:*

$$(f \cdot g)(x) = (fg)(x) \stackrel{\text{def}}{=} f(x) \cdot g(x), \quad \text{minden } x \in A \text{ elemre.}$$

3. *Számmal való szorzás:*

$$(\alpha \cdot f)(x) = (\alpha f)(x) \stackrel{\text{def}}{=} \alpha \cdot f(x), \quad \text{minden } x \in A \text{ elemre.}$$

4. *Hányados, feltéve, hogy a nevező nem nulla semmilyen  $x \in A$  elemre sem:*

$$\left(\frac{f}{g}\right)(x) \stackrel{\text{def}}{=} \frac{f(x)}{g(x)}, \quad \text{minden } x \in A \text{ elemre.}$$

5. Vegyünk most egy olyan  $f$  függvényt, amelyik az  $A$  halmazból a  $B$  halmazba képez, és egy olyan  $g$  függvényt, amelyik a  $B$  halmazból a  $C$  halmazba képez. Ekkor definiálni lehet az  $g \circ f$   $A$  halmazból a  $C$  halmazba képező közvetett függvény vagy kompozíciót a következőképpen:

$$(g \circ f)(x) \stackrel{\text{def}}{=} g(f(x)), \quad \text{minden } x \in A \text{ elemre.}$$

Már sokszor hangsúlyoztuk, de még egyszer megteesszük: A függvényekkel mint objektumokkal végzünk műveleteket, és az eredmények is mindig függvény objektumok. Emiatt például csak azonos értelmezési tartományú függvényeket adhatunk össze.

Nagyon fontos észrevétel az, hogy a műveletekkel függvényekből álló formulákat, új függvényeket tudunk készíteni, Így sokszor hivatkozunk úgy is ezekre, mint *függvényépítési eljárásokra*.

A pont hátralévő részében még két fogalmat definiálunk.

**Definíció 76** Legyen az  $A$  a valós számok egy nemüres részhalmaza. Azt mondjuk, hogy az  $f : A \rightarrow \mathbb{R}$  függvény monoton növekedő a nemüres  $B \subseteq A$  halmazon, ha minden

$$x_1, x_2 \in B, \quad x_1 \leq x_2$$

pontpárra

$$f(x_1) \leq f(x_2).$$

Ha az  $x_1 < x_2$  (szigorú) egyenlőtlenség (szigorú)  $f(x_1) < f(x_2)$  egyenlőtlenséget von maga után, akkor szigorú monoton növekedésről beszélünk. Teljesen analóg a monoton fogyás (csökkenés) és a szigorú monoton fogyás (csökkenés) definíciója.

Figyeljünk fel arra, hogy a monoton növekedő függvény olyan tulajdonságú leképezés, amelyik megőrzi a rendezés struktúráját abban az értelemben, hogy nagyobb számot nagyobbra képez; a szigorú monoton növekedés még jobban megőrzi a rendezés struktúráját, mivel szigorúan nagyobb számokat szigorúan nagyobbakra képez.

**Példa 2.15** *Vizsgáljuk meg az alábbi függvényeket monotonitási szempontból.*

1.  $x \mapsto x^2$ ,  $x \in \mathbb{R}$ .
2.  $x \mapsto |x|$ ,  $x \in \mathbb{R}$ .
3.  $x \mapsto \operatorname{sgn}(x)$ ,  $x \in \mathbb{R}$ .

Az első függvény a számegyenes pozitív részén szigorúan monoton nő, mivel nagyobb számok négyzete is nagyobb. A negatív része felett monoton fogy, hiszen a gráfja az  $y$  tengelyre szimmetrikus.

A második leképezés ugyanúgy viselkedik, mint az előző függvény.

A harmadik leképezés az egész egyenesen monoton növekedő, de nem szigorúan.

□

Szigorúan monoton növekedő függvények inverze is szigorúan monoton növekedő, a következő tétel pontos fogalmazása szerint.

**Állítás 77** *Legyen  $A \subseteq \mathbb{R}$ . Ha az  $f : A \rightarrow B$  szürjektív leképezés szigorúan monoton növekedő (fogyó), akkor létezik az  $f^{-1} : B \rightarrow A$  inverz-függvény, és az is szigorúan monoton növekedő (fogyó).*

**Bizonyítás.** Nézzük, mondjuk, a szigorú monoton növekedés esetét.

Az inverz létezése: Az  $f$  leképezés a feltétel szerint szurjektív, a szigorú monoton növekedése miatt pedig injektív. Ezek szerint az  $f$  bijektív és így van inverze.

Az  $f^{-1}$  szigorúan monoton növekedő: Legyen

$$y_1, y_2 \in B, \quad \text{és} \quad y_1 < y_2.$$

Ha  $f^{-1}(y_1) \geq f^{-1}(y_2)$  lenne, akkor az  $f$  monoton növekedése miatt

$$f(f^{-1}(y_1)) \geq f(f^{-1}(y_2)), \quad \text{azaz} \quad y_1 \geq y_2$$

adódnék, ellentétben az  $y_1 < y_2$  kiindulással, tehát

$$f^{-1}(y_1) < f^{-1}(y_2),$$

azaz az  $f^{-1} : B \rightarrow A$  szigorúan monoton növekedő. □

**Definíció 78** *Legyen az  $A$  egy tetszőleges nemüres halmaz, és az  $f : A \rightarrow \mathbb{R}$  egy valós értékű függvény. Az  $f$  függvényt felülről korlátosnak mondjuk, ha az  $f(A)$  értékkészlet felülről korlátos halmaz. Analóg az alulról való korlátosság fogalma. Korlátosnak mondjuk a függvényt, ha mind alulról, mind felülről korlátos.*

A felülről való korlátosság ezek szerint azt jelenti, hogy van olyan  $K$  szám, felső korlát, hogy

$$f(x) \leq K, \quad \text{minden } x \in A \text{ helyen.}$$

A korlátosság — ami a meghatározás szerint az  $f(A)$  valós számhalmaz korlátossága — azt jelenti, hogy van olyan  $K_1$  és  $K_2$  pozitív valós szám, amelyekre

$$-K_1 \leq f(x) \leq K_2, \quad \text{minden } x \in A \text{ helyen.}$$

A  $K_1$  és a  $K_2$  számok közül a nagyobbikat  $K$ -val jelölve a korlátosság így is írható:

$$|f(x)| \leq K, \quad \text{minden } x \in A \text{ helyen.}$$

**Példa 2.16** Vizsgáljuk meg a következő függvényeket korlátossági szempontból

1.  $x \mapsto 1/x, \quad 0 < x \leq 1.$
2.  $x \mapsto |x|, \quad -1 \leq x \leq 1.$
3.  $x \mapsto x^2, \quad x \in \mathbb{R}_+$

Az első függvény legalább 1, mivel a pozitív számok reciprokfüggvénye csökkenő függvény, ezért alulról korlátos. Felülről viszont nem korlátos, mert a nulla felé közeledve tetszőlegesen nagy lehet a függvény. Például az  $x = 1/n$  helyen az értéke  $n$ , ahol az  $n$  tetszőlegesen nagy egész szám.

A második függvény mind alulról, mind felülről korlátos, hiszen — nézve a grafikonját (2.9. ábra) — a megkötött szakasz felett egynél nem nagyobbak az értékei.

A harmadik függvény alulról korlátos, hiszen a függvénynek minden értéke nem kisebb a nullánál, mert a négyzet nemnegatív. Felülről viszont nem korlátos, mert az  $x^2$  tetszőlegesen nagy szám fölé nőhet, hiszen  $x < x^2$  ha  $x > 1$ .  $\square$

A valós számok korlátos részhalmazaira bevezettük az infimum (alsó határ) és szuprémum (felső határ) fogalmakat. Ezeknek a közvetlen alkalmazásaként, felülről korlátos illetve alulról korlátos függvényekre definiálni tudjuk az alábbi analóg fogalmakat.

**Definíció 79** Legyen az  $f : A \rightarrow \mathbb{R}$  egy felülről korlátos függvény. Az  $f$  függvény szuprémumának mondjuk az  $f(A)$  értékkészlet  $\sup f(A)$  felső határát, amire a

$$\sup_{x \in A} f(x)$$

jelölést használjuk, illetve ha félreértéstől nem kell tartanunk, akkor röviden csak  $\sup f$ -et írunk. Teljesen analóg az

$$\inf_{x \in A} f(x)$$

függvény-infimum definíciója alulról korlátos függvényre.

Nyilvánvalóan megállapíthatjuk, hogy egy korlátos függvénynek van mind infimuma mind szuprimuma.

A függvény szuprimumával és infimumával később gyakorta fogunk okoskodni, és a gondolat magja — például a felső határt véve — sok esetben a következő lesz:

A  $\sup_{x \in A} f(x)$  szám a következő tulajdonságokkal rendelkezik:

(a) Felső korlátja a függvény értékeinek, azaz

$$f(x) \leq \sup_{x \in A} f(x), \quad \text{minden } x \in A \text{ elemre.}$$

(b) A legkisebb felső korlát, azaz ha tetszőleges nála kisebb

$$\sup_{x \in A} f(x) - \epsilon, \quad \epsilon > 0$$

számot veszünk, akkor van olyan  $x' \in A$  hely, ahol a függvényérték nagyobb ennél az értéknél:

$$f(x') > \sup_{x \in A} f(x) - \epsilon.$$

**Példa 2.17** Legyenek az  $f : A \rightarrow \mathbb{R}$  és  $g : A \rightarrow \mathbb{R}$  függvények felülről korlátosak. Mutassuk meg, hogy

$$\sup_{x \in A} (f + g)(x) = \sup_{x \in A} (f(x) + g(x)) \leq \sup_{x \in A} f(x) + \sup_{x \in A} g(x).$$

Először is jegyezzük meg, hogy az  $f + g$  összeg függvény is felülről korlátos, hiszen egy felső korlát lesz a két függvény egy-egy felső korlátjának az összege, ezért lehet beszélni az összeg függvény szuprimumáról is.

Elégséges megmutatni, hogy a  $\sup f + \sup g$  felső korlátja az  $f + g$  függvénynek, mert ekkor a  $\sup(f + g)$  felső határ annál csak kisebb egyenlő lehet. A  $\sup f$  felső korlátja az  $f$ -nek, a  $\sup g$  pedig a  $g$ -nek, azaz

$$f(x) \leq \sup f \quad \text{és} \quad g(x) \leq \sup g,$$

minden  $x \in A$  pontra. Ezeket az egyenlőtlenségeket összeadva azt kapjuk, hogy minden  $x \in A$  elemre

$$f(x) + g(x) \leq \sup f + \sup g,$$

amivel be is láttuk amit akartunk.

Meg kell jegyeznünk, hogy általában  $\sup(f + g)$  kisebb, mint  $\sup f + \sup g$ , amint azt az alábbi példa mutatja: Legyen  $f : x \rightarrow -x^2 + 1$   $x \in (-1, 1)$  és  $g : x \rightarrow x - 1$   $x \in (-1, 1)$ . Akkor  $\sup f = 1$  és  $\sup g = 0$ , míg  $\sup(f + g) = \frac{1}{4}$ . Hasznos elnevezéseket rögzítünk a következő definícióban.

**Definíció 80** Ha egy  $f : A \rightarrow \mathbb{R}$  függvénynek van szuprémuma, és az eleme az  $f(A)$  értékkészletnek, akkor a szuprémumot maximumnak nevezzük, jelölésben:

$$\sup_{x \in A} f(x) = \max_{x \in A} f(x) = f(x_0), \quad x_0 \in A.$$

Hasonló a minimum fogalom definíciója.

Az  $f$  függvénynek pontosan akkor van maximuma, ha az  $f(A)$  értékkészlet felső határa benne van az értékkészletben, azaz van olyan  $x' \in A$  hely, ahol  $f(x') = \max f = \sup f$ . Ilyenkor az a szóhasználat is szokásos, hogy az  $f$  felveszi a maximumát az  $A$  halmazon.

## 2.4.2 Polinomok és racionális törtfüggvények

Az előző pontban definiáltunk néhány fontos valós függvényt, de — noha példákban használtunk később tárgyalandó leképezéseket — még nem foglalkoztunk azzal, hogy a valós számok test volta milyen függvények definiálását teszi lehetővé.

Az előző pontban definiált, függvényekre vonatkozó, műveleti szabályokat (75. definíció) a test-műveletek alapján definiáltuk, és ezek úgy foghatók fel, mint *függvény-építési szabályok*, amelyek segítségével egyszerű leképezésekből viszonylag bonyolult függvények egész seregét építhetjük fel.

Legyen az  $A$  a valós számok egy nemüres részhalmaza. Ahhoz, hogy a test adta lehetőségek kihasználásával felépíthető  $f : A \rightarrow \mathbb{R}$  függvényeket előállíthassuk, elégséges két egyszerű függvényből kiindulni:

- (1) Az azonosan 1, állandó (konstans) leképezés:

$$x \mapsto 1, \quad \text{minden } x \in A \text{ elemre.}$$

- (2) Az  $A \rightarrow A$  azonosság-leképezés ( $\text{id}_A$ ):

$$x \mapsto x, \quad \text{minden } x \in A \text{ elemre.}$$

Most pedig nézzük meg, hogy milyen függvények állíthatók elő ezen két függvényből a függvények közötti — összeadás, szorzás és osztás — műveletek segítségével.

A számmal való szorzás műveletét felhasználva az (1) alatti konstans leképezést tetszőleges  $c \in \mathbb{R}$  számmal szorozva adódik az

$$x \mapsto c, \quad \text{minden } x \in A \text{ elemre}$$

általános állandó (konstans) leképezés.

A függvények szorzási szabálya alapján a (2) azonosság-leképezést önmagával szorozva megkapjuk az  $x \mapsto x^2$  függvényt; ezt szorozva a (2) azonosság-leképezéssel adódik az  $x \mapsto x^3$  függvény. Folytatva ezt az eljárást teszőleges  $n$  pozitív egész számra megkaphatjuk az

$$x \mapsto x^n, \quad \text{minden } x \in A \text{ elemre}$$

*pozitív egész kitevős hatványfüggvényt.*

A hatványfüggvényeket számmal szorozva, és a kapott függvényeket összeadva eljuthatunk a következő alakú függvényekhez:

$$x \mapsto a_1 \cdot x + a_2 \cdot x^2 + \cdots + a_n \cdot x^n.$$

És ehhez már csak az  $x \mapsto a_0$  konstans leképezést kell hozzáadni, hogy megkapjuk az

$$x \mapsto a_0 + a_1 \cdot x + a_2 \cdot x^2 + \cdots + a_n \cdot x^n. \quad (2.23)$$

alakú függvényeket, amiket polinomoknak (polinom függvényeknek) szokás nevezni.

A harmónia kedvéért az  $1 = x^0$  megállapodással szokásos élni, ami könnyen beláthatóan összhangban van a hatványfüggvényeknél megszokott számolásunkkal. Ezzel a konvencióval a polinom a szumma jel segítségével így írható:

$$\sum_{i=0}^n a_i x^i.$$

**Definíció 81** Az  $a_0, a_1, \dots, a_n$  valós számok és az  $x$  szimbólum (független változó) segítségével

$$x \mapsto a_0 + a_1 \cdot x + a_2 \cdot x^2 + \cdots + a_n \cdot x^n$$

$$(x \mapsto a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \cdots + a_1 \cdot x + a_0)$$

formában megadott,  $\mathbb{R} \rightarrow \mathbb{R}$  valós függvényeket polinomoknak fogjuk nevezni. Az  $a_0, a_1, \dots, a_n$  valós számokat a polinom együtthatóinak mondjuk.

A legmagasabb fokú nemnulla együtthatójú hatvány  $a_n$  együtthatójának az elnevezése: főegyüttható.

Két polinomot akkor mondunk azonosnak, ha az összes megfelelő együtthatójuk megegyezik.

Egy polinom foksámának azt a legnagyobb kitevőt mondjuk, amelyiknek az együtthatója nem nulla, a konstans polinom foksámát nullának vesszük.

Azt a polinomot, amelynek mindegyik együtthatója nulla, nulla polinomnak nevezzük.

Ha egy  $\alpha$  helyen egy  $p$  polinom  $p(\alpha)$  helyettesítési értéke nulla, akkor az  $\alpha$  számot a  $p$  polinom gyökének mondjuk.

Az a megállapodás, hogy a konstans polinom fokszámát nullának vettük, összhangban van azzal, hogy az  $x$  változó nulladik hatványa megállapodás szerint egy, feltéve, hogy  $x \neq 0$ .

Lássuk most, hogy az osztás műveletét is használva milyen függvényeket állíthatunk elő.

**Definíció 82** Legyen  $a$

$$p(x) = a_n x^n + \cdots + a_1 x + a_0$$

és

$$q(x) = b_m x^m + \cdots + b_1 x + b_0$$

két polinom. Minden olyan  $x \in \mathbb{R}$  helyen, ahol  $q(x) \neq 0$  definiálható az

$$x \mapsto \frac{p(x)}{q(x)} = \frac{a_n x^n + \cdots + a_1 x + a_0}{b_m x^m + \cdots + b_1 x + b_0}$$

*hányados-függvény, amit racionális törtfüggvénynek szokás nevezni.*

A racionális törtfüggvények összessége a legbővebb olyan függvényosztály, amit a testbeli műveletekkel fel lehet építeni. Természetes, hogy ezek felépítése minden testben lehetséges, és emiatt még nem lett volna szükség arra, hogy a valós számok rendezésre nézve teljes testjét bevezessük.

Racionális törtfüggvény értelmezési tartománya nyilvánvalóan nem tartalmazhatja a nevező gyökeit.

### 2.4.3 A hatvány, exponenciális és logaritmus függvény.

Tetszőleges  $a$  valós szám és nemnegatív  $n$  egész szám esetében a test műveletei alapján definiált az  $a^n$  hatvány. Az  $n = 0$  esetben megállapodás szerint azt vesszük, hogy  $a^0 = 1$ , ha  $a \neq 0$ .

Ha az  $n$  egész számot negatívnak is meg akarjuk engedni, akkor az

$$a^{-n} \stackrel{\text{def}}{=} \frac{1}{a^n}$$

definíció szerint ki kell zárunk azt, hogy az  $a$  szám nulla lehessen.

Ha tovább akarjuk terjeszteni a kitevőben szereplő számok körét, akkor ki kell zárunk a negatív  $a$  számokat is, mert például az  $(-1)^{1/2}$  olyan szám lenne, aminek a négyzete  $-1$ , ilyen valós szám pedig nincs.

A célunk az, hogy valós kitevőre is ki tudjuk terjeszteni a hatványozást, amit az előzőek szerint csak pozitív alap mellett remélhetünk. Ilyen függvény egyelőre csak az egész számokra definiált. Lássuk most az egész kitevőjű hatványozás

legfontosabb tulajdonságait, hogy tudjuk, mit is szeretnénk elvárni a hatvány általánosabb fogalmától:

Először is van két fontos műveleti szabály:

$$\begin{aligned} a^n \cdot a^m &= a^{n+m}, \\ (a^n)^m &= a^{nm}. \end{aligned}$$

Ezen szabályok közül az első a fontosabb, mert a második már következik belőle, ezért az első szabályt fogjuk meghatározónak tekinteni.

Fontos tulajdonság még, hogy ha  $a > 1$ , akkor nagyobb kitevő mellett a hatvány is nagyobb:

$$n < m \quad \text{esetében} \quad a^n < a^m.$$

Ezt (a kitevő szerinti) szigorú monotonitásnak mondhatnánk. Ha  $0 < a < 1$ , akkor fordított irányú a monotonitás.

Első lépésben értelmezzük a racionális kitevőjű hatvány fogalmát:

Ha  $a$  pozitív valós szám  $n$  pedig pozitív egész, akkor legyen

$$a^{1/n} \stackrel{\text{def}}{=} \sqrt[n]{a},$$

Tetszőleges  $r \in \mathbb{Q}$  racionális szám előállítható  $r = \frac{m}{n}$  alakban, ahol  $m$  egész,  $n$  pedig pozitív egész és legnagyobb közös osztójuk 1. Az  $a$   $r$ -edik hatványán az

$$a^r = a^{\frac{m}{n}} = (\sqrt[n]{a})^m$$

egyenlőséggel adott számot értjük.

Középiskolai tanulmányainkból jól tudjuk, hogy a racionális kitevőjű hatványok ilyen értelmezése mellett a hatványozás fent említett azonosságai érvényesek maradnak, csakúgy mint a kitevő szerinti szigorú monotonitás.

Ezekután legyen  $\alpha$  tetszőleges valós szám. Az  $a$  szám  $\alpha$  kitevőjű hatványán az

$$a^\alpha = \sup\{a^r \mid r \in \mathbb{Q}, r \leq \alpha\}$$

valós számot értjük.

Legyen  $\alpha_1, \alpha_2 \in \mathbb{R}$ . Minden  $r, s \in \mathbb{Q}$  racionális számpárra  $a^r \cdot a^s = a^{r+s}$ , és a felső határ egyértelműségéből következik, hogy

$$a^{\alpha_1} \cdot a^{\alpha_2} = a^{\alpha_1 + \alpha_2}.$$

A valós kitevőjű hatványozás monotonitása ugyancsak megkapható a racionális kitevőjű hatványozás monotonitásának és a racionális számok azon tulajdonságának a felhasználásával, hogy bármely két különböző valós szám között van racionális szám.

A valós kitevőjű hatványozás értelmezése lehetővé teszi két függvény definiálását:

Vegyük először azt, amikor az  $a$  rögzített és az  $x$  változik. Ekkor a kitevőtől függő leképezést kapunk; az ehhez tartozó elnevezéseket a következő meghatározásban rögzítjük.

**Definíció 83** A rögzített pozitív  $a$  mellett létező  $\mathbb{R} \rightarrow \mathbb{R}_+^0$  és  $x \mapsto a^x$  hozzárendelési szabállyal megadott függvényt  $a$ -alapú exponenciális (kitevő) függvénynek nevezzük.

Megjegyezzük, hogy helyenként használjuk az

$$x \mapsto \exp_a(x)$$

jelölést is az exponenciális függvényre.

Gyakorta—megengedhetően laza módon—az  $x \mapsto a^x$  helyett egyszerűen csak azt mondjuk, hogy az “ $a^x$  függvény”.

Az exponenciális függvények már a középiskolában megismert legfontosabb tulajdonságait az alábbi állításban foglaljuk össze.

**Állítás 84** Legyen az  $a$  egy pozitív valós szám. Ekkor van olyan az egész  $\mathbb{R}$  valós egyenesen definiált, pozitív értékű  $x \mapsto a^x$  függvény, amelyik teljesíti az alábbi feltételeket.

- (1) Minden valós  $x$  és  $y$  számokra

$$a^{x+y} = a^x \cdot a^y,$$

- (2) és

$$(a^x)^y = a^{xy}.$$

- (3) Ha  $a > 1$ , akkor szigorúan növekedő; ha  $a < 1$ , akkor szigorúan csökkenő; ha pedig  $a = 1$ , akkor azonosan 1.

- (4) Ha  $a \neq 1$ , akkor szürjektív, vagyis értékkészlete az  $\mathbb{R}_+^\circ$ .

Azt is megtehetjük, hogy az  $x$  változót képzeljük rögzítve és az  $a$  számot vesszük független változónak; ekkor kapjuk a hatványfüggvényt, a hatványt mint az alap függvényét. Ezt már az előző, megszokott jelölés felhasználásával adjuk meg.

**Definíció 85** Legyen  $a$   $b$  tetszőleges valós szám. Az  $x \mapsto x^b$ ,  $\mathbb{R}_+^\circ \rightarrow \mathbb{R}_+^\circ$  leképezést hatványfüggvénynek ( $a$  hatvány mint az alap függvénye) fogjuk nevezni.

A definiált két függvény nyilvánvalóan igen szorosan összefügg, és elégséges az egyik tanulmányozása ahhoz, hogy a másiktól is legyenek ismereteink. Az exponenciális függvényt szokás általában vizsgálni és mi is ezt tesszük.

Most pedig bevezetjük a logaritmus függvényt, feltéve, hogy  $a \neq 1$ . Ez már könnyű feladat, mert az  $x \mapsto a^x$  bijektív függvény inverzeként definiáljuk. Eszerint tetszőleges  $y$  pozitív számra egyetlen megoldása van az

$$y = a^x$$

egyenletnek. Ennek a megoldásnak a szokásos jelölése:

$$x = \log_a y.$$

Az így definiált

$$y \mapsto \log_a y, \quad \mathbb{R}_+^\circ \rightarrow \mathbb{R}$$

leképezés egy olyan függvényt ad, amelynek a tulajdonságai már adódnak az előző tételekből.

A következő állításban összefoglaljuk ezzel az új függvénnyel kapcsolatos tudnivalókat.

**Állítás 86** *Legyen az  $a \neq 1$  egy rögzített pozitív szám. Ekkor az  $x \mapsto a^x$ ,  $\mathbb{R} \rightarrow \mathbb{R}_+^\circ$   $a$ -alapú bijektív exponenciális függvénynek van inverze, amit  $a$ -alapú logaritmus függvénynek mondunk. A  $\log_a : \mathbb{R}_+^\circ \rightarrow \mathbb{R}$  logaritmus függvény tulajdonságai:*

(i) *Az  $\exp_a$  és  $\log_a$  egymás inverzei, azaz formálisan:*

$$\exp_a \circ \log_a = id_{\mathbb{R}_+^\circ} \quad \text{és} \quad \log_a \circ \exp_a = id_{\mathbb{R}},$$

*vagy másképpen kifejezve ugyanezt*

$$a^{\log_a x} = x, \quad x \in \mathbb{R}_+^\circ \quad \text{és} \quad \log_a a^x = x, \quad x \in \mathbb{R}.$$

(ii) *Minden  $u, v$  pozitív számra*

$$\log_a uv = \log_a u + \log_a v.$$

(iii) *Ha  $a < 1$ , akkor a logaritmus függvény szigorúan monoton fogyó, ha pedig  $a > 1$  akkor szigorúan monoton növekedő.*

(iv) *A logaritmus függvény  $\mathbb{R}_+^\circ \rightarrow \mathbb{R}$  bijektív leképezés.*

**Bizonyítás.** Az állítások mind közvetlen következményei a 84. állításnak.

(i): Az exponenciális függvény szürjektivitása és szigorú monotonitása miatt bijektív, ezért létezik az inverze. Az (i) alatt csak azt fejeztük ki két különböző formalizmussal, hogy az exponenciális és logaritmus függvények egymás inverzei.

(ii): Az exponenciális függvény bijektivitása miatt az  $u$  és  $v$  pozitív számokhoz van olyan  $x$  és  $y$  szám, hogy

$$u = a^x, \quad \log_a u = x \quad \text{és} \quad v = a^y, \quad \log_a v = y. \quad (2.24)$$

Ebből kiindulva az exponenciális függvény (1)-es tulajdonsága alapján, kihasználva azt is, hogy az exponenciális és logaritmus függvények egymás inverzei, azt kapjuk, hogy

$$\log_a uv = \log_a(a^x \cdot a^y) = \log_a a^{x+y} = x + y,$$

amiből a (2.24) szerint azonnal adódik a kívánt  $\log uv = \log u + \log v$  egyenlőség.

(iii): A 84. (3) és a 77. állítások alapján nyilvánvaló.

(iv): Bijektív függvény inverze is nyilvánvalóan bijektív, ezért ez az exponenciális függvény  $\mathbb{R} \rightarrow \mathbb{R}_+^\circ$  bijektivitásából adódik.  $\square$

Befejezésül egy tételben felsorolunk két gyakran használt azonosságot a logaritmus függvényre.

**Állítás 87** *Legyen az  $a \neq 1$  egy pozitív valós szám.*

(I)  $\log_a b^x = x \cdot \log_a b$ , feltéve, hogy  $b$  egy pozitív valós szám.

(II) Ha  $a, b \neq 1$  is egy pozitív valós szám, akkor

$$\log_b x = \frac{\log_a x}{\log_a b}.$$

Bizonyítás. (I): Az exponenciális függvény (2)-es tulajdonsága alapján:

$$a^{x \log_a b} = (a^{\log_a b})^x.$$

Ebből figyelembe véve azt, hogy logaritmus függvény definíciója szerint  $a^{\log_a b} = b$ , azt kapjuk, hogy

$$a^{x \log_a b} = b^x,$$

amiből pedig mindkét oldal  $a$ -alapú logaritmusát véve adódik a bizonyítandó

$$x \log_a b = \log_a b^x$$

azonosság.

(II): A logaritmus függvény definíciója szerint  $x = b^{\log_b x}$ . Véve mindkét oldal  $a$ -alapú logaritmusát és használva az (I) azonosságot az alábbi módon számolhatunk

$$\log_a x = \log_a (b^{\log_b x}) = (\log_b x) \cdot (\log_a b),$$

amiből azonnal adódik, hogy

$$\log_b x = \frac{\log_a x}{\log_a b},$$

$\square$

#### 2.4.4 Trigonometrikus függvények és inverzeik

A most tárgyalásra kerülő, trigonometrikus függvények geometriai eredetűek. A szinusz és koszinusz függvények geometriai definíciójánál abból indulunk ki, hogy vesszük az origó körüli egységnyi sugarú körvonalat. A körvonalnak az  $x$  radián szögnél lévő első koordinátája a koszinusz függvény értéke az  $x$  helyen, a második

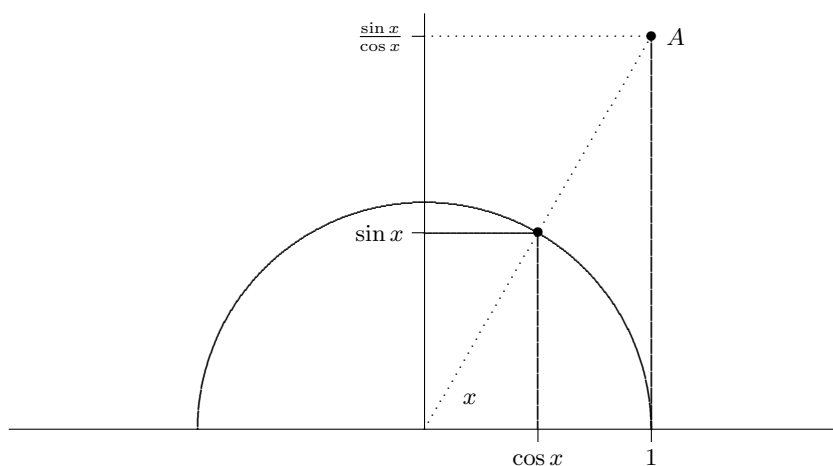
koordinátája pedig a szinusz függvénynek az  $x$  helyen felvett értéke. Ennek az alapján rajzoltuk meg a 2.10. ábrát.

Az ábrán egy egységi sugarú kört rajzoltunk fel, és egy  $x$  szögmértékű (radián) szöget, valamint a  $-x$  szöget. Ne feledjük, hogy az  $x$  voltaképpen a szög által megadott körív hosszát jelöli, tehát az  $(1, 0)$  és  $(\cos x, \sin x)$  koordinájú pontokat összekötő ív hossza  $x$ .

Az  $(1, 0)$  pontban a vízszintes tengelyre emelt merőleges az  $x$  szöget alkotó sugár meghosszabbítását az  $A$  pontban metszi. Fontos észrevétel, hogy a  $(1, 0)$  és  $A$  pontokat összekötő szakasz hossza  $\frac{\sin x}{\cos x}$ , amit azonnal beláthatunk abból, hogy a  $(0, 0)$ ,  $(\cos x, 0)$  és  $(\cos x, \sin x)$  pontok által meghatározott háromszög hasonló a  $(0, 0)$ ,  $(1, 0)$  és  $A$  pontok által meghatározott háromszöghöz, ezért a megfelelő oldalainak az aránya megegyezik:

$$(\cos x) : 1 = \sin x : \overline{(1, 0), A}.$$

Ezek szerint minden jelzést indokoltunk, amit az ábrára ráírtunk.



2.10. ábra: A szinusz és koszinusz definíciójának a szemléltetése

A szinusz és koszinusz függvények legfontosabb jellemzőit gyűjtöttük össze az alábbi állításban:

**Állítás 88** *A  $\sin$  és  $\cos$  szimbólumokkal jelölt, szinusz és koszinusznak nevezett,  $\mathbb{R} \rightarrow \mathbb{R}$  függvény, és a  $\pi$  valós számra teljesülnek a következő tulajdonságok.*

$$(1) \sin(x + 2k\pi) = \sin x \quad \text{és} \quad \cos(x + 2k\pi) = \cos x, \quad x \in \mathbb{R},$$

ahol a  $k$  tetszőleges egész szám.

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

$$(3) \sin(-x) = -\sin x \quad \text{és} \quad \cos(-x) = \cos x, \quad x \in \mathbb{R},$$

$$(4) \sin 0 = 0, \quad \cos 0 = 1 \quad \text{és} \quad \sin \frac{\pi}{2} = 1, \quad \cos \frac{\pi}{2} = 0,$$

$$(5) \sin(x + y) = \sin x \cos y + \cos x \sin y, \quad x, y \in \mathbb{R},$$

$$(6) \cos(x + y) = \cos x \cos y - \sin x \sin y, \quad x, y \in \mathbb{R},$$

$$(7) \sin^2 x + \cos^2 x = 1, \quad x \in \mathbb{R},$$

$$(8) 0 < \sin x < x < \frac{\sin x}{\cos x} \quad \text{ha} \quad 0 < x < \frac{\pi}{2}.$$

A szinusz és koszinusz függvényeket és az ezekből származtatott néhány függvényt hívják összefoglalólag *trigonometrikus függvényeknek*.

Most pedig sorra vesszük a tétel formuláit — kivéve az (5) és (6) addíciós formulákat, amiknek egy kicsit bonyolultabb a geometriai tartalma — és a 2.10. ábrára hivatkozva, rámutatunk a relációknak a geometriai definícióval való összefüggésére. Ezeket célszerű megjegyezni, mert itt viszonylagosan nagy a formulák száma, és ezen a módon segíthetünk a memóriánknak is.

**(1):** Ez az állítás, részletesebben írva, azt mondja, hogy a szinusz (koszinusz) függvény, tetszőleges  $x$  esetében azonos értéket vesz fel az

$$\dots, x - 2k\pi, \dots, x - 2\pi, x, x + 2\pi, \dots, x + 2k\pi, \dots$$

helyeken, azaz  $2\pi$  értékkel eltolva a helyettesítési helyeket a függvény értékei azonosak. Az ilyen tulajdonságú függvényre azt mondják, hogy  $2\pi$  szerint *periodikus*.

A  $2\pi$  szerinti periodicitás az ábránk szerint azt jelenti, hogy a szinusz és koszinusz függvények értéke nem függ attól, hogy hány teljes körbefordulás után jutunk az  $x$  szögnek megfelelő helyre.

**(2):** Eszerint a tulajdonság szerint a szinusz és koszinusz függvények szorosan összefüggenek, egyik igen egyszerűen megkapható a másiktól, csak a független változó egyszerű  $x \mapsto (\pi/2 - x)$  transzformációjára van szükség.

Emlékezzünk rá, hogy ez a függvények gráfjára nézve azt jelenti, hogy tükrözni kell a gráfot az origóra és el kell tolni vízszintesen  $(-\pi/2)$ -l. A (3) tulajdonság azonban azt mondja, hogy  $\sin(-x) = -\sin x$ , azaz a szinusz függvény gráfja szimmetrikus az origóra (az origóra való tükrözés önmagát adja), ezért a koszinusz függvény gráfja úgy kapható a szinusz gráfjából, hogy eltoljuk vízszintesen  $(-\pi/2)$ -l.

Az  $x$  szög kiegészítő szögét nézve, az állítás geometriailag teljesen evidens.

**(3):** A  $\sin(-x) = -\sin x$  azonosság azt jelenti, hogy a függvény gráfja szimmetrikus az origóra. Az ilyen függvényt *páratlan* függvénynek nevezzük. Ennek az elnevezésnek az az eredete, hogy a páratlan kitevőjű hatvány-függvények is ilyenek:  $(-x)^n = -x^n$ , ha az  $n$  páratlan.

A  $\cos(-x) = \cos x$  pedig azt jelenti a függvény gráfjára, hogy a függőleges koordináta-tengelyre nézve lesz szimmetrikus. Az ilyen függvényt *páros függvénynek*

mondják, aminek az eredete az, hogy ilyenek a páros kitevős hatvány-függvények:  $(-x)^n = x^n$ , ha az  $n$  páros.

Az ábránkon ez a reláció úgy jelenik meg, hogy ha az  $x$  szöget lefelé rajzoljuk, akkor kapjuk meg a  $-x$  szöget, és geometriailag világosan leolvasható az összefüggés.

**(4):** Ezeken a speciális helyeken felvett értékek a geometriai definíció alapján azonnal leolvashatóak az ábráról. Olvassuk le a  $\pi$ ,  $(3/2)\pi$  helyeken felvett értékeket is.

Ide tartozó megjegyzés, hogy a trigonometrikus függvények értékei általában nem számolhatók ki véges sok testbeli művelettel, sőt még a gyökjelek használatával sem, de néhány helyen megadhatóak a helyettesítési értékek gyökjelekkel, például az  $x = \pi/6$  helyen.

**(7):** A  $(0, 0)$ ,  $(\cos x, 0)$  és  $(\sin x, \cos x)$  derékszögű háromszögre alkalmazva a Pitagorasz tételt azonnal adódik.

**(8):** Az ábráról azonnal látszik, hogy a második koordináta, a  $\sin x$  pozitív, ha az  $x$  is a megadott számközben van.

Az egyenlőtlenség többi részének a következő szemléletes geometriai indoklásokat adhatjuk:

A  $(\cos x, 0)$  és  $(\cos x, \sin x)$  pontokat összekötő egyenes-szakasznál hosszabb az  $(1, 0)$  és  $(\cos x, \sin x)$  pontokat összekötő szakasz, mert ez utóbbi átfogója annak a derékszögű háromszögnek, amelyiknek az előző szakasz az egyik befogója. Az  $(1, 0)$  és  $(\cos x, \sin x)$  pontokat összekötő szakasznál hosszabb az  $(1, 0)$  és  $(\cos x, \sin x)$  pontokat összekötő ív  $x$  hossza, mert közismert (de nem pontos) geometriai állítás, hogy két pont között legrövidebb út az egyenes. Ezt az okoskodási menetet összefoglalva azt kaptuk, hogy

$$\sin x < x.$$

Emlékezzünk most arra, hogy az ábrán szereplő egységkör területe  $\pi$ , és így egy  $x$  nyílásszögű körcikk területe eléggé szemléletes elv alapján  $x/2$ . A  $(0, 0)$ ,  $(1, 0)$  és  $(\cos x, \sin x)$  pontok meghatározta körcikk, aminek a területe  $x/2$ , része a  $(0, 0)$ ,  $(1, 0)$  és  $A$  pontok által meghatározott háromszögnek. Ez utóbbinak a területe

$$\frac{1 \cdot \frac{\sin x}{\cos x}}{2} = \frac{1 \sin x}{2 \cos x},$$

és így, mivel a rész területe kisebb, azt kaptuk, hogy

$$\frac{x}{2} < \frac{1 \sin x}{2 \cos x},$$

azaz

$$x < \frac{\sin x}{\cos x};$$

ezzel a tulajdonságok geometriai szemléltetését és értelmezését befejeztük.

A következő állításban további nélkülözhetetlen azonosságokat sorolunk fel, amelyek beláthatók az előbbi tétel relációi alapján.

**Állítás 89** *Fennállnak a következő azonosságok.*

- (a)  $\sin 2x = 2 \sin x \cos x$ ,
- (b)  $\cos 2x = \cos^2 x - \sin^2 x$ ,
- (c)  $\sin^2 x = \frac{1 - \cos 2x}{2}$ ,
- (d)  $\cos^2 x = \frac{1 + \cos 2x}{2}$ ,
- (e)  $\sin x + \sin y = 2 \sin \frac{x+y}{2} \cos \frac{x-y}{2}$ ,
- (f)  $\sin x - \sin y = 2 \sin \frac{x-y}{2} \cos \frac{x+y}{2}$ ,
- (g)  $\cos x + \cos y = 2 \cos \frac{x+y}{2} \cos \frac{x-y}{2}$ ,
- (h)  $\cos x - \cos y = -2 \sin \frac{x+y}{2} \sin \frac{x-y}{2}$ ,

A bizonyításokat az olvasóra bízjuk.

A következő tételben a szinusz és koszinusz függvény előjel és nagysági viszonyairól mondunk ki egyszerű állításokat. Ezek az előző két tételre támaszkodva bizonyíthatók be. Az állítások geometriai szemlélet szerint meglehetősen evidensek.

**Állítás 90** *A szinusz és koszinusz függvényekre igazak az alábbi nagysági illetve előjel állítások.*

- (i)  $|\sin x| \leq 1$  és  $|\cos x| \leq 1$ .
- (ii) *A szinusz függvény pozitív, ha  $0 < x < \pi$ ; negatív ha  $\pi < x < 2\pi$ ; és nulla a 0 és  $\pi$  helyeken.*  
*A többi  $x$  helyen az előjelek a  $2\pi$  szerinti periodicitás miatt ebből már adódnak.*
- (iii) *A koszinusz függvény pozitív, ha  $0 < x < \pi/2$  vagy  $(3/2)\pi < x < 2\pi$ ; negatív ha  $\pi/2 < x < (3/2)\pi$ ; és nulla a  $\pi/2$  és  $(3/2)\pi$  helyeken.*  
*A többi  $x$  helyen az előjelek a  $2\pi$  szerinti periodicitás miatt ebből már adódnak.*
- (iv)  $|\sin x| < |x|$  ha  $-\pi/2 < x < \pi/2$ .

Most pedig fontoljuk meg, hogy mi a helyzet az esetlegesen létező inverz függvényekkel. Erre a problémára a válasz első lépésben teljesen negatív. A szinusz és koszinusz függvények periodikusak, ezért nincs inverzük. A szinusz függvény inverzének a létezése például azt jelentené, hogy egyértelmű  $x$  megoldása lenne az  $y = \sin x$  egyenletnek. Ha azonban egy  $x_0$  megoldás, akkor a  $2\pi$  szerinti periodicitás miatt az  $x_0 + 2\pi$  is megoldás lenne, sőt végtelen sok megoldást tudnánk mondani, ezért inverz nem létezik.

Ahhoz, hogy inverzről tudjunk beszélni, az  $\mathbb{R}$ -nél szűkebb értelmezési tartományt és értékkészletet kell választani, hogy a függvény bijektív legyen. Tekintve, hogy a  $-\pi/2 \leq x \leq \pi/2$  számközben a  $-1 \leq y \leq 1$  számközre bijektív a szinusz függvény, ezért ebben a számközben van inverze, amelyet arcsin-nak (arkusz szinusz) hívunk. Tehát  $\arcsin x$   $x \in [-1, 1]$  azt a  $[-\frac{\pi}{2}, \frac{\pi}{2}]$  intervallumba eső valós számot jelöli, amelynek szinusz  $x$ . Hasonlóan, a koszinusz függvény a  $[0, \pi]$  intervallumra való leszűkítése bijektív, amelynek inverze az arccos (arkusz koszinusz) függvény, azaz  $\arccos x$   $x \in [-1, 1]$  azt a  $[0, \pi]$  intervallumba eső valós számot jelenti, amelynek koszinusz  $x$ .

Az előző trigonometrikus függvényekkel további függvények definiálhatók. Még két függvény, a tangens és kotangens függvények használatosak leginkább. Mivel a kotangens függvény a tangens függvény reciproka, ezért fölösleges fogalom szaporításnak éreznénk mindkettőt bevezetni, így csak a tangens függvényt definiáljuk.

**Definíció 91** Minden  $x \neq (2k+1)\frac{\pi}{2}$  valós számra definiált a

$$x \mapsto \frac{\sin x}{\cos x}$$

függvény, amit a  $\tan$  szimbólummal fogunk jelölni.

A definíción értelmes, mert a definiáló tört nevezője, a koszinusz függvény pontosan a  $\pi/2$  páratlan számú többszöröseinél nulla. Erre a függvényre is kimondunk néhány azonosságot a következő tételben.

**Állítás 92** Ahol a tangens függvény értelmezve van, ott fennállnak a következő azonosságok.

1.  $1 + \tan^2 x = \frac{1}{\cos^2 x},$
2.  $\sin 2x = \frac{2 \tan x}{1 + \tan^2 x},$
3.  $\cos 2x = \frac{1 - \tan^2 x}{1 + \tan^2 x}.$

Az azonosságok bizonyítását az olvasóra bízunk.

Amint az jól ismert, a tangens függvény is periódikus, periódusa  $\pi$ , érték készlete  $\mathbb{R}$ , a  $(-\frac{\pi}{2}, \frac{\pi}{2})$  intervallumra való leszűkítése bijektív, amelynek inverze az arctan (arkusz tangens) függvény. Tehát  $\arctan x$   $x \in \mathbb{R}$  az a  $(-\frac{\pi}{2}, \frac{\pi}{2})$  intervallumba eső valós szám, amelynek tangense  $x$ .

## 3.

# Metrikus terek és leképezéseik

## 3.1 Metrikus terek

Ez a fejezet a metrikus terek és leképezéseik tulajdonságaival foglalkozik. Olyan tulajdonságokat gyűjtünk egybe, amelyek közös jellemzője, hogy távolság mérésével kapcsolatosak.

### 3.1.1 Példák metrikus terekre

Bevezetjük a távolság fogalmának absztrakt definícióját.

**Definíció 93** Legyen az  $X$  egy tetszőleges, nemüres halmaz és a  $d : X \times X \rightarrow \mathbb{R}$  egy olyan függvény, amelyre

- (1)  $d(x, y) \geq 0$  minden  $x, y \in X$  elemre, és pontosan akkor nulla, ha  $x = y$ .
- (2)  $d(x, y) = d(y, x)$ , minden  $x, y \in X$  elemre. (Szimmetria)
- (3)  $d(x, y) \leq d(x, z) + d(z, y)$ , minden  $x, y, z \in X$  elemre. (Háromszög-egyenlőtlenség).

Ekkor a  $d$  függvényt távolságnak (távolság-függvénynek) vagy metrikának mondjuk.

A  $d$  metrikával ellátott  $X$  halmazt metrikus térnek hívjuk, és röviden az  $(X, d)$  szimbólummal jelöljük.

**Példa 3.1** A valós számok  $\mathbb{R}$  halmaza metrikus tér a

$$d(x, y) = |x - y|$$

egyenlőséggel értelmezett metrikával.

Valóban, az abszolút érték tulajdonságai alapján a fenti  $d$  leképezés metrikát értelmez, hiszen az (1), (2) és (3) tulajdonságok könnyen ellenőrizhetők. Ezt a metrikát *euklideszi metrikának* nevezzük.

**Példa 3.2** *Mutassuk meg, hogy a*

$$s(x, y) \stackrel{\text{def}}{=} |x^3 - y^3|$$

*módon definiált leképezés metrika a valós számok  $\mathbb{R}$  halmazán.*

Ellenőrizzük először rendre a metrika tulajdonságainak a teljesülését.

- (1): Az  $s(x, y)$  nyilvánvalóan nemnegatív, és pontosan akkor nulla, ha  $x^3 = y^3$ . Ez utóbbi pedig pontosan akkor teljesül, ha  $x = y$ .  
 (2): A szimmetricitás nyilvánvaló, mivel  $|x^3 - y^3| = |y^3 - x^3|$ .  
 (3): A háromszög-egyenlőtlenség belátása az abszolút érték háromszög-egyenlőtlenségének a felhasználásával történik:

$$\begin{aligned} s(x, y) &= |x^3 - y^3| = |(x^3 - z^3) + (z^3 - y^3)| \\ &\leq |x^3 - z^3| + |z^3 - y^3| = s(x, z) + s(z, y). \end{aligned}$$

Tehát  $s$  valóban metrikát definiál. Ez a példánk azt mutatja, hogy egy halmazon többféle metrika is értelmezhető. (Lásd az előző példát.)  $\square$

**Példa 3.3** *Mutassuk meg, hogy tetszőleges nemüres  $X$  halmazon metrika a*

$$\delta(x, y) \stackrel{\text{def}}{=} \begin{cases} 1 & \text{ha } x \neq y \\ 0 & \text{ha } x = y \end{cases}$$

*módon definiált függvény. Ezt a metrikát diszkrét metrikának fogjuk nevezni.*

A metrika első és második tulajdonsága teljesen nyilvánvalóan teljesül. A háromszög-egyenlőtlenség indoklása: A

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

háromszög-egyenlőtlenségben a jobboldal csak akkor nulla, ha  $x = z = y$ , ekkor azonban a baloldal is nulla, tehát teljesül az egyenlőtlenség; ha pedig nem nulla a jobboldal, akkor legalább egy, és ezért teljesül.  $\square$

**Példa 3.4** Tekintsük a valós szám  $p$ -esek  $\mathbb{R}^p$  terét, azaz mindazon  $x = (x_1, \dots, x_p)$  vektorok terét, ahol az  $x_k$  elemek valós számok. Ezzel a térrel már találkoztunk az előző fejezetben. Vezessük be a

$$d(x, y) \stackrel{\text{def}}{=} \left( \sum_{k=1}^p (x_k - y_k)^2 \right)^{\frac{1}{2}} = \|x - y\|$$

metrikát. Ez a kifejezés valóban metrikát értelmez, hiszen az (1) és (2) tulajdonságok nyilvánvalóan teljesülnek, míg a (3) háromszög-egyenlőtlenség 68 folyománya. Ezt a metrikát az  $\mathbb{R}^p$  téren *euklideszi metrikának* nevezzük. További metrikára vonatkozó példákkal találkozunk az  $\mathbb{R}^p$  téren a következő szakaszokban.

Az alábbi definíció azt mutatja, hogy egy metrikus tér tetszőleges nem üres részhalmazát is metrikus térnek tekinthetjük.

**Definíció 94** Legyen az  $(X, d)$  egy metrikus tér, és a  $C$  az  $X$  halmaz egy nem üres részhalmaza. A  $d : X \times X \rightarrow \mathbb{R}$  távolság függvénynek az  $C \times C$  halmazra való  $d|_{C \times C}$  leszűkítése nyilvánvalóan egy metrika a  $C$  halmazon. Ezzel a metrikával ellátott  $C$  halmazt az  $X$  metrikus tér részterének nevezzük.

A leszűkített metrika függvényt is egyszerűen  $d$ -vel fogjuk jelölni, a  $d|_{C \times C}$  helyett, mert ez nem okoz félreértést, és így az  $(X, d)$  metrikus tér  $(C, d)$  részteréről (alteréről) fogunk beszélni.

### 3.1.2 Gömb-környezetek metrikus térben

Egy metrika segítségével — a geometriai analógia alapján — definiálhatjuk a “gömb” absztrakt fogalmát:

**Definíció 95** Legyen az  $(X, d)$  egy metrikus tér, az  $x$  egy tetszőleges pontja és  $r$  egy pozitív valós szám. Az  $x$  pont körüli  $r$  sugarú nyílt gömbnek mondjuk az

$$B^\circ(x, r) \stackrel{\text{def}}{=} \{y \in X : d(x, y) < r\}$$

halmazt. Az  $x$  pont körüli  $r$  sugarú zárt gömb pedig az

$$B(x, r) \stackrel{\text{def}}{=} \{y \in X : d(x, y) \leq r\}$$

halmaz.

**Példa 3.5** Mi a nyílt illetve zárt gömb a valós számok abszolút-érték hosszúsággal vett metrikus terében? (Lásd a 3.1 Példát.)

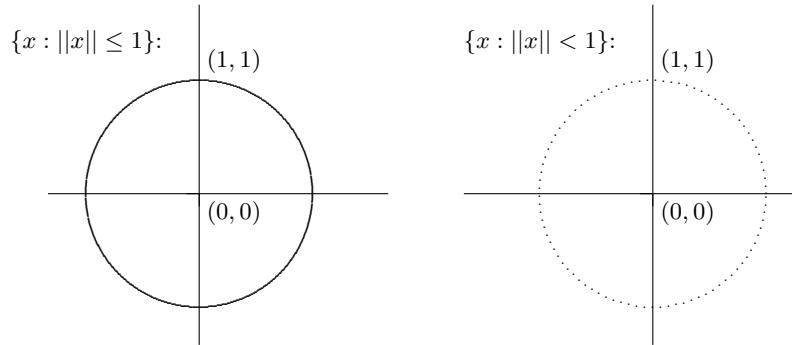
A  $d_e(x, y) = |x - y| < r$  egyenlőtlenség azzal ekvivalens, hogy  $-r < y - x < r$ , amiből átrendezéssel  $x - r < y < x + r$ . Az utóbbi egyenlőtlenség adja a választ: Az  $x$  pont körüli  $r$  sugarú nyílt gömb éppen az  $(x - r, x + r)$  nyílt intervallum.

Hasonlóan, a  $B(x, r)$  zárt gömb az  $x$  középpontú  $2r$  hosszúságú zárt intervallum:  $[x - r, x + r]$ .  $\square$

**Példa 3.6** Az  $\mathbb{R}^2$  euklideszi metrikus-térben mi lesz az 1 sugarú, origó középpontú, nyílt illetve zárt gömb? (Vesd össze a 3.4 Példával.)

A nyílt gömbhöz, definíció szerint, olyan  $y = (y_1, y_2)$  pontjait kell venni a síknak, amelyekre

$$\sqrt{(0 - y_1)^2 + (0 - y_2)^2} < 1 \quad \text{azaz} \quad y_1^2 + y_2^2 < 1.$$

3.1. ábra: Nyílt és zárt gömbök az  $\mathbb{R}^2$  euklideszi térben.

Eszerint az origó középpontú, egységnyi sugarú közösleges körlemezt kell venni, de el kell hagyni belőle az

$$\{(y_1, y_2) \in \mathbb{R}^2 : y_1^2 + y_2^2 = 1\}$$

körvonalat. Zárt gömb esetében a körvonal is beletartozik a gömbbe (3.1. ábra).  $\square$

**Példa 3.7** *Mi lesz az origó körüli 1 sugarú zárt gömb a síkon az alábbi metrika mellett?*

$$d((x_1, x_2), (y_1, y_2)) \stackrel{\text{def}}{=} |x_1 - y_1| + |x_2 - y_2|.$$

Azt, hogy a  $d$  függvény metrika, könnyen belátható. (Ennek belátását otthon feltétlenül végezzük el.) A definíció szerint a nulla körüli egységnyi sugarú zárt gömb a

$$B((0,0), 1) = \{(x_1, x_2) \in \mathbb{R}^2 : |x_1| + |x_2| \leq 1\}$$

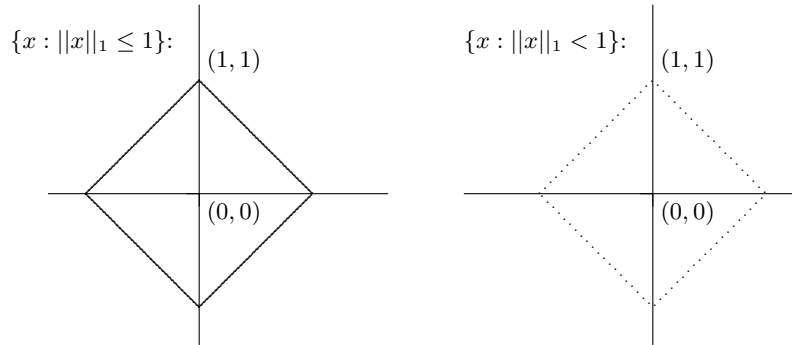
halmaz. Egy  $x = (x_1, x_2)$  pont pontosan akkor eleme ennek a halmaznak, ha

$$\begin{aligned} x_1 + x_2 &\leq 1, & \text{ha } x_1, x_2 &\geq 0 \\ -x_1 + x_2 &\leq 1, & \text{ha } x_1 &\leq 0 \text{ és } x_2 > 0 \\ -x_1 - x_2 &\leq 1, & \text{ha } x_1, x_2 &< 0 \\ x_1 - x_2 &\leq 1, & \text{ha } x_1 &\geq 0 \text{ és } x_2 < 0 \end{aligned}$$

. Ez alapján a zárt gömb az a négyzet, amelyiknek a csúcspontjai:

$$(1,0), (0,1), (-1,0), (0,-1).$$

A 3.2. ábrán rajzoltuk fel a nyílt és zárt gömböt, ami most—meglepetésre—négyzet alakú. Ez azt mutatja, hogy a definiált gömb fogalom olyan általános, hogy sok konkrét példa lehet mögötte.



3.2. ábra: Gömbök a  $d((x_1, x_2), (y_1, y_2)) = |x_1 - y_1| + |x_2 - y_2|$  metrika mellett.

**Definíció 96** Legyen az  $(X, d)$  egy metrikus tér. Az  $X$  egy  $x$  pontja nyílt (zárt) gömb-környezetének mondunk minden az  $x$  körüli nyílt (zárt) gömböt. Egy  $x$  pont nyílt (zárt) gömb-környezeteinek az összességét az  $x$  nyílt (zárt) gömb-környezet rendszerének fogjuk mondani.

Ha  $y \in B^\circ(x, r)$  ( $y \in B(x, r)$ ), akkor azt fogjuk mondani, hogy az  $y$  pont az  $x$  pont  $r$  sugarú nyílt (zárt) gömb-környezetében van.

Minden  $B^\circ(x, r)$  nyílt gömb-környezetekben van zárt gömb-környezet, például  $B(x, r/2) \subseteq B^\circ(x, r)$ . Megfordítva, minden  $B(x, r)$  zárt gömbi környezet tartalmaz nyílt gömb-környezetet, például:

$$B^\circ(x, r) \subseteq B(x, r).$$

Ezt az igen egyszerű észrevételt, fontossága miatt, egy állításban is megfogalmazzuk.

**Állítás 97** Bármely metrikus tér egy  $x$  pontjának a zárt és nyílt gömb-környezet-rendszere között a következő kapcsolat van:

Minden nyílt gömb-környezet tartalmaz zárt gömb-környezetet, és megfordítva: minden zárt gömb-környezet tartalmaz nyílt gömb-környezetet.

Ennek nyilvánvaló következménye, és csak gyakori használata miatt fogalmazzuk meg állításként:

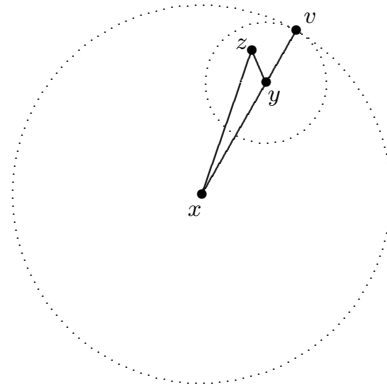
**Állítás 98** Ha az  $y$  az  $x$  egy nyílt gömb-környezetében van, akkor benne van egy zárt gömb-környezetében is, és megfordítva.

A következő tételben a nyílt gömbök egy fontos elemi tulajdonságát mondjuk ki.

**Állítás 99** Legyen az  $(X, d)$  egy metrikus tér. Ha az  $y$  egy tetszőleges pontja egy  $B^\circ(x, r)$  nyílt gömbnek, akkor van olyan  $y$  középpontú nyílt gömb, amelyik része a  $B^\circ(x, r)$  gömbnek.

Másképpen fogalmazva: *Nyílt gömb tartalmazza minden pontjának valamilyen gömb-környezetét.*

Az állítás nagyon szemléletes az  $\mathbb{R}$  és  $\mathbb{R}^2$  esetében, amikor az euklideszi távolságot vesszük. A 3.3. ábrán szemléltettük az állítást, amelyik szemléletesen meg is “adja” az igazolást, de a pontos bizonyítást is le kell írunk, mert olyan elvont fogalomról van szó, amelynél könnyen megtéveszthet a puszta szemlélet.



$$\begin{aligned} d(x, v) &= r \\ \alpha &= d(y, v) = r - d(x, y), \\ d(x, z) &\leq d(x, y) + d(y, z) < r. \end{aligned}$$

3.3. ábra: Nyílt gömb minden pontja körül tartalmaz nyílt gömböt.

**Bizonyítás.** A bizonyítást a már említett 3.3. ábrán követhetjük. Legyen a feltevés szerint  $y \in B^o(x, r)$ , azaz  $d(x, y) < r$ . Az  $\alpha = r - d(x, y)$  szám pozitív. Megmutatjuk, hogy a  $B^o(y, \alpha)$  gömb része a  $y \in B^o(x, r)$  gömbnek.

Ha  $z \in B^o(y, \alpha)$ , akkor  $d(z, y) < \alpha = r - d(x, y)$ . Ebből a  $d(z, x) < d(z, y) + d(y, x)$  háromszög egyenlőtlenség felhasználásával azt kapjuk, hogy

$$d(z, x) < r - d(x, y) + d(y, x) = r$$

a metrika szimmetriája miatt, és így  $z \in B^o(x, r)$ . □

Az előző alpont végén bevezettük a metrikus tér részterének a fogalmát. Most ennek a gömb fogalomra való kihatásával foglalkozunk:

Legyen az  $(X, d)$  egy metrikus tér, és  $C \subseteq X$  az  $X$  egy résztere. A metrika nyilvánvalóan nem változik a  $C$  részhalmazon, de változni fognak a gömbök, hiszen ha  $x \in C$ , akkor az  $(X, d)$  metrikus térben az  $x \in X$  pont körüli  $r$  sugarú nyílt gömb

$$\{y \in X : d(x, y) < r\},$$

a  $(C, d)$  metrikus térben pedig

$$\{y \in C : d(x, y) < r\} = \{y \in X : d(x, y) < r\} \cap C.$$

Nyilvánvaló, hogy a  $C$  halmazban vett gömb határozottan szűkebb is lehet az  $X$  halmazban vett gömbnél.

Ha a  $C$  zárt intervallum, akkor a végpontjai körüli gömbökre külön elnevezést is bevezetünk.

**Definíció 100** Vegyük a valós számok  $\mathbb{R}$  euklideszi terét. Az  $\mathbb{R}$  egy  $[a, b]$ , ( $a < b$ ) zárt intervallumán vett metrikus részterében az  $a$  pont körüli  $r$  ( $< b - a$ ) sugarú nyílt gömböt, az  $a$  pont körüli  $r$  sugarú jobboldali gömbnek fogjuk mondani. Hasonló a baloldali gömb definíciója. Jelölésként azt tesszük, hogy a  $B$  betű alsó indexének a  $B$  illetve  $J$  indexeket írjuk. Hasonló a jobb- és baloldali zárt gömbök definíciója.

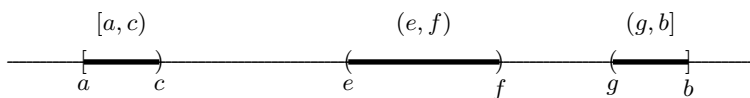
A definíció szerint az  $a$  pont körüli  $r$  sugarú jobboldali nyílt gömb:

$$B_J^\circ(a, r) = \{x : a \leq x < a + r\} = [a, a + r),$$

a  $b$  pont körüli baloldali nyílt gömb pedig:

$$B_B^\circ(b, r) = \{x : b - r < x \leq b\} = (b - r, b].$$

A 3.4. ábrán szemléltetjük a bal- és jobboldali gömböket.



3.4. ábra: Nyílt gömbök az  $[a, b]$  metrikus térben.

Vegyük észre, hogy — amíg az  $\mathbb{R}$  halmazon vett nyílt gömb mindig nyílt intervallum — a bal illetve jobboldali nyílt gömbök balról illetve jobbról zárt intervallumok. Ilyen módon most már mindenfajta korlátos intervallumot tekinthetünk, alkalmas módon nyílt gömbnek.

Gyakorta fogunk vizsgálni olyan függvényeket, amelyek olyan halmazon vannak értelmezve, amelyet úgy kapunk, hogy egy gömbből elvesszük a középpontját (ami baloldali illetve jobboldali gömbnél a jobboldali illetve baloldali végpont). Például egy ilyen eset: Az  $x \mapsto 1/x$  függvény értelmezhető azon a halmazon, amelyet úgy kapunk, hogy a nulla körüli 1 sugarú gömbből elvesszük a 0 középpontját.

**Definíció 101** Legyen az  $(X, d)$  metrikus tér. Egy  $x$  pont körüli  $r$  sugarú hiányos nyílt gömbnek nevezzük a

$$B^\circ(x, r) \setminus \{x\}$$

halmazt.

A fogalmat elsősorban a valós egyenesen fogjuk használni, erre az esetre külön is megfogalmazzuk a definíciókat.

AZ  $\mathbb{R}$ -beli hiányos nyílt gömb-környezet a

$$B^\circ(x, r) \setminus \{x\} = \{y \in X : |x - y| < r \text{ és } x \neq y\} = (x - r, x) \cup (x, x + r)$$

halmaz.

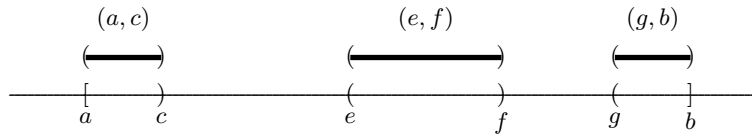
A jobboldali illetve baloldali hiányos nyílt gömb-környezetek definíciója:

$$B_J^\circ(x, r) \setminus \{x\} = \{y : x < y < x + r\} = (x, x + r)$$

illetve

$$B_B^\circ(x, r) \setminus \{x\} = \{y : x - r < y < x\} = (x - r, x).$$

Hasonló a hiányos zárt környezetek definíciója. A nyílt és zárt gömbi környezetek 98. Állításban leírt tulajdonsága miatt a nyílt illetve zárt jelzőket gyakorta elhagyjuk. A 3.5. ábrán illusztráljuk a fogalmat.



3.5. ábra: Hiányos nyílt gömbök az  $[a, b]$  térben.

**Definíció 102** Egy metrikus tér valamely nem üres részhalmazát korláatosnak nevezzük, ha benne van valamilyen gömbben.

Tekintsük például az az  $\mathbb{R}^p$  ( $p > 1$ ) teret az euklideszi metrikával. Ebben a térben az

$$S_p = \left\{ (x_1, \dots, x_p) : \sum_{k=1}^p x_k = 1, x_k \geq 0, k = 1, \dots, p \right\}$$

úgynevezett szimplex korlátos halmaz, hiszen az origó középpontú egységsugarú gömbben fekszik.

Ugyanakkor ebben a térben a

$$K = \left\{ (x_1, \dots, x_p) : \sum_{k=1}^p x_k = 1 \right\}$$

halmaz nem korlátos.

### 3.1.3 Nyílt és zárt halmazok metrikus térben

Egy metrikus tér részhalmazainak a pontjait különféleképpen jellemezhetjük.

Emlékeztetünk rá, hogy minden nyílt gömb-környezet tartalmaz zárt gömb-környezetet, és megfordítva, ezért a következő definíciókban — mindkettő helyett — mondhatunk egyszerűen gömb-környezetet.

**Definíció 103** Legyen az  $A$  egy részhalmaza valamilyen metrikus térnek, és  $x \in A$ . Ha az  $A$  halmaz tartalmazza az  $x$  pontnak valamilyen gömbkörnyezetét is, akkor azt mondjuk, hogy az  $x$  belső pontja az  $A$  halmaznak. Az  $A$  halmaz belső pontjainak az összességét az  $A$  halmaz belsejének (interiorjának) mondjuk, és  $\text{int}(A)$ -val jelöljük.

**Definíció 104** Legyen az  $A$  egy részhalmaza valamilyen metrikus-térnek. A tér egy  $x$  pontját az  $A$  halmaz érintkezési pontjának vagy limesz-pontjának mondjuk, ha tetszőleges gömb-környezetének van közös pontja az  $A$  halmazzal. Az  $A$  halmaz érintkezési pontjainak az összességét az  $A$  halmaz lezárásának mondjuk, és a  $\text{cl}(A)$  jelölést fogjuk használni.

**Példa 3.8** Határozzuk meg az  $\{1, 1/2, \dots, 1/n, \dots\} \subseteq \mathbb{R}$  halmaz összes érintkezési pontját.

Egyrészt egy halmaz elemei nyilván mind érintkezési pontok.

Másképp a  $0$  pont nem eleme a halmaznak, de érintkezési pont, hiszen tetszőleges  $\epsilon > 0$  mellett

$$\frac{1}{n} \in B(0, \epsilon), \quad \text{ha} \quad n \geq \frac{1}{\epsilon}.$$

□

Az érintkezési ponttal rokon a következő fogalom:

**Definíció 105** Az  $(X, d)$  metrikus-tér egy  $p$  pontját egy  $S$  részhalmaza torlódási pontjának mondjuk, ha a  $p$  pont minden gömb-környezete tartalmaz a  $p$ -től különböző pontot az  $S$  halmazból.

**Példa 3.9** Mik a torlódási- és érintkezési pontjai a valós számok következő részhalmazainak?

- (1)  $\{1, 2, \dots\}$ .
- (2)  $\{1, 1/2, 1/3, \dots, 1/n, \dots\}$ .
- (3) A racionális számok  $Q$  halmaza.

Az első halmaznak torlódási pontja nincs: Nem lehet ugyanis torlódási-pont a halmaznak egy pontja sem, mert azoknak egy eléggé kis sugarú (például  $1/2$ ) gömb-környezete a középponton kívül nem tartalmaz elemet a halmazból. A halmazon

kívüli elem pedig azért nem lehet torlódási pont, mert eléggé kis sugarú környezetet nem tartalmaz elemet a halmazból. Az előbb mondottakból világos, hogy a halmaz minden pontja érintkezési pont, de azon kívül nincsenek érintkezési pontok.

A második halmaznak a torlódási pontja könnyen láthatóan a 0 pont, az érintkezési pontjai pedig: a halmaz pontjai és a 0 pont.

A racionális számok torlódási és érintkezési pontjai kiadják a valós számok összességét, a lezártja a valós számok  $\mathbb{R}$  összessége. Ez indokolta azt az elnevezést, hogy a racionális számok sűrűek az  $\mathbb{R}$ -ben.  $\square$

Egy metrikus tér részhalmazai között alapvető szerepe van azoknak, amelyek a következő definícióban leírt tulajdonsággal bírnak.

**Definíció 106** Valamely  $(X, d)$  metrikus-tér egy  $O$  részhalmazát nyíltnak nevezzük, ha minden pontja belső pont, azaz minden  $x \in O$  pontjának van olyan gömb-környezete, amelyik része az  $O$  halmaznak..

A következő tételben a nyílt halmazok alapvető tulajdonságait soroljuk fel.

**Állítás 107** Legyen az  $(X, d)$  tetszőleges metrikus tér. A nyílt halmazok összessége teljesíti a következőket.

- (1) Az  $\emptyset$  és az  $X$  halmaz nyíltak.
- (2) Nyílt halmazok egyesítése is nyílt, azaz ha  $O_\gamma$ ,  $\gamma \in \Gamma$  nyílt halmazok egy tetszőleges családja, akkor az

$$\bigcup_{\gamma \in \Gamma} O_\gamma$$

halmaz is nyílt.

- (3) Véges sok nyílt halmaz metszete is nyílt, azaz ha az  $O_1, \dots, O_n$  halmazok nyíltak, akkor a

$$\bigcap_{i=1}^n O_i$$

halmaz is nyílt.

**Bizonyítás. (1):** Mindkét halmaz esetében teljesen nyilvánvaló az állítás.

**(2):** Ha az  $x$  egy tetszőleges pontja az  $O_\gamma$  halmazok uniójának, akkor eleme valamilyen  $O_{\gamma_0}$  halmaznak is, és mivel ez nyílt, ezért van olyan gömb az  $x$  pont körül, amelyik része az  $O_{\gamma_0}$  halmaznak, következésképpen része az  $O_\gamma$  halmazok uniójának is.

**(3):** Ha az  $x$  pont eleme az  $O_i$  halmazok metszetének, akkor minden  $i$  indexre van olyan  $r_i$  sugarú gömb, amelyik részhalmaza az  $O_i$  halmaznak, azaz

$$B^\circ(x, r_i) \subseteq O_i,$$

minden  $i = 1, 2, \dots, n$  indexre. A véges sok  $r_i$  pozitív számnak az  $r \stackrel{\text{def}}{=} \min_{1 \leq i \leq n} r_i$  minimuma is pozitív. Az  $x$  körüli  $r$  sugarú kör benne van valamennyi  $r_i$  sugarú körben, ezért benne van az  $O_i$  részhalmazok mindegyikében, tehát a metszetükben is, amivel beláttuk, hogy a metszet is nyílt.  $\square$

A következő állítás azt mutatja, hogy indokoltak a “nyílt” intervallum és “nyílt” gömb elnevezések.

**Állítás 108** (1) *Egy  $(X, d)$  metrikus térben minden nyílt gömb nyílt halmaz.*

(2) *A valós számegyenesen, az euklideszi metrika mellett, minden nyílt intervallum nyílt halmaz.*

**Bizonyítás. (1):** A 99. Állításban pontosan ezt mondtuk ki.

**(2):** Korlátos nyílt intervallumra adódik az előzőből, mivel az  $(a, b)$  korlátos nyílt intervallum az a nyílt gömb, amelyiknek a középpontja az intervallum  $(a + b)/2$  középpontja, a sugara pedig az intervallum hosszának a  $(b - a)/2$  fele.

Korlátlan nyílt intervallumra is belátható direktben, vagy azonnal adódik abból, hogy minden korlátlan nyílt intervallum korlátos nyílt intervallumok uniója:

$$(a, +\infty) = \bigcup_{n=1}^{+\infty} (a, a + n),$$

és nyílt halmazok uniója is nyílt.  $\square$

**Definíció 109** *Legyen az  $(X, d)$  egy metrikus-tér. Egy  $O$  nyílt halmaz komplementerét zárt halmaznak nevezzük.*

Mivel egy komplementer komplementere az eredeti halmaz, ezért zárt halmaz komplementere nyílt, tehát *egy halmaz pontosan akkor zárt, ha a komplementere nyílt.*

Fel kell hívni a figyelmet arra, hogy a “nyílt” és a “zárt” nem ellentétes fogalmak, hiszen egy halmaz lehet egyszerre nyílt és zárt. Például az  $X$  halmaz és az  $\emptyset$  üres halmaz mindketten nyíltak és egyben zártak is, mivel egymás komplementerei.

Az alábbi állítás könnyen adódik a definíciókból:

**Állítás 110** *Legyen az  $(X, d)$  egy metrikus-tér. Egy  $F \subseteq X$  halmaz pontosan akkor zárt, ha minden érintkezési pontját tartalmazza.*

**Bizonyítás.** Ha  $F$  zárt, akkor az  $F$  komplementeréből vett pontnak egy környezete is a komplementerhez tartozik, így az nem lehet az  $F$  érintkezési pontja. Fordítva, ha  $F$  minden érintkezési pontját tartalmazza, akkor a komplementer bármely elemének van olyan környezete, amely nem metszi az  $F$  halmazt.  $\square$

A zárt halmazok alaptulajdonságait mondja ki a következő tétel.

**Állítás 111** *Legyen az  $(X, d)$  egy metrikus-tér. A zárt halmazok összessége eleget tesz a következőknek.*

- (1) *Az  $\emptyset$  és  $X$  halmazok zártak.*
- (2) *Akárhány zárt halmaz metszete is zárt.*
- (3) *Véges sok zárt halmaz uniója is zárt.*

A zárt halmazok a nyílt halmazok komplementereiként lettek definiálva, és ennek megfelelően a fenti tulajdonságok — a halmazelmélet műveleteinél mondottak szerint (13. oldal) — duális viszonyban vannak a nyílt halmazok tulajdonságaival.

**Bizonyítás.** Az (1) állítás teljesen nyilvánvaló, a (2) és (3) teljesülése pedig a De Morgan formulák azonnali következménye.

(2): Ha a  $C_\gamma$ ,  $\gamma \in \Gamma \neq \emptyset$  zárt halmazok az  $O_\gamma$  nyílt halmazok komplementerei, akkor

$$\bigcap_{\gamma \in \Gamma} C_\gamma = \bigcap_{\gamma \in \Gamma} O_\gamma^c = \left( \bigcup_{\gamma \in \Gamma} O_\gamma \right)^c,$$

tehát — mivel nyílt halmazok egyesítése is nyílt — zártak rendszerének a metszete is zárt.

(3): Az előzővel teljesen hasonló módon: ha a  $C_i$ ,  $i = 1, \dots, n$  zárt halmazok az  $O_i$ ,  $i = 1, \dots, n$  nyílt halmazok komplementere, akkor

$$\bigcup_{i=1}^n C_i = \bigcup_{i=1}^n O_i^c = \left( \bigcap_{i=1}^n O_i \right)^c,$$

amivel be is láttuk az állítást.  $\square$

A következő állítás szerint jogosultak a “zárt” gömb és intervallum elnevezések.

**Állítás 112** (1) *Egy  $(X, d)$  metrikus-térben minden zárt gömb zárt halmaz.*

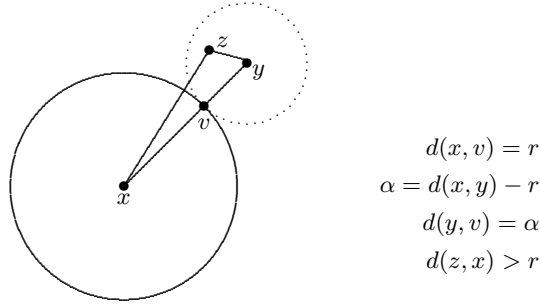
(2) *Az  $\mathbb{R}$  egyenes tetszőleges zárt intervalluma zárt halmaz.*

**Bizonyítás. (1):** A zárt halmaz definíciója szerint azt kell belátnunk, hogy egy  $B(x, r)$  zárt gömb komplementere nyílt halmaz.

A metrikus tér vizsgálatánál, mint minden más matematikai területen, hasznos valamilyen “fantázia rajzon” szemléltetni az állításokat, és követni a bizonyításokat, ahogyan azt most a 3.6. ábrán tesszük.

Legyen  $y \in (B(x, r))^c$ . Megmutatjuk, hogy az  $y$  körül van olyan gömb, ami szintén a  $B(x, r)$  zárt gömb komplementerében van.

Mivel  $y \notin B(x, r)$ , ezért  $d(x, y) > r$ . Vegyük az  $\alpha \stackrel{\text{def}}{=} d(x, y) - r$  pozitív számot. Az  $y$  körüli  $\alpha$  sugarú  $B^\circ(y, \alpha)$  gömb minden pontja a  $B(x, r)$  gömbön kívül van.



3.6. ábra: Zárt gömb komplementere nyílt.

Legyen ugyanis  $z \in B^\circ(y, \alpha)$ , akkor  $d(y, z) < \alpha$ . Mivel a háromszög egyenlőtlenség szerint  $d(x, y) \leq d(x, z) + d(y, z)$ , ezért

$$\begin{aligned} d(x, z) &\geq d(x, y) - d(y, z) > d(x, y) - \alpha = \\ &= d(x, y) - (d(x, y) - r) = r, \end{aligned}$$

tehát  $d(x, z) > r$ , azaz  $z \notin B(x, r)$ , amit bizonyítani kellett.

(2): Egy  $[a, b]$  korlátos, zárt intervallum komplementere két nyílt intervallum egyesítése, amely az eddigiek szerint nyílt halmaz.

A bizonyítás végtelen intervallumokra hasonlóan végezhető el.  $\square$

## 3.2 Folytonos függvények metrikus téren

Ebben a pontban metrikus-terek közötti leképezések egy alapvető összességével az u. n. folytonos függvényekkel fogunk foglalkozni.

### 3.2.1 Definíciók

Metrikus-terek közötti leképezésekkel kapcsolatban természetes célkitűzés olyan függvények vizsgálata, amelyek "közeli" pontokat "közeli" pontokba képeznek, ezek a folytonos függvények.

Lássuk ezek után a folytonos függvény definícióját:

**Definíció 113** Legyenek az  $(X, d_X)$  és  $(Y, d_Y)$  metrikus terek. Az  $f : X \rightarrow Y$  függvényt folytonosnak nevezzük egy  $b \in X$  pontban, ha az  $f(b)$  minden  $V_{f(b)}$  gömb-környezetéhez van olyan  $U_b$  gömb-környezete a  $b$ -nek, hogy

$$f(U_b) \subseteq V_{f(b)}. \quad (3.1)$$

A következőkben elsősorban  $\mathbb{R} \rightarrow \mathbb{R}$  és  $\mathbb{R}^p \rightarrow \mathbb{R}$  függvényekkel fogunk foglalkozni, ezért ismételjük el ezen esetekben a  $b$  pontban való folytonosság feltételét: Minden  $\epsilon$  pozitív számhoz van olyan  $\delta$  pozitív szám, hogy

$$|f(x) - f(b)| < \epsilon, \quad \text{hacsak} \quad |x - b| < \delta.$$

vagy másképpen:

$$|x - b| < \delta \implies |f(x) - f(b)| < \epsilon.$$

Az  $\mathbb{R}^p$  téren értelmezett függvényekre csak az az eltérés, hogy

$$|x - b| \text{ helyett } \|x - b\|$$

szerepel.

Egy függvény folytonossága egy adott pontban úgynevezett *lokális* tulajdonság, abban az értelemben, hogy a folytonosság csak az adott pont egy környezetében való viselkedéstől függ.

Ha egy konkrét függvény folytonosságát akarjuk belátni, akkor a 113. definíció a következőképpen alkalmazható:

- Feltételezzük, hogy adott egy  $\epsilon$  tetszőleges, de rögzített pozitív szám.
- Alkalmas módon meghatározunk az  $\epsilon$  számhoz egy olyan  $\delta$  pozitív számot, amelyikkel a definíció feltétele teljesül. Ilyen  $\delta$  szám sok lehet, hiszen ha egy már van annál minden kisebb is alkalmas.

Lássuk most ennek a módszernek — amit  $\epsilon$ - $\delta$ -*technikának* mondanak — az alkalmazásait.

**Példa 3.10** *Mutassuk meg, hogy az  $x \mapsto (3x + 1)$ ,  $\mathbb{R} \rightarrow \mathbb{R}$  leképezés minden  $x$  pontban folytonos.*

Legyen adva egy  $\epsilon$  pozitív szám. Az  $x$  pontban való folytonossághoz olyan  $\delta$  pozitív számot kell találnunk, hogy  $|(3u + 1) - (3x + 1)| < \epsilon$  legyen, ha  $|u - x| < \delta$ .

Az  $|(3u + 1) - (3x + 1)| = 3|u - x|$  egyenlőségből azonnal kapjuk, hogy

$$|u - x| < \epsilon/3 \implies |(3u + 1) - (3x + 1)| < \epsilon,$$

tehát a  $\delta = \epsilon/3$  szám alkalmas a célunkhoz (és persze minden ennél kisebb is).  $\square$

**Példa 3.11** *Lássuk be, hogy az  $x \mapsto x^2$ ,  $\mathbb{R} \rightarrow \mathbb{R}$  leképezés folytonos minden pontban. Milyen közel kell menni a  $b = 1$  illetve  $b = 10$  pontokhoz, ha azt akarjuk, hogy a függvényértékek  $1/100$  pontossággal közelítsék a függvény  $b$ -ben felvett értékét?*

Egy tetszőleges  $b$  helyen látjuk be a folytonosságot. Kiinduló környezetnek, ahol a függvényt vizsgáljuk, válasszuk a  $B^\circ(b, 1)$  gömböt, azaz az

$$b - 1 < x < b + 1 \tag{3.2}$$

halmazt.

Mivel

$$|x^2 - b^2| = |x - b| \cdot |x + b| \leq |x - b|(|x| + |b|),$$

ezért a (3.2) választás miatt

$$|x^2 - b^2| \leq |x - b|(|b| + 1 + |b|) = (2|b| + 1)|x - b|.$$

Ebből viszont könnyen láthatjuk, hogy

$$|x^2 - b^2| < \epsilon \quad \text{ha} \quad |x - b| < \frac{\epsilon}{2|b| + 1},$$

ami azt jelenti, hogy a  $\delta = \epsilon/(2|b| + 1)$  megfelel a kívánalmaknak.

Nézzük először a  $b = 1$  helyet. Ha  $1/100$  pontossággal akarjuk közelíteni az  $1^2 = 1$  értéket az  $x^2$  értékkel, akkor az  $x$  számnak elégséges a

$$\delta = \frac{1/100}{2 + 1} = \frac{1}{300}$$

pontossággal megközelíteni az 1 helyet.

Nézzük most a  $b = 10$  helyet. Ha  $1/100$  pontossággal akarjuk közelíteni az  $10^2 = 100$  értéket az  $x^2$  értékkel, akkor az  $x$  számnak elegendő a

$$\delta = \frac{1/100}{20 + 1} = \frac{1}{2100}$$

pontossággal megközelíteni a 10 helyet. □

Az utóbbi példából megállapíthatjuk, hogy adott  $\epsilon$  közelítési pontosság mellett a  $\delta$  szám függhet attól a helytől is, ahol a függvény közelítését vizsgáljuk. A megelőző példában a  $\delta$  a helytől nem függött, csak az  $\epsilon$ -től.

Most pedig a pontban való folytonosság lokális tulajdonságát globális tulajdonsággá terjesztjük ki:

**Definíció 114** Legyenek az  $(X, d_X)$  és  $(Y, d_Y)$  metrikus-terek. Egy  $f : X \rightarrow Y$  leképezést folytonosnak mondunk (az  $X$  halmazon), ha folytonos minden  $b \in X$  pontban.

**Állítás 115** Legyenek az  $(X, d_X)$  és  $(Y, d_Y)$  metrikus-terek. Egy  $f : X \rightarrow Y$  leképezés pontosan akkor folytonos, ha minden nyílt halmaz inverzképe nyílt.

**Bizonyítás.** Legyen  $f$  folytonos,  $G$  az  $Y$  nyílt részhalmaza és  $x \in f^{-1}(G)$ . Ekkor  $f(x) \in G$ , ezért van olyan  $V_{f(x)}$  környezete az  $f(x)$  pontnak, amelyre  $V_{f(x)} \subseteq G$ . Mivel  $f$  folytonos, található olyan  $U_x$  környezete az  $x$  pontnak, hogy  $f(U_x) \subseteq V_{f(x)}$ . Tehát

$$U_x \subseteq f^{-1}(V_{f(x)}) \subseteq f^{-1}(G).$$

Ez éppen azt jelenti, hogy  $f^{-1}(G)$  nyílt halmaz.

Fordítva, ha  $V$  az  $f(x)$  pont egy nyílt környezete, akkor  $f^{-1}(V)$  nyílt, így tartalmazza az  $x$  pont egy  $U$  környezetét, azaz  $f(U) \subseteq f(f^{-1}(V)) \subseteq V$ . Ez azt jelenti, hogy  $f$  folytonos az  $x$  pontban. □

**Állítás 116** *Legyenek az  $(X, d_X)$  és  $(Y, d_Y)$  metrikus-terek, az  $f : X \rightarrow Y$  folytonos leképezés és  $A \subseteq X$ . Ekkor az  $f|_A, A \rightarrow Y$  leszűkített leképezés is folytonos.*

Más szavakkal: Egy téren folytonos leképezés annak egy részhalmazán is folytonos. Például: ha az  $f : \mathbb{R} \rightarrow \mathbb{R}$  folytonos, akkor az  $f$  tetszőleges intervallumon (intervallumra leszűkítve) is folytonos.

**Bizonyítás.** Legyen  $b \in A$ . Az  $f : X \rightarrow Y$  folytonossága miatt adott  $\epsilon > 0$  számhoz van olyan  $\delta > 0$ , hogy

$$d_Y(f(b), f(x)) < \epsilon, \quad \text{ha} \quad d_X(b, x) < \delta.$$

Ebből nyilvánvalóan igaz az, hogy

$$d_Y(f(b), f(x)) < \epsilon, \quad \text{ha} \quad d_X(b, x) < \delta \quad \text{és} \quad x \in A,$$

ami az  $f|_A : A \rightarrow Y$  folytonosságát jelenti.  $\square$

A következő két állítás az összetett függvény folytonosságával foglalkozik.

**Állítás 117** *Legyenek az  $(X, d_X)$ ,  $(Y, d_Y)$  és  $(Z, d_Z)$  metrikus terek. Tegyük fel, hogy a  $g : X \rightarrow Y$  leképezés folytonos a  $b \in X$  pontban, az  $f : Y \rightarrow Z$  pedig a  $g(b) \in Y$  pontban.*

*Ekkor az  $f \circ g : X \rightarrow Z$  összetett leképezés (kompozíció) is folytonos a  $b$  pontban. Az első lokális jellegű.*

**Bizonyítás.** Az  $f$  függvény  $g(b)$  helyen lévő folytonossága miatt adott  $\epsilon > 0$  számhoz van olyan  $\delta > 0$  szám, hogy

$$d_Z(f(g(b)), f(g(x))) < \epsilon, \quad \text{ha} \quad d_Y(g(b), g(x)) < \delta.$$

A  $g$  függvénynek a  $b$  helyen való folytonossága miatt a  $\delta > 0$  számhoz van olyan  $\delta_1 > 0$ , hogy

$$d_Y(g(b), g(x)) < \delta, \quad \text{ha} \quad d_X(b, x) < \delta_1.$$

A két kiemelt sor alapján

$$d_Z(f(g(b)), f(g(x))) < \epsilon, \quad \text{ha} \quad d_X(b, x) < \delta_1,$$

ami éppen az  $f \circ g$  kompozíció  $b$  pontban való folytonosságát jelenti.  $\square$

Az előző tételből a folytonosság globális definíciója alapján azonnal adódik a globális tétel:

**Állítás 118** *Legyenek az  $(X, d_X)$ ,  $(Y, d_Y)$  és  $(Z, d_Z)$  metrikus terek, a  $g : X \rightarrow Y$  és  $f : Y \rightarrow Z$  leképezések folytonosak.*

*Ekkor a  $f \circ g : X \rightarrow Z$  összetett leképezés (kompozíció) is folytonos.*

### 3.2.2 Az alpműveletek folytonossága

A valós számok közötti műveletek is leképezések. Például az

$$(x_1, x_2) \mapsto x_1 \cdot x_2$$

a szorzás az  $\mathbb{R}^2$ -ből az  $\mathbb{R}$ -be képező függvény. Ilyen esetben, amikor az értelmezési tartomány számpárokból (kettesekből) áll, akkor azt szokás mondani, hogy kétváltozós a függvény. Eszerint a szorzás egy kétváltozós függvény. Hasonló mondható az összeadásra és az osztásra is.

Ebben a pontban be fogjuk látni, hogy az alpműveletek folytonos függvények. Ezeknek az állításoknak nagy jelentősége van, mert például a szorzat esetében azt állítják, hogy az  $x$  és  $y$  számok  $xy$  szorzatához előírt közelségbe kerül az  $u$  és  $v$  számok  $uv$  szorzata, ha az  $(u, v)$  számpár megfelelően közel van az  $(x, y)$  számpárhoz. Ennek az állításnak az alapján végezhetjük el "közelítő" módon a szorzásokat. Ne feledjük, hogy a valós számok zömével csak úgy tudunk számolni, hogy racionálissal közelítjük őket.

**Állítás 119** Vegyük az  $\mathbb{R}^2$  és  $\mathbb{R}$  tereket az euklideszi metrikákkal. Az

$$(x_1, x_2) \mapsto x_1 \cdot x_2$$

$\mathbb{R}^2 \rightarrow \mathbb{R}$  (szorzás) leképezés folytonos.

**Bizonyítás.** Megmutatjuk, hogy a leképezés tetszőleges  $b = (b_1, b_2)$  pontban folytonos. Legyen adott egy  $\epsilon$  pozitív szám.

Mivel a folytonosság lokális tulajdonság, ezért első lépésként megválaszthatjuk a  $b$  pontnak azt a környezetét, amelyben vizsgáljuk a leképezést. A  $b$  pont  $B^\circ(b, 1)$  környezetében fogunk dolgozni, azaz olyan  $x = (x_1, x_2)$  párosok jöhetnek szóba, amelyekre

$$(x_1 - b_1)^2 + (x_2 - b_2)^2 < 1. \quad (3.3)$$

Vegyük a szorzat-függvénynek az eltérését, és rendezzük alkalmas módon:

$$\begin{aligned} |b_1 b_2 - x_1 x_2| &= |b_1 b_2 - b_2 x_1 + b_2 x_1 - x_1 x_2| = \\ &= |b_2(b_1 - x_1) + x_1(b_2 - x_2)|. \end{aligned}$$

A Cauchy-Schwarz egyenlőtlenség alkalmazásával ebből azt kapjuk, hogy

$$|b_1 b_2 - x_1 x_2| \leq \sqrt{b_2^2 + x_1^2} \sqrt{(b_1 - x_1)^2 + (b_2 - x_2)^2},$$

ezért  $|b_1 b_2 - x_1 x_2| < \epsilon$ , ha

$$\sqrt{(b_1 - x_1)^2 + (b_2 - x_2)^2} \leq \frac{\epsilon}{\sqrt{b_2^2 + x_1^2}}. \quad (3.4)$$

Már csak az  $x_1$  változót kell kiküszöbölnünk a baloldaltól, és akkor tudunk mondani egy alkalmas, csak a  $b$ -től függő  $\delta$  számot.

A (3.3) megállapodás miatt  $(b_1 - x_1)^2 < 1$ , és így  $|x_1| < 1 + |b_1|$ , ezért a (3.4) egyenlőtlenségből adódik, hogy a

$$\|b - x\| = \sqrt{(b_1 - x_1)^2 + (b_2 - x_2)^2} \leq \frac{\epsilon}{\sqrt{b_2^2 + (1 + b_1)^2}} = \delta.$$

Összefoglalólag megállapíthatjuk, hogy

$$|b_1 b_2 - x_1 x_2| < \epsilon \quad \text{ha} \quad \|x - b\| < \delta = \frac{\epsilon}{\sqrt{b_2^2 + (1 + b_1)^2}}.$$

□

**Állítás 120** Vegyük az  $\mathbb{R}^2$  és  $\mathbb{R}$  tereket az euklideszi metrikával. Az

$$x = (x_1, x_2) \mapsto x_1 + x_2$$

$\mathbb{R}^2 \rightarrow \mathbb{R}$  (összeadás) leképezés folytonos.

**Bizonyítás.** Legyen a  $b = (b_1, b_2)$  egy tetszőleges pont. Becsüljük meg a  $(b_1 + b_2)$  és  $(x_1 + x_2)$  összegek eltérését, a Cauchy-egyenlőtlenség segítségével:

$$\begin{aligned} |(b_1 + b_2) - (x_1 + x_2)| &= |1 \cdot (b_1 - x_1) + 1 \cdot (b_2 - x_2)| \leq \\ &\leq \sqrt{1^2 + 1^2} \cdot \sqrt{(b_1 - x_1)^2 + (b_2 - x_2)^2} = \sqrt{2} \cdot \|b - x\|. \end{aligned}$$

Eszerint adott  $\epsilon > 0$  szám mellett

$$|(b_1 + b_2) - (x_1 + x_2)| \leq \epsilon, \quad \text{ha} \quad \|b - x\| \leq \frac{\epsilon}{\sqrt{2}} = \delta,$$

tehát az összeg-leképezés folytonos.

□

**Állítás 121** Vegyük az  $\mathbb{R}^2$  és  $\mathbb{R}$  tereket az euklideszi metrikával. Az

$$(x_1, x_2) \mapsto \frac{x_1}{x_2}$$

(hányados) leképezés, aminek az értelmezési tartománya az  $(x_1, x_2) \in \mathbb{R}^2$ ,  $x_2 \neq 0$  számpárokból áll, folytonos.

**Bizonyítás.** Legyen a  $b = (b_1, b_2)$  egy tetszőleges olyan pont, amelyre  $b_2 \neq 0$ .

Először becsüljük meg a függvény eltérését a Cauchy-egyenlőtlenség alapján:

$$\begin{aligned} \left| \frac{b_1}{b_2} - \frac{x_1}{x_2} \right| &= \left| \frac{b_1 x_2 - b_2 x_1}{b_2 x_2} \right| = \frac{|b_1 x_2 - b_1 b_2 + b_1 b_2 - b_2 x_1|}{|b_2 x_2|} = \\ &= \frac{|b_1(x_2 - b_2) + b_2(b_1 - x_1)|}{|b_2 x_2|} \leq \frac{\sqrt{b_1^2 + b_2^2} \cdot \sqrt{(x_2 - b_2)^2 + (b_1 - x_1)^2}}{|b_2 x_2|} = \end{aligned}$$

$$= \frac{||b|| \cdot ||b-x||}{|b_2 x_2|}.$$

Összefoglalva:

$$\left| \frac{b_1}{b_2} - \frac{x_1}{x_2} \right| \leq \frac{||b|| \cdot ||b-x||}{|b_2 x_2|}. \quad (3.5)$$

A jobboldal további becsléséhez egy alsó becslést kell adnunk az  $|x_2|$ -re, mivel a nevezőben van. Ezt megtehetjük, ha a  $b$  pont alkalmas környezetére szorítkozunk, amit a lokáltság miatt megtehetünk. Legyen

$$||b-x|| \leq \frac{|b_2|}{2}. \quad (3.6)$$

Ebből nyilvánvalóan  $|b_2 - x_2| \leq |b_2|/2$ , és így az abszolút értékre vonatkozó második háromszög-egyenlőtlenség alapján:

$$|x_2| = |b_2 - (b_2 - x_2)| \geq |b_2| - |b_2 - x_2| \geq |b_2| - \frac{|b_2|}{2} = \frac{|b_2|}{2}.$$

Ezt felhasználva a (3.5) tovább alakítható:

$$\left| \frac{b_1}{b_2} - \frac{x_1}{x_2} \right| \leq \frac{||b|| \cdot ||b-x||}{|b_2| \cdot \frac{|b_2|}{2}} = \frac{2||b||}{|b_2|^2} ||b-x||, \quad \text{ha} \quad ||b-x|| \leq \frac{|b_2|}{2}.$$

Ebből pedig — figyelembe véve a (3.6)-t is — megállapíthatjuk, hogy

$$\left| \frac{b_1}{b_2} - \frac{x_1}{x_2} \right| < \epsilon, \quad \text{ha} \quad ||b-x|| < \min \left\{ \frac{|b_2|^2}{2||b||} \epsilon, \frac{|b_2|}{2} \right\} = \delta,$$

amit bizonyítani kellett.  $\square$

A következő állításban egy  $\mathbb{R} \rightarrow \mathbb{R}^2$ -be függvény folytonosságát igazoljuk.

**Állítás 122** *Ha az  $f$  és  $g$  valós függvények folytonosak a  $b \in \mathbb{R}$  pontban, akkor az*

$$x \mapsto (f(x), g(x))$$

$\mathbb{R} \rightarrow \mathbb{R}^2$  leképezés is folytonos a  $b$  pontban.

**Bizonyítás.** Nézzük a függvény értékek távolságát a  $b$  szám egy környezetében:

$$||(f(b), g(b)) - (f(x), g(x))|| =$$

$$= ||(f(b) - f(x), g(b) - g(x))|| = \sqrt{(f(b) - f(x))^2 + (g(b) - g(x))^2}.$$

Az  $f$  és  $g$  folytonossága miatt, adott  $\epsilon > 0$  számhoz van olyan  $\delta_1 > 0$  és  $\delta_2 > 0$ , hogy

$$|f(b) - f(x)| < \frac{\epsilon}{\sqrt{2}}, \quad \text{ha} \quad |b-x| < \delta_1$$

és

$$|g(b) - g(x)| < \frac{\epsilon}{\sqrt{2}}, \quad \text{ha} \quad |b - x| < \delta_2.$$

Ezek alapján:

$$\begin{aligned} \|(f(b), g(b)) - (f(x), g(x))\| &\leq \\ &\leq \sqrt{(f(b) - f(x))^2 + (g(b) - g(x))^2} < \sqrt{\epsilon^2} = \epsilon, \end{aligned}$$

feltéve, hogy  $|b - x| < \min\{\delta_1, \delta_2\} = \delta$ .  $\square$

### 3.2.3 Formális szabályok, elemi függvények

Az előző alpontban megoldottunk néhány olyan feladatot, amelyben a függvény folytonosságát úgy láttuk be, hogy meghatároztunk az adott  $\epsilon$  számhoz egy alkalmas  $\delta$  számot. Ez némelykor nehéz feladat is lehet. Ebben a pontban általános módszert adunk valós függvények folytonosságának az igazolásához, amivel eléggé széles függvényosztály folytonosságát be tudjuk látni.

A leírandó eljárás két alapra támaszkodik:

- Megmutatjuk, hogy a folytonos függvények közötti műveletek (leképezés-építési eljárások) folytonos függvényekhez vezetnek.
- Néhány egyszerű függvényre belátjuk, hogy folytonosak. Ezek az  $x \mapsto 1$ ,  $x \mapsto x$ , exponenciális, logaritmus és a trigonometrikus függvények lesznek.

Ezek alapján az elemi függvényekből a függvényépítő eljárások segítségével nyerhető függvények eléggé széles összességére be tudjuk látni, hogy folytonosak.

**Állítás 123 (Formális szabályok)** *Legyenek az  $f$  és  $g$  valós értékű függvények egy metrikus téren, és az  $\alpha$  tetszőleges valós szám.*

(1) *Ha az  $f$  és  $g$  függvények folytonosak a  $b$  helyen, akkor az*

(a)  $f + g$ ,

(b)  $f \cdot g$ ,

(c)  $\alpha f$ ,

(d)  $\frac{f}{g}$ , *feltéve, hogy  $g(b) \neq 0$*

*leképezések is folytonosak az a pontban.*

(2) *Ha a  $g$  függvény folytonos a  $b$  pontban, az  $f$  pedig folytonos a  $g(b)$  pontban, akkor a  $f \circ g$  kompozíció (összetett) függvény is folytonos a  $b$  helyen.*

A bizonyításokból látszani fog, hogy a tétel értelemszerűen igaz marad akkor is, ha  $\mathbb{R}^p \rightarrow \mathbb{R}$  függvényekről van szó.

**Bizonyítás.** A bizonyítások az előző alpont eredményeire és a közvetett függvény folytonosságára támaszkodnak.

**(1): (a):** Az  $x \mapsto f(x) + g(x)$  függvényt, aminek a folytonosságát be kell látnunk, jelöljük  $h(x)$ -szel. A  $h$  a következőképpen kapható meg leképezések egymás utáni végrehajtásával:

$$x \xrightarrow{\phi} (f(x), g(x)) \xrightarrow{s} f(x) + g(x),$$

azaz  $h(x) = s(\phi(x))$ , tehát  $h = s \circ \phi$ . A  $h$  folytonossága azonnal adódik abból, hogy a  $\phi$  illetve  $s$  függvények folytonosak (122. illetve 120 állítások), és az összetett függvény képzés is örzi a folytonosságot (117. állítás).

**(b):** Azonos az előző bizonyítással, csak az összeg leképezés helyett a szorzat leképezést kell szerepeltetni, és a szorzás folytonosságát kell használni (lásd a 119. Állítást).

**(c):** A (b) közvetlen következménye, mivel a konstans leképezés folytonos.

**(d):** Azonos az (a) bizonyításával, csak az összeg leképezés helyett a hányados leképezést kell szerepeltetni, és a hányados folytonosságát kell használni (lásd a 121. Állítást).

**(2):** A 117. állításban általánosabban beláttuk. □

Az  $x \mapsto 1$  és  $x \mapsto x$  valós függvények folytonosak, ezért folytonosak azok a függvények is, amelyek belőlük a formális szabályokkal felépíthetők, azaz a polinomok és racionális törtfüggvények. Ezt fogalmazzuk meg a következő tételben.

**Állítás 124** (1) *Tetszőleges*

$$x \mapsto p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

*polinom folytonos.*

(2) *Tetszőleges*

$$x \mapsto \frac{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0}{b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0}$$

*racionális törtfüggvény folytonos.*

**Bizonyítás.** Először lássuk be azt, hogy a pozitív egész kitevős  $x \mapsto x^n$  hatványfüggvény folytonos. Ez teljes indukcióval történik. Az  $n = 1$  esetben igaz az állítás, hiszen csak annak kell teljesülni, hogy

$$|x - b| < \epsilon \quad \text{ha} \quad |x - b| < \delta (= \epsilon).$$

Ha feltesszük, hogy  $n$ -re igaz az állítás, akkor ebből azonnal adódik  $(n + 1)$ -re, mivel csak a folytonos  $x \mapsto x$  és  $x \mapsto x^n$  függvényeket kell összeszorozni.

A polinom folytonossága abból következik, hogy a hatvány (és állandó) függvényeket kell számmal szorozni és összeadni, és ezek a formális szabályok szerint folytonos függvényt eredményeznek.

A racionális törtfüggvény folytonossága a polinom folytonosságából és a hányadosra vonatkozó formális szabályból adódik.  $\square$

**Állítás 125** *A szinusz, koszinusz és tangens függvények értelmezési tartományukon folytonosak.*

**Bizonyítás.** Először a szinusz függvény folytonosságát látjuk be. Ehhez szükségünk lesz a 88. tételnek arra az állítására, hogy a  $-\frac{\pi}{2} < x < \frac{\pi}{2}$  számközben

$$|\sin x| \leq |x|. \quad (3.7)$$

Legyen a  $b$  tetszőleges valós szám. Közismert trigonometriai azonosság alapján (89. állítás):

$$|\sin x - \sin b| = 2 \left| \cos\left(\frac{x+b}{2}\right) \right| \cdot \left| \sin\left(\frac{x-b}{2}\right) \right| \leq$$

$$2 \left| \sin\left(\frac{x-b}{2}\right) \right| \leq |x-b|.$$

Ebből azonnal kapjuk, hogy adott  $\epsilon$  pozitív számhoz a  $\delta \doteq \epsilon$  pozitív számmal teljesül a következő:

$$|\sin x - \sin b| < \epsilon \quad \text{ha} \quad |x - b| < \delta.$$

A koszinusz függvény folytonossága jön a szinusz függvény folytonosságából a műveletek folytonossága és a kompozíció képzés szabálya alapján a  $\cos x = \sin(\frac{\pi}{2} - x)$  formula miatt.

A tangens függvény folytonossága a szinusz és koszinusz függvények folytonosságából következik a hányadosra vonatkozó szabály alapján.  $\square$

Az exponenciális és logaritmus függvények folytonosságát a határértékekkel foglalkozó szakaszban fogjuk igazolni.

**Példa 3.12** *Mutassuk meg, hogy a következő függvények folytonosak.*

$$(i) \quad x \mapsto \sin(x^2 - 7)$$

$$(ii) \quad x \mapsto \sin(\cos(4x + x^3)).$$

A szereplő függvények olyan függvényekből épülnek fel a formális szabályok segítségével, amelyek folytonosságát már beláttuk.  $\square$

### 3.2.4 Folytonos függvények alaptulajdonságai

Ebben az alponban a folytonos függvények alaptulajdonságaival foglalkozunk.

**Állítás 126** *Ha egy  $f : [a, b] \rightarrow \mathbb{R}$  függvény folytonos, akkor az*

$$\{x \in [a, b] : f(x) < \alpha\} \quad \text{és} \quad \{x \in [a, b] : f(x) > \alpha\}$$

*úgynevezett nívóhalmazok minden  $\alpha$  mellett nyíltak az  $[a, b]$ -ben.*

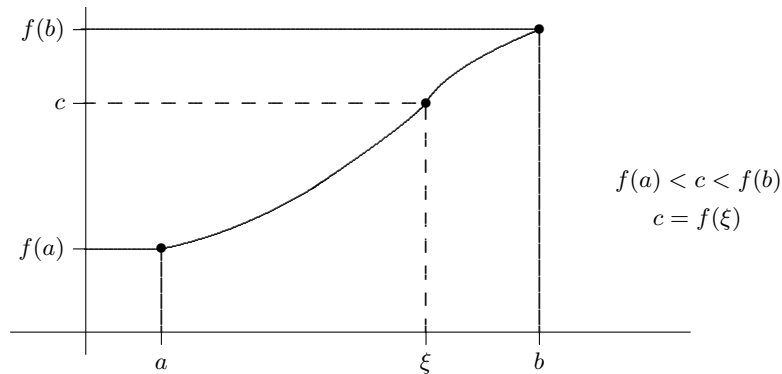
**Bizonyítás.** Mivel a  $H = (-\infty, \alpha)$  intervallum nyílt, azért

$$\{x \in [a, b] : f(x) < \alpha\} = f^{-1}(H)$$

nyílt halmaz az  $[a, b]$  metrikus térben a 115 Állítás alapján.  $\square$

**Állítás 127 (Bolzano-tétel)** *Ha egy  $f : [a, b] \rightarrow \mathbb{R}$  függvény folytonos, akkor minden értéket felvesz az  $f(a)$  és  $f(b)$  között.*

Az állítást a 3.7. ábrán szemléltettük.



3.7. ábra: Bolzano-tétel

**Bizonyítás.** Ha  $f(a) = f(b)$ , akkor nincs mit bizonyítani, és legyen— mondjuk—  $f(a) < c < f(b)$ . Megmutatjuk, hogy van olyan  $\xi \in (a, b)$ , hogy  $f(\xi) = c$ .

Legyen  $A \doteq \{x \in [a, b] : f(x) < c\}$ . Az  $A$  halmaz nem üres, ugyanis  $a \in A$ . Az  $A$  korlátos halmaz, mivel része az  $[a, b]$  intervallumnak, ezért létezik a

$$\xi \doteq \sup A \in [a, b]$$

felső határ, amiről belátjuk, hogy  $f(\xi) = c$ .

Ha  $f(\xi) < c$  lenne, akkor nyilván  $\xi \neq b$ , és így a  $\xi$  pontban folytonos az  $f$  függvény. Emiatt a  $\xi$  egy

$$[\xi, \xi + \delta), \quad \delta > 0$$

jobboldali környezetének az  $x$  pontjaira:  $f(x) < c$  (a megelőző állítás), és így lenne a  $\xi$ -nél nagyobb eleme is az  $A$  halmaznak, ellentétben azzal, hogy a  $\xi$  felső határ.

Ha pedig  $f(\xi) > c$  lenne, akkor  $\xi \neq a$ , és így a  $\xi$  pontban folytonos. Emiatt a megelőző állítás szerint van olyan

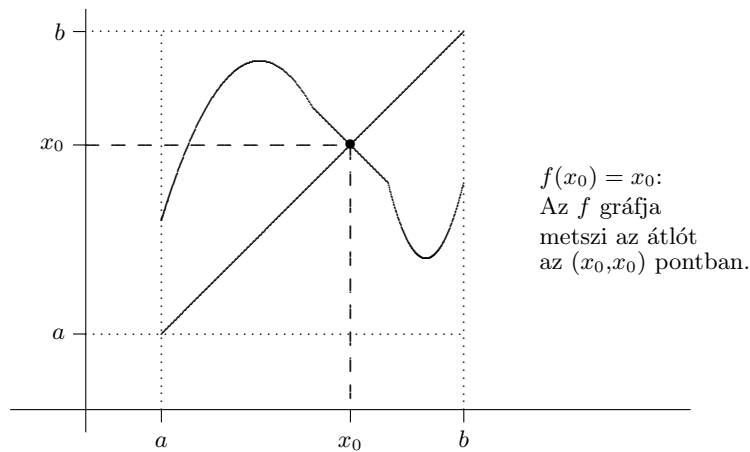
$$(\xi - \delta, \xi], \quad \delta > 0$$

baloldali környezete a  $\xi$  pontnak, amelynek az  $x$  pontjaiban  $f(x) > c$ , vagyis lenne a  $\xi$ -nél kisebb felső korlátja az  $A$  halmaznak, ellentétben azzal, hogy a  $\xi$  felső határ.

Az előzők szerint sem az  $f(\xi) < c$  sem az  $f(\xi) > c$  nem állhat fenn, tehát  $f(\xi) = c$ .  $\square$

Most pedig a Bolzano tételnek egy fontos alkalmazását fogjuk bemutatni. Ehhez azonban szükségünk lesz egy fogalomra.

**Definíció 128** Legyen az  $f : X \rightarrow X$  egy leképezés. Ha van olyan  $x_0 \in X$  pont, amelyre  $f(x_0) = x_0$ , akkor az  $x_0$  pontot az  $f$  leképezés fixpontjának mondjuk.



3.8. ábra: Az  $f : [a, b] \rightarrow [a, b]$  fixpontja.

Az elnevezés eléggé szemléletes, hiszen ha  $f(x_0) = x_0$ , akkor az  $x_0$  pontot helyben (fixen) hagyja az  $f$  leképezés. Általában fontos kérdés az, hogy egy függvénynek van-e fixpontja. A közgazdaságtanban elterjedt módszer az, hogy a gazdaság működését egyensúlyi folyamatként írják le. Az egyensúly létezése rendszerint megfelelő fixpont létezésével ekvivalens.

**Állítás 129** Ha az  $f$  egy  $[a, b] \rightarrow [a, b]$  folytonos leképezés, akkor van fixpontja, azaz létezik olyan  $x \in [a, b]$ , amelyre  $f(x) = x$ .

Az állítást és a bizonyítását a 3.8. ábrán szemléltetjük. A most kimondott tétel egy általánosabb tételnek az u. n. Brouwer-féle fixpont tételnek egy nagyon speciális esete.

**Bizonyítás.** Tekintsük a  $g(x) = f(x) - x$  függvényt. A  $g$  függvény folytonos, mert folytonosok különbsége. Mivel az  $f$  az  $[a, b]$  intervallumot önmagára képezi, ezért

$$g(a) = f(a) - a \geq 0 \quad \text{és} \quad g(b) = f(b) - b \leq 0,$$

és így, a Bolzano tétel miatt, van olyan  $x$  hely, ahol a  $g$  függvény nulla, azaz  $g(x) = f(x) - x = 0$ , tehát az  $x$  fixpontja az  $f$  függvénynek.  $\square$

**Állítás 130** *Ha egy  $f : [a, b] \rightarrow \mathbb{R}$  függvény folytonos, akkor korlátos.*

**Bizonyítás.** Tekintsük a

$$H = \{x \in [a, b] : f \text{ korlátos az } [a, x] \text{ intervallumon}\}$$

halmazt. Világos, hogy  $H$  nem üres, hiszen  $a \in H$ . Legyen  $\alpha = \sup H$ . Ha  $\alpha < b$  lenne, akkor a folytonosság miatt  $f$  korlátos az  $\alpha$  egy környezetében is, így  $\alpha$  nem lehetne a  $H$  felső korlátja. Tehát  $\alpha = b$ , azaz  $f$  korlátos az  $[a, b]$  intervallumon. Ez azt jelenti, hogy  $f$  korlátos az  $[a, b - \epsilon]$  intervallumon minden  $\epsilon > 0$  mellett.

Másrészt  $f$  folytonos a  $b$  pontban is, ezért korlátos a  $[b - \epsilon, b]$  intervallumon valamely  $\epsilon > 0$  mellett. Tehát  $f$  korlátos az egész  $[a, b]$  intervallumon.  $\square$

**Állítás 131 (Weierstrass-tétel)** *Ha egy  $f : [a, b] \rightarrow \mathbb{R}$  függvény folytonos, akkor van minimuma és maximuma.*

**Bizonyítás.** Lássuk például a maximum esetét. A megelőző tétel miatt az  $f$  korlátos, és így létezik az

$$S = \sup_{x \in [a, b]} f(x)$$

felső határa. Azt kell belátnunk, hogy valahol fel is veszi az  $S$  értéket. Ha ezzel ellentétben az  $f$  nem venné fel az  $S$  értéket, akkor  $f(x) < S$  lenne az  $[a, b]$  minden  $x$  pontjára, hiszen az  $S$  felső korlát. Így a

$$h(x) = \frac{1}{S - f(x)}, \quad x \in [a, b]$$

függvény folytonos lenne az intervallumon. Emiatt a megelőző tétel miatt korlátos lenne, azaz lenne olyan  $K > 0$  szám, hogy

$$\frac{1}{S - f(x)} \leq K, \quad \text{azaz} \quad f(x) \leq S - \frac{1}{K},$$

minden  $x \in [a, b]$  pontban, és így az  $S$  nem lenne felső határ, ellentétben a definíciójával.

A minimum esete hasonlóan bizonyítható.  $\square$

A következő tétel a Bolzano-tételnek egy szép és sokszor használt alkalmazása.

**Állítás 132** Legyen a  $J$  valamilyen intervallum. Ha az  $f : J \rightarrow f(J)$  függvény szigorúan monoton növekedő és folytonos, akkor az

$$f^{-1} : f(J) \rightarrow J$$

inverz függvényre igazak a következők:

- (i) Az  $f^{-1}$  inverz  $f(J)$  értelmezési tartománya is intervallum.
- (ii) Az  $f^{-1}$  is szigorúan monoton növekedő.
- (iii) Az  $f^{-1}$  is folytonos.

Hasonló állítás igaz szigorúan fogyó folytonos függvényre.

**Bizonyítás.** A 77. állítás szerint az inverz létezik és szigorúan monoton. A 127. állítás szerint a  $J$  intervallum  $f(J)$  folytonos képe is intervallum. Ezzel beláttuk az (i) és (ii) állításokat, és már csak a (iii) igazolása van hátra.

Az inverz folytonosságát belátjuk, ha megmutatjuk, hogy az  $x$  egy tetszőleges  $I$  gömb-környezetének az  $f(I)$  képében van gömb-környezete az  $f(x)$  pontnak. Hangsúlyozzuk, hogy a gömb-környezetek a  $J$ -re illetve  $f(J)$ -re nézve értendők, ezért például az  $I$  gömb-környezet olyan is lehet, hogy az  $x$  pont a bal szélére esik (jobboldali gömb).

A 127. állítás szerint az  $f(I)$  egy részintervalluma az  $f(J)$  intervallumnak. Emiatt elegendő azt igazolni, hogy az  $f(x)$  csak akkor eshet az  $f(I)$  valamelyik végpontjába, ha az  $x$  is egyik végpontja a  $I$  intervallumnak. Ennek az indoklása: Ha az  $x$  nem végpontja az  $I$ -nek, akkor  $x \in (c, d) \subseteq I$  valamilyen  $c < d$  mellett. A szigorú monoton növekedés miatt  $f(c) < f(x) < f(d)$ , és így

$$f(x) \in (f(c), f(d)) \subseteq f(I),$$

tehát az  $f(x)$  nem végpontja az  $f(I)$ -nek, így nyilván  $f(I)$  tartalmazza  $f(x)$ -nek egy alkalmas környezetét.  $\square$

### 3.3 Határérték

Ebben a pontban metrikus térben értelmezett függvények határértékének a fogalmát ismerjük meg.

#### 3.3.1 Véges határérték végesben

A függvény határérték fogalma ebből a kérdésből eredeztethető:

Ha egy  $A$  halmazon értelmezett függvényt értelmezni szeretnénk egy olyan  $b$  pontban, amelyik nincs feltétlenül benne a halmazban, de az  $A$  halmaz pontjaival tetszőlegesen megközelíthető, akkor minek kell választani a függvényt a  $b$  helyen, ha azt akarjuk, hogy ott folytonos legyen?

Először metrikus terek esetében definiáljuk a fogalmat, azután azonban, szűkítve az általánosságot, a valós függvények esetében megyünk csak részletekbe.

**Definíció 133** Legyenek  $(X, d_X)$  és  $(Y, d_Y)$  metrikus terek,  $A$  az  $X$  tér részhalmaza,  $a$  pedig az  $A$  halmaz egy torlódási pontja. Egy  $f : A \rightarrow Y$  függvény  $b$  pontban vett határértékének mondjuk az  $y \in Y$  elemet, ha az

$$\tilde{f}(x) \stackrel{\text{def}}{=} \begin{cases} f(x) & \text{ha } x \in A \setminus \{b\} \\ y & \text{ha } x = b \end{cases}$$

módon definiált  $\tilde{f} : A \cup \{b\} \rightarrow Y$  függvény  $a$  pontban folytonos.

Az  $y$  határértékre a

$$\lim_{x \rightarrow b, x \in A} f(x)$$

jelölést használjuk, és azt mondjuk, hogy az  $y$  az  $f$  határértéke (limesze)  $a$  pontnál (az  $A$  halmazra nézve), vagy: az  $f$  tart az  $y$ -hoz, ha az  $x$  tart (az  $A$ -ban)  $a$ -hez.

További jelölési formák még:

$$\lim_{\substack{x \rightarrow b \\ x \in A}} f(x) = y_0,$$

$$f(x) \rightarrow y_0, \quad \text{ha } x \rightarrow b, \quad x \in A.$$

Vessünk fel néhány egyszerű példát valós függvények esetében:

**Példa 3.13** Nézzük meg a következő, formulákkal értelmezett valós leképezéseket. Hol értelmezettek és hol nem; hol célszerű határértéket keresni.

$$1) \quad x \rightarrow \frac{(x+2)^2 - 4}{x}.$$

$$2) \quad x \rightarrow \frac{\sin x}{x}.$$

$$3) \quad \frac{x}{|x-1|}.$$

Az első függvény minden pontban folytonos, kivéve az  $x = 0$  helyet, ahol nincs is értelmezve. Ha a számlálót is megnézzük a nullánál, akkor azt találjuk, hogy az is nulla. A határérték meghatározása itt a nullánál lehet érdekes.

A második függvény a nullánál nincs értelmezve, mivel ott a nevező nulla. Itt kell majd határértéket keresni. Minden más pontban a függvény folytonos.

A harmadik függvény mindenhol folytonos, kivéve az  $x = 1$ -et. Ott a számláló 1, a nevező pedig nulla, ezért úgy vélhetjük, hogy az 1 “közelében igen nagy” a hányados.  $\square$

A következő állítás a 133. definíció átírása, figyelembe véve a folytonosság definícióját.

**Állítás 134** Legyenek  $(X, d_X)$  és  $(Y, d_Y)$  metrikus terek,  $A$  az  $X$  tér részhalmaza,  $a$  pedig az  $A$  halmaz egy torlódási pontja. Egy  $f : A \rightarrow Y$  függvénynek  $a$  pontban az  $y$  pontosan akkor limesze, ha az  $y$  in  $Y$  pont tetszőleges  $V$  gömb-környezetéhez van olyan  $U$  hiányos gömb-környezete a  $b \in A$  pontnak, hogy

$$f(U \cap A) \subseteq V.$$

Ekvivalens megfogalmazás a következő.

**Állítás 135** Legyenek az  $(X, d_X)$  és  $(Y, d_Y)$  metrikus terek,  $A$  az  $X$  egy részhalmaza, és  $b$  az  $A$  egy torlódási pontja. Egy  $f : A \rightarrow Y$  függvénynek az  $y$  pontosan akkor limesze a  $b$  pontban (helyen), ha tetszőleges pozitív  $\epsilon$  számhoz van olyan  $\delta$  pozitív szám, hogy

$$d_Y(f(x), y) < \epsilon, \quad \text{hacsak} \quad x \in A, \quad 0 < d_X(b, x) < \delta.$$

A határérték természetesen vagy létezik, vagy nem, de ha létezik, akkor abban az esetben egyetlen:

**Állítás 136** A 133. definíció szerinti határérték egyetlen.

Bizonyítás. Ha az  $f$  függvénynek a  $b \in A$  helyen az  $y$  és  $z$  pontok is határértékei lennének, akkor az  $\epsilon = d(y, z)/3$  sugarú,  $y$ , illetve  $z$  körüli gömbök egyaránt tartalmaznák a  $b$  valamely környezetének képét. Ez azonban lehetetlen, hiszen e gömbök diszjunktak.  $\square$

**Állítás 137** Legyen  $A \subseteq \mathbb{R}$ ,  $b$  az  $A$  egy torlódási pontja és  $f : A \rightarrow \mathbb{R}$ . Az  $\alpha \in \mathbb{R}$  a határértéke az  $f$  függvénynek a  $b$  pontban (az  $A$ -ra nézve), ha tetszőleges pozitív  $\epsilon$  számhoz van olyan  $\delta$  pozitív szám, hogy

$$|f(x) - \alpha| < \epsilon, \quad \text{hacsak} \quad x \in A, \quad 0 < |x - b| < \delta.$$

Az  $A$  halmaz, amin a függvény értelmezve van, a gyakorlatban rendszerint a következő három típusú, már definiált halmaz lesz:

I. Az  $A$  egy hiányos gömb a  $b$  pont körül, azaz valamilyen pozitív  $r$  számmal

$$\{x \in \mathbb{R} : 0 < |x - b| < r\} = \{x : b - r < x < b + r\} \setminus \{b\}.$$

Ekkor a határérték jelölése:

$$\lim_{x \rightarrow b} f(x).$$

II. Az  $A$  halmaz egy jobboldali hiányos gömb, azaz valamilyen pozitív  $r$  számmal

$$\{x \in \mathbb{R} : b < x < b + r\}.$$

Ekkor a  $b$  pontban vett jobboldali határértékről beszélünk, és a jelölés:

$$\lim_{x \searrow b} f(x).$$

III. Az  $A$  halmaz egy baloldali hiányos gömb, azaz valamilyen  $r$  pozitív számmal

$$\{x \in \mathbb{R} : b - r < x < b\}.$$

Ekkor baloldali határértékről fogunk beszélni, és a jelölés

$$\lim_{x \nearrow b} f(x).$$

Az általunk vizsgált függvények rendszerint valamilyen intervallumon értelmezettek, és a bal- illetve jobboldali határértéket rendszerint az intervallum végpontjaiban kell meghatározni. A ferde nyilak arra utalnak, hogy csökkenőleg vagy fogyólag tartunk-e a határérték helyéhez.

A definíció szerint igaz a következő állítás:

**Állítás 138** *Ha fennáll a*

$$\lim_{x \searrow b} f(x) = \lim_{x \nearrow b} f(x),$$

*akkor létezik a  $b$  pontban határérték is, éspedig*

$$\lim_{x \rightarrow b} f(x) = \lim_{x \searrow b} f(x) = \lim_{x \nearrow b} f(x).$$

A következő állítás teljesen magától értetődő.

**Állítás 139** *A 133. definíció jelöléseivel, ha az  $f : A \rightarrow Y$  függvény folytonos egy  $b \in A$  torlódási pontban, akkor a  $b$  helyen vett határértéke megegyezik az  $f(b)$  helyettesítési értékével:*

$$\lim_{x \rightarrow b} f(x) = f(b).$$

Foglalkozzunk a továbbiakban csak valós függvényekkel, habár az elmondottak, megfelelő változtatással általánosabban is igazak.

Teljesen evidens, hogy ha az  $f$  és  $g$  függvények megegyeznek a  $b$  pont egy hiányos környezetében, akkor

$$\lim_{x \rightarrow b} f(x) = \lim_{x \rightarrow b} g(x).$$

Ez a nyilvánvaló tény sok határérték kiszámításánál döntő eszköz. Ha ugyanis, mondjuk az  $f$  függvény nincs értelmezve a  $b$  helyen, de a  $g$  ott folytonos, akkor azonnal kész a számolás:

$$\lim_{x \rightarrow b} f(x) = g(b).$$

Erre alapozva most megoldunk néhány feladatot.

**Példa 3.14**

$$\lim_{x \rightarrow 0} \frac{(x+2)^2 - 4}{x} = ?$$

A tört nevezője és számlálója a nullánál nulla, de a függvény minden más helyen folytonos, tehát definiált a nulla egy hiányos környezetében. A nulla hiányos környezetében megengedett a következő átalakítás:

$$\frac{(x+2)^2 - 4}{x} = \frac{x^2 + 4x}{x} = x + 4.$$

Eszerint az a függvény, amelynek a nullánál a határértékét keressük, a nulla hiányos környezetében megegyezik az  $x \mapsto (x+4)$  folytonos függvénnyel, ezért a példát megelőző bekezdésben mondottak alapján

$$\lim_{x \rightarrow 0} \frac{(x+2)^2 - 4}{x} = (x+4)|_{x=0} = 0 + 4 = 4.$$

□

**Példa 3.15**

$$\lim_{x \rightarrow 3} \frac{x^2 - 2x - 3}{x^2 - 5x + 6} = ?$$

A függvénynek mind a számlálója mind a nevezője nulla a 3 helyen. A számláló és nevező másodfokú polinomok, ezért mindkettő felírható gyöktényezős alakban:

$$x^2 - 2x - 3 = (x-3)(x+1)$$

és

$$x^2 - 5x + 6 = (x-3)(x-2).$$

Ezek alapján:

$$\lim_{x \rightarrow 3} \frac{x^2 - 2x - 3}{x^2 - 5x + 6} = \lim_{x \rightarrow 3} \frac{(x-3)(x+1)}{(x-3)(x-2)} = \lim_{x \rightarrow 3} \frac{x+1}{x-2} = 4.$$

□

**Példa 3.16** A  $\operatorname{sgn}$  előjel függvénynek határozzuk meg a 0 pontban a jobboldali és baloldali határértékeit.

Az előjel függvény a pozitív számokon az 1 értéket veszi fel, ezért

$$\lim_{x \searrow 0} \operatorname{sgn}(x) = 1.$$

Hasonló indoklással:

$$\lim_{x \nearrow 0} \operatorname{sgn}(x) = -1.$$

□

Az alpont hátralévő részében egyrészt megvizsgáljuk, hogy miként viselkedik a határérték a függvények közötti műveletekkel (leképezés-építési eljárásokkal) szemben; másrészt kiszámolunk néhány nevezetes határértéket. Ezeknek a segítségével viszonylag nagyszámú határértéket tudunk majd meghatározni.

**Állítás 140 (Formális szabályok)** Tegyük fel, hogy az  $f$  és  $g$  valós függvényeknek van határértékük a  $b$  helyen. Ekkor fennállnak a következők.

- (1)  $\lim_{x \rightarrow b} (f + g)(x) = \lim_{x \rightarrow b} f(x) + \lim_{x \rightarrow b} g(x).$
- (2)  $\lim_{x \rightarrow b} (f \cdot g)(x) = \lim_{x \rightarrow b} f(x) \cdot \lim_{x \rightarrow b} g(x).$
- (3)  $\lim_{x \rightarrow b} (\alpha f)(x) = \alpha \cdot \lim_{x \rightarrow b} f(x)$ , ha az  $\alpha$  tetszőleges valós szám.
- (4)  $\lim_{x \rightarrow b} \frac{f}{g}(x) = \frac{\lim_{x \rightarrow b} f(x)}{\lim_{x \rightarrow b} g(x)}$ , feltéve, hogy  $\lim_{x \rightarrow b} g(x) \neq 0$ .

Pontosan ilyen állítások igazak az egyoldali határértékek esetében is.

**Bizonyítás.** Az állítások mindegyike egyszerű következménye a folytonos függvényekre vonatkozó formális szabályoknak és a 133. definíciónak. Példaképpen lássuk az (1) állítást.

**(1):** A 133. definícióban megadott  $\tilde{f}$  és  $\tilde{g}$  függvények folytonosak a  $b$  helyen, ezért folytonos az  $\tilde{f} + \tilde{g}$  összegfüggvény is; de ez azt jelenti, hogy a határérték megegyezik a helyettesítési értékkel:

$$\begin{aligned} \lim_{x \rightarrow b} (f(x) + g(x)) &= \lim_{x \rightarrow b} (\tilde{f}(x) + \tilde{g}(x)) = \\ &= \tilde{f}(b) + \tilde{g}(b) = \lim_{x \rightarrow b} f(x) + \lim_{x \rightarrow b} g(x). \end{aligned}$$

□

A közvetett függvényre vonatkozólag két egyszerű állítást is megfogalmazunk:

**Állítás 141** (a) Ha az  $f$  leképezésnek van határértéke a  $b$  helyen, a  $g$  függvény pedig folytonos a  $\lim_{x \rightarrow b} f(x)$  helyen, akkor

$$\lim_{x \rightarrow b} g(f(x)) = g\left(\lim_{x \rightarrow b} f(x)\right).$$

(b) Ha az  $f$  leképezés folytonos a  $b$  helyen, és injektív a  $b$  egy környezetében, a  $g$ -nek pedig van határértéke az  $f(b)$  helyen, akkor

$$\lim_{x \rightarrow b} g(f(x)) = \lim_{y \rightarrow f(b)} g(y).$$

**Bizonyítás. (a):** Mivel az  $\tilde{f}$  folytonos a  $b$  pontban, a  $g$  pedig az  $\tilde{f}(b)$  helyen, ezért a  $g \circ \tilde{f}$  leképezés is folytonos a  $b$ -nél, tehát

$$\lim_{x \rightarrow b} g(f(x)) = \lim_{x \rightarrow b} g(\tilde{f}(x)) = g(\tilde{f}(b)) = g\left(\lim_{x \rightarrow b} f(x)\right).$$

(b): Mivel az  $f$  folytonos a  $b$ -nél, a  $\tilde{g}$  leképezés pedig az  $f(b)$ -nél, ezért a  $\tilde{g} \circ f$  függvény is folytonos a  $b$  helyen, tehát

$$\lim_{x \rightarrow b} g(f(x)) = \lim_{x \rightarrow b} \tilde{g}(f(x)) = \tilde{g}(f(b)) = \lim_{y \rightarrow f(b)} g(y)$$

és ezzel indokoltuk az állításokat.  $\square$

#### Állítás 142

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1.$$

Bizonyítás. Elég az  $x > 0$  esetre szorítkozni. A 88. Tétel következő egyenlőtlenségét fogjuk használni

$$\sin x \leq x \leq \tan x. \quad (3.8)$$

Ebből az egyenlőtlenségből egyrészt  $(\sin x)/x \leq 1$ , másrészt  $\cos x \leq (\sin x)/x$ , összefoglalva:

$$\cos x \leq \frac{\sin x}{x} \leq 1.$$

Az egyenlőtlenség baloldala tart az egyhez, mert a koszinusz függvény folytonos és  $\cos 0 = 1$ , a jobboldal pedig 1, ezért a közrefogott  $\frac{\sin x}{x}$  függvény is tart az egyhez.  $\square$

Az előző bizonyítás végén szerepelt a következő könnyen belátható érvelés:

Ha két ugyanazon határértékhez tartó függvény közrefog egy harmadik függvényt, akkor a közrefogott függvény is ugyanazon határértékhez tart. Ezt az állítást (csak magyar körben) szellemesen “rendőr-elvnek” is mondják.

**Példa 3.17** *Határozzuk meg a következő limeszt.*

$$\lim_{x \rightarrow 0} \frac{\sin(2\pi x)}{x}.$$

Az előző feladathoz hasonlóan járunk el:

$$\lim_{x \rightarrow 0} \frac{\sin(2\pi x)}{x} = \lim_{x \rightarrow 0} 2\pi \frac{\sin(2\pi x)}{2\pi x} = 2\pi.$$

Ha a szög mérésénél úgy jártunk volna el, hogy a teljes körül-fordulást 1-gyel mérnénk, akkor a szinusz függvény helyett az  $\phi(x) \doteq \sin(2\pi x)$  függvénnyel kellene dolgoznunk, és erre a szóbanforgó limesz “bonyolultabb” szám lenne. Ez azt mutatja, hogy a  $\pi$  szám használata, harmonikusabbá, szebbé teszi a trigonometrikus függvények használatát.

Mondjuk el, hogy milyen általános recept van arra, hogy a rendelkezésre álló módszereinkkel határértékeket határozzunk meg.

A formális szabályok alkalmazásain kívül, lényegében véve két alaptrükk kombinációjával dolgozhatunk:

- A függvényt, aminek a határértékét meg akarjuk határozni úgy próbáljuk alakítani, hogy egy olyan függvénnyel egyezzék meg a hiányos környezetben, amelyik a környezet középpontjában folytonos. Ha ez sikerül, akkor a határértéket a helyettesítési értékkel kapjuk meg.
- Megfelelő átalakítással olyan formát igyekszünk találni, amelyik valamelyik nevezetes határértéket tartalmazza.

### 3.3.2 A végtelen szerepe a határértékekénél

A valós számok között nem szerepelnek a  $+\infty$  és  $-\infty$  végtelen objektumok. Ennek a fő oka az, hogy elrontanák a test struktúrát, amire a számolásoknál alapvetően támaszkodunk, hiszen nem tudnánk értelmet tulajdonítani például a  $(+\infty - \infty)$  különbségnek. A függvények viselkedésével kapcsolatban azonban valamilyen formában észszerű bevezetni a fogalmi rendszerünkbe a végtelen szimbólumokat is. Ezzel azonban, hangsúlyozottan, nem a számolások körét kívánjuk kibővíteni.

A definíciókat aszerint fogjuk kimondani, hogy a határérték helyére vagy értékre értelmezzük-e a végtelen “értékeket”. Ennek megfelelően a következő eseteket fogjuk megkülönböztetni:

- Véges határérték végtelenben.
- Végtelen határérték végtelenben.
- Végtelen határérték végesben.

Az első esetben azt kell pontosan megfogalmaznunk, hogy mit értsünk azon, hogy egy függvény a plusz vagy mínusz végtelenben egy véges  $\beta$  határértéket vesz fel.

**Definíció 143** *Tegyük fel, hogy az  $f$  valós függvény értelmezési tartománya felülről nem korlátos, Ha létezik olyan  $\beta$  szám, hogy tetszőleges  $\epsilon$  pozitív számhoz van olyan  $p$  szám, hogy*

$$|f(x) - \beta| < \epsilon, \quad \text{ha} \quad p < x,$$

*akkor azt mondjuk, hogy a  $\beta$  az  $f$  függvény határértéke a plusz végtelenben, azaz*

$$\lim_{x \rightarrow +\infty} f(x) = \beta.$$

Az előző alpontban részletezték szerint

$$|f(x) - \beta| \leq \epsilon, \quad \text{ha} \quad p \leq x$$

is írható lett volna a definícióban, és a “ $\leq$ ”, “ $<$ ” jelek egyéb “keverése” is megengedett.

Röviden így mondhatnánk a definíciót: A függvény előírt pontossággal megközelíti a végtelenben vett határértéket, ha a független változó értéke megfelelően nagy.

Természetes változtatásokkal kapjuk a minusz végtelenben vett véges határérték definícióját:

**Definíció 144** *Tegyük fel, hogy az  $f$  valós függvény értelmezési tartománya alulról nem korlátos. Ha létezik olyan  $\beta$  szám, hogy tetszőleges  $\epsilon$  pozitív számhoz van olyan  $p$  szám, hogy*

$$|f(x) - \beta| \leq \epsilon, \quad \text{ha} \quad p \leq x,$$

*akkor azt mondjuk, hogy a  $\beta$  az  $f$  függvény határértéke a minusz végtelenben, jelölésben:*

$$\lim_{x \rightarrow -\infty} f(x) = \beta.$$

Megjegyezzük, hogy a “ $+\infty$ ” és “ $-\infty$ ” szimbólumokkal ténylegesen bővíthetnénk az  $\mathbb{R}$  halmazt, és bevezethetnénk olyan metrikát, amelyik az  $\mathbb{R}$ -en megegyezik az euklideszi metrikával, és a  $+\infty$  nyílt gömb-környezetei az

$$(a, +\infty), \quad a \in \mathbb{R}$$

korlátlan intervallumok, a zárt gömb-környezetei pedig az

$$[a, +\infty), \quad a \in \mathbb{R}$$

intervallumok lennének. Hasonlót mondhatnánk a  $-\infty$  gömb-környezeteire.

Az ilyen módon kibővített  $\mathbb{R} = \mathbb{R} \cup \{-\infty, +\infty\}$  halmazt *kiterjesztett valós számoknak* nevezzük.

Ilyen tárgyalás mellett nem kellene külön foglalkozni a végesben és végtelenben vett határértékekkel, tömörebbek lehetnének. Ezzel a konvencióval az előző definícióink a 134. ekvivalens definícióra redukálódnak.

Most pedig két olyan egyszerű példát oldunk meg, amelyeken könnyű bemutatni a definíciót.

**Példa 3.18** *Határozzuk meg a*

$$\lim_{x \rightarrow +\infty} \frac{1}{x-1}$$

*határértéket.*

A későbbiekben mondandók alapján azonnal tudunk majd a kérdésre válaszolni, de itt közvetlenül a definíciót akarjuk használni.

A határértéket nullának gondolhatjuk, mivel a nevező nagy számokra nagy, ezért azt kell megmutatnunk, hogy  $\left| \frac{1}{x-1} \right| \leq \epsilon$ , ha az  $x$  megfelelően nagy. Az

$1 < x$  értékekre az abszolútérték jelet elhagyhatjuk, és így az egyenlőtlenséget a következőképpen alakíthatjuk:

$$\frac{1}{x-1} \leq \epsilon \Leftrightarrow 1 \leq (x-1) \cdot \epsilon, \Leftrightarrow x \leq \frac{1+\epsilon}{\epsilon},$$

tehát az  $\frac{1+\epsilon}{\epsilon}$  szám választható a definícióban szereplő  $p$  számnak, azaz

$$\left| \frac{1}{x-1} \right| \leq \epsilon, \quad \text{ha} \quad x > \frac{1+\epsilon}{\epsilon},$$

amivel beláttuk, hogy

$$\lim_{x \rightarrow +\infty} \frac{1}{x-1} = 0.$$

□

**Példa 3.19** Keressük meg az

$$\lim_{x \rightarrow -\infty} \left( \frac{2}{x+1000} - 3 \right)$$

határértéket.

Az előző példa megoldásával azonos módon járunk el, és egyszerű számolással kapható, hogy

$$\left| \frac{2}{x+1000} \right| \leq \epsilon, \quad \text{ha} \quad x < \frac{1+1000\epsilon}{\epsilon},$$

tehát

$$\lim_{x \rightarrow -\infty} \left( \frac{2}{x+1000} - 3 \right) = 3.$$

□

Most pedig a végtelenben vett végtelen határértékek definícióit soroljuk fel.

**Definíció 145** Tegyük fel, hogy az  $f$  valós függvény értelmezési tartománya felülről nem korlátos. Ha tetszőleges  $q$  számhoz van olyan  $p$  szám, hogy

$$f(x) > q, \quad \text{ha} \quad x > p,$$

akkor ezt mondjuk: az  $f$  függvény határértéke a plusz végtelenben plusz végtelen, jelölésben:

$$\lim_{x \rightarrow +\infty} f(x) = +\infty.$$

Ha pedig tetszőleges  $q$  számhoz van olyan  $p$  szám, hogy

$$f(x) < q, \quad \text{ha} \quad x > p,$$

akkor ezt mondjuk: az  $f$  függvény határértéke a plusz végtelenben mínusz végtelen, jelölésben:

$$\lim_{x \rightarrow +\infty} f(x) = -\infty.$$

A formális

“tetszőleges  $q$  számhoz van olyan  $p$  szám, hogy ...”

kifejezés pontosan ezt mondja:

“az  $f(x)$  előírt számnál nagyobb, ha az  $x$  megfelelően nagy”.

A mínusz végtelenben vett végtelen határértékek teljesen azonos módon történő definiálását az olvasóra bízuk. Lássuk végül a végesben vett végtelen határértéket:

**Definíció 146** *Ha az  $f$  függvény definiálva van a  $b$  pont egy hiányos környezetében, akkor, ha minden  $q$  számhoz van olyan  $\delta$  pozitív szám, hogy*

$$f(x) > q, \quad \text{ha} \quad 0 < |x - b| < \delta,$$

*akkor azt mondjuk, hogy az  $f$  határértéke a  $b$  helyen plusz végtelen, jelölésben:*

$$\lim_{x \rightarrow b} f(x) = +\infty.$$

*Ha pedig minden  $q$  számhoz van olyan  $\delta$  pozitív szám, hogy*

$$f(x) < q, \quad \text{ha} \quad 0 < |x - b| < \delta,$$

*akkor a  $b$  helyen mínusz végtelen a határérték.*

A végesben vett, egyoldali, végtelen határérték definícióját is az olvasóra bízuk.

**Példa 3.20** *Határozzuk meg az  $x \mapsto 1/x$  függvénynek a határértékét a nulla helyen.*

Az  $x$  pont környezetében a függvény nagy pozitív és nagy negatív értékeket felvesz, aszerint, hogy a jobb vagy bal oldalról közeledünk-e a nullához. Eszerint a jobboldali és baloldali határérték különböző lesz. Nézzük először a jobboldali limeszt. Azonnal felírhatjuk, hogy

$$\frac{1}{x} \geq q, \quad \text{ha} \quad 0 < x \leq \frac{1}{q} = p,$$

tehát

$$\lim_{x \searrow 0} \frac{1}{x} = +\infty.$$

A baloldali limesz esetében pedig

$$\frac{1}{x} \leq q, \quad \text{ha} \quad -\frac{1}{|q|} = p \leq x < 0,$$

tehát

$$\lim_{x \nearrow 0} \frac{1}{x} = -\infty.$$

A jobb- és baloldali határértékek különböznek, ezért határérték nincs a nulla helyen.  $\square$

A későbbi határérték számolásoknál nem kell mindig visszamenni a definíciókhoz, ahogyan most tettük, hanem támaszkodni fogunk a formális szabályokra és néhány speciális függvény ismert limeszére, ahogyan azt a végesben vett véges limesz esetében is tettük az előző alpontban.

A végtelenben vett véges határérték formális szabályai pontosan olyanok, mint a végesben vett véges határértékekénél (140. tétel). A  $+\infty$  és  $-\infty$  közül az első esetre fogalmazzuk meg az állítást, mert a másik eset teljesen azonos.

**Állítás 147 (Formális szabályok)** *Tegyük fel, hogy az  $f$  és  $g$  valós függvényeknek van határértékük  $a + \infty$ -ben. Ekkor fennállnak a következők.*

- (1)  $\lim_{x \rightarrow +\infty} (f + g)(x) = \lim_{x \rightarrow +\infty} f(x) + \lim_{x \rightarrow +\infty} g(x).$
- (2)  $\lim_{x \rightarrow +\infty} (f \cdot g)(x) = \lim_{x \rightarrow +\infty} f(x) \cdot \lim_{x \rightarrow +\infty} g(x).$
- (3)  $\lim_{x \rightarrow +\infty} (\alpha f)(x) = \alpha \cdot \lim_{x \rightarrow +\infty} f(x),$  ha az  $\alpha$  tetszőleges valós szám.
- (4)  $\lim_{x \rightarrow +\infty} \frac{f}{g}(x) = \frac{\lim_{x \rightarrow +\infty} f(x)}{\lim_{x \rightarrow +\infty} g(x)},$  feltéve, hogy  $\lim_{x \rightarrow +\infty} g(x) \neq 0.$

**Bizonyítás.** A végesben vett véges határértékekre vonatkozó szabályokat a folytonos függvények formális szabályaiból láttuk be. Itt viszont közvetlen igazolásokra van szükség. Mintaként az összegre vonatkozó szabályt látjuk be, a többi a gyakorlatokra hagyjuk.

**(1):** Legyenek az  $\epsilon$  tetszőleges pozitív szám, és a  $p_1$  és  $p_2$  számok olyan nagyok, hogy

$$|f(x) - \lim_{u \rightarrow +\infty} f(u)| \leq \epsilon/2, \quad \text{ha } p_1 \leq x$$

és

$$|g(x) - \lim_{u \rightarrow +\infty} g(u)| \leq \epsilon/2, \quad \text{ha } p_2 \leq x.$$

Ezt a kettőt felhasználva a következőképpen számolhatunk:

$$\begin{aligned} & \left| (f(x) + g(x)) - \left( \lim_{u \rightarrow +\infty} f(u) + \lim_{u \rightarrow +\infty} g(u) \right) \right| \leq \\ & \leq |f(x) - \lim_{u \rightarrow +\infty} f(u)| + |g(x) - \lim_{u \rightarrow +\infty} g(u)| \leq \\ & \leq \epsilon/2 + \epsilon/2 = \epsilon, \quad \text{ha } \max\{p_1, p_2\} \leq x. \end{aligned}$$

amivel az állítást beláttuk.  $\square$

A közvetett függvényre vonatkozólag is megfogalmazunk egy gyakorta használt egyszerű állítást:

**Állítás 148** *Ha az  $f$  leképezésnek van véges határértéke a  $+\infty$ -ben, a  $g$  függvény pedig folytonos a  $\lim_{x \rightarrow +\infty} f(x)$  helyen, akkor*

$$\lim_{x \rightarrow +\infty} g(f(x)) = g\left(\lim_{x \rightarrow +\infty} f(x)\right).$$

**Bizonyítás.** Legyen  $\epsilon$  egy pozitív szám. A  $g$  függvény  $\lim_{x \rightarrow +\infty} f(x)$  helyen való folytonossága miatt van olyan  $\delta$  pozitív szám, hogy

$$\left|g\left(\lim_{x \rightarrow +\infty} f(x)\right) - g(y)\right| < \epsilon, \quad \text{ha} \quad \left|\lim_{x \rightarrow +\infty} f(x) - y\right| < \delta.$$

A limesz definíciója szerint van olyan  $p$ , hogy

$$\left|\lim_{x \rightarrow +\infty} f(x) - f(x)\right| < \delta, \quad \text{ha} \quad p < x.$$

A két kiemelt sor alapján

$$\left|g\left(\lim_{x \rightarrow +\infty} f(x)\right) - g(f(x))\right| < \epsilon, \quad \text{ha} \quad p < x,$$

amivel az állítást beláttuk. □

Arra az esetre, amikor a határérték végtelen, akkor — a teljesség igénye nélkül — másképpen fogalmazzunk meg a formális szabályokat. A megfogalmazás részletezettségére nem törekszünk.

**Állítás 149** (1) *Ha az  $f$  és  $g$  limeszei egy pontban vagy a végtelenben véges, vagy plusz végtelen, akkor az  $f + g$  limesze a két limesz összege (plusz végtelen, ha legalább az egyik plusz végtelen).*

(2) *Ha az  $f$  limesze pozitív, a  $g$  limesze pedig plusz végtelen, akkor a szorzat limesze is plusz végtelen. Ha az  $f$  limesze negatív, akkor a szorzat limesze minusz végtelen.*

(3) *Ha az  $f$  limesze plusz vagy minusz végtelen, akkor a reciprokanak a limesze nulla.*

(4) *Ha az  $f$  limesze nulla, akkor a reciprokanak a limesze plusz végtelen, ha pozitív számokon keresztül tart a nullához, és minusz végtelen, ha negatív számokon keresztül.*

A bizonyításokat a gyakorlatra hagyjuk, és most lássuk néhány fontos egyszerű függvény limeszét a végtelenben.

**Állítás 150** *Fennállnak a következő limesz-relációk.*

$$(1) \quad \lim_{x \rightarrow +\infty} x^n = +\infty, \quad \text{ha} \quad n > 0,$$

$$(2) \quad \lim_{x \rightarrow +\infty} x^n = 0, \quad \text{ha } n < 0,$$

$$(3) \quad \lim_{x \rightarrow +\infty} a^x = +\infty, \quad \text{ha } a > 1.$$

$$(4) \quad \lim_{x \rightarrow +\infty} a^x = 0, \quad \text{ha } 0 < a < 1.$$

$$(5) \quad \lim_{x \rightarrow +\infty} \ln x = +\infty.$$

Bizonyítás. (1): Legyen a  $p$  tetszőleges szám. Az

$$x^n \geq 1 + n(x - 1), \quad \text{ha } 0 < x$$

Bernoulli egyenlőtlenségből

$$x^n \geq p, \quad \text{ha } n(x - 1) \geq p, \quad \text{azaz ha } x \geq p/n + 1.$$

(2): Az előzőből és abból a formális szabályból, hogy a végtelenhez tartó függvény reciproka a nullához tart.

(3): Az  $x \mapsto a^x$  monoton növekedő, ezért a Bernoulli egyenlőtlenséget is használva

$$a^x \geq a^{\text{eg}(x)} \geq 1 + (a - 1)\text{eg}(x),$$

ahol az  $\text{eg}(x)$  az  $x$  egész részét jelöli. Mivel az egyenlőtlenség jobboldala tart a plusz végtelenhez, ha az  $x$  tart a plusz végtelenhez, ezért a baloldala is.

(4): Ugyanúgy adódik az előzőből, mint az (1)-ből a (2).

(5): Az  $\ln x$  függvény monoton növekedő, és ha belátjuk hogy korlátlan, akkor készen is leszünk. Ha ezzel ellentétben korlátos lenne:

$$\ln x \leq K \quad \text{minden } x > 0\text{-ra,}$$

akkor  $x = e^{\ln x} \leq e^K$  adódnék minden pozitív  $x$  mellett, ami nyilvánvalóan lehetetlen.  $\square$

Az exponenciális és logaritmus függvény végtelenben való viselkedésével kapcsolatban látni fogjuk még, hogy minden  $n$ -re

$$\lim_{x \rightarrow +\infty} \frac{e^x}{x^n} = +\infty.$$

és ha az  $\alpha$  pozitív, akkor

$$\lim_{x \rightarrow +\infty} \frac{\ln x}{x^\alpha} = 0.$$

Ezeket is bebizonyíthatnánk most ezen a helyen, de kicsit körülményes lenne, és később könnyen fognak adódni.

**Példa 3.21** *Határozzuk meg a*

$$\lim_{x \rightarrow +\infty} \frac{3x^2 + 7x - 9}{-2x^2 - 6x + 12}$$

*limeszt.*

A számlálót és nevezőt  $x^2$ -tel osztva:

$$\frac{3x^2 + 7x - 9}{-2x^2 - 6x + 12} = \frac{3 + 7\frac{1}{x} - 9\frac{1}{x^2}}{-2 - 6\frac{1}{x} + 12\frac{1}{x^2}}.$$

A számlálóban és a nevezőben az  $1/x$  és  $1/x^2$  függvények tartanak a nullához, ha az  $x$  a végtelenbe tart, így

$$\lim_{x \rightarrow +\infty} \frac{3x^2 + 7x - 9}{-2x^2 - 6x + 12} = -\frac{3}{2}.$$

□

## 4.

# Sorozatok és sorok

Ez a fejezet sorozatok konvergenciájával foglalkozik metrikus terekben.

## 4.1 Sorozatok metrikus terekben

Emlékeztetünk arra, hogy egy  $X$  halmazbeli sorozat a természetes számok  $\mathbb{N}$  halmazán értelmezett  $X$ -be képező, azaz  $a : \mathbb{N} \rightarrow X$  alakú függvény. Az  $n \in \mathbb{N}$  helyen felvett értékét szokásosan  $a_n$ -nel jelöljük. Nevezetesen, ha  $X$  metrikus tér, beszélhetünk az  $a$  leképezés  $+\infty$ -ben vett határértékéről. Ezt fogalmazza meg az alábbi definíció. (Lásd a 3. fejezet 134. állítását.)

**Definíció 151** Legyen  $(X, d)$  metrikus tér. Azt mondjuk, hogy az  $a : \mathbb{N} \rightarrow X$  sorozat határértéke az  $y \in X$  pont, ha az  $y$  pont egy tetszőleges gömb-környezete tartalmazza a sorozat minden tagját egy indextől kezdve. Egy sorozatot konvergensnek nevezünk, ha van határértéke. Ellenkező esetben divergensnek nevezzük.

A definíció más módon megfogalmazva: tetszőleges  $\epsilon$  pozitív számhoz van olyan  $n_0$ , hogy

$$d(y, a_n) \leq \epsilon, \quad \text{ha} \quad n_0 < n.$$

Az  $n_0 < n$  helyett természetesen bárhol  $n_0 \leq n$  is szerepelhetne.

Az előző állításokban az  $\mathbb{R}$  illetve  $\mathbb{R}^p$  euklideszi terek esetében

$$d(y, a_n) = |y - a_n| \quad \text{illetve} \quad d(y, a_n) = \|y - a_n\|,$$

írandók.

Már láttuk, hogy a limesz, ha létezik, egyértelmű.

**Állítás 152** Ha egy  $(X, d)$  metrikus térnek egy  $(a_n)$  sorozata konvergens, akkor minden  $\epsilon$  pozitív számhoz van olyan  $n_0$  index, hogy

$$d(a_n, a_m) < \epsilon \quad \text{ha} \quad n, m \geq n_0.$$

**Bizonyítás.** Legyen adva a tételben szereplő  $\epsilon$  pozitív szám. Az alábbiakban megadjuk a tételben kívánt tulajdonságú  $n_0$  indexet.

Ha a sorozat limesze a  $p$  pont, akkor az  $\epsilon/2$  pozitív számhoz van olyan  $n_0$  index, hogy

$$d(p, a_n) < \epsilon/2 \quad \text{ha} \quad n \geq n_0.$$

Ennek a felhasználásával a háromszög egyenlőtlenség segítségével kapjuk, hogy

$$d(a_n, a_m) \leq d(a_n, p) + d(p, a_m) < \epsilon/2 + \epsilon/2,$$

feltéve hogy  $n, m \geq n_0$ , amivel be is láttuk a tételt.  $\square$

**Definíció 153** *A 152. állítás feltételét Cauchy-feltételnek fogjuk mondani, a neki eleget tevő sorozatokat pedig Cauchy-sorozatoknak nevezzük.*

Ezzel az elnevezéssel az állítás röviden: *Minden konvergens sorozat Cauchy-sorozat.*

**Definíció 154** *Egy metrikus teret teljesnek nevezünk, ha benne minden Cauchy-sorozat konvergens.*

**Példa 4.1** Tekintsük a racionális számok  $Q$  halmazát az euklideszi metrikával. Világos, hogy ebben a térben az

$$\begin{aligned} a_1 &= 1.4 \\ a_2 &= 1.41 \\ a_3 &= 1.414 \\ a_4 &= 1.4142 \\ &\vdots \end{aligned}$$

sorozat Cauchy-sorozat, hiszen bármely  $n < m$  indexekre  $|a_n - a_m| < 10^{-n}$ . Azonban  $a_n \rightarrow \sqrt{2}$ , így a sorozatnak nincs határértéke a  $Q$  térben.

**Definíció 155** *Tekintsünk egy  $a : \mathbb{N} \rightarrow X$  sorozatot, és legyen  $n : \mathbb{N} \rightarrow \mathbb{N}$  egy szigorúan monoton növekvő leképezés. Ekkor az  $a \circ n : \mathbb{N} \rightarrow X$  sorozatot az  $a$  részsorozatának nevezzük. E sorozat  $k$ -ik elemét az  $a_{n_k}$  szimbólummal jelöljük.*

**Állítás 156** *Ha egy Cauchy-sorozatnak létezik konvergens részsorozata, akkor az eredeti sorozat is konvergens.*

**Bizonyítás.** Tekintsük az  $(a_n)$  Cauchy-sorozatot, és tegyük fel, hogy az  $(a_{n_k})$  részsorozat konvergens,  $a_{n_k} \rightarrow x$ . Legyen  $\epsilon > 0$ . Ekkor van olyan  $n_0$  index, hogy egyrészt

$$d(a_n, a_m) < \frac{\epsilon}{2},$$

bármely  $n, m \geq n_0$  esetén, másrészt

$$d(x, a_{n_k}) < \frac{\epsilon}{2},$$

minden  $n_k \geq n_0$  mellett. Tehát tetszőleges  $n \geq n_0$  indexre

$$d(x, a_n) \leq d(x, a_{n_k}) + d(a_{n_k}, a_n) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Ez azt jelenti, hogy  $\lim a_n = x$ .  $\square$

**Definíció 157** Azt mondjuk, hogy egy sorozat korlátos, ha az értékkészlete korlátos halmaz.

Könnyen látható, hogy minden konvergens sorozat, illetve minden Cauchy-sorozat is korlátos.

**Definíció 158** Legyen az  $a_n$  az  $(X, d)$  metrikus térnek egy sorozata. Egy  $x \in X$  elemet a sorozat torlódási pontjának mondjuk, ha az  $x$  pont tetszőleges gömb-környezetébe végtelen sok tagja esik a sorozatnak.

Nyilvánvaló, hogy egy konvergens sorozat határértéke egyúttal torlódási pont is. Torlódási pont nem feltétlenül létezik, például az  $a_n = n$  valós sorozatnak nincs torlódási pontja. Másrészt az  $a_n = (-1)^n$  sorozatnak pontosan két torlódási pontja van:  $-1$  és  $+1$ .

Fontos megjegyezni, hogy egy sorozat torlódási pontjának fogalma különbözik az értékkészletének torlódási pontjától. Például az  $a_n = 1$  konstans sorozatnak az  $1$  torlódási pontja, de a sorozat értékkészletének, az  $\{1\}$  halmaznak persze nincs torlódási pontja.

**Állítás 159** A 158. definíció jelöléseivel, egy  $x$  pont pontosan akkor torlódási pont, ha az  $a_n$  sorozatnak van olyan  $a_{n_k}$  részsorozata, amelyik tart az  $x$ -hez.

**Bizonyítás.** Ha van olyan  $a_{n_k}$  részsorozat, amelyik tart az  $x$  ponthoz, akkor a konvergencia definíciója szerint az  $x$  tetszőleges gömb-környezetébe beleesik a részsorozat minden tagja, véges soktól eltekintve, és így végtelen sok tagja beleesik az  $a_n$  sorozatnak.

Ha pedig az  $x$  torlódási pontja az  $a_n$  sorozatnak, akkor a következőképpen választunk ki egy  $x$ -hez tartó részsorozatot:

Az  $x$   $1$  sugarú gömb-környezetében van egy  $a_{n_1}$  pontja a sorozatnak. Az  $x$   $1/2$  sugarú gömb-környezetében végtelen sok pontja van a sorozatnak, ezért van olyan  $a_{n_2}$  pontja, hogy  $n_1 < n_2$ . Általánosan: ha az

$$a_{n_1}, a_{n_2}, \dots, a_{n_k}, \quad n_1 < n_2 < \dots < n_k$$

tagokat már kiválasztottuk, akkor az  $a_{n_{k+1}}$  tagot kivehetjük az  $x$  pont körüli  $1/k$  sugarú körből, mivel oda végtelen sok tagja esik a sorozatnak, úgy hogy  $n_k < n_{k+1}$ . A kiválasztás módja miatt az  $a_{n_k}$ ,  $k = 1, 2, \dots$  részsorozat tart az  $x$ -hez.  $\square$

**Állítás 160** Egy Cauchy-sorozatnak legfeljebb egy torlódási pontja lehet.

**Bizonyítás.** Tekintsük az  $a_n$  Cauchy-sorozatot. Ha az  $x$  és  $y$  pontok egyaránt torlódási pontok lennének, úgy legyen

$$\epsilon = \frac{1}{3}d(x, y).$$

Mivel  $a_n$  Cauchy-tulajdonságú, található olyan  $n_0$ , hogy bármely  $n, m \geq n_0$  esetén  $d(a_n, a_m) < \epsilon$ . Ez ellentmond annak, hogy az  $x$  és  $y$  pontok  $\epsilon$  sugarú környezeteibe egyaránt végtelen sok elem esik.  $\square$

## 4.2 Számsorozatok

### 4.2.1 A Bolzano-Weierstrass tétel

Mivel egy  $(a_n)$   $\mathbb{R}$ -beli sorozat egy  $a : \mathbb{N} \rightarrow \mathbb{R}$  valós értékű függvény, és a valós értékű függvényekre már bevezettük a monotonitás és korlátosság fogalmát, ezért ezek a fogalmak valós sorozatokra is értelmezve vannak:

**Definíció 161** Egy  $(a_n)$   $\mathbb{R}$ -beli sorozat monoton növekedő (monoton fogyó), ha minden  $n \in \mathbb{N}$  esetén

$$a_n \leq a_{n+1} \quad (a_n \geq a_{n+1}).$$

**Definíció 162** Egy  $(a_n)$   $\mathbb{R}$ -beli sorozat alulról korlátos (felülről korlátos illetve korlátos), ha létezik olyan  $K_1$  valós szám ( $K_2$  valós szám), hogy minden  $n \in \mathbb{N}$  esetén

$$K_1 \leq a_n \quad (a_n \leq K_2, \text{ illetve } K_1 \leq a_n \leq K_2).$$

Vegyük észre, hogy egy  $(a_n)$   $\mathbb{R}$ -beli sorozat korlátossága ekvivalens azzal, hogy mint  $(\mathbb{R}, d_e)$  metrikus térbeli sorozat korlátos. A monoton és korlátos sorozatok nagyon jól viselkednek a határérték szempontjából:

**Állítás 163** Minden  $\mathbb{R}$ -beli monoton és korlátos sorozat konvergens.

Monoton növekedő  $(a_n)$  sorozat esetében:

$$\lim_{n \rightarrow +\infty} a_n = \sup_{n \in \mathbb{N}} a_n,$$

monoton csökkenő esetben pedig:

$$\lim_{n \rightarrow +\infty} a_n = \inf_{n \in \mathbb{N}} a_n.$$

**Bizonyítás.** Legyen az  $(a_n)$  sorozat — mondjuk — monoton növekedő, és  $s = \sup_{n \in \mathbb{N}} a_n$ . Mivel az  $s$  a legkisebb felső korlát, ezért tetszőleges  $\epsilon > 0$  számhoz van olyan  $N$  index, hogy

$$s - \epsilon < a_N \leq s.$$

A monoton növekedés szerint ebből azonnal adódik, hogy

$$s - \epsilon < a_n \leq s, \text{ tehát } |a_n - s| \leq \epsilon \quad \text{ha} \quad N \leq n,$$

ahogyan állítottuk.  $\square$

**Állítás 164** Minden  $\mathbb{R}$ -beli sorozatnak van monoton részsorozata.

Bizonyítás. A sorozat egy  $a_m$  tagját nevezzük "csúcsnak", ha

$$a_n \leq a_m, \quad \text{ha} \quad m \leq n.$$

Ha végtelen sok  $a_{n_k}$ ,  $k = 1, 2, \dots$  csúcs van, akkor a csúcs definíciója szerint  $a_{n_k} \geq a_{n_{k+1}}$ , tehát az  $a_{n_k}$  részsorozat monoton fogyó.

Azt az esetet kell még megnéznünk, amikor csak véges sok csúcs van. Ekkor van olyan  $n_1$  index, hogy  $a_{n_1}$  nem csúcs, és ennél nagyobb csúcselem sincs. A részsorozat első eleme legyen ez az  $a_{n_1}$ . Mivel ez nem csúcselem, ezért van olyan  $a_{n_2}$ , ( $n_1 < n_2$ ) elem, hogy  $a_{n_2} > a_{n_1}$ . Mivel az  $a_{n_2}$  sem csúcselem, ezért van olyan  $a_{n_3}$ , ( $n_2 < n_3$ ) elem, hogy  $a_{n_3} > a_{n_2}$ . Általában ha már  $a_{n_k}$ -ig kiválasztottuk a részsorozat elemeit, akkor mivel az  $a_{n_k}$  nem csúcselem, van olyan  $a_{n_{k+1}}$  elem, hogy  $a_{n_{k+1}} > a_{n_k}$  ( $n_k < n_{k+1}$ ). Az így kiválasztott sorozat nyilvánvalóan (szigorúan) monoton növekedő.  $\square$

A következő tétellel sokszor fogunk találkozni.

**Tétel 165 (Bolzano-Weierstrass)** Minden  $\mathbb{R}$ -beli korlátos sorozatnak van konvergens részsorozata.

Mivel láttuk, hogy konvergens részsorozat léte torlódási pontot garantál, így a fenti tételt az alábbi módon is szokták fogalmazni. Minden  $\mathbb{R}$ -beli korlátos sorozatnak van torlódási pontja.

Bizonyítás. Az előző két állítás közvetlen folyománya.  $\square$

### 4.2.2 Valós Cauchy-sorozatok

Az alábbi állítás azt mondja, hogy az  $(\mathbb{R}, d_e)$  euklideszi metrikus tér teljes.

**Tétel 166** Minden  $\mathbb{R}$ -beli Cauchy-sorozat konvergens.

Bizonyítás. Láttuk, hogy minden Cauchy-sorozat korlátos, így van konvergens részsorozata, de ha egy Cauchy-sorozatnak van konvergens részsorozata akkor az konvergens is.  $\square$

### 4.2.3 Limesz szuperior és limesz inferior

A valós sorozatok tulajdonságainak a jellemzésénél gyakorta használjuk a következőkben bevezetett fogalmakat.

Ha az  $a_n$  egy tetszőleges korlátos, valós számsorozat, akkor a

$$b_n \stackrel{\text{def}}{=} \sup\{a_n, a_{n+1}, \dots\}$$

módon definiál  $b_n$  sorozat monoton csökkenő, mivel nagyobb  $n$ -re kevesebb szám szuprémumát kell venni. Emiatt van véges határértéke.

Hasonlóan, ha az  $a_n$  egy tetszőleges korlátos, valós számsorozat, akkor a

$$c_n \stackrel{\text{def}}{=} \inf\{a_n, a_{n+1}, \dots\}$$

módon definiált  $c_n$  sorozat monoton növekedő, ezért van véges határértéke.

A mondottak miatt helyes a következő definíció:

**Definíció 167** Ha az  $a_n$  valós sorozat felülről korlátos, akkor a sorozat limesz szuperiorjának mondjuk a

$$\lim_{n \rightarrow +\infty} \sup a_n \stackrel{\text{def}}{=} \inf_n \sup\{a_n, a_{n+1}, \dots\}$$

valós számot, a jelölése:

$$\limsup_{n \rightarrow +\infty} a_n, \quad \text{vagy egyszerűen:} \quad \limsup a_n.$$

Hasonlóan, alulról korlátos sorozatra:

$$\lim_{n \rightarrow +\infty} \inf a_n \stackrel{\text{def}}{=} \sup_n \inf\{a_n, a_{n+1}, \dots\}$$

valós számot a sorozat limesz inferiorjának nevezzük, és a jelölése:

$$\liminf_{n \rightarrow +\infty} a_n, \quad \text{vagy egyszerűen:} \quad \liminf a_n.$$

A következőkben a limesz szuperiort és inferiort fogjuk más módon is jellemezni.

**Állítás 168** 1. A limesz szuperiornál nagyobb számnál csak véges sok nagyobb tagja lehet a sorozatnak.

2. A limesz inferiornál kisebb számnál legfeljebb véges sok kisebb tagja lehet a sorozatnak.

**Bizonyítás.** Csak a limesz szuperior esetét bizonyítjuk. Legyen  $s = \limsup a_n$  és  $\epsilon > 0$ . Ekkor van olyan  $n \in \mathbb{N}$  természetes szám, melyre

$$s + \epsilon \geq \sup\{a_n, a_{n+1}, a_{n+2}, \dots\}$$

ami azt jelenti, hogy az  $s + \epsilon$ -nál nagyobb sorozat érték csak az  $1, \dots, n-1$  indexek közül kerülhet ki.  $\square$

A limesz szuperior és inferior jó és szemléletes karakterizációját adja a következő állítás:

**Állítás 169** *Egy korlátos, valós sorozat limesz szuperiorja a sorozat torlódási pontjai összességének a maximuma, a limesz inferior pedig a minimuma.*

**Bizonyítás.** Lássuk a limesz szuperior esetét.

Legyen az  $(a_n)$  egy valós korlátos sorozat, és jelölje a limesz szuperiort  $s$ . Mivel a sorozatnak  $s + \epsilon$ -nál csak véges sok nagyobb eleme van, ez azt jelenti, hogy  $s + \epsilon$ -nál nagyobb torlódási pontja nem lehet, viszont  $\epsilon > 0$  tetszőleges volta garantálja, hogy  $s$ -nél nagyobb torlódási pont sem lehet.

Az van még hátra hogy belássuk: az  $s$  torlódási pont, azaz van hozzá tartó részsorozat. Legyen  $k \in \mathbb{N}$  természetes szám. A infimum definíciója miatt létezik  $n_k^* \in \mathbb{N}$  természetes szám, melyre

$$s + \frac{1}{k} > \sup\{a_{n_k^*}, a_{n_k^*+1}, \dots\}.$$

Viszont a szuprémum definíciója szerint létezik olyan  $n_k \geq n_k^*$  természetes szám, melyre

$$s - \frac{1}{k} \leq \sup\{a_{n_k^*}, a_{n_k^*+1}, \dots\} - \frac{1}{k} < a_{n_k}.$$

Az így kiválasztott  $a_{n_k}$  sorozat elemre

$$a_{n_k} \in \left(s - \frac{1}{k}, s + \frac{1}{k}\right),$$

tehát az  $a_{n_k}$  részsorozat tart  $s$ -hez.  $\square$

A fenti állítás speciálisan adja azt, hogy egyáltalán van torlódási pont, és így van egy ahhoz tartó részsorozat. Ezzel egy új bizonyítást adtunk a Bolzano-Weierstrass-féle kiválasztási tételre is.

A limesz szuperior és inferior szoros kapcsolatban van a konvergenciával:

**Állítás 170** *Egy valós sorozat pontosan akkor konvergens, ha a limesz szuperior és limesz inferior megegyeznek.*

**Bizonyítás.** Konvergens sorozatnak, könnyen láthatóan, nincs a határértékén kívül más torlódási pontja, ezért a határérték egyben a limesz szuperior és inferior is. Ha pedig a limesz szuperior és inferior megegyeznek, akkor a közös  $s$  értékének az  $[s - \epsilon, s + \epsilon]$ , ( $\epsilon > 0$ ) környezetén kívül csak véges sok tagja lehet a sorozatnak, tehát az  $s$  a limesze.  $\square$

#### 4.2.4 Határérték és műveletek

A sorozatokra vonatkozó műveleteket már definiáltuk tekinthetjük, mivel a valós értékű függvények közötti műveleteket már tanulmányoztuk. A határérték és a műveletek közötti kapcsolatokat most sorozatokra is felelevenítjük.

**Állítás 171** *Tartsanak az  $a_n$  illetve  $b_n$  valós sorozatok az  $a$  illetve  $b$  határértékekhez.*

(1) *Az  $(a_n + b_n)$  összegsorozat határértéke:*

$$\lim_{n \rightarrow +\infty} (a_n + b_n) = \lim_{n \rightarrow +\infty} a_n + \lim_{n \rightarrow +\infty} b_n.$$

(2) *Ha  $\lambda$  tetszőleges valós szám, akkor a  $\lambda a_n$  sorozat határértéke:*

$$\lim_{n \rightarrow +\infty} (\lambda a_n) = \lambda \lim_{n \rightarrow +\infty} a_n.$$

(3) *Az  $(a_n b_n)$  szorzatsorozat határértéke:*

$$\lim_{n \rightarrow +\infty} a_n b_n = \lim_{n \rightarrow +\infty} a_n \cdot \lim_{n \rightarrow +\infty} b_n.$$

(4) *Ha  $b \neq 0$ , akkor az  $a_n/b_n$  tört nevezője bizonyos tagtól kezdve nem nulla, és a hányados-sorozat határértéke:*

$$\lim_{n \rightarrow +\infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow +\infty} a_n}{\lim_{n \rightarrow +\infty} b_n}.$$

**Bizonyítás.** Mivel a sorozatok függvények, és most ezeknek a plusz végtelenben vett határértékéről van szó, már beláttuk az állításokat.  $\square$

#### 4.2.5 Sorozatok végtelen határértéke

A függvények végtelen határértékének az ismeretében nem jelent semmi újat a valós sorozatok végtelen határértéke:

**Definíció 172** *Az  $(a_n)$  valós sorozat határértéke plusz végtelen  $(+\infty)$ , ha tetszőleges  $K$  számhoz van olyan  $n_0$  index, hogy*

$$a_n > K, \quad \text{ha} \quad n_0 < n.$$

*Minusz végtelen  $(-\infty)$  a határértéke, ha tetszőleges  $K$  számhoz van olyan  $n_0$  index, hogy*

$$a_n < K, \quad \text{ha} \quad n_0 < n.$$

A végtelenbe tartó sorozatoknak is vannak megfelelő formális szabályaik. Például ha az  $a_n$  és  $b_n$  sorozatoknak van véges vagy plusz végtelen határértéke, akkor

$$\lim_{n \rightarrow +\infty} (a_n + b_n) = \lim_{n \rightarrow +\infty} a_n + \lim_{n \rightarrow +\infty} b_n.$$

Az egyes formális szabályok olyan egyszerűek, hogy minden adott esetben könnyen meggondolhatóak, ezért nem is részletezzük, mivel a függvények esetében is beszéltünk már róluk.

**Példa 4.2** *Igazoljuk, hogy*

$$\lim_{n \rightarrow +\infty} \sqrt[n]{n!} = +\infty.$$

A bizonyítás alapötlete a következő egyszerű becslés: az  $n!$  szorzat tényezőinek első felét elhagyjuk, a második felének pedig minden tényezőjét a legkisebbel helyettesítjük:

$$n! = 1 \cdot 2 \cdots \operatorname{eg}\left(\frac{n}{2}\right) \cdots n \geq \operatorname{eg}\left(\frac{n}{2}\right) \cdots n \geq \left(\frac{n}{2} - 1\right)^{\frac{n}{2}-2}.$$

Ebből

$$\sqrt[n]{n!} \geq \left(\frac{n}{2} - 1\right)^{\frac{1}{2}-1}.$$

Mivel a jobboldal tart a plusz végtelenhez, ezért igaz az állítás.  $\square$

#### 4.2.6 $\mathbb{R}^2$ -beli sorozatok

A valós számok sorozataira már konkretizáltuk a metrikus térben definiált konvergencia fogalmát. A következő két tételben megmutatjuk, hogy az  $\mathbb{R}^2$  vektortérben (az euklideszi metrika mellett) és a komplex számok testében (az abszolút érték által definiált metrika mellett) a konvergencia a valós számok sorozatainak a konvergenciájára vezethető vissza.

**Állítás 173** *Egy  $(\alpha_n, \beta_n)$ ,  $n = 1, 2, \dots$  vektorsorozat pontosan akkor konvergál egy  $(\alpha, \beta)$  vektorhoz az  $(\mathbb{R}^2, d_e)$  euklideszi térben, ha koordinátáinként konvergál, azaz*

$$\lim_{n \rightarrow +\infty} \alpha_n = \alpha \quad \text{és} \quad \lim_{n \rightarrow +\infty} \beta_n = \beta.$$

**Bizonyítás.** Tartson először az  $(\alpha_n, \beta_n)$  vektorsorozat az  $(\alpha, \beta)$  vektorhoz, azaz tetszőleges pozitív  $\epsilon$  számhoz van olyan  $n_0$  index, hogy

$$\sqrt{(\alpha_n - \alpha)^2 + (\beta_n - \beta)^2} < \epsilon, \quad \text{ha} \quad n > n_0.$$

Ebből nyilvánvalóan

$$|\alpha_n - \alpha| < \epsilon \quad \text{és} \quad |\beta_n - \beta| < \epsilon,$$

ha  $n > n_0$ . Ez pedig éppen azt mondja, hogy az  $\alpha_n$  sorozat az  $\alpha$ -hoz, a  $\beta_n$  pedig a  $\beta$ -hoz tart.

Az állítás másik irányának a bizonyításához tegyük fel, hogy a koordináták sorozata konvergál, azaz adott  $\epsilon$  pozitív számhoz van olyan  $n_0$  index, hogy

$$|\alpha_n - \alpha| < \frac{\epsilon}{\sqrt{2}} \quad \text{ha} \quad n' < n,$$

és

$$|\beta_n - \beta| < \frac{\epsilon}{\sqrt{2}}, \quad \text{ha} \quad n'' < n.$$

Ez alapján:

$$\sqrt{(\alpha_n - \alpha)^2 + (\beta_n - \beta)^2} < \sqrt{\frac{\epsilon^2}{2} + \frac{\epsilon^2}{2}} = \epsilon,$$

ha  $n_0 > \max\{n', n''\}$ , tehát a vektorsorozat tart az  $(\alpha, \beta)$  vektorhoz.  $\square$

Az  $\mathbb{R}^2$  vektortér, ezért a vektor összeadást és a számmal való szorzást kell megnéznünk:

**Állítás 174** Az  $(a_n, b_n)$  illetve  $(c_n, d_n)$  vektor-sorozatokat tartanak az

$$(a, b) \quad \text{illetve} \quad (c, d)$$

vektorokhoz.

(i) Az  $[(a_n, b_n) + (c_n, d_n)]$  összeg-sorozat határértéke:

$$\lim_{n \rightarrow +\infty} [(a_n, b_n) + (c_n, d_n)] = \lim_{n \rightarrow +\infty} (a_n, b_n) + \lim_{n \rightarrow +\infty} (c_n, d_n).$$

(ii) Legyen a  $\gamma$  tetszőleges valós szám. A  $\lambda(a_n, b_n)$  számmal-szorozott sorozat határértéke:

$$\lim_{n \rightarrow +\infty} \lambda(a_n, b_n) = \lambda \cdot \lim_{n \rightarrow +\infty} (a_n, b_n).$$

**Bizonyítás.** Az állítások mindegyike a 173. Állítás és 171. Állításokból következik. Lássuk például az (i) állítás igazolását.

Az 173. állítás szerint az  $a_n, b_n, c_n$  és  $d_n$  számsorozatok rendre az  $a, b, c$  és  $d$  számokhoz konvergálnak. Ezek alapján az

$$(a_n, b_n) + (c_n, d_n) = (a_n + c_n, b_n + d_n)$$

összeg-sorozat koordinátáinak a sorozata, a 171. (1) állítás felhasználásával az  $(a + c)$  és  $(b + d)$  számokhoz tart, és így ismét a 173. állítás szerint az összeg-sorozat tart az

$$(a + c, b + d) = (a, b) + (c, d)$$

vektorhoz, amit igazolni kellett.  $\square$

A Bolzano-Weierstrass-féle kiválasztási tétel igaz az  $\mathbb{R}^2$  euklideszi tér esetében is:

**Állítás 175** *Ha  $\mathbb{R}^2$ -beli vektorok egy sorozata korlátos, akkor van konvergens részsorozata.*

**Bizonyítás.** Legyen  $(x_n, y_n)$  a szóban forgó sorozat. A feltétel szerint van olyan  $K$  korlát, hogy

$$\sqrt{x_n^2 + y_n^2} \leq K, \text{ minden } n\text{-re,}$$

ezért  $|x_n| \leq K$  és  $|y_n| \leq K$ , minden  $n$ -re, az első és második koordináták sorozatai is korlátosak. A valós sorozatokra vonatkozó Bolzano-Weierstrass-tétel szerint az első koordináták sorozatának van valamilyen  $x$  számhoz tartó  $x_{n_k}$  konvergens részsorozata. Az  $(x_{n_k}, y_{n_k})$  sorozat második koordinátájára megismételve az előző lépést: van olyan  $y_{n_{k_i}}$  részsorozat, amelyik konvergál valamilyen  $y$  számhoz. Összefoglalva: A  $(x_{n_{k_i}}, y_{n_{k_i}})$  sorozat a 181. állítás szerint tart az  $(x, y)$  komplex számhoz.  $\square$

Most pedig megmutatjuk, hogy nemcsak a valós számok  $\mathbb{R}$  teste teljes metrikus tér az euklideszi metrika mellett, hanem az  $\mathbb{R}^2$  euklideszi tér is teljes:

**Állítás 176** *Az  $\mathbb{R}^2$  euklideszi térben minden Cauchy-sorozat konvergens.*

**Bizonyítás.** Legyen a  $z_n = (x_n, y_n)$  egy Cauchy-sorozat az  $\mathbb{R}^2$ -ben. Ekkor az  $x_n$  és  $y_n$  sorozatok Cauchy-sorozatok az  $\mathbb{R}$ -ben:

A feltétel szerint tetszőleges  $\epsilon > 0$  számhoz van olyan  $N$ , hogy

$$\sqrt{(x_n - x_m)^2 + (y_n - y_m)^2} \leq \epsilon, \quad \text{ha} \quad N \leq n, m,$$

amiből

$$|x_n - x_m| \leq \epsilon, \quad \text{ha} \quad N \leq n, m,$$

és azonos igaz a másik koordináta sorozatra.

A valós számok teljessége miatt a koordináták Cauchy-sorozata konvergens, és így a 173. állítás szerint maga a  $z_n \in \mathbb{R}^2$  sorozat is konvergens, tehát az  $\mathbb{R}^2$  euklideszi tér is teljes.  $\square$

Könnyen látható, hogy az alfejezet összes eddigi állítása az  $\mathbb{R}^p$  euklideszi térre is átvihető:

**Állítás 177** *Egy  $(\alpha_n^1, \alpha_n^2, \dots, \alpha_n^p)$ ,  $n = 1, 2, \dots$  vektorsorozat pontosan akkor konvergál egy  $(\alpha^1, \alpha^2, \dots, \alpha^p)$  vektorhoz az  $(\mathbb{R}^p, d_e)$  euklideszi térben, ha koordinátánként konvergál, azaz*

$$\lim_{n \rightarrow +\infty} \alpha_n^i = \alpha^i \text{ minden } i = 1, 2, \dots, p \text{ mellett.}$$

**Állítás 178** *Az  $a_n = (\alpha_n^1, \alpha_n^2, \dots, \alpha_n^p)$  illetve  $b_n = (\beta_n^1, \beta_n^2, \dots, \beta_n^p)$   $\mathbb{R}^p$ -beli vektorsorozatok tartanak az*

$$a = (\alpha^1, \alpha^2, \dots, \alpha^p) \text{ illetve } b = (\beta^1, \beta^2, \dots, \beta^p)$$

*vektorokhoz. Legyen továbbá gamma tetszőleges valós szám. Ekkor*

$$(i) \lim_{n \rightarrow \infty} (a_n + b_n) = a + b.$$

$$(ii) \lim_{n \rightarrow \infty} \gamma a_n = \gamma a.$$

**Állítás 179** Ha  $\mathbb{R}^p$ -beli vektorok egy sorozata korlátos, akkor van konvergens rész-sorozata.

**Állítás 180** Az  $\mathbb{R}^p$  euklideszi térben minden Cauchy-sorozat konvergens.

Mivel a komplex számok távolsága azonos az  $\mathbb{R}^2$ -beli euklideszi távolsággal, azért a fenti állítások a komplex számtest esetére is megfogalmazhatók.

**Állítás 181** Komplex számok egy  $z_n = x_n + y_n i$  sorozata pontosan akkor konvergál a  $w = x + y i$  komplex számhoz, ha a valós részek sorozata konvergál az  $x$ -hez, az imaginárius részek sorozata pedig az  $y$ -hoz, azaz ha

$$\lim_{n \rightarrow +\infty} x_n = x \quad \text{és} \quad \lim_{n \rightarrow +\infty} y_n = y.$$

A komplex számok test struktúra, ezért a következő állítás formálisan ugyanolyan, mint a 171. Állítás, csak a jelek tartalma más, például az “ $|\cdot|$ ” most a komplex számok abszolút értékét jelöli:

**Állítás 182** Tartsanak az  $z_n = x_n + y_n i$  illetve  $w_n = u_n + v_n i$  komplex sorozatok a  $z = x + y i$  illetve  $w = u + v i$  komplex számokhoz.

(a) Az  $(z_n + w_n)$  összeg-sorozat határértéke:

$$\lim_{n \rightarrow +\infty} (z_n + w_n) = \lim_{n \rightarrow +\infty} z_n + \lim_{n \rightarrow +\infty} w_n.$$

(b) Ha a  $c$  tetszőleges komplex szám, akkor a  $cz_n$  számmal szorzott sorozat határértéke:

$$\lim_{n \rightarrow +\infty} (cz_n) = c \lim_{n \rightarrow +\infty} z_n.$$

(c) Az  $(z_n w_n)$  szorzat-sorozat határértéke:

$$\lim_{n \rightarrow +\infty} z_n w_n = \lim_{n \rightarrow +\infty} z_n \cdot \lim_{n \rightarrow +\infty} w_n.$$

(d) Ha  $w \neq 0$ , akkor az  $z_n/w_n$  tört nevezője egy indextől kezdve nem nulla, és a hányados-sorozat határértéke:

$$\lim_{n \rightarrow +\infty} \frac{z_n}{w_n} = \frac{\lim_{n \rightarrow +\infty} z_n}{\lim_{n \rightarrow +\infty} w_n}.$$

Bizonyítás. Az állítások, az előző tételhez hasonlóan a valós és képzetes részek konvergenciájára vezethetők vissza a 181. és 171. állítások felhasználásával. Példaképpen lássuk, mondjuk a szorzat-sorozatra vonatkozó, (c) állítást (aminek a (b) nyilvánvalóan speciális esete).

A 181. állítás szerint a valós és képzetes részek  $x_n$ ,  $y_n$ ,  $u_n$  és  $v_n$  sorozatai rendre tartanak az  $x$ ,  $y$ ,  $u$  és  $v$  valós számokhoz. A 171. állítás alapján a

$$z_n w_n = (x_n u_n - y_n v_n) + (x_n v_n + y_n u_n)i$$

szorzat-sorozatnak a valós illetve képzetes részei tartanak az

$$(xu - yv), \quad \text{illetve} \quad (xv + yu)$$

valós számokhoz, amelyek éppen a  $zw$  valós illetve képzetes részei.  $\square$

A Bolzano-Weierstrass-féle kiválasztási tétel igaz a  $\mathbb{C}$  metrikus térben is:

**Állítás 183** *Ha  $\mathbb{R}^2$ -beli vektorok egy sorozata korlátos, akkor van konvergens rész-sorozata.*

Ugyanígy a  $\mathbb{C}$  teljessége:

**Állítás 184** *A komplex számok metrikus terében minden Cauchy-sorozat konvergens.*

## 4.3 Numerikus sorok

Ebben a pontban az összeadás műveletének végtelen sok tagra való kiterjesztésével fogunk foglalkozni. Az első alpont tartalmazza az alapfogalmakat és létezési (konvergencia) kritériumokat, a második alpontban pedig a sorok közötti műveleteket tárgyaljuk.

### 4.3.1 Alapfogalmak

Definícióinkat komplex számokra fogalmazzuk meg. Ha csak valós sorozatokat vizsgálunk, azt külön jelezni fogjuk.

**Definíció 185** *Legyen az  $(a_k)$  egy komplex sorozat. A szimbolikusan felírt*

$$a_1 + a_2 + a_3 + \cdots + a_k + \cdots,$$

*végtelen összeget sornak nevezzük, aminek a tagjai a sorozat elemei. Tömörebb felírási módok:*

$$\sum_{k=1}^{+\infty} a_k,$$

vagy, ha az összegezés határai magától értetődőek:

$$\sum_k a_k \quad \text{vagy} \quad \sum a_k.$$

A végtelen összegből képzett

$$s_n \stackrel{\text{def}}{=} a_1 + a_2 + \cdots + a_n = \sum_{k=1}^n a_k$$

véges összeget a sor  $n$ -edik részletösszegének (szeletének) mondjuk.

A definícióban felírt végtelen összegeket az 1 helyett bármilyen indextől indíthatnánk, például:

$$a_0 + a_1 + a_2 + \cdots + a_k + \cdots = \sum_{k=0}^{+\infty} a_k$$

vagy

$$a_n + a_{n+1} + \cdots = \sum_{k=n}^{+\infty} a_k.$$

Azzal, hogy a “szimbolikusan felírt” szavakat használtuk, azt akartuk kifejezni, hogy most még csak pusztán formális jelölésről van szó, hiszen eddigi ismereteink szerint végtelen tagú összeget nem tudunk kiszámolni.

**Definíció 186** Egy  $\sum_{k=1}^{+\infty} a_k$  sort konvergensnek mondunk, ha a részletösszegek  $s_n = a_1 + \cdots + a_n$  sorozata konvergens, és ezt a határértéket fogjuk a sor összegének nevezni, azaz formálisan:

$$\sum_{k=1}^{+\infty} a_k = \lim_{n \rightarrow +\infty} s_n.$$

A nem konvergens sort divergensnek nevezzük.

Más szavakkal: Egy konvergens sor összege a részletösszegek sorozatának a limesze.

Mivel a sor részletösszegeinek a képzési szabálya:

$$s_n = s_{n-1} + a_n, \tag{4.1}$$

ezért azt mondhatjuk, hogy a sor összegének a vizsgálata egy speciális típusú képzési szabállyal megadott sorozat konvergenciájának a vizsgálata. Ennek megfelelően minden, a sorozatoknál megismert állításból egy sorokra vonatkozó tételt fogunk tudni megfogalmazni, de a sorok új természetű problémákat is fel fognak vetni.

Meg kell jegyeznünk, hogy — az előző bekezdésben mondottak megfordításaként — *a sorozatok vizsgálata is visszavezethető a sorok vizsgálatára*. A (4.1) képzési szabályból ugyanis  $a_n = s_n - s_{n-1}$ , és így

$$a_n = a_1 + (s_2 - s_1) + \cdots + (s_n - s_{n-1}),$$

Eszerint az  $a_n, n = 1, 2, \dots$  sorozat az

$$a_1, (s_2 - s_1), (s_3 - s_2), \dots, (s_n - s_{n-1}), \dots$$

sorozat által meghatározott sor részletösszegeinek a sorozata. Ennek a segítségével minden sorra vonatkozó tételből meg tudunk fogalmazni egy sorozatra vonatkozó tételt. Ezek szerint: *a sorozatok és sorok vizsgálata teljesen párhuzamosan végezhető*.

Mielőtt a sorozatokra vonatkozó legfontosabb tételeket átfogalmaznánk sorokra, lássunk egy példát.

**Példa 4.3** *Melyik sor konvergens az alábbiak közül*

(i) *Legyen  $c$  egy valós szám, és a sorozat:*

$$c + c + \cdots + c + \cdots = \sum_{k=1}^{+\infty} c.$$

(ii)  $\sum_{k=1}^{+\infty} (-1)^{k+1} = 1 - 1 + 1 - 1 + \cdots + (-1)^{n+1} + \cdots$ .

Az (i) sor részletösszegei:  $s_n = cn$ , ami pontosan akkor konvergál, ha  $c = 0$ , és ekkor az összege: 0. Szavakban: az állandó tagokból álló sor csak akkor konvergál, ha a tagjai nullák, és ekkor az összeg is nulla.

A (ii) sor részletösszegeinek a sorozata:  $1, 0, 1, 0, 1, 0, \dots$ , ami nyilván divergens, ezért a sor is divergens.  $\square$

**Példa 4.4** *Határozzuk meg a*

$$\sum_{k=2}^{+\infty} \frac{1}{k(k-1)}$$

*sor összegét.*

Tekintsük a sor  $n$ -ik részletösszegét. Mivel

$$\frac{1}{k(k-1)} = \frac{1}{k-1} - \frac{1}{k},$$

azért minden  $n \geq 2$  esetén

$$s_n = \sum_{k=2}^n \frac{1}{k(k-1)} = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) = 1 - \frac{1}{n}.$$

Tehát  $s_n \rightarrow 1$ , azaz a sor konvergens, és az összege 1.

A sorozatoknál egy általános — komplex sorozatokra is érvényes — konvergencia kritériumot ismertünk meg, a Cauchy-kritériumot. Ennek a sorokra vonatkozó átfogalmazása:

**Állítás 187 (Cauchy-kritérium sorokra)** Egy  $\sum_{k=1}^{+\infty} a_k$  sor pontosan akkor konvergens, ha minden  $\epsilon > 0$  számhoz van olyan  $N$  index, hogy

$$\left| \sum_{k=m+1}^n a_k \right| \leq \epsilon \quad \text{ha} \quad N \leq m < n.$$

**Bizonyítás.** A részletösszegek sorozata a Cauchy-kritérium szerint pontosan akkor konvergál, ha minden  $\epsilon > 0$  számhoz van olyan  $N$  index, hogy

$$|s_n - s_m| = \left| \sum_{k=m+1}^n a_k \right| \leq \epsilon, \quad \text{ha} \quad N \leq m < n.$$

□

A Cauchy kritérium az  $m+1 = n$  esetben speciálisan azt adja, hogy

$$|a_n| \leq \epsilon \quad \text{ha} \quad N < n,$$

ami azt jelenti, hogy az  $a_n$  sorozat a nullához tart, ha a sor konvergens. Ezt a hangsúlyozás kedvéért állításban is megfogalmazzuk:

**Állítás 188** Ha egy sor konvergens, akkor a tagjai nullához tartanak.

Az állítás megfordítása nem igaz, ahogyan a következő példa is mutatja.

**Példa 4.5** A

$$\sum_{k=1}^{+\infty} \frac{1}{k}$$

harmonikusnak nevezett sor divergens, részletösszegeinek a sorozata a plusz végtelelnhez tart.

Vegyük észre, hogy az  $(n+1)$  és  $2n$  közötti tagok összege nagyobb, mint  $\frac{1}{2}$ :

$$\frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} > \overbrace{\frac{1}{2n} + \cdots + \frac{1}{2n}}^{n\text{-szer}} = \frac{1}{2}.$$

Ebből nyilvánvaló, hogy

$$\left| \sum_{k=n}^m \frac{1}{k} \right| > \frac{1}{2}, \quad \text{ha} \quad 2n < m,$$

ezért a Cauchy-kritérium az  $\epsilon = \frac{1}{2}$  (vagy ennél kisebb) számra nem teljesülhet, tehát a sor divergens.

Mivel a sor részletösszegei monoton nőnek, ezért csak úgy divergálhatnak, hogy korlátlanok, tehát tartanak a plusz végtelenhez.  $\square$

A szóbanforgó sor divergenciája (az összeg "végtelen értéke") meglepő, mert szemléletesen ezt mondja: ha úgy haladunk, hogy először egy lépést teszünk meg, a második esetben egy fél lépést, a harmadik esetben egy harmad lépést, a milliódik esetben pedig egy milliomod lépést és így tovább, akkor ezen a módon bármilyen messze eljuthatunk.

Egy új konvergencia fogalom a sorokra:

**Definíció 189** Egy  $\sum_{k=1}^{+\infty} a_k$  sort abszolút konvergensnek mondunk, amennyiben az  $\sum_{k=1}^{+\infty} |a_k|$  sor konvergens.

Az abszolút konvergencia erősebb, mint a közönséges konvergencia.

**Állítás 190** Ha egy sor abszolút konvergens, akkor konvergens is.

Bizonyítás. A sorokra vonatkozó Cauchy-kritérium szerint egy  $\sum a_k$  sor pontosan akkor konvergens, ha minden  $\epsilon > 0$  számhoz van olyan  $N$  index, hogy

$$\left| \sum_{k=m+1}^n a_k \right| \leq \epsilon, \quad \text{ha } N \leq m < n,$$

és akkor abszolút konvergens, ha

$$\sum_{k=m+1}^n |a_k| \leq \epsilon, \quad \text{ha } N \leq m < n.$$

Az utóbbi sorból következik a megelőző kiemelt sor, mivel az abszolút érték háromszög-egyenlőtlenségéből:

$$\left| \sum_{k=m+1}^n a_k \right| \leq \sum_{k=m+1}^n |a_k|.$$

$\square$

Állításunk megfordítása azonban nem érvényes. Mielőtt erre példát mutatnánk ismerkedjünk meg egy gyakran jól használható állítással.

**Állítás 191 (Lebnitz-típusú sor)** Legyen az  $a_k$  egy monoton fogyó sorozat, amelyre  $a_k \rightarrow 0$ . Ekkor a

$$a_1 - a_2 + a_3 - \dots + (-1)^{n+1} a_n + \dots = \sum_{n=1}^{+\infty} (-1)^{n+1} a_n$$

váltakozó előjelű sor konvergens. Az ilyen sort Leibnitz-típusúnak is nevezzük.

Bizonyítás. Legyen adott egy  $\epsilon > 0$  szám. Az  $N$  indexet megválaszthatjuk úgy, hogy  $a_n \leq \epsilon$ , ha  $N \leq n$ . Ebből azt kapjuk, hogy

$$|a_n - a_{n+1} + a_{n+2} - \cdots \pm a_m| \leq |a_n| \leq \epsilon,$$

ha  $n \leq N$ , ezért a Cauchy-kritérium miatt a sor konvergens.

Más bizonyítást adhat valaki abból az észrevételből kiindulva, hogy a páros illetve páratlan indexű részletösszegek sorozata (ellenkező irányban) monoton.  $\square$

**Példa 4.6** *Mutassuk meg, hogy az*

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots + (-1)^{n+1} \frac{1}{n} + \cdots$$

*sor konvergens, de nem abszolút konvergens.*

Valóban, a sor nyilvánvalóan Leibnitz-típusú, ezért konvergens, de nem abszolút konvergens, mivel a harmonikus sor divergens.  $\square$

### 4.3.2 Konvergenca kritériumok

A monoton sorozatokra vonatkozó konvergenca kritériumból azonnal adódik a következő tétel. (Ha monotonitásról, nemnegativitásról, stb. beszélünk, akkor valós sorozatokra gondolunk, hiszen a komplex számok teste nem rendezett.)

**Állítás 192 (Nemnegatív tagú sorok konvergenciája)** *A  $\sum_k a_k$  nemnegatív tagú sor pontosan akkor konvergens, ha a részletösszegek sorozata felülről korlátos.*

Bizonyítás. A (4.1) képzési szabály szerint egy sor részletösszegei pontosan akkor monoton növekedőek, ha a sor nemnegatív tagú, és így a monoton növekedő sorozatokra vonatkozó konvergenca kritérium azonnal adja az igazolást.  $\square$

A következő tétel arra ad lehetőséget, hogy konkrét sorok konvergenciáját ismerve más sorok konvergenciájára következtethessünk.

**Állítás 193 (Összehasonlító kritérium)** *Tegyük fel, hogy a  $\sum c_k$  valós sor konvergens, a  $\sum d_k$  nemnegatív tagú valós sor pedig divergens.*

- (1) *Ha egy  $a_k$  (komplex) sorozatra  $|a_k| \leq c_k$  egy bizonyos indextől kezdve, akkor a  $\sum a_k$  sor (abszolút) konvergens.*
- (2) *Ha egy  $b_k$  valós sorozatra egy bizonyos indextől kezdve  $b_k \geq d_k$ , akkor a  $\sum b_k$  sor divergens.*

**Bizonyítás.** (1): A  $\sum c_k$  konvergenciája miatt, a Cauchy-kritérium szerint adott  $\epsilon > 0$  számhoz van olyan  $N$  index, hogy

$$\left| \sum_{k=n}^m c_k \right| \leq \epsilon, \quad \text{ha} \quad N \leq n \leq m.$$

Emiatt a feltevés alapján

$$\left| \sum_{k=n}^m a_k \right| \leq \sum_{k=n}^m |a_k| \leq \sum_{k=n}^m c_k \leq \epsilon,$$

ha  $N \leq n \leq m$ , ami a Cauchy-kritérium szerint biztosítja az  $\sum a_k$  sor konvergenciáját.

(2): Ha az  $\sum b_k$  sor konvergálna, akkor a belátott (1) állítás alapján a  $\sum d_k$  sor is konvergálna, ellentétben a feltevéssel.  $\square$

**Példa 4.7** *Igazoljuk, hogy a*

$$\sum_{k=1}^{+\infty} \frac{1}{k^2}$$

*sor konvergens.*

A  $k = 2$  indextől kezdve

$$\frac{1}{k^2} < \frac{1}{k(k-1)}.$$

Másrészt a 4.4 Példa szerint a  $\sum_{k=2}^{+\infty} 1/k(k-1)$  sor konvergens, ezért az Összehasonlító kritérium szerint a  $\sum_{k=1}^{+\infty} 1/k^2$  sor is konvergens.

Ebből a megfontolásból az is látható, hogy

$$\sum_{k=1}^{+\infty} \frac{1}{k^2} < 2.$$

A pontos érték meghatározása (amely egyébként  $\pi^2/6$ ) már jóval komplikáltabb feladat, és ezzel itt nem foglalkozunk.

Fontos szerepet játszik a későbbiekben a mértani sor.

**Állítás 194** *A*

$$\sum_{k=0}^{+\infty} x^k$$

*mértani sor pontosan akkor konvergens, ha  $|x| < 1$ , és ekkor az összege:*

$$\sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}.$$

Bizonyítás. Az  $s_n$  részletösszeg:

$$s_n = 1 + x + x^2 + \cdots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

Ha  $|x| < 1$ , akkor  $\lim_{n \rightarrow +\infty} x^{n+1} = 0$ , ezért

$$\sum_{k=0}^{+\infty} x^k = \lim_{n \rightarrow +\infty} s_n = \frac{1}{1 - x}.$$

Ha  $|x| > 1$ , akkor  $\lim_{n \rightarrow +\infty} |x|^{n+1} = +\infty$ , ezért az  $|s_n|$  sorozat korlátlan, és így nem lehet konvergens.

Ha  $|x| = 1$ , akkor két esetet kell megkülönböztetni:

- 1)  $x = 1$ , ekkor közvetlen összegezés alapján (az összeg-képlet a nevező miatt nem alkalmazható):  $s_n = n + 1$ , ezért ekkor nyilván divergens a sor.
- 2) Valamilyen  $\phi \neq 0$  számra

$$x = \cos \phi + i \sin \phi$$

és így az összegképletben szereplő  $x^{n+1}$  hatványra

$$x^{n+1} = \cos(n+1)\phi + i \sin(n+1)\phi.$$

Ez a kifejezés pedig nem tart semmihez, hiszen a komplex számok szorzása szerint  $\phi$  szöggel ugrik az egységkörön, ha az  $n$  eggyel nő.  $\square$

**Példa 4.8** Legyen az  $\alpha$  és  $\beta$  két tetszőleges pozitív szám. Konvergens-e a

$$\sum_{k=1}^{+\infty} \left( \frac{\alpha\beta}{\alpha^2 + \beta^2} \right)^k \cos k\alpha$$

sor?

A számtani és mértani közép közötti egyenlőtlenségből:

$$\frac{\alpha^2 + \beta^2}{2} \geq \alpha\beta, \quad \text{ezért} \quad \frac{\alpha\beta}{\alpha^2 + \beta^2} \leq \frac{1}{2},$$

és így a

$$\left| \left( \frac{\alpha\beta}{\alpha^2 + \beta^2} \right)^k \cos k\alpha \right| \leq \left( \frac{1}{2} \right)^k,$$

egyenlőtlenség miatt az összehasonlító kritérium és az  $\frac{1}{2}$  hányadosú mértani sor konvergenciája igazolja a konvergenciát.  $\square$

A pozitív tagú mértani sor konvergenciájára és az összehasonlító kritériumra alapozva két kritériumot mondunk ki, amelyek gyorsabbá teszik az előző példában használt módszer alkalmazását.

**Állítás 195 (Gyökkritérium)** *Ha egy  $a_n$  sorozatra*

(1)  $\limsup_{k \rightarrow +\infty} \sqrt[k]{|a_k|} < 1$ , *akkor a  $\sum a_k$  sor konvergens;*

(2)  $\limsup_{k \rightarrow +\infty} \sqrt[k]{|a_k|} > 1$ , *akkor a  $\sum a_k$  sor divergens.*

A  $\limsup_{k \rightarrow +\infty} \sqrt[k]{|a_k|} = 1$  eset nem szerepel az állításban, mert ekkor semmit sem tudunk mondani.

**Bizonyítás.** (1): Legyen  $\alpha \stackrel{\text{def}}{=} \limsup_{k \rightarrow +\infty} \sqrt[k]{|a_k|}$ . Ha  $\alpha < 1$  és a  $\delta$  számot úgy választjuk meg, hogy  $\alpha < \alpha + \delta < 1$ , akkor a limesz superior tulajdonsága miatt

$$\sqrt[k]{|a_k|} < \alpha + \delta, \quad \text{azaz} \quad |a_k| < (\alpha + \delta)^n,$$

minden eléggé nagy  $n$ -re és így az összehasonlító kritérium és a mértani sor konvergencia tétele alapján adódik a  $\sum a_k$  konvergenciája.

(2): Legyen ismét  $\alpha \stackrel{\text{def}}{=} \limsup_{k \rightarrow +\infty} \sqrt[k]{|a_k|}$ . Mivel most  $\alpha > 1$ , azért végtelen sok  $k$  indexre  $\sqrt[k]{|a_k|} > 1$ , azaz  $|a_k| > 1$ , így a sor tagjai nem tarthatnak a nullához, ezért a sor nem is lehet konvergens.  $\square$

**Állítás 196 (Hányados kritérium)** *Ha egy  $a_k$  sorozatra*

1.  $\limsup_{k \rightarrow +\infty} \left| \frac{a_{k+1}}{a_k} \right| < 1$ , *akkor a  $\sum a_k$  sor konvergens;*

2. *ha valamilyen  $k_0$ -nál nagyobb minden  $k$ -ra*

$$\left| \frac{a_{k+1}}{a_k} \right| \geq 1,$$

*akkor a sor divergens. Ezt biztosítja például:*

$$\lim_{k \rightarrow +\infty} \left| \frac{a_{k+1}}{a_k} \right| > 1.$$

A

$$\lim_{k \rightarrow +\infty} \left| \frac{a_{k+1}}{a_k} \right| = 1$$

eset nem szerepel az állításban, mert ekkor semmit sem tudunk mondani.

**Bizonyítás.** (1): Legyen  $\alpha \stackrel{\text{def}}{=} \limsup_{k \rightarrow +\infty} \left| \frac{a_{k+1}}{a_k} \right| < 1$ . Ha a  $\delta$  számot úgy választjuk meg, hogy  $\alpha < \alpha + \delta < 1$ , akkor a limesz superior tulajdonsága miatt

$$\left| \frac{a_{k+1}}{a_k} \right| < \alpha + \delta, \quad \text{azaz} \quad |a_{k+1}| < (\alpha + \delta)|a_k|.$$

minden  $k \geq N$ -re. Ebből a rekurzív egyenlőtlenségből azt kapjuk, hogy

$$|a_{k+1}| \leq (\alpha + \delta)^{k+1-N} |a_N|.$$

Eszerint az összehasonlító kritérium és a mértani sor konvergenciája alapján adódik az állítás.

(2): Legyen most  $\left| \frac{a_{k+1}}{a_k} \right| \geq 1$ , ha  $k_0 \leq k$ . Ekkor

$$|a_{k+1}| \geq a_k,$$

minden eléggé nagy  $k$ -ra, és így a  $\sum a_k$  sor tagjai nem tartanak a nullához, ezért nem lehet konvergens.  $\square$

**Példa 4.9** Vizsgáljuk meg a következő sor konvergenciáját a gyök- és hányadoskritériummal

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{2^n} + \frac{1}{3^n} + \cdots.$$

A gyökkritériumhoz:

$$\limsup_{k \rightarrow +\infty} \sqrt[k]{a_k} = \lim_{k \rightarrow +\infty} \sqrt[2k]{\frac{1}{2^k}} = \frac{1}{\sqrt{2}},$$

ezért a sor konvergens.

A hányadoskritériumhoz: A limesz szuperior és inferior:

$$\begin{aligned} \limsup_{k \rightarrow +\infty} \frac{a_{k+1}}{a_k} &= \lim_{k \rightarrow +\infty} \left( \frac{3}{2} \right)^k = +\infty. \\ \liminf_{k \rightarrow +\infty} \frac{a_{k+1}}{a_k} &= \lim_{k \rightarrow +\infty} \left( \frac{2}{3} \right)^k = 0. \end{aligned}$$

Ez azt mutatja, hogy erre a sorra a gyökkritérium alkalmazható, de a hányadoskritérium nem.  $\square$

## 4.4 Sorozatok és függvények határértéke

A folytonosság fogalma a sorozatok fogalmán keresztül is bevezethető:

**Állítás 197** Legyenek az  $(X, d_X)$  és  $(Y, d_Y)$  metrikus-terek. Egy  $f : X \rightarrow Y$  leképezés pontosan akkor folytonos egy  $x$  pontban, ha tetszőleges  $x_n \rightarrow x$   $X$ -beli sorozatra

$$\lim_{n \rightarrow +\infty} f(x_n) = f(x).$$

**Bizonyítás.** A feltétel szükségességének az igazolásához, legyen az  $f$  folytonos az  $x$  pontban, és az  $x_n$  egy tetszőleges  $x$ -hez tartó sorozat.

Az  $f$  folytonossága miatt, adott  $\epsilon > 0$  számhoz van olyan  $\delta > 0$  szám, hogy

$$d_Y(f(x), f(u)) < \epsilon, \quad \text{ha} \quad d_X(x, u) < \delta.$$

Az  $x_n$   $x$ -hez való tartása miatt a  $\delta > 0$  számhoz van olyan  $n_0$  index, hogy

$$d(x, x_n) < \delta, \quad \text{ha} \quad n > n_0.$$

Így

$$d_Y(f(x), f(x_n)) < \epsilon, \quad \text{ha} \quad n > n_0,$$

tehát  $\lim_{n \rightarrow +\infty} f(x_n) = f(x)$ .

A feltétel elégségességének az igazolásához megmutatjuk, hogy ha az  $f$  nem folytonos az  $x$  pontban, akkor van olyan sorozat, amire nem teljesül a feltétel.

Ha az  $f$  nem folytonos az  $x$ -ben, akkor van olyan  $\epsilon > 0$  szám, ami mellett akármilyen (kicsi)  $\delta > 0$  számhoz van olyan  $x_\delta \in X$  pont, hogy

$$d_Y(f(x), f(x_\delta)) \geq \epsilon \quad \text{és} \quad d_X(x, x_\delta) < \delta.$$

Vegyük a  $\delta = \frac{1}{n}$ , ( $n = 1, 2, \dots$ ) sorozatot. Az ehhez létező

$$x_{\frac{1}{n}}, \quad n = 1, 2, \dots$$

sorozat nem teljesíti a tétel feltételét. Egyrészt ugyanis az  $x_{\frac{1}{n}}$  tart az  $x$ -hez:

$$d_X(x, x_{\frac{1}{n}}) \leq \frac{1}{n}, \quad \text{tehát} \quad \lim_{n \rightarrow +\infty} x_{\frac{1}{n}} = x;$$

másrészt az  $f\left(x_{\frac{1}{n}}\right)$  nem tart az  $f(x)$ -hez:

$$d_Y\left(f(x), f\left(x_{\frac{1}{n}}\right)\right) \geq \epsilon.$$

□

Mivel a tétel szükséges és elégséges feltételt fogalmaz meg a metrikus téren való folytonossághoz, ezért a folytonosság definíciójaként is használható.

Sorozatok konvergenciájára már megismertük a Cauchy-féle konvergencia kritériumot. Most ennek a függvényekre vonatkozó megfelelőjét bizonyítjuk be.

**Állítás 198 (Cauchy kritérium)** *Ha egy  $f$  valós változós és valós értékű függvény definiált a  $b$  pont egy hiányos környezetében, és minden  $\epsilon$  pozitív számhoz van olyan  $\delta$  pozitív szám, hogy*

$$|f(x) - f(y)| \leq \epsilon \quad \text{ha} \quad 0 < |x - b| < \delta \quad \text{és} \quad 0 < |y - b| < \delta, \quad (4.2)$$

*akkor az  $f$  függvénynek a  $b$  pontban van határértéke.*

Bizonyítás. A bizonyítás két lépésből áll.

*Első lépés:* Vesszünk egy tetszőleges olyan  $x_n, n = 1, 2, \dots$  sorozatot, amely benne van az  $f$  értelmezési tartományában, és tart a  $b$  ponthoz. Megmutatjuk, hogy ekkor az  $f(x_n), n = 1, 2, \dots$  sorozat tart valamilyen  $z$  számhoz.

A tétel feltétele szerint, adott  $\epsilon$  pozitív számhoz van olyan  $\delta$ , hogy a (4.2) teljesül. Mivel az  $x_n$  sorozat tart a  $b$ -hez, ezért van olyan  $n_0$  index, hogy

$$|x_n - b| < \delta, \quad \text{ha} \quad n_0 < n,$$

ezért (4.2) alapján:

$$|f(x_n) - f(x_m)| < \epsilon, \quad \text{ha} \quad n_0 < n, m.$$

Eszerint az  $f(x_n)$  számsorozat Cauchy-sorozat, és a számsorozatokra vonatkozó Cauchy-kritérium alapján, létezik limesze, amit  $z$ -vel jelölünk.

*Második lépés:* Most pedig megmutatjuk, hogy az első lépésben kapott  $z$  szám limesze az  $f$  függvénynek a  $b$  helyen.

Legyen adott egy tetszőleges  $\epsilon$  pozitív szám. A tétel feltételét az  $\epsilon/2$  számra alkalmazva:

$$|f(x) - f(y)| < \frac{\epsilon}{2} \quad \text{ha} \quad 0 < |x - b| < \delta \quad \text{és} \quad 0 < |y - b| < \delta. \quad (4.3)$$

Mivel  $\lim x_n = b$ , ezért van olyan  $n_1$ , hogy

$$|x_n - b| < \delta, \quad \text{ha} \quad n_1 < n,$$

ezért a (4.3) alapján:

$$|f(x) - f(x_n)| < \frac{\epsilon}{2}, \quad \text{ha} \quad 0 < |x - b| < \delta \quad \text{és} \quad n_1 < n. \quad (4.4)$$

A  $\lim f(x_n) = z$  miatt, van olyan  $n_2$ , hogy

$$|f(x_n) - z| < \frac{\epsilon}{2}, \quad \text{ha} \quad n_2 < n.$$

Ennek és a (4.4)-nek a felhasználásával:

$$|f(x) - z| \leq |f(x) - f(x_n)| + |f(x_n) - z| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

ha  $0 < |x - b| < \delta$  és  $\max\{n_1, n_2\} < n$ . Ez viszont azt jelenti, hogy

$$|f(x) - z| < \epsilon, \quad \text{ha} \quad 0 < |x - b| < \delta,$$

tehát

$$\lim_{x \rightarrow \beta} f(x) = z$$

ahogyan állítottuk.  $\square$

A kritériumra akkor lesz majd szükségünk, amikor be kell látnunk a határérték létezését, de magát a határértéket nem tudjuk meghatározni.

Most a Cauchy-kritérium megfelelőjét bizonyítjuk be a végtelen helyen vett véges határértékre.

**Állítás 199 (Cauchy kritérium)** Tegyük fel, hogy egy  $f$  valós függvény értelmezési tartománya, valamilyen  $\alpha$  szám mellett, tartalmazza az

$$(\alpha, +\infty) = \{x : \alpha < x\}$$

korlátlan intervallumot. Ha tetszőleges  $\epsilon$  pozitív számhoz van olyan  $p$  szám, hogy

$$|f(x_1) - f(x_2)| \leq \epsilon, \quad \text{ha} \quad p \leq x_1, x_2,$$

akkor az  $f$  függvénynek a plusz végtelenben van véges határértéke.

**Bizonyítás.** A bizonyítást a 198. véges pontban vett Cauchy kritériumra fogjuk visszavezetni.

Azt az alapvető észrevételt fogjuk felhasználni, hogy a  $+\infty$ -ben vett határérték a nullában vett jobboldali határértékre vezethető vissza a következők szerint. Ha az  $x$  az  $(\alpha, +\infty)$  intervallumban van, akkor az  $1/x$  a nullának a  $(0, 1/\alpha)$  jobboldali hiányos környezetében helyezkedik el, és megfordítva. Emiatt azt állíthatjuk, hogy

$$\lim_{x \rightarrow +\infty} f(x) = \lim_{u \searrow 0} f\left(\frac{1}{u}\right).$$

A jobboldali (és így a baloldali) limesz létezésére a 198. Állítás szerint elégséges feltétel: Minden  $\epsilon$  pozitív számhoz van olyan  $\delta$  pozitív szám, hogy

$$\left| f\left(\frac{1}{u_1}\right) - f\left(\frac{1}{u_2}\right) \right| \leq \epsilon, \quad \text{ha} \quad 0 < u_1, u_2 \leq \delta.$$

Ez viszont azzal ekvivalens, hogy minden  $\epsilon$  pozitív számhoz van olyan  $\beta$  szám, hogy

$$|f(x_1) - f(x_2)| \leq \epsilon, \quad \text{ha} \quad \beta \leq x_1, x_2,$$

( $\beta = 1/\delta$ ). Ezzel viszont éppen a most bizonyítandó tétel feltételét írtuk fel.  $\square$



## 5.

# Differenciálszámítás

### 5.1 Definíciók, értelmezések

Most elérkeztünk ahhoz a ponthoz, hogy az analízis egyik legfontosabb fogalmi apparátusát felépíthessük. Az egész matematikai analízisnek a tárgya — némi túlzással szólva — a függvény fogalmának széleskörű tanulmányozása. Azt mondhatnánk, hogy az eddig bevezetett legalapvetőbb fogalmak, a folytonosság és a határérték, a most sorra kerülő differenciálhányados- (derivált-) fogalom bevezetését készítették elő.

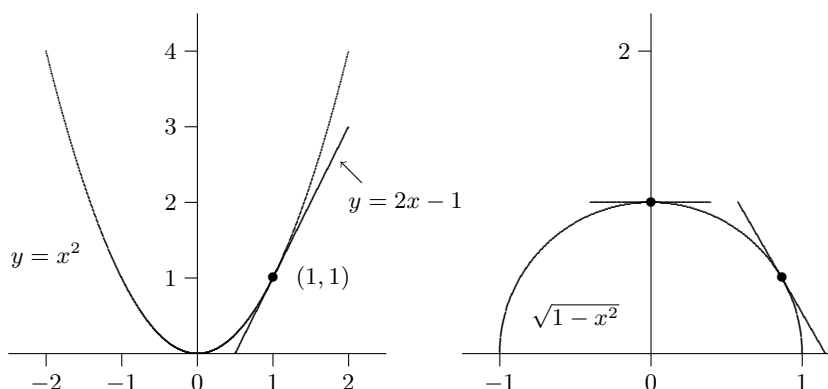
#### 5.1.1 Bevezetés, definíciók

A pontos definíció magadása előtt egy olyan intuitív megközelítéssel kezdjük, amelyik jó és szemléletes háttérrel ad a differenciálhányados (derivált) fogalmának a bevezetéséhez.

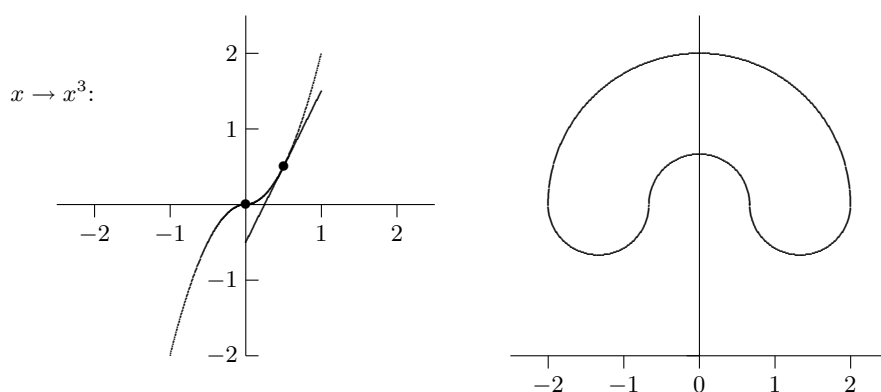
Már az antik görög korban foglalkoztatta a geometria művelőit az, hogy mit értsenek egy görbe “érintőjén”. A probléma megközelítéséhez vegyünk néhány példát. Az 5.1. és 5.2. ábrán négy görbét rajzoltunk fel az  $\mathbb{R}^2$  síkon.

A felrajzott görbék közül az első három közismert, egyszerű formulákkal megadott függvények gráfja. A kör esetében, annak az érdekében, hogy egy függvény gráfját kapjuk, csak a felső félkört rajzoltuk fel. Az 5.2. ábra jobbfelén olyan görbe, mondhatni: “kifli” van, ami nem illik bele jelen vizsgálódásainkba, mert nem egy függvény gráfja, hiszen a rajz szerint egy értékhez több függvényérték is tartozik. A jelenlegi anyagban az ilyen görbéket nem fogjuk vizsgálni, pedig a szemléletből kiindulva jogosan úgy érezhetjük, hogy az ilyen alakzatú görbéknek is lehetnek “érintőik”. Ezek tanulmányozása azonban jelenleg nehézséget okozna, és eltérítene bennünket tanulmányaink fő irányától. Csak az olyan görbékhez keressük az érintőfogalmat, amelyek függvények gráfjai.

A kör esetében elemi geometriai megfontolásokból ismerjük, hogy az érintő merőleges az érintési ponthoz húzott sugárra, ezért könnyen megszerkeszthető.



5.1. ábra: Kör és parabola.

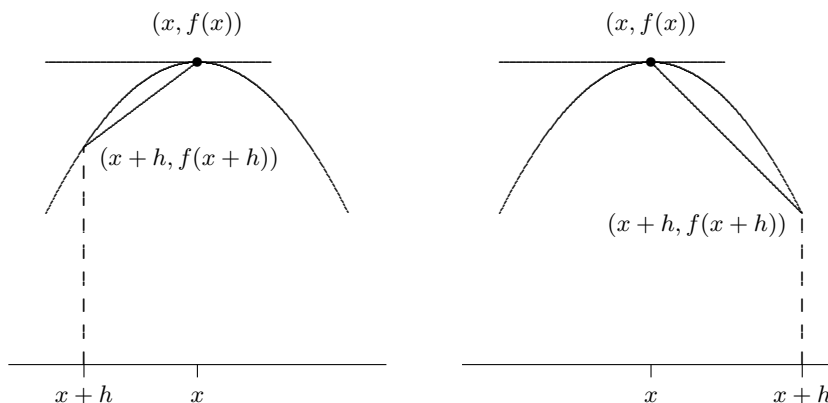
5.2. ábra: Az  $x \mapsto x^3$  függvény és egy "kifli".

Voltaképpen azonban ezt a jólismert állítást is csak a következőkben — az érintő fogalmának az egzakt megragadása, definiálása után — fogjuk majd pontosan igazolni.

Az  $x \mapsto x^2$  másodfokú polinom (parabola) esetében említjük meg, hogy a külön felrajzolt érintőn kívül az  $x$ -tengelyt is érintőnek gondolhatjuk a görbe  $(0, 0)$  pontjában.

Az  $x \mapsto x^3$  függvény esetében is érintőnek vélhetjük az  $x$ -tengelyt, bár érdekes, hogy az érintő ezen esetben átmetszi a görbét.

A szemléletes geometriai fantáziálás hozzásegíthet bennünket ahhoz, hogy az érintő fogalmát definiálni tudjuk. Meglepő — egy kicsit patetikusan fogalmazva: csodálatos — az, hogy erre az egyszerűen felvethető problémára csak az előzőekben bevezetett határérték fogalmát használva tudunk majd választ adni.



5.3. ábra: Szelők és érintő egy pontban.

A probléma szemléletes megközelítése közben, a rajzokat látva megfogalmazódhat bennünk valami olyasmi, hogy az érintőn egy olyan egyenest értsünk, amelyik “jól simul” a görbéhez az érintkezési pontban.

A “jól simul” pontosításával a következőképpen próbálkozhatunk: Olyan szelőt vesszünk az  $f$  függvény gráfjának az  $(x, f(x))$  pontban, amelyeknek a görbével vett másik  $(x+h, f(x+h))$  metszéspontja valamilyen értelemben közel van az  $(x, f(x))$  ponthoz, és úgy képzelhetjük, hogy ezek a szelők annál jobban megközelítik a keresett “érintőt”, minél közelebb van az említett két pont a görbén (5.3 ábra). (Azért szemléltetjük így a mondottakat, hogy hangsúlyozzuk azt a fontos tényt, hogy az  $(x+h, f(x+h))$  pont jobbra vagy balra egyaránt eshet az  $(x, f(x))$  ponttól, aszerint hogy  $h > 0$  vagy  $h < 0$ .)

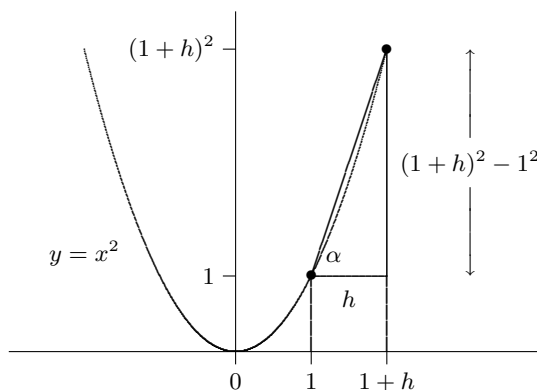
A későbbi definíció jobb megértéséhez példaképpen végezzük el a “simulásnak” egy teljesnek mondható pontosítását az  $x \mapsto x^2$  esetében, az  $(1, 1)$ , más szóval az  $x = 1$  pontban. A számolás geometriai háttere a 5.4. ábrán követhető.

Az  $(1, 1)$  és  $(1+h, (1+h)^2)$  pontokon átmenő szelőnek az  $x$  tengellyel bezárt  $\alpha$  szöge vizsgálatában sejthetjük a megközelítés kulcsát, mivel az változik akkor, amikor a  $h$  értéket változtatjuk, azaz amikor az  $(1+h, (1+h)^2)$  pontot közelítjük az  $(1, 1)$  ponthoz. A határérték fogalmának a birtokában most már pontosabban megfogalmazhatjuk a problémát: Keressük az  $\alpha$  szögnek a limeszét (határesetét) abban az esetben, amikor a  $h$  tart a nullához.

Teljes részletezettséggel számoljuk végig ezt az ötletet. A szóbanforgó  $\alpha$  szög tangensére az 5.4 ábra szerint azt kapjuk, hogy

$$\tan \alpha = \frac{(1+h)^2 - 1}{h}.$$

Ennek a kifejezésnek a nullánál nyilvánvalóan nincs értelme, de definiálva van minden nem nulla  $h$  esetében, tehát értelmezett a nulla egy hiányos környezetében,



5.4. ábra: A parabola érintője.

és így jogos a kérdés: Van-e az  $((1+h)^2 - 1)/h$  hányadosnak határtértéke és mi az, ha a  $h$  tart a nullához?

Ezt a határértéket könnyen meghatározhatjuk, mivel a nulla hiányos környezetében helyes az alábbi számolás:

$$\frac{(1+h)^2 - 1}{h} = \frac{(1+2h+h^2) - 1}{h} = \frac{2h+h^2}{h} = 2+h.$$

Ebből pedig adódik, hogy

$$\lim_{h \rightarrow 0} \frac{(1+h)^2 - 1}{h} = \lim_{h \rightarrow 0} (2+h) = 2.$$

Így azt kaptuk, hogy ésszerű ezt mondanunk: Az  $y = x^2$  parabola  $(1, 1)$  pontban vett “érintőjének” az iránytangense 2, azaz párhuzamos az  $y = 2x$  egyenessel.

Azért tettük idézőjelbe az “érintő” szót, mert *nem szabad azt gondolnunk, hogy van valami “érintő”, és mi most annak az iránytangensét számoltuk ki.*

A helyzet az, hogy a fogalomalkotás stádiumában vagyunk csak, és most próbáljuk alkalmasan meghatározni, hogy mit is értsünk majd az érintő fogalmán. A definíciót persze — habár az elvben bármilyen logikailag helyes meghatározás lehetne — szemléletes elvárásunkhoz illőnek szeretnénk látni.

Most már elérkeztünk ahhoz a ponthoz, hogy összefoglaljuk az eddigieket, és azután majd megadhatjuk a differenciálhányados pontos definícióját.

Vegyünk egy olyan  $f$  függvényt, ami értelmezve van az  $x$  pontnak egy környezetében. Ekkor a

$$h \mapsto \frac{f(x+h) - f(x)}{h} \quad (5.1)$$

leképezés értelmezve van a nulla egy hiányos környezetében. Elégké kis  $h$  mellett ugyanis az  $(x+h)$  beleesik az  $x$  azon környezetébe, ahol az  $f$  értelmezve van, és

így az 5.1 függvény definiált a nulla valamilyen környezetében, eltekintve magától a nullától, amikor is a nevező nulla.

A (5.1) függvényt definiáló törtnek egyszerű geometriai jelentése van: Az  $(x, f(x))$  és  $(x+h, f(x+h))$  pontokon átmenő szelőnek az iránytangense (5.5 ábra). Jogosan lehet olyan elvárásunk, hogy ennek a  $h=0$  helyen vett határértékét — ha van — az  $f$  függvény gráfja  $(x, f(x))$  pontban vett érintője iránytangensének tekintsük.

Ez az a kiinduló geometriai tartalom, amit a következő definícióban rögzítünk, és a differenciálhányados vagy derivált fogalmának a bevezetéséhez hozzásegít. Előbb azonban elnevezzük a (5.1) alatti függvényt.

**Definíció 200** Legyen az  $f$  valós változós és valós értékű függvény az  $x$  pont egy környezetében értelmezve. A

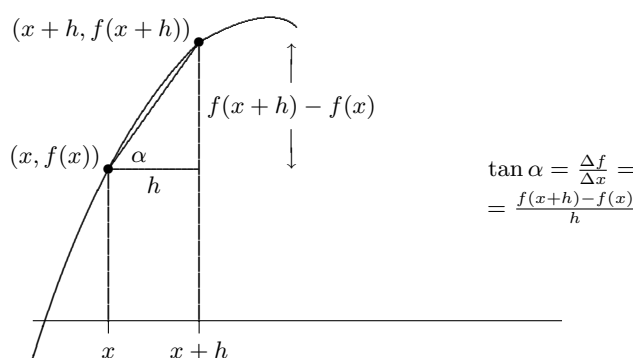
$$h \mapsto \frac{f(x+h) - f(x)}{h} \quad (5.2)$$

leképezést az  $f$  függvény  $x$  pontban vett differenciáhányados- (különbségihányados-) függvényének mondjuk.

Az (5.2) alatti differenciáhányadost más formában is szokásos írni. Ha az  $x+h$  helyett az  $y$  jelölést vezetjük be, akkor az

$$y \mapsto \frac{f(y) - f(x)}{y - x} \quad (5.3)$$

hányadoshoz jutunk. A definícióban szereplő alaknál az értelmezési tartomány a 0 egy hiányos környezete, az utóbbi esetben pedig az  $x$  egy hiányos környezete.



5.5. ábra: A differenciáhányados.

**Definíció 201** Legyen az  $f$  valós változós és valós értékű függvény az  $x$  pont egy környezetében értelmezve. Az  $f$  függvény  $x$  pontban vett

$$h \mapsto \frac{f(x+h) - f(x)}{h}$$

differenciahányadosának (differenciahányados-függvényének) a  $h = 0$  helyen vett határértékét — ha létezik — az  $f$  leképezés  $x$  pontban vett differenciálhányadosának vagy deriváltjának nevezzük és a következő jelölések valamelyikét szokás használni:

$$f'(x), \quad \frac{df(x)}{dx}, \quad \frac{d}{dx}f(x) \quad (5.4)$$

Ha létezik az  $f'(x)$  derivált, akkor az  $f$  függvényt differenciálhatónak vagy deriválhatónak mondjuk az  $x$  helyen.

Ha a (5.3) alatti differenciahányadossal dolgozunk, akkor ennek értelemszerűen az  $x$  helyen kell venni a határértékét (a  $h$  pontosan akkor tart a nullához, ha az  $y = x + h$  az  $x$ -hez). Így összefoglalva a definíció szerint:

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{y \rightarrow x} \frac{f(y) - f(x)}{y - x}.$$

A definíció szerint az  $f$  függvény az  $x$  pontnak valamilyen környezetében van értelmezve, ezért az  $x$  pont az  $f$  függvény értelmezési tartományának egy belső pontja. Ezt a tényt sohase felejtjük el!

A (5.4) jelölések történelmi, tradicionális eredetűek, és mindegyiknek megvan a maga előnye és hátránya, amik részletes magyarázatra szorulnak. Ezt most részletezzük is. A legrégebbi jelölés, ami az általunk használt  $f'(x)$ -nek felel meg, a Newtontól származó  $\dot{y}$  vagy  $\dot{f}(x)$ . Első pillantásra ez teljesen pontos jelölésnek látszik, de ez sajnos nem így van, amint azt a következő példa mutatja.

Legyen az  $f$  leképezés az  $(x^2 + 2xz - 1)$  formulával megadott függvény. Szándékosan pongyolán mondtuk meg, hogy mi a leképezés, hiszen lehet szó az  $x \mapsto (x^2 + 2xz - 1)$  függvényről, amire első látásra gondolunk, mert megszoktuk azt, hogy az  $x$  a független változó, de szóba jöhet a  $z \mapsto (x^2 + 2xz - 1)$  függvény is. Az előző megkülönböztetés nyilvánvalóan attól függ, hogy az  $x$  vagy a  $z$  változót képzeljük-e rögzítettnek. Mit jelentsen ekkor az  $f'(1)$ ? A  $z \mapsto (x^2 + 2xz - 1)$  vagy az  $x \mapsto (x^2 + 2xz - 1)$  leképezésnek a deriváltját az 1 helyen?

A “vesszős” jelölés erre a kérdésre nem ad pontos választ, tehát sajnos nem nyújt teljes információt a feladatról abban az esetben, ha a deriválandó függvényünkben több változó is szerepel, de persze csak az egyik a valódi változó, mert a másik rögzítettnek képzelt. Ebben az esetben a vesszős jelölést nem célszerű használni, de ha az itt leírt kétértelműség nem fordul elő, akkor mégis ez ajánlható leginkább, mert a legtömörebb.

A  $\frac{df(x)}{dx}$  és a  $\frac{d}{dx}f(x)$  jelölésekkel kapcsolatban először is egy fontos figyelmeztetés: Nem szabad a benne szereplő  $dx$  illetve  $df$  szimbólumokat a  $d$  és  $x$  illetve

a  $d$  és  $f$  szorzatának tekinteni, noha valóban annak látszanak. És nem szabad törtként sem felfogni, habár formálisan az, csak *egyetlen, egyszerűbb elemekre nem bontható szimbólumnak*. A formális eredetük az, hogy a (5.2) differenciahányadost a

$$\Delta x \doteq (x + h) - x = h \quad \text{és} \quad \Delta f(x) \doteq f(x + \Delta x) - f(x)$$

jelölésekkel az

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

alakba lehet írni. A határérték vételének a során a közönséges differenciákat jelölő  $\Delta f$  és  $\Delta x$  szimbólumokból lett a “végtelenül kicsi”  $df$  és  $dx$ . Ez az indoklás így persze egyáltalában nem pontos, bár alkalmas módon pontosítható.

Ha formalizmusnak megfelelően fogjuk fel (vagyis nem úgy, hogy az  $f(x)$  az  $f$ -nek az  $x$  helyen felvett értéke), akkor ez a jelölés nyújt a legteljesebb információt, mert megmondja — a  $dx$  révén — hogy melyik változó függvényének kell tekinteni az  $f$ -et, amikor a differenciáhányadosát keressük, ezért ezt célszerű használni az előző bekezdésben említett példa esetében, hiszen ekkor a

$$\frac{d(x^2 + 2zx - 1)}{dx} \quad \text{és} \quad \frac{d(x^2 + 2zx - 1)}{dz}$$

feladatok is pontosan definiáltak.

A legpontosabb — de kicsit körülményes — jelölést az  $f'(a)$ -ra talán a

$$\left. \frac{d}{dx} f(x) \right|_{x=a} \quad \text{vagy} \quad \frac{df}{dx}(a)$$

adna, mivel ez olvasható leginkább így: Vedd az  $f$  leképezésnek az  $x$  változó szerinti deriváltját az  $a$  helyen. Ha célszerű, akkor érdemes ezt is használni.

Az utóbbi jelöléshez és példához meg kell még említenünk, hogy abban az esetben, amikor egy függvényt több változót tartalmazó formula ad meg, mint az előbbi példánkban, akkor szokásos a “ $d$ ” betű helyett a “ $\partial$ ” jelet (stilizált görög delta) is használni, amit “gömbölyű  $d$ ”-nek lehet mondani. Ezzel a jelöléssel az előző példánk

$$\frac{\partial(x^2 + 2zx - 1)}{\partial x} \quad \text{és} \quad \frac{\partial(x^2 + 2zx - 1)}{\partial z}$$

formában is írható. Ezzel a szimbolikával a többváltozós analízis során fogunk újra találkozni.

A definiált fogalom elmélyítéséhez megoldunk négy feladatot, közvetlenül a meghatározásra támaszkodva.

**Példa 5.1** Az  $x \mapsto x^3$  függvény deriváltjának a meghatározása egy  $x$  helyen.

A differenciahányados átalakítása akkor, amikor az  $y$  az  $x$  egy hiányos környezetébe esik:

$$\frac{y^3 - x^3}{y - x} = y^2 + xy + x^2.$$

A határérték számolása:

$$\lim_{y \rightarrow x} \frac{y^3 - x^3}{y - x} = \lim_{y \rightarrow x} (y^2 + xy + x^2) = x^2 + xx + x^2 = 3x^2.$$

tehát

$$\frac{d}{dx} x^3 = 3x^2,$$

amit röviden  $(x^3)' = 3x^2$  módon is írhatunk.  $\square$

**Példa 5.2** Határozzuk meg  $k : x \mapsto \sqrt{1 - x^2}$  (úgy is nevezhetnénk: kör vagy félkör) leképezésnek a deriváltját egy tetszőleges  $b \in (-1, 1)$  helyen.

Nyilvánvaló, hogy a  $k$  függvény minden  $b \in (-1, 1)$  pontnak valamilyen környezetében definiált, ezért a kérdés értelmes. A  $b = 1$  és  $b = -1$  esetekben nem mondható el ugyanez, mert az 1 és  $-1$  nem belső pontjai az értelmezési tartománynak. Később majd úgy általánosítjuk a deriváltfogalmat, hogy az ilyen esetekben is lehessen deriváltról beszélni.

A különbséghányados-függvény a  $b$  helyen az

$$y \mapsto \frac{\sqrt{1 - y^2} - \sqrt{1 - b^2}}{y - b} \quad (y \neq b)$$

függvény. A számlálót és nevezőt  $(\sqrt{1 - y^2} + \sqrt{1 - b^2})$ -tel szorozva és felhasználva az  $(u + v)(u - v) = u^2 - v^2$  azonosságot, a  $b$  pont valamilyen hiányos környezetébe eső  $y$  mellett helyes az alábbi átalakítás:

$$\begin{aligned} \frac{\sqrt{1 - y^2} - \sqrt{1 - b^2}}{y - b} &= \frac{(1 - y^2) - (1 - b^2)}{(y - b)(\sqrt{1 - y^2} + \sqrt{1 - b^2})} = \\ &= \frac{b^2 - y^2}{(y - b)(\sqrt{1 - y^2} + \sqrt{1 - b^2})} = \frac{(b - y)(b + y)}{(y - b)(\sqrt{1 - y^2} + \sqrt{1 - b^2})} \\ &= -\frac{b + y}{\sqrt{1 - y^2} + \sqrt{1 - b^2}}. \end{aligned}$$

Így pedig az adódik, hogy

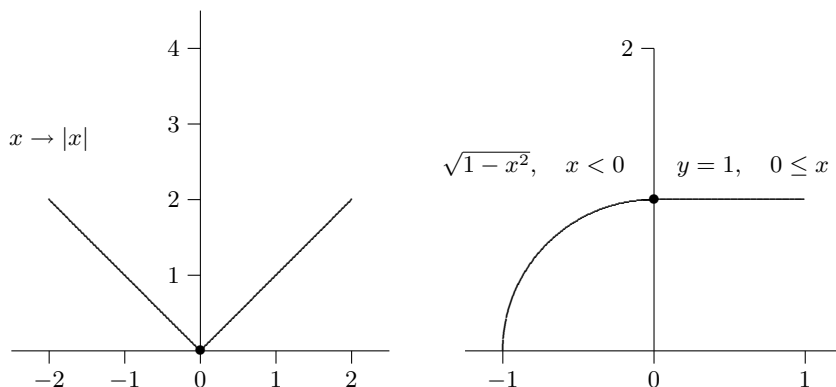
$$\begin{aligned} \lim_{y \rightarrow b} \frac{\sqrt{1 - y^2} - \sqrt{1 - b^2}}{y - b} &= \lim_{y \rightarrow b} \left( -\frac{b + y}{\sqrt{1 - y^2} + \sqrt{1 - b^2}} \right) = \\ &= -\frac{2b}{2\sqrt{1 - b^2}} = -\frac{b}{\sqrt{1 - b^2}}, \end{aligned}$$

tehát végül is azt kaptuk, hogy

$$\frac{d}{db} \sqrt{1 - b^2} = -\frac{b}{\sqrt{1 - b^2}},$$

feltéve, hogy  $b \in (-1, 1)$ . □

Ebből az eredményből kiindulva már beláthatjuk azt is, hogy az érintési ponthoz húzott sugár merőleges az érintőre.



5.6. ábra: Abszolút érték és egy “összetoldott” függvény.

**Példa 5.3** Az 5.6. ábra jobboldali részén az alábbi, a  $[-1, 1]$  zárt intervallumon értelmezett  $g$  függvényt ábrázoltuk.

$$g(x) = \begin{cases} \sqrt{1-x^2} & \text{ha } -1 \leq x < 0 \\ 1 & \text{ha } 0 \leq x \leq 1 \end{cases}.$$

Határozzuk meg a  $g$  függvény deriváltját az  $x = 0$  helyen.

Vegyük észre, hogy a görbét — szemléletesen szólva — egy negyedkörből és egy vízszintes egyenesből raktuk össze (5.6. ábra). Vegyük figyelembe azt, hogy a függvény a 0-tól balra illetve jobbra más-más előírással definiált, és nézzük a 0 helyhez tartozó differenciahányadost:

$$\frac{g(0+h) - g(0)}{h} = \frac{g(h) - 1}{h} = \begin{cases} \frac{\sqrt{1-h^2}-1}{h} & \text{ha } -1 < h < 0 \\ \frac{1-1}{h} = 0 & \text{ha } 0 < h \leq 1 \end{cases}.$$

Eszerint ha  $0 < h \leq 1$ , akkor a differenciahányados értéke nulla, ha pedig  $-1 \leq h < 0$ , akkor megengedett a következő átalakítás elvégzése:

$$\begin{aligned} \frac{\sqrt{1-h^2}-1}{h} &= \frac{(\sqrt{1-h^2}-1)(\sqrt{1-h^2}+1)}{h(\sqrt{1-h^2}+1)} = \\ &= \frac{(1-h^2)-1}{h(\sqrt{1-h^2}+1)} = \frac{-h^2}{h(\sqrt{1-h^2}+1)} = \frac{-h}{\sqrt{1-h^2}+1}. \end{aligned}$$

Az eddigi számolások eredményeinek a felhasználásával azt írhatjuk, hogy a következő folytonos függvény a nulla egy hiányos környezetében megegyezik a differenciáhányados-függvénnyel:

$$\begin{array}{ll} -\frac{h}{\sqrt{1-h^2}+1} & \text{ha } -1 < h < 0 \\ 0 & \text{ha } 0 \leq h < 1 \end{array} ,$$

amiből, véve ennek a  $h = 0$  helyettesítési értékét, az adódik, hogy

$$\lim_{h \rightarrow 0} \frac{g(0+h) - g(0)}{h} = 0.$$

Tehát végül is kiszámoltuk, hogy  $g'(0) = 0$ .  $\square$

**Példa 5.4** Határozzuk meg az  $x \mapsto (\sin 2x + x^2)$  függvény érintőjének az iránytangensét az  $x = 0$  pontban.

A definíció szerint:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{(\sin 2h + h^2) - (\sin 0 + 0^2)}{h} &= \lim_{h \rightarrow 0} \frac{\sin 2h + h^2}{h} = \\ &= \lim_{h \rightarrow 0} \left( 2 \frac{\sin 2h}{2h} + h \right) = 2 \cdot \lim_{h \rightarrow 0} \frac{\sin 2h}{2h} + \lim_{h \rightarrow 0} h = 2 \cdot 1 + 0 = 2. \end{aligned} \quad \square$$

A következő egyszerű észrevétel azt mutatja, hogy a deriválhatóság erősebb (többet követelő) fogalom a folytonosság fogalmánál, mert a deriválhatóságból következik a folytonosság.

**Állítás 202** Ha az  $f$  függvény deriválható egy pontban, akkor ott folytonos is.

**Bizonyítás.** A differenciáhányados (5.3) alakjával számolva azt kapjuk, hogy

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

Az összeg és szorzat határértékére vonatkozó számolási szabályok felhasználásával a következőhöz jutunk:

$$\begin{aligned} \lim_{y \rightarrow x} f(y) - f(x) &= \lim_{y \rightarrow x} (f(y) - f(x)) = \\ &= \lim_{y \rightarrow x} \frac{f(y) - f(x)}{y - x} \cdot \lim_{y \rightarrow x} (y - x) = f'(x) \cdot 0 = 0 \end{aligned}$$

tehát

$$f(x) = \lim_{y \rightarrow x} f(y).$$

Ez pedig pontosan azt jelenti, hogy az  $f$  függvény folytonos az  $x$  helyen.  $\square$

Az előző állítással ellenkező irányban: A folytonosság nem implikálja a deriválhatóságot, ahogyan azt az alábbi példa mutatja.

**Példa 5.5** *Vizsgáljuk meg az  $x \mapsto |x|$  abszolútérték-függvényt az  $x = 0$  helyen a deriválhatóság szempontjából.*

Az abszolútérték-függvény differenciahányadosa

$$h \mapsto \frac{|0+h| - |0|}{h} = \frac{|h|}{h},$$

és ez 1, ha a  $h$  pozitív, és  $-1$ , ha a  $h$  negatív, ezért a 0-ban nyilvánvalóan nincs határértéke. Nézzük meg figyelmesen, milyen az abszolút érték grafikonja az  $x = 0$  helyen (5.6. ábra).  $\square$

Legyen most az  $f$  függvény egy  $(a, b)$  nyílt intervallumban értelmezve. Mivel ekkor az  $(a, b)$  nyílt intervallum minden pontjának valamilyen környezetében értelmezett a függvény, ezért a 201. definíció szerint feltehető a kérdés: Létezik-e a derivált tetszőleges  $x \in (a, b)$  helyen? Ha a válasz igenlő, akkor minden  $x \in (a, b)$  pontban létezik az  $f'(x)$  érték, ezért az  $x \mapsto f'(x)$  leképezés az  $(a, b)$  nyílt intervallum minden pontjában definiált. Az elmondottak tetszőleges nyílt halmazra érvényesek, ezért helyes a következő definíció.

**Definíció 203** *Ha egy  $f : (a, b) \rightarrow \mathbb{R}$  függvény az  $O \subseteq (a, b)$  nyílt halmaz minden pontjában deriválható, akkor azt mondjuk, hogy az  $f$  függvény deriválható az  $O$  halmazon, és az eszerint létező,*

$$x \mapsto f'(x), \quad O \rightarrow \mathbb{R}$$

*függvényt az  $f$  függvény differenciálhányadosának vagy deriváltjának (differenciálhányados- vagy deriváltfüggvényének) nevezzük és az*

$$f', \quad \frac{df}{dx} \quad \text{vagy} \quad \frac{d}{dx}f$$

*jelölések valamelyikét használjuk.*

Hangsúlyozni kell, hogy az  $f$  függvénynek az 201. definíció szerint valamilyen  $x$  helyen vett deriváltja egy  $f'(x)$  szám, az előző definícióban meghatározott  $f'$  pedig egy függvény, aminek az  $x$  helyen felvett értéke az  $f'(x)$  szám. Erre a különbségre érdemes figyelni.

A gyakorlatban az  $O$  nyílt halmaz rendszerint egy  $(a, b)$  (korlátos vagy korlátlan) nyílt intervallum. Most pedig lássunk néhány példát a deriváltfüggvény meghatározására.

**Példa 5.6** *Határozzuk meg az  $x \mapsto x^2$ ,  $\mathbb{R} \rightarrow \mathbb{R}$  függvény deriváltját (deriváltfüggvényét).*

A jelen pont elején a bevezető mintapéldánk éppen e függvény deriváltjának a meghatározása volt az  $x = 1$  pontban. Az ott leírtakat az 1 helyett az  $x$  helyen végrehajtva azt kapjuk, hogy

$$(x^2)' = \lim_{y \rightarrow x} \frac{y^2 - x^2}{y - x} = \lim_{y \rightarrow x} (y + x) = 2x,$$

tehát az  $x \mapsto x^2$ ,  $\mathbb{R} \rightarrow \mathbb{R}$  függvény deriváltja az  $x \mapsto 2x$ ,  $\mathbb{R} \rightarrow \mathbb{R}$  függvény.  $\square$

**Példa 5.7** Az 5.2. és 5.1. példák alapján mondjuk meg, hogy mi az

$$x \mapsto x^3, \quad \mathbb{R} \rightarrow \mathbb{R} \quad \text{és} \quad x \mapsto \sqrt{1-x^2}, \quad (-1, 1) \rightarrow (0, 1)$$

függvények deriváltja.

Az  $x \mapsto x^3$ ,  $\mathbb{R} \rightarrow \mathbb{R}$  függvény deriváltja az  $x \mapsto 3x^2$ ,  $\mathbb{R} \rightarrow \mathbb{R}$  függvény. Az  $x \mapsto \sqrt{1-x^2}$ ,  $(-1, 1) \rightarrow \mathbb{R}$  függvény deriváltja pedig az

$$f'(x) = -\frac{x}{\sqrt{1-x^2}},$$

$(-1, 1) \rightarrow \mathbb{R}$  függvény.  $\square$

**Példa 5.8** Hol deriválható és hol nem az  $x \mapsto |x|$ ,  $\mathbb{R} \rightarrow \mathbb{R}_+$  függvény?

Az 5.5 példában már láttuk, hogy a nulla helyen nem deriválható a függvény, de az ábrája alapján (5.6. ábra) azt gondolhatjuk, hogy ezen a helyen kívül nincs baj. Valóban vegyük először a  $0 < x$  esetet. Ekkor a derivált meghatározása könnyű feladat:

$$\frac{d}{dx}|x| = \lim_{y \rightarrow x} \frac{|y| - |x|}{y - x} = \lim_{y \rightarrow x} \frac{y - x}{y - x} = 1,$$

ahol az  $y$  értéket olyan közel választottuk az  $x$ -hez, hogy az is pozitív legyen. Ha pedig  $x < 0$ , akkor — olyan közel választva az  $y$  számot az  $x$ -hez, hogy az is negatív legyen — azt kapjuk, hogy

$$\frac{d}{dx}|x| = \lim_{y \rightarrow x} \frac{|y| - |x|}{y - x} = \lim_{y \rightarrow x} \frac{(-y) - (-x)}{y - x} = -1.$$

Ezeket összefoglalva megállapíthatjuk, hogy az abszolútérték-függvény az  $x = 0$  hely kivételével mindenütt deriválható, ezért azt állíthatjuk, hogy az

$$x \mapsto |x|, \quad x \in \mathbb{R} \setminus \{0\}$$

függvény (ami az értelmezési tartomány különbözősége miatt nem azonos az abszolútérték-függvénnyel) deriválható, és a deriváltja az

$$s(x) = \begin{cases} 1 & \text{ha } 0 < x \\ -1 & \text{ha } x < 0 \end{cases}$$

függvény, azaz a  $\operatorname{sgn}$  függvény leszűkítése az  $\mathbb{R} \setminus \{0\}$  halmazra.  $\square$

Némelykor szükségünk lesz a zárt intervallumon való deriválhatóságra is. Ehhez először is általánosítanunk kell a deriváltfogalmat, amit — felhasználva a jobb- és baloldali határérték fogalmát — könnyen megtehetünk:

**Definíció 204** Legyen az  $f$  valós függvény értelmezve az  $x$  pont egy baloldali környezetében. Ha a

$$h \mapsto \frac{f(x+h) - f(x)}{h} \quad (h < 0)$$

differenciahányados-függvénynek létezik a baloldali határértéke a nulla helyen, akkor az  $x$  pontban balról deriválhatónak mondjuk, és erre a határértékre a következő jelölések valamelyikét használjuk.

$$f'_-(x), \quad \frac{df(x)}{dx_-}, \quad \frac{d}{dx_-}f(x).$$

Hasonlóan definiálható az

$$f'_+(x), \quad \frac{df(x)}{dx_+}, \quad \frac{d}{dx_+}f(x)$$

módok valamelyikével jelölhető jobboldali deriváltfogalom. Tehát

$$f'_+(x) = \lim_{h \searrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{y \searrow x} \frac{f(y) - f(x)}{y - x}.$$

$$f'_-(x) = \lim_{h \nearrow 0} \frac{f(x+h) - f(x)}{h} = \lim_{y \nearrow x} \frac{f(y) - f(x)}{y - x}.$$

A határértékfogalomról tanultak szerint, ha a baloldali és jobboldali határértékek megegyeznek, akkor létezik a közös értékkel megegyező határérték. Eszerint ha egy helyen a bal- és a jobboldali deriváltak megegyeznek, akkor ott a függvénynek létezik a deriváltja, a közös érték. Szépen látható ez a tény azon a példánkon, ahol a függvényt egy negyedkörből és egy vízszintes egyenesből raktuk össze. Az abszolútérték-függvény példája is mutatja, hogy ha léteznek a bal- és jobboldali deriváltak egy pontban, de különbözőek, akkor ott a függvény grafikonja "hegyesnek" látszik, nem "sima", mint a deriválhatósági pontokban.

A 203. definíció mintájára megfogalmazhatjuk azt, hogy mit fogunk érteni egy függvénynek egy zárt intervallumban való differenciálhatóságán.

**Definíció 205** Legyen  $f : [a, b] \rightarrow \mathbb{R}$  függvény. Azt fogjuk mondani, hogy az  $f$  függvény differenciálható az  $[a, b]$  zárt intervallumon, ha

- (1) az  $f$  differenciálható az  $(a, b)$  nyílt intervallumon, és
- (2) létezik az  $a$  pontban a jobboldali, a  $b$  pontban pedig a baloldali deriváltja.

Ekkor az  $f' : [a, b] \rightarrow \mathbb{R}$  deriváltfüggvényt az alábbi módon definiáljuk:

$$f'(x) \doteq \begin{cases} f'(x) & \text{ha } x \in (a, b) \\ f'_+(a) & \text{ha } x = a \\ f'_-(b) & \text{ha } x = b \end{cases}.$$

**Példa 5.9** Legyen az  $f : \mathbb{R} \rightarrow \mathbb{R}$  függvény a következőképpen definiálva:

$$f(x) \doteq \begin{cases} -2x & \text{ha } x < 0 \\ 0.5x & \text{ha } 0 \leq x \end{cases}.$$

Határozzuk meg a bal- és jobboldali deriváltakat az  $x = 0$  helyen.

A differenciahányados balról  $\frac{-2(0+h)-(-2 \cdot 0)}{h} = \frac{-2h}{h} = -2$ , és így a baloldali derivált  $-2$ . Hasonlóan a jobboldalról  $\frac{0.5(0+h)-0.5 \cdot 0}{h} = \frac{0.5h}{h} = 0.5$ , tehát a jobboldali derivált  $0.5$ .  $\square$

Ennek az alpontnak a végére végül is már csak egy igen természetes definíció maradt. Ha egy  $f$  függvény  $f'$  deriváltfüggvénye létezik az  $x$  pont valamilyen környezetében, akkor felvethető a kérdés: Az  $f'$  függvény deriválható-e az  $x$  pontban? Hangsúlyoznunk kell, hogy ez csak akkor kérdezhető, ha az  $f$  nem csak az  $x$  pontban, hanem annak egy egész környezetében deriválható. Ha a feltett kérdésre igenlő a válasz, akkor az  $f'$  függvény  $x$  pontban vett deriváltját az  $f$  második deriváltjának nevezzük, és az

$$f''(x), \quad f^{(2)}(x), \quad \left(\frac{d}{dx}\right)^2 f(x), \quad \frac{d^2 f(x)}{dx^2}, \quad \frac{d^2 f}{dx^2}(x)$$

jelölések valamelyikét használjuk.

A második derivált mintájára tetszőleges  $n$  pozitív egész szám esetében definiálható az  $n$ -edrendű derivált:

**Definíció 206** Tegyük fel, hogy az  $f$  függvény az  $x$  pont valamilyen környezetében (annak minden pontjában)  $(n-1)$ -szer deriválható. Ha az  $f^{(n-1)}$  függvény deriválható az  $x$  pontban, akkor a deriváltját  $f$  függvény  $n$ -edik vagy  $n$ -edrendű deriváltjának nevezzük az  $x$  pontban, és a következő jelölések valamelyikét használjuk:

$$f^{(n)}(x), \quad \left(\frac{d}{dx}\right)^n f(x), \quad \frac{d^n f}{dx^n} f(x).$$

Meg kell említenünk, hogy az eddigi elnevezésekkel és megállapodásokkal összhangban lesz az, ha magát az  $f$  függvényt a 0-adik deriváltnak nevezzük, amit néha meg is teszünk. Az  $n$ -edik derivált előző, teljes indukción alapuló definíciójában eszerint a megállapodás szerint az  $n = 0$  is lehet a kiinduló eset.

### 5.1.2 Érintő és érintőapproximáció

A derivált fogalmának a definíciójánál az érintő "hétköznapien elképzelt" szemléletéből indultunk ki. Nem szabad azonban azt hinnünk, hogy az érintő fogalmát valamilyen "tapasztalati tényként" ismerjük, hiszen azt csak a derivált birtokában tudjuk pontosan definiálni. Ebben az alpontban az érintőfogalommal foglalkozunk, abból a célból, hogy a differenciálszámítás tartalmi vonatkozásait elmélyítsük.

**Definíció 207** Legyen az  $f$  függvény deriválható az  $x$  pontban. Ekkor az  $f$  függvény gráfja  $(x, f(x))$  pontban vett érintőjének mondjuk azt az egyenest, amely

- (1) átmegy az  $(x, f(x))$  ponton, és
- (2) az iránytangense az  $f'(x)$  differenciálhányados.

Ezekután már egzaktan tekinthetjük az érintő fogalmát, és könnyen fel is írhatjuk a szóbanforgó érintőegyenes egyenletét. Ezt az egyszerű, de fontos tényt állításban is rögzítjük.

**Állítás 208** Ha az  $f$  függvény deriválható az  $a$  pontban, akkor az  $f$  leképezés gráfját az  $(a, f(a))$  pontban érintő egyenes egyenlete:

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

Bizonyítás. Mint tudjuk, az  $f'(a)$  iránytangensű egyenes egyenletének az általános alakja

$$y = f'(a)x + b,$$

és mivel az egyenesnek át kell mennie az  $(a, f(a))$  ponton, ezért

$$f(a) = f'(a)a + b.$$

A két utóbbi egyenlőséget kivonva egymásból, a  $b$  kiesik, és

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

ahogyan állítottuk. □

Az érintő fogalmával kapcsolatban szemléletesen helyes az a felfogás, hogy az érintő egy olyan egyenes, amelyik az érintési pontban “hozzásimul” a görbéhez. Most tovább vizsgáljuk ezt a “simulást”.

Az érintő nemcsak hogy megegyezik a görbével az érintési pontban, hanem az érintési pont közelében eléggé jól “approximálja”, közelíti is a görbét. Vegyük például az előző pont elején vizsgált  $x \mapsto x^2$  parabolát. Ennek az  $(1, 1)$  pontban az érintője az előző tétel szerint az  $y - 1 = 2(x - 1)$  egyenes, ami rendezve  $y = 2x - 1$ . Számoljuk ki néhány pontban az érintőnek és a parabolának az  $y$  koordinátáját az 1 ponthoz közel:

$x$	0.9	0.99	0.999	1	1.001	1.01	1.1
$2x - 1$	0.8	0.98	0.998	1	1.002	1.02	1.2
$x^2$	0.81	0.9801	0.99801	1	1.0021	1.0201	1.21

Elvárásunknak megfelelően megállapíthatjuk, hogy az  $x \mapsto x^2$  és  $x \mapsto (2x - 1)$  függvény az 1 pont közelében eléggé nagy pontossággal megegyeznek. Ez pedig nagy szerencsének mondható, hiszen a lineáris függvény — jelen esetben a  $(2x - 1)$  — számolása lényegesen egyszerűbb, mint a négyzetre emelés, nem is beszélve

az ennél bonyolultabb leképezésekről. Az itt vázolt észrevételt pontosítjuk a továbbiakban.

Most részletesen megvizsgáljuk az  $x$  pontban deriválható  $f$  függvény differenciáhányadosának a viselkedését az  $x$  egy környezetében. Párhuzamosan — konkrét példaként — az  $f(x) = x^2$  függvényre (amit a fentiekben is használtunk példaként) vonatkozó megfelelő formulákat, számolási eredményeket zárójelbe írjuk az általános eset alá.

Ha az  $f$  függvény differenciálható az  $a$  pontban, akkor a nulla valamilyen hiányos környezetében az alábbi módon definiált  $r$  függvény

$$r(h) \doteq \frac{f(a+h) - f(a)}{h} - f'(a) \quad (5.5)$$

$$\left( r(h) \doteq \frac{(a+h)^2 - a^2}{h} - 2a = h \right)$$

nullához tart, ha a  $h$ -val tartunk a nullához. Ennek az átrendezésével az

$$f(a+h) = f(a) + f'(a)h + hr(h) \quad (5.6)$$

$$((a+h)^2 = a^2 + 2ah + hr(h) = a^2 + 2ah + h^2)$$

előállítást kaptuk az  $f(a+h)$ -ra, azaz az  $f$  függvénynek az  $x$  pont egy hiányos környezetében felvett értékeire. Érdekes felírni a (5.6) formát abban az esetben is, ha a differenciáhányadosnak (5.5) helyett az

$$\frac{f(x) - f(a)}{x - a} \quad (5.7)$$

alakjával dolgozunk. Mivel ezen utóbbi esetben  $x = a + h$  azaz  $h = x - a$ , ezért az (5.6) formának megfelelő alak:

$$f(x) = f(a) + f'(a)(x - a) + (x - a) \cdot r(x - a) \quad (5.8)$$

$$(x^2 = a^2 + 2a(x - a) + (x - a) \cdot r(x - a) = a^2 + 2(x - a) + (x - a)^2).$$

Ha szemügyre vesszük a 208. állításban szereplő érintőegyenesnek az egyenletét, akkor észrevehetjük, hogy a (5.8) egyenlet jobboldalán éppen az

$$\begin{aligned} y &= f(a) + f'(a)(x - a) \\ (y &= a^2 + 2a(x - a)) \end{aligned}$$

érintőegyenesnek a függvényértéke és még az

$$\begin{aligned} &(x - a) \cdot r(x - a) \\ ((x - a) \cdot r(x - a) &= (x - a)^2) \end{aligned}$$

— mondjuk így: hibatag — szerepel.

Ha a parabola esetében nézzük a hibatagot, akkor azt láthatjuk, hogy az  $h^2 = (x - a)^2$ , és így megállapíthatjuk, hogy “kicsi”, ha a  $h$  kicsi, mégpedig a  $h$ -nál is “lényegesen” kisebb, hiszen annak a négyzete. Egészítsük ki az előző táblázatot szemléltetésképpen az itt kapott  $h^2 = (x - a)^2$  hibatag kiszámolásával az  $a = 1$  hely körül, ami szépen illusztrálja a magyarázatunkat:

$x$	0.9	0.99	0.999	1	1.001	1.01	1.1
$h = x - 1$	-0.1	-0.01	-0.001	0	0.001	0.01	0.1
$(x - 1)^2$	0.01	0.001	0.00001	0	0.00001	0.001	0.01

Térjünk most vissza az (5.7) és (5.8) alakokhoz. Az előző megfontolást összefoglalva: A  $h \mapsto f(a + h)$  illetve az  $x \mapsto f(x)$  a 0 illetve az  $a$  pont egy hiányos környezetében az

$$f(a + h) = f(a) + f'(a)h + hr(h) \quad (5.9)$$

illetve az

$$f(x) = f(a) + f'(a)(x - a) + (x - a)r(x - a). \quad (5.10)$$

alakba írható, ahol a  $h \mapsto r(h)$  illetve  $x \mapsto r(x - a)$  hibatagok nullához tartanak, ha a  $h$  nullához illetve ha az  $x$  az  $a$ -hoz tart. Emiatt a  $hr(h)$  hibatag olyan erősen tart a nullához, hogy még a szintén nullához tartó  $h$ -val osztva is nullához tart. Így az  $(f(a) + f'(a)h)$  lineáris függvény valóban jól közelíti az  $f(a + h)$  függvénynek az értékeit, ha a  $h$  eléggé kicsi. Végül is szemléletesen szólva ezt mondhatjuk: Az  $a$  pontban deriválható függvény közelítőleg egyenes (lineáris) az  $a$  pont valamilyen környezetében. Más szavakkal:

*A deriválhatósági pontnál a függvény “lokálisan egyenes”.*  
Más szóhasználat: “kicsiben lineáris”.

A felvetett gondolat pontos megfogalmazásához szükség lesz egy új fogalomra, amit a következő definícióban vezetünk be.

**Definíció 209** Ha a  $g$  valós leképezés értelmezve van a nulla egy környezetében és

$$\lim_{h \rightarrow 0} \frac{g(h)}{h} = 0$$

és

$$g(0) = 0,$$

akkor azt mondjuk, hogy a  $g$  leképezés erősebben tart a nullához, mint a  $h \mapsto h$  (röviden:  $h$ ) leképezés. Az ilyen  $g$  leképezéseket  $o(h)$  — olvasva: kis ordó  $h$  — módon fogjuk jelölni, és azt is szokás mondani, hogy a  $g$  függvény kis ordó  $h$  (kisrendű  $h$ ).

Ez a definíció igen szemléletes szóhasználatot fogalmaz meg. Tartalmának a megértéséhez még egy kis magyarázat: Ha a  $g(h)/|h|$  leképezést  $t(h)$ -val jelöljük,

akkor átrendezve azt kapjuk, hogy  $g(h) = |h|t(h)$ , ahol a definíció szerint a  $t(h)$  függvény a nullához tart, ha a  $h$  tart a nullához. Ez azt jelenti, hogy a  $g$  függvény — a  $h$  nullához tartásával — “sebesebben” tart a nullához, mint maga a  $h$  érték, hiszen még a nullához tartó abszolútérték-függvénnyel is meg van szorozva.

A “kis ordó” elnevezés arra kíván utalni, hogy az  $o(h)$  függvény a nullánál “nagyságrendileg” vagy “lényegesen” kisebb, mint a  $h$ . A definícióból világos, hogy ez a “kicsiség” limeszben értendő.

Most pedig a kezdeti részletes bevezető magyarázat után pontosan megfogalmazzuk az “érintőapproximáció” (érintőközelítés) tételét, és ezzel nemcsak a fentiekben vázolt “kicsiben lineáris” gondolatot tesszük precízzé, hanem a deriváltfogalomnak is egy új szemléletet, tartalmat adunk.

**Állítás 210 (Érintőapproximáció-tétel.)** *Az  $a \in \mathbb{R}$  szám valamilyen környezetében értelmezett  $f$  valós függvény pontosan akkor differenciálható az  $a$  pontban, ha van olyan  $\alpha$  szám, hogy a nulla egy környezetében*

$$f(a+h) - f(a) = \alpha h + o(h),$$

vagy

$$f(x) - f(a) = \alpha \cdot (x - a) + o(x - a),$$

és ekkor az  $\alpha$  szám megegyezik az  $f$  függvénynek az  $a$  pontban vett  $f'(a)$  deriváltjával, azaz

$$f(a+h) - f(a) = f'(a)h + o(h),$$

illetve

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

**Bizonyítás.** Az  $f$  függvény  $x$  helyen vett differenciáhányadosa pontosan akkor tart az  $f'(x)$  differenciáhányadoshoz, ha a

$$h \mapsto \frac{f(x+h) - f(x)}{h} - f'(x) = \frac{f(x+h) - f(x) - hf'(x)}{h}$$

leképezés a nullához tart a  $h$  nullához tartása mellett, amit úgy is mondhatunk, hogy a leképezést megadó tört számlálója  $h$ -val osztva nullához tart, és a nullában nulla.

Ez pedig az előző definícióban bevezetett kis ordó terminológiával pontosan azt jelenti, hogy a kifejezés baloldalának az

$$f(x+h) - f(x) - hf'(x)$$

számlálója kis ordó  $h$  függvény, azaz formálisan

$$f(x+h) - f(x) - hf'(x) = o(h),$$

amit átrendezve éppen a tétel egyenlőségéhez jutunk. □

A tétel egyszerű bizonyítása ellenére azért adtunk neki külön nevet, mert rendkívül fontos jellemzését adja a deriválhatóságnak és a deriváltnak. Ennek a karakterizációnak nagyon fontos előnyei vannak, amit majd a többváltozós analízisnél fogunk tapasztalni, ahol a deriváltfogalomnak ezt az értelmezését tudjuk továbbvinni.

Végezetül lássunk még egy olyan feladatot, ami egy nagyon szemléletes geometriai észrevételt bizonyít.

**Példa 5.10** *Mutassuk meg, hogy a parabola gráfja tetszőleges érintője felett helyezkedik el, ami formálisan azzal ekvivalens, hogy tetszőleges  $y = mx + b$  érintő esetében minden  $x$  pontra  $mx + b \leq x^2$ . Milyen  $x$  helyeken áll fenn egyenlőség?*

Az  $y = x^2$  parabola függvény deriváltja az  $a$  helyen, ahogyan már láttuk,  $2a$ , ezért a  $(a, a^2)$  ponton átmenő érintő egyenlete a 208. állítás szerint

$$y - a^2 = 2a(x - a),$$

ami rendezve  $y = 2ax - a^2$  alakba írható. Eszerint azt kell megmutatnunk, hogy minden  $x$ -re

$$2ax - a^2 \leq x^2.$$

Ez az egyenlőtlenség azonban pontosan akkor áll fenn, ha az

$$x^2 - 2ax + a^2 = (x - a)^2 \geq 0$$

egyenlőtlenség igaz, ami mindig határozottan pozitívan teljesül, kivéve az  $x = a$  érintési pontot.  $\square$

### 5.1.3 Sebesség

Ahogyan a bevezetésben is írtuk, a matematika egy precíz, egzakt nyelv, amiben olyan fogalmakat alakítunk ki, amelyek alkalmasak arra, hogy segítségükkel a hétköznapi életben használt — inkább csak érzett, mint tudott — fogalmaknak pontos “jelentést” tudjunk adni. A jelentésen persze nem érthetünk soha mást, mint egy alkalmasan alkotott matematikai fogalomnak a jelenséghez való hozzákapcsolását. Ezzel a hozzáállással lehetővé válik a “valóság” valamilyen értelemben vett matematikai “leírása”.

Mindenki a legtermészetesebb módon beszél “sebességről” vagy “gyorsaságról”. Például a fizikában: “ez az autó ilyen vagy olyan sebességgel megy.” Társadalmi vagy gazdasági jelenségek változására: “az emberek elmagányosodása sebesen nő.” “A pénz értéke romlásának a sebessége nagy.” Ezek a megállapítások meglehetősen semmitmondóak és rosszul definiáltak mindaddig, amíg a “sebességről” nincs pontosabb fogalmunk.

Meglehetősen meglepő az, hogy ennek a hétköznapi “sebesség”-fogalomnak csak a differenciálhányados fogalmának a felhasználásával tudunk pontos értelmet

adni. Meg kell említeni, hogy történetileg ez a probléma volt a deriváltfogalom felfedezésének a legfontosabb ösztönzője, mondhatni létrehozója.

Kezdjük egy fizikai példával. Tegyük fel, hogy egy anyagi pont indul el az origóból az  $x$  tengely pozitív irányába és  $t$  nagyságú idő alatt  $s(t)$  utat tesz meg. Hogyan tudnánk a pont mozgásának a "sebességét" megragadni? Csak az időt és a megtett utak hosszát tudjuk mérni. Kézenfekvő, hogy először egy átlagos sebesség meghatározásával próbálkozunk:

$$\begin{aligned} \text{átlagsebesség} &= \\ &= \frac{\text{A } t \text{ és } (t+h) \text{ időpontok között megtett út}}{\text{Az út megtételéhez szükséges időtartam}} = \frac{s(t+h) - s(t)}{h}. \end{aligned} \quad (5.11)$$

Az így számolt átlagsebesség nyilvánvalóan erősen függ a választott  $h$  időtartamtól, és ha a  $h$  értéke nagy, akkor nem jellemzi jól a mozgásnak a  $t$  időpontban való viselkedését.

Például: Ha éppen megálltunk az autóval és vesszük az utolsó negyedóra vonatkozó átlagsebességet, akkor az esetleg nagy pozitív érték lehet, holott már állunk, tehát a sebességünket jogosan nullának tartanánk. Az átlagsebesség annál jobban jellemző a szóbanforgó  $t$  időpontra, minél kisebb a  $h$  időtartam értéke, ezért természetes a következő definíció.

**Definíció 211** *Ha egy  $f$  valós függvény deriválható egy  $x$  pontban, akkor az  $f'(x)$  deriváltat az  $f$  függvény  $x$  pontban vett sebességének is fogjuk nevezni. Ha az  $f : (a, b) \rightarrow \mathbb{R}$  függvénynek van  $f' : (a, b) \rightarrow \mathbb{R}$  deriváltja, akkor azt az  $f$  függvény sebességének (sebességfüggvényének) is nevezzük.*

Vegyük észre, hogy ez a definíció semmi mást nem mond, csak azt, hogy a differenciálhányadost (deriváltat) némelykor sebességnek is nevezhetjük. Ha egy függvény valami olyan folyamatot ír le, aminek a "sebessége" lényeges tartalmat hordoz a számunkra, akkor célszerű ez a szószaporítás.

Nem szabad elfelejtenünk, hogy a sebesség a *függvény* "megváltozására jellemző" valami (azért nem növekedést írtunk, mert csökkenő is lehet a változás), de az a szóhasználat is elfogadható, hogy az "anyagi pont mozgásának a sebességéről" beszélünk, mert emögött az van, hogy a "mozgás" voltaképpen az azt leíró "függvénnyel" azonosított.

A további példákban más, a közgazdaságtanban használatos néhány elnevezéssel is találkozunk. Ezeknek a lényegi tartalma is csak az itt taglalt sebesség-konceptió, de az elnevezések — a tartalomhoz illeszkedve — különbözőek.

Tovább növelhetjük a deriváltnak és a függvény által leírt változás "sebességének" a kapcsolatát, ha az érintőapproximáció tételét nézzük. Az

$$f(x+h) - f(x) = f'(x)h + o(h)$$

egyenlőségből azt olvashatjuk ki, hogy a függvény  $f(x+h) - f(x)$  megváltozása az  $x$  pont megfelelő környezetében *közel arányos* az  $f'(x)$  deriválttal, hiszen a  $o(h)$  hibatag megfelelően kicsi.

Más szavakkal: Az  $x$  pontnál az argumentum egységnyi megváltozása mellett hozzávetőlegesen  $f'(x)$  “egységnyi” a függvény változása. Ezek alapján jogos azt mondani, hogy az  $f'(x)$  érték az  $f$  függvény vagy a függvény által leírt folyamat sebessége az  $x$  helyen, hiszen közel ezzel arányosan változik.

Ha egy szaktudomány hagyományai vagy érdekei úgy kívánják, akkor egyéb neveket is adnak a sebességnek. A közgazdaságtudományban például gyakori a “határ-” (marginal) elnevezés. Nem szabad azonban elfelejtenünk, hogy ezek csak a szaktudomány nyelvének a tartalomhoz való jobb illeszkedésére vagy a nyelv élénkítésére szolgálnak, és csak a derivált (differenciálhányados) szó szinonimái.

#### 5.1.4 Relatív sebesség, elaszticitás

A sebességfogalommal kapcsolatban egy lényeges észrevételt kell tennünk. A sebesség az (5.11) alatt bevezetett

$$\frac{\Delta s}{\Delta t} = \frac{s(t+h) - s(t)}{(t+h) - t} \quad (5.12)$$

átlagsebességnek a határértéke a  $h = 0$  helyen. Ha az  $s(t)$  utat mondjuk méterben, az időt pedig másodpercben (secundum) mérjük, akkor a 5.12 átlagsebesség mértékegysége a két mértékegység hányadosa: m/sec. Ennek megfelelően az  $s'(t)$  sebesség mértékegysége is m/sec. Ha méter helyett centiméterre térünk át, akkor az 5.12 számlálója százszorosára nő, és ennek megfelelően százszoros lesz a sebességet adó szám is, ami nem zavaró, mert az ennek megfelelő cm/sec mértékegység egyértelművé teszi a szám tartalmát.

Ha az  $s$  függvényt ábrázolnánk, akkor ez előző mértékegység-változás azzal jár, hogy “meredekebb” lesz a gráf akkor, ha az  $y$  tengelyen a méterről a centiméterre térünk át, pedig az  $s$  függvényünk bizonyos értelemben azonos. A mértékegységek megválasztása nyilvánvalóan nem befolyásolja a fizikai jelenség lényegét, és a fizikai mértékegységrendszer következetes használata mindig pontos információt ad.

A közgazdaságtan természetesen átveszi a fizikai és egyéb mértékegységeket, amikor például a termelt anyag mennyiségéről van szó. Specifikusan közgazdasági mértékegységek azonban nemigen mondhatók. Ezt elvileg is nehéz lenne elképzelni, hiszen milyen egységben tudnánk mérni például a hasznosságot, vagy milyen egységes pénzegység lenne elfogadható, amikor nem is létezik szilárd, mozdulatlan egység az értékelésre? A probléma egyik megoldási lehetőségét jelenti a következőkben bevezetésre kerülő fogalom.

Ha egy függvénynek az értékei valamilyen mértékegységben értendők, és úgy szeretnénk az  $x$  helyen vett megváltozását mérni, hogy az ne függjön a mértékegységtől, a differenciahányados (átlagsebesség) helyett más hányadost kell bevezetnünk. Az  $x$  független változó relatív megváltozása az  $x$  helyen

$$\frac{\Delta x}{x} = \frac{(x+h) - x}{x},$$

a függvény relatív megváltozása pedig

$$\frac{\Delta f}{f(x)} \doteq \frac{f(x+h) - f(x)}{f(x)}.$$

Ezeknek a hányadosoknak a felírásánál persze fel kell tételeznünk, hogy  $x \neq 0$  és  $f(x) \neq 0$ . Ezek a relatív számok persze függetlenek az  $x$  és  $f(x)$  mennyiségekre használt mértékegységek megválasztásától. Ezeknek a relatív megváltozásoknak a

$$h \mapsto \frac{\frac{\Delta f}{f(x)}}{\frac{\Delta x}{x}} = \frac{\frac{f(x+h) - f(x)}{f(x)}}{\frac{(x+h) - x}{x}} \quad (5.13)$$

hányadosát *relatív differenciahányadosnak* (relatív átlagsebességnek) nevezhetjük. A relatív differenciahányados egyszerűbb alakra is hozható:

$$\frac{\frac{\Delta f}{f(x)}}{\frac{\Delta x}{x}} = \frac{\frac{f(x+h) - f(x)}{f(x)}}{\frac{h}{x}} = \frac{x}{f(x)} \frac{f(x+h) - f(x)}{h}.$$

Ha az  $x \neq 0$  és  $f(x) \neq 0$  feltételek teljesülnek, és az  $f$  függvény deriválható az  $x$  helyen, akkor azonnal meg is határozhatjuk a relatív differenciahányados határértékét:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\frac{\Delta f}{f(x)}}{\frac{\Delta x}{x}} &= \lim_{h \rightarrow 0} \frac{x}{f(x)} \frac{f(x+h) - f(x)}{h} = \\ &= \frac{x}{f(x)} \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} = \frac{x}{f(x)} f'(x). \end{aligned}$$

Az előző megfontolások eredményére támaszkodva egy új elnevezést vezetünk be:

**Definíció 212** Legyen az  $f$  függvény az  $x \neq 0$  pont egy környezetében definiálva és  $f(x) \neq 0$ . Ha az  $f$  függvény deriválható az  $x$  helyen, akkor az (5.13) relatív differenciahányadosnak a  $h = 0$  helyen vett határértékét az  $f$  függvény  $x$  helyen vett elaszticitásának (rugalmasságának) nevezzük. Ennek értéke

$$\frac{x}{f(x)} f'(x).$$

**Példa 5.11** Kereslet-elaszticitás (elasticity of demand)

Valamilyen adott jószág esetében jelölje  $d(p)$  azt, hogy  $p$  egységár mellett mekkora a szóban forgó jószágra vonatkozó kereslet (keresett mennyiség). A  $d$  keresleti függvény

$$\frac{p}{d(p)} d'(p)$$

elaszticitását a kereslet elaszticitásának (rugalmasságának) szokásos nevezni.

Ez a mérőszám azt mutatja, hogy kis árváltozások mellett hogyan változik — a jószág mennyiségének a mérésére és az ár mérésére szolgáló egységek megválasztásától függetlenül — a jószágra vonatkozó kereslet (százalékos változás). Azt például sejthetjük, hogy a kereslet elaszticitása többnyire negatív, hiszen nagyobb ár mellett általában csökken a piaci kereslet.  $\square$

**Példa 5.12** *Vizsgáljuk meg az  $x \mapsto x^3$  ( $x > 0$ ) hatványfüggvény elaszticitását, és nézzük meg, hogy az  $x \mapsto (ax + b)$ ,  $a, b > 0$  lineáris függvénynek milyen határok között változik az elaszticitása.*

A példában szereplő függvények deriváltjait különböző példákban már kiszámoltuk. Nézzük először az  $x \mapsto x^3$  függvény elaszticitását, felhasználva, hogy a deriváltja  $x \mapsto 3x^2$ .

Az elaszticitás formulája szerint

$$\frac{x}{x^3}(x^3)' = \frac{x}{x^3}3x^{3-1} = 3,$$

tehát az elaszticitás pontosan a fokszaot adja. (Gondoljuk végig ugyanezt általában az  $x \mapsto x^n$  hatványfüggvényre!)

Most pedig nézzük a lineáris függvényt: ennek deriváltja  $a$ , ezért az elaszticitása

$$\frac{ax}{ax + b},$$

ami a következőképpen alakítható:

$$\frac{ax}{ax + b} = \frac{ax + b - b}{ax + b} = 1 - \frac{b}{ax + b}.$$

Mivel az  $b/(ax + b)$  az  $x = 0$  helyen 1, ugyanakkor a végtelenben a limesze 0, ezért a 0 közelében az elaszticitás közel nulla, a végtelenben pedig monoton növekedve tart az egyhez.  $\square$

## 5.2 Differenciálás, kalkulus

Ebben a pontban a differenciálhányados kiszámítására szolgáló rendszert, egy ún. "kalkulust" alakítunk ki. Voltaképpen innen ered az, hogy a bevezető analízist tárgyaló könyveket "differenciálszámítás" néven is szokták emlegetni.

### 5.2.1 Formális szabályok

A következő tételben felsoroljuk az összes formális szabályt.

**Állítás 213 (A differenciálás formális szabályai)**

1. Legyen az  $f$  deriválható az  $x$  pontban és az  $\alpha$  tetszőleges valós szám. Ekkor az  $\alpha f$  függvény is deriválható az  $x$  pontban, és

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

2. Legyen az  $f$  és  $g$  függvény differenciálható az  $x$  pontban. Ekkor az  $(f + g)$  függvény is differenciálható az  $x$  pontban, és

$$(f + g)'(x) = f'(x) + g'(x).$$

3. Legyen az  $f$  és  $g$  differenciálható az  $x$  pontban. Ekkor az  $fg$  függvény is deriválható az  $x$  pontban, és

$$(fg)'(x) = f'(x)g(x) + f(x)g'(x).$$

4. Ha az  $f$  és  $g$  függvények deriválhatóak az  $x$  pontban és  $g(x) \neq 0$ , akkor az  $f/g$  hányadosfüggvény is deriválható az  $x$  pontban, és

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

5. Legyen az  $f$  függvény deriválható az  $x$  pontban, a  $g$  függvény pedig deriválható az  $f(x)$  pontban. Ekkor az  $g \circ f$  kompozíció- (közvetett) függvény is deriválható az  $x$  pontban, és

$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x).$$

6. Tegyük fel, hogy

- (a) az  $f$  függvény deriválható az  $x$  pontban,
- (b)  $f'(x) \neq 0$ ,
- (c) az  $f$  szigorúan monoton és folytonos az  $x$  pont egy környezetében és
- (d) az  $f^{-1}$  inverz függvény folytonos az  $f(x)$  pontban.

Ekkor az  $f^{-1}$  inverz függvény deriválható az  $f(x)$  pontban, és

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

**Bizonyítás.** A bizonyítások általános vázlata: Vesszük a megfelelő különbségi hányadost és alkalmas algebrai rendezéssel olyan alakra hozzuk, amelyből határátmenettel már adódik a szóbanforgó szabály.

**1.** Számszoros deriválási szabálya. Az  $\alpha f$  függvény differenciahányadosa az  $x$  helyen

$$\frac{\alpha f(x+h) - \alpha f(x)}{h} = \alpha \frac{f(x+h) - f(x)}{h}$$

módon írható, amiből — mivel egy függvény számszorosának a limesze a limesz számszorosa — azonnal kapjuk, hogy a jobboldal limesze a  $h = 0$  helyen  $\alpha f'(x)$ . Emiatt a baloldalnak is van limesze, ami a definíció szerint  $(\alpha f)'(x)$ .

**2. Összeg deriválása.** Véve az  $(f + g)$  leképezés differenciahányadosát, egyszerű átalakítással adódik, hogy

$$\begin{aligned} & \frac{(f(x+h) + g(x+h)) - (f(x) + g(x))}{h} = \\ &= \frac{f(x+h) - f(x)}{h} + \frac{g(x+h) - g(x)}{h}, \end{aligned}$$

amiből, mivel összeg határértéke az összeadandók határértékeinek az összege, azonnal kapjuk, hogy a jobboldal határértéke az  $f'(x) + g'(x)$ , és így a baloldalnak is van határértéke, ami a definíció szerint  $(f + g)'(x)$ .

**3. Szorzat deriváltja.** A kiindulás itt is az, hogy az  $fg$  leképezés differenciahányadosát úgy alakítjuk, hogy az  $f$  és  $g$  leképezések differenciahányadosai jöjjenek be:

$$\begin{aligned} & \frac{f(x+h)g(x+h) - f(x)g(x)}{h} = \\ &= \frac{f(x+h)g(x+h) - f(x+h)g(x) + f(x+h)g(x) - f(x)g(x)}{h} = \\ &= f(x+h) \frac{g(x+h) - g(x)}{h} + g(x) \frac{f(x+h) - f(x)}{h}. \end{aligned}$$

A jobboldalon az  $f(x+h)$  limesze  $h \rightarrow 0$  mellett  $f(x)$ , mivel az  $f$  függvény deriválható, tehát folytonos is az  $x$  helyen (202. állítás). Alkalmazva a határérték formális tulajdonságait, az adódik, hogy a jobboldal limesze  $f(x)g'(x) + f'(x)g(x)$ , ami megegyezik a baloldal — emiatt létező — limeszével, ami a definíció szerint  $(fg)'(x)$ .

**4. Tört deriváltja.** Először belátjuk, hogy az  $1/g$  leképezés deriválható az  $x$  helyen és

$$\left(\frac{1}{g}\right)'(x) = -\frac{g'(x)}{g^2(x)}. \quad (5.14)$$

Mivel a deriválhatóság miatt a  $g$  folytonos az  $x$ -ben (202. állítás) és  $g(x) \neq 0$ , ezért  $g(x+h)$  nem nulla, ha a  $h$  elégé kicsi. Ennek a szem előtt tartásával minden elégé kicsi  $h$  mellett felírhatjuk az  $1/g$  differenciahányadosát, és megengedett az alábbi átalakítás:

$$\begin{aligned} & \frac{(1/g)(x+h) - (1/g)(x)}{h} = \frac{1}{h} \left( \frac{1}{g(x+h)} - \frac{1}{g(x)} \right) = \\ &= \frac{g(x) - g(x+h)}{hg(x+h)g(x)} = \frac{g(x) - g(x+h)}{h} \cdot \frac{1}{g(x+h)g(x)}. \end{aligned}$$

A jobboldal limesze a határérték formális szabályai szerint számolva  $-g'(x)/g^2(x)$ , ezért a baloldalnak is létezik a limesze, ami a definíció szerint  $(1/g)'(x)$ .

Rátérve a hányados deriválási szabályának a bizonyítására, a szorzat differenciálására már igazolt szabály és (5.14) felhasználásával azt kapjuk, hogy

$$\begin{aligned} \left(\frac{f}{g}\right)'(x) &= f'(x)\frac{1}{g}(x) + f(x)\left(\frac{1}{g}\right)'(x) = \\ &= \frac{f'(x)}{g(x)} - f(x)\frac{g'(x)}{g^2(x)} = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}. \end{aligned}$$

**5. Közvetett függvény deriválása.** Legyenek  $f$  és  $g$  valós változós valós értékű függvények, továbbá  $f$  legyen deriválható az  $x \in \mathbb{R}$  pontban és  $g$  deriválható  $f(x)$ -ben. Megmutatjuk, hogy a  $g \circ f$  függvény is deriválható  $x$ -ben és

$$(g \circ f)'(x) = g'(f(x)) \cdot f'(x).$$

A feltételek alapján léteznek olyan  $p$  és  $r$  a 0 egy környezetében értelmezett és a 0-ban kisrendű függvények, hogy bármely, a mondott környezetbeli  $u$ -ra és  $v$ -re

$$f(x+u) - f(x) = f'(x) \cdot u + p(u)$$

illetve

$$g(f(x)+v) - g(f(x)) = g'(f(x)) \cdot v + r(v).$$

Így az  $f$   $x$ -beli folytonossága miatt a 0-nak egy új alkalmas környezetébe eső  $u$  értékekre  $v = f(x+u) - f(x)$  választással  $v$ -re a fentiek teljesülnek, továbbá  $f(x+u) = f(x) + v$ , ezért

$$\begin{aligned} (g \circ f)(x+u) - (g \circ f)(x) &= g(f(x)+v) - g(f(x)) = \\ &= g'(f(x)) \cdot v + r(v) = g'(f(x)) \cdot (f(x+u) - f(x)) + r(f(x+u) - f(x)) = \\ &= g'(f(x)) \cdot (f'(x) \cdot u + p(u)) + r(f(x+u) - f(x)) = \\ &= g'(f(x)) \cdot f'(x) \cdot u + g'(f(x)) \cdot p(u) + r(f(x+u) - f(x)). \end{aligned}$$

Tekintsük a 0-nak a fenti mondott környezetében az alábbi módon megadott  $S$  függvényt:

$$S(u) = g'(f(x)) \cdot p(u) + r(f(x+u) - f(x)).$$

Ekkor nyilván az illető környezetben

$$(g \circ f)(x+u) - (g \circ f)(x) = g'(f(x)) \cdot f'(x) \cdot u + S(u).$$

Ezek szerint a tétel bizonyításához elegendő megmutatnunk, hogy az  $S$  függvény kisrendű a 0-ban.

Mivel  $p$  kisrendű, ezért egy számszorosa is az (ami az  $S$  definíciójában szereplő első tag), tehát elegendő megmutatnunk, hogy az  $u \mapsto r(f(x+u) - f(x))$  leképezés kisrendű a 0-ban. Nyilván a 0-beli értéke 0.

Az  $r$  kisrendűsége miatt  $r(0) = 0$  és

$$\lim_{y \rightarrow 0} \frac{r(y)}{y} = 0,$$

ami a határérték definíciója alapján azt jelenti, hogy van olyan  $q$  függvény, mely folytonos a 0-ban és  $q(0) = 0$ , továbbá a 0 egy környezetében minden  $y$ -ra

$$r(y) = y \cdot q(y).$$

Ezért  $u \neq 0$  esetén

$$\begin{aligned} r(f(x+u) - f(x)) &= (f(x+u) - f(x)) \cdot q(f(x+u) - f(x)) = \\ &= (f'(x) \cdot u + p(u)) \cdot q(f(x+u) - f(x)) = \\ &= u \cdot (f'(x) + \frac{p(u)}{u}) \cdot q(f(x+u) - f(x)), \end{aligned}$$

így

$$\frac{r(f(x+u) - f(x))}{u} = (f'(x) + \frac{p(u)}{u}) \cdot q(f(x+u) - f(x)),$$

ami  $u \rightarrow 0$  esetén tart a 0-hoz, hiszen az  $u \mapsto f(x+u) - f(x)$  leképezés és a  $q$  függvény folytonos a 0-ban, tehát a kompozíciójuk is folytonos, továbbá  $\frac{p(u)}{u}$  tart a 0-hoz. Ezzel megmutattuk, hogy a  $u \mapsto r(f(x+u) - f(x))$  leképezés kisrendű a 0-ban. Ezért  $S$  is kisrendű, és éppen ezt akartuk igazolni.

**6.** Az inverz függvény deriválása. A feltevések miatt van az  $x$ -nek egy  $B(x, r_1)$  környezete, melynek  $f$ -képe tartalmazza az  $f(x)$ -nek egy  $B(f(x), r_2)$  környezetét. Legyen a  $h$  a továbbiakban olyan kicsi, hogy

$$f(x) + h \in B(f(x), r_2).$$

Ekkor egyértelműen van olyan  $y \in B(x, r_1)$ , hogy  $f(x) + h = f(y)$ . Így az  $f^{-1}$  függvény  $f(x)$  pontbeli differenciáhányadosa a következőképpen alakítható:

$$\begin{aligned} \frac{f^{-1}(f(x) + h) - f^{-1}(f(x))}{h} &= \frac{f^{-1}(f(y)) - f^{-1}(f(x))}{h} = \\ &= \frac{y - x}{f(y) - f(x)} = \frac{1}{\frac{f(y) - f(x)}{y - x}}. \end{aligned} \quad (5.15)$$

Az utolsó lépésben használtuk ki azt, hogy  $f'(x) \neq 0$  miatt az  $f$   $x$ -beli differenciáhányadosa sem nulla, ha az  $|y - x|$  elég kicsi, ezért írhatjuk nevezőbe. A feltételezés szerint az  $f^{-1}$  folytonos az  $f(x)$ -ben, ezért az  $y = f^{-1}(f(x) + h)$  érték tart az  $f^{-1}(f(x)) = x$  számhoz, ha a  $h$  tart a 0-hoz. Így a 5.15 egyenlőtlenség jobboldala tart az  $1/f'(x)$  számhoz, az emiatt konvergens baloldal pedig definíció szerint tart az  $(f^{-1})'(f(x))$ -hez.  $\square$

### 5.2.2 Speciális függvények differenciálása

Gondolatmenetünk szerint most az következik, hogy meghatározzuk azoknak a függvényeknek a deriváltjait, amelyekből — mint alapelemekből, a formális szabályok segítségével — a vizsgálataink tárgyát képező függvényeket elő tudjuk állítani. A deriválási szabályokat azonban — praktikus szempontok miatt — a függvények bővebb körére készítjük el, de ezzel persze csak azt érhetjük el, hogy bizonyos fontos függvényeknél nem kell mindig az említett négy alapfüggvényhez visszamennünk.

#### Állítás 214 (Speciális függvények deriváltjai)

1. Az  $x \mapsto c$  ( $c \in \mathbb{R}$  adott) konstans függvény deriváltja minden  $x \in \mathbb{R}$  pontban nulla.
2. Az  $x \mapsto x^n$  ( $1 \leq n$  egész) hatványfüggvény az  $\mathbb{R}$  minden  $x$  pontjában deriválható és deriváltja  $nx^{n-1}$ .

Az  $x \mapsto x^n$  ( $n \leq -1$  egész) hatványfüggvény az  $\mathbb{R}$  minden  $x \neq 0$  pontjában deriválható, és a deriváltja  $nx^{n-1}$ .

3. A  $\sin$  és  $\cos$  trigonometrikus függvények minden  $x \in \mathbb{R}$  pontban deriválhatók, és

$$\sin' x = \cos x \quad \text{és} \quad \cos' x = -\sin x.$$

4. A tangensfüggvény az értelmezési tartománya minden pontjában deriválható, és ekkor

$$\tan' x = \frac{1}{\cos^2 x}.$$

5. A trigonometrikus függvények inverzeinek az  $x$  helyen vett deriváltjai:

$$\arctan' x = \frac{1}{1+x^2}, \quad \text{ha } x \in \mathbb{R}$$

és

$$\arcsin' x = \frac{1}{\sqrt{1-x^2}}, \quad \text{ha } |x| < 1.$$

Bizonyítás. (1) : A konstans függvény esetében a különbségi hányados nulla.

(2): Legyen először  $n$  pozitív. Az állítást  $n$ -re vonatkozó teljes indukcióval igazoljuk a szorzatfüggvény deriválási szabálya alapján.  $n = 1$ -re az állítás ismert. Most tegyük fel, hogy  $n$ -re igaz, azaz hogy az  $x \mapsto x^n$  függvény deriváltja az  $x \mapsto n \cdot x^{n-1}$  függvény, és lássuk be az állítást  $n + 1$ -re:

$$\begin{aligned} (x^{n+1})' &= (x \cdot x^n)' = x' \cdot x^n + x \cdot (x^n)' = \\ &= 1 \cdot x^n + x \cdot n \cdot x^{n-1} = (n+1) \cdot x^n. \end{aligned}$$

Ha  $n$  negatív, akkor a fentiek és a hányadosfüggvény deriválási szabálya alapján azonnal látható az állítás.

(3): Ismert trigonometrikus azonosság alapján

$$\begin{aligned}\lim_{h \rightarrow 0} \frac{\sin(x+h) - \sin x}{h} &= \lim_{h \rightarrow 0} \frac{2 \cos\left(\frac{2x+h}{2}\right) \sin\left(\frac{h}{2}\right)}{h} = \\ &= \lim_{h \rightarrow 0} \cos\left(\frac{2x+h}{2}\right) \cdot \lim_{h \rightarrow 0} \frac{\sin\left(\frac{h}{2}\right)}{\frac{h}{2}} = \cos x \cdot 1.\end{aligned}$$

Az utolsó lépésben a már bebizonyított  $\frac{\sin h}{h} \rightarrow 1$  ( $h \rightarrow 0$ ) határértéket és a koszinuszfüggvény folytonosságát használtuk.

A cos függvény deriváltjának a meghatározásához a

$$\cos x = \sin\left(\frac{\pi}{2} - x\right)$$

azonosságot és a közvetett függvény deriválási szabályát használjuk fel:

$$\begin{aligned}\cos' x &= \frac{d}{dx} \sin(\pi/2 - x) = \sin'(\pi/2 - x) \cdot (\pi/2 - x)' = \\ &= \cos(\pi/2 - x) \cdot (-1) = -\sin x.\end{aligned}$$

(4): A  $\tan = \frac{\sin}{\cos}$  összefüggés alapján, a tört deriválására vonatkozó szabály alkalmazásával adódik, hogy

$$\begin{aligned}\tan' x &= \frac{d}{dx} \left( \frac{\sin x}{\cos x} \right) = \frac{\sin' x \cos x - \sin x \cos' x}{\cos^2 x} = \\ &= \frac{\cos x \cos x - \sin x(-\sin x)}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x},\end{aligned}$$

feltéve persze, hogy  $\cos x \neq 0$ , amit az  $x \neq (2n+1)\frac{\pi}{2}$  feltétel biztosít.

(5): A bizonyítások természetesen az inverz függvény deriválására vonatkozó formális szabályt használják, valamint azt, hogy a szóban forgó függvényeknek az adott intervallumokon van inverzük és folytonosak (125. állítás).

Legyen először  $y = \arcsin x$  azaz  $x = \sin y$ . Ekkor a sin függvény deriváltjának az ismeretében, a  $\sin^2 x + \cos^2 x = 1$  azonosság felhasználásával számolunk:

$$\begin{aligned}\frac{d}{dx} \arcsin x &= \frac{1}{\cos y} \Big|_{y=\arcsin x} = \frac{1}{\sqrt{1-\sin^2 y}} \Big|_{y=\arcsin x} = \\ &= \frac{1}{\sqrt{1-\sin^2(\arcsin x)}} = \frac{1}{\sqrt{1-x^2}}, \quad |x| < 1.\end{aligned}$$

Legyen  $y = \arctan x$  azaz  $x = \tan y$ . Ekkor, hasonlóan járva el, mint az előbb:

$$\begin{aligned} \frac{d}{dx} \arctan x &= \frac{1}{\tan' y} \Big|_{y=\arctan x} = \cos^2 y \Big|_{y=\arctan x} = \\ &= \frac{1}{1 + \tan^2 y} \Big|_{y=\arctan x} = \frac{1}{1 + \tan^2(\arctan x)} = \frac{1}{1 + x^2}. \end{aligned}$$

Ezzel az összes szabályt beláttuk.  $\square$

Az exponenciális függvény differenciálhatóságának vizsgálatához szükségünk lesz a konvexitás és a differenciálhatóság összefüggéseinek ismeretére, így erre később térünk vissza.

### 5.2.3 A szélsőérték szükséges feltételei

Az előzőekben már megismertük egy  $f : A \rightarrow \mathbb{R}$  ( $A \subseteq \mathbb{R}$ ) függvény (globális) maximumának és minimumának a fogalmát; összefoglaló néven: a (globális) szélsőérték- (extrémum-) fogalmat. Most ezeknek a lokális megfelelőit fogjuk bevezetni.

**Definíció 215** Az  $f : A \rightarrow \mathbb{R}$  függvénynek lokális minimuma van az  $A$  halmaz egy  $b$  pontjában, ha

$$f(b) \leq f(x) \quad (5.16)$$

$a$   $b$  valamilyen ( $A$ -beli) környezetének minden  $x$  elemére. Ha az (5.16) egyenlőtlenség szigorúan teljesül  $a$   $b$  pont valamilyen hiányos környezetében, akkor szigorú lokális minimumról beszélünk. Az  $f(b)$  helyettesítési értéket a lokális minimum értékének mondjuk.

Szavakban röviden:  $A$   $b$  pontosan akkor (szigorú) lokális minimumhelye az  $f$  függvénynek, ha a  $b$  valamilyen környezetében (szigorú) globális minimumhelye. Teljesen hasonló a lokális maximum definíciója. A lokális minimumot és maximumot közös néven lokális szélsőértékeknek mondjuk.

A globális szélsőérték nyilvánvalóan lokális is, de megfordítva nem igaz. A lokális szélsőérték általában csak valamilyen környezetben globális.

A globális minimum nyilvánvalóan egyetlen (ha van), de több helyen is felveheti a függvény. Lokális szélsőérték viszont több is lehet különböző helyeken. Szigorú minimumot csak egyetlen helyen vehet fel a függvény, és szigorú lokális minimum esetében a minimumhely alkalmas környezetében csak a lokális minimum helyén veszi fel a minimumát a függvény. Azonosak mondhatók maximum esetében.

Általában nem könnyű eldönteni azt, hogy egy függvénynek van-e, s ha igen, hol van szélsőértéke.

Egy rendkívül fontos tételt már láttunk a szélsőértékekkel kapcsolatban: korlátos és zárt halmazon értelmezett folytonos valós függvény mindig felveszi mind a minimumát mind a maximumát (131. tétel).

Ez a tétel azonban csak a szélsőértékek létezését biztosítja, de nem ad módszert a meghatározásukra. A jelenlegi pont egyik fő célja az, hogy deriválható függvények esetén módszert adjunk a szélsőértékek kiszámítására.

Ha az eddig rajzolt ábráinkat nézzük, akkor azt vehetjük észre, hogy az értelmezési tartomány olyan belső pontjaiban, ahol lokális szélsőérték van, az érintő párhuzamosnak látszik az  $x$  tengellyel, azaz az irántangense nulla. A következő igen egyszerű tételben ezt az észrevételt fogalmazzuk meg, ami utat nyit a deriválható függvények szélsőértékeinek a meghatározásához.

**Állítás 216 (Szélsőérték, szükséges feltétel)** *Ha egy  $f$  leképezésnek lokális szélsőértéke van egy olyan  $b$  pontban, ahol a függvény deriválható, akkor ott az  $f'(b)$  derivált nulla.*

Emlékeztetünk rá, hogy a deriváltat csak az értelmezési tartomány belső pontjában definiáltuk, ezért a tételben szereplő  $b$  is belső pontja az  $A$  értelmezési tartománynak.

**Bizonyítás.** Lássuk be az állítást a lokális maximum esetére, a minimum esete hasonlóan tárgyalható.

Ha a  $b$  pontban differenciálható az  $f$  függvény, és a  $b$  pontban lokális maximuma van, akkor minden eléggé kicsi abszolútértékű  $h$  számra  $f(b+h) \leq f(b)$ , azaz átrendezve  $f(b+h) - f(b) \leq 0$ . Ebből az egyenlőtlenségből pedig adódik az, hogy

$$\frac{f(b+h) - f(b)}{h} \leq 0, \quad \text{ha} \quad 0 < h \quad (5.17)$$

és

$$\frac{f(b+h) - f(b)}{h} \geq 0, \quad \text{ha} \quad h < 0, \quad (5.18)$$

feltéve, hogy a  $h$  abszolút értéke eléggé kicsi (pontosan: valamilyen  $\delta$  pozitív számra  $|h| \leq \delta$ ). Ezek szerint az  $f$  függvény különbségi hányadosa a nulla környezetében balra nemnegatív, jobbra pedig nempozitív. Ennek a különbségi hányadosnak a feltevés szerint létező  $f'(b)$  limesze a differenciáhányados.

Ezek szerint  $f'(b)$ -nek minden környezetében van nempozitív és nemnegatív érték is. Ezért  $f'(b)$  csak 0 lehet.  $\square$

**Példa 5.13** *Van-e szélsőértéke az  $x \mapsto x^3$  függvénynek a nullában?*

A válasz nagyon egyszerű: Noha a  $3x^2$  derivált a nulla helyen nulla, ennek ellenére sincs szélsőértéke a függvénynek a nullában. Előtte ugyanis negatív, utána pedig pozitív, ezért a nulla érték nem lehet sem minimum sem maximum.  $\square$

Ez az egyszerű példa nagyon fontos, mert azt mutatja, hogy még ilyen “jó viselkedésű” függvény esetében sem elégséges a tétel feltétele, ezért hangsúlyozni kell: a szükséges feltétel csak a lehetőségét adja meg annak, hogy valahol szélsőérték lehessen. Csak azt mutatja meg, hogy hol kereshetjük az extrémumokat. Ahogyan azonban a következő példa is mutatja, ez a tétel a Weierstrass-tétellel (131) kombinálva már lehetőséget nyújt a szélsőértékek meghatározásához.

**Példa 5.14** *Vizsgáljuk meg a*

$$g: x \mapsto x + \frac{1}{x}, \quad x \in [\frac{1}{2}, 2]$$

*függvényt szélsőértékek szempontjából.*

Az  $[1/2, 2]$  zárt és korlátos intervallumon a szóbanforgó függvény folytonos, ezért van maximuma és minimuma (131 tétel), így van mit keresnünk. A szélsőértékek helyeinek és értékeinek a meghatározásához a következőképpen járunk el.

1) Megnézzük a  $g$  függvényt az intervallum két végpontjában:  $g(1/2) = 2.5$  és  $g(2) = 2.5$ .

2) Megvizsgáljuk a  $g$  értékeit az intervallum  $(1/2, 2)$  belsejében. Ezt persze nem tehetjük meg úgy, hogy mindenhol kiszámoljuk. Itt segít az előző tétel, ami szerint csak az olyan  $x$  helyeket kell megnézni, ahol  $g'(x) = 0$ . Mivel

$$g'(x) = 1 - \frac{1}{x^2},$$

ezért a derivált nulla helyei:  $x = 1$  és  $x = -1$ , amelyekből csak az 1 esik az értelmezési tartományba. Itt megnézve a függvény értékét azt kapjuk, hogy  $g(1) = 2$ .

Összevetve az előzőekben kapott értékeket, azt találjuk, hogy maximuma van a  $g$  függvénynek az  $x = 1/2$  és  $x = 2$  helyeken, ahol az értéke 2.5, és minimuma van az  $x = 1$  helyen, ahol az értéke 2.  $\square$

Érdemes összefoglalni a példában követett eljárást, mert nagyon általános esetekben használható:

Legyen az  $f$  függvény deriválható egy  $(a, b)$  korlátos intervallumban és folytonos az  $a$  és  $b$  végpontokban is. Mivel a deriválhatóságból következik a folytonosság, ezért az  $f$  az  $[a, b]$  zárt intervallumban folytonos.

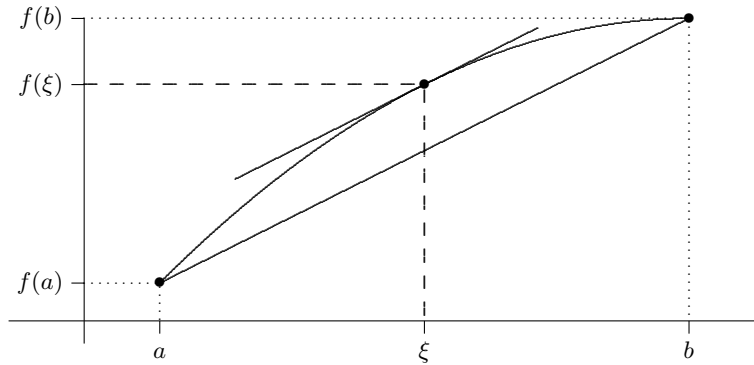
1. Észrevesszük, hogy van maximuma is és minimuma is a függvénynek az  $[a, b]$  intervallumban a Weierstrass-tétel alapján.
2. Deriváljuk az  $f$  függvényt és meghatározzuk a derivált nullhelyeit.
3. Kiszámoljuk a függvény értékeit az intervallum végpontjaiban és a derivált nullhelyeinél. Az így kiszámolt értékek között kell keresni a maximumot és a minimumot.

## 5.3 Középértéktételek

### 5.3.1 A differenciálszámítás középértéktétele

Mielőtt a jelen pont fő tételét kimondanánk, vetítsük előre a tétel szemléletes tartalmát (5.7. ábra). Ha egy  $f: [a, b] \rightarrow \mathbb{R}$  függvényről vesszük az  $(a, f(a))$  és

$(b, f(b))$  pontokon átmenő szelőt, akkor felvethetjük a kérdést, hogy van-e olyan érintője a gráfnak az intervallum belsejében, amelyik párhuzamos a szelővel. A válasz igenlő a következő tétel szerint.



5.7. ábra: A közéérték tétel geometriai tartalma.

**Állítás 217 (A deriválás közéértéktétele)** *Ha egy  $f : [a, b] \rightarrow \mathbb{R}$  leképezés deriválható az intervallum belsejében és az  $a$  és  $b$  végpontokban is folytonos, akkor van olyan  $\xi$  pont az intervallum belsejében, hogy*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}. \quad (5.19)$$

Bizonyítás. Vegyük a

$$h(x) \doteq (f(b) - f(a))x - (b - a)f(x)$$

leképezést. A  $h$  függvény a 202. állítás alapján folytonos az  $[a, b]$  zárt intervallumon és deriválható az intervallum belsejében, mivel ilyen tulajdonságokkal rendelkező függvények számszorosainak összege.

Az intervallum végpontjaiban azonos értékeket vesz fel a  $h$ , mert

$$h(a) = (f(b) - f(a))a - (b - a)f(a) = af(b) - bf(a)$$

és

$$h(b) = (f(b) - f(a))b - (b - a)f(b) = af(b) - bf(a).$$

Ha a  $h$  függvény a  $h(a) = h(b)$  állandó értéket veszi fel az egész  $[a, b]$  intervallumon, akkor az intervallum belsejében állandó, és így a deriváltja nulla.

Ha pedig a  $h$  nem állandó az egész intervallumon, akkor valahol az intervallumban nagyobb vagy kisebb mint a végpontokban felvett értékek. Ennélfogva a  $h$  leképezés a 131. tétel szerint létező maximumát vagy minimumát az intervallum

valamilyen  $\xi$  belső pontjában veszi fel, ahol a megelőző 216. tétel szerint a  $h'(\xi)$  derivált nulla. Tehát mindkét esetben van olyan  $\xi \in (a, b)$  pont, melyre

$$0 = h'(\xi) = (f(b) - f(a)) - (b - a)f'(\xi),$$

amiből

$$f'(\xi) = \frac{f(b) - f(a)}{b - a},$$

amit bizonyítani akartunk.  $\square$

A közéértéktétel viszonylagosan mély, az egész fejezet legközpontibb állítása. A “mélység” oka elsősorban az, hogy a bizonyítás magja a mélyebben fekvő 131. tétel.

**Példa 5.15** Vegyük az  $x \mapsto x^2$  függvényt, és keressük meg a  $(0, 0)$  és  $(a, a^2)$  pontokon átmenő szelővel párhuzamos érintőt a  $(0, a)$  intervallumban.

A szóbanforgó szelő iránytangense  $\frac{a^2-0}{a-0} = a$ . A deriváltfüggvény pedig az  $x \mapsto 2x$ . Van olyan  $\xi \in (0, a)$ , amelyre  $2\xi = a$ , mégpedig a  $\xi = a/2$ , tehát a szelővel párhuzamos érintő az intervallum  $a/2$  felezőpontjában érinti a görbét. A keresett érintő egyenlete könnyen fel is írható:

$$y - (a/2)^2 = 2(a/2)(x - a/2),$$

amit rendezve adódik az

$$y = ax - a^2/4$$

egyen-es-egyenlet. Rajzoljuk fel a példa eredményét.  $\square$

### 5.3.2 Magasabbrendű approximációk, Taylor-formula

Az érintőapproximáció-tételnek a tárgyalásakor megvizsgáltuk azt a kérdést, hogy miként lehet egy függvényt lokálisan (kicsiben) egy egyenessel, az érintővel megfelelően közelíteni. Ha egy  $f$  függvény deriválható egy  $b$  pontban, akkor a  $b$  pont valamilyen környezetében felvett  $f(x)$  függvényértékeket jól közelíti az

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

egyenletű egyenes abban az értelemben, hogy az eltérés kis ordó nagyságrendű, vagyis

$$f(x) = f(b) + f'(b)(x - b) + o(x - b). \quad (5.20)$$

Kétszer deriválható  $f$  függvény esetében a 5.20 egyenlőségben a  $o(x - b)$  kis ordó tagról többet is mondhatunk.

E célból élesítsük a feladatot. Feltesszük, hogy az  $f$  kétszer deriválható a  $b$  pontban. Ebből következik, hogy van olyan környezete a  $b$  pontnak, ahol az  $f'$  létezik. Csak a jobboldali felét nézve a környezetnek, van olyan  $c > b$  szám, hogy

az  $f'$  létezik az  $[b, c]$  zárt intervallumon. Feltesszük még, hogy az  $f'$  folytonos ezen az intervallumon. Ezekből már következik, hogy az  $f$  is folytonos a  $[b, c]$  intervallumon, hiszen deriválható a végpontokban is. Legyen az  $x$  egy tetszőleges, de rögzített eleme az  $(b, c)$  intervallumnak, ahol a  $o(b-x)$  tagot kiszámítjuk.

A  $o(x-b)$  leképezést

$$M \cdot (x-b)^2$$

alakban keressük, ahol az  $M$  egy meghatározandó szám, ami feltehetőleg függ az  $b$  és  $x$  számoktól és az  $f$  függvénytől. Eszerint egy

$$f(x) = f(b) + f'(b)(x-b) + M(x-b)^2 \quad (5.21)$$

formájú előállítást szeretnénk találni.

Természetesen semmi sem biztosítja előre, hogy találunk ilyen alakú maradéktagot, de ha találunk, akkor megoldottuk a feladatot. Ez a “fogás” gyakori a matematikában.

Vegyük a

$$g(y) \stackrel{\text{def}}{=} f(y) + f'(y)(x-y) + M(x-y)^2 \quad (5.22)$$

függvényt, amelyik a feltevések szerint folytonos az  $[b, x]$  intervallumon, és belül deriválható, ezért a középértéktétel szerint van olyan  $\xi$  szám az  $(b, x)$  intervallumban, amelyre

$$\frac{g(x) - g(b)}{x - b} = g'(\xi). \quad (5.23)$$

Már nincs más feladatunk, csak ki kell számolnunk ebben az egyenlőségben a tagokat. (5.21) és (5.22) szerint

$$g(b) = f(b) + f'(b)(x-b) + M(x-b)^2 = f(x),$$

és mivel (5.22) alapján nyilvánvalóan  $g(x) = f(x)$ , ezért a (5.23) bal oldala nulla, és így  $g'(\xi) = 0$ .

A  $g'$  kiszámolásával

$$g'(y) = f'(y) + f''(y)(x-y) - f'(y) - 2M(x-y) = f''(y)(x-y) - 2M(x-y)$$

adódik, amiből a  $g'(\xi) = 0$  alapján azonnal kapjuk, hogy

$$M = \frac{f''(\xi)}{2}. \quad (5.24)$$

Összefoglalva az eddigieket: van olyan  $b < \xi < x$  szám, amelyre

$$f(x) = f(b) + f'(b)(x-b) + \frac{f''(\xi)}{2}(x-b)^2. \quad (5.25)$$

Az előbbi gondolatmenetet természetesen egy alkalmas  $[c, b]$ ,  $c < b$  intervallumra is értelemszerűen végrehajthattuk volna. Az eredményt — maradv a  $b$  jobboldali környezetében — egy állításban is rögzítjük:

**Állítás 218** Legyen az  $f$  függvény deriválható a  $[b, c]$  ( $b < c$ ) intervallumon, az  $f'$  derivált függvény folytonos  $[b, c]$ -n és differenciálható a  $(b, c)$  intervallumban.

Ekkor tetszőleges  $x \in (b, c)$  számhoz van olyan  $\xi \in (b, x)$  szám, hogy

$$f(x) = f(b) + f'(b)(x - b) + \frac{f''(\xi)}{2}(x - b)^2.$$

Az

$$x \rightarrow f(b) + f'(b)(x - b)$$

elsőfokú polinom az  $f$  függvény elsőfokú  $(-)$ -rendű, lineáris) Taylor-közelítése a  $b$  pontban.

Az eredményt kielégítőnek mondhatjuk, hiszen most már nemcsak azt tudjuk, hogy az  $f$  függvénynek a közelítésétől való eltérése  $(x - b)$ -vel osztva nullához tart. hanem azt is, hogy az eltérés az  $(x - b)$  négyzetével arányos. Lássunk most alkalmazásként egy példát.

**Példa 5.16** Adjuk meg a  $\sin$  függvény elsőrendű közelítését a 0 helyen. Milyen pontosságú kiszámolását teszi ez lehetővé a szinuszfüggvénynek az 0.1 helyen?

A szinuszfüggvény első két deriváltja:

$$\sin' y = \cos y \quad \text{és} \quad \sin'' y = -\sin y,$$

ezért a 0-ban vett lineáris közelítésre

$$\sin x = \sin 0 + \sin' 0(x - 0) + \frac{1}{2} \sin'' \xi(x - 0)^2 = x - \frac{\sin \xi}{2} x^2,$$

ahol  $0 < \xi < x$ .

Az  $x = 0.1$  helyen a hiba értéke így nem nagyobb, mint

$$\left| \frac{\sin \xi}{2} \right| 0.1^2 \leq \frac{1}{200},$$

tehát 0.1-nek  $\sin 0.1$ -től való eltérése kisebb, mint öt ezred. □

Az általános eset vizsgálata előtt vezessünk be néhány elnevezést.

**Definíció 219** Legyen az  $f$  függvény  $n$ -szer deriválható a  $b$  pontban. A

$$T_n(b, x) \stackrel{\text{def}}{=} f(b) + \frac{f'(b)}{1!}(x - b) + \frac{f''(b)}{2!}(x - b)^2 + \cdots + \frac{f^{(n)}(b)}{n!}(x - b)^n$$

$n$ -edfokú polinomot az  $f$  függvény  $b$  pontban vett (vagy  $b$  pont körüli vagy  $b$  pontnál lévő)  $n$ -edik  $(n$ -edfokú,  $n$ -edrendű) Taylor-polinomjának (vagy Taylor-közelítésének) fogjuk mondani. Az

$$f(x) - T_n(b, x)$$

különbséget az  $n$ -edik Taylor-féle maradéktagnak (hibatagnak) nevezzük.

Hasznos, ha egy függvényt polinommal tudunk közelíteni. Látni fogjuk, hogy Taylor-polinommal, eléggé általános feltételek mellett jól approximálhatóak lesznek a megfelelően sokszor differenciálható függvények. A közelítés pontosságát a maradéktag méri, ezért nem meglepő, hogy erre sokféle előállítást találtak. Mi csak a legegyszerűbb formát tárgyaljuk. A 218. állításban az elsőrendű problémát vizsgáltuk, most lássuk az általános esetet:

**Állítás 220** *Legyen az  $f$  függvény  $n$ -szer deriválható a  $[b, c]$  ( $b < c$ ) intervallumon, az  $f^{(n)}$   $n$ -ik deriváltfüggvény folytonos  $[b, c]$ -n és differenciálható a  $(b, c)$  intervallumban.*

*Ekkor tetszőleges  $x \in (b, c)$  számhoz van olyan  $\xi$  szám a  $(b, x)$  intervallumban, amelyre*

$$f(x) = T_n(b, x) + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-b)^{n+1}. \quad (5.26)$$

Hangsúlyozottan fontos az az eset, amikor a  $b$  szám nulla, ekkor a közelítés (Taylor-polinom):

$$f(0) + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + \dots + \frac{f^{(n)}(0)}{n!}x^n.$$

A bizonyításból majd látni lehet, hogy a  $b < c$  feltétel nem lényeges, lehetne a  $b$  baloldali környezetében is dolgozni, és ekkor persze az  $x$  is kisebb lenne a  $b$ -nél. **Bizonyítás.** A bizonyítás menete pontosan megegyezik azzal, ahogyan a 218. tételt bebizonyítottuk.

Legyen  $M$  az az  $x$ -től függő szám, amelyre

$$f(x) = \sum_{i=0}^n \frac{f^{(i)}(b)}{i!} (x-b)^i + M(x-b)^{n+1}. \quad (5.27)$$

A  $g$  függvényt a következőképpen definiáljuk:

$$g(y) \stackrel{\text{def}}{=} \sum_{i=0}^n \frac{f^{(i)}(y)}{i!} (x-y)^i + M(x-y)^{n+1}. \quad (5.28)$$

A  $g$  függvény deriválható az  $(b, x)$  intervallumban és folytonos az  $[b, x]$  zárt intervallumon, ezért alkalmazható a középértéktétel, miszerint van olyan  $\xi \in (b, x)$ , amelyre

$$\frac{g(x) - g(b)}{x - b} = g'(\xi) = \left. \frac{d}{dy} g(y) \right|_{y=\xi}. \quad (5.29)$$

A  $g$  (5.28) definíciója és az  $M$  szám megválasztását rögzítő (5.27) alapján  $g(b) = f(b)$ , és egyszerű behelyettesítéssel  $g(x) = f(x)$ . Emiatt a középérték-egyenlőség az egyszerű

$$g'(\xi) = \left. \frac{d}{dy} g(y) \right|_{y=\xi} = 0 \quad (5.30)$$

egyenlőségbe megy át. A következő formulát fogjuk felhasználni, amit a bizonyítás végén számítottunk ki.

*Fennáll a következő*

$$\begin{aligned} \frac{d}{dy} \left( f(y) + \frac{f'(y)}{1!}(x-y) + \cdots + \frac{f^{(n)}(y)}{n!}(x-y)^n \right) &= \\ &= \frac{f^{(n+1)}(y)}{n!}(x-y)^n \end{aligned} \quad (5.31)$$

*deriválási formula.*

Eszerint a  $g$  függvény deriváltja

$$g'(y) = \frac{f^{(n+1)}(y)}{n!}(x-y)^n - (n+1)M(x-y)^n,$$

és így (5.30) szerint az  $M$  a következő egyenletnek tesz eleget:

$$\frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n - (n+1)M(x-\xi)^n = 0,$$

amiből adódik, hogy

$$M = \frac{f^{(n+1)}(\xi)}{(n+1)!}, \quad (5.32)$$

amivel a tételt be is láttuk.

Most már csak a felhasznált formula igazolása van hátra, ami egyszerű számolás:

$$\begin{aligned} \frac{d}{dy} \left( f(y) + \frac{f'(y)}{1!}(x-y) + \cdots + \frac{f^{(n)}(y)}{(n-1)!}(x-y)^n \right) &= \\ &= f'(y) + \left( \frac{f''(y)}{1!}(x-y) - \frac{f'(y)}{0!} \right) + \\ &+ \left( \frac{f^{(3)}(y)}{2!}(x-y)^2 - \frac{f^{(2)}(y)}{1!}(x-y) \right) + \cdots + \\ &+ \left( \frac{f^{(n)}(y)}{(n-1)!}(x-y)^{n-1} - \frac{f^{(n-1)}(y)}{(n-2)!}(x-y)^{n-2} \right) + \\ &+ \left( \frac{f^{(n+1)}(y)}{n!}(x-y)^n - \frac{f^{(n)}(y)}{(n-1)!}(x-y)^{n-1} \right) = \\ &= \frac{f^{(n+1)}(y)}{n!}(x-y)^n. \end{aligned}$$

Azon észrevétel alapján számoltunk, hogy a deriválás után nyert összeg olyan, hogy ami az első helyen szerepel az egyik tag deriváltjában, az a következő tagban második helyen szerepel negatív előjellel, ezért csak az utolsó pozitív előjellel vett tag marad meg, a többi az összeadásnál kiesik.  $\square$

### 5.3.3 Általánosított középértéktétel, L'Hospital-szabály

Ebben az alponban a deriválás középértéktételének egy természetes általánosítását bizonyítjuk be, és ennek a segítségével egy olyan tételt tárgyalunk, amelyik hatékony eszközt ad a határértékek kiszámításához.

A középértéktétel általánosításához a következő gondolatmenettel juthatunk el. A középértéktétel azt állítja, hogy az

$$\frac{f(b) - f(a)}{b - a} \quad (5.33)$$

hányados megegyezik az  $f$  deriváltjának valamilyen közbülső  $\xi$  helyen felvett  $f'(\xi)$  értékével. Vegyük észre, hogy a (5.33) hányados nevezőjében is egy függvénynek a megváltozása van, nevezetesen az  $x \mapsto x$  függvény  $b$  és  $a$  helyen vett értékének a különbsége. Emiatt felvetődik a kérdés, hogy tudunk-e valamit mondani az olyan általánosított különbségi hányadosról, amelyik egy  $f$  és egy  $g$  függvény megváltozásának a hányadosa, azaz

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

alakú. A válasz a következő tétel szerint igenlő, és azt állítja, hogy van olyan  $\xi$  az  $(a, b)$  intervallumban, amelyre a szóbanforgó hányados a  $f'(\xi)/g'(\xi)$  hányadossal egyezik meg. A megfogalmazás azonban bizonyos óvatosságot igényel, mert a  $g(b) - g(a)$  különbség esetleg nulla is lehet, ezért mondjuk ki az állítást “hányadosmentesen”:

**Állítás 221 (Általánosított középértéktétel)** *Ha az  $f, g : [a, b] \rightarrow \mathbb{R}$  függvények deriválhatók az intervallum belsejében és folytonosak az  $a$  és  $b$  végpontokban is, akkor van olyan  $\xi$  szám az intervallum belsejében, amelyre*

$$(f(b) - f(a)) \cdot g'(\xi) = (g(b) - g(a)) \cdot f'(\xi).$$

Az általánosított középértéktételt Cauchy-féle középértéktételnek is szokás nevezni.

**Bizonyítás.** A bizonyítás pontosan követi a középértéktétel igazolását. Vegyük a következőképpen definiált  $h$  függvényt:

$$h(t) = (f(b) - f(a)) \cdot g(t) - (g(b) - g(a)) \cdot f(t). \quad (5.34)$$

Azt fogjuk belátni, hogy van olyan  $\xi \in (a, b)$  szám, amelyre  $h'(\xi) = 0$ . Ha ugyanis ez teljesül, akkor a 5.34 deriválásából adódik, hogy

$$h'(\xi) = (f(b) - f(a)) \cdot g'(\xi) - (g(b) - g(a)) \cdot f'(\xi) = 0,$$

ami átrendezve pontosan a tétel állítása.

Most pedig belátjuk, hogy létezik a kívánt tulajdonságú  $\xi$  szám. A  $h$  függvény az intervallum két végpontjában azonos értéket vesz fel, hiszen

$$\begin{aligned} h(b) - h(a) &= (f(b) - f(a)) \cdot g(b) - (g(b) - g(a)) \cdot f(b) \\ &\quad - (f(b) - f(a)) \cdot g(a) + (g(b) - g(a)) \cdot f(a) = \\ &= (f(b) - f(a))(g(b) - g(a)) - (g(b) - g(a))(f(b) - f(a)) = 0 \end{aligned}$$

Ha a  $h$  függvény állandó, akkor az  $(a, b)$  intervallum minden pontja alkalmas lenne  $\xi$  pontnak. Ha nem állandó a  $h$ , akkor valahol az intervallumon belül kisebb vagy nagyobb, mint az intervallum végpontjaiban, és ezért felveszi az intervallumon belül a minimumát vagy a maximumát. Véve egy ilyen  $\xi$  szélsőértékhelyet, ott a  $h'(\xi)$  deriválnak nullának kell lenni.  $\square$

Az általánosított középértéktétel egyik haszna az, hogy ennek a segítségével tárgyalhatjuk a határérték meghatározásának egy hatékony módszerét.

A határértékek kiszámításánál gyakori az a helyzet, hogy olyan

$$f(x)/g(x)$$

törtünk van, amelyre egy  $b$  pontban

- 1) mind a számlálónak mind a nevezőnek az értéke (vagy a határértéke) nulla;
- 2) mind a számlálónak mind a nevezőnek a határértéke végtelen.

Egyszerűen szólva: az

$$\frac{0}{0} \quad \text{vagy} \quad \frac{\infty}{\infty}$$

“határozatlan”, nem értelmezhető hányadosok valamelyikével találkozunk. Az ilyen esetekre nyújt jó módszert az ún. L'Hospital-szabály:

**Állítás 222** *Tegyük fel, hogy az  $f$  és  $g$  függvények differenciálhatók az  $a$  pontnak egy hiányos jobboldali környezetében, és ott  $g'(x) \neq 0$ , továbbá létezik a*

$$\lim_{x \searrow a} \frac{f'(x)}{g'(x)} \tag{5.35}$$

(véges vagy végtelen) limesz. Ekkor igazak a következő állítások.

- (1) Ha

$$\lim_{x \searrow a} f(x) = \lim_{x \searrow a} g(x) = 0, \tag{5.36}$$

akkor

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \lim_{x \searrow a} \frac{f'(x)}{g'(x)}. \tag{5.37}$$

- (2) Ha

$$\lim_{x \searrow a} |g(x)| = +\infty, \tag{5.38}$$

akkor

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \lim_{x \searrow a} \frac{f'(x)}{g'(x)}. \tag{5.39}$$

**Bizonyítás.** Mielőtt a bizonyítást elkezdénénk, jegyezzük meg, hogy az igazolások során azt a feltételt, hogy  $g'(x) \neq 0$  az  $a$  pont valamilyen jobboldali környezetében, a következő állítás bizonyításához használjuk fel: Az  $a$  pont szóbanforgó jobboldali környezetében lévő  $x > y$  pontokra az általánosított középértéktétel az

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

hányados formájában írható fel. Ennek indoklása: Elegendő azt belátni, hogy a nevezők nem lehetnek nullák. A  $g'(\xi)$  esetében ez éppen a feltevés. A  $g(x) - g(y)$  pedig azért nem lehet nulla, mert ha nulla lenne, akkor alkalmazva a középértéktételt a  $g$  függvényre, a  $g$  deriváltja az  $(y, x)$  intervallumban valahol nulla lenne, ami ellentmondana a feltevésnek.

**(1):** A bizonyításban két esetet fogunk megkülönböztetni, aszerint hogy a (5.35) limesz véges vagy végtelen.

*Első eset:*  $\lim_{x \searrow a} \frac{f'(x)}{g'(x)} = \alpha \in \mathbb{R}$ .

A kiindulás szerint tetszőleges  $\epsilon$  pozitív számhoz van olyan  $\delta$  pozitív szám, hogy

$$\left| \frac{f'(\xi)}{g'(\xi)} - \alpha \right| < \epsilon \quad \text{ha} \quad \xi \in (a, a + \delta). \quad (5.40)$$

Legyen az  $x$  egy tetszőleges rögzített szám az  $(a, a + \delta)$  intervallumban. Ha az  $y$  egy tetszőleges szám az  $(a, x)$  intervallumban, akkor az általánosított középértéktétel szerint van olyan  $\xi$  szám, amelyre

$$\xi \in (y, x) \subseteq (a, a + \delta)$$

és

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

amiből 5.40 szerint

$$\left| \frac{f(x) - f(y)}{g(x) - g(y)} - \alpha \right| < \epsilon.$$

Ha az utóbbi egyenlőtlenségben az  $y$  értéke tart az  $a$ -hoz jobbról, akkor (5.36) alapján azt kapjuk, hogy

$$\left| \frac{f(x)}{g(x)} - \alpha \right| \leq \epsilon, \quad \text{ha} \quad x \in (a, a + \delta),$$

ami pontosan a bebizonyítandó (5.37) relációval ekvivalens.

*Második eset:*  $\lim_{x \searrow a} \frac{f'(x)}{g'(x)} = +\infty$ .

Azzal az esettel, amikor a limesz mínusz végtelen, nem kell külön foglalkoznunk, mert adódik ebből az esetből, ha az  $f$  helyett a  $-f$ -et vesszük.

A kiindulás szerint tetszőleges  $\beta$  pozitív számhoz van olyan  $\delta$  pozitív szám, hogy

$$\frac{f'(\xi)}{g'(\xi)} > \beta, \quad \text{ha } \xi \in (a, a + \delta). \quad (5.41)$$

Legyen az  $x$  egy tetszőleges rögzített szám az  $(a, a + \delta)$  intervallumban. Ha az  $y$  egy tetszőleges szám az  $(a, x)$  intervallumban, akkor az általánosított középértéktétel szerint van olyan  $\xi$  szám, amelyre

$$\xi \in (y, x) \subseteq (a, a + \delta)$$

és

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

amiből 5.41 szerint

$$\frac{f(x) - f(y)}{g(x) - g(y)} > \beta.$$

Ha az utóbbi relációban az  $y$  értéke tart az  $a$ -hoz jobbról, akkor (5.36) alapján azt kapjuk, hogy

$$\frac{f(x)}{g(x)} \geq \beta, \quad \text{ha } x \in (a, a + \delta),$$

ami pontosan a bebizonyítandó (5.37) relációval ekvivalens a plusz végtelen határérték esetében.

Vegyük észre, hogy a véges és végtelen eseteknek a fenti megkülönböztetése nem feltétlenül szükségszerű, és egyszerre is tárgyalható lenne. Mindkét esetben egyenlőtlenséget kell ugyanis kezelni a bizonyítás során (az abszolút értékre vonatkozó egyenlőtlenség két közönséges egyenlőtlenségre bontható). Ugyanez elmondható a következő esetek bizonyítására nézve is.

**(2):** Ennek az esetnek a bizonyítása lényegében véve az előzőekben követett menet szerint történik, csak a (5.38) feltétel kihasználása egy kicsit nehezebb, mint az (5.36) relációé.

A bizonyításban itt is két esetet fogunk megkülönböztetni, aszerint hogy az (5.35) limesz véges vagy végtelen.

*Első eset:*  $\lim_{x \searrow a} \frac{f'(x)}{g'(x)} = \alpha \in \mathbb{R}$ .

A kiindulás szerint tetszőleges  $\epsilon$  pozitív számhoz van olyan  $\delta$  pozitív szám, hogy

$$\left| \frac{f'(\xi)}{g'(\xi)} - \alpha \right| < \epsilon, \quad \text{ha } \xi \in (a, a + \delta). \quad (5.42)$$

Az  $x$  és  $y$  változók megválasztása most megfordított: Legyen az  $y$  egy tetszőleges rögzített szám az  $(a, a + \delta)$  intervallumban. Ha az  $x$  egy tetszőleges szám az  $(a, y)$  intervallumban, akkor az általánosított középértéktétel szerint van olyan  $\xi$  szám, amelyre

$$\xi \in (x, y) \subseteq (a, a + \delta)$$

és

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

amiből (5.42) szerint

$$\left| \frac{f(y) - f(x)}{g(y) - g(x)} - \alpha \right| < \epsilon, \quad \text{ha} \quad a < x < y < a + \delta. \quad (5.43)$$

Ha ebben az egyenlőtlenségben az  $a$  elem valamilyen jobboldali környezetében a különbségi hányadost az  $\frac{f(x)}{g(x)}$  hányadossal tudnánk helyettesíteni, akkor készen is lennénk. Ebből a célból kiindulva végezzük el a következő átalakítást:

$$\begin{aligned} \frac{f(y) - f(x)}{g(y) - g(x)} - \frac{f(x)}{g(x)} &= \frac{f(y) - f(x)}{g(y) - g(x)} + \frac{f(y) - f(x) - f(y)}{g(x)} = \\ &= \frac{f(y) - f(x)}{g(y) - g(x)} + \frac{f(y) - f(x)}{g(x)} - \frac{f(y)}{g(x)} = \\ &= (f(y) - f(x)) \left( \frac{1}{g(y) - g(x)} + \frac{1}{g(x)} \right) - \frac{f(y)}{g(x)} = \\ &= \frac{f(y) - f(x)}{g(y) - g(x)} \frac{g(y)}{g(x)} - \frac{f(y)}{g(x)}. \end{aligned}$$

Rögzítsük a most belátott azonosságot:

$$\frac{f(y) - f(x)}{g(y) - g(x)} - \frac{f(x)}{g(x)} = \frac{f(y) - f(x)}{g(y) - g(x)} \frac{g(y)}{g(x)} - \frac{f(y)}{g(x)}. \quad (5.44)$$

Most pedig a (5.43) relációból először is azt olvassuk le, hogy a differenciahányados korlátos, ha az  $x$  és  $y$  a leírtak szerint helyezkedik el, hiszen nem nagyobb, mint az  $|\alpha| + \epsilon$  szám.

Nézzük ezután az (5.44) azonosságnak a jobb oldalát. Ha az  $x$  értéke jobbról tart az  $a$ -hoz, akkor (5.38) alapján az első tag tart a nullához, mert a  $g(y)$  rögzített, a differenciahányados pedig korlátos. A második tag szintén tart a nullához, ha az  $x$  jobbról tart az  $a$ -hoz, tehát az

$$x \mapsto \left( \frac{f(y) - f(x)}{g(y) - g(x)} - \frac{f(x)}{g(x)} \right)$$

függvény tart a nullához, ha az  $x$  jobbról tart az  $a$ -hoz. Ez alapján pedig, az (5.43) reláció szerint van olyan  $(a, a + \delta_1)$  részkörnyezete az  $(a, a + \delta)$  környezetnek, hogy

$$\left| \frac{f(x)}{g(x)} - \alpha \right| < \epsilon, \quad \text{ha} \quad x \in (a, a + \delta_1),$$

ami pontosan a bebizonyítandó (5.39) relációval ekvivalens.

Második eset:  $\lim_{x \searrow a} \frac{f'(x)}{g'(x)} = +\infty$ .

A kiindulás szerint tetszőleges  $\beta$  pozitív számhoz van olyan  $\delta$  pozitív szám, hogy

$$\frac{f'(\xi)}{g'(\xi)} > 2\beta, \quad \text{ha } \xi \in (a, a + \delta). \quad (5.45)$$

Legyen az  $y$  egy tetszőleges rögzített szám az  $(a, a + \delta)$  intervallumban. Ha az  $x$  egy tetszőleges szám az  $(a, y)$  intervallumban, akkor az általánosított középértéktétel szerint van olyan  $\xi$  szám, amelyre

$$\xi \in (x, y) \subseteq (a, a + \delta)$$

és

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

amiből (5.45) szerint

$$\frac{f(y) - f(x)}{g(y) - g(x)} > 2\beta, \quad \text{ha } a < x < y < a + \delta. \quad (5.46)$$

A most is fennálló (5.44) azonosság szerint

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

amiből

$$\frac{f(x)}{g(x)} = \frac{f(y) - f(x)}{g(y) - g(x)} \left( 1 - \frac{g(y)}{g(x)} \right) + \frac{f(y)}{g(x)}.$$

Felhasználva az (5.46) egyenlőtlenséget, a jobboldal a következőképpen becsülhető: az első tag első tényezője nagyobb, mint  $2\beta$ , a másik tényező tart az 1-hez, mert a  $g(y)/g(x)$  nullához tart, mivel a  $|g(x)|$  tart a  $+\infty$ -hez; a második tag, ahogyan az előbb már indokoltuk, tart a nullához. Ezek miatt van olyan  $(a, a + \delta_1)$  részkörnyezete az  $(a, a + \delta)$  környezetnek, amelybe eső  $x$  értékekre

$$\frac{f(x)}{g(x)} > \beta,$$

amivel be is láttuk a tételt.  $\square$

Az előző tételt jobboldali határértékekre mondtuk ki. Nyilvánvaló a bizonyításokból, hogy megfelelő változtatásokkal azonos a bizonyítás a baloldali határértékek esetében is. A jobboldali és baloldali határértékekre kimondott tételek következményeként megfogalmazhatjuk a határértékekre vonatkozó tételt mint egyszerű következményt.

**Állítás 223 (L'Hospital-szabály.)** Tegyük fel, hogy az  $f$  és  $g$  függvények differenciálhatók az  $a$  pontnak egy hiányos környezetében, és ott a  $g'(x)$  nem 0, továbbá létezik a

$$\lim_{x \rightarrow a} \frac{f'(x)}{g'(x)} \quad (5.47)$$

limesz. Ekkor igazak a következő állítások.

(1) Ha

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0, \quad (5.48)$$

akkor

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}. \quad (5.49)$$

(2) Ha

$$\lim_{x \rightarrow a} |g(x)| = +\infty, \quad (5.50)$$

akkor

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}. \quad (5.51)$$

A pont hátralévő részében példákat mutatunk a L'Hospital-szabály alkalmazására. Először kezdjük egy egyszerű példával.

**Példa 5.17** Határozzuk meg a következő határértéket:

$$\lim_{x \rightarrow 1} \frac{x^3 - 3x^2 + 2}{x^3 - 4x^2 + 3}.$$

A tört az  $x = 1$  helyen  $\frac{0}{0}$  határozatlan alakú. A nevező  $3x^2 - 8x$  deriváltja az 1-ben  $-5$ , ezért nemnulla az 1 egy környezetében, tehát teljesülnek a L'Hospital-szabály feltételei. A számláló és nevező deriváltjai hányadosának van limesze:

$$\lim_{x \rightarrow 1} \frac{3x^2 - 6x}{3x^2 - 8x} = \frac{3}{5},$$

ezért a meghatározandó limesz is  $3/5$ . □

**Példa 5.18**

$$\lim_{x \rightarrow 0} \frac{\tan x - x}{x - \sin x} = ?$$

A tört a nulla helyen  $\frac{0}{0}$  határozatlan alakú. A nevező deriváltja nemnulla a 0 egy hiányos környezetében, ezért a szabály alkalmazható.

Először számoljuk ki a számláló és nevező deriváltjának a hányadosát:

$$\frac{(\tan x - x)'}{(x - \sin x)'} = \frac{\frac{1}{\cos^2 x} - 1}{1 - \cos x} = \frac{1 - \cos^2 x}{(1 - \cos x) \cos^2 x} = \frac{1 + \cos x}{\cos^2 x}.$$

Ez alapján a keresett limesz:

$$\left. \frac{1 + \cos x}{\cos^2 x} \right|_{x=0} = \frac{2}{1} = 2.$$

□

## 6.

# Monoton és konvex függvények

Ebben a fejezetben bebizonyítjuk a monoton és konvex függvények legismertebb tulajdonságait.

## 6.1 Alaptulajdonságok

Ebben a pontban először a monoton majd a konvex függvények azon tulajdonságait vizsgáljuk, amelyek függetlenek a differenciálhatóság fogalmától.

### 6.1.1 Monoton függvények

Most elsősorban a folytonossági kérdésekkel foglalkozunk. A monoton függvények a határérték szempontjából eléggé szabályosan viselkednek:

**Állítás 224** *Legyen az  $f$  függvény monoton (növekedő vagy fogyó) egy  $(a, b)$  intervallumon. Ekkor az intervallum minden  $c$  pontjában van mind jobboldali mind baloldali határértéke, mégpedig növekedő függvény esetében*

$$\lim_{x \nearrow c} f(x) = \sup_{a < x < c} f(x) \quad \text{és} \quad \lim_{x \searrow c} f(x) = \inf_{c < x < b} f(x).$$

*Fogyó függvény esetében a szuprérum és infimum jelek helyet cserélnek.*

A tétel szerint egy monoton függvény egy  $c$  pontban a következőképpen viselkedhet:

- A függvénynek van határértéke és az megegyezik a helyettesítési értékkel, azaz a  $c$  pontban a függvény folytonos.

- A függvénynek a bal- és jobboldali határértéke különböző, és a függvényérték is ezektől eltérő.
- A bal- és jobboldali határértékek különbözőek, és a függvény értéke megegyezik a baloldali (jobboldali) határértékkel; ekkor a leképezést balról (jobbról) folytonosnak is szokás mondani.

A felsorolt esetek közül az utolsó kettőre, amikor a függvény nem folytonos, azt is szoktuk mondani, hogy *a c helyen szakadása van a függvénynek*. A 6.1. ábrán illusztráljuk az állítást, és olyanfüggvényt vettünk, hogy a szakadások mindegyik típusa előforduljon.

**Bizonyítás.** Vizsgáljuk, mondjuk, a monoton növekedő függvény és a baloldali határérték esetét. A többi eset hasonlóan kezelhető.

Az  $\alpha = \sup_{a < x < c} f(x)$  felső határ létezik, mivel az  $f$  a monotonitása miatt korlátos a  $c$  egy baloldali környezetében, egy felső korlát például az  $f(c)$ .

Megmutatjuk, hogy  $\lim_{x \nearrow c} f(x) = \alpha$ . A felső határ tulajdonsága szerint tetszőleges  $\epsilon$  pozitív számhoz van olyan  $\delta$  pozitív szám, hogy

$$\alpha - \epsilon < f(c - \delta) \leq \alpha.$$

Mivel az  $f$  monoton növekedő, ezért ebből adódik, hogy

$$\alpha - \epsilon < f(x) \leq \alpha, \quad \text{ha} \quad c - \delta \leq x < c,$$

amit átrendezve kapjuk, hogy

$$-\epsilon < f(x) - \alpha \leq 0 < \epsilon, \quad \text{ha} \quad c - \delta < x < c.$$

Ez pedig pontosan azt jelenti, hogy

$$|f(x) - \alpha| < \epsilon, \quad \text{ha} \quad c - \delta < x < c,$$

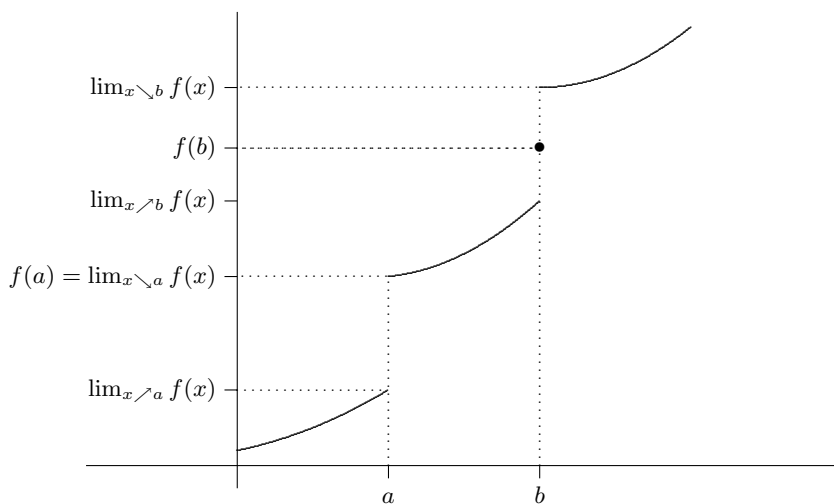
ezért  $\lim_{x \nearrow c} f(x) = \alpha$ . □

A következő tétel azt fogja mondani, hogy egy monoton függvénynek viszonylag kevés olyan helye van, ahol szakad, azaz ahol nem folytonos.

**Állítás 225** *Ha az  $f : (a, b) \rightarrow \mathbb{R}$  függvény monoton, akkor szakadási helyeinek a halmaza legfeljebb megszámlálhatóan végtelen számosságú.*

Az állítás alapján azt hihetné valaki, hogy a szakadások “izoláltan” diszkrét pontokban vannak, amilyen rajzot tudunk is készíteni. Ez azonban általában nem igaz, mert meg lehet adni olyan monoton függvényt, amelyiknek szakadása van az intervallum minden racionális pontjában.

Megjegyezzük még azt is, hogy tetszőleges valós függvényre igaz az, hogy az olyan pontoknak a halmaza, ahol a függvény nem folytonos, de létezik a baloldali és jobboldali határértéke, legfeljebb megszámlálhatóan végtelen számosságú. Ennek a bizonyítása azonban nem annyira egyszerű, mint a most kimondott tételé.



6.1. ábra: Monoton függvény szakadásai

Általában egy valós függvény esetében persze lehetnek olyan pontok is, ahol nem léteznek sem baloldali sem jobboldali határértékek, vagy a kettő közül csak az egyik létezik.

**Bizonyítás.** Monoton növekedő függvényre végezzük a bizonyítást. Legyen  $S$  azon pontok halmaza, ahol az  $f$  függvénynek szakadása van. A 224. állítás szerint az  $S$  pontjaiban különböző baloldali és jobboldali határértékek léteznek. Emiatt, mivel minden pozitív hosszúságú intervallumban van racionális szám, minden  $s \in S$  ponthoz van olyan  $r(s)$  racionális szám, hogy

$$\lim_{x \nearrow s} f(x) < r(s) < \lim_{x \searrow s} f(x).$$

A monoton növekedés miatt  $s_1 < y < s_2$  esetén

$$\lim_{x \searrow s_1} f(x) \leq f(y) \leq \lim_{x \nearrow s_2} f(x),$$

ezért az

$$\left( \lim_{x \nearrow s} f(x), \lim_{x \searrow s} f(x) \right)$$

intervallumok diszjunktak, következésképpen az  $r(s)$  racionális számok különbözőek, és így a számuk, azaz az  $S$  elemeinek a száma is, legfeljebb akkora, mint a racionális számok számossága, ami megszámlálható.  $\square$

A következő állítás röviden szólva azt mondja, hogy folytonos és injektív leképezés csak szigorúan monoton leképezés lehet.

**Állítás 226** *Tetszőleges  $f : (a, b) \rightarrow \mathbb{R}$  folytonos injektív leképezés egyúttal szigorúan monoton.*

Itt is tetszőleges intervallum szerepelhetne az  $(a, b)$  helyett.

**Bizonyítás.** Elegendő megmutatni, hogy minden  $a < \alpha < \beta < b$  esetén  $f$  szigorúan monoton  $[\alpha, \beta]$ -n. Nyilván  $f(\alpha) \neq f(\beta)$ . Tegyük fel például, hogy  $f(\alpha) < f(\beta)$ . Megmutatjuk, hogy  $f$  szigorúan monoton nő  $[\alpha, \beta]$ -n.

Indirekt módon tegyük fel, hogy létezik  $\alpha \leq x < y \leq \beta$ , hogy  $f(x) \geq f(y)$ . Ekkor nem lehet  $f(x) \geq f(\beta)$ , hiszen ezesetben  $f(\alpha) < f(\beta) \leq f(x)$  alapján a Bolzano-tétel miatt az  $[\alpha, x]$  intervallumon az  $f$  függvény felvenné az  $f(\beta)$  értéket, ellentmondásban az injektivitással. Így

$$f(y) \leq f(x) < f(\beta),$$

ezért ismét a Bolzano-tétel alapján az  $[y, \beta]$  intervallumon  $f$  felveszi az  $f(x)$  értéket. Ez ismét ellentmondásban van az  $f$  injektivitásával.

Tehát  $f$  szigorúan monoton nő  $[\alpha, \beta]$ -n. □

### 6.1.2 Konvex és konkáv függvények

A következő tételben a korlátos és zárt intervallum (szakasz) pontjainak egy sajátos megadási módját vezetjük be.

**Állítás 227** *Legyen az  $u < v$  két tetszőleges valós szám. Az  $[u, v]$  zárt intervallum (szakasz) tetszőleges  $x$  pontja egyértelműen írható fel*

$$x = \lambda u + (1 - \lambda)v, \quad \lambda \in [0, 1]$$

*formában. Az  $x$  pontot az  $u$  és  $v$  pontok  $\lambda$  és  $(1 - \lambda)$  súlyokkal vett konvex kombinációjának vagy súlyozott számtani közepének mondjuk.*

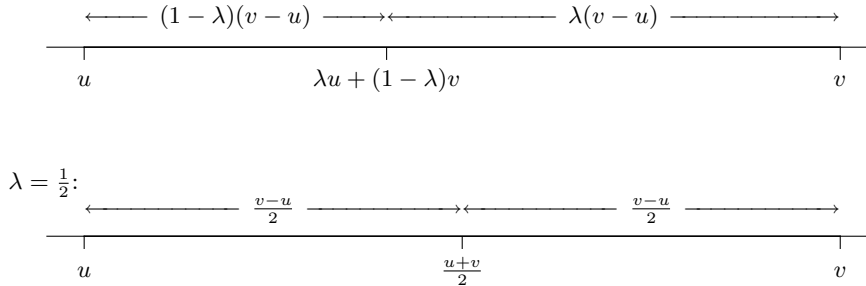
A konvex kombinációnak a következő fizikai interpretáció adható: ha az  $u$  illetve  $v$  pontokba  $\lambda$  illetve  $(1 - \lambda)$  tömegű anyagi pontokat képzelünk, akkor a súlypontjukat az  $x = \lambda u + (1 - \lambda)v$  konvex kombináció adja meg. A 6.2. ábra szemlélteti a geometriai tartalmat.

**Bizonyítás.** Legyen az  $x \in [u, v]$  egy tetszőleges pont. A kívánt előállítás létezését és egyértelműségét egyszerre igazoljuk azzal, hogy az  $x = \lambda u + (1 - \lambda)v$  egyenletből kiszámítjuk a  $\lambda$  számot:

$$x = \lambda u + (1 - \lambda)v \implies x - u = (1 - \lambda)(v - u) \implies \lambda = \frac{v - x}{v - u}.$$

Ennek megfelelően az  $x$  konvex kombinációként való előállítása:

$$x = \frac{v - x}{v - u}u + \frac{x - u}{v - u}v.$$

6.2. ábra: Az  $u$  és  $v$   $\lambda$  súllyal vett konvex kombinációja.

□

A konvex kombináció fogalmát kettő helyett véges sok tagra is általánosíthatjuk, de ekkor már nincs szó egyértelműségről.

**Definíció 228** Legyenek az  $x_1, x_2, \dots, x_n$  tetszőleges számok, legyen továbbá

$$\lambda_1, \lambda_2, \dots, \lambda_n \in [0, 1] \quad \text{és} \quad \lambda_1 + \lambda_2 + \dots + \lambda_n = 1.$$

A

$$\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = \sum_{i=1}^n \lambda_i x_i$$

módon definiált számot az  $x_i$ ,  $i = 1, \dots, n$  számok  $\lambda_i$ ,  $i = 1, \dots, n$  súlyokkal vett konvex kombinációjának vagy súlyozott számtani közepének nevezzük.

Abban az esetben, ha mindegyik súly azonos, azaz  $1/n$ , akkor a már ismert

$$\frac{x_1 + x_2 + \dots + x_n}{n}$$

számtani közepet kapjuk, amit az  $x_1, x_2, \dots, x_n$  számok átlagának is mondanak.

Egy igen egyszerű észrevétel a súlyozott számtani közép nagyságára:

**Állítás 229** Tetszőleges  $x_1, x_2, \dots, x_n$  számok tetszőleges  $\lambda_i \in [0, 1]$ ,  $i = 1, \dots, n$ ,  $\sum_{i=1}^n \lambda_i = 1$  súlyokkal vett

$$\sum_{i=1}^n \lambda_i x_i$$

súlyozott számtani közepe a számok minimuma és maximuma közé esik, azaz

$$\min_{1 \leq i \leq n} x_i \leq \sum_{i=1}^n \lambda_i x_i \leq \max_{1 \leq i \leq n} x_i.$$

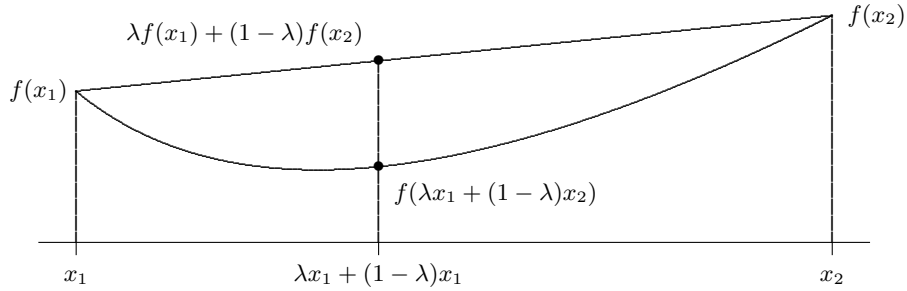
Bizonyítás. Egyrészt

$$\sum_{i=1}^n \lambda_i x_i \leq \sum_{i=1}^n \lambda_i \max_{1 \leq j \leq n} x_j = \max_{1 \leq j \leq n} x_j,$$

másrészt

$$\sum_{i=1}^n \lambda_i x_i \geq \sum_{i=1}^n \lambda_i \min_{1 \leq j \leq n} x_j = \min_{1 \leq j \leq n} x_j.$$

□



6.3. ábra: Konvex függvény.

Ha egy  $f$  függvényt valamilyen  $[u, v]$  szakasz pontjaiban vizsgáljuk, akkor ez a 227. állítás szerint az  $\lambda u + (1 - \lambda)v$  helyeket jelenti. Az  $l(x) = cx$  lineáris függvényre

$$l(\lambda u + (1 - \lambda)v) = \lambda l(u) + (1 - \lambda)l(v).$$

A következőkben olyan  $f$  függvényekkel fogunk foglalkozni, amelyek az egyenlőség helyett egyenlőtlenséggel teljesítik ezt a formát, azaz

$$f(\lambda u + (1 - \lambda)v) \leq \lambda f(u) + (1 - \lambda)f(v).$$

Ahogy látható, az ilyen függvény helyettesítési értéke a lineáris függvény helyettesítési értéke alatt marad, ezért jogosan nevezhetnénk “szublineáris” (lineáris alatti) függvénynek, de ehelyett a konvex elnevezés vált általánossá. Geometriailag ez nagyon szemléletes, ahogyan azt a 6.3 ábra is mutatja, a gráf mindig az  $(u, f(u))$  és  $(v, f(v))$  pontokat összekötő egyenes alatt marad az  $[u, v]$  intervallumban. Az elnevezéseket definícióban is rögzítjük:

**Definíció 230** Legyen az  $f$  egy intervallumon értelmezett függvény. Ha az intervallum minden  $x_1$  és  $x_2$  pontjára és  $\lambda \in [0, 1]$  számra

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2), \quad (6.1)$$

akkor az  $f$  függvényt konvexnek mondjuk.

Ha a (6.1) egyenlőtlenség fordított irányban teljesül, akkor konkávnak mondjuk a függvényt.

Ha a definícióban szereplő egyenlőtlenség szigorúan teljesül, kivéve az

$$x_1 = x_2 \quad \text{vagy} \quad \lambda = 1 \quad \text{vagy} \quad 0$$

eseteket, akkor szigorú konvexitásról illetve konkávitásról beszélünk.

A konkáv függvény tipikus ábráját a 6.4. rajzunk illusztrálja.

A következő állításban a konvex és konkáv függvények legelemibb tulajdonságait gyűjtöttük egybe.

#### Állítás 231

- (1) Egy  $f : [a, b] \rightarrow \mathbb{R}$  függvény pontosan akkor konvex, ha a  $-f$  függvény konkáv.
- (2) Ha az  $f$  és  $g : [a, b] \rightarrow \mathbb{R}$  függvények konvexek (konkávak), akkor az  $f + g$  függvény is konvex (konkáv).
- (3) Ha az  $f$  függvény konvex (konkáv) és az  $\alpha$  nemnegatív szám, akkor az  $\alpha f$  függvény is konvex (konkáv).

Az (1) állítás úgy is fogalmazható, hogy a konvex és konkáv függvények halmaza között az  $f \mapsto -f$  leképezés kölcsönösen egyértelmű (bijekció). Ezen egyszerű kapcsolat miatt valójában csak a konvex függvényekkel foglalkozunk részletesen, hiszen az állítások a kapcsolat révén megfelelően átvihetők konkáv függvényekre is.

**Bizonyítás.** (1): Az  $f$  konvexitása azt jelenti, hogy  $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$ , amit mínusz eggyel megszorozva a

$$(-f)(\lambda x + (1 - \lambda)y) \geq \lambda(-f)(x) + (1 - \lambda)(-f)(y)$$

egyenlőtlenséghez jutunk, ami éppen azt jelenti, hogy  $-f$  konkáv.

(2): Az  $f$  és  $g$  függvények konvexitását jelentő két  $f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$  és  $g(\lambda x + (1 - \lambda)y) \leq \lambda g(x) + (1 - \lambda)g(y)$  egyenlőtlenséget összeadva adódik, hogy

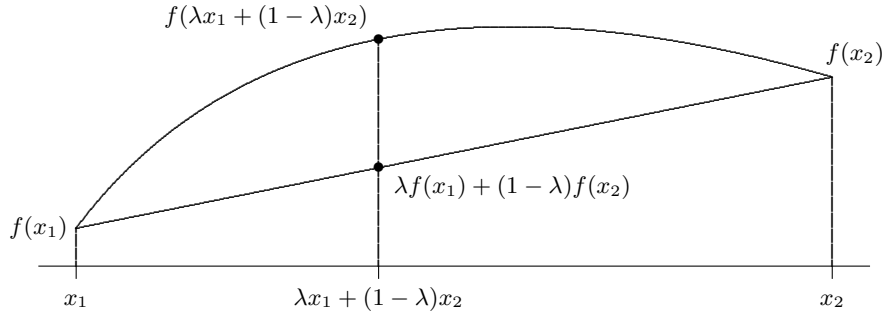
$$(f + g)(\lambda x + (1 - \lambda)y) \leq \lambda(f + g)(x) + (1 - \lambda)(f + g)(y).$$

(3): Az  $f$ -re vonatkozó definíció szerinti egyenlőtlenséget egy  $\alpha$  nemnegatív számmal megszorozva az

$$(\alpha f)(\lambda x + (1 - \lambda)y) \leq \lambda(\alpha f)(x) + (1 - \lambda)(\alpha f)(y)$$

egyenlőtlenséget kapjuk, tehát az  $\alpha f$  függvény is konvex.  $\square$

A következő tétel azt állítja, hogy két tagú konvex kombináció helyett tetszőleges számú tagot is vehetünk egy konvex (konkáv) függvénynél.



6.4. ábra: Konkáv függvény.

**Állítás 232 (Jensen-egyenlőtlenség)** *Ha egy  $f$  leképezés konvex egy  $[a, b]$  intervallumon, akkor az intervallum tetszőleges  $x_1, x_2, \dots, x_n$  elemeinek tetszőleges konvex kombinációjára fennáll az*

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i)$$

*egyenlőtlenség.*

**Bizonyítás.** Teljes indukcióval bizonyítjuk az állítást.

Kezdő lépés: Az állítás  $n = 2$  esetében igaz a konvex függvény definíciója szerint.

Indukciós lépés: Tegyük fel most, hogy fennáll az egyenlőtlenség  $(n - 1)$ -nél nem nagyobb tagszámú konvex kombinációk mellett. Ebből belátjuk, hogy igaz  $n$  tagú konvex kombinációkra is.

Ha egy  $n$  tagú konvex kombinációban valamelyik súly nulla, akkor arra az indukciós feltevés szerint igaz az egyenlőtlenség, hiszen ekkor legfeljebb  $(n - 1)$  tagja van a kombinációnak. Legyen ezért most az  $\lambda_1, \lambda_2, \dots, \lambda_n$  súlyok mindegyike pozitív.

Vegyük az  $f$  függvénynek az értékét egy ilyen kombinációnál és végezzünk el egy olyan célszerű átalakítást, amelyikkel kevesebb tagú konvex kombinációkat hozunk be:

$$\begin{aligned} & f(\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n) = \\ &= f\left[\lambda_1 x_1 + (1 - \lambda_1) \left(\frac{\lambda_2}{1 - \lambda_1} x_2 + \frac{\lambda_3}{1 - \lambda_1} x_3 + \dots + \frac{\lambda_n}{1 - \lambda_1} x_n\right)\right]. \end{aligned}$$

Amint látható, az átalakítással azt értük el, hogy az  $f$  argumentumában egy két tagú konvex kombinációt alakítottunk ki,  $\lambda_1$  és  $(1 - \lambda_1)$  együtthatókkal. A kombináció második tagja,

$$\frac{\lambda_2}{1 - \lambda_1} x_2 + \frac{\lambda_3}{1 - \lambda_1} x_3 + \dots + \frac{\lambda_n}{1 - \lambda_1} x_n$$

maga is egy  $(n - 1)$  tagú konvex kombináció, mivel

$$\begin{aligned} & \frac{\lambda_2}{1 - \lambda_1} + \frac{\lambda_3}{1 - \lambda_1} + \cdots + \frac{\lambda_n}{1 - \lambda_1} = \\ & = \frac{1}{1 - \lambda_1} (\lambda_2 + \lambda_3 + \cdots + \lambda_n) = \frac{1}{1 - \lambda_1} (1 - \lambda_1) = 1. \end{aligned}$$

Folytassuk most már az utóbbiak figyelembevételével az előbb elkezdett egyenlőtlenséget, felhasználva azt, hogy fennáll az egyenlőtlenség kettő és minden  $(n-1)$ -nél nem nagyobb tagszámú konvex kombinációra:

$$\begin{aligned} & f(\lambda_1 x_1 + (1 - \lambda_1) \left( \frac{\lambda_2}{1 - \lambda_1} x_2 + \frac{\lambda_3}{1 - \lambda_1} x_3 + \cdots + \frac{\lambda_n}{1 - \lambda_1} x_n \right)) \leq \\ & \lambda_1 f(x_1) + (1 - \lambda_1) f \left( \frac{\lambda_2}{1 - \lambda_1} x_2 + \frac{\lambda_3}{1 - \lambda_1} x_3 + \cdots + \frac{\lambda_n}{1 - \lambda_1} x_n \right) \leq \\ & \lambda_1 f(x_1) + (1 - \lambda_1) \left( \frac{\lambda_2}{1 - \lambda_1} f(x_2) + \frac{\lambda_3}{1 - \lambda_1} f(x_3) + \cdots + \frac{\lambda_n}{1 - \lambda_1} f(x_n) \right) = \\ & = \lambda_1 f(x_1) + \lambda_2 f(x_2) + \cdots + \lambda_n f(x_n), \end{aligned}$$

ami pontosan azt adja, hogy  $n$ -re is teljesül az egyenlőtlenség.  $\square$

A későbbiekben szükségünk lesz a következő tételre:

**Állítás 233 (Monotonitási kritérium)** *Ha az  $f : [a, b] \rightarrow \mathbb{R}$  függvény konvex, akkor a*

$$h \mapsto \frac{f(x+h) - f(x)}{h}, \quad x, (x+h) \in [a, b], \quad h \neq 0$$

*különbséghányados-függvény monoton növekedő.*

Megjegyezzük, hogy a tétel megfordítása is igaz: ha a szóban forgó hányados monoton növekedő, akkor az  $f$  függvény konvex. Az állításnak erre az irányára nem lesz szükségünk, ezért most nem tárgyaljuk.

**Bizonyítás.** Legyen  $h_1 < h_2$ . Azt kell megmutatnunk, hogy

$$\frac{f(x+h_1) - f(x)}{h_1} \leq \frac{f(x+h_2) - f(x)}{h_2}. \quad (6.2)$$

A bizonyítás abból áll, hogy a (6.2) egyenlőtlenséget ekvivalens (visszafelé is elvégezhető) átalakításokkal olyan alakra hozzuk, ami a függvény konvexitása miatt teljesül. Problémát csak az okoz, hogy a  $h_1$  illetve  $h_2$  előjele szerint eltérő módon kell elvégezni a rendezést. A következő eseteket különböztetjük meg:

$$1) \ h_1 < h_2 \leq 0, \quad 2) \ h_1 < 0 < h_2, \quad 3) \ 0 \leq h_1 < h_2.$$

Az átalakítást mindegyik esetben úgy kell elvégezni, hogy a közbülső helyhez tartozó tag maradjon a baloldalon. Ennek megfelelően az átrendezett alakok:

1)

$$f(x + h_2) \leq \frac{h_2}{h_1} f(x + h_1) + \left(1 - \frac{h_2}{h_1}\right) f(x).$$

Ez pedig egy konvexitási egyenlőtlenség, mivel

$$x + h_2 = \frac{h_2}{h_1}(x + h_1) + \left(1 - \frac{h_2}{h_1}\right)x.$$

2)

$$f(x) \leq \frac{-h_1}{h_2 - h_1} f(x + h_2) + \frac{h_2}{h_2 - h_1} f(x + h_1),$$

és ez is egy konvexitási egyenlőtlenség, mivel

$$x = \frac{-h_1}{h_2 - h_1}(x + h_2) + \frac{h_2}{h_2 - h_1}(x + h_1).$$

3) Az 1)-hez hasonlóan:

$$f(x + h_1) \leq \frac{h_1}{h_2} f(x + h_2) + \left(1 - \frac{h_1}{h_2}\right) f(x).$$

□

A közgazdaságtanban sok esetben feltételezik, hogy egy adott jelenséget leíró függvény konvex vagy konkáv. Éppen ezért fontos tudnunk, hogy ez a feltevés már maga után vonja a függvény korlátosságát és folytonosságát, ahogyan azt a következő tételekben állítjuk.

**Állítás 234** Ha egy  $f : [a, b] \rightarrow \mathbb{R}$  függvény konvex, akkor korlátos.

**Bizonyítás.** A felülről való korlátosság: Ha az  $x$  tetszőleges pontja az  $[a, b]$  intervallumnak, akkor van olyan  $0 \leq \lambda \leq 1$ , hogy  $x = \lambda a + (1 - \lambda)b$ , így a konvexitás szerint

$$f(x) = f(\lambda a + (1 - \lambda)b) \leq \lambda f(a) + (1 - \lambda)f(b) \leq \max\{f(a), f(b)\},$$

tehát az  $M \doteq \max\{f(a), f(b)\}$  felső korlát az  $f$  értékeire.

Az alulról való korlátosság: Az  $[a, b]$  intervallum tetszőleges  $x$  pontjához van olyan  $t$  szám, hogy  $x = \frac{a+b}{2} + t$ . A konvexitás felhasználásával:

$$\begin{aligned} f\left(\frac{a+b}{2}\right) &= f\left[\frac{1}{2}\left(\frac{a+b}{2} + t\right) + \frac{1}{2}\left(\frac{a+b}{2} - t\right)\right] \leq \\ &\leq \frac{1}{2}f\left(\frac{a+b}{2} + t\right) + \frac{1}{2}f\left(\frac{a+b}{2} - t\right) = \\ &= \frac{1}{2}f(x) + \frac{1}{2}f\left(\frac{a+b}{2} - t\right), \end{aligned}$$

amiből

$$f(x) \geq 2 \cdot f\left(\frac{a+b}{2}\right) - f\left(\frac{a+b}{2} - t\right) \geq 2 \cdot f\left(\frac{a+b}{2}\right) - M,$$

ahol az  $M$  a bizonyítás első felében megadott felső korlátja  $f$ -nek. Eszerint az

$$m \doteq 2 \cdot f\left(\frac{a+b}{2}\right) - M$$

szám egy alsó korlát. □

**Állítás 235** *Ha egy  $f : [a, b] \rightarrow \mathbb{R}$  függvény konvex, akkor az intervallum belsejében folytonos.*

**Bizonyítás.** Legyen  $c \in (a, b)$  egy rögzített pont. Láttuk, hogy  $f$  konvexitása mellett a  $c$  ponthoz tartozó

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

különbséghányados-függvény monoton nő. Ez azt jelenti, hogy

$$F_c(a) \leq F_c(x) \leq F_c(b), \quad x \in [a, b], x \neq c.$$

Így ha bevezetjük a  $K = \max\{|F_c(a)|, |F_c(b)|\}$  jelölést akkor  $|F_c(x)| \leq K$ , azaz minden  $x \in [a, b] \setminus \{c\}$  mellett

$$|f(x) - f(c)| \leq K|x - c|.$$

Ebből viszont az  $f$  függvény  $c$ -beli folytonossága már nyilvánvaló. □

## 6.2 Differenciálható függvények vizsgálata

Ebben az alponthban a differenciálhatóság feltevése mellett fogjuk vizsgálni a függvények monotonitását és konvexitását.

### 6.2.1 A monotonításra vonatkozó feltételek

A középértéktétel segítségével könnyen bebizonyíthatjuk a deriválható függvények monotonítására és állandó voltára vonatkozó alábbi állítást.

**Állítás 236** *Legyen az  $f : [a, b] \rightarrow \mathbb{R}$  függvény deriválható az intervallum belsejében, és folytonos az  $a$  és  $b$  végpontokban is. Ekkor fennállnak a következő állítások:*

- (1) *Az  $f$  függvény pontosan akkor monoton növekedő az  $[a, b]$  intervallumban, ha  $f'(x) \geq 0$  minden  $x \in (a, b)$  pontban.*

- (2) Az  $f$  függvény pontosan akkor monoton fogyó az  $[a, b]$  intervallumban, ha  $f'(x) \leq 0$  minden  $x \in (a, b)$  pontban.
- (3) Ha  $f'(x) = 0$  az  $(a, b)$  intervallum minden pontjában, akkor az  $f$  leképezés állandó az  $[a, b]$  intervallumban.
- (4) Ha a  $g$  leképezés is deriválható az  $(a, b)$  intervallumban, folytonos a végpontokban és  $g'(x) = f'(x)$  minden  $x \in (a, b)$  pontban, akkor az  $f$  és  $g$  függvények különbsége az  $[a, b]$  intervallumban állandó.

**Bizonyítás.** Az állítások mindegyike egyszerű következménye a középértéktételnek. Legyen az  $x < y$  két tetszőleges pontja az  $[a, b]$  intervallumnak. Az  $f$  függvény folytonos az  $[x, y]$  zárt és korlátos intervallumban, belül deriválható, ezért a középértéktétel szerint van olyan  $\xi$  pont az  $(x, y)$  intervallumban, hogy

$$f'(\xi)(y - x) = f(y) - f(x). \quad (6.3)$$

(1): Ha  $f'(\xi) \geq 0$  minden pontjára az intervallumnak, akkor (6.3)-ból következik, hogy

$$f(y) - f(x) \geq 0, \quad \text{ha} \quad x < y,$$

tehát az  $f$  monoton növekedő.

Az állításnak az a fele, hogy monoton növekedő és deriválható leképezésnek a deriváltja nemnegatív, azonnal jön abból, hogy ezen esetben a különbségi hányados nemnegatív.

(2): Hasonló a fenti indokláshoz, vagy következik az előző állításból, ha azt az  $f$  helyett a  $-f$ -re alkalmazzuk.

(3): Ha  $f'(\xi) = 0$  az intervallum minden pontjában, akkor a (6.3) szerint  $f(y) - f(x) = 0$ , és mivel az  $x$  és az  $y$  tetszőleges volt, ezért ez az  $f$  állandóságát jelenti.

(4): Az előző állítást az  $f - g$  függvényre alkalmazva azonnal adódik következményként.  $\square$

Értelmezzük egy kicsit bővebben a harmadik és negyedik állítást. A deriválás egyik egyszerű szabálya szerint az állandó (konstans) függvény deriváltja nulla. A harmadik állítás ennek az állításnak a megfordítása: Ha valamely függvény deriváltja egy intervallumban nulla, akkor ott a függvény állandó. Természetesen azt nem tudjuk, hogy milyen állandó a függvény értéke, hiszen például az  $x \mapsto 5$  és  $x \mapsto -2$  függvények deriváltja egyaránt nulla.

A negyedik állítás, ami a megelőző állítás következménye, lehetővé teszi a deriválás műveletének bizonyos értelemben való megfordítását. Ha tudjuk ugyanis, hogy egy intervallumon értelmezett  $F$  függvénynek az  $f$  függvény a deriváltja, akkor az összes olyan függvény, aminek az  $f$  a deriváltja,

$$F + \text{konstans}$$

alakú. A tétel szerint ugyanis ha  $G' = f$ , akkor  $G - F = \text{konstans}$ , tehát a deriválás művelete egy konstanstól eltekintve megfordítható.

Például tudjuk, hogy az  $x \mapsto 2x$  az  $x \mapsto x^2$  deriváltja, ezek szerint azt mondhatjuk, hogy az a függvény, amelyiknek a deriváltja az  $x \mapsto 2x$  függvény,

$$y = x^2 + \text{konstans}$$

alakú.

A mondottak alapján felvetődik a deriválás megfordításának, inverzének a kérdése. Ezzel az “antideriválásnak” nevezett eljárással (művelettel) a következő fejezettől kezdve fogunk foglalkozni.

**Példa 6.1** *Vizsgáljuk meg, hogy az  $f(x) = x^3 - \frac{9x^2}{2} + 6x + 1$  polinom hol nő és hol fog.*

Az  $f$  polinom deriváltja az  $f'(x) = 3x^2 - 9x + 6 = 3(x^2 - 3x + 2)$  polinom, aminek a nullahelyei az  $x_1 = 2$  és  $x_2 = 1$ . Ezek szerint a deriváltpolinom gyöktényezőssé alakja  $f'(x) = 3(x-1)(x-2)$ . Mivel  $0 \leq x-1$  pontosan akkor, ha  $1 \leq x$ , és  $0 \leq x-2$  pontosan akkor ha  $2 \leq x$ , ezért az  $f'$  derivált előjele:

1. nemnegatív, ha  $2 \leq x$  vagy  $x \leq 1$ ,
2. nempozitív, ha  $1 \leq x \leq 2$ ,

tehát az  $f$  polinom monoton növekedő a  $(-\infty, 1]$  és  $[2, +\infty)$  intervallumokban, és monoton fogyó az  $[1, 2]$  intervallumban.  $\square$

**Példa 6.2** *Határozzuk meg azt az  $f$  függvényt, amelyiknek a deriváltja az  $x \mapsto \cos x + 3$  függvény, és a nulla pontban az értéke egy.*

Könnyen kitalálhatjuk, hogy a  $\cos x + 3$  függvény a  $\sin x + 3x$  függvény deriváltja. Így az előző tétel (4) állítása szerint azoknak az  $f$  függvényeknek az általános alakja, amelyeknek a deriváltja a  $\cos x + 3$  függvény,

$$g(x) = \sin x + 3x + \text{konstans}$$

alakú. Azt is megköveteltük azonban, hogy  $f(0) = 1$ , ami már meghatározza a konstans értékét, hiszen az

$$1 = f(0) = \sin 0 + 3 \cdot 0 + \text{konstans}$$

egyenletből a konstans értéke 1. Ezek alapján a keresett megoldás:  $x \mapsto \sin x + 3x + 1$ .  $\square$

A monotonitáshoz hasonlóan a deriválható függvények szigorú monotonitása is vizsgálható a derivált előjele alapján:

**Állítás 237** *Legyen az  $f : [a, b] \rightarrow \mathbb{R}$  függvény deriválható az intervallum belsejében és folytonos az  $a$  és  $b$  végpontokban is. Ekkor fennállnak a következő állítások:*

- (1) Ha  $f'(x) > 0$  minden  $x \in (a, b)$  pontban, akkor az  $f$  függvény szigorúan monoton növekedő az  $[a, b]$  intervallumban.
- (2) Ha  $f'(x) < 0$  minden  $x \in (a, b)$  pontban, akkor az  $f$  függvény szigorúan monoton fogyó az  $[a, b]$  intervallumban.

Az állítások megfordítása itt nem igaz, szigorúan monoton függvény deriváltja nem feltétlenül pozitív. Példaképpen vehetjük az  $x \mapsto x^3$  függvényt.

**Bizonyítás.** Az állítások mindegyike egyszerű következménye a középértéktételnek. Legyen az  $x < y$  két tetszőleges pontja az  $[a, b]$  intervallumnak. Az  $f$  függvény folytonos az  $[x, y]$  zárt és korlátos intervallumban, belül deriválható, ezért a középértéktétel szerint van olyan  $\xi$  pont az  $(x, y)$  intervallumban, hogy

$$f'(\xi)(y - x) = f(y) - f(x). \quad (6.4)$$

(1): Ha  $f'(\xi) > 0$  minden pontjára az intervallumnak, akkor (6.4)-ből következik, hogy

$$f(y) - f(x) > 0, \quad \text{ha} \quad x < y,$$

tehát az  $f$  szigorúan monoton növekedő.

(2): Hasonló az előző indokláshoz, vagy következik az előző állításból, ha azt az  $f$  helyett a  $-f$ -re alkalmazzuk.  $\square$

### 6.2.2 A szélsőérték elégséges feltételei

A szélsőértékek meghatározásához hasznos az alábbi tétel, amely a szükséges feltételek mellett elégséges feltételeket is ad.

**Állítás 238 (Szélsőérték, elégséges feltételek)** Ha az  $f$  függvény deriválható a  $b$  pontnak egy környezetében, akkor a  $b$  pontban

- (1) lokális (szigorú) maximuma van, ha a  $b$  pont egy környezetének a baloldali felében (szigorúan) nő, a jobboldali felében pedig (szigorúan) csökken:  
 $f'(x) \geq 0$  ( $f'(x) > 0$ ) a  $b$  egy baloldali környezetében és  $f'(x) \leq 0$  ( $f'(x) < 0$ ) egy jobboldali környezetében;
- (2) lokális (szigorú) minimuma van, ha a  $b$  pont egy környezetének a baloldali felében (szigorúan) fogy, a jobboldali felében pedig (szigorúan) nő:  
 $f'(x) \leq 0$  ( $f'(x) < 0$ ) a  $b$  egy baloldali környezetében és  $f'(x) \geq 0$  ( $f'(x) > 0$ ) egy jobboldali környezetében.

**Bizonyítás.** (1): Nézzük például a maximum esetét. Nyilvánvaló, hogy ha az  $f$  a  $b$  pont egy baloldali környezetében nő és egy jobboldali környezetében csökken, akkor a  $b$  pontban maximuma van. A 236. tétel első két állítása szerint ez pontosan azt jelenti, hogy a  $b$  pont egy baloldali környezetében  $f'(x) \geq 0$ , egy jobboldali környezetében pedig  $f'(x) \leq 0$ . Pontosán azonos a bizonyítás a szigorú maximum esetében, csak akkor a 237. tételre kell hivatkozni.  $\square$

A következő tételben a kétszer differenciálható függvények szigorú szélsőértékeinek a meghatározásához adunk elégséges feltételt.

**Állítás 239 (Szélsőérték, másodrendű feltételek)** *Ha az  $f$  leképezés kétszer deriválható a  $b$  pontban és  $f'(b) = 0$ , akkor*

- (i) *ha  $f''(b) < 0$ , akkor az  $f$ -nek szigorú maximuma van a  $b$  helyen;*
- (ii) *ha  $f''(b) > 0$ , akkor az  $f$ -nek szigorú minimuma van a  $b$  helyen.*

Jegyezzük meg, hogy az  $f'(b) = 0$  és  $f''(b) = 0$  esetben semmit sem tudunk mondani az elégségeségről további feltételek teljesülése nélkül. A magasabb rendű deriváltakra vonatkozó újabb feltételek mellett azonban finomítható lenne az állítás.

**Bizonyítás.** Vizsgáljuk meg, mondjuk, az  $f''(b) < 0$  esetet. Az  $f''(b) < 0$  egyenlőtlenségből következik, hogy az

$$\frac{f'(b+h) - f'(b)}{h} = \frac{f'(b+h)}{h}$$

különbségi hányados negatív, ha a  $|h|$  eléggé kicsi (van olyan  $\delta$  pozitív szám, hogy  $|h| < \delta$  esetében negatív a tört). Ebből viszont adódik, hogy

$$f'(b+h) < 0, \quad \text{ha } h > 0 \text{ és eléggé kicsi,}$$

$$f'(b+h) > 0, \quad \text{ha } h < 0 \text{ és eléggé kicsi.}$$

Ez utóbbiakból a 237. tétel szerint azt kapjuk, hogy az  $f$  függvény a  $b$  pont egy baloldali környezetében szigorúan nő, egy jobboldali környezetében pedig szigorúan csökken, ezért nyilvánvalóan szigorú maximuma van a  $b$  helyen.  $\square$

### 6.2.3 Differenciálható konvex függvények

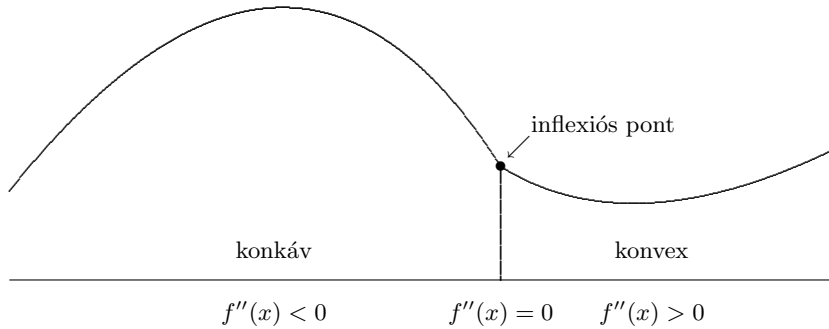
Most deriválható függvények esetében lehetőséget adunk arra, hogy a konvex illetve konkáv voltukat eldönthessük:

**Állítás 240** *Legyen az  $f$  leképezés folytonos az  $[a, b]$  intervallumon, az intervallum belsejében pedig deriválható.*

- (1) *Ha az  $f'$  derivált (szigorúan) monoton növekedő az intervallumon, akkor az  $f$  függvény (szigorúan) konvex az intervallumon.*
- (2) *Ha az  $f'$  derivált (szigorúan) monoton fogyó az intervallumon, akkor az  $f$  függvény (szigorúan) konkáv az intervallumon.*
- (3) *Ha az  $f$  függvény kétszer deriválható az intervallumon és az  $f''$  második derivált nemnegatív (pozitív) az intervallumban, akkor az  $f$  (szigorúan) konvex az intervallumon.*

- (4) Ha az  $f$  függvény kétszer deriválható az intervallumon és az  $f''$  második derivált nempozitív (negatív) az intervallumban, akkor az  $f$  (szigorúan) konkáv az intervallumon.

A differenciálható függvény monotonitására vonatkozó 236. tételben kerekébb volt az állításunk, mert ott szükséges és elégséges feltételt tudtunk mondani: Pontosán akkor monoton növekedő a differenciálható függvény, ha a deriváltja nemnegatív. A jelenlegi tételnek is van hasonló, erősebb alakja, amit egy következő tételként kimondunk.



6.5. ábra: Deriválható függvény konvexitása és konkávitása.

**Bizonyítás.** A bizonyítások előtt jegyezzük meg, hogy a (2) állítás az (1), a (4) állítás pedig a (3) állítás nyilvánvaló következménye, ezért csak az (1) és (3) szorulnak igazolásra.

(1):  $\lambda \in [0, 1]$  esetén vegyük az

$$\begin{aligned} & \lambda f(x) + (1 - \lambda)f(y) - f(\lambda x + (1 - \lambda)y) = \\ & = \lambda(f(x) - f(\lambda x + (1 - \lambda)y)) + (1 - \lambda)(f(y) - f(\lambda x + (1 - \lambda)y)) \end{aligned} \quad (6.5)$$

azonosságot. Megmutatjuk, hogy ennek az azonosságnak a jobb oldala nemnegatív (pozitív), ha az  $f'$  leképezés (szigorúan) monoton növekvő, ami pontosan azt adja, hogy az  $f$  függvény (szigorúan) konvex.

A bizonyítás gondolata egyszerűen az, hogy a (6.5) azonosság jobboldalán lévő két tag mindegyikét felírjuk a középértéktétel alapján, amivel behozzuk az  $f$  függvény deriváltjait.

A középértéktételnek megfelelően van olyan  $\xi_1$ , hogy

$$x < \xi_1 < \lambda x + (1 - \lambda)y, \quad (6.6)$$

és

$$f(\lambda x + (1 - \lambda)y) - f(x) = f'(\xi_1)(1 - \lambda)(y - x), \quad (6.7)$$

ahol a második egyenlőségnél felhasználtuk azt, hogy  $\lambda x + (1 - \lambda)y - x = (1 - \lambda)(y - x)$ . Hasonlóan járva el a (6.5) jobboldalának másik tagjára kapjuk, hogy van olyan  $\xi_2$ , amelyre

$$\lambda x + (1 - \lambda)y < \xi_2 < y, \quad (6.8)$$

és

$$f(y) - f(\lambda x + (1 - \lambda)y) = f'(\xi_2) \cdot \lambda(y - x), \quad (6.9)$$

ahol az utóbbi egyenlőség felírásánál felhasználtuk, hogy

$$y - (\lambda x + (1 - \lambda)y) = \lambda(y - x).$$

A (6.6) és (6.8) egyenlőségek alapján láthatjuk, hogy

$$\xi_1 < \xi_2. \quad (6.10)$$

A (6.5) azonosság jobboldalán a (6.7) és (6.9) szerint helyettesítve a tagokat a következőképpen számolhatunk:

$$\begin{aligned} \lambda f(x) + (1 - \lambda)f(y) - f(\lambda x + (1 - \lambda)y) &= \\ &= -\lambda(1 - \lambda)f'(\xi_1)(y - x) + \lambda(1 - \lambda)f'(\xi_2)(y - x) = \\ &= \lambda(1 - \lambda)(y - x)(f'(\xi_2) - f'(\xi_1)), \end{aligned}$$

amiből a (6.10) alapján azonnal kapjuk azt, hogy az  $f$  függvény (szigorúan) konvex, ha az  $f'$  függvény (szigorúan) monoton növekedő, amit bizonyítanunk kellett.

(3): Ha az  $f''$  létezik és (pozitív) nemnegatív, akkor a 237. állítás szerint az  $f'$  függvény (szigorúan) monoton, és így a jelen tétel (1) állítása alapján készen is vagyunk.  $\square$

Fontos szerepük van még azoknak a pontoknak, ahol a függvény konvexből konkávba (vagy megfordítva) megy át, ezért ezeket el is nevezzük:

**Definíció 241** *Ha egy  $f$  függvény az  $x$  pont valamilyen baloldali környezetében konkáv és valamilyen jobboldali környezetében konvex, vagy megfordítva, akkor az  $x$  pontot inflexiós pontnak mondjuk.*

Egyszer illetve kétszer deriválható függvény esetében sokszor eldönthetjük, hogy valamely pont inflexiós pont-e, a következő tétel segítségével.

**Állítás 242** *Az inflexiós pont meghatározására lehetőséget adnak a következő állítások:*

- (a) *Ha az  $f$  deriválható az  $x$  környezetében, és az  $f'$  monotonitást vált az  $x$  pontban (növekedőből fogyóba vagy fogyóból növekedőbe megy át), akkor az  $x$  pont inflexiós pont.*
- (b) *Ha az  $f$  kétszer deriválható az  $x$  pont egy környezetében, és az  $f''$  második derivált előjelet vált (pozitívból negatívba vagy negatívból pozitívba megy át), akkor az  $x$  inflexiós pont.*

Bizonyítás. Közvetlen következménye az előző tételnek.  $\square$

Az alpont hátralévő részében konvex függvények deriválásával kapcsolatos további tételeket tárgyalunk.

**Állítás 243** Legyen  $f : (a, b) \rightarrow \mathbb{R}$  egy konvex függvény. Ekkor tetszőleges  $x \in (a, b)$  pontban léteznek az

$$f'_-(x) = \lim_{h \searrow 0} \frac{f(x+h) - f(x)}{h}$$

baloldali és az

$$f'_+(x) = \lim_{h \nearrow 0} \frac{f(x+h) - f(x)}{h}$$

jobboldali deriváltak, és

$$f'_-(x) \leq f'_+(x).$$

Bizonyítás. A 233. állítás szerint az

$$y \mapsto \frac{f(y) - f(x)}{y - x}$$

differenciahányados monoton növekedő, ezért tetszőleges  $x \in (a, b)$  esetén létezik mind a baloldali  $f'_-(x)$ , mind a jobboldali  $f'_+(x)$  határértéke, mégpedig

$$f'_-(x) = \sup_{y < x} \frac{f(y) - f(x)}{y - x} \quad \text{és} \quad f'_+(x) = \inf_{y > x} \frac{f(y) - f(x)}{y - x}.$$

Ezekből azonnal adódik az állított egyenlőtlenség is.  $\square$

**Állítás 244** Legyen  $f : (a, b) \rightarrow \mathbb{R}$  nyílt intervallumon értelmezett differenciálható függvény. Az  $u \in (a, b)$  pontbeli érintőt jelölje  $l_u : (a, b) \rightarrow \mathbb{R}$ ,

$$l_u(x) := f(u) + f'(u)(x - u).$$

Ekkor az alábbi három állítás ekvivalens:

1. Az  $f$  függvény konvex.
2. Az  $f'$  deriváltfüggvény monoton növekvő.
3. Tetszőleges pontbeli érintő a függvény gráfja alatt fekszik, azaz minden  $u \in (a, b)$  és minden  $x \in (a, b)$  esetén

$$l_u(x) \leq f(x).$$

Bizonyítás. Először tegyük fel, hogy az  $f$  függvény konvex, és lássuk be, hogy  $f'$  deriváltfüggvény monoton növekvő. Mivel konvex függvény különbséghányados-függvénye monoton növekvő, így tetszőleges  $x_1 < x_2$  esetén

$$f'(x_1) = \lim_{x \rightarrow x_1} \frac{f(x) - f(x_1)}{x - x_1} = \inf_{x > x_1} \frac{f(x) - f(x_1)}{x - x_1} \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

és hasonlóan

$$f'(x_2) = \lim_{x \rightarrow x_2} \frac{f(x) - f(x_2)}{x - x_2} = \sup_{x < x_2} \frac{f(x) - f(x_2)}{x - x_2} \geq \frac{f(x_1) - f(x_2)}{x_1 - x_2}$$

amiből

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

Éppen ezt akartuk bizonyítani.

Most tegyük fel, hogy az  $f'$  deriváltfüggvény monoton növekvő, és lássuk be, hogy az érintők a gráf alatt fekszenek. Tetszőleges, de a továbbiakban rögzített  $u \in (a, b)$  esetén jelölje  $h : (a, b) \rightarrow \mathbb{R}$  a következő függvényt:

$$h := f - l_u.$$

Világos, hogy egyrészt  $h(u) = 0$ , másrészt  $h$  differenciálható az  $(a, b)$  minden pontjában, és  $f'$  monotonitása miatt  $h' \leq 0$  az  $(a, u]$  intervallumon, valamint  $h' \geq 0$  az  $[u, b)$  intervallumon. Ez viszont azt jelenti, hogy a  $h$  függvény monoton csökken az  $(a, u]$ , monoton nő az  $[u, b)$  intervallumon, és  $u$ -ban éppen nulla, azaz  $h \geq 0$  az egész  $(a, b)$  intervallumon, de éppen ezt kellett belátni.

Namármost tegyük fel, hogy az érintők a függvény gráfja alatt vannak, és mutassuk meg, hogy  $f'$  monoton növekvő. Először megmutatjuk, hogy minden  $x, u, y \in (a, b)$  és  $x < u < y$  esetén

$$F_u(x) \leq f'(u) \leq F_u(y), \quad (6.11)$$

ahol

$$F_u(x) := \frac{f(x) - f(u)}{x - u}$$

az  $u$  ponthoz tartozó különbséghányados-függvényt jelöli. Nézzük először az  $u$  és  $y$  vonatkozásában vett egyenlőtlenséget:

$$F_u(y) = \frac{f(y) - f(u)}{y - u} \geq \frac{l_u(y) - f(u)}{y - u} = \frac{f'(u)(y - u)}{y - u} = f'(u).$$

Az egyenlőtlenséget az indokolja, hogy  $l_u(y) \leq f(y)$ , azaz a függvény az érintő fölött fekszik, másrészt  $y - u \geq 0$ . Ehhez hasonlóan az  $x$  és  $u$  vonatkozásában

$$F_u(x) = \frac{f(u) - f(x)}{u - x} \leq \frac{f(u) - l_u(x)}{u - x} = \frac{-f'(u)(x - u)}{u - x} = f'(u),$$

ahol szintén azt használtuk ki, hogy  $l_u(x) \leq f(x)$ , valamint azt, hogy  $u - x \geq 0$ . Most már könnyen befejezhetjük a bizonyítást, hiszen tetszőleges  $x, y \in (a, b)$ ,  $x < y$  esetén az imént belátott (6.11) -et kétszer alkalmazva azt kapjuk, hogy

$$f'(x) \leq F_x(y) = F_y(x) \leq f'(y),$$

ami épp azt jelenti, hogy  $f'$  az  $(a, b)$  intervallumon monoton nő.

Az  $f'$  deriváltfüggvény növekedéséből  $f$  konvexitása már következik az ismert (240) tétel szerint.

□

### 6.2.4 Az Euler-szám és az exponenciális függvény

Jelölje  $a \in \mathbb{R}$ ,  $a > 0$ ,  $a \neq 1$  esetén  $f_a : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f_a(x) := a^x$ . A célunk az, hogy megmutassuk, hogy ez a függvény differenciálható az egész valós számegyenesen, és kiszámoljuk a derivált függvényt. Ehhez szükségünk van olyan, majd a továbbiakban  $e$ -nek (Euler) nevezett számra, melyre

$$\lim_{h \rightarrow 0} \frac{e^h - 1}{h}$$

létezik és egyenlő 0 -val.

Az  $e$  szám bevezetése az alábbi lépésekből áll:

1. Megmutatjuk, hogy az  $f_a$  konvex függvény.
2. Megmutatjuk, hogy az  $f_a$  differenciálható a  $0 \in \mathbb{R}$  pontban.
3. Megmutatjuk, hogy az  $f_a$  differenciálható az egész  $\mathbb{R}$ -en, és

$$f'_a(x) = f'_a(0) \cdot f_a(x) \tag{6.12}$$

minden  $x \in \mathbb{R}$  esetén.

4. Megmutatjuk, hogy van egyetlen olyan  $e \in \mathbb{R}$ , melyre

$$f'_e(0) = 1.$$

Most nézzük az egyes pontok részletezését.

**Állítás 245**  $f_a$  konvex.

**Bizonyítás.** Azt kell megmutatni, hogy tetszőleges  $x, y \in \mathbb{R}$  és tetszőleges  $\lambda \in [0, 1]$  esetén

$$a^{\lambda x + (1-\lambda)y} \leq \lambda a^x + (1-\lambda) a^y.$$

Mutassuk meg ezt először racionális  $\lambda$  mellett. Nyilván, ha  $\lambda \in [0, 1]$  racionális szám, akkor léteznek  $p, q \in \mathbb{N}$  egész számok, melyekre

$$\lambda = \frac{p}{p+q} \text{ és } 1 - \lambda = \frac{q}{p+q}.$$

A számtani és mértani közép közti egyenlőtlenség szerint

$$a^{px+qy} = \overbrace{a^x \cdot \dots \cdot a^x}^{p\text{-szor}} \cdot \overbrace{a^y \cdot \dots \cdot a^y}^{q\text{-szor}} \leq \left( \frac{pa^x + qa^y}{p+q} \right)^{p+q},$$

amiből  $p+q$  adik gyököt vonva kapjuk, hogy

$$a^{\frac{p}{p+q}x + \frac{q}{p+q}y} \leq \frac{p}{p+q}a^x + \frac{q}{p+q}a^y,$$

amit éppen bizonyítani kellett a racionális esethez.

Az általános eset bizonyításához tegyük fel, hogy  $a > 1$  és  $y < x$ . Ekkor olyan  $r$  racionális számra, melyre  $\lambda < r$ ,

$$\lambda x + (1 - \lambda)y < rx + (1 - r)y,$$

amiből a hatványfüggvény szigorú monoton növése, és a racionális esetre már bizonyított egyenlőtlenség miatt

$$\begin{aligned} a^{\lambda x + (1 - \lambda)y} &\leq \inf \left\{ a^{rx + (1 - r)y} : r > \lambda, r \in Q \right\} \leq \\ &\leq \inf \left\{ ra^x + (1 - r)a^y : r > \lambda, r \in Q \right\} = \\ &= \lambda a^x + (1 - \lambda)a^y, \end{aligned}$$

és éppen ezt kellett belátni. Az  $a < 1$  eset fentihez hasonló bizonyítását az olvasóra bízunk.  $\square$

**Állítás 246**  $f_a$  differenciálható 0-ban.

Bizonyítás. Mivel  $f_a$  konvex, ezért létezik bal- és jobboldali deriváltja. Megmutatjuk, hogy a baloldali derivált egyenlő a jobboldali deriválttal:

$$\begin{aligned} f'_-(0) &= \lim_{h \searrow 0} \frac{a^h - 1}{h} = \lim_{h \searrow 0} \frac{a^{-h} - 1}{-h} = \lim_{h \searrow 0} \frac{1 - a^h}{-ha^h} = \\ &= \lim_{h \searrow 0} \left( \frac{a^h - 1}{h} \cdot \frac{1}{a^h} \right) = \lim_{h \searrow 0} \frac{a^h - 1}{h} \cdot \lim_{h \searrow 0} \frac{1}{a^h} = f'_+(0) \cdot 1 = f'_+(0). \end{aligned}$$

$\square$

**Állítás 247**  $f_a$  differenciálható tetszőleges  $x \in \mathbb{R}$  pontban és

$$f'_a(x) = f'_a(0) \cdot f_a(x).$$

Bizonyítás. Vegyük észre, hogy

$$\frac{a^{x+h} - a^x}{h} = \frac{a^x a^h - a^x}{h} = a^x \cdot \frac{a^h - 1}{h}$$

amiből

$$\lim_{h \rightarrow 0} \frac{a^{x+h} - a^x}{h} = a^x \cdot \lim_{h \rightarrow 0} \frac{a^h - 1}{h}$$

ami azt jelenti, hogy  $f_a$  differenciálható minden  $x \in \mathbb{R}$  pontban és

$$f'_a(x) = f'_a(0) \cdot f_a(x).$$

□

**Állítás 248** Van olyan  $e \in \mathbb{R}$  szám, melyre

$$f'_e(0) = 1.$$

Bizonyítás. Rögzítsünk egy tetszőleges  $a$  pozitív valós számot, melyre  $a \neq 1$ . Világos, hogy minden  $t \in \mathbb{R}$  mellett

$$f_{a^t}(x) = (a^t)^x = a^{t \cdot x} = f_a(tx),$$

így a kompozíciófüggvény deriválási szabálya miatt

$$f'_{a^t}(x) = f'_a(tx) \cdot t$$

minden  $x \in \mathbb{R}$  esetén, így speciálisan  $x = 0$ -ra is

$$f'_{a^t}(0) = f'_a(0) \cdot t. \quad (6.13)$$

Namármost nyilván rögzített  $a \neq 1$  mellett  $f'_a(0) \neq 0$  (egyébként 6.12 miatt  $f_a$  konstans lenne), így  $t^* = \frac{1}{f'_a(0)}$  választás mellett

$$f'_{a^{t^*}}(0) = f'_a(0) \cdot t^* = 1.$$

Tehát ha bevezetjük az  $e = a^{t^*}$  jelölést, akkor találtunk  $e \in \mathbb{R}$  valós pozitív számot, melyre

$$f'_e(0) = 1.$$

□

**Állítás 249** Csak egy olyan  $e$  valós szám van, melyre  $f'_e(0) = 1$ .

Bizonyítás. Legyenek  $e_1, e_2$  valós számok, melyekre

$$f'_{e_1}(0) = 1 = f'_{e_2}(0).$$

Ekkor 6.12 miatt

$$f'_{e_1} = f_{e_1} \text{ és } f'_{e_2} = f_{e_2},$$

amiből

$$\left(\frac{f_{e_1}}{f_{e_2}}\right)' = \frac{f'_{e_1}f_{e_2} - f'_{e_2}f_{e_1}}{f_{e_2}^2} = \frac{f_{e_1}f_{e_2} - f_{e_2}f_{e_1}}{f_{e_2}^2} = 0,$$

ahonnan az következik, hogy valamely  $c \in \mathbb{R}$  mellett

$$f_{e_1} = cf_{e_2}.$$

De mivel  $f_{e_1}(0) = f_{e_2}(0) = 1$ , így  $c$  nyilván csak 1 lehet, tehát  $f_{e_1} = f_{e_2}$ , innen

$$e_1 = e_2$$

következik.  $\square$

Az  $e$  alapú hatványfüggvényt természetes alapú exponenciális függvénynek szokták nevezni és  $e^x$ -szel vagy  $\exp$ -pel szokták jelölni. A fentiek szerint ez differenciálható minden valós pontban és

$$\frac{d}{dx}e^x = e^x,$$

vagy ami ugyanazt jelenti,

$$\exp' = \exp.$$

Az exponenciális függvény szigorúan monoton, folytonos, sőt differenciálható, így a logaritmusfüggvény -  $\ln$  - mint az exponenciális függvény inverze is monoton, folytonos, sőt deriválható. Számoljuk ki a deriváltfüggvényét!

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

Azaz minden  $x \in \mathbb{R}$ ,  $x > 0$  mellett

$$\frac{d}{dx} \ln x = \frac{1}{x}.$$

Vegyük észre, hogy ha (6.13) egyenlőségbe  $t = \frac{1}{\ln a}$ -t helyettesítünk, akkor az

$$f'_a(0) = (\ln a) \cdot f'_e(0) = \ln a$$

egyenlőtlenséget kapjuk, ami az  $f_a$  függvény deriválási szabálya alapján azt jelenti, hogy minden  $a \in \mathbb{R}_+$ ,  $a \neq 1$  esetén

$$\frac{d}{dx}a^x = a^x \cdot \ln a.$$

Innen könnyen kiszámolható, hogy

$$\frac{d}{dx} \log_a x = \frac{1}{x \cdot \ln a}$$

Most megoldunk néhány példát, amely az exponenciális és logaritmusfüggvényre vonatkozik:

**Példa 6.3** Határozzuk meg azt a pontot az  $[1, 2]$  intervallumban, ahol az  $e^x$  függvény érintője párhuzamos az  $(1, e)$  és  $(2, e^2)$  pontokat összekötő szelővel.

Az intervallum végpontjai felett átmenő szelőnek az iránytangense az  $\frac{e^2 - e}{2 - 1} = e(e - 1)$  differenciahányados, ami megegyezik egy  $e^\xi$  ( $\xi \in (1, 2)$ ) deriválttal. Tehát  $e^\xi = e(e - 1)$ , amiből  $\xi = \ln(e(e - 1)) = 1 + \ln(e - 1)$ .  $\square$

**Példa 6.4** Adjuk meg az  $e^x$  függvény  $n$ -edik Taylor-közelítését a nulla körül, és becsüljük meg, hogy milyen pontossággal tudjuk meghatározni az  $e$  szám értékét a 9-edik közelítésből.

Az  $n$ -edik közelítést kevés függvéynél tudjuk könnyen kiszámolni, mivel a magasabb rendű deriváltak nagyon bonyolultak lehetnek. A jelen esetben szerencsések vagyunk, hiszen az  $e^x$  függvény minden deriváltja is  $e^x$ , ezért az első feltett kérdésre a válasz: Az  $e^x$  függvény  $n$ -edik Taylor-approximációja

$$1 + \frac{x}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}.$$

Az előző tétel szerint van olyan  $\xi \in (0, x)$ , hogy

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + \frac{e^\xi}{(n+1)!} x^{n+1}.$$

Ennek az egyenlőségnek a birtokában már tudunk adni egy előállítást az  $e$  számra, hiszen az előző szerint az  $x = 1$  helyen

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \frac{e^\xi}{(n+1)!},$$

ahol a  $\xi$  egy a 0 és 1 között lévő szám. Az

$$\frac{e^\xi}{(n+1)!}$$

hibatag becslésében csak az a probléma, hogy elvileg az  $e$  szám nagyságáról nincs információnk, csak közöltük a közelítő értékét. Az előző előállításból azonban könnyen kaphatunk egy felső becslést. Ha felírjuk az előző előállítást az  $n = 1$  esetben, akkor

$$e = 1 + \frac{1}{1!} + \frac{e^\xi}{2!} = 2 + \frac{e^\xi}{2},$$

és mivel  $\xi < 1$ ,  $e^\xi < e$ , és ezért  $e < 2 + e/2$ , innen azonnal kapjuk, hogy  $e < 4$ . Ennek a segítségével az  $e$  fenti előállításában a hibatag becslése az  $n = 9$  esetben (ahol  $\alpha \in (0, 1)$ ):

$$\frac{e^\alpha}{10!} < \frac{e}{10!} < \frac{4}{10!} < 4/3, 628, 800 < 0.000001,$$

az  $e$  szám értékét tehát egymilliomodnál kisebb hibával adja meg az

$$1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{9!}$$

összeg. □

**Példa 6.5** *Határozzuk meg a*

$$\lim_{x \rightarrow +\infty} \frac{x^n}{e^x}$$

*limeszt.*

A jelen esetben  $\frac{\infty}{\infty}$  “alakú” a tört. A számláló és nevező deriváltjának a hányadosa

$$\frac{nx^{n-1}}{e^x},$$

amelyik még ugyanolyan tulajdonságú, mint a kiinduló tört volt, feltéve, hogy  $n > 1$ , ezért ismételtén alkalmazhatjuk a L'Hospital szabályt. Végül is  $n$ -szer kell deriválni a számlálót és nevezőt, ami után az

$$\frac{n(n-1) \cdots 2 \cdot 1}{e^x}$$

alakhoz jutunk, aminek a limesze a plusz végtelenben nulla, ezért nulla a keresett limesz is. □

A következőkben két olyan limeszt határozunk meg, amelyik nem tört, de kedvező törtalakra hozható a logaritmusfüggvény segítségével.

**Példa 6.6** *Számoljuk ki a következő limeszt.*

$$\lim_{x \searrow 0} (\sin x)^{\sin x}$$

A  $(\sin x)^{\sin x}$  függvény a nullában “ $0^0$ ” “határozatlan alakú” (megállapodás szerint  $b^0 = 1$ , ha  $b \neq 0$ ). Ezt jól kezelhető törtalakra hozhatjuk, ha vesszük a logaritmusát, amit megtehetünk, mert a nulla jobboldali környezetében pozitív. Eszerint az

$$\ln(\sin x)^{\sin x} = \sin x \ln(\sin x) = \frac{\ln(\sin x)}{\frac{1}{\sin x}}$$

függvény jobboldali limeszét kell meghatározni. A jobboldali tört nevezőjének a deriváltja a

$$\frac{d}{dx} \frac{1}{\sin x} = -\frac{\cos x}{\sin^2 x},$$

ami nem nulla a 0 egy jobboldali környezetében. A számláló deriváltja

$$\frac{d}{dx} \ln(\sin x) = \frac{1}{\sin x} \sin' x = \frac{\cos x}{\sin x}.$$

A számláló és nevező deriváltjának a hányadosa

$$-\frac{\cos x}{\sin x} \bigg/ \frac{\cos x}{\sin^2 x} = -\sin x,$$

aminek van jobboldali limesze a nullában, nevezetesen a 0, ezért az  $\ln$  és  $\exp$  függvények folytonossága miatt

$$\lim_{x \searrow 0} (\sin x)^{\sin x} = \lim_{x \searrow 0} \exp \left( \frac{\frac{d}{dx} \ln(\sin x)}{\frac{d}{dx} \frac{1}{\sin x}} \right) = e^0 = 1.$$

□

**Példa 6.7**

$$\lim_{x \searrow 0} (1+x)^{\ln x} = ?$$

A függvény logaritmusa

$$(\ln x) \ln(1+x) = \frac{\ln(1+x)}{1/\ln x}.$$

Most egymásután többször alkalmazzuk a L'Hospital-szabályt. (Közben a tétel alkalmazhatóságát mindig ellenőrizni kellene.)

$$\begin{aligned} \lim_{x \searrow 0} \frac{\ln(1+x)}{1/\ln x} &= \lim_{x \searrow 0} \frac{1/(1+x)}{-1/(x \ln^2 x)} = - \lim_{x \searrow 0} \frac{x \ln^2 x}{x+1} = \\ &= - \lim_{x \searrow 0} \frac{\ln^2 x}{1+1/x} = - \lim_{x \searrow 0} \frac{(2 \ln x)/x}{-1/x^2} = 2 \lim_{x \searrow 0} \frac{\ln x}{1/x} = 2 \lim_{x \searrow 0} \frac{1/x}{-1/x^2} = 0. \end{aligned}$$

Tehát azt kaptuk, hogy a keresett limesz  $e^0 = 1$ .

□

A következő példa eredményét érdemes megjegyezni, ezért állításban fogalmazzuk meg.

**Állítás 250**

$$\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = 1.$$

Bizonyítás. Mivel

$$\frac{\ln(1+x)}{x} = \frac{\ln(1+x) - \ln 1}{x},$$

ezért a tört, aminek a limeszét keressük, differenciahányados. Emiatt a limesz értéke az  $\ln$  függvény deriváltjának az  $x = 1$  helyen vett értéke, azaz 1. □

**Példa 6.8** Határozzuk meg az  $f(x) = e^{-x^2}$  függvény szélsőértékeit.

A függvény deriváltja  $f'(x) = -2x \exp(-x^2)$ , aminek egyetlen nullahelye van az  $x = 0$ , mivel az exponenciális függvény mindig pozitív. A második derivált

$$f''(x) = -2 \exp(-x^2) + (-2x)(-2x \exp(-x^2)) = -2(1 - 2x^2)e^{-x^2},$$

ami az  $x = 0$  helyen negatív, hiszen  $f''(0) = -2$ . Ezek szerint a nullánál maximuma van a függvénynek, ahol a függvényérték 1. Több szélsőértéke nem lehet, mivel nincs több gyöke a deriválnak.  $\square$

**Példa 6.9** *Hol konvex és hol konkáv az  $y = e^{-x^2}$  függvény?*

Az első derivált  $y' = -2xe^{-x^2}$  és ebből a második derivált

$$y'' = 2(2x^2 - 1)e^{-x^2}.$$

Ennek a függvénynek nullhelyei az  $1/\sqrt{2}$  és a  $-1/\sqrt{2}$  számok. Ennek ismeretében:

$$y'' = 2(x\sqrt{2} - 1)(\sqrt{2}x + 1)e^{-x^2}.$$

Ebből azonnal látható, hogy az  $y = e^{-x^2}$  függvény az  $-1/\sqrt{2}$  pontig konvex, az  $[-1/\sqrt{2}, 1/\sqrt{2}]$  intervallumban konkáv, az  $1/\sqrt{2}$  ponttól jobbra pedig újra konvex. Az inflexiós pontok:  $\frac{\sqrt{2}}{2}$  és  $-\frac{\sqrt{2}}{2}$ .  $\square$

### 6.2.5 Konvexitási egyenlőtlenségek

Ebben az alponthban három nevezetes egyenlőtlenséget fogunk belátni, amelyek a logaritmusfüggvény konkavitásán alapszanak. Az első állítás ezt rögzíti.

**Állítás 251** *A logaritmusfüggvény a pozitív számok halmazán konkáv.*

Az  $\ln x$  függvény első deriváltja az  $x \mapsto 1/x$  függvény, a második derivált pedig

$$\frac{d}{dx} \frac{1}{x} = -\frac{1}{x^2}$$

függvény, ami minden helyen negatív, ezért a 240 tétel szerint az  $\ln$  függvény konkáv a pozitív számok halmazán.  $\square$

Ennek az egyszerűen kapott eredménynek a segítségével egy nevezetes egyenlőtlenséget bizonyítunk be:

**Állítás 252** *Tetszőleges  $x_1, x_2, \dots, x_n$  pozitív és  $\alpha_1, \alpha_2, \dots, \alpha_n, \sum_{i=1}^n \alpha_i = 1$  nemnegatív számokra teljesül az*

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \geq x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

*általánosított számtani és mértani közép közötti egyenlőtlenség.*

Az  $\alpha_i = \frac{1}{n}$ ,  $i = 1, \dots, n$  esetben az egyenlőtlenség:

$$\frac{x_1 + x_2 + \dots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \dots x_n}.$$

Ezt a speciális esetet a második fejezetben már igazoltuk.

**Bizonyítás.** Írjuk fel azt, hogy mit jelent a logaritmusfüggvény konkávitására a Jensen-egyenlőtlenség (232 tétel):

$$\ln(\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n) \geq \alpha_1 \ln x_1 + \alpha_2 \ln x_2 + \dots + \alpha_n \ln x_n,$$

amit rendezve az

$$\ln(\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n) \geq \ln(x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n})$$

egyenlőtlenséghez jutunk. Az  $\ln$  függvény inverze (az  $e^x$ ) szigorúan monoton, ezért az utóbbiból az

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \geq x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

egyenlőtlenséget kapjuk.  $\square$

**Állítás 253 (Young-egyenlőtlenség)** Legyen  $a$   $p$  és  $q$  két olyan pozitív szám, amelyekre

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Ekkor tetszőleges  $x$  és  $y$  pozitív számokra fennáll az

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q}$$

egyenlőtlenség.

**Bizonyítás.** A számtani és mértani közép közötti egyenlőtlenséget alkalmazva:

$$\frac{x^p}{p} + \frac{y^q}{q} \geq (x^p)^{\frac{1}{p}} \cdot (y^q)^{\frac{1}{q}} = xy.$$

$\square$

**Állítás 254 (Hölder-egyenlőtlenség)** Legyen  $a$   $p$  és  $q$  két olyan pozitív szám, amelyekre

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Ekkor tetszőleges

$$x_1, x_2, \dots, x_n \quad \text{és} \quad y_1, y_2, \dots, y_n$$

nemnegatív számok mellett fennáll a következő egyenlőtlenség:

$$\sum_{i=1}^n x_i y_i \leq \left( \sum_{i=1}^n x_i^p \right)^{\frac{1}{p}} \left( \sum_{i=1}^n y_i^q \right)^{\frac{1}{q}}.$$

Bizonyítás. A 253. állítás egyenlőtlenségét alkalmazzuk az

$$\frac{x_i}{\left(\sum_{j=1}^n x_j^p\right)^{\frac{1}{p}}} \quad \text{és} \quad \frac{y_i}{\left(\sum_{j=1}^n y_j^q\right)^{\frac{1}{q}}}, \quad i = 1, \dots, n$$

számokra és összegezzük  $i$  szerint:

$$\begin{aligned} \frac{\sum_{i=1}^n x_i y_i}{\left(\sum_{i=1}^n x_i^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^n y_i^q\right)^{\frac{1}{q}}} &= \sum_{i=1}^n \frac{x_i}{\left(\sum_{j=1}^n x_j^p\right)^{\frac{1}{p}}} \cdot \frac{y_i}{\left(\sum_{j=1}^n y_j^q\right)^{\frac{1}{q}}} \leq \\ &\leq \frac{1}{p} \sum_{i=1}^n \frac{x_i^p}{\sum_{j=1}^n x_j^p} + \frac{1}{q} \sum_{i=1}^n \frac{y_i^q}{\sum_{j=1}^n y_j^q} = \\ &= \frac{1}{p} \frac{\sum_{i=1}^n x_i^p}{\sum_{j=1}^n x_j^p} + \frac{1}{q} \frac{\sum_{i=1}^n y_i^q}{\sum_{j=1}^n y_j^q} = \\ &= \frac{1}{p} + \frac{1}{q} = 1. \end{aligned}$$

A számolássorozat elejét és végét nézve azonnal adódik az állítás.  $\square$

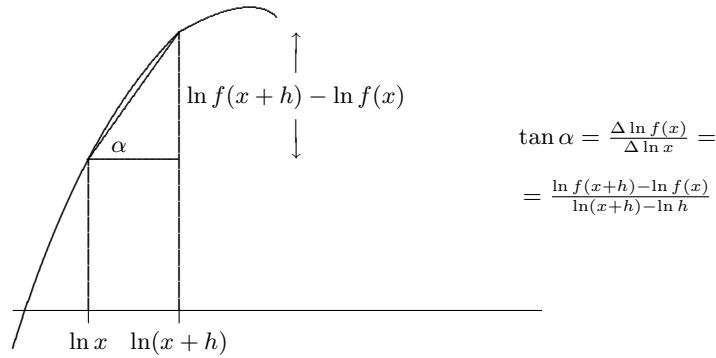
### 6.2.6 Az elaszticitás és a logaritmikus derivált

Valamely függvény megszokott ábrázolásánál az  $x$  tengelyre a független változó, az  $y$  tengelyre pedig a függvény  $y = f(x)$  értékeit rajzoljuk, és az  $(x, f(x))$  pontok összessége adja az  $f$  leképezés grafikonját.

Némelykor az  $x$  számunkra érdekes értékei a rendkívül széles tartományban változnak, és elfogadható méretű ábra elkészítése nehézséget jelent. Például ha mondjuk az  $x$  értékeit egytől tízmilliárdig akarjuk felrajzolni, akkor mindenképpen problémába ütközünk. Nem lehet az egynek az egy centimétert megfeleltetni, mert ekkor nem férnénk el, de vehetjük a milliárdot egy centiméternek, ekkor viszont az egységet nem tudjuk ábrázolni. A skálázásnak ezt a módját lineárisnak mondtuk, mivel azt jelenti, hogy a tengelyen felvett egységeket valamivel szorozzuk (az előbbi esetben az  $1/10^9$  értékkel).

Próbálkozhatunk viszont másképpen is: Az  $x$  értékei helyett azok tízes alapú logaritmusaikat rajzoljuk az  $x$  tengelyre, és ekkor is elfér a rajz tíz centiméter széles papíron. Ne feledjük azonban, hogy ekkor az  $x$  1, 10, 100, ...,  $10^9$  értékeihez a 0, 1, 2, ..., 9 értékeket rajzoljuk a papírra. Így az  $(x, f(x))$  pontok helyett a  $(\log x, f(x))$  pontok adják a számunkra alkalmas gráfot, ami nem a függvény gráfja, hanem annak egy transzformált alakja. Ezt a módszert logaritmikus skálázásnak szokás nevezni. Természetesen az  $y$  tengelyen is megtehetjük az előzőekben mondotakat, és ekkor a  $(\log x, \log(f(x)))$  pontok összessége adja a függvény egy transzformált gráfját. Mivel csak pozitív számnak van logaritmusa, természetes, hogy ekkor a  $0 < x$  és  $0 < f(x)$  feltételek elengedhetetlenek.

A következő tétel azt mutatja, hogy az elaszticitás bizonyos értelemben közönséges deriválttá válik, ha a logaritmikus skálát alkalmazzuk mind az  $x$  mind az  $y$  tengelyen.



6.6. ábra: Az elaszticitás mint differenciahányados limesze.

**Állítás 255** Legyen az  $f$  pozitív függvény definiálva az  $0 < x$  pont egy környezetében. Ha az  $f$  deriválható az  $x$  helyen, akkor az  $x$  helyen vett elaszticitása megegyezik a

$$\lim_{h \rightarrow 0} \frac{\ln(f(x+h)) - \ln(f(x))}{\ln(x+h) - \ln x}$$

limesszel, amit

$$\frac{d \ln(f(x))}{d \ln x}$$

módon szokás jelölni.

A 6.6. ábrán szemléltettük az állítást. Szavakban röviden a tétel:

*Az elaszticitás a függvény növekedésének egy olyan jellemzője, amelyik a függvény logaritmusa növekedésének a sebességét méri a független változó logaritmikus megváltozása mellett.*

**Bizonyítás.** Az igazolás könnyen adódik az alábbi számolás alapján:

$$\begin{aligned} \lim_{h \rightarrow 0} \frac{\ln(f(x+h)) - \ln(f(x))}{\ln(x+h) - \ln x} &= \lim_{h \rightarrow 0} \frac{\frac{\ln(f(x+h)) - \ln(f(x))}{h}}{\frac{\ln(x+h) - \ln x}{h}} = \\ &= \frac{(\ln \circ f)'(x)}{\ln' x} = \frac{x \cdot f'(x)}{f(x)}. \end{aligned}$$

□

### 6.2.7 Kalkulus-összefoglaló

A következőkben néhány táblázatban összefoglaljuk azokat a szabályokat, amelyek memorizálása szükséges a kalkulus alapos elsajátításához. A táblázatok emlékeztetők, és a pontos feltételeket a megfelelő tételek tartalmazzák.

#### FORMÁLIS SZABÁLYOK

$(\alpha f)'(x)$	$=$	$\alpha f'(x)$
$(f + g)'(x)$	$=$	$f'(x) + g'(x)$
$(fg)'(x)$	$=$	$f'(x)g(x) + f(x)g'(x)$
$\left(\frac{f}{g}\right)'(x)$	$=$	$\frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}$
$(f^{-1})'(f(x))$	$=$	$\frac{1}{f'(x)}$
$(g \circ f)'(x)$	$=$	$g'(f(x))f'(x)$

#### SPECIÁLIS FÜGGVÉNYEK DERIVÁLTJAI

$(x^n)'$	$=$	$nx^{n-1}$
$(e^x)'$	$=$	$e^x$
$\ln' x$	$=$	$\frac{1}{x}$
$\sin' x$	$=$	$\cos x$
$\cos' x$	$=$	$-\sin x$
$\tan' x$	$=$	$\frac{1}{\cos^2 x}$
$\cot' x$	$=$	$-\frac{1}{\sin^2 x}$

A következő táblázatban két olyan speciális függvény deriváltjára emlékeztetünk, amelyek ritkábban fordulnak elő, de igen fontosak lesznek a következő pontban. Voltaképpen nem az a legfontosabb, hogy tudjuk: az  $\arctan x$  függvény

deriváltja az  $1/(1+x^2)$ ; hanem megfordítva: az  $1/(1+x^2)$  függvény az  $\arctan x$  függvény deriváltja. Ez persze elvileg ugyanaz a tudás, de “fordított” irányban kell majd elsősorban használni.

$$\begin{aligned}\frac{d}{dx} \arctan x &= \frac{1}{1+x^2}, \quad (x \in \mathbb{R}) \\ \frac{d}{dx} \arcsin x &= \frac{1}{\sqrt{1-x^2}}, \quad (-1 < x < 1)\end{aligned}$$

A közvetett függvény deriválási szabálya és a hatvány-, exponenciális és logaritmusfüggvény deriválásával kaphatjuk a következő formulákat, amelyek igen hasznosak a rutinszerű számolásban.

$$\begin{aligned}\frac{d}{dx} (h(x))^n &= n(h(x))^{n-1} h'(x) \\ \frac{d}{dx} (e^{h(x)}) &= h'(x) e^{h(x)} \\ \frac{d}{dx} \ln(h(x)) &= \frac{h'(x)}{h(x)}, \quad (0 < h(x))\end{aligned}$$

**Példa 6.10** Mutassuk meg, hogy a  $\sin(\alpha x)$  és  $\cos(\alpha x)$  függvények eleget tesznek az alábbi egyenletnek:

$$f''(x) = -\alpha^2 f(x).$$

A  $\sin(\alpha x)$  első deriváltja  $\alpha \cos(\alpha x)$ , és ennek a további deriválásával kapjuk, hogy  $\sin''(\alpha x) = -\alpha^2 \sin(\alpha x)$ . Hasonlóan lehet eljárni a  $\cos$  függvény esetében.  $\square$

### 6.2.8 Függvények diszkussziója

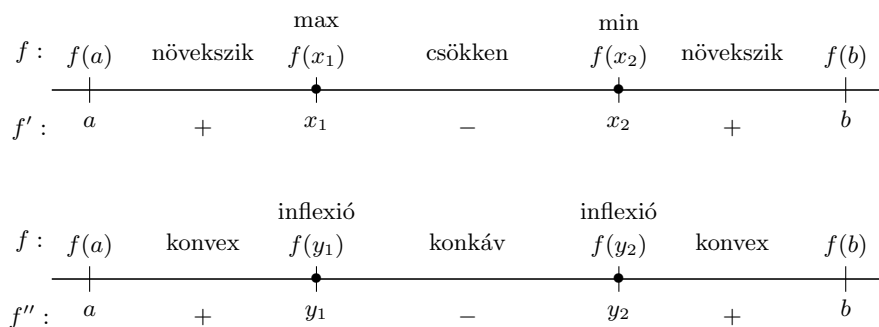
Ennek az alpontnak az a célja, hogy az előző alpontok eredményeit összefoglalva és rendezve egy alkalmas sémát nyújtson a differenciálható leképezések vizsgálatához. Ennek megfelelően voltaképpen semmi új fogalmat vagy állítást nem tartalmaz.

A függvényvizsgálat (diszkusszió) menetének az egyes lépéseit a következő pontokban soroljuk fel. A leírtak *általános tanácsok*, és ezeket mindig céljainknak megfelelően kell alkalmazni.

**A függvénydiszkusszió menetének a vázlata.**

- (I) A függvény általános megtekintése során először az értelmezési tartományt és az értékkészletet nézzük meg. Ha rendesen adjuk meg a függvényt, akkor az értelmezési tartomány adott, de némelykor ennek a meghatározása is feladat. Ezzel kapcsolatban az alábbiakat célszerű elvégezni:
- Az értelmezési tartomány meghatározása. A legtöbb függvény valamilyen formulával van megadva, és ekkor a formulák értelmezhetőségének a körét kell szemügyre venni.
  - Az értelmezési tartomány néhány speciális, hangsúlyozott helyén meghatározzuk a függvény értékeit.
  - Az értelmezési tartományon kívül, de ahhoz szorosan kapcsolódva lehetnek olyan pontok, ahol a függvény nem értelmezett, de a határérték létezik meghatározni. Ilyen pontok gyakorta a  $+\infty$  és  $-\infty$  helyek, és a nevezők zérushelyei.
- (II) Megnézzük, hogy hol folytonos a függvény. Amennyiben deriválható, akkor eleve igenlő a válasz, és áttérhetünk a következő pontra.
- (III) Ha differenciálható a függvény, akkor meghatározzuk az első deriváltját, és meghatározzuk annak a gyökeit és a gyökök közötti előjeleit. Ehhez ésszerű néha szorzattá alakítani a deriváltat.
- Amely intervallumon a derivált nemnegatív (pozitív), ott (szigorúan) monoton növekedő, ahol pedig nempozitív (negatív), ott (szigorúan) csökkenő a függvény.
  - A belső pontokban csak ott lehet szélsőérték, ahol a derivált nulla. Ha növekedőből fogyóba megy át a függvény, akkor maximuma van, ha pedig fogyóból növekedőbe megy át, akkor minimuma. Az extremáltság eldöntéséhez segítséget nyújthat a következőkben kiszámolandó második derivált is.
- (IV) Amennyiben létezik, kiszámoljuk a második deriváltat is.
- Amely intervallumon a második derivált nemnegatív (pozitív), ott a függvény (szigorúan) konvex, ahol pedig nempozitív (negatív), ott (szigorúan) konkáv.
  - Megkeressük az inflexiós pontokat, ahol konvexből konkávba vagy konkávból konvexbe megy át a gráf.
  - A szélsőértéket rendszerint már a függvény növekedése és fogyása segítségével meg tudjuk határozni, de segítségünkre lehet az is, hogy a második derivált milyen értéket vesz fel az első derivált zérushelyeinél. Ha a második derivált negatív, akkor szigorú maximum van, ha pedig pozitív, akkor szigorú minimum.

- (V) Az eddigiekben nyert információk felhasználásával, a vizualitás kedvéért felvázoljuk a gráfot. Olyan rajzra van szükség, amelyik jól tükrözi a kvalitatív megállapításokat. Nem a pontos számszerűség a lényeges.



6.7. ábra: Egy diszkussziós séma.

Egy sémát is ajánlunk az előzőek rögzítésére, amit a 6.7. ábrán rajzoltunk meg. A rajz önmagáért beszél és nem szorul magyarázatra. A példamegoldások során mi is el fogjuk készíteni.

Most pedig kidolgozunk néhány példát.

**Példa 6.11** *Diszkutáljuk a másodfokú polinommal megadott  $p : \mathbb{R} \rightarrow \mathbb{R}$ ,*

$$p(x) = ax^2 + bx + c \quad a \neq 0$$

*függvényt.*

Az  $a \neq 0$  feltételt azért kötöttük ki, hogy valódi másodfokú polinomunk legyen. Az értelmezési tartomány nyilvánvalóan az egész  $\mathbb{R}$ . A  $p$  első deriváltja

$$p'(x) = 2ax + b = 2a(x + b/2a),$$

amiből a derivált nullhelye és előjelei azonnal leolvashatók. Két esetet célszerű megkülönböztetni, aszerint, hogy az  $a$  pozitív vagy negatív. Vegyük mondjuk a pozitív  $a$  esetét, a másik eset teljesen azonosan tárgyalható. A derivált negatív, ha  $x < -\frac{b}{2a}$  és pozitív, ha  $x > -\frac{b}{2a}$ . Ennek megfelelően a függvény szigorúan fogy, ha  $x < -\frac{b}{2a}$  és szigorúan nő, ha  $-\frac{b}{2a} < x$ , és így az  $x = -\frac{b}{2a}$  helyen minimuma van. A második derivált  $p''(x) = 2a$ , és eszerint a függvény konvex, ha az  $a$  pozitív és konkáv, ha az  $a$  negatív. A gráfokat a 6.8. ábrán szemléltettük.

**Példa 6.12** *Diszkutáljuk a*

$$h(x) = ax^3 + bx^2 + cx + d \quad (a \neq 0)$$

*harmadfokú polinom által megadott függvényt.*

6.8. ábra: Az  $x \mapsto ax^2 + bx + c$  függvény gráfja.

Értelmezési tartományként nyilvánvalóan az egész számegyenes szóbajöhet. Az első derivált

$$h'(x) = 3ax^2 + 2bx + c.$$

Mivel ennek a gyökeit kell meghatároznunk, ezért értelemszerűen három esetet különböztetünk meg:

- 1) Nincs valós gyöke a deriválnak. Ekkor nyilvánvalóan állandó előjelű, és ezért a függvény szigorúan növekedő vagy fogyó az egész  $\mathbb{R}$ -en.
- 2) Egy  $\alpha \in \mathbb{R}$  gyöke van a deriválnak. Az  $\alpha$  nyilvánvalóan kétszeres gyök, és a derivált

$$h'(x) = 3a(x - \alpha)^2$$

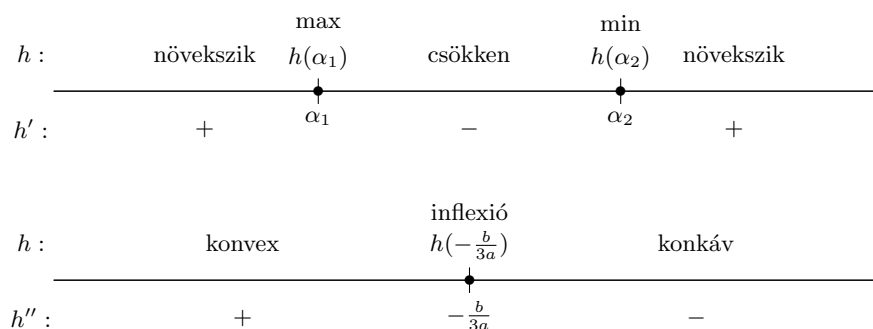
alakba írható, amiből azonnal látható, hogy az  $\alpha$  helytől eltekintve minden helyen pozitív, ha  $a > 0$ , és negatív, ha  $a < 0$ . Ennek megfelelően a  $h$  függvény szigorúan fogy vagy nő az  $a$  előjele szerint. Az  $\alpha$  helyen nulla ugyan a derivált, de itt szélsőérték nincs. Az inflexiós pontban húzott érintő párhuzamos az  $x$  tengellyel.

3) Két gyöke van a deriválnak. Ekkor a deriváltja, mint másodfokú polinom gyöktényezős alakba írható:

$$h'(x) = 3a(x - \alpha_1)(x - \alpha_2),$$

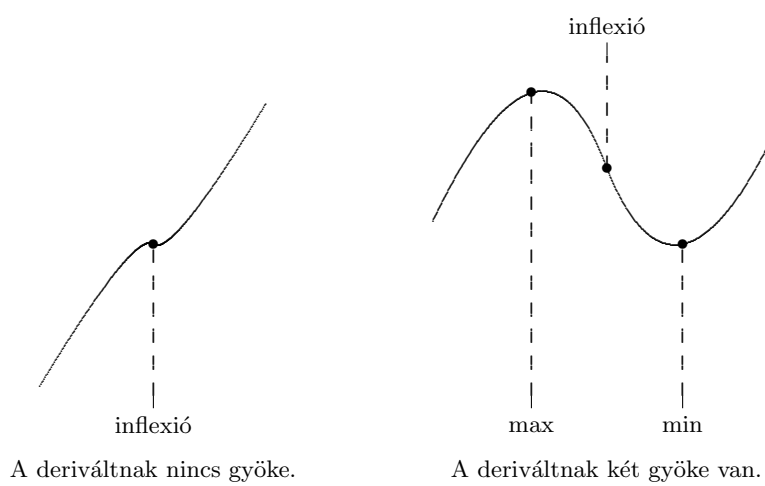
ahol feltehető, hogy  $\alpha_1 \leq \alpha_2$ . Ebből leolvasható, hogy a  $h'$  a két gyöknél vált előjelet. A gyöktényezők előjele a két gyök között ellenkező. A második derivált egy lineáris függvény, aminek egy gyöke van, ahol előjelet is vált.

Ennek az esetnek a diszkussziós sémáját is elkészítettük a 6.9. ábrán, pozitív  $a$  együttható mellett.



6.9. ábra: Harmadfokú polinom diszkussziós sémája.

A 5.2. ábra egy olyan esetet mutat, amikor egy kétszeres gyök van, nevezetesen a nulla. A 6.10. ábrán azokat az eseteket vázoltuk fel, amikor nincs gyök és amikor két gyök van. Az  $a$  együtthatót pozitívnak vettük.



6.10. ábra: Harmadfokú polinomok tipikus gráfjai.

## 7.

# Antiderivált és differenciálegyenletek

Ennek a fejezetnek a célja nagyon egyszerűen megadható: egy  $f$  függvény esetében arra szeretnénk válaszolni, hogy milyen függvény deriváltjaként kapható meg. Eszerint a feladat a differenciálás kalkulusa “megfordításának” a vizsgálata. A tárgyalt kalkulus felhasználására fontos példát adnak a második pontban tárgyalt differenciálegyenletek.

## 7.1 Az antiderivált, határozatlan integrál

Az első alpontban definiáljuk az antiderivált fogalmát, a következő három alpontban pedig módszereket ismerünk meg az antiderivált kiszámítására.

### 7.1.1 Definíciók, elemi tulajdonságok

A deriválás eljárásának a megfordítása nem egyértelmű, hiszen például az  $x^2$  deriváltja az

$$x \mapsto \frac{x^3}{3} \quad \text{és} \quad x \mapsto \frac{x^3}{3} + 5$$

függvényeknek. A 236. tétel (4) állítása szerint ha az  $F$  és  $G$  (*intervallumon értelmezett*), függvények deriváltja azonos, akkor az  $F$  és  $G$  függvények különbsége egy konstans függvény. Ez lehetővé teszi, hogy azt mondjuk: a deriválás megfordítása egy állandó (függvény) erejéig lehetséges, vagy azt is mondhatnánk: a deriválás megfordítása egy olyan függvényhalmazt eredményez, amelyben bármely két függvény különbsége egy konstans függvény. Ezt a gondolatot pontosítjuk a következőkben.

**Definíció 256** Ha az  $f : (a, b) \rightarrow \mathbb{R}$  függvény az  $F : (a, b) \rightarrow \mathbb{R}$  függvény deriváltja (derivált függvénye), akkor azt fogjuk mondani, hogy az  $F$  függvény egy antideriváltja vagy primitív függvénye az  $f$  függvénynek.

Az antiderivált függvények összességének a jelölésére az

$$\int f(x) dx, \quad \int dx f(x), \quad \int f$$

formák valamelyike a szokásos. Az antideriváltak  $\int f(x) dx$  összességét az  $f$  függvény határozatlan integráljának mondjuk, amelyet egy  $F$  antiderivált birtokában

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

módon adhatunk meg, ahol a  $c$  tetszőleges állandó.

A definíció utóbbi megadási módja az 236.(4) állítás szerint jogos. Hangsúlyozzuk, hogy a határozatlan integrál egy függvényhalmaz, az antiderivált pedig egy az elemei közül.

Az antiderivált elnevezés nem szorul indoklásra, hiszen a bevezetett művelet a deriválás megfordítása. A határozatlan integrál elnevezés a következő fejezetben kapja meg az indoklását, ahol majd látni fogjuk, hogy ez a művelet teszi lehetővé az u.n. (határozott) integrálok kiszámítását.

A jelölésnek tradicionális oka van, az “ $\int$ ” jel az “ $\mathcal{S}$ ” betűnek a stilizált alakja arra utal, hogy — amint majd látni fogjuk — a határozott integrál, bizonyos értelemben a közönséges összegezésnek (szumma) az általánosítása.

A jelölések közül az elsőt használjuk leginkább, mert a benne szereplő “ $dx$ ” szimbólum világosan megmondja: egy  $x$  változótól függő leképezés antideriváltját kell venni. A középső jelölés a leglogikusabb (de a legritkábban használt), mivel az antideriválendő függvény elé teszi, a deriválásnál megszokott módon, az utasítást adó szimbólumot.

A derivált függvényt tetszőleges nyílt halmazon értelmezett függvényre definiáltuk, ezért jogosan felvethető, hogy miért nem tettünk így az antiderivált esetében is. Ennek az oka: Nem igaz az, hogy ha két függvény deriváltja egy nyílt halmazon megegyezik, akkor a különbségük egy konstans függvény. Erre mutatunk most példát:

**Példa 7.1** Mutassuk meg, hogy az

$$f(x) = \frac{1}{x}, \quad \text{és} \quad g(x) = f(x) + \operatorname{sgn}(x), \quad x \in \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$$

függvények derivált függvénye azonos de a különbségük nem állandó.

Az  $f$  és  $g$  derivált függvénye:  $x \rightarrow -\frac{1}{x^2}$ , a különbségük azonban az előjel függvény, ami nem állandó.  $\square$

A részletes bevezetés után nézzük most az antiderivált-kalkulus alapjait. Az eddig is követett menet szerint két típusú műveleti szabály rendszert tárgyalunk:

- 1) A formális szabályokat, amelyek a függvények közötti műveleteknek az antideriválás műveletével szembeni viselkedését adják meg.
- 2) A legfontosabb speciális függvények antideriváltjait.

**Állítás 257 (Antideriválás formális szabályai)**

- (1) Ha az  $\alpha \neq 0$  tetszőleges szám, és az  $f : (a, b) \rightarrow \mathbb{R}$  függvénynek van antideriváltja, akkor az  $\alpha f$  függvénynek is van, és

$$\int \alpha f(x) dx = \alpha \int f(x) dx.$$

- (2) Ha az  $f : (a, b) \rightarrow \mathbb{R}$  és  $g : (a, b) \rightarrow \mathbb{R}$  függvényeknek van antideriváltja, akkor az összegüknek is van, és

$$\int (f(x) + g(x)) dx = \int f(x) dx + \int g(x) dx.$$

- (3) Ha az  $f$  és  $g, (a, b) \rightarrow \mathbb{R}$ , deriválható függvények, és az  $f'g$  függvénynek van antideriváltja, akkor az  $f'g'$  függvénynek is van, és

$$\int f'(x)g(x) dx = f(x)g(x) - \int f(x)g'(x) dx.$$

- (4) Ha a  $g : (a, b) \rightarrow \mathbb{R}$  függvény deriválható, az  $f$  pedig a  $g$  függvény  $g((a, b))$  (intervallum) értékkészletén deriválható, akkor

$$\int f'(g(x))g'(x) dx = f(g(x)).$$

- (5) Ha a  $g : (a, b) \rightarrow \mathbb{R}$  függvény deriválható, az  $f$  függvény pedig a  $g$  függvény  $g((a, b))$  értékkészletén definiált, és van antideriváltja, akkor  $f(g(x))g'(x)$  függvénynek is van antideriváltja, és

$$\int f(g(x))g'(x) dx = \int f(y) dy \Big|_{y=g(x)}.$$

- (6) Ha az előző állítás feltételei mellett a  $g$  függvénynek van inverze, akkor

$$\int f(y) dy = \int f(g(x))g'(x) dx \Big|_{x=g^{-1}(y)}.$$

A leírt formális szabályok rendre a számmal való szorzás, az összeadás, a szorzás és a közvetett függvény deriválási szabályainak felelnek meg. A közvetett függvény deriválási szabályából három antideriválási szabályt is eredeztettünk ((4)–(6)).

Meg kell vallani, hogy a jelölések zavart okozhatnak az állítás egyenlőségeinek a megértésében, ha nem látjuk pontosan, hogy miről is van szó. Például az

$$\int \alpha f(x) dx = \alpha \int f(x) dx, \quad \alpha \neq 0$$

formula értelmezése: Az egyenlőség mindkét oldalán függvény halmazok állnak és úgy kell érteni, hogy a második függvény halmaz úgy adódik az elsőből, hogy az elemeit rendre szorozzuk az  $\alpha$  számmal. Hasonló értelmezést adunk a többi egyenlőségnek is.

A szorzat deriválási szabályából származó (3) szabályt *parciális integrálásnak* szokás nevezni. Az (5) és (6) szabályt pedig *helyettesítéssel való integrálásnak*. Ezek az elnevezések a “határozatlan integrál” elnevezéshez kapcsolódnak.

**Bizonyítás.** Az állítások mindegyike egy egyenlőség, amelyeknek a bizonyítása — mivel antideriváltak egyenlőségéről van szó — úgy megy, hogy deriváljuk mindkét oldalt, és azt találjuk, hogy mindkét oldal deriváltja azonos.

(1): A határozatlan integrál definíciója szerint a baloldal

$$\frac{d}{dx} \int \alpha f(x) dx = \alpha f(x),$$

a jobboldal pedig a szorzat deriválási szabálya és a definíció alapján

$$\frac{d}{dx} \left( \alpha \int f(x) dx \right) = \alpha \frac{d}{dx} \int f(x) dx = \alpha f(x).$$

(2): A definíció szerint a baloldal deriváltja  $f(x) + g(x)$ , a jobboldal deriváltja pedig az összeg deriválásának a formális szabályával:

$$\left( \int f(x) dx + \int g(x) dx \right)' = \left( \int f(x) dx \right)' + \left( \int g(x) dx \right)' = f(x) + g(x).$$

(3): A baloldal deriváltja definíció szerint  $f'(x)g(x)$ , a jobboldal pedig az összeg és szorzat deriválási szabálya szerint:

$$\left( f(x)g(x) - \int f(x)g'(x) dx \right)' = f'(x)g(x) + f(x)g'(x) - f(x)g'(x) = f'(x)g(x).$$

(4): A közvetett függvény differenciálási szabályának az evidens következménye.

(5): A jobboldal deriváltja, a közvetett függvény deriválási szabálya alapján:

$$\frac{d}{dx} \int f(y) dy \Big|_{y=g(x)} = \frac{d}{dy} \int f(y) dy \cdot \frac{d}{dx} g(x) = f(y)g'(x),$$

ami megegyezik a baloldal  $x$  szerinti deriváltjával, az  $f(g(x))g'(x)$ -vel, ha az  $y$  helyére előírás szerint beírjuk a  $g(x)$ -et.

(6): Azonnal adódik az előző állítás egyenlőségéből, ha az  $x$  helyett a  $g^{-1}(y)$  értéket helyettesítjük.  $\square$

A következő tétel speciális függvények antideriváltjait adja meg. Igazolásra nincs is szükség, hiszen csak visszafelé, jobbról balra, kell olvasni a deriválási szabályokat.

### Állítás 258 (Speciális függvények antideriváltjai)

$$(1) \text{ Ha } 0 \leq n \text{ egész, akkor } \int x^n dx = \frac{x^{n+1}}{n+1} + c, \quad x \in \mathbb{R}.$$

$$(2) \text{ Ha az } n < -1 \text{ egész, akkor } \int x^n dx = \begin{cases} \frac{x^{n+1}}{n+1} + c_1, & \text{ha } x > 0, \\ \frac{x^{n+1}}{n+1} + c_2, & \text{ha } x < 0. \end{cases}$$

$$(3) \text{ Ha } 0 < x \text{ és } \alpha \neq -1, \text{ akkor } \int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + c.$$

$$(4) \int e^x dx = e^x + c.$$

$$(5) \text{ Ha az } x \text{ pozitív, akkor } \int \frac{1}{x} dx = \ln x + c.$$

$$(6) \int \sin x dx = -\cos x + c \quad \text{és} \quad \int \cos x dx = \sin x + c.$$

$$(7) \int \frac{1}{1+x^2} dx = \arctan x + c.$$

$$(8) \text{ Ha } x \in (-1, 1), \text{ akkor } \int \frac{1}{\sqrt{1-x^2}} dx = \arcsin x + c.$$

Végezetül, mielőtt példákat dolgoznánk ki, a számolási szabályokat tömören, táblázatokban is összefoglaljuk a következő oldalon.

**ANTIDERIVÁLÁS FORMÁLIS SZABÁLYAI**

$$\begin{aligned}
\int \alpha f(x) dx &= \alpha \cdot \int f(x) dx, (\alpha \neq 0) \\
\int (f + g)(x) dx &= \int f(x) dx + \int g(x) dx \\
\int f(x) g'(x) dx &= f(x) g(x) - \int f'(x) g(x) dx \\
\int f'(g(x)) g'(x) dx &= f(g(x)) + c \\
\int f(g(x)) g'(x) dx &= \int f(y) dy \Big|_{y=g(x)} \\
\int f(y) dy &= \int f(g(x)) g'(x) dx \Big|_{x=g^{-1}(y)}
\end{aligned}$$

**SPECIÁLIS FÜGGVÉNYEK ANTIDERIVÁLTJAI**

$$\begin{aligned}
\int 1 dx &= x + c \\
\int x^n dx &= \frac{x^{n+1}}{n+1} + c \\
\int x^\alpha dx &= \frac{x^{\alpha+1}}{\alpha+1} + c \\
\int e^x dx &= e^x + c \\
\int \frac{1}{x} dx &= \ln x + c, \quad (0 < x) \\
\int \sin x dx &= -\cos x + c \\
\int \cos x dx &= \sin x + c \\
\int \frac{1}{1+x^2} dx &= \arctan x + c \\
\int \frac{1}{\sqrt{1-x^2}} dx &= \arcsin x + c
\end{aligned}$$

A táblázatok emlékeztetők, és a pontos feltételeket a megfelelő állítások tartalmazzák.

A következő táblázatban néhány olyan formulát írunk le, amelyek a formális szabályok és a speciális függvények antideriváltjainak az ismeretében könnyen felírhatóak, és hasznosak lehetnek gyorsabb számolásnál.

$$\begin{array}{rcl} \int h'(x)h^n(x) dx & = & \frac{h^{n+1}(x)}{n+1} + c, \quad (n \neq -1) \\ \int h'(x)e^{h(x)} dx & = & e^{h(x)} + c \\ \int \frac{h'(x)}{h(x)} dx & = & \ln h(x) + c, \quad (0 < h(x)) \end{array}$$

A megismert szabályok közül a parciális és helyettesítéssel való integrálás formális szabályainak a használata, megfelelő gyakorlatot igényel. Emiatt ezek gyakorlásának két külön pontot szentelünk. Itt csak a közvetlenül megoldható feladatokkal foglalkozunk.

**Példa 7.2** *Keressük meg a tangens függvény egy antideriváltját.*

A tangens függvényre  $\tan x = \frac{\sin x}{\cos x}$ , ami továbbalakítva a

$$\tan x = -\frac{-\sin x}{\cos x}, \quad \text{ha} \quad \cos x > 0$$

és

$$\tan x = -\frac{\sin x}{-\cos x}, \quad \text{ha} \quad \cos x < 0$$

formákba írható, ahol a tört számlálója éppen a nevező deriváltja, ezért egy antiderivált függvény:

$$x \mapsto -\ln \cos x, \quad \text{ha} \quad \cos x > 0$$

és

$$x \mapsto -\ln(-\cos x), \quad \text{ha} \quad \cos x < 0.$$

□

**Példa 7.3** *Mi a határozatlan integrálja az  $x + 2 \sin x \cos^3 x$  függvénynek?*

Az integrálandó függvény a következő formába alakítható

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

ami alapján a második tagnál a harmadik táblázat első sorát használva az adódik, hogy

$$\int (x + 2 \sin x \cos^3 x) dx = \frac{x^2}{2} - \frac{1}{2} \cos^4 x + c.$$

□

**Példa 7.4**

$$\int \sin^3 x dx = ?$$

A

$$\sin^3 x = \sin x(1 - \cos^2 x) = \sin x + \frac{1}{3}(-\sin x)3\cos^2 x$$

azonosság alapján

$$\int \sin^3 x dx = \int \sin x dx + \frac{1}{3} \int (-\sin x)3\cos^2 x dx = -\cos x + \frac{1}{3} \cos^3 x.$$

**Példa 7.5** *Mi a határozatlan integrálja az*

$$x \mapsto \frac{3x^2 + x}{x^2 + 1}$$

*függvénynek?*

Az előző példák megoldásában láttuk, hogy a megoldás kulcsa az, hogy ügyes átalakítással olyan alakot hozzunk be, amelyekre a harmadik táblázat valamelyik formulája ráillik. Ez a kulcsa az összes ebben a pontban megoldott és kitűzött feladat megoldásának.

A jelenlegi esetben — ami egyes racionális törtfüggvényeknél tipikusnak tekinthető — a következő az átalakítás:

$$\frac{3x^2 + x}{x^2 + 1} = \frac{3(x^2 + 1) - 3 + x}{x^2 + 1} = 3 - 3\frac{1}{1 + x^2} + \frac{1}{2}\frac{2x}{x^2 + 1}.$$

A vizsgált törtet három tagú összegre bontottuk. Az első tag egy antideriváltja a  $3x$ , a második tagé  $-3 \arctan x$ . A harmadik tagban a számláló a nevező deriváltja, ezért egy antiderivált:  $\frac{1}{2} \ln(1 + x^2)$ . Összefoglalva:

$$\int \frac{3x^2 + x}{x^2 + 1} dx = 3x - 3 \arctan x + \frac{1}{2} \ln(1 + x^2) + c.$$

□

### 7.1.2 Parciális integrálás

Bevezetésként emlékeztetünk a parciális integrálás szabályára: Ha az  $f$  és  $g$  függvények deriválhatóak valamilyen intervallumon, és az  $f(x)g'(x)$  függvénynek van antideriváltja, akkor az  $f'(x)g(x)$  függvénynek is van, és

$$\int f(x)g'(x) dx = f(x)g(x) - \int f'(x)g(x) dx, \quad (7.1)$$

ahogyan azt a 257. tételben láttuk.

Gondoljuk meg alaposan, hogy mikor előnyös ennek a formális szabálynak a használata. A 7.1. egyenlőség akkor könnyíti meg a dolgunkat, ha egyrészt meg tudjuk mondani a  $g'$  függvénynek egy antideriváltját, a  $g$  függvényt; másrészt ha az  $f'g$  függvény “egyszerűbb” az  $fg'$  függvénynél. Lássunk előzetesként néhány tipikus példát az alapgondolat megértésére.

- (1) A “polinom  $\cdot e^x$ ” alakú függvények esetében kedvező a helyzet. Vegyük például az

$$\underbrace{(x^2 + 7)}_f \cdot \underbrace{e^x}_{g'},$$

függvényt, amiből az  $f'g$ -re a az

$$f'(x)g(x) = 2xe^x,$$

alak adódik, ami abban az értelemben egyszerűbb az eredeti formánál, hogy a polinom fokszáma eggyel csökkent. Ezt a fogást tovább lehet folytatni, és ismételt lépések után a polinom nulla fokszámú, azaz konstans lesz.

- (2) Teljesen azonos a helyzet a “polinom  $\cdot \sin x$ ” alakú függvényeknél, ahol a szinusz függvény helyett a koszinusz is szerepelhet.
- (3) Bizonyos értelemben eltérő szempont miatt vezet kedvező helyzetre a parciális integrálás a “polinom  $\cdot \ln x$ ” forma esetében. Ebben a

$$\underbrace{\text{polinom}}_{g'} \cdot \underbrace{\ln x}_f$$

szereposztásban, mivel polinom antideriváltja is polinom, az  $f'g$  alakja

$$\text{polinom} \cdot \frac{1}{x}$$

lesz, ami egyszerű racionális törtfüggvény, és könnyű mondani antideriváltját.

A pont további részében példákat oldunk meg a parciális integrálás módszerének a bemutatására.

**Példa 7.6** *Határozzuk meg az*

$$x \mapsto (x^2 + 7)e^{-x}$$

*függvény határozott integrálját.*

A szabály alkalmazását — az  $f$  és  $g$  leképezésekre vonatkozó “szereposztást” (eleinte) — a függvények alá való írásával fejezzük ki.

Felhasználjuk azt, hogy az  $e^{-x}$  egy antideriváltja a  $(-e^{-x})$  függvény.

$$\begin{aligned} \int \underbrace{(x^2 + 7)}_f \cdot \underbrace{e^{-x}}_{g'} dx &= \underbrace{(x^2 + 7)}_f \cdot \underbrace{(-e^{-x})}_g - \int \underbrace{2x}_{f'} \cdot \underbrace{(-e^{-x})}_g dx = \\ &= (x^2 + 7)(-e^{-x}) + \int 2x(-e^{-x}) dx. \end{aligned}$$

A jobboldali integrálra ugyanezt az eljárást folytatjuk tovább:

$$\begin{aligned} \int \underbrace{2x}_f \cdot \underbrace{e^{-x}}_{g'} dx &= \underbrace{2x}_f \cdot \underbrace{(-e^{-x})}_g - \int \underbrace{2}_{f'} \cdot \underbrace{(-e^{-x})}_g dx = \\ &= 2x(-e^{-x}) + 2 \int e^{-x} dx = -2xe^{-x} - 2e^{-x} + c. \end{aligned}$$

Összefoglalva az eddigieket:

$$\int (x^2 + 7)e^{-x} dx = -(x^2 + 7)e^{-x} - 2xe^{-x} - 2e^{-x} + c = -(x^2 + 2x + 9)e^{-x} + c.$$

□

**Példa 7.7**

$$\int (x^7 + 2x^2 + 1) \ln x \, dx = ?$$

Alkalmas szereposztással számolva egyszerűen nyerhető a válasz:

$$\begin{aligned} &\int \underbrace{(x^7 + 2x^2 + 1)}_{g'} \cdot \underbrace{\ln x}_f dx = \\ &= \underbrace{\left(\frac{x^8}{8} + \frac{2x^3}{3} + x\right)}_g \cdot \underbrace{\ln x}_f - \int \underbrace{\left(\frac{x^8}{8} + \frac{2x^3}{3} + x\right)}_g \cdot \underbrace{\frac{1}{x}}_{f'} dx. \end{aligned}$$

A jobboldali integrál kiszámolása már nagyon egyszerű:

$$\int \left( \frac{x^7}{8} + \frac{2x^2}{3} + 1 \right) dx = \frac{x^8}{64} + \frac{2x^3}{9} + x + c,$$

amiből végülis azt kapjuk, hogy

$$\int (x^7 + 2x^2 + 1) \ln x \, dx = \left( \frac{x^8}{8} + \frac{2x^3}{3} + x \right) \ln x - \frac{x^8}{64} - \frac{2x^3}{9} - x + c.$$

□

**Példa 7.8** Számítsuk ki az  $x \mapsto \sin^2 x$  függvény határozatlan integrálját.

Első pillantásra úgy tűnhetne, hogy ez a feladat nem kezelhető a parciális integrálás módszerével, hiszen nem is szorzat. Ennek ellenére — azzal az észrevétellel, hogy a négyzet is szorzat — megoldhatóvá válik a feladat ezen a módon, az alábbiak szerint.

$$\begin{aligned} \int \underbrace{\sin x}_f \cdot \underbrace{\sin x}_{g'} dx &= \underbrace{\sin x}_f \cdot \underbrace{(-\cos x)}_g - \int \underbrace{\cos x}_{f'} \cdot \underbrace{(-\cos x)}_g dx = \\ &= -\sin x \cos x + \int \cos^2 x \, dx = -\sin x \cos x + \int 1 - \sin^2 x \, dx \\ &= -\sin x \cos x + x - \int \sin^2 x \, dx, \end{aligned}$$

amiből, átrendezéssel az adódik, hogy  $2 \int \sin^2 x \, dx = -\sin x \cos x + x + c$ , tehát  $\int \sin^2 x \, dx = -\frac{1}{2} \sin x \cos x + \frac{x}{2} + c$ . □

A pontosan gondolkodók észrevehették, hogy egy furcsa dolgot tettünk akkor, amikor az  $\int \sin^2 x \, dx$  határozatlan integrált az egyenlőség egyik oldaláról átvittük a másik oldalára, mivel az nem egyetlen objektum, nem egy függvény, hanem egymástól egy konstansban különböző függvények halmaza. Ha alaposabban utána gondolunk, akkor beláthatjuk, hogy ez mégis megengedhető.

A határozatlan integrállal kapcsolatos számolásainknak egyébként van egy kitűnő ellenőrzése: deriválni kell az eredményt, és meg kell nézni, hogy az integrálandó függvényt kapjuk-e meg. Az előző példában hajtsuk végre ezt az ellenőrzést:

$$\begin{aligned} \frac{d}{dx} \left( -\frac{1}{2} \sin x \cos x + \frac{x}{2} \right) &= 1/2(-\cos x \cos x - \sin x(-\sin x) - 1) = \\ &= 1/2(-\cos^2 x + \sin^2 x + 1) = \sin^2 x. \end{aligned}$$

**Példa 7.9** Adjuk meg az  $x \mapsto x^3 \exp(-x^2)$  függvény egy antideriváltját.

Ha ebben az esetben abból indulnánk ki, hogy csak az  $x^3$  függvényt választhatjuk deriválnak, mert ennek tudjuk csak megmondani egy antideriváltját, az  $x^4/4$  függvényt, akkor csak megnehezítenénk a dolgunkat. A másik  $\exp(-x^2)$  függvény deriváltja ugyanis  $-2x \exp(-x^2)$ , és így az  $x^5 \exp(-x^2)$  függvényt kellene integrálnunk, és ez nem könnyebb, mint a kiinduló feladat. Egy “fantáziadúsabb” látásmódra van szükség, az alábbiak szerint.

$$\begin{aligned} \int x^3 \exp(-x^2) dx &= -\frac{1}{2} \int \underbrace{x^2}_f \cdot \underbrace{(-2x e^{-x^2})}_{g'} dx = \\ &= -\frac{1}{2} \underbrace{x^2}_f \cdot \underbrace{e^{-x^2}}_g + \frac{1}{2} \int \underbrace{2x}_{f'} \cdot \underbrace{e^{-x^2}}_g dx = -\frac{1}{2} e^{-x^2} (x^2 + 1) + c, \end{aligned}$$

tehát egy antiderivált az  $x \mapsto -(1/2)e^{-x^2}(x^2 + 1)$  függvény.  $\square$

**Példa 7.10** Határozzuk meg az

$$\frac{x^2}{(1+x)^3}$$

függvénynek egy antideriváltját.

A megoldás menete az eddigi tapasztalatokat követve:

$$\begin{aligned} \int \frac{x^2}{(1+x)^3} dx &= \int \underbrace{x^2}_f \cdot \underbrace{(1+x)^{-3}}_{g'} dx = \\ &= \underbrace{x^2}_f \cdot \underbrace{\frac{(1+x)^{-2}}{-2}}_g - \int \underbrace{2x}_{f'} \cdot \underbrace{\frac{(1+x)^{-2}}{-2}}_g dx = \\ &= -\frac{x^2}{2(1+x)^2} + \int \underbrace{x}_f \cdot \underbrace{(1+x)^{-2}}_{g'} dx = \\ &= -\frac{x^2}{2(1+x)^2} + \underbrace{x}_f \cdot \underbrace{\frac{(1+x)^{-1}}{-1}}_g - \int \underbrace{1}_{f'} \cdot \underbrace{\frac{(1+x)^{-1}}{-1}}_g dx = \\ &= -\frac{x^2}{2(1+x)^2} - \frac{x}{1+x} + \int (1+x)^{-1} dx = \\ &= -\frac{x^2}{2(1+x)^2} - \frac{x}{1+x} + \ln(1+x) + c. \end{aligned}$$

$\square$

A parciális integrálás formális szabálya a szorzat deriválási szabályának a megfordítása, de észrevehettük, hogy az alkalmazása lényegesen nehezebb. A lényeges lépés a fentiekben “szereposztásnak” nevezett probléma. Ha ez utóbbit helyesen oldjuk meg, akkor a továbbiak gyakran már maguktól mennek.

### 7.1.3 Helyettesítéssel való integrálás

A közvetett függvény deriválási szabályából három szabályt is levezettünk a határozatlan integrál meghatározására (257. (4)–(6) állítás). A legegyszerűbb az alábbi

$$\int f'(g(x))g'(x) dx = f(g(x)) + c. \quad (7.2)$$

Ezt a formulát már eddig is sokszor használtuk, amikor az integrálandó függvényben ezt az alakot igyekeztünk felfedezni, vagy átalakítással létrehozni. “Helyettesítéssel való integrálásnak” főként az idézett tétel (6) formulája alkalmazását szokás nevezni. Eszerint

$$\int f(y) dy = \int f(g(x))g'(x) dx \Big|_{x=g^{-1}(y)}. \quad (7.3)$$

Ne feledjük el, hogy itt az alkalmazhatóság feltétele az is, hogy a  $g$  függvénynek legyen inverze.

Oldjunk meg most egy feladatot, kétféle módon, nagy részletességgel, abból a célból, hogy annak az alapján a (7.3) formula alkalmazásának, nem kevés ötletességet kívánó, menetét lerögzíthessük.

**Példa 7.11** *Keressük meg az*

$$y \mapsto \frac{y^5}{\sqrt{y^3+1}}$$

*függvény határozatlan integrálját.*

Első megoldás: Az feltűnően szembevető, hogy az integrálandó kifejezés sokkal egyszerűbb lenne, ha a négyzetgyök helyett egyetlen  $x$  változó állna. Ekkor

$$x = \sqrt{y^3+1} \quad \text{és} \quad y = \sqrt[3]{x^2-1}$$

lenne. Az alkalmazandó formulában szereplő  $g$  függvény ezek szerint:  $y = g(x) = \sqrt[3]{x^2-1}$ . Figyeljünk fel azonban arra, hogy a  $g$  függvény  $x = g^{-1}(y) = \sqrt{y^3+1}$  inverz függvényéből indultunk ki. Azt választottuk ki először, és a  $g$  függvényt csak ezután számoltuk ki, kifejezve az  $y$  változót.

A  $g$  és  $g^{-1}$  függvények megválasztása után az látszik természetesnek, hogy kiszámoljuk az  $f(g(x))$  függvényt és a  $g'(x)$  deriváltat, hogy ezek alapján a jobboldali integrandust felírhassuk. Ebben a megoldásban valóban ezt az utat követjük, de találni fogunk majd más eljárást is a második megoldásban. A mondotaknak megfelelően:

$$\begin{aligned} f(g(x)) &= \left( \sqrt[3]{x^2-1} \right)^5 \frac{1}{x}. \\ g'(x) = \frac{d}{dx} \sqrt[3]{x^2-1} &= \frac{1}{3}(x^2-1)^{\frac{1}{3}-1} \cdot (2x) = \frac{2}{3}x \left( \sqrt[3]{x^2-1} \right)^{-2}. \end{aligned}$$

Így a jobboldali integrandus:

$$f(g(x))g'(x) = \left(\sqrt[3]{x^2-1}\right)^5 \frac{1}{x} \cdot \left(\frac{2}{3}x \left(\sqrt[3]{x^2-1}\right)^{-2}\right) = \frac{2}{3}(x^2-1).$$

Eszerint a helyettesítés (7.3) szabálya szerint

$$\begin{aligned} \int \frac{y^5}{\sqrt{y^3+1}} dy &= \frac{2}{3} \int (x^2-1) dx \Big|_{x=\sqrt{y^3+1}} = \\ &= \frac{2}{3} \left( \frac{x^3}{3} - x + c \right) \Big|_{x=\sqrt{y^3+1}} = \frac{2}{9}(y^3+1)\sqrt{y^3+1} - \frac{2}{3}\sqrt{y^3+1} + c. \end{aligned}$$

Második megoldás: Ezzel a megoldással két szempontra szeretnénk felhívni a figyelmet. Egyrészt arra, hogy többféle helyettesítés is célhoz vezethet; másrészt arra, hogy az  $f(g(x))g'(x)$  integrandus (integrálandó függvény) kiszámítására más út is van: Nem hajtjuk végre teljes mértékben az  $f(g(x))$  helyettesítést mindjárt az elején, és nem közvetlenül a  $g'(x)$  deriváltat számítjuk ki, hanem a  $g^{-1}(y)$  deriváltat, aminek — az inverz függvény deriválási szabálya szerint — a reciproka a  $g'$  derivált.

Helyettesítsük most az integrálandó kifejezésben a gyök alatti mennyiséget  $x$ -szel, azaz  $x = y^3 + 1$ . Eszerint  $y = g(x) = \sqrt[3]{x-1}$ . Ha az előző megoldásban követett úton haladnánk, akkor a  $g$  függvényt kellene deriválni, ami — gyökös kifejezés lévén — viszonylag bonyolult. Annál egyszerűbb viszont a  $g^{-1}$  függvény deriválása, hiszen  $g^{-1}(y) = y^3 + 1$ . Lássunk hozzá ennek megfelelően a baloldali integrandus kiszámolásához. Először a  $x = g^{-1}(y)$  deriválása:

$$\frac{dx}{dy} = \frac{d}{dy}(y^3 + 1) = 3y^2.$$

Ezek szerint a baloldali integrandus:

$$\frac{y^5}{\sqrt{x}} g'(x) = \frac{y^5}{\sqrt{x}} \frac{1}{3y^2} = \frac{1}{3} \frac{y^3}{\sqrt{x}},$$

és most már érdemes az  $y$  helyére is betenni az  $y = \sqrt[3]{x-1}$  kifejezést, amivel folytatva az előzőt az következik, hogy

$$= \frac{1}{3} \frac{x-1}{\sqrt{x}} = \frac{1}{3} (x^{1/2} - x^{-1/2}).$$

A helyettesítés formulája szerint tehát

$$\int \frac{y^5}{\sqrt{y^3+1}} dy = \frac{1}{3} \int (x^{1/2} - x^{-1/2}) dx \Big|_{x=y^3+1} =$$

$$= \frac{1}{3} \left( \frac{x^{3/2}}{3/2} - \frac{x^{1/2}}{1/2} + c \right) \Big|_{x=y^3+1} = \frac{2}{9}(y^3+1)\sqrt{y^3+1} - \frac{2}{3}\sqrt{y^3+1} + c.$$

Mielőtt összefoglalnánk az előzőeket ejtsünk pár szót egy lehetséges jelölésről. Mivel  $y = g(x)$  és  $g'(x) = \frac{dg(x)}{dx} = \frac{dy}{dx}$ , ezért a 4.2.4. pontban látottak szerint azt írhatjuk, hogy  $dg(x) = dy = g'(x)dx$ . Megjegyzés-technikai szempontokat figyelembe véve a (7.3) helyettesítési szabályban a  $g'(x)dx$  formát a következőképpen értelmezhetjük: A „ $dy$ ” helyére a  $\frac{dy}{dx} = g'(x)$  egyenlőségből formálisan kifejezett „ $g'(x)dx$ ” írandó, azaz a helyettesítés

$$dy = g'(x)dx.$$

Ha a  $g$  helyett a  $g^{-1}$  deriváltját számoljuk ki, akkor pedig nyilvánvalóan a

$$dy = \frac{1}{g^{-1}(y)} dx$$

a helyettesítés.

Részletezzük most, hogy mi az integrálás menete a (7.3) szabály alkalmazásánál:

- (1) Az  $f$  függvény alakja alapján eldöntjük azt, hogy milyen  $g(x)$  függvényt helyettesítünk az  $y$  helyébe. Erre nincs általános recept, csak az a szempont, hogy a helyettesítés utáni  $f(g(x))$  függvény minél egyszerűbb legyen. Mivel az  $f(y)$  függvény alakja alapján fantáziálunk, ezért nem meglepő, hogy a  $g$  függvény helyett rendszerint annak az inverzét találjuk meg.
- (2) Az  $f(g(x))g'(x)$  szorzat kiszámításánál kétféleképpen is eljárhatunk:
  - (a) Először elvégezzük az  $f(g(x))$  behelyettesítést, majd kiszámoljuk a helyettesítéshez szükséges  $g'(x)$  derivált függvényt, és a  $dy$  helyére a  $g'(x)dx$  formát tesszük.
  - (b) A  $g$  deriváltja helyett a  $g^{-1}(y)$  függvény deriváltját számoljuk ki.
 Az (a) módszer mindig követhető, a (b) esetben esetleg egyszerűbb a számolás.
- (3) Meghatározzuk az  $\int f(g(x))g'(x) dx$  határozatlan integrált.
- (4) Az előző lépésben kapott kifejezésben az  $x$  helyére a  $g^{-1}(y)$  formát helyettesítjük vissza.

#### Példa 7.12

$$\int \frac{e^{2t}}{e^{4t} + 1} dt = ?$$

Próbálkozzunk az alábbi helyettesítéssel:

$$u = e^{2t}, \quad \text{amiből} \quad t = (\ln u)/2.$$

A helyettesítéshez szükséges deriválás:

$$\frac{du}{dt} = 2e^{2t}, \quad \text{azaz} \quad dt = \frac{du}{2e^{2t}}.$$

Ebből a helyettesített integrandus:

$$\frac{e^{2t}}{e^{4t} + 1} \frac{1}{2e^{2t}} = \frac{1}{2} \frac{1}{e^{4t} + 1} = \frac{1}{2} \frac{1}{u^2 + 1}.$$

Ennek az  $u$ -szerinti határozatlan integrálja:

$$\int \frac{1}{2} \frac{1}{u^2 + 1} du = \frac{1}{2} \arctan u + c.$$

Az  $u = e^{2t}$  helyettesítéssel kapjuk is a keresett eredményt:

$$\int \frac{e^{2t}}{e^{4t} + 1} dt = \frac{1}{2} \arctan(e^{2t}) + c.$$

□

**Példa 7.13** Számoljuk ki az  $x \mapsto \sqrt{r^2 - x^2}$  ( $0 \leq x \leq r$ ) függvény határozatlan integrálját.

Nyilvánvalóan felvetődik bennünk, hogy jó lenne egy olyan helyettesítés, ami mellett a gyökjel eltűnne. Az előzőek alapján gondolhatnánk azt, hogy az  $y = \sqrt{r^2 - x^2}$  célhoz vezet. Ha azonban végigszámoljuk ezt, akkor azt tapasztaljuk, hogy nem jutunk könnyen integrálható alakhoz.

A most elvégzendő helyettesítés bizonyos értelemben tipikus: ha a gyökjel alatt  $a - x^2$  típusú kifejezés van, akkor az  $x$  helyébe  $\sqrt{a} \cos t$  függvényt téve  $a - x^2 = a - a \cos^2 t = a(1 - \cos^2 t) = a \sin^2 t$ , és így el tudjuk végezni a négyzetgyökvonást (a  $\sin t$  nemnegatív). Ez az ötlet hasonló esetekben gyakorta megoldja a problémát. Hasonlóan alkalmas persze az  $x = \sin t$  helyettesítés is.

Lássuk ezután a feladat megoldását. Az  $x = r \cos t$  helyettesítés esetében  $t = \arccos \frac{x}{r}$ , és

$$\frac{dx}{dt} = r(-\sin t), \quad \text{azaz} \quad dx = -r \sin t dt.$$

Ezek alapján a helyettesítés végrehajtása:

$$\begin{aligned} \int \sqrt{r^2 - x^2} dx &= \int \sqrt{r^2 - r^2 \cos^2 t} (-r \sin t) dt = \\ &= - \int r \sin t \cdot r \sin t dt = -r^2 \int \sin^2 t dt. \end{aligned}$$

Az  $\sin^2 t$  függvényt már integráltuk a 7.8. példában, ahol azt kaptuk, hogy

$$\int \sin^2 t \, dt = -\frac{1}{2} \sin t \cos t + \frac{t}{2} + c = -\frac{1}{2} \cos t \sqrt{1 - \cos^2 t} + \frac{t}{2} + c.$$

A példában kitűzött feladat megoldásához már csak az kell, hogy a  $t$  helyére az  $\arccos \frac{x}{r}$  függvényt helyettesítsük, figyelembe véve azt, hogy  $\cos(\arccos \frac{x}{r}) = \frac{x}{r}$ :

$$\begin{aligned} \int \sqrt{r^2 - x^2} \, dx &= \frac{r^2}{2} \cos t \sqrt{1 - \cos^2 t} - \frac{r^2 t}{2} + c = \\ &= \frac{r^2}{2} \cdot \frac{x}{r} \cdot \sqrt{1 - \left(\frac{1}{x}\right)^2} - \frac{r^2 \arccos \frac{x}{r}}{2} + c = \\ &= \frac{x \sqrt{r^2 - x^2}}{2} - \frac{r^2 \arccos x/r}{2} + c. \end{aligned}$$

□

## 7.2 Differenciálegyenletek

Az előző pontban tárgyalt és begyakorolt antideriválási kalkulus technikailag lényegesen nehezebb a deriválásnál, több számolási fogást, formális érzéket igényel. Természetes, hogy nem a pusztán számolási technika csábított bennünket, hiszen ahogyan a bevezetésben is említettük, ez a kalkulus több fontos alkalmazásra kerülő matematikai tárgykörnek is a számítási eszköze. Ezt a legközvetlenebbül a differenciálegyenletek megoldásával kapcsolatban lehet látni.

Alaposabb definíció nélkül is megérthetjük, hogy milyen egyenletet értünk differenciálegyenleten: Olyan egyenletet, amelyben egy függvény, a deriváltja és a független változó között valami "függvény kapcsolat" van megadva. Például:

$$y(x) = 3y'(x), \quad (y'(x))^2 = x^4, \quad a(x)y'(x) + b(x)y(x) + c(x) = d(x),$$

ahol az  $a$ ,  $b$ ,  $c$  és  $d$  valamilyen függvényeket jelölnek. Természetesen számtalan típusú függvénykapcsolat írható fel. Mi néhány olyan egyszerű kapcsolatot fogunk tanulmányozni, amelyik mellett könnyen megoldható a differenciálegyenlet, és ne törekszünk a differenciálegyenletek rendszeres elméletének a kiépítésére.

A legegyszerűbb formájú differenciálegyenlet, amivel foglalkozunk

$$y'(x) = a(x) \tag{7.4}$$

típusú egyenlet. Ilyen feladat megoldására már teljesen felkészültek vagyunk, hiszen, amennyiben az  $a(x)$  függvénynek van  $A(x)$  antideriváltja, a megoldás az

$$\int a(x) \, dx = A(x) + c, \quad \text{a } c \text{ tetszőleges állandó}$$

függvény összesség tetszőleges eleme. Ebből a függvény-halmazból csak további feltétel, megkötés segítségével lehet egyetlen függvényt kiszemelni. A leggyakrabban használt feltétel az, hogy megadjuk azt, hogy a keresendő  $y$  függvény milyen  $b$  értéket vesz fel valamilyen  $\beta$  helyen:  $y(\beta) = b$ . Ennek a segítségével már meghatározható, hogy mi a  $c$  állandó értéke:

$$y(\beta) = A(\beta) + c = b,$$

amiből  $c = b - A(\beta)$ , tehát  $y = A(x) + (b - A(\beta))$ .

A  $c$  konstans meghatározásához megadott

$$y(\beta) = b \tag{7.5}$$

egyenletet *kezdeti-érték (feltételnek)* szokás nevezni. A (7.4) differenciálegyenlet és a (7.5) kezdeti-érték — megfelelő feltételek mellett — egyértelműen meghatározzák a keresett  $y$  függvényt.

A következőkben néhány olyan differenciálegyenletet fogunk vizsgálni, amelyeknek a megoldása a következő tételen alapszik

**Állítás 259 (Szétválasztható változójú egyenlet)**

*Ha az  $(a, b)$  intervallumban deriválható  $y = y(x)$  leképezés kielégíti az*

$$u(y) \frac{dy}{dx} = v(x) \tag{7.6}$$

*differenciálegyenletet, ahol az  $x \mapsto v(x)$  antideriválható az  $(a, b)$  intervallumban, az  $y \mapsto u(y)$  pedig egy az  $y((a, b))$  értékkészletet tartalmazó nyílt intervallumban, akkor*

$$\int u(y) dy \Big|_{y=y(x)} = \int v(x) dx. \tag{7.7}$$

Az (7.7) egyenlőség azt jelenti, hogy ha az  $U$  egy antideriváltja az  $u$  függvénynek, a  $V$  pedig a  $v$  függvénynek, akkor

$$U(y(x)) = v(x) + c.$$

Az (7.6) alakú differenciálegyenletet *szétválasztható változójú differenciálegyenletnek* nevezzük. Az elnevezés oka az, hogy az egyenletben az  $y$  és  $x$  változóktól való függés — az  $u(y)$  és  $v(x)$  függvények formájában — a felírás szerint szétválasztható.

**Bizonyítás.** Mivel tudjuk, hogy pontosan az állandó függvénynek a deriváltja nulla, ezért a (7.7) egyenlőséget igazoljuk, ha megmutatjuk, hogy

$$\frac{d}{dx} \left( \int u(y) dy \Big|_{y=y(x)} - \int v(x) dx \right) = 0.$$

A kijelölt deriválás elvégzésével ezt könnyen beláthatjuk. A határozatlan integrál definíciója és a közvetett függvény deriválási szabálya alapján:

$$\begin{aligned} & \frac{d}{dx} \left( \int u(y) dy \Big|_{y=y(x)} - \int v(x) dx \right) = \\ &= \frac{d}{dy} \int u(y) dy \Big|_{y=y(x)} \cdot \frac{d}{dx} y(x) - \frac{d}{dx} \int v(x) dx = \\ &= u(y)|_{y=y(x)} \cdot y'(x) - v(x), \end{aligned}$$

ami a (7.6) szerint nulla.  $\square$

**Példa 7.14 (Normális szaporodás növekedés folyamata)** *Jelölje egy populáció (valamilyen egyedekből álló összesség) elemeinek a számát  $y$ , ami a  $t$  időtől függ. Tegyük fel, hogy a populáció elemei számának a növekedési sebessége arányos a populáció tagjainak a számával, azaz az  $y$  kielégíti a*

$$\frac{dy}{dt} = \alpha y \quad (7.8)$$

*differenciálegyenletet. Oldjuk meg a differenciálegyenletet az*

$$y(t_0) = \beta \quad (7.9)$$

*kezdeti érték mellett, és analizáljuk az így leírt folyamatot.*

A (259). tétel alkalmazásához írjuk a (7.8) differenciálegyenletet az

$$\frac{1}{y} \frac{dy}{dt} = \alpha$$

alakba. A tétel szerint ebből azt kapjuk, hogy teljesül a következő egyenlőség

$$\int \frac{dy}{y} dy \Big|_{y=y(t)} = \int \alpha dt,$$

amiből

$$\ln y(t) = \alpha t + c, \quad \text{azaz} \quad y(t) = e^{\alpha t + c}.$$

A (7.9) kezdeti feltétel szerint

$$y(t_0) = e^{\alpha t_0 + c} = \beta,$$

amiből  $c = \ln \beta - \alpha t_0$ , és így a keresett megoldást már kiszámolhatjuk:

$$y(t) = e^{\alpha t + c} = e^{\alpha t + \ln \beta - \alpha t_0} = \beta e^{\alpha(t-t_0)}.$$

$\square$

Az  $y(t) = \beta e^{\alpha(t-t_0)}$  megoldást elemezve elmondhatjuk, hogy az egyedek száma igen gyorsan, exponenciális mértékben nő (vagy csökken, ha az  $\alpha$  negatív).

**Példa 7.15 (A robbanás egyenlete)** Jelölje  $y$  az egyedek időtől függő számát egy populációban. Tegyük fel, hogy az egyedszám növekedésének a sebessége arányos az  $y^2$ -tel, azaz az  $y(t)$  függvény eleget tesz a

$$\frac{dy}{dt} = \alpha y^2 \quad (7.10)$$

differenciálegyenletnek. Keressük meg, és elemezzük a megoldásokat.

A (259). tétel alkalmazásához írjuk a (7.10) differenciálegyenletet az

$$\frac{1}{y^2} \frac{dy}{dt} = \alpha$$

alakba. A tétel szerint ebből azt kapjuk, hogy teljesül a következő egyenlőség

$$\int \frac{dy}{y^2} \Big|_{y=y(t)} = \int \alpha dt,$$

amiből

$$-\frac{1}{y(t)} = \alpha t + c, \quad \text{azaz} \quad y(t) = \frac{1}{-\alpha t - c}.$$

A megoldásnak nyilvánvalóan pozitívnek kell lennie, ezért csak olyan  $c$  konstans jöhet szóba, amelyre  $-\alpha t - c > 0$ , azaz  $-c > \alpha t$ . A megoldás függvény képe hiperbola, amelyiknek az  $t = -\frac{c}{\alpha}$  egyenes az asszimptotája. E függvény által leírt folyamat kis  $t$  értékek mellett lassabban nő, mint az exponenciális függvény, a  $-\frac{c}{\alpha}$  ponthoz közel viszont a folyamat robbanásszerűen nő, hiszen a függvény határértéke a  $-\frac{c}{\alpha}$  helyen végtelen. Ez utóbbiból ered a folyamat elnevezése.  $\square$

**Példa 7.16 (Korlátozott normális növekedés egyenlete)** Tegyük fel, hogy egy populáció egyedei  $y$  számának a növekedési sebessége arányos egy  $A$  határesetnek és a pillanatnyi populáció számának az  $(A - y)$  különbségével, azaz az  $y(t)$  populáció szám kielégíti az alábbi differenciálegyenletet.

$$\frac{dy}{dt} = \alpha(A - y), \quad (7.11)$$

ahol  $t$  az időt jelöli. Oldjuk meg az egyenletet és analizáljuk a megoldást.

A (259). tétel alkalmazásához írjuk a (7.11) differenciálegyenletet az

$$\frac{1}{A - y} \frac{dy}{dt} = \alpha$$

alakba. A tétel szerint ebből azt kapjuk, hogy teljesül a következő egyenlőség

$$\int \frac{1}{A - y} dy \Big|_{y=y(t)} = \int \alpha dt + c,$$

Itt most óvatosan kell eljárunk, mert az  $A - y$  mennyiség negatív és pozitív egyaránt lehet, aszerint hogy a folyamat felülről vagy alulról közelíti az  $A$  határpopuláció számot.

1. eset: Ha  $A - y > 0$ , akkor az előző számolás folytatása:

$$-\ln(A - y(t)) = \alpha t + c, \quad \text{azaz} \quad y(t) = A - e^{-\alpha t - c}.$$

2. eset: Ha  $A - y < 0$ , akkor pedig a kiinduló számolás folytatása:

$$\ln(y(t) - A) = -\alpha t + c, \quad \text{azaz} \quad y(t) = A + e^{-\alpha t + c}.$$

Ha alaposan megnézzük a két megoldást, akkor azok alakja

$$y(t) = A + Ce^{-\alpha t} \tag{7.12}$$

ahol a  $C$  konstans az első esetben  $-e^{-c}$ , a második esetben pedig  $e^c$ .

A kapott függvény görbéjének a viselkedése könnyen leírható. Természetesen feltételezhetjük, hogy az  $A$  növekedési határ pozitív. Három esetet különböztetünk meg. (Az elmondottakat a differenciálható függvények diszkussziójánál tanultak szerint ellenőrizzük.)

$C < 0$ : Ekkor az  $y$  függvény a  $t = 0$  helyen az  $A + C$  értéket veszi fel, onnan monoton növekedőleg, exponenciális mértékben tart az  $A$  számhoz, és konkáv.

$C > 0$ : Ekkor a függvény a  $t = 0$  helyen  $A + C$ , ettől kezdve monoton fogyóan tart az  $A$ -hoz, konvex.

$C = 0$ : A függvény állandó. □

A következő példánkban szereplő növekedési folyamat különösen fontos, sokszor használt a közgazdaságtanban. A normális szaporodás esetében feltételeztük, hogy a szaporodásnak nincs akadálya, a korlátozott szaporodásnál egy fix határ helyetthez tartott a populáció száma. Most pedig egy olyan szaporodási folyamatot képzeljünk el, ahol az egyedek szaporodását az élelemért vagy egyéb jószágért folytatott versengés gátolja. Ilyen esetben úgy gondolkozhatunk, hogy eleinte olyan a helyzet, mintha normális szaporodásról lenne szó, mert kis egyed számnál nem gátolják egymást a szaporodásban. A kezdeti növekedést gyors növekedés követi, majd pedig lassú növekedés áll be, és végül tart egy fix állapothoz. Az ilyen görbék gráfjai egy nagy "S" betűre hasonlítanak, és valóban szokták is ezeket  $S$ -görbéknek nevezni. Ezen heurisztikus bevezetés után lássuk a pontos folyamatot.

**Példa 7.17** *Tegyük fel, hogy egy populáció  $t$  időbeli egyedei számának a növekedési sebessége arányos az egyedek számának és egy határhelyzetnek és az egyedek száma különbségének a szorzatával.*

*Pontos formális fogalmazással: Az  $y(t)$  populációszám eleget tesz a következő differenciálegyenletnek*

$$\frac{dy}{dt} = \alpha(A - y)y. \quad (7.13)$$

*Ennek a differenciálegyenleteknek a megoldásait logisztikus függvényeknek szokás nevezni (S-görbék). Oldjuk meg az egyenletet és analizáljuk a megoldásokat.*

Megjegyezzük, hogy az egyenletben szereplő szorzat azt mutatja, hogy az egyedek  $y$  számát és az  $A - y$  különbséget a folyamatra “függetlenül” ható tényezőnek képzeljük el.

A (259). tétel alkalmazásához írjuk a (7.13) differenciálegyenletet az

$$\frac{1}{(A - y)y} \frac{dy}{dt} = \alpha$$

alakba. A tétel szerint ebből azt kapjuk, hogy teljesül a következő egyenlőség

$$\int \frac{dy}{(A - y)y} \Big|_{y=y(t)} = \int \alpha dt. \quad (7.14)$$

A baloldali integrál meghatározásához felhasználjuk a következő azonosságot

$$\frac{1}{(A - y)y} = \frac{1}{A} \left( \frac{1}{y} + \frac{1}{A - y} \right).$$

Ennek alapján a (7.14) egyenlőség az alábbi formába írható:

$$\frac{1}{A} \left( \int \frac{dy}{y} dy + \int \frac{dy}{A - y} dy \right) \Big|_{y=y(t)} = \alpha t + c.$$

A baloldali első integrál  $\ln y$ , mivel az  $y$ -ről feltehetjük, hogy pozitív, a másik integrál kiszámításánál gondolni kell arra, hogy az  $A - y$  pozitív és negatív egyaránt lehet. Ha  $A - y > 0$ , akkor

$$\int \frac{1}{A - y} dy = \ln(A - y) + c;$$

ha pedig  $A - y < 0$ , akkor

$$\int \frac{1}{A - y} dy = - \int \frac{1}{y - A} dy = -\ln(y - A) + c.$$

Számoljunk most tovább ezen két eset szerint: Ha  $(A - y) > 0$ , akkor

$$\frac{1}{A} (\ln y(t) - \ln(A - y(t))) = \alpha t + c.$$

Fejezzük ki ebből az  $y$  változót:

$$\ln\left(\frac{y}{A-y}\right) = \alpha At + Ac, \quad \text{azaz} \quad \frac{y}{A-y} = e^{\alpha At + Ac}.$$

Ebből egyszerű számolással kapjuk, hogy

$$y = \frac{Ae^{\alpha At + Ac}}{1 + e^{\alpha At + Ac}} = \frac{A}{1 + e^{-(\alpha At + Ac)}} \quad (7.15)$$

Ha  $(A - y) < 0$ , akkor

$$\frac{1}{A}(\ln y(t) - \ln(y(t) - A)) = \alpha t + c.$$

Ebből kifejezve az  $y$  változót, hasonlóan számolva, mint az előbb azt kapjuk, hogy

$$y(t) = \frac{Ae^{\alpha At + Ac}}{e^{\alpha At + Ac} - 1} = \frac{A}{1 - e^{-(\alpha At + Ac)}}. \quad (7.16)$$

□



## 8.

# Határozott integrál

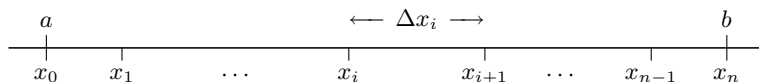
A matematikai analízis bevezetésével foglalkozó könyvek gyakori címe: “Bevezetés a differenciál- és integrálszámításba”. A differenciálszámítással és annak a megfordításával a határozatlan integrálnak a kiszámításával már foglalkoztunk. Ezzel túljutottunk a “kalkulus” jellegű témakörök döntő részén. Ebben a fejezetben a határozott integrál fogalmával foglalkozunk. A fejezet jórészt elmélet lesz, mert a számítási módszerek már rendelkezésünkre állnak. A határozatlan integrál voltaképpen a deriválás megfordítása, és így sokkal logikusabb az antiderivált szó használata. A határozott integrál az “igazi” integrálfogalom, és el is szokás hagyni a “határozott” jelzőt.

Az első pontban a Riemann-integrál elméletének az alapjait írjuk le. A második pontban a számítási módszereket gyűjtjük egybe, és feladatokat oldunk meg. A harmadik pontban a Riemann-integrál egy kiterjesztését, az improprius integrál fogalmát tárgyaljuk.

## 8.1 A Riemann integrál definíciója

A kiinduló geometriai problémánk most a következő. Vegyünk egy  $f : [a, b] \rightarrow \mathbb{R}$  leképezést, és tegyük fel a kérdést: Hogyan tudnánk megalkotni a függvény grafikonja és az  $x$  tengely közötti terület fogalmát? Szándékosan kerültük annak a szónak a használatát, hogy ki szeretnénk számítani a szóbanforgó területet, mert ilyen általános halmaz területét még nem definiáltuk. Éppen a tárgyalandó integrál definícióján keresztül fogjuk megadni a terület fogalmát, ahogyan az érintőt is a differenciálhányadoson keresztül definiáltuk.

Feltesszük, hogy az  $\alpha$  és  $\beta$  oldalú téglalap területe  $\alpha \cdot \beta$ . Ebből kiindulva szeretnénk megragadni a görbe alatti területet, a következő alapelv szerint: Olyan téglalapokkal töltjük ki a görbe alatti halmazt, amelyek a lehetőségekhez képest nagyon kis részt hagynak lefedetlenül, és egyre finomabb lefedést tesznek lehetővé (8.2. ábra).

8.1. ábra: Az  $[a, b]$  egy felosztása.

**Definíció 260** Legyenek az  $x_0, x_1, \dots, x_{n-1}$  és  $x_n$  olyan pontjai az  $[a, b]$  intervallumnak, hogy

$$x_0 = a, x_n = b \quad \text{és} \quad x_0 < x_1 < \dots < x_{n-1} < x_n.$$

Ekkor az  $\{x_0, x_1, \dots, x_{n-1}, x_n\}$  pontok az  $[a, b]$  intervallumnak az

$$[x_i, x_{i+1}], \quad (i = 0, 1, \dots, n-1)$$

részintervallumokra való felosztását határozzák meg. A felosztásban szereplő részintervallumok hosszát a

$$\Delta x_i = x_{i+1} - x_i, \quad (i = 0, 1, \dots, n-1)$$

szimbólum jelöli. A felosztást megadó  $x_0, x_1, \dots, x_n$  pontokat a felosztás osztópontjainak mondjuk.

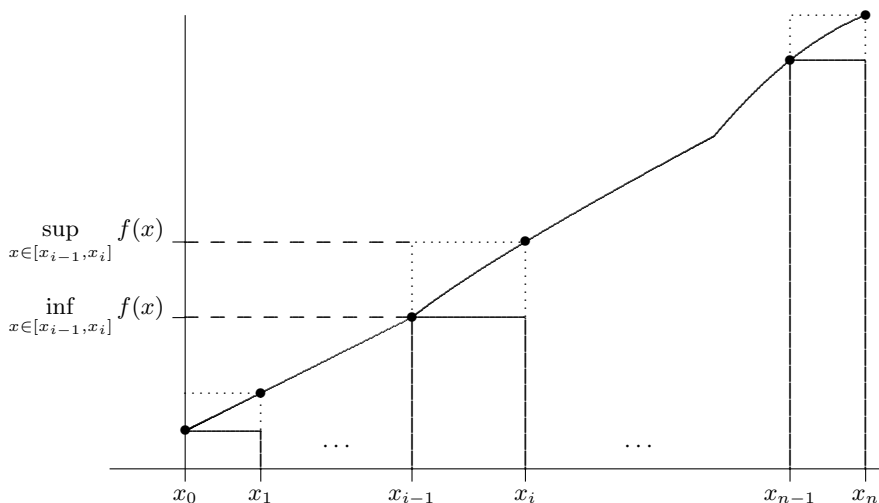
A 8.1. ábrán szemléltetünk egy felosztást. A definíció szerint egy felosztást az osztópontjai határoznak meg. A felosztások tömörebb jelölésére nagy betűket fogunk használni, és ha például az  $I$  jelöli az  $x_0, x_1, \dots, x_n$  osztópontok által megadott felosztást, akkor azt írjuk, hogy

$$I = \{x_0, x_1, \dots, x_n\}.$$

**Definíció 261** Az  $[a, b]$  intervallum egy  $J$  felosztására azt mondjuk, hogy finomabb, mint az  $I$  felosztás, ha az  $I$  felosztás minden osztópontja osztópontja a  $J$  felosztásnak is.

Az  $a$  és  $b$  pontok az  $[a, b]$  intervallum minden felosztásának osztópontjai, ezért az  $\{a, b\}$  felosztásnál minden felosztás finomabb, amit úgy is mondhatunk, hogy az  $\{a, b\}$  a legdurvább felosztás. A legfinomabb felosztásról nem beszélhetünk, hiszen tetszőleges felosztáshoz hozzávéve még egy osztópontot, finomabb felosztást kapunk. Ha az  $I$  és  $J$  két tetszőleges felosztás, akkor könnyű olyan felosztást mondani, amelyik mindkettőnél finomabb: például az  $a$  felosztás, amelyiknek az osztópontjai az  $I$  és  $J$  osztópontjainak az egyesítése. Ezt a felosztást az  $I$  és  $J$  felosztások közös finomításának fogjuk nevezni.

Most pedig, a görbe alatti terület alsó és felső megközelítéséhez, definiáljuk a felosztásokhoz tartozó közelítő összegeket. A 8.2. ábrán illusztráljuk a három összeget.



8.2. ábra: Az integrál közelítő összegek.

**Definíció 262** Legyen az  $f$  függvény az  $[a, b]$  intervallumon korlátos, és az  $I = \{x_0, x_1, \dots, x_n\}$  az intervallum egy felosztása. Az  $f$  függvényhez és az  $I$  felosztáshoz a következő közelítő összegeket vezetjük be:

1. Alsó közelítő összeg (röviden: alsó összeg):

$$s(f, I) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot \inf_{x \in [x_{i-1}, x_i]} f(x).$$

2. Felső közelítő összeg (röviden: felső összeg):

$$S(f, I) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot \sup_{x \in [x_{i-1}, x_i]} f(x).$$

3. Egy közbülső közelítő összeg (röviden: közbülső összeg):

$$m(f, I, \xi_1 \dots \xi_n) = m(f, I) = \sum_{i=1}^n (x_i - x_{i-1}) f(\xi_i).$$

A definíciók értelmességéhez először is jegyezzük meg, hogy a bennük szereplő alsó és felső határok léteznek, mivel az  $f$  függvény korlátos. Hangsúlyozzuk, hogy a közbülső közelítő összeg definíciójában lévő  $\xi_i$  szám tetszőleges az  $[x_{i-1}, x_i]$  intervallumban, tehát ilyen  $m(f, I)$  összeg — ellentétben az alsó és felső közelítő összegekkel — adott felosztás esetén is több van.

A következő állításban összefoglaljuk a közelítő összegek legfontosabb tulajdonságait.

**Állítás 263** A 262. definíció jelöléseivel, az integrál közelítő összegeknek a következő tulajdonságai vannak.

1. Minden  $I$  felosztás esetén bármely közbülső összegre

$$s(f, I) \leq m(f, I) \leq S(f, I).$$

2. Ha az  $I$  felosztás finomabb, mint a  $J$  felosztás, akkor

$$s(f, J) \leq s(f, I) \quad \text{és} \quad S(f, J) \geq S(f, I).$$

3. Tetszőleges  $I$  és  $J$  felosztásokra

$$s(f, I) \leq S(f, J).$$

4. Az alsó közelítő összegek halmazának van felső, a felső közelítő összegek halmazának pedig alsó határa, és

$$\sup_I s(f, I) \leq \inf_I S(f, I).$$

A tétel első három állítását célszerű szavakban is megfogalmazni:

1. Egy felosztáshoz tartozó alsó és felső közelítő összeg közrefogja a felosztáshoz tartozó közbülső közelítő összegeket.
2. A felosztások finomításával az alsó közelítő összegek monoton nőnek, a felső közelítő összegek pedig monoton csökkennek.
3. Minden alsó közelítő összeg kisebb minden felső közelítő összegnél, a felosztásoktól függetlenül.

**Bizonyítás. (1):** Mivel az

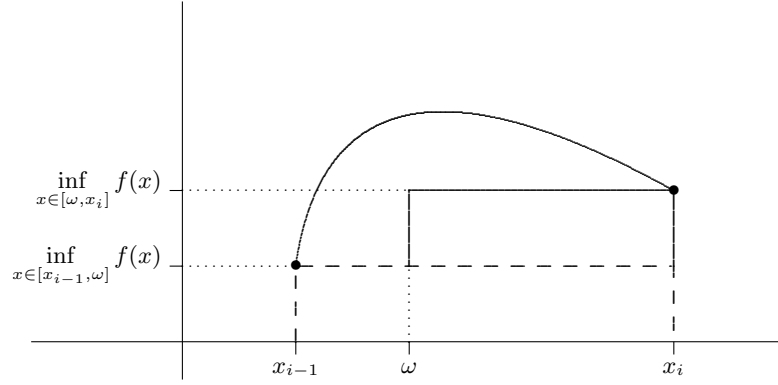
$$\inf_{x \in [x_{i-1}, x_i]} f(x) \leq f(\xi_i) \leq \sup_{x \in [x_{i-1}, x_i]} f(x)$$

egyenlőtlenség alapján

$$\inf_{x \in [x_{i-1}, x_i]} f(x) \Delta x_{i-1} \leq f(\xi_i) \Delta x_{i-1} \leq \sup_{x \in [x_{i-1}, x_i]} f(x) \Delta x_{i-1},$$

ezért

$$\sum_{i=1}^n \inf_{x \in [x_{i-1}, x_i]} f(x) \Delta x_{i-1} \leq \sum_{i=1}^n f(\xi_i) \Delta x_{i-1} \leq \sum_{i=1}^n \sup_{x \in [x_{i-1}, x_i]} f(x) \Delta x_{i-1}.$$



8.3. ábra: Egy osztópont közbeiktatásának a hatása.

**(2):** Lássuk be az alsó közelítő összegek esetét. A felső közelítő összegekre vonatkozó megfelelő állítás bizonyítását az olvasóra hagyjuk. Elégséges azt megmutatnunk, hogy egy új osztópont közbeiktatásával az alsó közelítő összeg növekedik (nem csökken), mivel ezen az úton egy felosztásból egy tetszőleges nála finomabb felosztáshoz eljuthatunk.

Legyen az  $I = \{x_0, x_1, \dots, x_n\}$  egy tetszőleges felosztása az  $[a, b]$  intervallumnak. Azt a felosztást, amelyik abban különbözik az  $I$  felosztástól, hogy van egy új  $\omega$  osztópontja, amelyre  $x_{i-1} < \omega < x_i$ , jelöljük  $J$ -vel. Ekkor az  $s(f, I)$  és  $s(f, J)$  összegek csak abban az összeadandóban térnek el, amelyikben az új osztópont van (8.3. ábra), mégpedig

$$s(f, I) - s(f, J) = (x_i - x_{i-1}) \cdot \inf_{x \in [x_{i-1}, x_i]} f(x) - (\omega - x_{i-1}) \cdot \inf_{x \in [x_{i-1}, \omega]} f(x) - (x_i - \omega) \cdot \inf_{x \in [\omega, x_i]} f(x). \quad (8.1)$$

Mivel

$$\inf_{x \in [x_{i-1}, x_i]} f(x) \leq \inf_{x \in [x_{i-1}, \omega]} f(x)$$

és

$$\inf_{x \in [x_{i-1}, x_i]} f(x) \leq \inf_{x \in [\omega, x_i]} f(x),$$

ezért a (8.1) egyenlőség alapján

$$\begin{aligned} s(f, I) - s(f, J) &\leq \\ &= (x_i - x_{i-1}) \inf_{x \in [x_{i-1}, x_i]} f(x) - (\omega - x_{i-1}) \cdot \inf_{x \in [x_{i-1}, x_i]} f(x) - (x_i - \omega) \cdot \inf_{x \in [x_{i-1}, x_i]} f(x) = \\ &= [x_i - x_{i-1} - (\omega - x_{i-1}) - (x_i - \omega)] \cdot \inf_{x \in [x_{i-1}, x_i]} f(x) = 0, \end{aligned}$$

tehát  $s(f, I) \leq s(f, J)$ , amit bizonyítanunk kellett.

(3): Jelöljük  $L$ -l az  $I$ -t felosztást, amelynek az osztópontjai az  $I$  és  $J$  felosztások osztópontjainak az egyesítése (tehát az  $I$  és  $J$  közös finomítását). Az  $L$  felosztás finomabb mind az  $I$  mind a  $J$  felosztásnál, ezért a jelenlegi tétel már belátott első és második állítása alapján

$$s(f, I) \leq s(f, L) \leq S(f, L) \leq S(f, J). \quad (8.2)$$

(4): Az alsó közelítő összegek halmaza nyilván nem üres és az előbb belátott állítás szerint felülről korlátos, hiszen egy tetszőleges felső közelítő összeg felső korlát, ezért az alsó közelítő összegeknek van felső határa. Hasonlóan indokolható az, hogy a felső közelítő összegeknek van alsó határa.  $\square$

Induljunk most ki az előbbi tétel harmadik és negyedik állításából. Ezek szerint az  $f$  alsó közelítő összegeinek a felső határa nem nagyobb, mint a felső közelítő összegek alsó határa, azaz:

$$\sup_I s(f, I) \leq \inf_I S(f, I).$$

Számunkra az az eset a döntő, amikor egyenlőség van, amit a következő definícióban rögzítünk.

**Definíció 264** Egy  $f : [a, b] \rightarrow \mathbb{R}$  korlátos függvényt Riemann szerint integrálhatónak mondunk, ha az alsó közelítő összegek felső határa megegyezik a felső közelítő összegek alsó határával. Az egyező értéket az  $f$  függvény  $[a, b]$  intervallumon vett Riemann-integráljának mondjuk, és a következőképpen jelöljük:

$$\int_a^b f(x) dx, \quad \text{vagy} \quad \int_a^b f$$

Megállapodunk abban, hogy

$$\int_a^a f(x) dx = 0$$

és

$$\int_b^a f(x) dx = - \int_a^b f(x) dx.$$

**Definíció 265** Az

$$\begin{aligned} \Omega(f, I) &= S(f, I) - s(f, I) \\ &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot \sup_{x \in [x_{i-1}, x_i]} f(x) - \sum_{i=1}^n (x_i - x_{i-1}) \cdot \inf_{x \in [x_{i-1}, x_i]} f(x) = \\ &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot \left( \sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x) \right) = \\ &= \sum_{i=1}^n (x_i - x_{i-1}) \sup_{x, y \in [x_{i-1}, x_i]} |f(x) - f(y)| \end{aligned}$$

kifejezést az  $I$  felosztáshoz tartozó oszcillációs összegnek szokás nevezni. A

$$\sup_{x,y \in [x_{i-1}, x_i]} |f(x) - f(y)|$$

érték az  $f$  függvény ingadozása, oszcillációja, az  $[x_{i-1}, x_i]$  intervallumon, és az oszcillációs összeg az  $f$  részintervallumonkénti ingadozásának az  $I$  felosztás intervallumaival súlyozott összege.

**Állítás 266** Egy  $f : [a, b] \rightarrow \mathbb{R}$  korlátos függvény pontosan akkor integrálható, ha tetszőleges  $\varepsilon$  pozitív számhoz van olyan

$$I = \{x_0, x_1, \dots, x_n\}$$

felosztása az  $[a, b]$  intervallumnak, hogy

$$\Omega(f, I) = S(f, I) - s(f, I) \leq \varepsilon.$$

Az állítás szerint az integrálhatóság szükséges és elégséges feltétele az, hogy az oszcillációs összeg “tetszőlegesen kicsi” (adott  $\varepsilon$ -nál kisebb) legyen alkalmas felosztás mellett.

**Bizonyítás:** Ha a tétel feltétele teljesül, akkor az  $f$  integrálható: Tegyük fel, hogy az  $\alpha \leq \beta$  számok az alsó és felső közelítő összegek közé esnek. Tetszőleges  $\varepsilon$  pozitív számhoz van olyan  $I$  felosztás, amelyre

$$\Omega(f, I) = S(f, I) - s(f, I) \leq \varepsilon.$$

Emiatt viszont  $\beta - \alpha \leq \varepsilon$ , ami csak úgy lehet — lévén az  $\varepsilon$  tetszőleges —, ha  $\alpha = \beta$ , tehát az alsó és felső közelítő összegek közé egyetlen szám esik, azaz a függvény integrálható.

Ha az  $f$  integrálható, akkor teljesül a tétel feltétele: Legyen az  $\varepsilon$  egy tetszőleges pozitív szám. Mivel most a felső közelítő összegek infimuma és az alsó közelítő szuprimuma megegyezik, ezért csak egy  $A$  szám esik az alsó és felső közelítő összegek közé, így van olyan  $s(f, I)$  alsó és  $S(f, J)$  felső közelítő összeg, hogy

$$S(f, J) - s(f, I) = (S(f, J) - A) + (A - s(f, I)) \leq \varepsilon. \quad (8.3)$$

Az az  $L$  felosztás, amelynek az osztópontjai az  $I$  és  $J$  osztópontjainak az egyesítése, finomabb mindkét felosztásnál, és így a 263. (2) állítás felhasználásával

$$S(f, J) - s(f, I) \geq S(f, L) - s(f, L),$$

amiből a (8.3) egyenlőtlenség szerint adódik, hogy

$$S(f, L) - s(f, L) \leq \varepsilon,$$

tehát az  $L$  felosztással teljesül a tétel feltétele.  $\square$

**Állítás 267** Egy  $f : [a, b] \rightarrow \mathbb{R}$  korlátos függvény pontosan akkor integrálható, ha létezik olyan  $A$  szám, hogy tetszőleges  $\varepsilon$  pozitív számhoz van olyan

$$I = \{x_0, x_1, \dots, x_n\}$$

felosztása az  $[a, b]$  intervallumnak, hogy

$$S(f, I) - \varepsilon \leq A \leq s(f, I) + \varepsilon.$$

Ez az  $A$  szám, ha létezik, éppen az  $f$  integrálja az  $[a, b]$  intervallumon, vagyis  $A = \int_a^b f(x) dx$ .

**Bizonyítás:** Az egyik irány az előző állítás alapján nyilvánvaló. Ha pedig az  $f$  integrálható, akkor szintén a fentiek miatt minden  $\varepsilon$  számhoz van olyan  $I$ , hogy

$$S(f, I) - \varepsilon \leq s(f, I) \leq S(f, I) + \varepsilon.$$

De ekkor

$$S(f, I) - \varepsilon \leq s(f, I) \leq \int_a^b f(x) dx \leq S(f, I) \leq s(f, I) + \varepsilon.$$

□

**Példa 8.1** Határozzuk meg az  $[a, b]$  intervallumon konstans függvény határozott integrálját.

Legyen a konstans  $c$ , és jelölje a függvényt a  $h$ . Tetszőleges

$$I = \{x_0, x_1, \dots, x_n\}$$

felosztás mellett

$$s(h, I) = S(h, I) = \sum_{i=1}^n c \cdot (x_{i+1} - x_i) = c(x_n - x_0) = c(b - a),$$

tehát a  $c$  értéket felvevő konstans függvény integrálja  $c(b - a)$ .

□

A következő feladat azért érdekes, mert példát ad olyan korlátos függvényre, amelyik definíciónk szerint nem integrálható.

**Példa 8.2** Legyen a  $g(x)$   $x \in [0, 1]$  függvény értéke 1, ha az  $x$  racionális, és 0, ha az  $x$  irracionális. Számoljuk ki az alsó illetve felső közelítő összegek felső illetve alsó határát.

Tetszőleges  $I$  felosztás mellett egy részintervallumon a függvény infimuma (minimuma) 0, a szuprémuma (maximuma) pedig 1, mivel minden (valódi) intervallumon van mind racionális mind irracionális szám. Emiatt az alsó közelítő összeg 0, a felső közelítő összeg pedig 1. Emiatt az alsó közelítő összegek felső határa 0, az alsó közelítő összegek alsó határa pedig 1, és mivel ezek különbözőek, ezért a függvény nem integrálható.

□

**Példa 8.3** Legyen a  $g(x)$   $x \in [0, 1]$  függvény értéke véges sok  $x_1, \dots, x_n$  ponttól eltekintve 0, az  $x_i$  ( $i = 1, \dots, n$ ) pontokban pedig az értéke legyen 1. Számoljuk ki az alsó illetve a felső közelítő összegek felső illetve alsó határát.

Mivel minden felosztás minden részintervallumában van olyan szám, amelyre a  $g$  függvény nulla értéket vesz fel, ezért nyilván minden alsó közelítő összeg 0, és így az alsó közelítő összegek felső határa is 0. Legyen  $\varepsilon > 0$  tetszőleges és legyen  $I$  egy olyan felosztás, amelyre a leghosszabb részintervallum is rövidebb mint  $\varepsilon/n$ . Mivel csak  $n$  olyan pont van, ahol a  $g$  értéke 1, és mivel egy intervallum hossza legfeljebb  $\varepsilon/n$  ezért a felosztásra a felső közelítő összeg értéke legfeljebb  $\varepsilon$ . Mivel  $\varepsilon$  tetszőleges ezért a felső közelítő összegek alsó határa szintén nulla, vagyis a  $g$  függvény integrálható, és  $\int_0^1 g(x) dx = 0$ .  $\square$

## 8.2 Integrálhatósági tételek

Ezt az alpontot két erősen összefüggő kérdés vizsgálatára szánjuk. Egyrészt megvizsgáljuk, hogyan viselkedik az integrálás a függvényekkel végzett operációkkal szemben, másrészt megnézzük, hogy a függvények milyen körére terjed ki az integrálhatóság. Az első kérdés — az előzőekben már használt elnevezésünk szerint — az integrálás formális szabályait jelenti. Ennek keretében először is belátjuk, hogy valamely  $[a, b]$  intervallumon integrálható függvények összessége vektortér. Ehhez szükségünk lesz a következő egyszerű állításra, aminek a bizonyítását az olvasóra bizzuk.

**Állítás 268** Ha  $f$  és  $g$  az  $[a, b]$  intervallumon korlátos függvények, akkor tetszőleges  $L$  felosztásra

$$s(f, L) + s(g, L) \leq s(f + g, L) \quad \text{és} \quad S(f, L) + S(g, L) \geq S(f + g, L).$$

Ha  $\lambda > 0$ , akkor

$$\lambda s(f, L) = s(\lambda f, L) \quad \text{és} \quad \lambda S(f, L) = S(\lambda f, L).$$

Ha  $\lambda < 0$ , akkor

$$\lambda s(f, L) = S(\lambda f, L) \quad \text{és} \quad \lambda S(f, L) = s(\lambda f, L).$$

**Állítás 269** Legyenek  $f$  és  $g$  az  $[a, b]$  intervallumon integrálható függvények, és  $\lambda$  tetszőleges valós szám. Ekkor az  $f + g$  és  $\lambda f$  függvények is integrálhatóak, és

$$(1) \quad \int_a^b (f(x) + g(x)) dx = \int_a^b f(x) dx + \int_a^b g(x) dx,$$

$$(2) \quad \int_a^b \lambda f(x) dx = \lambda \cdot \int_a^b f(x) dx.$$

**Bizonyítás. (1):** Mivel az  $f$  és  $g$  függvények integrálhatóak, ezért adott  $\varepsilon$  pozitív számhoz van olyan  $I$ , hogy ha  $L$  finomabb mint  $I$ , akkor

$$\int_a^b f(x) dx - s(f, L) \leq \varepsilon/2$$

és

$$S(f, L) - \int_a^b f(x) dx \leq \varepsilon/2.$$

Hasonlóan, van olyan  $J$  felosztás, hogy ha  $L$  finomabb mint  $J$ , akkor

$$\int_a^b g(x) dx - s(g, L) \leq \varepsilon/2$$

és

$$S(g, L) - \int_a^b g(x) dx \leq \varepsilon/2.$$

Ha  $L$  az  $I$  és  $J$  felosztások közös finomítása, akkor a fenti egyenlőtlenségek teljesülnek. A megfelelő egyenlőtlenségek összeadva

$$\begin{aligned} \left[ \int_a^b f(x) dx + \int_a^b g(x) dx \right] - [s(f, L) + s(g, L)] &\leq \varepsilon, \\ [S(f, L) + S(g, L)] - \left[ \int_a^b f(x) dx + \int_a^b g(x) dx \right] &\leq \varepsilon. \end{aligned}$$

Ezekből pedig, figyelembe véve a tételt megelőző állítást, kapjuk, hogy

$$\begin{aligned} \left[ \int_a^b f(x) dx + \int_a^b g(x) dx \right] - s(f+g, L) &\leq \varepsilon, \\ S(f+g, L) - \left[ \int_a^b f(x) dx + \int_a^b g(x) dx \right] &\leq \varepsilon. \end{aligned}$$

Ez pedig éppen azt jelenti, hogy az  $(f+g)$  integrálható, és az integrálja

$$\int_a^b f(x) dx + \int_a^b g(x) dx.$$

**(2):** Ha a  $\lambda$  szám nulla, akkor nyilvánvalóan igaz az állítás, ezért a továbbiakban feltesszük, hogy  $\lambda \neq 0$ . Mivel az  $f$  integrálható, ezért adott  $\varepsilon$  pozitív számhoz van olyan  $L$  felosztás, hogy

$$S(f, L) - \frac{\varepsilon}{|\lambda|} \leq \int_a^b f(x) dx \leq s(f, L) + \frac{\varepsilon}{|\lambda|}.$$

Az egyenlőtlenségeket a  $|\lambda|$  pozitív számmal beszorozva

$$|\lambda|S(f, L) - \varepsilon \leq |\lambda| \int_a^b f(x) dx \leq |\lambda|s(f, L) + \varepsilon.$$

Tovább alakítva, aszerint hogy a  $\lambda$  milyen előjelű, figyelembe véve az előző állítást a következőképpen számolhatunk:

Ha  $\lambda > 0$ , akkor:

$$S(\lambda f, L) - \varepsilon \leq \lambda \int_a^b f(x) dx \leq s(\lambda f, L) + \varepsilon,$$

amiből már következik, hogy

$$\lambda \int_a^b f(x) dx = \int_a^b \lambda f(x) dx.$$

Ha  $\lambda < 0$ , akkor:

$$-s(\lambda f, L) - \varepsilon \leq -\lambda \int_a^b f(x) dx \leq -S(\lambda f, L) + \varepsilon.$$

Ezt  $-1$ -gyel beszorozva az előző esethez hasonlóan kapjuk, hogy

$$\lambda \int_a^b f(x) dx = \int_a^b \lambda f(x) dx.$$

□

**Állítás 270** Ha  $f$  és  $g$  az  $[a, b]$  intervallumon integrálható függvények és minden  $x \in [a, b]$  pontban  $f(x) \leq g(x)$ , akkor  $\int_a^b f(x) dx \leq \int_a^b g(x) dx$ .

**Bizonyítás.** A tétel feltételei alapján nyilvánvalóan minden  $I$  felosztásra

$$\int_a^b f(x) dx \leq S(f, I) \leq S(g, I),$$

amiből

$$\int_a^b f(x) dx \leq \inf_I S(g, I) = \int_a^b g(x) dx.$$

□

**Állítás 271** Ha az  $f$  függvény az  $[a, b]$  intervallumon integrálható, akkor integrálható az  $|f|$  függvény is, és

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

**Bizonyítás.** Ha  $x, y$  tetszőleges elemei az  $[a, b]$  intervallumnak, akkor

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

Ezért tetszőleges  $I$  felosztáshoz tartozó oszcillációs összegre

$$\Omega(|f|, I) \leq \Omega(f, I).$$

Ebből az  $|f|$  integrálhatósága már világos, hiszen tetszőleges  $\varepsilon > 0$  számhoz az  $f$  integrálhatósága miatt létezik olyan  $I$  felosztás, amelyre

$$\Omega(f, I) \leq \varepsilon,$$

így az

$$\Omega(|f|, I) \leq \Omega(f, I) \leq \varepsilon$$

egyenlőtlenség is teljesül, amely éppen az  $|f|$  integrálhatóságát jelenti. Mivel

$$-|f(x)| \leq f(x) \leq |f(x)|,$$

ezért az előző állítás alapján

$$\int_a^b -|f(x)| dx \leq \int_a^b f(x) dx \leq \int_a^b |f(x)| dx.$$

Mivel a skalár kivihető az integrál elé ezért egyúttal a

$$-\int_a^b |f(x)| dx \leq \int_a^b f(x) dx \leq \int_a^b |f(x)| dx$$

is teljesül, ami éppen a bizonyítandó

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$$

egyenlőtlenség. □

**Állítás 272** Ha az  $f$  függvény integrálható az  $[a, b]$  intervallumon, és a  $g$  függvény véges számú ponttól eltekintve azonos az  $f$  függvényel, akkor a  $g$  függvény is integrálható és  $\int_a^b f(x) dx = \int_a^b g(x) dx$ .

**Bizonyítás.** Elegendő azzal az esettel foglalkozni, ha a  $g$  egyetlen  $x_0$  pontban tér el az  $f$  függvénytől. Az integrál linearitása miatt viszont elegendő azt megmutatni, hogy annak a  $h$  függvénynek az integrálja, amely az  $x_0$  pontban 1 és az  $[a, b]$  többi pontjában 0 az integrálja 0, hiszen  $g = f + g(x_0)h$ . Az pedig, hogy a  $h$  integrálja nulla következik az előző pont végén bemutatott harmadik példából. □

**Állítás 273** *Ha az  $f$  függvény integrálható az  $[a, b]$  intervallumon, és  $c \in [a, b]$  tetszőleges szám, akkor az  $f$  integrálható az  $[a, c]$  és  $[c, b]$  intervallumokon is és*

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

**Bizonyítás.** Először megmutatjuk, hogy ha az  $f$  integrálható az  $[a, b]$  intervallumon, és  $[d, e] \subseteq [a, b]$ , akkor az  $f$  integrálható a  $[d, e]$  intervallumon is. Mivel az  $f$  korlátos az  $[a, b]$  intervallumon nyilván korlátos a  $[d, e]$  részintervallumon is. Legyen  $\varepsilon > 0$  tetszőleges, és legyen  $I$  az  $[a, b]$  egy olyan felosztása, amelyre

$$\Omega(f, I, [a, b]) = S(f, I, [a, b]) - s(f, I, [a, b]) < \varepsilon.$$

Mivel további osztópontok hozzávételével a felső közelítő összeg csökken az alsó pedig nő, ezért feltehető, hogy az  $I$  osztópontjai között már szerepelnek a  $d$  és  $e$  pontok is. Mivel az oszcillációs összeg minden tagja nemnegatív, ezért a  $[d, e]$  halmazon kívüli intervallumokat elhagyva kapjuk, hogy

$$\Omega(f, I, [d, e]) \leq \Omega(f, I, [a, b]) < \varepsilon.$$

Tehát az  $f$  integrálható a  $[a, b]$  minden  $[d, e]$  részintervallumán. Jelölje  $\chi_{[d, e]}$  a  $[d, e]$  intervallum karakterisztikus függvényét. A bizonyítás második lépéseként belátjuk, hogy

$$\int_d^e f(x) dx = \int_a^b f(x) \chi_{[d, e]}(x) dx. \quad (8.4)$$

Legyen  $\varepsilon > 0$  tetszőleges és legyen  $I$  a  $[d, e]$  egy olyan felosztása melyre

$$S(f, I, [d, e]) - \int_d^e f(x) dx < \varepsilon$$

és

$$\int_d^e f(x) dx - s(f, I, [d, e]) < \varepsilon.$$

Mivel az  $I$  felosztás kiegészítve az  $a$  és  $b$  pontokkal egyúttal az  $[a, b]$  egy olyan  $J$  felosztása, melyre

$$S(f\chi, J, [a, b]) = S(f, I, [d, e])$$

és

$$s(f\chi, J, [a, b]) = s(f, I, [d, e])$$

ezért

$$S(f\chi, J, [a, b]) - \int_d^e f(x) dx < \varepsilon$$

$$\int_d^e f(x) dx - s(f\chi, J, [a, b]) < \varepsilon,$$

amiből a (8.4) már nyilvánvalóan következik. A bizonyítás harmadik lépéseként elegendő hivatkozni arra, hogy összeg integrálja az integrálok összege, és így

$$\begin{aligned}\int_a^b f(x) dx &= \int_a^b f(x) (\chi_{[a,c]}(x) + \chi_{(c,b]}(x)) = \int_a^b f(x) (\chi_{[a,c]}(x) + \chi_{[c,b]}(x)) = \\ &= \int_a^b f(x) \chi_{[a,c]}(x) dx + \int_a^b f(x) \chi_{[c,b]}(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.\end{aligned}$$

□

A következő két tételben a folytonos illetve a monoton függvények integrálhatóságát bizonyítjuk be.

**Állítás 274** *Ha az  $f : [a, b] \rightarrow \mathbb{R}$  függvény folytonos, akkor integrálható az  $[a, b]$  intervallumon.*

**Bizonyítás.** Először is jegyezzük meg, hogy egy zárt intervallumon folytonos függvény korlátos, mivel felveszi a maximumát és minimumát (131. tétel).

Legyen az  $\varepsilon > 0$  tetszőleges. Az  $f$  integrálhatóságának belátásához elegendő megmutatni, hogy létezik az  $[a, b]$  intervallumnak egy olyan  $I$  partíciója, amelyre az  $\Omega(f, I)$  oszcillációs összeg kisebb mint  $\varepsilon$ . Az állítást indirekt fogjuk belátni. Tegyük fel, hogy az állítással ellentétben ilyen partíció nincs, és tekintsük az  $[a, b]$  intervallum  $n$  egyenlő részintervallumra való felosztását, amikor két osztópont távolsága éppen  $(b - a)/n$ . Mivel a partícióhoz tartozó oszcillációs összeg nagyobb mint  $\varepsilon$ , ezért van olyan részintervallum, amelyen a függvény ingadozása nagyobb mint  $\varepsilon/(b - a)$ , és így van az intervallumnak olyan  $x_n$  és  $y_n$  pontja, melyek távolsága kisebb mint  $(b - a)/n$ , de a függvényértékek  $|f(x_n) - f(y_n)|$  távolsága nagyobb mint  $\varepsilon/(b - a)$ . A Bolzano-Weierstrass tétel alapján feltehetjük, hogy a  $(x_n)_{n=1}^\infty$  és az  $(y_n)_{n=1}^\infty$  sorozatoknak van konvergens részsorozata. Mivel az  $(x_n)_{n=1}^\infty$  és az  $(y_n)_{n=1}^\infty$  sorozatok  $n$ -dik tagjainak távolsága kisebb mint  $(b - a)/n$ , ezért a két részsorozat határértéke megegyezik. De ha két sorozat határértéke azonos, akkor az  $f$  feltételezett folytonossága alapján a függvényértékek határértékének is azonosnak kell lenni, ami viszont ellentmond annak, hogy az  $|f(x_n) - f(y_n)|$  távolság nagyobb mint  $\varepsilon/(b - a)$ . □

**Állítás 275** *Ha az  $f : [a, b] \rightarrow \mathbb{R}$  leképezés monoton, akkor integrálható az  $[a, b]$  intervallumon.*

**Bizonyítás.** Legyen az  $f$  függvény mondjuk monoton növekedő. Ekkor először is korlátos, mivel  $f(a) \leq f(x) \leq f(b)$ . Feltehető, hogy  $f(a) < f(b)$ , mivel ellenkező esetben az állandó függvény integrálásáról lenne szó, amiről egy példában már láttuk, hogy integrálható. Az integrálhatóság bizonyításához legyen az  $\varepsilon$  egy tetszőleges pozitív szám, és az  $I$  egy olyan felosztás, amelyben a részintervallumok hossza kisebb, mint

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

Az  $f$  egy  $[x_i, x_{i+1}]$  részintervallumon — monoton növekedése miatt — az  $x_i$ -nél veszi fel a minimumát, az  $x_{i+1}$ -nél pedig a maximumát, ezért az oszcillációs összeg:

$$\begin{aligned}\Omega(f, I) &= S(f, I) - s(f, I) = \\ &= \sum_{i=1}^n (x_i - x_{i-1}) \left( \sup_{x \in [x_{i-1}, x_i]} f(x) - \inf_{x \in [x_{i-1}, x_i]} f(x) \right) = \\ &= \sum_{i=1}^n (x_i - x_{i-1}) (f(x_i) - f(x_{i-1})) < \sum_{i=1}^n \frac{\varepsilon}{f(b) - f(a)} (f(x_i) - f(x_{i-1})) = \\ &= \frac{\varepsilon}{f(b) - f(a)} \cdot \sum_{i=1}^n (f(x_i) - f(x_{i-1})) = \frac{\varepsilon}{f(b) - f(a)} [f(b) - f(a)] = \varepsilon,\end{aligned}$$

amivel be is láttuk az állítást.  $\square$

### 8.3 A határozott integrál és a differenciálás

Ebben az alponban a határozott integrálnak és a differenciálszámításnak a kapcsolatát fogjuk megvizsgálni.

A következő tétel lehetőséget ad arra, hogy az antiderivált (határozatlan integrál) segítségével kiszámoljuk a határozott integrált. Emiatt ez a tétel a definíciók után talán a lehangsúlyosabb tudnivaló.

**Tétel 276 (Newton-Leibnitz tétel)** *Legyen az  $f$  függvény integrálható az  $[a, b]$  intervallumon, és tegyük fel, hogy van olyan  $F$  függvény, amely*

1. az  $f$  antideriváltja az  $(a, b)$  nyílt intervallumon, azaz

$$F'(x) = f(x), \quad x \in (a, b);$$

2. és amely folytonos az  $[a, b]$  zárt intervallumon.

Ekkor

$$\int_a^b f(x) dx = F(b) - F(a), \quad (8.5)$$

ahol az  $F(b) - F(a)$  különbségre az  $[F(x)]_a^b = [F]_a^b$  jelölés szokásos.

Az  $f$  függvény egy  $F$  antideriváltja

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

alakú, és így — ha a Newton-Lebnitz-tétel feltételei teljesülnek — azt írhatjuk, hogy

$$\int_a^b f(x) dx = [F]_a^b = \left[ \int f(x) dx + c \right]_a^b.$$

Ez szépen megmagyarázza azt, hogy miért is nevezik az antideriváltak halmazát határozatlan integrálnak. Hangsúlyozni kell azonban, hogy a mondottak csak akkor igazak, ha a Newton-Lebnitz-tétel feltételei teljesülnek. Határozott integrálja olyan függvénynek is lehet, aminek határozatlan integrálja nincs. Mivel deriválható függvény egyszersmind folytonos is, ezért az  $F$  függvényre tett két feltételt maga után vonja a következő erősebb, de egyetlen feltétel: Az  $F$  antideriváltja az  $f$  függvénynek egy olyan nyílt intervallumban, amelyiknek az  $[a, b]$  része.

**Bizonyítás.** Mivel az  $f$  integrálható az  $[a, b]$  intervallumon, ezért elégséges azt megmutatnunk, hogy tetszőleges

$$I = \{x_0, x_1, \dots, x_n\}$$

felosztáshoz van olyan  $m(f, I)$  közbülső közelítő összeg, amelyikre

$$m(f, I) = F(b) - F(a).$$

Ekkor ugyanis a közelítő összegek csak az  $F(b) - F(a)$  különbséget közelíthetik, tehát ez az integrál értéke. Alkalmazzuk a differenciálszámítás középértéktételét az  $F$  függvényre az  $[x_{i-1}, x_i]$  intervallumon. Ezt megtehetjük, mert az  $F$  folytonos ezeken az intervallumokon, és deriválható az intervallumok belsejében. Így minden  $i$ -re van olyan  $\xi_i \in (x_{i-1}, x_i)$ , amelyre

$$F(x_i) - F(x_{i-1}) = (x_i - x_{i-1}) \cdot f(\xi_i).$$

Ennek a felhasználásával egy közbülső közelítő összeg a következőképpen számolható:

$$\begin{aligned} m(f, I) &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot f(\xi_i) = \sum_{i=1}^n (F(x_i) - F(x_{i-1})) = \\ &= F(x_n) - F(x_0) = F(b) - F(a), \end{aligned}$$

és ezzel be is láttuk a tételt. □

Most pedig azt fogjuk megvizsgálni, hogy a határozott integrál segítségével hogyan tudunk megadni antiderivált függvényt. A következő tételek előtt bevezetünk egy fogalmat:

**Definíció 277** Legyen az  $f$  függvény integrálható az  $[a, b]$  intervallumon. Az

$$x \mapsto \int_a^x f(t) dt, \quad x \in [a, b]$$

függvényt az  $f$  integrálfüggvényének mondjuk.

A definícióban szereplő függvény létezik, hiszen beláttuk, hogy az  $[a, b]$  intervallumon való integrálhatóságból következik az  $[a, x]$  halmazon való integrálhatóság. Az integrálfüggvény két fontos tulajdonságát tartalmazza a következő tétel.

**Állítás 278** *Legyen az  $f$  függvény integrálható az  $[a, b]$  intervallumon.*

1. Az

$$x \mapsto \int_a^x f(t) dt$$

*integrálfüggvény folytonos az  $[a, b]$  intervallumon.*

2. Ha az  $f$  függvény folytonos egy  $x_0 \in [a, b]$  pontban, akkor ott az integrál függvénye deriválható, és

$$\frac{d}{dx} \left( \int_a^x f(t) dt \right) (x_0) = f(x_0).$$

Az első állítás röviden: Tetszőleges integrálható függvény integrálfüggvénye folytonos.

A második állítás különösen fontos: Folytonos függvény integrálfüggvénye a függvénynek egy antideriváltja.

Ebből következik, hogy folytonos függvény esetében teljesülnek a Newton-Lebnitz-tétel feltételei.

**Bizonyítás. (1):** Jelöljük az  $f$  integrálfüggvényét  $F$ -fel, és az  $f$  egy korlátját  $K$ -val. Ha  $a \leq x \leq y \leq b$ , akkor mivel az integrál az integrációs határ additív függvénye, ezért

$$F(y) - F(x) = \int_a^y f(t) dt - \int_a^x f(t) dt =$$

$$\left( \int_a^x f(t) dt + \int_x^y f(t) dt \right) - \int_a^x f(t) dt = \int_x^y f(t) dt.$$

Felhasználva az abszolút értékre vonatkozó egyenlőtlenséget

$$|F(y) - F(x)| = \left| \int_x^y f(t) dt \right| \leq \int_x^y |f(t)| dt \leq K \cdot |y - x|,$$

ami alapján az  $F$  folytonossága nyilvánvaló. Az  $x \leq y$  feltevés természetesen mellékes, hiszen két tetszőleges  $x, y$  pontra vagy  $x \leq y$  vagy megfordítva.

**(2):** Tegyük fel, hogy az  $f$  függvény folytonos az  $x_0$  pontban. Becsüljük meg az

$$\left| \frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) \right|$$

eltérést. Mivel az integrál additív mind a függvényekre, mind a határookra nézve, ezért

$$\begin{aligned} \left| \frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) \right| &= \left| \frac{\int_{x_0}^{x_0+h} f(t) dt - h \cdot f(x_0)}{h} \right| = \\ &= \left| \frac{\int_{x_0}^{x_0+h} (f(t) - f(x_0)) dt}{h} \right|. \end{aligned}$$

Mivel az  $f$  folytonos az  $x_0$  pontban ezért tetszőleges  $\varepsilon > 0$  számhoz létezik olyan  $\delta > 0$  szám, hogy

$$|f(t) - f(x_0)| < \varepsilon \quad \text{ha} \quad t \in (x_0 - \delta, x_0 + \delta).$$

Tegyük fel először, hogy  $\delta > h > 0$ . Felhasználva az abszolút értékre vonatkozó egyenlőtlenséget:

$$\left| \frac{\int_{x_0}^{x_0+h} (f(t) - f(x_0)) dt}{h} \right| \leq \frac{\int_{x_0}^{x_0+h} |f(t) - f(x_0)| dt}{h}.$$

Ebből viszont

$$\left| \frac{F(x_0 + h) - F(x_0)}{h} - f(x_0) \right| \leq \frac{\int_{x_0}^{x_0+h} \varepsilon dt}{h} = \varepsilon.$$

Ha  $-\delta \leq h < 0$ , akkor

$$\left| \frac{\int_{x_0}^{x_0+h} (f(t) - f(x_0)) dt}{h} \right| = \left| \frac{-\int_{x_0+h}^{x_0} (f(t) - f(x_0)) dt}{h} \right| \leq \frac{\int_{x_0+h}^{x_0} \varepsilon dt}{|h|} = \varepsilon.$$

Ez pedig éppen azt jelenti, hogy

$$\frac{dF}{dx}(x_0) = \lim_{h \rightarrow 0} \frac{F(x_0 + h) - F(x_0)}{h} = f(x_0).$$

□

**Példa 8.4** Határozzuk meg a  $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \dots$  sor összegét!

A sor nyilvánvalóan Leibnitz-típusú, ezért konvergens, de nem abszolút konvergens, mivel a harmonikus sor divergens.

Az összeg meghatározása nem rutin-feladat. A részletösszegek konvergenciája miatt elégséges kiszámolnunk azt, hogy az

$$s_{2n} = \sum_{i=1}^{2n} (-1)^{i+1} \frac{1}{i}$$

páros indexű sorozat mihez tart. Először hajtsuk végre a következő átalakítást:

$$\begin{aligned}\sum_{i=1}^{2n} (-1)^{i+1} \frac{1}{i} &= \sum_{i=1}^{2n} \frac{1}{i} - 2 \cdot \sum_{i=1}^n \frac{1}{2i} = \\ &= \sum_{i=1}^{2n} \frac{1}{i} - \sum_{i=1}^n \frac{1}{i} = \sum_{i=n+1}^{2n} \frac{1}{i}.\end{aligned}$$

A jobboldalon kapott alakot úgy alakítjuk, hogy egy integrál közelítő összeget kapjunk:

$$\sum_{i=n+1}^{2n} \frac{1}{i} = \frac{1}{n} \left[ \frac{1}{1 + \frac{1}{n}} + \frac{1}{1 + \frac{2}{n}} + \cdots + \frac{1}{1 + \frac{n}{n}} \right].$$

A jobboldal az  $x \mapsto \frac{1}{1+x}$  függvénynek az  $[0, 1]$  intervallumon, ekvidisztans ( $n$  részre osztás melletti) alsó közelítő összege, ami tart az

$$\int_0^1 \frac{1}{1+x} dx$$

integrál értékéhez, ami a Newton-Leibnitz szabály alapján egyszerű számolással:

$$\int_0^1 \frac{1}{1+x} dx = [\ln(1+x)]_0^1 = \ln 2 - \ln 1 = \ln 2.$$

## 8.4 Integrálszámítási szabályok, példák

A határozott integrál kiszámolásához a legalapvetőbb kiindulásunk természetesen a Newton-Leibnitz szabály, és ennek megfelelően a számítási eljárások magja egy antiderivált meghatározása. Az antiderivált kiszámolása két legfontosabb módszerének a parciális integrálásnak és a helyettesítésnek a szabályait fogalmazzuk át határozott integrálra. Ezek nélkül az átfogalmazások nélkül is lehet persze alkalmazni a Newton-Leibnitz szabályt, de a számolás menetét esetleg lerövidíthetjük a következő tételek segítségével.

**Állítás 279** *Ha az  $f$  és  $g$  függvények folytonosan differenciálhatók az  $[a, b]$  intervallumon akkor*

$$\int_a^b f'(x)g(x) dx = [f(x)g(x)]_a^b - \int_a^b f(x)g'(x) dx.$$

**Bizonyítás.** A tett feltételek biztosítják, hogy az  $f'g$  és  $fg'$  függvények integrálhatóak és létezik antideriváltjuk, ezért alkalmazható rájuk a Newton-Leibnitz szabály:

$$\left[ \int_a^b f'(x)g(x) dx \right]_a^b = \left[ f(x)g(x) - \int_a^b f(x)g'(x) dx \right]_a^b,$$

amiből

$$\int_a^b f'(x)g(x) dx = [f(x)g(x)]_a^b - \int_a^b f(x)g'(x)dx.$$

□

**Példa 8.5** Számítsuk ki az

$$\int_0^2 xe^x dx$$

integrált.

Azonnal látható, hogy teljesülnek az állítás feltételei, és így parciális integrálással

$$\int_1^2 xe^x dx = [xe^x]_1^2 - \int_1^2 e^x dx = [(x-1)e^x]_1^2 = e^2.$$

□

**Állítás 280** Ha az  $f$  függvény folytonos az  $[c, d]$  intervallumon és a  $g$  függvény folytonosan differenciálható leképezése a  $[a, b]$  intervallumnak az  $[c, d]$  intervallumba, akkor

$$\int_{g(a)}^{g(b)} f(x) dx = \int_a^b f(g(t))g'(t) dt.$$

Bizonyítás. Jelölje  $F$  az  $f$  integrálfüggvényét. Mivel az  $f$  folytonos, ezért a  $F$  a  $f$  antideriváltja. A Newton-Leibnitz szabály alapján

$$\int_{g(a)}^{g(b)} f(x) dx = [F]_{g(a)}^{g(b)} = F(g(b)) - F(g(a)).$$

Ugyanakkor mivel az  $F \circ g$  összetett függvény deriváltja az  $x$  pontban

$$f(g(x))g'(x),$$

amely szintén folytonos, és így ismételten a Newton-Leibnitz szabály alapján

$$\int_a^b f(g(t))g'(t) dt = [F \circ g]_a^b = F(g(b)) - F(g(a)).$$

□

**Példa 8.6** Határozzuk meg az

$$\int_1^2 \frac{e^{\frac{1}{x}}}{x^2} dx$$

integrált.

Az  $1/x$  helyébe jónak látszik  $t$ -t helyettesíteni, ezért a tétel jelöléseivel:

$$x = g(t) = \frac{1}{t}, \quad t = g^{-1}(x) = \frac{1}{x} \quad \text{és} \quad \frac{dx}{dt} = -\frac{1}{t^2}$$

és  $g^{-1}(1) = 1$  és  $g^{-1}(2) = 1/2$ . Ezek felhasználásával írhatjuk, hogy

$$\int_1^2 \frac{e^{\frac{1}{x}}}{x^2} dx = \int_1^{1/2} t^2 e^t \cdot \left(-\frac{1}{t^2} dt\right) = -\int_1^{1/2} e^t dt = [-e^t]_1^{1/2} = e - \sqrt{e}.$$

□

## 8.5 Impropius integrálok

A fejezet elején bevezettük a Riemann-integrál fogalmát. Ezzel a függvények egy megfelelő összessége Riemann-integrálható lett. A matematika tipikus problémája: hogyan lehet olyan integrálfogalmat bevezetni, amely mellett több függvény lesz integrálható. Ez nem öncélú törekvés, hiszen a valós számok bevezetését is ilyen ok motiválta: több szakaszt akartunk mérni, mint racionális számokkal lehetett.

Ebben a pontban egy olyan integrálfogalmat definiálunk, amely a Riemann-integrál egy általánosítása: az *impropius Riemann-integrált*. Az előzőekben olyan függvények egy összességére definiáltuk az integrálfogalmat, amelyek *korlátosak* egy *korlátos* intervallumon. Az alább bevezetett integrálfogalom segítségével általánosíthatjuk az integrál fogalmát bizonyos olyan függvényekre is, amelyek vagy

- (I) nem korlátosak az intervallum valamelyik végpontjának környezetében (vagy egyéb “bajuk” van ott);

vagy:

- (II) nem korlátos az intervallum, amin integrálni akarunk.

Az (I) esettel kezdjük. Korlátos intervallum esetében előfordulhat, hogy egy függvény az intervallum valamelyik pontjának környezetében nem korlátos, például az  $1/x^2$  a  $[0, 1]$  intervallum 0 pontjának környezetében nem korlátos. Olyan esetekkel fogunk foglalkozni, amikor az intervallum valamelyik (vagy mindkét) végpontja “problémás” pont. Ilyen függvények egy alkalmas összességére definiáljuk most a Riemann-integrál egy általánosítását, kezdve az intervallum jobboldali végpontjával:

**Definíció 281** Legyen az  $f$  függvény integrálható minden

$$[a, b - \varepsilon], \quad 0 < \varepsilon < b - a$$

intervallumon.

Ha  $a$

$$\lim_{t \nearrow b} \int_a^t f(x) dx$$

(véges) határérték létezik, akkor azt mondjuk, hogy az  $f$  improprius módon integrálható az  $[a, b]$  intervallumon, és az improprius (Riemann) integrálja a létező határérték. A jelölés a közönséges Riemann-integrállal azonos, azaz

$$\lim_{t \nearrow b} \int_a^t f(x) dx = \int_a^b f(x) dx.$$

Az improprius integrál határértékkel van definiálva, ezért ahelyett, hogy  
 “az  $f$  impropriusan integrálható az  $[a, b]$  intervallumon”

tradicionálisan azt is szokták mondani, hogy

$$\text{az } \int_a^b f(x) dx \text{ improprius integrál konvergens}.$$

Mivel a (közönséges) integrál is — lényegében véve — határértékkel van értelmezve, ezért annál is élhetnénk az utóbbi szóhasználattal, de ott nem szokás. A hagyományos elnevezések, sajnos nem mindig logikusak és következetesek.

Mielőtt definiálnánk azt az esetet, amikor az intervallum másik, (vagy mindkét) végpontjában szinguláris (nem “jó viselkedésű”) a függvény, jegyezzük meg, hogy a jelölés zavaró addig, amíg nem látjuk be azt, hogy a bevezetett új integrálfogalom a Riemann-integrál kiterjesztése.

**Állítás 282** Ha az  $f$  integrálható az  $[a, b]$  intervallumon, akkor impropriusan is integrálható, és a két integrál értéke megegyezik.

**Bizonyítás.** Az  $f$  függvény

$$F(t) = \int_a^t f(x) dx,$$

integrálfüggvénye folytonos az  $[a, b]$  intervallumon, ezért az improprius integrál létezik és az értéke:

$$\lim_{t \rightarrow b} F(t) = F(b) = \int_a^b f(x) dx,$$

ahol a jobboldali integrál most (közönséges) integrált jelöl. □

A jobboldali végponthoz hasonló az eljárás a baloldali végpont esetében. Lásuk mindjárt azt az esetet, amikor mindkét végpontot egyszerre vesszük figyelembe.

**Definíció 283** Legyen az  $f$  függvény integrálható az  $(a, b)$  intervallum minden zárt résintervallumán, és  $c \in (a, b)$  egy tetszőleges de rögzített pont. Az  $f$  függvényt impropriusan integrálhatónak mondjuk az  $[a, b]$  intervallumon, ha léteznek a

$$\lim_{t \searrow a} \int_t^c dx \quad \text{és} \quad \lim_{s \nearrow b} \int_c^s f(x) dx$$

(véges) határértékek, és az  $f$  improprius integráljának értéke a két limesz összege. Az improprius integrál jelölése megegyezik a közös integrál jelölésével.

A definíció helyességéhez be kellene látni, hogy a definíció független a  $c$  pont választásától. Ezt az egyszerű feladatot az olvasóra bízunk.

Ez a definíció természetesen tartalmazza a megelőző definíciót is, és csak a fokozatos bevezetés, a jobb megértés kedvéért kezdtük az egyik végponttal.

A 282. állítás itt is szóról-szóra kimondható, tehát az improprius integrál a Riemann-integrál általánosítása egy bővebb függvényosztályra.

**Példa 8.7** Mutassuk meg, hogy

$$\int_0^1 \frac{dx}{x^c} = \begin{cases} \frac{1}{1-c} & \text{ha } c < 1 \\ +\infty & \text{ha } c \geq 1. \end{cases}$$

A definíció szerint ha  $c \neq 1$

$$\int_0^1 \frac{1}{x^c} dx = \lim_{\alpha \searrow 0} \int_\alpha^1 \frac{1}{x^c} dx = \lim_{\alpha \searrow 0} \left[ \frac{x^{1-c}}{1-c} \right]_\alpha^1,$$

ezért folytatva a számolást

$$\int_0^1 \frac{1}{x^c} dx = \frac{1}{1-c} - \lim_{\alpha \searrow 0} \left( \frac{\alpha^{1-c}}{1-c} \right).$$

Ha  $(1-c) > 0$ , akkor a limesz 0, ha pedig  $(1-c) < 0$ , akkor  $-\infty$ .

A  $c = 1$  esetet kell még megnéznünk. Ekkor

$$\int_0^1 \frac{dx}{x} = \lim_{\alpha \searrow 0} \int_\alpha^1 \frac{dx}{x} = \lim_{\alpha \searrow 0} [\ln x]_\alpha^1 = \ln 1 - \lim_{\alpha \searrow 0} \ln \alpha = +\infty,$$

ahogyan állítottuk. □

**Definíció 284** Ha az  $f$  függvény integrálható minden  $[a, \beta]$ ,  $a \leq \beta$  intervallumon és létezik az

$$\lim_{\beta \rightarrow +\infty} \int_a^\beta f(x) dx$$

(véges) határérték, akkor az  $f$  függvényt impropriusan integrálhatónak mondjuk az  $[a, +\infty)$  korlátlan intervallumon, és az integrálja az előző határérték. A jelölés:

$$\int_a^{+\infty} f(x) dx.$$

Hasonló a balról korlátlan intervallumon vett integrál definíciója. A jobbról és balról korlátlan intervallumon definiáljuk még az improprius integrált:

**Definíció 285** Legyen az  $f$  függvény integrálható minden

$$[\alpha, \beta], \quad \alpha, \beta \in \mathbb{R}$$

intervallumon és legyen a  $c$  egy tetszőleges, de rögzített szám. Ha a

$$\lim_{t \rightarrow -\infty} \int_t^c f(x) dx \quad \text{és} \quad \lim_{s \rightarrow +\infty} \int_c^s f(x) dx$$

(véges) határértékek léteznek, akkor azt mondjuk, hogy az  $f$  impropriusan integrálható  $(-\infty, +\infty)$  korlátlan intervallumon, és az integrálja az előző két határérték összege. A jelölés:

$$\int_{-\infty}^{+\infty} f(x) dx.$$

Könnyű igazolni, hogy a definíció független a  $c$  megválasztásától, tehát a definíció korrekt.

Természetes, hogy még sokféle esetre definiálhatnánk az improprius integrált. Például elő is fog fordulni olyan eset, mikor az  $(a, +\infty)$  minden zárt intervallumban integrálható a függvény és az  $[a, +\infty)$  intervallumra kell értelmezni az improprius integrálját. Ez az eddigiek ismeretében nem okozhat problémát.

**Példa 8.8** Határozzuk meg az

$$\int_0^{+\infty} x e^{-x} dx$$

integrál értékét.

A 284. definíció szerint

$$\int_0^{+\infty} x e^{-x} dx = \lim_{\beta \rightarrow +\infty} \int_0^{\beta} x e^{-x} dx,$$

számoljuk ki ezért a limesz mögötti integrált. Parciális integrálással

$$\int x e^{-x} dx = x(-e^{-x}) - \int (-e^{-x}) dx = -x e^{-x} - e^{-x} = -(x+1)e^{-x},$$

ezért

$$\int_0^{\beta} f(x) dx = [-(x+1)e^{-x}]_0^{\beta} = -(\beta+1)e^{-\beta} + 1.$$

Ebből pedig

$$\lim_{\beta \rightarrow +\infty} (-(\beta+1)e^{-\beta} + 1) = - \lim_{\beta \rightarrow +\infty} (1+\beta)e^{-\beta} + 1 = 0 + 1,$$

tehát az integrál értéke 1. □

**Példa 8.9** *Mutassuk meg, hogy*

$$\int_1^{\infty} \frac{dx}{x^c} = \begin{cases} \frac{1}{c-1} & \text{ha } c > 1 \\ +\infty & \text{ha } c \leq 1. \end{cases}$$

A definíció szerint ha  $c \neq 1$  akkor

$$\begin{aligned} \int_1^{+\infty} \frac{dx}{x^c} &= \lim_{\alpha \rightarrow +\infty} \int_1^{\alpha} \frac{dx}{x^c} = \lim_{\alpha \rightarrow +\infty} \left[ \frac{x^{1-c}}{1-c} \right]_1^{\alpha} = \\ &= \lim_{\alpha \rightarrow +\infty} \left( \frac{\alpha^{1-c}}{1-c} \right) - \frac{1}{1-c}. \end{aligned}$$

Ha  $(1-c) < 0$ , akkor a limesz 0, ha pedig  $(1-c) > 0$ , akkor  $+\infty$ .

A  $c = 1$  esetet kell még megnéznünk. Ekkor

$$\int_1^{+\infty} \frac{dx}{x} = \lim_{\alpha \rightarrow +\infty} \int_1^{\alpha} \frac{dx}{x} = \lim_{\alpha \rightarrow +\infty} [\ln x]_1^{\alpha} = +\infty,$$

ahogyan állítottuk.  $\square$

A formális szabályokra vonatkozó állítások némelyike szóról szóra fennáll improprius integrálható függvényekre is. Például teljesül, hogy az új integrál fogalomra is vektorteret alkotnak az integrálható függvények. Tekintsük például a jobbról korlátlan intervallum esetét:

**Állítás 286** *Legyenek az  $f$  és  $g$  az  $[a, +\infty)$  intervallumon impropriusan integrálható függvények, és a  $\lambda$  egy tetszőleges valós szám. Ekkor az  $f+g$  és  $\lambda f$  függvények is impropriusan integrálhatóak, és*

$$(1) \int_a^{+\infty} (f(x) + g(x)) dx = \int_a^{+\infty} f(x) dx + \int_a^{+\infty} g(x) dx,$$

$$(2) \int_a^{+\infty} \lambda f(x) dx = \lambda \cdot \int_a^{+\infty} f(x) dx.$$

**Bizonyítás.** Az állítások a közönséges integrálra vonatkozó tételből, és a határérték formális szabályaiból adódnak. Lássuk például az összeg improprius integrálhatóságának a szabályát:

$$\begin{aligned} \lim_{\alpha \rightarrow +\infty} \int_a^{\alpha} (f(x) + g(x)) dx &= \lim_{\alpha \rightarrow +\infty} \left( \int_a^{\alpha} f(x) dx + \int_a^{\alpha} g(x) dx \right) = \\ &= \lim_{\alpha \rightarrow +\infty} \int_a^{\alpha} f(x) dx + \lim_{\alpha \rightarrow +\infty} \int_a^{\alpha} g(x) dx = \int_a^{+\infty} f(x) dx + \int_a^{+\infty} g(x) dx, \end{aligned}$$

amit bizonyítani kellett.  $\square$

A formális szabályoknak tekinthető állítások közül teljesül még a felső határ szerinti additivitást kimondó tétel is, azaz ha  $a \leq c$  és az  $f$  impropriusan integrálható az  $[a, +\infty)$  intervallumon, akkor

$$\int_a^{+\infty} f(x) dx = \int_a^c f(x) dx + \int_c^{+\infty} f(x) dx.$$

Hangsúlyos megjegyzés, hogy nem minden állítás vihető át Riemann integrálról improprius integrálra. Például az  $f$  improprius integrálhatóságából nem következik feltétlenül az abszolút értékének az improprius integrálhatósága (8.12. példa). Ha azonban az  $f$  és  $|f|$  függvények is impropriusan integrálhatóak, akkor

$$\left| \int_a^{+\infty} f(x) dx \right| \leq \int_a^{+\infty} |f(x)| dx.$$

Improprius integráloknál gyakori az, hogy nem tudjuk meghatározni az integrál értékét, de szükségünk van arra, hogy tudjuk: létezik-e egyáltalán az integrál. A következő tételek erre a problémára adnak választ. A két típusú improprius integrálnak megfelelően két konvergencia (létezési) kritériumot mondunk ki.

**Állítás 287** *Ha az  $f$  függvény integrálható minden  $[a, \beta]$ ,  $a \leq \beta$  intervallumon, akkor az*

$$\int_a^{+\infty} f(x) dx$$

*improprius integrál pontosan akkor létezik (konvergál), ha tetszőleges  $\varepsilon$  pozitív számhoz van olyan  $p$  szám, hogy*

$$\left| \int_{t_1}^{t_2} f(x) dx \right| \leq \varepsilon, \quad \text{ha } p \leq t_1, t_2,$$

**Bizonyítás.** Egyszerű következménye a függvény-konvergenciára vonatkozó Cauchy-kritériumnak.

Vegyük az

$$F(t) = \int_a^t f(x) dx, \quad a \leq t$$

integrálfüggvényt. Az említett tétel szerint az  $F$  függvénynek pontosan akkor van véges határértéke a plusz végtelenben, ha tetszőleges  $\varepsilon$  pozitív számhoz van olyan  $p$  szám, hogy

$$|F(t_2) - F(t_1)| \leq \varepsilon \quad \text{ha } p \leq t_1, t_2,$$

amiből — ha mondjuk  $t_1 \leq t_2$  — az

$$F(t_2) - F(t_1) = \int_{t_1}^{t_2} f(x) dx$$

egyenlőség miatt következik is az állítás. □

**Állítás 288** *Legyen az  $f$  függvény integrálható minden*

$$[a + \varepsilon, b], \quad 0 < \varepsilon < b - a$$

*intervallumon. Az*

$$\int_a^b f(x) dx = \lim_{t \searrow a} \int_t^b f(x) dx$$

*impropius integrál pontosan akkor létezik és véges, ha minden  $\varepsilon$  pozitív számhoz van olyan  $\delta$  pozitív szám, hogy*

$$\left| \int_{u_1}^{u_2} f(x) dx \right| \leq \varepsilon, \quad \text{ha } u_1, u_2 \in (a, a + \delta).$$

A bizonyítás pontosan azonos az előző állítás igazolásával, csak a véges helyen vett függvény határértékére vonatkozó Cauchy-kritériumot kell használnunk.

Az előző két kritérium és a 8.9. és 8.7. példa alapján egyszerű eszközt adunk arra, hogy az impropius integrál létezését beláthassuk.

**Állítás 289** *A 287. állítás jelöléseivel, ha minden eléggé nagy  $x$ -re*

$$|f(x)| \leq K \cdot \frac{1}{x^c},$$

*valamilyen  $c > 1$  számra, akkor az*

$$\int_a^{+\infty} f(x) dx$$

*impropius integrál létezik és véges.*

**Bizonyítás.** A 287. állításban leírt kritériumot alkalmazzuk, figyelembe véve azt, hogy egy függvény abszolút értékének az integrálja nagyobb (nem kisebb) a függvény integráljánál és nagyobb függvény integrálja is nagyobb:

$$\begin{aligned} \left| \int_{t_1}^{t_2} f(x) dx \right| &\leq \int_{t_1}^{t_2} |f(x)| dx \leq K \cdot \int_{t_1}^{t_2} \frac{1}{x^c} dx = \\ &= K \cdot \left[ \frac{x^{1-c}}{1-c} \right]_{t_1}^{t_2} = K \cdot \frac{1}{c-1} (t_1^{1-c} - t_2^{1-c}) \leq K \cdot t_1^{1-c}. \end{aligned}$$

Az “ $K \cdot t_1^{1-c}$ ” adott  $\varepsilon$  pozitív számnál kisebb, ha a  $t_1$  és  $t_2$  megfelelően nagy, tehát teljesül a 287. kritérium.  $\square$

Az előzőhöz teljesen hasonló módon kapható a következő állítás.

**Állítás 290** *A 288. állítás jelöléseivel, ha minden, az 0 számhoz eléggé közeli  $x$ -re*

$$|f(x)| \leq K \cdot \frac{1}{x^c},$$

*valamilyen  $0 < c < 1$  számra, akkor az*

$$\int_0^b f(x) dx$$

*impropius integrál létezik és véges.*

**Példa 8.10** *Mutassuk meg, hogy létezik és véges az*

$$\int_{-\infty}^{+\infty} x^n e^{-x^2} dx$$

*improprius integrál, ahol az  $n$  tetszőleges rögzített pozitív szám.*

Elég a nullától a plusz végtelenig vett integrállal foglalkozni, mert a mínusz végtelentől a nulláig vett integrál hasonlóan kezelhető. Mivel

$$\lim_{x \rightarrow +\infty} e^{-x^2} x^{n+2} = 0$$

minden  $n$  pozitív számra, ezért

$$x^n e^{-x^2} \leq K \cdot \frac{1}{x^2},$$

ha az  $x$  eléggé nagy, ezért a 290. alapján be is láttuk az állítást.  $\square$

**Példa 8.11** *Mutassuk meg, hogy az*

$$\int_1^{+\infty} \frac{\sin x}{x^2} dx \quad \text{és} \quad \int_1^{+\infty} \frac{\cos x}{x^2} dx$$

*integrálok léteznek és végesek.*

Az előző példa megoldásához hasonlóan az

$$\left| \frac{\sin x}{x^2} \right| \leq \frac{1}{x^2} \quad \text{és} \quad \left| \frac{\cos x}{x^2} \right| \leq \frac{1}{x^2}$$

egyenlőtlenségekből adódnak.  $\square$

**Példa 8.12** *Lássuk be, hogy az*

$$\int_1^{+\infty} \frac{\sin x}{x} dx$$

*integrál értéke véges, de az*

$$\int_1^{+\infty} \left| \frac{\sin x}{x} \right| dx$$

*integrál értéke már nem véges  $(+\infty)$ .*

Ez a példa azt mutatja, egy függvény improprius integrálhatóságából nem következik az abszolút értékének az improprius integrálhatósága.

Először lássuk be a konvergenciát. Parciális integrálással azt kapjuk, hogy

$$\int_1^\alpha \frac{\sin x}{x} dx = \left[ -\frac{1}{x} \cos x \right]_1^\alpha + \int_1^\alpha \frac{\cos x}{x^2} dx.$$

Ebből

$$\int_1^{+\infty} \frac{\sin x}{x} dx = \cos 1 + \int_1^{+\infty} \frac{\cos x}{x^2} dx.$$

A jobboldalon szereplő integrál konvergenciáját az előző példában beláttuk.

A másik integrál végtelenségéhez először is jegyezzük meg, hogy

$$\int_1^{+\infty} \left| \frac{\sin x}{x} \right| dx \geq \int_1^{+\infty} \frac{\sin^2 x}{x} dx,$$

és a jobboldali integrál végtelen voltát látjuk be. Egy ismert trigonometriaai azonosságot használva azt kapjuk, hogy

$$\int_1^\alpha \frac{\sin^2 x}{x} dx = \int_1^\alpha \frac{1 - \cos 2x}{2x} dx = [(1/2) \ln x]_1^\alpha - \int_1^\alpha \frac{\cos 2x}{2x} dx.$$

A jobboldali első tag a plusz végtelenhez tart, ha az  $\alpha$  tart a plusz végtelenhez, a jobboldali integrál értéke pedig a  $\frac{\sin x}{x}$  esetéhez hasonlóan véges.  $\square$

**Állítás 291 (Impropius integrál kritérium)** *Legyen az*

$$f : [1, +\infty) \rightarrow [0, +\infty)$$

*monoton fogyó függvény. Ekkor a*

$$\sum_{n=1}^{+\infty} f(n)$$

*sorozat pontosan akkor konvergens, ha az*

$$\int_1^{+\infty} f(x) dx$$

*impropius integrál konvergens.*

**Bizonyítás.** A feltételek alapján

$$\int_{m-1}^m f(x) dx \leq f(m) \leq \int_m^{m+1} f(x) dx, \quad \text{ha } m = 2, 3, \dots$$

Összegezve ezeket az  $m = 2, 3, \dots, n$  indexekre:

$$\sum_{m=2}^n \int_{m-1}^m f(x) dx = \int_1^n f(x) dx \leq \sum_{m=2}^n f(m) < \int_2^{n+1} f(x) dx.$$

Az  $\sum f(n)$  sor szeletei és a  $t \rightarrow \int_1^t f$  függvény monoton növekedők, és az előző kiemelt sor szerint ha az egyik konvergens, akkor a másik korlátos, és így egyszerre konvergensek.  $\square$

**Példa 8.13** Az  $\alpha$  szám milyen értékére konvergens illetve divergens a

$$\sum_{n=1}^{+\infty} \frac{1}{n^\alpha}$$

sor?

Az improprius integrál kritérium szerint a  $\sum \frac{1}{n^\alpha}$  sor egyszerre konvergens az

$$\int_1^{+\infty} \frac{1}{x^\alpha} dx$$

improprius integrállal, amiről láttuk, hogy konvergens, ha  $\alpha > 1$  és divergens, ha  $\alpha \leq 1$ .  $\square$

## 8.6 Hatványsorok

Az előzőekben már foglalkoztunk a  $\sum_{k=0}^{\infty} x^k$  végtelen sorral. Megállapítottuk, hogy a sor  $|x| < 1$  esetén konvergens, és ekkor

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}.$$

Most azzal a kérdéssel foglalkozunk, hogy vajon milyen függvények állíthatók elő

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \quad (8.6)$$

alakban, ahol  $a_k$  számsorozat, és a fenti sor milyen  $x$  értékekre konvergens.

### 8.6.1 Alapvető tulajdonságok

**Definíció 292** Legyen az  $a_k, k = 0, 1, 2, \dots$  valós számok egy tetszőleges sorozata. Ekkor a

$$\sum_{k=0}^{\infty} a_k x^k$$

sort *hatványsornak* nevezzük. Az  $a_k$  számok a hatványsor együtthatói. Néha a rövidebb  $\sum a_k x^k$  jelölést is használjuk.

**Állítás 293** Tekintsük a  $\sum a_k x^k$  hatványsort és legyen

$$r \stackrel{\text{def}}{=} \frac{1}{\limsup_{k \rightarrow +\infty} \sqrt[k]{|a_k|}}.$$

Ha a nevező nulla, úgy legyen  $r \stackrel{\text{def}}{=} +\infty$ , ha pedig a nevező  $+\infty$ , úgy legyen  $r \stackrel{\text{def}}{=} 0$ .

Ekkor a  $\sum a_k x^k$  hatványsor konvergens ha  $|x| < r$ , illetve divergens, ha  $|x| > r$ .

Az  $r$  számot a hatványsor *konvergencia sugarának* nevezzük. Egy hatványsor tehát konvergens a  $(-r, r)$  nyílt intervallumban. Ezt az intervallumot a hatványsor *konvergencia intervallumának* nevezzük.

**Bizonyítás.** Legyen  $x$  rögzített, és tekintsük a  $c_k = a_k x^k$  tagokból álló numerikus sort. Ekkor

$$\limsup_{k \rightarrow +\infty} \sqrt[k]{|c_k|} = \limsup_{k \rightarrow +\infty} \sqrt[k]{|a_k x^k|} = |x| \limsup_{k \rightarrow +\infty} \sqrt[k]{|a_k|} = \frac{|x|}{r}.$$

Tehát a gyökkritérium szerint  $|x|/r < 1$  esetén a  $\sum c_k$  sor konvergens, illetve  $|x|/r > 1$  esetén pedig divergens.  $\square$

**Példa 8.14** Milyen  $x$  pontokban konvergensek a

$$\sum_{k=1}^{+\infty} kx^k \quad \text{és} \quad \sum_{k=0}^{+\infty} \frac{x^k}{k!}$$

hatványsorok?

Számoljuk ki a konvergencia sugarak reciprokait. Az első példában

$$\lim_{k \rightarrow +\infty} \frac{\ln k}{k} = 0,$$

és így  $\lim_k \sqrt[k]{k} = 1$ , tehát a  $\sum kx^k$  hatványsor konvergencia sugara: 1.

Másrészt láttuk egy feladatban a numerikus sorozatoknál, hogy

$$\lim_{k \rightarrow +\infty} \sqrt[k]{k!} = +\infty,$$

ezért a  $\sum \frac{x^k}{k!}$  hatványsor konvergencia sugara végtelen.

**Állítás 294** Tekintsük a hatványsor összegfüggvényét

$$f(x) = \sum_{k=0}^{\infty} a_k x^k,$$

ahol a hatványsor konvergencia sugara  $r$ , és  $|x| < r$ . Akkor  $f$  a  $(-r, r)$  intervallum minden pontjában folytonos.

**Bizonyítás.** Legyenek  $x_0 \in (-r, r)$  és  $\epsilon > 0$  adottak. Ekkor van olyan  $c > 0$ , amelyre  $|x_0| < c < r$ . Válasszunk egy olyan  $n$  indexet, hogy

$$\sum_{k=n+1}^{\infty} |a_k| c^k < \frac{\epsilon}{3}. \quad (8.7)$$

hiszen ez a sor abszolút konvergens. Mivel az

$$s_n(x) = \sum_{k=0}^n a_k x^k$$

részletösszeg folytonos függvény (hiszen polinom), van olyan  $\delta > 0$ , hogy bármely  $x \in (x_0 - \delta, x_0 + \delta)$  mellett  $|s_n(x) - s_n(x_0)| < \epsilon/3$ . Természetesen föltehető, hogy  $|x_0| + \delta < c$ . Tehát (8.7) alapján

$$|f(x) - f(x_0)| \leq |f(x) - s_n(x)| + |s_n(x) - s_n(x_0)| + |s_n(x_0) - f(x_0)| < \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

bármely  $x \in (x_0 - \delta, x_0 + \delta)$  esetén. Ez éppen azt jelenti, hogy  $f$  folytonos az  $x_0$  pontban.  $\square$

**Állítás 295** Legyen  $r$  a hatványsor konvergencia sugara, és

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

az összegfüggvénye a  $(-r, r)$  intervallumon. Legyen  $s_n$  a hatványsor  $n$ -ik szelete. Akkor

$$\int_0^c f(x) dx = \lim_{n \rightarrow \infty} \int_0^c s_n(x) dx$$

bármely  $c \in (-r, r)$  esetén.

**Bizonyítás.** Valóban, legyen  $\epsilon > 0$  adott, és válasszuk meg az  $N$  indexet úgy, hogy (8.7) teljesüljön minden  $n \geq N$  mellett. Ekkor

$$\left| \int_0^c f(x) dx - \int_0^c s_n(x) dx \right| \leq \left| \int_0^c |f(x) - s_n(x)| dx \right| < \frac{\epsilon}{3} c,$$

hacsak  $n \geq N$ . Innen azonnal adódik az állítás.  $\square$

**Tétel 296** Legyen  $r$  a hatványsor konvergencia sugara, és

$$f(x) = \sum_{k=0}^{\infty} a_k x^k$$

az összegfüggvénye a  $(-r, r)$  intervallumon. Ekkor  $f$  differenciálható, és

$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$$

a  $(-r, r)$  intervallumban.

Bizonyítás. Mivel

$$\limsup \sqrt[k]{k|a_k|} = \limsup \sqrt[k]{|a_k|},$$

azért a  $\sum_{k=1}^{\infty} ka_k x^{k-1}$  hatványsor konvergencia sugara szintén  $r$ . Vezessük be a

$$g(x) = \sum_{k=1}^{\infty} ka_k x^{k-1}, \quad t_n(x) = \sum_{k=1}^n ka_k x^{k-1}$$

jelöléseket. Ha most  $x_0 \in (-r, r)$ , úgy egyrészt

$$\lim_{n \rightarrow \infty} \int_0^{x_0} t_n(x) dx = \int_0^{x_0} g(x) dx.$$

Másrészt  $\int_0^{x_0} t_n(x) dx = \sum_{k=1}^n a_k x_0^k$ , ezért

$$\lim_{n \rightarrow \infty} \int_0^{x_0} t_n(x) dx = f(x_0) - f(0).$$

Következésképpen  $f(x_0) = f(0) + \int_0^{x_0} g(x) dx$ . Azonban a 294. Állítás szerint  $g$  folytonos, így  $f$  differenciálható az  $x_0$  pontban, éspedig

$$f'(x_0) = g(x_0) = \sum_{k=1}^{\infty} ka_k x_0^{k-1},$$

és éppen ezt akartuk bizonyítani.  $\square$

Nyilvánvaló, hogy állításunk többszöri alkalmazásával adódik, hogy az  $f$  összegfüggvény a konvergencia intervallumban akárhányszor differenciálható, és a  $j$ -ik deriváltja

$$f^{(j)}(x) = \sum_{k=j}^{\infty} k(k-1) \dots (k-j+1) a_k x^{k-j}$$

minden  $x \in (-r, r)$  mellett.

### 8.6.2 Példák

**Példa 8.15** Határozzuk meg a  $\sum_{k=1}^{\infty} k/3^k$  sor összegét.

A mértani sorra vonatkozó összegformula alapján

$$\sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}.$$

a 296. Tételből. A baloldali sor konvergencia sugara  $r = 1$ , így speciálisan az  $x = 1/3$  pontban

$$\sum_{k=1}^{\infty} \frac{k}{3^k} = \frac{1}{3} \cdot \frac{1}{(1-1/3)^2} = \frac{3}{4}.$$

**Példa 8.16** Állítsuk elő a

$$\sum_{k=0}^{\infty} \frac{x^k}{k!}$$

hatványsor összegfüggvényét.

Mivel  $\lim \sqrt[k]{k!} = \infty$ , a hatványsor konvergencia sugara  $r = \infty$ , azaz a sor a számegyenesen mindenütt konvergens. Másrészt a Taylor-formula szerint bármely  $x$  ponthoz van olyan  $\xi$  a 0 és az  $x$  között, hogy

$$e^x = \sum_{k=0}^n \frac{x^k}{k!} + \frac{e^{\xi}}{(n+1)!} x^{n+1}.$$

Rögzített  $x$  esetén a maradéktag nullához tart, tehát

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

minden  $x$  pontban.

**Példa 8.17** Milyen pontokban konvergens a  $\sum_{k=1}^{\infty} x^k/k$  hatványsor, és határozzuk meg az összegét.

Mivel  $\lim \sqrt[k]{k} = 1$ , a hatványsor konvergencia sugara  $r = 1$ . Másrészt a sor divergens  $x = 1$  esetén, hiszen ekkor a harmonikus sorhoz jutunk, de konvergens az  $x = -1$  pontban, ugyanis a sor ekkor Leibniz-típusú. Tehát a konvergencia-halmaz a  $[-1, 1)$  félig nyílt intervallum.

Jelölje  $f$  e sor összegfüggvényét. Ekkor a 296. Tétel szerint

$$f'(x) = \sum_{k=1}^{\infty} x^{k-1} = \frac{1}{1-x}.$$

Továbbá nyilván  $f(0) = 0$ , ezért

$$f(x) = \int_0^x f'(t) dt = -\ln(1-x).$$

Ez azt is mutatja, hogy  $\sum_{k=1}^{\infty} (-1)^{k-1}/k = \ln 2$ , amely azonosságot más módon már beláttunk.

**Példa 8.18** Igazoljuk a

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots,$$

illetve a

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$$

egyenlőségeket.

Csak az első azonosságot igazoljuk, a második teljesen hasonlóan mutatható meg. A 8.16. Példához hasonlóan látható, hogy a jobboldali hatványsor mindenütt konvergens a számegyenesen. A Taylor-formula alapján bármely  $x$  ponthoz van olyan  $\xi$  a 0 és az  $x$  között, hogy

$$\sin x = x - \frac{x^3}{3!} + \dots + \sin^{(n+1)} \xi \frac{x^{n+1}}{(n+1)!}.$$

Nem nehéz belátni, hogy bármely rögzített  $x$  mellett a maradéktag nullához tart, és ezzel az azonosságot igazoltuk.

# Tárgymutató

- Abszolút érték
  - komplex 53
- Abszolútérték-függvény 78
- Abszolút konvergencia 151
- Algebra(i)
  - alaptétele 62
  - egyenlet 61
  - gyökei 62
  - stuktúra 69
- Alsó határ, infimum
  - függvény 81
  - halmaz 40
- Alsó korlát 40
- Alulról korlátos
  - függvény 80
  - halmaz 40
- Antideriválás 244
  - helyettesítéses 246, 255
  - parciális 246, 251
- Antiszimmetrikus 18, 20
  - szigorúan 18, 20
- Approximáció
  - érintő 178
  - lineáris 196
  - $n$ -ed rendű 196
  - Taylor-féle 196
- Archimedeszi tulajdonság 42
- Asszociativitás 70
- Állandó függvény 83
  - deriváltja 188
- Átlag sebesség 180
- Baloldali határérték 123
- Belsőpont 103
- Bijekció 26
- Bijektív 26
- Bolzano tétele 117
- Bolzano-Weierstrass tétel
  - komplexben 147
  - $\mathbb{R}^2$ -ben 145
  - $\mathbb{R}^p$ -ben 146
- Boole-algebra 13
- Brouwer fixpont tétele 119
- Cantor-féle tulajdonság 42
- Cauchy
  - kritériuma
    - függvényre 157
    - sorozatra 135
    - számsorra 150
- Cauchy-Schwarz egyenlőtlenség 68
- Csoport 73
  - kommutatív 73
- Deriválható
  - egyoldalról 173
  - egy pontban 166
  - $n$ -szer 174
  - nyílt intervallumon 171
  - parciálisan 167
  - zárt intervallumon 173
- Derivált
  - baloldali 173
  - egyoldali 173
  - egy pontban 166
  - elemi függvényeké 188
  - formális szabályok 183
  - jobboldali 173

- $n$ -edik 174
- nyílt intervallumon 171
- parciális 167
- relatív 182
- zárt intervallumon 173
- Descartes-szorzat 15
- Differenciahányados 165
- Differenciálegyenlet 259
  - szétválasztható változójú 260
- Differenciálható  $\rightarrow$  deriválható
- Differenciálhányados  $\rightarrow$  derivált
- Differenciálszámítás középértéktétele 193
- Direkt kép 29
- Diszjunkt 11
- Diszkrét metrika 96
- Disztributivitás 74
- Egyesítés, unió 11
- Egységgyök 59
- Ekvipotencia 32
- Ekvivalencia 21
  - osztály 21
  - reláció 21
- Elaszticitás 182
  - kereslet 182
- Elem 7
- Előjel, szignum függvény 78
- $\epsilon$ - $\delta$ -technika 108
- Euklideszi norma 68
- Euler-féle szám 226
- Exponenciális függvény 87
- Érintő 175
- Érintő approximáció tétel 178
- Értelmezési tartomány 23
- Értékkészlet 23, 29
- Felosztás (intervallum) 268
  - finomítása 268
  - közös finomítás 268
  - osztópontjai 268
- Felső határ, szuprémum 40
  - függvényé 81
  - halmazé 40
- Felső határ tulajdonság 42
- Felső korlát 40
- Felülről korlátos
  - függvény 80
  - halmaz 40
- Fixpont 118
- Folytonos 109
  - függvény  $\rightarrow$  függvény
  - hányados 115
  - összeg 115
  - szorzat 115
- Független változó 23
- Függvény, leképezés 22
  - abszolút érték 78
  - bijektív, egyértelmű 26
  - deriválható  $\rightarrow$  deriválható
  - deriváltja  $\rightarrow$  derivált
  - differenciálható  $\rightarrow$  deriválható
  - differenciálhányadosa  $\rightarrow$  derivált
  - direkt kép 29
  - diszkussziója 239
  - előjel 78
  - exponenciális 87
  - értelmezési tartomány 23
  - értékkészlet 23, 29
  - folytonos 109
  - független változó 23
  - gráf 24
  - határértéke, limesze 121, 122
    - formális szabályok 125, 131,
  - hatvány 87, 188
  - infimuma 81
  - injektív 26
  - inverz 30
  - inverz kép 29
  - kompozíció, összetétel 26
  - konvex 213
    - szigorúan 213
  - logaritmus 88
  - maximuma 83
  - minimuma 83
  - páratlan 91
  - páros 92

- periodikus 91
- szinusz, koszinusz 90
- szuprémuma 81
- szűrjektív 26
- tangens 94
- trigonometrikus 91
- valós 77
  - hányadosa 79
  - kompozíciója 79
  - monoton 79
  - növekedő 79
  - összege 79
  - szigorúan monoton 79
  - szorzata 79
  - szorzása számmal 79
- Gömb
  - baloldali nyílt 101
  - baloldali zárt 101
  - hiányos 102
  - jobboldali nyílt 102
  - jobboldali zárt 101
  - nyílt 97
  - zárt 97
- Gráf (függvény) 24
- Gyök
  - algebrai egyenleté 62
  - komplex számé 58
  - multiplicitás 62
  - tényező 62
- Gyökkritérium 155
- Halmaz 7
  - alsó határa, infimuma 40
  - alulról korlátos 40
  - belső pontja 103
  - Descartes szorzata 15
  - diszjunkt 11
  - egyesítés (unió) 11
  - ekvipotens 32
  - eleme 7
  - felső határa, szuprémum 40
  - felülről korlátos 40
  - komplementer 11
  - korlátos 40
  - limesz pontja 103
  - kontinuum számosságú 33
  - közös rész, metszet 11
  - különbség 11
  - megszámlálható 32
  - nyílt 104
  - részhalmaz 10
    - összesége,  $\mathcal{P}(X)$  10
  - számosság 32
  - szorzata 15
  - torlódási, érintkezési pont 104
  - üres 9
  - véges 32
  - Venn-diagram 7
  - zárt 105
- Határérték, limesz
  - függvény –
  - sorozat
- Hatvány-függvény 87, 188
- Hatványsor 296
  - konvergenca intervalluma 297
  - konvergenca köre 296
  - konvergenca sugara 296
- Hányados folytonossága 113
- Háromszög egyenlőtlenség 95
  - abszolút érték
  - komplex számok
  - metrika
  - Minkowski-egyenlőtlenség
  - norma
- Hiányos
  - gömb
  - intervallum
- Hölder egyenlőtlensége 234
- Imaginárius, képzetes rész 52
- Improprius integrál 287, 289, 290
  - konvergenca kritériumok 292
- Infimum
  - függvény 81
  - halmaz 40
- Inflexiós pont 223

- Injekció 26
- Injektív (függvény) 26
- Integrál 272, 276
  - függvény 282, 283
  - határozatlan  $\rightarrow$  antiderivált
  - határozott  $\rightarrow$  integrál
  - improprius  $\rightarrow$  improprius integrál
  - közelítő összegek  $\rightarrow$  közelítő összegek
  - Riemann 272
- Integrálható
  - folytonos függvény 280
  - Riemann szerint 272
- Intervallum
  - felosztás  $\rightarrow$  felosztás
  - hiányos 102
  - jobbról (balról) nyílt 102
  - jobbról (balról) zárt 107
- Inverz
  - elem 70
- Irreflexív (reláció) 18, 20
- Iteráció  $\rightarrow$  rekurzió
- Jensen egyenlőtlensége 214
- Kereslet elaszticitás 182
- Kezdeti érték 260
- Kommutatív csoport 73
- Kommutativitás 70
- Komplementer (halmaz) 11
- Komplex számok 47, 50, 52
  - abszolút érték 53
  - egységgyökök 59
  - gyöke 58
  - imaginárius, képzetes rész 52
  - konjugált 52
  - normál alak 50
  - teljes metrikus tér 147
  - trigonometrikus alak 56
  - valós rész 52
- Kompozíció (függvény) 26
- Konjugált (komplex) 52
- Konstans függvény 83
- Kontinuum számosság 33
- Konvex
  - függvény 213
  - folytonossága 217
  - kombináció 210
- Korlátos
  - függvény 80
  - halmaz 40
- Koszinusz függvény 90
- Környezet
  - gömb 101
- Közelítés  $\rightarrow$  approximáció
- Közelítő összegek
  - alsó 269
  - felső 269
  - közbülső 269
- Középértéktétel
  - differenciálszámítás 193
- Közös rész, metszet 11
- Közvetett függvény  $\rightarrow$  kompozíció
- Leibnitz-típusú sor 151
- Leképezés  $\rightarrow$  függvény
- Limesz
  - függvényé 122
  - sorozaté 135
- Limesz inferior 140
- Limesz pont halmazé 103
- Limesz szuperior 140
- Logaritmikus skála 235
- Logaritmus függvény 88
- Logisztikus görbe 263
- Lokális
  - maximum 190
  - szigorú 190
  - minimum 190
  - szigorú 190
  - szélsőérték 190
  - szigorú 190
- Marginális  $\rightarrow$  határ-
- Maximum
  - függvény 83
- Megszámálható 32

- Metrika, távolság 95
  - diszkrét 96
  - résztere, altere 97
- Metszet, közös rész (halmaz) 11
- Mértani, geometriai közép 45, 233
- Minimum
  - függvény 83
- Monoton
  - fogyó 79
  - növekedő 79
  - szigorúan 79
- Monoton függvény 79
  - határértékei 207
  - szakadásai 208
- Newton-Lebnitz tétel 281
- Norma, hosszúság 68
  - euklideszi 68
  - háromszög egyenlőtlenség 68
- Normál alak (komplex) 50
- Normális növekedés, szaporodás 261
- Növekedés
  - korlátozott normális 262
  - robbanásos 262
  - versengő 263
- Numerikus sor 147
  - abszolút konvergenciája 148
  - Cauchy-kritériuma 150
  - gyökkritérium 155
  - hányadoskritérium 155
  - integrál kritérium 287, 289, 290
  - konvergenciája 148
  - Leibnitz-típusú 151
  - nemnegatív tagú 152
  - összehasonlító kritérium 152
- Nyílt
  - gömb 97
  - halmaz 102
- Összeg folytonossága 112
- Összehasonlító kritérium 152
- Összetett függvény 26
- Parciális derivált 167
- Páratlan függvény 91
- Páros függvény 92
- Periodikus függvény 91
- Polár koordináták 55, 56
- Polinom 83
  - fokszám 84
  - együttható 84
  - gyökei 84
  - gyöktényezős alak 62
  - konstans 84
  - nulla 84
- Primitív függvény  $\rightarrow$  antiderivált
- Racionális számok 35
- Racionális törtfüggvény 85
- Reflexív (reláció) 18, 20
- Rekurzió 44
- Rekurzív definíció 44
- Relatív
  - átlagsebesség 181
  - derivált 182
  - differenciálhányados  $\rightarrow$  derivált
  - megváltozás 182
- Reláció
  - antiszimmetrikus 18, 20
  - szigorúan 18, 20
  - ekvipotencia 32
  - ekvivalencia 21
  - irreflexív 18, 20
  - reflexív 18, 20
  - szimmetrikus 18, 20
  - teljes 18, 20
  - transzitiv 18, 20
- Rendezett
  - párok 15
  - test 39
- Rendszer 7
- Reprodukáló elem 70
- Riemann integrál  $\rightarrow$  integrál
- Sebesség 180
- $\mathcal{S}$ -görbe 263
- Skáláris szorzat 67
- Sor 47

- geometriai 153
  - numerikus 147
- Sorozat
  - Cauchy 135
  - határérték 135
  - konvergenciája 135
  - részsorozat 136
- Sorozat konvergencia
  - komplexben 143, 145, 146
  - $\mathbb{R}$ -ben 135
  - $\mathbb{R}^p$ -ben 135
  - végtelenhez 142
- Struktúra
  - algebrai 69
- Szaporodás  $\rightarrow$  növekedés
- Számok
  - egész
  - komplex 47, 52
  - racióális
  - természetes
  - valós
- Számosság
  - kontinuum 33
  - megszámlálható 33
  - véges 32
- Számsorozatok 135
- Számtani, aritmetikai közép 45, 233
- Szélsőérték  $\rightarrow$  lokális
- Szigorúan konvex 213
- Szimmetrikus
  - reláció 18, 20
- Színusz függvény 90
- Szorzat
  - folytonossága 111
  - halmaz 15
  - számmal  $\rightarrow$  vektor
- Szupremum, felső határ
  - függvény 80
  - halmaz 40
- Szurjekció 26
- Szurjektív (függvény) 26
- Tangens függvény 94
- Taylor
  - approximáció 196
  - polinom 196
  - sor 196
- Távolság  $\rightarrow$  metrika
- Teljes
  - indukció 43
  - reláció 18, 20
  - rendezésre nézve (test) 42
- Természetes számok
- Test 74
  - rendezett 39
  - rendezésre nézve teljes 42
- Tér
  - metrikus 95
  - vektor 65
- Tranzitív (reláció) 18, 20
- Trigonometrikus alak 57
- Trigonometrikus függvény 91
- Unió, egyesítés (halmaz) 11
- Üres (halmaz) 9
- Valós
  - értékű függvény 77
  - függvény 77
  - rész (komplex) 52
  - számok,  $\mathbb{R}$  38
- Végtelen határérték  $\rightarrow$  határérték
- Végtelen mértani sor 153
- Vektor
  - kivonás 65
  - negatívja 65
  - norma  $\rightarrow$  norma
  - összeadás 65
  - számmal való szorzása 65
  - $\mathbb{R}^p$  64
- Venn-diagram 7
- Weierstrass
  - tétele 119
- Young egyenlőtlensége 234

Zárt

gömb 97

intervallum 107

halmaz 107